Comparative metabolite fingerprinting of legumes using LC-MS-based untargeted metabolomics

Rafael Llorach [a,b], Claudia Favari [a], David Alonso [a], Mar Garcia-Aloy [a,b], Cristina Andres-Lacueva [a,b], Mireia Urpi-Sarda [a,b,*]

[a] *Biomarkers and Nutrimetabolomics Laboratory, Department of Nutrition, Food Science and Gastronomy, Food Technology Reference Net (XaRTA), Nutrition and Food Safety Research Institute (INSA), Faculty of Pharmacy and Food Science, Campus Torribera, University of Barcelona, 08028 Barcelona, Spain*

[b] *CIBER Fragilidad y Envejecimiento Saludable (CIBERFES), Instituto de Salud Carlos III, 08028 Barcelona, Spain*

* Corresponding author at: Biomarkers and Nutrimetabolomics Laboratory, Department of Nutrition, Food Science and Gastronomy, Food Technology Reference Net (XaRTA), Nutrition and Food Safety Research Institute (INSA), Faculty of Pharmacy and Food Science, Campus Torribera, University of Barcelona, 08028 Barcelona, Spain.
E-mail address: murpi@ub.edu (M. Urpi-Sarda).

Keywords:
Legumes
Metabolomics
Authenticity
Quality control
Food analysis
Discriminant compounds
Mass spectrometry
Phytochemicals

ABSTRACT

Legumes are a well-known source of phytochemicals and are commonly believed to have similar composition between different genera. To date, there are no studies evaluating changes in legumes to discover those compounds that help to discriminate for food quality and authenticity. The aim of this work was to characterize and make a comparative analysis of the composition of bioactive compounds between Cicer arietinum L. (chickpea), Lens culinaris L. (lentil) and Phaseolus vulgaris L. (white bean) through an LC-MS-Orbitrap metabolomic approach to establish which compounds discriminate between the three studied legumes. Untargeted metabolomic analysis was carried out by LC-MS-Orbitrap from extracts of freeze-dried legumes prepared from pre-cooked canned legumes. The metabolomic data treatment and statistical analysis were realized by using MAIT R's package, and final identification and characterization was done using MSn experiments. Fold-change evaluation was made through Metaboanalyst 4.0. Results showed 43 identified and characterized compounds displaying differences between the three legumes. Polyphenols, mainly flavonol and flavanol compounds, were the main group with 30 identified compounds, followed by α-galactosides (n=5). Fatty acyls, prenol lipids, a nucleoside and organic compounds were also characterized. The fold-change analysis showed flavanols as the wider class of discriminative compounds of lentils compared to the other legumes; prenol lipids and eucomic acids were the most discriminative compounds of beans versus other legumes and several phenolic acids (such as primeveroside salycilic), kaempferol derivatives, coumesterol and α-galactosides were the most discriminative compounds of chickpeas. This study highlights the applicability of metabolomics for evaluating which are the characteristic compounds of the different legumes. In addition, it describes the future application of metabolomics as tool for the quality control of foods and authentication of different kinds of legumes.

1. Introduction

Legumes are a habitual part of the diet in several countries worldwide, especially as a source of dietary protein in the developing ones (Caprioli et al., 2016; Curiel et al., 2015; Kalogeropoulos et

al., 2010). In recent years, interest in legumes has increased due to their beneficial or protective effects on human health. Many studies have shown that a frequent consumption of legumes decreases the risk of cardiovascular disease, type 2 diabetes, some types of cancer, overweight and obesity (Curiel et al., 2015; Jenkins et al., 2012). These activities are attributed to the nutritional composition of pulses and their bioactive compounds (Margier et al., 2018). Legumes are known for their high levels in vegetable protein and fiber (Rebello, Greenway, & Finley, 2014). It is to highlight their wide composition of bioactive compounds, such as polyphenols - flavonoids, phenolic acids, tannins – and also triterpenic acids and saponins, among others (Ha et al., 2014). Flavanols have been reported to have nitric oxide-dependent arterial function and immune and inflammatory function modulation (Rodriguez-Mateos et al., 2015); α-galactooligosaccharides such as ciceritol or stachyose were reported to have immunomodulatory activity in vitro (Dai et al., 2018); naturally occurring eucomic acid has been reported to have cytochrome c oxidase activity and to stimulate respiratory functions in vitro in protective anti-aging skin therapies (Simmler, Antheaume, André, Bonté, & Lobstein, 2011). Among European countries, the highest legume consumption is observed around the Mediterranean, with a daily consumption of between 8 and 23 g/capita (Caprioli et al., 2016). The most consumed legumes (i.e. Leguminosae or Fabaceae) are lentils (Lens culinaris Medik.), beans (Phaseolus vulgaris L.) and chickpeas (Cicer arietinum L.). There are several studies using targeted analysis and focused on specific bioactive compounds found in legumes, such as the flavan-3-ols and procyanidins (Bittner, Rzeppa, & Humpf, 2013), flavonoids (Sumner, Paiva, Dixon, & Geno, 1996), isoflavones (Vila-Donat et al., 2015) and soyasaponins (Ha et al., 2014). Additionally, in the last few years, several works have been focused on a specific Leguminosae variety, showing its composition in terms of phytochemicals and major compounds by using mass spectrometry analytical techniques (Abu- Reidah, Arráez-Román, Lozano-Sánchez, Segura-Carretero, & Fernández-Gutiérrez, 2013; Lin, Harnly, Pastor-Corrales, & Luthria, 2008). To the best of our knowledge, there is very little information available on the complete phytochemical profile of common legumes, and additionally, no previous works have compared the phytochemical profile of several legumes using untargeted metabolomic approaches (Caprioli et al., 2016; Curiel et al., 2015; Kalogeropoulos et al., 2010). Over

the last few years, metabolomics approaches have emerged as powerful tools in the field of food sciences. Castro-Puyana et al., reviewed the application of metabolomics in food safety, food quality and food traceability, highlighting the need to develop and apply techniques such as metabolomics that enables to stay abreast with the new requirements of the food market (Castro-Puyana & Herrero, 2013). In addition, the authors concluded that based on their ability to detect new markers, the metabolomics approaches will allow the industry to analyse food quality. Likewise, Cubero-Leon et al., reviewed the application of metabolomics to food authentication. The authors concluded that it is very important to apply untargeted applications in order to enable us to detect new markers to fight against food fraud (Cubero-Leon, Peñalver, & Maquet, 2014). In this context, the aim of this work was to identify, characterize and perform a comparative analysis between Cicer arietinum L. (chickpea), Lens culinaris Medik. (lentil) and Phaseolus vulgaris L. (white bean) through an LC-MS-Orbitrap metabolomic approach to establish which compounds discriminate between the three studied legumes.

2. Materials and methods

2.1. Standards and reagents

The following chemicals were obtained commercially: gallic acid, protocatechuic acid, catechin, p-coumaric, taxifolin, kaempferol, sinapic acid, epigallocatechin and citric acid were purchased from Sigma- Aldrich (St Louis, MO); procyanidin B2, naringin, isoquercitrin and luteolin were purchased from Extrasynthese (Genay, France). HPLCgrade methanol, acetonitrile and formic acid were purchased from Scharlab S.L. (Barcelona, Spain). Ultra-pure water (Milli-Q) was obtained from a Milli-Q system (Millipore, Bedford, MA).

2.2. Sampling and sample preparation

Three pulse samples were selected according to the EU protected geographical indication (PGI). Lentils were “Lenteja Pardina de Tierra de Campos”, white beans were “Mongetes del Ganxet” and chickpeas were “Garbanzo de Fuentesauco”; all of these varieties came from the EU PGI. Pre-cooked

canned Cicer arietinum L. (chickpea from Legumer Precocinados S.L.), Lens culinaris (lentils from Legumbres La Auténtica S.L.) and Phaseolus vulgaris (white bean from Conserves Ferrer S.A.) were selected for use in an intervention study published in the framework of the JPI HDHL Foodball (Madrid-Gambin et al., 2018). A total of 430 g of lentils, 441 g of chickpeas and 564 g of three legumes were washed separately five times using Milli-Q water, then ground and homogenized. The resulting paste was weighed, saved into amber containers and stored for 24 h at −80 °C before the freeze-drying process. Subsequently, samples were placed in the freeze-dryer equipment (Telstar Cryodos, Spain) until dry. Then, each sample was placed in polyethylene bags and stored until analysis.

2.3. Extraction procedure

The extraction procedure was performed following previous methodology reported by Konar et al. and Abu- Reidah et al. with brief modifications (Abu-Reidah, del Mar Contreras, Arráez-Román, Fernández-Gutiérrez, & Segura-Carretero, 2014; Konar, Poyrazoğlu, Demir, & Artik, 2012). In quadruplicate, 1.5 g of each legume powder were mixed with 8 mL of MeOH/H2O (80:20) acidified with 0.5% of formic acid and sonicated for one hour to extract the components. The extracts were centrifuged at 4000 G for 13 min at 4 °C and the resulting supernatants were concentrated using a rotary evaporator under vacuum at 30 °C. Then a second extraction was applied. The residues were resuspended in 8 mL of acidulated MeOH/H2O (80:20), sonicated and centrifuged as before. The resulting supernatants were mixed with the first ones and concentrated up to a volume of 1 mL. The samples were centrifuged at 12000 G for 12 min at 4 °C before the analysis.

2.4. Metabolomics analysis

2.4.1. LC-ESI-LTQ-Orbitrap mass spectrometry

The analysis of bioactive compounds in legumes was carried out by LC-ESI-LTQ-Orbitrap mass spectrometry. Liquid chromatography (LC) was performed on an HPLC Agilent series 1200RR system equipped with a quaternary pump and a thermostatted autosampler. A Phenomenex RP 18

Luna column (50×2.0 mm, 5 μm) was used. A 10 μL full loop injection and a linear gradient elution were performed with a binary system consisting of [A] Milli-Q water with 0.1% HCOOH (v/v) and [B] acetonitrile 0.1% HCOOH (v/v), at a constant flow rate of 600 μL min−1. The gradient elution (v/v) of phase [B] used was as previously reported for a metabolomics approach (Llorach et al., 2013) with slight modifications as follow (time, min; B, %): (0; 1), (5; 40), (6.50; 70), (6.51; 100), (8; 100), (8.10; 1), (12; 1). The HPLC system was online-coupled with an LTQ Orbitrap Velos mass spectrometer (Thermo Scientific, Hemel Hempstead, UK) equipped with an electrospray ionization source working in negative mode (LC-ESI-LTQ-Orbitrap) and coupled to an Accela system (Thermo Scientific, Hemel Hempstead, UK). The ESI-LTQ-Orbitrap data were acquired in FTMS scan mode (scan range from 100 to 2000 m/z) with a resolution of 30,000 fwhm. Operation parameters were as follows: source voltage, 4 (kV); source current, 7 (μA); S-Lens RF level, 94 (%); sheath gas, 50 (arbitrary units); auxiliary gas, 20 (arbitrary units); sweep gas, 2 (arbitrary units); and capillary temperature, 375 °C. The maximum injection time was set at 100 ms with two micro scans for MS mode, and to 1000 ms with one micro scan for MSn mode. Samples were injected in a randomized order jointly with quality controls (QC1: Milli-Q water samples; QC2: standard mixture solution (1 ppm) consisting of gallic acid, protocatechuic acid, catechin, procyanidin B2, p-coumaric, taxifolin, naringin, genistein and kaempferol, QC3: reinjection of one sample for each legume) following the protocol published previously by the research group (Llorach, Urpi-Sarda, Jauregui, Monagas, & Andres- Lacueva, 2009). The coefficient of variation of QC2 (n=6) for all the compounds was lower than 13%. The between day precision (RSD, %) of six significant compounds was calculated (n=8) in different days as additional QC. In beans, stachyose and heliangin showed values of 5.1% and 6.0%, respectively; in lentils, (epi)gallocatechin-(epi)catechin I and megastigmadiene-diol -[apiosyl-glucoside] showed values of 13.3% and naturally occurring eucomic acid has been reported to have cytochrome c oxidase activity and to stimulate respiratory functions in vitro in protective anti-aging skin therapies (Simmler, Antheaume, André, Bonté, & Lobstein, 2011). Among European countries, the highest legume consumption is observed around the Mediterranean, with a daily consumption of between 8 and 23 g/capita (Caprioli et al., 2016). The most consumed legumes (i.e. Leguminosae or

Fabaceae) are lentils (Lens culinaris Medik.), beans (Phaseolus vulgaris L.) and chickpeas (Cicer arietinum L.). There are several studies using targeted analysis and focused on specific bioactive compounds found in legumes, such as the flavan-3-ols and procyanidins (Bittner, Rzeppa, & Humpf, 2013), flavonoids (Sumner, Paiva, Dixon, & Geno, 1996), isoflavones (Vila-Donat et al., 2015) and soyasaponins (Ha et al., 2014). Additionally, in the last few years, several works have been focused on a specific Leguminosae variety, showing its composition in terms of phytochemicals and major compounds by using mass spectrometry analytical techniques (Abu- Reidah, Arráez-Román, Lozano-Sánchez, Segura-Carretero, & Fernández-Gutiérrez, 2013; Lin, Harnly, Pastor-Corrales, & Luthria, 2008). To the best of our knowledge, there is very little information available on the complete phytochemical profile of common legumes, and additionally, no previous works have compared the phytochemical profile of several legumes using untargeted metabolomic approaches (Caprioli et al., 2016; Curiel et al., 2015; Kalogeropoulos et al., 2010). Over the last few years, metabolomics approaches have emerged as powerful tools in the field of food sciences. Castro-Puyana et al., reviewed the application of metabolomics in food safety, food quality and food traceability, highlighting the need to develop and apply techniques such as metabolomics that enables to stay abreast with the new requirements of the food market (Castro-Puyana & Herrero, 2013). In addition, the authors concluded that based on their ability to detect new markers, the metabolomics approaches will allow the industry to analyse food quality. Likewise, Cubero-Leon et al., reviewed the application of metabolomics to food authentication. The authors concluded that it is very important to apply untargeted applications in order to enable us to detect new markers to fight against food fraud (Cubero-Leon, Peñalver, & Maquet, 2014). In this context, the aim of this work was to identify, characterize and perform a comparative analysis between Cicer arietinum L. (chickpea), Lens culinaris Medik. (lentil) and Phaseolus vulgaris L. (white bean) through an LC-MS-Orbitrap metabolomic approach to establish which compounds discriminate between the three studied legumes.

2. Materials and methods

186 *2.1. Standards and reagents*

187 The following chemicals were obtained commercially: gallic acid, protocatechuic acid, catechin, p-
188 coumaric, taxifolin, kaempferol, sinapic acid, epigallocatechin and citric acid were purchased from
189 Sigma- Aldrich (St Louis, MO); procyanidin B2, naringin, isoquercitrin and luteolin were purchased
190 from Extrasynthese (Genay, France). HPLCgrade methanol, acetonitrile and formic acid were
191 purchased from Scharlab S.L. (Barcelona, Spain). Ultra-pure water (Milli-Q) was obtained from a
192 Milli-Q system (Millipore, Bedford, MA).

193 *2.2. Sampling and sample preparation*

194 Three pulse samples were selected according to the EU protected geographical indication (PGI).
195 Lentils were "Lenteja Pardina de Tierra de Campos", white beans were "Mongetes del Ganxet" and
196 chickpeas were "Garbanzo de Fuentesauco"; all of these varieties came from the EU PGI. Pre-cooked
197 canned Cicer arietinum L. (chickpea from Legumer Precocinados S.L.), Lens culinaris (lentils from
198 Legumbres La Auténtica S.L.) and Phaseolus vulgaris (white bean from Conserves Ferrer S.A.) were
199 selected for use in an intervention study published in the framework of the JPI HDHL Foodball
200 (Madrid-Gambin et al., 2018). A total of 430 g of lentils, 441 g of chickpeas and 564 g of three
201 legumes were washed separately five times using Milli-Q water, then ground and homogenized. The
202 resulting paste was weighed, saved into amber containers and stored for 24 h at −80 °C before the
203 freeze-drying process. Subsequently, samples were placed in the freeze-dryer equipment (Telstar
204 Cryodos, Spain) until dry. Then, each sample was placed in polyethylene bags and stored until
205 analysis.

206 *2.3. Extraction procedure*

207 The extraction procedure was performed following previous methodology reported by Konar et al.
208 and Abu- Reidah et al. with brief modifications (Abu-Reidah, del Mar Contreras, Arráez-Román,
209 Fernández-Gutiérrez, & Segura-Carretero, 2014; Konar, Poyrazoğlu, Demir, & Artik, 2012). In
210 quadruplicate, 1.5 g of each legume powder were mixed with 8 mL of MeOH/H2O (80:20) acidified

with 0.5% of formic acid and sonicated for one hour to extract the components. The extracts were centrifuged at 4000 G for 13 min at 4 °C and the resulting supernatants were concentrated using a rotary evaporator under vacuum at 30 °C. Then a second extraction was applied. The residues were resuspended in 8 mL of acidulated MeOH/H2O (80:20), sonicated and centrifuged as before. The resulting supernatants were mixed with the first ones and concentrated up to a volume of 1 mL. The samples were centrifuged at 12000 G for 12 min at 4 °C before the analysis.

2.4. Metabolomics analysis

2.4.1. LC-ESI-LTQ-Orbitrap mass spectrometry

The analysis of bioactive compounds in legumes was carried out by LC-ESI-LTQ-Orbitrap mass spectrometry. Liquid chromatography (LC) was performed on an HPLC Agilent series 1200RR system equipped with a quaternary pump and a thermostatted autosampler. A Phenomenex RP 18 Luna column (50×2.0 mm, 5 μm) was used. A 10 μL full loop injection and a linear gradient elution were performed with a binary system consisting of [A] Milli-Q water with 0.1% HCOOH (v/v) and [B] acetonitrile 0.1% HCOOH (v/v), at a constant flow rate of 600 μL min−1. The gradient elution (v/v) of phase [B] used was as previously reported for a metabolomics approach (Llorach et al., 2013) with slight modifications as follow (time, min; B, %): (0; 1), (5; 40), (6.50; 70), (6.51; 100), (8; 100), (8.10; 1), (12; 1). The HPLC system was online-coupled with an LTQ Orbitrap Velos mass spectrometer (Thermo Scientific, Hemel Hempstead, UK) equipped with an electrospray ionization source working in negative mode (LC-ESI-LTQ-Orbitrap) and coupled to an Accela system (Thermo Scientific, Hemel Hempstead, UK). The ESI-LTQ-Orbitrap data were acquired in FTMS scan mode (scan range from 100 to 2000 m/z) with a resolution of 30,000 fwhm. Operation parameters were as follows: source voltage, 4 (kV); source current, 7 (μA); S-Lens RF level, 94 (%); sheath gas, 50 (arbitrary units); auxiliary gas, 20 (arbitrary units); sweep gas, 2 (arbitrary units); and capillary temperature, 375 °C. The maximum injection time was set at 100 ms with two micro scans for MS mode, and to 1000 ms with one micro scan for MSn mode. Samples were injected in a randomized order jointly with quality controls (QC1: Milli-Q water samples; QC2: standard mixture solution (1

ppm) consisting of gallic acid, protocatechuic acid, catechin, procyanidin B2, p-coumaric, taxifolin, naringin, genistein and kaempferol, QC3: reinjection of one sample for each legume) following the protocol published previously by the research group (Llorach, Urpi-Sarda, Jauregui, Monagas, & Andres- Lacueva, 2009). The coefficient of variation of QC2 (n=6) for all the compounds was lower than 13%. The between day precision (RSD, %) of six significant compounds was calculated (n=8) in different days as additional QC. In beans, stachyose and heliangin showed values of 5.1% and 6.0%, respectively; in lentils, (epi)gallocatechin-(epi)catechin I and megastigmadiene-diol -[apiosyl-glucoside] showed values of 13.3% and 7.5%, respectively; and in chickpeas, kaempferol-diglucoside and ciceritol showed values of 7.7% and 5.3%, respectively. Therefore, these values meet the FDA recommendations for between runs precision (< 15%).

2.4.2. Data processing and statistical analysis

LC-MS data obtained by full scan analysis were processed using MAIT (Metabolite Automatic Identification Toolkit) for the untargeted metabolomic analysis (Fernández-Albert, Llorach, Andrés-Lacueva, and Perera, 2014). MAIT performed feature extraction by peak finding for each sample and alignment using mass and retention time windows for the peaks obtaining spectra for each compound. Then, the application of a non-negative matrix factorization, such as the peak aggregation method, produced a table where the variables were the detected pseudospectra instead of the single mass features (Fernández-Albert, Llorach, Andres-Lacueva, and Perera-Lluna, 2014). The peak picking parameters were: snthresh=5, mzdiff=0.3, retcorrMethod=loess, groupMethod=density, bw=3, mzWid=0.25, filter- Method=matchedFilter, step=0.03, minfrac=0.5. This table was exported to Metabolanalyst (Xia & Wishart, 2016) for the subsequent statistical analysis and metabolomics visualization. The data were log-transformed and pareto-scaled and differences between the metabolomic fingerprint of the three different legumes were analysed by ANOVA followed by the Fisher post hoc test. The metabolomic fingerprint of one legume compared to the other two legumes was also analysed by fold-change analysis followed by t-test analysis. A probability level of $p < .05$ was considered statistically significant. In addition, a principal component analysis (PCA) and a two-

way hierarchical cluster analysis (HCA) were carried out. The two-way HCA was carried out using Pearson's correlation, and aggregation of the observations was performed with Ward's method. A heatmap of intensities was obtained to visualize the legume metabolome differences. The most significant features between legumes went on to be identified and characterized by MSn Orbitrap experiments.

2.4.3. Identification of bioactive compounds by MSn Orbitrap experiments

A multistep procedure combining computational-assisted compound identification and LC-MS pattern analysis was applied. Phytochemicals were tentatively annotated on the basis of their exact mass (< 2 mDa and additional<5 ppm, following criteria based on Directive 2002/ 657/EC) (Gómez-Canela, Ventura, Caixach, & Lacorte, 2014), which was compared to those registered in freely available databases, namely FooDB (http://foodb.ca), MassBank (http://www.massbank.jp), PhytoHub (http://phytohub.eu), Phenol-Explorer 2.0. (http://phenolexplorer.eu) and an in-house database enriched with literature compounds present in legumes. The level of annotation of the compounds was stated in the results following criteria of the Metabolomics Standard Initiative (MSI) (Sumner et al., 2007). The more significant metabolites were characterized by MSn experiments in the Orbitrap with a resolution of 15,000 fwhm. These experiments were carried out by entering manually the parent ions and their main fragments observed in the spectra resulting from the FTMS scan mode analysis. Mass chromatograms and spectral data were acquired using XCalibur software 2.0 (Thermo Scientific, San Jose, CA). The mass spectra pattern was compared with metabolomic databases such as the Human Metabolome Database (HMDB) (www.hmdb.ca) and the in silico fragmentation behaviour using MetFrag (https://msbi.ipbhalle. de/MetFrag), MassBank (http://www.massbank.jp), FooDB (http://foodb.ca) and information from publications.

3. Results and discussion

A total of 43 compounds belonging to various phytochemical classes were tentatively annotated and characterized in the three different legume types using the LC-ESI-LTQ-Orbitrap-MS analytical technique after multi- and univariate statistical analysis (PCA, HCA and ANOVA with Fisher post

289 hoc statistical test). The MSn spectra and fragmentation patterns of these 43 compounds are shown
290 in Table 1. Six of the 43 compounds allowed an identification with level 1 and the other 37 allowed
291 level 2 or level 3 identification following MSI criteria (Table 1) (Sumner et al., 2007). These
292 compounds corresponded to six classes of phytochemicals: polyphenols, α-galactosides, fatty acyls,
293 prenol lipids, nucleosides and organic compounds. As far as we know, this is the first untargeted
294 metabolomic study identifying phytochemical differences between the most consumed legumes in
295 Spain, France and other countries (Marinangeli et al., 2017), and evaluating the fold-changes of the
296 individual legumes compared to the other two legumes for different purposes. Previously, one
297 metabolomic study determined different kinds of compounds in different fractions, such as lipid,
298 sugar, amino acids and amines, although only in mung bean seeds (Na Jom, Frank, & Engel, 2011),
299 and another identified a high number of compounds of different classes in soybean sprouts (Gu et al.,
300 2017). Moreover, from the 43 compounds that were tentatively annotated, six had not been previously
301 identified in these legumes (beans, chickpeas or lentils), however they had been identified in other
302 leguminous species, such as Lathyrus cicera L. (Ferreres et al., 2017), Medicago truncatula (Pollier
303 et al., 2013) or other plant foods.

304 *3.1. Metabolic fingerprinting visualization*

305 Supplementary Fig. 1 depicts the chromatograms of the three legumes and supplementary Fig. 2
306 shows the PCA results with samples coloured according to legume type. The PCA score plot revealed
307 a great separation between the legume classes and showed that in each class, the samples were tightly
308 clustered, but mainly in lentils and in chickpea classes. Another way to analyse and visualize the
309 metabolome differences is to perform an HCA with a heatmap plot (Supplementary Fig. 3). In this
310 context, the HCA using the ANOVA filter data ($p < .05$) classified the samples into two main clusters
311 corresponding to lentils and to the other two legume types. Subsequently, this last cluster was divided
312 into two cluster levels, separating the beans from the chickpeas. Investigation of the clustering
313 behaviour of the features showed that the three classes have specific biomarkers. Likewise, there is

314 an appreciable shared pattern of biomarkers between legume samples. The ANOVA and post hoc
315 results are included in Table 1.

316 *3.2. Polyphenols: identification and changes between legumes*

317 Table 1 shows several subclasses of polyphenols (including flavonols, flavones, flavanols,
318 flavononols or dihydroflavonols, flavones, phenolic acids and stilbenes) putatively annotated in this
319 study, which are different between legumes. Concerning the number of compounds, the flavanols
320 (n=9), phenolic acids (n=9) and flavonols (n=7), respectively, were the most important classes. It is
321 important to highlight the identification of monomers and dimers of flavanol compounds. Compound
322 9 showed an MS/MS behaviour similar to that proposed by the HMDB for prodelphinidin B, therefore
323 this compound can be annotated as prodelphinidin B. Compounds 11, 12, 14, 15 showed a loss of
324 162 amu corresponding to a loss of hexoside moiety (Ferreres et al., 2017). In this context, compounds
325 14 and 15 presented a mass that is 162 amu lower than compounds 11 and 12. In fact, compounds 11
326 and 14 showed a similar MS/MS pattern, presenting both ions at m/z 305 ((epi)gallocatechin moiety).
327 The compounds 12 and 15 also showed a similar MS/MS behaviour but in this case presenting the
328 ion at 289 amu ((epi)catechin moiety). According to this data, these compounds were labelled as
329 (epi)gallocatechin-dihexoside, (epi)catechin-dihexoside, (epi)gallocatechin-hexoside, and
330 (epi)catechin-hexoside, respectively. Nearly all characterized flavanol compounds (9–10 and 12–17)
331 were exclusively of lentils (Table 1) and showed significant differences between lentils and the other
332 two legumes (Fig. 1). In addition, four flavanol compounds in beans (11, 13, 14, 17) and another four
333 in chickpeas (10, 12, 15, 16) had inverse and significant fold-changes between them and lentils, these
334 compounds being useful to discriminate the lentils and their products from the other two legumes.
335 However, the flavonol class was shared in lentils (2, 3, 4, 7), beans (1, 4, 5) and chickpeas (2, 6). The
336 only annotated dihydroflavonol was the compound 19, a glucoside of aromadendrin or carthamidin,
337 which has been identified for the first time in these three legumes. Previously, this last compound
338 was identified in Rhamnus davurica Pall. (Chen, Li, Saleri, & Guo, 2016), however the aglycone
339 aromadendrin and its diglucoside were detected in pulses and carthamidin glycosides were found in

herbs and spices (FoodDB.ca). The compounds of Table 1 have been characterized by MSn experiments and confirmed with matches with spectra from FooDB and/ or the literature. In accordance with our results, the main polyphenols (~70%) in lentils were reviewed as being catechins and procyanidins, while flavonol compounds were present in 17% of total polyphenols in raw lentils and only 4% in pinto beans (Singh, Singh, Kaur, & Singh, 2017). In this sense, a previous study evaluating 20 Canadian lentil cultivars showed flavanol and flavonol compounds as the main phenolic compounds that contribute to the strong antioxidant activity of lentils (Zhang et al., 2015), and contributed to discriminating this legume from the others. Phenolic acids were the second major group of compounds identified in this study. Beans presented five characteristic phenolic acids, with hydroxyeucomic acid (22) and eucomic acid (25) being those compounds with higher fold-changes compared to the other two legumes. With regard to the characteristic compounds of chickpeas compared to the other two legumes, the presence of primeveroside salicylic acid (23), a sinapic isomer (26) and protocatechuic acid glucoside (21) should be highlighted. This is the first time that primeveroside salycilic acid has been identified as a discriminant compound of chickpeas, although it was previously identified in green beans (Abu- Reidah et al., 2013). However, the only significant phenolic acid in lentils compared to beans and chickpeas was the uralenneoside, a phenolic acid that was previously identified in herbs and species (FooDB.ca) and this is the first time it has been identified in lentils (Table 1; Fig. 1). In lentils, the presence should also be highlighted of the stilbene resveratrol glucoside, as has been previously published (Dueñas, Hernández, & Estrella, 2007). Nearly all putative annotated flavonoids and non-flavonoids were previously detected in some of these legumes except for compound 19, and 24, which have been characterized for the first time in these legumes. Nevertheless, this new compound 19 has been identified previously in other Leguminosae varieties, such as Afzelia bella (Binutu & Cordell, 2001).

3.3. α-Galactosides: identification and changes between legumes

This class is the second most important class identified and characterized in legumes with five putative annotated compounds. The levels of five α-galactosides showed statistically significant

differences between the three legumes. Three of them, putatively annotated as galactopyranosylciceritol 31, ciceritol 32, and galactopinitol 33 had significant higher levels in chickpeas than the other two legumes (Fig. 1), while stachyose (34) and raffinose (35) showed higher levels in beans and lower levels in lentils compared to the other legumes. In line with our results, Sanchez-Mata et al., also demonstrated the higher amounts of ciceritol in chickpeas than in lentils. They also observed higher levels of raffinose and stachyose in beans, which are responsible for the flatulence associated with legumes, which represented a 50% of the total sugar in white beans compared to 22% in chickpeas (Sánchez- Mata, Peñuela-Teruel, Cámara-Hurtado, Díez-Marqués, & Torija-Isasa, 1998).

3.4. Fatty acyls: identification and changes between legumes

Three fatty acyls were characterized in legumes for the first time in this metabolomic approach. Previously they had been described in some fruits such as loquat (FooDB.ca). Mass spectra were confirmed by comparison with those published in the FooDB database and losses of −132 amu and consecutive losses of 132 and 162 amu, corresponded to a loss of a pentose and two consequential losses of a pentose and a hexose, respectively. The compound 36 showed a significant fold-change for chickpeas compared to the other legumes. Moreover, the amount of its isomer (37) and the other fatty acyl (38) were significantly higher in lentils (Fig. 1). The 37 and 38 compounds showed negative fold-changes for beans compared to lentils and chickpeas.

3.5. Prenol lipids: identification and changes between legumes

The two putatively annotated compounds (39 and 40) were specific to beans and showed higher amounts in beans than in the other two legumes (Fig. 1). Their identification has been confirmed by a comparison of MSn spectra with those published mainly in FooDB. This is the first time that helinagin (39) has been putatively annotated in these legumes. Previously, heliangin, a sesquiterpene lactone, was found in the leaves of Helianthus tuberosus L. (Ahmed, El-Sakhawy, Soliman, & Abou-Hussein, 2005). Otherwise, gibberellin compounds (40) have been detected previously in peas, lentils and several species of beans, as stated in FooDB. They are a class of phytohormones involved in the

maturation of legume nodules and other biological processes (Hayashi, Gresshoff, & Ferguson, 2014).

3.6. Nucleosides and organic compounds: identification and changes between legumes

In this group of compounds, one nucleoside (41) and two organic compounds (42 and 43) were putatively annotated based on the fragmentation pattern published in HMDB or by comparison with authentic standard (citric acid). The levels of pseudouridine (41) in beans were found in higher amounts than in both lentils and chickpeas. Previous studies purified the enzymes catalyzing uridinediphosphate glucose in mung bean seedlings (Kaushal & Elbein, 1986), which could allow the presence of these compounds in plants. FooDB also showed that pseudouridine has been detected in these legumes previously. Although it had statistical significance, the fold-change obtained in beans was lower than flavonoids but higher than α-galactosides. With regard to organic compounds, citric acid (42) was identified and was statistically higher in beans than in the other legumes. Although it appears in the metabolome of beans, in this case it is an additive of canned beans added during the manufacturing. So, it was not incorporated in Fig. 1. Legumes contain high levels of proteins (22–29%) and lentils have been described as being rich in lysine and leucine along with other legumes (Roy, Boye, & Simpson, 2010). We found phenylalanyl-leucine (43) as a dipeptide exclusive to chickpeas. Although it had the higher fold-change in chickpeas with respect to the other legumes, it is a dipeptide that can be found in a high number of foods.

4. Conclusions

In conclusion, this is the first metabolomic study comparing the bioactive compounds of the three most consumed legumes: beans, chickpeas and lentils. A total of 43 compounds were identified, putatively annotated and characterized based mainly on their accurate mass measurement from LTQ-Orbitrap, MSn experiments, as well as comparison with reference standards when available and with specialized databases and literature. From the total annotated compounds, 40% were exclusive to lentils and 30% to beans, while only 26% was exclusive to chickpeas. The fold-change evaluation has shown flavanol derivatives as the main compounds that differentiate lentils from the other

legumes. In addition, resveratrol glucoside and two megastigmadiene-diol -[apiosyl-glucoside] compounds were also discriminant compounds of lentils. Beans showed higher changes for phenolic acids highlighting eucomic and hydroxyeucomic acids followed by the two prenol lipids heliangin and gibberellin. Chickpeas can be highlighted for their higher levels of phenylalanyl-leucine, primeveroside salicylic acid and two kaempferol derivates, but also for their levels of coumesterol compared to the other legumes. This comparative study helps to discriminate which compounds could be different among certain legume consumption and provides important information to contribute to building up the metabolomics databases. This study highlights metabolomics for future applications as a tool for the quality control of foods and the authentication of different kinds of legumes and their products.

Acknowledgements

This research was supported by the EU Joint Programming Initiative 'A Healthy Diet for a Healthy Life' on Biomarkers BioNHFOODBALL (PCIN-2014-133 Ministry of Economy and Competitiveness (MINECO) Spain) and the 2017SGR1546/AGAUR and ICREA Academia awards from the Generalitat de Catalunya.

Declaration of Competing Interest

The authors declare that they have no conflict of interest.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.foodres.2019.108666.

References

Abu-Reidah, I. M., Arráez-Román, D., Lozano-Sánchez, J., Segura-Carretero, A., & Fernández-Gutiérrez, A. (2013). Phytochemical characterisation of green beans (Phaseolus vulgaris L.) by using

441 high-performance liquid chromatography coupled with time-of-flight mass spectrometry.
442 Phytochemical Analysis, 24(2), 105–116.

443 Abu-Reidah, I. M., Arráez-Román, D., Quirantes-Piné, R., Fernández-Arroyo, S., Segura- Carretero,
444 A., & Fernández-Gutiérrez, A. (2012). HPLC-ESI-Q-TOF-MS for a comprehensive characterization
445 of bioactive phenolic compounds in cucumber whole fruit extract. Food Research International,
446 46(1), 108–117.

447 Abu-Reidah, I. M., Arráez-Román, D., Warad, I., Fernández-Gutiérrez, A., & Segura- Carretero, A.
448 (2017). UHPLC/MS2-based approach for the comprehensive metabolite profiling of bean (Vicia faba
449 L.) by-products: A promising source of bioactive constituents. Food Research International, 93, 87–
450 96.

451 Abu-Reidah, I. M., del Mar Contreras, M., Arráez-Román, D., Fernández-Gutiérrez, A., & Segura-
452 Carretero, A. (2014). UHPLC-ESI-QTOF-MS-based metabolic profiling of Vicia faba L. (Fabaceae)
453 seeds as a key strategy for characterization in foodomics. Electrophoresis, 35(11), 1571–1581.

454 Ahmed, M. S., El-Sakhawy, F. S., Soliman, S. N., & Abou-Hussein, D. M. R. (2005). Phytochemical
455 and biological study of Helianthus tuberosus L. Section Title: Plant Biochemistry, 18, 134–147.

456 Binutu, O. A., & Cordell, G. A. (2001). Constituents of Afzelia bella stem bark. Phytochemistry,
457 56(8), 827–830.

458 Bittner, K., Rzeppa, S., & Humpf, H. U. (2013). Distribution and quantification of flavan- 3-ols and
459 procyanidins with low degree of polymerization in nuts, cereals, and legumes. Journal of Agricultural
460 and Food Chemistry, 61(38), 9148–9154.

461 Caprioli, G., Giusti, F., Ballini, R., Sagratini, G., Vila-Donat, P., Vittori, S., & Fiorini, D. (2016).
462 Lipid nutritional value of legumes: Evaluation of different extraction methods and determination of
463 fatty acid composition. Food Chemistry, 192, 965–971.

464 Castro-Puyana, M., & Herrero, M. (2013). Metabolomics approaches based on mass spectrometry for
465 food safety, quality and traceability. TrAC, Trends in Analytical Chemistry, 52, 74–87.

466 Chahdoura, H., Barreira, J. C. M., Barros, L., Santos-Buelga, C., Ferreira, I. C. F. R., & Achour, L.
467 (2014). Phytochemical characterization and antioxidant activity of the cladodes of Opuntia
468 macrorhiza (Engelm.) and Opuntia microdasys (Lehm.). Food & Function, 5(9), 2129–2136.

469 Chen, G., Li, X., Saleri, F., & Guo, M. (2016). Analysis of flavonoids in Rhamnus davurica and its
470 antiproliferative activities. Molecules, 21(10), E1275.

471 Cubero-Leon, E., Peñalver, R., & Maquet, A. (2014). Review on metabolomics for food
472 authentication. Food Research International, 60, 95–107.

473 Curiel, J. A., Coda, R., Centomani, I., Summo, C., Gobbetti, M., & Rizzello, C. G. (2015).
474 Exploitation of the nutritional and functional characteristics of traditional Italian legumes: The
475 potential of sourdough fermentation. International Journal of Food Microbiology, 196, 51–61.

476 Dai, Z., Lyu, W., Xiang, X., Tang, Y., Hu, B., Ou, S., & Zeng, X. (2018). Immunomodulatory effects
477 of enzymatic-synthesized α-galactooligosaccharides and evaluation of the structure-activity
478 relationship. Journal of Agricultural and Food Chemistry, 66(34), 9070–9079.

479 Dueñas, M., Hernández, T., & Estrella, I. (2007). Changes in the content of bioactive polyphenolic
480 compounds of lentils by the action of exogenous enzymes. Effect on their antioxidant activity. Food
481 Chemistry, 101(1), 90–97.

482 Fernández-Albert, F., Llorach, R., Andrés-Lacueva, C., & Perera, A. (2014). An R package to analyse
483 LC/MS metabolomic data: MAIT (Metabolite Automatic Identification Toolkit). Bioinformatics,
484 30(13), 1937–1939.

485 Fernández-Albert, F., Llorach, R., Andres-Lacueva, C., & Perera-Lluna, A. (2014). Peak aggregation
486 as an innovative strategy for improving the predictive power of LC-MS metabolomic profiles.
487 Analytical Chemistry, 86(5), 2320–2325.

488 Ferreres, F., Magalhães, S. C. Q., Gil-Izquierdo, A., Valentão, P., Cabrita, A. R. J., Fonseca, A. J.
489 M., & Andrade, P. B. (2017). HPLC-DAD-ESI/MSn profiling of phenolic compounds from Lathyrus
490 cicera L. seeds. Food Chemistry, 214, 678–685.

491 Gómez-Canela, C., Ventura, F., Caixach, J., & Lacorte, S. (2014). Occurrence of cytostatic
492 compounds in hospital effluents and wastewaters, determined by liquid chromatography coupled to
493 high-resolution mass spectrometry. Analytical and Bioanalytical Chemistry, 406, 3801–3814.

494 Gu, E. J., Kim, D. W., Jang, G. J., Song, S. H., Lee, J. I., Lee, S. B., & Kim, H. J. (2017). Mass-based
495 metabolomic analysis of soybean sprouts during germination. Food Chemistry, 217, 311–319.

496 Ha, T. J., Lee, B. W., Park, K. H., Jeong, S. H., Kim, H. T., Ko, J. M., & Lee, J. H. (2014). Rapid
497 characterisation and comparison of saponin profiles in the seeds of Korean leguminous species using
498 ultra performance liquid chromatography with photodiode array detector and electrospray
499 ionisation/mass spectrometry (UPLC-PDA-ESI/MS) analysis. Food Chemistry, 146, 270–277.

500 Hayashi, S., Gresshoff, P. M., & Ferguson, B. J. (2014). Mechanistic action of gibberellins in legume
501 nodulation. Journal of Integrative Plant Biology, 56(10), 971–978.

502 Jenkins, D. J. A., Kendall, C. W. C., Augustin, L. S. A., Mitchell, S., Sahye-Pudaruth, S., Blanco
503 Mejia, S., & Josse, R. G. (2012). Effect of legumes as part of a low glycemic index diet on glycemic
504 control and cardiovascular risk factors in type 2 diabetes mellitus: A randomized controlled trial.
505 Archives of Internal Medicine, 172(21), 1653–1660.

506 Jin, A., Ozga, J. A., Lopes-Lutz, D., Schieber, A., & Reinecke, D. M. (2012). Characterization of
507 proanthocyanidins in pea (Pisum sativum L.), lentil (Lens culinaris L.), and faba bean (Vicia faba L.)
508 seeds. Food Research International, 46, 528–535.

509 Kalogeropoulos, N., Chiou, A., Ioannou, M., Karathanos, V. T., Hassapidou, M., & Andrikopoulos,
510 N. K. (2010). Nutritional evaluation and bioactive microconstituents (phytosterols, tocopherols,
511 polyphenols, triterpenic acids) in cooked dry legumes usually consumed in the Mediterranean
512 countries. Food Chemistry, 121(3), 682–690.

513 Kaushal, G. P., & Elbein, A. D. (1986). Purification and properties of UDPGlcNAc: Dolichyl-
514 pyrophosphoryl-GlcNAc GlcNAc transferase from mung bean seedling. Plant Physiology, 82(3),
515 748–752.

516 Konar, N., Poyrazoğlu, E. S., Demir, K., & Artik, N. (2012). Determination of conjugated and free
517 isoflavones in some legumes by LC-MS/MS. Journal of Food Composition and Analysis, 25(2), 173–
518 178.

519 Lin, L. Z., Harnly, J. M., Pastor-Corrales, M. S., & Luthria, D. L. (2008). The polyphenolic profiles
520 of common bean (Phaseolus vulgaris L.). Food Chemistry, 107(1), 399–410.

521 Llorach, R., Urpi-Sarda, M., Jauregui, O., Monagas, M., & Andres-Lacueva, C. (2009). An LC-MS-
522 based metabolomics approach for exploring urinary metabolome modifications after cocoa
523 consumption. Journal of Proteome Research, 8(11), 5060–5068.

524 Llorach, R., Urpi-Sarda, M., Tulipani, S., Garcia-Aloy, M., Monagas, M., & Andres- Lacueva, C.
525 (2013). Metabolomic fingerprint in patients at high risk of cardiovascular disease by cocoa
526 intervention. Molecular Nutrition & Food Research, 57(6), 962–973.

527 Madrid-Gambin, F., Brunius, C., Garcia-Aloy, M., Estruel-Amades, S., Landberg, R., & Andres-
528 Lacueva, C. (2018). Untargeted 1H NMR-based metabolomics analysis of urine and serum profiles
529 after consumption of lentils, chickpeas, and beans: An extended meal study to discover dietary
530 biomarkers of pulses. Journal of Agricultural and Food Chemistry, 66(27), 6997–7005.

531 Margier, M., Georgé, S., Hafnaoui, N., Remond, D., Nowicki, M., Du Chaffaut, L., ... Reboul, E.
532 (2018). Nutritional composition and bioactive content of legumes:

533 Characterization of pulses frequently consumed in France and effect of the cooking method.
534 Nutrients, 10, 1668.

Marinangeli, C. P. F., Curran, J., Barr, S. I., Slavin, J., Puri, S., Swaminathan, S., & Patterson, C. A. (2017). Enhancing nutrition with pulses: Defining a recommended serving size for adults. Nutrition Reviews, 75(12), 990–1006.

Na Jom, K., Frank, T., & Engel, K. H. (2011). A metabolite profiling approach to follow the sprouting process of mung beans (Vigna radiata). Metabolomics, 7(1), 102–117.

Pollier, J., Moses, T., González-Guzmán, M., De Geyter, N., Lippens, S., Vanden Bossche, R., ... Goossens, A. (2013). The protein quality control system manages plant defence compound synthesis. Nature, 504(7478), 148–152.

Rebello, C. J., Greenway, F. L., & Finley, J. W. (2014). A review of the nutritional value of legumes and their effects on obesity and its related co-morbidities. Obesity Reviews, 15(5), 392–407.

Rodriguez-Mateos, A., Hezel, M., Aydin, H., Kelm, M., Lundberg, J. O., Weitzberg, E., & Heiss, C. (2015). Interactions between cocoa flavanols and inorganic nitrate: Additive effects on endothelial function at achievable dietary amounts. Free Radical Biology and Medicine, 80, 121–128.

Roy, F., Boye, J. I., & Simpson, B. K. (2010). Bioactive proteins and peptides in pulse crops: Pea, chickpea and lentil. Food Research International, 43(2), 432–442.

Sánchez-Mata, M. C., Peñuela-Teruel, M. J., Cámara-Hurtado, M., Díez-Marqués, C., & Torija-Isasa, M. E. (1998). Determination of mono-, di-, and oligosaccharides in legumes by high-performance liquid chromatography using an amino-bonded silica column. Journal of Agricultural and Food Chemistry, 46(9), 3648–3652.

Simmler, C., Antheaume, C., André, P., Bonté, F., & Lobstein, A. (2011). Glucosyloxybenzyl eucomate derivatives from Vanda teres stimulate HaCaT cytochrome c oxidase. Journal of Natural Products, 74, 949–955.

Singh, B., Singh, J. P., Kaur, A., & Singh, N. (2017). Phenolic composition and antioxidant potential of grain legume seeds: A review. Food Research International, 101, 1–16.

559 Sumner, L. W., Amberg, A., Barrett, D., Beale, M. H., Beger, R., Daykin, C. A., & Viant, M. R.
560 (2007). Proposed minimum reporting standards for chemical analysis: Chemical Analysis Working
561 Group (CAWG) Metabolomics Standards Initiative (MSI). Metabolomics, 3(3), 211–221.

562 Sumner, L. W., Paiva, N. L., Dixon, R. A., & Geno, P. W. (1996). High-performance liquid
563 chromatography/continuous-flow liquid secondary ion mass spectrometry of flavonoid glycosides in
564 leguminous plant extracts. Journal of Mass Spectrometry, 31(5), 472–485.

565 Urpi-Sarda, M., Boto-Ordóñez, M., Queipo-Ortuño, M. I., Tulipani, S., Corella, D., Estruch, R., &
566 Andres-Lacueva, C. (2015). Phenolic and microbial-targeted metabolomics to discovering and
567 evaluating wine intake biomarkers in human urine and plasma. Electrophoresis, 36(18), 2259–2268.

568 Vila-Donat, P., Caprioli, G., Maggi, F., Ricciutelli, M., Torregiani, E., Vittori, S., & Sagratini, G.
569 (2015). Effective clean-up and ultra high-performance liquid chromatography- tandem mass
570 spectrometry for isoflavone determination in legumes. Food Chemistry, 174, 487–494.

571 Wang, S., Liu, L., Wang, L., Hu, Y., Zhang, W., & Liu, R. (2012). Structural characterization and
572 identification of major constituents in jitai tablets by high-performance liquid chromatography/diode-
573 array detection coupled with electrospray ionization tandem mass spectrometry. Molecules, 17(9),
574 10470–10493.

575 Xia, J., & Wishart, D. S. (2016). Using metaboanalyst 3.0 for comprehensive metabolomics data
576 analysis. Current Protocols in Bioinformatics, 55, 14.10.1–14.10.91.

577 Zhang, B., Deng, Z., Ramdath, D. D., Tang, Y., Chen, P. X., Liu, R., & Tsao, R. (2015). Phenolic
578 profiles of 20 Canadian lentil cultivars and their contribution to antioxidant activity and inhibitory
579 effects on α-glucosidase and pancreatic lipase. Food Chemistry, 172, 862–872.

580 Zhou, Y., Xu, D. S., Liu, L., Qiu, F. R., Chen, J. L., & Xu, G. L. (2016). A LC-MS/MS method for
581 the determination of stachyose in rat plasma and its application to a pharmacokinetic study. Journal
582 of Pharmaceutical and Biomedical Analysis, 123, 24–30.

583 **FIGURES**

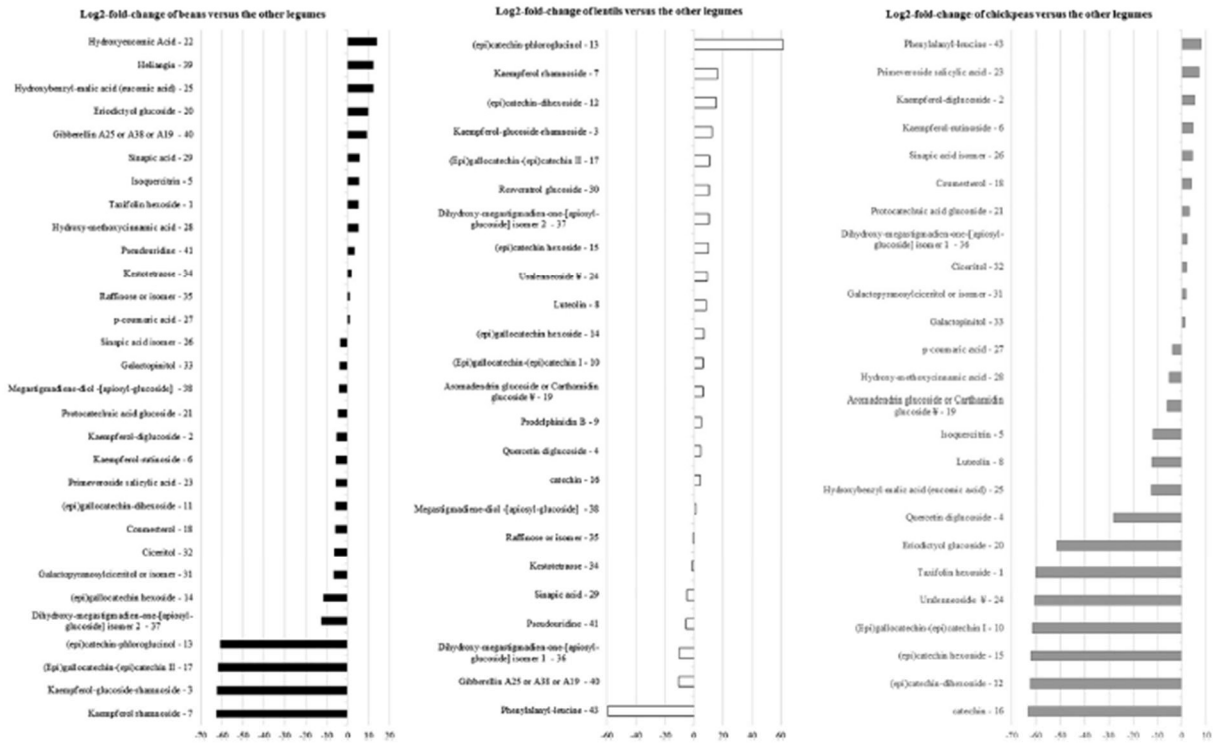

Fig. 1. Significant changes in the metabolome of legumes: a) beans compared to the other two legumes; b) lentils compared to the other two legumes; and c) chickpeas compared to the other two legumes.
Note: Statistical analysis: all the changes in the figures are significantly different, $P < .05$.

584

585

586

587

TABLES

Table 1
Identification of bioactive compounds in legumes.

n	Retention time	Observed Mass [M-H]-	Theoretical Mass [M-H]-	Absolute error (mDa)	Molecular formula [M-H]-	Potential compound (level of identification)	MS/MS and MSn	Statistical analysis	Identified based on spectra published by
Flavonoids: Flavonols									
1	2.81	465.1013	465.1033	2.0	$C_{21}H_{21}O_{12}$	Taxifolin hexoside (level 2)	Ms2: 303.0503 [M-162-H] (90); 285.0399 [M-162-18-H] (100) Ms3 of 285: 241.0501 (100); 217.0501 (20); 199.0395 (30); 175.0397 (70)	Beans [a] Lentils [b] Chickpeas [c]	(Abu-Reidah, Arráez-Román, Warad, Fernández-Gutiérrez, & Segura-Carretero, 2017)
2	2.92	609.1462	609.1455	−0.7	$C_{27}H_{29}O_{16}$	Kaempferol-diglucoside (level 2)	Ms2: 447.0909 [M-162-H]- (70); 285.0387 [M-162-162-H]- (100) MS3 of 285: 257.0439 (80); 241.0492 (40);229.0493 (20); 213.0544 (20); 151.0027 (100)	Chickpeas [a] Lentils [a] Beans [b]	(Ferreres et al., 2017)
3	3.25	593.1505	593.1506	0.1	$C_{27}H_{29}O_{15}$	Kaempferol-glucoside-rhamnoside (level 2)	Ms2: 447.0911 [M-146-H] (40); 431.0963 [M-162-H] (10); 327.0493 [M-266-H] (20); 285.0389 [M-146-162-H] (100) MS3 of 285: 267.0285 [M-146-162-18] (40); 257.0439 (100); 241.0493 (40); 229.0493 60); 213.0545 (30); 151.0029 (100)	Lentils [a] Chickpeas [b] Beans [c]	(Abu-Reidah et al., 2012)
4	3.28	625.1414	625.1405	−0.9	$C_{27}H_{29}O_{17}$	Quercetin diglucoside (level 2)	Ms2: 463.0857 [M-162-H] (100) Ms3 of 463: 301.0334 [M-162-162-H] (100)	Lentils [a] Beans [a] Chickpeas [b]	FooDB.ca
5	3.31	463.0874	463.0876	0.2	$C_{21}H_{19}O_{12}$	Isoquercitrin (level 1)	Confirmed with mass spectra and std.	Beans [a] Chickpeas [b] Lentils [b]	
6	3.47	593.1505	593.1506	0.1	$C_{27}H_{29}O_{15}$	Kaempferol-rutinoside (level 2)	Ms2: 285.0387 [M-146-H]-(100) MS3 of 285: 257.0441 (100); 241.0491 (30); 229.0494 (60); 213.0545 (30); 151.0028 (10)	Chickpeas [a] Lentils [a] Beans [b]	(Ferreres et al., 2017)
7	4.41	431.0960	431.0978	1.8	$C_{21}H_{19}O_{10}$	Kaempferol rhamnoside (level 2)	Ms2: 285.0400 [M-146-H] (100); 257.0452 (5); 241.0503 (2); 229.0484 (5); 213.0552 (5); 151.0034 (10)	Lentils [a] Chickpeas [b] Beans [b]	(Abu-Reidah et al., 2012)
Flavonoids: Flavones									
8	4.39	285.0388	285.0399	1.1	$C_{15}H_9O_6$	Luteolin (level 1)	Confirmed with mass spectra and std.	Lentils [a] Beans [a] Chickpeas [b]	
Flavonoids: Flavanols									
9	0.87	609.1235	609.1244	0.9	$C_{30}H_{25}O_{14}$	Prodelphinidin B (level 3)	Ms2: 591.1142 [M-18H] (15); 565.1385 [M-44-H] (2); 483.0925 [M-126-H] (40); 441.0819 [M-126-42-H] (100); 423.0714 [M-126-42-18-H] (80); 305.0658 (60)	Lentils [a] Chickpeas [b] Beans [b]	(Abu-Reidah et al., 2014)
10	1.1	593.1286	593.1294	0.8	$C_{30}H_{25}O_{13}$	(epi)gallocatechin-(epi)catechin I (level 2)	Ms2: 467.0961 [M-126-H] (50); 425.0858 (100); 407.0755 (40); 303.0494 (15); 289.0701 (80)	Lentils [a] Beans [b] Chickpeas [c]	(Abu-Reidah et al., 2014)

(continued on next page)

Table 1 (*continued*)

n	Retention time	Observed Mass [M-H]-	Theoretical Mass [M-H]-	Absolute error (mDa)	Molecular formula [M-H]-	Potential compound (level of identification)	MS/MS and MSn	Statistical analysis	Identified based on spectra published by
11	1.3	629.1718	629.17229	0.5	$C_{27}H_{34}O_{17}$	(epi)gallocatechin-dihexoside (level 2)	Ms2: 467.1170 [M-162-H] (100); 305.0648 [M-162-162-H] (60) Ms3 of 305: 261.0748 (30); 221.0439 (90); 219.0646 (60); 179.0337 (100); 125.0237 (25)	Lentils [a] Chickpeas [a] Beans [b]	(Abu-Reidah et al., 2014)
12	1.55	613.1774	613.1768	− 0.6	$C_{27}H_{34}O_{16}$	(epi)catechin-dihexoside (level 2)	Ms2: 451.1239 [M-162-H] (100); 299.0769 [M-162-152-H] (5); 289.0712 [M-162-162-H] (50) MS3 of 289: 245.0807 (100)	Lentils [a] Chickpeas [b] Beans [b]	(Abu-Reidah et al., 2014)
13	1.64	413.0853	413.0872	1.9	$C_{21}H_{17}O_{8}$	(epi)catechin-phloroglucinol (level 2)	Ms2: 287.0556 [M-126-H] (100); 261.0399 [M-152-H] (30); 161.0242 [M-126-126-H] (10); 125.0242 (20)	Lentils [a] Chickpeas [b] Beans [c]	(Jin, Ozga, Lopes-Lutz, Schieber, & Reinecke, 2012)
14	1.86	467.1169	467.1189	2.0	$C_{21}H_{23}O_{12}$	(epi)gallocatechin hexoside (level 2)	Ms2: 305.0657 [M-162-H] (100) Ms3 of 305: 287.0557 [M-162- 18-H] (10) 261.0765 [M-162-44-H] (40) 221.0452 (80); 179.0346 (100); 165.0190 (30); 137.0241 (25); 125.0242 (35)	Lentils [a] Chickpeas [a] Beans [b]	(Abu-Reidah et al., 2014)
15	1.89	451.1223	451.1240	1.7	$C_{21}H_{23}O_{11}$	(epi)catechin hexoside (level 2)	Ms2: 289.0712 [M-162-H] (100); 271.0606 [M-162-18-H] (40); 245.0815 [M-162-44-H] (10); 137.0241(50) Ms3 of 289: 245.0814 [M-162-44-H] (10); 205.0502 (35); 179.0343 (15); 137.0242 (5); 125.0241 (5)	Lentils [a] Beans [b] Chickpeas [c]	(Abu-Reidah et al., 2014)
16	2.09	289.0710	289.0712	0.2	C15H13O6	Catechin (level 1)	Confirmed with mass spectra and std.	Lentils [a] Beans [b] Chickpeas [c]	
17	2.32	593.1286	593.1294	0.8	$C_{30}H_{25}O_{13}$	(epi)gallocatechin -(epi)catechin II (level 2)	Ms2: 575.1171 (30); 467.0960 [M-126-H] (100); 449.0855 (50); 425.0856 (20); 407.0753 (30); 331.0805 (50); 289.0700 (90); 285.0389 (40); 261.0389 (70); 245.0805 (15)	Lentils [a] Chickpeas [b] Beans [b]	(Abu-Reidah et al., 2014)
Isoflavonoids: Coumestrans									
18	4.65	267.0294	267.0293	− 0.1	$C_{15}H_{7}O_{5}$	Coumesterol (level 2)	Ms2: 252.0414 [M-15-H] (10); 239.0337 [M-28-H] (60); 211.0389 (10); 223.0389 [M-44-H]- (100); 195.0440[M-44-28-H]- (30)	Chickpeas [a] Lentils [a] Beans [b]	HMDB.ca
Flavonoids: Flavononols or dihydroflavonols									
19	2.6	449.1085	449.1084	− 0.1	$C_{21}H_{21}O_{11}$	Aromadendrin glucoside or Garthamidin glucoside [e] (level 3)	Ms2: 287.0544 [M-162-H] (100) Ms3 of 287: 259.0594 (100); 243.0647 (20); 201.0543 (10)	Lentils [a] Beans [b] Chickpeas [c]	FooDB.ca (Chen et al., 2016; Wang et al., 2012)
Flavonoids: Flavanones									
20	3.72	449.1085	449.1084	− 0.1	$C_{21}H_{21}O_{11}$	Eriodictyol glucoside (level 2)	Ms2: 287.0555 [M-162-H] (100) Ms3 of 287: 151.0028 (100); 135.0443 (10); 107.0132 (5) Ms4 of 151: 107.0131 (100)	Beans [a] Lentils [b] Chickpeas [c]	FooDB.ca

(*continued on next page*)

Table 1 (*continued*)

n	Retention time	Observed Mass [M-H]-	Theoretical Mass [M-H]-	Absolute error (mDa)	Molecular formula [M-H]-	Potential compound (level of identification)	MS/MS and MSn	Statistical analysis	Identified based on spectra published by
Phenolic Acids									
21	1.12	315.0703	315.0716	1.3	$C_{13}H_{15}O_9$	Protocatechuic acid glucoside (level 2)	Ms2: 153.0191 [M-162-H] (100); 152.0113 (50); 109.0293 [M-162-44-H] (15)	Chickpeas [a] Lentils [b] Beans [c]	FooDB.ca
22	1.2	255.0499	255.0505	0.6	$C_{11}H_{11}O_7$	Hydroxyeucomic Acid (level 2)	Ms2: 193.0503 [M-18-44-H] (40); 165.0554 (100); 149.0605 (10)	Beans [a] Lentils [b] Chickpeas [b]	
23	1.33	431.1187	431.1194	0.7	$C_{18}H_{23}O_{12}$	Primeveroside salicylic acid (level 2)	Ms2: 299.0768 [M-132-H] (30) 137.0241[M-132-162-H] (100) Ms3 299: 179.0343[M-132-120-H] (100)	Chickpeas [a] Lentils [b] Beans [b]	(Abu-Reidah et al., 2013)
24	1.75	285.0607	285.0610	0.3	$C_{12}H_{13}O_8$	Uralenneoside [*] (level 2)	Ms2: 153.0190 [M-132-H] (100) 109.0292 [M-132-44-H]- (10)	Lentils [a] Beans [b] Chickpeas [c]	FooDB.ca
25	2.24	239.0549	239.0556	0.7	$C_{11}H_{11}O_6$	Hydroxybenzyl-malic acid (eucomic acid) (level 2)	Ms2: 195.0660 [M-44-H] (20); 179.0346 (100); 177.0555 [M-44-18-H]- (50); 149.0605 (70); 107.0500 (10) Ms3 of 179: 107.0499 (100) Ms3 of 177: 133.0656[M-18-44-44]- (100)	Beans [a] Lentils [b] Chickpeas [b]	(Chahdoura et al., 2014)
26	2.31	223.0604	223.0606	−0.2	$C_{11}H_{11}O_5$	Sinapic acid isomer (level 2)	Ms2: 208.0366 [M-15-H]- (100); 179.0704 [M-44-H]- (30); 164.0469 0704 [M-44-15-H]- (10)	Chickpeas [a] Lentils [b] Beans [b]	FooDB.ca
27	2.87	163.0394	163.0395	−0.1	$C_9H_7O_3$	p-coumaric acid (level 1)	Confirmed with mass spectra and std.	Beans [a] Lentils [a] Chickpeas [b]	
28	3.17	193.0497	193.0506	−0.9	$C_{10}H_9O_4$	Hydroxy- methoxycinnamic acid (level 2)	Ms2: 178.0267 [M-15-H] (80); 149.0604 [M-44-H] (100); 134.0370 [M-44-15-H] (30)	Beans [a] Lentils [b] Chickpeas [c]	FooDB.ca
29	3.23	223.0607	223.0612	−0.5	$C_{11}H_{11}O_5$	Sinapic acid (level 1)	Confirmed with mass spectra and std.	Beans [a] Chickpeas [b] Lentils [c]	
Stilbenes									
30	3.17	389.1225	389.1236	−1.1	$C_{20}H_{21}O_8$	Resveratrol glucoside (level 2)	Ms2: 227.0710 [M-162-H] (70); 209.1176 [M-180-H] (100) Ms3 227: 185.0603 [M-162-42-H] (100); 143.0496 [M-162- 42-42-H] (20)	Lentils [a] Chickpeas [b] Beans [b]	(Urpi-Sarda et al., 2015)
α-Galactosides									
31	0.26	679.2275	679.2296	−2.1	$C_{25}H_{44}O_{21}$	Galactopyranosylciceritol or isomer (level 3)	Ms2: 611.2177 (10); 517.1723 [M-162-H] (2); 499.1642 [M-180-H] (30); 485.1487 (2); 383.1175 (100); 341.1070 (20); 221.0654 (15)	Chickpeas [a] Lentils [b] Beans [c]	FooDB.ca
32	0.27	517.1759	517.1768	−0.9	$C_{19}H_{34}O_{16}$	Ciceritol (level 2)	Ms2: 499.1638 [M-18-H] (10); 337.1120 [M-180-H] (100); 263.0753 (10); 221.0652 (15); 193.0705 (10); 179.0549 (25); 161.0444 (10)	Chickpeas [a] Lentils [b] Beans [c]	FooDB.ca

(*continued on next page*)

Table 1 (*continued*)

n	Retention time	Observed Mass [M-H]-	Theoretical Mass [M-H]-	Absolute error (mDa)	Molecular formula [M-H]-	Potential compound (level of identification)	MS/MS and MSn	Statistical analysis	Identified based on spectra published by
33	0.28	355.1225	355.1240	−1.5	$C_{13}H_{23}O_{11}$	Galactopinitol (level 2)	Ms2: 337.1117 [M-18-H] (20); 323.0960 (20); 309.0644 [M-46-H] (20); 263.0752 [M-46-46-H] (20); 193.0703 [M-162-H] (100); 179.0549 [M-176-H] (35); 161.0444 [M-176-18-H] (45); 143.0339 [M-176-18-18-H] (20)	Chickpeas [a] Lentils [b] Beans [c]	FooDB.ca
34	0.28	665.2121	665.2140	−1.9	$C_{24}H_{41}O21$	Stachyose (level 2)	Ms2: 503.15970 [M-162-H] (10); 485.14886 [M-180-H] (40); 443.13824 [M-180-42-H] (10); 383.11743 [M-162-120-H] (100); 341.10699 [M-162-162-H] (30); 323.09656 [M-162-180-H] (10)	Beans [a] Lentils [b] Chickpeas [b]	(Zhou et al., 2016)
35	0.3	503.1595	503.16118	−1.7	$C_{18}H_{31}O_{16}$	Raffinose or isomer (level 3)	Ms2: 341.1074 [M-162-H] (30); 323.0968 [M-162- 18-H] (70); 221.0655 [M-162-120-H] (100); 179.0552 [M-162-162-H] (95); 161.0447 [M-162-180-H] (25)	Beans [a] Chickpeas [b] Lentils [c]	HMDB
Fatty acyls									
36	2.67	517.2291	517.22902	0.1	$C_{24}H_{37}O_{12}$	Dihydroxy-megastigmadien-one-[apiosyl-glucoside] isomer 1 [#] (level 2)	Ms2: 385.1851 [M-132-H]- (100); 223.1338 [M-132-162-H]- (30); 205.1233 [M-132-180-H]- (60); 153.0912 (20)	Chickpeas [a] Beans [a] Lentils [c]	FooDB.ca
37	2.77	517.2284	517.22902	−0.6	$C_{24}H_{37}O_{12}$	Dihydroxy-megastigmadien-one-[apiosyl-glucoside] isomer 2[#] (level 2)	Ms2: 385.1849 [M-132-H]- (100); 223.1326 [M-132-162-H]- (20); 205.1222 [M-132-180-H]- (35); 153.0912 (20)	Lentils [a] Chickpeas [a] Beans [b]	FooDB.ca
38	3.69	503.2491	503.2498	−0.7	$C_{24}H_{39}O_{11}$	Megastigmadiene-diol -[apiosyl-glucoside] [#] (level 2)	Ms2: 371.2059 [M-132-H]- (100); 209.1536 [M-132-162-H]- (5) Ms3 of 371: 209.1539 [M-132-162-H]- (50); 161.0446 (100)	Lentils [a] Chickpeas [a] Beans [b]	FooDB.ca
Prenol lipids									
39	3.09	361.1647	361.1656	−0.9	$C_{20}H_{25}O_{6}$	Heliangin [#] (level 2)	Ms2: 343.2118 [M-18-H]- (10); 317.1747[M-44-H] (100); 261.1487 [M-C5H8O2-H]- (60); 233.1543 [M-C5H8O2-C2H4-H]- (30); 203.1071 [M-158-H]- (35)	Beans [a] Lentils [b] Chickpeas [b]	FooDB.ca
40	4.28	361.1647	361.165	−0.3	$C_{20}H_{25}O_{6}$	Gibberellin A25 or A38 or A19 (level 3)	Ms2: 343.2122 [M-18-H]- (40); 317.1753 [M-44-H]- (100); 273.1855 [M-44-44-H]- (60) Ms3 of 317: 273.1855 [M-44-44-H]- (100)	Beans [a] Chickpeas [b] Lentils [b]	FooDB.ca
Nucleosides									
41	0.46	243.0618	243.0623	0.5	$C_9H_{11}N_2O_6$	Pseudouridine (level 2)	Ms2: 225.0878 [M-18-H]- (15); 200.0560 [M-43-H]- (100); 153.0302 (10); 152.0350 (20); 111.0197 [M-132-H]- (15); 110.0245 (25)	Beans [a] Chickpeas [b] Lentils [c]	HMDB
Other organic compounds									
42	0.43	191.0193	191.0192	−0.1	$C_6H_7O_7$	Citric acid (level 1)	Confirmed with mass spectra and std.	Beans [a] Lentils [b] Chickpeas [c]	

(*continued on next page*)

Table 1 (*continued*)

n	Retention time	Observed Mass [M-H]-	Theoretical Mass [M-H]-	Absolute error (mDa)	Molecular formula [M-H]-	Potential compound (level of identification)	MS/MS and MSn	Statistical analysis	Identified based on spectra published by
43	2.70	277.1550	277.1552	0.2	$C_{15}H_{21}N_2O_4$	Phenylalanyl-leucine (level 3)	Ms2: 216.1382 (20); 141.1024 (75); 130.0865 (100)	Chickpeas [a] Beans [b] Lentils [c]	HMDB

Notes: 15 corresponded to the loss of a CH_3; 18 corresponded to the loss of H_2O; 42 corresponded to the loss of C_2H_2O; 44 corresponded to the loss of a CO_2; 126 corresponded to 1,3,5-trihydroxybenzene structure in ring A of the extension unit or indicates that the A ring of the upper unit has a 1,3,5-trihydroxybenzene structure; 132 corresponded to the loss of $C_5H_8O_4$ (pentosides: i.e. ribose); 146 corresponded to the loss of $C_6H_{10}O_5$; 162 corresponded to the loss of $C_6H_{10}O_5$; 176 corresponded to the loss of feruloyl moiety or glucuronic acid; 180 corresponded to the loss of hexose +18 amu; 266 corresponded to the loss of 146 + 120 amu.
Std, standard; H: 1.00782; CH_3: 15.2347; H_2O: 18.01056; CO_2: 43.98982; $C_5H_8O_4$: 132.04225; $C_6H_{10}O_5$: 146.05790; $C_6H_{10}O_5$: 162.05281; $C_6H_{12}O_6$: 180.06337.
Statistical analysis: Legumes with different superscript letters are significantly different, $P < .05$ (Fisher post hoc test).
[v] These compounds have not been previously identified in beans, lentils or chickpeas.