

# Challenges in proteome analyses of tropical plants

Tiago S. Balbuena<sup>1+</sup>, Leonardo L. C. Dias<sup>2, 3+</sup>, Mariana L. B. Martins<sup>3</sup>, Tatiana B. Chiquieri<sup>4</sup>, Claudete Santa-Catarina<sup>5</sup>, Eny I. S. Floh<sup>6</sup>, Vanildo Silveira<sup>4\*</sup>

<sup>1</sup> Christopher S. Bond Life Sciences Center – University of Missouri- Columbia-USA.

<sup>2</sup> Tissue Culture Laboratory II – BIOAGRO – Federal University of Viçosa, Viçosa-MG, Brazil

<sup>3</sup> Genetic Biochemistry Laboratory – Department of Genetic and Immunology – Institute of Biological Sciences – Federal University of Minas Gerais (ICB-UFMG) Belo Horizonte – MG, Brazil

<sup>4</sup> Laboratório de Biotecnologia/Centro de Biociências e Biotecnologia/ Universidade Estadual do Norte Fluminense Darcy Ribeiro, Av. Alberto Lamego, 2000, Pq. Califórnia. CEP: 28013-602. Campos dos Goytacazes – RJ, Brazil

<sup>5</sup> Laboratório de Biologia Celular e Tecidual – CBB, Universidade Estadual do Norte Fluminense. Campos dos Goytacazes – RJ, Brazil

<sup>6</sup> Plant Cell Biology Laboratory (BIOCEL) Institute of Biosciences University of São Paulo (IB-USP), São Paulo – SP, Brazil

\* Corresponding author: Fone/Fax +55-22-2739-7031; e-mail: vanildo@uenf.br

+ These authors contributed equally to this work

Received: 29 March 2011; Accepted: 08 August 2011

## ABSTRACT

Genome sequencing of various organisms allow global analysis of gene expression, providing numerous clues on the biological function and involvement in the biological processes studied. Proteomics is a branch of molecular biology and biotechnology that has undergone considerable development in the post-genomic era. Despite the recent significant advancements in proteomics techniques, still there is much to be improved. Due to peculiarities to the plant kingdom, proteomics approaches require adaptations, so as to improve efficiency and accuracy of results in plants. Data generated by proteomics can substantially contribute to the understanding and monitoring of plant physiological events and development of biotechnological strategies. Especially for tropical species, challenges are even greater, in the light of the abundance of secondary metabolites, as well as of the lack of complete genome sequences. This review discusses current topics in proteomics concerning challenges and perspectives, with emphasis on the proteomics of tropical plant species.

**Key words:** 2-DE, mass spectrometry, recalcitrant species, proteomics, unknown genome

## RESUMO

O sequenciamento dos genomas de diversos organismos tornou praticável a análise global da expressão gênica, fornecendo numerosas pistas quanto à função biológica e envolvimento de inúmeros genes nos processos biológicos estudados. A proteômica é um dos ramos da biologia molecular e da biotecnologia que passou por expressivo desenvolvimento na era pós-genômica. Apesar dos recentes avanços nas técnicas em proteômica, muito resta a ser aprimorado. Devido às particularidades do reino vegetal, várias abordagens devem ser adaptadas no sentido de aumentar a eficiência e a acurácia de resultados das análises proteômicas em plantas. Nesse sentido, os dados gerados pela proteômica podem contribuir substancialmente para a compreensão e acompanhamento de eventos fisiológicos e para o desenvolvimento de estratégias biotecnológicas. No que tange

às espécies tropicais, as dificuldades são ainda maiores, devido ao fato de serem ricas em metabólitos secundários e ainda não terem seus genomas totalmente sequenciados. Esta revisão trata de tópicos atuais em estudos de proteômica de plantas, desafios e perspectivas, com ênfase em espécies tropicais.

**Palavras-chave:** 2-DE, espectrometria de massas, espécies recalcitrantes, genomas desconhecidos, proteômica.

## INTRODUCTION

In recent years several articles have reported the complete sequencing of genetic codes of various organisms. However, obtaining the nucleotide sequences of an organism's genome represents a mere first step in the more comprehensive understanding of gene expression regulation. In the "omics" context, great emphasis has been directed towards functional genomics, whose aim is to identify the sequence and function of genes and proteins based on the investigation of the transcriptome and the proteome.

Transcriptome analysis provides important information about the gene expression of an organism in a given state; however, it does not directly reflect the protein expression of this organism (Chen and Harmon, 2006). Several mechanisms are involved in the regulation of the protein synthesis (Figure 1). During protein synthesis, post-transcriptional and post-translational modifications may regulate the spatial and temporal expression and conformation of proteins, modulating distinct proteins classes which, biochemically and structurally, may play different roles in metabolic pathways and proteome composition of an organism.

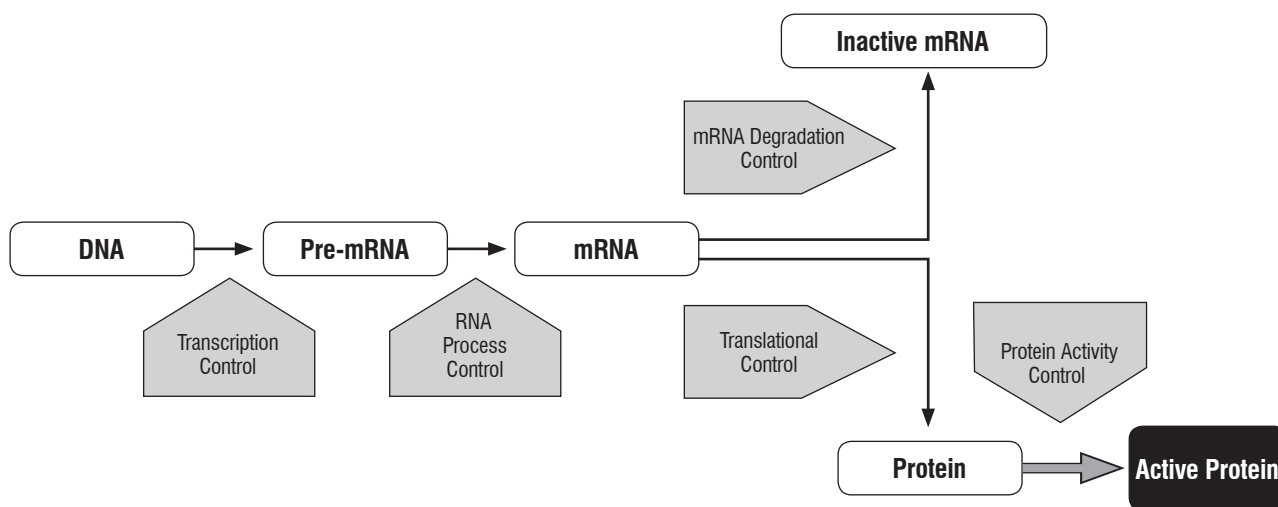


Figure 1. Control mechanisms that act in protein synthesis where a single gene gives rise to multiple protein conformation and different functions.

On the other hand, proteomics is the study of the proteins set in a cell in a given situation. Proteomics has gained the status of an essential approach towards the investigation of the roles that proteins play in cell metabolism. The term *proteome* was first defined by Mark Wilkins in 1994, at the first 2-DE meeting in Siena, Italy, and was subsequently mentioned in publication, in 1995 (Wasinger et al., 1995). Proteome refers to the "total protein complement of a genome" or, in the case of a pluricellular organism, to the protein complement of a tissue, expressed by a genome,

under certain physiological conditions (Wilkins et al., 1996). Proteomics has been regarded as one element of genomic investigations, the post-genomic era, as a direct result of the advancements made possible by large-scale DNA sequencing (Mann et al., 2001).

Proteomic analysis provides the elements required for the study of proteins' properties (expression level, post-translational modifications, interactions etc.) and for the systematic evaluation of protein expression in tissues

exposed to an array of different physiological scenarios, such as pathological and healthy, treated and untreated, resistant and susceptible conditions (Wasinger et al., 1995).

In its very beginning, the proteomics — like any other discovery in science technology, such as genomic sequencing, microarray analysis and metabolic profiling — was focused on ambitious projects that aimed at separating and detecting as many proteins as possible from one source, which could allow these proteins to be computationally catalogued. These studies were in vogue between 1995 and 2000. Nevertheless, in its second wave, proteomics now aims at the global characterization of the protein functions, contributing to a global understanding and integrated knowledge of different functions in living cells, *i.e.* functional study of proteins and their involvement in cellular pathways. This kind of information cannot be obtained only from the study of genes, since proteins are major agents responsible for the phenotypes of cells.

Moreover, proteomic analysis generates knowledge about cell signaling pathways as well as sets of regulatory proteins. Additionally, proteomics is used to identify, quantify and study the post-translational modifications of proteins in a cell, tissue or organism under a given physiological condition. Also, it may be employed to gather important supplementary information about the physiological and pathophysiological states of cells and organisms (Mann and Jensen, 2003; Newton et al., 2004).

Proteomic characterization is an important complementary tool of genomic studies, since the dynamics of proteins in a living system is influenced by a variety of internal and external factors that determine structural and conformational protein changes. Proteomic analysis offers the opportunity to simultaneously examine changes and classify temporal patterns of protein accumulation in different developmental stages and/or growth conditions.

Nowadays, proteomics has numerous applications, including: i) study of differential protein expression, providing important information on cell signaling and development of organisms; ii) study of post-translational modifications; iii) protein-protein interactions; iv) structural proteomics; v) functional proteomics and vi) computational proteomics.

Proteome analysis involves further difficulties beyond that of genomes, because the proteome of every living cell is

dynamic and changes in response to intra- and extracellular signal molecules. Understanding the proteome is a complex task, because in most cases the protein does not act alone, since it may interact with other proteins (Gazzana and Borlak, 2007). Besides, the number of expressed proteins is much larger than the number of genes, and there is no available methodology that allows the amplification of proteins with low number of copies. Proteins are also physically and chemically more complex than nucleic acids, with interactions and functions that depend on their three-dimensional structure (Jorin et al., 2006). Proteins can undergo post-translational modifications such as phosphorylation, methylation, acetylation and proteolytic cleavage, among others. For these reasons, predicting the final structure and function of different proteins from proteomics analysis is a new and exciting research challenge.

Recent studies indicate that the correlation between mRNA and protein levels is inexplicably low (Jorin et al., 2006). In fact, proteomics cannot provide complete information to understand the mechanism that sets the pattern of protein expression, in part because key regulatory proteins such as transcription factors and signaling proteins are hardly visualized in 2-DE gels due to low abundance. However, parallel studies of transcriptome and proteomics may provide key information about transcription processes and post translational events (Zivy and Vienne, 2000).

Plant proteomics studies were first carried out in maize (Touzet et al., 1996). The study of proteins expressed in plant tissues and organs has been applied to monitor changes or assess the influence of environmental stimulus on patterns of protein expression (Agrawal et al., 2005), variations in response to physiological events (Gallardo et al., 2001), variations in host-pathogen interaction, investigations for plant breeding improvements, including the study of defense proteins in plants (Elvira et al., 2008).

In recent years, several studies have been directed to the characterization of the dynamics of proteins involved in plant development, coupled with genome and transcriptome characterization (Roberts, 2002; Heazlewood and Millar, 2003; Chen and Harmon, 2006; Rossignol et al., 2006). Most plant proteomics studies have investigated species native to temperate regions, mostly grain species, while tropical species remain mainly underexplored.

This review addresses current topics in the proteomics of tropical species - the major challenges hindering the progress of the studies and the current scientific scenery of tropical plants proteomics.

## TROPICAL PLANTS PROTEOMICS: CHALLENGES AND LIMITATIONS

Previous proteomics studies have focused mostly on economically important species such as rice, wheat and grape, as well as on some model plants like *Arabidopsis*. However, genomic and proteomic studies in tropical plants are very scarce. Tropical plants species represent 2/3 of total number green plants of the world and have a huge potential for pharmaceutical, agronomical, reforestation and other applications.

For proteomics studies, tropical plants present a recalcitrant nature because they have many interfering

compounds and in general they do not have sequenced genomes. Tropical plants are exposed to high temperatures and excess of radiation, presenting an efficient apparatus to fight oxidative stress (Wahid et al., 2007). These plants can behave dynamically in response to environmental stimuli, with rapid induction and suppression of secondary metabolite production, as strategies to cope with oxidative stress (Metlen et al., 2005).

Although tropical plants proteomics has been poorly studied, some studies have been published whose aim was to separate and detect as many proteins as possible, in order to characterize protein profiles. Further studies have been developed towards understanding the proteome changes of these tropical plants, especially in the embryogenic process and response to pathogens, anaerobic conditions and abiotic stress, like low temperatures. Knowing how and when these events occur is a very important aspect concerning the understanding of the physiology/pathophysiology of tropical plants (Table 1).

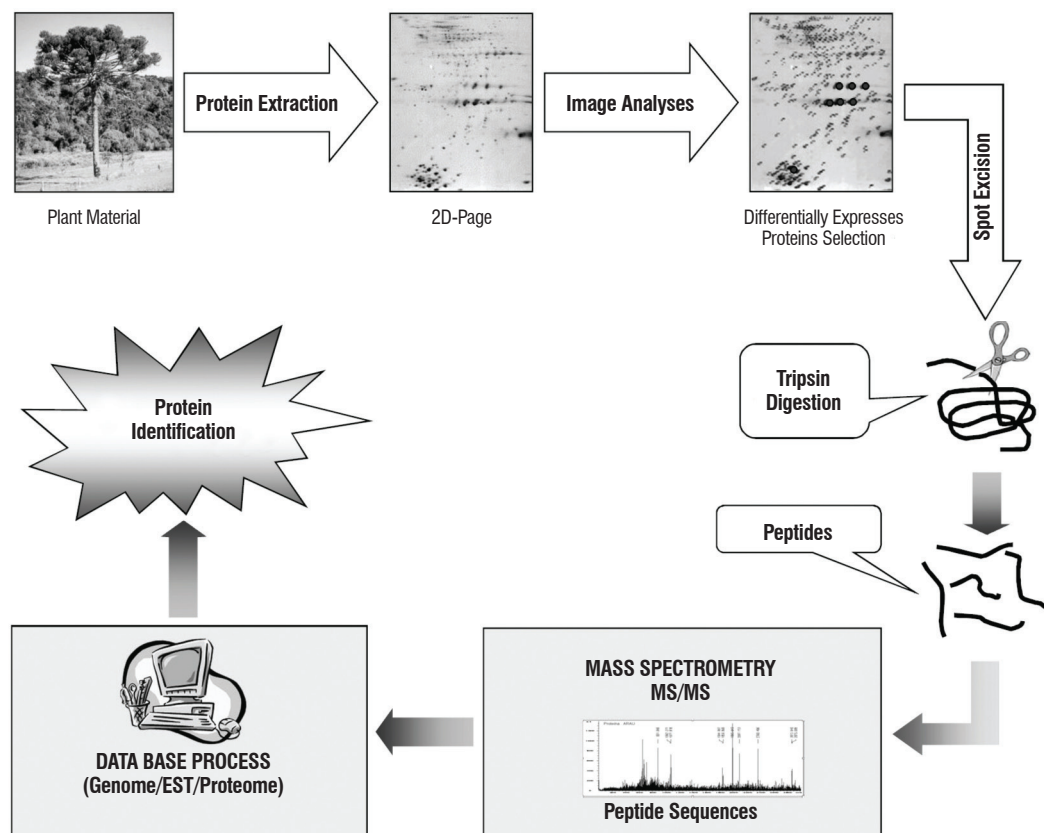
Table 1. Proteome studies in tropical species. Material employed, conditions studies and platforms.

Species	Material	Evaluated condition	Platform	Reference
<i>Acca sellowiana</i>	Somatic embryos culture	-	2-DE/MALDI-MS	Cangahuala-Inocente et al., 2009
<i>Aloe vera L.</i>	Leaves	-	2-DE	He and Wang 2008
<i>Araucaria angustifolia</i>	Embryo and gametophyte seeds	-	2-DE/LC-ESI-MSMS	Balbuena et al., 2009
<i>Araucaria angustifolia</i>	Zygotic embryos	-	2-DE/ MALDI-MS	Silveira et al., 2008
<i>Araucaria angustifolia</i>	Mature and germinated Seeds	-	2-DE/LC-ESI-MSMS	Balbuena et al., 2011
<i>Citrus clementina</i> cv <i>Tomatera</i>	Leaves	<i>Tetranychus urticae</i> infestation/ methyl jasmonate	2-DE/LC-ESI-MSMS	Maserti et al., 2010
<i>Citrus limon</i>	Flavedo tissue	-	2-DE/LC-ESI-MSMS	Pignataro et al., 2010
<i>Citrus paradisi</i> Macf and <i>Citrus reticulata</i> Blanco	Flavedo and juice vesicle tissue	Anaerobic stress	2-DE/LC-ESI-MS	Shi et al., 2008
<i>Citrus sinensis</i>	Leaves	Stress by plant natriuretic peptides of <i>Xanthomonas</i> <i>axonopodis</i> pv. <i>Citri</i>	2-DE/MALDI-MS	Garavaglia et al., 2010
<i>Citrus sinensis</i>	Embryogenic callus culture	Different stages of somatic embryos	2-DE/MALDI-MS	Pan et al., 2009
<i>Citrus sinensis</i>	Flesh fruit	Ripening time	2-DE/LC-ESI-MSMS	Muccilli et al., 2009
<i>Citrus reticulata</i>	Juice sacs	Low temperature storage	2-DE/MALDI-MSMS	Yun et al., 2010
<i>Cocos nucifera L.</i>	Wood shaving	Lethal yellowing	2-DE	Reyes-Martínez et al., 2007

Species	Material	Evaluated condition	Platform	Reference
<i>Euterpe edulis</i>	Embryo and endosperm seeds	-	-	Panza et al., 2007
<i>Hevea brasiliensis</i>	Latex	-	2-DE/LC-ESI-MSMS	D'Amato et al., 2010
<i>Hevea brasiliensis</i>	Latex	-	2-DE/MALDI-MSMS	Wang et al., 2010
<i>Hevea brasiliensis</i>	Seeds	Germinated and dry seeds	2-DE/MALDI-MS	Wong and Abubakar 2005
<i>Manihot esculenta</i>	Green cotyledons of somatic embryos	Undergoing secondary somatic embryogenesis	2-DE/MALDI-MSMS	Baba et al., 2008
<i>Manihot esculenta</i> Crantz	Somatic embryos, plantlets and tuberous roots ( <i>in vitro</i> )	-	SDS-PAGE/LC-ESI-MSMS	Li et al., 2010
<i>Manihot esculenta</i> Crantz	Fibrous and tuberous root tissues	-	2-DE/LC-ESI-MSMS	Sheffield et al., 2010
<i>Mentha arvensis</i>	Leaves	<i>Alternaria alternata</i> infection	2-DE/MALDI-MSMS	Sinha and Chattopadhyay 2011
<i>Musa spp.</i>	Shoot meristems ( <i>in vitro</i> )	-	2-DE/MALDI-MSMS	Samyn et al., 2006
<i>Musa spp.</i>	Meristem cultures	High sucrose medium	2-DE/MALDI-MSMS	Carpentier et al., 2007
<i>Nepenthes alata</i>	Pitcher fluid	-	2-DE/LC-ESI-MSMS	Hatano and Hamada 2008
<i>Ocotea catharinensis</i>	Embryo and endosperm seeds	Developing seeds	2-DE/MALDI-MSMS	Dias et al., 2010
<i>Ocotea catharinensis</i>	Seeds	Developing seeds	2-DE/ MALDI-MS and LC-ESI-MSMS	Moraes, 2006
<i>Phaseolus vulgaris</i>	Seeds	-	2-DE/MALDI-MSMS	De La Fuente et al., 2010
<i>Phoenix dactylifera</i>	Somatic embryos culture ( <i>in vitro</i> )	ABA and sucrose	2-DE/MALDI-MSMS	Sghaier-Hammami et al., 2010
<i>Phoenix dactylifera</i> L.	Sap	-	2-DE/MALDI-MSMS	Thabet et al., 2010
<i>Phoenix dactylifera</i> L.	Zygotic and somatic embryos	-	2-DE/MALDI-MSMS	Sghaier-Hammami et al., 2009
<i>Ricinus communis</i>	Endoplasmic reticulum from seeds	Developing and germinating seeds	2-D DIGE/ MALDI-MS and LC-ESI-MSMS	Maltman et al., 2007
<i>Ricinus communis</i> L.	Seeds	-	2-DE/MALDI-MSMS	Campos et al., 2008
<i>Sporobolus stapfianus</i>	Leaves	Dehydration	2-D DIGE/LC-ESI-MSMS	Oliver et al., 2010
<i>Theobroma cacao</i> L.	Leaves and meristems	<i>Moniliophthora perniciosa</i> infection	2-DE	Pirovani et al., 2008
<i>Vanilla planifolia</i>	Organogenic callus ( <i>in vitro</i> )	-	2-DE/MALDI-MSMS	Palama et al., 2010
<i>Vigna unguiculata</i>	Leaves	Manganese-toxicity response	2-DE/LC-ESI-MSMS	Führs et al., 2008
<i>Vigna unguiculata</i>	Embryogenic cell suspensions	-	2-DE/MALDI-MSMS	Nogueira et al., 2007

**Sample preparation challenges:** In general, most comparative studies on plant proteomics rely mainly on protein separation by two dimensional electrophoresis (2-DE) with subsequent protein identification by mass spectrometry

(MS/MS) (Figure 2). Although the platform based on 2-DE is still the most commonly used, gel-free and second-generation quantitative proteomic techniques have been increasingly used (Jorrín-Novo et al., 2009).



**Figure 2.** Stages of plant proteomics, using interface two-dimensional electrophoresis (2D-PAGE) and mass spectrometry (MS).

Depending on the gel size and pH gradient used, 2-DE can resolve around 2,000 protein spots simultaneously (Gorg et al., 2004). Although 2-DE protein separation can produce impressive protein maps, there are several problems associated with technical reproducibility related to correct spot matching. Even with the most advanced software technologies, high matching accuracy requires final manual validation of each computationally generated spot group (Hajduch et al., 2006). To simplify this labor intense task and maximize accuracy of spot matching, high quality 2-DE gels with minimal spot streaking and overlap are essential (Thelen and Peck, 2007). In the case of plant proteome studies, most of the problems associated with lack of resolution and reproducibility comes from inadequate protein isolation and sample preparation.

Protein extraction and solubilization are key steps for 2-DE based proteomics. Consistent, reproducible 2-DE

gels depend on maximum protein recovery, with minimum amounts of interfering compounds. Proteomic research on poorly characterized experimental systems, like tropical plants, is often hampered by the lack of routine sample preparation procedures. In this way, special attention should be devoted to preparing a preliminary, baseline sample of tropical plants, which could be used to test different extraction methods and isolate and fractionate proteins from plant material (Carpentier et al., 2005; Natarajan et al., 2005). Excellent reviews have indicated the most successful guidelines to be followed in order to perform the adequate protein extraction prior to 2-DE (Shaw et al., 2003; Rose et al., 2004; Gorg et al., 2004; Carpentier et al., 2005; Natarajan et al., 2005; Weiss and Gorg, 2008).

Tropical plant tissues are difficult to disrupt due to the presence of the cell wall, which is made of a complex assembly of polysaccharides. The cell wall and vacuole make up the

majority of the plant cell mass, while the cytosol represents not more than 1-2% of the total cell volume (Carpentier *et al.*, 2005). Moreover, plant tissues have relatively low protein/volume content, when compared to bacterial or animal tissues. The process gains complexity when proteins are extracted from mature plants, for which breaking the thick cell wall for protein extraction is a process that imposes special difficulties (Islam *et al.*, 2004).

In addition, the cell wall and vacuole are both associated with numerous substances that might be responsible for irreproducible and poor results, as among which proteolytic breakdown, streaking and charge heterogeneity (Carpentier *et al.*, 2005). Secondary metabolites, mainly polyphenolic compounds, reversibly combine with the proteins by covalent condensations (Carpentier *et al.*, 2005), leading to charge heterogeneity and streaks in gels. Extracted carbohydrates may block gel pores causing precipitation and extended focusing times, resulting in streaking and sample losses. Terpenoids, pigments and lipids also produce streaking and charge heterogeneity that negatively affects the reproducibility of the 2-DE gels (Carpentier *et al.*, 2005).

Moreover, extremely abundant proteins, like RuBisCO, may potentially mask the detection of proteins with lower abundance by MS, though reliable and robust methods to remove highly abundant proteins from plant tissue samples are still under development (Neilson *et al.*, 2010).

Numerous strategies have been developed in order to deal with interfering elements in tropical plants tissues, using salt solutions, buffers and/or organic solvents (Carpentier *et al.*, 2005). Maximizing high-quality protein extraction, tissue disruption, removal of interfering compounds and protein solubilization steps need to be optimized for each system studied (Wang *et al.*, 2008). The selection of the ideal extraction solution, which would be one that solubilizes higher protein amounts, depends on the species, tissues and proteins being studied.

It has been established that a given extracting solution would have affinities with one or more specific protein classes, which allows extracting proteins differentially, according to the method used (Carpentier *et al.*, 2005). Methods using chaotropic buffers that are used to extract total protein from plant cells and tissues are usually based on urea and thiourea, also containing nonionic and/or zwitterionic

detergents, reducing agents, carrier ampholytes and inhibitors of proteases, phosphatases and oxidoreductases (Saravanan and Rose, 2004; Natarajan *et al.*, 2005); however, this step of protein extraction from plant tissues usually results in poorly resolved protein maps.

Optimal conditions for sample preparation from plant tissues prior to 2-DE have also been reported (Jacobs *et al.*, 2001; Wang *et al.*, 2003a; Giavalisco *et al.*, 2003; Islam *et al.*, 2004; Saravanan and Rose, 2004; Natarajan *et al.*, 2005; Vasconcelos *et al.*, 2005; Carpentier *et al.*, 2005; da Silva *et al.*, 2010). Direct protein precipitation in acetone containing 10% trichloroacetic acid (Mechin *et al.*, 2007) and phenol partitioning (Hurkman and Tanaka, 1986) are the two main protein extraction procedures that have been adopted for 2-DE analysis in plant tissues. However, unfortunately, there is no single method of sample preparation that can be universally applied to all kind of samples analyzed by 2-DE (Weiss and Gorg, 2008). Problems concerning protein solubilization have to be addressed case by case, in order to achieve high-quality 2-DE protein maps and facilitate the spot matching procedure.

Physicochemical properties of some proteins may also interfere in solubilization efficiency. This is particularly true for integral membrane proteins and outer membrane proteins. Integral membrane proteins comprise one or several membrane-spanning regions that are formed from hydrophobic stretches of nonpolar amino acids that transverse the phospholipid bilayer as  $\alpha$ -helices (Braun *et al.*, 2007). Outer membrane proteins form  $\beta$ -barrels that transverse the membranes as pores, containing both polar and nonpolar amino acids that face either the aqueous channel or the hydrophobic lipid bilayer (Braun *et al.*, 2007). Due to their high hydrophobicity the solubilization and proteome analysis of membrane proteins remain a huge challenge: the physicochemical heterogeneity of 2-DE separations is not appropriate for the comprehensive mapping of membrane proteins; many hydrophobic proteins are not solubilized in the IEF sample buffer and precipitate at their isoelectric point and low-abundance of some membrane proteins are beyond the limit of detection of standard proteomic techniques (Ephritikhine *et al.*, 2004).

Alternative strategies have been proposed in order to overcome these problems. In non-plant organisms, 2-DE systems that use ionic detergents, such as SDS,

16-BAC (benzyltrimethyl-n-hexadecylammonium chloride) or CTAB (cetyl trimethyl ammonium bromide) in the first dimension have been capable of resolving a large number of high hydrophobic membrane protein spots (Braun et al., 2007). A comprehensive review on plant plasma membrane proteomics was described by Komatsu (2008) which suggests that analysis of membrane proteins still remains a major challenge for proteomic techniques based on SDS-PAGE, 2-DE or shotgun strategies. The author also states that, although these methods have been successful in identifying membrane proteins, complementary proteomic methods should be adopted, since they are less effective than a combination of quantitative analysis and do not afford an appropriate degree of separation of protein variants by 2-DE.

Difference gel electrophoresis (DIGE) represents an alternative protein separation approach that may be used to overcome issues concerning sensitivity, allowing the detection of less abundant proteins, and less subject to gel variability, which promotes improved reproducibility. Great quantitative accuracy may also be obtained from 2DE DIGE experiments due to the ability to run multiple samples on the same gel (Marouga et al., 2005). In DIGE, proteins are preincubated with activated fluorescent dyes to label lysine or cysteine residues with a sensitive tag that can be used to quantify the abundance of that protein in solution (Tong et al., 2001). When performed with charge-matched reactive dyes, the labeled protein will migrate to the same isoelectric position on a 2-DE map as the traditional staining methods, although a slight mass shift is possible below 20 kDa (Thelen and Peck, 2007).

The main advantage of prelabeling proteins with spectrally distinct fluorescent tags is the ability to combine different protein samples to be separated within the same gel, greatly simplifying spot matching and quantification. Technically, it could be possible to use as many dyes as possible, limited only by the number of fluorescence-emitting dyes with non-overlapping spectral patterns; however, today only three dyes are commercially available (Thelen and Peck, 2007). The linearity, sensitivity and wide dynamic range of these dyes have granted DIGE the status of a powerful quantitative technique, extremely attractive for comparative analysis. In addition, the ultrahigh sensitivity of CyDye (dye linked

to the cysteine residues) DIGE offers more advantages than silver staining, being able to detect proteins at pico to femtomolar levels (Marouga et al., 2005). Unfortunately, the high sensitivity of this strategy is matched by only a handful of commercial mass spectrometers. In light of this, today resolving a separate preparative 2-DE gel containing the Cy-labeled protein samples with larger amounts of unlabeled protein has become a common practice (Thelen and Peck, 2007). After fluorescent imaging, this preparative gel is overstained with Coomassie blue so that the spots of interest can be matched between fluorescent and Coomassie images and excised for mass spectral analysis (Thelen and Peck, 2007).

**Mass spectrometry-based quantitative proteomics analysis:** Although widely used, the 2-DE technique has limitations when dealing with very large or small proteins, the proteins at the extremes of the *pI* scale, membranes and low abundant proteins (Chelius and Bondarenko, 2002). Alternatives to 2-DE gel-based quantification of intact proteins are mass spectrometry (MS)-based approaches, in which, usually, proteins are enzymatically degraded to peptides subsequently used to identify and quantify the proteins from which they originated. Currently, there are two main approaches to comparative quantification by MS. The first approach is the incorporation of stable isotopes into one or more of the samples being studied. This procedure may be carried out *in vivo* using stable isotope labeling, such as the isotope-containing amino acids introduced in the cell culture media (SILAC) (Ong et al., 2002, 2004) or *in vitro*, by introduction of a label into the peptides during or after digestion of the proteins. The most common *in vitro* labeling is the  $^{18}\text{O}$  labeling during trypsin digestion (Mirgorodskaya et al., 2000) and the use of isobaric tags for relative and absolute quantitation (iTRAQ), in which primary amines are chemically modified to produce mass reporters that are used to quantify the peptides in the MS/MS scan (Ross et al., 2004).

The second approach to MS-based protein quantification is the label-free technique, which has become more commonly used in recent years. This strategy is very attractive, since it does not require additional sample processing steps that increase experimental cost and reduce detection sensitivity. Moreover, the limited numbers



of isotope labels that can be run concurrently impose limitations against the comparative analyses of large sample sets (Domon and Aebersold, 2006; Nesvizhskii et al., 2007; Mueller et al., 2008). Based on the information obtained with chromatographic data, MS and MS/MS assignments, label-free quantitative strategies are attractive alternatives for quantitative proteomics because of their simplicity, affordability and flexibility (Gao et al., 2008). Previous studies have shown that quantitation based on spectral ion intensity, peak area of peptide ions or spectral counting correlates well with protein abundance in complex samples (Bondarenko et al., 2002; Chelius and Bondarenko 2002; Wang et al., 2003b; Wiener et al., 2004; Liu et al., 2004, Venable et al., 2004; Old et al., 2005).

In a label-free approach based on relative comparison of peak areas, the use of dedicated software is essential for large scale matching of peptides from multiple raw LC-MS/MS files (Mueller et al., 2008). The basis for such approach is the assumption that under well controlled conditions, identical peptides across different LC-MS/MS experiments can be compared directly (Wong et al., 2007). This procedure is carried out using combinations of retention time and precursor mass characteristics to iteratively match peptides that elute from the chromatography gradient. Once matched, the peak areas (corresponding to the matched peptides) are compared to produce an expression ratio (Thelen and Peck, 2007). In this strategy, one should be very careful with the reproducibility of the chromatographic step, as errors drifts in retention time and in peak detection can influence the label-free quantification result.

An alternative form of label-free quantification is the spectral counting, a strategy that tabulates the number of MS/MS scans attributed to the same precursor ion (Thelen and Peck, 2007). The premise of the method is that the more abundant the peptide, the more likely it will be selected for MS/MS analysis. Liu et al. (2004), Old et al. (2005) and Gao et al. (2008) have shown that spectral counting is highly reproducible and sensitive to protein abundance changes, even when studying complex samples. To facilitate spectral count, all MS/MS spectra are first mined using database search programs and the whole spectra belonging to the same peptide ion or protein is tallied (Wong et al., 2007). Although the technique is easy to conduct, because it is compatible with any software

that exports the number of MS/MS spectra assigned to a specific peptide or protein, the major challenge lies in the computational and statistical analysis of the results in order to assign differential expression of the protein(s) of interest.

**Protein identification of tropical plants: Strategies to study organisms with unknown genomes:** Successful protein identification depends on several factors, including the database used for data mining. The choice of database is very important, since it is the database that determines the significance of results (Cañas et al., 2006). Currently, numerous plant proteome and genome databases are available online, as reviewed by Jorrín-Novo (2009), which can be used for data mining of “orphan” species. Database search can be carried out using protein sequences from an evolutionarily close-related organism: protein identification based on similarity with orthologues (Grossman et al., 2007). Currently, over 20 million green plant EST sequences are listed in the NCBI GenBank Nucleotide database. However, due to the single, automated, sequence reading from one or both ends of random cDNA clone insertions, EST sequences are short (typically about 400-600 bases) and relatively inaccurate (around 2%) (Pontius et al., 2003). The large volume, short lengths and lack of functional annotation are common obstacles to the use of EST sequences for gene modeling and gene or protein identification. One approach to reduce the redundancy and to consolidate EST sequence data, as well as to maximize the sequence information of the gene transcripts is to assemble EST sequences and, if possible, the available cDNA sequences using highly stringent assembly criteria (Childs et al., 2007).

The first efforts to cluster large numbers of nucleotide sequences in a transcription locus, each represented by a set of related but unassembled transcript sequences, and assign functional annotation to these sequences was performed by the NCBI UniGene project (Pontius et al., 2003). Later, the Gene Index project generated CAP3 assembled tentative consensus sequences for data mining of, originally, 30 plant species (Lee et al., 2005). The PlantGDB EST (Dong et al., 2004) and the TIGR (Childs et al., 2007) projects also provide an important source of EST assemblies for protein database search and novel gene identification (Table 2).

**Table 2.** Some sources of assembled and annotated nucleotide sequences for protein database search in plant organisms.

Project Name	Website	Species	Reference
TIGR Plant Transcript Assemblies	<a href="http://plantta.jcvi.org/index.shtml">http://plantta.jcvi.org/index.shtml</a>	254	Childs et al., 2006
PlantGDB EST Assemblies	<a href="http://www.plantgdb.org/prj/ESTCluster/">http://www.plantgdb.org/prj/ESTCluster/</a>	223	Dong et al., 2004
The Gene Index Project*	<a href="http://compbio.dfci.harvard.edu/tgi/">http://compbio.dfci.harvard.edu/tgi/</a>	60	Lee et al., 2005
NCBI UniGene*	<a href="http://www.ncbi.nlm.nih.gov/unigene">http://www.ncbi.nlm.nih.gov/unigene</a>	48	Pontius et al., 2003
HarvEST project	<a href="http://harvest.ucr.edu/">http://harvest.ucr.edu/</a>	09	Close et al., 2007

\* These projects include resources from non-plant organisms. For these projects, only the number of plant organisms is reported.

Another possible strategy is based on peptide sequencing extracting information directly from acquired spectra. Conventional protein identification approaches presume the identity between sequences of peptide precursors fragmented in MS/MS experiments and sequences produced by *in silico* digestion of protein database entries. However, this approach has limited capacity to identify proteins whose sequences remain unknown, heavily modified proteins, or even proteins from wild-bred species that often contain large sequence polymorphisms (Junqueira et al., 2008). Besides, protein sequence mining with a specific database is always biased towards the most conserved and well represented proteins which, in most cases, do not reflect the true composition in a protein mixture (Junqueira et al., 2008). An alternative approach to overcome these issues is to use direct signal information from acquired spectra to assign strings of the primary sequence of a peptide. The *de novo* sequencing by MS dates back more than 30 years and was firstly used in combination with Edman degradation, and then on its own (Biemann, 2002). The development of "soft" ionization techniques, such as the electrospray ionization, high resolution and accuracy mass spectrometry greatly improved the results from *de novo* sequencing (Standing, 2003, Seidler et al., 2010). With improved mass resolution capacity of FT-ICR instruments, detection of more fragments was made possible. Additionally, precision in MS data allows studying homeometric peptides (different peptides with nearly identical sets of *b*- and *y*- ion peaks), improving the quality of acquired data and facilitating both manual and automated interpretation (Frank et al., 2007). Since thousands of MS/MS spectra are generated in each mass spectrometry experiment, manual interpretation of the acquired spectra can be extremely hard and time-consuming. Currently, several engines exclusively dedicated to automated *de novo* sequencing have been developed in order to overcome this issue, such as PepNovo,

SeqMS, MSNovo, Lutefisk, Peaks, AUDENS and SHERENGA. Proteins of organisms with unknown genome often show sequence homologies to functionally related homologues in other completely sequenced organisms. Therefore, existing sequences databases may be helpful through homology-based searches of peptide candidates suggested by the automated *de novo* engines (Seidler et al., 2010). It is important to notice that peptide sequence candidates cannot be directly used to BLAST and FASTA engines without compromising search specificity. For this reason, several database searching tools have been developed to handle the large numbers of peptide sequences proposed by *de novo* sequencing engines (Shevchenko et al., 2009). Cidentify (Taylor and Johnson 1997), MSBLAST (Shevchenko et al., 2001), FASTS (Mackey et al., 2002), MS-Homology (Chalkley et al., 2005), OpenSea (Searle et al., 2004), among others, have already been successfully applied in numerous proteomics studies. For a recent and extensive review on homology-driven proteomics, please read Shevchenko et al. (2009).

## PROSPECTS OF TROPICAL PLANTS PROTEOMICS

The most important platform used in proteomic studies is still 2-DE, followed by MS/MS for separation and identification of the proteins of interest. This kind of analysis provides a convenient way to study the various proteins that are present in plants and to identify those that are regulated in response to different growth and/or stress conditions. However, the amount of proteins identified is modest, when compared to other commonly studied species. It is expected that the number of genomic databases will increase rapidly with the current advances in large-scale genome sequencing technologies, which may facilitate and provide large amounts of data to be used in proteomic experiments. Despite all these

various difficulties, in recent years several papers on plant proteomics have been published. Researchers in this field are becoming aware that the results produced by a well-designed proteomic technique validate the workload and the technical difficulties involved in these investigations (Nielson et al., 2010).

Tropical plant proteomics offers a contribution to several biotechnological applications in biomedicine (through the identification and characterization of allergens and potential protein/peptide drugs, derived from species used in popular medicine), agronomics (through studies of the equivalence of transgenic crops, genotyping, studies of heterosis, plant defense peptides and the environment) and food science (through studies of food quality control and traceability).

Studies addressing seed development play an essential role in conservation strategies for tropical species, since the propagation of tropical species is one of the main problems in terms of economic/agronomic use. These species present recalcitrant seeds, and so only 48% have seeds tolerant to dehydration (Berjak et al., 1984). Proteomics addresses this question, playing a fundamental role towards a better understanding of metabolism and genes expression in embryogenesis and seed filling.

Moreover, different global models predict that greenhouse gases will gradually increase the world's average environment temperature, and the knowledge about functional genomic of tropical plant species will be of great value considering global warming, increasing carbon dioxide concentrations, and UV radiation.

## REFERENCES

Agrawal GK, Yonekura M, Iwahashi Y, Iwahashi H, Rakwal R (2005) System, trends and perspectives of proteomics in dicot plants: Part I: Technologies in proteome establishment. *J Chromat B* 815: 109-123.

Baba A, Nogueira F, Pinheiro C, Brasil J, Jereissati E, Juca T, Soares A, Santos M, Domont G, Campos F (2008) Proteome analysis of secondary somatic embryogenesis in cassava (*Manihot esculenta*). *Plant Sci*. 175: 717-723.

Balbuena TS, Jo L, Pieruzzi FP, Dias LLC, Silveira V, Santa-Catarina C, Junqueira M, Thelen JJ, Shevchenko A, Floh EIS (2011). Differential proteome analysis of mature and germinated embryos of *Araucaria angustifolia*. *Phytochem*. 72: 302-311

Balbuena TS, Silveira V, Junqueira M, Dias LLC, Santa-Catarina C, Shevchenko A, Floh EIS (2009) Changes in the 2-DE protein profile during zygotic embryogenesis in the Brazilian Pine (*Araucaria angustifolia*). *J Proteomics* 72: 337-352.

Berjak P, Dini M, Pammenter NW (1984) Possible mechanisms underlying the differin dehydration responses in recalcitrant and orthodox seeds:

desiccation-associated subcellular changes in propagules of *Avicennia marina*. *Seed Sci Technol*. 12: 365-384.

Biemann K (2002) Four decades of structure determination by mass spectrometry: from alkaloids to heparin. *J American Soc Mass Spec*. 13: 1254-1272.

Bondarenko PV, Chelius D, Shaler TA (2002) Identification and relative quantitative of protein mixtures by enzymatic digestion followed by capillary reversed-phase liquid chromatography-tandem mass spectrometry. *Anal Chem*. 74: 4741-4749.

Braun RJ, Kinkl N, Beer M, Ueffing M (2007) Two-dimensional electrophoresis of membrane proteins. *Analyt Bioanalyt Chem*. 389: 1033-1045.

Cangahuala-Inocente GC, Villarino A, Seixas D, Dumas-Gaudot E, Terenzi H, Guerra MP (2009) Differential proteomic analysis of developmental stages of *Acca sellowiana* somatic embryos. *Acta Physiol Plant*. 31: 501-514.

Campos F, Nogueira F, Moura R, Cardoso C, Costa G, Stancato D, Domont GB, Silva MJ, Juca TL (2008) Proteome analysis of castor bean seeds. *J. Biotech*. 136: S101-S101.

Cañas B, López-Ferrer D, Ramos-Fernández A, Camafeita E, Calvo E (2006) Mass spectrometry technologies for proteomics. *Brief Funct Genomic Proteomic* 4: 295-320.

Carpentier SC, Witters E, Laukens K, Deckers P, Swennen R, Panis B (2005) Preparation of protein extracts from recalcitrant plant tissues: An evaluation of different methods for two-dimensional gel electrophoresis analysis. *Proteomics* 5: 2497-2507.

Carpentier SC, Witters E, Laukens K, Onckelen HV, Swennen R, Panis B (2007) Banana (*Musa* spp.) as a model to study the meristem proteome: Acclimation to osmotic stress. *Proteomics* 7: 92-105.

Chalkley RJ, Baker PR, Hansen KC, Medzhradzsky KF, Allen NP, Rexach M, Burlingame AL (2005) Comprehensive analysis of a multidimensional liquid chromatography mass spectrometry dataset acquired on a quadrupole selecting, quadrupole collision cell, time-of-flight mass spectrometer-how much of the data is theoretically interpretable by search engines? *Mol Cell Proteomics* 4: 1189-1193.

Chelius D, Bondarenko PV (2002) Quantitative profiling of proteins in complex mixtures using liquid chromatography and mass spectrometry. *J Proteome Res* 1: 317-323.

Chen S, Harmon AC (2006) Advances in plant proteomics. *Proteomics* 6: 5504-5516.

Childs KL, Hamilton JP, Zhu W, Ly E, Cheung F, Wu H, Rabinowicz PD, Town CD, Buell CR, Chan AP (2007) The TIGR plant transcript assemblies database. *Nucleic Acids Res* 35: D846-D851.

Close TJ, Wanamaker S, Roose ML, Lyon M (2007) HarVEST: an EST Database and viewing software. In: *Methods in Molecular Biology-Plant Bioinformatics: Methods and Protocols*. Ed.: Edwards, D. Humana Press, Totowa NJ, 161-177 pp.

Da Silva MAO, Garcia JS, de Souza GHMF, Eberlin MN, Gozzo FC, Arruda MAZ (2010) Evaluation of sample preparation protocols for proteomic analysis of sunflower leaves. *Talanta* 80: 1545-1551.

D'Amato A, Bachi A, Fasoli E, Boschetti E, Peltre G, Sénéchal H, Sutra JP, Citterio A, Righetti PG (2010) In-depth exploration of *Hevea brasiliensis* latex proteome and "hidden allergens" via combinatorial peptide ligand libraries. *J. Proteomics* 73:1368-80.

De La Fuente M, Borrajo A, Bermúdez J, Lores M, Alonso J, López M, Santalla M, De Ron AM, Zapata C (2011) 2-DE-based proteomic analysis of common bean (*Phaseolus vulgaris* L.) seeds. *J. Proteomics* 74: 262-267.

Dias LLC, Balbuena TS, Silveira V, Santa-Catarina C, Shevchenko A, Floh EIS (2009) Two-dimensional gel electrophoretic protein profile analysis during seed development of *Ocotea catharinensis*: a recalcitrant seed species. *Braz J Plant Physiol*. 22: 23-33.

Domon B, Aebersold R (2006) Mass spectrometry and protein analysis. *Science* 312: 212-217.

- Dong Q, Schlueter SD, Brendel V (2004) PlantGDB, plant genome database and analysis tools. *Nucl Acids Res.* 32: D354-D359.
- Ephritikhine G, Ferro M, Rolland N (2004) Plant membrane proteomics. *Plant Physiol Bioch.* 42: 943-962.
- Elvira MI, Galdeano MM, Gilardi P, García-Luque I, Serra MT (2008) Proteomic analysis of pathogenesis-related proteins (PRs) induced by compatible and incompatible interactions of pepper mild mottle virus (PMMoV) in *Capsicum chinense* L<sup>3</sup> plants. *J. Exp. Bot.* 59: 1253-1265.
- Frank AM, Savitski MM, Nielsen ML, Zubarev RA, Pevzner PA (2007) *De novo* peptide sequencing and identification with precision mass spectrometry. *J Proteome Res.* 6: 114-123.
- Fuhrs H, Hartwig M, Molina LEB, Heintz D, Dorselaer AV, Braun HP, Horst WJ (2008) Early manganese-toxicity response in *Vigna unguiculata* L. – a proteomic and transcriptomic study. *Proteomics* 8: 149-159.
- Gao BB, Stuart L, Feener EP (2008) Label-free quantitative analysis of one-dimensional PAGE LC/MS/MS proteome. Application on angiotensin II-stimulated smooth muscle cells secretome. *Mol Cell Proteomics* 7: 2399-2409.
- Gallardo K, Job C, Groot SPC, Puype M, Demol H, Vandekerckhove J, Job D (2001) Proteomic analysis of Arabidopsis seed germination and priming. *Plant Physiol.* 126: 835-848.
- Garavaglia BS, Thomas L, Zimaro T, Gottig N, Daurelio LD, Ndimba B, Orellano EG, Ottado J, Gehring, C (2010) A plant natriuretic peptide-like molecule of the pathogen *Xanthomonas axonopodis* pv. *citri* causes rapid changes in the proteome of its citrus host. *BMC Plant Biol.* 10: 51-61.
- Gazzana G, Borlak J (2007) Improved methods for proteome mapping of the Liver by 2-DE MALDI-TOF MS. *J Proteome Res.* 6: 3143-3151.
- Giavalisco P, Nordhoff E, Lehrach H, Gobom J, Klose J (2003) Extraction of proteins from plant tissues for two-dimensional electrophoresis analysis. *Electrophoresis* 24: 207-216.
- Gorg A, Weiss W, Dunn MJ (2004) Current two-dimensional electrophoresis technology for proteomics. *Proteomics* 4: 3665-3685.
- Grossmann J, Fischer B, Baerenfaller K, Owiti J, Buhmann JM, Gruissem W, Baginsky S (2007) A workflow to increase the detection rate of proteins from unsequenced organisms in high-throughput proteomics experiments. *Proteomics* 7: 4245-4254.
- Hajduch M, Casteel JE, Hurrelmeyer KE, Song KE, Agrawal GK, Thelen JJ (2006) Proteomic analysis of seed filling in *Brassica napus*: developmental characterization of metabolic isozymes using high-resolution two-dimensional gel electrophoresis. *Plant Physiol.* 141: 32-46.
- Hatano N, Hamada T (2008) Proteome analysis of pitcher fluid of the carnivorous plant *Nepenthes alata*. *J. Proteome Res.* 7: 809-816.
- He C, Wang Y (2008) Protein extraction from leaves of *Aloe vera* L., a succulent and recalcitrant plant, for proteomic analysis. *Plant Mol Biol Rep.* 26:292-300.
- Heazlewood JL, Millar AH (2003) Integrated plant proteomics – putting green genomes to work. *Funct Plant Biol.* 30: 471-482.
- Hurkman WJ, Tanaka CK (1986) Solubilization of plant membrane proteins for analysis by two-dimensional gel electrophoresis. *Plant Physiol.* 81: 802-806.
- Islam N, Lonsdale M, Upadhyaya NM, Higgins TJ, Hirano H, Akhurst R (2004) Protein extraction from mature rice leaves for two-dimensional gel electrophoresis and its application in proteome analysis. *Proteomics* 4: 1903-1908.
- Jacobs DI, van Rijssen MS, der Heijden R, Verpoorte R (2001) Sequential solubilization of proteins precipitated with trichloroacetic acid in acetone from cultured *Catharanthus roseus* cells yields 52% more spots after two-dimensional electrophoresis. *Proteomics* 1: 1345-1350.
- Jorrin JV, Rubiales D, Dumas-Gaudot E, Recobert G, Maldonado AM, Castillejo MA, Curto M (2006) Proteomics: a promising approach to study biotic interaction in legumes. A review. *Euphytica.* 147: 37-47.
- Jorrin-Novoa JV, Maldonado AM, Echevarría-Zomeño S, Valledor L, Castillejo MA, Curto M, Valeroa J, Sghaier B, Donoso G, Redondo I (2009) Plant proteomics update (2007–2008): Second-generation proteomic techniques, an appropriate experimental design, and data analysis to fulfill MIAPE standards, increase plant proteome coverage and expand biological knowledge. *J Proteomics.* 72: 285-314.
- Junqueira M, Spirin V, Balbuena TS, Warindel P, Surendranath V, Kryukov G, Adzhubei I, Thomas H, Sunyaev S, Shevchenko A (2008). Separating the wheat from the chaff: unbiased filtering of background tandem mass spectra improves protein identification. *J Proteome Res.* 7: 3382-3395.
- Komatsu S (2008). Plasma membrane proteome in *Arabidopsis* and rice. *Proteomics* 8: 4137-4145.
- Lee Y, Tsai J, Sunkara S, Karamycheva S, Perteu G, Sultana R, Antonescu V, Chan A, Cheung F, Quackenbush J (2005) The TIGR gene indices: clustering and assembling EST and known genes and integration with eukaryotic genomes. *Nucl Acids Res.* 33: D71-D74.
- Li K, Zhu W, Zeng K, Zhang Z, Ye J, Ou W, Rehman S, Heuer B, Chen S (2010) Proteome characterization of cassava (*Manihot esculenta* Crantz) somatic embryos, plantlets and tuberous roots. *Proteome Sci.* 8: 10-22.
- Liu H, Sadygov RG, Yates JR (2004) A model for random sampling and estimation of relative proteins abundance in shotgun proteomics. *Anal Chem.* 76: 4193-4201.
- Mackey AJ, Haystead TAJ, Pearson WR (2002) Getting more from less: algorithms for rapid protein identification with multiple short peptide sequences. *Mol Cell Proteomics* 1: 139-147.
- Maltman DJ, Gadd SM, Simon WJ, Slabas AR (2007) Differential proteomic analysis of the endoplasmic reticulum from developing and germinating seeds of castor (*Ricinus communis*) identifies seed protein precursors as significant components of the endoplasmic reticulum. *Proteomics* 7: 1513-1528.
- Mann M, Hendrickson RC, Pandey A (2001) Analysis of proteins and proteomes by mass spectrometry. *Ann Rev Biochem.* 70: 437-473.
- Mann M, Jensen ON (2003) Proteomic analysis of post-translational modifications. *Nature Biotech.* 21: 255-261.
- Marouga R, David S, Hawkins E (2005) The development of the DIGE system: 2D fluorescence difference gel analysis technology. *Anal Bioanal Chem.* 382: 669-678.
- Maserti BE, Del Carratore R, Croce CM, Podda A, Migheli Q, Froelicher Y, Luro F, Morillon R, Ollitrault P, Talon M, Rossignol M (2011) Comparative analysis of proteome changes induced by the two spotted spider mite *Tetranychus urticae* and methyl jasmonate in citrus leaves. *J. Plant Physiol.* 4: 392-402
- Mechin V, Damerval C, Zivy M (2007) Total protein extraction with TCA-acetone. In: *Methods in Molecular Biology-Plant Proteomics*. Eds.: Thiellement, H.; Zivy, M.; Damerval, G.; Mechin, V. Humana Press, Totowa NJ, 1-8 pp.
- Metten KL, Aschehoug ET, Callaway RM (2009) Plant behavioral ecology: dynamic plasticity in secondary metabolites. *Plant Cell Env.* 32: 641-653.
- Mirgorodskaya OA, Kozmin YP, Titov MI, Korner R, Sonksen CP, Roepstorff P (2000) Quantitation of peptides and proteins by matrix-assisted laser desorption/ionization mass spectrometry using (18)O-labeled internal standards. *Rapid Commun Mass Spec.* 14: 1226-1232.
- Moraes FM (2006) Análise proteômica da embriogênese somática e da aquisição de competência embriogênica de *Ocotea catharinensis* Mez (Lauraceae). Brasília, Universidade de Brasília. Master thesis.
- Muccilli V, Licciardello C, Fontanini D, Russo MP, Cunsolo V, Saletti R, Recupero GR, Foti, S (2009) Proteome analysis of *Citrus sinensis* L. (Osbeck) flesh at ripening time. *J. Proteomics* 73: 134-152.
- Mueller LN, Brusniak MY, Mani DR, Aebersold R (2008) An assessment of software solutions for the analysis of mass spectrometry based quantitative proteomics data. *J Proteome Res.* 7: 51-61.

- Natarajan S, Xu C, Carpena TJ, Garrett WM (2005) Comparison of protein solubilization methods suitable for proteomic analysis of soybean seed proteins. *Anal Biochem.* 342: 214-220.
- Neilson KA, Gammulla CG, Mirzaei M, Imin N, Haynes PA (2010) Proteomic analysis of temperature stress in plants. *Proteomics* 10: 828-845.
- Nesvizhskii AI, Vitek O, Aebersold R (2007) Analysis and validation of proteomic data generated by tandem mass spectrometry. *Nature Methods* 4: 787-797.
- Newton RP, Brenton AG, Smith CJ, Dudley R (2004) Plant proteome analysis by mass spectrometry: Principles, problems, pitfalls and recent developments. *Phytochem.* 65: 1449-1485.
- Nogueira FC, Gonçalves EF, Jereissati ES, Santos M, Costa JH, Oliveira-Neto OB, Soares AA, Domont GB, Campos FAP (2007) Proteome analysis of embryogenic cell suspensions of cowpea (*Vigna unguiculata*). *Plant Cell Rep.* 26: 1333-1343.
- Old WM, Meyer-Arendt K, Aveline-Wolf L, Pierce KG, Mendoza A, Sevinsky JR, Resing KA, Ahn NG (2005) Comparison of label-free methods for quantifying human proteins by shotgun proteomics. *Mol Cell Proteomics* 4: 1487-1502.
- Oliver MJ, Jain R, Balbuena TS, Agrawal G, Gasulla F, Thelen JJ (2010) Proteome analysis of leaves of the desiccation-tolerant grass, *Sporobolus stapfianus*, in response to dehydration *Phytochem.* 72: 1273-128.
- Ong SE, Blagoev B, Kristensen DB, Steen H, Pandey A, Mann M (2002) Stable isotope labeling by amino acids in cell culture, SILAC, as a simple and accurate approach to expression proteomics. *Mol Cell Proteomics* 1: 376-386.
- Ong SE, Mittler G, Mann M (2004) Identifying and quantifying *in vivo* methylation sites by heavy methyl SILAC. *Nature Methods* 1: 119-126.
- Palama TL, Menard P, Fock I, Choi YH, Bourdon E, Govinden-Soulangue J, Bahut M, Payet B, Verpoorte R, Kodja H (2010) Shoot differentiation from protocorm callus cultures of *Vanilla planifolia* (Orchidaceae): proteomic and metabolic responses at early stage. *BMC Plant Biol.* 10: 82-100.
- Pan Z, Guan R, Zhu S, Deng X (2009) Proteomic analysis of somatic embryogenesis in Valencia sweet orange (*Citrus sinensis* Osbeck). *Plant Cell Rep.* 28: 281-289.
- Panza V, Distéfano AJ, Carjuzaa P, Láinez V, del Vas M, Maldonado S (2007) Detection of dehydrin-like proteins in embryos and endosperm of mature *Euterpe edulis* seed. *Protoplasma* 231: 1-5.
- Pignataro V, Canton C, Spadafora A, Mazzuca S (2010) Proteome from lemon fruit flavedo reveals that this tissue produces high amounts of the Cit s1 germin-like isoforms. *J. Agr. Food Chem.* 58: 7239-7244.
- Pirovani CP, Carvalho HA, Machado RC, Gomes DS, Alvim FC, Pomella AW, Gramacho KP, Cascardo JCM, Pereira GAG, Micheli, F (2008) Protein extraction for proteome analysis from cacao leaves and meristems, organs infected by *Moniliophthora perniciosa*, the causal agent of the witches' broom disease. *Electrophoresis* 29: 2391-401.
- Pontius JU, Wagner L, Schuler GD (2003) UniGene: a unified view of the transcriptome. In: The NCBI Handbook. Bethesda MD: National Center for Biotechnology Information.
- Reyes-Mastínez C, Morales-Angeles D, Narvaez-Cab M, Osuna-Castro JA, Zizumbo-Villarreal D, Oropeza-Salín CM, Peña-Beltrán, E (2007) Protein extraction for *Cocos nucifera* with and without lethal yellowing phytoplasma: gel-based proteomics. *B. Insectol.* 60: 231-232.
- Roberts JKM (2002) Proteomics and a future generation of plant molecular biologists. *Plant Mol Biol.* 48: 143-154.
- Rose JKC, Bashir S, Giovanoni JJ, Jahn MM, Saravanan RS (2004) Tackling the plant proteome: practical approaches, hurdles and experimental tools. *Plant J.* 39: 715-733.
- Ross PL, Huang YN, Marchese JN, Williamson B, Parker K, Hattan S, Khainovski N, Pillai S, Dey S, Daniels S, Purkayastha S, Juhasz P, Martin S, Bartlett-Jones M, He F, Jacobson A, Pappin DJ (2004) Multiplex protein quantitation in *Saccharomyces cerevisiae* using amine-reactive isobaric tagging reagents. *Mol Cell Proteomics* 3: 1154-1169.
- Rossignol M, Peltier JB, Mock HP, Matros A, Maldonado AM, Jorriñ J (2006) Plant proteome analysis: a 2004 – 2006 update. *Proteomics* 6: 5529-5548.
- Samyn B, Sergeant K, Capentier S, Debysier G, Panis B, Swennen R, Beeumen JV (2006) Functional proteome analysis of the banana plant (*Musa spp.*) using de novo sequence analysis of derivatized peptides. *J Proteome Res.* 6: 70-80.
- Saravanan R, Rose JKC (2004) A critical evaluation of sample extraction techniques for enhanced proteomic analysis of recalcitrant plant tissues. *Proteomics* 4: 2522-2532.
- Searle BC, Dasari S, Turner M, Reddy AP, Choi D, Wilmarth PA, McCormack AL, David LL, Nagalla SR (2004) High-throughput identification of proteins and unanticipated sequence modifications using a mass-based alignment algorithm for MS/MS *de novo* sequencing results. *Anal Chem.* 76: 2220-2230.
- Seidler J, Zinn N, Boehm ME, Lehmann WD (2010). *De novo* sequencing of peptides by MS/MS. *Proteomics* 10: 634-649.
- Sghaier-Hammami B, Drira N, Jorriñ-Novo JV (2009) Comparative 2-DE proteomic analysis of date palm (*Phoenix dactylifera* L.) somatic and zygotic embryos. *J. Proteomics* 73: 161-177.
- Sghaier-Hammami B, Jorriñ-Novo JV, Gargouri-Bouazid R, Drira N (2010) Abscisic acid and sucrose increase the protein content in date palm somatic embryos, causing changes in 2-DE profile. *Phytochem.* 71: 1223-1236.
- Shaw MM, Riederer BM (2003). Sample preparation for two-dimensional gel electrophoresis. *Proteomics* 3: 1408-1417.
- Sheffield J, Taylor N, Fauquet C, Chen S (2006) The cassava (*Manihot esculenta* Crantz) root proteome: protein identification and differential expression. *Proteomics* 6: 1588-1598.
- Shevchenko A, Sunyaev S, Loboda A, Shevchenko A, Bork P, Ens W, Standing KG (2001) Charting the proteomes of organisms with unsequenced genomes by MALDI-quadrupole time-of-flight mass spectrometry and BLAST homology searching. *Anal Chem.* 73: 1917-1926.
- Shevchenko A, Valcu CM, Junqueira M (2009) Tools for exploring the proteome. *J Proteomics* 72: 137-144.
- Shi J, Chen S, Gollop N, Goren R, Goldschmidt E, Porat R (2008) Effects of anaerobic stress on the proteome of citrus fruit. *Plant Sci.* 175: 478-486.
- Silveira, V., C. Santa-Catarina, T. S. Balbuena, F. M. S. Moraes, C. A. O. Ricart, M. V. Sousa, M. P. Guerra, W. Handro and E. I. S. Floh (2008) Endogenous abscisic acid and protein contents during seed development of *Araucaria angustifolia*. *Biol. Plant.* 52: 101-104.
- Sinha R, Chattopadhyay S (2011) Changes in the leaf proteome profile of *Mentha arvensis* in response to *Alternaria alternata* infection. *J. Proteomics* 74: 327-336.
- Standing KG (2003) Peptide and protein *de novo* sequencing by mass spectrometry. *Cur Opin Struct Biol.* 13: 595-601.
- Taylor JA, Johnson RS (1997) Sequence database searches via *de novo* peptide sequencing by tandem mass spectrometry. *Rapid Commun Mass Spec.* 11: 1067-1075.
- Thabet IB, Francis F, de Pauw E, Besbes S, Attia H, Deroanne C, Blecker C (2010) Characterization of proteins from date palm sap (*Phoenix dactylifera* L.) by a proteomic approach. *Food Chem.* 123: 765-770.
- Thelen JJ, Peck SC (2007) Quantitative proteomics in plants: choices in abundance. *Plant Cell* 19: 3339-3346.
- Tong R, Shaw J, Middleton B, Rowlinson R, Rayner S, Young J, Pognan F, Hawkins E, Currie I, Davison M (2001) Validation and development of fluorescent two-dimensional gel electrophoresis proteomics technology. *Proteomics* 1: 377-396.
- Touzet P, Riccardi F, Morin C, Damerval C, Huet JC, Pernollet JC, Zivy D, Devienne D (1996) The maize two dimensional gel protein database: towards an integrated genome analysis program. *Theor. App. Genet.* 93: 997-1005.

- Vasconcelos EAR, Nogueira FCS, Abreu EFM, Gonçalves EF, Souza PAS, Campos FAP (2005) Protein extraction from cowpea tissues for 2-D gel electrophoresis and MS analysis. *Chromatographia* 62: 447-450.
- Venable JD, Dong MQ, Wohlschlegel J, Dillin A, Yates JR (2004) Automated approach for quantitative analysis of complex peptide mixtures from tandem mass spectra. *Nature Methods* 1: 39-45.
- Wahid A, Gelani S, Ashraf M, Foolad MR (2007) Heat tolerance in plants: A overview. *Env Exp Bot.* 61: 199-223.
- Wang W, Scali M, Vignani R, Spadafora A, Sensi E, Mazzuca S, Cresti M (2003a) Protein extraction for two-dimensional electrophoresis from olive leaf, a plant tissue containing high levels of interfering compounds. *Electrophoresis* 24: 2369-2375.
- Wang W, Zhou H, Lin H, Roy S, Shaler TA, Hill LR, Norton S, Kumar P, Anderle M, Becker CH (2003b) Quantification of proteins and metabolites by mass spectrometry without isotopic labeling or spiked standards. *Anal Chem.* 75: 4818-4826.
- Wang W, Tai F, Chen S (2008) Optimizing protein extraction from plant tissues for enhanced proteomics analysis. *J Separ Sci.* 31: 2032-2039.
- Wang X, Shi M, Lu X, Ma R, Wu C, Guo A, Peng M, Tian W (2010) A method for protein extraction from different subcellular fractions of laticifer latex in *Hevea brasiliensis* compatible with 2-DE and MS. *Proteome Sci.* 8: 35-45.
- Wasinger VC, Cordwell SJ, Cerpa-Poljak A, Yan JX, Gooley AA, Wilkins MR, DuncanMW, Harris R, Williams KL, Humphery-Smith I (1995) Progress with gene product mapping of the Mollicutes: *Mycoplasma genitalium*. *Electrophoresis* 16: 1090-1094.
- Weiss W, Gorg A (2008). Sample solubilization buffers for two-dimensional electrophoresis. In.: *Methods in Molecular Biology-2D PAGE: Sample Preparation and Fractionation*. Ed.: Posch, A. Humana Press, Totowa NJ, 35-42 pp.
- Wiener MC, Sachs JR, Deyanova EG, Yates NA (2004) Differential mass spectrometry: a label-free LC-MS method for finding significant differences in complex peptide and protein mixtures. *Anal Chem.* 76: 6085-6096.
- Wilkins MR, Sanchez JC, Williams KL (1996) Current challenges and future applications for protein maps and post-translational vectors maps in proteome projects. *Electrophoresis.* 17: 830-838.
- Wong P, Abubakar S (2005) Post-germination changes in seeds proteome. *Plant Sci.* 169: 303-311.
- Wong WHJ, Sullivan MJ, Cagney G (2007) Computational methods for the comparative quantification of proteins in label-free LC<sup>n</sup>-MS experiments. *Briefings Bioinformatics* 9: 156-165.
- Yun Z, Li W, Pan Z, Xu J, Cheng Y, Deng X (2010) Comparative proteomics analysis of differentially accumulated proteins in juice sacs of ponkan (*Citrus reticulata*) fruit during postharvest cold storage. *Postharvest Biol. Tec.* 56: 189-201.
- Zivy M, Vienne D (2000) Proteomics: a link between genomics, genetics and physiology. *Plant Mol Biol.* 44: 575-580.