JOURNAL OF AGRICULTURAL AND FOOD CHEMISTRY

Article

pubs.acs.org/JAFC

Proteomic Analysis of Pigeonpea (*Cajanus cajan*) Seeds Reveals the Accumulation of Numerous Stress-Related Proteins

Hari B. Krishnan,*[,†] Savithiry S. Natarajan,[‡] Nathan W. Oehrle,[†] Wesley M. Garrett,[#] and Omar Darwish[§]

[†]Plant Genetics Research Unit, Agricultural Research Service, U.S. Department of Agriculture, University of Missouri, Columbia, Missouri 65211, United States

[‡]Soybean Genomics and Improvement Laboratory, PSI, Agricultural Research Service, U.S. Department of Agriculture, Beltsville, Maryland 20705, United States

[#]Animal Biosciences and Biotechnology Laboratory, Agricultural Research Service, U.S. Department of Agriculture, Beltsville, Maryland 20705, United States

[§]Department of Computer and Information Sciences, Towson University, Towson, Maryland 21252, United States

Supporting Information

ABSTRACT: Pigeonpea is one of the major sources of dietary protein for more than a billion people living in South Asia. This hardy legume is often grown in low-input and risk-prone marginal environments. Considerable research effort has been devoted by a global research consortium to develop genomic resources for the improvement of this legume crop. These efforts have resulted in the elucidation of the complete genome sequence of pigeonpea. Despite these developments, little is known about the seed proteome of this important crop. Here, we report the proteome of pigeonpea seed. To enable the isolation of maximum number of seed proteins, including those that are present in very low amounts, three different protein fractions were obtained by employing different extraction media. High-resolution two-dimensional (2-D) electrophoresis followed by MALDI-TOF-TOF-MS/MS analysis of these protein fractions resulted in the identification of 373 pigeonpea seed proteins. Consistent with the reported high degree of synteny between the pigeonpea and soybean genomes, a large number of pigeonpea seed proteins exhibited significant amino acid homology with soybean seed proteins. Our proteomic analysis identified a large number of stress-related proteins, presumably due to its adaptation to drought-prone environments. The availability of a pigeonpea seed proteome reference map should shed light on the roles of these identified proteins in various biological processes and facilitate the improvement of seed composition.

KEYWORDS: *pigeonpea, proteomics, storage proteins, 7S vicilin, 11S legumin*

INTRODUCTION

Pigeonpea (*Cajanus cajan* (L.) Millspaugh) is an important legume crop commonly grown in semiarid tropical regions of the world. It is a diploid species ($2n = 2x = 22$) belonging to the tribe Phaseoleae in the subfamily Papilionideae.[1] Even though it has been considered as an orphan crop, it is widely grown in India, Myanmar, Malawi, and Uganda and ranks as the sixth most important legume food crop globally.[2] Importantly, this orphan crop is a major protein source for a large section of human population.[3] Unlike the major crops (maize, rice, and wheat), orphan crops such as pigeonpea, chickpea (*Cicer arietinum*), and cowpea (*Vigna ungiculata*), have received only limited scientific investigation.[2] Orphan crops are well adapted for marginal environments of Africa, Asia, and South America and are vital to food security in many low-income countries.

Pigeonpea, which has its origin in India, is widely grown in all tropical and semitropical regions of the world. The seeds are a rich source of protein and are widely consumed by people in developing countries to meet their caloric and protein needs.[3,4] Considerable research effort, mainly through the efforts of the International Crops Research Institute for the Semi-Arid Tropics (ICRISAT), has resulted in the release of several high-yielding pigeonpea varieties and hybrids that have improved resistance to drought and pests.[5,6] These achievements have been made possible due to the availability of genomic tools and advancements in trait mapping and molecular breeding.[7,8] Recently, a draft genome sequence of pigeonpea has been published.[2]

Genome sequence analysis resulted in the identification of 48,680 genes in the pigeonpea genome. This number is comparable to those of soybean (46,430) and *Medicago truncatula* (47,529) but slightly higher than that of *Lotus japonicas* (38,483). A high degree of synteny between the pigeonpea and soybean genomes was reported even though these two legumes diverged about 20−30 million years ago.[2] In addition to the availability of genome sequences, extensive information on the proteome and metabolome have also been generated in model legumes (*M. truncatula* and *L. japonicas*) and other important crops.[9] Researchers are taking advantage of the vast information gleaned through genome projects as well as proteomic and metabolomics databases to identify

Received: March 3, 2017
Revised: May 1, 2017
Accepted: May 22, 2017
Published: May 22, 2017

DOI: 10.1021/acs.jafc.7b00998
J. Agric. Food Chem. 2017, 65, 4572−4581

desirable traits and to incorporate those traits for crop improvement. Advancement in mass spectrometry (MS) has resulted in the identification of thousands of legume proteins. Proteome reference maps of several legumes are now available.[10,11] Given the importance of pigeonpea as a major protein source for people in the developing world, it is imperative that the available genome information be integrated with proteomic data. Until now, no systematic study of the proteome of pigeonpea has been carried out. Here, we have investigated the protein composition of pigeonpea seed. High-resolution two-dimensional (2-D) electrophoresis followed by MALDI-TOF-TOF-MS/MS analysis resulted in the identification of 373 pigeonpea seed proteins. Our study has identified a large number of proteins with potential role in abiotic stress.

MATERIALS AND METHODS

Plant Materials. Seeds of pigeonpea variety Georgia-2 (GA-2) were obtained from Dr. Srinivas Rao (USDA-ARS Grazinglands Research Unit, El Reno, OK, USA). An early-maturing (85 days) pigeonpea variety that was developed in India was purchased from Reimer Seeds (Saint Leonard, MD, USA). Soybean (*Glycine max* L. Merr.) cultivar Williams 82 seeds were from our laboratory collection.

Protein Extraction and Fractionation. For one-dimensional electrophoresis (1-DE) dried pigeonpea seeds were ground into a fine powder using a mortar and pestle, and 20 mg was placed into a tube for extraction using 1 mL of SDS-PAGE sample buffer [(60 mM Tris-HCl, pH 6.8, 2% SDS (w/v), 10% glycerol (v/v), and 5% 2-mercaptoethanol (v/v)]. Protein extraction was carried out for 15 min, with vigorous shaking, at 25 °C. Solution was clarified with centrifugation at 16,100*g* for 10 min. Supernatant was removed and placed into a clean tube, boiled for 5 min before electrophoresis, and designated total seed protein fraction.

To isolate low-abundant seed proteins, 100 mg of seed powder was separately extracted either with 1 mL of 50% isopropanol or with 30% ethanol on a 30 °C shaker for 30 min. The resulting slurry was clarified by centrifugation at 16,100*g* for 10 min. To the clear supernatant was added 2 volumes of ice-cold acetone and left at −20 °C overnight. Precipitated proteins were recovered by centrifugation as before, and the resulting pellet was air-dried and dissolved in 300 μL of the sample buffer.

1-D Electrophoresis. One-dimensional separation followed the method of Laemmli[12] using 13.5% gels run using a Hoeffer SE 250 mini-Vertical electrophoresis apparatus (GE Healthcare, Pittsburgh, PA, USA). Separation was achieved with a constant 20 mA per gel and run time of 1.5 h. Gels were removed from the cassette and placed immediately in Coomassie Blue R-250 staining solution. Typically, 20−40 μg of protein from each sample was loaded per well.

2-D Electrophoresis. Mature, dried pigeonpea seeds were ground into a fine powder. For total protein extraction, using a mortar and pestle, 500 mg of seed powder was ground further in the presence of 5 mL of extraction buffer (100 mM Tris-Cl, pH 8.8, 0.9 M sucrose, and 0.4% β-mercaptoethanol) containing protease inhibitor cocktail (G-Biosciences, St. Louis, MO, USA). Similarly, 50% isopropanol and 30% ethanol soluble proteins extracted from 5 g of seed powder (as described above) was dissolved in 5 mL of extraction buffer. To these samples was added an equal volume of saturated phenol. Solution was then mixed vigorously for 30 min at 25 °C followed immediately by centrifugation at 5000*g* for 20 min at 25 °C in a swing-bucket rotor. The phenolic upper phase was removed and added to 10 volumes of freshly prepared 100% methanol with 0.1 M ammonium acetate (chilled to −80 °C). Precipitation of the extracted proteins progressed for 2 h at −80 °C and was followed by centrifugation at 12,000*g* for 15 min at 4 °C. Supernatant was discarded, and the protein pellet was suspended vigorously in a freshly prepared solution of 100% methanol with 0.1 M ammonium acetate and 0.01 M DTT (chilled to −20 °C). Washing of the insoluble proteins was repeated three times with the same solution with incubation at −20 °C for 20 min followed by centrifugation at 12,000*g* for 10 min at 4 °C between each vigorous wash step. Washing of the insoluble proteins was repeated four more times with a freshly prepared solution of 100% acetone containing 0.01 M DTT with incubation at −20 °C for 20 min followed by centrifugation at 12,000*g* for 10 min at 4 °C between each vigorous wash step. After the final centrifugation, the protein pellet was allowed to air-dry slightly and then solubilized in a small volume of 7 M urea, 2 M thiourea, 1% CHAPS, and 2% C7BzO. Samples were placed on ice, and the protein concentration was estimated following the method of Bradford.

Approximately 300 μg of sample protein was loaded per strip using in-gel rehydration. Linear gradient, 13 cm IPG strips (GE Healthcare) were brought to a rehydration volume of 250 μL with 7 M urea, 2 M thiourea, 1% CHAPS, and 2% C7BzO with a final concentration of the following: 5% glycerol, 60 mM DTT, and 2.2% 2-HED. Strips were then passively rehydrated with the entire rehydration solution containing protein sample at 25 °C for 15 h prior to isoelectric focusing.

Isoelectric focusing was achieved at 8000 V until 60,000 Vh using a Protean II IEF (Bio-Rad, Hercules, CA, USA). Separation was finalized when the current was at or below 10 μA per strip and steady for 1 h. Prior to separation in the second dimension, IPG strips were first equilibrated with 5% SDS in a urea-based solution (50 mM Tris-Cl, pH 8.8, 6 M urea, 30% glycerol, and 0.1% bromophenol blue) containing 2% DTT for 15 min and again with the solution containing 2.5% iodoacetamide for 15 min. IPG strips were carefully placed onto a Hoeffer SE600 (GE Healthcare) 15%T gel for vertical second-dimension separation and secured into place with warm 1% agarose dissolved in SDS-PAGE running buffer (0.2% SDS). Gels were run at an initial 20 mA/gel for 1 h followed by 50 mA/gel for the remainder of the run (elimination of dye front; approximately 3.5 h). 2-DE gels were immediately removed and fixed in 5:4:1 (methanol/water/acetic acid) for 1 h, followed by staining in Coomassie G-250 for 24 h.

Image Acquisition. 1-DE and 2-DE Coomassie-stained gels were destained with multiple changes of ultrapure H_2O to remove background. All gels were scanned separately using an Epson V700 Perfection scanner controlled through Adobe Photoshop.

In-Gel Digestion of Protein Spots. For protein digestion, the protein spots were excised with a 1.5 mm Spot Picker (The Gel Co., San Francisco, CA, USA) and dehydrated using acetonitrile. Protein digestion was performed at 37 °C with modified porcine trypsin (Promega, Madison, WI, USA) in ammonium bicarbonate. The resulting peptides were extracted, dried, and dissolved in 50% acetonitrile and 0.1% trifluoroacetic acid.

Mass Spectrometry and Data Interpretation. The digested peptides were subjected to MALDI-TOF-TOF-MS/MS analysis for protein identification. An AB SCIEX TOF/TOF 5800 System (AB SCIEX, Framingham, MA, USA) operated in positive ion reflector mode was used to analyze tryptic peptides. Samples were cocrystallized with a 5 mg/mL content of α-cyanohydroxycinnamic acid (CHCA) matrix prepared in 70% acetonitrile containing 0.1% trifluoroacetic acid. For analysis of tryptic unknown peptides, the instrument was operated in batch mode, which entails first performing an MS survey scan on all spots of interest, followed by sequential MS/MS analysis of peaks detected in the MS scan. MS/MS acquisition was controlled by an interpretation method that acquired MS/MS spectra on the strongest precursors up to the first 100 precursors detected in the MS scan. An exclusion mass list was prepared to prevent MS/MS analysis of common human keratin contaminant and minor porcine trypsin autolysis peaks. MS spectra for both standards and unknowns were acquired in positive ion reflector mode with 400 shots of a 349 nm Nd:YAG laser operating at 404 Hz. MS/MS spectra were also acquired in positive ion reflector mode with 1000 laser shots firing at a rate of 1010 Hz. Collision energy was set to 1 kV, and collision-induced dissociation (CID) was enabled with air as the collision gas in the CID cell. When possible, MS spectra were internally calibrated using known trypsin autolysis peaks at m/z 842.51 and 2211.10.

Protein Identification. The MS/MS results were analyzed using Mascot version 2.3.02 (Matrix Science, Boston, MA, USA). A FASTA formatted database was constructed for Mascot searching by downloading the protein reference sequence of *Cajanus cajan* from

DOI: 10.1021/acs.jafc.7b00998
J. Agric. Food Chem. 2017, 65, 4572−4581

the National Center for Biotechnology Information nonredundant (NCBInr) protein database using the appropriate taxonomy filter (green plants). The following parameters were used for Mascot searches: monoisotopic mass, parent ion tolerance of 50 ppm, fragment ion tolerance of 0.5 Da, peptide charge state set to 1^+, trypsin as digesting enzyme with 1 missed cleavage allowed and variable modification of oxidation of methionine, N-terminal pyroglutamic acid from glutamic acid or glutamine. Scaffold (version Scaffold_4.3.0, Proteome Software Inc., Portland, OR, USA) was used to validate MS/MS-based peptide and protein identifications. Peptide identifications were accepted if they could be established at >80.0% probability as specified by the Peptide Prophet algorithm. Protein identifications were accepted if they could be established at >95.0% probability and contained at least two identified peptides. Protein probabilities were assigned by the Protein Prophet algorithm. The functional classification of the identified pigeonpea seed proteins was carried out on the basis of GO term retrieval using Blast2GO (https://www.blast2go.com/).

RESULTS AND DISCUSSION

1-D SDS-PAGE Analysis of Pigeonpea Seed Proteins. Seed storage proteins have been traditionally classified into four major groups (albumins, globulins, prolamins, and glutelins) on the basis of their solubility in different solvents. Proteins soluble in salt solutions are termed globulins and are the predominant group of seed proteins in legumes.[13] On a dry weight basis the protein content of pigeonpea seed is estimated to be about 22−28%.[14,15] On the basis of their sedimentation coefficients globulins are divided into two groups, 7S vicilin-type globulins and 11S legumin-like globulins. In pigeonpea, the globulin fraction represents about 54−60%; albumin, 10−15%; glutelin, 10−15%; and prolamin, 4−5% of the total protein.[14,15] Like other legumes, pigeonpea globulin also contains a low concentration of sulfur-containing amino acids.[16] Among legumes, the seed proteins of pea, soybean, broad bean, and French bean have been well studied.[13] In contrast, only limited studies on pigeonpea seed storage proteins have been conducted. First, we investigated the protein composition of pigeonpea seed proteins extracted with three different solvents and compared their profile to that of soybean (Figure 1). SDS-PAGE resolved the total seed proteins into several distinct bands ranging in molecular weight from 10 to 100 kDa. The most abundant proteins of pigeonpea had molecular weights of 64 and 47 kDa. These two prominent proteins represent the two subunits of the 7S vicilin.[17,18] The 11S legumin-like proteins are not abundant in pigeonpea total seed protein fraction (Figure 1). In contrast, in soybean both the 7S and 11S globulins are prominent (Figure 1). Soybean 7S globulin is made up of three subunits, α' (76 kDa), α (72 kDa), and β (53) kDa, and the 11S glycinin is represented by two abundant group of proteins with molecular weights of 37−44 and 17−22 kDa, respectively (Figure 1). We also examined the protein profile of the 50% isopropanol and 30% ethanol soluble proteins of pigeonpea seed proteins. Previously we have shown these solvents to enable the isolation and enrichment of low-abundant proteins in soybean.[19] Ethanol (30%) was more efficient in extracting low-abundant proteins than 50% isopropanol (Figure 1). A comparison of the pigeonpea and soybean seed proteins extracted by different solvents clearly indicates size heterogeneity among the major seed storage proteins between these legumes (Figure 1).

2-D SDS-PAGE Separation of Pigeonpea Seed Proteins. A recent study has examined three different protein extraction procedures for the separation of pigeonpea proteins from different tissues by 2-D gel electrophoresis.[20] Better resolution of pigeonpea seed proteins was achieved with a phosphate−TCA−acetone protocol. We had previously successfully used phenol-based protein extraction protocol for high resolution of soybean seed proteins by 2-D gel electrophoresis.[21,22] The same procedure was also employed for the extraction of pigeonpea seed proteins from three biological triplicates. Analysis of pigeonpea seed proteins was performed with immobilized p*I* gradient 3−10 and p*I* gradient 4−7. Most of the resolved proteins were clustered in the p*I* region ranging from 4 to 7, although several well-resolved protein spots with basic p*I* were also found (Figure 2A). The major seed proteins of pigeonpea were seen migrating as poorly resolved abundant isoelectric protein series (Figure 2A). To improve the resolution of these groups of proteins, we subjected pigeonpea seed proteins to isoelectric focusing using p*I* 4−7 IPG strips followed by SDS-PAGE analysis (Figure 2B). This resulted in drastically improved protein separation of most seed proteins. However, because of their abundance, the major seed storage proteins of pigeonpea were not distinctly separated into individual protein spots (Figure 2B). When significantly lower proteins were loaded (50 μg), these groups of seed proteins were resolved into a series of discrete multiple isoelectric spots (data not shown). A similar 2-D gel migration pattern of major seed proteins has also been reported in soybean.[23] The 7S globulins are glycoproteins,[13] and post-translational modification has been suggested to be responsible for the tight clustering of 7S globulin protein spots on 2-D gels.

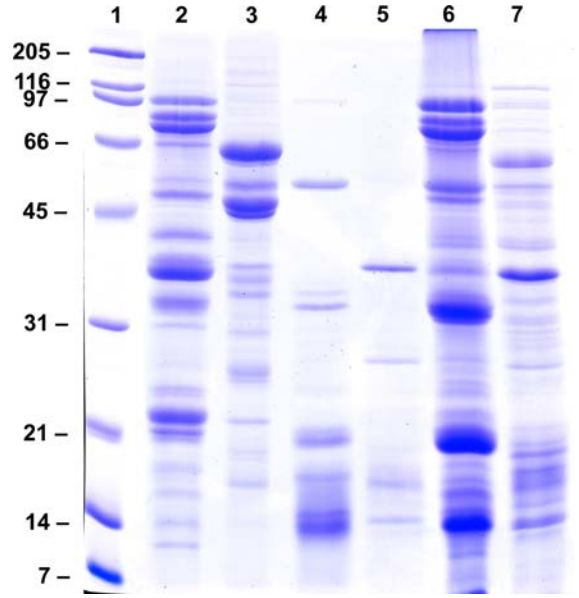

Figure 1. SDS-PAGE comparison of pigeonpea and soybean seed proteins. Pigeonpea (lanes 3, 5, and 7) and soybean (lanes 2, 4, and 6) seed proteins extracted by different methods were analyzed by SDS-PAGE. Seed proteins were separated using 15% SDS−polyacrylamide gels, and resolved proteins were detected by staining the gel with Coomassie Blue. Lanes: 1, molecular weight markers (sizes in kDa are shown on the left); 2 and 3, total seed proteins; 4 and 5, 50% isopropanol-extracted proteins; 6 and 7, 30% ethanol extracted proteins.

2-D Separation of Low Abundant Pigeonpea Seed Proteins. Proteomic analysis of seed proteins often provides challenges largely due to the presence of many abundant seed storage proteins, which limits the detection of low-abundant

DOI: 10.1021/acs.jafc.7b00998
J. Agric. Food Chem. 2017, 65, 4572−4581

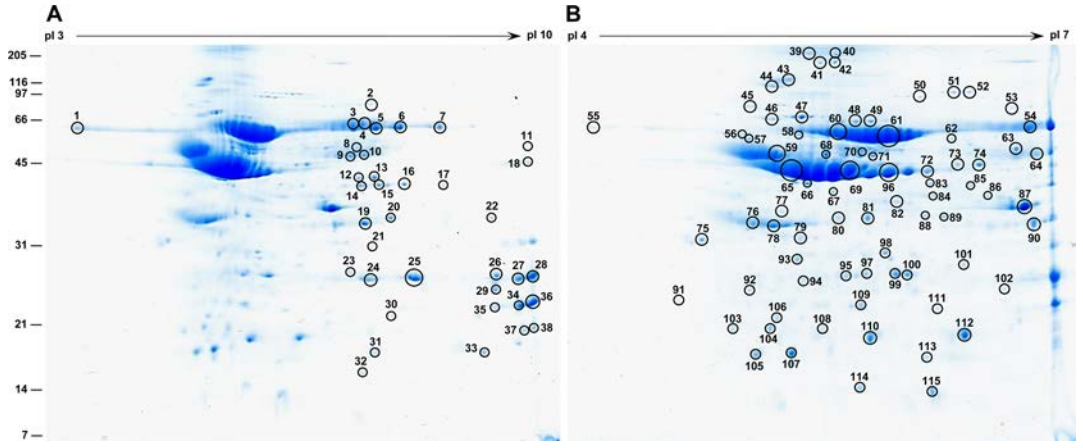

Figure 2. Two-dimensional gel electrophoresis of pigeonpea total seed proteins. Seed proteins (300 μg) were first separated by isoelectric focusing either on p*I* 3−10 strip (A) or p*I* 4−7 strip (B) followed by separation by SDS-PAGE on 16% gels. The gels were stained with Colloidal Coomassie Blue G-250. Protein spots picked from the gels for identification using MALDI-TOF-TOF-MS/MS analysis are circled. The positions and sizes of the molecular weight markers in kDa are shown on the left.

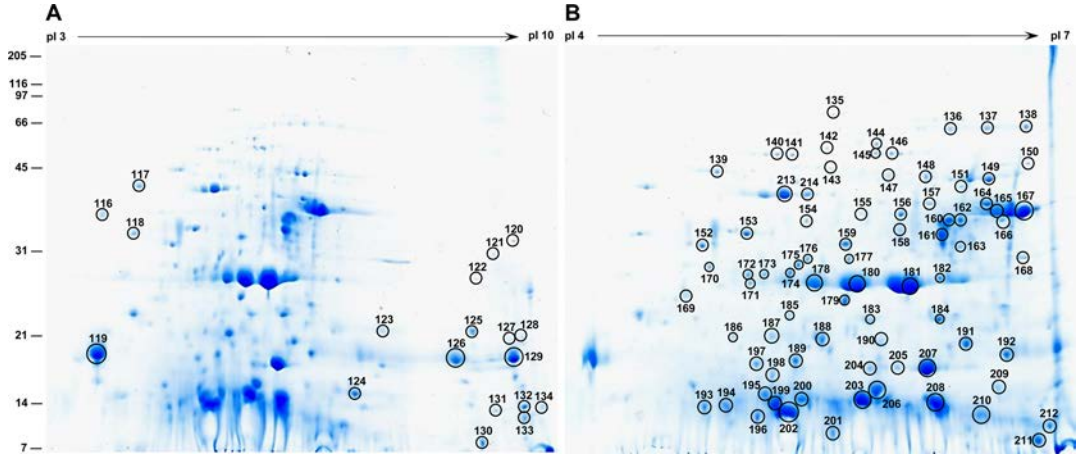

Figure 3. Two-dimensional gel electrophoresis of 50% isopropanol extracted pigeonpea seed proteins. Seed proteins (300 μg) were first separated by isoelectric focusing either on p*I* 3−10 strip (A) or p*I* 4−7 strip (B) followed by separation by SDS-PAGE on 16% gels. The gels were stained with Colloidal Coomassie Blue G-250. Protein spots picked from the gels for identification using MALDI-TOF-TOF-MS/MS analysis are circled. The positions and sizes of the molecular weight markers in kDa are shown on the left.

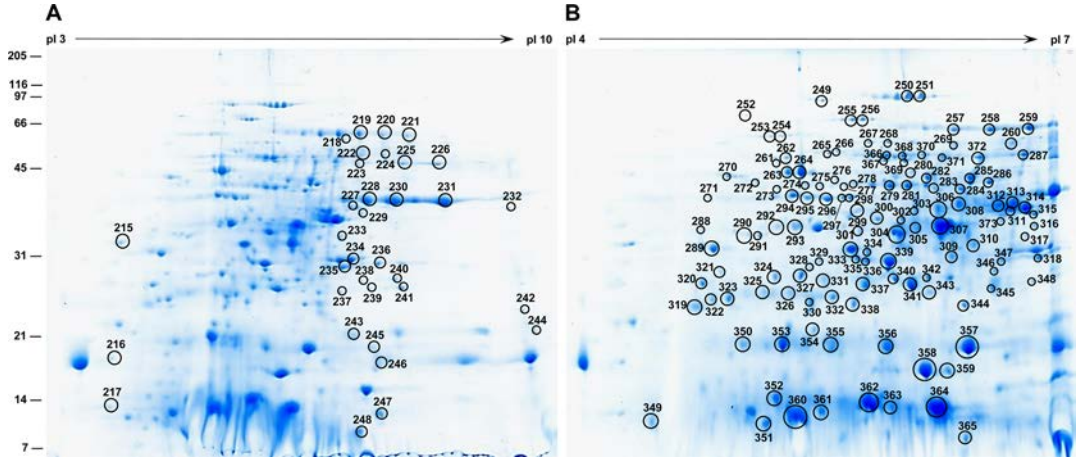

Figure 4. Two-dimensional gel electrophoresis comparison of 30% ethanol extracted pigeonpea seed proteins. Seed proteins (300 μg) were first separated by isoelectric focusing either on p*I* 3−10 strip (A) or p*I* 4−7 strip (B) followed by separation by SDS-PAGE on 16% gels. The gels were stained with Colloidal Coomassie Blue G-250. Protein spots picked from the gels for identification using MALDI-TOF-TOF-MS/MS analysis are circled. The positions and sizes of the molecular weight markers in kDa are shown on the left.

DOI: 10.1021/acs.jafc.7b00998
J. Agric. Food Chem. 2017, 65, 4572−4581

Table 1. Stress-Related Proteins Identified from Pigeonpea (*Cajanus cajan*) by MALDI-TOF-TOF-MS/MS Analysis[a]

						MW		pI	
spot no.	protein name [*Cajanus cajan*]	accession no.	protein score	% coverage	no. of peptides	theor	exptl	theor	exptl
5	embryonic protein DC-8	KYP73710.1	644	21	6	40514	62000	7.64	7.50
6	embryonic protein DC-8	KYP73710.1	556	21	6	40514	62000	7.64	7.80
7	embryonic protein DC-8	KYP73710.1	515	21	6	40514	62000	7.64	8.40
11	elongation factor 1-α	KYP42506.1	194	11	4	49452	50000	9.15	9.60
45	heat shock cognate protein 80	KYP49760.1	345	13	6	80888	80000	4.96	5.00
46	luminal-binding protein 4	KYP42326.1	559	21	8	73366	71000	5.10	5.10
47	heat shock cognate 70 kDa protein	KYP50771.1	387	11	4	71032	73000	5.10	5.30
95	1-Cys peroxiredoxin	KYP50274.1	461	55	7	24458	28000	5.62	5.50
97	1-Cys peroxiredoxin	KYP50274.1	94	5	1	24458	28000	5.62	5.70
99	1-Cys peroxiredoxin	KYP50274.1	681	55	7	24458	28000	5.62	5.90
100	1-Cys peroxiredoxin	KYP50274.1	709	58	8	24458	28000	5.62	6.00
103	desiccation protectant protein Lea14 isogeny	KYP77206.1	271	33	4	16795	21000	5.11	4.90
104	desiccation protectant protein Lea14 isogeny	KYP77206.1	420	51	6	16795	21000	5.11	5.10
105	18.2 kDa class I heat shock protein	KYP59373.1	155	22	3	17416	17000	5.40	5.00
107	18.2 kDa class I heat shock protein	KYP59373.1	436	51	5	17416	17000	5.40	5.20
110	18.2 kDa class I heat shock protein	KYP48731.1	570	48	5	17166	22000	5.50	5.70
113	superoxide dismutase [Cu-Zn]	KYP77391.1	231	33	3	15144	17000	6.21	6.05
114	EMB-1 protein	KYP49070.1	111	22	2	10920	14000	6.14	5.65
115	EMB-1 protein	KYP49070.1	209	42	3	10920	13000	6.14	6.10
130	10 kDa chaperonin	KYP40884.1	501	81	6	10510	9000	8.05	8.90
136	embryonic protein DC-8	KYP73710.1	358	20	5	40514	66000	7.64	6.30
137	embryonic protein DC-8	KYP73710.1	206	7	2	40514	66000	7.64	6.50
138	embryonic protein DC-8	KYP73710.1	472	20	5	40514	66000	7.64	6.75
147	embryonic protein DC-8	KYP73710.1	78	3	1	40514	45000	7.64	5.90
148	embryonic protein DC-8	KYP73710.1	337	17	4	40514	45000	7.64	6.10
178	1-Cys peroxiredoxin	KYP50274.1	554	52	7	24458	26000	5.62	5.40
180	1-Cys peroxiredoxin	KYP50274.1	702	60	8	24458	26000	5.62	5.70
181	1-Cys peroxiredoxin	KYP50274.1	726	60	8	24458	26000	5.62	6.00
186	desiccation protectant protein Lea14 isogeny	KYP77206.1	425	52	7	16795	22000	5.11	4.95
187	desiccation protectant protein Lea14 isogeny	KYP77206.1	162	23	3	16795	22000	5.11	5.10
189	18.2 kDa class I heat shock protein	KYP59373.1	417	38	4	17416	19000	5.40	5.15
193	1-Cys peroxiredoxin	KYP50274.1	67	5	1	24458	13000	5.62	4.70
195	EMB-1 protein	KYP49070.1	206	36	3	10920	13000	6.14	5.05
200	EMB-1 protein	KYP49070.1	74	12	1	10920	15000	6.14	5.35
203	EMB-1 protein	KYP49070.1	333	46	4	10920	15000	6.14	5.70
204	superoxide dismutase [Cu-Zn]	KYP77391.1	416	52	4	15144	17000	6.21	5.80
205	superoxide dismutase [Cu-Zn]	KYP77391.1	393	52	4	15144	17000	6.21	5.95
207	superoxide dismutase [Cu-Zn]	KYP77391.1	529	52	4	15144	17000	6.21	6.10
208	EMB-1 protein	KYP49070.1	422	46	4	10920	15000	6.14	6.10
219	embryonic protein DC-8	KYP73710.1	596	27	6	40514	65000	7.64	7.20
220	embryonic protein DC-8	KYP73710.1	564	27	6	40514	65000	7.64	7.50
221	embryonic protein DC-8	KYP73710.1	286	19	4	40514	60000	7.64	7.90
222	embryonic protein DC-8	KYP73710.1	516	20	5	40514	55000	7.64	7.20
223	embryonic protein DC-8	KYP73710.1	366	20	5	40514	49000	7.64	7.20
224	embryonic protein DC-8	KYP73710.1	457	20	5	40514	55000	7.64	7.50
242	embryonic protein DC-8	KYP73710.1	199	8	2	40514	60000	7.64	9.30
257	embryonic protein DC-8	KYP73710.1	237	13	3	40514	60000	7.64	6.30
258	embryonic protein DC-8	KYP73710.1	428	26	6	40514	60000	7.64	6.60
259	embryonic protein DC-8	KYP73710.1	795	31	7	40514	60000	7.64	6.85
260	embryonic protein DC-8	KYP73710.1	422	22	5	40514	50000	7.64	6.70
269	embryonic protein DC-8	KYP73710.1	416	19	4	40514	55000	7.64	6.30
280	Embryonic protein DC-8	KYP73710.1	59	3	1	40514	44000	7.64	6.10
282	embryonic protein DC-8	KYP73710.1	531	21	5	40514	45000	7.64	6.10
285	embryonic protein DC-8	KYP73710.1	288	17	4	40514	47000	7.64	6.05
303	embryonic protein DC-8	KYP73710.1	244	12	3	40514	37000	7.64	6.05
331	1-Cys peroxiredoxin	KYP50274.1	657	44	7	24458	27000	5.62	5.50
337	1-Cys peroxiredoxin	KYP50274.1	337	25	4	24458	28000	5.62	5.80
341	1-Cys peroxiredoxin	KYP50274.1	315	32	4	24458	28000	5.62	6.00
350	desiccation protectant protein Lea14 isogeny	KYP77206.1	377	46	6	16795	21000	5.11	4.95

DOI: 10.1021/acs.jafc.7b00998
J. Agric. Food Chem. 2017, 65, 4572−4581

Table 1. continued

spot no.	protein name [*Cajanus cajan*]	accession no.	protein score	% coverage	no. of peptides	MW theor	MW exptl	p*I* theor	p*I* exptl
353	desiccation protectant protein Lea14 isogeny	KYP77206.1	494	52	7	16795	21000	5.11	5.20
358	superoxide dismutase [Cu-Zn]	KYP77391.1	159	14	1	15144	17000	6.21	6.20
359	superoxide dismutase [Cu-Zn]	KYP77391.1	174	23	2	15144	18000	6.21	6.20
361	EMB-1 protein	KYP49070.1	376	55	4	10920	14000	6.14	5.40
362	EMB-1 protein	KYP49070.1	386	55	4	10920	14000	6.14	5.60
363	EMB-1 protein	KYP49070.1	210	45	3	10920	14000	6.14	5.80
364	EMB-1 protein	KYP49070.1	421	55	4	10920	12000	6.14	6.15
365	embryonic protein DC-8	KYP73710.1	254	9	2	40514	10000	7.64	6.50

[a]Spot identification numbers correspond to those proteins labeled in Figures 2, 3, and 4). Molecular weight (M_r) and isoelectric point (p*I*) values are given as theoretical and experimental values. Protein scores are raw values and represent the total ion scores. Peptides matched, % coverage, and accession numbers from within the *Cajanus cajan* database are given.

proteins.[24] This problem can be overcome by utilizing extraction procedures that will preclude the isolation of the majority of storage proteins. The abundant storage proteins of pigeonpea are globulins, which are easily soluble in salt solutions but have only limited solubility in organic solvents such as ethanol and isopropanol (Figure 1). We took advantage of this differential solubility property of seed proteins to isolate protein fractions that are enriched with low-abundant proteins. 2-D separation of 50% isopropanol soluble proteins revealed the presence of many well-resolved protein spots (Figure 3), most of which were either absent or poorly represented in protein samples prepared by general isolation protocol (Figure 2). The most abundant seed storage proteins were not detected in 50% isopropanol soluble protein fraction; instead, there was an enrichment of numerous proteins ranging in molecular weight from 10 to 40 kDa (Figure 3). Similarly, an examination of the protein profile of pigeonpea seed proteins soluble in 30% ethanol revealed many different protein spots not seen in previous fractions (Figure 4). The abundant seed proteins were also not detected in the 30% ethanol extracted protein fraction (Figure 4). Although there were several protein spots that were unique to 30% ethanol and 50% isopropanol extracted protein fractions, a significant number of proteins spots were also shared. A total of 475 protein spots (150 spots from 2-D gels using protein obtained by the standard extraction protocol (Figure 2), 128 spots from 50% isopropanol extracted fraction (Figure 3), and 197 protein spots from 30% ethanol extracted fraction (Figure 4) were excised from 2-D gels. These protein spots were digested with trypsin, and the digested peptides were subjected to MALDI-TOF-TOF-MS/MS analysis for protein identification

Identification of Pigeonpea Seed Proteins. The identity of pigeonpea seed proteins obtained by three protein extraction methods is shown in Table 1. Initially, the mass spectrometer spectra of pigeonpea peptides were searched against the NCBI nonredundant *Glycine max* database and the UniProtKB *Glycine max* database. This search resulted in the identification of 250+ protein spots, but a significant number of proteins had no match. This situation may have resulted due to differences in the sequence of peptides between closely related organisms, leading to misidentification or nonidentification of MS-acquired spectra.[25] The recent availability of the *Cajanus cajan* protein database enabled us to search against a homologous database, leading to the identification of a large number of spots with high-confident PSMs (Supporting Information Table 1).

Seed Storage Proteins. The most abundant protein spots seen in the 2D gels (Figure 2) were identified as members of vicilin- and legumin-type globulins (Supporting Information Table 1). This is predictable because globulins represent about 61% of pigeonpea total seed proteins.[17] Eighteen protein spots (1, 39, 40, 41, 42, 55, 60, 61, 65, 69, 70, 71, 72, 82, 84, 94, 96, and 143) were identified as β-conglycinin, a member of the 7S globulin family. The 7S globulins occur as trimeric proteins of 150−190 kDa and are generally devoid of cysteine residues.[13] In soybean, the 7S β-conglycinin is made up of three subunits, α', α, and β, with apparent molecular weights of 76, 72, and 53 kDa, respectively.[26,27] These proteins, like other 7S globulins, are glycosylated. A previous study has shown that pigeonpea vicilin is also a glycoprotein and composed of 72 and 57 kDa subunits.[17] The 20 protein spots identified as β-conglycinin belonged to just two accessions (KYP75005.1 β-conglycinin, α chain; KYP48930.1 β-conglycinin, β chain). We recovered the amino acid sequence of these β-conglycinin subunits and subjected them to a homology search against the UniProt database using BLAST (Basic Local Alignment Search Tool). This analysis revealed significant amino acid sequence homology between pigeonpea and soybean β-conglycinin subunits (Supporting Information Figures 1 and 2).

Accumulation of β-conglycinin is controlled by the nutritional status of the plant.[28] When plants are grown in the presence of a high concentration of nitrogen in the growth medium, a preferential accumulation of the β-subunit of β-conglycinin occurs. In general, β-conglycinin contains low amounts of sulfur-containing amino acids. In particular, the β-subunit of β-conglycinin contains no methionine and cysteine.[26,27] The paucity of sulfur-containing amino acids lowers the nutritional quality of 7S seed proteins. At present, it is not known if the accumulation of pigeonpea β-conglycinin is also regulated by the nitrogen status.

Eleven protein spots (24, 25, 26, 27, 28, 34, 35, 36, 76, 78, and 79) were identified as glycinin (Supporting Information Table 1). Glycinins are synthesized as precursor proteins that undergo post-translational processing resulting in acidic and basic subunits, which are held together by a single disulfide bond.[29] In soybean, glycinins are represented by two abundant groups of protein with molecular weights of 37−44 kDa (acidic subunit) and 17−22 kDa (basic subunit), respectively. These proteins are the products of the five-glycinin genes (*gy1*, *gy2*, *gy3*, *gy4*, and *gy5*). Just as in soybean, the pigeonpea glycinins are also represented by acidic (59, 76, 78, 79) and basic (24, 25, 26, 27, 28, 34, 35, 36) subunits (Figure 2). BLAST analysis also revealed high sequence homology between pigeonpea and soybean glycinins (Supporting Information Figure 3).

DOI: 10.1021/acs.jafc.7b00998
J. Agric. Food Chem. 2017, 65, 4572−4581

Protein spots 18, 31, and 33 have been identified as basic 7S globulins. The 7S globulin of pigeonpea has been previously identified as γ-protein, a protein rich in sulfur amino acids.[30] It is synthesized as larger precursor protein and processed into two subunits of molecular weights of 32 and 20 kDa.[30] In soybean the basic 7S globulin is synthesized as a precursor protein of 43 kDa and subsequently processed into α (27 kDa) and β (16 kDa) subunits.[31] Interestingly, basic 7S globulin is known to bind insulin and insulin-like growth factors, leading to the suggestion that this protein may function as a hormone receptor.[32] Basic 7S globulin is present in several legumes and could play a role in plant defense against phytopathogens.[33] Becaus of sequence similarity with other legumes, it is likely that pigeonpea basic 7S globulin may also have similar functions.

Plant Protease Inhibitors (PIs). Plant PIs are generally present at high concentration in seeds. In addition to reserve organs, they also accumulate in vegetative tissue in response to the attack of insects and pathogens.[34] In plants, PIs that inactivate serine, cysteine, and metallocarboxyproteases are ubiquitous. Interestingly, PIs that are specific against aspartic proteases have not been detected in seeds.[35] PIs are generally encoded by multigene families. Two types of protease inhibitors, Kunitz trypsin inhibitors (KTi) and Bowman−Birk inhibitors (BBi), accumulate in soybean seeds. KTi is a 21 kDa protein and specifically inhibits trypsin, whereas BBi is made up of a group of cysteine-rich small molecular weight proteins that inhibit trypsin, chymotrypsin, and elastase.[26,27] Soybean seeds contain three trypsin inhibitors (KTi1, KTi2, and KTi3) with KTi3 as the most abundantly expressed in seeds.[36] In our proteomic study we identified 10 protein spots (30, 112, 123, 125, 190, 191, 245, 355, 356, and 357) as KTi and 2 spots (210 and 248) as BBi (Supporting Information Table 1). In addition, protein spots (199 and 360) have been identified as subtilisin inhibitor 1 (Supporting Information Table 1). These proteins were more prevalent in the 50% isopropanol and 30% ethanol extracted protein preparations (Figures 3 and 4).

Soybean meal is a predominant source of protein meal (~67%) utilized worldwide. It is extensively used in animal feed mixtures for poultry (50%) and swine (26%) (http://soystats.com). However, soybean meal also contains a number of antinutritional factors that reduce animal weight gain. Although lectins and phytate could have negative effects on animal weight gain, the protease inhibitors KTi and BBi, because of their inhibitory effect on animal digestive enzymes, are considered to be primary antinutritional factors. Although pigeonpea is a main dietary protein source for over a billion people in Asia, it is also used as livestock feed.[37−39] The presence of trypsin inhibitors, as in soybean meal, could limit its use as livestock feed. Studies have shown that processing (heat inactivation) of the seeds significantly reduced the trypsin inhibitor activity, resulting in the enhancement of its nutritive value.[40] In developing countries pigeonpea meal has the potential to be substituted for soybean meal in ruminant diet.[37] Thus, pigeonpea meal could serve as an alternative protein source in animal feed, especially in countries such as India and eastern and southern Africa.

Stress-Related Proteins. During the seed desiccation phase, drastic metabolic changes occur, resulting in the accumulation of specialized proteins. This enables the seeds to maintain viability for long periods. On the basis of physiological function, these proteins can be grouped into molecular chaperones, antioxidative proteins, and late embryogenesis abundant (LEA) proteins.[41] Our proteomic analysis revealed several proteins that can be classified as stress-related proteins (Table 1). Eight protein spots (45, 46, 47, 105, 107, 110, 130, and 189) were grouped under molecular chaperones. Protein spots 45, 46, and 47 were identified as 80 kDa heat shock proteins (HSP), and spots 105, 107, 110, and 189 were identified as 18.2 HSP. Protein spot 130 was identified as a 10 kDa chaperonin. These proteins play an important role during stress by preventing the irreversible aggregation of denatured proteins.[42,43]

Plants have developed defense systems to counteract reactive O_2 species (ROS) that are produced under stress conditions. One common mechanism that involves the removal of ROS is increased enzymatic activity of antioxidant enzymes.[44] Superoxide dismutases (SODs) play an important role in negating the effect of ROS.[45,46] In our study we have identified six protein spots (113, 204, 205, 207, 358, and 359) as superoxide dismutases. Additionally, several proteins have been identified as peroxiredoxins (Prx). These proteins are important players in the antioxidant defense system and are found in organisms ranging from bacteria to humans.[47] They can contain either one or two conserved cysteine residues. Interestingly, the 1-Cys peroxiredoxin genes are expressed in seeds specifically in the embryo and the aleurone layer, and their expression is elevated by ABA and osmotic stresses.[47] In this study we identified 11 protein spots (95, 97, 99, 100, 178, 180, 181, 193, 331, 337, and 341) as 1-Cys peroxiredoxin (Table 1). Some of the protein spots (178, 180, and 181) are abundantly present in the 50% isopropanol extracted protein fraction (Figure 3). Even though several Prx (2-CysPrx, 1-CysPrx, and PrxQ) are present in higher plants, our proteomic analysis indicated that pigeonpea seeds contain only 1-CysPrx. These proteins have been shown to be involved in protecting the seeds against damaging reactive oxygen species during desiccation and maintenance of dormancy.[47]

Another group of proteins that accumulate to high levels during the last stage of seed maturation are the late embryogenesis abundant (LEA) proteins.[48] LEA proteins are expressed during late embryogenesis and under various stress conditions. The expression of LEA proteins occurs when plants are subjected to water deficiency as well as during seed dehydration. These proteins are hydrophilic because of high contents of Gly, Ala, and Ser.[49] A comprehensive computational reanalysis of LEA proteins has grouped these proteins into several different classifications.[50] These proteins have a protective function during desiccation.[51] Protein spots 103, 104, 186, 187, 350, and 353 were identified as LEA proteins. A total of nine protein spots (114, 115, 195, 200, 203, 208, 361, 362, and 364) were identified as EMB-1 proteins.[52] Twenty-five protein spots (5, 6, 7, 136, 137, 138, 147, 148, 219, 220, 221, 222, 223, 224, 242, 257, 258, 259, 260, 269, 280, 282, 285, 303, and 365) were identified as embryonic protein DC-8 (Table 1). These proteins accumulate in all monocot and dicot embryos[53,54] and are expressed primarily at the time of embryo desiccation or in response to ABA.[49] The abundance of these stress-related proteins in pigeonpea seeds suggests an important role of these proteins in protecting macromolecules, such as enzymes and lipids, during seed desiccation.[55]

Hypothetical Proteins. A total of 76 protein spots were identified as hypothetical proteins (Supporting Information Table 2). However, when the amino acid sequences of these hypothetical proteins were subjected to BLAST search against the National Center for Biotechnology Information non-

DOI: 10.1021/acs.jafc.7b00998
J. Agric. Food Chem. 2017, 65, 4572−4581

redundant (NCBInr) protein database, significant sequence homology to several soybean and *Medicago truncatula* seed proteins were observed (Supporting Information Table 2). Of the 76 protein spots classified as hypothetical proteins, 21 spots (87, 158, 159, 160, 161, 162, 165, 167, 169, 246, 306, 307, 308, 309, 310, 311, 313, 314, 319, 333, and 343) were found to have significant amino acid identity to soybean dehydrin. Seven additional proteins spots (9, 10, 124, 126, 129, 202, and 211) were identified as seed maturation proteins (Supporting Information Table 2). Dehydrins play an important role in plant responses and adaptation to stress. They belong to group 2 LEA proteins and accumulate during the late stages of embryogenesis.[56] A positive correlation between dehydrin gene expression or dehydrin protein accumulation and plant stress tolerance has been established in plants, leading to the suggestion that the accumulation of dehydrin can be used as a plant molecular marker for stress tolerance.[56] In our study we found that the dehydrins are enriched in 50% isopropanol and 30% ethanol extracted protein fractions. Our observation is consistent with an earlier study in which dehydrins were reported to be abundant in 40% isopropanol extract fraction in soybean.[19] The abundant accumulation of seed maturation proteins, dehydrins, and LEA proteins in pigeonpea seeds indicates a critical role of these proteins in stress tolerance because this legume is grown mostly in arid and semiarid conditions. Consistent with our observation, a previous study had reported that pigeonpea had a higher number of drought-responsive sequences in its genome than did *M. truncatula* and *L. japonicas*.[2]

Functional Classification of the Pigeonpea Seed Proteins. Pigeonpea seed proteins were functionally classified on the basis of gene ontology (GO) terms of the three main categories: biological process, molecular function, and cellular component. Pigeonpea proteins were classified in 31, 5, and 7 functional categories representing biological process, molecular function, and cellular component, respectively. All 31 functional categories referred to in the biological process category were well represented, with small molecule metabolic process, oxidation−reduction process, organonitrogen compound metabolic process, macromolecule metabolic process, organic substance catabolic process, and carbohydrate metabolic process being the dominant categories (Supporting Information Figure 4). Pigeonpea seed proteins representing hydrolase activity (29%), oxidoreductase activity (26%), and ion binding (19%) were the most frequent terms in the molecular function category (Supporting Information Figure 4). In the cellular component category, cytoplasm, proteasome complex, and catalytic complex were observed to occur most frequently. The abundant representation of pigeonpea proteins involved in catalytic activity, binding activity, nutrient reservoir activity, and carbohydrate metabolic activity is typical of seeds and mirrors what has been reported in other legumes.

During the past decade, through the efforts of the International Crops Research Institute for Semi-Arid Tropics (ICRISAT), significant progress has been made to develop large-scale genomics resources for pigeonpea, chickpea, and groundnut.[6] These efforts have led to the development of extensive genetic and physical maps, SSR markers, SNPs, high-throughput marker genotyping, and genome sequencing.[6] However, advances in pigeonpea genomics and trancriptomics should also be complemented with proteomics and metabolomics research. In contrast to pigeonpea, extensive proteomics work has been reported in soybean, common bean, *M. truncatula*, and *L. japonicas*.[9] To develop a more sustainable pigeonpea, it is imperative to develop a pigeonpea proteome atlas. Efforts are being made to fill this void. For example, a recent study compared the seed protein profile of nontransgenic and transgenic pigeonpea harboring two independent *cry* genes, by 2-D gel analysis.[57] This comparison detected 11 and 10 protein spots showing a difference in relative intensity of ≥2.5-fold between control and transgenic plants. Of these, five protein spots were successfully identified by MS/MS analysis.[57] Our proteomic analysis of pigeonpea seed should serve as a starting point to build a comprehensive pigeonpea proteome. It is important that proteomics research must be carried out from all organs of the plant to obtain a comprehensive pigeonpea proteome atlas. Also, the effect of biotic and abiotic stresses on the proteomic changes in different organs should also be investigated. Such a research effort should lead to the identification of unique proteins and metabolic pathways that could be exploited by researchers to improve the yield potential of pigeonpea. The availability of large-scale genomics resources along with emerging proteomics data should facilitate genomics-assisted breeding (GAB), leading to the development of improved pigeonpea cultivars that can perform well under marginal environments.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.jafc.7b00998.

Amino acid sequence alignments of 7S β-conglycinins and 11S glycinins; distribution of GO biological process and molecular function terms drawn from identified pigeonpea seed proteins; proteins identified from pigeonpea by MALDI-TOF-TOF-MS/MS analysis; hypothetical or uncharacterized proteins from pigeonpea (PDF)

■ AUTHOR INFORMATION

Corresponding Author

*(H.B.K.) Plant Genetics Research Unit, USDA-ARS, 108 Curtis Hall, University of Missouri, Columbia, MO 65211, USA. E-mail: Hari.Krishnan@ars.usda.gov.

ORCID

Hari B. Krishnan: 0000-0001-6437-3672

Notes

Mention of a trademark, vendor, or proprietary product does not constitute a guarantee or warranty of the product by the USDA and does not imply its approval to the exclusion of other products or vendors that may also be suitable.

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

We acknowledge the technical assistance of Farooq Khan.

■ REFERENCES

(1) Ratnaparkhe, M. B.; Gupta, V. S. In *Genome Mapping and Molecular Breeding in Plants: Pulses, Sugar and Tuber Crops*; Kole, C., Ed.; Springer: Berlin, Germany, 2007; pp 133−142.

(2) Varshney, R. K.; Chen, W.; Li, Y.; Bharti, A. K.; Saxena, R. K.; Schlueter, J. A.; et al. Draft genome sequence of pigeonpea (*Cajanus cajan*), an orphan legume crop of resource-poor farmers. *Nat. Biotechnol.* **2012**, *30*, 83−89.

DOI: 10.1021/acs.jafc.7b00998
J. Agric. Food Chem. 2017, 65, 4572−4581

(3) Salunkhe, D. K.; Chauvan, J. K.; Kadam, S. S. Pigeon pea as an important food source. *CRC Crit. Rev. Food Sci. Nutr.* **1986**, *23*, 103−145.
(4) Singh, U.; Jambunathan, R.; Saxena, K.; Subrahmanyam, N. Nutritional quality evaluation of newly developed high protein genotypes of pigeon pea (*Cajanus cajan* L.). *J. Sci. Food Agric.* **1990**, *50*, 201−209.
(5) Varshney, R. K.; Penmetsa, R. V.; Dutta, S.; Kulwal, P. L.; Saxena, R. K.; Datta, S.; et al. Pigeonpea genomics initiative (PGI): an international effort to improve crop productivity of pigeonpea (*Cajanus cajan* L.). *Mol. Breed.* **2010**, *26*, 393−408.
(6) Varshney, R. K. Exciting journey of 10 years from genomes to fields and markets: Some success stories of genomics-assisted breeding in chickpea, pigeonpea and groundnut. *Plant Sci.* **2016**, *242*, 98−107.
(7) Varshney, R. K.; Terauchi, R.; McCouch, S. R. Harvesting the promising fruits of genomics: applying genome sequencing technologies to crop breeding. *PLoS Biol.* **2014**, *12*, e1001883.
(8) Pazhamala, L.; Saxena, R. K.; Singh, V. K.; Sameerkumar, C. V.; Kumar, V.; Sinha, P.; et al. Genomics-assisted breeding for boosting crop improvement in pigeonpea (*Cajanus cajan*). *Front. Plant Sci.* **2015**, *6*, 50.
(9) Ramalingam, A.; Kudapa, H.; Pazhamala, L. T.; Weckwerth, W.; Varshney, R. J. Proteomics and metabolomics: two emerging areas for legume improvement. *Front. Plant Sci.* **2015**, *6*, 1116.
(10) Hummel, J.; Niemann, M.; Wienkoop, S.; Schulze, W.; Steinhauser, D.; Selbig, J. ProMEX: A mass spectral reference database for proteins and protein phosphorylation sites. *BMC Bioinf.* **2007**, *8*, 216.
(11) Wienkoop, S.; Staudinger, C.; Hoehenwarter, W.; Weckwerth, W.; Egelhofer, V. ProMEX: a mass spectral reference database for plant proteomics. *Front. Plant Sci.* **2012**, *3*, 125.
(12) Laemmli, U. K. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* **1970**, *227*, 680−685.
(13) Shewry, P. R.; Napier, J. A.; Tatham, A. S. Seed storage proteins: structures and biosynthesis. *Plant Cell* **1995**, *7*, 945−956.
(14) Singh, U.; Eggum, B. O. Factors affecting the protein quality of pigeonpea (*Cajanus cajan* L.). *Qual. Plant. - Plant Foods Hum. Nutr.* **1984**, *34*, 273−283.
(15) Saxena, K. B.; Sawargaonkar, S. L. Genetic enhancement of seed proteins in pigeonpea—methodologies, accomplishments, and opportunities. *Int. J. Sci. Res.* **2015**, *4*, 3−7.
(16) Gupta, S.; Sharma, S.; Gupta, A.; Chand, L. Biochemical studies of major storage proteins from the seeds of pigeonpea [*Cajanus cajan* (L.) Millsp.]. *Physiol. Mol. Biol. Plants* **2001**, *7*, 167−174.
(17) Gopala Krishnan, T.; Mitra, R. K.; Bhatia, C. R. Seed globulins of *Cajanus cajan*. *Plant Foods Hum. Nutr.* **1977**, *27*, 313−325.
(18) Krishna, T. G.; Bhatia, C. R. Vicilin from *Cajanus cajan* seeds. *Phytochemistry* **1985**, *24*, 2201−2203.
(19) Natarajan, S. S.; Krishnan, H. B.; Lakshman, S.; Garrett, W. M. An efficient extraction method to enhance analysis of low abundant proteins from soybean seed. *Anal. Biochem.* **2009**, *394*, 259−268.
(20) Singh, N.; Jain, N.; Kumar, R.; Jain, A.; Singh, N. K.; Rai, V. A comparative method for protein extraction and 2-D gel electrophoresis from different tissues of *Cajanus cajan*. *Front. Plant Sci.* **2015**, *6*, 606.
(21) Krishnan, H. B.; Nelson, R. L. Analysis of high protein soybean (*Glycine max*) accessions demonstrates the contribution of novel glycinin subunits. *J. Agric. Food Chem.* **2011**, *59*, 2432−2439.
(22) Krishnan, H. B.; Natarajan, S. S.; Mahmoud, A. A.; Nelson, R. L. Identification of glycinin and ß-conglycinin subunits that contribute to the increased protein content of high-protein soybean lines. *J. Agric. Food Chem.* **2007**, *55*, 1839−1845.
(23) Mooney, B.; Krishnan, H. B.; Thelen, J. J. High-throughput peptide masss fingerprinting of soybean seed proteins: automated workflow and utility of UniGene expressed sequence tag databases for protein identification. *Phytochemistry* **2004**, *65*, 1733−1744.
(24) 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*, 3174−3188.
(25) Balbuena, T. S.; He, R.; Salvato, F.; Gang, D. R.; Thelen, J. J. Large-scale proteome comparative analysis of developing rhizomes of the ancient vascular plant *Equisetum hyemale*. *Front. Plant Sci.* **2012**, *3*, 131.
(26) Nielsen, N. C. Soybean seed composition. In *Soybean: Genetics, Molecular Biology and Biotechnology*; Verma, D. P. S., Shoemaker, R. C., Eds.; CABI:Wallingford, UK, 1996.
(27) Krishnan, H. B. Biochemistry and molecular biology of soybean seed storage proteins. *J. New Seeds* **2000**, *2*, 1−25.
(28) Gayler, K. R.; Sykes, G. E. Effects of nutritional stress on the storage proteins of soybeans. *Plant Physiol.* **1985**, *78*, 582−585.
(29) Staswick, P. E.; Hermodson, M. A.; Nielsen, N. C. Identification of the cystines which link the acidic and basic components of the glycinin subunits. *J. Biol. Chem.* **1984**, *259*, 13431−13435.
(30) Krishna, T. G.; Bhatia, C. R. γ-Protein, a sulphur amino acid rich protein from pigeon pea (*Cajanus cajan* (L.) Millsp.). *J. Biosci.* **1986**, *10*, 57−65.
(31) Yamauchi, F.; Sato, K.; Yamagishi, T. Isolation and partial characterization of a salt-extractable globulin from soybean seeds. *Agric. Biol. Chem.* **1984**, *48*, 645−650.
(32) Watanabe, Y.; Barbashov, S. F.; Komatsu, S.; Hemmings, A. M.; Miyagi, M.; Tsunasawa, S.; Hirano, H. A peptide that stimulates phosphorylation of the plant insulin-binding protein. Isolation, primary structure and cDNA cloning. *Eur. J. Biochem.* **1994**, *224*, 167−172.
(33) Kagawa, H.; Yamauchi, F.; Hirano, H. Soybean basic 7S globulin represents a protein widely distributed in legume species. *FEBS Lett.* **1987**, *226*, 145−149.
(34) Ryan, C. A. Protease inhibitors in plants: genes for improving defenses against insects and pathogens. *Annu. Rev. Phytopathol.* **1990**, *28*, 425−449.
(35) De Leo, F.; Volpicella, M.; Licciulli, F.; Liuni, S.; Gallerani, R.; Ceci, L. R. PLANT-PIs: a database for plant protease inhibitors and their genes. *Nucleic Acids Res.* **2002**, *30*, 347−348.
(36) Jofuku, K. D.; Goldberg, R. B. Kunitz trypsin inhibitor genes are differentially expressed during the soybean life cycle and in transformed tobacco plants. *Plant Cell* **1989**, *1*, 1079−1093.
(37) Cheva-Isarakul, B. Pigeon pea as a ruminant feed. *Asian−Australas. J. Anim. Sci.* **1992**, *5*, 549−558.
(38) Tangtaweewipat, S.; Elliott, R. Nutritional value of pigeon pea (*Cajanas cajan*) meal in poultry diets. *Anim. Feed Sci. Technol.* **1989**, *25*, 123−135.
(39) Singh, F.; Diwakar, B. *Nutritive Value and Uses of Pigeonpea and Groundnut*; Skill Development Series 14; International Crops Research Institute for the Semi-Arid Tropics: Patancheru, Andhra Pradesh, India, 1993.
(40) Mekbungwan, A.; Yamauchi, K. Growth performance and histological intestinal alterations in piglets fed dietary raw and heated pigeon pea seed meal. *Histol. Histopathol.* **2004**, *19*, 381−389.
(41) Azarkovich, M. I. *Stress-Induced Proteins in Recalcitrant Seeds during Deep Dormancy and Early Germination, Abiotic and Biotic Stress in Plants—Recent Advances and Future Perspectives*; Shanker, A., Ed.; InTech: Rijeka, Croatia, 2016; DOI: 10.5772/61958.
(42) Sarkar, N. K.; Kim, Y. K.; Grover, A. Rice sHsp genes: genomic organization and expression profiling under stress and development. *BMC Genomics* **2009**, *10*, 393.
(43) Waters, E. R. The evolution, function, structure, and expression of the plant sHSPs. *J. Exp. Bot.* **2013**, *64*, 391−403.
(44) Alscher, R. G.; Hess, J. L. *Antioxidants in Higher Plants*; CRC Press: Boca Raton, FL, USA, 1993.
(45) Grene, R. Oxidative stress and acclimation mechanisms in plants. In *Arabidopsis Book*; American Society of Plant Biologists: Rockville, MD, USA, 2002; e0036, DOI: 10.1199/tab.0036.1.
(46) Alscher, R. G.; Erturk, N.; Heath, L. S. Role of superoxide dismutases (SODs) in controlling oxidative stress in plants. *J. Exp. Bot.* **2002**, *53*, 1331−1341.
(47) Aalen, R. B. Peroxiredoxin antioxidants in seed physiology. *Seed Sci. Res.* **1999**, *9*, 285−295.

DOI: 10.1021/acs.jafc.7b00998
J. Agric. Food Chem. 2017, 65, 4572−4581

(48) Battaglia, M.; Covarrubias, A. A. Late embryogenesis abundant (LEA) proteins in legumes. *Front. Plant Sci.* **2013**, *4*, 190.
(49) Dure, L. A repeating 11-mer amino acid motif and plant desiccation. *Plant J.* **1993**, *3*, 363−369.
(50) Wise, M. J. LEAping to conclusions: a computational reanalysis of late embryogenesis abundant proteins and their possible roles. *BMC Bioinf.* **2003**, *4*, 52−10.
(51) Bartels, D.; Salamini, F. Desiccation tolerance in the resurrection plant *Craterostigma plantagineum*. A contribution to the study of drought tolerance at the molecular level. *Plant Physiol.* **2001**, *127*, 1346−1353.
(52) Wurtele, E. S.; Wang, H.; Durgerian, S.; Nikolau, B. J.; Ulrich, T. H. (1993) Characterization of a gene that is expressed early in somatic embryogenesis of Daucus carota. *Plant Physiol.* **1993**, *102*, 303−312.
(53) Litts, J. C.; Colwell, G. W.; Chakerian, R. L.; Quatrano, R. S. The nucleotide sequence of a cDNA clone encoding the wheat Em protein. *Nucleic Acids Res.* **1987**, *15*, 3607−3618.
(54) Baker, J.; Steele, C.; Dure, L. Sequence and characterization of 6 Lea proteins and their genes from cotton. *Plant Mol. Biol.* **1988**, *11*, 277−291.
(55) Seki, M.; Narusaka, M.; Ishida, J.; Nanjo, T.; Fujita, M.; Oono, Y.; Kamiya, A.; Nakajima, M.; Enju, A.; Sakurai, T. Monitoring the expression profiles of 7000 *Arabidopsis* genes under drought, cold and high-salinity stresses using a full-length cDNA microarray. *Plant J.* **2002**, *31*, 279−292.
(56) Hanin, M.; Brini, F.; Ebel, C.; Toda, Y.; Takeda, S.; Masmoudi, K. Plant dehydrins and stress tolerance: versatile proteins for complex mechanisms. *Plant Signaling Behav.* **2011**, *6*, 1503−1509.
(57) Mishra, P.; Singh, S.; Rathinam, M.; Nandiganti, M.; Kumar, N. R.; Thangaraj, A.; Thimmegowda, V.; Krishnan, V.; Mishra, V.; Jain, N.; Rai, V.; Pattanayak, D.; Sreevathsa, R. Comparative proteomic and nutritional composition analysis of independent transgenic pigeon pea seeds harboring *cry1AcF* and *cry2Aa* genes and their nontransgenic counterparts. *J. Agric. Food Chem.* **2017**, *65*, 1395−1400.

DOI: 10.1021/acs.jafc.7b00998
J. Agric. Food Chem. 2017, 65, 4572−4581