



HAL
open science

Identification and quantification of key phytochemicals in peas – Linking compounds with sensory attributes

A. Cosson, E. Meudec, C. Ginies, A. Danel, P. Lieben, N. Descamps,
Veronique Cheynier, A. Saint-Eve, Isabelle Souchon

► To cite this version:

A. Cosson, E. Meudec, C. Ginies, A. Danel, P. Lieben, et al.. Identification and quantification of key phytochemicals in peas – Linking compounds with sensory attributes. *Food Chemistry*, 2022, 385, 10.1016/j.foodchem.2022.132615 . hal-03607206

HAL Id: hal-03607206

<https://hal.inrae.fr/hal-03607206v1>

Submitted on 13 Mar 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

1 **Title:** Identification and quantification of key phytochemicals in peas - linking compounds with
2 sensory attributes

3 **Authors:** A. Cosson^{a,b}, E. Meudec^{c,d}, C. Ginies^e, A. Danel^a, P. Lieben^a, N. Descamps^b, V.
4 Cheynier^{c,d}, A. Saint-Eve^a, I. Souchon^{e*}

5 **Affiliations:**

6 ^aUniv Paris Saclay, UMR SayFood, AgroParisTech, INRAE, F-78850 Thiverval Grignon,
7 France

8 ^bRoquette Frères, 10 rue haute loge, F-62136, Lestrem, France

9 ^cSPO, INRAE, Univ Montpellier, Institut Agro Montpellier Supagro, Montpellier, France

10 ^dINRAE, PROBE research infrastructure, Polyphenol Analytical Facility, Montpellier, France

11

12 ^eAvignon Univ, UMR SQPOV, INRAE, F-84000 Avignon, France

13 ***Corresponding author:** Isabelle Souchon, Phone: +33 (0) 1 30 81 45 31, Address:

14 UMR408 SQPOV - Sécurité et Qualité des Produits d'Origine Végétale, Domaine Saint Paul,
15 228, route de l'Aérodrome, Site Agroparc - CS 40509, 84914 Avignon Cedex 9, France

16 **Email:** isabelle.souchon@inrae.fr

17 **Email addresses for co-authors:**

18 audrey.cosson@inrae.fr

19 emmanuelle.meudec@inrae.fr

20 christian.ginies@inrae.fr

21 alice.danel@inrae.fr

22 pascale.lieben@inrae.fr

23 nicolas.descamps@roquette.com

24 veronique.cheynier@inrae.fr

25 anne.saint-eve@inrae.fr

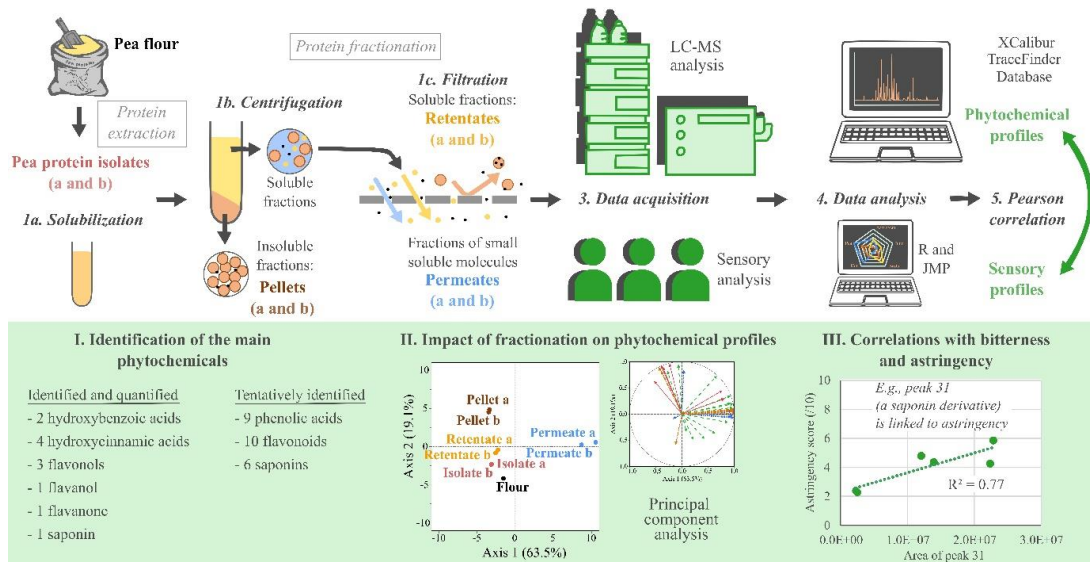
26

27 **Abstract :** Pea protein isolates contain high-quality plant protein. However, they have sensory
 28 drawbacks, notably bitterness and astringency, that have limited their use in commercial foods. This
 29 study's aim was thus to identify the main phytochemicals in pea-based samples and to examine
 30 associations with sensory attributes. The phytochemical profiles of pea flour, pea protein isolates, and
 31 pea protein isolate fractions were characterized via UHPLC-DAD-MS. Forty-eight phytochemicals
 32 were observed: 6 phenolic acids, 5 flavonoids, and 1 saponin were unambiguously identified and
 33 quantified, while another 9 phenolic acids, 10 flavonoids, and 6 saponins were tentatively identified.
 34 The impacts of protein extraction and fractionation were studied. These processes appear to have
 35 caused some compound degradation. It was found that 29 compounds were correlated with perceived
 36 bitterness and/or astringency. Therefore, these results show that certain phytochemicals can lead to
 37 negative sensory attributes in pea-protein-based products.

38 **Keywords:** Pulse, Flavonoids, Saponins, Bitterness, Astringency, Correlation

39

40 **Graphical Abstract**



41

42 **Highlights**

- 43 • 6 phenolic acids, 5 flavonoids, and 1 saponin were identified and quantified
- 44 • 9 phenolic acids, 10 flavonoids, and 6 saponins were tentatively identified
- 45 • Protein fractionation led to some compound degradation
- 46 • 29 compounds were correlated with perceived bitterness and/or astringency

47 **1. Introduction**

48 Over the past several years, the agrifood industry has taken on the challenge of developing
49 replacements for animal proteins. The latter have pronounced environmental impacts and may become
50 scarce in the future. Consequently, industry stakeholders and consumers have begun focusing on plant
51 proteins. For example, pea (*L. Pisum sativum*) protein isolates are increasingly being incorporated into
52 foods because of their functional properties, protein content, environmental sustainability, and low
53 price. However, there are sensory drawbacks to the use of plant proteins, notably when the source is
54 isolate fractions derived from raw plant materials. In particular, unpleasant olfactory and gustatory
55 sensations may arise, which presents a hurdle for the development of commercial plant-based foods.
56 Research on pea-based foods has largely focused on how volatile aroma compounds lead to perceived
57 beaniness. However, far less studied is why such products are also perceived as bitter and/or
58 astringent.

59 It has been suggested that a wide variety of phytochemicals could be responsible for the bitter and
60 astringent notes associated with pea protein isolates. Research has particularly highlighted the
61 potential role played by lipids, saponins, and phenolics. Scientists have also explored the contribution
62 of bitter lipid oxidation products formed either through enzymatic pathways or via autoxidation. Using
63 pea protein isolate fractions, recent work identified 14 lipids and lipid oxidation products that are
64 associated with greater perceived bitterness (Gläser *et al.*, 2020). Similarly, saponins have been found
65 to give rise to bitter and astringent notes. Such is namely the case for soyasaponin I and DDMP
66 saponins (2,3-dihydro-2,5-dihydroxy-6-methyl-4H-pyran-4-one) (Heng *et al.*, 2006; Reim & Rohn,
67 2015). The work by Gläser *et al.* examined soyasaponin I's influence on the overall bitter and
68 astringent notes associated with pea protein isolates by calculating dose-over-threshold factors (the
69 ratio of compound concentration to taste threshold for each test substance). For bitterness, the factor
70 value was below 1, indicating a limited sensory contribution. For astringency, the factor value was 1.8,
71 indicating a more pronounced sensory contribution. However, to date, there has been no detailed
72 research on the composition, content, or impact of other saponins on the sensory perception of pea-
73 based products.

74 Such research is similarly scarce for phenolic compounds. That said, several phenolic acids in plant-
75 based materials other than peas have been found to result in bitter and astringent notes. For example,
76 perceived bitterness and astringency appear to arise from caffeic acid and o-caffeoylquinic acid in
77 coffee (Frank *et al.*, 2006) and hydroxybenzoic and hydroxycinnamic acids in wine (Hufnagel &
78 Hofmann, 2008). Several flavonoids seem to have the same effect, such as catechin and gallic acid in
79 red wine (Robichaud & Noble, 1990) and flavan-3-ols and flavonol glycosides in cocoa (Stark *et al.*,
80 2006). A range of studies have found that the yellow pea contains at least 115 different phenolics
81 (Fahim *et al.*, 2019; Neugart *et al.*, 2015; Stanisavljevic *et al.*, 2015), mainly glycosylated flavonols,
82 although other flavonoids are also present, including flavanols, anthocyanins, and isoflavonoids
83 (Fahim *et al.*, 2019). Many kaempferol and quercetin 3-O-glycosides have also been characterized in
84 the pea (Neugart *et al.*, 2015; Stanisavljevic *et al.*, 2015). It is therefore possible that these compounds
85 also contribute to the perceived bitterness and astringency of pea protein isolates.

86 This study thus aimed to identify the key phytochemicals (phenolic acids, flavonoids, and saponins)
87 found in pea protein isolates and to explore their links with bitter and astringent sensory attributes. To
88 this end, pea flour, pea protein isolates, and pea protein isolate fractions were utilized. First, the
89 phytochemical profiles of these different sample types were characterized using ultra-high-
90 performance liquid chromatography–diode array detector–tandem mass spectrometry (UHPLC-DAD-
91 MS). Second, the impacts of processing (i.e., extraction and fractionation) on the phytochemical
92 profiles were examined. Third, the contribution of key phytochemical compounds to perceived
93 bitterness and astringency was explored.

94

95 **2. Materials and methods**

96 **2.1. Chemical reagents**

97 Based on past research, 12 phenolic compounds were chosen as standards: kaempferol-3-O-glucoside
98 (CAS 480-10-4, grade: analytical standard), soyasapogenol B (CAS 595-15-3, grade: $\geq 98\%$ HPLC),
99 caffeic acid (CAS 331-39-5, grade: $\geq 98.0\%$ HPLC), (+)-catechin (CAS 225937-10-0, grade: $\geq 98\%$
100 HPLC), *trans*-ferulic acid (CAS 537-98-4, grade: $\geq 99.0\%$ HPLC), gallic acid (CAS 149-91-7, grade:
101 97.5–102.5% titration), naringin (CAS 10236-47-2, grade: $\geq 95\%$ HPLC), *trans-p*-coumaric acid (CAS

102 501-98-4, grade: $\geq 98.0\%$ HPLC), *p*-hydroxybenzoic acid (CAS 99-96-7, grade: $\geq 99\%$ FG), quercetin-
103 3-O-glucoside (CAS 482-35-9, grade: $\geq 90\%$ HPLC), quercetin-3-O-rutinoside (CAS 207671-50-9,
104 grade: $\geq 94\%$ HPLC), and sinapic acid (CAS 530-59-6, grade: $\geq 98\%$). All 12 were purchased from
105 Sigma-Aldrich. Other reagents included acetonitrile (grade: Optima™ for HPLC-MS), water (grade:
106 Optima™ for HPLC-MS), formic acid (grade: Optima™ for HPLC-MS), and methanol (grade:
107 Optima™ for HPLC-MS); they were purchased from Thermo Fisher. Prior to performing the UHPLC-
108 DAD-MS analysis, the solutions were run through polytetrafluoroethylene (PTFE) filters (13 mm with
109 a porosity of 0.22 μm ; Fisherbrand). Stock solutions of the individual analytes and working solutions
110 were prepared and immediately placed in glass vials stored at -80°C (Vial N9: 1.5 mL, 11.6 x 32 mm;
111 MACHEREY-NAGEL).

112

113 **2.2. Pea sample types**

114 Pea flour and two commercial pea protein isolates (a and b) (protein content = nitrogen [N] content \times
115 6.25; 83% dry matter [DM] V/V) were obtained from Roquette Frères (Lestrem, France). These two
116 pea protein isolates differed because they were produced under different protein extraction and heat
117 treatment conditions.

118 Six fractions were obtained from these two isolates as fully described in Cosson *et al.* (2021): pellets a
119 and b; permeates a and b; and retentates a and b. In brief, the isolates were first dispersed in tap water
120 (4% DM content V/V) and kept under conditions of constant agitation (12 h, 3°C). Second, the
121 solutions were centrifuged (Jouan Kr4i and Sorvall Lynx 4000 [Thermo Scientific, Waltham, US];
122 6000 g, 10 min, 4°C). The supernatant was manually separated from the pellet, which was then
123 diluted with tap water (12.35% DM content V/V). Third, the supernatant was filtered with a tangential
124 filtration module (TIA, Bollene, France) equipped with two ST-3B-1812 PES Synder membranes (46-
125 mil spacer, 10-kDa MWCO) and a high-pressure diaphragm pump (Wanner Hydra-Cell G10, Wanner
126 International Ltd, Church Crookham, UK). Fourth, ultrafiltration was used to obtain 10 L of permeate;
127 the retentate was washed with 1 diavolume during diafiltration. During the latter process, the retentate
128 was at 13°C , inlet pressure (P1) was 1.5 bar, outlet retentate pressure (P2) was 1 bar, and mean
129 transmembrane pressure ($(P1 + P2)/2$) was 1.25 bar. All the fractionation steps were carried out at

130 4°C, except for membrane filtration, which was carried out at 13°C. Finally, the samples were stored
131 at -20°C in 50 mL glass flasks.

132 DM content was determined via drying (analysis system: prepASH@219). The nine sample types had
133 the following DM levels: flour—92%; isolates a and b—93.7%; pellets a and b—6.2%; retentates a
134 and b—1.7%; and permeates a and b—0.2%.

135

136 **2.3. Standard and sample preparation**

137 **2.3.1. Standard solutions**

138 For each standard, solutions of 0.1 g/L were prepared in methanol and water (70/30 [V/V]). For
139 quercetin-3-O-glucoside and quercetin-3-O-rutinoside, solubilization was promoted using 15 min of
140 sonication. Then, 0.01 g/L stock solutions were prepared by mixing each standard with methanol and
141 water (70/30 [V/V]). These stock solutions were diluted with methanol to arrive at working solutions
142 ranging in concentrations from 0.05 to 15 ng/μl for p-hydroxybenzoic acid and from 1 to 10 ng/μl for
143 the other standards. The stock and working solutions were run through a 0.22-μm PTFE filter and
144 stored in the dark at -80°C until analysis. Calibration curves were obtained by plotting the peak areas
145 of positive compound mass (see the conditions described in section 2.4.) as a function of standard
146 solution concentration using TraceFinder software (Thermo Scientific, USA). The calibration curves
147 were linear with R² values exceeding 0.99. The detection limit (DL) was 0.02 ng/μl, and the
148 quantification limit (QL) was 8.0 ng/μl. There were two exceptions: for soyasapogenol B, DL = 0.07
149 ng/g and QL = 3 ng/g, and for p-hydroxybenzoic acid, DL = 0.05 ng/g and QL = 30.0 ng/g.

150

151 **2.3.2. Sample types with higher dry matter content**

152 For the sample types with higher DM content (flour, isolates a and b, pellets a and b), 6.0 g of material
153 was placed in 150-mL glass vials (Schott vials, Dutscher, France) and extracted three times with
154 30 mL of a methanol and formic acid mixture (99/1 [V/V]). The extraction process lasted 2 h, used a
155 magnetic stirrer, and occurred at room temperature in the dark. The supernatants were then kept
156 at -20°C for 10 min before being separated via centrifugation (4,500 rpm, 20 min, 4°C; Eppendorf
157 5804R). The supernatant was evaporated down at 40°C using a vacuum concentrator (Jouan Thermo

158 Electron Corporation, RC 1022); the process took place in darkness. To remove certain precipitates
159 generated during evaporation, centrifugation was performed before the solvent evaporated entirely
160 (4,500 rpm, 20 min, 4°C). The evaporation process was then relaunched using the supernatants and
161 continued until the solvent was gone. The samples were solubilized in 1 mL of a methanol and water
162 mixture (80/20 [V/V]) and then kept at -20°C for 10 min to promote precipitation. At that point, the
163 samples were separated by centrifugation (4,500 rpm, 20 min, 4°C). The supernatants were run
164 through a 0.22- μ m PTFE filter and stored at -80°C in darkness until analysis. This series of filtration
165 and centrifugation steps was performed to obtain clear extracts that contained no precipitates. The
166 samples were prepared in triplicate.

167

168 **2.3.3. Sample types with lower dry matter content**

169 For the sample types with lower DM content (retentates a and b, permeates a and b), a similar
170 preparation process was used. Only the extraction step differed: 6.0 g of material was mixed with 1
171 mL of a methanol and formic acid mixture (99/1 [V/V]) for 10 min at room temperature in darkness.

172 **2.4. Ultra-high-performance liquid chromatography–diode array detector–tandem mass** 173 **spectrometry**

174 The samples were subject to UHPLC analysis (Ultimate 3000, Thermo Scientific, USA) using a
175 Hypersil GOLD column (100 mm x 2.1 mm x 1.9 μ m, Thermo Scientific). The mobile phase consisted
176 of (A) water + 0.1% formic acid and (B) acetonitrile + 0.1% formic acid. A gradient program with a
177 flow rate of 0.25 mL/min was used: there was 4 min of 98% (A); a linear gradient from 98% to 70% of
178 (A) over 26 min; a linear gradient from 70% to 2% of (A) over 6 min; and 9 min of 2% (A). Initial
179 gradient pressure was 280 bar. The injection volume was 5 μ L, and the injector temperature was 7°C.
180 The oven temperature was 25°C.

181 The UHPLC system was coupled with a high-resolution mass spectrometer (Q Exactive Orbitrap,
182 Thermo Scientific, USA) equipped with a heated electrospray ionization probe (HESI II, Thermo
183 Scientific, USA). The mass spectrometer was operated in both negative and positive ion modes. The
184 parameters for the ion source were as follows: sheath gas pressure = 2.4×10^5 Pa; auxiliary gas flow
185 rate = 10; sweep gas flow rate = null; spray voltage = 3 kV; capillary temperature = 300°C, S-lens

186 radio frequency = 50 V; and heater temperature = 300°C. The spectra (MS¹ and MS²) were acquired
187 using full MS¹ and full MS¹/ddMS² across a range from 85 to 1,000 amu at two resolution levels
188 (70,000 and 17,500, respectively). The system was also coupled with a diode array detector covering
189 the full range of acquisition (190–600 nm).
190 Phytochemicals were identified based on the following features: mass spectra, accurate mass,
191 characteristic fragmentation, UV spectrum, and characteristic retention time. Xcalibur (Thermo
192 Scientific, USA) was used for instrument control, data acquisition, and data analysis. Phytochemical
193 quantification was carried out by comparing the specific reconstructed ion current with the calibration
194 curves using TraceFinder software (Thermo Scientific, USA). To propose hypothetical compound
195 formulas and to identify peaks, Compound Discoverer software (Thermo Scientific, USA) was used
196 alongside several databases (PubChem, Phenol-Explorer, Flavonoid Database, Arita Database, NPASS
197 Database, KNApSAcK Database).

198

199 **2.5. Sensory analysis**

200 Pea solutions were characterized using static profiling performed by 17 trained panelists (13 women
201 and 4 men; mean age = 23 years old) as fully described in Cosson *et al.* (2021). Briefly, attribute
202 selection was carried out using a check-all-that-apply (CATA) questionnaire. The panelists were
203 trained to evaluate attribute intensity along an unstructured scale (range: 0–10) using external
204 references. The attributes were evaluated in blocks. The first block focused on attributes shaped by
205 olfactory perception in the nose (attribute block 1: pea, broth, nuts, almond, potato, and cereals). The
206 second block focused on attributes shaped by taste perception and mouthfeel (attribute block 2: salty,
207 sugar, bitter, astringent, mouthfeel, and granularity). The third block focused on attributes shaped by
208 olfactory perception in the mouth (attribute block 3: pea, broth, nuts, almond, potato, and cereals).
209 Each solution was evaluated in duplicate by the 17 panelists. In generally, they arrived at repeatable,
210 homogeneous scores, and there was no between-session drift in scoring. Additional details on attribute
211 selection, panelist training, and method characterization can be found elsewhere (Cosson *et al.*, 2021).
212 In this study, only the data for the bitter and astringent attributes perceived in mouth are discussed.

213

214 **2.6. Statistical analysis**

215 Analyses were performed using R (R Core Team, 2019) and JMP (v. 13.1.0; SAS Institute Inc., Cary,
216 SC, USA). For the inferential analyses, $\alpha = 0.05$ was the threshold for statistical significance. Principal
217 component analysis (PCA; centered reduced variables, Pearson's r) was used on a correlation matrix
218 to visually explore differences in the phytochemical profiles of the nine sample types. The
219 relationships between the phytochemical concentrations and the sensory attribute scores (i.e.,
220 bitterness and astringency) were also explored using a Pearson correlation matrix.

221

222 **3. Results and discussion**

223 **3.1. Phytochemical identification**

224 This research characterized the main soluble phytochemicals in pea flour, protein isolates, and isolate
225 fractions. The results of the UHPLC-DAD-MS analysis led to the tentative identification of 54
226 substances (Supplementary Figure 1), which are either native seed compounds or reaction products
227 formed during extraction and fractionation. By comparing the retention times and accurate MS¹ and
228 MS² data for the samples with those for the standards, 12 compounds were unambiguously identified:
229 2 hydroxybenzoic acids, 4 hydroxycinnamic acids, 3 flavonols, 2 flavanols, and 1 saponin. In addition,
230 based on the UV–vis, accurate MS¹, and MS² data; the web databases; and published research results,
231 26 compounds were tentatively identified: 9 phenolic acids, 10 flavonoids, 6 terpenoids, and 1
232 compound from another chemical family. A range of data were collected: retention times, assigned
233 identities, UV–vis absorption levels, molecular formulas, accurate masses, the main MS data, and the
234 web database(s) used in identification (Table 1).

235

236 **3.1.1. Phenolic acid identification**

237 Gallic acid (peak 1), p-hydroxybenzoic acid (peak 5), caffeic acid (peak 8), *trans-p*-coumaric acid
238 (peak 11), *trans*-ferulic acid (peak 12), and sinapic acid (peak 13) were identified by comparing the
239 retention times, the UV spectra, and MS spectra of the samples with those of the standards (in positive
240 and negative ion mode). ESI ionization in the negative ion mode showed better sensitivity. In addition,

241 nine phenolic acids were tentatively identified; the results obtained in the negative ion mode are
242 discussed below.

243 Peak 2 had two absorption bands characteristic of phenolic acids, one at 241sh-257 and one at 293 nm.
244 It displayed a major molecular ion at m/z 299.0773. A fragment at m/z 137.0235 was seen in the MS^2
245 spectra; it may indicate the presence of a hydroxybenzoic moiety and the loss of a hexoside residue (-
246 162). A fragment at m/z 93.0334 was also observed and was related to the loss of a carboxylic acid
247 functional group (-44). Peak 2 was thus tentatively identified as an hydroxybenzoic hexoside. Peak 3
248 had two absorption bands, one at 256 nm and one at 293 nm. It displayed a major molecular ion at m/z
249 153.0183. A fragment ion was seen in the MS^2 spectra at m/z 109.0284 and may correspond to the loss
250 of COO. This peak was thus tentatively identified as protocatechuic acid, which has previously been
251 observed in the pea (Klejdus *et al.*, 2008; Singh *et al.*, 2017).

252 Peak 6 had two absorption bands, one at 271 nm and one at 317 nm. It displayed a major molecular
253 ion at m/z 341.0874. A fragment ion was seen in the MS^2 spectra at m/z 179.0343; it may correspond
254 to [caffeic acid-H⁺]. The single secondary fragment at m/z 135 may correspond to [caffeic
255 acid-CO₂-H⁺] (Jaiswal *et al.*, 2014). This peak was thus tentatively identified as caffeoyl hexoside,
256 previously observed in pea leaves (Klopsch *et al.*, 2019; Neugart *et al.*, 2015). Peak 7 had an
257 absorption band at 295 nm and a major molecular ion at m/z 325.0928. The fragment ions at m/z
258 163.0396 and 119.0492 in the MS^2 spectra could correspond to [coumaric acid-H⁺] and [coumaric
259 acid-CO₂-H⁺] (Iswaldi *et al.*, 2013). The peak was thus tentatively identified as coumaroyl-4-O-
260 hexoside, also previously observed in pea leaves (Klopsch *et al.*, 2019).

261 Peaks 4, 9, 10, 14, and 15 all displayed even ion mass, the presence of an odd number of nitrogen
262 atoms, and a fragment at m/z 132.0293 that was attributed to [aspartic acid-H⁺]. They also had a
263 fragment ion resulting from the neutral loss of part of an aspartic acid (116 amu, C₄H₄O₄), which
264 indicates a link between the phenolic acid and the aspartic acid in the form of an amide bond.

265 Fragment ions corresponding to [phenolic acid -H⁺] were also observed, as described earlier for N-
266 caffeoylaspartic acid (Fayeulle *et al.*, 2019); they apparently arose following the fragmentation of the
267 peptides containing aspartic acid (Waugh *et al.*, 1991; Harrison and Young, 2006).

268 Peak 4 had an absorption band at 254 nm and a major molecular ion at m/z 252.0512. There were
269 fragment ions at m/z 132.0293, 136.0394, 137.0235, and 93.0334 in the MS^2 spectra that may
270 correspond to [aspartic acid– H^+], [hydroxybenzoyl amide – H^+] (neutral loss of 116 amu),
271 [hydroxybenzoic acid – H^+], and a [phenol – H^+], respectively. Peak 4 was tentatively identified as N-p-
272 hydroxybenzoyl aspartic acid based on the fragment at m/z 136.0394, which confirmed the presence of
273 an amide link, and the fragment at m/z 132.0293, which confirmed the presence of aspartic acid
274 (Clifford et Knight 2004).

275 Similarly, peaks 9, 10, 14, and 15 displayed even ion mass and a fragment at m/z 132.0293, which was
276 attributed to aspartic acid. Peaks 9 and 10 presented two absorption bands characteristic of
277 hydroxycinnamic acids at 295/293 nm and 310/320 nm, respectively. Peak 9 showed evidence of a
278 major molecular ion at m/z 278.067. The fragment ion at m/z 162.0552 could correspond to coumaroyl
279 amide following the loss of 116 amu. Peak 9 was thus tentatively identified as N-coumaroyl aspartic
280 acid (Oracz *et al.*, 2019).

281 Peak 10 displayed signs of a major molecular ion at m/z 308.0776. The fragment at m/z 192.0660
282 could correspond to [feruloyl amide – H^+]. The peak was thus tentatively identified as a N-feruloyl
283 aspartic acid. Peaks 14 and 15 had an absorption band at 250 nm. Peak 14 had a major molecular ion
284 at m/z 250.0720 and a fragment at m/z 134.0601, which may correspond to [phenylacetyl amide – H^+].
285 Peak 14 was thus tentatively identified as an N-phenylacetyl aspartic acid. Peak 15 had a major
286 molecular ion at m/z 266.0670 and was tentatively identified as a hydroxyphenylacetic acid-aspartic
287 acid conjugate (Waugh et al, 1991; Harrison and Young, 2006).

288

289 **3.1.2. Flavonoid identification**

290 The standard solutions of kaempferol 3-O-glucoside (peak 20), quercetin-3-O-rutinoside (peak 25),
291 quercetin-3-O-glucoside (peak 26), catechin (peak 28), and naringin (peak 29) were analyzed using the
292 ESI source in positive and negative ion modes. There was greater sensitivity in the negative ion mode
293 for these compounds. In addition, ten other flavonoids were tentatively identified and are discussed
294 below, based on the results obtained in negative ion mode. These compounds displayed two absorption
295 bands characteristic of flavonoids: one within the range of 211–267 nm, corresponding to the phenolic

296 core, and the second within the range of 347–373 nm, corresponding to the conjugated system (Mabry
297 *et al.*, 1970).

298 In the MS² spectra, peaks 16 to 22 had fragment ions at m/z 287.0545 (positive ion mode) and either
299 m/z 284.0237 or 285.0405 (negative ion mode), which suggests the presence of a kaempferol moiety.

300 Also in the MS² spectra, peaks 16, 17, 19, 21, and 22 had a fragment at m/z 609.1436, corresponding
301 to [kaempferol + 2 hexoses], and peaks 21 and 22 had a fragment at m/z 771.1990, corresponding to

302 [kaempferol + 3 hexoses]. Kaempferol derivatives have been observed in pea seed coats (Duenas *et*
303 *al.*, 2004; Stanisavljevic *et al.*, 2015), pea seeds (Jha *et al.*, 2019), pea leaves (Klopsch *et al.*, 2019;

304 Neugart *et al.*, 2015), and pea shoots (Ferrerres *et al.*, 1995; Santos *et al.*, 2014). Peak 16 displayed
305 signs of a major molecular ion at m/z 771.1978 and was tentatively identified as kaempferol

306 sophorotrioxide, which has previously been detected in pea shoots using nuclear magnetic resonance
307 (NMR) spectroscopy (Ferrerres *et al.*, 1995). Peak 19 had a major molecular ion at m/z 695.1442 and

308 was tentatively identified as kaempferol malonyl di-hexoside. Peak 22 had a major molecular ion at
309 m/z 947.2452 and was tentatively identified as kaempferol feruloyl tri-hexoside. This identification

310 was based on comparisons with data for quercetin feruloyl tri-hexoside, found in pea shoots using
311 NMR spectroscopy (Ferrerres *et al.*, 1995), and the results of Goupy *et al.* (2013). Peak 21 displayed

312 signs of a major molecular ion at m/z 977.2558 and was tentatively identified as kaempferol sinapoyl
313 tri-hexoside based on comparisons with data for quercetin sinapoyl tri-hexoside, also found in pea

314 shoots using NMR spectroscopy (Ferrerres *et al.*, 1995). Peaks 17 and 18 had major molecular ions at
315 m/z 753.1879 and 593.1510, respectively. They were tentatively identified as kaempferol derivatives.

316 In the MS² spectra, peaks 23 through 27 displayed a fragment ion at either m/z 303.0497 (positive ion
317 mode) or m/z 300.0276 (negative ion mode), which could indicate the presence of a quercetin moiety

318 or a flavone of the same mass. Also in the MS² spectra, peaks 23 and 24 had a fragment at m/z
319 445.0775, which could correspond to [quercetin + hexose - H₂O]. Like kaempferol derivatives,

320 quercetin derivatives have been found in pea seed coats (Duenas *et al.*, 2004; Stanisavljevic *et al.*,
321 2015), pea seeds (Jha *et al.*, 2019), pea leaves (Klopsch *et al.*, 2019; Neugart *et al.*, 2015), and pea

322 shoots (Ferrerres *et al.*, 1995; Santos *et al.*, 2014). Peak 23 had a major molecular ion at m/z 787.1926
323 and two absorption bands, at 256-267sh nm and 350 nm, respectively. The peak was tentatively

324 identified as quercetin tri-hexoside. Peak 24 had a major molecular ion at m/z 625.1401 and two
325 absorption bands, at 256 nm and at 355 nm, respectively. The peak was tentatively identified as
326 quercetin di-hexoside. Peak 27 had a major molecular ion at m/z 301.0352 and an absorption band at
327 370 nm; it was also associated with a relatively higher retention time (35.95 min). The peak was
328 tentatively identified as quercetin aglycone, which has been observed in pea seed coats (Stanisavljevic
329 *et al.*, 2015) and pea seeds (Jha *et al.*, 2019).

330 Peak 30 had two absorption bands, at 267 nm and 336 nm, respectively, and a major molecular ion at
331 m/z 431.0981. In the MS^2 spectra, fragment ions occurred at m/z 269.0456, potentially corresponding
332 to an apigenin moiety, and at m/z 164.0448, potentially corresponding to an O-hexoside. The peak was
333 tentatively identified as apigenin-7-O-glucoside, which has been found in pea roots (based on
334 comparison with a standard; Šibul *et al.*, 2016).

335

336 **3.1.3. Terpenoid identification**

337 The standard solution of soyasapogenol B (peak 37) was analyzed using the ESI source in positive and
338 negative ion modes. There was greater sensitivity in the positive ion mode for soyasapogenol B. The
339 peak had a single, small absorption band whose maximum was at 228 nm and a major ion at m/z
340 459.3844 in the MS^1 spectra (positive ion mode). There were two fragment ions at m/z 441.3723 and
341 423.3616 in the MS^2 spectra (positive ion mode).

342 In addition, seven peaks were tentatively identified as saponins (peaks 31 to 36). Sensitivity was better
343 in the positive ion mode for all seven; each peak displayed only one small absorption band whose
344 maximum occurred between 193 nm and 229 nm, a feature characteristic of saponins (Decroos *et al.*,
345 2005). The peaks' retention times were also high (between 35.9 and 37.6 min), confirming that they
346 could be extremely apolar compounds. In the MS^2 spectra (positive ion mode), the peaks displayed the
347 same set of fragment ions—located at m/z 85.0291, 141.0183, 365.3195, 423.3616, 581.3833, and
348 441.3723—which supports the interpretation that they belong to the same compound family.

349 Soyasapogenol B also had fragments at m/z 441.3723 and 423.3616. The fragments at m/z 581.3833
350 and 423.3616 have been observed for DDMP saponin (Daveby *et al.*, 1998). In addition, saponins
351 have been seen in pea seeds (Curl *et al.*, 1985; Daveby *et al.*, 1998; Heng *et al.*, 2006; Reim & Rohn,

352 2015), and the masses associated with peaks 31 through 36 correspond to *Pisum sativum* saponins in
353 the KNApSAcK database. Peak 33 specifically had a major molecular ion at m/z 943.5251 in the MS¹
354 spectra (positive ion mode) and was tentatively identified as saponin B (Heng *et al.*, 2006).
355 Finally, the last peak, peak 38, did belong to any of the above families. It displayed a major molecular
356 ion at m/z 203.0821 in the negative ion mode and a characteristic absorption band at 279–289sh nm. It
357 was tentatively identified as tryptophan.

358

359 **3.2. Pea phytochemical profiles and key compounds underlying bitterness and astringency**

360 **3.2.1. Pea phytochemical quantification and impacts of extraction and fractionation**

361 This study used six phenolic acids, five flavonoids, and one saponin as standards to characterize key
362 compounds in pea flour, isolates, and fractions. The analysis was based on the UHPLC-DAD-MS
363 results (mass and calibration curves). Table 2 shows the concentrations (\pm standard deviation) of the
364 12 standards. Quercetin-3-O-rutinoside and catechin were identified but could not be quantified
365 because they occurred at very low concentrations in all the sample types. Gallic acid, sinapic acid, and
366 quercetin-3-O-glucoside could only be quantified in the flour samples because their concentrations
367 were too low in the other sample types. In general, compound concentrations were higher in the flour
368 samples than in the isolates. The two exceptions were kaempferol-3-O-glucoside and soyasapogenol
369 B, for which the opposite was seen. In the case of p-hydroxybenzoic acid, the difference in
370 concentrations was relatively small. Furthermore, compound concentrations were generally higher in
371 the isolates than in the fractions, except in the case of soyasapogenol B. The latter had higher
372 concentrations in the pellets. It is difficult to interpret these findings because the samples differed in
373 nature; notably different MS rates and extraction protocols were employed.

374 Concentrations varied between 0 and 2,000 ng/g (i.e., 0.002 mg/g) for the different compounds and
375 sample types. The literature contains few sources of quantitative data on the phytochemicals in pea
376 flour and isolates. Dvorak *et al.* (2011) found that several phenolic acids occurred at varying
377 concentrations (range: 0–0.026 mg/g) in pea samples. Another study found that the total phenolic acid
378 concentrations in peas varied between 12 and 19 mg/g; for total flavonoids, the range was 0–9 mg/g
379 (Nithiyantham *et al.*, 2012). Research on pea flour found that DDMP saponin occurred at levels of

380 1.5 mg/g (Reim & Rohn, 2015) and of 0.7–1.90 mg/g (Heng *et al.*, 2006). Also in pea flour,
381 soyasaponin I occurred within a range of 0.82–2.5 mg/g (Curl *et al.*, 1985); in pea protein isolates,
382 levels reached 1.4 mg/g (1.1 mmol/kg; Gläser *et al.*, 2020). However, these values cannot be reliably
383 compared because they were obtained from different raw materials and underwent different extraction
384 and analysis procedures.

385 In this study, the ratio between peak area and DM content was established to better compare the
386 different pea protein sample types. When characterizing the compounds identified, data were used that
387 described the peak areas in positive mass. The PCA found that the phytochemical profiles of the flour,
388 isolate, and fraction (pellet, retentate, and permeate) solutions were well distributed along axes F1 and
389 F2, which accounted for 82.6% of the variance (Fig. 1). Thus, maps based on the first two axes
390 provided a high-quality representation of the initial multidimensional data. Most of the phytochemical
391 traits were clustered within one quarter of the correlation circles along axis 1 and thus are clearly
392 correlated. Examples can be seen in peak 7 (coumaroyl-4-O-hexoside), peak 10 (N-feruloyl aspartic acid
393 derivative), peak 30 (apigenin-7-O-glucoside), and peak 31 (a saponin derivative). In particular,
394 permeates had large areas associated with phenolic acids, such as peak 14 (a N-phenylacetyl aspartic
395 acid derivative), peak 5 (p-hydroxybenzoic acid), and peak 7 (coumaroyl-4-O-hexoside). A smaller
396 number of phytochemicals representing the different families present were orthogonal to this larger
397 group. For example, along axis 2 are peak 1 (gallic acid), peak 27 (quercetin aglycone), and peak 32 (a
398 saponin derivative). Accounting for DM content, the nine sample types varied in their phytochemical
399 profiles. Permeates a and b had the largest overall areas across all the phytochemical peaks. The flour
400 had the largest areas for the flavonoid peaks. Finally, pellets a and b had the largest terpenoid peaks.

401 Commercial pea protein isolates undergo significant temperature and pH changes during processing.
402 In particular, pea proteins are usually extracted via isoelectric precipitation, during which pH falls to
403 4–5. To promote flocculation, the raw extract can also be heated to increase protein denaturation
404 (Murat *et al.*, 2013). The resulting protein solutions subsequently undergo drum drying or spray
405 drying, which involves a rise in temperature. Flavonoids and especially phenolic acids are highly
406 unstable and easily degraded due to changes in pH, temperature, light conditions, or enzyme presence.
407 Thus, pulses subject to processing display significantly reduced levels of phenolics (Nithiyantham *et*

408 *al.*, 2012), as observed in our results (i.e., the differences between the flour and isolates). However,
409 different factors can result in synergistic or contradictory actions, which might explain why a small set
410 of phytochemicals could present different behaviors. For example, the following can occur during
411 processing: oxidative reactions; compound formation, breakdown, or leaching; losses of solids; and
412 interactions between proteins and other compounds (Nithiyantham *et al.*, 2012). According to the
413 review by Singh *et al.* (2017), saponins are also highly vulnerable to degradation due to modifications
414 in pH and temperature during processing. Furthermore, the large number of phytochemicals in the
415 permeates may have resulted from water-soluble compounds having leached away.

416

417 **3.2.2. Key compounds underlying bitterness and astringency**

418 The phytochemical profiles of the six fractions displayed correlations with the bitterness and
419 astringency scores (analyzed and discussed in Cosson *et al.* [2021]). The main results of this prior
420 research are described here, including a summary of the scores for the two isolates (4% DM), the two
421 permeates, the two retentates, and the two pellets (6% DM) as well as the results of a stepwise
422 multiple comparisons procedure (Newman-Keuls test, alpha level = 0.05) (Fig. 2). Bitterness scores
423 were highest for the retentates and isolates and lowest for the permeates. Furthermore, scores were
424 lower for sample types produced from isolate a than from isolate b. The astringency scores showed
425 less pronounced differences: pellet b had the highest score, and the permeates had the lowest scores.
426 Correlations were characterized between the areas of the phytochemical peaks and the attribute scores
427 (Pearson's r ; alpha level = 0.05). The same sample types were used in the sensory analysis and the
428 phytochemical analysis. Based on psychophysical curves, the perception of a compound in a product
429 depends on compound concentration (Chambers & Koppel, 2013). The relationship between
430 compound concentration and sensory intensity may be linear (i.e., above a threshold) or non-linear
431 (i.e., below a threshold). Consequently, this work explored both linear and logarithmic correlations.
432 Different degrees of correlation were seen between the phytochemical peak areas and the attribute
433 scores in the linear and logarithmic models (Table 3). The linear correlations with the attribute scores
434 are presented in Supplementary Figure 2. Overall, 11 phenolic acids, 2 flavonoids, 6 saponins, and 10

435 non-identified compounds were significantly correlated with bitterness and astringency (Table 3,
436 Supplementary Figure 2).

437 Among the phenolic acids, caffeic acid (peak 8) was positively correlated with bitterness ($R^2 = 0.90$)
438 and astringency ($R^2 = 0.87$). Caffeic acid is known to cause a persistent sensation of intense bitterness
439 in plant products such as coffee (Frank *et al.*, 2006). Caffeoyl hexoside (peak 6) was negatively
440 correlated with both bitterness and astringency. The degradation of caffeic acid derivatives may lead to
441 an increase in the concentration of free caffeic acid. The other phenolic acids—hydroxybenzoic
442 hexoside (peak 2), protocatechuic acid (peak 3), N-p-hydroxybenzoyl aspartic acid (peak 4), p-
443 hydroxybenzoic acid (peak 5), coumaroyl-4-O-hexoside (peak 7), N-phenylacetyl aspartic acid
444 (peak 14), and the aspartic acid derivative (peak 15)—were negatively correlated with astringency.
445 Bitterness was negatively correlated with peaks 4, 5, and 15. Finally, N-coumaroyl aspartic acid
446 (peak 9) and N-feruloyl aspartic acid (peak 10) were negatively correlated with bitterness. A variety of
447 hydroxybenzoic acid ethyl esters and hydroxycinnamic acid ethyl esters have been identified as bitter
448 compounds in wine (Hufnagel & Hofmann, 2008). However, these ethyl esters are less polar than the
449 acids, which could explain their bitterness. A range of N-phenylpropenoyl-L-amino acids have also
450 been identified as key astringent compounds in roasted cocoa (Stark *et al.*, 2006), and several
451 hydroxycinnamic acids act as precursors of off-flavors in fruit (Naim *et al.*, 1992). Thus, the negative
452 correlations above could also be explained by the release of related compounds (not identified here)
453 that are responsible for perceived bitterness and astringency. With regards to the threshold values,
454 various figures have been reported and are matrix dependent. For example, threshold values for
455 chlorogenic, caffeic, and p-coumaric acids were 40–90 mg/L in water; 520–690 mg/L in beer; and 10–
456 32 mg/L in wine (Boulet *et al.*, 2017). All these compound concentrations are much higher than those
457 in the isolates.

458 Among the flavonoids, a kaempferol derivative (peak 17) was negatively correlated with bitterness
459 and astringency ($R^2 = -0.93$ and $R^2 = -0.94$). Quercetin-3-O-glucoside (peak 26) was positively
460 correlated with astringency ($R^2 = 0.85$). Past research has shown that flavonoids can contribute to both
461 these attributes, although most of this work was focused on polyphenols in red wine (Hufnagel &
462 Hofmann, 2008). From a mechanistic perspective, the structural configuration of flavonoid compounds

463 plays a role in activating bitterness receptors. (+)-catechin can activate the TAS2R14 and TAS2R39
464 receptors, while (-)-epicatechin can activate the TAS2R4, TAS2R5, TAS2R14, and TAS2R39
465 receptors (Roland *et al.*, 2017). Moreover, the molecular size of polyphenol compounds also plays a
466 role in receptor activation dynamics. Larger polymers tend to result in less bitterness and more
467 astringency, whereas smaller polymers tend to result in more bitterness and less astringency (Sun *et*
468 *al.*, 2007). In addition, the presence of galloyl groups on epicatechin can affect receptor activation, and
469 the aglycone isomers of isoflavones, being more hydrophobic, are more compatible with the receptors
470 than are their glucoside counterparts. Thus, the negative correlations could be explained by the
471 presence of derivatives associated with peak 17 (e.g., more or less glycolyzed compound forms not
472 identified here) that could play a role in perceived bitterness and astringency. The other flavonoids
473 identified in this study were not correlated with either attribute. However, given that flavone threshold
474 values of 0.1–20 mg/L have been reported for red wine (Sáenz-Navajas *et al.*, 2010), it could be that
475 compound concentrations in the isolates were too low.

476 Finally, six saponins were positively correlated with astringency. Saponins are generally perceived as
477 bitter and astringent (Heng *et al.*, 2006). Here, the saponin concentrations in the pea protein isolates
478 were calculated based on the soyasapogenol B standard and then compared to values in the literature.
479 The disadvantage of this approach was that a single standard was employed for all the saponins instead
480 of a unique standard for each. However, no commercial standards are available for pea saponins given
481 that their purification remains challenging. The mean concentrations (\pm standard deviation) of the six
482 saponins were calculated, and their correlations with the astringency scores were examined (Fig. 3).
483 Saponin B (peak 33) occurred at the highest concentration (0.05 mg/g). The panelists performing the
484 sensory analysis were therefore exposed to saponin B levels of 20 mg/L, given the compound's
485 relative concentration (4%) in the pea protein isolates. Previous work found that saponin-mediated
486 bitterness could be perceived by panelists at very low concentrations in dry peas, at around 2 mg/L for
487 a saponin mixture (saponin B and DDMP saponin in a 1:4 ratio) and around 8 mg/L for saponin B
488 (Heng *et al.*, 2006). In this study, the concentration of saponin B should have been high enough to be
489 perceived by the panelists and to thus contribute to sensations of bitterness and astringency. In
490 contrast, soyasaponin I appears to have perception thresholds of 1.62 mmol/L (1,528 mg/L) for

491 bitterness and 0.64 mmol/L (604 mg/L) for astringency (Gläser *et al.*, 2020). Thus, this saponin was
492 not concentrated enough to contribute to bitterness and astringency on its own; it may, however, have
493 exerted an influence through interactions with other compounds.

494

495 **4. Conclusion**

496 In this study, UHPLC-DAD-MS was used to identify the main phytochemicals present in pea flour,
497 isolates, and fractions. Several key results emerged. First, 48 phytochemicals were observed. Fifteen
498 compounds were tentatively identified as phenolic acids, 15 flavonoids, and 7 saponins. Furthermore,
499 when the MS data were compared with the reference standards data, it was possible to unambiguously
500 identify and quantify 2 hydroxybenzoic acids, 4 hydroxycinnamic acids, 3 flavonols, 2 flavanols, and
501 1 saponin. Second, based on the peak areas for the compounds, larger amounts of phytochemicals
502 were present in the flour than in the isolates and fractions, suggesting compounds experienced
503 degradation during processing. However, when accounting for DM content, the permeates contained
504 larger amounts of phytochemicals, which could have resulted from the leaching away of water-soluble
505 compounds. Third, the peak areas of the compounds displayed different degrees of correlation with
506 perceived bitterness and astringency. A total of 29 compounds (phenolic acids, flavonoids, and
507 saponins) were correlated with one or both attributes. It is possible that the complex mixture of these
508 compounds (which includes other compounds as well, such as peptides) could influence overall
509 perceptions of bitterness and astringency. Consequently, at this stage, it remains difficult to make
510 concrete recommendations about which phytochemicals could be removed to improve the desirability
511 of commercial pea-protein-based products. That said, one promising strategy could be to explore
512 different phytochemical compositions, such as those resulting from a decrease in oxidative reactions or
513 the leaching of water-soluble compounds.

514 **5. CRediT author statement**

515 Audrey Cosson: Methodology, Investigation, Formal analysis, Writing - Original Draft. Emmanuelle
516 Meudec: Resources, Investigation. Christian Ginies: Investigation. Alice Danel: Resources,
517 Investigation. Pascale Lieben: Resources, Investigation. Nicolas Descamps: Funding acquisition.

518 Véronique Cheynier: Resources, Supervision, Writing - review & editing. Anne Saint-Eve:
519 Methodology, Writing - review & editing. Isabelle Souchon: Conceptualization, Supervision, Writing
520 - review & editing.

521 **6. Acknowledgments**

522 This work was funded by Roquette (Lestrem, France), the French National Research and Technology
523 Agency (ANRT-CIFRE 2017/0815), AgroParisTech (Paris, France), and the French National Research
524 Institute for Agriculture, Food, and Environment (INRAE). The authors thank the Polyphenol
525 Platform (<https://www6.montpellier.inrae.fr/spo/Structures-collectives/Plate-forme-Polyphenols>). The
526 authors are also grateful to David Forest for providing technical support.

527 **7. Declaration of interest**

528 This work was supported by Roquette (Lestrem, France) and the French National Research and
529 Technology Agency (ANRT-CIFRE 2017/0815). Roquette has provided the samples of pea flour and
530 pea protein isolates. They did not participate to the analysis and interpretation of results and agreed to
531 submit the article for publication.

532

533 **8. References**

- 534 Boulet, J.-C., Ducasse, M.-A., & Cheynier, V. (2017). Ultraviolet spectroscopy study of phenolic
535 substances and other major compounds in red wines: Relationship between astringency and the
536 concentration of phenolic substances. *Australian Journal of Grape and Wine Research*, 7.
537 <https://doi.org/10.1111/ajgw.12265>
- 538 Chambers, E., & Koppel, K. (2013). Associations of volatile compounds with sensory aroma and
539 flavor: The complex nature of flavor. *Molecules*, 18(5), 4887–4905.
540 <https://doi.org/10.3390/molecules18054887>
- 541 Clifford, M. (2004). The cinnamoyl-amino acid conjugates of green robusta coffee beans. *Food*
542 *Chemistry*, 87(3), 457–463. <https://doi.org/10.1016/j.foodchem.2003.12.020>

543 Cosson, A., Blumenthal, D., Descamps, N., Souchon, I., & Saint-Eve, A. (2021). Using a mixture
544 design and fraction-based formulation to better understand perceptions of plant-protein-based
545 solutions. *Food Research International*, *141*, 110151. <https://doi.org/10.1016/j.foodres.2021.110151>

546 Curl, C. L., Price, K. R., & Fenwick, G. R. (1985). The quantitative estimation of saponin in pea
547 (*Pisum sativum* L.) and soya (*Glycine max*). *Food Chemistry*, *18*(4), 241–250.
548 [https://doi.org/10.1016/0308-8146\(85\)90105-0](https://doi.org/10.1016/0308-8146(85)90105-0)

549 Daveby, Y. D., Åman, P., Betz, J. M., & Musser, S. M. (1998). Effect of storage and extraction on
550 ratio of soyasaponin I to 2,3-dihydro-2,5-dihydroxy-6-methyl-4-pyrone-conjugated soyasaponin I in
551 dehulled peas (*Pisum sativum*L). *Journal of the Science of Food and Agriculture*, *78*(1), 141–146.
552 [https://doi.org/10.1002/\(SICI\)1097-0010\(199809\)78:1<141::AID-JSFA169>3.0.CO;2-6](https://doi.org/10.1002/(SICI)1097-0010(199809)78:1<141::AID-JSFA169>3.0.CO;2-6)

553 Decroos, K., Vincken, J.-P., Heng, L., Bakker, R., Gruppen, H., & Verstraete, W. (2005).
554 Simultaneous quantification of differently glycosylated, acetylated, and 2,3-dihydro-2,5-dihydroxy-6-
555 methyl-4H-pyran-4-one-conjugated soyasaponins using reversed-phase high-performance liquid
556 chromatography with evaporative light scattering detection. *Journal of Chromatography A*, *1072*(2),
557 185–193. <https://doi.org/10.1016/j.chroma.2005.03.021>

558 Duenas, M., Estrella, I., & Hernandez, T. (2004). Occurrence of phenolic compounds in the seed coat
559 and the cotyledon of peas (*Pisum sativum* L.). *European Food Research and Technology*, *219*(2).
560 <https://doi.org/10.1007/s00217-004-0938-x>

561 Dvořák, R., Pechová, A., Pavlata, L., Filípek, J., Dostálová, J., Réblová, Z., Klejdus, B., Kovařík, K.,
562 & Poul, J. (2011). Reduction in the content of antinutritional substances in pea seeds (*Pisum sativum*
563 L.) by different treatments. *Czech Journal of Animal Science*, *50*(No. 11), 519–527.
564 <https://doi.org/10.17221/4257-CJAS>

565 Fahim, J. R., Attia, E. Z., & Kamel, M. S. (2019). The phenolic profile of pea (*Pisum sativum*): A
566 phytochemical and pharmacological overview. *Phytochemistry Reviews*, *18*(1), 173–198.
567 <https://doi.org/10.1007/s11101-018-9586-9>

568 Fayeulle, N., Meudec, E., Boulet, J. C., Vallverdu-Queralt, A., Hue, C., Boulanger, R., Cheynier, V.,
569 & Sommerer, N. (2019). Fast discrimination of chocolate quality based on average-mass-spectra

570 fingerprints of cocoa polyphenols. *Journal of Agricultural and Food Chemistry*, 67(9), 2723–2731.
571 <https://doi.org/10.1021/acs.jafc.8b06456>

572 Ferreres, F., Esteban, E., Carpena-Ruiz, R., Jiménez, M. A., & Tomás-Barberán, F. A. (1995).
573 Acylated flavonol sophorotriosides from pea shoots. *Phytochemistry*, 39(6), 1443–1446.
574 [https://doi.org/10.1016/0031-9422\(95\)00137-V](https://doi.org/10.1016/0031-9422(95)00137-V)

575 Frank, O., Zehentbauer, G., & Hofmann, T. (2006). Bioresponse-guided decomposition of roast coffee
576 beverage and identification of key bitter taste compounds. *European Food Research and Technology*,
577 222(5–6), 492–508. <https://doi.org/10.1007/s00217-005-0143-6>

578 Gläser, P., Dawid, C., Meister, S., Bader-Mittermaier, S., Schott, M., Eisner, P., & Hofmann, T.
579 (2020). Molecularization of bitter off-taste compounds in pea-protein Isolates (*Pisum sativum* L.).
580 *Journal of Agricultural and Food Chemistry*, acs.jafc.9b06663.
581 <https://doi.org/10.1021/acs.jafc.9b06663>

582 Goupy, P., Vian, M. A., Chemat, F., & Caris-Veyrat, C. (2013). Identification and quantification of
583 flavonols, anthocyanins and lutein diesters in tepals of *Crocus sativus* by ultra performance liquid
584 chromatography coupled to diode array and ion trap mass spectrometry detections. *Industrial Crops*
585 *and Products*, 44, 496–510. <https://doi.org/10.1016/j.indcrop.2012.10.004>

586 Harrison, A. G., & Young, A. B. (2006). Fragmentation reactions of deprotonated peptides containing
587 aspartic acid. *International Journal of Mass Spectrometry*, 255–256, 111–122.
588 <https://doi.org/10.1016/j.ijms.2005.12.037>

589 Heng, L., Vincken, J.-P., van Koningsveld, G., Legger, A., Gruppen, H., van Boekel, T., Roozen, J., &
590 Voragen, F. (2006). Bitterness of saponins and their content in dry peas. *Journal of the Science of*
591 *Food and Agriculture*, 86(8), 1225–1231. <https://doi.org/10.1002/jsfa.2473>

592 Hufnagel, J. C., & Hofmann, T. (2008). Orosensory-directed identification of astringent mouthfeel and
593 bitter-tasting compounds in red wine. *Journal of Agricultural and Food Chemistry*, 56(4), 1376–1386.
594 <https://doi.org/10.1021/jf073031n>

595 Iswaldi, I., Arráez-Román, D., Gómez-Caravaca, A. M., Contreras, M. del M., Uberos, J., Segura-
596 Carretero, A., & Fernández-Gutiérrez, A. (2013). Identification of polyphenols and their metabolites in

597 human urine after cranberry-syrup consumption. *Food and Chemical Toxicology*, 55, 484–492.
598 <https://doi.org/10.1016/j.fct.2013.01.039>

599 Jaiswal, R., Matei, M. F., Glembockyte, V., Patras, M. A., & Kuhnert, N. (2014). Hierarchical key for
600 the LC-MSn identification of all ten regio- and stereoisomers of caffeoylglucose. *J. Agric. Food*
601 *Chem.*, 14. <https://doi.org/10.1021/jf501210s>

602 Jha, A. B., Purves, R. W., Elessawy, F. M., Zhang, H., Vandenberg, A., & Warkentin, T. D. (2019).
603 Polyphenolic profile of seed components of white and purple flower pea lines. *Crop Science*, 59(6),
604 2711–2719. <https://doi.org/10.2135/cropsci2019.04.0279>

605 Klejdus, B., Vacek, J., Lojková, L., Benešová, L., & Kubáň, V. (2008). Ultrahigh-pressure liquid
606 chromatography of isoflavones and phenolic acids on different stationary phases. *Journal of*
607 *Chromatography A*, 1195(1–2), 52–59. <https://doi.org/10.1016/j.chroma.2008.04.069>

608 Klopsch, Baldermann, Voss, Rohn, Schreiner, & Neugart. (2019). Narrow-Banded UVB Affects the
609 Stability of Secondary Plant Metabolites in Kale (*Brassica oleracea* var. *sabellica*) and Pea (*Pisum*
610 *sativum*) Leaves Being Added to Lentil Flour Fortified Bread: A Novel Approach for Producing
611 Functional Foods. *Foods*, 8(10), 427. <https://doi.org/10.3390/foods8100427>

612 Mabry, T. J., Markham, K. R., & Mabry, H. (1970). *The Systematic Identification of Flavonoids*.
613 Springer Verlag.

614 Murat, Bard, M.-H., Dhalleine, C., & Cayot, N. (2013). Characterisation of odour active compounds
615 along extraction process from pea flour to pea protein extract. *Food Research International*, 53(1),
616 31–41. <https://doi.org/10.1016/j.foodres.2013.03.049>

617 Naim, M., Zehavi, U., Nagy, S., & Rouseff, R. L. (1992). Hydroxycinnamic acids as off-flavor
618 precursors in citrus fruits and their products. In C.-T. Ho, C. Y. Lee, & M.-T. Huang (Eds.), *Phenolic*
619 *Compounds in Food and Their Effects on Health I* (Vol. 506, pp. 180–191). American Chemical
620 Society. <https://doi.org/10.1021/bk-1992-0506.ch014>

621 Neugart, S., Rohn, S., & Schreiner, M. (2015). Identification of complex, naturally occurring
622 flavonoid glycosides in *Vicia faba* and *Pisum sativum* leaves by HPLC-DAD-ESI-MSn and the
623 genotypic effect on their flavonoid profile. *Food Research International*, 76, 114–121.
624 <https://doi.org/10.1016/j.foodres.2015.02.021>

625 Nithiyanantham, S., Selvakumar, S., & Siddhuraju, P. (2012). Total phenolic content and antioxidant
626 activity of two different solvent extracts from raw and processed legumes, *Cicer arietinum* L. and
627 *Pisum sativum* L. *Journal of Food Composition and Analysis*, 27(1), 52–60.
628 <https://doi.org/10.1016/j.jfca.2012.04.003>

629 Oracz, J., Nebesny, E., & Żyżelewicz, D. (2019). Identification and quantification of free and bound
630 phenolic compounds contained in the high-molecular weight melanoidin fractions derived from two
631 different types of cocoa beans by UHPLC-DAD-ESI-HR-MSn. *Food Research International*, 115,
632 135–149. <https://doi.org/10.1016/j.foodres.2018.08.028>

633 Reim, V., & Rohn, S. (2015). Characterization of saponins in peas (*Pisum sativum* L.) by HPTLC
634 coupled to mass spectrometry and a hemolysis assay. *Food Research International*, 76, 3–10.
635 <https://doi.org/10.1016/j.foodres.2014.06.043>

636 Robichaud, J. L., & Noble, A. C. (1990). Astringency and bitterness of selected phenolics in wine.
637 *Journal of the Science of Food and Agriculture*, 53(3), 343–353.
638 <https://doi.org/10.1002/jsfa.2740530307>

639 Roland, W. S. U., Pouvreau, L., Curran, J., van de Velde, F., & de Kok, P. M. T. (2017). Flavor
640 aspects of pulse ingredients. *Cereal Chemistry Journal*, 94(1), 58–65.
641 <https://doi.org/10.1094/CCHEM-06-16-0161-FI>

642 Sáenz-Navajas, M.-P., Tao, Y.-S., Dizy, M., Ferreira, V., & Fernández-Zurbano, P. (2010).
643 Relationship between nonvolatile composition and sensory properties of premium spanish red wines
644 and their correlation to quality perception. *Journal of Agricultural and Food Chemistry*, 58(23),
645 12407–12416. <https://doi.org/10.1021/jf102546f>

646 Santos, J., Oliveira, M. B. P. P., Ibáñez, E., & Herrero, M. (2014). Phenolic profile evolution of
647 different ready-to-eat baby-leaf vegetables during storage. *Journal of Chromatography A*, 1327, 118–
648 131. <https://doi.org/10.1016/j.chroma.2013.12.085>

649 Šibul, F., Orčić, D., Vasić, M., Anačkov, G., Nađpal, J., Savić, A., & Mimica-Dukić, N. (2016).
650 Phenolic profile, antioxidant and anti-inflammatory potential of herb and root extracts of seven
651 selected legumes. *Industrial Crops and Products*, 83, 641–653.
652 <https://doi.org/10.1016/j.indcrop.2015.12.057>

653 Singh, B., Singh, J. P., Kaur, A., & Singh, N. (2017). Phenolic composition and antioxidant potential
654 of grain legume seeds: A review. *Food Research International*, *101*, 1–16.
655 <https://doi.org/10.1016/j.foodres.2017.09.026>

656 Stanisavljevic, N., Ilic, M., Jovanovic, Z., Cupic, T., Dabic, D., Natic, M., Tesic, Z., & Radovic, S.
657 (2015). Identification of seed coat phenolic compounds from differently colored pea varieties and
658 characterization of their antioxidant activity. *Archives of Biological Sciences*, *67*(3), 829–840.
659 <https://doi.org/10.2298/ABS141204042S>

660 Stark, T., Bareuther, S., & Hofmann, T. (2006). Molecular definition of the taste of roasted Cocoa
661 Nibs (*Theobroma cacao*) by means of quantitative studies and sensory experiments. *Journal of*
662 *Agricultural and Food Chemistry*, *54*(15), 5530–5539. <https://doi.org/10.1021/jf0608726>

663 Sun, J., Liang, F., Bin, Y., Li, P., & Duan, C. (2007). Screening non-colored phenolics in red wines
664 using liquid chromatography/ultraviolet and mass spectrometry/mass spectrometry libraries.
665 *Molecules*, *12*(3), 679–693. <https://doi.org/10.3390/12030679>

666 Waugh, R. J., Bowie, J. H., & Hayes, R. N. (1991). Collision-Induced dissociations of deprotonated
667 peptides. Dipeptides containing aspartic or glutamic acids. *Organic Mass Spectrometry*, *26*(4), 250–
668 256. <https://doi.org/10.1002/oms.1210260413>

669 **Figure captions**

670

671 **Figure 1:** Principal component analysis (centered reduced variables, Pearson's r) examining the
672 phytochemical profiles of the different sample types: pea flour, isolates, and fractions (pellets,
673 retentates, and permeates). On the right is a loading plot showing the correlational relationships
674 between PCA axes 1 and 2 and the peak areas (accounting for the sample type's dry matter content)
675 for the 54 phytochemicals identified in the study: the phenolic acids are in blue (dotted line), the
676 terpenoids are in red (solid line), the flavonoids are in green (dashed line), and the other compounds
677 are in orange (thick dashed and dotted line). On the left is a PCA plot with the same two axes that
678 shows the relative similarity of the nine sample types.

679

680 **Figure 2:** Bitterness and astringency scores (out of 10) for the different sample types as determined
681 via static profiling by trained panelists (Cosson *et al.*, 2021). Significant differences between groups
682 are indicated by differences in letters (Newman-Keuls test, alpha level = 0.05).

683

684 **Figure 3:** Mean concentrations (\pm standard deviation; ng/g) of the six saponins that contributed to
685 perceived astringency.

686

687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703

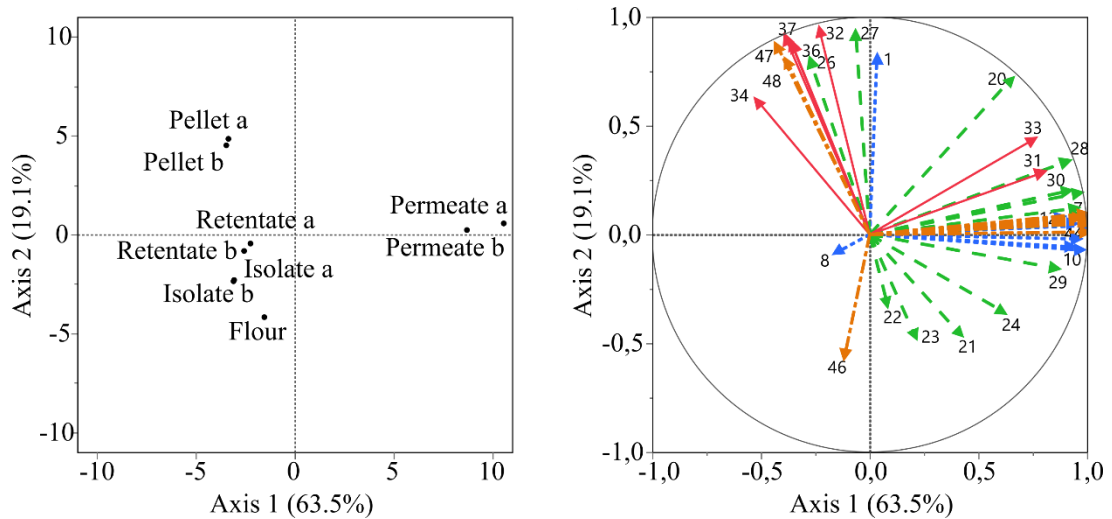
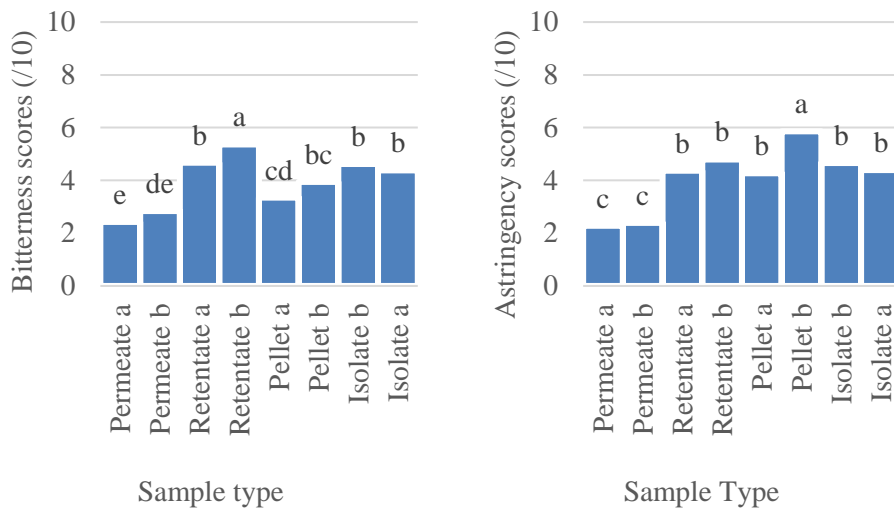


Figure 1: Principal component analysis (centered reduced variables, Pearson's n) examining the phytochemical profiles of the different sample types: pea flour, isolates, and fractions (pellets, retentates, and permeates). On the right is a loading plot showing the correlational relationships between PCA axes 1 and 2 and the peak areas (accounting for the sample type's dry matter content) for the 54 phytochemicals identified in the study: the phenolic acids are in blue (dotted line), the terpenoids are in red (solid line), the flavonoids are in green (dashed line), and the other compounds are in orange (thick dashed and dotted line). On the left is a PCA plot with the same two axes that shows the relative similarity of the nine sample types.

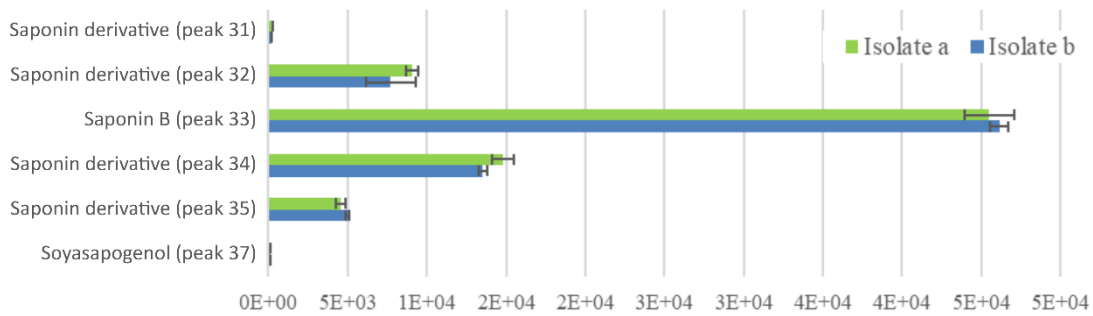
704



705

706 **Figure 2:** Bitterness and astringency scores (out of 10) for the different sample types as determined
707 via static profiling by trained panelists (Cosson *et al.*, 2021). Significant differences between groups
708 are indicated by differences in letters (Newman-Keuls test, alpha level = 0.05).

709



710

711 **Figure 3:** Mean concentrations (\pm standard deviation; ng/g) of the six saponins that contributed to
712 perceived astringency.

713

714

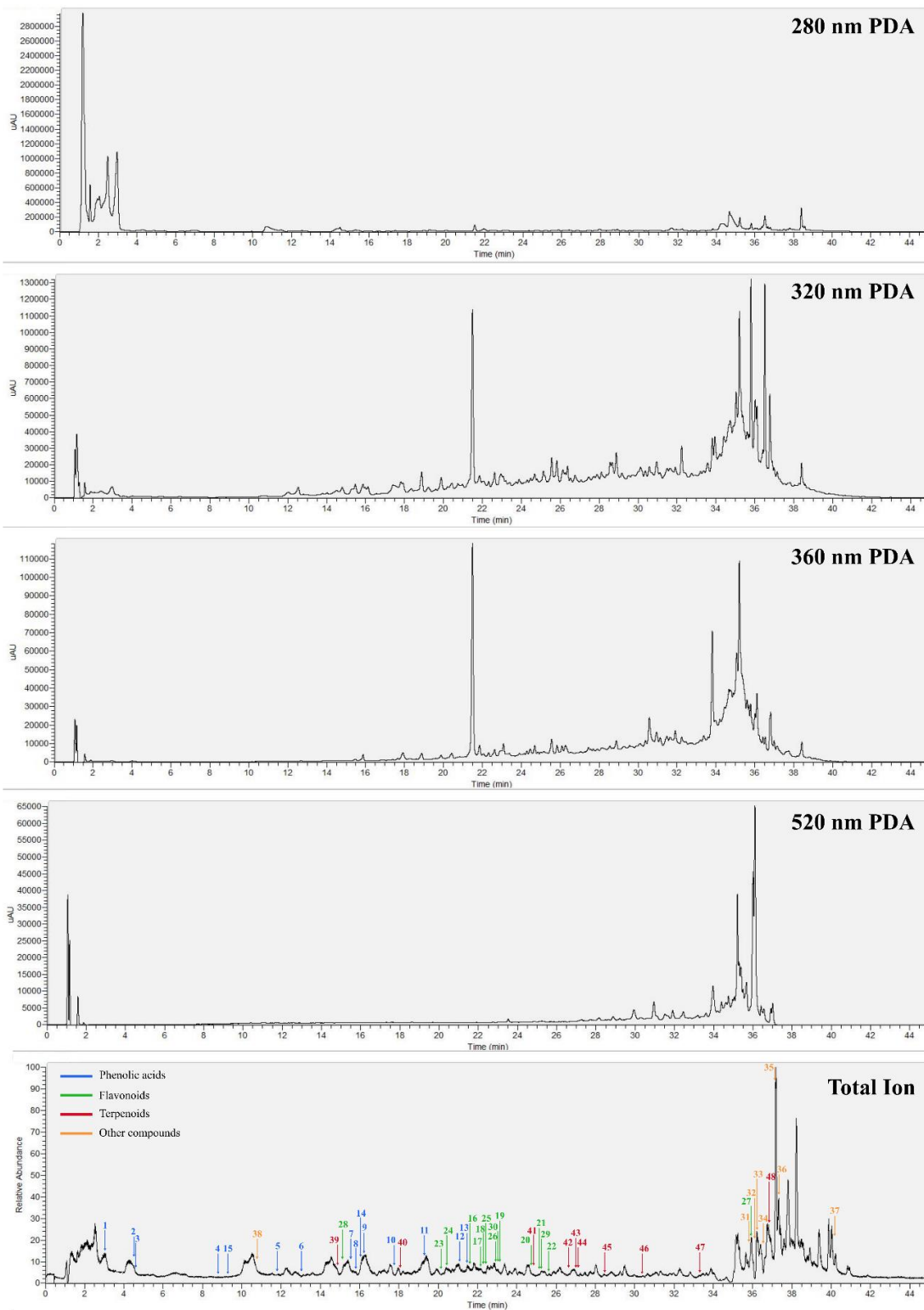
715

716

717

718 **Supplementary Figures**

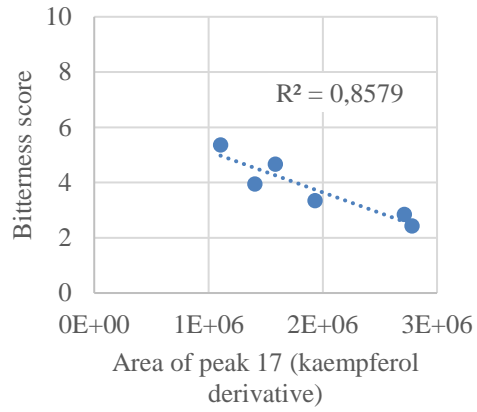
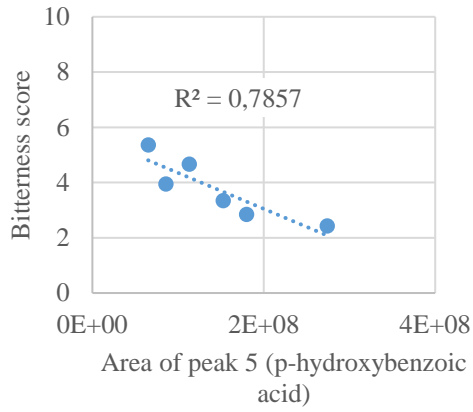
719 **Figure 1:** Total ion chromatogram and diode array detector (PDA) chromatogram at 280, 320, 360,
720 and 520 nm for isolate a.



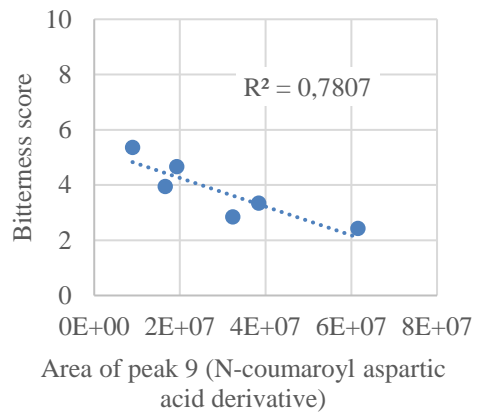
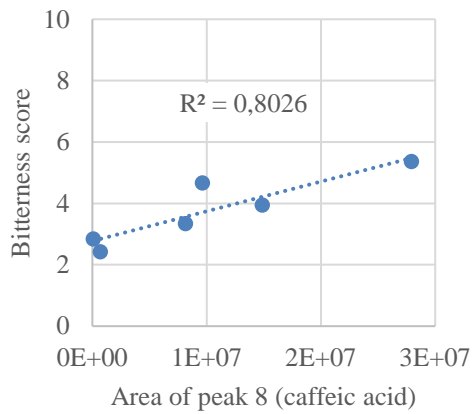
721

722

723 **Supplementary Figure 2:** Representative positive and negative linear correlations between the
724 phytochemicals (peak areas determined via UHPLC-DAD-MS) and the scores for bitterness or
725 astringency (out of 10); the R² values are indicated.



726



727

728

729 **Table captions**

730

731 **Table 1:** Phytochemicals identified in the pea flour, isolates, and fractions. Indicated are peak number,
732 RT [min], UV (nm), MS¹ [M+H]⁺, MS² [M+H]⁺, MS¹ [M-H]⁻, MS² [M-H]⁻, theoretical molecular
733 mass, theoretical [M+H]⁺, theoretical [M-H]⁻, hypothetical chemical formula, tentative of
734 identification, and the database used for identification. The compounds identified using standards are
735 in bold.

736

737 **Table 2:** Concentrations (\pm standard deviation) of the 12 standards (performed in triplicate) on 10-3
738 mg/g. The detection limit (DL) was 0.02 ng/ μ l, and the quantification limit (QL) was 8.0 ng/ μ l. There
739 were two exceptions: soyasapogenol B, where DL = 0.07 ng/g and QL = 3 ng/g, and p-
740 hydroxybenzoic acid, where DL = 0.05 ng/g and QL = 30.0 ng/g.

741

742 **Table 3:** The coefficients (Pearson's r) for the correlations between the phytochemical compounds (peak
743 areas determined via UHPLC-DAD-MS) and the scores for bitterness and astringency (out of 10)
744 determined using linear and logarithmic models. In bold are the negative correlation coefficients. Only
745 the statistically significant values are indicated (p-value < 0.05).

Table 1

Peak number	RT [min]	UV (nm)	MS ¹ [M+H] ⁺	MS ² [M+H] ⁺	MS ¹ [M-H] ⁻	MS ² [M-H] ⁻	Hypothetical formula	Expected [M+H] ⁺	Error (ppm)	Expected [M-H] ⁻	Error (ppm)	Hypothetical class	Hypothetical compound	Database	Relative quantification
Phenolic acids															
1	3.05	NA	171.0291	NA	169.0134	125.0234	C7H6O5	171.0288	-1.8	169.0142	5	Hydroxybenzoic acid	gallic acid	Standard	NEG
2	4.43	241sh, 257, 293	301.1021	NA	299.0773	137.0235; 93.0334	C13H16O8	301.0918	-34.2	299.0772	-0.2	Hydroxybenzoic acid	p-hydroxybenzoic hexoside or isomer	Arita	NEG
3	4.45	256; 293	155.0339	NA	153.0183	109.0284	C7H6O4	155.0339	-0.1	153.0193	6.7	Hydroxybenzoic acid	protocatechuic acid	Arita	NEG
4	8.71	254	254.0658	NA	252.0512	136.0394 (p-hydroxybenzoyl amide-H ⁺); 137.0235 (hydroxybenzoic -H ⁺); 132.0293 (aspartic acid-H ⁺); 93.0334 (phenol moiety)	C11H11NO6	254.0659	0.4	252.0514	0.6	Hydroxybenzoic acid	N-p-hydroxybenzoyl aspartic acid	PubChem	NEG
5	11.7	257	139.039	NA	137.0234	93.0334	C7H6O3	139.039	-0.2	137.0244	7.4	Hydroxybenzoic acid	p-hydroxybenzoic acid	Standard	NEG
6	13	271; 317	343.1023	NA	341.0874	135.0442 (caffeic acid-CO ₂ -H ⁺); 179.0343 (caffeic acid-H ⁺)	C15H18O9	343.1024	0.2	341.0878	1.2	Hydroxycinnamic acid	caffeoyl hexoside	PubChem	NEG
7	15.47	295	327.1085	NA	325.0928	163.0396 (coumaric acid-H ⁺); 119.0492 (coumaric acid-CO ₂ -H ⁺)	C15H18O8	327.1074	-3.2	325.0929	0.3	Hydroxycinnamic acid	coumaroyl- 4-O-hexoside	KNAPSAcK	NEG
8	15.61	320	181.0496	NA	179.0342	135.0442	C9H8O4	181.0495	-0.4	179.035	4.4	Hydroxycinnamic acid	caffeic acid	Standard	NEG
9	16.13	295sh, 310	280.0812	NA	278.067	132.0292 (aspartic acid -H ⁺); 162.0552 (coumaroyl amide -H ⁺); 163.0392 (coumaric acid-H ⁺)	C13H13NO6	280.0816	1.3	278.067	0	Hydroxycinnamic acid	N-coumaroyl aspartic acid derivative	NA	NEG
10	17.7	293sh, 320	310.0916	NA	308.0776	132.0292 (aspartic acid-H ⁺); 192.0660 (feruloyl amide -H ⁺); 193.0502 (ferulic acid -H ⁺)	C14H15NO7	310.0921	1.7	308.0776	-0.1	Hydroxycinnamic acid	N-feruloyl aspartic acid derivative	NA	NEG
11	19.34	310	165.0546	NA	163.0391	119.0491	C9H8O3	165.0546	0.1	163.0401	5.9	Hydroxycinnamic acid	trans-p-coumaric acid	Standard	NEG
12	21.08	324	195.0651	NA	193.05	134.0364; 178.0265; 149.0597	C10H10O4	195.0652	0.4	193.0506	3.3	Hydroxycinnamic acid	trans-ferulic acid	Standard	NEG
13	21.44	325	225.0757	NA	223.0608	208.0372	C11H12O5	225.0757	0.2	223.0612	1.8	Hydroxycinnamic acid	sinapic acid	Standard	NEG
14	16.02	250	252.0864	NA	250.072	132.0293 (aspartic acid-H ⁺); 135.0442 (phenylacetic acid -H ⁺); 134.0601 (phenylacetyl amide-H ⁺)	C12H13NO5	252.0866	1	250.0721	0.4	Phenylacetic acid	N-phenylacetyl aspartic acid derivative	NA	NEG
15	9.34	250	268.0814	NA	266.067	132.0293 (aspartic acid-H ⁺)	C12H13NO6	268.0816	0.6	266.067	0	Other phenolic acid	hydroxyphenylacetic acid aspartic acid conjugate	PubChem	NEG
Flavonoids															
16	21.55	266; 347	773.2117	287.0545 (kaempferol)	771.1978	609.1436 (kaempferol + 2 hexoses); 429.0825; 327.0507; 284.0237 (kaempferol);	C33H40O21	773.2135	2.3	771.1989	1.5	Flavonol (kaempferol derivative)	kaempferol sophorotrioxide	Arita, KNAPSAcK	NEG

						285.0393 (kaempferol); 255.0297; 771.2020										
17	22.13	345	755.201	287.0545 (kaempferol)	753.1879	193.0501; 255.0297; 285.0405 (kaempferol); 609.1436 (kaempferol + 2 hexoses)	C33H38O20	755.2029	2.5	753.1884	0.6	Flavonol (kaempferol derivative)	kaempferol derivative	Arita	NEG	
18	22.27	258; 345	595.1647	287.0545 (kaempferol)	593.151	488.1147; 389.2183; 285.0405 (kaempferol); 255.0297	C27H30O15	595.1657	1.8	593.1512	0.3	Flavonol (kaempferol derivative)	kaempferol derivative	Arita	NEG	
19	23.1	ND	697.1592	287.0545 (kaempferol)	695.1442	488.1147; 489.1039 (-44- hexose); 477.0941 (kaempferol + hexose); 389.2183; 285.0405 (kaempferol); 255.0297; 609.1436 (kaempferol + 2 hexoses); 447.0902 (- malonylhexose)	C30H32O19	697.1611	2.7	695.1465	3.3	Flavonol (kaempferol derivative)	kaempferol malonyl di-hexoside	Arita	NEG	
20	24.75	266; 347	449.1071	287.0545 (kaempferol)	447.0928	284.0323; 255.0297	C21H20O11	449.1078	1.6	447.0933	1.1	Flavonol (kaempferol derivative)	kaempferol-3-O- glucoside	Standard	NEG	
21	25.12	269, 346	979.2695	287.0545 (kaempferol)	977.2558	284.0237 (kaempferol); 255.0297; 609.1436 (kaempferol + 2 hexoses); 771.2020 (kaempferol + 3 hexoses)	C44H50O25	979.2714	1.9	977.2568	1.1	Flavonol (kaempferol derivative)	kaempferol sinapoyl triglucoside	Arita, KNApSAcK	NEG	
22	25.62	ND	949.2592	287.0545 (kaempferol)	947.2452	771.1990 (kaempferol + 3 hexoses); 609.1436 (kaempferol + 2 hexoses); 429.0825; 327.0507; 284.0237 (kaempferol); 255.0297; 300.0277; 389.2183	C43H48O24	949.2608	1.7	947.2463	1.1	Flavonol (kaempferol derivative)	kaempferol feruloyl tri-hexoside	Arita	NEG	
23	20.18	256, 267sh, 350	789.2068	303.0497 (quercetine)	787.1926	300.0276 (quercetine); 445.0775 (quercetine + hexose - H2O); 271.0230; 178.9980	C33H40O22	789.2084	2	787.1938	1.6	Flavonol (quercetin derivative)	quercetin tri- hexoside	KNApSAcK	NEG	
24	20.41	256, 355	627.154	303.0496 (quercetine)	625.1401	300.0276 (quercetine); 445.0775 (quercetine + hexose - H2O); 271.0230; 178.9980	C27H30O17	627.1556	2.5	625.141	1.5	Flavonol (quercetin derivative)	quercetin di- hexoside	KNApSAcK	NEG	
25	22.41	258; 355	611.1589	303.0496 (quercetine)	609.1456	300.0275 (quercetine); 178.9980	C27H30O16	611.1607	2.9	609.1461	0.8	Flavonol (quercetin derivative)	quercetin-3-O- rutinoside	Standard	NEG	
26	22.94	258; 355	465.1016	303.0496 (quercetine)	463.0885	300.0274 (quercetine); 271.0230; 178.9980	C21H20O12	465.1028	2.5	463.0882	-0.6	Flavonol (quercetin derivative)	quercetin-3-O- glucoside	Standard	NEG	
27	35.95	370	303.0277	NA	301.0352	NA	C15H10O7	303.0499	73.4	301.0354	0.6	Flavonol (quercetin derivative)	quercetin aglycone	Arita	NEG	
28	15.02	279	291.086	NA	289.0714	178.998; 271.0612; 151.0389	C15H14O6	291.0863	1	289.0718	1.4	Flavanol	catechin	Standard	NEG	
29	25.26	284	581.1872	273.0754	579.1717	271.0612; 151.0389	C27H32O14	581.1865	-1.2	579.1719	0.4	Flavanone	naringin	Standard	NEG	
30	23	267, 336	NA	NA	431.0981	269.0456 (apigenine); 164.0448; 271.0628	C21H20O10	433.1129	NA	431.0984	0.6	Flavone (apigenin derivative)	apigenine-7-O- glucoside	KNApSAcK	NEG	

Terpenoids

31	35.9	193- 222	925.5158	85.0291; 141.0183; 365.3195; 423.3616; 581.3833; 441.3723	923.4626	NA	C48H76O17	925.5155	-0.3	923.501	41.6	Saponin	saponin derivative	KNApSAcK	POS
32	36.06	193- 222	797.4666	85.0290; 141.0182; 365.3194; 423.3615; 581.3832; 441.3723	795.4531	NA	C42H68O14	797.4682	2	795.4536	0.7	Saponin	saponin derivative	KNApSAcK	POS

33	36.12	195-205	943.5251	85.0293; 141.0181; 365.3193; 423.3614; 581.3831; 441.3723	941.5092	NA	C48H78O18	943.5261	1.1	941.5115	2.5	Saponin	saponin B	PubChem, KNApSAcK	POS
34	36.53	228	941.5086	85.0291; 141.0182; 365.3191; 423.3615; 581.3833; 441.3723	939.4941	NA	C48H76O18	941.5104	2	939.4959	1.9	Saponin	saponin derivative	KNApSAcK	POS
35	37.22	226	971.5193	85.0290; 141.0184; 365.3194; 423.3617; 581.3832	969.5037	NA	C49H78O19	971.521	1.8	969.5065	2.8	Saponin	saponin derivative	KNApSAcK	POS
36	37.6	229	825.4613	85.0288; 141.0180; 365.3192; 423.3615; 581.3831	823.4426	NA	C43H68O15	825.4631	2.2	823.4485	7.2	Saponin	saponin derivative	KNApSAcK	POS
37	40.2	228	459.3844	441.3723; 423.3618	NA	NA	C30H50O3	459.3833	-2.5	457.3687	NA	Saponin	soyasapogenol B	Standard	POS

Others

38	10.79	279-289sh	205.0973	159.0917	203.0821	116.0495; 159.0920	C11H12N2O2	205.0972	-0.7	203.0826	2.5	Amino acid	tryptophan	PubChem	NEG
39	14.9	214; 280; 320sh	457.2171	NA	455.1774	293.1245; 179.0554	NA	NA	NA	NA	NA	NA	NA	NA	NEG
40	18.03	212; 283	389.1794	NA	387.1659	163.1120; 225.1129	NA	389.1806	3.1	387.1661	0.4	NA	NA	NA	NEG
41	24.82	193-218sh	549.2483	NA	547.2394	161.0446; 149.0445; 89.0232	NA	549.2542	10.7	547.2396	0.4	NA	NA	NA	NEG
42	26.6	349	741.222	309.2784; 599.5033; 703.5117	739.207	164.0448; 271.0628	NA	741.2237	2.2	739.2091	2.8	NA	NA	NA	NEG
43	27	280; 340	591.1696	285.0753	589.1566	149.0446; 283.0612; 178.9980	NA	591.1708	2.1	589.1563	-0.5	NA	NA	NA	NEG
44	27.15	193-218sh; 277	705.2587	NA	703.2447	161.0446; 149.0445; 89.0232	NA	705.26	1.9	703.2455	1.1	NA	NA	NA	NEG
45	28.48	250	591.2638	NA	589.2498	161.0446; 149.0445; 89.0232	NA	591.2647	1.6	589.2502	0.6	NA	NA	NA	NEG
46	30.39	258	619.2751	NA	617.2639	NA	NA	NA	NA	NA	NA	NA	NA	NA	NEG
47	33.38	270; 370	271.0598	NA	269.0453	223.0819; 161.0448	NA	271.0601	1.1	269.0455	0.9	NA	NA	NA	NEG
48	36.85	282; 333	313.0711	285.0753; 163.0391; 257.0805	311.0568	267.1967; 134.0463; 271.0612; 283.0627	NA	313.0707	-1.4	311.0561	-2.2	NA	NA	NA	NEG

747

748

749

750 **Table 2**

Compound	gallic acid	p-hydroxybenzoic acid	caffeic acid	coumaric acid	ferulic acid	sinapic acid	kaempferol-3-O-glucoside	quercetin-3-O-rutinoside	quercetin-3-O-glucoside	catechin	naringin	soyasapogenol B
<i>Peak number</i>	<i>1</i>	<i>2</i>	<i>8</i>	<i>11</i>	<i>12</i>	<i>13</i>	<i>20</i>	<i>25</i>	<i>26</i>	<i>28</i>	<i>29</i>	<i>37</i>
Isolate b	QL	1265.9 ± 6.9	DL	3.0 ± 1.7	3.0 ± 1.7	DL	108.9 ± 1.7	QL	QL	QL	5.6 ± 1.7	91.0 ± 10.0
Isolate a	QL	1682.0 ± 56.8	QL	8.0 ± 1.7	8.4 ± 1.7	DL	105.9 ± 1.7	QL	QL	QL	10.2 ± 1.7	93.9 ± 13.42
Permeate a	QL	170.8 ± 6.4	QL	DL	DL	QL	DL	QL	QL	QL	QL	QL
Permeate b	QL	258.5 ± 1.8	QL	2.0 ± 1.7	DL	QL	DL	QL	QL	QL	QL	QL
Pellet b	QL	72.0 ± 6.8	3.1 ± 2.3	DL	DL	QL	29.0 ± 1.7	QL	QL	QL	QL	175.6 ± 15.4
Pellet a	QL	126.5 ± 4.4	DL	1.7 ± 1.7	2.3 ± 1.7	QL	41.7 ± 1.7	QL	QL	QL	QL	188.6 ± 41.6
Retentate b	QL	54.5 ± 6.2	5.8 ± 2.1	DL	DL	QL	2.3 ± 1.7	QL	QL	QL	QL	9.4 ± 10.0
Retentate a	QL	94.2 ± 2.3	DL	DL	DL	QL	3.0 ± 1.7	QL	QL	QL	QL	15.7 ± 10.0
Flour	16.08 ± 1.08	1999.0 ± 18.4	90.7 ± 2.4	124.0 ± 2.2	151.3 ± 1.8	44.49 ± 1.7	62.9 ± 11.1	QL	14.8 ± 1.7	QL	81.7 ± 3.1	48.4 ± 10.0

751

Hypothetical family	Peak number	Hypothetical compound	Linear model - Bitterness		Linear model - Astringency		Logarithmic model - Bitterness		Logarithmic model - Astringency	
			p-value	R ²	p-value	R ²	p-value	R ²	p-value	R ²
Phenolic acids	2	hydroxybenzoic hexoside	NA	NA	<0.01	-0.94	NA	NA	<0.01	-0.96
	3	protocatechuic acid	NA	NA	<0.01	-0.96	NA	NA	<0.01	-0.99
	4	N-p-hydroxybenzoyl aspartic acid	0.04	-0.83	<0.01	-0.95	0.03	-0.85	<0.01	-0.95
	5	p-hydroxybenzoic acid	0.02	-0.89	0.03	-0.87	0.01	-0.92	0.03	-0.86
	6	caffeoyl hexoside	0.02	-0.88	0.01	-0.91	0.02	-0.89	0.01	-0.90
	7	coumaroyl-4-O-hexoside	NA	NA	<0.01	-0.96	NA	NA	<0.01	-0.98
	8	caffeic acid	0.02	0.90	NA	NA	NA	NA	0.02	0.87
	9	N-coumaroyl aspartic acid derivative	0.02	-0.88	NA	NA	0.01	-0.94	NA	NA
	10	N-feruoyl aspartic acid derivative	NA	NA	NA	NA	0.05	-0.82	NA	NA
	14	N-phenylacetyl aspartic acid derivative	NA	NA	0.01	-0.93	NA	NA	<0.01	-0.97
	15	aspartic acid derivative	0.03	-0.84	<0.01	-0.95	0.02	-0.87	<0.01	-0.95
Flavonoids	17	kaempferol derivative	0.01	-0.93	0.01	-0.92	0.01	-0.94	0.02	-0.88
	26	quercetin-3-O-glucoside	NA	NA	NA	NA	NA	NA	0.03	0.85
Terpenoids	31	saponin derivative	NA	NA	0.02	0.88	NA	NA	0.01	0.92
	32	saponin derivative	NA	NA	NA	NA	NA	NA	0.02	0.88
	33	saponin B	NA	NA	NA	NA	NA	NA	0.02	0.90
	34	saponin derivative	NA	NA	NA	NA	NA	NA	0.01	0.92
	35	saponin derivative	NA	NA	NA	NA	NA	NA	0.04	0.83
	37	soyasapogenol B	NA	NA	NA	NA	NA	NA	0.02	0.88
Others	39	NA	NA	NA	<0.01	-0.95	NA	NA	<0.01	-0.96
	42	NA	0.02	-0.90	0.01	-0.92	0.01	-0.90	0.01	-0.90
	43	NA	NA	NA	NA	NA	0.05	-0.81	NA	NA
	40	NA	NA	NA	<0.01	-0.97	NA	NA	<0.01	-0.97
	41	NA	<0.01	-0.97	0.03	-0.84	<0.01	-0.98	NA	NA
	44	NA	NA	NA	<0.01	-0.97	NA	NA	NA	NA
	45	NA	NA	NA	<0.01	-0.96	NA	NA	NA	NA
	46	NA	NA	NA	0.04	0.84	NA	NA	NA	NA
	47	NA	NA	NA	NA	NA	NA	NA	0.02	0.88
	48	NA	NA	NA	NA	NA	NA	NA	0.03	0.86