Journal of Proteomics xxx (2016) xxx–xxx

Contents lists available at ScienceDirect

Journal of Proteomics

journal homepage: www.elsevier.com/locate/jprot

Differential proteomics reveals the hallmarks of seed development in common bean (*Phaseolus vulgaris* L.)

J.R. Parreira [a], J. Bouraada [b], M.A. Fitzpatrick [b], S. Silvestre [c,d], A. Bernardes da Silva [c,d], J. Marques da Silva [c,d], A.M. Almeida [a,1], P. Fevereiro [a,d], A.F.M. Altelaar [b], S.S. Araújo [a,e,*]

[a] *Plant Cell Biotechnology Laboratory, Instituto de Tecnologia Química e Biológica António Xavier, Universidade Nova de Lisboa (ITQB-NOVA), Av. da República, Estação Agronómica Nacional, 2780-157 Oeiras, Portugal*
[b] *Biomolecular Mass Spectrometry and Proteomics, Bijvoet Centre for Biomolecular Research, Utrecht Institute for Pharmaceutical Sciences, University of Utrecht, Netherlands Proteomics Centre, Padualaan 8, 3584CH Utrecht, The Netherlands*
[c] *Plant Functional Genomics Group, Biosystems and Integrative Sciences Institute (BioISI), Faculdade de Ciências, Universidade de Lisboa, Campo Grande, 1749-016, Lisboa, Portugal*
[d] *Departamento de Biologia Vegetal, Faculdade de Ciências da Universidade de Lisboa, Campo Grande, 1749-016, Lisboa, Portugal*
[e] *Plant Biotechnology Laboratory, Department of Biology and Biotechnology, 'L. Spallanzani', University of Pavia, via Ferrata 1, 27100 Pavia, Italy*

ARTICLE INFO

Article history:
Received 29 January 2016
Received in revised form 29 February 2016
Accepted 1 March 2016
Available online xxxx

Keywords:
Legume
Shotgun proteomics
Seed development
Common bean

ABSTRACT

Common bean (*Phaseolus vulgaris* L.) is one of the most consumed staple foods worldwide. Little is known about the molecular mechanisms controlling seed development. This study aims to comprehensively describe proteome dynamics during seed development of common bean. A high-throughput gel-free proteomics approach (LC–MS/MS) was conducted on seeds at 10, 20, 30 and 40 days after anthesis, spanning from late embryogenesis until desiccation. Of the 418 differentially accumulated proteins identified, 255 were characterized, most belonging to protein metabolism. An accumulation of proteins belonging to the MapMan functional categories of "protein", "glycolysis", "TCA", "DNA", "RNA", "cell" and "stress" were found at early seed development stages, reflecting an extensive metabolic activity. In the mid stages, accumulation of storage, signaling, starch synthesis and cell wall-related proteins stood out. In the later stages, an increase in proteins related to redox, protein degradation/modification/folding and nucleic acid metabolisms reflect that seed desiccation-resistance mechanisms were activated. Our study unveils new clues to understand the regulation of seed development mediated by post-translational modifications and maintenance of genome integrity. This knowledge enhances the understanding on seed development molecular mechanisms that may be used in the design and selection of common bean seeds with desired quality traits.

Significance: Common bean (*P. vulgaris*) is an important source of proteins and carbohydrates worldwide. Despite the agronomic and economic importance of this pulse, knowledge on common bean seed development is limited. Herein, a gel-free high throughput methodology was used to describe the proteome changes during *P. vulgaris* seed development. Data obtained will enhance the knowledge on the molecular mechanisms controlling this grain legume seed development and may be used in the design and selection of common bean seeds with desired quality traits. Results may be extrapolated to other pulses.

1. Introduction

Common bean (*Phaseolus vulgaris* L.) seeds are one of the most consumed staple foods worldwide, with a global production estimated at around 12 million tons per year, mainly in Latin America and Africa [1]. Like other legumes, common bean seeds are rich in protein, carbohydrates, fibers, vitamins, minerals as well as other health-promoting phenolic compounds, being crucial for food security, nutrition and income source, particularly for local farmers [2–4].

Common bean is emerging as one of the best agricultural model crops to study seed biology. Its genome was recently released [5], several genomic and genetic resources are available [6], as well as, functional reverse genetics tools [7–9]. Moreover, this species has a high level of synteny with model legumes such as *Glycine max*, *Lotus japonicus* and *Medicago truncatula*, which is relevant for comparative genomic and genetic approaches [10,11]. Cumulative studies on common bean seed traits are emerging, addressing both seed phenotype markers for development, yield and quality, as well as environment interaction [3,12–17].

Seed quality traits depend on the accumulation of various storage molecules during the seed development (SD) process, which is

* Corresponding author at: Plant Cell Biotechnology Laboratory, Instituto de Tecnologia Química e Biológica António Xavier, Universidade Nova de Lisboa (ITQB-NOVA), Av. da República, Estação Agronómica Nacional, 2780-157 Oeiras, Portugal.
E-mail address: saraujo@itqb.unl.pt (S.S. Araújo).
[1] Present address: Ross University School of Veterinary Medicine, P.O. Box 334, Basseterre, Saint Kitts and Nevis.

http://dx.doi.org/10.1016/j.jprot.2016.03.002
1874-3919/

influenced in turn by the genotype and adaptive changes to environment [18]. In the first SD stage (histodifferentiation/embryogenesis) differentiation of seed embryo, endosperm and seed coat occurs concomitantly with a slow biomass accumulation rate [17]. During the second SD stage, known as maturation/filling phase, a fast continuous boost of biomass accumulation occurs and the embryo grows [17]. In the final stage, the seed starts to dehydrate, the accumulation of biomass ceases and mechanisms for embryo desiccation-tolerance are activated [19].

Seed development has been extensively studied in important crops such as rice [20,21], soybean [22–24] and maize [25,26] and model plants like Arabidopsis [27–29] and *M. truncatula* [30–35]. The *P. vulgaris* seed proteome is still poorly studied and there is a need for a comprehensive proteome resource for this species [36]. Changes in protein composition of *P. vulgaris* mature seeds lacking phaseolin were studied in 2010 by Marsolais et al. [37]. The authors described the compensatory molecular mechanisms for the absence of the major storage proteins, resulting in an increase in sulphur amino acid contents. A mature seed proteome map of *P. vulgaris*, usable as a reference for 2-DE assays was published in 2011 by De La Fuente [38]. In 2013, a more comprehensive proteome map of *P. vulgaris* mature seeds, identifying and functionally categorizing 141 protein spots, was published by Natarajan [39]. Nevertheless, the temporal changes occurring in the proteome during common bean seed development still remain poorly addressed.

Herein, a high-throughput gel-free liquid chromatography–tandem mass spectrometry (LC–MS/MS) proteomics study was conducted to extend our knowledge on the molecular mechanisms underlying SD in the common bean. The broadness, comprehensiveness and high-throughput nature of LC–MS/MS (for a recent review on label-free proteomics methodology, refer to Jorrín-Novo [40]) are advantages over 2-DE approaches. Four time-points were analyzed, spanning from late embryogenesis until desiccation stage. Specific temporal protein accumulation patterns for each time point were described, providing a comprehensive seed proteome atlas with new information for future development and selection of new common bean seeds with desired quality traits.

2. Materials and methods

2.1. Plant material

In this study, *P. vulgaris* seeds from the SER 16 genotype were kindly supplied by Dr. Steve Beebe from the International Center for Tropical Agriculture (CIAT-CGIAR, Cali, Colombia). Seeds were placed in water soaked paper in Petri dishes at 27 °C for 2 days, followed by 3 days at 23 °C, always in the dark. Seedlings were transferred to watered vermiculite trays where they were allowed to grow under controlled conditions: 50% humidity, photoperiod of 16/8-h day/night at 25/18 °C, respectively, and light intensity of 400 μmol m^{-2} s^{-1} (Fitoclima 5.000 EH, ARALAB, Portugal). A week later, seedlings were individually transferred to 2.5 l pots with a (3:1) mixture of commercial soil (Compo Sana S.A., Barcelona, Spain) and vermiculite, respectively. SER16 flowers were tagged and pods/seeds were harvested at 10, 20, 30 and 40 days after anthesis (DAA). Plants were kept in the environmental conditions described above during the assay and were watered 3 times per week. Harvested seeds were divided into two batches: one immediately frozen in liquid nitrogen and stored at −80 °C until further analysis and another one was used to measure seed length, fresh and dry weight to characterize the seed development process. Four biological replicates (individual plants) per seed developmental stage were used.

2.2. Protein sample extraction

One hundred milligrams of seed material were ground to a fine powder in liquid nitrogen using mortar and pestle. Ground material was solubilized in a buffer containing 8 M urea, 50 mM triethylammonium bicarbonate containing one tablet of cOmplete, Mini, EDTA-free Protease Inhibitor Cocktail Tablets in EASYpacks (Roche Diagnostics GmbH, Mannheim, Germany). The samples were centrifuged (20,000 × *g* for 10 min at 4 °C) and the supernatants were reduced with 2 mM dithiothreitol (Fluka, Buch, Switzerland) for 30 min at 56 °C and alkylated with 4 mM iodoacetamide (Sigma-Aldrich, Steinheim, Germany) for 30 min at room temperature in the dark. The proteins were then digested with Lys-C (enzyme/substrate ratio 1:100) for 4 h at 37 °C. Subsequently trypsin was added for an overnight digestion at 37 °C (enzyme/substrate ratio 1:100) (Promega, Madison, WI, USA). Peptides were separated and analyzed using a LC–MS/MS system with a label-free approach.

2.3. LC–MS/MS and data analysis

Samples were analyzed on a LTQ-Orbitrap Elite mass spectrometer (Thermo Scientific, Bremen, Germany) at The Netherlands Proteomic Center (Utrecht, Netherlands).

After being re-suspended in 10% formic acid, peptides were analyzed using nanoflow reversed phase liquid chromatography on a Proxeon Easy-nLC 100 (Thermo Scientific, Bremen, Germany) connected to an LTQ-Orbitrap Elite mass spectrometer (Thermo Scientific, Bremen, Germany). The peptides were first trapped with 100% buffer A (0.1% formic acid in water) on an in-house-made double-fritted trap column (Reprosil C18, 3 μm, 2 cm × 100 cm, Dr. Maisch, GmbH Ammerbuch, Germany) before being separated on an in-house-made analytical column (2.7 μm, 50 μm × 50 cm, Agilent Zorbax SB-C18, Amstelveen, The Netherlands). The peptides were separated with a gradient of an aqueous (water, A) and an organic (acetonitrile, B) solvent, both with 0.1% formic acid. The 120 min gradient starts with 7% buffer B and goes up to 30% in 90 min and then to 100% B in 30 min at a flow rate of 150 nl/min. After the separation, peptides were introduced through a nanoelectrospray ion source by using an in-house-made fused silica needle (10 μm inner diameter tip) applied at a voltage of 1.7 kV and analyzed using an Orbitrap Elite mass spectrometer. Full-scan MS spectra from *m/z* 375 to 1600 were acquired in the Orbitrap at a resolution of 60,000. The 20 most intense ions were selected for collision-induced fragmentation in the linear ion trap at a normalized collision energy of 35%. The MaxQuant software (version 1.5.2.8) was used for the analysis of raw data [41,42].

The data was searched against the *P. vulgaris* L. UniProt FASTA database (version of June 3rd, 2015): (http://www.uniprot.org/uniprot/?query=*&fil=organism:"*Phaseolus vulgaris* (Kidney bean) (French bean) [3885]"). The following parameters were used for database search: initial tolerance of 20 ppm and a final mass tolerance of 7 ppm for precursor masses, 0.6 Da for CID ion trap fragment ions, and two missed cleavages allowed. The thresholds used were as follows: Minimum number of peptides: one, Orbitrap MS/MS match tolerance: 20 ppm, Orbitrap MS/MS de novo tolerance: 10 ppm, Orbitrap MS/MS de-isotoping tolerance: 7 ppm, Orbitrap top peaks per 100 Da: 12. Ion trap MS/MS match tolerance: 0.5 Da. Ion trap MS/MS de-isotoping tolerance: 0.15 Da, ion trap top peaks per 100 Da: eight, maximum peptide mass: 4600 Da, minimum peptide length for unspecific search: eight, maximum peptide length for unspecific search is 25. Carbamidomethylation was used as a fixed modification, and protein *N*-terminal acetylation and oxidation as a variable modification. The false discovery rate was set at 1% for peptide and proteins. The minimum peptide length specified was seven amino acids.

Data acquired from MaxQuant was analyzed using Perseus software (version 1.5.1.6). Intensities normalization was performed by subtracting the median of the log-transformed intensities for each nano-LC–MS/MS run. A two-sided two-sample *t*-test was performed with a permutation-based false discovery rate of 0.05. The *t*-test was performed using four biological replicates with comparisons between each subsequent time point (10 vs. 20, 20 vs. 30, 30 vs. 40). Results

Please cite this article as: J.R. Parreira, et al., Differential proteomics reveals the hallmarks of seed development in common bean (*Phaseolus vulgaris* L.), J Prot (2016), http://dx.doi.org/10.1016/j.jprot.2016.03.002