Ruby chocolate : a study of its phytochemical composition and quantitative comparison with dark, milk and white chocolate

Ruby chocolate: a study of its phytochemical composition and quantitative comparison with dark, milk and white chocolate

Emmy TUENTER[a,*], Maria E. SAKAVITSI[b], Andrés RIVERA-MONDRAGÓN[a,c,d], Nina HERMANS[a], Kenn FOUBERT[a], Maria HALABALAKI[b], Luc PIETERS[a]

[a]*Natural Products & Food Research and Analysis (NatuRA), Department of Pharmaceutical Sciences, University of Antwerp, Antwerp, Belgium*

[b]*Department of Pharmacognosy and Natural Products Chemistry, Faculty of Pharmacy, National and Kapodistrian University of Athens, Athens, Greece*

[c]*Centre of Innovation and Technology Transfer, Institute of Scientific Research and High Technology Services (INDICASAT-AIP), Building 208, City of Knowledge, Panama, Republic of Panama.*

[d]*Department of Medicinal Chemistry and Pharmacognosy, Faculty of Pharmacy, University of Panama, Panama City, Republic of Panama*

E-mail addresses: emmy.tuenter@uantwerpen.be (E. Tuenter), mariaelenisakavitsi@gmail.com (M.E. Sakavitsi), arivera.qclabherbals@gmail.com (A. Rivera-Mondragón), nina.hermans@uantwerpen.be (N. Hermans), kenn.foubert@uantwerpen.be (K. Foubert), mariahal@pharm.uoa.gr (Maria Halabalaki), luc.pieters@uantwerpen.be (L. Pieters)

*Corresponding author, phone: +3232652731

Abstract

Ruby chocolate was introduced in 2017 as the fourth type of chocolate, in addition to white, milk and dark chocolate. However, until now not much is reported about its phytochemical composition. Therefore, we analyzed ruby chocolate by UPLC-HRMS, together with the three other types of chocolate. Feature-based molecular networking was carried out to aid in the identification, while a set of 51 reference compounds were analyzed simultaneously for targeted quantification. In this way, a total of 54 compounds could be (tentatively) identified in the chocolates, of which 43 were found in the ruby chocolate. Moreover, 19 compounds were quantified, of which 13 in the ruby chocolate. The compounds include flavan-3-ols, proanthocyanidins and methylxanthines, but also biogenic amines and alkaloids. In general, ruby chocolate contained lower levels of these constituents compared to dark chocolate. However, A-type proanthocyanidins were found to be characteristic for the ruby chocolate.

Keywords

Ruby chocolate; UPLC-HRMS; phytochemical analysis; feature-based molecular networking

1. Introduction

After dark, milk and white chocolate, ruby chocolate was introduced in 2017 as a fourth type of chocolate by the Belgian chocolate manufacturer Callebaut. Like other types of chocolate, ruby chocolate is produced from cocoa beans, the seeds of the cocoa tree *Theobroma cacao* L. (Malvaceae). The exact production process of ruby chocolate is not known. However, in 2012, Callebaut obtained a European patent for a specific "process for producing cocoa derived material". As opposed to regular chocolate production, either underfermented cocoa beans, not fermented for more than three days, or preferably, unfermented cocoa beans, the so-called "Lavados" beans, meaning "washed" beans, are used (Dumarche et al., 2012). Since fermentation was found to drastically reduce the levels of polyphenols in cocoa beans (Nazaruddin, Seng, Hassan, & Said, 2006; Wollgast & Anklam, 2000), these Lavados beans typically possess a higher polyphenol content than fermented beans. The origin of the cocoa beans used for the production of ruby chocolate can be Brazil, Ecuador or Ivory coast (Callebaut, n.d.).

Furthermore, a treatment of the cocoa nibs with an acidic solution is introduced as an additional step in the cocoa processing. When soaking in, or spraying or washing with an acidic solution (for example with citric acid or phosphoric acid) is carried out under the appropriate conditions (pH, time, temperature, etc.), this will result in cocoa nibs with a red or purple color and this will help to preserve the polyphenolic content (Dumarche et al., 2012). According to the patent, the level of polyphenols in the nibs should comprise at least 20 mg/g and most preferably 40 to 60 mg/g polyphenols, expressed as epicatechin equivalents (Dumarche et al., 2012). Further processing of these nibs can result in the production of red or purple chocolates, with a higher content of polyphenols than other types of chocolate. Most probably, the lack of fermentation and the acidification step, described in Callebaut's patent, are the key steps in the production of the ruby chocolate.

From a quantitative point of view, polyphenols form a major class of phytochemicals in cocoa and chocolate. For example, in fresh cocoa beans of the Forastero variety, the total amount of soluble polyphenols in the dried, fat-free mass is 15-20%, while the level in fermented beans is reduced to approximately 5% (Wollgast & Anklam,

2000). These polyphenols can be divided into three main groups: flavan-3-ols (about 37% of total polyphenols; mainly epicatechin), proanthocyanidins (about 58%) and anthocyanins (about 4%) (Aprotosoaie, Luca, & Miron, 2016; Wollgast & Anklam, 2000). Another important class of compounds are the methylxanthines theobromine, caffeine, and theophylline, of which theobromine is the most abundant, with levels between 5 to 7 mg/g in dark chocolate (Aprotosoaie et al., 2016). In addition, various other minor constituents were reported in cocoa and chocolate, like biogenic amines (tyramine, tryptamine, phenylethylamine), *N*-phenylpropenoyl-L-amino acids (clovamide), endocannabinoids (anandamide) and alkaloids (salsolinol) (Aprotosoaie et al., 2016; Tuenter, Foubert, & Pieters, 2018).

Up till now, only one study reported the quantification of the total phenolic content, total flavonoid content and total amount of proanthocyanidins in ruby chocolate (Šeremet et al., 2019). However, to the best of our knowledge, no detailed investigation of the composition of ruby chocolate was carried out yet. Thus, in this research, ruby chocolate was subjected to two different analyses. Firstly, the ruby chocolate was analyzed with a UPLC-hybrid quadrupole-TOF MS/MS method, together with a dark chocolate sample, with the aim of identifying a wide range of phytochemicals. A feature-based molecular network was created, using the GNPS (Global Natural Products Social molecular networking) platform, to aid in the identification. In addition, dark, milk, ruby and white chocolate were analyzed with a UPLC-Orbitrap-MS method to further strengthen the tentative identifications. Secondly, the ruby chocolate was analyzed on a UPLC-QTOF-MS system, together with white, milk and dark chocolate, in order to quantify a targeted set of natural products. In addition, the raw dataset obtained during this analysis was used for PCA (principal component analysis). By combining both approaches, a vast amount of information was obtained on the phytochemical composition of ruby chocolate and a comparison can be made between ruby chocolate and other types of chocolate.

2. Materials and methods

2.1 Chemicals and materials

White chocolate callets (lot No. 191070202), ruby RB1 chocolate callets (lot No. 183340601, 191880601), milk chocolate callets (Power 41) (lot No. 190700202) and dark chocolate callets (Recipe N^{o} 70-30-38, lot No. 183170102, 190870201), all from Callebaut, were purchased in a Belgian baking shop.

UPLC-grade acetonitrile, methanol, formic acid and ammonium formate were purchased from Biosolve (Valkenswaard, The Netherlands). HPLC-grade acetone was obtained from Acros Organics (Geel, Belgium), and HPLC-grade *n*-hexane and methanol came from Fisher Scientific (Loughborough, UK). Ultrapure water was generated with a Direct Pure Up system of Rephile. Amberlite XAD-7HP was supplied by Acros Organics (Geel, Belgium). Analytical standards were purchased from Extrasynthese (Lyon, France), Sigma-Aldrich (Bornem, Belgium), Santa Cruz Biotechnology (Heidelberg, Germany), or Carl Roth (Karlsruhe, Germany).

2.2 Sample preparation

Prior to extraction, chocolates were milled with a mortar and pestle, under cooling with liquid nitrogen, and sieved (sieve width: 2 mm). Next, 1 g of each sample was defatted with *n*-hexane, according to the previously reported method (Tuenter et al., 2020). The defatting was carried out in triplicate. Extraction of 0.1 g of the defatted samples was performed three consecutive times with 2 mL of a mixture of 70.0/29.8/0.2 acetone/water/acetic acid. Samples were vortex mixed and submitted to ultrasonication for 30 min, followed by centrifugation (5 min, 1370 x G) and collection of the supernatant. The combined extracts were diluted prior to the quantitative analysis (vide infra).

For identification purposes, an additional sample preparation step was carried out, with the aim of enriching phenolic compounds in the extract, while removing undesired compounds such as glucose and fructose. For this purpose, 5 mL of the obtained extract was dried and 0.15 g of Amberlite XAD-7HP resin was added to the dried material. Next, this mixture was suspended in 20 mL of water and left to stir overnight, on a stirring plate, protected from light. Then, filtration was carried out, using a sand filter, followed by rinsing of the residue with 3 x 10 mL of water. In order to obtain the purified extract, 10 mL of methanol was added to the resin three

consecutive times, and sonicated for 20 min/10 min/10 min, for the first, second and third extraction, respectively, followed by filtration and collection of the extract. The total methanolic extract was then dried under reduced pressure and dissolved in methanol/water (80/20) at a concentration of about 1 mg/mL, and analyzed as such.

2.3 Identification of phytochemical constituents

Experiments aimed at the identification of marker compounds in ruby and dark chocolate were performed in triplicate with an AB SCIEX TripleTOF 5600+ Hybrid Quadrupole-TOF LC/MS/MS Mass Spectrometer (AB SCIEX, Concord, ON, Canada).

Chromatographic separation was performed using an Exion LC (Binary Gradient AC Pump, AC Autosampler, AC Column Oven, System Controller, SCIEX) and a Fortis speedcore diphenyl analytical column (2.1 × 100 mm i.d., 2.6 µm particle size, Fortis) was kept at 40°C. The injected sample volume was 10 µl. The mobile phase consisted of 0.1% formic acid in water (eluent A) and acetonitrile (eluent B). The gradient elution was performed as follows: 0–2 min eluent B 5%; 2–15 min eluent B 5–95%; 16–18 min eluent B 100%; 18–20 min column equilibration-eluent B 5%. A flow rate of 0.4 mL/min was employed.

The TripleTOF 5600+ utilized a Duo Spray source with separated electro-spray ionization (ESI) and atmospheric-pressure chemical ionization (APCI) probes. The ESI was used for mass spectrometric detection of compounds, while the APCI probe worked as the second gas heater; APCI was also utilized for exact mass calibration of the instrument. The negative ESI mode parameters were as follows: temperature 550°C, curtain gas (CUR) 35, GS1 and GS2 at 50 and 50, respectively, ion-spray voltage floating (ISVF) at 3.5 kV and declustering potential: 65 V. The TripleTOF 5600+ was operated using an information dependent acquisition (IDA) method to collect full scan MS and MS/MS information simultaneously. The negative polarity method consisted of a period (0–20.07 min) including a product ion scan (mass range *m/z* 100–1000, CE -30V, CES 15V, accumulation time 100 ms) and a TOF-MS scan (mass range *m/z* 100–1000, CE -10, accumulation time 250 ms). The duty cycle time for this period

was 1.5005 sec and the number of cycles 800. Maximum number of candidate ions to monitor per cycle: 12 spectra and Ions Tolerance: 50.000 mDa. Switch Criteria exceeded 500 cps and exclusion of isotopes was carried out within: 4.0000 Da. Dynamic background subtraction was activated to automatically acquire EPI (Enhanced Product Ion) spectra of co-eluting compounds. The exact mass calibration was performed automatically before each analysis employing the automated Calibration Delivery System (CDS). The Software Version was Analyst TF 1.7.1.

The raw data in wiff/wiff.scan format were converted to mzXML format, using the MSConvertGUI tool of ProteoWizard (version 3.0.19325-0a79a5f45).(Chambers et al., 2012) Data were then processed with MZmine, V.2.51, and submitted to the GNPS platform for feature based molecular networking.(Wang et al., 2016) To visualize the data, the output was imported into Cytoscape, V.3.7.2.(Shannon et al., 2003) Manual annotation of peak ions was performed based on predicted molecular formulae, fragmentation patterns and comparison to literature.

In parallel, the dark, milk, ruby and white chocolate samples were analyzed once each by UPLC-Orbitrap-MS in both ESI- and ESI+ mode, with data-dependent fragmentation, to obtain additional data with the aim of strengthening the tentative identifications of the UPLC-TripleTOF analysis. Liquid chromatography analysis was performed on an Acquity® UPLC System (Waters, Milford, MA, USA). Detection was performed on an LTQ-Orbitrap® XL hybrid mass spectrometer equipped with an ESI source (Thermo Scientific, Waltham, MA, USA) for accurate mass measurements. Separation was achieved on an Acquity UPLC® Peptide BEH C18 column (2.1 x 100 mm, 1.7 µm, Waters corporation®, Wexford, Ireland) using a water gradient containing 0.1% (v/v) formic acid (A) and acetonitrile (B). The gradient elution was performed as follows: 0–2 min eluent B 2%; 2–18 min eluent B 2–100%; 18–20 min eluent B 100%; 21–25 min column equilibration-eluent B 2%. A flow rate of 0.4 mL/min was employed for elution. The column was maintained at 40 °C, the samples at 7 °C. All samples were analyzed in the full scan *m/z* range of 115–1000, in negative and positive mode at a resolving power of 30,000 and data-dependent MS/MS events were acquired. In both modes the data-dependent acquisition was simultaneously performed using a collision induced dissociation C-trap (CID) with normalized collision energy at 35 V and a mass

resolution of 10,000. In negative mode capillary temperature was set to 350 °C and the source voltage was 2.7 kV. Tube lens and capillary voltage were respectively tuned at −100 V and −30 V. In positive mode capillary temperature was set to 350 °C and the source voltage was 3.50 kV. Tube lens and capillary voltage were respectively tuned at +120 V and +40 V. In both modes the arbitrary units were used for sheath gas, auxiliary gas and sweep gas was nitrogen at (40,10,0 arbitrary units respectively). The control of the system and the spectral interpretation was performed using the XcaliburTM (Version 2.2, Thermo Scientific) software.

2.4 Quantification of phytochemical constituents

White, ruby, milk and dark chocolates were analyzed in triplicate on a UPLC-QTOF-MS system, as described previously (Tuenter et al., 2020). Detection was carried out in ESI+ mode and ESI- mode, *m/z* scan range 50-1500, with the capillary voltage set at 1.0 kV and 0.8 kV, respectively. A selected set of 51 reference compounds was prepared as described before (Tuenter et al., 2020) and analyzed under the same conditions in concentrations ranging from 2 ng/mL to 2 µg/mL. A complete list of these compounds can be found in the Supplementary information, S1. Calibration curves were constructed, in order to allow a quantitative determination of these compounds in the samples. All chocolate extracts were analyzed after a ten-, hundred- or thousand-fold dilution in a methanol/40 mM ammonium formate buffer (60/40). Data were processed using MassLynx V.4.1 and creation of calibration curves and statistical analysis were carried out with GraphPad Prism 8. Limit of detection (LOD) and limit of quantification (LOQ) were determined for each compound, according to the “blank sample” method, as described by Mani et al. (Mani, Abbatiello, & Carr, 2012). The unpaired t-test or one-way Anova (Analysis of variance) were carried out with GraphPad Prism 8 and were used to determine statistically significant differences between levels of two, or three or more chocolates, respectively.

2.5 Principal Component Analysis

The complete .raw dataset obtained with the UPLC-QTOF-MS for each sample in both ESI+ and ESI- mode was converted into .mzXML format using a dedicated R-script, Rstudio version 1.1.463 (Boston, MA, USA), as well as MassWolf V. 4.3.1 and MassLynx V. 4.1. Next, Workflow4Metabolomics, part of the Galaxy platform, was used for data pre-processing of the LC-MS data (Gatto & Lilley, 2012; Giacomoni et al., 2015; Kuhl, Tautenhahn, Böttcher, Larson, & Neumann, 2012; Smith, Want, O'Maille, Abagyan, & Siuzdak, 2006). Finally, MetaboAnalyst was used for performing the PCA (Chong, Wishart, & Xia, 2019) on the data matrices from both ionization modes. In order to reduce the number of features from 19,535 (ESI+) and 9,747 (ESI-) to 2,500 features, data-filtering was performed using the interquantile range. Next, normalization by sum and Pareto scaling were carried out, followed by PCA.

3. Results and discussion

3.1 Identification of phytochemical constituents

After analysis of ruby and dark chocolate by means of a UPLC-hybrid quadrupole-TOF-MS/MS method and feature based molecular networking, 41 compounds could be tentatively identified, 30 of which in ruby chocolate and 31 of which in dark chocolate. Detailed information regarding the identification is shown in Table 1.

A first class of compounds that was detected were the proanthocyanidins, which are a major type of polyphenols in cocoa, as mentioned earlier. Twelve different proanthocyanidins, ranging from dimers up till pentamers were observed. Identification of the dimers and trimers was supported by clustering in the feature-based molecular network (Figure 1). Two procyanidin B-type dimers were identified in both dark and ruby chocolate (**1**, **2**), according to their accurate *m/z* measurement, fragmentation pattern, and comparison to literature data (Cádiz-Gurrea et al., 2014; Hellström, Sinkkonen, Karonen, & Mattila, 2007; Rue, Rush, & van Breemen, 2018). Determination of the specific procyanidin B-type dimer compounds was not possible in this stage. However, the quantitative analysis revealed the presence of both procyanidin B1 and procyanidin B2 in ruby and dark

chocolate, and thus, the two procyanidin B-type dimers detected here were tentatively assigned as such. Also a procyanidin B-type trimer (**3**) was detected in the ruby and dark chocolate, but higher oligomers of the procyanidin B-type (tetramer (**4**) and pentamer (**5**)) could only be observed in the dark chocolate and were not found in the ruby chocolate.

Interestingly, the ruby chocolate contained various proanthocyanidins of the A-type (dimer, trimer, tetramer and pentamer (**6**-**9**)), which could not be found in the dark chocolate. Whereas proanthocyanidins of the B-type consist of flavan-3-ol monomers, linked via single bonds (4→8 or 4→6), proanthocyanidins of the A-type are built up of flavan-3-ol monomers with a double ligation, for example via 4→8 and 2→O→7 links. Previously, procyanidin A-type trimers were tentatively identified in cocoa, based on the observation of ions ($[M-H]^-$) with a *m/z* 861.1686, corresponding to a molecular formula of $C_{45}H_{34}O_{18}$ (Patras, Milev, Vrancken, & Kuhnert, 2014; Tuenter et al., 2020). However, due to the low abundance in the studied cocoa samples (raw fermented cocoa beans (Patras et al., 2014), cocoa liquor and dark chocolate (Tuenter et al., 2020)), no MS/MS data were available to further confirm this identification. In this study, a clear fragmentation of the molecular ion with *m/z* 861 was obtained in the ruby chocolate and its MS fragmentation spectrum can be found in the Supplementary data, S2. Several of the observed fragments can be explained, when applying fragmentation pathways similar to those reported for procyanidin A-type dimers, such as quinone methide reaction and heterocyclic ring fission (Figure 2) (Rue et al., 2018; Sui et al., 2016). Moreover, the link between the molecular ion with *m/z* 575.1199, tentatively identified as a procyanidin A-type dimer, and the molecular ion with *m/z* 861.1686, observed in the molecular network (Figure 1), supports this hypothesis. Thus, our observations provide additional proof for the tentative identification of the procyanidin A-type trimer, which was clearly detected in the ruby chocolate, while it was not found in the dark chocolate. Moreover, to the best of our knowledge, this is the first report of tetrameric and pentameric A-type proanthocyanidins in chocolate.

In dark chocolate, a compound with *m/z* value 591.1520 (**12**) was found, which could not be detected in the ruby chocolate. This compound was tentatively identified as bis-8,8-catechinylmethane, previously reported in cocoa liquor (Hatano et al., 2002). Fragments with *m/z* 439 and 289 are resulting from a retro Diels-Alder reaction

(RDA), with a characteristic neutral loss of 152 Da, and from fission resulting in a (epi)catechin-fragment (Supplementary data, S3), thus supporting its identification. Its MS fragmentation spectrum can be found in the Supplementary data, S4. In contrast, a compound with *m/z* value 423.0717 and most probable molecular formula $C_{22}H_{16}O_9$ (**13**) could only be detected in ruby chocolate and not in dark chocolate. Its complete structure could not be determined, but the observed fragment with *m/z* 285 indicates the presence of a flavan-3-ol moiety. Moreover, in the molecular network, this compound was part of the same cluster as other proanthocyanidins, indicating their resemblances regarding MS fragmentation. Possibly, this compound is catechinlactone A, previously reported in *Euonymus alatus* (Zhang et al., 2013), but no MS/MS data were found in literature, to compare with our experimental data and to further support this hypothesis.

With regard to the flavan-3-ol monomers, only one clear signal with *m/z* 289.0701 was found in dark and ruby chocolate (**14**). It is well-known that epicatechin is the most abundant flavan-3-ol in cocoa and chocolate, which was supported by the quantitative analysis (vide infra) and thus, this compound was tentatively identified as such. Although not observed during the current analysis, catechin could be identified and quantified during the quantitative analysis too. The fact that an additional clean-up step was used for the qualitative, but not for the quantitative analysis, as well as the different instrumental conditions applied, may explain why catechin could not be detected during both UPLC-MS analyses.

Table 1. Tentatively identified compounds in dark and ruby chocolate, based on chromatographic and spectral data, obtained with UPLC-hybrid quadrupole-TOF MS/MS analysis.

Compound number	Compound name	RT (min)	Ion	Measured *m/z*	Calculated *m/z**	Error (ppm)	RDB	Fragmentation pattern	Molecular formula	Dark chocolate	Ruby chocolate	Remarks	References
1	Procyanidin B-type dimer	4.58	$[M-H]^-$	577.1342	577.1346	-0.69	18.5	451.1005; 425.0862; 407.0766; 289.0731; 287.0569	$C_{30}H_{26}O_{12}$	x	x	Dark > ruby. Most probably procyanidin B1.	α,β,γ
2	Procyanidin B-type dimer	5.28	$[M-H]^-$	577.1380	577.1346	5.89	18.5	451.1067; 425.0893; 407.0681; 289.0729; 287.0567	$C_{30}H_{26}O_{12}$	x	x	Dark > ruby. Most probably procyanidin B2.	α,β,γ
3	Procyanidin B-type trimer	5.1	$[M-H]^-$	865.2003	865.1980	2.66	27.5	739.1736; 713.1581; 695.1446; 577.1358; 575.1151; 425.0880; 407.0812; 289.0747; 287.0562	$C_{45}H_{38}O_{18}$	x	x	Dark > ruby. Possibly procyanidin C1.	β,γ
4	Procyanidin B-type tetramer	5.16	$[M-2H]^{2-}$	576.1296	576.1270	4.51	37.0	Non fragmented	$C_{60}H_{50}O_{24}$	x		Possibly Cinnamtannin A2 (reported by Hatano et al. (2002).	β,δ
5	Procyanidin B-type pentamer	5.23	$[M-2H]^{2-}$	720.1604	720.1585	2.64	46.0	644.1333; 635.1309; 577.1413; 575.1138; 449.0837; 407.0765; 289.0703; 287.0518	$C_{75}H_{62}O_{30}$	x			δ,ε
6	Procyanidin A-type dimer	5.69	$[M-H]^-$	575.1199	575.1190	1.56	19.5	539.1006; 449.0882; 423.0726; 289.0704; 285.0403	$C_{30}H_{24}O_{12}$		x	Ruby >> dark.	γ,η
7	Procyanidin A-type trimer	5.99	$[M-H]^-$	861.1686	861.1667	2.21	29.5	825.1586; 735.1432; 575.1204; 571.0895; 539.1010; 449.0916; 289.0707; 285.0406	$C_{45}H_{34}O_{18}$		x	Ruby >> dark.	ε,η
8	Procyanidin A-type tetramer	6.30	$[M-2H]^{2-}$	573.1080	573.1035	7.85	39.5	735.1492; 571.0911; 539.0991; 449.0867; 289.0713; 285.0411	$C_{60}H_{44}O_{24}$		x		
9	Procyanidin A-type pentamer	6.45	$[M-2H]^{2-}$	716.1307	716.1270	5.17	46.5	571.0899; 539.1046; 449.0887; 289.0730; 285.0404	$C_{75}H_{60}O_{30}$		x		
10	Procyanidin A-type dimer - pentoside	5.84	$[M-H]^-$	707.1656	707.1612	6.22	20.5	581.1353; 539.1010; 449.0902; 407.0780; 289.0732; 287.0569	$C_{35}H_{32}O_{16}$	x	x	Pentoside: most probably arabinoside.	β,δ,ε
11	Procyanidin A-type dimer - hexoside	5.71	$[M-H]^-$	737.1697	737.1718	-2.85	20.5	611.1412; 539.0984; 449.0884; 407.0768; 289.0730	$C_{36}H_{34}O_{17}$	x	x		β
12	Bis-8,8-catechylmethane	5.71	$[M-H]^-$	591.1520	591.1503	2.88	18.5	439.1004; 301.0708; 289.0706; 245.0822; 215.0705	$C_{31}H_{28}O_{12}$	x			β
13	Not identified	5.69	$[M-H]^-$	423.0717	423.0716	0.24	15.5	285.0403; 257.0460; 137.0240	$C_{22}H_{16}O_9$		x		
14	Epicatechin§	4.87	$[M-H]^-$	289.0701	289.0712	-3.81	9.5	245.0818; 221.0816; 205.0517; 203.0703; 125.0247; 123.0441; 109.0293	$C_{15}H_{14}O_6$	x	x		ε,η,θ
15	(Epi)catechine sulfonic acid	4.52	$[M-H]^-$	369.0295	369.0280	4.06	9.5	289.0733; 245.0819; 216.9798; 205.0492; 203.0718; 137.0221	$C_{15}H_{14}O_9S$	x			ε,η
16	Apigenin§ (trihydroxy-flavone)	7.57	$[M-H]^-$	269.0455	269.0450	1.86	11.5	240.0442; 225.0561; 224.0474; 201.0561; 181.0649; 159.0432;	$C_{15}H_{10}O_5$	x	x		

								151.0030; 135.0466; 133.0275; 117.0344					
17	Luteolin[§] (tetrahydroxy-flavone	6.94	$[M-H]^-$	285.0404	285.0399	1.75	11.5	267.0287; 257.0440; 241.0499; 217.0490; 201.0165; 199.0407; 175.0408; 151.0039; 133.0301	$C_{15}H_{10}O_6$	x	x	Dark > ruby.	η,ι
18	Tetrahydroxy-flavone-pentoside	4.58	$[M-H]^-$	417.0821	417.0822	-0.24	12.5	285.0389; 284.0334	$C_{20}H_{19}O_{10}$	x	x	Most probably luteolin- or kaempferol-pentoside.	η
19	Tetrahydroxy-flavone-hexoside	5.79	$[M-H]^-$	447.0928	447.0927	0.22	12.5	285.0386	$C_{21}H_{20}O_{11}$	x		Most probably luteolin-or kaempferol-hexoside.	ι
20	Pentahydroxy-flavanone	5.36	$[M-H]^-$	303.0490	303.0505	-4.95	10.5	285.0375; 256.9118; 192.8947; 165.0181; 151.0417;137.0252; 125.0260	$C_{15}H_{12}O_7$		x	Most probably dihydroquercetin/ taxifolin.	
21	Quercetin[§] (3,3',4',5,7-pentahydroxy-flavone)	6.96	$[M-H]^-$	301.0338	301.0348	-3.32	11.5	273.0404; 245.0466; 178.9987; 151.0039; 121.0297	$C_{15}H_{10}O_7$	x	x		η,θ
22	Quercetin-pentoside	5.9	$[M-H]^-$	433.0744	433.0771	-6.23	12.5	300.0278; 271.0256; 255.0305; 243.0302; 178.9970; 151.0024	$C_{20}H_{18}O_{11}$	x	x	Most probably quercetin-3-O-arabinoside (avicularoside).	η,θ
23	Quercetin-hexoside	5.65	$[M-H]^-$	463.0865	463.0877	-2.59	12.5	300.0277; 271.0255; 255.0282; 178.9971; 151.0025	$C_{21}H_{20}O_{12}$	x	x	Most probably quercetin-3-O-glucoside (isoquercitrin) or quercetin-3-O-galactoside (hyperoside).	η,θ
24	Cinchonain I	6.55	$[M-H]^-$	451.1034	451.1029	1.11	15.5	388.192; 341.0662; 297.0790; 240.9993; 217.0135; 189.0187; 177.0197	$C_{24}H_{20}O_9$	x	x	Dark > ruby. Flavonolignan.	β
25	Clovamide	5.31	$[M-H]^-$	358.0937	358.0927	2.79	11.5	222.0405; 178.0512; 161.0239; 135.0448	$C_{18}H_{17}NO_7$	x	x	Dark > ruby.	β,ε,η
26	Deoxyclovamide (N-caffeoyl tyrosine)	5.66	$[M-H]^-$	342.0976	342.0978	-0.58	11.5	222.0409; 206.0476; 178.0516; 163.0384; 161.0229; 145.0293; 135.0436; 119.0475	$C_{18}H_{17}NO_6$	x			β,ε
27	Dideoxyclovamide (N-coumaroyl tyrosine)	6.09	$[M-H]^-$	326.1042	326.1028	4.29	11.5	282.1125; 206.0444; 119.0505	$C_{18}H_{17}NO_5$	x	x		β,ε
28	N-Coumaroyl-L-aspartic acid (= L-aspartic acid, N-[3-(4-hydroxyphenyl)-1-oxo-2-propenyl)	4.44	$[M-H]^-$	278.0670	278.0665	1.80	8.5	234.0801; 216.0668; 190.0852; 162.0565; 147.0451; 145.0308; 132.0301; 119.0487; 115.0010	$C_{13}H_{13}NO_6$	x	x		β,η
29	L-aspartic acid, N-[3-(4-hydroxyl-3-methoxyphenyl)-1-oxo-2-propenyl	4.91	$[M-H]^-$	308.0778	308.0770	2.60	8.5	Non fragmented	$C_{14}H_{15}NO_7$	x			β
30	Dihexose	0.7	$[M-H]^-$	341.1081	341.1084	-0.88	2.5	179.0558; 161.0459; 143.0376; 119.0354; 113.0244	$C_{12}H_{22}O_{11}$	x	x		β,δ
31	Not identified	6.38	$[M-H]^-$	427.1623	427.1604	4.45	7.5	325.1339; 261.0978; 221.0731; 179.0585; 161.0464; 143.0345; 125.0221; 101.0230	$C_{20}H_{28}O_{10}$	x	x	Dark > ruby. Several fragments (*m/z* 179, 161, 143) indicate the presense of a hexoside moiety.	δ
32	Protocatechuic acid[§]	1.86	$[M-H]^-$	153.0195	153.0188	4.57	5.5	Non fragmented	$C_7H_6O_4$	x	x		

33	Caffeic acid	4.37	$[M-H]^-$	179.0348	179.0344	2.23	6.5	135.0449; 134.0373	$C_9H_8O_4$		x		θ, κ
34	Vanillic acid	5.35	$[M-H]^-$	167.0348	167.0344	2.39	5.5	152.0092; 108.0197	$C_8H_8O_4$	x	x		κ
35	Vanillic acid - hexoside	4.25	$[M-H]^-$	329.0892	329.0873	5.77	6.5	310.9144; 292.9039; 274.8942; 167.0338; 123.0441	$C_{14}H_{18}O_9$		x	*m/z* 167 corresponds to $[M\text{-hexoside-H}]^-$	δ
36	Hydroxy-jasmonic acid sulphate	4.97	$[M-H]^-$	305.0700	305.0695	1.64	4.5	225.1135	$C_{12}H_{18}O_7S$	x	x	Dark > ruby	ε,η
37	Not identified	6.05	$[M-H]^-$	393.1774	393.1761	3.31	3.5	331.1725; 291.1471; 249.1342; 161.0444; 125.0235; 101.0228	$C_{17}H_{30}O_{10}$	x	x	Dark > ruby	β,δ,ε, η
38	Not identified	7.41	$[M-H]^-$	421.2065	421.2074	-2.14	3.5	359.2138; 319.1761; 277.1666; 125.0238; 101.0231	$C_{19}H_{34}O_{10}$	x	x	Difference with compound 37: 28 amu, corresponding to an additional C_2H_4. The same difference is reflected in several fragments. Fragments with *m/z* 125 and 101 observed for both compounds 37 and 38 → structural similarity likely.	δ,ε,η
39	Not identified	6.36	$[M-H]^-$	503.1756	503.1765	-1.79	7.5	419.1141; 233.0818; 209.0828; 125.0251	$C_{22}H_{32}O_{13}$	x	x		δ
40	Not identified	6.43	$[M-H]^-$	459.1283	459.1291	-1.74	12.5	423.088; 384.1480; 297.0793; 191.0343	$C_{23}H_{24}O_{10}$	x			δ
41	Not identified	6.46	$[M-H]^-$	516.2462	516.2445	3.29	6.5	237.0996; 184.0730	$C_{24}H_{39}NO_{11}$	x			β,δ

*https://www.lfd.uci.edu/~gohlke/molmass/. RDBeq: ring double bond equivalent. §Identification confirmed with analytical standard (during quantitative analysis). α: Hellstrom et al. (2007), β: Cádiz-Gurrea et al. (2014), γ: Rue et al.(2018), δ: D'Souza (2017), ε: Patras et al. (2014), ζ: Sui et al. (2016), η: Tuenter et al. (2020), θ: Kumar et al. (2017), ι: Guan et al. (2014), κ: Gruz et al. (2008)

Several other flavonoids, eluting between 4.5 and 7.6 min, were tentatively identified, based on the experimental MS and MS/MS data and comparison to literature (**15**-**23**) (Guan et al., 2014; Kumar, Singh, & Kumar, 2017; Patras et al., 2014; Tuenter et al., 2020). A tetrahydroxyflavone-pentoside (**18**) and –hexoside (**19**) were identified based on accurate mass. These compounds are structural isomers of the anthocyanidines cyanidin-3-*O*-arabinoside and cyanidin-3-*O*-galactoside (ideain), respectively, which are reported constituents of cocoa too (Aprotosoaie et al., 2016; Tuenter et al., 2020; Wollgast & Anklam, 2000). The MS/MS data did not provide additional proof for the identification as tetrahydroxyflavone-glycosides. However, the retention times of these compounds are in the same range as those of the other identified flavonoids. Since cyanidin and its glycosides are molecules with an intrinsic positive charge, it may be expected that these compounds would elute earlier in a reversed phase system, compared to the neutral flavonoids, which is in agreement with previously published experimental data (Tuenter et al., 2020). Thus, indeed, the compounds observed here are most probably tetrahydroxyflavone-glycosides, of which the –hexoside could only be detected in dark chocolate. In addition, a pentahydroxyflavanone (**20**) was only detected in ruby chocolate, which could correspond to dihydroquercetin/taxifolin. Apart from these flavonoids, one flavonolignan, cinchonain I (**24**) was identified in ruby and dark chocolate, after comparison to literature (Cádiz-Gurrea et al., 2014).

Furthermore, the *N*-containing compounds clovamide (**25**), deoxyclovamide (**26**) and dideoxyclovamide (**27**) were identified, as were *N*-coumaroyl-L-aspartic acid (**28**) and L-aspartic acid, *N*-[3-(4-hydroxyl-3-methoxyphenyl)-1-oxo-2-propenyl (**29**). However, compounds **27** and **29** were only found in dark chocolate.

Also several organic acids (**32**-**36**) were identified. Caffeic acid (**33**) was identified in ruby chocolate only, by comparison of its MS fragmentation to literature (Gruz, Novák, & Strnad, 2008; Kumar et al., 2017), and the quantitative analysis, using reference compounds, confirmed this finding (vide infra). Also a vanillic acid-hexoside (**35**) was identified by comparison to literature (D'Souza et al., 2017) and was only found in the ruby chocolate.

Compounds **37** and **38**, with *m/z* 393.1774 and 421.2065, were detected in cocoa and chocolate before, with most probable molecular formulas $C_{17}H_{30}O_{10}$ and $C_{19}H_{34}O_{10}$. According to D'Souza et al. (2017), they were

tentatively identified as hexenyl xylopyranosyl glucopyranoside and octenyl xylopyranosyl glucopyranoside (octen-3-yl primeveroside), but data on MS fragmentation were not reported. Cádiz-Gurrea et al. (2014) tentatively identified two compounds with *m/z* 393 as hexenyl xylopyranosyl glucopyranoside isomers in a *T. cacao* extract, with one MS fragment with *m/z* 249 detected, which was also found in our data. Given the similarities in the MS/MS data that we obtained, structural similarities between these two compounds may be expected indeed. However, since a logical explanation for the obtained MS/MS data in view of the proposed structures could not be found, unfortunately their identification remains doubtful.

Finally, three other peaks were detected (**39**-**41**), of which the *m/z* values and most probable molecular formulas were reported before, but their exact structures are not determined up until now (Cádiz-Gurrea et al., 2014; D'Souza et al., 2017).

In general, when looking at Table 1, it can be concluded that a majority of the identified compounds is present in both dark and ruby chocolate, and few compounds are present in either one of them. The most striking difference can be found in the occurrence of the A- and B-type proanthocyanidins, as mentioned earlier. Although this analysis was not carried out with the purpose of quantification, it must be mentioned that some clear differences in intensity of some of the identified peaks were observed, which is described in the "remarks" column in Table 1.

The data obtained with the UPLC-Orbitrap-MS corroborated the results of the UPLC-TripleTOF analysis. Except for the procyanidin A-type pentamer (**9**) and cinchonain I (**24**), all compounds listed in Table 1 were detected ESI- mode, and, in addition, 17 of the 41 compounds could also be detected in ESI+ mode, based on their retention time and accurate mass measurement. The results of this analysis can be found in the Supplementary data, S5. While the UPLC-TripleTOF analysis concerned the dark and ruby chocolate only, milk chocolate and white chocolate were also included in the UPLC-Orbitrap-MS analysis. It was found that all compounds present in dark chocolate, were also present in milk chocolate. This is not unexpected, since both chocolates contain cocoa powder, although their cocoa content differs (70.5% vs. 40.7% for dark and milk chocolate, respectively). Five of

294 the compounds could also be detected in white chocolate. These included the dihexose (**30**), which may be
295 expected, but also four flavonoids (**16**, **17**, **22**, **23**). Since white chocolate contains cocoa butter, but no cocoa
296 powder, the finding of the latter four compounds was rather unexpected.

297

298 *3.2 Quantification of phytochemical constituents*

299 In addition to the first part of this study, which was focused on the identification of phytochemicals in dark and
300 ruby chocolate, in this part the quantification of a targeted set of phytochemicals in dark, ruby, milk and white
301 chocolate is described. A total of 19 compounds could be quantified: 18 in dark chocolate, 15 in milk chocolate,
302 13 in ruby chocolate and 2 in white chocolate, as reported in Table 2. These compounds include flavan-3-ols,
303 proanthocyanidins, flavonoids, methylxanthines, amino acids, biogenic amines and organic acids (Figure 3).
304 Interestingly, 13 out of these 19 compounds were detected, in addition to the compounds already mentioned in
305 Table 1. These compounds mainly include *N*-containing substances, identified during the UPLC-HRMS analysis in
306 ESI+ mode, thus bringing the total of identified compounds at 54. The *N*-containing substances most probably
307 were not retained in sufficient amount by the Amberlite resin, applied in the sample preparation of the
308 identification analysis, and as a consequence, were only identified in this quantificational analysis.

309 When comparing the levels of the different compounds between the chocolates, almost all compounds are most
310 abundant in the dark chocolate. This is in correspondence to the percentage of cocoa powder present in each of
311 the chocolates: the dark chocolate is composed of at least 70.5% cocoa powder, while the analyzed ruby and
312 milk chocolate contain at least 47.3% and 40.7% cocoa powder, respectively. An exception is caffeic acid, which
313 could only be quantified in the ruby chocolate, while levels in the other chocolates were below the LOQ of 0.41
314 µg/g chocolate. On the other hand, only two compounds could be quantified in white chocolate, which may be
315 expected, given the fact that white chocolate contains cocoa butter, but no cocoa powder. The two
316 phytochemicals that were quantified were caffeine and L-tryptophan, and according to the results of the Anova,

their levels were lower in white chocolate compared to milk and dark chocolate, but not significantly different from the levels in ruby chocolate.

The most interesting results are found when comparing the levels of polyphenols in ruby and milk chocolate, which are the most similar regarding cocoa powder content. It was found that catechin and quercetin are more abundant in milk chocolate (34.9 and 1.74 μg/g in milk chocolate vs. 17.28 and <1.23 μg/g in ruby chocolate, respectively), while epicatechin and procyanidin B2 are present in higher concentrations in the ruby chocolate (120 and 67 μg/g in milk chocolate vs. 244 and 167 μg/g in ruby chocolate, respectively). The level of procyanidin B1 did not differ significantly between the ruby and dark chocolate.

A possible explanation for the difference in the level of epicatechin may be the absence of, or shorter duration of fermentation of the cocoa beans used in ruby chocolate vs. milk and dark chocolate. During fermentation the levels of flavan-3-ols are known to decrease significantly, due to diffusion into fermentation sweatings and due to oxidation and polymerization reactions, resulting in the formation of mostly insoluble tannins (Nazaruddin et al., 2006; Wollgast & Anklam, 2000). Thus, the higher level of epicatechin in ruby chocolate vs. milk chocolate is plausible, since less epicatechin is lost during the processing. However, the opposite is observed for the catechin-level in ruby and milk chocolate: the concentration of catechin is higher in milk chocolate than in ruby chocolate. This cannot be explained by the aforementioned processes, but may be the result of epimerization taking place during fermentation. In raw cocoa beans, (+)-catechin and (-)-epicatechin are the major stereoisomers present, but during fermentation, (-)-epicatechin may be partly converted into (-)-catechin, as reported by Hurst et al. (2011). In the cocoa beans used for the production of milk chocolate, the formation of catechin due to epimerization may have outweighed the reduction, due to oxidation, polymerization and diffusion, resulting in a catechin-level that was even higher than in the unfermented, raw beans (used for the ruby chocolate).

Table 2. Levels of various compounds in different types of chocolate. Values are reported in µg/g. Data are mean ± SD (n = 3).

	Compound	Mode	White chocolate	Ruby chocolate	Milk chocolate	Dark chocolate	LOQ (µg/g)
Polyphenols	***Catechin***	ESI+	<LOD	17.28±0.97^{c}	34.9±5.4^{b}	176.4±1.2^{a}	0.73
	Epicatechin	ESI+	<LOD	244±18^{b}	120±20^{c}	489±29^{a}	1.04
	Procyanidin B1	ESI+	<LOD	5.085±0.084^{b}	6.64±0.85^{b}	33.08±0.95^{a}	0.32
	Procyanidin B2	ESI+	<LOD	167±17^{b}	67±15^{c}	251±31^{a}	0.82
	Quercetin	ESI-	<LOD	<LOQ	1.74±0.33^{b}	5.41±0.28^{a}	1.23
	Luteolin	ESI-	<LOD	<LOD	<LOQ	1.581±0.076	0.61
	Naringenin	ESI-	<LOQ	<LOQ	<LOQ	0.3416±0.0079	0.21
Methyl-xanthines	***Theobromine***	ESI+	<LOQ	354.5±6.3^{c}	(1.18±0.19)·10$^{3\,b}$	(4.924±0.077)·10$^{3\,a}$	9.24
	Caffeine	ESI+	14.5±0.9^{c}	37.0±1.2^{c}	97±18^{b}	529.7±9.0^{a}	1.14
Amino acids	***L-tryptophan***	ESI+	0.42±0.06^{c}	1.297±0.068^{c}	6.5±1.2^{b}	23.75±0.42^{a}	0.32
	L-α-Phenylalanine	ESI+	<LOQ	2.076±0.026^{c}	44.2±7.6^{b}	154.5±3.2^{a}	0.70
	L-tyrosine	ESI+	<LOD	<LOQ	28.9±5.3$^{1.b}$	112.4±1.8^{a}	3.16
Biogenic amines	***Tryptamine***	ESI+	<LOQ	<LOQ	0.327±0.053^{b}	1.35±0.02^{a}	0.18
	5-hydroxytryptophan	ESI-	<LOD	<LOD	0.98±0.25^{b}	3.41±0.32^{a}	0.55
	Serotonin	ESI+	<LOD	0.55±0.21^{b}	<LOQ	4.26±0.12^{a}	0.24
	Tyramine	ESI+	<LOQ	<LOQ	0.95±0.11^{b}	2.79±0.14^{a}	0.32
	Salsolinol	ESI+	<LOQ	0.746±0.036^{c}	4.06±0.63^{b}	13.24±0.27^{a}	0.22
Phenolic acids	*Protocatechuic acid*	ESI-	<LOQ	1.345±0.079^{c}	4.95±0.95^{b}	28.37±0.86^{a}	0.32
	Chlorogenic acid	ESI-	<LOQ	0.528±0.041^{a}	<LOQ	0.64±0.20^{a}	0.33
	Caffeic acid	ESI-	<LOQ	0.419±0.052	<LOQ	<LOQ	0.41

Compounds in bold were identified during the quantitative analysis, while they were not detected during the identification analysis. Different superscripts in a row point out significant differences ($p < 0.05$).

When comparing the levels of procyanidin B1 and B2 in ruby and milk chocolate, the different fermentation conditions may play a role too. Wollgast and Anklam (2000) reported that the levels of procyanidin dimers may reduce two- to threefold during fermentation. Although this statement is in agreement with our data obtained for procyanidin B2, it does not explain why the levels of procyanidin B1 did not differ significantly between the ruby and milk chocolate. We hypothesize that an epimerization of procyanidin B2, similar to the epimerization of epicatechin, may take place during fermentation too. Although epimerization of procyanidin B2 was found during cocoa bean roasting (Kothe, Zimmermann, & Galensa, 2013), it has not been reported during fermentation yet. However, unpublished data of the analysis of roasted cocoa beans carried out in our lab, using

353 the same UPLC-QTOF-MS method, showed a reduction in procyanidin B2, while the level of procyanidin B1
354 seemed to increase, which did not correspond to data reported in literature (Kothe et al., 2013). A possible
355 explanation for that finding may be that the epimerization products of procyanidin B2 co-elute with procyanidin
356 B1. Since these compounds possess the same molecular weight, they cannot be distinguished using the applied
357 UPLC-QTOF-MS method and we hypothesize now that the same may have happened when analyzing the ruby
358 and milk chocolate. Thus, the level of procyanidin B1 in milk chocolate may be overestimated and the actual level
359 may be lower than reported, and possibly even lower than the level in ruby chocolate.

360 Šeremet et al. (2019) determined the total phenolic content, and levels of flavan-3-ols and proanthocyanidins in
361 different types of chocolate. They found that the total phenolic content was lower in ruby chocolate than in milk
362 chocolate and dark chocolate. However, the level of flavan-3-ols was higher in ruby chocolate, compared to milk
363 chocolate, but lower compared to dark chocolate. These results are in agreement with our data, which show that
364 epicatechin, which is the major flavan-3-ol in chocolate and thus of big influence on the total flavan-3-ol content,
365 is present in ruby chocolate in a concentration about twice as high as in milk chocolate (244 vs. 120 µg/g,
366 respectively). Moreover, based on our data, the sum of the catechin and epicatechin content in ruby chocolate
367 is higher than in milk chocolate (261 vs. 154 µg/g, respectively). Moreover, Šeremet et al. found a higher level of
368 proanthocyanidins in ruby chocolate, compared to milk chocolate. Although we only quantified two of the major
369 proanthocyanidins, procyanidin B1 and B2, our data are again in agreement with the finding of Šeremet et al.,
370 since we obtained a total level of procyanidin B-dimers in ruby chocolate (172 µg/g) compared to milk chocolate
371 (73 µg/g).

372 Since the ruby and milk chocolate contain similar levels of cocoa powder, the finding that all quantified
373 methylxanthines, amino acids and biogenic amines are less abundant in ruby compared to milk chocolate was
374 unexpected. Possibly these differences can be accounted to the different origin of the cocoa starting material,
375 and/or the production process. With regard to the different origin of the cocoa, a possible future investigation
376 could be directed to NMR (nuclear magnetic resonance) metabolomics, instead of MS metabolomics. NMR in

377 combination with multivariate data-analysis was recently reported as potent approach to distinguish dark
378 chocolates with different countries of origin (Le Gresley & Peron, 2019). It would be interesting to see if this
379 approach can be extended to ruby and milk chocolate.

380 Finally, two out of three quantified organic acids, namely chlorogenic acid and caffeic acid, were more abundant
381 in ruby chocolate, while their levels were <LOQ in milk chocolate. Although only a limited number of organic
382 acids were quantified during this assay, our findings are in line with information obtained from the webpage of
383 Callebaut (Callebaut, n.d.). According to this webpage, the taste profile of ruby chocolate is characterized by
384 fruitiness, sourness and sweetness, and the results obtained for the two organic acids may in part be responsible
385 for the sour flavor notes.

386 *3.3 Principal Component Analysis*

387 The two datasets used as input for PCA (one from LC-MS analysis in ESI+ and one in ESI- mode) resulted in two
388 score plots in which the four different types of chocolate can be clearly distinguished (Figure 4). Figure 4 also
389 shows the corresponding loading plots, with each feature encoded according to its *m/z* value (M) and retention
390 time (T) in seconds (general format: M…T…). For the data obtained in ESI+ mode, PC1 accounted for 50.8% of
391 the variability and PC2 for 8.8%. Dark chocolate is characterized by low values of PC1, and from the loading plot
392 it can be deduced that this corresponds to a high level of theobromine (feature M181T212_2), in particular. Milk
393 chocolate shows less typical features, while white chocolate corresponds to high values of PC1, influenced by
394 high levels of feature M365T47, which was not identified. Procyanidin B2 (M579T388_2) and epicatechin
395 (M291T1422_1) are markers for both dark and ruby chocolate. Some other features, typical for ruby chocolate
396 (*m/z* 129, *m/z* 183, both eluting after almost 26 minutes – feature codes M129T1559_1 and M183-T1556) could
397 not be assigned to certain chemical structures. The PCA of the ESI- LC-MS dataset resulted in a score plot with
398 PC1 representing 48.9% and PC2 18.8% of the variability. In agreement with the first PCA, procyanidin B2
399 (M577T388_2) and epicatechin (M289T422_3) were identified as markers of dark and ruby chocolate, while also
400 feature M191T64 was found as marker for these two types of chocolate, but its identity was not determined.

401 The same is true for feature M305T387_2, which is more abundant in dark chocolate. In addition to the possible
402 marker compounds for ruby chocolate detected in ESI+ mode, two additional possible features could be found
403 in ESI- mode, namely M343T40 and M377T43, whose identities could not be defined. The A-type
404 proanthocyanidins which were identified in the first part of this work in ruby, but not, or to a much lesser extent
405 in dark chocolate, did not appear as characteristic constituents in this PCA. Possibly, the four proposed marker
406 features, as deduced from the PCA, are even more characteristic for ruby chocolate than the A-type
407 proanthocyanidins. However, further research is required in order to determine their structures and to confirm
408 whether they are indeed useful markers for ruby chocolate. For white chocolate, three markers were found,
409 namely M387T58, M387T42_2, as well as M341T43_3. Of these, the latter most probably corresponds to a
410 dihexose, as deduced from its accurate *m/z* value and retention time.

411

412 **4. Conclusion**

413 Ruby chocolate was introduced as the fourth type of chocolate in 2017, but until now, not much is known about
414 its exact phytochemical composition. Therefore, in this work, ruby chocolate was extracted and analyzed with a
415 UPLC-hybrid quadrupole-TOF MS/MS method, together with a dark chocolate extract, and a feature-based
416 molecular network was created with the obtained data, with the aim of identifying a wide range of compounds.
417 Interpretation of this network and comparison of the raw data with literature, led to the tentative identification
418 of 41 compounds, of which 30 were represented in ruby, and 31 in dark chocolate. These compounds mainly
419 include proanthocyanidins and flavonoids, but also *N*-containing compounds and organic acids. The data show
420 that a wide range of phytochemicals, present in the “conventional” dark and milk chocolates are present in ruby
421 chocolate too. Most interesting is the finding that proanthocyanidins of the A-type appear to be characteristic
422 for ruby chocolate, while B-type proanthocyanidins were found mainly in the dark chocolate. In addition, dark,
423 milk, ruby and white chocolate were analyzed by UPLC-Orbitrap-MS and the results of the latter analysis
424 corroborated the results of the former analysis to a high extent.

In a second part of the work, a UPLC-QTOF-MS analysis of white, ruby, milk and dark chocolate was carried out, simultaneous with a range of concentrations of 51 reference compounds. This allowed for the quantification of 19 compounds in total, of which 13 could be quantified in ruby chocolate. Moreover, 13 out of these 19 compounds had not been identified in the first part of the work. Thus, overall, 54 compounds were identified in the four studied chocolates and 43 in the ruby chocolate in particular.

The levels of quantified polyphenols, methylxanthines, amino acids and biogenic amines were highest in the dark chocolate, which correlates to its higher content of cocoa powder (≥ 70.5%). Milk chocolate (≥40.7% cocoa powder) was richer in methylxanthines, amino acids and biogenic amines than ruby chocolate (≥47.3% cocoa powder) too. In contrast, ruby chocolate contained higher levels of epicatechin and procyanidin B2, compared to milk chocolate, which may be the result of a shorter, or no fermentation of the cocoa beans starting material used for the production of ruby chocolate. Moreover, the ruby chocolate was the only chocolate in which caffeic acid could be quantified. Finally, PCA showed that the different chocolates can be clearly distinguished and revealed several possible markers, but given the limited sample set, further research is required in order to confirm these findings. To the best of our knowledge this is the first detailed report on the phytochemical composition of ruby chocolate.

CRediT authorship contribution statement

Emmy Tuenter: Conceptualization, Investigation, Formal analysis, Visualization, Roles/Writing - original draft. **Maria E. Sakavitsi**: Investigation, Roles/Writing - original draft. **Andrés Rivera-Mondragón**: Formal analysis, Writing - review & editing. **Nina Hermans:** Writing - review & editing**, Kenn Foubert**: Conceptualization, Writing - review & editing. **Maria Halabalaki**: Conceptualization, Supervision, Writing - review & editing. **Luc Pieters**: Conceptualization, Supervision, Writing - review & editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Appendix a. Supplementary data

Supplementary data to this article can be found online at…

References

Aprotosoaie, A. C., Luca, S. V., & Miron, A. (2016). Flavor Chemistry of Cocoa and Cocoa Products-An Overview. *Comprehensive Reviews in Food Science and Food Safety*, *15*(1), 73–91. https://doi.org/10.1111/1541-4337.12180

Cádiz-Gurrea, M. L., Lozano-Sanchez, J., Contreras-Gámez, M., Legeai-Mallet, L., Fernández-Arroyo, S., & Segura-Carretero, A. (2014). Isolation, comprehensive characterization and antioxidant activities of Theobroma cacao extract. *Journal of Functional Foods*, *10*, 485–498. https://doi.org/10.1016/j.jff.2014.07.016

Callebaut. (n.d.). Finest Belgian Ruby Chocolate - RB1. Retrieved January 17, 2020, from https://www.callebaut.com/en-GB/chocolate-cocoa-nuts/chr-r35rb1/ruby

Chambers, M. C., MacLean, B., Burke, R., Amodei, D., Ruderman, D. L., Neumann, S., … Mallick, P. (2012). A cross-platform toolkit for mass spectrometry and proteomics. *Nature Biotechnology*, *30*(10), 918–920. https://doi.org/10.1038/nbt.2377

Chong, J., Wishart, D. S., & Xia, J. (2019). Using MetaboAnalyst 4.0 for Comprehensive and Integrative Metabolomics Data Analysis. *Current Protocols in Bioinformatics*, *68*(1), 1–128. https://doi.org/10.1002/cpbi.86

D'Souza, R. N., Grimbs, S., Behrends, B., Bernaert, H., Ullrich, M. S., & Kuhnert, N. (2017). Origin-based polyphenolic fingerprinting of Theobroma cacao in unfermented and fermented beans. *Food Research International*, *99*, 550–559. https://doi.org/10.1016/j.foodres.2017.06.007

Dumarche, A., Troplin, P., Bernaert, H., Lechevalier, P., Beerens, H., & Landuyt, A. (2012). *Process for producing cocoa-derived material (European patent)*.

Gatto, L., & Lilley, K. S. (2012). Msnbase-an R/Bioconductor package for isobaric tagged mass spectrometry data visualization, processing and quantitation. *Bioinformatics*, *28*(2), 288–289.

https://doi.org/10.1093/bioinformatics/btr645

Giacomoni, F., Le Corguillé, G., Monsoor, M., Landi, M., Pericard, P., Pétéra, M., … Caron, C. (2015). Workflow4Metabolomics: A collaborative research infrastructure for computational metabolomics. *Bioinformatics*, *31*(9), 1493–1495. https://doi.org/10.1093/bioinformatics/btu813

Gruz, J., Novák, O., & Strnad, M. (2008). Rapid analysis of phenolic acids in beverages by UPLC-MS/MS. *Food Chemistry*, *111*(3), 789–794. https://doi.org/10.1016/j.foodchem.2008.05.014

Guan, Z., Li, S., Lin, Z., Yang, R., Zhao, Y., Liu, J., … Chen, A. (2014). Identification and quantitation of phenolic compounds from the seed and pomace of perilla frutescens using HPLC/PDA and HPLC-ESI/QTOF/MS/MS. *Phytochemical Analysis*, *25*(6), 508–513. https://doi.org/10.1002/pca.2521

Hatano, T., Miyatake, H., Natsume, M., Osakabe, N., Takizawa, T., Ito, H., & Yoshida, T. (2002). Proanthocyanidin glycosides and related polyphenols from cacao liquor and their antioxidant effects. *Phytochemistry*, *59*(7), 749–758. https://doi.org/10.1016/S0031-9422(02)00051-1

Hellström, J., Sinkkonen, J., Karonen, M., & Mattila, P. (2007). Isolation and structure elucidation of procyanidin oligomers from Saskatoon berries (Amelanchier alnifolia). *Journal of Agricultural and Food Chemistry*, *55*(1), 157–164. https://doi.org/10.1021/jf062441t

Hurst, W. J., Krake, S. H., Bergmeier, S. C., Payne, M. J., Miller, K. B., & Stuart, D. A. (2011). Impact of fermentation, drying, roasting and Dutch processing on flavan-3-ol stereochemistry in cacao beans and cocoa ingredients. *Chemistry Central Journal*, *5*(1), 53. https://doi.org/10.1186/1752-153X-5-53

Kothe, L., Zimmermann, B. F., & Galensa, R. (2013). Temperature influences epimerization and composition of flavanol monomers, dimers and trimers during cocoa bean roasting. *Food Chemistry*, *141*(4), 3656–3663. https://doi.org/10.1016/j.foodchem.2013.06.049

Kuhl, C., Tautenhahn, R., Böttcher, C., Larson, T. R., & Neumann, S. (2012). CAMERA: An integrated strategy for compound spectra extraction and annotation of liquid chromatography/mass spectrometry data sets.

Analytical Chemistry, *84*(1), 283–289. https://doi.org/10.1021/ac202450g

Kumar, S., Singh, A., & Kumar, B. (2017). Identification and characterization of phenolics and terpenoids from ethanolic extracts of Phyllanthus species by HPLC-ESI-QTOF-MS/MS. *Journal of Pharmaceutical Analysis*, *7*(4), 214–222. https://doi.org/10.1016/j.jpha.2017.01.005

Le Gresley, A., & Peron, J. M. R. (2019). A semi-automatic approach to the characterisation of dark chocolate by Nuclear Magnetic Resonance and multivariate analysis. *Food Chemistry*, *275*(September 2018), 385–389. https://doi.org/10.1016/j.foodchem.2018.09.089

Mani, D. R., Abbatiello, S. E., & Carr, S. A. (2012). Statistical characterization of multiple-reaction monitoring mass spectrometry (MRM-MS) assays for quantitative proteomics. *BMC Bioinformatics*, *13*, S9. https://doi.org/10.1186/1471-2105-13-S16-S9

Nazaruddin, R., Seng, L. K., Hassan, O., & Said, M. (2006). Effect of pulp preconditioning on the content of polyphenols in cocoa beans (Theobroma Cacao) during fermentation. *Industrial Crops and Products*. https://doi.org/10.1016/j.indcrop.2006.03.013

Patras, M. A., Milev, B. P., Vrancken, G., & Kuhnert, N. (2014). Identification of novel cocoa flavonoids from raw fermented cocoa beans by HPLC-MSn. *Food Research International*, *63*, 353–359. https://doi.org/10.1016/j.foodres.2014.05.031

Rue, E. A., Rush, M. D., & van Breemen, R. B. (2018). Procyanidins: a comprehensive review encompassing structure elucidation via mass spectrometry. *Phytochemistry Reviews*, *17*(1), 1–16. https://doi.org/10.1007/s11101-017-9507-3

Šeremet, D., Mandura, A., Cebin, A. V., Oskomić, M., Champion, E., Martinić, A., & Komes, D. (2019). Ruby chocolate - bioactive potential and sensory quality characteristics compared with dark, milk and white chocolate. *Food in Health and Disease: Scientific-Professional Journal of Nutrition and Dietetics*, *8*, 89–96.

Shannon, P., Markiel, A., Ozier, O., Baliga, N. S., Wang, J. T., Ramage, D., … Ideker, T. (2003). Cytoscape: A

Software Environment for Integrated Models. *Genome Research*, *13*, 2498–2504. https://doi.org/10.1101/gr.1239303.metabolite

Smith, C. A., Want, E. J., O'Maille, G., Abagyan, R., & Siuzdak, G. (2006). XCMS: Processing mass spectrometry data for metabolite profiling using nonlinear peak alignment, matching, and identification. *Analytical Chemistry*, *78*(3), 779–787. https://doi.org/10.1021/ac051437y

Sui, Y., Zheng, Y., Li, X., Li, S., Xie, B., & Sun, Z. (2016). Characterization and preparation of oligomeric procyanidins from Litchi chinensis pericarp. *Fitoterapia*, *112*, 168–174. https://doi.org/10.1016/j.fitote.2016.06.001

Tuenter, E., Delbaere, C., De Winne, A., Bijttebier, S., Custers, D., Foubert, K., … Pieters, L. (2020). Non-volatile and volatile composition of West African bulk and Ecuadorian fine-flavor cocoa liquor and chocolate. *Food Research International*, *130*, 108943. https://doi.org/10.1016/j.foodres.2019.108943

Tuenter, E., Foubert, K., & Pieters, L. (2018). Mood Components in Cocoa and Chocolate: The Mood Pyramid. *Planta Medica*, *84*(12–13), 839–844. https://doi.org/10.1055/a-0588-5534

Wang, M., Carver, J. J., Phelan, V. V., Sanchez, L. M., Garg, N., Peng, Y., … Bandeira, N. (2016). Sharing and community curation of mass spectrometry data with Global Natural Products Social Molecular Networking. *Nature Biotechnology*, *34*(8), 828–837. https://doi.org/10.1038/nbt.3597

Wollgast, J., & Anklam, E. (2000). Review on polyphenols in Theobroma cacao: Changes in composition during the manufacture of chocolate and methodology for identification and quantification. *Food Research International*. https://doi.org/10.1016/S0963-9969(00)00068-5

Zhang, Y., Xiao, G., Sun, L., Wang, Y., Wang, Y., & Wang, Y. (2013). A new flavan-3-ol lactone and other constituents from Euonymus alatus with inhibitory activities on a-glucosidase and differentiation of 3T3-L1 cells. *Natural Product Research*, *27*(17), 1513–1520. https://doi.org/10.1080/14786419.2012.725400

Figure captions

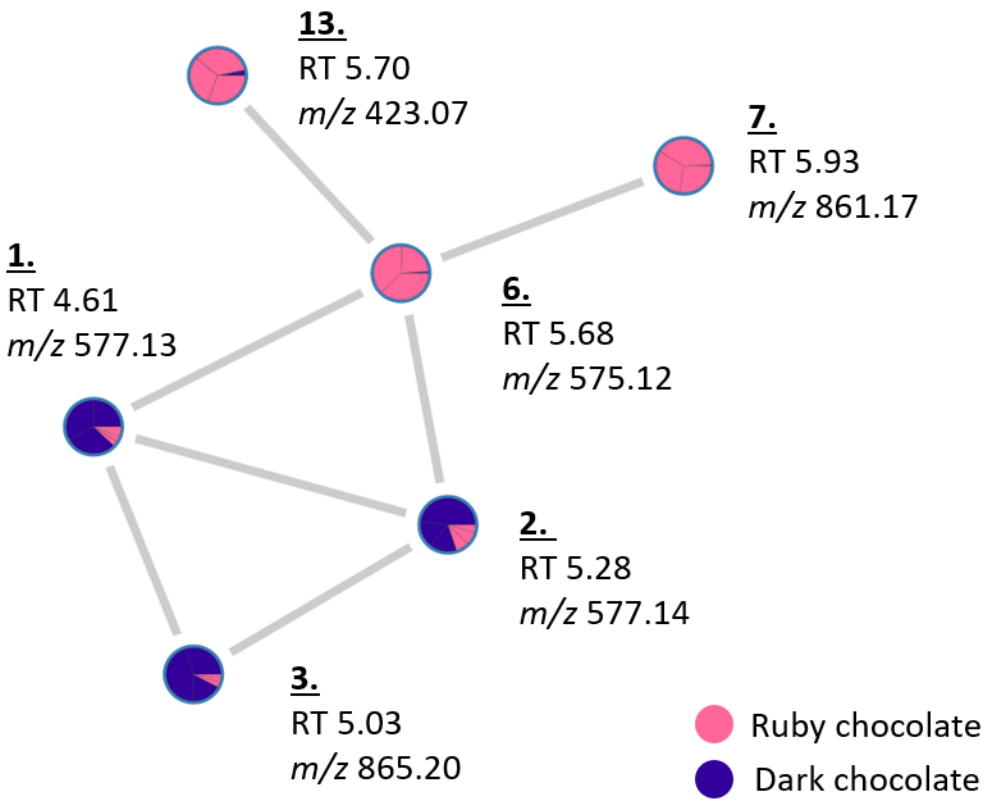
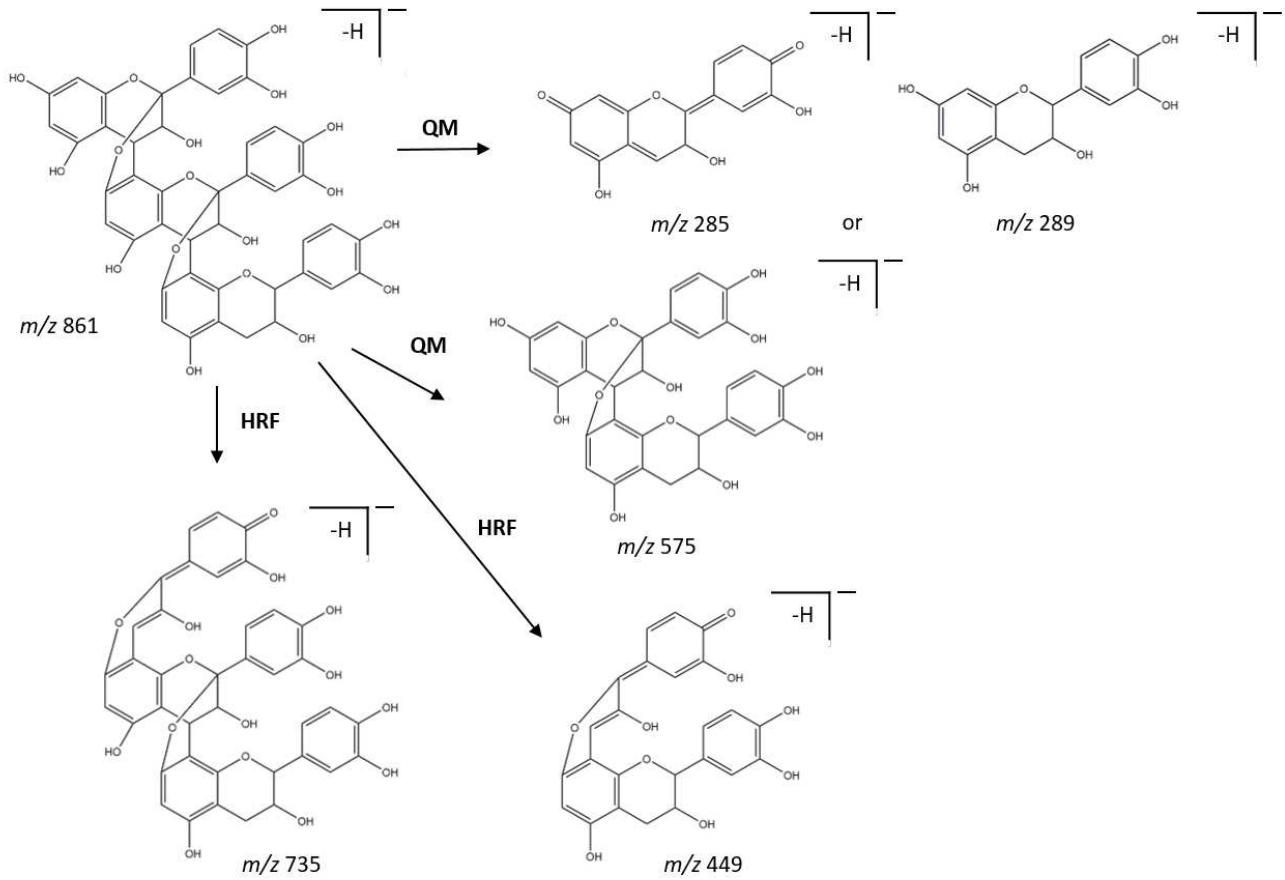


Fig. 1 Cluster of tentatively identified proanthocyanidin dimers and trimers, obtained with feature-based molecular networking using GNPS and visualization with Cytoscape. Compound numbers match with numbers in Table 1. Pie charts represent relative abundance of each compound in the ruby and dark chocolates (each analyzed in triplicate). RT: retention time (minutes).



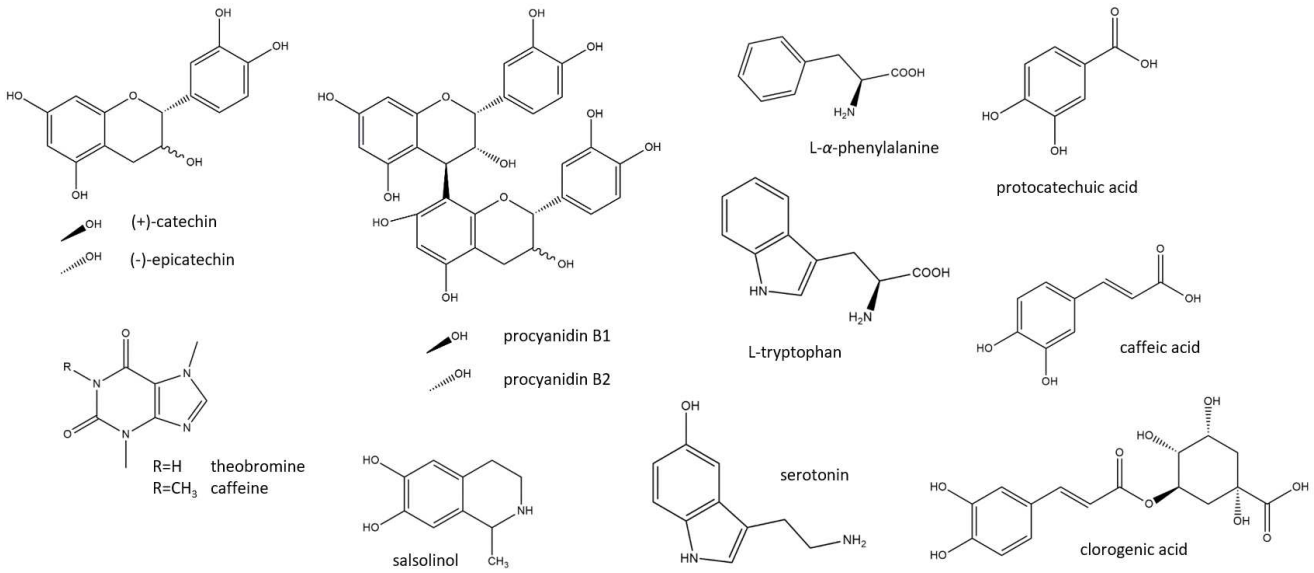
559

560 **Fig. 2** Proposed ESI-MS fragmentation pathway of an A-type procyanidin trimer, based in part on

561 fragmentation pathways reported for A-type procyanidin dimers (Rue et al., 2018; Sui et al., 2016). QM:

562 quinone methide reaction, HRF: Heterocyclic Ring Fission.

563

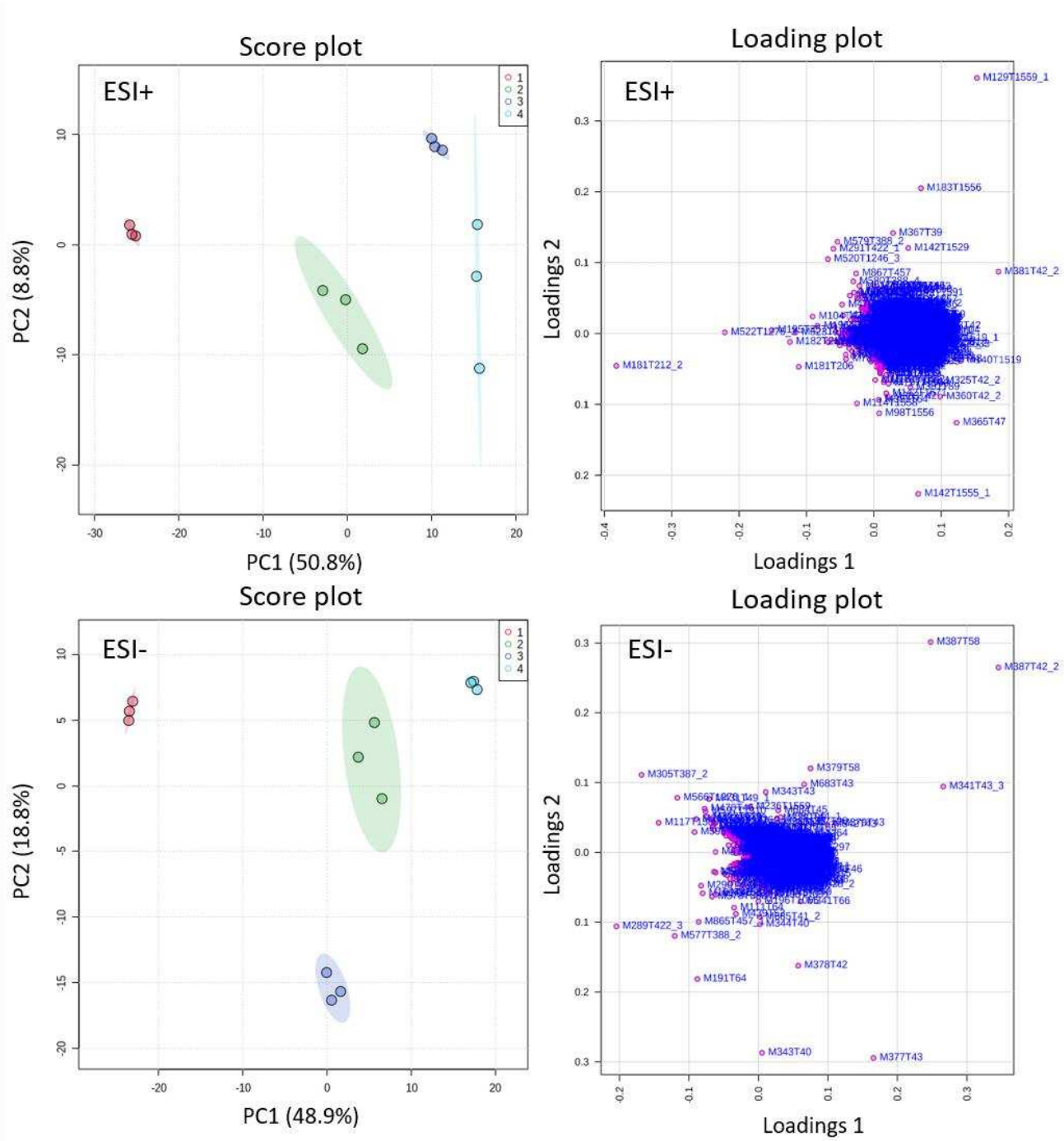


564

565 **Fig. 3** Compounds identified and quantified in ruby chocolate by means of UPLC-QTOF-MS analysis and

566 comparison to reference compounds.

567



568

569 **Fig. 4** Score plots (left) and loading plots (right) of PCA of UPLC-QTOF-MS analysis of the four chocolates in ESI+

570 (top) and ESI- mode (bottom). Score plots: Red dots (1): dark chocolate, green dots (2): milk chocolate, dark

571 blue dots (3): ruby chocolate, cyan dots (4): white chocolate. Loading plots: features are encoded according to

572 their *m/z* value (M) and retention time (T) in seconds.