Proteome Analysis of Pod and Seed Development in the Model Legume *Lotus japonicus*

Gitte Nautrup-Pedersen,[†,‡] Svend Dam,[†,‡] Brian S. Laursen,[§] Astrid L. Siegumfeldt,[‡] Kasper Nielsen,[||] Nicolas Goffard,[⊥] Hans Henrik Stærfeldt,[||] Carsten Friis,[||] Shusei Sato,[#] Satoshi Tabata,[#] Andrea Lorentzen,[▽] Peter Roepstorff,[▽] and Jens Stougaard*,[‡]

Centre for Carbohydrate Recognition and Signaling, and the Department of Molecular Biology, University of Aarhus, DK-8000 Aarhus C, Denmark, Center for Biological Sequence Analysis, Technical University of Denmark, DK-2800 Kgs Lyngby, Denmark, Australian Research Council Centre of Excellence for Integrative Legume Research, Genomic Interactions Group, Research School of Biological Sciences, Australian National University, Canberra, Australian Capital Territory 2601, Australia, Kazusa DNA Research Institute, Kisarazu, Chiba 292-0818, Japan, and Department of Biochemistry and Molecular Biology, University of Southern Denmark, DK-5230 Odense M, Denmark

Legume pods serve important functions during seed development and are themselves sources of food and feed. Compared to seeds, the metabolism and development of pods are not well-defined. The present characterization of pods from the model legume *Lotus japonicus*, together with the detailed analyses of the pod and seed proteomes in five developmental stages, paves the way for comparative pathway analysis and provides new metabolic information. Proteins were analyzed by two-dimensional gel electrophoresis and tandem-mass spectrometry. These analyses lead to the identification of 604 pod proteins and 965 seed proteins, including 263 proteins distinguishing the pod. The complete data set is publicly available at http://www.cbs.dtu.dk/cgi-bin/lotus/db.cgi, where spots in a reference map are linked to experimental data, such as matched peptides, quantification values, and gene accessions. Identified pod proteins represented enzymes from 85 different metabolic pathways, including storage globulins and a late embryogenesis abundant protein. In contrast to seed maturation, pod maturation was associated with decreasing total protein content, especially proteins involved in protein biosynthesis and photosynthesis. Proteins detected only in pods included three enzymes participating in the urea cycle and four in nitrogen and amino group metabolism, highlighting the importance of nitrogen metabolism during pod development. Additionally, five legume seed proteins previously unassigned in the glutamate metabolism pathway were identified.

Keywords: Legume • *Lotus japonicus* • two-dimensional gel electrophoresis • pod proteome • seed proteome

Introduction

The legume plant family contains approximately 20 000 species distributed across 700 genera,[1] of which the majority are able to form symbiosis with nitrogen-fixing rhizobia and/or phosphate-retrieving mycorrhiza. This unique feature makes legumes attractive crops, especially in regions low in nitrogen.

The fruit of a legume develops from a simple carpel into a pod that encloses and protects the seeds.[2] Evolutionarily, the pod is derived from a leaf, which is folded and joined by a dorsal suture; a ventral suture develops from the middle lamella of the leaf.[3] After seed maturation, the pods of many of the legumes dehisce and separate along the two sutures, freeing the seeds through a mechanism known as shattering. The pods are photosynthetically active until the desiccation phase when the tissue senesces and dries out.[4]

For production of food and feed in pulse crops, the legume fruits are grown to desiccation to harvest dry seeds that are stored for later consumption. However, a considerable amount of immature fruits are harvested and sold on the market as fresh, frozen, or canned food. Typical examples are the common bean (*Phaseolus vulgaris*) and the sugar pea (*Pisum sativum* var. macrocarpon). Several varieties of these crops have been bred particularly for the fleshiness, flavor, or sweetness of their pods. The flesh of such unripe pods contains lower amounts of starch and protein than mature desiccated seeds, but they are rich in vitamin C and contain pro-vitamin A.[5]

* To whom correspondence should be addressed. E-mail: stougaard@mb.au.dk.
† These authors contributed equally.
‡ University of Aarhus.
§ Present address: Danisco A/S, Edwin Rahrsvej 38, DK- 8220 Brabrand, Denmark.
|| Technical University of Denmark.
⊥ Australian National University.
Kazusa DNA Research Institute.
▽ University of Southern Denmark.

The seeds of legumes are nutrient rich, and high protein content makes them an inexpensive substitute for meat in the human diet. Additionally, some legumes produce oil-rich seeds from which oil is extracted for consumption and biofuel production. The high protein content in the seeds results from extensive accumulation of storage globulins, which are particularly rich in the nitrogen-containing amino acids glutamine and asparagine. During seed development, nitrogen can be mobilized into the seeds in the form of free amino acids. Furthermore, nitrogen-containing substances can be transported directly through the pod via the phloem and unloaded to the seed coat and then through active/passive transport into the embryo for incorporation into storage globulins. Alternatively, the nitrogen can be stored in the pod for later mobilization; hence, during progression of development the pod switches from sink to source tissue. For the pea (*Pisum sativum* L. cv Caméor), it was shown that 71% of the nitrogen accumulating in the seeds is mobilized from endogenous sources, including leaves (30%), pods (20%), roots (11%), and stems (10%).[6] During the seed filling phase, a decrease in abundant leaf proteins occurs, which illustrates this nitrogen mobilization process.[7]

Although several legume grains are rich in starch, protein, fatty acids, and micronutrients such as iron and zinc, seed improvement has been a major aim in breeding programs to further improve the nutritional value of the legume grains. Generally, seed development proceeds through three growth phases: the prestorage phase, the seed filling phase, and the desiccation phase. During the past decade, several proteomic studies of seed development have been carried out for different species. *Lotus japonicus* (*L. japonicus*)[8] and *Medicago truncatula*[9−11] are two commonly studied model legumes, and common crop plants like soybeans (*Glycine max*)[12−14] and peas (*Pisum sativum L.*)[15] have also been investigated. These studies have provided knowledge about pathways important for legume seed development, especially the synthesis of storage globulins and methionine metabolism. However, essentially no information exists on the proteome of the pod. Furthermore, a detailed analysis of the proteomes of the pods and seeds could contribute to efforts aimed at increasing the nutritional value of the immature pods marketed for human consumption.

Here we present a morphological characterization of the pods from *L. japonicus* at six developmental stages, as well as a two-dimensional (2D) gel proteomic analysis of both pods and seeds in five of these stages. In contrast to our previous GeLC−MS/MS analysis of seeds from two distinctly different developmental stages,[8] the present 2D gel proteomic analysis provides quantitative data from both pods and seeds in five consecutive developmental stages and hence represents a complementary study of *L. japonicus* both methodically and biologically. Additionally, the results from both the previous[8] and the present studies on *L. japonicus*, as well as results from *M. truncatula*[9,11] and the soybean,[12,13] are made available in a public database, accessible for comparative metabolic pathway analysis.

Materials and Methods

Plant Material. *Lotus japonicus* Gifu B-129 was grown at 70% relative humidity in a greenhouse. The plants were cultivated from December 2004 until June 2005, and the fruits were harvested at different developmental stages defined by days after flowering (DAF) from April until June 2005. Immediately after harvest, the fruits were separated into pods and seeds in

tubes on dry ice and stored at −80 C until use. For each developmental stage, fruits with similar lengths and colors were harvested, and the fruits contained 17−19 seeds. The plant material used in this study was collected in parallel with the material used by Dam et al.[8]

Characterization of the Developing Pod. The fresh weight of pods was measured immediately after harvest and the dry weight measured after drying at 65 C for 5 days. The water content was determined as the difference and expressed as a percentage of the fresh weight.

Fruits for histological studies were harvested in tubes with 70% (v/v) ethanol and stored at 4 C until use. The fruits were processed similar to a procedure described for seeds.[8]

Protein Isolation. Total protein was isolated by phenol−SDS extraction combined with several acetone and trichloroacetic acid (TCA) precipitations, as described by Wang et al.[16] Briefly, the pods or seeds were homogenized in liquid N_2 and washed with ice-cold acetone twice. After drying in a speed vac, the powder was weighed prior to further grinding in quartz sand. The powder was rinsed with 10% ice-cold TCA in acetone until the supernatant was colorless, then with aqueous 10% TCA twice, and finally twice with 80% ice-cold acetone. The pellet was dried in a speed vac prior to protein extraction. Proteins were extracted with a phenol−SDS mixture using a commercial Tris-HCl (pH 8.0) saturated phenol (AppliChem). The phenol phase was transferred to a new tube, and proteins were precipitated using five volumes of ice-cold methanol in 0.1 M ammonium acetate. Recovered pellet was washed twice with ice-cold methanol in 0.1 M ammonium acetate and twice with 80% ice-cold acetone. The final pellet was stored in 80% acetone at −80 C.

Two-Dimensional Gel Electrophoresis. For the 2D gel electrophoresis, proteins were first subjected to isoelectric focusing (IEF) and then separated according to size. To obtain high resolution, the IEF was performed on long, commercial, immobilized, pH-gradient (IPG) strips in two different pH ranges, pH 4 to 7 (Immobiline DryStrip, 24 cm, GE Healthcare) and pH 6 to 11 (Immobiline DryStrip, 18 cm, GE Healthcare), and the size separation was done on large (26 × 20 cm) 12.5% SDS PAGE.

The pellet from the total protein isolation was air-dried at room temperature and dissolved in IEF sample buffer as described by Zheng et al.[17] using 4% (w/v) CHAPS, 1% (w/v) IPG buffer (either IPG pH 4−7 or IPG pH 6−11; GE Healthcare), 10 mM Tris-HCl (pH 8.5), 10 mM dithiothreitol (DTT), and trace amounts of bromphenol blue. For pods and seeds, an amount corresponding to 25 and 10 mg of dry weight tissue powder was loaded on each strip, respectively. For each developmental stage, five 2D gels were made representing two biological samples and, for each of these, two or three technical replicates. Seeds at 7 and 13 DAF were represented with five technical replicates. The pH 4−7 strips were rehydrated with 450 μL protein samples for 21 h in a tray at room temperature, and IEF was performed at 20 C in the IPGphor II system (GE Healthcare). The pH 6−11 strips were rehydrated with 340 μL of IEF buffer without DTT and protein samples but with the addition of 12 μL mL^{-1} DeStreak (GE Healthcare), and the protein samples were anodic cup-loaded. IEF was performed at 20 C in the IPGphor II system (GE Healthcare).

Prior to the second dimension, each gel strip was reduced using DTT for 15 min and then alkylated by iodoacetamide for 15 min both at room temperature and under standard conditions.[18] For separation in the second dimension, IPG strips were

loaded on vertical 12.5% polyacrylamide (37.5:1) gels using an EttanTM DALTsix electrophoresis unit (Amersham Biosciences). Strips were overlaid with 0.5% agarose in electrophoresis buffer (25 mM Tris, 0.192 M glycine, and 0.01% (w/v) SDS). This buffer was also used in the anodic chamber of the unit, while the buffer in the cathodic chamber contained twice this concentration. The gels were run at 10 W/gel for 30 min followed by 16 W/gel for approximately 4.5 h and stopped before the dye migrated off the gels.

Gels were fixed in 50% (v/v) ethanol and 2.55% (v/v) phosphoric acid for 18 h. Subsequently, the gels were preincubated in 34% (v/v) methanol, 2.55% (v/v) phosphoric acid, and 17% (w/v) ammonium sulfate for 1 h before 350 mg L^{-1} of Coomassie Brilliant Blue G-250 was added and stained for 100 h. For destaining, the gels were washed three times with water before they were left in water for 20 h. Finally, the gels were washed two times with water prior to scanning. Image acquisition was done using a GS-800 Calibrated Densitometer scanner (BIO-RAD) with a resolution of 150 dpi, 8-bit gray scale. Afterward, gels were stored in 5% acetic acid at 4 C until use.

Analysis of the 2D Gel Images. The acquired 2D gel images were analyzed with the PDQuest Advanced software (version 8.0.1, BIO-RAD), where the gel images of biological and technical replicates were grouped in sets (each developmental stage with four or five gels in each set) prior to spot detection and spot matching across the different gels to create a master gel, which represented the spots detected in at least one of the developmental stages. The master gel was based on the image with the highest spot number and quality. The analysis was re-evaluated by manual inspection.

For quantification analysis, the spots matched to the master gel were normalized with options available within the PDQuest software. For the pods and seeds, two different methods were applied. The spots from pods were normalized using the method of local regression. For the seeds, the applied normalization method was total quantity in valid spots, and the spots containing globulins were rejected as valid spots prior to the normalization calculation. This alternative method was chosen for the seeds to avoid an excessively high contribution of the storage globulins to the overall protein intensity.

The quantification made by the PDQuest software was utilized to categorize the pod proteins into nine different expression patterns reflecting the dynamics in the protein amounts during development. The quantified data, together with their uncertainties, were fitted with a function of two straight lines, where the slope of each was compared to the uncertainty of the slope. If the uncertainty of the slope exceeded the slope, the function was considered constant. If both slopes had the same sign, then the pattern was considered continuously increasing or decreasing.

Protein Identification by Mass Spectrometry. Manually excised gel plugs were destained in water, dehydrated in 100% acetonitrile, and vacuum-dried. Proteins were in-gel digested and eluted on the MALDI target as described by Bak-Jensen et al.,[19] except that the proteins were eluted in 70% (v/v) acetonitrile and 0.1% (v/v) trifluoroacetic acid.

Peptide mass spectra were acquired by matrix-assisted laser desorption ionization mass spectrometry (MALDI) using either a 4700 Proteomics Analyzer MALDI TOF/TOF or a 4800 *Plus* MALDI TOF/TOF Analyzer (both Applied Biosystems) in the positive ion reflector mode. Depending on the sample analyzed, laser intensity and number of laser shots were varied to obtain optimal spectra. Internal apparatus calibration was performed

with trypsin-digested lactoglobulin, and MS profiles were collected in the m/z range 838−3500 Da. Ignoring MS peaks from trypsin and human keratin, the one to five most abundant peptide ions from the MS spectra were subjected to MS/MS analysis.

Both the MS and MS/MS spectra were converted into peak lists using the Data Explorer software (v4.4, Applied Biosystems) with default parameters. All spectra from a single sample were merged into one MASCOT generic file using an in-house made PERL script (courtesy of Jakob Bunkenborg, University of Southern Denmark) and utilized to search for matches in several databases using the Mascot Daemon 2.1.0 software (http://www.matrixscience.com).[20] Searches were performed with carbamidomethylation of cysteines as a fixed modification and oxidation of methionines as a variable modification. Peptide tolerance was set to 70 ppm, and one missed cleavage was accepted. The MS/MS tolerance was set to 0.5 Da. Searches were performed as in Dam et al.,[8] albeit using newer versions of the two public databases: Swiss Prot [SPROT, version 56.0 (392 667 sequences; 141 217 034 residues)] and the nonredundant proteins at the National Center for Biotechnology Information [NCBInr, version 20080826 (6 937 173 sequences; 2 395 285 404 residues)]. According to statistical analysis for the five databases (lotus_aa, lotus_nt, lotus_EST, SPROT, and NCBInr), the Mascot software assigned hits as significant ($p <$ 0.05) if the MOWSE score in the protein summary was higher than 61 for lotus_aa, 69 for lotus_nt, 65 for lotus_EST, 68 for SPROT, and 81 for NCBInr. However, MOWSE scores less than 90, 103, 97, 102, and 121 for the five databases, respectively, were manually inspected to verify the assignment by the Mascot software. Manually inspected spectra were only considered significant if they contained a peptide sequence tag of at least three consecutive b and/or y ions and all assignments based on one peptide were rejected. In total, 104 and 118 spots for pods and seeds with significant hits using Mascot were discarded using the above criteria. The analysis of some protein spots gave multiple identifications, which could not be distinguished based on the information given by the identified peptides as also discussed by Dam et al.[8]

Databases. The database "Experiment Database - Process & Analysis View" (http://www.cbs.dtu.dk/cgi-bin/lotus/db.cgi) was introduced by Dam et al.[8] and is publicly available. To support the information obtained with the present work, additional features were implemented. The selection of "Experiments" in the left side menu gives the user the possibility to choose between several proteomic experiments performed on *L. japonicus*, e.g., the 2D gel experiments presented here. By clicking one of the 2D gel experiments, a list of matching analyses appears: two 2D gel analyses together with numerous MS analyses. Choosing one of the 2D gel analyses leads to a list of spot IDs. In addition, a picture of the master gel is shown in the main field on the right, where spots are marked with turquoise plus signs. By moving a mouse across the spots in the picture, the spot ID and protein identifications pop up. To retrieve spot information, one can either click on a spot in the picture or in the list in the left-side menu. By either approach, the corresponding spot in the picture changes to a pink circle, and the experimental pI and mass are displayed together with quantification data when applicable. On the left, the accessions identified in this particular spot appear, and each one links to a page on the right side where one can choose a protein page displaying: (1) a link "KDRI" to the particular protein at the Kazusa DNA Research Institute homepage (http://www.
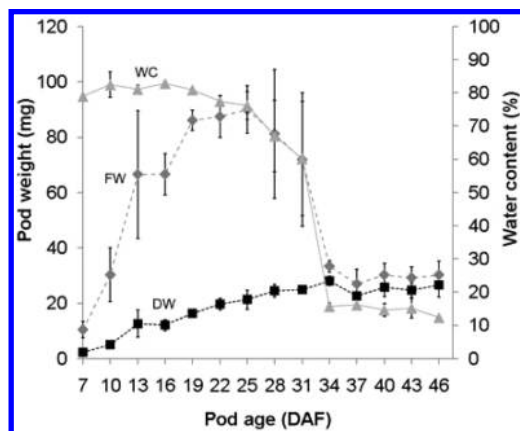
Figure 1. Pod characterization. The graph represents changes in pod fresh weight (FW), dry weight (DW), and water content (WC) from 7 to 46 DAF. FW and DW were measured for pod halves. Data represent the mean ± SD of 10 samples, each consisting of 2–6 pod halves.

kazusa.or.jp/lotus/index.html) when applicable, (2) the amino acid sequence, (3) the cDNA sequence, (4) "Experimental data for Protein *Accession*", where information is listed on whether this protein has been identified in the present and/or other proteomic analyses, and (5) a list of hits from a genomic BLAST search. Furthermore, the MS analyses are listed on the left, which are identical to the ones listed under the main menu point "Experiments", mentioned above. An MS analysis contains a peak list with experimental MS information (i.e., the matched peptides) from the different protein identifications.

The identified pod and seed proteins were Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) annotated as described by Dam et al.,[8] and the KEGG annotations were used for comparative pathway analysis. Data are available at http://bioinfoserver.rsbs.anu.edu.au/utils/PathExpress/pathexpress4legumes.php.

Results and Discussion

Characterization of the Developing Pod. The present morphological characterization of the *L. japonicus* pods, together with the detailed analyses of the proteomes of pods and seeds, paves the way for comparative pathway analysis and adds new important information on their metabolism. Prior to this study, essentially no information about the proteome of isolated pods existed since intact legume fruits containing pods and seeds together were used for EST libraries (*L. japonicus*[1,21,22]), as well as in transcriptomic (*M. truncatula*[1,21,22]) and proteomic analyses (*M. truncatula*[10]).

The development of the *L. japonicus* pod was analyzed from 7 to 46 DAF. After harvest, the fresh and dry weights were measured to determine the water content (Figure 1). Initially, the pods were green, and their maximum fresh weight was reached around 25 DAF. As they started to lose water, the color changed from green to yellow-brown. After 28 DAF, pods turned completely brown upon entering the senescence phase. Despite this, the amount of dry material increased until 34 DAF, after which a plateau was reached, as the fresh weight was only slightly higher than the dry weight (water content between 12 and 16%).

The progression of the fresh and dry weight of the pods was markedly different than that of the seeds.[8] Pods reached their maximum fresh weight a few days earlier than the seeds. Although the drying of pods was initially slower than the drying of the seeds, the pods were dry at 34 DAF compared to 37 DAF for the seeds. The slower initial drying of the pods presumably enables relocalization of nutrients from the pods to the seeds prior to desiccation.

Fruits at six developmental stages (7, 13, 19, 22, 31, and 43 DAF) were selected for histological studies. Cellular changes were visualized by toluidine blue staining of tissue sections (Figure 2). The structure of the actual wall of the *L. japonicus* pod is similar to other legume species and consists of three well-defined tissue layers: the outer exocarp, the middle
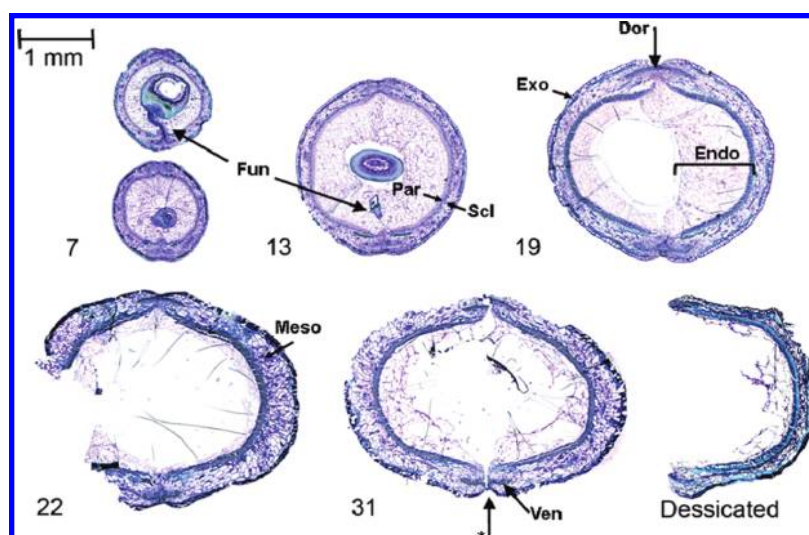
Figure 2. Histological studies of the fruit in different developmental stages. Thin sections of fruits from the five different developmental stages, 7, 13, 19, 22, and 31 DAF, together with one-half of a desiccated pod are shown. Fruits were harvested, fixed, sectioned into 5 μm slices, and stained with toluidine blue. The sections are oriented such that the ventral suture is pointing upwards and the dorsal suture downwards. For 7 DAF, one of the slices shows a seed connected to the pod via the funiculus, and in the other slice from 7 DAF and the one from 13 DAF, a part of a seed is seen inside the pod. The star indicates the weakening of the connection between the two halves of the pod at senescence. Endo, endocarp; Exo, exocarp; Meso, mesocarp; Fun, funiculus; Par, parenchyma; Scl, sclerenchyma cells; Ven, ventral sclerenchyma cords; and Dor, dorsal sclerenchyma cords.

mesocarp, and the inner endocarp. The exocarp consists of a single outer epidermis cell layer throughout the development. The mesocarp consists of several cell layers ranging from three to four layers at 7 DAF to six to seven layers at 22 DAF. During senescence, the mesocarp shrinks. Except for transition of vascular bundles (not shown), the mesocarp exhibits a uniform cell composition. The endocarp consists of three different cell types: a single layer of sclerenchyma cells, a single layer of parenchyma cells, and the inner epidermal layer surrounding the seeds. As observed in the desiccated stage, the inner epidermal layer dries out during senescence. The presence of an extended inner epidermal layer is similar to that of soybean[23] and pea[3] but different from *M. truncatula* that contains only a single layer of inner epidermal cells.[24] The connections between the two halves of the pod are tight in the earlier developmental stages; however, as indicated in Figure 2, the joining is weakened at 31 DAF leading to disconnection and shattering.

2D Gel Electrophoresis. To establish a framework for pathway analysis in legume fruits during development, we performed a comparative proteomic analysis on *L. japonicus* pods and seeds by analyzing five developmental stages, namely, 7, 13, 19, 22, and 31 DAF, which define characteristic changes during pod and seed development.[8] At each of these five developmental stages, plant material was collected and analyzed according to the methodology outlined in Figure 3. Total protein was extracted from the samples and applied to 2D gel electrophoresis. A protein extraction method utilizing trichloroacetic acid, acetone, and phenol−SDS[16] in combination with a loading buffer containing the chaotropic 2 M thiourea resulted in a high number of defined protein spots on the gels with good resolution in the molecular mass interval from 10 to 125 kDa (Figure 4). First dimensional isoelectric focusing (IEF) was performed in two different pH intervals, namely, pH 4−7 and pH 6−11. Representative master gels from each of the five developmental stages in both pH intervals are shown in Figures S1 and S2 (Supporting Information) for pods and seeds, respectively.

Total protein amounts in the two tissues during the analyzed period were strikingly different, emphasizing the senescence of the pods accompanying seed maturation. For the seeds, the protein amount was fairly constant until the last stage (31 DAF) where accumulation of storage globulins was evident (Figure S2, Supporting Information). There was a clear decline in the total amount of pod protein during development (Figure S1, Supporting Information), which correlated with decreasing intensity of several proteins as the pods approached senescence. The absolute protein amount was also much lower for the pods than for the seeds, requiring two and a half times the dry material used for the seeds to obtain equal total protein quantities on the 2D gels.

Protein Identification. In total, 1204 distinct protein spots were detected and verified on the master gels for the pods and 1773 for the seeds (Figures S1 and S2, Supporting Information). For protein identification of the individual spots, 1014 spots from pods and 1295 spots from seeds were excised. All the excised spots were numbered with a spot ID. Spots from seeds pH 6 to 11 were numbered from 1 to 598, and spots from pH 4 to 7 were given spot IDs from 2002 until 3070. For the pods, the spot ID numbers were 4001−5897 and 6001−6192 for the pH intervals 4−7 and 6−11, respectively. Excised protein spots were digested with trypsin and subjected to matrix-assisted laser desorption ionization time-of-flight tandem mass spec-
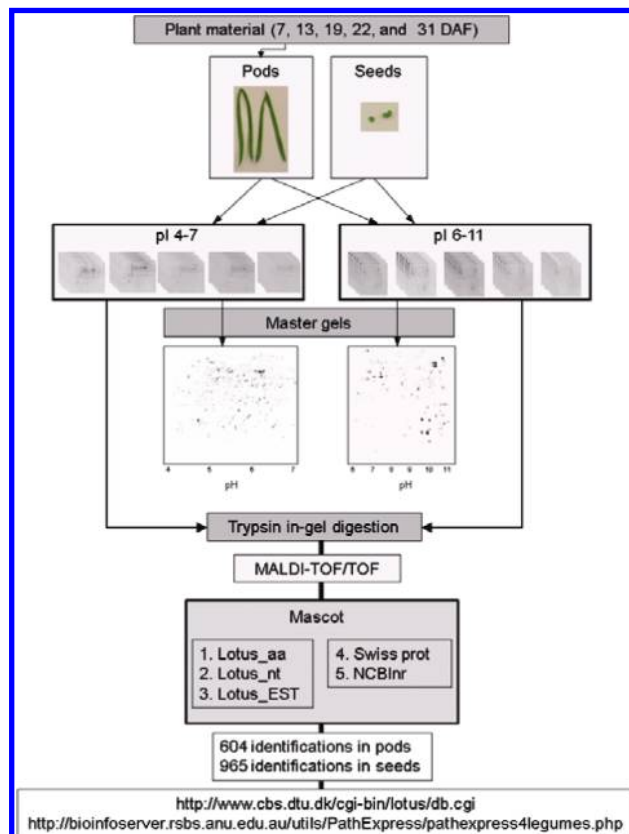
Figure 3. Proteome preparation and analysis. For each of the five developmental stages, five 2D gels were performed for each of the two pH intervals (pH 4−7 and pH 6−11). Images of the gels were collected and then matched with the PDQuest software. Spots were matched within categories from each developmental stage. Spots that were present in at least three of the replicates from a single developmental stage were utilized to produce a master gel using one of the individual gels as a template. Initially, spots were automatically matched to the master gel, and then every individual spot was verified by manual inspection. The gels were normalized and quantified. Spots were excised from the gels, digested with trypsin, and subjected to MALDI-TOF/TOF. MS results were compared against two sets of databases. Information on identified proteins was uploaded to two Web sites. The whole procedure was performed on both pods and seeds.

trometry (MALDI-TOF/TOF MS). Peptide information from MS spectral data were matched to known gene sequences using the Mascot software.[20] Initially matching the peptides to databases containing only *L. japonicus* sequences provided a high identification rate, and the public databases (SPROT and NCBInr) returned only a few additional identifications. Of the spots that were analyzed, we identified 604 pod proteins and 965 seed proteins; hence, the identification ratio for the complete analysis was 60% for pods and 75% for seeds, comparable to ratios obtained in similar studies.[9] Inspection of a subset of 63 samples producing high-quality MS spectra revealed 50 significant identifications in the lotus_aa and lotus_nt databases. Thus, we estimated that the genomic sequences covered approximately 80% of the proteome, which is close to the estimated 91% coverage of the transcriptome previously reported by Sato et al.[25] The total number of spots, the number of isolated spots, and identified proteins in the 2D gels are summarized in Table 1.
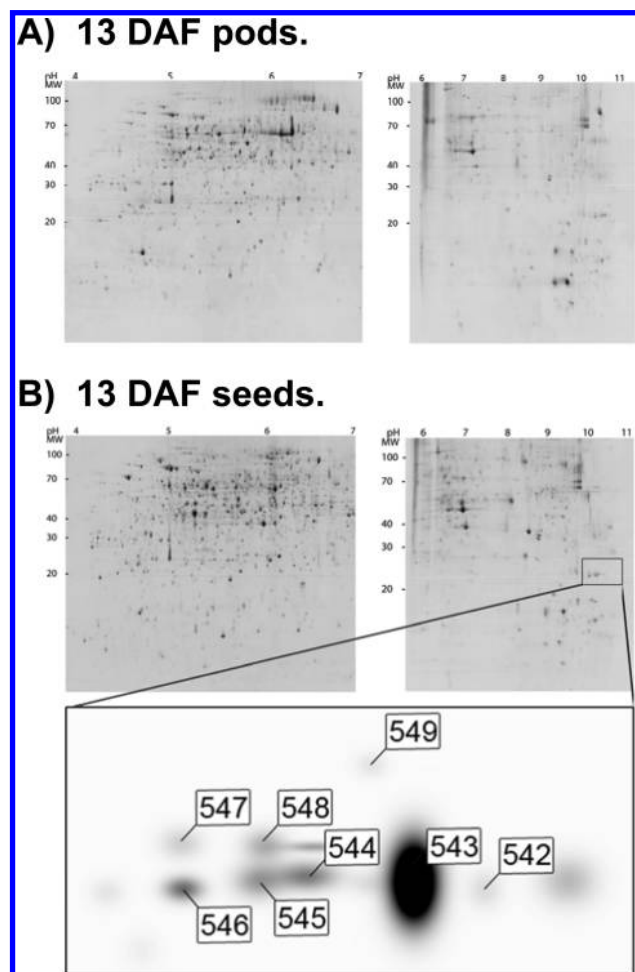
Figure 4. Pod and seed proteomes at 13 DAF. These 2D gels are representative images of the colloidal Commassie Brillant Blue stained gels from 13 DAF. (A) Gel from pod proteome analysis at 13 DAF. (B) Gel from seed proteome analysis at 13 DAF. The enlarged portion shows the corresponding segment from the master gel. The highlighted spots correspond to protein translation elongation factor 1-alpha (spots 128, 527, 562, and 563) and ribosomal proteins (40S S7 [spots 542 and 543], 40S S5 [spots 544, 545, 546, and 548], and 60S L9 [spot 553]). These were visible only when thiourea was included in the loading buffer. Gels on the left are from pH 4−7, and the gels on the right are from pH 6−11.

Table 1. Number of Spots Isolated and Proteins Identified in the 2D Gel Analysis

	pods	seeds
total number of spots	1204	1773
isolated spots	1014	1295
identified spots[a]	604	965
total protein identifications[a]	1075	1480
different identifications	567	738
intersection		304
uniprot homologues[b]	550	699
EC assignments[c]	289	344
enzymatic functions	138	159

[a] The number of total protein identifications exceeds the number of identified spots as some peptides match equally well to different gene accession numbers. [b] Protein homologues identified with the UniProt Knowledgebase. [c] EC: enzyme commission.

Protein identifications from the pods are listed according to their spot ID number in Table S1 (Supporting Information).

Accompanying each spot ID, the matching gene accessions are listed together with experimental and analytical data. The total number of pod protein identifications was 1075, which corresponded to 567 different gene accessions (Table S2, Supporting Information). Similarly, the results for seeds are presented in the Tables S3 and S4 (Supporting Information) with 1480 identifications corresponding to 738 different gene accessions. The intersection between the identifications in the pods and seeds constitutes 304 gene accessions.

A previous study on the *L. japonicus* seed proteome identified proteins corresponding to 1113 different gene accessions.[8] In this study, we independently identified proteins corresponding to 738 different accessions. The two different approaches contained a common set of proteins representing 341 accessions. Thus, the 2D gel approach described here adds 397 new identifications. The complete number of identified proteins in one or more developmental stages of *L. japonicus* seeds corresponds to 1510 accessions.

To predict biological functions for the obtained gene accessions, sequences were blasted against the UniProt Knowledgebase (UniProtKB)[26,27] using the probability value $e \leq 1 \times 10^{-6}$, and the obtained protein functions are listed in Table S5 (Supporting Information), where enzymatic functions are denoted by their Enzyme Commission annotation (EC number).[28] For the pods, a biological function was predicted for 550 of the 567 different gene accessions. Similarly, 699 of the 738 different seed gene accessions identified had a homologue in the UniProtKB.

Proteins Identified in the Pods. Approximately half of the 550 gene accessions in the pods that were assigned a biological function through their UniProtKB homologue had a predicted enzymatic function. The other half (256 identifications) included ribosomal subunits, protein translation factors, structural proteins such as actin and tubulin, structural integral membrane proteins (e.g., ion channels), RNA and DNA binding proteins (U6 snRNA, DNA repair protein, and RNA binding protein 8A), annexins, photosynthesis-associated proteins, protein inhibitors such as trypsin inhibitor and polygalacturonase inhibitor, nuclear transport-related proteins such as RAN GTPase and importin alpha, chaperones and heat shock proteins, ATP synthases, and transcription factors BTF3 and Pur-alpha 1.

In our analysis, there were 263 proteins that were only detected in the pod. Of these, 255 identifications could be assigned to a homologue in the UniProtKB, and 114 were assigned with an EC annotation, leading to the prediction of 50 different enzymatic functions. The 141 protein identifications without an assigned enzymatic function were evenly distributed throughout several different biological functions.

On the basis of the identifications from *L. japonicus* pods, the PathExpress interface (see below) allows pathway analysis of pods and comparison of pod and seed identifications. The 289 pod accessions that were assigned an enzymatic function represented 138 different enzymatic functions in 85 various pathways (summarized in Table S6, Supporting Information). Carbon fixation, glycolysis/gluconeogenesis, and pyruvate metabolism were prominent among these pathways. For each of these three pathways, only one or two assigned enzymatic functions were identified in pods and absent in seeds.

By comparing the identifications in pods with both our new identifications in *L. japonicus* seeds and all the previous identifications in other legume seed studies (*L. japonicus*,[8] *M. truncatula*,[9−11] and soybean[12,13,29]), we found 45 assigned
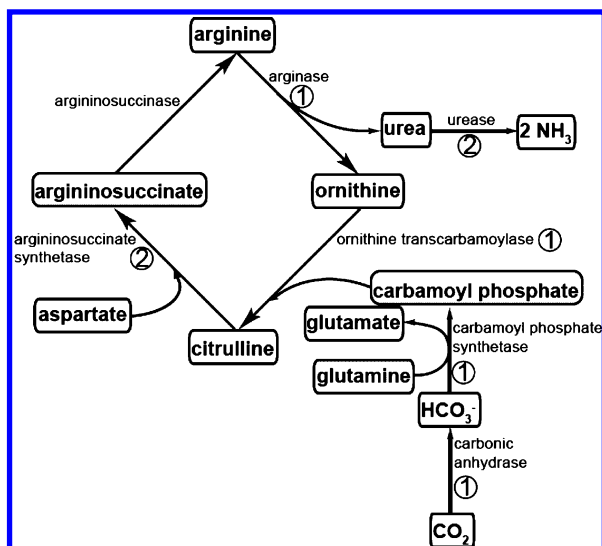
Figure 5. Comparative analysis of the urea cycle in pods and seeds. The bicarbonate ions (HCO_3^-) are produced by the enzyme carbonic anhydrase, and ammonium ions (NH_4^+) can be produced by deamination. After relocalization to the seeds, urea can be hydrolyzed by urease to produce ammonium, which can then be assimilated into amino acids. (1) Protein identifications in *L. japonicus* pods and (2) protein identifications in legume seeds.[8,9,11−13]

enzymatic functions that were characteristic for pods. In total, they represent 32 different metabolic pathways (Table S7, Supporting Information). Interestingly, "nitrogen metabolism" and "urea cycle and metabolism of amino groups" had three and four distinctive pod proteins, respectively.

In the urea cycle alone (Figure 5), ornithine transcarbamoylase (spot 4022), arginase (spot 5403), and glutamine-dependent carbamoyl-phosphate synthetase (spot 4426) are found in pods but not in seeds. Carbamoyl is donated to the urea cycle as carbamoyl-phosphate, which is synthesized by carbamoyl-phosphate synthetase using bicarbonate and ammonium ions. The bicarbonate ions are produced by the enzyme carbonic anhydrase, which was identified in the pods but not in seeds, and ammonium is obtained from deamination of glutamine. Ornithine transcarbamoylase was present at more than a 10-fold higher concentration during the later three stages of pod development (19, 22, and 31 DAF) than at 7 DAF. Arginase showed a high protein level throughout pod development. Taken together, this data alludes to the differential expression of proteins in the urea cycle between pods and seeds. After relocalization to the seeds, urea can be hydrolyzed by urease, a protein found in seeds (soybean[12,13,29]). Ammonium derived from the degradation of urea can then be assimilated into amino acids. Hence, urea may serve as a transfer unit for nitrogen from pods to seeds during maturation of the protein-rich legume seeds.

Further emphasizing the importance of active nitrogen metabolism in pods, two other proteins distinguishing pods were predicted to be enzymes active in nitrogen transfer: formamidase (spot 4446), which catalyzes the conversion of formamide to formate and ammonia, and the long acyl aminoacylase (spot 4569), which catalyzes an *N*-acyl-L-amino acid to a carboxylate and L-amino acid.

Dynamics of Pod Proteins. The quantification by the PDQuest software was used for a detailed assessment of the tendencies in the dynamics in protein amounts during pod

development. There were 411 identified protein spots with quantification from the 2D gel analysis of the pods. The progressions of the intensities of these spots were categorized to nine patterns as shown in Table 2. As expected, the majority of the proteins decreased in abundance during development and belonged to categories VII, VIII, and IX. For example, a 60-fold decreased intensity was found for IAA-amino acid hydrolase ILR1-like 4 (spot 4248), which is probably involved in the activation of auxin (indole-3-acetic acid, IAA) by hydrolyzing an amino acid conjugated IAA. Another protein showing decreasing intensity was the general transcription factor BTF3 (spot 4080), which decreased more than 16-fold during development.

The proteins that increased in intensity during development are particularly interesting. A 3-ketoacyl-CoA thiolase (spot 6150) belongs to category III with an increasing protein level until 19 DAF followed by a plateau for the remainder of the development. This enzyme catalyzes a key step in fatty acid β-oxidation breakdown and mobilization of lipids from stores of triacylglycerol in the pod.

Endopeptidase cucumisin belonging to the subtilisin family was identified in two of the spots from category IV (spots 6132 and 6134). The highest intensities of these spots were at 22 and 19 DAF, respectively. Cucumisin has only been identified in pods and not in any of the proteomic studies on legume seeds[8,9,11−13] indicating that cucumisin most likely plays an important role in protein degradation during pod senescence. In addition, peroxisomal glycolate oxidase (spots 6087, 6088, and 6089; the sum of their intensities would belong to the category IV with peak intensity at 13 DAF), ornithine carbamoyltransferase (spot 4022, category IV with peak intensity at 19 DAF), and Class-10 pathogenesis-related protein 1 (spot 4131, category III and constant after 22 DAF) were also only found in pods.

Pod Senescence. The pods dry out via a particular type of programmed cell death called senescence, which can be observed at various levels of biological organization ranging from individual cells to tissues, organs, and entire plants. Senescence in the *L. japonicus* pods was evident at 25 DAF, where pods changed color from green to yellow-brown, and 28 DAF, where they turn completely brown and dry out (data not shown).

As pods are evolutionarily derived from leaves, it is tempting to compare the pods with drying leaves. Leaf senescence is a slow process, and cell components are degraded while nutrients are mobilized and redirected from the dying cells to other parts of the plant. For the pods, nutrients are most likely relocated to the seeds (e.g., chickpea,[30] pea,[6,7,31] and bean[32,33]).

At the onset of senescence, the chloroplasts are degraded, while mitochondria and nuclei remain structurally and functionally intact until a later stage. A reduction in the protein level of the key stromal enzyme RuBisCO (EC4.1.1.39) has been associated with senescence, and its degradation is triggered by reactive oxygen species, such as ozone, already before visible symptoms of senescence (reviewed by Hortensteiner and Feller[34]).

In this study, we have identified both the RuBisCO small subunit (spots 6097 and 6098) and large subunit (spot 4549), together with RuBisCO-associated proteins such as RuBisCO activase (spots 4329, 4624, 4633, and 5874) and RuBisCO subunit binding protein alpha (spots 4334, 4339, and 5708) or beta (spots 4257, 4342, and 5862). During pod development, the majority of these proteins decreased in abundance (see

**Table 2.** Nine Categories of Expression Patterns during Pod Developmental Progression

Category	Expression pattern		Number of spots
I	/	Increasing	35
II	_/	Constant then increasing	9
III	/‾	Increasing then constant	19
IV	∧	Hill	64
V	—	Constant	49
VI	∨	Valley	36
VII	_	Decreasing then constant	44
VIII	‾\	Constant then decreasing	29
IX	\	Decreasing	126
Total			411

Figure 6A for the intensities of selected spots). Additionally, other proteins assigned as enzymes acting in the Calvin cycle were identified. Particularly, glyceraldehyde 3-phosphate dehydrogenase (EC 1.2.1.9) was identified in 19 distinct spots of which the protein level decreased during development for 12 spots (spots 6009, 4027, 6043, 4037, 6023, 4188, 4366, 4035, 6186, 6024, 5551, and 6179) and increased for seven spots (spots 6018, 6008, 5546, 6022, 4054, 6016, and 4026). Such distribution in several spots of identifications annotated to the same enzyme might be due to different post-translational modifications.

Other proteins from the photosystems (spots 6002, 6102, 6110, and 4324) and several chloroplastic proteins (spots 4140, 4097, 6078, 4103, and 5062) decreased in abundance. Further, structural components such as actin (spot 5203) also decreased, and the amount of the ferredoxin-dependent enzyme sulfite reductase (EC 1.8.7.1, spot 6068) and the chloroplastic ferredoxin-NADP reductase (EC 1.18.1.2, spots 4033, 5554, 6019, and 4024) showed reduction in their protein level. These results reflect the degradation of the photosynthetic machinery during the pod senescence.

The plastidial glutamine synthetase (EC 6.3.1.2) decreased during senescence (spots 4236 and 5227) and was earlier shown to be degraded even faster than RuBisCO.[35,36] A cytosolic counterpart of glutamine synthetase was detectable longer and seemed to increase in intensity (spot 5234), possibly reflecting that the cytosolic glutamine synthetase could be responsible for nitrogen recycling from the senescing tissue. The synthesized glutamine amino acids could relocate to the seeds.

During development, there was a general decrease in the intensity of proteins involved in protein biosynthesis. All the identified ribosomal subunits decreased in abundance (spots 6107, 6117, 6116, 6113, 6111, 6115, 6121, 6077, 6118, 4028, 5017, 6126, 6060, 6119, 6114, 6127, 6109, 6120, 6125, 6124, and 5020; see Figure 6B for the intensities of selected ribosomal subunits). There was also a lower level of several protein translation factors including translation initiation factor 5A (spots 5834, 5830, and 5829; see Figure 6C for their added intensities), initiation factor 4A (spot 4251), initiation factor 3G (spot 6073), initiation factor eIF2 (spot 4282), elongation factor 1 alpha

(spots 6138 and 6136) and 1 beta (spots 4303 and 4300), chloroplast elongation factor Tu (spot 5062), and the poly(A)-binding protein (spots 4157, 6041, and 6040). In contrast to these proteins, two spots representing the glutathione-S-transferases (EC 2.5.1.18, spots 4265 and 4267) increased in abundance. This might imply a function for these enzymes in protecting the plant against reactive oxygen radicals generated during peroxisomal lipid degradation at senescence.

Pod Shattering. For many seed-producing plants, pod senescence ends with pod shattering, which is important in nature as a seed dispersal mechanism but generally unwanted in crops. In fact, many crop plants are bred against this ability.[37−40] During pod shattering, the two halves of the pod detach due to a combination of the cell wall loosening in the dehiscence zone, which is a region no more than a few cells wide extending the entire length of the fruit at the boundaries along the two sutures, and the tensions established by the differential mechanical properties of drying cells in the pod.[41] At maturity, the dehiscence zone consists of a nonlignified separation layer.

A variety of biological activities are required in advance of pod shattering. First, early regulators of cell differentiation must act to mediate cell fate specification, and then several enzymatic activities must work to accomplish the associated processes, such as changes in cell wall composition, lignification, and disintegration of the middle lamella in the separation layer.[42] Polygalacturonase (EC 3.2.1.15) is known to degrade the major component of pectin and has been shown to be involved in shattering.[43−46] We have identified a putative polygalacturonase inhibitor (spot 6086), although it apparently does not change in abundance during the period analyzed (Figure 6D).

Storage Globulins and Late Embryogenesis Abundant Protein. Surprisingly, the proteins identified in both pods and seeds included both a storage globulin and a late embryogenesis abundant (LEA) protein. A putative pod storage globulin was identified from spot 4418, which is visible as early as 7 DAF. The peptides from spot 4418 matched two common *Lotus* legumin storage proteins: LLP2 and LLP3.[8] The significance of
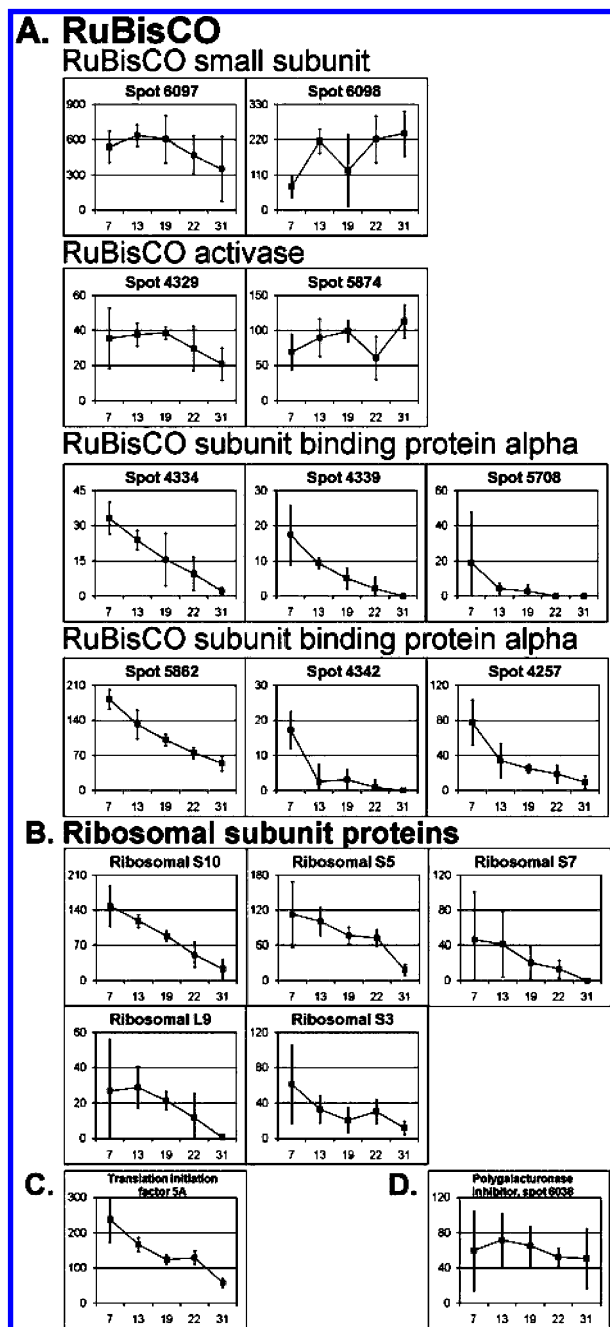
Figure 6. Expression patterns of selected identified pod proteins. (A) RuBisCO small subunit, RuBisCO activase, RuBisCO subunit binding protein alpha and beta. (B) Selected ribosomal subunits, where intensities for spots containing identical proteins were added, and the combined graphs are shown. Ribosomal subunit S10 (spots 6107, 6113, and 6114), ribosomal subunit S5 (spots 6117, 6118, 4028, 6119, and 6120), ribosomal subunit S7 (spots 6116 and 6115), ribosomal subunit L9 (spots 6121, 6122, and 6123), and the ribosomal subunit S3 (spots 6126, 6127, 6125, and 6124). (C) Translation initiation factor 5A (combination of spots 5834, 5830, and 5829). (D) Polygalacturonase inhibitor. The graphs show the intensity (OD*area) as a function of the developmental stages (DAF). Error bars indicate ± SD for five replicates.

the MOWSE scores[20] clearly document that this globulin was present in the pods and was not exclusively present in the seeds during the seed filling phase, as expected. The identification of a storage globulin in the pods at this early developmental stage prior to accumulation of storage globulins in the seeds suggests a role for pods in intermediate amino acid storage. A LEA protein was identified in the pod (spot 4288). The intensity of this spot decreased during the developmental stages (Figure S3, Supporting Information), which is in direct contrast to the pattern observed in seeds.

Glutamate Metabolism in Seeds. Comparative analysis of metabolic pathways in legume seeds performed with the PathExpress tool showed that the present study contributes several new assignments to glutamate metabolism. Glutamate plays a central role in assimilation and dissimilation of ammonia to other amino acids and is the precursor for arginine, proline, and γ-aminobutyric acid. In the previous GeLC−MS/MS study of *L. japonicus*,[8] three different enzymes were assigned to glutamate metabolism. The present 2D gel study predicts another seven enzymes (EC numbers) to this pathway. Summarizing the legume seed studies, 11 proteins with putative glutamate metabolism-related functions have been identified. Five of these were exclusively found in this study, namely, alanine amino transferase, γ-glutamylcysteine synthetase, glutathione reductase, glutamyl tRNA synthetase, and glutamate decarboxylase. At the homepage http://bioinfoserver.rsbs. anu.edu.au/utils/PathExpress/pathexpress4legumes.php, a complete record of enzymes involved in the glutamate metabolism can be found, as well as information about which enzymes have been identified in each study. Figure 7 shows the relatively small segment of the pathway where the five predicted enzymes, which were exclusively found in the present analysis, occur.

Experiment Database and the Database for Comparative Legume Proteomics. For further analysis, all the proteome data obtained were collected in two publicly available databases. The experimental data for both the pods and the seeds are available in the database "Experiment Database - Process & Analysis View" (http://www.cbs.dtu.dk/cgi-bin/lotus/db.cgi). This database also contains protein sequence information for all published *L. japonicus* sequences.[25] The options provided by the database are further described in the Material and Methods section. Briefly, the database offers the possibility to navigate via the spots on the master gels, to extract experimental data for individual spots, and to obtain links leading to DNA and protein sequences. Furthermore, the database encompasses the experimental data from the previous study in *L. japonicus*[8] and provides information about the study that identified a particular gene accession.

The seed proteome data from several legume studies, *L. japonicus*,[8] *M. truncatula*,[9,11] and soybean,[12,13] together with the new data for the *L. japonicus* pods and seeds was uploaded to the Web-based tool PathExpress (http://bioinfoserver.rsbs. anu.edu.au/utils/PathExpress/pathexpress4legumes.php), where comparative metabolic pathway analysis is possible. The analysis is based on assignment of enzymatic functions to the amino acid sequences using the EC nomenclature and extraction of information about metabolic networks using the KEGG LIGAND database (http://www.genome.jp/ligand/).[47,48] Only sequences that were assigned a complete EC number (all four levels of classification) were included.[49]

The PathExpress tool provides access to 567 identifications for pods together with more than 2300 for seeds. This difference in numbers reflects that more enzymatic functions were assigned in seed metabolic pathways than in pod pathways. Enzymatic functions identified only in pods may therefore indicate real differences between pod and seed metabolism.
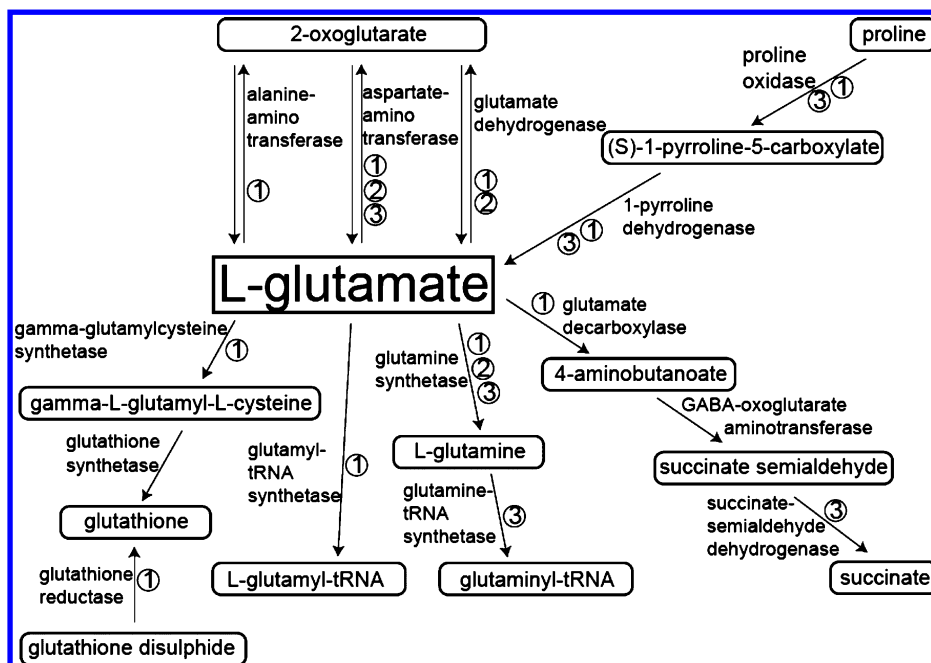
Figure 7. Glutamate metabolism in legume seeds. A part of the glutamate metabolism pathway showing glutamate as the substrate for several enzymatic reactions. At each of the arrows, the active enzyme is indicated. If the particular enzyme was identified in a legume seed proteomic study, the analysis is indicated with numbers: (1) the present 2D gel study in *L. japonicus*; (2) GeLC−MS/MS analysis in *L. japonicus*;[8] and (3) 2D gel analysis in soybean.[12,13]

For example, pods and seeds seem to utilize different strategies in nitrogen metabolism. Pods relocate the nitrogen, while seeds store the nitrogen for later use in germination, which were reflected by the major differences found for expression levels of enzymes participating in the urea cycle and in other aspects of nitrogen metabolism as mentioned above.

In combination, the legume seed studies provided 2341 identifications, of which 1741 proteins were assigned a function via the UniProtKB. Using the PathExpress interface, 789 accessions were assigned to 247 different enzymatic functions. In total, predicted enzymes from 103 pathways were identified, and for 42 pathways, five or more predicted enzymes were identified. None of the 103 identified pathways are completely covered, but comparing the identifications obtained in this study with the entire data set leads to 41 assigned enzymatic functions not previously identified in legumes (i.e., 17% are exclusively identified through this present study and contribute to increased knowledge of seed development in legumes). For example, with the present identification of sedoheptulose-1,7-biphosphatase (EC 3.1.3.37, spots 2087 and 2089) and phosphopentose epimerase (EC 5.1.3.1, spot 2465), 10 of the 11 assigned enzymes acting in the Calvin cycle, which is part of the carbon fixation pathway, were identified in legume seeds. However, the remaining assigned enzyme, phosphopentose isomerase (EC 5.3.1.6), so far unidentified in seeds, was identified in the pod proteome (spot 4158).

Conclusion

The results from our proteomic study on pods and seeds from the model legume *L. japonicus* are publicly available in two databases for additional studies on the comparative metabolic differences between the two tissues during development. Here we have focused on nitrogen metabolism in pods and seeds and also shown that typical seed proteins, such as storage globulins and LEA, were present in pods. The progres-

sion for a number of specific proteins in the pods (e.g., RuBisCO, chloroplastic proteins, and translation factors) agreed with previous observations for senescing leaves, and a further comparison with biological processes in dying leaves would be interesting. More in-depth studies could include dissection of tissue parts by laser capture, where isolation of the dehiscence zone from the pods would be particularly interesting.

Supporting Information Available: Seven supplemental tables and three supplemental figures. Figure S1: 2D gel electrophoresis of pods. Figure S2: 2D gel electrophoresis of seeds. Figure S3: The developmental changes in intensity for spot 4288 (LEA). Table S1: Protein identifications in pods. Table S2: Peptide identifications in pods. Table S3: Protein identifications in seeds. Table S4: Peptide identifications in seeds. Table S5: Protein function analysis. Two sheets are available, for pods and seeds, respectively. Table S6: Metabolic pathway analysis in *L. japonicus* pods. Table S7: Metabolic pathways with enzymatic functions identified in the pods and not in the seeds. This material is available free of charge via the Internet at http://pubs.acs.org.

References

(1) Firnhaber, C.; Pühler, A.; Küster, H. EST sequencing and time course microarray hybridizations identify more than 700 *Medicago truncatula* genes with developmental expression regulation in flowers and pods. *Planta* **2005**, *222* (2), 269–83.
(2) We have adopted the definition that the legume fruit consists of a pod and seeds. Note, however, that other definitions are used

in the literature, where for instance fruit is defined as a pod consisting of a pod wall and the seeds, for example.

(3) Ozga, J. A.; van Huizen, R.; Reinecke, D. M. Hormone and seed-specific regulation of pea fruit growth. *Plant Physiol.* **2002**, *128* (4), 1379–89.

(4) Atkins, C. A.; Kuo, J.; Pate, J. S.; Flinn, A. M.; Steele, T. W. Photosynthetic Pod Wall of Pea (*Pisum sativum L.*): Distribution of Carbon Dioxide-fixing Enzymes in Relation to Pod Structure. *Plant Physiol.* **1977**, *60* (5), 779–786.

(5) USDA Nutrient Database, USDA National Nutrient Database for Standard Reference: http://www.nal.usda.gov/fnic/foodcomp/search/, accessed 16 April, 2010.

(6) Schiltz, S.; Munier-Jolain, N.; Jeudy, C.; Burstin, J.; Salon, C. Dynamics of exogenous nitrogen partitioning and nitrogen remobilization from vegetative organs in pea revealed by 15N in vivo labeling throughout seed filling. *Plant Physiol.* **2005**, *137* (4), 1463–73.

(7) Schiltz, S.; Gallardo, K.; Huart, M.; Negroni, L.; Sommerer, N.; Burstin, J. Proteome reference maps of vegetative tissues in pea. An investigation of nitrogen mobilization from leaves during seed filling. *Plant Physiol.* **2004**, *135* (4), 2241–60.

(8) Dam, S.; Laursen, B. S.; Ørnfelt, J. H.; Jochimsen, B.; Stærfeldt, H. H.; Friis, C.; Nielsen, K.; Goffard, N.; Besenbacher, S.; Krusell, L.; Sato, S.; Tabata, S.; Thøgersen, I. B.; Enghild, J. J.; Stougaard, J. The proteome of seed development in the model legume *Lotus japonicus*. *Plant Physiol.* **2009**, *149* (3), 1325–40.

(9) Gallardo, K.; Le Signor, C.; Vandekerckhove, J.; Thompson, R. D.; Burstin, J. Proteomics of *Medicago truncatula* seed development establishes the time frame of diverse metabolic processes related to reserve accumulation. *Plant Physiol.* **2003**, *133* (2), 664–82.

(10) Watson, B. S.; Asirvatham, V. S.; Wang, L.; Sumner, L. W. Mapping the proteome of barrel medic (*Medicago truncatula*). *Plant Physiol.* **2003**, *131* (3), 1104–23.

(11) Gallardo, K.; Firnhaber, C.; Zuber, H.; Héricher, D.; Belghazi, M.; Henry, C.; Küster, H.; Thompson, R. A combined proteome and transcriptome analysis of developing *Medicago truncatula* seeds: evidence for metabolic specialization of maternal and filial tissues. *Mol. Cell. Proteomics* **2007**, *6* (12), 2165–79.

(12) Hajduch, M.; Ganapathy, A.; Stein, J. W.; Thelen, J. J. A systematic proteomic study of seed filling in soybean. Establishment of high-resolution two-dimensional reference maps, expression profiles, and an interactive proteome database. *Plant Physiol.* **2005**, *137* (4), 1397–419.

(13) Agrawal, G. K.; Hajduch, M.; Graham, K.; Thelen, J. J. In-depth investigation of the soybean seed-filling proteome and comparison with a parallel study of rapeseed. *Plant Physiol.* **2008**, *148* (1), 504–18.

(14) Krishnan, H. B.; Oehrle, N. W.; Natarajan, S. S. A rapid and simple procedure for the depletion of abundant storage proteins from legume seeds to advance proteome analysis: a case study using *Glycine max. Proteomics* **2009**, *9* (11), 3174–88.

(15) Bourgeois, M.; Jacquin, F.; Savois, V.; Sommerer, N.; Labas, V.; Henry, C.; Burstin, J. Dissecting the proteome of pea mature seeds reveals the phenotypic plasticity of seed protein composition. *Proteomics* **2009**, *9* (2), 254–71.

(16) Wang, W.; Scali, M.; Vignani, R.; Spadafora, A.; Sensi, E.; Mazzuca, S.; Cresti, M. Protein extraction for two-dimensional electrophoresis from olive leaf, a plant tissue containing high levels of interfering compounds. *Electrophoresis* **2003**, *24* (14), 2369–75.

(17) Zheng, X.; Hong, L.; Shi, L.; Guo, J.; Sun, Z.; Zhou, J. Proteomics analysis of host cells infected with infectious bursal disease virus. *Mol. Cell. Proteomics* **2008**, *7* (3), 612–25.

(18) Amersham Biosciences AB, *Ettan DIGE System, User Manual*; 2003.

(19) Bak-Jensen, K. S.; Laugesen, S.; Østergaard, O.; Finnie, C.; Roepstorff, P.; Svensson, B. Spatio-temporal profiling and degradation of alpha-amylase isozymes during barley seed germination. *FEBS J.* **2007**, *274* (10), 2552–65.

(20) Perkins, D. N.; Pappin, D. J.; Creasy, D. M.; Cottrell, J. S. Probability-based protein identification by searching sequence databases using mass spectrometry data. *Electrophoresis* **1999**, *20* (18), 3551–67.

(21) Asamizu, E.; Nakamura, Y.; Sato, S.; Tabata, S. Characteristics of the *Lotus japonicus* gene repertoire deduced from large-scale expressed sequence tag (EST) analysis. *Plant Mol. Biol.* **2004**, *54* (3), 405–14.

(22) Krusell, L.; Jochmann, G.; Szczyglowski, K.; Udvardi, M. Lotus japonicus seed and seed pod ESTs: tools for functional genomics: http://www.ncbi.nlm.nih.gov/, 2004.

(23) Dubbs, W. E.; Grimes, H. D. The mid-pericarp cell layer in soybean pod walls is a multicellular compartment enriched in specific lipoxygenase isoforms. *Plant Physiol* **2000**, *123* (4), 1281–8.

(24) Wang, H. L.; Grusak, M. A. Structure and development of *Medicago truncatula* pod wall and seed coat. *Ann. Bot. (London)* **2005**, *95* (5), 737–47.

(25) Sato, S.; Nakamura, Y.; Kaneko, T.; Asamizu, E.; Kato, T.; Nakao, M.; Sasamoto, S.; Watanabe, A.; Ono, A.; Kawashima, K.; Fujishiro, T.; Katoh, M.; Kohara, M.; Kishida, Y.; Minami, C.; Nakayama, S.; Nakazaki, N.; Shimizu, Y.; Shinpo, S.; Takahashi, C.; Wada, T.; Yamada, M.; Ohmido, N.; Hayashi, M.; Fukui, K.; Baba, T.; Nakamichi, T.; Mori, H.; Tabata, S. Genome structure of the legume *Lotus japonicus. DNA Res.* **2008**, *15* (4), 227–39.

(26) The UniProt Consortium, The universal protein resource (UniProt). Nucleic Acids Res. 2008, *36*, (Database issue), D190−5.

(27) The UniProt Consortium, The Universal Protein Resource (UniProt) 2009. *Nucleic Acids Res.* **2009**, *37* (Database issue), D169−74.

(28) Nomenclature Commitee-IUBMB, *Enzyme Nomenclature*; Academic Press: San Diego, CA, 1992.

(29) Hajduch, M.; Casteel, J. E.; Hurrelmeyer, K. E.; Song, Z.; Agrawal, G. K.; Thelen, J. J. Proteomic analysis of seed filling in *Brassica napus*. Developmental characterization of metabolic isozymes using high-resolution two-dimensional gel electrophoresis. *Plant Physiol.* **2006**, *141* (1), 32–46.

(30) Furbank, R. T.; White, R.; Palta, J. A.; Turner, N. C. Internal recycling of respiratory CO2 in pods of chickpea (*Cicer arietinum L.*): the role of pod wall, seed coat, and embryo. *J. Exp. Bot.* **2004**, *55* (403), 1687–96.

(31) Munier-Jolain, N.; Larmure, A.; Salon, C. Determinism of carbon and nitrogen reserve accumulation in legume seeds. *C. R. Biol.* **2008**, *331* (10), 780–7.

(32) Götz, K. P.; Staroske, N.; Radchuk, R.; Emery, R. J.; Wutzke, K. D.; Herzog, H.; Weber, H. Uptake and allocation of carbon and nitrogen in *Vicia narbonensis* plants with increased seed sink strength achieved by seed-specific expression of an amino acid permease. *J. Exp. Bot.* **2007**, *58* (12), 3183–95.

(33) Borisjuk, L.; Rolletschek, H.; Wobus, U.; Weber, H. Differentiation of legume cotyledons as related to metabolic gradients and assimilate transport into seeds. *J. Exp. Bot.* **2003**, *54* (382), 503–12.

(34) Hortensteiner, S.; Feller, U. Nitrogen metabolism and remobilization during senescence. *J. Exp. Bot.* **2002**, *53* (370), 927–37.

(35) Kamachi, K.; Yamaya, T.; Mae, T.; Ojima, K. A Role for Glutamine Synthetase in the Remobilization of Leaf Nitrogen during Natural Senescence in Rice Leaves. *Plant Physiol.* **1991**, *96* (2), 411–417.

(36) Kamachi, K.; Yamaya, T.; Hayakawa, T.; Mae, T.; Ojima, K. Changes in Cytosolic Glutamine Synthetase Polypeptide and its mRNA in a Leaf Blade of Rice Plants during Natural Senescence. *Plant Physiol.* **1992**, *98* (4), 1323–1329.

(37) Liljegren, S. J.; Ditta, G. S.; Eshed, Y.; Savidge, B.; Bowman, J. L.; Yanofsky, M. F. *SHATTERPROOF* MADS-box genes control seed dispersal in *Arabidopsis*. *Nature* **2000**, *404* (6779), 766–70.

(38) Østergaard, L.; Kempin, S. A.; Bies, D.; Klee, H. J.; Yanofsky, M. F. Pod shatter-resistant *Brassica* fruit produced by ectopic expression of the *FRUITFULL* gene. *Plant Biotechnol. J.* **2006**, *4* (1), 45–51.

(39) Liljegren, S. J.; Roeder, A. H.; Kempin, S. A.; Gremski, K.; Østergaard, L.; Guimil, S.; Reyes, D. K.; Yanofsky, M. F. Control of fruit patterning in *Arabidopsis* by INDEHISCENT. *Cell* **2004**, *116* (6), 843–53.

(40) Child, R. D.; Summers, J. E.; Babij, J.; Farrent, J. W.; Bruce, D. M. Increased resistance to pod shatter is associated with changes in the vascular structure in pods of a resynthesized *Brassica napus* line. *J. Exp. Bot.* **2003**, *54* (389), 1919–30.

(41) Spence, J.; Vercher, Y.; Gates, P.; Harris, N. 'Pod shatter' in *Arabidopsis thaliana, Brassica napus and B. juncea. J. Microsc.* **1996**, *181* (2), 195–203.

(42) Robles, P.; Pelaz, S. Flower and fruit development in *Arabidopsis thaliana. Int. J. Dev. Biol.* **2005**, *49* (5−6), 633–43.

(43) Petersen, M.; Sander, L.; Child, R.; van Onckelen, H.; Ulvskov, P.; Borkhardt, B. Isolation and characterisation of a pod dehiscence zone-specific polygalacturonase from *Brassica napus. Plant Mol. Biol.* **1996**, *31* (3), 517–27.

(44) Sander, L.; Child, R.; Ulvskov, P.; Albrechtsen, M.; Borkhardt, B. Analysis of a dehiscence zone endo-polygalacturonase in oilseed rape (*Brassica napus*) and *Arabidopsis thaliana*: evidence for roles in cell separation in dehiscence and abscission zones, and in stylar tissues during pollen tube growth. *Plant Mol. Biol.* **2001**, *46* (4), 469–79.

(45) González-Carranza, Z. H.; Elliott, K. A.; Roberts, J. A. Expression of polygalacturonases and evidence to support their role during

cell separation processes in *Arabidopsis thaliana*. *J. Exp. Bot.* **2007**, *58* (13), 3719–30.

(46) Ogawa, M.; Kay, P.; Wilson, S.; Swain, S. M. ARABIDOPSIS DEHISCENCE ZONE POLYGALACTURONASE1 (ADPG1), ADPG2, and QUARTET2 are Polygalacturonases required for cell separation during reproductive development in *Arabidopsis*. *Plant Cell* **2009**, *21* (1), 216–33.

(47) Goto, S.; Nishioka, T.; Kanehisa, M. LIGAND: chemical database for enzyme reactions. *Bioinformatics* **1998**, *14* (7), 591–9.

(48) Goto, S.; Okuno, Y.; Hattori, M.; Nishioka, T.; Kanehisa, M. LIGAND: database of chemical compounds and reactions in biological pathways. *Nucleic Acids Res.* **2002**, *30* (1), 402–4.

(49) Goffard, N.; Frickey, T.; Weiller, G. PathExpress update: the enzyme neighbourhood method of associating gene-expression data with metabolic pathways. *Nucleic Acids Res.* **2009**, *37* (Web Server issue), W335–9.

PR100511U