JOURNAL OF AGRICULTURAL AND FOOD CHEMISTRY

Article

pubs.acs.org/JAFC

Profiles of Phenolic Compounds and Purine Alkaloids during the Development of Seeds of *Theobroma cacao* cv. Trinitario

Gema Pereira-Caro,[†] Gina Borges,[†] Chifumi Nagai,[‡] Mel C. Jackson,[‡] Takao Yokota,[§] Alan Crozier,[†] and Hiroshi Ashihara*[,⊥]

[†]Joseph Black Building, School of Medicine, College of Medical, Veterinary and Life Sciences, University of Glasgow, Glasgow G12 8QQ, United Kingdom

[‡]Hawaii Agriculture Research Center (HARC), P.O. Box 100, Kunia, Hawaii, Hawaii 96759-0100, United States

[§]Department of Biosciences, Teikyo University, 1-1 Toyosatodai, Utsunomiya 320-8551, Japan

[⊥]Department of Biological Sciences, Graduate School of Humanities and Science, Ochanomizu University, Bunkyo-ku, Tokyo 112-8610, Japan

ABSTRACT: Changes occurring in phenolic compounds and purine alkaloids, during the growth of seeds of cacao (*Theobroma cacao*) cv. Trinitario, were investigated using HPLC-MS/MS. Extracts of seeds with a fresh weight of 125, 700, 1550, and 2050 mg (stages 1−4, respectively) were analyzed. The phenolic compounds present in highest concentrations in developing and mature seeds (stages 3 and 4) were flavonols and flavan-3-ols. Flavan-3-ols existed as monomers of epicatechin and catechin and as procyanidins. Type B procyanidins were major components and varied from dimers to pentadecamer. Two anthocyanins, cyanidin-3-*O*-arabinoside and cyanidin-3-*O*-galactoside, along with the *N*-phenylpropernoyl-L-amino acids, *N*-caffeoyl-L-aspartate, *N*-coumaroyl-L-aspartate, *N*-coumaroyl-3-hydroxytyrosine (clovamide), and *N*-coumaroyltyrosine (deoxyclovamide), and the purine alkaloids theobromine and caffeine, were present in stage 3 and 4 seeds. Other purine alkaloids, such as theophylline and additional methylxanthines, did not occur in detectable quantities. Flavan-3-ols were the only components to accumulate in detectable quantities in young seeds at developmental stages 1 and 2.

KEYWORDS: *cacao seed development, flavonols, flavan-3-ols, proanthocyanidins, anthocyanins, N-phenylpropenoyl amino acids, purine alkaloids*

INTRODUCTION

There have been extensive studies implicating the consumption of cacao-based products with reduced risks of cardiovascular disease and cancer that have recently been reviewed in depth.[1,2] Seeds of cacao, *Theobroma cacao* L. (Malvaceae), used in the manufacture of cocoa beverage and chocolate products, contain abundant quantities of secondary metabolites, including procyanidins and theobromine. Three major cacao varieties, Forastero, Criollo, and Trinitario, are grown commercially throughout the world. Trinitario is a hybrid of Criollo and Forastero, and its seeds are used in about 10% of chocolate products. Trinitario beans generally provide better flavor qualities to cocoa-based products than Forastero,[3,4] although because of substantial genetic variations[5] there may be exceptions to this rule. Nonetheless, demand for Trinitario beans is likely to increase as the market for high quality chocolate products expands.

Ripe cacao pods containing about 30−40 seeds, embedded in white pulp comprised mainly of sugars, are the starting material for the manufacture of cocoa products. The pods are harvested and broken open and the pulp and seeds formed into large mounds and covered with leaves before being fermented for 6−8 days. During this period, sucrose is converted to glucose and fructose by invertase, and the glucose is subsequently utilized in fermentation, yielding ethanol which is metabolized to acetic acid. As the tissues of the beans lose cellular integrity, storage proteins are hydrolyzed to peptides and amino acids while polyphenol oxidase converts phenolic components to quinones which polymerize, yielding the brown, highly insoluble compounds that give chocolate its characteristic color.[6] After fermentation, the seeds are dried in the sun, reducing the moisture content from ∼55% to ∼7.5%, before being packed for wholesale trade and subsequent conversion to a variety of chocolate products.[4]

Secondary metabolites in the fermented, dried cocoa beans have frequently been analyzed, and these data have appeared in food-related journals. However, metabolites in such fermented seeds vary greatly because of different postharvest circumstances, and it is also known that subsequent manufacturing steps to produce cocoa products lead to considerable losses of phytochemicals.[7,8]

It was against this background that the current study investigated changes occurring in the secondary metabolite profile during the growth and development of Trinitario cacao seed obtained from a tree in a test field of the Hawaii Agriculture Research Center that had been subject to genomic analysis.[9] Different sized cacao pods selected and seeds at different stages of growth were collected and frozen prior to lyophilization, extraction, and analysis of flavan-3-ols and

Received: October 16, 2012
Revised: December 2, 2012
Accepted: December 6, 2012
Published: December 6, 2012

ACS Publications © 2012 American Chemical Society

proanthocyanidins, flavonols, methylxanthines, anthocyanins, and *N*-phenylpropenoyl amino acids by HPLC with MS/MS and PDA or fluorometric detection.

MATERIALS AND METHODS

Plant Material. Four different sizes of fruit from an adult tree of *Theobroma cacao* L. cv. Trinitario (strain no. K25) planted at a farm of the Hawaii Agriculture Research Center at Kunia, Oahu Island, were used in the study. Seeds were isolated from the cacao pods and immediately frozen with liquid nitrogen prior to lyophilization. The seed samples in developmental stages 1 and 2 included small amounts of pulp tissue which was technically difficult to remove. In the case of stage 3 and 4 seeds, the seed coat was removed (see Figure 1). This

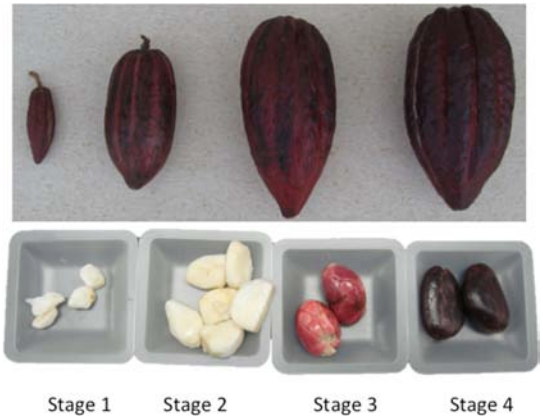

Figure 1. Developing pods of cacao (upper) and seeds (lower), left to right, at stages 1−4 of development.

was necessary in order to avoid degradation, as a ∼50% loss of epicatechin has been reported to occur over a 2 day period during sun-drying of fresh, unfermented cacao seeds.[7]

The length and width of the fruits and fresh weights of four different growth stages selected for the experiments were as follows: 90 × 38 mm, 60 g (stage 1); 125 × 60 mm, 220 g (stage 2); 180 × 80 mm, 560 g (stage 3); 195 × 90 mm, 673 g (stage 4) (Figure 1). The mean fresh weights of seeds in stages 1 to 4 were, respectively, 125, 700, 1550, and 2050 mg The weight of selected seeds varied within 5% of the mean.

Extraction of Cacao Seeds. Lyophilized, powdered seed weighing 20 mg was homogenized in 10 mL of methanol/formic acid (99:1, v/v) with an Ultra-Turrax T-25 (IKAR-Werke, Staufen, Germany) prior to centrifugation at 4000*g* for 15 min. The supernatant was decanted and the pellet re-extracted twice. The pooled supernatants were reduced to dryness under N_2 and resuspended in 0.5 mL of methanol/formic acid (99:1, v/v) prior to analysis by HPLC-MS/MS.

HPLC-PDA-MS^2 Analysis. Samples were analyzed on a Surveyor HPLC system comprising an HPLC pump, an autosampler operating at 4 °C, a PDA detector scanning from 250 to 600 nm (Thermo Electron Corporation, San Jose, CA), and a fluorescence detector (Jasco, Great Dunmow, Essex). Reverse phase separations were carried out using a 250 × 4.6 mm i.d. 4 μm C_{12} Synergi Max-RP column with a guard column of the same support (Phenomenex, Macclesfield, UK) maintained at 40 °C and eluted at a flow rate of 1.0 mL/min with (a) a 60 min linear gradient of 10−40% methanol in 1% aqueous formic acid for anthocyanins, (b) a 75 min gradient, 60 min of 10−40% methanol in 1% aqueous formic acid, followed by 60−75 min at 40%, for flavonols and *N*-phenylpropenoyl amino acids, and (c) a 75 min linear gradient of 10−40% methanol in 1% aqueous formic acid for methylxanthines. After passing through the flow cells of the PDA and fluorescence detectors, the column eluate was split and 0.3 mL/min directed to an LCQ Advantage ion trap mass spectrometer fitted with an electrospray interface operating in positive ionization mode for anthocyanins and methylxanthines, and in negative ionization for flavonols and *N*-phenylpropenoyl amino acids. Identifications were based on full scan, data-dependent MS/MS scanning from *m/z* 100−2000, and comparisons of the retention time and absorption λ_{max} with available standards. Identification of flavonols was also based on their MS/MS/MS fragmentation patterns. With ESI in positive ionization mode, capillary temperature was 300 °C, sheath gas was 50 units, auxiliary gas was 40 units, and source voltage was 3 kV for anthocyanins, while capillary temperature was 275 °C, sheath gas was 30 units, auxiliary gas was 10 units, and source voltage was 2 kV for methylxanthines. For negative ionization, capillary temperature was 250 °C, sheath gas was 60 units, auxiliary gas was 40 units, and source voltage was 4.5 kV.

Anthocyanins were quantitated on the basis of chromatographic peak areas acquired at 520 nm, while the quantification of other compounds were based on the absorbance response at 270 nm for methylxanthines, 320 nm for *N*-phenylpropenoyl amino acids, and 365 nm for flavonols, and the fluorescence response (excitation 230 nm, emission 320 nm) for the flavan-3-ols catechin and epicatechin.

Extraction of Procyanidins. Lyophilized, powdered seeds, 20 mg, were first defatted with hexane prior to extraction with a mixture of acetone:water:acetic acid (70:29.5:0.5, v/v/v). After centrifugation, extracts were sonicated at 50 °C and then passed through a SPE cartridge Strata SCX (Phenomenex) according to the procedures of Robbins and co-workers.[10]

Analysis of Procyanidins by HPLC-MS/MS. Extracts were analyzed with the Surveyor HPLC system described above with separations carried out using a Develosil Diol column 100 Å (Phenomenex, Macclesfield, UK).[10] After passing through the flow cell of the fluorescence detector (excitation 230 nm, emission 320 nm), the column eluate was split and 0.3 mL/min was directed to an LCQ Duo mass spectrometer fitted with an ESI. The capillary temperature was 300 °C, sheath gas and auxiliary gas were 30 and 20 units, respectively, and the source voltage was 7 kV. Samples were analyzed using full scan in negative ionization mode with a scan range from 200 to 2000 *m/z*. Procyanidins were identified on the basis of their MS/MS fragmentation patterns and quantified in (−)-epicatechin equivalents using fluorescence peak areas after which estimates were adjusted to account for the reduced fluorescence response for the procyanidin oligomers and polymers.[10]

RESULTS AND DISCUSSION

Identification of Secondary Metabolites in Developing Cacao Seeds. HPLC-absorbance profiles of extracts of developing cacao seeds indicated the presence of five different groups of compounds in significant amounts: flavan-3-ols and methylxanthines at 270 nm, anthocyanins at 520 nm, *N*-phenylpropenoyl amino acids at 320 nm, and flavonols at 365 nm (Figure 2) while MS/MS analysis facilitated the identification and quantification of 15 compounds as summarized in Table 1.

Peak 1 (t_R 11.0 min, λ_{max} 270 nm) had a positively charged molecular ion ($[M - H]^+$) at *m/z* 181 which on MS/MS produced a major ion at *m/z* 138. On the basis of the retention time, absorbance, and mass spectra data compared with an authentic standard, this compound is identified as theobromine (3,7-dimethylxanthine) (**1** in Figure 3).

Peak 2 (t_R 15.3 min, λ_{max} 280 nm) fluoresced and had a negatively charged molecular ion ($[M - H]^-$) at *m/z* 289 which yielded MS/MS fragments at *m/z* 245 and 205. Mass spectrum and cochromatography characteristics with an authentic standard identified this compound as catechin. It is probably (+)-catechin (**2**), but reverse phase HPLC, unlike chiral chromatography, does not separate the (+) and (−) enantiomers of catechin.[11]

Peak 3 (t_R 16.0 min, λ_{max} 320 nm) produced a $[M - H]^-$ at *m/z* 294 which on MS/MS yielded major fragments at *m/z* 179, 131, and 276. Chromatographic properties and fragmen-

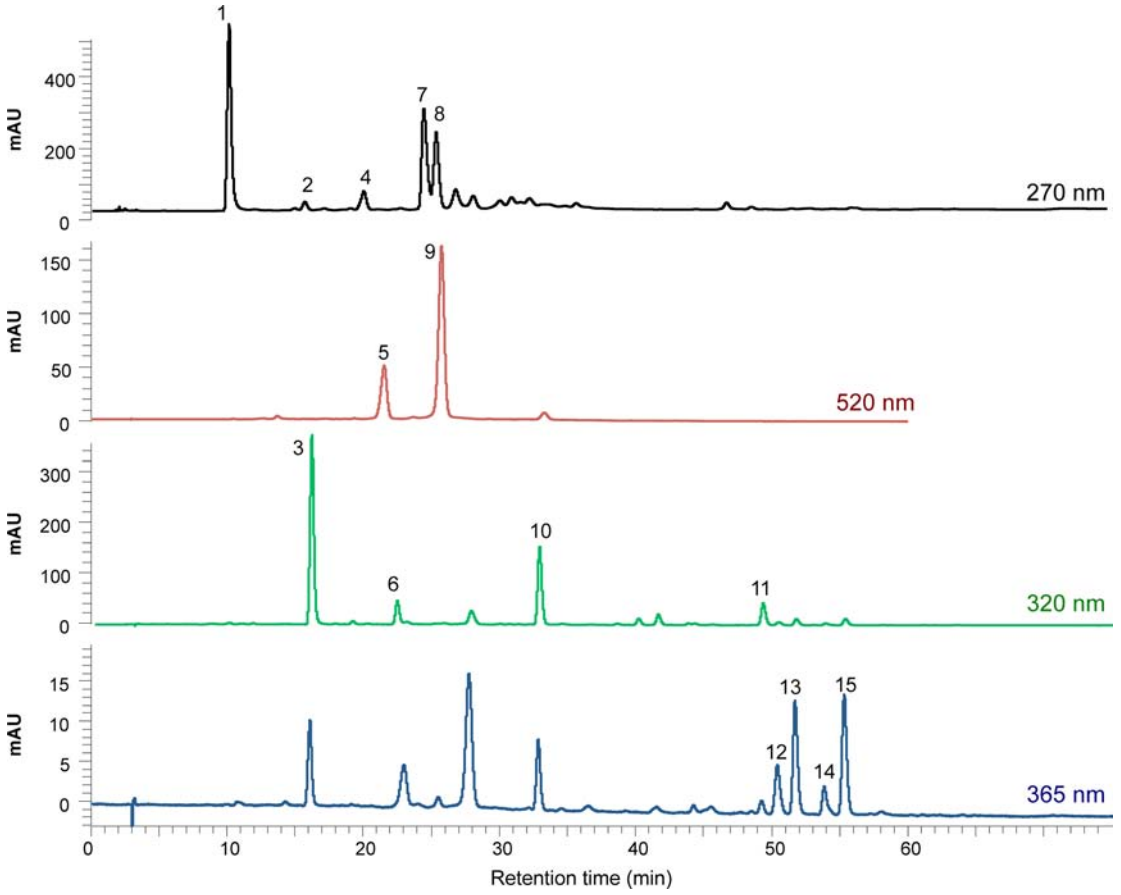

Figure 2. HPLC profiles of a methanolic extract of cocoa seeds. (A) 270 nm: methylxanthines and flavan-3-ols (flavan-3-ols were also monitored with fluorescence detection), (B) 520 nm: anthocyanins, (C) 320 nm: *N*-phenylpropenoyl amino acids, and (D) 365 nm, flavonols. For peak identification, see Table 1.

Table 1. HPLC-MS/MS Identification of Flavonoids and Related Compounds Detected in Cocoa Beans[a]

peak	t_R (min)	compound	$[M - H]^-$ (m/z)	MS/MS (m/z)
1	11.0	theobromine	181^{+} [b]	138
2	15.3	catechin	289	245, 205
3	16.0	*N*-caffeoyl-L-aspartate	294	179, 276, 131
4	19.5	procyanidin B2 dimer	577	425, 407, 289
5	21.5	cyanidin-3-*O*-galactoside	449^{+} [b]	287
6	22.3	*N*-coumaroyl-L-aspartate	278	163, 132, 119
7	24.3	epicatechin	289	245, 205, 179
8	25.3	caffeine	195^{+} [b]	138
9	25.8	cyanidin-3-*O*-arabinoside	419^{+} [b]	287
10	32.8	*N*-coumaroyl-3-hydroxytyrosine (clovamide)	358	222, 178
11	49.3	*N*-coumaroyltyrosine (deoxyclovamide)	326	282, 206, 145
12	50.4	quercetin-3-*O*-galactoside	463	301
13	51.6	quercetin-3-*O*-glucoside	463	301
14	53.7	kaempferol-3-*O*-rutinoside	593	447, 285
15	55.2	quercetin-*O*-pentoside	433	301

[a]For HPLC profiles and peaks, see Figure 2. [b]+ = positively charged molecular ion.

tation pattern of *N*-phenylpropenoyl amino acids indicate that this peak is *N*-caffeoyl-L-aspartate (**3**).[12]

Peak 4 (t_R 19.5 min, λ_{max} 280 nm) fluoresced and had a $[M - H]^-$ at m/z 577 which yielded MS/MS ions at m/z 425, 407, and 289. This spectrum and chromatographic elution time tentatively identifies peak 4 as B-type flavan-3-ol dimer, probably procyanidin B2 (**4**), a known component of cacao.[13,14]

Peak 5 (t_R 21.5 min, λ_{max} 520 nm) had a $[M - H]^+$ at m/z 449 which fragmented with a loss of 162 Da (hexose group) to produce an m/z 287 (cyanidin) daughter ion. This mass spectrum and the elution of this peak prior to that of a standard of cyanidin-3-*O*-glucoside tentatively identifies peak 5 as the known cacao anthocyanin cyanidin-3-*O*-galactoside (**5**).[15]

Peak 6 (t_R 22.3 min, λ_{max} 320 nm) produced a $[M - H]^-$ at m/z 278 which on MS/MS yielded major fragments at m/z 163, 132, and 119. In keeping with the findings of Stark and Hofmann,[12] this peak was identified as *N*-coumaroyl-L-aspartate (**6**).

Peak 7 (t_R 24.3 min, λ_{max} 280 nm) fluoresced and yielded a $[M - H]^-$ at m/z 289 which gave rise to MS/MS ions at m/z 245, 205, and 179. The cochromatography with an authentic standard supports the identification of peak 7 as epicatechin, most probably (−)-epicatechin (**7**), although the presence of its epimer (+)-epicatechin, with which it cochromatographs,[11] cannot be completely ruled out.

Peak 8 (t_R 25.3 min, λ_{max} 270 nm) had a $[M - H]^+$ at m/z 195 which on MS/MS produced a major ion at m/z 138. On

dx.doi.org/10.1021/jf304397m | J. Agric. Food Chem. 2013, 61, 427−434

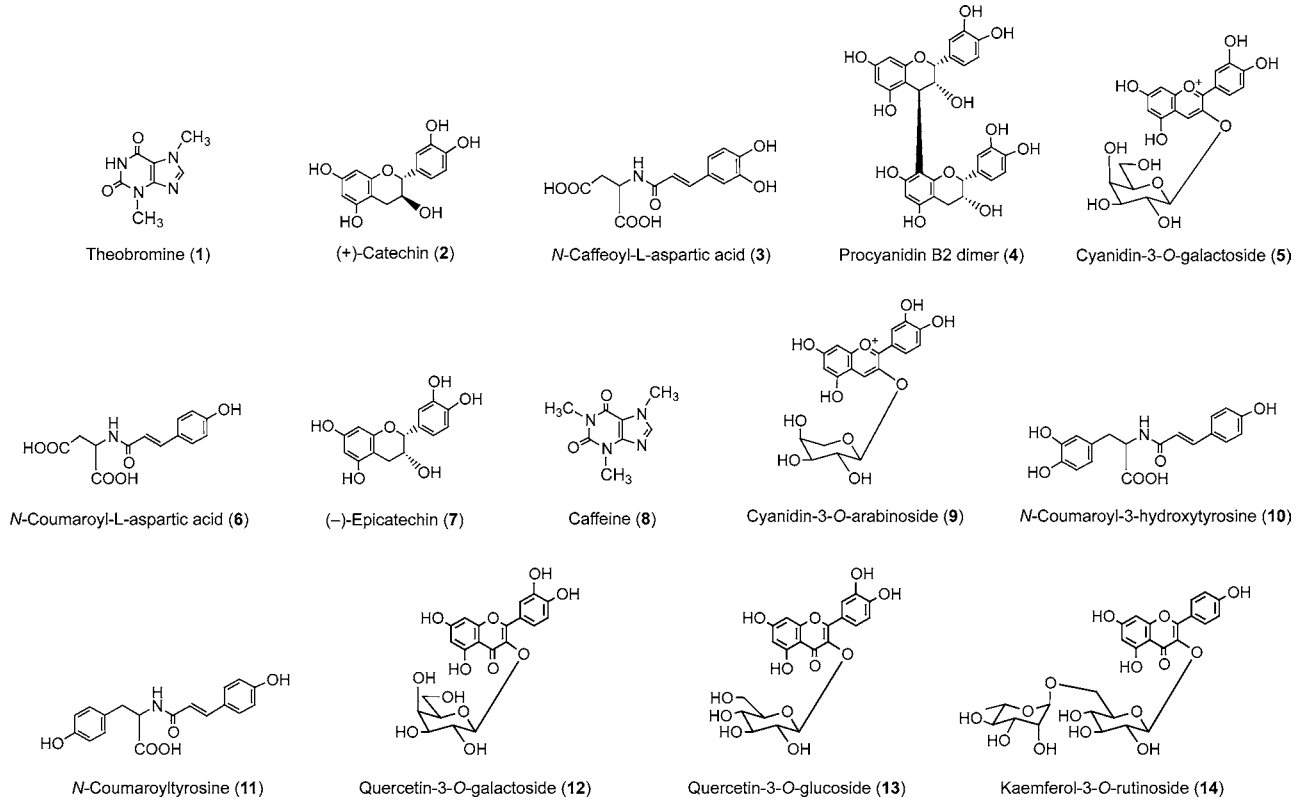

Figure 3. Structures of secondary metabolites detected in developing cacao seeds.

the basis of the retention time, λ_{max} and mass spectra data compared with an authentic standard, this compound is caffeine (1,3,7-trimethylxanthine) (**8**).

Peak 9 (t_R 25.8 min, λ_{max} 520 nm) had a $[M - H]^+$ at m/z 419 which upon MS/MS fragmented with a loss of 132 *amu* (pentose group) to produce a m/z 287 (cyanidin) daughter ion. Based on the fragmentation pattern this compound was tentatively identified as cyanidin-3-*O*-arabinoside (**9**) in keeping with the data of Forsyth and Quesnel [15]

Peak 10 (t_R 32.8 min, λ_{max} 320 nm) produced a $[M - H]^-$ at m/z 358 which yielded major MS/MS fragment at m/z 222 and a minor ion at m/z 178. On the basis of the mass spectrum and the HPLC elution profile, peak 10 was tentatively identified as *N*-coumaroyl-3-hydroxytyrosine (clovamine) (**10**), a known constituent of cacao beans.[12]

Peak 11 (t_R 49.3 min, λ_{max} 320 nm) had a $[M - H]^-$ at m/z 326, which upon MS/MS yielded a major ion at m/z 282 and minor fragments at m/z 206 and 145. On the basis of the mass spectrum and the HPLC elution profile, peak 11 was tentatively identified as *N*-coumaroyltyrosine (deoxyclovamine) (**11**) which has previously been detected in cacao.[12]

Peak 12 (t_R 50.4 min, λ_{max} 365 nm) and peak 13 (t_R 51.6 min, λ_{max} 365 nm) both yielded a $[M - H]^-$ at m/z 463 and a MS/MS spectrum with ions at m/z 301 ($[M - H-162]^-$), indicating the presence of an *O*-linked hexose moiety. Based on the mass spectrum and their elution sequence peak, peak 12 is tentatively identified as quercetin-*O*-galactoside (**12**), and peak 13 as quercetin-3-*O*-glucoside (**13**), on the basis of cochromatography with authentic standards, in keeping with the findings of Sanchez-Rabaneda et al.[16]

Peak 14 (t_R 53.7 min, λ_{max} 365 nm) had a $[M - H]^-$ at m/z 593, and MS/MS produced a minor fragment at m/z 447, a 146 amu cleavage of a rhamnose moitey, and a major ion at m/z 285, which is indicative of the flavonol aglycone kaempferol. On the basis of the mass spectral fragmentation pattern and cochromatography with an authentic standard, peak 14 was established as kaempferol-3-*O*-rutinoside (**14**).

Peak 15 (t_R 55.2 min, λ_{max} 365 nm) yielded a $[M - H]^-$ ion at m/z 433 and an MS/MS spectrum with an ion at m/z 301 ($[M - H - 132]^-$), indicating the presence of an O-linked pentose moiety. On the basis of the mass spectrum and λ_{max}, peak 15 is tentatively identified as a quercetin-*O*-pentoside.

Analysis of Procyanidins. Flavan-3-ols in cocoa comprise the monomeric (+)-catechin and (−)-epicatechin and their oligomeric and polymeric procyanidins. Because of the poor resolution for the higher oligomeric constituents on reversed phase HPLC columns, the cacao procyanidins with a degree of polymerization up to pentadecamers were analyzed by HPLC using a diol column (Figure 4) and quantified by fluorometry in (−)-epicatechin equivalents adjusted for the reduced fluorescence response of the procyanidins. Mass spectrometric data were used to identify some of the procyanidins. The ions observed for each procyanidin oligomer are shown in Table 2. Peaks with a $[M - H]^-$ at m/z 575, 863, 1151, 1439, and 1727 were identified as type-A dimeric, trimeric, tetrameric, pentameric, and hexameric procyanidins of (epi)catechin, while peaks with a $[M - H]^-$ at m/z at 577, 865, 1153, 1441, and 1729 corresponded to type-B dimeric, trimeric, tetrameric, pentameric, and hexameric procyanidins of (epi)-catechin. The fluorescent peaks 27 to 35 were outside the upper mass range of the mass spectrometer and did not yield MS data but were tentatively identified as heptameric to pentadecameric procyanidins of (epi)catechin, as previously described by Robbins and co-workers.[10]

Changes in Secondary Metabolite Content during Seed Development. Secondary metabolites quantified in

dx.doi.org/10.1021/jf304397m | J. Agric. Food Chem. 2013, 61, 427−434

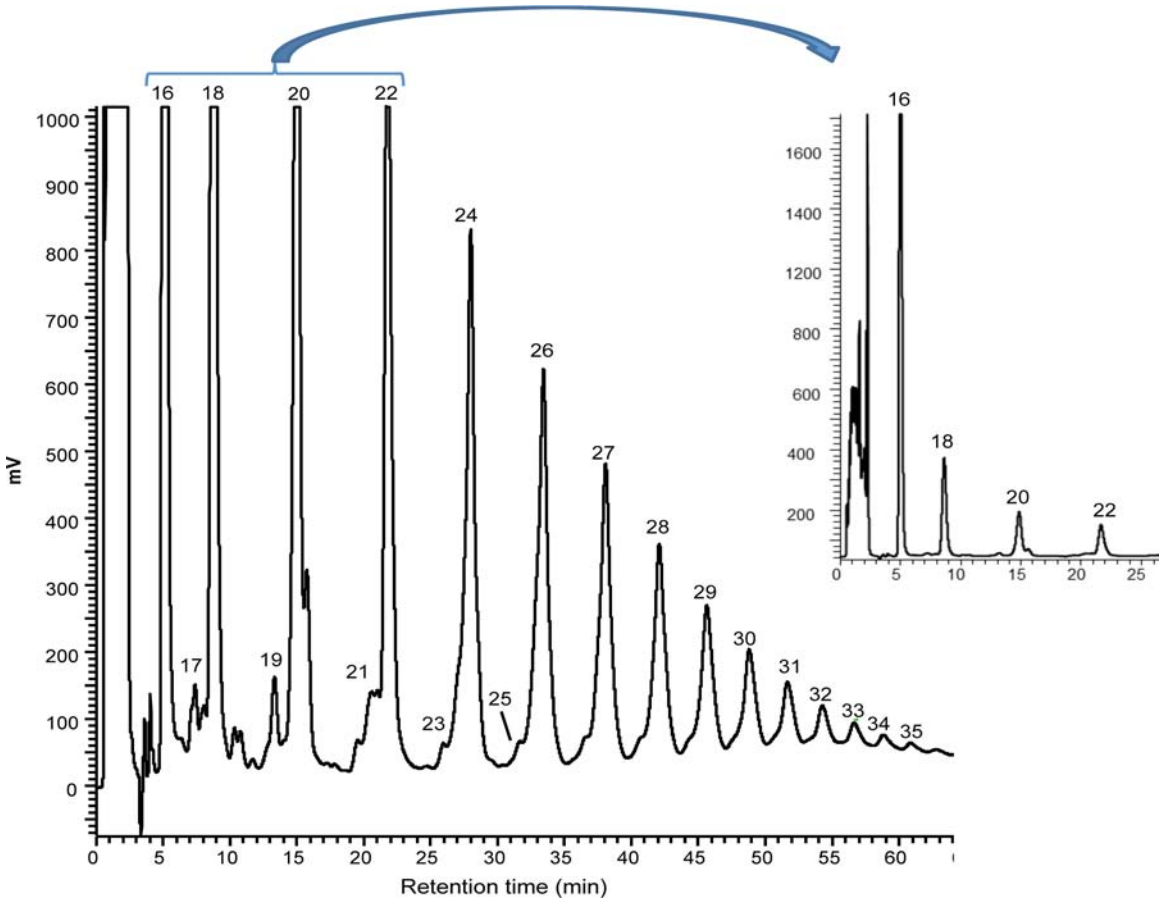

Figure 4. A diol HPLC profile with fluorescence detection of procyanidins in a stage 4 cacao seed extract. For MS data and peak identification see Table 2.

Table 2. Diol HPLC Identification of Flavan-3-ols and Procyanidins in Cacao Beans[a]

peak	t_R (min)	oligomer	$[M - H]^-$ (m/z)	MS/MS (m/z)
16	5.25	catechin[b]	289	245, 205, 179
16	5.25	epicatechin[b]	289	245, 205, 179
17	7.41	dimer, type A	575	423, 449, 289
18	8.79	dimer, type B	577	407, 451, 289
19	13.35	trimer, type A	863	575, 449, 289
20	14.76	trimer, type B	865	577, 407, 451
21	21.3	tetramer, type A	1151	863, 575, 423
22	21.48	tetramer, type B	1153	865, 577, 423
23	27.23	pentamer, type A	1439	863, 575, 449
24	27.75	pentamer, type B	1441	865, 577, 407
25	33.04	hexamer, type A	1727	1439, 863, 575
26	33.23	hexamer, type B	1729	1439, 1153, 863
27	37.87	heptamer	–	–
28	41.93	octamer	–	–
29	45.47	nonamer	–	–
30	48.67	decamer	–	–
31	51.48	undecamer	–	–
32	54.06	dodecamer	–	–
33	56.47	tridecamer	–	–
34	58.65	tetradecamer	–	–
35	60.61	pentadecamer	–	–

[a]For HPLC peaks, see Figure 4. [b]Analysis of catechin and epicatechin is based on C_{12} reverse phase HPLC which unlike diol HPLC is able to separate the two monomers.

developing cacao seeds were flavonols, flavan-3-ols, *N*-phenylpropenoyl-L-amino acids, anthocyanins, and the purine alkaloids theobromine and caffeine. The concentration of these compounds on a mg/g dry weight basis is presented in Table 3 while an overview of the amounts present on a per seed basis is illustrated in Figure 5.

Flavonols. Several quercetin *O*-glycosides, including quercetin-3-*O*-galactoside, quercetin 3-*O*-arabinoside, and quercetin-3-*O*-glucoside have previously been identified in cocoa powders.[17] In the present study with developing cacao beans, a quercetin-*O*-pentoside and kaempferol-3-*O*-rutinoside were also detected. Extremely high flavonol concentrations, 25 mg and 47 mg/g dry weight, were found in stage 3 and 4 seeds. The major components were quercetin-3-*O*-glucoside and quercetin-*O*-pentoside which represented 70% of the total flavonols in stage 4 seeds (Table 3). Earlier studies with processed cocoa powders detected much smaller quantities of flavonols (<0.1 mg/g dry weight) which have invariably been trace constituents compared to flavan-3-ol monomers and procyanidins.[17] This substantial difference in flavonol contents between the intact seeds and cacao powder may be due to the disappearance of flavonols during fermentation and/or commercial processing of the beans. A decrease in flavonols of at least 86% has been reported to occur during the alkalization process used in the manufacture of commercial cocoa powder.[18]

Flavan-3-ols. The cacao seeds contained two flavan-3-ol monomers, epicatechin and catechin. Epicatechin is always the predominant monomer (Table 3). In contrast to flavan-3-ol

dx.doi.org/10.1021/jf304397m | J. Agric. Food Chem. 2013, 61, 427−434

Table 3. Concentration of Secondary Metabolites in Developing Cacao Seeds[a]

compounds	stage 1	stage 2	stage 3	stage 4
N-Phenylpropenoyl Amino Acids				
N-caffeoyl-L-aspartate	n.d.	n.d.	0.07 ± 0.01	0.30 ± 0.02
N-coumaroyl-L-aspartate	n.d.	n.d.	0.05 ± 0.01	0.08 ± 0.00
clovamide	n.d.	n.d.	0.04 ± 0.01	0.16 ± 0.01
deoxyclovamide	n.d.	n.d.	0.02 ± 0.00	0.09 ± 0.00
total	n.d.	n.d.	0.18 ± 0.03	0.63 ± 0.02
Flavonols				
quercetin-3-*O*-galactoside	n.d.	n.d.	3.9 ± 0.3	8.2 ± 0.6
quercetin-3-*O*-glucoside	n.d.	n.d.	7.7 ± 1.2	16.3 ± 1.2
quercetin-*O*-pentoside	n.d.	n.d.	10.1 ± 1.3	16.4 ± 1.0
kaempferol-3-*O*-rutinoside	n.d.	n.d.	3.7 ± 0.6	6.0 ± 0.6
total	n.d.	n.d.	25.4 ± 3.4	46.9 ± 3.4
Flavan-3-ols and Procyanidins				
catechin[b]	0.05 ± 0.01	0.05 ± 0.01	0.11 ± 0.01	0.47 ± 0.01
epicatechin[b]	0.70 ± 0.10	0.84 ± 0.01	4.3 ± 0.0	3.9 ± 0.0
dimer, type A	n.d.	n.d.	0.11 ± 0.01	0.15 ± 0.03
dimer, type B	2.5 ± 0.3	3.5 ± 0.4	4.8 ± 0.1	5.4 ± 0.2
trimer, type A	n.d.	n.d.	0.15 ± 0.01	0.22 ± 0.05
trimer, type B	2.0 ± 0.1	2.6 ± 0.3	3.00 ± 0.02	3.0 ± 0.3
tetramer, type A	n.d.	n.d.	0.43 ± 0.03	0.52 ± 0.03
tetramer, type B	2.31 ± 0.01	3.0 ± 0.3	2.5 ± 0.1	3.0 ± 0.6
pentamer, type A	0.15 ± 0.03	n.d.	0.3 ± 0.1	0.37 ± 0.01
pentamer, type B	2.23 ± 0.04	3.1 ± 0.3	2.10 ± 0.04	2.60 ± 0.02
hexamer, type A	0.36 ± 0.02	0.34 ± 0.05	0.18 ± 0.02	0.15 ± 0.01
hexamer, type B	2.4 ± 0.1	3.7 ± 0.1	2.50 ± 0.04	2.69 ± 0.03
heptamer	1.85 ± 0.02	2.5 ± 0.2	1.5 ± 0.1	1.49 ± 0.06
octamer	1.20 ± 0.01	2.5 ± 0.3	1.3 ± 0.1	1.40 ± 0.08
nonamer	2.35 ± 0.05	3.0 ± 0.3	1.4 ± 0.1	1.60 ± 0.03
decamer	1.2 ± 0.1	1.40 ± 0.03	0.61 ± 0.07	0.8 ± 0.1
undecamer	0.97 ± 0.07	1.20 ± 0.05	0.5 ± 0.1	0.70 ± 0.01
dodecamer	0.68 ± 0.03	1.00 ± 0.06	0.30 ± 0.05	0.42 ± 0.04
tridecamer	0.55 ± 0.01	0.71 ± 0.03	0.11 ± 0.01	0.30 ± 0.08
tetradecamer	0.36 ± 0.07	0.50 ± 0.06	0.14 ± 0.04	0.20 ± 0.01
pentadecamer	0.33 ± 0.01	0.40 ± 0.04	0.08 ± 0.01	0.16 ± 0.06
	22.19 ± 1.08	30.34 ± 2.54	26.42 ± 1.06	29.54 ± 1.76
Anthocyanins				
cyanidin-3-*O*-galactoside	n.d.	n.d.	0.10 ± 0.01	0.16 ± 0.02
cyanidin-3-*O*-arabinoside	n.d.	n.d.	0.42 ± 0.01	0.44 ± 0.05
total	n.d.	n.d.	0.52 ± 0.01	0.60 ± 0.07
Purine Alkaloids				
theobromine	n.d.	n.d.	0.89 ± 0.02	2.3 ± 0.2
caffeine	n.d.	n.d.	0.69 ± 0.01	1.4 ± 0.1
total	n.d.	n.d.	1.58 ± 0.02	3.7 ± 0.3

[a]Data expressed as mean values in mg/g dry weight ± standard deviation (n = 3). [b]Catechin and epicatechin values determined by reverse phase HPLC, procyanidins by diol HPLC, n.d. = not detected.

monomers, no large differences in the concentration of procyanidin oligomers and polymers were observed in stages 1 to 4 (Table 3). Concentration of some flavan-3-ol polymers, such as the undecamer and dodecamer, was higher in young (stages 1 and 2) seeds than in developed (stages 3 and 4) seeds. The overall flavan-3-ol content on a per seed basis increased broadly in line with increased seed weight. In contrast to this increase, procyanidin levels have been shown to decrease 3- to 5-fold during fermentation.[19] Likewise, major losses of both (−)-epicatechin and (+)-catechin occur during fermentation alongside the formation of (−)-catechin, events attributable to heat, and which continue during natural and Dutch-processing of cacao.[20]

N-Phenylpropernoyl-L-amino Acids. These hydroxycinnamate derivatives are unique to cacao. The occurrence of *N*-coumaroyl-3-hydroxytyrosine (clovamide) and *N*-coumaroyl-tyrosine (deoxyclovamide) in cocoa liquor was first described by Sanbongi et al.,[21] and they have also been detected in unfermented cacao seeds and somatic embryos.[22] In the present study, two aspartate derivatives, *N*-caffeoyl-L-aspartate and *N*-coumaroyl-L-aspartate, were found in the stage 3 and 4 seeds. The total *N*-phenylpropernoyl-L-amino acid content in the stage 4 seed was ∼3.5 times higher than the concentration at stage 3. Therefore, accumulation of the hydroxycinnamate derivatives occurred in the later two stages of seed growth. The order of concentration was *N*-caffeoyl-L-aspartate > clovamide > deoxyclovamide > *N*-coumaroyl-L-aspartate (Table 3).

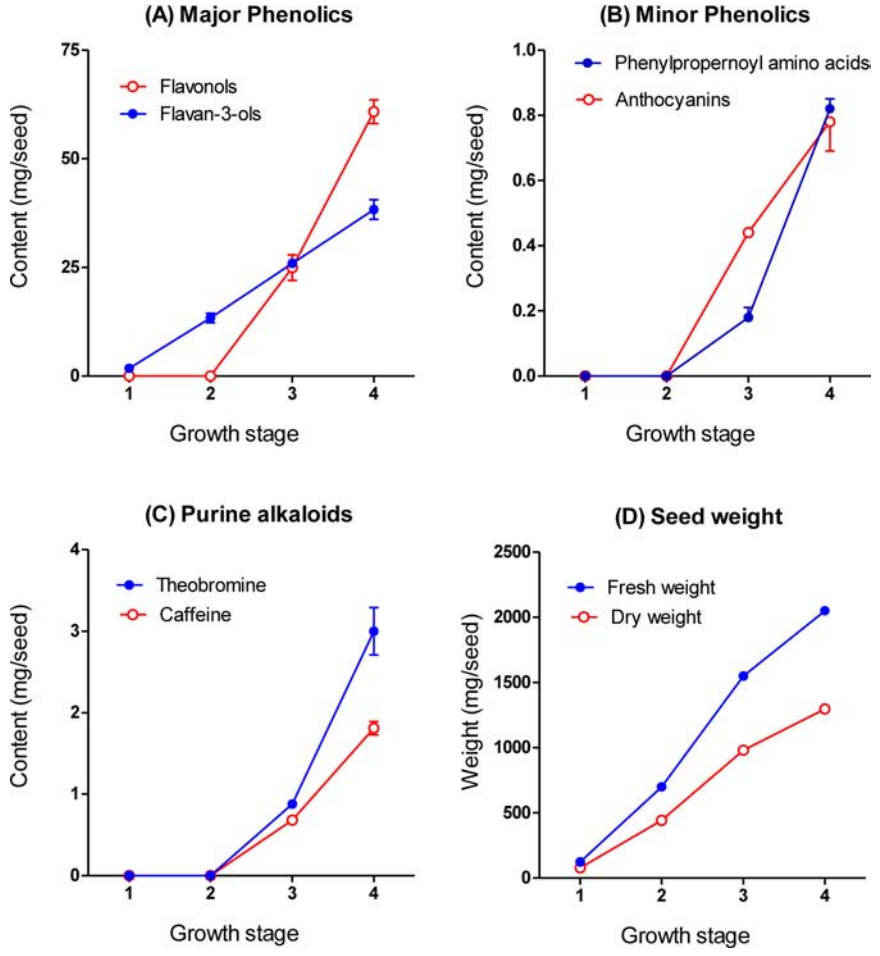

Figure 5. Changes in (A) major phenolics (flavonols and flavan-3-ols), (B) minor phenolics (*N*-phenylpropenoyl amino acids and anthocyanins), (C) purine alkaloids (theobromine and caffeine), and (D) fresh and dry weight of the developing cacao seeds. For A, B, and C, data are expressed as mg/seed with error bars indicating standard deviation (n = 3). Seed weight = mean values mg/seed.

Recently, *N*-phenylpropernoyl-L-amino acids were detected in all parts of the cacao fruit, with the husk and pulp containing clovamide as the dominant constituent.[23] In flowers and leaves, *N*-phenylpropernoyl-L-amino acids are absent although 2-*O*-caffeoyltartaric acid is the major caffeic acid metabolite in leaves. Elwers et al. reported that the *N*-caffeoyl-L-aspartate content of Criollo was higher than that of Forastero and Trinitario beans, and that the content was increased by fertilization of the soil.[24]

Anthocyanins. Two anthocyanins, cyanidin-3-*O*-arabinoside and cyanidin-3-*O*-glactoside, were detected in developed cacao seeds with the arabinoside being the major component at 78% and 73% of total anthocyanin content in stage 3 and stage 4 seeds, respectively. This contrasts with the findings of Elwers et al. who reported that Criollo cacao seeds do not contain anthocyanins.[24] However, many Criollo genotypes exist, and some would appear to contain anthocyanins, as the seeds have a pink coloration.

Purine Alkaloids. Two methylxanthines, theobromine and caffeine, were found in seeds of stages 3 and 4 but not in younger seeds. These results are consistent with the evidence that the biosynthetic activity of purine alkaloids is only found in developing cacao seeds from large cacao pods.[25] In contrast to tea leaves and coffee seeds,[25] the cacao seeds contained more theobromine than caffeine with the latter comprising 56−62% of the total purine alkaloids. The predominance of theobromine in cacao is probably a consequence of the *N*-methyltransferase-catalyzed metabolism of theobromine to caffeine being a rate-limiting conversion in the four-step caffeine biosynthesis pathway.[26] Examination of several cacao genotypes representing the three horticultural races, Criollo, Forastero, and Trinitario, revealed considerable variations in the purine alkaloid content of the seed with slightly higher levels found in Criollo seeds.[26,27]

In summary, this study investigated changes in secondary metabolite pools occurring during the maturation of Trinitario cacao seeds rather than that of some of the derived products that appear during fermentation and commercial processing in the manufacture of cocoa and chocolate. The data obtained indicate that the major phytochemicals in the developing seed are flavonols and flavan-3-ols accompanied by lower amounts of *N*-phenylpropernoyl-L-amino acids, anthocyanins, and the purine alkaloids, theobromine and caffeine. With the exception of flavan-3-ols, these compounds were not found in young stage 1 and 2 seed but accumulated in the later stages of development. In broad terms, the secondary metabolites increased in proportion to the increase in the dry weight of the developing seed. The profile of the metabolites in the intact cacao seeds observed in this study, most notably the high flavonol content, was markedly different to that of cocoa powders.[18] As noted in the text, there has been extensive work on the declines in cacao secondary metabolite content, most

dx.doi.org/10.1021/jf304397m | *J. Agric. Food Chem.* 2013, 61, 427−434

notably flavan-3-ols, occurring during fermentation, drying, roasting, and Dutching, as attempts are made to balance the health benefits of flavan-3-ol monomers with the beneficial effects of fermentation and roasting on chocolate flavor.[20] One key factor which complicates such studies is the large cultivar to cultivar variations that occur as a consequence of the genetic diversity of seemingly cacao similar trees. The data presented in this study were obtained using seeds from a single tree of known genetic background and reduced the variability that would have been evident if mixed seeds from different trees of undetermined genotype had been utilized. Such an approach would be of value when screening cacao seeds during development in attempts to identify high flavan-3-ol producers to propagate the clonal material.

AUTHOR INFORMATION

Corresponding Author

*Phone: +81-5700-4225. Fax: +81-5700-4225. E-mail: ashihara.hiroshi@ocha.ac.jp.

Funding

G.P.-C. was supported by a postdoctoral fellowship from IFAPA (Programa Operativo del Fondo Social Europeo 2007−2013 de Andalucia).

Notes

The authors declare no competing financial interest.

REFERENCES

(1) Donovan, J. L.; Holes-Lewis, K. A.; Chavin, K. D.; Egan, B. M. Cocoa and Health. In *Teas, Cocoa and Coffee, Plant Secondary Metabolites and Health*; Crozier, A.; Ashihara, H.; Tomas-Barberan, F., Eds.; Wiley-Blackwell: Oxford, 2012; pp 219−246.

(2) Del Rio, D.; Mateos, A. M.; Spencer, J. P. E.; Tognolini, M.; Borges, G.; Crozier, A. Dietary (poly)phenolics in human health and disease: structures, bioavailability, evidence of protective effects and potential mechanisms. *Antioxid. Redox Signal.* **2012**, in press (DOI: 10 1089/ars.2012.4581)

(3) Leung, A. T.; Foster, S. *Encyclopedia of Common Natural Ingredients: Used in Food, Drugs, and Cosmetics*, 2nd ed.; Wiley-Interscience: New York, 1996.

(4) Tomás-Barberán, F.; Borges, G.; Crozier, A. Phytochemicals in cocoa and flavan-3-ol bioavailability. In *Teas, Cocoa and Coffee: Plant Secondary Metabolites and Health*; Crozier, A.; Ashihara, H.; Tomas-Barberan, F., Eds.; Wiley-Blackwell: Oxford, 2012; pp 193−217.

(5) Motamayor, J. C.; Lachenaud, P.; da Silva e Mota, J. W.; Loor, R.; Kuhn, D. N.; Brown, J. S.; Schnell, R. J. Geographic and genetic population differentiation of the Amazonian chocolate tree (*Theobroma cacao* L). *PLoS One* **2008**, *3*, e3311.

(6) Haslam, E. *Practical Polyphenolics: From Structure to Molecular Recognition and Physiological Action*; Cambridge University Press: Cambridge, 1998.

(7) Wollgast, J.; Anklam, E. Review on polyphenols in *Theobroma cacao*: changes in composition during the manufacture of chocolate and methodology for identification and quantification. *Food Res. Int.* **2000**, *33*, 423−447.

(8) Gu, L.; House, S. E.; Wu, X.; Ou, B.; Prior, R. L. Procyanidin and catechin contents and antioxidant capacity of cocoa and chocolate products. *J. Agric. Food Chem.* **2006**, *54*, 4057−4061.

(9) Schnell, R. J.; Olano, C. T.; Brown, J. S.; Meerow, A. W.; Cervantes-Martinez, C.; Nagai, C.; Motamayor, J. C. Retrospective determination of the parental population of superior cacao (*Theobroma cacao* L.) seedlings and association of microsatellite alleles with productivity. *J. Am. Soc. Hortic. Sci.* **2005**, *130*, 181−190.

(10) Robbins, R. J.; Leonczak, J.; Johnson, J. C.; Li, J. L.; Kwik-Uribe, C.; Prior, R. L.; Gu, L. W. Method performance and multi-laboratory assessment of a normal phase high pressure liquid chromatography-fluorescence detection method for the quantitation of flavanols and procyanidins in cocoa and chocolate containing samples. *J. Chromatogr., A* **2009**, *1216*, 4831−4840.

(11) Clifford, M. N.; Crozier, A. Phytochemicals in teas and tisanes and their bioavailability. In *Teas, Cocoa and Coffee: Plant Secondary Metabolites and Health*; Crozier, A.; Ashihara, H.; Tomas-Barberan, F., Eds.; Wiley-Blackwell: Oxford, 2012; pp 45−98.

(12) Stark, T.; Hofmann, T. Isolation, structure determination, synthesis, and sensory activity of *N*-phenylpropenoyl-L-amino acids from cocoa (*Theobroma cacao*). *J. Agric. Food Chem.* **2005**, *53*, 5419−5428.

(13) Natsume, M.; Osakabe, N.; Yasuda, A.; Baba, S.; Tokunaga, T.; Kondo, K.; Osawa, T.; Terao, J. In vitro antioxidative activity of (−)-epicatechin glucuronide metabolites present in human and rat plasma. *Free Radic. Res* **2004**, *38*, 1341−1348.

(14) Tomás-Barberán, F. A.; Cienfuegos-Jovellanos, E.; Marín, A.; Muguerza, B.; Gil-Izquierdo, A.; Cerdá, B.; Zafrilla, P.; Morillas, J.; Mulero, J.; Ibarra, A.; Pasamar, M. A.; Ramón, D.; Espín, J. C. A new process to develop a cocoa powder with higher flavonoid monomer content and enhanced bioavailability in healthy humans. *J. Agric. Food Chem.* **2007**, *55*, 3926−3935.

(15) Forsyth, W.G. C.; Quesnel, V. C. *Cacao* polyphenol pigments 4. The anthocyanin pigments. *Biochem. J.* **1957**, *65*, 177−179.

(16) Sanchez-Rabaneda, F.; Jauregui, O.; Casals, I.; Andres-Lacueva, C.; Izquierdo-Pulido, M.; Lamuela-Raventos, R. M. Liquid chromatographic/electrospray ionization tandem mass spectrometric study of the phenolic composition of cocoa (*Theobroma cacao*). *J. Mass Spectrom.* **2003**, *38*, 35−42.

(17) Hammerstone, J. F.; Lazarus, S. A.; Mitchell, A. E.; Rucker, R.; Schmitz, H. H. Identification of procyanidins in cocoa (*Theobroma cacao*) and chocolate using high-performance liquid chromatography/mass spectrometry. *J. Agric. Food Chem.* **1999**, *47*, 490−496.

(18) Andres-Lacueva, C.; Monagas, M.; Khan, N.; Izquterdo-Pulido, M.; Urpi-Sarda, M.; Permanyer, J.; Lamuela-Raventós, R. M. Flavanol and flavonol contents of cocoa powder products: Influence of the manufacturing process. *J. Agric. Food Chem.* **2008**, *56*, 3111−3117.

(19) Kealey, K. S.; Snyder, R. M.; Romanczyk, L. J.; Geyer, H. M.; Myers, M. E.; Withcare, E. J.; Hammerstone, J. F.; Schmitz, H. H. Cocoa components, edible products having enhanced polyphenol content, methods of making same and medical uses. Patent Cooperation Treaty (PCT) WO 98/09533, Mars Incorporated, 1998.

(20) Hurst, W. J.; Krake, S. H.; Bergmeier, S. C.; Payne, M. J.; Miller, K. B; Stuart, D. A. Impact of fermentation, drying, roasting and Dutch processing on flavan-3-ol stereochemistry in cacao beans and cocoa ingredients. *Chem. Central J.* **2011**, *5*, 53.

(21) Sanbongi, C.; Osakabe, N.; Natsume, M.; Takizawa, T.; Gomi, S.; Osawa, T. Antioxidative polyphenols isolated from *Theobroma cacao*. *J. Agric. Food Chem.* **1998**, *46*, 454−457.

(22) Alemanno, L.; Ramos, T.; Gargadenec, A.; Andary, C.; Ferriere, N. Localization and identification of phenolic compounds in *Theobroma cacao* L. Somatic embryogenesis. *Ann. Bot.* **2003**, *92*, 613−623.

(23) Lechtenberg, M.; Henschel, K.; Liefländer-Wulf, U.; Quandt, B.; Hensel, A. Fast determination of *N*-phenylpropenoyl-L-amino acids (NPA) in cocoa samples from different origins by ultra-performance liquid chromatography and capillary electrophoresis. *Food Chem.* **2012**, *135*, 1676−1684.

(24) Elwers, S.; Zambrano, A.; Rohsius, C.; Lieberei, R. Differences between the content of phenolic compounds in Criollo, Forastero and Trinitario cocoa seed (*Theobroma cacao* L.). *Eur. Food Res. Technol.* **2009**, *229*, 937−948.

(25) Zheng, X.-Q.; Koyama, Y.; Nagai, C.; Ashihara, H. Biosynthesis, accumulation and degradation of theobromine in developing *Theobroma cacao* fruits. *J. Plant Physiol.* **2004**, *161*, 363−369.

(26) Ashihara, H.; Sano, H.; Crozier, A. Caffeine and related purine alkaloids: Biosynthesis, catabolism, function and genetic engineering. *Phytochemistry* **2008**, *69*, 841−856.

(27) Hammerstone, J. F.; Romanczyk, L. J.; Aitkent, W. M. Purine alkaloid distribution within *Herrania* and *Theobroma*. *Phytochemistry* **1994**, *35*, 1237−1240.