

# Mining seed proteome: from protein dynamics to modification profiles

**Andrej Frolov<sup>1,2</sup>, Tatiana Mamontova<sup>1,2</sup>, Christian Ihling<sup>3</sup>, Elena Lukasheva<sup>2</sup>, Mikhail Bankin<sup>4</sup>, Veronika Chantseva<sup>4</sup>, Maria Vikhnina<sup>1,2</sup>, Alena Soboleva<sup>1,2,5</sup>, Julia Shumilina<sup>2</sup>, Gregory Mavropolo-Stolyarenko<sup>2</sup>, Tatiana Grischina<sup>2</sup>, Natalia Osmolovskaya<sup>4</sup>, Vladimir Zhukov<sup>6</sup>, Wolfgang Hoehenwarter<sup>7</sup>, Andrea Sinz<sup>3</sup>, Igor Tikhonovich<sup>6,8</sup>, Ludger A. Wessjohann<sup>1</sup>, Tatiana Bilova<sup>1,4</sup>, Galina Smolikova<sup>4</sup>, and Sergei Medvedev<sup>4</sup>**

<sup>1</sup>Department of Bioorganic Chemistry, Weinberg 3, 06120 Halle/Saale DE, Leibniz Institute of Plant Biochemistry;

<sup>2</sup>Department of Biochemistry, Faculty of Biology, Saint Petersburg State University, Srednii prosp., 41–43, St. Petersburg, 199004, Russian Federation;

<sup>3</sup>Department of Pharmaceutical Chemistry and Bioanalytics, Institute of Pharmacy, Wolfgang-Langenbeck-Strasse 4, 06120 Halle/Saale DE, Martin-Luther Universität Halle-Wittenberg;

<sup>4</sup>Department of Plant Physiology and Biochemistry, Faculty of Biology, Saint Petersburg State University, Universitetskaya nab., 7–9, Saint Petersburg, 199034, Russian Federation;

<sup>5</sup>Department of Cell and Metabolic Biology, Weinberg 3, 06120 Halle/Saale DE, Leibniz Institute of Plant Biochemistry;

<sup>6</sup>All-Russia Research Institute for Agricultural Microbiology, Podbelsky chaussee 3, Pushkin, St. Petersburg, 196608, Russian Federation;

<sup>7</sup>Proteome Analytics Research Group, Weinberg 3, 06120 Halle/Saale DE, Leibniz Institute of Plant Biochemistry;

<sup>8</sup>Department of Genetics and Biotechnology, Faculty of Biology, Saint Petersburg State University, Universitetskaya nab. 7–9, St. Petersburg, 199034, Russian Federation

Address correspondence and requests for materials to Andrej Frolov, Andrej.Frolov@ipb-halle.de

**Citation:** Frolov, A., Mamontova, T., Ihling, C., Lukasheva, E. et al. 2018. Mining seed proteome: from protein dynamics to modification profiles. *Bio. Comm.* 63(1): 43–58. <https://doi.org/10.21638/spbu03.2018.106>

**Author's information:** Andrej Frolov, [orcid.org/0000-0002-7593-7717](https://orcid.org/0000-0002-7593-7717); Tatiana Mamontova, [orcid.org/0000-0002-9794-8089](https://orcid.org/0000-0002-9794-8089); Christian Ihling, Scopus ID: 55444090000; Elena Lukasheva, [orcid.org/0000-0002-0889-2395](https://orcid.org/0000-0002-0889-2395); Mikhail Bankin, [orcid.org/0000-0002-6240-2246](https://orcid.org/0000-0002-6240-2246); Veronika Chantseva, [orcid.org/0000-0002-3486-4232](https://orcid.org/0000-0002-3486-4232); Maria Vikhnina, [orcid.org/0000-0003-0120-3435](https://orcid.org/0000-0003-0120-3435); Alena Soboleva, [orcid.org/0000-0002-7318-8628](https://orcid.org/0000-0002-7318-8628); Julia Shumilina, [orcid.org/0000-0002-9747-5779](https://orcid.org/0000-0002-9747-5779); Gregory Mavropolo-Stolyarenko, [orcid.org/0000-0001-8712-0272](https://orcid.org/0000-0001-8712-0272); Tatiana Grischina [orcid.org/0000-0002-5306-2980](https://orcid.org/0000-0002-5306-2980); Ekaterina Romanovskaya, [orcid.org/0000-0001-8912-5886](https://orcid.org/0000-0001-8912-5886); Natalia Osmolovskaya, [orcid.org/0000-0001-8764-8552](https://orcid.org/0000-0001-8764-8552); Vladimir Zhukov, [orcid.org/0000-0002-2411-9191](https://orcid.org/0000-0002-2411-9191); Wolfgang Hoehenwarter, [orcid.org/0000-0002-7669-7524](https://orcid.org/0000-0002-7669-7524); Andrea Sinz, [orcid.org/0000-0003-1521-4899](https://orcid.org/0000-0003-1521-4899), Scopus ID: 6701513466; Igor Tikhonovich, [orcid.org/0000-0001-8968-854X](https://orcid.org/0000-0001-8968-854X); Ludger A. Wessjohann, [orcid.org/0000-0003-2060-8235](https://orcid.org/0000-0003-2060-8235), Scopus ID: 16164439600; Tatiana Bilova, [orcid.org/0000-0002-6024-3667](https://orcid.org/0000-0002-6024-3667); Galina Smolikova, [orcid.org/0000-0001-5238-1851](https://orcid.org/0000-0001-5238-1851); Sergei Medvedev, [orcid.org/0000-0003-1127-1343](https://orcid.org/0000-0003-1127-1343)

**Manuscript Editor:** Farida Minibayeva, Kazan Institute of Biochemistry and Biophysics, Kazan, Russia;

**Guest Editor:** Maria Shishova, Saint Petersburg State University, Saint Petersburg, Russia;

**Received:** January 30, 2018;

**Revised:** April 24, 2018;

**Accepted:** May 04, 2018;

**Copyright:** © 2018 Frolov et al. This is an open-access article distributed under the terms of the License Agreement with Saint Petersburg State University, which permits to the authors an unrestricted distribution and self-archiving free of charge.

**Funding:** This work was supported by the Russian Science Foundation (Grant No. 17-16-01042).

**Competing interests:** The authors have declared that no competing interests exist.

## Abstract

In the modern world, crop plants represent a major source of daily consumed foods. Among them, cereals and legumes — i.e. the crops accumulating oils, carbohydrates and proteins in their seeds — dominate in European agriculture, tremendously impacting global protein consumption and biodiesel production. Therefore, the seeds of crop plants attract the special attention of biologists, biochemists, nutritional physiologists and food chemists. Seed development and germination, as well as age- and stress-related changes in their viability and nutritional properties, can be addressed by a variety of physiological and biochemical methods. In this context, the methods of functional genomics can be applied to address characteristic changes in seed metabolism, which can give access to stress-resistant genotypes. Among these methods, proteomics is one of the most effective tools, allowing mining metabolism changes on the protein level. Here we discuss the main methodological approaches of seed proteomics in the context of physiological changes related to environmental stress and ageing. We provide a comprehensive comparison of gel- and chromatography-based approaches with a special emphasis on advantages and disadvantages of both strategies in characterization of the seed proteome.

**Keywords:** food safety, LC-MS, post-translational modifications, proteomics, seed quality.

## Introduction

Due to their relatively low costs, crop plants are the basis of the human diet in most countries and represent a major source of daily consumed foods (Altshul, 1962; Miller, Herman, Jahn, and Bradford, 2010; Bradford et al., 2018). Among them, cereals and legumes are characterized by the highest content of energetically valuable reserved molecules — proteins, polysaccharides and lipids (mostly represented by triacylglycerols — i.e. oils) (Duranti, 2006; Bourgeois et al., 2009; Yoshida et al., 2010; Lafiandra et al., 2012; Guerrier and Gavaletto, 2018) — and are, therefore, most intensively cultured worldwide (Goff and Salmeron, 2004; Cernay, Pelzer, and Makovski, 2016). During the last decades, soybeans (*Glycine max*), common (dry) beans (*Phaseolus vulgaris*) and peanuts (*Arachis hypogaea*) have become major sources of food protein (Messina, 1999; Asgar et al., 2010). Indeed, soybeans were cultivated on an area of approximately 100 million ha with crop yields of 238 million tons in the year 2010; soybeans account for the most pronounced portion of world legume yields (Nedumaran et al., 2015). The area share of this crop in the developed world (Europe, the United States and Canada) is much higher than in developing countries (Nedumaran et al., 2015). On the other hand, according to the Food and Agriculture Organization of the United Nations, cereals such as maize (*Zea mays*), wheat (mostly *Triticum aestivum*), rice (*Oryza sativa*) and barley (*Hordeum vulgare*) are the most prominent sources of food carbohydrates (FAO, 2018). World cereal production in 2017 was estimated at 2627 million tons, with maize, wheat and rice accounting for about 1070, 750 and 700 million tons, respectively (FAO, 2017). Finally, oilseed rape (*Brassica napus*), soybean and sunflower (*Helianthus annuus*) are the most important sources of food oils (Carre and Pouzet, 2012; Bates, Stymne, and Ohlrogge, 2013; Liu et al., 2014). Moreover, rapeseed oil is one of the most important sources of biodiesel (I. C. E. Committees, 2015).

Obviously, sustaining seed quality is of principle importance for effective agriculture and production of safe foods (Miller, Herman, Jahn, and Bradford, 2010). In this context, the mechanisms underlying seed viability and ensuring high quality of food protein need to be well understood. It is important to note that seed quality can be affected by prolonged storage (Sano et al., 2016), especially under high humidity and/or temperature conditions (Murthy, Kumar, and Sun, 2003; Walters, Ballasteros, and Vertucci, 2010; Sum, Sreenivasan, Singh, and Radhamani, 2013). Not less important are environmental conditions of seed development. Environmental stressors applied during seed maturation might compromise seed viability and shelf life (Ghassemi-Golezani and Mazloomi-Oskooyi, 2008; Pereira Lima et al., 2017).

For example, even a transient drought can suppress degradation of embryonic chlorophylls and thereby dramatically decrease the quality of formed seeds (Smolikova and Medvedev, 2016; Smolikova et al., 2017a; Smolikova et al., 2017b). For these reasons, the seeds of crop plants attract the particular attention of biologists, biochemists, nutritional physiologists and food chemists, addressing these two principle questions.

In general, seed quality (in the widest sense of these words) can be addressed by a rich pattern of physiological methods, including assessment of membrane integrity, germination tests, monitoring of morphological alterations and vigor (Hampton and TeKrony, 1995; McDonald, 1998; Don, 2009; Corbineau, 2012). As seed vigor is dependent on the environmental conditions of seed development and storage, knowledge about the mechanisms of seed ageing gives access to new methodological approaches to seed vigor characterization (Hampton and TeKrony, 1995; Smolikova, 2014; Filho, 2015). On the biochemical level, these experiments can be complemented by oxidative stress assays (Bailly and Kraner, 2011; Kocsy, 2015). However, today, methods of functional genomics and systems biology provide an excellent opportunity to address seed quality on a principally new level, assuming an integrated systematic and physiology-based approach (Rajjou et al., 2008; Fu, Ahmed, and Diederichsen, 2015; Nagel et al., 2015; Smolikova et al., 2016). Indeed, integration of transcriptomic, proteomic and metabolomic methodological platforms gives access to the whole network of metabolic events underlying gene expression (May et al., 2011; Mochida and Shinozaki, 2011; Brink-Jensen, Bak, Jorgensen, and Ekstrom, 2013). Due to the high specificity and sensitivity of modern analytical instrumentation, these so-called omics techniques allow scientists to consider plant metabolism on the level of organs (Lombardo et al., 2011; Palma, Corpas, and del Rio, 2011), tissues (Gupta et al., 2015), or even individual cell groups (Barkla and Vera-Estrella, 2015). Moreover, this approach gives insight in a time dimension — i.e. it allows seed germination and development to be addressed (Wang et al., 2016; Gupta, Bhaskar, Sriram, and Wang, 2017).

As proteins represent an important seed component critical for germination and seedling development, during the past decade proteomics has become an important tool of seed research. Importantly, proteomics makes it possible to address physiological changes by characterization of alteration in protein abundances (expression) simultaneously with assessment of post-translational modification patterns, which might be informative in the sense of antinutritive alterations in seeds. Here we present a short overview of the techniques currently established in this field and potentially applicable to assessment of seed quality, both in terms of viability and potential food safety aspects. We also consider the

methods of proteomics required for assessment of protein dynamics and characterization of PTM patterns.

## Methodological approaches in seed proteomics

The proteome is a complex system, representing the result of interconnected dynamic properties of individual proteins (Larance and Lamond, 2015). Accordingly, characterization of the proteome includes annotation of individual proteins, determination of their abundances, expression of isoforms, synthesis and degradation rates, sub-cellular localization, interactions and post-translational modifications. Although different analytical techniques can be applied to address these questions (e.g., immunochemical methods such as ELISA or western blotting) (Jimenez-Lopez et al., 2018), in the absolute majority of cases, proteome analysis relies on mass spectrometry (MS) and bioinformatics (Heyer et al., 2017; Takac, Samajova, and Samaj, 2017). Quasi-molecular ions (those less than 25kDa) of proteins can be either fragmented directly — the so-called top-down approach (Chmelik et al., 2009; Catherman, Skinner, and Kelleher, 2014) — or the macromolecules can be first subjected to limited proteolysis prior to MS analysis — the bottom-up strategy (Gillet, Leither, and Aebersold, 2016). In the latter case, annotation of individual proteins relies on sequence tags — i.e. information on exact monoisotopic mass, charge and fragment ions — obtained in tandem mass spectrometric (MS/MS) experiments with proteolytic peptides (Mortz et al., 1996). Because of the intrinsic limitations of the top-down strategy — relatively small molecular weight and a strong requirement for high sample purity (Soboleva et al., 2017) — it is only rarely applied in seed proteomics (Hummel, Wigger, and Brockmeyer, 2015). In contrast, bottom-up methods find application in all aspects of seed biology (Rathi et al., 2016).

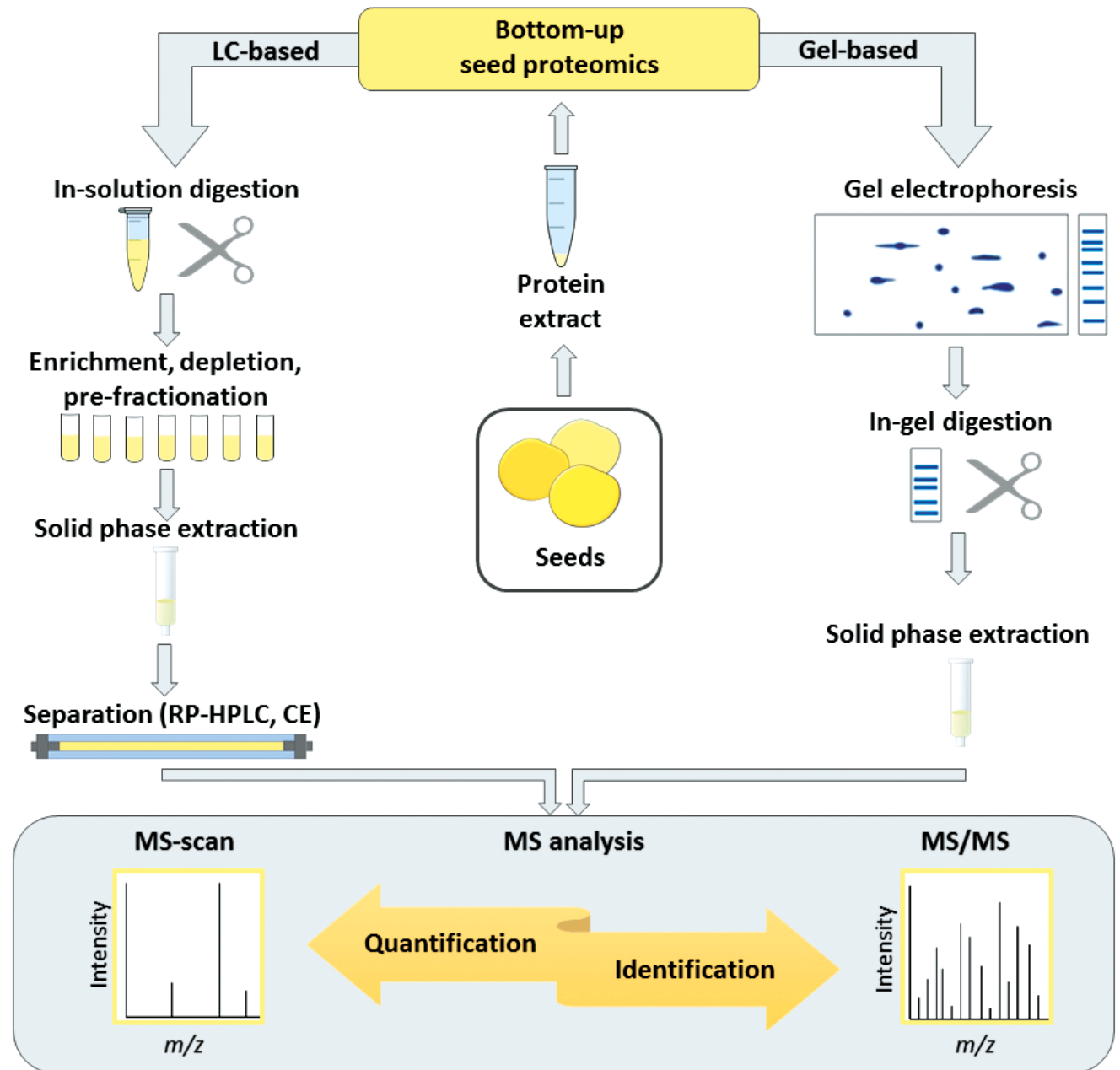
Due to the high complexity of living organisms in general and the plant proteome in particular, mass-spectrometric methods employed in its characterization need to be complemented by adequate separation techniques. Generally, based on the nature of these techniques, all bottom-up proteomic methods can be attributed to either gel-based or gel-free strategy (Tan, Lim, and Lau, 2017). The latter group of methods is also often referred to as liquid chromatography (LC)-based proteomics, although it also can rely on non-chromatographic techniques such as capillary electrophoresis (Chen, Cociorva, Norris, and Yates, 2017). Besides the applied separation techniques, gel-based and gel-free approaches differ in their workflow (Fig. 1): in the former, separation is done on the protein level (i.e. prior to enzymatic digestion), whereas in the latter, peptides are

subjected to separation after enzymatic digestion (Tan, Lim, and Lau, 2017).

## Gel-based bottom-up proteomic strategy

The gel-based strategy (Fig. 1) typically relies on two-dimensional gel electrophoresis (2D-GE), comprising two steps of protein separation: by isoelectric points and by molecular weight (SDS-PAGE). For the first step, isoelectrofocusing, proteins are solubilized in aqueous buffers containing chaotropic agents (7–8 mol/L urea and 2 mol/L thiourea) and detergents (typically 3-((3-cholamidopropyl)dimethylammonio)-1-propanesulfonate, CHAPS) to disturb polar and non-polar interactions in the protein molecule, respectively (Rabilloud, Chevallet, Luche, and Lelong, 2010). The samples can be supplemented with carrier-ampholytes if no immobilized pH gradients are used (mostly not the case today) (Rabilloud and Lelong, 2011). Generally, most of the proteins can be reconstituted in this buffer, although CHAPS is not the best detergent in this case (Rabilloud and Lelong, 2011). Protein isolation can rely on different methods, providing extraction of total protein, or only its water-soluble fraction. For protein extraction, plant tissues are typically shock-frozen in liquid nitrogen and ground sequentially with a mortar and ball mill (Paudel et al., 2016). In the easiest and most straightforward way, ground plant material is extracted directly with the sample buffer, containing urea and thiourea in presence or absence of CHAPS or optionally Triton X-100 (Gallardo, Kurt, Thompson, and Ochatt, 2003; Han, Yin, He, and Yang, 2013; Gallardo, Kurt, Thompson, and Ochatt, 2018). Thereby, to reduce interference with low molecular weight metabolites, pre-extraction with petroleum ether can be employed (Murad and Rech, 2012). In the most comprehensive way, the total protein fraction can be isolated by phenol or acetone extraction (Isaacson et al., 2006). Among these two approaches, phenol extraction provides lower protein yields but higher purities — i.e. it is advantageous when sufficient sample amounts are available. Alternatively, the soluble protein fraction can be extracted by aqueous buffers (typically Tris-HCl or HEPES), further precipitated by acetone and dried (Han, Yin, He, and Yang, 2013).

In the easiest way, individual protein-containing electrophoretic zones can be visualized with colloidal Coomassie Blue dye (Neuhoff, Arold, Taube, and Ehrhardt, 1988) or by formaldehyde-free silver staining (Chevallet et al., 2008). These methods, however, are not ideal, as the first approach suffers from low sensitivity; the second one — from low linearity (Rabilloud and Lelong, 2011). Alternatively, fluorescent staining (e.g., SYPRO™ Ruby gel stain), providing both good sensitivity and linearity of up to three orders of magnitude, can be applied (Zhou et al., 2002). Independently



**Fig. 1.** The overview of the experimental workflows for gel-based and LC-based (gel-free) proteomics.

from their nature, all these methods generate patterns of characteristic signals, so-called “spots” visible by eye or in fluorescence detection mode (Soboleva et al., 2017). The resolution of the method is high enough to distinguish up to 1500 spots in one 2D-GE experiment (Gygi et al., 2000). However, on the one hand, it is difficult to separate all proteins (some spots can contain several polypeptides); and on the other hand, different isoforms of one protein can be represented by individual spots. Thus, in a good case scenario, several hundred proteins might be reliably identified in one 2D-electrophoregram of a seed protein extract. Annotation of electrophoretic zones typically relies on MS analysis. For this, individual

spots are excised, destained, and proteins, immobilized in gel, and can be digested with site-specific proteases (most often with trypsin) (Rathi et al., 2016). The resulting digests are extracted from gel sections and analyzed by matrix-assisted laser desorption-ionization — time of flight (MALDI-TOF) — or electrospray ionization — (ESI)-MS.

In the simplest case, MALDI-TOF-MS analysis relies on the so-called peptide fingerprinting approach — i.e. identification of proteins by the patterns of proteolytic peptides (Clauser, Baker, and Burlingame, 1999). This technique, however, lacks both sensitivity and selectivity: only the major proteins can be identified in each spot,

and this method requires very high electrophoretic resolution. In a more sophisticated method, MALDI-TOF/TOF instrumentation and tandem mass spectrometry (MS/MS) can be applied to obtain more reliable results (Binita et al., 2014; Min et al., 2017). ESI-MS, coupled online to reversed phase high-performance liquid chromatography (RP-HPLC), has a much higher potential in identification of individual proteins composing electrophoretic zones, and it might thereby provide a higher sequence coverage of proteins immobilized in gel (Arun, 2015).

One of the most important features of 2D-GE is its applicability to characterization of protein dynamics — i.e., alterations in protein expression profiles in time or in response to application of certain experimental conditions. Most often, quantification of proteins separated by 2D-GE relies on the measurements of staining intensities (e.g., by fluorescent detection) or UV/VIS densities of corresponding protein spots. This typically relies on software analysis tools providing convenient access to spot detection, matching, normalization, analysis, and annotation, also often combined with robotized spot cutting. The most widely used softwares are based on numerical approaches. PDQuest (Bio-Rad Laboratories), Delta2D (Decodon), and Melanie (GeneBio) can be successfully complemented by home-made systems (Marengo et al., 2005). Integration of densitograms of intensity patterns gives direct access to quantitative profiles of protein abundances (also often referred to as protein expression). The obtained values allow clear and straightforward statistical interpretation of data (Jensen, Jessen, and Jorgensen, 2008). For example, hierarchical clustering provides classification of individual proteins by their expression profiles (Faria-Oliveira et al., 2015). Subsequent annotation of protein functions and localization gives an opportunity for reliable conclusion about affected functions and metabolic pathways, as well as the impact of certain sub-cellular structures in observed plant response (Gong et al., 2012). In this way, protein dynamics (i.e. changes in expression of individual polypeptide species) can be addressed in the context of protein functions, relations to other molecular partners (proteins and metabolites), as well as localization in organ, tissue and cell (Li et al., 2016).

However, it is important to mention that despite multiple advantages of 2D-GE — relatively low prices for instrumentation, ease of handling, availability of effective interpretation software tools — this method has several intrinsic disadvantages. It suffers from low inter-gel reproducibility — i.e. high variability between treatment groups, analytical replicates, experiment time points and performers. Fortunately, the inter-group variation can be overcome by sample multiplexing in terms of the difference gel electrophoresis (DIGE) approach (Beckett, 2012). For this, cyanine dyes (Cy3 and Cy5) are added

to alternating samples, whereas the third dye (Cy2) can be used for standardization (Unlu, Morgan, and Mindin, 1997). Mixing all three mixtures of protein derivatives in a 1:1:1 ratio gives direct access to protein expression (Unlu, Morgan, and Minde, 1997). The individual dyes contain different fluorophore moieties, so green, yellow and red fluorescence can be acquired from the same gel with minimal cross-talk and sensitivities of up to 150 pg per visualized electrophoretic zone (Tan, Lim, and Lau, 2017). Unfortunately, the DIGE approach does not allow reduction of inter-replicate variability or the personnel-related bias. Indeed, as most of the spots contain more than one protein, the reliability of quantitative protein profiles obtained by 2D-GE is often questionable. Moreover, in comparison to LC-based methods, 2D-GE has much lower analytical resolution and linear dynamic range. Hence, currently, gel-free methods are being actively optimized in analysis of the seed proteome.

### **Gel-free (LC-based) bottom-up proteomic strategy**

In comparison to gel-based techniques, LC-MS represents a much more powerful analytical tool, providing reliable identification of several thousand proteins in one separation experiment (Kocher, Pichler, Swart, and Mechtler, 2012). Obviously, due to its high analytical resolution, the LC-MS-based approach is preferable in analysis of complex proteomes. Generally, it relies on separation of individual proteolytic peptides after enzymatic proteolysis of the total protein mixture (Fig. 1). However, LC-MS proteomics has some limitations which need to be kept in mind when planning plant proteomic experiments. First, the conventional ionization technique used in LC-MS-based proteomics is ESI (Soboleva et al., 2017), although off-line LC-MALDI-TOF/TOF-MS can also be employed (Krokhin, Ens, and Standing, 2005). This technique is highly prone to matrix effects, which are most often manifested as suppression of low-intensity peptide signals by highly abundant coeluting species (Taylor, 2005). Taking into account the high numbers of proteolytic peptides formed during enzymatic digestion (dozens or even hundreds of thousands), high separation efficiency becomes a critical requirement for successful ionization and detection of all of them.

Secondly, ESI is absolutely incompatible with detergents (in particular SDS and CHAPS), and these substances directly hamper LC-MS (Vissers, Chervet, and Salzmann, 1996). Moreover, SDS, Triton X100 and CHAPS, conventionally used in gel-based proteomics, can be efficiently retained on reversed phase and broadly coeluted with proteolytic peptides, suppressing their detection (Chen, Cociorva, Norris, and Yates, 2007). Therefore, conventional detergents cannot be used for solubilization of proteins when ESI-MS experiments

are planned. In the simplest method, researchers target only the soluble part of the plant proteome by extraction with aqueous buffers (Bilova et al., 2016a). In this case, it is necessary to keep in mind that such extracts may contain low molecular weight metabolites, which might affect efficiency of the proteolysis and therefore need to be removed prior to digestion. For this, size-exclusion chromatography, performed on PD-10 columns, proved to be an efficient tool (Bilova et al., 2016b). Alternatively, enzymatic digest of the whole amount of protein in each particular sample can be performed. For this, isolated protein is dissolved in a sample buffer containing SDS (which is known to be one of the best detergents) and separated by SDS-PAGE. Afterwards, the whole lanes are excised and cut in several (at least about ten) segments for in-gel digestion with trypsin (Han, Yin, He, and Yang, 2013). The resulting digests are pooled, desalted by reversed phase solid phase extraction (RP-SPE), lyophilized, and analyzed by LC-MS after reconstitution in an appropriate solvent. Obviously, both these strategies have a disadvantage of incomplete recovery of the seed proteome.

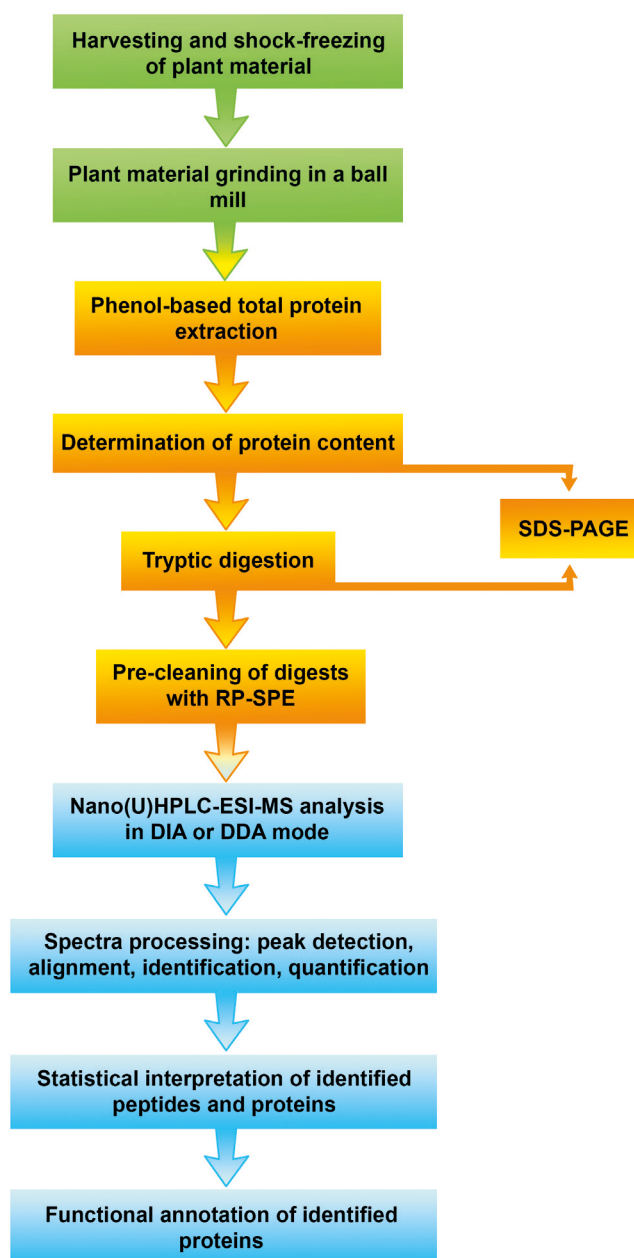
Another limitation of LC-based proteomics is so-called “undersampling” in LC-MS, related to data-dependent acquisition (DDA) experiments, which are still most widely used in LC-proteomics (Kalli, Smith, Sweredoski, and Hess, 2013). This type of MS experiment relies on acquisition of several (usually 5–10, or up to 20 in modern instruments) MS/MS spectra of the most intense signals observed in a so-called survey MS scan (which typically takes from 20 to 500 ms, depending on the instrument) (Majovsky et al., 2014). Thus, the number of MS/MS scans acquirable during the time of the whole DDA experiment is limited and usually essentially lower than the number of peptides detected on the MS level. Another issue worthy of consideration is the relatively low intra- and inter-batch precision of label-free quantification. Therefore, reliable inter- and intra-run normalization is strictly mandatory (Wilm, 2009). In this context, techniques relying on metabolic (Guo and Li, 2011), chemical (Yao, 2011) and/or stable isotope (Soboleva, 2017) labeling are advantageous, as they provide a possibility of sample multiplexing prior to (preferred) or after proteolysis. Finally, probably the most important factor limiting the application of LC-MS-based methods is the high costs of instrumentation and requirements of the skills of personnel.

Taken together, overcoming these limitations requires essential investments of work and resources. Because of this, 2D-GE is still the most widely used technique in plant protein research in general, and in seed proteomics in particular. However, several studies employing the LC-MS-based shotgun proteomics approach justify this technique as an advantageous one in comparison to other strategies. Indeed, the current state

of analytical science provides successful solutions for all the limitations listed above. First of all, the ultra-high performance chromatography (UHPLC) technique provides highly-efficient separations (Soboleva et al., 2017), minimizing matrix effects and suppression of low-intensity peptide signals by coeluting highly abundant counterparts. Secondly, quantitative reconstitution of hydrophobic proteins without any accompanying effects on the sensitivity has been achieved. For this, commercially available degradable detergents can be used (Zhang et al., 2013). Among them, RapiGest SF Surfactant (Waters Corporation) is probably the most widely used acid-labile detergent for blood plasma analysis, as it ensures complete reconstitution of isolated blood proteins and can be precipitated by hydrochloric acid upon enzymatic digestion (Jagadeeshaprasad et al., 2016). However, recently it was effectively applied to digestion of barley seed protein, resulting in identification of 226 proteins in shotgun data-independent experiments (Kaspar-Schoenefeld et al., 2016). Among many others, Progenta anionic acid labile surfactant (AALSII) is another option for enzymatic digestion of hydrophobic (e.g., membrane) proteins (Waas et al., 2016). We recently used this detergent for analysis of the total Arabidopsis leaf proteome (Frolov et al., 2016; Paudel et al., 2016). Later, we successfully applied this approach to germinating seeds of oilseed rape (*Brassica napus*) (Frolov et al., 2017). When combined with long chromatographic gradients, this digestion protocol results in identification of thousands of proteins (Paudel et al., 2016). To avoid analytical undersampling in DDA experiments (Kalli, Smith, Sweredoski, and Hess, 2013), enrichment, depletion and pre-fractionation procedures can be used. For example, glycosylated leaf protein can be selectively enriched by boronic acid chromatography (Bilova et al., 2016b), whereas removal of major legume reserve seed proteins by protamine sulfate is a good example of a depletion procedure that essentially increases sequence coverage of the seed proteome (Kim et al., 2015). Pre-fractionation can be reliably established on the chromatographic level by application of two-dimensional separations. For example, hydrophilic interaction liquid chromatography (HILIC) can be used as the first dimension (orthogonal to reversed phase chromatography). Several (typically 5–10) fractions are collected, dried, reconstituted in appropriate low-organic solvent and separated by RP-(U) HPLC. This design has proved to be applicable both for mammalian (human) (Bollineni, Hoffmann, and Fedorova, 2011) and plant (Paudel et al., 2016) samples. Alternatively, pre-fractionation can be performed at the level of multipole pre-filter in a mass-spectrometer. This technique, known as the gas phase fractionation approach (Canterbury et al., 2014), gives access to low-abundant peptides and, therefore, provides an essential increase of protein sequence coverage.

Summarizing and taking into account the considerations mentioned above, according to our experience, an efficient LC-MS-based proteomic workflow can be structured as follows (Fig. 2). After shock-freezing in liquid nitrogen, seed tissues are ground in a ball mill, and approximately 50–200 mg (depending on species) of grinded material are taken for phenol-based extraction procedure (Isaacson et al., 2006; Frolov et al., 2017). After several cleaning steps, proteins are precipitated from phenol phase with ice-cold methanolic ammonium acetate solution, sequentially washed with methanol and acetone, dried and completely reconstituted in a “shotgun buffer” containing chaotropic agents (urea and thiourea) and detergent (Progenta Anionic Acid-Labile Surfactant, AALSII) (Frolov et al., 2016). Afterwards, after determination of protein contents and verification of results with SDS-PAGE, solubilized proteins are digested by trypsin (Greifenhagen et al., 2016). When completeness of the proteolysis is verified with SDS-PAGE, AALS can be destroyed and the digests can be pre-cleaned with RP-SPE either in a cartridge or stage-tip format as described by Spiller et al. (2017).

After pre-cleaning, samples are analyzed by nano (U)HPLC–ESI–MS in a data-independent acquisition (DIA) or data-dependent acquisition (DDA) mode (Fig. 3). Identification of peptide sequences relies on tandem mass spectrometry (MS/MS, Fig. 3A and B), while label quantification is usually based on integration of extracted ion chromatograms (XICs, Fig. 3C, inserts). Whereas annotation of fragments by XICs is critical for DIA (Jamwal et al., 2017), the major challenge of DDA experiments is analytical undersampling (Kalli, Smith, Sweredoski, and Hess, 2013). Indeed, as only a limited number of quasi-molecular ions can be involved in fragmentation in each DDA cycle, an essential (if not the major) part of a proteome is left undetected. To avoid (or at least minimize) this problem, LC-MS analysis needs to rely either on enrichment/pre-fractionation techniques (when annotation of the maximum possible number of proteins is the main scope of the study) or prolonged gradients (when reliable label-free quantification is desired) (Bollineni, Hoffmann, and Fedorova, 2011; Paudel et al., 2016). The choice of mass analyzer is based on the principle goal of the research. When the experiment can be described as “discovery proteomics”, high mass accuracy at least at the MS level is a pre-requisite for reliable annotation of sequence tags. The high mass accuracy in MS/MS scans might significantly increase the reliability of protein annotation. Therefore, for DDA experiments, linear ion trap (LTQ)-Orbitrap and especially the modern quadrupole (Q)-Orbitrap instruments are the best choice (Bollineni, Hoffmann, and Fedorova, 2011; Kalli, Smith, Sweredoski, and Hess, 2013; Scheltema et al., 2014). Importantly, the acquired data are well compatible with the most well-established data process-

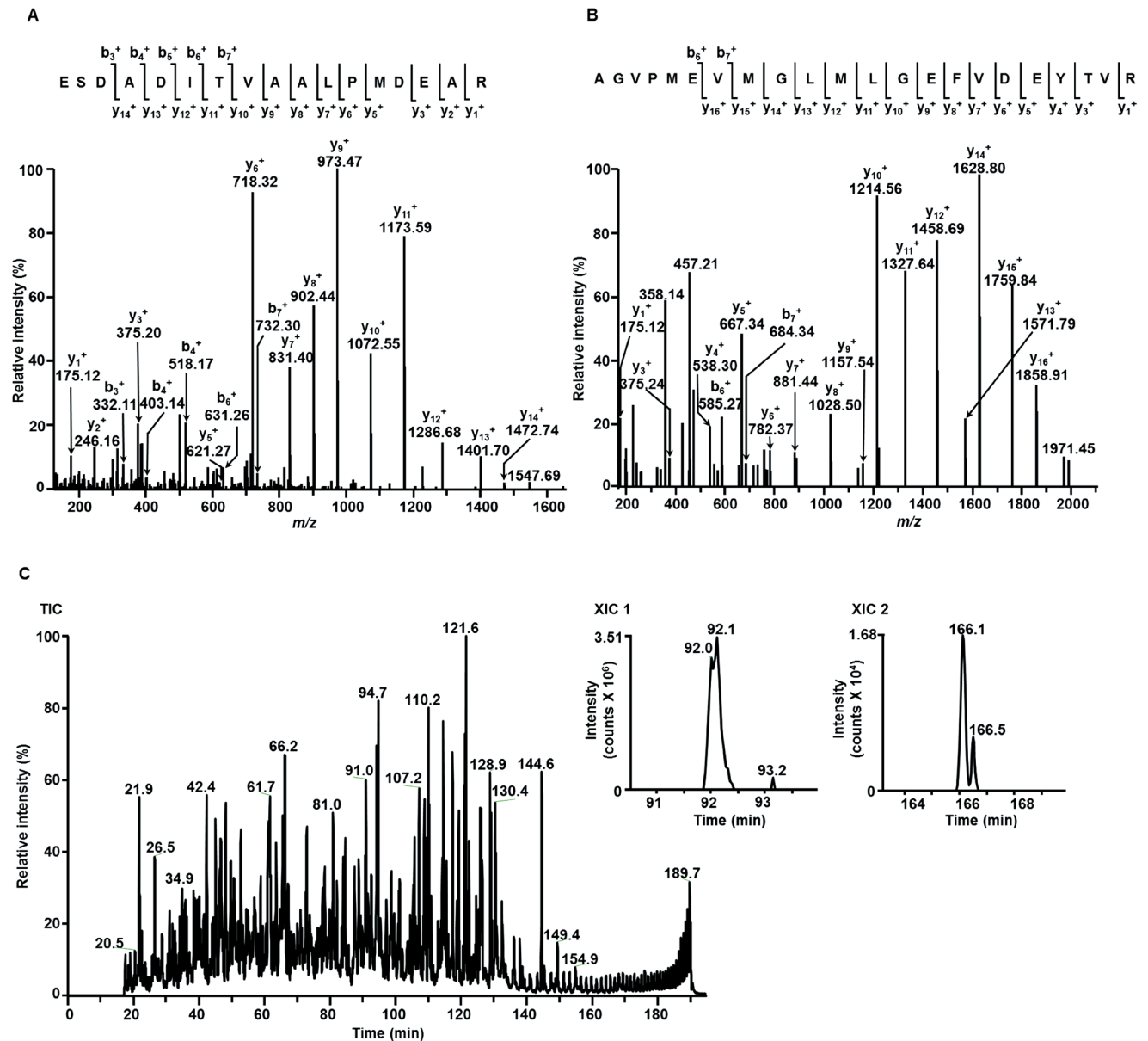


**Fig. 2.** Detailed workflow for LC-based proteomics: protein isolation, sample preparation, analysis and data processing.

ing pipelines, like OpenMS, Progenesis QI (Non-Linear Dynamics/Waters Corporation) and MaxQuant (Max Planck Institute of Biochemistry, Martinsried, Germany), compatible with various search machines (Weisser et al., 2013). The processed data can be subjected to statistical interpretation and used as input information for gene ontology tools (Tyanova et al., 2016).

### Profiles of post-translational modifications

Seed maturation is accompanied by accumulation of reserve biopolymers (Smolikova et al., 2017b). Although for some species reserve molecules are represented with



**Fig. 3.** Tandem mass spectra of  $m/z$  902.42 corresponding to the  $(M+2H)^{2+}$  of the tryptic peptide ESDADITVAALPMD<sup>240</sup>, representing glucose-1-phosphate adenylyltransferase (A), and of  $m/z$  1222.09 corresponding to the  $(M+2H)^{2+}$  of the tryptic peptide AGVPM<sup>72</sup> E V M G L M L G E F V D E Y T V R, representing 26S proteasome non-ATPase regulatory subunit-like protein (B), and annotation of these sequences by characteristic retention times ( $t_R$ ) from corresponding extracted ion chromatograms (C, inserts XIC1 and XIC2), obtained from a total ion chromatogram (C, TIC) of tryptic peptide mixture from *Pisum sativum* L. seed proteins: XIC 1 (for  $m/z$  902.42  $\pm$  0.02,  $t_R$  91.95); XIC 2 (for  $m/z$  1222.09  $\pm$  0.02,  $t_R$  166.15).

proteins (i.e. legumes), abundances of carbohydrates and lipids might be relatively high even in such seeds (Landry, Fuchs, and Hu, 2016). Although it is obviously favorable at the early steps of seed maturation, high contents of these metabolites might affect the patterns on non-enzymatic modifications of seed proteins. Indeed, oxidative stress, accompanying both seed maturation and ageing, in the presence of high amounts of reducing sugars (generated by de-polymerization of reserve polysaccharides) and unsaturated fatty acids or fatty acid residues in lipids, results in oxidative degradation of these biomolecules. Thereby, sugars are involved in met-

al-catalyzed monosaccharide autoxidation (Wolff and Dean, 1987), whereas fatty acid residues in lipids can be subjected to lipid peroxidation (Ayala, Munoz, and Arguelles, 2014). Resulting reactive carbonyl compounds (RCCs) readily interact with side chains of nucleophilic amino acid residues, such as lysine, arginine, histidine, and cysteine in reactions of glyco- and lipoxydation (Vistoli et al., 2013). Moreover, as the both processes are free radical-based, they are accompanied by enhanced production of reactive oxygen species, and they promote oxidative stress (Vistoli et al., 2013). Thus, glycation, lipid peroxidation and oxidation need to be addressed



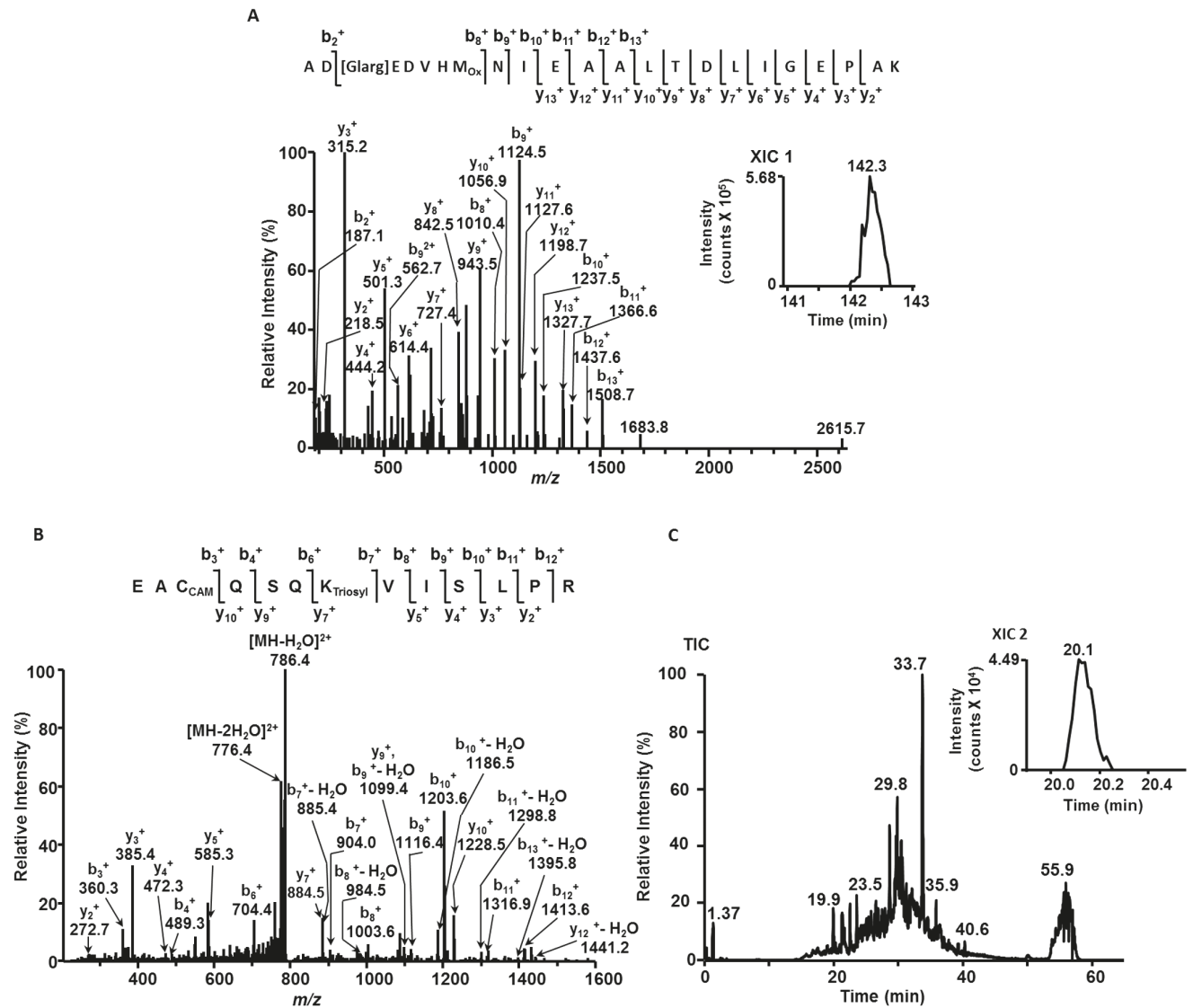
when studying seed development and ageing. Indeed, it was shown that seed proteins undergo this kind of modification during prolonged storage (Strelec, Ugaric-Hardi, and Hlevnjak, 2008) and accelerated ageing of seeds (Wettlaufer and Leopold, 1991). It is important to note that resulting advanced glycation end products (AGEs) and advanced lipoxidation end products (ALEs) are known to interact with immunoglobulin-like receptors for advanced glycation end-products (RAGEs), and to trigger inflammatory response in mammals. However, the phenomena of seed protein glyco- and lipoxydation were generally addressed only at the biochemical level (i.e. measurement of basic markers, like total contents of Amadori compounds and thiobarbituric acid-reactive substances) (Murthy, Kumar, and Sun, 2003) or profiles of prospective glyco- and lipoxidation products (Colville et al., 2012), whereas a comprehensive proteomics survey of modification profiles still needs to be done. In the most efficient way, such analyses might rely on the LC–MS-based approach with a special focus on in-depth analysis of low-abundant species.

As was mentioned before, the coverage of non-enzymatic modifications (which are typically represented with low-intensity signals in chromatograms) can be increased by implementation of enrichment/depletion and pre-fractionation procedures (Bilova et al., 2016b; Paudel et al., 2016). Besides, fragmentation of high-intensity signals of abundant unmodified peptides can be suppressed in DDA experiments by using the result-based exclusion strategy (Paudel et al., 2016). In terms of this approach, acquired DDA data are searched against a sequence database without consideration of specific variable post-translational modifications (typically, only methionine sulfoxide is used in this case). Based on these results, quasi-molecular ions of unmodified peptides are excluded from fragmentation, i.e. corresponding  $m/z$  values are entered in a so-called exclusion list, located in an instrument method. Analogously, peptide signals can be submitted to an inclusion list — in this case, target  $m/z$  value will be fragmented upon detection in the survey scan without consideration of relative signal intensity (Greifenhagen et al., 2014).

Identification of glyco- and lipoxydized peptides relies on their specific  $m/z$  increments and characteristic tandem mass spectrometric (MS/MS) fragmentation patterns (Frolov and Hoffmann, 2008; Milic, Hoffmann, and Fedorova, 2013). Thereby, either collision-activated dissociation (CAD) (Frolov et al., 2006) or electron capture/transfer dissociation (ECD/ETD) can be employed (Zhang et al., 2007). When fragmentation occurs under CAD conditions, sequence information can be derived by indicative b- and y-fragment ion series (Fig. 4). For most AGEs, ALEs and oxidation products, the positions of individual side chain modifications can be unambiguously assigned by characteristic mass increments of frag-

ment ions because these modifications are stable under CAD conditions (Ehrlich et al., 2009; Frolov, Blucher, and Hoffman, 2014; Greifenhagen, Frolov, and Hoffmann, 2014; Schmidt, Böhme, Singer, and Frolov, 2015). In contrast, early glycation products (i.e. Amadori and Heyns compounds) (Heyns and Noack, 1962; Hodge, 1955) are readily involved in fragmentation already at relatively low collision energies, yielding characteristic patterns of signals in their tandem mass spectra (Frolov, Singer, and Hoffmann, 2006). Thus, early glycated peptides can be easily recognized by a characteristic pattern of  $(M-H_2O)^{n+}$ ,  $(M-2H_2O)^{n+}$ ,  $(M-3H_2O)^{n+}$ ,  $(M-4H_2O)^{n+}$  signals and additional loss of one water and one formaldehyde molecule (Frolov and Hoffmann, 2010; Frolov et al., 2014). This characteristic pattern of oxonium, pyrylium and furylium ions (Horvat and Jakas, 2004) can be considered a signature of protein glycation in peptide sequence (Frolov et al., 2006). It is important to note that these patterns are less abundant when resonance excitation is used, i.e. the loss of one water molecule is usually missing in the spectra acquired with a Paul trap (Frolov et al., 2014). Remarkably, neutral losses related to glycation moiety accompany b- and y-fragment series and, therefore, can be used for peptide sequencing (Fig. 4) (Frolov, Singer, and Hoffmann, 2006). This is also true for complex molecules containing a monosaccharide part, e.g., nucleotides (Fedorova, Frolov, and Hoffmann, 2010). However, for oligosaccharide-derived Amadori compounds (e.g.,  $N^{\epsilon}$ -(lactosyl)lysine-containing peptides), CAD does not deliver any valuable sequence information (Milkovska-Stamenova and Hoffmann, 2016). Therefore, in such cases, ETD or ECD fragmentation capabilities are strongly desired.

Due to the high variability of non-enzymatic post-translational modifications, the high complexity of real proteomes and the different fragmentation behavior of individual PTMs, it seems logical to address the pathways of fragmentation in some simple models. In the most straightforward way, such experiments might rely on synthetic peptide models (Frolov, Blucher, and Hoffmann, 2014; Greifenhagen, Frolov, and Hoffmann, 2015). Indeed, synthesis of glycated peptides is well established for a large selection of early and advanced glycation products (Gruber and Hofmann, 2005; Frolov, Singer, and Hoffmann, 2006, 2007; Ehrlich et al., 2009; Greifenhagen, Frolov, and Hoffmann, 2014; Schmidt, Böhme, Singer, and Frolov, 2015). Based on these synthetic standards, efficient methods for structural characterization, as well as specific and sensitive detection, can be designed. Detection techniques might rely on neutral loss (Gadgil et al., 2007) or precursor ion (Frolov et al., 2006; Greifenhagen et al., 2014; Schmidt et al., 2015) scans. In the latter case, indicative characteristic fragments, related to modified amino acid residues, can serve as diagnostic signals. On the other hand, the knowledge of



**Fig. 4.** Tandem mass spectra of  $m/z$  855.42 corresponding to the  $(M+3H)^{3+}$  of the tryptic AGE-modified peptide AD(Gl a r g)E D V H M O<sub>x</sub>NIEAALT-DLIGEP A K<sup>159</sup>, representing argininosuccinate lyase from *Pisum sativum* L. seeds (A), and of  $m/z$  794.41 corresponding to the  $(M+2H)^{2+}$  of the tryptic Amadori/Heyns-modified peptide E A C<sub>CAM</sub>Q S Q K<sub>Triosyl</sub>V I S L P R<sup>501</sup> representing chloride channel protein CLC-d from *Arabidopsis thaliana* L. leaves (B), and annotation of these by the corresponding extracted ion chromatograms (XICs, inserts) at the characteristic retention times ( $t_{R}$ ), obtained from corresponding total ion chromatogram (TICs) (TIC at Fig. 3) XIC 1 (for  $m/z$  855.42  $\pm$  0.02,  $t_{R}$  142.13); C — TIC and XIC 2 (for  $m/z$  794.41  $\pm$  0.02,  $t_{R}$  20.08). Modifications: Gl a r g, glyoxal-derived hydroimidazolone, 1-(4-amino-4-carboxybutyl)2-imino-5-oxo-imidazolidine; M O<sub>x</sub>, methionine sulfoxide; C<sub>CAM</sub>, cysteine carbamidomethylation; K<sub>Triosyl</sub>, N<sup>ε</sup>-(triosyl)lysine.

fragmentation patterns allows reliable identification of individual oxidation or glycation products (Frolov et al., 2014; Greifenhagen, Frolov, and Hoffmann, 2015; Greifenhagen, Frolov, Bluher, and Hoffmann, 2016a, 2016b). Additionally, it provides valuable information for new quantification methods. Thus, based on the patterns of characteristic b- and y-fragments, multiple reaction monitoring (MRM)-based quantitative methods can be developed (Spiller, Frolov, and Hoffmann, 2017; Spiller et al., 2017). Not less importantly, synthetic peptide models provide access to valuable kinetic information. Indeed, both formation (Frolov et al., 2014) and degradation (Greifenhagen, Frolov, and Hoffmann, 2015) ki-

netics of individual glycation products can be addressed with this analytical tool.

## Conclusion

The seed proteome represents a highly complex system, only a minor part of which has been characterized so far. Changes in protein dynamics accompanying seed germination are characterized better, but still insufficiently, due to generally low overall numbers of annotated seed proteins. Moreover, the role of non-enzymatic post-translational modifications in regulatory events accompanying seed development and germination has also

been only minimally addressed so far. Thus, the quality and completeness of proteomics information available to date are insufficient and need to be improved. The most obvious reason for this is the relatively low analytical power of the methods routinely used for characterization of the seed proteome. Indeed, as 2D-GE is the most widely used analytical technique for this, additional enrichment/depletion or pre-fractionation steps need to be included in established workflows to increase their analytical power. However, one needs to keep in mind the limited potential of gel-based techniques to characterize the whole complexity of the seed proteome. Therefore, LC-based methods (which are currently rarely used), relying on highly-effective separation systems, are obviously advantageous. High resolution mass spectrometry needs to be employed for efficient identification and reliable quantification of individual proteins.

## References

- Altshul, A.M. 1962. Seed proteins and world food problems. *Economy Botany* 16:2–13. <https://doi.org/10.1007/BF02986049>
- Arun, K.A.V. 2015. Gel-based proteomics in plants: time to move on from the tradition. *Frontiers of Plant Science* 6:369. <https://doi.org/10.3389/fpls.2015.00369>
- Asgar, M.A., Fazilah, A., Huda, N., Bhat, R., and Karim, A.A. 2010. Nonmeat protein alternatives as meat extenders and meat analogs. *Comprehensive Reviews in Food Science and Food Safety* 9:513–529. <https://doi.org/10.1111/j.1541-4337.2010.00124.x>
- Ayala, A., Munoz, M.F., and Arguelles, S. 2014. Lipid peroxidation: production, metabolism, and signaling mechanisms of malondialdehyde and 4-hydroxy-2-nonenal. *Oxidative Medicine and Cellular Longevity* 2014:360–438. <https://doi.org/10.1155/2014/360438>
- Bailly, C. and Kraner, I. 2011. Analyses of ROS and antioxidants in relation to seed longevity and germination. *Methods in Molecular Biology* 773:343–367. [https://doi.org/10.1007/978-1-61779-231-1\\_20](https://doi.org/10.1007/978-1-61779-231-1_20)
- Barkla, B.J. and Vera-Estrella, R. 2015. Single cell-type comparative metabolomics of epidermal bladder cells from the halophyte *Mesembryanthemum crystallinum*. *Frontiers in Plant Science* 6:435. <https://doi.org/10.3389/fpls.2015.00435>
- Bates, P.D., Stymne, S., and Ohlrogge, J. 2013. Biochemical pathways in seed oil synthesis. *Current Opinion in Plant Biology* 16:358–364. <https://doi.org/10.1016/j.pbi.2013.02.015>
- Beckett, P. 2012. The basics of 2D DIGE. *Methods in Molecular Biology* 854:9–19. [https://doi.org/10.1007/978-1-61779-573-2\\_2](https://doi.org/10.1007/978-1-61779-573-2_2)
- Bilova, T., Greifenhagen, U., Paudel, G., Lukashova, E., Brauch, D., Osmolovskaya, N., Tarakhovskaya, E., Balcke, G.U., Tissier, A., Vogt, T., Milkowski, C., Birkemeyer, C., Wessjohann, L.A., and Frolov, A. 2016a. Glycation of plant proteins under environmental stress—methodological approaches, potential mechanisms and biological role; pp. 295–316 in Shanker, A. (Ed.). *Abiotic and biotic stress in plants — recent advances and future perspectives*. IntechOpen. ISBN 978-953-51-2250-0. <https://doi.org/10.5772/61860>
- Bilova, T., Lukashova, 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., and Frolov, A. 2016b. A Snapshot of the plant glycated proteome: structural, functional, and mechanistic aspects. *Journal of Biological Chemistry* 291:7621–7636. <https://doi.org/10.1074/jbc.M115.678581>
- Binita, K., Kumar, S., Sharma, V.K., Sharma, V., and Yadav, S. 2014. Proteomic identification of *Syzygium cumini* seed extracts by MALDI-TOF/MS. *Applied Biochemical Biotechnology* 172:2091–2105. <https://doi.org/10.1007/s12010-013-0660-x>
- Bollineni, R.C., Hoffmann, R., and Fedorova, M. 2011. Identification of protein carbonylation sites by two-dimensional liquid chromatography in combination with MALDI- and ESI-MS. *Journal of Proteomics* 74:2338–2350. <https://doi.org/10.1016/j.jprot.2011.07.002>
- Bourgeois, M., Jacquin, F., Savoie, V., Sommerer, N., Labas, V., Henry, C., and Burstin, J. 2009. Dissecting the proteome of pea mature seeds reveals the phenotypic plasticity of seed protein composition. *Proteomics* 9:254–271. <https://doi.org/10.1002/pmic.200700903>
- Bradford, K.J., Dahal, P., Asbrouk, J.V., Kunusoth, K., Bello, P., Thompson, J., and Wu, F. 2018. The dry chain: Reducing postharvest losses and improving food safety in humid climates. *Trends in Food Science and Technology* 71:84–93. <https://doi.org/10.1016/j.tifs.2017.11.002>
- Brink-Jensen, K., Bak, S., Jorgensen, K., and Ekstrom, C.T. 2013. Integrative analysis of metabolomics and transcriptomics data: a unified model framework to identify underlying system pathways. *PLoS One* 8:e72116. <https://doi.org/10.1371/journal.pone.0072116>
- Canterbury, J.D., Merrihew, G.E., MacCoss, M.J., Goodlett, D.R., and Shaffer, S.A. 2014. Comparison of data acquisition strategies on quadrupole ion trap instrumentation for shotgun proteomics. *Journal of American Society for Mass Spectrometry* 25:2048–2059. <https://doi.org/10.1007/s13361-014-0981-1>
- Carre, P. and Pouzet, A. 2012. Rapeseed market, worldwide and in Europe oilseeds and fats. *Crops and Lipids* 21:D102. <https://doi.org/10.1051/ocl/2013054>
- Catherman, A.D., Skinner, O.S., and Kelleher, N.L. 2014. Top Down proteomics: facts and perspectives. *Biochemical and Biophysical Research Communications* 445:683–693. <https://doi.org/10.1016/j.bbrc.2014.02.041>
- Cernay, C., Pelzer, E., and Makovski, D. 2016. A global experimental dataset for assessing grain legume production. *Scientific Data* 3:160084. <https://doi.org/10.1038/sdata.2016.84>
- Chen, D., Shen, X., and Sun, L. 2017. Capillary zone electrophoresis-mass spectrometry with microliter-scale loading capacity, 140 min separation window and high peak capacity for bottom-up proteomics. *Analyst* 142:2118–2127. <https://doi.org/10.1039/c7an00509a>
- Chen, E.I., Cociorva, D., Norris, J.L., and Yates, J.R., III. 2007. Optimization of mass spectrometry-compatible surfactants for shotgun proteomics. *Journal of Proteome Research* 6:2529–2538. <https://doi.org/10.1021/pr060682a>
- Chevallet, M., Luche, S., Diemer, H., Strub, J.M., Van, D.A., and Rabilloud, T. 2008. Sweet silver: a formaldehyde-free silver staining using aldoses as developing agents, with enhanced compatibility with mass spectrometry. *Proteomics* 8:4853–4861. <https://doi.org/10.1002/pmic.200800321>
- Chmelik, J., Zidkova, J., Rehulka, P., Petry-Podgorska, I., and Bobalova, J. 2009. Influence of different proteomic protocols on degree of high-coverage identification of non-specific lipid transfer protein 1 modified during malting.

- Electrophoresis* 30:560–567. <https://doi.org/10.1002/elps.200800530>
- Clauser, K. R., Baker, P., and Burlingame, A. L. 1999. Role of accurate mass measurement (+/- 10 ppm) in protein identification strategies employing MS or MS/MS and database searching. *Analytical Chemistry* 71:2871–2882. <https://doi.org/10.1021/ac9810516>
- Colville, L., Bradley, E. L., Lloyd, A. S., Pritchard, H. W., Castle, L., and Kranner, I. 2012. 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. *Journal of Experimental Botany* 63:6519–6530. <https://doi.org/10.1093/jxb/ers307>
- Corbineau, F. 2012. Markers of seed quality: from present to future. *Seed Science Research* 22:S61–S68. <https://doi.org/10.1017/S0960258511000419>
- Don, R. 2009. ISTA Handbook on seedling evaluation. 3rd Edition. By the ISTA Germination Committee <https://www.seedtest.org/>
- Duranti, M. 2006. Grain legume proteins and nutraceutical properties. *Fitoterapia* 77:67–82. <https://doi.org/10.1016/j.fitote.2005.11.008>
- Ehrlich, H., Hanke, T., Frolov, A., Langrock, T., Hoffmann, R., Fischer, C., Schwarzenbolz, U., Henle, T., Born, R., and Worch, H. 2009. Modification of collagen in vitro with respect to formation of Nepsilon-carboxymethyllysine. *International Journal of Biological Macromolecules* 44:51–56. <https://doi.org/10.1016/j.ijbiomac.2008.10.001>
- FAO Cereal Supply and Demand Brief. 2017. <https://www.fao.org/worldfoodsituation>
- Faria-Oliveira, F., Carvalho, J., Ferreira, C., Hernaez, M. L., Gil, C., and Lucas, C. 2015. Quantitative differential proteomics of yeast extracellular matrix: there is more to it than meets the eye. *BMC Microbiology* 15:271. <https://doi.org/10.1186/s12866-015-0550-1>
- Fedorova, M., Frolov, A., and Hoffmann, R. 2010. Fragmentation behavior of Amadori-peptides obtained by non-enzymatic glycosylation of lysine residues with ADP-ribose in tandem mass spectrometry. *Journal of Mass Spectrometry* 45:664–669. <https://doi.org/10.1002/jms.1758>
- Filho, J. M. 2015. Seed vigor testing: an overview of the past, present and future. *Scientia Agricola* 72:363–374. <https://doi.org/10.1590/0103-9016-2015-0007>
- Food and Agriculture Organization of the United Nations, 2018. <https://www.fao.org>
- Frolov, A. and Hoffmann, R. 2008. Analysis of Amadori peptides enriched by boronic acid affinity chromatography. *Annals of New York Academy of Science* 1126:253–256. <https://doi.org/10.1196/annals.1433.060>
- Frolov, A. and Hoffmann, R. 2010. Identification and relative quantification of specific glycation sites in human serum albumin. *Analytical and Bioanalytical Chemistry* 397:2349–2356. <https://doi.org/10.1007/s00216-010-3810-9>
- Frolov, A., Bilova, T., Paudel, G., Berger, R., Balcke, G. U., Birkenmeyer, C., and Wessjohann, L. A. 2016. Early responses of mature *Arabidopsis thaliana* plants to reduced water potential in the agar-based polyethylene glycol infusion drought model. *Journal of Plant Physiology* 280:70–83. <https://doi.org/10.1016/j.jplph.2016.09.013>
- Frolov, A., Bluhner, M., and Hoffmann, R. 2014. Glycation sites of human plasma proteins are affected to different extents by hyperglycemic conditions in type 2 diabetes mellitus. *Analytical and Bioanalytical Chemistry* 406:5755–5763. <https://doi.org/10.1007/s00216-014-8018-y>
- Frolov, A., Didio, A., Ihling, C., Chantzeva, V., Grishina, T., Hoehenwarter, W., Sinz, A., Smolikova, G., Bilova, T., and Medvedev, S. 2017. The effect of simulated microgravity on the Brassica napus seedling proteome. *Functional Plant Biology* 45(4):440–452. <https://doi.org/10.1071/FP16378>
- Frolov, A., Hoffmann, P., and Hoffmann, R. 2006. Fragmentation behavior of glycated peptides derived from D-glucose, D-fructose and D-ribose in tandem mass spectrometry. *Journal of Mass Spectrometry* 41:1459–1469. <https://doi.org/10.1002/jms.1117>
- Frolov, A., Schmidt, R., Spiller, S., Greifenhagen, U., and Hoffmann, R. 2014. Formation of arginine-derived advanced glycation end products in peptide-glucose mixtures during boiling. *Journal of Agricultural and Food Chemistry* 62:3626–3635. <https://doi.org/10.1021/jf4050183>
- Frolov, A., Singer, D., and Hoffmann, R. 2006. Site-specific synthesis of Amadori-modified peptides on solid phase. *Journal of Peptide Science* 12:389–395. <https://doi.org/10.1002/psc.739>
- Frolov, A., Singer, D., and Hoffmann, R. 2007. Solid-phase synthesis of glucose-derived Amadori peptides. *Journal of Peptide Science* 13:862–867. <https://doi.org/10.1002/psc.901>
- Fu, Y. B., Ahmed, Z., and Diederichsen, A. 2015. Towards a better monitoring of seed ageing under ex situ seed conservation. *Conservative Physiology* 3:cov026. <https://doi.org/10.1093/conphys/cov026>
- Gadgil, H. S., Bondarenko, P. V., Treuheit, M. J., and Ren, D. 2007. Screening and sequencing of glycated proteins by neutral loss scan LC/MS/MS method. *Analytical Chemistry* 79:5991–5999. <https://doi.org/10.1021/ac070619k>
- Gallardo, K., Kurt, C., Thompson, R., and Ochatt, S. 2018. In vitro culture of immature *M. truncatula* grains under conditions permitting embryo development comparable to that observed in vivo. *Plant Science* 170:1052–1058. <https://doi.org/10.1016/j.plantsci.2005.12.021>
- Gallardo, K., Le, S. C., Vandekerckhove, J., Thompson, R. D., and Burstin, J. 2003. Proteomics of *Medicago truncatula* seed development establishes the time frame of diverse metabolic processes related to reserve accumulation. *Plant Physiology* 133:664–682. <https://doi.org/10.1104/pp.103.025254>
- Ghassemi-Golezani, K. and Mazloomi-Oskooyi, R. 2008. Effect of water supply on seed quality development in common bean (*Phaseolus vulgaris* var.). *Journal of Plant Production* 2:117–124. <https://doi.org/10.22069/IJPP.2012.604>
- Gillet, L. C., Leitner, A., and Aebbersold, R. 2016. Mass spectrometry applied to bottom-up proteomics: entering the high-throughput era for hypothesis testing. *Annual Reviews of Analytical Chemistry (Palo Alto, California)* 9:449–472. <https://doi.org/10.1146/annurev-anchem-071015-041535>
- Goff, S. A. and Salmeron, J. M. 2004. Back to the future of cereals. Genomic studies of the world's major grain crops, together with a technology called marker-assisted breeding, could yield a new green revolution. *Scientific American* 291:42–49. PMID: 15298118
- Gong, C. Y., Li, Q., Yu, H. T., Wang, Z., and Wang, T. 2012. Proteomics insight into the biological safety of transgenic modification of rice as compared with conventional genetic breeding and spontaneous genotypic variation. *Journal of Proteome Research* 11:3019–3029. <https://doi.org/10.1021/pr300148w>
- Greifenhagen, U., Frolov, A., and Hoffmann, R. 2015. Oxidative degradation of N-fructosylamine-substituted peptides in heated aqueous systems. *Amino Acids* 47:1065–1076. <https://doi.org/10.1007/s00726-015-1940-2>
- Greifenhagen, U., Frolov, A., Bluhner, M., and Hoffmann, R. 2016a. Plasma proteins modified by Advanced Glycation End products (AGEs) reveal site-specific susceptibilities to glycemic control in patients with Type 2 diabetes.

- Journal of Biological Chemistry* 291:9610–9616. <https://doi.org/10.1074/jbc.M115.702860>
- Greifenhagen, U., Frolov, A., Bluhner, M., and Hoffmann, R. 2016b. Site-specific analysis of advanced glycation end products in plasma proteins of type 2 diabetes mellitus patients. *Analytical and Bioanalytical Chemistry* 408:5557–5566. <https://doi.org/10.1007/s00216-016-9651-4>
- Greifenhagen, U., Nguyen, V. D., Moschner, J., Giannis, A., Frolov, A., and Hoffmann, R. 2014. Sensitive and site-specific identification of carboxymethylated and carboxyethylated peptides in tryptic digests of proteins and human plasma. *Journal of Proteome Research* 14:768–777. <https://doi.org/10.1021/pr500799m>
- Gruber, P. and Hofmann, T. 2005. Chemoselective synthesis of peptides containing major advanced glycation end-products of lysine and arginine. *Journal of Peptide Research* 66:111–124. <https://doi.org/10.1111/j.1399-3011.2005.00279.x>
- Guerrier, N. and Gavaletto, M. 2018. Cereal proteins. Second edition: 223–244.
- Guo, G. and Li, N. 2011. Relative and accurate measurement of protein abundance using <sup>15</sup>N stable isotope labeling in *Arabidopsis* (SILIA). *Phytochemistry* 72:1028–1039. <https://doi.org/10.1016/j.phytochem.2011.01.007>
- Gupta, M., Bhaskar, P. B., Sriram, S., and Wang, P. H. 2017. Integration of omics approaches to understand oil/protein content during seed development in oilseed crops. *Plant Cell Reports* 36:637–652. <https://doi.org/10.1007/s00299-016-2064-1>
- Gupta, R., Min, C. W., Kim, S. W., Wang, Y., Agrawal, G. K., Rakwal, R., Kim, S. G., Lee, B. W., Ko, J. M., Baek, I. Y., Bae, D. W., and Kim, S. T. 2015. Comparative investigation of seed coats of brown- versus yellow-colored soybean seeds using an integrated proteomics and metabolomics approach. *Proteomics* 15:1706–1716. <https://doi.org/10.1002/pmic.201400453>
- Gygi, S. P., Corthals, G. L., Zhang, Y., Rochon, Y., and Aebersold, R. 2000. Evaluation of two-dimensional gel electrophoresis-based proteome analysis technology. *Proceedings of North Atlantic Academy of Science, U. S. A.* 97:9390–9395. <https://doi.org/10.1073/pnas.160270797>
- Hampton, J. H. and TeKrony, D. 1995. ISTA. Handbook of Vigor Test Methods. By the ISTA Germination Committee <https://www.seedtest.org/>
- Han, C., Yin, X., He, D., and Yang, P. 2013. Analysis of proteome profile in germinating soybean seed, and its comparison with rice showing the styles of reserves mobilization in different crops. *PLoS One*. 8:e56947. <https://doi.org/10.1371/journal.pone.0056947>
- Heyer, R., Schallert, K., Zoun, R., Becher, B., Saake, G., and Bendorf, D. 2017. Challenges and perspectives of metaproteomic data analysis. *Journal of Biotechnology* 261:24–36. <https://doi.org/10.1016/j.jbiotec.2017.06.1201>
- Heyns, K. and Noack, H. 1962. Die Umsetzung von D-Fructose mit L-Lysine und L-Arginin und deren Beziehung zu nichtenzymatischen Bräunungsreaktionen. *Chemische Berichte* 95:720–727. <https://doi.org/10.1002/cber.19620950323>
- Hodge, J. E. 1955. The Amadori rearrangement. *Advances of Carbohydrate Chemistry* 10:169–205. [https://doi.org/10.1016/S0096-5332\(08\)60392-6](https://doi.org/10.1016/S0096-5332(08)60392-6)
- Horvat, S. and Jakas, A. 2004. Peptide and amino acid glycation: new insights into the Maillard reaction. *Journal of Peptide Science* 10:119–137. <https://doi.org/10.1002/psc.519>
- Hummel, M., Wigger, T., and Brockmeyer, J. 2015. Characterization of mustard 2S albumin allergens by bottom-up, middle-down, and top-down proteomics: a consensus set of isoforms of Sin a 1. *Journal of Proteome Research* 14:1547–1556. <https://doi.org/10.1021/pr5012262>
- I. C. E. Committees, 2015. World agricultural supply and demand estimates. <https://www.usda.gov>
- Isaacson, T., Damasceno, C. M., Saravanan, R. S., He, Y., Catala, C., Saladie, M., and Rose, J. K. 2006. Sample extraction techniques for enhanced proteomic analysis of plant tissues. *Nature Protocols* 1:769–774. <https://doi.org/10.1038/nprot.2006.102>
- Jagadeeshaprasad, M. G., Batkulwar, K. B., Meshram, N. N., Tiwari, S., Korwar, A. M., Unnikrishnan, A. G., and Kulkarni, M. J. 2016. Targeted quantification of N-1-(carboxymethyl) valine and N-1-(carboxyethyl) valine peptides of beta-hemoglobin for better diagnostics in diabetes. *Clinical Proteomics* 13:7. <https://doi.org/10.1186/s12014-016-9108-y>
- Jamwal, R., Barlock, B. J., Adusumalli, S., Ogasawara, K., Simons, B. L., and Akhlaghi, F. 2017. Multiplex and label-free relative quantification approach for studying protein abundance of drug metabolizing enzymes in human liver microsomes using SWATH-MS. *Journal of Proteome Research* 16:4134–4143. <https://doi.org/10.1021/acs.jproteome.7b00505>
- Jensen, K. N., Jessen, F., and Jorgensen, B. M. 2008. Multivariate data analysis of two-dimensional gel electrophoresis protein patterns from few samples. *Journal of Proteome Research* 7:1288–1296. <https://doi.org/10.1021/pr700800s>
- Jimenez-Lopez, J. C., Foley, R. C., Brear, E., Clarke, V. C., Lima-Cabello, E., Florido, J. F., Singh, K. B., Alche, J. D., and Smith, P. M. C. 2018. Characterization of narrow-leaf lupin (*Lupinus angustifolius* L.) recombinant major allergen IgE-binding proteins and the natural beta-conglutin counterparts in sweet lupin seed species. *Food Chemistry* 244:60–70. <https://doi.org/10.1016/j.foodchem.2017.10.015>
- Kalli, A., Smith, G. T., Sweredoski, M. J., and Hess, S. 2013. Evaluation and optimization of mass spectrometric settings during data-dependent acquisition mode: focus on LTQ-Orbitrap mass analyzers. *Journal of Proteome Research* 12:3071–3086. <https://doi.org/10.1021/pr3011588>
- Kaspar-Schoenefeld, S., Merx, K., Jozefowicz, A. M., Hartmann, A., Seiffert, U., Weschke, W., Matros, A., and Mock, H. P. 2016. Label-free proteome profiling reveals developmental-dependent patterns in young barley grains. *Journal of Proteomics* 143:106–121. <https://doi.org/10.1016/j.jprot.2016.04.007>
- Kim, Y. J., Wang, Y., Gupta, R., Kim, S. W., Min, C. W., Kim, Y. C., Park, K. H., Agrawal, G. K., Rakwal, R., Choung, M. G., Kang, K. Y., and Kim, S. T. 2015. Protamine sulfate precipitation method depletes abundant plant seed-storage proteins: A case study on legume plants. *Proteomics* 15:1760–1764. <https://doi.org/10.1002/pmic.201400488>
- Kocher, T., Pichler, P., Swart, R., and Mechtler, K. 2012. Analysis of protein mixtures from whole-cell extracts by single-run nanoLC-MS/MS using ultralong gradients. *Nature Protocols* 7:882–890. <https://doi.org/10.1038/nprot.2012.036>
- Kocsy, G. 2015. Die or survive? Redox changes as seed viability markers. *Plant Cell Environment* 38:1008–1010. <https://doi.org/10.1111/pce.12515>
- Krokhin, O. V., Ens, W., and Standing, K. G. 2005. MALDI QqTOF MS combined with off-line HPLC for characterization of protein primary structure and post-translational modifications. *Journal of Biomolecular Technology* 16:429–440.
- Lafiandra, D., Masi, S., Sissons, M., Dornez, E., Delour, J. A., Ourtin, M., and Aboni, M. F. 2012. Kernel components of technological value. Second edition:85–124.
- Landry, E. J., Fuchs, S. J., and Hu, J. 2016. Carbohydrate composition of mature and immature faba bean seeds. *Jour-*

- nal of Food Component Analysis* 50:55–60. <https://doi.org/10.1016/j.jfca.2016.05.010>
- Larance, M. and Lamond, A. I. 2015. Multidimensional proteomics for cell biology. *Nature Review in Molecular Cell Biology* 16:269–280. <https://doi.org/10.1038/nrm3970>
- Li, Q. F., Xiong, M., Xu, P., Huang, L. C., Zhang, C. Q., and Liu, Q. Q. 2016. Dissection of brassinosteroid-regulated proteins in rice embryos during germination by quantitative proteomics. *Scientific Reports* 6:34583. <https://doi.org/10.1038/srep34583>
- Liu, J., Hua, W., Yang, H., Guo, T., Sun, X., Wang, X., Liu, G., and Wang, H. 2014. Effects of specific organs on seed oil accumulation in *Brassica napus* L. *Plant Science* 227:60–68. <https://doi.org/10.1016/j.plantsci.2014.06.017>
- Lombardo, V. A., Osorio, S., Borsani, J., Lauxmann, M. A., Bustamante, C. A., Budde, C. O., Andreo, C. S., Lara, M. V., Fernie, A. R., and Drincovich, M. F. 2011. Metabolic profiling during peach fruit development and ripening reveals the metabolic networks that underpin each developmental stage. *Plant Physiology* 157:1696–1710. <https://doi.org/10.1104/pp.111.186064>
- Majovsky, P., Naumann, C., Lee, C. W., Lassowskat, I., Trujillo, M., Dismeyer, N., and Hoehenwarter, W. 2014. Targeted proteomics analysis of protein degradation in plant signaling on an LTQ-Orbitrap mass spectrometer. *Journal of Proteome Research* 13:4246–4258. <https://doi.org/10.1021/pr500164j>
- Marengo, E., Robotti, E., Antonucci, F., Cecconi, D., Campos-trini, N., and Righetti, P. G. 2005. Numerical approaches for quantitative analysis of two-dimensional maps: a review of commercial software and home-made systems. *Proteomics* 5:654–666. <https://doi.org/10.1002/pmic.200401015>
- May, P., Christian, N., Ebenhoh, O., Weckwerth, W., and Walther, D. 2011. Integration of proteomic and metabolomic profiling as well as metabolic modeling for the functional analysis of metabolic networks. *Methods of Molecular Biology* 694:341–363. [https://doi.org/10.1007/978-1-60761-977-2\\_21](https://doi.org/10.1007/978-1-60761-977-2_21)
- McDonald, M. B. 1998. Seed quality assessment. *Seed Science Research* 8:265–275. <https://doi.org/10.1017/S0960258500004165>
- Messina, M. J. 1999. Legumes and soybeans: overview of their nutritional profiles and health effects. *American Journal of Clinical Nutrition* 70:439–450. <https://doi.org/10.1093/ajcn/70.3.439s>
- Milic, I., Hoffmann, R., and Fedorova, M. 2013. Simultaneous detection of low and high molecular weight carbonylated compounds derived from lipid peroxidation by electrospray ionization-tandem mass spectrometry. *Analytical Chemistry* 85:156–162. <https://doi.org/10.1021/ac302356z>
- Milkovska-Stamenova, S. and Hoffmann, R. 2016. Identification and quantification of bovine protein lactosylation sites in different milk products. *Journal of Proteomics* 134:112–126. <https://doi.org/10.1016/j.jprot.2015.07.021>
- Miller, J. K., Herman, E. M., Jahn, M., and Bradford, K. J. 2010. Strategic research, education and policy goals for seed science and crop improvement. *Plant Science* 179:645–652. <https://doi.org/10.1016/j.plantsci.2010.08.006>
- 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., and Kim, S. T. 2017. In-depth proteomic analysis of *Glycine max* seeds during controlled deterioration treatment reveals a shift in seed metabolism. *Journal of Proteomics* 169:125–135. <https://doi.org/10.1016/j.jprot.2017.06.022>
- Mochida, K. and Shinozaki, K. 2011. Advances in omics and bioinformatics tools for systems analyses of plant functions. *Plant Cell Physiology* 52:2017–2038. <https://doi.org/10.1093/pcp/pcr153>
- Mortz, E., O'Connor, P. B., Roepstorff, P., Kelleher, N. L., Wood, T. D., McLafferty, F. W., and Mann, M. 1996. Sequence tag identification of intact proteins by matching tandem mass spectral data against sequence data bases. *Proceedings of Northatlantic Academy of Science, U. S. A.* 93:8264–8267. <https://doi.org/10.1073/pnas.93.16.8264>
- Murad, M. M. and Rech, E. L. 2012. NanoUPL-MS<sup>E</sup> proteomic data assessment of soybean seeds using the Uniprot database. *BMC Biotechnology* 12:82. <https://doi.org/10.1186/1472-6750-12-82>
- Murthy, U. M., Kumar, P. P., and Sun, W. Q. 2003. Mechanisms of seed ageing under different storage conditions for *Vigna radiata* (L.) Wilczek: lipid peroxidation, sugar hydrolysis, Maillard reactions and their relationship to glass state transition. *Journal of Experimental Botany* 54:1057–1067. <https://doi.org/10.1093/jxb/erg092>
- Nagel, M., Kranner, I., Neumann, K., Rolletschek, H., Seal, C. E., Colville, L., Fernandez-Marin, B., and Borner, A. 2015. Genome-wide association mapping and biochemical markers reveal that seed ageing and longevity are intricately affected by genetic background and developmental and environmental conditions in barley. *Plant Cell Environment* 38:1011–1022. <https://doi.org/10.1111/pce.12474>
- Nedumaran, S., Abinaya, P., Jyosthna, P., Shraavya, B., Parthasarathy, R., and Bantilan, C. 2015. Grain legumes production, consumption and trade trends in developing countries. Working paper series No 60. *ICRISAT Research program, markets, institutions and politics*. <https://oar.icrisat.org/id/eprint/8991>
- Neuhoff, V., Arold, N., Taube, D., and Ehrhardt, W. 1988. Improved staining of proteins in polyacrylamide gels including isoelectric focusing gels with clear background at nanogram sensitivity using Coomassie Brilliant Blue G-250 and R-250. *Electrophoresis* 9:255–262. <https://doi.org/10.1002/elps.1150090603>
- Palma, J. M., Corpas, F. J., and del Rio, L. A. 2011. Proteomics as an approach to the understanding of the molecular physiology of fruit development and ripening. *Journal of Proteomics* 74:1230–1243. <https://doi.org/10.1016/j.jprot.2011.04.010>
- Paudel, G., Bilova, T., Schmidt, R., Greifenhagen, U., Berger, R., Tarakhovskaya, E., Stöckhardt, S., Balcke, G. U., Humbek, K., Brandt, W., Sinz, A., Vogt, T., Birkemeyer, C., Wessjohann, L., and Frolov, A. 2016. Osmotic stress is accompanied by protein glycation in *Arabidopsis thaliana*. *Journal of Experimental Botany* 67:6283–6295. <https://doi.org/10.1093/jxb/erw395>
- Pereira Lima, J. J., Buitink, J., Lalanne, D., Rossi, R. F., Pelletier, S., da Silva, E. A. A., and Leprince, O. 2017. Molecular characterization of the acquisition of longevity during seed maturation in soybean. *PLoS One* 12:e0180282. <https://doi.org/10.1371/journal.pone.0180282>
- Rabilloud, T. and Lelong, C. 2011. Two-dimensional gel electrophoresis in proteomics: a tutorial. *Journal of Proteomics* 74:1829–1841. <https://doi.org/10.1016/j.jprot.2011.05.040>
- Rabilloud, T., Chevallet, M., Luche, S., and Lelong, C. 2010. Two-dimensional gel electrophoresis in proteomics: past, present and future. *Journal of Proteomics* 73:2064–2077. <https://doi.org/10.1016/j.jprot.2010.05.016>
- Rajjou, L., Lovigny, Y., Groot, S. P., Belghazi, M., Job, C., and Job, D. 2008. Proteome-wide characterization of seed aging in *Arabidopsis*: a comparison between artificial and natural aging protocols. *Plant Physiology* 148:620–641. <https://doi.org/10.1104/pp.108.123141>

- Rathi, D., Gayen, D., Gayali, S., Chakraborty, S., and Chakraborty, N. 2016. Legume proteomics: progress, prospects, and challenges. *Proteomics* 16:310–327. <https://doi.org/10.1002/pmic.201500257>
- Sano, N., Rajjou, L., North, H. M., Debeaujon, I., Marion-Poll, A., and Seo, M. 2016. Staying alive: Molecular aspects of seed longevity. *Plant Cell Physiology* 57:660–674. <https://doi.org/10.1093/pcp/pcv186>
- Scheltema, R. A., Hauschild, J. P., Lange, O., Hornburg, D., Denisov, E., Damoc, E., Kuehn, A., Makarov, A., and Mann, M. 2014. The Q Exactive HF, a Benchtop mass spectrometer with a pre-filter, high-performance quadrupole and an ultra-high-field Orbitrap analyzer. *Molecular and Cellular Proteomics* 13:3698–3708. <https://doi.org/10.1074/mcp.M114.043489>
- Schmidt, R., Böhme, D., Singer, D., and Frolov, A. 2015. Specific tandem mass spectrometric detection of AGE-modified arginine residues in peptides. *Journal of Mass Spectrometry* 50:613–624. <https://doi.org/10.1002/jms.3569>
- Smolikova, G. N. 2014. Application of the method of accelerated aging to evaluate the stress tolerance of seeds. *Biological Communications [s.l.]* 2:82–93.
- Smolikova, G. N. and Medvedev, S. S. 2016. Photosynthesis in the seeds of Chloroembryophytes. *Russian Journal Plant Physiology* 63:1–12. <https://doi.org/10.1134/S1021443715060163>
- Smolikova, G. N., Shavarda, A. L., Alekseichuk, I. V., Chantzeva, V. V., and Medvedev, S. S. 2016. The metabolomic approach to the assessment of cultivar specificity of *Brassica napus*. *Russian Journal Genetic: Applied Research* 6:78–83. <https://doi.org/10.1134/S2079059716010147>
- Smolikova, G., Dolgikh, E., Vikhnina, M., Frolov, A., and Medvedev, S. 2017a. Genetic and hormonal regulation of chlorophyll degradation during maturation of seeds with green embryos. *International Journal of Molecular Science* 18:E1993. <https://doi.org/10.3390/ijms18091993>
- Smolikova, G., Kreslavski, V., Shiroglazova, O., Frolov, A., and Medvedev, S. 2017b. Photochemical activity changes accompanying the embryogenesis of pea (*Pisum sativum*) with yellow and green cotyledons. *Functional Plant Biology* 45(2):228–235. <https://doi.org/10.1071/FP16379>
- Soboleva, A., Modzei, M., Didio, A., Plociennik, H., Kijewska, H., Grishina, T., Karonova, T., Bilova, T., Stefanov, V., Stefanovic, P., and Frolov, A. 2017. Quantification of prospective type 2 diabetes mellitus biomarkers by stable isotope dilution with bi-labeled standard glycosylated peptides. *Analytical Methods* 9:409–418. <https://doi.org/10.1039/C6AY02483A>
- Soboleva, A., Schmidt, R., Vikhnina, M., Grishina, T., and Frolov, A. 2017. Maillard proteomics: Opening new pages. *International Journal of Molecular Science* 18: E2677. <https://doi.org/10.3390/ijms18122677>
- Spiller, S., Frolov, A., and Hoffmann, R. 2017. Quantification of specific glycation sites in human serum albumin as prospective type 2 diabetes mellitus biomarkers. *Protein and Peptide Letters* 24:887–896 <https://doi.org/10.2174/092986652466617020124120>
- Spiller, S., Li, Y., Bluhner, M., Welch, L., and Hoffmann, R. 2017. Glycosylated lysine-141 in haptoglobin improves the diagnostic accuracy for type 2 diabetes mellitus in combination with glycosylated hemoglobin HbA1c and fasting plasma glucose. *Clinical Proteomics* 14:10. <https://doi.org/10.1186/s12014-017-9145-1>
- Strelec, I., Ugarcic-Hardi, Z., and Hlevnjak, M. 2008. Accumulation of Amadori and Maillard products in wheat seeds aged under different storage conditions. *Croatica Chemica Acta* 81:131–137. <https://hrcak.srce.hr/23436>
- Suma, A., Sreenivasan, K., Singh, A. K., and Radhamani, J. 2013. Role of relative humidity in processing and storage of seeds and assessment of variability in storage behaviour in *Brassica* spp. and *Eruca sativa*. *Scientific World Journal* 2013:504141. <https://doi.org/10.1155/2013/504141>
- Takac, T., Samajova, O., and Samaj, J. 2017. Integrating cell biology and proteomic approaches in plants. *Journal of Proteomics* 169:165–175. <https://doi.org/10.1016/j.jprot.2017.04.020>
- Tan, B. C., Lim, Y. S., and Lau, S. E. 2017. Proteomics in commercial crops: an overview. *Journal of Proteomics* 169:176–188. <https://doi.org/10.1016/j.jprot.2017.05.018>
- Taylor, P. J. 2005. Matrix effects: the Achilles heel of quantitative high-performance liquid chromatography-electrospray-tandem mass spectrometry. *Clinical Biochemistry* 38:328–334. <https://doi.org/10.1016/j.clinbiochem.2004.11.007>
- Tyanova, S., Temu, T., Sinitcyn, P., Carlson, A., Hein, M. Y., Geiger, T., Mann, M., and Cox, J. 2016. The Perseus computational platform for comprehensive analysis of (prote) omics data. *Nature Methods* 13:731–740. <https://doi.org/10.1038/nmeth.3901>
- Unlu, M., Morgan, M. E., and Minden, J. S. 1997. Difference gel electrophoresis: a single gel method for detecting changes in protein extracts. *Electrophoresis* 18:2071–2077. <https://doi.org/10.1002/elps.1150181133>
- Vissers, J. P., Chervet, J. P., and Salzman, J. P. 1996. Sodium dodecyl sulphate removal from tryptic digest samples for on-line capillary liquid chromatography/electrospray mass spectrometry. *Journal of Mass Spectrometry* 31:1021–1027. [https://doi.org/10.1002/\(SICI\)1096-9888\(199609\)31:9<1021::AID-JMS384>3.0.CO;2-G](https://doi.org/10.1002/(SICI)1096-9888(199609)31:9<1021::AID-JMS384>3.0.CO;2-G)
- Vistoli, G., De, M. D., Cipak, A., Zarkovic, N., Carini, M., and Aldini, G. 2013. Advanced glycoxidation and lipoxidation end products (AGEs and ALEs): an overview of their mechanisms of formation. *Free Radical Research* 47 Suppl 1:3–27. <https://doi.org/10.3109/10715762.2013.815348>
- Waas, M., Bhattacharya, S., Chuppa, S., Wu, X., Jensen, D. R., Omasits, U., Wollscheid, B., Volkman, B. F., Noon, K. R., and Gundry, R. L. 2014. Combine and conquer: surfactants, solvents, and chaotropes for robust mass spectrometry based analyses of membrane proteins. *Analytical Chemistry* 86:1551–1559. <https://doi.org/10.1021/ac403185a>
- Walters, C., Ballesteros, D., and Vertucci, V. 2010. Structural mechanics of seed deterioration: Standing the test of time. *Plant Science* 179:565–573. <https://doi.org/10.1016/j.plantsci.2010.06.016>
- Wang, L., Fu, J., Li, M., Fagner, L., Weckwerth, W., and Yang, P. 2016. Metabolomic and proteomic profiles reveal the dynamics of primary metabolism during seed development of lotus (*Nelumbo nucifera*). *Frontiers in Plant Science* 7:750 <https://doi.org/10.3389/fpls.2016.00750>
- Weisser, H., Nahnsen, S., Grossmann, J., Nilse, L., Quandt, A., Brauer, H., Sturm, M., Kenar, E., Kohlbacher, O., Aebersold, R., and Malmstrom, L. 2013. An automated pipeline for high-throughput label-free quantitative proteomics. *Journal of Proteome Research* 12:1628–1644. <https://doi.org/10.1021/pr300992u>
- Wettlaufer, S. H. and Leopold, A. C. 1991. Relevance of Amadori and Maillard products to seed deterioration. *Plant Physiology* 97:165–169. <https://doi.org/10.1104/pp.97.1.165>
- Wilm, M. 2009. Quantitative proteomics in biological research. *Proteomics* 9:4590–4605. <https://doi.org/10.1002/pmic.200900299>

- Wolff, S. P. and Dean, R. T. 1987. Glucose autoxidation and protein modification. The potential role of 'autoxidative glycosylation' in diabetes. *Biochemical Journal* 245:243–250. <https://doi.org/10.1042/bj2450243>
- Yao, X. 2011. Derivatization or not: a choice in quantitative proteomics. *Analytical Chemistry* 83:4427–4439. <https://doi.org/10.1021/ac200925p>
- Yoshida, H., Tomiyama, Y., Yoshida, N., Shibata, K., and Mizushima, Y. 2010. Regiospecific profiles of fatty acids in triacylglycerols and phospholipids from adzuki beans (*Vigna angularis*). *Nutrients* 2:49–59. <https://doi.org/10.3390/nu20100049>
- Zhang, Q., Frolov, A., Tang, N., Hoffmann, R., van de Goor, T., Metz, T. O., and Smith, R. D. 2007. Application of electron transfer dissociation mass spectrometry in analyses of non-enzymatically glycosylated peptides. *Rapid Communications in Mass Spectrometry* 21:661–666. <https://doi.org/10.1002/rcm.2884>
- Zhang, Y., Fonslow, B. R., Shan, B., Baek, M. C., and Yates, J. R., III. 2013. Protein analysis by shotgun/bottom-up proteomics. *Chemical Reviews* 113:2343–2394. <https://doi.org/10.1021/cr3003533>
- Zhou, G., Li, H., DeCamps, D., Chen, S., Shu, H., Gong, Y., Flaig, M., Gillespie, J. W., Hu, N., Taylor, P. R., Emmert-Buk, M. R., Liotta, L. A., Petrioin, E. F., and Zhao, Y. 2002. 2D differential in-gel electrophoresis for the identification of esophageal cancer-associated protein markers. *Molecular Cell Proteomics* 1:117–123. <https://doi.org/10.1074/mcp.M100015-MCP200>