Review

Discovering seed proteome: challenges and perspectives

Galina Smolikova ¹, Elena Lukasheva ², Gregory Mavropolo-Stolyarenko ², Tatiana Bilova ^{1,3}, Alena Soboleva ^{2,3}, Alexander Tsarev ^{2,3}, Ekaterina Podolskaya ^{4,5}, Vladimir Zhukov ⁶, Igor Tikhonovich ^{6,7}, Sergei Medvedev ¹, Wolfgang Hoehenwarter ⁸ and Andrej Frolov ^{2,3*}

- ¹ Department of Plant Physiology and Biochemistry, St. Petersburg State University;
 - ² Department of Biochemistry, St. Petersburg State University;
 - ³ Department of Bioorganic Chemistry, Leibniz Institute of Plant Biochemistry;
 - ⁴ Institute of Analytical Instrumentation, Russian Academy of Science;
 - ⁵ Institute of Toxicology, Russian Federal Medical Agency;
 - ⁶ All-Russia Research Institute for Agricultural Microbiology;
 - ⁷ Department of Genetics and Biotechnology, St. Petersburg State University;
 - ⁸ Proteome Analytics Research Group, Leibniz Institute of Plant Biochemistry
- * Correspondence: andrej.frolov@ipb-halle.de; Tel.: +49(0)34555821350

Abstract: For centuries, crop plants have represented the basis of the daily human diet. Among them, cereals and legumes, accumulating oils, proteins and carbohydrates in their seeds, distinctly dominate modern agronomic practice. Indeed, these plants play an essential role in the food industry and fuel production. Therefore, the seeds of crop plants are intensively studied by food chemists, biologists, biochemists, and nutritional physiologists. Accordingly, not only seed development and germination, but also age- and stress-related alterations in seed vigor, longevity, nutritional value and safety can be addressed by a broad panel of analytical, biochemical and physiological methods. Currently, functional genomics is one of the most powerful tools, giving direct access to characteristic metabolic changes, accompanying plant development, senescence and response to biotic or environmental stress. Among individual methodological platforms, proteomics represents one of the most effective ones, giving access to cellular metabolism at the level of proteins. Here we discuss the main methodological approaches employed by seed proteomics in the context of physiological changes related to seed development, ageing and response to environmental stress.

Keywords: Food safety; gel-based proteomics; LC-based proteomics; post-translational modifications; proteomics; seed ageing; seed quality

1. Introduction

Seeds represent the basis of the human diet and contribute daily to consumed foods [1]. Hence, the rapidly growing human population requires a secured, continuous supply with foods of appropriate quality and safety [2]. Because of this, modern agriculture aims at sustainable production of high-quality seeds. On the one hand, due to the ongoing climate changes, it can be a challenging task [3]. On the other, improvement of crop plant productivity is desired [4,5]. Hence, the yields of crop biomass and tolerance to environmental stress need to be increased together. For this, understanding of the fundamental processes accompanying seed development, storage and germination, is required.

During development, seeds pass through the stages of embryogenesis, filling, and late maturation [6]. Importantly, the last steps of seed maturation are accompanied with development of desiccation tolerance and onset of dormancy – events, affecting resistance of seeds to stress and

storage, i.e. having a great impact on their quality [7]. Indeed, environmental stressors, like drought, high salinity and high light, directly affect maturation of seeds and might compromise their longevity, i.e. ability to maintain viability during storage [8]. The most probable reason for this is the general similarity of the adjustments of metabolic events that accompany seed maturation and stress-driven dehydration, which contribute to the majority of plant responses to environmental stress [9]. Thus, similarly to drought response [10], tissue dehydration at the late stages of seed maturation is hallmarked by accumulation of late embryogenesis abundant (LEA) proteins and carbohydrates (predominantly, oligosaccharides of raffinose family) in parallel to degradation of embryonic chlorophylls [11–15]. Accordingly, these effects are also mediated by abscisic acid (ABA) [6, 13, 16, 17] and the transcription factor ABSCISIC ACID INSENSITIVE5 (ABI5) as the main regulator [18].

One of the key properties of seeds, directly affecting their quality, is vigor, i.e. their ability to germinate and produce normally developed seedlings under varying environmental conditions [19]. According to International Seed Testing Organization rules [20], characterization of seed vigor relies on several markers: germination rate, seedling morphology, developmental kinetics and integrity of cellular membranes [21]. Biochemically, seed vigor can be characterized by a range of oxidative stress markers, i.e. up-regulation of reactive oxygen species, like the superoxide anion ($O_2^{\bullet-}$), hydrogen peroxide (H_2O_2), and the hydroxyl radical ($^{\bullet}OH$), accompanied by a decrease in the activities of antioxidant enzymes and ascorbate and glutathione contents [20, 21].

In general, kinetics of seed germination and seedling development greatly depend on the metabolic status of the plant. Biomolecular reserves (polysaccharides, oils and proteins), accumulated in seeds during their maturation, play a key role in both processes [24–26]. The chemical nature of these biomolecules can differ essentially between plant species. Thus, in cereals, the major seed storage tissue is endosperm, where starch can account up to 87 % of the total dry seed weight [24, 27]. In contrast, in legumes, the principle storage organs are cotyledons, which, for example, in soybean, can contain up to 40 and 20% (w/w) proteins and oils, respectively [27]. Therefore, due to high constitutive levels of protein biosynthesis, legume seeds represent an excellent platform for the large-scale production of transgenic proteins [28–30]. As legume seeds are widely used for human nutrition and animal feeding [31], their development attracts the special interest of plant biologists. The model plants, most commonly used for this are pea (*Pisum sativum*), faba bean (*Vicia faba*), soybean, lotus (*Lotus japonicus*) and annual barrel medic (*Medicago truncatula*) [24].

In this context, knowledge of the legume seed proteome, as well as its dynamics in response to environmental and biological stressors, may significantly impact the improvement of their quality and nutritional properties [32]. Therefore, such a versatile functional genomics tool as proteomics is widely used in seed research [33–38]. Currently, it is routinely employed in comprehensive profiling of complex protein extracts, and delivers valuable qualitative and quantitative information on protein dynamics in plant organisms. Implementation of cell fractionation techniques allows reasonable simplification of the protein sample matrix and provides better insight into the molecular organization of individual compartments, as well as macromolecular complexes like cytoskeleton, membranes and the cell wall [39–42]. More specifically, the global proteomic approach addresses alterations in multiple biological processes, occurring sequentially and in parallel at the tissue, cellular or sub-cellular level [35, 43]. On the basis of the acquired information, development of seeds and seedlings can be systematically characterized [44, 45]. Because of this, the molecular basis of seed vigor and its alterations in response to developmental or environmental stimuli was intensively studied during the last decade [26, 33, 46–52], although the key proteins associated with seed vigor are still poorly understood.

Therefore, in this review we address the proteomics techniques, currently established in plant science and potentially applicable to assessment of seed quality, both in terms of viability and potential food safety aspects. Here we discuss methodology aspects, with a special emphasis on sample preparation, data acquisition, processing and post-processing.

2. Seed proteomics: general methodology and applications

The proteome is a complex system, representing a result of interconnected dynamic properties of individual proteins [53]. In agreement with this, proteomics aims at qualitative and/or quantitative characterization of the proteome, with the depth defined by the type and complexity of the specific research aim. Accordingly, so-called discovery proteomics is focused on maximizing sampling depth, annotation of individual proteins, expression patterns of their isoforms, as well as identity and precise localization of post-translational modifications [37, 54]. Discovery proteomics is inherently hypothesis generating, implementing relative or absolute quantification of individual proteins (protein dynamics), in response to internal or external factors [55]. Although, to some extent, these tasks can be solved by other analytical techniques (e.g. immunochemical methods, like ELISA or Western blotting) [56], the most efficient way it can be achieved is by mass spectrometry coupled to high-resolution separation techniques [57]. Thereby, protein identity can be assigned by tandem mass spectrometry (MS/MS), i.e. fragmentation of quasi-molecular ions, pre-selected in so-called full-MS scans (data-dependent acquisition) or acquired independently (data-independent acquisition) [58, 59]. Further post-processing with versatile bioinformatics tools gives access to functions and intracellular localization of individual proteins, as well as their involvement in complex protein-protein interaction networks [60].

In general, proteomic analysis can rely either on top-down or bottom-up strategies [61]. In the first case, quasi-molecular ions of individual proteins with molecular weight below 25 kDa are measured directly, including fragmentation and MS/MS acquisition of protein fragment ions [62, 63], whereas the bottom-up strategy employs proteolytic digestion of polypeptides prior to MS and MS/MS analysis of the resulting peptide-specific ions [64]. Thereby, identification of proteins by bottom-up proteomics relies on sequence tags, comprising exact monoisotopic mass, charge and MS/MS fragmentation patterns acquired for peptide ions [65]. As application of top-down proteomics is restricted by molecular weight and purity requirements for analyte proteins as well as simply the protein size and charge abundance [61], it's use in seed proteomics is limited [66]. Thus, in the absolute majority of cases, seed proteomics relies on the bottom-up strategy[31].

Obviously, the high complexity of the eukaryotic proteome represents the greatest challenge in proteomics [67]. Therefore, powerful mass spectrometric techniques, used for proteome analysis, need to be complemented by high-resolution and high-throughput separation methods. In general, all separation methods employed in proteomics can be attributed to either (i) gel-based or (ii) gel-free strategies [68], clearly different in their methodological setup: gel-based methods assume separation on the protein level, whereas the gel-free approach relies on limited enzymatic hydrolysis prior to separation (Figure 1). Accordingly, the first group of methods relies on polyacrylamide gel electrophoresis in sodium dodecyl sulfate (SDS-PAGE) or two-dimensional gel electrophoresis (2D-GE), whereas the second one employs liquid chromatography (LC), or electrophoretic techniques, like free-flow or capillary electrophoresis[69].

Figure 1. The overview of the experimental workflows for gel-based and gel-free proteomics.

3. Gel-based bottom-up proteomics

Due to relatively low resolution of SDS-PAGE, gel-based proteomics typically relies on 2D-GE, i.e. a two-step procedure, sequentially employing isoelectrofocusing (IEF) for separation of proteins by isoelectric point, followed by SDS-PAGE for separation by molecular weight (Figure 1). To establish pH gradients, required for IEF, protein samples are solubilized in aqueous buffers supplemented with carrier-ampholytes, chaotropic agents and detergents[70]. Most often, 7 – 8 mol/L urea and 2 mol/L thiourea are used as chaotropic agents (i.e. compounds, disrupting the hydrogen bonding network of water) [71], whereas 3-(3-cholamidopropyl)dimethylammonio)- 1-propanesulfonate (CHAPS) [72] and Triton X-100 are typically used as detergents (i.e. substances, disrupting intra- and inter-molecule non-polar interactions) [73].

Protein isolation procedures are usually designed in agreement with this reconstitution scheme (Table 1). To achieve a high efficiency of protein extraction, plant tissues are usually shock-frozen in liquid nitrogen and homogenized with a mortar and pistil or/and in a ball mill [74]. Among the variety of available techniques, phenol extraction and trichloroacetic acid (TCA)/acetone precipitation are the most effective methods to achieve comprehensive protein isolation[75]. Both these techniques ensure quantitative solubilization of membrane proteins, especially, when detergents (SDS or Triton X-100) are added [76, 77]. Hence, application of these methods can be considered as a pre-requisite for complete coverage of seed proteome [78]. Technically TCA/acetone precipitation is easier – it relies on direct precipitation of proteins from plant tissues [79]. In contrast, phenol extraction is a two-step procedure, i.e. (i) solubilization of proteins with phenol and (ii) precipitation with methanol afterwards [80]. In comparison to other extraction methods, the TCA/acetone precipitation proved to

be highly-efficient in seed proteome analysis [81, 82]. On the other hand, although the phenol-based methods suffer from compromised recovery, this approach provides better purity, in comparison to acetone/TCA extraction. Indeed, it ensures reliable removal of anionic polysaccharides and nucleic acids, which might interfere with SDS-PAGE [83]. Importantly, when working with whole seeds, a special attention needs to be paid on removal of secondary metabolites, especially phenolics, which are typically abundant in seed coats and might react with proteins, affecting electrophoretic separation [84]. Oxidative damage of proteins by phenolics can be suppressed by addition of soluble or non-soluble polyvinylpyrrolidone (PVP) to the extraction buffer [85, 86] or (even better) by supplementation of PVP directly to plant samples prior tissue grinding [87]. Reducing agents, such as mercaptoethanol, DTT and ascorbic acid are added as well [88], although the latter can promote oxidation and glycation of amino acid side chains in presence of transition metals [89–91]. It is important to note, that despite its excellent performance, phenol extraction suffers from low sample throughput and requires some experience in handling.

Most other protein extraction methods provide lower protein yields, incomplete protein recovery and/or compromised purity. Easiest, ground plant tissues can be extracted directly with sample buffer (sodium or potassium phosphate with pH close to neutral) containing urea and usually (but not always) thiourea [81, 92]. To achieve higher recovery of membrane proteins, sample buffer is typically supplemented with CHAPS [73], Triton X-100 [25], Nonidet P40 (NP-40) [93], or sodium dodecyl sulfate (SDS) in presence of 20% (v/v) glycerol [94]. To remove lipids (which can reduce solubility of proteins and affect thereby IEF separation [95]), pre-cleaning with petroleum ether [96] or dichloromethane (DCM) [97] can be implemented in the protocol. Resulted pre-cleaned extracts can be analyzed directly [98] or further purified, e.g. by precipitation with TCA/acetone [99]. Alternatively, soluble protein fractions can be extracted by aqueous buffers (typically Tris-HCl or HEPES), further precipitated by acetone and dried [25].

In an electropherogram, separated proteins can be visualized as patterns of characteristic signals - so-called “spots”, visible with the eye or in fluorescence detection mode, respectively [61]. Obviously, due to the high number of detected signals, highly-sensitive methods ultimately require higher resolution. In the best case scenario, 2D-GE separation allows distinction of up to 10,000 spots per experiment [100, 101]. Remarkably, as 2D-GE analysis relies on sequential separation by isoelectric point and molecular weight, these two parameters can be easily derived from 2D-electropherograms, and provide, thereby, a further level of protein annotation, in addition to MS- and MS/MS analysis of individual spots [34]. Therefore, appropriate visualization of separated spots becomes critical for successful interpretation of 2D-GE data. Usually, when only major proteins are targeted, corresponding spots are visualized with colloidal Coomassie Blue dye [102], which is completely compatible with MS and provides sensitivity of several dozen nanogramms per electrophoretic zone [103]. When higher sensitivity is desired, various silver staining methods can be employed [104]. This technique is highly-sensitive and provides reliable protein detection in the femtomol range [105], and formaldehyde-free protocols allow reliable MS-based identification of separated proteins [106]. One must keep in mind, however, that this method suffers from low linearity [70]. Therefore, fluorescent dyes, like SYPRO™ Ruby [107] and Ruthenium red [108] stains, are advantageous, when quantification is desired. Indeed, such reagents provide sensitivity, close to silver staining, and a linear dynamic range of up to three orders of magnitude [109].

It is important to note, that the numbers of detected spots do not directly correspond to protein identification rates. Indeed, polypeptides with similar pI and Mr tend to co-migrate in both separation dimensions. Also proteins, represented by several isoforms usually appear as characteristic signal patterns, slightly different by one or both parameters. Although 2D-GE delivers valuable information about qualitative and quantitative isoform patterns [43], the overall numbers of non-redundant proteins, identified in seed isolates by 2D-GE, do not exceed several hundred [37].

After excision of individual spots, corresponding proteins can be subjected to limited proteolysis [31]. The corresponding protocols typically rely on a well-established procedure, only slightly varying depending on specific experimental setup (Figure 2). Thus, when Coomassie dye is used for gel staining, the workflow includes excision of spots, cutting them in small sections, de-staining by

repeated incubation in aqueous acetonitrile-containing buffers, and enzymatic digestion (Table 2) [110]. After de-staining, gels are dried at room temperature and rehydrated with a protease solution [111]. The gel pieces, swollen in protease solution efficiently absorb enzyme. It facilitates proteolysis and increases its efficiency. After completion of digestion, proteolytic peptides can be recovered by repeated dehydration and re-hydration of gel sections. For gels stained with silver nitrate, no de-staining procedure prior to mass spectrometry is necessary [112]. However, only formaldehyde-free techniques can be applied, when proteolysis and MS-based characterization of hydrolysates are desired [105]. Importantly, although reduction of disulfide bonds and alkylation of resulted sulfhydryls can be done during sample preparation (Table 2) [113], more often this step is implemented after excision of gel slices [41, 114, 115]. The main limitation of in-gel digestion workflow is its somewhat limited throughput. However, recently, this was successfully overcome by introduction of the high-throughput in-gel digestion (HiT-Gel) technique, based on a 96-well plate format (Table 2)[116].

Analysis of resulted digests most often relies on matrix-assisted laser desorption-ionization – time of flight (MALDI-TOF)- or TOF/TOF-MS. In the simplest case, MALDI-TOF-MS analysis employs so-called peptide mass fingerprinting, i.e. protein identification by the characteristic MS pattern of proteolytic peptides [117, 118]. This approach suffers, however from poor sensitivity and selectivity [61]. Indeed, only highly-abundant proteins generate patterns of proteolytic peptides sufficient for their unambiguous identification. At the same time, these abundant components suppress the signals of co-migrating minor proteins, making their annotation challenging. To some extent, this limitation can be overcome by tandem mass spectrometry (MS/MS): MALDI-TOF/TOF instrumentation can be applied to obtain more reliable results [115, 119]. In this case, MS/MS-based sequencing of at least three protein-specific (unique) peptides is recognized as a gold standard for reliable annotation of individual electrophoretic zones [120]. However, low resolution of the precursor selection algorithm in MALDI-TOF/TOF-MS instruments (which is based on separation in first short field-free region) essentially restricts the discovery potential of this technique. Implementation of liquid chromatography (LC) separation prior to MALDI-TOF/TOF-MS analysis allows overcoming this limitation and provides improved identification rates for proteins, co-migrating in one electrophoretic zone (spot) [121].

Figure 2. Overview of enzymatic digestion implemented in gel-based and gel-free experimental setup.

One of the most important features of 2D-GE is the possibility for relative quantification of individual electrophoretic zones [122]. This information provides direct access to quantitative profiles of protein abundance (also often referred to as protein expression) and allows clear and straightforward statistical interpretation of data [123]. In this context, protein dynamics, i.e. alterations in protein expression profiles in time or in response to application of certain experimental conditions, can be characterized [34, 115]. Thereby, UV or VIS optical density of stained proteins, or intensity of fluorescence signal, can be considered as quantitative parameters. The appropriate integration algorithms are usually implemented in automatized protocols and established software tools, providing convenient access to image acquisition, spot detection, matching, normalization, statistical analysis, and annotation, also often combined with robotized spot cutting and in-gel digestion. The most commonly applied tools rely on numerical approaches (PDQuest (Bio-Rad Laboratories), Delta2D (Decodon), and Melanie (GeneBio)) and can be successfully complemented by home-made systems [124].

Unfortunately, although 2D-GE instrumentation is relatively cheap, easy to handle and adequately supported by efficient data analysis software, this method has several intrinsic disadvantages. The most critical limitation is the relatively low inter-gel reproducibility, i.e. difficulties in spot detection and matching in parallel gels [125]. Indeed, each IPG strip and electrophoresis block represents an independent system with separation conditions, different from those of others. Although batch setups for casting and running gels are currently well-established, this inter-system error cannot be eliminated completely because of a local inhomogeneity of casting solutions. This error might also affect dispersion within treatment groups and significance of inter-group comparisons [126]. Fortunately, to some extent, this error can be corrected by sample multiplexing in terms of the difference gel electrophoresis (DIGE) approach [127]. For this, cyanine dyes (Cy3 and Cy5) can be added to alternating samples, whereas the third dye (Cy2) can be used for

standardization [128]. Due to the presence of different fluorophores, each dye can be detected independently without essential cross-talk with limits of detection (LODs) in the higher pg range[31]. Although the 2D-GE technique is being continuously improved, some limitations of this method are still not overcome. First, as most of the spots are composed of several proteins, the reliability of quantitative profiles, obtained by 2D-GE is often questionable. Because of this, additional quantitative Western blotting experiments for proteins of interest are desired[41]. Further, even when multiplexing with DIGE is performed, the method might still suffer from inter-replicate variability. And finally, in comparison to gel-free techniques, 2D-GE is characterized by poor resolution and a relatively low linear quantitative dynamic range. Therefore, gel-free techniques are being actively introduced into seed proteomics during the current decade.

4. Gel-free bottom-up proteomics

The gel-free approach assumes application of enzymatic proteolysis prior to separation, i.e. the entire protein extract is digested, and the resulting mixture of proteolytic peptides separated (Figure 1). The analysis typically employs nanoRP-HPLC with electrospray ionization (ESI)-MS coupled on-line [61], or with matrix-assisted laser desorption/ionization (MALDI)-MS coupled off-line [129]. Among these two techniques, LC-ESI-MS is more powerful, and may result in identification of up to several hundred thousand peptides [37]. Thus, the LC-MS approach provides impressive analytical resolution and represents a powerful tool, delivering protein identification rates approximately ten-fold higher, than conventional 2D-GE methods [130].

Unfortunately, despite of its high analytical power, LC-MS-based proteomics has some limitations, which need to be kept in mind when planning experiments. Thus, this technique is highly amenable to matrix effects, which are typically manifested as suppression of low-intensity peptide signals by highly abundant co-eluting species [131, 132]. This ultimately results in underestimation of low-abundant proteins and might compromise both dynamic range of proteome analysis and quality of label-free quantification [133]. Therefore, the separation efficiency of the LC-MS system, prefacing ESI-MS is of critical importance for successful identification of low-abundant peptide quasi-molecular ions and quantification of corresponding proteins.

Another issue is incompatibility of the ESI mechanism with detergents (e.g. SDS and CHAPS) [134]. Indeed, as desolvation of low molecular weight analyte ions relies on the solvent evaporation model [135], and detergents suppress evaporation of eluents in droplets, these compounds disrupt transfer of analytes into the gas phase and dramatically affect sensitivity of detection. Besides this, conventional detergents (like SDS, Triton X100 and CHAPS) are quantitatively retained in reversed phase chromatography (RPC) and broadly co-elute with proteolytic peptides, further suppressing their signals [136]. Although SDS can be removed by potassium dodecyl sulfate (KDS) precipitation method, this procedure requires time-consuming optimization (Table 2) [137]. Thus, the mentioned detergents cannot be employed in sample preparation for LC-MS. Because of this, isolation of proteins with detergent-free aqueous buffers gives access only to the soluble part of the plant proteome [138]. In this case, pre-cleaning of protein extract might be necessary to remove low molecular weight metabolites, which may interfere with enzymatic proteolysis (e.g. by inhibition of proteases) [139]. Whereas for human plasma samples ultrafiltration is sufficient to remove interfering metabolites and salts [140], for plant extracts an additional pre-cleaning by size-exclusion chromatography on PD-10 columns is required [139]. Alternatively, SDS-PAGE can be introduced as a sample preparation step. For this, the isolated protein mixture is reconstituted in SDS-sample buffer and separated by SDS-PAGE (Table 2). Afterwards, the whole lanes are excised and cut in several (at least, ten) segments for in-gel digestion with protease [25]. After pooling of individual digests, samples can be desalted and analyzed by LC-MS. However, both of the described strategies have two essential disadvantages. They require time-consuming sample preparation and might suffer from incomplete recovery of the seed proteome. Digestion recovery can be improved when buffers containing urea or guanidine chloride without addition of detergents are used [39, 92]. However, even these methods do not provide the highest possible proteome coverage.

Importantly, the numbers of identified proteins can be affected by intrinsic limitations of data-dependent acquisition (DDA) experiments, most widely used in LC-MS proteomics [141]. State of the art quadrupole-orbital trap (Q-Orbitrap) and quadrupole-time of flight (QqTOF) mass spectrometers are able to perform DDA experiments that comprise up to 20 MS/MS scans within 0.5-1 s, acquired for the most intense signals detected in a survey full-MS scan [142]. However, despite this high acquisition speed (scan rate up to 20 Hz), the number of peptides that can be sequenced in an MS/MS scan is still limited and usually essentially lower, than the number of peptides, detected at the MS level.

Another important issue is the relatively low inter-run reproducibility of spray performance. Because of this, intra- and inter-batch precision of label-free quantification is typically low. Therefore, analytical strategies need to employ inter- and intra-batch normalization [143]. In this context, techniques, relying on internal normalization and multiplexing, like metabolic [144], chemical [145] or/and stable isotope [146] labeling, are advantageous, when precise quantification is desired. Finally, probably, the most important factors limiting application of LC-MS-based methods are the high costs of instrumentation and limited availability of well-trained personnel. Due to this, 2D-GE is still the most popular analytical technique, employed in seed proteomics.

Fortunately, current developments in analytical science provide adequate solutions to all limitations of the gel-free approach, and allow efficient shotgun LC-MS analysis of the total seed proteome. Thus, highly-efficient separation can be achieved by ultra-high performance liquid chromatography (UHPLC) [61]. Elution of individual analytes in well-defined chromatographic zones minimizes matrix effects and attenuates suppression of low-intense signals by highly-abundant species [147].

Quantitative reconstitution of membrane proteins without any deleterious effects on sensitivity of peptide detection is another essential achievement. Thus, at the end of the last decade, Mann and co-workers proposed the so-called filter-aided sample preparation (FASP) method, assuming digestion of proteins in the presence of SDS on membrane filters, and exchanging it by urea afterwards (Table 2) [148]. This technique proved to be efficient for grape leaf [149] and cucumber seeds [150], and delivered superior protein identification rates in comparison to sample pre-cleaning by SDS-PAGE. An alternative approach relies on immobilization of protein in gel by mixing with acrilamide/bisacrilamide mixture after its complete solubilization in appropriate detergent solution (so-called Tube-gel digestion protocol) [151]. Immobilization in gel results in unfolding of hydrophobic (e.g. membrane) proteins, and promotes their effective digestion. A similar protocol, known as gel-aided sample preparation (GASP), employing immobilization of protein in gel, is also applicable to intact cells [152]. Both FASP and GASP yield up to five-fold higher intensities of peptide signals, in comparison to in-solution digestion in absence of detergents. Interestingly, application of aqueous solutions, containing high proportions of acetonitrile, results in superior digestion yields, when small protein amounts need to be hydrolyzed [153]. Surprisingly, the numbers of identified peptides were much less, when methanol was used [39].

Further improvement of proteome coverage could be achieved by application of commercially available detergents, forming insoluble precipitates under low pH values [154]. One such compound, sodium deoxycholate (SDC), is well-known from plasma proteomics [140]. Recently, SDC was successfully applied to solubilization and digestion of total barley leaf protein preparation [155] and, isolates from germinating maize seeds [156]. Although this reagent can be precipitated by acids upon digestion, this process is not always quantitative. The same is the case for sodium laurate [157].

An alternative solubilization/digestion strategy might rely on degradable acid-labile detergents. For example, zwitterionic PPS Silent Surfactant [158] and anionic Protease-MAX [159] surfactants propose a good alternative for SDS in digestion buffers. To date these detergents are only minimally involved in plant proteomics. Thus, Protease-MAX was recently successfully used for the characterization of *Solanum tuberosum* leaf tissue [160]. Unfortunately, both compounds yield hydrophobic cleavage products, which form films on the surface of aqueous phases thereby contaminating samples [161]. This ultimately requires additional intensive sample cleaning after proteolysis. The same is known for the anionic acid-labile detergent RapiGest SF Surfactant [162],

which is currently most often applied for digestion of human plasma and tissues [163]. Recently, this reagent was successfully employed in the digestion of barley seed protein isolates, resulting in annotation of 226 polypeptides in shotgun data-independent experiments [164].

Another option available from the company Progenta, is anionic, cationic and zwitterionic acid labile surfactants I and II (AALS I/II, CALS I/II, and ZALS I/II), among which AALS II, CALS II and ZALS II have higher protein solubilization potential, i.e. can be applied to hydrophobic (e.g. membrane) proteins [165]. Recently, in a systematic comparison of different degradable surfactants, AALS II (structurally mimicking SDS) showed the best performance in terms of solubilization efficiency and sequence coverage of digested and analyzed standard protein samples [166]. Importantly, cleavage of this detergent yields relatively hydrophilic products, which can be easily removed by solid phase extraction (SPE) on reversed phase [38]. This compound has tunable surfactant properties, i.e. adjustable critical micelle concentration (CMC) [161]. Therefore, recently we selected this surfactant for profiling of total Arabidopsis leaf proteome [10, 66]. Later, we successfully addressed proteome changes in oilseed rape (*Brassica napus*) seeds during germination [38].

Remarkably, when we combined application of AALS II with long chromatographic gradients of 1.5–3 h (which allow a higher numbers of cycles of DDA experiments per acquisition run), confident identification of thousands proteins could be achieved [74]. The effect of the under sampling phenomenon, known to accompany DDA experiments [141], can be further attenuated by sample enrichment, depletion or pre-fractionation. For instance, glycosylated plant proteins, can be selectively enriched by boronic acid chromatography (BAC) either at the protein or peptide level [139, 167], whereas precipitation of legume reserve seed proteins by protamine sulfate or by organic solvents might represent an example of a depletion procedure, giving a deeper insight in to the seed proteome [168]. In the most efficient way, the overall analytical resolution of the whole proteomic workflow can be increased by comprehensive two-dimensional chromatography [169]. Thus, in human proteomics, hydrophilic interaction liquid chromatography (HILIC) can be used as the first dimension, whereas collected fractions (typically 5 – 10) can be easily dried, reconstituted in appropriate low-organic solvent and separated by RP-(U)HPLC [170]. Recently, we have successfully established this approach for the plant proteome [74]. It is important to mention, that pre-fractionation techniques can also rely on multipole pre-filters of mass spectrometric instrumentation – the method, known as gas phase fractionation (GPF) [171]. In terms of this approach, the whole mass range of survey full-MS scans in DDA experiments is split into smaller m/z intervals, which are analyzed in independent analytical runs. This dramatically reduces the complexity of peptide mixtures analyzed in individual fractions, and ensures, thereby, the complete detection of low-abundant proteins. After combining the results obtained in all experiments, substantial increase in proteome coverage can be achieved.

To achieve the highest possible protein identification rates and precision of label-free quantification, we optimized our proteomics workflow for the highest protein recovery and reproducibility (Figure 3). Thus, after shock-freezing in liquid nitrogen, seeds or their parts (e.g. embryos or cotyledons) are ground in a ball mill and 50 – 200 mg (depending on species) of frozen powder are taken for phenol extraction [38, 75]. After drying, protein isolates are quantitatively reconstituted in shotgun buffer, containing chaotropic agents (urea and thiourea) and degradable detergent (Progenta Anionic Acid-Labile Surfactant, AALSII) [10]. If protein isolation is performed carefully, and no interphase (typically containing nucleic acids and polysaccharides) is co-precipitated with the target protein fraction, seed protein isolates can be reconstituted completely. If it is not the case, after a short centrifugation, protein contents are determined in supernatants. For this, we are using a two-step procedure, relying on protein determination by the 2D-Quant kit (which is based on specific binding of copper to protein [172]) with a subsequent cross-validation of results by SDS-PAGE with Coomassie staining [173]. Application of these two orthogonal approaches ensures reliable normalization of protein amounts taken for enzymatic digestion. According to our experience, digestion of solubilized proteins with trypsin was sufficient for complete proteolysis of seed samples [61]. The completeness of hydrolysis was confirmed by SDS-PAGE, as was earlier

shown for human samples [91]. Having this in hand, AALS can be destroyed and the digests can be pre-cleaned with RP-SPE either in a cartridge or in stage-tip format [63, 123].

Figure 3. Detailed workflow for LC-based proteomics: protein isolation, sample preparation, and analysis.

The pre-cleaned samples can be analyzed by nano(U)HPLC-ESI-MS employing either data-independent acquisition (DIA) or data-dependent acquisition (DDA) modes (Figure 3). In QqTOF-MS instruments, DIA relies on simultaneous fragmentation of multiple peptide ions within a wide range of m/z . In its simplest form, this m/z range covers the whole measurement mass range (MSE technology, employed by Waters instruments) [174]. Alternatively, the full mass range is split into 10–30 segments of 20–50 m/z each, wherein all ions are concomitantly fragmented- sequential window acquisition of all theoretical fragment ion spectra (SWATH) MS technology [175]. The strategy for analysis of SWATH data is based on peptide-centric scoring, which relies on querying chromatographic and mass spectrometric coordinates of the proteins and peptides of interest in form of so-called peptide query parameters (PQPs) [176]. In contrast, the DDA scan strategy relies on selective fragmentation of a defined number of the highest intensity signals, detected in survey full-MS scans [141]. To increase analytical reproducibility, this approach can be additionally mapped to extracted ion chromatogram (XIC) information [177]. Identification of individual peptide sequences relies on tandem mass spectrometry (MS/MS, Figure 3), whereas quantitative information can be derived from spectrum counting (PSM counting), peak heights and, in the most precise way, from the integration of extracted ion chromatograms [178].

Both DIA and DDA have advantages and disadvantages. Thus, DIA delivers rich fragment information, although correct annotation of fragments is critical for this technique [179]. On the other hand, the major challenge of DDA experiments is analytical under sampling [141]. Indeed, just a limited number of quasi-molecular ions can be fragmented in each DDA method cycle, whereas the major part of the proteome remains unidentified. In order to minimize this limitation, LC-MS analysis can be pre-faced with an additional chromatographic dimension or implementation of prolonged separation gradients is required [74, 170]. Thereby, selection of mass spectrometric instrumentation is generally governed by the overall goal of the specific research. In this context, high resolution and

high mass accuracy are critical for successful annotation of sequence tags in discovery proteomics. This is provided at the MS level by the linear ion trap (LTQ)-Orbitrap-MS technology, which assumes acquisition of MS and MS/MS spectra with high (up to 120000) and low (i.e. unit) resolution, respectively [71]. Q Exactive (quadrupole-orbital trap, Q-Orbitrap) and QqTOF mass analyzers rely on high resolution and high mass accuracy both in MS and MS/MS mode [180]. Therefore, for DDA experiments, LTQ-Orbitrap and especially Q-Orbitrap instruments represent the best option [141, 170, 181]. On the other hand, for targeted proteomics, triple quadrupole (QqQ) mass spectrometers, operating in selected reaction monitoring (SRM) mode, provide the highest sensitivity in the lower atomole range [133]. As QqTOF-MS can be considered as an extension of QqQ-MS, these instruments are also well suited for analysis in SRM mode [182]. A scan strategy called parallel reaction monitoring (PRM) extends the targeted SRM approach to the Q-Orbitrap mainly because of its two linear quadrupoles configured before the Orbitrap mass analyzer in analogy to the triple quadrupole configuration. It acquires MS/MS spectra of precursor ions isolated in the second quadrupole mass filter as opposed to monitoring individual transitions as in SRM/MRM, delivering higher specificity with comparable sensitivity [183]. Also, it circumvents many of the difficulties associated with SRM development such as optimization of collision energy and dwell times. The approach has also been demonstrated on the LTQ-Orbitrap with a limit of quantification of 100 amol in a complex plant leaf extract [142].

5. Data processing and post-processing

Large datasets generated in high-throughput proteomic experiments require powerful computational platforms, which enable reliable annotation of MS and MS/MS signal patterns to sequences of specific proteolytic peptides, give access to precise quantitative assignments and support further statistical analysis to retrieve biological information. Here we summarize some major recent achievements, facilitating efficient processing and post-processing of proteomics data.

In MS-based proteomics, identification of peptide sequences typically relies on so-called sequence tags (see above) [184]. In the simplest way, a sequence tag might contain only patterns of peptide signals (peptide mass fingerprinting, PMF) [185]. Extension of the approach to MS/MS data (tandem mass spectrometric fragment patterns) provides direct access to sequence information [186]. Moreover, due to the unique character of fragment patterns, tandem mass spectrometry allows analysis of complex protein mixtures [187]. However, interpretation of MS/MS spectra in terms of peptide structure remains a challenging task. Typically, it relies on comparison of MS signals and MS/MS fragment patterns with computed data, obtained by *in silico* digestion of defined proteins and subsequent theoretical fragmentation of matched signals, that is typically done by search machines [188]. Unfortunately, in the majority of cases, MS/MS spectra suffer from poor quality, i.e. from incomplete of fragment ion series and presence of signals not related to backbone fragmentation, i.e. other than b- y- or immonium ions [189]. Moreover, insufficient separation prior to MS might result in overlap of several fragmentation patterns and makes straightforward interpretation difficult and possibly ambiguous.

The first successful approach for high-throughput and reliable identification of peptides relied on a partial unique sequence of at least 2-4 residues, derived from characteristic mass increments in MS/MS spectra [190]. The next generation of search engines skipped this intermediate step of identification (tag sequence). Thereby, in one group of engines, having the so called 'SEQUEST' architecture (PepSearch, SEQUEST, Crux, etc.), matching of theoretical and measured MS/MS spectra relies on cross-correlation analysis [191]. The engines, representing another group (Mascot, OLAV, Andromeda, etc.), estimate the probability of matches between the experimentally measured and calculated fragment m/z values occurring by chance [192, 193].

In general, the error tolerance of the measured m/z and the size of the sequence search space are the factors defining resolution of computational searches. The values of parent and fragment ion mass tolerance are defined by mass analyzer type and contribute directly to scoring matches [194]. Since search engines perform multiple comparisons on large datasets with defined mass tolerance, this might lead to some amount of false positive identifications, occurring by chance. And thus, actual

percentage of such false identifications should be corrected by the false discovery rate (FDR) procedures. For this, tandem mass spectra are searched against a decoy database, which represents a false sequence dataset, typically representing a reversed sequence set [195, 196] or database, generated by permutation [197–199]. The size of the search sequence space is partially defined by database size. Therefore, the most reliable peptide/protein identification can be achieved when database size is small (e.g. top-down proteomics). Importantly, the size of the search space denotes not only protein sequences in a primary database, but also all possible sequence post-translational modifications (PTMs) [200]. PTMs can be typically recognized by characteristic mass increments, observed in tandem mass spectra [201]. This feature is exploited by implementation of PTMs in search algorithms [202–204]. Thereby, PTMs can be defined as static (i.e. ultimately present at specifically defined amino acid residues, e.g. carbamidomethyl, formed at all cysteine residues after treatment of proteins with iodoacetamide prior to digestion), or variable (optionally occurring at specific residues, e.g. methionine sulfoxide, often formed during sample preparation). Submission of variable modifications results in significantly increased effective search space and time, as the search engine tests all possible combinations of unmodified and modified residues [194, 200].

Another popular method of MS/MS-based peptide identification is *de novo* sequencing. It is designed to exert maximal useful information from experimental spectra translating to proper peptide sequences [205, 206]. Since the *de novo* sequencing does not require a sequence database, in broad perspective view this technique might be the method of choice to identify proteins which are not present in databases, for example mutation-modified proteins or proteins of poorly or/and unstudied organisms, whose genomes are not sequenced.

Originally, due to large differences in ionization efficiencies of individual peptides, quantitative perspectives of MS-based proteomics seemed limited. Indeed, a high variability in charge state, peptide length, amino acid composition and PTM patterns result in great differences in ion intensities of peptide signals in spectra, even if they represent the same protein. Thus, for accurate quantitation by intensities of quasi-molecular ions, comparisons between different samples can only be done for the same peptide mass-to-charge ratios (m/z) which were acquired under the same conditions in LC-MS experiments [207]. The strategies of quantitative comparison are usually categorized as label-based (most often - stable-isotope-labeling techniques) and label-free approaches [208]. Label-based techniques are based on the fact, that stable isotope-labeled peptides differ from their non-labeled counterparts only in mass, but demonstrate the same chromatographic behavior and ionization efficiency [209]. Label-free quantification, in turn, can rely on feature-based quantification, or spectral counting (SC). Thereby, individual signals of quasi-molecular ions are treated as features, which are selected during peak peaking and matched during feature map alignment. The SC quantitation methods count the number of MS/MS spectra that are assigned to the same protein, ignoring intensities of corresponding peptide ions [210]. Label-free methods are characterized by a high analytical resolution, and dynamic range. Label free methods are supported by numerous software tools, which make quantitative comparison of proteins across many samples and biological scenarios facile even for a non-expert (Table 3): they include SuperHirn [211], MaxQuant [212], Progenesis [213], and OpenMS [214], Proteome Discoverer [215] and the Trans-Proteomic Pipeline [216]. They implement various protein quantification indices based on both spectral counting and ion signal intensities such as emPAI [217], APEX [218], mSCI [219], TOPn [220] and MS stats [221]. For users working with standard workflows and not developing their own algorithms, Progenesis and MaxQuant, which are suitable for fast data analysis, might represent the best solution [222]. If more flexibility and development of task-specific algorithms are required, open-source packages such as Viper [223], mzMine2v [224] or OpenMS [214] represent the best solution. Large proteomics labs and core facilities might appreciate the modularity and automation provided by pipeline tools (Figure 4) [210].

Figure 4. Detailed workflow for LC-based proteomics: protein isolation, sample preparation, and analysis.

Information on protein identities and relative or absolute abundance is important, but not sufficient for comprehensive understanding of complex interactions within a biological system. To build a generalized image of multiple simultaneous processes, occurring in biological systems, bioinformatics tools are required. These tools allow assignment of an identified protein to a specific sub-cellular compartment, molecular function or biological process as well as establishing its protein-protein interactions (PPI) as a network. The information about protein sub-cellular localization might be found in a database of experimentally derived annotations, available at resource SWISS-PROT [225]. However, the experimental data collections of even well-studied organisms (e.g. Arabidopsis or yeast) are not nearly complete [226, 227]. To overcome this, *in silico* prediction approaches (e.g. Plant-mPLOC, BUSCA, LOCALIZER, etc.), relying on specific protein sequence features, were developed [228–233]. For example, these algorithms are able to predict localization of proteins to membranes by the characteristic positions of hydrophobic and hydrophilic residues in their sequence [234]. The information on protein functions can be retrieved from protein sequence databases, like UniProt [235], KEGG [236–238] and plant specific functional database tools such as MapMan [239]. The two first databases organize functional information in Gene Ontology (GO) tags with the following increasing hierarchy order: enzymatic reaction, metabolic pathway, functional process, biological process [240]. This functional hierarchy scheme, however, might differ from database to database. For example, in MapMan the top hierarchical level comprises 34 functional bins [239]. In general, biological function is indirectly associated with protein sequence, although similar sequences tend to share functionality [241].

Finally, analysis of PPIs provides essential information for understanding cellular processes. For example, this information can be obtained by structure-based simulation and statistical/machine learning prediction approaches [242–244]. Thereby, the first group of methods relies on docking and molecular dynamics methods. This kind of calculations requires high computational capacity and rich protein structural information, and cannot, therefore, be applied to the size of protein sets usually

obtained in proteomics studies. The second group of methods relies on protein molecular descriptors (sequence, domain occurrence, co-expression, etc.) usually retrieved from databases. Information on protein interactions (both experimental and predicted) can be obtained from multiple resources, like STRING (PPIs prediction), BioGRID (PPIs and protein-chemical interactions), Reactome/Plant Reactome (PPIs along with metabolic pathways) etc. (Table 3)

6. Future perspectives

Seed maturation is accompanied by accumulation of reserve molecules [13], to a great extent represented by carbohydrates and lipids [245]. Accumulation of fatty acids and carbohydrates at the early steps of seed maturation, and their mobilization during germination may affect the patterns on non-enzymatic modifications of seed proteins, which are a highly-abundant component of, for example, legume seeds. Indeed, in seeds, oxidative stress, accompanying both their maturation and ageing [22, 246–248], develops at the background of high contents of reducing sugars (generated by de-polymerization of reserve polysaccharides) and unsaturated fatty acids (mostly formed during mobilization of reserve lipids) [249–253]. These conditions might result in oxidative degradation of carbohydrates and lipids/fatty acids via metal-catalyzed monosaccharide autoxidation and lipid peroxidation, respectively, yielding in both cases reactive carbonyl compounds (RCCs) [254, 255]. These intermediates are readily involved in glyco- and lipoxidation of proteins, resulting in formation of advanced glycoxidation and lipoxidation end products (AGEs and ALEs, respectively), which are potentially pro-inflammatory in mammals and impact in ageing and metabolic diseases [256]. Hence, this phenomenon needs to be taken into account, when seeds are stored for long periods of time [257], or when the processes, accompanying storage, are simulated in models of accelerated seed ageing [258].

Although glyco- and lipoxidation in plants was addressed by biochemical methods [252, 259], a comprehensive analysis of non-enzymatic post-translational modifications (PTMs) in ageing seeds still needs to be done. In the most efficient way, such a profiling of glyco- and lipoxidative modifications might rely on LC-MS-based approach with a special focus on in-depth analysis of low-abundant species using enrichment/depletion, pre-fractionation, and result-based exclusion of non-modified peptides from fragmentation [74, 139]. Alternatively, modified peptides can be selectively detected by highly sensitive and specific precursor ion scanning and further identified by targeted MS/MS experiments [260, 261]. Such experiments might rely on MS/MS analyses with model oxidized, glycosylated, as well as glyco- and lipoxidized peptides [260–265], which, can be obtained by solid phase peptide synthesis (SPPS) with high yield and purity [266–268]. Thereby, formation and degradation of early and advanced glycation products can be dissected by glycation models, based on synthetic peptides. On the one hand, such an approach may give access to the pathways of AGE formation [90, 139, 269], on the other hand, it can deliver information about carbonyl compounds involved in glycation [89]. Based on this approach, highly sensitive and reliable LC-MS/MS-based quantification techniques can be developed [133, 146], although their transfer to seed research is still not accomplished. Although glycation and lipoxidation are well-known markers of ageing in mammals [270], this was only recently confirmed for plants. Indeed, AGEs were shown to impact ageing not only in plant leaves [271], but also in such specialized structures, as legume root nodules [173]. Thereby, multiple regulatory proteins and transcription factors were involved in age-dependent modification. It raises the question about a possible regulatory role of this type of modification in plants. Keeping in mind a pronounced pro-inflammatory effect of glyco- and lipoxidized proteins on organisms of heterotrophic mammalian consumers of seed-derived food, one can conclude, that this aspect is highly relevant for food chemistry as well [256]. Taking all these aspects into account, patterns of AGEs and ALEs in seed proteins of crop plants need to be comprehensively characterized.

7. Conclusions

The seed proteome represents a highly complex system, only a part of which is characterized so far. The patterns of non-enzymatic post-translational modifications in seed polypeptides also remain

largely unknown. Thus, the quality and completeness of proteomics information, available to date, is insufficient for complete understanding of the processes, underlying protein dynamics during seed maturation and germination. The most obvious reason for this is the relatively low analytical power of 2D-GE, routinely used in seed proteomics. To some extent this problem can be solved by implementation of additional enrichment/depletion or pre-fractionation steps in established workflows. However, gel-based techniques have limited analytical power, insufficient for characterization of the complexity of the seed proteome. Therefore, LC-based methods, based on efficient separation systems, are obviously advantageous, although rarely used so far. Thereby, high resolution mass spectrometry needs to be employed for efficient identification and reliable quantification of individual proteins and their post-translational modifications.

Author Contributions: GS wrote the Introduction part and contributed other chapters, EL and AS prepared figures and tables and contributed gel- and LC-based proteomics parts, GMS, TB and ATs wrote the data processing part, EP, VZh, IT and SM contributed wrighting and critical discussion of the manuscript, WH and AF supervised writing, contributed in all chapters, and prepared the final version of the manuscript.

Funding: This research was funded Russian Science Foundation (project № 16-16-00026).

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

Abbreviations

LEA	Late embryogenesis abundant
ABA	Absciscic acid
MS	Mass spectrometry
MS/MS	Tandem mass spectrometry
SDS-PAGE	Polyacrylamide gel electrophoresis in sodium dodecyl sulfate
2D-GE	Two-dimensional gel electrophoresis
LC	Liquid chromatography
IEF	Isoelectrofocusing
CHAPS	3-(3-cholamidopropyl)dimethylammonio)-1-propanesulfonate
TCA	Trichloroacetic acid
SDS	Sodium dodecyl sulfate
PVP	Polyvinylpyrrolidone
DTT	Dithiothreitol
DCM	Dichloromethane
HiT-Gel	High-throughput in-gel digestion technique
MALDI	Matrix-assisted laser desorption-ionization
TOF	Time of flight
IPG	Immobilized pH gradient
DIGE	Difference gel electrophoresis
LODs	Limits of detection
ESI	Electrospray ionization
RPC	Reversed phase chromatography
KDS	Potassium dodecyl sulfate
DDA	Data-dependent acquisition
UHPLC	Ultra-high performance liquid chromatography
FASP	Filter-aided sample preparation
GASP	Gel-aided sample preparation
SDC	Sodium deoxycholate
AALS I/II	Anionic acid labile surfactant I/II
CALS I/II	Cationic acid labile surfactant I/II
ZALSI/II	Zwitterionic acid labile surfactant I/II
SPE	Solid phase extraction
CMC	Critical micelle concentration

BAC	Boronic acid chromatography
HILIC	Hydrophilic interaction liquid chromatography
GPF	Gas phase fractionation
DIA	Data-independent acquisition
PQPs	Peptide query parameters
XIC	Extracted ion chromatogram
LIT	Linear ion trap
SRM	Selected reaction monitoring
PRM	Parallel reaction monitoring
MRM	Multiple reaction monitoring
PMF	Peptide mass fingerprinting
FDR	False discovery rate
PTMs	Post-translational modifications
PPI	Protein-protein interactions
RCCs	Reactive carbonyl compounds
AGEs	Advanced glycoxidation end products
ALEs	Advanced lipoxidation end products
SPPS	Solid phase peptide synthesis

643 Appendix A

644

Table 1. Overview of protein extraction techniques.

#	Extraction technique	Extraction buffer	Chaotropic agents	Detergents	Reducing and chelating additives	Further additives	Precipitation (Vprecipitant : Vextract)	Isolate cleaning	Reconstitution	Ref
1	Phenol extraction	0.5 mol/L Tris-HCl (pH 7.5)	none	none	2% (v/v) ME, 50 mmol/L EDTA	1-15% (w/v) PVPP, PIC	0.1 mol/L AmAc/MeOH (5:1)	MeOH (3x), acetone (3x)	SDS-PAGE SB, IEF buffer, SB for LC-MS	[75, 74]
2	TCA/acetone extraction	10% (w/v) TCA in acetone	none	none	2% (v/v) ME	1-15% (w/v) PVPP, PIC	precipitation at the extraction step	acetone (3x)	SDS-PAGE SB, IEF buffer	[75, 79, 82]
3	Extraction with urea/thiourea buffer	14 mmol/L Tris-HCl	7 mmol/L urea, 2 mmol/L thiourea	2% (v/v) Triton X-100, 58 mmol/L CHAPS	none	PIC, 18 mmol/L ampholytes	none	none	solubilization at the extraction step	[73]
4	Acetone precipitation	20 mmol/L Tris-HCl (pH 7.5)	none	1% (v/v) Triton X-100	10 mmol/L EGTA, 1 mmol/L DTT	1 mmol/L PMSF, 250 mmol/L sucrose	precipitation at the extraction step	acetone (3x)	SDS-PAGE SB	[25]
5	Extraction with SDS-Tris buffer	125 mmol/L Tris-HCl	none	4% (w/v) SDS	2% (v/v) ME	20% (v/v) glycerol	none	none	solubilization at the extraction step	[94]
6	Extraction HEPES buffer/ delipidation (DCM)	50 mmol/L HEPES buffer	none	none	1 mmol/L EDTA	1 mmol/L PMSF, 0.1 mmol/L nDHGA	acetone (1:5)	none	SDS-PAGE SB, IEF buffer	[97]
7	Extraction with urea/thiourea buffer	6 mmol/L Tris-HCl, 4.2 mmol/L Trizma R	7 mmol/L urea, 2 mmol/L thiourea	4% (w/v) CHAPS	3% (w/v) DTT	PIC, DNase I, RNase A	none	none	solubilization at the extraction step	[98]
8	MeOH/CHCl ₃ precipitation, delipidation (PE)	50 mmol/L Tris-HCl (pH 8.8)	none	1% (w/v) SDS	0.07 % (v/v) ME, 1.5 mmol/L KCl	PIC, delipidation (PE)	MeOH/CHCl ₃ /ddH ₂ O (4:1:3)	SPE	8 mol/L urea in 50 mmol/L ABC	[92]

9	TCA/acetone precipitation, delipidation (PE)	50 mmol/L Tris-HCl (pH 8.8)	none	1% (<i>w/v</i>) SDS	0.07 % (<i>v/v</i>) ME, 1.5 mmol/ L KCl	PIC , delipidation (PE)	acetone (1:4)	SPE	8 mol/L urea in 50 mmol/L ABC	[92]
10	Acetone precipitation, delipidation (PE)	50 mmol/L Tris-HCl (pH 8.8)	none	1% (<i>w/v</i>) SDS	0.07 % (<i>v/v</i>) ME, 1.5 mmol/ L KCl	PIC, delipidation (PE)	acetone/10% (<i>w/v</i>) TCA (1:4)	SPE	8 mol/L urea in 50 mmol/L ABC	[92]
11	Urea solubilization buffer	8 mol/L urea, 2% (<i>w/v</i>) ampholyte (pH 3–10)	8 mol/L urea	4% (<i>w/v</i>) CHAPS	none	none	none	2D cleanup kit (GE Healthcare)	solubilization at the extraction step	[81]
12	Thiourea/urea solubilization buffer delipidation (hexane)	5 mol/L urea, 2 mol/L thiourea, 0.8% (<i>w/v</i>) ampholytes (pH 3–10)	5 mol/L urea, 2 mol/L thiourea	4% (<i>w/v</i>) CHAPS	65 mmol/L DTT	delipidation (hexane)	none	none	solubilization at the extraction step	[81]
13	Phenol extraction	0.1 mol/L Tris-HCl (pH 8.8)	none	none	10 mmol/L EDTA, 0.4% (<i>v/v</i>) ME	none	AmAc/MeOH (5:1)	0.1mol/L AmAc/MeOH (2x) acetone (2x) MeOH (1x)	8 mol/L urea, 2 mol/L thiourea, 2% (<i>w/v</i>) CHAPS, 2% (<i>v/v</i>) Triton X-100, 50 mmol/L DTT, 0.5% (<i>w/v</i>) ampholytes (pH 3–10)	[81]
14	Modified TCA/acetone precipitation / Urea solubilization extraction	10% (<i>w/v</i>) TCA in acetone	none	none	0.07% (<i>v/v</i>) ME	none	precipitation at the extraction step	acetone (2-3x)	9 mol/L urea, 1% (<i>w/v</i>) CHAPS, 1% (<i>w/v</i>) ampholytes pH (3–10), 1% (<i>w/v</i>) DTT	[81]
15	Phenol extraction	0.5 mol/L Tris-HCl, (pH 7.5)	none	none	2% (<i>v/v</i>) ME, 50 mmol/L EDTA	10% (w/v) PVPP, 1 mmol/L PMSF	0.1 mol/L AmAc/MeOH (2x)	none	IEF buffer, SDS-PAGE SB	[82]
16	Tris/TCA extraction	100 mmol/L Tris, TCA in acetone (pH 8.5)	none	none	5 mmol/L DTT, 1 mmol/L EDTA, 0.07% (<i>v/v</i>) ME	1 mmol/L PMSF	Precipitation at the extraction step	0.07% (<i>v/v</i>) ME in acetone	IEF buffer, SDS-PAGE SB	[82]

17	Tris-base extraction	40 mmol/L Tris	5 mol/L urea, 2 mol/L thiourea	2% (<i>w/v</i>) CHAPS	2% (<i>v/v</i>) ME	5% (<i>w/v</i>) PVP	Precipitation at the extraction step	0.07% (<i>v/v</i>) ME in acetone	IEF buffer, SDS-PAGE SB	[82]
18	TCA/acetone extraction	10% (<i>v/v</i>) TCA in acetone	none	none	20 mmol/L DTT	none	Precipitation at the extraction step	Acetone or 20 mmol/L DTT in acetone, or 10% ddH ₂ O in acetone or 20 mmol/L DTT, 10% ddH ₂ O in acetone	IEF SB	[84]
19	Phenol extraction	50 mmol/L Tris-HCl (pH 7.5)	none	none	5 mmol/L EDTA, 5 mmol/L DTT	1% (<i>w/v</i>) PVPP, 1 mmol/L PMSF	Acetone : supernatant (5:1)	none	Urea buffer (50 mmol/L HEPES (pH 7.8), 8 mol/L urea) SB for LC-MS	[86]
20	TCA/acetone extraction	10% (<i>v/v</i>) TCA in acetone	none	none	10 mmol/L DTT	10% (<i>w/v</i>) PVPP, 55 mmol/L iodoacetamide, 0.5 mol/L TEAB	Precipitation at the extraction step	Acetone (3x)	TEAB buffer, IEF SB, SB for LC-MS	[87]
21	TCA/acetone and methanol washes and phenol extraction	Phenol (pH 8.0) : SDS (1:1)	none	none	none	none	0.1 mol/L AmAc/MeOH	10% (<i>v/v</i>) TCA/acetone 0,1 MAmAc 80% MeOH 80% acetone	SDS-PAGE SB, IEF buffer	[95]
22	Tris-HCl / TCA / acetone extraction	0.1 mol/L Tris-HCl (pH 6.8), 10% (<i>v/v</i>) TCA/acetone	none	1% (<i>w/v</i>) SDS	0.1 mol/L DTT	none	10% (<i>v/v</i>) TCA/acetone	10% (<i>v/v</i>) TCA/acetone aqueous 10% TCA (2x) dH ₂ O (1x) acetone (1x)	SDS-PAGE SB	[95]

645 ABC, ammonium bicarbonate buffer; AmAc, ammonium acetate; CHAPS, 3-[(3-cholamidopropyl)dimethylammonio]-1-propanesulfonate hydrate; DCM, dichloromethane; DTT, dithiothreitol;
 646 EDTA, ethylenediaminetetraacetic acid; EGTA, ethylene-bis(oxyethylenenitrilo)tetraacetic acid; HEPES, 4-(2-hydroxyethyl)piperazine-1-ethanesulfonic acid; IEF, isoelectric focusing; LC-MS, liquid
 647 chromatography-mass spectrometry; ME, β-mercaptoethanol; nDHGA, nordihydroguaiaretic acid; PE, petroleum ether; PIC, protein inhibitor cocktail; PMSF, phenylmethylsulfonyl fluoride;
 648 PVPP, polyvinylpyrrolidone; SB, sample buffer; SDS, sodium dodecyl sulfate; SDS-PAGE, sodium dodecyl sulfate polyacrylamide gel electrophoresis; SPE, solid phase extraction; TCA,
 649 trichloroacetic acid; TEAB, triethylammonium bicarbonate; Tris, tris(hydroxymethyl)aminomethane.

650

Table 2. Digestion strategies used in bottom-up proteomics.

#	Object/tissue	Methodology						Ref
		Protein Isolation	Detergent/ chaotropic agent	Reduction/ Alkylation	Protease	Chromatographic System	MS	
1	<i>Brassica napus</i> , seeds	detergent extraction, phenol extraction	none (in-gel digest)	DTT/IA	trypsin	none	MALDI-TOF/TOF-MS	[272]
2	<i>Lupinus luteus</i> L., seeds	delipidation with hexane, acetone precipitation	none (in-gel digest)	DTT/IA	trypsin	RP C18, L-column water-ACN grad., 0.1% (v/v) FA	ESI-IT-MS	[113]
3	<i>Cicer arietinum</i> L., plasma membrane aerial parts	chloroform/methanol (5:4) extraction	none (in-gel digest)	DTT/IA	trypsin	RP C18,	ESI-LTQ-Orbitrap-MS	[41]
			4% (w/v) SDS/none	DTT/IA	trypsin	water-ACN grad., 0.1% (v/v) FA		
4	<i>Glycine max</i> L., seeds	delipidation with hexane, extraction with SDS- PAGE SB	none (in-gel digest)	DTT/IA	trypsin	none	MALDI-TOF/TOF-MS	
				DTT/none		RP, BEH130C18, water-ACN grad. 0.1% (v/v) FA	ESI-QqTOF-MS	[114]
5	<i>Glycine max</i> L., seeds	2 steps extraction: 1)protamine sulfateprecipitation 2) TCA / acetone precipitation	none (in-gel digest)	DTT/IA	trypsin	none	MALDI-TOF/TOF-MS	
			4% (w/v) SDS/ 8 mol/L urea	DTT/IA	trypsin	RP, C18. water-ACN grad. 0.1% (v/v) FA	ESI-Q-Orbitrap-MS	[115, 148]
6	<i>Arabidopsis thaliana</i> , aerial parts	1) detergent extraction 2)aq. buffer extraction	none (in-gel digest)	DTT/IA	trypsin	RP, C18 PepMap water-ACN grad. 0.1% (v/v) FA	ESI-Q-Orbitrap-MS	[116]
7	Rat, liver*	membrane isolation, centrifugation in sucrose gradient	1) 1%(w/v) SDS/none 2) 1%(w/v) RGS/ none 3)none/8mol/L urea 4) 60% (v/v) methanol**	DTT/IA	trypsin	RP, C18PepMap column water-ACN grad. 0.1% (v/v) FA	ESI-IT-MS	[137]
8	<i>A. thaliana</i> , leaves	phenol extraction	0.5% (w/v) AALS/ 7 mol/L urea	TCEP/IA	trypsin	RP, water-ACN grad. 0.1% (v/v) FA	ESI-LIT-Orbitrap ESI- QqTOF-MS	[138]

9	<i>G. max</i> L., seeds	1) detergent extraction 2) acetone precipitation	none (in-gel digest)	none	trypsin	RP, C18 PepMap column water-ACN grad. 0.1% (v/v) FA	ESI-LIT-MS	[25]
10	<i>Chenopodium quinoa</i> , seeds	1) detergent extraction 2) methanol/chloroform or acetone precipitation	none / 8 mol/L urea	DTT/IA	trypsin	RP, Acclaim-C18 column water-ACN grad. 0.1% (v/v) FA	ESI-LIT-Orbitrap-MS	[92]
11	<i>Solanum esculentum</i> roots,	detergent extraction for isolation of cell microsomal fractions by centrifugation	1) methanol** 2) 0.2% PPS SilentSurfactant/ none 3) 0.2% RGS /none 4) none / 6 mol/L GdnCl	TCEP/IA	trypsin	RP, BEH C18 water-ACN grad. 0.1% (v/v) FA	ESI-QqTOF-MS	[39]
12	<i>Vitis riparia</i> , leaves	1) detergent extraction 2) methanol-chloroform extraction	none (in-gel digest) 50% TFE**	DTT/IA DTT/IA	trypsin Lys-C, trypsin	RP, Magic C18AQ resin water-ACN grad. 0.1% (v/v) FA	ESI-LIT-Orbitrap-MS	[149]
13	Cucumber, seeds	TCA/acetone (1:9w/v) precipitation	none/ 8 mol/L urea	DTT/IA	trypsin	RP, C18. water-ACN grad. 0.1% (v/v) FA	ESI-Q-Orbitrap-MS	[150]
14	BSA, ubiquitin, myoglobin, PC3 cells	cell lysis, isolation of cell membranes by centrifugation	none	DTT/IA	trypsin	RP, C18. water-ACN grad. 0.1% (v/v) FA	ESI-QqTOF-MS	[151]
15	<i>Rhodopseudomonas palustris</i> Mixture of protein standards	acid extraction and sucrose density fractionation none	0, 60 or 80% acetonitrile** in 50 mmol Tris-HCl, 10 mmol/L CaCl ₂ none/6 mol/LGdnHCl	DTT/none	trypsin	RP, Vydac C18 water-ACN grad. 0.1% (v/v) FA	ESI-IT-MS ESI-FT-ICR-MS	[153]
16	<i>Hordeum vulgare</i> , leaves	detergent extraction	1)none / 8 mol/L urea 2) 2% (w/v) SDS/ 8 mol/L urea 3) 1% SDC / none 4) 2% SDC /none	DTT/IA	trypsin	RP, Reposilpur 120 C18 water-ACN grad. 0.1% (v/v) FA	ESI-Q-Orbitrap-MS	[155]

17	<i>Zea mays</i> , seeds	1) detergent extraction 2) TCA/acetone (1:9w/v) extraction	none (in-geldigest)	DTT/IA	trypsin	RP, Eksigent C8-CL- 120 column water-ACN grad. 0.1% (v/v) FA	ESI-QqQ-MS	[273]
18	Myoglobin, bacteriorhodopsin, BSA	none	1) 0.1-1.0% RGS / none 2) 1.0% SDC / none 3) 0.1-1.0% SL / none	none	trypsin	none	MALDI-TOF/TOF-MS	[157]
	Rat, liver	isolation of cell membranes by centrifugation in the gradient of sucrose	1) 1.0% SDS / none 2) 1.0% SL / none 3) 1.0% RGS / none 4) 1.0% SDC / none	DTT/IA	trypsin	RP, C18 PepMap column water-ACN grad. 0.1% (v/v) FA	ESI-IT-MS	
19	<i>Solanum tuberosum</i> leaves,	2 steps: 1) detergent extraction 2)co- immunoprecipitation	0.1% ProteaseMAX™ surfactant / none	TCEP/MMTS	trypsin	RP, Reprosil C18-AQ, water-ACN grad. 0.1% (v/v) FA	ESI-Q-LIT-Orbitrap-MS	[274]
20	<i>H. vulgare</i> caryopses,	10% TCA, 0.07% (w/v) β- mercaptoethanol in acetone	0.1% RGS / none	DTT/IA	trypsin	RP, C18 water-ACN grad. 0.1% (v/v) FA	ESI-QqTOF-MS	[164]
21	<i>B. napus</i> , seedling	1) aq. buffer extraction 2) phenol extraction	0.02% (w/v) AALS / 8 mol/L urea, 2 mol/L thiourea	TCEP/IA	trypsin	RP, Acclaim PepMap 100 C18 column water-ACN grad. 0.1% (v/v) FA	ESI-Q-LIT--Orbitrap-MS	[38]

652 *not established in seeds or generally in plants, shown as a reference method, **non detergent or chaotropic reagent used for solubilization; AALS, anionic acid labile
653 surfactant;ACN, acetonitrile;BSA, bovine serum albumin;DTT, dithiothreitol;ESI, electrospray ionization;FA, formic acid;FT-ICR, Fourier transform-ion cyclotron
654 resonance;GdnHCl, guanidinium chloride;IA, iodoacetamide;IT, ion trap;LIT – linear ion trap;MALDI-TOF, matrix assisted laser desorption/ionization - time of flight;MMTS,
655 methyl methanethiosulfonate;MS, mass spectrometry;nanoLC, nano-scaled liquid chromatography;Q, quadrupole mass analyzer;QqTOF, quadrupole-time of flight;RGS,
656 RapiGest SF;RP, reversed phase;SDC, sodium deoxycholate;SDS, sodium dodecyl sulfate;SL, sodium laurate;TCA, trichloroacetic acid;TCEP, tris-(2-carboxyethyl)-
657 phosphine;TFE, trifluoroethanol

658

Table 3. Software tools for processing and post-processing for proteomics data.

#	Tool	Version	Supported platform	GUI/CMD	Open source	Input formats	Quantification technique	Ref
1	MaxQuant	v1.6.3.3	Windows, Linux	+/+	+	AB SCIEX (*.wiff), mzXML, Thermo (*.raw), Agilent & Bruker Daltonics (*.d), Uimf (*.uimf)	LFQ, label-based	[275]
2	OpenMS/TOPP	v2.0	Windows, Linux, Mac OS	+/+	+	mzML, mzXML, mzData	LFQ, labeled-based	[276]
3	Progenesis QI	v2.3	Windows			Agilent & Bruker Daltonics (*.d), AB SCIEX (*.wiff), mzML, mzXML, Thermo & Agilent (*.raw)	LFQ, labeled-based	[277]
4	Proteome discoverer	v2.2	Windows	+/-	-	mzXML, mzDATA, mzML, MSF	LFQ, labeled-based	[278]
5	Census	v2.3	Windows, Linux, Mac OS	+/-	-	mzXML	LFQ, labeled-based possibility to self-define mass tags	[279]
6	SILVER	V3.0	Windows	+/-	+	mzXML, *.raw	SILAC	[280]
7	msInspect	v3.1	Windows, Linux, Mac OS	+/+	+	mzXML	LFQ, labeled-based	[281, 210]
8	mzMine2	v2.6	Windows, Linux, Mac OS	+/-	+	mzML, mzXML, mzData, NetCDF, RAW (Thermo)	LFQ	[224]
9	MassChroQ	v2.2.12	Windows, Linux	-/+	+	mzXML, mzML	LFQ, labeled-based	[282]
10	Skyline	v4.1	Windows, Linux	+/+	+	.sky, .skyd, mzML, mzXML, major vendor formats	LFQ	[283]

11	DIA-Umpire	v2.0	Windows, Linux	-/+	+	mzXML	ICAT, ¹⁸ O	[284]
12	Viper	v3.49	Windows, Linux	+/-	+	PEK, .CSV (Decon2LS), .mzXML,.mzData	ICAT, ¹⁸ O	[223]
13	OpenSWATH	v2.2	Windows, Linux, Mac OS	-/+	+	mzML, mzXML, TraML	LFQ, labeled-based	[175]
14	TPP	v5.1.0	Windows, Linux, Mac OS	+/+	+	mzXML, .RAW (Thermo), wiff, baf (Brucker), pepXML	LFQ	[285, 216]
15	moFF	v2.0	Windows, Linux, Mac OS	+/+	+	Thermo (.raw), mzML	LFQ	[286]
16	Mascot Distiller	v2.7	Windows	+/+	-	mzML, mzXML, mzData, major vendors	LFQ, labeled-based	[287]
17	Corra	v3.1	Linux			mzML, pepXML	LFQ, labeled-based	[288, 210]
18	FlashLFQ	v0.1.61	Windows	-/+	+	MzML, raw	LFQ	[289]

660 ¹⁸O, ¹⁸O labeling by ¹⁸O-enriched water; CMD, command line; GUI, graphical user interface; ICAT, isotope-coded affinity tag; LFQ, label free quantification; SILAC, stable
661 isotope labeling by/with amino acids in cell culture
662

References

1. FAO. 2018. *Seeds Toolkit - Module 5: Seed Marketing*; Food and Agriculture Organization of the United Nations: Rome, 2018, p. 108, ISBN 9789251309537.
2. FAO 2013. *Statistical Book. Part 3. Feeding the World*; Food and Agriculture Organization of the United Nations, 2013, p. 199, ISBN 978-92-5-107396-4.
3. FAO, IFAD, UNICEF, WFP and WHO. 2018. *The State of Food Security and Nutrition in the World 2018. Building Climate Resilience for Food Security and Nutrition*; Food and Agriculture Organization of the United Nations: Rome, 2018 ISBN 9789251305713.
4. Bradford, K.J.; Dahal, P.; Van Asbrouck, J.; Kunusoth, K.; Bello, P.; Thompson, J.; Wu, F. The dry chain: reducing postharvest losses and improving food safety in humid climates. *Trends Food Sci. Technol.* **2018**, *71*, 84–93, doi: 10.1016/j.tifs.2017.11.002.
5. Miller, J.K.; Herman, E.M.; Jahn, M.; Bradford, K.J. Strategic research, education and policy goals for seed science and crop improvement. *Plant Sci.* **2010**, *179*, 645–652, doi: 10.1016/j.plantsci.2010.08.006.
6. Bewley, J.D.; Bradford, K.J.; Hillhorst, H.W.M.; Nonogaki, H. *Seeds: Physiology of Development, Germination and Dormancy*. **2013** doi: 10.1007/978-1-4614-4693-4.
7. Bradford, K.J. Development of seed quality: maturity and environmental effects. In *Seed Production and Quality*; 2004; pp. 52–57, .
8. Finch-Savage, W.E.; Bassel, G.W. Seed vigour and crop establishment: extending performance beyond adaptation. *J. Exp. Bot.* **2016**, *67*, 567–591, doi: 10.1093/jxb/erv490.
9. Leprince, O.; Pellizzaro, A.; Berriri, S.; Buitink, J. Late seed maturation: drying without dying. *J. Exp. Bot.* **2017**, *68*, 827–841, doi: 10.1093/jxb/erw363.
10. Frolov, A.; Bilova, T.; Paudel, G.; Berger, R.; Balcke, G.U.; Birkemeyer, C.; Wessjohann, L.A. Early responses of mature arabidopsis thaliana plants to reduced water potential in the agar-based polyethylene glycol infusion drought model. *J. Plant Physiol.* **2017**, *208*, 70–83, doi: 10.1016/j.jplph.2016.09.013.
11. Chatelain, E.; Hundertmark, M.; Leprince, O.; Gall, S. Le; Satour, P.; Deligny-Penninck, S.; Rogniaux, H.; Buitink, J. Temporal profiling of the heat-stable proteome during late maturation of medicago truncatula seeds identifies a restricted subset of late embryogenesis abundant proteins associated with longevity. *Plant, Cell Environ.* **2012**, *35*, 1440–1455, doi: 10.1111/j.1365-3040.2012.02501.x.
12. Verdier, J.; Lalanne, D.; Pelletier, S.; Torres-Jerez, I.; Righetti, K.; Bandyopadhyay, K.; Leprince, O.; Chatelain, E.; Vu, B.L.; Gouzy, J.; Gamas, P.; Udvardi, M.K.; Buitink, J. A regulatory network-based approach dissects late maturation processes related to the acquisition of desiccation tolerance and longevity of medicago truncatula seeds. *Plant Physiol.* **2013**, *163*, 757–774, doi: 10.1104/pp.113.222380.
13. Smolikova, G.; Kreslavski, V.; Shiroglazova, O.; Bilova, T.; Sharova, E.; Frolov, A.; Medvedev, S. Photochemical activity changes accompanying the embryogenesis of pea (pisum sativum) with yellow and green cotyledons. *Funct. Plant Biol.* **2018**, *45*, 228, doi: 10.1071/FP16379.
14. Smolikova, G.N.; Medvedev, S.S. Photosynthesis in the seeds of chloroembryophytes. *Russ. J. Plant Physiol.* **2016**, *63*, 1–12, doi: 10.1134/S1021443715060163.
15. Smolikova, G.N.; Laman, N.A.; Boriskevich, O. V. Role of chlorophylls and carotenoids in seed tolerance to abiotic stressors. *Russ. J. Plant Physiol.* **2011**, *58*, 965–973, doi: 10.1134/S1021443711060161.
16. Nambara, E.; Okamoto, M.; Tatematsu, K.; Yano, R.; Seo, M.; Kamiya, Y. Abscissic acid and the control of seed dormancy and germination. *Seed Sci. Res.* **2010**, *20*, 55, doi: 10.1017/S0960258510000012.
17. Smolikova, G.; Dolgikh, E.; Vikhnina, M.; Frolov, A.; Medvedev, S. Genetic and hormonal regulation of chlorophyll degradation during maturation of seeds with green embryos. *Int. J. Mol. Sci.* **2017**, *18*, 1993, doi: 10.3390/ijms18091993.
18. Zinsmeister, J.; Lalanne, D.; Terrasson, E.; Chatelain, E.; Vandecasteele, C.; Vu, B.L.; Dubois-Laurent, C.; Geoffriau, E.; Signor, C. Le; Dalmais, M.; Gutbrod, K.; Dörmann, P.; Gallardo, K.; Bendahmane, A.; Buitink, J.; Leprince, O. ABI5 is a regulator of seed maturation and longevity in legumes. *Plant Cell* **2016**, *28*, 2735–2754, doi: 10.1105/tpc.16.00470.
19. McDonald, M.B. Seed quality assessment. *Seed Science Research*, 1998, *8*, 265–275.
20. *Handbook of Vigour Test Methods*, 3rd Ed.; Hampton, J. C., TeKrony, D. M., Eds.; International Seed Testing Organization, 1995, p. 117, .
21. Corbineau, F. Markers of seed quality: from present to future. *Seed Sci. Res.* **2012**, *22*, S61–S68, doi: 10.1017/S0960258511000419.
22. Bailly, C.; Kraner, I. Analyses of ros and antioxidants in relation to seed longevity and germination. In *Seed Dormancy: Methods and Protocols*; Kermode, A. R., Ed.; Springer Science + Business Media, LLC, 2011; pp. 343–367, .

23. Kocsy, G. Die or survive? redox changes as seed viability markers. *Plant, Cell Environ.* **2015**, *38*, 1008–1010, doi: 10.1111/pce.12515.
24. Gallardo, K.; Thompson, R.; Burstin, J. Reserve accumulation in legume seeds. *Comptes Rendus - Biol.* **2008**, *331*, 755–762, doi: 10.1016/j.crvi.2008.07.017.
25. Han, C.; Yin, X.; He, D.; Yang, P. Analysis of proteome profile in germinating soybean seed, and its comparison with rice showing the styles of reserves mobilization in different crops. *PLoS One* **2013**, *8*, e56947, doi: 10.1371/journal.pone.0056947.
26. Wang, W.Q.; Liu, S.J.; Song, S.Q.; Möller, I.M. Proteomics of seed development, desiccation tolerance, germination and vigor. *Plant Physiol. Biochem.* **2015**, *86*, 1–15, doi: 10.1016/j.plaphy.2014.11.003.
27. Copeland, L.O.; McDonald, M.B. The chemistry of seeds. *Princ. Seed Sci. Technol.* **1999**, No. 3, 40–58, doi: 10.1007/978-1-4615-1783-2_3.
28. Miernyk, J.A.; Hajduch, M. Seed proteomics. *J. Proteomics* **2011**, *74*, 389–400, doi: 10.1007/978-1-62703-631-3_26.
29. Robić, G.; Farinas, C.S.; Rech, E.L.; Miranda, E.A. Transgenic soybean seed as protein expression system: aqueous extraction of recombinant β -glucuronidase. *Appl. Biochem. Biotechnol.* **2010**, *160*, 1157–1167, doi: 10.1007/s12010-009-8637-5.
30. Schmidt, M.A.; Herman, E.M. Proteome rebalancing in soybean seeds can be exploited to enhance foreign protein accumulation. *Plant Biotechnol. J.* **2008**, *6*, 832–842, doi: 10.1111/j.1467-7652.2008.00364.x.
31. Rath, D.; Gayen, D.; Gayali, S.; Chakraborty, S.; Chakraborty, N. Legume proteomics: progress, prospects, and challenges. *Proteomics* **2016**, *16*, 310–327, doi: 10.1002/pmic.201500257.
32. Shewry, P.R.; Casey, R. Seed proteins. *Seed Proteins* **1999**, 1–10, doi: 10.1007/978-94-011-4431-5_1.
33. Gallardo, K.; Job, C.; Groot, S.P.C.; Puype, M.; Demol, H.; Vandekerckhove, J.; Job, D. Proteomic analysis of arabidopsis seed germination and priming. *Plant Physiol.* **2001**, *126*, 835–848, doi: 10.1104/pp.126.2.835.
34. Gallardo, K.; Signor, C. Le; Vandekerckhove, J.; Thompson, R.D.; Burstin, J. Proteomics of medicago truncatula seed development establishes the time frame of diverse metabolic processes related to reserve accumulation. *PLANT Physiol.* **2003**, *133*, 664–682, doi: 10.1104/pp.103.025254.
35. Catusse, J.; Job, C.; Job, D. Transcriptome- and proteome-wide analyses of seed germination. *Comptes Rendus - Biol.* **2008**, *331*, 815–822, doi: 10.1016/j.crvi.2008.07.023.
36. Rajjou, L.; Lovigny, Y.; Groot, S.P.C.; Belghazi, M.; Job, C.; Job, D. Proteome-wide characterization of seed aging in arabidopsis: a comparison between artificial and natural aging protocols. *Plant Physiol.* **2008**, *148*, 620–641, doi: 10.1104/pp.108.123141.
37. Frolov, A.; Mamontova, T.; Ihling, C.; Lukasheva, E.; Bankin, M.; Chantseva, V.; Vikhnina, M.; Soboleva, A.; Shumilina, J.; Mavropolo-Stolyarenko, G.; et al. Mining seed proteome: from protein dynamics to modification profiles. *Biol. Commun.* **2018**, *63*, 43–58, doi: 10.21638/spbu03.2018.106.
38. Frolov, A.; Didio, A.; Ihling, C.; Chantseva, V.; Grishina, T.; Hoehenwarter, W.; Sinz, A.; Smolikova, G.; Bilova, T.; Medvedev, S. The effect of simulated microgravity on the brassica napus seedling proteome. *Funct. Plant Biol.* **2018**, *45*, 440, doi: 10.1071/FP16378.
39. Mbeunkui, F.; Goshe, M.B. Investigation of solubilization and digestion methods for microsomal membrane proteome analysis using data-independent lc-mse. *Proteomics* **2011**, *11*, 898–911, doi: 10.1002/pmic.200900698.
40. Balmer, Y.; Vensel, W.H.; DuPont, F.M.; Buchanan, B.B.; Hurkman, W.J. Proteome of amyloplasts isolated from developing wheat endosperm presents evidence of broad metabolic capability. *J. Exp. Bot.* **2006**, *57*, 1591–1602, doi: 10.1093/jxb/erj156.
41. Barua, P.; Subba, P.; Lande, N.V.; Mangalparthi, K.K.; Prasad, T.S.K.; Chakraborty, S.; Chakraborty, N. Gel-based and gel-free search for plasma membrane proteins in chickpea (*cicer arietinum* L.) augments the comprehensive data sets of membrane protein repertoire. *J. Proteomics* **2016**, *143*, 199–208, doi: 10.1016/j.jprot.2016.04.015.
42. Yadeta, K.A.; Elmore, J.M.; Coaker, G. Advancements in the analysis of the arabidopsis plasma membrane proteome. *Front Plant Sci* **2013**, *4*, 86, doi: 10.3389/fpls.2013.00086.
43. Bourgeois, M.; Jacquín, F.; Savoie, V.; Sommerer, N.; Labas, V.; Henry, C.; Burstin, J. Dissecting the proteome of pea mature seeds reveals the phenotypic plasticity of seed protein composition. *Proteomics* **2009**, *9*, 254–271, doi: 10.1002/pmic.200700903.
44. Wang, W.Q.; Möller, I.M.; Song, S.Q. Proteomic analysis of embryonic axis of pisum sativum seeds during germination and identification of proteins associated with loss of desiccation tolerance. *J. Proteomics* **2012**, *77*, 68–86, doi: 10.1016/j.jprot.2012.07.005.
45. Hajduch, M.; Casteel, J.E.; Hurrelmeyer, K.E.; Song, Z.; Agrawal, G.K.; Thelen, J.J.; W, T.G.E.; Hajduch, M.; Casteel, J.E.; Hurrelmeyer, K.E.; Song, Z.; Agrawal, G.K.; Thelen, J.J. Proteomic analysis of seed filling in brassica napus . developmental characterization of metabolic isozymes using high-resolution. *Plant*

- Physiol.* **2006**, *141*, 32–46, doi: 10.1104/pp.105.075390.
46. Zhang, H.; Wang, W.Q.; Liu, S.J.; Møller, I.M.; Song, S.Q. Proteome analysis of poplar seed vigor. *PLoS One* **2015**, *10*, 1–19, doi: 10.1371/journal.pone.0132509.
 47. Ventura, L.; Donà, M.; Macovei, A.; Carbonera, D.; Buttafava, A.; Mondoni, A.; Rossi, G.; Balestrazzi, A. Understanding the molecular pathways associated with seed vigor. *Plant Physiol. Biochem.* **2012**, *60*, 196–206, doi: 10.1016/j.plaphy.2012.07.031.
 48. Catusse, J.; Meinhard, J.; Job, C.; Strub, J.M.; Fischer, U.; Pestsova, E.; Westhoff, P.; Van Dorselaer, A.; Job, D. Proteomics reveals potential biomarkers of seed vigor in sugarbeet. *Proteomics* **2011**, *11*, 1569–1580, doi: 10.1002/pmic.201000586.
 49. Rajjou, L.; Duval, M.; Gallardo, K.; Catusse, J.; Bally, J.; Job, C.; Job, D. Seed germination and vigor. *Annu. Rev. Plant Biol.* **2012**, *63*, 507–533, doi: 10.1146/annurev-arplant-042811-105550.
 50. Hatzig, S. V.; Frisch, M.; Breuer, F.; Nesi, N.; Ducournau, S.; Wagner, M.-H.; Leckband, G.; Abbadi, A.; Snowdon, R.J. Genome-wide association mapping unravels the genetic control of seed germination and vigor in brassica napus. *Front. Plant Sci.* **2015**, *6*, 1–13, doi: 10.3389/fpls.2015.00221.
 51. Xin, X.; Lin, X.-H.; Zhou, Y.-C.; Chen, X.-L.; Liu, X.; Lu, X.-X. Proteome analysis of maize seeds: the effect of artificial ageing. *Physiol. Plant.* **2011**, *143*, 126–138, doi: 10.1111/j.1399-3054.2011.01497.x.
 52. Smolikova, G.N.; Shavarda, A.L.; Alekseichuk, I. V.; Chantseva, V. V.; Medvedev, S.S. The metabolomic approach to the assessment of cultivar specificity of brassica napus l. seeds. *Russ. J. Genet. Appl. Res.* **2016**, *6*, 78–83, doi: 10.1134/S2079059716010147.
 53. Larance, M.; Lamond, A.I. Multidimensional proteomics for cell biology. *Nat. Rev. Mol. Cell Biol.* **2015**, *16*, 269–280, doi: 10.1038/nrm3970.
 54. Plomion, C.; Meddour, H.; Kohler, A.; Jacob, D.; Bastien, C.; Dreyer, E.; Daruvar, A. De; Guehl, J.; Schmitter, J.; Martin, F.; Bonneau, M. Mapping the proteome of poplar and application to the discovery of drought-stress responsive proteins. *Proteomics* **2006**, *6*, 6509–6527, doi: 10.1002/pmic.200600362.
 55. Bantscheff, M.; Schirle, M. Quantitative mass spectrometry in proteomics : a critical review. *Anal. Bioanal. Chem.* **2007**, *389*, 1017–1031, doi: 10.1007/s00216-007-1486-6.
 56. Jimenez-lopez, J.C.; Foley, R.C.; Brear, E.; Clarke, V.C.; Lima-cabello, E.; Florido, J.F.; Singh, K.B.; Alché, J.D.; Smith, P.M.C. Characterization of narrow-leaf lupin (*lupinus angustifolius* l .) recombinant major allergen ige-binding proteins and the natural β -conglutinin counterparts in sweet lupin seed species. *Food Chem.* **2018**, *244*, 60–70, doi: 10.1016/j.foodchem.2017.10.015.
 57. Li, X.; Wang, W.; Chen, J. Recent progress in mass spectrometry proteomics for biomedical research. *Sci. China Life Sci.* **2017**, *60*, 1093–1113, doi: 10.1007/s11427-017-9175-2.
 58. Vowinkel, J.; Zelez, A.; Bruderer, R.; Mülleder, M.; Reiter, L.; Ralser, M. Cost-effective generation of precise label-free quantitative proteomes in high-throughput by microlc and data-independent acquisition. *Sci. Rep.* **2018**, *8*, 4346, doi: 10.1038/s41598-018-22610-4.
 59. Hart-Smith, G.; Reis, R.S.; Waterhouse, P.M.; Wilkins, M.R. Improved quantitative plant proteomics via the combination of targeted and untargeted data acquisition. *Front. Plant Sci.* **2017**, *8*, 1669, doi: 10.3389/fpls.2017.01669.
 60. Weisser, H.; Wright, J.C.; Mudge, J.M.; Gutenbrunner, P.; Choudhary, J.S. A flexible data analysis pipeline for high-confidence proteogenomics. *J. Proteome Res.* **2016**, *15*, 4686–4695, doi: 10.1021/acs.jproteome.6b00765.
 61. Soboleva, A.; Schmidt, R.; Vikhnina, M.; Grishina, T.; Frolov, A. Maillard proteomics: opening new pages. *Int. J. Mol. Sci.* **2017**, *18*, 2677, doi: 10.3390/ijms18122677.
 62. Catherman, A.D.; Skinner, O.S.; Kelleher, N.L. Top down proteomics: facts and perspectives. *Biochem. Biophys. Res. Commun.* **2014**, *445*, 683–693, doi: 10.1016/j.bbrc.2014.02.041.
 63. Chmelik, J.; Zidkova, J.; Rehulka, P.; Petry-Podgorska, I.; Bobalova, J. Influence of different proteomic protocols on degree of high-coverage identification of nonspecific lipid transfer protein 1 modified during malting. *Electrophoresis* **2009**, *30*, 560–567, doi: 10.1002/elps.200800530.
 64. Gillet, L.C.; Leitner, A.; Aebersold, R. Mass spectrometry applied to bottom-up proteomics : entering the high-throughput era for hypothesis testing. *Annu. Rev. ofAnalytical Chem.* **2016**, *9*, 449–472, doi: 10.1146/annurev-anchem-071015-041535.
 65. Mørtz, E.; O'Connor, P.B.; Roepstorff, P.; Kelleher, N.L.; Wood, T.D.; McLafferty, F.W.; Mann, M. Sequence tag identification of intact proteins by matching tandem mass spectral data against sequence data bases. *Proceedings of the National Academy of Sciences of the United States of America*, 1996, *93*, 8264–8267.
 66. Hummel, M.; Wigger, T.; Brockmeyer, J. Characterization of mustard 2s albumin allergens by bottom-up , middle-down and top-down proteomics: a consensus set of isoforms of sin a 1. *J. Proteome Res.* **2015**, *14*, 1547–1556, doi: 10.1021/pr5012262.

67. Nilsen, T.W.; Graveley, B.R. Expansion of the eukaryotic proteome by alternative splicing. *Nature* **2010**, *463*, 457–463, doi: 10.1038/nature08909.
68. Tan, B.C.; Lim, Y.S.; Lau, S. Proteomics in commercial crops: an overview. *J. Proteomics* **2017**, *169*, 176–188, doi: 10.1016/j.jprot.2017.05.018.
69. Chen, D.; Shen, X.; Sun, L. Capillary zone electrophoresis-mass spectrometry with microliter-scale loading capacity, 140-min separation window and high peak capacity for bottom-up proteomics. *Analyst* **2017**, *142*, 2118–2127, doi: 10.1039/C7AN00509A.
70. Rabilloud, T.; Lelong, C. Two-dimensional gel electrophoresis in proteomics: a tutorial. *J. Proteomics* **2011**, *74*, 1829–1841, doi: 10.1016/j.jprot.2011.05.040.
71. Gladilovich, V.; Greifenhagen, U.; Sukhodolov, N.; Selyutin, A. Immobilized metal affinity chromatography on collapsed langmuir-blodgett iron (iii) stearate films and iron (iii) oxide nanoparticles for bottom-up phosphoproteomics immobilized metal affinity chromatography on collapsed langmuir-blodgett iron (iii). *J. Chromatogr. A* **2016**, *1443*, 181–190, doi: 10.1016/j.chroma.2016.03.044.
72. Rabilloud, T.; Chevallet, M.; Luche, S.; Lelong, C. Two-dimensional gel electrophoresis in proteomics : past , present and future. *J. Proteomics* **2010**, *73*, 2064–2077, doi: 10.1016/j.jprot.2010.05.016.
73. Gallardo, K.; Kurt, C.; Thompson, R.; Ochatt, S. In vitro culture of immature m. truncatula grains under conditions permitting embryo development comparable to that observed in vivo. *Plant Sci.* **2006**, *170*, 1052–1058, doi: 10.1016/j.plantsci.2005.12.021.
74. Paudel, G.; Bilova, T.; Schmidt, R.; Greifenhagen, U.; Berger, R.; Tarakhovskaya, E.; Stöckhardt, S.; Balcke, G.U.; Humbeck, K.; Brandt, W.; Sinz, A.; Vogt, T.; Birkemeyer, C.; Wessjohann, L.; Frolov, A. Osmotic stress is accompanied by protein glycation in *arabidopsis thaliana*. *J. Exp. Bot.* **2016**, *67*, 6283–6295, doi: 10.1093/jxb/erw395.
75. Isaacson, T.; Damasceno, C.M.B.; Saravanan, R.S.; He, Y.; Catalá, C.; Saladié, M.; Rose, J.K.C. Sample extraction techniques for enhanced proteomic analysis of plant tissues. *Nat. Protoc.* **2006**, *1*, 769–774, doi: 10.1038/nprot.2006.102.
76. Wang, W.; Vignani, R.; Scali, M.; Cresti, M. A universal and rapid protocol for protein extraction from recalcitrant plant tissues for proteomic analysis. *Electrophoresis* **2006**, *27*, 2782–2786, doi: 10.1002/elps.200500722.
77. Wang, X.; Li, X.; Deng, X.; Han, H.; Shi, W.; Li, Y. A protein extraction method compatible with proteomic analysis for the euhalophyte *salicornia europaea*. *Electrophoresis* **2007**, *28*, 3976–3987, doi: 10.1002/elps.200600805.
78. Soboleva, A.; Vikhnina, M.; Grishina, T.; Frolov, A. Probing protein glycation by chromatography and mass spectrometry: analysis of glycation adducts. *Int. J. Mol. Sci.* **2017**, *18*, 2557, doi: 10.3390/ijms18122557.
79. Damerval, C.; Dominique, de V.; Zivy, M.; Thiellement, H. Technical improvements in two-dimensional electrophoresis increase the level of genetic variation detected in wheat-seedling proteins. *Electrophoresis*, **1986**, *7*, 52–54.
80. Schuster, A.M.; Davies, E. Ribonucleic acid and protein metabolism in pea epicotyls. *Plant Physiol.* **1983**, *73*, 809–816, doi: 10.1104/pp.73.3.822.
81. Natarajan, S.; Xu, C.; Caperna, T.J.; Garrett, W.M. Comparison of protein solubilization methods suitable for proteomic analysis of soybean seed proteins. *Anal. Biochem.* **2005**, *342*, 214–220, doi: 10.1016/j.ab.2005.04.046.
82. Vilhena, M.B.; Franco, M.R.; Schmidt, D.; Carvalho, G.; Azevedo, R.A. Evaluation of protein extraction methods for enhanced proteomic analysis of tomato leaves and roots. *An. Acad. Bras. Cienc.* **2015**, *87*, 1853–1863, doi: 10.1590/0001-3765201520150116.
83. Görg, A.; Weiss, W. Two-dimensional electrophoresis with immobilized ph gradients. *Proteome Res. Two-Dimensional Gel Electrophor. Identif. Methods* **2000**, 57–106, doi: 10.1007/978-3-642-57105-3_4.
84. Vâlcu, C.-M.; Schlink, K. Reduction of proteins during sample preparation and two-dimensional gel electrophoresis of woody plant samples. *Proteomics* **2006**, *6*, 1599–1605, doi: 10.1002/pmic.200500314.
85. Toth, G.B.; Pavia, H. Removal of dissolved brown algal phlorotannins using insoluble polyvinylpyrrolidone (pvpp). *Journal of Chemical Ecology*, **2001**, *27*, 1899–1910.
86. Ranjbar Sistani, N.; Kaul, H.; Desalegn, G.; Wienkoop, S. Rhizobium impacts on seed productivity, quality, and protection of *pisum sativum* upon disease stress caused by *didymella pinodes*: phenotypic, proteomic, and metabolomic traits. *Front. Plant Sci.* **2017**, *8*, doi: 10.3389/fpls.2017.01961.
87. Dong, Y.; Wang, Q.; Zhang, L.; Du, C.; Xiong, W. Dynamic proteomic characteristics and network integration revealing key proteins for two kernel tissue developments in popcorn. *PLoS One* **2015**, *10*, e0143181, doi: 10.1371/journal.pone.0143181.

88. Cremer, F.; van de Walle, C. Method for extraction of proteins from green plant tissues for two-dimensional polyacrylamide gel electrophoresis. *Analytical Biochemistry*, 1985, 147, 22–26.
89. Milkovska-Stamenova, S.; Schmidt, R.; Frolov, A.; Birkemeyer, C. GC-ms method for the quantitation of carbohydrate intermediates in glycation systems. *J. Agric. Food Chem.* **2015**, 63, 5911–5919, doi: 10.1021/jf505757m.
90. Frolov, A.; Schmidt, R.; Spiller, S.; Greifenhagen, U.; Hoffmann, R. Arginine-derived advanced glycation end products generated in peptide–glucose mixtures during boiling. *J. Agric. Food Chem.* **2014**, 62, 3626–3635, doi: 10.1021/jf4050183.
91. Greifenhagen, U.; Frolov, A.; Blüher, M.; Hoffmann, R. Plasma proteins modified by advanced glycation end products (ages) reveal site-specific susceptibilities to glycemic control in patients with type 2 diabetes. *J. Biol. Chem.* **2016**, 291, 9610–9616, doi: 10.1074/jbc.M115.702860.
92. Capriotti, A.L.; Cavaliere, C.; Piovesana, S.; Stampachiachiere, S.; Ventura, S.; Zenezini Chiozzi, R.; Laganà, A. Characterization of quinoa seed proteome combining different protein precipitation techniques: improvement of knowledge of nonmodel plant proteomics. *J. Sep. Sci.* **2015**, 38, 1017–1025, doi: 10.1002/jssc.201401319.
93. Nishimura, N.; Tsuchiya, W.; Moresco, J.J.; Hayashi, Y.; Satoh, K.; Kaiwa, N.; Irisa, T.; Kinoshita, T.; Schroeder, J.L.; Yates, J.R.; Hirayama, T.; Yamazaki, T. Control of seed dormancy and germination by dog1-ahg1 pp2c phosphatase complex via binding to heme. *Nat. Commun.* **2018**, 9, doi: 10.1038/s41467-018-04437-9.
94. Lorenz, C.; Brandt, S.; Borisjuk, L.; Rolletschek, H.; Heinzl, N.; Tohge, T.; Fernie, A.R.; Braun, H.-P.; Hildebrandt, T.M. The role of persulfide metabolism during arabidopsis seed development under light and dark conditions. *Front. Plant Sci.* **2018**, 9, 1–15, doi: 10.3389/fpls.2018.01381.
95. Wang, W.; Vignani, R.; Scali, M.; Sensi, E.; Tiberi, P.; Cresti, M. Removal of lipid contaminants by organic solvents from oilseed protein extract prior to electrophoresis. *Anal. Biochem.* **2004**, 329, 139–141, doi: 10.1016/j.ab.2004.02.044.
96. Murad, A.M.; Rech, E.L. NanoUPLC-mse proteomic data assessment of soybean seeds using the uniprot database. *BMC Biotechnol.* **2012**, 12, 82, doi: 10.1186/1472-6750-12-82.
97. Satour, P.; Youssef, C.; Châtelain, E.; Vu, B.L.; Teulat, B.; Job, C.; Job, D.; Montrichard, F. Patterns of protein carbonylation during medicago truncatula seed maturation. *Plant. Cell Environ.* **2018**, 41, 2183–2194, doi: 10.1111/pce.13194.
98. Tang, H.; Ming, Z.; Liu, R.; Xiong, T.; Grevelding, C.G.; Dong, H.; Jiang, M. Development of adult worms and granulomatous pathology are collectively regulated by t- and b-cells in mice infected with schistosoma japonicum. *PLoS One* **2013**, 8, e54432, doi: 10.1371/journal.pone.0054432.
99. Rakwal, R.; Hayashi, G.; Shibato, J.; Deepak, S.A.; Gundimeda, S.; Simha, U.; Padmanaban, A.; Gupta, R.; Han, S.; Kim, S.T.; Kubo, A.; Imanaka, T.; Fukumoto, M.; Agrawal, G.K.; Shioda, S. Progress toward rice seed omics in low-level gamma radiation environment in iitate village, fukushima. *J. Hered.* **2018**, 109, 206–211, doi: 10.1093/jhered/esx071.
100. Klose, J.; Kobalz, U. Two-dimensional electrophoresis of proteins: an updated protocol and implications for a functional analysis of the genome. *Electrophoresis* **1995**, doi: 10.1002/elps.11501601175.
101. Gygi, S.P.; Corthals, G.L.; Zhang, Y.; Rochon, Y.; Aebersold, R. Evaluation of two-dimensional gel electrophoresis-based proteome analysis technology. *Proc. Natl. Acad. Sci.* **2000**, 97, 9390–9395, doi: 10.1073/pnas.160270797.
102. Neuheff, V.; Arold, N.; Taube, D.; Ehrhardt, W. Dependence of particle and fiber properties. *Electrophoresis* **1988**, 9, 255–262, doi: 10.1002/elps.11501601175.
103. Frolov, A.; Blüher, M.; Hoffmann, R. Glycation sites of human plasma proteins are affected to different extents by hyperglycemic conditions in type 2 diabetes mellitus. *Anal. Bioanal. Chem.* **2014**, 406, 5755–5763, doi: 10.1007/s00216-014-8018-y.
104. Chevalier, F. Standard dyes for total protein staining in gel-based proteomic analysis. *Materials (Basel)*. **2010**, 3, 4784–4792, doi: 10.3390/ma3104784.
105. Mortz, E.; Krogh, T.N.; Vorum, H.; Görg, A. Improved silver staining protocols for high sensitivity protein identification using matrix-assisted laser desorption/ionization-time of flight analysis. *Proteomics* **2001**, 1, 1359–1363, doi: 10.1002/1615-9861(200111)1:11<1359::AID-PROT1359>3.0.CO;2-Q.
106. Chevallet, M.; Luche, S.; Strub, J.; Dorselaer, A. Van; Rabilloud, T. Sweet silver : a formaldehyde-free silver staining using aldoses as developing agents , with enhanced compatibility with mass spectrometry. *Proteomics* **2008**, 8, 4853–4861, doi: 10.1002/pmic.200800321.
107. Berggren, K.; Chernokalskaya, E.; Steinberg, T.H.; Kemper, C.; Lopez, M.F.; Diwu, Z.; Haugland, R.P.; Patton, W.F. Background-free, high sensitivity staining of proteins in one- and two-dimensional sodium dodecyl sulfate-polyacrylamide gels using a luminescent ruthenium complex. *Electrophoresis* **2000**, 21,

- 2509–2521, doi: 10.1002/1522-2683(20000701)21:12<2509::AID-ELPS2509>3.0.CO;2-9.
108. Rabilloud, T.; Strub, J.; Luche, S.; Dorsselaer, A. Van; Lunardi, J. A comparison between sypro ruby and ruthenium ii tris (bathophenanthroline disulfonate) as. *Proteomics* **2001**, *1*, 699–704, doi: 1615-9853/01/0505–699.
 109. Zhou, G.; Li, H.; DeCamp, D.; Chen, S.; Shu, H.; Gong, Y.; Flaig, M.; Gillespie, J.W.; Hu, N.; Taylor, P.R.; Emmert-Buck, M.R.; Liotta, L.A.; Petricoin, E.F.; Zhao, Y. 2D differential in-gel electrophoresis for the identification of esophageal scans cell cancer-specific protein markers. *Mol. Cell. Proteomics* **2002**, *1*, 117–123, doi: 10.1074/mcp.M100015-MCP200.
 110. Aslebagh, R.; Channaveerappa, D.; Arcaro, K.F.; Darie, C.C. Comparative two-dimensional polyacrylamide gel electrophoresis (2d-page) of human milk to identify dysregulated proteins in breast cancer. *Electrophoresis* **2018**doi: 10.1002/elps.201800025.
 111. Rosenfeld, J.; Capdevielle, J.; Guillemot, J.C.; Ferrara, P. In-gel digestion of proteins for internal analysis after one- or two-dimensional gel electrophoresis. *Analytical Biochemistry*, 1992, *203*, 173–179.
 112. Blum, H.; Beier, H.; Gross, H.J. Improved silver staining of plant proteins , rna and dna in polyacrylamide gels. *Electrophoresis*, 1987, *8*, 93–99.
 113. Ogura, T.; Ogihara, J.; Sunairi, M.; Takeishi, H.; Aizawa, T.; Olivos-Trujillo, M.R.; Maureira-Butler, I.J.; Salvo-Garrido, H.E. Proteomic characterization of seeds from yellow lupin (*lupinus luteus* l.). *Proteomics* **2014**, *14*, 1543–1546, doi: 10.1002/pmic.201300511.
 114. Gomes, L.S.; Senna, R.; Sandim, V.; Silva-Neto, M.A.C.; Perales, J.E.A.; Zingali, R.B.; Soares, M.R.; Fialho, E. Four conventional soybean [glycine max (l.) merrill] seeds exhibit different protein profiles as revealed by proteomic analysis. *J. Agric. Food Chem.* **2014**, *62*, 1283–1293, doi: 10.1021/jf404351g.
 115. Min, C.W.; Lee, S.H.; Cheon, Y.E.; Han, W.Y.; Ko, J.M.; Kang, H.W.; Kim, Y.C.; Agrawal, G.K.; Rakwal, R.; Gupta, R.; Kim, S.T. In-depth proteomic analysis of glycine max seeds during controlled deterioration treatment reveals a shift in seed metabolism. *J. Proteomics* **2017**, *169*, 125–135, doi: 10.1016/j.jprot.2017.06.022.
 116. Swart, C.; Martínez-Jaime, S.; Gorka, M.; Zander, K.; Graf, A. Hit-gel: streamlining in-gel protein digestion for high-throughput proteomics experiments. *Sci. Rep.* **2018**, *8*, 8582, doi: 10.1038/s41598-018-26639-3.
 117. Thiede, B.; Höhenwarter, W.; Krah, A.; Mattow, J.; Schmid, M.; Schmidt, F.; Jungblut, P.R. Peptide mass fingerprinting. *Methods.* **2005**, *35*(3), 237–47, doi: 10.1016/j.ymeth.2004.08.015
 118. Clauser, K.R.; Baker, P.; Burlingame, A.L. Role of accurate mass measurement (10 ppm) in protein identification strategies employing ms or ms/ms and database searching. *Analytical Chemistry*, 1999, *71*, 2871–2882.
 119. Binita, K.; Kumar, S.; Sharma, V.K.; Sharma, V.; Yadav, S. Proteomic identification of syzygium cumini seed extracts by maldi-tof / ms. *Appl. Biochem. Biotechnol.* **2014**, *172*, 2091–2105, doi: 10.1007/s12010-013-0660-x.
 120. Yaprak, E.; Kasap, M.; Akpınar, G.; Kayaaltı-Yüksek, S.; Sinanoğlu, A.; Guzel, N.; Demirturk Kocasarac, H. The prominent proteins expressed in healthy gingiva: a pilot exploratory tissue proteomics study. *Odontology* **2018**, *106*, 19–28, doi: 10.1007/s10266-017-0302-9.
 121. Vadivel, A.K.A. Gel-based proteomics in plants : time to move on from the tradition. *Front. Plant Sci.* **2015**, *6*, 2–5, doi: 10.3389/fpls.2015.00369.
 122. Maaß, S.; Becher, D. Methods and applications of absolute protein quantification in microbial systems. *J. Proteomics* **2016**, *136*, 222–233, doi: 10.1016/j.jprot.2016.01.015.
 123. Nedenskov Jensen, K.; Jessen, F.; Jørgensen, B.M. Multivariate data analysis of two-dimensional gel electrophoresis protein patterns from few samples †. *J. Proteome Res.* **2008**, *7*, 1288–1296, doi: 10.1021/pr700800s.
 124. Marengo, E.; Robotti, E.; Antonucci, F.; Cecconi, D.; Campostrini, N.; Righetti, P.G. Numerical approaches for quantitative analysis of two-dimensional maps : a review of commercial software and home-made systems. *Proteome Sci.* **2005**, *5*, 654–666, doi: 10.1002/pmic.200401015.
 125. Voss, T.; Haberl, P. Observations on the reproducibility and matching efficiency of two-dimensional electrophoresis gels : consequences for comprehensive data analysis. *Electrophoresis*, 2000, *21*, 3345–3350.
 126. Arentz, G.; Weiland, F.; Oehler, M.K.; Hoffmann, P. State of the art of 2d dige. *Proteomics - Clin. Appl.* **2015**, *9*, 277–288, doi: 10.1002/prca.201400119.This.
 127. Beckett, P. The basics of 2d dige. *Methods Mol. Biol.* **2012**, *854*, 9–19, doi: 10.1007/978-1-61779-573-2_2.
 128. Morgan, M.E.; Minden, J.S. Difference gel electrophoresis : a single gel method for detecting changes in protein extracts. *Electrophoresis*, 1997, *18*, 2071–2077.
 129. Krokhn, O. V; Ens, W.; Standing, K.G. MALDI qqt of ms combined with off-line hplc for characterization

- of protein primary structure and post-translational modifications. *Journal of Biomolecular Technique*, 2005, 16, 427–438.
130. Köcher, T.; Pichler, P.; Swart, R.; Mechtler, K. Analysis of protein mixtures from whole-cell extracts by single-run nanolc-ms/ms using ultralong gradients. *Nat. Protoc.* **2012**, 7, 882–890, doi: 10.1038/nprot.2012.036.
 131. Taylor, P.J. Matrix effects: the achilles heel of quantitative high-performance liquid chromatography–electrospray–tandem mass spectrometry. *Clin. Biochem.* **2005**, 38, 328–334, doi: 10.1016/j.clinbiochem.2004.11.007.
 132. Frolov, A.; Henning, A.; Böttcher, C.; Tissier, A.; Strack, D. An uplc-ms/ms method for the simultaneous identification and quantitation of cell wall phenolics in brassica napus seeds. *J. Agric. Food Chem.* **2013**, 61, 1219–1227, doi: 10.1021/jf3042648.
 133. Spiller, S.; Frolov, A.; Hoffmann, R. Quantification of specific glycation sites in human serum albumin as prospective type 2 diabetes mellitus biomarkers. *Protein Pept. Lett.* **2018**, 24, 887–896, doi: 10.2174/0929866524666170202124120.
 134. Vissers, J.P.C.; Chervet, J.; Salzmann, J. Sodium dodecyl sulphate removal from tryptic digest samples for on-line capillary liquid. *J. mass Spectrom.* **1996**, 31, 1021–1027, doi: 10.1076-51 74/96/091021-07.
 135. Cole, R.B. *Electrospray and MALDI Mass Spectrometry: Fundamentals, Instrumentation, Practicalities, and Biological Applications: Second Edition*. *Electrospray MALDI Mass Spectrom. Fundam. Instrumentation, Pract. Biol. Appl. Second Ed.* **2012**; doi: 10.1002/9780470588901.
 136. Chen, E.I.; Cociorva, D.; Norris, J.L.; Yates, J.R. Optimization of mass spectrometry-compatible surfactants for shotgun proteomics. *J. Proteome Res.* **2007**, 6, 2529–2538, doi: 10.1021/pr060682a.
 137. Lin, Y.; Jiang, H.; Yan, Y.; Peng, B.; Chen, J.; Lin, H.; Liu, Z. Shotgun analysis of membrane proteomes by an improved sds-assisted sample preparation method coupled with liquid chromatography–tandem mass spectrometry. *J. Chromatogr. B* **2012**, 911, 6–14, doi: 10.1016/j.jchromb.2012.10.016.
 138. Bilova, T.; Greifenhagen, U.; Paudel, G.; Lukasheva, E.; Brauch, D.; Osmolovskaya, N.; Tarakhovskaya, E.; Balcke, G.U.; Tissier, A.; Vogt, T.; Milkowski, C.; Birkemeyer, C.; Wessjohann, L.; Frolov, A. Glycation of plant proteins under environmental stress — methodological approaches, potential mechanisms and biological role. *Abiotic Biot. Stress Plants - Recent Adv. Futur. Perspect.* **2016**; doi: 10.5772/61860.
 139. Bilova, T.; Lukasheva, E.; Brauch, D.; Greifenhagen, U.; Paudel, G.; Tarakhovskaya, E.; Frolova, N.; Mittasch, J.; Balcke, G.U.; Tissier, A.; Osmolovskaya, N.; Vogt, T.; Wessjohann, L.A.; Birkemeyer, C.; Milkowski, C.; Frolov, A. A snapshot of the plant glycosylated proteome: structural, functional and mechanistic aspect. *J. Biol. Chem.* **2016**, 291, 7621–7636, doi: 10.1074/jbc.M115.678581.
 140. Greifenhagen, U.; Frolov, A.; Blüher, M.; Hoffmann, R. Site-specific analysis of advanced glycation end products in plasma proteins of type 2 diabetes mellitus patients. *Anal. Bioanal. Chem.* **2016**; doi: 10.1007/s00216-016-9651-4.
 141. Kalli, A.; Smith, T.; Sweredoski, M.J.; Hess, S. Evaluation and optimization of mass spectrometric settings during data-dependent acquisition mode: focus on ltq-orbitrap mass analyzers. *J. Proteome Res.* **2013**, 12, 3071–3086, doi: 10.1021/pr3011588.
 142. Majovsky, P.; Naumann, C.; Lee, C.; Lassowskat, I.; Trujillo, M.; Dissmeyer, N.; Hoehenwarter, W. Targeted proteomics analysis of protein degradation in plant signaling on an ltq-orbitrap mass spectrometer. *Journal of Proteome Research*, **2014**, 13, 4246–4258, doi: 10.1021/pr500164j
 143. Wilm, M. Quantitative proteomics in biological research. *Proteomics* **2009**, 9, 4590–4605, doi: 10.1002/pmic.200900299.
 144. Guo, G.; Li, N. Relative and accurate measurement of protein abundance using ¹⁵n stable isotope labeling in arabidopsis (silia). *Phytochemistry* **2011**, 72, 1028–1039, doi: https://doi.org/10.1016/j.phytochem.2011.01.007.
 145. Yao, X. Derivatization or not: a choice in quantitative proteomics. *Anal. Chem.* **2011**, 83, 4427–4439, doi: 10.1021/ac200925p.
 146. Soboleva, A.; Modzel, M.; Didio, A.; Płóciennik, H.; Kijewska, M.; Grischina, T.; Karonova, T.; Bilova, T.; Stefanov, V.; Stefanowicz, P.; Frolov, A. Quantification of prospective type 2 diabetes mellitus biomarkers by stable isotope dilution with bi-labeled standard glycosylated peptides. *Anal. Methods* **2017**, 9, 409–418, doi: 10.1039/C6AY02483A.
 147. Handrick, V.; Vogt, T.; Frolov, A. Profiling of hydroxycinnamic acid amides in arabidopsis thaliana pollen by tandem mass spectrometry. *Anal. Bioanal. Chem.* **2010**, 398, 2789–2801, doi: 10.1007/s00216-010-4129-2.
 148. Wiśniewski, J.R.; Zougman, A.; Nagaraj, N.; Mann, M. Universal sample preparation method for proteome analysis. *Nat. Methods* **2009**, 6, 359–362, doi: 10.1038/nmeth.1322.
 149. George, I.S.; Fennell, A.Y.; Haynes, P.A. Protein identification and quantification from riverbank grape,

- 1068 vitis riparia : comparing sds-page and fasp-gpf techniques for shotgun proteomic analysis. *Proteomics*
 1069 **2015**, *15*, 3061–3065, doi: 10.1002/pmic.201500085.
- 1070 150. Zhang, N.; Zhang, H.-J.; Sun, Q.-Q.; Cao, Y.-Y.; Li, X.; Zhao, B.; Wu, P.; Guo, Y.-D. Proteomic analysis
 1071 reveals a role of melatonin in promoting cucumber seed germination under high salinity by regulating
 1072 energy production. *Sci. Rep.* **2017**, *7*, 503, doi: 10.1038/s41598-017-00566-1.
- 1073 151. Lu, X.; Zhu, H. Tube-gel digestion. *Mol. Cell. Proteomics* **2005**, *4*, 1948–1958, doi: 10.1074/mcp.M500138-
 1074 MCP200.
- 1075 152. Fischer, R.; Kessler, B.M. Gel-aided sample preparation (gasp)-a simplified method for gel-assisted
 1076 proteomic sample generation from protein extracts and intact cells. *Proteomics* **2015**, *15*, 1224–1229, doi:
 1077 10.1002/pmic.201400436.
- 1078 153. Strader, M.B.; Tabb, D.L.; Hervey, W.J.; Pan, C.; Hurst, G.B. Efficient and specific trypsin digestion of
 1079 microgram to nanogram quantities of proteins in organic–aqueous solvent systems. *Anal. Chem.* **2006**,
 1080 *78*, 125–134, doi: 10.1021/ac051348l.
- 1081 154. Zhang, Y.; Fonslow, B.R.; Shan, B.; Baek, M.-C.; Yates, J.R. Protein analysis by shotgun/bottom-up
 1082 proteomics. *Chem. Rev.* **2013**, *113*, 2343–2394, doi: 10.1021/cr3003533.
- 1083 155. Wang, W.-Q.; Jensen, O.N.; Møller, I.M.; Hebelstrup, K.H.; Rogowska-Wrzesinska, A. Evaluation of
 1084 sample preparation methods for mass spectrometry-based proteomic analysis of barley leaves. *Plant*
 1085 *Methods* **2018**, *14*, 72, doi: 10.1186/s13007-018-0341-4.
- 1086 156. Kołodziejczyk, I.; Dzitko, K.; Szewczyk, R.; Posmyk, M.M. Exogenous melatonin expediently modifies
 1087 proteome of maize (*zea mays* l.) embryo during seed germination. *Acta Physiol. Plant.* **2016**, *38*, 146, doi:
 1088 10.1007/s11738-016-2166-y.
- 1089 157. Lin, Y.; Huo, L.; Liu, Z.; Li, J.; Liu, Y.; He, Q.; Wang, X.; Liang, S. Sodium laurate, a novel protease- and
 1090 mass spectrometry-compatible detergent for mass spectrometry-based membrane proteomics. *PLoS One*
 1091 **2013**, *8*, e59779, doi: 10.1371/journal.pone.0059779.
- 1092 158. Norris, J.L.; Porter, N.A.; Caprioli, R.M. Mass spectrometry of intracellular and membrane proteins using
 1093 cleavable detergents. *Anal. Chem.* **2003**, *75*, 6642–6647, doi: 10.1021/ac034802z.
- 1094 159. Zakowick, H.; Schagat, T.; Yoder, D.; Niles, A.L. Measuring cell health and viability sequentially by
 1095 same-well multiplexing using the glomax®-multi detection system. *Promega Notes*, 2008, *99*, 25–28.
- 1096 160. *Technical Bulletin: ProteaseMAX(TM) Surfactant, Trypsin Enhancer*; Promega corporation, 2015, p. 19,
 1097 ISBN 6082744330.
- 1098 161. Li, M.; Powell, M.J.; Razunguzwa, T.T.; O'doherty, G.A. A general approach to anionic acid-labile
 1099 surfactants with tunable properties. *J. Org. Chem.* **2010**, *75*, 6149–6153, doi: 10.1021/jo100954q.
- 1100 162. Ross, A.R.S.; Lee, P.J.; Smith, D.L.; Langridge, J.I.; Whetton, A.D.; Gaskell, S.J. Identification of proteins
 1101 from two-dimensional polyacrylamide gels using a novel acid-labile surfactant. *Proteomics* **2002**, *2*, 928–
 1102 936, doi: 1615-9853/02/0707-928.
- 1103 163. Jagadeeshaprasad, M.G.; Batkulwar, K.B.; Meshram, N.N.; Tiwari, S.; Korwar, A.M.; Unnikrishnan,
 1104 A.G.; Kulkarni, M.J. Targeted quantification of n-1-(carboxymethyl) valine and n-1-(carboxyethyl) valine
 1105 peptides of β -hemoglobin for better diagnostics in diabetes. *Clin. Proteomics* **2016**, *13*, 7, doi:
 1106 10.1186/s12014-016-9108-y.
- 1107 164. Kaspar-Schoenefeld, S.; Merx, K.; Jozefowicz, A.M.; Hartmann, A.; Seiffert, U.; Weschke, W.; Matros, A.;
 1108 Mock, H. Label-free proteome profiling reveals developmental-dependent patterns in young barley
 1109 grains. *J. Proteomics* **2016**, *143*, 106–121, doi: 10.1016/j.jpro.2016.04.007.
- 1110 165. *Increase Analytical Accuracy LC/MS: Solvents, Blends, Standards, Surfactants*; ThermoFisher Scientific, p.
 1111 <https://beta-static.fishersci.com/content/dam/fish/>.
- 1112 166. Waas, M.; Bhattacharya, S.; Chuppa, S.; Wu, X.; Jensen, D.R.; Omasits, U.; Wollscheid, B.; Volkman, B.F.;
 1113 Noon, K.R.; Gundry, R.L. Combine and conquer: surfactants, solvents, and chaotropes for robust mass
 1114 spectrometry based analyses of membrane proteins. *Anal. Chem.* **2014**, *86*, 1551–1559, doi:
 1115 10.1021/ac403185a.
- 1116 167. Frolov, A.; Hoffmann, R. Analysis of amadori peptides enriched by boronic acid affinity
 1117 chromatography. *Ann. N. Y. Acad. Sci.* **2008**, *1126*, 253–256, doi: 10.1196/annals.1433.060.
- 1118 168. Kim, Y.J.; Wang, Y.; Gupta, R.; Kim, S.W.; Min, C.W.; Kim, Y.C.; Park, K.H.; Agrawal, G.K.; Rakwal, R.;
 1119 Choung, M.-G.; Kang, K.Y.; Kim, S.T. Protamine sulfate precipitation method depletes abundant plant
 1120 seed-storage proteins: a case study on legume plants. *Proteomics* **2015**, *15*, 1760–1764, doi:
 1121 10.1002/pmic.201400488.This.
- 1122 169. Donato, P.; Cacciola, F.; Tranchida, P.Q.; Dugo, P.; Mondello, L. MASS spectrometry detection in
 1123 comprehensive liquid chromatography : basic concepts , instrumental aspects , applications and trends.
 1124 *Mass Spectrom. Rev.* **2012**, *31*, 523–559, doi: 10.1002/mas.
- 1125 170. Bollineni, R.C.; Hoffmann, R.; Fedorova, M. Identification of protein carbonylation sites by two-

- dimensional liquid chromatography in combination with maldi- and esi-ms. *J. Proteomics* **2011**, *74*, 2338–2350, doi: 10.1016/j.jprot.2011.07.002.
171. Canterbury, J.D.; Merrihew, G.E.; Maccoss, M.J.; Goodlett, D.R.; Shaffer, S.A. Comparison of data acquisition strategies on quadrupole ion trap instrumentation for shotgun proteomics. *J. Am. Soc. Mass Spectrom.* **2014**, *25*, 2048–2059, doi: 10.1007/s13361-014-0981-1.
172. Manual 2D Quant Kit https://www.gelifesciences.com/gehcls_images/GELS/RelatedContent/Files/1314729545976/litdoc28954714AE_20110830215136.pdf.
173. Matamoros, M.A.; Kim, A.; Peñuelas, M.; Ihling, C.; Griesser, E.; Hoffmann, R.; Fedorova, M.; Frolov, A.; Becana, M. Protein carbonylation and glycation in legume nodules. *Plant Physiol.* **2018**, *177*, pp.00533.2018, doi: 10.1104/pp.18.00533.
174. Plumb, R.S.; Johnson, K.A.; Rainville, P.; Smith, B.W.; Wilson, I.D.; Castro-Perez, J.M.; Nicholson, J.K. UPLC/mse; a new approach for generating molecular fragment information for biomarker structure elucidation. *Rapid Commun. Mass Spectrom.* **2006**, *20*, 1989–1994, doi: 10.1002/rcm.2550.
175. Ludwig, C.; Gillet, L.; Rosenberger, G.; Amon, S.; Collins, B.C.; Aebersold, R. Data- independent acquisition-based SWATH-MS for quantitative proteomics: a tutorial. *Mol. Syst. Biol.* **2018**, *14*, e8126, doi: 10.15252/msb.20178126.
176. Gillet, L.C.; Navarro, P.; Tate, S.; Ro, H.; Selevsek, N.; Reiter, L.; Bonner, R.; Aebersold, R. Targeted data extraction of the ms / ms spectra generated by data-independent acquisition: a new concept for consistent and accurate proteome analysis. *Mol. Cell. Proteomics* **2012**, *11*, 1–17, doi: 10.1074/mcp.O111.016717.
177. Bateman, N.W.; Goulding, S.P.; Shulman, N.J.; Gadok, A.K.; Szumlinski, K.K.; MacCoss, M.J.; Wu, C.C. Maximizing peptide identification events in proteomic workflows using data-dependent acquisition (dda). *Mol. Cell. Proteomics* **2014**, *13*, 329–338, doi: 10.1074/mcp.M112.026500.
178. Zhu, W.; Smith, J.W.; Huang, C. Mass spectrometry-based label-free quantitative proteomics. *J. Biomed. Biotechnol.* **2010**, *2010*, 1–6, doi: 10.1155/2010/840518.
179. Jamwal, R.; Barlock, B.J.; Adusumalli, S.; Ogasawara, K.; Simons, B.L.; Akhlaghi, F. Multiplex and label-free relative quantification approach for studying protein abundance of drug metabolizing enzymes in human liver microsomes using swath-ms. *J. Proteome Res.* **2017**, *16*, 4134–4143, doi: 10.1021/acs.jproteome.7b00505.
180. Kelstrup, C.D.; Bekker-jensen, D.B.; Arrey, T.N.; Hogrebe, A.; Harder, A.; Olsen, J. V. Performance evaluation of the q exactive hf - x for shotgun proteomics. *J. Proteome Res.* **2018**, *17*, 727–738, doi: 10.1021/acs.jproteome.7b00602.
181. Scheltema, R.A.; Hauschild, J.; Lange, O.; Hornburg, D.; Denisov, E.; Damoc, E.; Kuehn, A.; Makarov, A.; Mann, M.; Exactive, Q. The q exactive hf , a benchtop mass spectrometer with a pre-filter , high-performance quadrupole and an ultra-high-field orbitrap. *Mol Cell Proteomics* **2014**, *13*, 3698–3708, doi: 10.1007/978-94-011-4431-5_1.
182. Nakamura, K.; Hirayama-Kurogi, M.; Ito, S.; Kuno, T.; Yoneyama, T.; Obuchi, W.; Terasaki, T.; Ohtsuki, S. Large-scale multiplex absolute protein quantification of drug-metabolizing enzymes and transporters in human intestine, liver, and kidney microsomes by swath-ms: comparison with mrm/srm and hr-mrm/prm. *Proteomics* **2016**, *16*, 2106–2117, doi: 10.1002/pmic.201500433.
183. Peterson, A. C.; Russell, J.D.; Bailey, D.J.; Westphall, M.S.; Coon, J.J. Parallel reaction monitoring for high resolution and high mass accuracy quantitative, targeted proteomics. *Mol Cell Proteomics.*, **2012**, *11*(11), 1475–88, doi: 10.1074/mcp.O112.020131.
184. Mann, M.; Højrup, P.; Roepstorff, P. Use of mass spectrometric molecular weight information to identify proteins in sequence databases. *Biol. Mass Spectrom.* **1993**, *22*, 338–345, doi: 10.1002/bms.1200220605.
185. Henzel, W.J.; Watanabe, C.; Stults, J.T. Protein identification: the origins of peptide mass fingerprinting. *J. Am. Soc. Mass Spectrom.* **2003**, *14*, 931–942, doi: 10.1016/S1044-0305(03)00214-9.
186. Eng, J.K.; McCormack, A.L.; Yates, J.R. An approach to correlate tandem mass spectral data of peptides with amino acid sequences in a protein database. *J. Am. Soc. Mass Spectrom.* **1994**, *5*, 976–989, doi: 10.1016/1044-0305(94)80016-2.
187. Cottrell, J.S. Protein identification using ms/ms data. *J. Proteomics* **2011**, *74*, 1842–1851, doi: 10.1016/j.jprot.2011.05.014.
188. Stein, S.E.; Scott, D.R. Optimization and testing of mass spectral library search algorithms for compound identification. *J. Am. Soc. Mass Spectrom.* **1994**, *5*, 859–866, doi: 10.1016/1044-0305(94)87009-8.
189. Collisionally induced dissociation of protonated peptide ions and the interpretation of product ion spectra. *Protein Seq. Identif. Using Tandem Mass Spectrom.* **2005**, 64–116, doi: 10.1002/0471721980.ch4.
190. Mann, M.; Wilm, M. Error-tolerant identification of peptides in sequence databases by peptide sequence tags. *Anal. Chem.* **1994**, *66*, 4390–4399, doi: 10.1021/ac00096a002.

- 1184 191. Tabb, D.L. The sequest family tree. *J. Am. Soc. Mass Spectrom.* **2015**, *26*, 1814–1819, doi: 10.1007/s13361-
1185 015-1201-3.
- 1186 192. Perkins, D.N.; Pappin, D.J.C.; Creasy, D.M.; Cottrell, J.S. Probability-based protein identification by
1187 searching sequence databases using mass spectrometry data. *Electrophoresis* **1999**, *20*, 3551–3567, doi:
1188 10.1002/(SICI)1522-2683(19991201)20:18<3551::AID-ELPS3551>3.0.CO;2-2.
- 1189 193. Cox, J.; Neuhauser, N.; Michalski, A.; Scheltema, R.A.; Olsen, J. V.; Mann, M. Andromeda: a peptide
1190 search engine integrated into the maxquant environment. *J. Proteome Res.* **2011**, *10*, 1794–1805, doi:
1191 10.1021/pr101065j.
- 1192 194. Eng, J.K.; Searle, B.C.; Clauser, K.R.; Tabb, D.L. A face in the crowd: recognizing peptides through
1193 database search. *Mol. Cell. Proteomics* **2011**, *10*, R111.009522, doi: 10.1074/mcp.R111.009522.
- 1194 195. Moore, R.E.; Young, M.K.; Lee, T.D. Qscore: an algorithm for evaluating sequest database search results.
1195 *J. Am. Soc. Mass Spectrom.* **2002**, *13*, 378–386, doi: 10.1016/S1044-0305(02)00352-5.
- 1196 196. Peng, J.; Elias, J.E.; Thoreen, C.C.; Licklider, L.J.; Gygi, S.P. Evaluation of multidimensional
1197 chromatography coupled with tandem mass spectrometry (lc/lc–ms/ms) for large-scale protein analysis:
1198 the yeast proteome. *J. Proteome Res.* **2003**, *2*, 43–50, doi: 10.1021/pr025556v.
- 1199 197. Reiter, L.; Claassen, M.; Schrimpf, S.P.; Jovanovic, M.; Schmidt, A.; Buhmann, J.M.; Hengartner, M.O.;
1200 Aebersold, R. Protein identification false discovery rates for very large proteomics data sets generated
1201 by tandem mass spectrometry. *Mol. Cell. Proteomics* **2009**, *8*, 2405–2417, doi: 10.1074/mcp.M900317-
1202 MCP200.
- 1203 198. Savitski, M.M.; Wilhelm, M.; Hahne, H.; Kuster, B.; Bantscheff, M. A scalable approach for protein false
1204 discovery rate estimation in large proteomic data sets. *Mol. Cell. Proteomics* **2015**, *14*, 2394–2404, doi:
1205 10.1074/mcp.M114.046995.
- 1206 199. Elias, J.E.; Gygi, S.P. Target-decoy search strategy for increased confidence in large-scale protein
1207 identifications by mass spectrometry. *Nat. Methods* **2007**, *4*, 207–214, doi: 10.1038/nmeth1019.
- 1208 200. Sinitcyn, P.; Rudolph, J.D.; Cox, J. Computational methods for understanding mass spectrometry-based
1209 shotgun proteomics data. *Annu. Rev. Biomed. Data Sci.* **2018**, *1*, 207–234, doi: 10.1146/annurev-biodatasci-
1210 080917-013516.
- 1211 201. Angel, T.E.; Aryal, U.K.; Hengel, S.M.; Baker, E.S.; Kelly, R.T.; Robinson, E.W.; Smith, R.D. Mass
1212 spectrometry-based proteomics: existing capabilities and future directions. *Chem. Soc. Rev.* **2012**, *41*, 3912,
1213 doi: 10.1039/c2cs15331a.
- 1214 202. Hansen, B.T.; Davey, S.W.; Ham, A.-J.L.; Liebler, D.C. P-mod: an algorithm and software to map
1215 modifications to peptide sequences using tandem ms data. *J. Proteome Res.* **2005**, *4*, 358–368, doi:
1216 10.1021/pr0498234.
- 1217 203. Kertesz, V.; Connelly, H.M.; Erickson, B.K.; Hettich, R.L. PTMSearchPlus : software tool for automated
1218 protein identification and post-translational modification characterization by integrating accurate intact
1219 protein mass and bottom-up mass spectrometric data searches. *Anal. Chem.* **2009**, *81*, 8387–8395, doi:
1220 10.1021/ac901163c.
- 1221 204. Na, S.; Paek, E. Software eyes for protein post-translational modifications. *Mass Spectrom. Rev.* **2015**, *34*,
1222 133–147, doi: 10.1002/mas.21425.
- 1223 205. Medzihradszky, K.F.; Chalkley, R.J. Lessons in de novo peptide sequencing by tandem mass
1224 spectrometry. *Mass Spectrom. Rev.* **2015**, *34*, 43–63, doi: 10.1002/mas.21406.
- 1225 206. Jeong, K.; Kim, S.; Pevzner, P.A. UniNovo: a universal tool for de novo peptide sequencing. *Bioinformatics*
1226 **2013**, *29*, 1953–1962, doi: 10.1093/bioinformatics/btt338.
- 1227 207. Hu, J. The importance of experimental design in proteomic mass spectrometry experiments: some
1228 cautionary tales. *Briefings Funct. Genomics Proteomics* **2005**, *3*, 322–331, doi: 10.1093/bfpg/3.4.322.
- 1229 208. Schulze, W.X.; Usadel, B. Quantitation in mass-spectrometry-based proteomics. *Annu. Rev. Plant Biol.*
1230 **2010**, *61*, 491–516, doi: 10.1146/annurev-arplant-042809-112132.
- 1231 209. Oda, Y.; Huang, K.; Cross, F.R.; Cowburn, D.; Chait, B.T. Accurate quantitation of protein expression
1232 and site-specific phosphorylation. *Proc. Natl. Acad. Sci.* **1999**, *96*, 6591–6596, doi: 10.1073/pnas.96.12.6591.
- 1233 210. Nahnsen, S.; Bielow, C.; Reinert, K.; Kohlbacher, O. Tools for label-free peptide quantification. *Mol. Cell.*
1234 *Proteomics* **2013**, *12*, 549–556, doi: 10.1074/mcp.R112.025163.
- 1235 211. Mueller, L.N.; Rinner, O.; Schmidt, A.; Letarte, S.; Bodenmiller, B.; Brusniak, M.-Y.; Vitek, O.; Aebersold,
1236 R.; Müller, M. SuperHirn – a novel tool for high resolution lc-ms-based peptide/protein profiling.
1237 *Proteomics* **2007**, *7*, 3470–3480, doi: 10.1002/pmic.200700057.
- 1238 212. Cox, J.; Mann, M. MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass
1239 accuracies and proteome-wide protein quantification. *Nat. Biotechnol.* **2008**, *26*, 1367–1372, doi:
1240 10.1038/nbt.1511.
- 1241 213. Progenesis QI User Guide: Analysis workflow guidelines for DDA data

- http://www.nonlinear.com/progenesis/qi-for-proteomics/v4.1/user-guide/.
214. Bertsch, A.; Gröpl, C.; Reinert, K.; Kohlbacher, O. OpenMS and topp: open source software for lc-ms data analysis. *Data Min. Proteomics. Methods Mol. Biol. (Methods Protoc.)* **2011**, 353–367, doi: 10.1007/978-1-60761-987-1_23.
 215. <https://www.thermofisher.com>.
 216. <http://tools.proteomecenter.org/software.php>.
 217. Ishihama, Y.; Oda, Y.; Tabata, T.; Sato, T.; Nagasu, T.; Rappsilber, J.; Mann, M. Exponentially modified protein abundance index (empai) for estimation of absolute protein amount in proteomics by the number of sequenced peptides per protein. *Mol. Cell. Proteomics* **2005**, 4, 1265–1272, doi: 10.1074/mcp.M500061-MCP200.
 218. Braisted, J.C.; Kuntumalla, S.; Vogel, C.; Marcotte, E.M.; Rodrigues, A.R.; Wang, R.; Huang, S.-T.; Ferlanti, E.S.; Saeed, A.I.; Fleischmann, R.D.; Peterson, S.N.; Pieper, R. The apex quantitative proteomics tool: generating protein quantitation estimates from lc-ms/ms proteomics results. *BMC Bioinformatics* **2008**, 9, 529, doi: 10.1186/1471-2105-9-529.
 219. Sun, A.; Zhang, J.; Wang, C.; Yang, D.; Wei, H.; Zhu, Y.; Jiang, Y.; He, F. Modified spectral count index (msci) for estimation of protein abundance by protein relative identification possibility (ripro): a new proteomic technological parameter. *J. Proteome Res.* **2009**, 8, 4934–4942, doi: 10.1021/pr900252n.
 220. Silva, J.C.; Gorenstein, M. V.; Li, G.-Z.; Vissers, J.P.C.; Geromanos, S.J. Absolute quantification of proteins by lcms e. *Mol. Cell. Proteomics* **2006**, 5, 144–156, doi: 10.1074/mcp.M500230-MCP200.
 221. Clough, T.; Key, M.; Ott, I.; Ragg, S.; Schadow, G.; Vitek, O. Protein quantification in label-free lc-ms experiments. *J. Proteome Res.* **2009**, 8, 5275–5284, doi: 10.1021/pr900610q.
 222. Al Shweiki, M.R.; Mönchgesang, S.; Majovsky, P.; Thieme, D.; Trutschel, D.; Hoehenwarter, W. Assessment of Label-Free Quantification in Discovery Proteomics and Impact of Technological Factors and Natural Variability of Protein Abundance. *J. Proteome Res.* **2017**, 16(4), 1410–1424. doi: 10.1021/acs.jproteome.6b00645.
 223. Monroe, M.E.; Tolić, N.; Jaitly, N.; Shaw, J.L.; Adkins, J.N.; Smith, R.D. VIPER: an advanced software package to support high-throughput lc-ms peptide identification. *Bioinformatics* **2007**, 23, 2021–2023, doi: 10.1093/bioinformatics/btm281.
 224. Pluskal, T.; Castillo, S.; Villar-Briones, A.; Orešič, M. MZmine 2: modular framework for processing, visualizing, and analyzing mass spectrometry-based molecular profile data. *BMC Bioinformatics* **2010**, 11, 395, doi: 10.1186/1471-2105-11-395.
 225. Bairoch, A.; Apweiler, R. The swiss-prot protein sequence data bank and its supplement trembl in 1999. *Nucleic Acids Res.* **1999**, 27, 49–54, doi: 10.1093/nar/27.1.49.
 226. Simpson, J.C.; Pepperkok, R. Localizing the proteome. *Genome Biol.* **2003**, 4, 240, doi: 10.1186/gb-2003-4-12-240.
 227. Huh, W.-K.; Falvo, J. V.; Gerke, L.C.; Carroll, A.S.; Howson, R.W.; Weissman, J.S.; O'Shea, E.K. Global analysis of protein localization in budding yeast. *Nature* **2003**, 425, 686–691, doi: 10.1038/nature02026.
 228. Yu, C.-S.; Chen, Y.-C.; Lu, C.-H.; Hwang, J.-K. Prediction of protein subcellular localization. *Proteins Struct. Funct. Bioinforma.* **2006**, 64, 643–651, doi: 10.1002/prot.21018.
 229. Horton, P.; Park, K.-J.; Obayashi, T.; Fujita, N.; Harada, H.; Adams-Collier, C.J.; Nakai, K. WoLF psort: protein localization predictor. *Nucleic Acids Res.* **2007**, 35, W585–W587, doi: 10.1093/nar/gkm259.
 230. Briesemeister, S.; Rahnenfiß, J.; Kohlbacher, O. YLoc—an interpretable web server for predicting subcellular localization. *Nucleic Acids Res.* **2010**, 38, W497–W502, doi: 10.1093/nar/gkq477.
 231. Savojardo, C.; Martelli, P.L.; Fariselli, P.; Profiti, G.; Casadio, R. BUSCA: an integrative web server to predict subcellular localization of proteins. *Nucleic Acids Res.* **2018**, 46, W459–W466, doi: 10.1093/nar/gky320.
 232. Sperschneider, J.; Catanzariti, A.-M.; DeBoer, K.; Petre, B.; Gardiner, D.M.; Singh, K.B.; Dodds, P.N.; Taylor, J.M. LOCALIZER: subcellular localization prediction of both plant and effector proteins in the plant cell. *Sci. Rep.* **2017**, 7, 44598, doi: 10.1038/srep44598.
 233. Chou, K.-C.; Shen, H.-B. Plant-mploc: a top-down strategy to augment the power for predicting plant protein subcellular localization. *PLoS One* **2010**, 5, e11335, doi: 10.1371/journal.pone.0011335.
 234. Käll, L.; Krogh, A.; Sonnhammer, E.L.. A combined transmembrane topology and signal peptide prediction method. *J. Mol. Biol.* **2004**, 338, 1027–1036, doi: 10.1016/j.jmb.2004.03.016.
 235. Bateman, A.; Martin, M.J.; O'Donovan, C.; Magrane, M.; Alpi, E.; Antunes, R.; Bely, B.; Bingley, M.; Bonilla, C.; Britto, R.; et al. UniProt: the universal protein knowledgebase. *Nucleic Acids Res.* **2017**, 45, D158–D169, doi: 10.1093/nar/gkw1099.
 236. Kanehisa, M.; Goto, S.; Sato, Y.; Kawashima, M.; Furumichi, M.; Tanabe, M. Data, information, knowledge and principle: back to metabolism in kegg. *Nucleic Acids Res.* **2014**, 42, D199–D205, doi:

- 1300 10.1093/nar/gkt1076.
- 1301 237. Ogata, H.; Goto, S.; Sato, K.; Fujibuchi, W.; Bono, H.; Kanehisa, M. KEGG: kyoto encyclopedia of genes
1302 and genomes. *Nucleic Acids Res.* **1999**, *27*, 29–34, doi: 10.1093/nar/27.1.29.
- 1303 238. Klie, S.; Nikoloski, Z. The choice between mapman and gene ontology for automated gene function
1304 prediction in plant science. *Front. Genet.* **2012**, 3doi: 10.3389/fgene.2012.00115.
- 1305 239. USADEL, B.; POREE, F.; NAGEL, A.; LOHSE, M.; CZEDIK-EYSENBERG, A.; STITT, M. A guide to using
1306 mapman to visualize and compare omics data in plants: a case study in the crop species, maize. *Plant.*
1307 *Cell Environ.* **2009**, *32*, 1211–1229, doi: 10.1111/j.1365-3040.2009.01978.x.
- 1308 240. Kanehisa, M.; Sato, Y.; Kawashima, M.; Furumichi, M.; Tanabe, M. KEGG as a reference resource for
1309 gene and protein annotation. *Nucleic Acids Res.* **2016**, *44*, D457–D462, doi: 10.1093/nar/gkv1070.
- 1310 241. LOHSE, M.; NAGEL, A.; HERTER, T.; MAY, P.; SCHRODA, M.; ZRENNER, R.; TOHGE, T.; FERNIE,
1311 A.R.; STITT, M.; USADEL, B. Mercator: a fast and simple web server for genome scale functional
1312 annotation of plant sequence data. *Plant. Cell Environ.* **2014**, *37*, 1250–1258, doi: 10.1111/pce.12231.
- 1313 242. Saha, I.; Klingström, T.; Forsberg, S.; Wikander, J.; Zubek, J.; Kierczak, M.; Plewczynski, D. Evaluation
1314 of machine learning algorithms on protein-protein interactions. *Adv. Intell. Syst. Comput.* **2014**, 211–218,
1315 doi: 10.1007/978-3-319-02309-0_22.
- 1316 243. Hayashi, T.; Matsuzaki, Y.; Yanagisawa, K.; Ohue, M.; Akiyama, Y. MEGADOCK-web: an integrated
1317 database of high-throughput structure-based protein-protein interaction predictions. *BMC Bioinformatics*
1318 **2018**, *19*, 62, doi: 10.1186/s12859-018-2073-x.
- 1319 244. Liu, S.; Liu, C.; Deng, L. Machine learning approaches for protein-protein interaction hot spot
1320 prediction: progress and comparative assessment. *Molecules* **2018**, *23*, 2535, doi:
1321 10.3390/molecules23102535.
- 1322 245. Landry, E.J.; Fuchs, S.J.; Hu, J. Carbohydrate composition of mature and immature faba bean seeds. *J.*
1323 *Food Compos. Anal.* **2016**, *50*, 55–60, doi: 10.1016/j.jfca.2016.05.010.
- 1324 246. Oracz, K.; Bouteau, H.E.M.; Farrant, J.M.; Cooper, K.; Belghazi, M.; Job, C.; Job, D.; Corbineau, F.; Bailly,
1325 C.; El-Maarouf Bouteau, H.; Farrant, J.M.; Cooper, K.; Belghazi, M.; Job, C.; Job, D.; Corbineau, F.; Bailly,
1326 C. ROS production and protein oxidation as a novel mechanism for seed dormancy alleviation. *Plant J.*
1327 **2007**, *50*, 452–465, doi: 10.1111/j.1365-313X.2007.03063.x.
- 1328 247. Leymarie, J.; Vitkauskaitė, G.; Hoang, H.H.; Gendreau, E.; Chazoule, V.; Meimoun, P.; Corbineau, F.; El-
1329 Maarouf-Bouteau, H.; Bailly, C. Role of reactive oxygen species in the regulation of arabidopsis seed
1330 dormancy. *Plant Cell Physiol.* **2012**, *53*, 96–106, doi: 10.1093/pcp/pcr129.
- 1331 248. Bailly, C. Active oxygen species and antioxidants in seed biology. *Seed Sci. Res.* **2004**, *14*, 93–107, doi:
1332 10.1079/SSR2004159.
- 1333 249. Wolff, I. A. Seed lipids. *Science* **1966**, *154*, 1140–1149, doi: 10.1126/science.154.3753.1140.
- 1334 250. Ponquett, R.T.; Ross, G. Lipid autoxidation and seed ageing: putative relationships between seed
1335 longevity and lipid stability. *Seed Sci. Res.* **1992**, *2*, 51–54, doi: 10.1017/S0960258500001100.
- 1336 251. Ouzouline, M.; Tahani, N.; Demandre, C.; El Amrani, E.; Benhassaine-Kesri, G.; Serghini Caid, H. Effects
1337 of accelerated aging upon the lipid composition of seeds from two soft wheat varieties from morocco.
1338 *Grasas y Aceites* **2009**, *60*, 367–374, doi: 10.3989/gya.010409.
- 1339 252. Murthy, U.M.N.N.; Kumar, P.P.; Sun, W.Q. Mechanisms of seed ageing under different storage
1340 conditions for vigna radiata (l.) wilczek: lipid peroxidation, sugar hydrolysis, maillard reactions and
1341 their relationship to glass state transition. *J. Exp. Bot.* **2003**, *54*, 1057–1067, doi: 10.1093/jxb/erg092.
- 1342 253. Hoekstra, F. a; Golovina, E. a; Tetteroo, F. a; Wolkers, W.F. Induction of desiccation tolerance in plant
1343 somatic embryos: how exclusive is the protective role of sugars? *Cryobiology* **2001**, *43*, 140–150, doi:
1344 10.1006/cryo.2001.2358.
- 1345 254. Wolff, S.P.; Dean, R.T. Glucose autoxidation and protein modification. the potential role of ‘autoxidative
1346 glycosylation’ in diabetes. *Biochem. J.* **1987**, *245*, 243–250, doi: 10.1042/bj2450243.
- 1347 255. Ayala, A.; Muñoz, M.F.; Argüelles, S. Lipid peroxidation: production, metabolism, and signaling
1348 mechanisms of malondialdehyde and 4-hydroxy-2-nonenal. *Oxid. Med. Cell. Longev.* **2014**, *2014*, 1–31,
1349 doi: 10.1155/2014/360438.
- 1350 256. Vistoli, G.; De Maddis, D.; Cipak, A.; Zarkovic, N.; Carini, M.; Aldini, G. Advanced glycoxidation and
1351 lipoxidation end products (ages and ales): an overview of their mechanisms of formation. *Free Radic. Res.*
1352 **2013**, *47*, 3–27, doi: 10.3109/10715762.2013.815348.
- 1353 257. Strelec, I.; Hlevnjak, M. Accumulation of amadori and maillard products in wheat seeds aged under
1354 different storage conditions. *Croatica Chemica Acta*, 2008, *81*, 131–137.
- 1355 258. Wettlaufer, S.H.; Leopold, A.C. Relevance of amadori and maillard products to seed deterioration. *Plant*
1356 *Physiol.* **1991**, *97*, 165–169, doi: 10.1104/pp.97.1.165.
- 1357 259. Colville, L.; Bradley, E.L.; Lloyd, A.S.; Pritchard, H.W.; Castle, L.; Kranner, I. Volatile fingerprints of

- seeds of four species indicate the involvement of alcoholic fermentation, lipid peroxidation, and maillard reactions in seed deterioration during ageing and desiccation stress. *J. Exp. Bot.* **2012**, *63*, 6519–6530, doi: 10.1093/jxb/ers307.
260. Greifenhagen, U.; Nguyen, V.D.; Moschner, J.; Giannis, A.; Frolov, A.; Hoffmann, R. Sensitive and site-specific identification of carboxymethylated and carboxyethylated peptides in tryptic digests of proteins and human plasma. *J. Proteome Res.* **2015**, *14*, 768–777, doi: 10.1021/pr500799m.
 261. Schmidt, R.; Böhme, D.; Singer, D.; Frolov, A. Specific tandem mass spectrometric detection of age-modified arginine residues in peptides. *J. Mass Spectrom.* **2015**, *50*, 613–624, doi: 10.1002/jms.3569.
 262. Frolov, A.; Hoffmann, P.; Hoffmann, R. Fragmentation behavior of glycated peptides derived from d-glucose, d-fructose and d-ribose in tandem mass spectrometry. *J. Mass Spectrom.* **2006**, *41*, 1459–1469, doi: 10.1002/jms.1117.
 263. Zhang, Q.; Frolov, A.; Tang, N.; Hoffmann, R.; van de Goor, T.; Metz, T.O.; Smith, R.D. Application of electron transfer dissociation mass spectrometry in analyses of non-enzymatically glycated peptides. *Rapid Commun. Mass Spectrom.* **2007**, *21*, 661–666, doi: 10.1002/rcm.2884.
 264. Fedorova, M.; Frolov, A.; Hoffmann, R. Fragmentation behavior of amadori-peptides obtained by non-enzymatic glycosylation of lysine residues with adp-ribose in tandem mass spectrometry. *J. Mass Spectrom.* **2010**, n/a-n/a, doi: 10.1002/jms.1758.
 265. Ehrlich, H.; Hanke, T.; Frolov, A.; Langrock, T.; Hoffmann, R.; Fischer, C.; Schwarzenbolz, U.; Henle, T.; Born, R.; Worch, H. Modification of collagen in vitro with respect to formation of n ϵ -carboxymethyllysine. *Int. J. Biol. Macromol.* **2009**, *44*, 51–56, doi: 10.1016/j.ijbiomac.2008.10.001.
 266. Frolov, A.; Singer, D.; Hoffmann, R. Site-specific synthesis of amadori-modified peptides on solid phase. *J. Pept. Sci.* **2006**, *12*, 389–395, doi: 10.1002/psc.739.
 267. Frolov, A.; Hoffmann, R. Separation of amadori peptides from their unmodified analogs by ion-pairing rp-hplc with heptafluorobutyric acid as ion-pair reagent. *Anal. Bioanal. Chem.* **2008**, *392*, 1209–1214, doi: 10.1007/s00216-008-2377-1.
 268. Frolov, A.; Singer, D.; Hoffmann, R. Solid-phase synthesis of glucose-derived amadori peptides. *J. Pept. Sci.* **2007**, *13*, 862–867, doi: 10.1002/psc.901.
 269. Greifenhagen, U.; Frolov, A.; Hoffmann, R. Oxidative degradation of n ϵ -fructosylamine-substituted peptides in heated aqueous systems. *Amino Acids* **2015**, *47*, 1065–1076, doi: 10.1007/s00726-015-1940-2.
 270. Dammann, P.; Sell, D.R.; Begall, S.; Strauch, C.; Monnier, V.M. Advanced glycation end-products as markers of aging and longevity in the long-lived ansell's mole-rat (*fukomys anselli*). *Journals Gerontol. Ser. A* **2012**, *67A*, 573–583, doi: 10.1093/gerona/glr208.
 271. Bilova, T.; Paudel, G.; Shilyaev, N.; Schmidt, R.; Brauch, D.; Tarakhovskaya, E.; Milrud, S.; Smolikova, G.; Tissier, A.; Vogt, T.; Sinz, A.; Brandt, W.; Birkemeyer, C.; Wessjohann, L.A.; Frolov, A. Global proteomic analysis of advanced glycation end products in the arabidopsis proteome provides evidence for age-related glycation hot spots. *J. Biol. Chem.* **2017**, *292*, 15758–15776, doi: 10.1074/jbc.M117.794537.
 272. Yin, X.; He, D.; Gupta, R.; Yang, P. Physiological and proteomic analyses on artificially aged brassica napus seed. *Front. Plant Sci.* **2015**, *6*, 112, doi: 10.3389/fpls.2015.00112.
 273. Kołodziejczyk, I.; Dzitko, K.; Szewczyk, R.; Posmyk, M.M. Exogenous melatonin improves corn (*zea mays* l.) embryo proteome in seeds subjected to chilling stress. *J. Plant Physiol.* **2016**, *193*, 47–56, doi: 10.1016/j.jplph.2016.01.012.
 274. DeBlasio, S.L.; Johnson, R.S.; MacCoss, M.J.; Gray, S.M.; Cilia, M. Model system-guided protein interaction mapping for virus isolated from phloem tissue. *J. Proteome Res.* **2016**, *15*, 4601–4611, doi: 10.1021/acs.jproteome.6b00715.
 275. Tyanova, S.; Temu, T.; Carlson, A.; Sinitcyn, P.; Mann, M.; Cox, J. Visualization of lc-ms/ms proteomics data in maxquant. *Proteomics* **2015**, *15*, 1453–1456, doi: 10.1002/pmic.201400449.
 276. Cox, B.; Kislinger, T.; Wigle, D.A.; Kannan, A.; Brown, K.; Okubo, T.; Hogan, B.; Jurisica, I.; Frey, B.; Rossant, J.; Emili, A. Integrated proteomic and transcriptomic profiling of mouse lung development and nmyc target genes. *Mol. Syst. Biol.* **2007**, *3*, doi: 10.1038/msb4100151.
 277. Röst, H.L.; Sachsenberg, T.; Aiche, S.; Bielow, C.; Weissner, H.; Aicheler, F.; Andreotti, S.; Ehrlich, H.-C.; Gutenbrunner, P.; Kenar, E.; et al. OpenMS: a flexible open-source software platform for mass spectrometry data analysis. *Nat. Methods* **2016**, *13*, 741–748, doi: 10.1038/nmeth.3959.
 278. *Proteome Discoverer. User Guide. Software Version 2.2*; Thermo Fisher Scientific Inc, 2017.
 279. Park, S.K.; Yates, J.R. Census for proteome quantification. *Curr. Protoc. Bioinforma.* **2010**, *29*, 13.12.1–13.12.11, doi: 10.1002/0471250953.bi1312s29.
 280. Chang, C.; Zhang, J.; Han, M.; Ma, J.; Zhang, W.; Wu, S.; Liu, K.; Xie, H.; He, F.; Zhu, Y. SILVER: an efficient tool for stable isotope labeling lc-ms data quantitative analysis with quality control methods. *Bioinformatics* **2014**, *30*, 586–587, doi: 10.1093/bioinformatics/btt726.

- 1416 281. Bellew, M.; Coram, M.; Fitzgibbon, M.; Igra, M.; Randolph, T.; Wang, P.; May, D.; Eng, J.; Fang, R.; Lin,
1417 C.; Chen, J.; Goodlett, D.; Whiteaker, J.; Paulovich, A.; McIntosh, M. A suite of algorithms for the
1418 comprehensive analysis of complex protein mixtures using high-resolution lc-ms. *Bioinformatics* **2006**,
1419 *22*, 1902–1909, doi: 10.1093/bioinformatics/btl276.
- 1420 282. Valot, B.; Langella, O.; Nano, E.; Zivy, M. MassChroQ: a versatile tool for mass spectrometry
1421 quantification. *Proteomics* **2011**, *11*, 3572–3577, doi: 10.1002/pmic.201100120.
- 1422 283. MacLean, B.; Tomazela, D.M.; Shulman, N.; Chambers, M.; Finney, G.L.; Frewen, B.; Kern, R.; Tabb, D.L.;
1423 Liebler, D.C.; MacCoss, M.J. Skyline: an open source document editor for creating and analyzing
1424 targeted proteomics experiments. *Bioinformatics* **2010**, *26*, 966–968, doi: 10.1093/bioinformatics/btq054.
- 1425 284. Tsou, C.-C.; Avtonomov, D.; Larsen, B.; Tucholska, M.; Choi, H.; Gingras, A.-C.; Nesvizhskii, A.I. DIA-
1426 umpire: comprehensive computational framework for data-independent acquisition proteomics. *Nat.*
1427 *Methods* **2015**, *12*, 258–264, doi: 10.1038/nmeth.3255.
- 1428 285. Deutsch, E.W.; Mendoza, L.; Shteynberg, D.; Farrah, T.; Lam, H.; Tasman, N.; Sun, Z.; Nilsson, E.; Pratt,
1429 B.; Prazan, B.; Eng, J.K.; Martin, D.B.; Nesvizhskii, A.I.; Aebersold, R. A guided tour of the trans-
1430 proteomic pipeline. *Proteomics* **2010**, *10*, 1150–1159, doi: 10.1002/pmic.200900375.
- 1431 286. Argentini, A.; Goeminne, L.J.E.; Verheggen, K.; Hulstaert, N.; Staes, A.; Clement, L.; Martens, L. MoFF:
1432 a robust and automated approach to extract peptide ion intensities. *Nat. Methods* **2016**, *13*, 964–966, doi:
1433 10.1038/nmeth.4075.
- 1434 287. <http://www.matrixscience.com/distiller.htm>.
- 1435 288. Brusniak, M.-Y.; Bodenmiller, B.; Campbell, D.; Cooke, K.; Eddes, J.; Garbutt, A.; Lau, H.; Letarte, S.;
1436 Mueller, L.N.; Sharma, V.; Vitek, O.; Zhang, N.; Aebersold, R.; Watts, J.D. Corra: computational
1437 framework and tools for lc-ms discovery and targeted mass spectrometry-based proteomics. *BMC*
1438 *Bioinformatics* **2008**, *9*, 542, doi: 10.1186/1471-2105-9-542.
- 1439 289. Millikin, R.J.; Solntsev, S.K.; Shortreed, M.R.; Smith, L.M. Ultrafast peptide label-free quantification with
1440 flashlfq. *J. Proteome Res.* **2018**, *17*, 386–391, doi: 10.1021/acs.jproteome.7b00608.