Available online at www.sciencedirect.com
SCIENCE DIRECT®
Journal of Chromatography B, 815 (2005) 109–123
JOURNAL OF CHROMATOGRAPHY B
www.elsevier.com/locate/chromb

Review

System, trends and perspectives of proteomics in dicot plants Part I: Technologies in proteome establishment

Ganesh Kumar Agrawal[a,1], Masami Yonekura[b], Yumiko Iwahashi[c], Hitoshi Iwahashi[d], Randeep Rakwal[a,d,*]

[a] *Research Laboratory for Agricultural Biotechnology and Biochemistry (RLABB), G.P.O. Box 8207, Kathmandu, Nepal*
[b] *Food Function Laboratory, School of Agriculture, Ibaraki University, Ami, Ibaraki 300-0393, Japan*
[c] *National Food Research Institute (NFRI), Kannondai 2-1-12, Tsukuba, Ibaraki 305-8642, Japan*
[d] *Human Stress Signal Research Center (HSS), National Institute of Advanced Industrial Science and Technology (AIST), Central 6, 1-1-1 Higashi, Tsukuba, Ibaraki 305-8566, Japan*

Received 26 April 2004; accepted 2 November 2004
Available online 9 December 2004

Abstract

The first 3 years of the 21st century have seen the impact of plant proteomics on functional genomics that has enhanced our understanding, not only on the plant genome(s), but also more importantly, on the functional aspect of proteins. This is mainly due to availability of the complete genome sequence of the *Arabidopsis thaliana*—a dicotyledoneous (dicot) model plant—and technological advancements in proteomics. Proteomic analyses of a variety of dicot plants, including both *Arabidopsis* and the model legume species, barrel medic (*Medicago truncatula*), have greatly helped in an efficient separation, identification and cataloguing of a large number of proteins, and thereby defining their proteomes. Therefore, we have composed an inclusive review on dicot plant materials, as of February 2004, that provides system, trends and perspectives of proteomics in growth and development and the environment. The review is summarized and discussed as three individual, but interlinked, entities: Part I, technologies in proteome establishment (this review), Part II, proteomes of the complex developmental stages [G.K. Agrawal, M. Yonekura, Y. Iwahashi, H. Iwahashi, R. Rakwal, J. Chromatogr. B (2004)], and Part III, unraveling the proteomes influenced by the environment, and at the levels of function and genetic relationships [G.K. Agrawal, M. Yonekura, Y. Iwahashi, H. Iwahashi, R. Rakwal, J. Chromatogr. B (2004)]. This review deals with the diverse proteomic technologies being used in proteome development of different dicot plants.

Keywords: Dicotyledoneous plants; Sample preparation; Techniques; 2-DGE; MS; Protein identification; Bioinformatics

Contents

* Corresponding author. Fax: +81 29 861 6066.
E-mail addresses: gkagrawal@onebox.com (G.K. Agrawal), rakwal-68@aist.go.jp (R. Rakwal).
[1] Co-corresponding author.

1570-0232/$ – see front matter © 2004 Elsevier B.V. All rights reserved.
doi:10.1016/j.jchromb.2004.11.024

1. Introduction

This is the age (21st century) of plant functional genomics. With the availability of the complete genome information of model plant species, *Arabidopsis thaliana* (L.) Heyhn [1,2], and rice (*Oryza sativa* L.) [2–5], we are in a position to use multiparallel approaches, and to apply a range of new technologies, to the functional analysis of plant genomes in a high-throughput mode. This inevitably also leads to an unprecedented pace (and efficiency) in analysis and deduction of gene (one or many) function in a short time, at the transcript (transcriptomics), protein (proteomics) and metabolite (metabolomics) levels.

Proteomics, one of the "rapidly emerging and expanding fields" of the functional genomics era, involves a systematic and detailed analysis of the protein population in a cell, subcellular compartment, tissue, and whole organisms, and complements genomics and metabolomics [6–8]. A graduate student at Macquarie University in Australia, Marc Wilkins, coined the term "proteome"—PROTEin complement of the genome—in 1994 [9]. It has only been a few years (2000 onwards) since the term proteomics "as a whole" is being applied to investigate the plant proteome. Plant proteomics is still in its infancy compared to the proteomic analyses of prokaryotes, such as yeast, and humans. This is partly due to the lack of availability of complete genomic or cDNA sequences from plants. Nevertheless, plant proteomics is gaining momentum, and is poised to become a major area of research and development in the field of plant biology (for reviews, see [10–18]). It has been elegantly stated by Watson and co workers "as we seek to better understand the gene function and to study the holistic biology of systems, it is inevitable that we study the proteome" [19].

To proteomic analyses, *A. thaliana* and barrel medic (*Medicago truncatula*) are currently the models of choice among the dicotyledonous (dicot) plants (Fig. 1). *Arabidopsis*—a long-day flowering plant—was the first flowering plant to have its genome decoded, and its complete genome sequence carrying approximately 25,000 genes is freely available and almost perfectly annotated [1,2,20]. *M. truncatula* is an excellent model for studying the symbiotic root nodule formation [21–24]. Proteomic analyses of these two plant materials will reveal the complete set of proteins encoded by their respective genomes, and at the same time will lay a foundation for comparative proteomic analysis with other dicot plants. Nonetheless, proteomic approaches have already been undertaken on a variety of other dicot plants, such as beans, flax, grape, oilseed rape, pea, poppy, potato, tobacco, tomato, spinach, etc. Their studies have provided a number of improved methodologies including sample preparation and techniques, and proteomes of plant materials from developmental stages to environmental stresses, and to plant–microbe interactions. Moreover, improvements in two core technologies—the classical two-dimensional gel electrophoresis (2-DGE) and mass spectrometry (MS)—have helped advance plant proteomics. There are some good reviews on plant proteomics as a whole [10–18], and in general deal with the proteomics of dicot plants. Therefore, this review along with two other reviews [25,26] in series is intended to cover all the studies conducted as of February 2004 on dicot plant materials from proteomics viewpoint in order to reveal the system, trends and perspectives therein. This review describes the technologies (old and new) used for establishing proteomes of dicot plants, their limitations and the challenges ahead.

2. Technology

2.1. Sample preparation

Good sample preparation—"extraction of a maximum number of proteins from a given cell, tissue, organ or organism"—is the most important step for subsequent separation, resolution, and identification of proteins. It reminds us of the statement "the key to good sample preparation is efficient protein solubilization with a minimum of handling" [27]. Several sample preparation methods are now available for different plant materials (Figs. 2 and 3; and for details see [28–30]). The most popular one is the use of tricholoroacetic acid (TCA) and acetone (TCA/acetone) for the direct precipitation of proteins from a given cell material [31]. An advantage is the immediate precipitation of proteins and si-

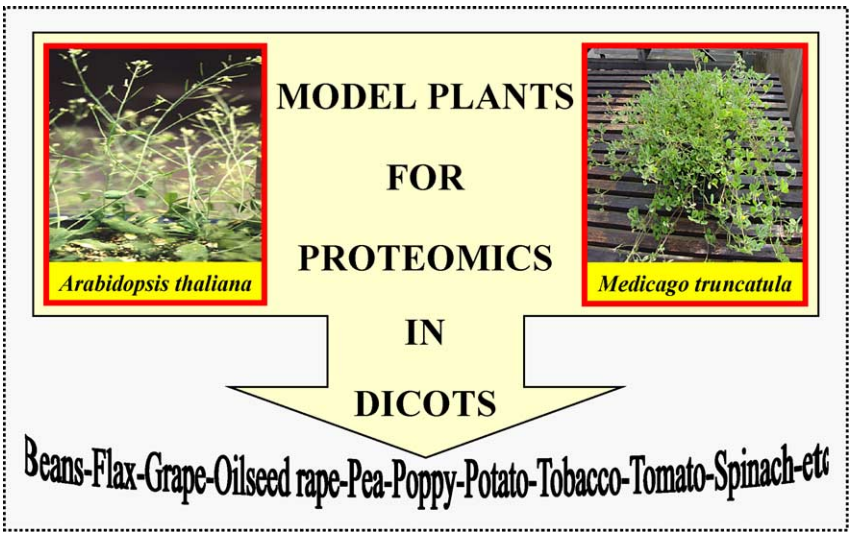

Fig. 1. Dicot plant proteomics. *Arabidopsis thaliana* (ecotype Columbia) and *Medicago truncatula* genotype Jemalong A17 serve as model plants for dicots. The common names of the other dicot plants used for proteome analyses are given below. Photographs of *Arabidopsis* and *Medicago* were kindly provided by Dr. Akihiro Kubo (NIES) and Prof. Richard Oliver (Murdoch University), respectively.

multaneous inactivation of components involved in protein degradation, such as the proteases. Additionally, the method removes several compounds (salts, pigments and polyphenolics) that interfere with isoelectric focusing (IEF) from the samples. Nonetheless, protein precipitation usually results in protein losses and also causes difficulties in resolubilization of proteins. To overcome this problem, sequential solubilization technique can be used [32–34]. Another one involves the solubilization of proteins in phenol, with or without sodium dodecyl sulfate (SDS), and subsequently precipitation with methanol and ammonium acetate [35]. Its benefit lies in the fact that high quality protein extracts can be generated with minimal contamination. These methods are schematically depicted in Fig. 2. The sequential solubilization technique along with recently developed other sample preparation techniques, like sequential fractionation [36] and extraction of recalcitrant plant tissues [37] are discussed below.

2.1.1. Sequential solubilization

Sequential solubilization, one of the powerful sample preparation techniques [32,33], allows fractionation of the sample based on solubility, molecular mass (M_r) and isoelectric point (p*I*). Sequential solubilized fractions carry more proteins than separated without sequential solubilization. The

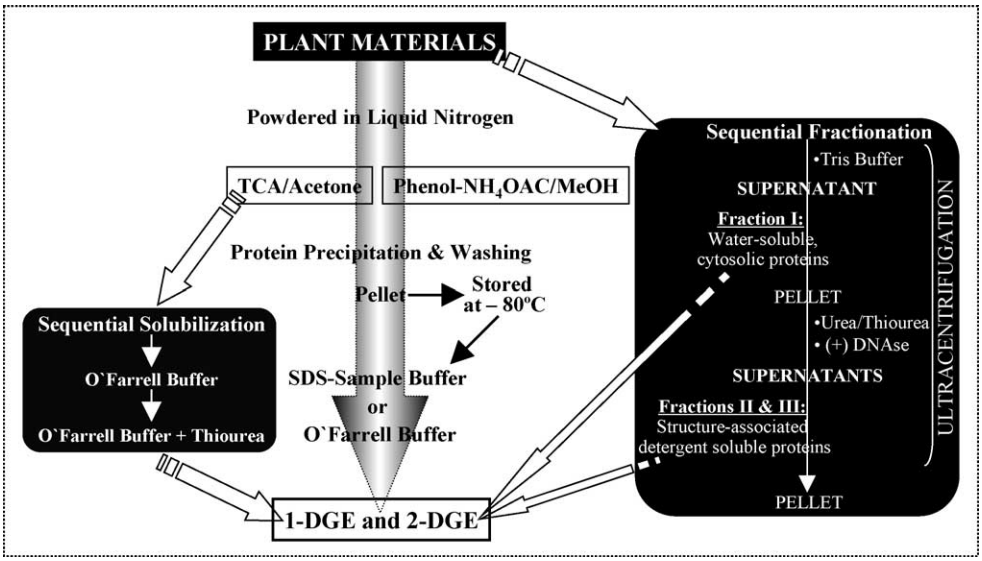

Fig. 2. Sample preparation techniques. The flow chart of TCA/acetone (the most commonly used sample preparation method), Phenol-NH_4OAC/MeOH (method suitable for resistant and hard tissues, like wood and olive leaf, etc.), sequential solubilization, and sequential fractionation is outlined. The final samples prepared in SDS-sample buffer can be used for both 1- and 2-DGE, whereas the samples solubilized/prepared in O'Farrell buffer are used for 2-DGE only. For details on each method, especially sequential fractionation, please see the text (and the original articles).

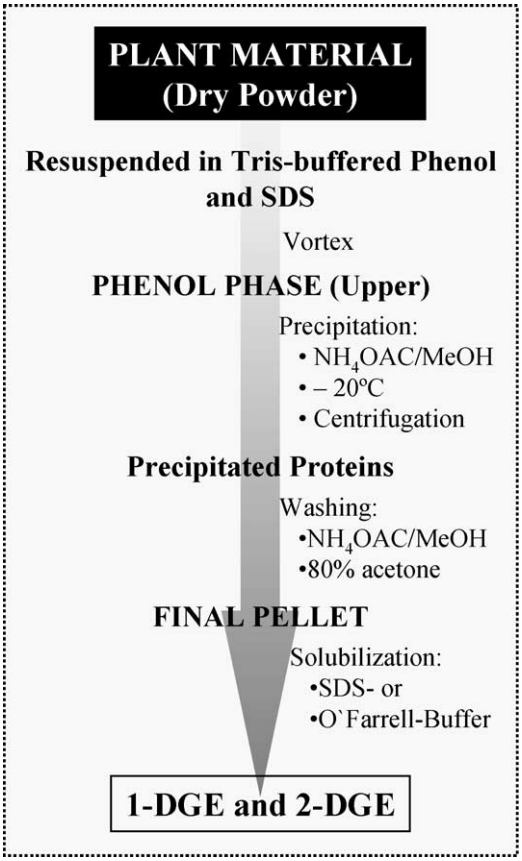

Fig. 3. Sample preparation from recalcitrant plant tissues. The flow chart schematically depicts each step involved in sample preparation. Briefly, the dry power obtained upon grinding in liquid nitrogen is resuspended in Trsi-buffered phenol, containing SDS, which after vortexing results in protein-rich phenol phase. Proteins are precipitated by NH_4OAC/MeOH, and solubilized in SDS- or O'Farrell-buffer.

method was modified for the proteome research of *Catharanthus roseus* (Madagascar periwinkle) suspension-cultured cells (cell line A11) and seedlings [34]. The improvement includes solubilization of proteins with a conventional urea buffer followed by a stronger solubilizing buffer containing thiourea after precipitation of proteins in TCA/acetone (Fig. 2). The 2-DGE protein profiles of sequential solubilized fractions were reported to be very different, where only 10% of the total number of protein spots was detected in both samples. This modification increased the number of protein spots by 52% compared to the proteins detected after solubilization in a single step. However, even this method has its limitations; for example, an insoluble residue still remains after extraction with thiourea buffer.

2.1.2. Sequential fractionation

The suitability of this technique was demonstrated on extraction of total proteins from *Arabidopsis* (ecotype Columbia). A three-fold increase in protein spots was revealed by the use of 2-DGE [36]. Total proteins (leaf and stem) were extracted in three fractions: the first fraction containing the cytoplasmatic, water-soluble proteins, second and third fractions containing the structure-associated, detergent-soluble proteins, including the membrane and nucleic acid-associated proteins (Fig. 2). By doing this, the complexity of proteins in each fraction was decreased, and rare proteins were enriched. The method also avoided loss of proteins by omitting technical steps like precipitation, washing of cell pellets, dialysis and lyophilization of protein samples.

2.1.3. Extraction of recalcitrant plant tissues

Some of the plant materials, such as olive leaf, are notoriously recalcitrant to common protein extraction methods due to high levels of interfering compounds. In a recent paper, it was shown that addition of phenol extraction to the TCA/acetone precipitation protocol greatly improved protein extraction from olive leaf [37]. The method basically involved: (a) preparation of a very fine dry acetone powder of leaf tissue; (b) extensive washing with organic solvent and aqueous TCA to remove pigments, lipids, etc. and water soluble contaminants, respectively; and (c) phenol extraction of proteins in the presence of SDS (Fig. 3).

2.1.4. Membrane and hydrophobic proteins

Membrane and hydrophobic proteins are difficult to solubilize completely, and hence attempts have been made by several groups to tackle this problem. For example, the plasma membrane (PM) is one of the examples for low recovery of hydrophobic proteins. Several methods have been tested to enrich a membrane sample in hydrophobic proteins [38–40]. One method involved purification of the PM from *Arabidopsis* leaf by the phase partitioning method [41]. It was also shown that PM treatment with carbonate at high pH and solubilization with C8Ø detergent favors the isolation of integral proteins and the release of peripheral proteins [38]. In another study, Santoni and co-workers tested four different extraction procedures (Triton X-100, Triton X-114, carbonate treatment, chloroform/methanol treatment) for hydrophobic protein isolation, and six different lysis buffers [7 M urea, 2 M thiourea, 0.5% Triton X-100, 1.2% pharmalytes (pH 3–10), 20 mM DTT, and 2% (w/v) of detergent (C5Ø, C6Ø, C7Ø, C8Ø, and ASB14) or 4% (w/v) CHAPS] for solubilization of the isolated hydrophobic proteins (for details see [39,40]). These studies reached to the conclusion that first, the efficiency of detergents to solubilize hydrophobic proteins is dependent on the lipid content of the samples, and second, the need for a preliminary study to optimize the solubilization conditions for individual experiments/samples.

Recently, using *Arabidopsis* leaf membrane proteins, Luche and co-workers [42] investigated the solubilizing power of various non-ionic and zwitterionic detergents as membrane protein solubilizers for 2-DGE. Among the commercially available non-ionic detergents, dodecyl maltoside and decaethylene glycol mono hexadecyl ether proved most efficient. Though this progress has been able to find new detergents for membrane protein solubilization, solubilization of hydrophobic proteins or their elution from the gel matrices, and ionization for subsequent MS analyses [43,44], still

remains a challenge. One way is the direct use of the prepared samples in MS. A good example comes from the isolation, separation and identification of *Arabidopsis* integral membrane carrier proteins [45].

2.1.5. Subcellular fractionation

Current technology does not favor a single-step characterization of the complete proteome of a cell. This is the consequence of the large number of cellular proteins with varying levels of abundance and diverse p*I*(s), hydrophobicities and M_r(s). Therefore, subcellular fractionation is needed to reduce complexity and increase resolution of proteomic experiments (for review, see [46]). As elegantly stated, "plant proteomics exemplifies perfectly this functional dimension with recent explosion of proteomic initiatives, which are more and more focused on the analysis of subcellular compartments" [12,13]. Moreover, it allows the characterization of individual organelle proteomes, and to know protein location in relation to their function. Cell wall [47], chloroplast [48], endoplasmic reticulum [49,50], mitochondria [51–53], nucleus [54], peroxisome [55], and PM [40,56] have been isolated for creating their proteomes. For details on the individual isolation methods, the readers are referred to original articles [40,47–56].

2.2. Techniques

2.2.1. One- and two-dimensional gel electrophoresis

One- (1) and 2-DGE separate complex protein mixtures on the basis of their molecular masses of approximately 10–300 kDa (10^2–10^5 molecular masses) and by charge in the first dimension (IEF) and molecular masses in the second dimension (SDS–polyacrylamide gel electrophoresis, SDS–PAGE), respectively. For protein visualization, though Coomassie brilliant blue (CBB), colloidal CBB, and silver staining are the most commonly used stains in plant proteomic studies, fluorescence-based protein detection methods have recently begun to surpass these conventional stains, due to quantitative accuracy, detection sensitivity, and compatibility with modern downstream protein identification and characterization procedures, such as MS. 2-DGE established in the 1970s [57–59], is a popular technique and commonly used by the average proteomics researcher for high quality protein resolution and dynamic range. "It is cost-effective, affordable, and accessible to labs world-wide" [60,61]. Introduction of immobilized pH gradient (IPG) strips further increased the utility of 2-DGE in the proteomic era [62,63]. Separation of protein complexes by 1-DGE in the native state allows one to couple it with SDS–PAGE in the second dimension or 2-DGE (IEF and SDS–PAGE), now referred to as three-dimensional gel electrophoresis (3-DGE [64–68]). Importantly, the 2-DGE profiles provide us with a "proteome signature", and is one of the key technologies [61,63,69–71] in the proteomics workflow (Fig. 4).

2.2.2. Amino acid sequencing

The usual and most widely used approach for obtaining primary sequence analysis is to sequence the amino-terminal (N-terminus or Edman reaction) of the intact protein, by which the order of amino acids of proteins or peptides is determined. For this, proteins must be transferred onto polyvinylidene difluoride (PVDF) membranes from gels (Fig. 4). The Edman sequencing is however slow (typically no more than one to two proteins maybe identified per day) and requires rather large sample amounts, which restricts analysis to the most abundant proteins. Moreover, many plant proteins are blocked at the N-terminus probably most likely due to carrier ampholyte gels not washed prior to use (in contrast to pre-washed IPG gels), which means that contaminants can remain in the gel and may modify proteins [72]. In case the N-terminus is blocked, deblocking is necessary by means of enzymatic treatment, such as use of pyroglutamate amino peptidase, or fragmentation of the protein by enzymatic (trypsin) and/or chemical (cyanogen bromide) methods [73,74]. However, internal amino acid sequencing becomes even more laborious and requires even larger protein amounts than needed for N-terminal Edman sequencing. On the other hand, carboxy-terminal (C-terminus) is used for the direct confirmation of the C-terminal sequence of native and expressed proteins, for detection and characterization of protein processing at the C-terminus, for identification of post-translation proteolytic cleavages, and for obtaining partial sequence information on N-terminally blocked protein samples.

2.2.3. Metal affinity shift assay

Although, various methods are available for detecting metal binding proteins, most of them generally require a purified or semi-purified target of interest, and do not facilitate identification of unknown targets from complex protein mixtures. To this end, Kameshita and Fujisawa developed a method—metal affinity shift assay—suitable for identification of divalent metal cation binding proteins [75]. The method is based on the simple principle that binding of metal ions to proteins changes both the charges characteristics and the confirmation of proteins altering mobility during electrophoresis. The usefulness of this method has been demonstrated in the analysis of the *Arabidopsis* mitochondrial proteome [76].

2.2.4. Mass spectrometry

MS-based techniques play important roles in this proteomic era and are most commonly used to identify proteins either separated or visualized on 2-DGE gels or directly from complex pepetide mixtures (Fig. 4; [60,77–80]). The MS techniques includes matrix-assisted laser desorption ionization (MALDI)-time-of-flight (TOF)-MS, electrospray ionisation-MS (ESI-MS), ESI tandem-MS (ESI-MS/MS), ESI-quadrupole-TOF-MS (ESI-Q-TOF-MS), reverse phase-high performance liquid chromatography ESI-MS (RP-HPLC-ESI-MS), gas chromatography-TOF-tandem

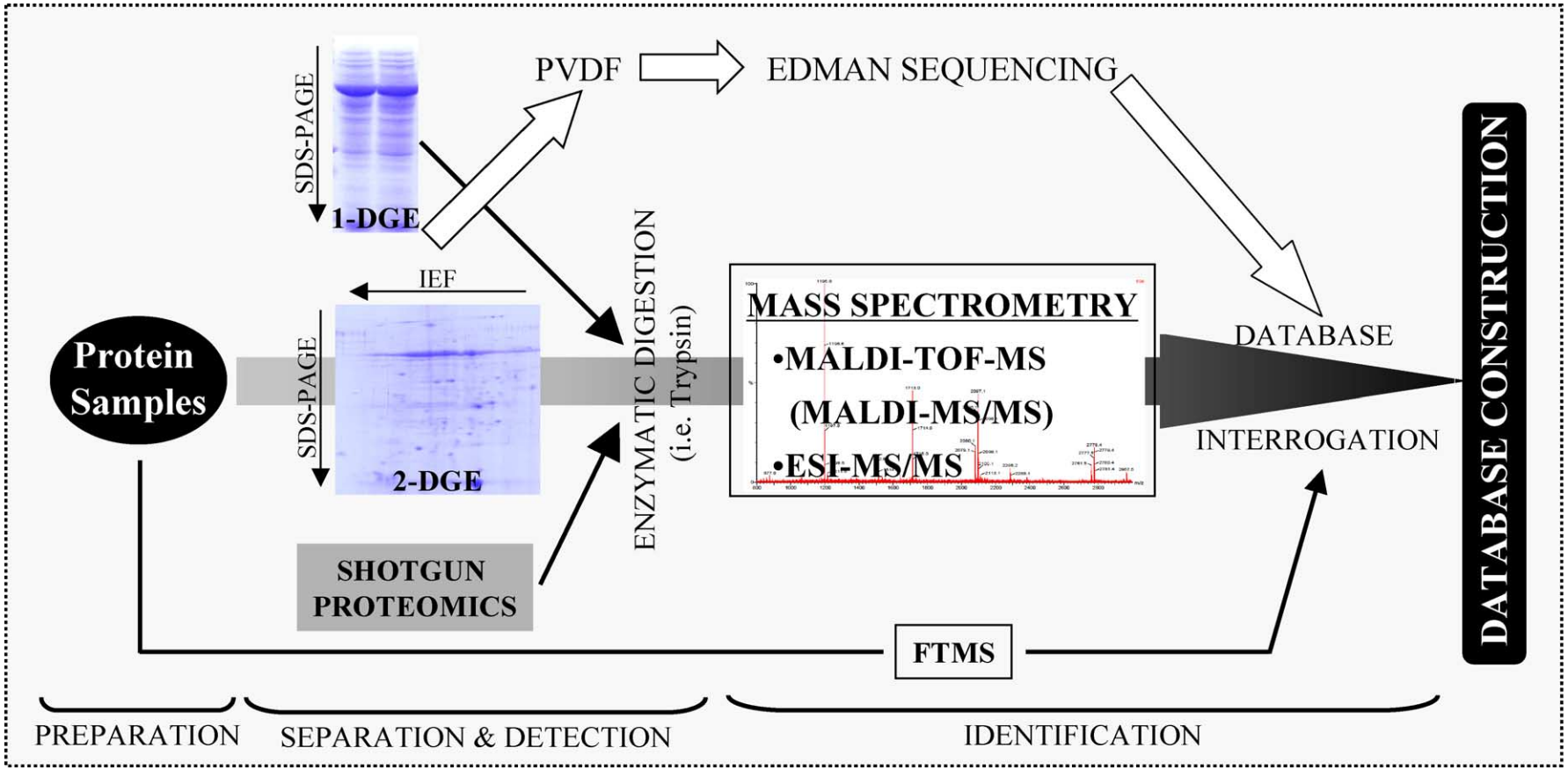

Fig. 4. Schematic presentation of the proteomics workflow in dicot plants. Protein samples can be processed either through gel-based (1- and 2-DGE) approaches or shotgun proteomics followed by mass spectrometric analyses and database interrogation to identify the proteins. Complex protein mixtures can be also directly subjected to FTMS without extensive purification and digestion. 1- and 2-DGE separated proteins upon transfer to PVDF membrane can also be identified by Edman sequencing. The final step is the database construction, once proteins are unambiguously identified.

MS (GC-TOF-MS), and the recently developed "top-down" Fourier transform MS (FTMS). FTMS has uniquely valuable attributes over the commonly used "bottom up" MS ([81–83], and reviewed in [84,85]). The MS techniques, MALDI-TOF-MS, which is normally used to analyze relatively simple peptide mixtures, and ESI-MS/MS, where complex protein samples are analyzed, have their own advantages and disadvantages. For MALDI-TOF-MS, the instrument is user-friendly and robust, compatible with new robotic sample preparation devices designated to aid high-throughput proteomics, has improved accuracy and resolution, and high sensitivity in the fentomolar to attomolar range. Moreover, sample preparation can be performed more easily for MALDI-TOF-MS, as it is more tolerant towards low amounts of contaminants (salts and low molecular weight chemicals) from the biological environments or buffers. The biggest disadvantages of the MALDI-generated peptide mass fingerprints (PMFs) are ambiguity in protein identification, mainly because of peptide mass redundancy, and that it requires relatively pure protein samples for the analysis. On the other hand, ESI-MS/MS is quite amenable for complex protein mixtures, where the total protein extract after proteolytic digestion can be directly subjected to MS. Characteristically, ESI results in multiply charged ions, effectively lowering the *m*/*z* values. Using a nanorange LC separation of proteins prior to MS, very pure samples can be obtained, an advantage over MALDI-TOF-MS. Moreover, the low flow rates possible with nanospray ionization reduce the amount of sample consumed and increase the time available for analysis. Recently, an alternative technology, termed multidimensional protein identification technology (MudPIT), has been developed that allows automated analyses of peptide mixtures generated from complex protein samples [86–88]. Moreover, isotope-coded affinity tag (ICAT) in the LC–MS/MS system has also emerged as a tool for quantitative proteomics [89].

Homology-based identification with MS data is possible but generally requires a large amount of experimentally determined protein sequence tags or PMFs. One limitation of MS is the inability to match the major fraction of the spectra to predicted open reading frames. Two possible factors might be attributed to it. First, post-translational modifications (PTMs) of proteins result in alteration of apparent masses of peptides that differ to the database (DB) entries. Although not perfect, such modifications can be recognized by MS/MS-based analysis of peptide sequences. Second the lack of accurate determination of intron-exon boundaries of individual genes in the genome. "To this regard, full-length cDNAs are desperately needed". It should be noted that most of the above mentioned MS-based techniques are in common practice in plant proteomics, except for RP-HPLC-ESI-MS and FTMS, which are recently getting more attention in proteomic analysis of certain dicot plant materials, such as the chloroplast proteome. These two techniques have been mentioned in some detail.

2.2.4.1. Reverse phase-high performance liquid chromatography electrospray ionization mass spectrometry. RP-HPLC-ESI-MS has evolved into a highly powerful tool for accurate mass measurement of proteins (Fig. 4; [90–94]). Application of this technology was demonstrated on photosystem I/II (PSI and PSII) proteins of various plant species [95–100]. One important benefit of this technique is that, identification of proteins from plants whose genome sequence information is not available can be done by comparison with the mass range expected from the known genes of other plant species. Moreover, the HPLC methods used in these studies were found to be highly reproducible, and it was suggested that the chromatograms might serve as a highly confident fingerprint for comparison within a single and among different species for future studies of PSI and PSII. This technique provides an attractive means to monitor physiological changes in covalent status across the entire complement of thylakoid proteins and in subfractions from different membrane domains as a function of light and other stresses, providing significant benefits to functional genomics.

2.2.4.2. Fourier transform mass spectrometry. In this method, the protein mixture, without extensive purification and digestion, is introduced directly into the FTMS instrument using ESI (Fig. 4). This approach was applied to find accurate ($\pm$1 Da) M_r values for 22 proteins and to identify and characterize 7 proteins, all from the three soluble proteomes (thylakoid peripheral, thylakoid lumen, and stroma) of the chloroplast of *Arabidopsis* (ecotype Columbia) [101] identified previously by conventional MS [102,103]. However, the authors conclude that the "bottom up" methodology remains the better choice for the first identification of the precursor from a genome, such as *Arabidopsis*, mainly due to full automation [101]. However, the future use of "top down" MS in plant proteomics depends on development of new automation methods for sample preparation, MS, and data analysis [101,104,105].

2.2.5. *Affinity chromatography*

Affinity-based chromatography (Fig. 5) is a powerful protein separation method based on the specific interaction between immobilized ligands and target proteins (for review, see [106]). "This technique can be used reduce the complexity of protein or peptide mixtures as a part of the traditional 2-DGE in conjunction with N-terminal sequencing and MS, or in the identification of protein–protein interactions in combination with MS" [106,107]. Several affinity ligands have been used for a variety of target proteins, helping in generating proteome maps. An added benefit is that it can enrich low abundance proteins, identify protein–protein interactions, and find the cellular location of proteins. In the future, more applications of affinity-based purification can be expected, including increasing the resolution in 2-DGE, improving the sensitivity of MS quantification, and incorporating purification as part of the MudPIT.

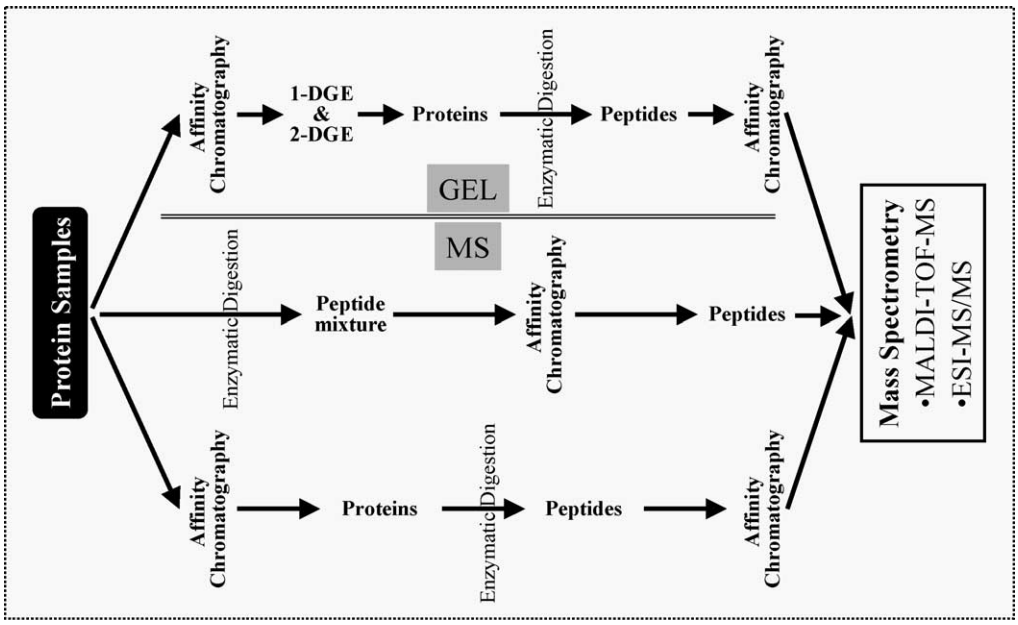

Fig. 5. Affinity chromatography in plant proteomics. The affinity chromatography can be employed in both the gel- and MS-based approaches. In the gel-based approach, it can be used prior to and/or after 1- and 2-DGE. In the MS-based approach, it can be used after the enzymatic digestion of the protein samples, or protein samples can be affinity purified followed by enyzymatic digestion and affinity chromatography of the resulting peptides. In all cases, mass spectrometry is used for protein identification.

2.2.5.1. Immobilized metal affinity chromatography. Specific capture of phosphopeptides is possible by affinity chromatography with immobilized metal ions (IMAC)—a simple technique used in several phosphoproteomic studies, and having great promise for large-scale studies ([108,109], and recently reviewed in [110,111]). The conventional IMAC procedures, with either Fe(III) [112] or Ga(III) [113] suffer from non-specific binding of peptides containing multiple carboxylic acid groups. To overcome this problem, a slight modification was done, which involves the conversion of carboxylic acids to methyl esters, thereby allowing enrichment of the phosphorylated peptides [114]. The modified IMAC technique was successfully used to enrich the phosphopeptides from the tryptic thylakoid peptides, isolated from *Arabidopsis* (ecotype Wassilewskija-2), and sequenced using ESI-Q-TOF-MS/MS [115]. Results revealed the identification of three new phosphopeptides in addition to the five known phosphorylation sites in PSII proteins. All phosphopeptides are found phosphorylated at threonine residues implementing a strict threonine specificity of the thylakoid kinases. The finding of these novel phosphoproteins extends involvement of the redox-regulated protein phosphorylation in photosynthetic membranes beyond the PSII and its light-harvesting antennae. Additionally, Nuhse and co-workers investigated the potential of IMAC in combination with LC-Q-TOF-MS/MS for identification of over 75% pure PM phosphoproteins of *Arabidopsis* [116]. Using a strong anion exchange chromatography prior to IMAC, they could be able to decrease the complexity of IMAC-purified phosphopeptides and resulted in far greater yields of monophosphorylated peptides. The IMAC technique used for phosphoprotein analyses in these studies has been schematically depicted (Fig. 6).

2.2.5.2. Thiol affinity chromatography. Redox regulation is important in many biological processes, and the identification of targets for thioredoxin (Trx) and glutaredoxin is of considerable interest in plant biology. Hence, it becomes essential to isolate and characterize the proteins possessing redox-regulated cysteine residues. An easy, robust and comprehensive method involving thiol affinity chromatography technique was used to define the "plant disulfide proteome" using *Arabidopsis* (ecotype Columbia) as a model [117]. Stem and leaf tissues were used for extracting proteins. In this method, free thiols in proteins are fully blocked by alkylation, following which disulfide cysteines are converted to sulfhydryl groups by reduction. Finally, proteins with sulfhydryls are isolated by thiol-affinity (thiol-Sepharose 4B) chromatography (for details see [117]). The method is unique in the sense that membrane as well as water-soluble proteins can be examined for their disulfide nature. Use of this method resulted in the identification of 65 putative disulphide proteins, including 20 novel proteins with function in redox regulation, such as violaxanthin de-epoxidase, two oxygen-evolving enhancer proteins, carbonic anhydrase, PSI reaction center subunit N, PSI subunit III, *S*-adenosyl-L-methionine carboxyl methyltransferase, guanylate kinase, and bacterial mutT homolog. Interestingly among these novel proteins, only a single protein, carbonic anhydrase, was found in the list of recently identified Trx-target proteins [117]. This might be due to the fact that SDS/Tris buffer was used to solubilize proteins from stem and leaf, whereas Balmer

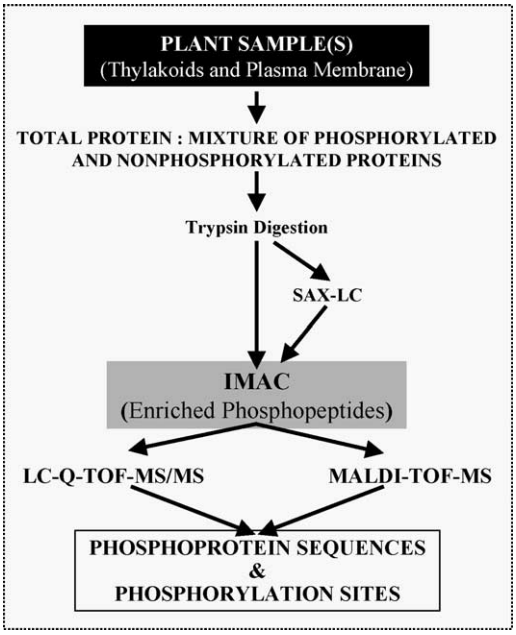

Fig. 6. IMAC in plant phosphoproteomics. Total proteins, isolated from a plant sample, are usually trypsin digested, before being subjected to IMAC. Alternatively, a strong anion exchanger liquid chromatography (SAX-LC) can be used prior to IMAC, to decreases the complexity of IMAC-purified phosphopeptides and to yield a far greater coverage of monophosphorylated peptides. The phosphoprotein sequences and the phosphorylation sites are determined using LC-Q-TOF-MS/MS, whereas the protein MS spetrum data can be obtained using MALDI-TOF-MS.

and co-workers [118] targeted water-soluble stromal proteins only.

2.2.5.3. Thioredoxin affinity chromatography. Trx affinity chromatography has also been used to capture cytosolic Trx(s) from cell lysate of dark-grown *Arabidopsis* whole tissues [119]. For this, a mutant of cytosolic Trx, in which an internal cysteine at the active site was substituted with serine, was immobilized on CNBr activated resin. Identified proteins were found to be involved in anti-oxidative stress response, protein biosynthesis and degradation, metabolic pathways and chloroplast.

2.2.6. ^{15}N-isotope labeling

Structural proteomics play an important role in understanding protein–protein interactions in living systems at a molecular level [120]. In vivo nuclear magnetic resonance (NMR) spectroscopy and high-resolution solution NMR spectroscopy are two powerful tools for studying intermolecular interactions in complex environments. The latter technique is useful in determining detailed conformational changes in individual proteins. A major prerequisite for NMR applications in proteomics is the need to isotopically label proteins with ^{15}N- and/or ^{13}C-isotopes to achieve spectral dispersion of chemical shifts in multiple dimensions and to obtain a higher intrinsic sensitivity for detection. Using potato (*Solanum tuberosum* L.) cultivar Elkana as a model, it was demonstrated that it is possible to uniformly label (>98%)

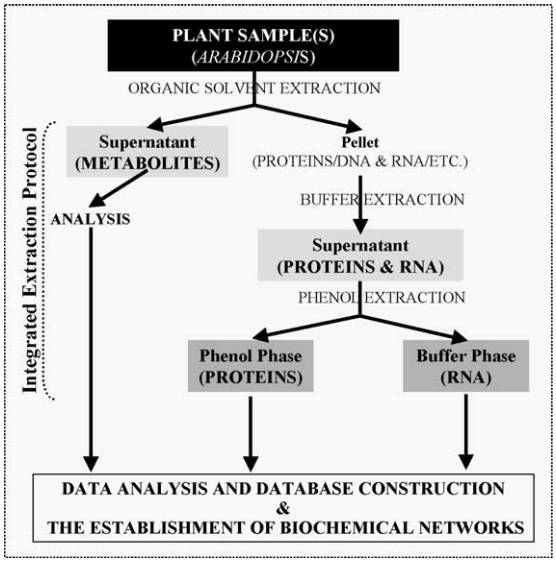

Fig. 7. Integrated extraction protocol for plant metabolites, proteins and RNA. Sequential extraction from the same sample is convenient, easy to replicate, and overcomes the question of inherent biological variation of independent samples.

a plant with ^{15}N-isotope ([^{15}N]-nitrate as the sole labeling source) for use in proteomics [121]. The success of the labeling procedure was demonstrated by the NMR results obtained on the complete proteome of potato sap and on an isolated protease inhibitor. The advantage of such in vivo labeling of higher organisms is that all constituting proteins are labeled and become available as functional, post-translationally modified, correctly folded proteins.

2.2.7. Integrated extraction procedure for metabolites, proteins and RNA

A novel extraction protocol has been recently described with which metabolites, protein and RNA can be sequentially extracted from the same sample, thereby providing a convenient procedure for the analysis of replicates as well as exploiting inherent biological variation of independent samples for multivariate data analysis (Fig. 7; [122]). Using 30–100 mg of *Arabidopsis* (ecotype Columbia) leaf as the source material, a total of 652 metabolites (GC coupled to TOF-MS), 297 proteins (2-D LC coupled to MS, LCQ DecaXP ion trap MS/MS), and clear RNA bands (Northern analysis) were validated. A subset of the most abundant proteins and metabolites from replicate analysis of different *Arabidopsis* accessions was merged to form an integrative dataset allowing both classification of different genotypes and the unbiased analysis of the hierarchical organization of proteins and metabolites with a real biochemical network. A similar strategy was taken to isolate proteins (proteome) and mRNA (transcriptome) simultaneously from single root samples from *M. truncatula* genotype J5 (Jemalong) inoculated with or without the arbuscular mychorrhizal fungus, *Glomus mosseae* (Nicol. and Gerd.) Gerdemann and Trappe (isolate BEG 12) [123].

3. Bioinformatics

Bioinformatics is the next essential tool to link the proteome to its genome. Recently, a comparative proteomics resource database of *Arabidopsis* proteins has been developed and is available at "Proteins of *Arabidopsis thaliana* (PAT) Database [http://www.pat.sdsc.edu/]" [124]. In this database, an integrative genome annotation pipeline (iGAP) for proteome-wide protein structure and functional domain assignment has been used. This database serves the *Arabidopsis* and plant proteomics community through the provision of structure and functional assignment to all identified proteins in the *Arabidopsis* genome. In addition, the MIPS (Institute for Bioinformatics) *Arabidopsis* database MatDB (http://www.mips.gsf.de/proj/thal/db) was also constructed with an aim to provide comprehensive resource for *Arabidopsis* as a genome model that serves as a primary reference for research in plants and is suitable for transfer of knowledge to other plants, especially crops [125]. On the other hand, proteome profiles/representative 2-DGE gel images for stems, leaves, seedpods, roots, flowers, tissues, and suspension cell cultures are available for *M. truncatula* (http://www.noble.org/medicago/ and Website/2DPAGE/search.asp). Recently, a relational database system, called DOME, has been developed for *M. truncatula* functional genomics and bioinformatics (http://www.medicago.vbi.vt.edu/dome.html). Most of the bioinformatics tools required for the proteomic analysis are available from the ExPASy (www server, http://www.us.expasy.org/).

The usefulness and importance of bioinformatics in proteome research was demonstrated in a comparative proteome bioinformatics study in *Arabidopsis*. A whole complement of putative protein tyrosine kinases in *Arabidopsis* was identified using slightly degenerate PROSITE (http://www.expasy.ch/prosite) protein serine-threonine kinases (PS00108) and protein tyrosine kinase (PS00109) signatures as sequence probes [126]. Furthermore, based on immunological evidence, it was found that the number of *Arabidopsis* proteins specifically phosphorylated on tyrosine residues is much higher than in yeast.

Recently, the universal protein knowledgebase (UniProt) consortium has been established in which major protein databases; Swiss-Prot, TrEMBL and PIR have joined together to handle the increasing volume and variety of protein sequences and functional information [127]. This will be of great benefit for scientist actively involved in modern plant biology.

4. Conclusions and perspectives

Genomic information has been a prerequisite for the developments of "OMICS", including proteomics. Therefore, the dicot model *A. thaliana*, whose genome has been completely sequenced and annotated, and *M. truncatula*, which is the focus of current genomic projects, have been the plant materials of choice for proteomic studies in dicots, including those involving the development/establishment of new techniques/technologies. Proteomic technologies have been evolving over the past 20 years. During the past few years, use, adaptation and refinements of these developing technologies have led to remarkable achievements in large-scale protein separation by 2-DGE and their analysis by high-

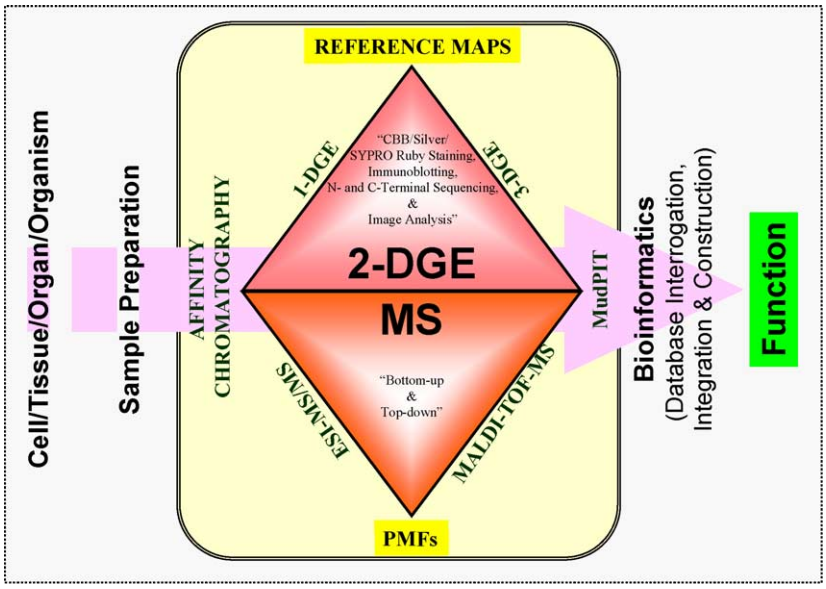

Fig. 8. Schematic illustration of technologies in proteomics of dicot plant materials. 2-DGE and MS are two core techniques involved in the separation of protein samples, prepared from a variety of plant materials (cell, tissue, organ, organism), and identification using both the Edman sequencing and MS (MALDI-TOF-MS, ESI-MS/MS), followed by database interrogation to assign protein function. Detection of protein spots on 2-DGE is usually carried out using organic (CBB, silver) and fluorescent stains (SYPRO Ruby). The use of affinity chromatography in protein enrichment is now gaining prominence in proteome analysis. Database integration and construction of the outcoming data (2-DGE reference maps and PMFs) form a valuable resource for the plant proteomics community.

throughput MS-based techniques in plants (Fig. 8). A number of techniques available to date are a clear indication of impressive progress in dicot plant proteomics (summarized and discussed in [25,26]), and as exemplified by the increasing use of affinity chromatography-based protein purification as an integral part of Trx proteomics. The field of MS has been transformed into a key technology in proteome research, due to increased sensitivity, more efficient ionization techniques and better detectors [128]. "Not only, are the protein spots of 2-DGE separated samples quantitatively sufficient for unequivocal protein identification, but it is also possible to take a closer look at PTMs. For example, it has been speculated that modifications like phosphorylation or glycosylation exist on every second protein, which are essential for the protein function". In addition, the need for a faster and more efficient method to identify frequently observed proteins on 2-DGE gels resulted in the use of PMFs, which laid the foundation for high-throughput, high-sensitivity methods in proteomics [129]. We discuss below the immediate challenges and the future directions needed to take the art of protein separation, detection, and identification to its full potential by the dicot (and plant) proteomic community.

(1) Identification of total proteins: current sample preparation methods have been successful in the identification of a large number of proteins; however, limitations still exist for the extraction/isolation of low abundance proteins, membrane and hydrophobic proteins, organelle proteins, and basic proteins. Thus, the focus should be on further improvements in sequential solubilization/fractionation by the use of more powerful solubilizing buffers or their combinations, including the use of new detergents. Moreover, the use of non-gel based chromatographic protein separation methods, including affinity chromatography, will also help to improve the separation of proteins from different samples and further the experimentally identifiable proteome. On the other hand, by constructing and comparing the proteome across developmental stages and following environmental stimuli, it will be possible to identify even more proteins.

(2) 2-DGE reference maps and PMF databases: databases created on 2-DGE and PMFs are the platforms for comparing proteomes. 2-DGE maps of different plant species have been generated and will be needed to obtain a better insight into the interspecies correlation of 2-DGE protein patterns. As for example, 2-DGE protein profiles of chloroplast envelope proteins of *Arabidopsis* and spinach, obtained using the same chloroform/methanol extraction method [130], revealed the presence of 15 and 20 proteins exclusively in *Arabidopsis* and spinach, respectively [131], suggesting that several plant models may be required to identify chloroplast envelope proteins. This will also be true for other proteomic studies. Furthermore, comparing the 2-DGE reference maps among wild type and defined genetic mutants, biotically and abiotically challenged plants and/or environmental stimuli, will help for a more complete understanding of protein function. The outcome will result in an integrated global view of protein networks. However, for this, a consistent proteomic approach is necessary to establish good proteome reference maps on the World Wide Web. This will allow for constant updates and comparison of proteomes and its associated data with other studies. But, the question remains, "how far it is possible world-wide".

(3) PTMs and isoforms: "sequence analysis of proteins and peptides is not limited to the elucidation of the primary structure of a protein, and therefore the analysis of PTMs is an important task of protein chemistry in proteome research" [132]. PTMs generate tremendous diversity, complexity and heterogeneity of gene products [133]. Therefore, it is expected that for a given plant species, the number of proteins will exceed manifold the number of genes. Phosphorylation of proteins is an important PTM, and identification of phosphoproteins are needed to characterize entire phosphorylation cascades involved in a broad range of biological function(s). As stated by Peck and co-workers, "it will be necessary to develop the in planta phospholabeling assay such that the analysis of phosphoproteins can be performed using whole plants and not just cell cultures" [134]. Recently, two new commercially available products, namely the "PhosphoProtein Purification Kit" (Qiagen, Cat. No. 37101) and the "Pro Q Diamond phosphoprotein gel stain" (Molecular Probes, Cat. No. P-33356) have given new impetus to phosphoproteomics. The PhosphoProtein Purification Kit is based on an affinity chromatography process and provides complete separation of phosphorylated and unphosphorylated proteins from a cell lysate, and therefore facilitating investigation of the phosphorylation status of both entire cells and specific proteins. The "Pro Q Diamond phosphoprotein gel stain" is a breakthrough technology that provides a simple method for selectively staining phosphoproteins directly in polyacrylamide gels [135]. Here, we demonstrate the validity of the Pro Q Diamond stain in plant phosphoproteomics, using total proteins isolated from germinating *Brassica napus* seeds as an example (Fig. 9). The visualization of a large number of phosphoproteins, followed by staining of the total proteins using by SYPRO Ruby dye, on 2-DGE gels, is a first such demonstration of the power of the phosphoprotein gel stain in plants. Isoforms can represent post-translationally modified forms of the same protein or could be translated from the same gene (spliced variants) or different genes from multigene families. A combination of DNA- and protein-based functional approaches, including radioisotope labeling, immunoassaying with specific antibodies, affinity enrichment/tagging, mass tagging and MS, will be powerful tools to test the correlation between gene transcription and translation, allowing for the identification of PTMs and an estimation of the number of proteins translated

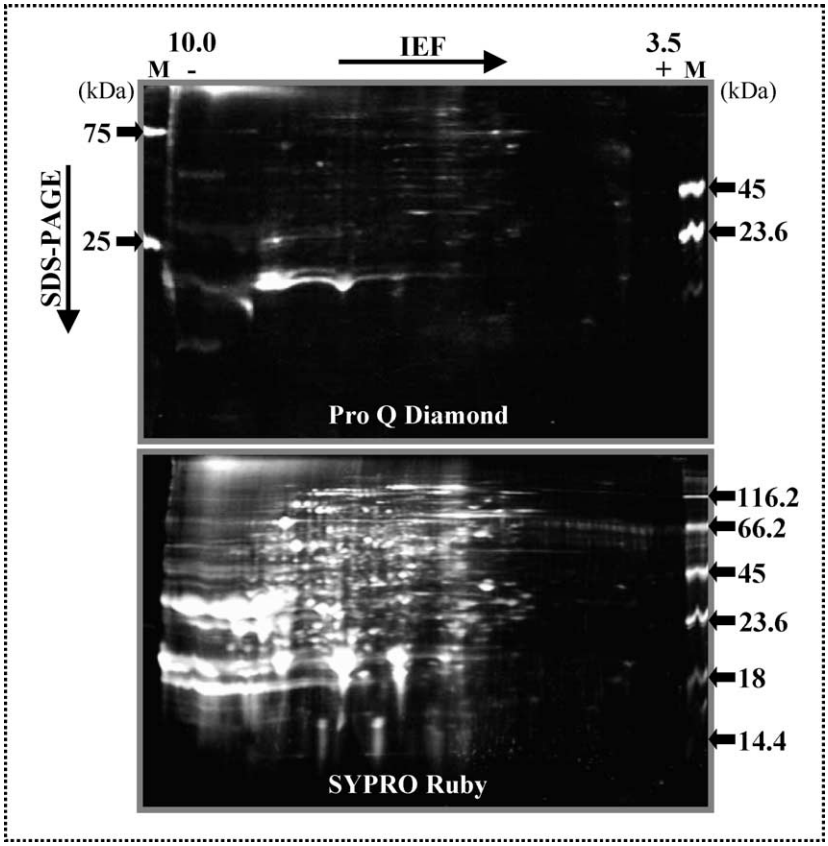

Fig. 9. Visualization of phosphoproteins by Pro Q Diamond dye. *Brassica napus* seeds were germinated in the dark on a wet Kimwipe tissue (in covered petri dish) placed in a growth incubator set at 25 °C. After 12 h, the seeds were removed and homogenized in lysis buffer (urea/thiourea buffer containing Tris–HCl/Trizma base). Approximately 300 μg total protein extract was separated by 2-DGE (for details on methods, see [137]), and the separated proteins were visualized using fluorescent stains, Pro Q Diamond (for phosphoproteins) and SYPRO Ruby (for total protein) using a UV-transilluminator (ATTO, Tokyo Japan). M: molecular mass markers.

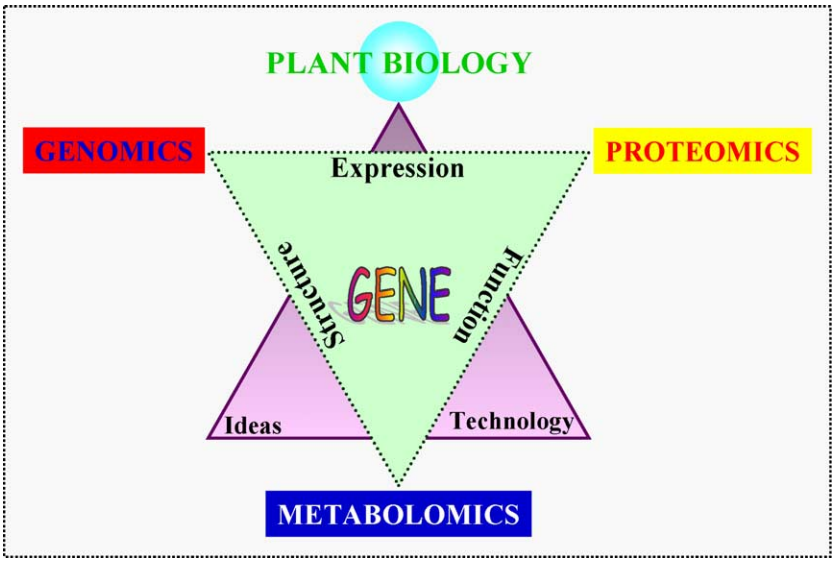

Fig. 10. In our quest towards understanding plant biology: proteomics in plant functional genomics. The perfect union between these "omic" technologies, creative ideas and advances in technology, will be essential in our quest for the "holy grail" of plant biology—defining the function and interaction of each and every gene in the genome.

from the same/different genes (isoforms). Their determination remains a big challenge.

(4) Construction of prediction programs for plants: to date, the available prediction programs are largely based on non-plant sequences, and thus the robustness of these tools for prediction in plants remains uncertain. A comprehensive proteomic study in at least one model plant species, such as *Arabidopsis* or rice (and/or *M. truncatula*) will certainly assist in constructing efficient prediction programs suitable for plants.

(5) Robotic 2-DGE and automation workflow: the existing 2-DGE technologies are still time-consuming, laborious and messy. To overcome these problems, and to facilitate improved protein resolution and reproducibility, NextGen Sciences Ltd., UK (http://www.nextgensciences.com) have designed and developed an automated 2-DGE proteomic—"inject sample and walk away"—system called a2DE. Once available, this automated system will greatly aid 2-DGE proteomic analyses. Furthermore, as the frequency and scope of proteomics increases, and to minimize sample contamination and facilitate sample processing and identification, a workflow employing robotic automation starting after 2-DGE is also undoubtedly needed for high-throughput proteomics. The study on high-throughput PMFs of soybean seed proteins is an excellent example, where automated workflow and utility of UniGene EST databases for protein identification was demonstrated [136].

Finally, and as aforementioned (Section 1), functional genomics is undoubtedly the next frontier, which will require systematic and multi-parallel high-throughput approaches in the analyses and deduction of gene (one or many) function (Fig. 10). Proteomics is one of the pillars of functional genomics, and will only be most useful when combined with other "omics" pillars, on which the functional genomics stands. Collective information from these pillars will answer the most fundamental question on how biological systems work and interact in their environment.

5. Nomenclature

CBB	Coomassie brilliant blue
ESI-MS/MS	electrospray ionization tandem-mass spectrometry
ESI-Q-TOF-MS	ESI-quadrupole-time-of-flight tandem MS
FTMS	top-down Fourier transform MS
GC-TOF-MS	gas chromatography-TOF-tandem MS
HPLC	high performance liquid chromatography
IEF	isoelectric focusing
LC–MS/MS	LC-tandem MS
MALDI	matrix-assisted laser desorption/ionization
MALDI-TOF-MS	matrix-assisted laser desorption ionization-time-of-flight-mass spectroscopy
M_r	molecular mass
MS	mass spectrometry
NMR	nuclear magnetic resonance
1-DGE	one-dimensional gel electrophoresis
PAGE	polyacrylamide gel electrophoresis
p*I*	isoelectric point
PMF	peptide mass fingerprinting
PTM	post-translational modification
PROTEOME	PROTEins expressed by a genOME
SDS	sodium dodecyl sulfate
3-DGE	three-dimensional gel electrophoresis
TCA	trichloroacetic acid
2-DGE	two-dimensional gel electrophoresis

Acknowledgements

We are grateful to Ms. Junko Shibato (HSS, AIST) for her help during the preparation of this manuscript. We also thank Prof. Richard Oliver (Murdoch University, Australia) for allowing us to use the photograph of the *Medicago* plant, and Dr. Jay J. Thelen, Associate Director, Proteomics Center, University of Missouri-Columbia, MO, USA, for providing the *B. napus* seeds. Authors highly appreciate the vision and collaboration established with Drs. Vishwanath Prasad Agrawal (RLABB, a private non-profit research organization, Kathmandu, Nepal), Masami Yonekura (School of Agriculture, Ibaraki University, Ami, Japan), Shigeru Tamogami (Akita Prefectural University, Akita, Japan), Akihiro Kubo (National Institute for Environmental Studies, Tsukuba, Japan), and Nam-Soo Jwa (Sejong University, Seoul, Korea), in proteomics of plant materials, including the dicot model *Arabidopsis*.

References

[1] The Arabidopsis Genome Initiative, Nature 408 (2000) 796.
[2] A. Bernal, U. Ear, N. Kyrpides, Nucleic Acids Res. 29 (2001) 126.
[3] S.A. Goff, D. Ricke, T.-H. Lan, G. Presting, R. Wang, M. Dunn, J. Glazebrook, A. Sessions, P. Oeller, H. Varma, D. Hadley, D. Hutchison, C. Martin, F. Katagiri, B.M. Lange, T. Moughamer, Y. Xia, P. Budworth, J. Zhong, T. Miguel, U. Paszkowski, S. Zhang, M. Colbert, W.L. Sun, L. Chen, B. Cooper, S. Park, T.C. Wood, L. Mao, P. Quail, R. Wing, R. Dean, Y. Yu, A. Zharkikh, R. Shen, S. Sahasrabudhe, A. Thomas, R. Cannings, A. Gutin, D. Pruss, J. Reid, S. Tavtigian, J. Mitchell, G. Eldredge, T. Scholl, R.M. Miller, S. Bhatnagar, N. Adey, T. Rubano, N. Tusneem, R. Robinson, J. Feldhaus, T. Macalma, A. Oliphant, S. Briggs, Science 296 (2002) 92.
[4] J. Yu, S. Hu, J. Wang, G.K. Wong, S. Li, B. Liu, Y. Deng, L. Dai, Y. Zhou, X. Zhang, M. Cao, J. Liu, J. Sun, J. Tang, Y. Chen, X. Huang, W. Lin, C. Ye, W. Tong, L. Cong, J. Geng, Y. Han, L. Li, W. Li, G. Hu, X. Huang, W. Li, J. Li, Z. Liu, L. Li, J. Liu, Q. Qi, J. Liu, L. Li, T. Li, X. Wang, H. Lu, T. Wu, M. Zhu, P. Ni, H. Han, W. Dong, X. Ren, X. Feng, P. Cui, X. Li, H. Wang, X. Xu, W. Zhai, Z. Xu, J. Zhang, S. He, J. Zhang, J. Xu, K. Zhang, X.

Zheng, J. Dong, W. Zeng, L. Tao, J. Ye, J. Tan, X. Ren, X. Chen, J. He, D. Liu, W. Tian, C. Tian, H. Xia, Q. Bao, G. Li, H. Gao, T. Cao, J. Wang, W. Zhao, P. Li, W. Chen, X. Wang, Y. Zhang, J. Hu, J. Wang, S. Liu, J. Yang, G. Zhang, Y. Xiong, Z. Li, L. Mao, C. Zhou, Z. Zhu, R. Chen, B. Hao, W. Zheng, S. Chen, W. Guo, G. Li, S. Liu, M. Tao, J. Wang, L. Zhu, L. Yuan, H. Yang, Science 296 (2002) 79.
[5] S. Kikuchi, K. Satoh, T. Nagata, N. Kawagashira, K. Doi, N. Kishimoto, J. Yazaki, M. Ishikawa, H. Yamada, H. Ooka, I. Hotta, K. Kojima, T. Namiki, E. Ohneda, W. Yahagi, K. Suzuki, C.J. Li, K. Ohtsuki, T. Shishiki, Y. Otomo, K. Murakami, Y. Iida, S. Sugano, T. Fujimura, Y. Suzuki, Y. Tsunoda, T. Kurosaki, T. Kodama, H. Masuda, M. Kobayashi, Q. Xie, M. Lu, R. Narikawa, A. Sugiyama, K. Mizuno, S. Yokomizo, J. Niikura, R. Ikeda, J. Ishibiki, M. Kawamata, A. Yoshimura, J. Miura, T. Kusumegi, M. Oka, R. Ryu, M. Ueda, K. Matsubara, J. Kawai, P. Carninci, J. Adachi, K. Aizawa, T. Arakawa, S. Fukuda, A. Hara, W. Hashizume, N. Hayatsu, K. Imotani, Y. Ishii, M. Itoh, I. Kagawa, S. Kondo, H. Konno, A. Miyazaki, N. Osato, Y. Ota, R. Saito, D. Sasaki, K. Sato, K. Shibata, A. Shinagawa, T. Shiraki, M. Yoshino, Y. Hayashizaki, A. Yasunishi, Rice Full-Length cDNA Consortium, National Institute of Agrobiological Sciences Rice Full-Length cDNA Project Team, Foundation of Advancement of International Science Genome Sequencing & Analysis Group, RIKEN, Science 301 (2003) 376.
[6] A. Pandey, M. Mann, Nature 405 (2000) 837.
[7] S.D. Patterson, R.H. Abersold, Nat. Genet. Suppl. 33 (2003) 311.
[8] E. Phizicky, P.I. Bastiaens, H. Zhu, M. Snyder, S. Fields, Nature 422 (2003) 208.
[9] M.R. Wilkins, J.-C. Sanchez, A.A. Gooley, R.D. Appel, I. Humphrey-Smith, D.F. Hochstrasser, K.L. Williams, Biotechnol. Genet. Eng. Rev. 13 (1995) 19.
[10] H. Thiellement, N. Bahrman, C. Damerval, C. Plomion, M. Rossignol, V. Santoni, D. de Vienne, M. Zivy, Electrophoresis 20 (1999) 2013.
[11] M. Rossignol, Curr. Opin. Biotechnol. 12 (2001) 131.
[12] K.J. van Wijk, Trends Plant Sci. 5 (2000) 420.
[13] K.J. van Wijk, Plant Physiol. 126 (2001) 501.
[14] B. Kersten, L. Burkle, E.J. Kuhn, P. Giavalisco, Z. Konthur, A. Lueking, G. Walter, H. Eickhoff, U. Schneider, Plant Mol. Biol. 48 (2002) 133.
[15] J.K.M. Roberts, Plant Mol. Biol. 48 (2002) 143.
[16] U. Mathesius, N. Imin, S.H.A. Natera, B.G. Rolfe, Methods in Molecular Biology. Plant Functional GEnomics: Methods and Protocols, vol. 236, 2004, p. 395.
[17] O.K. Park, J. Biochem. Mol. Biol. 37 (2004) 133.
[18] F.M. Canovas, E. Dumas-Gaudot, G. Recorbet, J. Jorin, H.P. Mock, M. Rossignol, Proteomics 4 (2004) 285.
[19] B.S. Watson, V.S. Asirvathan, L. Wang, L.W. Summer, Plant Physiol. 131 (2003) 1104.
[20] P. Hilson, I. Small, M.T.R. Kuiper, Curr. Opin. Plant Biol. 6 (2003) 426.
[21] D.G. Barker, S. Bianchi, F. Blondon, Y. Dattee, Plant Mol. Biol. Rep. 8 (1990) 40.
[22] D.R. Cook, K. Vandenbosch, F.G. de Bruijn, T. Huguet, Plant Cell. 9 (1997) 275.
[23] D.R. Cook, Curr. Opin. Plant Biol. 2 (1999) 301.
[24] A.T. Trieu, S.H. Burleigh, I.V. Kardailsky, I.E. Maldonado-Mendoza, W.K. Versaw, L.A. Blaylock, H. Shin, T.J. Chiou, H. Katagi, G.R. Dewbre, D. Weigel, M. Harrison, Plant J. 22 (2000) 531.
[25] G.K. Agrawal, M. Yonekura, Y. Iwahashi, H. Iwahashi, R. Rakwal, J. Chromatogr. B 815 (2005) 125.
[26] G.K. Agrawal, M. Yonekura, Y. Iwahashi, H. Iwahashi, R. Rakwal, J. Chromatogr. B 815 (2005) 137.
[27] M.R. Wilkins, J.-C. Sanchez, K.L. Williams, D.F. Hochstrasser, Electrophoresis 17 (1996) 830.
[28] T. Rabilloud, Electrophoresis 17 (1996) 813.
[29] C. Damerval, M. Zivy, F. Granier, D. deVienne, in: A. Chrambach, M.J. Dunn, B.J. Radola (Eds.), Adavnces in Electrophoresis, VCH, Weinheim, 1998, p. 263.
[30] D.I. Jacobs, R. van der Heijden, R. Verpoorte, Phytochem. Anal. 11 (2000) 277.
[31] F. Granier, Electrophoresis 9 (1988) 712.
[32] W. Weiss, W. Postel, A. Gorg, Electrophoresis 13 (1992) 770.
[33] M.P. Molloy, B.R. Herbert, B.J. Walsh, M.I. Tyler, M. Traini, J.C. Sanchez, D.F. Hochstrasser, K.L. Williams, A.A. Gooley, Electrophoresis 18 (1998) 837.
[34] D.I. Jacobs, M.S. van Rijssen, R. van der Heijden, R. Verpoorte, Proteomics 1 (2001) 1345.
[35] W.J. Hurkman, C.K. Tanaka, Plant Physiol. 81 (1986) 802.
[36] P. Giavalisco, E. Nordhoff, H. Lehrach, J. Gobom, J. Klose, Electrophoresis 24 (2003) 207.
[37] W. Wang, M. Scali, R. Vignani, A. Spadafora, E. Sensi, S. Mazzuca, M. Cresti, Electrophoresis 24 (2003) 2369.
[38] V. Santoni, T. Rabilloud, P. Doumas, D. Rouquie, M. Manison, S. Kieffer, J. Garin, M. Rossignol, Electrophoresis 20 (1999) 705.
[39] V. Santoni, N. Molly, T. Rabilloud, Electrophoresis 21 (2000) 1054.
[40] V. Santoni, S. Kieffer, D. Desclaux, F. Masson, T. Rabilloud, Electrophoresis 21 (2000) 3329.
[41] M. Chevallet, V. Santoni, A. Poinas, D. Rouquie, A. Fuchs, S. Kieffer, M. Rossignol, J. Lunardi, J. Garin, T. Rabilloud, Electrophoresis 19 (1998) 1901.
[42] S. Luche, V. Santoni, T. Rabilloud, Proteomics 3 (2003) 249.
[43] M. Ferro, D. Seigneurin-Berny, N. Rolland, A. Chapel, D. Slavi, J. Garin, J. Joyard, Electrophoresis 21 (2000) 3517.
[44] M.P. Molloy, Anal. Biochem. 280 (2000) 1.
[45] A.H. Millar, J.L. Heazelwood, Plant Physiol. 131 (2003) 443.
[46] E. Jung, M. Heller, J.C. Sanchez, D.F. Hochstrasser, Electrophoresis 21 (2000) 3369.
[47] S. Chivasa, B.K. Ndimba, W.J. Simon, D. Robertson, X.L. Yu, J.P. Knox, P. Bolwell, A.R. Slabas, Electrophoresis 23 (2002) 1754.
[48] J.B. Peltier, G. Frisco, D.E. Kalume, P. Roepstorff, F. Nilsson, I. Adamska, K.J. van Wijk, Plant Cell. 12 (2000) 319.
[49] T. Prime, D. Sherrier, P. Mahon, L. Packman, P. Dupree, Electrophoresis 21 (2000) 3488.
[50] D.J. Maltman, W.J. Simon, C.H. Wheeler, M.J. Dunn, R. Wait, A.R. Slabas, Electrophoresis 23 (2002) 626.
[51] V. Kruft, H. Eubel, L. Jansch, W. Werhan, H.P. Braun, Plant Physiol. 127 (2001) 1694.
[52] A.H. Millar, L.J. Sweetlove, P. Giege, C.J. Leaver, Plant Physiol. 127 (2001) 1711.
[53] J. Bardel, M. Louwagie, M. Jaquinod, A. Jourdain, S. Luche, T. Rabilloud, D. Macherel, J. Garin, J. Bourguignn, Proteomics 2 (2002) 880.
[54] M.S. Bae, E.J. Cho, E.Y. Choi, O.K. Park, Plant J. 36 (2003) 652.
[55] Y. Fukao, M. Hayashi, M. Nishimura, Plant Cell. Physiol. 43 (2002) 689.
[56] V. Santoni, D. Rouquie, P. Doumas, M. Manison, M. Boutry, H. Degand, P. Dupree, L. Packman, J. Sherrier, T. Prime, G. Bauw, E. Posada, P. Rouze, P. Dehais, I. Sahnoun, I. Barlier, M. Rossignol, Plant J. 16 (1998) 633.
[57] P.F. O'Farrell, J. Biol. Chem. 250 (1975) 4007.
[58] J. Klose, Humangenetik 26 (1975) 231.
[59] G.A. Scheele, J. Biol. Chem. 250 (1975) 5375.
[60] B.R. Herbert, J.L. Harry, N.H. Packer, A.A. Gooley, S.K. Pedersen, K.L. Williams, Trends Biotech. 19 (2001) S3.
[61] P.G. Righetti, A. Stoyanov, M. Zhukov, The Proteome Revisited: Theory and Practice of the Relevant Electrophoretic Steps, Elsevier, Amsterdam, 2001.
[62] A. Gorg, Nature 349 (1991) 545.
[63] A. Gorg, C. Obermaier, G. Boguth, A. Harder, B. Scheibe, R. Wildgruber, W. Weiss, Electrophoresis 21 (2000) 1037.
[64] H. Schagger, G. von Jagow, Anal. Biochem. 199 (1991) 223.
[65] H. Schagger, Methods Cell Biol. 65 (2001) 231.

[66] H. Eubel, L. Jansch, H.P. Braun, Plant Physiol. 133 (2003) 274.
[67] B.J. Hanson, B. Schulenberg, W.F. Patton, R.A. Capaldi, Electrophoresis 22 (2001) 950.
[68] W. Werhahn, H.P. Braun, Electrophoresis 23 (2002) 640.
[69] J.M. Hille, A.L. Freed, H. Watzig, Electrophoresis 22 (2001) 4035.
[70] T. Rabilloud, Proteomics 2 (2002) 3.
[71] W.F. Patton, B. Schulenberg, T. Steinberg, Curr. Opin. Biotechnol. 13 (2002) 321.
[72] M. Hamdan, M. Galvani, E. Bordini, P.G. Righetti, Electrophoresis 22 (2001) 1633.
[73] D.W. Cleveland, S.G. Fischer, M.V. Kirschner, U.K. Laemmli, J. Biol. Chem. 252 (1977) 1102.
[74] P.J. Matsudaria, J. Biol. Chem. 262 (1987) 10035.
[75] I. Kameshita, H. Fujisawa, Anal. Biochem. 249 (1997) 252.
[76] V.L. Herald, J.L. Heazlewood, D.A. Day, A.H. Millar, FEBS Letts. 537 (2003) 96.
[77] M. Karas, F. Hillenkamp, Anal. Chem. 60 (1988) 2299.
[78] J.B. Fenn, M. Mann, C.K. Meng, S.F. Wong, C.N. Whitehouse, Science 246 (1989) 64.
[79] J. Li, S.M. Assmann, Plant Physiol. 123 (2000) 807.
[80] R. Aebersold, M. Mann, Nature 422 (2003) 198.
[81] K.D. Henry, E.R. Williams, B.H. Wang, F.W. McLafferty, J. Shabanowitz, D.F. Hunt, Proc. Natl. Acad. Sci. U.S.A. 86 (1989) 9075.
[82] S.C. Beu, M.W. Senko, J.P. Quinn, F.M. Wampler III, F.W. McLafferty, J. Am. Soc. Mass Spectrom. 4 (1993) 557.
[83] N.L. Kelleher, H.Y. Lin, G.A. Valaskovic, D.J. Aasured, E.K. Fridriksson, F.W. McLafferty, J. Am. Chem. Soc. 121 (1999) 806.
[84] T.H. Mize, W.J. Simonsick Jr., I.J. Amster, Eur. J. Mass Spectrom. 9 (2003) 473.
[85] L.K. Zhang, D. Rempel, B.N. Pramanik, M.L. Gross, Mass Spectrom. Rev., in press.
[86] A.J. Link, J. Eng, D.M. Schicltz, E. Carmack, G.J. Mizc, D.R. Morris, B.M. Garvik, J.R. Yates III, Nat. Biotechnol. 17 (1999) 676.
[87] M.P. Washburn, D. Wolters, J.R. Yates III, Nat. Biotechnol. 19 (2001) 242.
[88] D.A. Wolters, M.P. Washburn, J.R. Yates III, Anal. Chem. 73 (2001) 5683.
[89] D.K. Han, J. Eng, H. Zhou, R. Aebersold, Nat. Biotechnol. 19 (2001) 946.
[90] T.R. Covey, R.F. Bonner, B.I. Shushan, J. Henion, Rapid Commun. Mass Spectrom. 2 (1988) 249.
[91] B.T. Chait, S.B. Kent, Science 257 (1992) 1885.
[92] J.P. Whitelegge, J. le Coutre, J.C. Lee, C.K. Angel, G.G. Prive, K.F. Faull, H.R. Kaback, Proc. Natl. Acad. Sci. U.S.A. 96 (1999) 10695.
[93] A. Premstaller, H. Oberacher, W. Walcher, A.M. Temperio, L. Zolla, J.P. Chervet, N. Cavusoglu, A. Van Dorsellaer, C.G. Huber, Anal. Chem. 73 (2001) 2390.
[94] J.P. Whitelegge, J. le Coutre, Am. J. Pharmacogen. 1 (2001) 29.
[95] J.P. Whitelegge, C.V. Gundersen, K.F. Faull, Protein Sci. 7 (1998) 1423.
[96] L. Zolla, A.M. Timperio, Proteins 41 (2000) 398.
[97] C.G. Huber, A.M. Timperio, L. Zolla, J. Biol. Chem. 276 (2001) 45755.
[98] S.M. Gomez, J.N. Nishio, K.F. Faull, J.P. Whitelegge, Mol. Cell Proteomics 1 (2002) 46.
[99] L. Zolla, S. Rinalducci, A.M. Timperio, C.G. Huber, Plant Physiol. 130 (2002) 1938.
[100] L. Zolla, A.M. Timperio, W. Walcher, C.G. Huber, Plant Physiol. 131 (2003) 198.
[101] V. Zabrouskov, L. Giacomelli, K.J. van Wijk, F.W. McLafferty, Mol. Cell. Proteomics 2 (2003) 1253.
[102] J.B. Peltier, O. Emanuelsson, D.E. Kalume, J. Ytterberg, G. Frisco, A. Rudella, D.A. Liberles, L. Soderberg, P. Roepstorff, G. von Heijne, K.J. van Wijk, Plant Cell. 14 (2002) 211.
[103] M. Schubert, U.A. Petersson, B.J. Haas, C. Funk, W.P. Schroder, T. Kieselbach, J. Biol. Chem. 277 (2002) 8354.
[104] F. Meng, B.J. Cargile, S.M. Patrie, J.R. Johnson, S.M. McLoughlin, N.L. Kelleher, Anal. Chem. 74 (2002) 2923.
[105] G.K. Taylor, Y. Kim, A.J. Forbes, F. Meng, R. McCarthy, N.L. Kelleher, Anal. Chem. 75 (2003) 4081.
[106] W.C. Lee, K.H. Lee, Anal. Biochem. 324 (2004) 1.
[107] D. Figeys, L.D. McBroom, M.F. Moran, Methods 24 (2001) 230.
[108] J. Porath, J. Carlsson, I. Olsson, G. Belfrage, Nature 258 (1975) 598.
[109] A.V. Vener, A. Harms, M.R. Sussman, R.D. Vierstra, J. Biol. Chem. 276 (2001) 6959.
[110] V. Gaberc-Porekar, V.J. Menart, Biochem. Biophys. Methods 49 (2001) 335.
[111] E.K. Ueda, P.W. Gout, L. Morganti, J. Chromatogr. A 988 (2003) 1.
[112] L. Andersson, J. Porath, Anal. Biochem. 154 (1986) 250.
[113] M.C. Posewitz, P. Tempst, Anal. Chem. 71 (1999) 2883.
[114] S.B. Ficarro, M.L. McCleland, P.T. Stukenberg, D.J. Burke, M.M. Ross, J. Shabanowitz, D.F. Hunt, F.M. White, Nat. Biotechnol. 20 (2002) 301.
[115] M. Hansson, A.V. Vener, Mol. Cell. Proteomics 2 (2003) 550.
[116] T.S. Nuhse, A. Stenballe, O.N. Jensen, S.C. Peck, Mol. Cell. Proteomics 2 (2003) 1234.
[117] K. Lee, J. Lee, Y. Kim, D. Bae, K.Y. Kang, S.C. Yoon, D. Lim, Electrophoresis 25 (2004) 532.
[118] Y. Balmer, A. Koller, G. de Val, W. Manieri, P. Schumann, B.B. Buchanan, Proc. Natl. Acad. Sci. U.S.A. 100 (2003) 370.
[119] D. Yamazaki, K. Motohashi, T. Kasama, Y. Hara, T. Hisabori, Plant Cell Physiol. 45 (2004) 18.
[120] M. Norin, M. Sundstrom, Trends Biotechnol. 20 (2002) 79.
[121] J.H. Ippel, L. Pouvreau, T. Kroel, H. Gruppen, G. Versteeg, P. van den Putten, P.C. Struik, C.P.M. van Mierlo, Proteomics 4 (2004) 226.
[122] W. Weckwerth, K. Wenzel, O. Fiehn, Proteomics 4 (2004) 78.
[123] E. Dumas-Gaudot, N. Amiour, S. Weidmann, G. Bestel-Corre, B. Valot, S. Lenogue, V. Gianinazzi-Pearson, S. Gianinazzi, Proteomics 4 (2004) 451.
[124] W.W. Li, G.V. Quinn, N.N. Alexandrov, P.E. Bourne, I.N. Shindyalov, Genome Biol. 4 (2003) R51.1.
[125] H. Schoof, R. Ernst, V. Nazarov, L. Pfeifer, H.W. Mewes, K.F.X. Mayer, Nucleic Acid Res. 32 (2004) D373.
[126] A. Carpi, G. Di Maira, M. Vedovato, V. Rossi, T. Naccari, M. Floriduz, M. Terzi, F. Filippini, Proteomics 2 (2002) 1494.
[127] R. Apweiler, A. Bairoch, C.H. Wu, Curr. Opin. Chem. Biol. 8 (2004) 76.
[128] A. Sickmann, M. Mreyen, H.E. Meyer, Adv. Biochem. Eng. Biotechnol. 83 (2003) 141.
[129] W.J. Henzel, C. Watanabe, J.T. Stults, J. Am. Soc. Mass Spectrom. 14 (2003) 931.
[130] M. Ferro, D. Seigneurin-Berny, N. Rolland, A. Chapel, D. Slavi, J. Garin, J. Joyard, Electrophoresis 21 (2000) 3517.
[131] M. Ferro, D.D. Slavi, S. Brugiere, S. Miras, S. Kowalski, M. Louwagie, J. Garin, J. Joyard, N. Rolland, Mol. Cell. Proteomics 2 (2003) 325.
[132] A. Sickmann, M. Mreyen, H.E. Meyer, IUBMB Life 54 (2002) 51.
[133] O.N. Jensen, Curr. Opin. Chem. Biol. 8 (2004) 33.
[134] S.C. Peck, T.S. Nuhse, D. Hess, A. Iglesias, F. Meins, T. Boller, Plant Cell. 13 (2001) 1467.
[135] T.H. Steinberg, B.J. Agnew, K.R. Gee, W.Y. Leung, T. Goodman, B. Schulenberg, J. Hendrickson, J.M. Beechem, R.P. Haugland, W.F. Patton, Proteomics 3 (2003) 1128.
[136] B.P. Mooney, J.J. Thelen, Phytochemistry 65 (2004) 1733.
[137] G.K. Agrawal, R. Rakwal, M. Yonekura, A. Kubo, H. Saji, Proteomics 2 (2002) 947.