

Back to Osborne. Sequential Protein Extraction and LC-MS Analysis for the Characterization of the Holm Oak Seed Proteome

M. Cristina Romero-Rodríguez, Ana M. Maldonado-Alconada, Luis Valledor, and Jesus V. Jorrin-Novo

Abstract

It is impossible to capture in just one experiment all or most of the total set of protein species that constitute the cell's proteome. Thus, according to our results, and even considering that they depend on the experimental system carried out (plant, yeast, fungi, or bacteria), the best protein extraction protocol yielded less than 20 % of the total amount of proteins, as determined by the Kjeldahl method. For this reason, protein cataloguing and the whole proteome characterization require the use of firstly, fractionation techniques at the cellular, subcellular, protein, or peptide level, and secondly, the use of complementary approaches.

Within our current research on Holm oak (*Quercus ilex* subsp. *ballota*), we aim to characterize its seed proteome. For that we have optimized an experimental workflow in which the Osborne sequential protein extraction (Osborne, Science 28:417–427, 1908) is combined with downstream electrophoretic protein separation or shotgun MS analysis. In general, it can be used to study any plant seed, as well as to investigate on seed maturation and germination, genotype characterization, allergens identification, food traceability, and substantial equivalence, among others.

Key words Holm oak proteomics, Seed proteomics, Osborne

1 Introduction

Osborne (1924) classified seed proteins into four categories according to their solubility in various solutions: (1) albumins, soluble in water, which primarily correspond to enzymes; (2) globulins, insoluble in water but soluble in salt solutions (i.e., 1 M NaCl); (3) prolamins, soluble in ethanol; and (4) glutelins, which can be extracted with either very acidic or very basic solutions [1, 2].

The major seed storage proteins are albumins, globulins, and prolamins. All storage protein fractions are mixtures of components that exhibit polymorphism both within single genotypes and among genotypes of the same species according to species [3]. These protein classes are not present systematically in the seeds of all

plant species. For example, prolamins exist only in the herbaceous plant species and cereals as the major storage proteins, whereas only fractions of albumin, globulin, and glutelin are found in sugarbeet [4].

Eighteen years after the Osborne work, in 1994, Marc R. Wilkins coined the term proteome [5, 6]; this being the beginning of a paradigm shift in the protein studies from protein chemistry/biochemistry to proteomics. This change was possible thanks to the progress in genome sequencing projects, the applicability of mass spectrometry to the analysis of peptide and proteins, and the development of bioinformatic tools for protein identification and data analysis. The potential of proteomics in plant biology research is extensively validated with the number of original papers and revisions, some of them dealing with plant seeds, devoted to different purposes: descriptive studies, cataloguing of genotypes, phylogenetic studies, seed germination, development, and responses against biotic and abiotic stresses [7–12]. The full potential of proteomics, however, is far from being fully exploited in biological research, due to the inherent difficulty of research with proteins, together with techniques and equipment limitations. We should highlight that the number of protein species is much higher than the number of genes, they are quite different physicochemically from each other, its dynamic range is of up nine orders of magnitude, and there is no possibility of being amplified, as it occurs with DNA. For all the above reasons, it is impossible to capture in just one experiment all or most of the total set of protein species that constitute the cell's proteome. Thus, according to our results, and considering that they depend on the experimental system carried out (plant, yeast, fungi, or bacteria), the best protein extraction protocol yielded less than 20 % of the total amount of proteins as determined by the Kjeldahl method [13]. For this reason, the proteome characterization and the protein species cataloguing require the use of firstly fractionation techniques at the cellular, subcellular, protein or peptide level; and secondly, the use of complementary approaches.

We have optimized for Holm oak seeds an experimental workflow in which the Osborne sequential protein extraction is combined with downstream electrophoretic separation or shotgun MS analysis. Up to now, most analyses of plant seed proteomic have been based on a total protein extract, and also with some examples of a similar strategy [4, 14–17].

In our current project on *Quercus ilex* subsp. *ballota*, we aim to obtain a full description of the seed proteome, in order to: (1) search for proteins of practical interest (i.e., enzymes); (2) identify possible allergens; (3) characterize the acorn nutritional quality; (4) catalog populations; and (5) carry out studies of seed germination and responses to stresses. The optimized workflow is schematized in Fig. 1.

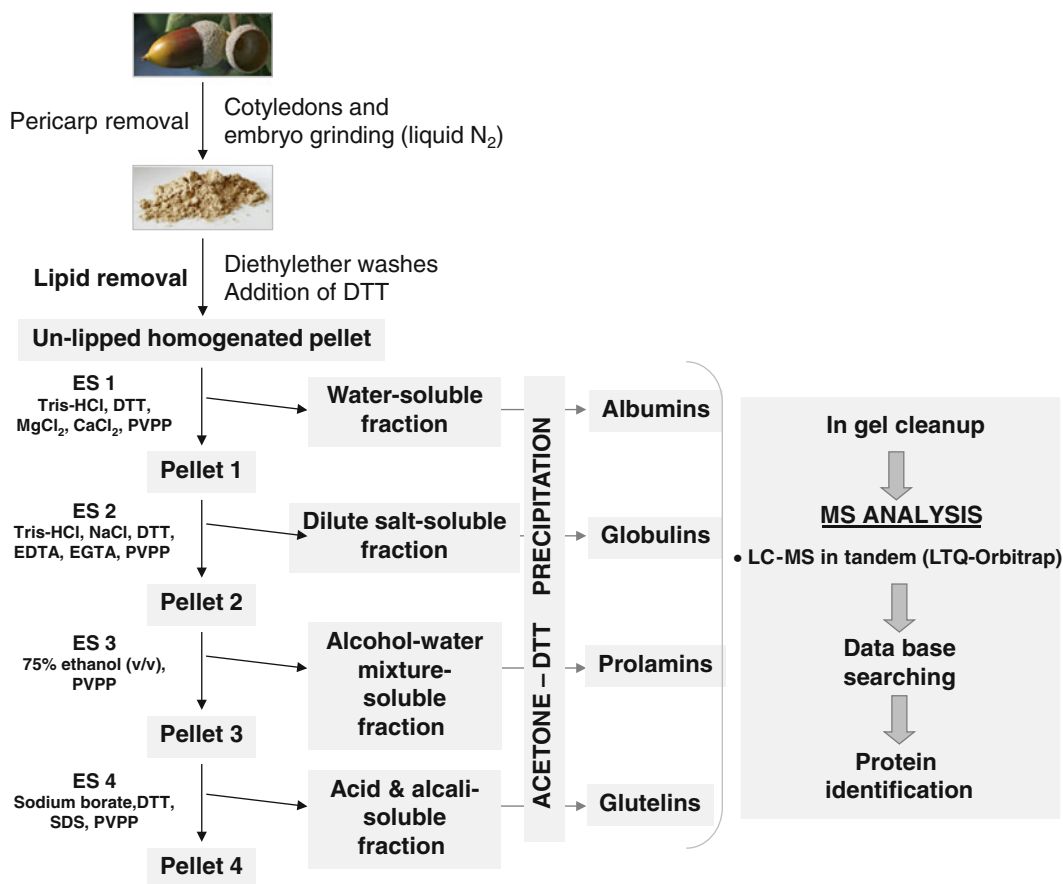


Fig. 1 Flowchart of the protein extraction procedure and MS analysis

2 Materials

2.1 Equipment (See Note 1)

1. Analytical Mill, Model A-10 (115 V), Janke & Kunkel IKA Labortechnik.
2. Mortar, pestle, and knife.
3. *Ultrasonic Homogenizer 4710 series*, Cole-Parmer Instrument Co.
4. Refrigerated centrifuge Model J2-21; Beckman.
5. Microcentrifuge Model 5415 D, Eppendorf.
6. BioPhotometer, Model 22331, Eppendorf.
7. Vertical Electrophoresis equipment. Mini PROTEAN II, Bio Rad.
8. Power supply, PowerPac HC, Bio-Rad

2.2 Reagents and Solutions

Prepare all solutions in distilled water (conductivity lower than $3 \mu\text{S}\cdot\text{cm}^{-1}$). Store all solutions at 4°C (unless otherwise indicated).

1. Extraction solution for albumins (ES1): 10 mM Tris-HCl, pH 7.5, 0.1 % (w/v) DTT, 10 mM MgCl_2 , 10 mM CaCl_2 .
2. Extraction solution for globulins (ES2): 10 mM Tris-HCl, pH 7.5, 1 M NaCl, 0.1 % (w/v) DTT, 10 mM EDTA, 10 mM EGTA.
3. Extraction solution for prolamins (ES3): 75 % (v/v) ethanol.
4. Extraction solution for glutelins (ES4): 50 mM sodium borate buffer, pH 10, 2 % (w/v) DTT, 1 % (w/v) SDS. Store this solution at room temperature.
5. Solubilization solution: 7 M urea, 2 M thiourea, 4 % (w/v) CHAPS, 2 % (v/v) Tritón-X-100, 100 mM DTT. Add DTT to 10 mM just before use.
6. Bradford reagent (Sigma-Aldrich).
7. SDS-PAGE running buffer: 25 mM Tris-HCl, pH 8, 192 mM glycine, 0.1 % (w/v) SDS.
8. Acrylamide/Bis-acrylamide solution (30 %) commercial solution (Bio-Rad).
9. Ammonium persulfate: 10 % (w/v).
10. SDS: 10 % (w/v).
11. Resolving gel buffer: 1.5 M Tris-HCl, pH 8.8.
12. Stacking gel buffer: 0.5 M Tris-HCl pH 6.8.
13. Colloidal Coomassie staining gel: 10 % (w/v) ammonium sulfate, 3 % (w/v) phosphoric acid, 0.1 % (w/v) Coomassie Brilliant Blue (CBB) G250, 20 % (v/v) methanol.
14. Tris-phosphoric solution: 0.1 M Tris-phosphoric acid, pH 6.5.
15. Ammonium sulfate: 20 % (w/v).
16. Methanol: 25 % (v/v).
17. TEMED, commercial product (Sigma) (*see* **Note 2**).

3 Methods

3.1 Plant Material

Holm oak mature acorns (with brown pericarp) were used in our research. Once harvested, acorns were selected, cleaned and stored in airtight polyethylene bags at $4 \pm 1^\circ\text{C}$. Acorns were scarified with a knife by making transversal and longitudinal cuts, thus permitting the pericarp to be rapidly removed. Once peeled off, cotyledons (including embryo axes) were triturated in an analytical mill to obtain a homogeneous fine flour (Fig. 1). The obtained powder was weighed and stored in a desiccator at $4 \pm 1^\circ\text{C}$, until protein extraction [18].

3.2 Protein Extraction (See Note 3)

3.2.1 Sample Delipidation

1. Grind acorn flour (10 g) in liquid nitrogen with a precooled mortar and pestle until a fine powder is formed (*see Note 4*).
2. Add 60 mL of diethylether and shake the mixture for 1 h at 4 °C (*see Note 5*).
3. Centrifuge at 15,000×*g* for 15 min (4 °C) and decant the supernatant (*see Note 6*).
4. Repeat **steps 2 and 3** two more times.
5. Add 60 mL of acetone containing 0.07 % (w/v) DTT to the above pellet, and shake the mixture for 10 min at 4 °C (*see Note 7*).
6. Centrifuge at 15,000×*g* for 15 min (4 °C) and decant the supernatant (*see Note 8*).
7. Allow the pellet to dry at room temperature until no traces of acetone remained.

3.2.2 Solubilization of the Albumin Fraction

1. Add 50 mL of ES1 and 1 g of PVPP to the pellet obtained in the previous step.
2. Sonicate the mixture (while keeping on ice) for 15 s (50 W, amplitude 60), and shake in an orbital shaker for 2 h at 4 °C.
3. Centrifuge at 15,000×*g* for 15 min (4 °C) and transfer the supernatant to a new tube (*see Note 9*).
4. Precipitate proteins in the supernatant by adding four volumes (approximately 100 mL for each tube) of 0.07 % (w/v) DTT containing cold acetone, and keep 1 h at 20 °C.
5. Centrifuge at 15,800×*g* for 15 min (4 °C) and decant the supernatant.
6. Dry the pellet at room temperature and store it at –80 °C if not immediately used (*see Notes 10 and 11*).

3.2.3 Solubilization of the Globulin Fraction

1. Add 50 mL of ES2 to the pellet obtained in **step 3** (Subheading 3.2.2).
2. Sonicate the mixture (while keeping on ice) for 15 s (50 W, amplitude 60), and shake in an orbital shaker for 2 h at 4 °C.
3. Centrifuge at 15,000×*g* for 15 min (4 °C) and transfer the supernatant to a new tube (*see Note 9*).
4. Repeat **steps 4–6** of the albumin fraction solubilization to obtain the globulin fraction (*see Note 12*).

3.2.4 Solubilization of the Prolamin Fraction

1. Add 50 mL of ES3 to the pellet obtained in **step 3** (Subheading 3.2.3).
2. Sonicate the mixture (while keeping on ice) for 15 s (50 W, amplitude 60), and shake in an orbital shaker for 2 h at 4 °C.

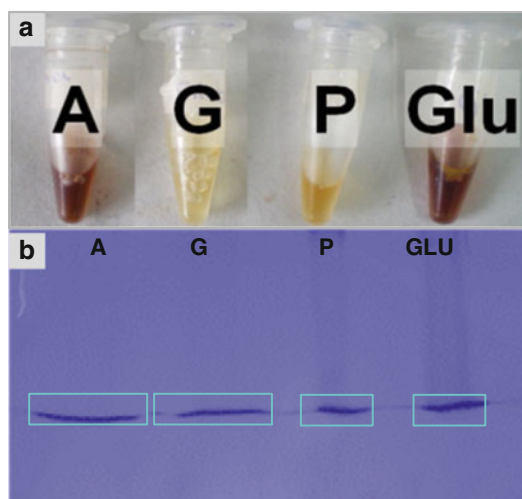


Fig. 2 (a) Aspects of each protein fraction obtained. *A* Albumins, *G* Globulins, *P* Prolamins, *Glu* Glutelins. (b) Gel precleaning. 100 μ g of each fraction (Albumins, A; Globulins, G; Prolamins, P; Glutelins, Glu) obtained with the protein were cleaned in a 5 % SDS-PAGE that was further stained with Coomassie Brilliant Blue. Labeled proteins bands were cut from the gel and analyzed by LC coupled to MS/MS

3. Centrifuge at $15,000\times g$ for 15 min (4°C), and transfer the supernatant to a new tube (*see Note 9*).
4. Repeat **steps 4–6** of the albumin fraction solubilization to obtain the prolamin fraction.

3.2.5 Solubilization of the Glutelin Fraction

1. Add 50 mL of ES4 to the pellet obtained in the **step 3** (Subheading 3.2.4).
2. Sonicate the mixture (while keeping on ice) for 15 s (50 W, amplitude 60), and shake in an orbital shaker for 2 h at room temperature.
3. Centrifuge at $15,000\times g$ for 15 min (4°C), and transfer the supernatant to a new tube (*see Note 9*).
4. Repeat **steps 4–6** of the albumin fraction solubilization to obtain the pellet corresponding to the glutelin fractions (*see Note 13*). (*see Note 14*)

3.3 Protein Solubilization

In all cases dried pellets resulting from each protein fraction precipitation were weighed and suspended in 7 M urea, 2 M thiourea, 4 % (w/v) CHAPS, 2 % (v/v) Triton X-100, 100 mM DTT at a (w/v) ratio of 1:2 (*see Note 15*). Then, they were shaken for 1 h to facilitate protein solubilization. Once the pellet was solubilized and the insoluble material eliminated by centrifugation (*see Note 16*) (Fig. 2a), the protein content was quantified by the Bradford methods [19], using bovine albumin as standard.

3.4 Protein Sample Preparation for LC-MS

One hundred μg of proteins corresponding to the different fractions were subjected to one-dimensional, denaturing SDS-PAGE according to Laemmli [20], with minor modifications. We used 7 cm stacking (5 % polyacrylamide) and 3 cm resolving (12.5 % polyacrylamide) gels, these run at 80 V in a Mini PROTEAN II cell (*see* Note 17) (Fig. 2b). Gel was stained by colloidal Coomassie [21].

Bands were excised using a clean scalpel, transferred to Multiwell 96 plates and digested with modified porcine trypsin by using a ProGest (Genomics Solution) digestion station. The digestion protocol used is that of Schevchenko et al. [22], with minor variations. Gel plugs were destained by incubation (twice for 30 min) with 200 mM ammonium bicarbonate in 40 % ACN at 37 °C, then subjected to three consecutive dehydration–rehydration cycles with pure acetonitrile and 25 mM ammonium bicarbonate in 40 % ACN, respectively; and, finally, dried at room temperature for 10 min. Trypsin, at a final concentration of 12.5 ng/ μL in 25 mM ammonium bicarbonate, was added to the dry gel pieces and the digestion proceeded at 37 °C for 12 h. Peptides were extracted from gel plugs with 0.5 % TFA, 50 % ACN (15 min incubation), dried under vacuum and resuspended in 0.1 % TFA.

3.5 LC-MS and Protein Identification

Tryptic peptides (obtained according Subheading 3.4) were injected into the LC system (i.e., a Finnigan Surveyor HPLC system). Peptides were detected in a LTQ-Orbitrap equipped with a nanoelectrospray ion source (nESI). The acquired data can be analyzed with Proteome Discoverer v1.3 software (Thermo Fisher Scientific, USA) and MASCOT (<http://www.matrixscience.com>) or SEQUEST (<http://fields.scripps.edu/sequest/>) algorithms. Protein identification was conducted by combining search (MS plus MS/MS) to the entries of a nonredundant protein Viridiplantae database downloaded from the National Center for Biotechnology Information. Search parameters were the enzyme trypsin, taxonomy restrictions to Viridiplantae, ± 50 ppm peptide mass tolerance in MS, ± 0.8 Da for MS/MS data and, carbamidomethyl (C) as a fixed modification, whereas methionine oxidation as a variable modification (*see* Note 18).

3.6 Main Results

Considering the particular composition of the acorn flour, rich in non-protein material [23, 24], the main challenge was to obtain protein extracts of enough quality for the subsequent analyses and identification of proteins by MS. In Holm oak acorns, glutelin was the major fraction, corresponding approximately to 80 % of the total protein extracted with this protocol.

The analysis by LC coupled with MS/MS allowed the identification of 509 protein species, belonging to the following functional categories: storage, energy production and metabolism, defense, antioxidant system, signaling and gene expression regulation,

DNA and protein metabolism, cell wall biogenesis, cytoskeleton and, cellular cycle and translocation. The percentage of total proteins identified which were present in one fraction was: 25, 23, 15, and 23 % for the albumin, globulin, prolamin, and glutelin fractions, respectively. We have established the first *Q. ilex* acorn reference proteome that can be used in subsequent analysis, namely, the identification of protein markers to discriminate between populations, studies on the germination process, etc.

The main limitation of this approach is that the genus *Quercus* is an orphan plant species poorly characterized at the genetic and physiological levels. The low number of well annotated genomic sequences available for genus *Quercus*, and for *Q. ilex* in particular, in public databases limits the confident protein identification. Moreover, most software tools have been developed to search in a non-error tolerant way against a database of known proteins [10, 23]. A higher number of identifications which are more accurate would be obtained by searching against EST sequences [25, 26]. These sequences can be obtained from the following EST public database (<http://compbio.dfci.harvard.edu/tgi/>; <http://www.fagaceae.org/>; <http://www.ncbi.nlm.nih.gov/nucest/>); currently, our group is constructing an EST database for *Quercus*.

4 Notes

1. References to specific companies, equipment or pieces of equipment are not mandatory and do not represent an endorsement by the authors
2. Standard reagents and items of current use, such as Bovine serum albumin (BSA), liquid nitrogen and polyvinylpyrrolidone, are not specified.
3. Protein extraction and fractionation procedures were based on the differential solubility criteria described by Osborne (1924), and later adapted to *Quercus* subsp. by Fonseca et al. (1997) and Martin et al. (2009) [2, 27, 28].
4. The extraction is performed with 10 g of fresh tissue because fractions of albumins, globulins and prolamins are minority proteins, with 1 g of tissue the small amount of protein isolated in the mentioned fraction is not enough to analyse.
5. This step is to reduce lipid content, which allows the obtaining of a higher quality extract for subsequent analysis.
6. Be careful that the pellet does not stick tightly to the bottom of the tube.
7. This step is to remove traces of diethylether.

8. Remove the supernatant with a pipette to avoid losing the pellet. Be careful: do not move the pellet because it is very easy to disintegrate
9. Remove the supernatant with a pipette to avoid disaggregating the pellet and carry contaminant.
10. The pellet of the albumins fraction appearance may be viscous; this aspect is normal, indicating that it may contain the remainders of lipids and polysaccharides [29].
11. Dry the pellet to ensure that the acetone was completely removed, if it lasts for a long time is difficult to resuspend the protein.
12. The pellet of the globulins fraction has two phases, namely, one solid and another oily. Remove the last one and use the solid phase to resuspend proteins.
13. The glutelins solubilization is performed at room temperature to prevent the insolubilization of SDS.
14. **Steps 1–3** of each fraction solubilization should be repeated in order to make sure that the maximum amount of proteins corresponding to each fraction are solubilized.
15. The volume of buffer to resuspend proteins depends on the amount of precipitate obtained. It is recommended that samples are well concentrated.
16. The insoluble material is abundant in the case of the glutelin fraction.
17. The extracts are cleaned in this step, removing contaminants that may interfere with the identification by MS such a salt, detergents, etc.
18. For more details you can consult the following review [30].

Acknowledgments

This work has been financed by the Spanish Ministry of Innovation and Science, project AGL2009-12243-C02-02 “Variability, cataloguing, responses to stresses and clonal propagation of holm oak (*Quercus ilex* subsp. *ballota*) (DECOVA)” (FEDER Co-financed) and the Andalusian Government (Junta de Andalucía). MC Romero Rodríguez has been supported by a grant from the “Fundación Parque Tecnológico Itaipu-Binacional” and “Centro Multidisciplinario de Investigaciones Tecnológicas - Dirección General de Investigación Científica y Tecnológica-Universidad Nacional de Asunción - Paraguay”.

References

1. Osborne TB (1908) Our present knowledge of plant proteins. *Science* 28:417–427
2. Osborne TB (1924) The vegetable proteins. Logmans, London
3. Shewry PR, Napier JA, Tatham AS (1995) Seed storage proteins: structures and biosynthesis. *Plant Cell* 7:945–956
4. Catusse J, Pestsova E, Bechetoille C et al (2007) Betaomics: a combined proteome and transcriptome profiling approach to characterize seed germination and vigour in sugarbeet seeds. In: Adkins SW, Ashmore S, Navie SC (eds) *Seeds: biology, development and ecology*. CAB International, London, pp 169–178
5. Wilkins M (2009) Proteomics data mining. *Expert Rev Proteomics* 6:599–603
6. Wilkins MR, Pasquali C, Appel RD et al (1996) From proteins to proteomes: large scale protein identification by two-dimensional electrophoresis and amino acid analysis. *Nat Biotechnol* 14:61–65
7. Catusse J, Rajjou L, Job C et al (2008) Proteome of Seed Development and Germination. In: Agrawal G, Rakwal R (eds) *Plant Proteomics: Technologies, Strategies, and Applications*. John Wiley & Sons, Inc., pp 191–206
8. González-Fernández R, Prats E, Jorrín-Novo JV (2010) Proteomics of plant pathogenic fungi. *J BiomedBiotechnol*. doi:[10.1155/2010/932527](https://doi.org/10.1155/2010/932527)
9. Hossain Z, Nouri MZ, Komatsu S (2012) Plant cell organelle proteomics in response to abiotic stress. *J Proteome Res* 11:37–48
10. Jorrín-Novo JV, Maldonado AM, Echevarría-Zomeño S et al (2009) Plant proteomics update (2007–2008): second-generation proteomic techniques, an appropriate experimental design, and data analysis to fulfill MIAPE standards, increase plant proteome coverage and expand biological knowledge. *J Proteomics* 72:285–314
11. Jorrín-Novo JV, Maldonado AM, Castillejo MA (2007) Plant proteome analysis: a 2006 update. *Proteomics* 7:2947–2962
12. Rossignol M, Peltier JB, Mock HP et al (2006) Plant proteome analysis: a 2004–2006 update. *Proteomics* 6:5529–5548
13. Valero-Galván J, Jorrín-Novo JV, Cabrera A et al (2012) Population variability based on the morphometry and chemical composition of the acorn in Holm oak (*Quercus ilex* subsp. *ballota* [Desf.] Samp.). *Eur J For Res* 131:893–904
14. Agrawal GK, Hajdуч M, Graham K et al (2008) In-depth investigation of the soybean seed-filling proteome and comparison with a parallel study of rapeseed. *Plant Physiol* 148: 504–518
15. Hajdуч M, Ganapathy A, Stein JW et al (2005) A systematic proteomic study of seed filling in soybean. establishment of high-resolution two-dimensional reference maps, expression profiles, and an interactive proteome database. *Plant Physiol* 137:1397–1419
16. Krishnan HB, Oehrle NW, Natarajan SS (2009) A rapid and simple procedure for the depletion of abundant storage proteins from legume seeds to advance proteome analysis: a case study using *Glycine max*. *Proteomics* 9:3174–3188
17. Miernyk JA, Hajdуч M (2011) Seed proteomics. *J Proteomics* 74:389–400
18. Valero-Galván J, Villedor L, Cerrillo RM et al (2011) Studies of variability in Holm oak (*Quercus ilex* subsp. *ballota* [Desf.] Samp.) through acorn protein profile analysis. *J Proteomics* 74:1244–1255
19. Bradford MM (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal Biochem* 7: 248–254
20. Laemmli UK (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* 227:680–685
21. Neuhoff V, Arold N, Taube D et al (1988) Improved staining of proteins in polyacrylamide gels including isoelectric focusing gels with clear background at nanogram sensitivity using Coomassie Brilliant Blue G-250 and R-250. *Electrophoresis* 9:255–262
22. Schevchenko A, Wilm M, Vorm O, Mann M (1996) Mass spectrometric sequencing of proteins silver-stained polyacrylamide gels. *Anal Chem* 68:850–858
23. Abril N, Gion JM, Kerner R et al (2011) Proteomics research on forest trees, the most recalcitrant and orphan plant species. *Phytochemistry* 72:1219–1242
24. Bonner FT, Vozzo JA (1987) *Seed biology and technology of Quercus*. Gen Tech Rep SO-66. New Orleans
25. Choudhary JS, Blackstock WP, Creasy DM et al (2001) Matching peptide mass spectra to EST and genomic DNA databases. *Trends Biotechnol* 19:S17–S22
26. Edwards NJ (2007) Novel peptide identification from tandem mass spectra using ESTs and sequence database compression. *Mol Syst Biol* 3:102. doi:[10.1038/msb4100142](https://doi.org/10.1038/msb4100142)

27. Fonseca PA, Ferreira RB, Teixeira AR (1997) Seed proteins from *Quercus suber*. *J Agric Food Chem* 45:3443–3447
28. Martin MA, Navarro-Cerrillo RM, Ortega P et al (2009) The use of cotyledon proteins to assess the genetic diversity in sweet holm oak. *J Fort Sci* 55:526–531
29. Görg A, Weiss W, Dunnet MJ (2004) Current two-dimensional electrophoresis technology for proteomics. *Proteomics* 4:3665–3685
30. Schulze WX, Usadel B (2010) Quantitation in mass-spectrometry-based proteomics. *Annu Rev Plant Biol* 61:491–516