JOURNAL OF AGRICULTURAL AND FOOD CHEMISTRY

Article

pubs.acs.org/JAFC

Structural Annotation and Elucidation of Conjugated Phenolic Compounds in Black, Green, and White Tea Extracts

Justin J. J. van der Hooft,*[,†,‡,⊥] Moktar Akermi,[†] Fatma Yelda Ünlü,[†] Velitchka Mihaleva,[†,⊥] Victoria Gomez Roldan,[‡,§,⊥] Raoul J. Bino,[#] Ric C. H. de Vos,[‡,⊗,⊥] and Jacques Vervoort[†,⊥]

[†]Laboratory of Biochemistry, Wageningen University and Research Centre, Dreijenlaan 3, 6703 HA Wageningen, The Netherlands

[‡]Plant Research International, BU Biosciences, Wageningen University and Research Centre, P.O. Box 16, 6700 AA Wageningen, The Netherlands

[§]Netherlands Consortium for Systems Biology, P.O. Box 94215, 1090 GE Amsterdam, The Netherlands

[#]Laboratory of Plant Physiology, Wageningen University and Research Centre, P.O. Box 658, 6700 AR Wageningen, The Netherlands

[⊗]Centre for BioSystems Genomics, P.O. Box 98, 6700 AB Wageningen, The Netherlands

[⊥]Netherlands Metabolomics Centre, Einsteinweg 55, 2333 CC Leiden, The Netherlands

Supporting Information

ABSTRACT: Advanced analytical approaches consisting of both LC-LTQ-Orbitrap Fourier transformed (FT)-MS and LC-time-of-flight-(TOF)-MS coupled to solid-phase extraction (SPE) NMR were used to obtain more insight into the complex phenolic composition of tea. On the basis of the combined structural information from (i) accurate mass fragmentation spectra, derived by using LC-Orbitrap FTMS^n, and (ii) proton NMR spectra, derived after LC-TOFMS triggered SPE trapping of selected compounds, 177 phenolic compounds were annotated. Most of these phenolics were glycosylated and acetylated derivatives of flavan-3-ols and flavonols. Principal component analysis based on the relative abundance of the annotated phenolic compounds in 17 commercially available black, green, and white tea products separated the black teas from the green and white teas, with epicatechin-3,5-di-*O*-gallate and prodelphinidin-*O*-gallate being among the main discriminators. The results indicate that the combined use of LC-LTQ-Orbitrap FTMS and LC-TOFMS-SPE-NMR leads to a more comprehensive metabolite description and comparison of tea and other plant samples.

KEYWORDS: *ellagic acid, MS, MS^n, metabolite identification, NMR, phenolic conjugates, polyphenols, profiling, tea*

INTRODUCTION

Tea drink is prepared as a hot water infusion from tea material, that is, leaves or buds of the *Camellia sinensis* plant, and is one of the most consumed beverages in the world; its consumption has been positively associated with human health.[1−3] The chemical content of the tea drink and, thereby, its flavor, taste, and health characteristics depend on many factors such as postharvest treatment, genotype, growing conditions, and plant processing.[4] White tea is an infusion of the dried young leaf buds, green tea derives from steamed and dried mature green leaves, and black tea is an infusion of the fermented and heated leaves. Polyphenols, polyphenol conjugates, and polymerized phenolic structures are the main constituents of the tea drink. Due to the fermentation process, polymerized phenolic structures such as theaflavins and thearubiginins and larger related phenolic complexes can be found in black tea.[5] In addition, a diverse mixture of conjugated flavonoids is present in black, green, and white teas.[6] The number of glycoside moieties attached to phenolic aglycones such as kaempferol, quercetin, and myricetin ranges from one to four, and also acyl groups such as *p*-coumaroyl can be attached, forming complex conjugates.[7] The complete structural identification of these conjugates is important to improve the link between tea quality and nutritional value and its chemical variation.[8]

Hyphenated analytical techniques, such as liquid chromatography (LC) coupled to mass spectrometry (MS),[9] have been used to study the chemical content of tea samples.[1,6,7,10−13] Quality assessments and influences of production and processing effects on specific tea metabolites have also been described.[14,15] In view of their potential health benefits, most of these studies have been directed toward the class of phenolic compounds, which include polyphenols and hydroxycinnamic acids.[16] Because authentic standards are scarcely available, complete structural elucidation is often not achieved by using only LC-MS or LC-MS/MS. Thus, metabolite annotation and identification remain major challenges in tea metabolomics.

In the present research we applied LC coupled to an on-the-fly multistage accurate mass Orbitrap Fourier transformed MS (FTMS^n) approach paired with LC coupled to a time-of-flight (TOF) MS-based solid-phase extraction (SPE)-NMR method[17,18] to (i) get a more comprehensive picture of the complex metabolome of tea products and to (ii) fully identify selected

Special Issue: 5th International Conference on Polyphenols and Health

Received: January 20, 2012
Revised: March 20, 2012
Accepted: April 3, 2012
Published: April 3, 2012

dx.doi.org/10.1021/jf300297y | *J. Agric. Food Chem.* 2012, 60, 8841−8850

tea metabolites. Here, we focused on the annotation of the main phenolic acids and polyphenols present. Specifically, we studied the complex conjugation patterns of the flavan-3-ols catechin and epicatechin and the flavonols kaempferol and quercetin. The structural information from accurate mass LC-MS^n, that is, elemental formulas and fragmentation patterns, was combined with diagnostic signals in the one-dimensional (1D)-^{1}H NMR spectra for their chemical elucidation, leading to a more comprehensive insight into the metabolome of tea products.

MATERIALS AND METHODS

Tea Samples. Black, green, and white tea samples were obtained from different origins: a local store (bought on June 5, 2011) and a dedicated tea shop (bought on June 4, 2011) that also ordered several special tea samples at http://www.wollenhaupt.com/ (article no. in brackets) with a packaging date of May 19, 2011, indicated with an asterisk. The tea products were stored under dark conditions at room temperature until use (about 6 weeks). Tea samples were labeled as follows: green tea (GT), China Yunnan (Y) [article no. 00518]*, China Sencha (S) [article no. 00515]*, Chinese Huangshan Green Tea from laboratory (GL) (local store), China Bancha (B) [article no. 00568], China Chun Mee (CCM) [article no. 00500]*, China Lung Ching (CLC), China Gunpowder (CG); white tea (WT), China Mao Feng (CMF) [article no. 00519]*, China Pai Mu Tan (PMT) [article no. 00509]*, Jasmijn and Oranjebloesem (WTJ) (local store); black tea (BT), Darjeeling FTGFOP1 (D) [article no. 00312]*, Ceylon OP Adawatte (CAD) [article no. 00203]*, Assam TGFOP1 Hazelbank (A) [article no. 00109]*, English Earl Gray−Ceylon Black Tea with bergamot (CB) [article no. 10861], Irish Breakfast (IB) [article no. 00418], Ceylon OP Pettiagalla (CP) [article no. 00209], Pickwick English Tea Blend (BL) (local store).

Chemicals. Acetonitrile (HPLC grade) was obtained from Biosolve (Valkenswaard, The Netherlands), methanol (HPLC grade) from Merck-Schuchardt (Hohenbrunn, Germany), formic acid (99−100) from VWR International SAS (Briare, France), acetonitrile NMR chromasolve (for LC-NMR; SPE device) from Riedel-de Haën, Seelze, Germany, and deuterated methanol (99.8% pure) from CIL Inc., Andover, MA, USA. Ultrapure water was obtained from an in-house Millipore Milli-Q water purification system (A10 gradient, Millipak 40, 0.22 mm).

Metabolite Extraction. The methanol/water extraction was performed as follows: a few grams of dry tea was ground into a fine powder, and 100 mg was weighed and extracted with 10.00 mL of methanol/water (60:40, v/v) using a sonicator (40 kHz, 100 W) for 60 min at room temperature. The slurry mixture was centrifuged at 2500 rpm for 15 min. The supernatant was filtered through a 0.45 μm PVDF syringe filter prior to the SPE step.

Solid Phase Extraction. HLB SPE cartridges (3 cm^3, OASIS, Waters) were activated under vacuum with 4 mL of methanol and washed with 6 mL of H_2O. Then, 1.5 mL of extracted tea sample was diluted 1:1 with H_2O containing 4% phosphoric acid (final pH about 3.0) and deposited on the SPE cartridges, followed by a washing with 4 mL of H_2O. Elution of tea compounds was performed with 4 mL of 100% methanol. Subsequently, the eluates were dried overnight in a Speedvac protected from light at 35 °C and then stored at −20 °C until further use (within 5 days).

Sample Preparation for Analysis. For LC-MS^n the freeze-dried samples were redissolved in 200 μL of 75% methanol in H_2O containing 0.1% formic acid, whereas for LC-MS-SPE-NMR, the freeze-dried samples of four SPE extractions (thus 12 mL of sample) were combined and redissolved in 800 μL of 75% MeOH in H_2O containing 0.1% formic acid. Samples were then sonicated for 5 min and filtered through a 0.45 μm filter before transfer to an HPLC vial.

LC-MS^n Conditions. The C18-reversed phase LC-MS^n setup consisted of an Accela HPLC tower connected to a photodiode array (PDA) detector and an LTQ/Orbitrap hybrid mass spectrometer (Thermo Fisher Scientific). LC conditions were as described earlier,[19] using a binary eluent solvent system of degassed ultrapure water and acetonitrile, both containing 0.1% v/v FA (solvents A and B, respectively), with an injection volume of 5 μL, a flow rate of 0.19 mL/min, and a column temperature of 40 °C. Eluting compounds were trapped within an LTQ Ion Trap followed by accurate mass MS^n fragmentation. MS settings and spectral tree topologies were as previously described.[18]

HPLC-MS-SPE-NMR Conditions. The HPLC-MS-SPE-NMR system was used as described in ref 17. In short, it consisted of the online-coupled Agilent 1200 quaternary solvent delivery pump, Agilent 1200 degasser, Agilent 1200 autosampler, Bruker Daltonics MicrO-TOF ESI mass spectrometer, Knauer K120 pump for postcolumn water delivery, Spark Prospekt 2 SPE device containing Oasis (Waters Co.) HLB cartridges (2.0 × 10 mm i.d.), and the Bruker Advance III 600 spectrometer. Chromatographic separation was on an Alltima HP column (Alltima, 4.6 × 150 mm i.d., particle size = 3 mm) with a precolumn of the same material. For black (Pickwick English Tea Blend) and green teas (China Bancha), 4 and 5, respectively, cumulative trappings of selected compounds were performed, on the basis of their LC-MS base peak, using 40 μL injection volume for each trapping event.

NMR Experiments. The content of each cartridge was transferred to the NMR spectrometer by elution with 227 μL of MeOD, followed by the acquisition of 1D-^{1}H NMR measurements of 128 scans and a d_1 of 4 s (i.e., almost 15 min) at a receiver gain of 128 in MeOD at 600 MHz NMR (Bruker Daltronics) using standard pulse sequences. Two-dimensional (2D) ^{1}H measurements (COSY) were conducted for structural confirmation if the purity and signal/noise ratio of the obtained 1D-^{1}H spectra were sufficient (i.e., ≥8). Quantification of the complex conjugates in the NMR probe was performed by comparing proton integrals to those obtained from a 3 mm NMR tube filled with 200 μL of a standard flavonoid (rutin) dissolved in MeOD at a concentration of 13 μg/mL.

Metabolite Identification. Metabolites were identified by querying the assigned elemental formulas in Web-based databases such as Chemspider (http://www.chemspider.com/) and Scifinder (https://scifinder.cas.org/scifinder/). After the MS^n fragmentation patterns had been studied, the chemical structures of the flavonoid backbones were used as restraint. If NMR data were available, further candidate selection was performed by comparing observed NMR spectral data to those in the literature. The Dictionary of Natural Products (http://dnp.chemnetbase.com/) and KNApsSAcK (http://kanaya.naist.jp/KNApSAcK/) databases were used with the restraint of biological source (*Camellia sinensis*) to check if the annotated metabolites detected in our tea samples have been described before as being present in tea. In addition, we checked for phenolic compounds that were previously unambiguously identified in tea infusions using NMR[20] but were not present in the above-mentioned *Camellia* metabolite databases.

Data Analysis. Annotated metabolites were relatively quantified on the basis of the peak intensity (peak height) of the specific mass signal of the parent ions. For this purpose, the raw data files were first all preprocessed (peak picking by baseline correction and noise correction) using MetAlign software[21] (www.metalign.nl).[22,23] The 19 preprocessed data files (17 tea and 2 control samples) were used for ultrafast mass retention time searching using the Search_LCMS module of MetAlign.[24] The Excel compatible output contained observed accurate mass, retention time, scan number, ppm error with respect to the expected mass, and intensity for each target compound in each sample. Eleven annotated compounds were missed upon this automatic peak picking, due to too close elution of isomeric species, and were therefore not taken into account. The data matrix was subsequently used for unsupervised principal component analysis (PCA) using Genemaths XT v. 1.6 software (www.applied-maths.com). Metabolite intensities were normalized using $\log_2$ transformation and standardized using range scaling. Metabolites contributing to the separation between the black and green tea samples were selected using an ANOVA ($p < 0.01$). Relative quantification of 51 selected complex polyphenol conjugates (Supporting Information, Supplemental Table 2) was done manually

Table 1. Compounds Identified up to MSI Level 1 or 2[a]

metabolite common name	trapped from	MSI level ID	RT (min)	$[M - H]^-$ (m/z)	EF $[M - H]^-$ (CHON)	lit. data NMR
caffeine	BT	1	17.14	193.0732	$C_{16}H_{13}O_9$	
catechin	GT, BT	1	15.61	289.0719	$C_{15}H_{13}O_6$	
epicatechin	GT, BT	1	19.53	289.0718	$C_{15}H_{13}O_6$	
gallocatechin	GT	1	9.27	305.0669	$C_{15}H_{13}O_7$	
epigallocatechin	GT	1	13.96	305.0667	$C_{15}H_{13}O_7$	
p-coumaroylquinic acid-3-*O trans*	BT	1	15.65	337.0932	$C_{16}H_{17}O_8$	
p-coumaroylquinic acid-5-*O trans*	BT	1	20.02	337.0932	$C_{16}H_{17}O_8$	43
p-coumaroylquinic acid-4-*O cis*	GT	1	20.54	337.0933	$C_{16}H_{17}O_8$	
p-coumaroylquinic acid-4-*O trans*	BT	1	21.13	337.0930	$C_{16}H_{17}O_8$	
p-coumaroylquinic acid-5-*O cis*	BT	1	24.19	337.0930	$C_{16}H_{17}O_8$	43
theogallin (3-*O*-galloylquinic acid)	BT	1	4.88	343.0673	$C_{14}H_{15}O_{10}$	41
epicatechin-3-*O*-gallate	GT	1	26.63	441.0832	$C_{22}H_{17}O_{11}$	
kaempferol-3-*O*-glucoside	BT,GT	1	31.98	447.0938	$C_{21}H_{19}O_{11}$	17
epigallocatechin-3-*O*-gallate	GT	1	21.59	457.0781	$C_{22}H_{17}O_{11}$	
quercetin-3-*O*-galactoside	BT	1	28.18	463.0886	$C_{21}H_{19}O_{12}$	
quercetin-3-*O*-glucoside	BT	1	28.67	463.0888	$C_{21}H_{19}O_{12}$	17
Myricetin-3-*O*-galactoside	GT, BT	1	23.63	479.0835	$C_{21}H_{19}O_{13}$	20
myricetin-3-*O*-glucoside	BT	1	24.13	479.0836	$C_{21}H_{19}O_{13}$	20
kaempferol-3-*O*-rutinoside	BT	1	30.81	593.1518	$C_{27}H_{29}O_{15}$	17
phloretin-3,5-di-*C*-glucoside	BT	1	28.97	597.1830	$C_{27}H_{33}O_{15}$	
(epi)gallocatechin dimer 2	GT	2	10.67	609.1256	$C_{30}H_{25}O_{14}$	
(epi)gallocatechin dimer 3	GT	2	12.77	609.1249	$C_{30}H_{25}O_{14}$	
quercetin-3-*O*-gal-1,6-rhm	BT	1	27.14	609.1464	$C_{27}H_{29}O_{16}$	
quercetin-3-*O*-glc-1,6-rhm	BT	1	27.71	609.1469	$C_{27}H_{29}O_{16}$	17
(epi)gallocatechin dimer 1	GT	2	7.20	625.1249	$C_{30}H_{25}O_{14}$	
myricetin-3-*O*-(rhm-1,6-gal)	GT	1	23.53	625.1411	$C_{27}H_{29}O_{17}$	20
kaempferol-3-(glc-(1,6-rhm))-4′-rhm	BT	1	29.05	739.2100	$C_{33}H_{39}O_{19}$	
quercetin-3-*O*-(glc-(1,3-rhm-1,6-rhm))	GT	2	26.12	755.2049	$C_{33}H_{39}O_{20}$	
kaempferol-3-*O*-(glc-(1,3-rhm-1,6-gal))	BT	1	27.71	755.2045	$C_{33}H_{39}O_{20}$	20
kaempferol-3-*O*-(glc-(1,3-rhm-1,6-glc))	GT, BT	1	28.43	755.2044	$C_{33}H_{39}O_{20}$	20
quercetin-3-*O*-(glc-(1,3-rhm-1,6-gal))	GT, BT	1	24.84	771.1989	$C_{33}H_{39}O_{21}$	20
quercetin-3-*O*-(glc-(1,3-rhm-1,6-glc))	GT, BT	1	25.68	771.1993	$C_{33}H_{39}O_{21}$	
kaemp-pC-glycoside	GT	2	40.66	901.2402	$C_{42}H_{45}O_{22}$	
querc-3-(2G-pC-tr-3G)-2G-ara-3R-rhm-rut	GT	1	38.54	1033.2825	$C_{47}H_{53}O_{26}$	
kaemp-3-(2G-pC-tr-3G)-2G-ara-3R-glc-rut	GT	1	40.11	1033.2827	$C_{47}H_{53}O_{26}$	
kaemp-3-(2G-pC-cis-3G)-2G-ara-3R-glc-rut	GT	1	41.12	1033.2826	$C_{47}H_{53}O_{26}$	
querc-3-(2G-pC-tr-3G)-2G-ara-3R-glc-rut	GT	1	38.11	1049.2777	$C_{47}H_{53}O_{27}$	
querc-3-(2G-pC-cis-3G)-2G-ara-3R-glc-rut	GT	2	39.35	1049.2783	$C_{47}H_{53}O_{27}$	
kaemp-3-(2G-pC-tr-3G)-2G-glc-3R-glc-rut	GT	2	38.92	1063.2933	$C_{48}H_{55}O_{27}$	

[a]Metabolites are sorted based on their m/z value followed by their retention time order. The short notation for the complex conjugates is as follows: 3-*O*-(2G-*p*-coumaroyl(*trans*)-3G-*O*-*β*-L-arabinosyl-3*R*-*O*-*β*-D-glucosylrutinoside = 3-(2G-pC-tr-3G)-2G-ara-3R-glc-rut. Table column labels are follows: compound names, trapped from black tea, BL (BT) and/or green tea B (GT) extract, MSI MI level, RT in the LC-MS^n system, m/z $[M - H]^-$ EF $[M - H]^-$, and references to literature NMR data with MeOD as solvent.

using Xcalibur software and was based on the areas under the curve of the specific mass signal of the parent ions.

RESULTS

To identify phenolic compounds present in teas and profile 17 commercially available teas for the presence of the identified phenolic compounds, we extracted the tea material with aqueous methanol. Tea phenolic compounds are usually analyzed from hot water extracts of the dry leaf product. Our aim was to identify metabolites present in tea products, rather than analyzing the tea drinks. On the basis of preliminary results comparing aqueous methanol extraction[7,25] with hot water extraction (tea drink), we decided to use the aqueous methanol extraction, as it better represents the complete set of phenolic compounds and their conjugates present in the tea material. Moreover, the methanol/water extraction protocol is a robust and general extraction method of secondary metabolites from most plant samples,[22] enabling the direct comparison of profiles with minimal technical variation in extracted metabolites. To test the reproducibility of this aqueous methanol extraction for our tea samples, we performed three independent extractions for two contrasting tea products, that is, one black tea and one green tea. These extracts were subsequently analyzed by LC-Orbitrap FTMS, and the peak areas and heights were determined and coefficients of variation of 10 specific tea metabolites were calculated. For all of these selected metabolites, the coefficients of variation were on average well within 10% for both the black and green tea samples.

A total of 17 different aqueous methanol extracts from tea, including 7 black, 7 green, and 3 white teas, were subsequently analyzed using LC with online Orbitrap-FT-MS^n.[18] The Orbitrap FTMS was set to perform both full-scan MS (m/z

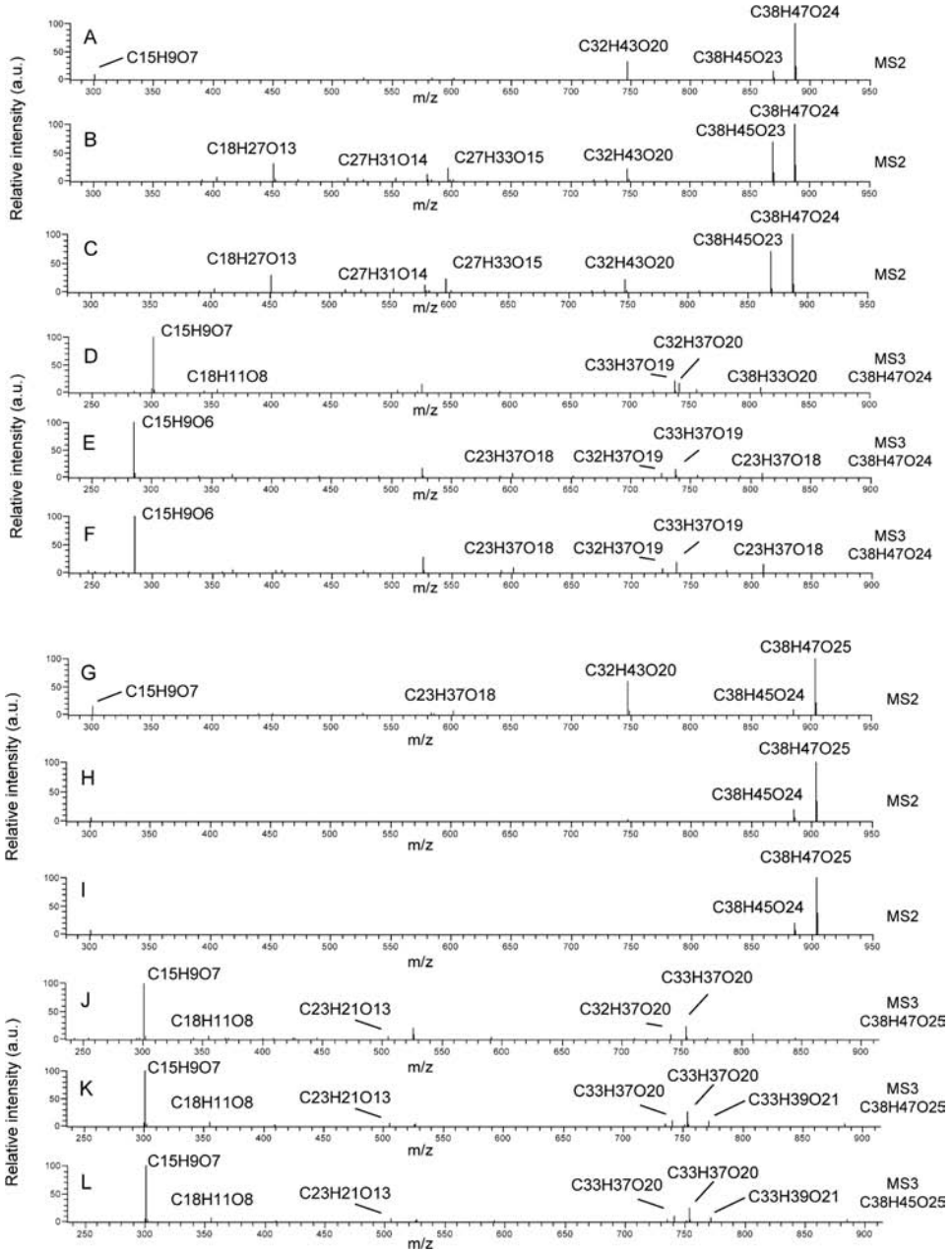

Figure 1. LC-MS^n spectra of three $C_{47}H_{54}O_{26}$ isomers (A–F) and three $C_{47}H_{54}O_{27}$ isomers (G–L). The MS^2 and MS^3 spectra of the most intense ion in MS^2 are shown for the $C_{47}H_{54}O_{26}$ isomers at RT 38.54 (A, D), at RT 40.11 (B, E), and RT 41.12 (C, F), as well as for the $C_{47}H_{54}O_{27}$ isomers at RT 36.38 (G, J), at RT 38.11 (H, K), and RT 39.83 (I, L). The m/z values that are visible but not assigned (e.g., around m/z 525) are likely to be derived of coeluting as well as cofragmenting metabolites, as no meaningful EF could be assigned to them within 5 ppm.

90–1200) and online fragmentation up to MS^3 for the three highest mass signals present in MS^2 scans. The raw data of GT China Bancha and BT Pickwick English Tea Blend teas were subsequently manually inspected, using Xcalibur software, for the presence of accurate masses (within a window of 3 ppm mass accuracy) of both parent ions and fragment ions of tea compounds previously reported in the literature, which provided a good starting point for the annotation of the phenolic content of the tea metabolome. This approach resulted in the characterization of a series of 177 phenolic compounds, mainly consisting of conjugates of the flavonols quercetin and kaempferol and the flavan-3-ols epicatechin/catechin (Supporting Information, Supplemental Table 1). The MS^n spectra did not reveal differences between catechin and epicatechin.[26] Many flavonols were substituted on their 3-hydroxyl position, which was concluded from both the presence of radical fragment ions, which are characteristic of 3-*O*-glycosides, and the fragmentation pattern of the aglycone daughter ions.[18] During initial screening for the quercetin aglycone fragment ($C_{15}H_9O_7$, m/z of 301.0354 $[M - H]^-$) in the raw LC-MS^n data of the black tea extract, another compound with an almost identical mass of m/z 300.9990, corresponding to the elemental formula (EF) of $C_{14}H_5O_8$ $[M - H]^-$, was spotted. By comparison of the fragmentation data of this molecule to available literature data we were able to identify this compound as ellagic acid.[27,28] To evaluate the relative presence of ellagic acid and the large phenolic conjugates in the different tea samples, these metabolites

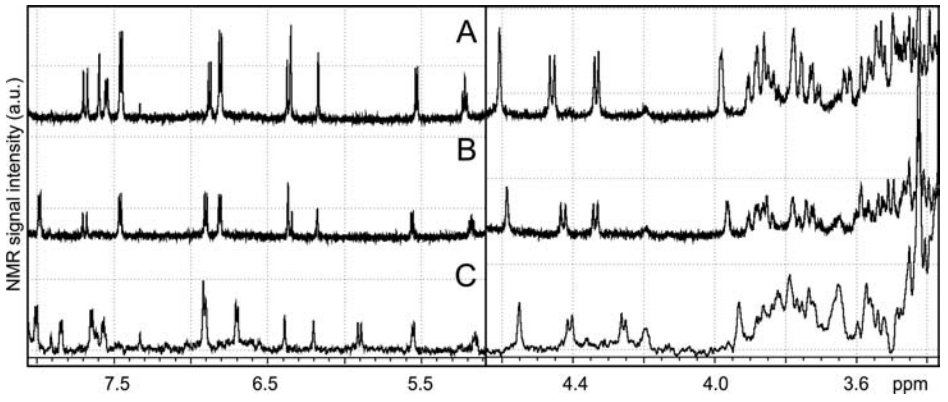

Figure 2. Two parts of the NMR spectra (8.05−5.0 ppm, left; 4.65−3.35 ppm) of quercetin 3-*O*-(2G-*p*-coumaroyl(*trans*)-3G-*O*-*β*-L-arabinosyl-3*R*-*O*-*β*-D-glucosylrutinoside (A), kaempferol 3-*O*-(2G-*p*-coumaroyl(*trans*)-3G-*O*-*β*-L-arabinosyl-3*R*-*O*-*β*-D-glucosylrutinoside (B), and kaempferol 3-*O*-(2G-*p*-coumaroyl(*cis*)-3G-*O*-*β*-L-arabinosyl-3*R*-*O*-*β*-D-glucosylrutinoside (C). Spectrum C was processed differently to enhance signal intensities (line broadening = 1.5), and the signal intensities were 4 times multiplied compared to those in spectra A and B.

were manually quantified on the basis of the area under curve of their accurate mass parent ion signal (Supplemental Table 2 in the Supporting Information).

To unambiguously elucidate the chemical structures of LC-MS^n detectable metabolites, the black tea extract of Pickwick English Tea Blend and the green tea extract of China Bancha were also used for injection into an LC-TOFMS-SPE-NMR system using the compound-specific mass signal as trigger for SPE trapping of chromatographic peaks.[17] In this manner, we trapped 23 compounds selected from black tea and 26 compounds selected from green tea, which were subsequently eluted from the SPE cartridge using deuterated methanol and subjected to 1D-^{1}H NMR analysis. Different levels of metabolite identification exist, and here we adopt the levels as defined by the Metabolomics Standards Initiative (MSI).[29,30] Tea metabolite annotations with MSI identification levels resulting from the LC-SPE-NMR approach are presented in Table 1, and additional spectral data of those metabolites can be found in Supplemental Table 1 in the Supporting Information. The combination of their MS^n and 1D-^{1}H NMR spectral data led to the full structure identification of 33 of the 49 compounds selected for LC-SPE trapping (Table 1), including compounds previously identified in tea such as epigallocatechin-3-*O*-gallate, kaempferol-3-*O*-glucoside, and myricetin-3-*O*-galactoside, as well as cis and trans isomers of five *p*-coumaroyl quinic acid isomers, which could not be differentiated on the basis of MS^n only.[31] Thus, NMR data were required for full identification and confirmation of the *p*-coumaroyl trans and cis isomers. In some cases, mixtures of two or more compounds were observed in the 1D-^{1}H NMR spectrum of a trapped LC-MS peak. However, comparison of the NMR signal integrals and using the structural information of the LC-MS^n data led to additional identification of the (minor) coeluting compounds.

We observed four isomers with an EF of $C_{27}H_{30}O_{15}$ present in all 17 tea samples. Two of them (at RT 29.42 and 30.85 min) could be annotated as kaempferol-3-*O*-(rhamnosyl-(1,6)-*O*-galactoside) and kaempferol-3-*O*-rutinoside (= rhamnosyl-1,6-glucoside), respectively, on the basis of their MS^n fragmentation patterns, elution order, and previous SPE-NMR-based identification of kaempferol-3-*O*-rutinoside in tomato fruit.[17] Compared to these two kaempferol-*O*-glycosides, the other two isomers (RT 20.71 and 25.41 min) showed completely different fragmentation spectra. Their MS^2 spectra showed common fragments such as *m*/*z* of 563.1408 (EF $C_{26}H_{27}O_{14}$, $[M - H]^-$), indicating that these two metabolites are structurally related. As their NMR spectra showed aromatic and sugar proton NMR resonances indicative of a flavonoid-*C*-glycoside, we could tentatively identify these compounds as apigenin-6,8-C-diglycosides.[7]

In the retention time window of 37−42 min, the LC-MS^n data of several tea extracts revealed many doubly charged ions around *m*/*z* 510 and their corresponding singly charged ions above *m*/*z* 1000. Although the presence of these doubly charged species partly hampered fragmentation of the singly charged parent metabolites, we could obtain MS^n spectra for three $C_{47}H_{53}O_{26}$ isomers (Figure 1A−F) and three $C_{47}H_{53}O_{27}$ isomers (Figure 1G−L). From the similarities in the MS^n spectra and the small differences in elution times, we could conclude that these compounds share many structural elements. This is illustrated by spectra 1B and 1E that show similar fragmentation patterns compared to spectra 1C and 1F, indicating that these kaempferol-based conjugates are structurally highly related. From the MS^n data, it can be observed that the third $C_{47}H_{53}O_{26}$ isomer (spectra 1A and 1D) is a conjugate of quercetin. The same systematic analysis was performed for the three quercetin-based $C_{47}H_{53}O_{27}$ isomers (Figure 1G−L).

To fully identify these complex conjugates, we trapped them using mass-based SPE and subsequently obtained almost pure NMR spectra for three of them (Figure 2). The SPE-NMR fraction of the $C_{47}H_{53}O_{27}$ isomer eluting at RT 38.11 min yielded NMR signals that were in very good agreement with those previously reported for quercetin-3-*O*-(2G-*p*-coumaroyl-(*trans*)-3G-*O*-*β*-L-arabinosyl-3*R*-*O*-*β*-D-glucosylrutinoside[32] (Figure 2A). Additional 2D-^{1}H COSY experiments confirmed its structure. Moreover, on the basis of the similarities observed in both MS^n fragmentation patterns and NMR signals (Figures 1 and 2), we could unambiguously elucidate the structures of two acylated kaempferol tetraglycosides: kaempferol-3-*O*-(2G-*p*-coumaroyl(*trans*)-3G-*O*-*β*-L-arabinosyl-3*R*-*O*-*β*-D-glucosylrutinoside and its *p*-coumaroyl cis isomer (Figure 2B,C). On the basis of its RT and LC-MS^n pattern alone, we could tentatively assign the quercetin conjugate eluting at RT 39.83 min as the cis isomer of quercetin-3-*O*-(2G-*p*-coumaroyl(*trans*)-3G-*O*-*β*-L-arabinosyl-3*R*-*O*-*β*-D-glucosylrutinoside.

The $C_{47}H_{53}O_{27}$ isomer eluting at 36.38 min is most likely a positional isomer of quercetin-3-*O*-(2G-*p*-coumaroyl(*trans*)-3G-*O*-*β*-L-arabinosyl-3*R*-*O*-*β*-D-glucosylrutinoside, because it

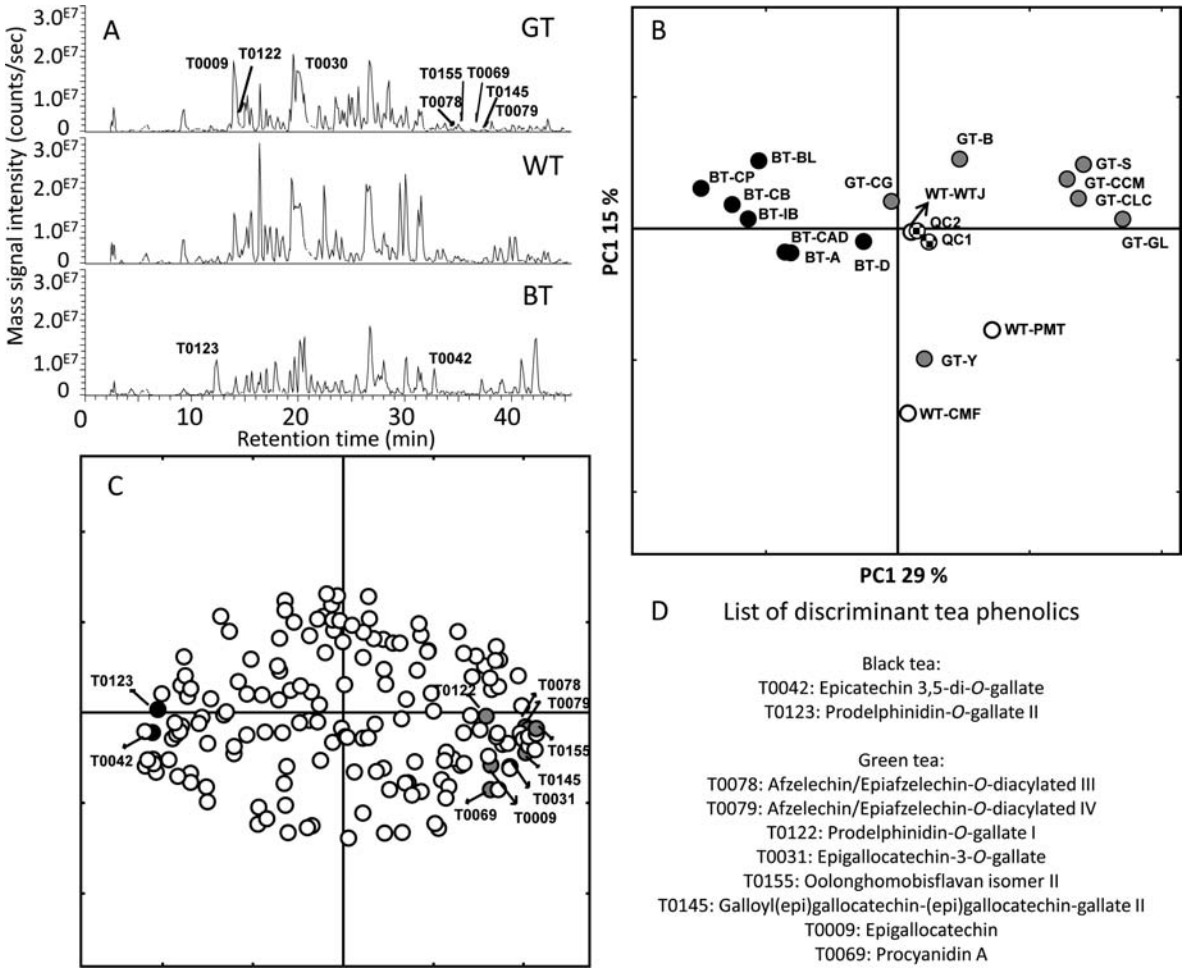

Figure 3. (A) RT 0−45 min window of three representative full-scan LC-MS profiles of green tea (GT-CG), white tea (WT-CMF), and black tea (BT-IB) with discriminative metabolites indicated. The mass signal intensities were all set to a fixed scale. Scoring (B) and loading (C) plots of unsupervised principal component analysis of 166 annotated metabolites, in which the significant discriminative metabolites are indicated in gray. (D) List of discriminative phenolic metabolites for black and green tea found in our study.

yields similar fragment ions, but in different relative intensities, and upon fragmentation of the $C_{32}H_{43}O_{20}$ fragment, the loss of its coumaroyl group was observed. On the basis of both MS^n fragmentation patterns and diagnostic NMR signals, we identified the $C_{47}H_{53}O_{27}$ isomer eluting at RT 38.54 min as being quercetin 3-*O*-(2G-*p*-coumaroyl(*trans*)-3G-*O*-*β*-L-arabinosyl-3*R*-*O*- *β*-D-rhamnosylrutinoside. Furthermore, we identified kaempferol-3-*O*-(2G-*p*-coumaroyl(*trans*)-3G-*O*-*β*-D-glucosyl-3*R*-*O*-*β*-D-glucosylrutinoside because it eluted slightly earlier than its arabinose-substituted analogue and the conjugate had an observed mass of *m*/*z* 1063.2933 (EF $C_{48}H_{55}O_{27}$ $[M - H]^-$) corresponding to the difference of an arabinose and a glucose moiety.

The $LC\text{-}MS^n$ data of all acylated complex polyphenol conjugates in the tea extracts revealed two losses of $C_9H_6O_2$ (*m*/*z* 146.0372) and $C_9H_8O_3$ (*m*/*z* 164.0478) (Figure 1). The latter is indicative for a coumaroyl loss of complex acylated conjugates.[33] Analogously, the loss of $C_9H_6O_2$ appeared to be indicative for cinnamic acid conjugation. For example, two quercetin-based acylated glycosides were observed with *m*/*z* 609.1248 (EF $C_{30}H_{25}O_{14}$ $[M - H]^-$) and *m*/*z* 771.1779 (EF $C_{36}H_{35}O_{19}$ $[M - H]^-$), together with their kaempferol analogues. All of them were tentatively annotated on the basis of similar fragmentation but mass difference corresponding to one oxygen atom and their elution order, with quercetin analogues eluting earlier than kaempferol analogues[18] and coumaryol analogues eluting earlier than cinnamic analogues.

Representative LC-MS chromatograms of the three different types of tea are shown in Figure 3A. The chromatogram of black tea clearly differs from those of green and white tea, and also differences exist between the white and green tea extracts. To compare the phenolic profiles and pinpoint those compounds that are mostly responsible for the differences between the 17 teas analyzed within this study, the abundance (peak height) of all annotated compounds (Table 1) was automatically extracted from the raw LC-MS data using the Search-LCMS module of Metalign software, on the basis of the accurate mass of the parent ions and their specific retention times. In this manner, the relative abundance of 166 of the 177 annotated compounds could be retrieved; 11 compounds were missed as they were partly coeluting with an isomeric compound (having the same accurate mass), which did not allow automatic peak picking. The relative intensities of these 166 compounds were subsequently used as variables to perform PCA on the extracts from the 17 different tea products (Figure 3B). The two technical replicates, which consisted of replicate analysis of a mix of all extracts, were positioned close to each other in the middle of the PCA plot, which indicates good

technical reproducibility of the LC-MS analysis followed by the automatic peak picking and annotation procedure. The PCA plot showed a separation of, on the one hand, all seven black teas and, on the other hand, the green and white teas (PC1, explaining 29% of the total metabolite variation). Within the group of green and white teas, a group of four green tea extracts clustered separately from the others, whereas the three white teas were spread between the green teas (Figure 3B). The metabolites that were significantly differential between the three tea types (ANOVA, $p < 0.01$) are provided in Figure 3D, and their presence is indicated in the chromatograms of the corresponding teas (Figure 3A). No significant differences were found between the white teas and the green teas. Eighty-two compounds were present (signal-to-noise ratio > 3) in all 17 extracts, and the relative abundance over the samples ranged from a factor 1.86 for gallocatechin-*O*-gallate to a factor 2690 for quercetin-3-*O*-(glucosyl(1−3)rhamnosyl(1−6)rhamnoside, indicating large differences between teas in their levels of individual phenolic compounds.

DISCUSSION

Using an on-the-fly multistage accurate mass MS (MS^n) approach paired with LC coupled to MS-based SPE-NMR, we aimed to improve the current tea metabolome coverage by improving the identification levels of a large series of polyphenols without the need for reference compounds. We were able to annotate a total of 177 phenolic metabolites in black and green tea samples (Supporting Information, Supplemental Table 1). Of these annotated compounds, 82 have not been described in black and green tea extracts before, at least not at their present level of structural elucidation. We subsequently determined the variation in relative abundance of the identified phenolic compounds in 17 different tea samples, representing black, green, and white teas, by using the mass specific LC-MS peak of selected large acylated conjugates (Supporting Information, Supplemental Table 2) as well as unsupervised multivariate analysis (PCA) based on 166 annotated phenolic metabolites. The results indicate highly variable levels in investigated tea samples for most compounds.

Full-scan accurate mass is helpful in the MS analysis due to fast assignment of elemental formulas.[34] Moreover, mass fragmentation will help to obtain more structural information to enable annotation of metabolites to the correct compound class.[18] For example, Zhao et al.[35] observed an unknown compound of m/z 499 during their UPLC-DAD-MS analysis. Our analysis using Orbitrap FTMS indicated three isomers of m/z 499.0883 in the retention time region of 33−50 min, and their fragmentation patterns using $LC\text{-}MS^n$ led to their tentative annotation as *p*-coumaroyl-caffeoylquinic acid conjugates.[18] Due to relatively low levels of these quinic acid conjugates in our samples and the large amounts of compounds coeluting within the specific RT region, it was not possible to obtain NMR spectra for their full identification. The power of accurate mass $LC\text{-}MS^n$ is further illustrated by the fragmentation of a GT China Chun Mee metabolite (although not observed or identified by Zhao et al.[35]) at an m/z of 269.0455 (EF $C_{15}H_9O_5$ $[M - H]^-$) that matched with the fragmentation of apigenin.[17]

Unequivocal identification of metabolites remains a bottleneck in many metabolomics studies,[36,37] as often large numbers of tentatively identified compounds are reported. Even though many tea metabolites have previously been fully identified, it is still hard to unambiguously annotate them. Currently, there is no single experimental database that provides all chemical knowledge of a sample, which hampers a fast identification based on a combination of observed retention time, UV absorption, MS spectra, and NMR signals. For example, the acylated quercetin tetraglycoside with observed $[M - H]^-$ mass of 1033.2827 was already identified in an oolong tea extract in 2004,[32] but was not fully annotated in later studies, most likely due to limited structural information.[7] We have previously shown that our HPLC-MS-SPE-NMR approach enabled structural elucidation of quercetin and kaempferol glycosides, including several isomers.[17] Indeed, in the present tea study we obtained good-quality 1D-^{1}H NMR data for 25 flavonoid glycosides and for three complex acylated flavonoid tetraglycosides. Nearly all flavonoids detected were specifically conjugated at the 3-*O* position. Only one kaempferol-triglycoside (observed m/z 739.2200; EF $C_{33}H_{39}O_{19}$ $[M - H]^-$) showed a different fragmentation behavior: its triglycoside MS^2 spectrum displayed a fragment ion of $C_{27}H_{29}O_{15}$ that corresponds to a single deoxyhexoside substitution at either the 7-*O* or the 4′-*O* position.[18] Combined with its 1D-^{1}H NMR data, generated after LC-TOFMS-SPE trapping, its chemical structure was elucidated as kaempferol-3-*O*-(glucoside-1,6-*O*-rhamnoside)-4′-*O*-rhamnoside.

The combination of $LC\text{-}MS^n$ and NMR spectral data proved to be of great help in unraveling the structures of highly complex conjugated phenolics, also of compounds with m/z >800 amu (Figures 1 and 2). Although these large molecules were partly detected as doubly charged ion species by MS, our $LC\text{-}MS^n$ approach was able to collect sufficient MS^n data for structural elucidation purposes. Upon combination of analytical data such as the elution order from the reversed phase column, their specific MS^n fragmentation patterns and their characteristic proton NMR signals, a few of these large molecules could be unambiguously identified (i.e., up to MSI metabolite identification level 1). In a similar manner, we could identify (MSI level 1) three coeluting compounds that showed characteristic MS^n fragmentation patterns while being visible as only minor compounds in the 1D-^{1}H NMR spectra. This shows that the combination of $LC\text{-}MS^n$ and HPLC-MS-SPE-NMR platforms enables the analysis of small amounts of compounds present in a highly complex extract such as tea. For unambiguous identification we currently used 100 mg of dried tea product, which is much less than in other tea metabolite identification studies taking a few grams up to kilograms of plant sample.[20,32,38,39] In fact, for a few metabolites, depending upon their ionization efficiency in negative electrospray ionization mode and the intensity of specific NMR signals obtained within the 15 min of scan time, we needed <2 μg of compound for full structural characterization.

The success of structural elucidation of oligomers of catechin and epicatechin, as well as of theaflavins, was relatively limited in our current study. Although these compounds were clearly detectable by LC-MS and their presence in tea has been reported,[4] our ^{1}H NMR measurements did not show any diagnostic aromatic signals for these phenolic molecules. Nevertheless, on the basis of their MS^n spectra, conjugated dimers of catechin or epicatechin, such as oolonghomobisflavan isomers[40] (observed m/z 927.1636; EF $C_{45}H_{35}O_{22}$ $[M - H]^-$), and isomeric digalloyl substituted (epi)gallocatechin dimers (observed m/z 911.1678; EF $C_{45}H_{35}O_{21}$ $[M - H]^-$), could be annotated. Perhaps, to obtain good NMR data, an SPE sorbent different from the HLB material used in the present study could be an alternative for concentration and trapping purposes, as

these compounds at least showed retention on the analytical C18 HPLC column.

The LC-MS-SPE-NMR approach could trap even very polar compounds, such as 3-*O*-galloylquinic acid, in a sufficient manner to obtain ^{1}H NMR signals to compare with literature data.[41] Additionally, we could obtain NMR data for a series of *p*-coumaroylquinic acids. These compounds, especially the trans (*E*) isomers, have been previously studied by MS^n, and the obtained fragmentation data match well with those reported in the literature.[42−44] However, the cis (*Z*) isomers are less well described and show the same fragmentation behavior as the corresponding *E* isomers.

The power of our analytical approach, using LC-MS^n structural information combined with ^{1}H NMR of MS triggered LC-SPE trapped compounds, is emphasized by the fact that no time-consuming 2D NMR spectra were necessary to assign the diagnostic proton signals of NMR spectra of the complex acylated conjugates, even at the low amount of analyte in the NMR probe of 2 and 1 μg, respectively, for the quercetin and kaempferol analogues substituted by 3-*O*-(2G-*p*-coumaroyl-(*trans*)-3G-*O*-*β*-L-arabinosyl-3*R*-*O*-*β*-D-glucosylrutinoside. Moreover, the *p*-coumaroyl cis isomer of kaempferol was trapped at even a much lower amount (about 0.2 μg); nevertheless, full identification of this molecule was possible due to the high similarity of its MS^n fragmentation patterns (Figure 1) and NMR signals (Figure 2) with its trans isomer and the observation of the diagnostic *α* proton of the *p*-coumaroyl moiety at lower ppm resonance (5.93 vs 6.38 ppm) and a smaller coupling constant (13 vs 16 Hz), as was seen in related acylated quercetin glycosides.[45] Literature data also reported the presence of more of these complex conjugates in tea based on LC-MS,[7,35] as well as NMR.[32,38] In our study, 1D-^{1}H NMR spectra were needed to fully unravel and confirm their structures and enable MSI level 1 annotations.

Ellagic acid was described in tea in 1941,[46] but has never been reported in tea samples since then until now. Ellagic acid was detected in all teas, but was present in relatively higher amounts in all black teas compared to all green and white teas except two. Additionally, in all tea extracts we observed a molecule closely related to ellagic acid, which was tentatively identified as galloyl-bis-HHDP-glucose (observed *m*/*z* 633.0737; EF $C_{27}H_{21}O_{18}$ $[M - H]^-$). Upon MS^n fragmentation of this molecule its fragment ion $C_{14}H_5O_8$ had a similar fragmentation pattern as ellagic acid,[47] showing their metabolite relationship.[18]

Classification of different teas is greatly influenced by their phenolic content, as was reported before on the basis of flavan-3-ols such as epigallocatechin.[35,48,49] Four isomeric digalloyl substituted (epi)gallocatechin dimers and three oolonghomo-bisflavan isomers were detected in all tea extracts (Supplemental Table 2 of the Supporting Information). The smallest variations in phenolic compounds between the 17 tea samples analyzed (factor 2 or lower) were found for gallocatechin-3-*O*-gallate, caffeine, and epicatechin-3-*O*-gallate, suggesting that these compounds are widely distributed and do not greatly differ in their abundance among different commercial samples. The largest variations between samples (factor 480 or higher) were observed for quercetin-3-*O*-glucosyl-(1−3)-rhamnosyl-(1−6)-rhamnoside, kaempferol-3-*O*-(2G-*p*-coumaroyl(*trans*)-3G-*O*-L-arabinosyl-3*R*-*O*-D-glucosylrutinoside, and quercetin-based acylated triglycoside $C_{42}H_{42}O_{22}$ V (present in all 17 tea extracts) and for quercetin-3-*O*-glucosyl-(1−3)-rhamnosyl-(1−6)-galactoside, *p*-*trans*-coumaroylquinic acid-5-*O*, and chalcane flavan-3-ol (gambiriin) I (present in 16 of the 17 tea extracts).

Small flavan-3-ol conjugates such as epigallocatechin-3-*O*-gallate and epigallocatechin were characteristic for green tea (Figure 3). In fact, the metabolite differences that separate black and green/white teas can be correlated to changes occurring during postharvest fermentation, as is supported by previous studies showing a decrease in free flavan-3-ols and their gallated conjugates during fermentation from green to black tea.[50,51] Interestingly, the isomers prodelphinidin-*O*-gallate II and prodelphinidin-*O*-gallate I were most characteristic for green and black tea, respectively. Unfortunately, we were not able to elucidate the structural differences between both isomeric compounds. The relative similarity between white and green teas in their metabolic composition (Figure 3B) may be partly due to the fact that both types are not fermented during after harvest.

In conclusion, our analytical approach generated highly comprehensive tea phenolic profiles and enabled annotation of 177 metabolites, of which 82 have not previously been reported at such an annotation level in black or green tea. Without the need for reference compounds and using only 100 mg of sample, we could unambiguously structurally elucidate 33 tea metabolites, using a combination of specific MS^n fragmentation patterns and NMR signals. In the present study, the identified metabolites were used to compare the phenolic profiles of 17 tea samples and to pinpoint differential metabolites. We are confident that these advanced metabolite identification approaches will increase our knowledge of the differences in chemical composition of tea samples, for instance, in relation to genotype, growth conditions, postharvest processing, and preparation of tea drinks, as well as chemical profiles of material from other biological sources.

SAFETY

There were no special safety concerns for this study; all chemicals in this study were used according their safety precautions.

ASSOCIATED CONTENT

Supporting Information

Supplemental Table 1: 177 annotated metabolites. In cases when NMR spectra were obtained (column D, SPE-NMR), additional spectral data are provided (columns L-O). Metabolites were sorted on the basis of their accurate mass ($[M - H]^-$ (*m*/*z*), column G) and finally on their retention times (RT Orbitrap, column E). Supplemental Table 2: 51 metabolites that were relatively quantified in all 17 tea extracts on the basis of their area under the curve of the specific parent mass signal and their relative areas in the different samples. Metabolites were sorted on the basis of their accurate mass ($[M - H]^-$ (*m*/*z*), column D) and finally on their retention times (RT Orbitrap, column B). This material is available free of charge via the Internet at http://pubs.acs.org.

AUTHOR INFORMATION

Corresponding Author

*Phone: +31 317 482 620. Fax: +31 317 484 801. E-mail: justin.vanderhooft@wur.nl.

Funding

This research was granted by The Netherlands Metabolomics Centre and the Centre for BioSystems Genomics, both of

which are part of The Netherlands Genomics Initiative/Netherlands Organization for Scientific Research.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

We thank the tea shop Delicious in Hoogland, The Netherlands, for providing batches of tea. We also thank Dr. Saiko Yoshida for her help.

■ ABBREVIATIONS USED

BT, black tea; DAD, diode array detector; EF, elemental formula; FT, Fourier transformed; GAL, galactose moiety; GLC, glucose moiety; GT, green tea; HHDP, 6,6′-dicarbonyl-2,2′,3,3′,4,4′-hexahydroxybiphenyl moiety; *m/z*, mass to charge; MS, mass spectrometry; NMR, nuclear magnetic resonance; MI, metabolite identification; MSI, Metabolomics Standards Initiative; PCA, principal component analysis; TOF, time of flight; RT, retention time; RHM, rhamnose moiety; UV, ultraviolet; WT, white tea.

■ REFERENCES

(1) Sang, S.; Lambert, J. D.; Ho, C.-T.; Yang, C. S. The chemistry and biotransformation of tea constituents. *Pharmacol. Res.* **2011**, *64*, 87−99.

(2) Hodgson, J. M.; Croft, K. D. Tea flavonoids and cardiovascular health. *Mol. Asp. Med.* **2010**, *31*, 495−502.

(3) Yang, C. S.; Wang, H.; Li, G. X.; Yang, Z.; Guan, F.; Jin, H. Cancer prevention by tea: evidence from laboratory studies. *Pharmacol. Res.* **2011**, *64*, 113−122.

(4) Peterson, J.; Dwyer, J.; Bhagwat, S.; Haytowitz, D.; Holden, J.; Eldridge, A. L.; Beecher, G.; Aladesanmi, J. Major flavonoids in dry tea. *J. Food Compos. Anal.* **2005**, *18*, 487−501.

(5) Kuhnert, N.; Drynan, J. W.; Obuchowicz, J.; Clifford, M. N.; Witt, M. Mass spectrometric characterization of black tea thearubigins leading to an oxidative cascade hypothesis for thearubigin formation. *Rapid Commun. Mass Spectrom.* **2011**, *24*, 3387−3404.

(6) Drynan, J. W.; Clifford, M. N.; Obuchowicz, J.; Kuhnert, N. The chemistry of low molecular weight black tea polyphenols. *Nat. Prod. Rep.* **2010**, *27*, 417−462.

(7) Lin, L.-Z.; Chen, P.; Harnly, J. M. New phenolic components and chromatographic profiles of green and fermented teas. *J. Agric. Food Chem.* **2008**, *56*, 8130−8140.

(8) Manach, C.; Hubert, J.; Llorach, R.; Scalbert, A. The complex links between dietary phytochemicals and human health deciphered by metabolomics. *Mol. Nutr. Food Res.* **2009**, *53*, 1303−1315.

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

(10) Del Rio, D.; Stewart, A. J.; Mullen, W.; Burns, J.; Lean, M. E. J.; Brighenti, F.; Crozier, A. HPLC-MS^n analysis of phenolic compounds and purine alkaloids in green and black tea. *J. Agric. Food Chem.* **2004**, *52*, 2807−2815.

(11) Wang, D.; Lu, J.; Miao, A.; Xie, Z.; Yang, D. HPLC-DAD-ESI-MS/MS analysis of polyphenols and purine alkaloids in leaves of 22 tea cultivars in China. *J. Food Compos. Anal.* **2008**, *21*, 361−369.

(12) Wu, C.; Xu, H.; Héritier, J.; Andlauer, W. Determination of catechins and flavonol glycosides in Chinese tea varieties. *Food Chem.* **2012**, *132*, 144−149.

(13) Dou, J.; Lee, V. S. Y.; Tzen, J. T. C.; Lee, M.-R. Identification and comparison of phenolic compounds in the preparation of oolong tea manufactured by semifermentation and drying processes. *J. Agric. Food Chem.* **2007**, *55*, 7462−7468.

(14) Peterson, J.; Dwyer, J.; Jacques, P.; Rand, W.; Prior, R.; Chui, K. Tea variety and brewing techniques influence flavonoid content of black tea. *J. Food Compos. Anal.* **2004**, *17*, 397−405.

(15) Wang, Y.; Li, Q.; Wang, Q.; Li, Y.; Ling, J.; Liu, L.; Chen, X.; Bi, K. Simultaneous determination of seven bioactive components in oolong tea *Camellia sinensis*: a quality control by chemical composition and HPLC fingerprints. *J. Agric. Food Chem.* **2011**, *60*, 256−260.

(16) Karaköse, H.; Jaiswal, R.; Kuhnert, N. Characterization and quantification of hydroxycinnamate derivatives in *Stevia rebaudiana* leaves by LC-MS^n. *J. Agric. Food Chem.* **2011**, *59*, 10143−10150.

(17) Van der Hooft, J. J. J.; Mihaleva, V.; Bino, R. J.; de Vos, R. C. H.; Vervoort, J. A strategy for fast structural elucidation of metabolites in small volume plant extracts using automated MS-guided LC-MS-SPE-NMR. *Magn. Reson. Chem.* **2011**, *49* (S1), S55−S60.

(18) Van der Hooft, J. J. J.; Vervoort, J.; Bino, R. J.; de Vos, R. C. H. Spectral trees as a robust annotation tool in LC-MS based metabolomics. *Metabolomics* **2011**, DOI: 10.1007/s11306-011-0363-7.

(19) Moco, S.; Bino, R. J.; Vorst, O.; Verhoeven, H. A.; de Groot, J.; van Beek, T. A.; Vervoort, J.; de Vos, C. H. R. A liquid chromatography-mass spectrometry-based metabolome database for tomato. *Plant Physiol.* **2006**, *141*, 1205−1218.

(20) Scharbert, S.; Holzmann, N.; Hofmann, T. Identification of the astringent taste compounds in black tea infusions by combining instrumental analysis and human bioresponse. *J. Agric. Food Chem.* **2004**, *52*, 3498−3508.

(21) Lommen, A. MetAlign: interface-driven, versatile metabolomics tool for hyphenated full-scan mass spectrometry data preprocessing. *Anal. Chem.* **2009**, *81*, 3079−3086.

(22) deVos, C. H. R.; Moco, S.; Lommen, A.; Keurentjes, J. J.; Bino, R. J.; Hall, R. D. Untargeted large-scale plant metabolomics using liquid chromatography coupled to mass spectrometry. *Nat. Prot.* **2007**, *2*, 778−791.

(23) Tikunov, Y. M.; de Vos, R. C. H.; Paramás, A. M. G.; Hall, R. D.; Bovy, A. G. A role for differential glycoconjugation in the emission of phenylpropanoid volatiles from tomato fruit discovered using a metabolic data fusion approach. *Plant Physiol.* **2010**, *152*, 55−70.

(24) Lommen, A.; Gerssen, A.; Oosterink, J. E.; Kools, H. J.; Ruiz-Aracama, A.; Peters, R. J. B.; Mol, H. G. J. Ultra-fast searching assists in evaluating sub-ppm mass accuracy enhancement in U-HPLC/Orbitrap MS data. *Metabolomics* **2011**, *7*, 15−24.

(25) Le Gall, G.; Colquhoun, I. J.; Defernez, M. Metabolite profiling using H-1 NMR spectroscopy for quality assessment of green tea, *Camellia sinensis* (L.). *J. Agric. Food Chem.* **2004**, *52*, 692−700.

(26) Van der Hooft, J. J. J.; Vervoort, J.; Bino, R. J.; Beekwilder, J.; De Vos, R. C. H. Polyphenol identification based on systematic and robust high-resolution accurate mass spectrometry fragmentation. *Anal. Chem.* **2011**, *83*, 409−416.

(27) Lee, J.-H.; Johnson, J. V.; Talcott, S. T. Identification of ellagic acid conjugates and other polyphenolics in muscadine grapes by HPLC-ESI-MS. *J. Agric. Food Chem.* **2005**, *53*, 6003−6010.

(28) Mullen, W.; Yokota, T.; Lean, M. E. J.; Crozier, A. Analysis of ellagitannins and conjugates of ellagic acid and quercetin in raspberry fruits by LC-MS^n. *Phytochemistry* **2003**, *64*, 617−624.

(29) Sumner, L. W.; Amberg, A.; Barrett, D.; Beale, M. H.; Beger, R.; Daykin, C. A.; Fan, T. W. M.; Fiehn, O.; Goodacre, R.; Griffin, J. L.; Hankemeier, T.; Hardy, N.; Harnly, J.; Higashi, R.; Kopka, J.; Lane, A. N.; Lindon, J. C.; Marriott, P.; Nicholls, A. W.; Reily, M. D.; Thaden, J. J.; Viant, M. R. Proposed minimum reporting standards for chemical analysis. *Metabolomics* **2007**, *3*, 211−221.

(30) Molyneux, R. J.; Schieberle, P. Compound identification: a *Journal of Agricultural and Food Chemistry* perspective. *J. Agric. Food Chem.* **2007**, *55*, 4625−4629.

(31) Clifford, M. N.; Kirkpatrick, J.; Kuhnert, N.; Roozendaal, H.; Salgado, P. R. LC-MS^n analysis of the cis isomers of chlorogenic acids. *Food Chem.* **2008**, *106*, 379−385.

(32) Mihara, R.; Mitsunaga, T.; Fukui, Y.; Nakai, M.; Yamaji, N.; Shibata, H. A novel acylated quercetin tetraglycoside from oolong tea (*Camellia sinensis*) extracts. *Tetrahedron Lett.* **2004**, *45*, 5077−5080.

(33) Harbaum, B.; Hubbermann, E. M.; Wolff, C.; Herges, R.; Zhu, Z.; Schwarz, K. Identification of flavonoids and hydroxycinnamic acids in pak choi varieties (*Brassica campestris* L. ssp. *chinensis* var.

communis) by HPLC-ESI-MS^n and NMR and their quantification by HPLC-DAD. *J. Agric. Food Chem.* **2007**, *55*, 8251−8260.
(34) Kind, T.; Fiehn, O. Advances in structure elucidation of small molecules using mass spectrometry. *Bioanal. Rev.* **2010**, *2* (1), 23−60.
(35) Zhao, Y.; Chen, P.; Lin, L.; Harnly, J. M.; Yu, L.; Li, Z. Tentative identification, quantitation, and principal component analysis of green pu-erh, green, and white teas using UPLC/DAD/MS. *Food Chem.* **2011**, *126*, 1269−1277.
(36) Moco, S.; Bino, R. J.; De Vos, R. C. H.; Vervoort, J. Metabolomics technologies and metabolite identification. *TrAC−Trends Anal. Chem.* **2007**, *26*, 855−866.
(37) Scalbert, A.; Andres-Lacueva, C.; Arita, M.; Kroon, P.; Manach, C.; Urpi-Sarda, M.; Wishart, D. Databases on food phytochemicals and their health-promoting effects. *J. Agric. Food Chem.* **2011**, *59*, 4331−4348.
(38) Lee, V. S.-Y.; Chen, C.-R.; Liao, Y.-W.; Tzen, J. T.-C.; Chang, C.-I. Structural determination and DPPH radical-scavenging activity of two acylated flavonoid tetraglycosides in oolong tea (*Camellia sinensis*). *Chem. Pharm. Bull.* **2008**, *56*, 851−853.
(39) Lu, J.-h.; Liu, Y.; Zhao, Y.-y.; Tu, G.-z. New flavonoids from *Oxytropis myriophylla*. *Chem. Pharm. Bull.* **2004**, *52*, 276−278.
(40) Hashimoto, F.; Nonaka, G. I.; Nishioka, I. Tannins and related compounds. XC. 8-*C*-Ascorbyl (−)-epigallocatechin 3-*O*-gallate and novel dimeric flavan-3-ols, oolonghomobisflavans A and B, from oolong tea (3). *Chem. Pharm. Bull.* **1989**, *37*, 3255−3263.
(41) Nishimura, H.; Nonaka, G.-I.; Nishioka, I. Seven quinic acid gallates from *Quercus stenophylla*. *Phytochemistry* **1984**, *23*, 2621−2623.
(42) Clifford, M. N.; Johnston, K. L.; Knight, S.; Kuhnert, N. Hierarchical scheme for LC-MS^n identification of chlorogenic acids. *J. Agric. Food Chem.* **2003**, *51*, 2900−2911.
(43) Lu, Y.; Sun, Y.; Foo, L. Y.; McNabb, W. C.; Molan, A. L. Phenolic glycosides of forage legume *Onobrychis viciifolia*. *Phytochemistry* **2000**, *55*, 67−75.
(44) Plazonić, A.; Bucar, F.; Maleš, Ž.; Mornar, A.; Nigović, B.; Kujundžić, N. Identification and quantification of flavonoids and phenolic acids in burr parsley (*Caucalis platycarpos* L.), using high-performance liquid chromatography with diode array detection and electrospray ionization mass spectrometry. *Molecules* **2009**, *14*, 2466−2490.
(45) Diaz, J. G.; Herz, W. Acylated flavonol tetraglycosides from *Delphinium gracile*. *Phytochemistry* **2010**, *71*, 463−468.
(46) Tsujimura, M. Ellagic acid in green tea. *Sci. Pap. IPCR* **1941**, *38*, 487−489.
(47) Kajdžanoska, M.; Gjamovski, V.; Stefova, M. HPLC-DAD-ESI-MS^n identification of phenolic compounds in cultivated strawberries from Macedonia. *Macedon. J. Chem. Chem. Eng.* **2010**, *29*, 181−194.
(48) Li, J.-H.; Nesumi, A.; Shimizu, K.; Sakata, Y.; Liang, M.-Z.; He, Q.-Y.; Zhou, H.-J.; Hashimoto, F. Chemosystematics of tea trees based on tea leaf polyphenols as phenetic markers. *Phytochemistry* **2011**, *71*, 1342−1349.
(49) Xie, G.; Ye, M.; Wang, Y.; Ni, Y.; Su, M.; Huang, H.; Qiu, M.; Zhao, A.; Zheng, X.; Chen, T.; Jia, W. Characterization of pu-erh tea using chemical and metabolic profiling approaches. *J. Agric. Food Chem.* **2009**, *57*, 3046−3054.
(50) Lee, J.-E.; Lee, B.-J.; Chung, J.-O.; Shin, H.-J.; Lee, S.-J.; Lee, C.-H.; Hong, Y.-S. ^{1}H NMR-based metabolomic characterization during green tea (*Camellia sinensis*) fermentation. *Food Res. Int.* **2011**, *44*, 597−604.
(51) Ku, K. M.; Kim, J.; Park, H.-J.; Liu, K.-H.; Lee, C. H. Application of metabolomics in the analysis of manufacturing type of pu-erh tea and composition changes with different postfermentation year. *J. Agric. Food Chem.* **2010**, *58*, 345−352.