8-11-2007

Proteome and phosphoproteome dynamic change during cell dedifferentiation in Arabidopsis thaliana

Brahmananda Reddy Chitteti

PROTEOME AND PHOSPHOPROTEOME DYNAMIC CHANGE DURING CELL DEDIFFERENTIATION IN *ARABIDOPSIS THALIANA*

By

Brahmananda Reddy Chitteti

A Dissertation
Submitted to the Faculty of
Mississippi State University
in Partial Fulfillment of the Requirements
for the Degree of Doctor of Philosophy
in Molecular Biology
in the Department of Biochemistry and Molecular Biology

Mississippi State, Mississippi

August 2007

PROTEOME AND PHOSPHOPROTEOME DYNAMIC CHANGE DURING CELL DEDIFFERENTIATION IN *ARABIDOPSIS THALIANA*

By

Brahmananda Reddy Chitteti

Approved:

Zhaohua Peng Assistant Professor of Biochemistry and Molecular Biology (Director of Dissertation)	Din-Pow Ma Professor of Biochemistry and Molecular Biology (Director of Graduate Studies, Member of Committee)
John A. Boyle Professor of Emeritus (Member of Committee)	Jiaxu Li Assistant Professor of Biochemistry and Molecular Biology (Member of Committee)
Daniel Peterson Assistant Professor of Plant and Soil Sciences (Member of Committee)	Vance H. Watson Dean of the College of Agriculture and Life Sciences

Name: Brahmananda Reddy Chitteti

Date of Degree: August 11, 2007

Institution: Mississippi State University

Major Field: Molecular Biology

Major Professor: Dr. Zhaohua Peng

Title of Study: PROTEOME AND PHOSPHOPROTEOME DYNAMIC CHANGE DURING CELL DEDIFFERENTIATION IN *ARABIDOPSIS THALIANA*

Pages of Study: 161

Candidate for Degree of Doctor of Philosophy

Cell dedifferentiation is a cell fate switching process in which a differentiated cell reverts to a status with competence for cell division and organ regeneration like an embryonic stem cell. Although the phenomenon of cell dedifferentiation has been known for over two and a half centuries in plants, little is known of the underlying mechanisms. Here, the proteome map of *Arabidopsis* cotyledons has been established and investigated the dynamic change of the cotyledon proteome in the time course of cell dedifferentiation. Among the 353 distinct genes, corresponding to 500 2-DE gel protein spots identified with high confidence, 12% have over twofold differential regulations within the first 48 h of induction of cell dedifferentiation. The distributions of these genes among different Gene Ontology categories and gene differential regulations within each of the categories have been examined. In addition, the cotyledon phosphoproteome has been investigated using Pro-Q Diamond Phosphoprotein in Gel Stain followed by mass spectrometry analyses. Among the 53 identified putative

phosphoproteins, nine are differentially regulated during cell dedifferentiation. *Arabidopsis* cotyledon proteome at four different time points after the induction of cell dedifferentiation with MudPIT approach has been investigated and analyzed the protein quantity change using two label-free methods, the Spectral Count (SC) and SEQUEST Cross Correlation Coefficient (ΣXcorr) methods. Among the 662 MudPIT identified proteins, one hundred forty eight displayed differential regulation. The up-regulated proteins include transcription factors, calmodulins, translational regulators, and stress response proteins. The Spectral Count and the cross correlation coefficient quantification results are highly consistent in over 81% of the differentially regulated proteins.

These studies have provided significant new insight into cell dedifferentiation process in *Arabidopsis thaliana* and also enhanced the *Arabidopsis* cotyledon proteome database established using gel based and non gel based methods. The results show that cell dedifferentiation involves extensive protein quantitative and qualitative changes in almost every cellular compartment and cellular process. Proteins like 14-3-3 proteins, Translational controlled tumor protein (TCTP) and its possible interaction protein-Translational elongation factor eEF1 alpha chain, GTP binding nuclear protein RAN2, GTP binding protein SAR1B and several other hypothetical and expressed proteins and nine other phosphoproteins showed significant differential expression during early dedifferentiation. Deciphering the molecular mechanisms regulating the cellular dedifferentiation certainly enhances the understandings and mechanisms of reprogramming all types of differentiated cells including animal cells.

DEDICATION

I would like to dedicate this work to my parents Chitteti Lakshmi and Bhooma Reddy and my grand father late Chitteti Ganga Reddy.

ACKNOWLEDGEMENTS

I would like to express my humble gratitude to my major Professor Dr. Zhaohua Peng for his continuous encouragement and support. Without his technical and critical advice, this work would not have been completed. His encouragement made me to involve in various aspects of research that gave me multi faceted experience in conducting independent research. I am thankful to him forever.

My humble thanks also goes to Dr. Din-Pow Ma for serving as a co-major Professor on my committee and for letting us to use his lab facilities and his kind help is critical throughout my graduate study. I am very thankful to Dr. Boyle, whose continuous support fostered me throughout my graduate work and his unconditional help making me to grow in my scientific career. I am also thankful to Dr. Jiaxu Li and Dr. Daniel Peterson for serving on my committee.

I am very grateful to Tibor Pechan, Senior technician at Life Sciences and Biotechnology Institute (LSBI), for his help with Mass Spectrometry analysis which is very critical for this dissertation. I am also thankful to Juliet Tang at LSBI who taught me iTRAQ methodology for quantitative Proteomics. I owe a special note of gratitude to Dr. Susan Bridges, Department of Computer Science, and her student Bryce Magee for their collaboration in analyzing the non labeled Proteome data.

I am also thankful to Dr. Peng's lab members – Dr. Guosheng Li, Dr. Kyu-Ri Choi, Aparna Kakani, Feng Tan, Babi Ramesh Nallamilli and Hui Wang for their

friendship and understanding. I am grateful to my friend Udaya Bhaskar Gunturu for his help in diversified technical aspects of the research.

Last, but not least, I would like to thank my family for supporting all my endeavors in life. I owe to my sister Shashikala, Brother Pratap Reddy for their endless love and affection.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER I

INTRODUCTION

Cellular differentiation is a concept in Developmental Biology describing the process by which cells acquire specialization. The series of changes that occur in cells during development permits a greater variety of organs. During this process the morphology of a cell changes dramatically in numerous aspects such as size, shape, polarity, metabolic activity etc., but the genetic material remains the same, with few exceptions like epigenetic modifications. This acquired or assigned cell fate is memorized in subsequent growth and cell divisions.

A cell that is able to differentiate into all cell types ultimately giving the new entire organism is known as totipotent. In mammals, only the zygote and early embryonic cells are totipotent, while in plants, many differentiated cells can display totipotency by undergoing dedifferentiation followed by redifferentiation. Dedifferentiation is a reversal process to cellular differentiation where regression of a differentiated cell into embryonic, unspecialized stem cell state. It may occur before the regeneration of appendages in plants, animals and in the development of some cancers. When a plant cell undergoes dedifferentiation, the cell fate memory is erased, the genome is reprogrammed, totipotency is recovered, and callus production occurs. Production of callus from differentiated cell is always associated with the dedifferentiation phenomenon. These

dedifferentiated cells are now able to redifferentiate into other organs/tissues. In vivo cell dedifferentiation is related to wounding damage repair in both plants and animals.

However, it is still obscure at the molecular level what are the changes in the genome and how it is reprogrammed in such a short period of time. Studying dynamic changes in the proteome certainly provide critical insight into the molecular mechanism of cell dedifferentiation. In addition, studies on cell dedifferentiation may serve as an alternative approach for investigating cell differentiation.

Due to its small size and short life cycle (six weeks from seed to seed), Arabidopsis was chosen as an ideal model experimental organism to study the dedifferentiation phenomenon in this part of the research. *Arabidopsis* has five chromosomes (2n=10) and has a haploid genome size about 125Mbp. Its complete genome was sequenced by *Arabidopsis* Research Group in December 2000 along with plastid and mitochondrial genomes.

In the course of cell dedifferentiation, the cells remove the entire proteins specific to the previous metabolic status and reprogram the genome to gain cell division and totipotency competence. Genomics could provide the data of all possible proteins that could present in a cell, but only proteomics gives the information what proteins are really present under that particular biological condition. Because, many studies have shown there is poor correlation between mRNA and protein expression levels. In order to study the proteome dynamics, traditional proteomic methods, Two-Dimensional Gel Electrophoresis (2-DE) and Multidimensional Protein Identification Technology (MudPIT) have been employed.

In 2-DE, proteins are separated according to their pI values followed by molecular mass. All interesting protein spots are then excised from the gel and proteolytically digested and mass analyzed using Matrix assisted laser desorption ionization Time-of-Flight (MALDI-TOF). MudPIT is a complex protein/peptides separation and identification technique. It separates peptides in liquid chromatography, so that the separation is directly linked to the mass spectrometer. It uses strong cation exchange (SCX) and reversed phase (RP) columns. The chromatography proceeds in cycles each comprising an increase in salt concentration to "bump" peptides off of the SCX column followed by a gradient of increasing hydrophobicity to progressively elute peptides from the RP column into the ion source. The eluted peptides will be subjected to Collision-Induced Dissociation, followed by recording the fragment ions in a tandem mass spectrometer. The spectral data is matched against the protein databases using SEQUEST algorithm.

The main goal of my research is to understand the dynamic change of proteome during the cell dedifferentiation in *Arabidopsis thaliana*. The proteins or protein complexes involved in the dedifferentiation process are identified and these proteins and their corresponding genes have been characterized using Gene Ontology tools.

The main objectives of this research are:

1. The global proteome of cotyledon of *Arabidopsis thaliana* will be generated using Two-Dimensional Gel Electrophoresis (2-DE) and Multidimensional Protein Identification technique (MudPIT).

2. To investigate the process of cell dedifferentiation in *Arabidopsis thaliana*, the dynamic changes of the total proteome at different time points - 0, 12, 24, and 48hr will be examined and corresponding genes will be characterized according to Gene Ontology rules.

3. Post translational modifications of proteins are crucial for the regulation of cell functions. Among numerous protein modifications, phosphorylation is well known to play a key role in many cellular processes. Phospho- proteome of *Arabidopsis* cotyledon cells and its regulation during cell dedifferentiation are studied using phosphoprotein specific binding stains.

4. The differential regulation of the MudPIT analyzed proteins are studied and compared with the proteins identified by 2-DE.

CHAPTER II

LITERATURE REVIEW

Totipotency

The ability of a single cell to divide and produce all differentiated cells and tissues, thereby, giving an individual organism is called totipotency. Totipotency is one of the main properties of plant cells, which ensures a small part of the plant such as leaf, root, or pollen etc., to grow as an entire plant on the appropriate culture media under *in vitro* conditions. Practically any living cell with a nucleus, after the process of dedifferentiation, can initiate the cell proliferation turning into undifferentiated callus into shoots, roots or an entire plant (Ezhova, 2003). It is evident that the study of genetic control of totipotency is important for both development and improvement of plant biotechnology and also for better understanding of this most important characteristic of plant cells (Ezhova, 2003). In case of human development, in the first hours after fertilization, a fertilized egg creates a single totipotent cell and this cell divides into identical totipotent cells. Approximately four days after fertilization and after several cycles of cell division, these totipotent cells begin to specialize. They specialize into pluripotent cells that can give rise to most, but not all, of the tissues necessary for embryo development. Pluripotent cells undergo further specialization into multi potent cells that are committed to give rise to cells that have a particular function. Thus, totipotent cells

are known as zygotes (Salo, 2006). This totipotency phenomenon is present both in plants and animals.

This remarkable ability to generate a complete organism from a single cell has fascinated biologists for more than a century. What are the molecular mechanisms that underlie totipotency? Are these mechanisms unique to the germline or do they also exist in somatic cells? Does gamesomeness involve mechanisms fundamentally different from those driving somatic differentiation? These questions are still abscure (Seydoux and Braun, 2006). The molecular mechanisms that maintain totipotency are also not well understood (Ciosk et al., 2006).

Two biologists, Schleiden in 1838 and Schwann in 1839, concluded from their work that a plant cell is capable of autonomy and is totipotent. Except terminally differentiated cells, all other cells have an ability to redisplay their full genetic program, leading to formation of new plants (Alberts et al., 1994). The concept of totipotency includes a two-step process. The first is acquiring stem cell morphogenic potential in response to stimuli, called dedifferentiation, and the second is expressing the potential during morphogenesis called redifferentiation/regeneration. There is another phenomenon, where a particular cell type switches into another cell type, called transdifferentiation. It is believed that dedifferentiation might be an intermediate step during transdifferentiation. These three phenomenons capture most of the stem cell spotlight (Grafi and Avivi, 2004). Throughout this dissertation the main focus is on cellular dedifferentiation phenomenon.

Dedifferentiation

In higher organisms performing sexual reproduction, the complicated molecular structure is derived from a single fertilized cell - zygote. The developmental process requires progressive cell division and cell differentiation. During differentiation, each cell type is assigned to a specific role and only expresses a set of genes that match with its function. The differentiated cells are capable of maintaining their expression profiles and have their expression profiles inherited in subsequent growth and mitotic divisions. The differentiated living cells with an intact nucleus can loose their cell fate memory completely and regain an ability to display their full genetic program under certain environmental stimuli. This phenomenon is called dedifferentiation. In plants, the newly acquired cell division activities of dedifferentiated cells usually lead to the formation of amorphous, undifferentiated cell mass called callus (Bhojwani and Razdan, 1983). Therefore, callus induction is often used to represent cell dedifferentiation. At molecular level, the genome undergoes reprogramming to restore the stem cell status, and then the genes in the reprogrammed genome are expressed following the orderly pattern of zygotic cell during regeneration.

Callus formation was initially observed by Seigneur du Monaceau in the middle of eighteenth century. By removing a small ring of cortex from an elm tree, he observed the development of a swelling above the area of decortications, while buds developed on the lower part. Another aspect is meristematic cells from which shoots and roots are produced throughout the life. These meristematic cells acquire new fates, thereby, giving the new organs of the plant. An entire plant can be regenerated *in vitro* from a single somatic cell (Vasil and Vasil, 1974). The generation of embryos from cells rather than

zygote reveals the ability of totipotency in plant cells. The full zygotic potential of somatic plant cells is observed in some species where shoots and roots are generated from the margins of the leaf (Kerstetter and Hake, 1997). It has long been believed that wounding the plant also triggers dedifferentiation, and during this process rapid expression of the cdc2aAt gene which encodes PSTAIRE CDK of *Arabidopsis* has been found (Hemerly et al., 1993).

Recently, few genes have been identified which are thought to be involved in cell dedifferentiation. Polycomb group (PcG) proteins are highly conserved regulatory factors responsible for transcriptional silence through cell divisions (Ringrose and Paro, 2004; Klebes et al., 2005; Lee et al., 2005). Polycomb repressive complex PRC1 and PRC2 groups have been identified in animal cells. It has been found that PRC1 is responsible for locking chromatin in repressive state that is for long term cellular memory. PRC2 is responsible for depositing epigenetic marks to be identified by PRC1 (Ringrose and Paro, 2004). These PCG proteins are conserved between animals and plants and it is believed that PCG proteins in plants involve in dedifferentiation. Mutations in PRC2 showed the reorganization of heterochromatin in *Arabidopsis* (Schubert et al., 2005). The double mutant curly leaf swinger (clf swn) forms masses of undifferentiated cells on the plantlet tissues, thereby, giving somatic embryos. This clf swn is homologues to Zeste E(Z), a PRC2 component in Drosophila (Chanvivattana et al., 2004). Another outstanding group in plants is the MADS box protein family, which consists of more than 80 members and regulates not only meristem and floral organ identities but also the timing of flowering and cell-type specification in floral organs (Jack, 2001). AGAMOUS is a MADS box protein involved in floral organ identity control. The FACIATA genes play a critical role

in the organization of shoot and root epical meristems during polyembryonic development. LEAFY is another key transcription regulator which interacts with a putative chromatin remodeling ATPase (Wagner and Meyerowitz, 2002).

Dedifferentiation in other organisms

Cellular dedifferentiation is very obvious in other organisms also. For example, a unicellular slime mold, *Dictostelium discoideum,* switches from unicellular to multi cellular depending on food availability. Under starving conditions, undifferentiated single cells aggregate into multi cellular organism which can find food and it is able to dedifferentiate back to unicellular state. Another example is that a somatic cell of hydra gives rise to a new individual. Somatic cell nuclear transfer (SCNT) is an obvious demonstration of dedifferentiation. These SCNT experiments revealed that differentiated cells keep the same genome, but they reprogram it in order to undergo dedifferentiation and to acquire new fate and it has been also demonstrated that the somatic embryos acted same as the zygotic embryos (Briggs and King, 1952).

External injury or positional information induces transdifferentiation thus creating entire missing part of the body in many amphibians. For example, neural cells can transdifferentiate into muscle and cartilage during tail regeneration (Echeverri and Tanaka, 2002). A newt is capable of regenerating its limbs, spinal cord, heart ventricle, tail, retinas, eye lenses, and upper and lower jaws (Becker et al., 1974; Davis et al., 1990; Brockes, 1997). Teleost fish can regenerate their fins and spinal cord (Zottoli et al., 1994; Johnson and Weston, 1995). The molecular mechanisms in these two organisms are also not clearly understood. But, it has been observed that epithelial cells begin migrating

across amputation site, forming wound epithelium (WE) within 24 hours and start making apical epithelial cap (AEC) (Christensen and Tassava, 2000). The internal stump cells underlying the WE-AEC begin to dedifferentiate in response to undefined signals found in the early limb regeneration (Thornton, 1957; Bodemer, 1959; Hay and Fischman, 1961; Thornton and Thronton, 1965; Steen, 1968; Lo et al., 1993; Kumar et al., 2000). These dedifferentiated cells then proliferate to give progenitor and pluripotent cells, which will later redifferentiate to form the regenerated limb (McGann et al., 2001).

Cell fate switch is also observed recently in Drosophila in imaginal discs – larval structures in which cells are destined to give rise to defined body parts in the adult. After mechanical fragmentation of the imaginal discs, some cells switched from leg fate to a wing fate (Maves and Schubiger, 2003). In mammals transdifferentiation is seen in the liver (Taub, 2004) and in the Schwann cells of the peripheral nervous system (Harrisingh et al., 2004). Two research groups were able to produce mice from the nuclei of olfactory neurons. During this process, fully differentiated olfactory neuron cells completely lost their cell fate memory, and redifferentiated to a complete mouse (Eggan et al., 2004; Li et al., 2004).

With an efficient dedifferentiation process, it is conceivable that healthy, abundant and easily accessible adult cells could be used to generate different types of functional cells for the repair of damaged tissues and organs (Ding and Schultz, 2005). The research group at The Scripps Research Institute screened about 50000 compounds and found a purine analogue called ‘reversine’ which was found to induce dedifferentiation activity (Chen et al., 2007). It has been assumed that reversine inhibits myotube formation and treated myoblasts continue to grow to form a confluent culture of

mononucleated cells, which can redifferentiate into osteoblasts and adipocytes upon exposure to appropriate differentiation conditions (Warashina et al., 2006).

Another research group identified a glycol protein from serum extract with unknown function that can trigger cell cycle re-entry from the differentiated state (Straube et al., 2004). Oct-4 plays an essential role as a central regulator of the undifferentiated state. Recently, it has been demonstrated that Oct-4 also has the ability to reprogram committed somatic cells, inducing their dedifferentiation. So, Oct-4 might be the master regulator of the pluripotent state in mammalian cells (Buitrago and Roop, 2007). All these examples indicate the ability of reprogramability among vertebrates (Briggs and King, 1952; Gurdon, 1962; Campbell et al., 1996).

Genetic material reprogramming during dedifferentiation is usually associated with chromatin reorganization (Byrne et al., 2003; Gonda et al., 2003; Weimann et al., 2003). During the protoplast culture, upon removal of cell wall, chromatin decondensation and disruption of nucleolus (Grafi, 2004) have been observed.

Chromatin role in cellular dedifferentiation

Chromatin is an essential structure in eukaryotic cells. It not only stores the genetic information but also controls when, where and how a gene is expressed by working together with its interacting proteins.

It has long been assumed that epigenetic mechanisms that take place on chromatin are somehow linked to cellular dedifferentiation and regeneration. Chromatin modifications such as Lysine at the amino terminal ends of the core histones are the primary sites for reversible post translational modifications, including acetylation,

methylation, phosphorylation, ubiquitination, and ADP-ribosylation (Richards and Elgin, 2002; Strahl et al., 2002). Still we don't know exactly how the genome reprogramming takes place because of these epigenetic changes. Though, several remodeling factors and complexes involving these processes are identified (Cairns, 2005; Hsieh and Gage, 2005; Bultman et al., 2006).

The DNA of active genes is preferentially associated with highly acetylated histones while DNA of inactive genes is associated with hypo acetylated histones. The modification of positively charged lysine residues provides a powerful device to unfold chromatin for gene expression (Wolffe and Pruss, 1996). Heterochromatin is characterized by methylation of cytosine nucleotides of the DNA, the methylation of histone H3 at lysine 9, and the specific binding of hetero chromatin protein 1 (HP1) to methylated H3 lysine 9. It has been shown that deacetylation of lysine 9 at the amino terminus of H3 is a prerequisite for methylation of this same lysine. Methylation of H3 lysine 9 in turn recruits the binding of HP1 that helps to establish highly compacted and transcriptionally inactive heterochromatin (Rice and Allis, 2001).

Recently, it has been shown that two rounds of chromatin decondensation are required for tobacco cells to undergo dedifferentiation and re enter the cell division (Zhao et al., 2001). The first phase takes place during enzymatic digestion of cell wall in the course of protoplast isolation; whereas, the second decondensation occurs only after protoplasts are induced with phytohormones for re-enter of the cell cycle (Grafi, 2004). In the absence of hormonal application, protoplasts undergo cycles of chromatin condensation/decondensation and die (Zhao et al., 2001). There is increasing evidence

that chromatin components are the key players in establishing and maintaining the spatial and temporal gene expression profile in both plants and animals (Meyerowitz, 2002).

In the presence of Auxin and Cytokinin, about 10% of protoplasts approach S phase within 48 hours and 30-40% cells approach S to G2 phase within 72 hours. Further studies in tobacco revealed that condensation of 18S rDNA accompany with acquisition of competence for cell fate switch, suggesting that changes in chromatin during cell fate switch may be better defined as chromatin reorganization instead of just decondensation (Williams et al., 2003). In addition, the ubiquitin proteolytic system was found to be indispensable for protoplast progression into S phase, being required for the second but not the first phase of chromatin decondensation (Zhao et al., 2001). A specific inhibitor of the 26S proteosome, MG132, has no effect on the occurrence of the first phase chromatin de condensation but interferes with the entry of protoplasts into S phase (Zhao et al., 2001). In agreement, it has been reported that genes coding for ubiquitin are induced during cell dedifferentiation in tobacco. Recently, hypomethylation of ribosomal RNA gene is found to be associated with cell dedifferentiation in tobacco cell (Koukalova et al., 2005). Although these studies have provided some insight into plant cell dedifferentiation, it is still unknown what kind of changes are associated with dedifferentiation at the protein level.

Transcription in the embryonic germline appears to be regulated at the level of chromatin modification (Seydoux and Braun, 2006). For example, in *Drosophila* and *C. elegans*, early germ cells have reduced levels of H3-K4me2 (dimethylation of lysine 4 on histone H3), a methyl mark linked with transcription (Schaner and Kelly, 2006). No signal transduction pathway components that neither perceive the environmental stimuli

nor mediate the genome reprogramming process have been identified. Recent progress in genome technologies made it possible to examine the cell dedifferentiation course in a large scale. Compared to other genomic methods, proteomics is unique in that it is capable of illustrating the course of removal of existing proteins and identifying possible post translational modifications associated with dedifferentiation.

Overall, dedifferentiated plant cells might be more similar to pluripotent animal stem cells in their ability to continuously perceive extra cellular signals and in maintaining a chromatin organization that allows a fast response to the signals. Understanding the molecular mechanism regulating the cellular dedifferentiation in plant cells could provide a new perspective on cellular dedifferentiation in animal cells (Costa and Shaw, 2006). In this study, the model organism *Arabidopsis thaliana* was used to explore the cell dedifferentiation mechanism.

Arabidopsis thaliana

Arabidopsis thaliana, commonly called *Arabidopsis*, Thale cress, or mouse-ear cress, is a small plant belonging to Brassicaceae (Mustard or Crucifer) family. Currently the genus *Arabidopsis* has nine species and eight subspecies recognized. This plant was named by a scientist named Hasno Johnson. Although *Arabidopsis thaliana* has little direct significance for agriculture, it has several advantages that made it the model for understanding the Genetic, Cellular, and Molecular Biology of flowering plants. It is most extensively used species for over forty years in research for identifying genes and in the field of plant functional genomics (Schmidt et al., 1995; Bevan et al., 1998). It is 15-20 cm high when matured and seeds are about 0.5 mm long. It can grow on Petri dish

with appropriate media in the lab even up to 1000 seedlings per plate and can produce over 5,000 seeds per plant. Because of *Arabidopsis*' small size and short life cycle (six weeks from seed to seed) it was chosen as an ideal model experimental organism by plant biologists.

The earliest non-taxonomic appearance of *Arabidopsis* in the literature appears to be a paper by Alexander Braun in 1873, describing a mutant (*AGAMOUS* gene) plant found in a field near Berlin (Braun, 1873; Meyerowitz, 2001). Friedrich Laibach published the chromosome number of *Arabidopsis* in 1907. The *Arabidopsis* research community started in 1964 with a newsletter called *Arabidopsis* Information Service (AIS). The first International *Arabidopsis* Conference was held in 1965, in Göttingen, Germany. In the 1980s *Arabidopsis* started to become widely used in plant research laboratories around the world. The breakthrough year for *Arabidopsis* as the preferred model plant came in 1986 when T-DNA mediated transformation was first published (An et al., 1986) and this coincided with the first gene to be cloned and published (Chang and Meyerowitz, 1986).

Arabidopsis has five chromosomes (2n=10) and has haploid genome size about 125Mbp of DNA. Its complete genome was sequenced by *Arabidopsis* Research Group in December(2000) along with plastid and mitochondrial genomes sequence. It was the first flowering plant genome sequenced and deciphered the genes necessary for a plant to function. Though, only one third of the genes' functions are hypothesized. It is estimated that the *Arabidopsis* genome has ~25,000 genes representing 11,000 gene families. All the above mentioned properties made us to choose *Arabidopsis* as a material to conduct cell dedifferentiation studies.

Callus induction

Plant cell uses external signals to differentiate and to maintain or to change the differentiated state. A cell-cell signaling and positional information strategy seem to play a key role in plant development (Bai et al., 2000). It has been well established that phytohormones (or plant growth regulators, or PGRs), such as auxins and cytokinins play a crucial role in cell dedifferentiation in plants (Skoog and Miller, 1957). These PGRs are signal molecules produced at specific locations, and cause altered processes in target cells at other locations. When the conditions of the auxin were correctly known, the first success of callus formation was obtained by Ball (Ball, 1950) with *Sequuoia.* Later Skoog and Miller extracted a substance from yeast called kinetin, which in the presence of indole acetic acid, induced proliferation of tobacco cells and bud formation. The bud promoting properties of Kinetin suggested a broad interaction of this substance with auxin. Innumerable experiments after 1951 revealed that multiple factors control organ formation. Phytohormones especially Auxin and Cytokinin and other hormones such as Gibberellic acid, Abscisis acid, and ethylene seem to play a key role in differentiation and in dedifferentiation. Skoog and Miller indicated that differentiation of tobacco tissue in culture was determined by the cytokine/auxin ratios in the medium. The absolute concentrations of the two hormones are not critical (Skoog and Miller, 1957; Krikorian, 1995). It reveals that the concentrations can be in an appropriate range. When the ratios are kept at one, a callus mass is produced and the callus continuously proliferate. When the level of auxin relative to that of cytokinin is high, root forms. In contrast, shoot forms when the level of cytokinin relative to that of auxin is high. The device of adjusting auxin/cytokinin ratios to induce shoots and roots is now a well established practice for a

variety of plants in both research and industry. Although callus formation and cell dedifferentiation have been discovered over two centuries ago, very little is known about molecular mechanisms underlying the process.

Proteomics - a versatile tool in life sciences

Proteomics is defined as the systematic study and characterization of proteins, particularly their structure, functions, quantitative and qualitative analyses at different expression levels. Proteomics, the term is coined in 1994 analogous with Genomics. Genomics could provide the data of potential proteins that could present in a cell, but only proteomics gives the information what proteins are really present under that particular biological conditions. Many studies have shown that there is poor correlation between mRNA and protein expression levels (Anderson and Seilhamer, 1997; Gygi et al., 1999a). Based on micro array experiments studies, it has been noticed that the expression level of about 50% of genes do not match with the protein studies. Therefore characterization of proteins and their expression levels at a given time under different cellular conditions have to be studied as a separate branch.

Proteomics has become a vital tool in the field of functional genomics, because only proteins are directly related to the functions. In *Arabidopsis*, tremendous progress has been made in the past few years in generating large sets of data to understand the protein-protein interactions, in elucidating global and organelle proteomes, and in hypothesizing the gene functions. Sensitive technological improvements, studying the sub cellular proteomes, and pooling the proteomic data together from across the world and making it open access to all researchers etc., are necessary to fulfill the potential of

proteomics (Tyers and Mann, 2003). Applications of Proteomics lays a new path to analyze and characterize the complex functions of proteins at different levels in an organism or cell (Wilkins et al., 1996). The most extensively used method in the field of proteomics is Two-Dimensional Gel Electrophoresis (2-DE), developed by O' Farrell (O'Farrell, 1975). Present proteomics research aims at both identifying new proteins in relation to their function and ultimately explaining how their expressions are controlled within regulatory networks (Xi, 2006).

Gel based protein separation methods

The traditional approach for proteomic analyses is separating protein mixtures on Two-Dimensional Polyacrylamide Gel Electrophoresis (2D-PAGE) gels followed by characterization of individual proteins/peptides by Mass Spectrometry (MS). 2-DE takes place in two steps: Isoelectric focusing as first step and Gel Electrophoresis as second step as shown in the flow chart (Figure 2.2). In Isoelectric focusing the pI is the pH at which a protein do not migrate in an electric field. When a protein is placed in a medium with a pH gradient and subjected to an electric field, it initially moves towards the electrode with the opposite charge. During migration through the pH gradient, the protein pick up or lose protons. As it migrates, the net charge and the mobility will decrease and the protein will slow down. Eventually the protein will arrive at the point in the pH gradient which is equal to its pI and stops. A molecule with a net charge can move on a solid support under electric field, this phenomenon is termed as electrophoresis. This powerful technique can be employed to separate proteins based on their molecular mass and charge. The velocity (ν) of a charged molecule on a solid support depends on the

electric field strength (E), the net charge (Z) on the molecule, and the frictional co-efficient (f). $\nu = ZE / f$

Recent significant advances in proteomics made it possible to distinctly separate thousands of proteins with higher resolution and with higher reproducibility. These advances include immobilized pH gradients in the first dimension (Gorg et al., 2000), multiple gel running chambers, and accurate image analysis software etc. 2D gel that contains more than 10,000 polypeptide spots of mouse tissue has been obtained (Klose and Kobalz, 1995). Sample preparation is critical for better separation of proteins on a gel. Sample has to be denatured, disaggregated, and reduced to ensure each spot represents an individual polypeptide (Meyer et al., 1988). In order to detect the low abundant proteins, it is better to prefractionate the high abundance proteins like Ribulose-1,5-bisphosphate carboxylase/oxygenase (Kim et al., 2001). Prefractionation methods include sequential extraction with increasingly stronger solubilization solution, sub cellular fractionation, and size prefractionation (Pasquali et al., 1999). 2-DE is not appropriate to separate membrane, hydrophobic, and highly basic proteins efficiently (Gygi et al., 2000). To overcome these limitations, increasing attention is being focused on an alternative approaches such as MudPIT, protein chips etc., (Service, 2001).

Several types of dyes are currently used to visualize the protein/polypeptide spots on gels. Coomassie Brilliant Blue (CBB) and silver staining are commonly used dyes. However, the binding of these dyes to the amount of protein is not in linear ratio. Low abundant spots are not detected by CBB due to low sensitivity. Though, silver staining is very sensitive, there is nonspecific binding of the dye to gel and it is difficult to stop the chemical reaction on multiple gels at the same time. The recent development of

fluorescent dyes such as SYPRO Ruby overcomes these limitations (Patton, 2000). The binding of SYPRO Ruby to the amount of protein is in linear ratio. Besides these dyes, Phosphoproteome specific dye, Pro-Q Diamond phosphoprotein gel stain, (Schulenberg et al., 2003) specifically bind to phosphoproteins and Pro-Q Emerald glycoprotein gel stain which specifically bind to Glycoproteins (Hart et al., 2003). These dyes are currently commercially available (Molecular Probes).

Another advanced 2-DE method is Differential in Gel Electrophoresis (DIGE) (Unlu et al., 1997). 2-D DIGE is an approach where protein samples are labeled with spectrally distinct fluorescent dyes. Up to three labeled proteins are mixed and resolved simultaneously on the Isoelectric focusing strips followed by vertical SDS-PAGE Gel Electrophoresis. Advantages of this technique include: Variation between the gels that are being analyzed is removed. Relative concentration of the protein spots in a gel will be unchanged, because protein loss during the entire process will be the same for proteins on the same gel. Moreover, gel documentation, matching and quantification etc., becomes easier. Highly sensitive CyDye fluors can also detect low abundance proteins effectively, and no additional dyes are required to visualize the protein spots on the gel.

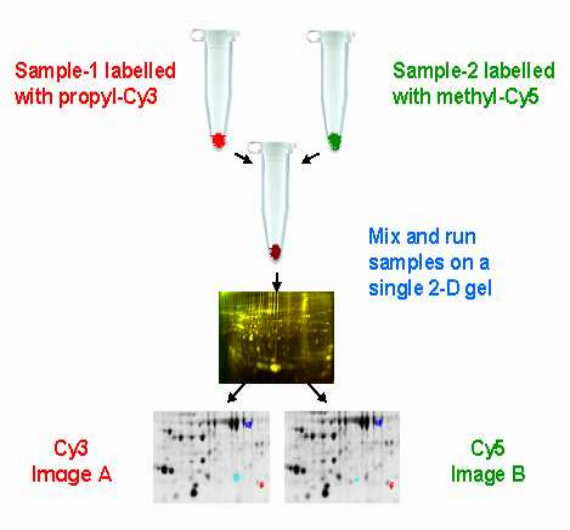

Figure 2.1: Graphical representation of DIGE method.

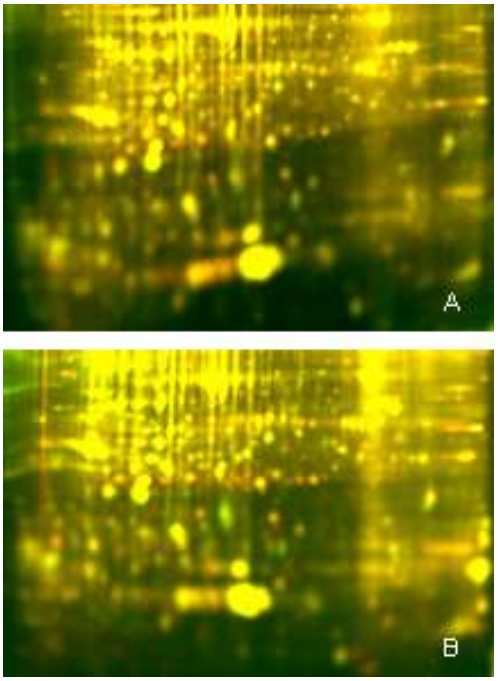

Figure 2.2: Illustration of CyDye labeled DIGE gels.
(Chitteti et al., 2004)

Note: Figure A: 12hr sample was lebeled with Cy5 (red) and 0hr sample was labeled with Cy3 (green). Figure B: 24hr sample was labeled with Cy5 (red) and 0hr sample was labeled with Cy3 (green). 0hr sample specific proteins appear in Red and 12 and 24hr samples appear in green. Common proteins are in yellow color.

After running the sets of gels, it is necessary to analyze all the detected spots for their quantification and to see the expression differences. A number of software programs are commercially available for 2-DE pattern image analysis such as PDQuest 7.3.1 (Bio-Rad Laboratories, Hercules, CA, USA), Image Master 2D Platinum 5.0 (Amersham Pharmacia Biotech, Uppsala, Sweden), Phoretix2D (Nonlinear Company, UK), and DeCyder to analyze the DIGE images. These 2D images of different organelles or different samples can be posted or compared with other database images such as SWISS-2D-PAGE established in 1993 and can be accessed through ExPASy proteomics server (Hoogland et al., 2000). All interesting protein spots can be excised manually or with the help of spot cutters and subjected for mass analysis. The robotic spot picker cuts the spots from the gel very precisely in light mode and UV mode at speed of approximately 100 spots/hour without user intervention. The Investigator ProPrep is a fully integrated system for automated in-gel enzymatic digestion of proteins. The instrument automates sample cleanup procedures of the peptide fragment samples.

Mass spectrometry had great advantage over other analytical techniques because it measures an intrinsic property of a bio molecule, its mass/charge ratio (m/z) with very high sensitivity up to femto mole range. MS based methods usually identify proteins by analyzing the peptides derived from proteolytic digestion of proteins. Usually, a protein sample is digested with sequence specific trypsin which cleaves proteins at lysine (Lys) and arginine (Arg) residues. After several hours of incubation, proteolytic peptides are extracted and concentrated and subjected to MS analysis. The essential step in mass spectrometry analysis is converting an analyte molecule into gas phase ionic species followed by separation of these molecular ions and their charged fragments according to

their m/z ratio. The ion current is detected by a detector and displayed in the form of a mass spectrum. These masses of peptides measured by mass spectrometer are compared with theoretical masses of peptides from all known protein sequences present in a databases such as SwissProt, TReMBL, and NCBI etc., (Pappin et al., 1993). The process of comparing the experimental peptide masses with theoretical peptide masses and identifying a protein is called 'Protein Finger Printing'. Alternatively, if an organism's genome is not completely sequenced, the protein identification can be carried out by de novo sequencing the short amino acid sequences by tandem MS (MS/MS) (Mann et al., 2001). These sequences are called peptide sequence tags and these can be compared to the protein or EST (Expressed sequence tags) data bases.

Non Gel based protein separation methods

In order to separate large samples and for automatization, non gel based techniques like MudPIT and protein chip methods are getting popularity (Wolters et al., 2001). The revolutionary development in mass spectrometry techniques involving Matrix Assisted Laser Desorption Ionization Time-of- Flight Mass Spectrometry (MALDI-TOF-MS) (Karas and Hillenkamp, 1988), Electro Spray Ionization – Tandem Mass Spectrometry (ESI-MS/MS) (Fenn et al.), and multidimensional separation strategies in conjunction with automated ESI tandem mass spectrometry (2D LC MS/MS) for analyzing complex protein mixtures (Opiteck et al., 1997; Peng and Gygi, 2001) are substituting 2-DE approaches (Gerber et al., 2003). In this way, total protein mixture can be sequenced in a single run, and this method allows detection of low abundance proteins such as transcription factors, hormone receptors, etc. The 2-DE limitations such as poor

separation of high abundant proteins and absence of very low abundant proteins, difficulty in separation of very acidic and basic proteins, and gel artifacts are almost overcome by adopting non gel based separation methods.

MudPIT

MudPIT (Multidimensional Protein Identification Technology) technique is used for the separation and identification of complex protein and peptide mixtures using liquid chromatography and it is directly connected with the ion source of a mass spectrometer (Link et al., 1999; Washburn et al., 2001).

All proteolytically digested peptides will be separated on strong cation exchange (SCX) column. It is a biphasic liquid chromatography column consisting of SCX material back-to-back with reversed phase (RP) C18 material inside fused silica capillaries. The chromatography proceeds in cycles. As the first step, increasing concentration of salt is used to free peptides from the cation-exchange resin after which they bind to a reversed phase resin. In the second step, a gradient of increasing hydrophobicity is used to progressively elute peptides from the RP into the ion source. The ionized peptides undergo Collision-Induced Dissociation, and the fragment ions are recorded in a tandem mass spectrometer. The spectral information is matched with databases using SEQUEST algorithm and protein identification is done by DTASelect algorithm (Link et al., 1999; Yates et al., 1999; Washburn et al., 2001; Tabb et al., 2002). Quantifying the identified proteins and studying the expression profiles among the experimental samples are the limiting factors in using MudPIT. But, this can be overcome by adopting ICAT and iTRAQ techniques.

Isotope-coded Affinity Tagging (ICAT)

ICAT utilizes stable isotope labeling to perform quantitative analysis of protein samples, followed by separation and identification of protein mixtures by LC/MS. Tags usually consists of biotin tag (used during avidin affinity chromatography), a reactive group and deuterated linker (heavy or light). These tags are covalently bound to sulfhydryl groups of cysteine residues within a protein. In order to quantify or compare samples, two protein mixtures are labeled with heavy and light deuterium tags (d_8 and d_0 respectively). These two samples are combined and digested to peptides with trypsin, and the ICAT labeled peptides will be isolated utilizing the biotin tag. These peptides will be separated by micro capillary – high performance liquid chromatography and measured quantitatively by comparing peaks from scanning mass spectrometry. The identification of these proteins will be accomplished by switching the instrument to MS/MS mode and the protein identification is done by searching database (Gygi et al., 1999b; Patterson, 2000; Han et al., 2001). After labeling, the separation can also be done on standard 2-DE gels (Smolka et al., 2002). The limitation of this technique is that it requires a free cysteine group on the protein.

iTRAQ

iTRAQ (Isotope tags for Relative and Absolute Quantification) is the most recent gel free quantitative method based upon chemically tagging the N-terminus of peptides generated from protein digests. Each peptide mixture from four samples can be labeled with four isobaric tags termed 114, 115, 116, and 117, respectively. The labeled samples are then combined, fractionated by nano liquid chromatography and analyzed by tandem

mass spectrometer. Protein identification is done by searching databases using the fragmentation data generated by the peptides. Fragmentation of the tag attached to the peptides generates a low molecular mass reporter ion that is unique to the tag. The relative quantification is determined by comparing the peak intensities of same peptide from four samples, and fold difference is calculated (Chong et al., 2006; Gan et al., 2007; Wiese et al., 2007).

Protein Arrays

Protein arrays (also biochip, protein chip) are revolutionizing the field of proteomics. A protein chip is a piece of glass where different protein molecules are affixed precisely making a microscopic array. Most common protein chips are either glass slide chips or nano-well arrays. Protein profile arrays are used to monitor proteomes and protein quantities (Schweitzer et al., 2002). The functional protein arrays are used to examine protein-protein interactions, to identify the substrates of protein kinases, and to identify the targets of biologically active small molecules (Kuruvilla et al., 2002; Schweitzer et al., 2003). The most common protein microarray is the antibody microarrays. Recently, nucleic acids, receptors, enzymes, and proteins have been spotted onto chips and used as capture molecules. This will revolutionize the experiments on protein-protein interactions, and all other protein binding substrates (MacBeath and Schreiber, 2000; Jones et al., 2006b). Still, several technical hurdles have to be overcome in order to utilize protein chips effectively.

Major applications of Proteomics

1. One of the largest applications is protein profiling and studying the protein expression levels under different biological conditions.
2. Another important application is the characterization of post translational modifications such as phosphorylation, methylation, glycosylation, acetylation etc.
3. Proteomics can also be employed to study the composition of protein complexes, cellular localization of proteins, and to understand the interactions with other biomolecules.
4. Another application of proteomics is proteome mining. It is a method to identify the protein targets of drugs and identifying drugs that interact with specific proteins.

CHAPTER III

PROTEOME PROFILING OF *ARABIDOPSIS THALIANA* COTYLEDON

Abstract

Cotyledon proteome analysis of *Arabidopsis thaliana* was investigated using 2-Dimensional Gel Electrophoresis (2-DE) and Multidimensional Protein Identification Technology (MudPIT). Total proteins were isolated from 10 days old fully expanded cotyledons and separated on 2-DE gels. From SYPRO Ruby stained 2-D gels, among the 748 excised protein spots, 583 were annotated. Of these annotated spots, 500 proteins were identified with Confidence Intervals (C.I.%) over 95% and the other 83 annotations had a Cross Confidence Interval below 95%. These 500 identified proteins correlate to 353 distinct genes. From MudPIT results, using ESI MS/MS, 662 proteins have been identified with the parameters: Xcorr values = +1-1.9; +2-2.2; +3-3.75, Delta Corr = 0.1 and used pFactor - 0.001. By combining the data from both methods, the proteome map of cotyledon was constructed with 1023 proteins. The genes corresponding to all these proteins were characterized using the Gene Ontology tools. They were grouped into different levels and presented in pie charts. Gene Ontology studies reveal that most cotyledon proteins are involved in photosynthesis and energy metabolic pathways, consistent with the role of cotyledons in supporting the young seedlings with photosynthesis products and storage nutrients.

Introduction

Despite its little agricultural significance, *Arabidopsis thaliana* is a model organism for studying Genetic, Cellular and Molecular Biology of plants due to its small genome size and short life cycle. After the sequencing of *Arabidopsis* genome, scientists have focused on the plant functional genomics very extensively. Identification and functional categorization of proteins that found in specific tissues, their tissue specific expression, and differential time point expression during the development give an insight for understanding the functions and protein networks. To do this, proteomics has become an indispensable and versatile tool.

The field of proteomics can be divided into two major types. a. Cell map proteomics strives to define all the proteins within a particular organelle to discern the cellular architecture and their function (Fountoulakis et al., 2002). Functional proteomics strives to study the protein profile change in a cell or tissue or organism in response to a specific biological condition. These changes could be protein modifications, proteolysis, sub cellular localization, or interaction with other proteins (Graves and Haystead, 2002). To date, several tissue specific proteomes like Seed proteome (Gallardo et al., 2001), Vacuole Proteome (Carter et al., 2004), Pollen proteome (Noir et al., 2005), and organelle proteomes like Plasma Membrane Proteome (Alexandersson et al., 2004), cell wall proteome (Robertson et al., 1997), endoplasmic reticulum proteome (Chivasa et al., 2002), chloroplast envelope membrane proteome (Ferro et al., 2003), leaf peroxisome proteome (Fukao et al., 2002), chloroplast proteome (Kleffmann et al., 2004), mitochondria proteome (Kruft et al., 2001) etc., have been established. Here the unprecedented *Arabidopsis* cotyledon proteome map has been established.

Cotyledon is defined as a embryo leaf of a seed. Plant cotyledon is a specialized organ with terminally differentiated cell fate and its function is to support the growth of young seedlings with storage nutrients and photosynthetic products. Differentiated cotyledon cells can dedifferentiate and regenerate efficiently under appropriate stimuli. To induce callus from cotyledons, the cells have to undergo cell dedifferentiation. Therefore, cotyledons are ideal explants for the studies on cell dedifferentiation. In addition, the cotyledon is abundant and easy to collect.

Two analytical methods, Two-Dimensional Polyacrylamide Gel Electrophoresis (2D-PAGE) and Multidimensional Protein Identification Technology (MudPIT) are currently being used for cell map and functional proteomics. The traditional approach for proteomic analyses is separating protein mixtures on Two-Dimensional Polyacrylamide Gel Electrophoresis (2D-PAGE) followed by characterization of individual proteins/peptides by mass spectrometry (MS). Recent significant advances in proteomics made it possible to distinctly separate thousands of proteins with higher resolution and with higher reproducibility. Another revolutionary development is mass spectrometry techniques such as Matrix assisted laser desorption ionization Time-of- Flight mass spectrometry (MALDI-TOF-MS) (Karas and Hillenkamp, 1988), Electro spray ionization – tandem mass spectrometry (ESI-MS/MS) (Fenn et al.), and the development of multidimensional separation strategies in conjunction with automated ESI tandem mass spectrometry (2D LC MS/MS) for analyzing complex protein mixtures (Opiteck et al., 1997; Peng and Gygi, 2001). The latter technique is an alternative strategy to 2DE (Gerber et al., 2003). In this way, total protein mixture can be sequenced in a single run,

and this method allows detection of low abundance proteins such as transcription factors, hormone receptors, etc.

Materials and methods

Cotyledon collection

Arabidopsis thaliana variety Columbia seeds were sterilized using 30% Clorox and vernalized for three days at 4^{o}C. The seeds were plated on germination media (Murashige & Skoog basal salt media) and allowed to grow for 10 days at 22^{0}C in a light and air controlled incubator (Percival). Cotyledons were harvested after 10 days.

Protein extraction, sample preparation and concentration determination

The cotyledon tissue was ground in liquid nitrogen with mortar and pestle into very fine powder and total proteins were extracted using a phenol based protocol with modifications (Hurkman and Tanaka, 1986a) as follows: Grounded tissue was dissolved in an extraction buffer (0.9M sucrose, 0.5M tris-HCl, 0.05M EDTA, 0.1M KCl and added 2% β mercaptoethanol freshly, pH 8.7), and an equal volume of saturated phenol, pH – 8.0 was added and homogenized for 10 minutes. The homogenate was centrifuged at 2,500g for 10 minutes and phenol phase was collected. This process was repeated three more times to ensure not to collect any nucleic acid and starch content. Lastly, the phenol phase was collected into a fresh tube and added five volumes of precipitation solution (methanol, 0.1M ammonium acetate and 1% β-mercaptoethanol). Precipitation was carried out at -70^{o}C overnight. Precipitate was collected by centrifuging at 13,400g for 10

minutes and the pellet was washed with cold precipitation solution three times and another three times with cold 70% ethanol. The protein pellet was lyophilized to powder in a speed vacuum (LABCONCO, model LYPH-LOCK 6) and stored at -70^{o}C. Another two independent protein extractions were also carried out in the same manner for running replica gels.

For Two-Dimensional Polyacrylamide Gel Electrophoresis (2D PAGE), proteins were dissolved thoroughly in rehydration buffer (7M urea, 2M thio urea, 4% CHAPSO, 1% DTT, and 0.2% Ampholines) and centrifuged at 10,000 rpm for 10 minutes to remove any undissolved content and supernatant was quantified using Bio-Rad Rc Dc protein assay kit according to the protocol provided by the manufacturer.

Two-Dimensional Polyacrylamide Gel Electrophoresis (2D PAGE)

Isoelectric focusing (IEF) was carried out using Bio-Rad PROTEAN IEF cell on 24cm 3-10 pH non linear IPG strips (Bio-Rad Hercules, CA). Protein (1mg) in 400µl of rehydration buffer was loaded into IEF tray and active rehydration was carried out for 12h at 23^{o}C followed by 250V for 2h, and then voltage linear increase up to 10,000V for 4h. Focusing was performed at 10,000V to a total of 90kVh. The strips were equilibrated in a buffer containing 6M urea, 0.375M Tris-HCl - pH 6.8, 20% glycerol, 2% SDS, and 2% dithiothreitol for 15 minutes and followed by equilibration in another buffer containing 6M urea, 0.375M Tris-HCl - pH 6.8, 20% glycerol, 2% SDS, 0.1% bromophenol blue, and 2.5% iodoacetamide for 15 minutes.

The separated proteins according to their pH on IPG strips were separated again according to their molecular mass on horizontal slab gels (25 X 20.5 X 1.5mm). Slab gels

were cast in the multicasting chamber (Bio-Rad Hercules, CA) containing 12% separating gel and 4% stacking gel. Electrophoresis was carried out in Bio-Rad PROTEAN PLUS horizontal Dodeca cell at 20 mA/gel.

The gels were stained with SYPRO Ruby (Bio-Rad) according to the manual provided by the manufacturer and scanned with VersaDoc 4000 (Bio-Rad), and images were analyzed with PDQuest software (Bio-Rad, Hercules, CA).

In gel digestion and mass spectrometry

After PDQuest analysis**,** clearly separated spots were made as a "pick list" and were excised by robotic spot cutter (Proteome works; Bio-Rad) and deposited into 96-well plates. Gel plugs were processed using an Investigator ProPrep 4 block system (Genomic solutions Ann Arbor, MI). The excised spots were reduced with 10 mM DTT (Sigma) for 5 minutes and alkylated with 100 mM iodoacetamide (Sigma) for 30 minutes within the robotic digester. The spots were then digested with sequencing-grade trypsin (Promega, Madison, WI) for 16h at 37^0C. Peptides were desalted using C18 Zip tips and then mixed in 1:1 ratio with matrix (α-cyano-4-hydroxycinnamicacid dissolved in 70% acetonitrile and 0.1% trifluoroacetic acid) and spotted on MALDI plate.

MS spectra were collected in MALDI TOF/TOF (ABI 4700 proteomics Analyzer, Applied Biosystems, CA), and protein identification was performed using the Result Dependent Analysis (RDA) of ABI GPS explorer software, version 3.5. The parameters set for analysis were - MS peak filtering: 800-4000m/z interval, mono isotopic, minimum S/N = 10, mass tolerance = 150ppm. MS/MS peak filtering: 0-105% m/z of parent ion, monoisotopic, minimum S/N = 3, MS/MS fragment tolerance = 0.2Da. Proteins with

Cross confidence Intervel C.I. % > 95 were automatically selected for “in silico” digestion and their three strongest corresponding parent ions present in the MS spectra were selected for MS/MS analysis (First RDA). For proteins below 95% C.I. % or after 1st RDA, 20 strongest parent ions were selected for MS/MS analysis (second RDA). The initial MS scan data, 1st and 2nd RDA MS/MS data were together analyzed by using the MASCOT algorithm (Pappin et al., 1993). The proteins with total score C.I. % >95 were considered as positive IDs.

Confidence Interval (C.I. %) Formulas

The Protein Score and Ion Score values displayed by the GPS Explorer™ software and the Mascot® software are rounded. However, the GPS Explorer software uses the full precision values to calculate the Confidence Intervals (C.I. %). If the C.I. % is manually calculated using the rounded value from GPS or Mascot, the value may vary from the C.I. % displayed by the GPS Explorer software.

Protein Score C.I. % Formula

$$\text{C.I. \%} = 100 - \left(\frac{\text{sequences\_after\_tax} \times 100}{10^{(\text{ProteinScore}/10)}} \right)$$

where:
sequences_after_tax = from Mascot .DAT result file
ProteinScore = from Mascot

Ion Score C.I. % Formula (Also used for Total Ion Score C.I.% and Best Ion Score C.I.%)

$$\text{C.I. \%} = 100 - \left(\frac{\text{qmatch\_avg} \times \text{weighting factor}}{10^{(\text{IonScore}/10)}} \right)$$

where:

qmatch_avg = Sum of all qmatch values from Mascot .DAT result file / Number of queries (queries value from Mascot .DAT result file)

weighting factor = 5 for MS/MS analyses; 100 for combined analyses

IonScore = from Mascot

Sample preparation for MudPIT and MS analysis

The dried pellets of three protein samples that were extracted separately (replicas) were dissolved in 50µl of 6M urea with 100 mM Tris. The samples were centrifuged at 16,000g for 10 minutes, and any undissolved pellet was discarded. The supernatant was quantified using Bio-Rad Rc Dc kit according to the manual provided by the manufacturer. Protein (100µg) from each replica was reduced with 20µl reducing agent (200 mM DTT and 100 mM Tris, pH 7.8) for 1hr at room temperature and then alkylated in dark with 20µl of the alkylating reagent (200 mM IAA and 100 mM Tris) for 1hr. The urea concentration was then diluted down to 0.6M, a concentration at which the trypsin retains its activity. Trypsin solution (10µl) was added to the sample giving a protease to substrate ratio 1 to 50. The digestion was carried out overnight (15h) at 37^0C. The reaction was stopped by adding 10µl of lysine, adjusted to a pH less than 6 and vacuum dried for ~2.5h to make the final volume to 25µl. The peptides were desalted with a peptide macro trap (Michrom Bioresources, Inc., Auburn, CA) using a protocol provided

by the manufacturer and eluted with 0.1% trifluoroacetic acid and 95% acetonitrile. The eluted peptides were vacuum dried and the pellet re-dissolved in 0.1% formic acid and 5% acetonitrile.

Peptide samples were analyzed by strong cation exchange column (SCX BioBasic 0.32 × 100 mm) and followed by RP-LC (BioBasic C18, 0.18 × 100 mm Thermo Hypersil-Keystone, Bellefonte, PA) coupled directly in-line with electro spray ionization ion trap tandem MS (ProteomeX workstation Thermo Finnigan). A flow rate of 3µl/min was used for both SCX and reverse phase columns. For SCX, a salt gradient of 0, 10, 15, 25, 30, 35, 40, 45, 50, 57, 64, 90, and 700 mM ammonium acetate in 5% ACN, 0.1% formic acid was applied. For RP, the gradient of 0.1% formic acid in acetonitrile was used and the ACN concentration increased in a linear gradient from 5% to 30% in 30 minutes, 30% to 65% in 9 minutes and then followed by 95% for 5 minutes and 5% for 15 minutes.

The spectrum collection time was 59 minutes for each strong cation exchange step. The LCQ Deca XP ion trap mass spectrometer was configured to optimize the duty cycle length with the quality of data acquired by alternating between a single full MS scan followed by three tandem MS scans on the three most intense precursor masses from the full scan. Dynamic mass exclusion windows were set at 2 minutes. In addition, MS spectra for all samples were measured with an overall mass/charge (m/z) range of 200 to 2,000. The mass spectra and tandem mass spectra produced were used to search the non redundant protein database (nrPDB) downloaded from the National institute for biotechnology information by using the TURBOSEQUEST, BIOWORKS 3.1 SRI (Thermo Finnigan). TURBOSEQUEST cross correlates experimentally acquired mass

spectra with theoretical mass spectra that is generated in silico. The idealized spectra were weighed with a, b, and y fragment ions. Trypsin digestion was applied to generate the "precursor ions" and the database included mass changes due to cystine carbamidomethylation and methionine oxidation. The peptide (precursor) ion mass tolerance was 2.5Da, and the fragment ion (MS2) tolerance was 0.2. For a protein to be considered to be identified in the sample two or more peptides from this protein meet the following criteria: more than two peptides with X-correlation >1.7 (+1 charge), >2.2 (+2 charge), >3.75 (+3charge); delta correlation values of > 0.1 (Durr et al., 2004).

Functional categorization of proteins (Gene Ontology)

Functional categorization of proteins was carried out according to the GO rules (Berardini et al., 2004a) using Gene Ontology browser at http://www.arabidopsis.org/. Gene Ontology can be described as "a controlled vocabulary used to describe the Biology of a gene product in any organism". There are three independent sets of vocabularies or ontologies: the molecular function (MF) of a gene product, the biological process (BP) in which the gene product participates, and the cellular component (CC) where the gene product can be found. After analyses, these genes were grouped into different levels and pie charts were generated.

Results

Protein identification by 2-DE

The function of the cotyledon is to support the growth of young seedlings with storage nutrients and photosynthetic products, and the cells in cotyledons have a terminated cell fate. To induce callus, the cells have to undergo cell dedifferentiation. In addition, the cotyledon tissues are abundant and easy to collect. Therefore, cotyledons are ideal materials for the studies on cell dedifferentiation in plants.

The proteome of cotyledons has not been explored in plants. Construction of a proteome map of cotyledons in plants will significantly advance our understanding of the biological functions of cotyledons and set a foundation to examine the dynamic change of the cotyledon proteome during cell dedifferentiation. Therefore, the cotyledon proteome in *Arabidopsis* was mapped in this study. The cotyledon proteins of 10 days old seedlings were separated on 2-D gels, stained with SYPRO Ruby, visualized using a BIO-RAD VersaDoc4000 image system, and analyzed with BIO-RAD PDQuest 7.3.1 software. Over 748 protein spots had been consistently resolved on 2-D gels in three biological replicas. The prominent protein spots were excised, processed for in gel trypsin digestion, and analyzed with a MALDI-TOF/TOF mass spectrometer. Among the 748 excised protein spots, 583 were annotated. Of these annotated spots, 500 proteins were identified with Confidence Intervals (C.I.%) over 95% as shown in Figure 3.1 and listed in Table 3.1, the other 83 annotations had a Cross Confidence Interval below 95%. These 500 identified proteins correlate to 353 distinct genes. The rest 165 spots failed to obtain any protein annotation. Out of 165 unidentified spots, 36 spots were excised several times,

but always failed to get the identity. The remaining 129 spots were unable to be picked by spot cutter or unidentified by MALDI TOF/TOF.

Gene Ontology (GO) analyses of the corresponding genes revealed that 18.9% of the identified genes encoded chloroplast proteins and 8.1% encoded mitochondria proteins. In contrast, the genes for nucleus, cytosol, plasma membrane, ER, cell wall, and Golgi apparatus were 2.7%, 2.6%, 1.1%, 0.3%, 0.2%, and 0.2%, respectively (Figure 3.2A). These results were consistent with the primary role of cotyledons in photosynthesis and digestion of storage nutrients in mitochondria. Proteins in other organelles and cellular compartments were also shown in Figure 3.2A. Gene Ontology analyses based on biological processes indicated that 20.9% of the genes were involved in Metabolic Process and 21.4% genes were involved in Other Physiological Processes (Figure 3.2B). There were 7.9% genes in Protein Metabolic Pathways and only 0.6% genes in the DNA or RNA Metabolic Pathways (Figure 3.2B). In addition, the analyses based on Molecular Function showed that 21.6% of the proteins possessed Other Enzyme Activities, 10.1% were Molecular Function Unknown Proteins, 9.2% had Nucleotide Binding Activities, 5.7% possessed protein binding activity, and 3.2% were kinases. Proteins in other biological processes and molecular function categories were presented in detail in Figures 3.2B and 3.2C.

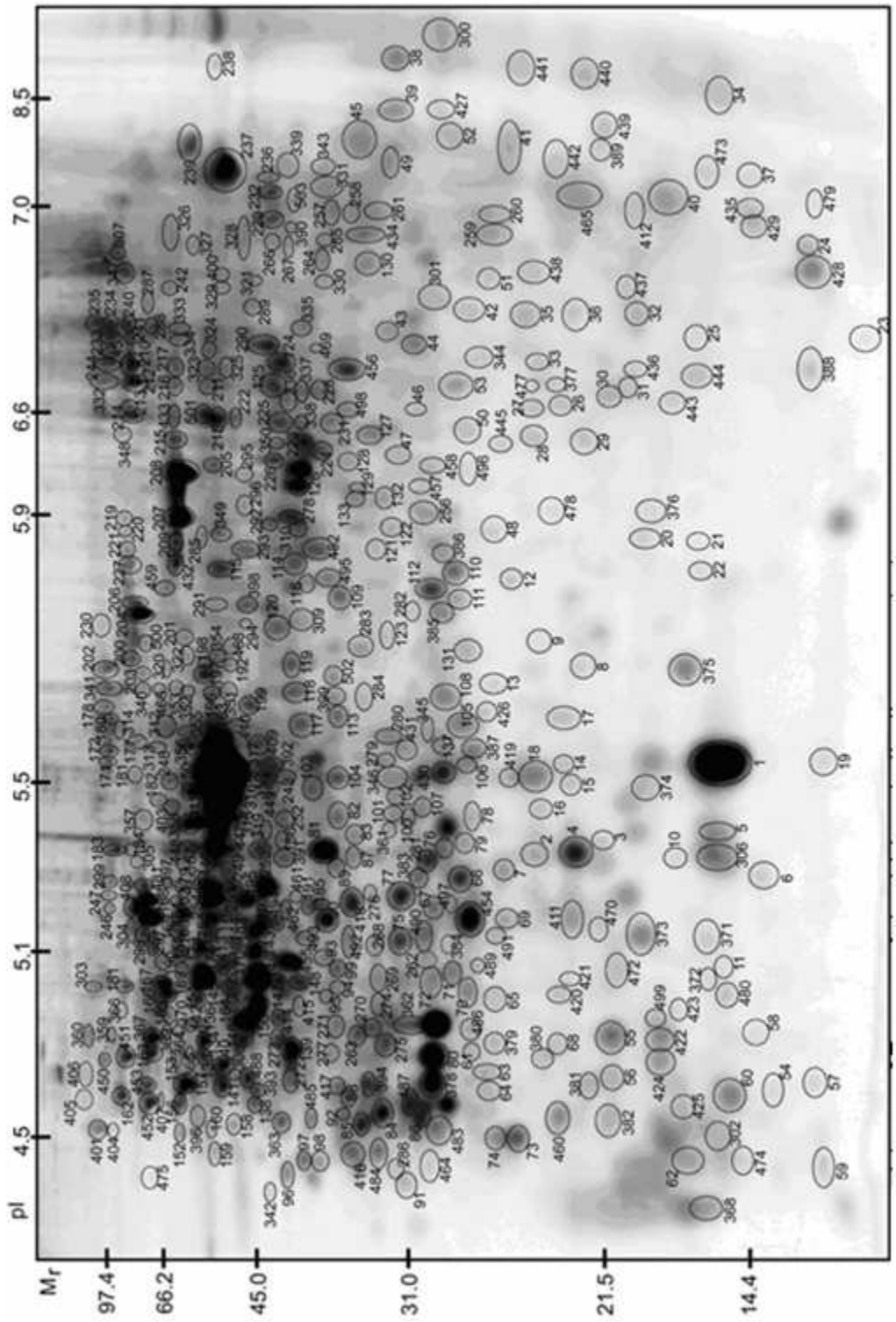

Figure 3.1: Master 2D map of *Arabidopsis thaliana* cotyledon proteome.

Table 3.1: Proteins Identified using MALDI TOF/TOF in *Arabidopsis* cotyledons

Spot #	Protein Identification	Exp. pI	Exp./ Theo. MW (kDa)	Pept--ides #	C.I.%*	Gene/locus
1	RuBisCO small subunit 1B, chloroplast precursor (EC 4.1.1.39)	5.5	15 / 48.5	15	100.00	AT5G38420
2	Ribulosebisphosphate carboxylase	5.4	23 / 28.3	17	100.00	ATCG00490
3	Peptidylprolyl isomerase ROC4 (EC 5.2.1.8) (Rotamase)	5.4	19.6 / 16.7	05	100.00	AT3G62030
4	Hypothetical protein / F3P11.21	5.3	19.6 / 46.4	15	100.00	AT2G19610
5	Nucleoside diphosphate kinase type 2	5.4	15.3 / 33.8	09	100.00	AT5G63310
6	Unknown protein (Hypothetical protein)	5.3	11.2 / 81.9	15	100.00	AT2G44650
7	Genomic DNA, chromosome 3, TAC clone:K13N2 (K13N2_9)	5.3	24.6 / 58.3	13	99.45	AT3G25770
8	Expressed protein	5.6	20.7 / 24.1	11	100.00	AT5G23720
9	Putative cytochrome P450 monooxygenase	5.7	22.7 / 57.6	16	100.00	AT1G13100
10	ATP synthase CF1 epsilon chain	5.3	16.7 / 42.9	10	100.00	AT1G05260
11	F15O4.36	5.1	14.3 / 148.1	20	99.63	AT1G35580
12	DNA mismatch repair protein, MutS family	5.7	24.2 / 36.9	04	100.00	AT1G65070
13	Similarity to senescence-associated protein	5.6	25.1 / 27.4	06	98.69	AT3G21600
14	(3R)-hydroxymyristoyl-[acyl carrier protein] dehydratase -like protein	5.5	21.4 / 47.2	09	98.19	AT5G10160
15	A_TM021B04.7 protein	5.5	21.1 / 65.5	12	100.00	AT5G27220
16	A_TM021B04.7 protein	5.4	22.7 / 54.3	16	100.00	AT5G27220
17	Hypothetical protein	5.6	21.5 / 23.5	07	96.70	AT3G62010
18	Germin-like protein	5.5	23 / 24	07	100.00	AT5G20630
19	Photosystem I iron-sulfur center (Photosystem I subunit VII) (PSI-C) (PsaC)	5.5	7.3 / 36.6	11	100.00	AT3G51680
20	Similarity to senescence-associated protein	5.8	18 / 15	05	98.59	AT3G21600
21	A_TM021B04.7 protein	5.8	15.9 / 43.5	19	100.00	AT5G27220
22	RuBisCO small subunit 1B, chloroplast precursor (EC 4.1.1.39)	5.7	15.8 / 53.3	13	100.00	AT5G38410
23	A_TM021B04.7 protein	7.2	5.6 / 5.3	12	100.00	AT5G27220
24	F5I14.5 protein	7.9	8.2 / 41.9	18	100.00	AT1G65520
25	Senescence-associated protein sen1	7.3	15.9 / 39.8	11	100.00	AT4G35770
26	Protein kinase family	6.7	21.5 / 35.8	12	100.00	AT1G03740
27	MAG2_6 (Hypothetical protein)	6.7	23 / 22	12	100.00	AT3G14110
28	ATP-dependent Clp protease proteolytic subunit (EC 3.4.21.92) (Endopeptidase Clp)	6.5	23 / 55.3	19	100.00	AT1G11750
29	Ubiquitin-conjugating enzyme E2 (EC 6.3.2.19)	6.5	20.5 / 45.1	14	100.00	AT5G50870
30	F5I14.5 protein	6.8	19.4 / 44.4	14	100.00	AT1G65520
31	Peptidyl-prolyl cis-trans isomerase (EC 5.2.1.8) (Cyclosporin A-binding protein)	6.8	18.6 / 39.8	10	100.00	AT4G38740
32	Non-symbiotic hemoglobin 1 (Hb1) (ARAth GLB1)	7.4	18.2 / 55.3	16	100.00	AT2G16060
33	F28I16_90 (Hypothetical protein)	7.0	22.7 / 68.3	12	100.00	AT5G19940
34	F13B4.9 protein	9.0	15.1 / 56.5	15	99.88	AT1G13610
35	50S ribosomal protein L21, chloroplast precursor (CL21)	7.4	23.4 / 51.6	11	100.00	AT1G35680
36	Hypothetical protein (Fragment)	7.4	21 / 38.5	09	100.00	AT3G21070
37	F15O4.36	8.4	12.2 / 30.4	08	99.62	AT1G35580
38	ADP, ATP carrier protein 1, mitochondrial precursor (ADP/ATP translocase 1) (ANT 1)	9.5	31.3 / 66	15	100.00	AT3G08580
39	ADP, ATP carrier protein 1, mitochondrial precursor (ADP/ATP translocase 1) (ANT 1)	8.9	31.4 / 43.4	08	100.00	AT3G08580
40	T19E23_1	8.3	17 / 49.6	07	100.00	AT1G31330
41	40S ribosomal protein S5-2	8.7	24.3 / 52.9	27	100.00	AT3G11940
42	Ribosome recycling factor, chloroplast precursor (RRFHCP)	7.4	26.5 / 34.6	03	96.77	AT3G63190

Table 3.1. Continued.

Spot #	Protein Identification	Exp. pI	Exp./ Theo. MW (kDa)	Pept--ides #	C.I.%*	Gene/locus
43	Hypothetical protein	7.3	31.9 / 27.4	09	100.00	AT4G33945
44	Putative L-ascorbate peroxidase, chloroplast precursor (EC 1.11.1.11)	7.2	30 / 37.7	09	100.00	AT4G09010
45	Putative oxoglutarate / malate translocator protein	8.7	34.6 / 39	09	99.99	AT5G19760
46	12S seed storage protein (CRA1)	6.7	29.6 / 36.9	18	100.00	AT3G51680
47	Hypothetical protein	6.4	31.2 / 43	15	99.99	AT1G47260
48	P-glycoprotein, multi-drug resistance related, ABC transporter-like protein	5.8	25.2 / 55.5	15	99.79	AT3G28345
49	Putative oxoglutarate / malate translocator protein	8.5	31.9 / 29.5	06	99.99	AT5G19760
50	Delta subunit of mitochondrial F1-ATPase	6.5	26.8 / 28.2	07	100.00	AT5G13450
51	Germin-like protein subfamily 3 member 1 precursor (AtGER1)	7.6	25.5 / 52.9	11	100.00	AT1G72610
52	Chloroplast 30S ribosomal protein S2	8.7	28.2 / 39.8	14	100.00	ATCG00160
53	Expressed protein	6.9	27.3 / 40	10	100.00	AT1G73850
54	Late embryogenesis abundant (LEA) domain-containing protein	4.6	10.3 / 33.7	11	100.00	AT5G44310
55	Putative peroxiredoxin protein	4.8	19.3 / 65.8	13	100.00	AT3G52960
56	Putative peroxiredoxin protein	4.7	19.3 / 47.8	13	100.00	AT3G52960
57	Cytochrome b-559 alpha subunit (Fragment)	4.7	7.8 / 34.4	16	100.00	ATCG00580
58	Ribulose bisphosphate carboxylase large chain precursor (EC 4.1.1.39)	4.8	11.7 / 58.5	17	100.00	ATCG00490
59	Autophagy 12a (Autophagy APG12)	4.2	7.3 / 72.4	21	100.00	AT1G54210
60	Profilin 2	4.6	13.9 / 40.9	13	100.00	AT4G29350
61	Hypothetical protein	4.8	26.4 / 53.5	10	100.00	AT3G05625
62	Genomic DNA, chromosome 5, P1 clone:MRO11	4.3	16.1 / 40.3	10	99.97	AT5G23820
63	2-cys peroxiredoxin BAS1, chloroplast precursor	4.7	25.5 / 18	07	100.00	AT3G11630
64	F3N23_3	4.6	25.4 / 42.8	16	100.00	AT1G72830
65	Expressed protein	4.9	25 / 31	11	100.00	AT5G38660
66	2,4-D-inducible glutathione S-transferase, putative	5.3	27.1 / 42.4	11	100.00	AT1G78370
67	Chlorophyll a/b-binding like protein	5.3	28.7 / 25	12	100.00	AT4G10340
68	ATP synthase D chain, mitochondrial (EC 3.6.3.14)	4.8	21.8 / 141.5	20	99.95	AT3G52300
69	Putative elongation factor P (EF-P)	5.2	24.4 / 29.3	04	99.99	AT3G08740
70	Hypothetical protein (Fragment)	5.0	26.5 / 49.5	15	100.00	AT3G21070
71	Ferritin 1, chloroplast precursor (AtFer1)	5.0	27.5 / 14.7	05	100.00	AT5G01600
72	Heat shock protein 70	5.0	28.8 / 56.1	14	100.00	AT5G49910
73	Translationally controlled tumor protein-like protein	4.4	23.9 / 45.7	16	100.00	AT3G16640
74	Translationally controlled tumor protein-like protein	4.4	25.1 / 52.9	17	100.00	AT3G16640
75	Carbonic anhydrase 2 (EC 4.2.1.1) (Carbonate dehydratase 2)	5.2	31 / 40.3	16	100.00	AT5G14740
76	Carbonic anhydrase 2 (EC 4.2.1.1) (Carbonate dehydratase 2)	5.3	30.9 / 49.4	10	100.00	AT5G14740
77	Expressed protein	5.4	28.1 / 37	08	100.00	AT1G73850
78	Expressed protein	5.4	26.4 / 41.2	09	100.00	AT3G62010
79	Multicatalytic endopeptidase complex, proteasome precursor, beta subunit (EC 3.4.99.46)	5.4	26.7 / 52	12	100.00	AT4G31300
80	Chlorophyll A-B binding protein 2, (LHCII type I CAB-2) (CAB-140) (LHCP)	4.7	28.8 / 43	11	100.00	AT1G29910
81	Fructose-bisphosphate aldolase, putative	5.4	37.6 / 35.5	13	100.00	AT4G38970
82	Putative ferredoxin-$NADP^+$ reductase	5.4	36.4 / 37.7	08	99.97	AT5G66190
83	Hypothetical protein	5.4	34.9 / 28.9	09	100.00	AT5G24490
84	14-3-3-like protein GF14 epsilon (General regulatory factor 10)	4.6	32.4 / 52.9	19	100.00	AT1G22300

Table 3.1. Continued.

Spot #	Protein Identification	Exp. pI	Exp./ Theo. MW (kDa)	Pept--ides #	C.I.%*	Gene/locus
85	F23J3_30 (14-3-3 protein GF14chi) (Grf1)	4.5	34.2 / 31.6	07	100.00	AT4G09000
86	Putative RNA-binding protein	4.6	29.4 / 42.5	08	100.00	AT2G37220
87	Expressed protein	5.3	35.1 / 27.4	11	100.00	AT4G11670
88	T32G9_30	4.6	35.6 / 46.2	09	100.00	AT1G35160
89	Putative ferredoxin-NADP+ reductase	5.3	36.6 / 35.7	07	99.97	AT5G66190
90	Putative fructose bisphosphate aldolase	5.2	37.4 / 49.1	12	100.00	AT2G21330
91	Nascent polypeptide associated complex alpha chain (T2E22_130)	4.2	30.7 / 17	06	99.80	AT3G12390
92	Disease resistance protein (TIR class), putative	4.5	36 / 35	10	100.00	AT4G19920
93	Pyruvate dehydrogenase E1 component beta subunit,mitochondrial precursor (EC 1.2.4.1)	5.1	37.7 / 24.4	14	100.00	AT5G50850
94	Hypothetical protein	5.0	36.5 / 54	14	100.00	AT5G48180
95	Diaminopimelate epimerase-like protein	5.0	37.7 / 18.5	03	99.80	AT3G53580
96	Putative glucanase	4.2	41.5 / 59.2	15	100.00	AT2G17390
97	AFT protein (Fragment)	4.3	39.8 / 37.3	11	100.00	AT4G35450
98	RNA-binding protein cp33 (F22O6_240)	4.3	38.3 / 63.3	10	100.00	AT3G52380
99	Pyruvate dehydrogenase E1 component beta subunit, mitochondrial precursor (EC 1.2.4.1)	5.0	36.4 / 80.2	14	100.00	AT5G50850
100	Multicatalytic endopeptidase complex, proteasome component, alpha subunit	5.4	30.4 / 29.9	19	100.00	AT5G35590
101	Multicatalytic endopeptidase complex, proteasome component, alpha subunit	5.4	31.5 / 28.1	14	100.00	AT5G35590
102	F7G19.1 protein	5.5	31.4 / 50.1	13	100.00	AT1G09130
103	AT10D17_100	5.5	38.6 / 80.2	12	100.00	AT3G44300
104	Cysteine synthase (EC 2.5.1.47) (O-acetylserine)	5.5	36.4 / 72.4	24	100.00	AT4G14880
105	Glutathione S-transferase 6 (EC 2.5.1.18) (GST class phi)	5.6	26.9 / 39.2	08	100.00	AT2G47730
106	Hypothetical protein	5.5	26.6 / 36.5	08	100.00	AT4G25360
107	L-ascorbate peroxidase, cytosolic (EC 1.11.1.11)	5.4	29.6 / 40.6	07	100.00	AT1G07890
108	Glutathione transferase, putative	5.6	28 / 40.4	11	100.00	AT4G02520
109	T20H2.20 protein	5.7	36.2 / 149.9	27	99.65	AT1G20020
110	Glutathione S-transferase (At2g30860/F7F1.7)	5.7	27.4 / 39.5	08	100.00	AT2G30860
111	Proteasome subunit beta type 1 (EC 3.4.25.1)	5.7	27.1 / 73.6	15	100.00	AT3G60820
112	Putative vegetative storage protein	5.7	28.8 / 44	06	99.95	AT5G44020
113	Hypothetical protein	5.5	36.8 / 15.1	05	100.00	AT1G20020
114	Putative glyceraldehyde-3-phosphate dehydrogenase (F13B4_8)	5.8	40.6 / 10.6	07	99.92	AT1G13440
115	F26F24_4	5.7	51.5 / 22	07	100.00	AT1G70580
116	ATP synthase gamma chain 1, chloroplast precursor (EC 3.6.3.14)	5.7	39.2 / 44.2	13	100.00	AT4G04640
117	Malate dehydrogenase, cytoplasmic 1 (EC 1.1.1.37)	5.6	40.1 / 42.1	13	100.00	AT1G04410
118	Fructose bisphosphate aldolase-like protein (F8J2_100)	5.6	40.7 / 44.4	07	99.84	AT3G52930
119	mRNA binding protein-like (T20O10_240)	5.6	41 / 135.7	11	97.61	AT3G63140
120	Hypothetical protein	5.7	42.6 / 37.7	11	100.00	AT4G19880
121	Apocytochrome f precursor	5.8	32.9 / 34.3	14	100.00	ATCG00540
122	Ascorbate peroxidase (EC 1.11.1.11) (L-ascorbate peroxidase)	5.8	31.8 / 52.9	26	100.00	AT4G35000
123	Hypothetical protein	5.7	32 / 51.6	07	100.00	AT2G25150
124	Hypothetical protein	7.0	42 / 63.4	15	100.00	AT4G25360
125	Aminomethyltransferase, mitochondrial precursor (EC 2.1.2.10)	6.9	43.2 / 81.9	15	100.00	AT1G11860

Table 3.1. Continued.

Spot #	Protein Identification	Exp. pI	Exp./ Theo. MW (kDa)	Pept--ides #	C.I.%*	Gene/locus
126	Putative glyceraldehyde-3-phosphate dehydrogenase (F13B4_8)	6.1	40.3 / 59.1	09	99.91	AT1G13440
127	Expressed protein (T31E10.20) (Hypothetical protein)	6.4	35.7 / 59.8	27	100.00	AT2G34460
128	Cinnamyl-alcohol dehydrogenase (CAD), putative	6.3	35.7 / 38.8	13	100.00	AT5G19440
129	30S ribosomal protein S5	6.1	35 / 26.1	15	100.00	AT2G33800
130	F9D12.7 protein	7.8	33.6 / 63.7	16	100.00	AT5G26280
131	Uracil phosphoribosyltransferase-related protein	5.6	28.7 / 44.4	13	100.00	AT3G53900
132	12S seed storage protein (cruciferin), putative	6.0	32.4 / 31.6	13	100.00	AT4G28520
133	12S seed storage protein (cruciferin), putative	6.0	35.1 / 35.3	12	100.00	AT4G28520
134	Putative phosphorethanolamine N-methyltransferase 2 (EC 2.1.1.103)	4.8	55 / 38.5	07	100.00	AT1G48600
135	Tubulin beta-2 / beta-3 chain	4.8	55.2 / 37.7	13	100.00	AT5G62690
136	Hypothetical protein	4.8	55 / 24.1	09	100.00	AT4G17610
137	Glutathione transferase, putative	5.5	28.2 / 24.1	08	100.00	AT4G02520
138	Fructose-1,6-bisphosphatase, chloroplast precursor (EC 3.1.3.11)	4.6	45.4 / 70	13	100.00	AT3G54050
139	Sedoheptulose-1,7-bisphosphatase, chloroplast precursor (EC 3.1.3.37)	4.8	40.9 / 84.3	25	100.00	AT3G55800
140	T5I7.3	4.7	50.8 / 20.3	07	100.00	AT2G39730
141	COP8 (Constitutive photomorphogenic) homolog (CSN complex subunit 4)	4.7	51.1 / 24.2	07	99.96	AT5G42970
142	Putative RNA-binding protein LAH1	4.9	58.1 / 28.1	10	100.00	AT4G32720
143	T5I7.3 (Hypothetical protein)	5.1	45 / 67.8	26	100.00	AT2G39730
144	Actin 8	5.2	48.5 / 41.4	12	100.00	AT1G49240
145	Glutamate--cysteine ligase, chloroplast precursor (EC 6.3.2.2)	5.0	54.8 / 63.4	16	100.00	AT4G23100
146	Phosphoglycerate kinase (EC 2.7.2.3) (MBK21_14)	4.8	47.3 / 98.1	17	99.87	AT3G12780
147	Phosphoribulokinase, chloroplast precursor (EC 2.7.1.19) (Phosphopentokinase)	5.0	42.3 / 71.1	18	100.00	AT1G32060
148	Putative pyruvate dehydrogenase E1 beta subunit	5.0	40.2 / 67	12	100.00	AT2G34590
149	Phosphoribulokinase, chloroplast precursor (EC 2.7.1.19) (Phosphopentokinase)	5.0	44.5 / 85.9	22	100.00	AT1G32060
150	Phosphoribulokinase, chloroplast precursor (EC 2.7.1.19) (Phosphopentokinase)	4.9	44.8 / 72.7	13	100.00	AT1G32060
151	ATP synthase alpha chain (EC 3.6.3.14)	4.9	60.4 / 42.8	16	100.00	ATCG00120
152	Hypothetical protein	4.5	60.4 / 34.3	13	100.00	AT5G39570
153	Hypothetical protein	4.7	61.3 / 22.5	05	100.00	AT1G20260
154	ATP synthase beta chain (EC 3.6.3.14)	5.2	59.5 / 84.3	12	100.00	ATCG00480
155	ATP synthase alpha chain (EC 3.6.3.14)	5.1	59.8 / 51.6	11	100.00	ATCG00120
156	F2N1_31	4.6	60.5 / 59.7	17	100.00	AT4G01050
157	Trigger factor-like protein	4.7	59.2 / 23.8	04	99.96	AT5G55220
158	Fructose-1,6-bisphosphatase, chloroplast precursor (EC 3.1.3.11)	4.5	49 / 72.7	10	100.00	AT3G54050
159	DNA repair protein RAD23 homolog (Hypothetical protein)	4.3	52.8 / 42.4	08	99.98	AT5G38470
160	Tubulin beta-5 chain	4.5	53.4 / 36.9	19	100.00	AT1G20010
161	Putative elongation factor G protein	5.0	90.3 / 14.7	05	100.00	AT1G62750
162	Heat shock protein 70	4.6	83.8 / 21.5	02	95.00	AT5G49910
163	Dihydroxyacetone kinase -related	4.8	71.1 / 38.2	10	100.00	AT3G17770
164	T5I7.3	4.9	49.9 / 29.1	07	100.00	AT2G39730
165	ATP synthase alpha chain (EC 3.6.3.14)	4.9	63.9 / 46.7	17	100.00	ATCG00120
166	ATP synthase alpha chain (EC 3.6.3.14)	4.9	66.1 / 27.5	07	100.00	ATCG00120
167	ATP synthase alpha chain (EC 3.6.3.14)	5.0	65.7 / 39.8	14	100.00	ATCG00120

Table 3.1. Continued.

Spot #	Protein Identification	Exp. pI	Exp./ Theo. MW (kDa)	Pept--ides #	C.I.%*	Gene/locus
168	Transketolase-like protein	5.3	72.8 / 20.8	10	100.00	AT3G60750
169	GloEL protein, chaperonin, 60 kDa	5.2	71.5 / 56.8	11	99.79	AT3G13470
170	Actin 7	5.1	49.2 / 42.8	11	100.00	AT5G09810
171	Ribulose-bisphosphate carboxylase	5.5	89.7 / 27.1	11	100.00	ATCG00490
172	Ribulose-bisphosphate carboxylase	5.5	91 / 18.9	09	100.00	ATCG00490
173	Jasmonate inducible protein ISOLOG (Putative lectin protein)	5.1	55.5 / 35.3	14	100.00	AT3G16400
174	Glucose-1-phosphate adenylyltransferase small subunit, (EC 2.7.7.27)	5.1	52.7 / 41.2	11	100.00	AT5G48300
175	Eukaryotic initiation factor 4A-2 (eIF4A-2) (eIF-4A-2)	5.2	53.4 / 40.3	12	100.00	AT1G54270
176	Eukaryotic initiation factor 4A-1 (eIF4A-1) (eIF-4A-1)	5.2	52.8 / 40.4	19	100.00	AT3G13920
177	Phosphoglucomutase	5.5	76.9 / 39.7	06	100.00	AT1G23190
178	5-methyltetrahydropteroyltriglutamate--homocysteine Methyltransferase (EC 2.1.1.14)	5.6	93.5 / 41.7	13	100.00	AT5G17920
179	H+-transporting ATP synthase beta chain (Mitochondrial) -like protein	5.3	59.9 / 32.2	11	99.99	AT5G08680
180	Glycyl-tRNA synthetase (EC 6.1.1.14) (Glycine--tRNA ligase) (GlyRS)	5.6	88.3 / 77.1	20	100.00	AT1G29880
181	Ribulose bisphosphate carboxylase large chain precursor (EC 4.1.1.39)	5.5	78.2 / 35.6	10	100.00	ATCG00490
182	Ribulose bisphosphate carboxylase large chain precursor (EC 4.1.1.39)	5.5	76.4 / 32.6	09	100.00	ATCG00490
183	GcpE	5.4	87.2 / 50.9	10	100.00	AT5G60600
184	Eukaryotic translation initiation factor 3 subunit 7	5.4	80.5 / 89.1	20	99.97	AT5G44320
185	Putative lectin	5.3	38.2 / 52.9	29	100.00	AT1G45214
186	Proline iminopeptidase	5.4	42.1 / 20.6	09	100.00	AT2G14260
187	ATP synthase beta chain (EC 3.6.3.14)	5.3	59.7 / 84.3	05	100.00	ATCG00480
188	ATP synthase alpha chain (EC 3.6.3.14)	5.2	62.3 / 38.2	11	100.00	ATCG00120
189	ATP synthase alpha chain (EC 3.6.3.14)	5.0	59.4 / 33.8	09	100.00	ATCG00120
190	GloEL protein, chaperonin, 60 kDa	5.1	67.7 / 53	11	99.77	AT3G13470
191	Fructose-bisphosphate aldolase, putative	5.3	39.3 / 38.5	23	100.00	AT4G38970
192	F24J8.12 protein	5.6	49.8 / 33.3	05	99.64	AT1G21510
193	T5I7.3 (Hypothetical protein)	5.2	44.6 / 77.2	18	100.00	AT2G39730
194	Aspartate-semialdehyde dehydrogenase (EC 1.2.1.11) (Fragment)	5.2	43.6 / 37.5	08	99.95	AT1G14810
195	Auxin-regulated protein	5.3	46.5 / 47.7	12	100.00	AT2G39730
196	Tubulin alpha-3/alpha-5 chain	5.6	53.5 / 57.4	14	100.00	AT5G19770
197	F26F24_4	5.6	54.6 / 52.9	25	100.00	AT1G70580
198	Hypothetical protein	5.6	52.4 / 17.5	10	100.00	AT4G13430
199	uroporphyrinogen decarboxylase (UPD), putative	5.6	46 / 53.3	19	100.00	AT3G14930
200	T3F20.18 protein (DTDP-D-glucose 4,6-dehydratase, putative 102946-105028)	5.6	80 / 72.7	10	99.64	AT1G53500
201	ATP synthase alpha chain, mitochondrial (EC 3.6.3.14)	5.7	59 / 42.9	12	100.00	ATMG01190
202	5-methyltetrahydropteroyltriglutamate--homocysteine Methyltransferase (EC 2.1.1.14)	5.6	90.6 / 47.6	14	100.00	AT5G17920
203	Gb|AAF00675.1	5.6	75.3 / 44.3	13	99.23	AT3G20350
204	MPI7_60	5.7	77.8 / 76.9	21	100.00	AT5G17920
205	Putative fumarase (T30B22.19)	6.3	53.3 / 68	13	100.00	AT2G47510
206	5-methyltetrahydropteroyltriglutamate--homocysteine Methyltransferase (EC 2.1.1.14)	5.7	79.3 / 51.6	16	100.00	AT5G17920
207	F18O14.27	5.9	60 / 183.4	23	99.66	AT1G19520
208	F20D10_50 (EC 2.1.2.1) (Serinehydroxylmethyltransferase)	6.2	59.8 / 35.7	21	100.00	AT4G37930
209	Lipoamide dehydrogenase	5.8	62.1 / 183.4	28	95.12	AT1G48030

Table 3.1. Continued.

Spot #	Protein Identification	Exp. pI	Exp./ Theo. MW (kDa)	Pept--ides #	C.I.%*	Gene/locus
210	Heat shock cognate 70 kDa protein 3 (Hsc70.3)	7.0	79 / 52.9	15	100.00	AT3G09440
211	Putative elongation factor G protein	6.9	54.3 / 31.9	10	100.00	AT1G62750
212	Putative poly-A binding protein	6.9	78.8 / 38.3	08	99.90	AT1G49760
213	Heat shock cognate 70-1	6.7	80.8 / 40.3	17	100.00	AT4G24280
214	Putative poly-A binding protein	6.6	79 / 135.7	10	99.90	AT1G49760
215	F20D10_50 (EC 2.1.2.1) (Serinehydroxylmethyltransferase)	6.5	62.7 / 28.4	12	100.00	AT4G37930
216	Catalase 3 (EC 1.11.1.6)	6.9	62.4 / 38.4	10	100.00	AT1G20620
217	Unknown protein	7.0	61.9 / 17.9	05	100.00	AT1G10070
218	Expressed protein	6.6	51.8 / 33.3	15	100.00	AT1G64180
219	Putative DEAD/DEAH box RNA helicase	5.9	82.3 / 47.6	16	100.00	AT2G42520
220	Putative poly(A)-binding protein (Putative polyA-binding protein)	5.8	82 / 49.4	11	100.00	AT2G23350
221	ATP-dependent RNA helicase-like protein	5.8	81.2 / 25.6	07	96.68	AT3G58510
222	Putative GTP-binding protein	6.6	48.7 / 28.2	07	100.00	AT1G49300
223	Dbj|BAA87936.1 (MQC12_13) (Hypothetical protein)	6.8	38.6 / 27.2	10	99.30	AT3G20370
224	Glyceraldehyde 3-phosphate dehydrogenase A subunit	6.4	38.4 / 29.1	13	100.00	AT3G26650
225	Glyceraldehyde 3-phosphate dehydrogenase A, chloroplast precursor (EC 1.2.1.13)	6.6	42.5 / 57.4	20	100.00	AT3G26650
226	Putative glyceraldehyde-3-phosphate dehydrogenase (F13B4_8)	6.3	43 / 55.6	12	99.91	AT1G13440
227	ATP-dependent RNA helicase-like protein	5.8	80.9 / 27.1	09	100.00	AT2G34590
228	Glyceraldehyde 3-phosphate dehydrogenase A subunit	6.4	39.6 / 22.9	08	100.00	AT3G26650
229	Polygalacturonase inhibitor-like protein (Hypothetical protein)	8.1	43 / 59.3	12	99.40	AT3G20820
230	5-methyltetrahydropteroyltriglutamate--homocysteine Methyltransferase (EC 2.1.1.14)	5.7	94.3 / 56.6	14	100.00	AT5G17920
231	Ferredoxin--NADP+ reductase-like protein	6.6	37.1 / 36.6	07	97.32	AT4G05390
232	MOA2_2	8.3	43.4 / 14	04	100.00	AT3G14420
233	5-methyltetrahydropteroyltriglutamate--homocysteine Methyltransferase (EC 2.1.1.14)	7.2	89 / 46.7	14	100.00	AT5G17920
234	5-methyltetrahydropteroyltriglutamate--homocysteine Methyltransferase (EC 2.1.1.14)	7.3	89.3 / 46.7	14	100.00	AT5G17920
235	T6H22_13 (Elongation factor EF-2)	7.3	98.9 / 13.7	03	100.00	AT1G56070
236	MOA2_2	8.4	43.7 / 26.8	11	100.00	AT3G14420
237	Probable (S)-2-hydroxy-acid oxidase, peroxisomal 1 (EC 1.1.3.15)	8.5	50.4 / 60.3	12	99.49	AT3G14415
238	T6D22_3	8.7	57.5 / 25.6	09	100.00	AT1G07930
239	Putative translation elongation factor eEF-1 alpha chain (Gene A4)	8.6	58.9 / 44.8	13	100.00	AT5G60390
240	Luminal binding protein 1 precursor (BiP1) (AtBP1)	7.3	81.2 / 71.9	12	99.73	AT5G13820
241	Heat shock protein, putative	7.2	79.6 / 51.7	11	99.99	AT1G56410
242	K17E12_12	7.6	64.5 / 20.1	05	100.00	AT3G27300
243	Fructose bisphosphate aldolase-like protein (F8J2_100)	5.5	41.4 / 23.4	13	99.84	AT3G52930
244	5-methyltetrahydropteroyltriglutamate--homocysteine Methyltransferase (EC 2.1.1.14)	7.0	92.1 / 42.1	14	100.00	AT5G17920
245	5-methyltetrahydropteroyltriglutamate--homocysteine Methyltransferase (EC 2.1.1.14)	7.1	91.4 / 63.1	09	100.00	AT5G17920
246	Jasmonate inducible protein ISOLOG (Putative lectin protein)	5.3	90 / 44.4	11	100.00	AT3G16460
247	Oligopeptidase A	5.3	92.6 / 57.5	10	99.41	AT5G65620
248	Ribulose bisphosphate carboxylase large chain precursor (EC 4.1.1.39)	5.5	63.1 / 38.2	11	100.00	ATCG00490
249	Elongation factor Tu, chloroplast precursor (EF-Tu)	5.3	50 / 55	15	100.00	AT4G20360

Table 3.1. Continued.

Spot #	Protein Identification	Exp. pI	Exp./ Theo. MW (kDa)	Pept--ides #	C.I.%*	Gene/locus
250	Ribulose bisphosphate carboxylase large chain precursor (EC 4.1.1.39)	5.5	57.9 / 25.1	05	100.00	ATCG00490
251	S-adenosylmethionine synthase (EC 2.5.1.6)	5.3	48.9 / 138.5	19	99.11	AT3G17390
252	Fructose-bisphosphate aldolase - like protein	5.4	41.7 / 52.9	17	100.00	AT4G26530
253	S-adenosylmethionine synthase (EC 2.5.1.6)	5.4	49.6 / 23.2	09	98.88	AT3G17390
254	Putative chloroplast translation elongation factor EF-Tu precursor	5.3	52.1 / 50.3	13	100.00	AT4G20360
255	Eukaryotic initiation factor 4A-2 (eIF4A-2) (eIF-4A-2)	5.3	54.1 / 84.3	20	100.00	AT1G54270
256	Multicatalytic endopeptidase	5.9	29.1 / 28.9	12	100.00	AT3G22110
257	Lipase/hydrolase, putative (F15D2_22)	8.2	37.1 / 23.2	06	99.99	AT1G29670
258	F9D12.7 protein	8.1	34.9 / 55.3	13	100.00	AT5G26280
259	Putative 60S ribosomal protein	8.0	24.5 / 36.2	16	100.00	AT3G24830
260	Germin-like protein subfamily 3 member 1 precursor (AtGER1)	8.1	24.4 / 47.6	11	100.00	AT1G72610
261	cp29	8.1	33.2 / 31.1	10	100.00	AT1G60000
262	Hypothetical protein	5.0	30.1 / 25	06	100.00	AT1G36070
263	Glutamyl-tRNA reductase 2 (GluTR) (HEMA2)	4.8	34.5 / 36.5	10	100.00	AT1G09940
264	MKM21_30 (60S ribosomal protein L5)	7.8	38.1 / 39.2	13	100.00	AT3G25520
265	MKM21_30 (60S ribosomal protein L5)	8.0	37.4 / 44	20	100.00	AT3G25520
266	Protein Polygalacturonase inhibitor-like protein	7.9	43 / 40	03	99.40	AT3G20820
267	Protochlorophyllide reductase B, chloroplast precursor (EC 1.3.1.33) (PCR B)	7.9	41.2 / 54	22	100.00	AT4G27440
268	Carbonic anhydrase 2 (EC 4.2.1.1) (Carbonate dehydratase 2)	5.2	33.1 / 49.5	13	100.00	AT5G14740
269	Carbonic anhydrase 2 (EC 4.2.1.1) (Carbonate)	5.0	32.9 / 89.1	18	100.00	AT5G14740
270	33 kDa polypeptide of oxygen-evolving complex (OEC) in photosystem II	4.8	33.4 / 84.4	14	100.00	AT5G66570
271	Expressed protein (Hypothetical protein)	4.9	36.5 / 32.1	08	100.00	AT2G44040
272	Fructokinase, putative	4.7	41.1 / 42.8	12	99.99	AT1G66430
273	Sedoheptulose-bisphosphatase	4.8	42.6 / 44.4	16	100.00	AT3G55800
274	Putative poly-A binding protein	4.9	32.8 / 87.3	18	99.89	AT1G49760
275	At2g45990 protein (Hypothetical protein) (Expressed protein)	4.8	32.2 / 63.3	23	100.00	AT2G45990
276	Protein kinase, putative :CRA|225000014647594_3702	5.3	33.2 / 36.9	09	100.00	AT1G48490
277	Expressed protein (Hypothetical protein)	4.7	37.2 / 66.2	18	100.00	AT2G44040
278	ATP synthase gamma chain 1, chloroplast precursor (EC 3.6.3.14)	5.8	40.9 / 36.9	13	100.00	AT4G04640
279	Putative esterase D (T32G6.5) (S-formylglutathione hydrolase) (EC 3.1.2.12)	5.5	32.4 / 65.1	15	100.00	AT2G41530
280	V-type proton-ATPase	5.5	32 / 28.7	02	100.00	AT2G18960
281	L-ascorbate peroxidase, cytosolic (EC 1.11.1.11)	5.4	30 / 60.6	12	100.00	AT1G07890
282	T1K7.5 protein	5.7	30.4 / 55.8	13	99.88	AT1G26580
283	Kinesin-related protein katA (fragment)	5.7	34.3 / 28.2	07	100.00	AT4G21370
284	Indole-3-glycerol phosphate synthase, chloroplast precursor (EC 4.1.1.48)	5.6	34.1 / 53.9	20	100.00	AT2G04400
285	Ribulose bisphosphate carboxylase large chain precursor (EC 4.1.1.39)	5.8	35.5 / 43.7	09	100.00	ATCG00490
286	Putative lyase	4.3	31.2 / 151.7	20	100.00	AT5G10920
287	Polyadenylate-binding protein 2 (Poly(A)-binding protein 2) (PABP 2)	7.5	72.7 / 93.8	25	100.00	AT4G34110
288	ATP citrate lyase, putative (ATP citrate lyase, putative 38389-41775)	7.3	71.4 / 37.5	08	99.99	AT3G06650
289	GTP-binding protein ATGB2	7.4	45.6 / 79.9	15	100.00	AT4G35860
290	Alanine:glyoxylate aminotransferase (EC 2.6.1.44)	7.3	46.3 / 53.9	16	100.00	AT2G13360

Table 3.1. Continued.

Spot #	Protein Identification	Exp. pI	Exp./ Theo. MW (kDa)	Pept--ides #	C.I.%*	Gene/locus
291	Expressed protein	5.7	53.2 / 40.9	14	100.00	AT4G36105
292	Putative pyruvate dehydrogenase e1 alpha subunit	5.8	43 / 26.9	12	100.00	AT1G59900
293	Glyceraldehyde 3-phosphate dehydrogenase, cytosolic (EC 1.2.1.12)	5.8	42.2 / 56.7	14	100.00	AT3G04120
294	At4g05190	5.7	49 / 19.4	12	100.00	AT4G05190
295	Glutathione-dependent formaldehyde dehydrogenase	6.2	48.3 / 19.2	04	100.00	AT5G43940
296	Aspartate aminotransferase	6.0	47.1 / 34.8	11	100.00	AT4G31990
297	F21M12_17	5.1	71.1 / 8.9	03	100.00	AT1G09780
298	Putative 2,3-bisphosphoglycerate-independent phosphoglycerate mutase	5.2	70.5 / 26.5	12	100.00	AT1G09780
299	Oligopeptidase A	5.3	89.1 / 28.3	07	99.43	AT5G65620
300	60S ribosomal protein L10a-2	9.6	28.2 / 41.4	12	100.00	AT2G27530
301	Hypothetical protein	7.5	28.8 / 54.9	12	100.00	AT2G32870
302	Thylakoid membrane phosphoprotein 14 kDa, chloroplast precursor	4.5	14.8 / 55.3	18	100.00	AT2G46820
303	Cell division control protein 48 homolog E (AtCDC48e)	5.0	97.2 / 33.3	06	97.38	AT5G04590
304	Transketolase-like protein	5.2	79.4 / 11.6	05	100.00	AT3G60750
305	Soluble starch synthase	5.3	75.5 / 26.2	10	100.00	AT5G24300
306	F1O19.10/F1O19.10	5.3	14 / 24.7	10	100.00	AT1G67040
307	Luminal binding protein 1 precursor (BiP1) (AtBP1)	7.9	86.3 / 50.3	09	99.73	AT5G13820
308	Hypothetical protein (Putative ribosomal protein S1)	4.7	49.2 / 28.3	13	100.00	AT5G30510
309	F14J9.13 protein	5.7	40.2 / 66.2	16	100.00	AT1G09470
310	Glyceraldehyde 3-phosphate dehydrogenase A subunit	5.9	41 / 31.9	12	100.00	AT3G26650
311	Ribulose bisphosphate carboxylase large chain precursor (EC 4.1.1.39)	5.5	52.2 / 54	14	100.00	ATCG00490
312	Putative phosphoribosylaminoimidazolecarboxamide formyltransferase	5.6	65.2 / 33.4	13	100.00	AT2G35040
313	T-complex polypeptide 1 homologue	5.5	67.3 / 37.5	10	100.00	AT3G20050
314	Glycyl-tRNA synthetase (EC 6.1.1.14) (Glycine--tRNA ligase) (GlyRS)	5.5	84.1 / 71.3	20	100.00	AT1G29880
315	Probable elongation factor 1-gamma 2 (EF-1-gamma) (eEF-1B gamma)	5.4	51.6 / 43.7	05	99.95	AT1G57720
316	Ribulose-bisphosphate carboxylase	5.5	48.5 / 25.6	14	100.00	ATCG00490
317	Glyceraldehyde 3-phosphate dehydrogenase B, chloroplast precursor (EC 1.2.1.13)	5.5	45.3 / 47.8	11	100.00	AT3G26650
318	Glutamate-1-semialdehyde 2,1-aminomutase 1, chloroplast precursor (EC 5.4.3.8)	5.5	45.1 / 84.3	20	100.00	AT5G63570
319	glyceraldehyde-3-phosphate dehydrogenase (NADP) (EC 1.2.1.13) B precursor	5.4	45.8 / 32.1	11	100.00	AT1G42970
320	4-coumarate-CoA ligase-like protein	5.6	65.7 / 52.9	22	100.00	AT3G48990
321	Probable (S)-2-hydroxy-acid oxidase, peroxisomal 1 (EC 1.1.3.15)	7.7	45.8 / 35.6	10	99.50	AT3G14415
322	ATPase subunit 1	5.6	59 / 22	07	100.00	ATMG01190
323	6-phosphogluconate dehydrogenase, putative, 13029-14489 (EC 1.1.1.44)	7.0	53.9 / 14.5	06	99.93	AT3G02360
324	Tubulin beta-5 chain	7.1	53.7 / 42.5	15	100.00	AT1G20010
325	Elongation factor Tu, chloroplast precursor (EF-Tu)	7.0	50.3 / 55	12	100.00	AT4G20360
326	Unknown protein	8.0	63.9 / 8.6	02	100.00	AT4G01690
327	F20M17.15	7.9	56.6 / 39.4	19	100.00	AT2G31810
328	Putative geranyl reductase	7.9	50.3 / 37.7	11	100.00	AT1G74470
329	F22O6_120	7.6	51.4 / 63.9	14	100.00	AT3G52500
330	Putative dTDP-glucose 4-6-dehydratase	7.6	37.9 / 64	15	100.00	AT2G47650
331	Ribosomal protein L1	8.3	37.7 / 45.2	11	98.26	AT3G63490

Table 3.1. Continued.

Spot #	Protein Identification	Exp. pI	Exp./ Theo. MW (kDa)	Pept--ides #	C.I.%*	Gene/locus
332	5-methyltetrahydropteroyltriglutamate--homocysteine Methyltransferase (EC 2.1.1.14)	6.9	92.8 / 84.3	07	100.00	AT5G17920
333	Unknown protein	7.3	61.8 / 29.1	10	100.00	AT1G10070
334	Chaperonin subunit, putative	7.2	61.7 / 57.7	11	100.00	AT3G18190
335	Ribulose 1,5-bisphosphate carboxylase/oxygenase	7.3	39.8 / 47.6	16	100.00	ATCG00490
336	Glyceraldehyde 3-phosphate dehydrogenase A subunit	6.8	41.5 / 31.6	04	100.00	AT3G26650
337	Ribulose-bisphosphate carboxylase	6.7	40.5 / 33.3	10	100.00	ATCG00490
338	Glyceraldehyde 3-phosphate dehydrogenase A subunit	6.6	40 / 30	13	100.00	AT3G26650
339	MKM21_30 (60S ribosomal protein L5)	8.5	41.4 / 52.9	24	100.00	AT3G25520
340	Asparagine synthetase	5.6	75 / 41.1	15	100.00	AT5G10240
341	Glycyl-tRNA synthetase (EC 6.1.1.14) (Glycine--tRNA ligase) (GlyRS)	5.6	86.4 / 73.4	19	100.00	AT1G29880
342	Late embryogenesis abundant (LEA) domain-containing protein	4.1	43.3 / 26.6	09	100.00	AT5G44310
343	Ribosomal protein, putative	8.5	37.9 / 73.6	27	100.00	AT3G25520
344	Chloroplast 30S ribosomal protein S4	7.1	26.2 / 28.3	13	100.00	ATCG00380
345	Vegetative storage protein-like	5.5	29 / 67	17	100.00	AT5G44020
346	Proteasome subunit alpha type 3 (EC 3.4.25.1) (20S proteasome alpha subunit G)	5.5	31.8 / 61.9	07	100.00	AT2G27020
347	Putative 5-methyltetrahydropteroyltriglutamate—homocysteine S-methyl transferase	7.7	90.1 / 47.6	14	100.00	AT5G17920
348	Hypothetical protein	6.5	84.2 / 66.6	13	96.70	AT3G62010
349	Hypothetical protein	5.8	52.3 / 70.8	14	99.80	AT5G01060
350	Glyceraldehyde-3-phosphate dehydrogenase (NADP) (EC 1.2.1.13) A precursor	6.4	42.6 / 30.2	16	100.00	AT3G26650
351	Similar to NADP-specific isocitrate dehydrogenase	5.6	50.8 / 47.6	17	100.00	AT1G65930
352	Ribulose bisphosphate carboxylase large chain precursor (EC 4.1.1.39)	5.6	58.5 / 50.7	17	100.00	ATCG00490
353	Threonine synthase	5.6	61.7 / 24.1	11	100.00	AT4G29830
354	Ribulose bisphosphate carboxylase large chain precursor (EC 4.1.1.39)	5.6	52.5 / 24.1	13	100.00	ATCG00490
355	Ribulose bisphosphate carboxylase large chain precursor (EC 4.1.1.39)	5.5	58.2 / 45.1	07	100.00	ATCG00490
356	Ribulose bisphosphate carboxylase large chain precursor (EC 4.1.1.39)	5.5	56.7 / 42.4	15	100.00	ATCG00490
357	Myosin heavy chain -related protein	5.4	75.1 / 43.9	11	100.00	AT3G49055
358	Ribulose bisphosphate carboxylase large chain precursor (EC 4.1.1.39)	5.4	57.1 / 52	13	100.00	ATCG00490
359	Mitochondrial RNA helicase-like protein	4.8	88.6 / 25.1	02	99.96	AT5G39840
360	Probable RAD50 DNA repair protein	4.8	100.8 / 24.2	08	100.00	AT2G31970
361	Multicatalytic endopeptidase complex, proteasome component, alpha subunit	5.4	31.8 / 30.7	10	100.00	AT5G35590
362	CAB binding protein,photosystem II (Fragment)	4.9	30.8 / 41.1	16	100.00	AT2G34420
363	Gb|AAF00675.1	4.6	43.5 / 49.9	09	99.28	AT3G20350
364	Putative glucose-6-phosphate isomerase	4.8	63.9 / 56.7	12	100.00	AT4G24620
365	H+-exporting ATPase (EC 3.6.3.6) 57K chain	4.8	61 / 20.3	12	100.00	AT1G76030
366	Luminal binding protein 1 precursor (BiP1) (AtBP1)	4.9	83.6 / 166.2	24	99.71	AT5G13820
367	Cell division protein ftsH homolog 2, chloroplast precursor (E.C.3.4.24.-)	4.8	71.9 / 52.3	13	99.97	AT5G42270
368	Calmodulin-7	4.0	15.2 / 35.1	12	100.00	AT3G43810
369	Putative glucose-6-phosphate isomerase	4.8	65 / 57.7	15	100.00	AT4G24620
370	ATP synthase alpha chain (EC 3.6.3.14)	4.8	61.5 / 34.5	13	100.00	ATCG00120
371	Thylakoid lumenal 15 kDa protein, chloroplast precursor	5.2	15.5 / 55.3	21	100.00	AT2G44920

Table 3.1. Continued.

Spot #	Protein Identification	Exp. pI	Exp./ Theo. MW (kDa)	Pept--ides #	C.I.%*	Gene/locus
372	RuBisCO small subunit 1B, chloroplast precursor (EC 4.1.1.39)	5.0	15.5 / 49.1	20	100.00	AT5G38430
373	Genomic DNA, chromosome 5, P1 clone:MDC12	5.2	18 / 41.4	10	99.96	AT5G63280
374	Hypothetical protein (F5K20.290)	5.5	18.1 / 22.7	03	95.75	AT3G53990
375	Unknown protein (Hypothetical protein)	5.6	16.3 / 31	08	100.00	AT4G09320
376	F15O4.13	5.9	17.8 / 207.8	23	99.56	AT1G35647
377	F28I16_90 (Hypothetical protein)	6.9	21.9 / 67.8	13	100.00	AT5G19940
378	Putative photosystem II type I chlorophyll a/b binding protein	4.7	28.8 / 47.6	16	100.00	AT2G34430
379	Glyoxalase I, putative	4.8	25 / 84.3	18	100.00	AT1G08110
380	AIG2-like protein (Hypothetical protein)	4.7	22.5 / 30.2	04	99.97	AT5G39730
381	Non-LTR retroelement reverse transcriptase -related	4.7	20.3 / 40.7	08	100.00	AT2G06290
382	50S ribosomal protein L12-1, chloroplast precursor (CL12-A)	4.5	19.3 / 84.3	09	100.00	AT3G27830
383	Putative 6-phosphogluconolactonase	5.4	31 / 47.6	14	100.00	AT5G24400
384	Thylakoid lumenal 21.5 kDa protein, chloroplast precursor	5.1	27.7 / 85.9	22	100.00	AT4G15510
385	Carbonic anhydrase, chloroplast precursor (EC 4.2.1.1)	5.7	28.1 / 65	12	100.00	AT3G01500
386	GTP-binding nuclear protein RAN-2	5.8	28.4 / 84.3	21	100.00	AT5G20020
387	Putative superoxide dismutase [Mn/Fe] (EC 1.15.1.1)	5.5	26.2 / 62.2	12	99.98	AT4G25100
388	Hypothetical protein	7.0	8 / 75.2	13	99.99	AT1G63270
389	40S ribosomal protein S10-3	8.6	19.8 / 52.9	07	99.33	AT5G52650
390	Polygalacturonase inhibitor-like protein (Hypothetical protein)	8.0	41 / 56.8	14	99.37	AT3G20820
391	Reversibly glycosylated polypeptide-1	5.3	41.8 / 40.7	10	100.00	AT3G02230
392	Glyceraldehyde 3-phosphate dehydrogenase B,chloroplast precursor (EC 1.2.1.13)	5.5	43 / 56.7	13	100.00	AT3G26650
393	Inorganic pyrophosphatase-like protein (EC 3.6.1.1)	4.6	44.7 / 20.1	07	98.10	AT5G09650
394	Phosphoribosylformylglycinamidine cyclo-ligase, chloroplast precursor (EC 6.3.3.1)	4.6	33.7 / 60.5	18	100.00	AT3G55010
395	Tubulin alpha-6 chain	5.0	49.3 / 48.6	12	100.00	AT4G14960
396	Hypothetical protein	4.5	56.1 / 65.1	10	100.00	AT5G39570
397	Ribosomal protein L17-like protein (F24B22_170)	5.3	65.4 / 135.7	09	95.44	AT3G54210
398	Putative GDP-mannose pyrophosphorylase (F9E10_24)	5.7	47.4 / 22.8	07	99.99	AT1G74910
399	Gb|AAD20127.1 (K1L20_3)	5.6	36.8 / 20.1	05	99.97	AT5G42270
400	3-deoxy-D-arabino-heptulosonate 7-phosphate synthase	7.7	52.2 / 27.6	10	100.00	AT4G33510
401	Heat-shock protein	4.5	97.4 / 15.9	07	100.00	AT1G79930
402	GloEL protein, chaperonin, 60 kDa	5.3	60.5 / 43.3	06	99.77	AT3G13470
403	Phosphoglycerate dehydrogenase-like protein	5.4	59.6 / 79.9	20	100.00	AT4G34200
404	Genomic DNA, chromosome 3, P1 clone: MVC8 (Hypothetical protein)	4.5	92.7 / 59.3	10	99.43	AT3G15950
405	T6D22.14	4.7	118.1 / 82.5	14	99.69	AT1G08060
406	Cytosolic factor, putative	4.7	118.8 / 31.4	09	99.99	AT1G72150
407	Expressed protein	4.6	66.3 / 39.2	11	100.00	AT1G73850
408	Jasmonate inducible protein ISOLOG (Putative lectin protein)	5.3	79 / 36.2	20	100.00	AT3G16460
409	Isocitrate dehydrogenase (EC 1.1.1.42)	5.5	44.7 / 63.3	13	100.00	AT2G17130
410	Protein kinase	5.6	47.7 / 28.2	08	100.00	AT2G39660
411	Gb|AAF07790.1	5.2	21.1 / 71.6	12	99.97	AT5G47690
412	Probable guanylate kinase	8.1	18.4 / 25.1	08	100.00	AT2G41880
413	Expressed protein	5.1	51.2 / 32.1	10	100.00	AT1G73850
414	Genomic DNA, chromosome 3, P1 clone: MOE17	4.9	42.1 / 138.5	18	99.37	AT3G20760
415	Putative ATP-dependent RNA helicase A	4.9	40.3 / 50.3	12	100.00	AT2G01130

Table 3.1. Continued.

Spot #	Protein Identification	Exp. pI	Exp./ Theo. MW (kDa)	Pept--ides #	C.I.%*	Gene/locus
416	Hypothetical protein	4.4	35.2 / 66.6	13	99.84	AT5G16110
417	At1g24030	4.6	36.4 / 81.9	17	100.00	AT1G24030
418	Hypothetical protein (F3P11.21)	5.3	33 / 45.7	12	100.00	AT2G19610
419	Hypothetical protein	5.5	24.3 / 40.6	14	100.00	AT2G03810
420	Plant transposase (Ptta/En/Spm) family	5.0	21.6 / 42.4	17	100.00	AT2G10070
421	Putative cytochrome P450 monooxygenase	5.0	21.1 / 51.8	16	100.00	AT1G13100
422	TL17_ARATH Thylakoid lumenal 17.4 kDa protein	4.8	17.2 / 30.1	12	100.00	AT5G53490
423	Major latex-like protein (T23E23_22)	4.9	16.6 / 27.4	10	100.00	AT1G24020
424	Cytochrome P450, putative	4.7	17.3 / 35.8	10	100.00	AT5G61320
425	F14J9.3 protein	4.6	16.4 / 82.1	19	100.00	AT1G09370
426	Proteasome subunit beta type 2-1 (EC 3.4.25.1) (20S proteasome alpha subunit D1)	5.6	25.5 / 73.4	26	100.00	AT3G22630
427	BELONGS to the L5P family of ribosomal proteins (Putative L5 ribosomal protein)	8.9	28.3 / 55.3	16	100.00	AT4G01310
428	F15O4.36	7.6	7 / 23.8	08	99.61	AT1G35580
429	Hypothetical protein	8.1	12.1 / 63.3	15	100.00	AT4G21820
430	L-ascorbate peroxidase, cytosolic (EC 1.11.1.11)	5.5	29.6 / 49.8	12	100.00	AT1G07890
431	RING finger-like protein (T12E18_50)	5.5	30.8 / 18.4	05	95.84	AT3G54360
432	Lipoamide dehydrogenase	5.8	57.8 / 34.6	10	95.12	AT1G48030
433	F1O19.9 protein	6.6	63.1 / 67.6	11	99.87	AT1G67035
434	T19G15_130	8.0	34.1 / 52.9	25	100.00	AT5G26280
435	Hypothetical protein	8.2	12.3 / 94.1	25	100.00	AT5G60960
436	At2g47710 protein (Hypothetical protein)	7.0	18.2 / 55.3	20	100.00	AT2G47710
437	Cytosolic cyclophilin ROC3 (EC 5.2.1.8)	7.6	18.8 / 89.9	17	100.00	AT2G16600
438	Hypothetical protein	7.7	23.1 / 44.5	13	99.84	AT5G16110
439	Expressed protein	8.8	19.6 / 42.3	13	100.00	AT5G66250
440	Ribosomal protein L17-like protein	9.3	20.5 / 15.7	07	100.00	AT3G54210
441	40S ribosomal protein S9	9.4	23.7 / 40.3	11	99.96	AT5G39850
442	Hypothetical protein	8.5	21.7 / 63.3	23	100.00	AT5G47190
443	Glycosyl hydrolase family 85	6.8	16.8 / 33.8	13	100.00	AT3G61010
444	self-incompatibility protein-related	6.9	15.8 / 32.3	09	100.00	AT3G16970
445	GTP-binding protein SAR1B	6.9	23 / 68.6	13	100.00	AT1G56330
446	Ribulose bisphosphate carboxylase large chain precursor (EC 4.1.1.39)	5.4	53.1 / 44.9	11	100.00	ATCG00490
447	Ribulose bisphosphate carboxylase large chain precursor	5.4	52.8 / 49.1	15	100.00	ATCG00490
448	Ketol-acid reductoisomerase	5.4	58.5 / 43	13	100.00	AT3G58610
449	Glyceraldehyde-3-phosphate dehydrogenase (NADP) (phosphorylating) (EC 1.2.1.13)	5.5	43.6 / 40.3	13	100.00	AT1G42970
450	Glutamyl-tRNA reductase 2 (GluTR) (HEMA2)	4.7	92.5 / 48	16	100.00	AT1G09940
451	Hypothetical protein	4.8	88.3 / 43.5	15	99.99	AT1G35480
452	Hypothetical protein	4.6	71.6 / 53.8	15	99.99	AT1G25682
453	Heat shock protein 70	4.7	72.8 / 47.5	11	99.15	AT5G49910
454	Oxygen-evolving enhancer protein 2-1, chloroplast precursor (OEE2)	5.2	26.5 / 73.6	15	100.00	AT1G06680
455	Phosphoribulokinase, chloroplast precursor (EC 2.7.1.19) (Phosphopentokinase)	5.1	42.5 / 48.6	18	100.00	AT1G32060
456	Guanine nucleotide-binding protein beta subunit-like protein (WD-40 repeat auxin-dependent protein ARCA)	7.0	35.5 / 68	13	100.00	AT1G18080
457	F17O7.5 protein (Carbonic anhydrase, putative)	6.0	29.9 / 99.4	14	100.00	AT1G70410
458	Disease resistance protein AIG1, 916-2572	6.1	29.4 / 86.2	17	99.95	AT1G33960

Table 3.1. Continued.

Spot #	Protein Identification	Exp. pI	Exp./ Theo. MW (kDa)	Pept--ides #	C.I.%*	Gene/locus
459	Probable phospholipid hydroperoxide glutathione peroxidase (EC 1.11.1.9) (AtGPX1)	5.7	73.8 / 55.3	16	100.00	AT4G11600
460	Plastid ribosomal protein S6, putative	4.5	21.7 / 45.7	14	100.00	AT1G64510
461	Putative flavonol sulfotransferase	5.3	40.6 / 54.7	21	100.00	AT1G74100
462	T28K15.4 protein (Expressed protein)	5.3	41.6 / 17	05	99.92	AT1G12230
463	Ribulose bisphosphate carboxylase large chain precursor (EC 4.1.1.39)	5.6	52.9 / 41.9	17	100.00	ATCG00490
464	Expressed protein	4.3	29 / 24.6	08	100.00	AT5G26880
465	Cyclophilin-like protein (EC 5.2.1.8) (PPIase) (Rotamase)	8.0	19.8 / 29	05	97.92	AT3G55920
466	Glycyl tRNA synthetase, putative	5.6	67.8 / 75.2	15	100.00	AT1G29880
468	Acetolactate synthase-like protein	5.7	49.3 / 44.5	13	99.97	AT5G16290
469	Putative chloroplast translation elongation factor EF-Tu precursor	7.2	38.6 / 59.7	12	100.00	AT4G20360
470	Calreticulin	5.2	20 / 22.6	11	100.00	AT1G09210
472	Hypothetical protein	5.0	18.9 / 40.6	12	100.00	AT4G00695
473	Hypothetical protein common family	8.4	15.5 / 42.4	13	100.00	AT3G30848
474	Putative RAS superfamily GTP-binding protein	4.3	12.8 / 37.4	14	100.00	AT2G31680
475	Expressed protein	4.5	69.3 / 27.1	13	100.00	AT1G73850
477	GTP-binding protein SAR1A	6.4	24.7 / 53.9	14	100.00	AT4G02080
478	Major latex protein (Major latex protein like) (dl3070w)	5.9	21.9 / 45.1	08	100.00	AT4G14060
479	Hypothetical protein	8.2	7.8 / 85	30	100.00	ATMG00970
480	Glycine-rich RNA-binding protein (clone A81)	5.0	14.3 / 35	06	100.00	AT4G13850
481	Asparaginyl-tRNA synthetase, cytoplasmic 1 (EC 6.1.1.22)	5.3	73.4 / 53.3	11	100.00	AT5G56680
482	ATP synthase gamma chain, (EC 3.6.1.34) (Fragment)	5.8	38.7 / 91.6	11	100.00	AT4G04640
483	Type III chlorophyll A/B binding protein	4.5	28.3 / 94.5	10	99.62	AT5G54270
484	Elongation factor 1B alpha-subunit	4.4	32.7 / 10.8	03	100.00	AT5G19510
485	Spermidine synthase 1 (EC 2.5.1.16)	4.5	38.9 / 29.4	09	100.00	AT1G23820
486	3-isopropylmalate dehydratase, small subunit	4.8	26.4 / 8.6	03	100.00	AT2G43090
487	14-3-3 protein homolog RCI2	4.6	30.5 / 18.6	02	100.00	AT5G10450
488	F5J5.1	4.7	46.3 / 37.9	07	99.71	AT1G36035
489	Proteasome subunit beta type 3-1 (EC 3.4.25.1)	5.1	26.1 / 57.7	06	99.99	AT1G21720
490	Triose-phosphate isomerase (EC 5.3.1.1)	5.2	29.3 / 51.6	07	100.00	AT3G55440
491	T22H22.19 protein (T22H22_19)	5.2	24.9 / 58.3	07	100.00	AT1G54780
492	Putative lectin	5.2	35.5 / 52.9	19	100.00	AT1G45214
493	Putative fructose bisphosphate aldolase (F3K23.9)	5.2	39.9 / 40.4	10	100.00	AT2G21330
494	Actin related protein 2	5.2	42.7 / 23	06	99.23	AT3G27000
495	Putative aldolase	5.7	37.3 / 26.5	08	100.00	AT2G01140
496	F14L17.18 protein (Hypothetical protein)	6.2	26.6 / 45	08	99.90	AT1G14410
497	Putative vegetative storage protein	5.7	28.8 / 44	06	99.56	AT5G44020
498	Endo-XYLOGLUCAN transferase-like protein	6.7	35.5 / 46.3	09	99.60	AT4G37800
499	Actin depolymerizing factor 2	4.9	17.5 / 41.3	05	100.00	AT3G46000
500	10-formyltetrahydrofolate synthetase (EC 6.3.4.3)	5.7	72.4 / 66.6	13	99.99	AT1G50480
501	F5F19.16 protein	6.4	54.7 / 66.6	11	100.00	AT1G52100
502	Malate dehydrogenase, (EC 1.1.1.37)	5.6	36.8 / 19.5	07	100.00	AT1G53240
503	Aldose 1-epimerase-like protein (EC 5.1.3.3)	8.2	40.7 / 45.5	12	100.00	AT3G47800

Note: Spot #: The spot number is based on the order of mass analysis. Exp. pI: Experimental pI; Exp./Theo. MW: Experimental/Theoretical molecular weight; Identified Peptides #: Number of peptides that matched with the identified protein in mass analyses. C.I. %: Cross Confidence Interval %; Proteins were extracted from 10 days old cotyledons and separated using 2-D gels. The protein spots were excised, in gel digested, and analyzed using MALDI TOF/TOF mass spectrometer (ABI 4700 proteomics Analyzer, Applied Biosystems, CA). Protein identification was performed using the Result dependent Analysis (RDA) of ABI GPS explorer software (version 3.5).

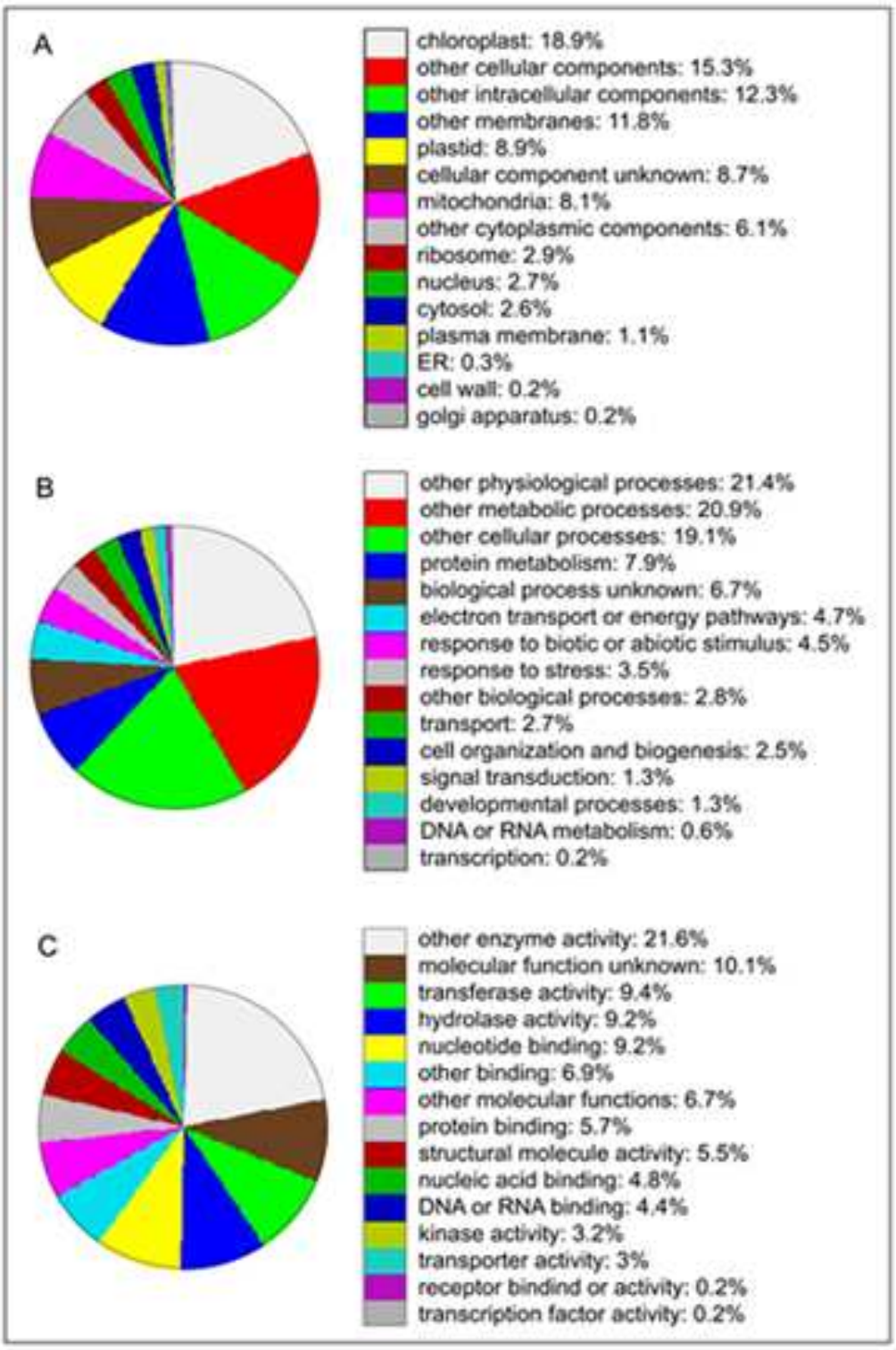

Figure 3.2: Functional classification of cotyledon 2-DE proteome.

Note: (A) Cellular localization, (B) Biological process, (C) Molecular function.

Protein identification by MudPIT

In order to increase the proteome coverage of cotyledon, Multidimensional Protein Identification Technology (MudPIT) has been employed to analyze the same protein sample. Protein samples with three biological replicates were prepared. The proteins with stringent filter parameters (Xcorr = +1-1.9; +2-2.2; +3-3.75, Delta Corr = 0.1, pFactor -0.001), unique peptide were considered as positive hits and obtained 662 protein identities and listed in **Table 3.3**. The genes corresponding to all these proteins were characterized using Gene Ontology rules according to their Cellular localization, Biological process and Molecular function. They were grouped into different levels and generated pie charts (**Figure 3.3**).

Protein classification based on cellular localization indicated that 12.6% of the identified genes encoded chloroplast proteins and 8.8% encoded mitochondria proteins. Proteins in other organelles and cellular compartments were also shown in Figure 3.3A. Gene Ontology analyses based on biological processes indicated that 21.2% of the genes were involved in Other Physiological Processes and 20.6% genes were involved in Other Metabolic Process (Figure 3.3B). In addition, the analyses based on Molecular Function showed that 23.7% of proteins possessed Other Enzyme Activities and 11.7% with structural molecule activity (Figure 3.3C). Proteins in other biological processes and molecular function categories were presented in detail as shown in Figures 3.3B and 3.3C respectively.

Table 3.2: List of Proteins identified by MudPIT in *Arabidopsis thaliana* cotyledon

S.No.	Protein Identification	pI / MW	P (pro)	Gene/locus
1	B Chain B, Crystal Structure Of Thi1 Protein	6 / 30049.5	3.15E-05	At5g54770
2	B Chain B, X-Ray Structure Of Gene Product	6 / 27058.6	1.77E-04	At5g02240
3	A Chain A, Crystal Structure Of The *Arabidopsis Thaliana* O-Acetylserine	4.3 / 33747.8	1.80E-04	At4g14880
4	B Chain B, Gdp-Mannose-3', 5' –Epimerase K178R, With G	6 / 42914.4	5.66E-07	At5g28840
5	B Chain B, Gdp-Mannose-3', 5' –Epimerase With Gdp-Alph	6 / 42886.4	5.66E-07	At5g28840
6	B Chain B, Gdp-Mannose-3', 5' –Epimerase Y174F	6 / 42870.4	5.66E-07	At5g28840
7	B Chain B, Gdp-Mannose-3', 5' –Epimerase K217A	6 / 42829.3	5.66E-07	At5g28840
8	Phosphoglycerate Kinase	4.3 / 23868.5	4.54E-06	AT1G56190
9	Gf14 Chi Chain	4.3 / 29901.4	7.40E-07	At4g09000
10	Adp Glucose Pyrophosphorylase Small Subunit	6 / 56598.1	1.36E-05	At5g48300
11	Heat-Shock Protein	4.25 / 80035.7	4.45E-04	At5g56010
12	Beta-Glucosidase	6 / 59986.6	3.08E-08	AT3G09260
13	Germin-Like Protein	9.8 / 21057.2	1.11E-04	At1g72610
14	Phosphoglycerate Kinase	4.3 / 41906	3.04E-11	At1g56190
15	Beta-Glucosidase	6 / 60052.7	5.36E-05	AT1G66270
16	Strong Similarity To 60S Ribosomal Protein L17	10.5 / 16998.1	6.80E-04	AT1G04480
17	Putative Transketolase Precursor	6 / 68851.3	1.03E-04	At2g45290
18	Ribosomal Protein L23A	10.5 / 17472.6	3.64E-05	At2g39460
19	Photosystem Ii Oxygen-Evolving Complex 23K Protein, Putative	10.1 / 27720.9	5.71E-05	At2g30790
20	Clpc	6 / 103455	1.34E-05	AT5G50920
21	Mitochondrial Chaperonin (Hsp60)	4.3 / 55253.3	1.08E-05	At2g33210
22	Chaperonin 10	10.1 / 26928.8	1.69E-05	At5g20720
23	Catalase 1	6 / 56860.8	3.32E-07	At1g20630
24	Actin 3	4.3 / 37189.4	1.29E-04	At2g37620
25	Manganese Superoxide Dismutase	9.435 / 25491.8	5.10E-04	At3g10920
26	Acc Oxidase	4.3 / 36168.8	4.87E-04	AT1G62380
27	Aft Protein	4.25 / 39993.6	6.75E-05	At4g35450
28	Catalase 3	8.3 / 56696.1	3.32E-07	At1g20620
29	Glyoxalase Ii Cytoplasmic Isozyme	6 / 28826.9	4.09E-09	At3g10850
30	Glutathione S-Transferase (Gst6)	6 / 24076.6	1.25E-05	At2g47730
31	Putative Pectin Methylesterase	10.1 / 64358.7	2.88E-07	AT4G03926
32	Glyceraldehyde 3-Phosphate Dehydrogenase A Subunit	6 / 37674.8	5.22E-07	At3g26650
33	Glyceraldehyde 3-Phosphate Dehydrogenase B Subunit	6 / 42795.6	3.94E-12	At1g42970
34	Ankyrin Repeat-Containing Protein 2	4.25 / 36968	6.75E-05	At4g35450
35	Putative Ty3-Gypsy-Like Retroelement Pol Polyprotein	4.3 / 67931	2.22E-04	At2g06170
36	Strong Similarity To Gb	4.3 / 44049.4	2.04E-08	At1g31180
37	Belongs To The Pf	6 / 47729.2	3.72E-04	AT1G54010
38	Putative Alpha-Carboxyltransferase	6 / 88515.6	2.38E-05	At2g38040
39	T10O24.18	10.53 / 17414.6	4.05E-04	AAD39578
40	Putative 60S Acidic Ribosomal Protein, 5' Partial	4.3 / 24295.3	5.42E-06	At3g09200
41	Phosphoglycerate Kinase	10.5 / 33883.9	1.25E-06	At1g56190
42	Elongation Factor Ef-2	6 / 94246.1	3.28E-05	AT1G56070
43	F22C12.4	6 / 35836.9	1.18E-04	AT1G64200
44	T19E23.15	10.5 / 15751.5	2.28E-06	AT1G31355
45	Heat Shock Protein 70	4.3 / 77105.2	1.59E-06	AT5G49910
46	Ribulose-1,5-Bisphosphate Carboxylase	4.3 / 3641.1	4.88E-08	ATCG00490

Table 3.2. Continued.

S.No.	Protein Identification	pI / MW	P (pro)	Gene/locus
47	F12M16.14	8.3 / 36991.5	8.07E-10	At1g53240
48	F3F9.11	5.125 / 76998.6	8.46E-06	AT1G78380
49	F18O14.33	6 / 50160	9.40E-07	AT1G19310
50	F21D18.28	6 / 53785.4	1.19E-10	At1g48030
51	T6D22.2	10.1 / 107174.4	3.60E-11	AT1G07920
52	Contains Similarity To Ferredoxin-Nadp+ Reductase	8.3 / 38829.2	4.80E-05	AT1G20020
53	F2D10.11	8.3 / 116681.4	3.32E-07	At1g20620
54	F10A5.6	10.1 / 14501.2	2.59E-08	At1g75750
55	F10A5.19	10.5 / 26772.1	7.82E-07	At1g75600
56	Profilin	4.25 / 14013.8	5.67E-05	At4g29350
57	Hypothetical Protein	10.1 / 31880.5	7.56E-04	AT1G12250
58	At4G28750	10.5 / 11710.2	8.24E-07	AT4g28750
59	At1G76180	4.3 / 20782.2	4.11E-07	At1g76180
60	At1G67090	8.3 / 20315.1	7.53E-06	At1g67090
61	At3G62030	9 / 28208.1	8.10E-06	At3g62030
62	Putative Photosystem I Subunit Psi-E	10.5 / 12253.8	8.24E-07	At4g28750
63	Ribosomal Protein L9, 5' Partial	10.5 / 16953.6	1.42E-04	AT5G53070
64	F1O19.10/F1O19.10	5.1 / 14698.8	7.53E-06	At1g67090
65	At5G66570/K1F13_25	4.3 / 35128.1	4.28E-06	At5G66570
66	Unknown Protein	10.5 / 10561	6.10E-06	At3g47070
67	Putative Elongation Factor	6 / 73989.2	5.28E-07	At1g56075
68	Putative Fructose Bisphosphate Aldolase	6 / 42888.6	1.64E-08	At2g21330
69	Unknown Protein	8.3 / 30091.6	7.56E-04	At1g12250
70	Putative Dnak-Type Molecular Chaperone Hsc70.1 Protein	4.3 / 23515	3.15E-04	At5g02500
71	Unknown Protein	4.3 / 85999	5.60E-05	At1g62750
72	Putative Methionine Synthase	6 / 84610.3	3.52E-04	At3g03780
73	Alanine Aminotransferase-Like Protein	4.3 / 42174	4.64E-04	AT1G23310
74	Putative Cytosolic O-Acetylserine(Thiol)Lyase	5.1 / 33641.7	1.80E-04	At4g14880
75	Putative Dihydrolipoamide S-Acetyltransferase	7.2 / 38103.5	2.82E-04	At3g25860
76	At1G07930/T6D22_3	10.1 / 49451	3.60E-11	At1G07930
77	At5G56010/Mda7_5	4.25 / 60844.8	4.45E-04	At5g56010
78	Putative Peroxiredoxin Protein	10.1 / 24712.2	6.78E-05	At3g52960
79	At2G39730/T5I7.3	4.3 / 52039	3.53E-08	At2G39730
80	33 Kda Polypeptide Of Oxygen-Evolving Complex	4.3 / 35188.2	4.28E-06	At5g66570
81	Catalase	6 / 56887.9	9.62E-06	At4g35090
82	At5G13450/T22N19_100	10.1 / 26294.3	1.62E-05	At5g13450
83	At5G26000/T1N24_7	6 / 61172.4	5.04E-05	At5g26000
84	At2G36530/F1O11.16	6 / 47704	8.51E-05	At2G36530
85	At4G37930/F20D10_50	8.6 / 57399.3	8.31E-08	At4g37930
86	At5G17920/Mpi7_60	6 / 84313	1.54E-07	At5G17920
87	At4G30920/F6I18_170	6 / 61119.6	6.13E-05	At4g30920
88	At5G09660/F17I14_150	8.3 / 37383	2.15E-05	At5G09660
89	At2G36530/F1O11.16	4.3 / 34913.4	8.51E-05	At2G36530
90	At4G14960/Dl3520C	4.3 / 49502.8	8.38E-09	At4G14960
91	At4G38970/F19H22_70	7.2 / 42941.5	3.28E-08	At4G38970
92	At3G44310/T10D17_100	6 / 38178	6.30E-07	At3G44310
93	At3G14420/Moa2_2	10.1 / 40313.1	2.38E-05	At3g14420

Table 3.2. Continued.

S.No.	Protein Identification	pI / MW	P (pro)	Gene/locus
94	At1G23310/F26F24_4	7.2 / 53299.9	3.24E-05	At1G23310
95	At5G26000/T1N24_7	5.1 / 61106.3	5.04E-05	At5g26000
96	At1G35160/T32G9_30	4.3 / 30178.7	7.40E-07	At1G35160
97	At1G02500/T14P4_22	9.9 / 40471.1	2.62E-07	At1g02500
98	Ubiquitin / Ribosomal Protein Cep52	4.3 / 35672.9	6.13E-05	At3g52590
99	G5Bf Protein	9.6 / 42618.5	3.03E-10	At1g09340
100	Putative Glycine-Rich Protein	10.1 / 9869.1	4.30E-04	At4g39260
101	Putative Heat Shock Protein 90	4.25 / 80085.5	4.45E-04	At5g56010
102	At4G33010/F26P21_130	6 / 112943.6	2.72E-04	At4g33010
103	Putative Enolase (2-Phospho-D-Glycerate Hydroylase)	5.1 / 47728	8.51E-05	At2g36530
104	Glycine-Rich Rna Binding Protein	10.1 / 10152.3	8.51E-07	At2g21660
105	Unknown Protein	8.3 / 15050.4	1.05E-04	At2g44650
106	Unknown Protein	6 / 57526.9	1.73E-04	At5g14260
107	At4G22240/T10I14_70	4.3 / 33731.3	6.48E-05	At4g22240
108	At5G14200/Mua22_20	5.1 / 44131.6	1.97E-04	At5g14200
109	At2G47730/F17A22.12	4.3 / 19741.5	1.25E-05	At2G47730
110	At2G19940/F6F22.3	8.3 / 44163.6	2.03E-05	At2g19940
111	At4G28520/F20O9_210	6 / 58243.9	1.36E-06	At4G28520
112	Glycolate Oxidase	7.2 / 19073.8	2.38E-05	At3g14420
113	Germin-Like Protein	10.1 / 21528.9	1.11E-04	At1g72610
114	Enolase (2-Phospho-D-Glycerate Hydroylase)	4.3 / 47777	8.51E-05	At2g36530
115	Similar To Nadp-Specific Isocitrate Dehydrogenase	6 / 45731.9	1.37E-05	At1g65930
116	Fructose-Bisphosphate Aldolase-Like Protein	8.3 / 38657.8	2.74E-04	AT5G03690
117	At2G04030/F3C11.14	4.3 / 88613.3	1.05E-05	At2G04030
118	Unknown Protein	4.25 / 48827.3	2.08E-04	At4g29060
119	At3G14210/Mag2_18	8.3 / 44090	8.29E-04	At3G14210
120	At1G03480/F21B7_24	5.1 / 34173.2	2.01E-05	At1g03475
121	Phosphoglycerate Dehydrogenase-Like Protein	6 / 63309.6	6.66E-06	AT1G17745
122	Phosphoglycerate Kinase, Putative	4.3 / 42147.5	4.59E-05	AT1G56190
123	Putative Alanine Aminotransferase	6 / 53476.2	4.64E-04	AT1G17290
124	Ribosomal Protein L4	10.1 / 30584.5	1.39E-05	At1g07320
125	Glycine-Rich Rna Binding Protein 7	5.1 / 16888.8	7.47E-05	At2g21660
126	Fructose-Bisphosphate Aldolase-Like Protein	10.1 / 43002.9	8.49E-10	AT2G01140
127	Putative (Mn) Superoxide Dismutase	9.435 / 25416.7	5.10E-04	AT3G56350
128	Unknown	4.3 / 16188.5	1.68E-06	AT1G13930
129	Transketolase-Like Protein	6 / 79924.8	2.97E-07	AT2G34590
130	Unknown	8.3 / 22077.9	1.26E-07	AT4G21860
131	Nuclear Rna Binding Protein A-Like Protein	10.07 / 37956.2	1.14E-04	AT4G16830
132	Endomembrane-Associated Protein	4.3 / 24540.6	1.41E-07	AT4G20260
133	Putative Rubisco Subunit Binding-Protein Alpha Subunit	4.3 / 62129.6	5.40E-08	AT2G28000
134	Profilin 2	4.25 / 13982.8	5.67E-05	AT4G29350
135	Putative 60S Acidic Ribosomal Protein P0	4.3 / 34089.8	5.42E-06	At3g09200
136	Ribosomal Protein L9, Putative	10.5 / 22043.6	1.42E-04	AT1G33120
137	Photosystem I Reaction Center Subunit Iv B, Chloroplast Precursor	10.5 / 14987	8.24E-07	AT2G20260
138	Atpm24.1 Glutathione S Transferase	6 / 24052.3	4.86E-08	AT4G02520
139	Peptidylprolyl Isomerase Roc4	8.3 / 28196	8.10E-06	At3g62030

Table 3.2. Continued.

S.No.	Protein Identification	pI / MW	P (pro)	Gene/locus
140	Unknown	4.25 / 16614.1	3.40E-06	T1G62480
141	60S Ribosomal Protein L23A	10.5 / 17382.5	3.64E-05	At2g39460
142	Glutathione Transferase, Putative	6 / 25749.5	8.46E-06	AT1G02920
143	Inorganic Pyrophosphatase-Like Protein	4.3 / 33393.7	6.81E-08	At5g09650
144	Mitochondrial Nad-Dependent Malate Dehydrogenase	8.3 / 35776.1	8.07E-10	At1g53240
145	Phosphoserine Aminotransferase	8.33 / 47312.9	8.87E-04	AT1G08490
146	3-Methyladenine Dna Glycosylase, Putative	4.3 / 43889.2	2.04E-08	AT1G31180
147	Cinnamyl-Alcohol Dehydrogenase Cad1	6 / 38932.5	1.79E-06	AT5G44070
148	Unknown	10.1 / 33853.4	6.76E-04	At2g20890
149	Unknown	7.2 / 56723	4.42E-05	At1g20620
150	Putative Rna-Binding Protein	8.3 / 42604.5	3.03E-10	AT1G32790
151	33 Kda Polypeptide Of Oxygen-Evolving Complex (Oec) In Photosystem II	4.3 / 35158.1	4.28E-06	AT3G50820
152	Glyceraldehyde-3-Phosphate Dehydrogenase C Subunit (Gapc)	6 / 36989.1	2.78E-07	AT3G04120
153	Putative Porin	10.1 / 29439.2	1.36E-04	AT3G01280
154	Cytosolic Malate Dehydrogenase	7.2 / 35661.9	6.90E-05	AT5G43330
155	Actin 4	4.3 / 41826.6	1.29E-04	At5g59370
156	Isocitrate Dehydrogenase, Putative	6 / 45759	1.37E-05	AT1G54340
157	Cinnamyl Alcohol Dehydrogenase	4.3 / 39095.9	4.09E-05	AT3G19450
158	Putative Lectin	5.1 / 32115.7	7.73E-05	AT1G45258
159	Putative Photosystem I Reaction Center Subunit Ii Precursor	10.5 / 22613.7	4.17E-07	At4g02770
160	Translationally Controlled Tumor Protein-Like Protein	4.3 / 18880.2	3.74E-04	At3g16640
161	40S Ribosomal Protein S5 [	10.3 / 22891.2	1.09E-04	At2g37270
162	Putative Rna-Binding Protein	4.3 / 30745.7	5.35E-11	AT3G26932
163	Gast1-Like Protein	9.2 / 10800.8	2.59E-08	AT1G75750
164	Putative 60S Ribosomal Protein L17	10.5 / 15012.7	6.80E-04	At1g04480
165	Myrosinase-Associated Protein, Putative	8.3 / 42981.8	3.72E-04	AT1G54010
166	Water Stress-Induced Protein, Putative	6 / 10793.7	3.75E-06	AT1G54410
167	Fibrillin Precursor-Like Protein	5.1 / 33664.2	6.48E-05	At4g22240
168	Sedoheptulose-Bisphosphatase Precursor	6 / 42442.2	9.58E-07	At3g55800
169	P-Protein-Like Protein	6 / 112954.6	2.72E-04	At4g33010
170	Phosphoserine Aminotransferase	8.33 / 47387.1	8.87E-04	At4g35630
171	Expressed Protein	9.2 / 10730.6	2.59E-08	At1g75750
172	P-Protein — Like Protein	6 / 104012.5	2.72E-04	At4g33010
173	At3G60750/T4C21_160	6 / 50299.2	2.97E-07	At3G60750
174	At2G39730/T5I7.3	4.3 / 52004.9	2.01E-06	At2G39730
175	Putative Ferredoxin—Nitrite Reductase	6 / 65535.6	5.60E-04	At2g15620
176	Putative Ferredoxin—Nitrite Reductase	6 / 65405.4	5.60E-04	At2g15620
177	Putative Chloroplast Translation Elongation Factor Ef-Tu Precursor	6 / 51655.9	3.71E-07	At4g20360
178	Putative 5-Methyltetrahydropteroyltriglutamate	6 / 84483.2	9.78E-05	At5g17920
179	Putative Elongation Factor	7.2 / 74232.5	5.28E-07	At1g56075
180	Unknown Protein	4.3 / 48256.6	1.41E-05	At5g08670
181	Putative Photosystem I Subunit Psi-E Protein	10.5 / 14764.6	2.60E-04	At4g28750
182	Unknown Protein	4.3 / 18350.7	9.13E-05	AT4G20260
183	Unknown Protein	4.3 / 25404.7	9.54E-04	At5g26000
184	Unknown Protein	7.9 / 10779.7	3.75E-06	AT1G54410
185	Atpase Alpha Subunit	6.3 / 12283.9	9.32E-06	ATCG00120

Table 3.2. Continued.

S.No.	Protein Identification	pI / MW	P (pro)	Gene/locus
186	Ribulose 1,5-Bisphosphate Carboxylase/Oxygenase Large Chain	6 / 47493.6	3.42E-08	ATCG00490
187	Ribulose 1,5-Bisphosphate Carboxylase/Oxygenase Large Chain	6 / 47463.6	3.68E-08	ATCG00490
188	Putative Elongation Factor 1B Alpha-Subunit	4.3 / 22758	2.48E-04	At5g12110
189	Unknown Protein	6 / 79953.8	2.97E-07	AT3G60750
190	Putative Glutamine Synthetase	5.125 / 38652.3	3.48E-05	At3g17820
191	At5G49910	10.1 / 40538.5	1.59E-06	At5g49910
192	Putative Cytosolic Malate Dehydrogenase	7.2 / 13155.6	5.18E-06	At1g04410
193	Chloroplast Polyprotein Of Elongation Factor Ts Precursor	4.25 / 76926.7	2.08E-04	AT4G29060
194	At3G17820	6 / 38594.3	3.48E-05	At3g17820
195	T4G22240	4.3 / 33655.1	6.48E-05	At4g22240
196	Actin 8	5.1 / 38590.9	1.29E-04	AT1G49240
197	Cp33	4.3 / 35056.2	6.13E-05	AT3G52380
198	Luminal Binding Protein (Bip)	4.25 / 73491.7	5.68E-06	At5g42020
199	Monodehydroascorbate Reductase	8.3 / 53275.4	5.15E-07	At1g63940
200	Delta Subunit Of Mitochondrial F1-Atpase	10.5 / 26206.3	1.62E-05	At5g13450
201	Luminal Binding Protein	4.25 / 73621.8	5.68E-06	At5g28540
202	Luminal Binding Protein	4.25 / 73590.8	5.68E-06	At5g42020
203	Ribulosebisphosphate Carboxylase	6 / 47638.7	3.42E-08	AtCg00490
204	Beta Subunit Of Coupling Factor One	10.5 / 8916.3	9.05E-08	ATCG00480
205	Larger Subunit Of Rubisco	7.2 / 5455.1	6.83E-04	AtCg00490
206	Atclpc	6 / 105738.7	1.34E-05	AT3G48870
207	Atpase Alpha Subunit	4.3 / 55327.9	3.02E-06	AtCg00120
208	Atpase Beta Subunit	4.3 / 53933.5	3.71E-08	AtCg00480
209	Large Subunit Of Riblose-1,5-Bisphosphate Carboxylase/Oxygenase	6 / 52954.7	3.42E-08	AtCg00490
210	Cytochrome F	8.7 / 35356.5	2.53E-04	AtCg00540
211	Dihydrolipoamide S-Acetyltransferase	6 / 45075.5	2.82E-04	AT1G34430
212	Trigger Factor-Like Protein	4.3 / 65160.1	7.74E-05	
213	Shepherd	4.3 / 94190.2	2.60E-07	AT4G24190
214	Putative 3-Isopropylmalate Dehydrogenase	4.3 / 34813.7	2.04E-08	At1g31180
215	Carboxyltransferase Alpha Subunit	4.3 / 85287.5	2.38E-05	At2g38040
216	Endomembrane-Associated Protein	4.3 / 14494.1	1.41E-07	At4g20260
217	Elongation Factor G	4.3 / 44945	5.60E-05	At1g62750
218	Carbonic Anhydrase, Chloroplast Precursor	4.3 / 28184.2	1.01E-10	At3g01500
219	Dead Box Rna Helicase Rh15 - Like Protein	4.3 / 23768.9	2.27E-05	At5g11170
220	Glyceraldehyde 3-Phosphate Dehydrogenase A Subunit	4.3 / 17627	5.22E-07	At3g26650
221	Thioglucosidase 3D Precursor	10.1 / 24321.4	5.67E-04	At3g09260
222	Formate Dehydrogenase	4.3 / 17212.6	9.33E-07	At5g14780
223	Vacuolar-Type H+-Atpase Subunit A	4.3 / 36802.5	1.87E-04	At1g78900
224	Nadp Specific Isocitrate Dehydrogenase Like Protein	4.3 / 9767.9	1.37E-05	At1g65930
225	Adenosine Kinase Like Protein	5.125 / 19759.5	5.00E-06	At3g09820
226	Adpg Pyrophosphorylase Small Subunit	4.3 / 24801.2	1.36E-05	At5g48300
227	Phosphoglycerate Dehydrogenase - Like Protein	4.3 / 27442.5	6.66E-06	At4g34200
228	Hypothetical Protein	4.3 / 40159	5.28E-07	At1g56070
229	Adpg Pyrophosphorylase Small Subunit	5.1 / 13699.6	1.36E-05	At5g48300
230	Hypothetical Protein	6 / 60067.2	5.28E-07	At1g56070
231	Glutamate-Ammonia Ligase Precursor	5.1 / 19798.2	1.09E-06	At5g35630
232	Putative Leucine Aminopeptidase	4.3 / 22236.4	6.13E-05	At2g24200

Table 3.2. Continued.

S.No.	Protein Identification	pI / MW	P (pro)	Gene/locus
233	Hsp90-Like Protein	4.3 / 36976	2.60E-07	At4g24190
234	Fructose Bisphosphate Aldolase Like Protein	4.3 / 5251.8	1.92E-05	At2g21330
235	Putative Cytosolic Factor Protein	10.1 / 25598.1	1.02E-04	At1g72150
236	Enolase	4.3 / 28219.6	2.05E-04	At2g36530
237	Putative Glyceraldehyde-3-Phosphate Dehydrogenase	9.9 / 17308.6	1.11E-07	At1g13440
238	Heat-Shock Protein	4.3 / 17037.8	2.81E-04	At3g09440
239	Putative Rubisco Subunit Binding-Protein Alpha Subunit	4.3 / 35384.7	1.44E-06	At2g28000
240	Legumin-Like Protein	10.3 / 13073.6	7.53E-04	At5g44120
241	Rna Helicase	6 / 51136.3	2.27E-05	At5g11170
242	2-Oxoglutarate Dehydrogenase, E3 Subunit	6 / 50018.2	1.19E-10	At3g17240
243	Chlorophyll A/B Binding Protein (Lhcp Ab 180)	4.3 / 24994.2	1.24E-04	At1g29930
244	Atpb Gene Product (1 Is 3Rd Base In Codon)	6.6 / 1519.8	4.05E-04	AtCg00480
245	Ribulose Bisphosphate Carboxylase	8.3 / 20461.3	7.53E-06	At1g67090
246	Unnamed Protein Product	5.1 / 51770.6	3.53E-08	At2g39730
247	12S Seed Storage Protein	8.3 / 52623.7	5.68E-07	At5g44120
248	Ribulose Bisphosphate Carboxylase	8.3 / 20316.2	1.20E-09	At5g38420
249	Ribulose Bisphosphate Carboxylase	8.3 / 20314.3	1.20E-09	At5g38410
250	33 Kda Oxygen-Evolving Protein	5.1 / 35135	4.28E-06	At5g66570
251	Nitrilase I	6 / 38178	6.30E-07	At3g44310
252	Catalase	6 / 56884.9	9.62E-06	At4g35090
253	Heat Shock Protein 70 Cognate	4.3 / 71371.3	1.60E-06	At5g02500
254	Heat Shock Cognate 70-1	4.3 / 70070.9	1.60E-06	At5g02500
255	Heat Shock Cognate 70-2	9.005 / 10727	7.44E-04	At5g02490
256	Cysteine Synthase	4.3 / 33862.9	1.80E-04	At4g14880
257	60S Ribosomal Protein L2	11.505 / 27739.7	4.54E-05	At2g18020
258	Mitochondrial Elongation Factor Tu	4.3 / 51384.4	3.40E-04	At4g02930
259	Thioglucoside Glucohydrolase	6 / 59746.3	3.08E-08	AT3G09260
260	Germin1	10.3 / 21618	1.11E-04	At1g72610
261	60S Ribosomal Protein L9	10.3 / 22357.9	1.42E-04	At1g33120
262	V-Type Proton-Atpase	6 / 26069.9	1.68E-06	At4g11150
263	Catalase	6 / 56916.9	9.62E-06	At4g35090
264	Gst6	6 / 23704.2	1.25E-05	At2g47730
265	Glyceraldehyde-3-Phosphate Dehydrogenase (Nadp+) (Phosphorylating)	7.2 / 42507.3	5.22E-07	At3g26650
266	Hsc70-G7 Protein	4.3 / 10779.9	2.81E-04	AT5G49910
267	Plastid Ribosomal Protein	10.5 / 22004.6	2.69E-05	At5g54600
268	G5Bf	8.3 / 42679.6	3.03E-10	AT1G09340
269	Heat Shock Protein	4.25 / 79992.5	4.45E-04	At5g56010
270	Heat Shock Protein	4.3 / 87486.2	1.05E-05	AT2G20550
271	Chloroplast Nad-Mdh	8.3 / 42421.2	4.56E-08	At3g47520
272	Ribosomal Protein L4	10.1 / 30524.4	1.39E-05	At1g07320
273	Putative Rna Binding Protein	8.3 / 42157.8	3.03E-10	AT1G09340
274	Psi 9Kd Protein	7.165 / 9038.5	3.28E-04	ATCG01060
275	Chaperonin Hsp60	4.3 / 61350.3	1.08E-05	At3g23990
276	16 Kda Polypeptide Of Oxygen-Evolving Complex	10.5 / 23005	1.81E-07	At4g21280
277	Atp Synthase Gamma Chain, Chloroplast Precursor	4.3 / 33324.3	7.05E-06	AT4G04640
278	Ferredoxin-Nadp+ Reductase	8.3 / 40164.1	8.04E-04	AT1G65960

Table 3.2. Continued.

S.No.	Protein Identification	pI / MW	P (pro)	Gene/locus
279	Photosystem I Subunit Ii Precursor	10.5 / 22348.5	4.17E-07	At1g03130
280	Putative Transposon Protein	6 / 52770.9	8.77E-04	AT4g08020
281	Leucyl Aminopeptidase-Like Protein	6 / 55058.8	6.13E-05	At4g30920
282	Putative Fructose-Bisphosphate Aldolase	8.7 / 37143.2	3.28E-08	AT4g38970
283	Transketolase-Like Protein	6 / 81474.5	2.97E-07	AT3G60750
284	Putative Protein	6 / 41597.1	1.77E-04	At5g02240
285	Putative Protein	10.1 / 32129	8.78E-04	At3g63190
286	Putative Protein	4.25 / 60127.4	1.73E-04	AT5G14260
287	Actin (Act3)	4.3 / 47027.5	1.29E-04	At2g37620
288	40S Ribsomomal Protein	10.5 / 13115.3	4.00E-07	At5g62300
289	Rieske Fes Protein	8.3 / 24349.9	3.34E-07	At4g03280
290	Mitochondrial F1 Atp Synthase Beta Subunit	6 / 63370.6	1.41E-05	At5g08670
291	Coproporphyrinogen Iii Oxidase	6 / 43795.9	2.01E-05	At1g03475
292	Isocitrate Dehydrogenase	6 / 45725.9	1.37E-05	AT4G35260
293	Putative O-Acetylserine Thiol Lyase	7.2 / 32033.9	1.80E-04	AT4G14880
294	Cinnamoyl Alcohol Dehydrogenase	4.3 / 33430.6	4.09E-05	AT1G09500
295	Cinnamoyl Alcohol Dehydrogenase	4.3 / 33429.6	4.09E-05	AT1G09500
296	Cinnamoyl Alcohol Dehydrogenase	4.3 / 34234.5	4.09E-05	AT1G09500
297	Cobalamin-Independent Methionine Synthase	6 / 84314	1.54E-07	AT5G17920
298	Thioglucoside Glucohydrolase	6 / 53281.6	5.04E-05	AT5G26000
299	Thioglucoside Glucohydrolase	6 / 52767	5.04E-05	AT5G26000
300	Thioglucoside Glucohydrolase	5.1 / 54377.6	5.04E-05	AT5G26000
301	Atpase Alpha Subunit	4.3 / 19681.3	6.63E-05	ATCG00120
302	Atpase Alpha Subunit	4.3 / 19768.4	6.63E-05	ATCG00120
303	Atpase Beta Subunit	9.8 / 16687.2	9.05E-08	ATCG00480
304	Chaperonin 60 Beta Precursor	6 / 63830.8	5.91E-05	AT1G33740
305	Ca1 (Carbonic Anhydrase 1); Carbonate Dehydratase/ Zinc Ion Binding	6 / 31347.7	1.01E-10	At3g01500
306	Structural Constituent Of Ribosome	10.5 / 23593.9	2.31E-07	At3g07110
307	Atplc2 (Phospholipase C 2); Phospholipase C	6 / 66121.9	4.80E-05	At3g08510
308	Msd1 (Manganese Superoxide Dismutase 1); Manganese Superoxide Dismutase	9.435 / 25344.6	5.10E-04	At3g10920
309	Glycolate Oxidase/ Oxidoreductase	10.1 / 38166.7	2.38E-05	AT3G14420
310	Structural Constituent Of Ribosome	12.48 / 12676	4.33E-04	At3g53740
311	Rpl4; Structural Constituent Of Ribosome	10.3 / 30142	1.39E-05	At1g07320
312	Apx1; L-Ascorbate Peroxidase	8.3 / 21175.9	2.60E-11	At1g07890
313	Apx1; L-Ascorbate Peroxidase	6 / 27520	2.60E-11	At1g07890
314	Cat3 (Catalase 3); Catalase	8.3 / 57019.4	4.42E-05	At1g20620
315	Cat3 (Catalase 3); Catalase	7.2 / 55967.2	4.42E-05	At1g20620
316	Ggt1 (Alanine-2-Oxoglutarate Aminotransferase 1)	8.3 / 48549.6	4.64E-04	At1g23310
317	Aoat2 (Alanine-2-Oxoglutarate Aminotransferase 2); Alanine Transami	6 / 53444.1	4.64E-04	At1g70580
318	Vha-A; Atp Binding / Hydrogen-Transporting Atp Synthase, Rotational	4.3 / 68812.1	1.87E-04	At1g78900
319	Fructose-Bisphosphate Aldolase	6 / 33323	1.64E-08	At2g21330
320	Fructose-Bisphosphate Aldolase	6 / 41807.4	1.64E-08	AT2G21330
321	Unknown Protein	4.3 / 24711.8	1.41E-07	At4g20260
322	Calcium Ion Binding	10.5 / 23866	1.61E-07	At4g21280
323	Emb2726; Rna Binding / Structural Constituent Of Ribosome	4.25 / 76855.6	2.08E-04	AT4G29060

Table 3.2. Continued.

S.No.	Protein Identification	pI / MW	P (pro)	Gene/locus
324	Cad5 (Cinnamyl Alcohol Dehydrogenase 5)	4.3 / 38716.4	2.61E-05	At4g34230
325	Cat2 (Catalase 2); Catalase	6 / 54994.7	9.62E-06	At4g35090
326	Oxidoreductase/ Zinc Ion Binding	7.2 / 33253.1	1.79E-06	AT4G39330
327	Pmdh2	7.2 / 34975.3	2.15E-05	At5g09660
328	Unknown Protein	10.5 / 21083.3	1.62E-05	At5g13450
329	3-Isopropylmalate Dehydrogenase/ Oxidoreductase	5.1 / 41891.9	1.97E-04	At5g14200
330	Ca2 (Carbonic Anhydrase 2); Carbonate Dehydratase/ Zinc Ion Binding	8.3 / 36527.7	2.78E-10	At5g14740
331	Ca2 (Carbonic Anhydrase 2); Carbonate Dehydratase/ Zinc Ion Binding	8.3 / 36485.6	2.78E-10	At5g14740
332	Homoserine Dehydrogenase	6 / 40669.2	5.56E-04	At5g21060
333	Gs2 (Glutamine Synthetase 2); Glutamate-Ammonia Ligase	6 / 47410.3	1.09E-06	At5g35630
334	Ribulose-Bisphosphate Carboxylase	8.3 / 19395.3	1.20E-09	At5g38410
335	Kas I (3-Ketoacyl-Acyl Carrier Protein Synthase I)	8.3 / 44729.5	1.54E-06	At5g46290
336	Atp Binding / Protein Binding	4.3 / 63324.1	5.91E-05	AT5G56500
337	Calmodulin Binding / Translation Elongation Factor	8.3 / 44355.8	3.60E-11	AT5G60390
338	Psi-N; Calmodulin Binding	10.1 / 16553.9	2.17E-06	At5g64040
339	Hypothetical Protein Atmt_P24	9.6 / 12664.4	5.75E-08	AtMg00280
340	Unknown Protein	10.5 / 22306.5	4.17E-07	At1g03130
341	Lin2 (Lesion Initiation 2); Coproporphyrinogen Oxidase	8.3 / 40761.7	2.01E-05	AT1G03475
342	Malate Dehydrogenase/ Oxidoreductase	6 / 35570.7	1.57E-07	At1g04410
343	Tua4	4.3 / 49540.6	8.38E-09	At1g50010
344	Psbp (Oxygen-Evolving Enhancer Protein 2); Calcium Ion Binding	7.2 / 28095	5.20E-05	At1g06680
345	Unknown Protein	8.3 / 42619.5	3.03E-10	At1g09340
346	Aminomethyltransferase	8.3 / 44444.5	4.15E-07	At1g11860
347	Nad Binding / Glyceraldehyde-3-Phosphate Dehydrogenase	8.3 / 42846.5	5.22E-07	At1g12900
348	Dna Binding	12.5 / 15357.8	7.82E-07	At1g13370
349	Nad Binding / Glyceraldehyde-3-Phosphate Dehydrogenase	6 / 36913	2.78E-07	At1g13440
350	Glutathione Dehydrogenase (Ascorbate)	6 / 17061.4	9.40E-07	AT1G19550
351	Glutathione Dehydrogenase (Ascorbate)	6 / 23640.8	9.40E-07	At1g19570
352	Oxidoreductase	8.3 / 41167.9	4.80E-05	At1g20020
353	Atpdil1-1; Electron Transporter/ Isomerase	4.25 / 55601.2	4.83E-05	At1g21750
354	Transporter	4.25 / 76007.5	2.47E-04	At1g22530
355	Cab1 (Chlorophyll A/B Binding Protein 1); Chlorophyll Binding	4.3 / 28240.9	1.24E-04	At1g29910
356	3-Isopropylmalate Dehydrogenase/ Oxidoreductase	4.3 / 43847.1	2.04E-08	AT1G31180
357	Atp Binding / Kinase/ Phosphoribulokinase/ Uridine Kinase	6 / 44463.3	8.80E-08	At1g32060
358	Glycine Dehydrogenase (Decarboxylating)	4.3 / 17897	1.57E-06	At1g32470
359	Gapb (Glyceraldehyde-3-Phosphate Dehydrogenase B Subunit)	6 / 47659.2	3.94E-12	At1g42970
360	Dihydrolipoyl Dehydrogenase	6 / 53987.7	1.19E-10	At1g48030
361	Icdh (Isocitrate Dehydrogenase); Isocitrate Dehydrogenase (Nadp+)	8.3 / 47233.8	3.39E-05	At1g54340
362	Unknown Protein	6 / 10795.6	3.75E-06	At1g54410
363	Phosphoglycerate Kinase	9 / 49938.3	1.25E-06	At1g56190
364	Hsp70T-1; Atp Binding	4.3 / 68356.2	1.64E-06	At1g56410
365	Aco2 (Acc Oxidase 2)	4.3 / 36182.8	4.87E-04	At1g62380
366	Vha-E3; Hydrogen-Transporting Atp Synthase	6 / 27084.9	1.18E-04	AT1G64200
367	Isocitrate Dehydrogenase (Nadp+)/ Oxidoreductase	6 / 45745.9	1.37E-05	At1g65930
368	Hydrolase, Hydrolyzing O-Glycosyl Compounds	6 / 59663.3	5.36E-05	At1g66270
369	Ribulose-Bisphosphate Carboxylase	8.3 / 20216	7.53E-06	At1g67090

Table 3.2. Continued.

S.No.	Protein Identification	pI / MW	P (pro)	Gene/locus
370	Carbonate Dehydratase/ Zinc Ion Binding	6 / 30836.1	4.88E-04	At1g70410
371	Patl1 (Patellin 1); Transporter	4.3 / 64045.9	1.02E-04	At1g72150
372	Glp1 (Germin-Like Protein 1); Nutrient Reservoir	10.1 / 21558.9	1.11E-04	At1g72610
373	Dna Binding	11.5 / 15401.9	2.52E-06	At1g75600
374	Erd14 (Early Response To Dehydration 14)	4.3 / 20786.2	4.11E-07	At1g76180
375	Atgstu23; Glutathione Transferase	4.25 / 25680.4	7.07E-05	At1g78320
376	Unknown Protein	10.1 / 14585.6	3.46E-04	At1g79040
377	Pgk (Phosphoglycerate Kinase)	4.3 / 42131.5	4.59E-05	At1g79550
378	3-Isopropylmalate Dehydrogenase/ Oxidoreductase	5.1 / 43370.5	1.97E-04	At1g80560
379	Atgstf3 (Glutathione S-Transferase 16); Glutathione Transferase	6 / 24120.4	4.86E-08	AT2G02930
380	Cr88; Atp Binding / Unfolded Protein Binding	4.3 / 88662.3	1.05E-05	At2g04030
381	Atp Binding / Hydrogen-Transporting Atp Synthase	4.3 / 85932.7	2.48E-05	AtMg01200
382	Nir1; Ferredoxin-Nitrate Reductase	6 / 65504.6	5.60E-04	At2g15620
383	Roc3; Peptidyl-Prolyl Cis-Trans Isomerase	8.3 / 18491.8	4.82E-06	At2g16600
384	Phosphoserine Transaminase/ Transaminase	8.33 / 46633.2	8.87E-04	At2g17630
385	Emb2296; Structural Constituent Of Ribosome	11.505 / 27858.9	4.54E-05	At2g18020
386	Structural Constituent Of Ribosome	10.5 / 15895.4	3.73E-05	At2g19730
387	Unknown Protein	10.5 / 15189.3	8.24E-07	At2g20260
388	Tim (Triosephosphate Isomerase); Triose-Phosphate Isomerase	8.3 / 33345.5	2.69E-05	At2g21170
389	Atgrp7; Rna Binding	5.1 / 16889.7	7.47E-05	At2g21660
390	Pmdh1; Malate Dehydrogenase/ Oxidoreductase	8.3 / 37465.3	2.15E-05	At2g22780
391	Aminopeptidase	5.1 / 54509	6.13E-05	At2g24200
392	Glycine Dehydrogenase (Decarboxylating)	6 / 113775	2.72E-04	At2g26080
393	Cpn60A; Atp Binding / Protein Binding	4.3 / 62071.6	5.40E-08	At2g28000
394	Catalytic	6 / 71986.3	7.86E-05	At2g29630
395	Calcium Ion Binding	10.1 / 28181.5	5.71E-05	AT2G30790
396	Asp1 (Aspartate Aminotransferase 1)	8.3 / 47757.5	3.35E-04	At2g30970
397	Structural Constituent Of Ribosome	10.1 / 32644.9	2.47E-05	At2g33800
398	Gdch	4.3 / 17947.1	6.10E-06	At2g35370
399	Fructose-Bisphosphate Aldolase	7.2 / 38386.5	8.49E-10	At2g36460
400	Los2; Phosphopyruvate Hydratase	4.3 / 47719	8.51E-05	At2g36530
401	Atp Binding / Methionine Adenosyltransferase	6 / 42497.2	2.62E-07	At2g36880
402	Structural Constituent Of Ribosome	10.3 / 17941.7	3.19E-06	At2g37190
403	Rna Binding / Nucleic Acid Binding	4.3 / 30717.7	5.35E-11	At2g37220
404	Atrps5B (Ribosomal Protein 5B); Structural Constituent Of Ribosome	10.5 / 22990.4	1.09E-04	At2g37270
405	Structural Constituent Of Ribosome	12.48 / 12736.3	4.33E-04	At2g37600
406	Lp1; Lipid Binding	8.3 / 11754.9	2.27E-04	At2g38540
407	Atrpl23A (Ribosomal Protein L23A); Rna Binding / Structural Constituen	10.5 / 17440.6	3.64E-05	At2g39460
408	S-Adenosylmethionine-Dependent Methyltransferase	4.3 / 25288.5	2.96E-04	At2g43910
409	Voltage-Gated Ion-Selective Channel	10.1 / 29425.1	1.36E-04	At3g01280
410	Ca1 (Carbonic Anhydrase 1); Carbonate Dehydratase/ Zinc Ion Binding	5.1 / 37449.6	1.01E-10	At3g01500
411	5-Methyltetrahydropteroyltriglutamate-Homocysteine S-Methyltransferase	6 / 84583.2	1.54E-07	At3g03780
412	Gapc (Glyceraldehyde-3-Phosphate Dehydrogenase C Subunit)	6 / 36914	2.78E-07	At3g04120
413	Structural Constituent Of Ribosome	10.5 / 23465.7	2.31E-07	At3g07110

Table 3.2. Continued.

S.No.	Protein Identification	pI / MW	P (pro)	Gene/locus
414	Vha-E2; Hydrogen-Exporting Atpase, Phosphorylative Mechanism	10.07 / 26852.7	2.21E-04	At3g08560
415	Structural Constituent Of Ribosome	4.3 / 34132.8	5.42E-06	At3g09200
416	Pyk10; Hydrolase, Hydrolyzing O-Glycosyl Compounds	6 / 59720.2	3.08E-08	At3g09260
417	Atp Binding	4.3 / 71147.2	1.64E-06	At3g09440
418	Adk1 (Adenosine Kinase 1)	4.25 / 37835.9	5.00E-06	At3g09820
419	Structural Constituent Of Ribosome	10.5 / 7370.5	3.50E-04	At5g03850
420	Glx2-2 (Glyoxalase 2-2); Hydroxyacylglutathione Hydrolase	6 / 28791.5	4.09E-09	At3g10850
421	Msd1 (Manganese Superoxide Dismutase 1); Manganese Superoxide Dismutase	9.435 / 25443.7	5.10E-04	At3g10920
422	Atrps5A (Ribosomal Protein 5A); Structural Constituent Of Ribosome	10.3 / 22921.3	1.09E-04	At3g11940
423	Unknown Protein	3.86 / 21982.1	2.10E-06	At3g12390
424	Hsp70; Atp Binding	4.3 / 71101	1.64E-06	At3g12580
425	Pgk1 (Phosphoglycerate Kinase 1); Phosphoglycerate Kinase	6 / 50111.4	3.04E-11	At3g12780
426	Atp Binding / Protein Binding	4.3 / 63341.2	5.91E-05	At3g13470
427	Carboxylic Ester Hydrolase/ Hydrolase, Acting On Ester Bonds	8.3 / 44060	8.29E-04	At3g14210
428	Atpme3	10.1 / 64255.5	2.88E-07	At3g14310
429	Malate Dehydrogenase/ Oxidoreductase	8.3 / 35875.1	9.57E-06	At3g15020
430	Athm4; Electron Transporter/ Thiol-Disulfide Exchange Intermediate	10.5 / 21172.1	4.19E-04	At3g15360
431	Tctp (Translationally Controlled Tumor Protein)	4.3 / 18910.2	3.74E-04	At3g16640
432	Mto3; Methionine Adenosyltransferase	5.1 / 42795.3	2.62E-07	At3g17390
433	Cad4 (Cinnamyl Alcohol Dehydrogenase 4); Cinnamyl-Alcohol Dehydrogenase	4.3 / 39097.9	4.09E-05	At3g19450
434	Hsp60; Atp Binding / Protein Binding	4.3 / 61280.2	1.08E-05	At3g23990
435	Structural Constituent Of Ribosome	10.5 / 23458.8	2.31E-07	At3g24830
436	Lta2 (Plastid E2 Subunit Of Pyruvate Decarboxylase)	9.9 / 50080	2.82E-04	At3g25860
437	Antioxidant [	10.1 / 23677.8	4.24E-08	At3g26060
438	Cam7 (Calmodulin 7); Calcium Ion Binding A]	3.9 / 16847.6	3.41E-07	At3g43810
439	Structural Constituent Of Ribosome	10.5 / 13878.2	4.00E-07	At3g45030
440	Act12 (Actin-12); Structural Constituent Of Cytoskeleton	4.3 / 41794.6	1.29E-04	At3g46520
441	Unknown Protein	10.5 / 10530	6.10E-06	At3g47070
442	Mdh; Malate Dehydrogenase/ Oxidoreductase	8.6 / 42405.3	4.56E-08	At3g47520
443	Psbo-2/Psbo2; Oxygen Evolving	6 / 35019	7.22E-06	At3g50820
444	Cp33; Rna Binding	4.3 / 35744	6.13E-05	At3g52380
445	Fructose-Bisphosphate Aldolase	6 / 38539.5	8.49E-10	At3g52930
446	Antioxidant	10.1 / 24684.1	6.78E-05	At3g52960
447	Structural Constituent Of Ribosome	10.3 / 17969.7	3.19E-06	At3g53430
448	Rna Binding / Structural Constituent Of Ribosome	10.5 / 17395.6	3.64E-05	At3g55280
449	Atctimc (Cytosolic Triose Phosphate Isomerase)	4.3 / 27168.9	1.74E-06	At3g55440
450	Sbpase; Phosphoric Ester Hydrolase	6 / 42414.1	9.58E-07	At3g55800
451	Cch	4.3 / 12970.6	7.92E-06	At3g56240
452	Atcysc1; L-3-Cyanoalanine Synthase/ Cysteine Synthase	9.9 / 39927	5.87E-05	At3g61440
453	Roc4; Peptidyl-Prolyl Cis-Trans Isomerase	8.3 / 28208	8.10E-06	At3g62030
454	Mrna Binding	9 / 43929.5	3.44E-04	At3g63140
455	Atp Binding / Methionine Adenosyltransferase	6 / 43254.8	2.62E-07	At4g01850
456	Atgstf2; Glutathione Transferase	6 / 24128.4	4.86E-08	At4g02520
457	Unknown Protein	10.5 / 22597.7	4.17E-07	At4g02770
458	Gtp Binding / Translation Elongation Factor	6 / 49409.5	3.40E-04	At4g02930

Table 3.2. Continued.

S.No.	Protein Identification	pI / MW	P (pro)	Gene/locus
459	Petc (Photosynthetic Electron Transfer C)	8.3 / 24365.9	3.34E-07	At4g03280
460	Structural Molecule	4.3 / 34948.2	1.92E-07	At4g04020
461	Calcium Ion Binding	10.5 / 24642.8	1.81E-07	At4g05180
462	Tuf (Tuff)	6 / 26059.8	1.68E-06	At4g11150
463	Structural Constituent Of Ribosome	10.5 / 23628.9	2.31E-07	At4g13170
464	Oasa1	5.1 / 33804.9	1.80E-04	At4g14880
465	Tua6	4.3 / 49537.6	8.38E-09	At4g14960
466	Chli1 (Chlorina 42); Magnesium Chelatase	6 / 46269.6	9.59E-04	At4g18480
467	Atp Binding / Gtp Binding / Translation Elongation Factor	6 / 51629.8	3.71E-07	At4g20360
468	Calcium Ion Binding	10.5 / 23795	1.61E-07	At4g21280
469	Unknown Protein	6 / 17517.7	1.16E-07	AT4g23670
470	Shd (Shepherd); Atp Binding / Unfolded Protein Binding	4.3 / 94203.2	2.60E-07	At4g24190
471	Cphsc70-1; Atp Binding	4.3 / 76507.5	1.59E-06	AT4g24280
472	Acp4	4.3 / 14544.3	1.82E-05	AT4g25050
473	Fructose-Bisphosphate Aldolase	4.3 / 38293.4	8.49E-10	AT4g26530
474	Cru3 (Cruciferin 3); Nutrient Reservoir	6 / 58234.9	1.36E-06	AT4g28520
475	Prf2 (Profilin 2); Actin Binding	4.25 / 13997.8	5.67E-05	At4g29350
476	Aminopeptidase	6 / 61306.8	6.13E-05	At4g30920
477	Glycine Dehydrogenase (Decarboxylating)	6 / 112924.6	2.72E-04	At4g33010
478	Amino Acid Binding / Oxidoreductase	6 / 63324.6	6.66E-06	At4g34200
479	Cad5 (Cinnamyl Alcohol Dehydrogenase 5); Cinnamyl-Alcohol Dehydrogenase	4.3 / 38743.5	2.61E-05	At4g34230
480	Roc5; Peptidyl-Prolyl Cis-Trans Isomerase	8.7 / 18377.9	2.68E-05	At4g34870
481	Cat2 (Catalase 2); Catalase	6 / 56930.9	9.62E-06	At4g35090
482	Psat; Phosphoserine Transaminase/ Transaminase	8.33 / 47359	8.87E-04	AT4G35630
483	Shm1 (Serine Hydroxymethyltransferase 1)	8.3 / 57400.3	8.31E-08	At4g37930
484	Atgrp8 (Glycine-Rich Protein 8); Rna Binding / Nucleic Acid Binding	4.3 / 16578.5	2.47E-04	At4g39260
485	Oxidoreductase/ Zinc Ion Binding	6 / 38933.5	1.79E-06	At4g39330
486	Structural Constituent Of Ribosome	12.48 / 12189.4	4.33E-04	At5g02450
487	Atp Binding	4.3 / 71386.3	1.64E-06	At5g02490
488	Hsc70-1; Atp Binding	4.3 / 71357.3	1.64E-06	At5g02500
489	Adk2 (Adenosine Kinase 2); Kinase	4.25 / 37845.9	5.00E-06	At5g03300
490	Nucleic Acid Binding	6 / 33820.5	1.48E-05	At5g04430
491	Rna Binding / Structural Constituent Of Ribosome	10.53 / 16948	7.39E-04	At5g08180
492	Inorganic Diphosphatase/ Magnesium Ion Binding / Pyrophosphatase	5.1 / 33379.8	6.81E-08	At5g09650
493	Pmdh2; Malate Dehydrogenase	8.3 / 37369	2.15E-05	At5g09660
494	Translation Elongation Factor	4.3 / 24788.3	2.48E-04	AT5G12110
495	Hydrogen-Transporting Atp Synthase, Rotational Mechanism	10.3 / 26321.4	1.62E-05	At5g13450
496	Structural Constituent Of Ribosome	10.1 / 24735.6	4.66E-07	At5g13510
497	Fdh (Formate Dehydrogenase); Oxidoreductase, Acting On The Ch-Oh Group	6 / 42409.3	9.33E-07	At5g14780
498	Atcims (Cobalamin-Independent Methionine Synthase)	6 / 84356	1.54E-07	At5g17920
499	Tua3	4.3 / 49653.8	8.38E-09	At5g19770
500	Cpn20 (Chaperonin 20); Calmodulin Binding	10.1 / 26801.7	1.69E-05	At5g20720
501	5-Methyltetrahydropteroyltriglutamate-Homocysteine S-Methyltransferase	8.3 / 90593.5	1.54E-07	At5g20980
502	Unknown Protein	10.5 / 41285.2	1.88E-04	At5g23060
503	Hcf136	7.2 / 44103.2	4.40E-04	At5g23120

Table 3.2. Continued.

S.No.	Protein Identification	pI / MW	P (pro)	Gene/locus
504	Gdp-Mannose 3,5-Epimerase/ Catalytic	6 / 42758.2	5.66E-07	At5g28840
505	Cam1 (Calmodulin 1); Calcium Ion Binding	3.9 / 16861.7	3.41E-07	At5g37780
506	Ribulose-Bisphosphate Carboxylase	8.3 / 20284.2	1.20E-09	At5g38410
507	Ribulose-Bisphosphate Carboxylase	8.3 / 20350.2	1.20E-09	At5g38420
508	Ribulose-Bisphosphate Carboxylase	8.3 / 20286.2	6.39E-05	At5g38430
509	Atphb3 (Prohibitin 3)	6 / 30399.6	6.70E-04	At5g40770
510	Bip; Atp Binding	4.25 / 67399.9	5.68E-06	At5g42020
511	Malate Dehydrogenase/ Oxidoreductase	6 / 35674.9	6.90E-05	At5g43330
512	Cra1 (Cruciferina); Nutrient Reservoir	8.3 / 52594.7	5.68E-07	At5g44120
513	Unknown Protein	4.3 / 132581	2.45E-04	At5g45520
514	Kas I (3-Ketoacyl-Acyl Carrier Protein Synthase I); Catalytic	8.3 / 50413	1.54E-06	At5g46290
515	Rna Binding	10.07 / 37999.2	1.14E-04	At5g47210
516	Adg1 (Adp Glucose Pyrophosphorylase Small Subunit 1)	6 / 56650.3	1.36E-05	At5g48300
517	Structural Constituent Of Ribosome	10.5 / 23590	2.31E-07	At5g48760
518	Cphsc70-2 (Heat Shock Protein 70-7); Atp Binding	4.3 / 76996	1.59E-06	At5g49910
519	Unknown Protein	4.25 / 25573	7.40E-05	At5g53490
520	Lhcb3 (Light-Harvesting Chlorophyll Binding Protein 3)	4.25 / 28706.4	5.73E-04	At5g54270
521	Structural Constituent Of Ribosome	10.5 / 19753	2.69E-05	At5g54600
522	Thi1 (Thiazole Requiring)	6 / 36664	3.15E-05	At5g54770
523	Peptidyl-Prolyl Cis-Trans Isomerase	4.3 / 61733.1	7.74E-05	At5g55220
524	Hsp81-3; Atp Binding / Unfolded Protein Binding	4.25 / 80051.5	4.45E-04	At5g56010
525	Atp Binding / Protein Binding	4.3 / 63241	5.91E-05	AT5G56500
526	Structural Constituent Of Ribosome	10.5 / 17842.6	3.19E-06	At5g60670
527	Psi-N; Calmodulin Binding	9.7 / 18429	2.51E-07	At5g64040
528	Dna Binding	12.5 / 15591.1	7.82E-07	At5g65350
529	Oxidoreductase	8.3 / 40326.2	8.04E-04	At5g66190
530	Psbo-1 (Oxygen-Evolving Enhancer 33)	4.3 / 35142.1	4.28E-06	At5g66570
531	Voltage-Gated Ion-Selective Channel	10.1 / 29594.4	4.26E-05	At5g67500
532	Unknown Protein	10.5 / 18834.5	8.53E-10	At1g03600
533	Structural Constituent Of Ribosome	10.5 / 15026.7	6.80E-04	At1g04480
534	Rpl4; Structural Constituent Of Ribosome	10.1 / 30558.4	1.39E-05	At1g07320
535	Dna Binding	12.5 / 11409.3	4.48E-07	At1g07820
536	Calmodulin Binding / Translation Elongation Factor	10.1 / 49502	3.60E-11	At1g07940
537	Structural Constituent Of Ribosome	10.5 / 18652.6	2.28E-06	At1g09690
538	Unknown Protein	8.3 / 30064.5	7.56E-04	At1g12250
539	Unknown Protein	4.3 / 16163.5	3.22E-09	At1g13930
540	Unknown Protein	6 / 17859.5	1.92E-05	At1g23130
541	Ggt1 (Alanine-2-Oxoglutarate Aminotransferase 1)	6 / 53300.9	4.64E-04	At1g23310
542	Cab3 (Chlorophyll A/B Binding Protein 3); Chlorophyll Binding	4.3 / 28226.9	1.24E-04	At1g29910
543	Structural Constituent Of Ribosome	10.5 / 22017.5	1.42E-04	At1g33140
544	Grf4; Protein Phosphorylated Amino Acid Binding	4.3 / 30193.7	7.40E-07	At1g35160
545	Unknown Protein	10.5 / 24160.8	1.55E-04	At1g51100
546	Malate Dehydrogenase	8.3 / 35804.2	8.07E-10	At1g53240
547	Carboxylic Ester Hydrolase/ Hydrolase, Acting On Ester Bonds	8.3 / 43143	3.72E-04	At1g54010
548	Structural Constituent Of Ribosome	10.5 / 18708.7	2.28E-06	At1g57860
549	Unknown Protein	4.25 / 16628.1	3.40E-06	At1g62480
550	Gtp Binding / Translation Elongation Factor/ Translation Factor	4.3 / 86057	5.60E-05	At1g62750

Table 3.2. Continued.

S.No.	Protein Identification	pI / MW	P (pro)	Gene/locus
551	Disulfide Oxidoreductase/ Oxidoreductase	7.2 / 52501.5	5.15E-07	At1g63940
552	Gasa1	9.2 / 10744.7	2.59E-08	At1g75750
553	Grf2; Protein Phosphorylated Amino Acid Binding	4.3 / 29161.5	9.91E-09	At1g78300
554	Atgstu19 (Glutathione Transferase 8); Glutathione Transferase	6 / 25650.4	8.46E-06	At1g78380
555	Unknown Protein	6 / 21416.2	6.84E-06	At2g04039
556	Thf1	10.1 / 33795.4	6.76E-04	At2g20890
557	Fructose-Bisphosphate Aldolase	6 / 42930.6	1.64E-08	At2g21330
558	Lhb1B2; Chlorophyll Binding	4.3 / 28053.7	1.24E-04	At2g34420
559	Lhb1B1; Chlorophyll Binding	4.3 / 28169.8	3.05E-04	At2g34430
560	Unknown Protein	10.5 / 46696.3	4.59E-04	At2g35880
561	Unknown Protein	8.7 / 34879.5	2.18E-05	At2g37660
562	Cac3; Acetyl-Coa Carboxylase	4.3 / 85305.6	2.38E-05	At2g38040
563	Rca (Rubisco Activase)	5.1 / 51980.9	3.53E-08	At2g39730
564	Chl-Cpn10	9 / 15049.4	1.05E-04	At2g44650
565	Unknown Protein	8.3 / 23778.6	8.38E-09	At2g44920
566	Transketolase	6 / 79921.8	1.03E-04	AT2G45290
567	Gapa (Glyceraldehyde 3-Phosphate Dehydrogenase A Subunit)	7.2 / 42489.2	5.22E-07	At3g26650
568	Unknown Protein	4.3 / 13757.6	1.19E-04	At3g29170
569	Unknown Protein	10.1 / 54357.9	1.27E-04	At3g46780
570	Atclpc; Atp Binding / Atpase/ Dna Binding / Nuclease	6 / 105770.8	1.34E-05	At3g48870
571	Structural Constituent Of Ribosome	12.48 / 11679.8	4.33E-04	At3g53740
572	Unknown Protein	6 / 17793.4	6.30E-06	At3g53990
573	Transketolase	6 / 79967.8	2.97E-07	At3g60750
574	Unknown Protein	10.5 / 30422	8.78E-04	At3g63190
575	Cysteine-Type Endopeptidase/ Cysteine-Type Peptidase	6 / 39417.4	1.20E-04	At4g01620
576	Atpc1	8.3 / 40910.8	7.05E-06	At4g04640
577	Grf1; Protein Phosphorylated Amino Acid Binding	4.3 / 29931.4	7.40E-07	At4g09000
578	Unknown Protein	8.3 / 21967.8	1.26E-07	At4g21860
579	Psae1 (Psa E1 Knockout)	10.5 / 14966.9	8.24E-07	At4g28750
580	Emb2726; Rna Binding / Structural Constituent Of Ribosome	4.25 / 103781.4	2.08E-04	At4g29060
581	Fructose-Bisphosphate Aldolase	7.2 / 42987.5	3.28E-08	At4g38970
582	Unknown Protein	6 / 27102.7	1.77E-04	At5g02240
583	Fructose-Bisphosphate Aldolase	10.1 / 42905.7	8.49E-10	AT5G03690
584	Atp Binding / Hydrogen-Exporting Atpase, Phosphorylative Mechanism	6 / 59670.6	1.41E-05	At5g08670
585	Atp Binding / Hydrogen-Exporting Atpase, Phosphorylative Mechanism	6 / 59712.6	1.41E-05	At5g08690
586	Atp Binding / Atp-Dependent Helicase/ Helicase/ Nucleic Acid Binding	4.3 / 48324.2	2.27E-05	At5g11170
587	Atp Binding / Atp-Dependent Helicase/ Helicase/ Nucleic Acid Binding	4.3 / 48337.2	2.27E-05	At5g11170
588	Ca2 (Carbonic Anhydrase 2); Carbonate Dehydratase/ Zinc Ion Binding	7.2 / 36614.7	2.78E-10	At5g14740
589	Translation Elongation Factor	4.3 / 24200.9	2.48E-04	At5g19510
590	Unknown Protein	10.5 / 7755.9	1.15E-04	At5g24165
591	Clpc; Atp Binding / Atpase	6 / 103451.8	1.34E-05	At5g50920
592	Atp Binding / Hydrogen-Exporting Atpase, Phosphorylative Mechanism	6 / 59858.7	1.41E-05	At5g08680
593	Cysteine-Type Endopeptidase/ Cysteine-Type Peptidase	6 / 39344.4	1.20E-04	At4g01610
594	Petc (Photosynthetic Electron Transfer C)	8.3 / 22532.9	3.34E-07	At4g03280

Table 3.2. Continued.

S.No.	Protein Identification	pI / MW	P (pro)	Gene/locus
595	Tua6	8.3 / 47234.9	8.38E-09	At4g14960
596	Unknown Protein	4.3 / 24583.6	1.41E-07	AT4g20260
597	Cru3 (Cruciferin 3); Nutrient Reservoir	6 / 43723.3	1.36E-06	AT4G28520
598	Akr2 (Ankyrin Repeat-Containing Protein 2); Protein Binding	4.25 / 36984.1	6.75E-05	At4g35450
599	Akr2 (Ankyrin Repeat-Containing Protein 2); Protein Binding	4.25 / 33073.9	6.75E-05	At4g35450
600	Atgrp8 (Glycine-Rich Protein 8); Rna Binding	4.3 / 12806.7	2.47E-04	At4g39260
601	Atgrp8 (Glycine-Rich Protein 8); Rna Binding	4.3 / 10230.2	2.47E-04	At4g39260
602	Atgrp8 (Glycine-Rich Protein 8); Rna Binding	4.3 / 10863.6	2.47E-04	At4g39260
603	Sam1; Methionine Adenosyltransferase	5.1 / 43157.8	2.62E-07	At1g02500
604	Rpl4; Structural Constituent Of Ribosome	10.1 / 30414.3	1.39E-05	At1g07320
605	Apx1; L-Ascorbate Peroxidase	6 / 27561	2.60E-11	At1g07890
606	Atpdil1-1; Electron Transporter/ Isomerase	4.25 / 54158.8	4.83E-05	At1g21750
607	Cpn60B (Chaperonin 60 Beta); Atp Binding	6 / 63808.7	5.91E-05	At1g55490
608	Los1; Gtp Binding / Translation Elongation Factor	6 / 93890.7	5.28E-07	At1g56075
609	Disulfide Oxidoreductase/ Oxidoreductase	8.3 / 53301.5	5.15E-07	At1g63940
610	Disulfide Oxidoreductase/ Oxidoreductase	9 / 45023	5.15E-07	At1g63940
611	Disulfide Oxidoreductase/ Oxidoreductase	8.3 / 52115.1	5.15E-07	At1g63940
612	Hydrolase, Hydrolyzing O-Glycosyl Compounds	6 / 59493.1	5.36E-05	At1g66270
613	Carbonate Dehydratase/ Zinc Ion Binding	6 / 28412.2	4.88E-04	At1g70410
614	Cr88; Atp Binding / Unfolded Protein Binding	4.3 / 88255.9	1.05E-05	At2g04030
615	Unknown Protein	6 / 18334.9	6.84E-06	At2g04039
616	N-Acetyl-Gamma-Glutamyl-Phosphate Reductase/ Nad Binding	6 / 39600.3	2.03E-05	At2g19940
617	Atgrp7; Rna Binding	4.3 / 15548.4	7.47E-05	At2g21660
618	Cam2 (Calmodulin-2); Calcium Ion Binding	3.9 / 12868.3	1.30E-05	AT2G27030
619	Cam2 (Calmodulin-2); Calcium Ion Binding	4.3 / 20575.9	3.41E-07	At2g27030
620	Atp Binding / Protein Binding	6 / 61978.1	1.08E-05	At2g33210
621	Lhb1B2	4.3 / 26585.9	1.24E-04	At2g34420
622	Rca (Rubisco Activase)	7.2 / 49099.8	3.53E-08	At2g39730
623	Rca (Rubisco Activase)	8.3 / 48500.2	3.53E-08	At2g39730
624	Atcal4; Calcium Ion Binding	3.9 / 16819.6	3.41E-07	At2g41110
625	Atgstf8 (Glutathione S-Transferase 8); Glutathione Transferase	9.6 / 29231.4	1.25E-05	At2g27720
626	Ca1 (Carbonic Anhydrase 1); Carbonate Dehydratase/ Zinc Ion Binding	4.3 / 29503.8	1.01E-10	At3g01500
627	Ca1 (Carbonic Anhydrase 1); Carbonate Dehydratase/ Zinc Ion Binding	4.3 / 36144	1.01E-10	At3g01500
628	Glycolate Oxidase/ Oxidoreductase	10.1 / 40341.2	2.38E-05	At3g14420
629	Glycolate Oxidase/ Oxidoreductase	10.1 / 40316.3	2.38E-05	At3g14420
630	Unknown Protein	5.1 / 32157.8	7.73E-05	At3g16420
631	Unknown Protein	6 / 32233.1	1.48E-04	At3g16430
632	Structural Constituent Of Ribosome	10.5 / 13693	4.00E-07	At3g47370
633	Fructose-Bisphosphate Aldolase	8.3 / 38629.7	2.74E-04	At5g03690
634	Nucleic Acid Binding	6 / 36018	1.48E-05	At5g04430
635	Atp Binding / Atp-Dependent Helicase/ Helicase	6 / 39080.1	2.27E-05	At5g11170
636	Cam6 (Calmodulin 6); Calcium Ion Binding	3.9 / 16833.6	3.41E-07	At5g21274
637	Rna Binding / Structural Constituent Of Ribosome	4.3 / 45110.1	6.24E-07	At5g30510
638	Lpd2 (Lipoamide Dehydrogenase 2); Fad Binding	6 / 53985.7	1.19E-10	At3g17240
639	Nit1 (Nitrilase 1)	6 / 38151.9	6.30E-07	At3g44310
640	Unknown Protein	6 / 57584.9	1.73E-04	At5g14260

Table 3.2. Continued.

S.No.	Protein Identification	pI / MW	P (pro)	Gene/locus
641	Glycine Hydroxymethyltransferase	10.07 / 57341.4	5.27E-05	At5g26780
642	Glycine Hydroxymethyltransferase	10.07 / 59129.5	5.27E-05	At5g26780
643	Bip; Atp Binding	4.25 / 73560.8	5.68E-06	At5g42020
644	Cra1 (Cruciferina); Nutrient Reservoir	9.9 / 31645.2	5.68E-07	AT5G44120
645	Cra1 (Cruciferina); Nutrient Reservoir	6 / 41031.5	5.68E-07	At5g44120
646	Unknown Protein	4.25 / 25644.1	7.40E-05	At5g53490
647	Structural Constituent Of Ribosome	10.5 / 21976.5	2.69E-05	At5g54600
648	Histone H3	12.5 / 15267.8	7.82E-07	
649	Aldh11A3; Aldehyde Dehydrogenase/ Oxidoreductase	6 / 53060	4.01E-08	At2g24270
650	Cat3 (Catalase 3); Catalase	6 / 48898.3	4.42E-05	AT1G20620
651	Ribulose-Bisphosphate Carboxylase	5.1 / 14558.7	7.53E-06	At1g67090
652	Adk1 (Adenosine Kinase 1)	6 / 33325.9	5.00E-06	AT3G09820
653	Unknown Protein	5.1 / 14249.4	6.30E-06	AT3G53990
654	Shd (Shepherd); Atp Binding / Unfolded Protein Binding	4.3 / 94148.2	2.60E-07	At4g24190
655	Fructose-Bisphosphate Aldolase	9.6 / 41343.9	3.28E-08	At4g38970
656	Ca2 (Carbonic Anhydrase 2); Carbonate Dehydratase	4.3 / 28344.2	2.78E-10	At5g14740
657	Glutamate-Ammonia Ligase (Ec 6.3.1.2)	4.25 / 40729.6	3.48E-05	At3g17820
658	Catalase (Ec 1.11.1.6) 3	8.3 / 56688.1	3.32E-07	AT1G20630
659	Hypothetical Protein At2G44920 [Imported]	8.3 / 20295.8	8.38E-09	At2G44920
660	Glycine Hydroxymethyltransferase (Ec 2.1.2.1) A_Ig002P16.3	10.07 / 59304	5.27E-05	AT4G13890
661	Hypothetical Protein At2G37660	6 / 36084.8	2.18E-05	At2g37660
662	Histone H3	12.5 / 15406	7.82E-07	

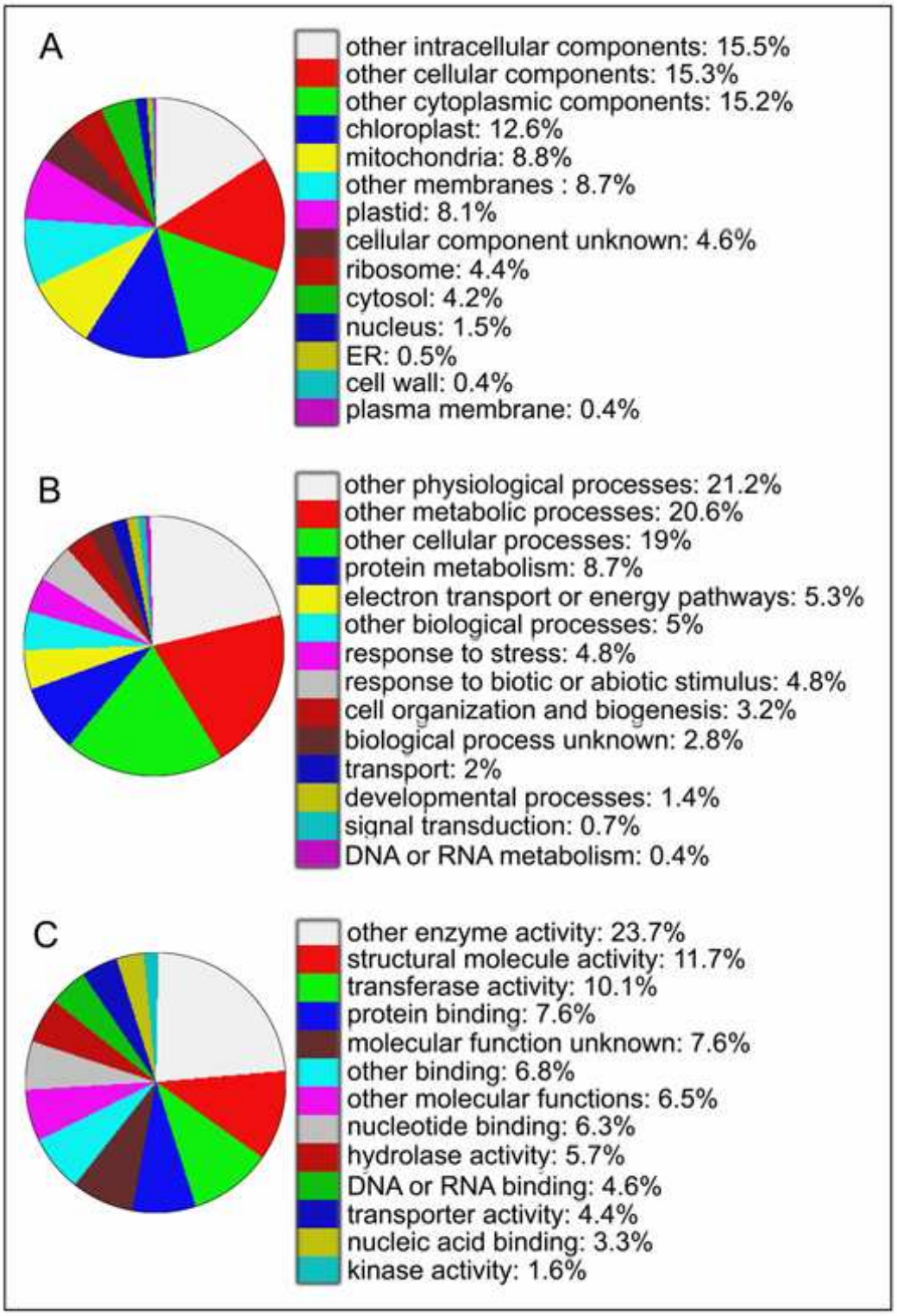

Figure 3.3: Functional classification of cotyledon MudPIT proteome.

Note: (A) Cellular localization, (B) Biological process, (C) Molecular function.

Discussion

The cotyledons were selected to carry out cell dedifferentiation studies in *Arabidopsis* because cotyledon is an embryonic organ whose cells have underwent differentiation with a terminated cell fate. For callus induction from cotyledons, cell dedifferentiation has to occur. Therefore, it is an excellent system to reveal the proteome dynamic change associated with cell dedifferentiation. To help understanding protein differential regulation during cell differentiation, the cotyledon proteome has been examined as the first step. Cotyledon development is different from most of the visible plant organs. It initiates during embryogenesis and its preliminary structure is completed during seed maturation. Cotyledon shares many common features with leaves in structure and function, including the morphology for efficient light harvesting and the photosynthetic capability. On the other hand, cotyledons are different from leaves. While leaves are rich in trichomes, for example, there is no trichome in cotyledons. In addition, there are a lot of storage nutrients, which are partially digested and mobilized during seed germination, in cotyledon cells but not in leaf cells. The cotyledon proteome has not been studied in plants thus far. In this study 500 *Arabidopsis* cotyledon protein spots have been successfully identified, corresponding to 353 distinct genes, with high confidence using 2-DE followed by MALDI-TOF/TOF. From MudPIT results, using ESI MS/MS, 662 proteins have been identified with the parameters: Xcorr values = +1-1.9; +2-2.2; +3-3.75, Delta Corr = 0.1 and used pFactor - 0.001. By combining the data from both methods, the proteome map of cotyledon was constructed with 1023 proteins. There are only 139 proteins commonly found both in 2-DE and MudPIT. This protein distribution is represented in Figure 3.4 using a Venn diagram. The genes corresponding to all these

proteins were characterized using the Gene Ontology tools. They were grouped into different levels and presented in pie charts (Figure 3.5). Our results show that the chloroplast genes occupy the largest number of identified genes, about 13.7% of the total identified genes. Other membrane proteins represented 9.6% and mitochondrial proteins with 8.4%. Ribosome, cytosol, nucleus, plasma membrane, ER, cell wall, and Golgi apparatus only occupy 3.9%, 3.3%, 2.1%, 0.7%, 0.5%, 0.3%, and 0.1% of the total identified genes, respectively (Figure 3.5A). These results strongly substantiate that the primary role of cotyledon is photosynthesis and energy metabolism. In addition, analyses based on Biological Process reveal that 21% of the identified proteins are involved in Other physiological processes, 20.5% are involved in Other metabolic processes in addition to those involved in protein, nucleic acids, and energy metabolic pathways, providing further evidence of the important role of metabolism and photosynthesis in the cotyledon cells. Proteins in other biological processes and molecular function categories were presented in detail in Figures 3.5B and 3.5C. The establishment of the cotyledon proteome map has provided new insight into the function of cotyledon at the proteome level, enriched the plant proteome database, and set a foundation for further research using cotyledon as materials, including comparative proteomics during cell dedifferentiation as presented in this study.

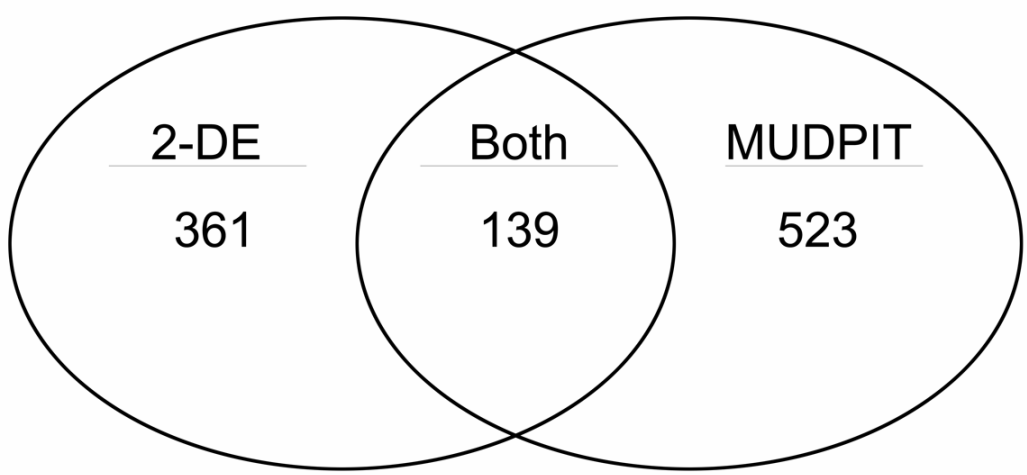

Fig 3.4: Venn diagram representing the proteins distribution identified using 2-DE and MudPIT.

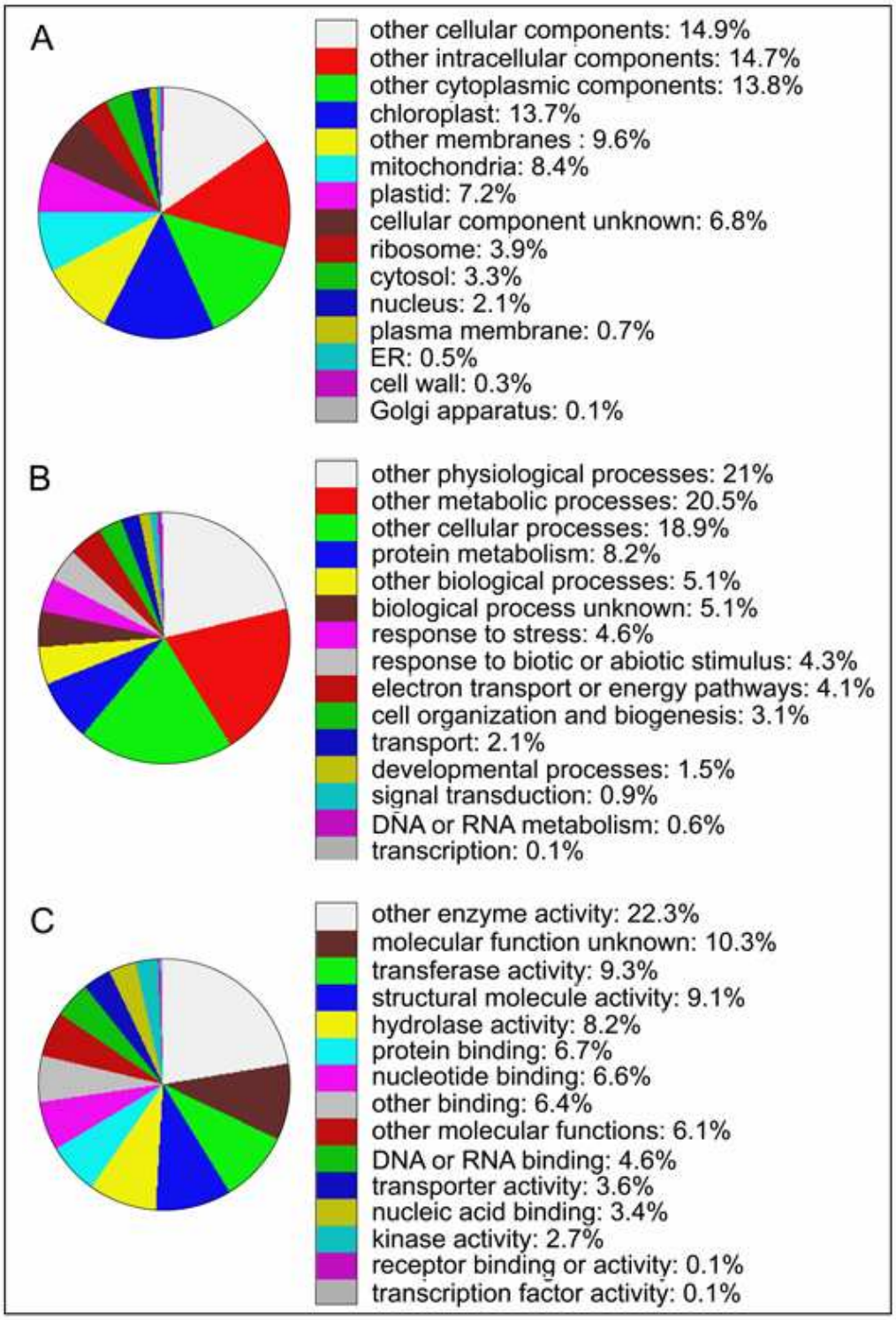

Figure 3.5: Functional classification of cotyledon proteins identified by 2-DE and MudPIT.

Note: (A) Cellular localization, (B) Biological process, (C) Molecular function.

In annotating proteins many of them represent in more than one category. For 500 2-DE identified proteins about 829 Cellular component GO terms were found and classified into higher hierarchies. For 662 MudPIT identified proteins, about 1275 Cellular component GO terms were found and also classified to higher hierarchies. Figure 3.6A represents the percentage of identified proteins by 2-DE and MudPIT in each cellular compartment.

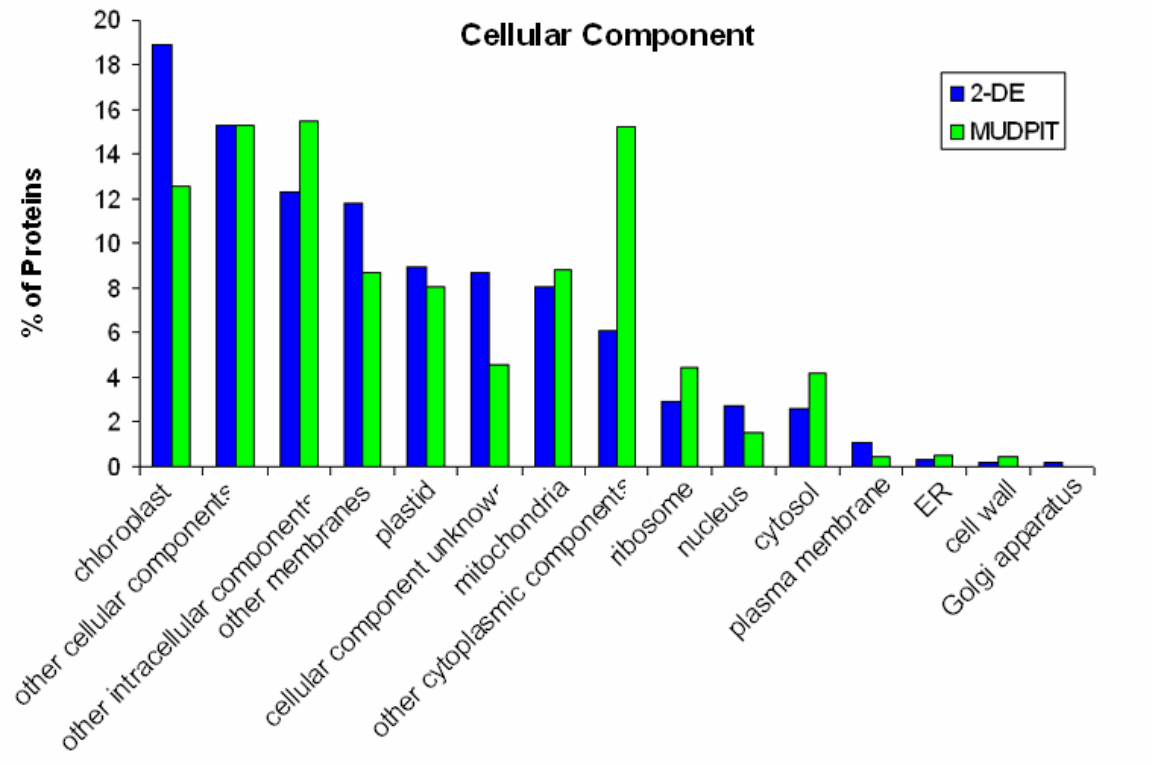

Fig 3.6A: Percentage of proteins distribution from two data sets according to their Cellular Component.

About 911 GO terms for Biological process were found for 500 2-DE identified proteins and about 1276 found for 662 MudPIT identified proteins. Their classification was clearly depicted in Figure 3.6B.

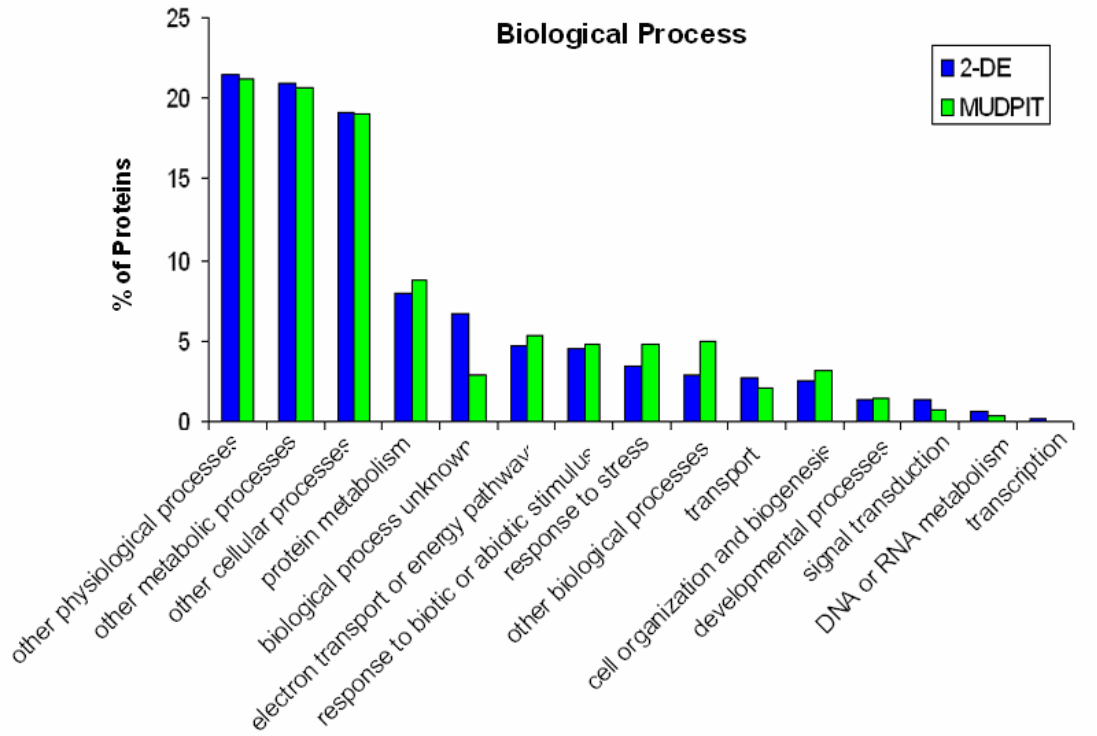

Fig 3.6B: Percentage of proteins distribution from two data sets according to their Biological Process.

There were 1153 GO terms found for Molecular Function annotations from 500 2-DE identified proteins and 1339 GO terms found for MudPIT identified proteins under Molecular Function category. These GO terms classified into higher level functional categories and numerically showed in Figure 3.6C.

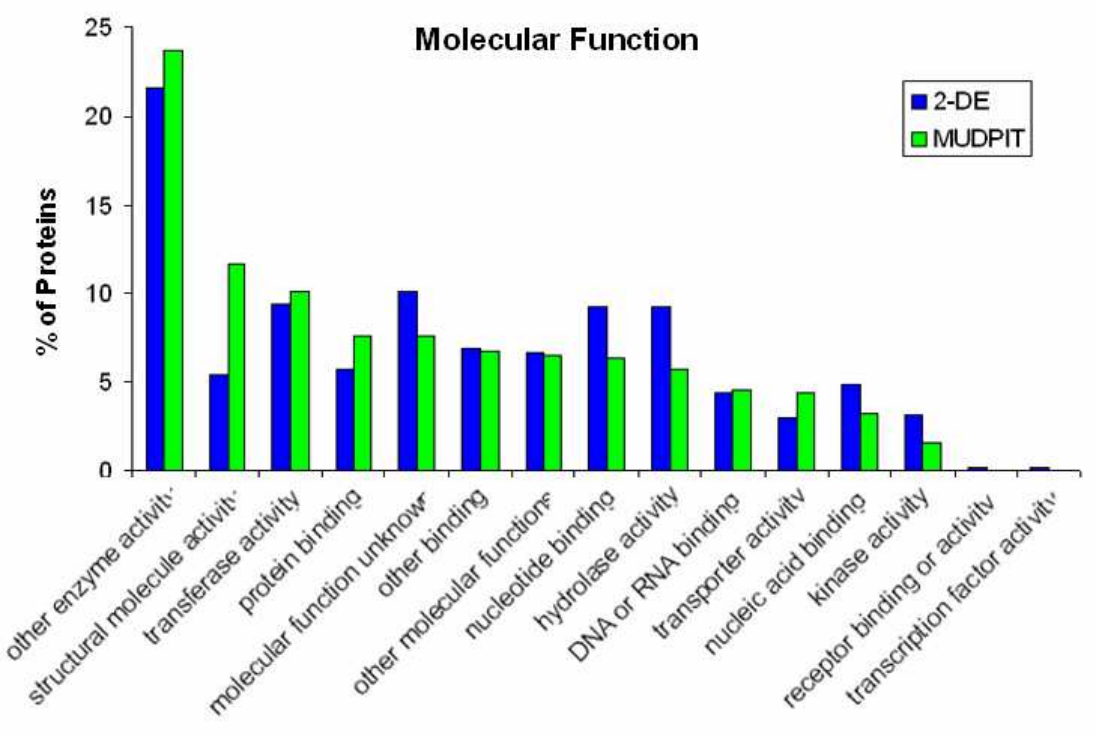

Fig 3.6C: Percentage of proteins distribution from two data sets according to their Molecular Function.

Molecular mass distribution of 500 high confident proteins identified by 2-DE are ranged from 5.3KDa to 183.4KDa. Most of these proteins, about 441, are distributed within 20-60 KDa. These 2-D identified 500 proteins distribution according to molecular mass is clearly depicted in Figure 3.7. Whereas, the molecular mass distribution is ranged from 1.5KDa to 132.6KDa for the 662 MudPIT identified proteins (Figure 3.7). Most of these proteins, almost 553, are within the range of 10-60 KDa. It has also been observed that low molecular mass proteins coverage is high in MudPIT analysis. Because, low molecular mass proteins are very poorly stained with staining dyes on 2-DE gels and there are more chances to miss them at time of spot picking. Five proteins with molecular mass over 150KDa are identified on 2-DE gels. In contrast no protein over 150KDa is

identified in MudPIT samples. It is better to employ MudPIT technique when dealing with low molecular mass proteins.

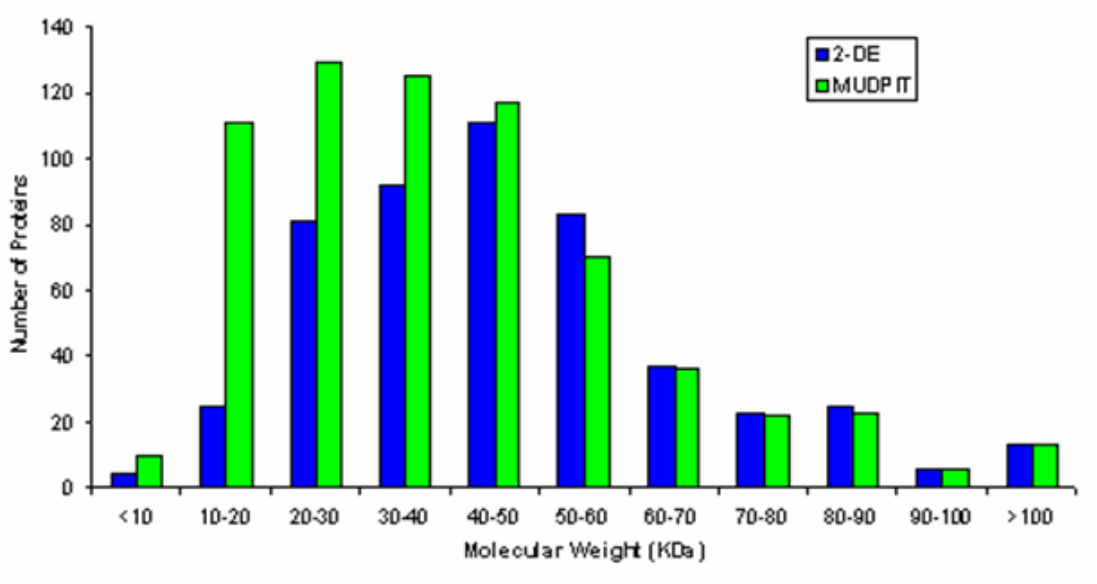

Fig 3.7: 2-DE and MudPIT identified proteins distribution according to their Molecular weight.

The pI distribution of 500 high confident proteins identified by 2-DE are ranged from 4 to 9.5. Almost 299 proteins are distributed within the range of pI 5-7 (Figure 3.8). Whereas, the pI distribution from 662 MudPIT identified proteins is ranged from pI 3.9 to 12.5. About 396 MudPIT identified proteins are within the pI range of 4-7. This distribution is clearly depicted in Figure 3.8.

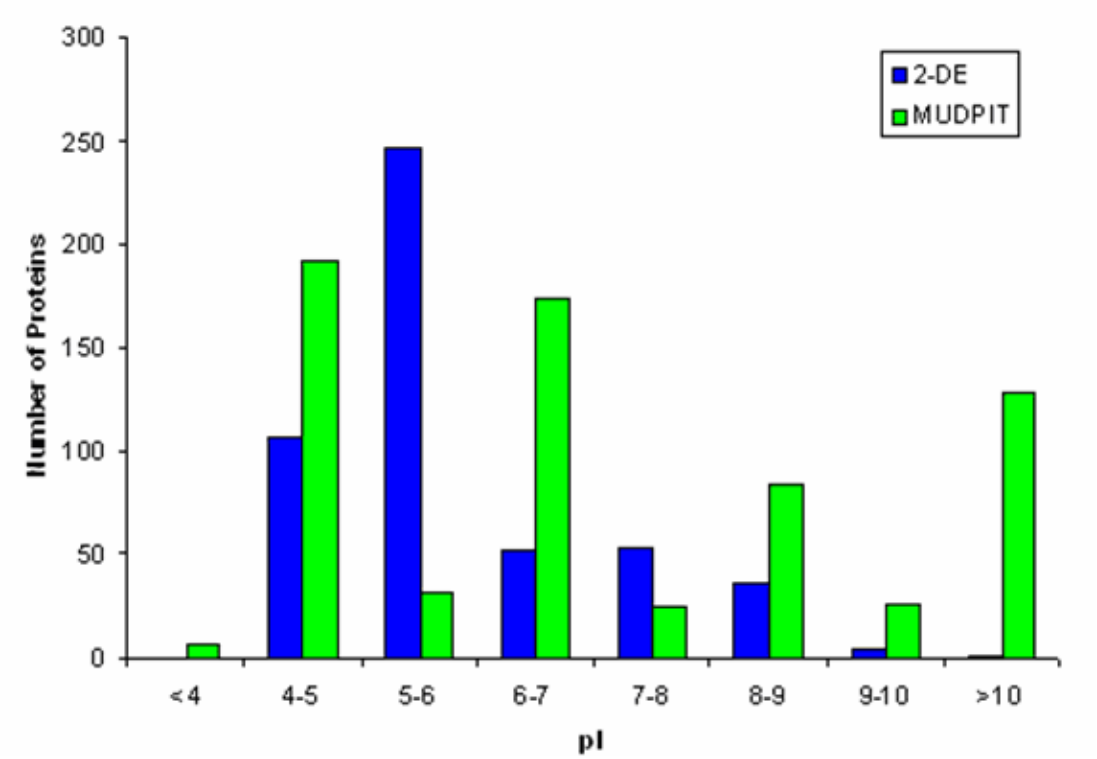

Fig 3.8: 2-DE and MudPIT identified proteins distribution according to their pI.

128 basic proteins with pI over 10 are identified by MudPIT analysis. It has been observed that it is important to employ MudPIT technique in order to increase the basic proteins coverage.

CHAPTER IV

PROTEOME DYNAMIC CHANGE DURING CELL DEDIFFERENTIATION IN *ARABIDOPSIS THALIANA*

Abstract

Cell dedifferentiation is a cell fate switching process in which a differentiated cell reverts to a status with competence for cell division and organ regeneration like an embryonic stem cell. Although the phenomenon of cell dedifferentiation has been known for over two and a half centuries in plants, little is known of the underlying mechanisms. In this study the dynamic change of the *Arabidopsis* cotyledon proteome in the time course of cell dedifferentiation has been investigated. Among the 353 distinct genes, corresponding to 500 2-DE gel protein spots identified with high confidence, 12% have over two fold differential regulations within the first 48 hour induction of cell dedifferentiation. The distributions of these genes among different Gene Ontology categories and gene differential regulations within each of the categories have been examined. These studies have provided significant new insight into protein differential expression during cell dedifferentiation in *Arabidopsis thaliana.*

Introduction

Most plant cells possess totipotency. The concept of totipotency involves a two-step process. The first is acquiring, in response to appropriate stimuli, stem cell morphogenetic potential (dedifferentiation), and the second is expressing this potential during morphogenesis (regeneration) (Bhojwani and Razdan, 1983; Vasil et al., 1984; Alberts et al., 1994; Sugiyama, 1999). At the molecular level, the genome undergoes reprogramming to restore the stem cell status during dedifferentiation, and then the genes in the reprogrammed genome are expressed following the orderly pattern of a zygotic cell during regeneration. Specifically, the proteins exclusive to the prior cell status have to be removed and proteins required for the dedifferentiated status have to be synthesized in a timely fashion during dedifferentiation. Recent success in mammalian cloning suggests that the differentiated mammalian nucleus is also capable of undergoing dedifferentiation, leading to the formation of new organisms from mature somatic cells if the conditions are suitable (Solter, 2000).

The phenomenon of cell dedifferentiation has been known for over two and a half centuries in plants (Duhamel du Monceau, ; Vasil et al., 1984). It has been well established that phytohormones, such as auxin and cytokinin, play a critical role in cell dedifferentiation in plants (Bhojwani and Razdan, 1983; Vasil et al., 1984). Studies, which were initiated in tobacco pith tissue culture and confirmed in other organisms, have demonstrated that an appropriate cytokinin/auxin ratio in the medium is critical to the fate of a cell in culture (Krikorian, 1995). The absolute concentrations of the two hormones are not critical, providing that they are within an appropriate concentration range (Skoog and Miller, 1957; Krikorian, 1995). Callus is a group of relatively

undifferentiated dividing cells produced following cell dedifferentiation in plants. When the ratio of cytokinin/auxin is about one, a callus mass is produced and the callus continuously proliferates. When the level of auxin relative to that of cytokinin is high, roots regenerate; when cytokinin relative to that of auxin is high, shoots regenerate. The device of adjusting auxin/cytokinin ratios to induce shoots and roots and to maintain cell propagation is now a well-established practice for a variety of plants in both research and industry.

In tobacco protoplast culture, it has been known that the ubiquitin proteolytic pathway is indispensable for protoplast progression into S phase. A specific inhibitor of the 26S proteasome, MG132, interferes with the entry of protoplasts into S phase (Zhao et al., 2001). In agreement, it has also been reported that genes coding for ubiquitin are induced during cell dedifferentiation in tobacco (Jamet et al., 1990). Recently, hypomethylation of the ribosomal RNA genes is found to be associated with cell dedifferentiation in tobacco cells (Koukalova et al., 2005).

Cell dedifferentiation can occur rapidly when an appropriate environment is provided. In the presence of auxin and cytokinin, about 10% of tobacco protoplasts approach S phase within 48 hours and 30-40% cells approach S to G2 phase within 72 hours (Zhao et al., 2001). Studies on *Arabidopsis* cotyledon protoplast culture reveal that the first cell division has occurred within 24 hours while the majority of the cells divide between 48 to 72 hours (Dovzhenko et al., 2003). Since cells have to go through dedifferentiation before re-entering the cell cycle, these observations suggest that cell dedifferentiation can occur within less than 24 hours in some plant cells.

Although the above studies have provided some insight into the plant cell dedifferentiation process, little is known of the event at the molecular level. No cellular components that perceive the environmental stimuli and mediate the genome reprogramming process have been identified. In this chapter, cotyledon proteome dynamic change in the time course of cell dedifferentiation using comparative proteomics has been reported.

Materials and Methods

Plant materials and induction of cell dedifferentiation

Arabidopsis thaliana (Columbia) seeds were plated on germination media (Murashige & Skoog basal salt with 10 g/L sucrose, pH 5.7) after sterilization. Following 4 days of vernalization treatment at 4 ^{0}C, the plates were transferred to an incubator for germination at 22 ^{0}C with 16 hours of light and 8 hours of dark. Cotyledons harvested from 10 days old seedlings were chopped into small pieces (1 to 2 mm in diameter) for induction of cell dedifferentiation with 2 mg/l 2,4-D (2,4-Dichlorophenoxyacetic acid) and 0.2 mg/l KT (Kinetin) in B5 medium (pH 5.7) supplemented with 10 g/l sucrose.

Protein extraction

Cotyledon tissues, including control and cell dedifferentiation treated samples, were ground in liquid nitrogen with mortar and pestle into fine powders. Proteins were extracted using a modified phenol extraction protocol (Hurkman and Tanaka, 1986a) as follows: Ground tissues were suspended in an extraction buffer (0.9M sucrose, 0.5M tris-

HCl, 0.05M EDTA, 0.1M KCl and 2% G mercaptoethanol added freshly, pH 8.7), mixed with an equal volume of saturated phenol (pH 8.0) and then homogenized for 10 minutes. The homogenate was centrifuged at 2,500g for 10 minutes, and the phenol phase was recovered. The phenol extraction was repeated three times, and the final collection of phenol was mixed with five volumes of precipitation buffer (methanol with 0.1 M ammonium acetate and 1% G-mercaptoethanol). Precipitation was carried out at -70 ^{0}C overnight. The precipitant was recovered by centrifugation at 13,400g for 10 minutes and the pellet was washed three times with cold precipitation buffer and followed by three more washes with ice cold 70% ethanol. The protein pellet was lyophilized to powder in a speed vacuum (LABCONCO, model LYPH-LOCK 6) and stored at -70 ^{0}C. At least three biological replicas were extracted for each treatment.

Two-Dimensional Polyacrylamide Gel Electrophoresis (2D PAGE)

Proteins were dissolved thoroughly in a rehydration buffer (7M urea, 2M thiourea, 4% CHAPSO, 1% DTT, and 0.2% Ampholines) and centrifuged at 12,000 g for 10 minutes to remove undissolved content. The supernatant was quantified using a Bio-Rad Rc Dc protein assay kit according to the instructions of the manufacturer. The quantified proteins were then used for 2D PAGE. Isoelectric focusing (IEF) was carried out using a Bio-Rad PROTEAN IEF cell on 24 cm 3-10 pH non linear IPG strips (Bio-Rad). One milligram of protein in 400 µl of rehydration buffer was loaded into the IEF tray and active rehydration was carried out at 23 ^{0}C for 12 hours, followed by 250 V for 2 hours, and a linear increase of voltage to 10,000V for 4 hours. The isoelectric focusing was performed at 23 ^{0}C for a total of 90,000 VH. After the IEF, the strips were

equilibrated in an equilibrating buffer containing 6M urea, 0.375M Tris-HCl (pH 6.8), 20% glycerol, 2% SDS, and 2% dithiothreitol for 15 minutes and followed by equilibration for another 15 minutes in a buffer containing 6M urea, 0.375M Tris-HCl (pH 6.8), 20% glycerol, 2% SDS, 0.1% bromophenol blue, and 2.5% iodoacetamide.

The equilibrated IPG strips were loaded on horizontal slab gels (25 X 20.5 X 1.5mm) containing 12% (w/v) separating gel and 4% stacking gel (w/v). Electrophoresis was carried out in a Bio-Rad PROTEAN PLUS horizontal Dodeca cell at 20 mA/gel.

The gels were stained with SYPRO Ruby fluorescence stain (Bio-Rad) according to the protocols provided by the manufacturer and scanned with a VersaDoc4000 image system (Bio-Rad). For each treatment, at least three 2-DE gels representing the three biological repeats were used for data analyses. The images were analyzed with PDQuest 7.3.1. software (Bio-Rad, Hercules, CA), including gel cropping, anchor spots selection, alignment, subtracting background, and spot volume normalization for differences in staining intensities. The spots with consistent size and shape within replicate group were chosen for quantitative analysis. Meanwhile, the protein spots were checked manually to ensure that all analyzed spots were true protein spots and the gel alignment was appropriate. The cut off for differentially regulated proteins was two fold of change in all biological replicas. For quantification analysis, only well separated spots were included. In case two proteins were identified in a single protein spot that displayed differential expression, these proteins were not included in the list of differentially regulated proteins because it was unknown which one of them was differentially regulated.

In gel digestion and Mass Spectrometry, Gene Ontology analysis were carried out as explained in materials and methods section of Chapter 3.

Results

Cell dedifferentiation is a cell fate switching process in which little is known at the protein level. To gain insight, we examined the protein expression profile dynamic change during cell dedifferentiation using excised *Arabidopsis* cotyledons as explants. The first cell division of cultured cotyledon protoplasts has been reported to occur within 24 hours after protoplast isolation and hormone induction and the majority of protoplasts divide between 48 to 72 hours in *Arabidopsis* (Dovzhenko et al., 2003). Since a mature somatic cell reenters cell division only if it has gone through the cell dedifferentiation process, these results suggests that cell dedifferentiation can be completed within 48 hours in most cotyledon cells. Therefore, the protein expression profile at 0, 12, 24, and 48 hour time points after induction of cell dedifferentiation with 2 mg/l 2, 4-D and 0.2 mg/l KT has been examined in *Arabidopsis* cotyledons. A representative gel at each time point of induction is shown in Figure 4.1, where the up-regulated protein spots are marked with black arrows and down-regulated spots were marked with white arrows. The cut off of differential regulation was two fold differences in expression compared with 0 hour induction in all three biological replicas. The analyses were done using PDQuest 7.3.1 software. Among the 353 identified distinct genes, 30 were up-regulated and 12 were down-regulated. The total differential regulated genes were 42, which was about 12% of the identified cotyledon genes. The molecular identities and the corresponding protein spot numbers of these differentially regulated proteins are presented in Table 4.1, which are grouped according to induction time as well as up and down regulations. In addition, the relative protein quantities (the average of

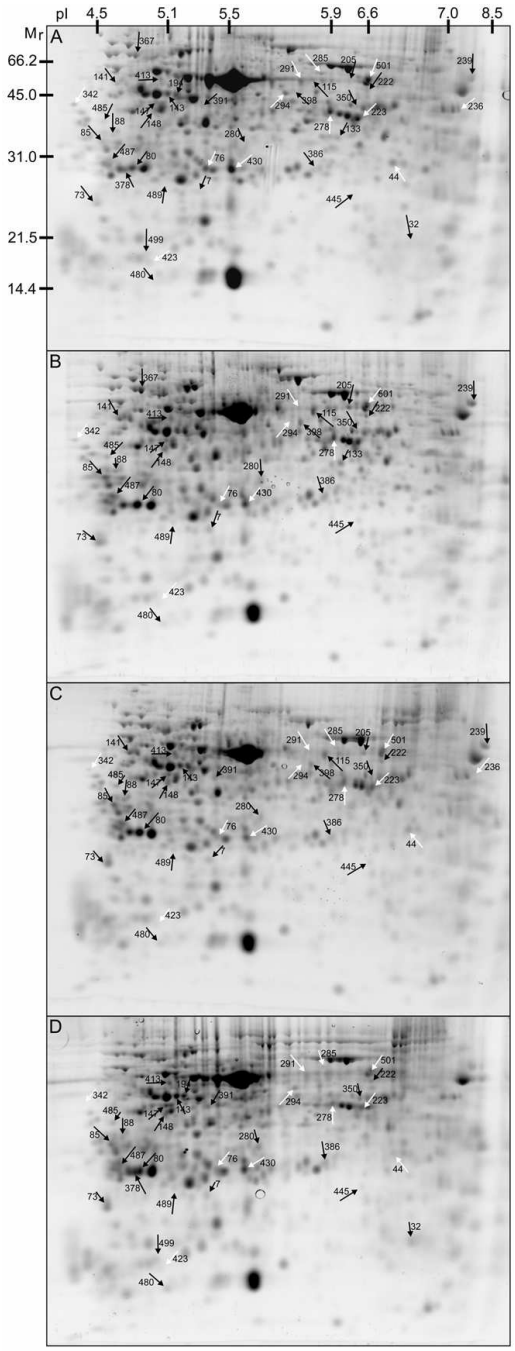

Figure 4.1: Protein differential regulation in the time course of cell dedifferentiation.

Note: The differentially up-regulated proteins are marked by black arrows and the down-regulated proteins are marked by white arrows. (A) Proteins without hormone induction. All differential regulated proteins within the first 48 h of induction are marked on the 2-DE gel. (B) Proteins with 12 h of hormone induction. Proteins with twofold change between 0 and 12 h are indicated by arrows. (C) Proteins with 24 h of hormone induction. Proteins with twofold change between 0 and 24 h are indicated by arrows. (D) Proteins with 48 h of hormone induction. Proteins with twofold change between 0 and 48 h are indicated by arrows.

Table 4.1: Differentially expressed proteins in the time course of cell dedifferentiation

Proteins with at least Two fold up regulation after 12 hour induction

Spot #	Protein Identified	No. of Peptides	C.I.%	Express. Graph	Ratio	Gene ID
7	Genomic DNA, chromosome 3, TAC clone:K13N2 (K13N2_9)	10	100		3.02	AT3G25770
73	Translationally controlled tumor like protein	09	100		2.05	AT3G16640
80	Chlorophyll A-B binding protein 2, (LHCII type I CAB-2)	11	100		2.21	AT1G29910
85	F23J3_30 (14-3-3 protein GF14chi) (Grf1)	19	100		2.01	AT4G09000
88	T32G9_30	16	100		2.91	AT1G35160
115	F26F24_4	07	100		2.28	AT1G70580
133	12S seed storage protein (cruciferin), putative	09	100		7.27	AT4G28520
141	COP8 (Constitutive photomorphogenic) homolog	11	100		2.15	AT5G42970
147	Phospho ribulokinase, chloroplast precursor (EC 2.7.1.19)	11	100		2.2	AT1G32060
148	Putative pyruvate dehydrogenase E1 beta subunit	06	99.952		2.03	AT2G34590
205	Putative fumarase (T30B22.19)	11	99.989		3.48	AT2G47510
222	Putative GTP-binding protein	13	99.842		2.09	AT1G30580
239	Putative translation elongation factor eEF-1 alpha chain (Gene A4)	13	100		2.42	AT5G60390
280	V-type proton-ATPase	15	100		2.6	AT4G11150
350	Glyceraldehyde-3-phosphate dehydrogenase (EC 1.2.1.13)	08	100		2.34	AT3G26650
367	Cell division protein ftsH homolog 2, (E.C.3.4.24.)	15	100		2.53	AT5G42270
386	GTP-binding nuclear protein RAN-2	06	100		2.78	AT5G20020
398	Putative GDP-mannose pyrophosphorylase (F9E10_24)	07	99.985		2.51	AT1G74910
413	Expressed protein	10	100		2.01	AT1G73850
445	GTP-binding protein SAR1B	07	100		3.51	AT1G56330
480	Glycine-rich RNA-binding protein (clone A81)	03	99.995		3.02	AT4G13850
485	Spermidine synthase 1 (EC 2.5.1.16)	10	100		2.11	AT1G23820
487	14-3-3 protein homolog RCI2	14	100		2.12	AT5G10450
489	Proteasome subunit beta type 3-1 (EC 3.4.25.1)	07	99.985		2.31	AT1G21720

Proteins with at least Two fold up regulation after 24 hour induction

Spot #	Protein Identified	No. of Peptides	C.I.%	Express. Graph	Ratio	Gene ID
7	Genomic DNA, chromosome 3, TAC clone:K13N2 (K13N2_9)	10	100		3.69	AT3G25770
73	Translationally controlled tumor like protein	09	100		3.09	AT3G16640
80	Chlorophyll A-B binding protein 2, (LHCII type I CAB-2)	11	100		2.5	AT1G29910
85	F23J3_30 (14-3-3 protein GF14chi) (Grf1)	19	100		3.14	AT4G09000
88	T32G9_30	16	100		6.64	AT1G35160
115	F26F24_4	07	100		2.01	AT1G70580
141	COP8 (Constitutive photomorphogenic) homolog	11	100		2.8	AT5G42970
143	T5I7.3 (Hypothetical protein)	26	100		2.02	AT2G39730
147	Phospho ribulokinase, chloroplast precursor (EC 2.7.1.19)	11	100		4.91	AT1G32060
148	Putative pyruvate dehydrogenase E1 beta subunit	06	99.952		2.52	AT2G34590
205	Putative fumarase (T30B22.19)	11	99.989		2.22	AT2G47510
222	Putative GTP-binding protein	13	99.842		4.08	AT1G30580
239	Putative translation elongation factor eEF-1 alpha chain (Gene A4)	13	100		7.75	AT5G60390
280	V-type proton-ATPase	15	100		9.59	AT4G11150
350	Glyceraldehyde-3-phosphate dehydrogenase (EC 1.2.1.13)	08	100		3.31	AT3G26650
386	GTP-binding nuclear protein RAN-2	06	100		3.5	AT5G20020
391	Reversibly glycosylated polypeptide-1	14	100		3.82	AT3G02230
398	Putative GDP-mannose pyrophosphorylase (F9E10_24)	07	99.985		2.1	AT1G74910
413	Expressed protein	10	100		2.05	AT1G73850
445	GTP-binding protein SAR1B	07	100		3.53	AT1G56330

Table 4.1. Continued.

Spot #	Protein Identified	No. of Peptides	C.I.%	Express. Graph	Ratio	Gene ID
480	Glycine-rich RNA-binding protein (clone A81)	03	99.995		7.73	AT4G13850
485	Spermidine synthase 1 (EC 2.5.1.16)	10	100		2.74	AT1G23820
487	14-3-3 protein homolog RCI2	14	100		3.67	AT5G10450
489	Proteasome subunit beta type 3-1 (EC 3.4.25.1)	07	99.985		3.8	AT1G21720

Proteins with at least Two fold up regulation after 48 hour induction

Spot #	Protein Identified	No. of Peptides	C.I.%	Express. Graph	Ratio	Gene ID
7	Genomic DNA, chromosome 3, TAC clone:K13N2 (K13N2_9)	10	100		2.69	AT3G25770
32	Non-symbiotic hemoglobin 1 (Hb1) (GLB1)	07	99.995		3.52	AT2G16060
73	Translationally controlled tumor like protein	09	100		2.33	AT3G16640
80	Chlorophyll A-B binding protein 2, (LHCII type I CAB-2)	11	100		2.09	AT1G29910
85	F23J3_30 (14-3-3 protein GF14chi) (Grf1)	19	100		2.2	AT4G09000
88	T32G9_30	16	100		2.12	AT1G35160
143	T5I7.3 (Hypothetical protein)	26	100		2.17	AT2G39730
147	Phosphoribulokinase, chloroplast precursor (EC 2.7.1.19)	11	100		5.24	AT1G32060
148	Putative pyruvate dehydrogenase E1 beta subunit	06	99.952		2.98	AT2G34590
194	Aspartate-semialdehyde dehydrogenase (EC 1.2.1.11)	08	99.95		2.62	AT1G14810
222	Putative GTP-binding protein	13	99.842		3.14	AT1G30580
280	V-type proton-ATPase	15	100		7.34	AT4G11150
350	Glyceraldehyde-3-phosphate dehydrogenase (EC 1.2.1.13)	08	100		3.41	AT3G26650
378	Putative photosystem II type I chlorophyll a/b binding protein	16	100		2.14	AT2G34430
386	GTP-binding nuclear protein RAN-2	06	100		2.43	AT5G20020
391	Reversibly glycosylated polypeptide-1	14	100		3.05	AT3G02230
413	Expressed protein	10	100		2.29	AT1G73850
445	GTP-binding protein SAR1B	07	100		4.12	AT1G56330
480	Glycine-rich RNA-binding protein (clone A81)	03	99.995		7.83	AT4G13850
485	Spermidine synthase 1 (EC 2.5.1.16)	10	100		2.6	AT1G23820
487	14-3-3 protein homolog RCI2	14	100		3.8	AT5G10450
489	Proteasome subunit beta type 3-1 (EC 3.4.25.1)	07	99.985		2.42	AT1G21720
499	Actin depolymerizing factor 2	05	100		2.87	AT3G46000

Proteins with at least Two fold down regulation after 12 hour induction

Spot #	Protein Identified	No. of Peptides	C.I.%	Express. Graph	Ratio	Gene ID
76	Carbonic anhydrase 2 (EC 4.2.1.1) (Carbonate dehydratase 2)	10	100		2.03	AT5G14740
278	ATP synthase gamma chain 1, (EC 3.6.3.14)	13	100		2.01	AT4G04640
291	Expressed protein	14	100		3.71	AT4G36105
294	At4g05190	12	100		2.76	AT4G05190
342	Late embryogenesis abundant (LEA) domain-containing protein	06	99.869		2.19	AT5G44310
423	Major latex-like protein (T23E23_22)	10	100		2.1	AT1G24020
430	L-ascorbate peroxidase, cytosolic (EC 1.11.1.11)	12	100		2.16	AT1G07890
501	F5F19.16 protein	13	100		2.02	AT1G52100

Proteins with at least Two fold down regulation after 24 hour induction

Spot #	Protein Identified	No. of Peptides	C.I.%	Express. Graph	Ratio	Gene ID
44	Putative L-ascorbate peroxidase, chloroplast precursor (EC 1.11.1.11)	09	100		2.15	AT4G09010
76	Carbonic anhydrase 2 (EC 4.2.1.1) (Carbonate dehydratase 2)	10	100		2.1	AT5G14740
223	Hypothetical protein (MQC12_13)	10	99.296		2.3	AT3G20370
236	MOA2_2	11	100		2.13	AT3G14420

Table 4.1. Continued.

Spot #	Protein Identified	No. of Peptides	C.I.%	Express. Graph	Ratio	Gene ID
278	ATP synthase gamma chain 1, (EC 3.6.3.14)	13	100		2.34	AT4G04640
285	RUBISCO large chain precursor (EC 4.1.1.39)	11	99.151		2.13	ATCG00490
291	Expressed protein	14	100		4.39	AT4G36105
294	At4g05190	12	100		2.88	AT4G05190
342	Late embryogenesis abundant (LEA) domain-containing protein	06	99.869		3.05	AT5G44310
423	Major latex-like protein (T23E23_22)	10	100		2.26	AT1G24020
430	L-ascorbate peroxidase, cytosolic (EC 1.11.1.11)	12	100		2.37	AT1G07890
501	F5F19.16 protein	13	100		3.31	AT1G52100

Proteins with at least Two fold down regulation after 48 hour induction

Spot #	Protein Identified	No. of Peptides	C.I.%	Express. Graph	Ratio	Gene ID
44	Putative L-ascorbate peroxidase, chloroplast precursor (EC 1.11.1.11)	09	100		2.58	AT4G09010
76	Carbonic anhydrase 2 (EC 4.2.1.1) (Carbonate dehydratase 2)	10	100		2.1	AT5G14740
223	Hypothetical protein (MQC12_13)	10	99.296		2.45	AT3G20370
278	ATP synthase gamma chain 1, (EC 3.6.3.14)	13	100		2.61	AT4G04640
285	RUBISCO large chain precursor (EC 4.1.1.39)	11	99.151		3.5	ATCG00490
291	Expressed protein	14	100		2.03	AT4G36105
294	At4g05190	12	100		2.14	AT4G05190
342	Late embryogenesis abundant (LEA) domain-containing protein	06	99.869		8.34	AT5G44310
423	Major latex-like protein (T23E23_22)	10	100		2.26	AT1G24020
430	L-ascorbate peroxidase, cytosolic (EC 1.11.1.11)	12	100		2.29	AT1G07890
501	F5F19.16 protein	13	100		4	AT1G52100

Note: Spot #: the spot number was identical with those in Table 1; No. of Peptides: Number of peptides matched with identified protein in mass analyses. C.I. %: cross confidence Interval obtained using the MASCOT algorithm. Expres Graph: The four bars represent relative expression level of the protein at 0, 12, 24, 48 hour induction from left to right. Ratio: the up or down regulation ratio of the protein at the time point indicated in the table compared with 0 hour induction.

three biological replicas) of each protein spot at the four tested time points were presented as a small graph embedded in Table 4.1. The GO category distributions of these differentially regulated proteins were analyzed and presented as a pie chart in Figure 4.2A, which were given as percentage distribution of these proteins among the various GO categories.

As shown in Figure 4.2A, 10.2% of the differentially regulated proteins were chloroplast proteins and 6.2% were nuclear protein. Since much more chloroplast proteins (18.9 %) were identified than nuclear proteins (2.7%) (Figure 3.2A), the nuclear

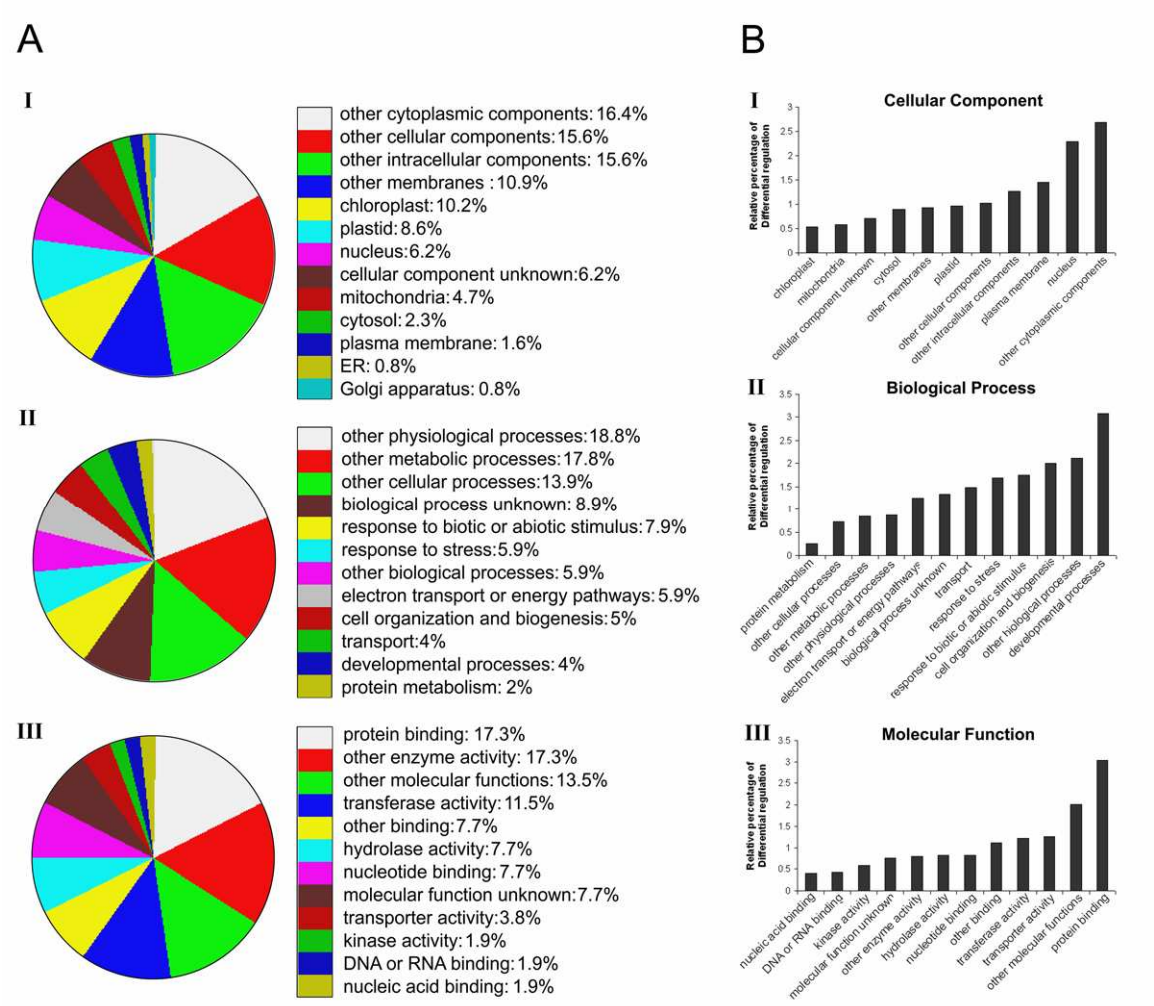

Figure 4.2: Distribution of differentially regulated proteins among different GO categories.

Note: The cut off for differential regulation was two fold of change. All differentially regulated proteins within the 48 hour induction period were included. **A)** Pie chart presentation of the differentially regulated proteins among different GO categories. Percentage distribution of the differentially regulated proteins was used. The pie charts were generated using the program provided by TAIR (The *Arabidopsis* Information Resource). **A-I,** Protein distribution based on Cellular Localization. **A-II,** Protein distribution based on Biological Processes. **A-III,** Protein distribution based on Molecular Functions. **B)** Relative ratios of differentially regulated proteins in different GO categories. The average relative ratio of differentially regulated proteins should be 1.0. **B-I,** Relative ratio of differentially regulated proteins in categories classified based on Cellular Localizations. **B-II,** Relative ratio of differentially regulated proteins in categories classified based on Biological Processes. **B-III,** Relative ratio of differentially regulated proteins in categories classified based on Molecular Functions.

proteins had a much higher ratio of protein differential regulation than chloroplast proteins. However, this feature was not reflected in the distribution pie chart shown in Figure 4.2A. For a better illustration of protein differential regulation in each GO category, Relative differential regulation ratio (Rdif) was defined as percentage of differential regulated protein in a category (as shown in Figure 4.2A) divided by percentage of all identified proteins in the same category (as shown in Figure 3.2). With this definition, the expected average Rdif should be 1.0 for these GO categories. If a GO category has a Rdif above 1.0, it means that the proteins in this category have an above average protein differential regulation during cell dedifferentiation and suggests that proteins in this GO category are either actively involved in cell dedifferentiation or highly affected by the cell dedifferentiation process. The Rdif values of all these GO categories were presented in Figure 4.2B. Analyses based on cellular components showed that Other Cytoplasmic Components and Nucleus had the two highest ratios of protein differential regulation (Fig 4.2B-I). In contrast, chloroplast and mitochondrion had the two lowest ratios of protein differential regulation, indicating that these two organelles play a minor role in the early stage of cell dedifferentiation. When proteins were analyzed according to the Biological Processes (Fig 4.2B-II), the Developmental Process, Other Biological Processes, Cell Organization and Biogenesis, and Response to Stresses had high ratios of differentially regulated proteins in the order presented. In contrast, Protein Metabolism and Metabolic Processes had the two lowest ratios of differentially regulated proteins. When the proteins were analyzed according to Molecular Function (Figure 4.2B-III), proteins in the categories of Protein Binding and Other Molecular Function had the two highest ratios of differentially regulated proteins.

Discussion

Cell dedifferentiation is a genome reprogramming process in which little is known at the molecular level in both plants and animals, including human beings. In this study protein expression profile dynamics in the time course of cell dedifferentiation has been investigated using *Arabidopsis* cotyledons as explants. Our results show that cell dedifferentiation involves extensive protein quantitative and qualitative changes in almost every cellular compartment and cellular process. Among the 353 identified genes 42 (12%) have changed over two fold within the first 48 hour induction period. Although these 42 genes are distributed in all cellular compartments and biological processes, Gene Ontology analyses have revealed that proteins in Nucleus and Other cytoplasmic components have the top two highest ratios to be differentially regulated. Meanwhile, analyses based on molecular function have shown that proteins involved in Protein Binding and Other Molecular Function had high ratios of protein differential regulation. Finally, studies based on Biological Process have found that proteins in Developmental Processes, Other Biological Process, and Cell Organization and Biogenesis all have high ratio of differential regulated proteins. These results indicate that cell dedifferentiation requires substantial change in multiple biological processes. These observations have lead to the following preliminary conclusions: 1) Nuclear proteins are actively involved in cell dedifferentiation; 2) Development Process and Cell Organization and Biogenesis are essential to the cell dedifferentiation. 3) Proteins in those not well defined GO categories, such as Other Molecular Function, Other Biological Process, and Other Cytoplasmic Components, are also essential to cell dedifferentiation.

To gain a better understanding on the cell dedifferentiation process, several groups of differentially regulated proteins are listed here and discussed their possible roles in cell dedifferentiation.

14-3-3 proteins: The 14-3-3 proteins are a large group of phosphoserine / phosphothreonine-binding proteins that play a critical role in various signal transduction pathways (Dougherty and Morrison, 2004; Ferl, 2004). The 14-3-3 protein homologs RCI2, T32G9_30, and 14-3-3-like protein GF14 chi (Grf1) were up-regulated with the maximal expression level at 24 hours (Table 4.1). In addition, the GF14 Chi and T32G9_30 were both stained by Pro-Q Diamond dye. Although the biological functions of these three proteins still remain to be explored in plants, the up regulation of these proteins suggests the involvement of protein phosphorylation in the cell dedifferentiation process. Identification of the binding partners of these three proteins might lead to the discovery of cellular signaling pathways critical to cell dedifferentiation.

Proteases: Several proteases or proteins in the protein degradation pathways were up-regulated during cell dedifferentiation. These proteins included cell division protein ftsH homolog 2 (an ATP-dependent zinc metallopeptidase), COP8, and proteasome subunit beta type 3-1 (Table 1). The ftsH homolog 2 was also stained by Pro-Q Diamond stain. Since auxin plays an essential role in cell dedifferentiation and auxin is known to accomplish its role via regulating ubiquitin mediated protein degradation in plants (Parry and Estelle, 2006), the up regulation of COP8 and proteasome subunit is consistent with the essential role of auxin in cell dedifferentiation. On the other hand, only 12 proteins are down-regulated over two fold within the first 48 hour induction period. One possible reason is that 48 hours is too short a period to observe the change in protein level unless

the protein is directly involved in cell dedifferentiation process. Another possibility is that the down-regulated proteins are low abundance proteins. The down regulation of these proteins is beyond the detection limit of the methods used here.

Translational controlled tumor protein (TCTP) and its possible interaction protein*:* The TCTP protein was identified from two distinct protein spots (Spot #73 and 74). The # 73 spot was also stained by the Pro-Q Diamond dye and was up-regulated during cell dedifferentiation with the expression peak at 24 hours. Consistent with our result, the mammalian TCTP protein has been reported to be phosphorylated by mitotic polo-like kinase at serine residues *in vitro* (Yarm, 2002). TCTP has been shown to play a role in (i) cell cycle control and tumourigenesis, (ii) act as growth factor, (iii) stress response, (iv) microtubule stabilizing, (v) calcium binding, (vi) interacting with translational elongation factors eEF1A (Cans et al., 2003; Bommer and Thiele, 2004). Interestingly, a putative TCTP interaction protein "putative translational elongation factor eEF-1 Alpha chain" was also up-regulated and had a very similar up regulation profile as TCTP with a maxima at 24 hours (Table 4.1) , suggesting the possible involvement of this pathway in cell dedifferentiation. Tumourigenesis requires partial cell dedifferentiation to enable a cell to reenter the cell cycle, and cell dedifferentiation has been shown to be involved in carcinogenesis (Sell, 1993). Therefore, the involvement of TCTP in plant cell dedifferentiation is highly interesting. Understanding the molecular mechanism of TCTP action during cell dedifferentiation may provide insight into the general mechanism of cell dedifferentiation.

Stress response proteins: The relationship between cell dedifferentiation and stress response is unknown. L-ascorbate peroxidase (cytosolic), putative L-ascorbate peroxidase (Chloroplast), and late embryogenesis abundant domain-containing protein were down-regulated during the induction of cell dedifferentiation (Table 4.1). Meanwhile, the nonsymbiotic hemoglobin 1 (Hb1), Glycine-rich RNA binding protein (Clone A81), and spermidine synthase were up-regulated. The late embryogenesis abundant domain containing protein was also stained by Pro-Q Diamond dye. Since the above proteins are either directly or indirectly related to cell stress responses, the up and down regulation of different sets of stress related proteins indicates that the cell dedifferentiation process exposes the cells to a different type of stress compared with the seedlings.

GTP binding protein: GTP-binding nuclear protein RAN-2, a putative GTP binding protein, and GTP-binding protein SAR1B were up-regulated during cell dedifferentiation (Table 4.1). These proteins are involved in diversified cellular processes. Up regulation of these proteins indicate that cell dedifferentiation involving broad cellular processes.

Hypothetical and Expressed proteins*:* Among the differentially regulated proteins, the hypothetical proteins and function unknown proteins were a major group, including AT2G39730, AT1G73850, AT4G36105, AT1G52100, AT4G05190, and AT3G25770. The differential regulation of these proteins is consistent with the fact that the mechanisms of cell dedifferentiation are still poorly explored. Further investigation on the biological functions of these genes could significantly contribute to the research on dissecting cell dedifferentiation pathways.

CHAPTER V

PHOSPHOPROTEOME DIFFERENTIAL EXPRESSION DURING CELL DEDIFFERENTIATION IN *ARABIDOPSIS THALIANA*

Abstract

Phosphoproteome differential expression during cell dedifferentiation has been investigated in *Arabidopsis* using cotyledon as a starting material. Among the 118 putative phosphoprotein spots, 53 proteins have been identified with a Confidence Interval (C.I. %) over 95% additionally 19 proteins identified with a Confidence Interval less than 95% using MALDI-TOF/TOF. Among these 53 high confident proteins nine protein spots displayed differential regulation after initiating the cell dedifferentiation when examined using Pro-Q Diamond Phosphoprotein Gel stain following 2-D gel separation. Among the 53 proteins 29 proteins were found with phosphopeptides, suggesting that Pro-Q Diamond Phosphoprotein Gel Stain is a useful tool in exploring phospho- proteome. These studies have provided significant new insight into phosphoprotein differential expression during cell dedifferentiation in plants.

Introduction

Plant cotyledon is a specialized organ with terminally differentiated cell fate. Differentiated cotyledon cells can dedifferentiate and regenerate efficiently under appropriate stimuli. Therefore, it is an ideal material for cell dedifferentiation studies. The completion of genome sequencing in both *Arabidopsis* and rice and the advances in genomic and proteomic technologies have significantly facilitated the studies of plant response to various stimulus.

Studies in Molecular Biology, Biochemistry, and Genetics have indicated that protein phosphorylation plays a critical role in plant response to hormonal stimulus. However, phosphoprotein differential regulation during cell dedifferentiation has not been well studied using proteome approaches in *Arabidopsis*. Pro-Q Diamond in Gel Stain has been used extensively to identify phosphoproteins following 2-DE gel separation due to its specificity to phosphoproteins (Steinberg et al., 2003; Schulenberg et al., 2004; Agrawal and Thelen, 2005; Stasyk et al., 2005; Agrawal and Thelen, 2006b). Here putative phosphoprotein differential regulation during cell dedifferentiation using the Pro-Q Diamond dye in *Arabidopsis* has been investigated. The majority of proteins stained heavily by Pro-Q Diamond dye have a relatively low pI. This observation is consistent with the fact that protein phosphorylation results in acidification of the protein. In this study several phosphopeptides for the identified putative phosphopeptides have been successfully identified, suggesting that the Pro-Q Diamond in gel Stain is a useful tool in identifying putative phosphoproteins. It is also found that using an appropriate concentration of Pro-Q Diamond dye is critical for reducing non specific background. Direct visualization of the putative phosphoprotein spots using Pro-Q Diamond dye

enables us to select the potentially interesting proteins for mass analyses. However, the limited protein quantity, from an excised protein spot of 2-DE gel, makes mapping the phosphorylated residue(s) extremely challenging. Recently, several phosphoproteome analysis methods have been developed, including Titanium Dioxide microcolumn (Larsen et al., 2005), immobilized metal-affinity chromatography method (Ficarro et al., 2002), phosphopeptide enrichment by IEF (Maccarrone et al., 2006), and the phosphoprotein extraction kit of QIAGEN (Jones et al., 2006a). The combination of the 2-DE gel based Pro-Q Diamond stain and the phosphopeptide enriching microcolumn may help mapping the phosphorylation sites of the putative phosphoproteins in the future.

Using Pro-Q Diamond Phosphoprotein gel stain, one can follow quantitative changes of particular protein spots on 2-DE gels in a time course under a treatment or at different developmental stages. It is a useful tool in revealing differential regulation of phosphoproteins.

In this study, the protein differential regulation in the time course of cellular dedifferentiation has been investigated in *Arabidopsis thaliana* cotyledons using the Pro-Q Diamond Phosphoprotein Gel Stain. This study mailnly focused on protein differential regulation within the first 48 hours of time after inducing the dedifferentiation process with both 2.4.D and Kinetin. Nine differentially regulated putative phosphoproteins have been identified among 53 high confident putative phosphoproteins by Pro-Q Diamond stain followed by mass spectrometry analyses.

Materials and Methods

Cotyledon collection and induction of cell dedifferentiation, Protein extraction, 2D PAGE analysis were carried out as explained in materials and methods section of Chapter 4.

Identification of phosphoproteins using Pro-Q Diamond Phosphoprotein Gel Stain

Detection of phosphoproteins after separation on 2-DE gels followed the instructions from the manufacturer except that only half of the recommended concentration was used (Molecular Probes). In brief, 2-DE gels were fixed in solution containing 50% methanol and 10% acetic acid, washed with several changes of water to remove SDS, and stained with the Pro-Q Diamond dye. After destaining, the gel images were recorded using VersaDoc4000 (Bio-RAD). The spots consistently stained with Pro-Q Diamond dye in all three biological replicas were considered as putative phosphoproteins. The proteins with two fold differential regulation in all three biological replicas were considered as differentially regulated putative phosphoproteins. Protein gels of the same protein samples were also stained with SYPRO Ruby to correlate the protein spots revealed by these two different dyes.

In gel digestion and Mass Spectrometry, Gene Ontology analysis were carried out as explained in materials and methods section of Chapter 3.

Results

Reversible protein phosphorylation plays an essential role in multitude biological processes in plants, including developmental regulation and hormone responses. To investigate the role of protein phosphorylation in cell dedifferentiation, the cotyledon phosphoproteome has been examined using Pro-Q Diamond Phosphoprotein in Gel Stain and performed MS/MS analysis of the stained proteins using a MALDI-TOF/TOF mass spectrometer. The Pro-Q Diamond Phosphoprotein in Gel Stain is a phosphoprotein stain that has been used widely in detecting phosphoproteins (Steinberg et al., 2003; Stasyk et al., 2005; Agrawal and Thelen, 2006a). As shown in Figure 5.1, the majority of the Pro-Q Diamond dye stained protein spots had a relatively low pI, and it was very obvious for those heavily stained protein spots. In contrast, the SYPRO Ruby stained proteins were more evenly distributed on the 2-DE gel (Figure 5.1B). Many protein spots that were heavily stained with SYPRO Ruby were not stained or weakly stained by Pro-Q Diamond dye. Conversely, many spots that were weakly stained by SYPRO Ruby were heavily stained by Pro-Q Diamond dye. Since SYPRO Ruby stain intensity has a linear correlation to protein concentration in a broad range (Berggren et al., 1999), the results suggest that Pro-Q Diamond dye stain intensity is not proportional to protein concentration. One possibility is that the stain intensity of the Pro-Q Diamond dye is proportional to the level of phosphorylation although further tests are required to confirm this. After fixation, the 2-DE gels were stained with Pro-Q Diamond dye as instructed by the manufacturer. The 2-DE gel images were acquired using VersaDoc4000. As shown in Figure 5.1A, about 103 putative phosphoprotein spots were detected using the Pro-Q

Diamond in gel stain. A section of the image was enlarged and presented in Figure 5.1B and the SYPRO Ruby stain of the corresponding region was shown in Figure 5.1C.

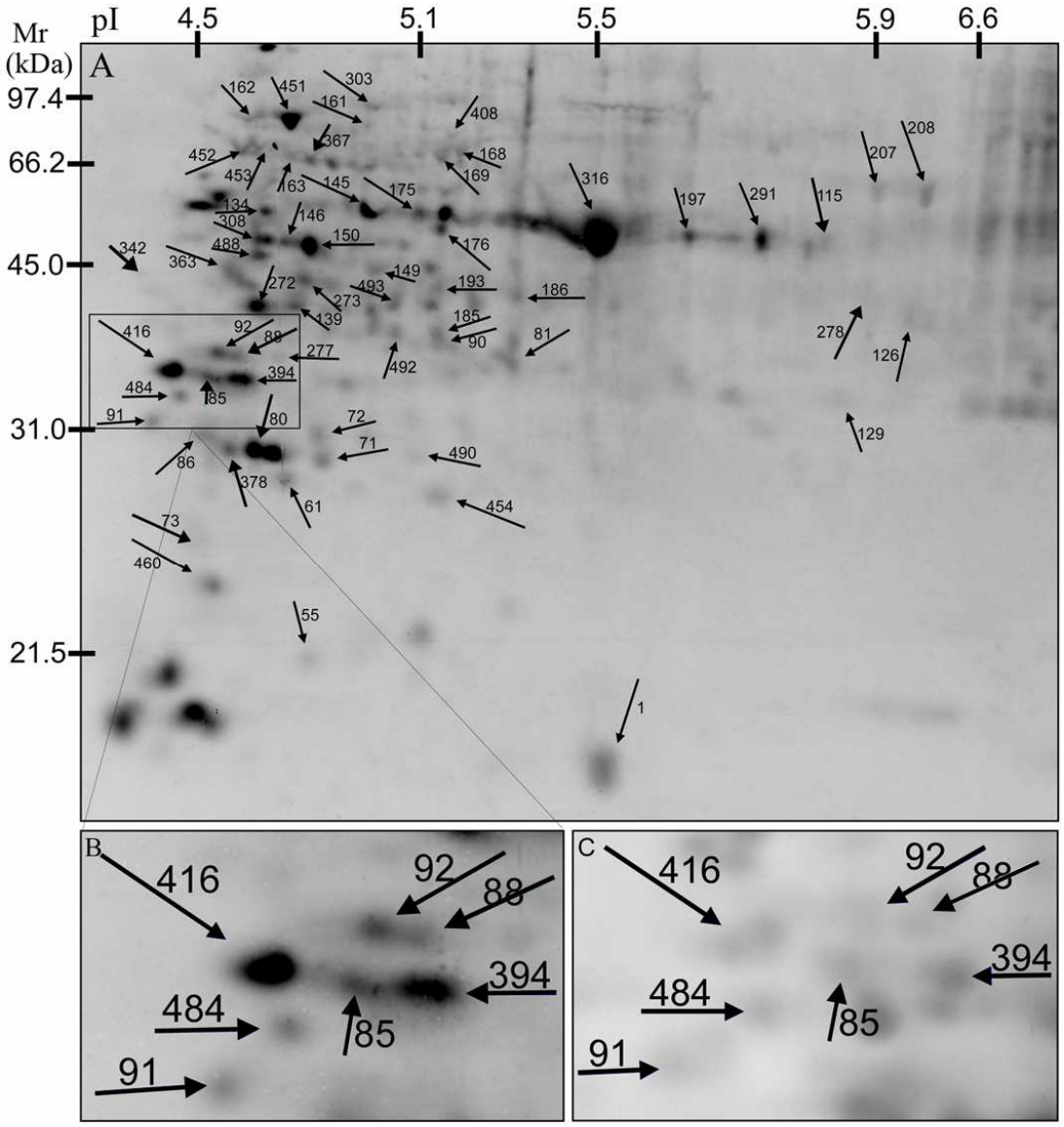

Figure 5.1: 2-DE gel image of putative phosphoproteins revealed by Pro-Q Diamond Phosphoprotein in Gel Stain.

Note: The first dimension pH gradient is indicated on the top and the molecular mass (KDa) are indicated on the left. Proteins identified with high confidence are marked by arrows. The spot numbers are correlated with those used in Figure 3.1. **A)** 2-DE phosphoprotein image revealed by Pro-Q Diamond dye. **B)** An enlarged section of the phosphoproteome image stained by Pro-Q Diamond dye. **C)** SYPRO Ruby stain image of the enlarged region corresponding to **B**.

As revealed via comparison between 5.1C and 5.1B, the Pro-Q Diamond stain displayed high affinity to some protein spots but did not stain with other protein spots, indicating the specificity of the dye. Mass spectrometry analyses identified 53 proteins with a Confidence Interval (C.I. %) over 95% as shown in Table 5.1 and annotated additionally 19 proteins with a Confidence Interval less than 95%. Other protein spots failed to provide enough mass information for protein identification. As shown in Table 5.2 and Figure 5.2, 9 of the 53 (17%) identified putative phosphoproteins were differentially regulated during cell dedifferentiation based on the criteria of two fold change in all three biological replicas. Among the 500 identified proteins, only 53 (11 %) were stained by Pro-Q Diamond dye in the 2-DE gels. The percentage was lower than the estimated 30% in other organisms (Ficarro et al., 2002). For many of the proteins stained with Pro-Q Diamond dye, it has been found literatures reporting those proteins or proteins in the same protein families to be phosphorylated. Of the 9 differentially regulated putative phosphoproteins, for example, 6 of them or their family members have been reported to be regulated by phosphorylation as indicated in Table 5.2. Of the other three proteins, one was a functionally unknown protein.

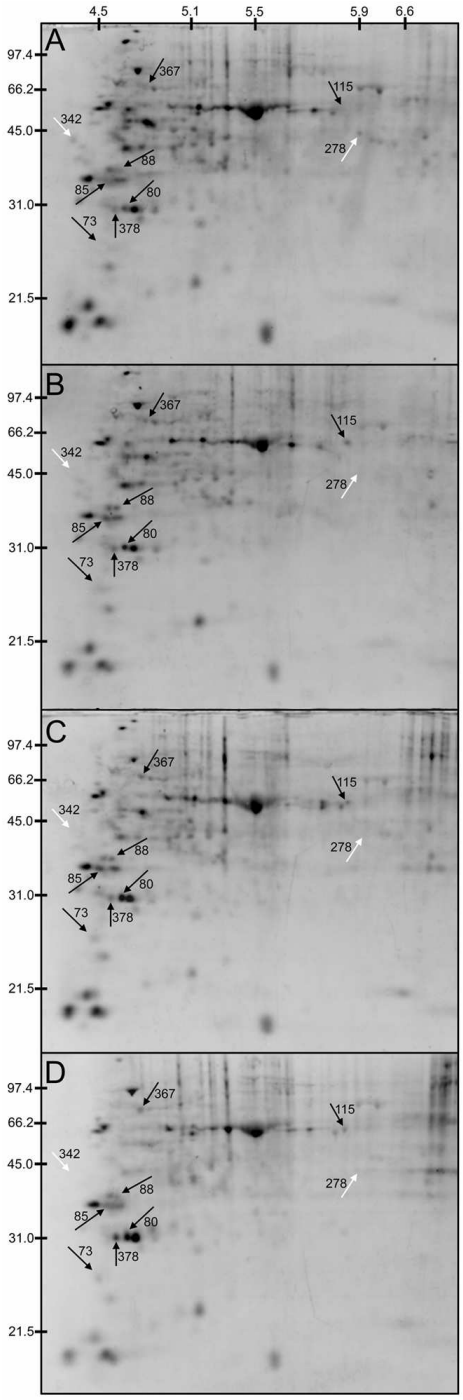

Figure 5.2: 2-DE gel image of differentially regulated putative phosphoproteins revealed by Pro-Q Diamond Phosphoprotein in Gel Stain.

Note: Proteins identified with high confidence are marked by arrows. The spot numbers are correlated with those used in Fig. 1. (A) 2-DE phosphoprotein image revealed by Pro-Q Diamond dye. (B) An enlarged section of the phosphoproteome image stained by Pro-Q Diamond dye. (C) SYPRO Ruby stain image of the enlarged region corresponding to (B).

Table 5.1: Putative phosphoproteins identified by Pro-Q Diamond in Gel Stain

Spot #	Protein Identification	Gene ID	Putative Phosphopeptides
1	RuBisCO small subunit 2B (EC 4.1.1.39)	AT5G38420	1. YWTMWK - (Y)[1]
55	Putative peroxiredoxin protein	AT3G52960	1. SFATTPVTASISVGDK - (S/T)[1], (S/T)[4,5,8]
61	Hypothetical protein	AT3G05625	1. EAYNLFK - (Y)[3]; 2. FTAISPFK - (S/T)[2]; 3. RFTAISPFK - (S/T)[3]
73	Translationally controlled tumor like protein	AT3G16640	1. GFIAYIKK - (Y)[5]
80	Chlorophyll A-B binding protein 2, (CAB-140) (LHCP)	AT1G29910	
81	Fructose-bisphosphate aldolase, putative	AT4G38970	1. TYDVAEK - (Y)[2]; 2. TVGGGGNTQLER - (S/T)[1]; 3. ASTSLLKASPVLDK - (S/T)[2,4], (ST)[3]; 4. GQSVLFRQPSSASVVLR – (S/T)[3,10,11]
85	F23J3_30 (14-3-3 protein GF14chi)	AT4G09000	1. GNDDHVSLIR - (S/T)[7]
86	Putative RNA-binding protein	AT2G37220	
88	T32G9_30	AT1G35160	
90	Putative fructose bisphosphate aldolase	AT2G21330	1. TYDVAEK - (Y)[2]; 2. YTGEGESEEAK - (Y)[1]
91	Nascent polypeptide associated complex alpha chain	AT3G12390	
92	Disease resistance protein (TIR class), putative	AT4G19920	
115	F26F24_4	AT1G70580	
126	Putative glyceraldehyde-3-phosphate dehydrogenase (F13B4_8)	AT1G13440	1. VPTVDVSVVDLTVR - (S/T)[7], (S/T)[3]; 2. VPTVDVSVVDLTVR - (S/T)[7], (S/T)[3,12]
129	30S ribosomal protein S5	AT2G33800	1. SAIDARR - (S/T)[1]; 2. SSGIGGKPKK - (S/T)[1]; 3. SSGIGGKPKK - (S/T)[1,2]; 4. FSSLTLVK - (S/T)[2,3], (S/T)[5]; 5. RNIVQVPMTK - (S/T)[9]; 6. SIVSFSSFLNR - (S/T)[1]
139	Sedoheptulose-1,7-bisphosphatase, (EC 3.1.3.37)	AT3G55800	1. LAPKSQLK - (S/T)[5]; 2. GIFTNVTSPTAK - (S/T)[8], (S/T)[4,7]; 3. SSTLVSPPSYSTSSSFK – (S/T)[1,2], (S/T)[3],(Y)[10]
145	Glutamate--cysteine ligase, (EC 6.3.2.2) (Gamma-ECS)	AT4G23100	
146	Phosphoglycerate kinase (EC 2.7.2.3) (MBK21_14)	AT3G12780	1. GVVSMAKK - (S/T)[4]; 2. LSELLGIEVTK - (S/T)[2], (S/T)[10]; 3. MASAAASSAFSLLK - (S/T)[3]; 4. VSSKIGVIESLLEK - (S/T)[2]
149	Phosphoribulokinase (EC 2.7.1.19 (PRKASE) (PRK)	AT1G32060	
161	Putative elongation factor G protein	AT1G62750	
162	Heat shock protein 70	AT5G49910	1. RDAIDTK - (S/T)[6]; 2. IASGSTQEIK - (S/T)[3,5]; 3. HIETTLTRGK - (S/T)[4]; 4. VVDWLASTFK - (S/T)[7], (S/T)[8]
163	Dihydroxyacetone kinase -related	AT3G17770	1. DSLNEWDGK - (S/T)[2]; 2. LTSQRFLTK - (S/T)[2]; 3. IPVPVPPSRSIK - (S/T)[8,10]

Table 5.1. Continued

Spot #	Protein Identification	Gene ID	Putative Phosphopeptides
168	Transketolase-like protein	AT3G60750	1. GSLPAFSGLK - (S/T)[2,7]; 2. IAQSMTKNR - (S/T)[4], (S/T)[6]; 3. VSIEAASTFGWGK - (S/T)[2]; 4. AFGDFQKATPEER - (S/T)[9]; 5. EAKTVTDKPTLIK - (S/T)[4,6]; 6. AISHHGSDQRGSLPAFS – (S/T)[3,7]
169	GloEL protein, chaperonin, 60 kDa	AT3G13470	1. NVVLESK - (S/T)[6]; 2. DGTTIRK - (S/T)[3]; 3. LADLVGVTLGPKGR - (S/T)[8]
175	Eukaryotic initiation factor 4A-2 (eIF4A-2)	AT1G54270	1. KFMSKPVR - (S/T)[4]; 2. DELTLEGIK - (S/T)[4]
176	Eukaryotic initiation factor 4A-1 (eIF4A-1)	AT3G13920	1. KFMSKPVR - (S/T)[4]; 2. DELTLEGIK - (S/T)[4]
185	Putative lectin	AT3G16420	
186	Proline iminopeptidase	AT2G14260	
193	T5I7.3 (Hypothetical protein)	AT2G39730	
207	F18O14.27	AT1G19520	
208	F20D10_50 (EC 2.1.2.1) (Serine hydroxymethyltransferase)	AT4G37930	
272	Fructokinase, putative	AT1G66430	1. SAHISAAK - (S/T)[1]; 2. YYTKDFSGR - (S/T)[3], (Y)[1,2]; 3. LLLVTEGPEGCRYYTK - (S/T)[5], (Y)[13,14]
273	Sedoheptulose-bisphosphatase	AT3G55800	
277	Expressed protein (Hypothetical protein)	AT2G44040	
278	ATP synthase gamma chain 1, (EC 3.6.3.14)	AT4G04640	1. LTTKEGK - (S/T)[2] 2. YLEAGTLPTAK - (S/T)[6], Y[1]
303	Cell division control protein 48 homolog E (AtCDC48e)	AT5G04590	1. SNLRVR - (S/T)[1]; 2. SIGVKPPK - (S/T)[1]; 3. GGMRSVEFK - (S/T)[5]; 4. RSVSDADIR - (S/T)[2,4]; 5. SPVAKDVDVTALAK - (S/T)[1]
308	Hypothetical protein (Putative ribosomal protein S1)	AT5G30510	
342	Late embryogenesis abundant domain-containing protein	AT5G44310	1. DSRADLAYDSK - (S/T)[2,10], (Y)[8]
363	Gb|AAF00675.1 (expressed protein)	AT3G20350	
367	Cell division protein ftsH homolog 2, (E.C.3.4.24.)	AT5G42270	1. QVTVDRPDVAGRVK - (S/T)[3]; 2. KPSFPFSFVSRAK - (S/T)[3,7]; 3. DYSMATADIVDAEVRELV - (Y)[2]
378	Putative photosystem II type I chlorophyll a/b binding protein	AT2G34430	1. LSPAASEVFGTGRITMR - (S/T)[2,6]
393	Inorganic pyrophosphatase-like protein (EC 3.6.1.1)	AT5G09650	
408	Jasmonate inducible protein ISOLOG (Putative lectin protein)	AT3G16460	1. FDYVKGGVTK - (S/T)[9]; 2. IYASYGGEGIQYVK - (Y)[2,5]; 3. IYASYGGEGIQYVK - (S/T)[4], (Y)[2]; 4. TAGPFGIVSGTKFEFK - (S/T)[1]
416	Hypothetical protein	AT5G16110	
452	Hypothetical protein	AT1G25682	
453	Heat shock protein 70	AT5G49910	1. VVDWLASTFK - (S/T)[7], (S/T)[8]; 2. EKIASGSTQEIK - (S/T)[5,7]
454	Oxygen-evolving enhancer protein 2-1	AT1G06680	1. MAYSACFLHQSALASSAAR - (S/T)[4,11], (Y)[3]
484	Elongation factor 1B alpha-subunit	AT5G19510	

Table 5.1. Continued

Spot #	Protein Identification	Gene ID	Putative Phosphopeptides
488	F5J5.1	AT1G36035	
490	Triose-phosphate isomerase (EC 5.3.1.1), cytosolic	AT3G55440	1. AAEVNKSA - (S/T)[7]
492	Putative lectin	AT3G16420	1. NGQPEQAPLRGTK - (S/T)[12]
493	Putative fructose bisphosphate aldolase	AT2G21330	1. TYDVAEK - (Y)[2]; 2. ALQNTCLK - (S/T)[5]; 3. ASSTATMLK - (S/T)[2,3], (S/T)[4,6]; 4. ATPEQVASYTLK - (Y)[9]
460	Plastid ribosomal protein S6, putative	AT1G64510	

a) Spot #: The spot numbers were based on the order of mass analyses (corresponding to Table 1 in Chapter 3). The corresponding Pro-Q Diamond stain image of these spots is shown in Figure 5.
b) Putative Phosphopeptides: Peptide whose mass matched with the calculated phosphopeptide mass in mass fingerprinting.

Table 5.2: Differentially regulated putative phosphoproteins in the time course of cell dedifferentiation

Spot #	Protein Identification	Expres Graph	Gene ID	Putative Phosphopeptides	Reference
73	Translationally controlled tumor like protein		AT3G16640	1. GFIAYIKK - (Y)[5]	Yarm, 2002
80	Chlorophyll A-B binding protein 2, (CAB-140) (LHCP)		AT1G29910		Liu and Shen, 2005
85	F23J3_30 (14-3-3 protein GF14chi)		AT4G09000	1. GNDDHVSLIR – (S/T)[7]	Aitken, 1995
88	T32G9_30		AT1G35160		Aitken, 1995
115	F26F24_4		AT1G70580		
278	ATP synthase gamma chain 1, (EC 3.6.3.14)		AT4G04640	1. LTTKEGK – (S/T)[2] 2. YLEAGTLPTAK – (S/T)[6], Y[1]	
342	Late embryogenesis abundant domain-containing protein		AT5G44310	1. DSRADLAYDSK – (S/T)[2,10], (Y)[8]	Jones, 2006
367	Cell division protein ftsH homolog 2, (E.C.3.4.24.)		AT5G42270	1. QVTVDRPDVAGRVK - (S/T)[3]; 2. KPSFPFSFVSRAK - (S/T)[3,7]; 3. DYSMATADIVDAEVRELV - (Y)[2]	
378	Putative photosystem II type I chlorophyll a/b binding protein		AT2G34430	1. LSPAASEVFGTGRITMR - (S/T)[2,6]	Larsson, 1983

Note: Spot #: The spot number is identical with Table 1(Chapter 3). Exprees Graph: the four bars in the expression graph represent the relative expression level of the protein after 0, 12, 24, and 48 hour induction. Putative Phosphopeptides: Peptides whose mass matched with theoretical phosphopeptide mass in mass fingerprinting. References: literature that reports phosphorylation of the protein or its family members.

To understand the biological processes and cellular components involved in cellular dedifferentiation in perspective of identified phosphoproteome in *Arabidopsis,* Gene Ontology analyses were performed using the all identified phosphoproteins, including both differentially up- and down-regulated proteins. As shown in Figure 5.3, the top two cellular components for the identified proteins were Chloroplast (20.5%), Other cellular components (17.9%). The Top two biological processes for the identified proteins were Other metabolic processes (27.3%) and Other cellular processes (21.8%). When analyzed based on molecular functions, the top two categories for the Pro-Q Diamond stain identified proteins were Other molecular functions (19.3%) and Other enzyme activity (17.5%). The detailed distribution of these putative phosphoproteins is clearly depicted in Figure 5.3. These results indicated that diversified cellular processes and cellular components are involved in the process of cell dedifferentiation. However, we should keep in mind that one protein could be grouped into more than one GO category sometimes. Therefore, the total percentage in each pie is over 100% when added together. In addition, the current GO system has not included all possible functions for proteins with multiple functions. Therefore, the GO analysis results presented here should be taken with above information in mind.

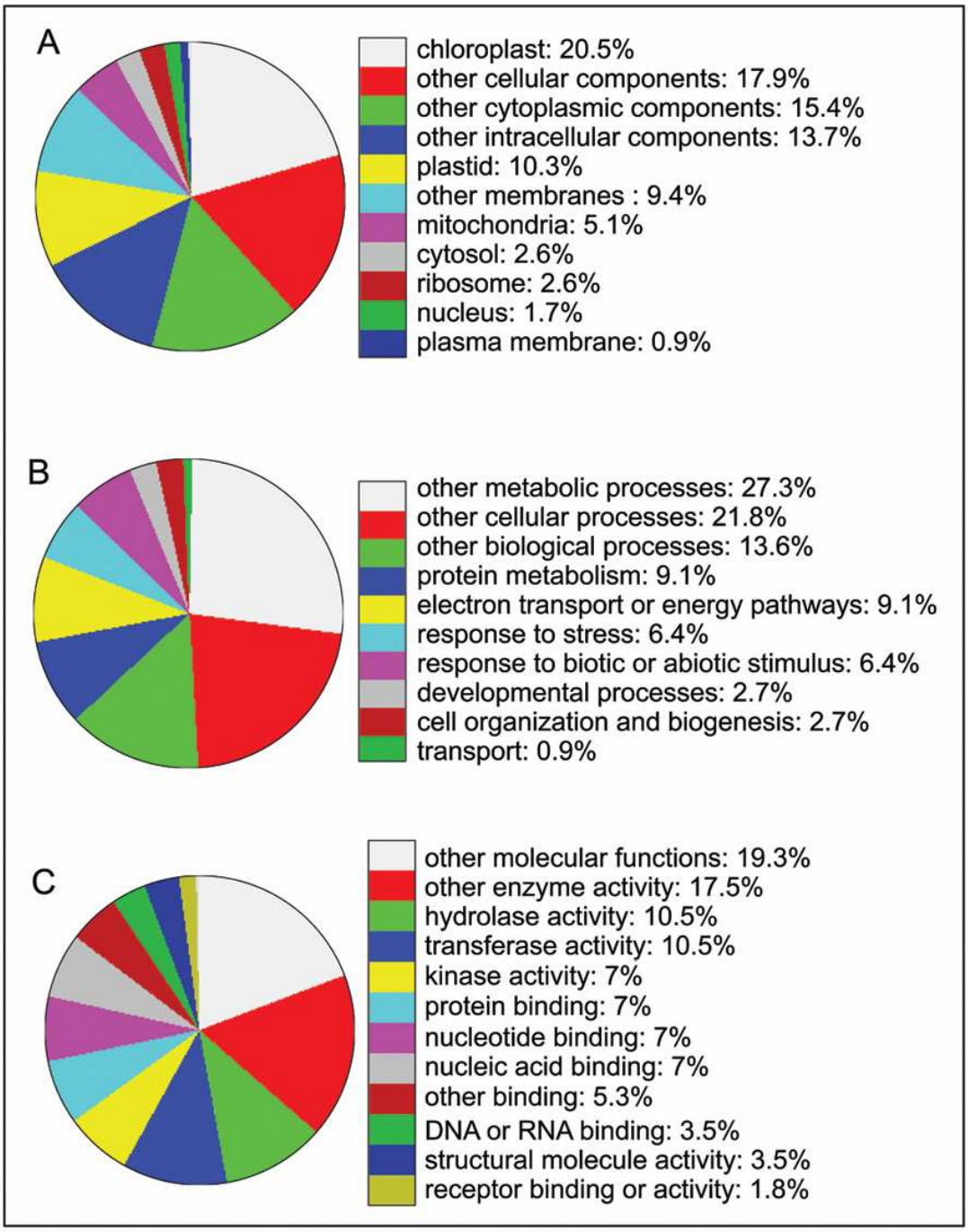

Figure 5.3: Distribution of the differentially regulated phosphoproteins among different GO categories.

Note: (A) Cellular localization, (B) Biological process, (C) Molecular function.

Discussion

To study the phosphoproteome in cotyledons, the 2-DE gels were stained with a Pro-Q Diamond Phosphoprotein in Gel Stain. The Pro-Q Diamond dye has been widely used to selectively stain the phosphoproteins on 2-DE gel (Steinberg et al., 2003). One significant advantage of this method is that one can follow the quantitative and qualitative changes of a particular putative phosphoprotein spot in a time course and under different treatments. Special attention can be applied to particular protein spots that are interesting. Our results show that the Pro-Q Diamond in gel stain displays good specificity to about 17 % protein spots. Among 53 putative phosphoproteins identified with high confidence, most of them have been reported that either the proteins themselves or their close family members are regulated by phosphorylation in published literatures. Among the 9 differentially regulated and Pro-Q Diamond dye stained proteins, for example, 6 of them or their protein family members have been reported to be phosphorylated. One of the three proteins that have no reference about phosphorylation is a functionally unknown protein. These results substantiate that the Pro-Q Diamond in Gel Stain can be used to identify phosphoprotein candidates on 2-DE gel effectively and it provides significant new insight into the role of few cotyledon phosphoproteins during cell dedifferentiation process in plants.

CHAPTER VI

PROTEOME DIFFERENTIAL REGULATION DURING CELL DEDIFFERENTIATION IN *ARABIDOPSIS*: COMPARISON OF THE MudPIT AND 2-DE GEL BASED METHODS

Abstract

Cell dedifferentiation is a cell fate switching process in which differentiated mature cells undergo genome reprogramming to regain the competency of cell division and organ regeneration. The molecular mechanism underlying the cell dedifferentiation process is basically unknown. It is believed that proteins specific to the original cell fate has to be removed and proteins required for new cell fate are synthesized timely during cell dedifferentiation. *Arabidopsis* cotyledon proteome at four different time points after the induction of cell dedifferentiation with MudPIT approach has been investigated and analyzed the protein quantity change using two label-free quantitation methods, the Spectral Count (SC) and SEQUEST Cross Correlation Coefficient (ΣXcorr) methods. Among the 662 MudPIT identified proteins, one hundred forty eight display differential regulation. The up-regulated proteins include transcription factors, calmodulins, translational regulators, and stress response proteins. The Spectral Count and the cross correlation coefficient quantification results are highly consistent in over 81% of the differentially regulated proteins. In addition, the 2-DE gel based quantification results are

significantly different from or controversial to the results of the label-free quantitation methods in over half of the differentially regulated proteins commonly identified. Further analysis finds that most of the proteins with discrepancy have been identified in more than one protein spots on 2-DE gels and some have been shown to be phosphoproteins, suggesting that the 2-DE gel based quantification reflects the change of a specific modification form of a protein instead of the total of that protein. Our results indicate that the MudPIT and 2-DE gel approaches complement each other in both protein identification and quantification.

Introduction

The remarkable cell dedifferentiation and regeneration abilities fascinated biologists in both plants and animals for centuries. However, little is known of the molecular mechanism underlying this biological process. In plants, it has been demonstrated that phytohormones play a critical role in cell dedifferentiation and organ regeneration. The newly acquired cell division activities of dedifferentiated cells usually lead to the formation of amorphous, relatively undifferentiated cell mass called callus in plants (Bhojwani and Razdan, 1983). Although the properties of callus vary considerably according to the callus induction condition and the genetic and epigenetic nature of the explants, cell dedifferentiation has to occur before the formation of visible callus. The role of auxin and cytokinin in callus induction and growth has been discovered over half a century (Gautheret, 1985). Studies, which were initiated in tobacco pith tissue culture and confirmed in other organisms, have demonstrated that an appropriate cytokinin/auxin ratio in the medium is critical to the cell fate in culture (Krikorian, 1995). The absolute

concentrations of the two hormones are not critical, providing that they are within an appropriate concentration range (Skoog and Miller, 1957). When the ratio of cytokinin/auxin is about the same, a callus mass is usually produced from explants and the callus in culture can proliferates continuously. When the level of cytokinin relative to that of auxin is low, roots regenerate from the tissues; when cytokinin relative to that of auxin is high, shoots regenerate.

Recently, it has been shown in tobacco that acquisition of competence for cell fate switch correlates with chromatin reorganization, redistribution of HP1 protein, and increase in acetylated histone H3 at K9 and K14 (Williams et al., 2003). Meanwhile, hypomethylation of the ribosomal RNA genes is found to be associated with cell dedifferentiation in tobacco cells (Koukalova et al., 2005). In addition, it has been known that the ubiquitin proteolytic pathway is indispensable for protoplast progression into S phase. A specific inhibitor of the 26S proteasome, MG132, interferes with the entry of protoplasts into S phase (Zhao et al., 2001). In agreement, it has also been reported that genes coding for ubiquitin are induced during cell dedifferentiation in tobacco (Jamet et al., 1990).

Another research group (Che et al., 2006) studied the change of gene expression profiles during shoot, root, and callus development and found that over 21% of the genes displayed over two fold differential regulation in *Arabidopsis* roots after four days of callus induction. Previous studies (Chitteti and Peng, 2007) had shown the proteome and phosphoproteome dynamic change in *Arabidopsis* within the first 48 hours of callus induction by using the 2-DE gel based comparative proteomic approach. Among the 500 proteins identified in *Arabidopsis* cotyledons, 42 proteins displayed over two folds of

differential regulation and nine proteins displayed differential phosphorylation. These studies indicate that both transcriptional and posttranslational regulations are involved in the regulation of cell dedifferentiation. Cellular protein concentration and modification status mainly determine the activity of cells. In the cell fate switch during cell dedifferentiation, the proteins exclusive to the prior cell status have to be removed and proteins required for the dedifferentiated status have to be synthesized accordingly. Therefore, a thorough understanding of protein differential regulation following the time course of cell dedifferentiation is vital to our understanding of the mechanisms underlying cell dedifferentiation. Studies have shown that cell dedifferentiation can occur rapidly when an appropriate environment is provided. Experimental results from *Arabidopsis* cotyledon protoplast culture revealed that the first cell division has occurred within 24 hours while the majority of the cells divide between 48 to 72 hours (Dovzhenko et al., 2003). Since cells have to go through dedifferentiation before re-entering the cell cycle, these observations suggest that cell dedifferentiation can occur within 24 hours in *Arabidopsis*.

Recently, label-free quantification methods, coupled with shotgun proteomics, have shown great potential in comparative proteomic studies. The label-free methods are high throughput and completely eliminate the labor intensive 2-DE gel separation and sample labeling steps. In addition, it can separate very wide range of proteins and overcome the protein solubility problem that is often encountered in 2-DE method. The label-free methods include peptide counts, sequence coverage, peak area intensity measurements, spectral counts, and the sum of the SEQUEST cross correlation coefficient of a protein (ΣXcorr). The Spectral Count, the total number of MS/MS spectra taken on peptides from

a given protein in LC/LC-MS/MS analysis, is linearly correlated with the protein abundance over a dynamic range of 2 orders of magnitude (Old et al., 2005; Zhang et al., 2006; Schmidt et al., 2007). It has been reported that the Spectral Count method provides more sensitivity in detecting proteins undergo change in abundance and the results match well with 1-D gel staining intensities (Old et al., 2005). The SEQUEST algorithm has been widely used in peptide and protein identification. It calculates a cross correlation to quantitatively measure the relatedness of experimental mass spectra to the in silico generated tandem mass spectra based on protein sequence. The cross correlation coefficient is determined by factors including the number of fragment ions in the mass spectrum, their relative abundance, continuity of ion series, and presence of immonium ions for certain amino acids in the spectrum, that are proportional to the concentration of the precursor ion. It is found that the sum of SEQUEST cross correlation coefficient (ΣXcorr) correlated well with the concentration of tested proteins (Nanduri et al., 2005). Because SEQUEST is widely used in peptide and protein identification, using the sum of cross correlation coefficient score for protein quantification is easy to implement and should have broad applications.

It is reported that proteins identified are dependent on the mode of ionization used for peptide analysis (Bodner et al., 2003). They splitted a complex peptide mixture and analyzed the mixture with LC/ESI/MS/MS and LC/MALDI/MS/MS. While 63% overlap was observed, unique peptides and thus unique proteins were observed by each method. Previously *Arabidopsis* cotyledon proteome using 2-DE gel separation followed with MALDI/MS/MS has been studied (Chitteti and Peng, 2007). Here *Arabidopsis* cotyledon proteome using LC/LC/MS/MS has been investigated and examined the protein

differentiation regulation during cell dedifferentiation using Spectral Count and SEQUEST cross correlation coefficient quantitation methods. 662 proteins are identified in cotyledons and 148 proteins displayed differential expression during cell dedifferentiation. The quantification results of the Spectral Count and SEQUEST cross correlation methods match very well in over 81% of the differentially regulated proteins. However, there is only about 21% overlap of the 148 proteins with the proteins identified by 2-DE-MALDI-MS/MS. In addition, over half of the differentially regulated proteins identified by 2-DE gel stain do not match the label-free quantification results because the 2-DE method reflects the change in some specific modification forms of a protein instead of the total of the protein. Our results indicate that the shotgun approach and the traditional 2-DE gel method complement each other in both protein identification and quantification.

Materials and Methods

Cotyledon collection and induction of cell dedifferentiation, Protein extraction were carried out as explained in materials and methods section of chapter four.

Sample preparation and LC/LC-MS/MS Mass analysis

The dried pellets of protein samples were dissolved in 50µl of 6M urea with 100 mM Tris (pH 7.8). The samples were centrifuged at 16,000g for 10 minutes. The supernatant was quantified using Bio-Rad Rc Dc kit according to the instructions provided by the manufacturer. Protein samples (100µg each), from each replica of the treatments, were reduced with 20µl reducing agent (200 mM DTT and 100 mM Tris, pH 7.8) for 1hr at

room temperature, alkylated in dark with 20µl of the alkylating reagent (200 mM iodoacetamide, 100 mM Tris, pH 7.8) for 1 hr, and diluted to a urea concentration of 0.6M, a concentration at which the trypsin retains its activity. Trypsin solution was added to a final ratio of enzyme to substrate at 1/50 and the digestion was carried out overnight (15h) at 37^0C. The reaction was stopped by adding 10µl lysine; adjusted to the pH below 6.0, and vacuum dried to a final volume of 25 µl. The peptides were desalted using a peptide macro trap (Michrom Bioresources, Inc., Auburn, CA) using a protocol provided by the manufacturer and eluted with 0.1% trifluoroacetic acid in 95% acetonitrile. The eluted peptides were vacuum dried to pellet and re-dissolved in a solution with 0.1% formic acid and 5% acetonitrile.

The peptide mixtures were analyzed by a strong cation exchange column (SCX BioBasic 0.32 × 100 mm) followed by RP-LC (BioBasic C18, 0.18 × 100 mm Thermo Hypersil-Keystone, Bellefonte, PA) coupled directly in-line with electro spray ionization ion trap tandem MS (ProteomeX workstation Thermo Finnigan). A flow rate of 3µl/min was used for both SCX and reverse phase columns. For SCX, a salt gradient of 0, 10, 15, 25, 30, 35, 40, 45, 50, 57, 64, 90, and 700 mM ammonium acetate in 5% acetonitrile and 0.1% formic acid was applied. The resultant peptides were loaded directly into the sample loop of the RP-LC column equilibrated with 0.1% formic acid and 5.0% acetonitrile. The peptides were eluted by acetonitrile gradient (in 0.1% formic acid) as follows: 5%-30% for 30 minutes, 30%-65% for 9 minutes, 95% for 5 minutes, 5% for 15 minutes, with a total of 59 minutes elution.

The LCQ Deca XP ion trap mass spectrometer was configured to optimize the duty cycle length with the quality of data acquired by alternating between a single full

MS scan followed by three tandem MS/MS scans on the three most intense precursor masses (as determined by XCALIBER mass spectrometer software in real time) from the full scan. The collision energy was normalized to 35%. Dynamic mass exclusion windows were 2 minutes. In addition, MS spectra for all samples were measured with an overall mass/charge (m/z) range of 200 to 2,000. The mass spectra and tandem mass spectra produced were used to search the *Arabidopsis thaliana* nonredundant protein database (NCBInrPDB) downloaded from the National Center for Biotechnology Information by using TurboSEQUEST, BIOWORKS BROWSER 3.1 SR1 (Thermo Finnigan). TurboSEQUEST cross-correlates experimentally acquired mass spectra with theoretical mass spectra generated in silico. The idealized spectra were weighted with a, b, and y fragment ions. Trypsin digestion was applied to generate the "precursor ions" and the database included mass changes due to cystine carbamidomethylation and methionine oxidation. The peptide (precursor) ion mass tolerance was 2.5 Da, and the fragment ion (MS2) tolerance was 0.2. For a protein to be considered to be identified in the sample, two or more peptides from this protein must meet the following criteria: more than two peptides with X-correlation >1.9 (+1 charge), >2.2 (+2 charge), >3.75 (+3 charge); delta correlation value ≥ 0.1 (Durr et al., 2004).

Protein quantification

The Spectral Count quantification (SC) was essentially the same as described by Zhang et al., (2006). But, in this study a software recently developed by Dr. Bridges group at Mississippi State has been used. The ΣXcorr quantification method was carried out as reported by Nanduri et al., (Nanduri et al., 2005) except that a software, ProtQuant,

specifically developed by the above group was used for the calculation of the sum of SEQUEST cross correlation coefficient (ΣXcorr) to perform the calculation.

Gene Ontology analysis was carried out as explained in materials and methods section of chapter three.

Results

Cotyledon proteome revealed by shotgun proteomics

Cotyledon is an embryonic organ with terminated cell fate. For callus induction from cotyledons, cell dedifferentiation has to occur. In addition, cotyledon can be easily obtained. Therefore, it is an excellent system to study cell dedifferentiation. Since a thorough understanding of the cotyledon proteome is critical to reveal the proteome dynamic change associated with cell dedifferentiation, Here the cotyledon proteome has been examined using a shotgun approach. The protein mixture isolated from 10-day-old cotyledon using the phenol extraction method was digested with trypsin and followed with LC/LC-MS/MS analysis. 662 proteins were identified with high confidence as listed in Table 3.1. These proteins were analyzed following the Gene Ontology rules based on: (i) biological process (BP) in which the gene product participates; (ii) molecular function (MF) that describes the gene product activities, such as catalytic or binding activities, at the molecular level; (iii) the cellular component (CC) where the gene product can be found. Protein distributions within these three categories are shown in Figure 3.3. The top 5 cellular components were other intracellular components (15.5%), other cellular components (15.3%), other cytoplasmic components (15.2%),

chloroplast, and mitochondria. The top 5 biological process components were other physiological processes (21.2%), other metabolic processes (20.6%), other cellular processes (19%), protein metabolism (8.7%), and electron transport or energy pathway (5%). The top five molecular function components were other enzyme activity, structural molecule activity, transferase activity, protein binding, and unknown molecular function.

Protein differential regulation during cell dedifferentiation revealed by ΣXcorr and Spectral Count quantification methods

To gain insight of the cell dedifferentiation process, protein differential regulation following the time course of cell dedifferentiation has been examined using two label-free quantification methods, the ΣXcorr and Spectral Count methods. Ten-day-old cotyledons were excised and treated with 2.0 mg/l 2, 4-D and 0.2 mg/l kinetin for four different time periods as reported previously to induce cell dedifferentiation (Chitteti and Peng, 2007). The cotyledon protein samples, with three biological replicas for each treatment was digested with trypsin and then followed by LC/LC-MS/MS analysis. A software ProtQuant, specifically developed for the calculation of the sum of SEQUEST cross correlation coefficient (ΣXcorr), was used to quantify and compare the relative protein level based on the principle that the ΣXcorr of a protein was proportional to its quantity (Nanduri et al., 2005). Using this quantitation method, it has been found that 84 proteins were up-regulated over two folds and 39 were down-regulated over two folds within the first 48 hours of induction. Meanwhile, the protein quantity change has been examined using the Spectral Count method with an algorithm developed by Dr. Bridges group at Mississippi State. Ninety-four proteins were found to be up-regulated and fifty

proteins were down-regulated. Comparison of the results obtained with these two methods revealed that about 81% of the differentially regulated proteins identified by these two methods were common (Figure 6.1). In addition, the relative protein expression levels quantified by these two methods at the four different time points were also very similar as shown in Tables 6.1 and 6.2. Meanwhile, the Spectral Count quantitation method identified 13 unique up-regulated proteins and 14 unique down-regulated proteins as shown in Figure 6.1. The ΣXcorr quantitation method identified 3 unique up-regulated proteins and 2 unique down-regulated proteins. The differentially regulated proteins include calcium binding protein, oxidative stress response protein, translational regulation proteins, etc.

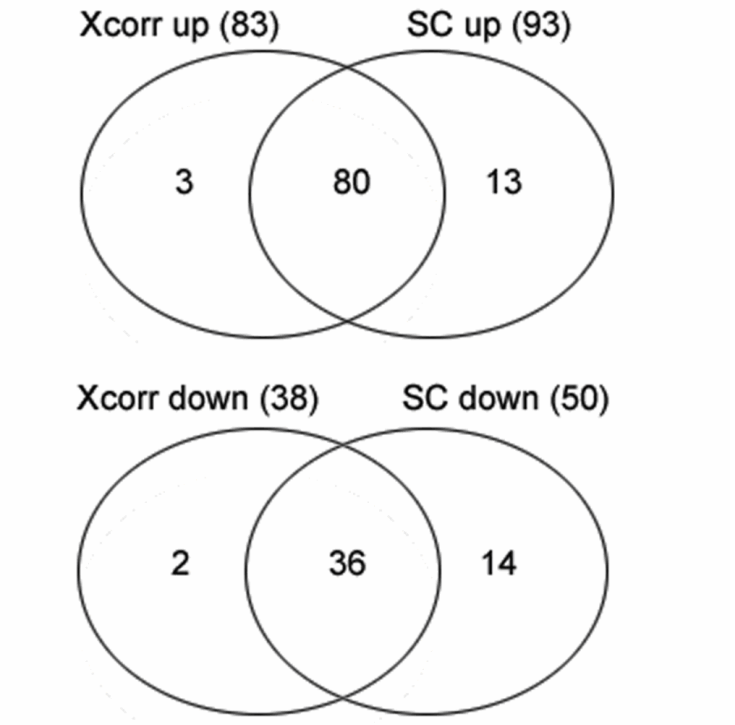

Figure 6.1: Venn diagram representing the identification of differentially regulated proteins using ΣXcorr and Spectral Count method.

Table 6.1: Differentially Up regulated proteins based on Sequest Cross Correlation Coefficient quantitation method

Protein name	0hr	12hr	24hr	48hr	Gene/Locus
LIN2 (LESION INITIATION 2); coproporphyrinogen oxidase	4.44	14.18	14.18	14.18	AT1G03475
Malate dehydrogenase/ oxidoreductase	30.63	71.37	71.37	71.37	AT1G04410
RPL4; structural constituent of ribosome	3.85	15.10	15.10	15.10	AT1G07320
APX1; L-ascorbate peroxidase	4.51	11.98	11.98	11.98	AT1G07890
Phosphoserine aminotransferase	3.62	6.50	6.50	28.62	AT1G08490
Structural constituent of ribosome	14.13	28.46	28.46	28.46	AT1G09690
Putative glyceraldehyde-3-phosphate dehydrogenase	8.97	44.70	44.70	44.70	AT1G13440
ATPDIL1-1; electron transporter/ isomerase/ protein disulfide isomeras	3.17	0.00	5.92	10.95	AT1G21750
Transporter	2.83	17.13	17.13	17.13	AT1G22530
GGT1 (ALANINE-2-OXOGLUTARATE AMINOTRANSFERASE 1)	7.15	14.73	14.73	14.73	AT1G23310
CAB1 (CHLOROPHYLL A/B BINDING PROTEIN 1)	10.39	24.63	24.63	24.63	AT1G29910
Chlorophyll a/b binding protein (LHCP AB 180)	10.39	24.63	24.63	24.63	AT1G29930
T19E23.15	14.13	28.46	28.46	28.46	AT1G31355
Putative lectin	34.21	71.56	71.56	71.56	AT1G45258
Elongation factor EF-2	3.47	16.40	16.40	16.40	AT1G56070
Structural constituent of ribosome	14.13	28.46	28.46	28.46	AT1G57860
ACO2 (ACC OXIDASE 2)	5.73	19.67	19.67	19.67	AT1G62380
Monodehydroascorbate reductase	25.75	52.00	52.00	52.00	AT1G63940
Hydrolase, hydrolyzing O-glycosyl compounds	3.17	21.27	21.27	21.27	AT1G66270
Carbonate dehydratase/ zinc ion binding	2.71	2.88	2.88	8.70	AT1G70410
GLP1 (GERMIN-LIKE PROTEIN 1); nutrient reservoir	8.88	26.02	26.02	26.02	AT1G72610
Unknown protein	2.94	6.56	6.56	6.56	AT1G79040
3-isopropylmalate dehydrogenase/ oxidoreductase	4.07	15.81	15.81	15.81	AT1G80560
Putative Ty3-gypsy-like retroelement pol polyprotein	2.51	2.85	5.48	5.48	AT2G06170
Phosphoserine transaminase/ transaminase	3.62	3.84	3.84	12.61	AT2G17630
EMB2296; structural constituent of ribosome	3.76	10.93	10.93	10.93	AT2G18020
At2g19940/F6F22.3	2.52	0.00	5.20	5.20	AT2G19940
ALDH11A3; aldehyde dehydrogenase/ oxidoreductase	5.05	18.65	18.65	18.65	AT2G24270
Calcium ion binding	8.98	22.62	22.62	22.62	AT2G30790
ASP1 (ASPARTATE AMINOTRANSFERASE 1)	5.95	13.96	13.96	13.96	AT2G30970
LHB1B2	10.39	24.63	24.63	24.63	AT2G34420
LHB1B1; chlorophyll binding	7.57	15.65	15.65	15.65	AT2G34430
Structural constituent of ribosome	2.85	2.96	8.93	8.93	AT2G37600
CAC3; acetyl-CoA carboxylase	2.63	12.57	12.57	12.57	AT2G38040
VHA-E2; hydrogen-exporting ATPase, phosphorylative mechanism	2.99	8.76	8.76	14.41	AT3G08560
Structural constituent of ribosome	9.31	20.76	20.76	20.76	AT3G09200
ADK1 (ADENOSINE KINASE 1)	4.10	14.40	14.40	14.40	AT3G09820
Unknown protein	4.78	9.72	9.72	9.72	AT3G12390
Unknown protein	34.21	71.56	71.56	71.56	AT3G16420
Unknown protein	10.92	28.15	28.15	28.15	AT3G16430
TCTP (TRANSLATIONALLY CONTROLLED TUMOR PROTEIN)	3.52	9.24	9.24	12.08	AT3G16640
Putative glutamine synthetase	4.79	13.28	13.28	13.28	AT3G17820
Structural constituent of ribosome	3.53	7.07	7.07	10.34	AT3G45030
unknown protein	8.96	20.10	20.10	20.10	AT3G46780

Table 6.1. Continued.

Protein name	0hr	12hr	24hr	48hr	Gene/Locus
Structural constituent of ribosome	3.53	7.07	7.07	10.34	AT3G47370
MDH; malate dehydrogenase/ oxidoreductase	7.29	23.91	23.91	23.91	AT3G47520
Fructose-bisphosphate aldolase	16.88	41.03	41.03	41.03	AT3G52930
Structural constituent of ribosome	2.85	2.96	8.93	8.93	AT3G53740
mRNA binding	5.46	11.11	11.11	14.13	AT3G63140
T01759 glycine hydroxymethyltransferase (EC 2.1.2.1) A_IG002P16.3	2.79	5.80	5.80	5.80	AT4G13890
Nuclear RNA binding protein A-like protein	3.00	9.28	9.28	9.28	AT4G16830
CHLI1 (CHLORINA 42); magnesium chelatase	3.48	10.51	10.51	10.51	AT4G18480
AT4g28520/F20O9_210	7.96	36.88	36.88	36.88	AT4G28520
Putative photosystem I subunit PSI-E protein	3.71	15.81	15.81	15.81	AT4G28750
EMB2726; RNA binding / structural constituent of ribosome	2.83	10.62	10.62	10.62	AT4G29060
PSAT; phosphoserine transaminase/ transaminase	3.62	6.50	6.50	28.62	AT4G35630
Putative glycine-rich protein	7.73	19.62	19.62	19.62	AT4G39260
Unknown protein	4.61	17.36	17.36	17.36	AT5G02240
Structural constituent of ribosome	2.85	2.96	8.93	8.93	AT5G02450
Heat shock cognate 70-2	3.21	9.28	13.09	13.09	AT5G02490
ADK2 (ADENOSINE KINASE 2); kinase	4.10	14.40	14.40	14.40	AT5G03300
Fructose-bisphosphate aldolase-like protein	4.78	12.48	22.46	22.46	AT5G03690
Nucleic acid binding	3.87	8.72	8.72	8.72	AT5G04430
ATP binding / ATP-dependent helicase/ helicase/ nucleic acid binding	2.58	9.15	9.15	9.15	AT5G11170
AT5g14200/MUA22_20	4.07	15.81	15.81	15.81	AT5G14200
Unknown protein	3.56	8.38	8.38	8.38	AT5G14260
FDH ; oxidoreductase, acting on the CH-OH group	6.25	15.40	15.40	15.40	AT5G14780
Putative 5-methyltetrahydropteroyltriglutamate	9.53	40.98	40.98	40.98	AT5G17920
Translation elongation factor	2.98	7.32	7.32	12.43	AT5G19510
5-methyltetrahydropteroyltriglutamate-homocysteine S-methyltransferase	9.70	23.38	23.38	23.38	AT5G20980
CAM6 (CALMODULIN 6); calcium ion binding	3.93	11.95	11.95	11.95	AT5G21274
Glycine hydroxymethyltransferase	2.79	5.80	5.80	5.80	AT5G26780
Luminal binding protein	3.98	10.60	10.60	10.60	AT5G28540
GDP-mannose 3,5-epimerase/ catalytic	3.02	12.32	12.32	15.35	AT5G28840
Luminal binding protein (BiP)	3.98	13.37	13.37	13.37	AT5G42020
Malate dehydrogenase/ oxidoreductase	6.46	25.98	25.98	25.98	AT5G43330
CRA1 (CRUCIFERINA); nutrient reservoir	11.79	64.33	64.33	64.33	AT5G44120
Unknown protein	2.56	2.77	5.36	5.36	AT5G45520
RNA binding	3.00	9.28	9.28	9.28	AT5G47210
AT5g56010/MDA7_5	3.98	4.43	8.47	8.47	AT5G56010
40S ribsomomal protein	3.53	7.07	7.07	10.34	AT5G62300
PSI-N; calmodulin binding	4.63	10.72	10.72	10.72	AT5G64040
PSI 9kD protein	7.48	33.21	33.21	33.21	ATCG01060

Table 6.1. Continued.

Differentially Down regulated proteins based on Sequest Cross Correlation Coefficient quantitation method

Protein Name	0hr	12hr	24hr	48hr	Gene/Locus
DNA binding	25.61	6.74	6.74	6.74	AT1G07820
Unknown protein	70.60	28.37	28.37	28.37	AT1G09340
Cinnamoyl alcohol dehydrogenase	24.57	11.69	11.69	23.78	AT1G09500
Aminomethyltransferase	98.32	48.66	48.66	48.66	AT1G11860
Unknown protein	30.64	11.96	11.96	11.96	AT1G13930
Unknown protein	15.04	3.09	3.09	3.09	AT1G23130
ATP binding / kinase/ phosphoribulokinase/ uridine kinase	64.79	27.48	27.48	27.48	AT1G32060
Putative RNA-binding protein	70.60	28.37	28.37	28.37	AT1G32790
JT0901 chaperonin 60 beta precursor - Arabidopsis thaliana	96.45	41.08	41.08	41.08	AT1G33740
At1g35160/T32G9_30	18.96	8.47	8.47	8.47	AT1G35160
Ribulose-bisphosphate carboxylase	33.83	14.26	14.26	14.26	AT1G67090
F10A5.19	38.66	19.05	19.05	19.05	AT1G75600
Putative fructose bisphosphate aldolase	30.93	12.86	12.86	12.86	AT2G21330
catalytic	6.87	3.19	3.19	3.19	AT2G29630
Mitochondrial chaperonin (HSP60)	9.15	0.00	0.00	3.85	AT2G33210
Unknown protein	7.48	3.19	11.05	11.05	AT2G44650
Unknown protein	11.53	3.32	3.32	3.32	AT2G44920
Putative transketolase precursor	47.81	22.60	22.60	22.60	AT2G45290
CA1 ; carbonate dehydratase/ zinc ion binding	111.58	46.27	46.27	46.27	AT3G01500
ATP binding / protein binding	95.00	40.71	40.71	40.71	AT3G13470
ATHM4; electron transporter/ thiol-disulfide exchange intermediate	7.29	3.63	3.63	3.63	AT3G15360
CAD4 ; cinnamyl-alcohol dehydrogenas	24.57	11.69	11.69	23.78	AT3G19450
HSP60; ATP binding / protein binding	9.15	0.00	0.00	3.85	AT3G23990
SBPASE; phosphoric ester hydrolase	5.85	2.61	2.61	2.61	AT3G55800
CCH	35.29	0.00	3.77	3.77	AT3G56240
At3g60750/T4C21_160	82.44	38.70	38.70	38.70	AT3G60750
Calcium ion binding	124.98	53.77	53.77	53.77	AT4G21280
Fructose-bisphosphate aldolase	25.26	11.02	11.02	11.02	AT4G26530
CAD5 ; cinnamyl-alcohol dehydrogenas	19.42	6.69	11.24	11.24	AT4G34230
Unknown	9.55	3.41	3.41	3.41	AT5G26000
Ribulose-bisphosphate carboxylase	19.28	7.10	10.84	10.84	AT5G38430
Unknown protein	8.18	3.73	3.73	3.73	AT5G53490
Peptidyl-prolyl cis-trans isomerase	9.99	2.92	2.92	2.92	AT5G55220
ATP binding / protein binding	78.37	31.13	31.13	31.13	AT5G56500
ATPase alpha subunit	114.72	51.94	51.94	51.94	ATCG00120
Beta subunit of coupling factor one	52.78	18.21	18.21	18.21	ATCG00480
Cytochrome f	24.04	6.74	6.74	6.74	ATCG00540
Larger subunit of Rubisco	99.14	45.17	45.17	45.17	ATCG00490

Table 6.2: Differentially Up regulated proteins based on Spectral Count method

Protein Name	0hr	12hr	24hr	48hr	Gene/Locus
LIN2 (LESION INITIATION 2); coproporphyrinogen oxidase	1	3	3	3	AT1G03475
Malate dehydrogenase/ oxidoreductase	8	18	18	18	AT1G04410
Structural constituent of ribosome	1	1	2	2	AT1G04480
RPL4; structural constituent of ribosome	1	4	4	4	AT1G07320
APX1; L-ascorbate peroxidase	1	3	3	3	AT1G07890
Phosphoserine aminotransferase	1	2	2	7	AT1G08490
Structural constituent of ribosome	4	8	8	8	AT1G09690
Putative glyceraldehyde-3-phosphate dehydrogenase	3	14	14	14	AT1G13440
F18O14.33	2	4	4	4	AT1G19310
Glutathione dehydrogenase (ascorbate)	2	4	4	4	AT1G19550
Glutathione dehydrogenase (ascorbate)	2	4	4	4	AT1G19570
ATPDIL1-1; electron transporter/ isomerase/ protein disulfide isomeras	1	0	2	3	AT1G21750
Transporter	1	5	5	5	AT1G22530
GGT1 (ALANINE-2-OXOGLUTARATE AMINOTRANSFERASE 1)	2	4	4	4	AT1G23310
CAB1 (CHLOROPHYLL A/B BINDING PROTEIN 1); chlorophyll binding	3	7	7	7	AT1G29910
Chlorophyll a/b binding protein (LHCP AB 180)	3	7	7	7	AT1G29930
T19E23.15	4	8	8	8	AT1G31355
Elongation factor EF-2	1	5	5	5	AT1G56070
Structural constituent of ribosome	4	8	8	8	AT1G57860
ACO2 (ACC OXIDASE 2)	2	6	6	6	AT1G62380
Unknown protein	1	2	2	2	AT1G62480
Monodehydroascorbate reductase	7	14	14	14	AT1G63940
VHA-E3; hydrogen-transporting ATP synthase	2	3	4	4	AT1G64200
Hydrolase, hydrolyzing O-glycosyl compounds	1	6	6	6	AT1G66270
Carbonate dehydratase/ zinc ion binding	1	1	1	3	AT1G70410
GLP1 (GERMIN-LIKE PROTEIN 1); nutrient reservoir	3	7	7	7	AT1G72610
Unknown protein	1	2	2	2	AT1G79040
3-isopropylmalate dehydrogenase/ oxidoreductase	1	3	3	3	AT1G80560
Putative Ty3-gypsy-like retroelement pol polyprotein	1	1	2	2	AT2G06170
Phosphoserine transaminase/ transaminase	1	1	1	3	AT2G17630
EMB2296; structural constituent of ribosome	1	3	3	3	AT2G18020
At2g19940/F6F22.3	1	0	2	2	AT2G19940
ALDH11A3; aldehyde dehydrogenase/ oxidoreductase	2	5	5	5	AT2G24270
ATGSTF8 (GLUTATHIONE S-TRANSFERASE 8)	2	5	5	5	AT2G27720
Photosystem II oxygen-evolving complex 23K protein, putative	3	7	7	7	AT2G30790
ASP1 (ASPARTATE AMINOTRANSFERASE 1)	2	4	4	4	AT2G30970
LHB1B2	3	7	7	7	AT2G34420
Structural constituent of ribosome	1	1	3	3	AT2G37600
CAC3; acetyl-CoA carboxylase	1	3	3	3	AT2G38040
At2g47730/F17A22.12	2	5	5	5	AT2G47730
VHA-E2; hydrogen-exporting ATPase, phosphorylative mechanism	1	3	3	5	AT3G08560
Structural constituent of ribosome	3	7	7	7	AT3G09200
Beta-glucosidase	7	14	14	14	AT3G09260
ADK1 (ADENOSINE KINASE 1)	1	3	3	3	AT3G09820
Unknown protein	1	3	3	3	AT3G12390

Table 6.2. Continued.

Protein Name	0hr	12hr	24hr	48hr	Gene/Locus
Unknown protein	3	7	7	7	AT3G16430
TCTP (TRANSLATIONALLY CONTROLLED TUMOR PROTEIN)	1	3	3	4	AT3G16640
Putative glutamine synthetase	1	3	3	3	AT3G17820
LTA2 (PLASTID E2 SUBUNIT OF PYRUVATE DECARBOXYLASE)	1	1	1	2	AT3G25860
Structural constituent of ribosome	1	2	2	3	AT3G45030
Unknown protein	3	6	6	6	AT3G46780
Structural constituent of ribosome	1	2	2	3	AT3G47370
MDH; malate dehydrogenase/ oxidoreductase	2	7	7	7	AT3G47520
Fructose-bisphosphate aldolase	5	13	13	13	AT3G52930
Structural constituent of ribosome	1	1	3	3	AT3G53740
mRNA binding	2	3	3	4	AT3G63140
T01759 glycine hydroxymethyltransferase (EC 2.1.2.1) A_IG002P16.3	1	2	2	2	AT4G13890
Nuclear RNA binding protein A-like protein	1	3	3	3	AT4G16830
CHLI1 (CHLORINA 42); magnesium chelatase	1	3	3	3	AT4G18480
SHD (SHEPHERD); ATP binding / unfolded protein binding	1	2	2	2	AT4G24190
AT4g28520/F20O9_210	2	11	11	11	AT4G28520
Putative photosystem I subunit PSI-E protein	1	5	5	5	AT4G28750
EMB2726; RNA binding / structural constituent of ribosome	1	3	3	3	AT4G29060
PSAT; phosphoserine transaminase/ transaminase	1	2	2	7	AT4G35630
Putative glycine-rich protein	3	8	8	8	AT4G39260
Unknown protein	1	3	3	3	AT5G02240
Structural constituent of ribosome	1	1	3	3	AT5G02450
Heat shock cognate 70-2	1	2	3	3	AT5G02490
ADK2 (ADENOSINE KINASE 2); kinase	1	3	3	3	AT5G03300
Fructose-bisphosphate aldolase-like protein	1	3	5	5	AT5G03690
Nucleic acid binding	1	2	2	2	AT5G04430
ATP binding / ATP-dependent helicase/ helicase/ nucleic acid binding	1	3	3	3	AT5G11170
structural constituent of ribosome	2	4	4	4	AT5G13510
AT5g14200/MUA22_20	1	3	3	3	AT5G14200
Unknown protein	1	2	2	2	AT5G14260
FDH ; oxidoreductase, acting on the CH-OH group	2	5	5	5	AT5G14780
Putative 5-methyltetrahydropteroyltriglutamate	4	10	10	10	AT5G17920
Translation elongation factor	1	2	2	4	AT5G19510
5-methyltetrahydropteroyltriglutamate-homocysteine S-methyltransferase	2	6	6	6	AT5G20980
CAM6 (CALMODULIN 6); calcium ion binding	1	3	3	3	AT5G21274
Glycine hydroxymethyltransferase	1	2	2	2	AT5G26780
Luminal binding protein	1	3	3	3	AT5G28540
GDP-mannose 3,5-epimerase/ catalytic	1	4	4	5	AT5G28840
Luminal binding protein (BiP)	1	4	4	4	AT5G42020
Malate dehydrogenase/ oxidoreductase	2	7	7	7	AT5G43330
CRA1 (CRUCIFERINA); nutrient reservoir	3	16	16	16	AT5G44120
Unknown protein	1	1	2	2	AT5G45520
RNA binding	1	3	3	3	AT5G47210
ADG1 (ADP GLUCOSE PYROPHOSPHORYLASE SMALL SUBUNIT 1)	1	2	2	2	AT5G48300
AT5g56010/MDA7_5	1	1	2	2	AT5G56010

Table 6.2. Continued.

Protein Name	0hr	12hr	24hr	48hr	Gene/Locus
40S ribsomomal protein	1	2	2	3	AT5G62300
PSI-N; calmodulin binding	1	3	3	3	AT5G64040
PSI 9kD protein	2	8	8	8	ATCG01060

Differentially Down regulated proteins based on Spectral Count method

Protein Name	0hr	12hr	24hr	48hr	Gene/Locus
DNA binding	8	2	2	2	AT1G07820
Unknown Protein	19	7	7	7	AT1G09340
Cinnamoyl Alcohol Dehydrogenase	6	3	3	6	AT1G09500
Aminomethyltransferase	29	14	14	14	AT1G11860
Unknown Protein	8	3	3	3	AT1G13930
Oxidoreductase	10	5	5	5	AT1G20020
Unknown	2	1	1	3	AT1G20620
Unknown Protein	4	1	1	1	AT1G23130
ATP binding / kinase/ phosphoribulokinase/ uridine kinase	19	8	8	8	AT1G32060
Putative Rna-Binding Protein	19	7	7	7	AT1G32790
JT0901 chaperonin 60 beta precursor - Arabidopsis thaliana	31	13	13	13	AT1G33740
Unknown Protein	6	3	3	3	AT1G54410
CPN60B (CHAPERONIN 60 BETA); ATP binding / protein binding	34	17	17	17	AT1G55490
Putative Cytosolic Factor Protein	2	1	1	2	AT1G72150
F10A5.19	14	7	7	7	AT1G75600
Putative Fructose Bisphosphate Aldolase	10	4	4	4	AT2G21330
Glycine Dehydrogenase (Decarboxylating)	9	4	4	4	AT2G26080
Catalytic	2	1	1	1	AT2G29630
Mitochondrial Chaperonin (Hsp60)	2	0	0	1	AT2G33210
Unknown Protein	7	3	4	4	AT2G37660
Unknown Protein	2	1	3	3	AT2G44650
Unknown Protein	3	1	1	1	AT2G44920
Putative Transketolase Precursor	15	7	7	7	AT2G45290
CA1 ; carbonate dehydratase/ zinc ion binding	30	12	12	12	AT3G01500
ATP binding / protein binding	30	12	12	12	AT3G13470
ATHM4; electron transporter/ thiol-disulfide exchange intermediate	2	1	1	1	AT3G15360
CAD4 (CINNAMYL ALCOHOL DEHYDROGENASE 4)	6	3	3	6	AT3G19450
HSP60; ATP binding / protein binding	2	0	0	1	AT3G23990
SBPASE; phosphoric ester hydrolase	2	1	1	1	AT3G55800
CCH	11	0	1	1	AT3G56240
AT3G60750/T4C21_160	27	13	13	13	AT3G60750
Calcium Ion Binding	35	17	17	17	AT4G21280
Fructose-Bisphosphate Aldolase	8	3	3	3	AT4G26530
Aminopeptidase	2	1	3	3	AT4G30920
CAD5 (CINNAMYL ALCOHOL DEHYDROGENASE 5)	5	2	3	3	AT4G34230
Putative Fructose-Bisphosphate Aldolase	12	6	6	6	AT4G38970
Structural Constituent Of Ribosome	2	1	1	2	AT5G03850
CA2 ; carbonate dehydratase/ zinc ion binding	22	8	8	8	AT5G14740
Unknown	3	1	1	1	AT5G26000

Table 6.2. Continued.

Protein Name	0hr	12hr	24hr	48hr	Gene/Locus
Glutamate-Ammonia Ligase Precursor	7	3	3	3	AT5G35630
Ribulose-Bisphosphate Carboxylase	5	2	3	3	AT5G38430
ATPHB3 (PROHIBITIN 3)	2	1	1	1	AT5G40770
KAS I; catalytic/ fatty-a	2	0	1	2	AT5G46290
Unknown Protein	2	1	1	1	AT5G53490
Peptidyl-Prolyl Cis-Trans Isomerase	3	1	1	1	AT5G55220
ATP binding / protein binding	25	9	9	9	AT5G56500
ATPase alpha subunit	34	17	17	17	ATCG00120
Beta Subunit Of Coupling Factor One	14	6	6	6	ATCG00480
Larger Subunit Of Rubisco	35	16	16	16	ATCG00490
Cytochrome F	7	2	2	2	ATCG00540

The distribution of differentially regulated proteins into different Gene Ontology categories were represented in Figure 6.2 and revealed very close similarity between the two non label methods.

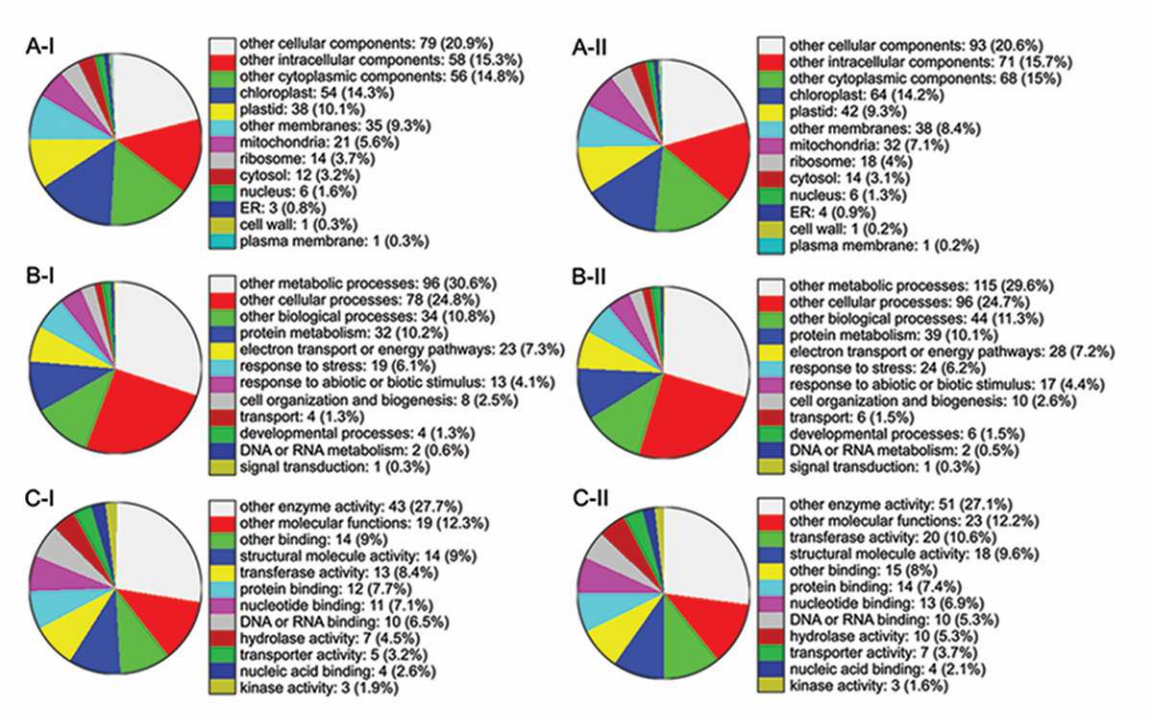

Figure 6.2: Distribution of differentially regulated proteins of MudPIT analyzed samples among different GO categories.

Note: (A) Pie chart presentation of the proteins based on ΣXcorr method. (A-I) based on Cellular Localization (CC). (A-II) based on Biological Processes (BP). (A-III) based on Molecular Functions (MF). (B) Differentially regulated proteins based on SC method (B-I) Based on CC. (B-II) Based on BP. (B-III). Based on MF.

Comparison of differentially regulated proteins identified by label-free quantitation methods and 2-DE gel based method

Since the protein samples used for above shotgun proteomic studies were prepared exactly in the same way as the reported 2-DE gel based cotyledon proteome studies (Chitteti and Peng, 2007), here the quantification results of the two label-free methods were compared with those of the 2-DE gel stain density measurements. Among the 42 differentially regulated proteins identified based on 2-DE gel stain density, 21 proteins were also identified in shotgun proteomics (Table 6.3). Out of the 21 commonly

found proteins, 10 had similar regulation pattern in all three quantification methods as shown in Table 6.3 (S.No. from 1 to 10). Proteins numbered 11, 12, and 13 were upregulated based on 2DE data, but down-regulated using Xcorr and SC quantitation methods. However, eight of the 21 differentially regulated proteins numbered from 14 to 21 (Table 6.3) showed no significant quantity change when analyzed using the Spectral Count and ΣXcorr methods. Among these eight proteins, five of them (14, 16, 18, 19, 20) had been found in more than one protein spot on 2-DE gels (Figure 6.3 and Table 6.4). For example, the glyceraldehyde 3- phosphate dehydrogenase A subunit protein was identified in nine different protein spots, and hypothetical protein (T517.3) were identified in four different proteins spots. Two of the five proteins with multiple spots on 2-DE gel had been found to be phosphorylated (Chitteti and Peng, 2007). In addition, we could not exclude that the other two proteins might also exist in other protein spots since few visible protein spots have not been identified. Furthermore, it has been found that one of the three proteins (Protein numbered 12), which displayed significant up regulation based on 2-DE gel method but down regulation based on label-free quantitation methods, was identified in four protein spots and this protein is a phosphorylated protein. These observations indicated a possibility that while a protein spot displayed quantity increase, the other protein spots with the same identity did not or displayed opposite quantity change. The shotgun coupled label-free quantification methods can not distinguish proteins of different modification status because most of the peptides released from these proteins are identical except the one with amino acid modification. In contrast, 2-DE gel can separate a single protein into multiple spots according to modification status, thus the differentially regulated proteins identified by 2-DE gel method only represent one of the

modification forms instead of the total of the protein. Therefore, the label-free quantification methods did not fit with the 2-DE gel quantification results in many proteins.

Table 6.3: Relative quantities of differentially expressed proteins that are commonly found both in 2-DE and MudPIT analyzed samples

		Sum of X CORR - Quantitation	Peptide Count - Quantitation	2DE Quantitation
S.No.	**Gene/Locus**	**0hr / 12hr / 24hr / 48hr**	**0hr / 12hr / 24hr / 48hr**	**0hr / 12hr / 24hr / 48hr**
1	At4g09010	0 / 3.53 / 3.53 / 3.53	0 / 1 / 1 / 1	234848.37 / 246568.8 / 109231.8 / 91026.5
2	At3g16640	3.52 / 9.24 / 9.24 / 12.08	1 / 3 / 3 / 4	57018.1 / 117107.8 / 176160.5 / 132908.4
3	At1g29910	10.389999 / 24.63 / 24.63 / 24.63	3 / 7 / 7 / 7	179897.5 / 397573.5 / 451542.8 / 375985.7
4	AT4g28520	7.96 / 36.88 / 36.88 / 36.88	2 / 11 / 11 / 11	7037.2 / 51180.5 / 23949.3 / 21833.1
5	At2g34430	7.57 / 15.65 / 15.65 / 15.65	2 / 3 / 3 / 3	21286.1 / 24568.6 / 26038.9 / 45552.4
6	At5g20020	0 / 7.15 / 7.15 / 7.15	0 / 2 / 2 / 2	23193.2 / 64478 / 81150.8 / 56331.2
7	At1g07890	4.51 / 11.98 / 11.98 / 11.98	1 / 3 / 3 / 3	281813.145 / 130469.0486 / 118908.5 / 123062.5087
8	At4g13850	0 / 0 / 0 / 2.98	0 / 0 / 0 / 1	5000.9 / 15087.3 / 38647.4 / 39153.3
9	At5g14740	74.8 / 30.02 / 30.02 / 30.02	22 / 8 / 8 / 8	82385.3 / 41256.2 / 42856.4 / 40583.8
10	AtCg00490	99.14 / 45.170002 / 45.170002 / 45.170002	35 / 16 / 16 / 16	20626.4 / 13023.9 / 9681 / 5901.1
11	At1G35160	18.96 / 8.469999 / 8.469999 / 8.469999	5 / 3 / 3 / 3	7283.5 / 21165.6 / 48355.2 / 15432.5
12	At1g32060	64.79 / 27.480001 / 27.480001 / 27.480001	19 / 8 / 8 / 8	6272.5 / 13862.2 / 30797.9 / 32867.9
13	AT2G34590	95.05 / 52.879997 / 52.879997 / 52.879997	31 / 17 / 17 / 17	40735 / 83089 / 117708.7 / 121708
14	At3g26650	184.91 / 136.1 / 136.1 / 136.1	54 / 40 / 40 / 40	7509 / 17628.2 / 24880.5 / 25584.7
15	At5g42270	0 / 0 / 0 / 3.45	0 / 0 / 0 / 1	11528.3 / 29206.1 / 40792.4 / 32441.3
16	At4g04640	13.26 / 19.58 / 19.58 / 19.58	4 / 6 / 6 / 6	446462 / 222120.398 / 190795.7 / 171058.2375
17	AT5G60390	39.71 / 49.95 / 49.95 / 49.95	12 / 15 / 15 / 15	15638.5 / 37907.6 / 121173.3 / 25839.4
18	At3g14420	10.64 / 13.620001 / 13.620001 / 13.620001	3 / 4 / 4 / 4	111629.679 / 138458.5 / 52408.3 / 74589.6
19	At2G39730	17.1 / 17.35 / 17.35 / 17.35	6 / 6 / 6 / 6	60203.3 / 83406.4 / 121610.8 / 130641.1
20	At1g70580	7.15 / 11.8 / 11.8 / 11.8	2 / 3 / 3 / 3	80618.6 / 183810.4 / 162043.4 / 93458.6
21	At4g09000	30.15 / 28.130001 / 28.130001 / 28.130001	9 / 9 / 9 / 9	38716 / 77856 / 121446.2 / 76983.8

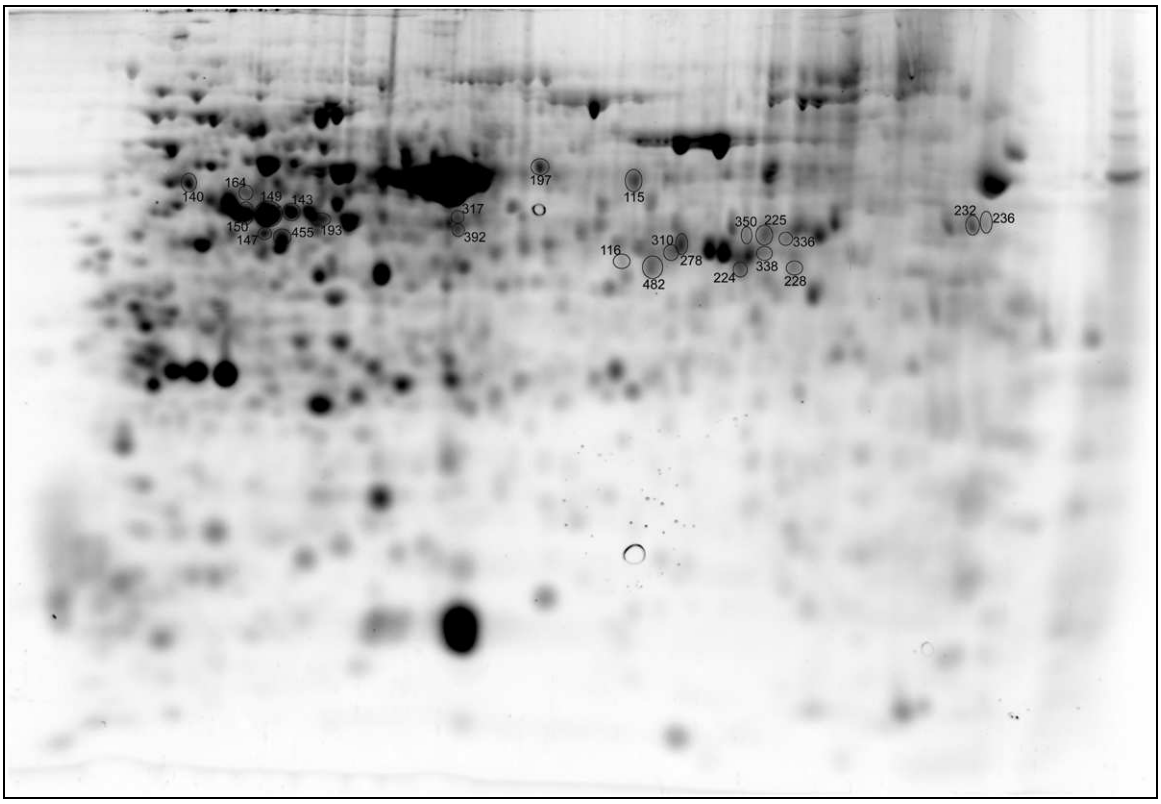

Figure 6.3: A single MudPIT identified protein was separated into multiple spots on 2-DE gel. Six proteins were separated into 24 spots.

Note: The details of these proteins spots were explained in the Table 6.4.

Table 6.4: A single MudPIT identified protein separated into multiple protein spots on 2-DE gels.

	Spot #	Protein name	pI	MW	Peptides	C.I.%	Gene/Locus
12	147	Phosphoribulokinase, chloroplast precursor (EC 2.7.1.19) (Phosphopentokinase)	5.0	42.3	18	100.00	AT1G32060
	149	Phosphoribulokinase, chloroplast precursor (EC 2.7.1.19) (Phosphopentokinase)	5.0	44.5	22	100.00	AT1G32060
	150	Phosphoribulokinase, chloroplast precursor (EC 2.7.1.19) (Phosphopentokinase)	4.9	44.8	13	100.00	AT1G32060
	455	Phosphoribulokinase, chloroplast precursor (EC 2.7.1.19) (Phosphopentokinase)	5.1	42.5	18	100.00	AT1G32060
14	224	Glyceraldehyde 3-phosphate dehydrogenase A subunit	6.4	38.4	13	100.00	AT3G26650
	225	Glyceraldehyde 3-phosphate dehydrogenase A, chloroplast precursor (EC 1.2.1.13)	6.6	42.5	20	100.00	AT3G26650
	228	Glyceraldehyde 3-phosphate dehydrogenase A subunit	6.4	39.6	8	100.00	AT3G26650
	310	Glyceraldehyde 3-phosphate dehydrogenase A subunit	5.9	41.0	12	100.00	AT3G26650
	317	Glyceraldehyde 3-phosphate dehydrogenase B, chloroplast precursor (EC 1.2.1.13)	5.5	45.3	11	100.00	AT3G26650
	336	Glyceraldehyde 3-phosphate dehydrogenase A subunit	6.8	41.5	4	100.00	AT3G26650
	338	Glyceraldehyde 3-phosphate dehydrogenase A subunit	6.6	40.0	13	100.00	AT3G26650
	350	Glyceraldehyde-3-phosphate dehydrogenase (NADP) (EC 1.2.1.13) A precursor	6.4	42.6	16	100.00	AT3G26650
	392	Glyceraldehyde 3-phosphate dehydrogenase B,chloroplast precursor (EC 1.2.1.13)	5.5	43.0	13	100.00	AT3G26650
16	116	ATP synthase gamma chain 1, chloroplast precursor (EC 3.6.3.14)	5.7	39.2	13	100.00	AT4G04640
	278	ATP synthase gamma chain 1, chloroplast precursor (EC 3.6.3.14)	5.8	40.9	13	100.00	AT4G04640
	482	ATP synthase gamma chain, (EC 3.6.1.34)	5.8	38.7	11	100.00	AT4G04640
18	232	MOA2_2	8.3	43.4	4	100.00	AT3G14420
	236	MOA2_2	8.4	43.7	11	100.00	AT3G14420
19	140	T5I7.3	4.7	50.8	7	100.00	AT2G39730
	143	T5I7.3	5.1	45.0	26	100.00	AT2G39730
	164	T5I7.3	4.9	49.9	7	100.00	AT2G39730
	193	T5I7.3	5.2	44.6	18	100.00	AT2G39730
20	115	F26F24_4	5.7	51.5	7	100.00	AT1G70580
	197	F26F24_4	5.6	54.6	25	100.00	AT1G70580

Discussions

The 2-DE/MALDI/MS/MS and LC/LC/MS/MS methods complement each other in protein identification

The detailed comparison of the 500 proteins identified by 2-DE/MALDI/MS/MS methods and the 662 proteins identified by LC/LC/MS/MS was done and found that the number of overlapped proteins identified with the two methods is very limited, 139 out of the 1023 identified proteins. Both the ionization method difference and the bias of 2-DE gel for specific proteins may have contributed to significant difference in protein coverage. Further examination of the identified proteins indicates that the LC/LC/MS/MS method can cover protein in a wide range of pIs including those with pI higher than 10 and lower than 4. In contrast, the optimal pI peak of 2-DE/MALDI/MS/MS identified proteins was around 5 to 6. In addition, the LC/LC/MS/MS method had a better coverage on proteins with molecular mass below 40 kDa, particularly those proteins below 20 kDa. Since each of the two methods identified many proteins which are not covered by the other method, the two methods complement each other very well.

Comparison of differential regulated proteins identified by the two label-free methods

The cotyledon proteome differential regulation upon the induction of cell dedifferentiation has been studied using two different label-free methods at 4 different time points. The Spectral Count method identified 143 differentially regulated proteins and the cross correlation coefficient method identified 121 differentially regulated

proteins. Analysis of the identified proteins indicates that over 81% of the differentially regulated proteins are common to these two methods. In addition, the patterns of differential regulation at the four induction time points were also very similar. Since identified peptide abundance is a critical factor in protein quantity determination, the results are understandable. Furthermore, each of the two methods identified some unique proteins. The Spectral Count method measures the Spectral Count only and thus has a broader coverage. The cross correlation coefficient is determined by several factors including the number of fragment ions in the mass spectrum, their relative abundance, continuity of ion series, and presence of immonium ions for certain amino acids in the spectrum. The differentially regulated proteins identified with the two methods include calcium binding protein, oxidative stress response protein, and so on.

The label-free quantification methods and 2-DE gel stain method can identify different types of regulation in protein quantity

Fourty-two differentially expressed proteins using 2-DE/MALDI/MS/MS method have been identified in previous studies. Among them, 21 have also been identified using LC/LC/MS/MS in this study. Surprisingly, 11 of the 21 proteins had very different quantity profile when analyzed using label-free method, including 8 proteins had no expression change and three proteins had opposite expression profile change. Further analysis find that 6 of the 11 proteins had been identified in more than one protein spots and three of them are phosphorylated when examined by Pro-Q Diamond Phosphoprotein Stain. Because many protein spots have not been identified, we can not exclude that some of the other 5 proteins can also be found in more than one protein spots. These

observations suggest that protein differential expression revealed by 2-DE gel may represent the change in one of modification forms. The label-free method can not detect such a change because other spots with the same identity can release the same set of peptide. These observations suggest that the 2-DE gel based method may have advantage in detecting quantitative change of a specific form of the protein when the modification results in detectable pI and molecular mass change on 2-DE gel. On the other hand, the 2-DE method has major disadvantage in the quantification of protein expression change because many protein spots contain more than one protein and we can not pinpoint the change in which protein.

Summary

The cell dedifferentiation has been investigated using LC/LC/MS/MS approach in *Arabidopsis*. Protein differential expression at four induction time points were analyzed with two label-free quantification methods and compared with the 2-DE gel based method. While the two label-free methods generated very similar results, about half of the differential expressed proteins identified by 2-DE gel does not match with the quantification results obtained using label-free methods. Further analyses indicate that the 2-DE gel based method reveals the change in one modification form of proteins with multiple status. In contrast, the label-free method can not distinguish the different forms generated by modification or close gene family members. Our results suggest that the 2-DE gel and LC/LC/MS/MS approaches complement very well to each other in both protein identification and differential expression studies.

From MudPIT results, using ESI MS/MS, 662 proteins have been identified with the parameters: Xcorr values = +1-1.9; +2-2.2; +3-3.75, Delta Corr = 0.1 and used pFactor - 0.001. By combining the data from both 2 DE and MudPIT methods, the proteome map of cotyledon was constructed with 1,162 proteins. There are only 139 proteins commonly found both in 2-DE and MudPIT. This protein distribution is represented in Figure 3.57 using a Venn diagram. The genes corresponding to all these proteins were characterized using the Gene Ontology tools. They were grouped into different levels and presented in pie charts (Figure 3.5). The results show that the chloroplast genes occupy the largest number of identified genes, about 13.7% of the total identified genes. Other membrane proteins represented 9.6% and mitochondrial proteins represented 8.4%. Ribosome, cytosol, nucleus, plasma membrane, ER, cell wall and Golgi apparatus only occupy 3.9%, 3.3%, 2.1%, 0.7%, 0.5%, 0.3% and 0.1% of the total identified genes, respectively (Figure 3.5A). These results strongly substantiate that the primary role of cotyledon is photosynthesis and energy metabolism. In addition, analyses based on Biological Process reveal that 21% of the identified proteins are involved in Other physiological processes, 20.5% are involved in Other metabolic processes in addition to those involved in protein, nucleic acids, and energy metabolic pathways, providing further evidence of the important role of metabolism and photosynthesis in the cotyledon cells. Proteins in other biological processes and molecular function categories were presented in detail in Figures 3.5B and 3.5C. The establishment of the cotyledon proteome map has provided new insight into the function of cotyledon at the proteome level, enriched the plant proteome database, and set a foundation for further research using cotyledon as materials.

CHAPTER VII

SUMMARY

In Developmental Biology, differentiation is a process where cells acquire specialization, thereby, giving greater variety of organs. This acquired or assigned cell fate is memorized in subsequent growth and development. Until recently it was thought that once a cell had acquired a stable differentiated state it could not change its phenotype. But, recent studies prove that cells can switch their cell fate and can express the functional characteristics of the other tissues through dedifferentiation. So, dedifferentiation is a reversible process where regression of a differentiated cell into an embryonic unspecialized state. During this process the previous memory of the cell will be erased, genome will be reprogrammed and totipotency is recovered. It is also been identified that dedifferentiation is an intermediate step in several transdifferentiation processes. For example, during the transdifferentiation of the iris pigment epithelial (IPE) cells into lens cells, dedifferentiation is involved in the formation of intermediate cells. This dedifferentiation phenomenon was observed in plants for centuries.

Seed is a resting state embryo. Upon germination it is differentiated into a plant consisting of leaves, roots, shoots etc. During the development, several mitotic divisions take place, and a type of cell gives the same type of cell during these divisions. So, the stability of differentiation is a corner stone for the normal development. During the

normal development, cells don't loose their assigned specific memory and retain the plasticity. They can acquire the new fates and switch their roles. Since they have this capability, Cells from any part of the plant (explants) upon apply of the hormones can undergo dedifferentiation and gives undifferentiated cell mass called callus. The callus cells are completely totipotent. Any part of the plant or a complete plant can be regenerated from the callus. It is also possible to divide this callus into several pieces which can induce multiple plants. During this processes, hormones such as auxins (2.4.D., Indole acetic acid) and cytokinins (Kinetin, 6-Benzoic acid) play a critical role. For example, in the presence of 1mg/l 2.4.D and 0.2mg/l Kinetin, the differentiated tissue undergoes dedifferentiation and gives callus. When this ratio is reverted, dedifferentiated callus are induced to form a plant. High concentrations of auxins promote the growth of only roots and high concentrations of cytokinins promote the growth of only shoots. This dedifferentiation phenomenon was also observed in animal kingdom.

Although these remarkable phenomenons were observed in many species for centuries, the molecular mechanisms of cell dedifferentiation at proteome and genome levels and what signal transduction components involved during these processes are not clearly understood yet. In order to understand the molecular mechanism of cell dedifferentiation and to identify the key factors involved during this process, the *Arabidopsis* cotyledon tissue and phosphoproteomes were established using gel based and non gel based methods. The differential expression of the total and phosphoproteins was extensively studied in this work using Proteomics as a tool. The corresponding genes were also characterized using Gene Ontology tools.

In this study, the traditional Two-Dimensional Gel Electrophoresis (2-DE) and Multidimensional Protein Identification Technology (MudPIT) methods have been used to establish the cotyledon proteome map of *Arabidopsis*. Mostly phenol based protocols were used to extract the proteins and lyophilized them to powders in a speed vacuum. The samples were dissolved in rehydration buffer and isoelectric focusing was carried out in Bio-Rad IEF cell to separate the proteins according to their pI values. Bio-Rad Dodeca cell was used to carry out the second dimensional gel electrophoresis where proteins separate according to their molecular mass. Using this unit up to 12 gels can be run at a time. Later gels were stained with SYPRO Ruby and scanned using VersaDoc 4000 image system. The gels were analyzed using PDQuet 7.4.0 software. At least three replicas were used in all experiments to get the statistical significance of the change. All significant spots were excised using Bio-Rad Robotic Spot cutter. It can cut up to 100 spots in an hour very precisely without user interference. Washing, reduction, alkylation, and in-gel proteolytic digestion of the spots were carried out on Genomic Solutions robotic digester and spotter. The peptides were desalted using C18 zip tips and spot on MALDI plates after mixing the peptides with CHCA matrix. Mass spectra were collected on the ABI 4700 MALDI TOF TOF mass spectrometer, and protein identification was done using ABI GPS explorer software and NCBI nr database. Proteins score C.I. % over 95% were considered as positive ID. Among the 748 excised protein spots, 603 were annotated after MS/MS mass analyses. Of these annotated spots, 500 proteins were identified with confidence intervals (C.I. %) over 95%. Functional categorization of proteins was carried out after getting the corresponding gene IDs. The gene distribution

was grouped on the basis of Cellular localization, biological process and Molecular functions.

In order to overcome the limitations of gel based methods, the protein samples were also analyzed using another approach called Multidimensional Protein Identification Technology. In this method the complex protein mixture was quantified, reduced, alkylated, trypsin digested and desalted off line using peptide macro trap and analyzed on Thermo Finnigan Proteome X workstation. Here the chromatography proceeds in cycles eluting the peptides off the SCX column followed by RP column. The eluted peptides were analyzed in Tandem Mass Spectrometer. The spectral data was analyzed using the Bioworks 3.1 software. Using the MudPIT method, 662 proteins were identified and analyzed according to the Gene Ontology rules as described above.

Most of the 2-DE identified proteins were distributed within the pI range of 5-7. Yet 128 basic proteins with pI over 10 were identified by MudPIT technique. Therefore, in order to increase the basic protein coverage, it is important to employ MudPIT technique. When this analysis was carried out based on Molecular mass of the proteins, it was observed that low molecular mass proteins coverage was high in MudPIT analysis. Because the low molecular mass proteins are poorly stained with dyes and more chances to miss them at the time of spot picking. Five proteins with molecular mass 150 KDa were identified by the gel based method, but 20 proteins with over 150 KDa were identified by MudPIT. The purpose of adopting these two techniques is there are several advantages and disadvantages of each method. The advantages of the gel based method include visualization. If we find a protein differentially expressed or newly expressed compared with the gel based method, we can simply excise it and get the identification. It

is fairly simple. We can also use the gels for making blots for further analysis. Disadvantages are: it is labor intensive, gel to gel variation makes us to repeat the experiments several times until we get the similar replicas. The pH is another limitation. The advantages of Non Gel based methods include: we can eliminate the complicated 2-DE step, using this high throughput technique to separate thousands of peptides in a single run. Furthermore, there is no pH limitation. Disadvantage is that quantitative information is not obtained unless fluorescent dyes are used.

Protein expression profile at 0, 12, 24, and 48hr time points after induction of cell dedifferentiation with hormones - auxins and cytokinins on *Arabidopsis* cotyledons has been explored. A two fold difference in expression compared to control in all three biological replicas was used as cut off for the identification of differentially regulated proteins. Among the 500 2-DE identified distinct proteins, 30 were up-regulated and 12 were down-regulated. The total differentially regulated proteins were 42. The differentially expressed proteins were then calculated in different categories. The relative percentages were obtained by dividing the number of differentially regulated proteins in one category by the total number of proteins identified in that compartment of the cell. This analysis revealed that nuclear proteins had the highest ratios of differential expression. The proteins involved in developmental processes and in cell organization and biogenesis had highest ratios of differential expression.

Dr. Bridges' group at Mississippi State have recently developed a program called "ProtQuant" which can quantify the MudPIT analyzed data based on the sum of Cross Correlation values and on the Spectral Count method. The differential expression of protein samples at different time points was also identified using these two methods.

Based on the sum of cross correlation method, 83 proteins were up-regulated and 38 proteins were down-regulated. Based on the Spectral Count method, 93 proteins were up-regulated and 50 proteins were down-regulated. This analysis revealed that these two non label quantification methods work in the same way. A protein that is considered as differentially expressed using sum of the cross correlation method is also considered as differentially expressed in the Spectral Count method.

Reversible protein phosphorylation plays an essential role in multitude biological processes, including developmental regulation and hormonal responses. To investigate the role of protein phosphorylation in cell dedifferentiation, the cotyledon phospho proteome has been examined using Pro Q Diamond Phosphoprotein gel stain from Molecular Probes followed by Mass Spectrometry analysis. Pro Q Diamond stain displayed high affinity to Phosphoproteins. Of the 102 putative phosphoprotein spots detected with the dye 53 proteins identified by Mass analysis had a C.I, % over 95. Only 11% of the total identified proteins were stained by the phosphoprotein specific dye. Nine of the 53 identified phosphoproteins were differentially regulated during cell dedifferentiation based on the criteria of two fold change in all three biological replicas. These proteins or protein families have been reported to be phosphorylated (Larsson et al., 1983; Aitken et al., 1995; Yarm, 2002; Jones et al., 2006). Of the nine differentially regulated phosphoproteins, six of them are reported to be regulated by phosphorylation.

Several identified proteins and group of proteins that were differentially regulated have been found to play a key role in cell dedifferentiation. Translational controlled tumor protein (TCTP) was significantly differentially expressed during cell dedifferentiation and also stained by the phosphoprotein specific dye. Consistent with our

results mammalian TCTP protein has also been reported to be phosphorylated by miotic polo like kinase at serine residues. It plays a key role in cell cycle control, tumerigenesis, and microtubule stabilization. Another protein called “Putative translational elongation factor eEF1 alpha chain” was also up-regulated in a similar profile with TCTP. When putting all differentially regulated proteins together to find their possible interactions using ingenuity pathway program, it was found that TCTP affects the activation of eEF1 alpha chain via protein-protein interaction. Therefore, the involvement of TCTP in cell dedifferentiation is highly interesting. Understanding the molecular mechanisms of TCTP will provide the critical insight into the general mechanism of cell dedifferentiation.

The 14-3-3 proteins are a large group of phospho serine / threonine binding proteins that play a critical role in various signal transduction pathways. The 14-3-3 protein homologs RC12, T32G9_30, and 14-3-3 like protein GF14 chi were up-regulated with the maximum expression level at 24 hours. In addition, GF14 chi and T32G9_30 were identified as putative phosphoproteins. Although the biological functions of these three proteins still remain to be explored, identification of the binding patterns of these three proteins might lead to the discovery of cellular signaling pathways that are critical to cell dedifferentiation.

GTP binding nuclear protein RAN-2 and GTP binding protein SAR1B were both up-regulated during cell dedifferentiation. It is reported that these proteins are involved in diversified cellular processes. Up regulation of these proteins indicate that cell dedifferentiation affect the broad range of cellular processes.

Several hypothetical proteins and function unknown proteins were differentially regulated. Further investigation on these genes could significantly contribute to the research on dissecting cell dedifferentiation pathways.

These studies have provided significant new insight into protein and phosphoprotein differential expression during cell dedifferentiation in plants, and the insights gained through non animal systems might one day lead to the development of methodologies that could aid in reprogramming all sorts of cells.

REFERENCES

The Arabidopsis Initiative (2000). Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. Nature 408, 796-815.

Agrawal, G.K., and Thelen, J.J. (2005). Development of a simplified, economical polyacrylamide gel staining protocol for phosphoproteins. Proteomics 5, 4684-4688.

Agrawal, G.K., and Thelen, J.J. (2006a). Large Scale Identification and Quantitative Profiling of Phosphoproteins Expressed during Seed Filling in Oilseed Rape. Mol Cell Proteomics 5, 2044-2059.

Aitken, A., Howell, S., Jones, D., Madrazo, J., and Patel, Y. (1995). 14-3-3 alpha and delta are the phosphorylated forms of raf-activating 14-3-3 beta and zeta. In vivo stoichiometric phosphorylation in brain at a Ser-Pro-Glu-Lys MOTIF. J Biol Chem 270, 5706-5709.

Alberts, B., Bray, D., Lewis, J., and Raff, M. (1994). Molecular Biology of the Cell. Garland Pub., New York, 330-331.

Alexandersson, E., Saalbach, G., Larsson, C., and Kjellbom, P. (2004). *Arabidopsis* plasma membrane proteomics identifies components of transport, signal transduction and membrane trafficking. Plant Cell Physiol 45, 1543-1556.

An, G., Watson, B.D., and Chiang, C.C. (1986). Transformation of Tobacco, Tomato, Potato, and *Arabidopsis thaliana* Using a Binary Ti Vector System. Plant Physiol 81, 301-305.

Anderson, L., and Seilhamer, J. (1997). A comparison of selected mRNA and protein abundances in human liver. Electrophoresis 18, 533-537.

Bai, S., Chen, L., Yund, M.A., and Sung, Z.R. (2000). Mechanisms of plant embryo development. Curr Top Dev Biol 50, 61-88.

Ball, E. (1950). Differentiation in a callus culture of Sequoia sempervirens. Growth 14, 295-325.

Becker, R.O., Chapin, S., and Sherry, R. (1974). Regeneration of the ventricular myocardium in amphibians. Nature 248, 145-147.

Berardini, T.Z., Mundodi, S., Reiser, L., Huala, E., Garcia-Hernandez, M., Zhang, P., Mueller, L.A., Yoon, J., Doyle, A., Lander, G., Moseyko, N., Yoo, D., Xu, I., Zoeckler, B., Montoya, M., Miller, N., Weems, D., and Rhee, S.Y. (2004). Functional annotation of the *Arabidopsis* genome using controlled vocabularies. Plant Physiol 135, 745-755.

Berggren, K., Steinberg, T.H., Lauber, W.M., Carroll, J.A., Lopez, M.F., Chernokalskaya, E., Zieske, L., Diwu, Z., Haugland, R.P., and Patton, W.F. (1999). A luminescent ruthenium complex for ultrasensitive detection of proteins immobilized on membrane supports. Anal Biochem 276, 129-143.

Bevan, M., Bancroft, I., Bent, E., Love, K., Goodman, H., Dean, C., Bergkamp, R., Dirkse, W., Van Staveren, M., Stiekema, W., Drost, L., Ridley, P., Hudson, S.A., Patel, K., Murphy, G., Piffanelli, P., Wedler, H., Wedler, E., Wambutt, R., Weitzenegger, T., et al., (1998). Analysis of 1.9 Mb of contiguous sequence from chromosome 4 of *Arabidopsis thaliana*. Nature 391, 485-488.

Bhojwani, S.S., and Razdan, M.K. (1983). Plant tissue culture : theory and practice. (New York, N.Y.).

Bodemer, C.W. (1959). Observations on the mechanism of induction of supernumerary limbs in adult Triturus viridescens. J Exp Zool 140, 79-99.

Bodnar, W.M., Blackburn, R.K., Krise, J.M., and Moseley, M.A. (2003). Exploiting the complementary nature of LC/MALDI/MS/MS and LC/ESI/MS/MS for increased proteome coverage. J Am Soc Mass Spectrom 14, 971-979.

Bommer, U.A., and Thiele, B.J. (2004). TheA translationally controlled tumour protein (TCTP). Int J Biochem Cell Biol 36, 379-385.

Braun, A. (1873). Gesell. Naturforsch. Freunde z. Berlin, 75.

Briggs, R., and King, T.J. (1952). Transplantation of Living Nuclei From Blastula Cells into Enucleated Frogs' Eggs. Proc Natl Acad Sci USA 38, 455-463.

Brockes, J.P. (1997). Amphibian limb regeneration: rebuilding a complex structure. Science 276, 81-87.

Buitrago, W., and Roop, D.R. (2007). Oct-4: the almighty POUripotent regulator? J Invest Dermatol 127, 260-262.

Bultman, S.J., Gebuhr, T.C., Pan, H., Svoboda, P., Schultz, R.M., and Magnuson, T. (2006). Maternal BRG1 regulates zygotic genome activation in the mouse. Genes Dev 20, 1744-1754.

Byrne, J.A., Simonsson, S., Western, P.S., and Gurdon, J.B. (2003). Nuclei of adult mammalian somatic cells are directly reprogrammed to oct-4 stem cell gene expression by amphibian oocytes. Curr Biol 13, 1206-1213.

Cairns, B.R. (2005). Chromatin remodeling complexes: strength in diversity, precision through specialization. Curr Opin Genet Dev 15, 185-190.

Campbell, K.H., McWhir, J., Ritchie, W.A., and Wilmut, I. (1996). Sheep cloned by nuclear transfer from a cultured cell line. Nature 380, 64-66.

Cans, C., Passer, B.J., Shalak, V., Nancy-Portebois, V., Crible, V., Amzallag, N., Allanic, D., Tufino, R., Argentini, M., Moras, D., Fiucci, G., Goud, B., Mirande, M., Amson, R., and Telerman, A. (2003). Translationally controlled tumor protein acts as a guanine nucleotide dissociation inhibitor on the translation elongation factor eEF1A. Proc Natl Acad Sci USA 100, 13892-13897.

Carter, C., Pan, S., Zouhar, J., Avila, E.L., Girke, T., and Raikhel, N.V. (2004). The vegetative vacuole proteome of *Arabidopsis thaliana* reveals predicted and unexpected proteins. Plant Cell 16, 3285-3303.

Chang, C., and Meyerowitz, E.M. (1986). Molecular cloning and DNA sequence of the *Arabidopsis thaliana* alcohol dehydrogenase gene. Proc Natl Acad Sci USA 83, 1408-1412.

Chanvivattana, Y., Bishopp, A., Schubert, D., Stock, C., Moon, Y.H., Sung, Z.R., and Goodrich, J. (2004). Interaction of Polycomb-group proteins controlling flowering in *Arabidopsis*. Development 131, 5263-5276.

Che, P., Lall, S., Nettleton, D., and Howell, S.H. (2006). Gene expression programs during shoot, root, and callus development in *Arabidopsis* tissue culture. Plant Physiol 141, 620-637.

Chen, S., Takanashi, S., Zhang, Q., Xiong, W., Zhu, S., Peters, E.C., Ding, S., and Schultz, P.G. (2007). Reversine increases the plasticity of lineage-committed mammalian cells. Proc Natl Acad Sci USA 104, 10482-10487.

Chitteti, B.R., Zhang, L., and Peng, Z. (2004). Proteome Dynamics during Cell Dedifferentiation Revealed Using Difference Gel Electrophoresis (DIGE) in *Arabidopsis thaliana*. American Society of Plant Biologists, Orlando, Fl.

Chitteti, B.R., and Peng, Z. (2007). Proteome and phosphoproteome dynamic change during cell dedifferentiation in *Arabidopsis*. Proteomics 7, 1473-1500.

Chivasa, S., Ndimba, B.K., Simon, W.J., Robertson, D., Yu, X.L., Knox, J.P., Bolwell, P., and Slabas, A.R. (2002). Proteomic analysis of the *Arabidopsis thaliana* cell wall. Electrophoresis 23, 1754-1765.

Chong, P.K., Gan, C.S., Pham, T.K., and Wright, P.C. (2006). Isobaric tags for relative and absolute quantitation (iTRAQ) reproducibility: Implication of multiple injections. J Proteome Res 5, 1232-1240.

Christensen, R.N., and Tassava, R.A. (2000). Apical epithelial cap morphology and fibronectin gene expression in regenerating axolotl limbs. Dev Dyn 217, 216-224.

Ciosk, R., DePalma, M., and Priess, J.R. (2006). Translational regulators maintain totipotency in the Caenorhabditis elegans germline. Science 311, 851-853.

Costa, S., and Shaw, P. (2006). 'Open minded' cells: how cells can change fate. TRENDS in Cell Biology Article in Press.

Davis, B.M., Ayers, J.L., Koran, L., Carlson, J., Anderson, M.C., and Simpson, S.B., Jr. (1990). Time course of salamander spinal cord regeneration and recovery of swimming: HRP retrograde pathway tracing and kinematic analysis. Exp Neurol 108, 198-213.

Ding, S., and Schultz, P.G. (2005). Small molecules and future regenerative medicine. Curr Top Med Chem 5, 383-395.

Dougherty, M.K., and Morrison, D.K. (2004). Unlocking the code of 14-3-3. J Cell Sci 117, 1875-1884.

Dovzhenko, A., Dal Bosco, C., Meurer, J., and Koop, H.U. (2003). Efficient regeneration from cotyledon protoplasts in *Arabidopsis thaliana.* Protoplasma 222, 107-111.

Duhamel du Monceau, H.L. La physique des Arbres, ou Il Est Traite de l'Anatomie des plantes et de l'Economie Vegetale pour servir d'Introduction au Traite Complet des Bois et des Forets. P. H. L. Guerin Pub.

Durr, E., Yu, J., Krasinska, K.M., Carver, L.A., Yates, J.R., Testa, J.E., Oh, P., and Schnitzer, J.E. (2004). Direct proteomic mapping of the lung microvascular endothelial cell surface in vivo and in cell culture. Nat Biotechnol 22, 985-992.

Echeverri, K., and Tanaka, E.M. (2002). Ectoderm to mesoderm lineage switching during axolotl tail regeneration. Science 298, 1993-1996.

Eggan, K., Baldwin, K., Tackett, M., Osborne, J., Gogos, J., Chess, A., Axel, R., and Jaenisch, R. (2004). Mice cloned from olfactory sensory neurons. Nature 428, 44-49.

Ezhova, T.A. (2003). Genetic control of totipotency of plant cells *in vitro*. Ontogenez 34, 245-252.

Fenn, J.B., Mann, M., Meng, C.K., Wong, S.F., and Whitehouse, C.M. (1989). Electrospray ionization for mass spectrometry of large biomolecules. Science. 246, 64-71.

Ferl, R.J. (2004). 14-3-3 proteins: regulation of signal-induced events. Physiol Plant 120, 173-178.

Ferro, M., Salvi, D., Brugiere, S., Miras, S., Kowalski, S., Louwagie, M., Garin, J., Joyard, J., and Rolland, N. (2003). Proteomics of the chloroplast envelope membranes from *Arabidopsis thaliana.* Mol Cell Proteomics 2, 325-345.

Ficarro, S.B., McCleland, M.L., Stukenberg, P.T., Burke, D.J., Ross, M.M., Shabanowitz, J., Hunt, D.F., and White, F.M. (2002). Phosphoproteome analysis by mass spectrometry and its application to Saccharomyces cerevisiae. Nat Biotechnol 20, 301-305.

Fountoulakis, M., Berndt, P., Langen, H., and Suter, L. (2002). The rat liver mitochondrial proteins. Electrophoresis 23, 311-328.

Fukao, Y., Hayashi, M., and Nishimura, M. (2002). Proteomic analysis of leaf peroxisomal proteins in greening cotyledons of *Arabidopsis thaliana.* Plant Cell Physiol 43, 689-696.

Gallardo, K., Job, C., Groot, S.P., Puype, M., Demol, H., Vandekerckhove, J., and Job, D. (2001). Proteomic analysis of *Arabidopsis* seed germination and priming. Plant Physiol 126, 835-848.

Gan, C.S., Chong, P.K., Pham, T.K., and Wright, P.C. (2007). Technical, Experimental, and Biological Variations in Isobaric Tags for Relative and Absolute Quantitation (iTRAQ). J Proteome Res 6, 821-827.

Gautheret, R.J. (1985). History of Plant Tissue and Cell Culture: A Personal Account. In Cell Culture and Somatic Cell Genetics of Plants. (Orlando, Florida: Academic Press Inc.,).

Gerber, S.A., Rush, J., Stemman, O., Kirschner, M.W., and Gygi, S.P. (2003). Absolute quantification of proteins and phosphoproteins from cell lysates by tandem MS. Proc Natl Acad Sci USA 100, 6940-6945.

Gonda, K., Fowler, J., Katoku-Kikyo, N., Haroldson, J., Wudel, J., and Kikyo, N. (2003). Reversible disassembly of somatic nucleoli by the germ cell proteins FRGY2a and FRGY2b. Nat Cell Biol 5, 205-210.

Gorg, A., Obermaier, C., Boguth, G., Harder, A., Scheibe, B., Wildgruber, R., and Weiss, W. (2000). The current state of two-dimensional electrophoresis with immobilized pH gradients. Electrophoresis 21, 1037-1053.

Grafi, G. (2004). How cells dedifferentiate: a lesson from plants. Dev Biol 268, 1-6.

Grafi, G., and Avivi, Y. (2004). Stem cells: a lesson from dedifferentiation. Trends Biotechnol 22, 388-389.

Graves, P.R., and Haystead, T.A. (2002). Molecular biologist's guide to proteomics. Microbiol Mol Biol Rev 66, 39-63; table of contents.

Gurdon, J.B. (1962). The developmental capacity of nuclei taken from intestinal epithelium cells of feeding tadpoles. J Embryol Exp Morphol 10, 622-640.

Gygi, S.P., Rochon, Y., Franza, B.R., and Aebersold, R. (1999a). Correlation between protein and mRNA abundance in yeast. Mol Cell Biol 19, 1720-1730.

Gygi, S.P., Corthals, G.L., Zhang, Y., Rochon, Y., and Aebersold, R. (2000). Evaluation of Two-dimensional Gel Electrophoresis-based proteome analysis technology. Proc Natl Acad Sci USA 97, 9390-9395.

Gygi, S.P., Rist, B., Gerber, S.A., Turecek, F., Gelb, M.H., and Aebersold, R. (1999b). Quantitative analysis of complex protein mixtures using isotope-coded affinity tags. Nat Biotechnol 17, 994-999.

Han, D.K., Eng, J., Zhou, H., and Aebersold, R. (2001). Quantitative profiling of differentiation-induced microsomal proteins using isotope-coded affinity tags and mass spectrometry. Nat Biotechnol 19, 946-951.

Harrisingh, M.C., Perez-Nadales, E., Parkinson, D.B., Malcolm, D.S., Mudge, A.W., and Lloyd, A.C. (2004). The Ras/Raf/ERK signalling pathway drives Schwann cell dedifferentiation. Embo J 23, 3061-3071.

Hart, C., Schulenberg, B., Steinberg, T.H., Leung, W.Y., and Patton, W.F. (2003). Detection of glycoproteins in polyacrylamide gels and on electroblots using Pro-Q Emerald 488 dye, a fluorescent periodate Schiff-base stain. Electrophoresis 24, 588-598.

Hay, E.D., and Fischman, D.A. (1961). Origin of the blastema in regenerating limbs of the newt Triturus viridescens. An autoradiographic study using tritiated thymidine to follow cell proliferation and migration. Dev Biol 3, 26-59.

Hemerly, A.S., Ferreira, P., de Almeida Engler, J., Van Montagu, M., Engler, G., and Inze, D. (1993). cdc2a expression in *Arabidopsis* is linked with competence for cell division. Plant Cell 5, 1711-1723.

Hoogland, C., Sanchez, J.C., Tonella, L., Binz, P.A., Bairoch, A., Hochstrasser, D.F., and Appel, R.D. (2000). The 1999 SWISS-2DPAGE database update. Nucleic Acids Res 28, 286-288.

Hsieh, J., and Gage, F.H. (2005). Chromatin remodeling in neural development and plasticity. Curr Opin Cell Biol 17, 664-671.

Hurkman, W.J., and Tanaka, C.K. (1986). Solubilization of Plant Membrane Proteins for Analysis by Two-Dimensional Gel Electrophoresis. Plant Physiol 81, 802-806.

Jack, T. (2001). Plant development going MADS. Plant Mol Biol 46, 515-520.

Jamet, E., Durr, A., Parmentier, Y., Criqui, M.C., and Fleck, J. (1990). Is ubiquitin involved in the dedifferentiation of higher plant cells? Cell Differ Dev 29, 37-46.

Johnson, S.L., and Weston, J.A. (1995). Temperature-sensitive mutations that cause stage-specific defects in Zebrafish fin regeneration. Genetics 141, 1583-1595.

Jones, A.M., Bennett, M.H., Mansfield, J.W., and Grant, M. (2006). Analysis of the defence phosphoproteome of *Arabidopsis thaliana* using differential mass tagging. Proteomics 6, 4155-4165.

Jones, R.B., Gordus, A., Krall, J.A., and MacBeath, G. (2006). A quantitative protein interaction network for the ErbB receptors using protein microarrays. Nature 439, 168-174.

Karas, M., and Hillenkamp, F. (1988). Laser desorption ionization of proteins with molecular masses exceeding 10,000 daltons. Anal Chem 60, 2299-2301.

Kerstetter, R.A., and Hake, S. (1997). Shoot Meristem Formation in Vegetative Development. Plant Cell 9, 1001-1010.

Kim, S.T., Cho, K.S., Jang, Y.S., and Kang, K.Y. (2001). Two-dimensional electrophoretic analysis of rice proteins by polyethylene glycol fractionation for protein arrays. Electrophoresis 22, 2103-2109.

Klebes, A., Sustar, A., Kechris, K., Li, H., Schubiger, G., and Kornberg, T.B. (2005). Regulation of cellular plasticity in Drosophila imaginal disc cells by the Polycomb group, trithorax group and lama genes. Development 132, 3753-3765.

Kleffmann, T., Russenberger, D., von Zychlinski, A., Christopher, W., Sjolander, K., Gruissem, W., and Baginsky, S. (2004). The *Arabidopsis thaliana* chloroplast proteome reveals pathway abundance and novel protein functions. Curr Biol 14, 354-362.

Klose, J., and Kobalz, U. (1995). Two-dimensional electrophoresis of proteins: an updated protocol and implications for a functional analysis of the genome. Electrophoresis 16, 1034-1059.

Korenberg, J.R., and Freedlender, E.F. (1974). Giemsa technique for the detection of sister chromatid exchanges. Chromosoma 48, 355-360.

Koukalova, B., Fojtova, M., Lim, K.Y., Fulnecek, J., Leitch, A.R., and Kovarik, A. (2005). Dedifferentiation of tobacco cells is associated with ribosomal RNA gene hypomethylation, increased transcription, and chromatin alterations. Plant Physiol 139, 275-286.

Krikorian, A. (1995). Hormones in Tissue Culture and Micropropagation. In: Plant Hormones. (Norwell, MA.: Kluwer Academic Publishers).

Kruft, V., Eubel, H., Jansch, L., Werhahn, W., and Braun, H.P. (2001). Proteomic approach to identify novel mitochondrial proteins in *Arabidopsis*. Plant Physiol 127, 1694-1710.

Kumar, A., Velloso, C.P., Imokawa, Y., and Brockes, J.P. (2000). Plasticity of retrovirus-labelled myotubes in the newt limb regeneration blastema. Dev Biol 218, 125-136.

Kuruvilla, F.G., Shamji, A.F., Sternson, S.M., Hergenrother, P.J., and Schreiber, S.L. (2002). Dissecting glucose signalling with diversity-oriented synthesis and small-molecule microarrays. Nature 416, 653-657.

Laibach, F. (1907). Bot Centbl Beihefte (I), 191–210.

Larsson, U.K., Jergil, B., and Andersson, B. (1983). Changes in the lateral distribution of the light-harvesting chlorophyll-a/b--protein complex induced by its phosphorylation. Eur J Biochem 136, 25-29.

Larsen, M.R., Thingholm, T.E., Jensen, O.N., Roepstorff, P., and Jorgensen, T.J. (2005). Highly selective enrichment of phosphorylated peptides from peptide mixtures using titanium dioxide microcolumns. Mol Cell Proteomics 4, 873-886.

Lee, N., Maurange, C., Ringrose, L., and Paro, R. (2005). Suppression of Polycomb group proteins by JNK signalling induces transdetermination in Drosophila imaginal discs. Nature 438, 234-237.

Lewin, B. (1994). Chromatin and gene expression: constant questions, but changing answers. Cell 79, 397-406.

Li, J., Ishii, T., Feinstein, P., and Mombaerts, P. (2004). Odorant receptor gene choice is reset by nuclear transfer from mouse olfactory sensory neurons. Nature 428, 393-399.

Link, A.J., Eng, J., Schieltz, D.M., Carmack, E., Mize, G.J., Morris, D.R., Garvik, B.M., and Yates, J.R., 3rd. (1999). Direct analysis of protein complexes using mass spectrometry. Nat Biotechnol 17, 676-682.

Liu, X.D., and Shen, Y.G. (2005). Salt-induced redox-independent phosphorylation of light harvesting chlorophyll a/b proteins in Dunaliella salina thylakoid membranes. Biochim Biophys Acta 1706, 215-219.

Lo, D.C., Allen, F., and Brockes, J.P. (1993). Reversal of muscle differentiation during urodele limb regeneration. Proc Natl Acad Sci USA 90, 7230-7234.

MacBeath, G., and Schreiber, S.L. (2000). Printing proteins as microarrays for high-throughput function determination. Science 289, 1760-1763.

Maccarrone, G., Kolb, N., Teplytska, L., Birg, I., Zollinger, R., Holsboer, F., and Turck, C.W. (2006). Phosphopeptide enrichment by IEF. Electrophoresis 27, 4585-4595.

Mann, D.L., Chase, G.W., Jr., and Eitenmiller, R.R. (2001). Liquid chromatographic analysis of vitamin B6 in soy-based infant formula. J AOAC Int 84, 1593-1599.

Marsden, M.P., and Laemmli, U.K. (1979). Metaphase chromosome structure: evidence for a radial loop model. Cell 17, 849-858.

Maves, L., and Schubiger, G. (2003). Transdetermination in Drosophila imaginal discs: a model for understanding pluripotency and selector gene maintenance. Curr Opin Genet Dev 13, 472-479.

McGann, C.J., Odelberg, S.J., and Keating, M.T. (2001). Mammalian myotube dedifferentiation induced by newt regeneration extract. Proc Natl Acad Sci USA 98, 13699-13704.

Meyer, Y., Grosset, J., Chartier, Y., and Cleyet-Marel, J.C. (1988). Preparation by two-dimensional electrophoresis of proteins for antibody production: antibodies

against proteins whose synthesis is reduced by auxin in tobacco mesophyll protoplasts. Electrophoresis 9, 704-712.

Meyerowitz, E.M. (2001). Prehistory and history of *Arabidopsis* research. Plant Physiol 125, 15-19.

Meyerowitz, E.M. (2002). Plants compared to animals: the broadest comparative study of development. Science 295, 1482-1485.

Nanduri, B., Lawrence, M.L., Vanguri, S., and Burgess, S.C. (2005). Proteomic analysis using an unfinished bacterial genome: the effects of subminimum inhibitory concentrations of antibiotics on Mannheimia haemolytica virulence factor expression. Proteomics 5, 4852-4863.

Noir, S., Brautigam, A., Colby, T., Schmidt, J., and Panstruga, R. (2005). A reference map of the *Arabidopsis thaliana* mature pollen proteome. Biochem Biophys Res Commun 337, 1257-1266.

O'Farrell, P.H. (1975). High resolution two-dimensional electrophoresis of proteins. J Biol Chem 250, 4007-4021.

O'Kane Jr, S.L., and Al-Shehbaz, I.A. (1997). A synopsis of *Arabidopsis* (Brassicaceae). Novon 7, 323–327.

Old, W.M., Meyer-Arendt, K., Aveline-Wolf, L., Pierce, K.G., Mendoza, A., Sevinsky, J.R., Resing, K.A., and Ahn, N.G. (2005). Comparison of label-free methods for quantifying human proteins by shotgun proteomics. Mol Cell Proteomics 4, 1487-1502.

Olins, A.L., and Olins, D.E. (1974). Spheroid chromatin units (v bodies). Science 183, 330-332.

Opiteck, G.J., Lewis, K.C., Jorgenson, J.W., and Anderegg, R.J. (1997). Comprehensive on-line LC/LC/MS of proteins. Anal Chem 69, 1518-1524.

Pappin, D.J., Hojrup, P., and Bleasby, A.J. (1993). Rapid identification of proteins by peptide-mass fingerprinting. Curr Biol 3, 327-332.

Parry, G., and Estelle, M. (2006). Auxin receptors: a new role for F-box proteins. Curr Opin Cell Biol 18, 152-156.

Pasquali, C., Fialka, I., and Huber, L.A. (1999). Subcellular fractionation, electromigration analysis and mapping of organelles. J Chromatogr B Biomed Sci Appl 722, 89-102.

Patterson, S.D. (2000). Proteomics: the industrialization of protein chemistry. Curr Opin Biotechnol 11, 413-418.

Patton, W.F. (2000). A thousand points of light: the application of fluorescence detection technologies to Two-Dimensional Gel Electrophoresis and proteomics. Electrophoresis 21, 1123-1144.

Peng, J., and Gygi, S.P. (2001). Proteomics: the move to mixtures. J Mass Spectrom 36, 1083-1091.

Rice, J.C., and Allis, C.D. (2001). Code of silence. Nature 414, 258-261.

Richards, E.J., and Elgin, S.C. (2002). Epigenetic codes for heterochromatin formation and silencing: rounding up the usual suspects. Cell 108, 489-500.

Ringrose, L., and Paro, R. (2004). Epigenetic regulation of cellular memory by the Polycomb and Trithorax group proteins. Annu Rev Genet 38, 413-443.

Robertson, D., Mitchell, G.P., Gilroy, J.S., Gerrish, C., Bolwell, G.P., and Slabas, A.R. (1997). Differential extraction and protein sequencing reveals major differences in patterns of primary cell wall proteins from plants. J Biol Chem 272, 15841-15848.

Salo, E. (2006). The power of regeneration and the stem-cell kingdom: freshwater planarians (Platyhelminthes). Bioessays 28, 546-559.

Schaner, C.E., and Kelly, W.G. (2006). Germline Chromatin. (The C. elegans Research Community, WormBook). 10.1895/wormbook.1.73.1.

Schmidt, R., West, J., Love, K., Lenehan, Z., Lister, C., Thompson, H., Bouchez, D., and Dean, C. (1995). Physical map and organization of *Arabidopsis thaliana* chromosome 4. Science 270, 480-483.

Schmidt, M.W., Houseman, A., Ivanov, A.R., and Wolf, D.A. (2007). Comparative proteomic and transcriptomic profiling of the fission yeast Schizosaccharomyces pombe. Mol Syst Biol 3, 79.

Schubert, D., Clarenz, O., and Goodrich, J. (2005). Epigenetic control of plant development by Polycomb-group proteins. Curr Opin Plant Biol 8, 553-561.

Schulenberg, B., Arnold, B., and Patton, W.F. (2003). An improved mechanically durable electrophoresis gel matrix that is fully compatible with fluorescence-based protein detection technologies. Proteomics 3, 1196-1205.

Schulenberg, B., Goodman, T.N., Aggeler, R., Capaldi, R.A., and Patton, W.F. (2004). Characterization of dynamic and steady-state protein phosphorylation using a

fluorescent phosphoprotein gel stain and mass spectrometry. Electrophoresis 25, 2526-2532.

Schweitzer, B., Predki, P., and Snyder, M. (2003). Microarrays to characterize protein interactions on a whole-proteome scale. Proteomics 3, 2190-2199.

Schweitzer, B., Roberts, S., Grimwade, B., Shao, W., Wang, M., Fu, Q., Shu, Q., Laroche, I., Zhou, Z., Tchernev, V.T., Christiansen, J., Velleca, M., and Kingsmore, S.F. (2002). Multiplexed protein profiling on microarrays by rolling-circle amplification. Nat Biotechnol 20, 359-365.

Sell, S. (1993). Environ. Health Perspect. Suppl 5, 15-26.

Service, R.F. (2001). Proteomics. Searching for recipes for protein chips. Science 294, 2080-2082.

Seydoux, G., and Braun, R.E. (2006). Pathway to totipotency: lessons from germ cells. Cell 127, 891-904.

Skoog, F., and Miller, C.O. (1957). Chemical regulation of growth and organ formation in plant tissues cultured *in vitro*. Symp Soc Exp Biol 54, 118-130.

Smolka, M., Zhou, H., and Aebersold, R. (2002). Quantitative protein profiling using Two-Dimensional Gel Electrophoresis, isotope-coded affinity tag labeling, and mass spectrometry. Mol Cell Proteomics 1, 19-29.

Solter, D. (2000). Mammalian cloning: advances and limitations. Nat Rev Genet 1, 199-207.

Stasyk, T., Morandell, S., Bakry, R., Feuerstein, I., Huck, C.W., Stecher, G., Bonn, G.K., and Huber, L.A. (2005). Quantitative detection of phosphoproteins by combination of two-dimensional difference gel electrophoresis and phosphospecific fluorescent staining. Electrophoresis 26, 2850-2854.

Steen, T.P. (1968). Stability of chondrocyte differentiation and contribution of muscle to cartilage during limb regeneration in the axolotl (Siredon mexicanum). J Exp Zool 167, 49-78.

Steinberg, T.H., Agnew, B.J., Gee, K.R., Leung, W.Y., Goodman, T., Schulenberg, B., Hendrickson, J., Beechem, J.M., Haugland, R.P., and Patton, W.F. (2003). Global quantitative phosphoprotein analysis using Multiplexed Proteomics technology. Proteomics 3, 1128-1144.

Strahl, B.D., Grant, P.A., Briggs, S.D., Sun, Z.W., Bone, J.R., Caldwell, J.A., Mollah, S., Cook, R.G., Shabanowitz, J., Hunt, D.F., and Allis, C.D. (2002). Set2 is a

nucleosomal histone H3-selective methyltransferase that mediates transcriptional repression. Mol Cell Biol 22, 1298-1306.

Straube, W.L., Brockes, J.P., Drechsel, D.N., and Tanaka, E.M. (2004). Plasticity and reprogramming of differentiated cells in amphibian regeneration: partial purification of a serum factor that triggers cell cycle re-entry in differentiated muscle cells. Cloning Stem Cells 6, 333-344.

Sugiyama, M. (1999). Organogenesis *in vitro*. Curr Opin Plant Biol 2, 61-64.

Tabb, D.L., McDonald, W.H., and Yates, J.R., 3rd. (2002). DTASelect and Contrast: tools for assembling and comparing protein identifications from shotgun proteomics. J Proteome Res 1, 21-26.

Taub, R. (2004). Liver regeneration: from myth to mechanism. Nat Rev Mol Cell Biol 5, 836-847.

Thoma, F., Koller, T., and Klug, A. (1979). Involvement of histone H1 in the organization of the nucleosome and of the salt-dependent superstructures of chromatin. J Cell Biol 83, 403-427.

Thornton, C.S. (1957). The effect of apical cap removal on limb regeneration in Amblystoma larvae. J Exp Zool 134, 357-381.

Thornton, C.S., and Thronton, M.T. (1965). The regeneration of accessory limb parts following epidermal cap transplantation in urodeles. Experientia 21, 146-148.

Tyers, M., and Mann, M. (2003). From genomics to proteomics. Nature 422, 193-197.

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

Vasil, V., and Vasil, I.K. (1974). Regeneration of tobacco and petunia plants from protoplasts and culture of corn protoplasts. *In Vitro* 10, 83-96.

Wagner, D., and Meyerowitz, E.M. (2002). SPLAYED, a novel SWI/SNF ATPase homolog, controls reproductive development in *Arabidopsis*. Curr Biol 12, 85-94.

Warashina, M., Min, K.H., Kuwabara, T., Huynh, A., Gage, F.H., Schultz, P.G., and Ding, S. (2006). A synthetic small molecule that induces neuronal differentiation of adult hippocampal neural progenitor cells. Angew Chem Int Ed Engl 45, 591-593.

Washburn, M.P., Wolters, D., and Yates, J.R., 3rd. (2001). Large-scale analysis of the yeast proteome by multidimensional protein identification technology. Nat Biotechnol 19, 242-247.

Weimann, J.M., Johansson, C.B., Trejo, A., and Blau, H.M. (2003). Stable reprogrammed heterokaryons form spontaneously in Purkinje neurons after bone marrow transplant. Nat Cell Biol 5, 959-966.

Wiese, S., Reidegeld, K.A., Meyer, H.E., and Warscheid, B. (2007). Protein labeling by iTRAQ: A new tool for quantitative mass spectrometry in proteome research. Proteomics 7, 340-350.

Wilkins, M.R., Sanchez, J.C., Gooley, A.A., Appel, R.D., Humphery-Smith, I., Hochstrasser, D.F., and Williams, K.L. (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.

Williams, L., Zhao, J., Morozova, N., Li, Y., Avivi, Y., and Grafi, G. (2003). Chromatin reorganization accompanying cellular dedifferentiation is associated with modifications of histone H3, redistribution of HP1, and activation of E2F-target genes. Dev Dyn 228, 113-120.

Wolffe, A.P., and Pruss, D. (1996). Targeting chromatin disruption: Transcription regulators that acetylate histones. Cell 84, 817-819.

Wolters, D.A., Washburn, M.P., and Yates, J.R., 3rd. (2001). An automated multidimensional protein identification technology for shotgun proteomics. Anal Chem 73, 5683-5690.

Xi. (2006). Recent Developments of *Arabidopsis thaliana* Proteomics Research. the Journal of American Science 2.

Yarm, F.R. (2002). Plk phosphorylation regulates the microtubule-stabilizing protein TCTP. Mol Cell Biol 22, 6209-6221.

Yates, J.R., 3rd, Carmack, E., Hays, L., Link, A.J., and Eng, J.K. (1999). Automated protein identification using microcolumn liquid chromatography-tandem mass spectrometry. Methods Mol Biol 112, 553-569.

Zhang, B., VerBerkmoes, N.C., Langston, M.A., Uberbacher, E., Hettich, R.L., and Samatova, N.F. (2006). Detecting differential and correlated protein expression in label-free shotgun proteomics. J Proteome Res 5, 2909-2918.

Zhao, J., Morozova, N., Williams, L., Libs, L., Avivi, Y., and Grafi, G. (2001). Two phases of chromatin decondensation during dedifferentiation of plant cells:

distinction between competence for cell fate switch and a commitment for S phase. J Biol Chem 276, 22772-22778.

Zottoli, S.J., Bentley, A.P., Feiner, D.G., Hering, J.R., Prendergast, B.J., and Rieff, H.I. (1994). Spinal cord regeneration in adult goldfish: implications for functional recovery in vertebrates. Prog Brain Res 103, 219-228.