J Plant Biotechnol (2010) 37:196–204
DOI:10.5010/JPB.2010.37.2.196

Review

Comprehensive proteome analysis using quantitative proteomic technologies

Abu Hena Mostafa Kamal · Jong-Soon Choi · Yong-Gu Cho · Hong-Sig Kim · Beom-Heon Song · Chul-Won Lee · Sun-Hee Woo

Received: 3 April 2010 / Accepted: 13 April 2010
ⓒ Korean Society for Plant Biotechnology

Abstract With the completion of genome sequencing of several organisms, attention has been focused to determine the function and functional network of proteins by proteome analysis. The recent techniques of proteomics have been advanced quickly so that the high-throughput and systematic analyses of cellular proteins are enabled in combination with bioinformatics tools. Furthermore, the development of proteomic techniques helps to elucidate the functions of proteins under stress or diseased condition, resulting in the discovery of biomarkers responsible for the biological stimuli. Ultimate goal of proteomics orients toward the entire proteome of life, subcellular localization, biochemical activities, and their regulation. Comprehensive analysis strategies of proteomics can be classified as three categories: (i) protein separation by 2-dimensional gel electrophoresis (2-DE) or liquid chromatography (LC), (ii) protein identification by either Edman sequencing or mass spectrometry (MS), and (iii) quanitation of proteome. Currently MS-based proteomics turns shiftly from qualitative proteome analysis by 2-DE or 2D-LC coupled with off-line matrix assisted laser desorption ionization (MALDI) and on-line electrospray ionization (ESI) MS, respectively, to quantitative proteome analysis. Some new techniques which include top-down mass spectrometry and tandem affinity purification have emerged. The *in vitro* quantitative proteomic techniques include differential gel electrophoresis with fluorescence dyes, protein-labeling tagging with isotope-coded affinity tag, and peptide-labeling tagging with isobaric tags for relative and absolute quantitation. In addition, stable isotope labeled amino acid can be *in vivo* labeled into live culture cells through metabolic incorporation. MS-based proteomics extends to detect the phosphopeptide mapping of biologically crucial protein known as one of post-translational modification. These complementary proteomic techniques contribute to not only the understanding of basic biological function but also the application to the applied sciences for industry.

Keywords Chromatography, Gel electrophoresis, Mass spectrometry, Plant proteomics, Quantitative proteomics

Introduction

Sequence analysis of genomic DNA, which started in 1990s on a full scale, developed rapidly during the last decade. The advent of genomic technology facilitated to overlook the entire gene information of many organisms including higher plants and enable us to design the tailored transformed organism using genetic tools. The genome itself gives us a sort of guide map where we depart and where we arrive. The actual life events happen in the coordination and cross-talk of cellular proteins. However, there left behind a technical gap to bridge between our understanding of genome sequence and cellular behaviors. Thus, understanding what proteins are transiently and specially expressed can be extended from the snap picture showing where proteins are to the nearly real movie showing how cells work inside. The terms "proteome" and "proteomics" were tossed up first by Mark Wil-

A. H. M. Kamal · H.-S. Kim · B.-H. Song · C.-W. Lee · S.-H. Woo (✉)
Dept. of Crop Science, Chungbuk National University, Cheong-ju 361-763, South Korea
e-mail: shwoo@chungbuk.ac.kr

J.-S. Choi
Division of Life Science, Korea Basic Science Institute, Daejeon 305-333, Korea

J.-S. Choi
Graduate School of Analytical Science and Technology, Chungnam National University, Daejeon 305-764, Korea

kins in the Siena Two-dimensional Electrophoresis Meeting to explain to complement “gene” and “genomics”, respectively (Wilkins et al. 1996). Three technical breakthroughs behind the birth of proteomics include (i) the full collection of genes and protein-sequence databases, (ii) the user-friendly bioinformatic tools to handle the tremendous plenty of biological data, and (iii) the microchip array platform technology to interpret the genome-wide expression. Currently many useful tools are placed on our hands to overview the whole system of life, where biologists in turn start to think big. Although a number of reviews on proteome research in plants are available (van Wijk 2001; Thiellement et al. 2002; Woo et al. 2002; Woo et al. 2003; Islam and Hirano 2003; Rakwal and Agrawal 2003), comprehensive information on proteome research encompassing most of the plants for which genome sequence are available are scant.

In this review, the enabling and challenging proteomic techniques will be reviewed and considered to hardness the disciplines whose research will be applied for the purpose of industry.

Consideration about biological system

Proteomics considered being closer to a technology rather than a science seems to be a strong and powerful prism to screen the multi-proteins in a cell, of which is to characterize not the behavior of the single protein molecule but the network of biological system. In this context, the final goal for proteomists to pursuit is totally different from those who are tamed in protein chemistry. Protein chemistry primarily concerns the complete sequences of target protein in order to speculate the cellular function, where its function can be deduced from the determination of three-dimensional structure. In contrary to classical protein chemistry, proteomics involves the high-throughput identification of expressed proteins by partial sequence analysis and, ultimately, the quantitation of proteins to interpret cellular function systematically. Thus, proteomics directs to investigate the cellular dynamics with proteomes under the biological environment, of which characteristics is a completely different feature compared with a static and fixed picture taken by genome and genomics. Regardless of cell type, the observed proteins are constitutively expressed proteins in common for biologically essential functions such as DNA replication, protein synthesis, and energy metabolism. The protein with unique cell-specific functions like rhodopsin in retinal epithelium can be a biomarker target protein responsible for specific function in a specialized biosource. In particular, proteins involved in signal transduction, transcription factor, and cell-cycle control turn over rapidly leading to play a key role of specific function in a specialized cell. Though the settlement of oligonucleotide chip analysis renders to interpret the whole genome-wide expression pattern in a cell (DeRisi et al. 1997), we can not say the corresponding proteins directly. There are three possible reasons to explain the difference between transcriptome and proteome. The stability of mRNA and the translational efficiency exert subsequently to affect the protein expression qualitatively and/ or quantitatively. Furthermore, only mRNA levels can’t singly tell the status about the elaborate regulation of a cell. In previous literature, there are many evidences to show the poor correlation between mRNA copies and its coding protein quantity (Gygi et al. 1999a, b). Proteomics explains what is happening in a cell, whereas transcriptomics and genomics tell us what might be happening and what could be happening, respectively. Taken for an extraordinary example with human α-antitypsin, 22 different types of isoforms exist in human plasma (Hoogland et al. 1999). Such the complexity of a protein can be shown only by proteomic study and hidden by transcriptomic and genomic studies. After the completion of human genome sequencing (Venter et al. 2001), many reductionists were surprised at the finding that the gene numbers of human comprising 10^{13} cells with ~26,000 genes have fewer twice as many as the 959-cell nematode, *Caenorhabditis elegans* with ~18,000 genes (The *C. elegans* Sequencing Consortium 1998; Online Datasets at http://www.integratedgenomics.com/ GOLD). In summary, proteomics seeks for the identification and quantitation of proteins first and subsequently for the determination of their localization, modification, interaction, activity, and their cellular function (Fields 2001).

Protein separation based on 2-dimensional gel electrophoresis

Analytical proteomics generally begins with separation of proteins or peptides. As shown in Fig. 1, the simple two strategies to analytical proteomics by MS can be divided with (i) the separation of complex protein mixture followed by tryptic digestion and reversely, (ii) the whole digestion of protein mixture followed by separation for the identification of proteins by on-line MS. In proteome analysis, first a number of proteins are separated 2-DE (O’ Farrel 1975; Klose 1975) is often used for the separation of proteins. The former procedure can be simply exerted by two-dimensional gel electrophoresis (2-DE) whereas the latter one, called as shot-gun proteomics, can be done by the tandemly linked two-dimensional liquid chromatography (2D-LC) with on-

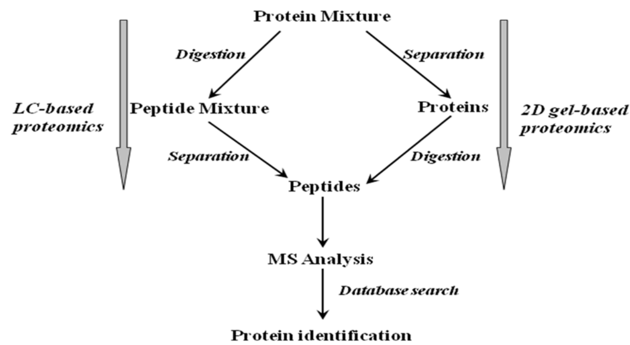

Fig. 1 Flowchart of analytical proteomics

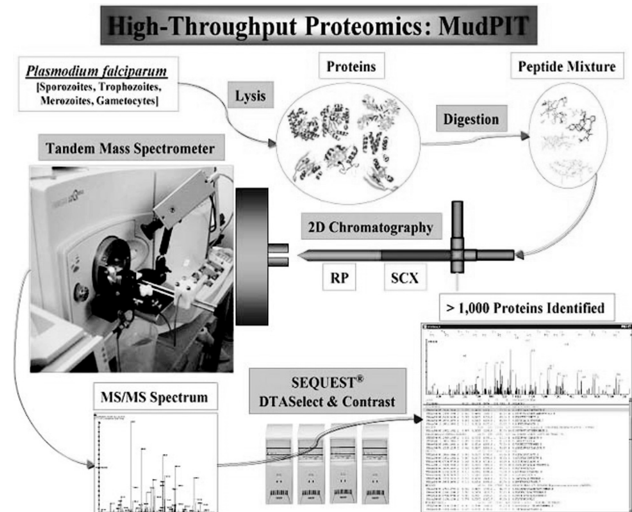

Fig. 2 The scheme of MudPIT. Denatured complex proteins are digested as a mixture and the digested peptides are separated by randomly connected columns, i.e., strong cation exchange and reverse-phase column. The subsequently MS-MS fragmented peptides are submitted in protein databases using SEQUEST (Bardi 2002)

line MS (MacCoss et al. 2002). The 2-DE is well-known protein separation method according to their isoelectric points (p*I*) and molecular weight (M*r*). The immobilized pH gradient strip gels provide commercially the stable pH gradients with 2-DE experiment users (Görg et al. 1988). Advantages of using 2-DE are simply able to display the protein profiles visually on 2-DE gel and to detect some isoforms of post-translational modification. Typically phosphorylated and non-phosphorylated proteins can be horizontally resolved on 2-DE due to the changed pI values by phosphorylation (Lewis et al. 2000). Recently, the improvement of 2-DE is introduced as the name of differential gel electrophoresis (DIGE) that applies two different fluorescence dyes (Cy5 and Cy3) for two states of protein samples and subsequently subjects to running on one same 2-D gel (Unlu et al. 1997).

Though the 2-DE based proteomics are relatively simple and low-cost, however, it is still labor-intensive and time-consuming, in particular, the drawback of automation. In addition, the proteins with extreme pH such as strong acidic or basic p*I*'s and hydrophobic proteins are hard to separate and visualize on conventional 2-D gel. To overcome this limitation, shot-gun proteomic method to bypassing 2-DE is developed by converting the entire peptide mixture after whole digestion of proteins with trypsin in combination with liquid chromatographic techniques (Washburn et al. 2001). Multidimensional protein identification (MudPIT) technique is a recently developed 2D-LC separation method by tandemly linked with cation-exchange column and reverse phase column, which this MudPIT applied to analyze several whole proteomes with enhanced computing power (Fig. 2).

Mass spectrometry and proteome analysis methods

One of the oldest methods to identify the unknown protein is N-terminal sequencing by Edman degradation chemistry (Edman 1949). The Edman reagent phenylisothiocyanate chemically modifies the free amino termini of polypeptide to form phenylthiocarbamyl (PTC) polypeptide. The addition of anhydrous acid to PTC amino-terminal residue rapidly cleaved from the polypeptide chain to yield anilinothiazolinone (ATZ) amino acid and *n-1* polypeptide. The unstable ATZ derivative amino acid converted to a more stable phenylthiohydantoin (PTH)-amino acid, in which this derivative allows to run in reverse-phase column to identify the amino acid by comparing it with the standard PTH amino acid. Though MS primarily replaces Edman sequencing in modern proteomics, it has still useful in distinguishing between isoleucine and leucine with the same molecular weight of 113 Da, obtaining N-terminal processed information after translation, and the identification of protein lack of genomic information by homology search (Gooley et al. 1997).

Besides the direct sequencing of protein, the information of partially digested peptide fragments from original protein can provide some clues by comparing the collection of experimentally obtained molecular weight data with genomic database. Theoretically each protein can produce a peptide mass fingerprinting (PMF) that is a peculiar pattern of given protein by specific protease (Pappin et al. 1993). Taken for the example of trypsin, the post-tryptic digestion products can be expected as the peptide fragments cleaved after carboxy-termini of arginine or lysine. Thus, the protein source from either separated protein on 2-DE or eluted solution of liquid chromatography can be a substrate for trypsin. As shown in Fig. 3, the purified tryptic digests

Fig. 3 Flowchart of peptide mass finger printing. The protein spots of interest were picked, after following tryptic digestion, analyzed by mass spectrometry

dissolved with 50% (v/v) aqueous acetonitrile and 0.1% (v/v) formic acid were added with a matrix solution of α -cyano-4-hydroxy cinnamic acid in 50% acetonitrile. The mixture is applied on MALDI plate to allow the crystals to be ionizable by nanosecond laser pulses, in which the ionized peptides in turn evaporate and mobilize into a vacuum tube to reach at detector. The molecular weight can be calculated as a form of mass-to-charge (*m/z*) by the mode of time-of-flight (TOF) analyzer from the applied kinetic energy, $E = \frac{1}{2}mv^2$. The collected data of *m/z* values are given to search engine, i.e., MASCOT (http://www.matrixscience.com) to make reports of the statistical summary of PMF in the order of best fitting scores. The identification methods such as Edman sequencing and PMF provide complementary protein information under the given physiological condition (Choi et al. 2000; Woo et al. 2002).

When the protein identification methods described above are unsatisfied, much higher degree of protein information can be obtained from the further fragmentation of the prevalent parent ion molecules to generate daughter ions. During this MS/MS analysis, the bidirectional peptide sequences of single peptide can be obtained by analyzing *b*- and *y*-ion series from database. The *de novo* peptide sequence data by MS/MS is very useful for protein identification with high confidence by searching peptide sequence information across both nucleotide and protein databases. The protein identification by MS/MS analysis like Edman sequencing is most adequate for organisms with less fully annotated databases by incomplete genome sequencing. The automated protein identification by MS-MS on a large-scale can be performed with on-line nanoscale-LC based electrospray ionization quadrupole TOF-MS or off-line LC-connected with MALDI-TOF/TOF-MS. These tandem MS-MS analyzers are compatible with shot-gun proteomics such as MudPIT.

Proteomic approach to protein phosphorylation is one of challenging technology. Protein phosphorylation is one of the most abundant and important post-translational modification and regulates biological function (Hunter 1995). Phosphorylation happens in various cellular dynamic processes such as cell division, metabolic regulation, and signal transduction. Regarding with signaling pathway process, phosphorylated proteins are known to occupy 1~2% of entire proteome and encountered with the dynamic characteristics by kinases and phosphatases (Schlessinger 1993). A common detection of phosphoproteins is by radioactive labelling of the phosphate groups with inorganic ^{32}P and subsequent autoradiography. Another detection of phosphoproteins is by Western-blotting using phosphospecific antibodies, in which anti-phosphoserine, -phosphothreonine, and -phosphotyrosine antibodies are commercially provided (Kaufmann et al. 2001).

Systematic quantitative proteomics techniques

Gel-based comparative proteomics such as DIGE has some limited drawbacks to full-automation and underrepresented low abundant proteins, poor resolution of highly acidic/basic proteins and proteins with extreme molecular weights and hydrophobicity. Contrary to this gel-based proteomics, LC-based quantitative techniques provide various feasibilities to high-throughput and automation of proteome by downsizing proteome complexicity to optimal proteins or peptides. Multi-dimensional LC approach abbreviated as MDLC is the basic analysis template, in which MDLC can be in turn applied to the systematically quantitative analysis of proteome in two states of cells. Chemical tagging of proteins and/or peptides in two proteomes under study gives directly relative quantitation of proteins by LC-MS analyses. The surface enhanced laser desorption ionization (SELDI)-MS is used in protein differential display analysis without using 2-DE. In SELDI MS, the protein is immobilized in 2 mm diameter of holes on the surface of metal affinity, antibody affinity, receptor ligand interaction, etc (Issaq et al. 2002).

The prototype of chemical tagging is the isotope-coded affinity tag (ICAT) developed by Aebersold and colleagues (Gygi et al. 1999a). ICAT labeling is also used in protein differential display analysis without 2-DE. The ICAT reagent consists of biotin group to bind avidin column, polyether linker with eight hydrogen (D_0) or deuterium (D_8), and reactive group specifically labeling to cysteine residues. The combined proteomes labeled with each light and heavy ICAT reagent allow to be digested by trypsin followed by

capturing biotin-containing peptides to avidin column. After washing the unbound peptides with buffer, the labeled cysteine-containing peptides are eluted and quantified with MS. Doublet peaks with 8 Da difference are targets to quantify the peak intensity implementing relative abundance of two states of cell populations. The identification of the target peaks is carried out by the bidirectional readout of amino acid sequences by MS/MS analysis. Recently, new cleavable ICAT reagent with ^{13}C isotope linker and acid-cleavable biotin group is introduced to overcome the limitations of ICAT such as co-elution of D_0 and D_8-labeled peaks and difficulty of identification due to biotin group (Hansen et al. 2003).

Another striking quantitative method, isobaric tags for relative and absolute quantitation (iTRAQ) was in public to employ four-plex labeling of amine-reactive isobaric tags containing an amine group of N-terminus and side chain of lysine, in which these tagging is capable of covering the whole proteome (Fig. 4). The iTRAQ labeling is based on the labeling of peptide level, thereby labeling after trypsinizing all the proteins. Peptides labeled with four-plex isobaric tags are impossible to identify each state of proteome in MS.

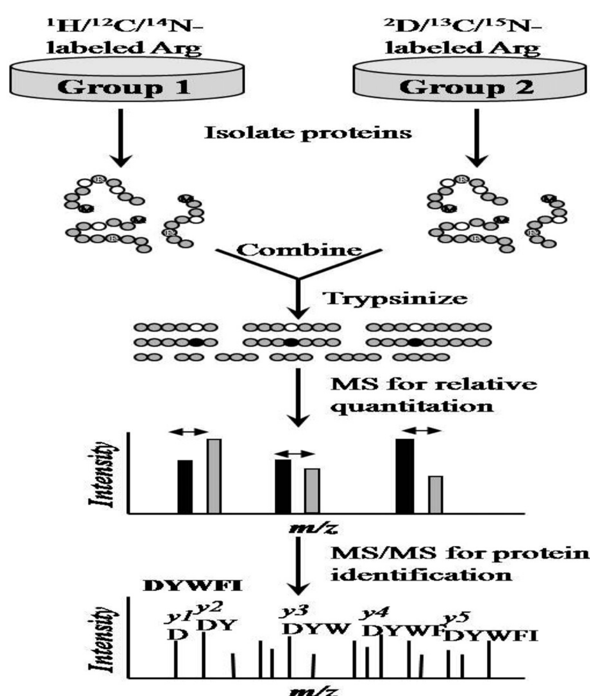

Fig. 4 Workflow of quantitative proteomics techniques. Proteins are labelled metabolically by culturing cells in media that are isotopically enriched (for example, containing $^1H/^{12}C/^{14}N$, or $^2D/^{13}C/^{15}N$ labeled Arg) or isotopically depleted. In each case, labeled proteins or peptides are combined, separated and analyzed by mass spectrometry and/or tandem mass spectrometry for the purpose of identifying the proteins contained in the sample and determining their relative abundance. The patterns of isotopic mass differences generated by each method are indicated schematically.

However, when the collected four-plex tagging peptides are fragmented by MS/MS analysis, signature ion peaks from *m/z* 114 to 117 are produced to give the relative quantitative information and concomitantly MS/MS spectra are given to identify the protein to be targeted (Ross et al. 2004). The chemical tagging methods such as ICAT, cICAT, and iTRAQ allow labeling the proteins or peptides *in vitro*, thus employing labeling, enrichment, and purification followed by MS and MS/MS analysis. The serial chemical reaction processes include the poor reproducibility and low confidence. Thus, labeling stable isotope labeling by essential amino acid culture (SILAC) is another innovative approach to bypass the unnecessary chemical reactions (Ong et al. 2002). Moreover, SILAC is *in vivo* labeling into the live cells through metabolic incorporation, in which the proteomes in two states of cultured cell populations are able to be used for direct quantitative analysis from purified proteins in cultured cells under study.

Plant proteomics and application

This review focused on general and more aspects of proteomics and cover issues such as protein separation, image analysis, mass spectrometry (MS), and bioinformatics. These issues are more relevant to plant biology as crop science. The array of proteomics applications varies from straightforward investigation of proteins to characterization of post-translation modification, comparative proteomics, and protein-protein interactions. In plant, the most widespread techniques for soluble and insoluble proteins is two-dimensional-SDS-PAGE, using immobilized pH gradient strip in the first dimensional for comparative proteomics and SDS-PAGE, Tricine 1D-SDS-PAGE for protein profiling including membrane proteomics using high throughput mass spectrometry such as LCQ-ESI-MS/MS, LTQ-ESI-FTICR and so on.

Mass spectrometer generates a large amount of numerical data and bioinformatics tools therefore are essential to match the MS data to protein, EST, and genome sequence databases. Most search engine have been developed in academic laboratories and some of those now been commercialized. Most of them plant are still un-sequenced without model plant *Arabidopsis* and Rice such as wheat, maize crops, and also have no identical database in NCBInr and SWISSPROT. For example, wheat (*Triticum* sp.) is one of the most important cereal crops for the global food supply. Many kind of wheat cultivars have been wheat bred and used for commercial foods such as bread, noodles, pasta, cakes and so on. Kamal et al. (2009a) reported that pre-harvesting wheat cultivars proteins profile using proteomics followed by

MALDI-TOF-MS. Out of the 73 protein spots submitted to proteomics analysis, identified 482 proteins for majority of the unique proteins with isoforms. Based on functional distribution, the total identified proteins were categorized into 9 categories: Metabolism (19%), storage (18%), photosynthesis (11%), amino acid (2%), allergy (1%), stress (16%), protein synthesis (16%), enzyme (14%), hypothetical (3%) in Jinpum and Metabolism (26%), storage (17%), photosynthesis (9%), amino acid (0%), allergy (4%), stress (31%), protein synthesis (6%), enzyme (6%), hypothetical (1%) in Keumgang.

Glutenin is a major determinant of baking performance and viscoelasticity, which are responsible for high-quality bread with a light porous crumb structure of a well-leavened loaf. Gluten protein composition determines the rheological characteristics (strength and extensibility) of flour dough and is the key constituent responsible for differences in end-use suitability (Butow et al. 2003). The gluten proteins consist of the monomeric gliadins and polymeric glutenins, where glutenins contain high and low molecular weight glutenin subunits. Wheat grain research has focused on the detailed analysis of gluten proteins to better understand those aspects of protein composition accounting for the unique properties of flour (Skylas et al. 2005). When treated with a reducing agent, glutenins dissociate into subunits of differing molecular weight: the high molecular weight subunits (HMW-GS) and the low molecular weight subunits (LMW-GS) (Payne et al. 1981). The glutenin consisted of two types of subunits containing LMW-GS (10–70 kDa) and HMW-GS (80–130 kDa) (Bietz and Wall 1972). When glutenins are further reduced, two types of subunits are released into high molecular weight subunit of 70–90 kDa and low molecular weight subunit of 20–45 kDa. However, actual molecular weight of glutenin calculated from derived amino acid sequences indicated 60–90 kDa, lower molecular weight rather than ever expected (Anderson and Green 1989; Anderson et al. 1989). A total of 217 proteins spots were subjected to tryptic digestion and identified by matrix assisted laser desorption/ionization–time of flight mass spectrometry. As shown in Figure 5, HMW-GS (43 isoforms) and LMW-GS (seven isoforms) are directly responsible for producing high-quality bread and noodles. Likewise, all the seed storage proteins are digested to provide nutrients for the embryo during seed germination and seedling growth (Kamal et al. 2009b). Out of 52 glutenin proteins, using peptide fragmentation method, 13 proteins were identified in Keumgang followed by 11 in Jinpum, 7 in China-108, 6 in Yeonnon-78, 10 in Norin-61, and 5 in Kantou-107. Two linked genes, puroindoline a (PINA) and puroindoline b (PINB), control most of the genetic variation in wheat grain texture,

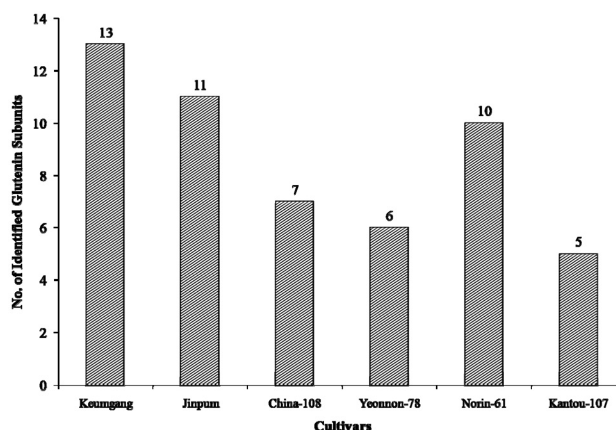

Fig. 5 Distribution of total identified glutenin subunits (GSs) protein among wheat cultivars (Kamal et al. 2009b).

which was examined to identify PINA and PINB gene using two Korean pre-harvest sprouting wheat cultivars; Jinpum (resistant) and Keumgang (susceptible), following proteins were separated and identified by two-dimensional electrophoresis with MALDI-TOF/MS (Kamal et al. 2009c).

Technical advances are still needed at a number of points in strategies for comparative or functional proteomics. In order to simplify mixtures of proteins from plant disease cells, many protein scientist presented an improved method to isolate membrane proteins and discuss the separation of membrane proteins by high throughput mass spectrometer. We reviewed the proteolytic labeling method introduced by this laboratory to provide global labeling of proteolytic peptide products from protein mixtures and demonstrate this approach in a comparative study of protein changes in acquired abiotic and biotic resistance. The identification of proteins revealed on 2D gels is useful means to examine the metabolic changes induced by treatment or any environmental variation. This approach was used to study effect of drought (Peng et al. 2009), salinity (Aghaei et al. 2009), water stress, heavy metal stress (Schützendübel and Polle 2002), and biotic stress (Wang et al. 2005) on protein expression in different crops.

Identification of multi-subunit complexes using non-denaturing membrane proteins such as chloroplast (D'Amici 2009), mitochondria (Haezlewood et al. 2003), plasma membrane (Jonas et al. 2004), and so on using blue-native gel electrophoresis (BN-PAGE). Plastids are essential organelles of prokaryotic origins present in nearly every plant cell. Plastids are responsible for the synthesis and storage of key molecules required for the basic architecture and functions of plant cells. Chloroplasts are the best known plastid type and contain a Thylakoid membrane system, which is carried the photosynthesis electron transferring process converting

light energy into chemical energy. Purified membrane proteomes are typically fractioned to improve dynamic resolution using two phases and three phase partitioning method (Peltier et al. 2000). This can be fractionation directly coupled to MS, via 1-D and 2-D on-line chromatography as MUDPIT, and/or different off-line fractionation techniques such as SDS-PAGE, 2D-PAGE, BN-PAGE. The chloroplast sub-oraneller, all nuclear coded proteins identified on the 1D-PAGE gels were used to test predictions for chloroplast localization and transit peptides, transmembranes made by using some bioinformatics programs such as PSORT, TargetP (ChloroP and SignalP), Predotar, Mitoport (especially mitochondria), PredSL, TMHMM and ARAMEMNON. After correcting to possible isoforms, all identified proteins will be modified by post-translational modifications by using dbPTM and sumoylation site prediction (SUMOsp).

Towards integrated biology

Due to quickly developing proteome technologies, in particular, in mass spectrometry, biologists begin to build protein legos to make dynamic movies what is happening in the cell because the cell itself is not static or fixed at a single moment. Currently, it tends to integrate bunch of data forms concerning quantitative information (proteome dynamics), proteome localization (subproteome atlas), and protein-protein interaction (proteome society), in which it enables to overview and integrate the intracelullar happenings such as biomolecular networking systematically by computing power with mathematical and bioinformatic techniques (Kislinger et al. 2006). Taken together with many breakthroughs for the previous technical barriers, recent proteomics are turning rapidly into the quantitative technical platform in order to express the biological function with the numerical format. Thus, we hope to deal intractable cells with mathematics that is termed as integrated biology or systems biology.

Acknowledgement

This work was supported by a grant from the BioGreen 21 Program (20070301034016 & 20070301034043), Rural Development Administration, Republic of Korea

References

Aghaei K, Ehsanpour AA, Shah AH (2009) Proteome analysis of soybean hypocotyl and root under salt stress. Amino acids 36: 91-98

Anderson OD, Green FC (1989) The characterization and comparative analysis of high-molecular-weight glutenin genes from genomes A and B of a hexaploid bread wheat. Theor Appl Genet 77:689-700

Anderson OD, Green FC, Yip RE, Halford NG, Shewry PR, Malpica-Romero JM (1989) Nucleotide sequences of the two high-molecular weight glutenin genes from the D-genome of a hexaploid bread wheat, *Triticum aestivum* L. cv Cheyenne. Nucleic Acids Res 17:461-462

Bardi JS (2002) Scientists Identify Thousands of Proteins Associated with the Deadliest Form of Malaria.The scripps research institute 2:29 (http://www.scripps.edu)

Bietz JA, Wall JS (1972) Wheat gluten subunits: Molecular weights determined by sodium sulfate-polyacrylamide gel electrophoresis. Cereal Chem 49:416-430

Butow BJ, Ma BW, Gale KR, Cornish GB, Rampling L, Larroque O, Morrell MK, Bekes F (2003) Molecular discrimination of Bx7 alleles demonstrates that a highly expressed high-molecular-weight glutenin allele has a major impact on wheat flour dough strength. Theor Appl Genet 107:1524-1532

Choi JS, Kim DS, Lee J, Kim SJ, Kim SI, Kim, YH, Jong J, Yoo JS, Suh KH, Park YM (2000) Proteome analysis of light-induced proteins in Synechocystis sp. PCC 6803: Identification of proteins separated by 2D-PAGE using N-terminal sequencing and MALDI-TOF-MS. Mol Cells 10:705-711

D'Amici GM, Huber GC, Zolla L (2009) Separation of Thylakoid Membrane proteins by Sucrose Gradient Ultracentrifuge or Blue Native-SDS-PAGE Two-Dimensional Electrophoresis. Matthew J. Peirce, Robun Waits (eds.), Membrane Proteomics: methods and protocols 528:61-70

DeRisi JL, Iyer VR, and Brown PO (1997) Exploring the metabolic and genetic control of gene expression in a genomic scale. Science 278:680-686.

Edman P (1949) A method for the determination of the amino acid sequence in peptides. Arch Biochem 22:475-476

Fields S (2001) Proteomics. Proteomics in genomeland. Science 291:1221-1224

Gooley AA, Ou K, Russell J, Wilkins MR, Sanchez JC, Hochstrasser DF, Williams KL (1997) A role of Edman degradation in proteome studies. Electrophoresis 18:1068-1072

Görg A, Postel W, Gunther S, Weser J, Strahler JR, Hanash SM, Somerlot L, Kuick R (1988) Approach to staionary two-dimensional pattern: influence of focusing time and immbiline/carrier ampholytes concentrations. Electrophoresis 9:37-46

Gygi SP, Rist B, Gerber SA, Turecek F, Gelb MH, Aebersold R. (1999a) Quantitative analysis of complex protein mixtures using isotope-coded affinity tags. Nature Biotech 17:994-999

Gygi SP, Rochon Y, Franza BR, and Aebersold R (1999b) Correlation between protein and mRNA abundance in yeast. Mol Cell Biol 19:1720-1730

Haezewood JL, Howell KA, Whelllm J, Millar AH (2003) Towards an analysis of the rive mitochondrial proteome. Plant physiol 132:230-242

Hansen KC, Schmitt-Uhlms G, Chalkley RJ, Hirsch J, Baldwin MA, Burlingame AL (2003) Mass spectrometric analysis of protein mixtures at low levels using cleavable ^{13}C-isotope-coded affinity tag and multidimensional chromatography. Mol Cell Proteomics 2:299-314

Hoogland C, Sanchez JC, Walther D, Baujard V, Baujard O, Tonella L, Hochstrasser DF, and Appel RD. (1999) Two-dimensional electrophoresis resources available from ExPASy. Electrophoresis 20:3568-3571

Hunter T (1995) Protein kinases and posphatases: the yin and yang of protein phosphorylation and signaling. Cell 80:225-236

Islam N, Hirano H (2003) Exploring protein–protein interactions in wheat proteome by manipulating genome composition: a perspective. J Electrophoresis 47:33-42

Issaq HJ, Veenstra TD, Cornard TP, Felschow D (2002) The SELDI-TOF MS approach to proteomics: protein profiling and biomarker identification. Biochem Biophys Res Commun 292:587-592

Jonas K, Allan G R, Hakan L, Susanne W (2004) Protein complexes of the plant plasma membrane resolved by Blue Native PAGE. Physiol Plantarum 121:546-555

Kamal AHM, Kim KH, Shin DH, Seo HS, Shin KH, Park CS, Heo HY, Woo SH (2009a) Proteomics profile of pre-harvest sprouting wheat by using MALDI-TOF Mass Spectrometry. Plant Omics J 2:110-119

Kamal AHM, Kim KH, Shin KH, Seo HS, Tsujimoto H, Heo HY, Choi JS, Park CS, Woo SH (2009b) Diversity of Novel Glutenin Subunits in Bread Wheat (*Triticum aestivum* L.). J Plant Biol 52:533-542

Kamal AHM, Park CS, Heo HY, Chung KY, Cho YG, Kim HS, Song BH, Lee CW, Woo SH (2009c) Proteomics Approach on Puroindoline Gene of Pre-harvest Sprouting Wheat. Kor J Breed Sci 41:205-212

Kaufmann H, Bailey JE, Fussenegger M (2001) Use of antibodies for detection of phosphorylated proteins separated by two-dimensional gel electrophoresis. Proteomics 1:194-199

Kislinger T, Cpx B, Kannan A, Chung C, Hu P, Ignatchenko A, Scott MS, Gramolinoi AO, Morris Q, Hallet MT, Rossant J, Highes TR, Frey B, Emili A (2006) Global survey of organ and organelle protein expression in mouse: combined proteomic and transcriptomic profiling. Cell 125:173-186

Klose J (1975) Protein mapping by combined isoelectric focusing and electrophoresis in mouse tissues. A novel approach to testing for induced point mutations in mammals. Humangenetik 26:231-243

Lewis TS, Hunt JB, Aveline LD, Jonscher KR, Louie DF, Yeh JM, Nahreini TS, Resing KA, Ahn NG (2000) Identification of novel MAP kinase pathway signaling targets by functional proteomics and mass spectrometry. Mol Cell 6:1343-1354

MacCoss MJ, McDonald WH, Saraf A, Sadygov R, Clark JM, Tasto JJ, Gould KL, Wolters D, Washburn M, Weiss A, Clark JI, Yates JR 3rd. (2002) Shotgun identification of protein modifications from protein complexes and lens tissue. Proc Nat'l Acad Sci USA 99:7900-7905

O`Farrell PH (1975) High resolution two-dimensional electrophoresis of proteins. J Biol Chem 250:4007-4021

Ong SE, Blagoev B, Kratchmarova I, Kristensen DB, Steen H, Pandey A, Mann M (2002) Stable isotope labeling by amino acids in cell culture, SILAC, as a simple and accurate approach to expression proteomics. Mol Cell Proteomics 1:376-386

Pappin DJ, Hojrup P, Bleasby AJ (1993) Rapid identification of proteins by peptide-mass fingerprinting. Curr Biol 3:327-332

Payne PI, Corfield KG, Blackman JA (1981) Correlation between the inheritance of certain high-molecular-weight subunits of glutenin and bread making quality in progenies of six crosses of bread wheat. J Sci Food Agric 32:51-60

Peltier JB, Friso G, Kalume DE, Roepstorff P, Nilsson F, Adamska I, van Wijk KJ (2000) Proteomics of the chloroplast: Systematic identification and targeting analysis of lumenal and peripheral thylakoid proteins. Plant Cell 12:319-342

Peng Z, Wang M, Li F, Lv H, Li C, Xia G (2009) Aproteomics study of the response to salinity and drought stress in an introgression strain of bread wheat. Mol Cell Proteomics 8: 2676-2686

Rakwal R, Agrawal GK (2003) Rice proteomics: current status and future perspectives. Electrophoresis 24:3378-3389

Ross PL, Huang YN, Marchese JN, Williamson B, Parker K, Hattan S, Khainovski N, Pillai S, Dey S, Daniels S, Purkayastha S, Juhasz P, Martin S, Barlet-Jones M, He F, Jacobson A, Pappin DJ (2004) Multiplexed protein quantitation in Saccharomyces cervisiae using amine-reactive isobaric tagging reagents. Mol Cell Proteomics 3:1154-1169

Schlessinger J (1993) Cellular signaling by receptor tyrosine kinases. Harvey Lect 89:105-123

Schützendübel A, Polle A (2002) Plant responses to abiotic stresses: heavy metal-induced oxidative stress and protection by mycorrhization. J Exp Bot 53:1351-1365

Skylas DJ, Van Dyk D, Wrigley CW (2005) Proteomics of wheat grain. Cereal Sci 41:165-179

The *C. elegans* Sequencing Consortium (1998) Genome sequence of the nematode *Caenorhabditis elegans*. A platform for investigating biology. Science 282:2012-2018

Thiellement H, Zivy M, Plomion C (2002) Combining proteomic and genetic studies in plants. J Chromatogr B Analyt Technol Biomed Life Sci 782:137-149

Unlu M, Morgan ME, Minden JS (1997) Difference gel electrophoresis: a single gel method for detecting changes in protein extracts. Electrophoresis 18:2071-2077

van Wijk KJ (2001) Challenges and prospects of plant proteomics. Plant Physiol 126:501-508

Venter JC, Adams MD, Myers EW, Li PW, Mural RJ, Sutton GG, Smith HO, Yandell M, Evans CA, Holt RA, et al. (2001) The sequence of the human genome. Science 291:1304-1351

Wang Y, Yang L, Xu H, Li Q, Ma Z, Chu C (2005) Differential proteomic analysis of proteins in wheat spikes induced by *Fusarium graminearum*. Proteomics 5:4496-4503

Washburn MP, Wolters D, Yates JR 3rd (2001) Large-scale analysis of the yeast proteome by multidimensional protein identification technology. Nat Biotechnol 19:242-247

Wilkins MR, Sanchez JC, Gooley AA, Appel RD, Hum-

phery-Smith I, Hochstrasser DF, Williams KL (1996) Progress with proteome projects: why all proteins expressed by a genome should be identified and how to do it. Biotechnol Genet Eng Rev 13:19-50
Woo SH, Fukuda M, Islam N, Takaoka M, Kawasaki H, Hirano H (2002) Efficient peptide mapping and its application to identification of embryo proteins in the rice proteome analysis. Electrophoresis2 3:647-654
Woo SH, Higa A, Kimura M, Jong SK, Yamaguchi I (2003) Proteome analysis of wheat lemma. Biosci Biotechnol Biochem 67:2486-2491