Journal of Food Composition and Analysis 139 (2025) 107145

Contents lists available at ScienceDirect

Journal of Food Composition and Analysis

journal homepage: www.elsevier.com/locate/jfca

Characterization of flavone *C*-glycosides and phenolic compounds in tara (*Caesalpinia spinosa*) seed germ flour by HPLC coupled with high resolution mass spectrometry

Gianluca Picariello [a,*], Olga Fierro [a], Ermanno Vasca [b], Francesco Siano [a]

[a] Istituto di Scienze dell'Alimentazione, Consiglio Nazionale delle Ricerche, Via Roma 64, Avellino 83100, Italy
[b] Dipartimento di Chimica e Biologia "A. Zambelli", Università degli Studi di Salerno, Via Giovanni Paolo II 132, Fisciano, SA 84084, Italy

ARTICLE INFO

Keywords:
Tara germ flours
Flavone *C*-glycosides
Mixed *O*-glycosyl-*C*-glycosyl flavones
O-Acylated-di-*C*-glycosides
High-resolution HPLC-MS/MS
Flavonoid glycoconjugates

ABSTRACT

The cotyledon or germ of tara (*Caesalpinia spinosa*, *Fabaceae* family) is a protein- (45 %, w/w) and polyphenol-rich (14.9 g kg^{-1}) co-product of the extraction of tara seed gum (E417). In this work, the polyphenols of tara seed germ (TSG) were comprehensively characterized for the first time using HPLC-diode array detector and nanoflow HPLC coupled with high-resolution tandem mass spectrometry (MS/MS) in switching polarity ion mode and stepped collision energy. Apigenin di-*C*-glycoside derivatives were the largely predominant phenolic compounds of TSG. The concentrations of vicenin-2, vicenin-1 and vicenin-3 were 6.7, 1.4 and 1.5 g kg^{-1}, respectively, as determined by HPLC-UV. Several complex flavone *C*-glycoconjugates, *O*-glycosyl-*C*-glycosyl mixed derivatives and *O*-acylated-di-*C*-glycosides, among which benzoylated vicenin-2, were putatively assigned based on MS/MS spectra. The applied MS method was particularly effective to distinguish between possible isomeric acylated derivatives and to confirm their structural difference from multi-glycosylated isobaric compounds. Arguably, TSG is the richest source of flavone *C*-glycosides described so far. Because of its content of macro- and micro-nutrients, TSG is a multifunctional food ingredient obtained from co-products of neglected species although its safety profile remains to be assessed. In the meanwhile, pinpointing the phenolic compounds of TSG offers a valuable analytical reference for its monitoring.

1. Introduction

Caesalpinia spinosa, also known as tara or "Peruvian carob", is a leguminous shrub endemic to the tropical South American Andes, also introduced to Eastern Africa and China. The presence of tara trees is mainly depending on weather conditions regardless of soil origin and composition. In fact, they can thrive in dry forests on marginal soils and in semi-arid environments, due to a high level of adaptability to water scarcity and limited nutritional needs. In several areas of the Andean regions, tara trees are cultivated by local communities, integrated into agroforestry systems, to accomplish two important ecological functions: i) soil water retention; ii) phytoremediation and phytostabilization practices for metal contaminated soils, as they are metallophyte plants (Orchard et al., 2009; Balaguer et al., 2011; Murga-Orillo et al., 2023). Tara trees produce pods with high economic value thanks to multipurpose uses and rapidly increasing international demand (Rojas et al. 2007). Ground dried tara pods contain high amounts of tannic and gallic acids and gallotannins widely used in the leather, pharmaceutical, cosmetic, chemical, or paint industries. Interesting food applications exploiting the antioxidant properties of tara pod polyphenols have recently been described (Pedreschi et al., 2018; Pedreschi et al., 2022). Peru is the world's largest producer and exporter of tara-derived products accounting for approximately 85 % of overall production (Murga-Orillo et al., 2023).

Around 35 % by weight of the tara pod is constituted by seeds. The endosperm represents 35–40 % by weight of the tara seed, and is almost completely constituted by a galactomannan, also known as tara gum (Sangay-Tucto and Duponnois, 2018), which is an excellent hydrocolloid that can be used at a 0.5–2 % (w/w) concentration as a stabilizer, thickener, viscosity modifier, fat replacer, and emulsifier agent for food, pharmaceutical and cosmetic formulations (Gelling, 2019; Mukherjee et al., 2023). Although its use began at least 50 years ago, tara gum is currently the food additive ranking as first among the fastest-growing ingredients for ice creams, desserts, and confectionery, with a

* Corresponding author.
E-mail address: picariello@isa.cnr.it (G. Picariello).

https://doi.org/10.1016/j.jfca.2024.107145
Received 27 June 2024; Received in revised form 11 December 2024; Accepted 18 December 2024
Available online 19 December 2024

forecasted Compound Annual Growth Rate of 13.6 % expected over the period 2024–2031 (https://issuu.com/reportprime-2, accessed on April 2024). Like locust bean gum (E410), which is extracted from *Ceratonia siliqua* (Mediterranean carob) seeds, tara gum is a food additive approved in the European Union as E417 and permitted by the food safety authorities of the USA, Japan, and Canada (Desai et al., 2022).

The germ or embryo makes up around 30 % of the tara seed weight and is a co-product of the industrial extraction of E417. Tara seed germ (TSG) is only barely used as a protein source for animal feed (Re-Jiménez and Amadó., 1989), while the considerable amount of this waste product causes relevant storage and disposal costs for manufacturing companies. Until recently, available information on TSG composition was fragmentary, mainly based on rather dated determinations (Re-Jiménez and Amadó., 1989; Borzelleca et al., 1993). However, the composition data complemented with those lately published by us point to TSG as a candidate ingredient for functional foods. In particular, TSG is a rich source of vegetable proteins (ca. 45 %, w/w) of good biological value and microelements (Fierro et al., 2024).

An exhaustive characterization of TSG composition has become a hot topic also for food safety reasons. In recent years, some North American companies have introduced TSG flour as an ingredient in high-protein crumbles and "superfood" smoothies. Shortly thereafter, TSG has been indicated as the alleged causative agent of an outbreak of serious acute adverse health events in hundreds of people who were affected by sharp abdominal pain, black urine, liver damage, and gallbladder failure. The outbreak has forced the companies involved to issue a recall of their food products, while the USA Food and Drug Administration (FDA) is still investigating to establish the cause of the adverse events. Some nonprotein amino acids constitutively occurring in tara germ, such as baikiain, have been identified as the most plausible toxic agents of TSG (Chittiboyina et al., 2023). On the other hand, based on the capability of tara germ to accumulate large amounts of metals, we speculated that TSG hazards could be some rare earth elements that contaminate those few batches of flour deriving from plants grown on polluted soils (Fierro et al., 2024). Pending the safety assessment and to prevent needless risks, in May 2024 the FDA has instituted screening at ports of entry for TSG used as an ingredient in imported foods (FDA, 2024). Soon after, also the European Food Safety Authority (EFSA) has included TSG among the emerging food risks (Georganas et al., 2024).

TSG is intensely yellow-colored, suggesting the presence of phenolic compounds structurally correlated to those found in the germ of *Leguminosae* plants taxonomically close to *C. spinosa*, such as *Prosopis* spp. and *Ceratonia siliqua* (Picariello et al., 2017). In fact, preliminary high-performance liquid chromatography-diode array detector (HPLC-DAD) analyses of polyphenol extracts from TSG have highlighted the occurrence of gallic acid and flavonoid *C*-glycosides (Fierro et al., 2024; Chittiboyina et al., 2023). However, total polyphenol content and antioxidant capacity of TSG are noticeably higher than the carob and *Prosopis* spp. counterparts (Fierro et al., 2024). Despite their potential interest, phenolic compounds of TSG have not been described in detail so far.

The polyphenol profile of authentic TSG, such as the sample analysed in this study, compared to that of the specific TSG sample believed responsible for the adverse events, which has been published by Chittiboyina et al. (2023), can be diagnostic to exclude the hypothesis that plants other than tara were used for the preparation of the toxic flour.

In this work, we present for the first time a comprehensive characterization of TSG phenolic compounds based on HPLC-DAD and nanoflow HPLC coupled with high resolution tandem mass spectrometry (MS/MS).

2. Materials and methods

2.1. Chemicals

All chemicals, solvents, and HPLC-MS grade water used in this work were purchased from Merck-Sigma-Aldrich and were of the highest purity degree available (St. Luis, Mo, USA). Pure schaftoside and isoschaftoside (purity > 99 %) were from Merck-Sigma-Aldrich too.

2.2. Tara germ flour sample

TSG flour was provided by Silvateam SpA (San Michele Mondovì, Cuneo, Italy), which is one of the largest companies committed in the commercialization of Peruvian tara products. Tara trees were grown in the Huánuco district, Peru, and pods were collected in 2023. The germ flour results as a co-product from the gum extraction from the seeds. TSG flour was received vacuum packed in a refrigerated box and was stored at −26 °C until analysis. Carob (*Ceratonia siliqua*) germ flour, used as a reference sample in this study, was donated by Carob Ingredients Co., Casablanca, Morocco (https://www.carobingredients.com).

2.3. Polyphenol extraction

TSG and reference carob seed germ extracts were obtained from 1.0 g of flour treated sequentially with (i) 10 mL methanol, (ii) 10 mL 80 % aqueous methanol (v/v) and (iii) 10 mL 80 % aqueous ethanol (v/v). The slurry was centrifuged at 10,000 $g \times 15$ min at room temperature between consecutive extraction steps and the supernatants were combined. Extracts were prepared in triplicate and the yellow-colored extracts were filtered using 0.22 μm PVDF disposable syringe filters (Millex, Millipore, Badford, MA, USA) and analyzed immediately or stored at −26 °C.

2.4. Total polyphenols, radical scavenging activity and antioxidant capacity

The concentration of total phenolic compounds (TPC) in the extracts was determined by the Folin-Ciocalteu colorimetric method, using the general procedure recommended by the European Pharmacopoeia for the determination of total tannins (European Pharmacopoeia, 2007). The absorbance of the Folin-Ciocalteu complexes was assayed at 760 nm using a UV-Vis spectrophotometer (Amersham Ultrospec 2100 Pro UV/Vis, GE Healthcare, Uppsala, Sweden). TPC was referred to a standard curve built with gallic acid within the 20–250 mg kg^{-1} concentration range, and expressed as g of gallic acid equivalents (GAE) kg^{-1}. TPC values are average of three independent determinations. Radical scavenging activity of the polyphenol extracts from TSG and, in comparison, from carob seed germ (CSG) flour was assessed with the DPPH assay against a calibration curve built with 2–25 mg L^{-1} gallic acid. Values are expressed as g_{GAE} kg^{-1}. DPPH inhibition of 1 % (w/w) aqueous methanolic (80 %, v/v) solutions of TSG and CSG was determined according to a previously detailed procedure (Siano et al., 2023 - a), while the antioxidant capacity was determined coulometrically (as CDAC, coulometrically determined antioxidant capacity, mmol e^- kg^{-1}) based on a coulometric method recently developed by our group (Siano et al., 2023 – b). Standard curves for the determination of TPC, DPPH radical scavenging activity and DPPH radical inhibition (%) are reported in **Supporting information Figure S1**.

2.5. HPLC-DAD analysis

TSG hydroalcoholic extracts were separated using a modular HP 1100 chromatographer (Agilent Technologies, Paolo Alto, CA, USA) equipped with a diode array detector (DAD). The stationary phase was a 250 × 2.1 mm i.d. C18 reversed-phase Aeris core-shell column, 3.6 μm particle diameter (Phenomenex, Torrance, CA, USA). The column temperature was maintained at 37 °C during the HPLC analyses. Runs were performed at a constant flow rate of 0.2 mL min^{-1} applying a 5–60 % linear gradient of the solvent B (acetonitrile/0.1 % TFA) after 5 min of isocratic elution at 5 % B. Solvent A was 0.1 % TFA in HPLC-grade water. For each analysis, 10 μL of combined extracts 10-fold diluted with 0.1 %

TFA were injected. Samples were run in triplicate, at least. DAD acquired an UV-Vis spectra every second between 200 and 600 nm. The HPLC separations were monitored at 520, 360, 320 and 280 nm. Chromatograms were elaborated with the ChemStation software vers. A.07.01 (Agilent Technologies). Amounts of vicenin-2 and vicenin-1/vicenin-3 were estimated using an external calibration curve built with pure schaftoside. To consider matrix effects TSG extracts were spiked with known amounts of pure schaftoside, which is practically missing in TSG extracts, as the internal standard. Validation parameters of the HPLC-UV based quantification of vicenins are reported in **Supporting information Figure S2**.

To correlate peaks between HPLC-DAD and nanoflow HPLC-MS chromatograms, the most intense HPLC-DAD peaks were manually collected, dried in a speed-vac, resuspended in nearly the same eluting volume of 0.1 % formic acid. The resulting solutions (2 μL) were individually analyzed by nanoflow HPLC-MS/MS using the same conditions for TSG extracts as detailed below.

2.6. Nanoflow HPLC-MS/MS analysis

Hydroalcoholic (80 % methanol, v/v) TSG extracts were ten-fold diluted with 0.1 % formic acid and analyzed by nanoflow-HPLC coupled with a high-resolution Q Exactive mass spectrometer (Dionex/Thermo Scientific, San Jose, CA, USA). For each analysis, 2 μL of the resulting solution were loaded through a 20 mm× 75 μm i.d. pre-column (LC Packings, USA) using a Famos (Dionex/Thermo Scientific) autosampler. Phenolic compounds were separated with an EASY-Spray™ C18 column (2 μm, 25 cm×75 μm i.d.) 3 μm particles, 100 Å pore size (Thermo Scientific), integrated with an in-source emitter tip with a capillary voltage 2.4 kV at 250 °C. Eluent A was 0.1 % formic acid (v/v) in water; eluent B was 0.1 % formic acid (v/v) in 80 % acetonitrile. Separation was carried out applying a 4–60 % linear gradient of B over 60 min at the flow rate of 300 nL min^{-1}. Spectra were acquired in positive and negative ionization polarity switching mode (1 s) and scanned the 120–1200 *m/z* range. The instrument operated in top-5 data-dependent acquisition for MS/MS, in automatically stepped collision energy (18, 25, and 30 arbitrary units), with 10 s of dynamic exclusion. MS1 and MS2 spectra were acquired at a 70,000 and 17,500 Full Width at Half Maximum (FWHM) resolution, respectively.

To confirm the assignment of gallic acid, the eluate from HPLC-UV at retention time (t_R) 7.6 min was manually collected and analysed by MS and MS/MS by flow direct injection, infusing it into the mass spectrometer through a syringe pump and a capillary nanoemitter (Thermo Scientific) at the 1 μL flow rate. Spectra were elaborated using the Xcalibur Software 3.1 version (Thermo Scientific), matched with literature data and open-source databases of phenolic compounds, and manually validated.

2.7. Hydrolysis of the extract and GC-MS analysis

To identify the glycoside with $[M-H]^-$= 443.1891, the TSG extracts were separated by HPLC using a semipreparative C18 column 250 × 10 mm i.d. (Phenomenex) with the same elution gradient as above at a 3 mL min^{-1} constant flow rate. Eluates at t_R 19.4 and 22.9 (λ_{MAX} = 270 nm) were manually collected, dried in speed-vac and lyophilized. The residue was hydrolyzed with 3 mL of a 2 M TFA solution, for 6 h at 90 °C, neutralized and the aglycone was extracted in ethyl acetate. The aglycone ($[M-H]^-$ = 149.0583, calculated molecular formula $C_9H_{10}O_2$) was analysed by GC-MS using a 7890 A gas chromatograph (Agilent Technologies, Palo Alto, CA, USA) directly interfaced with a 5975 C VL MSD Triple-Axis Detector mass-spectrometer (Agilent). The apparatus was equipped with an autosampler system Model G4513A with a 10 μL syringe. Analysis of hydrolyzed compounds was carried out by injecting 1 μL samples on an SPB-50 capillary column (Supelco), 30 m × 0.25 mm i.d. × 0.25 μm thickness film, through a split/splitless injector (splitless mode). The carrier gas was helium at a constant flow rate of 1.0 mL min^{-1}, and the make-up gas was nitrogen at a flow rate of 30 mL min^{-1}. The initial oven temperature was set at 60 °C for 5 min, and it was programmed from 60 to 300 °C at 15 °C min^{-1} and hold for 5 min. The injector temperatures were both set at 250 °C. The GC-MS interface temperature was 280 °C, and the electron ionization energy was 70 eV. Full-scan mass spectra were recorded in the 600–50 *m/z* range. The compound was assigned by matching the experimental mass spectra with reference spectra in the NIST14 library. Data were recorded and processed using the MassHunter Workstation Software (Agilent Technologies).

2.8. α-Amylase and α-glycosidase inhibition assays

The inhibition activity of TSG extracts against porcine pancreatic α-amylase and small intestinal α-glycosidases was assayed using the amylase activity assay kit (cod. MAK009, Merck-Sigma) and the α-glycosidase activity assay kit (cod. MAK123, Merck-Sigma), respectively, according to the manufacturer's instruction, as previously detailed (Siano et al., 2023 – a). Acarbose (Merck-Sigma) was used as the positive control for both α-amylase and α-glycosidase inhibition assays.

3. Results

3.1. Total polyphenols of TSG extracts and antioxidant capacity

The pigments of TSG are soluble in hydroalcoholic solutions suggesting the presence of flavonoid glycosides like the CSG counterpart. As recently reported by us (Fierro et al., 2024), total polyphenols of TSG determined with the Folin-Ciocalteu method were 14.9 g_{GAE} kg^{-1}, which largely exceeds the polyphenol amount found for carob and *Prosopis* spp. Germ (Picariello et al., 2017). Consistently, the DPPH radical scavenging activity of TSG (16.8 g_{GAE} kg^{-1}) was much higher than the one determined for CSG (7.6 g_{GAE} kg^{-1}). The concentration of the most abundant flavone *C*-glycoside of TSG, *i.e.*, vicenin-2 (see below), was 6.7 g kg^{-1} while vicenin-1 and vicenin-3 were 1.4 and 1.5 g kg^{-1}, respectively, as estimated by HPLC-UV against an external calibration curve built with pure schaftoside and spiking the TSG extracts with authentic schaftoside used as the internal standard. The DPPH radical inhibition of a 1 % (w/w) hydroalcoholic solution (80 % methanol, v/v) of TSG was 43.1 %, almost double than the carob counterpart (23.9 %). The antioxidant capacity of the two extracts was comparatively determined with a recently developed coulometric method, referred to as CDAC (Siano et al., 2023 - b). CDAC of TSG was 624 mmol e^- kg^{-1}, which is a value comparable to many common pure antioxidants and was 3.5 times higher than CSG (178 mmol e^- kg^{-1}). Although the relatively high antioxidant capacity determined *in vitro* should not be intended as the evidence of health-promoting properties *in vivo*, our data pointed to particularly abundant flavonoids in TSG.

3.2. HPLC-UV analysis of TSG hydroalcoholic extracts and MS assignment of most abundant compounds

TSG hydroalcoholic extracts were analysed by reversed phase HPLC-DAD and compared to the counterpart from CSG, which has been previously characterized (Picariello et al., 2017) (Fig. 1). The most abundant components of TSG were collected and individually analysed by high resolution MS and MS/MS. The HPLC-UV (280 nm) of the hydroalcoholic extract from TSG is dominated by gallic acid eluting at t_R 7.6 min, as confirmed by comparative analysis with the authentic standard and flow direct injection MS and MS/MS analysis. The UV spectrum of the intense sharp peak eluting at t_R 24.7 min (inset of Fig. 1) is related to those of glycosylated derivatives of apigenin, although its t_R does not correspond to the schaftoside and isoschaftoside that prevail in the chromatogram of CSG. The MS analysis of the component at t_R 24.7 min showed a $[M-H]^-$ = 593.1479 while its MS/MS spectrum exhibited the typical fragmentation pattern of flavonoid di-*C*-glycosides

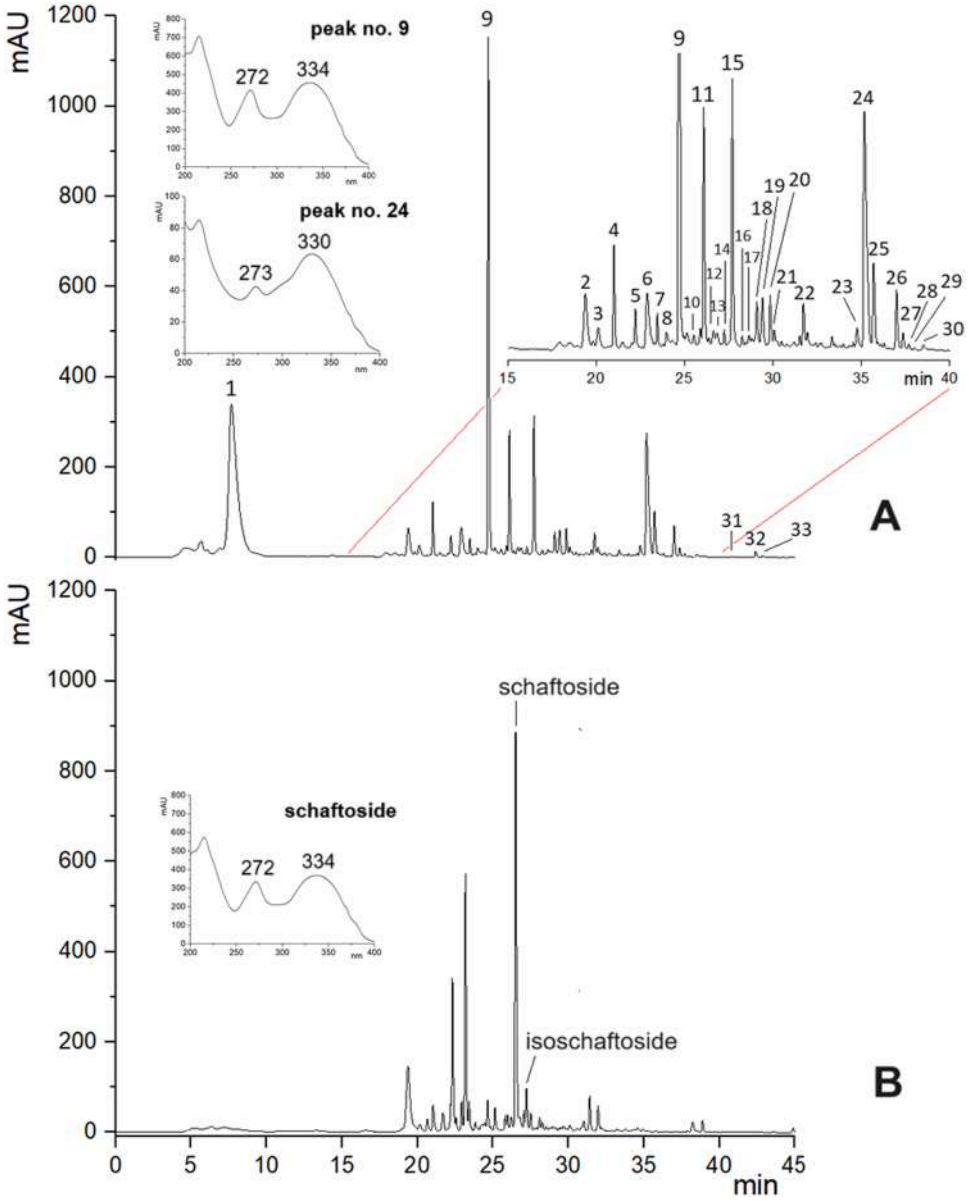

Fig. 1. RP-HPLC-DAD (280 nm) chromatogram of the hydroalcoholic extract from TSG (A) compared to the CSG counterpart (B). Labelled peaks are assigned in Table 1. The UV-Vis spectra of peaks no. 9 and 24 in the TSG extract and schaftoside in the CSG extract are shown in the insets.

with mass difference of 30 uma between adjacent signals arising from intraglycosidic cleavages, as schematized in Fig. 2 (Cao et al., 2014). The pair of signals at 353.0641 and 383.0754 is diagnostic for the apigenin as they correspond to the mass of the $[Agl+83]^-$ and $[Agl+113]^-$, where Agl is the molecular weight of the aglycone that in the case of apigenin is 270.0528. Based on these data, the most abundant phenolic compound of TSG was identified as vicenin-2, that is apigenin 6, 8-di-*C*-β-D-glucopyranoside, in line with a very recent report (Chittiboyina et al., 2023). The intense HPLC peaks at t_R 26.1 and 27.7 min were isobaric with schaftoside and isoschaftoside, although their t_R did not correspond to either.

3.3. Nanoflow HPLC-MS/MS analysis

Because of the heterogeneous composition of the TSG extracts, the nanoflow HPLC-MS/MS analysis was carried out in switching polarity positive/negative ion mode. After some exploratory analyses at fixed (27 arbitrary units) collision energy, an automatically stepped collision energy was used for MS/MS fragmentation, which allows to optimize the balance between parent and daughter ions in the spectra. In fact, TSG contains simultaneously *O*- and *C*-glycosides, which require significantly different collision energy to produce informative MS/MS spectra. As expected, the high resolution/high sensitivity HPLC-MS chromatogram highlighted a more complex and component-rich profile of the TSG extracts (Fig. 3). In the chromatographic configuration of this study, gallic acid was not detected as it was lost during the wash step of the trap column. Table 1 reports the list of compounds in the TSG extracts assigned based on HPLC-MS/MS analysis. In general, phenolic compounds exhibit higher ionization efficiency in the negative than in the positive ion mode. In the specific case of flavonoid-*C*-glycosides, the negative ion mode MS/MS spectra are more informative than the positive ones since the latter are complicated by multiple fragmentation signals arising from loss of water and formaldehyde (Vukics and Guttman, 2010). Nevertheless, the measured mass and relevant main fragments in both the ion modes are listed in Table 1. Flavonoid *O*-glycosides generally fragment breaking the *O*-glycoside bond and generate signals corresponding to the neutral loss of sugar moiety(es) with production of prominent Y_0 (negative ion mode) or Y_0^+ (positive ion mode) aglycone

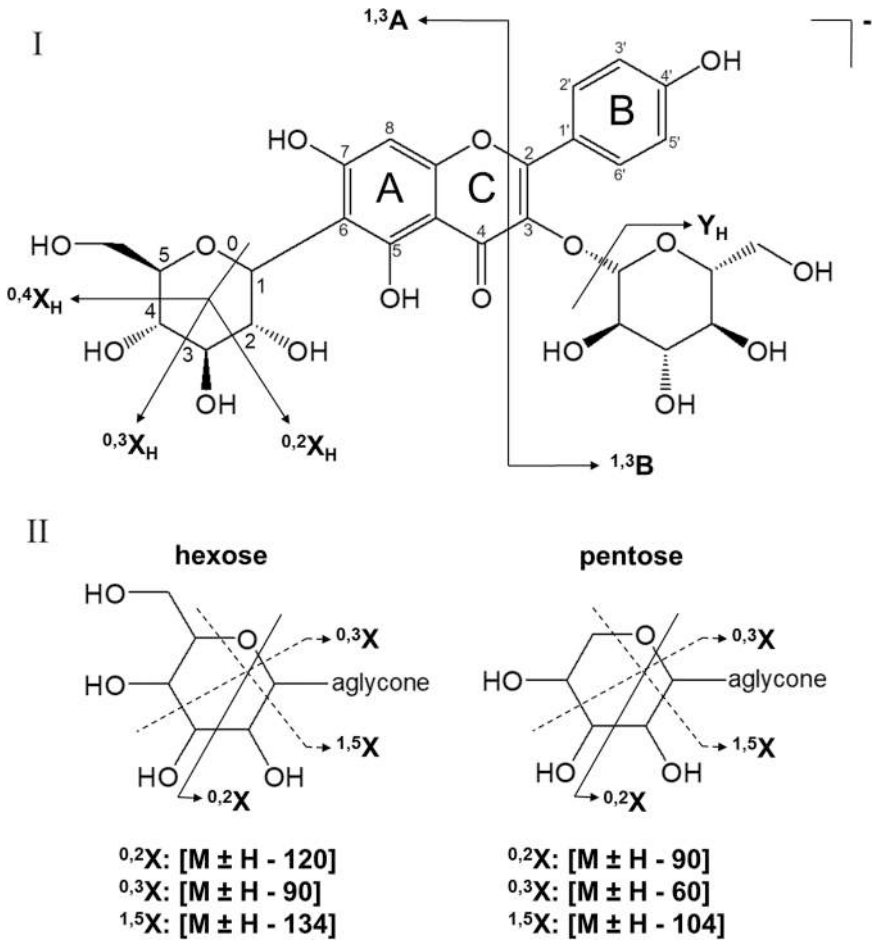

Fig. 2. Scheme of the main fragmentation routes of *C*- and *O*-glycosides (I) and intraglycosidic cleavage observed in *C*-glycosides (II). Typically, *O*-glycosides breakdown at level of the *O*-glycoside bonds with neutral loss of glycoside moieties. This fragmentation pathway yields the so called Y_0^- or Y_0^+ fragments with direct or stepwise formation of the aglycone ion. The fragment ions of *C*-glycosides generally result from fragmentation within the glycoside moieties. Figure adapted from Cao et al., 2014.

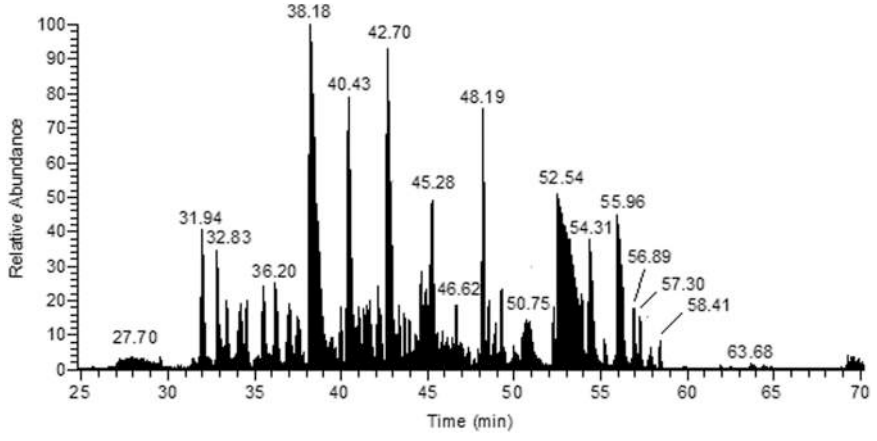

Fig. 3. Nanoflow HPLC-MS Total ion current (TIC) chromatogram of TSG hydroalcoholic extract. The main peaks have been assigned through the t_R value in Table 1.

signals. In contrast, the *C-C* glycoside-aglycone bond of flavonoid-*C*-glycosides survives even at high collision energy, while fragmentation preferentially involves breakdown pairs of intraglycosidic bonds (Fig. 2) (Cao et al., 2014). Thus, second order mass spectra (MS/MS) allow a straightforward identification of *C*-glycosides through the typical cross-ring decay that generates fragments differing by 30 uma, corresponding to CH-OH moieties, such as $[(M-H) - 90]^-$ ($[^{0,3}X_0-H]^-$ ion) and $[(M-H) - 120]^-$ ($[^{0,2}X_0-H]^-$ ion) signals. For the nomenclature of fragments, refer to Domon and Costello (1988). It has been observed that in asymmetric flavone di-*C*-glycosides, *i.e.*, those carrying different glycoside moieties at positions 6 and 8 of the aglycone, the intensity of the $[(M-H) - 60]^-$ ($[^{0,4}X_0-H]^-$ ion) signal differs between the isomeric species due to hydrogen bonding with two proxime OH groups of the flavone aglycone. Specifically, the fragment

Table 1
High resolution MS/MS-based assignment of the glycosides of TSG. Pentose and hexose moieties have been indicated (*e.g.*, arabinosyl, xylosyl, glucosyl) when a confident assignment was possible, otherwise they have been generically indicated as pentosyl or hexosyl moieties. A confident assignment of a glycosyl moiety to glucose is indicated as a glucosyl or glucoside derivative. Selected diagnostic signals are evidenced in bold. Gallic acid was assigned by HPLC-DAD comparison with the authentic standard and flow direct injection MS and MS/MS. HPLC-MS/MS peaks were correlated to the HPLC-DAD analysis based on UV spectra, peak intensity, elution order and, in some cases, manual collection of isolated peaks and individual HPLC-MS analysis. Compounds not assigned by HPLC-DAD were low-abundant species and were detected only by HPLC-MS/MS.

HPLC-MS t_R (min)	HPLC-UV Peak no.	$[M-H]^-$	MS/MS fragments $[M-H]^-$	$[M+H]^+$	MS/MS fragments $[M+H]^+$	Putative assignment
n.d.	1	169.0139	125.0241			Gallic acid
31.14		917.2543	755.2008 635.1555 515.1122 473.1051 383.0738 353.0641	919.2693	n.f.	apigenin 6-*C*-(*O*-hexosyl)-glucosyl−8-*C*-(*O*-hexosyl)-glucoside*
31.94	2	443.1896	149.0585	n.d.		2-methoxy−4-vinyl phenol *O*-(*O*-pentosyl)-hexoside
32.02	3	495.0753	343.0643 325.0557 169.0118	497.0939	327.0711 309.0606 **153.0184**	digalloyl quinic acid
32.81	4	755.1999	635.1587 **593.1450** 545.1219 515.1146 473.1046 383.0728 353.0634	757.2171	619.1671 541.1356 457.1131 379.0819	apigenin 4'-*O*-hexosyl−6,8-di-*C*-glucoside (O-glycosyalted vicenin−2)
33.48		495.0756	343.0645 325.0541 191.0537 169.0119	497.0929	327.0715 309.0608 **153.0185**	digalloyl quinic acid isomer
33.90	5	755.2004	635.1598 575.1384 515.1158 473.1053 383.0736 353.0639	757.2165	619.1661 541.1358 457.1142 379.0816	apigenin 6-*C*-glucosyl−8-*C*-(*O*-hexosyl)-glucoside*
35.19	6	459.1481	165.0533	n.d.		epoxy 2-methoxy−4-vinyl phenol *O*-(*O*-pentosyl)-hexoside
34.51	7	725.1900	665.1733 635.1571 605.1464 **563.1380** 545.1282 515.1166 473.1079 443.0953 383.0742 353.0643	727.2080	589.1566 511.1240 427.1029 409.0914 379.0818	apigenin 4'-*O*-hexosyl−8-*C*-xylosyl−6-*C*-glucoside (*O*-glycosylated vicenin−1 isomer)
35.64	8	725.1907	635.1610 605.1473 563.1380 545.1254 515.1190 473.1079 383.0737 353.0626	727.2076	619.1680 559.1472 511.1234 451.1027 379.0816	apigenin 6-C-xyloside-(O-hexosyl)−8-C–glucoside isomer*
36.21		609.1433	519.1097 489.1015 429.0790 399.0695 269.0590	611.1602	557.1323 527.1197 473.1089 443.0975 425.0875 395.0766 341.0664	luteolin 6,8-di-*C*-hexoside isomer
36.24		725.1895	635.1580 605.1448 **563.1392** 545.1271 515.1158 473.1061 443.0933 383.0739 353.0639	727.2071	589.1574 559.1432 511.1241 427.1021 409.0923 379.0815	apigenin 4'-*O*-hexosyl−6-*C*-xyloside−8-*C*-glucoside (O-glycosylated vicenin−3 isomer)
37.71		609.1432	519.1106 489.1004 429.0788	611.1604	557.1277 527.1187 473.1081 443.0975	luteolin 6,8-di-*C*-glucoside (lucenin−2)

(*continued on next page*)

Table 1 (*continued*)

HPLC-MS t_R (min)	HPLC-UV Peak no.	$[M-H]^-$	MS/MS fragments $[M-H]^-$	$[M+H]^+$	MS/MS fragments $[M+H]^+$	Putative assignment
			399.0690		425.0875	
			369.0587		395.0768	
					341.0659	
38.06		725.1895	635.1573	727.2074	619.1643	apigenin 6-*C*-pentosyl−8-*C*-(*O*-hexosyl)-glucoside isomer*
			605.1454		559.1454	
			545.1245		511.1229	
			515.1160		427.1026	
			473.1065		409.0914	
			443.0951		397.0927	
			383.0739		379.0818	
			353.0639			
38.18	9	593.1479	503.1163	595.1648	541.1363	apigenin 6,8-di-*C*-glucoside (vicenin−2)
			473.1055		457.1136	
			413.0842		409.0921	
			383.0754		379.0816	
			353.0641		325.0706	
40.04	10	647.0864	495.0759	649.1061	429.0495	3,4,5-tri-*O*-galloyl quinic acid
			343.0645		153.0185	
			325.0568			
			169.0119			
40.43	11	563.1375	503.1174	565.1552	529.1346	apigenin 6-C-xyloside−8-C-glucoside
			473.1064		511.1230	(vicenin−1)
			443.0956		427.1028	
			383.0747		409.0923	
			353.0642		391.0806	
					379.0809	
					325.0714	
40.54		647.0859	495.0751	649.1059	429.0499	3,4,5-tri-*O*-galloyl quinic acid
			343.1644		153.0185	
			325.0534			
			169.0119			
41.30	12	563.1378	503.1134	565.1559	529.1343	apigenin 6-C-pentosyl−8-C-hexoside isomer
			473.1055		511.1230	
			443.0955		427.1033	
			383.0751		409.0918	
			353.0648		379.0818	
					325.0711	
41.59	13	563.1383	503.1159	565.1566	529.1343	apigenin 8-C-pentosyl−6-C-hexoside isomer
			473.1056		511.1228	
			443.0957		427.1026	
			383.0753		409.0922	
			33.0644		379.0813	
					325.0708	
41.78	14	593.1485	503.1133	595.1659	559.1450	apigenin 6,8-di-*C*-hexoside (vicenin−2 isomer)
			473.1052		481.1157	
			413.0857		457.1130	
			383.0746		409.0919	
			353.0642		379.0816	
					325.0705	
42.15		593.1486	503.1180	595.1663	541.1350	apigenin 6,8-di-*C*-hexoside (vicenin−2 isomer)
			473.1054		457.1136	
			413.0856		409.0923	
			383.0759		379.0707	
			353.0637		325.0708	
42.70	15	563.1375	473.1045	565.1551	529.1352	apigenin 8-C-xylosyl−6-C-glucoside
			443.0951		457.1157	(vicenin−3)
			383.0754		397.0934	
			353.0648		379.0818	
					325.0705	
43.44	16	679.1486	635.1582	681.1661	561.1221	apigenin 6-*C*-(O-malonyl)-glucosyl−8-*C*-glucoside
			593.1484		457.1135	
			515.1171		379.0808	
			473.1055		355.0821	
			383.0749		337.0713	
			353.0641		325.0710	
43.80		577.1522	503.1161		543.1105	apigenin 8-C-deoxyhexosyl−6-C-glucoside*
			473.1053		525.1387	
			457.1109		423.1086	
			383.0745		379.0815	
			353.0638		325.0709	
44.60	17	487.1425	443.1516	489.1551	n.f.	digalloyl quinic acid
			401.1425			
			269.1011			
			161.0431			

(*continued on next page*)

Table 1 (*continued*)

HPLC-MS t_R (min)	HPLC-UV Peak no.	$[M-H]^-$	MS/MS fragments $[M-H]^-$	$[M+H]^+$	MS/MS fragments $[M+H]^+$	Putative assignment
44.85		917.2321	n.f.	919.2498	589.1323 457.1130 379.0805 325.0694 **163.0392**	apigenin *O*-caffeoyl−6-*C*-glucosyl−8-*C*-(*O*-hexosyl)-glucoside*
45.03		563.1368	503.1145 473.1053 443.0948 383.0749 353.0638	565.1559	529.1374 511.1248 409.0912 397.0915 379.0816 325.0713	apigenin 6-C-arabinosyl−8-c-glucoside (isoschaftoside)
45.08		431.0957	341.0655 311.0539	433.1130	415.1027 397.0921 379.0827 337.0707 303.0708	apigenin 8-*C*-glucoside (vitexin)
45.13	18	609.1426	300.0251 301.0331	611.1605	303.0502	quercetin−3-*O*-rutinoside (rutin)
45.28	19	533.1251	473.1039 443.0941 413.1853 383.0744 353.0636	535.1448	499.1242 481.1143 427.1018 409.0927 379.0823	apigenin 6,8-di-*C*-pentoside isomer
45.54	20	901.2360	811.2039 781.1932 575.1358 473.1049 455.0955 383.0742 353.0642	903.2553	457.1133 379.0815 325.0711 **147.0443**	apigenin 6-*C*-(*O*-*p*-coumaryl)-glucosyl−8-*C*-(*O*-hexosyl)-glucoside*
45.60	21	431.0959	341.0652 3110540	433.1134	415.1031 379.0923 379.0821 337.0707 313.0709	apigenin 6-*C*-glucoside (isovitexin)
45.75		931.2472	841.2158 811.2047 575.1368 455.0953 383.0744 353.0639	933.2651	457.1135 379.0130 325.0707 **177.0549**	apigenin 6-*C*-(*O*-feruloyl)-glucosyl−8-*C*-(*O*-hexosyl)-glucoside*
46.16		577.1527	473.1048 457.1110 383.0743 353.0634	579.1726	543.1483 525.1368 457.1139 393.0971 325.0704	apigenin 8-C-deoxyhexosyl−6-C-glucoside*
46.27		533.1277	473.1070 443.0964 383.0750 353.0645	535.1460	499.1255 481.1133 427.1025 409.0922 379.0815	apigenin 6,8-di-*C*-pentoside isomer
46.62		463.0865	300.0225 301.0330	465.1043	303.0501	quercetin 3-*O*-hexoside
47.89		961.2572	871.2258 841.2135 755.2009 635.1575 **593.1431** 575.1380 545.1265 515.1162 383.0750 353.0639	963.2578	619.1655 541.1364 517.1323 457.1136 379.0814 355.0816 325.0703 **207.0655** 175.0393	apigenin *O*-sinapoyl−6-C-(O-glycosyl)−8-*C*-glucoside isomer
48.19	22	665.2792	647.2723 615.2456 474.1664 345.1451 303.1338 204.0660 140.0709	667.2969	635.2708 478.1984 452.2179 337.1184 311.1391 287.1409	Unknown – probable luoteolin O-glycoside derivative
48.70		961.2576	871.2244 841.2138 755.1956 635.1572	963.2575	825.2198 619.1674 517.1354 457.1140	apigenin 6-*C*-(O-sinapoyl)-hexosyl−8-*C*-(*O*-hexosyl)-glucoside*

(*continued on next page*)

Table 1 (*continued*)

HPLC-MS t_R (min)	HPLC-UV Peak no.	$[M-H]^-$	MS/MS fragments $[M-H]^-$	$[M+H]^+$	MS/MS fragments $[M+H]^+$	Putative assignment
			575.1356		421.0920	
			545.1274		355.0816	
			515.1167		**207.0655**	
			383.0655		175.0393	
			353.0637			
50.17		447.0921	285.0390	449.1081	287.0540	luteolin *O*-glucoside
50.47		931.2467	841.2133	933.2657	589.1555	apigenin 6-*C*-(O-sinapoyl)-pentosyl−8-*C*-(*O*-hexosyl)-glucoside*
			769.1923		571.1451	
			605.1455		427.1023	
			545.1261		409.0921	
			515.1156		379.0821	
			383.0761		**207.0655**	
			353.0641			
50.69		431.0971	269.0457	433.1128	271.0591	apigenin *O*-glucoside
50.72		447.0903	300.0251	449.1073	303.0500	quercetin 3-O-deoxyhesoside
			301.0329			
52.54	23	799.2050	709.1717	801.2229	457.1130	apigenin 6-*C*-(*O*-sinapoyl)-glucosyl−8-*C*-glucoside isomer
			679.1599		379.0815	
			575.1360		355.0812	
			473.1046		**207.0655**	
			455.0964		175.0393	
			383.0743			
			353.0642			
53.67	24	799.2037	679.1662	801.2242	457.1135	apigenin *O*-sinapoyl−6,8-di-*C*- glucoside isomer
			593.1501		379.0815	
			575.1362		325.0705	
			473.1065		**207.0656**	
			455.0970		175.0394	
			383.0751			
			353.0642			
53.89		739.1844	649.1550	741.2039	573.1423	apigenin *O*-*p*-coumaryl−6,8-di-*C*- glucoside isomer*
			619.1389		457.1141	
			593.1478		379.0816	
			575.1363		355.0815	
			473.1062		325.0706	
			455.0964		**147.0444**	
			383.0742			
			353.0644			
54.31	25	769.1955	679.1581	771.2140	457.1133	apigenin *O*-feruloyl−6,8-di-*C*-glucoside isomer*
			649.1519		379.0819	
			593.1488		355.0827	
			575.1372		325.0707	
			473.1065		**177.0550**	
			451.3595		145.0289	
			383.0746			
			353.0636			
55.37		799.2056	709.1732	801.2253	457.1131	apigenin 6-*C*-(O-sinapoyl)-glucosyl−8-C-glucoside isomer
			679.1655		439.1017	
			575.1356		379.0812	
			473.1063		355.0809	
			455.0946		325.0706	
			413.0855		**207.0655**	
			383.0743		175.0393	
			353.0641			
55.96	26	769.1958	709.1711	771.2147	457.1136	apigenin *O*-feruloyl−6-*C*-glucosyl−8-C-glucoside isomer
			679.1629		409.0920	
			593.1484		379.0814	
			575.1376		325.0708	
			545.1266		207.0655	
			473.1033		**177.0550**	
			455.0928		145.0283	
			383.0742			
			353.0640			
56.81		769.1945	679.1667	771.2130	663.16.10	apigenin *O*-sinapoyl−6-*C*-xylosyl−8-*C*-glucoside isomer
			563.1372		427.1030	
			545.1257		379.0816	
			455.0953		337.0697	
			443.0952		325.0710	
			383.0743		**207.0655**	
			353.0638		175.0393	
56.89	27	739.1861	679.1605	741.2028	427.1028	apigenin 6-*C*-glucosyl-(O-feruloyl)−8-*C*-xyloside isomer
			649.1523		409.0926	
			545.1239		379.0814	
			455.0958		325.0697	

(*continued on next page*)

Table 1 (*continued*)

HPLC-MS t_R (min)	HPLC-UV Peak no.	$[M-H]^-$	MS/MS fragments $[M-H]^-$	$[M+H]^+$	MS/MS fragments $[M+H]^+$	Putative assignment
			443.0956		**177.0550**	
			383.0751		145.0288	
			353.0639			
57.30	28	697.1745	575.1387	699.1928	585.1382	apigenin−6-*C*-(O-benzoyl) glucosyl−8-C-glucoside (benzoyl vicenin−2)*
			473.1037		531.1309	
			455.0976		457.1151	
			383.0753		379.0811	
			353.0643		325.0708	
					105.0341	
57.63		739.1861	679.1580	741.2028	427.1030	apigenin *O*-feruloyl−6-*C*-xylosyl−8-*C*-glucoside isomer
			649.1570		409.0935	
			563.1373		379.0819	
			545.1249		325.0713	
			455.0972		**177.0549**	
			443.0959		145.0288	
			383.0741			
			353.0642			
58.41	29	769.1951	679.1564	771.2140	663.1725	apigenin *O*-sinapoyl−6-*C*-xylosyl−8-*C*-glucoside isomer
			563.1414		493.1137	
			545.1277		421.0924	
			473.1086		379.0812	
			443.0938		337.0712	
			413.0851		**207.0655**	
			383.0750		175.0392	
			353.0641			
59.86	30	667.1646	577.1309	669.1814	615.1508	apigenin−6-*C*-(O-benzoyl)-pentosyl−8-C-glucoside*
			545.1262		531.1297	
			443.0963		427.1026	
			455.0927		409.0925	
			383.0738		391.0815	
			535.0637		379.0817	
					337.0709	
					325.0714	
					105.0341	
61.87	31	723.1898	633.1552	725.2062	653.1653	apigenin−6-*C*-(O-cinnamoyl)-hexosyl−8-C-glucoside*
			603.1465		557.1422	
			575.1392		457.1140	
			473.1052		379.0818	
			455.0931		325.0709	
			383.0745		**131.0496**	
			353.0644		103.0549	
63.68	32	1005.2613	**799.2021**	1007.2765	729.1807	apigenin 6-*C*-(*O*-sinapoyl)-glucosyl−8-*C*-(O-sinapoyl)-glucoside isomer
			781.1943		663.1709	
			709.1700		523.1234	
			679.1628		457.1133	
			485.1032		421.0916	
			473.1042		379.0816	
			455.0953		355.0811	
			383.0743		**207.1654**	
			353.0636		175.0392	
64.83	33	975.2502	885.2774	977.3281	797.2621	apigenin 6-*C*-(*O*-feruloyl)-glucosyl−8-*C*-(O-sinapoyl)-glucoside isomer
			799.2094		699.1695	
			795.2451		523.1249	
			769.1926		457.1125	
			705.2163		421.0926	
			473.1060		379.0814	
			443.0956		355.0817	
			383.0747		**207.0655**	
			353.0641		**177.0550**	

* One of the possible isomeric structures has been indicated – n.d. = not detected n.f. = not fragmented.

$[(M-H) - 60]^-$ is more intense if the pentose is attached to the 6 position compared to that of the isomer carrying the pentose at the 8 position, thus allowing discrimination of the isomers through comparison of the MS/MS spectra (Bucar et al., 2021). However, the differences in fragmentation patterns may depend on the type of instrument, collision energies and experimental parameters (Abad-García et al., 2008). In our conditions, with steeping collision energy, we did not observe significant and stable differences of intensity between the $[(M-H) - 60]^-$ fragments of the isomeric species with $[M-H]^- = 563.1378$. Thus, these isomers eluting in the HPLC-UV chromatogram at t_R 26.1 and 27.7 min, were assigned to vicenin-1 (apigenin-6-*C*-β-D-xyloside-8-*C*-β-D-glucoside) and vicenin-3 (apigenin-6-*C*-β-D-glucoside-8-*C*-β-D-xyloside), respectively, based on their previously reported elution order (Wang et al., 2019). Notably, the MS/MS spectra of vicenin-2 and its apigenin 6, 8-di-*C*-hexoside did not exhibit detectable $[(M-H) - 60]^-$ fragment signals.

3.4. Flavone di-C-glycosides and mono-C-glycosides

As mentioned above, the most abundant glycoside of TSG was

identified as vicenin-2, while other relatively abundant compounds were vicenin-1 and vicenin-3. Isoschaftoside, which is isobaric with vicenin-1 and vicenin-3, was also detected at minor amounts and confirmed across the chromatographic t_R by comparison with the authentic standard. A compound corresponding to apigenin 6,8-di-*C*-pentoside ($[M-H]^-$ = 533.1269) was detected at low intensity in HPLC-UV (Fig. 1), in line with a previous characterization of this compound in *Viola yedoensis* validated through ion trap multi-collisional MS^n spectra (Cao et al., 2014). The HPLC-MS/MS analysis highlighted the presence of at least two low-abundance apigenin di-*C*-pentoside isomers. The MS/MS spectra of these compounds exhibited a prominent [(M-H) - 60]$^-$ fragment (**Supporting information, Figure S3**). Two minor isobaric components with measured $[M-H]^-$ = 577.1538 eluting closely to apigenin di-*C*-pentoside were assigned to the apigenin 6-*C*-deoxyhexose-8-*C*-hexose/apigenin 6-*C*-hexose-8-*C*-deoxyhexose couple (**Supporting information, Figure S4**), based on the characteristic [(M-H) - 74]$^-$ = 503.1161 and [(M-H) - 104]$^-$ = 473.1053 fragments (Kachlicki et al., 2016).

Low abundant apigenin mono-*C*-glycoside isomers ($[M-H]^-$ = 431.0958), most likely vitexin (apigenin-8-*C*-β-D-glucoside) and isovitexin (apigenin-6-*C*-β-D-glucoside), were detected at higher retention times compared to di-*C*-glycosides, alongside trace amounts of apigenin and luteolin *O*-glycoside. While the mono-*C*-glycosides exhibited the diagnostic $[Agl+41]^-$ and $[Agl+71]^-$ pair of fragments, apigenin *O*-hexoside was easily assigned due to the neutral loss of 162 uma that generated the prominent Y_0 Agl fragment ($[M-H]^-$ = 269.0457).

3.5. Complex flavone C-glycosides and mixed di-C,O-glycosides

Both mixed *O*-glycosyl-*C*-glycosyl flavonoid derivatives and multi-glycosylated compounds in which a glycosyl or other complex oligosaccharides are linked to any hydroxy group of a *C*-glycoside moiety have been well described (Talhi and Silva, 2012). The HPLC-UV chromatogram of TSG contained a variegate series of compounds more polar than vicenin-2, eluting in the t_R range 21–23 min and exhibiting UV-spectra overlapping with those of apigenin di-*C*-glycosides. Two isomeric compounds among these, one majority (Fig. 2, t_R 31.81 min) and the other occurring at trace level (Fig. 2, t_R 33.90 min), had $[M-H]^-$ = 755.2004, thus corresponding to vicenin-2 further glycosylated with an hexose moiety, also compatible with the increased polarity of the molecule. The MS/MS spectrum of the majority compound showed the prominent fragment $[M-H]^-$ = 593.1450, corresponding to vicenin-2, that was missing in the MS/MS spectrum of the minority one (**Supporting information, Figure S5**). These MS/MS spectra are suggestive of two distinct glycosidic structures, the majority one corresponding to a mixed *O*-glycosyl-di-*C*-glycoside apigenin derivative and the minority one to a *C*-(*O*-glycosyl)-glycoside. Interestingly, only one of the two possible isomers carrying the di-hexose moiety alternatively at the *C*-6 or *C*-8 positions of the apigenin skeleton was detected. In the positive ion mode, mixed di-*C*,*O*-glycosides undergo simultaneous fragmentation according to the typical pathways of both *C*- and *O*-glycosydes, but the fragments produced by the interglycosidic cleavage of *C*-glycosyl rings overcome those generated by the *O*-glycosides, so complicating their assignment (Kachlicki et al., 2016).

The extraction of the ion $[M-H]^-$ = 725.1895 from the HPLC-MS/MS chromatogram showed the presence of at least four closely eluting different peaks, among which the two majority ones were mixed apigenin *O*-glycosyl-di-*C*-glycosides. In particular, based on the MS/MS spectrum, these compounds were tentatively assigned to the *O*-glycosyl derivative of vicenin-1 and vicenin-3. The early eluting isomer of this pair was assigned to *O*-glycosyl-vicenin-1 due to the simultaneous presence of the fragment at $[M-H]^-$ = 563.1380, which is suggestive of a neutral loss of a hexoside according to the decay pathway typical for *O*-glycosides (Kachlicki et al., 2016), and $[M-H]^-$ = 665.1733, which corresponds to the neutral loss of 60 uma suggestive of a pentose unit at the *C*-6 position (**Supporting information, Figure S6**). The MS/MS spectrum of the lower-abundance isomer with $[M-H]^-$ = 725.1895 lacked the signal at $[M-H]^-$ = 563.1380, so pointing to a *C*-(*O*-glycosyl)-vicenin-1/vicenin-3 derivative. Second order fragmentation MS spectra do not allow to establish the *O*-hexosyl position on the trihydroxyflavone skeleton of mixed *O*-glycosyl-di-*C*-glycosides nor the nature of the hexose sugar. However, previous characterizations based on NMR structural analysis have demonstrated that the 4' OH is the most frequently derivatized in apigenin mixed *O*-glycosyl-di-*C*-glycoside (Kachlicki et al., 2016). For this reason, in Table 1 the dominant species with $[M-H]^-$ = 725.1895 have been tentatively assigned to vicenin-1/vicenin-3 4'-*O*-glucoside. Although the relatively intense signals $[M-H]^-$ = 665.2792 and $[M+H]^+$ = 667.2969 could be assigned to an apigenin 6-*C*-(*O*-pentosyl)-pentosyl-8-*C*-pentosyl isomer derivative based on mass values, this assignment does not match the MS/MS patterns of flavone di-*C*-glycosides derivatives and this compound is reported as unknown in Table 1.

3.6. Acylated flavone di-C-glycosides

Acylation of flavonoid glycosides with aliphatic (*e.g.*, acetic or malonic) or aromatic (*e.g.*, hydroxycinnamic acids) carboxylic acids is a frequently observed structural feature. Despite the proximity on the nominal mass of some acylating groups with sugar residues (*e.g.*, caffeic acid/hexose, *p*-coumaric acid/and deoxyexose), acylated species are straightforwardly distinguished from simple flavonoid glycoconjugates based on high-resolution MS, longer HPLC t_R, and UV-Vis spectra, which are modified by the presence of aromatic chromophores (Wang et al., 2019), as evident by the UV-Vis spectrum of peak no. 24 of the TSG chromatogram in the inset of Fig. 1. The placement of the acyl groups cannot be univocally established based on these techniques alone. In line with similar compounds characterized from other plant sources, likely acylation occurs at the OH group of a sugar moiety, most often at the 6''-OH position, although the direct acylation of any aglycone -OH to yield *O*-acylated-*C*-glycosylflavones is rather frequent as well (Talhi and MS Silva, 2012). In Table 1, acylated derivatives that showed a MS/MS fragment corresponding to the "basic" structure of flavone di-*C*-glycosides (e.g. MS/MS fragment $[M-H]^-$ = 593 for vicenin-2) have been assigned to *O*-acylated-di-*C*-glycosylflavones, while those in which the fragment of the basic structure was missing have been assigned to any flavone 6-(*O*-acyl)glycosyl-8-*C*-glycoside isomer. Though not definitive, those assignments seem to be the most reasonable ones. At variance from CSG in which the feruloyl derivatives prevail, the most abundant acylated flavone di-*C*-glycosides in TSG extracts are the sinapoyl derivatives, such as that of vicenin-2 (HPLC-UV t_R 35.2 min, $[M-H]^-$ = 799.2075) identified by the mass increment of 206 uma compared to the non-acylated counterparts. Isomeric sinapoyl derivatives of vicenin-1 (t_R 35.8 min) and vicenin-3 (t_R 37.0) min were differentiated from the much less intense isobaric feruloyl derivatives of vicenin-2 based on the MS/MS fragmentation either in the negative or in the positive ion mode. More in detail, in the negative ion mode the feruloyl vicenin-2 was distinguished from isobaric sinapoyl vicenin-1/vicenin-3 by fragments corresponding to the neutral loss of the acylating groups (**Supporting information, Figure S7**). Remarkably, prominent signals in the positive ion mode MS/MS spectra corresponding to the fragmentation of the ester bond, *i.e.*, 147, 177 and 207, were diagnostic for the acylating *p*-coumaryl, feruloyl and sinapoyl moieties, respectively. The use of polarity switching MS analysis was particularly effective for distinguishing between possible isomeric acylated derivatives and to confirm their structural difference from multi-glycosylated isobaric compounds. Two low-abundance isobaric signals at $[M-H]^-$ = 739.1868 were assigned to *p*-cumaroyl vicenin-2 and to feruloyl vicenin-1/vicenin-3, this latter also isobaric with feruloyl derivative of (iso)schaftoside, and distinguished due to the neutral loss of the acylating groups (**Supporting information, Figure S8**). Traces of early eluting malonyl-derivative of vicenin-2 were putatively assigned based on the matching with previously reported MS/MS transitions (Ferreres

et al., 2011). The relatively intense signals with $[M-H]^-$ = 697.1746 and $[M+H]^+$ = 699.1929 were putatively assigned to the benzoylated derivate of vicenin-2 (expected $[M-H]^-$ = 697.1768 and $[M+H]^+$ = 699.1925), also based on the intense MS/MS fragment at $[M+H]^+$ = 105.0341 (**Supporting information, Figure S9**). Benzoylated glycosides have been described in several plant extracts, but, to the best of our knowledge, this is the first time that they are putatively assigned for *C*-glycosyl flavones and were missing in the CSG extracts. Similarly, low-abundance signals at $[M-H]^-$ = 723.1898 and $[M+H]^+$ = 725.2062 were assigned to the cinnamoyl derivative of vicenin-2 based on the intense MS/MS fragment at $[M+H]^+$ = 131.0496.

The caffeoyl derivative of glycosyl-vicenin-2 was barely detected with $[M-H]^-$ = 917.2391 and $[M+H]^+$ = 919.2498, and assigned thanks to the MS/MS fragment at $[M+H]^+$ = 163.0392.

Interestingly, double acylated derivatives, such as di-sinapoyl vicenin-2 ($[M-H]^-$ = 1005.2609) and di-sinapoyl vicenin-1/vicenin-3 ($[M-H]^-$ = 975.2502) were detected by HPLC-MS/MS at longer t_R than the mono-acylated species. The sinapoyl-feruloyl derivative of vicenin-2 was differentiated from the possible isobaric di-sinapoyl derivative of vicenin-1/vicenin-3 based on the MS/MS spectrum that exhibited simultaneously the fragments at $[M-H]^-$ = 799.2094 and $[M-H]^-$ = 769.1926 corresponding to the alternative neutral loss of the feruloyl or sinapoyl moieties, respectively (**Supporting information, Figure S10**), also confirmed by the two prominent MS/MS fragmentation ions $[M+H]^+$ = 177.0550 and $[M+H]^+$ = 207.0655 in the positive ion mode. Overall, these data demonstrate the occurrence of a large variety of glycoconjugates in TSG at variable abundance.

3.7. Putative structural assignment of non-flavonoid phenolic glycosides

The HPLC-UV components eluting at t_R 19.4 and 22.9 min, relatively more prominent in CSG than in TSG, were detected only at 280 nm and exhibited UV spectra with sharp absorption bands centred at λ_{MAX} = 267 and 278 nm, respectively. These compounds were collected and individually analysed by HPLC-MS/MS and correlated to the corresponding peaks in the HPLC-MS chromatogram. They were better detected in the negative ion mode and had measured $[M-H]^-$ = 443.1891 and $[M-H]^-$ = 459.1481 differing by an oxygen atom. As a result of the fragmentation, both molecules were subject to a neutral loss of 294 with an intermediate neutral loss of 132, suggesting the presence of *O*-hexose-pentose di-glycoside moieties linked to the aglycones with molecular weights of $[M-H]^-$ = 149.0583 and $[M-H]^-$ = 165.0531 (**Supporting information, Figure S11**), respectively. The calculated molecular formulas for these two aglycones were $C_9H_{10}O_2/C_9H_{10}O_3$, or less likely $C_{10}H_{14}O/C_{10}H_{14}O_2$, which, in this latter case, would correspond to possible phenolic terpenes such as thymol or carvacrol. The whole TSG ethyl acetate extract subjected to hydrolysis did not contain detectable amount of thymol and carvacrol, as assessed by GC-MS and GC-FID in comparison with the analysis of authentic standards. In the attempt to elucidate the structure of these compounds and to support the HPLC-MS based hypothesis, the HPLC peak of the compound eluting at t_R 19.4, which is more abundant and better resolved than the one eluting at t_R 22.9, was collected and subjected to acidic hydrolysis. The GC-MS analysis of the hydrolyzed compound showed a prominent radical cation $M^{\cdot+}$ = 150 and a fragmentation spectrum matching 4-vinyl guaiacol (4-VG) with an estimated > 90 % probability score (Fig. 4). The UV spectrum of this compound is compatible with that of 4-VG exhibiting an absorption band with λ_{MAX} at 270 nm and a small shoulder around 300 nm (Luo et al., 2021). Considered the higher chromatographic retention in reversed phase HPLC, the compounds eluting at t_R 22.9 has been putatively assigned to the epoxide derivative of 4-VG. Differently from the 4-VG aglycone, which is a metabolite of the ferulic acid found in beer and in human gut as a product of the biotransformation by microbiota (Luo et al., 2021), the glycoside derivatives of 4-VG have not been described yet to the best of our knowledge. However, our assignments of the compounds $[M-H]^-$ = 443.1891 and $[M-H]^-$ = 459.1481 as diglycosides of 4-VG and related epoxide, respectively, are not definitive and await opportune confirmation with NMR-based methods. It would be interesting to definitely elucidate the structure of these compounds and quantify them, as either 4-VG or phenolic terpenes seem endowed with presumed health promoting properties such as preventive effects on colon cancer (Luo et al., 2021) or cardiovascular diseases (Garg and Sahu, 2024), respectively. Notably, the active aglycones could be released in the gastrointestinal tract upon hydrolysis of the *O*-glycoside bond.

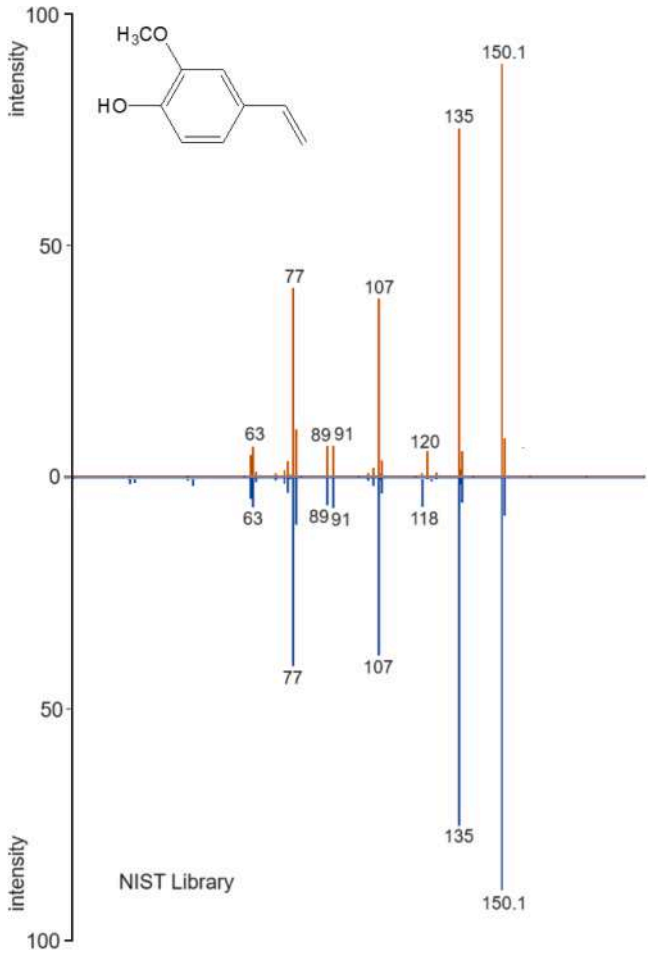

Fig. 4. Experimental and expected electron ionization MS spectrum of the aglycone released by acidic hydrolysis and putatively identified as 4-methyl guaiacol.

3.8. Other phenolic compounds

The two isobaric compounds with $[M-H]^-$ = 647.0873 that eluted close to each other in the proximity of vicenin-1 in the HPLC-MS chromatogram were assigned to tri-*O*-galloylquinic acid isomers, based on the MS/MS spectra with stepwise neutral losses of gallic acid (152 uma) (**Supporting information, Figure S12**) and on the intense MS/MS fragment ion $[M+H]^+$ = 153.0185. Likely, the most abundant among the two isomers is 3,4,5-tri-*O*-galloylquinic acid, already described in other plants (Moore et al., 2005). Similarly, two early eluting isobaric compounds with $[M-H]^-$ = 495.0768 were assigned to di-*O*-galloylquinic acid isomers (**Supporting information, Figure S13**). The HPLC-UV peak eluting at t_R = 29.8 min exhibited the typical UV spectrum of quercetin glycosides with absorption max at 257 and 356 nm. This compound was identified by HPLC-MS/MS as rutin (quercetin 3-*O*-rutinoside, [$(M-H)^-$ = 609.1445]) and confirmed with the comparative analysis of the authentic standard. Based on the negative ion mode

MS/MS spectra, rutin was distinguished from two less abundant and early eluting isobaric compounds, which were assigned to luteolin di-*C*-glucoside isomers based on the diagnostic signals $[Agl+83]^+ = 369.0590$ and $[Agl+113]^+ = 399.0695$ (**Supporting information, Figure S13**). Traces of quercetin 3-*O*-hexoside and kaempferol 3-*O*-hexoside were also detected by HPLC-MS/MS at longer t_R than rutin (**Supporting information, Figure S14**).

4. Discussion

The analytical strategy based on the combination of HPLC-DAD and nanoflow HPLC-MS/MS in polarity switching mode with stepped collision energy enabled the comprehensive putative assignment of polyphenols in TSG. They predominantly were flavone di-*C*-glycosides and their further glycosylated and/or acylated derivatives.

While seemingly confined to a limited number of plant species until a few decades ago, flavonoid *C*-glycosides are emerging as an important class of plant-derived bioactive compounds (Kong, 2012). Currently, they have been described in more than 180 plant species belonging to almost 40 families, also including very popular cereal crops, legumes, and medicinal herbs (Wang et al., 2020). *C*-glycoside derivatives of several aglycones, especially with polyketide structures, are biosynthesized as secondary metabolites by microbes and insects as well (Hultin, 2005). *C*-glycosylated phenolics have been detected in many plant compartments, although they appear particularly represented in the embryo of some species, such as *Triticaceae* and other cereals (Brazier-Hicks et al., 2009; Geng et al., 2016). Overall, more than 750 different *C*-glycoside derivatives have been characterized in plants, most of which are flavone *C*-glycosides (Zhang, 2022).

Citrus fruits have been reported as major food sources of flavonoid *C*-glycosides, with 2.9–15.4 mg 100 g^{-1} of fresh orange, mandarin, bergamot, and lemon juices. Their concentration in commercial fruit orange and pink grapefruit (2.9–4.5 mg 100 mL^{-1}) juices might be lower (Bucar et al., 2021), while they could be slightly more abundant in pericarp of some citrus species (flavedo) (Siano et al., 2023 – c). Considered the estimated total polyphenol amount of 14.9 g_{GAE} kg^{-1} and taken into account that a relevant part of them is represented by gallic acid, TSG is a much richer source of flavonoid *C*-glycosides as it contains up to 20–80 times higher amount than citrus fruits.

The biosynthetic pathways yielding flavonoid *C*-glycosides in higher plants have been recently investigated in detail (Wang et al., 2020) and an increasing number of *C*-glycosyltransferases that exhibit varying catalytic mechanisms are emerging from these studies (Zhang, 2022). Clearly, *C*-glycosylation of flavonoids provides carbohydrate mimetics with improved stability toward glycosidases and hydrolytic conditions compared to aglycones or *O*-glycosylated flavonoids (Rawat et al., 2009).

On the other hand, the physiological role of these compounds still remains in the field of hypothesis. Why do they accumulate at particularly abundant concentration in the germ of some plants, such as some *Leguminosae* species that thrive of semi-arid soils? Are they phytochemical modulators of the quiescence-germination cycle of the seeds or are they deputed to the defence of the germinating part of the seeds against phytophagic insects or microbial phytopathogens?

It is plausible that the effects of these compounds on human health, such as the inhibition of α-amylase, are related to the physiological role that they have in the plants and appear to possess in many *in vitro* and *in vivo* contexts.

In general, flavonoid *C*-glycosides show an ample range of bioactive effects *in vitro*, including antioxidant, anticancer, anti-inflammatory, hepatoprotective, antiviral, antibacterial, anti-diabetes, and antifungal activity, which might be related to their physiological role in the plants. Flavone *C*-glycosides have been described as hypoglycaemic natural compounds since they inhibit the activity of digestive α-amylase and α-glycosidases. To this purpose, seeds of *Trigonella foenum-graecum* (*Fabaceae* family) also known as fenugreek, are rich sources of flavone *C*-glycosides and have well described potential in lowering fasting blood sugar levels and improving glucose tolerance (Shabil et al., 2023; Singh et al., 2023).

Durum wheat pasta samples fortified with 5–10 % CSG delayed the release of reducing sugars from starch during simulated gastrointestinal digestion (Siano et al., 2023 – a). The inhibiting activity demonstrated *in vitro* has high chance to occur *in vivo* as well because it is exerted locally in the gastrointestinal tract. However, the inhibiting activity of flavonoids *C*-glycosides depends on the structure and dimension of glycoside moieties (Li et al., 2009). In the current study, we did not detect any significant inhibition activity of TSG extracts against α-amylase and α-glycosidases comparable to CSG, probably due to the relatively high content of large glycoconjugates, or, more likely, due to a lower affinity of these enzymes for vicenins than for (iso)schaftosides. However, high concentrations of phenolic compounds can decrease the *in vitro* α-amylase and α-glycosidases inhibition activity (Yang and Kong, 2016; Cisneros-Yupanqui et al., 2023).

Regarding the probability that flavone *C*-glycosides molecules carry out systemic actions, while flavonoid *C*-monoglycosides are poorly absorbed in humans and are for the most deglycosylated and degraded by microflora in the colon (Mori et al., 2021), flavonoid *C*-multiglycosides appear at least in part absorbed unchanged and distributed through blood circulation (Xiao et al., 2016). Thus, the antidiabetic properties of flavone *C*-glycosides could also occur through mechanisms different than the delay of carbohydrate digesting enzymes action, such as the inhibition of regulators of the insulin signalling pathways (Rampadarath et al., 2022).

Other health-promoting effects are potentially exerted by the major phenolic compounds of TSG, as vicenin-2 has been described as an hepatoprotective and potential anti-inflammatory agent based on *in vivo* assays (Marrassini et al., 2011) and vicenin-1 inhibits platelet aggregation with no appreciable *in vivo* toxicity in mice even at high dosage (Kandhare et al., 2016). Vicenins are angiotensin-converting enzyme inhibitors and can contribute to lower blood pressure (Zhang et al., 2015). Particularly interesting are the anti-glycation properties of vicenin-2, which could contrast the effects of diabetes and metabolic disorders (Islam et al., 2014). However, data on bioavailability, kinetics in human body and *in vivo* health-promoting effects of flavonoid *C*-glycosides are still limited and require further assessment. Moreover, tri-*O*-galloylquinic acid is classified as a natural molecule with anti-HIV (Kamng'ona et al., 2011) and DNA polymerase inhibitor (Parker et al., 1989) properties.

5. Conclusion

In this work, the phenolic compounds of TSG have been comprehensively characterized for the first time. The findings highlight the occurrence of a complex array of flavone *C*-glycosides. In the light of the high abundance of phenolic compounds endowed with potential health-promoting properties, as well as other compositional traits such as the large quantity of plant-derived proteins with good biological value, TSG would meet the requirements to be considered as a functional food ingredient. In fact, TSG had already been introduced in the North American market as a superfood ingredient. Nevertheless, all the hints have pointed out to TSG as the causative agent of the outbreak of foodborne diseases recorded in USA in 2022, despite it has been used for both animal and human consumption in the past. The phenolic profile of TSG investigated in this study closely matches that of "toxic" TSG published by Chittiboyina et al. (2023), confirming that the plant species responsible for the health injury actually is *C. spinosa*. The dominant presence of flavone *C*-glycoside derivatives already identified in many largely consumed foods allows to reasonably exclude that the alleged harmful component of TSG belongs to the class of phenolic compounds. The profile of phenolic compounds should be investigated across a higher number of samples and production batches. However, the availability of TSG is limited since tara is grown only in a few regions of

the world and many food companies have discontinued supplying flour.

TSG is an agro-industrial co-product, which could support the economy of producers and growers of marginalized areas. The high level and diversity of phenolic compounds and derived glycoconjugates suggest the use of TSG also as a nutraceutical, although there remains an urgent need of systematic investigations to definitely assess the toxicological profile of TSG and the risk associated with TSG consumption. In the meanwhile, the characterization of the phenolic profile of TSG can offer a valuable analytical reference for its monitoring.

Funding

This work was supported by the ALIFUN (PON MUR 2018; ARS01_00783) and by the NutrAge FOE projects of Italian National Research Council (CNR).

CRediT authorship contribution statement

Ermanno Vasca: Writing – review & editing, Supervision, Formal analysis, Data curation. **Olga Fierro:** Methodology, Investigation, Formal analysis, Data curation. **GIANLUCA PICARIELLO:** Writing – original draft, Funding acquisition, Formal analysis, Conceptualization. **Francesco Siano:** Writing – review & editing, Software, Methodology, Formal analysis, Data curation.

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.

Acknowledgement

The authors gratefully thank Silvateam SpA (San Michele Mondovì, Cuneo, Italy) for providing tara seed germ flour and Carob Ingredients Co., Casablanca, Morocco for providing carob seed germ flour.

G. P. wishes to thank Prof. Ikhlas Khan and Dr. Bharathi Avula (The University of Mississippi, USA) for discussion and advice.

Associated content

Supporting Information.

Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at doi:10.1016/j.jfca.2024.107145.

Data availability

Data will be made available on request.

References

Abad-García, B., Garmón-Lobato, S., Berrueta, L.A., Gallo, B., Vicente, F., 2008. New features on the fragmentation and differentiation of C-glycosidic flavone isomers by positive electrospray ionization and triple quadrupole mass spectrometry. Rapid Commun. Mass Spectrom. 22 (12), 1834–1842.

Balaguer, L., Arroyo, R., Jiménez, P., Jiménez, M.D., Villegas, L., Cordero, Rubio de Casas, I.R., Fernández-Delgado, R., Ron, M.E., Manrique, E., Vargas, P., Cano, E., Pueyo, J.J., Aronson, J., 2011. Forest restoration in a fog oasis: Evidence indicates need for cultural awareness in constructing the reference. PLoS One 6, 1–10.

Borzelleca, J.F., Ladu, B.N., Senti, F.R., Egle, J.L.Jr, 1993. Evaluation of the safety of tara gum as a food ingredient: A review of the literature. J. Am. Coll. Toxicol. 12 (1), 81–89.

Brazier-Hicks, M., Evans, K.M., Gershater, M.C., Puschmann, H., Steel, P.G., Edwards, R., 2009. The C-glycosylation of flavonoids in cereals. J. Biol. Chem. 284, 17926–17934.

Bucar, F., Xiao, J., Ochensberger, S., 2021. Flavonoid C-glycosides in diets. Handb. Diet. Phytochem. 117–153.

Cao, J., Yin, C., Qin, Y., Cheng, Z., Chen, D., 2014. Approach to the study of flavone di-C-glycosides by high performance liquid chromatography-tandem ion trap mass spectrometry and its application to characterization of flavonoid composition in Viola yedoensis. J. Mass Spectrom. 49 (10), 1010–1024.

Chittiboyina, A.G., Ali, Z., Avula, B., Khan, S.I., Mir, T.M., Zhang, J., Aydoğan, F., Zulfiqar, F., Techen, N., Parveen, I., Pandey, P., Adams, S.J., Wang, Y.H., Zhao, J., Marshall, G.D., Pugh, N.D., Khan, I.A., 2023. Is baikiain in tara flour a causative agent for the adverse events associated with the recalled Frozen French lentil & leek crumbles food product? - A working hypothesis. Chem. Res. Toxicol. 36 (6), 818–821.

Cisneros-Yupanqui, M., Lante, A., Mihaylova, D., Krastanov, A.I., Rizzi, C., 2023. The α-amylase and α-glucosidase inhibition capacity of grape pomace: A review. Food Bioprocess Technol. 16 (4), 691–703.

Desai, S., Prajapati, V., Chandarana, C., 2022. Chemistry, biological activities, and uses of tara gum. Chemistry, Biological Activities and Uses. In: Gums, Resins and Latexes of Plant Origin. Springer International Publishing, Cham, pp. 1–25.

Domon, B., Costello, C.E., 1988. A systematic nomenclature for carbohydrate fragmentations in FAB-MS/MS spectra of glycoconjugates. Glycoconj. J. 5, 397–409.

European Directorate for the Quality of Medicines. Council of Europe. 2007. Determination of tannins in herbal drugs. European Pharmacopoeia (6th ed.). France: Strasbourg. A286.

FDA, 2024. FDA Update on the Post-market Assessment of Tara Flour Available from: ⟨https://www.fda.gov/food/cfsan-constituent-updates/fda-update-post-market-assessment-tara-flour⟩.

Ferreres, F., Gil-Izquierdo, A., Vinholes, J., Grosso, C., Valentão, P., Andrade, P.B., 2011. Approach to the study of C-glycosyl flavones acylated with aliphatic and aromatic acids from Spergularia rubra by high-performance liquid chromatography-photodiode array detection/electrospray ionization multi-stage mass spectrometry. Rapid Commun. Mass Spectrom. 25 (6), 700–712.

Fierro, O., Siano, F., Bianco, M., Vasca, E., Picariello, G., 2024. Comprehensive molecular level characterization of protein- and polyphenol-rich tara (*Caesalpinia spinosa*) seed germ flour suggests novel hypothesis about possible accidental hazards. Food Res. Int. 181, 114119.

Garg, V., Sahu, B.D., 2024. Carvacrol and its effect on cardiovascular diseases: From molecular mechanism to pharmacological modulation. Food Biosci. 57, 103444.

Gelling, T., 2019. A chef's guide to gelling, thickening, and emulsifying agents. Boca Raton: CRC Press.

Geng, P., Sun, J., Zhang, M., Li, X., Harnly, J.M., Chen, P., 2016. Comprehensive characterization of C-glycosyl flavones in wheat (*Triticum aestivum* L.) germ using UPLC-PDA-ESI/HRMS^n and mass defect filtering. J. Mass Spectrom. 51 (10), 914–930.

Georganas, A., Maggiore, A., Bottex, B., 2024. Emerging chemical risks in food and feed. EFSA Suppl. Publ. 21 (8), 8992E.

Hultin, P.G., 2005. Bioactive C-glycosides from bacterial secondary metabolism. Curr. Top. Med. Chem. 5, 1299–1331.

Islam, M.N., Ishita, I.J., Jung, H.A., Choi, J.S., 2014. Vicenin 2 isolated from Artemisia capillaris exhibited potent anti-glycation properties. Food Chem. Toxicol. 69, 55–62.

Kachlicki, P., Piasecka, A., Stobiecki, M., Marczak, Ł., 2016. Structural Characterization of Flavonoid Glycoconjugates and Their Derivatives with Mass Spectrometric Techniques. Molecules 21 (11), 1494.

Kamng'ona, A., Moore, J.P., Lindsey, G., Brandt, W., 2011. Inhibition of HIV-1 and M-MLV reverse transcriptases by a major polyphenol (3, 4, 5 tri-O-galloylquinic acid) present in the leaves of the South African resurrection plant, Myrothamnus flabellifolia. J. Enzym. Inhib. Med. Chem. 26 (6), 843–853.

Kandhare, A.D., Bodhankar, S.L., Mohan, V., Thakurdesai, P.A., 2016. Acute and repeated doses (28 days) oral toxicity study of Vicenin-1, a flavonoid glycoside isolated from fenugreek seeds in laboratory mice. Regul. Toxicol. Pharmacol. 81, 522–531.

Kong, D.Y., 2012. In: Xu, R.;, Ye, Y.;, Zhao, W. (Eds.), Introduction to Natural Products Chemistry. Taylor & Francis CRC Press, NY.

Li, H., Song, F., Xing, J., Tsao, R., Liu, Z., Liu, S., 2009. Screening and structural characterization of α-glucosidase inhibitors from hawthorn leaf flavonoids extract by ultrafiltration LC-DAD-MSn and SORI-CID FTICR MS. J. Am. Soc. Mass Spectrom. 20, 1496–1503.

Luo, Y., Wang, C.Z., Sawadogo, R., Yuan, J., Zeng, J., Xu, M., Tan, T., Yuan, C.S., 2021. 4-Vinylguaiacol, an active metabolite of ferulic acid by enteric microbiota and probiotics, possesses significant activities against drug-resistant human colorectal cancer cells. ACS Omega 6 (7), 4551–4561.

Marrassini, C., Davicino, R., Acevedo, C., Anesini, C., Gorzalczany, S., Ferraro, G., 2011. Vicenin-2, a potential anti-inflammatory constituent of Urtica circularis. J. Nat. Prod. 74 (6), 1503–1507.

Moore, J.P., Westall, K.L., Ravenscroft, N., Farrant, J.M., Lindsey, G.G., Brandt, W.F., 2005. The predominant polyphenol in the leaves of the resurrection plant Myrothamnus flabellifolius, 3, 4, 5 tri-O-galloylquinic acid, protects membranes against desiccation and free radical-induced oxidation. Biochem. J. 385 (1), 301–308.

Mori, T., Kumano, T., He, H., Watanabe, S., Senda, M., Moriya, T., Adachi, N., Hori, S., Terashita, Y., Kawasaki, M., Hashimoto, Y., Awakawa, T., Senda, T., Abe, I., Kobayashi, M., 2021. C-Glycoside metabolism in the gut and in nature: Identification, characterization, structural analyses and distribution of C-C bond-cleaving enzymes. Nat. Commun.. 12 (1), 6294.

Mukherjee, K., Dutta, P., Badwaik, H.R., Saha, A., Das, A., Giri, T.K., 2023. Food Industry applications of Tara Gum and its modified forms. Food Hydrocoll. Hlth. 3, 100107.

Murga-Orrillo, H., Lobo, F.D., Santos Silva Amorim, R., Fernandes Silva Dionisio, L., Nuñez Bustamante, E., Chu-Koo, F.W., López, L.A., Arévalo-Hernández, C.O., Abanto-Rodriguez, C., 2023. Increased production of tara (*Caesalpinia spinosa*) by edaphoclimatic variation in the altitudinal gradient of the Peruvian Andes. Agronomy 13 (3), 646.

Orchard, C., León-Lobos, P., Ginocchio, R., 2009. Phytostabilization of massive mine wastes with native phytogenetic resources: potential for sustainable use and conservation of the native flora in north-central Chile. Cienc. Investig. Agrar 36 (3), 329–352.

Parker, W.B., Nishizawa, M., Fisher, M.H., Ye, N., Lee, K.H., Cheng, Y.C., 1989. Characterization of a novel inhibitor of human DNA polymerases: 3, 4, 5-tri-O-galloylquinic acid. Biochem. Pharmacol. 38 (21), 3759–3765.

Pedreschi, F., Saavedra, I., Bunger, A., Zuñiga, R.N., Pedreschi, R., Chirinos, R., Campos, D., Mariotti-Celis, M.S., 2018. Tara pod (*Caesalpinia spinosa*) extract mitigates neo-contaminant formation in Chilean bread preserving their sensory attributes. LWT-Food Sci. Technol. 95, 116–122.

Pedreschi, F., Matus, J., Bunger, A., Pedreschi, R., Huamán-Castilla, N.L., Mariotti Celis, M.S., 2022. Effect of the integrated addition of a red tara pods (Caesalpinia Spinosa) extract and NaCl over the neo-formed contaminants content and sensory properties of crackers. Molecules 27, 1020.

Picariello, G., Sciammaro, L., Siano, F., Volpe, M.G., Puppo, M.C., Mamone, G., 2017. Comparative analysis of *C*-glycosidic flavonoids from Prosopis spp. and Ceratonia siliqua seed germ flour. Food Res. Int. 99, 730–738.

Re-Jiménez, B.L.D., Amadó, R., 1989. Comparative study of the chemical composition of germ meals from carob, guar and Tara seeds. Food Hydrocoll. 3 (2), 149–156.

Rampadarath, A., Balogun, F.O., Pillay, C., Sabiu, S., 2022. Identification of Flavonoid *C*-Glycosides as Promising Antidiabetics Targeting Protein Tyrosine Phosphatase 1B. J. Diabetes Res., 6233217

Rawat, P., Kumar, M., Sharan, K., Chattopadhyay, N., Maurya, R., 2009. Ulmosides A and B: flavonoid 6-*C*-glycosides from Ulmus wallichiana, stimulating osteoblast differentiation assessed by alkaline phosphatase. Bioorg. Med. Chem. Lett. 19, 4684–4687.

Rojas, O., Rojas, N., Díaz, P., 2007. La tara y condiciones de reforestación en el Alto Jequetepeque, Microcuenca de San Juan-Cajamarca. Ind. Data 10, 38–46.

Sangay-Tucto, S., Duponnois, R., 2018. In: Gorawala, P., Mandhatri, S. (Eds.), Agricultural Research Updates, 22. Nova Science Publishers, Agricultural research updates, pp. 189–208.

Shabil, M., Bushi, G., Bodige, P.K., Maradi, P.S., Patra, B.P., Padhi, B.K., Khubchandani, J., 2023. Effect of fenugreek on hyperglycemia: A systematic review and meta-analysis. Medicina 59 (2), 248.

Siano, F., Mamone, G., Vasca, E., Puppo, M.C., Picariello, G., 2023. a. Pasta fortified with *C*-glycosides-rich carob (*Ceratonia siliqua* L.) seed germ flour: Inhibitory activity against carbohydrate digesting enzymes. Food Res. Int. 70, 112962.

Siano, F., Sammarco, A.S., Fierro, O., Castaldo, D., Caruso, T., Picariello, G., Vasca, E., 2023. b. Insights into the Structure-Capacity of Food Antioxidant Compounds Assessed Using Coulometry. Antioxidants 12 (11), 1963.

Siano, F., Picariello, G., Castaldo, D., Cautela, D., Caruso, T., Vasca, E., 2023. c. Monitoring antioxidants by coulometry: Quantitative assessment of the strikingly high antioxidant capacity of bergamot *(Citrus bergamia* R.) by-products. Talanta 251, 123765.

Singh, S., Chaurasia, P.K., Bharati, S.L., 2023. Hypoglycemic and hypocholesterolemic properties of Fenugreek: A comprehensive assessment. Appl. Food Res. 3 (2), 100311.

Talhi, O., MS Silva, A., 2012. Advances in C-glycosylflavonoid research. Curr. Org. Chem. 16 (7), 859-896. Villanueva, C., 2007. La Tara, El oro verde de los Incas. Lima. Ed. AGRUM. 1° edición. Universidad Nacional Agraria La Molina. 163-167.

Vukics, V., Guttman, A., 2010. Structural characterization of flavonoid glycosides by multi-stage mass spectrometry. Mass Spectrom. Rev. 29 (1), 1–16.

Wang, Y., Liang, Z., Liao, X., Zhou, C., Xie, Z., Zhu, S., Wei, G., Huang, Y., 2019. Identification of C-glycosyl flavones by high performance liquid chromatography electrospray ionization mass spectrometry and quantification of five main C-glycosyl flavones in Flickingeria fimbriata. BMC Chem. 13 (1), 1–20.

Wang, Z.L., Gao, H.M., Wang, S., Zhang, M., Chen, K., Zang, Y.Q., Wang, H.D., Han, B.Y., Xu, L.L., Song, T.Q., Yun, C.H., 2020. Dissection of the general two-step di-*C*-glycosylation pathway for the biosynthesis of (iso)schaftosides in higher plants. PNAS 117 (48), 30816–30823.

Xiao, J., Capanoglu, E., Jassbi, A.R., Miron, A., 2016. Advance on the Flavonoid C-glycosides and Health Benefits. Crit. Rev. Food Sci. Nutr. 56, S29–S45.

Yang, X., Kong, F., 2016. Effects of tea polyphenols and different teas on pancreatic α-amylase activity *in vitro*. LWT-Food Sci. Technol. 66, 232–238.

Zhang, Y. Plant C-glycosides. Mendeley Data, V1, 2022. ⟨https://data.mendeley.com/datasets/44grnj742t/1⟩.

Zhang, Y.Q., Luo, J.G., Han, C., Xu, J.F., Kong, L.Y., 2015. Bioassay-guided preparative separation of angiotensin-converting enzyme inhibitory C-flavone glycosides from *Desmodium styracifolium* by recycling complexation high-speed counter-current chromatography. J. Pharm. Biomed. Anal. 102, 276–281.