Journal of proteome research

pubs.acs.org/jpr

Article

Proteomic Profiling of Fast Neutron-Induced Soybean Mutant Unveiled Pathways Associated with Increased Seed Protein Content

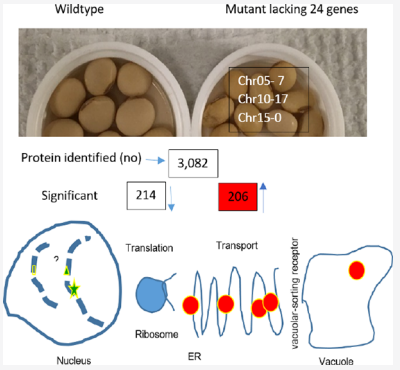
Nazrul Islam, Hari B. Krishnan, and Savithiry Natarajan*

Cite This: *J. Proteome Res.* 2020, 19, 3936−3944

Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: Mutagenesis through fast neutron (FN) radiation of soybean resulted in a mutant with a 15% increase in seed protein content. A comparative genomic hybridization analysis confirmed that the mutant is lacking 24 genes located at chromosomes 5 and 10. A tandem mass tag-based proteomic profiling of the wild type and the FN mutant revealed 3,502 proteins, of which 206 proteins exhibited increased abundance and 214 proteins showed decreased abundance. Among the abundant proteins, basic 7S globulin increased fourfold, followed by vacuolar-sorting receptor and protein transporters. The differentially expressed proteins were mapped on the global metabolic pathways. It was observed that there was an enrichment of 29 ribosomal proteins, 16 endoplasmic reticular proteins, and several proteins in export metabolic pathways. The deletion of the sequence-specific DNA binding transcription factor along with 23 other genes may have altered the negative regulation of protein syntheses processes, resulting in an increase in the overall protein content of the mutant seed. This mutant is a valuable resource for researchers to understand the metabolic pathways that may affect an increase in seed protein content (the mass spectrometry data files were submitted to massive.ucsd.edu # MassIVE MSV000084228).



KEYWORDS: *protein, mass spectrometry, proteome, soybean, pathways, seed*

INTRODUCTION

Soybean (*Glycine max*[L.] Merr) seeds are very rich in protein (~40%) and oil (~20%) content. There is a renewed interest in research to increase the percent of seed proteins in soybean, especially for the food industries that depend on soybean protein as a major source of their raw materials.[1−4] Genetic research with soybean has been augmented by the submission of the reference genome sequence in 2010, which greatly enhanced the scope of exploration of niches to improve soybean seed quality.[4]

Approaches to genetically alter soybean include mutagenesis, for example, transposon tagging,[5] chemical treatments,[6] radiation mutagenesis,[7] and genetic transformation, for example, gene editing.[8,9] However, genetic alteration by radiation mutagenesis, particularly fast neutron (FN) radiation, is preferred over other methods, mainly because of a wide range of variation through deletions, duplications, translocations, and inversions without inserting a foreign gene. In this regard, Bolon et al.,[7] using FN bombardment, produced more than 27,000 unique soybean mutants with various amounts of protein, oil, and carbohydrate (https://www.soybase.org/mutants/about.php). This group also adopted an array-comparative genomic hybridizations (CGHs) along with next-generation sequencing technologies to characterize the genome wide variation in a subset of FN-mutated soybeans.[10] Based on these genetic materials, several targeted investigations have been conducted to improve value added seed quality in soybean such as high-sucrose and low oil seed phenotype[11] and vitamin-E added seed phenotype.[12] A distinct FN-induced sequence rearrangement at an NAP1 gene model associated with stunted trichrome development in soybean has also been reported.[13] Recently, an FN mutant, where approximately 5.4 Mb (from position 15.7−21.1 Mb) of Chr 12 has been deleted, exhibited higher protein content measured by near infrared (NIR) spectroscopy.[14] This investigation has provided a useful information about the chromosomal deletion and increase in total seed protein content. However, it is not known how the deleted chromosomal segment affected overall protein synthesis processes using high-throughput quantitative proteomics analyses. Currently, several hundred FN mutants are available as a community resource for informative genome analyses.[10] To make the best use of the FN mutants, a comprehensive characterization for their biochemical content is essential for further analyses. In our previous studies, we characterized a subset of 10 mutants for their seed composition such as fatty acid methyl esters, oil, and protein profile.[15] In this study, we selected a mutant that lacks 24 genes located at Chr 5 and Chr 10, which produced 15% more protein compared to the wild type. A

Received: March 12, 2020
Published: August 21, 2020

proteomic profiling of this mutant using high throughput tandem mass tag (TMT) analysis will provide new information for functional gene expression to produce value added nutritional seed traits in soybean.

EXPERIMENTAL PROCEDURE

Mutant Materials and Initial Seed Screening

The mutant materials were developed using FN bombardment[7,10] in a project led by Robert Stupar at the University of Minnesota (St. Paul, MN).[15] The information about the doses of radiation and the pertinent procedure were reported in previous publications.[7,10] The seed stock of cultivar "MN1302" was used to develop the soybean line "M92-220".[16] From a large population, a screening of the field-harvested seeds was performed using NIR spectroscopy to analyze seed composition.[7] The mutant, 5R10C28Decfbar241aMN15, was later named M03 and selected for its higher protein content and used for the subsequent analyses along with the wild type (M92-220).

Protein Extraction

Soybean seed protein was extracted as previously reported.[17] Briefly, the ground soybean powder (200 mg) was subjected to defatting by hexane.[18] The proteins were extracted from the residue with 1 mL of a buffer containing sucrose (0.7 M), tris (0.5 M), ethylenediaminetetraacetic acid (50 mM), KCl (0.1 M), dithiothreitol (DTT) (25 mM), and PMSF (2 mM) by mixing for 30 min at room temperature with shaking. This was followed by centrifugation at 8000*g* for 30 min and the supernatant was collected. An equal amount of water-saturated phenol was added to the supernatant and mixed well for 10 min. After centrifugation at 4 °C for 30 min, the phenol phase, which contains the proteins, was collected. A fivefold volume of ammonium acetate (0.1 M) in methanol was added and kept at −20 °C overnight to precipitate the proteins. The protein pellets were collected after centrifugation at 15,000*g* for 30 min and washed with cold acetone (three times) and re-suspended in 6 M urea, 100 mM Tris-HCl. The protein concentration was estimated using bicinchoninic acid assay (Pierce, Rockford, IL).

Protein Digestion and TMT Labeling of Peptides

Protein digestion and peptide labeling were performed as described earlier.[19] After cleaning the protein pellets with methanol/chloroform precipitation, the pellets were suspended in lysis solution (8 M urea, 1% sodium dodecyl sulfate, 50 mM Tris-HCl pH 8.5, protease and phosphatase inhibitors). The micro-BCA method (Pierce, Thermo Fisher Scientific) was used to measure the protein concentration. Reduction and alkylation of the protein were performed with 30 min of 10 mM DTT followed with 30 min of 10 mM iodoacetamide. Lys-C protease (Wako, 129-02541, 2 mg/mL Stock) at a 100:1 protein-to-protease ratio was used to digest the protein by shaking overnight at room temperature, followed by trypsin (Pierce, 90305, 1 mg/mL stock) digestion at a 100:1 protein-to-protease ratio for 6 h at 37 °C. The resultant peptides were purified by reverse phase chromatography and the peptide concentration was measured using the Pierce Quantitative Colorimetric Peptide Assay as per the manufacturer's protocol. About 50 g of digested peptide was used for TMT experiments as per the manufacturer's instructions (Thermo Scientific; Pierce Biotechnology, Germany). For TMT analysis, an equal amount (50 μg) of peptide was labeled with TMT 6-plex reagents as per the manufacturer's instructions (Thermo Scientific; Pierce Biotechnology, Germany). A quality control check was performed using 2 μg of each sample from each 6-plex experiment in 100 μL of 1% formic acid (FA). After following several successful quality checks, the samples were subjected to TMT labeling. A ratio test was performed using an equal amount of peptide samples. The normalized intensities for each peptide sample from the ratio check were mostly similar. The manual adjustments were made before mixing based on the ratio check. The mixed labeled samples were further fractionated. A detailed procedure was described elsewhere.[20] Twelve fractions out of 24 were used for the LC−MS3 data collection strategy.[21] An Orbitrap Fusion mass spectrometer (Thermo Fisher Scientific) equipped with a Proxeon Easy nLC 1000 was used for online sample handling and peptide separations.

Mass Spectrometry and Data Analysis

Approximately 5 μg of peptides was suspended in 5% FA with 5% acetonitrile for further analysis. The samples were loaded onto a fused-silica micro capillary (100 μm inner diameter) with a needle tip pulled to an internal diameter less than 5 μm. The column packing and other procedures were described previously.[22] Briefly, the column was packed to a length of 35 cm with a C_{18} reverse phase resin (GP118 resin 1.8 μm, 120 Å, Sepax Technologies). The peptides were separated using a 180 min linear gradient from 5 to 25% buffer B containing 90% ACN + 0.125% FA equilibrated with buffer A containing 5% ACN + 0.1% FA with 400 nL/min flow rate. For peptide scanning, which leads to subsequent sequence determination, a Fusion Orbitrap with an MS1 spectrum (Orbitrap analysis, resolution 120,000, 350−1350 m/z scan range with quadrupole isolation, AGC target 1×10^6, maximum injection time 100 ms, dynamic exclusion of 90 s) was used. The top 10 fragment ion precursors from MS2 scan were selected for MS3 analysis (synchronous precursor selection), in which precursors were fragmented by HCD prior to Orbitrap analysis (collision energy 55%, max AGC $\times$ 10^5, maximum injection time 110 ms, resolution 50,000).

For protein quantification, an in-house Thermo Scientific Pinpoint Software tools developed by Thermo Center at Harvard Medical School was used. This software builds instrument methods, process data, generates reports, which are further integrated with Proteome Discoverer and Protein Center software to generate the accurate mass of isotopes or MS/MS fragment ions automatically based on the identified peptide sequence. All TMT reporter ion counts across all matching peptide-spectral matches were summed to retrieve the quantitative values after removal of reverse and contaminant proteins. A detailed procedure is reported earlier.[19,23] Briefly, every individual peptide was used for quantitation that contributed toward achieving maximum sufficient TMT reporter ions. In this software, an additional isolation specificity filter was used to minimize peptide co-isolation. The raw file processing (RAW file) was followed by controlling peptide; protein level false discovery rates; assembling proteins from peptides; protein quantification from peptides. The data were interrogated against Phytozome database for *G. max* (william 82.a4.v1, 2018) for MS/MS spectral search with both forward and reverse sequences. Database search criteria were used as follows: tryptic peptide with two missed cleavages, a precursor mass tolerance of 50 ppm, fragment ion mass tolerance of 1.0 Da, static alkylation of cysteine (57.02146 Da), static TMT labeling of lysine residues and N-termini of peptides (229.162932 Da), and variable oxidation of methionine

https://dx.doi.org/10.1021/acs.jproteome.0c00160
J. Proteome Res. 2020, 19, 3936−3944

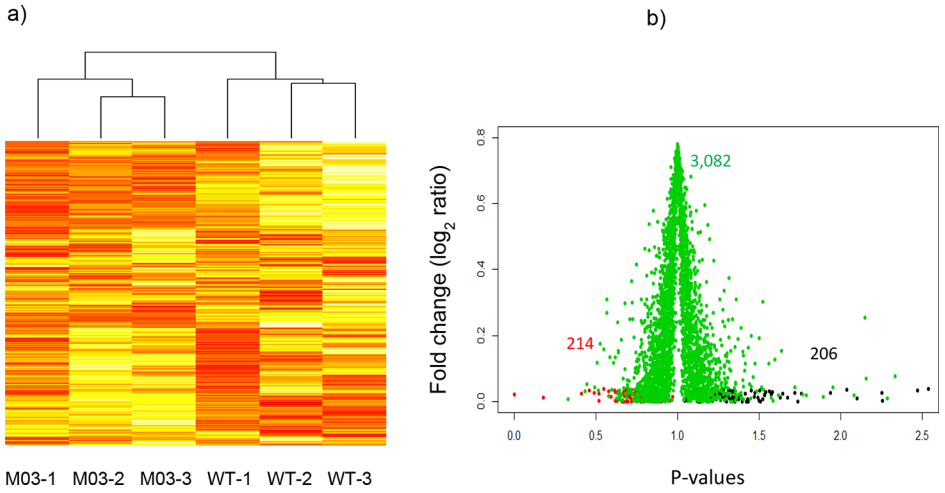


Figure 1. (a) Expression patterns of total number of proteins between replicated mutant (M03-1, M03-2, M03-3) and the wild type (WT-1, WT-2, WT-3). Deep red indicates higher abundance while yellow indicates lower abundance. (b) Matrix plot analyses of the identified proteins: 3082 (green) not significant; 206 (black) significantly higher abundance and 214 (red) significantly lower abundance.

Figure 2. GO terms of the high abundant proteins analyzed by AgriGO (http://bioinfo.cau.edu.cn/agriGO/search.php).

(15.99491 Da). The intensity of TMT reporter ion was calculated with a 0.003 Da window around the theoretical m/z for each reporter ion in the MS3 scan. For quantification, poor quality of MS3 spectra were discarded (<100 summed signal-to-noise across six channels and <0.5 precursor isolation specificity). These mass spectrometry data files were submitted to massive.ucsd.edu (accession # MassIVE MSV000084228).

Peptide spectra were considered for quantitative analyses if the peptide was detected in each of the three independent biological replicates. The protein quantitative values for each channel was initially normalized based on the total spectral count in each channel and the normalized value was then scaled to 100. A *t*-test was performed to identify differentially expressed proteins identified with more than one spectrum and the Benjamini and Hochberg correction was applied to limit false discovery to ≤0.05.

Annotations for the uncharacterized proteins were extracted by submitting the full-length sequences to Kyoto Encyclopedia of Genes and Genomes (KEGG; http://www.genome.jp/kegg) databases and to the MEta Sever for Sequence Analysis

https://dx.doi.org/10.1021/acs.jproteome.0c00160
J. Proteome Res. 2020, 19, 3936−3944

(MESSA) (http://prodata.swmed.edu/MESSA/MESSA.cgi) to determine their functional annotation.[24] With a given protein sequence, MESSA utilizes several tools to predict the local sequence properties such as signal peptides and transmembrane helices. In addition, it predicts homologous proteins along with their functional annotations.

Gene Ontology Term Analyses

The singular enrichment analysis (SEA) of gene ontology (GO) analyses was performed by a web-based tool, AgriGO v2.0 (http://systemsbiology.cau.edu.cn/agriGOv2/).[25] The list of differentially expressed proteins IDs were submitted for SEA analyses against the background of *G. max*. The enrichment GO terms were visualized by hierarchical graph result for cellular component. The statistically significant GO terms were classified into 10 levels, of which level 10 (red) is the highest and level 1 (yellow) is the lowest enrichment.

Mapping the Differentially Expressed Protein in Metabolic Pathways

The NCBI database (https://www.ncbi.nlm.nih.gov/protein/) was used to retrieve the full-length sequences of the differentially expressed proteins. The exported FASTA files were submitted to KEGG Pathways to annotate sequence (http://www.kegg.jp/kegg/tool/annotate_sequence.html) using 3803 as family and the subsequent pathways were visualized for enrichment using KEGG identifier. Proteins/enzymes with no equivalent KEGG identifier could not be mapped on the metabolic pathways.

RESULTS

The seeds of the mutant and the wild-type soybeans were collected from the plants grown in the field (University of Minnesota, St. Paul, MN). The mutant seeds did not exhibit any morphological differences compared to the wild type.[15] The deletion of the chromosomal segment of this mutant was identified by CGH and the deleted genes were further confirmed by PCR analysis as reported earlier.[15] The protein content of the seeds was 55.2% while the wild type, contained 40.8%, as reported in our earlier publication.[15] In this study, we performed TMT based mass spectrometry analyses and quantified 4338 proteins.

The heat map below shows a grouping of protein expression pattern between the wild type and the mutant replicates (Figure 1a). Out of 4,338 proteins detected, 3,502 were identified with more than one peptide and therefore considered for statistical analyses. A matrix plot analyses of identified protein is shown in Figure 1b, where black represents higher abundance, red represents lower abundance, and green represents not statistically significant (Figure 1b). Overall, 206 proteins showed higher abundance (Table S1) and 214 proteins exhibited ($\leq$0.05) lower abundance (Table S1). A list of rest of the proteins with their statistical significance is presented in the supporting tables (Tables S2−S5).

The SEA of GO analyses was performed by a web-based tool, AgriGO v2.0 (http://systemsbiology.cau.edu.cn/agriGOv2/). The list of differentially expressed proteins IDs were submitted for SEA analyses against the background of soybean (*G. max*). The enrichment GO terms were visualized by hierarchical graph result for cellular component. The statistically significant GO terms were classified into 10 levels. Among the 10 levels, red denotes the highest (level 10) and the yellow denotes the lowest (level 1) enrichment.

As shown in Figure 2, the higher GO enrichment of the abundant proteins was observed in ribosome and endoplasmic reticulum (ER). On the other hand, the lower abundant proteins did not show a significant enrichment in any of the cellular components (Figure S1). The enrichment of proteins in ribosome and ER could enhance protein synthesis and processing mechanisms and thus increase the total protein content.

To understand the mechanism of increased protein content in the mutant, we retrieved the KO equivalent of the differentially expressed proteins and mapped on the mRNA surveillance pathway. As evident from Figure 3, three translation initiation factors, eIF3, eIF2 and poly(A)-binding protein (PABP), were significantly increased in the mutant when compared to the wild type.

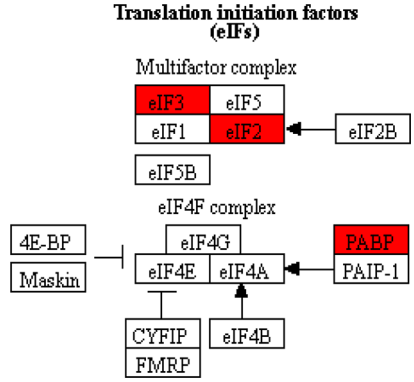


Figure 3. Mapping of the total significant protein on the translation initiation factor pathways in KEGG. Red indicates higher abundant proteins.

The differential expressed proteins were mapped on the ribosome. Protein associated with 29 ribosomal units increased significantly, of which 16 belong to large ribosomal subunits and 13 to the small subunits (Figure S2). A quantitative analysis of these subunits revealed that the large subunit ribosomal protein L34e increased 1.49-fold ($\log_2$), followed by large subunit ribosomal protein LP1 (1.34) and large subunit ribosomal protein L37Ae (1.30) (Table 1).

Similar to the ribosomal proteins, the higher enrichment of four exporters, SEC61a, Bip, SRPRB, and SPCS2 were observed in the protein export pathways. We also checked the enrichment of proteins processing in ER. As demonstrated in Table 2 (also Figure S3), 16 ER proteins exhibited higher abundance and only one protein (p97) showed lower abundance. This protein (p97) is associated with ubiquitin degradation. The higher abundant proteins include protein transporters (SEC61A, SSR2, SSR1, SEC31), sugar transporters (OST1 and SWP1), protein signaling (SEL1 and PRKCSH), protein folding (PDIA6 and CANX), and several heat shock proteins (HSPA5, HSPA5, and HSP90B).

Similar to ribosomal and ER proteins, the higher enrichment of vacuolar proteins was also observed. This includes several vacuolar sorting receptors and transporters (Table 3). Interestingly, among the top five differentially expressed proteins, two were basic 7S globulin.

The segments of deletions were analyzed by jbrowse (https://Phytozome.jgi.doe.gov/jbrowse/index.html) of Phytozome 12 using *G. max* Wm 82 v1.1 to retrieve the corresponding genes. We retrieved 24 genes that were deleted from Chr 5 and Chr 10 and could not find any gene deleted from Chr 15 (Table 4).

https://dx.doi.org/10.1021/acs.jproteome.0c00160
J. Proteome Res. 2020, 19, 3936−3944

Table 1. List of Higher Abundant Proteins in the Ribosome

accession	$\log_2$ ratio	*p*-value	protein name
NP.001347355.1	1.49	0.00913	RP-L34e;large subunit
XP.003526758.1	1.34	0.00318	RP-LP1;large subunit
XP.006573704.1	1.30	0.01329	RP-L37Ae;large subunit
NP.001240945.1	1.26	0.02990	RP-L12e;large subunit
NP.001238194.1	1.22	0.02308	RP-S7;small subunit
NP.001242462.1	1.22	0.03859	RP-SAe;small subunit
NP.001235812.2	1.19	0.00238	RP-S7e;small subunit
NP.001236040.1	1.19	0.02559	RP-L23Ae;large subunit
XP.003543439.2	1.18	0.01206	RP-LP2;large subunit
NP.001236502.1	1.18	0.03222	RP-L9e;large subunit
NP.001237243.2	1.17	0.02383	RP-S6e;small subunit
XP.003536811.1	1.17	0.01733	RP-L30e;large subunit
XP.006594498.1	1.16	0.01001	RP-S14e;small subunit
NP.001238278.1	1.16	0.00525	RP-L32e;large subunit
NP.001238100.1	1.16	0.02136	RP-S18e;small subunit
NP.001240176.1	1.14	0.03887	RP-S24e;small subunit
NP.001236216.2	1.13	0.02042	RP-S19e;small subunit
XP.003530295.1	1.13	0.01661	RP-S12e;small subunit
XP.003529684.1	1.13	0.01068	RP-S15e;small subunit
XP.003535989.1	1.13	0.01410	RP-S3e;small subunit
XP.003537257.1	1.12	0.03887	RP-S23e;small subunit
NP.001237255.1	1.12	0.01807	RP-L37e;large subunit
XP.003537248.1	1.11	0.00133	RP-L4e;large subunit
NP.001235068.1	1.10	0.01569	RP-L24e;large subunit
NP.001235533.1	1.10	0.01874	RP-S16e;small subunit
NP.001238740.2	1.08	0.02654	RP-S9e;small subunit
XP.003536417.1	1.08	0.00753	RP-L3e;large subunit
XP.006582827.1	1.07	0.03405	RP-L7e;large subunit
XP.003521308.1	1.07	0.03605	RP-L23e;large subunit
XP.003531105.1	1.07	0.03139	RP-L27e;large subunit

Table 2. List of Higher Abundant Proteins in the ER

accession	$\log_2$ ratio	*p*-value	protein name
XP.003527201.1	1.42	0.00743	SEC61A; protein transport protein
XP.003530964.2	1.39	0.02818	SEL1; signal transduction protein
XP.003536348.1	1.28	0.02164	OST1; oligosaccharyltransferase
NP.001238351.1	1.26	0.01520	SSR2; translocon protein subunit β
XP.003537760.3	1.20	0.03209	SSR1; translocon protein subunit α
XP.006583948.1	1.19	0.02886	PRKCSH; protein kinase
XP.003551850.1	1.17	0.01897	RAD23; UV excision repair protein
NP.001236300.2	1.15	0.00904	PDIA6; protein disulfide-isomerase A6
XP.006584879.1	1.13	0.01115	UFD1; ubiquitin fusion degradation protein 1
XP.003554291.1	1.13	0.01826	SWP1; oligosaccharyltransferase
XP.003534381.1	1.13	0.02022	SEC31; protein transport protein
NP.001238736.2	1.13	0.02373	HSPA5; heat shock 70 kDa
XP.003545030.1	1.12	0.02343	HSP90B; heat shock protein
NP.001242208.1	1.11	0.02737	EIF2S1; translation initiation factor 2
XP.003525166.1	1.10	0.03093	CANX; calnexin

As per GO terms analyses (http://bioinfo.cau.edu.cn/agriGO/search.php), among the 24 genes predicted, six were involved in a membrane-associated cellular component. These are CPK; calcium-dependent protein kinase, ATP-binding cassette family G25 (XP_003535544.1), plasma membrane intrinsic protein 1; 4 (Glyma.05G208700), sucrose-proton symporter 2, SLC45A1_2_4; solute carrier family 45, member 1/2/4(NP_001335852.1), FUT13; α-1,4-fucosyltransferase

Table 3. List of Differentially Expressed Proteins Located in the Vacuoles

accessions	$\log_2$ ratio	*p*-value	protein name
GLYMA_08G188000, XP_006585494.1	2.255	0.0001	Tcp11 peptidase_S46
GLYMA_20G210100, XP_003556387.1	2.253	0.0237	basic 7S globulin
GLYMA_02G148200, XP_003518926.1	1.675	0.0117	basic 7S globulin
GLYMA_01G242800, XP_003517598.2	1.428	0.0029	vacuolar-sorting receptor 1
GLYMA_06G227200, XP_003527201.1	1.42	0.0074	SEC61A; transport protein
GLYMA_03G040000, XP_003522128.1	1.372	0.0097	non-specific lipid-transfer protein
GLYMA_10G210300, XP_003536348.1	1.282	0.0216	OST1; oligosaccharyltransferase
GLYMA_02G204000, XP_003518220.1	1.261	0.0323	CARP; leucyl aminopeptidase
GLYMA_20G133800, XP_003555992.1	1.218	0.0186	vacuolar-sorting receptor
GLYMA_10G257300, XP_003536576.1	1.156	0.0361	vacuolar-sorting receptor
GLYMA_19G249900, XP_006604865.1	1.118	0.0162	protein ILITYHIA
XP_003536525.1	1.056	0.0038	trichohyalin
GLYMA_18G213400, XP_003552313.1	1.053	0.0228	vacuolar-sorting receptor

[EC:2.4.1.65] (XP_003536391.1), CPK; calcium-dependent protein kinase [EC:2.7.11.1] (XP_006589459.1). WoLF PSORT (https://wolfpsort.hgc.jp/) analyses have revealed the localization of five proteins in the vacuoles and nine in the nucleus. The vacuolar proteins are plasma membrane intrinsic protein, cytochrome P450, ATP-binding cassette family G25, PDIA1; protein disulfide-isomerase A1, and sucrose-proton symporter. The nuclear proteins that are deleted include acid−amino acid ligases, hydroxyproline-rich glycoprotein, exocyst complex component 2, protein tyrosine phosphatase, and a 1-phosphatidylinositol-3-phosphate 5-kinase. It is apparent that a wide range of proteins associated with membrane trafficking, vacuolar and nuclear associated proteins were deleted from the region of Chr 5 and Chr 10. The nuclear proteins that are deleted from Chr 10 also include a sequence-specific DNA binding transcription factors; transcription regulators (NP_001236272.1).

DISCUSSION

In our earlier publication, we showed that the FN mutant (line M03), lacking 24 genes, exhibited high protein content.[15] The GO term analyses of the significantly more abundant proteins demonstrate that some ribosomes and ER proteins were significantly enriched (Figure 2).

Protein synthesis is a complex process mediated by several organelles and molecules. This includes, but is not limited to, the ribosome, mRNA, tRNAs, translation initiation, and elongation and release factors. To begin the translation processes, an initiation factor, methionyl tRNA, is transferred to the ribosome to form a complex called a ternary complex, which then binds to the translation initiation factor 2 (eIF2). A cascade of binding of several initiation factors including eIF3, eIF4, and eIF5 ensure the formation of a 43S preinitiation complex.[26] The initiation factors also serve as a means of preparing the small ribosomal subunit for binding mRNA in coordination with eIF4F to produce the eIF4E. eIF4G, a derivative of eIF4E, then binds with

https://dx.doi.org/10.1021/acs.jproteome.0c00160
J. Proteome Res. 2020, 19, 3936−3944

Table 4. Name of the Proteins Deleted from the M03 because of FN Radiation

accession	protein name
XP_003524456.1/Glyma.05G208600	acid-amino acid ligases; ligases; ATP binding; ATP binding; ligases
Glyma_05G208700/Glyma.05G208700	plasma membrane intrinsic protein
XP_003525392.1/Glyma.05G208800	NAD(P)-binding Rossmann-fold superfamily protein
XP_003525390.1/Glyma.05G208900	cytochrome P450, family 86, subfamily A, polypeptide 8
XP_003525389.2/Glyma.05G209000	BTB/POZ domain with WD40/YVTN repeat-like protein
XP_003525388.1/Glyma.05G209100	hydroxyproline-rich glycoprotein family protein
XP_006580571.1/Glyma.05G209200	peptidase M1 family protein
Glyma.10G216800/Glyma.10G216800	UbiA prenyltransferase family
NP_001335884.1/Glyma.10G216900	CPK; calcium-dependent protein kinase [EC:2.7.11.1]
XP_006589459.1/Glyma.10G217000	CPK; calcium-dependent protein kinase [EC:2.7.11.1]
NP_001236272.1/Glyma.10G217100	sequence-specific DNA binding transcription factors; transcription regulators
XP_006589460.1/Glyma.10G217200	EXOC2; exocyst complexcomponent 2
XP_003535544.1/Glyma.10G217300	ATP-binding casette family G25
NP_001237721.1/Glyma.10G217400	deoxynucleoside triphosphate triphosphohydrolase
XP_003536389.1/Glyma.10G217500	RNF13; E3 ubiquitin-protein ligase RNF13 [EC:2.3.2.27]
XP_003536390.1/Glyma.10G217600	PDIA1; protein disulfide-isomerase A1 [EC:5.3.4.1]
NP_001236979.2/Glyma.10G217700	cca; tRNA nucleotidyltransferase (CCA-adding enzyme)
XP_003536391.1/Glyma.10G217800	FUT13; α-1,4-fucosyltransferase [EC:2.4.1.65]
NP_001335852.1/Glyma.10G217900	sucrose-proton symporter 2, SLC45A1_2_4; solute carrier family 45, member 1/2/4
XP_006589463.1/Glyma.10G218000	protein of unknown function (DUF668)
XP_006588511.1/Glyma.10G218100	protein tyrosine phosphatase 1
XP_003556187.1/Glyma.10G218200	exostosin family protein
XP_003536392.1/Glyma.10G218300	TRM112; multifunctional methyltransferase subunit TRM112
XP_006589464.1/Glyma.10G218400	PIKFYVE; 1-phosphatidylinositol-3-phosphate 5-kinase [EC:2.7.1.150]

a PABP. These translation factors (TFs) are considered a major driving force in mRNA recruitment and accommodation.[27−29] Aligned with this process, we discovered that three translation initiation factors, eIF3, eIF2 and PABP, exhibited significantly higher abundance in the M03 mutant, indicating that the deleted genes might have a regulatory effect on processes involved in protein translation.

As mentioned above, one of the first steps toward formation of a fully functional 80× ribosomal complex is the binding of eIF3 into the small ribosomal subunit (40s).[30,31] In our study, 13 of the 29 ribosomal proteins that exhibited significantly higher abundance were part of the small ribosomal subunit. The interactions among three of the small ribosomal subunit, RP-S15e, RP-S3e, RP-S9e, are reported elsewhere.[26,32−35] While the small ribosomal subunit proteins help to stabilize the mRNA, the processes of decoding the mRNA start with the large subunit (60s ribosomal subunit). In our study, a quantitative analysis of the large subunit ribosomal proteins revealed an increased abundance in several large subunit proteins, RP-L34e, RP-LP, RP-L37Ae, RPL-12e, RPL-23e, RPL-P2, RPL-9e, RPL-30e, RPL-32e, and others. A functional association among the RP-L34e, RPL-30e, and RPL-32e in several organisms is reported in the STRING database (https://string-db.org/network).

Through the processes of mRNA translation, the mRNA is decoded to produce polypeptides/proteins, a major portion of which undergoes further processing through the ER. Most of the water-soluble proteins simply diffuse through the ER for further processing; however, the hydrophobic proteins need a special process of recognition and integration into the lipid bilayer of the ER through a translocon apparatus.[36] In this process, signaling protein (SP) recognizes the presence of a transmembrane domain in the secreted protein. Eventually the SP binds with the secreted protein and releases them in the ER.[37−39] In our study, we identified a SP, SEL1, exhibited $\log_2$ 1.39-fold higher abundance in the mutant. This protein could have similar roles of SP as reported earlier.[36] In addition, we observed higher abundance of several translocon/transporters: SEC61A (SecY (α), SSR2, SSR1, SEC31, OST1, and SWP1). Among the translocons, SecY (α) exhibited the highest abundance ($\log_2$ 1.42). The SEC61 is a heterotrimeric complex consisting of SecY (α), SecE (Y), and SecG (β) subunits. The largest subunit of this complex is the SecY (α), which is considered as the main pore-forming channel in the ER for transporting transmembrane proteins.[40−42] The higher abundance of these proteins, in our study, implies that the gene deletion has a cascaded effect on the operation of SecY (α) apparatus that could lead to processing of more trans-membrane proteins into the vacuoles.

The third most abundant protein in our study is the oligosaccharyltransferase complex, which exhibited $\log_2$ 1.28-fold increase (2.4-fold increase) in the mutant. Also identified in the study was SWP1, which is a form of the above enzyme but having a $\log_2$ 1.13-fold abundance in the mutant. These enzymes are located in the ER and act as key members of N-linked glycosylation.[43] The higher abundance of these proteins indicate that glycan associated protein modification was enhanced because of the gene deletion in the mutant.

Once the targeted proteins are transported through the ER membrane to the ER lumen, they undergo several processing steps such as protein folding, assembly of multi-subunit proteins, disulfide bond formation, and others. A pre-requisite to initiate this process is the formation of disulfide bond, mostly accomplished by isomerase. Consistent with this process, we also identified an isomerase (PDIA6; protein disulfide-isomerase A6) that exhibited $\log_2$ 1.15-fold increase in the mutant. We, therefore, anticipate that the deleted genes might have an inhibitory effect on the overall processes of protein synthesis network, starting from translation to the processing by the ER. After processing by ER, the proteins are distributed to the different parts of the cells or organelles to carry out their

https://dx.doi.org/10.1021/acs.jproteome.0c00160
J. Proteome Res. 2020, 19, 3936−3944

functions. The unused proteins are directed to the vacuoles as a reserve for the future uses. Similar to the process, we also identified several vacuolar proteins, receptors, and storage (7S globulin) that exhibited significant abundance when compared to the wild type (Table 3).

Although it is not known how the gene deletion contributed to the higher protein content through the protein synthesis pathways, it is likely that the deleted genes might have some inhibitory influence throughout the protein synthetic processes. The question might also arise whether the deletion of gene in soybean (paleopolyploid) is not a complete deletion; it might be compensated by the paralog-gene. To confirm this possibility, we performed a BLAST search of the deleted proteins against reference sequence of *G. max* to find the nearest paralog of the deleted genes. We could not find a paralog of the genes listed in Table 4 ,which shows that the effect is not from the paralog genes. Interestingly, most of the nearest paralog of Chr 10 deleted genes are located on the Chr 20,[44] indicating a synteny between these two chromosomes. The synteny between Chr 10 and Chr 20 are well documented from evolution perspective of paleopolyploid soybean from the diploid progeny.[45] As indicated, although parts of the Chr 10 and Chr 20 exhibited synteny, most of the quantitative trait loci (QTL) for seed proteins are located on the Chr 20, which could be related to genome duplications that occurred two times. The genome duplications led to the gene diversification and loss, and, in some cases, numerous chromosome rearrangements. Therefore, we anticipate that, the Chr 10 might have some inhibitory gene/s that control the protein QTL on the Chr 20. Similar to our studies, deletion of the 137 putative genes by FN in Chr 12 elevated total seed protein content in soybeans.[14] The authors concluded that the deleted genes included several enzymes of catabolic processes, transcription factors, and transporters.

In our study, among the 24 genes deleted, a nuclear protein that is deleted from Chr 10 also includes a sequence-specific DNA binding transcription factor, transcription regulator (NP_001236272.1). This protein has been annotated as SOYBN BHLH transcription factor. Interestingly, the gene encoding this protein or its paralog could not be found in Chr 20, probably because of its negative regulation to the other protein TFs; it was eliminated in Chr 20. While the direct effect of this TF is not known, our results showed that the deletion of this regulator along with other 23 genes significantly increased total seed protein contents. Although, the functions of the transcription factor in soybean are not known, it has a basic domain helix-loop-helix (HLH) that might inhibit other transcription factors. The inhibitory effect of this TF is not known in plants, but a mouse twist protein containing the basic HLH is reported as a negative regulator of the MEF2 transcription factor and myogenic regulatory factors.[46]

To understand the relationship of protein abundance and gene expression, we performed RT-qPCR using matured seeds. However, we could not conclude the relationship because, in the dry dormant seeds, the concentration of mRNA is very low and, in addition, the turnover of mRNA is very minimal.

CONCLUSIONS

In this study, we selected one of the promising FN mutant line that showed approximately 15% increase in soybean seed proteins to unveil the mechanism of protein increase. A higher enrichment of protein synthesis pathways starting from translation initiation factors to the storage vacuole was observed. We anticipate that in the deletion of the sequence-specific DNA binding transcription factor, three kinase genes along with other genes might have a cascaded effect of protein synthesis, resulting in an overall increase in seed protein content.

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.jproteome.0c00160.

Total proteins identified with at least one significant spectrum and their relative abundance (XLSX)

Total proteins identified with at least one significant spectrum and their normalized abundance (XLSX)

Number of proteins that were upregulated along with their peptide spectrum matches and their p-values (XLSX)

Number of proteins that showed significantly lower abundance along with their peptide spectrum matches and p-values (XLSX)

Number of proteins that showed no significant difference along with their peptide spectrum matches and p-values (XLSX)

GO terms of the lower abundant proteins analyzed by AgriGO (http://bioinfo.cau.edu.cn/agriGO/search.php); proteins associated with 29 ribosomal units; and mapping of ER proteins (PDF)

AUTHOR INFORMATION

Corresponding Author

Savithiry Natarajan − *Soybean Genomics and Improvement Laboratory, USDA-ARS, Beltsville, Maryland 20705, United States;* orcid.org/0000-0002-6969-9791; Phone: 301-504-5258; Email: savi.natarajan@ars.usda.gov; Fax: 301-504-5728

Authors

Nazrul Islam − *Soybean Genomics and Improvement Laboratory, USDA-ARS, Beltsville, Maryland 20705, United States*

Hari B. Krishnan − *Plant Genetics Research Unit, USDA-ARS, University of Missouri, Columbia, Missouri 65211, United States*

Complete contact information is available at: https://pubs.acs.org/10.1021/acs.jproteome.0c00160

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

The authors would like to thank Thermo Center at Harvard Medical School for performing the quantitative protein analyses and Prof. Robert Stupar, University of Minnesota, for providing the mutant seeds. The authors would also like to thank Dr. Mark Tucker for critically reviewing the manuscript. Funding for this research was provided by Agricultural Research Service, USDA. Mention of trade name, proprietary product, or vendor does not constitute a guarantee or warranty of the product by the U.S. Department of Agriculture or imply its approval to the exclusion of other products or vendors that also may be suitable.

REFERENCES

(1) Flores, T.; Karpova, O.; Su, X.; Zeng, P.; Bilyeu, K.; Sleper, D. A.; Nguyen, H. T.; Zhang, Z. J. Silencing of GmFAD3 gene by siRNA leads

to low alpha-linolenic acids (18:3) of fad3-mutant phenotype in soybean [Glycine max (Merr.)]. *Transgenic Res.* **2008**, *17*, 839−850.
(2) Chennareddy, S.; Cicak, T.; Clark, L.; Russell, S.; Skokut, M.; Beringer, J.; Yang, X.; Jia, Y.; Gupta, M. Expression of a novel bi-directional Brassica napus promoter in soybean. *Transgenic Res.* **2017**, *26*, 727−738.
(3) Demorest, Z. L.; Coffman, A.; Baltes, N. J.; Stoddard, T. J.; Clasen, B. M.; Luo, S.; Retterath, A.; Yabandith, A.; Gamo, M. E.; Bissen, J.; Mathis, L.; Voytas, D. F.; Zhang, F. Direct stacking of sequence-specific nuclease-induced mutations to produce high oleic and low linolenic soybean oil. *BMC Plant Biol.* **2016**, *16*, 225.
(4) Schmutz, J.; Cannon, S. B.; Schlueter, J.; Ma, J.; Mitros, T.; Nelson, W.; Hyten, D. L.; Song, Q.; Thelen, J. J.; Cheng, J.; Xu, D.; Hellsten, U.; May, G. D.; Yu, Y.; Sakurai, T.; Umezawa, T.; Bhattacharyya, M. K.; Sandhu, D.; Valliyodan, B.; Lindquist, E.; Peto, M.; Grant, D.; Shu, S.; Goodstein, D.; Barry, K.; Futrell-Griggs, M.; Abernathy, B.; Du, J.; Tian, Z.; Zhu, L.; Gill, N.; Joshi, T.; Libault, M.; Sethuraman, A.; Zhang, X.-C.; Shinozaki, K.; Nguyen, H. T.; Wing, R. A.; Cregan, P.; Specht, J.; Grimwood, J.; Rokhsar, D.; Stacey, G.; Shoemaker, R. C.; Jackson, S. A. Genome sequence of the palaeopolyploid soybean. *Nature* **2010**, *463*, 178−183.
(5) Mathieu, M.; Winters, E. K.; Kong, F.; Wan, J.; Wang, S.; Eckert, H.; Luth, D.; Paz, M.; Donovan, C.; Zhang, Z.; Somers, D.; Wang, K.; Nguyen, H.; Shoemaker, R. C.; Stacey, G.; Clemente, T. Establishment of a soybean (Glycine max Merr. L) transposon-based mutagenesis repository. *Planta* **2009**, *229*, 279−289.
(6) Li, Z.; Jiang, L.; Ma, Y.; Wei, Z.; Hong, H.; Liu, Z.; Lei, J.; Liu, Y.; Guan, R.; Guo, Y.; Jin, L.; Zhang, L.; Li, Y.; Ren, Y.; He, W.; Liu, M.; Htwe, N. M. P. S.; Liu, L.; Guo, B.; Song, J.; Tan, B.; Liu, G.; Li, M.; Zhang, X.; Liu, B.; Shi, X.; Han, S.; Hua, S.; Zhou, F.; Yu, L.; Li, Y.; Wang, S.; Wang, J.; Chang, R.; Qiu, L. Development and utilization of a new chemically-induced soybean library with a high mutation density. *J. Integr. Plant Biol.* **2017**, *59*, 60−74.
(7) Bolon, Y.-T.; Haun, W. J.; Xu, W. W.; Grant, D.; Stacey, M. G.; Nelson, R. T.; Gerhardt, D. J.; Jeddeloh, J. A.; Stacey, G.; Muehlbauer, G. J.; Orf, J. H.; Naeve, S. L.; Stupar, R. M.; Vance, C. P. Phenotypic and genomic analyses of a fast neutron mutant population resource in soybean. *Plant Physiol.* **2011**, *156*, 240−253.
(8) Demorest, Z. L.; Coffman, A.; Baltes, N. J.; Stoddard, T. J.; Clasen, B. M.; Luo, S.; Retterath, A.; Yabandith, A.; Gamo, M. E.; Bissen, J.; Mathis, L.; Voytas, D. F.; Zhang, F. Direct stacking of sequence-specific nuclease-induced mutations to produce high oleic and low linolenic soybean oil. *BMC Plant Biol.* **2016**, *16*, 225.
(9) Haun, W.; Coffman, A.; Clasen, B. M.; Demorest, Z. L.; Lowy, A.; Ray, E.; Retterath, A.; Stoddard, T.; Juillerat, A.; Cedrone, F.; Mathis, L.; Voytas, D. F.; Zhang, F. Improved soybean oil quality by targeted mutagenesis of the fatty acid desaturase 2 gene family. *Plant Biotechnol. J.* **2014**, *12*, 934−940.
(10) Bolon, Y.-T.; Stec, A. O.; Michno, J.-M.; Roessler, J.; Bhaskar, P. B.; Ries, L.; Dobbels, A. A.; Campbell, B. W.; Young, N. P.; Anderson, J. E.; Grant, D. M.; Orf, J. H.; Naeve, S. L.; Muehlbauer, G. J.; Vance, C. P.; Stupar, R. M. Genome resilience and prevalence of segmental duplications following fast neutron irradiation of soybean. *Genetics* **2014**, *198*, 967−981.
(11) Dobbels, A. A.; Michno, J.-M.; Campbell, B. W.; Virdi, K. S.; Stec, A. O.; Muehlbauer, G. J.; Naeve, S. L.; Stupar, R. M. An Induced Chromosomal Translocation in Soybean Disrupts a KASI Ortholog and Is Associated with a High-Sucrose and Low-Oil Seed Phenotype. *G3: Genes, Genomes, Genet.* **2017**, *7*, 1215−1223.
(12) Stacey, M. G.; Cahoon, R. E.; Nguyen, H. T.; Cui, Y.; Sato, S.; Nguyen, C. T.; Phoka, N.; Clark, K. M.; Liang, Y.; Forrester, J.; Batek, J.; Do, P. T.; Sleper, D. A.; Clemente, T. E.; Cahoon, E. B.; Stacey, G. Identification of Homogentisate Dioxygenase as a Target for Vitamin E Biofortification in Oilseeds. *Plant Physiol.* **2016**, *172*, 1506−1518.
(13) Campbell, B. W.; Hofstad, A. N.; Sreekanta, S.; Fu, F.; Kono, T. J. Y.; O'Rourke, J. A.; Vance, C. P.; Muehlbauer, G. J.; Stupar, R. M. Fast neutron-induced structural rearrangements at a soybean NAP1 locus result in gnarled trichomes. *Theor. Appl. Genet.* **2016**, *129*, 1725−1738.
(14) Prenger, E. M.; Ostezan, A.; Mian, M. A. R.; Stupar, R. M.; Glenn, T.; Li, Z. Identification and characterization of a fast-neutron-induced mutant with elevated seed protein content in soybean. *Theor. Appl. Genet.* **2019**, *132*, 2965−2983.
(15) Islam, N.; Stupar, R. M.; Qijian, S.; Luthria, D. L.; Garrett, W.; Stec, A. O.; Roessler, J.; Natarajan, S. S. Genomic changes and biochemical alterations of seed protein and oil content in a subset of fast neutron induced soybean mutants. *BMC Plant Biol.* **2019**, *19*, 420.
(16) Orf, J. H.; Denny, R. L. Registration of "MN1302" Soybean. *Crop Sci.* **2004**, *44*, 693.
(17) Singh, N.; Jain, N.; Kumar, R.; Jain, A.; Singh, N. K.; Rai, V. A comparative method for protein extraction and 2-D gel electrophoresis from different tissues of Cajanus cajan. *Front. Plant Sci.* **2015**, *6*, 606.
(18) Hill, R. C.; Oman, T. J.; Wang, X.; Shan, G.; Schafer, B.; Herman, R. A.; Tobias, R.; Shippar, J.; Malayappan, B.; Sheng, L.; Xu, A.; Bradshaw, J. Development, Validation, and Interlaboratory Evaluation of a Quantitative Multiplexing Method To Assess Levels of Ten Endogenous Allergens in Soybean Seed and Its Application to Field Trials Spanning Three Growing Seasons. *J. Agric. Food Chem.* **2017**, *65*, 5531−5544.
(19) Weekes, M. P.; Tomasec, P.; Huttlin, E. L.; Fielding, C. A.; Nusinow, D.; Stanton, R. J.; Wang, E. C. Y.; Aicheler, R.; Murrell, I.; Wilkinson, G. W. G.; Lehner, P. J.; Gygi, S. P. Quantitative Temporal Viromics: An Approach to Investigate Host-Pathogen Interaction. *Cell* **2014**, *157*, 1460−1472.
(20) Nusinow, D. P.; Szpyt, J.; Ghandi, M.; Rose, C. M.; McDonald, E. R., 3rd; Kalocsay, M.; Jané-Valbuena, J.; Gelfand, E.; Schweppe, D. K.; Jedrychowski, M.; Golji, J.; Porter, D. A.; Rejtar, T.; Wang, Y. K.; Kryukov, G. V.; Stegmeier, F.; Erickson, B. K.; Garraway, L. A.; Sellers, W. R.; Gygi, S. P. Quantitative Proteomics of the Cancer Cell Line Encyclopedia. *Cell* **2020**, *180*, 387−402 e16.
(21) McAlister, G. C.; Nusinow, D. P.; Jedrychowski, M. P.; Wühr, M.; Huttlin, E. L.; Erickson, B. K.; Rad, R.; Haas, W.; Gygi, S. P. MultiNotch MS3 Enables Accurate, Sensitive, and Multiplexed Detection of Differential Expression across Cancer Cell Line Proteomes. *Anal. Chem.* **2014**, *86*, 7150−7158.
(22) Islam, N.; Bates, P. D.; Maria John, K. M.; Krishnan, H. B.; Z, J. Z.; Luthria, D. L.; Natarajan, S. S. Quantitative Proteomic Analysis of Low Linolenic Acid Transgenic Soybean Reveals Perturbations of Fatty Acid Metabolic Pathways. *Proteomics* **2019**, *19*, 1800379.
(23) Pease, B. N.; Huttlin, E. L.; Jedrychowski, M. P.; Talevich, E.; Harmon, J.; Dillman, T.; Kannan, N.; Doerig, C.; Chakrabarti, R.; Gygi, S. P.; Chakrabarti, D. Global analysis of protein expression and phosphorylation of three stages of Plasmodium falciparum intra-erythrocytic development. *J. Proteome Res.* **2013**, *12*, 4028−4045.
(24) Cong, Q.; Grishin, N. V. MESSA: MEta-Server for protein Sequence Analysis. *BMC Biol.* **2012**, *10*, 82.
(25) Tian, T.; Liu, Y.; Yan, H.; You, Q.; Yi, X.; Du, Z.; Xu, W.; Su, Z. agriGO v2.0: a GO analysis toolkit for the agricultural community, 2017 update. *Nucleic Acids Res.* **2017**, *45*, W122−W129.
(26) Valášek, L. S.; Zeman, J.; Wagner, S.; Beznosková, P.; Pavlíková, Z.; Mohammad, M. P.; Hronová, V.; Herrmannová, A.; Hashem, Y.; Gunišová, S. Embraced by eIF3: structural and functional insights into the roles of eIF3 across the translation cycle. *Nucleic Acids Res.* **2017**, *45*, 10948−10968.
(27) Hinnebusch, A. G. Structural Insights into the Mechanism of Scanning and Start Codon Recognition in Eukaryotic Translation Initiation. *Trends Biochem. Sci.* **2017**, *42*, 589−611.
(28) Valásek, L. S. "Ribozoomin"–translation initiation from the perspective of the ribosome-bound eukaryotic initiation factors (eIFs). *Curr. Protein Pept. Sci.* **2012**, *13*, 305−330.
(29) Hinnebusch, A. G. eIF3: a versatile scaffold for translation initiation complexes. *Trends Biochem. Sci.* **2006**, *31*, 553−562.
(30) des Georges, A.; Dhote, V.; Kuhn, L.; Hellen, C. U. T.; Pestova, T. V.; Frank, J.; Hashem, Y. Structure of mammalian eIF3 in the context of the 43S preinitiation complex. *Nature* **2015**, *525*, 491−495.
(31) Emanuilov, I.; Sabatini, D. D.; Lake, J. A.; Freienstein, C. Localization of eukaryotic initiation factor 3 on native small ribosomal subunits. *Proc. Natl. Acad. Sci. U.S.A.* **1978**, *75*, 1389−1393.

https://dx.doi.org/10.1021/acs.jproteome.0c00160
J. Proteome Res. 2020, 19, 3936−3944

(32) Cuchalová, L.; Kouba, T.; Herrmannová, A.; Dányi, I.; Chiu, W. L.; Valásek, L. The RNA recognition motif of eukaryotic translation initiation factor 3g (eIF3g) is required for resumption of scanning of posttermination ribosomes for reinitiation on GCN4 and together with eIF3i stimulates linear scanning. *Mol. Cell. Biol.* **2010**, *30*, 4671−4686.
(33) Hussain, T.; Llácer, J. L.; Fernández, I. S.; Munoz, A.; Martin-Marcos, P.; Savva, C. G.; Lorsch, J. R.; Hinnebusch, A. G.; Ramakrishnan, V. Structural changes enable start codon recognition by the eukaryotic translation initiation complex. *Cell* **2014**, *159*, 597−607.
(34) Liu, Y.; Neumann, P.; Kuhle, B.; Monecke, T.; Schell, S.; Chari, A.; Ficner, R. Translation initiation factor eIF3b contains a nine-bladed beta-propeller and interacts with the 40S ribosomal subunit. *Structure* **2014**, *22*, 923−930.
(35) Chiu, W.-L.; Wagner, S.; Herrmannová, A.; Burela, L.; Zhang, F.; Saini, A. K.; Valášek, L.; Hinnebusch, A. G. The C-terminal region of eukaryotic translation initiation factor 3a (eIF3a) promotes mRNA recruitment, scanning, and, together with eIF3j and the eIF3b RNA recognition motif, selection of AUG start codons. *Mol. Cell. Biol.* **2010**, *30*, 4415−4434.
(36) Spiess, M.; Junne, T.; Janoschke, M. Membrane Protein Integration and Topogenesis at the ER. *Protein J.* **2019**, *38*, 306−316.
(37) Walter, P.; Blobel, G. Translocation of proteins across the endoplasmic reticulum. II. Signal recognition protein (SRP) mediates the selective binding to microsomal membranes of in-vitro-assembled polysomes synthesizing secretory protein. *J. Cell Biol.* **1981**, *91*, 551−556.
(38) Gilmore, R.; Blobel, G.; Walter, P. Protein translocation across the endoplasmic reticulum. I. Detection in the microsomal membrane of a receptor for the signal recognition particle. *J. Cell Biol.* **1982**, *95*, 463−469.
(39) Meyer, D. I.; Krause, E.; Dobberstein, B. Secretory protein translocation across membranes-the role of the "docking protein". *Nature* **1982**, *297*, 647−650.
(40) Moreno, F. J.; Clemente, A. 2S Albumin Storage Proteins: What Makes them Food Allergens? *Open Biochem. J.* **2008**, *2*, 16−28.
(41) Osborne, A. R.; Rapoport, T. A.; van den Berg, B. Protein translocation by the Sec61/SecY channel. *Annu. Rev. Cell Dev. Biol.* **2005**, *21*, 529−550.
(42) Cheng, Z. Protein translocation through the Sec61/SecY channel. *Biosci. Rep.* **2010**, *30*, 201−207.
(43) Lindholm, P.; Kuittinen, T.; Sorri, O.; Guo, D.; Merits, A.; Törmäkangas, K.; Runeberg-Roos, P. Glycosylation of phytepsin and expression of dad1, dad2 and ost1 during onset of cell death in germinating barley scutella. *Mech. Dev.* **2000**, *93*, 169−173.
(44) Lee, S.; Van, K.; Sung, M.; Nelson, R.; LaMantia, J.; McHale, L. K.; Mian, M. A. R. Genome-wide association study of seed protein, oil and amino acid contents in soybean from maturity groups I to IV. *Theor. Appl. Genet.* **2019**, *132*, 1639−1659.
(45) Lestari, P.; Van, K.; Lee, J.; Kang, Y. J.; Lee, S. H. Gene divergence of homeologous regions associated with a major seed protein content QTL in soybean. *Front. Plant Sci.* **2013**, *4*, 176.
(46) Hamamori, Y.; Wu, H. Y.; Sartorelli, V.; Kedes, L. The basic domain of myogenic basic helix-loop-helix (bHLH) proteins is the novel target for direct inhibition by another bHLH protein, Twist. *Mol. Cell. Biol.* **1997**, *17*, 6563−6573.

https://dx.doi.org/10.1021/acs.jproteome.0c00160
J. Proteome Res. 2020, 19, 3936−3944