Proteomic analysis of non-toxic *Jatropha curcas* byproduct cake: Fractionation and identification of the major components

Andrés León-Villanueva[a,b,1], José A. Huerta-Ocampo[a,c,1], Alberto Barrera-Pacheco[a], Sergio Medina-Godoy[b,*], Ana P. Barba de la Rosa[a,*]

[a]*IPICyT, Instituto Potosino de Investigación Científica y Tecnológica A.C., Camino a la Presa San José No. 2055, Lomas 4a Sección, San Luis Potosí, SLP CP 78216, México.*
[b]*Instituto Politécnico Nacional, Centro Interdisciplinario de Investigación para el Desarrollo Integral Regional, Unidad Sinaloa. Laboratorio de Alimentos Funcionales, Departamento de Biotecnología Agrícola, Blvd Juan de Dios Bátiz Paredes 250, Guasave, Sinaloa, CP 81101, México.*
[c]*CONACyT-Centro de Investigación en Alimentación y Desarrollo, A.C. Apartado Postal 1735, Hermosillo, Sonora, Mexico.*

*Corresponding author.

E-mail: apbarba@ipicyt.edu.mx (A.P. Barba de la Rosa); smedinam@ipn.mx (S. Medina-Godoy).

[1]These authors contributed equally

ABSTRACT

Jatropha curcas non-toxic genotypes have been reported in Mexico and the press-cake, after oil extraction, represents a potential of new source of protein for food and feed uses. However, the characterization of the press-cake proteins is still unknown. The aim of this work was to carry out the molecular characterization of *J. curcas* seed storage proteins. Proteins in press-cake were pre-fractionated according to the classical Osborne procedure. Main protein fraction in *J. curcas* cake was represented by glutelins, the electrophoretic analysis showed that glutelins and globulins have the same profile, indicating that oil extraction process could have effect on globulins agglomeration. Protein fractions were analyzed by two-dimensional gel electrophoresis and mass spectrometry, results provide a new dataset of protein species or proteoforms that are accumulated in *J. curcas* endosperm. The identification of toxic proteins such as curcin in the non-toxic variety could represent that this protein have important roles in seeds. Regulatory proteins such as proteasome subunits and 14-3-3 were identified. A group of different heat shock and stress defense protein species was detected. Proteases related with inhibitory activity against DPPIV were also detected; this could support the potential use of *J. curcas* cake as nutraceutical food.

Keywords: *Jatropha curcas*; LC-MS/MS; mass spectrometry; non-toxic; proteome; seed storage proteins

1. Introduction

Jatropha curcas L. is a stem-succulent tree native to America, which as *Ricinus* and *Manihot* belongs to the family *Euphorbiaceae*. *J. curcas* was spread by Portuguese sailors to Africa and Asia (Fairless, 2007; Maghuly and Laimer, 2013). This subtropical plant can grow through different climatic and soil conditions, it has several uses in different agricultural systems such as barriers against wind and soil erosion as well as a source of firewood (Dias et al., 2012; Maghuly and Laimer, 2013). *J. curcas* main attraction is due to its high potential as plant for biodiesel production (Maghuly and Laimer, 2013). *J. curcas* seeds contains 30-45% toxic oil, with a high amounts of oleic and linoleic acids which make it more suitable for fuel purposes, as compared with other vegetable oils, due to its high rate fuel consumption and its higher oxidation stability (Fairless, 2007; Gübitz et al., 1999; Openshaw, 2000; Pramanik, 2003). *J. curcas* oil has therefore been used for long time as a raw material for paints and soap production as well as for lamp and lubricating oils (Kumar and Sharma, 2008). Besides, different parts of *J. curcas* plant contain a range of interesting metabolites and bioactive compounds, which is taking great attention for research as medicinal plant (Sabandar et al., 2013).

The mature seeds of *J. curcas* have a tick endosperm representing more than 90% of their weight in which is embedded a small embryo (Liu et al., 2009). The seeds are rich on protein (25-30%) and oil (55-62%), values that depends on the region agroclimatic characteristics. After oil extraction from dehulled kernels, the oil is converted into biodiesel and the resulting press-cake is a rich source of protein (60-63%) that could be an excellent protein source. The proteins present in the press-cake are rich of essential amino acids even with higher values (except

lysine) than those reported by the Food and Agriculture Organization reference protein (Haas and Mittelbach, 2000; Martinez-Herrera et al., 2012). However, due to the highly toxic and anti-nutritional compounds presents in *J. curcas* seeds, make the press-cake and oil unsuitable for the use as feedstuff or for human consumption (Maghuly and Laimer, 2013).

In Mexico, *J. curcas* L. grows wild in semitropical and tropical climates, however only Mexico has reported a nontoxic *J. curcas* genotype (He et al., 2011; Perea-Domínguez et al., 2017). The seeds of this nontoxic genotype are traditionally used to prepare a range of traditional dishes in Veracruz, Puebla, and Hidalgo States of Mexico (Makkar and Becker, 2009; Martinez-Herrera, 2012).

The exploitation of *J. carcus* press-cakes has been limited by the little knowledge about some aspects of the biochemistry of these seeds such as the type and functionality of proteins concentrated in the press-cake. Therefore, it is necessary to generate information on these aspects that would help to design innovative biotechnological approaches for the use of the high amount of proteins present in press-cake for future applications in the food industry.

The availability of the *J. curcas* genome and transcriptome data reported in the public databases (Costa et al., 2010; King et al., 2011) have allowed the increase of studies toward the seed proteome characterization. Shah et al. (2015) have used the label-free quantitative proteome analysis in order to analyse the proteins presents in the whole *J. curcas* seed endosperm in searching for information on phorbol esters biosynthesis, mechanism which still is not elucidated. Pinheiro et al. (2013) and Soares et al. (2014) focused in the proteomic analysis of the inner integument from developing seeds, while Shah et al. (2016) carried out

the proteome analysis on gerontoplasts isolated from the inner integument of developing seeds. Liu et al. (2015) carried out the proteomics analysis of oil body associated protein species by using gels-based proteomic technique, whereas Liu et al. (2013) used the comparative proteomic approach to profile the protein changes during seed development.

However, to date there is no information about the physical properties of *J. curcas* seed storage proteins presents in press-cakes. It is well known that the functional and molecular characterization of protein-enriched fractions is the first step in designing strategies that allow them to be integrated as additives or food ingredients (Rezig et al., 2013). In this sense, Osborne´s protein classification (Osborne, 1908), which is based on protein solubility characteristics, is the most common method for seed storage protein characterization and seems to be an excellent pre-fractionation step towards proteome characterization. For these reasons the aim of the present work was to carry out the characterization of proteins present in press-cake obtained after oil extraction of non-toxic *J. curcas.* Gel-based proteomics and LC-MS/MS were used in order to obtain a new description of *J. curcas* press-cake proteins; such valuable information will help in the way to design technologies for development of new food products.

2. Material and methods

2.1. J. curcas press-cake preparation

Non-toxic *Jatropha curcas* seeds (Puebla's ecotype, voucher specimen numbered 53203) were obtained from ripe fruits from plants cultivated in "La Esmeralda" Experimental Field of the Agricultures Association of West Sinaloa

River (AARSP, Guasave, Mexico). *J. curcas* seed press cake was produced after oil pressing from dehulled seeds (kernels) using a screw-press. Two press cake samples (500 g each) were collected and air-dried under ambient conditions. The combined air-dried press cakes were milled to a small particle size (40-mesh) using a Cyclotec 1093 mill (FOSS Tecator, Hilleroed, Denmark) and then defatted with n-hexane (1:10, w/v) for 4 h with agitation and air-dried in a fume hood after decantation of the hexane. Dry meals were kept in plastic bags at 4°C until used.

2.2. Meal protein sequential fractionation based on Osborne´s solubility

Protein fractions were sequentially extracted according to a modified Osborne´s procedure (Ribeiro et al., 2004). Briefly, the defatted meal was extracted with water containing 10 mM $CaCl_2$ and 10 mM $MgCl_2$ (1:10, w/v) pH 8.0, with constant stirring for 4 h at 4°C. The slurry was centrifuged at 30,000g for 1 h at 4°C. The supernatant was recovered and extensively dialyzed (12.4 kDa MW cut-off) against distilled water. After dialysis, the suspension was centrifuged at 15,000g for 15 min at 4°C and supernatant (albumins) was collected and freeze-dried.

The resulting pellet from albumins extraction was resuspended with 100 mM Tris-HCl, pH 7.5, containing 10% NaCl (w/v), 10 mM ethylenediaminetetraacetic acid (EDTA), and 10 mM ethylene glycol-bis(2-aminoethylether)-N,N,N',N'-tetraacetic acid (EGTA) (1:10, w/v). The suspension was stirred for 4 h at 4°C. The insoluble proteins were removed by centrifugation at 30,000g for 1 h at 4°C. The precipitate (globulins fraction) was dialyzed, centrifuged and freeze-dried.

146 The insoluble pellet resulting from globulins extraction was resuspended in
147 75% aqueous ethanol (1:10, w/v) and centrifuged (30,000g at 4°C for 15 min). The
148 supernatant was dialyzed against distilled water. The supernatant (prolamins) was
149 recovered and freeze-dried. The insoluble pellet from the previous extraction was
150 resuspended (1:10, w/v) with 50 mM sodium borate buffer, pH 10, containing 1%
151 (v/v) β-mercaptoethanol, and 1% (w/v) sodium dodecyl sulphate (SDS). The
152 suspension was stirred at room temperature for 2 h and centrifuged at 30,000g for
153 15 min at 20°C. After dialysis against distilled water, the fraction (glutelins) was
154 freeze-dried. All samples were kept at -70°C until use. A scheme of the whole
155 extraction procedure is presented in **Supplementary Figure S1**.
156
157 *2.3. Protein sample preparation*
158 Freeze-dried protein fractions (five grams) were mixed with 50 mL of its
159 respective extraction solution and 1% (w/v) of polyvinylpolypyrrolidone (PVPP).
160 The suspensions were mixed in a vortex for 1 min and sonicated for 2.5 min at 20
161 kHz with 35% of amplitude (GE-505, Ultrasonic Processor, Sonics & Materials,
162 Inc., Newtown, CT, USA). Samples were maintained at 4^{0}C during suspension
163 steps and centrifuged at 15,000g for 15 min at 4°C (Beckman Avanti J26-XP,
164 Beckman Coulter, Brea, CA, USA). Supernatants were transferred to new tubes
165 and proteins were precipitated by adding four volumes of ice-cold acetone and
166 incubated overnight at -20°C. After 15 min of centrifugation at 15,000g for 15 min
167 at 4°C, the supernatant was decanted and discarded, with the residual pellet being
168 washed twice with ice-cold acetone and allowed to dry at room temperature. The
169 resulting dried pellets were suspended in rehydration buffer (8 M urea, 2% (w/v) 3-

[(3-cholamidopropyl)-dimethylammonio]-1-propanesulfonate (CHAPS), 0.56% (w/v) dithiothreitol (DTT), 0.002% bromophenol blue) mixed in vortex for 30 s and sonicated for 80 s. The suspensions were centrifuged under the previous conditions and the supernatants were recovered. Protein concentration was determined by using protein assay (Bio-Rad, Hercules, CA, USA) with bovine serum albumin (BSA) used as standard.

2.4. Two-dimensional gel electrophoresis (2-DE) and gel image analysis.

Proteins (900 µg) were suspended in 450 µL of rehydration buffer, containing 0.5% (v/v) IPG buffer (Bio-Rad) in the range of 5-8 (albumins) and 3-10 for globulins and glutelins fractions. Proteins were loaded onto 24 cm linear immobilized pH gradient (IPG, Bio-Rad) strips 5-8 (albumins) or 3-10 (globulins and glutelins). Passive rehydration was carried out at room temperature for 14-16 h. The IEF was conducted at 50 mA per IPG strip and 20°C in an Ettan IPGphor 3 system (GE Healthcare, Piscataway, NJ, USA) under the following conditions: (I) 150 V gradient for 2 h (II) 200 V gradient for 2 h, (III) 400 V gradient for 2 h, (IV) 1500 V gradient for 2 h, (V) 4500 V gradient for 3 h, (VI) 10,000 V gradient for 3 h; and (VII) holding at 10,000 V for 10 h. After IEF, the IPG strips were stored at -20 °C or immediately equilibrated for 15 min in equilibration buffer [50 mM Tris–HCl pH 8.8, 6 M Urea, 30% (v/v) glycerol, 2% (w/v) SDS, 0.002% (w/v) bromophenol blue, 65 mM DTT].

Strips were placed directly onto 13% polyacrylamide-SDS slab gels and the second dimension was performed using an EttanTM DALT-six Electrophoresis Unit (GE Healthcare), using SDS electrophoresis buffer [25 mM Tris pH 8.8, 192 mM

194 glycine, and 0.1% (w/v) SDS] and resolved at 20 mA/gel until the dye
195 (bromophenol blue) reached the bottom of the gels. Three different extractions
196 were prepared for gel replicates using the same method. After SDS-PAGE gels
197 were stained with PhastGelTM Blue R-350 (GE Healthcare) and scanned at 100
198 µm resolution using a Pharos FX Plus Molecular Imager (Bio-Rad). Image analysis
199 was performed with PDQuest 2-D Analysis Software v8.0 (Bio-Rad). The molecular
200 masses of proteins in gels were determined by co-electrophoresis of molecular
201 weight standards (BenchMark Protein Ladder, Invitrogen. Carlsbad, CA, USA),
202 while isoelectric point (p*I*) was determined by migration of protein spots on the
203 linear IPG strips.
204
205 *2.5. Liquid chromatography-tandem mass spectrometry analyses (LC-MS/MS)*
206 Protein spots were carefully excised from the 2-DE gels and distained,
207 reduced with 10 mM DTT in 25 mM ammonium bicarbonate followed by protein
208 alkylation with 55 mM iodoacetamide. Protein digestion was carried out overnight
209 at 37°C with sequencing grade trypsin (Promega, Madison, WI, USA). Nanoscale
210 LC separation of tryptic peptides was performed with a nanoACQUITY UPLC
211 System (Waters, Milford, MA, USA) and LC-MS/MS analysis was carried out in a
212 SYNAPT-HDMS Q-TOF (Waters) as previously reported (Huerta-Ocampo et al.,
213 2014) with brief modifications: Accurate mass data were collected in an alternating
214 Data Dependent Acquisiton mode (DDA). In low energy mode, data were collected
215 at constant collision energy of 3 eV. In elevated-energy mode, the collision energy
216 was ramped from 15 to 45 eV during 3 s of integration.

2.6. Protein identification

MS/MS spectra datasets were used to generate PKL files (.pkl) using ProteinLynx Global Server v2.4 (Waters). Proteins were then identified using PKL files and the MASCOT search engine v2.3 (Matrix Science, London, UK). Searches were conducted using the *Viridiplantae* subset of the NCBInr protein database (2391213 sequences, December 2013) and an in-house database containing the *Jatropha curcas* nucleotide coding sequences available at the NCBI database (77340 sequences, December 2013). Trypsin was used as the specific protease and one missed cleavage was allowed. The mass tolerance for precursor and fragment ions was set to 100 ppm and 0.1 Da respectively and peptide charges were set at +2, +3, and +4. Carbamidomethylation of cysteine was set as fixed modification and oxidation of methionine was set as variable modification. Identifications were considered successful when significant MASCOT scores (>33 for *J. curcas* nucleotide in-house and >51 for *Viridiplantae* protein database) were obtained, indicating the identity or extensive homology at $p<0.01$ and the presence of at least two peptides were considered necessary for reliable identification. Identified proteins were classified into different categories according to Gene Ontology (http://www.geneontology.org/). A biological pathway diagram showing the identified enzymes involved in a metabolic network was generated using the visualization and analysis tool PathVisio *v*3.2.4 (Kutmon et al., 2015).

3. Results

3.1 Protein extraction

Although Osborne´s classification based on solubility criteria is an ambiguous system, due that all depends in solutions used for extractions, is the most useful method for seed storage protein characterization and is a very useful method for protein pre-fractionation towards seeds storage proteins proteomics analysis (Jorrin-Novo, 2014). Hence, the proteins in press-cake *J. curcas* were subjected to sequential extraction. It was observed that globulins fraction corresponds only to 20.17%, while glutelins represented the main fraction with 42.03% (**Supplementary Table S1**), which is different to all legumes where the main fraction is globulins. The representative electrophoretic profile in one dimension gel (SDS-PAGE) of different *J. curcas* press-cake protein fractions is shown in **Figure 1A**.

3.2. 2-DE maps of seed protein fractions

The water-soluble protein fraction (albumins) was separated by 2-DE in the range of linear pH 5-8, while the salt soluble proteins (globulins) and glutelins were resolved in the range of linear pH 3-10. In the albumins fraction 434 spots were separated in 2-DE gels (**Figure 1B, Supplementary Figure S2)**, while 310 spots for globulins (**Figure 1C, Supplementary Figure S3)** and 175 spots for glutelins (**Figure 1D, Supplementary Figure S4)** were resolved. From 84.1 to 96.0% of the protein spots from the three seed storage protein fractions were successfully identified by LC-MS/MS analysis and homology database search (**Table 1**). All identified proteins were classified into its functional classes. In albumins fraction were identified more than 100 protein species associated with well-known storage reservoir (**Figure 2A**). Enzymes involved in glycolysis, TCA cycle. Enzymes related

with sugars, nitrogen, and lipid metabolisms were also represented, as well proteins related with oxide-reduction function (SOD, GST). Proteasome subunits, 14-3-3 proteins, and enzymes related with defence and detoxification (LEA, glyoxalase, cyclophilin, and lactoylglutathione lyase) were also identified. Several heat shock proteins and proteases were identified (SERPIN-ZX-like, cysteine proteinase, leucine aminopeptidase). Interestingly in albumins fraction were detected enzymes related with the THF pathway and purine metabolism that have not been reported before (**Table 2, Supplementary File S1**).

Globulins and glutelins fractions showed 24 and 7 protein species, respectively (**Figure 2B and 2C**) with several proteoforms representing each protein species. Globulins fraction was composed mainly by legumin B-like, 2S albumin-like and germin-like protein, curcin 1 and enzymes related to energy production such as ATP synthase and luminal-binding protein (**Table 3 and Supplementary File S2**). Glutelins, although the most abundant fraction in *J. curcas* press-cake was composed mainly of high molecular weight globulins (**Table 4, Supplementary File S3**). In **Supplementary File S4** is shown the list of theoretical and experimental molecular weight/isoelectric point corresponding to all identified proteins.

4. Discussion

4.1. *J. curcas* seeds protein fractionation

It is recognized that the total set of protein species and their corresponding proteoforms (frequently abundant) that constitute the cell´s proteome could not be

captured in just one experiment (Romero-Rodríguez et al., 2014). Therefore, the use of fractionation techniques is necessary in order to obtain a more complete proteome characterization. Osborne´s method has been long used for seed protein fractionation and is still one of the preferred procedures for protein pre-fractionation before proteomics profiling of seeds.

The high lipid content in seeds turns *J. curcas* into a high value crop, however the press-cake obtained from non-toxic *J. curcas* has the potential to be used as source of food. In this sense, seed protein storage proteins are the basis for seed protein characterization; hence the seed storage protein fractions in *J. curcas* non-toxic seeds were characterized. Most of enzymatic proteins identified were found in albumins fraction (water soluble-proteins). The network of metabolic activities of identified enzymes is represented in Figure 3. Globulins and glutelins were represented by nutrient reservoir proteins (**Table 3 and 4**).

4.2. Protein species in albumins fraction

4.2.1. Protein species related to nutrient reservoir

Several proteoforms of legumin B and legumin A were detected (45 and 36, spots, respectively). 11S globulin was detected in 21 spots and globulin-1 S in 9 spots (**Table 2**). Globulin 1 (GLB1) is one of the most abundant proteins accumulated in maize seed tissues and has been designated as an excellent indicator for the scrutiny of genetic variation (Hilton and Gaut, 1998). Two isoforms of vicilin-like antimicrobial peptides 2-2 (AMPs) were detected in 88 spots while 38 spots corresponded to 2S-albumins (**Table 2**). Antimicrobial peptides have been frequently isolated from seeds among other plant tissues (Franco et al., 2006;

311 Lipkin et al., 2005; Pelegrini et al., 2006 and 2008). It has been reported that *J.*
312 *curcas* is a good source of cyclic peptides with diverse biological functions such as
313 antimalarial, antiplatelet, antiproliferative, and cytotoxic (Pinto et al., 2015). Xiao et
314 al. (2011) reported the JCpep7 peptide that displayed antimicrobial activity against
315 Gram-negative pathogens like *Shigella dysenteriae*, *Pseudomonas aeruginosa* and
316 *Salmonella typhimurium*. These bioactive peptides found in *J. curcas* could
317 contribute in the development of antimicrobial strategies based in the use of novel
318 natural sources (Mandal et al., 2009).

319 One LEA protein (Em-like) was detected in 2 spots. LEAs (Late
320 embryogenesis abundant proteins) are widely recognized as associated with
321 desiccation tolerance in seeds. Shih et al. (2010) identified an Em-like protein in
322 rice, which was not been identified previously in cereals, by other hand Vicient et
323 al. (2010) reported two *Em* genes in Arabidopsis.

324 He et al. (2011) described that non-toxic *J. curcas* does not contain phorbol
325 esters, nevertheless contains curcins. In agreement with this statement, four faint
326 spots were detected in the *J. curcas* 2-DE map, which are not enough to cause the
327 toxic effects. Lin et al. (2010) have indicated that curcin hemagglutinating activity is
328 presented at more than 7.8 mg/L and LD_{50} for oral semi-lethal dose was reported
329 to be 104.73 mg/kg. Curcins type-1 are proteins that specifically are accumulated
330 in the endosperm (Gu et al., 2015; Lin et al., 2003a) and are classified as Type I
331 ribosome-inactivating proteins (RIP). Zhang et al. (2017) reported a novel RIP from
332 *J. curcas*, however it is not well documented its function but it may be part of the
333 response to biotic and abiotic stresses (Qin et al., 2009). Curcins have also related
334 with antitumor activity (Lin et al., 2003b, Luo et al., 2006).

335 *4.2.2. Protein species involved in sugar metabolism*

336 Glycolysis is well represented in *J. curcas* albumins fraction,
337 glyceraldheyde-3P-DH was detected in 14 spots, two isoforms of
338 phosphoglucomutase and triose phosphate isomerase and three isoforms of
339 enolase were identified. The pyruvate dehydrogenase E1 component subunit beta-
340 1, part of the complex that carried out the conversion of pyruvate to Acetyl-CoA,
341 was detected in four spots. Aldehyde dehydrogenase, phosphoglycerate mutase,
342 and enzymes related to the TCA cycle (aconitase, two isoforms of citrate synthase
343 and two isoforms of malate dehydrogenase) were identified (**Table 2**).

344 Uridyltransferase (UTP-glucose-1-phosphate uridylytransferase) or glucose-
345 1-phosphate uridylytransferase (detected in 2 spots) is an enzyme involved in
346 carbohydrate metabolism with an important role in glycogenesis and cell wall
347 synthesis (Dai et al., 2006). In relation to the sugars metabolism, beta-
348 galactosidase 8 and beta-xylosidase/alpha-L-arabinofuranosidase 1 were detected
349 in *J. curcas*. Aldose-1-epimarase, which catalyses the first step in galactose
350 metabolism was also detected in one spot. Two proteoforms of sorbitol
351 dehydrogenase enzyme, which converts sorbitol into fructose, were also detected.

352

353 *4.2.2. Protein species related fatty acid metabolism*

354 Proteins related to fatty acid metabolism were identified in albumins fraction
355 (Table 2). Among them, the hydroxyacyl-ACP-dehydrase, enzyme involved in fatty
356 acid synthesis is the key enzyme of the fatty acid synthesis (FAS) system (Lung et
357 al., 2016) was identified in one spot. One proteoform of the acyl-Co-A-binding
358 protein was identified in *J. curcas* albumins fraction (Table 2), this is an enzyme

359 that binds medium- and long- chain acyl CoA esters with very high affinity. Acyl-
360 Co-A-binding protein may function as maintenance, protection, and transport of the
361 Acyl-CoA pool (Yurchenko and Weselake, 2011).

362 In *J. curcas* seeds the triacylglycerol's (TAGs) are accumulated in oil bodies,
363 which consist of a central core of neutral lipids, delimited by a monolayer of
364 phospho- and glico-lipids including sterols and proteins associated to the oil bodies
365 surface. One proteoform of the oil body-associated 2B-like protein in albumins
366 fraction was identified. In maize (*Zea mays* L.), the oil body associated protein 1
367 gene (*obap1*) was found mainly expressed in scutellum during maturation (López-
368 Ribera et al., 2014).

369

370 *4.2.3. Protein species related to amino acids and purine metabolism*

371 Nitrogen in seeds does not only come from root nitrogen uptake during the
372 later growing period, but also from the redistribution from the vegetative organs, in
373 addition, the recycling and redistribution of nitrogen in plants is important for the
374 environmental stress response (Zhang et al., 2010). Four proteoforms related to
375 amino acids and purine metabolism were identified in albumins fraction:
376 Ureidoglycolate hydrolase, Alanine aminotransferase 2-like, Nitrogen regulatory
377 protein P-II, and Reactive Intermediate Deaminase A, chloroplastic.
378 Ureidoglycolate (two proteoforms) is an intermediate of purine catabolism
379 catalyses the final step of ureide degradation in which inorganic nitrogen is re-
380 assimilated (Li et al., 2015; Werner et al., 2013). The beta-propionase enzyme,
381 detected in *J. curcas* albumins fraction (one spot), acts on carbon-nitrogen bonds,

specifically in linear amides and participates in pyrimidine, beta-alanine metabolism, and pantothenate biosynthesis (Shen et al., 2014).

Alanine amino transferase, an enzyme belonging to the L-alanine degradation and the nitrogen regulatory protein P-II, a constituent of the adenylation cascade involved in the regulation of glutamine synthetase (GS) activity were detected in one isoform (**Table 2**). In nitrogen-limiting conditions, P-II is uridylylated establishing a complex that allows the deadenylation of GS, thus activating the enzyme. On the contrary, under nitrogen surplus, this complex is deuridylated promoting adenylation and the consequent inactivation of GS (Brown et al., 1971; Jonsson and Nordlund, 2007). In plants, the function of reactive intermediate deaminase A is still unclear, however recent reports allowed to established its contribution in the biosynthesis of branched-chain amino acids (Niehaus et al., 2014), one proteoform was detected in *J. curcas* press-cake albumins fraction.

4.2.4. Protein species involved in detoxification and plant resistance

4-hydroxy-4-methyl-2-oxoglutarate aldolase (4HMO), dienelactone hydrolase, lactoylglutathione lyase, and cyclophilin were detected in *J. curcas* albumins fraction. The enzyme 4HMO cleaves 4-carboxy-4-hydroxy-2-oxoglutarate giving as a product pyruvate. Participates in benzoate degradation via hydroxylation, which links aromatic catabolism to central cellular metabolism (Tack et al., 1972). Dienelactone hydrolases play a crucial role in the bacterial degradation of chloroaromatic compounds. Many representatives of this group of xenobiotic compounds are converted to chlorosubstituted catechols by the initial

406 enzymes of aromatic catabolism (Schlöman et al., 1990). Two isoforms of
407 lactoylglutathone lyase were detected, this is an enzyme that catalyses the
408 isomerization of hemithioacetal adducts, which are cytotoxic compounds generated
409 under abiotc stresses, including metal toxicity (Nahar et al., 2017). Cyclophilin (two
410 isoforms) belongs to the immunophilin superfamily with peptidyl-prolyl cis-trans
411 isomerase (PPIase activity). Catalyse the interconversion of the cis- and trans-
412 rotamers of the peptidyl-prolyl amide bond of peptides. The interaction of soybean
413 cyclophylin GmCYP1 with the isoflavonoid regulator GmMYB176 and 14-3-3
414 proteins suggests its role in defence in soybean (Mainali, 2017). Interestingly, 14-3-
415 3 proteins were also detected in *J. curcas*.
416
417 *4.2.6. Protein species related to plant growth regulation*
418 Thiamine thiazole synthase (TH1, two proteoforms) plays a central in
419 thiamine biosynthesis, but it is also important in abiotic stress responses and
420 mitochondrial DNA damage tolerance (Li et al., 2016). It has also been associated
421 with heat tolerance in rice (Chen et al., 2011) and is involved in ABA-regulated
422 stomatal movement and in the plant´s drought response (Li et al., 2016).
423 Phosphatidylethanolamine binding protein (one isoform) includes various functions
424 such as lipid binding, neuronal development, control of the switch between shoot
425 growth and flower structures and the regulation of signalling such as the MPA
426 kinase and the NF-kappaB pathways (Vallée et al., 2003; Yeung et al., 2001).
427 Luminal-binding protein detected in three proteoforms, is involves in storage of
428 calcium pool inside the endoplasmic reticulum lumen, accumulation of this protein
429 confers resistance to drought (Valente et al., 2009). In soybean root tips the

induction of the luminal-binding protein 5 was observed during flood conditions (Komatsu et al., 2012).

4.2.6. Proteases and oxidoreductases

Cysteine protease RD19a-like, leucine aminopeptidase (three proteoforms), and serine proteinase inhibitor (SERPIN-ZX-like) were detected in *J. curcas* albumins fraction. Cysteine proteases RD19a-like (one proteoform) or cathepsin F-like proteases are members of the Papain-like cysteine protease, a large class of proteolytic enzymes associated with several plant processes (Richau et al., 2012). Serine protease inhibitor (Serpin, one proteoform) is a family of proteases inhibitors widely distributed. Members of the group of leucine aminopeptidase (three proteoforms), are implicated in transcriptional regulation (Asqui et al., 2017).

Glutathione-S-transferases (GSTs), detected in *J. curcas* albumins (four proteoforms), are a family of phase II detoxification enzymes. The acquired resistance to chemotherapy, herbicides, insecticides and microbial antibiotics has been attributed to the presence of GST (Townsend and Tew, 2003). Glutaredoxin (13 isoforms) have similar functions of thioredoxin (TRX), both are proteins involved in cell protection against oxidative stress damage (Sánchez-Riego et al., 2016). Protein disulphide isomerase (four proteoforms) participates throughout the maturation of extracellular proteins adding disulphide bonds to stabilize it or to covalently join it to other proteins (Wang and Tsou, 1993). Glutathione reductase (GR, spot 23) or glutathione-disulphide reductase (GSR) catalyses the reduction of glutathione disulphide (GSSG) to the sulfhydryl form glutathione (GSH), which is a

critical molecule in resisting oxidative stress and the maintenance of the cell reduced environment (Lüersen et al., 2013).

Iron is an essential nutrient for all cells, but its excess is harmful to cells, and so iron homeostasis must be controlled. Ferritins play important roles in sequestering or releasing iron as needed (Borg et al., 2012) and are exclusively targeted to plastids and mitochondria (Borg et al., 2012; Briat et al., 2010; Zancani et al., 2004). In this work, a ferritin was identified in albumins fraction (one proteoform). A recent report indicates that ferritins from wheat (*Triticum aestivum* L.) play important roles in enhancing tolerance in stresses associated with ROS (Zang et al., 2017).

4.2.8. Heat shock protein, ribosome, energy, and proteasome

A set of several proteoforms of heat shock proteins, small heat shock proteins, and chaperonins were identified in albumins fraction (Table 2). The nucleoside diphosphate kinase B (NDPKs, three isoforms) is an enzyme that catalyses, in a reversible manner, the production of different nucleoside diphosphates (NDP) and triphosphates (NTP) (Kihara et al., 2011). The regulator of ribonuclease (one proteoform) activity acts as a regulator of the endonuclease RNase E and inhibiting the RNA processing. Ribonuclease 3 is a family of ribonucleases that recognizes dsRNA and cleaves it to transform them into mature RNAs, this enzyme was detected in one proteoform. Ubiquitin-fold modifier conjugating (UBL, one isoform) is covalently linked to target proteins but its physiological functions are still not known (Daniel and Liebau, 2014).

4.3. Protein species in globulins and glutelins fraction

Globulins fractions were represented by legumin A and B, 2S albumins, vicilin-like antimicrobial peptides, Germin-like protein 5-1, and curcin 1. Germins (GER) and germin-like proteins (GLPs), together with sucrose-binding proteins and seed globulins, are part of protein superfamily called cupin. Large numbers of GERs and GLPs have been functionally characterized from diverse plant species, their accumulation is related to biotic and abiotic stress response (Ilyas et al., 2016). Despite their solubility properties, globulins and glutelins are structurally similar, even evidences indicate that both protein share a common gene origin (Okita et al., 1989). Our results indicate that similar proteins were identified in both fractions (Table 3 and 4).

4. Conclusions

In the present study, the protein composition of seed storage proteins of non-toxic *J. curcas* press-cake after oil extraction was investigated using proteomics tools. Results have shown that albumins or water-soluble proteins are rich on different protein species related with carbohydrate metabolism. Other proteins involved in lipid, nitrogen, and purine metabolism were also identified. Proteins related with detoxification of xenobiotics were detected. Interestingly, the 14-3-3 proteins and NDPK, which are components of signaling pathways as well as cyclophillin that is one of protein targets of 14-3-3 were detected in *J. curcas* water soluble fraction. Although toxic protein such as curcin was detected, the amount present in the non-toxic J*. curcas* seeds may be not enough to cause toxic effects. This study contributes to the information of *J. curcas* seed proteome.

Acknowledgments

Thank to CONACYT-MEXICO GRANT 56787 (Laboratory for Nanoscience and Nanotechnology Research-LINAN), INFRA-2013-01 No. 204373 and CONACYT-MEXICO GRANT CB-2008-103601.

Supplementary Information

Supplementary Table S1. Proportion of different protein fractions of non-toxic *Jatropha curcas* defatted meal.

Supplementary Figure S1. Scheme of non-toxic *Jatropha curcas* protein fractions (albumins, globulins, prolamins, and glutelins) extracted according to a modified Osborne procedure.

Supplementary Figure S2. Representative 2-DE gel of *Jatropha curcas* albumins fraction. Numbers indicates the spot number assigned for LC-MS/MS analysis reported in Table 2 and Supplementary File S1.

Supplementary Figure S3. Representative 2-DE gel of *Jatropha curcas* globulins fraction. Numbers indicates the spot number assigned for LC-MS/MS analysis reported in Table 3 and Supplementary File S2.

Supplementary Figure S4. Representative 2-DE gel of *Jatropha curcas* glutelins fraction. Numbers indicates the spot number assigned for LC-MS/MS analysis reported in Table 4 and Supplementary File S3.

Supplementary Files

Supplementary File S1. Identification of the albumins protein fraction extracted from non-toxic *Jatropha curcas* meal by LC-ESI-MS/MS analysis.

Supplementary File S2. Identification of the globulins protein fraction extracted from non-toxic *Jatropha curcas* meal by LC-ESI-MS/MS analysis.

Supplementary File S3. Identification of the glutelins protein fraction extracted from non-toxic *Jatropha curcas* meal by LC-ESI-MS/MS analysis.

Supplementary File S4. A list of theoretical and experimental values of molecular weight (PM) and isoelectric points (p*I*) from all identified protein spots from *J. curcas* seed storage proteins fractions (albumins, globulins, and glutelins) identified by LC-ESI-MS/MS

References

Asqui, S.L, Vercammen, D., Serrano, I., Valls, M., Rivas, S., Van Breusegem, F., Conlon, F.L., Dangl, J.L., Coll, N.S., 2017. AtSERPIN1 is an inhibitor of the metascapase AtMC1-mediated cell death and autocatalytic processing *in planta*. New Phytol. doi: 10.1111/nph.14446.

Borg, S., Brinch-Pedersen, H., Tauris, B., Madsen, L.H., Darbani, B., Noeparvar, S., Holm, P.B., 2012. Wheat ferritins: Improving the iron content of the wheat grain. J. Cereal Sci. 56, 204–213.

Briat, J.F., Duc, C., Ravet, K., Gaymard, F., 2010. Ferritins and iron storage in plants. Biochim. Biophys. Acta. 1800, 806–814.

Brown, M.S., Segal, A., Stadtman, E.R., 1971. Modulation of glutamine synthetase adenylylation and deadenylylation is mediated by metabolic transformation of the PII-regulatory protein. Proc. Nat. Acad. Sci. USA 68, 2949-2953.

Chen, X., Zhang, W., Zhang, B., Zhou, J., Wang, Y., Yang, Q., Ke, Y., He, H., 2011. Phosphoproteins regulated by heat stress in rice leaves. Proteome Sci. 9, 37.

Costa, G.G., Cardoso, K.C., Del Bem, L.E., Lima, A.C., Cunha, M.A.S., de Campos-Leite, L.; Vicentini, R., Papes, F., Moreira, R.C., Yunes, J.A., Campos, F.A.P., Da Silva, M.J., 2010. Transcriptome analysis of the oil-rich seed of the bioenergy crop *Jatropha curcas* L. BMC Genomics 11, 462.

Dai, N., Petreikov, M., Portnoy, V., Katzir, N., Pharr, D., Schaffer, A.A., 2006. Cloning and expression analysis of a UDP-galactose/glucose pyorphosphorylase from melon fruit provides evidence for the major metabolic pathway of galactose metabolism in raffinose oligosaccharide metabolizing plants. Plant Physiol. 142, 294-304.

Daniel, J., Liebau, E., 2014. The Ufm1 cascade. Cells 3, 627-638.

Dias, L.A., Missio, R.F., Dias, D.C., 2012. Antiquity, botany, origin and domestication of *Jatropha curcas* (*Euphorbiaceae*), a plant species with potential for biodiesel production. Genet. Mol. Res. 11, 2719–2728.

Fairless, D., 2007. Biofuel: the little shrub that could–maybe. Nature 449, 652–655. doi: 10.2038/449652a.

Franco, O.L., Murad, A.M., Leite, J.R., Mendes, P.A.M., Prates, M.V., Bloch, JrC., 2006. Identification of a cowpea g-thionin with bactericidal activity. FEBS J. 273, 3489–3497.

Gu, K., Tian, D., Mao, H., Wu, L., Yin, Z., 2015. Development of marker-free transgenic *Jatropha curcas* producing curcin-deficient seeds trhough endosperm-specific RNAi-mediated gnee silencing. BMC Plant Biol. 15, 242.

Gübitz, G.M., Mittelbach, M., Trabi, M., 1999. Exploitation of the tropical oil seed plant *Jatropha curcas* L. Bioresour. Technol. 67, 73–82.

Haas, W., Mittelbach, M., 2000. Detoxification experiments with the seed oil from *Jatropha curcas* L. Ind. Crops Prod. 12, 111–118.

He, W., King, A.J., Khan, M.W., Cuevas, J.A., Ramiaramanana, D., Graham, I.A., 2011. Analysis of seed phorbol-ester and curcin content together with genetic diversity in multiple prevenances of *Jatropha curcas* L. from Madagascar and Mexico. Plant Physiol. Biochem. 49, 1183-1190.

Hilton, H., Gaut, B.S., 1998. Speciation and domestication in maize and its wild relatives: Evidence from the *Globulin-1* Gene. Genetics 150, 863-872.

Huerta-Ocampo, J.A., Barrera-Pacheco, A., Mendoza-Hernández, C.S., Espitia-Rangel, E., Mock, H.P., Barba de la Rosa, A.P., 2014. Salt stress-induced

584 alterations in the root proteome of *Amaranthus cruentus* L. J. Proteome Res.13,
585 3607-3627.
586 Ilyas, M., Rasheed, A., Mahmood, T., 2016. Functional characterization of germin
587 and germin-like protein genes in various plant species using transgenic
588 approaches. Biotechnol. Lett. 38, 1405-1421.
589 Jonsson, A., Nordlund, S., 2007. In vitro studies of the uridylylation of the three PII
590 protein paralogs from *Rhodospirillum rubrum*: the transferase activity of R.
591 Rubrum GlnD is regulated by a-ketoglutarate and divalent cations but not by
592 glutamine. J. Bacteriol. 189, 3471-3478.
593 Jorrin-Novo, J.V., 2014. Plant Proteomics: Methods and protocols. In Methods in
594 Molecular Biology, vol. 1072, Chapter 1, pp 3-13. Springer. DOI 10.1007/978-1-
595 62703-631-3-1.
596 Kihara, A., Saburi, W., Wakuta, Sh., Kim, M.-H., Hamada, Sh., Ito, H., Imai, R.,
597 Matsui, H., 2011. Physiological and biochemical characterization of three
598 nucleoside diphospahte kinase isozymes from rice (*Oryza sativa* L.). Biosci.
599 Bitoechnol. Biochem. 75, 1740-1745.
600 King, A.J., Li, Y., Graham, I.A., 2011. Profiling the developing *Jatropha curcas* L.
601 Seed transcriptome by pyrosequencing. Bioenerg Res. 4, 211-221.
602 Komatsu, S., Kuki, R., Nanjo, Y., Hiraga, S., Furukawa, K., 2012. Comprehensive
603 analysis of endoplasmic reticulum-enriched fraction in root tips of soybean under
604 flooding stress using proteomics techniques. J. Proteomics 77, S31-S60,
605 Kumar, A., Sharma, S., 2008. An evaluation of multipurpose oil seed crop for
606 industrial uses (*Jatropha curcas* L.): A review. Ind. Crops Prod. 28, 1–10.
607 Kutmon, M., van Iersel, M.P., Bohler, A., Kelder T, Nunes, N., Pico, A.R., Evelo,
608 C.T., 2015. PathVisio 3: An Extendable Pathway Analysis Toolbox PLos
609 Computational Biology 11(2):e1004085. doi: 10.1371/journal.pcbi.1004085.
610 Li, Ch.-L., Wang, M., Wu, X.-M., Chen, D.-H., Lv, H.-J., Shen, J.-L., Qiao, Z.,
611 Zhang, W., 2016. THI1, a thiamine thiazole synthase, interacts with Ca^{2+}-
612 dependent protein kinase CPK33 and modulates the S-type anion channels and
613 stomatal closure in Arabidopsis. Plant Physiol. 170, 1090-1104.
614 Li, J., Qin, R.-Y., Li, H., Xu, R.-F., Yang, Y.-Ch., Ni, D.-H., Li Li, H.-M., Wei, P.-Ch.,
615 Yang, J.-B., 2015. Low-temperature-induced expression of rice ureidoglycoalte
616 amidohydrolase is mediated by a C-repeat/dehydration-responsive element that
617 specifically interacts with rice C-repeat-binding factor 3. Front. Plant Sci. 6:1011.
618 Doi:10.33389/fpls.2015.01011.
619 Lin, J., Chen, Y., Xu, Y., Yan, F., Tang, L., Chen, F., 2003a. Cloning and
620 expression of curcin, a ribosome-inactivating protein from the seeds of *Jatropha*
621 *curcas*. Acta Bot Sin. 45, 858–863.
622 Lin, J., Yan, F., Tang, L., Chen, F., 2003b. Antitumor effects of curcin from seeds
623 of *Jatropha curcas*. Acta Pharmacol Sin. 24, 241–246.
624 Lin, J., Zhou, X., Wang, J., Jiang, P., Tang, K., 2010. Purification and
625 characterization of curcin, a toxic lectin from the seed of Jatropha curcas. Prep
626 Biochem Biotechnol. 40, 107-118.

627 Lipkin, A., Anisimova, V., Nikonorova, A., Babakov, A., Krause, E., Bienert, M.,
628 Grishin, E., Egorov, T., 2005. An antimicrobial peptide Ar-AMP from amaranth
629 (*Amarnathus retroflexus* L.) seeds. Phytochem. 66, 2426-2431.
630 Liu, H., Liu, Y.J., Yang, M.F., Shen, S.H., 2009. A comparative analysis of embryo
631 and endosperm proteome from seeds of *Jatropha curcas*. J Integr Plant Biol. 51,
632 850–857.
633 Liu, H., Wang, C., Chen, F., Shen, S., 2015. Proteomic analysis of oil bodies in
634 mature *Jatropha curcas* seeds with different lipid content. J. Proteomics 113,
635 403-414.
636 Liu, H., Wang, C., Komatsu, S., He, M., Liu, G., Sheri, S., 2013. Proteomic analysis
637 of the seed development in *Jatropha curcas*: From carbon flux to the lipid
638 accumulation. J. Proteomics 91, 23-40.
639 López-Ribera, I., La Paz, J.L., Repiso, C., García, N., Miguel M., Hernández, M.L.,
640 Martínez-Rivas, J.M., Vicient, C.M., 2014. The evolutionary conserved oil body
641 associated protein OBAP1 participates in the regulation of oil body size. Plant
642 Physiol. 164, 1237-1249.
643 Lüersen, K., Stegehake, D., Daniel, J., Drscher, M., Arjonina, I., Ajonina, C., Hertel,
644 P., Woltersdorf, Ch., Liebau, E., 2013. The glutathione reductase GSR-1
645 determines stress tolerance and longevity in Caenorhabditis elegans. PloS One,
646 8(4):e60731. doi: 10.1371/journal.pone.0060731.
647 Lung, S.-Ch., Chye, M.-L. 2016., Acyl-Co-A-binding proteins (ACBPs) in plant
648 development. 2016. Subcell. Biochem. 86, 363-404.
649 Luo, M.J., Yang, X.Y., Liu, W.X., Xu, Y., Huang, P., Yan, F., Chen, F., 2006.
650 Expression, purification and anti-tumor activity of curcin. Acta Biochim. Biophys.
651 Sin. 38, 663–668.
652 Maghuly, F., Laimer, M., 2013. *Jatropha curcas*, a biofuel crop: Functional
653 gneomcis for understanding metabolic pathways and genetic improvement.
654 Biotechnol. J. 8, 1172-1182.
655 Mainali, H.R., Anguraj-Vadivel, A.K., Li, X., Gijzen, M., Dhaubhadel, S., 2017.
656 Soybean cyclophiln GmCYP1 interacts with an isoflavonoid regulator
657 GmMYB176. Sci. Rep. 7:39550. doi:10.1038/srep39550.
658 Makkar, H.P.S., Becker, K., 2009. *Jatropha curcas*, a promising crop for the
659 generation of biodiesel and value-added coproducts. Eur. Lipid Sci. Technol.
660 111, 773-787.
661 Mandal, S.M., Dey, S., Mandal, M., Sarkar, S., Maria-Neto, S., Franco, O.L., 2009.
662 Identification and structural insights of three novel antimicrobial peptides isolated
663 from green coconut water. Peptides 30, 633-637.
664 Martínez-Herrera, J., Jimenez-Martínez, Ch., Martínez-Ayala, A., Garduño-
665 Siciliano, L., Mora-Escobedo, R., Davila-Ortíz, G., Chamorro-Cevallos, G.,
666 Makkar, H., Becker, K., 2012. Evaluation of the nutritional quality of nontoxic
667 kernel flour from *Jatropha curcas* L. in rats. J. Food Qual. 35, 152-158.
668 Nahar, K., Hasanuzzaman, M., Suzuki, T., Fujita, M., 2017. Polyamines-induced
669 aluminum tolerance in mung bean: A study on antioxidant defense and
670 methyulglyoxal detoxificasiton systems: Ecotoxicol. 26, 58-73.
671 Niehaus, T.D., Nguyen, T.N.D., Gidda, S.K., ElBadawi-Sidhu, M., Lambrecht, J.A.,
672 McCarty, D.R., Downs, D.M., Cooper, A.J.L., Fiehn, O., Mullen, R.T., Hanson,
673 A.D., 2014. *Arabidopsis* and maize RidA proteins preempt reactive

674 enamine/imine damage to branched-chain amino acid biosynthesis in plastids.
675 Plant Cell 26, 3010-3022.
676 Okita, T.W., Hwang, Y.S., Hnilo, J., Kim, W.T., Aryan, A.P., Larson, R., Krishnan,
677 H.B., 1989. Structure and expression of the rice glutelin multigene family. J.
678 Biol. Chem. 264, 12573-12581.
679 Openshaw, K., 2000. A review of *Jatropha curcas*: An oil plant of unfulfilled
680 promise. Biomass Bioenergy 19, 1–15.
681 Osborne, T.B., 1908. Our present knowledge of plant proteins. Science 28, 417–
682 427.
683 Pelegrini, P.B., Murad, A.M., Silva, L.P., dos Santos, R.C.P., Costa, F.T., Tagliari,
684 P.D., Bloch, C.Jr., Noronha, E.F., Miller, R.N., Franco, O.L., 2008. Identification
685 of a novel storage glycine-rich peptide from guava (*Psidium guajava*) seeds with
686 activity against Gram-negative bacteria. Peptides 29, 1271–1279.
687 Pelegrini, P.B., Noronha, E.F., Muniz, M.A., Vasconcelos, I.M., Chiarello, M.D.,
688 Oliveira, J.T., Franco, O.L., 2006. An antifungal peptide from passion fruit
689 (*Passiflora edulis*) seeds with similarities to 2S albumin proteins. Biochem.
690 Biophys. Acta 1764, 1141–1146.
691 Perea-Domínguez, X.P., Espinosa-Alonso, L.G., Hosseinian, F., HadiNezhad, M.,
692 Valdez-Morales, M., Medina-Godoy, S., 2017. Phenolic profile and antioxidant
693 activity form non-toxic Mexican *Jatropha curcas* L. shell methanolic extracts.
694 Natural Prod. Res. 31, 610-614.
695 Pinheiro, C.B., Shah, M., Soares, E.L., Nogueira, F.C.S., Carvalho, P.C.,
696 Junqueira, M., Araújo, G.D.T., Soares, A.A., Domont, G.B., Cmapos, F.A.P.,
697 2013. Proteome analysis of plastids from developing seeds of *Jatropha curcas*
698 L. J. Proteome Res.12, 5137-5145.
699 Pinto, M.E., Batista, J.M.Jr., Koehbach, J., Gaur, P., Sharma, A., Nakabashi, M.,
700 Cilli, E.M., Gissel, G.M., Verli, H., Gruber, C.W., Bianch, E.W., Tavares, J.F., da
701 Silva, M.S., Garcia, C.R., Bolsani, V.S., 2015. Ribifolin, and orbitide from
702 Jatropha ribifolia, and its potential antimalarial activity. J. Nat. Prod. 27, 374-380.
703 Pramanik, A, 2003. Properties and use of *Jatropha curcas* oil and diesel fuel
704 blends in compression ignition engine. Renew Energy 28, 239–248.
705 Qin, X., Zheng, X., Shao, C., Gao, J., Jiang, L., Zhu, X., Yan, F., Tang, L., Xu, Y.,
706 Chen, F., 2009. Stress-induced curcin-L promoter in leaves of *Jatropha curcas*
707 L. and characterization in transgenic tobacco. Planta 230, 387–395.
708 Rezig, L., Chibani, F., Chouaibi, M., Dalgalarrondo, M., Hessini, K., Guéguen, J.,
709 Hamdi, S., 2013. Pumpkin (*Cucurbita maxima*) seed proteins: sequential
710 extraction processing and fraction characterization. J. Agric. Food Chem. 61,
711 7715-7721.
712 Ribeiro, A.C., Teixeira, A.R., Ferreira, R.B., 2004. Characterization of globulins
713 from common vetch (*Vicia sativa* L.). J. Agric. Food Chem. 52, 4913-4920.
714 Richau, K.H., Kaschani, F., Verdoes, M., Pansuriya, T.C., Niessen, S., Stüber, K.,
715 Colby, T., Overkleeft, H.S., Bogyo, M., Van der Hoorn, R.A.L., 2012.
716 Subclassification and biochemical analysis of plant papain-like cysteine
717 proteases display subfamily-specific characteristics. Plant Physiol. 158, 1583-
718 1599,

719 Romero-Rodríguez, M.C., Maldonado-Alconada, A.M., Valledor, L., Jorrin-Novo,
720 J.V., 2014. Back to Osborne, sequential protein extraction and LC-MS analysis
721 for the characterization of the Holm Oak seed proteome. In: Plant Proteomics:
722 Methods and Protocols, Methods in Molecular Biology, vol. 1072, Jesus V.
723 Jorrin-Novo et al. (eds.). Springer Science+Business Media, LLC. Chapter 27,
724 pp. 379-389, DOI 10.1007/978-1-62703-631-3_27.
725 Sabandar, C.W., Ahmat, N., Jaafar, F.M., Sahidin, I., 2013. Medicinal property,
726 phytochemistry and pharmacology of several *Jatropha* species (*Euphorbiaceae*):
727 A review. Phytochem. 85, 7–29.
728 Sánchez-Riego, A.M., Mata-Cabana, A., Galmozzi, C.V., Florencio, F.J., 2016.
729 NADPH-Thioredoxin reductase C mediates the response to oxidative stress and
730 thermotolerance in the Cyanobacterium Anabaena sp. PCC7120. Front.
731 Microbiol. 18, 7:1283. doi:10.3389/fmicb.2016.01283.
732 Schlöman, M., Fischer, P., Schmidt, E., 1990. Enzymatic formation, stability, and
733 spontaneous reactions of 4-fluoromuconolactone, a metabolite of the bacterial
734 degradation of 4-fluorobenzoate. J. Bacteriol. 172, 5119-5129.
735 Shah, M., Soares, E.L., Carvalho, P.C., Soares, A.A., Domont, G.B., Nogueira,
736 F.C.S., Campos F.A.P., 2015. Proteomic analysis of the endosperm ontogeny of
737 *Jatropha curcas* L. seeds. J. Proteome Res. 14, 2557-2568.
738 Shah, M., Soares, E.L., Lima, M.L.B., Pinheiro, C.B., Soares, A.A., Domont, G.B.,
739 Nogueira, F.C.S., Campos, F.A.P., 2016. Deep proteome analysis of
740 gerntoplasts from the inner integument of developing seeds of *Jatropha curcas*.
741 J. Proteomics 143, 346-352.
742 Shen, Y.-H., Chen, Y.-H., Liu, H.-Y., Chiang, F.-Y., Wang, Y.-Ch., Hou, L.-Y., Lin,
743 J.-Sh., L, Ch.-Ch., Lin, H.-H., Lai, H.-M., Jeng, Sh.-T., 2014. Expression of a
744 gene encoding β-ureidopropionase is critical for pollen germination in tomatoes.
745 Physiol. Plantarum 150, 425-435.
746 Shih, M.-D., Huang, L.-T., Wei, F.J., Wu, M.-T., Hoekstra, F.A., Hsing, Y.I., 2010.
747 OsLEA1a, a new Em-like protein of cereal plants. Plant Cell Physiol. 51, 2132-
748 2144.
749 Soares, E.L., Shah, M., Soares, A.A., Costa, J.H., Carvalho, P., Domont, G.B.,
750 Nogueira, F.C.S., Campos, F.A.P., 2014. Proteome analysis of the inner
751 integument from developing *Jatropha curcas* L. seeds. J. Proteome Res. 13,
752 3562-3570.
753 Tack, B.F., Chapman, P.J., Dagley, S., 1972. Purification and properties of 4-
754 hydroxy-4-methyl-2-oxoglutarate aldolase. J. Biol. Chem. 247, 6444-6449.
755 Townsend, D.M., Tew, K.D., 2003. The role of glutathione-S-transferase in anti-
756 cancer drug resistance. Oncogene 22, 7369-7375.
757 Valente, M.A., Faria, J.A., Soares-Ramos, J.R., Reis, P.A., Pinheiro, G.L.,
758 Piovesan, N.D., Morais, A.T., Menezes, C.C., Cano, M.A., Fietto, L.G., Loureiro,
759 M.E., Aragao, F.J., Fontes, E.P., 2009. The ER luminal binding protein (BiP)
760 mediates an increase in drought tolerance in soybean and delays drought-
761 induced leaf senescence in soybean and tobacco. J. Exp. Bot. 60, 533-546.
762 Vallée, B.S., Coadou, G., Labbé, H., Sy, D., Vovelle, F., Schoentgen, F., 2003.
763 Peptides corresponding to the N- and C-terminal parts of PEBP are well-
764 structured in solution: New insights into their possible interaction with partners

765 in vivo. J. Peptide Res. 61, 47-57.
766 Vicient, C.M., Hull, G., Guilleminot, J., Devic, M., Delseny, M., 2000. Differential
767 expression of the *Arabidopsis* genes coding for Em-like proteins. J. Exp. Bot. 51,
768 1211-1220.
769 Wang, C.C., Tsou, C.L., 1993. Protein disulfide isomerase is both an enzyme and
770 a chaperone. FASEB J. 7, 1515-1517.
771 Werner, A.K., Medina-Escobar, N., Zulawski, M., Sparkes, I.A., Cao, F.Q.,
772 Witte,C.-P., 2013. The ureide-degrading reactions of purine ring catabolism
773 employ three amido hydrolases and one amino hydrolase in *Arabidopsis*.
774 soybean, and rice. Plant Physiol. 163, 672–681. doi:10.1104/pp.11 3.224261.
775 Xiao, J., Zhang, H., Niu, L., Wang, X., 2011. Efficient screening of a novel
776 antimicrobial peptide from *Jatropha curcas* by cell membrane affinity
777 chromatography. J. Agric. Food Chem. 59, 1145-1151.
778 Yeung, K.C., Rose, D.W., Dhillon, A.S., Yaros, D., Gustafsson, M., Chatterjee, D.,
779 McFerran, B., Wyche, J., Kolch, W., Sediviy, J.M., 2001. Raf Kinase inhibitor
780 protein interacts with NF-κB-inducing kinase and TAK1 and inhibits NF-κB
781 activation. Mol. Cell. Biol. 21, 7207-7217.
782 Yurchenko, O.P., Weselake, R.J., 2011. Involvement of low molecular mass
783 soluble acyl-CoA-binding protein in seed oil biosynthesis. N. Biotechnol. 28,
784 97–109.
785 Zancani, M., Peresson, C., Biroccio, A., Federici, G., Urbani, A., Murgia, I.,
786 Soave,C., Micali, F., Vianello, A., Macri, F., 2004. Evidence for the presence of
787 ferritin in plant mitochondria. Eur. J. Biochem. 271, 3657–3664.
788 Zang, Y., Geng, X., Wang, F., Liu, Z., Zhang, L., Zhao, Y., Tian, X, Ni, Zh., Yao, Y.,
789 Xin, M., Hu, Zh., Sun, Q., Peng, H., 2017. Overexpression of wheat ferritin gene
790 *TaFER-5B* enhances tolerance to heat stress and other abiotic stresses
791 associated with the ROS scavenging. BMC Plant Biol. 17:14, doi:
792 10.1186/s12870-016-0958-2.
793 Zhang, Y., Yang, Q., Li, Ch., Ding, M., Lv, X., Tao, Ch., Yu, H., Chen, F., Xu, Y.,
794 2017. Curcin C, a novel type I ribosome-inactivating protein from the post-
795 germination cotyledons of *Jatropha curcas*. Amino acids, 2456-8,
796 doi:10.1007/s00726-017-2456-8.
797
798
799
800
801
802
803
804

Figure legends

Fig. 1. A) One-dimension electrophoretic profile of *Jatropha curcas* seed storage proteins. Representative 2-DE gels of *J. curcas* seed storage proteins for B) albumins; C) globulins; D) glutelins. Spots numbers indicate the protein spots analyzed by LC-MS/MS.

Fig. 2. Diagram indicating the number of proteofroms from different protein species detected in seed storage proteins. A) albumins, B) globulins, C) glutelins.

Fig. 3. Metabolic network of non-toxic *J. curcas* seeds. The identified protein spots are indicated in black spots (Table 2). Diagram was generated using the visualization and analysis tool PathVisio 3.2.4.

Table 1
Protein identification of *Jatropha curcas* seed storage protein fractions (albumins, globulins and glutelins) by LC-MS/MS and homology database search.

Fraction	Spots	Identified (%)	Non-identified (%)	Unique
Albumins	434	365 (84.1)	69 (15.9)	100
Globulins	310	296 (95.5)	14 (4.5)	24
Glutelins	175	168 (96.0)	7 (5.0)	7