Available online at www.sciencedirect.com
ScienceDirect
Phytochemistry 68 (2007) 997–1007
PHYTOCHEMISTRY
www.elsevier.com/locate/phytochem

Combined 2D electrophoretic approaches for the study of white lupin mature seed storage proteome

Chiara Magni [a], Alessio Scarafoni [a], Anita Herndl [a,1], Fabio Sessa [a], Bhakti Prinsi [b], Luca Espen [b], Marcello Duranti [a,*]

[a] *Department of AgriFood Molecular Sciences, State University of Milan, Via Celoria 2, 20133 Milano, Italy*
[b] *Department of Crop Production, State University of Milan, Via Celoria 2, 20133 Milano, Italy*

Received 28 November 2006; received in revised form 8 January 2007
Available online 23 February 2007

Abstract

Seed proteome analysis by 2D IEF/SDS–PAGE techniques is challenging for the intrinsic difficulties related to quantitative disparity of the seed proteins, i.e. storage and non-storage proteins, their polymorphic nature, the extensive post-translational modifications and the paucity of deposited primary structures available. Conversely, 2D maps of seed proteomes can be extremely useful for a number of fundamental and applied investigations.

In this work, we have used a combination of two experimental approaches to identify the main protein components of an emerging protein-rich legume seed, that is white lupin seed (*Lupinus albus*, L.). One is the canonical proteomic approach including 2D electrophoretic separation and mass spectrometry of selected trypsin-digested polypeptides; the other approach is a group comparative 2D electrophoretic analysis of cotyledonary protein families. To this second purpose, the three main families of lupin seed proteins, namely α-conglutins, the 11S globulin fraction, β-conglutins, the 7S globulin fraction, and γ-conglutin, a basic 7S protein, were isolated by conventional biochemical techniques and their 2D reference maps were compared with the total protein map.

With the first approach 37 out of 40 spots, making up about 35% of total spot volumes in the 2D map, were found to belong to the main seed protein families. Thanks to cDNA-deduced lupin storage protein sequences, determined on purpose and deposited, most of the identification statistical parameters were very good. Moreover, it was possible to identify several endogenously proteolysed subunits in the map. The second comparative approach, beside confirming these attributions, allowed to allocate 124 polypeptides within the three main lupin protein families.

These two approaches proved to be mutually validating and their combined use was effective for the establishment of a seed proteome map even in the case of sequence and protein post-translational processing lack of information. The results obtained also extend our knowledge of the seed storage protein polymorphism of white lupin.

Keywords: Liquid chromatography/electrospray ionisation tandem mass spectrometry (LC/ESI-MS/MS); *Lupinus albus*; Legume seed; Proteome analysis; Seed storage proteins; 2D Electrophoresis

1. Introduction

The quali-quantitative 2D electrophoretic analysis of economically relevant grain seeds can have a number of fundamental and applied uses. Proteomic analysis can be useful for studies on storage protein deposition in seeds, evaluation of the effects of growing conditions, technological treatments and other factors, identification of species, varieties and cultivars, monitoring of desirable or undesirable specific protein components, and assessment of many other parameters affecting the overall seed (protein) quality. However, these approaches are complicated by the peculiarity of the seed as a biological entity: indeed, most

* Corresponding author. Tel.: +39 02 50316817; fax: +39 02 50316801.
E-mail address: marcello.duranti@unimi.it (M. Duranti).
[1] On leave of absence from: Institut für Angewandte Mikrobiologie, Universität für Bodenkultur, Muthgasse 18, A-1190 Wien, Austria.

0031-9422/$ - see front matter © 2007 Elsevier Ltd. All rights reserved.
doi:10.1016/j.phytochem.2007.01.003

seeds, especially those of the Leguminosae, are extremely rich in proteins, most of which belong to the family of storage proteins. This means that all other metabolic, structural, regulatory and defence proteins are by far less represented. As a consequence, most, if not all, polypeptides identified from a seed protein extract by 2D electrophoretic techniques belong to the storage protein families.

The canonical proteomic approaches are further complicated by the presence of several isoforms of the same protein as the products of a multigene family and the occurrence of extensive post-translational modifications in the cotiledonary tissues. Both phenomena are typical of the seed storage proteins, resulting in extensive polypeptide heterogeneity. Indeed, while the dynamics of physiopathologically modified protein profiles, with emphasis to specific key proteins, have extensively been studied in many species and under differing seed ontogenesis time frames by 2D electrophoretic techniques (Rajjou et al., 2006), the systematic analysis of crop seed storage proteomes is less frequent. This lack can also be attributed to the insufficient sequence information available in plant genomic and ESTs databases. As a matter of facts, 2D maps are available only for seeds of model species, including *Arabidopsis* (Gallardo et al., 2001, 2002) and *Medicago* (Gallardo et al., 2003), and few major crop plants, such as *Triticum* (Skilas et al., 2001; Islam et al., 2002; Majoul et al., 2003), *Hordeum* (Finnie et al., 2004; Østergaard et al., 2004), and *Glycine* (Hajduch et al., 2005), whose genomes have fully or extensively been characterised.

For this experimental work, we have selected lupin (*Lupinus albus*, L.) seeds, an emerging minor crop which, along with soybean, is one of the most rich in protein (up to 30–35% dry weight basis). This legume seed is an excellent candidate to fulfil the new needs of plant proteins for food, especially in Europe, as witnessed by various ongoing projects on the subject. The availability of lupin seed 2D maps, depicting the profile of seed storage proteome, can thus be of interest for the expected developments related to its use as a valuable protein source. However, from the analytical viewpoint, lupin seed proteins are characterised by an extreme polypeptide heterogeneity due to the storage protein multigenic origin and a very pronounced post-translational processing of the main protein components (Cerletti et al., 1978). The identification of the mature polypeptides in lupin seed 2D map is therefore crucial for a better understanding of the many aspects related to their presence and function in the seed.

Lupin seeds, as other leguminous seeds, contain two major types of storage globulins (Blagrove and Gillespie, 1975). The first group includes proteins with sedimentation coefficients around 11S. They are named α-conglutins and belong to the "legumin-like" or 11S globulin family. The mature protein consists of hexamers of disulphide linked subunits referred to as acidic subunits, with greater M_r, and basic subunits, with lower M_r. The acidic subunits of α-conglutin, contrary to most other 11S globulins, are glycosylated (Duranti et al., 1995). The second group of proteins has sedimentation coefficients around 7S. They are named β-conglutins and are also referred to as "vicilin-like" or 7S globulins. They are trimeric proteins in which the monomers consist of a number of polypeptides ranging from 16 to over 70 kDa, with no disulphide bridges linking them. The endogenous cleavage of this family of proteins is known since long, such as in pea vicilins (Duranti et al., 1987), but in lupin seed this phenomenon is greatly enhanced (Duranti et al., 1992). In addition to these main protein families, lupin seeds contain significant amounts (5–6%) of γ-conglutin, a peculiar basic 7S protein (Duranti et al., 1981), which consists of two heterogenous disulphide linked subunits of 17 and 30 kDa (Scarafoni et al., 2001), and a monomeric low M_r protein similar to other 2S albumins, named δ-conglutin (Duranti et al., 1981; Salmanowicz and Weder, 1997), made by two disulphide linked chains of about 4 and 9 kDa. The subunits of all these four mentioned classes of proteins arise from the proteolytic cleavage of precursor molecules (Derbyshire et al., 1976; Müntz, 1998; Müntz et al., 2002).

In this work, white lupin (*Lupinus albus*, L, var. Multitalia) seed storage proteins have been identified through a combination of canonical proteomic analysis, based also on newly characterised cDNA sequences, and group comparative analysis of reference maps generated with purified lupin main protein families.

2. Results

Total protein extracts from mature dry *Lupinus albus* seeds were used to generate 2D IEF/SDS–PAGE maps. The map shown in Fig. 1 is a representative one out of 8 maps, arising from 2 independent total protein extracts. The similarity among the different maps was extremely high, suggesting that no significant variation both in the raw materials nor in the separation procedures has occurred. The digital image processing of the map detected 357 spots. The map denotes the intrinsically complex pattern of lupin storage proteins with several spots of same M_rs and different pIs, suggesting charge heterogeneity of same or similar gene products, as it will be detailed below. A number of 40 spots in the 2D gel was excised, digested with trypsin and analysed by mass spectrometry, as detailed in Section 4. Table 1 shows the corresponding attributions and reports the experimental features of the 37 identified spots, which account to 35% of total spot volume. At least one member of each of the mature lupin seed protein families, *i.e.* α-, β-, γ- and δ-conglutins, was identified. The identified proteins are named according to the database entries. With only few exceptions, the identification procedure was very reliable, as estimated by the calculated statistical parameters (Table 1). This result was made possible thanks to nucleotide sequencing of lupin storage protein cDNAs carried out in this work. With this step, which is detailed in Section 4, novel cDNA-deduced storage protein sequences, belonging to

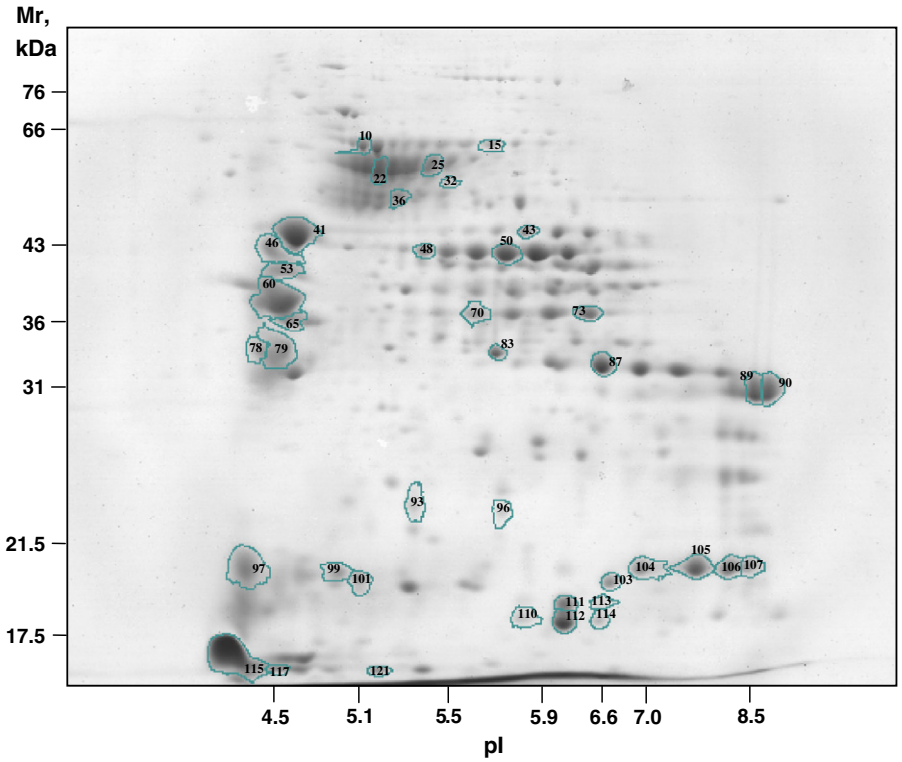

Fig. 1. Lupin seed total protein 2D reference map. The spots analysed by trypsin-digestion and mass spectrometry are outlined and numbered. The full list of the identified spots is reported in Table 1. Gels and running conditions are detailed under Section 4.

the α-(EMBL accession numbers: AJ938033, AJ938034 and AJ938035), β- (EMBL accession number: AJ966470) and δ-conglutins (EMBL accession numbers: AM156845 and AM150790), were determined and deposited in the EMBL database.

Nevertheless, some spots, such as those from number 103 to 107, were assigned through the identification of homologous gene products from another legume seeds, *i.e. Pisum sativum*, allowing only relatively low amino acid coverage values. However, the statistical values (*i.e. P*-value and score) for these peptides were highly significant. This finding suggests the presence of other lupin 11S globulin cDNAs not identified and sequenced as yet. This may also be the case of two other relevant spots, *i.e.* no. 41 and 53, which were not identified at all by mass spectrometry analyses of the corresponding tryptic fragments. On the other hand, it is intriguing that peptides arising from various spots in this region were found to fit to only one of the determined α-conglutin gene sequences, thus suggesting that the other genes may not be significantly expressed nor the protein products did accumulate to a significant extent.

As far as the 7S globulin family is concerned, the unique cDNA sequence determined and deposited in this work (EMBL accession number: AJ966470) did allow the best matching of 17 out of 22 β-conglutin spots, while the remaining 5 spots assigned where best fitted with a cDNA-deduced protein sequence already existing in the database (EMBL accession number: AAS97865).

In this work, the post-translational modifications of the storage proteins have also been taken in consideration. As mentioned in the Introduction, most seed proteins are synthesized as pro-proteins, which are proteolytically processed before or during the deposition within the storage vacuoles to give rise to the mature subunits. These events are responsible for the multi-subunit organisation of this kind of proteins (Müntz, 1998). When the cleavage sites of *L. albus* seed pro-proteins, or those of related species, were known (see references in Table 2), an *in silico* analysis made the identification of the corresponding mature protein subunits possible, by taking in consideration the position of the detected peptides along the precursor sequence. Table 2 shows the identification of the mature subunits in proteolytically processed α-, γ- and δ-conglutins. In the second column of the table, the complete deduced amino acid sequences of the three identified protein precursors are reported and the peptides experimentally determined by tandem mass spectrometry are evidenced in red colour. The experimental and theoretical masses and pIs were by far more consistent and the sequence coverages have significantly increased with respect to the figures of Table 1, where the whole pre-pro-protein sequences, as found in the database, were used. In most cases, fitting between measured and calculated masses and pIs is nearly perfect. For example, amino acid coverage of spot no. 111 changed from 25% (Table 1) to 73% (Table 2), when the γ-conglutin small subunit is taken in consideration. As

Table 1
List of identified spots by mass spectrometry and bioinformatic analysis

Spot	Vol.%	Protein identification	Accession number	*P*-value	Score	M_r[a]	*M*W[b]	pI[a]	pI[b]	A.a. cov.[c]%
10	0.44	β-Conglutin precursor (*Lupinus albus*)	AJ966470	4.7×10^{-10}	128	63,800	62,032	5.11	6.08	31.1
15	0.23	β-Conglutin precursor (*Lupinus albus*)	AAS97865	1.7×10^{-11}	140	64,400	62,130	5.67	6.43	33.8
22	0.64	β-Conglutin precursor (*Lupinus albus*)	AJ966470	6.4×10^{-11}	170	58,200	62,032	5.20	6.08	36.9
25	0.50	β-Conglutin precursor (*Lupinus albus*)	AJ966470	3.4×10^{-11}	168	59,400	62,032	5.42	6.08	37.9
32	0.11	β-Conglutin precursor (*Lupinus albus*)	AJ966470	6.4×10^{-9}	140	56,000	62,032	5.49	6.08	35.0
36	0.18	β-Conglutin precursor (*Lupinus albus*)	AJ966470	2.8×10^{-9}	168	52,900	62,032	5.30	6.08	37.5
41	3.06	No identification				45,400		4.67		
43	0.25	β-Conglutin precursor (*Lupinus albus*)	AAS97865	2.2×10^{-9}	100	46,000	62,130	5.80	6.43	25.1
46	1.30	No identification				44,200		4.50		
48	0.37	β-Conglutin precursor (*Lupinus albus*)	AJ966470	1.9×10^{-10}	180	44,000	62,032	5.40	6.08	41.1
50	1.25	β-Conglutin precursor (*Lupinus albus*)	AJ966470	6.5×10^{-11}	138	43,500	62,032	5.72	6.08	27.5
53	0.62	No identification				41,900		4.58		
60	3.73	α-Conglutin precursor (*Lupinus albus*)	AJ938034	3.3×10^{-7}	50	38,500	60,792	4.58	5.90	13.1
65	0.61	β-Conglutin precursor (*Lupinus albus*)	AJ966470	1.2×10^{-10}	88	37,000	62,032	4.62	6.08	22.0
		α-Conglutin precursor (*Lupinus albus*)	AJ938034	6.7×10^{-8}	30	37,000	60,792	4.62	5.90	10.9
70	0.47	β-Conglutin precursor (*Lupinus albus*)	AJ966470	1.6×10^{-9}	130	37,600	62,032	5.60	6.08	31.1
73	0.63	β-Conglutin precursor (*Lupinus albus*)	AJ966470	1.8×10^{-10}	160	37,600	62,032	6.46	6.08	35.4
78	0.59	β-Conglutin precursor (*Lupinus albus*)	AJ966470	2.6×10^{-9}	78	34,300	62,032	4.43	6.08	20.9
79	1.86	β-Conglutin precursor (*Lupinus albus*)	AJ966470	1.7×10^{-8}	58	34,500	62,032	4.56	6.08	14.1
83	0.42	Putative TAG-associated factor (*Lupinus angustifolius*)	AAN75426	3.5×10^{-4}	10	34,000	32,214	5.68	6.33	5.1
87	1.16	β-Conglutin precursor (*Lupinus albus*)	AAS97865	1.2×10^{-11}	110	32,900	62,130	6.59	6.43	24.3
89	1.28	γ-Conglutin precursor (*Lupinus albus*)	CAC16394	1.4×10^{-7}	60	30,800	49,219	8.65	8.40	18.1
90	1.49	γ-Conglutin precursor (*Lupinus albus*)	CAC16394	5.4×10^{-9}	60	30,900	49,219	8.85	8.40	23.0
		β-Conglutin precursor (*Lupinus albus*)	AAS97865 AJ966470	1.1×10^{-8}	70	30,900	62,130	8.85	6.43	17.8
93	0.30	α-Conglutin precursor (*Lupinus albus*)	AJ938034	2.6×10^{-8}	30	23,600	60,792	5.36	5.90	14.1
96	0.20	α-Conglutin precursor (*Lupinus albus*)	AJ938034	1.7×10^{-8}	40	23,200	60,792	5.71	5.90	14.4
97	2.17	β-Conglutin precursor (*Lupinus albus*)	AJ966470	1.2×10^{-8}	28	20,400	62,032	4.33	6.08	7.2
99	0.48	α-Conglutin precursor (*Lupinus albus*)	AJ938034	6.3×10^{-9}	20	20,200	60,792	4.96	5.90	10.7
		β-Conglutin precursor (*Lupinus albus*)	AAS97865	3.7×10^{-7}	50	20,200	62,130	4.96	6.43	8.6
101	0.37	Small HSP (*Retama raetam*)	AAL32036	1.3×10^{-6}	20	19,900	17,891	5.09	5.82	20.3
103	0.41	Legumin K (*Pisum sativum*)	S26688	1.5×10^{-5}	10	19,700	56,276	6.67	5.65	2.8
104	0.77	Legumin K (*Pisum sativum*)	S26688	6.9×10^{-5}	6	20,300	56,276	7.09	5.65	2.8
105	1.33	Legumin K (*Pisum sativum*)	S26688	1.5×10^{-7}	10	20,300	56,276	7.85	5.65	2.8
106	0.87	Legumin K (*Pisum sativum*)	S26688	6.5×10^{-6}	4	20,300	56,276	8.3	5.65	2.8
107	0.45	Legumin K (*Pisum sativum*)	S26688	9.8×10^{-6}	8	20,500	56,276	8.57	5.65	2.8

110	0.29	γ-Conglutin precursor (*Lupinus albus*)	CAC16394	1.1×10^{-8}	40	18,300	49,219	5.78	8.40	16.8
111	0.65	γ-Conglutin precursor (*Lupinus albus*)	CAC16394	7.1×10^{-14}	70	18,900	49,219	6.21	8.40	24.6
112	1.35	γ-Conglutin precursor (*Lupinus albus*)	CAC16394	1.2×10^{-7}	40	18,000	49,219	6.19	8.40	15.3
113	0.09	γ-Conglutin precursor (*Lupinus albus*)	CAC16394	2.7×10^{-10}	30	18,900	49,219	6.57	8.40	14.4
114	0.17	γ-Conglutin precursor (*Lupinus albus*)	CAC16394	6.0×10^{-14}	50	18,200	49,219	6.56	8.40	14.4
115	6.47	δ-Conglutin precursor (*Lupinus albus*)	CAJ42100 CAJ43922	1.8×10^{-9}	30	17,300	17,138	4.21	5.47	26.4
		β-Conglutin precursor (*Lupinus albus*)	AJ966470	6.0×10^{-8}	30	17,300	62,032	4.21	6.08	7.2
117	0.38	β-Conglutin precursor (*Lupinus albus*)	AJ966470	1.0×10^{-9}	68	16,500	62,032	4.56	6.08	17.7
121	0.21	β-Conglutin precursor (*Lupinus albus*)	AJ966470	6.2×10^{-10}	70	16,400	62,032	5.2	6.08	17.9

Volume of identified spots with respect to total detected spots: 34.80%.
[a] Experimental data.
[b] Theoretical data.
[c] Amino acid coverage.

far as the β-conglutins is concerned, the more extensive endogenous proteolytic cleavage of the respective precursors (Duranti et al., 1992) and the lack of knowledge of the specific cleavage sites prevented the identification of the constituent subunits.

In addition to the spot-by-spot mass spectrometry identification, a protein family identification was carried out with a different experimental set. In this alternative approach, group comparative analysis was carried out by using 2D maps generated with purified protein samples of each protein family. To this purpose, the Coomassie blue stained 2D gels of the three purified main lupin protein families, namely α-, β- and γ-conglutins, were used as a reference (Fig. 2, panels A–C). Altogether, the three protein families represented 83% of the globulin fraction and 76% of the total lupin seed proteins (data from Duranti et al., 1981). The computer-assisted matching and manual refining of each reference protein 2D gel with the total lupin protein map (Fig. 2, panels A′–C′) allowed the group attribution of a large number of spots to each protein family. Indeed, the two groups of acidic and basic α-conglutins, three groups of high, intermediate and low M_r β-conglutins, and the high and low M_r γ-conglutin subunits were localised (these groups of spots are outlined in the panels A′–C′). The spots allocated to the various protein families were 124, that is the 35% of the spots in the total protein map. When the spot volume (obtained multiplying area by intensity) was considered, the lupin polypeptides allocated to their corresponding family rose to 67%. This indicated that our identification did cover the quantitatively most relevant lupin seed protein components. Furthermore, this approach did allow the identification of those spots for which a low score (spots from 103 to 107) or even no identification at all (spots 43 and 51) were obtained with the canonical proteomic approach.

Interestingly, the attributions obtained by the group comparison of the standard proteins with the total protein extracts are always consistent with those determined by mass spectrometry.

3. Discussion

The application of 2D electrophoretic approaches to the identification of seed polypeptide components is challenging. In dry mature seeds, which are storage and metabolically dormant tissues, the overwhelming presence of storage proteins frequently hidden all other catalytic, structural, regulatory and defence proteins. Therefore, an effect of this wide range of protein relative abundancies is that the main spots visible in a 2D electrophoretic seed map belong almost exclusively to the storage protein family. One example of redundant identification of specific storage polypeptides is the work by Gianazza et al. (2003), which revealed a monotonous presence of same polypeptides in the soybean seed and protein isolate maps. In a more recent paper, Wait et al. (2005) attributed a number of spots,

mainly consisting of β-conglutin polypeptides, in 2D gels of lupin protein industrial preparations, with the aim of identifying biologically active polypeptides homologous to soybean hypocholesterolemic 7S globulin α′ subunits (Duranti et al., 2004). These papers emphasise the suitability of 2D gel proteomic approaches to the study of various fundamental and applied research aspects. However, the analysis of seed storage proteomes is still underperformed, especially for minor crops. As mentioned, this is mainly, but not exclusively, due to the lack of genomic and EST database in this area. For this reason we have undertaken a cloning and sequencing program in order to improve the previous spot identifications. The sequencing activity is still ongoing. Altogether, the new deposited sequences, along with the previous information on the proteolytic cleavage sites giving rise to the mature protein subunits, have allowed their identification in the lupin seed protein map with extremely good statistical figures. Nonetheless, in the case of some α-conglutin spots, the new gene sequences available still did not allow their identification, suggesting that further sequences are crucial to complete the 2D attribution work.

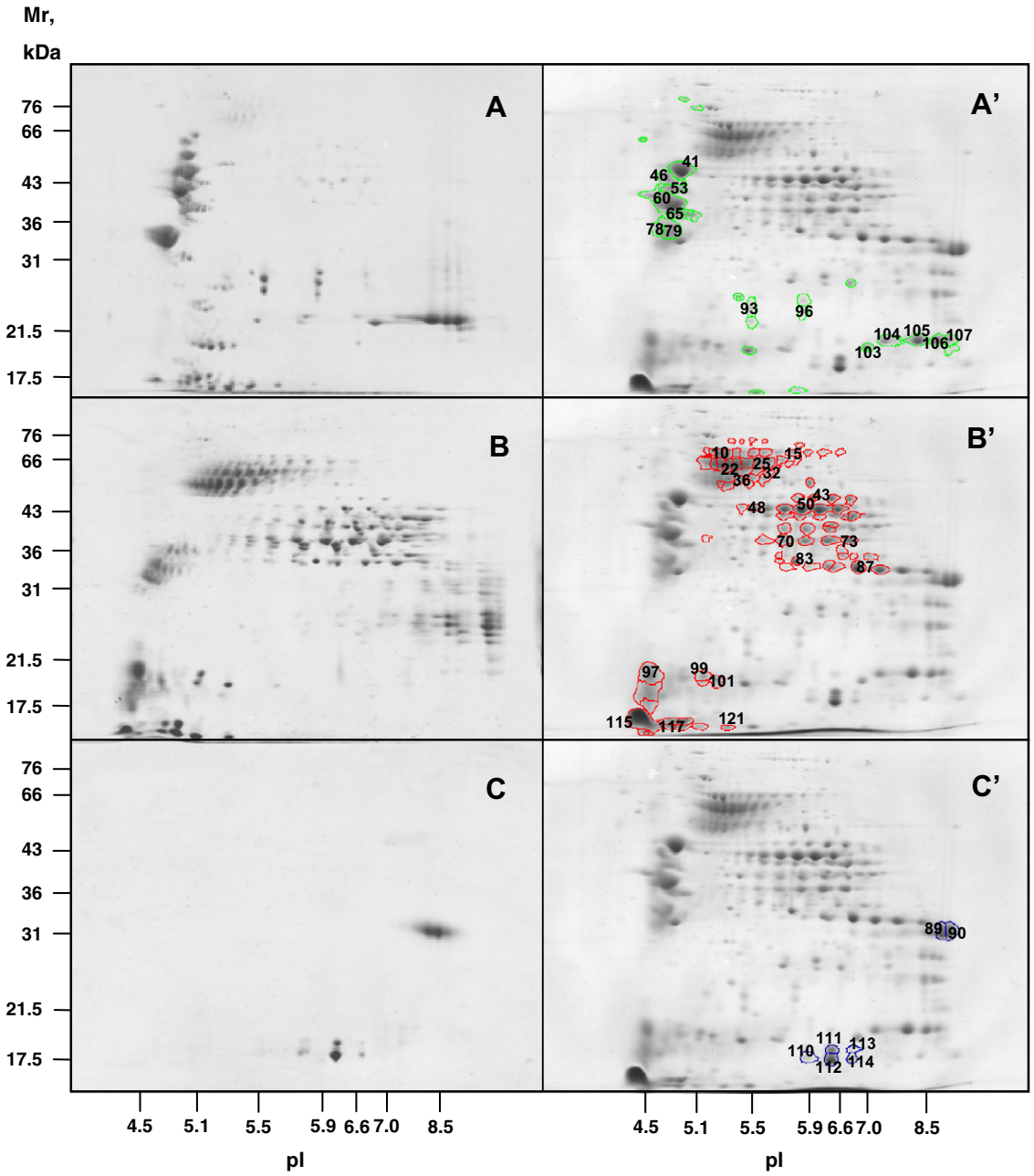

Fig. 2. Group comparative electrophoretic analysis of purified cotyledonary protein families. The three reference protein family maps, *i.e.* α-, β- and γ-conglutins, are shown in the panels A, B and C, respectively. A′–C′ indicate the same total protein 2D map where the spots allocated to α-, β- and γ-conglutin families are outlined, respectively. Numbers are those of the spots identified by mass spectrometry in Fig. 1.

Table 2
Identification of mature subunits in proteolytically processed proteins

Spot number	*Full length deduced aminoacid sequence*	*Subunit identification*	*MW*[a]	*MW*[b]	*pI*[a]	*pI*[b]	*Amino acid cov, %*
60	MSNKLLALSLFLPFLLLFFCGCFASTFR**QQPQENECQFQRLNALEPDNTVQSEAGTIETWNPK**NDEL RCAGVALSRCTIQRNGLR**RPFYTNAPQEIYIQQGR**GIFGMIFPGCGETYEEPQESEKGQGPRPQDRH QKVEHFKEGDIIAVPTGIPFWMYNDGQTPVVAITLIDTTNLDNQLDQIPR**RFYLSGNQEQEFLQYQE K**EGGQGQQQPEGGNVLSGFDDEFLEEALSVNKEIVRNIKGKNDDREGGIVEVKGGLKVIIPPTMRPRH GREEEEEEEEEDERRGDRRRRHPHHHHHEEEEEEEEEWSHQVRRVRRPHRHHHHRKDR**NG**LEETLCT MKLRHNIGESTSPDAYNPQAGRFKTLTSIDFPILGWLGLAAEHGSIYKNALFVPYYNVNANSILYVL NGSAWFQVVDCSGNAVFNGELNEGQVLTIPQNYAAAIKSLSDNFRYVAFKTNDIPQIATLAGANSEI SALPLEVVAHAFNLNRDQARQLKNNNPYKFLVPPPQSQLRAVAKGELP	α-Conglutin acidic subunit	38,500	35,461	4.58	5.34	23.1
65	MSNKLLALSLFLPFLLLFFCGCFASTFRQQPQENECQFQR**LNALEPDNTVQSEAGTIETWNPK**NDEL RCAGVALSRCTIQRNGLR**RPFYTNAPQEIYIQQGR**GIFGMIFPGCGETYEEPQESEKGQGPRPQDRH QKVEHFKEGDIIAVPTGIPFWMYNDGQTPVVAITLIDTTNLDNQLDQIPR**RFYLSGNQEQEFLQYQE K**EGGQGQQQPEGGNVLSGFDDEFLEEALSVNKEIVRNIKGKNDDREGGIVEVKGGLKVIIPPTMRPRH GREEEEEEEEEDERRGDRRRRHPHHHHHEEEEEEEEEWSHQVRRVRRPHRHHHHRKDR**NG**LEETLCT MKLRHNIGESTSPDAYNPQAGRFKTLTSIDFPILGWLGLAAEHGSIYKNALFVPYYNVNANSILYVL NGSAWFQVVDCSGNAVFNGELNEGQVLTIPQNYAAAIKSLSDNFRYVAFKTNDIPQIATLAGANSEI SALPLEVVAHAFNLNRDQARQLKNNNPYKFLVPPPQSQLRAVAKGELP	α-Conglutin acidic subunit	36,900	35,461	4.62	5.34	19.1
89	MAKNMAPILHILVISLSYSFLFVTSSSQNSQSLYHNSQPTSSSK**PNLLVLPIQQDASTKLHWGNILK** R**TPLMQVPVLLDLNGK**HLWVTCSQHYSSSTYQAPFCHSTQCSR**ANTHQCFTCTDSTTSR**PGCHNNTC GLISSNPVTQESGLGELAQDVLALHSTHGSKLGSLVK**IPQFLFSCAPTFLTQK**GLPNNVQGALGLGH APISLPNQLFSHFGLKRQFTMCLSSYPTSNGAILFGDINDPNNNNYIHNSLDVLHDMVYTPLTISK**Q GEYFIQVSAIR**VNKHMVIPTKNPSMFPSSSSS**SY**HESSEIGGAMITTNPYTVLRHSIFEVFTQVFA NNVPKQAQVKAVGPFGLCYDTKKISGGVPSVDLIMDKSDVVWRISGENLMVQAQDGVSCLGFVDGGV HTRAGIALGTHQLEENLVVFDLARSRVGFNTNSLKSHGKSCSNLFDLNNP	γ-Conglutin acidic subunit	30,800	29,209	8.65	8.73	31.1
90	MAKNMAPILHILVISLSYSFLFVTSSSQNSQSLYHNSQPTSSSK**PNLLVLPIQQDASTK**LHWGNILK R**TPLMQVPVLLDLNGK**HLWVTCSQHYSSSTYQAPFCHSTQCSR**ANTHQCFTCTDSTTSR**PGCHNNTC GLISSNPVTQESGLGELAQDVLALHSTHGSKLGSLVK**IPQFLFSCAPTFLTQKGLPNNVQGALGLGH APISLPNQLFSHFGLK**RQFTMCLSSYPTSNGAILFGDINDPNNNNYIHNSLDVLHDMVYTPLTISK**Q GEYFIQVSAIR**VNKHMVIPTKNPSMFPSSSSS**SY**HESSEIGGAMITTTNPYTVLRHSIFEVFTQVFA NNVPKQAQVKAVGPFGLCYDTKKISGGVPSVDLIMDKSDVVWRISGENLMVQAQDGVSCLGFVDGGV HTRAGIALGTHQLEENLVVFDLARSRVGFNTNSLKSHGKSCSNLFDLNNP	γ-Conglutin acidic subunit	30,900	29,209	8.85	8.73	38.9
93	MSNKLLALSLFLPFLLLFFCGCFASTFRQQPQENECQFQRLNALEPDNTVQSEAGTIETWNPKNDEL RCAGVALSRCTIQRNGLRRPFYTNAPQEIYIQQGRGIFGMIFPGCGETYEEPQESEKGQGPRPQDRH QKVEHFKEGDIIAVPTGIPFWMYNDGQTPVVAITLIDTTNLDNQLDQIPRRFYLSGNQEQEFLQYQE KEGGQGQQQPEGGNVLSGFDDEFLEEALSVNKEIVRNIKGKNDDREGGIVEVKGGLKVIIPPTMRPRH GREEEEEEEEEDERRGDRRRRHPHHHHHEEEEEEEEEWSHQVRRVRRPHRHHHHRKDR**NG**LEETLCT MKLR**HNIGESTSPDAYNPQAGR**FK**TLTSIDFPILGWLGLAAEHGSIYK**NALFVPYYNVNANSILYVL NGSAWFQVVDCSGNAVFNGEL	α-Conglutin acidic subunit	23,600	20,799	5.36	6.33	39.5
96	MSNKLLALSLFLPFLLLFFCGCFASTFRQQPQENECQFQRLNALEPDNTVQSEAGTIETWNPKNDEL RCAGVALSRCTIQRNGLRRPFYTNAPQEIYIQQGRGIFGMIFPGCGETYEEPQESEKGQGPRPQDRH QKVEHFKEGDIIAVPTGIPFWMYNDGQTPVVAITLIDTTNLDNQLDQIPRRFYLSGNQEQEFLQYQE KEGGQGQQQPEGGNVLSGFDDEFLEEALSVNKEIVRNIKGKNDDREGGIVEVKGGLKVIIPPTMRPRH GREEEEEEEEEDERRGDRRRRHPHHHHHEEEEEEEEEWSHQVRRVRRPHRHHHHRKDR**NG**LEETLCT MK**LRHNIGESTSPDAYNPQAGR**FK**TLTSIDFPILGWLGLAAEHGSIYK**NALFVPYYNVNANSILYVL NGSAWFQVVDCSGNAVFNGELNEGQVLTIPQNYAAAIKSLSDNFRYVAFK**TNDIPQIATLAGANSEI SALPLEVVAHAFNLNR**DQARQLKNNNPYKFLVPPPQSQLRAVAKGELP	α-Conglutin acidic subunit	23,200	20,799	5.71	6.33	40.5
99	MSNKLLALSLFLPFLLLFFCGCFASTFRQQPQENECQFQRLNALEPDNTVQSEAGTIETWNPKNDEL RCAGVALSRCTIQRNGLRRPFYTNAPQEIYIQQGRGIFGMIFPGCGETYEEPQESEKGQGPRPQDRH QKVEHFKEGDIIAVPTGIPFWMYNDGQTPVVAITLIDTTNLDNQLDQIPRRFYLSGNQEQEFLQYQE KEGGQGQQQPEGGNVLSGFDDEFLEEALSVNKEIVRNIKGKNDDREGGIVEVKGGLKVIIPPTMRPRH GREEEEEEEEEDERRGDRRRRHPHHHHHEEEEEEEEEWSHQVRRVRRPHRHHHHRKDR**NG**LEETLCT MKLRHNIGESTSPDAYNPQAGRFK**TLTSIDFPILGWLGLAAEHGSIYK**NALFVPYYNVNANSILYVL NGSAWFQVVDCSGNAVFNGELNEGQVLTIPQNYAAAIKSLSDNFRYVAFK**TNDIPQIATLAGANSEI SALPLEVVAHAFNLN**RDQARQLKNNNPYKFLVPPPQSQLRAVAKGELP	α-Conglutin acidic subunit	20,200	20,799	4.96	6.33	30.0
103*	VSKPSLSLFSLSLLLFASACLATRSEFDRLNQCQLDNINALEPDHRVESEAGLTETWNPNNPELKCA GVSLIRRTIDPNGLHLPSFSPSPQLIFIIQGKGVLGLSLPGCPETYEEPRSSQSRQGSRQQQGDSHQ KIRRFRKGDIIAIPSGIPYWTYNHGDEPLVAISLLDTSNIANQLDSTPRVFYLGGNPETEFPETQEE QQGRHRQKHSYPVGRRSGHHQQEEESEEQNEGNSVLSGVSSEFLAQTFNTEEDTAKRLRSPRDERSQ IVRVEGGLRIINPKGKEEEEEEKEQSHSHSHREEEEEEEEDEEKQRSEERK**NG**LEETICSAKIRENIA DAAGADLYNPRAGRIRTANSLTLPVLRYLRLSAEYVRLYRNGIYAPHWNINANSLLYVIRGEGRVRI VNFQGDAVFDNKVRKGQLVVVPQNFVVAEQAGEEEGLEYVVFKTNDRAAVSHVQQVLR**ATPAEVLAN AFGLR**QRQVTELKLSGNRGPLVHPQSQSQSH	α-Conglutin acidic subunit	19,700	20,135	6.67	9.55	15.5
104*	VSKPSLSLFSLSLLLFASACLATRSEFDRLNQCQLDNINALEPDHRVESEAGLTETWNPNNPELKCA GVSLIRRTIDPNGLHLPSFSPSPQLIFIIQGKGVLGLSLPGCPETYEEPRSSQSRQGSRQQQGDSHQ KIRRFRKGDIIAIPSGIPYWTYNHGDEPLVAISLLDTSNIANQLDSTPRVFYLGGNPETEFPETQEE QQGRHRQKHSYPVGRRSGHHQQEEESEEQNEGNSVLSGVSSEFLAQTFNTEEDTAKRLRSPRDERSQ IVRVEGGLRIINPKGKEEEEEEKEQSHSHSHREEEEEEEEDEEKQRSEERK**NG**LEETICSAKIRENIA DAAGADLYNPRAGRIRTANSLTLPVLRYLRLSAEYVRLYRNGIYAPHWNINANSLLYVIRGEGRVRI VNFQGDAVFDNKVRKGQLVVVPQNFVVAEQAGEEEGLEYVVFKTNDRAAVSHVQQVLR**ATPAEVLAN AFGLR**QRQVTELKLSGNRGPLVHPQSQSQSH	α-Conglutin acidic subunit	20,300	20,135	7.09	9.55	7.7
105*	VSKPSLSLFSLSLLLFASACLATRSEFDRLNQCQLDNINALEPDHRVESEAGLTETWNPNNPELKCA GVSLIRRTIDPNGLHLPSFSPSPQLIFIIQGKGVLGLSLPGCPETYEEPRSSQSRQGSRQQQGDSHQ KIRRFRKGDIIAIPSGIPYWTYNHGDEPLVAISLLDTSNIANQLDSTPRVFYLGGNPETEFPETQEE QQGRHRQKHSYPVGRRSGHHQQEEESEEQNEGNSVLSGVSSEFLAQTFNTEEDTAKRLRSPRDERSQ IVRVEGGLRIINPKGKEEEEEEKEQSHSHSHREEEEEEEEDEEKQRSEERK**NG**LEETICSAKIRENIA DAAGADLYNPRAGRIRTANSLTLPVLRYLRLSAEYVRLYRNGIYAPHWNINANSLLYVIRGEGRVRI VNFQGDAVFDNKVRKGQLVVVPQNFVVAEQAGEEEGLEYVVFKTNDRAAVSHVQQVLR**ATPAEVLAN AFGLR**QRQVTELKLSGNRGPLVHPQSQSQSH	α-Conglutin acidic subunit	20,300	20,135	7.85	9.55	7.7
106*	VSKPSLSLFSLSLLLFASACLATRSEFDRLNQCQLDNINALEPDHRVESEAGLTETWNPNNPELKCA GVSLIRRTIDPNGLHLPSFSPSPQLIFIIQGKGVLGLSLPGCPETYEEPRSSQSRQGSRQQQGDSHQ KIRRFRKGDIIAIPSGIPYWTYNHGDEPLVAISLLDTSNIANQLDSTPRVFYLGGNPETEFPETQEE QQGRHRQKHSYPVGRRSGHHQQEEESEEQNEGNSVLSGVSSEFLAQTFNTEEDTAKRLRSPRDERSQ IVRVEGGLRIINPKGKEEEEEEKEQSHSHSHREEEEEEEEDEEKQRSEERK**NG**LEETICSAKIRENIA DAAGADLYNPRAGRIRTANSLTLPVLRYLRLSAEYVRLYRNGIYAPHWNINANSLLYVIRGEGRVRI VNFQGDAVFDNKVRKGQLVVVPQNFVVAEQAGEEEGLEYVVFKTNDRAAVSHVQQVLR**ATPAEVLAN AFGLR**QRQVTELKLSGNRGPLVHPQSQSQSH	α-Conglutin acidic subunit	20,300	20,135	8.30	9.55	7.7

(*continued on next page*)

Table 2 (*continued*)

Spot number	*Full length deduced aminoacid sequence*	*Subunit identification*	*MW*[a]	*MW*[b]	*pI*[a]	*pI*[b]	*Amino acid cov, %*
107*	VSKPSLSLFSLSLLLFASACLATRSEFDRLNQCQLDNINALEPDHRVESEAGLTETWNPNNPELKCA GVSLIRRTIDPNGLHLPSFSPSPQLIFIIQGKGVLGLSLPGCPETYEEPRSSQSRQGSRQQQGDSHQ KIRRFRKGDIIAIPSGIPYWTYNHGDEPLVAISLLDTSNIANQLDSTPRVFYLGGNPETEFPETQEE QQGRHRQKHSYPVGRRSGHHQQEEESEEQNEGNSVLSGVSSEFLAQTFNTEEDTAKRLRSPRDERSQ IVRVEGGLRIINPKGKEEEEEKEQSHSHSHREEEEEEEEDEEKQRSEERKNGLEETICSAKIRENIA DAAGADLYNPRAGRIRTANSLTLPVLRYLRLSAEYVRLYRNGIYAPHWNINANSLLYVIRGEGRVRI VNFQGDAVFDNKVRKGQLVVVPQNFVVAEQAGEEEGLEYVVFKTNDRAAVSHVQQVLRATPAEVLAN AFGLRQRQVTELKLSGNRGPLVHPQSQSQSH	α-Conglutin basic subunit	20,500	20,135	8.57	9.55	7.7
	MAKNMAPILHILVISLSYSFLFVTSSSQNSQSLYHNSQPTSSSKPNLLVLPIQQDASTKLHWGNILK RTPLMQVPVLLDLNGKHLWVTCSQHYSSSTYQAPFCHSTQCSRANTHQCFTCTDSTTSRPGCHNNTC GLISSNPVTQESGLGELAQDVLALHSTHGSKLGSLVKIPQFLFSCAPTFLTQKGLPNNVQGALGLGH						
110	APISLPNQLFSHFGLKRQFTMCLSSYPTSNGAILFGDINDPNNNNYIHNSLDVLHDMVYTPLTISKQ GEYFIQVSAIRVNKHMVIPTKNPSMFPSSSSSSYHESSEIGGAMITTTNPYTVLRHSIFEVFTQVFA NNVPKQAQVKAVGPFGLCYDTKKISGGVPSVDLIMDKSDVVWRISGENLMVQAQDGVSCLGFVDGGV HTRAGIALGTHQLEENLVVFDLARSRVGFNTNSLKSHGKSCSNLFDLNNP	γ-Conglutin small subunit	18,300	16,167	5.78	6.20	50.0
111	MAKNMAPILHILVISLSYSFLFVTSSSQNSQSLYHNSQPTSSSKPNLLVLPIQQDASTKLHWGNILK RTPLMQVPVLLDLNGKHLWVTCSQHYSSSTYQAPFCHSTQCSRANTHQCFTCTDSTTSRPGCHNNTC GLISSNPVTQESGLGELAQDVLALHSTHGSKLGSLVKIPQFLFSCAPTFLTQKGLPNNVQGALGLGH APISLPNQLFSHFGLKRQFTMCLSSYPTSNGAILFGDINDPNNNNYIHNSLDVLHDMVYTPLTISKQ GEYFIQVSAIRVNKHMVIPTKNPSMFPSSSSSSYHESSEIGGAMITTTNPYTVLRHSIFEVFTQVFA NNVPKQAQVKAVGPFGLCYDTKKISGGVPSVDLIMDKSDVVWRISGENLMVQAQDGVSCLGFVDGGV HTRAGIALGTHQLEENLVVFDLARSRVGFNTNSLKSHGKSCSNLFDLNNP	γ-Conglutin small subunit	18,900	16,167	6.21	6.20	73.1
112	MAKNMAPILHILVISLSYSFLFVTSSSQNSQSLYHNSQPTSSSKPNLLVLPIQQDASTKLHWGNILK RTPLMQVPVLLDLNGKHLWVTCSQHYSSSTYQAPFCHSTQCSRANTHQCFTCTDSTTSRPGCHNNTC GLISSNPVTQESGLGELAQDVLALHSTHGSKLGSLVKIPQFLFSCAPTFLTQKGLPNNVQGALGLGH APISLPNQLFSHFGLKRQFTMCLSSYPTSNGAILFGDINDPNNNNYIHNSLDVLHDMVYTPLTISKQ GEYFIQVSAIRVNKHMVIPTKNPSMFPSSSSSSYHESSEIGGAMITTTNPYTVLRHSIFEVFTQVFA NNVPKQAQVKAVGPFGLCYDTKKISGGVPSVDLIMDKSDVVWRISGENLMVQAQDGVSCLGFVDGGV HTRAGIALGTHQLEENLVVFDLARSRVGFNTNSLKSHGKSCSNLFDLNNP	γ-Conglutin small subunit	18,100	16,167	6.19	6.20	45.4
113	MAKNMAPILHILVISLSYSFLFVTSSSQNSQSLYHNSQPTSSSKPNLLVLPIQQDASTKLHWGNILK RTPLMQVPVLLDLNGKHLWVTCSQHYSSSTYQAPFCHSTQCSRANTHQCFTCTDSTTSRPGCHNNTC GLISSNPVTQESGLGELAQDVLALHSTHGSKLGSLVKIPQFLFSCAPTFLTQKGLPNNVQGALGLGH APISLPNQLFSHFGLKRQFTMCLSSYPTSNGAILFGDINDPNNNNYIHNSLDVLHDMVYTPLTISKQ GEYFIQVSAIRVNKHMVIPTKNPSMFPSSSSSSYHESSEIGGAMITTTNPYTVLRHSIFEVFTQVFA NNVPKQAQVKAVGPFGLCYDTKKISGGVPSVDLIMDKSDVVWRISGENLMVQAQDGVSCLGFVDGGV HTRAGIALGTHQLEENLVVFDLARSRVGFNTNSLKSHGKSCSNLFDLNNP	γ-Conglutin small subunit	18,900	16,167	6.57	6.20	42.8
114	MAKNMAPILHILVISLSYSFLFVTSSSQNSQSLYHNSQPTSSSKPNLLVLPIQQDASTKLHWGNILK RTPLMQVPVLLDLNGKHLWVTCSQHYSSSTYQAPFCHSTQCSRANTHQCFTCTDSTTSRPGCHNNTC GLISSNPVTQESGLGELAQDVLALHSTHGSKLGSLVKIPQFLFSCAPTFLTQKGLPNNVQGALGLGH APISLPNQLFSHFGLKRQFTMCLSSYPTSNGAILFGDINDPNNNNYIHNSLDVLHDMVYTPLTISKQ GEYFIQVSAIRVNKHMVIPTKNPSMFPSSSSSSYHESSEIGGAMITTTNPYTVLRHSIFEVFTQVFA NNVPKQAQVKAVGPFGLCYDTKKISGGVPSVDLIMDKSDVVWRISGENLMVQAQDGVSCLGFVDGGV HTRAGIALGTHQLEENLVVFDLARSRVGFNTNSLKSHGKSCSNLFDLNNP	γ-Conglutin small subunit	18,200	16,167	6.56	6.20	56.6
115	MAKLTILIALVAALVLVVHTSASRSSQQSCKSQLQQVNLNHCENHIIQRIQQQEEEEEGRARLRGIK HVIRRNRSSQESEELDQCCEQLNELNSQRCQCRALQQIYENQSEQCQGRQEEQLLEQELENLPRTCG FGPLRRCNVNPDEE	δ-Conglutin large subunit	17,300	9,103	4.21	4.47	51.3

Red: identified trypsin fragments.
Yellow: intersubunit cleaved peptide bond (Müntz, 1998; Scarafoni et al., 2001).
Green: cleaved internal peptide (Gayler et al., 1990).
(For interpretation of the references to colour in this table legend, the reader is referred to the web version of this article.)
[a] Experimental data.
[b] Theoretical data.

Therefore, the need of a less accurate but more effective approach for practical applications is acute. In this work, we have combined both approaches with the aim of complementing the information that can be obtained by each of the two methods. As a matter of facts, the correspondence between the spots identified by mass spectrometry and those allocated by group comparative analysis is very good, with only a few minor exceptions mainly related to overlapping of closely migrating spots in the map, as in the case of spots no. 65, 90, 99 and 115. Moreover, the identity of some spots which were only equivocally attributed by mass spectrometry was confirmed by map comparative analysis. The reliability of this latter approach strongly depended on the nature and quality of the standard proteins used as reference. For example, the best attribution was obtained with γ-conglutin, which is the lupin protein with the most peculiar and less heterogeneous subunit composition (Duranti et al., 1981). Quantitative data for this protein (Table 1) also indicate a good agreement with previous estimations, which amounted γ-conglutin to the 5–6% of total lupin globulins (Duranti et al., 1981).

Though not fully reliable at the level of each single spot, the application of group comparative analysis to seed storage proteins did prevent tedious and redundant mass spectrometric analyses of the numerous seed polypeptide isoforms and naturally cleaved forms. Successful examples of application of this latter approach are the interspecies identification of immune cross-reacting polypeptides both by specific anti-arachin basic subunits IgGs and lupin-sensitised patients' IgEs (Magni et al., 2005a,b), which led to the identification of γ-conglutin and the basic subunits of the 11S globulin as potential allergenic polypeptides. The lupin storage protein 2D map, as implemented by the use of specific antibodies on membrane-blotted 2D maps, proved to be extremely useful also for the identification of biologically active polypeptides in the seed flour and industrial derivatives, such as protein concentrate and isolates (Sironi et al., 2005). These were the cases of the plasma glucose-controlling γ-conglutin (Magni et al., 2004) and the putative cholesterol-lowering 7S globulin subunit, homologous to the soybean one (Duranti et al., 2004).

In conclusion, the two approaches used in this work to establish lupin seed 2D electrophoretic maps have a series of mutually-balancing advantages and limitations. Their combined or else alternative use, according to the aims of the research work, the availability of plant gene sequences and the knowledge of the post-translational modifications of these proteins, can greatly help fundamental and applied investigations. Various areas, including several protein applications for food, nutrition and dietetics, can benefit of these results through the improvement of our knowledge on seed storage protein biosynthesis, maturation, deposition and polymorphism.

4. Experimental

4.1. General

Mature dry seeds of white lupin (*Lupinus albus* L. var. Multitalia) were kindly provided by Dr. Massimo Fagnano, University of Naples, Italy.

4.2. Lupin total protein extract preparation

The defatted lupin flour was extracted with a solution consisting of 7 M urea, 2 M thiourea, 2% CHAPS, 65 mM 1,4-dithiothreitol (DTT) in the ratio 1/30 w/v under stirring at room temperature for 2 h. The slurry was centrifuged at 10,000*g* for 30 min and the extracted proteins were immediately analysed or kept frozen at −80 °C until use.

4.3. 2D IEF/SDS–PAGE and image processing

Isoelectric focusing was performed on 18 cm, pH 3–10 non linear IPG strips (GE Healthcare, Milan, Italy). The strips were rehydrated overnight in a solution consisting of 7 M urea, 2 M thiourea, 2% CHAPS, 65 mM 1,4-dithiothreitol (DTT), 2% IPG buffer pH 3–10 (GE Healthcare, Milan, Italy) containing the protein sample. For the total protein extract, an estimated amount of 600 μg of protein was loaded, while for the α-, β- and γ-conglutins 350, 450 and 40 μg, respectively were analysed. These amounts were optimised for the best electrophoretic performance and reflected the relative distribution of the three protein families in the lupin seeds (Duranti et al., 1981). Strips were focussed at 56,000 Vh, with a maximum of 2900 V, at 20 °C using the Multiphor II electrophoresis unit (GE Healthcare, Milan, Italy). Prior to the second dimension, strips were incubated in equilibration buffer (375 mM Tris–HCl, pH 8.8, 6 M urea, 2% SDS, 20% glycerol) with 65 mM DTT for 15 min, then with 243 mM iodoacetamide in the same buffer without DTT for 10 min.

The separation was performed in 20 × 25 × 1 cm 12% acrylamide gels, containing 0.2% bisacrylamide, using an Ettan DALTtwelve system (GE Healthcare, Milan, Italy). The running buffer consisted of 2.5 mM Tris-base, 19.2 mM glycine and 0.01% SDS. The electrophoresis conditions were: 2 mA × gel for 2 h, 5 mA × gel for 2 h and 15 mA × gel overnight at 25 °C. The gels were stained with Coomassie Brilliant Blue (CBB). Two-D SDS–PAGE Standards (BioRad, Milan, Italy) were used as internal standards in selected gels for the evaluation of the relative molecular size and pI of the spots. To minimize experimental variability, two independent total protein extracts were obtained and quadruplicate gels from each protein extract were analysed. Standard lupin proteins were run in triplicate. The gels were digitalised in an Epson Expression 1680 Pro Scanner and analyzed with ImageMaster 2D Platinum Software (GE Healthcare, Milan, Italy). Master gels were generated for each sample by matching all the available gels with the same software, with the following spot detection parameters: smooth 4, saliency 1.8, minimal area 26. When needed, computer-assisted matching was complemented by manual matching.

4.4. Protein in-gel digestion

Protein spots were excised from Coomassie Brilliant Blue stained 2D gels by modified Gilson pipette tips, transferred to sterilised 0.5 ml tubes and stored in 50% ethanol at 4 °C until digestion. After removing ethanol solution, gel pieces were incubated in 100 μl distilled H_2O (HPLC-grade) for 15 min at room temperature and then in 40 μl 50% acetonitrile (ACN) for the same time. This step was repeated three times. Afterwards, the supernatants were removed and excised gel fragments were incubated for 5 min in 40 μl 100% ACN, for 5 min in 40 μl 100 mM ammonium bicarbonate and finally in 80 μl 50 mM ammonium bicarbonate/50% ACN for 15 min. After removing the supernatants, gel pieces were dried under vacuum on a centrifugal evaporator. For the protein digestion, 20 μl trypsin solution [Sequencing Grade Modified Trypsin V5111, Promega, Madison; 12.5 ng/μl in digestion buffer (25 mM ammonium bicarbonate containing 2.5 mM $CaCl_2$, pH 7.8)] was added to each sample. After an incubation of 45 min at 4 °C, the supernatants were replaced with 20 μl digestion buffer and samples were incubated for 16 h at 37 °C. In order to extract the tryptic fragments, the gel pieces were sonicated for 5 min in a cold water bath. After collecting the supernatants, gel pieces were incubated in 20 μl of 25 mM ammonium bicarbonate for 15 min (occasionally the samples were sonicated for 2–3 min) and then, after the addition of the same amount of ACN, incubated for further 15 min. The supernatants were collected and the samples were incubated for 15 min in 20 μl of 5% formic acid and then, after the addition of the same amount of ACN, incubated for further 15 min. After the repetition of these two last steps, 100 mM DTT was added to the pooled supernatants (about 140 μl) to give a final concentration of 1 mM DTT. Finally, the samples were dry under vacuum on a centrifugal evaporator and the resulting tryptic fragments were redissolved in 10 μl of 0.1% formic acid and stored at −80 °C.

4.5. Mass spectrometry analysis and protein identification

The extracted tryptic fragments were resuspended and analysed by LC/ESI-MS/MS. For all experiments a Finningan LCQ Deca XP MAX spectrometer equipped with a Finningan Surveyor MS HPLC system (Thermo Electron Corporation, CA, USA) was used. Chromatography separations were conducted on a BioBasic C18 column (150 μm I.D. × 150 mm length and 5 μm particle size; Thermo Electron Corporation, USA), using a linear gradient from 5% to 80% ACN containing 0.05% formic acid with a flow rate of 2.5 μl/min. One run lasted 90 min, including the regeneration step. Acquisitions were performed in data-dependent MS/MS scanning mode (full MS scan range of 400–2000 m/z followed by full MS/MS scan for the most intense ions from the MS scan) and enabling a dynamic exclusion window of 3 min. Protein identifications were conducted by correlation of uninterpreted tandem mass spectra to the entries of a non-redundant protein database downloaded from the National Center for Biotechnology Information (NCBI) using TurboSEQUEST Bioworks™ 3.2 software (Thermo Electron Corporation, CA, USA). The software was set to allow two missed cleavages per peptide and considering cysteine carbamidomethylation and methionine oxidation. The precursor ion tolerance was set to 1.4 AMU. In order to identify proteins, only peptides with X-correlation > 1.5 (+1 charge), 2.0 > (+2 charge), >2.5 (+3 charge) respectively and with peptide probability $<1.0 \times 10^{-003}$ were considered. Finally, the proteins were identified as the entries with the best P-value and Score.

4.6. Preparation of lupin reference protein families

The standard protein families, namely the α-, β- and γ-conglutins were purified by using a combination of gel permeation chromatography, for the removal of undesired components, and ion exchange chromatography for the fractionation of the main lupin proteins (Cerletti et al., 1978). This established procedure, as adapted and updated by various authors (Duranti et al., 1981; Melo et al., 1994), allowed the isolation of the three main lupin protein classes, without the separation of single isoforms in each protein family. The whole procedure is detailed here below.

Dry lupin seeds were dehulled and ground to a meal with a coffee grinder. The flour was defatted in a Soxhlet apparatus by extraction with n-pentane at 37 °C for 4 h and it was sieved through a 60 mesh metal sieve. The defatted flour was suspended in distilled water adjusted at pH 5.0 in the ratio of 1:10 (w/v) and stirred for 4 h at 4 °C. The suspension was centrifuged at 10,000g for 30 min at 4 °C. The supernatant, containing the albumin fraction and other minor soluble constituents, was discarded and the pellet was resuspended in the ratio 1:20 (w/v) in 50 mM sodium phosphate buffer pH 7.5 containing 0.5 M NaCl. After stirring for 4 hours at 4 °C, the suspension was centrifuged at 10,000g for 30 min at 4 °C. The supernatant was subsequently desalted on a Sephadex G-50 column (GE Healthcare, Milan, Italy) equilibrated with 50 mM sodium phosphate buffer pH 7.5. The desalted extract was immediately fractionated on Whatman DE 52 DEAE-cellulose equilibrated with 50 mM sodium phosphate buffer pH 7.5. In these conditions, the raw γ-conglutin fraction was not retained on the column; the α- and β-conglutins were eluted by stepwise addition of 0.15 M and 0.25 M NaCl, respectively. The α- and β-conglutins were further purified on TSK DEAE 5PW column (GE Healthcare, Milan, Italy) equilibrated in 50 mM sodium phosphate buffer pH 7.5, to avoid mutual cross-contamination. The protein solutions were eluted with a NaCl linear gradient from 0 to 0.3 M.

The γ-conglutin protein solution was adjusted to pH 4.5 with acetic acid, loaded on a Whatman CM-cellulose column equilibrated in 50 mM sodium acetate buffer, pH 4.5 and then eluted by addition of 0.35 M NaCl in the same buffer. For further purification of γ-conglutin, an insulin-agarose immobilised affinity chromatography step (Sigma-Aldrich, Milan, Italy) was used (Magni et al., 2004). γ-conglutin bound to the matrix was eluted with 0.25 M NaCl.

The isolated protein groups were dialyzed at 4 °C with frequent changes of milli-Q water and freeze dried.

4.7. Lupin storage protein cDNA sequencing

The recombinant DNA techniques were performed by following standard procedures (Sambrook and Russell, 2001). mRNA was extracted from developing lupin seeds (40 days after flowering) by using the FastTrack 2.0 Kit (Invitrogen, Milan, Italy), following the manufacturer's instructions. Prior to use, to test the quality and to calculate the concentrations of the RNA solutions, either a spectrophotometric scan with record of the absorbances at 230, 260 and 280 nm and an electrophoresis analysis on 1% agarose gel were performed. The purified mRNA (250 ng) was used for full-length, RNA ligase-mediated amplification of 5′ cDNA ends (RLM-RACE) with the GeneRacer kit (Invitrogen, Milan, Italy), according to the manufacturer's instructions. To this purpose, gene specific primers (GSP) were designed from previously determined partial nucleotide sequences of some seed storage protein cDNAs (Scarafoni A., Capraro J., and Duranti M., unpublished data) and chemically synthesised. GSPs for two vicilin-like variants, conglutin δ and legumin-like genes were 5′-AGTGCTCAAAATGGAAGAAATGGGAC, 5′-AGC CACTGCTCTGAGTTGAGACTG, 5′-TCCAAACCCA-CAAGTCCTAGGCA and 5′-AAATGGAAGAGATGG-GACCCCTTCT, respectively. After amplification and purification from agarose gel, the RACE PCR products have been cloned into the pCR4-TOPO vector (Invitrogen, Milan, Italy). The resulting plasmids were sequenced at least twice in both directions by the dideoxy method (Sanger et al., 1977), using specifically designed internal primers when needed.

Acknowledgements

Grant sources for this work were from the European Commission, Fifth Framework Programme, Quality of Life and Management of Living Resources Programme, Healthy-Profood QLRT 2001–2003 and from MIUR of Italy, FIRB Projects (2001) N. RBAU01JS5C and N. RBNE01TYZF.

References

Blagrove, R.J., Gillespie, J.M., 1975. Isolation, purification and characterization of the seed globulins of *Lupinus albus*. Aust. J. Plant Physiol. 2, 13–27.

Cerletti, P., Fumagalli, A., Venturin, D., 1978. Protein composition of seeds of *Lupinus albus*. J. Food Sci. 43, 1409–1414.

Derbyshire, E., Wright, D.J., Boulter, D., 1976. Legumin and vicilin, storage protein of legumes seeds. Phytochemistry 15, 3–24.

Duranti, M., Gatehouse, J.A., Boulter, D., Cerletti, P., 1987. *In vitro* proteolytic processing of pea and jack bean storage proteins by an endopeptidase from lupin seed. Phytochemistry 26, 627–631.

Duranti, M., Horstmann, C., Gilroy, J., Croy, R.R.D., 1995. The molecular basis for *N*-glycosylation in the 11S globulin (legumin) of lupin seed. J. Protein Chem. 14, 107–110.

Duranti, M., Lovati, M.R., Dani, V., Barbiroli, A., Scarafoni, A., Castiglioni, S., Ponzone, C., Morazzoni, P., 2004. The α′ subunit from soybean 7S globulin lowers plasma lipids and upregulates liver β-VLDL receptors in rats fed a hipercholesterolemic diet. J. Nutr. 134, 1334–1339.

Duranti, M., Restani, P., Poniatowska, M., Cerletti, P., 1981. The seed globulins of *Lupinus albus*. Phytochemistry 20, 2071–2075.

Duranti, M., Sessa, F., Carpen, A., 1992. Identification, purification and properties of the precursor of conglutin β, the 7S storage globulin of *Lupinus albus* L. seeds. J. Exp. Bot. 43, 1373–1378.

Finnie, C., Maeda, K., Østergaard, O., Bak-Jensen, K.S., Larsen, J., Svensson, B., 2004. Aspects of the barley seed proteome during development and germination. Biochem. Soc. Trans. 32, 517–519.

Gallardo, K., Job, C., Groot, S.P.C., Puype, M., Demol, H., Vandekerckhove, J., Job, D., 2001. Proteomic analysis of *Arabidopsis* seed germination and priming. Plant Physiol. 126, 835–848.

Gallardo, K., Job, C., Groot, S.P.C., Puype, M., Demol, H., Vandekerckhove, J., Job, D., 2002. Proteomics of *Arabidopsis* seed germination. A comparative study of wild-type and gibberellin-deficient seeds. Plant Physiol. 129, 823–837.

Gallardo, K., Le Signor, C., Vandekerckhove, J., Thompson, R.D., Burstin, J., 2003. Proteomics of *Medicago truncatula* seed development establishes the time frame of diverse metabolic processes related to reserve accumulation. Plant Physiol. 133, 664–682.

Gayler, K.R., Kolivas, S., Macfarlane, A.J., Lilley, G.G., Baldi, M., Blagrove, R.J., Johnson, E.D., 1990. Biosynthesis, cDNA and amino acid sequences of a precursor of conglutin delta, a sulphur-rich protein from *Lupinus angustifolius*. Plant Mol. Biol. 15, 879–893.

Gianazza, E., Eberini, I., Arnoldi, A., Wait, R., Sirtori, C.R., 2003. A proteomic investigation of isolated soy proteins with variable effects in experimental and clinical studies. J. Nutr. 133, 9–14.

Hajduch, M., Ganapathy, A., Stein, J.W., Thelen, J.J., 2005. 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. 137, 1397–1419.

Islam, N., Woo, S.H., Tsujimoto, H., Kawasaki, H., Hirano, H., 2002. Proteome approaches to characterize seed storage proteins related to ditelocentric chromosomes in common wheat (*Triticum aestivum* L.). Proteomics 2, 1146–1155.

Magni, C., Ballabio, C., Restani, P., Sironi, E., Scarafoni, A., Ponesi, C., Duranti, M., 2005a. Two-dimensional electrophoresis and Western blotting analyses with anti Ara h 3 basic subunit IgG evidence the cross-reacting polypeptides of *Arachis hypogaea*, *Glycine max* and *Lupinus albus* seed proteomes. J. Agric. Food Chem. 53, 2275–2281.

Magni, C., Herndl, A., Sironi, E., Scarafoni, A., Ballabio, C., Restani, P., Bernardini, R., Novembre, E., Vierucci, A., Duranti, M., 2005b. Mono- and two-dimensional electrophoretic identification of IgE-binding polypeptides of *Lupinus albus* and other legume seeds. J. Agric. Food Chem. 53, 4567–4571.

Magni, C., Sessa, F., Accardo, E., Vanoni, M., Morazzoni, P., Scarafoni, A., Duranti, M., 2004. Conglutin γ, a lupin seed protein, binds insulin *in vitro* and reduces plasma glucose levels of hyperglycemic rats. J. Nutr. Biochem. 15, 646–650.

Majoul, T., Bancel, E., Triboï, E., Ben Hamida, J., Branlard, G., 2003. Proteomic analysis of the effect of heat stress on hexaploid wheat grain: characterization of heat-responsive proteins from total endosperm. Proteomics 3, 175–183.

Melo, T.S., Ferriera, R.B., Teixeira, A.N., 1994. The seed storage proteins from *Lupinus albus*. Phytochemistry 37, 641–648.

Müntz, K., 1998. Deposition of storage proteins. Plant Mol. Biol. 38, 77–99.

Müntz, K., Blattner, F.R., Shutov, A.D., 2002. Legumains, a family of asparagine-specific cysteine endopeptidases involved in propolypeptide processing and protein breakdown in plants. J. Plant Physiol. 159, 1287–1293.

Østergaard, O., Finnie, C., Laugesen, S., Roepstorff, P., Svensson, B., 2004. Proteome analysis of barley seeds: identification of major proteins from two-dimensional gels (p*I* 4–7). Proteomics 4, 2437–2447.

Rajjou, L., Gallardo, K., Job, C., Job, D., 2006. Proteome analysis for the study of developmental processes in plants. In: Finnie, C. (Ed.), Plant Proteomics, Annual Plant Reviews, vol. 28. Blackwell Publishing, Oxford, pp. 151–184.

Salmanowicz, B.P., Weder, J.K.P., 1997. Primary structure of 2S albumin from seeds of *Lupinus albus*. Z. Lebensm. Unters. Forsch. A 204, 129–135.

Sambrook, J., Russell, D.W., 2001. Molecular Cloning: A Laboratory Manual. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.

Sanger, F., Nicklen, S., Coulson, A.R., 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA 74, 5463–5467.

Scarafoni, A., Di Cataldo, A., Vassilevskaia, T.D., Bekman, P., Rodriguez-Pousada, C., Ceciliani, F., Duranti, M., 2001. Cloning, sequencing and expression in the seeds and radicles of two *Lupinus albus* conglutin γ genes. Bioch. Bioph. Acta 1519, 147–151.

Skilas, D.J., Copeland, L., Rathmell, W.G., Wrigley, C.W., 2001. The wheat-grain proteome as a basis for more efficient cultivar identification. Proteomics 1, 1542–1546.

Sironi, E., Sessa, F., Duranti, M., 2005. A simple procedure of lupin seed protein fractionation for selective food applications. Eur. Food Res. Technol. 221, 145–150.

Wait, R., Gianazza, E., Brambilla, D., Eberini, I., Moranti, S., Arnoldi, A., Sirtori, C., 2005. Analysis of *Lupinus albus* storage proteins by two-dimentional electrophoresis and mass spectrometry. J. Agric. Food Chem. 53, 4599–4606.