Simulated gastrointestinal biotransformation of chlorogenic acid, flavonoids, flavonolignans and triterpenoid saponins in Cecropia obtusifolia leaf extract

Simulated gastrointestinal biotransformation of chlorogenic acid, flavonoids, flavonolignans and triterpenoid saponins in *Cecropia obtusifolia* leaf extract

Andrés Rivera-Mondragón [1,2], Laura Peeters [1], Anastasia Van der Auwera [1], Annelies Breynaert [1], Catherina Caballero-George [2], Luc Pieters [1], Nina Hermans [1] and Kenn Foubert [1]

Affiliation

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

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

Correspondence

Dr. Andrés Rivera-Mondragón, 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: arivera.qclabherbals@gmail.com. Phone number: +507 6578-0602

Abstract

It is well known that biotransformation processes in the human body are crucial to form potentially bioactive metabolites from particular classes of natural products. However, little research has been conducted concerning the bioavailability of polyphenols, especially in the colon. The gastrointestinal stability and colonic biotransformation of the crude extract of the leaves of *C. obtusifolia*, rich in flavone *C*-glycosides, was investigated under *in vitro* conditions, and the processing and interpretation of results were facilitated by using an automated machine-learning model. This investigation revealed that flavone *C*-glycosides and flavonolignans from *C. obtusifolia* were stable throughout their passage in the simulated gastrointestinal tract including the colon phase. On the other hand, the colon bacteria extensively metabolized chlorogenic acid, flavonol and triterpenoid *O*-glycosides. This investigation revealed that the colonic microbiota has an important role in the biotransformation of some chemical constituents of this extract.

Key words

Cecropia obtusifolia, Urticaceae, UPLC-DAD-QTOF, machine learning model (MLM), gastrointestinal model, gut microbiota

Abbreviations

CfA:	Caffeic acid
GIM:	gastrointestinal model
C1:	3-(3,4-dihydroxyphenyl) propionic acid
C2:	3-(3-hydroxyphenyl) propionic acid
C3:	3-(4-hydroxyphenyl) propionic acid
R1:	quercetin
T1:	2α,3α,19α,23-tetrahydroxy-12-ursen-28-oic acid
T2:	buergericic acid
T3:	euscaphic acid and tormentic acid.
MB:	method blank sample
MLM:	machine learning model
NC:	negative control sample
PC-ChA:	positive control of chlorogenic acid
PC-Q:	positive control of quercetin
QAc:	quinic acid
WCB:	Wilkins Chalgren Broth
WCA:	Wilkins Chalgren Agar

Introduction

Polyphenols are a family of approximately 8000 naturally occurring compounds commonly found in plants, and they can be classified, according to their chemical structure, into two main classes: flavonoids (e.g. flavonols, flavones, flavanols, flavanones, chalcones, anthocyanins, isoflavones and neoflavonoids) and non-flavonoids (e.g. phenolic acids, lignans, coumarins, stilbenes, xanthones, and tannins) [1,2]. Flavonoids are frequently described as the most abundant constituents of plants, sharing the same basic fifteen-carbon skeleton of diphenyl propane derivatives [3].

Flavonoids in general have been studied for more than 70 years in in *vivo* and *in vitro* experiments [4]. They have been associated with a broad spectrum of beneficial effects including anti-oxidant [5], anti-inflammatory [6], anti-carcinogenic [7], anti-viral [3], anti-bacterial [3], and hepatoprotective activities [8], as well combating cardiovascular diseases [9]. Furthermore, polyphenols have been reported for their capacity to modulate/inhibit several important enzyme functions, such as alpha-glucosidases (AG), cyclo-oxygenases (COX-1 and COX-2), xanthine oxidase (XO), acetylcholinesterase (AChE), lipoxygenases (LOX), phosphoinositide 3-kinase and cytochrome P450 [3,4,10].

Cecropia Loefl.(Urticaceae) is a native genus from Mexico, Central and South America that comprises 61 species [11]. It is popularly called 'yarumo', 'guarumo', 'guarumbo', 'embauba', 'ambay' and 'trumped tree', and its leaves are traditionally consumed as a water infusion in Latin America for its broad range of therapeutic properties [11]. *C. obtusifolia* Bertol and *C. peltata* L. (Mexico, Central America, the Caribbean region, Colombia and Ecuador) are reported for the treatment of type 2 diabetes and hypertention [12,13]. On the other hand, *C. glaziovii* Snethl., *C. pachystachya* Trécul. and *C. hololeuca* Miq. (South America, mainly in Brazil and Argentina) are frequently used to treat respiratory conditions, such as asthma; as well as anti-inflammatory, anti-hypertensive, and anti-malarial agents [14]. The broad range of therapeutic applications associated with *Cecropia* has been correlated with their content of chlorogenic acid, *C*-glycosyl flavones (orientin, isoorientin, vitexin and isovitexin), flavonols and their *O*-glycosides (quercetin, rutin and isoquercitrin) and triterpenoids (ursolic, tormentic, euscaphic, isoarjulonic and pomolic acid). Catechin, proanthocyanidins and steroids have also been reported [11].

In recent years, there has been an increasing interest in research on biotransformation of natural products by intestinal microbiota [15]. It is well known that gut microflora plays an important role in the bioavailability of chemical compounds that are orally administered [16]. In this sense, phytochemicals that are not absorbed in the small intestine and are able to reach the colon might be biotransformed by colonic bacteria before being absorbed or excreted without further conversion.

The biotransformation of flavone *C*-glycosides, the main compounds in *Cecropia* extract, has been little investigated in comparison with their *O*-glycoside analogues, which are more common in nature and more easily hydrolyzed than *C*-glycosides [17]. For that reason, it is of great importance to conduct investigations on the biotransformation of flavonoid glycosides in the human body in order to provide further understanding of their beneficial impact on health [17]. As a first step toward the elucidation of the behavior of the crude extract of the leaves of *C. obtusifolia*, rich in flavone *C*-glycosides [14], the stability after oral ingestion in gastrointestinal conditions and colonic biotransformation was investigated. This is the first time that the biotransformation of the extract of *C. obtusifolia*, an important source of compounds known for their health-promoting effects, has been addressed under gastrointestinal conditions.

Results and discussion

Biotransformation experiments were carried out for chlorogenic acid (10 mg) and quercetin (150 mg) (pure standards) and on a crude extract of *Cecropia obtusifolia* (300 mg) in a previously developed and validated *in vitro* gastrointestinal model (GIM) with colon phase [18]. Additionally, a NC [also containing 300 mg of the crude extract and culture broth (WCB) but lacking fecal suspension] and a MB containing fecal suspension without extract were included. An overview of the experiments is depicted in Fig. 1.

Changes in the initial chemical composition of the extract, and formation of metabolites were investigated as a function of time. Chlorogenic acid, flavonoids, flavonolignans and triterpenoid saponins were detected in the sample, and their identities had previously been assigned [14]. The main compounds present in the crude extract are summarized in Table 1. Concentrations of chlorogenic acid, flavonoids and flavonolignans in the crude leaf extract of *Cecropia obtusifolia* are depicted in Table 1S.

Using a suspect screening approach, a list of possible biotransformation products described in literature was obtained (including molecular formula and corresponding exact mass). Identification was carried out based on accurate mass obtained in the full MS mode, fragmentation pattern (MS/MS) of the product ions and comparison with analytical standards if available and/or fully identified compounds isolated from *C. obtusifolia*. Acceptance requirements for identification of biotransformation products were set as in similar studies: mass tolerance of ± 5 ppm between theoretical and measured monoisotopic mass; product ions may not exceed the maximal mass variation of ± 10 ppm; compounds detected in the samples that were also present in the MB and NC were ignored (allowing the interpretation of only those specific metabolites produced during the GIM and by action of microbiota), and the detected metabolites were present in at least two out of the three replicate samples [19]. Biotransformation products are shown in Table 2.

In this investigation, a metabolomics workflow combining two statistical models was used in order to find different biotransformation profiles over time. The first model was performed via the EDGE method [20,21]. This approach provides two *p*-values for every feature, one for sample vs MB and one for sample vs NC, which is not easy to interpret and challenging to associate. The second statistical model consisted on a machine learning model (MLM) in order to obtain a single score for each feature (see Materials and Methods section). All metabolites described here are ranked in the top 650 out of 10,534 features in negative ion mode and describe a tinderesting score over 0,890. This demonstrates that the use of the MLM is adequate for predicting interesting features (see Table 2).

Similar to previous reports in literature, it was observed that biotransformation of polyphenols described in this paper occurred mainly in the colon phase [18,22]. With regard to chlorogenic acid (**1**) in the samples, its biotransformation was interpreted based on a comparison with the positive control sample containing the pure chlorogenic acid (reference standard). It was observed that chlorogenic acid present in test samples was completely transformed during the colonic phase, also when it was administrated as a pure compound. Even though the crude extract (1.0 mg) had a lower concentration of chlorogenic acid than the positive control (10.0 mg), the metabolism rate of compound 1 was a lot slower compared to the single-compound incubation experiment (See Fig. 2 and 3). The complete digestion for chlorogenic acid in the colon phase was 12 h in samples, while it took 2 h in the positive control. A possible explanation for these results may be that the presence

of diverse chemicals in the crude extract (flavonoids, flavonolignans and triterpenoid saponins) could negatively interfere with the biotransformation rate of chlorogenic acid.

Furthermore, three molecular ions at *m/z* 181.0503 $[M - H]^-$, 165.0553 $[M - H]^-$ and 165.0554 $[M - H]^-$ were detected in both samples and positive control. The first two metabolites were unequivocally identified as 3-(3,4-dihydroxyphenyl) propionic acid (**C1**) and 3-(3-hydroxyphenyl) propionic (**C2**) acid by comparison with authentic reference standards, while the latter was tentatively characterized as 3-(4-hydroxyphenyl) propionic acid (**C3**) (See Fig. 4 and Table 2). **C1**, **C2** and **C3** were formed from the catabolism of caffeic acid, a product of the de-esterification of chlorogenic acid. **C1** was produced by the hydrogenation (reduction) of caffeic acid, while **C2** and **C3** was formed by the subsequent de-hydroxylation of **C1**. These findings are in agreement with those reported previously in the literature [18,22]. **C1**, **C2** and **C3** were rated as interesting by the MLM with a ranking score of 529, 107 and 418 out of 10,534 features, respectively. The intensity of these compounds increased after 2 h with a maximum relative abundance after 4-6 h for **C1**, and 18-22 h for **C2** and **C3**. Subsequently, these metabolites gradually disappeared after 30-48 h, except for **C2**, which remained stable and accumulated after 48 h. A targeted search for other potential metabolites was performed. Contrary to expectations, metabolites such as dihydro-chlorogenic acid, caffeic acid, quinic acid, 3,4-dihydroxyphenyl acetic acid, 3-hydroxyphenyl acetic acid, 4-hydroxyphenyl acetic acid, 4-hydroxybenzoic acid, 3-hydroxyphenyl ethanol and hydroxyl benzaldehyde were not detected [18,22]. It seems possible that these results are due to a fast conversion of chlorogenic acid to the final metabolites detected in this investigation.

In Fig. 5 and supplementary Fig. 1S the degradation profile of rutin (**15**), a flavonol *O*-diglycoside, in the GIM is shown. As expected, the peak intensity of **15** decreased rapidly during the first 2 h of incubation in the colon phase due to the hydrolysis of rutinoside. Simultaneously, the intensity of a deprotonated molecule at *m/z* 301.0348 $[M - H]^-$ (**R1**) started to increase after 2 h with a maximum relative abundance until 6 h, followed by a decrease until were undetectable after 22 h. **R1** was identified as quercetin by comparison with an authentic standard. In the same way as the interpretation of chlorogenic acid biotransformation, quercetin metabolism was investigated by comparison with a positive control containing a pure reference standard. Interestingly, the presence of **C3** was detected when quercetin was incubated as a single compound (positive control of quercetin). It is proposed that **C3** 3-(4-hydroxyphenyl propionic acid) was formed by C-ring fission

of quercetin and subsequent dehydroxylation of the B ring. Additional potential metabolites such as taxifolin, protocatechuic acid, 3,4-dihydroxyphenyl acetic acid, 3-hydroxyphenyl acetic acid, phloroglucinol were not detected [23–27].

On the other hand, flavone di-*C*,*O*-glycosides such as luteolin *C*-hexoside-*O*-hexosides (**2, 4**), isoorientin-2″-*O*-xyloside (**3**), isoorientin-4″-*O*-xyloside (**5**), isoorientin-2″-*O*-rhamnoside (**7**), isovitexin-2″-*O*-glucoside (**8**), isovitexin-2″-*O*-xyloside (**11**), diosmetin *C*-hexoside-*O*-pentoside (**12**), isovitexin 2″-*O*-rhamnoside (**13**) and diosmetin *C*-hexoside-*O*-deoxyhexoside (**16**) did not show significant differences between samples and NC. This means that the intensities of the peaks of these compounds from the samples containing colonic bacteria were similar to those with only the crude extract and no bacteria (see Fig. 6 and supplementary Fig. 2S). This indicates that di-*C,O*-glycosides of apigenin and luteolin are stable during the GIM with colon phase under the experimental conditions tested in this investigation. However, it was observed that the intensity of only one peak, assigned to a flavone di-*C*-*O*-glycoside, tentatively identified as apigenin *C*-hexoside-*O*-hexoside (**9**), decreased during the colonic phase. This finding indicates that this compound is particularly prone to biotransformation. A possible explanation for this might be that the *O*-glycosidic bond is probably linked to the 5-OH, 7-OH or 4´-OH position of the aglycone. It can therefore be assumed that a different sugar placement in this molecule could be more susceptible to biotransformation in comparison to its structural isomer isovitexin-2″-*O*-glucoside (**8**) (Fig. 6).

With regard to flavone mono-*C*-glycosides, it was observed that the peak intensity of isoorientin (**6**) and vitexin (**10**) did not decrease during the whole experiment and this was confirmed by the absence of *C*-deglycosylated aglycones (see Fig. 6 and supplementary Fig. 3S). In fact, a similar behavior described for most flavone di-*C,O*-glycosides in this research was observed. On the other hand, the intensity of the isovitexin peak (**14**) increased during the experiment (Supplementary Fig. 3S). This confirmed not only the stability of this compound for biotransformation, but in addition, it also supports the conversion of compound **9** into its mono-*C*-glycoside. A targeted search for potential metabolites such as apigenin, luteolin, eriodictyol was carried out [17,28]. However, none of these metabolites was detected. This result is explained by the fact that luteolin and apigenin *C*-glycosides were particularly stable in this GIM.

With respect to the pharmacokinetics of flavonoids, there is an important amount of literature on flavonoid *O*-glycosides and their effects on health. These compounds are reported to be deglycosylated to their aglycones by *β*-glucosidases, in part in the small intestine, and further by colon bacteria [17]. However, it seemed that *C*-glycosides showed a different pharmacokinetic behavior compared with their *O*-glycosidic analogues [17]. More recently, contradictory findings have been published about the metabolic fate of flavonoid di-*C,O*-glycosides and *C*-glycosides. Some research work pointed out that in spite of the low bioavailability of flavonoid *C*-multiglycosides (such as vitexin 4″-*O*-glucoside, and vitexin 2″-*O*-rhamnoside), deglycosylation is not a prerequisite for their oral absorption in rats [29–33]. However, excretion studies in male Wistar rats revealed a high recovery of unchanged vitexin-2″-*O*-rhamnoside and vitexin-4″-*O*-glucoside in feces (64 and 88%, respectively). This may indicate the relative high stability and non-absorption of these flavonoid di-*C,O*-glycosides in the gastrointestinal tract.

In contrast to our findings, other studies have revealed that intestinal bacteria can cleave the *C*-glucosyl bond in different flavonoid *C*-monoglycosides such as puerarin, isoorientin, orientin, vitexin and isovitexin [17,34]. Furthermore, an investigation on the biotransformation of isoorientin by human intestinal bacteria revealed that this compound was completely biotransformed to 6-*C*-glucosyl eriodictyol (by a reduction reaction), which was subsequently converted to eriodictyol (by cleaving the *C*-glycosyl bond), and finally transformed into 3,4-dihydroxyphenylpropionic acid and phloroglucinol via C-ring fission [34]. Recent literature focused on the ability of certain bacterial species to biotransform these flavonoids, since two isolated bacteria from human fecal samples (*Lactococcus* sp. and an *Enterococcus* sp.) were able to completely convert puerarin (an isoflavone *C*-glycoside) to daidzein (aglycone) within 16 h. Interestingly, vitexin was not converted by any of the isolated bacteria at all [35]. On the other hand, *Enterococcus faecalis* (a bacterial strain, named W12-1) isolated from human fecal sources was able to deglycosylate orientin, vitexin, and isovitexin. However, it did not transform isoorientin and puerarin to their corresponding aglycones [36].

This inconsistency may be due to differences in bacterial composition of the cultures used. Only a fraction of human intestinal bacteria responsable for the conversion of flavonoids has been identified so far, most of which have been able to cleave the *O*-glycosidic bond of flavonoids. On

the other hand, cleavage of the *C*-glycosides seems to be more specific for some bacterial strains, which may result in individually different host metabotypes [28].

Flavonolignans seemed to be stable during most of the experiment (see Fig. 4S). Only after 39 h of the colon phase, the relative intensities of **17** and **18** decreased. In reviewing the literature, no data were found on the gastrointestinal stability of compounds related to mururin A or vaccinin A. Research on bioavaibility of flavonolignans has been mostly restricted to a limited number of compounds like silybin. The metabolism of silybin, the main compound of *Silybum marianum* fruit extract (silymarin), has been investigated in several studies. For instance, the effective membrane permeability values of silybin were evaluated both *in vitro* and *in silico* [37]. According to Diukendjieva et al. (2019), silybin is considered to be highly permeable in the gastrointestinal tract, suggesting that this compound can be absorbed unchanged and has a relatively high bioavailability. On the other hand, the approach used in this investigation is similar to that used by other researchers. Valentová et al. studied the biotransformation of silymarin at two concentration levels (200 mg/L and 10 mg/L) in human fecal microbiota. The experiment using the highest concentration showed that silybin was resistant to the metabolic action of intestinal microbiota after 24 h incubation; while at lower concentration, this compound was found to be primarily demethylated with a final amount of 1.5-2.0% of the initial concentration [38]. In the same line, Pferschy-Wenzig et al. investigated that after 24 h of incubation of silymarin (0.21 mg/mL) the concentration of flavonolignans was remarkably decreased. Metabolites were found to be produced by cleavage of their dioxane ring and demethylation [39].

In spite of these recent findings, *in vivo* experiments were found contradictory. Silymarin has poor water solubility and low oral bioavailability (0.95%) with only 20-50% absorption by the gastrointestinal tract [40]. Despite its low bioavability, it was found that 3-8% of silymarin was excreted unchanged in urine after oral administration, suggesting that its conversion in the gastrointestinal tract was not a prerequisite for its absorption. Silymarin undergoes extensive enterohepatic recirculation. It is assumed that 20–40% is recovered from the bile as glucuronide and sulphate conjugates (phase II), whereas the remaining part is excreted in feces (unchanged) [41–43].

Similarly to *ent*-vaccinin A (**17**) and *ent*-mururin A (**18**), which are composed of a flavonoid moiety (catechin) and a phenylpropanoid part (caffeic acid), the structure of silybin consist of taxifolin (a

flavonoid) attached to a second unit consisting of a coniferyl alcohol moiety (phenylpropanoid unit). Their related structures suggest that *ent*-mururin A and *ent*-vaccinin A might go through a similar metabolic pathway to that of silybin. This suggestion implies that a small portion of these flavonolignans could be absorbed before they can reach the colon (the same for silybin), while the rest remains unchanged and is further excreted in feces. This hypothesis is in agreement with the *in vitro* gastrointestinal experiments observed in this investigation.

Triterpenoid saponins (no steroidal structure) contained in the crude extract of *C. obtusifolia* can be classified in two categories: oleanane (**20**) and ursane (**19**, **21** and **22**) type saponins. All these compounds are characterized by a glucose moiety attached to C-28 through an *O*-glycosidic bond. The proposed biotransformation pathway for triterpenoid saponins is depicted in Fig. 7 and supplementary Fig. 5S. A detailed information on the metabolites is summarized in Table 2. The relative abundance of **19**-**22** decreased gradually during the experiment until they were undetectable after 48 h. Among these four saponins, compound 22 which was completely biotransformed after 18 h of the colon phase. Metabolites **T1** (*m/z* 549.3424 $[M - H + HCOOH]^-$), **T2** (m/z 549.3431 $[M - H + HCOOH]^-$) and **T3** (*m/z* 533.3481$[M - H + HCOOH]^-$) were formed by the deglycosylation of saponin glucosides. Based on a comparison of similar retention time patterns to those of **19** and **20**, **T1** and **T2** were tentatively identified as 2*α*,3*α*,19*α*,23-tetrahydroxy-12-ursen-28-oic acid and buergericic acid, respectively. On the other hand, **T3** was tentatively characterized as tormentic acid and euscaphic acid. It is suggested that this pair of isomers co-eluted at the same retention time. Similar to findings reported in literature, the deglycosylation by intestinal microflora in *in vivo* and *in vitro* experiments, was the major metabolic pathway of triterpenoid saponin glucosides before absorption [15,16,20].

In conclusion, the *in vitro* biotransformation of the crude extract of *C. obtusifolia*, relevant for its content of health-promoting compounds, was assessed in a gastrointestinal model with colon phase for the first time. The use of the automated data analysis workflow facilitated the processing and interpretation of results obtained from the biotransformation experiments of this complex herbal extract. This investigation revealed that flavone *C*-glycosides and flavonolignans in this extract were stable throughout their passage in the simulated gastrointestinal tract including the colon phase. On the other hand, the colon bacteria extensively metabolized chlorogenic acid and flavonol *O*-glycosides in the extract, but in a slower rate than when were analyzed individually. Besides,

triterpenoid glucosides were transformed by elimination of sugar moieties to produce aglycones with no further conversions, suggesting that the released aglycone forms might be subsequently absorbed by passive diffusion.

The results of this study support the evidence that two factors play an important role in the biotransformation of natural products: the variation on colonic microbiota and the presence of multiple compounds in an extract. Thus, the identification of the microorganisms involved in the conversion of phytochemicals, the determination of transformation pathways and the understanding of their interaction (modulation and/or inhibition) are important for the determination of the underlying biological mechanisms. We believe that the findings of this research are key to elucidate the pharmacological effects of phytochemicals from *C. obtusifolia* leaf extract.

Materials and Methods

Solvents and reagents

All solvents, including MeOH, ACN, HCl (37%) and glacial acetic acid (all analytical reagent grade), were acquired from Fisher Scientific. FA (98+%, pure, analytical reagent grade) was obtained from Acros Organics. FA and ACN (both ULC/MS grade) for LC-MS analysis were acquired from Biosolve Chimie. Ultrapure water with a resistivity of 18.2 × MΩ × cm at 25 °C, dispensed by a Milli-Q system from Millipore, was used as solvent and for mobile phase preparation.

Pepsin (P-7000, from porcine stomach mucosa, 800-2500 U/mg protein), bile extract porcine (B-8631), pancreatin (76190, from hog pancreas, 149 USP U/mg amylase, 8 USP U/mg lipase and 100 USP U/mg protease), ethylene diamine tetra-acetate, sodium thioglycolate broth, glycerol (99.5%), $NaH_2PO_4 \cdot 2H_2O$, L-cysteine (98%), tween 80 and resazurine sodium salt were purchased from Sigma-Aldrich. Phosphoric acid (H_3PO_4, 85%), $NaHCO_3$, and NaOH were obtained from Merck. WCA and WCB were acquired from Oxoid.

Reference standards

Chlorogenic acid (99%), quercetin (99%), rutin (97%), protocatechuic acid (97%), 3-(3,4-dihydroxyphenyl) propionic acid (98%), and 3-(3-hydroxyphenyl) propionic acid (98%) were obtained from Sigma-Aldrich, while 4-hydroxy-benzoic acid (99%) from Sigma-Aldrich Chemie GmbH.

Plant material

Leaves of *Cecropia obtusifolia* Bertol. were collected in Cerro Azul, Republic of Panama (9°12'33" N, 79°24'49" W). The taxonomical classification was carried out by the botanist Orlando O. Ortiz and deposited at the Herbarium of the University of Panama (voucher specimen Ortiz, O.O. 2519).

Crude extraction preparation from *Cecropia obtusifolia*

The plant material was air-dried in a general protocol oven (HerathermTM, Thermo Scientific) at 40 °C and subsequently grounded using a mill (1.0 mm mesh size, MF 10 Basic, IKA). The dried leaves (0.5 kg) were defatted with n-hexane and consecutively macerated with 70% EtOH (v/v) at room temperature. The extracts were filtered through Whatman No. 1 filter paper. The solvent was reduced using a rotary evaporator under reduced pressure below 40 °C. The resulting reduced filtrate was lyophilized (52.2 g).

Gastrointestinal model with colon phase

An *in vitro* experimental gastrointestinal model with colon phase formerly developed and validated in the research group NatuRA was used to study the human biotransformation of phenolic compounds [18,20].

Preparation of fecal suspension

A fecal suspension was prepared as described by Peeters *et* al. (2019) [20]. Human fecal donors (n = 3) were selected, aged 25 to 58 years, non-smokers, non-vegetarians, without any history of gastrointestinal disease, not treated with antibiotics in the last three months and with normal defecation. A suspension of 10% (w/v) feces was prepared by homogenizing the pooled fecal sample with a sterile phosphate buffer solution (0.1 M, pH 7.0) in a stomacher (Lab-blender 400,

Seward Medical) during 3 min. The phosphate buffer solution consisted of $NaH_2PO_4 \cdot 2H_2O$ (1.03% w/v), sodium thioglycolate broth (3.45% v/v) and glycerol (17%). Before use, the buffer solution was sonicated and autoclaved (1 bar, 121 °C). Particulate food material was removed from the slurry by using a sterile filter bag consisting of a full-page filter (Bagpage® R/25 400 mL, VWR International). The mixed fecal pool was stored at –80 °C.

Prior to use, the fecal suspension was cultivated. To 10% (v/v) pooled fecal suspension in phosphate buffer 90% (v/v) sterile adapted WCB (Wilkins Chalgren Broth, Oxoid) was added in an anaerobic glove box (5% H_2, 5% CO_2 and 90% N_2) (Jacomex Globe Box T3). The composition of the adapted growth medium was: WCB (33 g/L), L-cysteine (0.5 g/L), Tween 80 (2.8 mL/L) and resazurine sodium salt (1.4 mg/L). The medium was autoclaved at 121 °C for 15 min. The bacteria were cultivated in an anaerobic environment for 17 h at 37 °C under continuously stirring. Afterwards, 80% sterile adapted WCB was added and the suspension was incubated for another hour in order to obtain a bacteria suspension of 10^8 CFU/mL. The viable cell concentrations were determined by means of decimal dilution series of the bacteria samples plated onto WCA (Wilkins Chalgren Agar, Oxoid).

Simulation of the stomach, small intestine and colon

Biotransformation experiments were performed for chlorogenic acid (10 mg) and quercetin (150 mg) as positive controls and on a crude ethanolic extract of *Cecropia obtusifolia* in triplicate (300 mg) (samples). A MB (without test compound) containing fecal suspension without extract was included. Additionally, the experiments were carried out for a NC [containing 300 mg of the crude extract and culture broth (WCB) but lacking fecal suspension].

The simulation of the gastric stage was performed in 200 mL plastic jars outside the anaerobic glove box. All samples were mixed with 47 mL of pure water and 3 ml of pepsin (Sigma-Aldrich) solution [5.4% (w/v) in 0.1 M HCl (2000 FIP-U mL−1)]. Subsequently, the pH was adjusted to 2.0 with 6 M HCl. Afterwards, the samples were placed in a shaking water bath during 1 h (120 strokes/min at 37 °C).

The small intestinal phase was mimicked by adding 50 mL of pure water to the gastric digest. An amount of 1 M $NaHCO_3$ was added to the NC to obtain a pH of 7.5. A small dialysis bag containing the same amount of 1 M $NaHCO_3$ was added to the three samples in order to obtain a gradual pH

change to 7.5. The samples were continuously stirred at 37 °C for 1.5 h. After 30 min, 15 mL pancreatin-bile solution was added to each sample and stirring was continued for an additional hour. This solution was prepared dissolving 0.4% (w/v) of pancreatin from porcine pancreas (143,600 FIP-U/L amylase, 32,000 FIP-U/L lipase, 6400 FIP-U/L protease) and 0.8% (w/v) of bile extract porcine in 0.1 M $NaHCO_3$[18] (for amylase = 1 FIP-unit = 4,15 USP-unit; for lipase = 1 FIP-unit = 1 USP-unit; for protease = 1 FIP-unit = 62,5 USP-unit). Pancreatin from porcine pancreas and bile extract porcine were obtained from Sigma-Aldrich.

The large intestinal phase was performed in an anaerobic glove box (5% CO_2, 5% H_2 and 90% N_2) creating an anaerobic environment at 37 °C. WCB lacking fecal suspension (50 mL) was added to the NC, so no colonic bacteria were present to metabolize the compounds. The pH was adjusted to 5.8 with 1 M HCl. The same volume of 1 M HCl was added to the three samples, the MB and positive control sample. A volume of 50 mL of the fecal suspension in WCB (10^8 CFU/mL) was added to the samples, positive control and the MB. All mixtures were continuously stirred for 48 h. An aliquot of each sample was taken at several time points: before biotransformation (t0), after the gastric stage (S), after the small intestinal phase (SI) and at several time points during the colon phase [C2 (2 hours), C4, C6, C10, C14, C18, C22, C24, C30, C39 and C48]. Every sample was diluted with methanol (1:2) and centrifuged for 5 min at 3500 rpm (approximately 1450 g). The supernatant was collected and samples were 10 times diluted with methanol:water (60:40).

UPLC-DAD-QTOF analysis

For accurate mass measurements of samples, a Xevo G2-XS QTof spectrometer (Waters) coupled with an ACQUITY LC system equipped with MassLynx version 4.1 software was used. For analysis, 5 µL of samples were injected on an ACQUITY UPLC® BEH Shield RP18 column (100 mm × 2.10 mm, 1.7 µm, Waters). The mobile phase solvents consisted of H_2O + 0.1% FA (A) and ACN + 0.1% FA (B), and the gradient was set as follows (min/B%): 0.0/2, 1.0/2, 14.0/26, 24/65, 26.0/100, 29.0/100, 31.0/2, 41.0/2. The temperature of the column was kept at 40 °C. The flow rate was 0.4 mL/min. During the first analysis, full scan data were recorded in ESI^+ and ESI^- ionization mode from *m/z* 50 to 2000 and the analyzer was set in sensitivity mode (approximate resolution: 22,000 FWHM). The spray voltage was set either at +1.5 kV and -1.0 kV; cone gas flow and desolvation gas flow at 50.0 L/h and 1000.0 L/h, respectively; and source temperature and desolvation temperature at 120 °C and 550 °C, respectively. Data were also recorded using MS^e in

the positive and negative ionization modes, and a ramp collision energy from 20 until 30 V was applied. Leucine-Encephalin was used as lock mass. UV chromatograms were recorded at 200 and 280 nm.

Data analysis

Data analysis was performed as previously described by Peeters *et al.* (2019) and Beirnaert *et al.* (2019) [20,21]. All analyzes were performed using the RStudio (version 1.2.1335) from the R software version 3.5.1 (R CORE TEAM, 2018). Briefly, LC-MS data were converted to the mzXML format to allow XCMS processing. The XCMS CentWave algorithm was used to convert the raw data into features via peak-picking followed by grouping. EDGE function was used to see if there are any differences between two groups (samples vs the NC and samples vs MB), regarding their biotransformation profiles over time. These differences were expressed as *p*-values.

However, this approach generates two *p*-values for every feature, one for sample vs MB and one for sample vs NC, which is non-trivial to interpret or combine. To avoid this problem, a MLM that emulates the human revision process, was constructed to obtain a single score for each feature. A random forest model was chosen. To train the MLM, labeled data are needed. These labeled data originate from the tinderesting Shiny app [21], which is used by the experts to review the features. Each training sample consisted of 70 features: 14 time points per sample, 3 replicates per sample, 1 NC and 1 MB.

The experts have three options for labeling a feature: interesting, uninteresting, or unknown. These correspond to the tinderesting labels. The unknown label indicates training cases where the expert is uncertain. Features with the unknown label are omitted when constructing the final random forest model. Thus, the final model is built by using the features that have a tinderesting label that is interesting or not. Biotransformation profiles are then classified as interesting when a change in intensity of the signal for the sample is observed, which at the same time is not detected in the NC or the MB. These features serve as the samples for the MLM (every sample to classify corresponds to a metabolic time profile rated by tinderesting) and the machine learning features are the individual elements of those metabolic features. Cross validation (10-fold) is used to obtain an unbiased estimate of the performance of the MLM (performance is evaluated on the hold-out fold).

After the training procedure, the MLM can be used to detect new metabolites fast and unbiased [21].

Supporting information

Concentrations and biotransformation profiles of chlorogenic acid, flavonoids and flavonolignans in the crude leaf extract of *Cecropia obtusifolia* are available as Supporting Information.

Acknowledgements

The authors gratefully acknowledge the scholarship received for Mr. Rivera's Doctoral studies from the National Secretariat for Science, Technology and Innovation (SENACYT) and the Institute for Training and Development of Human Resources (IFARHU) of the Republic of Panama. The authors also thank SENACYT for financial support to Dr. Caballero-George through the incentive program of the National Investigation System (SNI) as well as through grant FID14-116. We would like to thank the Panamanian Ministry of Environment (MiAMBIENTE) and the Panama Canal Authority (ACP) for facilitating the acquisition of plant material (permit # SC-P-4-15).

Conflict of Interest

The authors declare no conflict of interest.

References

[1] Belščak-Cvitanović A, Durgo K, Huđek A, Bačun-Družina V, Komes D. Overview of polyphenols and their properties. In: Galanakis CM (Hrsg.). Polyphenols: Properties, Recovery, and Applications. Cambridge, MA: Woodhead Publishing, 2018: 456

[2] Durazzo A, Caiazzo E, Lucarini M, Cicala C, Izzo AA, Novellino E, Santini A. Polyphenols : A concise overview on the chemistry , occurrence , and human health. Phyther Res 2019; 33: 1–23

[3] Kumar S, Pandey A. Chemistry and biological activities of Flavonoids: An Overview. Sci World J 2013; 2013: 162750

[4] Yáñez JA, Remsberg CM, Takemoto JK, Vega-villa KR, Andrews PK, Sayre CL, Martinez SE, Davies NM. Polyphenols and flavonoids: an overview. In: Davies NM, Yáñez JA (Hrsg.). Flavonoid pharmacokinetics: Methods of Analysis, Preclinical and Clinical Pharmacokinetics, Safety, and Toxicology. New Jersey: John Wiley & Sons, Inc, 2012: 1–69

[5] Pietta PG. Flavonoids as antioxidants. J Nat Prod 2000; 63: 1035–1042

[6] Rotelli AE, Guardia T, Juárez AO, De La Rocha NE, Pelzer LE. Comparative study of flavonoids in experimental models of inflammation. Pharmacol Res 2003; 48: 601–606

[7] Bisol Â, de Campos PS, Lamers ML. Flavonoids as anticancer therapies: A systematic review of clinical trials. Phyther Res 2020; 34: 568–582

[8] Tong J, Yao X, Zeng H, Zhou G, Chen Y, Ma B, Wang Y. Hepatoprotective activity of flavonoids from *Cichorium glandulosum* seeds *in vitro* and *in vivo* carbon tetrachloride-induced hepatotoxicity. J Ethnopharmacol 2015; 174: 355–363

[9] Rees A, Dodd GF, Spencer JPE. The effects of flavonoids on cardiovascular health: A review of human intervention trials and implications for cerebrovascular function. Nutrients 2018; 10: 1852

[10] Panche AN, Diwan AD, Chandra SR. Flavonoids: an overview. J Nutr Sci 2016; 5: 1–15

[11] Rivera-Mondragón A, Ortiz O, Bijttebier S, Vlietinck A, Apers S, Pieters L, Caballero-George C. Selection of chemical markers for the quality control of medicinal plants of the genus *Cecropia*. Pharm Biol 2017; 55: 1500–1512

[12] Nicasio P, Aguilar-Santamaría L, Aranda E, Ortiz S, González M. Hypoglycemic effect and chlorogenic acid content in two *Cecropia species*. Phyther Res 2005; 19: 661–664

[13] Vidrio H, García-Márquez F, Reyes J, Soto RM. Hypotensive activity of *Cecropia obtusifolia*. J Pharm Sci 1982; 71: 475–476

[14] Rivera-Mondragón A, Bijttebier S, Tuenter E, Custers D, Ortiz OO, Pieters L, Caballero-George C, Apers S, Foubert K. Phytochemical characterization and comparative studies of four *Cecropia* species collected in Panama using multivariate data analysis. Sci Rep 2019; 9: 1763

[15] Fu J, Wu H, Wu H, Deng R, Li F. Chemical and metabolic analysis of *Achyranthes bidentate* saponins with intestinal microflora-mediated biotransformation by ultra-performance liquid chromatography-quadrupole time-of-flight mass spectrometry coupled with metabolism platform. J Pharm Biomed Anal 2019; 170: 305–320

[16] Wang HY, Hua HY, Liu XY, Liu JH, Yu BY. *In vitro* biotransformation of red ginseng extract by human intestinal microflora: Metabolites identification and metabolic profile elucidation using LC-Q-TOF/MS. J Pharm Biomed Anal 2014; 98: 296–306

[17] Xiao J, Capanoglu E, Jassbi AR, Miron A. Advance on the flavonoid *C*-glycosides and health benefits. Crit Rev Food Sci Nutr 2015; 56: S29–S45

[18] Breynaert A, Bosscher D, Kahnt A, Claeys M, Cos P, Pieters L, Hermans N. Development and Validation of an in vitro Experimental GastroIntestinal Dialysis Model with Colon Phase to Study the Availability and Colonic Metabolisation of Polyphenolic Compounds. Planta Med 2015; 81: 1075–1083

[19] Mortelé O, Vervliet P, Gys C, Degreef M, Cuykx M, Maudens K, Covaci A, van Nuijs ALN, Lai FY. In vitro Phase I and Phase II metabolism of the new designer benzodiazepine cloniprazepam using liquid chromatography coupled to quadrupole time-of-flight mass spectrometry. J Pharm Biomed Anal 2018; 153: 158–167

[20] Peeters L, Beirnaert C, Van der Auwera A, Bijttebier S, De Bruyne T, Laukens K, Pieters L, Hermans N, Foubert K. Revelation of the metabolic pathway of hederacoside C using an innovative data analysis strategy for dynamic multiclass biotransformation experiments. J Chromatogr A 2019; 1595: 240–247

[21] Beirnaert C, Peeters L, Meysman P, Bittremieux W, Foubert K, Custers D, Van der Auwera A, Cuykx M, Pieters L, Covaci A, Laukens K. Using Expert Driven Machine Learning to Enhance Dynamic Metabolomics Data Analysis. Metabolites 2019; 9: 54

[22] Mortelé O, Iturrospe E, Breynaert A, Verdickt E, Xavier BB, Lammens C, Malhotra-

Kumar S, Jorens PG, Pieters L, van Nuijs ALN, Hermans N. Optimization of an in vitro gut microbiome biotransformation platform with chlorogenic acid as model compound: from fecal sample to biotransformation product identification. J Pharm Biomed Anal 2019; 175: 1127768

[23] Hein EM, Rose K, Van't Slot G, Friedrich AW, Humpf HU. Deconjugation and degradation of flavonol glycosides by pig cecal microbiota characterized by fluorescence in situ hybridization (FISH). J Agric Food Chem 2008; 56: 2281–2290

[24] Olthof MR, Hollman PCH, Buijsman MNCP, van Amelsvoort JMM, Katan MB. Chlorogenic acid, quercetin-3-rutinoside and black tea phenols are extensively metabolized in humans. J Nutr 2003; 133: 1806–1814

[25] Justesen U, Arrigoni E. Electrospray ionisation mass spectrometric study of degradation products of quercetin, quercetin-3-glucoside and quercetin-3-rhamnoglucoside, produced by in vitro fermentation with human faecal flora. Rapid Commun Mass Spectrom 2001; 15: 477–483

[26] Jaganath IB, Mullen W, Lean MEJ, Edwards CA, Crozier A. *In vitro* catabolism of rutin by human fecal bacteria and the antioxidant capacity of its catabolites. Free Radic Biol Med 2009; 47: 1180–1189

[27] Aura AM, O'Leary KA, Williamson G, Ojala M, Bailey M, Puupponen-Pimiä R, Nuutila AM, Oksman-Caldentey KM, Poutanen K. Quercetin derivatives are deconjugated and converted to hydroxyphenylacetic acids but not methylated by human fecal flora *in vitro*. J Agric Food Chem 2002; 50: 1725–1730

[28] Braune A, Blaut M. Bacterial species involved in the conversion of dietary flavonoids in the human gut. Gut Microbes 2016; 7: 216–234

[29] Angelino D, Berhow M, Ninfali P, Jeffery EH. Caecal absorption of vitexin-2-*O*-xyloside and its aglycone apigenin, in the rat. Food Funct 2013; 4: 1339–1345

[30] Gao Y, Du Y, Ying Z, Leng A, Zhang W. Hepatic, gastric and intestinal first-pass effects of vitexin-2″-*O*-rhamnoside in rats by ultra-high-performance liquid chromatography. Biomed Chromatogr 2016; 30: 111–116

[31] Liang M, Xu W, Zhang W, Zhang C, Liu R, Shen Y, Li H, Wang X, Wang X, Pan Q, Chen C. Quantitative LC/MS/MS method and in vivo pharmacokinetic studies of vitexin rhamnoside, a bioactive constituent on cardiovascular system from hawthorn. Biomed

Chromatogr 2007; 21: 422–429

[32] Cai S, Chen Y, Zhang W, Ying X. Comparative study on the excretion of vitexin-4″- *O* -glucoside in mice after oral and intravenous administration by using HPLC. Biomed Chromatogr 2013; 27: 1375–1379

[33] Ma LY, Liu RH, Xu XD, Yu MQ, Zhang Q, Liu HL. The pharmacokinetics of *C*-glycosyl flavones of Hawthorn leaf flavonoids in rat after single dose oral administration. Phytomedicine 2010; 17: 640–645

[34] Hattori M, Shu Y-Z, El-Sedawy AI, Namba T, Kobashi K, Tomimori T. Metabolism of homoorientin by human intestinal bacteria. J Nat Prod 1988; 5: 874–878

[35] Kim M, Lee J, Han J. Deglycosylation of isoflavone *C*-glycosides by newly isolated human intestinal bacteria. J Sci Food Agric 2015; 95: 1925–1931

[36] Zheng S, Geng D, Liu S, Wang Q, Liu S, Wang R. A newly isolated human intestinal bacterium strain capable of deglycosylating flavone *C*-glycosides and its functional properties. Microb Cell Fact 2019; 18: 94

[37] Diukendjieva A, Alov P, Tsakovska I, Pencheva T, Richarz A, Kren V, Cronin MTD, Pajeva I. *In vitro* and *in silico* studies of the membrane permeability of natural flavonoids from *Silybum marianum* (L.) Gaertn. and their derivatives. Phytomedicine 2019; 53: 79–85

[38] Valentová K, Havlík J, Kosina P, Papoušková B, Jaimes JD, Káňová K, Petrásková L, Ulrichová J, Křen V. Biotransformation of silymarin flavonolignans by human fecal microbiota. Metabolites 2020; 10: 1–16

[39] Pferschy-Wenzig EM, Kunert O, Moissl-Eichinger C, Bauer R. Metabolization of Silymarin constituents by human fecal bacteria *in vitro*. In: 66th International Congress and Annual Meeting of the Society for Medicinal Plant and Natural Product Research (GA) jointly with the 11th Shanghai TCM conference. 2018: 1–2

[40] Sornsuvit C, Hongwiset D, Yotsawimonwat S, Toonkum M, Thongsawat S, Taesotikul W. The Bioavailability and Pharmacokinetics of Silymarin SMEDDS Formulation Study in Healthy Thai Volunteers. Evidence-based Complement Altern Med 2018; 2018: 1507834

[41] Bijak M. Silybin, a Major Bioactive Component of Milk Thistle (*Silybum marianum* L. Gaernt.)—Chemistry, Bioavailability, and Metabolism. Molecules 2017; 22: 1942

[42] Usman M, Ahmad M, Madni AU, Akhtar N, Asghar W, Akhtar M, Atif M, Qamar-uz-zaman M. *In-vivo* kinetics of silymarin (Milk Thistle) on healthy male volunteers. Trop J

Pharm Res 2009; 8: 311–316

[43] Saller R, Brignoli R, Melzer J, Meier R. An updated systematic review with meta-analysis for the clinical evidence of silymarin. Forsch Komplementmed 2008; 15: 9–20

[44] Schymanski EL, Jeon J, Gulde R, Fenner K, Ruff M, Singer HP, Hollender J. Identifying small molecules via high resolution mass spectrometry: Communicating confidence. Environ Sci Technol 2014; 48: 2097–2098

Legends for figures

Figure 1. Overview of the biotransformation experiment and sampling points. Method blank (MB), positive control of chlorogenic acid (PC-ChA), positive control of quercetin (PC-Q), crude extract samples (S1-S3) and negative control (NC).

Figure 2. Stability of (A) chlorogenic acid (**1**) and its metabolites: (B) 3,4-dihydroxyphenyl propionic acid (**C1**), (C) 4-hydroxyphenyl propionic acid (**C3**), (D) 3-hydroxyphenyl propionic acid (**C2**) in samples (FF), negative control (NCFF) and method blank (MB).

Figure 3. Stability of (A) chlorogenic acid (**1**) and its metabolites: (B) 3,4-dihydroxyphenyl propionic acid (**C1**), (C) 4-hydroxyphenyl propionic acid (**C3**), (D) 3-hydroxyphenyl propionic acid (**C2**) in the positive control.

Figure 4. *In vitro* biotransformation of chlorogenic acid. **1** = chlorogenic acid; **C1** = 3-(3,4-dihydroxyphenyl) propionic acid; **C2** = 3-(3-hydroxyphenyl) propionic acid, **C3** = 3-(4-hydroxyphenyl) propionic acid. Caffeic acid (CfA) and quinic acid (QAc) were not detected.

Figure 5. *In vitro* biotransformation of rutin. **15** = rutin; **R1** = quercetin; **C3** = 3-(4-hydroxyphenyl) propionic acid.

Figure 6. *In vitro* biotransformation of flavone di-*C*,*O*-glycosides and flavone *C*-glycosides. **3** = isoorientin-2″-*O*-xyloside; **5** = isoorientin-4″-*O*-xyloside; **6** = isoorientin; **7** = isoorientin-2″-*O*-rhamnoside; **8** = isovitexin-2″-*O*-glucoside; **9** = luteolin *C*-hexoside-*O*-hexoside; **11**= isovitexin-2″-*O*-xyloside; **13** = isovitexin 2″-*O*-rhamnoside; **14** = isovitexin; LT = luteolin; AP = apigenin.

Figure 7. *In vitro* biotransformation of saponin triterpenoids (**19**-**22**). **19** = niga-ichigoside F2; **20** = buergericic acid 28-*O*-glucoside; **21** = kaji-ichigoside F1; **22** = tormentoside; **T1** = 2*α*,3*α*,19*α*,23-tetrahydroxy-12-ursen-28-oic acid; **T2** = buergericic acid, **T3** = euscaphic acid and tormentic acid.

Table 1. Main compounds identified in the crude extract of *Cecropia obtusifolia*.

No.	Compound identification	Molecular formula	ESI negative full MS	RT (min)
	Phenolic acids			
1	Chlorogenic acid	$C_{16}H_{18}O_9$	353.0859 $[M - H]^-$	6.41
	Flavonoids			
2	Luteolin *C*-hexoside-*O*-hexoside	$C_{27}H_{30}O_{16}$	609.1459 $[M - H]^-$	8.49
3	Isoorientin-2″-*O*-xyloside	$C_{26}H_{28}O_{15}$	579.1354 $[M - H]^-$	8.70
4	Luteolin *C*-hexoside-*O*-hexoside	$C_{27}H_{30}O_{16}$	609.1454 $[M - H]^-$	8.45
5	Isoorientin-4″-*O*-xyloside	$C_{26}H_{28}O_{15}$	579.1355 $[M - H]^-$	8.91
6	Isoorientin	$C_{21}H_{20}O_{11}$	447.0928 $[M - H]^-$	9.22
7	Isoorientin-2″-*O*-rhamnoside	$C_{27}H_{30}O_{15}$	593.1510 $[M - H]^-$	9.16
8	Isovitexin-2″-*O*-glucoside	$C_{27}H_{30}O_{15}$	593.1509 $[M - H]^-$	9.49
9	Luteolin *C*-hexoside-*O*-hexoside	$C_{27}H_{30}O_{15}$	593.1510 $[M - H]^-$	9.76
10	Vitexin	$C_{21}H_{20}O_{10}$	431.0975 $[M - H]^-$	9.78
11	Isovitexin-2″-*O*-xyloside	$C_{26}H_{28}O_{14}$	563.1448 $[M - H]^-$	9.90
12	Diosmetin *C*-hexoside-*O*-pentoside	$C_{27}H_{30}O_{15}$	593.1507 $[M - H]^-$	10.52
13	Isovitexin 2″-*O*-rhamnoside	$C_{27}H_{30}O_{14}$	577.1560 $[M - H]^-$	10.34
14	Isovitexin	$C_{21}H_{20}O_{10}$	431.0978 $[M - H]^-$	10.68
15	Rutin	$C_{27}H_{30}O_{16}$	609.1459 $[M - H]^-$	11.09
16	Diosmetin-*C*-hexoside-*O*-deoxyhexoside	$C_{28}H_{32}O_{15}$	607.1662 $[M - H]^-$	10.47
	Flavonolignans			
17	Ent-vaccinin A	$C_{24}H_{16}O_9$	447.0716 $[M - H]^-$	15.73
18	Ent-mururin A	$C_{24}H_{16}O_9$	447.0716 $[M - H]^-$	16.34
	Triterpenoid saponins			
19	Niga-ichigoside F2	$C_{36}H_{58}O_{11}$	711.3953 $[M - H + HCOOH]^-$	13.55
20	Buergericic acid 28-*O*-glucoside	$C_{36}H_{58}O_{11}$	711.3952 $[M - H + HCOOH]^-$	14.48
21	Kaji-ichigoside F1	$C_{36}H_{58}O_{10}$	695.4006 $[M - H + HCOOH]^-$	16.75
22	Tormentoside	$C_{36}H_{58}O_{10}$	695.4001 $[M - H + HCOOH]^-$	17.59

Table 2. Metabolites identified and ranked via tinderesting. Level of confirmation as proposed by Schymanski et al.: Level 1 (L1): Structure confirmed by reference standard or by comparison with full identified compounds previously isolated from *C. obtusifolia*; level 3 (L3): tentative candidates based on MS, MS^2 experimental data; level 4 (L4): Unequivocal molecular formula based on MS1 and isotope ratios [44].

No.	Compound identification	RT (min)	Molecular formula	Identification confidence	ESI negative full MS	Delta ppm	ESI negative mode MS^E	*p*-value sample vs MB	*p*-value sample vs NC	Tinderesting score	Ranking tinderesting
	Phenolic acids										
1	Chlorogenic acid	6.41	$C_{16}H_{18}O_9$	L1	353.0869 $[M-H]^-$	-1.1	191.0553 $[M-H-162(CFA)]^-$, 179.0343 $[M-H-174(QA)]^-$	7.32E-09	2.07E-09	0.984	72
C1	3-(3,4)-dihydroxyphenyl propionic acid	4.83	$C_9H_{10}O_4$	L1	181.0503 $[M-H]^-$	1.1	137.0604 $[M-H-44(COO)]^-$	2.62E-01	2.62E-01	0.91	529
C3	3-(4-hydroxyphenyl) propionic acid	6.45	$C_9H_{10}O_3$	L4	165.0553 $[M-H]^-$	0.6	N/A	1.37E-06	8.49E-08	0.928	418
C2	3-(3-hydroxyphenyl) propionic acid	7.42	$C_9H_{10}O_3$	L1	165.0554 $[M-H]^-$	1.2	121.0653 $[M-H-44(COO)]^-$	0.00E+00	0.00E+00	0.974	107
	Flavonoids										
15	Rutin	11.09	$C_{27}H_{30}O_{16}$	L1	609.1459 $[M-H]^-$	0.5	301.0332 $[M-H-146(rha)-162(glu)]^-$,	2.98E-04	1.97E-03	0.986	66
R1	Quercetin	16.42	$C_{15}H_9O_7$	L1	301.0348 $[M-H]^-$	0.0	151.0035 $[M-H-150]^-$,	1.58E-01	1.58E-01	0.894	641
	Triterpenoid saponins										
19	Niga-ichigoside F2	13.55	$C_{36}H_{58}O_{11}$	L1	711.3953 $[M-H+HCOOH]^-$	-0.4	503.3381 $[M-H-162(glu)]^-$	0.00E+00	4.24E-13	1	16
20	Buergericic acid 28-*O*-glucoside	14.48	$C_{36}H_{58}O_{11}$	L1	711.3952 $[M-H+HCOOH]^-$	-0.6	503.3373 $[M-H-162(glu)]^-$	0.00E+00	1.14E-14	1	15
21	Kaji-ichigoside F1	16.75	$C_{36}H_{58}O_{10}$	L1	695.4006 $[M-H+HCOOH]^-$	-0.1	487.3415 $[M-H-162(glu)]^-$	2.28E-13	2.06E-10	1	13
22	Tormentoside	17.59	$C_{36}H_{58}O_{10}$	L1	695.4001 $[M-H+HCOOH]^-$	-0.9	487.3416 $[M-H-162(glu)]^-$	2.58E-10	8.02E-05	0.994	39
T1	2α,3α,19α,23-Tetrahydroxy-12-ursen-28-oic acid	18.9	$C_{30}H_{48}O_6$	L3	549.3424 $[M-H+HCOOH]^-$	-0.5	503.3365 $[M-H]^-$	8.21E-07	5.08E-09	0.996	29
T2	Buergericic acid	19.61	$C_{30}H_{48}O_6$	L3	549.3431 $[M-H+HCOOH]^-$	0.7	503.3376 $[M-H]^-$	1.62E-03	6.13E-08	0.966	144
T3	Euscaphic acid or tormentic acid	22.04	$C_{30}H_{48}O_5$	L3	533.3481 $[M-H+HCOOH]^-$	0.6	487.3429 $[M-H]^-$	3.90E-08	8.47E-11	0.994	37