diversity

MDPI

Article

Proteomic Analysis of the Seeds of Four Wild Mexican *Lupinus* Species: Focus on Storage Proteins

Ruoxian Xu [1], Elena Sirtori [1,2], Giovanna Boschin [1], Kalina Bermudez Torres [3], Anna Arnoldi [1,*] and Gilda Aiello [4]

[1] Department of Pharmaceutical Sciences, University of Milan, Via Mangiagalli 25, 20133 Milan, Italy
[2] Axxam Spa, Via Meucci 3, 20091 Bresso, Italy
[3] Centro de Desarrollo de Productos Bióticos, Departamento de Biotecnología Apdo, Instituto Politécnico Nacional, Postal 24, Yautepec 62731, Mexico
[4] Department of Human Science and Quality of Life Promotion, Telematic University San Raffaele, 00166 Rome, Italy
* Correspondence: anna.arnoldi@unimi.it

Abstract: *Lupinus* is a wide genus, comprising between 300 and 500 species, most of them represented in America. Mexico is a secondary distribution center with more than 100 species growing along the highlands. Due to morphological similarities, the taxonomy of wild *Lupinus* species is still incomplete. It is, therefore, useful to collect morphological, chemical, and molecular data for the correct differentiation of these plants. In the present work, the composition of the seed proteins of four species: *Lupinus aschenbornii* Schauer, *Lupinus campestris* Cham and Schlecht, *Lupinus hintonii* C.P. Smith, and *Lupinus montanus* Kunth were analyzed. Seeds were collected at Iztaccihuatl—Popocatepetl National Park. Both total proteins and single protein families, purified by chromatographic procedures, were analyzed by SDS-PAGE and 2D-electrophoresis and by LC-MS/MS analysis. Data were compared with those of domesticated species whose proteomes had been already described in the literature. The protein profile may be useful for species identification since they have specific characteristics in each single species.

Keywords: wild lupin; *Lupinus*; 2D-electrophoresis; conglutins; seed proteins; anion-exchange chromatography; protein purification; LC-MS/MS

Citation: Xu, R.; Sirtori, E.; Boschin, G.; Torres, K.B.; Arnoldi, A.; Aiello, G. Proteomic Analysis of the Seeds of Four Wild Mexican *Lupinus* Species: Focus on Storage Proteins. *Diversity* **2022**, *14*, 814. https://doi.org/10.3390/d14100814

Academic Editor: Mario A. Pagnotta

Received: 31 August 2022
Accepted: 25 September 2022
Published: 29 September 2022

1. Introduction

Lupinus is a genus of the Fabaceae family whose seeds are characterized by a high protein content. The number of species belonging to this genus is still uncertain ranging from 300 to 500 and more [1] and it comprises either annual and perennial plants. Despite this great diversity, only four species have been domesticated and cultivated so far: *Lupinus albus* (white lupin), *Lupinus angustifolius* (narrow-leafed lupin), *Lupinus luteus* (yellow lupin), and *Lupinus mutabilis* (Andean lupin). Owing to their commercial interest, the chemical characterization of these species has been widely investigated, whereas this is not true for wild lupin species.

Mexico is a region of utmost importance for the biodiversity of this genus because it represents a main center of diversification [2–5]. However, owing to historical reasons, the duplicity of names and synonyms makes the taxonomy of Mexican *Lupinus* species still very problematic: any new chemical characterization may thus be useful for a better understanding of the phylogenetic relationship between species.

We have concentrated our attention on four wild Mexican lupin species, which were selected since they are easy to identify: *Lupinus aschenbornii* Schauer, *Lupinus campestris* Cham and Schlecht, *Lupinus hintonii* C.P. Smith, and *Lupinus montanus* Kunth. They are characteristic of the Mexican Iztaccihuatl–Popocatepetl National Park. *L. aschenbornii*

is part of the subalpine vegetation of the Central Mexican highlands and grows at altitudes between 3500 and 4000 m a.s.l., *L. campestris* grows in maize fields at altitudes between 2500 and 3000 m a.s.l., *L. hintonii* grows in the Pinus—Quercus forests at altitudes of 2800–3200 m a.s.l., and *L. montanus* has the widest distribution since it grows in the Pinus–Quercus forests as well as in the subalpine vegetation at altitudes between 2800 and 4200 m a.s.l.

A molecular phylogenic study based on barcode nucleotide sequences of internal transcribed spacer (ITS) indicates that these species belong to the sub-clade of taxa from Mexico included in the first clade gathering of West American species [5]. This classification shows that *L. aschenbornii* and *L. campestris* have a higher affinity, whereas *L. montanus* and *L. hintonii* fall at the opposite borders of this sub-clade.

A recent review collects data on the nutritional and bioactive compounds of some wild Mexican lupin beans focusing the attention on proteins, lipids, minerals, dietary fibers, and other bioactive compounds [2], but a characterization of the protein profile is still pending. Available information on lupin seed proteins derives mainly from studies on domesticated species that have been analyzed in detail using proteomic techniques [6–10]. Seed proteins are mainly represented by albumins and globulins in an approximate one to nine ratio [11–13]. The major protein families are named α-, β-, γ-, and δ-conglutins. α-Conglutin is a legumin-like globulin, belonging to the 11S family with a storage function. It has an oligomeric structure consisting of hexamers: each monomeric unit is composed of acidic and basic subunits deriving from the cleavage of a pro-polypeptide precursor and linked by a disulfide bridge [13]. β-conglutin is a vicilin-like protein, belonging to the 7S family, that is often the most abundant protein and has a storage function [11]. It is a trimeric protein, in which monomers consist of several polypeptides deriving from the hydrolysis of a common glycosylated precursor [11]. Its high heterogeneity is due to the expression of multigene families, whose individual genes are very closely related [6,14]. γ-conglutin is an unusual basic 7S protein, equally soluble in water and salt solutions, that in *L. albus* and *L. angustifolius* corresponds to 4–5% of the total proteins [11]. It is particularly stable to hydrolysis and gives strong interactions with metal ions [15]. Depending on the pH, it may be a tetramer, hexamer, or a monomer: each monomer is composed of two disulfide-linked subunits (17 kDa and 29 kDa), deriving from the post-translational proteolytic cleavage of a pro-polypeptide [16]. The role in the seed of this protein is still elusive [11]. Δ-conglutin is a low-molecular-weight 2S sulfur-rich albumin [11] with a monomeric structure, consisting of two subunits linked by two inter-chain disulfide bonds. The large subunit contains an intra-chain disulfide bridge, which impairs the digestibility [10]. Its structure suggests a potential defensive role.

The objective of the present work was the investigation of the proteome of these lupin species with the goal of providing useful information to integrate other chemical, molecular, and morphological data to improve the taxonomic identification of these species [5], offering new opportunities for plant systems biology. The investigation was performed by integrating classical analysis methods, such as SDS-PAGE and 2D-electrophoresis, with chromatographic purification and mass spectrometry analysis by comparing the results with proteomic data of domesticated species.

2. Materials and Methods

2.1. Sampling

Seeds were collected between June 2009 and June 2010 in the Popocatepetl—Iztaccihuatl National Park (Mexico) from mature pods, close before opening (fruits of these species are indehiscent). Only mature pods easily opened by a slight pressure were collected. The four Mexican lupin species were picked up in nature: *L.* grows mainly in cornfields, *L. hintonii* in *Pinus* forest, and *L. montanus* and *L. aschenbornii* are sympatric species associated with *Muhlenbergia* species in open subalpine meadows. Herbarium vouchers were collected from three flowering individuals per species and deposited at the Herbario Nacional MEXU: *L. aschenbornii* (voucher No. 1297311) and *L. montanus* (voucher No. 1297279) at altitude of

3889 m a.s.l., *L. campestris* (voucher No. 1297299) at 2781 m a.s.l., and *L. hintonii* (voucher No. 1344434) at 2960 m a.s.l. The identification was based on the morphological characteristics and by comparison with herbarium specimens [17–19].

2.2. Separation of Total Protein Extracts (TPEs) from Lupin Seeds

The soluble proteins (albumins and globulins) were extracted from defatted lupin flour, obtained from the seeds by using a lab mill with 100 mM Tris-HCl/0.5 M NaCl and pH 8.0 for 2 h at room temperature (RT) with gentle stirring. The slurry was centrifuged at 6000× *g* for 30 min, 4 °C, and the supernatant was dialyzed against 100 mM Tris-HCl, pH 8.0, for 24 h at 4 °C. Dialyzed TPEs were immediately analyzed or aliquoted and stored at −20 °C. The protein content was assessed by the colorimetric method of Bradford [20].

2.3. Purification of Globulins

The globulins were purified according to a published method [21], using a preparative HPLC 1200 with a diode-array detector (DAD) (Agilent Technologies, Palo Alto, Santa Clara, CA, USA). In brief, the TPE was filtered through sterile membrane Econofilter filters, 0.2 µm (Agilent Technologies), and loaded onto a DEAE-Sepharose Fast Flow column (1.6 × 2.5 cm, 15–70 µm, 5 mL column volume) (GE Healthcare Bio-Sciences AB, Uppsala, Sweden). Bound proteins were eluted with a linear salt gradient (0–100% NaCl over 16 column volumes) in 100 mM Tris-HCl pH 8.0, and the fractions were collected every 20 s. The total DEAE elution was further fractionated onto a MonoQ HR 5/5 anion-exchange column (Amersham Biosciences Europe GmbH, Freiburg, Germany) with a linear salt gradient (0–100% NaCl) in 100 mM Tris-HCl, pH 8.0, and the fractions were collected every 15 s. Samples were dialyzed against 100 mM Tris-HCl, pH 8.0, using a dialysis tubing made of regenerated cellulose with a cut-off of 7 kDa (Millipore, Burlington, MA, US), under gentle stirring for 24 h at 4 °C. Protein samples were immediately analyzed or aliquoted and kept frozen at −20 °C until use. The protein content was assessed according to the Bradford procedure [20].

2.4. SDS-PAGE Analysis

Desalted column fractions from the MonoQ column were separated via a classical reducing SDS-PAGE: samples were mixed 1:1 with 2X SDS-loading buffer (50 mM Tris-HCl pH 6.8, 2% SDS, 10% glycerol, and 0.1% bromophenol blue, 5% β-mercaptoethanol to add immediately before use) and heated for 7 min in Eppendorf tubes in boiling water. Samples were run on 13% linear gradient SDS-gels, and 2–10 µg of protein were applied to each well. A precision plus protein standard (Bio-Rad Laboratories Inc., Hercules, CA, USA), ranging from 250 to 10 kDa, was used as a reference. The cathodic and anodic compartments were filled with Tris–glycine buffer, pH 8.3, containing 0.1% m/v SDS. The electrophoresis was conducted at 100 V until the dye front reached the gel bottom. The gels were stained with Bio-Safe Coomassie (Bio-Rad), scanned in a Versa Doc 3000 Imaging System (Bio-Rad), and analyzed using Quantity One software (Bio-Rad).

2.5. Two-Dimensional-Electrophoresis

TPEs or desalted column fractions (20 µL, about 7 µg/µL) were diluted in IEF sample buffer (7 M urea, 2 M thiourea, 3% (*w*/*v*) 3-[(3-cholamidopropyl)-dimethylammonio]-1-propanesulfonate hydrate (CHAPS), 1% ampholyte pH 3–10 and pH 4–8). The proteins were reduced with 65 mM dithiothreitol (DTT) and alkylated with 200 mM 2-iodoacetamide (IAM); both steps for 1 h at RT in the dark [10]. Isoelectric focusing was performed on 7-cm, pH 3–10 non-linear IPG strips (Bio-Rad) using a gradient voltage of 100 V for the first hour, 200 V for the second hour, and 500 V for the last half hour. The second dimension was performed on 13% SDS-PAGE. The gels were stained with Bio-Safe Coomassie (Bio-Rad), which detects proteins without the use of methanol or acetic acid. This staining produces a white background on the gel, with greater spot intensity, allowing a good qualitative/quantitative analysis by the software. Gels were scanned in a Versa Doc 3000

Imaging System (Bio-Rad). The software used to compare the 2D maps was PD Quest Basic 2-D Analysis V7.0 (Bio-Rad). Each sample was run in duplicate.

2.6. LC-MS/MS Analysis

The proteins separated on the SDS-PAGE were sliced, reduced with 5 mM DL-dithiothreitol (DTT, Sigma-Aldrich, St. Louis, MO, USA) for 30 min at 52 °C, centrifuged at 500 rpm, and alkylated with 15 mM iodoacetamide (Sigma-Aldrich) for 20 min in the dark at room temperature. Proteins were digested with trypsin (E:S ratio equal to 1:20) overnight at 37 °C and then identified by LC-MS/MS. Aliquots of 5 µL of tryptic peptides were injected into a nano-chromatographic system, HPLC-Chip (Agilent Palo Alto, Santa Clara, CA, USA). The analysis was conducted on an SL IT mass spectrometer. LC-MS/MS analyses were performed in data-dependent acquisition Auto MS(n) mode. To increase the number of identified peptides, three technical replicates (LC-MS/MS runs) were run for each hydrolyzed band. The Spectrum Mill Proteomics Workbench (Rev B.04.00, Agilent), consulting *Lupinus* (205,529 entries of which 123 were reviewed by Swiss-Prot) database (downloaded on February 2022 from UniProtKB), was used for the automated peptide identification from tandem mass spectra. The MS/MS spectra were then searched against *Viridiplantae* UniProtKB database confirming the previous identifications (data not shown). Trypsin was selected as cutting enzyme with two allowed missed cleavages; carbamidomethylation was chosen as a fixed modification. The mass tolerance was set at 1.0 Da and 0.8 Da for MS1 and MS2, respectively. Protein identification was performed by applying filters with Score > 8, SPI (Score Peak Intensity) > 60%. Proteins identified with at least two different peptide sequences were studied in more detail. The characteristics of these proteins, such as the score, molecular weight, pI, number, and sequence of peptides, are summarized in Supporting Information Table S1.

3. Results and Discussion

3.1. Technical Approach

The investigation of the proteomes of the four wild lupin species was performed using a bottom-up proteomic approach, including: the separation of the total protein extracts (TPEs) from the dry seeds, the purification of the main storage proteins by a chromatographic procedure (ion-exchange chromatography on a DEAE-FF column combined with a Mono Q column), and the analysis by 2D-electrophoresis in reducing conditions. In parallel, the proteins were submitted to tryptic hydrolysis and analysis by LC-MS/MS. The identification of the lanes by LC-ESI-MS/MS was, however, impaired by the absence of the protein sequences of these wild lupin species in the international protein databases. The MS/MS data analysis was thus accomplished by a comparison with the data of the two domesticated species *L. albus* and *L. angustifolius*, previously investigated in our laboratory and for which numerous sequences are reported in international databases, such as SWISS-PROT.

For clarity, we report here some relevant features of the proteins of these domesticated species. The HPLC chromatograms of the TPEs are characterized by the presence of the following main peaks that are found also in the wild species: one small peak of γ-conglutin with retention time (RT) 1.5 min, a jagged large peak of β-conglutin with RT between 25 and 32 min that derives from the superimposition of the many peaks of this heterogeneous class of proteins, and an intense simple peak of α-conglutin with RT 35—43 min. The β-conglutin/α-conglutin ratios depend on the species: in *L. angustifolius*, this value is approximately equal to 1:2, whereas in *L. albus*, it is equal to 1:0.6. The 2D-electrophoresis of the total protein extracts and of the fractions separated by preparative HPLC as well as the mass spectrometry data are reported in published papers [8,10,22]. These data were the basis for the interpretation of the results obtained by analyzing the wild lupin proteins.

In the wild lupins, the LC-MS/MS approach allowed the identification of seed storage proteins as well as of some cytosolic proteins. From this point on, the discussion will be limited to storage proteins alone which will be described in the following paragraphs,

whereas the information concerning the proteins identified as a whole are reported in the Supplementary Materials (Table S1).

3.2. L. hintonii Protein Profile

Figure 1A shows the HPLC chromatogram of the TPE from *L. hintonii* and Figure 1B, the SDS-PAGE of the collected fractions. The first eluting peak (peak one, RT 5 min) may be assigned to γ-conglutin, the three major peaks (peaks two-four) between 25 and 35 min to β-conglutin, and the largest peak (peak five, 37–40 min) to α-conglutin. The following peaks (peaks six and seven) remained unidentified, but the SDS-page indicates that they do not contain proteins. These peaks were detected in all species and will not be further discussed. Both the TPE and purified fractions were further characterized by 2D-electrophoresis. Figure 1C shows the 2D map of the reduced TPE, whereas panels D, E, and F, respectively, show the 2D maps of γ-conglutin, β-conglutin, and α-conglutin. Immobilized pH gradient (IPG) runs were performed in the presence of 7 M urea, hence non-covalently bound oligomers were dissociated in all cases and the total spots, detected by the PD Quest Software, were 116. In most cases, rows rather than single spots were detected, in particular series of polypeptides with similar molecular weights (MW), but varying pI values with different degrees of post-translational phosphorylation and glycosylation. In fact, all the lupin globulins derive from a unique common ancestor polypeptide, which undergoes proteolytic cleavage, giving a complex mixture of polypeptides, which aggregates to form globulins [11]. Owing to these phenomena, globulins are an extremely heterogeneous class of proteins. Since neither gene nor protein sequences are available in the literature and in the recent database for *L. hintonii* so far, the protein identification by LC-MS/MS was based on the hypothesis of existing homologies with other well-known lupin species, i.e., *L. albus* and *L. angustifolius* [6,8,10,23].

Figure 1D shows the 2D maps of γ-conglutin, where three faint spots at 30 kDa and some other spots at 16–18 kDa appear, corresponding to the two disulfide linked subunits of the monomer. Unfortunately, the identification by LC-MS/MS analysis failed probably owing to a low sequence homology in respect to known species. In addition, γ-conglutin is resistant to tryptic hydrolysis, therefore, difficult to be observed in this data set.

Numerous spots correspond to β-conglutin (Figure 1E), the major seed storage protein in the genus Lupinus: 154 spots were identified, ranging from MW 72 to 13 kDa and from pI 5.2 to 8.5 A group of spots around 50–75 kDa corresponds to mature β-conglutin subunits, whereas all other spots from MW 15 to 40 kDa correspond to different proteolytic forms. Often these spots are further arranged into clusters of rows with closely spaced MW and pI values due to their different phosphorylation degrees. In detail, by mass spectrometry the following β-conglutin isoforms were detected: conglutin beta (B8Q5G0), conglutin beta (Fragment) (B0YJF7), conglutin beta 1 (F5B8V9), conglutin beta 5 (F5B8W3), and conglutin beta 7 (F5B8W5) of *L. angustifolius*, as well as conglutin beta 2 (Q6EBC1) of *L. albus*, with molecular weights ranging from 72 to 61 kDa and from pI 5.3 to 6.4. Table 1 lists the protein and relative identified peptides for each isoform. Among conglutin beta isoforms, the % A.A. coverage ranged from 4 to 14.4%, with the highest one observed for conglutin beta 2 (Q6EBC1). These isoforms show a high degree of identity with conservation often occurring in hydrophilic domains that are enriched in the amino acids Glu, Gln, and Arg. Such a high variety of isoforms among family members is justified by the fact that often insertions/deletions of repeated amino acid stretches of predominantly glutamic acid (E), glutamine (Q), serine (S), glycine (G), and arginine (R) are involved in the sequences. Despite the high similarity between the amino acid sequences across the isoforms of β-conglutin, our study confirmed the high micro-heterogeneities due to their polymorphism which ranges from 1% to 26% [24], mainly due to their multigenic origin. Moreover, there are obvious differences between β-conglutin across the species. In the case of *L. albus*, the 2D analysis of β-conglutin revealed that the mature protein is composed of 10 to 12 major types of subunits, with MW ranging from 15 to 20 kDa with acidic pI, from 17 to 30 kDa with basic pI, a big group of phosphorylated spots from 30 to 50 kDa, and from 50 to 70 kDa,

as well as a considerable number of minor constituents. In contrast, the protein from *L. angustifolius* contains two groups of polypeptides: a heavier and more abundant group (50 to 72 kDa) and a lighter group (15 to 40 kDa), whereas β-conglutins from *L. luteus* and *L. mutabilis* are essentially composed by heavy polypeptides (50 to 70 kDa) [25]. Considering these different aspects and the data collected so far, in our experimental conditions, the *L. hintonii* β-conglutin profile appears to be more similar to that of *L. angustifolius*, a fact confirmed by the protein identified by mass spectrometry.

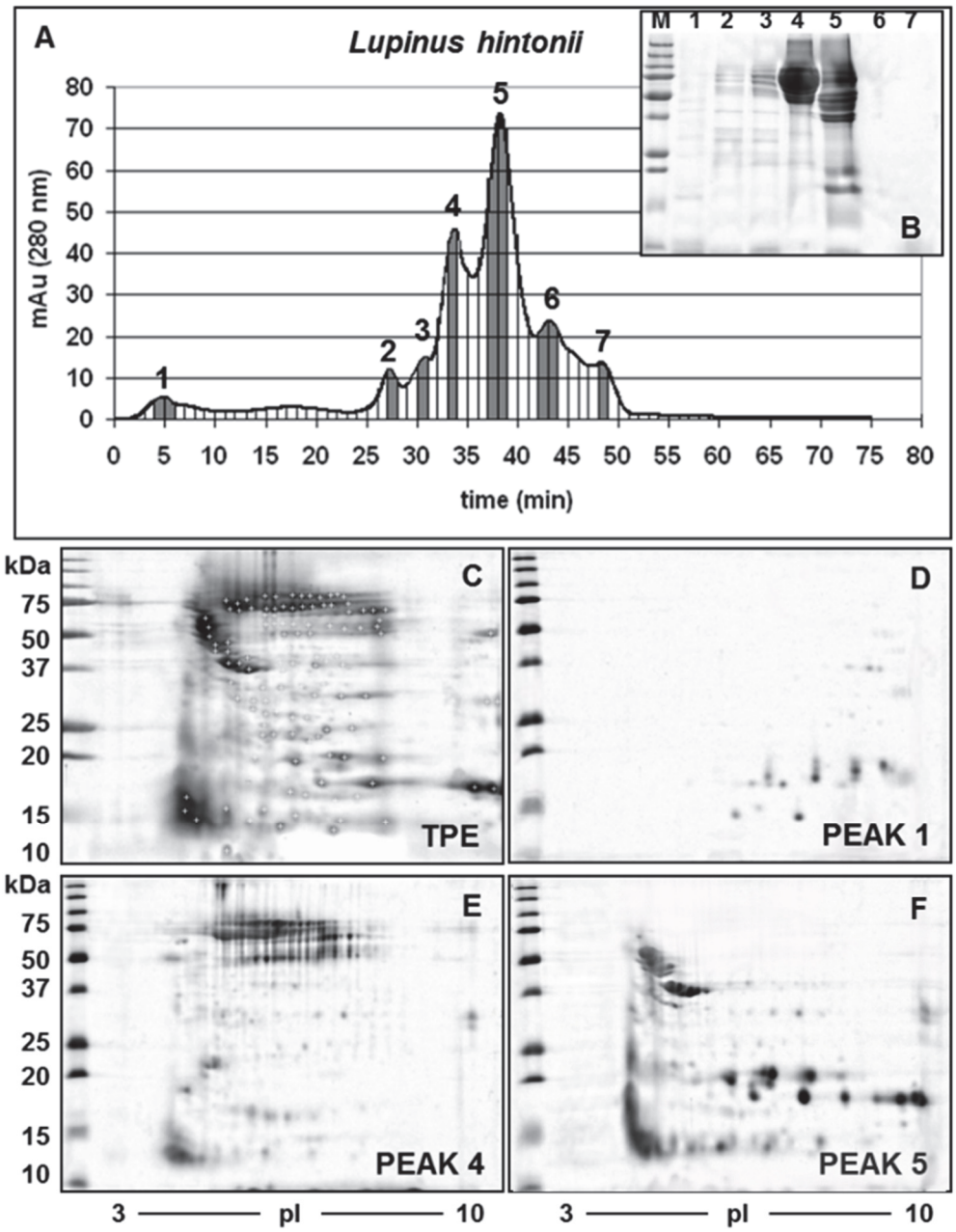

Figure 1. Electrophoretic gels of proteins of *L. hintonii*. (**A**) HPLC chromatogram of the total protein extract (TPE) on a MonoQ column. (**B**) SDS-PAGE of preparative HPLC collected fractions. (**C**) Proteome reference map of reduced TPE (proteins were run in the first dimension on 7-cm, pH 3–10 non-linear IPG strips; in the second dimension on 13% denaturing SDS–PAGE gel), where in total 116 spots were detected. (**D**) Two-dimensional map of purified γ-conglutin. (**E**) Two-dimensional map of purified β-conglutin. (**F**) Two-dimensional map of purified α-conglutin. The standard marker (Precision Plus Protein) is indicated in kDa on the left.

Table 1. Identified storage proteins in the seeds of the four wild lupin species by LC-MS/MS.

Protein Name (Accession N.)	Species	% AA Coverage	Protein PI	Peptide Sequences	*m*/*z* Measured (Da)	MH^+ Matched (Da)	Peptide PI
L. hintonii							
Conglutin beta (Fragment) ***(B0YJF7)***	*L. angustifolius*	8.7	5.3	(R)**LPAGTTSYILNPDDNQNLR**(V) *	701.67	2102.051	4.21
				(R)**LLGFGINANENQR**(N)	723.85	1445.750	6.00
Conglutin beta ***(B8Q5G0)***		11	5.8	(R)NFLAGSEDNVISQLDREVK(E)	712.34	2134.077	4.32
				(R)NFLAGSEDNVISQLDR(E)	593.58	1777.872	4.03
Conglutin beta 7 ***(F5B8W5)***		8.9	5.6	(R)TNRLENLQNYR(I)	711.20	1420.729	8.41
				(R)LENLQNYR(I)	525.41	1049.537	6.00
Conglutin beta 1 ***(F5B8V9)***	*L. angustifolius*	4	5.8	(R)**NPYHFSSNR**(P)	560.75	1121.512	6.5
				(R)LENLQNYR(I)	525.41	1049.537	6.00
Conglutin beta 2 ***(Q6EBC1)***	*L. albus*	14.4	6.4	(R)**LLGFGINADENQR**(N)	482.72	1446.734	4.37
				(K)**INEGALLLPHYNSK**(A)	523.92	1568.843	6.75
				(R)TNRLENLQNYR(I)	711.20	1420.729	8.41
				(R)IVEFQSKPNTLILPK(H)	576.55	1727.010	8.59
				(R)LENLQNYR(I)	525.41	1049.537	6.00
Conglutin beta 5 **(F5B8W3)**	*L. angustifolius*	11.2	5.8	(K)HSDADYILVVLNGR(A)	524.75	1571.818	5.21
				(R)LLGFGINADENQR(N)	482.72	1446.734	4.37
				(R)**LPAGTTSYILNPDDNQDLR**(V)	701.99	2103.035	3.93
				(R)**ILLGYEDEQEDEEQR**(R)	933.90	1865.840	3.71
				(R)TNRLENLQNYR(I)	711.20	1420.729	8.41
				(R)LENLQNYR(I)	525.41	1049.537	6.00
Conglutin alfa 1 ***(F5B8V6)***	*L. angustifolius*	6.9	5.3	(R)**RFYLSGNQEQEFLQYQQK**(E)	769.66	2306.120	6.14
				(R)**FYLSGNQEQEFLQYQQK**(E)	717.62	2150.019	5.43
				(R)RPFYTNAPQEIYIQQGR(G)	694.39	2081.056	8.59
Legumin like-protein ***(Q53I54)***	*L. albus*	6.6	5.5	(R)**FYLSGNQEQEFLQYQEK**(E)	717.85	2151.003	4.25
				(R)RPFYTNAPQEIYIQQGR(G)	694.39	2081.056	8.45
				(R)**RFYLSGNQEQEFLQYQEK**(E)	769.73	2307.104	4.79

Table 1. *Cont.*

Protein Name (Accession N.)	Species	% AA Coverage	Protein PI	Peptide Sequences	*m*/*z* Measured (Da)	MH^+ Matched (Da)	Peptide PI
L. campestris							
Conglutin beta 1 ***(F5B8V9)***	*L. angustifolius*	8.8	5.8	(R)ILLGYEDEQEDEEQR(R)	622.90	1865.840	3.71
				(R)TNRLENLQNYR(I)	474.46	1420.729	8.41
				(R)LLGFGINADENQR(N)	724.10	1446.734	4.37
				(R)SNEPIYSNK(F)	526.39	1051.506	5.72
				(R)ATITIVNPDKR(Q)	409.99	1227.706	8.79
				(R)LENLQNYR(I)	524.96	1049.537	6.00
Conglutin beta 2 ***(F5B8W0)***	*L. angustifolius*	11.4	5.6	(R)IVEFQSKPNTLILPK(H)	576.58	1727.010	8.59
				(R)LPAGTTSYILNPDDNQNLR(V)	701.59	2102.051	4.21
				(R)ATITIVNPDKR(Q)	409.99	1227.706	8.79
				(R)LLGFGINADENQR(N)	723.94	1446.734	4.37
Conglutin beta 3 ***(F5B8W1)***	*L. angustifolius*	14.1	5.7	(R)IVEFQSKPNTLILPK(H)	576.58	1727.010	8.59
				(R)TNRLENLQNYR(I)	474.46	1420.729	8.41
				(R)LPAGTTSYILNPDDNQNLR(V)	701.59	2102.051	4.21
				(R)LLGFGINADENQR(N)	723.94	1446.734	4.37
				(R)LENLQNYR(I)	525.39	1049.537	6.00
				(K)AIFVVVVDEGEGNYELVGIRDQQR(Q)	902.40	2705.389	4.18
Conglutin beta 4 ***(F5B8W2)***	*L. angustifolius*	9.3	5.9	(R)ILLGYEDEQEDEEQR(R)	622.82	1865.840	3.71
				(R)TNRLENLQNYR(I)	474.46	1420.729	8.41
				(R)LLGFGINADENQR(N)	724.10	1446.734	4.37
				(R)SNEPIYSNK(F)	526.39	1051.506	5.72

Table 1. *Cont.*

Protein Name (Accession N.)	Species	% AA Coverage	Protein PI	Peptide Sequences	*m*/*z* Measured (Da)	MH^+ Matched (Da)	Peptide PI
Conglutin beta 5 ***(F5B8W3)***	*L. angustifolius*	12.2	5.8	(R)TNRLENLQNYR(I)	474.46	1420.729	8.41
				(R)ILLGYEDEQEDEEQR(R)	622.82	1865.840	3.71
				(R)LLGFGINADENQR(N)	724.10	1446.734	4.37
				(R)SNEPIYSNK(F)	526.39	1051.506	5.72
				(R)ATITIVNPDKR(Q)	409.99	1227.706	8.79
				(R)LENLQNYR(I)	525.39	1049.537	6.00
				(R)**LPAGTTSYILNPDDNQDLR**(V)	701.71	2103.035	3.93
Conglutin beta 6 ***(A0A6A4QEF0)***	*L. albus*	13.7	5.4	(R)ILLGYEDEQEDEEQR(R)	622.90	1865.840	3.71
				(R)TNRLENLQNYR(I)	474.46	1420.729	8.41
				(R)LLGFGINADENQR(N)	724.10	1446.734	4.37
				(R)SNEPIYSNK(F)	526.39	1051.506	5.72
				(R)ATITIVNPDKR(Q)	613.93	1227.706	8.79
				(R)LENLQNYR(I)	524.96	1049.537	6.00
Conglutin beta (Fragment) ***(B0YJF7)*** Conglutin beta 7 ***(F5B8W5)*** Conglutin beta ***(B8Q5G0)***	*L. angustifolius*	6.3 12.5 5.5	5.3 5.6 5.8	(R)NFLAGSEDNVISQLDR(E)	889.93	1777.872	4.03
				(R)LPAGTTSYILNPDDNQNLR(V)	701.59	2102.051	4.21
				(R)LLGFGINANENQR(N)	723.71	1445.750	6.00
				(R)ATITIVNPDKR(Q)	613.93	1227.706	8.79
Conglutin alpha 3 ***(F5B8V8)***	*L. angustifolius*	2	5.3	(R)**ADLYNPTAGR**(I)	539.27	1077.532	5.88
L. aschenbornii							
Conglutin beta ***(B8Q5G0)***	*L. angustifolius*	9.4	5.8	(R)LPAGTTSYILNPDDNQNLR(V)	701.68	2102.051	4.21
				(R)**IIEFQSKPNTLILPK**(H)	581.31	1741.026	8.59
				(R)TNRLENLQNYR(I)	711.18	1420.729	8.41
				(R)LLGFGINANENQR(N)	723.65	1445.750	6.00

Table 1. *Cont.*

Protein Name (Accession N.)	Species	% AA Coverage	Protein PI	Peptide Sequences	*m*/*z* Measured (Da)	MH^+ Matched (Da)	Peptide PI
Conglutin beta (Fragment)***(B0YJF7)*** Conglutin beta 1 ***(F5B8V9)***	*L. angustifolius*	14.5	5.8	(R)**NFLAGSEDNVISQLDR**(E)	889.93	1777.872	4.03
				(R)LPAGTTSYILNPDDNQNLR(V)	701.68	2102.051	4.21
				(R)NFLAGSEDNVISQLDREVK(E)	712.35	2134.077	4.32
				(R)TNRLENLQNYR(I)	711.18	1420.729	8.41
				(R)LLGFGINANENQR(N)	723.65	1445.750	6.00
Conglutin beta 2 ***(Q6EBC1)***	*L. albus*	14.4	6.4	(K)INEGALLLPHYNSK(A)	523.92	1568.843	6.75
				(R)LLGFGINADENQR(N)	724.26	1446.734	4.37
				(R)TNRLENLQNYR(I)	711.18	1420.729	8.41
				(R)IVEFQSKPNTLILPK(H)	576.54	1727.010	8.59
				(R)LSEGDIFVIPAGYPISINASSNLR(L	845.14	2533.330	4.37
Conglutin beta 5 ***(F5B8W3)***	*L. angustifolius*	8.3	5.8	(R)ILLGYEDEQEDEEQR(R)	933.88	1865.840	3.71
				(R)LLGFGINADENQR(N)	724.26	1446.734	4.37
				(R)TNRLENLQNYR(I)	711.18	1420.729	8.41
				(R)**LPAGTTSYILNPDDNQDLR**(V)	701.83	2103.035	3.93
Conglutin beta 6 ***(A0A6A4QEF0)***	*L. albus*	10.8	5.8	(R)IVEFQSKPNTLILPK(H)	575.65	1727.010	8.59
				(R)TNRLENLQNYR(I)	711.18	1420.729	8.41
				(R)SNEPIYSNK(Y)	526.39	1051.506	5.72
				(R)LENLQNYR(I)	525.40	1049.537	6.00
				(R)**LIENQQQSYFANALPQQQQQSEK**(E)	907.49	2720.327	4.53
Conglutin beta 7 ***(F5B8W5)***	*L. angustifolius*	9.4	5.6	(R)**NFLAGSEDNVISQLDR**(E)	889.93	1777.872	4.03
				(R)LPAGTTSYILNPDDNQNLR(V)	701.68	2102.051	4.21
				(R)NFLAGSEDNVISQLDREVK(E)	712.35	2134.077	4.32
				(R)TNRLENLQNYR(I)	711.18	1420.729	8.41
				(R)LLGFGINANENQR(N)	723.65	1445.750	6.00
Conglutin alpha 1 ***(F5B8V6)***	*L. angustifolius*	5.5	5.3	(R)**FYLSGNQEQEFLQYQQK**(E	717.60	2150.019	4.53
				(R)**LNALEPDNSVK**(S)	600.56	1199.627	4.37

Table 1. *Cont.*

Protein Name (Accession N.)	Species	% AA Coverage	Protein PI	Peptide Sequences	*m*/*z* Measured (Da)	MH^+ Matched (Da)	Peptide PI
Conglutin alpha 2 ***(F5B8V7)***	*L. angustifolius*	3.2	5.1	(K)**TNDLAATSPVK**(Q)	558.92	1116.590	5.50
				(R)**LLENIAKPSR**(A)	570.88	1140.674	8.75
Conglutin alpha 3 ***(F5B8V8)***	*L. angustifolius*	4.9	5.3	(R)**ENIADPSRADLYNPTAGR**(I)	654.29	1959.952	4.56
				(R)**ADLYNPTAGR**(I)	539.39	1077.532	5.88
Legumin-like protein ***(Q53I54)***	*L. albus*	3.5	5.5	(R)**FYLSGNQEQEFLQYQEK**(E)	717.95	2151.003	4.25
				(R)RFYLSGNQEQEFLQYQEK(E)	769.71	2307.104	4.79
L. montanus							
Conglutin beta ***(B8Q5G0)*** Conglutin beta (Fragment)***(B0YJF7)***	*L. angustifolius*	9.4 14.5	5.8 5.3	(R)NFLAGSEDNVISQLDR(E)	889.91	1777.872	4.00
				(R)LLGFGINANENQR(N)	723.71	1445.750	6.00
				(R)NFLAGSEDNVISQLDREVK(E)	712.25	2134.077	4.34
				(R)TNRLENLQNYR(I)	474.44	1420.729	8.41
				(K)HSDADYILVVLNGR(A)	524.75	1571.818	5.21
Conglutin beta 1 ***(F5B8V9)***	*L. angustifolius*	7	5.8	(K)**ELTFPGSIEDVER**(N)	745.86	1491.732	3.91
				(R)NFLAGSEDNVISQLDR(E)	889.91	1777.872	4.03
				(R)LLGFGINANENQR(N)	723.71	1445.750	6.00
Conglutin beta 2 ***(Q6EBC1)***	*L. albus*	9.9	6.4	(R)IVEFQSKPNTLILPK(H)	864.48	1727.010	8.59
				(K)INEGALLLPHYNSK(A)	523.89	1568.843	6.75
				(R)TNRLENLQNYR(I	474.44	1420.729	8.41
				(R)LLGFGINADENQR(N)	482.72	1446.734	4.37
Conglutin beta 3 ***(F5B8W1)***	*L. angustifolius*	6.7	5.7	(R)IVEFQSKPNTLILPK(H)	864.48	1727.010	8.59
				(K)INEGALLLPHYNSK(A)	523.89	1568.843	6.75
				(R)TNRLENLQNYR(I)	474.44	1420.729	8.41
				(R)LLGFGINADENQR(N)	482.72	1446.734	4.37
Conglutin beta 4 ***(F5B8W2)***	*L. angustifolius*	9.8	5.9	(K)**FGNFYEITPNR**(N)	679.08	1358.650	5.7
				(R)TNRLENLQNYR(I)	474.44	1420.729	8.41
				(R)LLGFGINADENQR(N)	482.72	1446.734	4.37

Table 1. *Cont.*

Protein Name (Accession N.)	Species	% AA Coverage	Protein PI	Peptide Sequences	*m*/*z* Measured (Da)	MH^+ Matched (Da)	Peptide PI
Conglutin beta 5 ***(F5B8W3)***	*L. angustifolius*	8.3	5.8	(R)ILLGYEDEQEDEEQR(R)	933.78	1865.840	3.71
				(R)TNRLENLQNYR(I)	474.44	1420.729	8.41
				(K)HSDADYILVVLNGR(A)	524.75	1571.818	5.21
				(R)LLGFGINADENQR(N)	482.72	1446.734	4.37
Conglutin beta 6 ***(F5B8W4)***	*L. angustifolius*	9.7	6.2	(R)IVEFQSKPNTLILPK(H	864.48	1727.010	8.59
				(R)**TDRLENLQNYR**(I)	474.44	1421.713	5.73
				(R)LLGFGINADENQR(N)	482.72	1446.734	4.30
Conglutin beta 7 ***(F5B8W5)***	*L. angustifolius*	9.4	5.6	(R)NFLAGSEDNVISQLDR(E)	889.91	1777.872	4.03
				(R)LLGFGINANENQR(N)	723.71	1445.750	6.00
				(R)NFLAGSEDNVISQLDREVK(E)	712.25	2134.077	4.32
				(R)TNRLENLQNYR(I)	474.44	1420.729	8.41
				(K)HSDADYILVVLNGR(A)	524.75	1571.818	5.21
Conglutin alpha 1 ***(F5B8V6)***	*L. angustifolius*	9.1	5.3	(R)RPFYTNAPQEIYIQQGR(G)	694.67	2081.056	8.59
				(R)RFYLSGNQEQEFLQYQQK(E)	769.68	2306.120	6.14
				(R)**LNALEPDNSVK**(S)	600.35	1199.627	4.37
Conglutin alpha 3 ***(F5B8V8)***	*L. angustifolius*	2.2	5.3	(R)**ADLYNPTAGR**(I)	539.43	1077.532	5.88
Legumin-like protein ***(Q53I54)***	*L. albus*	3.5	5.5	(R)**FYLSGNQEQEFLQYQEK**(E)	717.95	2151.003	4.25
				(R)RFYLSGNQEQEFLQYQEK(E)	769.71	2307.104	4.79

* Peptides marked in bold are unique peptides for each protein isoform.

The other main storage protein is α-conglutin deriving from a precursor (pre-pro-polypeptides), comprising an N-terminal acidic alpha-chain (with greater MW) and a C-terminal basic beta-chain (with lower MW) [26]. Several spots belonging to α-conglutin were identified in the 2D-electrophoresis (Figure 1F). They are resolved at distinct positions in the 2D-gel under the denaturating conditions used. In particular, the acidic subunit (40–52 kDa) has a pI value of 4.6–5.0, while the basic subunit (17–22 kDa) has a pI value of 6.2–8.6. Although a high heterogeneity was observed by the 2D-gel, the LC-MS analysis allowed the identification of peptides belonging only to the acidic subunit at 58 kDa and pI 5.3 of conglutin alpha 1 (F5B8V6) of *L. angustifolius* and legumin-like protein (Q53I54) of *L. albus* with % A.A. of 6.9 and 6.6, respectively. In the seed of *L. albus*, the proteolytic cleavage of the 72 and 67 kDa protomers gives origin to processed precursor polypeptides of either 50–52 or 37–44 kDa, linked through disulfide bonds to the β-polypeptide of 20–22 kDa, typical of the mature legumin. We hypothesized a similar maturation cascade also for *L. hintonii*, which contributes to the significant micro-heterogeneity of α-conglutin. Instead, a major acidic spot at 65–67 kDa appears in the *L. angustifolius* 2D map, completely absent in all the other protein profiles described here.

3.3. L. campestris Protein Profile

Figure 2A shows the chromatogram of the TPE from *L. campestris*. The first eluting peak (peak one) corresponds to γ-conglutin, the five partially overlapped peaks (peaks two-six) between 25 and 38 min to β-conglutin, and the highest peak (peak seven) to α-conglutin (RT 39–41 min). The collected peaks were loaded on an SDS-PAGE after reduction (Figure 2B). Similarly, in this case, 2D-electrophoresis analyses were performed on the TPE and the main eluted peaks. In the TPE (Figure 2C), a very complex mixture of polypeptides derived from these heterogeneous classes of storage proteins are visible and in total about 100 spots were detected, the major ones belonging either to β-conglutin or α-conglutin. The 2D-electrophoresis of peak one (panel 3D) shows spots related to γ-conglutin: a spot triplet at 30 kDa (large subunit) and single spots at 15 and 20 kDa (small subunits), reflecting the SDS-PAGE profile. Again, however, the identification of this protein by mass spectrometry failed.

The 2D-electrophoresis analysis of β-conglutin allowed the detection of numerous distinct polypeptides (Figure 2E), deriving from the proteolysis of the protein precursor. The majority of these polypeptides may be grouped into distinct classes, which differ in their mass to charge ratio and may be designated by increasing order of pI: acidic polypeptides characterized by pI between 4.5 and 5.5 and MW of 15 to 20 kDa; moderately acidic polypeptides, with pI between 5 and 6 and MW between 45 and 70 kDa; neutral to basic polypeptides, with pI between 6 and 8 and MW in the range 25 to 37 kDa. Unfortunately, the last class of spots is not visualized on the 2D-gel because peaks two-four were not loaded on the IPG strip due to their low protein concentration. Anyway, a clear distribution of basic spots is shown in Figure 2C. The MS analysis permitted the identification of numerous β-conglutin isoforms: conglutin beta (fragment) (B0YJF7) with 14%, conglutin beta 1 (F5B8V9) with 8.8%, conglutin beta 2 (F5B8W0) with 11.4%, conglutin beta 3 (F5B8W1) with 14.1%, conglutin beta 4 (F5B8W2) with 9.3%, conglutin beta 5 (F5B8W3) with 12.2%, conglutin beta 7 (F5B8W5) with 8.9% of A.A. coverage from *L. angustifolius*, and conglutin beta 6 (A0A6A4QEF0) with 13.7% of A.A. coverage from *L. albus*.

Finally, in panel 2F, peak seven was resolved in several spots belonging to α-conglutin, with the typical distribution already observed in other lupin species: a big cluster of spots including acidic polypeptides characterized by pI between 5 and 5.5 and MW of 37 to 55 kDa, and a row of basic spots comprising the basic polypeptides with pI from 6.5 to 9 and MW in the range 17 to 20 kDa. Similarly, in this case, the α-conglutin precursor undergoes proteolytic cleavage that produces α-polypeptides of either 52 or 44 kDa linked through disulfide bonds to a β-polypeptide of 21 kDa, typical of the mature legumin. The MS analysis allowed the identification of conglutin alpha 3 (F5B8V8) from *L. angustifolius* with a low % A.A. coverage of about 1.7%.

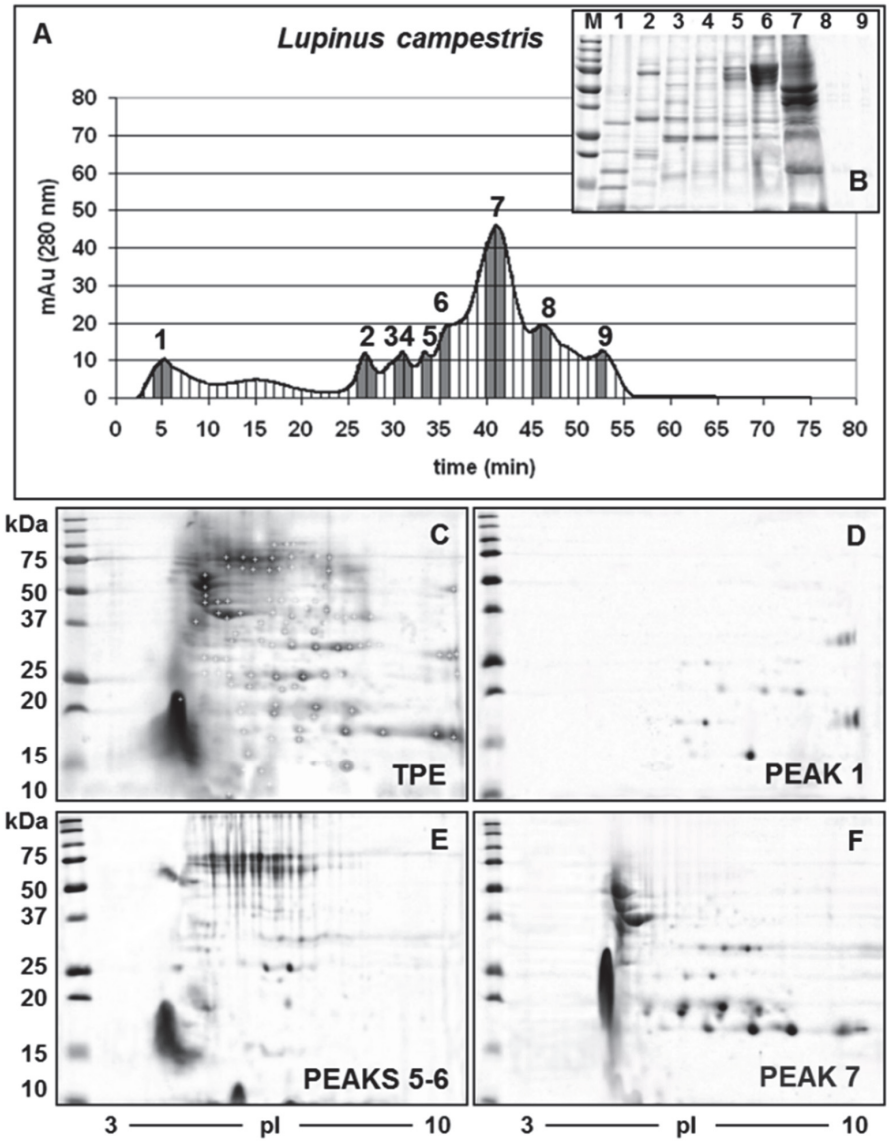

Figure 2. Electrophoretic gels of proteins of *L. campestris*. (**A**) HPLC chromatogram of the total protein extract (TPE) on a MonoQ column. (**B**) SDS-PAGE of the preparative HPLC collected fractions. (**C**) Proteome reference map of reduced TPE where in total 100 spots were detected. (**D**) Two-dimensional map of purified γ-conglutin. (**E**) Two-dimensional map of purified β-conglutin. (**F**) Two-dimensional map of purified α-conglutin. The standard marker (Precision Plus Protein) is indicated in kDa on the left.

3.4. L. aschenbornii Protein Profile

The chromatogram of the TPE from *L. aschenbornii* (Figure 3A) is slightly different since the peaks between 26 and 35 min corresponding to β-conglutin are much less resolved than in the just described species. Thus, only peak two was considered for further analysis. Another particularity is that the peak corresponding to α-conglutin (peak three) is here particularly high (RT 39–40 min). The SDS-PAGE after reduction in these fractions is shown in Figure 4B. After this preliminary analysis, each protein fraction was visualized on 2D-gels. Figure 4C shows that the 2D map of the TPE contains a very large number of distinct polypeptides, deriving from the proteolytic process of the main storage proteins. The spots detected were in total 126. The analysis of peak one (panel 4D) shows a triplet of spots at 30 kDa together with single intense spots at lower MW (17 kDa), a basic pI, corresponding to the large and small γ-conglutin subunits, respectively. On panel 3E, several isoforms of

β-conglutin polypeptides are evident, divided in two separate clusters: the first with MW in the range 50 to 75 kDa and pI values 5 to 7, the second at lower MW (around 15 kDa) and acidic pI. Since in this case only one eluted peak (peak two) was collected and analyzed, other β-conglutin isoforms with intermediate MW and pI values may have been possibly missed and consequently not visualized on the 2D-gel. The proteins identified by mass spectrometry were: conglutin beta (B8Q5G0) with 9.4%, conglutin beta (Fragment) (B0YJF7) and conglutin beta 1 (F5B8V9) with 14.5%, conglutin beta 5 (F5B8W3) with 8.3%, conglutin beta 7 (F5B8W5) with 9.4% of A.A. coverage from *L. angustifolius*, as well as conglutin beta 2 (Q6EBC1) with 14.4% and conglutin beta 6 (A0A6A4QEF0) with 8.8% of A.A coverage from *L. albus*.

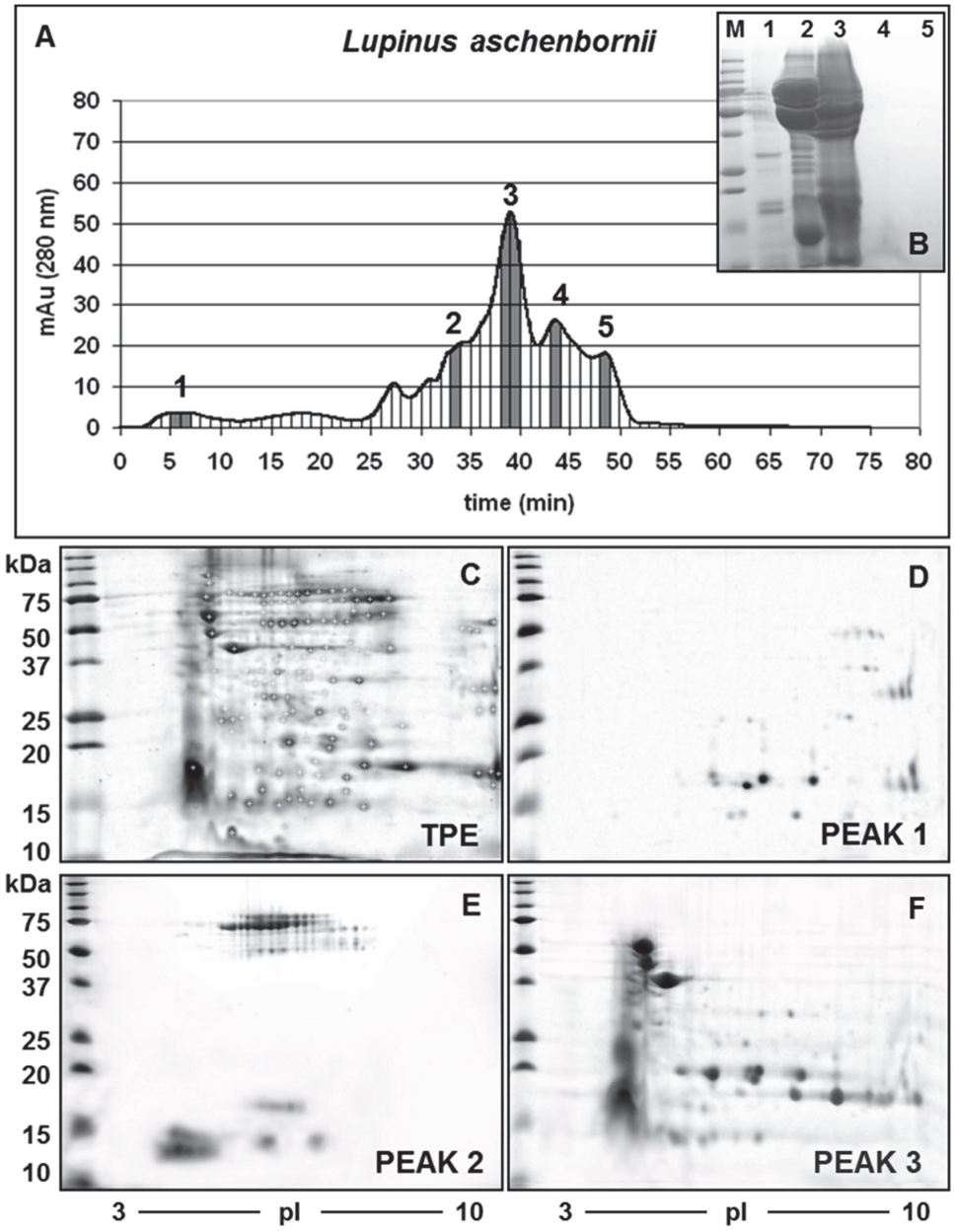

Figure 3. Electrophoretic gels of proteins of *L. aschenbornii*. (**A**) HPLC chromatogram of the total protein extract (TPE) on a Mono Q column. (**B**) SDS-PAGE of the preparative HPLC collected fractions. (**C**) Proteome reference map of reduced TPE, where in total 126 spots were detected. (**D**) Two-dimensional map of purified γ-conglutin. (**E**) Two-dimensional map of purified β-conglutin. (**F**) Two-dimensional map of purified α-conglutin. The standard marker (Precision Plus Protein) is indicated in kDa on the left.

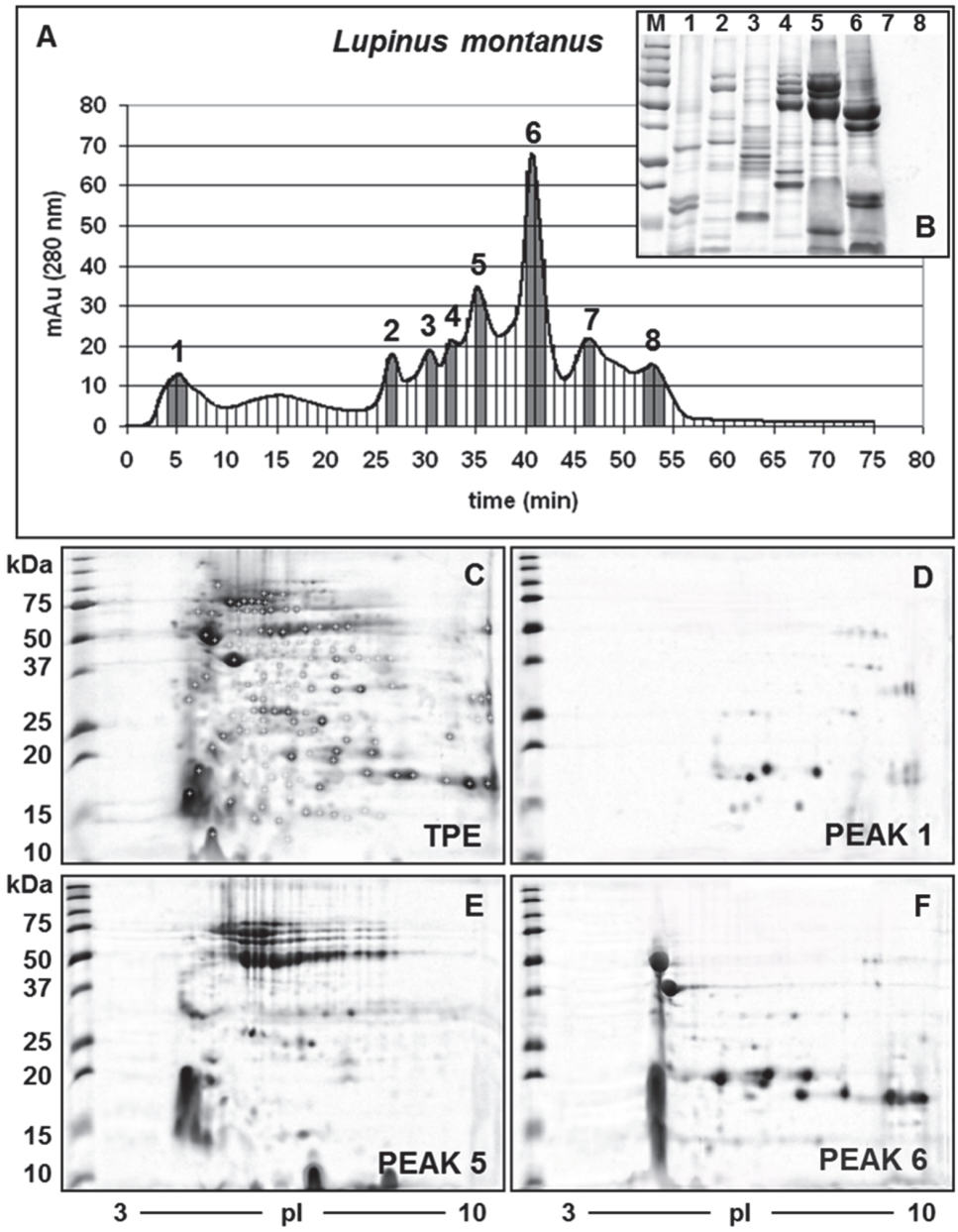

Figure 4. Electrophoretic gels of proteins of *L. montanus*. (**A**) HPLC chromatogram of the total protein extract (TPE) on a Mono Q column. (**B**) SDS-PAGE of the preparative HPLC collected fractions. (**C**) Proteome reference map of reduced TPE, where in total 134 spots were detected. (**D**) Two-dimensional map of purified γ-conglutin. (**E**) Two-dimensional map of purified β-conglutin. (**F**) Two-dimensional map of purified α-conglutin. The standard marker (Precision Plus Protein) is indicated in kDa on the left.

Panel 3F shows the distribution of the α-conglutin polypeptides resolved at distinct positions in 2D-gel: major spots belong to acidic subunit with MW from 37 to 55 and pI of 5–5.5, whereas the basic subunit displays MW in a range 15 to 20 kDa and pI of 6–8.5. The proteins identified by mass spectrometry were: conglutin alpha 1 (F5B8V6) covered by % A.A of 5.5%, conglutin alpha 2 (F5B8V7) covered by % A.A of 3.2%, and conglutin alpha 3 (F5B8V8) covered by % A.A of 4.9% from *L. angustifolius*, as well as legumin-like protein (Q53I54) covered by % A.A of 3.5% of *L. albus*. *L. aschenbornii*, showing a wide heterogeneity in alpha isoforms covered by MS/MS analysis with respect to the other wild species.

3.5. L. montanus Protein Profile

Figure 4A shows the preparative HPLC chromatogram of the TPE from *L. montanus*. In analogy with the chromatograms of the domesticated species, peak one may be attributed to γ-conglutin, the four partially overlapped peaks between 25 and 37 min (peaks two-five) to β-conglutin, and the highest peak (peak six) corresponds to α-conglutin. Each peak was collected during the elution and loaded on an SDS-PAGE gel in reducing conditions (Figure 4B). Panels 1C, 1D, 1E and 1F show the 2D-electrophoresis (in dissociating conditions) of the TPE, peak one, peak five, and peak six, respectively. The spots detected in the TPE (Figure 4C) were in total 134. The 2D map of γ-conglutin (Figure 4D) shows two series of spots around 30 kDa and between 15 and 20 kDa, belonging to the large and small subunits of this tetrameric protein. The proteolytic trimming of the terminal regions is likely a cause of the heterogeneity of the subunits [6,11]. It is important to note that in other lupin species, while the C-terminus of the protein is a unique peptide [27], three major N-terminus variants in the native protein small subunit are present. This implies that the small subunit variants should have different theoretical monoisotopic molecular masses [16] and could be submitted to many different post-translational modifications. Possibly, a similar phenomenon is responsible for the "triplets" of spots clearly visible in Figure 4D. However, although these interpretations may be reasonable in analogy to domesticated species, they remain tentative, since the identification by mass spectrometry failed, possibly owing to the lack of the correct sequences in the available databases.

Several spots corresponding to β-conglutin (Figure 4E) were identified, mostly arranged into clusters and defined as "trains of spot", with closely spaced molecular weight (MW) and pI values. Prominent spots of the mature β-conglutin subunits span from 75 to 50 kDa, whereas a group of spots around 15–25 kDa corresponds to proteolytic forms. This β-conglutin profile is similar to that of *L. angustifolius*. The heterogeneity was confirmed by LC-MS analysis, since numerous isoforms were identified with a sequence coverage of: conglutin beta (B8Q5G0) with 9.4%, conglutin beta fragment (B0YJF7) with 14.5%, conglutin beta 1 (F5B8V9) with 7%, conglutin beta 3 (F5B8W1) of 6.7%, conglutin beta 4 (F5B8W2) with 9.8%, conglutin beta 5 (F5B8W3) with 8.3%, conglutin beta 6 (F5B8W4) with 9.7%, conglutin beta 7 (F5B8W5) with 9.4% of A.A coverage based on *L. angustifolius* sequences, as well as conglutin beta 2 (Q6EBC1) with 9.9% of A.A. coverage based on *L. albus* sequences.

The α-conglutin 2D map (panel F) shows the different subunits of this hexameric protein that are resolved at distinct positions under denaturating conditions. In particular, the acidic subunit (37–50 kDa) has a pI value of 4.6–5.0, and the basic subunit (17–20 kDa) has pI values of 6.0–9.0. The fact that few isoforms of the acidic subunit are present in *L. montanus* may explain the different HPLC profile of this lupin species (see, for example, *L. hintonii*). The peptides detected by mass spectrometry belong either to conglutin alpha 1 (F5B8V6), whose % A.A coverage was 9.1%, and conglutin alpha 3 (F5B8V8), whose % A.A coverage was 2.2% of *L. angustifolius* or legumin-like protein (Q53I54) of *L. albus*, confirming the correct protein identification with the % A.A. coverage of 3.5%.

3.6. General Considerations

Despite the large heterogeneity observed among these species in what concerns the polypeptide composition of the storage proteins, they suffer a similar fate in terms of proteolytic processing and post-translational modifications (phosphorylation and glycosylation). In fact, the SDS-PAGE analyses revealed that the mature proteins are composed of numerous polypeptides differing in terms of MW and pI values, whereas the 2D-electrophoretic analyses further indicate that the individual subunits of each conglutin derive from the assembling of numerous distinct polypeptides. In particular, β-conglutin is characterized by a rather broad micro-heterogeneity that appears to be very characteristic of this protein, with almost a continuum of polypeptides ranging in molecular mass from 15 to 75 kDa and with pI values from 5 to 9. These data may be certainly useful for a deeper characterization of these wild species.

This proteomic study was helpful for screening and studying these *Lupinus* species. The applied approach was essential for characterizing the seed protein diversity, which goes far beyond gene diversity, since only at the protein level is it possible to investigate different proteolytic cleavages and diverse arrays of glycosylation forms and sites that make protein expression so heterogeneous. Knowing the protein sequences will allow directing future studies of the structure–activity relationship on different molecular targets.

4. Conclusions

In this report, we provide a detailed characterization of the major storage proteins of *L. hintonii*, *L. campestris*, *L. aschenbornii*, and *L. montanus*, underlining similarities and differences with respect to the main domesticated species *L. albus* and *L. angustifolius*. In general, it seems possible to affirm that their profiles are more similar to the latter than to the former. Through our integrated approach, proteins and peptide sequences were identified with a good degree of confidence for these wild lupin species. These pieces of information are beneficial for fundamental and applied studies concerning the storage proteins of the seeds. Moreover, to the best of our knowledge, this is the first catalog of *L. hintonii*, *L. campestris, L. aschernbornii*, and *L. montanus* storage proteins, allowing a considerable expansion of the data available for wild lupin species: this catalog may contribute to fundamental and applied studies on these seeds.

Finally, it is useful to observe that the literature presents scattered information demonstrating some interest in using some of these species in food or animal nutrition. In particular, there is interest in the exploitation of *L. campestris*, owing to its adaptation to grow at very high altitudes. Rodríguez-Ambriz et al [28] has optimized a procedure for protein purification, demonstrating that isoelectric precipitation is more effective than micellization to obtain high yields and effective separation of bitter and toxic quinolizidine alkaloids [29]. Another paper has instead investigated the possibility of using the seed of the same species to produce a yogurt-like product [30]. This evidence demonstrates that the interest for the proteins of this wild species is not only speculative.

Supplementary Materials: The following supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/d14100814/s1, Table S1: Identified proteins by LC-MS/MS.

Author Contributions: Conceptualization, A.A., E.S., K.B.T. and G.A.; investigation, K.B.T., E.S., G.B., G.A. and R.X.; validation, E.S., K.B.T., G.A. and A.A.; resources, K.B.T. and A.A. All authors have read and agreed to the published version of the manuscript.

Funding: The collection of plant material was financed with funds from CONACYT project CB 100808 and the visit to the Department of Endocrinology, Pathophysiology and Applied Biology, Faculty of Pharmacy, University of Milan, Italy, supported by the ERASMUS-MUNDUS program. The China Scholarship Council provided a fellowship to R.X.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Data are available on request.

Acknowledgments: The authors are grateful to G. M. Scigliuolo for precious help in experimental work.

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

A.A. amino acid; a.s.l.: above sea level; CHAPS, 3-[(3-cholamidopropyl)-dimethylammonio]-1-propanesulfonate hydrate; DAD, diode-array detector; DTT, dithiothreitol; IAM, 2-iodoacetamide; IEF, isoelectric focusing; LC, liquid chromatography; MS, mass spectrometry; MW, molecular weight; IPG strip, immobilized pH gradient strip; RT, room temperature; SDS-PAGE, sodium dodecyl sulfate-polyacrylamide gel electrophoresis; SPI, Score Peak Intensity; TPEs, total protein extracts.

References

1. Abraham, E.M.; Ganopoulos, I.; Madesis, P.; Mavromatis, A.; Mylona, P.; Nianiou-Obeidat, I.; Parissi, Z.; Polidoros, A.; Tani, E.; Vlachostergios, D. The Use of Lupin as a Source of Protein in Animal Feeding: Genomic Tools and Breeding Approaches. *Int. J. Mol. Sci.* **2019**, *20*, 851. [CrossRef] [PubMed]
2. Ruiz-López, M.A.; Barrientos-Ramírez, L.; García-López, P.M.; Valdés-Miramontes, E.H.; Zamora-Natera, J.F.; Rodríguez-Macias, R.; Salcedo-Pérez, E.; Bañuelos-Pineda, J.; Vargas-Radillo, J.J. Nutritional and Bioactive Compounds in Mexican Lupin Beans Species: A Mini-Review. *Nutrients* **2019**, *11*, 1785. [CrossRef] [PubMed]
3. Ramírez-Betancourt, A.; Hernández-Sánchez, A.M.; Salcedo-Morales, G.; Ventura-Zapata, E.; Robledo, N.; Wink, M.; Bermúdez-Torres, K. Unraveling the Biosynthesis of Quinolizidine Alkaloids Using the Genetic and Chemical Diversity of Mexican Lupins. *Diversity* **2021**, *13*, 375. [CrossRef]
4. Ferval, M.; Legal, L.; Gers, C.; Winterton, P.; Torres, K.B. Genomic fingerprinting versus nuclear gene sequences: A comparative approach for studying the Lupinus montanus (Fabaceae) species complex. *S. Afr. J. Bot.* **2013**, *89*, 106–110. [CrossRef]
5. Bermúdez-Torres, K.; Ferval, M.; Hernández-Sánchez, A.; Tei, A.; Gers, C.; Wink, M.; Legal, L. Molecular and Chemical Markers to Illustrate the Complex Diversity of the Genus *Lupinus* (Fabaceae). *Diversity* **2021**, *13*, 263.
6. Magni, C.; Scarafoni, A.; Herndl, A.; Sessa, F.; Prinsi, B.; Espen, L.; Duranti, M. Combined 2D electrophoretic approaches for the study of white lupin mature seed storage proteome. *Phytochemistry* **2007**, *68*, 997–1007. [CrossRef]
7. Ogura, T.; Ogihara, J.; Sunairi, M.; Takeishi, H.; Aizawa, T.; Olivos-Trujillo, M.R.; Maureira-Butler, I.J.; Salvo-Garrido, H.E. Proteomic characterization of seeds from yellow lupin (*Lupinus luteus* L.). *Proteomics* **2014**, *14*, 1543–1546. [CrossRef]
8. Wait, R.; Gianazza, E.; Brambilla, D.; Eberini, I.; Morandi, S.; Arnoldi, A.A.; Sirtori, C.R. Analysis of *Lupinus albus* Storage Proteins by Two-Dimensional Electrophoresis and Mass Spectrometry. *J. Agric. Food Chem.* **2005**, *53*, 4599–4606. [CrossRef]
9. Czubinski, J.; Grygier, A.; Siger, A. Lupinus mutabilis seed composition and its comparison with other lupin species. *J. Food Compos. Anal.* **2021**, *99*, 103875. [CrossRef]
10. Sirtori, E.; Resta, D.; Brambilla, F.; Zacherl, C.; Arnoldi, A. The effects of various processing conditions on a protein isolate from Lupinus angustifolius. *Food Chem.* **2010**, *120*, 496–504. [CrossRef]
11. Duranti, M.; Consonni, A.; Magni, C.; Sessa, F.; Scarafoni, A. The major proteins of lupin seed: Characterisation and molecular properties for use as functional and nutraceutical ingredients. *Trends Food Sci. Technol.* **2008**, *19*, 624–633. [CrossRef]
12. Shrestha, S.; Hag, L.V.; Haritos, V.S.; Dhital, S. Lupin proteins: Structure, isolation and application. *Trends Food Sci. Technol.* **2021**, *116*, 928–939. [CrossRef]
13. Mane, S.P.; Johnson, S.K.; Duranti, M.; Pareek, V.K.; Utikar, R.P. Lupin seed γ-conglutin: Extraction and purification methods—A review. *Trends Food Sci. Technol.* **2018**, *73*, 1–11. [CrossRef]
14. Brambilla, F.; Resta, D.; Isak, I.; Zanotti, M.; Arnoldi, A. A label-free internal standard method for the differential analysis of bioactive lupin proteins using nano HPLC-Chip coupled with Ion Trap mass spectrometry. *Proteomics* **2008**, *9*, 272–286. [CrossRef] [PubMed]
15. Duranti, M.; Scarafoni, A.; di Cataldo, A.; Sessa, F. Interaction of metal ions with lupin seed conglutin gamma. *Phytochemistry* **2001**, *56*, 529–533. [CrossRef]
16. Schiarea, S.; Arnoldi, L.; Fanelli, R.; De Combarieu, E.; Chiabrando, C. In-Depth Glycoproteomic Characterization of γ-Conglutin by High-Resolution Accurate Mass Spectrometry. *PLoS ONE* **2013**, *8*, e73906. [CrossRef]
17. Sousa, S.M.; Delgado, A. Leguminosas mexicanas: Fitogeografía, endemismo y orígenes. In *Diversidad Biológica de México: Orígenes y Distribución*; Instituto de biología, UNAM: Mexico City, Mexico, 1998; pp. 449–500.
18. Bermúdez-Torres, K.; Herrera, J.M.; Brito, R.F.; Wink, M.; Legal, L. Activity of quinolizidine alkaloids from three Mexican Lupinus against the lepidopteran crop pest Spodoptera frugiperda. *BioControl* **2008**, *54*, 459–466. [CrossRef]
19. Rzedowski, G.C.; Rzedowski, J. *Flora Fanerogámica del Valle de México*, 2nd ed.; Comisión Nacional para el Conocimiento y Uso de la Biodiversidad: Pàtzcuaro, Mexico, 2005; Volume 146.
20. Bradford, M. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal. Biochem.* **1976**, *72*, 248–254. [CrossRef]
21. Dooper, M.M.; Holden, L.; Faeste, C.K.; Thompson, K.M.; Egaas, E. Monoclonal antibodies against the candidate lupin allergens alpha-conglutin and beta-conglutin. *Int. Arch. Allergy Immunol.* **2007**, *143*, 49–58. [CrossRef]
22. Sirtori, E.; Resta, D.; Arnoldi, A.; Savelkoul, H.F.; Wichers, H.J. Cross-reactivity between peanut and lupin proteins. *Food Chem.* **2011**, *126*, 902–910. [CrossRef]
23. Foley, R.C.; Gao, L.-L.; Spriggs, A.; Soo, L.Y.; E Goggin, D.; MC Smith, P.; A Atkins, C.; Singh, K.B. Identification and characterisation of seed storage protein transcripts from Lupinus angustifolius. *BMC Plant Biol.* **2011**, *11*, 59. [CrossRef] [PubMed]
24. Lima-Cabello, E.; Robles-Bolivar, P.; Alché, J.D.; Jimenez-Lopez, J.C. Narrow leafed lupin beta-conglutin proteins epitopes identification and molecular features analysis involved incross-allergenicity to peanut and other legumes. *Genom. Comput. Biol.* **2016**, *2*, e29. [CrossRef]
25. Monteiro, S.; Freitas, R.; Rajasekhar, B.T.; Teixeira, A.R.; Ferreira, R.B. The Unique Biosynthetic Route from Lupinus β-Conglutin Gene to Blad. *PLoS ONE* **2010**, *5*, e8542.
26. Shutov, A.; Kakhovskaya, I.; Braun, H.; Bäumlein, H.; Müntz, K. Legumin-like and vicilin-like seed storage proteins: Evidence for a common single-domain ancestral gene. *J. Mol. Evol.* **1995**, *41*, 1057–1069. [CrossRef]

27. Scarafoni, A.; di Cataldo, A.; Vassilevskaia, T.; Bekman, E.; Rodrigues-Pousada, C.; Ceciliani, F.; Duranti, M. Cloning, sequencing and expression in the seeds and radicles of two *Lupinus albus* conglutin gamma genes. *Biochim. Biophys. Acta* **2001**, *1519*, 147–151. [CrossRef]
28. Rodríguez-Ambriz, S.L.; Martínez-Ayala, A.L.; Millán, F.; Dávila-Ortíz, G. Composition and Functional Properties of Lupinus campestris Protein Isolates. *Mater. Veg.* **2005**, *60*, 99–107. [CrossRef]
29. Boschin, G.; Annicchiarico, P.; Resta, D.; D'Agostina, A.; Arnoldi, A. Quinolizidine Alkaloids in Seeds of Lupin Genotypes of Different Origins. *J. Agric. Food Chem.* **2008**, *56*, 3657–3663. [CrossRef]
30. Jiménez-Martínez, C.; Hernández-Sánchez, H.; Dávila-Ortiz, G. Production of a yogurt-like product from *Lupinus campestris* seeds. *J. Sci. Food Agric.* **2003**, *83*, 515–522. [CrossRef]