González-García, Estefanía et al., 2016. Identification of plum and peach seed proteins by nLC-MS/MS via combinatorial peptide ligand libraries. Journal of proteomics, 148, pp.105–112.

IDENTIFICATION OF PLUM AND PEACH SEED PROTEINS BY nLC-MS/MS VIA COMBINATORIAL PEPTIDE LIGAND LIBRARIES

Estefanía González-García[(1)], María Luisa Marina[(1)], María Concepción García[(1)], Pier Giorgio Righetti[(2)], Elisa Fasoli[(2,*)]

[(1)] Departamento de Química Analítica, Química Física e Ingeniería Química, Universidad de Alcalá, Ctra. Madrid-Barcelona Km. 33.600, 28871 Alcalá de Henares (Madrid), Spain.

[(2)] Department of Chemistry, Materials and Chemical Engineering "Giulio Natta", Politecnico di Milano, 20131 Milan, Italy.

*Corresponding author (e-mail:elisa.fasoli@polimi.it; telephone: +39-02-23994749; fax: +39-02-23993080)

ABSTRACT

Plum (*Prunus domestica* L.) and peach (*Prunus persica* (L.) Batsch) seed proteinsare a source of bioactive peptides. These seeds, though,are usual residues produced during canning and beverage preparation that, in most cases, are irreversibly lost. The recovery and identification of these proteins might be ofimportance in human nutrition. This workemploysthe combinatorial peptide ligand libraries (CPLLs) technology as a tool to reduce the proteins dynamic concentration range.The most suitable extraction and CPLL capture conditions have been obtained and applied for the comprehensive identification of seed proteins. The analysis of recovered speciesby nLC-MS/MS has allowed the identification of 141 and 97unique gene products from plum and peach seeds,respectively. It was possible to identify16proteins belonging to the *Prunus* genus. Moreover, a high number of histones and seed storage proteins were identified. Additionally, 21 and 14 bioactive peptides previously identified were found within protein sequences in plum and peach seeds, respectively.

Key words: seed proteins; plum; peach; proteomics, combinatorial peptide ligand library; nano-liquid chromatography-mass spectrometry

1. INTRODUCTION

Plums (*Prunus domestica* L.) and peaches (*Prunus persica* (L.) Batsch) are among the most produced fruits in the world, with 11,530 and 21,640 thousand tons in 2013, respectively [1].The processing of plums and peaches for the canning and beverageindustries generates a great amount of residues, where peel and stones comprise 10-25 % of the raw material in the case of plums and 22-38% in the case of peaches[2].It has been previously reported that those stones contain seeds inside with about 40% of proteins[3,4], able torelease bioactive (antioxidant and Angiotensin-converting enzyme (ACE)-inhibitory) peptidesafter enzymatic digestion[3-6].Nevertheless, no information about the proteins present in these seeds has been published.

Reliability of proteomic analysis depends largely on the protein sample preparation process[7-9].This issue becomes even more significant in the case of vegetable samples, which contain high levels of proteases and non-protein compoundssuch as phenolic compounds, lipids or secondary metabolites, able to interfere with theprotein extraction and separation[7-12].Moreover, protein extraction representsa great challenge considering the presence of low-abundance proteins and the difficulty tosolubilize them due to the presence of vacuoles and rigid cell walls[8,10,12,13].Additionally, identification of proteins is another limiting step in proteomicsdue to: (1) the presence of high-abundance proteins which mask the detection of the low-abundance species, and (2) the lack of databases available for plant organisms since the vast majority of genomes is still not sequenced[14,15].

The combinatorial ligand libraries (CPLLs) technology, emerged in 2005[16],is an excellent tool to reduce the dynamic concentration range of proteins and to allow the detection and identification of low-abundance proteins (“hidden” proteome)[17-19];

61 [20].Due to its extraordinary advantages, this technology hasbeen widely employed in
62 proteomic studies to identify proteins and peptides with either positive (antioxidant,
63 antimicrobial...) or negative (allergens) effects[21], or to detect possible frauds in food
64 products[22,23].Some proteomes studied by using this technology have been those in
65 olive[14,24], avocado[25], banana[15], mango[26], and also goat milk[27].

66 The aim of this work was to comprehensively identify proteins inplum and peach
67 seeds by nLC-MS/MS and to evaluate their potential for obtaining bioactive peptides,
68 after the application of a suitable method for the proteinextraction and a CPLLs
69 treatment for reduction of dynamic range in the samples.

70

2. MATERIALS AND METHODS

2.1. Chemicals and samples

All reagents used were of analytical grade. Hexane, acetone, methanol (MeOH), Tris(hydroxymethyl)aminomethane (Tris), hydrochloric acid (HCl), sodium chloride (NaCl), sodium dodecyl sulphate (SDS), dithiothreitol (DTT),3-[3-cholamidopropyl dimethylammonio]-1-propanesulfonate(CHAPS),ethylenediaminetetraacetic acid (EDTA), β-mercaptoethanol, glycine, ammonium sulphate, phosphoric acid, acetonitrile (ACN), trifluoroacetic acid (TFA), formic acid (FA), sodium hydroxide (NaOH), trichloroacetic acid (TCA), ammonium persulphate and ammonium bicarbonate were acquired at Sigma-Aldrich (Saint Louis, MO, USA). Blue Coomassie, Laemmli buffer, 40% acrylamide/Bis solution, N,N,N',N'-tetramethylethylenediamine (TEMED), Precision Plus Protein Standards (recombinant proteins expressed by *E. Coli* with molecular mass values of 10, 15, 20, 25, 37, 50, 75, 100, 150, and 250 kDa) and ProteoMiner^{TM}(PM) (combinatorial hexapeptide ligand library beads) were purchased from Bio-Rad Laboratories (Hercules, CA, USA).Water and acetonitrile (ACN) (OPTIMA® LC/MS grade) for LC/MS analyses were acquired at Fisher Scientific (Leicestershire, UK). Home-made-CPLLs (HM-CPLLs) were synthetized in our laboratory. Complete protease inhibitor cocktail tablets and sequencing grade trypsin were from Roche Diagnostics (Basel, Switzerland). Seeds were obtained from plums and peaches from a local supermarket (Madrid, Spain).

2.2.Plum and peach protein extraction

Protein extraction from plum and peach seeds was carried out following the procedure of González-García et al. and Vásquez-Villanueva et al.[3,4] with some modifications.

Plums and peaches were cut, stones were open with a nutcracker and seeds were ground and kept at -20ºC until use.In order to defat seeds, approximately 20 mL of hexane were used per half gram of milled seeds followed by shaking three times for 30 min. Afterwards, 200 mg of defatted seeds were washed twice with 4 mL MeOH/H_2O (80:20) and twice more with 4 mL acetone/H_2O (80:20), both solutions at -20ºC. Two different extracting buffers (10 mL) were employed: a native buffer and a denaturing one. The native buffer consisted of 50 mM Tris-HCl pH 7.4, 15 mM NaCland protease inhibitor cocktail.The denaturing buffer was constituted, in addition, by 1 % SDS and 25 mM DTT.Extraction was made by sonication for 10 min and gentle shaking overnight. Then, protein precipitation with acetone was carried out by employing an extract:acetone ratio of 1:2 and storing in the freezer overnight. Afterwards, the solution was centrifuged (30 min, 13400 rpm) and the resultant pellet evaporated until dried.

2.3.Protein capture with CPLLs

The precipitated proteins were redissolved by employing 10 mL of a solubilizing buffer consisting of 50 mM Tris-HCl pH 7.4, 15 mM NaCl, 0.5 % CHAPS, and 1 mM EDTA.Afterwards, the extract was divided into two parts and pH adjusted to optimal pHs for CPLLs: pH 7.0 for PM-CPLLs and 2.2 for HM-CPLLs. To each half, 100 µL of the corresponding libraries were added and gently shaken overnight. The CPLLs beads were recovered by filtering 3 min at 13400rpm through Micro Bio-Spin chromatographic columns (Bio-Rad). Before protein desorption, the beads were washed with water to remove all possible contaminants and unbound proteins. Protein desorption was performed by elution with 100 µL of 4 % SDS containing 20mM DTT, under boiling conditions.

2.4. SDS-PAGE

The proteins in control samples and CPLL eluates were separated by SDS-PAGE.Control samples were prepared by dissolving the pellet in 50 µL of Laemmli buffer containing 5% (v/v)β-mercaptoethanol. The CPLL eluateswere prepared by mixing 5 µL of desorbed proteins with 5 µL of Laemmli buffer. Samples were boiled for 5 min and loaded onto a home-made gel composed by a stacking gel (4%polyacrylamide, 125 mM Tris-HCl pH 6.8, 0.1% (m/v) SDS, TEMED, and APS) cast over a running gel (12% polyacrylamide, 375 mM Tris-HCl pH 8.8, 0.1% (m/v) SDS, TEMED, and APS). The gels were run in a Bio-Rad Mini-Protean system employing a Tris-glycine buffer (pH 8.3) containing 0.1% (m/v) SDS and a Tris buffer (pH 8.8) as cathodic and anodic buffers, respectively. Electrophoresis was carried out by applying 50 V until the dye front entered the running gel, followed by150 V until complete separation. For the estimation of molecular mass values,a marker ladder was used. Gels were stained with Colloidal Coomassie Blue followed by destaining witha 7 % acetic acid solution.Scanning of gelswas performed by a VersaDoc imaging system (Bio-Rad) and the images were treated with the software Quantity One (Bio-Rad).

2.5. Mass spectrometry and data analysis

The sample bands obtained by SDS-PAGE were cut out and destainingperformed by washing with acetonitrile and 50 mM ammonium bicarbonate (AmBic) at 56 ºC. Afterwards,the gel pieces were reduced and alkylated with 1.5 mg/mL DTT (in 50 mM AmBic) at 56ºC and 10 mg/mL iodoacetamide (in 50 mM AmBic) at room temperature, respectively. Finally, proteins were digested with 0.02 µg/µL trypsin (in 25 mM AmBic) at 37ºC overnight.

The tryptic digests were acidified with FA up to a final concentration of 10 % (v/v) and 8 µL of this mixture were loaded on a nano chromatographic system, UltiMate 3000 RSLCnano System (Dionex, Sunnyvale, CA, USA). Prior to the chromatographic

144 separation, samples were cleaned up and pre-concentratedusing a reversed-phase trap
145 column Acclaim® PEPMap100 (C18, 100 Å, 10 µm i.d. x 2 cm) from Dionex. Next, the
146 trap column was located in series with the separation column by switching the 2-
147 position valve. The separation column was a fused silica reversed-phase PicoFrit (C18,
148 2.7 µm) from New Objective (Woburn, MA, USA). Peptide elution was performed with
149 the following chromatographic conditions: mobile phase A, H_20/ACN (98/2) with 0.1%
150 FA; mobile phase B, H_20/ACN (2/98) with 0.1% FA; elution gradient, 4-60% B in 30
151 min; flow rate, 300 nL/min; and temperature, 25ºC. The liquid chromatographic system
152 was connected to a LTQ-XL mass spectrometer (Thermo Scientific) equipped with a
153 nano spray ion source. Full scan mass spectra were acquired in the mass range from350
154 to 1800 *m/z*and the fivemost intense ions were automatically selected and fragmented in
155 the ion trap. The targeted ions already selected for mass spectrometry (MS/MS) and
156 fragmented were dynamically excluded for 30 s.The mass runs were performed in
157 triplicate for the final search and protein identification.MS/MS data was processed by
158 the Proteome Discoverer software (v. 1.2.0 Thermo) and by using Mascot search engine
159 (Matrix Science, London, UK, version 2.3.01). Data were searched against
160 Uniprot_*Viridiplantae* database (937204 residues, 321079394 sequences). Cysteine
161 alkylation and oxidation of methionine were set as variable modifications. Peptide mass
162 tolerance was fixed to 1 Da, fragment mass tolerance was set to 0.8 Da, and an ion
163 source cut-off of 20 was selected. The false discovery rate was less than 1%.

164 Gene Ontology (GO) searches were performed using the QuickGO software
165 (www.ebi.ac.uk/QuickGo).

166

3. RESULTS

Proteins were firstly extracted directly from plum and peach seeds using denaturing or native extraction conditions.

Figure 1.A shows the SDS-PAGE profiles of proteins obtained from plum seeds. Lane 1 displays the pattern obtained by using the denaturing solution as extracting buffer.It was possible to observe several bands between 75 and 30 kDa and two very intense bands below 25 kDa. The CPLLs captures at pH 7.0 and at pH 2.2(lanes 2 and 3, respectively)showed few bands corresponding to high-abundance proteinsalso observed in the control obtained under denaturing conditions. Nevertheless, no low-abundance proteinswere observed in addition. This behavior was due to the interference of SDS in the protein capture. Therefore, a previous acetone precipitation step was inserted to remove this interference and the new patterns were shown in lanes 4 and 5. Especially at pH 2.2, proteins displayed additional bands between 35 and 25 kDa that were probably due to low-abundance species. In addition, plum proteins were extracted with a native buffer (lane 6) exhibitinga similar profileas theone observedunder the denaturing conditions. Lanes 7-8 and 9-10 showproteins patterns of the CPLLscaptures without or withthe acetone precipitation step. In both cases,a higher number of protein bands were obtained fromthe CPLLs captures at pH 2.2. All bands were cut out along the lanes (by the brackets) and, after trypsin digestion, analyzed by nLC-MS/MS.For the identification of proteins, Uniprot_*Viridiplantae*database was used.

TheVenn diagram presented in **Figure 1.B**and the **Supplementary Table 1** display the number of identified proteins both in controls (native and denaturing buffers) and CPLLs captures (pH 7.0 and 2.2) obtained by comparing the MS data against the Uniprot_*Viridiplantae* database. A total of 134 proteinswere totally identified, 14 of them specifically observed in the CPLLs eluates.

An additional experiment was addressed to study the effect of the proteins defatting of seeds in the number of identified proteins. **Figure 2.A** shows the proteins profiles obtained after hexane extractionto remove interfering lipids. Unlike in previous experiments, the extraction using native conditions in absence of SDS showed a higher number of extracted proteins. Control (lane 1) wassimilar to the previous one without the hexane protocol, while the CPLL eluates provided a slight increase of intensity at pH 7.0 (lane 3) and a huge increase at 2.2 (lane 2). Ten gel pieces of each lane were cut out, digested and analyzed by MS in tandem.

Figures2.Band **2.C** display the Venn diagrams obtained by comparing the identified proteins both in controls and CPLLs eluates, before and after defatting, using the Uniprot_*Viridiplantae* database. Up to 36 proteins were identified in defatted seeds, 13 species being specific of CPLLs eluates. Furthermore, 4 and 9 unique gene products that were,respectively,recognized in defatted control and CPLLs captures, could not be found infatty samples.

Supplementary Table 2shows protein identifications in both defatted controls and related CPLL eluates by employing Uniprot_*Viridiplantae* database.

The next sample analyzed was the peach seed. Since defattedsamples have resulted clearer solutions, defatted extraction protocols were usedinnative and denaturing conditions.

Figure 3.Adisplays the protein profiles obtained from the peach seeds, after extraction with both native and denaturing buffers. Only in the case of the native buffer, CPLL captureswere characterized by numerous intensebands, probably due to

thecapture of low-abundance proteins. After MS analysis performedusing Uniprot_*Viridiplantae* database, 97 unique gene products were identified.

TheVenn diagrams,reported in **Figure 3.B**, show that 39 of these species could be identified only after CPLLs treatment. A complete list of identified peach proteins is shown in **Supplementary Table 3**.

Figure 4 displays the Venn diagram, comparing total number of proteinsidentified in plum seeds (141) versus those identified in peach seeds (97).Thirty-sixproteins werepresentin both fruits.

Figure 5.A represents the Pie Chart of molecular functions related to identified proteins found both in plum and in peach seeds, while in **Figure 5.B** a diagram graph compares the molecular functions of unique gene products specifically found in plum or peach seeds.

Moreover,**Table 1** shows a list of the 16unique gene products belonging to*Prunus*genus identified in both seeds.Finally, **Table 2** shows the bioactive peptides previously observed in plum and peach[3-6]andwere now assigned to proteins identified in this work.

4. DISCUSSION

Although plum and peach seeds are considered important sources of proteins and energy for plant, the knowledge of their specific proteome is poor also due to the fact that recent literature has spent efforts into a metabolomics search involved in organoleptic properties and nutritional properties of fruits[28] and into an exploration of antioxidant activity and bioactivity of fruits components[29-32]. As regards seed, only few studies are published on peach or plum kernels connected with their bioactivity[33,34] and with the possibility of using the oil obtained from stones in biodiesel production.

Because of paucity of proteomic research[36], it could be important to deeply investigate the seeds protein content in order to understand the plant biological role and also to evaluate the potential reuse of kernels, normally discarded. The CPLL technology was applied for a deeper detection of the "hidden" proteome: even if CPLLs capture was not able to increase substantially the number of identified proteins in both seeds, we had the possibility to detect 14 and 39 more proteins in plum and peach, respectively. As regards plum seeds, we have initially evaluated the possibility that fats could prevent the proteome exploitation, interfering with CPLLs interaction. For this reason, we performed the CPLLs incubation in plum seed's extract before (**Figure 1B**) and after (**Figure 2B**) the application of defatting protocol. Even if the results obtained for plum samples were not encouraging, we have decided to investigate peach seeds proteome removing any fatty contaminants. In this case CPLLs methodology has increased of 40% the number of identified proteins (Fig. 3B). The different number of identified proteins in CPLLs eluates was probably due to specific components able to interact with beads preventing the binding with proteins. On the other hand, all stones of species belonging to *Rose* family, like plum and peach, are well known for their content

of cyanogenic glycosides, like amygdalin, able to decompose into sugar and hydrogen cyanide gas, hazardous to human health, while our research was the first exploitation on such seeds proteome, important to understand proteins role on plant growth.After CPLLs incubationand MS analysis,the final data elaboration was aimed to investigate the biological roles of proteins and to find out precursors of bioactive peptides. The Gene Ontology analysis, performed in proteins commonly present in plum and peach seeds, has demonstrated an enrichment in GO categories related to nutrient reservoir activity (27%), to sugar metabolism (14% GO: 0004332) and to energy release (9% GO: 0005524).Our results have revealed the presence of common protein classification found in fruit mesocarp like oxidoreductase activity and energy metabolism as reported by Almeida et al. [36]: the correlation between seeds, primary source of energy for plant growth, and fruits, products of plants was evident. Another recent manuscript [28] has described the modification of metabolites levels, like sugars and amino acids, during ripening. Our protein identifications have elucidated the contribution of proteome in the role of kernels during plant's growth as revealed recent differential proteomic studies [37]. Parreira et al. [37] have found that in later stages of bean seed development proteins, belonging to redox, catalytic processes, protein modification and nucleic acid metabolism, were highly expressed reflecting the activation of seed desiccation-resistance mechanism. Their results were similar to our profile of molecular functions (Fig. 5B), probably due to accumulation of fruits storage.The GO analysis was applied to proteins identified in one type of seed. Only in plum seeds, a high percentage of species has shown GTP and NAD binding activity. The GTP binding proteins, often connected with NADH oxidase activity of plant plasma membrane[38], are important switches between an active and an inactive state of metabolic cycles. There are many GTPases present in various plant-signaling pathways able to show their regulatory

function in many plant processes[39]. In peach seeds, the most enriched GO categories were DNA binding activity, calcium-dependent phospholipid binding, lactoylglutathione lyase activity and protein heterodimerization activity. Lactoylglutathione lyase activity is fundamental for detoxification of methylglyoxal, normally produced by cell and cytotoxic at middle concentration. The modulation of such activity is very important to control cell growth. In 1993 Paulus et al. [40]proposed a protein with such function as a general biomarker for soybean cell proliferation and differentiation. Also proteins with calcium-dependent phospholipid-binding activity could be connected in plant signal transduction pathways regulating stress-induced cell growth. In 2011, De Silva et al. [41]have discovered in Arabidopsis Thaliana a calcium-dependent lipid-binding protein able to confer high salt tolerance and resistance to enhanced drought.

Only few proteins (10% in the case of plum and 11% for the peach seeds)were identified by searching against *Prunus* genus database, while the vastmajoritywere identifiedbyhomology toother plant species: vine (*Vitis vinifera*), soy (*Glycine max*), poplar (*Populus trichocarpa*), and corn (*Zea mays*), among many more species detected with lower percentages.

Regarding the biological functions, we could highlight the presence of various isoforms of histones and globulins which are typical proteins in seeds. Histones are basic proteins able to pack DNA, while globulins are storage proteinsalways present in seeds. A total of 2 and 5 histones, belonging to genus different from *Prunus,* were recognizedin plums and peaches, respectively: a histone H3 (from *Zea mays*) and a histone H3-like 5 (from *Arabidopsis thaliana*) in plum seeds and a histone H4 (from *glycine max*) and four isoforms of histone H2A (from *Vitis vinifera*, *Sorghum bicolor*, *Selaginella moellendorffii*, and *Zea mays*) in peach seeds. With respect to globulins,

four 11S globulin isoforms from *Prunus dulcis* were identified in both seeds, which are the seed storage proteins: Prunin 1, Prunin 2, Pru 2 protein, and Prunin (fragment).Furthermore, several globulins from other species were also identified:two 11S globulin isoforms and a 7S globulin isoform belonging to *Ficus awkeotsang*; five 11S globulins from *Glycine max*, and others from *Castanea crenata*, *Tritic aestivum*, *Lupinus angustifolius* or *Glycine soja*. In total, 16 different globulins were identified.

Regardingthe 16identified proteins from the *Prunus*genus, a total of 8 were attributed to the almond (*Prunus dulcis*), 7 to the black cherry (*Prunus serotina*),and 1 to the peach (*Prunus persica* (L.) Batsch).Considering *Prunus dulcis* proteins identified in both plum and peach seeds, one of them was a well-known allergen: a lipid transfer protein. Recent literature[41] has reported the relationship between severe allergic reaction to peach and tomato, demonstrating the presence of anti-rPrup 3 IgE may be an indicator of mild tomato allergy.The other gene productswere proteins with nutrient reservoir activity, lyases involved in the cyanogenesis, antimicrobial peptides and a glycosidase.The protein belonging to peach was a superoxide dismutase, which is an important antioxidant defense in cells exposed to oxygen.

Moreover, previous works have proposed the valorization of peach and plum seeds proteins through the evaluation of the presence of bioactive peptides [3-6]. Although the direct study of the seeds did not show significant bioactivity, the digestion of peach and plum seeds proteins with non specific enzymes such as Alcalase and Thermolysin generated short amino acid sequences that demonstrated significant antioxidant and antihypertensive activities. Bioactive peptides were also identified by *de novo* sequencing but it was not possible to assign these peptides to any peach or plum seed proteins since they were unknown and they have never been sequenced. This work has made possible the assignment of peptides with antioxidant and antihypertensive

activities, that were previously described, to specific plum and peach seed proteins. As listed in **Table 2**, 21 bioactive peptides from the 35 peptides exerting bioactivity [3,5] were assigned to specific plum seed proteins: 12 peptides exerting *in vitro* antioxidant capacity and 9 peptides with *in vitro* ACE-inhibitory capacity. Similarly, in the case of peach seeds [4,6], 14 bioactive peptides from the 23 peptides exerting bioactivity were assigned to specific peach seed proteins: 10 antioxidant peptides, 2 ACE-inhibitory peptides, and two additional peptides showing both activities. All these peptides, in both seeds, were concentrated in a short number of proteins and especially in the three big seed storage proteins (see **Table 2**) (21). Similarly, in the case of the peach seeds (14 peptides), most bioactive peptides were found in six of the proteins. Moreover, some of these peptides were simultaneously identified within the sequence of two or even three different proteins.

To conclude, the present research project has been focused in the recovery and identification of proteins from plum and peach seeds. Those proteins have been extracted by employing two different buffers in order to find out the best extraction protocol for the incubation with CPLLs.The best extraction protocol requires the previous elimination of fat and the final precipitation of proteins with acetone to remove SDS. Although both controls have revealed similar protein profiles, after CPLLs treatment the protein bands appeared more intense.The CPLL capture allowed an increase inthe number of identified proteins up to 67% in the case of peach seeds.A total of 141and 97unique gene products were identified in plum and peach seeds, respectively, and 16proteins belonged to the *Prunus*genus. Among the identified proteins there were several histones and seed storage proteins. Several sequences of antioxidant and antihypertensive peptides, previously reported in plum and peach seeds, have been found within the identified proteins sequences.These results suggest that it

would be worth to set up a large-scale industrial process for recovery of these valuable proteins and for extraction of thebioactive peptides, which could be profitably adopted in human nutrition.

FUNDING SOURCES

This work was supported by the Ministry of Economy and Competitiveness (AGL2012-36362) and the Comunidad de Madrid and European funding from FEDER program (ref. S2013/ABI-3028, AVANSECAL). E.G.-G. thanks the Universidadde Alcalá for her predoctoral contract and mobility grant to stay at the Politecnico di Milano.

REFERENCES

[1] FAOSTAT (Food and Agricultural Organization of the United Nations), http://faostat3.fao.org., 2015 (accessed 27.01.2016).

[2] Ros M, Pascual JA, Ayuso M, Morales AB, Miralles JR, Solera C. Salidas valorizables de los residuos y subproductos orgánicos de la industria de los transformados de frutas y hortalizas: proyecto Life+ Agrowaste. Residuos: Revista Técnica 2012;22:28-35.

[3] González-García E, Marina ML, García MC. Plum (*Prunus Domestica* L.) by-product as a new and cheap source of bioactive peptides: Extraction method and peptides characterization. J Funct Foods2014;11:428-37.

[4] Vásquez-Villanueva R, Marina ML, García MC. Revalorization of a peach (*Prunus persica* (L.) Batsch) byproduct: Extraction and characterization of ACE-inhibitory peptides from peach stones. J Funct Foods 2015;18:137-46.

[5] González-García E, Puchalska P, Marina ML, García MC. Fractionation and identification of antioxidant and angiotensin-converting enzyme-inhibitory peptides obtained from plum (*Prunus domestica L.*) stones. J Funct Foods2015;19:376-84.

[6] Vásquez-Villanueva R, Marina ML, García MC. Identification by hydrophilic interaction and reversed-phase liquid chromatography-tandem mass spectrometry of peptides with antioxidant capacity in food residues. Journal of Chromatogr A 2016;1428;1428:185-92.

[7] Wang W, Tai F, Chen S. Optimizing protein extraction from plant tissues for enhanced proteomics analysis. J Sep Sci 2008;31:2032-9.

[8] Wu X, Gong F, Wang W. Protein extraction from plant tissues for 2DE and its application in proteomic analysis. Proteomics 2014;14:645-58.

[9] Wang W, Vignani R, Scali M, Cresti M. A universal and rapid protocol for protein extraction from recalcitrant plant tissues for proteomic analysis. Electrophoresis 2006;27:2782-6.

[10] Wongpia A, Mahatheeranont S, Lomthaisong K, Niamsup H. Evaluation of Sample Preparation Methods from Rice Seeds and Seedlings Suitable for Two-Dimensional Gel Electrophoresis. Appl Biochem Biotechnol 2015;175:1035-51.

[11] Banasiak A. Evolution of the cell wall components during terrestrialization. Acta Soc Bot Pol 2014;83:349-62.

[12] Zheng Q, Song J, Doncaster K, Rowland E, Byers DM. Qualitative and quantitative evaluation of protein extraction protocols for apple and strawberry fruit suitable for two-dimensional electrophoresis and mass spectrometry analysis. J Agric Food Chem 2007;55:1663-73.

[13] Garcia MC, Puchalska P, Esteve C, Marina ML. Vegetable foods: A cheap source of proteins and peptides with antihypertensive, antioxidant, and other less occurrence bioactivities. Talanta 2013;106:328-49.

[14] Esteve C, D'Amato A, Marina ML, García MC, Citterio A, Righetti PG. Identification of olive (*Olea europaea*) seed and pulp proteins by nLC-MS/MS via combinatorial peptide ligand libraries. J Proteomics 2012;75:2396-403.

[15] Esteve C, D'Amato A, Marina ML, García MC, Righetti PG. In-depth proteomic analysis of banana (*Musa spp.)* fruit with combinatorial peptide ligand libraries. Electrophoresis 2013;34:207-14.

[16] Thulasiraman V, Lin SH, Gheorghiu L, Lathrop J, Lomas L, Hammond D, Boschetti E. Reduction of the concentration difference of proteins in biological liquids using a library of combinatorial ligands. Electrophoresis 2005;26:3561-71.

[17] Righetti PG, Candiano G, Citterio A, Boschetti E. Combinatorial peptide ligand libraries as a "Trojan Horse" in deep discovery proteomics. Anal Chem 2015;87:293-305.

[18] Boschetti E, Righetti PG. Optimized sample treatment protocol by solid-phase peptide libraries to enrich for protein traces. Amino Acids 2013;45:1431-42.

[19] Righetti PG, Boschetti E, Kravchuk AV, Fasoli E. The proteome buccaneers: how to unearth your treasure chest via combinatorial peptide ligand libraries. Expert Rev Proteomics 2010;7:373-85.

[20] Boschetti E, Righetti PG. Hexapeptide combinatorial ligand libraries: the march for the detection of the low-abundance proteome continues. BioTechniques 2008;44:663-5.

[21] Righetti PG, Esteve C, D'Amato A, Fasoli E, Marina ML, García MC. A sarabande of tropical fruit proteomics: Avocado, banana, and mango. Proteomics 2015;15:1639-45.

[22] [22] Fasoli E, Colzani M, Aldini G, Citterio A, Righetti PG. Lemon peel and Limoncello liqueur: A proteomic duet. Biochim Biophys Acta-Proteins and Proteomics2013;1834:1484-91.

[23] Lerma-García MJ, D'Amato A, Fasoli E, Simó-Alfonso EF, Righetti PG. According to the CPLL proteome sheriffs, not all aperitifs are created equal!Biochim Biophys Acta-Proteins and Proteomics 2014;1844:1493-9.

438 [24] Esteve C, D'Amato A, Marina ML, García MC, Righetti PG. Analytical
439 Approaches for the Characterization and Identification of Olive (Olea europaea)
440 Oil Proteins. J Agric Food Chem 2013;61:10384-91.

441 [25] Esteve C, D'Amato A, Marina ML, García MC, Righetti PG. Identification of
442 avocado (*Persea americana*) pulp proteins by nano-LC-MS/MS via combinatorial
443 peptide ligand libraries. Electrophoresis 2012;33:2799-805.

444 [26] Fasoli E, Righetti PG. The peel and pulp of mango fruit: A proteomic samba.
445 Biochim Biophys Acta-Proteins and Proteomics2013;1834:2539-45.

446 [27] Cunsolo V, Fasoli E, Saletti R, Muccilli V, Gallina S, Righetti PG, Foti S. Zeus,
447 Aesculapius, Amalthea and the proteome of goat milk. J Proteomics 2015:69-82.

448 [28] Monti LL, Bustamante CA, Osorio S, Gabilondo J, Borsani J, Lauxmann MA,
449 Maulion E, Valentini G, Budde CO, Fernie AR, Lara MV, Drincovich MF.
450 Metabolic profiling of a range of peach fruit varieties reveals high metabolic
451 diversity and commonalities and differences during ripening. Food Chem
452 2016;190:879-88.

453 [29] Belhadj F, Somrani I, Aissaoui N, Messaoud C, Boussaid M, Marzouki MN.
454 Bioactive compounds contents, antioxidant and antimicrobial activities during
455 ripening of *Prunus persica* L. varieties from the North West of Tunisia. Food
456 Chem 2016;204:29-36.

457 [30] Morabbi Najafabad A, Jamei R. Free radical scavenging capacity and antioxidant
458 activity of methanolic and ethanolic extracts of plum (*Prunus domestica* L.) in
459 both fresh and dried samples. Avicenna J Phytomed 2014;4:343-53.

460 [31] Kayano SI, Kikuzaki H, Yamada NF, Aoki A, Kasamatsu K, Yamasaki Y, Ikami
461 T, Suzuki T, Mitani T, Nakatani N. Antioxidant properties of prunes (*Prunus*
462 *domestica* L.) and their constituents. Biofactors 2004;21:309-13.

463 [32] Popov SV, Ovodova RG, Golovchenko VV, Khramova DS, Markov PA, Smirnov
464 VV, Shashkov AS, Ovodov YS. Pectic polysaccharides of the fresh plum *Prunus*
465 *domestica* L. isolated with a simulated gastric fluid and their anti-inflammatory
466 and antioxidant activities. Food Chem 2014;143:106-13.

467 [33] Loizzo MR, Pacetti D, Lucci P, Nunez O, Menichini F, Frega NG, Tundis R.
468 *Prunus persica var. platycarpa* (Tabacchiera Peach): bioactive compounds and
469 antioxidant activity of pulp, peel and seed ethanolic extracts. Plant Food Hum
470 Nutr 2015;70:331-7.

471 [34] Gornas P, Rudzinska M, Raczyk M, Misina I, Soliven A, Lacis G, Seglina D.
472 Impact of species and variety on concentrations of minor lipophilic bioactive
473 compounds in oils recovered from plum kernels. J Agric Food Chem
474 2016;64:898-905.

475 [35] Kostic MD, Velickovic AV, Jokovic NM, Stamenkovic OS, Veljkovic VB.
476 Optimization and kinetic modeling of esterification of the oil obtained from waste
477 plum stones as a pretreatment step in biodiesel production. Waste Manage
478 2016;48:619-29.

479 [36] Miyasaka Almeida A, Urra C, Moraga C, Jego M, Flores A, Meisel L, González
480 M, Infante R, Defilippi BG, Campos-Vargas R, Orellana A. Proteomic analysis of
481 a segregant population reveals candidate proteins linked to mealiness in peach. J
482 Proteomics 2016;131:71-81.

[37] Parreira J. R., Bouraada J., Fitzpatrick M. A., Silvestre S., Bernardes da Silva A., Marques da Silva J., Almeida A. M., Fevereiro P., Altelaar A. F. M., Araujo S. S. Differential proteomics reveals the hallmarks of seed development in common bean (*Phaseolus vulgaris L.*). J. Proteomics 2016 doi: 10.1016/j.jprot.**2016**.03.002.

[38] Morre DJ, Brightman AO, Barr R, Davidson M, Crane FL. Nadh Oxidase activity of plasma-membranes of soybean hypocotyls is activated by guanine-nucleotides. Plant Physiol 1993;102:595-602.

[39] Bischoff F, Molendijk A, Rajendrakumar CSV, Palme K. GTP-binding proteins in plants. Cell Mol Life Sci 1999;55:233-56.

[40] Paulus C, Kollner B, Jacobsen HJ. Physiological and biochemical-characterization of glyoxalase-I, a general marker for cell-proliferation, from a soybean cell-suspension. Planta 1993;189:561-6.

[41] de Silva K, Laska B, Brown C, Sederoff HW, Khodakovskaya M. Arabidopsis thaliana calcium-dependent lipid-binding protein (AtCLB): a novel repressor of abiotic stress response. J Exp Bot 2011;62:2679-89.

[42] Mascheri A, Farioli L, Pravettoni V, Piantanida M, Stafylaraki C, Scibilia J, Mirone C, Preziosi D, Nichelatti M, Pastorello EA. Hypersensitivity to tomato (*Lycopersicon esculentum*) in peach-allergic patients: rPru p3 and rPru p1 are predictive of symptom severity. J Investig Allergol Clin Immunol 2015;25:183-9.

Figure Captions

Figure 1. A. SDS-PAGE profiles of proteinsextracted from plum seeds using different extracting buffers (native buffer (NB) (lanes 1-5) and denaturing buffer (DB) (lanes 6-10)), with or withoutprecipitation step, and different CPLLs employed (pH 7.0 and pH 2.2). Brackets represent cut gel pieces, submitted to digestion and MS analysis. **B.** Venn diagram displaying the number of identified proteins in controls (native and denaturing buffers) versus identified in CPLL captures (pH 7.0 and 2.2) by Uniprot_*Viridiplantae* database.

Figure 2. A. SDS-PAGE profiles of proteins extracted from defatted plum seeds under native conditions (NB) and the CPLL at pH 7.0 and 2.2. Brackets symbolize cut-out bands, submitted to digestion and MS analysis. **B.** Venn diagram showing the number of identified proteins in control versus CPLL eluates (pH 7.0 and 2.2) byUniprot_*Viridiplantae* database. **C.** Venn diagram displaying proteins identified in non-defatted seeds versus defatted seeds (Uniprot_*Viridiplantae* database).

Figure 3. A. SDS-PAGE profiles of proteins extracted from peach seeds using different extracting buffers (native buffer (NB) (lanes 1-3) and denaturing buffer (DB) (lanes 4-6)) and different CPLL eluates (pH 7.0 and 2.2). Brackets represent cut-out bands, submitted to digestion and MS analysis. **B.** Venn diagram displaying the number of identified proteins in controls (native and denaturing buffers) versus those identified in CPLL captures (pH 7.0 and 2.2) (Uniprot_*Viridiplantae* database).

Figure 4. Venn diagram revealing the total number of proteins identified in plum seeds versus peach seedsin both controls and CPLLs eluates (Uniprot_*Viridiplantae* and Uniprot_*Prunus* databases).

Figure 5. A. Gene Ontology (GO) analysis related to identified proteins from both plum and peach seeds. **B.** Diagram graph comparing the molecular functions of proteins specially found in plum or peach seeds.

532 **Table 1.** List of proteins (accession number, name, mascot score and molecular weight) belonging to the *Prunus* genus identified in plum and

533 peach seeds.

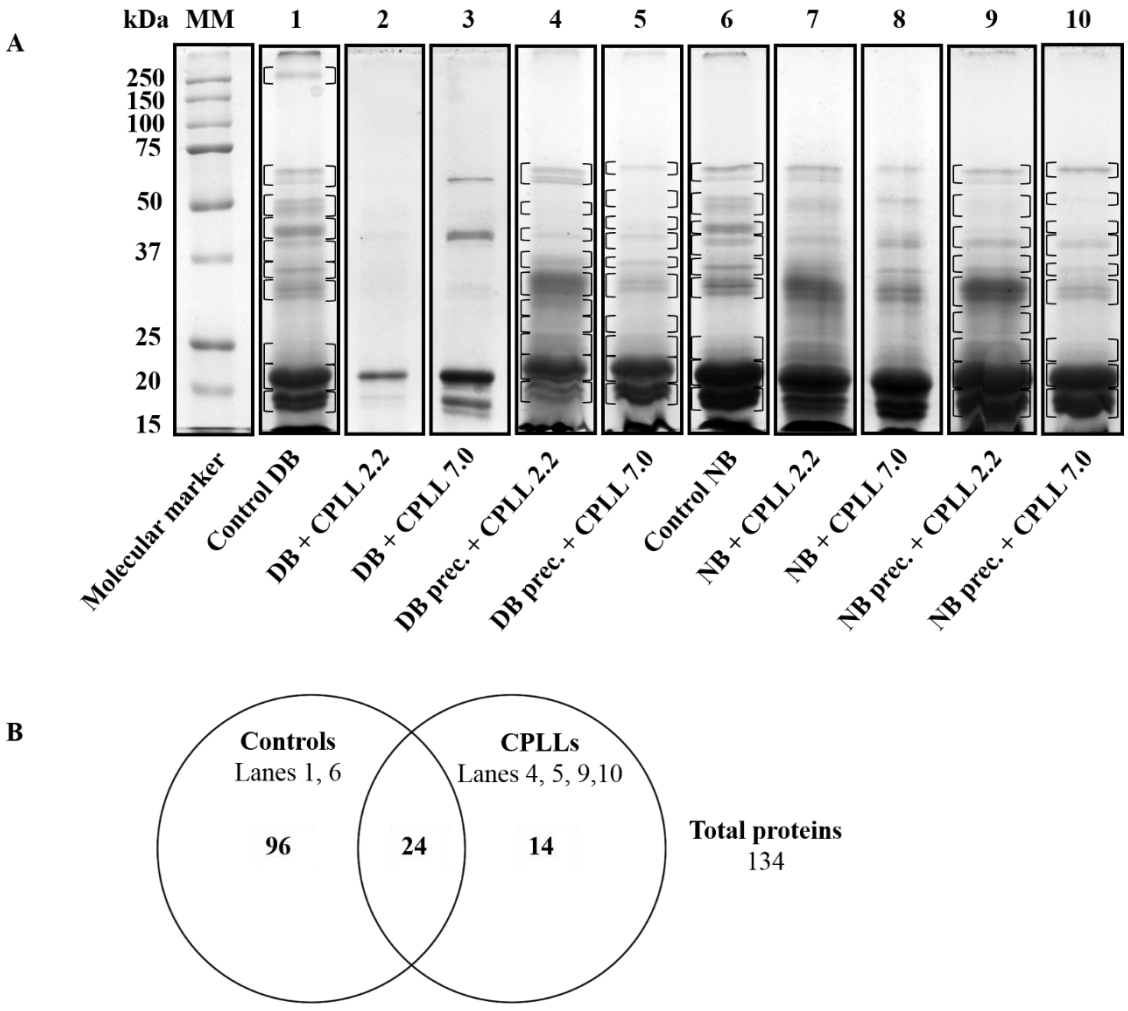
Fruit/s	Accession number	Protein name	Mascot score	Mw
Plum	tr|Q945G5|Q945G5_PRUSE	Prunasin hydrolase isoform PH I (Fragment) OS=Prunus serotina GN=PH-S1 PE=3 SV=1	402	58677
	tr|Q8W1W7|Q8W1W7_PRUSE	Prunasin hydrolase isoform PH B OS=Prunus serotina GN=PH-L3 PE=2 SV=1	158	61822
	tr|Q9M5X5|Q9M5X5_PRUSE	Prunasin hydrolase isoform PHA OS=Prunus serotina PE=2 SV=1	109	61348
	tr|A7Y7H1|A7Y7H1_PRUDU	Putative uncharacterized protein (Fragment) OS=Prunus dulcis PE=2 SV=1	70	19416
	tr|Q945G7|Q945G7_PRUSE	Amygdalin hydrolase isoform AH I (Fragment) OS=Prunus serotina GN=AH1 PE=3 SV=1	54	60531
Peach	SODM_PRUPE	Superoxide dismutase [Mn], mitochondrial OS=Prunus persica GN=SOD PE=2 SV=1	84	25439
	tr|Q9M5X4|Q9M5X4_PRUSE	Putative prunasin hydrolase isoform PH-L1 OS=Prunus serotina GN=PH-L1 PE=2 SV=2	81	62142
Plum and peach	tr|E3SH29|E3SH29_PRUDU	Prunin 2 (Fragment) OS=Prunus dulcis PE=2 SV=1	5306	57199
	tr|Q43608|Q43608_PRUDU	Pru2 protein (Fragment) OS=Prunus dulcis GN=pru2 PE=2 SV=1	2999	57245
	tr|E3SH28|E3SH28_PRUDU	Prunin 1 OS=Prunus dulcis PE=2 SV=1	2498	63356
	sp|P52706|MDL1_PRUSE	(R)-mandelonitrile lyase 1 OS=Prunus serotina GN=MDL1 PE=1 SV=1	927	61447
	tr|Q945K2|Q945K2_PRUDU	R-oxynitrile lyase isoenzyme 1 OS=Prunus dulcis GN=hnl1 PE=1 SV=1	829	61405
	tr|Q8W594|Q8W594_PRUSE	Prunasin hydrolase isoform PH C OS=Prunus serotina GN=PH-L4 PE=2 SV=1	289	61784
	tr|B5LXD2|B5LXD2_PRUDU	Prunin (Fragment) OS=Prunus dulcis PE=4 SV=1	275	9932
	tr|A7Y7K3|A7Y7K3_PRUDU	Putative lipid transfer protein (Fragment) OS=Prunus dulcis PE=2 SV=1	88	7255
	tr|B5LXC9|B5LXC9_PRUDU	Antimicrobial peptides 2-1 (Fragment) OS=Prunus dulcis PE=4 SV=1	83	17513

534 **Table 2.** List of peptides sequences found within the protein sequences from the *Prunus*

535 genus identified in plum and peach seeds.

Accession number	Peptides identified			
	Plum[3,5]		**Peach**[4,6]	
	Antioxidant	**Antihypertensive**	**Antioxidant**	**Antihypertensive**
tr|Q9M5X4|Q9M5X4_PRUSE	KGVI		-	-
tr|E3SH29|E3SH29_PRUDU	LPAGV; DQVPR; IVRVQ; HLPIL; HLPILR; KGVL; KGVLY	LLAQA; IYTPH	LHLPS; VLYN; IYTPH; VITQ; FEYI; ALPDEV	IYTPH
tr|Q43608|Q43608_PRUDU	LPAGV; DQVPR; IVRVQ; HLPIL; HLPILR; KGVL; KGVLY	LLAQA; IYTPH; ILNDE	VLYN; IYTPH; VITQ; FEYI; ALPDEV; ILNDE	IYTPH
tr|E3SH28|E3SH28_PRUDU	IPAGV; HNLPIL; NLPIL	VAVNL; LANGPENE; LLAQA; IFSPR; IYSPH; GIYSPH; VVYV	LHLPS; LANGPENE; IYSPH; ILDQE; ALPDEV	FNTQ; IYSPH
sp|P52706|MDL1_PRUSE	-	-	-	ILGIPLPK
tr|Q945K2|Q945K2_PRUDU	YLSF	-	VLYI	ILGIPLPK
tr|B5LXD2|B5LXD2_PRUDU	-	-	LHLPS	-

536



537

538 **Figure 1**

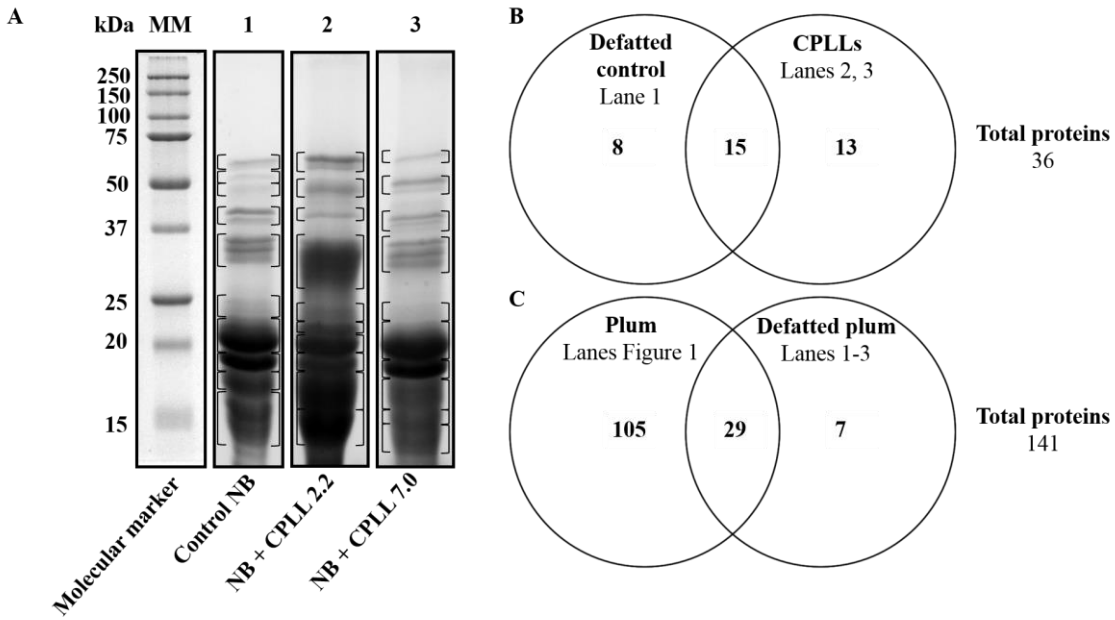
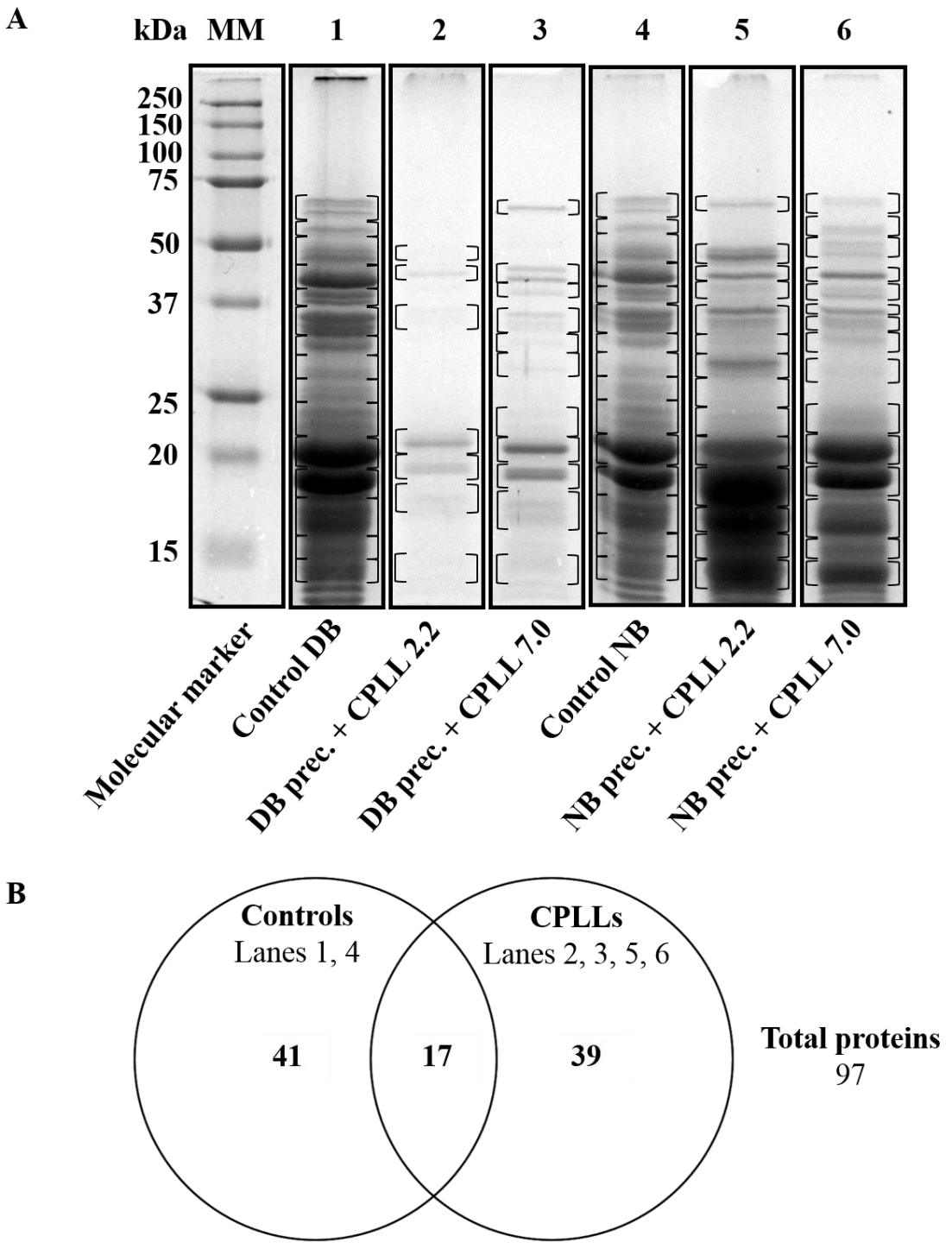
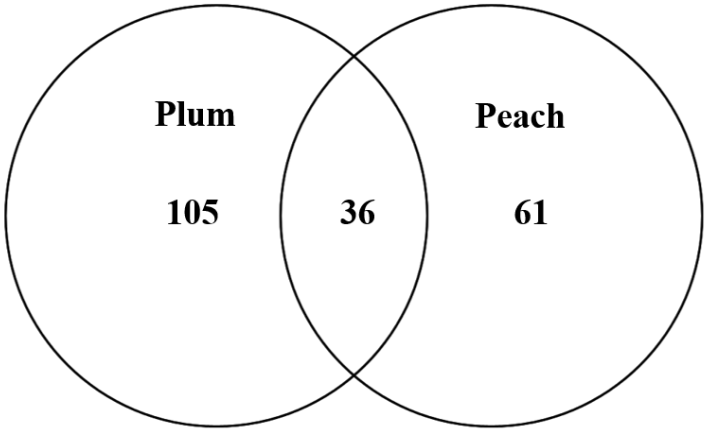

A
kDa MM 1 2 3
250
150
100
75
50
37
25
20
15
Molecular marker
Control NB
NB + CPLL 2.2
NB + CPLL 7.0
B
Defatted control
Lane 1
8
15
CPLLs
Lanes 2, 3
13
Total proteins
36
C
Plum
Lanes Figure 1
105
29
Defatted plum
Lanes 1-3
7
Total proteins
141

Figure 2



541

542 **Figure 3**



543

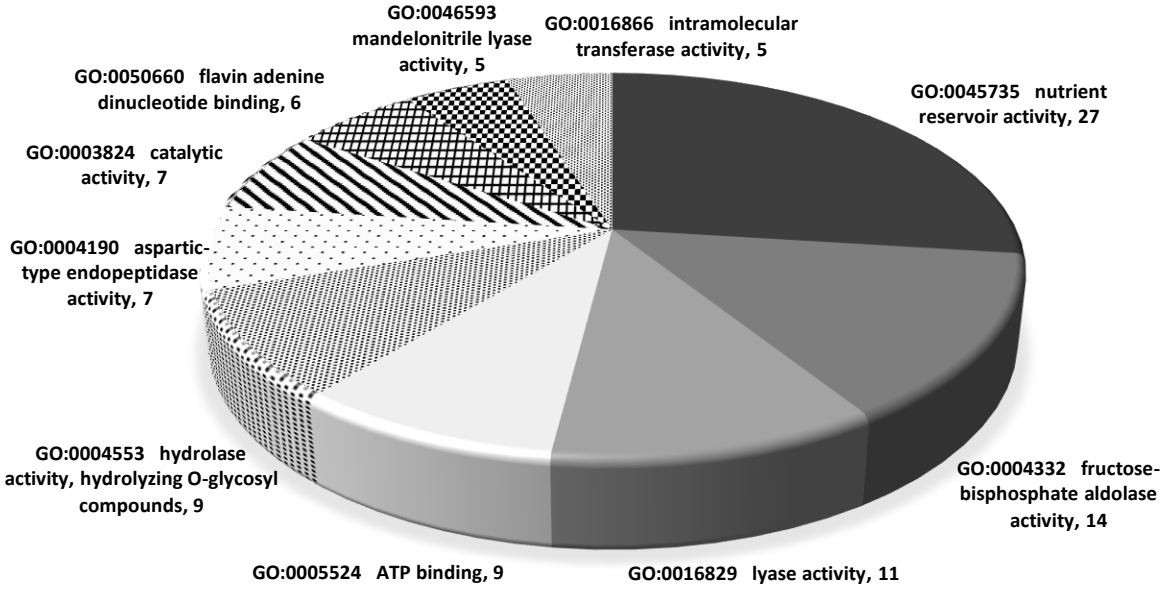
544 **Figure 4**

545

546

547

A


GO:0046593 mandelonitrile lyase activity, 5
GO:0016866 intramolecular transferase activity, 5
GO:0050660 flavin adenine dinucleotide binding, 6
GO:0045735 nutrient reservoir activity, 27
GO:0003824 catalytic activity, 7
GO:0004190 aspartic-type endopeptidase activity, 7
GO:0004553 hydrolase activity, hydrolyzing O-glycosyl compounds, 9
GO:0004332 fructose-bisphosphate aldolase activity, 14
GO:0005524 ATP binding, 9
GO:0016829 lyase activity, 11

B

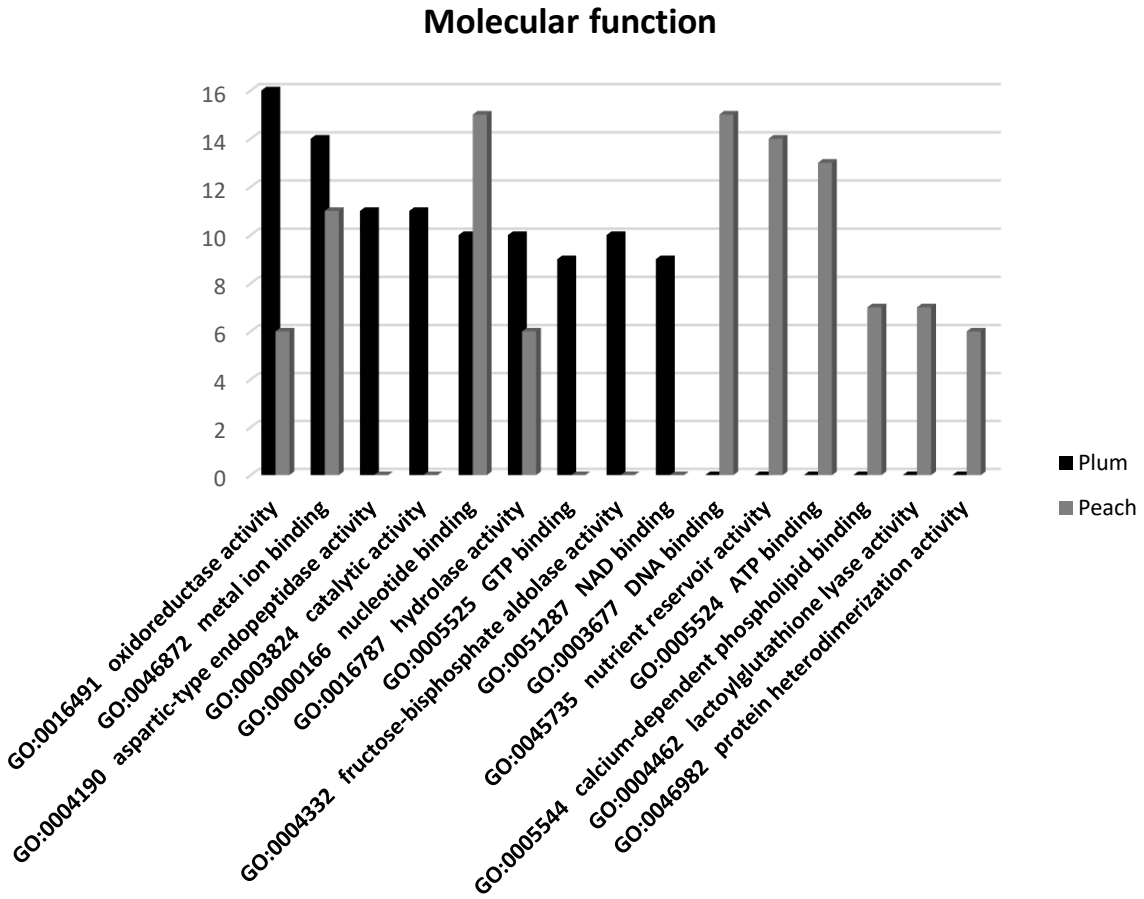

Molecular function
16
14
12
10
8
6
4
2
0
Plum
Peach
GO:0016491 oxidoreductase activity
GO:0046872 metal ion binding
GO:0004190 aspartic-type endopeptidase activity
GO:0003824 catalytic activity
GO:0000166 nucleotide binding
GO:0016787 hydrolase activity
GO:0005525 GTP binding
GO:0004332 fructose-bisphosphate aldolase activity
GO:0051287 NAD binding
GO:0003677 DNA binding
GO:0045735 nutrient reservoir activity
GO:0005524 ATP binding
GO:0005544 calcium-dependent phospholipid binding
GO:0004462 lactoylglutathione lyase activity
GO:0046982 protein heterodimerization activity

Figure 5