

Publications

Publications:

1. **Shekhar, S.**; Mishra, D.; Buragohain, A.K.; Chakraborty, S; Chakraborty, N. Comparative analysis of phytochemicals and nutrient availability in two contrasting cultivars of sweet potato (*Ipomoea batatas* L.). Food Chem., 2015, 173: 957-965.
2. **Shekhar, S.**; Agrawal, L.; Buragohain, A.K.; Datta, A.; Chakraborty, S.; Chakraborty, N. (2013) Genotype independent regeneration and agrobacterium-mediated genetic transformation of sweet potato (*Ipomoea batatas* L.). Plant Tissue Cult.Biotech., 2013, 23: 87-100.
3. Agrawal, L.; Narula, K.; Basu, S.; **Shekhar, S.**; Ghosh, S.; Datta, A.; Chakraborty, N.; Chakraborty, S. Comparative proteomics reveals a role for seed storage protein, AmA1 in cellular growth, development and nutrient accumulation. J. Proteome Res., 2013, 12, 4904–4930.
4. Subba, P.; Kumar, R.; Gayali, S.; **Shekhar, S.**; Parveen, S.; Pandey, A.; Datta, A.; Chakraborty, S.; Chakraborty, N. Characterization of the nuclear proteome of a dehydration-sensitive cultivar of chickpea and comparative proteomic analysis with a tolerant cultivar. Proteomics, 2013, 13, 1973-1992.
5. Chakraborty, S.; Chakraborty, N.; Agrawal, L.; Ghosh, S.; Narula, K.; **Shekhar, S.**; Naik, P.S.; Pande, P.C.; Chakraborti, S.K.; Datta, A. Next-generation protein-rich potato expressing the seed protein gene AmA1 is a result of proteome rebalancing in transgenic tuber. Proc. Natl. Acad. Sci. USA, 2010, 107, 17533-17538.



Comparative analysis of phytochemicals and nutrient availability in two contrasting cultivars of sweet potato (*Ipomoea batatas* L.)



Shubhendu Shekhar^{a,b}, Divya Mishra^a, Alak Kumar Buragohain^b, Subhra Chakraborty^{a,*}, Niranjan Chakraborty^{a,*}

^a National Institute of Plant Genome Research, Jawaharlal Nehru University Campus, Aruna Asaf Ali Marg, New Delhi 110067, India

^b Department of Molecular Biology and Biotechnology, Tezpur University, Assam, India

ARTICLE INFO

Article history:

Received 10 June 2014

Received in revised form 11 September 2014

Accepted 19 September 2014

Available online 30 October 2014

Keywords:

Anthocyanins

Flavonoids

Nutritional diversity

Phenolics

Phytochemicals

Proteome profiling

Storage stress

Transcript abundance

ABSTRACT

Sweet potato ranks as the world's seventh most important food crop, and has major contribution to energy and phytochemical source of nutrition. To unravel the molecular basis for differential nutrient availability, and to exploit the natural genetic variation(s) of sweet potato, a series of physiochemical and proteomics experiment was conducted using two contrasting cultivars, an orange-fleshed sweet potato (OFSP) and a white-fleshed sweet potato (WFSP). Phytochemical screening revealed high percentage of carbohydrate, reducing sugar and phenolics in WFSP, whereas OFSP showed increased levels of total protein, flavonoids, anthocyanins, and carotenoids. The rate of starch and cellulose degradation was found to be less in OFSP during storage, indicating tight regulation of gene(s) responsible for starch-degradation. Comparative proteomics displayed a cultivar-dependent expression of proteins along with evolutionarily conserved proteins. These results suggest that cultivar-specific expression of proteins and/or their interacting partners might play a crucial role for nutrient acquisition in sweet potato.

© 2014 Elsevier Ltd. All rights reserved.

1. Introduction

Sweet potato is a dicotyledonous species belonging to the morning glory family, Convolvulaceae. It is a perennial crop which serves as one of the major sources of food, animal feed and industrial raw materials. It has a significant contribution as energy supplement and phytochemical source of nutrition. It is widely cultivated in the tropics, subtropics and even in some temperate zones of the developing world (Ahn et al., 2010). Contribution of sweet potato towards health is acknowledged due to high nutrient content and its anti-carcinogenic and cardiovascular disease preventing properties (Teow et al., 2007). Almost all cultivars of sweet potato are excellent source of vitamin C, B₂, B₆ and E, as well

as dietary fibre, potassium, copper, manganese and iron, and are low in fat and cholesterol. Sweet potato is the seventh most important crop in terms of global production. In developing countries, it ranks third in value of production and fifth in caloric contribution to the human diet (Mohan, 2011). Based on the variability in flesh colour, white-fleshed to cream-fleshed sweet potato is widespread in the Pacific, whereas the yellow- to orange-fleshed sweet potato is predominantly found in the United States (Mohan, 2011).

Improvement in nutritional value for crop species has been a major long-term goal of plant breeding programs. Humans require a diverse and nutritionally well-balanced diet for the maintenance of optimal health. Therefore global scientific research is focused upon improving the nutritional qualities of food crops. This has become an increasingly critical issue in developing countries, particularly in India, where plants are the major primary nutritional support in the human diet and animal feed. Diversity in nutritional status amongst different varieties of crop has considerable impact on nutritional adequacy. Different cultivars of the same crop adapt very differently in response to environmental cues, even though they are accustomed to precisely the same environment. These altogether may invariably affect the nutritional status of closely related genotypes of a species (Huang, Tanudjaja, & Lum, 1999;

Abbreviations: BCE, β -carotene equivalents; GAE, gallic acid equivalent; QE, quercetin equivalent; OFSP, orange-fleshed sweet potato; TCC, total carotenoid contents; TFC, total flavonoids content; TPC, total phenolic content; TAC, total anthocyanin content; WHC, water holding capacity; WFSP, white-fleshed sweet potato.

* Corresponding authors. Tel.: +91 11 26735186; fax: +91 11 26741658 (S. Chakraborty). Tel.: +91 11 26735178; fax: +91 11 26741658 (N. Chakraborty).

E-mail addresses: subhrac@hotmail.com (S. Chakraborty), nchakraborty@nipgr.ac.in (N. Chakraborty).

Kennedy & Burlingame, 2003). Therefore, to address the above ambiguities, it is pertinent to unravel the extent and basis of nutrient diversity within the species, particularly amongst underutilized crops. The information(s) available on the diversity amongst these crops are very little and fragmented. Information on cultivar-specific nutrient content can be applied in breeding programs, which can significantly reduce the dependence on transgenic modifications of a cultivar. Hence, there is a need to exploit the existing potential available in the nutritionally rich germplasm of sweet potato for improved nutritional implications. Consequently, the objective of this study was to investigate the genotypic variability with regard to the biochemical composition and differential expression of gene(s). These might be anticipated to cause differences in nutritive value of two closely related cultivars of sweet potato.

2. Materials and methods

2.1. Maintenance of plants, tissue harvesting and preparation of samples

The slips of OFSP (*Ipomoea batatas* (L.) Lam. cv. SP-6) were obtained from the Central Tuber Crops Research Institute (CTCRI), Thiruvananthapuram, while those of WFSP, a close relative, were obtained from Tezpur, Assam, India. The cultivars were grown in parallel in four replicated plots of experimental field under identical environmental conditions and field management to eliminate the environmental influence(s), if any. Slips were planted in three-row plots 8.0 × 3.6 m, with 1.2 m spacing between beds and 0.5 m between plants. Mature tubers were harvested and pooled from eight different plants (two from each experimental plot) to normalise the growth and developmental effects.

2.2. Assessment of agrophysiological traits

The leaf area of the 3–5th leaf from the shoot apex was measured and calculated using LeafJ with ImageJ software (<http://rsb.info.nih.gov/ij/>) as described earlier (Maloof, Nozue, Mumbach, & Palmer, 2013). Rate of photosynthesis was quantified with a portable photosynthesis measurement system (GFS3000; Waltz). The photosynthetic potential was determined on the basis of single leaf measurements of 5–7 leaves from each plant, and was evaluated after 8–10 weeks of plantation under standard atmospheric (360 ppm CO₂) and light conditions (750 μmol m⁻²s⁻¹). A fully expanded second or third leaf from the shoot apex was held in the chamber for 2–3 min until the rate of photosynthesis was in a steady-state condition. Identical sized mature tubers were harvested from both the cultivars. The average weight and diameter were measured.

2.3. Evaluation of tuber colour difference

The colour of tubers was determined colorimetrically using high-performance colour measurement spectrophotometer (HunterLab Ultrascan Vis) coupled with EasyMatch QC software. The measurements were performed over 360–780 nm wavelengths, as per manufacturer's recommendation.

2.4. Biochemical and proximate analysis

For proximate analyses, tubers of both the cultivars were washed, peeled individually, sliced into smaller pieces and oven-dried at 40 °C for 18 h. Each sample was then grounded into powder using a laboratory-scale grinder. Total moisture, ash and crude fibre were determined in triplicates according to the

standard method of the Association of Official Analytical Chemists (AOAC, 2000). Moisture content was determined gravimetrically by the weight loss after drying the samples in a hot-air oven until constant weight was achieved. Total ash content was measured after igniting in a muffle furnace at 550–600 °C to acquire a constant weight. Crude fibres were determined according to the standard AOAC method. In brief, 2 g dried samples were boiled in 1.25% H₂SO₄ for 30 min. The samples were then filtered through gooch crucible and the residues obtained were subsequently washed with the boiling water repeatedly to eliminate the acid content. Residues were again boiled with 1.25% NaOH for 30 min, filtered and dried. The dried samples were placed in a crucible and ignited in the furnace at 600 ± 15 °C for 2–3 h and then allowed to cool. The weight of the residues was recorded and the percent crude fibre content was calculated via the loss in weight on and before ignition. The total protein was quantified by micro-Kjeldahl method and protein was calculated from nitrogen content multiplied by a factor of 6.25.

To carry out the biochemical analyses, tubers were lyophilized and grounded into fine powders. Total carbohydrate content in both the cultivars was determined by Anthrone method, using standard protocol with D-glucose as standard. Reducing sugar was determined by Nelson-Somogyi method (Sadasivam & Manickam, 1992).

2.5. Water holding capacity (WHC)

WHC was determined using a method previously described (Robertson et al., 2000) with few modifications. Briefly, 1 g of lyophilized tissue, in triplicate, was hydrated in 30 ml Milli-Q water containing 0.02% azide as a bacteriostat. Samples were centrifuged at 3000 × g for 20 min at room temperature after equilibrating it for 18 h. The supernatant was removed carefully and residual water was drained out. Sample fresh weight was recorded before drying. WHC was calculated as the amount of water retained by the pellet (g/g dry weight) after residual dry weight.

2.6. Evaluation of starch and cellulose degradation

The tubers were harvested from eight different plants (two from each experimental plot). To evaluate degradation of starch and cellulose during storage, harvested tubers were stored at room temperature and measured at every 15-d interval. Total starch content was determined by Anthrone method using standard method with few modifications (Sadasivam & Manickam, 1992). Glucose content in the samples was extrapolated using D-glucose as standard and the starch content was determined by multiplying it with a factor of 0.9. Total cellulose content was determined by modifying the method of acidolysis with acetic/nitric reagent followed by the treatment with H₂SO₄. Acid hydrolysed cellulose was then quantified by Anthrone method using standard procedures as described by Sadasivam and Manickam (1992).

2.7. Determination of phytochemicals and carotenoids

To determine the phytochemicals, the extracts were prepared from lyophilized tuber samples. Total phenolic content (TPC) was determined by the method based on oxidation–reduction reaction by Folin–Ciocalteu reagent using gallic acid as a standard.

Total flavonoids content (TFC) was determined by colorimetric method as described by Kim, Jeong, and Lee (2003). The alcoholic tuber extract was diluted to a final volume of 5 ml with Milli-Q water and 0.3 ml 5% NaNO₂. The mixture was then incubated at room temperature for 5 min. Subsequently, 0.3 ml 10% AlCl₃ was added followed by the addition of 2 ml 1 N NaOH after 6 min. The solution was mixed thoroughly and the absorbance was

measured at 510 nm against Milli-Q water as blank. Quercetin was used as standard and flavonoid contents were measured as quercetin equivalent.

Total anthocyanin content (TAC) was determined by following the procedures of Mancinelli, Hoff, and Cottrell (1988) with few modifications. Approximately 0.5 g of tuber sample was extracted with 20 ml of acidified methanol (1% v/v HCl) for 24 h in dark at 4 °C with occasional shaking. Samples were then centrifuged at 5,000 × g for 15 min at 4 °C. The supernatant was recovered and diluted to an appropriate concentration. The absorbance was read at 530 and 657 nm and total anthocyanins was estimated using the formula, $A_{530} - 0.25A_{657}$ to recompense the absorption of chlorophyll degradation products at A_{530} .

Total carotenoid contents (TCC) were determined by the modified method of Koala et al. (2013). A known weight of dry powder was extracted with acetone-hexane (50:50, v/v) for TCC assay. Extracts were stirred vigorously, filtered and stored at 4 °C in dark. Absorbance of suitably diluted extracts was measured at 455 nm with β -carotene as standard.

2.8. qRT-PCR analysis of flavonoid pathway genes

The expression of key flavonoid pathway genes, chalcone synthase (*CHS*), chalcone flavanone isomerase (*CHI*), flavanone 3-hydroxylase (*F3H*), dihydroflavonol 4-reductase (*DFR*), anthocyanidin synthase (*ANS*) and UDP-glucose flavonoid 3-O-glucosyl transferase (*UFGT*) was analysed by qRT-PCR. Total RNA was isolated by TriPure Isolation Reagent (Roche Diagnostics) following manufacturer's recommendation. cDNAs were prepared using SuperScript[®] VILO[™] cDNA Synthesis Kit (Invitrogen). qRT-PCR analysis was performed using gene-specific primers: *CHSF* 5'-TGGG CCTGGGCTTACAATC-3' and *CHSR* 5'-CTTGGGCGGGCTTAA-3' (*CHS*); *CHIF* 5'-GCGGAGGAGTTGACGGATT-3' and *CHIR* 5'-TTCTCAA AGGGACCCGTAACG-3' (*CHI*); *F3HF* 5'-TATTCAAGGTGGCCGGACAA-3' and *F3HR* 5'-CAGCAGTTTGCATGCCAAGT-3' (*F3H*); *DFRF* 5'-TTAT CGGCTCTGGTGGT-3' and *DFRR* 5'-TGTCCGCTTCGGTAGTTC-3' (*DFR*); *ANSF* 5'-GCGTCCCGAACTCCATCAT-3' and *ANSR* 5'-CTTGCCG TTGCTGAGGATCT-3' (*ANS*); *UFGTF* 5'-GCCGCCACTCCAAACG-3' and *UFGTR* 5'-CATTCTGGGATTACTTTCAGCTT-3' (*UFGT*). The expression data were normalised against the expression levels of actin as an internal control using actin specific primers *ActinF* 5'-CTCCCTAATGAGTGTGATGTGAT-3' and *ActinR* 5'-GAGCCCCATGA GAACATTACCA-3'. All the primers were designed using the Primer Express Software v3.0.1. The qRT-PCR was performed, in two biological and three technical replicates, by an ABI PRISM 7700 Sequence Detection System (Applied Biosystems) using SYBR green dye. Mean of Ct values for target and endogenous control was considered for calculating the relative quantitation (RQ) value using comparative Ct ($2^{-\Delta\Delta Ct}$) method.

2.9. 2-DE analysis

Proteins were isolated from the mature tubers as described earlier (Jiang, Chen, Tao, Wang, & Zhang, 2012) with few modifications. The tubers were peeled, sliced and grounded to fine powder with liquid nitrogen and suspended in acetone containing 10% trichloroacetic acid and 0.07% β -mercaptoethanol, and precipitated overnight at -20 °C. The precipitates were recovered by centrifugation at 10,000 × g at 4 °C for 10 min and washed twice with 0.07% β -mercaptoethanol in acetone and then dried. The dried pellets were resuspended in resuspension buffer (100 mM Tris-HCl buffer (pH 8.0), 10 mM EDTA, and 30% sucrose, 2% CHAPS, 2% SDS, 10 mM PMSF and 2% β -mercaptoethanol), and further extracted with an equal-volume of Tris-saturated phenol. Proteins were precipitated overnight at -20 °C by the addition of 0.1 M ammonium acetate in methanol, 4 times the volume of sample.

The pellets were washed 2–3 times with 0.1 M ammonium acetate in methanol and then with 0.07% β -mercaptoethanol in acetone. The protein pellets were then solubilised using IEF sample buffer (8 M urea, 2 M thiourea and 4% w/v CHAPS) and the concentration of protein was determined using 2-D Quant kit (GE Healthcare). Aliquots of 250 μ g protein were diluted with 2D rehydration buffer (8 M urea, 2 M thiourea, 4% w/v CHAPS, 20 mM DTT, 0.5% v/v pharmalyte (pH 4–7) and 0.05% w/v bromophenol blue) followed by rehydration of IPG strips (13 cm; pH 4–7) with 250 μ l of the buffer. Electrofocusing was performed using the IPGphor system (GE Healthcare) at 20 °C for 32,000 VhT. The focused strips were subjected to reduction with 1% w/v DTT in 10 ml of equilibration buffer (6 M urea, 50 mM Tris-HCl (pH 8.8), 30% v/v glycerol and 2% w/v SDS), followed by alkylation with 2.5% w/v iodoacetamide in the same buffer. The strips were then loaded on 12.5% polyacrylamide gels for SDS-PAGE. To reduce gel-to-gel variation, each protein preparation was analysed on at least three parallel 2-D gels, representing three technical replicates. The electrophoresed proteins were stained with Silver Stain Plus Kit (Bio-Rad).

2.10. Image acquisition and comparative proteomics analysis

The gel images were scanned using the Fluor-S Multimager system (Bio-Rad), and processed and analysed with PDQuest gel analysis software version 7.2.0 (Bio-Rad). The pI and experimental molecular mass were calculated from the scanned images using standard molecular mass marker proteins. To compare spots across gels, a match set representing a standard image of three replicates, representing three biological replicates, was created for each sample. Each spot on the standard gel was quantified by several criteria for being consistent in size and shape for all the replicate gels and being within the linear range of detection. The spots detected by the software program were further verified manually to eliminate any possible artifacts, such as gel background or streaks. In addition to quantify scores, the PDQuest software was used to assign quality scores to each gel spot. The spots with a quality score less than 30 were eliminated from further analysis. The high-quality spot quantities were used to calculate the mean value for a given spot, and the value was used as the spot quantity on the standard gel. The correlation coefficient, representing the association between the spot intensities on replicates, was maintained at a minimum of 0.8 between gel images. The spot densities on the standard gel were normalised against the total density in the gel image. To facilitate the comparison of the standard gels for each of the cultivars, the spot volumes were further normalised using three unaltered protein spots across all the gels.

2.11. Statistical analysis

Statistical significance of the data was analysed by the unpaired student's *t*-test method using GraphPad Prism 5. $P < 0.05$ was considered to be statistically significant, and the results are expressed as mean \pm SE.

3. Results and discussion

3.1. Genetic diversity and assessment of agrophysiological traits

To investigate the extent of nutritional miscellany between OFSP and WFSP, a comparative phytochemical analysis was performed. Though the cultivars were grown side-by-side in the experimental field under the same environmental conditions, genetic differences between them was evident at morphological level (Fig. 1A and B). The difference in average yield and weight of tubers was quite distinct. Average weight and diameter of the

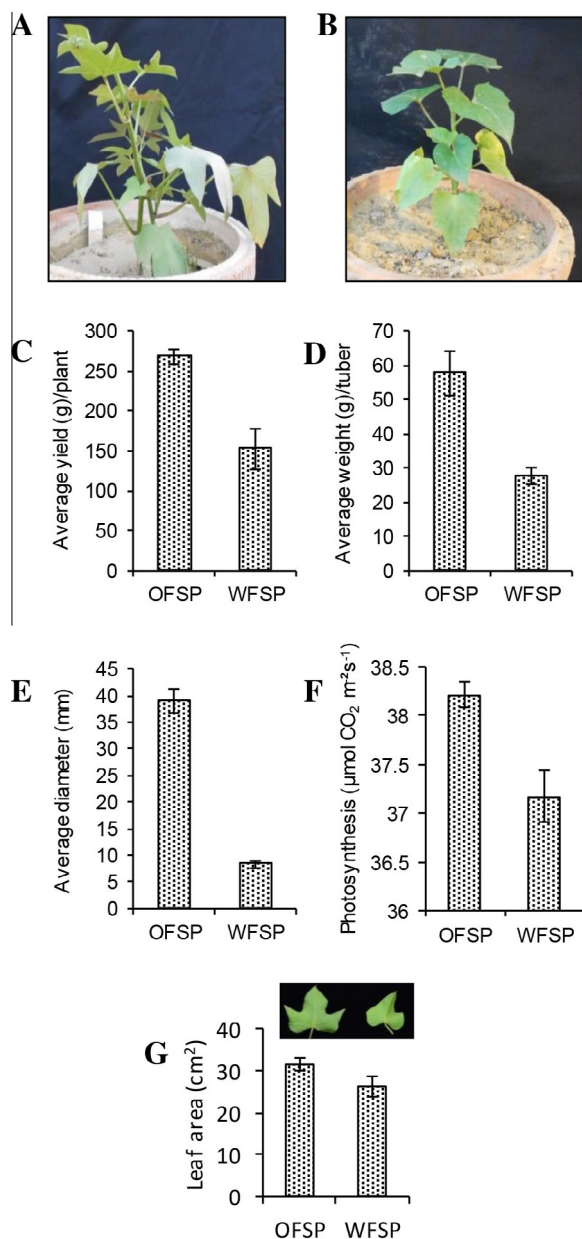


Fig. 1. Comparative analyses of agrophysiological traits. Representative photographs displaying morphological differences in aerial parts of OFSP (A) and WFSP (B). The variations are also shown in average yield (C), weight (D) and diameter (E). Each bar indicates the mean values \pm SE in triplicates. Photosynthetic rates (F) and leaf area (G) along with representative photograph of leaves from OFSP and WFSP were compared. The experiments were carried in five measurements and the average mean values \pm SE were plotted.

tubers of OFSP was found to be more than two- and fourfold higher when compared with WFSP. Moreover, the average yield was about twofold higher in OFSP (Fig. 1C–E). Increased tuber yield is closely associated with the photosynthetic carbon metabolism, which is considered to be a crucial factor for plant growth and productivity (Sweetlove, Kossmann, Riesmeier, Trethewey, & Hill, 1998). We therefore investigated the photosynthetic efficiency of both the cultivars. The photosynthetic CO_2 fixation was slightly higher in OFSP as compared to WFSP (Fig. 1F).

Leaf area is an important determinant of photosynthetic efficiency, exhibiting a strong correlation between them (Horsley & Gottschalk, 1993). There was a positive correlation between leaf size and photosynthetic efficiency as the leaf area (Fig. 1G) and

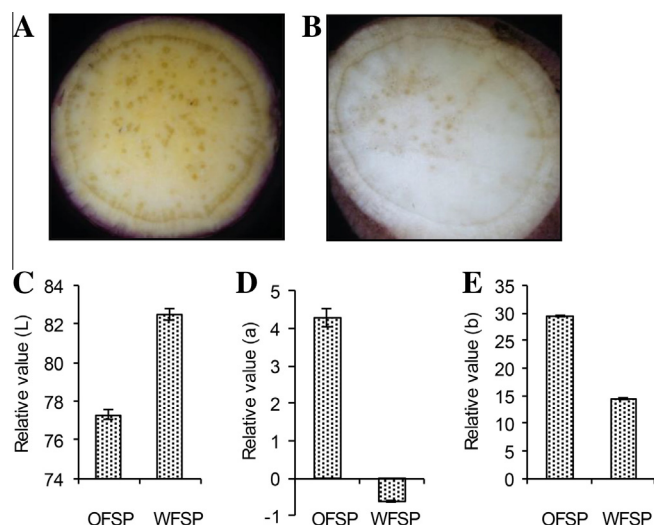


Fig. 2. Evaluation of colour differences. Cross sections of tubers displaying colour differences in OFSP (A), and WFSP (B). The relative Hunter 'L' 'a' 'b' values were determined (C–E). Data represent mean values \pm SE of three measurements.

photosynthetic CO_2 fixation (Fig. 1F) showed a proportional increase in OFSP.

3.2. Evaluation of colour difference

Hunter 'L' 'a' 'b' values demonstrates distinct colours in two sweet potato cultivars which corroborated the visual appearance of the tuber flesh (Fig. 2A and B). The 'L' (white) value for WFSP (82.51) was significantly higher than that of OFSP (77.31) (Fig. 2C). However, intense orange colour of the OFSP was confirmed by its higher 'a' and 'b' values when compared with WFSP. The 'a' (red) values of OFSP (4.28) was significantly higher ($p < 0.05$) than that of WFSP (-0.6) (Fig. 2D). Furthermore, the 'b' (yellow) values varied a remarkable twofold higher in OFSP (29.59) than that of WFSP (14.46) (Fig. 2E). The higher positive 'a' and 'b' values suggest that yellow-orange colour of OFSP might be due to the higher carotenoid pigments. A positive correlation have been earlier established in carotenoid contents and 'a' and 'b' colour space values in different orange-fleshed cultivars (Ameny & Wilson, 1997).

3.3. Differences in biochemical and proximate compositions

The proximate and biochemical analysis imparted a better understanding of the commonalities and diversities between the cultivars studied. The proximate analysis of the cultivars revealed no significant differences of moisture, ash and fibre contents ($p > 0.05$). Nevertheless, the analysis displayed high moisture content of 71.42% and 68.93% in OFSP and WFSP, respectively. These values were found to be well within the range of 60–84%, as reported in many sweet potato cultivars (Collins & Walter, 1985). The high-moisture content is an indicator of low dry matter and lower storage quality in tubers. The moisture content amongst cultivars might be attributed to the difference in their genetic composition and agro climatic conditions (Rose & Vasanthakalam, 2011). The ash and fibre contents were higher in OFSP (4.7% and 2.35%, respectively) as compared to WFSP (4.46% and 2.31%, respectively) (Supplementary Table S1). Interestingly, total protein content (4.85 g/100 g) of OFSP was considerably higher than that of WFSP (2.93 g/100 g). These results suggest that OFSP contain more nitrogenous substances as compared to WFSP (Fig. 3A). A previous study had reported a variable range of protein

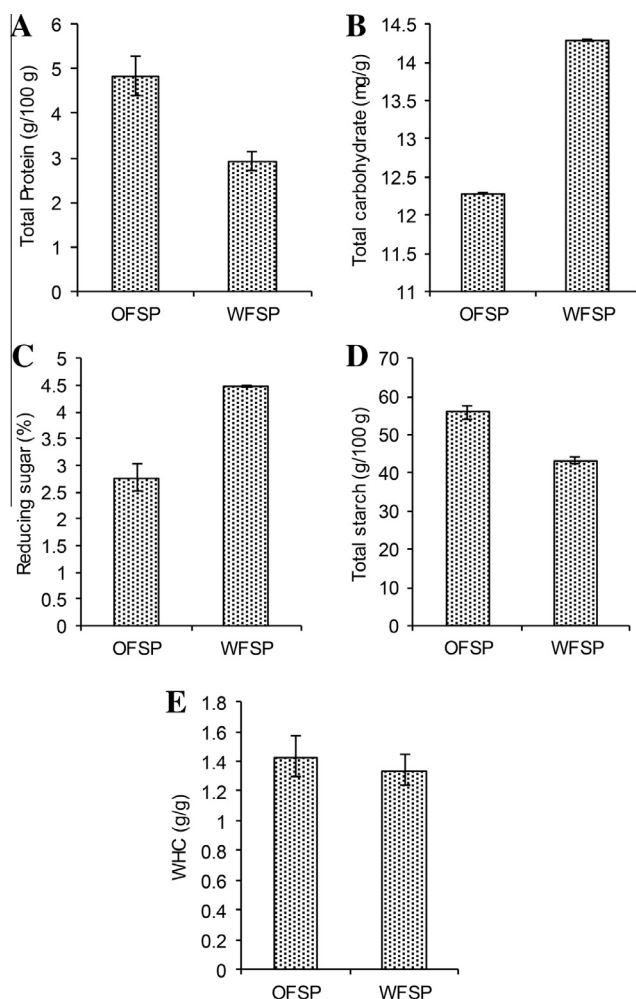


Fig. 3. Comparative biochemical analyses of OFSP and WFSP. Total protein (A), total carbohydrate (B), reducing sugar (C), and starch content (D) were determined on dry basis. The quantitative determination of WHC is shown (E). Data represent mean values \pm SE of three independent measurements.

(2.95–7.1985 g/100 g) in different sweet potato cultivar (Ravindran, Ravindran, Sivakanesan, & Rajaguru, 1995). In general, sweet potato is considered as a low protein crop ranging from 1% to > 10% (Bovell-Benjamin, 2007). The variations in protein content in sweet potato have been attributed to their genetic background and environmental niche (Walter, Collins, & Purcell, 1984). Several earlier reports revealed that the increase in protein biosynthesis is a consequence of increased rate of photosynthesis, which is eventually a crucial factor for higher yield (Agrawal et al., 2013; Chakraborty et al., 2010). It is increasingly clear that the tuberization is a complex process involving various metabolic cues, leading to massive accumulation of starch and proteins. Tubers, the storage organs, act as a sink and compete for the available photoassimilates. The allocation of photoassimilates is an important criterion for plant productivity, the harvest index. In tuberous crops, a higher harvest index indicates an efficient diversion of photoassimilates sink (Lemoine et al., 2013; Rukundo, Shimelis, Laing, & Gahakwa, 2013). Intriguingly, increase in the protein content via transgene approach has also showed a profound effect on the rate of photosynthesis (Agrawal et al., 2013; Chakraborty et al., 2010). Our study revealed a strong correlation amongst the leaf areas, photosynthetic rate and total protein content of OFSP.

In most sweet potato cultivars, the percentage of dry matter is reported up to 44%, out of which approximately 90% is

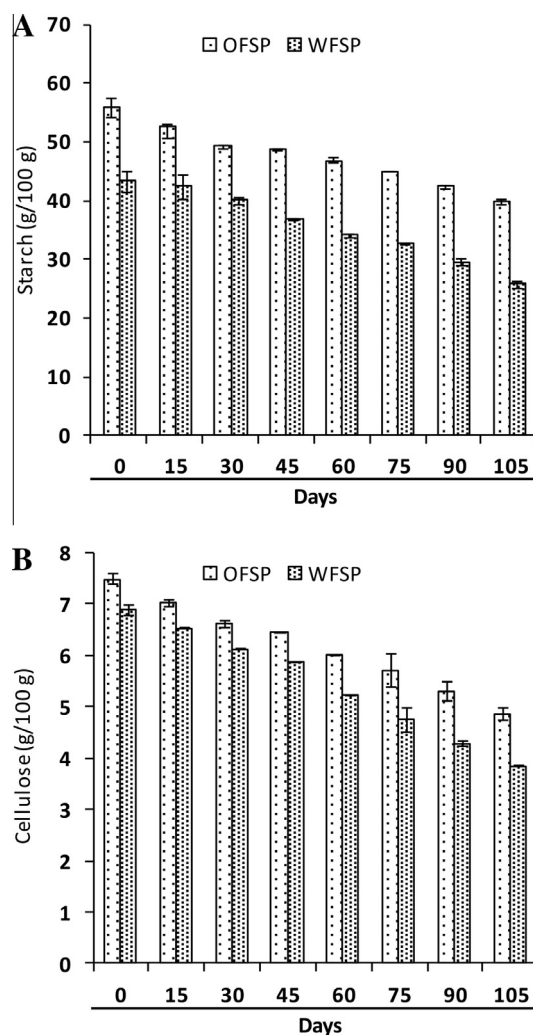


Fig. 4. Evaluation of storage stress response. Tubers were stored at room temperature and sampled at every 15 d interval until 105-d. The starch (A) and cellulose (B) contents were measured on dry basis at a succession of storage time points in both the cultivars. Data represent mean values \pm SE of three measurements.

carbohydrate (Woolfe, 1992). The present study revealed a considerably higher carbohydrate content in WFSP (14.28 mg/g) as compared to OFSP (12.29 mg/g) (Fig. 3B). Several earlier reports proclaimed a dynamic range of total carbohydrate. Ravindran et al. (1995) reported carbohydrate content in a range of 3.74–9.89%. However, Emmanuel, Vasanthakalam, Ndirigwe, and Mukwantali (2012) reported 7.65% and 8.7% in orange-fleshed and yellow-fleshed cultivars, respectively but 65% and 74%, respectively on dry basis. The reducing sugar content was also higher in WFSP (4.47%) than that of OFSP (2.76%) (Fig. 3C); however, starch content was higher of OFSP (55.83 g/100 g) as compared to WFSP (43.25 g/100 g) (Fig. 3D). The comparative analysis of different cultivars showed a variable range of starch content *i.e.* 33% and 64% in various cultivars on the dry basis (Senanayake, Ranaweera, Gunaratne, & Bamunuarachchi, 2013). The disparity between the starch contents and the status of carbohydrate and reducing sugar in the both cultivars suggests a possible degradation of starch and cellulose, presumably via hydrolysis of starch and conversion into reducing sugars during storage (Rose & Vasanthakalam, 2011).

3.4. Quantitative determination of water holding capacity

WHC signifies the amount of water retained by a known weight of the sample. It is related to the protein and carbohydrate

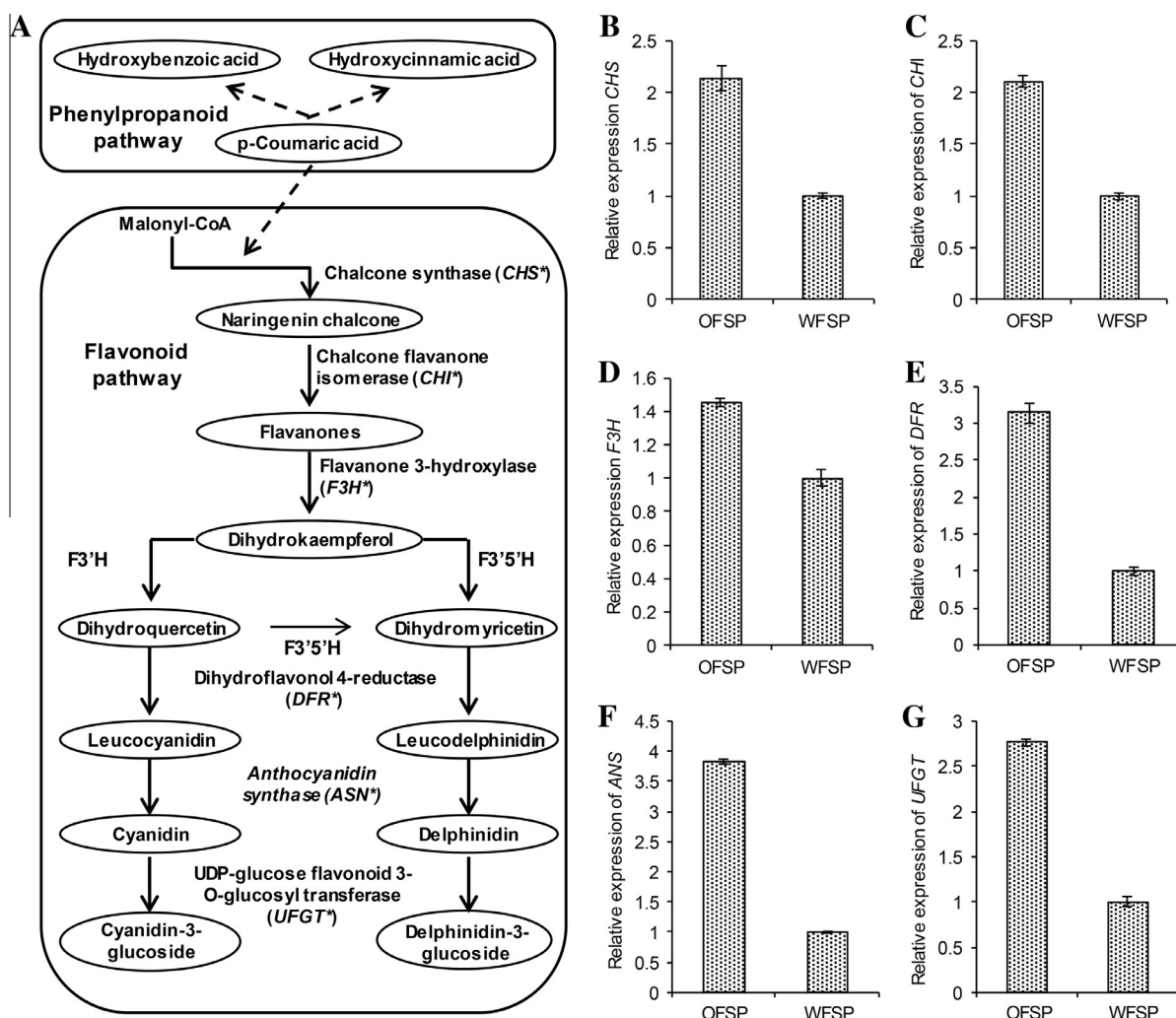


Fig. 5. Transcript abundance of flavonoid pathway genes. The candidate genes selected for expression analysis are shown with asterisk (*) (A). The relative abundance of chalcone synthase (*CHS*) (B), chalcone flavanone isomerase (*CHI*) (C), flavanone 3-hydroxylase (*F3H*) (D), dihydroflavonol 4-reductase (*DFR*) (E), anthocyanidin synthase (*ANS*) (F), and UDP-glucose flavanoid 3-O-glucosyl transferase (*UGFT*) (G) was determined by qRT-PCR. The mean values of three replicates were normalised using actin as internal control. Values represent mean \pm SE.

contents, especially the starch (Harijono, Estiasih, Saputri, & Kusnadi, 2013). We established a direct correlation amongst the starch, protein and WHC as these were found to be higher in OFSP than that of WFSP (Fig. 3E), though the fold increase varied.

3.5. Evaluation of storage stress response

To elucidate the impact of storage stress on carbohydrates, especially starch and cellulose, the rate of degradation was investigated. The disparity in the status of total carbohydrate, reducing sugar and starch contents impelled us to examine the degradation pattern of starch in both the cultivars during post-harvest storage. The starch content showed progressive decline till 105-d, and about 41% loss was observed in WFSP. However, 29% decline in starch content was observed until 105-d in OFSP (Fig. 4A).

The degradation of cellulose content showed a similar trend. The cellulose content in both the cultivars decreased progressively until 105-d during the storage, but degradation was more prominent in WFSP. The degradation of cellulose was 44% in WFSP as against 35% in OFSP (Fig. 4B). Post-harvest decline in starch content amongst different cultivars might be due to differential metabolism, especially the respiratory intensity and the increase in amylase activity during the storage (Jiang et al., 2012). OFSP

maintained a constant high starch and cellulose content during the storage stress and even the extent of degradation was found to be lower than that of WFSP. The degradation of starch and cellulose altogether indicates the better adaptability of OFSP during storage stress.

3.6. Determination of phytochemicals and carotenoids

Sweet potatoes are rich in antioxidants, such as phytochemicals and carotenoids which also provide distinct flesh colours. To examine the disparity at the level of metabolites, a comprehensive biochemical analyses was carried out. Detailed analyses revealed contrasting trends at the level of phytochemicals and carotenoids (Supplementary Table S2). While, TPC in WFSP was 1.58 mg GAE/g, it was 1.49 mg GAE/g in OFSP. Intriguingly, TFC was 0.705 mg QE/g in OFSP as against 0.617 mg QE/g in WFSP, which was significantly higher ($p < 0.05$).

Anthocyanin is the important group of phenolic compounds that contribute to the characteristic colour. As these constitute a subcategory of flavonoids, the TAC of both the cultivars was expected to follow the same trend of TFC. The TAC of OFSP was significantly higher (0.016 mg/g) than that of WFSP (0.012 mg/g). A higher TCC was expected in OFSP as the intensity of orange colour

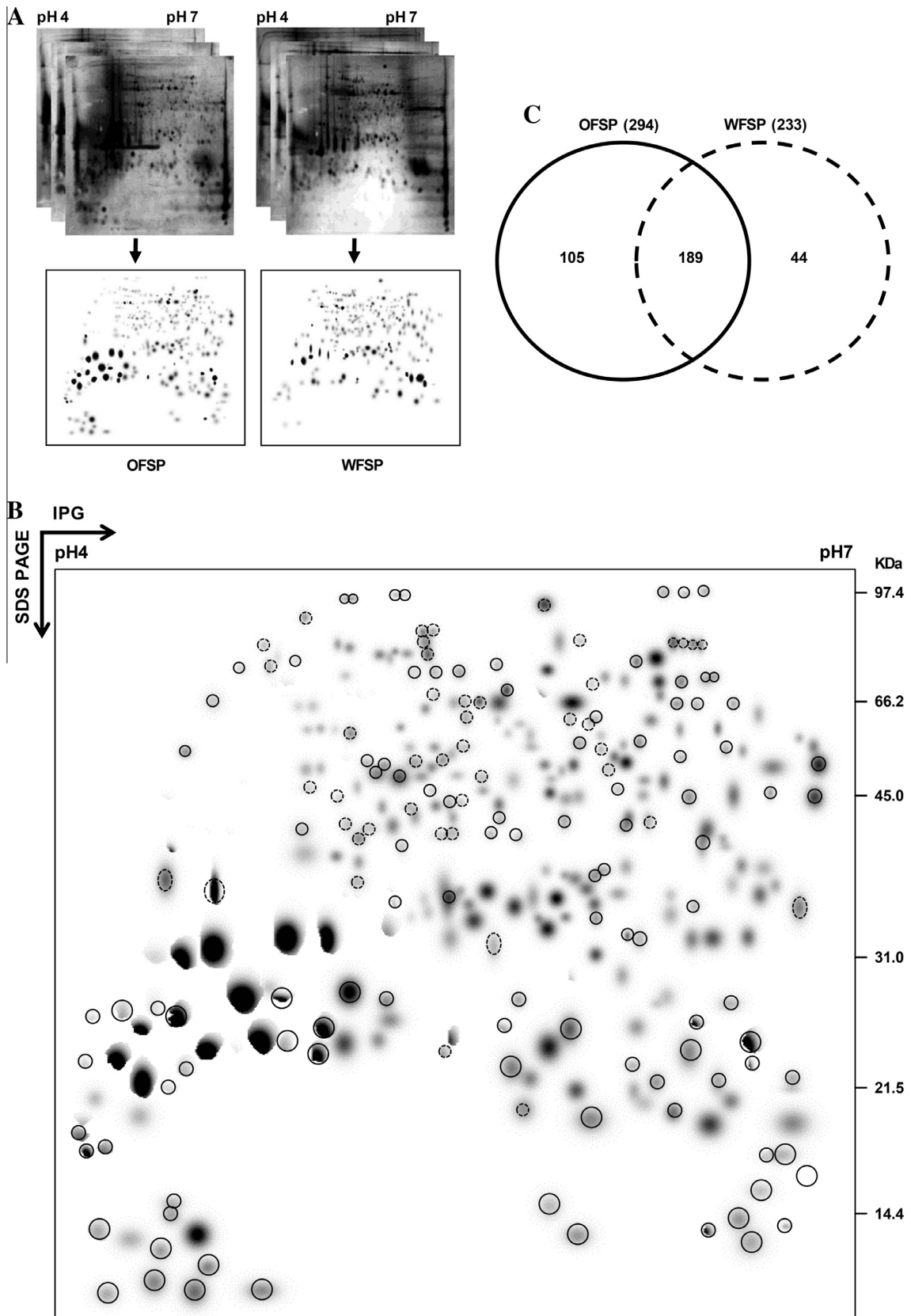


Fig. 6. Comparative proteome profile of OFSP and WFSP tubers. The proteomes displaying three replicate gels for each cultivar were computationally integrated into the “standard gel” (A). The differential proteome was developed from the “standard gels”. The exclusive protein spots displayed by cultivars are shown with solid and dotted circles for OFSP and WFSP, respectively (B). The Venn diagram shows the common and exclusive proteins in OFSP and WFSP tubers (C).

is attributed to the carotenoid content of an individual cultivar (Ameny & Wilson, 1997; Hagenimana, Carey, Gichuki, Oyunga, & Inumgi, 1998). TCC of OFSP was found to be 0.047 mg BCE/g, which was eightfold higher than that of WFSP (0.006 mg BCE/g).

The phytophenols and carotenoids often vary across the genotypes and are thought to be associated with genetic factors, which play a vital role in the formation of secondary metabolites (Teow et al., 2007). Padda and Picha (2007) have earlier reported a higher phenolic content in small-sized roots of potato than that of larger-sized roots. The decrease in phenolic content with the development of tubers in root crops has been attributed to a dilution effect resulting from tuber bulking. Higher contents of anthocyanin and carotenoids in OFSP substantiated its higher 'a' and 'b' colour space values. Orange-fleshed cultivars are known for their high β -carotene content, yet high genetic variation has been observed (Ameny & Wilson, 1997; Koala et al., 2013; Manrique & Hermann, 2001; Woolfe, 1992) in different cultivars. The higher acquisition of carotenoids in OFSP is likely to be a value added signature of this cultivar.

3.7. Differential expression of flavonoid pathway genes

As flavonoids are ubiquitous in nature and major contributors of colour and flavour, we investigated the relative transcript accumulation of the key flavonoid pathway genes viz., *CHS*, *CHI*, *F3H*, *DFR*, *ANS* and *UFGT* to elucidate the underlying molecular mechanism (Fig. 5A). Transcript analyses revealed a low transcript abundance of the early pathway genes viz., *CHS* and *CHI* (Fig. 5B and C) and the late pathway genes viz., *DFR*, *ANS* and *UFGT* in WFSP (Fig. 5E–G). Interestingly, though the *F3H* transcript was slightly induced in WFSP, the TFC was found to be low (Fig. 5D and Supplementary Table S2). This discrepancy might be due to the post-transcriptional and/or post-translational modification(s) of the candidate genes. In recent years, a mounting interest in producing food crops with increased levels of flavonoids necessitated the need to combine targeted metabolomic and transcriptomic approaches that might possibly help in determining gene-metabolite relationships. Carvalho, Cavaco, Carvalho, and Duque (2010) showed a strong correlation between the expressions of genes in flavonoid pathway. Furthermore, anthocyanin was found to be strongly associated with *DFR* gene. The downregulation of *DFR* gene has been shown to inhibit the accumulation of anthocyanin and proanthocyanidin, but increase flavonol influx (Wang et al., 2013).

Nutritional diversity amongst the cultivars is a distinguished and established fact, and are often observed in different transgenic lines, when compared with their wild-type counterparts, has been reported earlier. Several studies on genetic diversity between closely related genotypes had focused on the similarities as well as the difference in the metabolites owing to genetic factors and/or the expression of transgene (Chakraborty et al., 2013; Gayen, Sarkar, Datta, & Datta, 2013; Harrigan et al., 2007). We also found a significant difference in metabolites between the two closely related but contrasting cultivars grown in identical conditions. Changes in the accumulation of different metabolites may be attributed to pleiotropic and/or multigenic effect(s).

3.8. Comparative tuber proteome profiling

Comparative proteome profiling has remained an attractive tool for unambiguous comparison between cultivars of a particular crop species, even to categorise single mutations with multiple effects. However the outcome is affected by genotypic divergence, organ and tissues, developmental stages besides different environmental cues. The proteomics studies revealed the influence and efficacy of such approach to distinguish the landraces, populations, varieties and even species (Jaiswal et al., 2013; Subba et al., 2013).

To understand the diversity between OFSP and WFSP at protein level, comparative 2-DE analysis was carried out. Differential proteomic analysis of OFSP and WFSP revealed several contrasting trends as well as the commonalities in the proteomes (Fig. 6A and B). While 189 protein spots were found to be common to both the cultivars, 105 spots were exclusive to OFSP and 44 to WFSP (Fig. 6C). The proteomic analysis of OFSP and WFSP unraveled several common as well as cultivars-specific expressions of proteins, suggesting that subtle changes in the genome might lead to distinct proteome thereby influencing the acquisition of nutrients.

4. Conclusions

This study explores the nutritional diversity of two closely related but contrasting cultivars of sweet potato, so as to elucidate the factors pertaining to the varietal differences. This is a comprehensive report reflecting the contribution of metabolites and differential regulation of genes that might influence the nutrient acquisition. The intimate networking may be a major factor for genotype-specific disparity. The phytochemical analyses indicate that the relative performance and nutritional adequacy seems to be better in OFSP as compared to WFSP. The metabolite, transcript as well as proteomic analyses revealed commonality and diversity that might be assigned as distinct nutritional signature of the two cultivars. It is expected that nutrient rich crop would be beneficial to the targeted population, i.e. those dependent on plant based food, especially in the developing world. Furthermore, the introduction of new varieties would facilitate the valuable increase in the dietary intake of key nutrients such as carotenoids and protein.

Author's contributions

N.C., S.S., D.M., and S.C. designed research; S.S. and D.M., performed the experiments; S.S., D.M., N.C., S.C. and A.K.B. analysed data; and S.S. and N.C. wrote the article.

Acknowledgements

This work was supported by grants (BT/PR3123/16//250) from the Department of Biotechnology (DBT), Govt. of India. We kindly acknowledge the University Grant Commission (UGC), Govt. of India for providing predoctoral fellowship to D.M. We also thank CTCRI, India for providing planting materials of sweet potato. We thank Department of Food Engineering and Technology, Tezpur University, for extending their facilities. Assistance of Mr. Jasbeer Singh for illustrations and graphical representations, and Mr. Shankar Acharya and Mr. C. Ravishankar during field management is highly acknowledged.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at <http://dx.doi.org/10.1016/j.foodchem.2014.09.172>.

References

- Agrawal, L., Narula, K., Basu, S., Shekhar, S., Ghosh, S., Datta, A., et al. (2013). Comparative proteomics reveals a role for seed storage protein Am A1 in cellular growth, development, and nutrient accumulation. *Journal of Proteome Research*, 12, 4904–4930.
- Ahn, Y. O., Kim, S. H., Kim, C. Y., Lee, J. S., Kwak, S. S., & Lee, H. S. (2010). Exogenous sucrose utilization and starch biosynthesis among sweet potato cultivars. *Carbohydrate Research*, 345, 55–60.
- Ameny, M. A., & Wilson, P. W. (1997). Relationship between hunter color values and β -carotene contents in white-fleshed African sweet potatoes (*Ipomoea batatas* Lam.). *Journal of the Science of Food and Agriculture*, 73, 301–306.

- AOAC (2000). *Official methods of analysis* (17th ed.). Washington, DC, USA: Association of Official Analytical Chemists.
- Bovell-Benjamin, A. C. (2007). Sweet potato: A review of its past, present, and future role in human nutrition. *Advances in Food and Nutrition Research*, 52, 1–59.
- Carvalho, I. S., Cavaco, T., Carvalho, L. M., & Duque, P. (2010). Effect of photoperiod on flavonoid pathway activity in sweet potato (*Ipomoea batatas* (L.) Lam.) leaves. *Food Chemistry*, 118, 384–390.
- Chakraborty, N., Ghosh, R., Ghosh, S., Narula, K., Tayal, R., Datta, A., et al. (2013). Reduction of oxalate levels in tomato fruit and consequent metabolic remodeling following overexpression of a fungal oxalate decarboxylase. *Plant Physiology*, 162, 364–378.
- Chakraborty, S., Chakraborty, N., Agrawal, L., Ghosh, S., Narula, K., Shekhar, S., et al. (2010). Next-generation protein-rich potato expressing the seed protein gene Am A1 is a result of proteome rebalancing in transgenic tuber. *Proceedings of the National Academy of Sciences of the United States of America*, 107, 17533–17538.
- Collins, W. W., & Walter, W. M. Jr., (1985). Fresh roots for human consumption. In J. C. Bouwkamp (Ed.), *Sweet potato products: A natural resource for the tropics* (pp. 176–200). Boca Raton, Florida: CRC Press Inc.
- Emmanuel, H., Vasanthakalam, H., Ndirigwe, J., & Mukwantali, C. (2012). A comparative study on the β -carotene content its retention in yellow and orange fleshed sweet potato flours. *Association for strengthening Agricultural Research in Eastern and Central Africa*, 17, 41.
- Gayen, D., Sarkar, S. N., Datta, S. K., & Datta, K. (2013). Comparative analysis of nutritional compositions of transgenic high iron rice with its non-transgenic counterpart. *Food Chemistry*, 138, 835–840.
- Hagenimana, V., Carey, E. E., Gichuki, S. T., Oyunga, M. A., & Imungi, J. K. (1998). Carotenoid contents in fresh, dried and processed sweet potato products. *Ecology of Food and Nutrition*, 37, 455–473.
- Harijono, T. E., Saputri, D. S., & Kusnadi, J. (2013). Effect of blanching on properties of water yam (*Dioscorea alata*) flour. *Advance Journal of Food Science and Technology*, 5, 1342–1350.
- Harrigan, G. G., Stork, L. G., Riordan, S. G., Reynolds, T. L., Ridley, W. P., Masucci, J. D., et al. (2007). Impact of genetics and environment on nutritional and metabolite components of maize grain. *Journal of Agricultural and Food Chemistry*, 55, 6177–6185.
- Horsley, S. B., & Gottschalk, K. W. (1993). Leaf area and net photosynthesis during development of *Prunus serotina* seedlings. *Tree Physiology*, 12, 55–69.
- Huang, A. S., Tanudjaja, L., & Lum, D. (1999). Content of alpha-, beta-carotene, and dietary fiber in 18 sweet potato varieties grown in Hawaii. *Journal of Food Composition and Analysis*, 12, 147–151.
- Jaiswal, D. K., Ray, D., Choudhary, M. K., Subba, P., Kumar, A., Verma, J., et al. (2013). Comparative proteomics of dehydration response in the rice nucleus: New insights into the molecular basis of genotype-specific adaptation. *Proteomics*, 13, 3478–3497.
- Jiang, Y., Chen, C., Tao, X., Wang, J., & Zhang, Y. (2012). A proteomic analysis of storage stress responses in *Ipomoea batatas* (L.) Lam. tuberous root. *Molecular Biology Reports*, 39, 8015–8025.
- Kennedy, G., & Burlingame, B. (2003). Analysis of food composition data on rice from a plant genetic resources perspective. *Food Chemistry*, 80, 589–596.
- Kim, D. O., Jeong, S. W., & Lee, C. Y. (2003). Antioxidant capacity of phenolic phytochemicals from various cultivars of plums. *Food Chemistry*, 81, 321–326.
- Koala, M., Hema, A., Some, K., Pale, E., Sereme, A., Belem, J., et al. (2013). Evaluation of eight orange fleshed sweet potato (OFSP) varieties for their total antioxidant, total carotenoid and polyphenolic contents. *Journal of Natural Sciences Research*, 3, 67–72.
- Lemoine, R., La Camera, S., Atanassova, R., Dedaldechamp, F., Allario, T., Pourtau, N., et al. (2013). Source-to-sink transport of sugar and regulation by environmental factors. *Frontiers in Plant Science*, 4, 272.
- Maloof, J. N., Nozue, K., Mumbach, M. R., & Palmer, C. M. (2013). LeafJ: An ImageJ plugin for semi-automated leaf shape measurement. *Journal of Visualized Experiments*, 71, e50028.
- Mancinelli, A. L., Hoff, A. M., & Cottrell, M. (1988). Anthocyanin production in chl-rich and chl-poor seedlings. *Plant Physiology*, 86, 652–654.
- Manrique, K., & Hermann, M. (2001). Effects of GxE interaction on root yield and β -carotene content of selected sweet potato (*Ipomoea batatas* (L.) Lam) varieties and breeding clones. *CIP Program Report, 1999–2000*, 281–287.
- Mohan, C. (2011). Tropical tuber crops. In H. P. Singh & V. A. Parthasarathy (Eds.), *Advances in Horticultural biotechnology: Molecular markers and marker assisted selection-vegetables, ornamentals and tuber crops* (pp. 187–230). India: Westville Publishing House, New Delhi.
- Padda, M. S., & Picha, D. H. (2007). Antioxidant activity and phenolic composition in 'Beauregard' sweet potato are affected by root size and leaf age. *Journal of the American Society for Horticultural Science*, 132, 447–451.
- Ravindran, V., Ravindran, G., Sivakanesan, R., & Rajaguru, S. B. (1995). Biochemical and nutritional assessment of tubers from 16 cultivars of sweet potato (*Ipomoea batatas* (L.) Lam). *Journal of Agricultural and Food Chemistry*, 43, 2646–2651.
- Robertson, J. A., de Monredon, F. D., Dyssele, P., Guillon, F., Amado, R., & Thibault, J.-F. (2000). Hydration properties of dietary fibre and resistant starch: A European collaborative study. *LWT – Food Science and Technology*, 33, 72–79.
- Rose, I. M., & Vasanthakalam, H. (2011). Comparison of the nutrient composition of four sweet potato varieties cultivated in Rwanda. *American Journal of Food and Nutrition*, 1, 34–38.
- Rukundo, P., Shimelis, H., Laing, M., & Gahakwa, D. (2013). Storage root formation, dry matter synthesis, accumulation and genetics in sweetpotato. *Australian Journal of Crop Science*, 7, 2054–2061.
- Sadasivam, S., & Manickam, A. (1992). *Biochemical methods for agricultural sciences*. India: Wiley Estern Limited.
- Senanayake, S. A., Ranaweera, K. K., Gunaratne, A., & Bamunuarachchi, A. (2013). Comparative analysis of nutritional quality of five different cultivars of sweet potatoes (*Ipomea batatas* (L.) Lam) in Sri Lanka. *Food Science & Nutrition*, 1, 284–291.
- Subba, P., Kumar, R., Gayali, S., Shekhar, S., Parveen, S., Pandey, A., et al. (2013). Characterisation of the nuclear proteome of a dehydration-sensitive cultivar of chickpea and comparative proteomic analysis with a tolerant cultivar. *Proteomics*, 13, 1973–1992.
- Sweetlove, L. J., Kossmann, J., Riesmeier, J. W., Trethewey, R. N., & Hill, S. A. (1998). The control of source to sink carbon flux during tuber development in potato. *The Plant Journal*, 15, 697–706.
- Teow, C. C., Truong, V. D., McFeeters, R. F., Thompson, R. L., Pecota, K. V., & Yencho, G. C. (2007). Antioxidant activities, phenolic and β -carotene contents of sweet potato genotypes with varying flesh colours. *Food Chemistry*, 103, 829–838.
- Walter, W. M., Collins, W. W., & Purcell, A. E. (1984). Sweet potato protein: A review. *Journal of Agricultural and Food Chemistry*, 32, 695–699.
- Woolfe, J. A. (1992). *Sweet potato: An untapped Food Resource*. United Kingdom: Cambridge University Press.
- Wang, H., Fan, W., Li, H., Yang, J., Huang, J., & Zhang, P. (2013). Functional characterization of dihydroflavonol-4-reductase in anthocyanin biosynthesis of purple sweet potato underlies the direct evidence of anthocyanins function against abiotic stresses. *PLoS One*, 8, e78484.

Genotype Independent Regeneration and *Agrobacterium*-mediated Genetic Transformation of Sweet Potato (*Ipomoea batatas* L.)

Shubhendu Shekhar¹, Lalit Agrawal, Alak Kumar Buragohain¹, Asis Datta, Subhra Chakraborty and Niranjana Chakraborty^{1*}

National Institute of Plant Genome Research, Aruna Asaf Ali Marg, New Delhi-110067, India

Key words: Sweet potato, Somatic embryogenesis, Selectable marker, GUS expression

Abstract

Development of an efficient genotype independent regeneration and genetic transformation system in sweet potato continues to be of great interest. *Agrobacterium*-mediated genetic transformation protocol was established in two different cultivars of sweet potato using *Agrobacterium* strain EHA105 harbouring binary plasmid pBI121 containing *GUS* and *nptII* genes. The internodal stem segments from 30-day-old micropropagated plants were used as explant with different combinations of media and hormones. MS and LS media with various concentrations of growth regulators proved to be non-responsive and the infecundity was severe with the addition of cytokinins. Nonetheless, MS with 2,4-D and TDZ gave a good percentage of callusing but with low differentiation. In different concentrations of NAA, significant amount of callusing was observed but percentage of rooting remained low in both the genotypes. Gamborg's B5 supplemented with NAA proved to be the most suitable media and hormone combination, which yielded shoot formation after 8 - 10 weeks with a regeneration efficiency of 40 - 70%. Stable integration of transgene was confirmed by PCR analysis. Furthermore, qRT-PCR analysis was performed to assess the transcript accumulation in addition to the *GUS* enzymatic assay in the transgenic lines.

Introduction

Sweet potato is the world's seventh most important food crop and is grown in more than hundred countries world-wide covering tropical, subtropical and temperate zones. As an efficient biomass-producing plant for starch, especially

*Author for correspondence: <nchakraborty@nipgr.res.in>. ¹Department of Molecular Biology and Biotechnology, Tezpur University, Assam, India.

for bio-alcohol, sweet potato plays a crucial role and has become an attracted target for genetic improvements (Newell et al. 1995, Noh et al. 2010, Chen et al. 2010). Sweet potato ranks fifth after rice, wheat, maize and white potato, in terms of total global food production (Janson and Raman 1991). Asia ranks first in sweet potato production followed by Africa and South America, and it is one of the main food sources for the poorest class of population in underdeveloped and developing countries (Peirce 1987).

Highly efficient and reproducible transformation technologies are essential for genetic improvement program(s) for major food and fodder crops including the tuber species. An efficient plant regeneration and transformation system is highly desirable for the successful application of genetic engineering in sweet potato as it is a hexaploid, highly male sterile and self incompatible crop (Dhir et al. 1998). The demand for efficient transformation systems to facilitate transgene expression or RNAi-mediated silencing of target genes has been greatly increasing in sweet potato. The use of such technology in sweet potato may have future applications in improving the ability of this tuberous crop to withstand different adverse environmental factors, in increasing yield, and enhancing nutritional quality. Sweet potato is found to be relatively easy to micropropagate; however, it is quite recalcitrant to genetic transformation (Sihachakr and Ducreux 1993). In addition, the regeneration frequency in sweet potato seems to be genotype-dependent (Jarret et al. 1984, Wang et al. 1998, Aloufa 2002, Santa-Maria et al. 2009). In most of the cases, regeneration at a high frequency has been restricted to one or a few genotypes. The critical hurdle for genetic transformation systems in sweet potato is to overcome genotype dependence. Therefore, an ideal transformation system would be genotype independent but this remains one of the key challenges for sweet potato. In this study, we made an attempt to establish a simple, robust and efficient regeneration and genetic transformation system in sweet potato via somatic embryogenesis from culture of internodal explants.

Materials and Methods

Two cultivars of sweet potato (*Ipomoea batatas* L., cv. SP-6 and SP-17) were obtained from the Central Tuber Crop Research Institute (CTCRI), India. Plants were micropropogated in MS supplemented with IAA and maintained at $22 \pm 2^\circ\text{C}$ with $270 \mu\text{mol/m}^2/\text{s}$ light intensity under 16 hrs photoperiod. The internodal stem segments (3 - 5 mm) from four-week-old micropropogated plants were used as explants.

Agrobacterium tumefaciens strain EHA105 harboring binary plasmid pBI121 (Fig. 1) containing the *uidA* and *nptII* genes was used for genetic transformation. *Agrobacterium* strain containing the gene of interest was first grown in YEP plates supplemented with kanamycin (50 mg/l) and rifampicin (50 mg/l). The *Agrobacterium* suspension culture was prepared by picking a single colony and inoculated in YEP liquid medium containing the same antibiotics and allowed to grow at 28°C for 16 - 18 hrs. One-fiftieth volume of primary culture was inoculated and allowed to grow at 28°C until the optical density (OD600) reached 0.6 - 0.8. The *Agrobacterium* culture was pelleted at 5000 g for 5 min at ambient temperature (~25°C) by centrifugation and resuspended in same volume of MS liquid medium. About 50 explants were incubated for 45 min in a saturated culture of *Agrobacterium* with occasional swirling. The explants were retrieved from the culture and blotted on sterile Whatman filter paper to drain away extra inoculum. The explants were then co-cultivated on Gamborg's B5 basal medium supplemented with NAA (0.4 mg/l) and 0.8% agar for 48 hrs. The media was supplemented with bacteriostatic agent cefotaxime (250 mg/l) and selectable antibiotic kanamycin (100 mg/l). Growth room conditions maintained throughout were 22 ± 2°C and 16 hrs photoperiod. All cultures were examined periodically and morphological changes were recorded. The regenerated putative transformants were transferred to rooting media i.e., MS containing IAA (0.1 mg/l) and requisite antibiotics.

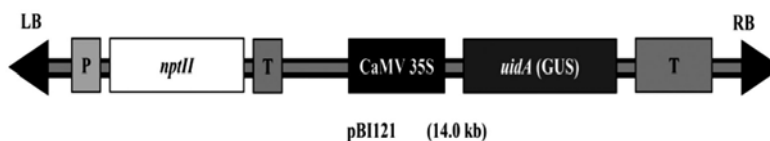


Fig. 1. Schematic representation of the plant expression vector pBI121 containing *nptII* and *uidA*, (encoding neomycin phosphotransferase II and *GUS*, respectively). LB, T-DNA left border; P, NOS promoter; T, NOS terminator; RB, T-DNA right border.

Integrity of the selectable marker gene *nptII* was assessed at the genome level by PCR analysis using genomic DNA isolated from the putative transformants and the wild-type plants as described earlier (Dellaporta et al. 1983). The presence of transgene was confirmed by PCR using gene specific primers (*nptIIF* 5' - ATGATTGAACAAGATGGATTGCACGCAGG -3' and *nptIIR* 5'- GAAGAA CTCGTCAAGAAGGCGATA -3'), which delimits 0.8 kb fragment from the *nptII* coding region. PCR analysis was performed in 20 µl reaction mix containing 1X PCR buffer, 2 mM MgCl₂, 200 µM dNTP mix, 0.5 µM each primer and 0.5 units of Taq polymerase. The cycling conditions employed were 3 min at 94°C denaturation and 30 amplification cycles using 94°C for 1 min, 58°C for 45 sec annealing and 72°C for 45 sec followed by 5 min extension at 72°C.

For qRT-PCR analysis of *GUS* expression, total RNA was isolated by phenol and guanidine thiocyanate method using the monophasic TriPure Isolation Reagent (Roche) as per the manufacturer's instructions. After determining RNA concentration by nanodrop, cDNA were prepared using 3'-RACE kit (Invitrogen). The *GUS*-specific primers (*GUSF* 5'-TGGTAATTACCGACGAAAACGGC-3' and *GUSR* 5'-ACGCGTGGTTACAGTCTTGCG-3') and the primers for housekeeping gene actin (*ActinF* 5'-CTCCCCTAATGAGTGTGATGTGAT-3' and *ActinR* 5'-GAGCCCCATGAGAACATTACCA-3') were designed using the primer express software. The quantitative real-time RT-PCR was performed with the ABI PRISM 7700 sequence detection system (Applied Biosystems) using SYBR green dye (Applied Biosystems). The analyses were done with two biological and three technical replicates. Mean of Ct values for target and endogenous control was considered for calculating the relative quantitation (RQ) value using comparative Ct ($2^{-\Delta\Delta Ct}$) method. Two putative transformants and the wild-type counterparts from both the genotypes were selected for quantitative enzyme assay. *GUS* (β -glucuronidase, EC 3.2.1.31) activity was determined using the fluorometric method described earlier (Jefferson 1987) with few modifications. The tissues were ground and homogenized with extraction buffer (50 mM sodium phosphate, pH 7.0, 10 mM beta-mercaptoethanol, 10 mM EDTA, 0.5 mM PMSF, 0.1% sodium lauryl sarcosine and 0.1% Triton X-100). The homogenates were then centrifuged (10000 g for 5 min at 4°C) and the supernatants were further used for assay. The reaction mixture in triplicate, consisted of 50 mM sodium phosphate, pH 7.0, 1 mM 4-methylumbelliferyl- β -D-glucuronide (MUG) and the tissue extract was then incubated at 37°C for appropriate time intervals (5 - 60 min). The reaction was terminated by the addition of 0.2 M Na₂CO₃. Fluorescence was measured at 455 nm using a Varian Cary Eclipse Fluorescence Spectrophotometer (Agilent) set at an excitation wavelength of 365 nm.

Results and Discussion

In the present investigation *in vitro* regeneration experiments were carried out using internodal explants from two different genotypes of sweet potato. In the past, there have been many attempts to produce transgenic sweet potato especially using the electroporation of protoplasts (Nishiguchi et al. 1992, Garcia et al. 2000) and particle bombardment (Prakash and Varadarajan 1992, Yang et al. 2005). In particular, the *A. tumefaciens* mediated transformation system has been widely used because of its efficiency, simplicity and stability of the introduced gene. The first successful such transformation protocol for sweet potato was based on the formation of hairy roots using leaf-discs explants by *A. rhizogenes*

(Dodds et al. 1991, Otani et al. 1993). However, morphological abnormalities shown by regenerated transgenic plants were a big question mark for the method. *A. tumefaciens* mediated transformation in sweet potato was well established by several workers (Al-Juboory and Skirvin 1991, Carelli et al. 1991, Prakash and Varadarajan 1991, Lowe et al. 1994, Otani et al. 2001). In general, these procedures have been mostly genotype-dependent with lower transformation efficiency (Otani et al. 2003, Song et al. 2004, Shimada et al. 2006, Luo et al. 2006), and often difficult to reproduce (Lowe et al. 1994). *Agrobacterium*-mediated transformation in sweet potato has also been applied for regeneration via somatic embryogenesis using somatic embryos or organs as explants by several workers (Pido et al. 1995, Newell et al. 1995, Gama et al. 1996, Otani et al. 1998, Luo et al. 2006).

To develop a simple regeneration and genetic transformation protocol, two different genotypes of sweet potato were considered as explant source. Over the years, *E. coli uidA* gene encoding *GUS* has been the most chosen reporter gene and widely used to assess the transient and stable transformation in plants. The *GUS* gene fusion system has found extensive application in plant gene expression studies due to the enzyme stability and high sensitivity (Fior and Gerola 2009). One of the most critical factors to develop a high efficiency transformation protocol is the use of a hypervirulent strain of *A. tumefaciens* and an appropriate medium for optimum infection of explants. Therefore, to mobilize the binary expression plasmid pBI121 harbouring *uidA* and *nptIII* genes into *A. tumefaciens* strain EHA105, triparental mating strategy was used.

In plant cell culture, growth and morphogenesis are significantly affected by the type(s) of media and the concentrations of growth regulators. Plant growth regulators are conceivably the most important components affecting shoot regeneration capacity of explants. Thus, optimization of the correct combinations of auxins and cytokinins is indispensable for high frequency shoot regeneration. Also, the genotype of the transformed explants is a deciding factor for their growth and development. To introduce the expression plasmid into sweet potato, diverse media and hormone combinations were examined. Some hormones and media combinations showed only callus formation, while some showed both shooting and rooting. During the first week of inoculation, little morphological differentiation was observed. The cut-ends began to swell and slight increase in tissue volume was observed. Callus formation took place from both side of the explants and began to increase in size in the following weeks. Root formation was evident from the third week depending on the plant genotypes and the concentrations of the hormone used. Use of MS with 2, 4-D and TDZ yielded significantly good percentage of calli, but showed no shoot or

root development. Similar results were observed when BAP and/or Zeatin were used alone or with combination of NAA and GA₃ (Table 1). The inhibitory action of cytokinin on regeneration of sweet potato has been well documented and contradicts the common practice of using cytokinin in combination with an auxin in regeneration trials. Several co-workers had previously reported inhibitory effect of cytokinin in shoot regeneration (Carswell and Locy 1984, Ozias-Akins and Perera 1992, Otani and Shimada 1996, Pido et al. 1995). Addition of either of these cytokinins inhibited root and shoot development in such a way that higher the cytokinin concentration, greater was the inhibitory effect; however, a corresponding increase in callus growth was also observed. Roots are the primary organ for the synthesis of cytokinins and thus rooted explants should conceivably be able to synthesize a little amount of cytokinins (Pido et al 1995). In addition, several tissues are considered to be cytokinin independent, and the findings of the present study also suggest that it is not necessary to supplement the regeneration medium with cytokinin for the regeneration of sweet potato. A similar response was observed in LS media in both the genotypes. However, overall performance of cv. SP-6 was found to be better when compared with SP-17, suggesting its better adaptability and efficacy. In different concentration of NAA, fair amount of callusing was observed in the explants in four weeks duration, but percentage of root development remained low in both the genotypes suggesting that media containing NAA alone might induce regeneration (Table 1). An alternate regeneration procedure with a single hormone was then employed wherein explants were first kept on callus inducing media (CIM) [3.2 g/l Gamborg's B5 with minimal organics, 30 g/l sucrose, 0.8% agar type A and 0.4 mg/l NAA, pH 5.6-5.8]. Within 3 weeks, callusing was observed from both the corners of the explants followed by initiation of roots. Once the complete rooting took place, the transformed explants were transferred to low auxin shoot induction media (SIM) [3.2 g/l Gamborg's B5 with minimal organics, 30 g/l sucrose, 0.8% agar type A and 0.2 mg/l NAA, pH 5.6-5.8]. Microshoots were produced within 8 - 10 weeks of subculturing in SIM (Fig. 2A-D). Regenerating shoots were harvested and transferred to root induction media (RIM) [4.4 g/l MS with vitamins (M5519, Sigma), 30 g/l sucrose, 0.8% agar type A and 0.1 mg/l IAA, pH 5.6-5.8) with cefotaxime (250 mg/l) and kanamycin (100 mg/l) (Table 2).

In *Agrobacterium*-mediated transformation system, activation of the virulence genes in the Ti plasmid is modulated by molecular signals provided by wounded tissue that is mimicked by phenolic compounds such as acetosyringone or hydroxyacetosyringone. The use of acetosyringone has been reported earlier in various sweet potato transformation systems (Luo et al. 2006, Yu et al. 2007).

Table 1. Differential response of intermodal explants in various culture conditions.

Genotype	Callusing media	No. of explants	Callus formation	% callus formation	Subculturing media	Root formation	% root formation
SP-6	BNG, MS	170	119	70	BNG (combination of different conc. of BAP, NAA, and GA ₃), MS	-	-
	2,4 D-BAP, MS	42	35	83.3	2,4 D-BAP, MS	-	-
	2,4 D-TDZ, MS	80	69	86.25	2,4 D-TDZ, MS	-	-
	Different conc. of NAA (0.1-0.8 mg/l), MS	70	65	92.8	MS, NAA (in different conc.)	15	21.42
SP-17	BNG, LS	65	11	16.9	BNG, LS	-	-
	Different conc. of NAA (0.1-0.8 mg/l), LS	60	27	45	LS, NAA (in different conc.)	12	20
	BNG, MS	42	5	11.9	BNG, MS, BAP (in different conc.)	-	-
	2,4 D-BAP, MS	40	24	60	BNG, MS	-	-
	2,4 D-TDZ, MS	60	30	50	BNG, BAP (in different conc.)	-	-
	Different conc. of NAA(0.1-0.8 mg/l), MS	45	36	80.0	NAA (in different conc.), MS	8	17.7
	BNG, LS	50	9	18	BNG, LS	-	-
	Different conc. of NAA(0.1-0.8 mg/l), LS	62	21	35	LS, NAA (in different conc.)	10	16.12

However, we developed the stable transformation system without the addition of exogenous signaling compound, considering the predominant presence of phenolic compounds in sweet potato. The regeneration efficiency was calculated as the number of stable transgenic plants obtained from the total number of explants that responded to rooting and callusing. The overall regeneration efficiency ranged between 40 - 70%, which appeared to be the best for sweet potato. The regeneration efficiency of cv. SP-6 was significantly better than that

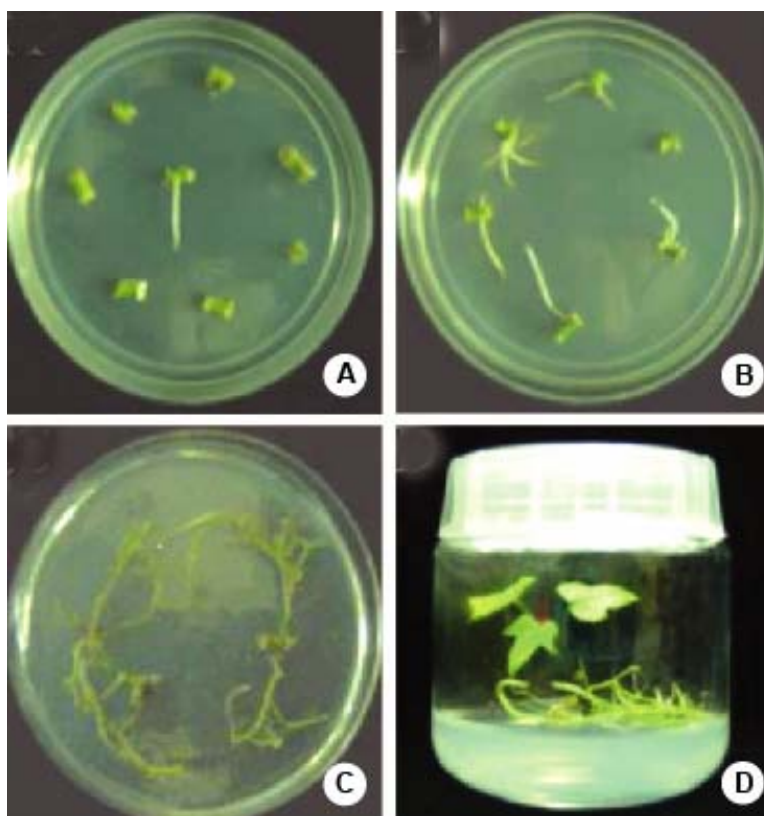


Fig. 2. Regeneration stages of putative transformants. The internodal explants infected by *A. tumefaciens* (strain EHA105) harbouring the expression plasmid pBI121 responded in a stage wise manner *viz.* callusing (A) followed by rooting (B and C) and shooting (D) in CIM and SIM, respectively.

of SP-17. Intriguingly, the induction of somatic embryogenic tissues at high frequencies has been found to be limited to a few genotypes, and when attempts were made to extend this into a wide range of genotypes, the majority was found to be recalcitrant or to respond at low frequencies. Al-Mazrooei et al. (1997) had optimized the culture condition for somatic embryogenesis in 14 genotypes and obtained the regeneration frequency up to 68%. It is evident that there has been a

gap between the regeneration frequency and the transformation efficiencies. However, transformation efficiencies of up to 20% reported by Luo et al. (2006) from leaf petioles were higher than those of several reports (Newell et al. 1995, Otani et al. 1998, Moran et al. 1998). A relatively rapid regeneration system (12 - 16 weeks) with a higher transformation frequency of 30.8% was reported by Song et al. (2004) from stem explants. In comparison with prolonged regeneration steps required in the somatic embryogenesis (Newell et al. 1995, Otani et al. 1998, Zhang et al. 2000), the rapid *de novo* organogenesis is likely to lower somaclonal variations (Luo et al. 2006). We observed the regeneration in time span of 8-12 weeks, starting from the day of infection with *Agrobacterium* which also strengthen the claim for the development of a relatively more rapid regeneration system for sweet potato.

Table 2. Combination of various media used for the induction of calli, roots and shoots from internodal explants.

No. of variants	Media code	Media composition
1	CIM	3.2 g/l Gamborg's B5, 30 g/l sucrose, 0.8% agar, and 0.4 mg/l NAA, pH 5.6 - 5.8.
2	SIM	3.2 g/l Gamborg's B5, 30 g/l sucrose, 0.8% agar, and 0.2 mg/l NAA, pH 5.6 - 5.8.
3	RIM	4.4 g/l MS, 30 g/l sucrose, 0.8% agar, and 0.1 mg/l IAA, pH 5.6 - 5.8 and supplemented with 250 mg/l cefotaxime and 100 mg/l kanamycin.

Selectable marker genes have a pivotal role in validating the plant transformation technologies as the marker genes allow to distinguish between the transformed and untransformed plant(s). While untransformed escapes are the common phenomena in transgenic technology of plants (Zang et al. 2000), a suitable concentration of test antibiotic can reduce such risk. To screen the true transformation events, high concentration of kanamycin at 100 mg/l was used and the kanamycin-resistant plants were considered for downstream analysis. While 18 out of 25 regenerated plants in cv. SP-6 were found to be positive, 9 out of 15 regenerated plants were positive in cv. SP-17. As stated above, morphological abnormality is one of the major constrains in the development of a successful regeneration and transformation system. We observed no detectable morphological difference between the wild-type and the transgenic plants.

Successful integration of the transgene at genome level was examined by PCR analysis in 12 putative transformants using *nptII* gene specific primers. The presence of 0.8 kb amplicons in the putative transformants confirmed the successful integration of the transgene (Fig. 3A, B).

Out of these 12 transgenic plants, 4 (2 from both the genotypes) were further selected to assess the transcript accumulation and relative *GUS* expression. The transcript accumulation of *GUS* showed 3 to 12-fold expression compared to the wild-type. Accumulation pattern also revealed more *GUS* transcript in the transformants of cv. SP-6 when compared with that of SP-17 transformants (Fig. 4A). The *GUS* activity of each transgenic plant was significantly higher than that of wild-type and untransformed plant. In addition, there was a positive correlation between the transcript accumulation and enzyme activity (Fig. 4B).

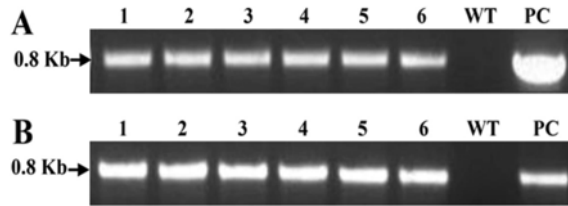


Fig. 3. Confirmation of transgene integration. Successful integration of the transgene was confirmed by PCR analysis in 6 putative transformants, each from cv. SP-6 (A) and SP-17 (B), respectively. An amplicon of 0.8 Kb confirmed the successful transgene integration in putative transformants which was absent in wild-type (WT) plants. Expression plasmid pBI121 was used as positive control (PC).

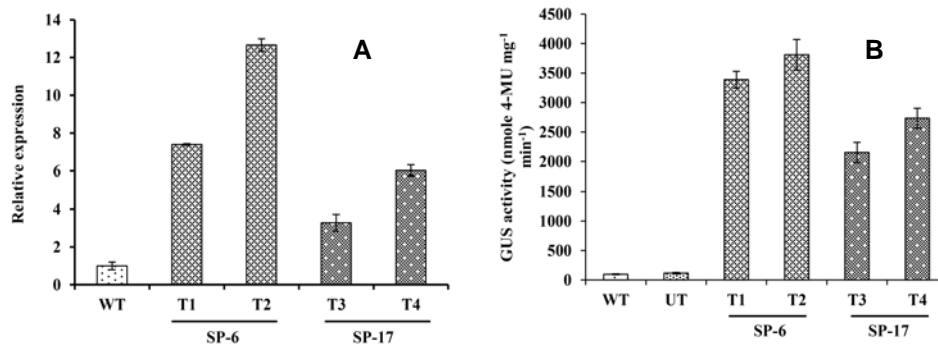


Fig. 4. Transcript accumulation and enzymatic activity of *GUS*. The transcript accumulation of *GUS* showed a 3 to 12-fold increase in expression than that of wild-type (WT) plant (A). Higher *GUS* activity was observed in the transgenic plants than WT and UT (untransformed plant) (B). Each analysis was performed in 4 transgenic plants (2 from each of the genotypes).

The genetic transformation protocol reported here may be used to incorporate agriculturally and/or industrially important candidate gene(s) in sweet potato. However, it is likely that, by manipulating the regeneration process or better susceptibility to *Agrobacterium*, it will be possible to address issues of

genotype-dependence and to improve transformation efficiencies in sweet potato further.

Acknowledgements

This work was supported by grants (BT/PR/3123/16/250/02) from the Department of Biotechnology (DBT), Govt. of India. The authors thank Divya Rathi for critical reading of the manuscript, and Jasbeer Singh for technical assistance.

References

- Aloufa MA** (2002) Some factors affecting the callus induction and shoot formation in two cultivars of sweet potato (*Ipomoea batatas* L. POIS). *Cienc. Agrotec.* **26**: 964-969.
- Al-Juboory KH** and **Skirvin RM** (1991) *In vitro* regeneration of *Agrobacterium*-transformed sweet potato (*Ipomoea batatas* L.). *PGRSA Quarterly* **19**: 82-89.
- Al-Mazrooei S**, **Bhatti MH** and **Henshaw GG** (1997). Optimisation of somatic embryogenesis in fourteen cultivars of sweet potato [*Ipomoea batatas* (L.) Lam.]. *Plant Cell. Rep.* **16**: 710-714.
- Carelli MLD**, **Skirvin RM** and **Harry DE** (1991) Transformation and regeneration studies of "Jewel" sweet potato. *In*: Hill WA, Bonsi CK and Loreatan PA (eds.) Sweet Potato Technology for the 21st Century. Tuskegee University, Tuskegee. pp. 52-60.
- Carswell GK** and **Locy RD** (1984) Root and shoot initiation by leaf, stem, and storage root explants of sweet potato. *Plant Cell Tiss. Org. Cult.* **3**: 229-236.
- Chen L**, **Du Z**, **Nishimura Y**, **Hamaguchi T**, **Sugita T**, **Nagata R**, **Terao H** and **Tsuzuki E** (2010) Approach to establishment of plant regeneration and transformation system in sweet potato (*Ipomoea batatas*) by culture of leaf segments. *Bull. Minamikyushu University* **40 A**: 59-63.
- Dhir SK**, **Oglesby J** and **Bhagsari AS** (1998) Plant regeneration via embryogenesis and transient gene expression in sweet potato protoplasts. *Plant Cell Rep.* **17**: 665-669.
- Dellaporta SL**, **Wood J** and **Hicks JB** (1983). A plant DNA miniprep: version II *Plant Mol. Biol. Rept.* **1**: 19-21.
- Dodds JH**, **Merzdorf C**, **Zambrano V**, **Sigüenñas C** and **Jaynes J** (1991) Potential use of *Agrobacterium*-mediated gene transfer to confer insect resistance in sweet potato *In*: Jansson RK, Raman KV (Eds.) Sweet potato pest management: a global perspective. West View Press, Oxford, U.K. pp. 203-219.
- Fior S** and **Gerola PD** (2009) Impact of ubiquitous inhibitors on the *GUS* gene reporter system: evidence from the model plants *Arabidopsis*, tobacco and rice and correction methods for quantitative assays of transgenic and endogenous *GUS*. *Plant Methods* **5**: 19.
- Gama MIC**, **Leite JRP**, **Cordeiro AR** and **Cantliffe DJ** (1996) Transgenic sweetpotato plants obtained by *Agrobacterium tumefaciens* mediated transformation. *Plant Cell Tiss. Org. Cult.* **46**: 237-244.

- García R, Morán R, Mena J, Somontes D, Zaldúa Z, López A and García M** (2000) Sweet potato (*Ipomoea batatas* L.) regeneration and transformation technology to provide weevil (*Cylas formicarius*) resistance. Field trial results. *In: Arencibia AD* (eds) Plant Genetic Engineering: Towards the Third Millennium. Elsevier Science B.V. pp. 112-117.
- Jansson RK and Raman KV** (1991) Sweet potato pest management: a global overview. *In: Jansson RK, Raman KV* (eds) Sweet potato pest management: a global perspective. Westview, Boulder. pp. 1-12.
- Jarret RL, Salazar S and Fernandez RZ** (1984) Somatic embryogenesis in sweet potato. *Hortsci.* **19**: 397-398.
- Jefferson RA** (1987). Assaying chimeric genes in plants: the GUS gene fusion system. *Plant Mol. Biol. Rep.* **5**: 387-405.
- Lowe JM, Hamilton WDO and Newell CA** (1994) Genetic transformation in *Ipomoea batatas* (L.) Lam. (Sweet potato). *In: Bajaj YPS* (ed.) Biotechnology in Agriculture and Forestry. Volume 29: Plant Protoplasts and Genetic Engineering V. Springer, Heidelberg, pp. 308-320.
- Luo HR, Santa-Maria M, Benavides J, Zhang DP, Zhang YZ and Ghislain M** (2006). Rapid genetic transformation of sweet potato (*Ipomoea batatas* (L.) Lam.) via organogenesis. *Afr. J. Biotechnol.* **5**: 1851-1857.
- Morán R, García R, López A, Zaldúa Z, Mena J, García M, Armas R, Somonte D, Rodríguez J, Gómez M and Pimentel E** (1998) Transgenic sweetpotato plants carrying the delta-endotoxin gene from *Bacillus thuringiensis* var. *tenebrionis*. *Plant Sci.* **139**: 175-184.
- Newell CA, Lowe JM, Merryweather A, Rooke LM and Hamilton WDO** (1995) Transformation of sweet potato (*Ipomoea batatas* (L.) Lam.) with *Agrobacterium tumefaciens* and regeneration of plants expressing cowpea trypsin inhibitor and snowdrop lectin. *Plant Sci.* **107**: 215-227.
- Nishiguchi M, Uehara Y and Komaki K** (1992) Stable transformation of sweet potato by electroporation. *In Vitro Cell Dev. Biol. Plant* **28**: 126.
- Noh SA, Lee HS, Huh EJ, Huh GH, Paek KH, Shin JS and Bae JM** (2010) *SRD1* is involved in the auxin-mediated initial thickening growth of storage root by enhancing proliferation of metaxylem and cambium cells in sweet potato (*Ipomoea batatas*). *J. Exp. Bot.* **61**: 1337-1349.
- Otani M, Mii M, Handa T, Kamada H and Shimada T** (1993) Transformation of sweet potato (*Ipomoea batatas* (L.) Lam.). *Plant Sci.* **94**: 151-159.
- Otani M and Shimada T** (1996). Efficient embryogenic callus formation in sweet potato (*Ipomoea batatas* (L.) Lam.). *Breed. Sci.* **46**: 257-260.
- Otani M, Shimada T, Kimura T and Saito A** (1998) Transgenic plant production from embryogenic callus of sweet potato (*Ipomoea batatas* (L.) Lam.) using *Agrobacterium tumefaciens*. *Plant Biotechnol.* **15**: 11-16.

- Otani M, Wakita Y and Shimada T** (2001) Genetic transformation of sweet potato (*Ipomoea batatas*(L.) Lam.) by *Agrobacterium tumefaciens*. *Acta Hort. (ISHS)* **560**: 193-196.
- Otani M, Wakita Y and Shimada T** (2003) Production of herbicide-resistant sweet potato (*Ipomoea batatas* (L.) Lam.) plants by *Agrobacterium tumefaciens*-mediated transformation. *Breed. Sci.* **53**: 145-148.
- Ozias-Akins P and Perera S** (1992) Regeneration of sweet potato plants from protoplast-derived tissues. *In*: Hill WA and Bonsick Loretan PA (Eds.) *Sweet Potato: Technology for the 21st Century*. Tuskegee University, Tuskegee. pp. 61-66.
- Pido N, Kowyama Y, Kakeda K and Shiotani I** (1995) Plant regeneration from leaf discs and stem segments of sweet potato using only NAA as supplementary regulator. *Plant Tiss. Cult. Lett.* **12**: 289-296.
- Pierce LC** (1987) Sweet potato. *In*: *Tuber and Tuberous Rooted Crop*, John Wiley, New York. pp. 300-308.
- Prakash CS and Varadarajan U** (1991) Expression of foreign genes in transgenic sweet potatoes. *Proceedings of the International Society and Plant Molecular Biology*. Tucson, AZ, USA, October 8-12.
- Prakash CS and Varadarajan U** (1992) Genetic transformation of sweet potato by particle bombardment. *Plant Cell Rep.* **11**: 53-57.
- Santa-Maria M, Pecota KV, Yencho CG, Allen G and Sosinski B** (2009) Rapid shoot regeneration in Industrial 'high starch' sweet potato (*Ipomoea batatas* L.) genotypes. *Plant Cell Tiss. Organ Cult.* **97**: 109-117.
- Shimada T, Otani M, Hamada T and Kim SH** (2006) Increase of amylose content of sweet potato starch by RNA interference of the starch branching enzyme II gene (*IbSBE II*). *Plant Biotechnol.* **23**: 85-90.
- Sihachakr D and Ducreux G** (1993) Regeneration of plants from protoplasts of sweet potato [*Ipomoea batatas* (L.) Lam.]. *In*: Bajaj, Y.P.S. (Eds.) *Biotechnology in Agriculture and Forestry. Volume 23: Plant Protoplast and Genetic Engineering IV*. Springer, New York. pp. 43-59.
- Song GQ, Honda H and Yamaguchi K** (2004) Efficient *Agrobacterium tumefaciens*-mediated transformation of sweet potato (*Ipomoea Batatas* (L.) Lam) from stem explants using a two-step kanamycin-hygromycin selection method. *In Vitro Cell. Dev. Biol. Plant* **40**: 359-365.
- Wang JS, Sato M and Kokubu T** (1998) Efficient embryogenic callus formation and plant regeneration in shoot tip cultures of sweet potato. *Mem. Fac. Agr. Kagoshima University.* **34**: 61-64.
- Yang KY, Yi GB, Kim YM, Choi G, Shin BC, Shin YM and Kim KM** (2005) Production of transgenic sweet potato (*Ipomoea batatas* (L.) Lam.) lines via microprojectile bombardment. *Korean J. Breed.* **37**: 236-240.
- Yu B, Zhai H, Wang Y, Zang N, He S and Liu Q** (2007) Efficient *Agrobacterium tumefaciens*-mediated transformation using embryogenic suspension cultures in sweet potato, *Ipomoea batatas* (L.) Lam. *Plant Cell Tiss. Organ Cult.* **90**: 265-273.

Zhang D, Cipriani G, Rety I, Golmirzaie A, Smit S and Michaud D (2000) Expression of protease inhibitors in sweet potato. *In: Recombinant Protease Inhibitors in Plants*. Michaud D (Eds.), Landes Bioscience, Georgetown, USA. pp. 167-178.

Zhang L, Rybczynski JJ, Langerberg W, Mitra A and French R (2000) An efficient wheat transformation procedure: transformed calli with long-term morphogenic potential for plant regeneration. *Plant Cell Rep.* **19**: 241-250.

Comparative Proteomics Reveals a Role for Seed Storage Protein AmA1 in Cellular Growth, Development, and Nutrient Accumulation

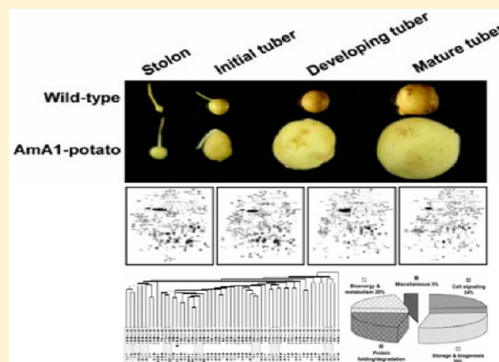
Lalit Agrawal,^{†,§} Kanika Narula,^{‡,§} Swaraj Basu,[‡] Shubhendu Shekhar,[†] Sudip Ghosh,[‡] Asis Datta,[‡] Niranjan Chakraborty,^{*,†} and Subhra Chakraborty^{*,‡}

[†]Laboratory 104 and [‡]Laboratory 105, National Institute of Plant Genome Research, Aruna Asaf Ali Marg, New Delhi 110067, India

Supporting Information

ABSTRACT: Seed storage proteins are known to be utilized as carbon and nitrogen source for growing seedlings and thus are considered as potential candidates for nutritional improvement. However, their precise function remains unknown. We have earlier shown that ectopic expression of a seed storage protein, AmA1, leads to increase in protein besides high tuber yield in potato. To elucidate the AmA1-regulated molecular mechanism affecting increased protein synthesis, reserve accumulation, and enhanced growth, a comparative proteomics approach has been applied to tuber life-cycle between wild-type and AmA1 potato. The differential display of proteomes revealed 150 AmA1-responsive protein spots (ARPs) that change their intensities more than 2.5-fold. The LC-ESI-MS/MS analyses led to the identification of 80 ARPs presumably associated with cell differentiation, regulating diverse functions, viz., protein biogenesis and storage, bioenergy and metabolism, and cell signaling. Metabolome study indicated up-regulation of amino acids paralleling the proteomics analysis. To validate this, we focused our attention on anatomical study that showed differences in cell size in the cortex, premedullary zone and pith of the tuber, coinciding with AmA1 expression and localization. Further, we interrogated the proteome data using one-way analysis of variance, cluster, and partial correlation analysis that identified two significant protein modules and six small correlation groups centered around isoforms of cysteine protease inhibitor, actin, heat shock cognate protein 83 and 14-3-3, pointing toward AmA1-regulated overlapping processes of protein enhancement and cell growth perhaps through a common mechanism of function. A model network was constructed using the protein data sets, which aim to show how target proteins might work in coordinated fashion and attribute to increased protein synthesis and storage reserve accumulation in AmA1 tubers on one hand and organ development on the other.

KEYWORDS: seed storage protein, AmA1, potato, comparative proteomics, metabolomics, nutrient accumulation, 2-DE, mass spectrometry, protein network



INTRODUCTION

Nutritional quality and agricultural productivity are the two key issues to sustainable food production worldwide. This is exemplified several times in the context of an ever-growing population where human health largely depends on plants. Storage organs display diverse nutritional quality and complex multistep development and act as sinks in plants. The composition of nutrients in the storage organs, carbon (C) and nitrogen (N) in particular, greatly influence the organ development and determine the nutritive quality. Both C and N metabolites are known to act as signals that influence many cellular processes, for example, development, metabolism related to N assimilation, and amino acid synthesis, in addition to their essential role as macronutrients in living organisms, including plants.^{1–11} Seed storage proteins in plants, rich in essential amino acids, are thought to serve as C and N source for the growing seedling¹² and meet the major dietary protein requirement of over half of the world population.¹³ Thus, they presumably play an essential role in productivity and protein quality and have received considerable attention due both to

their postulated dual role in growth and importance as a component of the human diet. Intensive research directed toward isolation and characterization of these seed storage proteins not only has allowed advances in our understanding of the synthesis, accumulation, processing, and transcriptional control of their genes^{14–17} but also has facilitated the targeted genetic engineering of crop plant's nutritional status.^{18–22} Despite the fact that there are many studies concerning the genetic control and spatial and temporal regulation of seed storage proteins, our current knowledge about their exact functional role and physiological relevance remains unknown. Only recently focus has been given to study crop plants overexpressing storage protein;^{23,24} however, the *modus operandi* of cellular network and physiological consequences toward sensing a storage protein has not been elucidated. Such

Special Issue: Agricultural and Environmental Proteomics

Received: August 2, 2013

finding will not only impact plant biology but in the near future would be useful for identifying biomarkers, prioritizing molecular targets, and pathway bioengineering for crop improvement.

Cellular physiology is based on the level of proteins, their activation or deactivation by post-translational modifications, and the metabolite pool. The identification of a protein and its site of modification can be obtained using MS/MS analyses, while the tandem development of metabolite profiling by gas chromatography–mass spectrometry (GC–MS) paves the way to a more substantial base of phenome study. Proteomics allows comprehensive characterization of proteomes, and despite the development of novel gel-free technologies, two-dimensional gel electrophoresis (2-DE) coupled with mass spectrometry allows evaluation of expression of hundreds of proteins and their isoforms or post-translational modifications and thus appears to be the technique of choice for protein expression studies.^{25–28} Isoforms are highly regulated gene products that can differ in their biological activity, regulatory properties, temporal and spatial expression, intracellular location, or any combination thereof. The significance of their existence can be attributed to their additive and/or differential function, which ultimately culminates in a cumulative and rapid response to physiological state. Therefore, isoforms are generally considered to diversify the function of a protein. The ability of this discovery approach to produce completely unpredictable and novel findings by application of a systematic process has been authenticated in many studies.^{29,30} Understanding the biological complexities in a cell upon genetic, epigenetic, and external perturbations can be improved and expanded by exploiting proteomic studies that provide comprehensive evaluation of molecular phenotypes and additional information about gene function and cellular pathways.³¹ Proteomic analyses of plant organs or tissues have been used to monitor developmental changes in seed^{32,33} and tuber,^{34,35} environmental stress responses,^{36–38} comparison of plant varieties, and mutant characterization³⁹ in addition to recent study of transgenic crops.^{31,40–43} Metabolomics is an imperative tool to recognize new protein function for functional proteomics. When combined with proteome studies, metabolite profiling reveals unanticipated insights into the diverse regulatory pathways to draw a comprehensive and integrated outlook of cellular metabolism.^{44,45} The large amount of data generated from such studies requires the use of appropriate statistical methods to extract the information of interest. Clustering algorithms and principle component analysis (PCA) identify patterns of expression that may suggest co-regulation, while Pearson partial correlation is one of the ways to verify the biological relevance of individual protein in the inferred correlation network.⁴⁶ In addition, the emergence of a postgenomic view expands the protein's role, regarding it as an element in a network of 'contextual' or 'cellular' functions within functional modules.⁴⁷

AmA1, a storage protein, was reported to be expressed in seed tissue of pseudocereal *Amaranthus hypochondriacus* and disappear during germination and seedling growth.¹⁵ The protein was synthesized during the grain development like other storage proteins.⁴⁸ It was one of the first characterized protein encoded by a nonallergenic seed storage albumin gene.^{15,20,49,50} In addition, we and others have shown that AmA1 regulates the biology of nutrition, and its ectopic expression in storage organs such as tuber and seed helps increase protein and amino acid accumulation.^{20,21} Further,

AmA1 plays an important physiological role during organ development and is crucial for the homoblastic growth of potato tuber, the gradual transition cascade from stolon to mature tuber.^{20,22} Thus we earlier hypothesized that induction of protein synthesis or mitogenic activity might be the result of overexpression of AmA1 or its turned-over products as signal molecule.²⁰ In a recent study, our preliminary data suggested that proteome rebalancing due to AmA1 expression might lead to nutrient enhancement and increased yield.²² It is thus conceivable that AmA1 might play a crucial role during seed germination and seedling growth as a nutrient source and growth-promoting substance. Further, seed storage proteins apart from their possible role in serving storage function have also been implicated in defense and environmental stress response.^{51–54} Likewise, we hypothesize that AmA1 might have similar functions in *Amaranthus hypochondriacus*. Despite the above research linking AmA1 with nutritional enhancement and increased growth, its biological function remains to be clarified. Following these observations, and given that we have a long-standing interest in understanding the role of the seed storage protein in plants, we decided to investigate the role of seed protein AmA1 on the observed homoblastic developmental changes and nutrient enhancement.

Here, we report the comparative proteome analysis of AmA1 potato and that of the wild-type at four different stages of tuberization. It was reasoned that a differential proteome analysis of AmA1 potato might broaden our understanding of functional protein and signaling networks involved in storage protein regulated pathways in nutrient signaling and growth in plants. In addition, we provide a global view of how protein networks are modulated in response to AmA1 sensing. We combined the quantitative models describing the protein expression changes and correlation network of the cell in response to AmA1 that allowed us to globally identify a set of protein subnetworks affected by the storage protein. These results are discussed in the context of current theories of the cellular and metabolic cues underlying nutrient accumulation and development. The differences in protein expression pattern and function appeared to encompass diverse metabolic and signaling pathways that provide new insights into the underlying mechanisms, which might contribute to altered nutritive values and increased tuber yield. Furthermore, the pathways identified by comparative proteomics were validated by analyzing the metabolome.

■ EXPERIMENTAL SECTION

Plant Material and Experimental Design

The transgenic potato genotypes 3SS-AmA1 (A16/1) and GBSS-AmA1 (A16/6), overexpressing a seed storage protein AmA1 constitutively and tuber-specifically, respectively, along with the wild-type genotype, A16 described earlier,²⁰ were used in the present study. A complete randomized design was used for growing size normalized seed tubers of wild-type and AmA1 potato in three replicate plots with 45 tubers per replicate for 100–110 days, until tuber maturation. The wild-type and AmA1 expressing lines were grown side-by-side in replicated plots at identical conditions to eliminate the environmental influence on the varieties, if any, and therefore should reflect the effects of the transgene only. Rows were spaced 60 cm apart, and plants were spaced 20 cm apart within each row in the replication block. Standard cultivation and management practices were followed throughout the growing season. Tubers

were planted in the month of October when the temperature is approximately 20–24 °C in the daytime and 15–18 °C at night. Tuberization occurred 45 days after planting at low temperature (15–18 °C in the daytime and 8–10 °C at night) and short-day photoperiod condition. The following developmental stages were selected and harvested: stolons at 6–8 weeks, stage 1 (S1); swollen stolons or initial tubers at 8–10 weeks, stage 2 (S2); developing tubers at 10–12 weeks, stage 3 (S3); and mature tubers at 14–15 weeks, stage 4 (S4). To maintain uniformity among the replicates, tubers were separated according to plantation time, morphology, tuber weight, and diameter of the wild-type and AmA1 potatoes for further analysis as described earlier.^{34,35} The average weights of each of the aforesaid developmental stages were 0.075, 0.510, 0.935, and 2.165 g with average diameter of 3.125, 5.75, 9.125, and 14.00 mm, respectively, in the wild-type tubers. However, the average weight of the AmA1 tubers at these four developmental stages in 35S-AmA1 and GBSS-AmA1 were 0.20 ± 0.025 , 1.40 ± 0.151 , 5.30 ± 0.512 , and 16.0 ± 1.425 g and the average diameters were 6.5 ± 0.635 , 12.5 ± 1.170 , 22.5 ± 2.121 , and 30.0 ± 3.55 mm, respectively (Figure 1A–C). Tubers were collected from three replicate plots and pooled to normalize the effect of variation in the biological replicates.

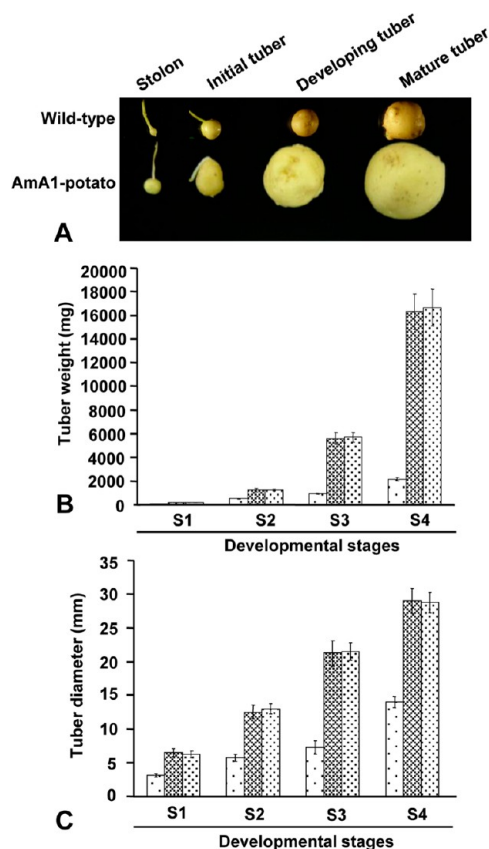


Figure 1. AmA1-induced growth response in potato tuber. (A) Photographs of various stages of tuber development: stolon, initial tuber, developing tuber, and mature tuber. The developmental stages were based on tuber weight and diameter as detailed in the Results section. The fresh weight (B) and the tuber diameter (C) of the harvested tubers were determined and plotted against each stage of tuberization process. Data represent means \pm SD of three measurements: (left bar) wild-type, (middle bar) 35S-AmA1, and (right bar) GBSS-AmA1.

Each biological replicate of a specific developmental stage for either wild-type or AmA1 potato lines consisted of 30–50 tubers for S1, 20–30 tubers for S2, 10–15 tubers for S3, and 10 tubers for S4 obtained from 10 different plants (Supporting Information Figure 1). Tubers were stored at -80 °C for further use after quick-freezing in liquid nitrogen.

Physiological and Morphological Characterization

We measured the photosynthetic activity in the transgenic and wild-type plants at 8 weeks of plantation under standard atmospheric (360 ppm CO₂) and light (750 μ mol/m²/s) conditions. Photosynthetic activity, leaf area, fresh weight and dry weight of wild-type and AmA1 plants were measured as described earlier.²²

Isolation of Tuber Proteins and 2-Dimensional Gel Electrophoresis

Soluble proteins were isolated from pooled potato tubers from each developmental stages of wild-type and AmA1 potato from three replicate plots to normalize the effect of variations in the biological replicates as described earlier.³⁵ Protein concentration was determined by Bradford assay (Bio-Rad). The tuber proteins were diluted in dilution buffer [100 mM, Tris-Cl (pH 8.5), 20% (v/v) glycerol, 8% (w/v) SDS, 20 mM DTT, 1 mM PMSF] and boiled for 5 min.⁵⁵ Protein samples were allowed to cool at room temperature (25 °C) and precipitated with 9 vol of 100% chilled acetone overnight at -20 °C. The precipitates were recovered by centrifugation at 10,000g at 4 °C, for 10 min. Protein pellets were washed twice with 80% acetone to remove excess SDS, air-dried, and resuspended in 2-D rehydration buffer [8 M urea, 2 M thiourea, 4% (w/v) CHAPS, 20 mM DTT, 0.5% (v/v) Pharmalyte (pH 4–7) and 0.05% (w/v) bromophenol blue]. Isoelectric focusing was carried out with 250 μ g of protein. Protein was loaded by in-gel rehydration method onto 13-cm IEF strips (pH 4–7), and electrofocusing was performed using an IPGphor system (Amersham Biosciences, Bucks, U.K.) at 20 °C for 25,000 Vh. The focused strips were subjected to reduction with 1% (w/v) DTT in 10 mL of equilibration buffer [6 M urea, 50 mM Tris-HCl (pH 8.8), 30% (v/v) glycerol and 2% (w/v) SDS], followed by alkylation with 2.5% (w/v) iodoacetamide in the same buffer. The strips were then loaded on top of 12.5% polyacrylamide gels for SDS-PAGE. The electrophoresed proteins were stained with a silver stain plus kit (Bio-Rad).

Image Acquisition and Data Analysis

After two-dimensional gel electrophoresis and gel staining, gel images were scanned using the Bio-Rad FluorS system equipped with a 12-bit camera. PDQuest version 7.2.0 (Bio-Rad) was used to assemble the first level match set (master image) from three replicate 2-DE gels. For each developmental stage, at least three 2-DE gels, representing three biological replicates, were used for the data analysis. The detailed data analyses were carried out as described previously.^{35,36} Assessment of protein spot quality, molecular mass, and pI of individual protein was determined as described previously.³⁵ Each spot included on the standard gel met several criteria: it was present in at least two of the three gels and was qualitatively consistent in size and shape in the replicate gels. The low quality spots scoring less than 30 quality score were eliminated from further analysis. The remaining high-quality spots were used to calculate the mean value for a given spot, and this value was used as the spot quantity on the standard gel (Supporting Information Document 1). The filtered spot

quantities from the standard gels were assembled into a data matrix of high quality spots from four developmental stages for further analysis. Protein spot detection and quantification were obtained using normalized spot volumes given by PDQuest software using the total spot volume normalization procedure to discard experimental variations in 2-DE gels.

Protein Identification and Expression Clustering

The differentially expressed protein spots were excised mechanically using pipet tips, and in-gel digested with trypsin, and peptides were extracted according to standard techniques.^{35,38} For LC-MS/MS analysis, trypsin-digested peptides were loaded onto a C₁₈PepMap100 column (3 μm, 100 Å, 75 μm i.d., 15 cm) at 300 nL/min (LCPackings), separated with a linear gradient of water/acetonitrile/0.1% formic acid (v/v), and analyzed by electrospray ionization using the ultimate 3000 nano HPLC system (Dionex) coupled to either a 4000 Q-TRAP mass spectrometer (Applied Biosystems) or a Q-Star Pulsar *i* time-of-flight mass spectrometer (Applied Biosystems). The peptides were eluted with a gradient of 10–40% acetonitrile (0.1% formic acid) over 60 min. The MS/MS data were extracted using Analyst software, version 1.5.1 (Applied Biosystems). Peptide analysis was performed through data-dependent acquisition of MS scan (m/z 400–1800) followed by MS/MS scans. Peptides were identified by searching the peak-list against the Potato Genome Sequence Consortium (PGSC_DM_v3.4) (56218 sequences, 16895844 residues) available at <http://potatogenomics.plantbiology.msu.edu/index.html> using MASCOT v.2.1 (<http://www.matrixsciences.com>) search engine. The database search criteria were as follows: taxonomy, all entries; peptide tolerance, ± 1.2 Da; MS/MS tolerance, ± 0.6 Da; peptide charge +1, +2, or +3; maximum allowed missed cleavage, 1; fixed modification, cysteine carbamidomethylation; variable modification, methionine oxidation; instrument type, Default. Protein scores were derived from ion scores as a nonprobabilistic basis for ranking protein hits and as the sum of a series of peptide scores. The score threshold to achieve $P < 0.05$ was set by the Mascot algorithm and was based on the size of the database used in the search. We considered only those protein spots whose MOWSE score was above the significant threshold level determined by Mascot. Proteins with a confidence interval percentage of greater than 95% were considered to represent a positive identification and were also evaluated on the basis of various parameters, such as the number of peptides matched, MOWSE score, and percent coverage of the matched protein. In all of the protein identifications, the probability scores were greater than the score fixed by MASCOT as significant with a P value < 0.05 . The abundance of each identified protein was estimated by determining the protein abundance index (PAI)⁵⁶ and the emPAL.⁵⁷ The corresponding protein content in mol % was calculated as described previously.⁵⁷ For the total number of observed peptides per protein, the unique sequences were counted and were imported to Microsoft Excel (Supporting Information Table 1). Where there was more than one accession number for the same peptide, the match was considered in terms of the putative function. The protein functions were assigned using the Pfam (<http://pfam.sanger.ac.uk>) and InterPro (<http://www.ebi.ac.uk/interpro>) protein function databases. Self-organizing tree algorithm (SOTA) clustering was performed on the log-transformed fold induction expression values across different developmental stages of tuberization using Multi Experiment Viewer (MEV) software.⁵⁸

The clustering was done with the Pearson correlation as distance with 10 cycles and a maximum cell diversity of 0.8.⁵⁹

Isoform Clustal Analysis

The web interface SPECLUST (<http://bioinfo.thep.lu.se/speclust.html>) was used to construct dendrogram based on LC-MS/MS data for potato proteins. For statistical analysis, the processed mass lists were converted into SPECLUST input format files. This program compares m/z peak lists by measuring a distance between each pair of lists. The web interface calculates the mass difference between two peaks taken from different peak lists and determines if the two peaks are identical after taking into account a certain measurement uncertainty (σ) and peak match score (s). The peak match score represents the probability that two peaks with measured masses m and m' have a mass difference equal or larger than $|m - m'|$, given that the mass difference is due only to measurement errors. In this online tool, each m/z peak list is initially assigned to its own cluster. Distances are calculated between each pair of m/z peak lists. The closest pair is found and merged to a new cluster. Distances between the new cluster and each of the old clusters are calculated. This procedure is repeated until there is only one single cluster. Then, a score for matching m/z peaks is created, where the similarity between two m/z peak lists is calculated. This is then translated into a distance measure, which varies between 1 for a completely different set of m/z peaks and zero for a perfect match. This is the starting point in clustering the m/z peak lists and building a dendrogram.

One-Way ANOVA Analysis

One-way ANOVA ($p < 0.05$) with Bonferroni post hoc correction was performed on the stage-specific expression values of both 35S-AmA1 and GBSS-AmA1 potato taking into consideration the four stages to identify significantly changed proteins. Consequently these were used in the software MultiExperiment Viewer (MeV) package⁵⁸ (<http://www.tm4.org/mev/>) for visualization of the data in the heat maps.

Protein Correlation Network

The protein expression data across four developmental stages of potato tubers in both 35S-AmA1 and GBSS-AmA1 were merged together to generate a partial Pearson correlation matrix of the first order with 0.01 alpha value cutoff,⁶⁰ and the resultant SIF file was uploaded in Cytoscape.⁶¹ A partial Pearson correlation matrix of zero order was prepared separately for both overexpressor with 0.0001 cutoff. The common correlations between the two matrices were extracted by uploading the resultant SIF files in Cytoscape and finding the network intersection between the two networks. The initial “combined” network and the “intersection” network were then merged to obtain the final correlation network. All self-correlations and duplicate protein entries were removed to maintain the data nonredundancy. The calculation was performed using the binaries available at <http://mendes.vbi.vt.edu/tikiindex.php?page=Software>.

Study of Tuber Cell

Three 1 mm thick pieces of developing tubers of size between 10.0 and 22.0 mm in diameter from wild-type and AmA1 potato were sampled from 10 weeks of field grown plant and sectioned transversely. The sections were fixed in 2.5% glutaraldehyde and 2.5% PFA in 0.1 M phosphate buffer (pH 7.2) for 2 h, washed in the phosphate buffer for 4 × 15 min followed by H₂O for another 2 × 15 min, dehydrated in a series

of ethanol, and embedded in Technovit 7100. The first set of the sections were stained with 0.01% (w/v) safranin to visualize cell walls. Data collected from cross sections of tubers included cell area of the following regions: (a) cortex, (b) perimedullary zone, and (c) pith.

Isolation, Extraction, Derivatization, and GC–MS Analysis of Tuber Metabolites

For metabolite analyses tubers were harvested at mature stage from wild-type, 35S-AmA1, and GBSS-AmA1 potato. The samples were immediately frozen in liquid nitrogen and stored at -80°C until further analyses. The experiments were performed at least in four replicates. Each replicate consisted of a pool of nine tubers per genotype. Metabolites were extracted and derivatized as described by Roessner et al.⁶² In brief, 100 mg of tuber was homogenized in 1400 μL of 100% methanol with 50 μL of ribitol as internal standard (2 mg mL^{-1}) and extracted for 15 min at 70°C . The extract was mixed with 1 vol of water and centrifuged at 2200g. Subsequently, the methanol/water supernatant was aliquoted to 1 mL and dried *in vacuo* for 9–16 h. The dried residue was redissolved and derivatized using 80 μL of 20 mg mL^{-1} methoxyamine hydrochloride in pyridine for 90 min at 30°C followed by a 30 min treatment with 80 μL of MSTFA at 37°C . Next, 40 μL of retention time standard mixture was added prior to trimethylsilylation. The derivatized extracts were diluted 10-fold in *n*-heptane, and a sample volume of 1 μL was injected in splitless mode into a Shimadzu GCMS-QP 2010 plus. The mass spectrometer was tuned according to the manufacturer's recommendations. GC was performed on an RtxSMS-30 m column with 0.25 mm i.d. and df 0.25 (Restek). The injection temperature was set at 260°C , interface was set at 270°C , and ion source was adjusted to 230°C . Helium was used as the carrier gas at a flow rate of 1 mL min^{-1} . The analysis was performed using the temperature program described in Roessner et al.⁶² Mass spectra were recorded at 2 scan s^{-1} with an m/z 40–600 scanning range. Peaks were assigned and quantified, and all data were normalized to the mean response calculated for the wild-type control of each replicate; to allow comparison between the samples, individual wild-type values were normalized in the same way as per Roessner et al.⁶³ The recovery of small representative amounts of each metabolite through the extraction, derivatization, storage, and quantification procedures has been followed and documented as detailed previously.⁶² Targeted compounds were analyzed and identified by comparing their retention times and mass spectra with those in the NIST or Wiley library.

Gene-Reporter Fusion Construct, Plant Transformation and Histochemical Analysis

The chimeric gene construct pSB3 was made for *Agrobacterium* mediated gene transfer. In brief, the 3.9 kb *HindIII*-*EcoRI* fragment of pSB2²⁰ carrying CamV 35S promoter-*AmA1*- β -GUS-nos terminator cassette was reconstituted into the binary vector pBI121⁶⁴ by replacement cloning. The 0.8 kb CamV 35S promoter from pSB3 was replaced by a 0.8 kb *HindIII*-*XbaI* fragment of the GBSS promoter from pPGB1, and the resulting plasmid was named pSB3G. The pSB3G construct was mobilized into *Agrobacterium* strain LBA 4404, and potato (*Solanum tuberosum* L. var. A16) was transformed as described earlier.²⁰ *In situ* GUS assays were performed by vacuum infiltration of potato tuber slices with the colorimetric substrate X-gluc (5-bromo-4-chloro-3-indolyl β -D-glucuronic acid, cyclohexylammonium salt) as detailed⁶⁴ after few modifications.

Infiltrated slices were incubated overnight at 37°C before they were cleared with 70% ethanol and visualized using Leica microscope.

Statistical Analyses

Analyses of the statistical significance of the data set of normalized spot volumes comparing the three sets of proteomic data (wild-type, 35S-AmA1, and GBSS-AmA1 potato tuber) were performed by unpaired Student's *t* tests using Multi-Experiment Viewer (MeV) v4.8⁵⁸ as described previously.^{65,66} The statistically significant differences with regard to metabolite analysis, tuber diameter, and weight measurements were determined by unpaired Student's *t* tests in Microsoft Excel, while the cell area measurements were performed by ImageJ software (<http://rsb.info.nih.gov/ij/>). $P < 0.05$ was considered statistically significant. The values of all parameters analyzed are from three biological replicates per sample except for the analysis of metabolites where four replicates were used.

RESULTS

Monitoring Changes in Plant Growth and Development

Our earlier findings on AmA1 potato revealed that the expression of the seed storage protein led to a striking increase in tuber yield in addition to the increased nutritive value, particularly protein quantity and quality, irrespective of promoters used to drive the expression of the transgene.²⁰ To further characterize, the following detailed analyses were performed between field grown wild-type and two independent AmA1 lines, 35S-AmA1 and GBSS-AmA1, expressing the transgene constitutively and tuber-specifically. The AmA1 tubers showed almost a week prior germination, and transgenic potato plants were phenotypically distinguishable from the wild-type plants. Eight weeks after plantation, AmA1-plants showed vigorous aerial growth compared to wild-type plants. The plant height and total fresh weight in 35S-AmA1 and GBSS-AmA1 were found to be 1.5-fold higher compared to wild-type (Supporting Information Figure 2A,B). Further, we found that the leaf area of the AmA1-plant was 50% greater than those of the wild-type plants but with no change in the number of nodes (Supporting Information Figure 2C,D). These results were in agreement with the corresponding increase in total biomass by 40% and 60% in the case of 35S-AmA1 and GBSS-AmA1, respectively (Supporting Information Figure 2E). The status of photosynthetic carbon metabolism is recognized as the major determinant of crop growth and yield.⁶⁷ The AmA1 plants displayed up to 35% higher photosynthetic CO₂ fixation (Supporting Information Figure 2F) in addition to substantial increase in tuber size and total yield at maturity (Supporting Information Figure 3).

AmA1 Responsive Changes and 2-DE Analysis

To understand the physiological role of the storage protein AmA1, we carried out the comparative proteomic analysis of the transgenics and their corresponding wild-type at various developmental stages during the tuber life cycle. For each of the tuber stages three replicate 2-DE gels were run that were then computationally combined into a representative standard gel, the first level matchset (Supporting Information Figure 4). The replicates had a correlation coefficient of variation (CV) > 0.8 and $P < 0.05$. Further, the gels showed more than 90% high quality protein spots, suggesting high reproducibility among the replicates (Supporting Information Document 2). For example, 248 spots were detected in the wild-type potato from stolon

stage, but 227 were classified as high quality, while in the case of AmA1 potato 216 spots were classified as high quality spots from 239 and 237 spots, respectively. The normalized volume values and standard deviations of all reproducible spots in three independent biological replicates for wild-type, 35S-AmA1, and GBSS-AmA1 potato lines at each tuber developmental stages have been detailed in Supporting Information Document 1. A second level matchset was then developed, which allowed comparison of the standard gels from each of the stage and a second normalization was done with a set of three unaltered spots identified from across the stages. The filtered spot quantities from the standard gels were assembled into a data matrix of high-quality spots from all the representative gels for further analysis (Figure 2A).

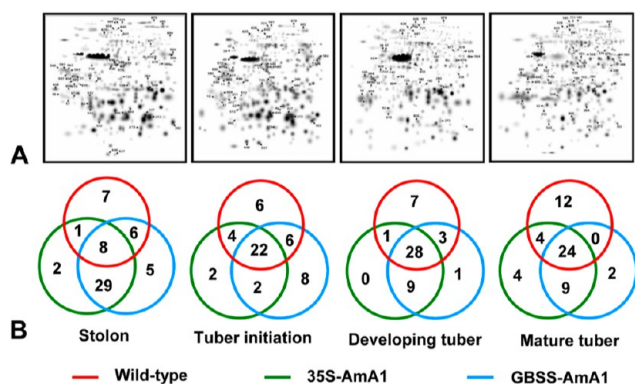


Figure 2. Higher level matchsets and Venn diagrams of AmA1 responsive protein spots (ARPs) detected by 2-DE. (A) The matchset for each individual tuber developmental stage was created *in silico* from three standard gels of each wild-type and AmA1 potato lines. The numbers correspond with the spot ID mentioned in Table 1. (B) Venn diagrams showing the common and exclusive differentially expressed proteins from wild-type and AmA1 potato lines in different stages of tuber development. The areas shown in the diagram are not proportional to the number of proteins in the groups.

Quantitative image analysis revealed around 53 protein spots that changed their intensities by more than 2.5-fold at least in one AmA1 potato line in any of the stages. Figure 2B shows the Venn diagram of the proteins found in different stages. In stolon stage 7 protein spots were exclusive to wild-type potato. The numbers of protein spots exclusive in wild-type were 6, 7, and 12 in the subsequent stages. Twenty-nine protein spots were commonly differential for both AmA1-expressing lines during the stolon stage. Eight protein spots were present in wild-type as well as both AmA1-expressing lines in stage 1 and showed a mixed pattern of either up-regulation or down-regulation. Also, 22, 28, and 24 protein spots in stage 2, stage 3, and stage 4, respectively, were common in wild-type and AmA1 tubers. A total of 80 differentially expressed AmA1-responsive protein (ARPs) spots from all stages were identified by ion trap LC-MS/MS, which are indicated on the higher level matchset of each stage independently (Figure 2A). These 80 protein spots actually account for 45 distinct proteins (Table 1), suggesting 57% unique protein identifications, while the remaining 43% of the identified ARPs either correspond to post-translationally modified forms or may be members of multigene families. Of the 80 ARPs, 3 protein spots were clearly up-regulated and 5 were down-regulated, while 72 of the protein spots showed a mixed pattern of development stage-dependent expression.

Many seemingly well-resolved 2-DE spots were found to contain more than a single protein. In an attempt to minimize co-migration of multiple protein species, we also used a relatively narrow pH range (4–7) for IEF, since the use of narrow range IPG strips facilitates higher resolution.^{28,68} The top-ranked hit resulting from LC-MS/MS analysis has been shown typically to correspond to the most abundant protein among multiple proteins present in a spot.⁶⁹ The spot intensities of different protein constituents were determined using the protein abundance index (PAI) and exponentially modified protein abundance index (emPAI) (Supporting Information Table 2), which have been routinely applied in proteomics workflows.^{57,69–73} Similarly, in this study, if more than one protein was identified in a spot, the relative abundance of each protein was determined by calculating the emPAI from the MS/MS data, and the assumption was made that the most abundant protein would account for the observed regulation. Taken together, the effects of co-migrating proteins on the protein expression ratios observed in 2-DE analyses were deemed negligible. In cases where more than one protein was indicated with a significant score for the MS/MS-derived peptide sequence, the match was considered in terms of the highest ranked hit, molecular mass matches, and emPAI, an approach that has been routinely used for proteomics studies.^{70,74,75} The details of the MS/MS analyses, including protein identification, the score of the identified protein, threshold score, number of matched peptides, % sequence coverage, organism, Potato Genome Sequence Consortium (PGSC_DM_v3.4) <http://potatogenomics.plantbiology.msu.edu/index.html> accession number, theoretical and experimental M_r , and pI values for each protein, peptide sequence, and corresponding peptide score, are provided in Supporting Information Table 1. This observation is common in 2-DE studies for several reasons, such as proteolytic processing, the existence of multiple isoforms, and posttranslational modification(s) leading to changes in the pI, M_r , or both for identical proteins. If the same protein was identified in multiple spots that contained several shared peptides, a differential accumulation pattern was observed for each of the protein species, such as for actin (StC-17, 88, 260, 309, 592, and 625), patatin (StC-39, 42, 45, 86, 471, 512, and 513), and 14-3-3 proteins (StC-11, 602, 604, 605, and 606). In such cases these are listed as independent entities.

Functional Distribution of AmA1-Responsive Proteins

To understand the function of the proteins associated with AmA1-induced changes in the tubers, the differentially expressed ARPs were sorted into different functional categories. Of all the differentially expressed proteins identified, 80 ARPs could be assigned to five functional classes based on their putative roles in the AmA1-response (Figure 3 and Table 1). The differentially expressed proteins appeared to be involved in different key pathways, viz., amino acid metabolism, glycolysis, sucrose and starch synthesis, and cell signaling. Among the identified ARPs, the major functional category corresponded to proteins involved in storage and biogenesis (29%) followed by proteins involved in cell signaling (24%) and protein folding and degradation (24%), while 20% were involved in bioenergy and metabolism (Figure 3).

As expected, proteins involved in storage and biogenesis were found to be the most abundant. Eight ARP spots were identified as patatin, of which StC-39, 86, 471, 472, 512, and 513 showed low expression, while two (StC-42 and 45) showed

Table 1. List of AmA1-Responsive Proteins (ARPs) Identified by MS/MS Analysis

Spot ID	Spot No. ^a	Score ^b	Sol Id ^c	No. of peptides	Stage kinetics ^d S1 S2 S3 S4	% coverage	Th Mw/PI	Exp Mw/PI
Cell signaling								
14-3-3 16R	Stc-11	204	PGSC0003DMT400041810/ PGSC0003DMP400028334	6		20	28.9/ 4.74	31.2/ 4.52
14-3-3 protein isoform 20R	Stc-602	317	PGSC0003DMT400060654/ PGSC0003DMP400040823	6		23	28.7/ 4.72	25.6/ 4.51
14-3-3 16R	Stc-604	446	PGSC0003DMT400041810/ PGSC0003DMP400028334	10		33	28.9/ 4.74	24.5/ 4.59
14-3-3 protein isoform 20R	Stc-605	273	PGSC0003DMT400060654/ PGSC0003DMP400040823	5		13	28.7/ 4.72	27.6/ 4.59
14-3-3 16R	Stc-606	588	PGSC0003DMT400041810/ PGSC0003DMP400028334	13		33	28.7/ 4.72	27.6/ 4.57
Actin-101	Stc-17	54	PGSC0003DMT400076499/ PGSC0003DMP400051821	2		14	21.2/ 5.77	29.0/ 4.57
Actin-101	Stc-88	55	PGSC0003DMT400076499/ PGSC0003DMP400051821	2		14	21.22/ 5.77	37.8/ 5.15
Actin	Stc-260	101	PGSC0003DMT400077954/ PGSC0003DMP400052766	2		8	41.7/ 5.37	39.8/ 5.64
Actin 101	Stc-309	52	PGSC0003DMT400076499/ PGSC0003DMP400051821	2		14	21.22/5. 77	48.5/ 6.72
Actin	Stc-592	100	PGSC0003DMT400047481/ PGSC0003DMP400032139	3		9	41.7/ 5.31	40.5/ 4.26
Actin-101	Stc-625	101	PGSC0003DMT400076500/ PGSC0003DMP400051822	3		21	16.92/ 5.46	75.6/ 5.31
Annexin p34	Stc-655	721	PGSC0003DMT400045665/ PGSC0003DMP400030948	17		47	35.8/ 5.39	49.0/ 5.76
Phospholipase A1	Stc-555	449	PGSC0003DMT400081244/ PGSC0003DMP400054973	12		21	44.9/ 5.97	60.0/ 6.21

Table 1. continued

Spot ID	Spot No. ^a	Score ^b	Sol Id ^c	No. of peptides	Stage kinetics ^d				% coverage	Th Mw/PI	Exp Mw/PI
					S1	S2	S3	S4			
phospholipase A1	Stc-557	582	PGSC0003DMT400081244/ PGSC0003DMP400054973	15					23	44.9/ 5.97	64.8/ 6.53
Phospholipase A1	Stc-558	262	PGSC0003DMT400081243/ PGSC0003DMP400054972	8					17	44.9/ 5.62	60.1/ 6.1
Phospholipase A1	Stc-707	488	PGSC0003DMT400081244/ PGSC0003DMP400054973	10					19	43.9/ 5.71	52.1/ 5.87
Nucleoside diphosphate kinase	StC-352	224	PGSC0003DMT400080042/ PGSC0003DMP400054296	7					32	16.21/6. 32	43.2/ 5.70
Protein Pop3	Stc-637	54	PGSC0003DMT400069960/ PGSC0003DMP400047285	2					13	11.9/ 5.41	15.1/ 5.38
Protein SKIP34	Stc-697	59	PGSC0003DMT400041764/ PGSC0003DMP400028297	2					17	11.4/ 9.18	45.9/ 6.18
Storage and biogenesis											
Patatin-05	StC-39	96	PGSC0003DMT400022567/ PGSC0003DMP400015381	2					7	38.2/ 5.07	54.95/ 4.87
Patatin-05	StC-42	139	PGSC0003DMT400022567/ PGSC0003DMP400015381	4					6	38.2/ 5.07	54.32/ 4.93
Patatin-05	StC-45	104	PGSC0003DMT400022567/ PGSC0003DMP400015381	2					6	38.2/ 5.07	47.00/ 4.79
Patatin-05	Stc-86	105	PGSC0003DMT400022567/ PGSC0003DMP400015381	3					7	38.2/ 5.07	37.5/ 5.19
Patatin-05	StC-471	88	PGSC0003DMT400044052/ PGSC0003DMP400029876	2					21	06.33/7. 98	24.46/ 4.84
Patatin-05	Stc-512	89	PGSC0003DMT400044052/ PGSC0003DMP400029876	2					21	6.3/ 7.98	25.8/ 5.01
Patatin	Stc-513	109	PGSC0003DMT400044052/ PGSC0003DMP400029876	2					21	6.3/ 7.98	24.5/ 5.01

Table 1. continued

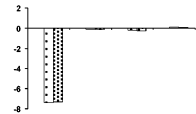
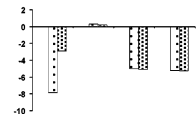
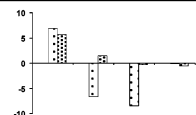
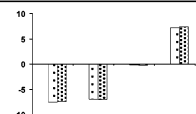
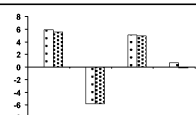

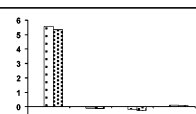
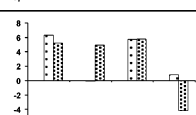
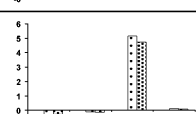
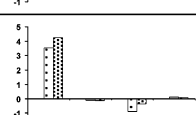
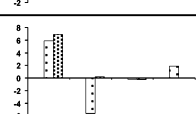
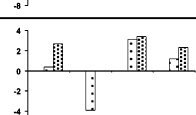
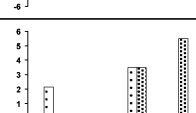
Spot ID	Spot No. ^a	Score ^b	Sol Id ^c	No. of peptides	Stage kinetics ^d S1 S2 S3 S4	% coverage	Th Mw/PI	Exp Mw/PI
Patatin-2-Kuras 4	StC-472	75	PGSC0003DMT400036586/ PGSC0003DMP400024808	3		7	25.6/ 4.89	23.34/ 4.85
Aspartic protease inhibitor 8	Stc-62	217	PGSC0003DMT400024598/ PGSC0003DMP400016822	5		23	24.0/ 7.42	24.7/ 4.83
Kunitz-type enzyme inhibitor S9C11	StC-18	211	PGSC0003DMT400026290/ PGSC0003DMP400017954	7		22	24.6/ 4.60	25.04/ 4.34
Kunitz-type enzyme inhibitor S9C11	Stc-542	362	PGSC0003DMT400026290/ PGSC0003DMP400017954	6		28	24.6/ 4.60	20.4/ 4.99
Kunitz-type enzyme inhibitor S9C11	Stc-609	299	PGSC0003DMT400026290/ PGSC0003DMP400017954	5		25	24.6/ 4.60	28.8/ 4.35
Cysteine protease inhibitor 9	Stc-540	185	PGSC0003DMT400026288/ PGSC0003DMP400017952	4		22	13.8/ 9.17	23.5/ 5.43
Cysteine protease inhibitor 1	Stc-612B	113	PGSC0003DMT400026276/ PGSC0003DMP400017944	2		9	24.7/ 6.89	25.5/ 4.33
Cysteine protease inhibitor 1	Stc-742	258	PGSC0003DMT400026276/ PGSC0003DMP400017944	4		18	24.7/ 6.89	43.2/ 4.92
Cysteine protease inhibitor 1	Stc-743	169	PGSC0003DMT400026276/ PGSC0003DMP400017944	3		14	24.7/ 6.89	44.51/ 4.55
Aminotransferase ybdL	Stc-677	243	PGSC0003DMT400074645/ PGSC0003DMP400050560	5		15	45.1/ 5.96	63.1/ 6.33
Methionine synthase	Stc-682	753	PGSC0003DMT400072279/ PGSC0003DMP400048869	19		21	84.5/ 6.19	85.5/ 6.36
Lipoxygenase	Stc-161	321	PGSC0003DMT400054145/ PGSC0003DMP400036451	11		13	96.7/ 5.56	89.3/ 5.61
Lipoxygenase	Stc-545	1159	PGSC0003DMT400054145/ PGSC0003DMP400036451	24		27	96.7/ 5.56	89.4/ 5.53

Table 1. continued

Spot ID	Spot No. ^a	Score ^b	Sol Id ^c	No. of peptides	Stage kinetics ^d S1 S2 S3 S4	% coverage	Th Mw/PI	Exp Mw/PI
Lipoxygenase	Stc-721	770	PGSC0003DMT400054145/ PGSC0003DMP400036451	18		20	96.7/ 5.56	98.9/ 5.63
Acyl-CoA-binding protein	Stc-648	99	PGSC0003DMT400079309/ PGSC0003DMP400053771	6		50	10.05/5. 82	15.1/ 5.51
Pepper esterase	Stc-552	153	PGSC0003DMT400048357/ PGSC0003DMP400032745	4		11	36.9 / 5.40	55.4/ 5.9
Protein folding/ degradation								
Heat-shock protein	Stc-189	193	PGSC0003DMT400062882/ PGSC0003DMP400042327	5		27	21.91/ 6.86	31.9/ 5.85
Heat-shock protein	Stc-804	85	PGSC0003DMT400053402/ PGSC0003DMP400036029	3		14	38.2/ 5.07	24.7/ 5.99
Small heat shock protein, chloroplastic	StC-375	233	PGSC0003DMT400008351/ PGSC0003DMP400005812	7		24	25.8/ 6.97	18.97/ 6.10
17.6 kD class I small heat shock protein	Stc-272	171	PGSC0003DMT400078006/ PGSC0003DMP400052807	7		25	17.6/ 6.19	20.8/ 6.29
17.6 kD class I small heat shock protein	StC-377	142	PGSC0003DMT400078201/ PGSC0003DMP400052949	4		26	17.6/ 5.83	21.25/ 6.17
Heat shock cognate protein 80	Stc-590	252	PGSC0003DMT400057003/ PGSC0003DMP400038314	5		6	80.2/ 4.96	36.7/ 4.85
Heat shock cognate protein 80	Stc-591	289	PGSC0003DMT400057003/ PGSC0003DMP400038314	7		8	80.2/ 4.96	41.8/ 4.20
Heat shock cognate protein 80	Stc-810	241	PGSC0003DMT400057003/ PGSC0003DMP400038314	4		6	80.2/ 4.96	46.7/ 4.77
Heat shock cognate protein 80	Stc-811	319	PGSC0003DMT400057003/ PGSC0003DMP400038314	6		8	80.2/ 4.96	48.9/ 4.85
Heat shock cognate protein 80	Stc-812	214	PGSC0003DMT400057003/ PGSC0003DMP400038314	4		5	80.2/ 4.96	48.9/ 4.90

Table 1. continued

Spot ID	Spot No. ^a	Score ^b	Sol Id ^c	No. of peptides	Stage kinetics ^d				% coverage	Th Mw/PI	Exp Mw/PI
					S1	S2	S3	S4			
Heat shock cognate protein 80	Stc-813	258	PGSC0003DMT400057003/ PGSC0003DMP400038314	8					9	80.2/ 4.96	48.9/ 4.51
Heat shock protein 83	Stc-593	208	PGSC0003DMT400074374/ PGSC0003DMP400050372	4					13	34.8/ 4.78	41.8/ 4.22
Heat shock protein 83	Stc-809	118	PGSC0003DMT400024594/ PGSC0003DMP400016818	6					8	80.8/ 4.99	46.7/ 4.87
Oligopeptidase A	Stc-628	332	PGSC0003DMT400078667/ PGSC0003DMP400053312	9					12	91.4/ 6.37	85.9/ 5.43
Oligopeptidase A	Stc-640	1406	PGSC0003DMT400078667/ PGSC0003DMP400053312	28					34	91.4/ 6.37	85.9/ 5.43
Oligopeptidase A	Stc-641	713	PGSC0003DMT400078667/ PGSC0003DMP400053312	16					18	91.4/ 6.37	86.0/ 5.49
Disulfide-isomerase	Stc-653	288	PGSC0003DMT400007185/ PGSC0003DMP400004986	6					15	39.5/ 5.62	48.5/ 5.70
Proteasome subunit alpha type	Stc-723	286	PGSC0003DMT400018852/ PGSC0003DMP400012942	9					31	27.1/ 5.63	28.1/ 5.61
Proteasome subunit beta type-6	Stc-680	174	PGSC0003DMT400025311/ PGSC0003DMP400017256	3					14	25.2/ 5.51	26.7/ 5.85
Bioenergy and metabolism											
Glucose-1-phosphate adenylyltransferase small subunit, chloroplastic/amyloplastic	Stc-128	269	PGSC0003DMT400079823/ PGSC0003DMP400054142	6					10	57.2/ 6.73	67.3/ 5.73
Pyruvate dehydrogenase E1 component subunit alpha, mitochondrial	Stc-583	386	PGSC0003DMT400058223/ PGSC0003DMP400039178	9					20	43.2/ 7.63	64.1/ 6.78
Aldo/keto reductase	Stc-551	268	PGSC0003DMT400016484/ PGSC0003DMP400011423	7					18	38.8/ 5.55	55.8/ 5.6
Aldo/keto reductase	Stc-553	97	PGSC0003DMT400016484/ PGSC0003DMP400011423	3					8	38.8/ 5.55	57.5/ 5.43

Table 1. continued

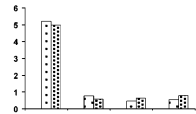
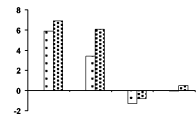
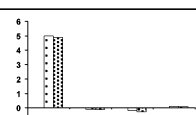
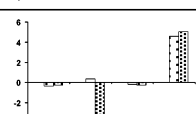
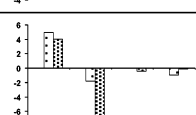
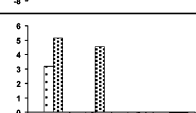
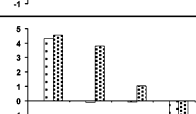
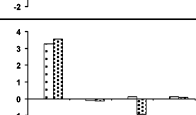
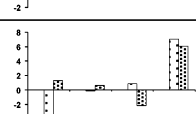
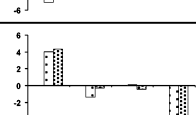
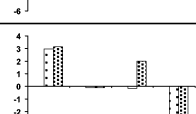
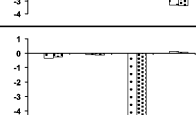
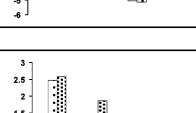
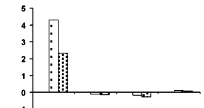
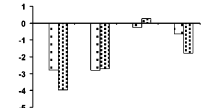
Spot ID	Spot No. ^a	Score ^b	Sol Id ^c	No. of peptides	Stage kinetics ^d				% coverage	Th Mw/PI	Exp Mw/PI
					S1	S2	S3	S4			
Fructokinase	Stc-618	199	PGSC0003DMT400069198/ PGSC0003DMP400046726	2		9	35.0/ 5.61	40.3/ 5.19			
Fructokinase	Stc-644	1032	PGSC0003DMT400069198/ PGSC0003DMP400046726	27		52	35.0/ 5.61	47.3/ 5.51			
Fructokinase	Stc-656	685	PGSC0003DMT400069198/ PGSC0003DMP400046726	18		41	35.0/ 5.61	45.4/ 5.64			
Glyceraldehyde 3-phosphate dehydrogenase	Stc-586	289	PGSC0003DMT400044944/ PGSC0003DMP400030463	8		23	36.6/ 6.34	52.7/ 6.30			
Glyceraldehyde 3-phosphate dehydrogenase	Stc-679	299	PGSC0003DMT400044944/ PGSC0003DMP400030463	9		19	36.6/ 6.34	49.2/ 6.52			
Aconitase	Stc-670	1263	PGSC0003DMT400074489/ PGSC0003DMP400050451	35		25	108.2/ 7.29	99.3/ 6.29			
Enolase	Stc-643	384	PGSC0003DMT400007048/ PGSC0003DMP400004893	10		19	47.8/ 5.58	68.1/ 5.43			
Adenylosuccinate synthetase, chloroplastic	Stc-678	396	PGSC0003DMT400060883/ PGSC0003DMP400040972	9		22	54.7/ 7.12	63.0/ 6.40			
Formate dehydrogenase, mitochondrial	Stc-584	237	PGSC0003DMT400001303/ PGSC0003DMP400000968	10		16	42.02/ 6.64	56.2/ 6.9			
Formate dehydrogenase, mitochondrial	Stc-676	197	PGSC0003DMT400001303/ PGSC0003DMP400000968	5		12	42.0/ 6.64	56.5/ 6.38			
catalase	Stc-684	83	PGSC0003DMP400051216 / PGSC0003DMT400075611	2		4	56.41/ 6.56	70.2/ 6.60			
Ascorbate peroxidase	Stc-192	58	PGSC0003DMT400004360/ PGSC0003DMP400047959	2		7	27.4/ 5.78	28.08/ 5.91			
Miscellaneous											
Ripening regulated protein DDTR10	Stc-595	125	PGSC0003DMT400057934/ PGSC0003DMP400039008	3		22	25.30/ 4.46	35.5/ 4.7			

Table 1. continued

Spot ID	Spot No. ^a	Score ^b	Sol Id ^c	No. of peptides	Stage kinetics ^d				% coverage	Th Mw/PI	Exp Mw/PI
					S1	S2	S3	S4			
P40	Stc-597	49	PGSC0003DMT400069333/ PGSC0003DMP400046814	2					6	32.4/ 4.95	46.6/ 4.49
PR10	Stc-805	230	PGSC0003DMT400003774/ PGSC0003DMP400002693	5					34	17.34/ 5.3	24.7/ 5.99

^aSpot number as given on the two dimensional gel images. The first letters (St) represent the source plant *Solanum tuberosum* followed by the fraction cytoplasm (c). The numerals indicate the spot numbers corresponding to Figure 2. ^bThe significance score ($P < 0.05$) of a protein, as produced by the Mascot algorithm. ^cGene identification number as in PGSC. ^dProtein expression profile represents the average change in spot density at various developmental stage kinetics S1 (stolon), S2 (tuber initiation), S3 (developing tuber), S4 (mature tuber) of 35S-AmA1 (left bar) and GBSS-AmA1 (right bar). The data were taken in terms of fold expression with respect to the control value and were log transformed to the base 2 in order to level the scale of expression and to reduce the noise.

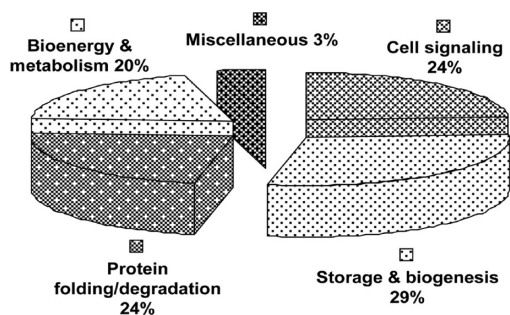


Figure 3. Functional cataloging of AmA1-responsive proteins (ARPs) in potato tuber. The identified ARPs were assigned a putative function using protein function databases and functionally grouped as represented in the pie chart.

mixed pattern. Isoforms of kunitz-type enzyme inhibitors (StC-18, 542, and 609) and cysteine protease inhibitor (StC-540, 612B, 742, and 743) also showed mixed expression. Among others, a very high expression was observed in ARPs encoding lipoxygenase proteins (StC-161, 545, and 721) in mature tubers of both AmA1 plants as compared to their wild-type counterpart. ARPs spots StC-723 and StC-653 were identified to be proteasome subunit alpha type and protein disulfide isomerase (PDI). PDI showed an elevated expression along with GAPDH (StC-679) in the stolon stage. A few other interesting ARPs such as adenylosuccinate synthase (StC-678) and methionine synthase (MetE; StC-682) were also found to be differentially expressed in AmA1 tubers as compared to wild-type. Acetyl Co-A binding protein (ACBP; StC-648), a highly conserved cytosolic lipid-binding protein that binds long-chain acyl-CoA esters with high affinity were also identified in this category.

The second largest category of differentially regulated ARPs comprised of proteins involved in cellular signaling. 14-3-3 proteins (Stc-11, 602, 604, 605, and 606) were the predominant proteins found in this category followed by annexin (StC-655), which showed a higher expression in both type of AmA1 tubers. We also observed differential expression of many of the actin isoforms in wild-type and AmA1 tubers. Few of the actins (StC-592 and 625) were highly expressed during stolon formation and tuber induction, while others (StC-17, 88, and 309) were more actively participating during tuber maturation (Table 1).

A third important class of ARPs identified is presumably known to be involved in protein folding and degradation and includes HSP80 and small HSPs (sHSPs). In this study, HSP80 showed much higher expression in the AmA1 tubers as compared to wild-type. StC-590 and 591 were up-regulated in stolons, while the others StC-810, 811, 812, and 813 showed up to 50-fold increase in developing tubers. Besides HSP80, small HSPs (sHSPs) (StC-272 and 377) also showed differential expression in AmA1 tubers. Oligopeptidase A (OpdA) (StC-628, 640, and 641) was among other in this class.

Proteins involved in bioenergy and metabolism represent the fourth set of ARPs that exhibited differential regulation. A high induction of carbohydrate metabolic enzymes such as fructokinase (FK; StC-618, 644, and 656), aldo/keto reductase (Stc-551 and 553), and glyceraldehyde-3-phosphate dehydrogenase (GAPDH; StC-586 and 679) were observed in AmA1 tubers. The expression profile of FKs showed a significant increase in stolons, while aldo/keto reductase, which showed up to 25-fold increase in AmA1 tubers, is known to catalyze the reduction of aldehyde and carbonyl including monosaccharide to glucose sugar alcohols. GAPDH, which also showed a significant increase in expression, reversibly catalyzes the conversion of GAP into 1,3-bis PGA. Aconitase (StC-670) is an enzyme that catalyzes the stereospecific isomerization of citrate to isocitrate via *cis*-aconitate in the tricarboxylic acid cycle along with glucose-1-PO₄ adenylyl transferase (StC-128) that belongs to the family of transferases specifically transferring phosphorus from nucleotide to glucose and thus participates in starch metabolism. Another spot was identified as formate dehydrogenase (FDH; StC 584 and 676), a soluble mitochondrial enzyme capable of oxidizing formate into CO₂ and abundantly found in nongreen tissues and scarce in photosynthetic tissues showed 20-fold up-regulation in the AmA1 tubers. Another important TCA cycle enzyme, pyruvate dehydrogenase E1 (StC-583), which is the first component of the PDC complex that contributes to transferring pyruvate into acetyl CoA to carry cellular respiration, is up-regulated in the mature AmA1 tuber.

A few differentially expressed ARPs were also found to be associated with miscellaneous functions. Protein spot StC-597 represent P40, while spot StC-805 was found to be PR10 having role in plant protection.

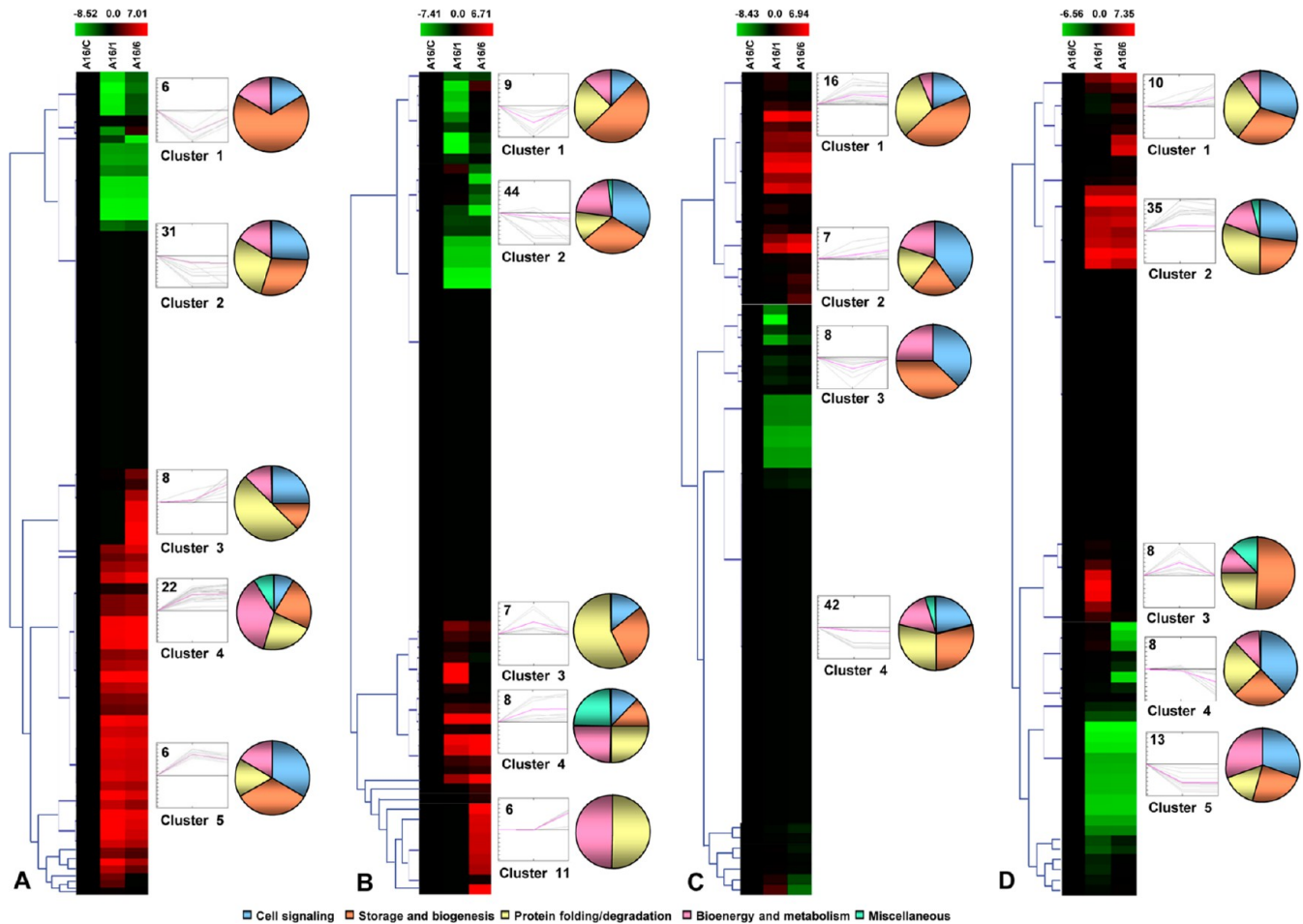


Figure 4. Clustergram of expression profiles of AmA1-responsive potato proteins at various stages of tuberization. The 80 differentially expressed proteins in all four stages were grouped into 11 clusters based on their expression profiles. The self organizing tree algorithm (SOTA) cluster trees are shown at the right side, and the expression profiles are shown in parallel. Each protein is represented by a single row of colored boxes, and each potato line is represented by a single column. Induction (or repression) ranges from pale to saturated red (or green). The expression profile of each individual protein in a cluster is depicted by *gray lines*, while the mean expression profile is marked in *pink* for each cluster. The number of proteins in each cluster is given in the *left upper* corner, and the cluster number is given *below* each expression profile. The cluster with $n \geq 5$ was taken into consideration for the study of co-expression patterns for functionally similar proteins. (A) Stolon, (B) tuber initiation, (C) developing tuber, and (D) mature tuber. Detailed information on proteins within each cluster can be found in Supporting Information Figure 6.

Categorization of AmA1-Responsive Protein Isoforms

Next, we manually analyzed possible isoforms, i.e., similar proteins present in more than one spot. The multiple spots of an identical protein are frequently reported in proteomic studies, presumably due to post translational modification and existence of different isoforms.⁷⁶ These modifications often introduce a variation in the molecular mass and net charge of the protein. The most efficient technique to separate protein isoforms thus remains 2-DE.^{76–78} A closeup of possible isoforms detected by 2-DE and their expression pattern showed 15 unipros appearing as 56 identities. Interestingly, 10 unipros (phospholipase A1, kunitz-type enzyme inhibitor, cysteine proteinase inhibitor, lipoxygenase, formate dehydrogenase, glyceraldehyde3-phosphate dehydrogenase, heat shock cognet protein 80, 17.6 kD classI sHSP, and oligopeptidaseA1) representing 29 identities (isoforms) showed that each set of isoforms had the mixed-regulated change in patterns in abundance in response to AmA1, whereas the other 5 unipros (14-3-3 16R, actin, patatin05, HSP 83, fructokinase) appeared as 23 isoforms that exhibited mixed expression patterns or either up-regulation or down-regulation within each set of

isoforms. In order to confirm whether this set of identified proteins could be possible isoforms to each other, a set of mass spectra representing these 56 putative isoforms from 15 different potato proteins were subjected to cluster analysis by the web interface SPECLUST.⁷⁹ The resulted dendrogram revealed that clusters were dominated by isoforms (Supporting Information Figure 5), e.g., heat shock protein 80 (cluster A and C), oligopeptidase A1 (cluster B), heat shock protein (cluster D), actin 101 (cluster E), patatin-05 (cluster F), phospholipase A1 (cluster G), and lipoxygenase (cluster H). Some of the mass spectra, which were supposed to contain isoforms, did not cluster together as nicely as expected due to the well-known fact that in 2-DE gels multiple proteins are present together in a single spot.⁷⁹ Isoforms are almost always either the products of one gene or of multiple genes that evolved from a single ancestor gene mostly as splice variants. It is generally accepted that multiple isoforms result from sequence-related proteins encoded by distinct genes and/or polypeptide variants encoded by the same gene [splice variants and/or posttranslational modifications (PTMs)]. Therefore, an amino acid sequence alignment of the 15 unipros was

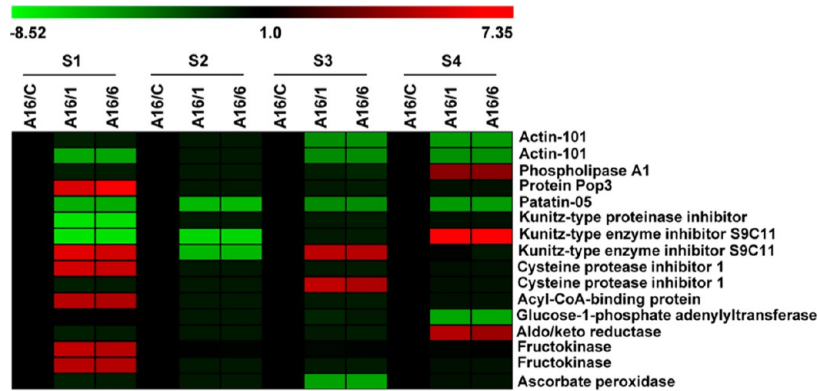


Figure 5. Heat map and one-way ANOVA analysis of ARPs. Heat map of a hierarchical cluster of the 16 protein spots that show significant differences ($P < 0.05$) between stages by Student's t test. The red color represents relatively high expression, and green color represents relatively low expression level.

performed using Bioedit to identify the amino acid sequence variation among the similar proteins present in different spots. The result demonstrated that 8 unipros, namely, 14-3-3, actin, formate dehydrogenase, HSP 17.6, HSP 83, patatin, phospholipase A1, oligopeptidase A1, and proteasome, showed either amino acid substitution, deletion, addition, or inversion (Supporting Information Document 3). In addition, apart from MS/MS peak alignment and amino acid sequence alignment of similar proteins, *in silico* analysis of possible post-translational modifications using different databases were performed. Data analyses revealed that phosphorylation site prediction using NetPhos 2.0, KinasePhos 2.0, DISPHOS 1.3, pKaPs prediction, pepbasepeptide spectra, and P3DB databases showed that phosphorylation sites are present in all the 15 unipros, but phosphorylation could occur in 14 unipros except kunitz-type enzyme inhibitor with respect to the observed pI and MW. Further, glycosylation site prediction using NetNGlyC 2.0, GlyMod, IsoGlyP, and NetOGlyC 3.0 showed that O-glycosylation was absent in all 15 unipros, while N-glycosylation might be present in 14-3-3, phospholipase A1, patatin, kunitz-type enzyme inhibitor, lipoxygenase, HSP 80, HSP 83, and fructokinase. Moreover, only actin variant showed the putative acetylation site based on the prediction of acetylation using Automotif server 2. However, no glycopospho inosides were detected in any of these proteins when searched against Big-PI plant predictor (Supporting Information Table 3). Our results suggest that amino acid sequence variation, phosphorylation, glycosylation, and acetylation might be involved in the generation of these isoforms. Likewise many similar phenomena were also observed in other previously reported proteomics studies.^{80,81} Different isoforms of a protein may function in either a similar or identical ways but sometimes may impart different functions.^{82,83} The function of each of the identified protein isoform thus was analyzed in view of the role of the candidate protein.

Dynamics of AmA1-Regulated Protein Network

To achieve a comprehensive overview of the comparative protein profile during tuber developmental stages, the relative expression profiles of the 80 differentially expressed protein spots were subjected to cluster analysis using the SOTA algorithm.⁵⁹ To group relative protein expression profiles on the basis of similar trends and not of similar expression levels, the Pearson correlation coefficient was used as the distance function. The data were taken in terms of fold expression with respect to the wild-type potato protein expression value.

Furthermore the data sets were log-transformed to base 2 to level the scale of expression and to reduce the noise. Figure 4A–D shows hierarchical clustering of protein accumulation and different relative expression patterns observed in the AmA1 lines versus their wild-type at four tuberization stages. The analysis yielded 11 expression clusters in each stage where only the clusters with $n \geq 5$ were taken into consideration for the study of co-expression patterns for functionally similar proteins. The most abundant group in the early stage of tuber development, viz., the stolon stage, was Cluster 4 with 22 proteins. This group consisted primarily of proteins involved in bioenergy and metabolism followed by storage and biogenesis and showed higher expression in AmA1 lines. The proteins in cluster 5 in stolon stage included 6 up-regulated proteins that may be involved in cell signaling, storage and biogenesis, and bioenergy and metabolism. Clusters 2 and 3 also consist of a majority of proteins of storage and biogenesis class. The largest cluster in stage 2 (cluster 2), stage 3 (cluster 4), and stage 4 (cluster 2) consist of 44, 42, and 35 proteins, respectively. Interestingly, in all four developmental stages proteins from storage and biogenesis had significant contribution to the clusters, which showed up-regulated proteins in the AmA1 lines, although the other classes also had remarkable contributions to these clusters. Detailed information on the proteins within each cluster can be found in Supporting Information Figure 6.

Statistical Testing of Protein Expression

We asked which, if any, ARPs exhibited overall differential expression during tuberization. We used one-way ANOVA to answer the question. A heat map and the expression graph of the developmental stage significant proteins are shown in Figure 5. Based on the ANOVA, 16 differentially expressed ARP spots were found to be significant across the tuber developmental stages that included patatin isomers (StC-86), proteinase inhibitor (StC-472, 542, 609, 612, and 743), and actin isomers (StC-88 and 309).

Correlation Network of AmA1-Responsive Proteins

To elucidate the biological significance of the responses to AmA1, we performed a correlation analysis to assign significance level to the proteins whose expression depends on the AmA1 irrespective of its spatial expression due to different promoters. To address statistical properties of these networks, we quantified correlations between connectivity of interacting nodes using Pearson's correlation coefficients

(PCC). We obtained a protein correlation network of 26 nodes and 35 edges. The resultant network incorporated 26 out of 80, 45 proteins eligible for the analysis. The network at an alpha value of 0.01 contains two modules (M) and six small correlation groups (SC) of just 2 and 3 proteins (Figure 6).

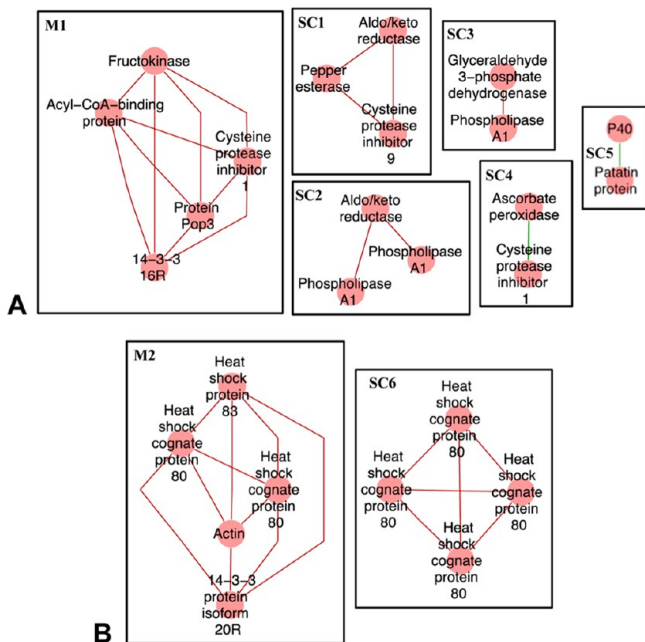


Figure 6. Functional correlation network of ARPs. The ARPs were subjected to partial Pearson correlation analysis using a matrix of the first order with 0.01 alpha value cutoff, and the resultant SIF file was uploaded in Cytoscape for visualization. Red and green edges correspond to positive and negative partial correlation, respectively. Boxes indicate modules (M) and small correlation groups (SC) of known function. (A) M1 and SC1–SC5 represent the nutritional subnetwork, and (B) M2 along with SC6 denote the other subnetwork related to growth and development.

We succeeded in extracting functional relationships using Pearson correlation method from the concatenated data sets based on the physiological role of the AmA1-responsive proteins that categorized the modules and small correlation groups into two subnetworks, viz., nutritional network (Figure 6A) and growth and development related network (Figure 6B).

The nutritional subnetwork was composed of module 1 (M1) and five small correlation groups (SC1–SC5) with cysteine proteinase inhibitor, aldo/keto reductase, ascorbate peroxidase, actin, patatin, acyl-CoA-binding protein, esterase, glyceraldehyde 3-phosphate dehydrogenase, fructokinase, and phospholipase A1. Proteins in M1 are related to storage protein accumulation, assembly of fatty acid residue and proteins related to calvin cycle of PSII. In M1, cysteine protease inhibitor1 and fructokinase were found to be positively correlated, while both cysteine protease inhibitor and 14-3-3 16R showed positive correlation with acyl coA binding protein. Interestingly, grouped with these proteins is P40 in SC5 that is negatively correlated with patatin, indicating its role in storage and biogenesis, and thus we hypothesize that it might have similar such function. The proteins in SC1 are involved in fatty acid metabolism and protein storage. Esterase was positively correlated with one of the cysteine protease inhibitor isoform. There is aldo keto reductase, a calvin cycle protein positively correlated with esterase in this group that could thus be co-

regulated. SC2, SC3, and SC4 showed candidates related to storage protein, cytoskeleton, fatty acid catabolism, and redox homeostasis. Phospholipase A1 was positively correlated with aldo/keto reductase, while cysteine proteinase inhibitor was negatively correlated with ascorbate peroxidase in an individual association.

The other subnetwork related to growth and development including module 2 (M2), and one small correlation group (SC6) consists of heat shock protein 83, 14-3-3 16R, actin, heat shock protein 80, pop3, GAPDH. The cascade in M2 contains proteins involved in cell growth and osmotic regulation, tuber maturity, and development and was positively correlated among each other. Proteins found in M2 were positively correlated with each other, viz., actin has direct positive correlation with hsp, whereas SC6 contains a cascade of heat shock protein 80 involved in protein homeostasis and morphological evolution.

Comparative Analysis of AmA1-Responsive Metabolites

Having found the differential proteome response in AmA1 tuber, we further extended the evaluation and analyzed the metabolite pools using GC–MS to understand the impact of ARPs on the primary metabolism in 35S-AmA1 and GBSS-AmA1 tubers. Targeted compounds were identified and analyzed to corroborate with the enzymatic pathways based on the proteomic analysis. The majority of the compounds detected were found to be altered within AmA1 tubers, in agreement with the data obtained from proteomic study. Sixty-two metabolites were identified in wild-type, and 62 and 56 metabolites detected in 35S-AmA1 and GBSS-AmA1 tubers, respectively, showed differential expression with high level of certainty. Of the 56 common metabolites between wild-type and transgenic tubers, 42 (75%) were up-regulated, whereas 14 (25%) were down-regulated. The primary metabolites related to AmA1-responsive pathways are shown in Figure 7. Consistent with the proteomic data, there was remarkable difference in metabolite profile with high accumulation of amino acids in mature AmA1 tubers. Metabolites related to amino acid biosynthesis were up-regulated. A notable trend in the levels of amino acids was the increase in the concentration of hydrophobic amino acids such as isoleucine, glycine, leucine, and alanine by a factor of 3.30, 2.74, 4.00, and 1.50 in 35S-AmA1 and 2.12, 1.53, 3.73, and 4.49 in GBSS-AmA1. Furthermore, hydroxyl amino acids displayed a small but significant increase in concentration in 35S-AmA1 tubers. Methionine, a sulfur amino acid, showed striking increase of 6–9 times, which indicates the mobilization of protein biosynthesis pathway in addition to the accumulation of S-adenosyl methionine (SAM). By contrast, glutamic acid showed a transient accumulation of 1.83-fold, which likely suggests its rapid mobilization as substrate for acidic amino acid biosynthesis. Aspartate represents an important connection between amino acid and carbohydrate metabolism. Increase in aspartate shifts the metabolic pathway toward the citric acid cycle by the conversion of aspartic acid to oxaloacetate. Nevertheless, level of aspartic acid was relatively lower than other amino acids, reflecting the decrease in carbohydrate metabolism (Figure 7A). Indeed, metabolites related to carbohydrate biosynthesis were significantly down-regulated or in steady state level. As expected, the levels of monosaccharides or disaccharide (glucose, fructose, and sucrose) were decreased or maintained a steady state in AmA1 tubers, in accordance with the proteome data (Figure 7B). Moreover, some metabolites, namely, saturated and unsaturated fatty acids, responded in opposite

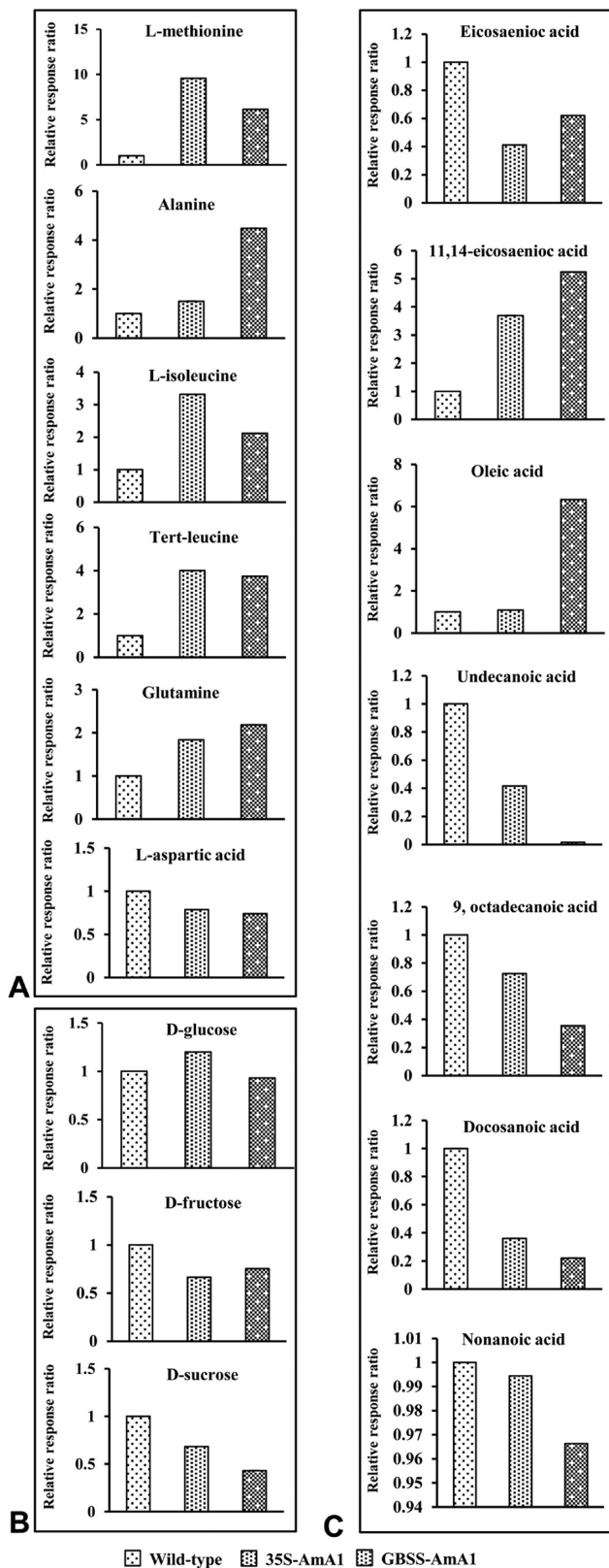


Figure 7. Comparison of primary metabolite levels in mature AmA1 tubers with those in tubers of wild-type. (A) Amino acids, (B) Sugars, and (C) Fatty acids. Data are normalized to the mean response calculated for wild-type levels of each replicate (to allow comparison between replicates, individual wild-type values were normalized in same way). Values presented are the mean \pm SE of four independent determinants.

directions probably due to conversion of free fatty acid to esterified fatty acids in AmA1 tubers. Unsaturated fatty acids such as eicosaenoic acid was down-regulated, while 11,14-eicosaenoic acid and oleic acid showed a significant increase of 3.69–5.25-fold in both 35S-AmA1 and GBSS-AmA1 tubers. In the case of the saturated fatty acids, undecanoic acid, 9, octadecanoic acid, docosanoic acid, and nonanoic acid showed reduced level in the AmA1 tubers (Figure 7C).

Investigation of Tuber Cell Architecture

To investigate the underlying mechanism that might have resulted in increased tuber yield in AmA1 plants, the cell architectures were studied in different cell layers at tuber maturity and compared with that of wild-type tuber (Figure 8A). The cell area was found to be identical in the epidermal region of the tubers. However, there was an increase in cell area in the cortex and perimedullary region, which was increased further in the pith tissue. The cell area in cortex region is 36.21% higher, while in the perimedullary and pith region it was 76.62% and 98.03% higher in transgenic tuber (Figure 8B). To further test this observation, we used AmA1 fused to β -glucuronidase reporter gene construct driven by GBSS promoter (Figure 8C). The AmA1 tubers displayed high level expression of AmA1-GUS fusion protein in the cortex, medullary region, and pith (Figure 8D). The *in planta* localization of the AmA1 protein in potato tuber was found to be correlated with cell growth in this study. Taken together, these results suggest that increase in cell division especially in the perimedullary region might contribute to more storage tissue in AmA1 tubers.

DISCUSSION

Quality and productivity of agricultural crops is a complex function of the acquisition of resources and their distribution within the plant to harvestable components. Photosynthesis and nutrient capture are the primary elements of overall biomass production, but it is the allocation of assimilated resources within the developing plant that determines the proportion of biomass that can be utilized. Plants are programmed to maintain a constant rate of photoassimilate synthesis, translocation, and supply to storage organs, i.e., the sink tissue, which is competitive, and thus the photoassimilate is partitioned to the active sinks. Sink strength is one of the significant factors that determine the direction of photoassimilate translocation, while sink size represents the total mass of sink primarily composed of carbon and nitrogen compounds, including storage proteins.⁸⁴ Storage proteins account for 10–60% of total dry weight in plants.⁸⁵ Further, it is known that sink strength is closely related to growth and in turn productivity and is influenced by cell turgor and hormones.

Our earlier findings and the result in this work showed a definite role of seed storage protein AmA1 toward nutritional enhancement and growth. Thus, our aim primarily was to investigate the regulatory and functional protein network operating in response to AmA1 sensing. We demonstrate that AmA1-regulated functional protein network and its combinatorial effect cause the protein enhancement and determine the organ development, tuber in particular. The introduction of the AmA1 gene by means of a constitutive and tuber-specific promoter may lead to numerous changes within the plant proteome that may be related to many pathways or general expression variation of individual proteins. Although the

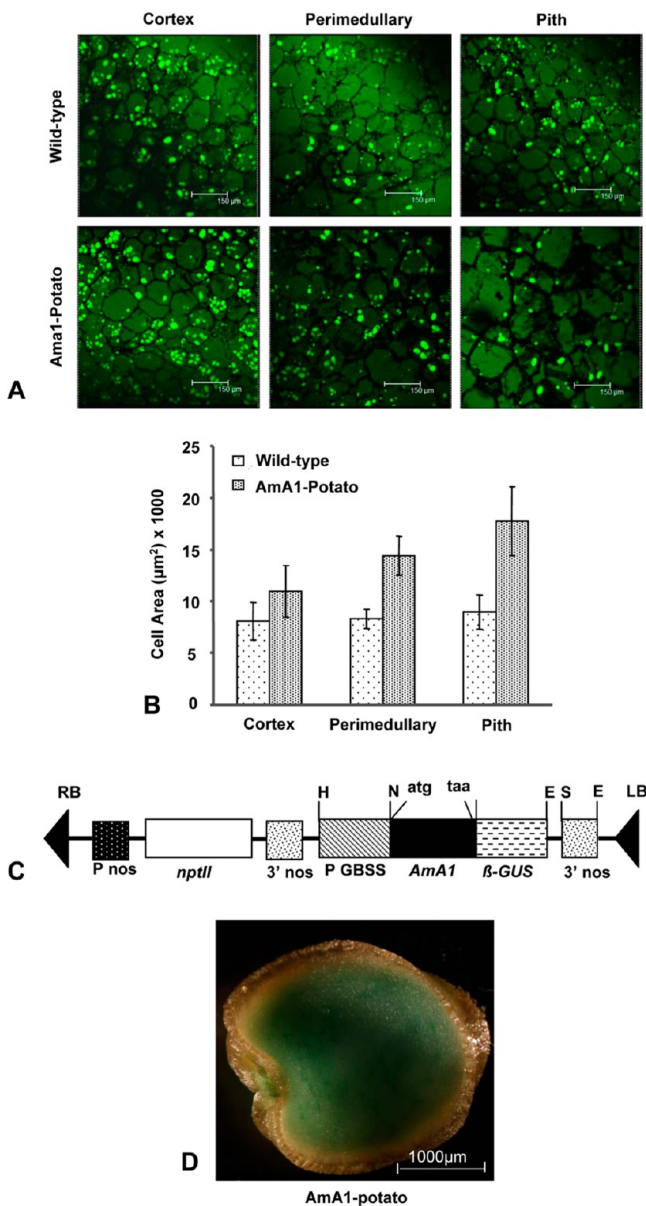


Figure 8. Cell architectures and accumulation of AmA1 protein. (A) Photomicrographs showing cells in the cortex, perimedullary, and pith regions of the wild-type and AmA1 potato tuber. (B) Graph showing cell area of different regions along the transverse axis of potato tubers. Data represent mean \pm SD of three measurements. (C) Schematic representation of the AmA1- β -GUS chimeric gene construct pSB3G. (D) Cross sections of potato tubers of wild-type and AmA1 lines. Each cross-section was treated with β -GUS stain for histochemical analysis. Blue-colored signals indicate accumulation of AmA1 protein in AmA1 potato. Most of AmA1 was colocalized in medullary and pith tissue.

comparative proteome analysis revealed the set of AmA1-target proteins differentially expressed in both 35S-AmA1 and GBSS-AmA1 tubers, the level of expression cannot be related to the presence of AmA1 with utmost confidence for all. Hence, we chose to interpret the results by taking into consideration the expression of AmA1 target proteins at four developmental stages for the two AmA1 potato genotypes together, considering the spots that show a correlation pattern in both, and also separately to check the correlation between spots individually in both AmA1 lines to identify the common

correlations. The final correlation network obtained by merging the initial “combined” network and the “intersection” network revealed a total picture of all sets of protein spots that may be correlated in both of the data sets with a decent statistical confidence. Verifying the biological relevance of the recovered networks is difficult, since many interactions between proteins are currently unknown. These correlation networks are also known to be phenomenological, i.e., many connections do not correspond to direct physical interactions between gene product and promoter elements, but to a complicated action through more complex regulatory pathways involving the proteome and metabolome.^{86,87} We verified the biological relevance of the inferred networks by using the Gene Ontology Term Finder to investigate if the subnetworks contain a high proportion of functionally related proteins. Indeed, the inferred networks had high significance scores, implying that the probability of grouping them by chance is less. Revealing the modular structures in the obtained biological networks helps us to understand cell function. A model representing the regulatory and functional network of AmA1 target proteins is depicted in Figure 9.

Metabolic and Functional Protein Network Involved in AmA1-Associated Nutritional Enhancement

The synthesis and accumulation of nutrients is a sophisticated and highly regulated process, wherein many different proteins act to ensure that a gene is translated and the products are protected at the right time in a cell. This requires synthesis of different energy molecules and amino acids, the building block of proteins, and participation of different molecules with a protective role in the formation and stabilization of the newly synthesized proteins.

Several ARPs known to be involved in the synthesis of energy molecule and amino acids were identified in the present study. One of the major regulators MetE (StC-682) involved in *de novo* synthesis of methionine and regeneration of the methyl group SAM by catalyzing the transfer of a methyl group from methyltetrahydrofolate to homocysteine was highly up-regulated in AmA1 tuber at the stolon stage. In plant, it has been estimated that about 20% of the Met is incorporated into proteins, while 80% is converted to SAM.^{80,81} Furthermore, it was interesting to note the increased accumulation of methionine in AmA1 mature tubers, which provides further support to our proteome data. From the many changes determined following metabolite analysis, the one of particular interest was the increase in aliphatic amino acids, namely, alanine, leucine, and isoleucine. This small group of branched-chain amino acids is essential for humans.^{88,89} Aminotransferase (StC-677) has an important role in the biosynthetic pathways of these amino acids and the methionine chain elongation cycle of aliphatic glucosinolate formation. Methionine occupies a central position in cellular metabolism: as a protein constituent, in the initiation of mRNA translation, and as a component of the regulatory molecule, SAM. AmA1 being a storage protein acts as sink to accumulate essential amino acids such as methionine and cysteine, and thus its expression requires a higher level of MetE, the key enzyme in methionine biosynthetic pathway.⁹⁰ SAM functions as a primary methyl group donor and most interestingly, as a precursor for metabolites, such as ethylene,⁹¹ that activates adenylosuccinate synthase (StC-678). We observed elevated expression of adenylosuccinate synthase in stolon, the key enzyme for *de novo* synthesis of adenosine monophosphate (AMP), suggest-

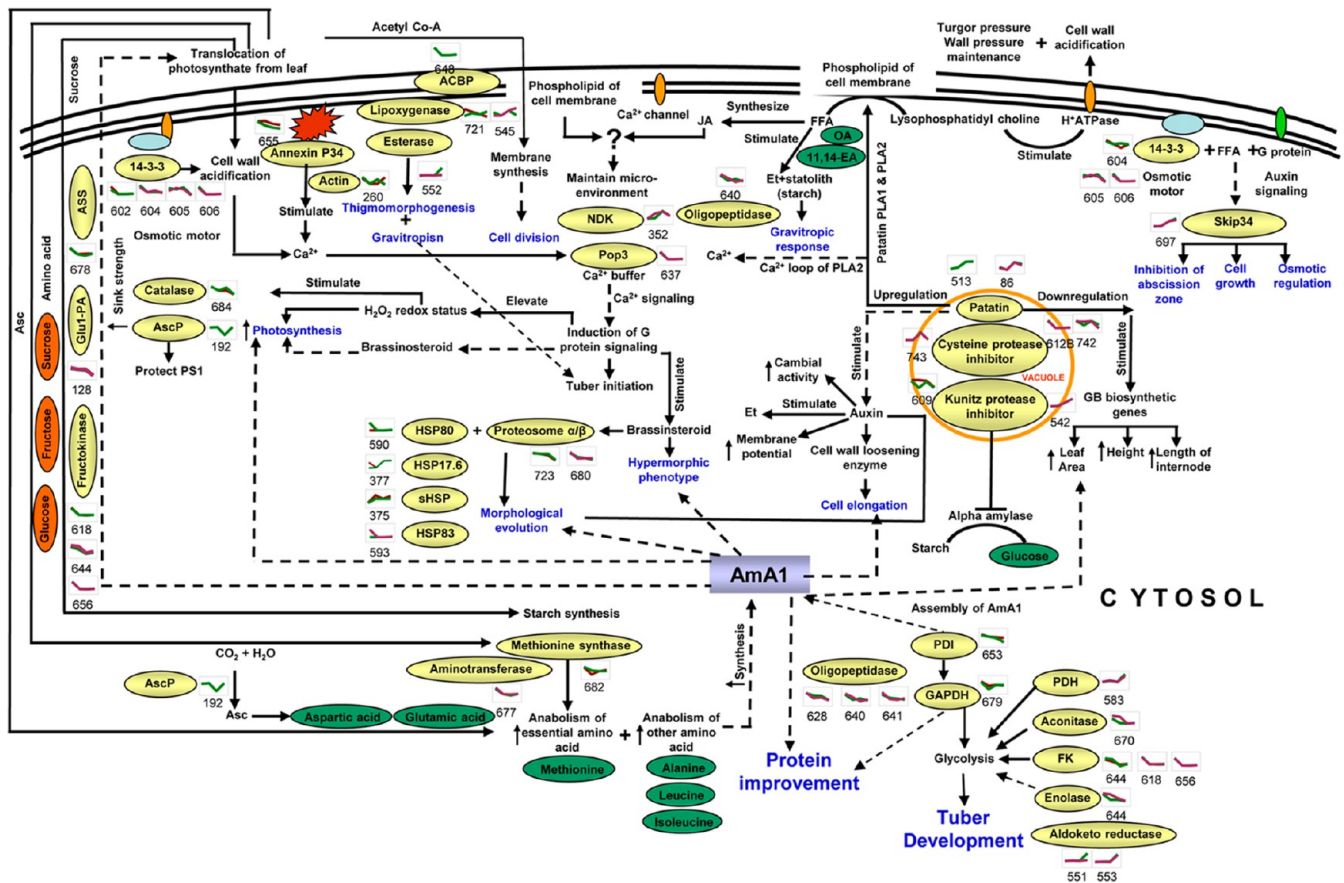


Figure 9. Pathway involved in AmA1-activated regulatory and functional network. Proteins identified in this study are indicated in the yellow boxes. Green circles represent metabolites up-regulated, and orange circles indicate down-regulated metabolites. Graphs are the representatives of expression profile of individual protein, and number given below in each graph indicates the protein identification number. ASS, adenylylsuccinate synthase; AscP, ascorbate peroxidase; ACBP, acyl CoA binding protein; HSP80, 80 kDa heat shock protein; HSP17.6, sHSP; HSP83, 83 kDa heat shock protein; PDH, pyruvate dehydrogenase; Glu 1-PA, Glucose 1- PO_4 adenyl transferase; NDK, Nucleotide dikinase; FK, fructokinase; GAPDH, glyceraldehyde 3-phosphate dehydrogenase; PDI, protein disulfide isomerase; 11,14-EA, 11,14-eicosadienoic acid; OA, oleic acid.

ing its role in increasing the tuber primordial formation that might lead to increased sink strength. It has been documented that increased synthesis of AMP is advantageous for the production of high energy donor molecules such as ADP and ATP for starch synthesis.⁹² In AmA1 potato increased level of glutamate, a key molecule in protein synthesis and the five carbon chain precursor of all tetrapyrroles including 5-aminolevulinic acid,⁹³ presumably bind to tRNA much efficiently, thereby increasing the rate of protein synthesis.

As sucrose is the major transport form of fixed photo-assimilates in developing potato tubers, the mechanism responsible for the immediate metabolism of sucrose is of particular interest. Sucrose cleavage after unloading out of the phloem into the sink organ, such as potato tuber, is known to be mediated by invertase that cleaves sucrose to glucose and fructose, which are transported to the cytosol for further metabolism. The differential proteome data also revealed the presence of a number of glycolytic and starch biosynthetic pathway enzymes. These proteins are known to have bioenergy-related function. While glucose 1- PO_4 adenyl transferase (StC-128) might be involved in starch synthesis using glucose and ADP as substrate, fructose is acted upon by fructokinase (FK; StC-618, 644, and 656) to form fructose-6-phosphate to maintain a balance between sucrose synthesis and starch formation. FK has little impact on glycolysis and

subsequently on metabolites downstream to glycolysis. Therefore, in AmA1 tubers glycolytic proteins such as aldo/keto reductase (StC-551 and 553), enolase (StC-643), pyruvate dehydrogenase (StC-583), and GAPDH (StC-586 and 679) are highly induced to supply inorganic phosphate for starch synthesis. The effect of rigorously modulating fructokinase in potato tuber has been well established,⁹⁴ and potato plants exhibiting reduced FK are characterized by fewer tuber numbers and reduced yield.⁹⁵ Thus, highly up-regulated expression of fructokinase- and fructokinase-like proteins (StC-618, 644, and 656) clearly supports our data of increased tuber yield. The metabolite profiling of carbohydrate complement provides further support to our claim that concentration of sucrose and fructose were low in AmA1 tuber, whereas glucose maintained a steady state level. It has long been known that reductases regulate glycolysis and starch synthesis, the most important component of carbon assimilation, by detoxifying glycolysis-derived reactive carbonyl in rapidly metabolizing cells. Aldo/keto reductase (AKR) superfamily (StC-551, 553) catalyzes mainly the reduction of carbonyl groups or carbon-carbon double bonds of a wide variety of substrates, including steroids, mainly brassinosteroid. Monomeric cytosolic form of AKR plays a crucial role in osmoregulation, an important process for the acquisition of desiccation tolerance in plants. Besides this, members of the

AKR family have been shown to be effective in the detoxification of lipid peroxidation and/or glycolysis-derived reactive carbonyls.⁹⁶

While actin filaments may enhance protein synthesis by providing a scaffold for specific assembly of the translational machinery, a recent study indicates that hexokinase (HXK)-dependent glucose signaling might require a previously unrecognized interaction with the actin cytoskeleton.⁹⁷ Our observation on differentially regulated actins (StC-88, 260, 592, 625) might provide a novel gateway to understand an important interface between protein synthesis, glucose signaling, and tuber development. Actin is known to be encoded by a multigene family with high divergence.⁹⁸ Further, different isoforms are likely to be expressed at a particular time depending on the physiological status of the cell. Therefore, highly regulated expression of actin genes as revealed in this study and their subsequent post-translational modifications are probably required during tuber development. Although class I (StC-272 and 377) and class II (StC-371) sHSPs were down-regulated at the initial developmental stages of AmA1 tubers, their subsequent up-regulation at tuber maturity implicates their role in reserve accumulation as has been reported earlier in pea seeds.⁹⁹ Among the five conserved families of HSPs, HSP83 in solanaceae, the ortholog of HSP90 are the most prevalent in plant. Plants synthesize multiple sHSPs encoded by six nuclear multigene families; each gene family encodes proteins that are destined to a specific cellular compartment (i.e., cytosol, chloroplast, ER, and mitochondria). Diversification of sHSPs and presence of their isoforms probably reflect a molecular adaptation to stress conditions unique to plants. Although HSP83 chaperones are constitutively expressed in most organisms, their expression increases in response to stress in both prokaryote and eukaryote. Expression of HSP90 isoforms play a substantial role in the subunit–subunit interactions that stabilize homo- and/or heterotrimers of developmentally regulated proteins in *Arabidopsis*.¹⁰⁰ Post-translational regulation as opposed to transcriptional or translational regulation allows a quick adjustment of protein activity. Amino acid sequence variation as revealed in our study might attribute to the diverse functions of the orthologs belonging to these multigene family.

The 14-3-3 proteins (StC-11, 602, 604, 605, and 606) have risen to a position of importance in biology, having been shown to regulate many crucial cellular processes because of specific phosphoserine/phosphothreonine binding activity for target protein that shuttle from the cytoplasm to the plasma membrane or nucleus,¹⁰¹ modulating their localization.^{102,103} The first plant 14-3-3 protein was identified as part of a protein–DNA complex,¹⁰⁴ indicating a role of 14-3-3s in the regulation of transcription. It is well established from the studies of animal and plant that isoforms have differential expression patterns, different subcellular localization, and/or functions.^{83,105–107} We identified isoforms of 14-3-3 protein known to be the product of a multigene family, with many organisms having 10 or more family members.¹⁰⁸ It is known that diversity of 14-3-3 is due to alternative splicing.¹⁰⁹ One possible reason for isoform diversity is simply to ensure fundamental 14-3-3 presence in all cell types, where 14-3-3 function is required. Differential expression pattern of 14-3-3 and actin isoforms in response to AmA1 might result from post-translational modification with phosphorylation or dephosphorylation state. These results suggest that isoforms of certain unipros may play the same or different roles in AmA1

tubers. Additional studies showed that 14-3-3 proteins (StC-11, 602, 604, 605, and 606) are known to function in regulating the activities of metabolic enzymes and transcription factors,¹¹⁰ besides their role as osmotic motor.¹¹¹ Increase in protein content presumably elevates the turgor pressure in AmA1 expressing cells. Upregulation of 14-3-3 proteins, known to act as osmotic motor might stimulate H⁺ ATPase pump and cell wall acidification, thereby balancing the turgor pressure that leads to cell elongation and induction of Ca²⁺ and G-protein signaling in the AmA1 potato. We observed increased expression of pop3 (StC-637) in GBSS-AmA1 tuber whose cross-talk with brassinosteroid is reported earlier.^{112,113} It is interesting to note BR-induced H₂O₂ may alter cellular redox status and induce changes in expression of photosynthetic genes ultimately leading to the increased photoassimilate. Catalase (StC-684) showed higher expression in early stage with subsequent decrease in the later stage of tuberization, suggesting that this enzyme might try to degrade H₂O₂. However, in later stage of tuberization H₂O₂ is utilized to maintain redox status of photosynthetic gene like thioredoxin. Expression of kunitz-type enzyme inhibitor (StC-609) known to inhibit activity of amylase,¹¹⁴ which cleaves starch to glucose that help in maintaining turgor pressure during tuber development.

One of the major regulations in the process of formation and stabilization of newly synthesized amino acids and proteins is carried out by the highly polymorphic low-molecular weight protease inhibitors, such as cysteine proteinase and kunitz-type enzyme inhibitors (StC-540, 542, 609, 612B, and 743), the second major class of storage proteins in tubers that showed maximum abundance in AmA1 mature tuber and were found to exist as product of multigene family having N-type of glycoform. They have been documented with multiple functions as enzyme inhibitors^{115,116} or agents of defense against pathogens or insects.¹¹⁷ Protein disulfide isomerase (PDI; StC-653), a multifunctional protein that catalyzes the formation of disulfide bonds during protein biogenesis and is considered to be of fundamental importance for stabilizing the tertiary and quaternary structures of many proteins that are processed through the endomembrane system,¹¹⁸ showed an elevated expression in stolon. The wheat PDI reported to be colocalized with storage proteins in protein bodies is involved in the assembly of storage proteins within the ER.¹¹⁹ Also, PDI markedly increases reactivation of GAPDH and prevents their aggregation.¹²⁰ Thus the identical expression profile of PDI and GAPDH in stolon in the transgenics is indicative of their cooperative role in protein enhancement during tuber development. Oligopeptidase A1 (StC-628, 640, and 641) may play a specific role in the degradation of signal peptides after they are released from precursor forms of secreted proteins.^{121,122} It is a thiol-dependent metalloproteinase, which can cleave and thereby modulate the activity of many proteins and is involved in seed germination, recycling damaged proteins, regulating aging processes, and modifying proteins to perform specific purposes in plant cell.¹²³ Some patatin isoforms (StC-42 and 45) showed elevated expression in stages 3 and 4. It is known that proteolysis of vegetative proteins facilitates accumulation of storage proteins with release of cell wall loosening enzymes for increased assimilation of C and N macromolecules.

Regulatory and Signaling Network Involved in AmA1-Induced Growth and Organ Development

Numerous studies have shown that the productivity of agricultural crops is a complex function of the acquisition of resources and their distribution within the plant to harvestable components. Photosynthesis and nutrient capture are the primary elements of overall biomass production, but it is the allocation of assimilated resources within the developing plant that determine the proportion of biomass that can be utilized. Regulation and coordination of the communication within the molecular signaling network is essential for organ growth and development, including potato tuber. It is an expression of the genetic program that directs the activities and interaction of individual cells. Growth, an irreversible increase in volume and manifestation of new protoplast, is accomplished by increase in fresh and dry weight from stolon to tuber, and this gradual transition is referred to as homoblastic development.¹²⁴ Functional sorting of ARPs identified 8 spots as patatin, the primary storage protein (40% of soluble protein) that is considered a critical nutritional component in potato tuber. Patatin, a group of glycoproteins in potato, is encoded by a multigene family.^{125,126} Isoforms of patatin separated in 2-DE analysis contained proteins with two predicted molecular masses of approximately 41.17 and 42.56 kDa, respectively. The size of this difference in the molar mass (1390 Da) is on the order of one carbohydrate moiety. This clearly emphasizes the possibility of the existence of a patatin glycoform (Supporting Information Table 3). These are mostly expressed in developing tubers and highly activated after tuber initiation. Patatin has been shown to play an important role in plant defense and antioxidant activity due to their enzymatic function of lipid acyl hydrolase that exhibits phospholipase A1 and A2 (PLA1 and PLA2) activity.¹²⁷ Increasing evidence suggests that patatin genes exhibit alteration in chromatin state and differential transcriptional regulation during the developmental transition from stolon to tuber.¹²⁸ The reaction catalyzed by this protein generates free fatty acids (FFA) from its substrate phospholipids, which can subsequently be converted to jasmonic acid (JA) and its derivatives.¹²⁹ FFA along with statolith containing starch and the release of ethylene from the JA-mediated catabolic pathway stimulate dia-gravitropic response that results in the growth of stolons at right angle to the pull of gravity, which reduce mutual shading and efficient capture of photosynthate making even large size tubers capable of capturing sucrose.¹³⁰ Metabolite profiling of fatty acids also confirmed similar change in stolon, wherein most of the saturated and unsaturated fatty acids showed up-regulation, implying that FFA is utilized for gravitropic response, while in mature tuber the concentration of fatty acid is less, suggesting that it is utilized for making fat in plasma membrane. Patatin isoforms (StC-42, 455, 512, 513, 471, and 472) showed down-regulation in AmA1 potato. Growth is an interplay of many phytohormones, and it is known that down-regulation of patatin increases the synthesis of gibberellin's¹³⁰ transcription factors and might lead to the observed increase in leaf area, height, and length of internode in AmA1 plants (Supporting Information Figure 2A–F).

Lysophosphatidylcholine (LPC), the other product of PLA activity of patatin (StC-42, 455, 512, 513, 471, and 472), is biologically active in a number of important cellular signaling pathways involving auxin-induced cell elongation that stimulates the H⁺-ATPase pump in the plasma membrane, which is known to be a key regulator in the induction of cell wall

acidification leading to elongation.¹²⁸ Therefore, this activity possibly is one of the factors in increasing cell division and cell enlargement that contribute to the increased growth of tubers in AmA1 potato. It is widely accepted that the growth of the perimedullary zone produces the largest portion of the tissue in mature tuber^{131,132} as has been documented in this study (Figure 8A). Of the phospholipase A, PLA2 has a PA2c domain that contains the highly conserved Ca²⁺-binding loop that is important for Ca²⁺-dependent membrane translocation.^{133–135} Ca²⁺ initiate signal transduction that proceeds in a linear fashion and cross-talk with pathway initiated by POP3 (StC637), a Ca²⁺ receptor and representative of EF-hand proteins classified as Ca²⁺ buffer,¹³⁶ which act as sensor to interact with specific targets and stimulate G-protein signaling. The up-regulation of phospholipase A1 (StC-555, 557, 558, and 707) in AmA1 tuber clearly indicates its involvement in tuber maturation. It has also been shown to induce brassinosteroid biosynthetic genes resulting in a hypermorphic phenotype with increased vegetative growth,¹³⁴ as was observed in the current study. Further, phospholipase A1 is known to play a pivotal role in Ca²⁺ signaling in eukaryotes and is considered to be multifunctional because of its ability to interact and regulate the activity of a number of other proteins. Ca²⁺ is a universal signal and the dynamic changes in its release and entry trigger a plethora of cellular responses. Central to this schema are members of the phospholipase superfamily, which relay information from the activated receptor to downstream signal cascades by production of second-messenger molecules. Recent studies reveal that, in addition to its enzymatic activity, phospholipase regulates Ca²⁺ entry via the formation of an intermolecular lipid-binding domain with canonical transient receptor potential ion channels. This complex, in turn, controls trafficking. Thus, ion channels are functionally linked to both lipase-dependent and -independent activities of phospholipase.¹³⁷ Similarly, annexins, another multigene and multifunctional family of soluble proteins, play a crucial role in Ca²⁺-dependent and Ca²⁺-independent signaling responses. They are ubiquitous soluble proteins capable of binding endomembranes and plasma membranes and have been implicated to function in diverse cellular processes, including signal transduction, channel activity, vesicular transport and fusion, and DNA replication.^{138,139} Annexins are found to bind with actin in a Ca²⁺- and pH-dependent manner. Annexin34 (StC-655), one of the calcium binding proteins, is also involved in the homeostatic regulation of intracellular Ca²⁺ in the presence of phospholipids and thus control structural changes to the membrane cytoskeleton, regulation of Ca²⁺ channels, and stimulation of G-protein signaling. The up-regulation of annexin (StC-655) in AmA1 potato tubers might cause an overall hypermorphic phenotype along with actin, as it maintains the cell structural microenvironment for survival, turgor, and wall pressure.¹³⁸

Actin (StC- 592 and 625) depolymerization resulted in alteration in cell division in a very slow manner but reduces the cell elongation dramatically,¹⁴⁰ which might lead to a burst of cell division. H⁺-ATPase stimulated by auxin is a key player in turgor regulation, since it provides the driving force for ion uptake, followed by water influx through osmosis. 14-3-3 protein is emerging as a regulator of many key ion transporters such as the plasma membrane H⁺-ATPase and outward rectifying K⁺ channel and functions as osmotic motor in plant cell.¹¹⁰ Therefore, increased expression of 14-3-3 (StC-11, 602, 604, 605, and 606) in AmA1 potato presumably translocates excess sucrose and simultaneously maintains

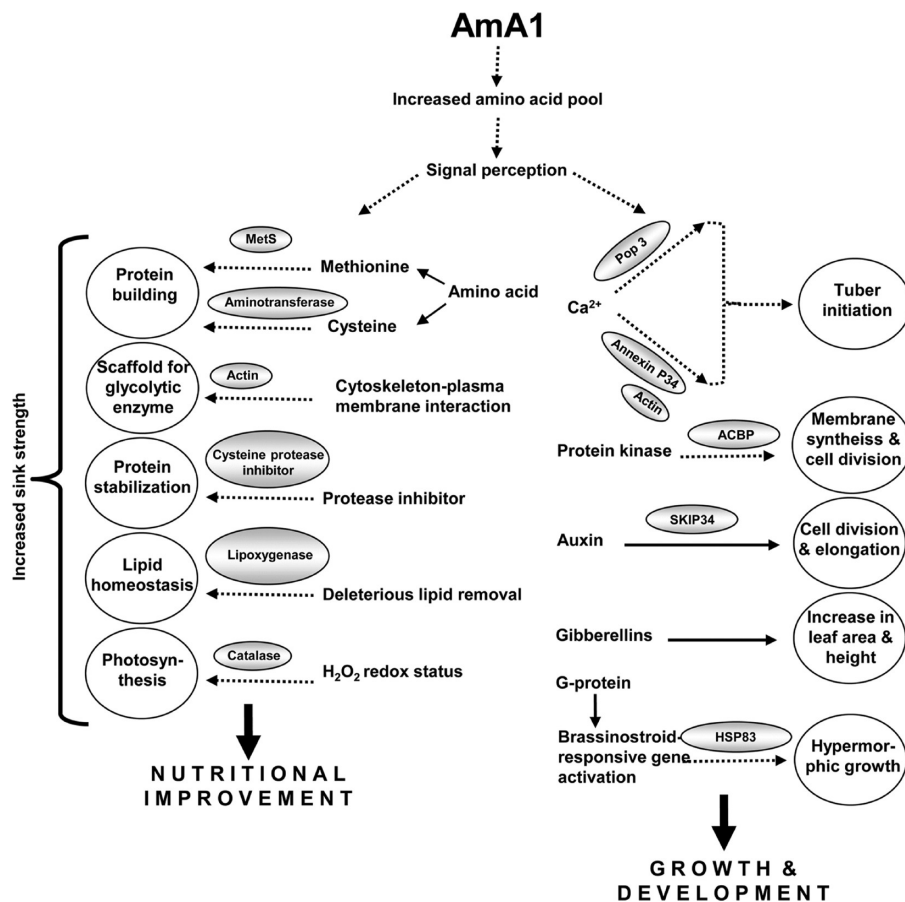


Figure 10. Model elucidating the summary of cellular pathways involving AmA1-regulated proteins. Proteins indicated in the colored boxes and the possible pathways depicted by dashed lines were identified in this study.

osmotic relation in different plant organs. Together with FFA and G protein, 14-3-3 in turn stimulates SKIP 34 (StC-697), a key regulator in auxin signaling. Auxin interacts with various F-box proteins and affects plant morphology, floral development, root and hypocotyls growth.¹⁴¹ Aconitase (StC-670), which documented to prevent senescence and promote growth,¹⁴² showed a mixed expression followed by a slight increase in the later stage of development in AmA1 tuber.

Plant lipoxygenases (LOXs: StC-161, 545 and 721), another lipid degrading enzyme belonging to functionally diverse class of dioxygenases, have been implicated in various physiological processes, such as growth, senescence, and stress-related responses.¹⁴³ Several reports have shown the involvement of LOXs in tuber growth and development in potato,^{143,144} during leaf development in soybean,¹⁴⁵ seed germination in cucumbers,¹⁴⁶ seedling development in *Arabidopsis*,¹⁴⁷ and fruit ripening in tomato.¹⁴⁸ The involvement of LOXs in tuber growth and development can be explained by the fact that LOX-derived products, such as jasmonic acid, methyl jasmonate, and tuberonic acid, have strong tuber inducing activity.¹⁴⁹ It is interesting to note individual lipoxygenase genes were shown to be induced by jasmonates and have a role in thigmomorphogenesis and gravitropism in wheat.¹²⁹ Thus higher expression of LOX in AmA1 tuber suggests its possible role in tuber growth and development. It might function by regulating the reorientation of microtubules that allows radial cell expansion leading to tuber enlargement.^{150,151}

It is becoming apparent that protein degradation is necessary not only for the removal of aberrant or damaged proteins but

also for altering the balance of proteins during development or stress adaptation. Proteasome subunit alpha type and subunit beta type-6 (StC-723 and 680), thought to be present in all eukaryotic cells,^{152,153} degrade specific proteins that have been targeted for proteolysis by ubiquitination. In animals and yeast, proteasomes also play a role in regulating the cell cycle and other developmental processes, via control of the levels of regulatory proteins, such as cyclins and transcription factors. In addition, it might also play a role in morphological evolution. A recent report showed that HSP80 along with sHSPs (StC-371, 377) interacts with the 26S proteasome¹⁵⁴ and plays a principal role in its assembly and maintenance. One of the HSP90 family of proteins, HSP83 (StC-593), is developmentally regulated and responds to phytohormones.¹⁵⁵ It is reported that HSP90 family proteins, being molecular chaperones, manage protein folding and play a key role in signal-transduction networks, protein degradation, and protein trafficking in *Arabidopsis*.¹⁵⁶ Moreover, HSP83 acts as a 'buffer' to sustain the functions of mutated proteins that participate in the signaling pathways of development and morphogenesis. HSP83 is unique in its function. Not only is it required for maturation and maintenance of most proteins *in vivo*, but also most of its cellular targets are signal transducer and developmental and cell cycle regulators. HSP83 keep these unstable signaling protein poised for activation until stabilized by conformational changes. It also might buffer against developmental noise caused by random microenvironmental effect with no genetic basis. Simultaneously, it also acts as genetic capacitor for developmentally regulated genes. In the present study its up-

regulation in AmA1 potato clearly suggests genetic variation might be expressed to a greater extent affecting many developmental pathways causing morphological evolution, supporting the basis of natural selection for the AmA1-plant. The higher expression of HSP83 might contribute in the development of AmA1-plant as they might bind to a range of target proteins, such as steroid hormone receptors, signaling kinases, and a variety of transcription factors^{157,158} that are crucial for regulating growth and development. It is known that HSP83 interacts with the 26S proteasome and plays a principal role in its assembly and maintenance.¹⁵⁹

Acetyl Co-A binding protein (ACBP; StC-648) is a highly conserved cytosolic lipid-binding protein that binds long-chain acyl-CoA esters with high affinity and is expressed in developing embryo, flowers, cotyledons, and other tissues.¹⁶⁰ It has been suggested that ACBP may mediate the translocation of acyl-CoA from the source organ to the sink organ during membrane synthesis.¹⁶¹ The increase in ACBP expression in AmA1 tubers could be a requirement for increased acyl-CoA transport during membrane synthesis that might facilitate high rate of cell division. However, under stress FDH (StC-584 and 676) accumulate in leaves, and leaf mitochondria acquire the ability to use formate as a respiratory substrate.¹⁶² FDH expression was found to be increased in AmA1 tubers, which might reflect an alternative source of photorespiration during the growth of tuber.

CONCLUSION

In summary, we describe the application of molecular phenotyping to gain new insights into the storage protein function by performing large-scale proteomic analysis on AmA1-expressed potato tubers and analyzing the data set using a network approach. Data presented here both confirmed our earlier hypothesis and, more importantly, identified novel AmA1-regulated proteins and pathways. We propose a working model on AmA1-mediated regulation involving novel pathways and cross-talk between various cellular pathways in nutrient-dependent storage accumulation and organ development. The results obtained from the present investigation have been illustrated to summarize the molecular events in AmA1 potato (Figure 10). The increased sink strength due to the combinatorial effect of the ARPs discussed might contribute to the concomitant increase in tuber growth. Thus, our observation suggests that a nutrient-responsive metabolic pathway overlaps with that of growth and development not only at the level of protein expression but also functionally in AmA1 potato. The exact function of these proteins can be predicted only after extensive studies and might be interesting for future investigations. Such a system based proteomic study further extends our understanding about the storage protein-mediated cellular physiology and may ultimately lead to development of more efficient strategies for plant biotechnology. In this study, we have identified a number of proteins previously unknown as AmA1 targets. Some of them are involved in the control of cellular gene expression and phenotype and are probable mediators of protein accumulation and the growth action of AmA1. The resulting lists of proteins were analyzed for the presence of members of known molecular networks. These networks showed substantial overlap and connectivity with the biological "theme" of nutrient accumulation and cell growth. Furthermore, integration of proteomics data with metabolomic analysis allowed a comprehensive system-level understanding of storage protein mediated

regulation in plant. It is expected that as functional genomics data in potato accumulate, the significance of the isoforms would unravel the post-translational regulation toward the AmA1-regulated protein network. It would be interesting to determine the possible interaction of these differentially expressed proteins among themselves as well as with AmA1, which could contribute to the increased protein and amino acid contents and tuber yield in the transgenics. Our study illustrated for the first time the molecular mechanism of a seed storage protein-regulated functional interaction structure of a cellular network involved in increased protein synthesis, storage reserve accumulation, and organ growth leading to enhanced productivity.

ASSOCIATED CONTENT

Supporting Information

Additional figures, tables, and data as described in the text. This material is available free of charge via the Internet at <http://pubs.acs.org>.

AUTHOR INFORMATION

Corresponding Authors

*Tel: 91-11-26735178. Fax: 91-11-26741658. E-mail: nchakraborty@hotmail.com.

*Tel: 91-11-26735186. Fax: 91-11-26741658. E-mail: subhrac@hotmail.com.

Author Contributions

[§]These authors contributed equally to this work.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

We thank Dr. Evert Jacobsen for providing the plasmid pPGB1. This work was supported by grants (BT/PR/11676/PBD/16/831/2008) from the Department of Biotechnology (DBT), Government of India and the National Institute of Plant Genome Research, New Delhi, India. L.A. and K.N. are the recipients of predoctoral fellowship from the Council of Scientific and Industrial research (CSIR), Govt. of India. Authors also thank Mr. Jasbeer Singh for illustrations and graphical representations in the manuscript.

REFERENCES

- (1) Wang, R.; Guegler, K.; LaBrie, S. T.; Crawford, N. M. Genomic analysis of a nutrient response in *Arabidopsis* reveals diverse expression patterns and novel metabolic and potential regulatory genes that are induced by nitrate. *Plant Cell* **2000**, *12*, 1419–1510.
- (2) Stitt, M.; Müller, C.; Matt, P.; Gibon, Y.; Carillo, P.; Morcuende, R.; Scheible, W. R.; Krapp, A. Steps towards an integrated view of nitrogen metabolism. *J. Exp. Bot.* **2002**, *53*, 959–970.
- (3) Wang, R.; Okamoto, M.; Xing, X.; Crawford, N. M. Microarray analysis of the nitrate response in *Arabidopsis* roots and shoots reveals over 1,000 rapidly responding genes and new linkages to glucose, trehalose-6-phosphate, iron, and sulfate metabolism. *Plant Physiol.* **2003**, *132*, 556–567.
- (4) Palenchar, P.; Kouranov, A.; Lejay, L.; Coruzzi, G. Genome-wide patterns of carbon and nitrogen regulation of gene expression validate the combined carbon and nitrogen (CN)-signaling hypothesis in plants. *Genome Biol.* **2004**, *5*, R91.
- (5) Schnepfer, L.; Duvel, K.; Broach, J. R. Sense and sensibility: nutritional response and signal integration in yeast. *Curr. Opin. Microbiol.* **2004**, *7*, 624–630.

- (6) Forchhammer, K. Global carbon/nitrogen control by PII signal transduction in cyanobacteria: from signals to targets. *FEMS Microbiol. Rev.* **2004**, *28*, 319–333.
- (7) Corbesier, L.; Lejeune, P.; Bernier, G. The role of carbohydrates in the induction of flowering in *Arabidopsis thaliana*: comparison between the wild type and a starchless mutant. *Planta* **1998**, *206*, 131–137.
- (8) Lynch, J. Root architecture and plant productivity. *Plant Physiol.* **1995**, *109*, 7–13.
- (9) Forde, B. G. Local and long-range signaling pathways regulating plant responses to nitrate. *Annu. Rev. Plant Biol.* **2002**, *53*, 203–224.
- (10) Koch, K. E. Carbohydrate-modulated gene expression in plants. *Annu. Rev. Plant Physiol. Plant Mol. Biol.* **1996**, *47*, 509–540.
- (11) Stitt, M. Nitrate regulation of metabolism and growth. *Curr. Opin. Plant Biol.* **1999**, *2*, 178–186.
- (12) Schmidt, W.; Barbazuk, B.; Sandford, M.; May, G.; Song, Z.; Zhou, W.; Basil, J.; Herman, E. M. Silencing of soybean seed storage proteins results in a rebalanced protein composition preserving seed protein content without major collateral changes in the metabolome and transcriptome. *Plant Physiol.* **2011**, *156*, 330–345.
- (13) Mandal, S.; Mandal, R. K. Seed storage proteins and approaches for improvement of their nutritional quality by genetic engineering. *Curr. Sci.* **2000**, *79*, 576–589.
- (14) Bechtel, D. B.; Gaines, R. L.; Pomeranz, Y. Changes in protein bodies of maturing wheat. *Cereal Foods World* **1980**, *25*, 516–517.
- (15) Raina, A.; Datta, A. Molecular cloning of a gene encoding a seed-specific protein with nutritionally balanced amino acid composition from *Amaranthus*. *Proc. Natl. Acad. Sci. U.S.A.* **1992**, *89*, 11774–11778.
- (16) Shewry, P. R.; Napier, J. A.; Tatham, A. S. Seed storage proteins: structures and biosynthesis. *Plant Cell* **1995**, *7*, 945–956.
- (17) Crofts, A. J.; Crofts, N.; Whitelegge, J. P.; Okita, T. W. Isolation and identification of cytoskeleton-associated prolamine mRNA binding proteins from developing rice seeds. *Planta* **2010**, *231*, 1261–1276.
- (18) Molvig, L.; Tabe, L. M.; Eggum, B. O.; Moore, A. E.; Craig, S.; Spencer, D.; Higgins, T. J. V. Enhanced methionine levels and increased nutritive value of seeds of transgenic lupins (*Lupinus angustifolius* L.) expressing a sunflower seed albumin gene. *Proc. Natl. Acad. Sci. U.S.A.* **1997**, *94*, 8393–8398.
- (19) Bellucci, M.; Lazzari, B.; Viotti, A.; Arcioni, S. Differential expression of a g-zein gene in *Medicago sativa*, *Lotus corniculatus* and *Nicotiana tabacum*. *Plant Sci.* **1997**, *127*, 161–169.
- (20) Chakraborty, S.; Chakraborty, N.; Datta, A. Increased nutritive value of transgenic potato by expressing a nonallergic seed albumin gene from *Amaranthus hypochondriacus*. *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97*, 3724–3729.
- (21) Tamas, C.; Kisgyorgy, B. N.; Rakszegi, M.; Wilkinson, M. D.; Yang, M.-S.; Lang, L.; Tamas, L.; Bedo, Z. Transgenic approach to improve wheat (*Triticum aestivum* L.) nutritional quality. *Plant Cell Rep.* **2009**, *28*, 1085–1094.
- (22) Chakraborty, S.; Chakraborty, N.; Agrawal, L.; Ghosh, S.; Narula, K.; Shekhar, S.; Naik, P. S.; Pande, P. C.; Chakraborty, S. K.; Datta, A. Next generation protein rich potato by expressing a seed protein gene *AmA1* as a result of proteome rebalancing in transgenic tuber. *Proc. Natl. Acad. Sci. U.S.A.* **2010**, *107*, 17533–17538.
- (23) Oszwald, M.; Balazs, G.; Polya, S.; Tömösközi, S.; Appels, R.; Bekes, F.; Tamas, L. Wheat storage proteins in transgenic rice endosperm. *J. Agric. Food Chem.* **2013**, *61*, 7606–7614.
- (24) Schmidt, M. A.; Barbazuk, W. B.; Sandford, M.; May, G.; Song, Z.; Zhou, W.; Nikolau, B. J.; Herman, E. M. Silencing of soybean seed storage proteins results in rebalanced protein composition preserving seed protein content without major collateral changes in the metabolome and transcriptome. *Plant Physiol.* **2011**, *156*, 330–345.
- (25) Burks, C. Molecular biology database list. *Nucleic Acids Res.* **1999**, *27*, 1–9.
- (26) Curreem, S. O. T.; Watt, R. M.; Lau, S. K. P.; Woo, P. C. Y. Two-dimensional gel electrophoresis in bacterial proteomics. *Protein Cell* **2012**, *3*, 346–363.
- (27) Hecker, M.; Antelmann, H.; Büttner, K.; Bernhardt, J. Gel-based proteomics of Gram-positive bacteria: A powerful tool to address physiological questions. *Proteomics* **2008**, *8*, 4958–4975.
- (28) McGregor, E.; Dunn, M. J. Proteomics of the heart: unraveling disease. *Circ. Res.* **2006**, *98*, 309–321.
- (29) Jungblut, P. R.; Zimny, A. U.; Zeindl-Eberhart, E.; Stulik, J.; Koupilova, K.; Pleibner, K. B.; Otto, A.; Müller, E. C.; Sokolowska-Kohler, W.; Grabber, G.; Stoffer, G. Proteomics in human disease: cancer, heart and infectious diseases. *Electrophoresis* **1999**, *20*, 2100–2110.
- (30) Steppan, C.; Bailey, S.; Bhat, S.; Brown, E.; Bannerjee, R.; Wright, C.; Patel, H.; Ahima, R.; Lazar, M. The hormone resistin links obesity to diabetes. *Nature* **2001**, *409*, 307–312.
- (31) Ruebelt, M. C.; Leimgruber, N. K.; Lipp, M.; Reynolds, T. L.; Nemeth, M. A.; Astwood, J. D.; Engel, K. H.; Jany, K. D. Application of two-dimensional gel electrophoresis to interrogate alterations in the proteome of genetically modified crops. 1. Assessing analytical validation. *J. Agric. Food Chem.* **2006**, *54*, 2154–2161.
- (32) Gallardo, K.; Job, C.; Groot, S. P. C.; Puype, M.; Demol, H.; Vandekerckhove, J.; Job, D. Proteomics analysis of *Arabidopsis* seed germination and priming. *Plant Physiol.* **2001**, *126*, 835–848.
- (33) 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*, 32–46.
- (34) Lehesranta, S. J.; Davies, H. V.; Shepherd, L. V. T.; Koistinen, K. M.; Massot, N.; Nunan, N.; McNicol, J. W.; Karenlampi, S. O. Proteomic analysis of the potato tuber life cycle. *Proteomics* **2006**, *6*, 6042–6052.
- (35) Agrawal, L.; Chakraborty, S.; Jaiswal, D.; Gupta, S.; Datta, A.; Chakraborty, N. Comparative proteomics of tuber induction, development and maturation reveal the complexity of tuberization process in potato (*Solanum tuberosum* L.). *J. Proteome Res.* **2008**, *7*, 3803–3817.
- (36) Bhushan, D.; Pandey, A.; Choudhary, M. K.; Datta, A.; Chakraborty, S.; Chakraborty, N. Comparative proteomic analysis of differentially expressed proteins in chickpea extracellular matrix during dehydration stress. *Mol. Cell. Proteomics* **2007**, *6*, 1868–1884.
- (37) Pandey, A.; Chakraborty, S.; Datta, A.; Chakraborty, N. Proteomics approach to identify dehydration responsive nuclear proteins from chickpea (*Cicer arietinum* L.). *Mol. Cell. Proteomics* **2008**, *7*, 88–107.
- (38) Choudhary, M. K.; Basu, D.; Datta, A.; Chakraborty, N.; Chakraborty, S. Dehydration-responsive nuclear proteome of rice (*Oryza sativa* L.) illustrates protein network, novel regulators of cellular adaptation, and evolutionary perspective. *Mol. Cell. Proteomics* **2009**, *8*, 1579–1598.
- (39) Herbig, A.; Giritich, A.; Horstmann, C.; Becker, R.; Blazer, H. J.; Baumlein, H.; Stephan, U. W. Iron and copper nutrition dependent changes in protein expression in a tomato wild type and the nicotinamine-free mutant chloronerva. *Plant Physiol.* **1996**, *111*, 533–540.
- (40) Chakraborty, N.; Ghosh, R.; Ghosh, S.; Narula, K.; Tayal, R.; Datta, A.; Chakraborty, S. Reduction of oxalate levels in tomato fruit and consequent metabolic remodeling following overexpression of a fungal oxalate decarboxylase. *Plant Physiol.* **2013**, *162*, 364–378.
- (41) Lehesranta, S. J.; Davies, H. V.; Shepherd, L. V. T.; Nunan, N.; McNicol, J. W.; Aureola, S.; Koistinen, K. M.; Suomalainen, S.; Kokko, K. I.; Karenlampi, S. O. Comparison of tuber proteomes of potato varieties, landraces and genetically modified lines. *Plant Physiol.* **2005**, *138*, 1690–1699.
- (42) Zolla, L.; Rinalducci, S.; Antonioli, P.; Righetti, P. G. Proteomics as a complementary tool for identifying unintended side effects occurring in transgenic maize seeds as a result of genetic modifications. *J. Proteome Res.* **2008**, *7*, 1850–1861.
- (43) Scossa, F.; Chingcuanco, D. L.; Anderson, O. D.; Vensel, W. H.; Lafiandra, D.; D'Ovidio, R.; Masci, S. Comparative proteomic and transcriptional profiling of a bread wheat cultivar and its derived

transgenic line overexpressing a low molecular weight glutenin subunit gene in the endosperm. *Proteomics* **2008**, *8*, 2948–2966.

(44) Sobhanian, H.; Motamed, N.; Jazii, F. R.; Nakamura, T.; Komatsu, S. Salt stress induced differential proteome and metabolome response in the shoots of *Aeluropus lagopoides* (Poaceae), a halophyte C4 plant. *J. Proteome Res.* **2010**, *9*, 2882–2897.

(45) Lu, J.; Fernandes, E. A.; Cano, A. E. P.; Vinitwatanakhun, J.; Boeren, S.; van Hooijdonk, T.; van Kneegsel, A.; Vervoort, J.; Hettinga, K. A. Changes in milk proteome and metabolome associated with dry period length, energy balance, and lactation stage in postparturient dairy cows. *J. Proteome Res.* **2013**, *12*, 3288–3296.

(46) Mutwil, M.; Usadel, B.; Schuttee, M.; Loraine, A.; Ebenhoh, O.; Persson, S. Assembly of an interactive correlation network for the Arabidopsis genome using a novel heuristic clustering algorithm. *Plant Physiol.* **2010**, *152*, 29–43.

(47) Albert, R. Scale-free networks in cell biology. *J. Cell Sci.* **2005**, *118*, 4947–4957.

(48) Gupta, R. B.; Masci, S.; Lafiandra, D.; Bariana, H. S.; MacRitchie, F. Accumulation of protein subunits and their polymers in developing grains of hexaploid wheats. *J. Exp. Bot.* **1996**, *47*, 1377–1385.

(49) Chakraborty, S.; Sarmah, B.; Chakraborty, N.; Datta, A. Premature termination of RNA polymerase II mediated transcription of a seed protein gene in *Schizosaccharomyces pombe*. *Nucleic Acids. Res.* **2002**, *30*, 2940–2949.

(50) Sarmah, B.; Chakraborty, N.; Chakraborty, S.; Datta, A. Plant pre-mRNA splicing in fission yeast, *Schizosaccharomyces pombe*. *Biochem. Biophys. Res. Commun.* **2002**, *293*, 1209–1216.

(51) Dussert, S.; Davey, M. W.; Laffargue, A.; Doubeau, S.; Swennen, R.; Etienne, H. Oxidative stress, phospholipid loss and lipid hydrolysis during drying and storage of intermediate seeds. *Physiol. Plant.* **2006**, *127*, 192–204.

(52) Pernas, M.; Monge, S.; Salcedo, B. Biotic and abiotic stress can induce cystatin expression in chestnut. *FEBS Lett.* **2000**, *467*, 206–210.

(53) Grudkowska, M.; Zagdańska, B. Multifunctional role of plant cysteine proteinases. *Acta Biochim. Pol.* **2004**, *51*, 609–624.

(54) Chen, Z. Y.; Brown, R. L.; Cleveland, T. E. Evidence for an association in corn between stress tolerance and resistance to *Aspergillus flavus* infection and aflatoxin contamination. *Afr. J. Biotechnol.* **2004**, *3*, 693–699.

(55) Hurkman, W. J.; Tanaka, C. K. Solubilization of plant membrane proteins for analysis by two-dimensional gel electrophoresis. *Plant Physiol.* **1986**, *81*, 802–806.

(56) Liu, H.; Sadygov, R. G.; Yates, J. R. A model for random sampling and estimation of relative protein abundance in shotgun proteomics. *Anal. Chem.* **2004**, *76*, 4193–4201.

(57) Ishihama, Y.; Oda, Y.; Tabata, T.; Sato, T.; Nagasu, T.; Rappsilber, J.; Mann, M. Exponentially modified protein abundance index (emPAI) for estimation of absolute protein amount in proteomics by the number of sequenced peptides per protein. *Mol. Cell. Proteomics* **2005**, *4*, 1265–1272.

(58) Saeed, A. I.; Sharov, V.; White, J.; Li, J.; Liang, W.; Bhagabati, N.; Braisted, J.; Klapa, M.; Currier, T.; Thiagarajan, M. TM 4: a free, open-source system for microarray data management and analysis. *Biotechniques* **2003**, *34*, 374–378.

(59) Herrero, M. Ovary signals for directional pollen tube growth. *Sex. Plant Reprod.* **2001**, *14*, 3–7.

(60) de la Fuente, A.; Bing, N.; Hoeschele, I.; Mendes, P. Discovery of meaningful associations in genomic data using partial correlation coefficients. *Bioinformatics* **2004**, *20*, 3565–3574.

(61) Shannon, P.; Markiel, A.; Ozier, O.; Baliga, N. S.; Wang, J. T.; Ramage, D.; Amin, N.; Schwikowski, B.; Ideker, T. Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res.* **2003**, *11*, 2498–2504.

(62) Roessner, U.; Wagner, C.; Kopka, J.; Trethewey, R. N.; Willmitzer, L. Technical advance: simultaneous analysis of metabolites in potato tuber by gas chromatography-mass spectrometry. *Plant J.* **2000**, *23*, 131–42.

(63) Roessner, U.; Willmitzer, L.; Fernie, A. R. High-resolution metabolic phenotyping of genetically and environmentally diverse potato tuber systems. Identification of phenocopies. *Plant Physiol.* **2001**, *127*, 749–764.

(64) Jefferson, R. A.; Kavanagh, T. A.; Bevan, M. W. GUS fusions: betaglucuronidase as a sensitive and versatile gene fusion marker in higher plants. *EMBO J.* **1987**, *61*, 3901–3907.

(65) Chaiyavit, S.; Thongboonkerd, V. Changes in mitochondrial proteome of renal tubular cells induced by calcium oxalate monohydrate crystal adhesion and internalization are related to mitochondrial dysfunction. *J. Proteome Res.* **2012**, *11*, 3269–3280.

(66) Sanchez-Bel, P.; Egea, I.; Sanchez-Ballesta, M. T.; Sevillano, L.; Del Carmen Bolarin, M.; Flores, F. B. Proteome changes in tomato fruits prior to visible symptoms of chilling injury are linked to defensive mechanisms, uncoupling of photosynthetic processes and protein degradation machinery. *Plant Cell Physiol.* **2012**, *53*, 470–484.

(67) Sweetlove, L. J.; Kossmann, J.; Riesmeier, J. W.; Trethewey, R. N.; Hill, S. A. The control of source to sink carbon flux during tuber development in potato. *Plant J.* **1998**, *15*, 697–706.

(68) Sghaier-Hammami, B.; Drira, N.; Jorrin-Novo, J. V. Comparative 2-DE proteomic analysis of date palm (*Phoenix dactylifera* L.) somatic and zygotic embryos. *J. Proteomics* **2009**, *73*, 161–177.

(69) Yang, Y.; Thannhauser, T. W.; Li, L.; Zhang, S. Development of an integrated approach for evaluation of 2-D gel image analysis: impact of multiple proteins in single spots on comparative proteomics in conventional 2-D gel/MALDI workflow. *Electrophoresis* **2007**, *28*, 2080–2094.

(70) 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*, 3551–3567.

(71) Zhang, L. J.; Wang, X. E.; Peng, X.; Wei, Y. J.; Cao, R.; Liu, Z.; Xiong, J. X.; Yin, X. F.; Ping, C.; Liang, S. Proteomic analysis of low-abundant integral plasma membrane proteins based on gels. *Cell. Mol. Life Sci.* **2006**, *63*, 1790–1804.

(72) Cilia, M.; Tamborindeguy, C.; Rolland, M.; Howe, K.; Thannhauser, T. W.; Gray, S. Tangible benefits of the aphid *Acyrtosiphon pisum* genome sequencing for aphid proteomics: enhancements in protein identification and data validation for homology-based proteomics. *J. Insect Physiol.* **2011**, *57*, 179–190.

(73) Taylor, N. L.; Heazlewood, J. L.; Millar, A. H. The *Arabidopsis thaliana* 2-D gel mitochondrial proteome: refining the value of reference maps for assessing protein abundance, contaminants and post-translational modifications. *Proteomics* **2011**, *11*, 1720–1733.

(74) Poetsch, A.; Wolters, D. Bacterial membrane proteomics. *Proteomics* **2008**, *8*, 4100–4122.

(75) Yoon, J. H.; Yea, K.; Kim, J.; Choi, Y. S.; Park, S.; Lee, H.; Lee, C. S.; Suh, P.-G.; Ryu, S. H. Comparative proteomic analysis of the insulin-induced L6 myotube secretome. *Proteomics* **2009**, *9*, 51–60.

(76) Chevalier, F. Highlights on the capacities of "Gel-based" proteomics. *Proteome Sci.* **2010**, *8*, 23.

(77) Jin, M.; Diaz, P. T.; Bourgeois, T.; Eng, C.; Marsh, C. B.; Wu, H. M. Two-dimensional gel proteome reference map of blood monocytes. *Proteome Sci.* **2006**, *4*, 16–23.

(78) Beckner, M. E.; Chen, X.; An, J.; Day, B. W.; Pollack, I. F. Proteomic characterization of harvested pseudopodia with differential gel electrophoresis and specific antibodies. *Lab. Invest.* **2005**, *85*, 316–327.

(79) Alm, R.; Johansson, P.; Hjerno, K.; Emanuelsson, C.; Ringner, M.; Hakkinen, J. Detection and identification of protein isoforms using cluster analysis of MALDI-MS mass spectra. *J. Proteome Res.* **2006**, *5*, 785–792.

(80) Yan, S. P.; Zhang, Q. Y.; Tang, Z. C.; Su, W. A.; Sun, W. N. Comparative proteomic analysis provides new insights into chilling stress responses in rice. *Mol. Cell. Proteomics* **2006**, *5*, 484–496.

(81) Ahsan, N.; Lee, D. G.; Lee, S. H.; Kang, K. Y.; Lee, J. J.; Kim, P. J.; Yoon, H. S.; Kim, J. S.; Lee, B. H. Excess copper induced physiological and proteomic changes in germinating rice seeds. *Chemosphere* **2007**, *67*, 1182–119.

- (82) Moreira, J. M. A.; Shen, T.; Ohlsson, G.; Gromov, P.; Gromova, I.; Celis, J. E. A combined proteome and ultrastructural localization analysis of 14-3-3 proteins in transformed human amnion (AMA) cells: definition of a framework to study isoform-specific differences. *Mol. Cell. Proteomics* **2008**, *7*, 1225–1240.
- (83) Dai, S. J.; Chen, T. T.; Chong, K.; Xue, Y. B.; Liu, S. Q.; Wang, T. Proteomic identification of differentially expressed proteins associated with pollen germination and tube growth reveals characteristics of germinated *Oryza sativa* pollen. *Mol. Cell. Proteomics* **2007**, *6*, 207–230.
- (84) Herbers, H.; Sonnewald, U. Molecular determinants of sink strength. *Curr. Opin. Plant Biol.* **1998**, *1*, 207–216.
- (85) Weber, H.; Sreenivasulu, N.; Weschke, W. Molecular Physiology of Seed Maturation and Seed Storage Protein Biosynthesis. *Plant Dev. Biol.* **2003**, *2*, 83–104.
- (86) Brazhnik, P.; de la Fuente, A.; Mendes, P. Gene networks: how to put the function in genomics. *Trends Biotechnol.* **2002**, *20*, 467–472.
- (87) de la Fuente, A.; Mendes, P. Quantifying gene networks with regulatory strengths. *Mol. Biol. Rep.* **2002**, *29*, 73–77.
- (88) Singh, B. K. Biosynthesis of Valine, Leucine and Isoleucine. In *Plant Amino Acids: Biochemistry and Biotechnology*; Singh, B. K., Ed.; Marcel Dekker: New York, NY, 1999; pp 227–247.
- (89) Knill, T.; Schuster, J.; Reichelt, M.; Gershenzon, J.; Binder, S. Arabidopsis branched-chain aminotransferase 3 functions in both amino acid and glucosinolate biosynthesis. *Plant Physiol.* **2008**, *146*, 1028–1039.
- (90) Giovanelli, J.; Mudd, S. H.; Datko, A. H. Quantitative analysis of pathways of methionine metabolism and their regulation in *Lemna*. *Plant Physiol.* **1985**, *78*, 555–560.
- (91) Larrainzar, E.; Wienkoop, S.; Weckwerth, W.; Ladrera, R.; Arrese-Igor, C.; González, E. M. *Medicago truncatula* root nodule proteome analysis reveals differential plant and bacteroid responses to drought stress. *Plant Physiol.* **2007**, *144*, 1495–1507.
- (92) Hesse, H.; Kreft, O.; Maimann, S.; Zeh, M.; Hoefgen, R. Current understanding of the regulation of methionine biosynthesis in plants. *J. Exp. Bot.* **2004**, *55*, 1799–1808.
- (93) Do, A.; Yang, S. Methionine metabolism in apple tissue: implication of S-adenosylmethionine as an intermediate in the conversion of methionine to C2H4. *Plant Physiol.* **1977**, *60*, 892–896.
- (94) Li, L.; Modolo, L. V.; Escamilla-Trevino, L. L.; Achnine, L.; Dixon, R. A.; Wang, X. Crystal structure of *Medicago truncatula* UGT85H2 - insights into the structural basis of a multifunctional (Iso)flavonoid glycosyltransferase. *J. Mol. Biol.* **2007**, *370*, 951–963.
- (95) Levican, G.; Katz, A.; Valenzuela, P.; Soll, D.; Orellana, O. A tRNA^{Glu} that uncouples protein and tetrapyrrole biosynthesis. *FEBS Lett.* **2005**, *579*, 6383–6387.
- (96) Oberschall, A.; Deák, M.; Török, K.; Sass, L.; Vass, I.; Kovács, I.; Fehér, A.; Dudits, D.; Horváth, G. V. A novel aldose/aldehyde reductase protects a transgenic plants against lipid peroxidation under chemical and drought stresses. *Plant J.* **2000**, *24*, 437–446.
- (97) Zrenner, R.; Salanoubat, M.; Willmitzer, L.; Sonnewald, U. Evidence for the crucial role of sucrose synthase for sink strength using transgenic potato plants (*Solanum tuberosum* L.). *Plant J.* **1995**, *7*, 97–107.
- (98) Meagher, R. B.; McKinney, E. C.; Kandasamy, M. K. Isovariant dynamics expand and buffer the responses of complex systems: the diverse plant actin gene family. *Plant Cell* **1999**, *11*, 995–1005.
- (99) Fisher, D. I.; Safrany, S. T.; Strike, P.; McLennan, A. G.; Cartwright, J. L. Nudix hydrolases that degrade dinucleoside and diphosphoinositol polyphosphates also have 5-phosphoribosyl 1-pyrophosphate (PRPP) pyrophosphatase activity that generates the glycolytic activator ribose 1,5-bisphosphate. *J. Biol. Chem.* **2002**, *277*, 47313–47317.
- (100) Sangster, T. A.; Queitsch, C. The HSP90 chaperone complex, an emerging force in plant development and phenotypic plasticity. *Curr. Opin. Plant Biol.* **2005**, *8*, 86–92.
- (101) Balasubramanian, R.; Karve, A.; Kandasamy, M.; Meagher, R. B.; Moore, B. D. A role for F-actin in hexokinase-mediated glucose signaling. *Plant Physiol.* **2007**, *145*, 1423–1434.
- (102) DeRocher, A.; Vierling, E. Developmental control of small heat shock protein expression during pea seed maturation. *Plant J.* **1994**, *5*, 93–102.
- (103) O'Kelly, I.; Butler, M. H.; Zilberberg, N.; Goldstein, S. A. Forward transport. 14-3-3 binding overcomes retention in endoplasmic reticulum by dibasic signals. *Cell* **2002**, *111*, 577–588.
- (104) Muslin, A. J.; Tanner, J. W.; Allen, P. M.; Shaw, A. S. Interaction of 14-3-3 with signaling proteins is mediated by the recognition of phosphoserine. *Cell* **1996**, *84*, 889–897.
- (105) Luque, C. M.; Correas, I. A constitutive region is responsible for nuclear targeting of 4.1R: modulation by alternative sequences results in differential intracellular localization. *J. Cell Sci* **2000**, *113*, 2485–2495.
- (106) Lockhart, D. J.; Winzler, E. A. Genomics, gene expression and DNA arrays. *Nature* **2000**, *405*, 827–836.
- (107) Sanchez-Quiles, V.; Mora, M. L.; Segura, V.; Greco, A.; Epstein, A. L.; Foschini, M. G.; Dayon, L.; Sanchez, J. C.; Prieto, J.; Corrales, F. J.; Santamaria, E. HSV-1 Cgal infection promotes quaking RNA binding protein production and induces nuclear-cytoplasmic shuttling of quaking I-5 isoform in human hepatoma cells. *Mol. Cell. Proteomics* **2011**, *10*, 1–14.
- (108) Paul, A.-L.; Sehnke, P. C.; Ferl, R. J. Isoform-specific subcellular localization among 14-3-3 proteins in *Arabidopsis* seems to be driven by client interactions. *Mol. Biol. Cell* **2005**, *16*, 1735–1743.
- (109) Philip, N.; Acevedo, S. F.; Skoulakis, E. M. Conditional rescue of olfactory learning and memory defects in mutants of the 14-3-3zeta gene leonardo. *J. Neurosci.* **2001**, *21*, 8417–8425.
- (110) Muslin, A. J.; Xing, H. 14-3-3 proteins: regulation of subcellular localization by molecular interference. *Cell. Signaling* **2000**, *12*, 703–709.
- (111) Lu, G.; DeLisle, A. J.; De Vetten, N. C.; Ferl, R. J. Brain proteins in plants: an *Arabidopsis* homolog to neurotransmitter pathway activators is part of a DNA binding complex. *Proc. Natl. Acad. Sci. U.S.A.* **1992**, *89*, 11490–11494.
- (112) Sehnke, P. C.; DeLille, J. M.; Ferl, R. J. Consummating signal transduction: the role of 14-3-3 proteins in the completion of signal-induced transitions in protein activity. *Plant Cell* **2002**, *14*, S339–S354.
- (113) de Boer, A. H. Plant 14-3-3 proteins assist ion channels and pumps. *Biochem. Soc. Trans.* **2002**, *4*, 416–420.
- (114) Bellincampia, D.; Camardellab, L.; Delcour, J. A.; Desseauxd, V.; D'Ovidioe, R.; Durandf, A.; Elliotf, G.; Gebruerse, K.; Giovaneg, A.; Jugef, N.; Sørenseni, J. F.; Svenssonj, B.; Vairo, D. Potential physiological role of plant glycosidase inhibitors. *Biochim. Biophys. Acta* **2004**, *1696*, 265–274.
- (115) Shewry, P. R. Enzyme Inhibitors of Seed: Types and Properties. In *Seed Proteins*; Shewry, P. R., Casey, R., Eds.; Kluwer Academic Publishers: Dordrecht, 1999; pp 587–615.
- (116) Glacinski, H.; Hibges, A.; Salmini, F.; Gebhardt, C. Members of the Kunitz type inhibitor gene family of potato inhibit soluble tuber invertase *in vitro*. *Potato Res.* **2003**, *15*, 213–218.
- (117) Schuler, T. H.; Poppy, G. M.; Kerry, B. R.; Denholm, I. Insect-resistant transgenic plants. *Trends Biotechnol.* **1998**, *16*, 168–175.
- (118) Freedman, R. B. Protein disulfide isomerase: multiple roles in the modification of nascent secretory proteins. *Cell* **1989**, *57*, 1069–1072.
- (119) Shimoni, Y.; Zhu, X.; Levanony, H.; Segal, G.; Galili, G. Purification, characterization and intracellular localization of glycosylated protein disulfide isomerase from wheat grains. *Plant Physiol.* **1995**, *108*, 327–335.
- (120) Cai, H.; Wan, C.; Tsou, C. Chaperone-like activity of protein disulfide isomerase in the refolding of a protein with no disulfide bonds. *J. Biol. Chem.* **1994**, *269*, 24550–24552.
- (121) Novak, P.; Dev, I. K. Degradation of a Signal Peptide by Protease IV and Oligopeptidase. *J. Bacteriol.* **1988**, *170*, 5067–5075.
- (122) Conlin, C. A.; Vimr, E. R.; Miller, C. G. Oligopeptidase is required for normal phage p22 development. *J. Bacteriol.* **1992**, *174*, 5869–5880.

- (123) Mc Cool, S.; Pierotti, A. R. Expression of the Thimet Oligopeptidase gene is regulated by positively and negatively acting elements. *DNA Cell Biol.* **2000**, *19*, 729–738.
- (124) Sonnewald, U.; Schaeven, A. V.; Willmitzer, L. Expression of mutant patatin protein in transgenic tobacco plants: role of glycans and intracellular location. *Plant Cell* **1999**, *2*, 345–355.
- (125) Pots, A. M.; Gruppen, H.; Hessing, M.; van Boekel, M. A. J. S.; Voragen, A. G. J. Isolation and characterization of patatin isoforms. *J. Agric. Food Chem.* **1999**, *47*, 4587–4592.
- (126) Weeda, S. M.; Kumar, G. N. M.; Knowles, N. R. Developmentally linked changes in proteases and protease inhibitors suggest a role for potato multicystatin in regulating protein content of potato tubers. *Planta* **2009**, *230*, 73–84.
- (127) Stupar, R. M.; Beaubien, K. A.; Jin, W.; Song, J.; Lee, M. K.; Wu, C.; Zhang, H. B.; Han, B.; Jiang, J. Structural diversity and differential transcription of the patatin multicopy gene family during potato tuber development. *Genetics* **2006**, *172*, 1263–1275.
- (128) Ryu, S. B. Phospholipid-derived signaling mediated by phospholipase A in plants. *Trends Plant Sci.* **2004**, *9*, 229–235.
- (129) Mauch, F.; Kmecl, A.; Schaffrath, U.; Volrath, S.; Gorch, J.; Ward, E.; Ryals, J.; Dudler, R. Mechanosensitive expression of a lipoxygenase gene in wheat. *Plant Physiol.* **1997**, *114*, 1561–1566.
- (130) Kloosterman, B.; Vorst, O.; Hall, R. D.; Visser, R. G. F.; Bachem, C. W. Tuber on a chip: differential gene expression during potato tuber development. *Plant Biotechnol. J.* **2005**, *3*, 505–519.
- (131) Reeve, R. M.; Hautala, E.; Weaver, M. L. Anatomy and compositional variations within potatoes. I. Developmental histology of the tuber. *Am. Potato J.* **1969**, *46*, 361–373.
- (132) Peterson, R. L.; Barker, W. G.; Howarth, M. J. Development and Structure of Tubers. In *Potato Physiology*; Li, P. H., Ed.; Academic Press: London, 1985; pp 123–147.
- (133) Du, L.; Poovaiah, B. W. Ca²⁺/calmodulin is critical for brassinosteroid biosynthesis and plant growth. *Nature* **2005**, *437*, 741–745.
- (134) Zhang, L.; Lu, Y.-T. Calmodulin-binding protein kinases in plants. *Trends Plant Sci.* **2003**, *8*, 123–128.
- (135) Rietza, S.; Dermendjiev, G.; Oppermann, E.; Tafesseb, F. G.; Effendib, Y.; Holkb, A.; Parkera, J. E.; Teigec, M.; Schere, G. F. E. Roles of arabidopsis patatin-related phospholipases A in root development are related to auxin responses and phosphate deficiency. *Mol. Plant* **2010**, *3*, 524–538.
- (136) Neer, E. J. Heterotrimeric G proteins: organizers of transmembrane signals. *Cell* **1995**, *80*, 249–257.
- (137) Pappan, K.; Wang, X. Plant phospholipase D is an acidic phospholipase active at near-physiological Ca²⁺ concentrations. *Arch. Biochem. Biophys.* **1999**, *368*, 347–353.
- (138) Calvert, C. M.; Gant, S. J.; Bowles, D. J. Tomato annexins p34 and p35 bind to F-actin and display nucleotide phosphodiesterase activity inhibited by phospholipid binding. *Plant Cell* **1996**, *8*, 333–342.
- (139) Blanchoin, L.; Boujemaa-Paterski, R.; Henty, J. L.; Khurana, P.; Staiger, C. J. Actin dynamics in plant cells: a team effort from multiple proteins orchestrates this very fast-paced game. *Curr. Opin. Plant Biol.* **2010**, *13*, 714–723.
- (140) Hussey, P. J.; Ketelaar, T.; Deeks, M. J. Control of the actin cytoskeleton in plant cell growth. *Annu. Rev. Plant Biol.* **2006**, *57*, 109–125.
- (141) Hagen, G.; Guilfoyle, T. Auxin-responsive gene, promoters and regulatory factors. *Plant Mol. Biol.* **2002**, *49*, 373–385.
- (142) Luce, K.; Osiewacz, H. D. Increasing organismal healthspan by enhancing mitochondrial protein quality control. *Nat. Cell Biol.* **2009**, *11*, 852–871.
- (143) Kolomiets, M. V.; Hannapel, D. J.; Chen, H.; Tymeson, M.; Gladon, R. J. Lipoxygenase is involved in the control of potato tuber development. *Plant Cell* **2001**, *13*, 613–626.
- (144) Bachem, C. W. B.; van der Hoeven, R. S.; de Bruijn, S. M.; Vreugdenhil, D.; Zabeau, M.; Visser, R. G. F. Visualization of differential gene expression using a novel method of RNA finger-printing based on AFLP: Analysis of gene expression during potato tuber development. *Plant J.* **1996**, *9*, 745–753.
- (145) Saravitz, D. M.; Siedow, J. N. The lipoxygenase isozymes in soybean [*Glycine max* (L.) Merr.] leaves: Changes during leaf development, after wounding, and following reproductive sink removal. *Plant Physiol.* **1995**, *107*, 535–543.
- (146) Matsui, K.; Irie, M.; Kajiwara, T.; Hatanaka, A. Developmental changes in lipoxygenase activity in cotyledons of cucumber seedlings. *Plant Sci.* **1992**, *85*, 23–32.
- (147) Melan, M. A.; Nemhauser, J. L.; Peterman, T. K. Structure and sequence of the *Arabidopsis thaliana* lipoxygenase 1 gene. *Biochim. Biophys. Acta* **1994**, *1210*, 377–380.
- (148) Kausch, K. D.; Handa, A. K. Molecular cloning of a ripening-specific lipoxygenase and its expression during wild-type and mutant tomato fruit development. *Plant Physiol.* **1997**, *113*, 1041–1050.
- (149) Pelacho, A. M.; Mingo-Castel, A. M. Jasmonic acid induces tuberization of potato stolons cultured *in vitro*. *Plant Physiol.* **1991**, *97*, 1253–1255.
- (150) Matsuki, T.; Tazaki, H.; Fujimori, T.; Hogetsu, T. The influences of jasmonic acid methyl ester on microtubules in potato cells and formation of potato tubers. *Biosci. Biotechnol. Biochem.* **1992**, *56*, 1329–1330.
- (151) Fujino, K.; Koda, Y.; Kikuta, Y. Reorientation of cortical microtubules in the subapical region during tuberization in single-node stem segments of potato in culture. *Plant Cell Physiol.* **1995**, *36*, 891–895.
- (152) Palmer, A.; Rivett, A. J.; Thomson, S.; Hendil, K. B.; Butcher, G. W.; Fuertes, G.; Knecht, E. Subpopulations of proteasomes in rat liver nuclei, microsomes and cytosol. *Biochem. J.* **1996**, *316*, 401–407.
- (153) Bahrami, A. R.; Gray, J. E. Expression of a proteasome α -type subunit gene during tobacco development and senescence. *Plant Mol. Biol.* **1999**, *39*, 325–333.
- (154) Koning, A. J.; Rose, R.; Comai, L. Developmental expression of tomato heat-shock cognate protein 80. *Plant Physiol.* **1992**, *100*, 801–811.
- (155) Rutherford, S. L.; Lindquist, S. Hsp90 as a capacitor for morphological evolution. *Nature* **1998**, *396*, 336–342.
- (156) Wang, W.; Vinocur, B.; Shoseyov, O.; Altman, A. Role of plant heat-shock proteins and molecular chaperones in the abiotic stress response. *Trends Plant Sci.* **2004**, *9*, 244–252.
- (157) Young, J. C.; Moarefi, I.; Hartl, F. U. Hsp90: a specialized but essential protein-folding tool. *J. Cell Biol.* **2001**, *154*, 267.
- (158) Neckers, L. Hsp90 inhibitors as novel cancer chemotherapeutic agents. *Trends Mol. Med.* **2002**, *8*, S55–61.
- (159) Imai, J.; Maruya, M.; Yashiroda, H.; Yahara, I.; Tanaka, K. The molecular chaperone Hsp90 plays a role in the assembly and maintenance of the 26S proteasome. *EMBO J.* **2003**, *22*, 3557–3567.
- (160) Brown, A. P.; Johnson, P.; Rawsthorne, S.; Hills, M. J. Expression and properties of acyl-CoA binding protein from *Brassica napus*. *Plant Physiol. Biochem. (Paris, Fr.)* **1998**, *36*, 629–635.
- (161) Suzui, N.; Nakamura, S.; Fujiwara, T.; Hayashi, H.; Yoneyama, T. A putative acyl-CoA-binding protein is a major phloem sap protein in rice (*Oryza sativa* L.). *J. Exp. Bot.* **2006**, *57*, 2571–2576.
- (162) Ambard-Bretteville, F.; Sorin, C.; Rébeillé, F.; Hourton-Cabassa, C.; Colas des Francs-Small, C. Repression of formate dehydrogenase in *Solanum tuberosum* increases steady-state levels of formate and accelerates the accumulation of proline in response to osmotic stress. *Plant Mol. Biol.* **2003**, *52*, 1153–1168.

RESEARCH ARTICLE

Characterisation of the nuclear proteome of a dehydration-sensitive cultivar of chickpea and comparative proteomic analysis with a tolerant cultivar

Pratigya Subba*, Rajiv Kumar*, Saurabh Gayali, Shubhendu Shekhar, Shaista Parveen, Aarti Pandey, Asis Datta, Subhra Chakraborty** and Niranjan Chakraborty

National Institute of Plant Genome Research, New Delhi, India

Water deficit or dehydration hampers plant growth and development, and shrinks harvest size of major crop species worldwide. Therefore, a better understanding of dehydration response is the key to decipher the regulatory mechanism of better adaptation. In recent years, nuclear proteomics has become an attractive area of research, particularly to study the role of nucleus in stress response. In this study, a proteome of dehydration-sensitive chickpea cultivar (ICCV-2) was generated from nuclei-enriched fractions. The LC-MS/MS analysis led to the identification of 75 differentially expressed proteins presumably associated with different metabolic and regulatory pathways. Nuclear localisation of three candidate proteins was validated by transient expression assay. The ICCV-2 proteome was then compared with that of JG-62, a tolerant cultivar. The differential proteomics and in silico analysis revealed cultivar-specific differential expression of many proteins involved in various cellular functions. The differential tolerance could be attributed to altered expression of many structural proteins and the proteins involved in stress adaptation, notably the ROS catabolising enzymes. Further, a comprehensive comparison on the abiotic stress-responsive nuclear proteome was performed using the datasets published thus far. These findings might expedite the functional determination of the dehydration-responsive proteins and their prioritisation as potential molecular targets for better adaptation.

Received: August 22, 2012

Revised: March 6, 2013

Accepted: March 19, 2013

Keywords:

Adaptive strategies / Comparative proteomics / Dehydration / Nuclei-enriched fraction / Pulse legume / ROS catabolising enzymes



Additional supporting information may be found in the online version of this article at the publisher's web-site

1 Introduction

Throughout their lifecycle, plants are exposed to a myriad of environmental factors leading them to develop mech-

anisms to sensitively counteract such conditions, and accordingly modify their physiology. It is increasingly clear that water deficit or dehydration is the main constrain on crop yield potential as it greatly influences the growth and productivity of crop worldwide [1, 2]. With water scarcity hitting the agricultural sector, it has been speculated that the yield reduction will only worsen with the unpredictable climate changes in the near future. Dehydration, in general, alters the normal physiological processes leading to disruption of water potential gradients, loss of turgor, disruption of

Correspondence: Prof. Niranjan Chakraborty, National Institute of Plant Genome Research, Aruna Asaf Ali Marg, New Delhi 110067, India

E-mail: nchakraborty@nipgr.ac.in

Fax: +91-11-26716658

Abbreviations: DRPs, dehydration-responsive proteins; GAPDH, glyceraldehyde-3-phosphate dehydrogenase; GRPs, glycine-rich proteins; MDH, malate dehydrogenase

*These authors contributed equally to this work.

**Additional corresponding author: Dr. Subhra Chakraborty, E-mail: subhrac@hotmail.com

membrane integrity and denaturation of proteins [3]. Plants respond and resort to various adaptive strategies to cope up with dehydration at both cellular and molecular levels; for instance, by the accumulation of compatible osmolytes and proteins, specifically involved in stress tolerance. Indeed, alteration of protein synthesis or degradation is one of the fundamental metabolic processes that greatly influence dehydration tolerance [4, 5]. Increasing evidence indicates both quantitative and qualitative changes in protein expression and a synergistic relationship between the accumulation of dehydration-induced proteins and physiological adaptations of plants to water deficit [6–8].

Legumes play an important role in rain-fed and irrigated agriculture by improving physical, chemical and biological properties of soil, and are considered excellent crops for natural resource management and crop diversification. Chickpea is the third most important pulse legume in the world, and is grown mainly in the arid and semi-arid regions. It possesses deep tap root system, and displays relatively high levels of dehydration tolerance [9–11]. These properties may thus render chickpea as an important legume to understand dehydration tolerance mechanisms in plants. Also, to its advantage, is its taxonomic proximity with the model legume, *Medicago truncatula*. Nonetheless, the global production of chickpea is not only currently stagnant, but also negatively affected by moderate to severe water-deficit condition (40–50% reduction) [12, 13].

Cultivars adapt very differently in response to environmental stresses at different growth stages even though they maybe accustomed to exactly the same environment. Therefore, a better understanding of genotypic control of stress tolerance mechanisms is required to maintain their yield potentials. Several studies have drawn attention to genotypic differences between dehydration-tolerant and dehydration-sensitive cultivars. However, most of these reports have focussed on biochemical analyses for enzymes involved in various pathways or stress-induced transcript accumulation patterns [14–18]. Analysis of two closely related genetic lines differing in their dehydration tolerance through sub-cellular proteomics provides an opportunity to determine if any of the differentially regulated proteins with dehydration severity are related genetically to dehydration tolerance. In recent times, 2DE has been effectively used to characterise and distinguish cultivars and/or genotypes in many important crops, and even to identify single mutations with multiple effects [19–22]. All these studies demonstrate the power and the usefulness of proteomic approach to distinguish lines, populations, varieties and even species; and for addressing important questions in stress biology.

The nucleus is an important sub-cellular compartment that contains almost all genetic information, and is essential for gene expression and regulation. In a previous study, we had screened several commercial chickpea cultivars for dehydration tolerance, which revealed ICCV-2 as a relatively sensitive cultivar while JG-62 as relatively tolerant cultivar [23]. Further, we had developed the differential nuclear proteome of cv.

JG-62, and identified an array of dehydration-responsive proteins (DRPs) [24]. To better understand the molecular mechanism of dehydration response, we developed a nucleus-specific proteome of cv. ICCV-2 and identified 75 DRPs. This proteome was then compared with that of JG-62, the tolerant cultivar. The differential display of the dehydration-responsive nuclear proteomes showed unique as well as overlapping proteins between the two cultivars, clearly indicating their significant contribution towards conferring dehydration tolerance.

2 Materials and methods

2.1 Plant growth, dehydration treatment and tissue harvest

Chickpea seeds (*Cicer arietinum* L.) cv. ICCV-2 were sown in pots containing a mixture of soil and soilrite (2:1 w/w; 10–12 plants/1.5 L capacity pots with a diameter of 18 cm) in an environmentally controlled growth room. The seedlings were maintained at $25 \pm 2^\circ\text{C}$, $50 \pm 5\%$ relative humidity, with $270 \mu\text{mol m}^{-2} \text{s}^{-1}$ light intensity, under 16 h photoperiod. The soil moisture content was maintained approximately to 30% by providing 100 mL of water per pot every day. Three-week-old seedlings were subjected to gradual dehydration by withdrawing water, and aerial tissues from 10 to 12 pots were harvested every 24 h up to 144 h in two independent experiments, referred to as biological replicates. Samples were collected for each time point as pool of seedlings treated at the same time. The dehydration symptoms of the seedlings were monitored by leaflet wilting. There were no visible changes in the seedlings until 48 h of dehydration. Visible wilting of the leaflets occurred at 48 h of dehydration, and the damage aggravated further during 48–144 h (Supporting Information Fig. 1). The samples from the unstressed seedlings were collected at each time points, and were finally pooled to normalise the growth and developmental effects, if any. The harvested unstressed and stressed tissues were instantly frozen in liquid nitrogen and stored at -80°C until further use. Experiments were performed with at least three technical replicates from two biological replicates for each time points. In total, we generated and analysed 42 highly reproducible 2D gels (three technical replicates, two biological and seven time points).

2.2 Nuclear protein extraction and 2DE

Nuclei were isolated, as described earlier, and the nuclear integrity was assessed using DAPI dihydrochloride staining [25]. Proteins were extracted from the nuclei-enriched fraction using TriPure Reagent (Roche) according to the manufacturer's instructions. Enrichment of nuclear proteins was analysed using immunoblot, wherein anti-histone antibody (Abcam Limited, UK) was used to probe the histone variants.

Further, contamination from non-nuclear proteins was examined using antibodies against cytochrome c oxidase and Rubisco large subunit, both obtained from Agrisera, and visualised using HRP conjugated secondary antibodies (Santa Cruz Biotechnology).

For 2DE, the protein pellet was solubilised using IEF sample buffer (8 M urea, 2 M thiourea and 4% w/v CHAPS) and the protein concentration was determined using 2D Quant kit (GE Healthcare). Aliquots of 150 µg protein were diluted with 2D rehydration buffer (8 M urea, 2 M thiourea, 4% w/v CHAPS, 20 mM DTT, 0.5% v/v pharmalyte (pH 4–7) and 0.05% w/v bromophenol blue) followed by rehydration of 13 cm IPG strips, pH 4–7 and pH 6–11 with 250 µL of the solution. Electrofocusing was performed using the IPGphor system (GE Healthcare) at 20°C for 30 000 VhT. The focused strips were subjected to reduction with 1% w/v DTT in 10 mL of equilibration buffer (6 M urea, 50 mM Tris-HCl (pH 8.8), 30% v/v glycerol and 2% w/v SDS), followed by alkylation with 2.5% w/v iodoacetamide in the same buffer. The strips were then loaded on top of 12.5% polyacrylamide gels for SDS-PAGE. To reduce gel-to-gel variation, each protein preparation was analysed on at least three parallel 2D gels, representing three technical replicates. The electrophoresed proteins were stained with MS compatible Silver Stain Plus Kit (Bio-Rad Laboratories).

2.3 Digital image acquisition and data analysis

The gels were scanned using a FluorS Imaging System (Bio-Rad Laboratories) equipped with a 12-bit camera. The gel images were then analysed using PDQuest v.7.2.0 (Bio-Rad Laboratories). Three representative 2D gels, for each of the time points, were used to generate a standard gel. The entire set of seven standard gels, designated as the first-level matchset wherefrom spots were matched across the time points. The replicate gels had at least correlation coefficient value of 0.8. The standard gels comprised of spots that met with several criteria: each spot was present in at least two of the three replicate gels, and was qualitatively consistent in size and shape. The “low-quality” spots with a quality score < 30 were eliminated from further analysis. The spot densities in the first-level matchset were normalised against the total density of the gel image. Further, an ANOVA was performed on these replicate gels to select out the statistically significant expression data using MultiExperiment Viewer [26] at a stringency level of 0.05. A second-level matchset was generated that allowed a comparison of the standard gels from each of the time points. Three unaltered spots identified from across the time points were used for the second normalisation. The filtered spot quantities from the standard gels were assembled into a data matrix of high-quality spots from the seven time points for further analysis. The software MrpI parameter, along with the use of standard molecular mass marker proteins was used to calculate the experimental molecular mass and *pI*. As a final step, this second-level matchset of ICCV-2 was compared with that of JG-62 [24].

2.4 Mass spectrometric analysis

The differentially expressed protein spots were excised mechanically, destained and in-gel digested with trypsin, and the peptides were extracted according to standard techniques [27]. Peptides were analysed by ESI MS using an Ultimate 3000 HPLC system (Dionex) coupled to either Q-Star Pulsar *i* or Q-Trap 4000 mass spectrometer (Applied Biosystems). Tryptic peptides were loaded onto a C18 PepMap 100, 3 m (LC Packings) and separated with a linear gradient of water, acetonitrile and 0.1% formic acid. The MS/MS data were extracted using Analyst Software v.1.4.1 (Applied Biosystems). The spectra were analysed to identify the proteins using MASCOT sequence matching software v.2.1 (www.matrixscience.com) with either Ludwig NR database (version Q309; 9870917 sequences; 3412714483 residues) or MSDB 200050929 (2344227 sequences; 779380795 residues) or MSDB 20060831 (3239079 sequences; 1079594700 residues), and taxonomy set to Viridiplantae (Green Plants) and a homology-based search was performed. The database search criteria were peptide tolerance, ± 1.2 Da; fragment mass tolerance, ± 0.6 Da; maximum allowed missed cleavage, 1; variable modifications, methionine oxidation; instrument type, ESI-QUAD-TOF for Q-Star Pulsar *i*, and ESI-TRAP for Q-Trap 4000 mass spectrometer. Protein scores were derived from ions scores on a non-probabilistic basis for ranking protein hits, and the protein scores as the sum of a series of peptide scores. The commonly accepted score threshold $p < 0.05$ was set by MASCOT algorithm, and was based on the size of the database. Only proteins identified with two or more peptides were considered for further analysis. The protein IDs were considered based on several aspects of the MS/MS analysis that include the number of identified proteins, their matched peptides, sequence coverage and protein-hit ranks. In the case of same protein being identified in multiple spots where several peptides were found to be shared, differential expression pattern was scrutinised for each of these candidates, and the proteins were listed as independent entities. The putative function(s) of each of the identified protein was analysed in view of the metabolic role of the candidate protein(s) in the nucleus. The protein functions were assigned using a protein function database Pfam (<http://www.sanger.ac.uk/software/Pfam/>) or Inter-Pro (<http://www.ebi.ac.uk/interpro/>). As functional annotation is based on Pfam and Interpro, the functional redundancy, if any, is thus greatly minimised.

2.5 Bioinformatic analysis

The computational prediction for sub-cellular localisation of DRPs was carried out using the following sets of programs: NucPred (<http://www.sbc.su.se/~maccallr/nucpred/>) [28], WoLFPSORT (<http://wolfpsort.org/>) [29, 30], BaCelLo (<http://gpcr.biocomp.unibo.it/bacello/>) [31], YLoc+

(www.multiloc.org/YLoc) [32, 33] and Cello (<http://cello.life.nctu.edu.tw/>) [34].

The co-expression profile of the common DRPs in cv. ICCV-2 and JG-62 was represented by a heat map using the MultiExperiment Viewer software (The Institute of Genome Research, TIGR) [26]. The heat map was generated based on the log transformed fold-change expression values of the proteins over 144 h dehydration period. Further, statistical analysis was performed using PCA functions of the XLStat-Pro Version 2012.4.03 (www.xlstat.com) software.

2.6 Quantitative real-time PCR analysis

Total RNA from the unstressed and stressed chickpea seedlings was isolated using RNeasy Plant Mini Kit (Qiagen) as per the manufacturer's instruction. Reverse transcription was performed with 2 µg RNA using SuperScript[®] VIL0™ cDNA Synthesis Kit (Invitrogen). The gene specific [Dna]F-ATGAAACCCCTCCGACAAC, [Dna]R- TCCGAAGGGTTA CGAAGAGTGT; MDHF- CGCGATTGGGTTCTTGGA, MD HR-CCGTCAGAATACTCCATTG; GAPDHF-ACAAA GGGAGCAAGGCAGTTAG, GAPDHR- TGTGGTTGGTGT-CAATGCTGAT; 14–3–3F- GGCAAGGTTACACGCACGAT, 14–3–3R- CGCTCAACTTTTCGGTGTCT], and housekeeping gene primers for elongation factor-1 (EF-1) [EF1F- AGC-CCAAGAGACCCTCAGACA, EF1R- CACGTCCGATTG-GCACAGT] were designed with the help of the primer express software using the sequences retrieved from the Chickpea Transcriptome Database (CTDB). The qRT-PCR was done using ABI PRISM 7700 sequence detecting system (Applied Biosystems) with SYBR green dye.

2.7 Sub-cellular localisation

The cDNA sequences of three of the identified DRPs, 2'-hydroxy isoflavonoreductase (CaSN-291), a putative abscisic acid (ABA)-responsive protein (CaSN-621) and malate dehydrogenase (CaSN-293) were downloaded and amplified using the following set of primers (IRF-CACCATGGCCTCACAAAAC AGGATCTTG, IRR-AACGAATTGATCCAAATATTCATCA GCAGTG), (ABAF-CACCATGGGTGTTTTTCTTTTCGATG ATGAAC, ABAR-GTAA TCAGGATTAGCCAAAACGTAAC-CCTC), (MDF-CACCATGGCCAAAGATCCAGTTCCG TG, MDR-AGAGAGGCAAGAGTAAGCCAGATTC) and cloned into pENTR/D-TOPO vector (Invitrogen) and subsequently introduced into pGWB441 (a generous gift from Dr. T. Nakagawa, Shimane University, Japan). The C-terminus enhanced yellow fluorescent protein (EYFP)-linked fusion proteins were transiently expressed in tobacco leaf epidermal cells, and the fluorescent signals were captured using a confocal laser scanning microscope (Leica Microsystems).

3 Results and discussion

3.1 Nuclear isolation and purity assessment

A purified fraction of intact nuclei, with no appreciable level of contamination from other cellular organelles, was isolated according to the method described earlier [25]. The integrity of the nuclei was assessed using DAPI, a blue-fluorescent nucleic acid stain that preferentially stains dsDNA (Fig. 1A and B). The enrichment of nuclear proteins was further evaluated by analyzing 1D profile of histone core proteins, which served as the organelle marker protein during the enrichment procedure (Fig. 1C and D). Non-nuclear proteins were assessed using anti-cytochrome c oxidase and anti-Rubisco large subunit antibodies (Fig. 1E and F). These results altogether demonstrate the enrichment of the nuclear fraction and that the proteins had no detectable non-nuclear contamination.

3.2 Altered protein expression in the nucleus of cv. ICCV-2

In order to elucidate the molecular mechanism governing the dehydration response, a comprehensive dehydration-responsive nuclear proteome was developed from the sensitive cultivar ICCV-2. As a first step, the nuclear proteome was developed under unstressed condition, and changes in the proteome were further monitored under dehydration (0–144 h). The proteins from the nuclei-enriched fraction were resolved onto IPG strips (13 cm, pH 4–7 and 6–11). Repetitive analyses yielded similar and reproducible patterns of protein expression. Three representative experimental replicate gels (as defined in Section 2) were computationally combined using PDQuest software to generate the first-level matchset (Fig. 2A). Only the spots that survived several stringent criteria (classified as “high-quality spots”) were used to estimate spot quantities, although otherwise, a large number of protein spots were included in the matchset. For example, 375 spots were detected under unstressed condition, whereas 359 spots were classified as high quality. The spot densities were then normalised against the total density present in the respective gel to overcome the experimental errors introduced due to differential staining. Further, in order to analyse the abundance changes for each spot across the time points, a second-level matchset was created. The higher-level matchset contained a total of 511 unique spots (Fig. 2B). The spots were of high quality and the technical replicates revealed a reproducibility of about 91% (Table 1). From the higher-level matchset, the filtered spot quantities were assembled into a data matrix, indicating changes in the intensity of each spot during the course of dehydration. The second normalisation was done with a set of three unaltered spots, identified across the time points. The quantitative proteome data were analysed by creating statistical and quantitative analysis sets. The results revealed 189 DRPs, which showed more than 2.5-fold difference in expression values, at least at one time point.

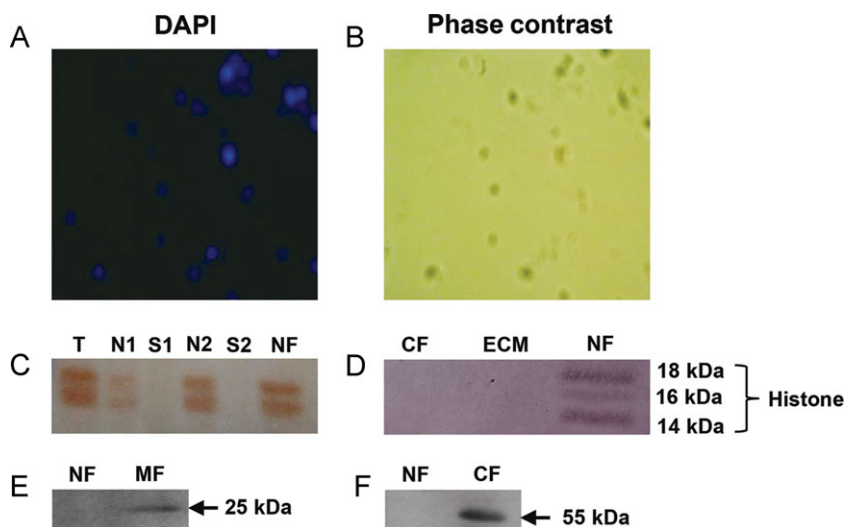


Figure 1. Purity analysis of nuclear fraction. The purified nuclear fraction of chickpea was stained with DAPI, and visualised by fluorescence microscopy. The DAPI-stained nuclei are shown in the left panel (A) and the phase-contrast micrograph of the nuclei is shown in the right panel (B). The enrichment of histone core complex which served as the nuclear marker was established by analysing 1DE profile (C). Western blot analysis of the fractionated protein was analysed using anti-histone (D), anti-cytochrome c oxidase (E) and anti-Rubisco large subunit (F) antibodies. T indicates total cell extract; N1 and S1 represent the nuclear and supernatant proteins, respectively; N2 and S2 represent the nuclear and supernatant protein after the first wash, respectively; and NF represents the purified nuclear fraction. MF indicates mitochondrial fraction whereas CF indicates the chloroplast fraction.

Table 1. Reproducibility of 2D gels at various dehydration treatment time points

Time (h)	Average no. of spots ^{a)}	High-quality spots ^{b)}	Reproducibility (%)
0	375	359	95.82
24	321	296	92.41
48	434	397	91.47
72	395	361	91.56
96	478	430	89.95
120	408	372	91.17
144	361	330	91.41
Total	2772	2545	91.97

a) Average number of spots present in three replicate gels of each time point.

b) Spots having quality score more than 30 assigned by PDQuest.

A total of 125 DRPs were subjected to MS/MS analysis, whereby the identities of 75 proteins were successfully established. Several of the analysed spots contained multiple proteins and in such cases the top-ranked hit was considered for their protein IDs. Protein spots identified with multiple IDs are listed in Supporting Information Document 1. The protein spots have been denoted as CaSN referring to the species *C. arietinum*, the cultivar sensitive and the organelle nucleus, and the accompanying numerals represent the spot number (Fig. 2B).

3.3 Functional classification of dehydration-responsive nuclear proteins in cv. ICCV-2

The DRPs were grouped into ten functional categories based on their putative functional roles (Fig. 3 and Table 2). “Molecular chaperones” represented 14% of the total protein spots whereas; proteins related to “Gene transcription and replication” comprised 21% of the identified proteins.

The other categories included “Cell signaling” (8%), “ROS catabolising enzymes” (5%), “Nucleocytoplasmic transport” (3%), “Chromatin structure and remodeling” (4%), “Protein translation” (3%) and “Protein degradation” (3%). The rest of the DRPs were grouped as “Miscellaneous” (19%). The proteins for which no known function could be assigned were grouped as “Unknown,” accounting for 20% of the dehydration-responsive nuclear proteins. To assign putative functions to these so called “orphan” proteins, domain analyses were carried out using the Conserved Domain Database (CDD, <http://www.ncbi.nlm.nih.gov/Structure/cdd/cdd.shtml>), and the results have been summarised in Supporting Information Table 1. A significant amount of redundancy was observed for several proteins, such as glyceraldehyde-3-phosphate dehydrogenase (GAPDH), glycine-rich RNA binding protein, among others. These DRPs were displayed as more than one isoelectric species with a change in MW and/or *pI*, indicating their probable PTM(s), or in some cases, they are likely to be the degraded products as evident from the gel-based size. The expression pattern of some of these proteins, such as cytosolic phosphoglycerate kinase, GAPDH and H⁺-transporting two-sector ATPase are shown in Fig. 4.

3.4 Statistical analysis

PCA is a mainstay of multivariate data analysis reducing the dimensionality of input datasets to reveal hidden dynamics that often underlie it. To assess the clustering behaviour of the identified ICCV-2 nuclear proteins in response to dehydration, PCA was performed using Pearson correlation matrix (*n*) (Supporting Information Fig. 2 and Supporting Information Table 2). Samples were accurately grouped by F1, which explains approximately 63% of the overall protein expression variability, whereas F2 explains approximately 18% of the variability. The first three PCs represented approximately 89% of the initial variability. In this analysis, proteins

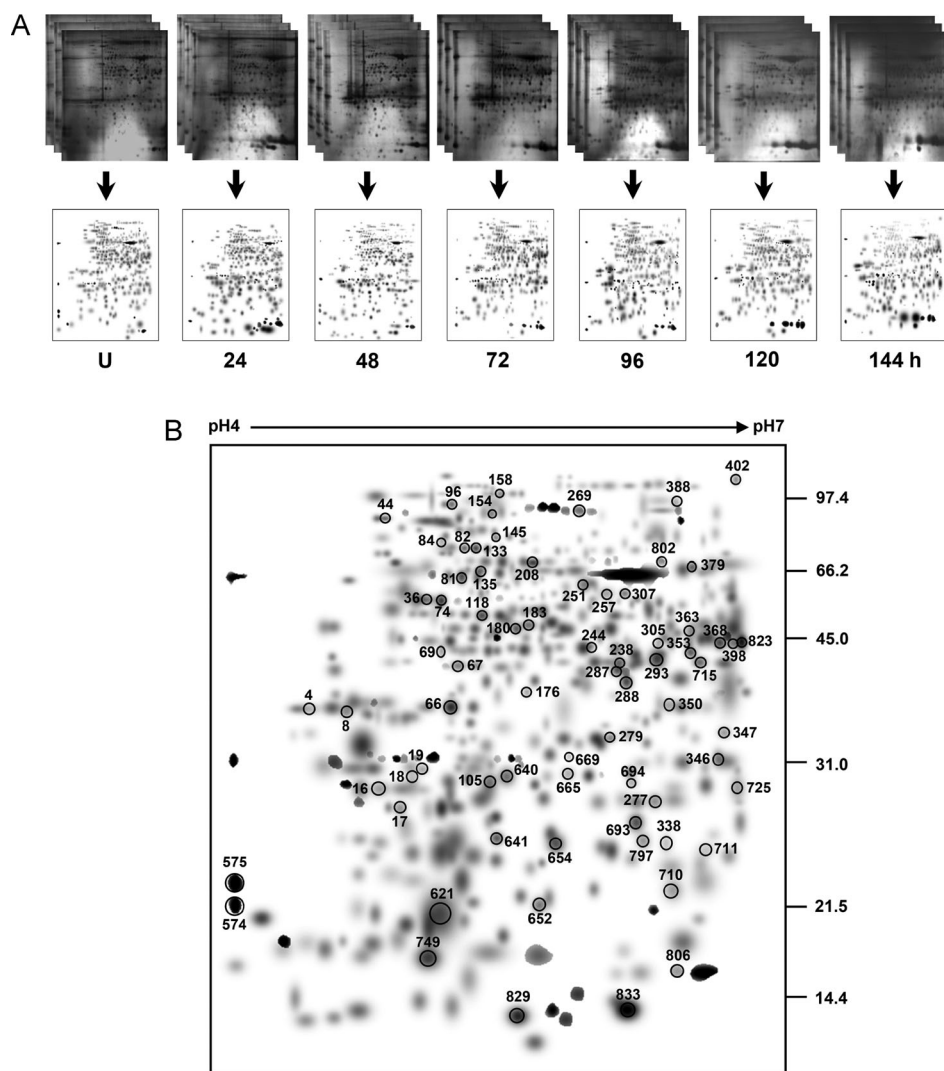


Figure 2. Dehydration-responsive nuclear proteome of the cv. ICCV-2 and the representative silver stained 2D gels. (A) Three-week-old seedlings were subjected to gradual dehydration, and tissues were harvested at 24 h interval for 144 h. Nuclear proteins were extracted from unstressed and stressed chickpea seedlings and separated on 13 cm IPG strips (pH 4–7) using IEF in the first dimension, followed by 12.5% SDS-PAGE in the second dimension. Three representative gels (selected from the technical and biological replicates) for each time point were computationally combined to generate the “standard gel” using PDQuest software. (B) Higher-level matchset of the dehydration-responsive proteins. The matchset was created in silico by integrating data from the “standard gels” for each time points as depicted in (A). The identified spots are encircled, and the numbers correspond to the spot IDs as mentioned in Table 2. U represents unstressed samples.

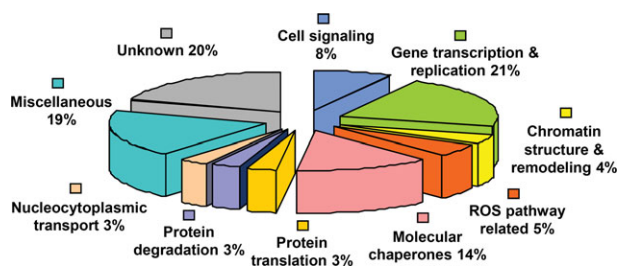


Figure 3. Functional classification of differentially expressed nuclear proteins of cv. ICCV-2. The putative functions of the dehydration-responsive proteins (DRPs) were assigned using Pfam and InterPro databases and the functional classes are represented by a pie chart.

were found to form distinct clusters on the biplot, for example, two members from the “Cell signaling” class, viz., similarity to receptor-like protein kinase (CaSN-693) and hybrid type

histidine kinase (CaSN-640) clustered together in the upper right quadrant. Similarly, proteins belonging to the class of “Chromatin structure and remodelling,” viz., histone H3 (CaSN-574) and histone H2B (CaSN-575), were found to cluster together. Interestingly, a group of proteins belonging to three different functional categories were also found to show a clustering pattern in the upper left quadrant due to their similar expression profiles, viz., WRKY-DNA binding domain protein (CaSN-338), flowering time control (FCA) protein (CaSN-711) and hypothetical protein OSJNBb0035114.1 (CaSN-652), belonging to the categories “Gene transcription and replication,” “Cell signaling,” and “Unknown” proteins.

3.5 qRT-PCR analysis

To understand the correlation pattern of the transcript and protein accumulation, qRT-PCR analysis was carried out on

Table 2. Dehydration-responsive nuclear proteins of cv. ICCV-2 and their relative changes

Functional category	Spot ID ^{a)}	Score	Identification	gi no. ^{b)}	No. of peptides	Protein expression profile ^{c)}	Percent coverage	Adjusted p-value	Thr MW/pI (kDa)	Exp MW/pI (kDa)	
Cell signalling	CaSN-711	43	FCA protein (fragment)	32482057	2		5	3.25E-04	29.37/8.63	21.93/6.65	
	CaSN-8	173	14–3–3-like protein	4775555	4		13	3.08E-04	29.42/4.71	36.76/4.60	
	CaSN-305	42	Serine-threonine kinase	38194927	2		5	0.002503714	45.58/5.48	46.78/6.39	
	CaSN-640	54	Hybrid-type histidine kinase	82466315	3		2	1.74E-04	113.17/8.65	29.70/5.50	
	CaSN-693	40	Similarity to receptor-like protein kinase	9758282	2		4	5.70E-09	71.94/9.44	23.10/6.26	
	CaSN-797	48	Predicted protein (diacylglycerol kinase domain)	226453803	3		2	3.19E-07	43.61/6.03	22.02/6.29	
	Gene transcription and replication	CaSN-287	41	Aspartate carbamoyl-transferase	15796550	2		3	1.05E-04	42.57/6.06	42.20/6.15
		CaSN-18	36	Putative transcription factor	12328532	2		1	2.24E-04	64.97/6.55	28.65/5.02
		CaSN-36	66	AAA ATPase, central region; homeodomain like	92874675	2		3	0.015551564	52.13/5.42	59.75/5.10
		CaSN-180	106	Cytosolic phosphoglycerate kinase	9230771	3		9	0.008415771	42.26/5.73	51.57/5.58
CaSN-183		239	Phosphoglycerate kinase (EC 2.7.2.3) precursor	1161600	5		14	7.85E-04	50.15/8.48	52.71/5.64	
CaSN-208		43	Retrotransposon protein, putative, Ty1-copia sub-class	77555208	2		2	1.36E-04	91.96/9.33	76.28/5.62	

Table 2. Continued

Functional category	Spot ID ^{a)}	Score	Identification	gi no. ^{b)}	No. of peptides	Protein expression profile ^{c)}	Percent coverage	Adjusted p-value	Thr MW/pi (kDa)	Exp MW/pi (kDa)
	CaSN-244	328	Fructose-bisphosphate aldolase (EC 4.1.2.13) precursor	169037	5		15	0.001977178	38.63/5.83	46.61/6.00
	CaSN-251	255	Transketolase (fragment)	4586600	5		32	1.86E-04	17.12/5.84	64.47/5.97
	CaSN-269	548	Glyceraldehyde-3-phosphate dehydrogenase (GAPDH; phosphorylating)	309671	11		24	1.56E-05	48.06/7.57	98.51/5.91
	CaSN-363	211	GAPDH (phosphorylating)	12159	4		10	3.05E-04	43.31/8.80	49.49/6.60
	CaSN-398	303	GAPDH (phosphorylating)	12159	6		14	1.24E-09	43.31/8.80	46.39/6.85
	CaSN-682	417	GAPDH, type I	92875110	10		17	0.031197106	47.91/6.76	56.07/5.83
	CaSN-823	313	GAPDH	462138	11		34	0.001299108	36.59/6.55	46.63/6.76
	CaSN-338	37	AY045676 NID (WRKY-DNA binding domain)	27363252	2		3	1.07E-05	55.78/8.18	22.30/6.42
	CaSN-368	101	G5bf protein	17064988	2		5	2.49E-04	42.59/8.65	45.00/6.76
	CaSN-833	67	DNA-directed RNA polymerase (fragment)	67810699	9		1	1.82E-05	65.52/7.06	12.00/6.2
Chromatin remodeling	CaSN-118	177	Actin (fragment)	1498334	4		11	0.001552693	37.14/5.47	55.73/5.37
	CaSN-574	42	Histone H3, putative, expressed	77555106	2		14	1.08E-04	15.42/11.06	18.40/4.03

Table 2. Continued

Functional category	Spot ID ^{a)}	Score	Identification	gi no. ^{b)}	No. of peptides	Protein expression profile ^{c)}	Percent coverage	Adjusted p-value	Thr MW/pI (kDa)	Exp MW/pI (kDa)
	CaSN-575	74	A15g59910-histone H2B	16323079	3		16	2.59E-06	16.44/10.00	20.09/4.05
ROS pathway related	CaSN-158	41	NADPH oxidase	87116554	2		2	0.002456	105.19/9.12	105.77/5.49
	CaSN-176	62	Glyoxalase/bleomycin resistance protein/dioxygenase	92887944	2		6	1.02E-05	32.22/5.13	39.69/5.64
	CaSN-293	247	Malate dehydrogenase (EC 1.1.1.37)	10334493	5		17	0.00741007	35.47/5.92	43.05/6.37
	CaSN-829	116	Thioredoxin h1	74058512	6		30	7.45E-14	12.79/5.77	12.1/5.64
Molecular chaperones	CaSN-16	73	Putative glycine-rich RNA-binding protein 2	6911146	2		15	6.68E-04	16.25/7.82	28.58/4.80
	CaSN-67	52	Putative glycine-rich RNA-binding protein 2	6911146	2		14	0.010621212	16.25/7.82	43.08/5.23
	CaSN-277	79	Putative glycine-rich RNA-binding protein 2	6911146	3		22	0.04200395	16.25/7.82	27.41/6.33
	CaSN-379	74	Putative glycine-rich RNA-binding protein 2	6911146	2		15	1.35E-04	16.25/7.82	70.76/6.60
	CaSN-17	106	Beta-conglutin	46451223	2		5	0.020016968	62.09/6.43	26.75/4.90
	CaSN-44	433	dnaK-type molecular chaperone CSS1 precursor	169023	9		13	0.007065051	75.47/5.22	93.45/4.77
	CaSN-715	45	HSP DnaJ, N-terminal	92896166	2		7	7.69E-04	53.40/7.12	42.96/6.58

Table 2. Continued

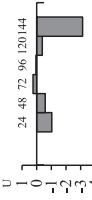
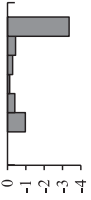
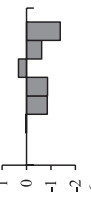


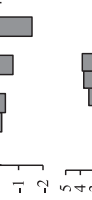
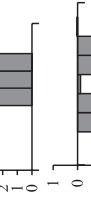

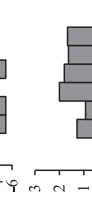

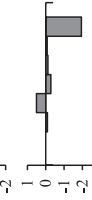

Functional category	Spot ID ^{a)}	Score	Identification	gi no. ^{b)}	No. of peptides	Protein expression profile ^{c)}	Percent coverage	Adjusted p-value	Thr MW/pI (kDa)	Exp MW/pI (kDa)
Protein translation	CaSN-82	213	Probable chaperonin 60 beta chain	806808	5		9	2.26E-06	62.94/5.85	81.07/5.27
	CaSN-133	276	Probable chaperonin 60 beta chain	806808	6		10	3.09E-05	62.94/5.85	81.15/5.36
Protein translation	CaSN-641	249	HSP 18.2	19618	6		27	0.001409249	18.15/5.84	22.87/5.50
	CaSN-654	73	HSP Hsp20	87240494	3		16	3.34E-08	18.18/5.81	23.90/5.79
	CaSN-257	69	Putative elongation factor 1-gamma	46806490	2		4	3.22E-05	46.90/6.31	63.49/6.16
	CaSN-279	158	Elongation factor 1-gamma	3868758	4		11	2.56E-04	47.45/6.10	33.75/6.12
Protein degradation	CaSN-84	440	Putative FtsH-like protein Ptf	52075838	8		11	0.010708383	72.49/5.54	83.12/5.19
	CaSN-350	41	Hypothetical protein (kelch repeat-containing F-box family protein)	48210029	2		4	0.019803619	56.81/4.93	37.23/6.47
Nucleocyto-plasmic transport	CaSN-347	131	ATRAN2	1668706	4		19	6.06E-09	25.02/6.65	34.24/6.79
	CaSN-19	44	Putative sorbitol transporter	34393630	2		4	0.034973074	54.38/9.27	29.84/5.09
Miscellaneous	CaSN-66	116	Oxygen-evolving complex protein 1 precursor	344004	3		8	6.51E-04	34.87/6.25	37.33/5.20
	CaSN-81	233	ATP synthase beta chain	69214424	5		9	0.012890223	53.36/5.23	67.06/5.25

Table 2. Continued

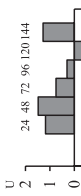
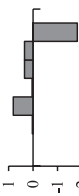


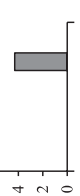

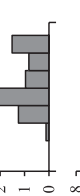


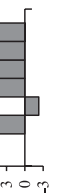
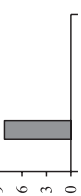
Functional category	Spot ID ^{a)}	Score	Identification	gi no. ^{b)}	No. of peptides	Protein expression profile ^{c)}	Percent coverage	Adjusted p-value	Thr MW/pI (kDa)	Exp MW/pI (kDa)
	CaSN-105	47	atpB (Fragment)	6110504	3		88	0.03228964	10.72/12.00	29.33/5.38
	CaSN-135	366	H ⁺ -transporting two-sector ATPase (EC 3.6.3.14) beta chain	18831	5		9	0.004398435	60.22/5.95	69.46/5.39
	CaSN-353	101	H ⁺ -transporting two-sector ATPase (EC 3.6.3.14) gamma chain precursor	19785	2		7	3.72E-08	41.42/8.16	44.10/6.60
	CaSN-238	37	NBS-LRR type disease resistance protein Rps1-k-1	62632823	2		1	0.001870041	139.71/7.29	43.42/6.09
	CaSN-291	54	2'-hydroxy isoflavonoreductase-transcriptional repressor activity	17949	3		9	2.47E-05	35.38/5.94	43.26/6.23
	CaSN-800	269	Isoflavonoreductase	1708425	6		22	0.001243575	35.39/5.94	44.34/6.24
	CaSN-346	102	Carbonic anhydrase (EC 4.2.1.1)	20502881	2		5	8.25E-05	35.28/6.96	30.63/6.75
	CaSN-388	136	Methionine synthase	71000469	3		3	5.68E-08	87.75/6.05	98.50/6.49
	CaSN-402	153	Glycine dehydrogenase (decarboxylating) component P precursor	20741	4		3	0.003205912	114.61/7.17	108.0/6.85
	CaSN-710	46	Nitrate reductase (EC 1.6.6.1) (EC 1.6.6.2) (EC 1.6.6.3)	6573214	2		1	6.49E-06	97.71/5.92	18.20/6.48
	CaSN-749	48	Putative uncharacterised protein At4g13990 (exostosin family-tumour suppressor genes)	42566766	2		3	4.49E-05	60.95/5.66	13.08/5.07

Table 2. Continued

Functional category	Spot ID ^{a)}	Score	Identification	gi no. ^{b)}	No. of peptides	Protein expression profile ^{c)}	Percent coverage	Adjusted p-value	Thr MW/pI (kDa)	Exp MW/pI (kDa)
							18	1.37E-05	21.65/6.09	28.68/6.21
Unknown	CaSN-694	162	Benzoquinone reductase	124488474	3					
							7	4.29E-06	42.22/5.64	47.07/5.19
	CaSN-69	164	OSJNBa0042F-21.13 protein	38347311	3					
							3	4.85E-05	51.97/5.69	59.96/5.18
	CaSN-74	119	At2g39730/T517.3	23308421	2					
							2	0.003824203	85.94/5.48	97.0/5.21
	CaSN-96	58	Hypothetical protein At1g62750	14532624	2					
							12	0.021685235	15.91/9.61	86.61/5.48
	CaSN-145	41	Hypothetical protein P0031A09.11	53792900	2					
							2	1.43E-05	64.81/7.57	94.75/5.40
	CaSN-154	45	Hypothetical protein P0033D06.7	52353656	2					
							4	0.003094569	40.55/8.70	41.24/6.20
	CaSN-288	68	VFU14956 NID	729479	2					
							2	0.00435662	93.70/5.87	62.09/6.18
	CaSN-307	41	OSJNBa0049H08.9 protein	38347658	2					
							48	2.01E-05	16.66/5.17	16.74/5.18
	CaSN-621	158	Putative ABA-responsive protein	6469115	5					
							3	0.006158838	23.07/11.70	17.33/5.72
	CaSN-652	39	Hypothetical protein OSJNBb0035114.1	54290266	2					
							1	5.11E-04	139.43/9.46	30.02/5.88
	CaSN-665	46	Hypothetical protein At2g23880 [imported]	75220227	2					
							1	1.76E-04	132.93/7.13	31.64/5.87
	CaSN-669	44	AC027036 NID	14475950	2					

Table 2. Continued

Functional category	Spot ID ^{a)}	Score	Identification	gi no. ^{b)}	No. of peptides	Protein expression profile ^{c)}	Percent coverage	Adjusted p-value	Thr MW/pI (kDa)	Exp MW/pI (kDa)
	CaSN-677	44	OSJNBa0089-N06.25 protein	21741073	2		3	0.005678163	16.21/11.70	49.00/6.07
	CaSN-725	47	Hypothetical protein A12g38370 (imported)	3395435	3		3	2.95E-06	60.29/9.07	28.33/6.81
	CaSN-802	46	Putative uncharacterised protein	218186008	2		2	3.51E-06	34.14/4.44	71.69/6.38
	CaSN-806	44	Putative uncharacterised protein	218186008	2		2	2.29E-09	34.14/4.44	12.42/6.45

a) Spot number as given on the 2D gel images. The first letters (Ca) represent the source plant, *Cicer arietinum*, followed by the sensitive cultivar (S) and sub-cellular fraction, Nuclear (N). The numerals indicate the spot numbers corresponding to Fig. 2B.

b) Gene identification number as in GenBank.

c) Protein expression profile represents the average change in spot density at various time points [U (unstressed), 24–144 h of dehydration]. The data were taken in terms of fold expression with respect to the unstressed value and were log transformed to the base two in order to level the scale of expression and to reduce the noise.

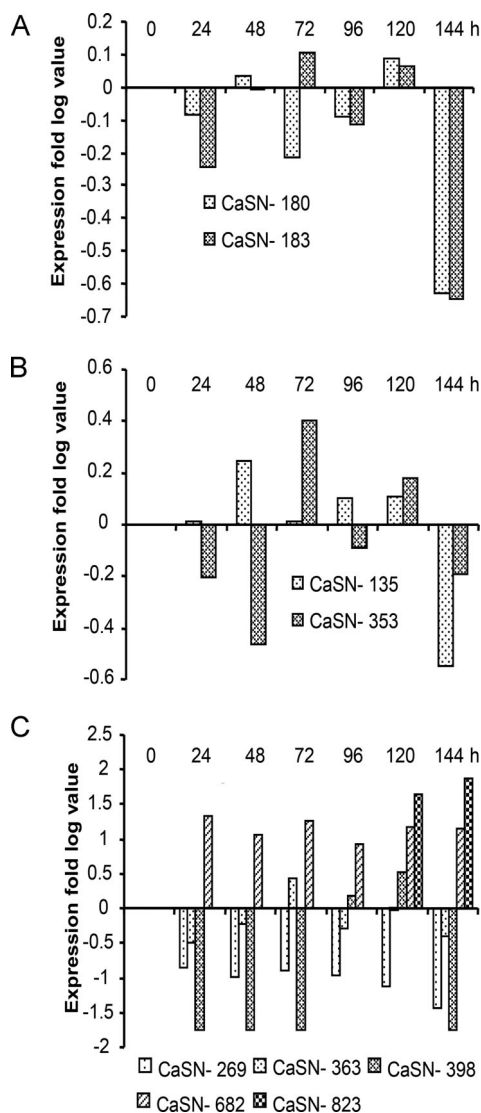


Figure 4. Nine differentially expressed proteins detected by 2DE matched to three unique proteins, and grouped into three isoform classes (A, B and C). The numbered spots correspond to the proteins listed in Table 2 and their log-transformed expression values were plotted against different time points during dehydration treatment for their expression profiles.

four randomly selected candidate genes. These genes represented three categories of protein expression patterns {down-regulated [Dna] N-terminal, CaSN-715 and malate dehydrogenase (MDH), CaSN-293], mixed pattern [GAPDH, CaSN-363] and upregulated [14–3–3-like protein, CaSN-8]}. While the transcript accumulation pattern of Dna] N-terminal and MDH was found to be in concordance with the protein expression pattern, the expression of GAPDH and 14–3–3 like protein was different from that of 2DE pattern (Fig. 5). The differences between RNA expression and protein accumulation might be due to the complexities of protein expression and multistep processes [35, 36], and the presence of multi-gene families [37, 38].

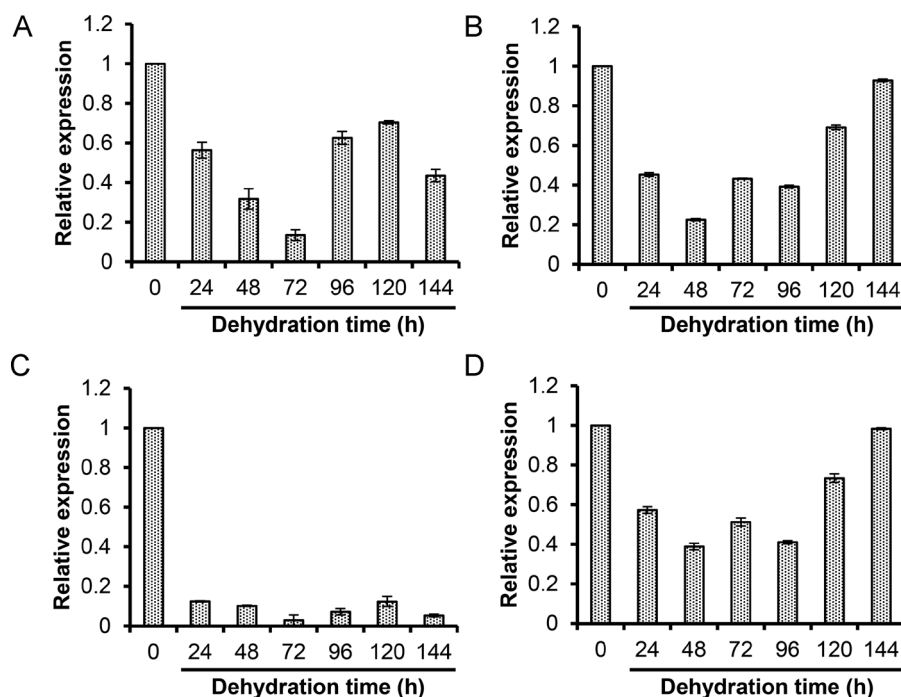


Figure 5. Effect of dehydration on the transcript accumulation of four candidate proteins of cv. ICCV-2. Four candidate proteins were selected randomly representing all the three categories of proteins having downregulated (DnaJ, N-terminal, CaSN-715 and MDH, CaSN-293), mixed (GAPDH, CaSN-363) and upregulated (14-3-3-like protein, CaSN-8) expression pattern. The transcript accumulation pattern shows the similar expression with reference to its protein expression for malate dehydrogenase (MDH) (A) and DnaJ (B). Whereas GAPDH (C) and 14-3-3 (D) showed a different expression pattern. The mean values of three replicates were normalised using elongation factor-1 (EF-1) as the internal control.

3.6 Prediction of nuclear localisation by bioinformatic predictors and validation

Protein localisation to the nucleus is accomplished by a complex battery of activities that help the cell to maintain homeostasis. Based on the localisation signature motifs, several software programs have been devised to predict the probability of a protein's nuclear localisation. To validate the nuclear localisation of the DRPs, five software programs were used as detailed in Section 2. NucPred analyses a eukaryotic protein sequence and predicts if the protein spends at least some time in the nucleus or spends no time. WoLFPSORT, on the other hand, predicts the sub-cellular localisation sites of proteins based on their amino acid sequences. BaCellO (balanced sub-cellular localisation predictor) is a predictor for the sub-cellular localisation of proteins in eukaryotes. While YLoc+ is an interpretable server that determines the sub-cellular localization of a protein based upon its biological property, Cello (sub-cellular localisation predictor) uses four types of sequence coding schemes: the amino acid composition, the dipeptide composition, the partitioned amino acid composition and the sequence composition based on the physico-chemical properties of amino acids. To remove redundancy, if any, the data were first filtered in terms of gi (genInfo identifier) number. In the non-redundant dataset, the proteins were considered to be nuclear localised if the output result was positive using the above-mentioned programs. About 48% of the DRPs were predicted to be localised in the nucleus (Supporting Information Document 2).

To experimentally validate the nuclear localisation, the cDNAs of three of the identified DRPs, viz., putative ABA-responsive protein (CaSN-621), MDH (CaSN-293) and 2'-

hydroxy isoflavone reductase (CaSN-291) were cloned, and the EYFP-fusion proteins were transiently expressed in tobacco leaf epidermal cells. The expression analysis of EYFP-fusion proteins showed fluorescence in the nucleus confirming the nuclear localisation of the candidate proteins (Fig. 6).

3.7 Comparative proteomics of nucleus of cv. ICCV-2 and JG-62

Several contrasting trends, as well as commonalities, were noticed in the proteomes of the two cultivars under unstressed condition in the context of 2DE spots. Whereas, 250 protein spots were found to be common to both the cultivars, 108 spots were exclusive to cv. ICCV-2 and 88 to JG-62 in the pH range 4–7 (Supporting Information Fig. 3A). A comparison of the proteome in the pH range 6–11 revealed 122 common protein spots, while 48 and 23 spots were exclusive to cv. JG-62 and ICCV-2, respectively (Supporting Information Fig. 3B). Spots were randomly chosen for identification from the pH range 4–7, and the results have been summarised in Table 2. The list of the common spots identified from both the cultivars is enlisted in Supporting Information Table 3. A few proteins showed redundancy as more than one spot were identified as the same protein; for example glycine-rich RNA binding protein1, aldolase and GAPDH were present as multiple spots in both the cultivars. Chaperonin 60, FCA, actin and kinesin were some of the other proteins which were common between the two cultivars. ROS catabolizing enzymes, for example NADPH oxidase, superoxide dismutase, ascorbate peroxidase as well as glyoxalase were also found to be common. Since these proteins were earlier found to be

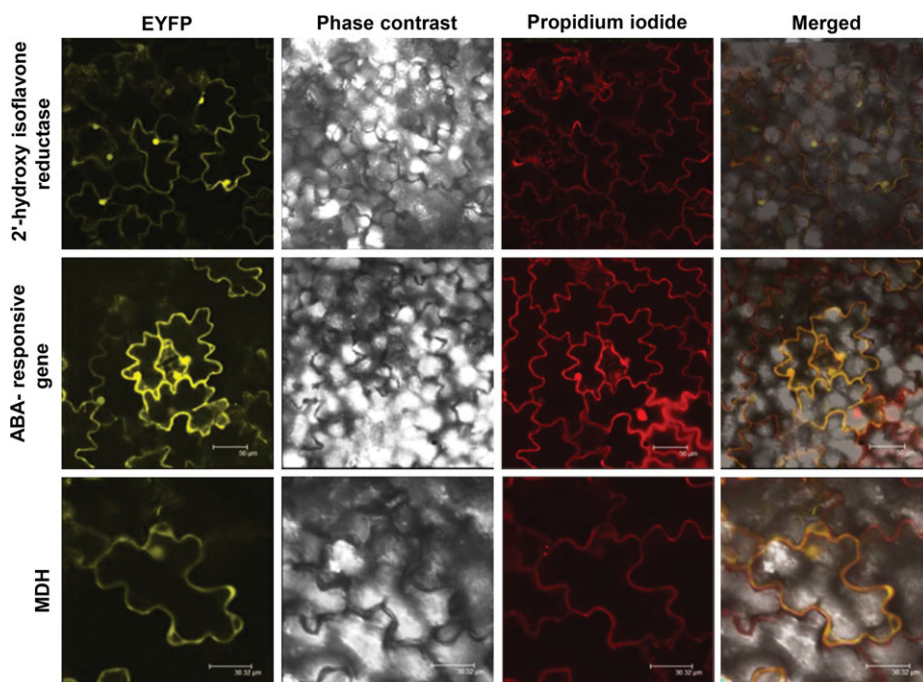


Figure 6. Validation of sub-cellular localisation of dehydration-responsive candidate proteins. The cDNAs of three dehydration-responsive proteins (DRPs), viz., 2'-hydroxy isoflavone reductase, ABA-responsive protein and malate dehydrogenase were fused in-frame to the 5'-end of EYFP and transiently expressed in tobacco leaf epidermal cells through agroinfiltration. The localisation of the proteins was visualised using confocal microscopy. The EYFP (yellow) and phase-contrast micrographs of the cells are shown in the first and second panels, respectively. The counter-staining with propidium iodide (red) is shown in third panel, while the overlaps of the cell fluorescence are shown in fourth panel.

dehydration responsive in cv. JG-62, it may be concluded that the basic machinery associated with stress tolerance in these cultivars remains the same.

Cultivar-specific protein expression is a complex phenomenon which is governed by several factors, such as different gene loci, multiple alleles, different subunit interaction, different splice variants or different PTMs [21]. We observed a cultivar-specific expression of certain structural proteins (e.g. histone H2B and H3), and proteins that are involved in the stress adaptation (e.g. sHsp, DnaJ) in cv. ICCV-2. On the other hand, 2-cys peroxiredoxin, 14–3–3, AP2/EREBP transcription factor BABY BOOM, among others, were found to be specific to cv. JG-62.

3.8 Comparative proteomics of dehydration-responsive nuclei

The identification of suitable plant characters for screening large numbers of genotypes, in a short time, at critical stages of growth, with the aim of selecting dehydration-tolerant cultivars, remains a major challenge. Proteomic tools and approaches may expedite breeding of genotypes that respond favourably to specific agro-climatic conditions, and guide the identification of proteins that protect crop performance under such conditions. With this aim, the dehydration-responsive proteome of cv. ICCV-2 was generated, and a comparison was made with the set of differentially expressed dehydration-responsive nuclear proteins of cv. JG-62. The analysis yielded unique, as well as overlapping proteins in the two cultivars, clearly indicating their significant contribution towards conferring tolerance. It must be noted that a higher number of DRPs (205) were detected in cv. JG-62 as compared to that of

189 in ICCV-2. Also, differences in kinetics of expression for the same protein between the two cultivars were found to be significant. A heat map highlighting the differential response for such common proteins is shown in Fig. 7. A comparison of the different functional classes of the DRPs however did not show much variation in the number of proteins represented in each class, except the class comprising of “ROS catabolizing enzymes” (Fig. 8).

A number of proteins involved in stress signalling were found to be dehydration responsive, most of them being common to both the cultivars. Hybrid type histidine kinase (CaSN-640) and receptor like protein kinase (CaSN-693) showed similar expression profiles in both the cultivars under dehydration, while protein kinase 2 was found to be upregulated in cv. JG-62 and non-differential in ICCV-2. The DRP with diacylglycerol kinase (CaSN-797) domain, on the other hand, was exclusively detected in cv. ICCV-2. The class “Gene transcription and replication” accounted for 21 and 17% of the DRPs in the sensitive and tolerant cultivars, respectively. GAPDH was detected as multiple isoforms in both the cultivars. Phosphoglycerate kinase (CaSN-180 and 183), G5bf protein (CaSN-368) and DNA-directed RNA polymerase (CaSN-833) were found to be differentially expressed exclusively in cv. ICCV-2, while fructose-1,6-bisphosphatase, AP2/EREBP transcription factor BABY BOOM and homoeobox protein HB3 were differentially expressed in cv. JG-62 (Supporting Information Table 4). Among the common proteins, AAA ATPase, central region; homeodomain like (CaSN-36) were downregulated in cv. ICCV-2, but upregulated in JG-62 under dehydration. In contrast, transketolase (CaSN-251) was upregulated in cv. ICCV-2 and downregulated in JG-62. WRKY-DNA binding domain protein (CaSN-338)

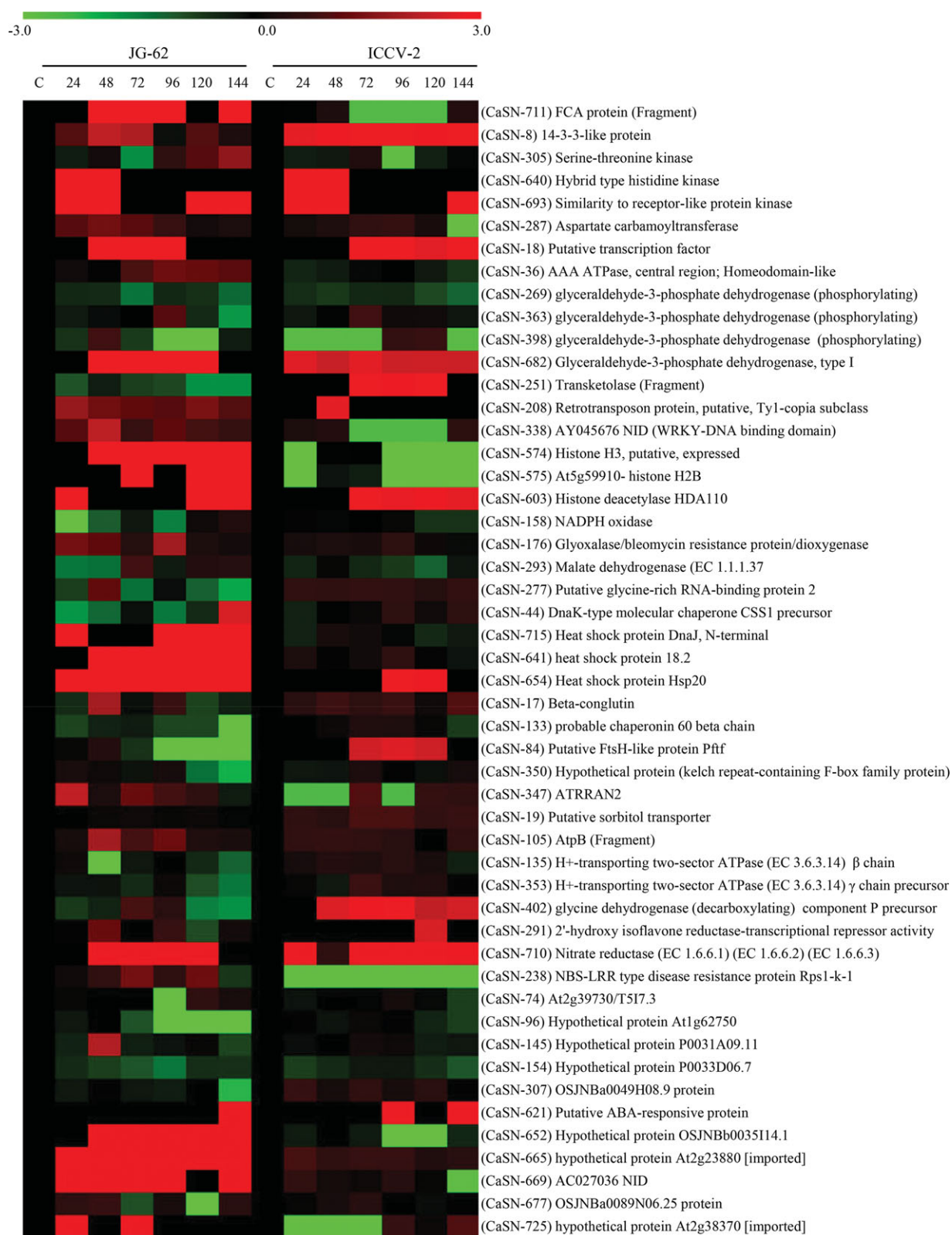


Figure 7. Heat map representation of the common dehydration-responsive proteins (DRPs) between cv. ICCV-2 and JG-62. The heat map was generated on the log transformed fold induction expression values using MultiExperiment Viewer software. Each protein is represented by a single row of coloured boxes and each time point is represented by a single column. Induction or repression ranges from pale to saturated shades of red and green, respectively.

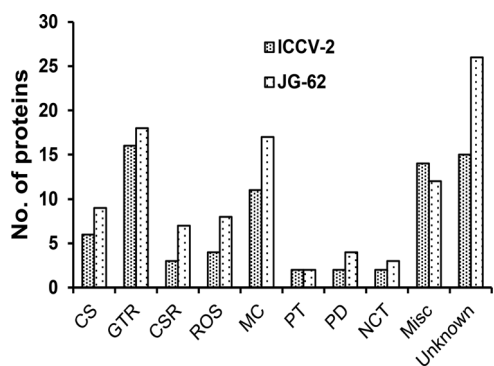


Figure 8. Major functional distributions of the dehydration-responsive proteins (DRPs) in cv. ICCV-2 and JG-62. The percentage of the identified DRPs categorised into different functional classes, such as cell signalling (CS); gene transcription and replication (GTR); chromatin structure and remodeling (CSR); ROS pathway related; molecular chaperones (MC); protein translation (PT); protein degradation (PD); nucleocytoplasmic transport (NCT); miscellaneous (Misc) and unknown function (Unknown), were analysed for both cultivars.

also showed dissimilar or rather contrasting profile between the two cultivars. The members of “Chromatin remodelling” class also showed contrasting expression profiles. Histone H1 has long been known to be induced by water deficit [39]. However, interestingly, we found other dehydration-responsive histone variants as well, for example histone H3 (CaSN-574) and H2B (CaSN-575). While both the proteins were found to be dehydration responsive, they were downregulated in cv. ICCV-2, but were highly induced in JG-62. DNA cytosine methyltransferase Zmet3 (CaN-240) was found to be induced exclusively in the tolerant variety.

The concept that redox signals are key regulators of plant metabolism, morphology and development is widely accepted. H_2O_2 is central to cross-tolerance phenomena and is a key component in the stress survival network. In response to any stress, the flux of H_2O_2 generation is increased. Relatively small increases or localised bursts of H_2O_2 influence only part of the network and modify gene expression in such a way to strengthen plant defence responses. In contrast, large increases in H_2O_2 trigger a distinct localised sequence of events in gene expression that leads inevitably to programmed cell death. Thus, removal of excess free radicals from the cell is necessary through specific detoxification mechanisms. Intriguingly, ROS catabolizing enzymes superoxide dismutase, ascorbate peroxidase and GSH peroxidase could not be detected in cv. ICCV-2, but were found to be induced in JG-62. Furthermore, though glyoxalase (CaSN-176) showed a downward trend in cv. ICCV-2, it was upregulated in JG-62 in response to progressive dehydration. The event of non-induction of major ROS catabolizing enzymes in cv. ICCV-2 might render its hypersensitivity to dehydration-induced oxidative damage.

Stress-induced alteration of proteolytic activities in plants is well documented [40]. This reorganization is necessary for

acclimation to stress, as well as to improve protein turnover, whereby the released amino acids may be recycled for synthesis of new proteins. A set of proteins involved in proteolytic degradation was found to be differentially expressed in both the cultivars. These include the kelch repeat-containing F-box family protein (CaSN-350) and putative FtsH-like protein Pff (CaSN-84). To protect the cellular system against stress-induced damage and to maintain functional protein conformations, a wide range of proteins with chaperone activity are known to swing into action [41]. The DnaK (CaSN-644), together with its co-chaperone DnaJ (CaSN-715) and GrpE, constitutes the KJE (DnaK, DnaJ and GrpE) system and prevents aggregation of misfolded proteins, thus, facilitating their subsequent refolding [24]. The small HSPs (sHSPs; CaSN-641 and 654) have been proposed to bind unfolded proteins, providing a reservoir of substrates available for subsequent refolding by the KJE system [42]. These proteins were downregulated in cv. ICCV-2 contrary to their upregulated expression in JG-62. Glycine-rich proteins (GRPs), the most abundant protein in this category, were identified from different spots and they displayed varied expression profile. The role of GRP has been implicated in plant responses to environmental stress, however, its exact function under stress remains largely unknown [43]. The upregulated expression of GRP2 in response to dehydration may stabilise mRNA, and enhance protein synthesis. Conglutin (CaSN-17) was another member found to be differentially expressed in both the cultivars.

Proteins known to play crucial roles in nucleocytoplasmic transport also showed differential response to dehydration. The nucleocytoplasmic transport is brought about mainly through the combined action of transport receptors, the importins and exportins [44]. Among the proteins identified in this category was ATRRAN2 (CaSN-347). A number of ATPases were also found to be dehydration responsive in both the cultivars. Among the proteins catalogued in the “Miscellaneous” category, carbonic anhydrase (CaSN-346), methionine synthase (CaSN-388) and benzoquinone reductase (CaSN-694) were found to be differentially expressed only in cv. ICCV-2. Dissection of the role for these differentially expressed proteins under dehydration would be very helpful in elucidating the mechanism of dehydration tolerance in plants.

3.9 In silico proteome analysis

In order to catalogue and compare the datasets published on the abiotic stress-responsive nuclear proteome in various plant species, a comparative study was carried out. The dataset of nuclear proteins identified in chickpea (cv. ICCV-2 and JG-62) was compared with that of the dehydration-responsive nuclear proteomes of *Oryza sativa* [45] and *Xerophyta viscosa* [46]. Also, included in this analysis was the cold-responsive nuclear proteome dataset of *Arabidopsis* [47] and those associated with seed filling process in *Medicago* [48]. Along with

identification of similarities and differences in the proteomes, an attempt was put forward so as to better understand the stress response in plants. A Venn diagram, inclusive of the organisms studied, was constructed to compile the results. While a large number of proteins remained exclusive to each of the datasets studied [chickpea (78), *O. sativa* (63), *X. viscosa* (12), *Medicago* (103) and *Arabidopsis* (34)], none of the proteins were found to be common in these species (Supporting Information Fig. 4A and Supporting Information Document 3). The inter-species comparison of overall distribution of functional classes was found to be similar, and no significant species-specific differences were observed in functional super classes (Supporting Information Fig. 4B). Nevertheless, these species showed differences mainly in the number of unique proteins within the super classes (Supporting Information Fig. 4C). It is possible that the crop species respond to adverse environmental conditions often causing a similar set of physicochemical reactions, and dynamic nature of such reactions dictate the tolerance levels. These findings suggest that multi-stress response in the nucleus constitutes an important adaptive trait for plants to withstand dynamic environments, and is also a potential mechanism to counteract such stress.

4 Concluding remarks

Crop plants exhibit different adaptive and acclimatisation strategies to adverse environmental conditions, including dehydration that range from seemingly simple phenotypic to complex physicochemical traits. These traits, often represented by differentially expressed proteins, may serve as important stress tolerance markers. Exploration of such traits is possible in crop species because different cultivars are available with differing degrees of tolerance, which provide correlative evidence for candidates involved in stress response. However, information available on the stress adaptive mechanism(s) shown by such tolerant genotypes has been fragmentary. One of the major factors is that the phenotypic traits that provide stress tolerance are often associative, and though they serve as important breeding tools, an understanding of the molecular mechanism of tolerance remains elusive.

The present study summarises the differential response of two contrasting cultivars of chickpea to dehydration in terms of the dynamic changes in the nucleus. In a previous proteomic study, we had established cv. JG-62 as a relatively tolerant cultivar, and identified an array of DRPs [24]. Now, with the development of dehydration-responsive nuclear proteome of a sensitive cultivar, ICCV-2, the differential physiological state at cellular level was reflected in the variance of the proteomes between the contrasting cultivars. The natural variation in the proteomes is important in the comparative analysis of dehydration sensitive versus tolerance mechanisms because the molecular differences in the cultivars might be the reason for differential response that would help in dissecting the underlying mechanisms of dehydration tolerance.

The major classes of DRPs, for example proteins involved in cell signalling, gene transcription and regulation, molecular chaperons, chromatin structure and remodelling, protein degradation, and nucleocytoplasmic transport were generally similar to those observed previously in other species. More specifically, our data highlights the dehydration-induced expression of enzymes, including those associated with the ROS catabolising network, exclusively in cv. JG-62. Further, several of the DRPs upregulated in cv. JG-62 were found to be downregulated in ICCV-2, which could contribute for its hypersensitive nature. The greater stress tolerance in cv. JG-62 might be due to its superior ability to maintain better water status and less oxidative damage as compared to ICCV-2. The presence of the unique proteins can be correlated to the differential dehydration response of the cultivars, which needs further investigation to obtain causative evidence. Further, cultivar-specific variation in terms of protein isoforms was also observed for some of the common proteins, and a detailed study in this panorama would be helpful in enhancing our understanding of the difference in these two cultivars with respect to dehydration response.

This work was supported by grants (BT/PR/10677/PBD/16/795) from the Department of Biotechnology (DBT), Government of India. We thank the Council of Scientific and Industrial Research (CSIR), Government of India and the DBT for providing predoctoral fellowship to PS, RK, SG and SP. We thank Dr. Suchismita Dass for critical reading of the manuscript and Mr. Jasbeer Singh for illustrations and graphical representation in the manuscript.

The authors have declared no conflict of interest.

5 References

- [1] Bhushan, D., Jaiswal, D. K., Ray, D., Basu, D. et al., Dehydration-responsive reversible and irreversible changes in the extracellular matrix: comparative proteomics of chickpea genotypes with contrasting tolerance. *J. Proteome Res.* 2011, 10, 2027–2046.
- [2] Sreenivasulu, N., Sopory, S. K., Kavi Kishor, P. B., Deciphering the regulatory mechanisms of abiotic stress tolerance in plants by genomic approaches. *Gene* 2007, 388, 1–13.
- [3] Ingram, J., Bartels, D., The molecular basis of dehydration tolerance in plants. *Annu. Rev. Plant Physiol. Plant Mol. Biol.* 1996, 47, 377–403.
- [4] Xiong, L., Wang, R. G., Mao, G., Koczan, J. M., Identification of drought tolerance determinants by genetic analysis of root response to drought stress and abscisic acid. *Plant Physiol.* 2006, 142, 1065–1074.
- [5] Bray, E. A., Plant responses to water deficit. *Trends Plant Sci.* 1997, 2, 48–54.
- [6] Seki, M., Narusaka, M., Ishida, J., Nanjo, T. et al., 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.

- [7] Shinozaki, K., Yamaguchi-Shinozaki, K., Gene expression and signal transduction in water-stress response. *Plant Physiol.* 1997, *115*, 327–334.
- [8] Shinozaki, K., Yamaguchi-Shinozaki, K., Molecular responses to drought and cold stress. *Curr. Opin. Biotechnol.* 1996, *7*, 161–167.
- [9] Leport, L., Turner, N. C., French, R. J., Barr, M. D. et al., Physiological responses of chickpea genotypes to terminal drought in a Mediterranean-type environment. *Eur. J. Agron.* 1999, *11*, 279–291.
- [10] Kumar, J., Abbo, S., in: Sparks, D. (Ed.), *Advances in Agronomy*, Elsevier, USA 2001, pp. 107–138.
- [11] Deokar, A. A., Kondawar, V., Jain, P. K., Karuppayil, S. M. et al., Comparative analysis of expressed sequence tags (ESTs) between drought-tolerant and -susceptible genotypes of chickpea under terminal drought stress. *BMC Plant Biol.* 2011, *11*, 70.
- [12] Ryan, J., in: Asthana, A., Ali, M. (Eds.), A global perspective on pigeonpea and chickpea sustainable production systems: present status and future potential, *Recent Advances in Pulses Research*, Indian Society for Pulses Research and Development, Kanpur, India 1997, pp. 1–31.
- [13] Ahmad, F., Gaur, P., Croser, J., in: Singh, R. J., Jauhar, P. P. (Eds.), *Chickpea (Cicer arietinum L.), Genetic Resources, Chromosome Engineering and Crop Improvement: Grain Legumes*, CRC Press, USA 2005, pp. 187–217.
- [14] Chen, D., Liang, M.-X., DeWald, D., Weimer, B. et al., Identification of dehydration responsive genes from two non-nodulated alfalfa cultivars using *Medicago truncatula* microarrays. *Acta Physiol. Plant.* 2008, *30*, 183–199.
- [15] Kathiresan, A., Lafitte, H. R., Chen, J., Mansueto, L. et al., Gene expression microarrays and their application in drought stress research. *Field Crops Res.* 2006, *97*, 101–110.
- [16] Tanyolaç, D., Ekmekçi, Y., Unalan, S., Changes in photochemical and antioxidant enzyme activities in maize (*Zea mays* L.) leaves exposed to excess copper. *Chemosphere* 2007, *67*, 89–98.
- [17] Chernikova, T., Robinson, J. M., Lee, E. H., Mulchi, C. L., Ozone tolerance and antioxidant enzyme activity in soybean cultivars. *Photosynth. Res.* 2000, *64*, 15–26.
- [18] Fazeli, F., Ghorbanli, M., Niknam, V., Effect of drought on biomass, protein content, lipid peroxidation and antioxidant enzymes in two sesame cultivars. *Biol. Plant.* 2007, *51*, 98–103.
- [19] Lehesranta, S. J., Davies, H. V., Shepherd, L. V. T., Nunan, N. et al., Comparison of tuber proteomes of potato varieties, landraces, and genetically modified lines. *Plant Physiol.* 2005, *138*, 1690–1699.
- [20] Carpentier, S. C., Witters, E., Laukens, K., Van Onckelen, H. et al., Banana (*Musa* spp.) as a model to study the meristem proteome: acclimation to osmotic stress. *Proteomics* 2007, *7*, 92–105.
- [21] Vincent, D., Ergül, A., Bohlman, M. C., Tattersall, E. A. et al., Proteomic analysis reveals differences between *Vitis vinifera* L. cv. Chardonnay and cv. Cabernet Sauvignon and their responses to water deficit and salinity. *J. Exp. Bot.* 2007, *58*, 1873–1892.
- [22] Kottapalli, K. R., Rakwal, R., Shibato, J., Burow, G. et al., Physiology and proteomics of the water-deficit stress response in three contrasting peanut genotypes. *Plant Cell Environ.* 2009, *32*, 380–407.
- [23] Bhushan, D., Pandey, A., Choudhary, M. K., Datta, A. et al., Comparative proteomics analysis of differentially expressed proteins in chickpea extracellular matrix during dehydration stress. *Mol. Cell. Proteomics* 2007, *6*, 1868–1884.
- [24] Pandey, A., Chakraborty, S., Datta, A., Chakraborty, N., Proteomics approach to identify dehydration responsive nuclear proteins from chickpea (*Cicer arietinum* L.). *Mol. Cell. Proteomics* 2008, *7*, 88–107.
- [25] Pandey, A., Choudhary, M. K., Bhushan, D., Chattopadhyay, A. et al., The nuclear proteome of chickpea (*Cicer arietinum* L.) reveals predicted and unexpected proteins. *J. Proteome Res.* 2006, *5*, 3301–3311.
- [26] Saeed, A. I., Bhagabati, N. K., Braisted, J. C., Liang, W. et al. TM4 microarray software suite. *Methods Enzymol.* 2006, *411*, 134–193.
- [27] Casey, T. M., Arthur, P. G., Bogoyevitch, M. A., Proteomic analysis reveals different protein changes during endothelin-1- or leukemic inhibitory factor-induced hypertrophy of cardiomyocytes *in vitro*. *Mol. Cell. Proteomics* 2005, *4*, 651–661.
- [28] Brameier, M., Krings, A., MacCallum, R. M., NucPred-predicting nuclear localization of proteins. *Bioinformatics* 2007, *23*, 1159–1160.
- [29] Horton, P., Park, K. J., Obayashi, T., Nakai, K., in: Jiang, T., Yang, U.-C., Chen, Y.-P. (Eds.), *Proceedings of the 4th Annual Asia Pacific Bioinformatics Conference*, Imperial College Press, London 2006, pp. 39–48.
- [30] Horton, P., Park, K.-J., Obayashi, T., Fujita, N. et al., WoLF PSORT: protein localization predictor. *Nucleic Acids Res.* 2007, *35*, W585–W587.
- [31] Pierleoni, A., Martelli, P. L., Fariselli, P., Casadio, R., BaCellO: a balanced subcellular localization predictor. *Bioinformatics* 2006, *22*, e408–e416.
- [32] Briesemeister, S., Rahnenführer, J., Kohlbacher, O., Going from where to why—interpretable prediction of protein subcellular localization. *Bioinformatics* 2010, *26*, 1232–1238.
- [33] Briesemeister, S., Rahnenführer, J., Kohlbacher, O., YLoc-an interpretable web server for predicting subcellular localization. *Nucleic Acids Res.* 2010, *38*, W497–W502.
- [34] Yu, C. S., Lin, C. J., Hwang, J. K., Predicting subcellular localization of proteins for Gram-negative bacteria by support vector machines based on n-peptide compositions. *Protein Sci.* 2004, *13*, 1402–1406.
- [35] Tian, Q., Stepaniants, S. B., Mao, M., Weng, L. et al., Integrated genomic and proteomic analyses of gene expression in mammalian cells. *Mol. Cell. Proteomics* 2004, *3*, 960–969.
- [36] Igwe, E. I., Essler, S., Al-Furoukh, N., Dehne, N. et al., Hypoxic transcription gene profiles under the modulation of nitric oxide in nuclear run on-microarray and proteomics. *BMC Genomics* 2009, *10*, 408.
- [37] Manjunath, S., Sachs, M. M., Molecular characterization and promoter analysis of the maize cytosolic glyceraldehyde 3-

- phosphate dehydrogenase gene family and its expression during anoxia. *Plant Mol. Biol.* 1997, 33, 97–112.
- [38] Xu, W. F., Shi, W. M. Expression profiling of the 14–3–3 gene family in response to salt stress and potassium and iron deficiencies in young tomato (*Solanum lycopersicum*) roots: analysis by real-time RT-PCR. *Ann. Bot.* 2006, 98, 965–974.
- [39] Scippa, G. S., Di Michele, M., Onelli, E., Patrignani, G. et al., The histone-like protein H1-S and the response of tomato leaves to water deficit. *J. Exp. Bot.* 2004, 55, 99–109.
- [40] Martinelli, T., Whittaker, A., Bochicchio, A., Vazzana, C. et al., Amino acid pattern and glutamate metabolism during dehydration stress in the ‘resurrection’ plant *Sporobolus stapfianus*: a comparison between desiccation-sensitive and desiccation-tolerant leaves. *J. Exp. Bot.* 2007, 58, 3037–3046.
- [41] Wiśniewski, K., Zagdańska, B., Genotype-dependent proteolytic response of spring wheat to water deficiency. *J. Exp. Bot.* 2001, 52, 1455–1463.
- [42] Lee, G. J., Vierling, E., A small heat shock protein cooperates with heat shock protein 70 systems to reactivate a heat-denatured protein. *Plant Physiol.* 2000, 122, 189–198.
- [43] Singh, U., Deb, D., Singh, A., Grover, A., Glycine-rich RNA binding protein of *Oryza sativa* inhibits growth of M15 *E. coli* cells. *BMC Res. Notes* 2011, 4, 18.
- [44] la Cour, T., Kiemer, L., Molgaard, A., Gupta, R. et al., Analysis and prediction of leucine-rich nuclear export signals. *Protein Eng. Des. Sel.* 2004, 17, 527–536.
- [45] Choudhary, M. K., Basu, D., Datta, A., Chakraborty, N. et al., Dehydration-responsive nuclear proteome of rice (*Oryza sativa* L.) illustrates protein network, novel regulators of cellular adaptation, and evolutionary perspective. *Mol. Cell. Proteomics* 2009, 8, 1579–1598.
- [46] Abdalla, K., Baker, B., Rafudeen, M. S., Proteomic analysis of nuclear proteins during dehydration of the resurrection plant *Xerophyta viscosa*. *Plant Growth Regul.* 2010, 62, 279–292.
- [47] Bae, M. S., Cho, E. J., Choi, E.Y., Park, O. K., Analysis of the *Arabidopsis* nuclear proteome and its response to cold stress. *Plant J.* 2003, 36, 652–663.
- [48] Repetto, O., Rogniaux, H., Firnhaber, C., Zuber, H. et al., Exploring the nuclear proteome of *Medicago truncatula* at the switch towards seed filling. *Plant J.* 2008, 56, 398–410.

Next-generation protein-rich potato expressing the seed protein gene *AmA1* is a result of proteome rebalancing in transgenic tuber

Subhra Chakraborty^{a,1,2}, Niranjan Chakraborty^{a,1}, Lalit Agrawal^a, Sudip Ghosh^a, Kanika Narula^a, Shubhendu Shekhar^a, Prakash S. Naik^b, P. C. Pande^c, Swarup Kumar Chakraborti^b, and Asis Datta^{a,2}

^aNational Institute of Plant Genome Research, New Delhi 110067, India; ^bCentral Potato Research Institute, Shimla, Himachal Pradesh 171001, India; and ^cCentral Potato Research Institute Campus, Modipuram, Uttar Pradesh 250110, India

Edited* by Eugene W. Nester, University of Washington, Seattle, WA, and approved August 17, 2010 (received for review May 5, 2010)

Protein deficiency is the most crucial factor that affects physical growth and development and that increases morbidity and mortality especially in developing countries. Efforts have been made to improve protein quality and quantity in crop plants but with limited success. Here, we report the development of transgenic potatoes with enhanced nutritive value by tuber-specific expression of a seed protein, AmA1 (Amaranth Albumin 1), in seven genotypic backgrounds suitable for cultivation in different agro-climatic regions. Analyses of the transgenic tubers revealed up to 60% increase in total protein content. In addition, the concentrations of several essential amino acids were increased significantly in transgenic tubers, which are otherwise limited in potato. Moreover, the transgenics also exhibited enhanced photosynthetic activity with a concomitant increase in total biomass. These results are striking because this genetic manipulation also resulted in a moderate increase in tuber yield. The comparative protein profiling suggests that the proteome rebalancing might cause increased protein content in transgenic tubers. Furthermore, the data on field performance and safety evaluation indicate that the transgenic potatoes are suitable for commercial cultivation. In vitro and in vivo studies on experimental animals demonstrate that the transgenic tubers are also safe for human consumption. Altogether, these results emphasize that the expression of AmA1 is a potential strategy for the nutritional improvement of food crops.

allergenicity | essential amino acids | nutritional health

Humans require a diverse and nutritionally well-balanced diet to maintain optimal health and depend largely on plants for their daily nutritional requirements. Moreover, a large proportion of the world's population is undernourished. Thus, nutritional improvement of crop plants is an urgent worldwide health issue as basic nutritional requirements for much of the world's population are still not met. Proteins, one of the principal constituents of a balanced diet, impart nutritional value to food due to their structural constituents, amino acids. However, it is very rare in nature to find all of the essential amino acids in a single food crop. Protein malnutrition is essentially caused by poor quality diets that include a high intake of staple crops with less protein and/or low-quality proteins in terms of amino acid composition. Protein deficiency lowers resistance to disease, delays physical growth and development, and may cause permanent impairment of the brain in infants and young children. A major effort has been to improve the amino acid composition of plant protein because animals, including humans, are incapable of synthesizing 10 of the 21 amino acids required for protein synthesis, and these "essential amino acids" must therefore be obtained from the diet (1). Because of the importance of dietary protein and the fact that plants are its major source, development of strategies to increase protein levels and the concentration of essential amino acids in food crops is of primary importance in a crop improvement program. There have been several attempts through mutant selection and engineering genes encoding key amino acid biosynthesis pathway enzymes to increase free essential amino acids in crop plants (2), but with limited success (3). A promising strategy is the genetic engineering of genes encoding proteins with high nutri-

tional value into food crops (1, 4–6). However, despite promises that genetically modified (GM) crops could make a significant contribution to achieving global food security, the new-generation GM varieties are primarily used for industrial crops, such as cotton and animal fodder (7), and few have been commercialized.

Although cereals and starchy food crops contribute to more than 80% of the calories in a diet, noncereal crops are becoming more popular throughout the world with the continuing change in food habits. The demand for noncereal crops will continue to increase as a consequence of the expanding human population. Potato is the most important noncereal food crop and ranks fourth in terms of total global food production. It is also used as animal feed and in industrial products. Today, potatoes are grown in nearly 125 countries and more than a billion people worldwide consume them on a daily basis (8). The total value of the crop is estimated at 40 billion dollars for the top 10 producing countries, which account for two-thirds of global potato production (9). The United Nations declared 2008 as the "International Year of the Potato," affirming the need to focus on the role that the potato can play in providing food security (10). Although in developing economies the majority of potato is used for direct consumption, a shift toward the use of potato in convenience foods, for example, in potato chips and fries, has dramatically increased in developed countries (11). Unfortunately, the nutritional quality of potato tubers is greatly compromised because they contain less protein and are deficient in lysine, tyrosine, and the sulfur-containing amino acids (12). To guarantee a sufficient supply of quality protein in a diet consisting mainly of staple foods such as potato, specific interventions in genetic engineering are an absolute necessity. However, there is currently public concern about the use of genetically engineered foods in contemporary agriculture, particularly when genes are from a nonplant source. During the past decade, several potential candidate genes have been targeted for the nutritional improvement of protein content in crops, namely: Brazil nut 2S albumin (13), *AmA1* (Amaranth Albumin 1) (14), β -phaseoline (15), HS-7 zein (3), cruciferin (16), sunflower seed albumin (17), and S-rich zein (18). However, excepting *AmA1* (14, 19, 20), introduction of these genes in target plants has often resulted in an increase in one of the amino acids at the expense of others, leading to an imbalance of the amino acid profile in transgenic crops.

AmA1 has great agricultural importance because it is a well-balanced protein in terms of amino acid composition, possessing even better values than recommended by the World Health Or-

Author contributions: S.C., N.C., and A.D. designed research; S.C., N.C., L.A., S.G., K.N., S.S., P.S.N., P.C.P., and S.K.C. performed research; S.C., N.C., L.A., S.G., K.N., S.S., P.S.N., P.C.P., S.K.C., and A.D. analyzed data; and S.C., N.C., and A.D. wrote the paper.

The authors declare no conflict of interest.

*This Direct Submission article had a prearranged editor.

¹S.C. and N.C. contributed equally to this article.

²To whom correspondence may be addressed. E-mail: subhrac@hotmail.com or asisdatta@hotmail.com.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1006265107/-DCSupplemental.

ganization for a nutritionally rich protein. More importantly, because it is a nonallergenic protein that originated from an edible crop, the transgenic crops expressing AmA1 would have greater acceptability. In a previous study, we showed that the *AmA1* cDNA can be expressed functionally in fission yeast (21) and in food crops (1). The name “protato” has been coined to describe the genetically modified potatoes that contain an increased quantity of protein with improved quality in terms of amino acid composition.

Geographical distributions of potato plants are greatly influenced by different agro-climatic conditions. In this work, we translated the approach of genetic modification into commercial potato varieties suitable for cultivation in different agro-climatic regions to overcome protein deficiency in potato tubers. Here, we describe the tuber-specific expression of *AmA1* into seven economically important potato genotypes and the results obtained from 2-y field trials in an advanced generation of the transgenic varieties. This study provides unequivocal evidence on the biosafety assessment and benefits of the genetically modified tubers, which would accelerate the platform for rapidly bringing products to consumers. This is a comprehensive report of translational research toward protein improvement programs in crop plants.

Results

Transgenic Plants Display a Wild-Type Growth and Developmental Phenotype. To develop new-generation potato cultivars, which might be suitable for practical use, the pSB8G construct (1) having full-length *AmA1* under the control of a tuber-specific, granule-bound starch synthase promoter, *GBSS*, was used for large-scale transformation into seven commercial potato cultivars by *Agrobacterium tumefaciens*-mediated transformation. We developed a robust, genotype independent, and reproducible transformation system using internodes as explant on the basis of our earlier protocol (1) with few modifications. Because the copy number and the site of integration greatly influence the transgene expression, dozens of primary independent transformants were developed and screened in each category to obtain the desired phenotype. Moreover, the genetic differences between potato varieties and the growing environment could also contribute to differences in performance. The putative transformants were selected on the basis of their growth on medium containing kanamycin as a selectable marker. We then transferred 1–2 dozen independent transgenic events from each category to the green house. There were no visible phenotypic changes in morphology either in the primary transgenic populations or in the subsequent generations.

Integration, Tissue-Specific Expression, and Accumulation of AmA1 in Transgenic Tubers. To identify the successful transgenic events, genomic PCR analysis was carried out on all kanamycin-positive transformants that revealed an amplicon of 1.02 kb, corresponding to the size of the *AmA1*-coding region. Furthermore, the transgene copy number in the PCR-positive plants was determined by real-time PCR using Taqman chemistry, which revealed the presence of a single copy of the transgene in most of the transgenic events with very few having two to three copies. In a step further, all of the kanamycin-selected, PCR-positive, and copy-number-detected transgenic events across the populations were transferred to the experimental plot for generating potato tubers. RT-PCR analysis revealed that the *AmA1* transcript was most abundant in tubers followed by the stolon and stem; the leaf showed the least *AmA1* transcript (Fig. S1A), suggesting the tissue-specific expression of the transgene. Our data support the fact that the *GBSS* promoter is most active in tubers, as reported earlier (22). Although the transcript was found in different organs, the protein products could be detected only in the field-grown tuber. Immunodetection of AmA1 in tubers showed a 35-kDa band in all of the transgenic events but at varying levels (Fig. S1B), suggesting that the transgene was actively producing the desired protein in the alien environment.

Overexpression of AmA1 Results in Increased Tuber Protein and Amino Acids. The protein content in crops is determined on the basis of nitrogen content, and the Kjeldahl method has been al-

Table 1. Comparison of total protein content of wild-type and AmA1-transgenic tubers

Genotype		Year I	Year II
K Chipsona 1	C	87.90 ± 1.30	83.60 ± 5.10
	21	131.70 ± 3.30 (49.83)	127.90 ± 5.60 (52.99)
K Chipsona 2	C	103.0 ± 1.0	103.40 ± 10.9
	15	138.50 ± 2.10 (34.47)	138.70 ± 6.00 (34.13)
K Jyoti	C	106.30 ± 1.00	102.70 ± 1.60
	16	131.70 ± 1.70 (23.89)	129.6 ± 9.8 (26.19)
K Sutlej	C	94.80 ± 3.70	88.00 ± 2.10
	3	148.60 ± 5.80 (56.75)	140.10 ± 7.50 (59.20)
K Badsah	C	89.40 ± 1.90	84.40 ± 7.70
	9	121.20 ± 0.80 (35.57)	115.10 ± 2.30 (36.37)
K Bahar	C	97.60 ± 5.70	98.20 ± 3.80
	5	124.90 ± 2.30 (27.97)	126.90 ± 12.50 (29.22)
K Pukhraj	C	95.00 ± 5.30	86.90 ± 7.20
	1	126.60 ± 7.70 (33.26)	118.00 ± 3.90 (35.79)

Potato plants, in field trials, were grown to maturity and tubers were harvested. Three tubers from each replication for each wild-type and transgenic events were randomly collected. The pooled tubers were divided into three groups, oven-dried at 100 °C, and powdered. Total protein content (mg protein g⁻¹ dry weight) of tubers was determined by standard Kjeldahl method in duplicate from three independent groups and compared. Values are presented as the mean ± SE. Numerical represents the number of transgenic event against its corresponding wild-type background. The values in parentheses represent percent increase in protein in transgenic tubers when compared with that of wild-type tubers (C).

most universally applied to determine total nitrogen content. We determined the amount of total protein on the basis of the nitrogen content in transgenic tubers and compared that with the respective control background. A detailed chemical analysis revealed that the overexpression of AmA1 leads to an increased level of tuber protein compared with wild type. There was a 35–60% increase in total protein content in moderately expressed transgenic events of all varieties (Table 1), which was higher than we had reported earlier (up to 35%) for a diploid cultivar (1). It was expected that the increased levels of protein in modified tuber could have a direct bearing on the amino acid ratio. The balance of amino acid is a key determinant for nutritive proteins because plant proteins typically do not provide the optimum amino acid ratios required for efficient protein synthesis in animals, including humans. Analysis of the pool sizes of the various amino acids revealed a significant increase in amino acids, notably lysine, tyrosine, and sulfur amino acids, which are otherwise limited in potato tubers. Also, there was an increase in aspartic acid, glutamic acid, arginine, leucine, and isoleucine levels (Table S1). These changes were notably more prominent in transgenic events exhibiting a higher level expression of AmA1 (Fig. S1B).

It seems likely that the increase in nitrogen storage in sink tissue may result in alteration of total biomass production. Nitrogen storage and carbohydrate metabolism are closely related. It is postulated that carbohydrate distribution within the plant is affected by the nitrogen supply, which strongly influences the processes of carbon metabolism, and thus nitrogen status has a great impact on the postharvest performance of plants (23). To confirm this, the dry weight of transgenic as well as wild-type plants was determined. Indeed, the dry weight of most transgenic plants significantly increased when compared with their wild-type counterparts. On average, the increases in biomass of most promising transgenic events were found to be in the range of 7–20% (Fig. S2A), in-

dicating a positive correlation between the increased nitrogen storage and the biomass in the transgenic plants.

2D Electrophoresis Analysis Shows Increase in Protein Content. Earlier it was reported that transgene expression may lead to a rebalancing of the proteome toward an increase in protein content (24). To examine the impact of AmA1 toward an increase in protein content, the changes in the tuber proteome of transgenic potatoes were monitored using 2D electrophoresis (2-DE) analysis. The gels showed more than 90% high-quality protein spots and had a correlation coefficient of variation above 0.8 (Fig. S3), suggesting high reproducibility among the replicates (Fig. 1, Table S2). PDQuest analysis showed an average of 323 and 362 spots in wild-type and transgenic events, respectively, substantiating the biochemical findings of increased protein content in transgenic events by Kjeldahl analysis (Table 1).

Field Performance of AmA1 Plants in Different Agro-Climatic Conditions. Field trials of transgenic plants provide a relevant test for understanding the *in planta* biological role of transgenes and proteins and also for verifying their potential use in commercial applications. The planning and execution of these trials are labor intensive but a useful task because agronomic performance of the transgenic crops in the field does not necessarily match the conditions maintained in the greenhouse or experimental plot. To assess the field performance of transgenic plants expressing AmA1 in different agro-climatic conditions, 3–10 independent transgenic events from each genotypic background were selected for further evaluation. The AmA1 plants from size-normalized seed tubers were grown in the field alongside the wild type and randomly replicated three, five, and seven times for four seasons in different field experiments. We selected the best independent transgenic event from each category in which the *AmA1* was stably expressed over generations. In all field trials, determination of tuber yield was carried out on fully senescent plants at about 100 d from the date of plantation. There were no significant changes in the earliness of the crop in any of the transgenic events. However, a moderate increase (≈ 15 – 25%) in the tuber yield was found in a few of the promising transgenic events (Table S3). These results are of great importance, given the twin problems of environmental deterioration and the progressive increase in world population.

Because of the increased tuber yield in the transgenic plants and the fact that photosynthetic carbon metabolism is one of the major determinants for crop growth and yield (25), we measured the photosynthetic efficiency of the transgenic and wild-type plants under standard atmospheric (360 ppm CO_2) and light ($750 \mu\text{mol}\cdot\text{m}^{-2}\cdot\text{s}^{-1}$) conditions. Compared with wild-type plants, pho-

tosynthetic CO_2 fixation of the most promising transgenic plants was up to 27% higher (Fig. S2B). This was corroborated by the corresponding increase in total biomass of the transgenic plants (Fig. S2A). Indeed, these results represent a unique example in which expression of a seed storage protein results in improved carbon fixation and growth in transgenic plants.

AmA1 Potato Tubers Are Nontoxic, Nonallergenic, and Safe for Consumption. In recent years, the use of animal models has become an integral part of the safety evaluation process for genetically modified crops. To examine the toxicity and allergenicity, if any, associated with the AmA1 potato, tubers of transgenic events along with the corresponding wild-type tubers were tested in two different animal models (rat and rabbit) following the recommendation laid down by the Food and Agriculture Organization and WHO (2006) (26). The potato tubers of transgenic events were found to be nontoxic to rats when administered orally as a single dose at a maximum dose level of 5 g of test substance per rat. The given samples were also found to be nonirritating to the skin of the rabbits. Furthermore, the samples of transgenic tubers were found to be nonirritating to the vaginal mucous membrane of the rabbits.

Mean values of hematological and serum biochemical parameters obtained from both male and female rats fed with wild-type and AmA1 potato for 21 and 90 d are presented in Table S4 and Table 2. No differences were observed in a blood biochemistry analysis. Counts of WBC, RBC, and hemoglobin were at same level in both male and female rats of each group. Total plasma protein, sugar, and urea levels were similar for transgenic and control group rats. Serum glutamic pyruvic transaminase (SGPT) and serum glutamic oxaloacetic transaminase (SGOT) levels were comparable in both groups in AmA1 and wild-type potato, indicating normal liver function. Furthermore, the heart and kidney functions of transgenic and wild-type potato-fed rats were also comparable, suggesting normal cellular metabolism.

Histopathological analysis of hematoxylin- and eosin-stained gut tissues of rat-fed AmA1 potato revealed normal structure with no distortion in gut lining. The stomach layer showed no inflammation and/or infiltrations of inflammatory cells into surrounding tissues (Fig. 2A and B). Intestinal villi of AmA1 potato-fed rats showed normal surface area with numerous absorptive enterocytes and mucous-secreting cells and no sign of ulcer, bleeding, or disruption-like symptoms and were comparable to that of the control group (Fig. 2C and D). Furthermore, no abnormal structures were observed in liver and kidney cells of rats fed with wild-type and AmA1 potato (Fig. 2E and F). Transgenic potato-fed rats showed normal lung structure with defined bronchial epithelial lining and alveoli. No histopathological differences were observed in brain and heart tissues (Fig. S4). Thus, under the conditions of this study, long-term oral feeding for 90 d did not show any detectable clinical and histopathological changes or observable toxic effects in any groups of the animals tested.

It is recognized that most allergenic proteins tend to have characteristic sequence stretches. The amino acid sequence of AmA1, however, did not show any homology with known allergenic candidates when searched against the allergen online database (www.allergenonline.com). In a previous study, we had shown that the AmA1 protein in its purified form does not evoke any IgE response, negating the possibility of the protein being allergenic (1). To further confirm whether potato tubers expressing AmA1 are allergenic or not, the hypersensitivity test of the transgenic tubers was carried out. In this experiment, serum from the laboratory animals fed with genetically modified AmA1 tubers was subjected to ELISA study. Although the IgG level was quite high, the IgE level could not be detected, suggesting that the transgenic tubers did not evoke any allergic response in the animals. These results conclusively proved that AmA1 potato tubers are nontoxic and nonallergenic.

Intraperitoneal (i.p.) sensitization with wild-type and AmA1 potato protein produced low IgE titer on day 15, which remained unchanged until day 45 (Fig. 3A), whereas ovalbumin immunization produced high IgE response by the i.p. route. Compared with ovalbumin, i.p. sensitization with both wild-type and AmA1 potato

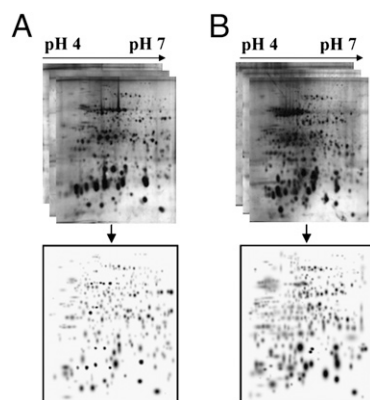


Fig. 1. Proteomic analysis of the mature potato tubers of (A) wild-type (KC1) and (B) the transgenic line (KC1/21). Proteins were extracted from same mass of mature tuber, and an equal volume ($250 \mu\text{L}$) of protein was separated by 2-DE as described in *SI Materials and Methods*. Three replicate silver-stained gels for each stage were computationally combined using PDQuest software, and four representative standard gel images were generated.

Table 2. Serum biochemistry indices of rats fed with control potato and AmA1 potato for 21 and 90 d

Parameter	90 d				21 d			
	Male		Female		Male		Female	
	C1	T1	C1	T1	C1	T1	C1	T1
Blood sugar (mg %)	103.60 ± 6.64	106.60 ± 6.06	103.60 ± 6.60	104.40 ± 7.20	103.60 ± 6.35	101.00 ± 6.20	103.60 ± 6.35	102.40 ± 5.46
BUN (mg %)	27.70 ± 4.15	28.48 ± 4.06	26.15 ± 4.42	26.77 ± 2.82	25.92 ± 3.81	26.98 ± 2.71	20.06 ± 1.57	26.42 ± 2.46
Total protein (g %)	6.546 ± 0.34	6.196 ± 0.19	6.484 ± 0.43	6.218 ± 0.19	6.86 ± 0.25	7.03 ± 0.48	6.90 ± 0.28	7.14 ± 0.22
SGPT (IU)	38.58 ± 7.67	33.96 ± 4.82	35.98 ± 8.05	26.58 ± 4.20	29.90 ± 3.54	31.90 ± 2.94	30.36 ± 1.78	29.46 ± 2.13
SGOT (IU)	39.38 ± 12.83	42.55 ± 13.49	34.12 ± 12.57	37.20 ± 9.30	33.98 ± 2.54	33.74 ± 2.81	33.32 ± 3.44	33.82 ± 2.93
Albumin (g %)	4.02 ± 0.22	4.05 ± 0.25	4.00 ± 0.25	4.02 ± 0.19	4.01 ± 0.25	4.13 ± 0.12	4.10 ± 0.08	4.05 ± 0.17
SAP (U/L)	86.00 ± 8.81	91.37 ± 6.72	88.75 ± 7.96	86.86 ± 7.42	92.80 ± 7.46	88.58 ± 10.73	93.78 ± 6.53	91.56 ± 7.64

BUN, blood urea nitrogen; SAP, serine alkaline phosphatase.

showed a high IgG1 response (Fig. 3B). Both wild-type and AmA1 potato produced a similar IgG2a level on day 15, which increased by day 30 (Fig. 3C). No significant difference ($P > 0.05$) was observed between the IgG1 and IgG2a antibody response induced by wild-type and AmA1 potato proteins.

Simulated Gastric Fluid- and Simulated Intestinal Fluid-Induced Digestibility of AmA1 Potatoes. The potent food allergens are known to be very stable in in vitro pepsin digestion, whereas most dietary proteins are readily digestible (27). The pepsin digestibility assay is thus considered one of the major ways to identify food allergens (28–30). To determine the relative stability of AmA1 to the extremes of pH and pepsin protease encountered in the mammalian digestive tract, the protein extracts of transgenic tubers were subjected to pepsin digestibility. The in vitro digestibility of tuber protein was confirmed by immunodetection assay. The results revealed that the high-molecular-weight proteins, including AmA1, were readily digested within 20 min in simulated gastric fluid (SGF) and also were completely degraded within 15 min in simulated intestinal fluid (SIF). The SGF- and SIF-mediated digestibility altogether suggests that the transgenic potatoes may not be allergenic when consumed as food.

Discussion

A global scientific emphasis has been laid on food security and on improving the nutritional qualities of food crops through contemporary science. Thus, the International Food Policy Research Institute launched 2020 Vision, a call for a new “Green Revolution.” A total of 1.02 billion people, residing mostly in the developing world, suffer from chronic under-nutrition, and 200 million children are affected due to lack of essential energy and protein (31). Therefore, improvement of nutritional value in food crops, particularly improvement in protein content, poses a major challenge in developing countries where plants provide most of the protein in the human diet and in animal feed. The biotechnology-based strategies to modify proteins and/or amino acids have always focused either on increasing the concentration of the direct precursors for biosynthesis of the targeted product (2) or on the concentration of some unique proteins (1, 32, 33). Although many of these strategies have yielded positive results, many others have failed despite changes in some or all of the required metabolic precursors (34). Our approach of using *AmA1* to increase the nutritional quality in tuber crops, particularly in potato, has been successful. We have developed an efficient, rapid, and genotype-independent transformation method for potato. The protocol involves a single medium composition throughout, and transformants were obtained within 6–8 wk. To our knowledge, the increase in protein content of genetically modified tubers in this study is one of the highest increases observed in any transgenic crop. Although the expression level of AmA1 in transgenic tubers was not high enough to be directly correlated with the protein increase, comparative proteome analysis showed a positive correlation with the finding of biochemical analysis. The increase in total protein content as shown by Kjeldahl and 2-DE analyses may

be explained as due to de novo synthesis and accumulation of unique proteins and/or quantitative change in the expression level or redistribution of proteins as a result of AmA1 expression. It is known that synthesis of seed storage protein depletes free amino acid in the storage organ that leads to an increase in the rate of photosynthesis (35). *AmA1*, a seed storage protein gene, when engineered in potato tuber might lead to the depletion of the endogenous free amino acid pool for its synthesis and accumulation. It is likely that the depletion of endogenous free amino acids in the transgenic tuber is then sensed by the photosynthetic machinery, causing an increase in the rate of photosynthesis. Earlier studies have also shown that protein synthesis increases as a consequence of photosynthesis, which is the ultimate basis of yield (35, 36). Thus, the increase in photosynthesis in transgenic potato, as shown in Fig. S2B, in turn enhances the total protein content as well as the yield. Seed storage

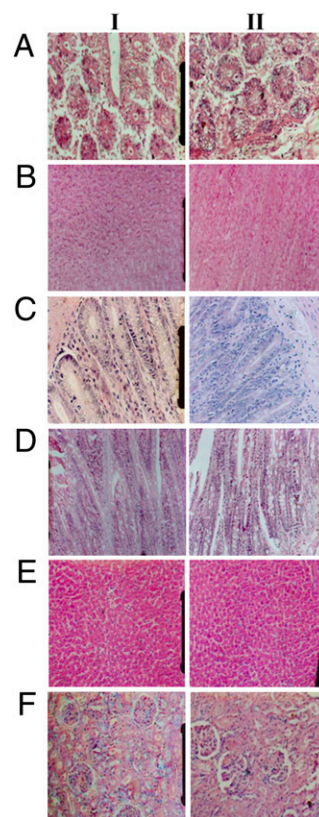


Fig. 2. Histopathological analysis of gut tissues. (I) Rats fed wild-type potato and (II) rats fed AmA1 potato: sections showing hematoxylin and eosin staining of (A) stomach (nonglandular), (B) stomach (glandular), (C) small intestine, (D) large intestine, (E) liver, and (F) kidney.

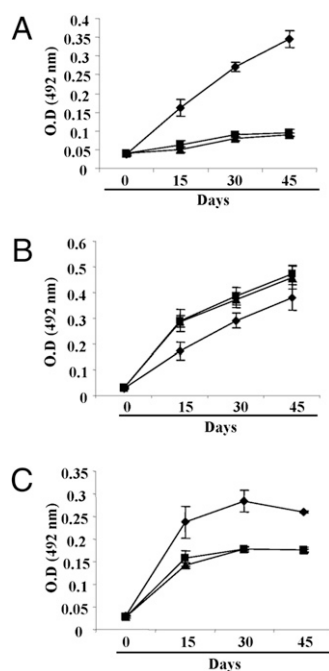


Fig. 3. Specific antibody response in Balb/C mice: serum antibody following i.p. administration of ovalbumin (◆), wild-type (■), and transgenic (▲) potato protein extract. (A) specific IgE, (B) specific IgG1, and (C) specific IgG2a. A preimmune serum was taken as the control, and the values were subtracted before plotting. Data are reported as mean \pm SD.

proteins are known to serve as the sink to regulate the movement of photosynthate into developing organs (35). It is conceivable that AmA1 as a storage protein might act as a sink protein in transgenics, thereby regulating the movement of metabolites, including the amino acids, into the developing tuber where they are fixed into newly synthesized proteins and consequently enhance the level of essential amino acids.

Potatoes are grown as vegetable for consumption and also as raw material for many processed foods. It is well documented that dry matter, specific gravity, and viscosity of the tuber are the major parameters that determine the cooking quality and viscosity and firmness reflect the processing grade of potato (37, 38). To find the cooking quality, processing, and palatability of AmA1 potato, we performed the comparative physio-chemical and rheological analyses of the transgenic and the corresponding wild-type tubers. Specific gravity of most of the transgenic lines was found to be 1.68–6.86% higher (Fig. S5A), indicating better quality of the processed product. With reference to dry matter, KC1/21 and KC2/15 showed the maximum increase, having 23.7 and 19.5%, respectively, as compared with 12.5–18.8% in the transgenic lines of other genetic backgrounds (Fig. S5B). Furthermore, transgenic potatoes upon frying showed less browning (Fig. S6A and B) with higher firmness after boiling (Fig. S6C and D), suggesting better palatability. Proteins are known to be major water absorbers, and they increase viscosity significantly (39). Viscoamylograph analyses revealed that transgenic potatoes reach pick viscosity either more slowly or at the same time in comparison with their respective wild types, maintaining a similar pasting temperature (Fig. S6E and F). These results demonstrate that enhanced protein content in transgenic potatoes neither changed pasting temperature nor increased peak viscosity to affect crispiness and texture of fried potatoes. Altogether, analyses of these variances imply better cooking and processing quality of the transgenic tubers. Our strategy thus allows the achievement of developing nutritionally enhanced potato with quality traits by judicious manipulation of storage protein.

Recently, there have been efforts to express 2S albumin from Brazil nut and sunflower seed albumin to increase the nutritive

value of transgenic crops. However, these attempts were not successful because, when introduced in target plants, the transgene resulted in a dramatic increase only in methionine along with a significant decrease in cysteine (13, 17). Other examples include the soybean glycinin, expression of which in potato tubers did not lead to any remarkable increase either in protein content or in amino acid concentration (40). Furthermore, plans for commercialization of these transgenic crops were abandoned because the donor proteins were allergenic (41). On the contrary, an intensive literature search did not reveal any allergenicity associated with the source of the transgene, amaranth grain or amaranth forage (SI Text). Feeding trials of the AmA1-transgenic tubers in rat models did not show any toxicity. Blood and serum biochemical indices indicated normal functioning of metabolic organs in the test animals. Postmortem examination demonstrated normal appearance of stomach, intestine, liver, kidney, and other organs. Thus, it seems that consumption of AmA1 potato has no detrimental effect on the growth and metabolic function of rat. The hypersensitivity test in animal models also showed that potato tubers expressing AmA1 do not evoke any IgE response, which negates the possibility of the protein being allergenic. The *in vitro* digestibility test showed that the accumulated AmA1 is readily degraded by SGF, confirming that the transgenic tubers are nonallergenic. The findings are significant because the ability of the food allergens to reach the intestinal mucosa is the prerequisite to allergenicity. It is apparent that the ability of food allergens to reach the intestinal mucosa necessarily implies their survival to gastric digestion by pepsin secreted in the stomach. Because the expressed protein is efficiently digested and degraded, the AmA1 protein as a whole would not reach the intestinal mucosa and cause allergenic reaction.

In summary, the tuber-specific expression of *AmA1* was associated with increased protein and a concomitant increase in several of the essential amino acids that are otherwise limited in potato. This study represents a major technological advance in translational research in which the engineering of a seed storage protein has led to nutritional improvement with essentially no negative collateral effects on crop yield or quality. The commercial potential of genetically modified plants depends on stable integration and expression of the transgene under the different genotypic backgrounds of the host species, on their wider environmental applicability, and on sustainable production, including food safety. Therefore, our strategy may prove to be more acceptable to the general public than currently used genetically modified crops because *AmA1* is an edible crop-derived sequence. The benefits are magnified particularly considering that this occurs without any yield penalty, but rather in combination with increased harvestable biomass. Because potato constitutes an important part of the diet of many people in developed as well as developing countries, it is apparent that this can add value to potato-based products with enhanced benefits for better human health. Our strategy also offers unique opportunities for the genetic engineering of unique traits into the next-generation crop to accrue nutritional benefits.

Materials and Methods

For detailed descriptions of materials and methods, see SI Materials and Methods and associated references. Transgenic potato lines were developed and molecular analyses of transgenics were performed as described in SI Text. Expression of AmA1 protein was detected and confirmed at maturity of potato tubers using anti-AmA1 antibody as described in SI Text. Chemical analysis and proteomic study of transgenics were carried out using extracts from wild-type and transgenic tubers as described in SI Text. Field trials, morphological characterization, measurement of photosynthetic activity, digestibility and allergenicity analyses, food safety assessment and toxicity studies, palatability, and cooking and processing quality analyses are also described in SI Text.

ACKNOWLEDGMENTS. We thank Shankar Acharya and C. Ravi Shankar for assistance during the field trials and Jasbeer Singh for illustration and graphical representation. We thank the Shriram Institute of Industrial Research, New Delhi, for support in animal studies and the Director, G.S. Shekhawat and S.M. Paul Khurana, Central Potato Research Institute, Shimla, India, and the Joint Director, B.P. Singh, Central Potato Research Institute

Campus, Modipuram, India, for their continued support throughout the study. This work was supported by the Department of Science and Technology, Department of Biotechnology, Ministry of Science and Technology, the

Government of India, and the National Institute of Plant Genome Research, India. L.A. and, K.N. are the recipients of a predoctoral fellowship from the Council of Scientific and Industrial Research (CSIR), Government of India.

1. Chakraborty S, Chakraborty N, Datta A (2000) Increased nutritive value of transgenic potato by expressing a nonallergenic seed albumin gene from *Amaranthus hypochondriacus*. *Proc Natl Acad Sci USA* 97:3724–3729.
2. Matthews BF, Hughes CA (1993) Nutritional improvement of the aspartate family of amino acids in edible crop plants. *Amino Acids* 4:21–34.
3. Falco SC, et al. (1995) Transgenic canola and soybean seeds with increased lysine. *Biotechnology (NY)* 13:577–582.
4. Stöger E, Parker M, Christou P, Casey R (2001) Pea legumin overexpressed in wheat endosperm assembles into an ordered paracrystalline matrix. *Plant Physiol* 125:1732–1742.
5. Yang P, et al. (2002) Inherent and apparent scattering properties of coated or uncoated spheres embedded in an absorbing host medium. *Appl Opt* 41:2740–2759.
6. Tamás C, et al. (2009) Transgenic approach to improve wheat (*Triticum aestivum* L.) nutritional quality. *Plant Cell Rep* 28:1085–1094.
7. Qaim M, Zilberman D (2003) Yield effects of genetically modified crops in developing countries. *Science* 299:900–902.
8. Mullins E, Milbourne D, Petti C, Doyle-Prestwich BM, Meade C (2006) Potato in the age of biotechnology. *Trends Plant Sci* 11:254–260.
9. Report FAO (2006) Food and Agricultural Organization of the United Nations. Available at <http://www.fao.org/magazine/0611sp1.htm>. Accessed March 8, 2010.
10. United Nations General Assembly Resolution (2005). Available at: <http://www.un.org/depts/dhl/resguide/r60.htm> (Resolution No. A/RES/60/191). Accessed March 8, 2010.
11. FAOSTAT (2005) Food and Agricultural Organization of the United Nations Statistical Database. Available at <http://www.faostat.fao.org>. Accessed March 8, 2010.
12. Jaynes JM, Yang MS, Espinoza N, Dodds JH (1986) Plant protein improvement by genetic engineering: Use of synthetic genes. *Trends Biotechnol* 4:314–320.
13. Altenbach SB, Pearson KW, Meeker G, Staraci LC, Sun SM (1989) Enhancement of the methionine content of seed proteins by the expression of a chimeric gene encoding a methionine-rich protein in transgenic plants. *Plant Mol Biol* 13:513–522.
14. Raina A, Datta A (1992) Molecular cloning of a gene encoding a seed-specific protein with nutritionally balanced amino acid composition from *Amaranthus*. *Proc Natl Acad Sci USA* 89:11774–11778.
15. Zheng Z, Sumi K, Tanaka K, Murai N (1995) The bean seed storage protein [beta]-phaseolin is synthesized, processed, and accumulated in the vacuolar type-II protein bodies of transgenic rice endosperm. *Plant Physiol* 109:777–786.
16. Kohno-Murase J, Murase M, Ichikawa H, Imamura J (1995) Improvement in the quality of seed storage protein by transformation of *Brassica napus* with an antisense gene for cruciferin. *Theor Appl Genet* 91:627–631.
17. Molvig L, et al. (1997) Enhanced methionine levels and increased nutritive value of seeds of transgenic lupins (*Lupinus angustifolius* L.) expressing a sunflower seed albumin gene. *Proc Natl Acad Sci USA* 94:8393–8398.
18. Bellucci M, Lazzari B, Viotti A, Arcioni S (1997) Differential expression of a γ -zein gene in *Medicago sativa*, *Lotus corniculatus* and *Nicotiana tabacum*. *Plant Sci* 127:161–169.
19. Datta A, Raina A, Biswas S (1997) Seed storage protein with nutritionally balanced amino acid composition. US Patent 5,670,635.
20. Datta A, Raina A, Biswas S (1998) Method of making seed specific DNA. US Patent 5,846,736.
21. Chakraborty S, Sarmah B, Chakraborty N, Datta A (2002) Premature termination of RNA polymerase II mediated transcription of a seed protein gene in *Schizosaccharomyces pombe*. *Nucleic Acids Res* 30:2940–2949.
22. Visser RGF, Stolte A, Jacobsen E (1991) Expression of a chimaeric granule-bound starch synthase-GUS gene in transgenic potato plants. *Plant Mol Biol* 17:691–699.
23. Druge U, Zerche S, Kadner R (2004) Nitrogen- and storage-affected carbohydrate partitioning in high-light-adapted *Pelargonium* cuttings in relation to survival and adventitious root formation under low light. *Ann Bot (Lond)* 94:831–842.
24. Schmidt MA, Herman EM (2008) Proteome rebalancing in soybean seeds can be exploited to enhance foreign protein accumulation. *Plant Biotechnol J* 6:832–842.
25. Sweetlove LJ, Kobmann J, Riesmeier JW, Trethewey RN, Hill SA (1998) The control of source to sink carbon flux during tuber development in potato. *Plant J* 15: 697–706.
26. FAO/WHO Food Standard Programme, Codex Alimentarius Commission, 25th Session (2003) *Guideline for the Conduct of Food Safety Assessment of Foods Derived From Recombinant-DNA Plants*, pp 47–60. Available at: ftp://ftp.fao.org/codex/alinorm03/al03_34e.pdf. Accessed March 8, 2010.
27. Asero R, et al. (2000) Lipid transfer protein: A pan-allergen in plant-derived foods that is highly resistant to pepsin digestion. *Int Arch Allergy Immunol* 122:20–32.
28. Astwood JD, Leach JN, Fuchs RL (1996) Stability of food allergens to digestion in vitro. *Nat Biotechnol* 14:1269–1273.
29. Buchanan BB, et al. (1997) Thioredoxin-linked mitigation of allergic responses to wheat. *Proc Natl Acad Sci USA* 94:5372–5377.
30. Thomas K, et al. (2004) A multi-laboratory evaluation of a common in vitro pepsin digestion assay protocol used in assessing the safety of novel proteins. *Regul Toxicol Pharmacol* 39:87–98.
31. FAO (2009) Food and Agricultural Organization Hunger Report. Available at <http://www.fao.org/news/story/en/item/36207/icode/>. Accessed March 8, 2010.
32. Galili G (1995) Regulation of lysine and threonine synthesis. *Plant Cell* 7:899–906.
33. Tabe L, Higgins TJV (1998) Engineering plant protein composition for improved nutrition. *Trends Plant Sci* 3:282–286.
34. Regierer B, et al. (2002) Starch content and yield increase as a result of altering adenylate pools in transgenic plants. *Nat Biotechnol* 20:1256–1260.
35. Moutot F, Huet J-CH, Morot-Gaudry J-F, Pernollet J-C (1986) Relationship between photosynthesis and protein synthesis in maize. I. Kinetics of translocation of the photoassimilated carbon from the ear leaf to the seed. *Plant Physiol* 80:211–215.
36. Zhu X-G, Long SP, Ort DR (2010) Improving photosynthetic efficiency for greater yield. *Annu Rev Plant Biol* 61:235–261.
37. Haase NU (2004) Estimation of dry matter and starch concentration in potatoes by determination of under-water weight and near infrared spectroscopy. *Potato Res* 46: 117–127.
38. Nourian F, Ramaswamy HS, Kushalappa AC (2003) Kinetic changes in cooking quality of potatoes stored at different temperatures. *J Food Eng* 60:257–266.
39. Cuevas RP, Fitzgerald M (2007–08) Linking starch structure to rice cooking quality. *IREC Farmers' News* 177:16–17.
40. Hasimoto W, et al. (1999) Safety assessment of genetically engineered potatoes with designed soybean glycinin: Compositional analyses of the potato tuber and digestibility of the newly expressed protein in transgenic potatoes. *J Sci Food Agric* 79:1607–1612.
41. Nordlee JA, Taylor SL, Townsend JA, Thomas LA, Bush RK (1996) Identification of a Brazil-nut allergen in transgenic soybeans. *N Engl J Med* 334:688–692.