2-DE-based proteomic analysis of common bean (*Phaseolus vulgaris L.*) seeds

M. De La Fuente[a,*], A. Borrajo[b], J. Bermúdez[b], M. Lores[c], J. Alonso[d], M. López-Pedrouso[b], M. Santalla[a], A.M. De Ron[a], C. Zapata[b], G. Alvarez[b]

[a]Department of Plant Genetic Resources, Misión Biológica de Galicia-CSIC, Pontevedra, Spain

[b]Department of Genetics, Santiago de Compostela University, Spain

[c]Department of Analytic Chemistry, Nutrition and Bromatology, Santiago de Compostela University, Spain

[d]Proteomics Laboratory, CHUS, Santiago de Compostela, Spain

*corresponding autor: mfuente@mbg.cesga.es

Abstract:

Common bean (Phaseolus vulgaris L.) is the most important grain legume for direct human consumption. Proteomic studies in legumes have increased significantly in the last years but few studies have been performed to date in P. vulgaris. We report here a proteomic analysis of bean seeds by two-dimensional electrophoresis (2-DE). Three different protein extraction methods (TCA–acetone, phenol and the commercial clean-up kit) were used taking into account that the extractome can have a determinant impact on the level of quality of downstream protein separation and identification. To demonstrate the quality of the 2-DE analysis, a selection of 50 gel spots was used in protein identification by mass spectrometry (MALDI-TOF MS and MALDI-TOF/TOF). The results showed that a considerable proportion of spots (70%) were identified in spite of incomplete genome/protein databases for bean and other legume species. Most identified proteins corresponded to storage protein, carbohydrate metabolism, defense and stress response, including proteins highly abundant in the seed of P. vulgaris such as the phaseolin, the phytohemagglutinin and the lectinrelated α-amylase inhibitor.

Keywords: Phaseolus vulgaris, Seed proteome, Two-dimensional electrophoresis, Mass spectrometry, Protein extraction methods

1. Introduction

Common bean, Phaseolus vulgaris L., is one of the most ancient crops in the world and the most important source of proteins for many countries, fundamentally in America, Africa and Asia. Bean seed appears, moreover, to be a source of vitamins, dietetic fiber, and minerals, without unsaturated fatty acids. This is one of the reasons why bean seed intake is related to a decrease in the risk of suffering from heart stroke, colon cancer and gastrointestinal diseases [1,2]. Common bean was suggested to be a diploid model species within the legume family [3], since it displays particular characteristics with regard to its breeding (it is considered an extreme selfing species, about 95% of selfing on average), and its ability to fix atmospheric nitrogen through a symbiotic relationship with Rhizobium bacteria [4], apart from its specific model of evolution where multiple domestications took place in America from wild populations. In this respect, the genetic diversity in common bean is organized in two major geographic gene pools, the Mesoamerican and Andean gene pools. Independent domestication events in Mesoamerica and the Andes have led to two distinct cultivated gene pools which are the result of the evolution under both natural conditions and cultivation [5,6].

In recent years, large-scale profiling and identification of proteins based on two-dimensional electrophoresis (2-DE) and mass spectrometry (MS) have been carried out in legume species principally in the model legume Medicago truncatula [7–10]. In a less extent, proteomic studies have been also performed in other legume species such as Lotus japonicus [11], Pisum sativum [12,13],

Glycine max [14,15]and Pithecellobium dulce [16]. Nevertheless, to our knowledge no systematic study of the proteome of the common bean P. vulgaris

has yet been performed, although in this species proteomic technologies have been used for studying specific problems relative to host–pathogen molecular interactions and seed storage protein deficiency [17–20]. On the other hand, gene cloning and sequencing in the common bean have produced a number of complete coding sequences (CDSs) for genes involved in insect and disease resistance and important traits for common bean productivity. Also the number of expressed sequence tags (ESTs) available in EST collections has increased markedly in the recent years [3,21]. At this moment, two international projects for sequencing the whole genome of Phaseolus vulgaris are in progress: PhasIbeAm (Latin-American Science & Technology Development Programme) and the Bean CAP (USA). Synergy between proteome and genome information is expected to produce a more integrated knowledge for this species.

In this work, an approximation to the seed proteome of the common bean is performed by using two-dimensional gel electrophoresis (2-DE) to separate seed proteins from ICA Pijao genotype (domesticated genotype from Colombia belonging to the Mesoamerican gene pool).

2. Material and Methods

Three different protein extraction methods were used: trichloroacetic acid–acetone (TCA–acetone) [22], phenol [23] and thecommercial clean-up kit (GEHealthcare). Each extraction protocol was replicated four times from a tissue resulting from pooling three single seeds. Approximately 250 μg of the total protein was loaded on 24-cm Reading Strips IPGs (immobilized pH gradients, Bio-Rad Laboratories) with 4–7 pH linear gradients for the first-dimension separation (see Supplementary material). The SDS-polyacrylamide gel electrophoresis (SDS-PAGE) for the second-dimension separation was carried out in an Ettan Dalt six gel system(GE Healthcare, Uppsala, Sweden). Gels were stained with Sypro Ruby® stain (Bio-Rad Laboratories) and digitalized using the Fujifilm LAS-3000 Imager system. Subsequently, 2-DE gels were stained with silver stain for visual inspection. Optimization of the 2-DE run conditions was accomplished in order to obtain gels with the highest resolution, separation of the protein spots, the lowest background staining and the highest reproducibility.

3. Results and discussion

Fig. 1 shows representative 2-DE gels of seed proteins for each of the three extraction protocols. All three methods resulted in good quality, well-resolved gels, with reduced vertical and horizontal streaking and smearing. However the phenol method gave place to the clearest gel background, increased resolution of protein spots and less streaking on both horizontal and vertical dimensions as compared to the TCA–acetone and clean-up methods. Image analysis with PDQuest™ Advanced software v8.0.1 (Bio-Rad Laboratories) allowed us to generate a total number of 571 spots matched among gels. The spatial distribution of the matched spots along the 2-DE gel showed differences among

protocols. Thus, when gel images were split into four sections or quadrants (Q1–4) numbered 1 to 4 clockwise, starting from the top left corner, a higher number of spots were observed in quadrant Q2 for the phenol method (ANOVA test: $F=7.27$, $P=0.013$; Supplementary Fig. S1 and Supplementary Table S1). Therefore, the phenol method captured more proteins with less acidic pI and higher Mr than TCA–acetone and clean-up methods. This observation adds to the evidence suggesting that the phenol method increases the number of high Mr protein spots for the more neutral and basic gel region from protein extracts of cotton, apple, tomato and the euhalophyte S. europaea, as compared to TCA–acetone [24–26]. However, Saravanan and Rose [23] did not detect a consistent differential pattern for pI or Mr between TCA–acetone and phenol methods.

A multivariate analysis for normalized volume of the matched spots was performed by means of a principal component analysis (PCA). Fig. 2A shows the score plot for the principal components (PCs) one and two, which explained 27.4% and 15.0% of the variation, respectively. The first principal component clearly separates phenol samples from the other two samples, while TCA–acetone and clean-up samples are separated by the second principal component. Fig. 2B shows the loading plot for PC1 and PC2 where those variables (spots) responsible for the separation observed in the score plot can be seen. A large number of spots presented high positive loadings for PC1 which indicated that all these spots are highly correlated and they are involved in the separation of the phenol samples from the other two extraction protocols. An ordinary statistical analysis was also performed in order to detect quantitative differences for normalized volume of matched spots between extraction methods. This analysis revealed that 322 out of 571 spots (56.4%) presented a fold change higher than

2 between protocols being these differences statistically significant by the Mann–Whitney test ($P<0.05$), while the remaining 249 spots (43.6%) were shared by all three extraction protocols (Supplementary Tables S2 and S3, and Supplementary Fig. S2). There were 150 spots unique in phenol which presented a significant enhanced volume in this protocol as compared to 25 unique spots in TCA–acetone and39 in clean-up. For any protocol, the great majority of unique spots were located in quadrant Q2 corresponding to proteins of higher pI and Mr. Overall, our results suggest the phenol method as the first choice for protein extraction in bean seed proteomic approaches because it gives rise to high-quality 2-DE gels, the largest amount of spots for less acidic proteins of high molecular weight and the largest number of spots with increased volume. However, TCA–acetone and clean-up extractions could add enriched proteome coverage.

The compatibility of gel-based bean seed proteomic analyses with mass spectrometry (MS) was tested from the analyses of 50 excised gel spots differentially or not differentially extracted by the three protein extraction protocols assayed (Supplementary Fig. S3). A combination of matrix-assisted laser-desorptionionization time-of-flight mass spectrometry (MALDI-TOF MS) and MALDI-TOF/TOF was used in protein identification of tryptic digested spots (see Supplementary material). MS data were obtained in an automated analysis loop using 4800 MALDI-TOF/ TOF analyzer (Applied Biosystems). Peptide mass fingerprinting (PMF) and peptide fragmentation spectra data of each sample were combined through the GPS Explorer Software v3.6 (Applied Biosystems) using Mascot software v2.1. (Matrix Science) to search against a non-identical protein database (NCBInr release data 20100526). Our study shows that separation of

TCA–acetone, phenol and clean-up extractomes by 2-DE was efficient in achieving optimal identification of bean seed proteins (Table 1 and Supplementary Table S4). We were able to identify up to 70% of 50 excised gel spots. From the 35 resolved spots, 40 different proteins were successfully identified because four spots (26, 66, 83 and 486) yielded more than one confident identification. Most of the identified proteins corresponded to P. vulgaris (67.5%) or other legume species (15%): Phaseolus acutifolius, L. japonicus, P. sativum, Daucus carota and Vigna unguiculata. This efficiency of identification is within the range of efficiency reported in previous proteomic studies carried out in seeds of a variety of plant species including the model plant Arabidopsis thaliana (Supplementary Table S5), in spite of the limited number of available sequences in databases when bean genome sequencing projects are still in progress. Several factors contributed to increase the number of identifications. First, a large number of excised gel spots (40%) were identified as being phaseolins (PHS), the lectin phytohemagglutinin (PHA) or the lectin-related α-amylase inhibitor (αAI), well characterized bean proteins that make up to about 70% of total bean seed protein [20]. Second, information on genomes of other plant species was also very useful for the identification of highly conserved metabolism enzymes. Finally, most of the identifications were made using high-quality spectra derived fromMSMSrather than through PMFof the tryptic digested spots.

Anumber of protein spots exhibited remarkable discrepancies between the theoretical and experimental values inMr and/or pI. In addition, different spots were identified as the same protein. These discrepancies in Mr and/or pI could correspond to isoforms generated by a variety of causes, including post-

translational modifications, alternative splicing and the occurrence of multigene families. The following two examples show, however, that proteolysis can be a major factor at work here. First, PHS is encoded by a small gene family of 6–10 tightly linked sequences that consists of two subfamilies (α and β) encoding polypeptide size classes α and β. These polypeptide size classes were previously resolved by 2-DE within a range in Mr from 52.0 to 45.0 kDa and or pIs near pH 5 [27,28], whereas the fifteen phaseolins spots (62, 66, 81, 83, 105, 110, 111, 224, 296, 311, 354, 355, 356, 368, 496) identified in our study ranged in Mr from51.1 to 17.7 and in pI from 6.5 to 4.6. It has been reported that no substantial changes in the molecular structure of PHS are detectable before 7–10 days following seed germination when PHSis thendegraded by proteolytic enzymes into discrete clusters of fragments of 27–23 kDa [29]. In contrast, our study suggests the occurrence of partial proteolysis of the PHS prior to seed germination. Second, αAI is considered to be a truncated form of PHA that is synthesized as a precursor polypeptide of approximately 40 kDa and then post-translationally processed to several polypeptides in the Mr range from 15 to 18 kDa [30,31].

Accordingly, we identified αAI spots with an experimental Mr of approximately 27.2 and 15.1 kDa, which might correspond to polypeptide fragments arising from successive proteolysis of the precursor polypeptide. Overall, the great reproducibility of spots over replicates and extraction protocols together with the fact that phenol-based method minimizes the proteolysis during extraction,make in vitro proteolysis less likely.

The proteins identified by MS were classified in different groups corresponding to their presumed biological function: storage protein, carbohydrate metabolism,

defense, stress response, detoxification, growth and development, protein transport and nitrogen metabolism (Table 1; Supplementary Fig. S4).

Most identified protein spots corresponded to storage protein (37.5%), carbohydrate metabolism (22.5%), defense (17.5%) and stress response (10.0%). Nevertheless, each extraction method provided a differentiated profile of the bean seed proteome when identified proteins were grouped into functional categories: the TCA–acetone and clean-up methods extracted the highest amount of storage and defense proteins whereas the phenol method extracted the highest amount of carbohydrate metabolism proteins. In addition, remarkable differences were also detected among protocols when identified proteins were classified in glycosylated (phaseolins and lectins family) vs. the remainder non-glycosylated proteins (Supplementary Fig. S4).

Thus, the number of glycosylated/non-glycosylated protein spots showing significant enhanced extraction was 15/3, 1/8 and 13/1 for TCA–acetone, phenol and clean-up, respectively. It follows that TCA–acetone and clean-up gave enriched seed extracts of glycoproteins as compared to phenol (Fisher's exact test, $P<0.001$). On the contrary, previous studies with tomato and some species of fruits have shown that the phenol-based method yielded greater number of glycoproteins than TCA-based methods [23,32]. However, our results must be cautiously interpreted because pI and Mr experimental values were generally lower for glycosylated than for non-glycosylated proteins (Mann–Whitney test, $P<0.05$ and $P<0.001$ for pI and Mr, respectively) and the aforementioned observations suggest that the phenol method was less efficient to extract acidic proteins of low molecular weight. Therefore, differences in pI and Mr among glycosylated and non-glycosylated proteins might be underlying glycosylation-

dependent extraction patterns over protocols. In addition, many identified glycoproteins (64%) were actually degradation products of PHS with unknown levels of glycosylation. In this regard, it has been reported that PHS has three different N-linked oligosaccharide side chains attached to Asn amino acid residues: two glycosylation sites at position 252 and one at position 341, the numbering startingwith the N-terminalmethionine of the signal peptide [28]. But MSMS analyses revealed that most (64.3%) PHS fragments (spots 105, 224, 296, 311, 354, 355, 356, 368, and 496) identified in our study corresponded to the C-terminal sequence (positions from 257 to 421) that does not contain two canonical glycosylation sites. Therefore, the differential extraction patterns of bean seed glycoproteins detected among protocols might be more apparent than real.

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.jprot.2010.10.004.

Acknowledgments

The research was supported by AGL2008-02091/AGR project from the Spanish Government and EU-FEDER Funds. The authors are grateful to the Diputacion de Pontevedra (Spain) for providing farm facilities. M. De La Fuente and José Bermúdez are grateful to the Xunta de Galicia for awarding them fellowship grants, which allowed them to carry out this study.

References

[1] Celleno L, Tolaini MV, D'Amore A, Perricone NV, Preuss HG. A Dietary supplement containing standardized Phaseolus vulgaris extract influences body composition of overweight men and women. Int J Med Sci 2007;4:45–52.

[2] Helmstädter A. Beans and diabetes: Phaseolus vulgaris preparations as antihyperglycemic agents. J Med Food 2010;13:251–4.

[3] McClean PE, Lavin M, Gepts P, Jackson SA. Phaseolus vulgaris: a diploid model for soybean. In: Stacey Gary, editor. Genetics and genomics of soybean. Plant genetics/GenomicsSpringer; 2008.

[4] Sessitsch A, Hwieson JG, Perret X, Antoun H, Martinez-Romero E. Advances in Rhizobium research. Cr Rev Plant Sci 2002;21:323–78.

[5] Chacón MI, Pickersgill SB, Debouck DG. Domestication patterns in common bean (Phaseolus vulgaris L.) and the origin of the Mesoamerican and Andean cultivated races. Theor Appl Genet 2005;110:432–44.

[6] Kwak M, Gepts P. Structure of genetic diversity in the two major gene pools of common bean (Phaseolus vulgaris L., Fabaceae). Theor Appl Genet 2009;118:979–92.

[7] Gallardo K, Le Signor C, Vandekerckhove J, Thompson RD, Burstin J. Proteomics of Medicago truncatula seed development establishes the time frame of diverse metabolic processes related to reserve accumulation. Plant Physiol 2003;133: 664–82.

[8] Gallardo K, Finhaber C, Zuber H, Hérichert D, Belghazi M, Henry C, et al. A combined proteome and transcriptome analysis of developing Medicago truncatula. Mol Cell Proteomics 2007;6: 2165–79.

[9] Jorrín JV, Rubiales D, Dumas-Gaudot E, Recorbet G, Maldonado AM, Castillejo MA, et al. Proteomics: a promising approach to study biotic interaction in legumes. A review. Euphytica 2006;147:37–47.

[10] Castillejo MA, Maldonado AM, Dumas-Gaudot E, Fernández-Aparicio M, Susín R, Diego R, et al. Differential expression proteomics to investigate responses and resistance to Orobanche crenata in Medicago truncatula. BMC Genomics 2009;10:294.

[11] Dam S, Laursen BS, Ørnfelt JH, Jochimsen B, Staerfeldt HH, Friss C, et al. The proteome of seed development in the model legume Lotus japonicus. Plant Physiol 2009;149:1325–39.

[12] Schiltz S, Gallardo K, Huart M, Negroni L, Sommerer N, Burstin J. Proteome reference maps of vegetative tissues in pea. An investigation of nitrogen mobilization from leaves during seed filling. Plant Physiol 2004;35:2241–60.

[13] Curto M, Camafeita E, López JA, Maldonado AM, Rubiales D, Jorrín JV. A proteomic approach to study pea (Pisum sativum) responses to powdery mildew (Erysiphe pisi). Proteomics 2006;6:S163–74.

[14] Hajduch M, Ganapathy A, Stein JW, Thelen JJ. A systematic proteomic study of seed filling in soybean: establishment of high-resolution two-dimensional reference maps, expression profiles, and an interactive proteome database. Plant Physiol 2005;137:1397–419.

292 [15] Cheng L, Gao X, Li S, Shi M, Javeed H, Jing X, et al. Proteomic analysis of
293 soybean [Glycine max (L.) Meer.] seeds during imbibition at chilling temperature.
294 Mol Breed 2010;26:1–17.

295 [16] Sawasdipuksa N, Sumner LLW, Leib Z, Sangvanich P. Proteome analysis of
296 Pithecellobium dulce seeds using two-dimensional gel electrophoresis and
297 tandem mass spectrometry. J Sci Food Agric 2009;89:1284–91.

298 [17] Brizard JP, Carapito C, Delalande F, Van Dorsselaer A, Brugidou C.
299 Proteome analysis of plant–virus interactome: comprehensive data of virus
300 multiplication inside their hosts. Mol Cell Proteomics 2006;5:2279–97.

301 [18] Bernal C, Galindo I, Pérez D, Diez N. Aplicación de la proteómica
302 comparativa para la identificación de proteínas en Phaseolus vulgaris asociadas
303 a resistencia a plagas. Agronomía Trop 2006;56:555–9.

304 [19] Lee J, Feng J, Campbell KB, Scheffler BE, Garret WM, Thibivilliers S, et al.
305 Quantitative proteomic analysis of bean plants infected by a virulent and avirulent
306 obligate rust fungus. Mol Cell Proteomics 2009;8:19–31.

307 [20] Marsolais F, Pajak A, YinF,TaylorM,Gabriel M,MerinoDM, et al. Proteomic
308 analysis of common bean seed with storage protein deficiency reveals up-
309 regulation of sulfur-rich proteins and starch and raffinose metabolic enzymes and
310 down-regulation of the secretory pathway. J Proteomics 2010;73:1587–600.

311 [21] Ramírez M, Graham MA, Blanco-López L, Silvente S, Medrano-Soto A, Blair
312 MW, et al. Sequencing and analysis of common bean ESTs. Building a foundation
313 for functional genomics. Plant Physiol 2005;137:1211–27.

[22] Santoni V, Bellini C, Caboche M. Use of 2-dimensional protein pattern analysis for the characterization of Arabidopsis thaliana mutants. Planta 1994;192:557–66.

[23] Saravanan RS, Rose JKC. A critical evaluation of sample extraction techniques for enhanced proteomic analysis of recalcitrant plant tissues. Proteomics 2004;4:2522–32.

[24] Carpentier SC,Witters E, Laukens K,Deckers P, Swennen R, Panis B. Preparation of protein extracts fromrecalcitrant plant tissues: an evaluation of different methods for two-dimensional gel electrophoresis analysis. Proteomics 2005;5:2497–507.

[25] Wang XC, Li XF, Deng X, Han HP, Shi WL, Li YX. A protein extraction method compatible with proteomic analysis for the euhalophyte Salicornia europaea. Electrophoresis 2007;28: 3976–87.

[26] Xie C,Wang D, Yang X. Protein extraction methods compatible with proteomic analysis for the cotton seedling. Crop Sci 2009;49:395–402.

[27] Slightom JL, Drong RF, Klassy RC, Hoffman LM. Nucleotide sequences from phaseolin cDNA clones: the major storage proteins from Phaseolus vulgaris are encoded by two unique gene families. Nucleic Acids Res 1985;13:6483–98.

[28] Sturm A, Van Kuik JA, Vliegenthart JFG, Chrispeels MJ. Structure, position, and biosynthesis of the high mannose and the complex oligosaccharide side chains of the bean storage protein phaseolin. J Bio Chem 1987;262:13392–403.

335 [29] BoylanMT, Sussex IM. Purification of an endonuclease involved with storage-
336 protein degradation in Phaseolus vulgaris L. cotyledons. Planta 1987;170:343–
337 52.

338 [30] Ceriotti A, Vitale A, Bollini R. Lectin-like proteins accumulate as fragmentation
339 products in bean seed protein bodies. FEBS Lett 1989;250:157–60.

340 [31] Nasi A, Picariello G, Ferranti P. Proteomic approaches to study structure,
341 functions and toxicity of legume seeds lectins. Perspectives for the assessment
342 of food quality and safety. J Proteomics 2009;72:527–38.

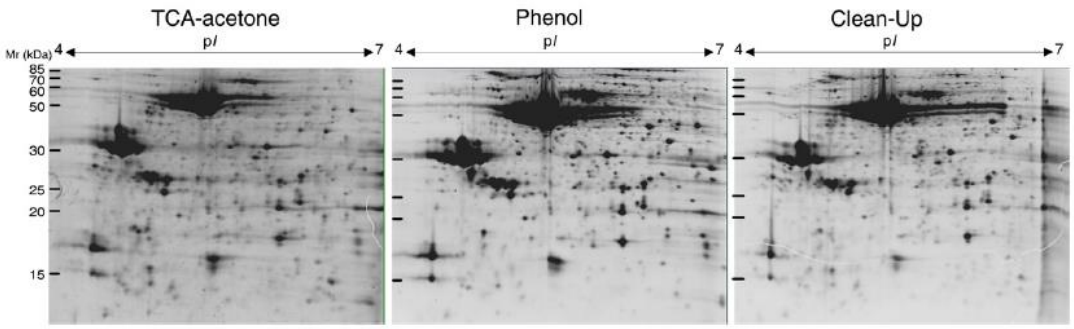
343 [32] Zheng Q, Song J, Doncaster K, Rowland E, Byers DM. Qualitative and
344 quantitative evaluation of protein extraction protocols for apple and strawberry
345 fruit suitable for two-dimensional electrophoresis and mass spectrometry
346 analysis. J Agric Food Chem 2007;55:1663–73.

347

CAPTIONS TO FIGURE

Figure 1. Representative 2-DE gels of proteins from bean seed extracted by three different protocols: TCA–acetone, phenol and clean-up. Proteins were separated on a linear pH 4–7 gradient in the first dimension and visualized using Sypro Ruby® staining. The approximate positions of molecular mass markers (kDa) are indicated on the left of the figure.

Figure 2. Principal component analysis for normalized volume of spots. A) Score plot for principal components (PCs) one and two. B) Loading plot for PC1 and PC2 (loadings were computed only for those spots presenting statistically significant differences between protein extraction methods by the Mann–Whitney tests).



360

361 **Figure 1.**

362

363

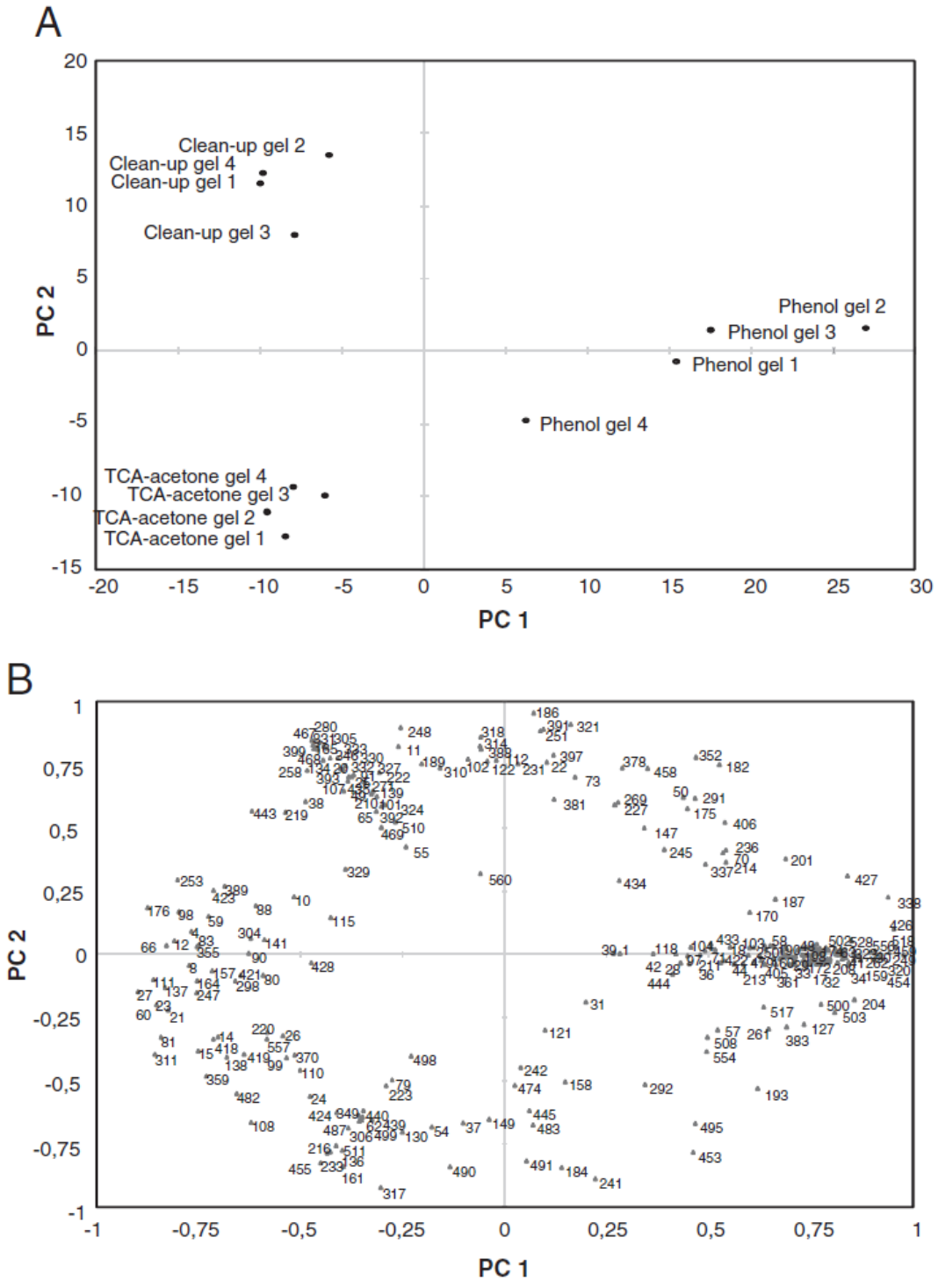

A
Clean-up gel 2
Clean-up gel 4
Clean-up gel 1
Clean-up gel 3
Phenol gel 2
Phenol gel 3
Phenol gel 1
Phenol gel 4
TCA-acetone gel 4
TCA-acetone gel 3
TCA-acetone gel 2
TCA-acetone gel 1
PC 2
PC 1
B
PC 2
PC 1

Figure 2.

Table 1. List of 2-DE-based identified P.vulgaris seed proteins by MALDI-TOF and MALDI-TOF/TOF

Spot no.	Protein identity	Mascot score	Match/ % Cov.	Mr (kDa)		pI	
				Theo.	Exp.	Theo.	Exp.
Storage protein							
62	Phaseolin, α-type (*P. vulgaris*)	563	13/33	49.2	18.3	5.3	4.9
66	Phaseolin — (*P. vulgaris*)	174	9/23	48.4	26.4	5.4	4.6
81	Phaseolin, α type — (*P. vulgaris*)	556	20/60	49.2	39.4	5.3	5.0
83	Phaseolin — (*P. vulgaris*)	251	15/43	47.5	51.1	5.0	4.9
105	Phaseolin — (*P. vulgaris*)	60	5/14	47.5	19.6	5.4	5.3
110	Phaseolin, chain A — (*P. vulgaris*)	97	9/32	45.0	17.7	5.2	5.0
111	Phaseolin, chain A — (*P. vulgaris*)	420	12/30	45.0	18.5	5.2	5.0
224	Phaseolin — (*P. vulgaris*)	388	12/36	48.4	20.7	5.3	5.7
296	Phaseolin — (*P. vulgaris*)	444	12/29	47.5	25.3	5.4	6.1
311	Phaseolin — (*P. vulgaris*)	335	12/34	47.5	30.4	5.4	5.9
354	Phaseolin , α-type — (*P. vulgaris*)	391	16/36	49.2	21.1	5.3	6.3
355	Phaseolin, α-type — (*P. vulgaris*)	302	12/34	49.2	23.8	5.3	6.3
356	Phaseolin — (*P. vulgaris*)	509	12/34	47.5	19.9	5.4	6.1
368	Phaseolin, α-type — (*P. vulgaris*)	348	15/31	49.2	26.0	5.3	6.1
496	Phaseolin — (*P. vulgaris*)	69	5/14	48.8	25.6	5.4	6.5
Carbohydrate metabolism							
121	α-1,4 glucan phosphorylase L isozyyme, chloroplastic/amyloplastic — (*Triticum aestivum*)	81	6/13	52.1	29.7	5.0	5.3
183	Glucose and ribitol dehydrogenase — (*Daucus carota*)	64	3/8	31.7	18.6	6.3	5.6
325	Malate dehydrogenase [NADP], chloroplastic– (*Zea mays*)	162	16/44	35.9	39.8	5.8	6.0
404	Ribulose 1,5-bisphosphate carboxylase-oxigenase, large subunit — (*Callaeum septentrionale*)	754	25/53	52.1	64.1	6.1	6.2
460	Ribulose 1,5-bisphosphate carboxylase-oxigenase, large subunit — (*Piptanthus nepalensis*)	136	9/24	50.8	60.5	6.3	6.3
472	Ribulose 1,5-bisphosphate carboxylase-oxigenase, large subunit — (*Dipogon lignosus*)	676	22/53	49.9	63.4	6.3	6.1
426	Granule-bound starch synthase I — (*P. vulgaris*)	316	5/11	67.6	18.6	6.4	6.5
427	Granule-bound starch synthase I — (*P. vulgaris*)	414	9/19	67.6	18.0	6.4	6.5
457	Phosphoglycerate kinase — (*Arabidopsis thaliana*)	63	4/13	42.2	43.5	5.5	6.3
Defense							
5	α-Amylase inhibitor β subunit — (*P. vulgaris*)	359	6/69	15.4	15.1	4.7	4.4
26	α-Amylase inhibitor β subunit — (*P. vulgaris*)	134	6/68	15.4	27.2	4.7	4.4
83	α-Amylase inhibitor like protein — (*P. vulgaris*)	95	3/17	28.9	51.1	4.95	4.9
26	Phytohemagglutinin — (*P. vulgaris*)	443	11/53	29.7	27.2	4.8	4.4
66	Phytohemaggutinin — (*P. vulgaris*)	104	9/43	29.8	26.4	5.0	4.6
27	Lectin — (*P. vulgaris*)	462	8/42	29.6	26.8	4.8	4.6
66	Lectin — (*P. vulgaris*)	100	4/20	29.6	26.4	4.8	4.6
Stress response							
360	Heat-shock cognate — (*D. carota*)	136	4/17	18.5	18.2	10.6	6.1
523	Pv42p — (*P. vulgaris*)	105	6/21	41.6	43.1	6.5	6.5
486	Superoxide dismutase [Mn] — (*Pisum sativum*)	98	4/20	26.7	25.6	7.2	5.7
554	Dehydrin — (*Vigna unguiculata*)	92	1/10	26.5	27.3	6.0	6.9
Detoxification							
43	Putative glutathione S-transferase — (*Phaseolus acutifolius*)	86	2/13	24.8	38.0	5.6	6.0
486	NAD-dependent formate dehydrogenase — (*Oryza sativa*)	87	4/10	41.4	25.6	6.9	6.5
Growth and development							
441	IAA-protein conjugate — (*P. vulgaris*)	413	23/53	35.5	36.4	6.2	6.4
Protein transport							
500	GTP-binding nuclear protein Ran1A — (*Lotus japonicus*)	260	10/52	24.1	27.4	6.7	6.7
Nitrogen metabolism							
524	Glutamine synthetase β2, cytosolic — (*P. vulgaris*)	57	3/42	39.3	41.1	6.1	6.4