Original paper Open Access

Proteomic profiling of quality protein Maize kernels using mass spectrometry

Benito Minjarez[1], Yury Rodríguez-Yañez[1], Eiko Osawa-Martínez[1], Oscar García-Leal[2], Jonathan Buriticá[2],Juan Pedro Luna-Arias[3], Aida Longán-Zarzoso[2], Raúl Páez-Quiñones[2], Laurent Ávila-Chauvet[2], Milagros Ascencio-Guirado[2], Moisés Martín Morales-Rivera[1], Salvador Mena-Munguía1*

[1] Universidad de Guadalajara, Centro Universitario de Ciencias Biológicas y Agropecuarias, 2100 Camino Ramón Padilla Sánchez, Nextipac, Zapopan, Jalisco, México, 45200.
[2] Centro de Estudios e Investigaciones en Comportamiento (Universidad de Guadalajara, Centro Universitario de Ciencias Biológicas y Agropecuarias), Universidad de Guadalajara, 180 Francisco de Quevedo Street, Col. Arcos Vallarta, Guadalajara, Jalisco, México, 44130.
[3] Centro de Investigación y de Estudios Avanzados del Instituto Politécnico Nacional, Departamento de Biología Celular, 2508 Instituto Politécnico Nacional Avenue, Gustavo A. Madero, San Pedro Zacatenco, México city, México, 07360.

* Corresponding author: E-mail: smena@cucba.udg.mx
Keywords: *Amino acids; bioinformatics; kernels; mass spectrometry; proteomics; quality protein maize*

Abstract

Maize (*Zea mays* L.) is the third most crucial crop worldwide and is of paramount importance in both humans and livestock diets. Conventional maize varieties have less than half of the amino acids recommended for human nutrition, and this deficiency results in an imbalance of amino acids and low protein content, which has been associated with several pathologies, including malnutrition. Thus, different countries have focused on research on fortified foods, such as quality protein maize (QPM) noting that these improved varieties may contain up to 100% more essential amino acids residues than conventional maize. Hence, this study aimed to characterize through tandem mass spectrometry and bioinformatics analysis, relative expression of polypeptides contained in a hybrid variety of QPM, which allow to identify potential markers with implications in the management and improvement of this crops maintaining their intrinsic characteristics. We identify 262 polypeptides, highlighting those related to molecular function (catalytic activity, structural molecule activity, and binding) and biological process (cellular and metabolic process). These results provide the necessary information, not only for the characterization of the QPM proteome through novel tools such as proteomics, but also to describe mechanisms related to different biological processes such as the embryogenesis, development and growth of grains and eventually plants. Potentially It promotes the discovery of molecular markers (biomarkers) that would allow the improvement of agronomical processes.

Introduction

Maize (*Zea mays* L.) showed the highest level of world production in history during the 2016/17 cycle with an approximate amount of 1,025.6 million tons, ranking as the mainsource of food for humans and livestock in many regions around the globe (FIRA, 2016). Corn consumption per capita in many developing countries is approximately 330 grams per day, providing between 15 and 56% of all calories ingested, especially in those low incomes (FAOSTAT, 2014).

Additionally, in emerging countries, corn supplies different vitamins and minerals that are important in the human diet, particularly for children, elderly and pregnant women, in addition to being an essential source of protein, accounting for up to 60% of daily intake of proteins supply. Therefore, the amino acid composition and amount of seed storage proteins are related to the nutritional quality (Musila et al., 2010). However, corn proteins are characterized by low nutritional values compared with those of animal origin, as in many other types of cereals (Liu et al., 2015).

The corn kernel is mainly constituted by endosperm (82%), germ (12%) and pericarp (6%). The endosperm contains starch, minerals and proteins, the latter being found in limited quantities (8%-15%). Between 60% and 70% of the protein content corresponds to a group of prolamin storage proteins known as zeins that belong to a family of alcohol-soluble prolamin storage proteins, which are considered the most abundant on kernels maize, representing up to 50% of the total protein content (Prasanna, 2001; Wang et al., 2016). The remaining percent of endosperm proteins include glutelins, globulins and albumins (Vasal, 2000), all of them with differences in amino acid composition and therefore, their structure (Wilson, 1987). These types of proteins have a high content of lysine (5%-6%), unlikezeins,

which are characterized by low concentrations of lysine (Lys) and tryptophan (Trp) (0.1%) but more abundant in glutamine/glutamic acid (21%) than three other proteins (11%). Nevertheless, glutamine, proline, leucine, and alanine make up approximately 65% of amino acid residues in zeins (Lawton and Wilson, 2003). The low content of Lys and Trpproduces an amino acids imbalance in proteins, which would have less than half of the recommended amino acids for human nutrition, which puts people at risk of nutritional deficiency (Wu et al., 2010; Tien et al., 2016).

On the other hand, Mertz et al. (1964) described a transcription factor belonging to the bZIP class (basic leucine zipper domain) known as *opaque-2 (o2)* in corn. This gene is expressed in the subaleurone layers of endosperm cells during seed development from 10 to 45 days after pollination and primarily regulates the expression of α- and β-zein genes. Furthermore, expression of theo2 gene can modify chromatin states and could storage proteins rich in lysine and tryptophan, decreasing the synthesis of zein proteins (Locatelli et al., 2009; Wu et al., 2010; Wu, & Messing, 2012).

These findings allowed development of so-called quality protein maize (QPM) varieties in the late 1990s. This improved variety maintains the phenotype and yield of normal corn, but significantly increases the content of lysine and tryptophan residues.

As mentioned above, there is a nutritional deficiency among the most abundant storage proteins in cereals, including corn; nevertheless, QPM could play a very important role in alleviating malnutrition and providing valuable nutritional improvement. To date, QPM has not been fully implemented due to multiple agronomic defects, i.e., low yield and susceptibility to pests and diseases. The high number of *o2* modifiers and its widespread distribution among all ten chromosomes increases the technical complexity. Moreover, the lack of studies on genetic and proteomic diversity, the content of specific residues in proteins in several QPM varieties, as well as their genetic relationships with conventional varieties of corn, also contribute to theseproblems (Hasjim et al., 2009; Liu et al., 2015).

Diverse efforts have been made related to the characterization of different QPM varieties and marker-assisted selection for the improvement of agronomically important traits. However, those genes that are expressed in the proteome of different varieties of corn,both conventional and improved, such as QPM,have yet to be identified. Thus, we decided to analyze QPM2006 because the characteristic "quality protein maize" (QPM) is determined by the expression of recessive genes, which has limited the expansion of its plantation and its intensive use in human and animal nutrition. In this situation, we can carry out an analysis from the molecular and proteomic point of view that allows us to go one step further in identifying markers with potential use in the management and improvement of crops that maintain certain characteristics that avoid the limitations that present the crossover method as it currently occurs. For this reason, it is not considered to include a conventional corn counterparty to test. Well, this is not the objective, where in addition, the differences between both conventional maize and QPM has already been sufficiently demonstrated.

Therefore, the main objective of this study was to characterize the proteins present in the kernels of a QPM variety using liquid chromatography, mass spectrometry, and bioinformatics analysis. The proteome profile of these kernels could support the enrichment of databases information to improve accessibility of the different bioinformatics platforms for users with specific needs, such as the discovery and generation biomarkers of potential quality as well as new approaches for the improvement in the production of food of high nutritional value.

Materials and Methods

Recommended agronomic practices were followed tobreed a variety of corn used in this study. Kernel proteins were obtained from plants grown at the University of Guadalajara, La Huerta, Jalisco, Mexico (19° 30′ 17.3″ N y 104° 32′ 56.09″ W), at an altitude of 1,700 m, in predominantly haplic phaeozem soil under semiarid climatewith warm and dry winters and springs.The average annual temperature is 25.2 °C, the maximum average is 32.8 °C, and the minimum average is 17.6 °C.We used a variety of quality protein maize, QPM 2006, which is a hybrid variety obtained from crossing CML 501 and CML 491 (female and male, respectively). The germplasm bank of the University of Guadalajara at Zapopan Jalisco, Mexico, kindly donated the sample. Pollination was controlled between the inbred progenitor lines and QPM in order to avoid the crossing of lines. As a nitrogen sourcewe added diammonium phosphate (DAP) (18-46-00) at a fertilization dose of 200-92-00 and urea (46-00-00), under irrigation conditions; one day after sowing, we applied 3 liters/ha of pre-emergent herbicide Harness Xtra (Acetochlor + Atrazine) and 1 kg/ha of the herbicide Caliber 90 (Atrazine). The corn sample kernels were ground through a grain milling machine (Thomas Model 4 Wiley) and stored at 4 °C and controlled humidity until use

Extraction and quantification of polypeptides

The proteomic analysis was carried out in the Proteomics Unit of the University of Valencia, Valencia, Spain

(a member of the PRB2-ISCIII ProteoRed Proteomics Platform). A sample of the ground kernel (100 mg) was used to obtain a total homogenate for the extraction and quantification of polypeptides. The sample was re-suspended in 300 µl of a solution containing 8 M urea with a portable tissue dispensator (a Sigma-Aldrich cordless motor pellet pestle), and then homogenized by sonication three times, each 1 min, at 60 watts of amplitude and keeping on ice for one min, using a CPX130PB Ultrasonic Processor (Cole-Parmer). The sample was delipidated three times with a mixture of methanol-chloroform-water (4:1:3; v:v:v), vortexed 1 minute at the highest speed, centrifuged at 12,000 xg for 5 min at room temperature (RT), and methanol precipitated (Minjarez et al., 2013). The pellet was partially dissolved with 400 µl of U/T/C mixture (7 M Urea, 2 M Thiourea, 2 % CHAPS). Due to the low solubility of the sample, it was loaded on a gel and stained with Coomassie Brilliant R250 Blue (Bio-Rad) to visualize proteins bands. All the 1D SDS-PAGE lanes (four in total) were sliced and the proteins digested for their analysis.

In-gel trypsin digestion of proteins

We used the method described by Shevchenko et al. (1996) and Tran et al. (2016) with some modifications. Briefly, each slide was cut and small pieces of approximately 1 mm in size, transferred into 1.5 ml tubes, and washed in 50 mM ammonium bicarbonate (ABC) in water pH 8.0 (Sigma). Proteins were reduced with 10 mM dithiothreitol (DTT) in 50 mM ABC for 20 minutes at 60 °C and alkylated with 50 mM iodoacetamide in 50 mM ABC for 30 minutes at RT in the dark. The gel slices were then dehydrated in acetonitrile (ACN) (Fisher Scientific) and the proteins in each sample were digested with 500 ng of trypsin in 200 µl 50 mM ABC at 37 °C overnight. Proteinase digestions were stopped by adding 10% Trifluoroacetic acid (TFA) to a final concentration of 1% and the supernatants were carefully removed. For additional peptide extraction, 200 µl ACN were added to each tube and incubated for 15 min at 37 °C in a shaker. The supernatants containing the peptide mixtures were dried in a speed vacuum (ISS 110 SpeedVac System, Thermo Savant, ThermoScientific, Langenselbold, Germany) for 20 min and then resuspended in 15 µl 2% ACN and 0.1% TFA before liquid chromatography (LC) and mass spectrometry (MS) analysis.

LC-MS/MS analysis

The peptides mixtures were separated by LC using a NanoLC Ultra 1-D plus Eksigent® (Eksigent Technologies, Dublin, CA, U.S.A.), which was directly connected to a SCIEX TripleTOF 5600 Mass Spectrometer (SCIEX, Framingham, MA, U.S.A.) in direct injection mode. Briefly, 5 µl from each digested sample were trapped on a NanoLC pre-column (3 µm particle size C18-CL, 350 µm in diameter x 0.5 mm long; Eksigent) and desalted with 0.1% TFA at a flow rate of 3 µl/min for 5 min. Then, the digested peptides were separated using an analytical C18-LC column (3 µm particle size, 75 µm in diameter x12 cm long, NikkyoTechnos Co®, Tokyo, Japan) and equilibrated in 5% ACN and 0.1% formic acid (FA) (Fisher Scientific). The peptides were eluted from the column with a linear gradient from 5% to 40% solvent B at a constant flow rate of 300 nl/min for 120 min. Solvent A was 0.1% FA in water and solvent B was 0.1% FA in ACN. Eluted peptides were ionized applying 2.8 kV to the spray emitter on an ESI Nanospray III (SCIEX). The analysis was carried out in data-dependent acquisition (DDA) mode. Survey MS1 scans were acquired from 350–1250 m/z for 250 milliseconds. The quadrupole resolution was set to 'UNIT' for MS2 experiments, which were acquired from 100–1500 m/z for 50 milliseconds in 'high sensitivity' mode. The following dynamic exclusion criteria were used: charge, 2+ to 5+; minimum intensity, 70 counts per second (cps). Up to 50 ions were selected for fragmentation after each survey scan. Dynamic exclusion time was set to 15 seconds. Collision energy was automatically set by the instrument according to the equation: |CE|= (slope) x (m/z) + (intercept) with Charge = 2; Slope = 0.0575 and Intercept = 9.

Protein identification

The MS/MS information was sent to the PARAGON algorithm through Protein Pilot v 4.5 software (ABSciex). The default Protein Pilot parameters were employed to generate the peak lists directly from the 5600 TripleTOF. The Paragon algorithm of Protein Pilot was used to search the UniProtKB protein database with the following parameters: trypsin specificity, cys-alkylation, no taxonomy restriction, and the search effort set to thorough. After this search, it was evident that the information of both fractions (supernatant and pellet) must be combined. Then, the combined information for each sample was applied to search the NCBInr database with the taxonomy established to *Zea mays* L.

To avoid using the same spectral evidence for more than one protein, the Protein-Pilot Pro Group algorithm grouped the identified proteins based on MS/MS spectra. A protein group in a Pro Group Report is a set of proteins that share some physical evidence. Unlike sequence alignment analyses, where full-length theoretical sequences are compared, protein group formation in the Pro Group is entirely guided by observed peptides only. Since the observed peptides are actually determined from experimentally acquired spectra, clu-

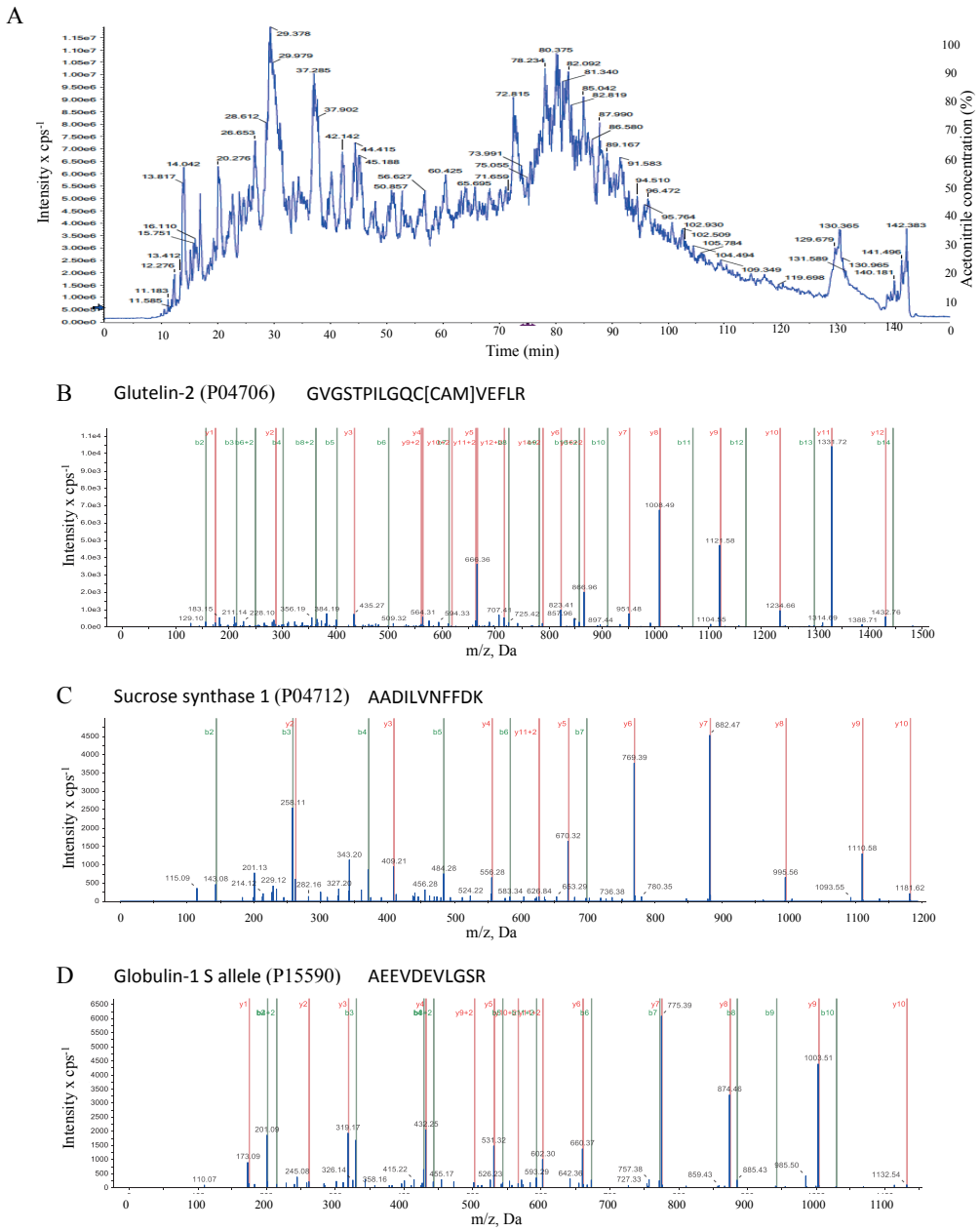
Table 1 - List of polypeptides identified in the proteomics analysis of whole protein extracts of Zea mays L. by tandem mass spectrometry.

Protein name	UniProt/KB_ID[a] Acc. No.	Gene[b]	MW[c] (kDa)	pId	Pep.[e] (> 95)	% Cov.[f] (≥ 95)
1. 40S ribosomal protein S3	K7TT73	100274348	25.44	9.55	11	44.7
2.Aminopeptidase M1-B	B4FSK5	100274215	38.25	8.61	4	3.8
3.60S ribosomal protein L5-1	B6TYP7	N/A	34.37	9.39	3	21.5
4.Shrunken2	P55241	SH2	52.07	5.86	34	43.5
5.60S ribosomal protein L9	B6SI39	00001d035901	21.46	9.71	5	21.5
6. Pyrophosphate-energized vacuolar membrane proton pump	P49087	N/A	61.95	5.88	7	7.9
7. Pyruvate kinase	B4F9G8	100191614	57.34	6.44	4	7.2
8. Glutelin-2 precursor	P04706	N/A	21.82	8.2	109	78
9. Sucrose synthase 1 isoform X1	P04712	SH-1	91.73	5.96	57	50.2
10. Granule-bound starch synthase 1, chloroplastic/amyloplastic	P04713	WAXY	58.57	5.49	35	40.4
11.Endochitinase A precursor	P29022	N/A	26.8	8.16	44	78.5
12. Endosperm ADP-glucose pyrophosphorylase	P55240	GLG1	13.24	5.56	35	52.4
13.Globulin-1 S allele precursor	P15590	GLB1	55.13	6.75	33	59.8
14. Alcohol dehydrogenase 1 isoform X1	P00333	ADH1	40.98	6.28	22	39.5
15. TPA: malate synthase, glyoxysomal	P49081	LIP	61.63	6.18	11	18.2
16. Sucrose synthase2	P49036	SUS1	92.93	6.03	18	16.1
17. Phosphoglucomutase, cytoplasmic 1	P93804	N/A	63.09	5.46	8	12.6
18. Glutathione transferase30	Q9FQA3	N/A	24.87	6	7	18.6
19. Enolase 1	P26301	ENO1	48.06	5.2	7	18.1
20.Tubulin beta-7 chain	Q41784	TUBB7	50.09	4.72	5	7.3
21.Actin-2	B6SI11	100280540	41.67	5.31	5	18
22.TPA: pyruvate decarboxylase3	Q05327	PDC3	22.29	5.21	3	8.2
23.Triosephosphate isomerase, cytosolic	P12863	N/A	26.89	5.52	3	13.8
24.60S ribosomal protein L4	B6SK79	N/A	44.44	10.51	4	6.4
25.14-3-3-like protein GF14-6	P49106	GRF1	29.66	4.76	6	23
26.Actin-depolymerizing factor 3	Q41764	ADF3	15.89	5.46	3	24.4
27. Sec24/Sec24 protein transport family protein	K7UW20	103635100	93.35	6.74	2	2.5
28. Heat shock 70 kDa protein-like	P11143	HSP70	70.57	5.22	3	3.7
29.Eukaryotic peptide chain release factor GTP-binding subunit ERF3A	P55876	EIF5	48.91	5.52	2	3.7
39.1-Cys peroxiredoxin PER1	A2SZW8	PER1	24.9	6.31	2	10
40.Guanine nucleotide-binding protein beta subunit-like protein	P49178	GB1	41.71	7.15	2	5.9
41. Brassinosteroid biosynthesis-like protein	Q5YFA2	DWF1	65	8.23	2	5.1
42. Phosphoenolpyruvate carboxykinase homolog	Q9SLZ0	N/A	73.31	6.57	2	4
43. ATP synthase subunit alpha, mitochondrial	P19023	ATPB	54.06	5.19	2	4.9

[a]UniProtKB Acc. No., UniProt Knowledgebase Accession Number; [b] Gene name; [c]MW (kDa), Molecular Weight; [d]pI, Isoelectric point; [e]Pep. (>95), number of peptides having at least 95% in confidence in their identification by mass spectrometry; [f] %Cov. (>95), Percentage of coverage (with peptides having a confidence with at least 95%).

stering can be considered a search guided by the use of spectra. Then, unobserved regions of the protein sequence play no role in explaining the data. Data were also exported to a spreadsheet for manual analysis.

Fig. 1 - Identification of peptides QPM2006 sample by tandem mass spectrometry. Solubilized samples were fractionated in a reverse phase column and eluted peptides were analyzed in a mass spectrometer as described in Materials and Methods. A) Total ion count found in QPM with 8 M urea. B) Spectrum of a peptide from the glutelin 2 obtained by tandem mass spectrometry. C) Mass spectrum of a selected peptide from the sucrose synthase 1. D) Fragmentation spectrum obtained for a specific peptide from the globulin 1 s allele. We included the sequence of identified peptides in their corresponding spectra. Intensity of peaks is in Counts Per Second or CPS.

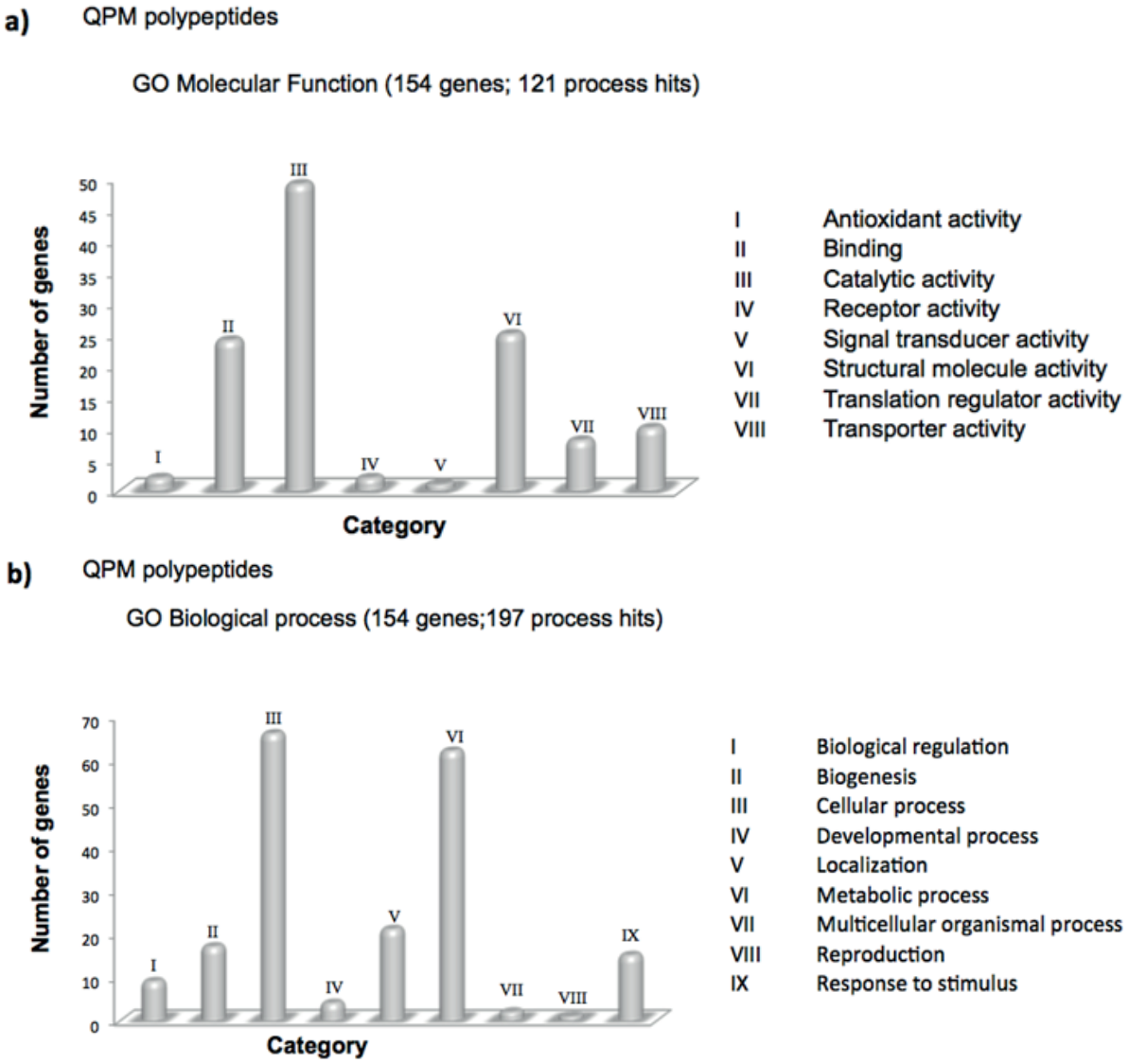


Bioinformatics analysis

Briefly, polypeptides were grouped and classified using the PANTHER (Protein Annotation Through Evolutionary Relationship) Classification System v10.0 (http://www.pantherdb.org/), according to their role in molecular function and biological processes (Mi et al., 2012). Additionally, the identification of the functional interaction networks was performed with the STRING (Search Tool for the Retrieval of Interacting Genes/Proteins) program v10.0 (http://string-db.org/). The active prediction methods selected were neighborhood, co-expression, gene fusions, experiments, concurrency, databases and text mining, using a high confidence value (0.7).

Results and discussion

Fig. 2 - Classification of the polypeptides identified by tandem mass spectrometry. Proteins identified specifically in QPM samples were classified according to their class (a) or molecular functions and (b) biological process using the PANTHER program and Zea maysL. databases. Numbers in parenthesis correspond to the percentage of the total polypeptides with the indicated function. Proteins identified in the UniProtKB/TrEMBL database were not included in this analysis.



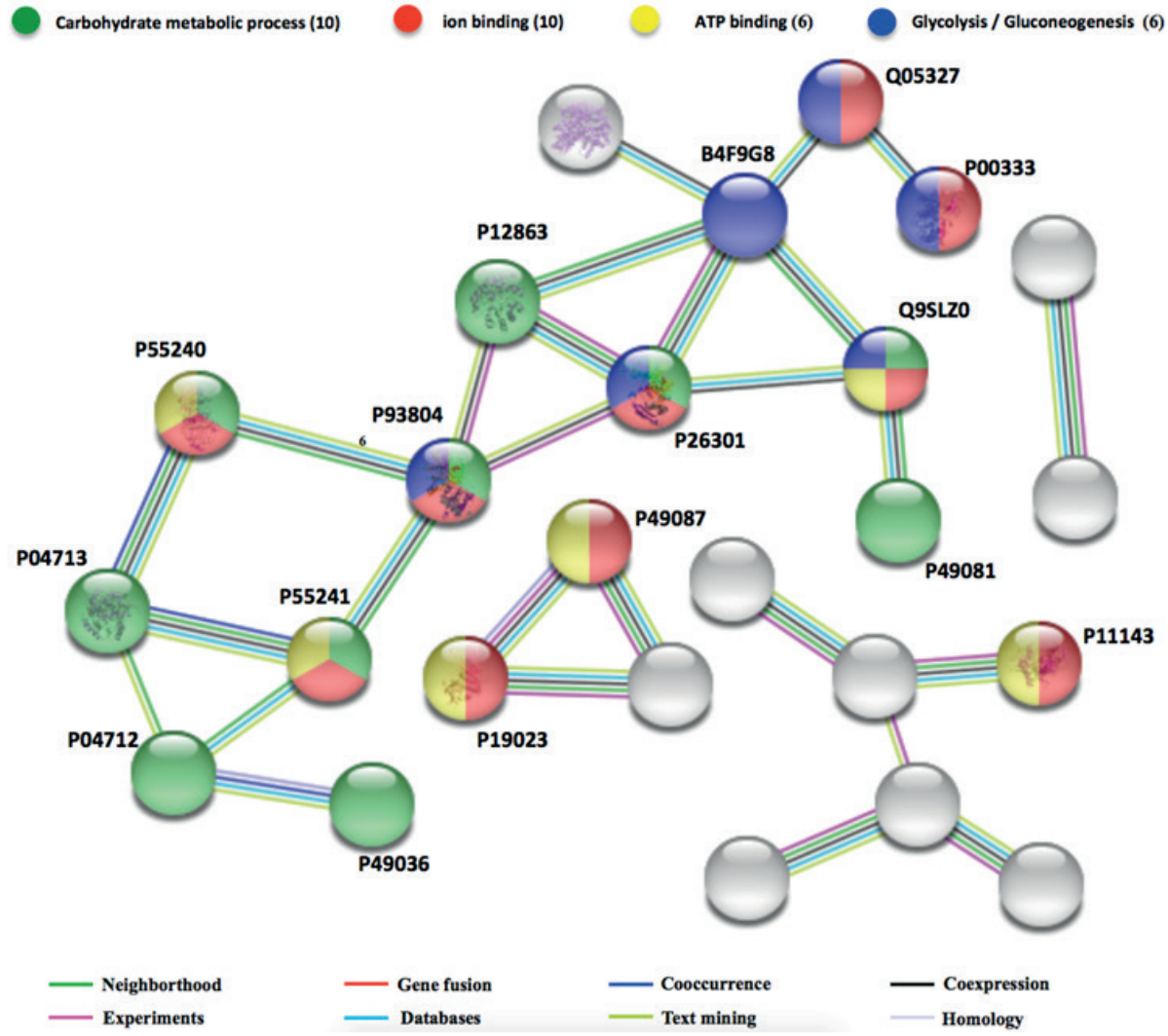
Proteins identified by tandem mass spectrometry

To better understand the protein components of the QPM2006 corn kernels, we developed a strategy through LC and mass spectrometry. First, proteins were solubilized in 8M urea using sonication as described above. Then, those biomolecules like carbohydrates and lipids that could interfere with the precipitated proteins and future analysis were removed. After the reduction and alkylation of proteins, we decreased salt concentrations and performed the trypsin digestion of polypeptides and analysis on a hybrid quadrupole time-of-flight mass spectrometer.

Peptides bound to the C18 reverse phase column were eluted using an acetonitrile gradient as described before and immediately analyzed by mass spectrometry for identification. We obtained 32,151 tandem mass spectra with at least 95% confidence (Fig. 1a). We show three tandem mass spectra as examples, corresponding to the GVGSTPILGQC[CAM]VEFLR, AADILVNFFDK, and AEEVDEVLGSR peptides identified in the UniProt Knowledgebase, which belong to Glutelin-2, Sucrose synthase 1, and Globulin-1 S allele polypeptides, respectively (Fig. 1 b–d). A total of 262 proteins were identified in the total homogenate sample, which has at least one identified peptide with greater than 95% confidence (see Supplementary Table 1).

To have a better panorama of the proteins identified,

Fig. 3 - Interactome of polypeptides found in QPM sample.UniProtKB accession numbers were submitted to the String program to identify the predicted functional network. Lines in color represent different evidences for each identified interaction. The polypeptides involved in carbohydrate metabolic process, ion biding, ATP biding and glycolysis/gluconeogenesis were denoted by a pink circle in green, red, yellow and blue, respectively. Proteins had a statistical significance (p-Value ≤0.05) in the different biological processes that were related.



we classified them according to their molecular functions and their biological (Table 1) processes using the PANTHER Classification System (Fig. 2). Polypeptides were submitted for classification, first according to their molecular function.Approximately 36% (97 polypeptides) were classified in three different classes1) Catalytic activity (GO:0003824), with 49 members, highlighting oxidoreductases, such as 1-cys-peroxiredoxin (A2SZW8)which has a strongly regulated expression in early germination of the grain (Stacy et al., 1996), and protects against reactive oxygen species (ROS) or sulfhydryl radical damage during grain desiccation; likewise, this protein acts as a nuclear DNA protector in seed cells during oxidative stress, thus playing a role as an antioxidant (Kim et al., 2011).

Another member within this class is the enolase 14-3-3 like-protein(P49106) whichplays a crucial role in growth and development through cell cycle regulation, hormone signaling, regulation of carbon and nitrate metabolism, as well as ion homeostasis (Diaz et al., 2011). Its highest level of expression occurs in early stages (3-10 d after pollination, DAP). In contrast, low levels are observed in the maturity stage of the grain (10 DAP), which is in agreement with its biological function (Yu et al., 2016). Besides, the expression pattern of 14-3-3 proteins has been associated with starch content in *Arabidopsis* leaves, where its overexpression may decrease the activity of enzymes participating in car-

bohydrate metabolism (Diaz et al., 2011). Therefore, in corn kernel, the synthesis and the starch content are expected to start increasing due to the low expression of the 14-3-3 like-protein.

2) Structural molecule activity (GO:0005198), with 25 members related mainly to the structural constituent of the cytoskeleton (GO:0005200) and ribosomes. They include the Tubulin beta-7 chain (Q41784), a component of the cytoskeleton involved in seed and plant growth, cell division, cell motility, intracellular transport, and maintenance of the cell shape. Previous studies have implied the involvement of beta-tubulin in seed dormancy of different crops such as dormant dry seeds of tomato (de Castro et. al., 1995) and *L. chinensis* (Liu et. al., 1997) that showed a low beta-tubulin accumulation. In contrast, seed dormancy abolishing and germination are characterized by increases of alpha-and beta-tubulin (Pawłowski et. al., 2004; Chibani et. al., 2006). Thus, the increase and accumulation of beta-tubulin is important to cell expansion, division and growth of the radicle through the seed coat (Pawłowski et. al., 2004).

Furthermore, among the proteins identified with structural activity also are standing out 40S- (with three members BSSHZ1, K7TT73, P2540) and 60S- ribosomal proteins (with three members B6SI39, BSTYP7, BSK79). In eukaryotes, the cytosolic ribosome is built by a large 60S subunit that contains three ribosomal RNAs (rRNA) and a small 40S subunit with two rRNA (Layat et al., 2011). The dissociation of polysomes takes place before fixation of the preformed messenger RNA (mRNA) occurs, leading to protein synthesis in germinating embryos (App et al., 1971). Interestingly, both aging and prolonged grain storage may cause loss and decreased mRNA synthesis. (Weidner &Zalewski, 1982). Some deleterious aspects, such as water inhibition or drought, might affect the germination process with the consequent loss of seed viability. In such a case, the ribosomes do not dissociate, and protein synthesis is delayed (Bray & Chow, 1976).

3) Binding proteins (GO:0005488), with 22 members related to nucleic acid binding activity (GO:0003676) and biding proteins activity (GO:0005515) that comprise nine and six proteins, respectively, where brassinosteroid biosynthesis like-protein (Q5YFA2) protrudes; this is known as DIM/DFW1 and is involved in gene expression regulation, ethylene biosynthesis, stem elongation, pollen tube growth, leaf bending, and epinasty (see Hola, 2019 for review). The knockout or down-expression ofbrassinosteroid (BR) biosynthetic genes in rice might produce infertility and reduced seed yield because of the downsizing seeds (Mori et al., 2002). However, only BR-deficient mutants could be rescued to wild type phenotype through BR exogenous administration as treatment. In maize, ZmDWF1 is expressed in all tissues but is highly expressed in the leaf collars, limitedly expressed in aged areas. Biosynthesis of active BR occurs primarily in maize roots, mainly in young root tips (Tao et al., 2004; Kim et al., 2005). Furthermore, BR is generally considered a positive factor in drought-stressed plants. In this sense, the drought-resistant genotype corn has a higher total BR content compared to the drought-sensitive genotype, which could be related to water-deficiency stress (Tůmová et al. 2018). Then, the remaining polypeptides were grouped into five other classes (Fig. 2a).

Meanwhile, when the proteins were classified according to biological processes, we observed two main groups that contained 63 members and represented up to 47% (126 proteins) of the total identified proteins. The first group corresponds to cellular processes (GO:0009987), mainly related to cell communication (GO:0007154) and cell cycle (GO:0007049). We found the Actin-depolymerizing factor 3 (Q41764) which participates in the assembly of the actin cytoskeleton and plays a crucial role in development, cell expansion, and tip growth, via the regulation of F-actin and therefore, cell elongation (Dong et al., 2001; Augustine et. al., 2008), in response to stressors of temperature, drought, and salinity (Baisakh&Subudi, 2009), as well as to pathogens and pests (Tian et al., 2009; Mondal et al., 2018).Itscorresponding gene in maize *ZmABP3(ZmADF3)* is expressed in all tissues, except pollen. Eventually,ZmABP3 can bind to G-and F-actin and the Ser6 phosphorylation is a conserved regulatory mechanism of ADF activity in F-actin disassembly (Ressad et al., 1998). Finally, the remaining polypeptides participate in seven different processes (Fig. 2b).

Putative interaction networks determined by String analysis

The putative interaction networks of the identified polypeptides were determined with the String software. The interactome was obtained based on evidence of interaction, as reported in the literature. In this way, the main processes involved were: 1) biological process, 2) cellular process, 3) cellular biosynthetic process, 4) organic substance biosynthetic process, 5) cellular macromolecule biosynthetic process, and 6) primary metabolic process, among others. We grouped some of the proteins according to their catalytic activity because this group was the most abundant according to Panther classification. Thereby, the protein network involved 16 molecular functions, all with statistical significance ($p \leq 0.05$), being the main processes identified: 1) Carbohydrate metabolic process (10 members in green), 2) ion binding (10 members in red), 3) ATP

binding (six members in yellow), and 4) Glycolysis / Gluconeogenesis (six members in blue) (Fig. 3). The interactome also provides information related to the specific interactions reported according to different pieces of evidence, such as an experimental level, text mining, homology, and co-expression, allowing to link proteins sharing functions, biochemical pathways, or neighborhoods (colored lines). For details, see table S1.

Agricultural and nutritional perspectives based on QPM analysis

Many countries around the world have already adopted new strategies related to the production of different agricultural products, both quality and quantity, to meet future demands and achieve self-sufficiency and food security (Weiler et al., 2014). Various strategies have been developed for the improvement of the main food products for humans and livestock to obtain an optimal calorie and protein intake (Sands et al., 2009). Many of the cereal grains have proteins with a relatively low nutritional value when they are compared to the proteins in meat. In the case of maize, most of the proteins are found in the endosperm. Still, they are deficient in the essential amino acids lysine and tryptophan (Hasjim et al., 2009), resulting in amino acids imbalance and low protein content. Common maize varieties have less than half the recommended amounts of amino acids for human nutrition. However, the discovery and development of QPM have provided a new alternative to develop a better grain, helping to improve the nutritional quality of food and nutrition. Thereby, the challenge is to innovate and identify those tools that are useful to enhance production and quality control mechanisms to increase the nutritional and commercial values of maize.

In this study, we identified 262 proteins in QPM maize kernel that were classified according to their molecular functions (catalytic activity, structural molecule activity and binding) and biological processes (cellular and metabolic processes). These proteins are involved in seven biological processes, mainly related to the stress response and catabolic processes(Kim et al., 2011). Because a wide range of physical and chemical stressors induce oxidative stress in grains and plant cells, it is essential to study those related to redox-dependent signaling pathways for better management, not only of the redox state but also of the metabolic processes that are susceptible to modification under those stressors, which often affect plantations. This knowledge would also provide excellent ways to increase stress resistance and crop performance.

As mentioned above, we found some proteins related to aspects of crop/plant growth, development and damage due to water-deficiency stress (Ashikari et al., 1999; Mori et al., 2002; Tůmová et al. 2018). In this sense, the study and knowledge of these proteins, the biosynthesis and the endogenous content (*i.e.* brassinosteroid biosynthesis like-protein) collectively with a water supply and cultivation conditions, could generate significant agronomic advances that could allow manipulating gene expression/protein activities in plants and crops to improve their quality and productivity.

On the other hand, it should be noted that those proteins related to structural and binding activities are highly relevant given their functions (i.e.beta-tubulin, ribosomal proteins or ADF). Accordingly, under detrimental conditions involving energy, nutritional or environmental stress on plants or grains, the adaptive response would appear through the dormancy of the seeds, as well as an inhibition of growth and development (Baisakh&Subudi, 2009; Tian et al., 2009; Mondal et al., 2018). Thereby, the research based on the regulation of ribosomes, translational activity and microtubule organization is a challenge that would provide possibilities to improve kernel development and plant growth.

Typically, it is assumed that mRNA transcript levels should be correlated with protein concentrations. However, there is a discrepancy between both rates. To date, there is no conclusive explanation about this behavior. Still, different contributors have been proposed, for example, the half-life of RNAs and proteins, as well as the transport of proteins within tissues (Greenbaum et al., 2003). In fact, as end effectors, mature proteins perform most of the biological functions. In this sense, proteomics offers the complete characterization of a set of proteins produced by an organism (known as a proteome), which includes quantification, activities, and interactions.

Noteworthy, there are also many other important polypeptides identified in this work that were not discussed. Further studies are required to fully characterize and quantify the level of expression of these proteins in QPM and common maize varieties through quantitative proteomics. These studies could allow the identification of novel molecules to establish a better understanding of the nutritional contribution of each of the agricultural products, favoring the development of new selection and quality control mechanisms.

Conclusions

In conclusion, we have standardized on several different approaches to use in proteomic analysis that could be used to improve agronomic performance. Therefore, in this first evidence, we could propose that proteomic

analysis may be used to characterize maize and other crops, having a direct impact on improving agronomic performance. One of the most popular applications of proteomics is the discovery of biomarkers that could be applied, especially through quality approaches, as predictors of prognosis and outcomes. These predictors would allow manipulation of metabolic control, as well as signaling pathways that handle embryogenesis, growth, and development of seeds. Finally, the impact on diet and its relationship with the quality of protein content requires further investigation and should be elucidated in subsequent studies.

Acknowledgements

The authors would like to thank Secretary of Public Education- Secretary of Higher Education (SEP-SES) of Mexico for the support to YRY (PRODEP 241286). Also, the authors would like to thank CONACYT México for the support to EEOM (724639/634015 CONACYT). The authors declare no competing financial interest and take part and approved the final manuscript.

References

App AA, Bulis, MG, McCarthy WJ, 1971. Dissociation of ribosomes and seed germination. Plant physiology 47(1):81-86.

Augustine RC, Vidali L, Kleinman KP, Bezanilla M, 2008. Actin depolymerizing factor is essential for viability in plants, and its phosphoregulation is important for tip growth. The Plant Journal 54(5):863-875.

Baisakh N, Subudhi PK, 2009. Heat stress alters the expression of salt stress induced genes in smooth cordgrass (Spartina alterniflora L.). Plant Physiology and Biochemistry 47(3):232-235.

Bray CM, Chow TY, 1976. Lesions in post-ribosomal supernatant fractions associated with loss of viability in pea (Pisum arvense) seed. Biochimica et Biophysica Acta (BBA)-Nucleic Acids and Protein Synthesis 442(1):1-13.

Chibani K, Ali-Rachedi S, Job C, Job D, Jullien M, Grappin P, 2006. Proteomic analysis of seed dormancy in Arabidopsis. Plant physiology 142(4):1493-1510.

de Castro RD, Zheng X, Bergervoet JH, De Vos CR, Bino RJ, 1995. [beta]-Tubulin accumulation and DNA replication in imbibing tomato seeds. Plant Physiology 109(2): 499-504.

Diaz C, Kusano M, Sulpice R, Araki M, Redestig H, Saito K, Stitt M, Shin R, 2011. Determining novel functions of Arabidopsis 14-3-3 proteins in central metabolic processes. BMC systems biology 5:192.

Dong CH, Xia GX, Hong Y, Ramachandran S, Kost B, Chua NH, 2001. ADF proteins are involved in the control of flowering and regulate F-actin organization, cell expansion, and organ growth in Arabidopsis. The Plant Cell 13(6):1333-1346.

FAOSTAT, 2014. Food Supply-Crops Primary Equivalent http://faostat3.fao.org/browse /FB/CC/S.

FIRA, 2016. Panorama agroalimentario mexicano: maíz 2016. Dirección de investigación y evaluación económica y sectorial. https://www.gob.mx/cms/uploads/attachment/file/200637/Panorama_Agroalimentario_Ma_z_2016.pdf

Greenbaum D, Colangelo C, Williams K, Gerstein M, 2003. Comparing protein abundance and mRNA expression levels on a genomic scale. Genome biology 4(9):117.

Hasjim J, Srichuwong S, Scott MP, Jane JL, 2009. Kernel composition, starch structure, and enzyme digestibility of opaque-2 maize and quality protein maize. Journal of agricultural and food chemistry 57(5):2049-2055.

Hola D, 2019. Role of Brassinosteroids in the Plant Response to Drought: Do We Know Anything for Certain? In Brassinosteroids: Plant Growth and Development (pp. 101-168). Springer, Singapore.

Kim SY, Paeng SK, Nawkar GM, Maibam P, Lee ES, Kim KS, Lee DH, Park DJ, Kang SB, Kim MR, Lee JH, Kim YH, Kim WY, Kang CH, 2011. The 1-Cys peroxiredoxin, a regulator of seed dormancy, functions as a molecular chaperone under oxidative stress conditions. Plant science 181(2):119-124.

Kim YS, Kim TW, Kim SK, 2005. Brassinosteroids are inherently biosynthesized in the primary roots of maize, Zea mays L. Phytochemistry 66(9):1000-1006.

Lawton JW, Wilson CM, 2003. Proteins of the kernel. Corn chemistry and technology. 2nd edn. AACC, St. Paul, MN.

Layat E, Sáez-Vásquez J, Tourmente S, 2011. Regulation of Pol I-transcribed 45S rDNA and Pol III-transcribed 5S rDNA in Arabidopsis. Plant and Cell Physiology 53(2):267-276.

Liu L, Jeffers D, Zhang Y, Ding M, Chen W, Kang MS, Fan X, 2015. Introgression of the crtRB1 gene into quality protein maize inbred lines using molecular markers. Molecular breeding 35(8):154.

Liu Y, Hilhorst HWM, Groot SPC, Bino RJ, 1997. Amounts of nuclear DNA and internal morphology of gibberellin-and abscisic acid-deficient tomato (Lycopersicon esculentum

Mill.) seeds during maturation, imbibition and germination. Annals of botany 79(2):161-168.

Locatelli S, Piatti P, Motto M, Rossi V, 2009. Chromatin and DNA modifications in the Opaque2-mediated regulation of gene transcription during maize endosperm development. The Plant Cell 21(5):1410-1427.

Mertz ET, Bates LS, Nelson OE, 1964. Mutant gene that changes protein composition and increases lysine content of maize endosperm. Science 145(3629):279-280.

Mi H, Muruganujan A, Thomas PD, 2012. PANTHER in 2013: modeling the evolution of gene function, and other gene attributes, in the context of phylogenetic trees. Nucleic acids research 41(D1):D377-D386.

Minjarez B, Valero Rustarazo ML, Sanchez del Pino MM, González-Robles A, Sosa-Melgarejo JA, Luna-Muñoz J, Mena R, Luna-Arias JP, 2013. Identification of polypeptides in neurofibrillary tangles and total homogenates of brains with Alzheimer's disease by tandem mass spectrometry. Journal of Alzheimer's Disease 34(1):239-262.

Mondal HA, Louis J, Archer L, Patel M, Nalam VJ, Sarowar S, Sivapalan V, Root DD, Shah J, 2018. Arabidopsis actin-depolymerizing factor3 is required for controlling aphid feeding from the phloem. Plant physiology 176(1):879-890.

Mori M, Nomura T, Ooka H, Ishizaka M, Yokota T, Sugimoto K, Okabe K, Kajiwara H, Satoh K, Yamamoto K, Hirochika H, Kikuchi S, 2002. Isolation and characterization of a rice dwarf mutant with a defect in brassinosteroid biosynthesis. Plant Physiology 130(3):1152-1161.

Musila RN, Diallo AO, Makumbi D, Njoroge K, 2010. Combining ability of early-maturing quality protein maize inbred lines adapted to Eastern Africa. Field crops research 119(2-3):231-237.

Pawłowski TA, Bergervoet JHW, Bino RJ, Groot SPC, 2004. Cell cycle activity and β-tubulin accumulation during dormancy breaking of Acer platanoides L. seeds. Biologia Plantarum 48(2):211-218.

Prasanna BM, Vasal SK, Kassahun B, Singh NN, 2001. Quality protein maize. Current science 1308-1319.

Ressad F, Didry D, Xia GX, Hong Y, Chua NH, Pantaloni D, Carlier MF, 1998. Kinetic Analysis of the Interaction of Actin-depolymerizing Factor (ADF)/Cofilin with G-and F-Actins comparison of plant and human adfs and effect of phosphorylation. Journal of Biological Chemistry 273(33):20894-20902.

Rhee SG, 2016. Overview on Peroxiredoxin. Molecules and cells 39(1):1–5.

Sands DC, Morris CE, Dratz EA, Pilgeram AL, 2009. Elevating optimal human nutrition to a central goal of plant breeding and production of plant-based foods. Plant Science 177(5):377-389.

Shevchenko A, Jensen ON, Podtelejnikov AV, Sagliocco F, Wilm M, Vorm O, Mortensen P, Shevchenko A, Boucherie H, Mann M, 1996. Linking genome and proteome by mass spectrometry: large-scale identification of yeast proteins from two dimensional gels. Proceedings of the National Academy of Sciences 93(25):14440-14445.

Stacy RA, Munthe E, Steinum T, Sharma B, Aalen RB, 1996. A peroxiredoxin antioxidant is encoded by a dormancy-related gene, Per1, expressed during late development in the aleurone and embryo of barley grains. Plant molecular biology 31(6):1205-1216.

Tao Y, Zheng J, Xu Z, Zhang X, Zhang K, Wang G, 2004. Functional analysis of ZmDWF1, a maize homolog of the Arabidopsis brassinosteroids biosynthetic DWF1/DIM gene. Plant Science 167(4):743-751.

Tian M, Chaudhry F, Ruzicka DR, Meagher RB, Staiger CJ, Day B, 2009. Arabidopsis actin-depolymerizing factor AtADF4 mediates defense signal transduction triggered by the Pseudomonas syringae effector AvrPphB. Plant physiology 150(2):815-824.

Tien Lea D, Duc Chua H, Quynh Lea N, 2016. Improving nutritional quality of plant proteins through genetic engineering. Current genomics 17(3):220-229.

Tran DT, Cavett VJ, Dang VQ, Torres HL, Paegel BM, 2016. Evolution of a mass spectrometry-grade protease with PTM-directed specificity. Proceedings of the National Academy of Sciences 113(51):14686-14691.

Tůmová L, Tarkowská D, Řehořová K, Marková H, Kočová M, Rothová O, Čečetka P, Holá, D, 2018. Drought-tolerant and drought-sensitive genotypes of maize (Zea mays L.) differ in contents of endogenous brassinosteroids and their drought-induced changes. PloS one 13(5):e0197870.

Wang G, Wang G, Wang J, Du Y, Yao D, Shuai B, Han L, Tang Y, Song, R, 2016. Comprehensive proteomic analysis of developing protein bodies in maize (Zea maysL.) endosperm provides

novel insights into its biogenesis. Journal of experimental botany 67(22):6323–6335.

Weidner S, Zalewski K, 1982. Changes in ribosomal proteins in wheat embryos in the course of grain development and maturation. Acta SocietatisBotanicorumPoloniae 51(2):283-290.

Weiler AM, Hergesheimer C, Brisbois B, Wittman H, Yassi A, Spiegel JM, 2014. Food sovereignty, food security and health equity: a meta-narrative mapping exercise. Health policy and planning 30(8):1078-1092.

Wilson CM, 1987. Proteins of the kernel. Corn: chemistry and technology 5:273-310.

Wu Y, Messing J, 2012. Rapid divergence of prolamin gene promoters of maize after gene amplification and dispersal. Genetics 192(2):507-519.

Wu Y, Holding DR, Messing J, 2010. γ-Zeins are essential for endosperm modification in quality protein maize. Proceedings of the National Academy of Sciences 107(29):12810-12815.

Yu T, Li G, Dong S, Liu P, Zhang J, Zhao B, 2016. Proteomic analysis of maize grain development using iTRAQ reveals temporal programs of diverse metabolic processes. BMC plant biology 16(1):241.