Volume 5 | Issue 1 | June 2015 ISSN 2182-0287

Journal of Integrated OMICS

a methodological journal

JIOMICS

Journal of Integrated OMICS

Focus and Scope

Journal of Integrated OMICS, JIOMICS, provides a forum for the publication of original research papers, preliminary communications, technical notes and critical reviews in all branches of pure and applied "-omics", such as genomics, proteomics, lipidomics, metabolomics or metallomics. The manuscripts must address methodological development. Contributions are evaluated based on established guidelines, including the fundamental nature of the study, scientific novelty, and substantial improvement or advantage over existing technology or method. Original research papers on fundamental studies, and novel sensor and instrumentation development, are especially encouraged. It is expected that improvements will also be demonstrated within the context of (or with regard to) a specific biological question; ability to promote the analysis of molecular mechanisms is of particular interest. Novel or improved applications in areas such as clinical, medicinal and biological chemistry, environmental analysis, pharmacology and materials science and engineering are welcome.

Associated editors

AUSTRALIA AND NEW ZEALAND

Xuequn Chen
Department of Molecular & Integrative Physiology, University of Michigan, Ann Arbor, USA
Ye Fang
Biochemical Technologies, Science and Technology Division, Corning Incorporated, USA
Ying Qu
Microdialysis Experts Consultant Service, San Diego, USA
Ying Xu
Department of Biochemistry and Molecular Biology, Institute of Bioinformatics, University of Georgia, Life Sciences Building Athens, GA, USA

JOURNAL OF INTEGRATED OMICS

A methodological Journal

CONTENTS OF VOLUME 5 | ISSUE 1 | JUNE 2015

REVIEW ARTICLES

ORIGINAL ARTICLES

JOURNAL OF INTEGRATED OMICS
A METHODOLOGICAL JOURNAL
HTTP://WWW.JIOMICS.COM

REVIEW ARTICLE | DOI: 10.5584/jiomics.v5i1.179

Arabidopsis thaliana and omics approaches: a review

Josiane Meire Toloti Carneiro*[1,2], Katherine Chacón-Madrid[1,2], Bruna Caroline Miranda Maciel[1,2], Marco Aurelio Zezzi Arruda[1,2]

[1]Department of Analytical Chemistry, Group of Spectrometry, Sample Preparation and Mechanization – GEPAM, Institute of Chemistry, P.O. Box 6154, University of Campinas – UNICAMP, 13083-970, Campinas, SP, Brazil; [2]National Institute of Science and Technology for Bioanalytics, Institute of Chemistry, P.O. Box 6154, University of Campinas – UNICAMP, 13083-970, Campinas, SP, Brazil.

Received: 12 November 2014 **Accepted:** 21 December 2014 **Available Online**: 28 January 2015

ABSTRACT

Arabidopsis thaliana is a small, flowering plant that is widely used as a model organism in plant biology, mainly because it is the first plant to have its entire genome sequenced. It has since proven to be an ideal organism for studying plant development. *Arabidopsis* is commonly used as a model plant for genomics, metabolomics and proteomics studies, and more recently it has been utilized in metallomic studies. Because of its widespread applications, many methods for *Arabidopsis* sample preparation, analytes separation and data quantification have been explored. This review briefly describes the *Arabidopsis thaliana* characteristics, the developed researches and the primary methods using this plant in different fields of OMICS. In the future, the availability of *Arabidopsis* genomic information may result in its continuous development for nanoparticles and metallomics studies.

Keywords: *Arabidopsis thaliana*; OMICS; genomics; proteomics; metabolomics; metallomics.

1. Introduction

Arabidopsis thaliana (*A. thaliana*) is a plant discovered in the sixteenth century by Johannes Thal in the Harz Mountains (Germany). It is a member of the mustard (*Brassicaceaeor Cruciferae*) family of dicotyledonous plants, which includes species such as cabbage and radish. *A. thaliana* is an annual herbaceous plant native to Europe, Central Asia and Northwest Africa [1], although it has been naturalized in many other places [2]. The location of its growth is responsible for the observed differences in *A. thaliana* life cycles that are reflective of genetic variation [3], affecting characteristics such as flowering time, natural variation, plant growth, among others [4,5]. *A. thaliana* has a rapid life cycle corresponding to approximately 6 weeks from germination to maturity.

According to Pigliucci, flowering time and seed dormancy are key traits that determine the timing and length of the *A. thaliana* natural life cycle [6]. Alvarez-Buylla *et al.* studied processes and stages of *A. thaliana* flower development using molecular genetic studies and genomic studies [7]. Based on research involving comparative and evolutionary approaches derived from *A. thaliana* studies, it is possible to establish a method for studying the molecular basis of diverse floral morphologies. When different *A. thaliana* species are grown together under similar environmental conditions, genetic variation can be observed for many traits [8].

Many different natural accessions of *Arabidopsis thaliana* have been collected, and researchers from around the world are using these to uncover complex genetic interactions, such as those underlying the plant's responses to its environment and the evolution of morphological traits. The phenotypic variation for morphological and physiological traits is abundant and enables almost every *Arabidopsis* accession to be distinguished from accessions collected at different locations. These genetically distinct variants are commonly referred to as ecotypes in the scientific literature. The distribution range of *Arabidopsis* is limited by low spring and autumn temperatures and high temperatures with low precipitation in summer [9]. Thus, this plant is an ideal model system for studying

***Corresponding author:** Josiane Meire Toloti Carneiro. Phone: (55) 19-35212133. P.O. Box 6154, University of Campinas – UNICAMP, 13083-970, Campinas, SP, Brazil. E-mail address: josianemeire1@hotmail.com

natural variation.

Considering the ideal characteristics of *A. thaliana* , together with the fact that it was the first plant to have its genome completely sequenced [10,11], it is easy to understand why the field of OMICS technologies uses *A. thaliana* as a model plant for biological, biochemical, physiological, toxicological and others researches. Because of the canonical relationship of gene to transcript to protein, the three OMICS platforms involving genomics, proteomics and metabolomics are inherently complementary, facilitating the detection and identification of many molecules that are expressed in different organisms [12]. Applying these integrated OMICS platforms, DNA, RNA, proteins, peptides, lipids and metabolites are currently detected and measured in different samples

The present review highlights *A. thaliana* research, specifically taking into account OMICS approaches, such as genomics, proteomics and metabolomics. Additionally, some trends regarding the application of this plant for metallomics and nanoparticles studies are briefly discussed, suggesting that both basic and applied science and all up-to-date technologies are needed to gain new insight and the most accurate information from a studied system. Fig. 1 shows a general scheme of commonly applied procedures for OMICS analysis using *A. thaliana*, of which the most important steps are as follows: 1) extraction and purification of the analyte (genes, proteins and metabolites); 2) separation of interesting species by Polymerase Chain Reaction (PCR), two-dimensional polyacrylamide gel electrophoresis (2-D PAGE), two-dimensional difference gel electrophoresis (2-D DIGE), chromatography, etc.; and 3) identification of interesting species by mass spectrometry (MS), nuclear magnetic resonance (NMR), Fourier transform infrared spectroscopy (FTIR), etc. The extraction and purification method as well as the analytical technique used for separation and identification are chosen according to the objective of the research.

2. Genomics studies

Arabidopsis thaliana was the first plant and the third multicellular organism after *Caenorhabditis elegans* [13] and *Drosophila melanogaster* [14] whose genome was completely sequenced [10,11]. To assist biological investigations and to define chromosomal structure, a coordinated effort to sequence the *A. thaliana* genome was initiated in late 1996, led by a consortium of researchers based mainly out of academic institutions in the USA, Europe and Japan (AGI) [11]. Separate teams within the consortium worked on different chromosomes, using distinct procedures [15]. Today, the *Arabidopsis* community is a diverse group of scientists and involves universities, research institutes and private companies.

Knowledge of the complete genomic sequence and a huge collection of gene disruptions provides a research resource that is unique for higher plants [16]. Three papers presenting the DNA sequence of the gene-rich regions on chromosomes 1, 3 and 5 of *A. thaliana* [17–19] were published first, followed by papers in which chromosomes 2 and 4 were described, altogether providing an overview on the *A. thaliana* sequence [20,21].

The initial identification of transcriptional units in the *A. thaliana* genome sequence was carried out largely by *ab initio* gene predictions, sequence homology, sequence motif analysis, and other non-experimental methods [11,17–21]. The Institute for Genomic Research (TIGR) launched a reannotation effort [11], employing the latest annotation tools and resources and applying uniform annotation protocols across the entire genome, with the goal of improving annotation by

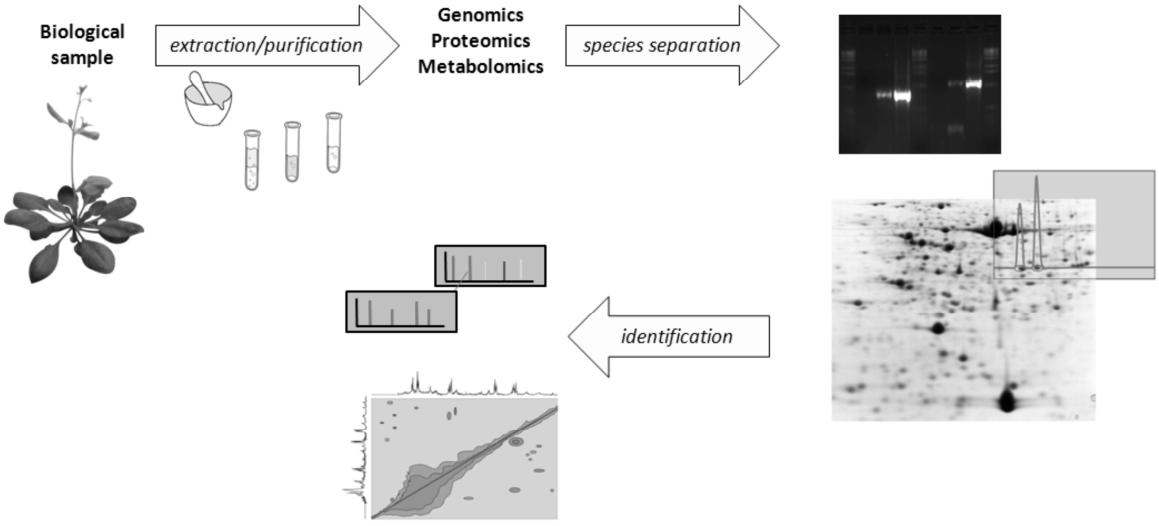

Figure 1. General scheme of common procedures applied for OMIC analysis using *A. thaliana*.

refining gene structure and gene function assignments. The final TIGR genome reannotation release contains annotations for 26,207 protein-coding genes [22]. The completed sequence of a plant nuclear genome yielded a high number of insights, particularly when comparing it with the completed genomic sequences of other species available at the time, namely *Caenorhabditis elegans* and *Drosophila melanogaster*. *A. thaliana* had many families of new proteins but also lacked several common protein families, indicating that these sets of common proteins had undergone differential expansion and contraction in the three multicellular eukaryotes [11].

Although its agronomic significance is little, *A. thaliana* has been widely used in plant biology, offering important advantages for basic research in genetics and molecular biology. It was chosen as a genetic plant model because of its short generation time, abundance of seeds, conveniently short height and solid history in genetics studies. Furthermore, this species has a small nuclear genome (114.5 Mb/125 Mb total), extensive genetic and physical maps of all 5 chromosomes, low repetitive DNA content and simple genetic transformation using *Agrobacterium tumefaciens* [23]. Sequence analysis of the 125-Mb nuclear genome of *A. thaliana* has uncovered 26,207 protein-coding genes, representing approximately 11,000 gene families. Of these genes, approximately 40% have unknown cellular roles, and an established phenotypic function has only been found in approximately 5% [10]. Thus, the wide use of *A. thaliana* in genetic and molecular studies has generated an extensive collection of point mutations, knockouts, knockdowns, over-expressers and other mutant lines.

Quantitative information for the identified proteins was used to establish correlations between transcript and protein accumulation in different plant organs. A proteome map for *A. thaliana* was assembled from high-density, organ-specific proteome catalogs generated for different organs, developmental stages, and undifferentiated cultured cells. The 86,456 unique peptides were matched to 13,029 proteins, providing the expression evidence for 57 gene models. Moreover, proteome analysis identified organ-specific biomarkers and enabled the compilation of an organ-specific set of proteotypic peptides for 4,105 proteins [24].

The flowering plant *A. thaliana* has been an important model system for identifying genes and determining their functions. Analysis of the genetic magnitude of natural variation within *A. thaliana* led to the discovery of novel functions of genes regarding a particular trait and the further characterization of previously identified genes [3]. Until recently, *Arabidopsis* was considered to have low levels of terpenoids (approximately 30 terpene synthase genes) [25], however, recent analysis has revealed the presence of sesquiterpenes in its flowers and monoterpenes in its roots [26]. Although the levels of terpenes are very low, the presence of these genes indicates that *A. thaliana* remains a suitable genetic model, especially for the study of the central pathways of terpene biosynthesis [27].

The definition of gene functions requires the phenotypic characterization of genetic variants. The availability of the *A. thaliana* genome sequence, increased use of large-scale sequencing, and improvements in the resolution of phylogenetic relationships make it an appropriate time to begin developing additional resources. The *Arabidopsis* proteome map provides information about genome activity and proteome assembly and it is available as a resource for plant systems biology [24].

In this sense, computational modeling has an important role in revealing genome-wide regulatory mechanisms. Using these programs, several-thousand new genes and pseudo -genes were added, and approximately one-third of the originally annotated gene models were significantly refined, yielding improved gene structure annotations. Additionally, each protein-coding gene was manually inspected and classified using Gene Ontology terms [22]. Complete and partial gene structures identified by this method were used to improve The Institute for Genomic Research *Arabidopsis* genome annotation (TIGR release v.4.0).

Access to the *A. thaliana* genomic sequence afforded a better understanding of the plant's developmental and environmental responses and allowed the structure and dynamics of plant genomes to be assessed [28–31]. This popular model plant is increasingly used to investigate questions in evolution and ecology; therefore, it is essential to understand patterns of natural genetic variation and to understand the dynamics of wild populations at a scale relevant to single plants [32,33]. The sequencing of complete genomes has advanced the understanding of biological systems and established a series of technologies for the analysis of gene functions, increasing information about the theoretical protein-coding capacity of organisms. The *A. thaliana* genome has been mined for clues to numerous important metabolic pathways and biological processes, many of which are documented in peer-reviewed publications, including the *Arabidopsis* Book [34]. One review [35] summarized the progress made during the past five years and speculated on the future developments in *A. thaliana* research and the implications of these developments for crop science.

3. Proteomics studies

Gene sequence information is not enough to provide significant biological knowledge regarding an organism. Proteomics, which is defined as the quantitative and exhaustive analysis of proteins expressed in a given organ, tissue, or cell, is becoming a more powerful and indispensable technology in the study of biological systems. The analysis of all expressed proteins provides complementary information about genome structure, activity, and regulation. Additionally, proteomics can provide information about post-translational protein modifications involved in developmental control and environmental responses. Thus, proteomic approaches are helpful for answering questions of protein function [36,37].

In recent years, the rapid progress in the determination,

quantification, identification and comprehension of proteins has been possible due to the use of model organisms such as *A. thaliana*. Improvements in the techniques for proteomics, including plant proteomics, based on existing platforms such as 2-DE, 2-LC and MS and some new techniques, including tandem affinity and protein chips, have been observed [38–47].

The availability of the entire genomic sequence of *A. thaliana* provides unique opportunities for the use of a post-genomic tool such as proteomics in its full capacity [11,17–22,48,49]. Research involving *A. thaliana* proteomics has made progress in the past few years, analyzing the proteome of the whole plant and at the level of organs, tissues and organelles. This progress has generated important data sets characterizing the protein-protein interactions, organelle composition, protein activity patterns and protein profiles of this plant [24,28,50–54]. Proteome analysis has proven to be an effective tool not only for analyzing the responses of plants to environmental stresses, including drought, salt, and high and low temperatures but also for allowing the analysis of differential gene expression at the protein level [38,55–57].

The number of distinct proteins that can be identified from complex samples has been increased by the use of shotgun proteomics, a gel-free liquid chromatography tandem mass spectrometry (LC-MS/MS) analysis, compared to traditional gel-based approaches [36,46,58–60]. However, there is no single standardized procedure for the analysis of all proteins and metabolites because these are highly diverse and biochemically heterogeneous [37].

One of the major problems in analyzing a complex material such as a plant leaf sample is the dynamic range of protein abundance and the lack of similarity of the protein content in various cell types of one organism. This difference is responsible for a great diversity of cells and can occur in response to various stimuli and in different cellular compartments [61,62]. The use of multiple model organisms increases proteome research; furthermore, insight into plant proteome dynamics and cell functions are rapidly increased with the use of model plants, such as *A. thaliana* and rice (*Oryza sativa*), that have relatively small genomes. *A. thaliana* has been applied in several of the most comprehensive studies using differential-relative and absolute-quantitative strategies to enhance genome annotation, profile organelles, tissues, cells or sub-cellular proteomes, and investigate developmental processes and responses to biotic and abiotic stresses [36,37,62].

This contribution as a model organism for plants and the increasing impact of proteome research is reflected in the recent increase of proteomic studies using different proteomic techniques to accomplish the separation and evaluation of proteins from crude tissue extracts to further analyze this plant. These techniques are 2-dimensional electrophoresis (2 -DE), 2-dimensional difference gel electrophoresis (2-D DIGE) and liquid chromatography (LC), followed by the identification and characterization of the proteins by mass spectrometric techniques (MS) [38,39,41–43,63–65].

It is important to remember that the results of any experiment are dependent on the condition of the starting material. Therefore, choosing the appropriate sample preparation, based on the subsequent analytical technique and the research objectives, is crucial for obtaining significant and trustworthy results. Sample preparation is of particularly great importance in comparative proteomics because there are often only minor differences between experimental and control samples [61,66–69]. Currently, methods to simplify complex protein mixtures prior to using separation techniques have been proposed, enabling more discrete samples to be analyzed. These methods include sample fractionation and protein enrichment techniques, such as profiling isolated cell organelles and sequential extraction for the selective removal of the most abundant proteins or interfering compounds [41,66–70].

One of the most commonly applied techniques in proteomic analysis is the traditional 2-DE. This method is based on orthogonal separation of proteins according to their pI, molecular weight, solubility and relative abundance. The number, resolution and reproducibility of spots visualized on a 2-DE map depends to a great extent on the tissue sample and the protein extraction protocol [39,43,56,62,63,71,72]. For plant proteomic analysis, the presence of large amounts of non-protein components and lower protein content (compared to bacterial or animal tissues) requires customized experimental strategies for each plant to avoid compromising 2-DE separations. For this reason, simple protein extraction protocols are advisable [38,56,71,73–78].

The most universal protocol for plant tissue analysis recommends protein precipitation after tissue homogenization. TCA-acetone, TCA-phenol or TCA-methanol precipitation methods or a protocol based on a combination of TCA-acetone precipitation followed by methanol washing and phenol extraction have been used and reported [74–81]. Variants of these methods have been previously used in the analysis of the *A. thaliana* proteome. For example, Maldonado *et al.* evaluated changes in the proteome of *A. thaliana* leaves as a response to *Pseudomonas syringae* by comparing three precipitation protocols for protein extraction using 2-DE: TCA-acetone, TCA-acetone + phenol, and phenol only. The quantity and intensity of observed spots were dependent on the protocol used. The TCA-acetone + phenol protocol provided the best results in terms of reproducibility as well as the ability to focus and resolve the intensity of spots and to detect the presence of a single spot [77].

A number of modifications related to 2-DE extraction methods have been published, focusing on reproducible results and how to obtain a good extraction of proteins from plant samples, remove interferences, and preserve proteins in solution [41,62,67,68,71,73,81]. For example, the low abundance of a protein present in a plant leaf sample may be interfered with by the presence of ribulose bisphosphate carboxylase/oxygenase (RuBisCo), the most abundant plant leaf protein. The presence of this protein not only limits the dy-

namic resolution and yield but also affects the electrophoretic migration of neighboring protein species, hampering a deep analysis of the leaf proteome [41,67,68,82]. However, there are some methods for removing RuBisCo, such as those utilizing polyethylene glycol (PEG) [68,82], DTT [83], or the immunocapture of RuBisCo (RuBisCo-IgY affinity) [82], as well as Ca^{2+} plus phytate for its precipitation [66]. Kim *et al.* tested the efficiency of the protamine sulfate precipitation (PSP) method for the depletion of large and small RuBisCo subunits (LSU and SSU) in *A. thaliana*, rice, and maize leaf proteins and provided a novel method for RuBisCo depletion [67]. Espagne *et al.* described a simple mobility shift method for the large subunit of RuBisCo in the first dimension. Using a mixture of ampholine-buffer containing both 4–7 and 3–10 immobilines enabled the characterization of previously undetected protein spots [71].

Profiling isolated cell organelles is another method for simplifying complex protein mixtures before their separation by 2-DE. This strategy has been essential for understanding the biogenesis and function of these plant organelles and for learning that each compartment is enclosed by a unique complement of proteins. To achieve this profiling, it is necessary to have reliable isolation and purification techniques for the cell compartment because many proteins may be lost during these procedures [41,84,85]. Using Triton X-114 phase partitioning, Prime *et al.* characterized the presence of peripheral and integral membrane proteins in a callus culture of *A. thaliana*. A database of mitochondrial, endoplasmic reticulum, golgi/prevacuolar compartment and plasma membrane markers were generated with these results, enabling the definition of specific proteins at the *A. thaliana* callus culture plasma membrane [84].

In another study, *Arabidopsis thaliana* seedlings grown in liquid culture were used to recover proteins secreted from the whole plant. The inclusion of water-insoluble polyvinylpolypyrrolidone (PVPP) in the protocol for the purification of secreted proteins in the culture media led to the identification of a new set of apoplastic proteins, which may have been lost during classical extraction procedures. The role of PVPP was to trap phenolic compounds and to prevent their unspecific interactions with proteins [41]. Fukao *et al.* reported a method of isolating leaf peroxisomes using 2-DE for understanding the tissue-specific expression of leaf peroxisomal proteins. A protein map of leaf peroxisomes from greening cotyledons of *A. thaliana* was built from different cotyledons protein fractions obtained after the extraction procedure. The activities of catalase and cytochrome c oxidase and the content of chlorophyll (Chl) were obtained from each fraction, which were proposed as markers of leaf peroxisomes, mitochondria and chloroplasts, respectively. Additionally, leaf peroxisomes were well separated from mitochondria and chloroplasts, which were present at a high purity and concentration [85].

The high resolution of 2-DE separations makes this methodology the most-used platform for proteomic studies. However, some difficulties with this method have been reported, such as poor reproducibility, the necessity of skilled analysts, and the subjective interpretation of the data obtained through the digitized images of the spots. The lack of reproducibility is frequently attributed to the sample preparation method and natural variations of biological samples and also to the electrophoretic system itself [40,65]. In such cases, the use of 2-D DIGE increases sensitivity and repeatability compared with 2-DE. Two different samples can be run together on the same gel, minimizing the problems mentioned above. Furthermore, 2-D DIGE enables the detection of low-abundance proteins because it is based on fluorescent cyanine dyes, which have higher sensitivity compared with other dyes, such as Coomassie Brilliant Blue (CBB) and silver staining [38,40,60,65,86,87].

In recent years, an increase in 2-D DIGE application as a supporting proteomic method in expression profiling has been observed. Following the manufacturer's instructions, sample protein extraction protocols are similar to those applied in 2-DE. General applications are focusing on the discovery of biomarkers in a wide variety of situations [88] as well as assessing proteomic changes based on stress conditions, genetic modifications [38,89,90], salt [91–93], drought [94], high and low temperatures [95], and metal addition [38,96,97], among others [40,86,98,99].

Studies from Casasoli *et al.*, using 2D-DIGE separation and MS identification, showed that oligogalacturonides (OGs) induced changes in nuclear protein abundance and in the apoplastic proteins of *A. thaliana* seedlings because the plants perceived the OGs as indicators of the presence of pathogens. The nuclear proteins responding to the OG treatment were mainly involved in the protein translation machinery and translation regulation, suggesting a general reprogramming of the plant cell metabolism in response to OGs. Additionally, the differentially expressed apoplastic proteins identified, obtained by a vacuum infiltration-based protocol, included proteins involved in the recognition of OGs and proteins whose post-translational modifications (PTMs) are regulated by OGs [60,86]. Ge *et al.* proposed a model that detailed the possible mechanisms for apoplastic proteins in pollen germination and pollen tube growth of *A. thaliana* pollen grains. Through the results produced by 2-D DIGE, LC–MS/MS and bioinformatics tools, the authors observed and identified global changes of the apoplast proteome during *A. thaliana* pollen germination and pollen tube growth. Additionally, the subcellular localization of three randomly selected differentially expressed proteins was also determined [40]. Holzmeister *et al.* infected wild-type and S-nitrosoglutathione (GSNO-reductase) knock-out *A. thaliana* plants with both avirulent and virulent pathogenic strains of *Pseudomonas syringae*. The authors investigated the importance of nitric oxide (NO) in the plant defense response through a proteomic analysis of the above-mentioned system. The use of 2-D DIGE and MS enabled the identification of proteins that are differentially accumulated during the infection process and a detailed proteomic analysis of the plant defense response [99].

The use of High Performance Liquid Chromatography (HPLC) followed by protein identification and characterization by MS in plant proteomic analysis demonstrated that basic, hydrophobic and membrane-spanning proteins have a greater chance of being separated, provided that they can be obtained for analysis. HPLC separates analytes using two immiscible phases or layers, one of which is held stationary while the other moves over it [47,100]. A variety of chromatographic modes have been developed and are used depending on the analytes: protein isolation and purification using ion-exchange (AEX), hydrophobic interaction chromatography (HIC), affinity chromatography (AC), reversed-phase (RP-HPLC) and/or size exclusion chromatography (SEC) [101–103]. The AEX, HIC and RP-HPLC modes are also used for peptide analysis [59,72,101,104–107]. The chromatographic modes are based on several different mechanisms; RP-HPLC is used for the separation of neutral species on the basis of hydrophobicity, AEX is used for the separation of ionic solutes on the basis of charge, SEC is used for the separation of molecules on the basis of differences in molecular size, and AC is used for the separation of biomolecules on the basis of the lock-and-key mechanism prevalent in biological systems [100].

Zolla *et al*. proved that the use of intact mass measurements (IMMs), performed by coupling RP-HPLC on-line with electrospray ionization mass spectrometry (ESI-MS), is an attractive alternative for monitoring the subtle changes that often accompany physiological adaptations of plants in terms of the concentration of components, measured by the integration of the chromatographic peak. This study reported the relative molecular mass (Mr) for all photosystem I (PSI) proteins in ten plant species, including *A. thaliana*, separated by RP-HPLC and identified by either in-solution trypsin digestion with peptide fragment fingerprinting or the close correspondence between the actual IMMs and those predicted from the DNA sequence [103].

Typically, the most common application of HPLC is high-throughput peptide analysis, due to its coupling with MS/MS. In this case, the protein content of a biological mixture is digested prior to separation and analysis. The MS/MS spectra obtained are searched against a protein database to identify peptides in the sample. Shotgun proteomics performed on a subcellular organelle enables the definition of the organelle proteome and can lead to novel insights into intracellular protein trafficking and sorting [59,69,104,105]. For example, Mitra *et al.* developed an effective chloroform extraction method to improve plasma-membrane protein identification. Rather than traditional solid-phase extraction (SPE), the authors used chloroform extraction prior to off-line AXC and RP-HPLC tandem LC/MS/MS analysis, facilitating the removal of chlorophyll a and b and trypsin used in the digestion and increasing the number of unique peptides for plasma-membrane protein identification [69].

Multidimensional separations are emerging methods designed to increase the resolution power of protein separation, which use off-line or on-line systems, each with specific advantages and limitations [104,108]. For example, a method combining sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) gel separation with RP-HPLC-MS/MS has been used. In this method, proteins are first separated by size on standard polyacrylamide gels or by isoelectric point on IPG strips. After separation, the gel slice is treated similarly to spots excised from 2D gels, and the peptides are separated on an RP column coupled with MS/MS. Variants of these methods have been used in the analysis of the *A. thaliana* proteome [59,106,109]. Batailler *et al.* carried out a proteomic survey of the phloem exudates of *A. thaliana*, collected by the ethylenediaminetetra-acetic acid (EDTA)-facilitated method. Phloem sap proteins of *A. thaliana* were separated by SDS–PAGE. The gel was stained, and bands distributed along the entire length of the lane were excised and subjected to manual in-gel digestion. After separation, the extracts were analyzed by HPLC-MS/MS [109]. To gain insight into the systemic responses of plants to local viral infection or wounding, Niehl *et al*. performed protein profiling of distal, virus-free leaves four and five days after local inoculation of *A. thaliana* plants with either oilseed rape mosaic virus or inoculation buffer alone. Using the system described above, they revealed biomarkers for systemic signaling in response to wounding and viral infection in *A. thaliana* [106].

There has recently been increased attention paid to plant proteomics, exemplified by the above-mentioned efforts concerning plant-specific tissues and organelles. Additionally, the search for possible biomarkers includes evaluating the responses to various biotic and abiotic factors in biological systems, using genetic-model plants such as *A. thaliana*. For this task, proteomics requires three key steps: high-quality extraction, separation, and visualization of complex protein mixtures from crude extracts; identification and characterization of the separated proteins by MS; and database searches. In conclusion, the use of proteomic techniques is critical for plant studies and helps elucidate several key aspects of the metabolic regulation of essential processes. The generation of plant proteome maps, including the identification of low-abundance proteins, requires effort in the most critical proteomic steps, protein extraction and sample preparation, as well as the integration of data obtained through the technologies developed for high-resolution protein separation and rapid, automated protein identification.

4. Metabolomics studies

The metabolome represents the collection of all metabolites in a cell, tissue, organ or organism, which are considered the end products of cellular processes [110]. Thus, metabolomics is the study of chemical processes involving metabolites, intermediates and products of metabolism. The large-scale analysis of metabolites in biological samples (metabolomics) has received increased attention in recent years as a complement to the large-scale analysis of gene transcription and proteins. The usual aim of metabolomic

studies is to quantify the entire metabolome in biological samples; because metabolomic correlations complement the information about changes in metabolite levels, these data may help elucidate the organization of metabolically functional modules.

Plant metabolomics is the study of predominantly low-molecular weight metabolites within cells, tissues or organisms, and it is a widely applied approach for the elucidation of gene function in a wide range of plant species [111]. In plant-based metabolomics, it is common to refer to "primary" and "secondary" metabolites. Primary metabolites are directly involved in the normal growth, development, and reproduction of the plant and are extremely essential to live. Unlike primary metabolites, secondary metabolites are not directly involved in those processes, and their absence does not result in a sudden death; however, secondary metabolites can influence the long-term survivability of the organism, impair fecundity, or affect aesthetics [112].

Plants produce and accumulate a wide variety of secondary metabolites via processes in which precursor structures are modified through biochemical steps driven by different classes of enzymes. Small fluctuations in the metabolome across independent plants may provide information regarding the build-up of a metabolic network [113–116].

Deciphering the metabolome is essential for a better understanding of cellular metabolism as a system. Metabolomics has been utilized not only to investigate plant metabolism but also to identify unknown gene functions by comparing the profiles of wild-type and genetically altered plants and of plants during various developmental stages [113,114,117–119]. Metabolomics studies have demonstrated their robustness in metabolic engineering, process engineering, biomarker discovery, and the functional characterization of novel genes. Furthermore, metabolomics represents one of the most powerful tools to probe the overall effects of gene down-regulation and knockout in transgenic plants at all stages of growth and development.

In plant species, 50,000 metabolites have been characterized, and *Medicago truncatula* and *A. thaliana* are the main models regarding metabolomics projects [113,116,120–122]. The physical and chemical properties of metabolites are highly variable because metabolites include many different types of compounds, such as amino acids, fatty acids, carbohydrates, and organic acids.

In metabolomics studies, efficient and reproducible protocols for the extraction and analysis of metabolites are applied to maximize the number and amounts of metabolites extracted and minimize analytical variations. These well-developed protocols have led to the acquisition of large amounts of information on the composition of *A. thaliana* metabolites [123]. The most common method used for the extraction of metabolites in this plant is one based on shaking the sample at low or high temperatures in organic solvents or in mixtures of solvents [6–8]. For polar metabolites, methanol, ethanol, and water are often used, while chloroform is the most commonly applied solvent for lipophilic compounds.

Metabolome analysis has already been reported using nuclear magnetic resonance (NMR) [5,6], Fourier transform infrared spectroscopy [7], pyrolysis/electron impact-mass spectrometry (pyrolysis/ EI-MS) [8], gas chromatography/ electron impact-mass spectrometry [9], electrospray mass spectrometry (ESI-MS) [12] and ESI-MS coupled with liquid chromatography (LC/MS) [10,11,13]. Analysis by GC-TOF/MS (gas chromatography time-of-flight mass spectrometry) and GC-EI/MS (gas chromatography electron ionization mass spectrometry) are the most applied techniques in *A. thaliana* metabolic studies, and a reliable protocol for analysis has been generated with a relatively limited number of experiments [68,69,70]. These approaches are invaluable for the study of metabolomics in *Arabidopsis*, due to their high reproducibility and the short, constant time between sample preparation and analysis, and have led to the identification of many metabolites.

The application of analytical methods using *A. thaliana* as a model plant is mainly focused on genetic studies [114,124,125], gene function elucidation [111,126], and, in most applications, understanding the expansion of metabolite correlation to gene-expression correlation and studying mediated defenses against biotic and abiotic stresses [114,119–121,126-128].

The metabolomics-based screening method is useful for the rapid characterization of novel genes in both *A. thaliana* and rice [129]. Screening *A. thaliana* lines over-expressing rice full-length (FL) cDNAs (rice FOX *A. thaliana* lines) with gas chromatography was carried out to identify rice genes that caused metabolic changes. Using this technique, it was discovered that the function of LBD37/ASL39 is likely conserved between the dicot and monocot model plant species (*A. thaliana* and rice). For more details on gene expression, a review discussing the study of gene-function relations using the over-expression of *Saccharomyces cerevisiae* and *Escherichia coligenes* cDNAs in *A. thaliana* is suggested [130].

Using *A. thaliana* as a model plant, statistical methods have been performed on metabolomics data. A large amount of microarray data is available, making it easier to build gene coexpression databases [131] and to survey the organization of the transcriptome [132–134]. For example, similarities and dissimilarities in metabolomics correlations were investigated by GCTOF/MS in the aerial parts of 3 *A. thaliana* genotypes: Col-0 wild type (WT), methionine-over accumulation 1 (mto1) and transparent testa 4 (tt4). Multivariate statistical analyzes showed the distinct metabolomes of these plants, provided complementary information on metabolomic correlations about changes in the main metabolite levels, and helped elucidate the organization of metabolically functional modules [120,135]. Regarding the transcript levels, microarray data have collaborated in the evolution of metabolomic studies. The application of MANOVA (Multivariate Analysis of Variance) has allowed researchers to handle multifactorial experimental designs and has revealed clear trends of biological interest. For example,

MANOVA has been applied to analyze *A. thaliana* metabolomic data from factorially designed experiments. This application was demonstrated by a metabolomic investigation using two different factorial designs, *A. thaliana* ethylene signaling mutants and their wild-type counterparts [136]. In this work, the putative *A. thaliana* FLS gene family was studied using a combination of genetic and metabolic analyzes. Although several of the FLS gene family members were expressed, only FLS1 appeared to influence flavonoid biosynthesis in this plant species.

Flavonol synthase (FLS) was the first flavonoid enzyme identified that may be encoded by a gene family in *A. thaliana* plants [137]. In addition to the characterized gene FLS1 (At5g08640), five putative FLS genes (FLS2–FLS6) have been identified in the *Arabidopsis* genome [128]. Studies based on the putative *A. thaliana* FLS gene family revealed that although several of the FLS gene family members were expressed, only flavonol synthase 1 (FLS1) influenced flavonoid biosynthesis. Seedlings of an *A. thaliana* FLS1 null mutant (FLS1-2) showed enhanced anthocyanin levels, a drastic reduction in flavonol glycoside content, and concomitant accumulation of glycosylated forms of dihydroflavonols (a substrate of the FLS reaction). Using a leucoanthocyanidin dioxygenase (LDOX) FLS1-2 double mutant, it was found that the remaining flavonol glycosides found in the FLS1-2 mutant are synthesized in the plant by the FLS-like side-activity of the LDOX enzyme [126]. The results revealed that the *A. thaliana* genome contains at least 24 flavin-containing monooxygenase genes, 272 cytochrome P450 genes, and more than 20 *S*-adenosylmethionine-dependent methyltransferase genes [133].

In most applications involving metabolomic studies and during the development of analytical approaches, *A. thaliana* has been used as a generic plant model to understand mediated defenses against biotic and abiotic stress [114,119–121,127,128]. For example, determining the responses to toxic heavy metals at the level of metabolomics was carried out using *A. thaliana* seedlings as a bio-indicator of Cd pollution [138]. The purpose of these studies is an understanding of the metabolic answer and the adaptation of plants towards heavy metal exposure. The study of Cd impact on the plant metabolome using multivariate statistical analyzes was carried out to compare the metabolic fingerprints and to isolate and identify some discriminating metabolites. *A. thaliana* cell suspensions were treated with different Cd concentrations at different time intervals, and then metabolites present in *A. thaliana* cells grown on Murashige and Skoog media were extracted and injected into the chromatographic system coupled to MS. Three types of data, pretreatment, multivariate statistical analysis (PCA, PLS and PLS-DA) and the PLS methods, proved to be appropriate for the classification of samples and for the extraction of discriminating variables. Additionally, an OSC-PLS2 approach enabled researchers to visualize time-induced and Cd dose-induced changes on the metabolism of *A. thaliana* cells [120].

Isolated *A. thaliana* cells were also used to study the intracellular localization and the biochemical effects of Cs in plant cells [43]. The incorporation and localization of ^{133}Cs in a plant cellular model and the induced metabolic response were analyzed as a function of external K concentration using a multidisciplinary approach. The cellular response to the Cs stress was also analyzed using proteomic and metabolic profiling.

A study involving cultures of *A. thaliana* subjected to high CO_2 stress was carried out to validate a systems biology methodological framework for the analysis of stress-induced molecular interaction networks in the context of plant primary metabolism [139]. An enhanced gas chromatography-mass spectrometry (GC-MS) metabolomic data correction strategy and a new algorithm for the significance analysis of time-series OMICs data were used to extract information about the transcriptional and metabolic plant response. The framework involved the application of time-series integrated full-genome transcriptomic and polar metabolic analyzes on liquid plant cultures. The treatment indicated changes in both transcriptional and metabolic activity, and the identified pathways through which these activities changed revealed insights regarding regulatory processes.

A diversity of metabolites was found by studying the response of *A. thaliana* [124,140–142] to varying light and temperature conditions. The culture was exposed to different environmental conditions in light intensity and/or temperature, and the resulting data sets were subjected to a number of statistical analyzes [143]. In similar studies, metabolome exploration by GC-MS of contrast ecotypes of *A. thaliana* showed that the highest natural variation for plant tolerance existed at lower temperatures than for acclimatory processes [144]. Finally, the resistance of *A. thaliana* plants to the damaging effects of ultraviolet (UV) radiation was investigated through *in vivo* biochemical changes using integrated physiological and metabolic responses. The ability to metabolize xenobiotic compounds was investigated over the entire life cycle of the plant. The results of this metabolic profiling showed that changes in the phenyl propanoid pathway was the key mechanism in both acclimation and plant development [142].

5. Trends: metallomics and nanoparticles

To complement areas such as genomics, proteomics and metabolomics, studies involving metals and metalloids (sometimes linked to the structure of proteins, sometimes free in equilibrium) became important. Metallomics characterizes the metal species present in metalloproteins and tries to elucidate their functions in living organisms [145–147]. Metallomic studies can be classified according to whether they are ionomic and/or metalloproteomic [148]. The first aims to determine free or elemental species in tissue samples, and the second aims to selectively define the metals/metalloids associated with different proteins, protein conformations and protein functions.

Recently, metallomic and proteomic studies performed on

A. thaliana leaves showed the effects of transgenesis and the effects promoted by addition of excess of selenium (Se) [37]. The detection of differentially expressed proteins was carried out by 2-D DIGE. Images of the distribution of Se and sulfur (S) in the leaves were obtained by laser ablation imaging inductively coupled plasma mass spectrometry [LA(i)-ICP-MS]. As a result, 68 species of differentially expressed proteins were detected, of which 27 were identified by ESI-Q-TOF MS/MS. The main biological events modified by these proteins were the glycolysis pathway, photosystems I and II, and the Calvin cycle. The images obtained by LA(i)-ICP-MS showed that added Se was translocated to the leaves and that transgenic plants absorbed higher amounts of Se compared with non-transgenic plants. The results indicated that genetic modification did not influence the production of differential protein species, but it did confer some resistance to the plant regarding abiotic oxidative stress induced by the presence of Se.

Nanoparticles (NPs) are classified as particles sized between 1 and 100 nm that can be dispersed in gaseous, liquid or solid media [149]. Because of the variety of potential applications in biomedical, manufacturing and materials, environmental, energy, optical and electronic fields, the production and characterization of these materials have been widely reported in the literature [150]. Some examples of the applications of NPs are as follows: the use of cerium oxide nanoparticles act as an antioxidant to remove oxygen free radicals that are present in a patient's bloodstream following a traumatic injury, a synthetic skin manufactured with nickel NPs and polymer used in prosthetics, the use of iron nanoparticles (Fe-NPs) to clean up carbon tetrachloride pollution in ground water, and silicon nanoparticles (Si-NPs) coating anodes of Li-ion batteries to increase battery power and reduce recharge time [149].

The scientific literature also contains studies involving NPs and plants. In general, these studies are focused on the effects of NPs on plant germination and growth, aimed at their potential use in agricultural fields [151]. In recent years, some researchers have produced interesting results; Lodeiro *et al.* [152] showed the use of NPs as chemosensors, including a revision about the ability of NP devices to detect metal ions.

In terms of constitution, there are different types of NPs and nanomaterials that have been used in plant science [153], and there is agreement that the effects produced by NPs are dependent on this type, along with the plant species and substrate (*i.e.,* soil, hydroponics, culture medium). Stress response to NPs is a field that appears in an extensive number of studies on metal response in plants. An increasing number of publications have recently considered the interactions of NPs with plants, and most of these studies are focused on the phytotoxicity, uptake and accumulation of NPs in plants [154–156].

Studies involving *A. thaliana* have evaluated the exposure of this plant to some NPs in different categories, such as metal oxides (nAl_2O_3, $nSiO_2$, nFe_3O_4 and nZnO) [157]. Seedlings of this plant were used for reporting the phytotoxicity of silver nanoparticles (Ag-NPs) [158], which was observed at low concentrations. Another study evaluated the impact of citrate-stabilized AgNPs on *A. thaliana* at three levels—physiological phytotoxicity, cellular accumulation and subcellular transport [159]. The phytotoxic effects of AgNPs could not be fully explained by the release of silver ions. Plants exposed to AgNP suspensions bioaccumulated a higher silver content than plants exposed to $AgNO_3$ solutions (Ag^+ representative), indicating AgNP uptake by plants. At three levels, the impacts of AgNPs differed from equivalent dosages of $AgNO_3$. In summary, the studies cited show that phytotoxicity is dependent upon the concentration and particle size of the NPs.

Changes in *A. thaliana* phenotype, at both the cellular and macroscopic level, were also observed. These changes were dependent on the distribution of NPs in the tissue, thereby revealing their bioaccumulative effect. Based on these findings, the researchers stressed that the exact mechanisms remained unclear and required elucidation, as was also observed in another group's research paper. Recently, changes in gene expression in *A.* thaliana exposed to polyvinylpyrrolidone-coated AgNPs and silver ions were evaluated by Kaveh *et al.* [160]. Many genes differentially expressed by AgNPs and Ag^+ were found to be involved in the response of plants to various stresses, providing insights into the molecular mechanisms of the response of plants to AgNPs and Ag^+. Exposure to gold nanoparticles (GNPs) significantly improved the seed germination rate, vegetative growth and antioxidant potential of *A. thaliana*. This was the first report showing GNPs as a promising tool to enhance the seed yield of plants [161].

The characterization of NPs is essential to obtain more information about their properties as well as their applications when focusing on toxicological studies. Responses to NPs would also be a key element in identifying mechanisms involved in stress tolerance and NP toxicity. Many subjects, studies and challenges involving the biological effects of NPs are still unresolved, and their interactions with plant-soil-microorganisms systems still need to be investigated.

In Table 1, the OMICS studies using *A. thaliana* cited in this review are summarized, including the target study, comments and reference number.

6. Conclusions and final remarks

Today, *A. thaliana* remains the standard reference plant for all of biology and it is an efficient tool for the analysis of plant functioning, combining classical genetics with molecular biology. The continuous advancement of *A. thaliana* knowledge enhances its value for plant biology. This plant offers important advantages for OMICS research; it was the first plant to have its entire genome sequenced, making it an ideal model system and a powerful tool for the development in this field. This review highlighted advances in OMICS studies, particularly genomics, proteomics and metabolom-

Table 1. Summary of the OMICS studies cited in the present review using *A. thaliana* as a model plant.

GENOMICS		
Target study	**Comments**	**Ref.**
General characteristics	Classification, flowering time, plant growth, diverse floral morphologies and seed dormancy. Studies of variation observed in life cycles due to genetic variation	[1–10,24]
Genome sequence	Definition of the chromosome structure, effort to sequence the complete genome	[11,15,17–22]
Studies based on genomic sequencing	Genetic transformation using *Agrobacterium tumefaciens*	[23]
	Suitable model, especially, for study of the central pathways of terpene biosynthesis	[25–27]
	Use of computational modeling for revealing genome-wide regulatory mechanisms.	[22]
	Assessment of the structure and dynamics of plant genomes, enabling a better understanding of plant development and environmental responses	[28–31]
	Investigation of evolution and ecology, for understanding patterns of natural genetic variation and the dynamics of wild populations	[32–34]
PROTEOMICS		
Target study	**Comments**	**Ref.**
Methodologies to simplify complex protein mixtures prior to the use of 2-DE	Evaluation of changes in the proteome of *A. thaliana* leaves in response to *Pseudomonas syringae* by comparing three precipitation protocols	[77]
	The selective removal of RuBisCo using protamine sulfate precipitation (PSP) and the mobility shift method	[67,71]
	Profiling of isolated cell organelles; characterized the presence of peripheral and integral membrane proteins in callus culture, the secreted proteins in culture media (apoplastic proteins), and leaf peroxisomes	[41,84,85]
Comparative proteomics studies using 2D-DIGE	Evaluation of proteomic changes based on different stress conditions, genetic modifications, salt, drought, high and low temperatures, and metal addition, among others	[38,40,60,86,91,93, 95,96,98,99]
The use of HPLC in plant proteomic analysis, followed by protein identification and characterization by MS, facilitates greater chance of the separation of basic, hydrophobic and membrane-spanning proteins	Monitoring the subtle changes in the proteome, produced by physiological adaptations of the plants, performed by intact mass measurements (IMMs), using RP-HPLC -ESI-MS.	[103]
	Shotgun proteomics on a subcellular organelle for the definition of the organelle proteome. The protein content of a subcellular organelle is digested prior to separation and analysis.	[59,69,104]
	Use of multidimensional separations, such as combining SDS-PAGE with HPLC-MS/MS, for the analysis of the phloem sap proteins and to gain insight into systemic responses to local virus infection or wounding. The proteins are digested after SDS-	[106,109]
METABOLOMICS		
Target study	**Comments**	**Ref.**
Metabolite profiling	*A. thaliana* is used as plant model for metabolomic projects	[113,116,117]
Effects on metabolites caused by biotic and abiotic stress	Investigation into the effects on *A. thaliana* metabolites following exposure to metals, pathogens, light and temperature changes	[43,120,122, 124,127,138, 139,142–144]
Genes studies, gene expression, functions and annotation	Demonstrating the robustness of metabolomics studies in the functional characterization of novel genes	[111,114,121,126,128–131]
Methods for extraction and derivation for metabolomic analysis	Investigations of extraction and derivation protocols for metabolomic studies	[6–8,123]

Analytical methods more commons in metabolomic studies	Nuclear magnetic resonance (NMR), Fourier transform infrared spectroscopy, pyrolysis/electron impact-mass spectrometry (pyrolysis/ EI-MS), gas chromatog-electron impact-mass spectrometry, electrospray mass spectrometry (ESI-MS) and ESI-MS coupled with liquid chromatography (LC/MS)	[5–13,68–70,113] [125]
Statistical methods performed in metabolomics	*A. thaliana* is used as model plant for development of statistical methods in the area of metabolomics. Evaluation of data pretreatment methods for further statistical analyzes	[120,131–136,139]

METALLOMICS AND NANOPARTICLES

Target study	**Comments**	**Ref.**
Metallomic studies	Study of the distribution of Se and S in the leaves through LA(i)-ICP-MS	[38]
Nanoparticles studies	Evaluation of the exposure of plants to NPs in different categories, and AgNPs phytotoxicity studies	[157–159]
	Study of the bioaccumulation effect of AgNPs and changes in gene expression due to exposure with GNPs	[160,161]

ics, focusing on the use of *A. thaliana* as a versatile plant model. These applications are responsible for the development of numerous methods for *Arabidopsis thaliana* analysis in different OMICS fields.

Considering genomics, *A. thaliana* has been an important model system primarily for identifying genes and determining their functions, thus providing information about genome activity. Having the complete *A. thaliana* genome sequence allowed further understanding of the structure and dynamics of plant genomes. The proteomics approach is helpful for answering questions regarding the functional analysis of proteins. The rapid progress in the determination, quantification, identification and comprehension of proteins has been possible due to the use of model organisms such as *A. thaliana* and their role in improving the existing techniques for proteomics. The proteome map of *A. thaliana* provides information about proteome assembly and is available as a resource for plant systems biology. The contribution of this plant as a model organism for plants increases the impact of proteome research and is reflected in the increase of proteomics studies. With regard to the metabolomic field, *Arabidopsis* has been utilized not only to investigate plant metabolism but also to identify unknown gene functions by comparing the metabolic profiles of nontransgenic and transgenic species.

As a final remark, because there are already defined genomic and proteomic databases available for this plant, *A. thaliana* is useful as a model plant for metallomic studies aiming to elucidate the physiological and biological functions related to the bioactive metallomes of proteins. Regarding the prospects of NPs, studies involving *A. thaliana* proteomics, genomics and metabolomics will be helpful for those researchers who decide to better understand the mechanisms involved in the interactions of NPs with plants, and, in future studies, this plant may become a great tool to clarify the phenomena of phytotoxicity, uptake and bioaccumulation.

Acknowledgments

The authors thank the Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq, Brasília, Brazil), the Fundação de Amparo a Pesquisa do Estado de São Paulo (FAPESP, São Paulo, Brazil), and the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES, Brazil).

References

[1] M. Nordborg, D. Weigel, Nature, 456 (2008) 720-723. doi: 10.1038/nature07629

[2] I.A. Al-Shehbaz, S.L. O'Kane, Jr., Arabidopsis Book, 1 (2002) e0001. doi: 10.1199/tab.0001

[3] M. Koornneef, C. Alonso-Blanco, D. Vreugdenhil, Annu Rev Plant Biol, 55 (2004) 141-172. doi: 10.1146/annurev.arplant.55.031903.141605

[4] V. Le Corre, F. Roux, X. Reboud, Molecular Biology and Evolution, 19 (2002) 1261-1271. doi: 10.1093/oxfordjournals.molbev.a004187

[5] U. Johanson, Science, 290 (2000) 344-347. doi: 10.1126/science.290.5490.344

[6] M. Pigliucci, Arabidopsis Book, 1 (2002) e0003. doi: 10.1199/tab.0003

[7] E.R. Alvarez-Buylla, M. Benitez, A. Corvera-Poire, A. Chaos Cador, S. de Folter, A. Gamboa de Buen, A. Garay-Arroyo, B. Garcia-Ponce, F. Jaimes-Miranda, R.V. Perez-Ruiz, A. Pineyro-Nelson, Y.E. Sanchez-Corrales, Arabidopsis Book, 8 (2010) e0127. doi: 10.1199/tab.0127

[8] Y. Zhen, M.C. Ungerer, New Phytol, 177 (2008) 419-427. doi: 10.1111/j.1469-8137.2007.02262.x

[9] M.H. Hoffmann, Journal of Biogeography, 29 (2002) 125-134. doi: 10.1046/j.1365-2699.2002.00647.x

[10] D.W. Meinke, Science, 282 (1998) 662-682. doi: 10.1126/science.282.5389.662

[11] I. Arabidopsis Genome, Nature, 408 (2000) 796-815. doi: 10.1038/35048692

[12] D.D. Myrold, P. Nannipieri, Classical Techniques Versus Omics Approaches, P. Nannipi, 2014.

[13] C. The, nbsp, C. elegans Sequencing, Science, 282 (1998) 2012-2018. doi: 10.1126/science.282.5396.2012

[14] M.D. Adams, Science, 287 (2000) 2185-2195. doi: 10.1126/science.287.5461.2185

[15] G.P. Copenhaver, Science, 286 (1999) 2468-2474. doi: 10.1126/science.286.5449.2468

[16] The Arabidopsis Biological Resource Center, http://www.arabidopsis.org/ (2013).

[17] A. Theologis, J.R. Ecker, C.J. Palm, N.A. Federspiel, S. Kaul, O. White, J. Alonso, H. Altafi, R. Araujo, C.L. Bowman, S.Y. Brooks, E. Buehler, A. Chan, Q. Chao, H. Chen, R.F. Cheuk, C.W. Chin, M.K. Chung, L. Conn, A.B. Conway, A.R. Conway, T.H. Creasy, K. Dewar, P. Dunn, P. Etgu, T.V. Feldblyum, J. Feng, B. Fong, C.Y. Fujii, J.E. Gill, A.D. Goldsmith, B. Haas, N.F. Hansen, B. Hughes, L. Huizar, J.L. Hunter, J. Jenkins, C. Johnson-Hopson, S. Khan, E. Khaykin, C.J. Kim, H.L. Koo, I. Kremenetskaia, D.B. Kurtz, A. Kwan, B. Lam, S. Langin-Hooper, A. Lee, J.M. Lee, C.A. Lenz, J.H. Li, Y. Li, X. Lin, S.X. Liu, Z.A. Liu, J.S. Luros, R. Maiti, A. Marziali, J. Militscher, M. Miranda, M. Nguyen, W.C. Nierman, B.I. Osborne, G. Pai, J. Peterson, P.K. Pham, M. Rizzo, T. Rooney, D. Rowley, H. Sakano, S.L. Salzberg, J.R. Schwartz, P. Shinn, A.M. Southwick, H. Sun, L.J. Tallon, G. Tambunga, M.J. Toriumi, C.D. Town, T. Utterback, S. Van Aken, M. Vaysberg, V.S. Vysotskaia, M. Walker, D. Wu, G. Yu, C.M. Fraser, J.C. Venter, R.W. Davis, Nature, 408 (2000) 816-820. doi: 10.1038/35048500

[18] M. Salanoubat, K. Lemcke, M. Rieger, W. Ansorge, M. Unseld, B. Fartmann, G. Valle, H. Blocker, M. Perez-Alonso, B. Obermaier, M. Delseny, M. Boutry, L.A. Grivell, R. Mache, P. Puigdomenech, V. De Simone, N. Choisne, F. Artiguenave, C. Robert, P. Brottier, P. Wincker, L. Cattolico, J. Weissenbach, W. Saurin, F. Quetier, M. Schafer, S. Muller-Auer, C. Gabel, M. Fuchs, V. Benes, E. Wurmbach, H. Drzonek, H. Erfle, N. Jordan, S. Bangert, R. Wiedelmann, H. Kranz, H. Voss, R. Holland, P. Brandt, G. Nyakatura, A. Vezzi, M. D'Angelo, A. Pallavicini, S. Toppo, B. Simionati, A. Conrad, K. Hornischer, G. Kauer, T.H. Lohnert, G. Nordsiek, J. Reichelt, M. Scharfe, O. Schon, M. Bargues, J. Terol, J. Climent, P. Navarro, C. Collado, A. Perez-Perez, B. Ottenwalder, D. Duchemin, R. Cooke, M. Laudie, C. Berger-Llauro, B. Purnelle, D. Masuy, M. de Haan, A.C. Maarse, J.P. Alcaraz, A. Cottet, E. Casacuberta, A. Monfort, A. Argiriou, M. flores, R. Liguori, D. Vitale, G. Mannhaupt, D. Haase, H. Schoof, S. Rudd, P. Zaccaria, H.W. Mewes, K.F. Mayer, S. Kaul, C.D. Town, H.L. Koo, L.J. Tallon, J. Jenkins, T. Rooney, M. Rizzo, A. Walts, T. Utterback, C.Y. Fujii, T.P. Shea, T.H. Creasy, B. Haas, R. Maiti, D. Wu, J. Peterson, S. Van Aken, G. Pai, J. Militscher, P. Sellers, J.E. Gill, T.V. Feldblyum, D. Preuss, X. Lin, W.C. Nierman, S.L. Salzberg, O. White, J.C. Venter, C.M. Fraser, T. Kaneko, Y. Nakamura, S. Sato, T. Kato, E. Asamizu, S. Sasamoto, T. Kimura, K. Idesawa, K. Kawashima, Y. Kishida, C. Kiyokawa, M. Kohara, M. Matsumoto, A. Matsuno, A. Muraki, S. Nakayama, N. Nakazaki, S. Shinpo, C. Takeuchi, T. Wada, A. Watanabe, M. Yamada, M. Yasuda, S. Tabata, C. European Union Chromosome 3 Arabidopsis Sequencing, R. Institute for Genomic, D.N.A.R.I. Kazusa, Nature, 408 (2000) 820-822. doi: 10.1038/35048706

[19] S. Tabata, T. Kaneko, Y. Nakamura, H. Kotani, T. Kato, E. Asamizu, N. Miyajima, S. Sasamoto, T. Kimura, T. Hosouchi, K. Kawashima, M. Kohara, M. Matsumoto, A. Matsuno, A. Muraki, S. Nakayama, N. Nakazaki, K. Naruo, S. Okumura, S. Shinpo, C. Takeuchi, T. Wada, A. Watanabe, M. Yamada, M. Yasuda, S. Sato, M. de la Bastide, E. Huang, L. Spiegel, L. Gnoj, A. O'Shaughnessy, R. Preston, K. Habermann, J. Murray, D. Johnson, T. Rohlfing, J. Nelson, T. Stoneking, K. Pepin, J. Spieth, M. Sekhon, J. Armstrong, M. Becker, E. Belter, H. Cordum, M. Cordes, L. Courtney, W. Courtney, M. Dante, H. Du, J. Edwards, J. Fryman, B. Haakensen, E. Lamar, P. Latreille, S. Leonard, R. Meyer, E. Mulvaney, P. Ozersky, A. Riley, C. Strowmatt, C. Wagner-McPherson, A. Wollam, M. Yoakum, M. Bell, N. Dedhia, L. Parnell, R. Shah, M. Rodriguez, L.H. See, D. Vil, J. Baker, K. Kirchoff, K. Toth, L. King, A. Bahret, B. Miller, M. Marra, R. Martienssen, W.R. McCombie, R.K. Wilson, G. Murphy, I. Bancroft, G. Volckaert, R. Wambutt, A. Dusterhoft, W. Stiekema, T. Pohl, K.D. Entian, N. Terryn, N. Hartley, E. Bent, S. Johnson, S.A. Langham, B. McCullagh, J. Robben, B. Grymonprez, W. Zimmermann, U. Ramsperger, H. Wedler, K. Balke, E. Wedler, S. Peters, M. van Staveren, W. Dirkse, P. Mooijman, R.K. Lankhorst, T. Weitzenegger, G. Bothe, M. Rose, J. Hauf, S. Berneiser, S. Hempel, M. Feldpausch, S. Lamberth, R. Villarroel, J. Gielen, W. Ardiles, O. Bents, K. Lemcke, G. Kolesov, K. Mayer, S. Rudd, H. Schoof, C. Schueller, P. Zaccaria, H.W. Mewes, M. Bevan, P. Fransz, D.N.A.R.I. Kazusa, H. Cold Spring, C. Washington University in St Louis Sequencing, C. European Union Arabidopsis Genome Sequencing, Nature, 408 (2000) 823-826. doi: 10.1038/35048507

[20] X. Lin, S. Kaul, S. Rounsley, T.P. Shea, M.I. Benito, C.D. Town, C.Y. Fujii, T. Mason, C.L. Bowman, M. Barnstead, T.V. Feldblyum, C.R. Buell, K.A. Ketchum, J. Lee, C.M. Ronning, H.L. Koo, K.S. Moffat, L.A. Cronin, M. Shen, G. Pai, S. Van Aken, L. Umayam, L.J. Tallon, J.E. Gill, M.D. Adams, A.J. Carrera, T.H. Creasy, H.M. Goodman, C.R. Somerville, G.P. Copenhaver, D. Preuss, W.C. Nierman, O. White, J.A. Eisen, S.L. Salzberg, C.M. Fraser, J.C. Venter, Nature, 402 (1999) 761-768. doi: 10.1038/45471

[21] K. Mayer, C. Schuller, R. Wambutt, G. Murphy, G. Volckaert, T. Pohl, A. Dusterhoft, W. Stiekema, K.D. Entian, N. Terryn, B. Harris, W. Ansorge, P. Brandt, L. Grivell, M. Rieger, M. Weichselgartner, V. de Simone, B. Obermaier, R. Mache, M. Muller, M. Kreis, M. Delseny, P. Puigdomenech, M. Watson, T. Schmidtheini, B. Reichert, D. Portatelle, M. Perez-Alonso, M. Boutry, I. Bancroft, P. Vos, J. Hoheisel, W. Zimmermann, H. Wedler, P. Ridley, S.A. Langham, B. McCullagh, L. Bilham, J. Robben, J. Van der Schueren, B. Grymonprez, Y.J. Chuang, F. Vandenbussche, M. Braeken, I. Weltjens, M. Voet, I. Bastiaens, R. Aert, E. Defoor, T. Weitzenegger, G. Bothe, U. Ramsperger, H. Hilbert, M. Braun, E. Holzer, A. Brandt, S. Peters, M. van Staveren, W. Dirske, P. Mooijman, R. Klein Lankhorst, M. Rose, J. Hauf, P. Kotter, S. Berneiser, S. Hempel, M. Feldpausch, S. Lamberth, H. Van den Daele, A. De Keyser, C. Buysshaert, J. Gielen, R. Villarroel, R. De Clercq, M. Van Montagu, J. Rogers, A. Cronin, M. Quail, S. Bray-Allen, L. Clark, J. Doggett, S. Hall, M. Kay, N. Lennard, K. McLay, R. Mayes, A. Pettett, M.A. Rajandream, M. Lyne, V. Benes, S. Rechmann, D. Borkova, H. Blocker, M. Scharfe, M. Grimm, T.H. Lohnert, S. Dose, M. de Haan, A. Maarse, M. Schafer, S. Muller-Auer, C. Gabel, M. Fuchs, B. Fartmann, K. Granderath, D. Dauner, A. Herzl, S. Neumann, A. Argiriou, D. Vitale, R. Liguori, E. Piravandi, O. Massenet, F. Quigley, G. Clabauld, A. Mundlein, R. Felber, S. Schnabl, R. Hiller, W. Schmidt, A. Lecharny, S. Aubourg, F. Chefdor, R. Cooke, C. Berger, A. Montfort, E. Casacuberta, T. Gibbons, N. Weber, M. Vandenbol, M.

Bargues, J. Terol, A. Torres, A. Perez-Perez, B. Purnelle, E. Bent, S. Johnson, D. Tacon, T. Jesse, L. Heijnen, S. Schwarz, P. Scholler, S. Heber, P. Francs, C. Bielke, D. Frishman, D. Haase, K. Lemcke, H.W. Mewes, S. Stocker, P. Zaccaria, M. Bevan, R.K. Wilson, M. de la Bastide, K. Habermann, L. Parnell, N. Dedhia, L. Gnoj, K. Schutz, E. Huang, L. Spiegel, M. Sehkon, J. Murray, P. Sheet, M. Cordes, J. Abu-Threideh, T. Stoneking, J. Kalicki, T. Graves, G. Harmon, J. Edwards, P. Latreille, L. Courtney, J. Cloud, A. Abbott, K. Scott, D. Johnson, P. Minx, D. Bentley, B. Fulton, N. Miller, T. Greco, K. Kemp, J. Kramer, L. Fulton, E. Mardis, M. Dante, K. Pepin, L. Hillier, J. Nelson, J. Spieth, E. Ryan, S. Andrews, C. Geisel, D. Layman, H. Du, J. Ali, A. Berghoff, K. Jones, K. Drone, M. Cotton, C. Joshu, B. Antonoiu, M. Zidanic, C. Strong, H. Sun, B. Lamar, C. Yordan, P. Ma, J. Zhong, R. Preston, D. Vil, M. Shekher, A. Matero, R. Shah, I.K. Swaby, A. O'Shaughnessy, M. Rodriguez, J. Hoffmann, S. Till, S. Granat, N. Shohdy, A. Hasegawa, A. Hameed, M. Lodhi, A. Johnson, E. Chen, M. Marra, R. Martienssen, W.R. McCombie, Nature, 402 (1999) 769-777. doi: 10.1038/47134

[22] B.J. Haas, J.R. Wortman, C.M. Ronning, L.I. Hannick, R.K. Smith, Jr., R. Maiti, A.P. Chan, C. Yu, M. Farzad, D. Wu, O. White, C.D. Town, BMC Biol, 3 (2005) 7. doi: 10.1186/1741-7007-3-7

[23] M. Bevan, G. Murphy, Trends in Genetics, 15 (1999) 211-214. doi: 10.1016/s0168-9525(99)01744-8

[24] J.C. D'Auria, J. Gershenzon, Curr Opin Plant Biol, 8 (2005) 308-316. doi: 10.1016/j.pbi.2005.03.012

[25] S. Aubourg, A. Lecharny, J. Bohlmann, Mol Genet Genomics, 267 (2002) 730-745. doi: 10.1007/s00438-002-0709-y

[26] K. Baerenfaller, J. Grossmann, M.A. Grobei, R. Hull, M. Hirsch-Hoffmann, S. Yalovsky, P. Zimmermann, U. Grossniklaus, W. Gruissem, S. Baginsky, Science, 320 (2008) 938-941. doi: 10.1126/science.1157956

[27] J.L. Ward, J.M. Baker, A.M. Llewellyn, N.D. Hawkins, M.H. Beale, Proc Natl Acad Sci U S A, 108 (2011) 10762-10767. doi: 10.1073/pnas.1018875108

[28] J.L. Bennetzen, Nat Genet, 27 (2001) 3-5. doi: 10.1038/83726

[29] T. Nishiyama, T. Fujita, I.T. Shin, M. Seki, H. Nishide, I. Uchiyama, A. Kamiya, P. Carninci, Y. Hayashizaki, K. Shinozaki, Y. Kohara, M. Hasebe, Proc Natl Acad Sci U S A, 100 (2003) 8007-8012. doi: 10.1073/pnas.0932694100

[30] M. Kirst, A.F. Johnson, C. Baucom, E. Ulrich, K. Hubbard, R. Staggs, C. Paule, E. Retzel, R. Whetten, R. Sederoff, Proc Natl Acad Sci U S A, 100 (2003) 7383-7388. doi: 10.1073/pnas.1132171100

[31] S. Stracke, S. Sato, N. Sandal, M. Koyama, T. Kaneko, S. Tabata, M. Parniske, Theor Appl Genet, 108 (2004) 442-449. doi: 10.1007/s00122-003-1438-2

[32] K. Bomblies, L. Yant, R.A. Laitinen, S.T. Kim, J.D. Hollister, N. Warthmann, J. Fitz, D. Weigel, PLoS Genet, 6 (2010) e1000890. doi: 10.1371/journal.pgen.1000890

[33] T. Mitchell-Olds, J. Schmitt, Nature, 441 (2006) 947-952. doi: 10.1038/nature04878

[34] S.D. Clouse, Arabidopsis Book, 1 (2002) e0009. doi: 10.1199/tab.0009

[35] K. Yamada, J. Lim, J.M. Dale, H. Chen, P. Shinn, C.J. Palm, A.M. Southwick, H.C. Wu, C. Kim, M. Nguyen, P. Pham, R. Cheuk, G. Karlin-Newmann, S.X. Liu, B. Lam, H. Sakano, T. Wu, G. Yu, M. Miranda, H.L. Quach, M. Tripp, C.H. Chang, J.M. Lee, M. Toriumi, M.M. Chan, C.C. Tang, C.S. Onodera, J.M. Deng, K. Akiyama, Y. Ansari, T. Arakawa, J. Banh, F. Banno, L. Bowser, S. Brooks, P. Carninci, Q. Chao, N. Choy, A. Enju, A.D. Goldsmith, M. Gurjal, N.F. Hansen, Y. Hayashizaki, C. Johnson-Hopson, V.W. Hsuan, K. Iida, M. Karnes, S. Khan, E. Koesema, J. Ishida, P.X. Jiang, T. Jones, J. Kawai, A. Kamiya, C. Meyers, M. Nakajima, M. Narusaka, M. Seki, T. Sakurai, M. Satou, R. Tamse, M. Vaysberg, E.K. Wallender, C. Wong, Y. Yamamura, S. Yuan, K. Shinozaki, R.W. Davis, A. Theologis, J.R. Ecker, Science, 302 (2003) 842-846. doi: 10.1126/science.1088305

[36] S. Baginsky, W. Gruissem, J Exp Bot, 57 (2006) 1485-1491. doi: 10.1093/jxb/erj130

[37] S. Wienkoop, S. Baginsky, W. Weckwerth, J Proteomics, 73 (2010) 2239-2248. doi: 10.1016/j.jprot.2010.07.012

[38] B.C. Maciel, H.S. Barbosa, G.S. Pessoa, M.M. Salazar, G.A. Pereira, D.C. Goncalves, C.H. Ramos, M.A. Arruda, Proteomics, 14 (2014) 904-912. doi: 10.1002/pmic.201300427

[39] A.A. Soler-Garcia, D. Johnson, Y. Hathout, P.E. Ray, Clin J Am Soc Nephrol, 4 (2009) 763-771. doi: 10.2215/CJN.0200608

[40] W. Ge, Y. Song, C. Zhang, Y. Zhang, A.L. Burlingame, Y. Guo, Biochim Biophys Acta, 1814 (2011) 1964-1973. doi: 10.1016/j.bbapap.2011.07.013

[41] S. Charmont, E. Jamet, R. Pont-Lezica, H. Canut, Phytochemistry, 66 (2005) 453-461. doi: 10.1016/j.phytochem.2004.12.013

[42] B.B. Sadi, A.P. Vonderheide, J.M. Gong, J.I. Schroeder, J.R. Shann, J.A. Caruso, J Chromatogr B Analyt Technol Biomed Life Sci, 861 (2008) 123-129. doi: 10.1016/j.jchromb.2007.11.004

[43] P. Le Lay, M.P. Isaure, J.E. Sarry, L. Kuhn, B. Fayard, J.L. Le Bail, O. Bastien, J. Garin, C. Roby, J. Bourguignon, Biochimie, 88 (2006) 1533-1547. doi: 10.1016/j.biochi.2006.03.013

[44] T.L. Delatte, H. Schluepmann, S.C. Smeekens, G.J. de Jong, G.W. Somsen, Anal Bioanal Chem, 400 (2011) 1137-1144. doi: 10.1007/s00216-011-4837-2

[45] M. Fouad, M. Jabasini, N. Kaji, K. Terasaka, M. Tokeshi, H. Mizukami, Y. Baba, Electrophoresis, 29 (2008) 2280-2287. doi: 10.1002/elps.200700635

[46] J. Lee, W.M. Garrett, B. Cooper, J Sep Sci, 30 (2007) 2225-2230. doi: 10.1002/jssc.200700151

[47] S. Mounicou, J. Szpunar, R. Lobinski, Eur J Mass Spectrom (Chichester, Eng), 16 (2010) 243-253. doi: 10.1255/ejms.1059

[48] T.W. Liu, B. Fu, L. Niu, J. Chen, W.H. Wang, J.X. He, Z.M. Pei, H.L. Zheng, J Proteome Res, 10 (2011) 2579-2589. doi: 10.1021/pr200056a

[49] M. Bevan, S. Walsh, Genome Res, 15 (2005) 1632-1642. doi: 10.1101/gr.3723405

[50] S. Chivasa, B.K. Ndimba, W.J. Simon, D. Robertson, X.-L. Yu, J.P. Knox, P. Bolwell, A.R. Slabas, Electrophoresis, 23 (2002) 1754. doi: 10.1002/1522-2683(200206)23:11<1754::aid-elps1754>3.0.co;2-e

[51] U. Basu, J.L. Francis, R.M. Whittal, J.L. Stephens, Y. Wang, O.R. Zaiane, R. Goebel, D.G. Muench, A.G. Good, G.J. Taylor, Plant and Soil, 286 (2006) 357-376. doi: 10.1007/s11104-006-9048-9

[52] A.H. Millar, J.L. Heazlewood, Plant Physiol, 131 (2003) 443-453. doi: 10.1104/pp.009985

[53] T. Shimaoka, Plant and Cell Physiology, 45 (2004) 672-683. doi: 10.1093/pcp/pch099

[54] S. Komatsu, Proteomics, 8 (2008) 4137-4145. doi: 10.1002/pmic.200800088

[55] H. Wang, H.Q. Zheng, W. Sha, R. Zeng, Q.C. Xia, J Exp Bot, 57 (2006) 827-835. doi: 10.1093/jxb/erj066

[56] J.E. Sarry, L. Kuhn, P. Le Lay, J. Garin, J. Bourguignon, Electrophoresis, 27 (2006) 495-507. doi: 10.1002/elps.200500561

[57] Y. Higashi, M.Y. Hirai, T. Fujiwara, S. Naito, M. Noji, K. Saito, Plant J, 48 (2006) 557-571. doi: 10.1111/j.1365-313X.2006.02900.x

[58] C.L. Hsieh, K.W. Yeh, L.J. De Kok, R.N. Pan, Y.H. Kuo, M.H. Tseng, Phytochem Anal, 23 (2012) 324-331. doi: 10.1002/pca.1361

[59] E. Oliveira, I. Amara, D. Bellido, M.A. Odena, E. Dominguez, M. Pages, A. Goday, J Mass Spectrom, 42 (2007) 1485-1495. doi: 10.1002/jms.1292

[60] M. Casasoli, S. Spadoni, K.S. Lilley, F. Cervone, G. De Lorenzo, B. Mattei, Proteomics, 8 (2008) 1042-1054. doi: 10.1002/pmic.200700523

[61] S.K. Mitra, J.A. Gantt, J.F. Ruby, S.D. Clouse, M.B. Goshe, J Proteome Res, 6 (2007) 1933-1950. doi: 10.1021/pr060525b

[62] A. Bodzon-Kulakowska, A. Bierczynska-Krzysik, T. Dylag, A. Drabik, P. Suder, M. Noga, J. Jarzebinska, J. Silberring, J Chromatogr B Analyt Technol Biomed Life Sci, 849 (2007) 1-31. doi: 10.1016/j.jchromb.2006.10.040

[63] R.M. Collins, M. Afzal, D.A. Ward, M.C. Prescott, S.M. Sait, H.H. Rees, A.B. Tomsett, PLoS One, 5 (2010) e10103. doi: 10.1371/journal.pone.0010103

[64] S. Chivasa, J.M. Hamilton, R.S. Pringle, B.K. Ndimba, W.J. Simon, K. Lindsey, A.R. Slabas, J Exp Bot, 57 (2006) 1553-1562. doi: 10.1093/jxb/erj149

[65] S.C. Arruda, S. Barbosa Hde, R.A. Azevedo, M.A. Arruda, Analyst, 136 (2011) 4119-4126. doi: 10.1039/c1an15513j

[66] H.B. Krishnan, S.S. Natarajan, Phytochemistry, 70 (2009) 1958-1964. doi: 10.1016/j.phytochem.2009.08.020

[67] Y.J. Kim, H.M. Lee, Y. Wang, J. Wu, S.G. Kim, K.Y. Kang, K.H. Park, Y.C. Kim, I.S. Choi, G.K. Agrawal, R. Rakwal, S.T. Kim, Proteomics, 13 (2013) 2176-2179. doi: 10.1002/pmic.201200555

[68] S.T. Kim, K.S. Cho, Y.S. Jang, K.Y. Kang, Electrophoresis, 22 (2001) 2103-2109. doi: 10.1002/1522-2683(200106)22:10<2103::aid-elps2103>3.0.co;2-w

[69] S.K. Mitra, B.T. Walters, S.D. Clouse, M.B. Goshe, J Proteome Res, 8 (2009) 2752-2767. doi: 10.1021/pr801044y

[70] C. Albenne, H. Canut, E. Jamet, Front Plant Sci, 4 (2013) 111. doi: 10.3389/fpls.2013.00111

[71] C. Espagne, A. Martinez, B. Valot, T. Meinnel, C. Giglione, Proteomics, 7 (2007) 3788-3799. doi: 10.1002/pmic.200700346

[72] A. Rashid, A. Badhan, M. Deyholos, N. Kav, Plant Molecular Biology Reporter, 31 (2012) 464-469. doi: 10.1007/s11105-012-0498-3

[73] M.A. da Silva, J.S. Garcia, G.H. de Souza, M.N. Eberlin, F.C. Gozzo, M.A. Arruda, Talanta, 80 (2010) 1545-1551. doi: 10.1016/j.talanta.2009.06.016

[74] T. Isaacson, C.M. Damasceno, R.S. Saravanan, Y. He, C. Catala, M. Saladie, J.K. Rose, Nat Protoc, 1 (2006) 769-774. doi: 10.1038/nprot.2006.102

[75] C. Damerval, D. De Vienne, M. Zivy, H. Thiellement, Electrophoresis, 7 (1986) 52-54. doi: 10.1002/elps.1150070108

[76] S.C. Carpentier, E. Witters, K. Laukens, P. Deckers, R. Swennen, B. Panis, Proteomics, 5 (2005) 2497-2507. doi: 10.1002/pmic.200401222

[77] A.M. Maldonado, S. Echevarria-Zomeno, S. Jean-Baptiste, M. Hernandez, J.V. Jorrin-Novo, J Proteomics, 71 (2008) 461-472. doi: 10.1016/j.jprot.2008.06.012

[78] W. Wang, F. Tai, S. Chen, J Sep Sci, 31 (2008) 2032-2039. doi: 10.1002/jssc.200800087

[79] R.S. Saravanan, J.K. Rose, Proteomics, 4 (2004) 2522-2532. doi: 10.1002/pmic.200300789

[80] W. Wang, R. Vignani, M. Scali, M. Cresti, Electrophoresis, 27 (2006) 2782-2786. doi: 10.1002/elps.200500722

[81] P. Giavalisco, E. Nordhoff, H. Lehrach, J. Gobom, J. Klose, Electrophoresis, 24 (2003) 207-216. doi: 10.1002/elps.200390016

[82] I. Widjaja, K. Naumann, U. Roth, N. Wolf, D. Mackey, J.L. Dangl, D. Scheel, J. Lee, Proteomics, 9 (2009) 138-147. doi: 10.1002/pmic.200800293

[83] J.-H. Cho, H. Hwang, M.-H. Cho, Y.-K. Kwon, J.-S. Jeon, S.H. Bhoo, T.-R. Hahn, Journal of Plant Biology, 51 (2008) 297-301. doi: 10.1007/bf03036130

[84] T.A. Prime, D.J. Sherrier, P. Mahon, L.C. Packman, P. Dupree, Electrophoresis, 21 (2000) 3488-3499. doi: 10.1002/1522-2683(20001001)21:16<3488::aid-elps3488>3.0.co;2-3

[85] Y. Fukao, Plant and Cell Physiology, 43 (2002) 689-696. doi: 10.1093/pcp/pcf101

[86] M. Casasoli, I. Meliciani, F. Cervone, G. De Lorenzo, B. Mattei, International Journal of Mass Spectrometry, 268 (2007) 277-283. doi: 10.1016/j.ijms.2007.07.007

[87] J.S. Minden, S.R. Dowd, H.E. Meyer, K. Stuhler, Electrophoresis, 30 Suppl 1 (2009) S156-161. doi: 10.1002/elps.200900098

[88] A. Sussulini, H. Dihazi, C.E. Banzato, M.A. Arruda, W. Stuhmer, H. Ehrenreich, O. Jahn, H.D. Kratzin, Proteomics, 11 (2011) 261-269. doi: 10.1002/pmic.201000371

[89] S.C. Arruda, H.S. Barbosa, R.A. Azevedo, M.A. Arruda, J Proteomics, 93 (2013) 107-116. doi: 10.1016/j.jprot.2013.05.039

[90] A.R. Brandao, H.S. Barbosa, M.A. Arruda, J Proteomics, 73 (2010) 1433-1440. doi: 10.1016/j.jprot.2010.01.009

[91] S. Shi, W. Chen, W. Sun, Proteomics, 11 (2011) 4712-4725. doi: 10.1002/pmic.201100042

[92] A. Kumar Swami, S.I. Alam, N. Sengupta, R. Sarin, Environmental and Experimental Botany, 71 (2011) 321-328. doi: 10.1016/j.envexpbot.2010.12.017

[93] B.K. Ndimba, S. Chivasa, W.J. Simon, A.R. Slabas, Proteomics, 5 (2005) 4185-4196. doi: 10.1002/pmic.200401282

[94] S. Cui, J. Hu, S. Guo, J. Wang, Y. Cheng, X. Dang, L. Wu, Y. He, J Exp Bot, 63 (2012) 711-726. doi: 10.1093/jxb/err296

[95] S. Amme, A. Matros, B. Schlesier, H.P. Mock, J Exp Bot, 57 (2006) 1537-1546. doi: 10.1093/jxb/erj129

[96] P. Kieffer, J. Dommes, L. Hoffmann, J.F. Hausman, J. Renaut, Proteomics, 8 (2008) 2514-2530. doi: 10.1002/pmic.200701110

[97] S. Alvarez, B.M. Berla, J. Sheffield, R.E. Cahoon, J.M. Jez,

L.M. Hicks, Proteomics, 9 (2009) 2419-2431. doi: 10.1002/pmic.200800478

[98] Y. Wang, A.R. Slabas, S. Chivasa, J Plant Physiol, 169 (2012) 1690-1697. doi: 10.1016/j.jplph.2012.06.013

[99] C. Holzmeister, A. Frohlich, H. Sarioglu, N. Bauer, J. Durner, C. Lindermayr, Proteomics, 11 (2011) 1664-1683. doi: 10.1002/pmic.201000652

[100] J.S. Garcia, C.S. Magalhaes, M.A. Arruda, Talanta, 69 (2006) 1-15. doi: 10.1016/j.talanta.2005.08.041

[101] L.R. Mataveli, M. Fioramonte, F.C. Gozzo, M.A. Arruda, Metallomics, 4 (2012) 373-378. doi: 10.1039/c2mt00186a

[102] Q. Chan, S.E. Afton, J.A. Caruso, Metallomics, 2 (2010) 147 -153. doi: 10.1039/b916194e

[103] L. Zolla, S. Rinalducci, A.M. Timperio, Proteomics, 7 (2007) 1866-1876. doi: 10.1002/pmic.200500053

[104] J. Lee, B. Cooper, Mol Biosyst, 2 (2006) 621-626. doi: 10.1039/b612326k

[105] H.M. Yuan, K.L. Li, R.J. Ni, W.D. Guo, Z. Shen, C.P. Yang, B.C. Wang, G.F. Liu, C.H. Guo, J. Jiang, Mol Biol Rep, 38 (2011) 3045-3054. doi: 10.1007/s11033-010-9971-y

[106] A. Niehl, Z.J. Zhang, M. Kuiper, S.C. Peck, M. Heinlein, J Proteome Res, 12 (2013) 2491-2503. doi: 10.1021/pr3010698

[107] Y. Fukao, M. Yoshida, R. Kurata, M. Kobayashi, M. Nakanishi, M. Fujiwara, K. Nakajima, A. Ferjani, Plant Cell Physiol, 54 (2013) 808-815. doi: 10.1093/pcp/pct033

[108] L.R. Mataveli, M.A. Arruda, J Proteomics, 104 (2014) 94-103. doi: 10.1016/j.jprot.2014.02.031

[109] B. Batailler, T. Lemaitre, F. Vilaine, C. Sanchez, D. Renard, T. Cayla, J. Beneteau, S. Dinant, Plant Cell Environ, 35 (2012) 1258-1273. doi: 10.1111/j.1365-3040.2012.02487.x

[110] K.W. Jordan, J. Nordenstam, G.Y. Lauwers, D.A. Rothenberger, K. Alavi, M. Garwood, L.L. Cheng, Dis Colon Rectum, 52 (2009) 520-525. doi: 10.1007/DCR.0b013e31819c9a2c

[111] L.W. Sumner, P. Mendes, R.A. Dixon, Phytochemistry, 62 (2003) 817-836. doi: 10.1016/s0031-9422(02)00708-2

[112] R. Bentley, Crit Rev Biotechnol, 19 (1999) 1-40. doi: 10.1080/0738-859991229189

[113] R.J. Bino, R.D. Hall, O. Fiehn, J. Kopka, K. Saito, J. Draper, B.J. Nikolau, P. Mendes, U. Roessner-Tunali, M.H. Beale, R.N. Trethewey, B.M. Lange, E.S. Wurtele, L.W. Sumner, Trends Plant Sci, 9 (2004) 418-425. doi: 10.1016/j.tplants.2004.07.004

[114] R.D. Hall, New Phytol, 169 (2006) 453-468. doi: 10.1111/j.1469-8137.2005.01632.x

[115] M. Stitt, R. Sulpice, J. Keurentjes, Plant Physiol, 152 (2010) 428-444. doi: 10.1104/pp.109.150821

[116] K. Saito, F. Matsuda, Annu Rev Plant Biol, 61 (2010) 463-489. doi: 10.1146/annurev.arplant.043008.092035

[117] O. Fiehn, Plant Molecular Biology, 48 (2002) 155-171. doi: 10.1023/a:1013713905833

[118] J. Kopka, A. Fernie, W. Weckwerth, Y. Gibon, M. Stitt, Genome Biol, 5 (2004) 109. doi: 10.1186/gb-2004-5-6-109

[119] K.M. Oksman-Caldentey, K. Saito, Curr Opin Biotechnol, 16 (2005) 174-179. doi: 10.1016/j.copbio.2005.02.007

[120] C. Ducruix, D. Vailhen, E. Werner, J.B. Fievet, J. Bourguignon, J.-C. Tabet, E. Ezan, C. Junot, Chemometrics and Intelligent Laboratory Systems, 91 (2008) 67-77. doi: 10.1016/j.chemolab.2007.08.002

[121] K. Kai, H. Takahashi, H. Saga, T. Ogawa, S. Kanaya, D. Ohta, Plant Biotechnology, 28 (2011) 379-385. doi: 10.5511/plantbiotechnology.11.0704b

[122] K. Hahlbrock, P. Bednarek, I. Ciolkowski, B. Hamberger, A. Heise, H. Liedgens, E. Logemann, T. Nurnberger, E. Schmelzer, I.E. Somssich, J. Tan, Proc Natl Acad Sci U S A, 100 Suppl 2 (2003) 14569-14576. doi: 10.1073/pnas.0831246100

[123] J. Gullberg, P. Jonsson, A. Nordstrom, M. Sjostrom, T. Moritz, Anal Biochem, 331 (2004) 283-295. doi: 10.1016/j.ab.2004.04.037

[124] F. Kaplan, J. Kopka, D.W. Haskell, W. Zhao, K.C. Schiller, N. Gatzke, D.Y. Sung, C.L. Guy, Plant Physiol, 136 (2004) 4159-4168. doi: 10.1104/pp.104.052142

[125] M. Kusano, A. Fukushima, M. Arita, P. Jonsson, T. Moritz, M. Kobayashi, N. Hayashi, T. Tohge, K. Saito, BMC Syst Biol, 1 (2007) 53. doi: 10.1186/1752-0509-1-53

[126] R. Stracke, R.C. De Vos, L. Bartelniewoehner, H. Ishihara, M. Sagasser, S. Martens, B. Weisshaar, Planta, 229 (2009) 427-445. doi: 10.1007/s00425-008-0841-y

[127] W. Maksymiec, Z. Krupa, Environmental and Experimental Botany, 57 (2006) 187-194. doi: 10.1016/j.envexpbot.2005.05.006

[128] M. Pelletier, Plant Physiology, 113 (1997) 1437-1445. doi: 10.1104/pp.113.4.1437

[129] D. Albinsky, M. Kusano, M. Higuchi, N. Hayashi, M. Kobayashi, A. Fukushima, M. Mori, T. Ichikawa, K. Matsui, H. Kuroda, Y. Horii, Y. Tsumoto, H. Sakakibara, H. Hirochika, M. Matsui, K. Saito, Mol Plant, 3 (2010) 125-142. doi: 10.1093/mp/ssp069

[130] A.R. Fernie, R.N. Trethewey, A.J. Krotzky, L. Willmitzer, Nat Rev Mol Cell Biol, 5 (2004) 763-769. doi: 10.1038/nrm1451

[131] K. Saito, M.Y. Hirai, K. Yonekura-Sakakibara, Trends Plant Sci, 13 (2008) 36-43. doi: 10.1016/j.tplants.2007.10.006

[132] W.I. Mentzen, E.S. Wurtele, BMC Plant Biol, 8 (2008) 99. doi: 10.1186/1471-2229-8-99

[133] A. Fukushima, M. Wada, S. Kanaya, M. Arita, DNA Res, 15 (2008) 367-374. doi: 10.1093/dnares/dsn025

[134] L. Mao, J.L. Van Hemert, S. Dash, J.A. Dickerson, BMC Bioinformatics, 10 (2009) 346. doi: 10.1186/1471-2105-10-346

[135] A. Fukushima, M. Kusano, H. Redestig, M. Arita, K. Saito, BMC Syst Biol, 5 (2011) 1. doi: 10.1186/1752-0509-5-1

[136] H.E. Johnson, A.J. Lloyd, L.A.J. Mur, A.R. Smith, D.R. Causton, Metabolomics, 3 (2007) 517-530. doi: 10.1007/s11306-007-0065-3

[137] M. Pelletier, Plant Physiology, 111 (1996) 339-345. doi: 10.1104/pp.111.1.339

[138] W. Liu, Y.S. Yang, D. Francis, H.J. Rogers, P. Li, Q. Zhang, Chemosphere, 73 (2008) 1138-1144. doi: 10.1016/j.chemosphere.2008.07.020

[139] B. Dutta, H. Kanani, J. Quackenbush, M.I. Klapa, Biotechnol Bioeng, 102 (2009) 264-279. doi: 10.1002/bit.22036

[140] L. Vaclavik, A. Mishra, K.B. Mishra, J. Hajslova, Anal Bioanal Chem, 405 (2013) 2671-2683. doi: 10.1007/s00216-012-6692-1

[141] P. Nair, S. Kandasamy, J. Zhang, X. Ji, C. Kirby, B. Benkel, M.D. Hodges, A.T. Critchley, D. Hiltz, B. Prithiviraj, BMC Genomics, 13 (2012) 643. doi: 10.1186/1471-2164-13-643

[142] J.A. Lake, K.J. Field, M.P. Davey, D.J. Beerling, B.H. Lomax, Plant Cell Environ, 32 (2009) 1377-1389. doi: 10.1111/j.1365-3040.2009.02005.x

[143] C. Caldana, T. Degenkolbe, A. Cuadros-Inostroza, S. Klie, R. Sulpice, A. Leisse, D. Steinhauser, A.R. Fernie, L. Willmitzer, M.A. Hannah, Plant J, 67 (2011) 869-884. doi: 10.1111/j.1365-313X.2011.04640.x

[144] D. Cook, S. Fowler, O. Fiehn, M.F. Thomashow, Proc Natl Acad Sci U S A, 101 (2004) 15243-15248. doi: 10.1073/pnas.0406069101

[145] H. Haraguchi, Journal of Analytical Atomic Spectrometry, 19 (2004) 5. doi: 10.1039/b308213j

[146] J.S. Becker, N. Jakubowski, Chem Soc Rev, 38 (2009) 1969-1983. doi: 10.1039/b618635c

[147] R. Lobinski, D. Schaumloffel, J. Szpunar, Mass Spectrom Rev, 25 (2006) 255-289. doi: 10.1002/mas.20069

[148] S. Mounicou, J. Szpunar, R. Lobinski, Chem Soc Rev, 38 (2009) 1119-1138. doi: 10.1039/b713633c

[149] C. Buzea, I.I. Pacheco, K. Robbie, Biointerphases, 2 (2007) MR17. doi: 10.1116/1.2815690

[150] R. Taylor, S. Coulombe, T. Otanicar, P. Phelan, A. Gunawan, W. Lv, G. Rosengarten, R. Prasher, H. Tyagi, Journal of Applied Physics, 113 (2013) 011301. doi: 10.1063/1.4754271

[151] L.R. Khot, S. Sankaran, J.M. Maja, R. Ehsani, E.W. Schuster, Crop Protection, 35 (2012) 64-70. doi: 10.1016/j.cropro.2012.01.007

[152] C. Lodeiro, J.L. Capelo, J.C. Mejuto, E. Oliveira, H.M. Santos, B. Pedras, C. Nunez, Chem Soc Rev, 39 (2010) 2948-2976. doi: 10.1039/b819787n

[153] C.M. Rico, S. Majumdar, M. Duarte-Gardea, J.R. Peralta-Videa, J.L. Gardea-Torresdey, J Agric Food Chem, 59 (2011) 3485-3498. doi: 10.1021/jf104517j

[154] D. Lin, B. Xing, Environ Pollut, 150 (2007) 243-250. doi: 10.1016/j.envpol.2007.01.016

[155] S. Lin, J. Reppert, Q. Hu, J.S. Hudson, M.L. Reid, T.A. Ratnikova, A.M. Rao, H. Luo, P.C. Ke, Small, 5 (2009) 1128-1132. doi: 10.1002/smll.200801556

[156] W.-M. Lee, Y.-J. An, H. Yoon, H.-S. Kweon, Environmental Toxicology and Chemistry, 27 (2008) 1915. doi: 10.1897/07-481.1

[157] C.W. Lee, S. Mahendra, K. Zodrow, D. Li, Y.C. Tsai, J. Braam, P.J. Alvarez, Environ Toxicol Chem, 29 (2010) 669-675. doi: 10.1002/etc.58

[158] X. Ma, J. Geisler-Lee, Y. Deng, A. Kolmakov, Sci Total Environ, 408 (2010) 3053-3061. doi: 10.1016/j.scitotenv.2010.03.031

[159] J. Geisler-Lee, Q. Wang, Y. Yao, W. Zhang, M. Geisler, K. Li, Y. Huang, Y. Chen, A. Kolmakov, X. Ma, Nanotoxicology, 7 (2013) 323-337. doi: 10.3109/17435390.2012.658094

[160] R. Kaveh, Y.S. Li, S. Ranjbar, R. Tehrani, C.L. Brueck, B. Van Aken, Environ Sci Technol, 47 (2013) 10637-10644. doi: 10.1021/es402209w

[161] V. Kumar, P. Guleria, V. Kumar, S.K. Yadav, Sci Total Environ, 461-462 (2013) 462-468. doi: 10.1016/j.scitotenv.2013.05.018

JOURNAL OF INTEGRATED OMICS
A METHODOLOGICAL JOURNAL
HTTP://WWW.JIOMICS.COM

REVIEW ARTICLE | DOI: 10.5584/jiomics.v5i1.186

Proteomic profiling of the HSPB chaperonome: Mass spectrometric identification of small heat shock proteins in stressed skeletal muscles

Sandra Murphy[1], Kay Ohlendieck*[1]

[1]Muscle Biology Laboratory, Department of Biology, Maynooth University, National University of Ireland, Maynooth, Co. Kildare, Ireland.

Received: 29 April 2015 **Accepted:** 30 June 2015 **Available Online**: 30 June 2015

ABSTRACT

The continuing maintenance of protein homeostasis and the protection of proteomic integrity is essential for the survival of complex cellular systems under stressful conditions. Proteostasis is maintained by a complex system of protective pathways that involve several classes of molecular chaperones, now referred to as the chaperonome. The elaborate interplay of these components averts detrimental protein aggregation and supports proteins in resuming their functional fold. In skeletal muscle tissues, molecular chaperones protect contractile functions throughout fibre adaptations to changed physiological demands and prevent tissue damage during acute phases of protein misfolding or prolonged periods of harmful protein accumulation. This results in considerable changes in the expression profile of individual members of the large family of heat shock proteins. Systematic proteomic surveys of skeletal muscle tissues have revealed that the concentration of small heat shock proteins is especially affected following strenuous exercise, in various neuromuscular disorders and during the natural aging process. Of the 10 identified members of the small heat shock protein HSPB family, HSPB1 (Hsp25), HSPB2 (MKBP), HSPB3 (Hsp27), HSPB4 (αA-crystallin), HSPB5 (αB-crystallin), HSPB6 (Hsp20), HSPB7 (cardiovascular cvHsp) and HSPB8 (Hsp22) are clearly present in skeletal muscle tissues. This review outlines the proteomic identification of small heat shock proteins and their muscle-specific expression and induction patterns in health and disease. Since HSPB molecules are of relatively low molecular mass, belong to the markedly soluble type of proteins and represent critical pro-survival proteins that are intrinsically involved in the prevention of stress-induced fibre damage, they present ideal muscle-associated biomarker candidates for the establishment of superior diagnostic and therapy-monitoring approaches to assess stress-related skeletal muscle degeneration.

Keywords: Biomarker discovery; Heat shock protein; Molecular chaperone; Muscle disease; Neuromuscular disorder; Stress response.

Abbreviations

αBC: alpha-B-crystallin; **cvHSP**: cardiovascular heat shock protein; **CAL**: calreticulin; **CAX**: calnexin; **HSP**: heat shock protein; **MS**: mass spectrometry; **sHSP**: small heat shock protein; **PDI**: protein disulfide isomerase; **PPI**: peptidyl-prolyl cis-trans isomerase.

1. Introduction

The detailed analysis of puffing activity patterns in the salivary gland chromosomes of *Drosophila melanogaster* in relation to protein synthesis during heat shock [1-5] led to the foundation of the new scientific area of stress biology [6-8]. Following the initial identification of heat shock-inducible proteins, comprehensive investigations into the cellular response to stressful conditions has established the concept of molecular chaperoning and the classification of the large family of heat shock proteins (HSP) [9-11]. A large variety of HSP molecules have been categorized based on their molecular masses and protective functions following exposure of cells to oxidative stress, heat shock or toxic insults [12-14]. Molecular chaperones provide a variety of essential functions in relation to general cytoprotection and the prevention of deleterious side effects on protein function during stress. A rapid response to acute stressors or the up-regulation of sus-

***Corresponding author:** Prof. Kay Ohlendieck, Department of Biology, Maynooth University, National University of Ireland, Maynooth, Co. Kildare, Ireland. Phone number: (353) (1) 708-3842; Fax number: (353) (1) 708-3845; Email address: kay.ohlendieck@nuim.ie

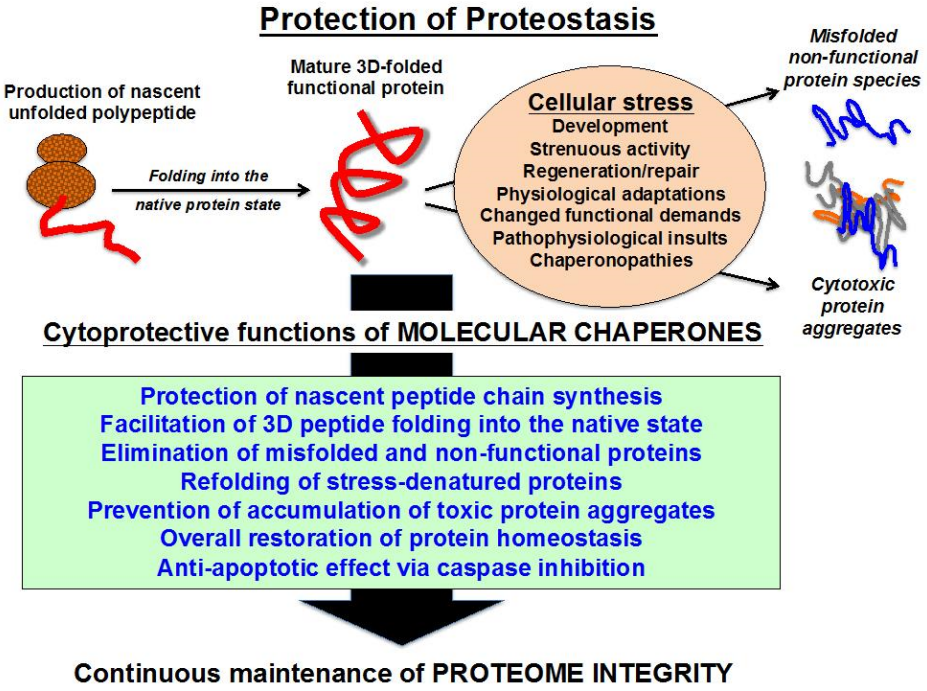

Figure 1. Overview of the cytoprotective roles of molecular chaperones.

tained protective pathways is provided by HSPs and related chaperoning biomolecules [10]. This includes the (i) protection of nascent peptide chain synthesis, (ii) the facilitation of proper peptide folding into the native protein state, (iii) the swift elimination of misfolded and non-functional protein species, (iv) the refolding of stress-denatured proteins, (v) the continuous prevention of the accumulation of misfolded proteins that might otherwise form toxic aggregates within cellular structures, (vi) anti-apoptotic effects via inhibition of the caspase system, and (vii) the restoration of proteostasis and thereby maintenance of proteome integrity [14]. Figure 1 outlines the concept of the cytoprotective role of molecular chaperones under stressful conditions in skeletal muscles, such as cellular development, strenuous activity, tissue regeneration, fibre repair, physiological adaptations to changed functional demands, secondary pathophysiological insults or primary abnormalities due to chaperonopathies [15].

2. The chaperonome and heat shock proteins

The human genome project has identified over 100 different molecular chaperones and based on OMICS-type investigations and systems biological approaches [16], the concept of the chaperonome has been developed [15]. The proteome-wide distribution of chaperones and their roles in maintaining and stabilizing the protein constituents of cells by facilitating the synthesis, transportation and macromolecular assembly of proteins, as well as peptide refolding following stress, essential protein degradation and proteotoxic aggregate dissociation, ensures cellular survival. The presence of molecular chaperones and their swift up-regulation during cellular stress balances protein synthesis and protein degradation, thereby providing proteostasis and proteome stability. In the past, HSPs have been classified based on their tissue-specific expression patterns, their molecular mass, their constitutive presence and/or their stress-related inducibility. In addition to the main HSPs, a variety of endoplasmic reticulum-associated proteins, co-chaperones and modifying enzymes are involved in cytoprotective pathways. The new nomenclature, as summarized by Kampinga et al. [13], categorizes molecular chaperones into several distinct protein families, including HSPA (HSP70) and co-chaperones, HSPB (αB-crystallin like small HSP), HSPC (HSP90), HSPD (HSP60), HSPE (HSP10), HSPH (HSP110), DNAJ proteins (HSP40), calreticulin (CAL), calnexin (CAX), peptidyl-prolyl cis-trans isomerases (PPI) and protein disulfide isomerases (PDI). As recently reviewed in detail in relation to skeletal muscle tissues, the different types of molecular chaperones provide a variety of protective functions [17].

HSPA molecules have a major chaperoning role by ensuring the correct folding of newly synthesized muscle proteins, by associating with misfolded and/or aggregated contractile proteins and by supporting the correct protein re-folding following cellular stress. The HSP70 family of proteins is widely distributed throughout muscle fibres, including inducible HSP70/72 and constitutive HSP73/Hsc70, as well as mitochondrial HSP75 and the HSP78 (GRP78) chaperone of the sarcoplasmic reticulum [18-20]. The large ATP-dependent molecular chaperones of the HSPC class function

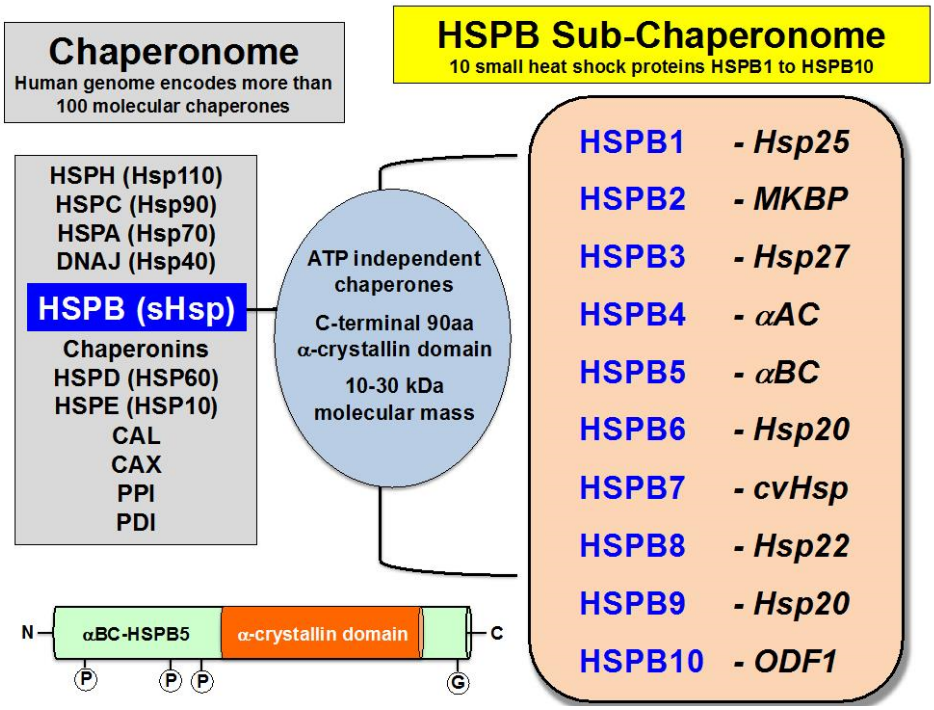

Figure 2. *Listing of the identified members of the chaperonome and the HSPB sub-chaperonome consisting of the family of small heat shock proteins.* The biochemical criteria for the classification of HSPB molecules are outlined and the domain structure of the characteristic α-crystallin domain within the prototype of a small heat shock protein, αB-crystallin (HSPB5), is shown.

down-stream of HSPA and bind to hormone receptors and kinases, whereby the interactions of HSP90α and HSP90β with co-regulators and co-chaperones is involved in the activation and stabilization of signalling proteins [21-23]. The HSP70-like unfolding proteins belonging to the category of disaggregating HSPH/HSP110 chaperones stabilize substrate proteins and actively dissociate stress-induced protein aggregates [24]. DNAJ/HSP40 proteins are primary co-chaperones that regulate the complex formation between HSPA and client proteins, and facilitate protein translation, protein folding, protein unfolding, protein translocation and protein degradation [25]. Another group of chaperonins that co-operate with HSPA molecules are HSPD/HSP60 molecules that support protein folding and re-folding patterns [26, 27]. The CAL/CAX chaperone system of the sarcoplasmic reticulum facilitates the correct folding of newly synthesized glycoproteins, especially those displaying N-linked glycan moieties, making it an essential part of the glycoprotein quality control system [28]. PPI enzymes preserve the correct conformation of distinct protein segments via catalysing the cis/trans isomerization of peptide-bonds besides proline residues [29] and PDI enzymes promote the correct disulfide-bridge formation and re-organization of disulphide -bridges in target muscle proteins [30].

3. Small heat shock proteins

The systems biological concept of the chaperonome is outlined in Figure 2 [15, 16], which specifically summarizes the biochemical criteria for the classification of the chaperone family of small heat shock proteins (sHSP). Several members of the sHSP/HSPB class of molecular chaperones respond swiftly to stressful stimuli during strenuous exercise or path-ophysiological insults [31-33]. The efficient disintegration of poly-disperse protein assemblies into smaller subunits is a key cytoprotective function of HSPB molecules and helps to counter-act the potentially harmful side effects from toxic protein aggregates [11]. The grouping of 10 distinct proteins, named HSPB1 to HSPB10 [13, 34], is based on (i) exhibiting distinct ATP-independent chaperoning activities [35-37], (ii) the presence of a conserved α-crystallin domain towards the carboxy-terminal region that spans approximately 90 residues [38-41], (iii) relatively low molecular masses ranging from approximately 10 to 30 kDa [34], and (iv) the capability of forming high-molecular-mass oligomers [42-44]. The 10 HSPB molecules and their tissue distribution are listed in Table 1. HSPB1, HSPB5, HSPB6 and HSPB8 are ubiquitously expressed throughout the body and HSPB9 and HSPB10 are restricted to testis [44]. HSPB1 to HSPB8, also referred to as HSP25, MKBP, Hsp27, αAC, αBC, HSP20, cvHSP, HSP22, respectively, are present at various concentrations in skeletal muscles and provide high-affinity binding platforms for partially misfolded or unfolded muscle proteins [45-48].

The expression of the cardiovascular cvHSP/HSPB7 chaperone is restricted to cardiac and skeletal muscles [47] and the highest concentration of αBC/HSPB5 among the non-lenticular cell types is in slow-twitching oxidative muscle fibres [45]. Major HSPB molecules are induced to prevent

Table 1. Summary of the human protein family of small heat shock proteins

Small heat shock protein	Protein Accession	Tissue distribution
HSPB1 (Hsp25, Hsp27)	P04792	Ubiquitous; high levels in muscle tissues
HSPB2 (MKBP)	Q16082	Skeletal, cardiac and smooth muscles
HSPB3 (Hsp27)	Q12988	Skeletal, cardiac and smooth muscles
HSPB4 (aA-Crystallin; aAC)	P02489	Highly abundant in the eye lens; low concentration in skeletal muscle tissues
HSPB5 (aB-crystallin; aBC)	P02511	Ubiquitous; high levels in skeletal and cardiac muscles
HSPB6 (Hsp20)	O14558	Ubiquitous; high levels in skeletal muscles
HSPB7 (cvHsp)	Q9UBY9	Skeletal and cardiac muscles
HSPB8 (Hsp22)	Q9UJY1	Ubiquitous; moderate levels in muscles
HSPB9 (Hsp20)	Q9BQS6	Testis
HSPB10 (ODF1)	Q14990	Testis

detrimental protein aggregation and are intrinsically linked to the association and modulation of the highly organized assembly of cytoskeletal protein networks, including actin, desmin, tubulin and vimentin [49-52]. HSPBs play a crucial role during myogenesis and the differentiation of mature contractile fibres in adult motor units [53-55]. Complex formation between HSPB2 and HSPB3 regulate myogenic differentiation steps [56] and modulate the muscle-specific transcription factor MyoD during fibre development [57]. In mature skeletal muscles, HSP molecules play a key role in preventing tissue damage during extensive repeats of excitation-contraction-relaxation cycles, which represents a major type of physiological stressor [58]. Strenuous exercise usually results in a robust and sustained up-regulation of molecular chaperones [59, 60], including especially HSPB molecules in skeletal muscle tissues [61, 62].

4. Proteomic profiling of small heat shock proteins from skeletal muscle

The highly adaptive neuromuscular system is heterogeneous in its molecular and cellular composition, extremely plastic in response to altered physiological demands, sensitive to mechanical unloading, vulnerable to traumatic injury and sensitive to altered metabolic states [63]. This dynamic nature of the musculature and the fact that muscle cells are highly abundant in the body makes contractile fibres and their supportive tissues exceedingly susceptible to various physiological and pathophysiological stressors [17]. Skeletal muscles therefore require a sophisticated and dynamic chaperoning system to prevent extensive cellular damage via protein unfolding and/or toxic protein aggregation [18-20]. The ATP-independent chaperones of the HSPB family of low-molecular-mass HSPs play a central role in these cytoprotective mechanisms [64] and provide considerable levels of stress tolerance by efficiently targeting misfolded muscle proteins for peptide refolding or degradation [17]. Comprehensive proteomic cataloguing studies that have focused on skeletal muscle preparations have identified a large number of molecular chaperones, including major representatives of the HSPB protein network [65-69]. This is illustrated here by the listings of major chaperoning molecules present in mouse skeletal muscle, as judged by the routine mass spectrometric evaluation of total muscle preparations versus the microsomal fraction.

An Ultimate 3,000 NanoLC system (Dionex Corporation, Sunnyvale, CA, USA) coupled to a Q-Exactive mass spectrometer (Thermo Fisher Scientific, Dublin, Ireland) was used for the label-free liquid chromatography mass spectrometric analysis of hind limb muscles from 6-month old C57BL6 mice, as recently described in detail [70]. Processing of the raw data generated from LC-MS/MS analysis was carried out using Progenesis QI for Proteomics software (version 3.1; Non-Linear Dynamics, a Waters company, Newcastle upon Tyne, UK). Data alignment was based on the LC retention time of each sample. The data was filtered using certain criteria prior to exporting the MS/MS data files to Proteome Discoverer 1.4 (Thermo Scientific): (i) peptide features with ANOVA < 0.05 between experimental groups, (ii) mass peaks with charge states from +1 to +5 and (iii) greater than one isotope per peptide [70]. A PepXML generic file was generated from all exported MS/MS spectra from Progenesis software. This file was used for peptide identification using Proteome Discoverer 1.4 against Mascot (version 2.3, Matrix Science, Boston, MA, USA) and Sequest HT (SEQUEST HT algorithm, licence Thermo Scientific, registered trademark University of Washington, USA) and searched against the UniProtKB-SwissProt database (taxonomy: *Mus musculus*). The following search parameters were used for protein identification: (i) peptide mass

Table 2. Routine identification of major chaperones in mouse hind limb muscle as revealed by the label-free mass spectrometric analysis of total tissue extracts..

Molecular chaperone	Protein Accession	Coverage (%)	Unique peptides	Molecular mass (kDa)
Heat shock protein HSPA1A (Hsp70, Hsp70A1A)	Q61696	9.98	1	70.0
Heat shock protein HSPA4 (Hsp70, HspA4)	Q61316	7.02	4	94.1
Heat shock protein HSPA5 (GRP78, mitochondrial HspA5)	P20029	30.08	15	72.4
Heat shock protein HSPA8 (Hsp70, Hsp71, HspA8)	P63017	49.69	25	70.8
Heat shock protein HSPA9 (GRP75, mitochondrial HspA9)	P38647	43.15	21	73.4
Heat shock protein HSPB1 (Hsp25, HspB1, Hsp beta-1)	P14602	51.20	8	23.0
Heat shock protein HSPB2 (MKBP, HspB2, Hsp beta-2)	Q99PR8	37.36	4	20.4
Heat shock protein HSPB6 (Hsp20, HspB6, Hsp beta-6)	Q5EBG6	59.26	6	17.5
Heat shock protein HSPB7 (cvHsp, HspB7, Hsp beta-7)	P35385	11.24	1	18.6
Heat shock protein HSPCA (Hsp90, Hsp90aa1, HspCA)	P07901	15.01	1	84.7
Heat shock protein HSPCB (Hsp90, Hsp90ab1, HspCB)	P11499	42.96	15	83.2
Heat shock protein HSPD1 (Hsp60, mitochondrial)	P63038	49.56	19	60.9
Heat shock protein HSPE1, (Hsp10, mitochondrial)	Q64433	52.94	6	11.0
DnaJ protein DNJA2 (Hsp40)	Q9QYJ0	6.80	2	45.7
Calreticulin (CAL, ERp60)	P14211	20.43	5	48.0
Protein disulfide isomerase A1 (PDI-A1, ERp59)	P09103	34.38	12	57.0
Protein disulfide-isomerase A3 (PDI-A3, ERp57)	P27773	10.69	4	56.6
Peptidyl-prolyl cis-trans isomerase A (PPI-A, PPIase A)	P17742	49.39	5	18.0
Peptidyl-prolyl cis-trans isomerase B (PPI-B, PPIase B)	P24369	10.19	2	23.7

tolerance set to 10 ppm, (ii) MS/MS mass tolerance set to 0.02 Da, (iii) up to two missed cleavages were allowed, (iv) carbamidomethylation set as a fixed modification and (v) methionine oxidation set as a variable modification [70]. For re-importation back into Progenesis LC-MS software for further analysis, only peptides with either ion scores of 40.00 or more (from Mascot) and peptides with XCorr scores >1.9 for singly charged ions, >2.2 for doubly charged ions and >3.75 for triply charged ions or more (from Sequest HT) were selected [70].

Crude muscle homogenates were shown to contain 19 major chaperones, including HSPA1A, HSPA4, HSPA5, HSPA8, HSPA9, HSPCA, HSPCB, HSPD1, HSPE1, DNJA2, CAL, PDI-A1, PDI-A3, PPI-A and PPI-B, as well as the sHSP molecules HSPB1, HSPB2, HSPB6 and HSPB7 (Table 2). The microsomal fraction, isolated by an optimized standard subcellular fractionation procedure for the depletion of the contractile apparatus and enrichment of the membrane fraction [70, 71], contained 28 chaperoning protein species. This included the sHSP molecules HSPB1, HSPB2, HSPB6 and HSPB7, as well as HSPA1A, HSPA1L, HSPA4, HSPA5, HSPA8, HSPA9, HSPCA, HSPCB, HSP90b1, HSPD1, HSPE1, Trap1, DNJA2, DNJA3, DJB11, DNJC3, DJC11, CAL, PPI-A, PPI-B, PDI-D, PDI-A1, PDI-A3 and PDI-A6 (Table 3). The fact that a considerable number of DNAJ and GRP proteins from the sarcoplasmic reticulum and mitochondria were identified to be present in the membrane fraction [25] underlines the successful application of subcellular fractionation [72]. The proteomic hits from the microsomal study were further characterized by standard bioinformatics using the STRING database of known and predicted protein interactions that include direct physical and indirect functional protein associations [73]. Figure 3 outlines the close interaction network of chaperoning proteins from skeletal muscles, including HSPB1, HSPB2, HSPB6 and HSPB7, emphasising the crucial importance of sHSP molecules for cytoprotection in tissues with a high degree of vulnerability to environmental, physiological or pathological stressors [74].

5. Comparative proteomic profiling of HSPB in skeletal muscles

Molecular chaperones are involved in a variety of neurodegenerative diseases, neuromuscular pathologies and the

Table 3. Identification of major chaperones in mouse hind limb muscle as revealed by the label-free mas spectrometric analysis of microsomes.

Molecular chaperone	Protein Accession	Coverage (%)	Unique peptides	Molecular mass (kDa)
Heat shock protein HSPA1A (Hsp70, Hsp70A1A)	Q61696	9.98	1	70.0
Heat shock protein HSPA1L (Hsp70, Hsp70A1L)	P16627	8.89	1	70.6
Heat shock protein HSPA4 (Hsp74, Hsp70RY, HspA4)	Q61316	7.25	4	94.1
Heat shock protein HSPA5 (GRP78, mitochondrial HspA5)	P20029	49.77	28	72.4
Heat shock protein HSPA8 (Hsp70, Hsp71, HspA8)	P63017	50.77	23	70.8
Heat shock protein HSPA9 (GRP75, mitochondrial HspA9)	P38647	40.35	20	73.4
Heat shock protein HSPB1 (Hsp25, HspB1, Hsp beta-1)	P14602	70.81	10	23.0
Heat shock protein HSPB2 (MKBP, HspB2, Hsp beta-2)	Q99PR8	43.96	5	20.4
Heat shock protein HSPB6 (Hsp20, HspB6, Hsp beta-6)	Q5EBG6	59.26	6	17.5
Heat shock protein HSPB7 (cvHsp, HspB7, Hsp beta-7)	P35385	38.46	3	18.6
Heat shock protein HSPCA (Hsp90, Hsp90aa1, HspCA)	P07901	9.28	1	84.7
Heat shock protein HSPCB (Hsp90, Hsp90ab1, HspCB)	P11499	30.39	12	83.2
Heat shock protein Hsp90b1 (Endoplasmin, GRP94)	P08113	30.92	15	92.4
Heat shock protein HSPD1 (Hsp60, mitochondrial)	P63038	58.12	23	60.9
Heat shock protein HSPE1, (Hsp10, mitochondrial)	Q64433	52.94	6	11.0
Trap1 (TNFR-associated protein 1, mitochondrial Hsp75)	Q9CQN1	4.11	1	80.2
DnaJ protein DNJA2 (Hsp40, DNJ3)	Q9QYJ0	7.77	1	45.7
DnaJ protein DNJA3 (mitochondrial Hsp40, mTid-1)	Q99M87	3.96	2	52.4
DnaJ protein DJB11 (ER Hsp40 co-chaperone of HSPA5)	Q99KV1	9.22	2	40.5
DnaJ protein DNJC3 (Hsp40, protein kinase inhibitor p58)	Q91YW3	5.95	1	57.4
DnaJ protein DJC11 (mitochondrial Hsp40, DNAJC11)	Q5U458	3.22	1	63.2
Calnexin (CAL, CALX, CNX)	P35564	11.34	4	67.2
Peptidyl-prolyl cis-trans isomerase A (PPI-A, PPIase A)	P17742	55.49	6	18.0
Peptidyl-prolyl cis-trans isomerase B (PPI-B, PPIase B)	P24369	10.19	2	23.7
Peptidyl-prolyl cis-trans isomerase D (PPI-D, PPIase D)	Q9CR16	4.32	1	40.7
Protein disulfide isomerase A1 (PDI-A1, ERp59)	P09103	44.79	15	57.0
Protein disulfide-isomerase A3 (PDI-A3, ERp57)	P27773	33.86	15	56.6
Protein disulfide-isomerase A6 (PDI-A6, TXNDC7)	Q922R8	12.05	3	48.1

natural aging process [17, 75-78]. Besides their essential neuroprotective functions, HSP molecule are associated with primary chaperonopathies and secondary alterations during pathophysiological insults. Neuromuscular disorders, metabolic diseases and neuropathies that are closely linked to altered expression levels in the HSPB family of molecular chaperones are myotonic dystrophy, myofibrillar myopathies, Duchenne muscular dystrophy, the dysferlinopathies

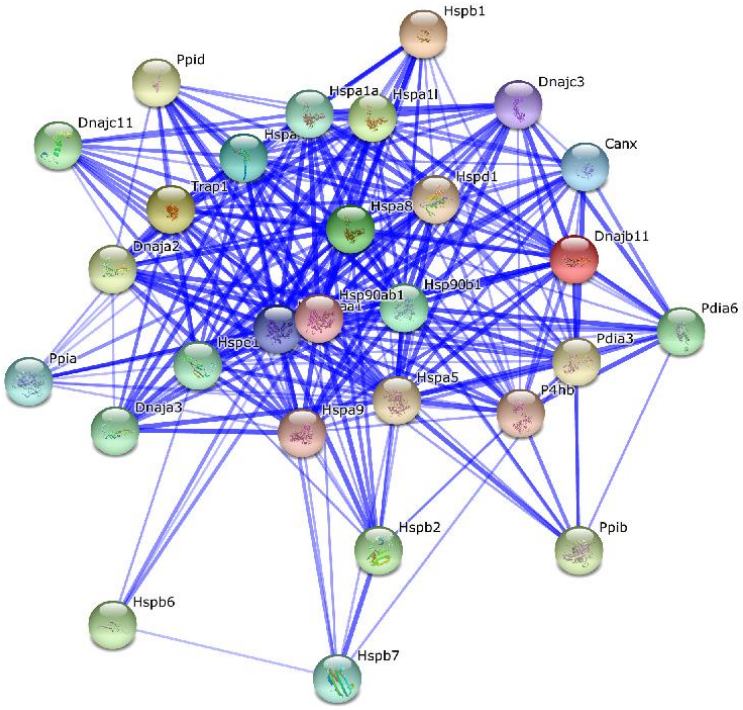

Figure 3. *Interaction map of the chaperoning protein system from mouse hind leg muscles.* The bioinformatics STRING database [73] was used to generate a protein interaction map with known and predicted protein associations that include direct physical and indirect functional protein linkages of mass spectrometrically identified molecular chaperones, including the sHSP molecules HSBP1, HSPB2, HSPB6 and HSPB7, as well as HSPA1A, HSPA1L, HSPA4, HSPA5, HSPA8, HSPA9, HSPCA, HSPCB, HSP90b1, HSPD1, HSPE1, Trap1, DNJA2, DNJA3, DJB11, DNJC3, DJC11, CAL, PPI-A, PPI-B, PDI-D, PDI-A1, PDI-A3 and PDI-A6 (Table 3).

termed Miyoshi myopathy and limb-girdle muscular dystrophy type LGMD2B, myotonia-related hyperexcitability, collagen VI myopathy, Charcot-Marie-Tooth disease type 2, distal hereditary motor neuropathy, motor neuron disease, type 2 diabetes-related muscle weakness and sarcopenia of old age, as listed in Table 4. Certain desmin-related myopathies (α-crystallinopathy) are characterized by abnormalities in HSPB5 [79-81]. HSPB1, HSPB3 and HSPB8 are primarily affected in distal motor neuropathy and the axonal form of Charcot-Marie-Tooth disease type 2 [82-86]. Figure 4 summarizes changes in HSPB molecules during physiological adaptations and in response to pathophysiological insults. A large number of comparative proteomic studies have focused on the molecular fate of HSP chaperones during stressful conditions. This has included the systematic analysis of myoblast differentiation and protein secretion during myogenesis [87-89], the fibre type specification in fast versus slow muscles [90-92], the physiological modifications following endurance exercise or resistance training [93-96], fast-to-slow muscle transformation following chronic low-frequency stimulation [97-99], skeletal muscle hypertrophy [100], muscular atrophy following immobilization or denervation [101-103], hypoxia-related stress [104] and sarcopenia of old age [105-110]. Exercise, fibre transitions, disuse atrophy and hypertrophy are clearly related to distinct changes in HSPB molecules.

Systematic surveys of changes in HSP chaperones in neuromuscular pathologies has included dysferlinopathy [111], Duchenne muscular dystrophy [17, 70, 71, 112-126], myofibrillar myopathies with abnormalities in HSPB5, desmin, filamin or myotilin [127-130], myasthenia gravis [131], myotonia-related hyperexcitability [132], motor neuron diseases including amyotrophic lateral sclerosis [133-137], hypokalemic myopathy [138], obesity-related muscular weakness and/or type 2 diabetes-associated insulin resistance [139-144], burn sepsis-related stress [145] and *post mortem* changes in skeletal muscle samples [146]. The differential expression of HSPB chaperones in pathological skeletal muscles is clearly related to the requirement of swiftly dispersing toxic protein aggregates and facilitating the degradation of misfolded muscle proteins under stressed conditions. The bioanalytical strategies used in the comparative proteomic profiling of HSP molecules in stressed skeletal muscles have relied on sophisticated labeling methods, efficient protein separation techniques and sensitive mass spectrometry. Often detergent- or urea-based extraction methods were used to prepare crude muscle extracts and the identification of changed proteins was conducted with various label-based or label-free approaches [17, 63].

In general, quantitative proteomic studies are routinely carried out with chemical or metabolic labeling and frequently include the isotopic tagging of peptides and proteins prior to mass spectrometric analysis. Cellular proteomics employs a variety of relative quantitation methods, including

Table 4. Major neuromuscular pathologies associated with primary or secondary abnormalities in small heat shock proteins.

Small heat shock protein	Pathological involvement
HSPB1 (Hsp25)	Charcot-Marie-Tooth disease type 2 Distal hereditary motor neuropathy Motor neuron disease Type 2 diabetes-associated muscle weakness Collagen VI myopathy Dysferlinopathy
HSPB2 (MKBP)	Myotonic dystrophy
HSPB3 (Hsp27)	Motor neuropathy Duchenne muscular dystrophy Myotonia-related hyperexcitability Sarcopenia of old age
HSPB5 (αBC)	Myofibrillar myopathies (desminopathy, crystallinopathy) Duchenne muscular dystrophy Sarcopenia of old age Motor neuron disease Muscular atrophy Type 2 diabetes-associated muscle weakness
HSPB6 (Hsp20)	Sarcopenia of old age
HSPB7 (cvHsp)	Duchenne muscular dystrophy Sarcopenia of old age
HSPB8 (Hsp22)	Charcot-Marie-Tooth disease type 2 Distal hereditary motor neuropathy

iTRAQ (isobaric Tags for Relative and Absolute Quantitation), ICAT (Isotope-Coded Affinity Tag) and SILAC (Stable Isotope Labeling with Amino acids in Cell culture) [147-149]. Subproteomic studies that focus on isolated organelles for the reduction of sample complexity involve advanced methods such as LCM (laser capture microscopy) [150] and LOPIT (Localization of Organelle Proteins by Isotope Tagging) [151]. Large-scale protein separation is usually achieved by gel electrophoretic methods and/or liquid chromatography. An efficient pre-electrophoretic labeling method for the comparative analysis of isolated proteomes is two-dimensional fluorescence difference in-gel electrophoresis (DIGE) [152-154]. Importantly, the extraordinary improvements of mass spectrometers in relation to mass accuracy, sensitivity, resolving power, dynamic range, throughput capacity and available fragmentation modes has greatly increased the coverage of the assessable proteome. Modern peptide mass analyzers rely on time-of-flight technology, linear ion traps, quadrupole, orbitrap or fourier transform ion cyclotron resonance methodologies [155-158]. Studies with a focus on HSP molecules have used many of these standard proteomic techniques. Since HSPB molecules are mostly low-molecular-mass proteins and relatively soluble as compared to many other protein species in skeletal muscles, they can be easily separated by gel electrophoretic procedures or liquid chromatography. For example, the application of the fluorescence DIGE technique has resulted in the identification of considerably increased levels of cvHSP/HSPB7 and αBC/HSPB5 in both dystrophic and senescent muscle fibres [107, 108]. Thus, although two-dimensional gel electrophoresis underestimates the presence of integral membrane proteins, high-molecular-mass proteins and low copy number proteins, it is highly suited for studying the large family of HSP molecules [17]. Label-free mass spectrometry has also been successfully employed to compare normal versus stressed muscle specimens and identified distinct changes in HSP molecules [70, 125, 126]. Since skeletal muscles contain some of the largest proteins in the human body, such as nebulin, obscurin and titin, alternative on-membrane digestion approaches have been developed to study these giant proteins [159-161]. In future studies, gradient gel electrophoretic separation in combination with on-

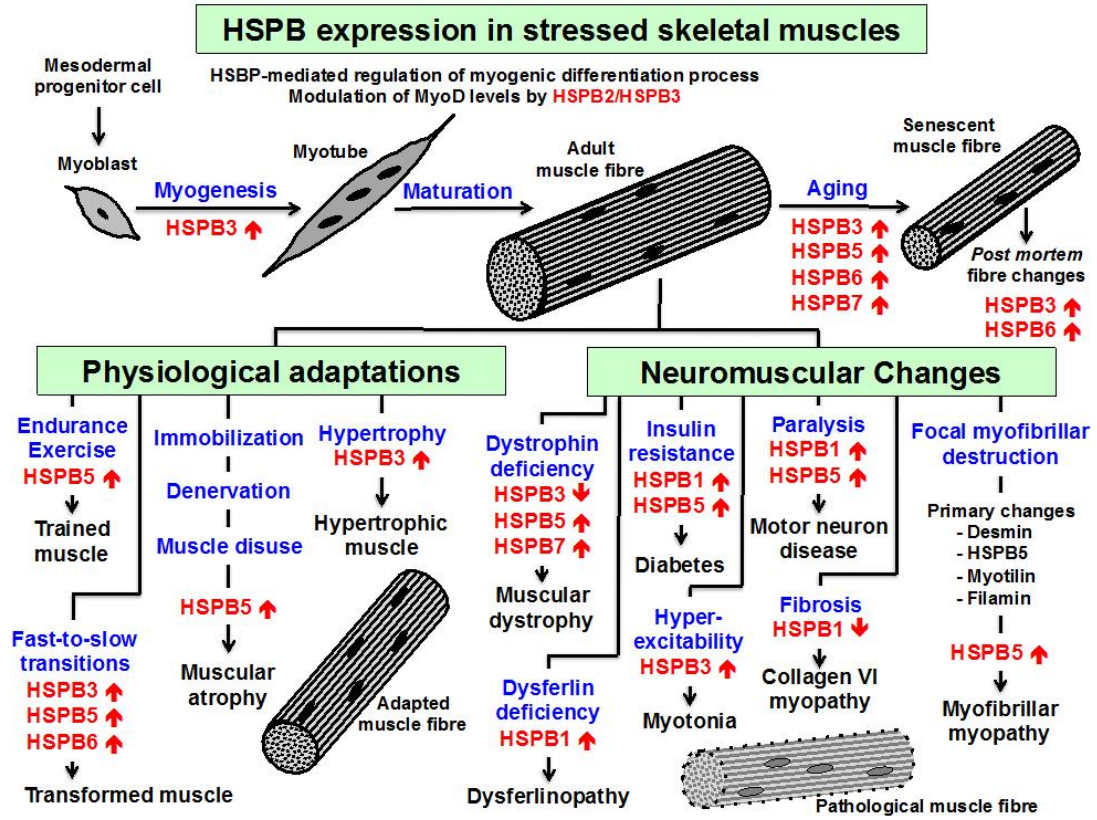

Figure 4. Overview of changes in HSPB molecules during physiological adaptations and in response to pathophysiological insults.

membrane digestion might also be useful to investigate the interaction patterns between HSPB species within supramolecular protein assemblies, such as the actomyosin apparatus or the Z-disc region of sarcomeres.

6. Concluding Remarks

The intrinsic role of HSPB molecules as chaperoning proteins makes them key components involved in the fine regulation of cellular proteostasis. In stressed skeletal muscles, HSPB chaperones are of central importance for the facilitation of various adaptive processes and cellular pro-survival mechanisms. This makes changes in their concentration, oligomerization, post-translational modifications and/or subcellular re-localization a characteristic feature of both stress tolerance and muscle damage pathways. These distinct alterations can be potentially exploited for the establishment of superior biomarker signatures of physiological adaptations and pathological changes [162-165]. Muscle-specific HSPB molecules are of low molecular mass and belong to the relatively soluble type of proteins. These biochemical and physicochemical properties make them good candidates for developing simple bioassays to predict, diagnose and evaluate stress-related skeletal muscle degeneration. Since HSPB chaperones are critical cytoprotective factors that prevent and reverse stress-related fibre damage, they are also potentially useful as therapy-monitoring biomarkers. In the future, the evaluation of disease model systems or routine testing in preclinical studies may profit from the usage of HSPB molecules as reliable and robust muscle-associated markers of cellular stress. This gives the detailed characterization of the HSPB chaperonome considerable importance for advancing several biomedical areas, including basic myology, applied physiology, sports medicine and neuromuscular pathology.

Acknowledgements

The authors would like to thank Muscular Dystrophy Ireland, the Hume scholarship programme of Maynooth University and the Deutsche Duchenne Stiftung *aktion benni & co e.V.* for funding of our proteomics research projects focusing on skeletal muscle diseases. We thank Dr. Paula Meleady and Mr. Michael Henry (Dublin City University), Prof. Dieter Swandualla (University of Bonn), Prof. Thomas Schmitt-John (Aarhus University) and Prof. Heinrich Brinkmeier (University of Greifswald) for their continued support of our research initiative to identify proteomic biomarkers of neuromuscular disorders.

References

[1] F. Ritossa, Experientia, 18 (1962) 571-573. http://dx.doi.org/10.1007/bf02172188

[2] A. Tissiéres, H.K. Mitchell, U.M. Tracy, Journal of Molecu-

lar Biology, 84 (1974) 389-398. http://dx.doi.org/10.1016/0022-2836(74)90447-1

[3] S.C. Wadsworth, E.A. Craig, B.J. McCarthy, Proceedings of the National Academy of Sciences, 77 (1980) 2134-2137. http://dx.doi.org/10.1073/pnas.77.4.2134

[4] J. Vazquez, D. Pauli, A. Tissieres, Chromosoma, 102 (1993) 233-248. http://dx.doi.org/10.1007/bf00352397

[5] M. Ashburner, J.J. Bonner, Cell, 17 (1979) 241-254. http://dx.doi.org/10.1016/0092-8674(79)90150-8

[6] E.A. Craig, M.J. Schlesinger, Critical Reviews in Biochemistry and Molecular Biology, 18 (1985) 239-280. http://dx.doi.org/10.3109/10409238509085135

[7] S. Lindquist, Annu Rev Biochem, 55 (1986) 1151-1191. http://dx.doi.org/10.1146/annurev.bi.55.070186.005443

[8] R.J. Ellis, S.M. van der Vies, Annu Rev Biochem, 60 (1991) 321-347. http://dx.doi.org/10.1146/annurev.bi.60.070191.001541

[9] R.I. Morimoto, M.P. Kline, D.N. Bimston, J.J. Cotto, Essays Biochem, 32 (1997) 17-29

[10] R. Voellmy, F. Boellmann, Adv Exp Med Biol, 594 (2007) 89-99. http://dx.doi.org/10.1007/978-0-387-39975-1_9

[11] Y.E. Kim, M.S. Hipp, A. Bracher, M. Hayer-Hartl, F.U. Hartl, Annu Rev Biochem, 82 (2013) 323-355. http://dx.doi.org/10.1146/annurev-biochem-060208-092442

[12] M.J. Vos, J. Hageman, S. Carra, H.H. Kampinga, Biochemistry, 47 (2008) 7001-7011. http://dx.doi.org/10.1021/bi800639z

[13] H.H. Kampinga, J. Hageman, M.J. Vos, H. Kubota, R.M. Tanguay, E.A. Bruford, M.E. Cheetham, B. Chen, L.E. Hightower, Cell Stress Chaperones, 14 (2009) 105-111. http://dx.doi.org/10.1007/s12192-008-0068-7

[14] K. Richter, M. Haslbeck, J. Buchner, Mol Cell, 40 (2010) 253-266. http://dx.doi.org/10.1016/j.molcel.2010.10.006

[15] V. Kakkar, M. Meister-Broekema, M. Minoia, S. Carra, H.H. Kampinga, Dis Model Mech, 7 (2014) 421-434. http://dx.doi.org/10.1242/dmm.014563

[16] R.I. Morimoto, Cold Spring Harb Symp Quant Biol, 76 (2011) 91-99. http://dx.doi.org/10.1101/sqb.2012.76.010637

[17] H. Brinkmeier, K. Ohlendieck, Proteomics Clin Appl, 8 (2014) 875-895. http://dx.doi.org/10.1002/prca.201400015

[18] M.B. Smolka, C.C. Zoppi, A.A. Alves, L.R. Silveira, S. Marangoni, L. Pereira-Da-Silva, J.C. Novello, D.V. Macedo, American Journal of Physiology-Regulatory Integrative and Comparative Physiology, 279 (2000) R1539-R1545

[19] Y. Liu, Frontiers in Bioscience, 11 (2006) 2802. http://dx.doi.org/10.2741/2011

[20] S.M. Senf, Front Physiol, 4 (2013) 330. http://dx.doi.org/10.3389/fphys.2013.00330

[21] B. De Paepe, K.K. Creus, J.J. Martin, J. Weis, J.L. De Bleecker, Ann N Y Acad Sci, 1173 (2009) 463-469. http://dx.doi.org/10.1111/j.1749-6632.2009.04812.x

[22] A. Wagatsuma, M. Shiozuka, N. Kotake, K. Takayuki, H. Yusuke, K. Mabuchi, R. Matsuda, S. Yamada, Mol Cell Biochem, 358 (2011) 265-280. http://dx.doi.org/10.1007/s11010-011-0977-0

[23] J. Li, J. Buchner, Biomed J, 36 (2013) 106-117. http://dx.doi.org/10.4103/2319-4170.113230

[24] D.P. Easton, Y. Kaneko, J.R. Subjeck, Cell Stress & Chaperones, 5 (2000) 276. http://dx.doi.org/10.1379/1466-1268(2000)005<0276:thagsp>2.0.co;2

[25] V. Kakkar, L. C. B. Prins, H. H. Kampinga, Current Topics in Medicinal Chemistry, 12 (2013) 2479-2490. http://dx.doi.org/10.2174/1568026611212220004

[26] Y. Oishi, K. Taniguchi, H. Matsumoto, A. Ishihara, Y. Ohira, R.R. Roy, J Appl Physiol (1985), 92 (2002) 1097-1103. http://dx.doi.org/10.1152/japplphysiol.00739.2001

[27] F. Cappello, A. Marino Gammazza, A. Palumbo Piccionello, C. Campanella, A. Pace, E. Conway de Macario, A.J. Macario, Expert Opin Ther Targets, 18 (2014) 185-208. http://dx.doi.org/10.1517/14728222.2014.856417

[28] D.B. Williams, J Cell Sci, 119 (2006) 615-623. http://dx.doi.org/10.1242/jcs.02856

[29] A. G. Erlejman, M. Lagadari, D. C. Harris, M. B. Cox, M. D. Galigniana, Current Protein & Peptide Science, 15 (2014) 205-215. http://dx.doi.org/10.2174/1389203715666140331113753

[30] C.A. Ruiz, R.L. Rotundo, J Biol Chem, 284 (2009) 31753-31763. http://dx.doi.org/10.1074/jbc.M109.038471

[31] I.J. Benjamin, Journal of Biological Chemistry, 271 (1996) 24089-24095. http://dx.doi.org/10.1074/jbc.271.39.24089

[32] P.D. Neufer, G.A. Ordway, R.S. Williams, Am J Physiol, 274 (1998) C341-346

[33] G.C. Melkani, A. Cammarato, S.I. Bernstein, J Mol Biol, 358 (2006) 635-645. http://dx.doi.org/10.1016/j.jmb.2006.02.043

[34] G. Kappé, E. Franck, P. Verschuure, W.C. Boelens, J.A.M. Leunissen, W.W. de Jong, Cell Stress & Chaperones, 8 (2003) 53. http://dx.doi.org/10.1379/1466-1268(2003)8<53:thgecs>2.0.co;2

[35] R.V. Montfort, C. Slingsby, E. Vierlingt, 59 (2001) 105-156. http://dx.doi.org/10.1016/s0065-3233(01)59004-x

[36] F. Narberhaus, Microbiology and Molecular Biology Reviews, 66 (2002) 64-93. http://dx.doi.org/10.1128/mmbr.66.1.64-93.2002

[37] W.W. de Jong, J.A. Leunissen, C.E. Voorter, Mol Biol Evol, 10 (1993) 103-126

[38] K.B. Merck, J. Horwitz, M. Kersten, P. Overkamp, M. Gaestel, H. Bloemendal, W.W. de Jong, Molecular Biology Reports, 18 (1993) 209-215. http://dx.doi.org/10.1007/bf01674432

[39] R. Stamler, G. Kappe, W. Boelens, C. Slingsby, J Mol Biol, 353 (2005) 68-79. http://dx.doi.org/10.1016/j.jmb.2005.08.025

[40] E.V. Baranova, S.D. Weeks, S. Beelen, O.V. Bukach, N.B. Gusev, S.V. Strelkov, J Mol Biol, 411 (2011) 110-122. http://dx.doi.org/10.1016/j.jmb.2011.05.024

[41] W.C. Boelens, Prog Biophys Mol Biol, 115 (2014) 3-10. http://dx.doi.org/10.1016/j.pbiomolbio.2014.02.005

[42] H.H. Kampinga, C. Garrido, Int J Biochem Cell Biol, 44 (2012) 1706-1710. http://dx.doi.org/10.1016/j.biocel.2012.06.005

[43] A.P. Arrigo, FEBS Lett, 587 (2013) 1959-1969. http://dx.doi.org/10.1016/j.febslet.2013.05.011

[44] R. Bakthisaran, R. Tangirala, C.M. Rao, Biochim Biophys Acta, 1854 (2015) 291-319. http://dx.doi.org/10.1016/j.bbapap.2014.12.019

[45] T. Iwaki, A. Iwaki, J. Goldman, Acta Neuropathologica, 85 (1993). http://dx.doi.org/10.1007/bf00230485

[46] P. Verschuure, C. Tatard, W.C. Boelens, J.F. Grongnet, J.C. David, Eur J Cell Biol, 82 (2003) 523-530. http://dx.doi.org/10.1078/0171-9335-00337

[47] N. Golenhofen, M.D. Perng, R.A. Quinlan, D. Drenckhahn, Histochem Cell Biol, 122 (2004) 415-425. http://dx.doi.org/10.1007/s00418-004-0711-z

[48] R.L. Neppl, M. Kataoka, D.Z. Wang, J Biol Chem, 289 (2014) 17240-17248. http://dx.doi.org/10.1074/jbc.M114.549584

[49] I.D. Nicholl, R.A. Quinlan, EMBO J, 13 (1994) 945-953

[50] T. Sakurai, Y. Fujita, E. Ohto, A. Oguro, Y. Atomi, FASEB J, 19 (2005) 1199-1201. http://dx.doi.org/10.1096/fj.04-3060fje

[51] H. Jee, T. Sakurai, S. Kawada, N. Ishii, Y. Atomi, J Physiol Sci, 59 (2009) 149-155. http://dx.doi.org/10.1007/s12576-008-0014-6

[52] P. Graceffa, Biochem Res Int, 2011 (2011) 901572. http://dx.doi.org/10.1155/2011/901572

[53] H. Ito, K. Kamei, I. Iwamoto, Y. Inaguma, K. Kato, Exp Cell Res, 266 (2001) 213-221. http://dx.doi.org/10.1006/excr.2001.5220

[54] Q. Zhang, H.G. Lee, S.K. Kang, M. Baik, Y.J. Choi, Biotechnol Lett, 36 (2014) 1225-1231. http://dx.doi.org/10.1007/s10529-014-1489-2

[55] A.S. Adhikari, K. Sridhar Rao, N. Rangaraj, V.K. Parnaik, C. Mohan Rao, Exp Cell Res, 299 (2004) 393-403. http://dx.doi.org/10.1016/j.yexcr.2004.05.032

[56] Y. Sugiyama, A. Suzuki, M. Kishikawa, R. Akutsu, T. Hirose, M.M.Y. Waye, S.K.W. Tsui, S. Yoshida, S. Ohno, Journal of Biological Chemistry, 275 (2000) 1095-1104. http://dx.doi.org/10.1074/jbc.275.2.1095

[57] B.N. Singh, K.S. Rao, M. Rao Ch, Biochim Biophys Acta, 1803 (2010) 288-299. http://dx.doi.org/10.1016/j.bbamcr.2009.11.009

[58] E.G. Noble, K.J. Milne, C.W. Melling, Appl Physiol Nutr Metab, 33 (2008) 1050-1065. http://dx.doi.org/10.1139/H08-069

[59] K.A. Huey, S. Burdette, H. Zhong, R.R. Roy, Exp Physiol, 95 (2010) 1145-1155. http://dx.doi.org/10.1113/expphysiol.2010.054692

[60] M. Folkesson, A.L. Mackey, H. Langberg, E. Oskarsson, K. Piehl-Aulin, J. Henriksson, F. Kadi, Acta Physiol (Oxf), 209 (2013) 26-33. http://dx.doi.org/10.1111/apha.12124

[61] G. Paulsen, K.E. Hanssen, B.R. Ronnestad, N.H. Kvamme, I. Ugelstad, F. Kadi, T. Raastad, Eur J Appl Physiol, 112 (2012) 1773-1782. http://dx.doi.org/10.1007/s00421-011-2132-8

[62] K.T. Cumming, G. Paulsen, M. Wernbom, I. Ugelstad, T. Raastad, Acta Physiol (Oxf), 211 (2014) 634-646. http://dx.doi.org/10.1111/apha.12305

[63] K. Ohlendieck, Skelet Muscle, 1 (2011) 6. http://dx.doi.org/10.1186/2044-5040-1-6

[64] T.M. Treweek, S. Meehan, H. Ecroyd, J.A. Carver, Cell Mol Life Sci, 72 (2015) 429-451. http://dx.doi.org/10.1007/s00018-014-1754-5

[65] K. Hojlund, Z. Yi, H. Hwang, B. Bowen, N. Lefort, C.R. Flynn, P. Langlais, S.T. Weintraub, L.J. Mandarino, Mol Cell Proteomics, 7 (2008) 257-267. http://dx.doi.org/10.1074/mcp.M700304-MCP200

[66] K.C. Parker, R.J. Walsh, M. Salajegheh, A.A. Amato, B. Krastins, D.A. Sarracino, S.A. Greenberg, J Proteome Res, 8 (2009) 3265-3277. http://dx.doi.org/10.1021/pr800873q

[67] J.G. Burniston, J. Connolly, H. Kainulainen, S.L. Britton, L.G. Koch, Proteomics, 14 (2014) 2339-2344. http://dx.doi.org/10.1002/pmic.201400118

[68] M. Murgia, N. Nagaraj, A.S. Deshmukh, M. Zeiler, P. Cancellara, I. Moretti, C. Reggiani, S. Schiaffino, M. Mann, EMBO Rep, 16 (2015) 387-395. http://dx.doi.org/10.15252/embr.201439757

[69] A.S. Deshmukh, M. Murgia, N. Nagaraj, J.T. Treebak, J. Cox, M. Mann, Mol Cell Proteomics, 14 (2015) 841-853. http://dx.doi.org/10.1074/mcp.M114.044222

[70] S. Murphy, M. Henry, P. Meleady, M. Zweyer, R.R. Mundegar, D. Swandulla, K. Ohlendieck, Biology (Basel), 4 (2015) 397-423. http://dx.doi.org/10.3390/biology4020397

[71] S. Carberry, M. Zweyer, D. Swandulla, K. Ohlendieck, Anal Biochem, 446 (2014) 108-115. http://dx.doi.org/10.1016/j.ab.2013.08.004

[72] K. Ohlendieck, Journal of Integrated OMICS, 2 (2012). http://dx.doi.org/10.5584/jiomics.v2i2.111

[73] A. Franceschini, D. Szklarczyk, S. Frankild, M. Kuhn, M. Simonovic, A. Roth, J. Lin, P. Minguez, P. Bork, C. von Mering, L.J. Jensen, Nucleic Acids Res, 41 (2013) D808-815. http://dx.doi.org/10.1093/nar/gks1094

[74] E.V. Mymrikov, A.S. Seit-Nebi, N.B. Gusev, Physiol Rev, 91 (2011) 1123-1159. http://dx.doi.org/10.1152/physrev.00023.2010

[75] R.I. Morimoto, Genes Dev, 22 (2008) 1427-1438. http://dx.doi.org/10.1101/gad.1657108

[76] S.K. Calderwood, A. Murshid, T. Prince, Gerontology, 55 (2009) 550-558. http://dx.doi.org/10.1159/000225957

[77] F. Romi, G. Helgeland, N.E. Gilhus, Eur Neurol, 66 (2011) 65-69. http://dx.doi.org/10.1159/000329373

[78] E.E. Benarroch, Neurology, 76 (2011) 660-667. http://dx.doi.org/10.1212/WNL.0b013e31820c3119

[79] P. Vicart, A. Caron, P. Guicheney, Z. Li, M.C. Prevost, A. Faure, D. Chateau, F. Chapon, F. Tome, J.M. Dupret, D. Paulin, M. Fardeau, Nat Genet, 20 (1998) 92-95. http://dx.doi.org/10.1038/1765

[80] C.S. Clemen, H. Herrmann, S.V. Strelkov, R. Schroder, Acta Neuropathol, 125 (2013) 47-75. http://dx.doi.org/10.1007/s00401-012-1057-6

[81] A. Sanbe, Biological & Pharmaceutical Bulletin, 34 (2011) 1653-1658. http://dx.doi.org/10.1248/bpb.34.1653

[82] J. Irobi, K. Van Impe, P. Seeman, A. Jordanova, I. Dierick, N. Verpoorten, A. Michalik, E. De Vriendt, A. Jacobs, V. Van Gerwen, K. Vennekens, R. Mazanec, I. Tournev, D. Hilton-Jones, K. Talbot, I. Kremensky, L. Van Den Bosch, W. Robberecht, J. Van Vandekerckhove, C. Van Broeckhoven, J. Gettemans, P. De Jonghe, V. Timmerman, Nat Genet, 36 (2004) 597-601. http://dx.doi.org/10.1038/ng1328

[83] S. Capponi, A. Geroldi, P. Fossa, M. Grandis, P. Ciotti, R. Gulli, A. Schenone, P. Mandich, E. Bellone, J Peripher Nerv Syst, 16 (2011) 287-294. http://dx.doi.org/10.1111/j.1529-8027.2011.00361.x

[84] A.M. Rossor, J.M. Polke, H. Houlden, M.M. Reilly, Nat Rev Neurol, 9 (2013) 562-571. http://dx.doi.org/10.1038/nrneurol.2013.179

[85] K. Nakhro, J.M. Park, Y.J. Kim, B.R. Yoon, J.H. Yoo, H. Koo, B.O. Choi, K.W. Chung, Neuromuscul Disord, 23 (2013) 656-663. http://dx.doi.org/10.1016/j.nmd.2013.05.009

[86] K. Maeda, R. Idehara, A. Hashiguchi, H. Takashima, Internal Medicine, 53 (2014) 1655-1658. http://dx.doi.org/10.2169/internalmedicine.53.2843

[87] N.S. Tannu, V.K. Rao, R.M. Chaudhary, F. Giorgianni, A.E. Saeed, Y. Gao, R. Raghow, Mol Cell Proteomics, 3 (2004) 1065-1082. http://dx.doi.org/10.1074/mcp.M400020-MCP200

[88] F. Gonnet, B. Bouazza, G.A. Millot, S. Ziaei, L. Garcia, G.S. Butler-Browne, V. Mouly, J. Tortajada, O. Danos, F. Svinartchouk, Proteomics, 8 (2008) 264-278. http://dx.doi.org/10.1002/pmic.200700261

[89] J. Henningsen, K.T. Rigbolt, B. Blagoev, B.K. Pedersen, I. Kratchmarova, Mol Cell Proteomics, 9 (2010) 2482-2496. http://dx.doi.org/10.1074/mcp.M110.002113

[90] N. Okumura, A. Hashida-Okumura, K. Kita, M. Matsubae, T. Matsubara, T. Takao, K. Nagai, Proteomics, 5 (2005) 2896-2906. http://dx.doi.org/10.1002/pmic.200401181

[91] D. Capitanio, A. Vigano, E. Ricci, P. Cerretelli, R. Wait, C. Gelfi, Proteomics, 5 (2005) 2577-2586. http://dx.doi.org/10.1002/pmic.200401183

[92] H.C. Drexler, A. Ruhs, A. Konzer, L. Mendler, M. Bruckskotten, M. Looso, S. Gunther, T. Boettger, M. Kruger, T. Braun, Mol Cell Proteomics, 11 (2012) M111 010801. http://dx.doi.org/10.1074/mcp.M111.010801

[93] J.G. Burniston, Biochim Biophys Acta, 1784 (2008) 1077-1086. http://dx.doi.org/10.1016/j.bbapap.2008.04.007

[94] W. Yamaguchi, E. Fujimoto, M. Higuchi, I. Tabata, J Biochem, 148 (2010) 327-333. http://dx.doi.org/10.1093/jb/mvq073

[95] P.G. Gandra, R.H. Valente, J. Perales, A.G. Pacheco, D.V. Macedo, Proteomics, 12 (2012) 2663-2667. http://dx.doi.org/10.1002/pmic.201200137

[96] A. Roca-Rivada, O. Al-Massadi, C. Castelao, L.L. Senin, J. Alonso, L.M. Seoane, T. Garcia-Caballero, F.F. Casanueva, M. Pardo, J Proteomics, 75 (2012) 5414-5425. http://dx.doi.org/10.1016/j.jprot.2012.06.037

[97] P. Donoghue, P. Doran, P. Dowling, K. Ohlendieck, Biochim Biophys Acta, 1752 (2005) 166-176. http://dx.doi.org/10.1016/j.bbapap.2005.08.005

[98] P. Donoghue, P. Doran, K. Wynne, K. Pedersen, M.J. Dunn, K. Ohlendieck, Proteomics, 7 (2007) 3417-3430. http://dx.doi.org/10.1002/pmic.200700262

[99] K. Ohlendieck, Muscles Ligaments Tendons J, 1 (2011) 119-126

[100] M. Hamelin, T. Sayd, C. Chambon, J. Bouix, B. Bibe, D. Milenkovic, H. Leveziel, M. Georges, A. Clop, P. Marinova, E. Laville, J Anim Sci, 84 (2006) 3266-3276. http://dx.doi.org/10.2527/jas.2006-162

[101] Y. Seo, K. Lee, K. Park, K. Bae, I. Choi, J Biochem, 139 (2006) 71-80. http://dx.doi.org/10.1093/jb/mvj007

[102] H. Sun, J. Liu, F. Ding, X. Wang, M. Liu, X. Gu, J Muscle Res Cell Motil, 27 (2006) 241-250. http://dx.doi.org/10.1007/s10974-006-9067-4

[103] H. Sun, M. Li, L. Gong, M. Liu, F. Ding, X. Gu, Mol Cell Biochem, 364 (2012) 193-207. http://dx.doi.org/10.1007/s11010-011-1218-2

[104] C.A.t. Bosworth, C.W. Chou, R.B. Cole, B.B. Rees, Proteomics, 5 (2005) 1362-1371. http://dx.doi.org/10.1002/pmic.200401002

[105] I. Piec, A. Listrat, J. Alliot, C. Chambon, R.G. Taylor, D. Bechet, FASEB J, 19 (2005) 1143-1145. http://dx.doi.org/10.1096/fj.04-3084fje

[106] K. O'Connell, J. Gannon, P. Doran, K. Ohlendieck, International Journal of Molecular Medicine, (2007). http://dx.doi.org/10.3892/ijmm.20.2.145

[107] P. Doran, J. Gannon, K. O'Connell, K. Ohlendieck, Eur J Cell Biol, 86 (2007) 629-640. http://dx.doi.org/10.1016/j.ejcb.2007.07.003

[108] P. Doran, K. O'Connell, J. Gannon, M. Kavanagh, K. Ohlendieck, Proteomics, 8 (2008) 364-377. http://dx.doi.org/10.1002/pmic.200700475

[109] J. Feng, H. Xie, D.L. Meany, L.V. Thompson, E.A. Arriaga, T.J. Griffin, The Journals of Gerontology Series A: Biological Sciences and Medical Sciences, 63 (2008) 1137-1152. http://dx.doi.org/10.1093/gerona/63.11.1137

[110] P. Doran, P. Donoghue, K. O'Connell, J. Gannon, K. Ohlendieck, Proteomics, 9 (2009) 989-1003. http://dx.doi.org/10.1002/pmic.200800365

[111] S. De Palma, L. Morandi, E. Mariani, S. Begum, P. Cerretelli, R. Wait, C. Gelfi, Proteomics, 6 (2006) 379-385. http://dx.doi.org/10.1002/pmic.200500098

[112] Y. Ge, M.P. Molloy, J.S. Chamberlain, P.C. Andrews, Electrophoresis, 25 (2004) 2576-2585. http://dx.doi.org/10.1002/elps.200406013

[113] P. Doran, P. Dowling, P. Donoghue, M. Buffini, K. Ohlendieck, Biochim Biophys Acta, 1764 (2006) 773-785. http://dx.doi.org/10.1016/j.bbapap.2006.01.007

[114] D. Gardan-Salmon, J.M. Dixon, S.M. Lonergan, J.T. Selsby, Eur J Appl Physiol, 111 (2011) 2763-2773. http://dx.doi.org/10.1007/s00421-011-1906-3

[115] S. Rayavarapu, W. Coley, E. Cakir, V. Jahnke, S. Takeda, Y. Aoki, H. Grodish-Dressman, J.K. Jaiswal, E.P. Hoffman, K.J. Brown, Y. Hathout, K. Nagaraju, Mol Cell Proteomics, 12 (2013) 1061-1073. http://dx.doi.org/10.1074/mcp.M112.023127

[116] P. Doran, G. Martin, P. Dowling, H. Jockusch, K. Ohlendieck, Proteomics, 6 (2006) 4610-4621. http://dx.doi.org/10.1002/pmic.200600082

[117] P. Doran, S.D. Wilton, S. Fletcher, K. Ohlendieck, Proteomics, 9 (2009) 671-685. http://dx.doi.org/10.1002/pmic.200800441

[118] S. Carberry, M. Zweyer, D. Swandulla, K. Ohlendieck, Biology (Basel), 2 (2013) 1438-1464. http://dx.doi.org/10.3390/biology2041438

[119] C.Y. Matsumura, B. Menezes de Oliveira, M. Durbeej, M.J. Marques, PLoS One, 8 (2013) e65831. http://dx.doi.org/10.1371/journal.pone.0065831

[120] S. Carberry, H. Brinkmeier, Y. Zhang, C.K. Winkler, K. Ohlendieck, Int J Mol Med, 32 (2013) 544-556. http://dx.doi.org/10.3892/ijmm.2013.1429

[121] S. Carberry, M. Zweyer, D. Swandulla, K. Ohlendieck, J Biomed Biotechnol, 2012 (2012) 691641. http://dx.doi.org/10.1155/2012/691641

[122] S. Duguez, W. Duddy, H. Johnston, J. Laine, M.C. Le Bihan, K.J. Brown, A. Bigot, Y. Hathout, G. Butler-Browne, T. Partridge, Cell Mol Life Sci, 70 (2013) 2159-2174. http://dx.doi.org/10.1007/s00018-012-1248-2

[123] L. Guevel, J.R. Lavoie, C. Perez-Iratxeta, K. Rouger, L. Dubreil, M. Feron, S. Talon, M. Brand, L.A. Megeney, J Proteome Res, 10 (2011) 2465-2478. http://dx.doi.org/10.1021/pr2001385

[124] R. Ramadasan-Nair, N. Gayathri, S. Mishra, B. Sunitha, R.B. Mythri, A. Nalini, Y. Subbannayya, H.C. Harsha, U. Kolthur-Seetharam, M.M. Srinivas Bharath, J Biol Chem, 289 (2014) 485-509. http://dx.doi.org/10.1074/jbc.M113.493270

[125] A. Holland, P. Dowling, P. Meleady, M. Henry, M. Zweyer, R.R. Mundegar, D. Swandulla, K. Ohlendieck, Proteomics, (2015). http://dx.doi.org/10.1002/pmic.201400471

[126] A. Holland, M. Henry, P. Meleady, C.K. Winkler, M. Krautwald, H. Brinkmeier, K. Ohlendieck, Molecules, 20 (2015) 11317-11344. http://dx.doi.org/10.3390/molecules200611317

[127] I. Sela, I. Milman Krentsis, Z. Shlomai, M. Sadeh, R. Dabby, Z. Argov, H. Ben-Bassat, S. Mitrani-Rosenbaum, PLoS One, 6 (2011) e16334. http://dx.doi.org/10.1371/journal.pone.0016334

[128] S. Feldkirchner, J. Schessl, S. Muller, B. Schoser, F.G. Hanisch, Proteomics, 12 (2012) 3598-3609. http://dx.doi.org/10.1002/pmic.201100559

[129] R.A. Kley, A. Maerkens, Y. Leber, V. Theis, A. Schreiner, P.F. van der Ven, J. Uszkoreit, C. Stephan, S. Eulitz, N. Euler, J. Kirschner, K. Muller, H.E. Meyer, M. Tegenthoff, D.O. Furst, M. Vorgerd, T. Muller, K. Marcus, Mol Cell Proteomics, 12 (2013) 215-227. http://dx.doi.org/10.1074/mcp.M112.023176

[130] A. Maerkens, R.A. Kley, M. Olive, V. Theis, P.F. van der Ven, J. Reimann, H. Milting, A. Schreiner, J. Uszkoreit, M. Eisenacher, K. Barkovits, A.K. Guttsches, J. Tonillo, K. Kuhlmann, H.E. Meyer, R. Schroder, M. Tegenthoff, D.O. Furst, T. Muller, L.G. Goldfarb, M. Vorgerd, K. Marcus, J Proteomics, 90 (2013) 14-27. http://dx.doi.org/10.1016/j.jprot.2013.04.026

[131] A.M. Gomez, A. Vanheel, M. Losen, P.C. Molenaar, M.H. De Baets, J.P. Noben, N. Hellings, P. Martinez-Martinez, J Neuroimmunol, 261 (2013) 141-145. http://dx.doi.org/10.1016/j.jneuroim.2013.05.008

[132] L. Staunton, H. Jockusch, C. Wiegand, T. Albrecht, K. Ohlendieck, Mol Biosyst, 7 (2011) 2480-2489. http://dx.doi.org/10.1039/c1mb05043e

[133] L. Staunton, H. Jockusch, K. Ohlendieck, Biochem Biophys Res Commun, 406 (2011) 595-600. http://dx.doi.org/10.1016/j.bbrc.2011.02.099

[134] D. Capitanio, M. Vasso, A. Ratti, G. Grignaschi, M. Volta, M. Moriggi, C. Daleno, C. Bendotti, V. Silani, C. Gelfi, Antioxid Redox Signal, 17 (2012) 1333-1350. http://dx.doi.org/10.1089/ars.2012.4524

[135] C.A. Mutsaers, D.J. Lamont, G. Hunter, T.M. Wishart, T.H. Gillingwater, Genome Med, 5 (2013) 95. http://dx.doi.org/10.1186/gm498

[136] A. Holland, T. Schmitt-John, P. Dowling, P. Meleady, M. Henry, M. Clynes, K. Ohlendieck, Biosci Rep, 34 (2014). http://dx.doi.org/10.1042/BSR20140029

[137] K. Ohlendieck, A. Holland, Journal of Integrated OMICS, 4 (2014). http://dx.doi.org/10.5584/jiomics.v4i2.171

[138] V. Thongboonkerd, R. Kanlaya, S. Sinchaikul, P. Parichatikanond, S.T. Chen, P. Malasit, J Proteome Res, 5 (2006) 3326-3335. http://dx.doi.org/10.1021/pr060136h

[139] Ohlendieck, International Journal of Molecular Medicine, 25 (2010). http://dx.doi.org/10.3892/ijmm_00000364

[140] H. Hwang, B.P. Bowen, N. Lefort, C.R. Flynn, E.A. De Filippis, C. Roberts, C.C. Smoke, C. Meyer, K. Hojlund, Z. Yi, L.J. Mandarino, Diabetes, 59 (2010) 33-42. http://dx.doi.org/10.2337/db09-0214

[141] E. Mullen, E. O'Reilly, K. Ohlendieck, Mol Med Rep, 4 (2011) 229-236. http://dx.doi.org/10.3892/mmr.2011.437

[142] T.E. Thingholm, S. Bak, H. Beck-Nielsen, O.N. Jensen, M. Gaster, Mol Cell Proteomics, 10 (2011) M110 006650. http://dx.doi.org/10.1074/mcp.M110.006650

[143] K. Ohlendieck, E. Mullen, Journal of Integrated OMICS, 1 (2011). http://dx.doi.org/10.5584/jiomics.v1i1.42

[144] H. Yuan, Y. Niu, X. Liu, F. Yang, W. Niu, L. Fu, PLoS One, 8 (2013) e53887. http://dx.doi.org/10.1371/journal.pone.0053887

[145] X. Duan, F. Berthiaume, D. Yarmush, M.L. Yarmush, Biochem J, 397 (2006) 149-158. http://dx.doi.org/10.1042/BJ20051710

[146] X. Jia, K. Hollung, M. Therkildsen, K.I. Hildrum, E. Bendixen, Proteomics, 6 (2006) 936-944. http://dx.doi.org/10.1002/pmic.200500249

[147] T. Geiger, J.R. Wisniewski, J. Cox, S. Zanivan, M. Kruger, Y. Ishihama, M. Mann, Nat Protoc, 6 (2011) 147-157. http://dx.doi.org/10.1038/nprot.2010.192

[148] S. Holper, A. Ruhs, M. Kruger, Methods Mol Biol, 1188 (2014) 95-106. http://dx.doi.org/10.1007/978-1-4939-1142-4_8

[149] G. Dittmar, M. Selbach, Proteomics Clin Appl, 9 (2015) 301-306. http://dx.doi.org/10.1002/prca.201400112

[150] T. Can, L. Faas, D.A. Ashford, A. Dowle, J. Thomas, P. O'Toole, G. Blanco, Proteome Sci, 12 (2014) 25. http://dx.doi.org/10.1186/1477-5956-12-25

[151] T.P. Dunkley, R. Watson, J.L. Griffin, P. Dupree, K.S. Lilley, Mol Cell Proteomics, 3 (2004) 1128-1134. http://dx.doi.org/10.1074/mcp.T400009-MCP200

[152] S. Viswanathan, M. Unlu, J.S. Minden, Nat Protoc, 1 (2006) 1351-1358. http://dx.doi.org/10.1038/nprot.2006.234

[153] M. Von Bergen, F. Dautel, S. Kalkhof, S. Trump, I. Lehmann, A. Beyer, Journal of Integrated OMICS, 1 (2011). http://dx.doi.org/10.5584/jiomics.v1i1.50

[154] G. Arentz, F. Weiland, M.K. Oehler, P. Hoffmann, Proteomics Clin Appl, 9 (2015) 277-288. http://dx.doi.org/10.1002/prca.201400119

[155] J.R. Yates, C.I. Ruse, A. Nakorchevsky, Annu Rev Biomed Eng, 11 (2009) 49-79. http://dx.doi.org/10.1146/annurev-bioeng-061008-124934

[156] A.F. Altelaar, A.J. Heck, Curr Opin Chem Biol, 16 (2012) 206-213. http://dx.doi.org/10.1016/j.cbpa.2011.12.011

[157] T.E. Angel, U.K. Aryal, S.M. Hengel, E.S. Baker, R.T. Kelly, E.W. Robinson, R.D. Smith, Chem Soc Rev, 41 (2012) 3912-3928. http://dx.doi.org/10.1039/c2cs15331a

[158] Z. Zhang, S. Wu, D.L. Stenoien, L. Pasa-Tolic, Annu Rev Anal Chem (Palo Alto Calif), 7 (2014) 427-454. http://dx.doi.org/10.1146/annurev-anchem-071213-020216

[159] C. Lewis, K. Ohlendieck, Anal Biochem, 404 (2010) 197-203. http://dx.doi.org/10.1016/j.ab.2010.05.017

[160] L. Staunton, K. Ohlendieck, Protein & Peptide Letters, 19 (2012) 252-263. http://dx.doi.org/10.2174/092986612799363208

[161] K. Ohlendieck, Journal of Membrane and Separation Technology, 2 (2013) 1-12

[162] C.R. Jimenez, S. Piersma, T.V. Pham, Biomark Med, 1 (2007) 541-565. http://dx.doi.org/10.2217/17520363.1.4.541

[163] D. Calligaris, C. Villard, D. Lafitte, J Proteomics, 74 (2011) 920-934. http://dx.doi.org/10.1016/j.jprot.2011.03.030

[164] K. Ohlendieck, Biomark Med, 7 (2013) 169-186. http://dx.doi.org/10.2217/bmm.12.96

[165] M. Stastna, J.E. Van Eyk, Proteomics, 12 (2012) 722-735. http://dx.doi.org/10.1002/pmic.201100346

JOURNAL OF INTEGRATED OMICS
A METHODOLOGICAL JOURNAL
HTTP://WWW.JIOMICS.COM

ORIGINAL ARTICLE | DOI: 10.5584/jiomics.v4i2.172

Proteomic analysis of low quantities of cellular material in the range obtainable from scarce patient samples

Katja Parapatics[1,2], Marie L. Huber[1], Dorit Lehmann[3], Christian Knoll[1], Giulio Superti-Furga[1], Keiryn L. Bennett[1], Elena L. Rudashevskaya*[1]

[1]CeMM Research Center for Molecular Medicine of the Austrian Academy of Sciences, Vienna, Austria. kparapatics@cemm.oeaw.ac.at (K.P.), mhuber@cemm.oeaw.ac.at (M.L.H.), gsuperti@cemm.oeaw.ac.at (G.S.F.), kbennett@cemm.oeaw.ac.at (K.L.B.); [2]Department of Ophthalmology, Medical University of Vienna, Vienna, Austria. [3]Department of Vascular Biology and Thrombosis Research, Medical University of Vienna, Vienna, Austria.

Received: 16 June 2014 **Accepted:** 21 December 2014 **Available Online**: 28 January 2015

Abstract

The application of proteomics to patient material is increasingly widespread, however, a major shortcoming still are the number of cells or protein material that can be obtained. This study explores the lower limit of cell numbers that can be successfully analysed by liquid chromatography mass spectrometry to determine the protein expression profile that is specific to, and indicative of, the investigated cell type. The aim was to analyse an equivalent quantity of cellular material that can be obtained from, *e.g.*, a fine-needle aspiration biopsy (FNAB). Fifteen thousand and 30,000 cells from adherent (HEK293) and suspension (U937) cell lines were lysed under two different conditions: a 'native' and a denaturing buffer. To extend the study to clinical material, human whole PBMCs were also lysed under identical conditions. Proteins from 5,000 and 10,000 cells were analysed by both 1D and 2D-LC-MSMS on an LTQ Orbitrap XL mass spectrometer. In total, 3,219; 1,693 and 659 unique proteins were identified from HEK293, U937 and total PBMCs, respectively. Additionally, an iTRAQ 4-plex experiment was performed to determine the relative quantity of the proteins in the three cell types. In this study, we show that it is feasible to obtain a deep, yet cell-specific protein profile from a very low number of cultured and primary cells. This advancement will enable proteomic-profiling of cellular material from fine needle aspiration biopsies that ultimately can assist cytopathologists in the diagnosis of disease.

Keywords: mini-proteome; iTRAQ; FNAB; PBMC; HEK293; U937.

Abbreviations

1D-LC-MSMS: one-dimensional liquid chromatography tandem mass spectrometry; **1D-SDS-PAGE**: one-dimensional sodium dodecyl-sulphate polyacrylamide gel electrophoresis; **2D-LC-MSMS**: two-dimensional liquid chromatography tandem mass spectrometry; **CID**: collision-induced dissociation; **DMEM:** Dulbecco's modified eagle medium; **FASP**: filter-aided sample preparation; **FCS**: fetal calf serum; **FNAB**: fine-needle aspiration biopsy; **GO**: gene ontology; **HCD**: higher-energy collision-induced dissociation; **iTRAQ**: isobaric tag for relative and absolute quantitation; **LCM:** laser-capture microdissection; **LCMS**: liquid chromatography mass spectrometry; **LTQ**: linear trap quadrupole; **PBMC**: peripheral blood mononucleocytes; **PBS**: phosphate-buffered saline; **RPMI**: Roswell Park Memorial Institute medium; **SD**: standard deviation; **SDS**: sodium dodecylsulfate; **TEAB**: triethylammonium bicarbonate.

1. Introduction

Clinical samples from patients are extremely valuable and the amount of material and cells available is very low, often difficult to obtain and thus extremely precious. A fine-needle aspiration biopsy (FNAB) is a relatively painless and straightforward minor surgical diagnostic procedure used to extract cellular material from percutaneous masses and/or drain fluid -filled cysts. A very fine gauge needle is inserted and the cells and/or fluid are aspirated (suctioned) from the mass into the

***Corresponding author:** Institute of Pharmacology, Center for Physiology and Pharmacology, Medical University of Vienna, Waehringerstrasse 13A. 1090 Vienna, Austria. Tel: +43 1 40160 31366. E-Mail address: elena.rudashevskaya@meduniwien.ac.at

needle, stained, examined by a cytopathologist and a clinical diagnosis made. The number of cells obtained via FNAB is exceptionally low and proteomic analysis of the material in the aspirated sample is extremely challenging. Thus, only a limited number of studies exist on proteomic profiling of FNAB [1-5]. In a study by Rapkiewicz et al. [6], fine-needle aspiration samples from breast tumors were analyzed by quantitative protein microarray technology. These researchers calculated that the number of cells in frozen aspirate samples was in the range of 1,000–50,000 cells.

The analysis of 'core' proteomes by liquid chromatography mass spectrometry (LCMS) has been performed on large quantities (30-50 μg) of protein from cell lysates of immortalised, cultured cells [7-9]. Protein numbers ranged from approximately 1,800 [7] to between 2,000-4,000 [8] and up to 10,500 [9] non-redundant proteins from a single cell line. The historical increase in protein numbers reflects the combination of improved MS instrumentation technology and sample preparation methodology. With limited protein quantities, e.g., from clinical samples, older generation mass spectrometers were unable to delve deeply into the proteome of a low number of patient cells. The advent of the hybrid linear trap quadrupole (LTQ) and quadrupole only Orbitrap series of mass spectrometers has meant that it is now possible to analyse low numbers of cells from primary sources. More importantly, such analyses do not just result in the identification of the usual abundant, 'uninteresting' housekeeping proteins but actually lead to the identification of cell-specific proteins. Some key examples from the literature include: 3,800 proteins identified from an equivalent of 5,000 FACS-sorted colon stem cells [10]; single islets of Langerhans and ~10,000 laser-capture microdissected (LCM) mouse kidney glomeruli containing 2,000-4,000 and ~2,400 proteins, respectively [11]; 900-1,900 proteins from 250-10,000 FACS-sorted short-term culture melanoma cells [12, 13]; ~1,000 proteins from 3,000 LCM-derived breast carcinoma tumor cells [14]; and 3,600-4,400 proteins from 20,000 microdissected formalin-fixed, paraffin-embedded colon carcinoma cells [15].

The current study was designed to ascertain the depth that could be obtained from a low number of cells using our sample preparation approach and MS instrumentation. In particular, we were interested in establishing appropriate experimental conditions for a 'native' lysis buffer with the goal of analysing cellular proteomes from small quantities of material. The data generated would be used to assess the feasibility of a planned projection into routine proteomic analyses of FNAB. The 'native' lysis buffer, that was ultimately chosen, closely resembled buffers commonly used for investigating native protein complexes and protein interactions. Considering the medical impact and relevance for translational research, there is an increasing interest in investigating alterations in protein complex formation or pathway composition and also protein-drug interactions directly from clinical material [16]. In all cases, such experiments require cellular lysis conditions that maintain and preserve native protein conformations and interactions, i.e., non-denaturing. In parallel, denaturing conditions that utilise buffers containing urea can supplement the data by 'releasing' proteins from cellular compartments that are not completely accessible with a 'native', non-denaturing buffer alone. Furthermore, urea buffers are included in commercially-available kits for enrichment of post-translationally-modified (PTM) peptides (e.g., phosphotyrosine-enrichment). It has also been reported that urea-based buffers have a certain advantage over SDS-containing buffers in the analysis of cancer tumors [17].

The well-known and extensively-studied adherent HEK-293 cell line was chosen as a model, the methodology extended to a human macrophage suspension cell line (U937) and finally to readily-obtainable primary human total peripheral blood mononucleocytes (PBMCs) to mimic a clinical setting. Two lysis conditions were chosen and evaluated on the three different cell types: (i) a 'native' buffer consisting of 50 mM HEPES, 150 mM NaCl, 5 mM EDTA, 0.5% NP-40, 1mM PMSF, pH 8.0; and (ii) a denaturing buffer comprised of 9 M urea, 20 mM HEPES, 1mM PMSF, pH 8.0. Fifteen thousand and 30,000 cells were lysed and the protein equivalent of 5,000 or 10,000 cells analyzed by one-dimensional liquid chromatography tandem mass spectrometry (1D-LC-MSMS). The peptide digests from 15,000 and 30,000 lysed cells were also separated by reversed-phase chromatography, 20 fractions collected and analyzed by LC-MSMS. Finally, to add relative quantitative information to the proteins identified from the different cell types, iTRAQ labeling coupled to 2D-LC-MSMS was used to compare the three different cell types lysed under two different conditions. Functional annotation analysis was utilised to evaluate the obtained protein lists to determine the specificity of the individual cellular proteomes.

2.Material and Methods

2.1 *Reagents*

Iodoacetamide, dithiothreitol (DTT), HEPES (4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid), 1 M triethylammonium bicarbonate (TEAB), formic acid, urea, NaCl, EDTA (ethylenediaminetetraacetic acid), thiourea, DMEM, protease inhibitor cocktail, PMSF (Phenylmethanesulfonyl fluoride), ≥99.0% (SIGMA-Aldrich, St. Louis, MO); trypsin (Promega Corp., Madison, WI); RPMI, PBS, penicillin/streptomycin (PAA, Pasching, Austria), iTRAQ (ABI, Framingham, MA); FCS (Gibco, Grand Island, NY), NP-40 Alternative (CALBIOCHEM, San Diego, CA), Krypton™ Protein Stain (Thermo Scientific, Austria, Vienna).

2.2 *Collection of HEK293 and U937 cells*

HEK293 and U937 cells were grown as adherent (in DMEM medium) and suspension cultures (in RPMI medium) respectively. Each media was complemented with 10%

foetal calf serum (FCS) and 1% penicillin/streptomycin. Cells were harvested in posphate-buffered saline (PBS) solution and counted with a CASY® cell counting system (Roche Diagnostics, Rotkreuz, Switzerland). Multiple aliquots of 15,000 and 30,000 cells were collected in 1.5 mL eppendorf tubes and centrifuged at 1,800×g for 5 min at 4°C. The supernatant was removed and discarded; and the cell pellets stored at -80°C until required.

2.3 *Collection of peripheral blood mononucleocytes (PBMCs)*

Venous blood from a healthy donor was collected into BD Vacutainer tubes containing EDTA and diluted 1:1 in PBS. PBMCs were separated using LSM 1077 lymphocyte separation media (PAA Austria) by density gradient centrifugation at 2,100 r.p.m. for 30 min (Sorvall RT6000B) and washed with PBS by centrifugation for 10 min at 1,200 r.p.m. (Sorvall RT6000B). Contaminating erythrocytes were removed by incubation in a lysis buffer (containing 0.15 M NH_4Cl, 0.01 M $NaHCO_3$ and 0.1 mM EDTA, pH 7.2) for 10 min at 4°C and washing with PBS by 20 min centrifugation at 800 r.p.m. (Sorvall RT6000B), which also eliminated remaining platelets. PBMCs (4×10^7 cells) were counted with a Neubauer counting chamber (Karl Roth GmbH, Karlsruhe, Germany) and aliquots of 15,000 and 30,000 cells collected. Cells were centrifuged, the supernatant removed, and pellets frozen at -80°C and stored until required.

2.4 *Assessing lysis conditions for large quantities of cells*

HEK293 cells (8.7×10^6) were collected and lysed in 1 mL under the following six different conditions: (i) 50 mM HEPES, 150 mM NaCl, 5 mM EDTA, 0.5% NP-40, 1 mM PMSF, pH 8.0 (no sonication). Cells were incubated for 30 min at 4°C; and then centrifuged at 20,000 × g for 15 min at 4°C. (ii) 50 mM HEPES, 150 mM NaCl, 5 mM EDTA, 1 mM PMSF, pH 8.0 (+ sonication). Cells were sonicated at 70% output with 3 bursts of 30 s; and cooled on ice for 30 s between each burst. The resultant lysates were centrifuged at 20,000 × g for 15 min at 4°C. (iii) 20 mM HEPES, 9 M urea, pH 8.0 (no sonication). Cells were incubated for 30 min at room temperature, with vortexing every 10 min, and centrifuged at 20,000 × g for 15 min at RT. (iv) 20 mM HEPES, 9 M urea, pH 8.0 (+ sonication). Cells were incubated for 30 min at room temperature, with vortexing every 10 min, and with sonication at 70% output with 3 bursts of 30 s; and centrifuged at 20,000 × g for 15 min at RT. (v) 20 mM HEPES, 7 M urea, 2 M thiourea, pH 8.0 (+ sonication). Cells were incubated for 30 min at RT, with vortexing every 10 min, and with sonication at 70% output with 3 bursts of 30 s; and centrifuged at 20,000 × g for 15 min at RT.

2.5 *Protein concentration measurement*

The total protein content of HEK293 cell lysates generated from 8.7×10^6 cells was determined by the Bradford assay (using bovine serum albumin as a protein standard).

2.6 *'Native' buffer lysis and in situ tryptic digestion*

Fifteen thousand and 30,000 HEK293, U937 and human total PBMC cells were individually lysed in 100 µL 'native' buffer: 50 mM HEPES, 150 mM NaCl, 5 mM EDTA, 0.5% NP-40, 1 mM PMSF, pH 8.0 and incubated for 30 min at 4° C. Cell lysates were centrifuged at 20,000 × g for 15 min at 4° C and the protein extracts (supernatants) were collected. Laemmli buffer was added, and the samples were reduced and alkylated with dithiothreitol and iodoacetamide, respectively. 1D-SDS-PAGE of the lysates was performed on Novex Bis-Tris 4-12% gels at 200 V for 15 min only. Gels were stained with colloidal Coomassie blue. Stained gel regions containing the proteins were excised from the gel and digested in situ with trypsin at 37°C overnight. Tryptic digests were concentrated and purified by solid phase extraction (SPE) (UltraMicroSpin columns 3-30 µg capacity, Nest Group Inc., Southboro, MA, USA), the volumes reduced to approximately 2 µL in a vacuum centrifuge and reconstituted in 26 µL 5% formic acid for 1D-LC-MSMS triplicate analysis.

2.7 *Denaturing lysis and solution tryptic digestion*

Fifteen thousand and 30,000 HEK293, U937 and human total PBMC cells were individually lysed in 100 µL denaturing buffer: 20 mM HEPES, 9 M urea, pH 8.0 for 30 min at room temperature. Samples were vortexed every 10 min. Cell lysates were centrifuged at 20,000 × g for 15 min at RT and the protein extracts (supernatants) were collected. Samples were diluted with 100 mM TEAB to a final concentration of 1.4 M urea, reduced with dithiothreitol, alkylated with iodoacetamide and digested with 0.5 µg modified porcine trypsin at 37°C overnight. Tryptic digests were concentrated and purified by solid phase extraction (SPE) (UltraMicroSpin columns 3-30 µg capacity, Nest Group Inc., Southboro, MA, USA), the volumes reduced to approximately 2 µL in a vacuum centrifuge and reconstituted in 26 µL 5% formic acid for 1D-LC-MSMS triplicate analysis.

2.8 *iTRAQ derivatisation*

The tryptic digests from 30,000 cells lysed in either the 'native' or denaturing buffer were derivatised with the 4-plex iTRAQ reagent (ABI, Framingham, MA) [18] and labelled according to the instructions provided by the manufacturer. Two iTRAQ 4-plex experiments were prepared. One experiment was performed on cells lysed in the 'native' buffer. iTRAQ labels 114, 115, corresponded to 30,000 HEK293 cells; 30,000 U937 cells; and 116 and 117 corresponded each to 30,000 PBMC cells, respectively. In the other experiment, the same labeling was performed but on the tryptic digests from the denaturing buffer.

2.9 *Reversed-phase reversed-phase (RPRP) separation for 2D-LC-MSMS [19]*

Tryptic digests from the 'native' and denaturing individual cell lysates and also for the two iTRAQ experiments were concentrated and purified by SPE (UltraMicroSpin columns 3-30 µg capacity, Nest Group Inc., Southboro, MA, USA) and reconstituted in 23 µL 100 mM TEAB, pH 10, prior to injection onto a Phenomenex column (150 × 2.0 mm Gemini-NX 3 µm C18 110Å, Phenomenex, Torrance, CA, USA) on an Agilent 1200 series HPLC (Agilent Biotechnologies, Palo Alto, CA). Twenty and 40 fractions were collected for expression and quantitative proteomics, respectively. All fractions were acidified with 5 µL 5% formic acid and the volumes reduced to approximately 2 µL in a vacuum centrifuge. Samples were reconstituted to 26 µL with 5% formic acid and analysed as technical duplicates by LC-MSMS. Details of the methodology are as previously described [20].

2.10 *Liquid chromatography mass spectrometry*

Mass spectrometry was performed on a hybrid linear trap quadrupole (LTQ) Orbitrap XL mass spectrometer (Thermo Fisher Scientific, Waltham, MA) using the Xcalibur version 2.0.7 (expression proteomics). iTRAQ experiments were analysed on a hybrid LTQ Orbitrap Velos (ThermoFisher Scientific, Waltham, MA) using the Xcalibur version 2.1.0.1140 (relative iTRAQ quantitation). The Orbitrap mass spectrometers were coupled to an Agilent 1200 HPLC nanoflow system (dual pump system with one precolumn and one analytical column) (Agilent Biotechnologies, Palo Alto, CA) via a nanoelectrospray ion source using liquid junction (Proxeon, Odense, Denmark). Solvents for LCMS separation of the digested samples were as follows: solvent A consisted of 0.4% formic acid in water and solvent B consisted of 0.4% formic acid in 70% methanol and 20% isopropanol. From a thermostatted microautosampler, 8 µL of the tryptic peptide mixture were automatically loaded onto a trap column (Zorbax 300SB-C18 5 µm, 5×0.3 mm, Agilent Biotechnologies, Palo Alto, CA) with a binary pump at a flow rate of 45 µL/min. 0.1% TFA was used for loading and washing the pre-column. After washing, the peptides were eluted by back-flushing onto a 16 cm fused silica analytical column with an inner diameter of 50 µm packed with C18 reversed phase material (ReproSil-Pur 120 C18-AQ, 3 µm, Dr. Maisch GmbH, Ammerbuch-Entringen, Germany). The peptides were eluted from the analytical column with a 27 minute gradient ranging from 3 to 30% solvent B, followed by a 25 minute gradient from 30 to 70% solvent B and, finally, a 7 minute gradient from 70 to 100% solvent B at a constant flow rate of 100 nL/min [20]. The analyses were performed in a data-dependent acquisition mode using a top 6 collision-induced dissociation (CID) method for peptide identification alone (LTQ Orbitrap XL); or a top 10 high-energy collision-induced dissociation (HCD) method for peptide identification plus relative quantitation of iTRAQ reporter ions (LTQ Orbitrap Velos). Dynamic exclusion for selected ions was 60s. A single lock mass at m/z 445.120024 was employed on the LTQ OrbitrapVelos [21], but no lock mass was used on the LTQ Orbitrap XL. Maximal ion accumulation time allowed on the LTQ Orbitrap in CID mode was 150 ms for MS^n in the LTQ and 1,000 ms in the C-trap. Automatic gain control was used to prevent overfilling of the ion traps and were set to 5,000 (CID) in MSn mode for the LTQ, 10^6 ions for a full FTMS scan and 10^5 ions for HCD. Maximum ion time for HCD was set to 1,000 ms for acquiring 1 microscan at a resolution of 7,500. Intact peptides were detected in the Orbitrap at 100,000 resolution for CID fragmentation and 30,000 for HCD fragmentation experiments. The threshold for switching from MS to MSMS was 2,000 counts.

2.11 *Data analysis*

The acquired raw MS data files were converted into Mascot generic format (mgf) files with msconvert (ProteoWizard Library v2.1.2708). The resultant peak lists were searched against the human SwissProt database version v2010.09_20100812 (35,149 sequences, including isoforms as obtained from varsplic.pl) with the search engines Mascot (v2.3.02, MatrixScience, London, UK, www.matrixscience.com) and Phenyx (v2.6, GeneBio, Geneva, Switzerland) [22]. Submission to the search engines was via a Perl script that performs an initial search with relatively broad mass tolerances (Mascot only) on both the precursor and fragment ions (±10 ppm and ±0.6 Da, respectively). High-confidence peptide identifications are used to recalibrate all precursor and fragment ion masses prior to a second search with narrower mass tolerances (±4 ppm and ±0.3 Da for CID and ±4 ppm and ±0.025 Da for HCD, respectively). One missed tryptic cleavage site was allowed. Carbamidomethyl cysteine and N-terminal and lysine residue iTRAQ labelling were set as fixed modifications, and oxidised methionine was set as a variable modification. For the two-dimensional LC-MSMS samples, 10 individual analyses were merged into a single .mgf file prior submission to the search engines.

To validate the proteins, Mascot and Phenyx output files were processed by internally developed parsers. Proteins with ≥2 unique peptides above a score T1, or with a single peptide above a score T2, were selected as unambiguous identifications. Additional peptides for these validated proteins with score >T3 were also accepted. For Mascot and Phenyx, T1, T2 and T3 were equal to 12, 45, 10 and 5.5, 9.5, 3.5, respectively (p-value $<10^{-3}$). Following the selection criteria, proteins were grouped on the basis of shared peptides, and only the group reporters are considered in the final output of identified proteins. Spectral conflicts between Mascot and Phenyx peptide identifications were discarded. The whole procedure was repeated against a reversed database to assess the protein group false discovery rate (FDR). Peptide and protein group identifications were<0.1 and <1% FDR, respectively.

Comparisons between analytical methods involved comparisons between the corresponding sets of identified proteins. This was achieved by an internally-developed program that simultaneously computes the protein groups in all samples and extracts statistical data such as the number of distinct peptides, number of spectra, and sequence coverage.

2.12 *iTRAQ Quantitation*

The quantitation module of Proteome Discoverer 1.4, version 1.4.0.288 (Thermo Fisher Scientific, Waltham, MA) was used to assess the ratios for the individually-tagged tryptic digests of 30,000 HEK293, U937 and PBMCs. The intensity of the iTRAQ 4-plex reporter ions were integrated using the default settings for centroid peak detection at the highest confidence and a mass tolerance of 20 ppm. Correction for isotopic impurities was not performed. In addition, spectra with reporter ion intensities below 100 counts; and spectra with co-isolation of contaminating peptides exceeding 40% of the selected precursor ion were excluded from the protein ratio calculations. The median ratios for all peptides was calculated for each pair of cell lines (114/115, 114/116 114/117, 115/114, 115/116, 115/117, 116/114, 116/115, 117/114, 117/115), but shared peptides were excluded from quantitation. Protein ratios for the two combined technical replicates were calculated using the arithmetic mean of the protein ratios (median ratio of all used peptide ratios) for each replicate.

2.13 *Functional annotation analysis*

Functional annotation analysis of the identified proteins was performed using the DAVID Bioinformatics Resources 6.7 [23, 24]. Enrichment of GENETIC_ASSOCIATION_DB_DISEASE categories were performed with the thresholds: Count = 2, EASE = 0.05. Enrichment of KEGG_PATHWAY and UP_TISSUE categories were performed with the thresholds: Count = 2, EASE = 0.01. REVIGO was used to summarize the GO analysis [25].

3. Results and Discussion

3.1 *Selection of the optimal cell lysis conditions.*

Standard laboratory buffers for lysing cells usually include solubilising agents (e.g., NP-40 or other detergents) and/or denaturing agents (e.g., urea, thiourea or SDS) to facilitate protein extraction. Additionally, sonication is often advised to aid in clarification of the cell lysate. To determine the optimal conditions for the present study, a series of standard laboratory lysis conditions were evaluated on cultured, adherent HEK293 cells. The chosen conditions were: (i) 50 mM HEPES, 150 mM NaCl, 5 mM EDTA, 0.5% NP-40, pH 8.0; (ii) 50 mM HEPES, 150 mM NaCl, 5 mM EDTA, pH 8.0 (without NP-40) but with sonication; (iii) and (iv) 9 M urea; 20 mM HEPES, pH 8.0 ± sonication; and (v) 7 M urea, 2 M thiourea, 20 mM HEPES, pH 8.0 + sonication. In 1 mL of each of the six buffers described above, 8.7×10^6 cells were lysed. Following lysis, aliquots corresponding to 1,000 cells were removed to simulate low quantities of cellular material. Samples lysed in buffers containing NP-40 (i.e., 'native' buffers) were analysed by 1D-gel-LC-MSMS and samples lysed in urea (i.e., denaturing buffers) were analysed as technical duplicates by gel-free 1D-LC-MSMS.

The total number of proteins identified by 1D-LC-MSMS from the aliquots of cell lysate corresponding to 1,000 cells was evaluated (Supplementary Figure S1 A). The two buffers that gave the highest number of identified proteins were: (i) HEPES with NP-40; and (iii) urea buffer. For the proteins identified in these samples, the Gene Ontology classification for cellular component (GOTERM_CC) was performed (P-value <0.01, Supplementary Figure S1 B). As anticipated, the protein extracts generated with the buffer containing NP-40 were enriched for the GO terms: membrane-bound organelles, mitochondria, endoplasmic reticulum, and organelle lumen. Conversely, analysis of the extracts from the urea lysis buffer complemented the 'native' buffer data set by providing additional ribosomal and nuclear proteins. Thus, these two lysis conditions: (i) a 'native' buffer comprised of 50 mM HEPES, 150 mM NaCl, 5 mM EDTA with 0.5% NP-40); and (iii) a denaturing buffer containing 9 M urea (both without sonication) generated the highest number of identified proteins that together cover the broadest range of cellular components. Subsequently, both buffers were chosen for further evaluation.

3.2 *Determining the protein content from low cell numbers*

All the standard laboratory methods for determining protein concentration failed to provide accurate and reproducible values for the low numbers of cells analysed in this study. Nevertheless, we were curious to attempt to estimate the quantity of protein from a low number of cells. Thus, the quantity of protein in 10,000 HEK293, U937 and human primary PBMCs was determined as follows. Thirty thousand cells of each type were lysed in 'native' buffer; one third of each sample resolved by 1D-SDS-PAGE (simulation of 10,000 lysed cells), stained with fluorescent Krypton Protein Stain and visualised with ChemiDoc™ MP Imaging System (BIO-RAD) (Figure 1A). In addition, one third of the extract from 30,000 HEK293 cells was loaded onto a gel together with serial dilutions of the cell lysate from 8.7×10^6 HEK293. Both samples were lysed under identical conditions. For the larger sample, an accurate determination of the protein content was achieved using the Bradford assay. A dilution series was used to calibrate the protein content, and thus, the protein quantity in the lysate of 10,000 cells was estimated by fluorescence intensity of the gel lanes analysed with Image Lab 5.0 software (BIO-RAD) (Figure 1B). For HEK293, 10,000 cells equated to 780 ng total protein. Perhaps as a consequence of the smaller size (ranging from 7-20 μm in diameter) and mixed cell population of human primary

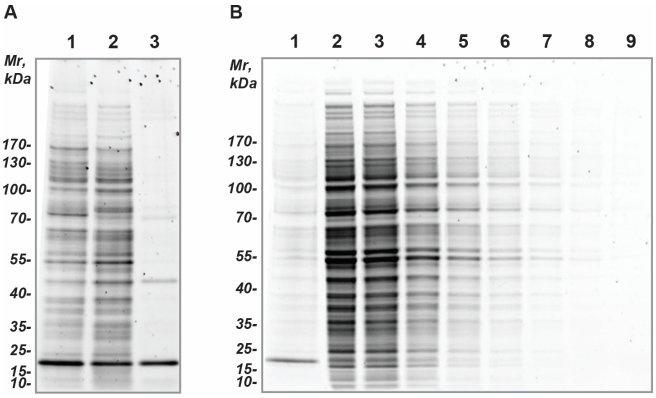

Figure 1. Calibration and determination of the protein content in 10,000 lysed cells. Samples were separated on a 4-12% Bis-Tris gel and proteins visualized by Krypton™ Protein Stain. (A) Protein extract from 30,000 lysed cells with one third of the extract loaded. Lane 1: 10,000 HEK293 cells; lane 2: 10,000 U937 cells; and lane 3: 10,000 primary human PBMCs. (B) Comparison of 10,000 HEK-293 cells from 30,000 cells extract with serial dilution of the protein extract from 8.7×10^6 lysed HEK293 cells. Protein content in the larger sample was determined by the Bradford assay. Lane 1: ⅓ protein extract from 30,000 cells (10,000 cells); Lane 2: 10 µg; Lane 3: 5 µg; Lane 4: 2.5 µg; Lane 5: 1.25 µg; Lane 6: 0.625 µg; Lane 7: 0.313 µg; Lane 8: 0.156 µg; Lane 9:0.078 µg.

PBMCs (consisting of neutrophils, eosinophils, basophils, lymphocytes, monocytes); the quantity of protein for 10,000 PBMCs is approximately 4 times less than for both HEK293 and U937 cells, as estimated by comparison of fluorescense intensity of the lanes on the gel (Figure 1A). All subsequent experiments were based on the number of cells rather than on the quantity of protein in the lysate.

3.3 *Determining the proteome of low cell quantities*

To emulate the situation of low sample availability, e.g., human FNAB and/or core needle biopsies, aliquots of 3,000, 15,000 and 30,000 HEK293 cells were lysed under (i) 'native' and (iii) denaturing conditions. Each sample was analysed as technical triplicates by 1D-LC-MSMS (each replicate thus corresponds to the analysis of 1,000, 5,000 and 10,000 cells). Interestingly, the result obtained from 5,000 cells lysed in the urea denaturing buffer (191 proteins) (Supplementary Figure S2) was comparable to a 1,000 cell aliquot from a higher number of cells lysed in the same buffer (148 proteins) (Supplementary Figure S1 A). The result from the 1,000 cells lysed under these conditions was unsatisfactory due to the extremely low number of identified proteins (Supplementary Figure S2). This is a critical point when reviewing data from the literature and designing proteomic experiments involving low cell numbers. It is important to distinguish between the types of approaches used in terms of the quantity of cellular material analysed. The total lysate of a low quantity of cells results in less protein identifications, than the same quantity of cells taken as an aliquot from a significantly larger number of lysed cells. This observation is simply explained by the fact that larger relative losses from small starting quantities are incurred during the sample preparation. Based on these findings, further experiments were performed on aliquots of 15,000 and 30,000 cells only.

Fifteen thousand and 30,000 HEK293 cells, U937 cells and primary human PBMCs were lysed in the two chosen buffers: (i) 'native' and (iii) denaturing. One third of each sample corresponding to 5,000 and 10,000 cells, respectively, was analysed by 1D-gel-based ('native' lysis) and 1D-gel-free (denaturing lysis) LC-MSMS. A schematic of the experimental design is shown in Figure 2. In parallel, a second set of samples were analysed by 2D-LC-MSMS. The proteins identified in all conditions from HEK293, U937 and PBMC cells are presented in Supplementary Tables S1, S2 and S3. In almost all instances, fractionation of the samples prior to LC -MSMS analysis gave a higher number of identified proteins compared to the 1D approach (Figure 3A). Only the PBMCs lysed in the urea denaturing buffer gave a poorer result (Figure 3A, left panel). Obviously, with a very low quantity of starting material (estimated below 100 ng protein from Fig.1) the additional fractionation step only resulted in fur-

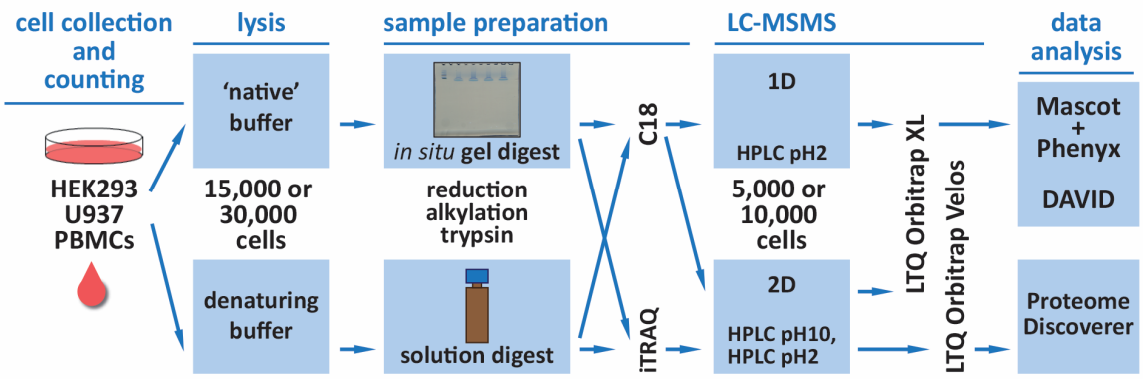

Figure 2. Schematic of the experimental design.

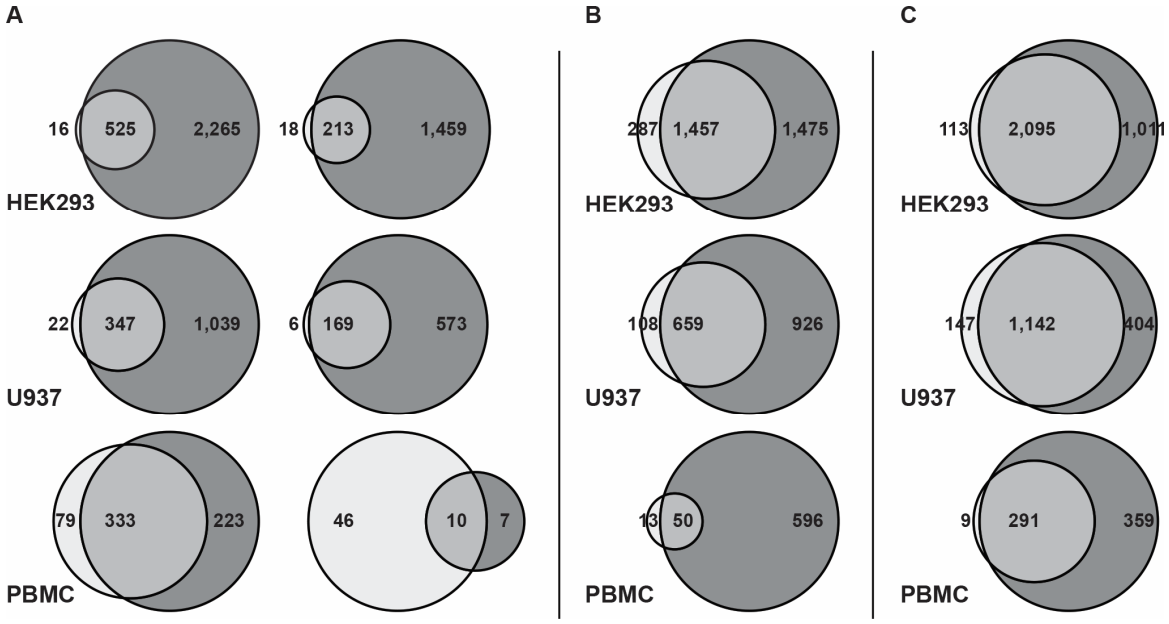

Figure 3. Comparison of the number of unique proteins identified from lysates of the three different cell lines: HEK293, U937, and PBMCs, according to additional sample fractionation before LC-MSMS analysis, the two lyses conditions and the amount of starting cellular material. (A) Comparison of 1D- and 2D-LC-MSMS of 10,000 cells (proteins extracted from 30,000 lysed cells). Left panel: protein extractions obtained with the 'native' NP-40 buffer. Right panel: protein extractions obtained with the denaturing urea buffer. Light grey circle, 1D-LC-MSMS; dark grey circle, 2D-LC-MSMS. (B) Comparison of the denaturing urea buffer (light grey circle) versus the 'native' buffer (dark grey circle). For each cell line, the number of unique protein identifications was generated from the analysis of 5,000 and 10,000 cells (protein extract obtained from 15,000 and 30,000 lysed cells, respectively) combined with the analyses by 1D- and 2D-LC-MSMS. (C) Comparison of 5,000 (light grey circle) and 10,000 cells (dark grey circle) (protein extract obtained from 15,000 and 30,000 lysed cells, respectively). For each cell line, the number of unique protein identifications was generated from the two lyses conditions ('native' and denaturing) combined with the analyses by 1D- and 2D-LC-MSMS.

ther dilution of the sample and/or sample loss leading to the observed decrease in the number of identified proteins compared to the 1D-LC-MSMS analysis.

As is often the case for proteomic experiments, using two conditions leads to an additive effect. The results of the different lysis buffers with respect to the number of proteins identified by 1D- and 2D- analyses of 5,000 and 10,000 cells are shown in Fig. 3B. As can be seen from this figure, the 'native' extraction conditions were more successful than the urea denaturing conditions; as evidenced by the higher number of proteins identified from the cells. Nevertheless, the results of the two lysis conditions are complementary with respect to the number of unique proteins identified.

When the quantity of material used for each experiment was compared, naturally 10,000 cells (initially 30,000 cells lysed) resulted in a higher number of identified proteins in all three cell types (Figure 3C). Approximately two thirds of the total number of proteins identified in HEK293 cells was common to the 5,000 and 10,000 cells; and three quarters of the proteins from the U937 cells were also shared. This appeared to occur even when applying a 2D approach that utilised additional sample fractionation. In contrast, only approximately one third of all the identified proteins from the PBMCs were evident for the results of different cell quantities. It is possible to surmise that for this cell type, 5,000 cells challenges the lower limits of successfully extracted protein material and sample fractionation and does not lead to an improvement in the number of proteins identified.

In total, 3,219; 1,693 and 659 unique proteins were identified from HEK293, U937 and total PBMCs, respectively. From 5,000 cells, 2,208; 1,289; and 300 proteins were identified and from 10,000 cells, 3,107; 1,546; and 650 proteins were identified. The most ideal condition for sample processing and subsequent analysis was to lyse 30,000 cells with the 'native' buffer followed by 2D-LC-MSMS. From this choice of experimental design alone, 2,790; 1,387; and 556 proteins were identified for the HEK293, U937 and PBMCs, respectively.

3.4 *'Mini-proteomes' complement known cell line proteomes and provide additional proteins*

Some of the earlier studies on HEK293 and U937cultured cell lines were performed utilising comparative studies of the 2D-gel approach combined with LC-MSMS [26-28] or MALDI-TOF analysis [26-28]. The earliest profile of HEK-293 proteins analysed by 1D-SDS-PAGE combined with LC-MSMS reported 1,111 and 1,063 proteins from the cytoplasmic and nuclear fractions, respectively [7]. A more recent publication identified 8,543 proteins from HEK293 cells. To date, this is the largest reported proteome of the eleven cell lines that were compared [9]. Another study that aimed at

describing the 'core' or 'central' proteome of different cultured cell lines, reported 4,154 and 2,073 protein groups for HEK293 and U937 cells, respectively [8]. The overall, central 'core' proteome of the five cell lines analysed in this work [8] consisted of 1,124 proteins. To the best of our knowledge, this is the only publication where HEK293 and U937 cell lines were proteomically-compared. As the cell lysis conditions were similar to those used in the current study, our HEK293 and U937 proteomes were evaluated against this work [8]. The 'mini-proteome' of both cell lines did not cover the 'core' proteome in entirety (Fig. 4 A, C). At the same time, however, the 'mini-proteome' generated from 5,000 and 10,000 HEK293 cells lysed in the 'native' buffer and analysed by 2D-LC-MSMS revealed that 54% and 63%, respectively, of the proteins identified were not part of the 'core' proteome. Under the same conditions, 42% and 45% of the proteins identified from 5,000 and 10,000 U937 cells, respectively, also did not overlap with 'core' proteome. This observation suggested that these non-'core' proteins represent a specific component of the proteome profile for each of the two cell lines. Interestingly, even from the proteome of 5,000 cells many proteins were identified that were not apparent from the proteomes generated from a larger amount (6 µg) of material (Fig. 4 B, D). Not unexpectedly, the total number of proteins identified in the individual HEK293 and U937 'mini-proteomes' decreased proportionally when compared to the larger proteomes. Namely, a decrease of 1.7 and 1.8 times for 10,000 HEK293 and U937 cells, respectively; and a decrease of 2.3 and 2.0 times for 5,000 cells (Fig 4 E). Comparison of each 'mini-proteome' with the 'core' proteome revealed that the U937 cells have a lower total number of protein identifications overlapping with the 'core' proteome than the HEK293 cells. For the HEK293 cells, however, the portion of the 'core' proteome proteins increased from 24% in the large proteome [8] to 46% in the 5,000 cell 'mini-proteome'. For the U937 cells, 46% of the large proteome overlaps with the 'core' proteome. This overlap increased to 58% when a lower amount of material was analysed (Fig 4 F). Thus, overall the smaller U937 proteome retained a larger portion of the 'core' proteome when the amount of analysed material was decreased.

It is important to note that although decreasing the amount of analysed material leads to an expected decrease in the total number of proteins identified, the obtained 'mini-proteome' still included many proteins from the 'core' proteome but also still remained representative of the cells under investigation. The relative abundance of the proteins expressed as emPAI values (Supplementary Table S4) showed that the most abundant proteins in the samples are from the 'core' proteome and many of the non-'core' proteins have a low relative abundance. At the same time, the relative abundance values have a large degree of variation for the 'core' proteins and also for the cell-specific proteins. As such, no direct connection between the abundance of a protein and the specificity of that particular protein for certain cell type can be made. As many of the cell-specific proteins have an abundance similar to the 'core' proteins, the former thus have an equal opportunity to be identified from low amounts of cellular material and therefore contribute to the cell-specific proteome.

There are a few publications describing the number of proteins identified in proteomic studies of PBMCs by 1D-LC-MSMS. These include: 1,432 [29]; and 514 proteins from 50 µg protein [30]. In our work, a total of 652 proteins were identified. Nevertheless, it is important to note here that this number of protein identifications was generated from a low amount of material. From just one of the analysis conditions (i.e., 10,000 cells, 'native' buffer, 2D-LC-MSMS) 556 proteins were identified. Interestingly, 17% of the PBMC proteome (115 protein identifications) from this current study did not correspond with the results from the two previously mentioned publications [29, 30] (Fig. 4 G).

As PBMCs are primary cells and by extrapolation, primary cells of any origin are of relevance and interest in clinical studies, the total proteome of these cells will be described and characterized in more detail. In particular, we will focus on the newly-identified proteins.

The mentioned above, 115 proteins identified from PBMCs were analysed using DAVID and the results revealed enrichment of the term IMMUNE as GENETIC_ASSOCIATION_DB_DISEASE_CLASS (Count = 13, p-value 0.01); and hsa05322:Systemic lupus erythematosus as KEGG_PATHWAY (Count = 5, p-value 0.05) (Supplementary Table S5). These categories indicate a possible relevance of the proteins involved in immune responses for clinical studies. Gene Ontology analysis also revealed features that are characteristic of the function of blood cells. The most highly-enriched GO biological process categories in the DAVID analyses with subsequent enrichment in REVIGO were: immune system process, immune response, cellular component assembly, response to stimulus, and multi-organism process (Fig 4H) (Supplementary Table S5). Response to biotic stimulus, killing cells of other organisms, defense response to bacterium were also revealed.

3.5 Identifying salient features of specific cell proteomes

Summarised in Fig. 5A are the combined, complementary results for the 30,000 cells lysed under both the 'native' and denaturing conditions and analyzed by 2D-LC-MSMS. The largest and smallest proteome coverage was obtained for the HEK293 and PBMC cells, respectively. The overlap between all three cell types contained the majority of the PBMC proteome, with around one third of the U937 proteome and almost one seventh of the HEK293 proteome. A similar outcome was obtained for the three different proteomes when the results from all the experimental conditions applied to each cell line were combined (Supplementary Figure S3). At the same time, both HEK293 and U937 have a specific and separate overlap with the proteome of the PBMCs. This was confirmed by functional annotation of the obtained proteomes for the 'pathways' (Fig. 5B; Supplementary Table S6)

and 'disease' categories (Fig. 5C; Supplementary Table S7).

For the HEK293 cell proteome, among the specific categories of diseases that were enriched are bladder cancer, head and neck cancer, and leukemia. This is in accordance with the nature of the HEK cells being originally endothelial or epithelial cells immortalized from embryonic kidney tissue

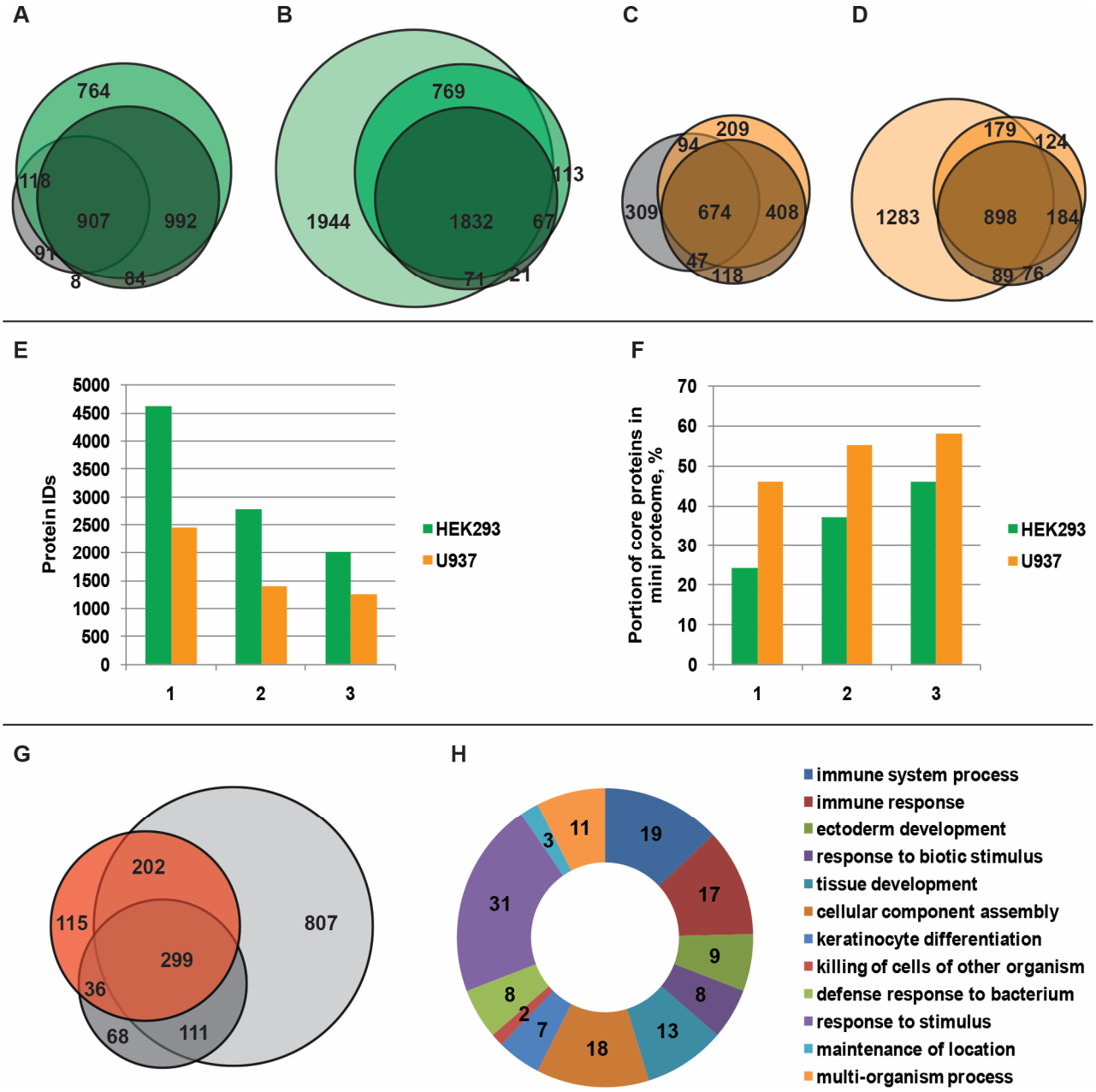

Figure 4. Comparison of the 'mini-proteomes' of HEK293, U937 and PBMC cells with the proteomes generated from larger amounts of protein. (A) (B) (C) (D) Comparison of the number of protein identifications from this current study with Burkard et al.[8] 'Core' proteome from 5 cell lines[8] (grey); large proteome of HEK293 cells (light green) and U937 cells (light orange) (50 µg protein lysate, 6 µg analysed;[8]) 'mini-proteomes' of 10,000 HEK293 (medium green) and 10,000 U937 cells (medium orange) (30,000 cells lysed in the 'native' buffer and material from 10,000 cells analysed by 2D-LC-MSMS); 'mini-proteomes' of 5,000 HEK293 (dark green) and 5,000 U937 cells (dark orange) (15,000 cells lysed in the 'native' buffer and material from 5,000 cells analysed by 2D-LC-MSMS). (A) (C) HEK293 proteomes and the 'core' proteome from;[8] (B) (D) U937 proteomes and the 'core' proteome from.[8] (E) Comparison of the number of proteins identified from (1) large proteome; (2) 'mini-proteome' from 10,000 cells; and (3) 'mini-proteome' from 5,000 HEK293 (green) and U937 cells (orange). (F) Comparison of the portion of the 'core' proteome observed in HEK293 (green) and U937 (orange) proteome obtained from the same amount of analysed material as indicated in (E). (G) Number of proteins identified in the proteome of PBMCs from three independent studies. Current study (red) representing the total number of protein identified from all combined experimental conditions; Haudek et al. (light grey);[29] Maccarrone et al. (dark grey).[30] (H) Enriched biological processes for the part of the PBMC 'mini-proteome' identified in our work that does not overlap with published data. GO analysis from DAVID and REVIGO.

[31]. Enrichment of neural tube defects and ALS/amyotrophic lateral sclerosis, both neurodegenerative diseases [32], is in accordance with the opinion that HEK cells originate from the transformation of neuronal cells of kidney tissue [33]. In the ALS/amyotrophic lateral sclerosis disease category, SOD1 (SODC_HUMAN) was found and this is known to be associated with ALS [34].

U937 cells originate from blood cells [35], so it is not surprising that these cells share common features with PBMCs. Both cell types were enriched for the disease category 'catalase activity'. Superoxide dismutase, catalase and glutathione peroxidase (SODM_HUMAN, CATA_HUAMN and GPX1_HUMAN) are important in the protection against oxidative stress. Several HLA (human leukocyte antigene) histocompatibility antigens were found in the PBMCs and U937 cells (1A68, 1A03, 1B53, 1B51, 1B58, 1B73, 1C12_HUMAN) that led to the enrichment of three other disease categories: hypothyroidism, spondyloarthropathies, and vitiligo. Although the PBMC proteome is the smallest of the three, many more disease categories were enriched from the data set (Fig. 5B, Supplementary Table S7).

In the pathway analysis, the lower number of pathway categories enriched for U937 and PBMCs in comparison with HEK293 cells may be a reflection of the size of the identified proteome (Fig 5C). Nevertheless, U937 cells were specifically-enriched for lysosome and Fc gamma R-mediated phagocytosis pathways (Supplementary Table S6). This is in accordance with the origin of the cell line and coincides with the monocytic (in particular macrophage) cell characteristics [36]. Interestingly, the tight junction pathway was enriched in HEK293 which may reflect the adherent growth pattern of these cells in culture compared to the U937 cell line.

The main energy releasing metabolic pathways (e.g., glycolysis, pentose phosphate pathway, pyruvate metabolism and oxidative phosphorylation) were enriched in PBMCs. Excluding glutathione metabolism, however, nucleotide and amino acid metabolism pathways were not enriched. This coincides with the important role of glutathione metabolism in reductive processes to counteract oxidative stress. The decrease in many metabolic processes and also DNA and translation-related processes is complementary to the status of PBMCs as highly-differentiated cells. Despite the lower number of proteins identified from PBMCs, there are several pathways specifically enriched in this cell type. These included cell motility, cell communication and immune system pathways. From the PBMC 'min-proteome' a range of proteins were also identified that are specific for certain cell types [37] (Supplementary Figure S5). These included monocytes, T cells, platelets and neutrophils.

It is important to highlight that from such a low number of cells combined with the described cell lyses conditions and mass spectrometry approaches; it was possible to identify specific characteristics inherent to each cell type. This was particularly evident for the human PBMCs, which are primary, differentiated cells, in contrast to the two immortalised cultured cells.

3.6 *Quantitative differences between specific cell proteomes*

For the quantitative analysis of the cell proteomes, 30,000 cells were lysed with denaturing or 'native' buffers as described above and shown in Fig. 2. After tryptic digestion, peptides from HEK293 and U937 were labelled with the 114 and 115 iTRAQ reagents, respectively. In addition, two independent PBMC preparations were labeled with the 116 and 117 iTRAQ reagents (Fig. 6A). Samples were prepared and injected in amounts equivalent to those in the experiment without labeling. At the same time, the quantitative MSMS analysis of the labeled samples was performed by implementation of HCD ion fragmentation in contrast to the CID fragmentation for the non-quantitative approach.

Taking into account the results from the two lysis conditions, the total number of proteins identified in these experiments was 1,647 (Supplementary Table S8). In contrast, in the unlabelled experiments analysed by collision-induced dissociation (CID), a total of 3,374 proteins were identified

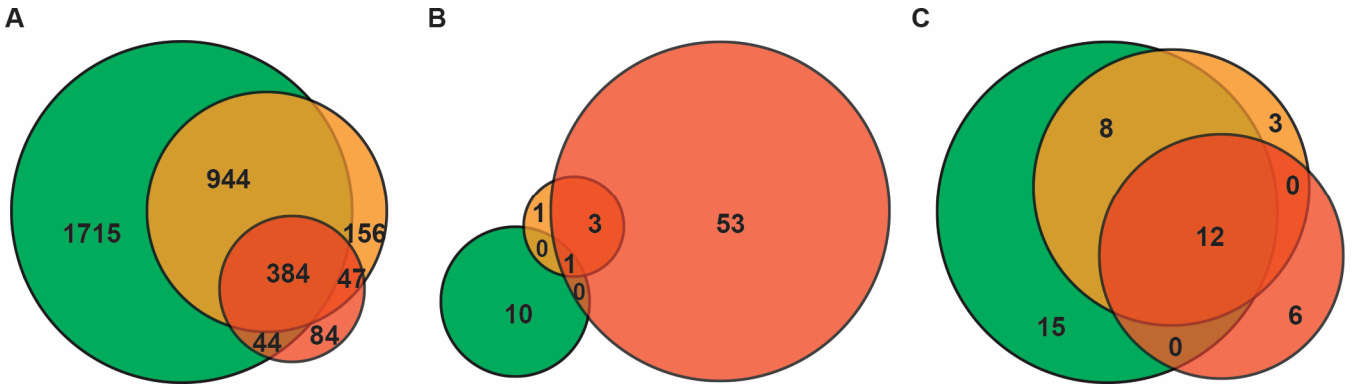

Figure 5. Comparison of the three cell types: HEK293 (green); U937 (orange); and human primary PBMCs (red). Data were generated on the proteins identified from 10,000 cells lysed in the two buffers and analysed by 2D-LCMSMS. (A) Number of unique proteins identified in the cell lines. (B) Number of disease categories. (C) Number of pathways enriched via functional annotation analysis.

from 10,000 cells lysed under the two conditions and analysed by 2D-LC-MSMS. The observed decrease in the total number of identified proteins between the two approaches can be explained by: (i) additional losses incurred during the extra sample preparation steps; and (ii) the slower duty cycle of high-energy collision-induced dissociation (HCD) compared to the duty cycle of CID. It has already been shown that the number of identified proteins in plasma decreased around 1.5 times in TMT-labeled samples compared to the unlabelled counterpart [38]. Care should obviously be taken when a quantitative, chemical labelling approach is performed on small amounts of material.

Shown in Fig. 6B and 6C are the distributions of the iTRAQ relative ratio quantitation between the three cell lines

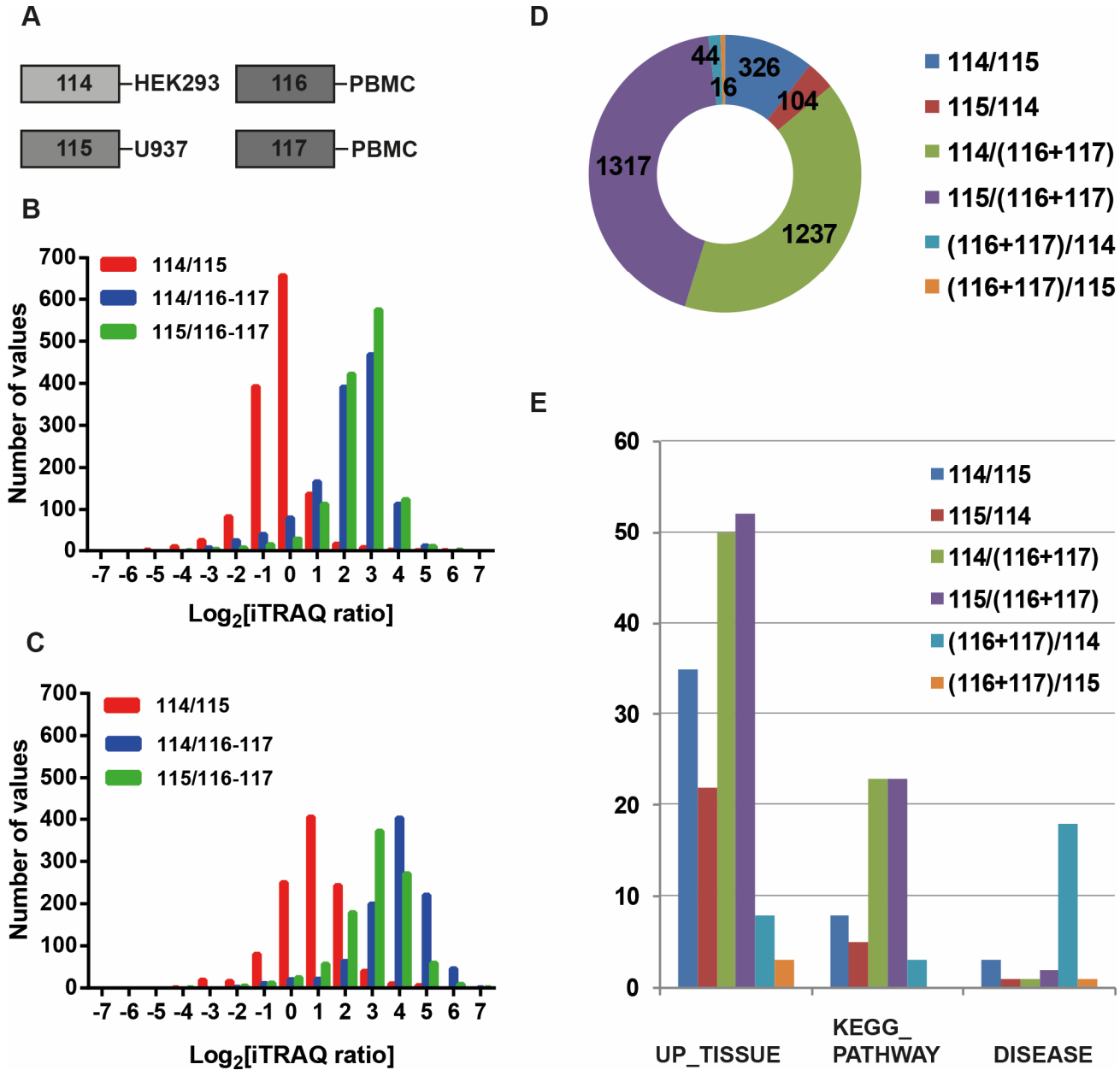

Figure 6. Quantitative analysis of the 'mini-proteomes' of the three cell lines: HEK293, U937 and PBMCs. (A) Scheme of the iTRAQ labelling. (B) Distribution of iTRAQ ratios in the cells lysed with the NP-40 'native' buffer. (C) Distribution of iTRAQ ratios in the cells lysed with the urea denaturing buffer. (D) Number of proteins identified that were more abundant in each cell line in a pair-wise analysis. Mean of the ratios and SD were calculated using GraphPad Prism 6. Proteins are discussed as more abundant if the ratios are 2-fold different and ratios are calculated in more than one replica. (E) Functional annotation analysis of the proteins from (B). Comparison of UP_TISSUE, KEGG_PATHWAY and GENETIC_ASSOCIATION_DB_DISEASE enrichment using DAVID Bioinformatics Resources 6.7. For UP_TISSUE and KEGG_PATHWAY, Count = 2, EASE = 0.01; for disease categories, Count = 2, EASE = 0.05; for the KEGG_PATHWAY of the PBMCs, Count = 2, EASE = 0.05.

for the 'native' buffer and the denaturing urea buffer, respectively. Data is expressed as a $\log_2$[iTRAQ ratio]-fold difference in protein abundance between the cell types. For both lyses conditions, the differences between the cultured and blood-derived cells are much greater than between the two cultured cell lines. From the 114/115 ratio (HEK293/U937), there are proteins more abundant in HEK293 cells and vice versa. Although the majority of the proteins obviously have a higher abundance in HEK293 and U937 cells compared to the PBMCs, there are still proteins that are more abundant in the PBMCs than either the HEK293 or the U937 cells. Some of these proteins exhibit up to a 3-fold difference in expression levels. This observation was particularly evident for the 'native' buffer (Fig. 6B).

A pair-wise comparison between the three cell proteomes was performed to determine the total number of identified proteins that are more abundant in each cell type (Fig. 6D). Again, it can be seen that the greatest difference between the proteomes is observed between the cultured cell lines and the primary PBMC cells. Nevertheless, 44 and 16 proteins from the PBMCs had a higher abundance compared to the HEK293 and U937 cells, respectively. The smaller difference between PBMCs and U937 might potentially reflect the closer origin between U937 cells and PBMCs. Numerous proteins showed differing degrees of abundance between the two types of cultured cells (Supplementary table S9).

Figure 6E further shows the observed differences between the cell lines and also that it is feasible to distinguish the specificity of each cell type. There are particular pathways enriched for each cell line. Due to the points noted earlier in this section that led to a decrease number of proteins identified, the total number of enriched categories is also lower than from the unlabelled experiments. Additionally, for each cell type, only the proteins with a significantly-higher expression (Supplementary table S9) were used for the enrichment analysis (Supplementary table S10). The quantitative data confirmed an enrichment of metabolic pathways (Supplementary table S11) in the cultured cells compared to the PBMCs. With a lower confidence (EASY>0.1), only a few categories were enriched in the blood-derived cells. These were ECM receptor interaction, regulation of actin cytoskeleton, and hematopoietic cell lineage. Although the number of proteins that were shown to be different in the PBMCs are relatively low, these are disease-related; and overall, there are more disease categories for the PBMCs in comparison to the other cells assessed in this study (Supplementary table S12). Among the proteins with an increased ratio in PBMCs are trombospondin-1 (P07996), gelsolin (P06396), several antigens CD41 (P08514), CD61 (P05106), CD11b (P11215), fibrinogens (P06271, P06279), arachidonate 5-lipoxygenase-activating protein (P20292). Thus, the approach confirms the applicability of this type of study in analysing and determining salient features of proteomes from samples of low cell number.

4. Concluding Remarks

Comparative proteomic analysis of three different cell types from low quantities of cellular material revealed that the data obtained contained information that was specific to, and representative of, the individual proteomes. The proteomes obtained from a limited amount of cellular material ('mini-proteomes') were smaller with respect to the total number of proteins identified than those obtained under similar conditions from larger quantities of cellular material. Although these 'mini-proteomes' contained a portion of the 'core' proteome shared between many different cell lines, the data also revealed many cell type specific proteins.

The specificity of the smallest possible proteome from a cell type and the minimal (critical) amount of material required to enable identification is highly-dependent on the source of the cells. A portion of specific cell line proteins was reduced in the smaller proteome, e.g., U937 in comparison with HEK293. From the same number of cells as the cultured cell lines, the highly-differentiated PBMCs provided less protein material which was reflected in the lower total number of identified proteins. Despite this, however, the PBMCs proteome contained a larger percentage of specific proteins than either the HEK293 or the U937 cells. Thus, the limitation caused by the lower amount of material was compensated in these differentiated cells by a higher relative level of cell-specific proteins. This finding provides enormous potential and relevance to clinical studies from low amounts of cellular material, e.g., the proteomic analysis of scarce patient samples such as a fine-needle aspiration biopsy (FNAB). Although out of the main focus of this study, our data from a small amount of cells supplements the characterisation of the PBMC proteome previously analysed by 1D-LC-MSMS. We identified and characterised an additional 115 proteins for this cell type.

Quantitation by chemical labelling and mass spectrometry is somewhat compromised when using small amounts of cells and protein material. At the same time, even for PBMCs it was possible to detect specific proteins when comparing the cell types to each other. It can be envisaged that with next generation approaches that combine improvements in protein solubilisation, sample preparation techniques and MS instrumentation (e.g., Q-Exactive) there will be an even deeper proteome of between 5,000-10,000 cells. Potentially, the analysis of material from even fewer cells will be facilitated and even deeper proteomic analyses, e.g., improved quantitative or posttranslational modifications will be feasible.

5. Supplementary material

Supplementary data and information is available at: http://www.jiomics.com/index.php/jio/rt/suppFiles/172/0

Supplementary-material.pdf file contains the table of content of all supplementary files, Supplementary Figures S1-S4, and Supple-

mentary Tables S6,S7 and S10-S12. Supplementary Tables S1-S5, S8 and S9 can be found in the separate Excel files with corresponding names.

Supplementary Figure S1 - Evaluation of buffers and lysis conditions for analysis of proteomes from small amounts of cellular material.
Supplementary Figure S2 - Number of unique protein identifications from different amounts of HEK293 cells lysed with the urea denaturing buffer.
Supplementary Figure S3 - Comparison of number of unique proteins identified in three cell types.
Supplementary Figure S4 - Proteins in PBMCs that are specific for certain blood cell types.
Supplementary Table S1 - List of proteins identified from HEK293 cells.
Supplementary Table S2 - List of proteins identified from U937 cells.
Supplementary Table S3 - List of proteins identified from PBMCs.
Supplementary Table S4 - emPAI values extracted from Mascot for the 30,000 cells analysed by 2D-LC-MSMS.
Supplementary Table S5 - List of proteins identified in our work from PBMCs (summary of all experimental approaches). These do not overlap with data previously published by Haudek et al., 2008 and Maccarrone et al., 2013; and Gene Ontology analysis of these proteins. Excel file.
Supplementary Table S6 - Comparison of functional annotation results for the three different cell types: HEK293, U937 and PBMCs. Enrichment of KEGG_PATHWAY categories.
Supplementary Table S7 - Comparison of functional annotation results for 3 different cell types: HEK293, U937 and PBMCs. Enrichment of GENETIC_ASSOCIATION_DB_DISEASE categories.
Supplementary Table S8 - Results from the iTRAQ quantitative experiments for the three cell types: HEK293, U937 and PBMCs. The table was generated from Proteome Discoverer. iTRAQ channels: 114, HEK293; 115, U937; 116 and 117, PBMCs. A and B, cells were lysed in denaturing buffer; C and D, cells were lysed in 'native' buffer. Excel file.
Supplementary Table S9 - Lists of proteins with a significantly higher level in each cell type. Sheet 'summary' contains calculation of fold differences in the level of all proteins identified in the iTRAQ experiment for the three cell types. Other sheets in the file contain lists of protein with significantly higher protein content in each pair of cell types.
Supplementary Table S10 - List of proteins with a significantly higher level in each cell type in comparison with the others in a pairwise comparison.
Supplementary Table S11 - Comparison of functional annotation results for quantitative analysis of the three different cell types: HEK293, U937 and PBMCs. Enrichment of KEGG_PATHWAY categories.
Supplementary Table S12 - Comparison of functional annotation results for quantitative analysis of HEK293, U937 and PBMCs. Enrichment of GENETIC_ASSOCIATION_ DB_DISEASE categories.

Acknowledgements

The authors would like to thank Elisabeth Salzer from the Boztug laboratory at CeMM for blood collection, and Melanie Planyavsky and André C. Müller from the Bennett laboratory for technical support with sections of the experimental work and assistance with the analysis of the iTRAQ data, respectively. The authors are grateful to members of the Colinge Bioinformatic group at CeMM: Florian P. Breitwieser and Jacques Colinge for assistance with the analysis of the iTRAQ data, Alexey Stukalov for advice on the DAVID analyses, and Peter Májek for technical assistance in data processing. In addition, the authors would like to thank Ulrike Resch from the Department of Vascular Biology and Thrombosis Research at the Medical University of Vienna and Edit Szodorai from Lubec laboratory at the Medical University of Vienna for consultation on the preparation of human PBMCs and assistance with the ChemiDoc™ MP Imaging System (BIO-RAD) and Image Lab software, respectively.

Funding Sources

Research in our laboratory is supported by the Austrian Academy of Sciences, the Austrian Federal Ministry for Science and Research (Gen-Au projects, APP-III and BIN-III), and the Austrian Science Fund FWF. E.L.R. is supported by the Gen-Au APP-III program (No. 820965).

References

[1] L. Giusti, P. Iacconi, F. Ciregia, G. Giannaccini, F. Basolo, G. Donatini, P. Miccoli, A. Lucacchini, Journal of Endocrinological Investigation, 30 (2014) 865-869. doi: bf03349229

[2] L. Giusti, P. Iacconi, F. Ciregia, G. Giannaccini, G.L. Donatini, F. Basolo, P. Miccoli, A. Pinchera, A. Lucacchini, J Proteome Res, 7 (2008) 4079-4088. doi: 10.1021/pr8000404

[3] L. Giusti, P. Iacconi, A. Lucacchini, Proteomics Clin Appl, 5 (2011) 24-29. doi: 10.1002/prca.201000091

[4] P. Iacconi, L. Giusti, Y. Da Valle, F. Ciregia, G. Giannaccini, L. Torregrossa, A. Proietti, G. Donatini, S. Mazzeo, F. Basolo, A. Lucacchini, J Med Case Rep, 6 (2012) 103. doi: 10.1186/1752-1947-6-103

[5] F. Ciregia, L. Giusti, A. Molinaro, F. Niccolai, P. Agretti, T. Rago, G. Di Coscio, P. Vitti, F. Basolo, P. Iacconi, M. Tonacchera, A. Lucacchini, PLoS One, 8 (2013) e72911. doi: 10.1371/journal.pone.0072911

[6] A. Rapkiewicz, V. Espina, J.A. Zujewski, P.F. Lebowitz, A. Filie, J. Wulfkuhle, K. Camphausen, E.F. Petricoin, 3rd, L.A. Liotta, A. Abati, Cancer, 111 (2007) 173-184. doi: 10.1002/cncr.22686

[7] M. Schirle, M.A. Heurtier, B. Kuster, Mol Cell Proteomics, 2 (2003) 1297-1305. doi: 10.1074/mcp.M300087-MCP200

[8] T.R. Burkard, M. Planyavsky, I. Kaupe, F.P. Breitwieser, T. Burckstummer, K.L. Bennett, G. Superti-Furga, J. Colinge, BMC Syst Biol, 5 (2011) 17. doi: 10.1186/1752-0509-5-17

[9] T. Geiger, A. Wehner, C. Schaab, J. Cox, M. Mann, Mol Cell Proteomics, 11 (2012) M111 014050. doi: 10.1074/

mcp.M111.014050

[10] S. Di Palma, D. Stange, M. van de Wetering, H. Clevers, A.J. Heck, S. Mohammed, J Proteome Res, 10 (2011) 3814-3819. doi: 10.1021/pr200367p

[11] L.F. Waanders, K. Chwalek, M. Monetti, C. Kumar, E. Lammert, M. Mann, Proc Natl Acad Sci U S A, 106 (2009) 18902-18907. doi: 10.1073/pnas.0908351106

[12] M. Maurer, A.C. Muller, C. Wagner, M.L. Huber, E.L. Rudashevskaya, S.N. Wagner, K.L. Bennett, J Proteome Res, 12 (2013) 1040-1048. doi: 10.1021/pr301009u

[13] M. Maurer, A.C. Muller, K. Parapatics, W.F. Pickl, C. Wagner, E.L. Rudashevskaya, F.P. Breitwieser, J. Colinge, K. Garg, J. Griss, K.L. Bennett, S.N. Wagner, J Proteome Res, 13 (2014) 2830-2845. doi: 10.1021/pr401270y

[14] A. Umar, T.M. Luider, J.A. Foekens, L. Pasa-Tolic, Proteomics, 7 (2007) 323-329. doi: 10.1002/pmic.200600293

[15] J.R. Wisniewski, P. Ostasiewicz, M. Mann, J Proteome Res, 10 (2011) 3040-3049. doi: 10.1021/pr200019m

[16] E.B. Haura, R. Sacco, J. Li, A.C. Muller, F. Grebien, G. Superti -Furga, K.L. Bennett, J Integr OMICS, 2 (2012) 55-68. doi: 10.5584/jiomics.v2i1.81

[17] L.C. Ngoka, Proteome Sci, 6 (2008) 30. doi: 10.1186/1477-5956-6-30

[18] P.L. Ross, Y.N. Huang, J.N. Marchese, B. Williamson, K. Parker, S. Hattan, N. Khainovski, S. Pillai, S. Dey, S. Daniels, S. Purkayastha, P. Juhasz, S. Martin, M. Bartlet-Jones, F. He, A. Jacobson, D.J. Pappin, Mol Cell Proteomics, 3 (2004) 1154-1169. doi: 10.1074/mcp.M400129-MCP200

[19] M. Gilar, P. Olivova, A.E. Daly, J.C. Gebler, Journal of Separation Science, 28 (2005) 1694-1703. doi: 10.1002/jssc.200500116

[20] K.L. Bennett, M. Funk, M. Tschernutter, F.P. Breitwieser, M. Planyavsky, C. Ubaida Mohien, A. Muller, Z. Trajanoski, J. Colinge, G. Superti-Furga, U. Schmidt-Erfurth, J Proteomics, 74 (2011) 151-166. doi: 10.1016/j.jprot.2010.10.002

[21] J.V. Olsen, L.M. de Godoy, G. Li, B. Macek, P. Mortensen, R. Pesch, A. Makarov, O. Lange, S. Horning, M. Mann, Mol Cell Proteomics, 4 (2005) 2010-2021. doi: 10.1074/mcp.T500030-MCP200

[22] J. Colinge, A. Masselot, M. Giron, T. Dessingy, J. Magnin, Proteomics, 3 (2003) 1454-1463. doi: 10.1002/pmic.200300485

[23] W. Huang da, B.T. Sherman, R.A. Lempicki, Nat Protoc, 4 (2009) 44-57. doi: 10.1038/nprot.2008.211

[24] D.W. Huang, B.T. Sherman, Q. Tan, J. Kir, D. Liu, D. Bryant, Y. Guo, R. Stephens, M.W. Baseler, H.C. Lane, R.A. Lempicki, Nucleic Acids Res, 35 (2007) W169-175. doi: 10.1093/nar/gkm415

[25] F. Supek, M. Bosnjak, N. Skunca, T. Smuc, PLoS One, 6 (2011) e21800. doi: 10.1371/journal.pone.0021800

[26] M. Qasim, H. Rahman, M. Oellerich, A.R. Asif, Proteome Sci, 9 (2011) 57. doi: 10.1186/1477-5956-9-57

[27] S. Bartel, J. Doellinger, K. Darsow, D. Bourquain, R. Buchholz, A. Nitsche, H.A. Lange, Virol J, 8 (2011) 380. doi: 10.1186/1743-422X-8-380

[28] L. Minafra, G. Di Cara, N.N. Albanese, P. Cancemi, Leuk Res, 35 (2011) 226-236. doi: 10.1016/j.leukres.2010.07.040

[29] V.J. Haudek, N.C. Gundacker, A. Slany, H. Wimmer, E. Bayer, K. Pable, C. Gerner, J Proteome Res, 7 (2008) 5138-5147. doi: 10.1021/pr800438f

[30] G. Maccarrone, C. Rewerts, M. Lebar, C.W. Turck, D. Martins-de-Souza, Proteomics, 13 (2013) 893-897. doi: 10.1002/pmic.201200377

[31] F.L. Graham, J. Smiley, W.C. Russell, R. Nairn, Journal of General Virology, 36 (1977) 59-72. doi: 10.1099/0022-1317-36-1-59

[32] V. Meininger, Arch Ital Biol, 149 (2011) 29-37. doi: 10.4449/aib.v149i1.1265

[33] G. Shaw, S. Morse, M. Ararat, F.L. Graham, FASEB J, 16 (2002) 869-871. doi: 10.1096/fj.01-0995fje

[34] L.I. Bruijn, Science, 281 (1998) 1851-1854. doi: 10.1126/science.281.5384.1851

[35] C. Sundström, K. Nilsson, International Journal of Cancer, 17 (1976) 565-577. doi: 10.1002/ijc.2910170504

[36] J.S. Passmore, P.T. Lukey, S.R. Ress, Immunology, 102 (2001) 146-156. doi: 10.1046/j.1365-2567.2001.01164.x

[37] V.J. Haudek, A. Slany, N.C. Gundacker, H. Wimmer, J. Drach, C. Gerner, J Proteome Res, 8 (2009) 3834-3843. doi: 10.1021/pr801085g

[38] L. Dayon, M. Kussmann, EuPA Open Proteomics, 1 (2013) 8-16. doi: 10.1016/j.euprot.2013.08.001

JOURNAL OF INTEGRATED OMICS
A METHODOLOGICAL JOURNAL
HTTP://WWW.JIOMICS.COM

ORIGINAL ARTICLE | DOI: 10.5584/jiomics.v5i1.184

Proteomic and lipidomic analysis of primary mouse hepatocytes exposed to metal and metal oxide nanoparticles

Sara Tedesco[#1], Narges Bayat[#2], Gabriela Danielsson[2], Xabier Buque[3], Patricia Aspichueta[3], Olatz Fresnedo[3], Susana Cristobal[*1,3,4]

[1]Department of Clinical and Experimental Medicine, Health Science Faculty, Linköping University, Linköping, Sweden. [2]Department of Biochemistry and Biophysics, Arrhenius laboratories, Stockholm University, Stockholm, Sweden. [3]Department of Physiology, Faculty of Medicine and Dentistry, University of the Basque Country UPV/EHU, Leioa, Spain. [4]IKERBASQUE, Basque Foundation for Science, Bilbao, Spain. [#]Both authors have equally contributed

Received: 09 January 2015 **Accepted:** 09 March 2015 **Available Online**: 26 April 2015

Abstract

The global analysis of the cellular lipid and protein content upon exposure to metal and metal oxide nanoparticles (NPs) can provide an overview of the possible impact of exposure. Proteomic analysis has been applied to understand the nanoimpact however the relevance of the alteration on the lipidic profile has been underestimated. In our study, primary mouse hepatocytes were treated with ultra-small (US) TiO_2-USNPs as well as ZnO-NPs, CuO-NPs and Ag-NPs. The protein extracts were analysed by 2D-DIGE and quantified by imaging software and the selected differentially expressed proteins were identified by nLC-ESI-MS/MS. In parallel, lipidomic analysis of the samples was performed using thin layer chromatography (TLC) and analyzed by imaging software. Our findings show an overall ranking of the nanoimpact at the cellular and molecular level: TiO_2-USNPs<ZnO-NPs<Ag-NPs<CuO-NPs. CuO-NPs and Ag-NPs were cytotoxic while ZnO-NPs and CuO-NPs had oxidative capacity. TiO_2-USNPs did not have oxidative capacity and were not cytotoxic. The most common cellular impact of the exposure was the down-regulation of proteins. The proteins identified were involved in urea cycle, lipid metabolism, electron transport chain, metabolism signaling, cellular structure and we could also identify nuclear proteins. CuO-NPs exposure decreased phosphatidylethanolamine and phosphatidylinositol and caused down-regulation of electron transferring protein subunit beta. Ag-NPs exposure caused increased of total lipids and triacylglycerol and decrease of sphingomyelin. TiO_2-USNPs also caused decrease of sphingomyelin as well as up-regulation of ATP synthase and electron transferring protein alfa. ZnO-NPs affected the proteome in a concentration-independent manner with down-regulation of RNA helicase. ZnO-NPs exposure did not affect the cellular lipids. To our knowledge this work represents the first integrated proteomic and lipidomic approach to study the effect of NPs exposure to primary mouse hepatocytes in vitro.

Keywords: nanoparticles; hepatocytes; proteomics; lipidomics; mass spectrometry; toxicity.

Abbreviations

2D-DIGE: two-dimensional difference gel electrophoresis; **NPs**: nanoparticles; **USNPs**: ultra-small nanoparticles; **ROS**: reactive oxygen species; **DLS**: dynamic light scattering.

1. Introduction

The rapid development of nanotechnology and its applications has led to a growing and widespread use of products containing NPs in a myriad of areas as diverse as electronics, cosmetics, food additives, and medicine [1]. Metal and metal oxide nanoparticles (NPs) such as Silver (Ag) titanium (IV) dioxide (TiO_2), zinc oxide (ZnO), and copper oxide (CuO) are some of the most common industrial NPs additives for various applications [2, 3]. We have previously shown the cytotoxicity as well as the cellular ultra-structural effects of these NPs on *Saccharomyces cerevisia*e [4]. In this study we focus on the effects of the mentioned NPs on hepatocytes considering that for those NPs that succeed in entering the

***Corresponding author:** Prof. Susana Cristobal, Department of Clinical and Experimental Medicine, Cell Biology, level 13, Faculty of Health Science Linköping University SE-581 85 Linköping, Sweden. Email address: Susana.Cristobal@liu.se Tel: +46-10-1030881

bloodstream, either after inhalation, via the gastrointestinal tract or dermal absorption, the liver is one of the most important targets. Previous studies have demonstrated high accumulation and retention of NPs in liver after injection and digestion respectively [5-7]. TiO_2–NPs are one of the most studied NPs due to their extensive application in paints, cosmetics, and sunscreens [8, 9]. The interest on ultra-small NPs (USNPs), size range between 1-3 nm, has increased enormously for its applicability to optics and theranostics [10, 11]. The uniqueness of USNPs arises from possessing an extremely large surface area to volume ratio. This property enables them to be regarded as large molecules and accentuating the properties derived from interfacial interactions of the surface atoms with the solvent [12, 13]. A previous study has shown that gold USNPs were able to penetrate deeply into tumor spheroids, showed high levels of accumulation in tumor tissue in mice, and were distributed throughout the cytoplasm and nucleus of cancer cells in vitro and in vivo, whereas at 15 nm, they were found only in the cytoplasm, where they formed aggregates [14]. However, information about the toxicity and effects of TiO_2-USNPs on the cellular response is scarce.

Another NPs of great interest are ZnO- NPs, which due to their remarkable ultra-violet (UV) absorption and optical properties, are included in personal care products such as toothpaste, cosmetics, and textiles [15]. However exposure to ZnO-NPs through inhalation has been shown to cause toxicity through a battery of mechanism including cell stress and inflammation [16]. It has been observed that ZnO-NPs elucidate their toxicity by release of ions which alter Zn homeostasis [17, 18]. This is particularly important in hepatocytes as Zn is an essential trace element required for normal cell growth and function, and Zn deficiency/altered metabolism is observed in many types of liver diseases [19, 20]. CuO-NPs are extensively applied due to their potential applications as gas sensors, catalysts, and superconductors [21]. Cu ions are essential and function as cofactor of many enzymatic reactions and would be cycling between the two redox states. This process can be the source of reactive oxygen species (ROS) [22]. Indeed as hepatocytes are responsible for the Cu ions balance of the body, they are a major target of exposure and line of defense in the case of exposure to CuO-NP. Previous studies have shown that toxicity of CuO-NPs as well as their interference with the Cu ion homeostasis in hepatocytes [23, 24]. Exposure to CuO-NPs has been shown to affect the fatty acid composition *Tetrahymena thermophila* [25]. Toxicity associated with CuO-NPs has been connected with release of Cu ions as well as with oxidative stress. Ag-NPs have been widely used in personal products, food service, medical instruments, and textiles because of their antibacterial effects [26, 27]. Internalized Ag-NPs can release ions which may lead to cellular metabolism and mitochondrial dysfunction, inducing directly and indirectly ROS generation [2, 28]. Previous studies have also shown the toxicity of Ag-NPs in hepatocytes by affecting homeostasis and reducing albumin release [5] or by stimulating glycogenolysis [29]. Numerous studies have demonstrated that the NPs interaction with serum proteins and cell membranes receptors is determined by the NPs design, affecting cellular uptake, gene and protein expression, and toxicity [30]. It has been reported the interaction of NPs with proteins, lipoproteins and plasma membrane might compromise its fluidity and integrity and/or facilitate the entry of the NPs [31]. However most of the studies showing NPs uptake have been mainly conducted on immortalized cell lines, whereas little is known those effects on primary cells [30]. Primary hepatocytes cultures represent a powerful *in vitro* system, as these cells are directly isolated from the animal keeping the parental specific properties of the liver (*in vivo*) from which they are derived unaltered. The aim of this study is to provide a functional understanding of the impact of the studied NPs in primary hepatocytes. The strategy is to apply a combined OMICs approach, lipidomics and proteomics that could integrated the functional role of lipids in the cellular response. Therefore, the differentially expressed proteins identified in combination with the changes in the lipid composition of the membranes may contribute to understanding the possible effects and exposure risks of the selected NPs. The field of nanotoxicology is aiming to fill gaps on the NP impact and system biology strategies could lead to evaluate possible outcome adverse pathways for human, animals and the environment.

2. Material and Methods

NPs characterization

The following NPs were used in this study: titanium (IV) oxide, 14027, dry nanopowder, rutile, average particle size: 1-3 nm (Plasmachem GmbH, Münster, Germany), ZnO nano powder, 544906, average size <100 nm, Copper (II) oxide nano powder, 544868, average size <50 nm, Ag-NPs aqueous colloidal solution, 0.1 mg/mL, and average particle size: 10 nm were purchased by Sigma (St. Louis, MO, USA). All NPs stock suspensions were prepared by suspending NPs in hepatocytes culture medium. The suspensions were prepared freshly, sonicated in a water bath sonicator for 30 min and vortexed vigorously before each assessment. The average hydrodynamic size by DLS measurement and the zeta potential were determined using a Malvern Zetasizer Nano series V5.03 (PSS0012-16 Malvern Instruments, Worcestershire. UK) and the analysis program DTS (dispersion technology software, Malvern Instruments). Two concentrations of NPs were used in order to assess their size and zeta potential: 5 and 500 ppm that correspond to the exposure and the stock suspension concentration, respectively. The measurements were conducted in clear disposable capillary cells (DTS1060).

Cell-free dichlorofluorescein (DCFH) assay

The study of the oxidative potential of NPs was measured by a cell free method described by Foucaud *et al.* [32] and

modified for this study. Briefly, 2′,7′ dichlorofluoroscein diacetate (DCFH-DA, Molecular Probes D-399) at 2.2 mM was hydrolyzed to DCFH at pH 7.0 with 0.01 N NaOH. The solution was put in the dark for 30 min at room temperature and the chemical reactions was stopped by adding ice cold 0.1 M PBS. Then, horse radish peroxidase (HRP, Sigma P8125) at 20U/ml was added to each sample. To facilitate the comparison between a cellular and cell free system, the solutions were incubated at 37°C in the dark. The fluorescence generated by the DCFH oxidation was measured using a microplate reader at 485 nm excitation and 530 nm emission after 120 min. Freshly diluted hydrogen peroxide (10μM) was used as a positive control. The data were recorded as arbitrary fluorescence units. Two technical and three biological replicates were performed.

Isolation and exposure of primary mice hepatocytes to NPs

Hepatocytes were isolated from C57/6J mice by a collagenase (Roche Diagnostics, Barcelona, Spain) perfusion technique, as described previously [33]. Cells were seeded on fibronectin-coated dishes (3.5 μg/cm2) (2.5 x 10^6 viable cells per plate) and cultured at 37 °C and 5% CO_2 as described by Palacios *et al.* [34]. The culture medium was Ham's F-12/ Leibovitz L-15 (1/1, v/v) supplemented with 2% newborn calf serum, 2 mM L-glutamine, 5 mM glucose, 5 U/mL penicillin, 5 mg/mL streptomycin, 50 mg/L gentamycin, 0.2% fatty acid-free bovine serum albumin (BSA), and 10 nM insulin. After 1 h of adhesion, the medium was changed and the hepatocytes were exposed to different types of NPs for 48 h, frozen in liquid nitrogen and stored at -80 °C. In this study, primary cultures of mouse hepatocytes were treated with the previously described metal and metal oxide NPs (TiO_2, ZnO, CuO, and Ag-NPs) at 1 and 5 ppm concentrations for 48 h. The choice of the concentrations was based on a previous *in vitro* study of catfish primary hepatocytes and human cells exposed to metal oxide NPs with some modifications [35]. All the experiments were conducted in compliance with institutional guidelines, and the analyses were performed on at least four biological replicates for each treatment (control included) unless specified otherwise. Animal procedures were approved by the University of the Basque Country and Animal Care and Use Committees.

Cell viability assay

The cytotoxicity of NPs was determined using standard MTT assay described previously with slightly modifications [36]. Briefly, primary mouse hepatocyte cells were plated in two 96-well culture plates in 200 μl of culture medium at a density of 1 x 10^5 cells/ml. After incubation for 24 h, NPs at concentrations of 1 and 5 ppm were added to respective cells. The cells were then cultivated for an additional 48 h with NPs containing medium changed every day. On the third day, 20 μl of tetrazolium dye MTT solution (5 mg/ml) was added to each well and was further incubated for 4 h. The supernatants were then removed and 200 μl of dimethyl sulfoxide (DMSO) was added to dissolve the formazan crystal at 37 °C. The absorbance was measured with a VICTOR3™ multi-labeled microplate reader (Perkinelmer Inc., Waltham, MA USA) at 560 nm. The assay was performed twice with three replicates for each sample in each assay.

Preparation of protein extracts

Hepatocytes media was carefully discarded and cells pellets (~ 1.5 x 10^6 cells per sample) were re-suspended in cell washing buffer solution (10mM Tris-base pH 8, 5mM of magnesium acetate) centrifuged at 12,000 g at 4°C for 4 min for three times according to the manufacturer's instructions (GE Healthcare). Later, hepatocytes were re-suspended in lysis buffer (2% ASB14, 8M urea, 5mM magnesium acetate, 20mM Tris-base pH 8.5)[37], left on ice for 10 min, and sonicated intermittently on ice until cells were lysed. Cell debris was removed by centrifugation at 12,000 g at 4°C for 10 min while the supernatant was transferred in new tubes followed by 20% of trichloroacetic acid (TCA) in cold acetone at -20°C overnight.

The protein precipitates were collected by centrifugation at 12,000 g for 5 min, and then the proteins were solubilized again in lysis buffer. Cycles of intermittent sonication followed by centrifugation at 10,000 g for 10 min were performed until all proteins were solubilized in the buffer and no evidence of precipitate was observed. All these steps were carried at 4 °C. Before DIGE labeling, protein concentrations were measured according to Bradford method [38].Bovine serum albumin was used as standard.

Cy-Dye labeling and separation of proteins by 2DE

Protein CyDye labeling and DIGE analysis were performed according to the manufacturer's instructions (GE Healthcare). Samples containing 25μg of solubilized proteins were labeled by 200 pmol of reconstitute CyDye. The quenched Cy3- and Cy5-labeled samples for each experimental sample were then combined with the quenched Cy2-labeled pool internal standard. These samples were then quenched by the addition of 1 μl 10 mM lysine followed by incubation on ice for 10 min. The total proteins (75μg) were mixed and denatured in sample buffer (7M urea, 2M thiourea, 2% ASB 14, 2% DTT, 2% IPG buffer (pH 3-10)), and then rehydrated with rehydration buffer (7M urea, 2M thiourea, 2% ASB 14, 0.2% DTT, 1% IPG buffer (pH 3-10)) and trace amounts of bromophenol blue. A final volume of 200 μl of sample was then distributed evenly along IPG strip pH 3–10NL, 11 cm, covered by mineral oil and passively rehydrated for at least 12 h in dark conditions. Isoelectric focusing was performed on a Protean IEF Cell (Bio-Rad) at 20°C using wet wicks inserted between the IPG strips and the electrodes. The first dimension was carried using the following program as recommended by the manufacturer's instructions (Bio-Rad): rapid voltage slope at all the steps; step 1, 250 V for 15 min; step 2, 8000 V

for 2.5 h, and step 3 at 8000 V until 35000 Vh was reached. After focusing the strips were equilibrated for 15 min in equilibration buffer (6 M urea, 0.375 M Tris, pH 8.8, 2% SDS, 20% glycerol) containing 2% DTT and then for 15 min in equilibration buffer containing 2.5% iodoacetamide. The second dimension was carried out on homogeneous 12.5% T Criterion precast gels (Bio-Rad, Hercules, CA) at 120 V for 2h using a Criterion Cell (Bio-Rad). DIGE gels were fixed in 10% methanol and 7.5% acetic acid for 1h in the dark and washed with bi-distilled water for 15 min before image acquisition. After image acquisition the gels were stained by colloidal Coomassie blue staining for subsequent spot picking and protein identification.

Image acquisition and analysis

DIGE gels were scanned using FLA-5100 Fluorescence Image Analyzer (Fuji Medical, Stamford, CT) according to manufacturer's recommendation. DIGE images (16 bit TIFF, 600 PMT) were analyzed by REDFIN software (Ludesi, Malmö, Sweden, http://www.ludesi.com) for spot detection, spot quantification and normalization, spot matching and statistical analysis. The comparison of test spot volumes (Cy3 or Cy5 labelled) with the corresponding internal standard spot volume (Cy2 labeled) gave normalization for each matched spot. This allows a satisfactory quantification and comparison of different gels. Differential expression of proteins was defined on the basis of ≥1.5-fold change between group averages and one-way ANOVA $p \leq 0.05$.

Protein identification by mass spectrometry

Mass spectrometry analysis for protein identification was performed on nano-LC-MS/MS (Bruker Daltonics, Bremen, Germany) after protein spot excision and trypsin in-gel digestion. Briefly, differentially expressed spots excised proteins were treated with 25mM of NH_4HCO_3 in 50% of acetonitrile (ACN) until complete de-staining, dried with 99.5% ACN, and digested with sequencing grade modified trypsin in 25mM NH_4HCO_3 for 16 hours at 37°C. The peptides were extracted twice with 5% formic acid (FA) in 50% ACN and dried in Speed Vac concentrator (THERMO SAVANT, Holbrook, NY, USA). The fractions were desalted using C18 ZipTip (Millipore) following the manufacturer's instructions and the nano-electrospray capillaries were loaded with 6 µl of peptide solutions in 50% ACN in water with 0.1% FA. A 20 mm × 100 µm pre column followed by a 100 mm × 75 µm analytical column both packed with reverse-phase C18 were used for separation at a flow rate of 300 nl/min. The gradient buffers used were 0.1% formic acid in water (A) and 0.1% formic acid in 100% acetonitrile (B). Separation was performed with a linear gradient for 60 min (100-0% sol. A in 60 min, 0-100% sol. B in 60 min). Automated online tandem MS analyses were performed when peptide ions were sequenced using two alternating fragmentation techniques: collision induced dissociation (CID) and electron transfer dissociation (ETD). The data obtained were analyzed by Bruker Daltonics DataAnalysis 3.4 and the resulting MGF files where used to search for protein in Swissprot (*Mus musculus*) using Mascot Server (2.3) (www.matrixscience.com). The search parameters allowed mass error up to 0.8 Da for MS data and up to two missed trypsin cleavage. Peptide modifications searched for included carbamidomethyl (Cys) as the only fixed modification, and up to two variable modifications from among the following: oxidation (Met), acetyl (N-term), pyroglutamate (Gln) and Met-loss (N-term). Significance threshold in the MASCOT searches was set as $p<0.01$. Peptides were considered reliable if the MS/MS spectra had a MASCOT score above 35 and an expect value below 0.01.

Molecular weight and pI of the identified proteins were calculated with the Expasy compute pI/Mw tool (http://www.expasy.ch/tools/pi_tool.html).

Extraction, separation and quantification of lipids

After quantification of the amount of cellular protein by the bicinchoninic acid method following manufacturer (PIERCE) instructions, lipids were extracted from 2 mg of cellular protein following the method of Folch *et al.* [39]. Briefly, eight volumes of chloroform/methanol/water (2:1:0.0075, v:v:v) were added and the methanol phase was re-extracted with four volumes of the same mixture. The chloroform phases were aspirated, combined, and washed with 1.5 ml of 0.88% KCl. Different species of lipids were separated using a thin-layer chromatography system composed of six sequential mobile phases as described by Ruiz and Ochoa [40]. Standard curves for all lipid classes were run in each plate. The lipid spots were quantified as detailed previously [41] using Quantity One software (Bio-Rad). Analysis was carried out at least twice per extract.

Statistical analysis

Statistical analysis was performed using GraphPad Prism version 5.02 (GraphPad Software, San Diego, CA). Paired comparisons were made using Student's t-test while the comparison of multiple treatments to a common control was performed using one-way analysis of variance (ANOVA) with Dunnett's test, and $p < 0.05$ was considered significant.

3. Results

NPs characterization

The results of NPs characterization in powder form and dispersed in the cell media are represented in Table 1. Information about the properties of the NPs in powder form was obtained from the manufacturer. NPs in the hepatocyte culture media showed agglomeration and/or aggregation. The NPs hydrodynamic size was characterized using Dynamic light scattering (DLS) which showed, in general, a bimodal distributions at concentrations 5 and 500 ppm. The hydrody-

Table 1. Characterization of Nanoparticles (NPs). NPs properties in powder form and dispersed in hepatocytes media. Ag-NPs: Zeta-potential values are not showed (-) due to several aggregations. SEM images of the largest NPs (i.e. CuO- and ZnO-NPs). Information about NPs properties from the powder (or liquid form for Ag-NPs) was provided from the manufacturing companies.

			Powder		Suspension		
NPs	Purity (%)	Crystal structure	Size (nm)	Specific surface area (M^2/g)	Concentration ppm	Size (nm)	Z-potential (mV)
TiO_2	99+	Rutile	1-3	470	5 500	$6.6e^4$ $1034e^5$	-0.5±0.1 -0.9±0.6
ZnO	79.8	Hexagonal Wurtzite	<100	15-25	5 500	440.7±110.7 747.4±3.9	-4.6±1.0 -8.2±0.4
CuO	77.3	Monoclinic Crystals	<50	29	5 500	- 939.6±10.6	4.0±5.6 -7.4±2.7
Ag	99+	Spheres	10	60	5 500	85.4±5.6	-8.5±2.5

namic size of CuO-NPs could not be obtained at 5 ppm due to high noise to signal ratio. Generally, a stable suspension has a zeta potential value higher or lower than +/-30 mV (Malvern) and therefore none of the NPs were in stable suspension.

NPs oxidative ability and impact in cell viability

The oxidative ability of the metal and metal oxide NPs was investigated by cell-free dichlorofluorescein (DCFH) assay using 5 and 1 ppm after 2 h exposure (Figure 1A). Our results evidenced that only ZnO-NPs and CuO-NPs at 5 ppm had significant oxidative activity ($p<0.01$) while Ag-NPs and TiO_2 -USNPs at 5 ppm showed a significantly low fluorescent intensity ($p<0.01$), remarking their negligible oxidizing activity. The cell viability has been assessed by MTT assay after NPs exposure for 48 h. Hepatocytes exposed to low and high concentration of TiO_2-USNPs, and ZnO-NPs, and to low concentration of CuO- and Ag-NPs did not show effects in the cell viability. However, the viability of the hepatocytes exposed to high concentration of CuO-NPs and Ag-NPs significantly decreased by 50% compared to non-treated cells (Figure 1B).

Proteomic analysis of impact of NPs exposure

Two dimensional DIGE (2D-DIGE) images of the protein extracts from hepatocytes (NPs treated and untreated) were imported to REDFIN software that detected 998 spots per gel (Supplementary Figure S1) evenly distributed along the whole range of pH (3-10) but more abundant between 24-150 kDa. Comparisons between several groups control versus all treated or each treatment were taking in consideration for the statistical analysis of the data. The comparison control versus all NPs treatments revealed a total of84 spots differentially expressed ($p<0.05$, fold change ratio≥1.5) (Figure 2A). In particular exposure to CuO-NPs and Ag-NPs at 5 ppm showed the largest number of modified proteins. ZnO-NPs exposure showed similar number of differentially expressed proteins at both concentrations, underlining a concentration-independent response. The TiO_2-USNPs exposures caused the least modified protein profiles (Figure 2B). We found the highest number of unique spots at the high concentration exposure for all NPs. However, the concentration-dependent response varied among the NPs studied. The CuO-NPs and Ag-NPs exposures duplicated and triplicated respectively, the number of differentially expressed spots from low to high concentration whereas a very low increase of concentration–dependent response was observed at TiO_2-USNPs and ZnO-NPs exposures. The impact at the protein level of the NP exposures was characterized by down-regulation. In hepatocytes exposed to Ag-NPs, most of the differentially expressed proteins were down-regulated underlining the strongest effects on the proteome. The changes in protein expression profile ($p<0.05$, fold change ratio ≥2) caused by exposure to the studied type and concentration of NPs were summarized in the supplementary material (Supplementary Figure S1 and S2).

Identification of differentially expressed proteins

Considering the analytical method applied, 2D-DIGE, and the results showing a general response based on down-regulation, many differentially expressed spots were under the expression level required for identification. For those spots, additional trials were performed after pooling the same spot from all the DIGE gels but unfortunately some excised and selected spots analyzed by mass spectrometry remained still unidentified. The identified proteins were selected among the proteins differentially expressed ($p<0.05$ and with fold change ≥1.5) and in common with at least two NPs exposures included the comparison control versus all NPs treatments (Figure 3, Table 2). Most of the identified proteins were common among all the exposures but some NPs had specific effect on the expression of unique proteins. The protein (ID25) carbamoyl-phosphatase synthase (CSP1) was the most commonly differentially expressed protein being up-

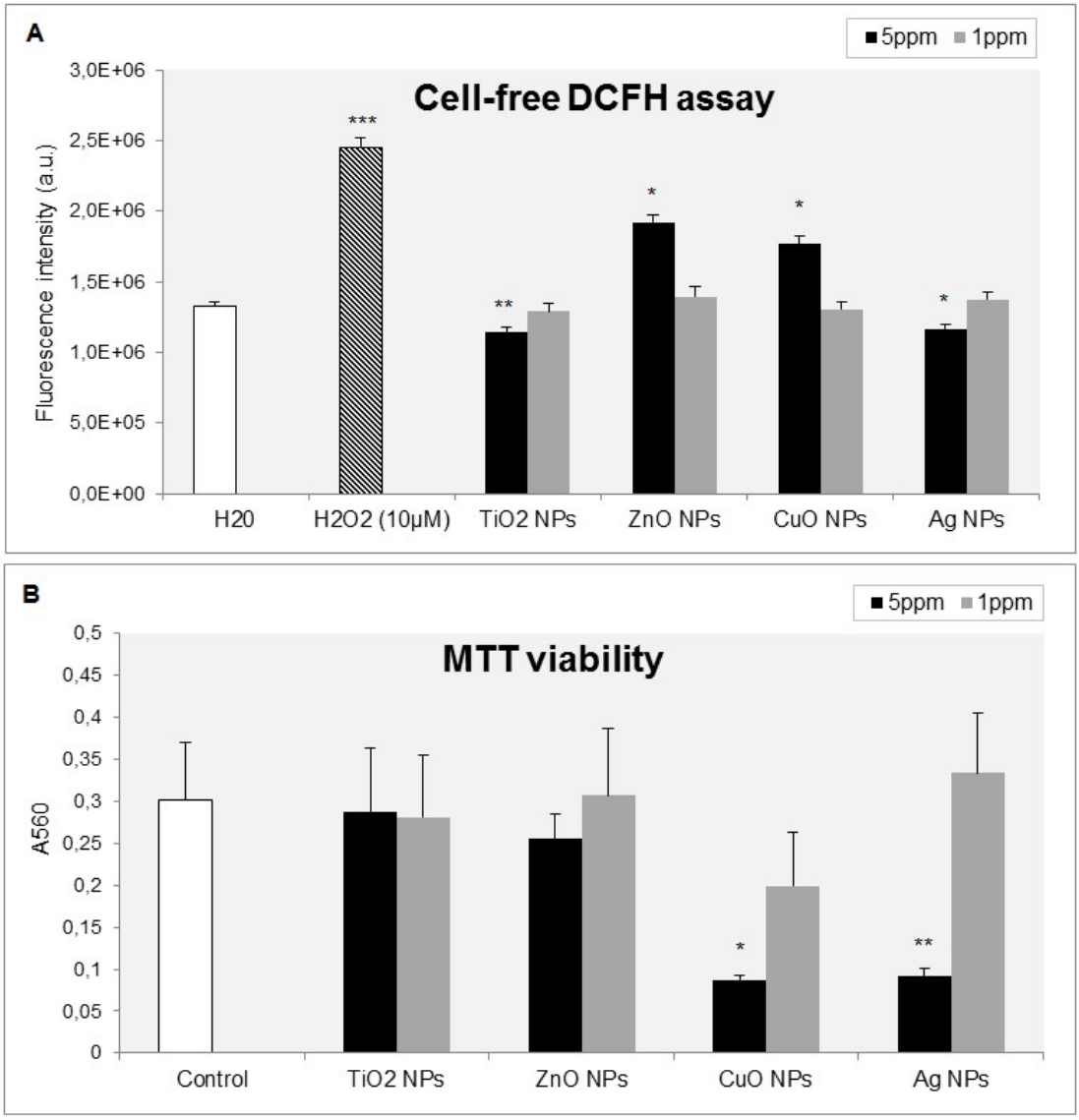

Figure 1. A) Oxidative potential assay. Fluorescence intensity [arbitrary units (a.u.)] of the NPs after incubation with DCFH for 2 h at 37°C. Values are the mean ± SEM from three experiments. For each treatment, two concentrations were used 1 and 5 *** $p < 0.001$. B) MTT assay for estimation of cell viability, expressed as absorbance at 560 nm. *$p< 0.01$ and *$p<0.001$.

regulated in CuO-NPs (5 ppm), ZnO-NPs (5 and 1 ppm) and Ag NP (5 ppm). TiO_2-USNPs caused the up-regulation ATP-Synthase and ETF protein subunit alpha while CuO-NPs (1 ppm) caused the down-regulation of ETF protein subunit beta as well as Tubulin beta-6 chain (ID497) at both concentrations. ZnO-NPs caused the down-regulation of RNA helicase (Figure 3). Approximately 50% of the identified proteins are localized in the specific organelles such as mitochondria (including matrix and membrane) while the remaining proteins belong to cytoplasm and also with the exception of alpha-enolase (ID49 and ID102) and guanine nucleotide-binding protein (G-Protein) subunit beta-2-like 1 (ID 572) which can also be from cell membranes. The only nuclear protein identified was heterogeneous nuclear ribonucleo-protein F (HNRPF) (ID222) (Table 2). The only protein with unclear subcellular localization was helicase eIF4A (ID 273) which can be both in the nucleus and in the cytoplasm.

Post-translational modifications

The main post-translational modification found in numerous proteins was the oxidation of methionine residues which causes small change of pI from the theoretical value (Table 2). It is significantly in the mitochondrial ATP synthase subunit alpha (ID209), (ATPA) that showed a big difference in pI from the theoretical value (Table 2). However the sequence found by mass spectrometry (the pI value was 6.1), which is close to that observed by 2DE, would match with the main chain of this protein without transit peptide.

A

Control vs Treated	Spots number
TiO_2 NPs (5ppm)	32
TiO_2 NPs (1ppm)	23
ZnO NPs (5ppm)	51
ZnO NPs (1ppm)	55
CuO NPs (5ppm)	103
CuO NPs (1ppm)	35
Ag (5ppm)	84
Ag (1ppm)	41

B

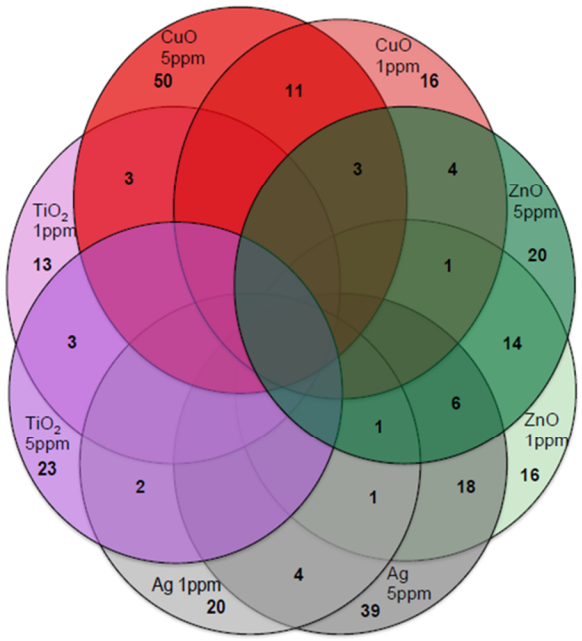

Figure 2. A) Differentially expressed proteins comparing control (untreated hepatocytes) versus each NPs exposure and B) Venn diagram representing differentially proteins among the exposures. The protein expression modification was considered significant for $p<0.05$ and fold change ratio≥1.5.

Lipidomics

Details on the lipid composition of hepatocytes from control and exposed to NPs at 5 ppm are represented in Figure 4. Interestingly, a significant decrease in the percentage of sphingomyelin (SM) was found in the cells exposed to Ag-NPs ($p<0.001$) but also exposed to TiO_2-USNPs ($p<0.05$) (Figure 4A). CuO-NPS exposure caused a decrease in the percentage of PI and PE (Figure 4A) which made the PC/PE ratio decreased (Figure 4B), a predictor of altered membrane fluidity. In the cells exposed to Ag-NPs changes in the total lipid quantities were observed with a significant increase of triacylglycerol (TG) cell content (Figure 4C).

4. Discussion

The application of quantitative proteomics in combination with lipidomics can be a useful method to illustrate the effects of NPs in cell lines. In this study the effects of exposure to TiO_2-USNPs, ZnO-NPs, CuO-NPs and Ag-NPs for 48 h were studied on primary mouse hepatocytes. After characterization of the physicochemical properties of the NPs, their cytotoxicity was assessed followed by quantitative proteomic and lipidomic analysis. Based on the cellular and molecular effects on the primary mouse hepatocytes, the overall ranking of the impact of the NPs exposures is as follows: TiO_2<ZnO<Ag<CuO.

Cytotoxicity of NPs

TiO_2-USNPs (1-3 nm) used in this study were not cytotoxic (Figure 1B) at 1 or 5 ppm. They did not produce significant ROS (Figure 1A) and the insoluble nature of TiO_2-NPs has been shown in previous studies [42]. Thus effects observed upon exposure to TiO_2-USNPs can be solely due to their size and direct interactions with cellular components. ZnO-NPs exposures did not affect to the cellular viability, although high concentration exposures could cause cytotoxicity in *in vitro* [15, 43]. However, despite lack of toxicity, these NPs produced significant ROS (Figure 1A) and based on a previous study conducted by this group, ZnO-NPs and CuO-NPs had the highest capacity of ions leakage [4]. Previous studies have illustrated the importance of Zn ions in progression of alcoholic liver disease and hepatic lipid homeostasis where it was shown that Zn supplementation reverses alcoholic steatosis by inhibiting oxidative stress [19]. Therefore the impact of ZnO-NPs exposure on the proteome could be related to the disruption of Zn homeostasis and in combination with the increase of ROS levels cause cytotoxicity. As mentioned, similar to ZnO-NPs, CuO-NPs produced ROS (figure 1A) and leaked ions. However the exposure to CuO-NPs caused the most severe effects at the cellular and molecular level with significant reduction of cell viability. The severe toxicity of CuO-NPs has been shown previously [23, 24]. Since the amount of ROS produced alone

Table 2. List of identified proteins by nano-LC-MS/MS after selection from the differentially expressed proteins ($p<0.05$ and with fold change ≥1.5) and in common with at least two NPs exposures included the comparison control versus all NPs treatments.

Spot no	Acession no	ID protein	Theor./ Obs. pI	Theor. Mr (Da)	Obs. Mr (Da)	Mascot score	SC (%)	Peptide sequence (if only one peptide)	Functional pathway	Subcellular location
25	gi|124248512	Carbamoyl-phosphate synthase	6.48/~6	165711	~150000	2125	48		Urea cycle	Mitochondrion
26	gi|183396771	60 kDa heat shock protein	5.91/~4.8	61088	52000-76000	1677	57		Chaperone	Mitochondrion matrix
34	gi|1352250	Aldehyde dehydrogenase	7.53/~6.2	57015	38000-52000	309	13		Alcohol metabolism, Aldehydes oxidation	Mitochondrion matrix
49 102	gi|13637776	Alpha-enolase	6.37/~5.8	47453	38000-52000	943 267	67 32		Carbohydrate degradation, glycolysis	Cytoplasm; Cell membrane
91	gi|61252474	Hydroxy-methylglutaryl-CoA synthase	8.65/~7	57300	38000-52000	310	31		Lipid synthesis	Mitochondrion
209	gi|416677	ATP synthase subunit alpha	9.22/~5.8	59830	38000-52000	176	12		ATP synthesis, Transport	Mitochondrion
222	gi|81918016	Heterogeneous nuclear ribonucleo-protein F	5.31/~5.2	46043	~38000	116	9	K.ITGEAFVQFA QFASQELAEK.A	Nucleotide binding, single-stranded RNA binding	Nucleus
225	gi|341941780	Cytochrome b-c1 complex subunit 1	5.81/~4.8	53446	31000-38000	115	26		Mitochondrial electron transport	Mitochondrion inner membrane
227	gi|342187137	Mitochondrial 3-oxoacyl-CoA thiolase	8.33/~9	42260	31000-38000	1355	74		Lipid metabolism	Mitochondrion
230	gi|55977481	Tubulin beta-4B chain	4.79/~4.2	50255	38000-52000	2337	64		Structural molecule activity	Cystoplasm, cytoskeleton
249	gi|92090596	Electron transfer flavoprotein subunit beta	8.24/~8.2	27834	17000-24000	480	45		Electron carrier activity	Mitochondrion matrix
273	gi|46397464	ATP-dependent RNA helicase elF4A-1	5.32/~5.8	46353	~31000	64	4	K.TATFAISILQQ IELDLK.A	Helicase	Cystoplasm, cytoskeleton
342	gi|146345417	Electron transfer flavoprotein subunit alpha	8.62/~7	35330	17000-24000	2282	66		Electron carrier activity	Mitochondrion matrix
497	gi|66775966	Tubulin beta-6 chain)	4.79/~4.2	50255	52000-76000	45	12	K.GHYTEGAELV DSVLDVVR.K	Structural molecule activity	Cystoplasm, cytoskeleton
572	gi|54037181	Guanine nucleotide-binding protein subunit beta-2-like 1	7.60/~5.8	35511	12000-17000	68	24		Developmental protein	Cell membrane, cell projection cytoplasm, cytoskeleton, nucleus

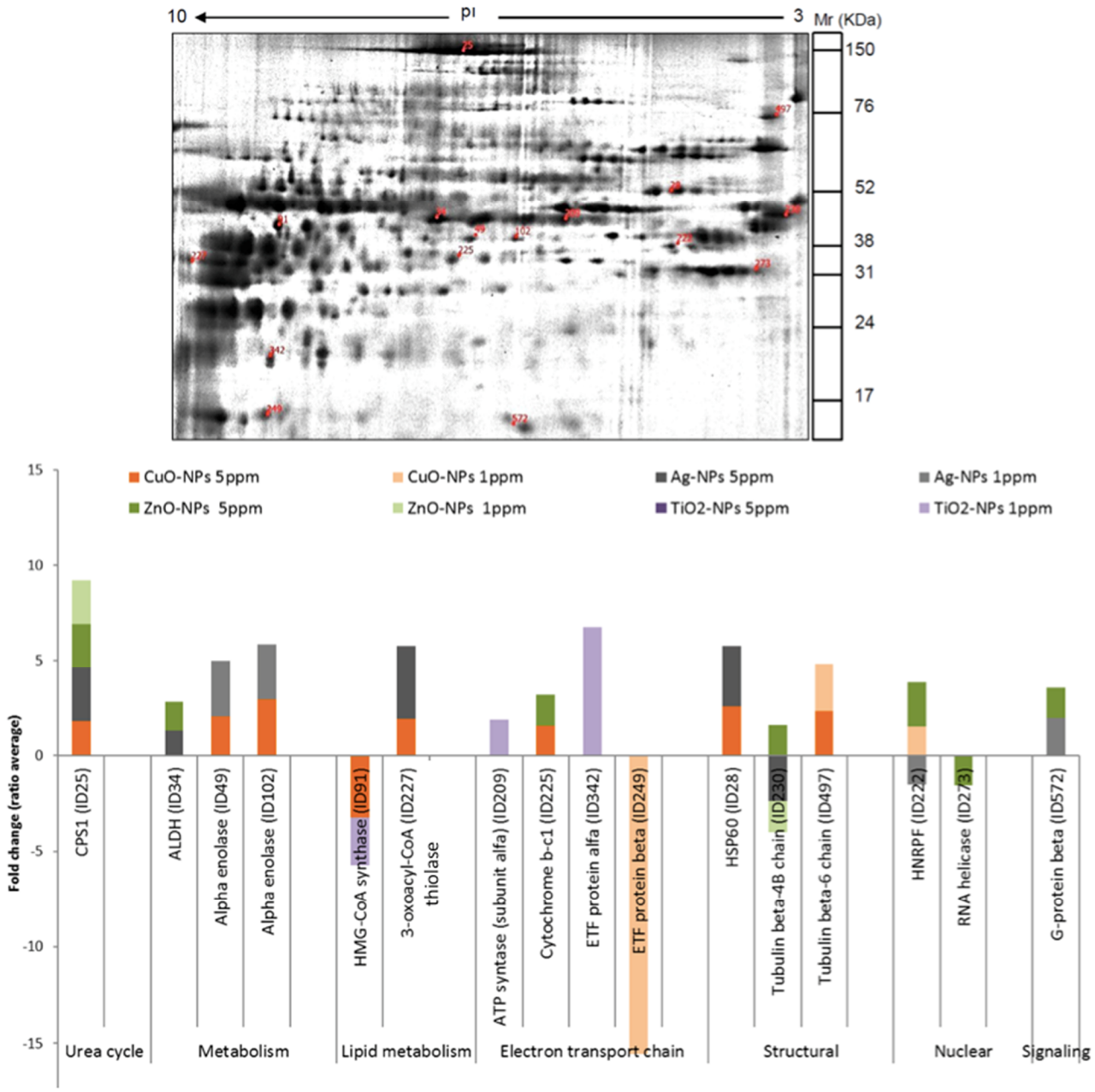

Figure 3. A) Representative 2D-DIGE with identified proteins and correspondent ID spot number. B) The protein expressions of the identified ID spots are illustrated as mean ± SEM based on fold change ratio value for the differentially expressed proteins and classified according to biological functions.

could not be the unique cytotoxic input (as shown for ZnO-NPs), it is likely that the released ions had actively contributed to the cytotoxicity. The importance of the intracellular solubility of NPs has arisen from understanding the Trojan horse-type mechanism of intracellular dissolution and its impact on the release of ions inside the cells leading to toxicity [44]. It has recently been reported that the intracellular solubility of CuO-NPs has the most critical role on the cytotoxicity [45]. Another type of NPs with great impact on the hepatocytes viability was Ag-NPs. These NPs however did not produce ROS. Previous studies have shown the uptake of the Ag-NPs despite different pattern of agglomeration as well as release of ions, both contributing to toxicity [46, 47].

Global impact of the NPs exposure to hepatocytes

The cellular impact of the NPs exposure was globally studied by combining proteomics and lipidomics. The differentially expressed proteins identified were involved in lipid metabolism, electron transport chain, structure of the cell, signaling, metabolism as well as nuclear proteins.

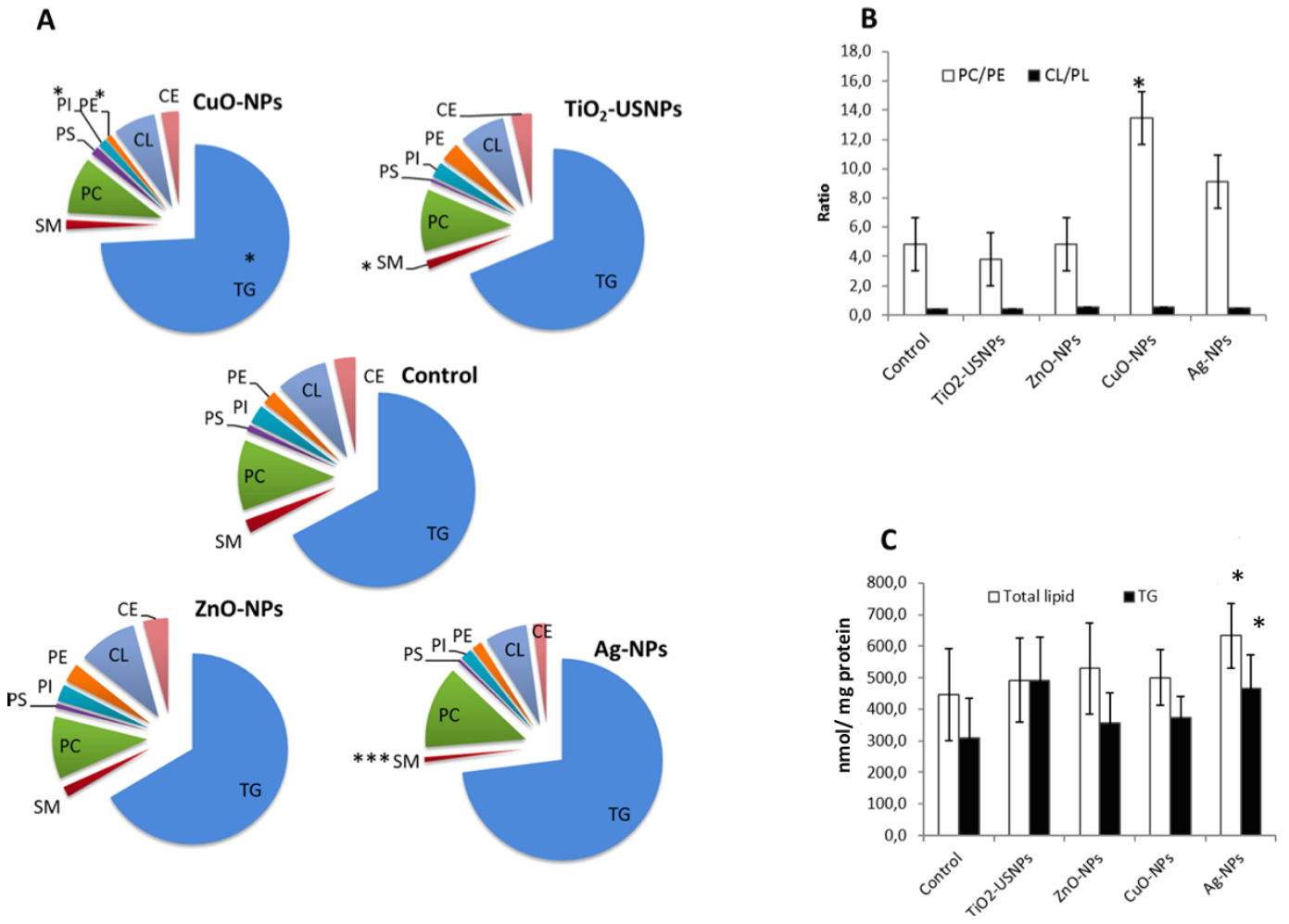

Figure 4. Distribution of total lipid content in control and exposures to NPs. A) Pie charts from percentages of lipid species; B) Ratio phosphatidylcholine/phosphatidylethanolamine and cholesterol/ phospholipid; C) Total lipid and total triacylglycerol in nmol/ mg protein TG, triacylglycerol; CL, cholesterol, CE, cholesteryl ester; PC, phosphatidylcholine; PE, phosphatidylethanolamine; PS, phosphatidylserine; PI, phosphatidylinositol; SM, sphingomyelin. Total lipid value corresponds to the summation of all measured lipid species, which are expressed as the percentage of the summation. Total phospholipid (PL) value corresponds to the summation of PC, PE, SM, PS and PI and total CL to the summation of FC and CE. Data are expressed as the mean ± SEM and correspond to the results obtained using 5 ppm concentration of NPs in the culture medium. Control vs. treated: *P ≤ 0.05, ***P ≤ 0.001.

Impact on lipids and fatty acid metabolism

One of the common cellular responses observed was variation of the cellular lipids (i.e. CuO-NPs, Ag-NPs and TiO_2-USNPs) and differential expression of proteins involved in fatty acid and lipid metabolism was also observed. The lipidomic results showed a significant decrease of percentage of SM in the hepatocytes exposed to TiO_2-USNPs at 5 ppm, although the PC/PE and CL/PL values indicated that the membrane fluidity was not affected (Figure 4). Lipid rafts, defined as cholesterol- and sphingolipid-enriched membrane micro-domains, might be altered by TiO_2-USNPs exposure in plasma membrane, triggering ROS release by enzymes localized in the membrane rafts. These ROS stimulate ceramide-releasing enzymes (e.g. acid sphingomyelinase) which are responsible for converting SM into phosphorylcholine and ceramide, increase the ceramide-enriched membrane platforms [48, 49]. It has been reported that carbon-based NPs treatment in lung epithelial cells led to an increase of ceramides in lipid rafts [50]. This feed-forward mechanism can justify the decrease of SM in the TiO_2-NPs exposure. The exposure to CuO-NPs caused significant increase of the ratio PC/PE and a decrease percentage of some PE and PI as well as increase in concentration of TG. The effect of Cu on the cellular lipid droplets has been shown previously [4]. Damage of the cellular plasma membrane has been shown to be one of the primary events in heavy metal (Cu and Zn) toxicity in plants [51, 52]. Previous studies have shown heavy metal stress increased PE, decreased PI, and PG [53], although the decrease in PE values observed in our study has also been

shown in other studies [54]. Cu deficiency has been shown to increase *in vivo* hepatic synthesis of fatty acids, TG, and PL in rats [55]. Therefore the decrease of this lipid class could be correlated to Cu overload. Cells exposed to Ag-NPs had decrease in SM but increase in the number of TG and total lipids. The increase in total lipids due to exposure to Ag-NPs has been observed previously [56]. Proteomic data in this study showed that mitochondrial HMG-CoA synthase was down-regulated in the cells exposed to TiO_2-USNPs at 1 ppm and to CuO-NPs at 5 ppm. This enzyme has a key function in regulating the ketogenesis, pathway involved in the biosynthesis of ketones bodies, metabolic fuel during starvation [57]. Another mitochondrial protein involved in lipid and fatty acid metabolism, 3-oxoacyl-coA thiolase was up-regulated in CuO-NPs, and particularly, in Ag-NPs treatment. This enzyme catalyzes the last step in mitochondrial and peroxisomal β-oxidation [58]. The increase the total lipids and TAG observed in cells exposed to Ag-NPs could have led to an increase in 3-oxoacyl-coA thiolase involved in beta oxidation and lipid metabolism.

Impact on proteins involved in electron transport chain

The differential expression of protein involved in the electron transport chain could reflect the increase in cellular energy demand upon exposure to NPs. CuO-NPs at both concentrations, TiO_2-USNPs (1 ppm) and ZnO-NPs (5ppm) affected these proteins. However proteins involved in this pathway were mostly affecting to one type of NPs exposure. The up-regulation of ATP synthase was only found in the hepatocytes exposed to TiO_2-USNPs. This protein is one of the most abundant proteins in the inner mitochondrial membrane which is involved in H^+ transport at the mitochondrial membrane and provides ATP [59, 60]. Another protein uniquely affected by TiO_2-USNP exposure was ETF subunit alpha which are heterodimers and function as electron shuttles between primary flavoprotein dehydrogenases involved in mitochondrial fatty acid and amino acid catabolism and the membrane-bound electron transfer flavoproteins ubiquinone oxidoreductase [61]. In cells exposed to CuO-NPs a remarkable reduction of the expression of ETFs subunit beta was detected. An imbalance of these "housekeeping" proteins can have serious repercussions especially in the oxidation of fatty acids [62]. ZnO-NPs and CuO-NPs at 5 ppm evidenced an increase of ROS and the up-regulation of the subunit 1 of cytochrome b-c1 complex or Complex III, protein. Complex III is the major ROS production site among all mitochondrial electron transport chain complexes, and it is the only complex that generates $^{\cdot}O2^{\cdot}$ in the mitochondrial inter-membrane space [63, 64]. Xia et al.[65] observed mitochondrial contribution to ZnO-NPs-induced ROS production, through the ultra-structural, and thereby membrane potential changes in this organelle. They also suggest that the release of Zn ions from NPs may exert extra-mitochondrial effects contributing to ROS generation, including NO production and generation of peroxynitrite (ONOO-). We have previously shown the significant release of Zn ions from ZnO-NPs [4].

Impact on proteins from urea cycle

CPS1, a mitochondrial enzyme involved in ATP-dependent formation of carbamoyl phosphate from glutamine or ammonia and bicarbonate in the first step of the urea cycle. This protein was over-expressed in the cells exposed to ZnO-NPs (5 and 1 ppm), Ag-NPs (5 ppm) and CuO-NPs (5 ppm). Generally, an increase of CPS1 expression has been observed in the case of liver damage or during acute hepatitis, as disorders induced by oxidative stress [66] and it is one of the main potential toxicity markers found in rat liver cells [67]. Previous studies have reported the effect of Zn in urea cycle and increased of activities of CPS1 in the liver of zinc-deficient rats[68]. It is interesting that the possible Zn ions released by the NPs in this study have caused the up regulation of CPS1.

Impact on nuclear proteins

ZnO-NPs were the only NPs that affected both RNA helicase, and hnRNP. It has been described how ultrafine NPs could affect the expression of nuclear proteins [69]. We observed that ZnO-NPs exposure specifically caused the down-regulation of the ATP-dependent RNA helicase (elF4) which plays important roles in the unwinding and remodeling of structured RNA as well as virtually all aspects of nucleic acid metabolism, and regulation, possibly enhancing the biosynthesis of altered proteins [70]. Previous study has shown that down-regulation in helicase is associated with cell cycle perturbations and in apoptosis which in this case might be an indication of oxidative stress and early stages of apoptosis experienced by the cells [71] .

Among all identified differentially expressed proteins, only one nuclear protein, the hnRNP F, was affected by NPs treatment and was down-regulated by treatment with Ag-NPs and up-regulated by ZnO-NPs, and CuO-NPs treatment. The hnRNP complexes are known to play a role in the regulation of the splicing events but they have also been shown to function in the regulation of cell proliferation. Overexpression of hnRNP F has been shown to promote cell proliferation while reverse effect was observed upon knockdown of hnRNP F [72]. Disruption in this protein therefore could lead to genotoxicity as well as disruption in cell proliferation. It is possible that the cytotoxicity observed in Ag-NPs exposed cells was due to down-regulation of this protein.

Impact on structural proteins

Another modified protein in hepatocytes exposed to ZnO-NPs or Ag-NPs (at 5 ppm) was ß-tubulin IV (TBB4B) which was down-regulated especially for the Ag-NPs treatment. This protein is the main constituent of microtubules, key

components of the cytoskeleton of eukaryotic cells and has an important role in various cellular functions such as intracellular migration and transport, cell shape maintenance, polarity, and cell signaling. Previous *in vitro* studies showed that metal and metal oxide NPs can directly bind functional groups of microtubules [73, 74]. In particular, Ag-NPs interacting with tubulin in correspondence of -SH residue may be responsible of ineffective mitotic spindle function [75][76]. Tubulin is the first non-receptor protein found to be phosphorylated by G-protein receptor kinases [77]. Interestingly both ZnO-NPs (5ppm) and Ag-NPs (1 ppm) induced an increase of G-protein expression involved in many cellular signaling pathways, including the ubiquitination and proteasome-mediated degradation [70]. The isotype of ß-tubulin (TBB6) was significantly up-regulated in hepatocytes exposed to CuO -NPs at 5 and 1 ppm which can contribute to an adaptation to oxidative stress conditions and drug resistance [78]. A compensatory mechanism from the hepatocytes exposed to CuO-NPs might occur to overwhelm the structural damages in the cytoskeleton, especially in the case of the highest concentration. HSPs function in important intracellular tasks such as protein folding and transport acting as chaperones under stress to prevent protein denaturation and loss of function [79]. HSP60 is a mitochondrial expressed stress protein that can be translocated to the cytosol and, later, transported to the cell surface. The HSP60 stress response is correlated with apoptosis and exacerbation of the disease state [80]. This protein was over-expressed in the two cytotoxic NPs i.e. Ag-NPs and CuO-NPs illustrating the apoptotic response of the cells.

Impact on cellular metabolism

Mitochondrial ALDH (ID34), and Alpha-enolase (ID49 or ID102) were found up-regulated in NPs treatments and can be considered as an early cellular defense response to general stress conditions. ALDH catalyzes the oxidation of various aliphatic and aromatic aldehydes to the corresponding acids and is in cellular defenses against toxic aldehydes [81]. Also it has been shown that mitochondria-located alpha-enolase stabilizes mitochondrial membrane and its' displacement may involve in activation of the intrinsic cell death pathway [82].

5. Concluding Remarks

Characterization of the NPs, classical toxicity assays and quantitative proteomics in combination with lipidomics could provide a detailed overview of the effects of NPs on primary hepatocytes. Most proteins identified to be differentially expressed were in common for the different NPs exposures and were involved in lipid metabolism, electron transport chain, cellular structure, metabolism, signaling as well nuclear proteins. CuO-NPs produced ROS, were cytotoxic, affected the PL and caused the down-regulation of ETF protein beta. Ag-NPs did not produce ROS but were cytotoxic, affected the SM as well as increasing total cellular lipids and TG. ZnO-NPs despite producing significant ROS were not cytotoxic and did not affect the cellular lipids but affected the RNA helicase. TiO_2-USNP did not produce ROS, were not cytotoxic yet affected the SM and affected ATP-synthase as well as ETF protein alpha. This work showed that some of our gaps for understanding the NP impact at the cellular level could be filled by combining data from alterations on lipidomic profiles with proteomic profiles. This OMICs methods or any extension to other OMICs methodologies would lead to a system biology understanding of NP impact and possible adverse outcome pathway.

6. Supplementary material

Supplementary data and information is available at: http://www.jiomics.com/index.php/jio/rt/suppFiles/184/0

Figure S1 - Representative 2D-DIGE proteins from hepatocytes exposed to NPs. A total of 998 spots were detected by REDFIN software.

Figure S2 - Proteins up- and down-regulated by NPs along with fold change (F.C.).

Table S1 - Lipidomics. TG, triacylglycerol; CE, cholesteryl ester; FC, free cholesterol; PC, phosphatidylcholine; PE, phosphatidylethanolamine; PS, phosphatidylserine; PI, phosphatidylinositol; SM, sphingomyelin. Total lipid quantities correspond to the summation of all measured lipid species, which are expressed as the percentage of the summation. Total phospholipid quantities correspond to the summation of PC, PE, SM, PS and PI and total cholesterol to the summation of FC and CE. Data are expressed as the mean ± SEM and correspond to the results obtained using 5 ppm concentration of NPs in the culture medium. Control vs. treated: $^{*}P \leq 0.05$, $^{***}P \leq 0.001$.

Acknowledgements

This work was supported by grants from the Swedish Research Council-Natural Science (VR-N) (SC), Carl Trygger Foundation (SC), VINNOVA-Vinnmer program (SC), Oscar Lilli Lamms Minne Foundation (SC), Längmanska kulturfonden (SC), Lars Hiertas Minne foundation (SC), IKER-BASQUE, Basque Foundation for science (SC), Ångpanneförening foundation (SC). We would like to thank Dr. Itsaso Garcia-Arcos for her help with our preliminary experiments with cell cultures, to Mr. Bengt-Arne Fredriksson for the TEM analysis, Jacob Kuruvilla, Christine Gallampois for their helpful suggestions on the elaboration of MS data.

References

[1] L. Mazzola, Nat Biotechnol, 21 (2003) 1137-1143. http://dx.doi.org/10.1038/nbt1003-1137

[2] C.N. Lok, C.M. Ho, R. Chen, Q.Y. He, W.Y. Yu, H. Sun, P.K.

Tam, J.F. Chiu, C.M. Che, J Proteome Res, 5 (2006) 916-924. http://dx.doi.org/10.1021/pr0504079

[3] E.J. Park, J. Yi, Y. Kim, K. Choi, K. Park, Toxicol In Vitro, 24 (2010) 872-878. http://dx.doi.org/10.1016/j.tiv.2009.12.001

[4] N. Bayat, K. Rajapakse, R. Marinsek-Logar, D. Drobne, S. Cristobal, Nanotoxicology, 8 (2014) 363-373. http://dx.doi.org/10.3109/17435390.2013.788748

[5] B.K. Gaiser, S. Hirn, A. Kermanizadeh, N. Kanase, K. Fytianos, A. Wenk, N. Haberl, A. Brunelli, W.G. Kreyling, V. Stone, Toxicol Sci, 131 (2013) 537-547. http://dx.doi.org/10.1093/toxsci/kfs306

[6] S. Hirn, M. Semmler-Behnke, C. Schleh, A. Wenk, J. Lipka, M. Schaffler, S. Takenaka, W. Moller, G. Schmid, U. Simon, W.G. Kreyling, Eur J Pharm Biopharm, 77 (2011) 407-416. http://dx.doi.org/10.1016/j.ejpb.2010.12.029

[7] C. Schleh, M. Semmler-Behnke, J. Lipka, A. Wenk, S. Hirn, M. Schaffler, G. Schmid, U. Simon, W.G. Kreyling, Nanotoxicology, 6 (2012) 36-46. http://dx.doi.org/10.3109/17435390.2011.552811

[8] N. Serpone, A. Salinaro, A.V. Emeline, S. Horikoshi, H. Hidaka, J. Zhao, Photochem Photobiol Sci, 1 (2002) 970-981

[9] N. Ambalavanan, A. Stanishevsky, A. Bulger, B. Halloran, C. Steele, Y. Vohra, S. Matalon, Am J Physiol Lung Cell Mol Physiol, 304 (2013) L152-161. http://dx.doi.org/10.1152/ajplung.00013.2012

[10] C.K. Kim, T. Kim, I.Y. Choi, M. Soh, D. Kim, Y.J. Kim, H. Jang, H.S. Yang, J.Y. Kim, H.K. Park, S.P. Park, S. Park, T. Yu, B.W. Yoon, S.H. Lee, T. Hyeon, Angew Chem Int Ed Engl, 51 (2012) 11039-11043. http://dx.doi.org/10.1002/anie.201203780

[11] H.S. Choi, W. Liu, P. Misra, E. Tanaka, J.P. Zimmer, B. Itty Ipe, M.G. Bawendi, J.V. Frangioni, Nat Biotechnol, 25 (2007) 1165-1170. http://dx.doi.org/10.1038/nbt1340

[12] B.H. Kim, N. Lee, H. Kim, K. An, Y.I. Park, Y. Choi, K. Shin, Y. Lee, S.G. Kwon, H.B. Na, J.G. Park, T.Y. Ahn, Y.W. Kim, W.K. Moon, S.H. Choi, T. Hyeon, J Am Chem Soc, 133 (2011) 12624-12631. http://dx.doi.org/10.1021/ja203340u

[13] L. Shang, S. Dong, G.U. Nienhaus, Nano Today, 6 (2011) 401-418. http://dx.doi.org/10.1016/j.nantod.2011.06.004

[14] K. Huang, H. Ma, J. Liu, S. Huo, A. Kumar, T. Wei, X. Zhang, S. Jin, Y. Gan, P.C. Wang, S. He, X. Zhang, X.J. Liang, ACS Nano, 6 (2012) 4483-4493. http://dx.doi.org/10.1021/nn301282m

[15] Y. Wang, W.G. Aker, H.M. Hwang, C.G. Yedjou, H. Yu, P.B. Tchounwou, Sci Total Environ, 409 (2011) 4753-4762. http://dx.doi.org/10.1016/j.scitotenv.2011.07.039

[16] M.J. Osmond-McLeod, R.I. Osmond, Y. Oytam, M.J. McCall, B. Feltis, A. Mackay-Sim, S.A. Wood, A.L. Cook, Part Fibre Toxicol, 10 (2013) 54. http://dx.doi.org/10.1186/1743-8977-10-54

[17] Y.Y. Kao, Y.C. Chen, T.J. Cheng, Y.M. Chiung, P.S. Liu, Toxicol Sci, 125 (2012) 462-472. http://dx.doi.org/10.1093/toxsci/kfr319

[18] T. Turney, M. Duriska, V. Jayaratne, A. Elbaz, S. O'Keefe, A. Hastings, T. Piva, P. Wright, B. Feltis, Chem Res Toxicol, 25 (2012) 2057 - 2066

[19] X. Kang, W. Zhong, J. Liu, Z. Song, C.J. McClain, J. Kang, Z. Zhou, Hepatology, 50 (2009) 1241-1250

[20] M.K. Mohammad, Z. Zhou, M. Cave, A. Barve, C.J. McClain, Nutr Clin Pract, 27 (2012) 8-20. http://dx.doi.org/10.1177/0884533611433534

[21] P.T. Leung, Y. Wang, S.S. Mak, W.C. Ng, K.M. Leung, Aquat Toxicol, 105 (2011) 49-61. http://dx.doi.org/10.1016/j.aquatox.2011.05.010

[22] B. Halliwell, J.M. Gutteridge, Methods Enzymol, 186 (1990) 1-85

[23] M. Cuillel, M. Chevallet, P. Charbonnier, C. Fauquant, I. Pignot-Paintrand, J. Arnaud, D. Cassio, I. Michaud-Soret, E. Mintz, Nanoscale, 6 (2014) 1707-1715. http://dx.doi.org/10.1039/c3nr05041f

[24] M.A. Siddiqui, H.A. Alhadlaq, J. Ahmad, A.A. Al-Khedhairy, J. Musarrat, M. Ahamed, PLoS One, 8 (2013) e69534. http://dx.doi.org/10.1371/journal.pone.0069534

[25] M. Mortimer, K. Kasemets, M. Vodovnik, R. Marinsek-Logar, A. Kahru, Environ Sci Technol, 45 (2011) 6617-6624. http://dx.doi.org/10.1021/es201524q

[26] J.S. Kim, E. Kuk, K.N. Yu, J.H. Kim, S.J. Park, H.J. Lee, S.H. Kim, Y.K. Park, Y.H. Park, C.Y. Hwang, Y.K. Kim, Y.S. Lee, D.H. Jeong, M.H. Cho, Nanomedicine, 3 (2007) 95-101. http://dx.doi.org/10.1016/j.nano.2006.12.001

[27] T.M. Benn, P. Westerhoff, Environmental Science & Technology, 42 (2008) 4133-4139. http://dx.doi.org/10.1021/es7032718

[28] M. Horie, H. Kato, K. Fujita, S. Endoh, H. Iwahashi, Chem Res Toxicol, 25 (2012) 605-619. http://dx.doi.org/10.1021/tx200470e

[29] A. Massarsky, J. Labarre, V.L. Trudeau, T.W. Moon, Aquat Toxicol, 147 (2014) 68-75. http://dx.doi.org/10.1016/j.aquatox.2013.11.026

[30] A. Albanese, P.S. Tang, W.C. Chan, Annu Rev Biomed Eng, 14 (2012) 1-16. http://dx.doi.org/10.1146/annurev-bioeng-071811-150124

[31] Colloids and Surfaces A: Physicochemical and Engineering Aspects, (2014)

[32] L. Foucaud, M.R. Wilson, D.M. Brown, V. Stone, Toxicol Lett, 174 (2007) 1-9. http://dx.doi.org/10.1016/j.toxlet.2007.08.001

[33] E. Isusi, P. Aspichueta, M. Liza, M.L. Hernandez, C. Diaz, G. Hernandez, M.J. Martinez, B. Ochoa, Atherosclerosis, 153 (2000) 283-294

[34] L. Palacios, B. Ochoa, M.J. Gomez-Lechon, J.V. Castell, O. Fresnedo, Biochim Biophys Acta, 1761 (2006) 698-708. http://dx.doi.org/10.1016/j.bbalip.2006.05.005

[35] Z. Wang, N. Li, J. Zhao, J.C. White, P. Qu, B. Xing, Chem Res Toxicol, 25 (2012) 1512-1521. http://dx.doi.org/10.1021/tx3002093

[36] J. Carmichael, W.G. DeGraff, A.F. Gazdar, J.D. Minna, J.B. Mitchell, Cancer Res, 47 (1987) 943-946

[37] J.E. Swatton, S. Prabakaran, N.A. Karp, K.S. Lilley, S. Bahn, Mol Psychiatry, 9 (2004) 128-143. http://dx.doi.org/10.1038/sj.mp.4001475

[38] M.M. Bradford, Anal Biochem, 72 (1976) 248-254

[39] J. Folch, M. Lees, G.H. Sloane Stanley, J Biol Chem, 226 (1957) 497-509

[40] J.I. Ruiz, B. Ochoa, J Lipid Res, 38 (1997) 1482-1489

[41] I. Garcia-Arcos, P. Gonzalez-Kother, P. Aspichueta, Y. Rueda, B. Ochoa, O. Fresnedo, Lipids, 45 (2010) 1101-1108. http://dx.doi.org/10.1007/s11745-010-3492-2

[42] S. Pakrashi, N. Jain, S. Dalai, J. Jayakumar, P.T. Chandrasekaran, A.M. Raichur, N. Chandrasekaran, A. Mukherjee, PLoS One, 9 (2014) e87789. http://dx.doi.org/10.1371/journal.pone.0087789

[43] J.X. Yu, T.H. Li, Cell Biosci, 1 (2011) 19. http://

dx.doi.org/10.1186/2045-3701-1-19

[44] L.K. Limbach, P. Wick, P. Manser, R.N. Grass, A. Bruinink, W.J. Stark, Environ Sci Technol, 41 (2007) 4158-4163

[45] A.M. Studer, L.K. Limbach, L. Van Duc, F. Krumeich, E.K. Athanassiou, L.C. Gerber, H. Moch, W.J. Stark, Toxicol Lett, 197 (2010) 169-174. http://dx.doi.org/10.1016/j.toxlet.2010.05.012

[46] A.R. Gliga, S. Skoglund, I.O. Wallinder, B. Fadeel, H.L. Karlsson, Part Fibre Toxicol, 11 (2014) 11. http://dx.doi.org/10.1186/1743-8977-11-11

[47] B. Fadeel, N. Feliu, C. Vogt, A.M. Abdelmonem, W.J. Parak, Wiley Interdiscip Rev Nanomed Nanobiotechnol, 5 (2013) 111-129. http://dx.doi.org/10.1002/wnan.1206

[48] C.A. Dumitru, Y. Zhang, X. Li, E. Gulbins, Antioxid Redox Signal, 9 (2007) 1535-1540. http://dx.doi.org/10.1089/ars.2007.1692

[49] A. Soloviev, E.M. Schwarz, M. Darowish, R.J. O'Keefe, J Orthop Res, 23 (2005) 1258-1265. http://dx.doi.org/10.1016/j.orthres.2005.03.019.1100230604

[50] H. Peuschel, U. Sydlik, S. Grether-Beck, I. Felsner, D. Stockmann, S. Jakob, M. Kroker, J. Haendeler, M. Gotic, C. Bieschke, J. Krutmann, K. Unfried, Part Fibre Toxicol, 9 (2012) 48. http://dx.doi.org/10.1186/1743-8977-9-48

[51] M. Janicka-Russak, K. Kabala, M. Burzynski, G. Klobus, J Exp Bot, 59 (2008) 3721-3728. http://dx.doi.org/10.1093/jxb/ern219

[52] J. Harwood, Recent Environmental Concerns and Lipid Metabolism, in: J.-C. Kader, P. Mazliak (Eds.) Plant Lipid Metabolism, Springer Netherlands, 1995, pp. 361-368.

[53] Amal Ahmed Morsy, K.H.A. Salama, H.A. Kamel, M.M.F. Mansour, EurAsian Journal of BioSciences, 6 (2012) 1-10

[54] J. Zhang, H. Tang, H. Sun, S. Xu, Dongbei Nongye Daxue Xuebao, 42 (2011) 91-94

[55] A.A. al-Othman, F. Rosenstein, K.Y. Lei, Proc Soc Exp Biol Med, 204 (1993) 97-103. http://dx.doi.org/10.3181/00379727-204-43640

[56] S.M. Ouda, Research Journal of Microbiology, 9 (2014) 34-42

[57] N. Shafqat, A. Turnbull, J. Zschocke, U. Oppermann, W.W. Yue, J Mol Biol, 398 (2010) 497-506. http://dx.doi.org/10.1016/j.jmb.2010.03.034

[58] B. Middleton, Biochem J, 132 (1973) 717-730

[59] S. Rexroth, A. Poetsch, M. Rogner, A. Hamann, A. Werner, H.D. Osiewacz, E.R. Schafer, H. Seelert, N.A. Dencher, Biochim Biophys Acta, 1817 (2012) 381-387. http://dx.doi.org/10.1016/j.bbabio.2011.11.006

[60] D. Degli Esposti, J. Hamelin, N. Bosselut, R. Saffroy, M. Sebagh, A. Pommier, C. Martel, A. Lemoine, Biochem Res Int, 2012 (2012) 387626. http://dx.doi.org/10.1155/2012/387626

[61] F.L. Crane, H. Beinert, J Biol Chem, 218 (1956) 717-731

[62] M. Weidenhaupt, P. Rossi, C. Beck, H.M. Fischer, H. Hennecke, Arch Microbiol, 165 (1996) 169-178

[63] J.F. Turrens, A. Boveris, Biochem J, 191 (1980) 421-427

[64] C. He, S. Murthy, M.L. McCormick, D.R. Spitz, A.J. Ryan, A.B. Carter, J Biol Chem, 286 (2011) 15597-15607. http://dx.doi.org/10.1074/jbc.M110.187377

[65] T. Xia, M. Kovochich, M. Liong, L. Madler, B. Gilbert, H. Shi, J.I. Yeh, J.I. Zink, A.E. Nel, ACS Nano, 2 (2008) 2121-2134. http://dx.doi.org/10.1021/nn800511k

[66] Y. Takami, H. Uto, T. Tamai, Y. Sato, Y. Ishida, H. Morinaga, Y. Sakakibara, A. Moriuchi, M. Oketani, A. Ido, T. Nakajima, T. Okanoue, H. Tsubouchi, Hepatol Res, 40 (2010) 438-445. http://dx.doi.org/10.1111/j.1872-034X.2009.00615.x

[67] B. Thome-Kromer, I. Bonk, M. Klatt, G. Nebrich, M. Taufmann, S. Bryant, U. Wacker, A. Kopke, Proteomics, 3 (2003) 1835-1862. http://dx.doi.org/10.1002/pmic.200300552

[68] Z.T. Cossack, A.S. Prasad, Nutrition Research, 7 (1987) 1161-1167. http://dx.doi.org/10.1016/s0271-5317(87)80041-6

[69] Y. Gao, N.V. Gopee, P.C. Howard, L.R. Yu, J Proteomics, 74 (2011) 2745-2759. http://dx.doi.org/10.1016/j.jprot.2011.08.009

[70] K.H. Moon, B.L. Hood, P. Mukhopadhyay, M. Rajesh, M.A. Abdelmegeed, Y.I. Kwon, T.P. Conrads, T.D. Veenstra, B.J. Song, P. Pacher, Gastroenterology, 135 (2008) 1344-1357. http://dx.doi.org/10.1053/j.gastro.2008.06.048

[71] R.J. Szczesny, H. Obriot, A. Paczkowska, R. Jedrzejczak, A. Dmochowska, E. Bartnik, P. Formstecher, R. Polakowska, P.P. Stepien, Biol Cell, 99 (2007) 323-332. http://dx.doi.org/10.1042/BC20060108

[72] E.T. Goh, O.E. Pardo, N. Michael, A. Niewiarowski, N. Totty, D. Volkova, I.R. Tsaneva, M.J. Seckl, I. Gout, J Biol Chem, 285 (2010) 17065-17076. http://dx.doi.org/10.1074/jbc.M109.078782

[73] X. Wu, Y. Tan, H. Mao, M. Zhang, Int J Nanomedicine, 5 (2010) 385-399

[74] J.C. Zhou, X. Wang, M. Xue, Z. Xu, T. Hamasaki, Y. Yang, K. Wang, B. Dunn, Materials Science and Engineering: C, 30 (2010) 20-26. http://dx.doi.org/10.1016/j.msec.2009.08.003

[75] R. Kuriyama, H. Sakai, J Biochem, 76 (1974) 651-654

[76] C.M. Rico, S. Majumdar, M. Duarte-Gardea, J.R. Peralta-Videa, J.L. Gardea-Torresdey, J Agric Food Chem, 59 (2011) 3485-3498. http://dx.doi.org/10.1021/jf104517j

[77] N. Yoshida, K. Haga, T. Haga, Eur J Biochem, 270 (2003) 1154-1163

[78] L. Cicchillitti, R. Penci, M. Di Michele, F. Filippetti, D. Rotilio, M.B. Donati, G. Scambia, C. Ferlini, Mol Cancer Ther, 7 (2008) 2070-2079. http://dx.doi.org/10.1158/1535-7163.MCT-07-2370

[79] R. Milkman, J Gen Physiol, 45 (1962) 777-799

[80] G. Pfister, C.M. Stroh, H. Perschinka, M. Kind, M. Knoflach, P. Hinterdorfer, G. Wick, J Cell Sci, 118 (2005) 1587-1594. http://dx.doi.org/10.1242/jcs.02292

[81] K.H. Lee, H.S. Kim, H.S. Jeong, Y.S. Lee, Biochem Biophys Res Commun, 298 (2002) 216-224. http://dx.doi.org/http://dx.doi.org/10.1016/S0006-291X(02)02423-3

[82] S. Gao, H. Li, Y. Cai, J.T. Ye, Z.P. Liu, J. Lu, X.Y. Huang, X.J. Feng, H. Gao, S.R. Chen, M. Li, P.Q. Liu, Arch Biochem Biophys, 542 (2014) 46-55. http://dx.doi.org/10.1016/j.abb.2013.12.008

JOURNAL OF INTEGRATED OMICS
A METHODOLOGICAL JOURNAL
HTTP://WWW.JIOMICS.COM

ORIGINAL ARTICLE | DOI: 10.5584/jiomics.v5i1.185

Improved reconstitution of Trizol derived protein extracts provides high quality samples for comprehensive proteomic characterization of cell cultures

Cristina Pop[1,2]; Sabine Ameling[2]; Dhople Vishnu[2]; Felicia Loghin[3]; Uwe Völker[2]; Elke Hammer*[2]

[1]Department of Pharmacology, Physiology and Pathophysiology, Faculty of Pharmacy, University of Medicine and Pharmacy "Iuliu Hatieganu", Cluj-Napoca, Romania. [2]Department of Functional Genomics, Interfaculty Institute for Genetics and Functional Genomics, University Medicine Greifswald. [3]Department of Toxicology, Faculty of Pharmacy, University of Medicine and Pharmacy "Iuliu Hatieganu", Cluj-Napoca, Romania

Received: 20 February 2015 **Accepted:** 29 April 2015 **Available Online**: 9 June 2015

ABSTRACT

Background: The study of RNA, DNA, and protein from the same sample is a great advantage but can be challenging. Using Trizol, one can simultaneously extract RNA, DNA, and protein, leading to efficient sample use and more comprehensive analysis. Although it is used routinely for RNA extraction, the frequency of use of Trizol extracts for proteomics applications is low. The aim of our study was to evaluate the results of a simple modification to the Trizol protocol in terms of extraction and protein recovery efficacy and compatibility of the extracts with proteomics technologies in comparison to our standard extraction protocol including freeze/thaw cycles in urea/ thiourea. *Method:* We used the human airway epithelial cell line S9 and extracted proteins either with a modified Trizol protocol or by freeze/thaw cycles in 8M urea/ 2M thiourea. Extracted proteins were quantified and subjected to 1D- and 2D-gel electrophoresis, Western Blotting and LC-coupled tandem mass spectrometry analysis. *Results:* Compared to urea/ thiourea extraction, the Trizol-extracted proteins exhibited a similar protein composition and identification rate in LC-coupled tandem mass spectrometry experiments. 1D- and 2D-PAGE of Trizol-extracted proteins revealed excellent protein resolution with better coverage of proteins in the low MW range than urea/ thiourea extraction. *Conclusion:* The modified Trizol-protocol enabled excellent protein extraction from cell culture samples and high compatibility with proteomics technologies, especially with LC-tandem mass spectrometry.

Keywords: Trizol; cell culture; nano-HPLC-MS/MS; 1D PAGE; 2D PAGE; Western Blot Analysis.

Abbreviations

1D: one dimensional; **2D:** two dimensional; **DTT:** dithiothreitol; **ESI:** electro spray ionization; **FTICR:** fourier transform ion cyclotron resonance; **GAPDH:** glyceraldehyde 3-phosphate dehydrogenase; **HPLC:** high performance liquid chromatography; **LTQ:** linear ion trap; **MS:** mass spectrometry; **MS/MS:** tandem mass spectrometry; **PSMs:** peptide spectrum matches; **SDS:** sodium dodecyl sulfate; **TBS-T:** Tris-buffered saline-Tween; **UT:** urea/ thiourea ; **AP-NBT:** alkaline phosphatase-nitrotetrazolium blue/5-bromo-4-chloro-3-indolyl-phosphate.

1. Introduction

In-depth physiology and pathophysiology studies profit from complementing analyses of gene expression and protein abundance patterns of various conditions [1]. While transcriptomics provides comprehensive information on gene expression, proteomics allows direct analysis of protein levels accounting for regulation of protein stability and post translational modifications as well [2]. However, both techniques are complementing, suggesting that combined analyses offer a much more comprehensive view of the physiology or pathophysiology of the sample at the molecular level [3]. For practical reasons, RNA and protein extraction from the same sample would be ideal, ensuring the same

***Corresponding author:** Elke Hammer, University Medicine Greifswald, Department of Functional Genomics, Interfaculty Institute for Genetics and Functional Genomics, Friedrich-Ludwig-Jahn-Str. 15A,D-17475 Greifswald, Germany. Phone: +49-3834-865872, Fax: +49-3834-86795872, E-mail Address: hammer@uni-greifswald.de

physiological state to be reviewed. Trizol (Invitrogen) or Qiazol (Qiagen) are commercial solutions of ready-to-use reagents containing a monophasic solution of phenol, guanidine isothiocyanate and chloroform to isolate nucleic acids and protein from different biological sources in a multiple-step method from the same biological sample [4,5]. Trizol is being used primarily and successfully for RNA extraction and analysis because it acts as RNA stabilizing agent by non-specifically denaturing proteins and disrupting enzyme activity, including RNases, thus also yielding low abundance and labile mRNAs [6,7]. Trizol reagent is not yet routinely used, when protein profiling is the main aim of a study, mainly due to difficulties with the resolubilization of the precipitated protein [8]. Modified Trizol protocols that improve the resolubilization of proteins have been developed [8–11]. They have been applied to various protein sources for different proteomics applications. However, as they use detergents such as CHAPS or SDS, they are compatible with 2D-PAGE but not with LC-tandem mass spectrometry.

In the present study we compare a modified Trizol protocol with urea/ thiourea protein extraction applied to the extraction and solubilization of proteins from a human airway epithelial cell line as a proof-of-principle study. We demonstrate the efficacy of the Trizol-protocol for protein extraction and the high compatibility of the extracted proteins with proteomics methods like 1D- and 2D-PAGE, Western Blotting and especially ESI-LC-tandem mass spectrometry.

2. Material and Methods

2.1. Biological samples

The biological model used was the adeno12-SV40-immortalized human airway epithelial cell line S9 (ATCC® number CRL-2778) cultured in an adapted minimal essential medium (MEM; PromoCell, Heidelberg, Germany) as described earlier [12]. Cells were cultured in six culture plates and independently harvested at a cell density of approximately 5x106 cells. All subsequent processing steps and experiments were performed independently for each sample to be able to judge overall variation of the procedure. Three samples were used for Trizol (Invitrogen, Darmstadt, Germany) protein extraction (T1-3), while the other three samples (UT1-3) were lysed in urea/ thiourea as described earlier [13].

2.2. Protein extraction and quantification

2.2.1. UT protein extraction protocol

Cell culture samples were lysed separately in 1000µl 8M urea/ 2M thiourea (UT) by subjecting them to 5 cycles of freezing in liquid nitrogen and subsequent shaking (1500 rpm; 10 min; 37oC). Afterwards, high molecular weight nucleic acids were fragmented by sonication on ice three times for 3s, each with nine cycles at 80% energy using a Sonoplus (Bandelin, Berlin, Germany). Cell debris was removed by centrifugation (21000 g; 30 min; 4^{0}C) and the supernatant was collected for further analyses [13].

2.2.2. Modified Trizol protocol

Protein extraction with Trizol reagent (Invitrogen) was performed according to the manufacturer's protocol, with a particular modification in the reconstitution of the protein pellet, as previously described for protein extraction from heart biopsies [14]. In detail as illustrated in Fig. 1, samples were homogenized and cells were disrupted by pipetting up and down after adding 700 µl Trizol in each sample. Next, samples were incubated (room temperature; 5 min), chloroform was added, the vials were centrifuged (12000 g; 15 min; 4^{0}C) and the resulting upper aqueous phase, containing RNA, was aspirated and stored. DNA was precipitated by adding 210 µl 100% ethanol and sedimented by centrifugation (2000 g; 15 min; 4^{0}C). The supernatant was collected and 100% isopropanol was added for protein precipitation. After incubation (room temperature; 10 min), samples were centrifuged (12000 g; 10 min; 4^{0}C) and the supernatant was discarded. The resulting protein pellets were washed three times with 0.3M guanidine hydrochloride in 95% ethanol, each step being followed by centrifugation (7500 g; 5 min; 4^{0}C). Finally, protein pellets were washed with 100% ethanol and left to air dry for 5-10 min, carefully avoiding extensive (over-)drying. To ensure maximum protein reconstitution, 400 µl UT was added to each sample with multiple dispensing/aspirating cycles, followed by shaking (800 rpm; 20-40 min; 20^{0}C). Samples were stored at -80^{0}C until further use.

A Bradford assay kit (Pierce, Thermo Scientific, Bonn, Germany) and bovine serum albumin as standard protein [15] were used for determination of protein concentration in the samples.

2.3. 1D SDS-PAGE and Western Blotting

Protein samples (20µg) were resolved on 12.5% SDS (sodium dodecyl sulphate) polyacrylamide gels and the patterns were visualized by staining with Coomassie brilliant blue R-250. For specific protein detection (Western Blot analysis), proteins were transferred from the gel onto a 0.45µm pore diameter PVDF (polyvinilidene fluoride) membrane using a semidry Milliblot apparatus (Merck Millipore, Billerica, MA, USA). To control blotting efficiency on the membrane, proteins were visualized using ink in 1% acetic acid and TBS-T (Tris-buffered saline-Tween containing 20mM Tris-HCl, 137 mM NaCl and 0.1% Tween-20). After visualization, ink was removed with TBS-T, the membrane was blocked with 5% powdered milk for 90 min and incubated with the primary antibody mouse anti-α-GAPDH (Cell Signaling Technology, Boston, Massachusetts, USA; dilution 1:50000) over night at 4^{0}C. Detection was

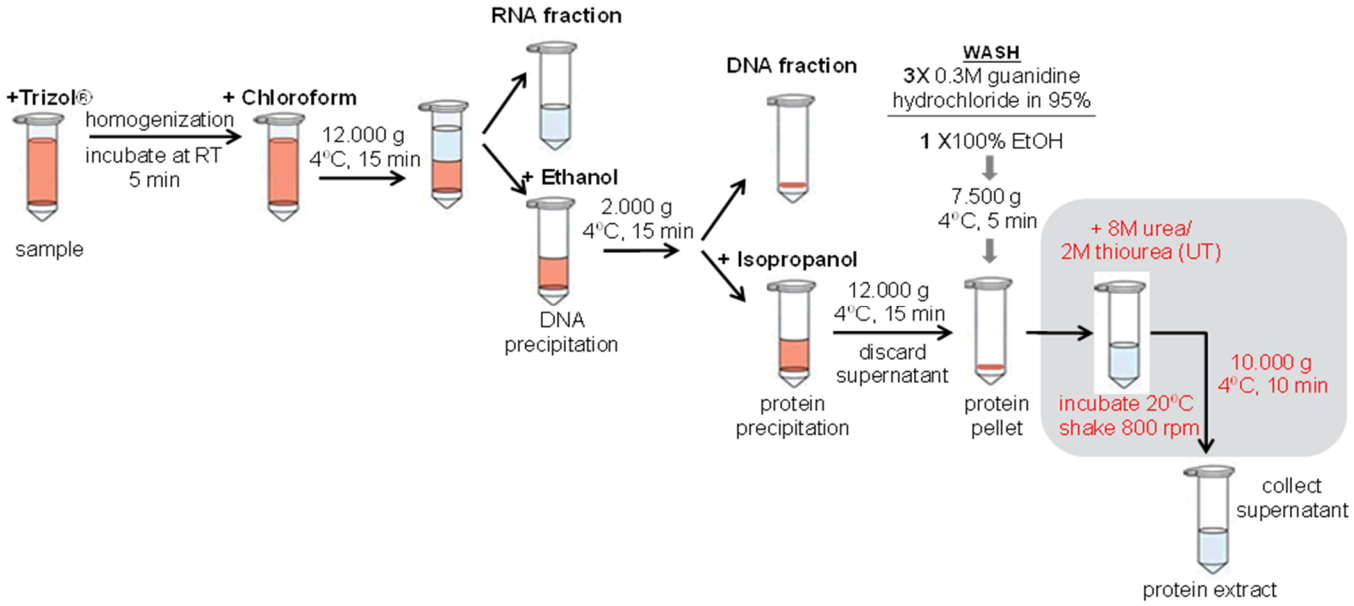

Fig. 1 Modified Trizol protocol for mass spectrometry-compatible protein extraction. The protocol retains the steps suggested by the manufacturer for preparation of RNA-and DNA-fractions, but uses modified steps for an improved reconstitution of the protein pellet in 8M urea/2M thiourea (UT). The modification is highlighted by blue background.

performed after incubation with alkaline phosphatase conjugated goat anti-rabbit IgG (Biorad, Munich, Germany, 1:5000) as secondary antibody for 60 min using the AP-NBT/BCIP in situ detection system.

2.4. 2D SDS-PAGE

Isoelectric focusing was performed using 7cm immobilized pH gradient (IPG) strips (Bio-Rad) with a pI range of 3-10. Strips were loaded with 30 µg proteins in rehydration buffer (8M urea, 2M thiourea, 2% CHAPS, 30mM DTT, 2% pharmalyte and bromophenol blue) and subjected to isoelectric focusing. As described previously [16], equilibration buffers were used for reduction and alkylation of proteins on the strips. Proteins were separated in the second dimension as previously described [17] on 12.5% SDS polyacrylamide gels in low fluorescent glass plates. Finally, Coomassie brilliant blue R-250 staining was used for protein spots visualization.

2.5 Mass spectrometry analysis

Mass spectrometry analysis was performed on three biological replicates for each extraction method (UT and Trizol), using the LTQ-FTICR mass spectrometer (Thermo Scientific, Bremen, Germany) after pre-fractionation of peptides by reverse phase nano-UPLC (Waters, Manchester, U.K.). In total, 4µg protein from each sample was first reduced (2.5mM dithiothreitol; 1 h; 60°C), then alkylated (10 mM iodoacetamide; 30 min; 37°C) and subsequently digested with trypsin (Promega, Madison, USA) in a 1:10 ratio (overnight – 16h; 37°C) as previously described [16]. Proteins were identified using the SEQUEST algorithm with Proteome discoverer 1.3 (Thermo Scientific). MS spectra were searched against a UniProt Swiss-Prot database (rel. 2010_11) limited to human entries with a mass tolerance of 10 ppm for peptide identifications and 0.6 Da fragment tolerance. Methionine oxidation was set as variable, carbamidomethylation at cysteine as fixed modification and up to two missed tryptic cleavages were considered (for details see Supplemental Table S1).

3. Results and discussion

Various modifications of the original Trizol protocol have been reported (Table 1). The majority of them implemented detergents in order to improve protein recovery efficiency for subsequent global protein profiling by 2D-PAGE. Predominantly detergents like CHAPS or SDS were considered to improve protein reconstitution [9,10,18–21]. However, such detergents are incompatible with the nano-HPLC coupled ESI-LC tandem mass spectrometry, except when particular sample clean-up steps are applied. Therefore, here we present the results of a simple modification of the Trizol protocol avoiding the use of detergents but enhancing protein recovery by reconstituting the protein pellet not only in the denaturing chemical urea (8M) but in the presence of 2M thiourea which especially supports the resolubilization of hydrophobic proteins [22]. Thiourea is a non-chaotropic compound, which has been frequently used in 2D-PAGE applications due to its high capacity to re-solubilize membrane proteins [20,22]. Thus, addition of UT followed by incubation at room temperature and shaking at 800 rpm resulted in a rapid and almost

Table 1 Published Trizol protocols modified with the aim of improving protein recovery.

Re-suspension of protein pellets in :	Author	Ref.
1:1 solution of 1% SDS and 8 M urea in Tris-HCl 1 M, pH 8.0, followed by 5 cycles of 15 sec sonication and 30 sec ice incubation	Simões et al.	[10]
7 M urea, 2 M thiourea, 4% (w/v) Chaps, 1 mM phenylmethane sulfonyl fluoride and 30 mM Tris-HCl, pH 8.5	Yamaguchi et al.	[20]
8 M urea, 4% (w/v) Chaps and 2% (w/v) DTT, followed by sonication (10 min, 4°C) and incubation at room temperature for 2 h	Xiong et al.	[19]
1% SDS, followed by incubation at 50°C for longer than 10 min with intermittent vortexing	Likhite et al.	[9]
8 M urea	Ham et al.	[21]
9.5 M urea and 2% (w/v) Chaps, pH 9.1 or 10% acetonitrile, pH 4.8 or 1% triton, pH 5.3	Man et al.	[11]
250 mM glycerol, 10 mM triethanolamine and 4% (w/v) Chaps	Kirkland et al.	[18]
Sonication in methanol and reconstitution of the powder in 0.2% Rapigest	Kline et al.	[26]

complete dissolution of the protein pellet. Due to the strong denaturing conditions used, proteins lose their native conformation and cannot be used for studies of natural activity. However, they are well suited for proteomics studies, including protein quantitation.

3.1 Protein extraction, reconstitution and quantitation

Standard extraction of proteins from 5 x 10^6 S9 cells was performed with 1000µl UT. In contrast, 400 µl UT were used for the reconstitution of the pellet obtained by precipitation of protein with Trizol. Due to the lower volume, protein concentrations of the Trizol derived protein extracts were similar to UT protein extraction (T1=1.75µg/µl, T2=1.37µg/µl, T3=0.67µg/µl; UT1=1.15µg/µl, UT2=0.98µg/µl, UT3=2.32 µg/µl). The total amount of protein extracted with UT was larger compared to Trizol extraction (T1=698.2µg, T2=547.32µg, T3=268.11µg and UT1=1150µg, UT2=983µg, UT3=2324.85µg). Lower protein yield with Trizol protein extraction was also previously reported [20]. However both extraction methods yielded sufficient protein for further proteomics and biochemical analysis methods.

3.2 Resolution and antigenic stability testing of the proteins extracted

Separation of all protein extracts on a 1D-gel revealed a similar complex protein pattern for both extraction methods (Fig. 1A). Minor differences were only observed in the staining intensity of particular protein bands in the low molecular weight (MW) range. Previous reports indicated that Trizol extraction might be more efficient for proteins with low-MW in comparison to other methods of extraction [20,23]. In order to assess if higher amounts of low MW proteins are indeed accessible after Trizol extraction or if the increased band intensity (Fig. 2) resulted from degradation of high MW proteins, the low MW regions were cut from the gel and subjected to in-gel-tryptic digestion and mass spectrometry. However, no differences in the spectral counts per protein between the UT and Trizol derived protein extracts and no indication of increased levels of degradation products were found. Thus, the stronger intensity of bands of low MW proteins after Trizol extraction is likely caused

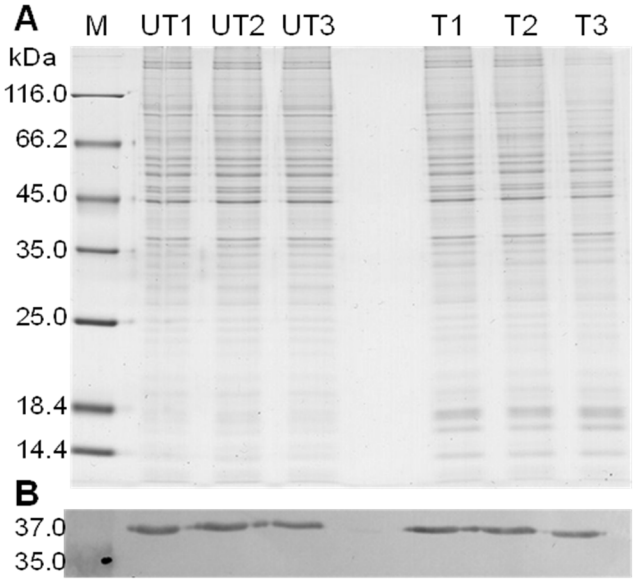

Fig. 2 (A) Visualization of UT and Trizol extracted proteins on a 1D 12.5% SDS polyacrylamide gel by Coomassie staining; (B) Determination of the antigenic stability of the extracted proteins through Western Blotting for the presence of α-GAPDH (at 37kDa). (M-protein marker)

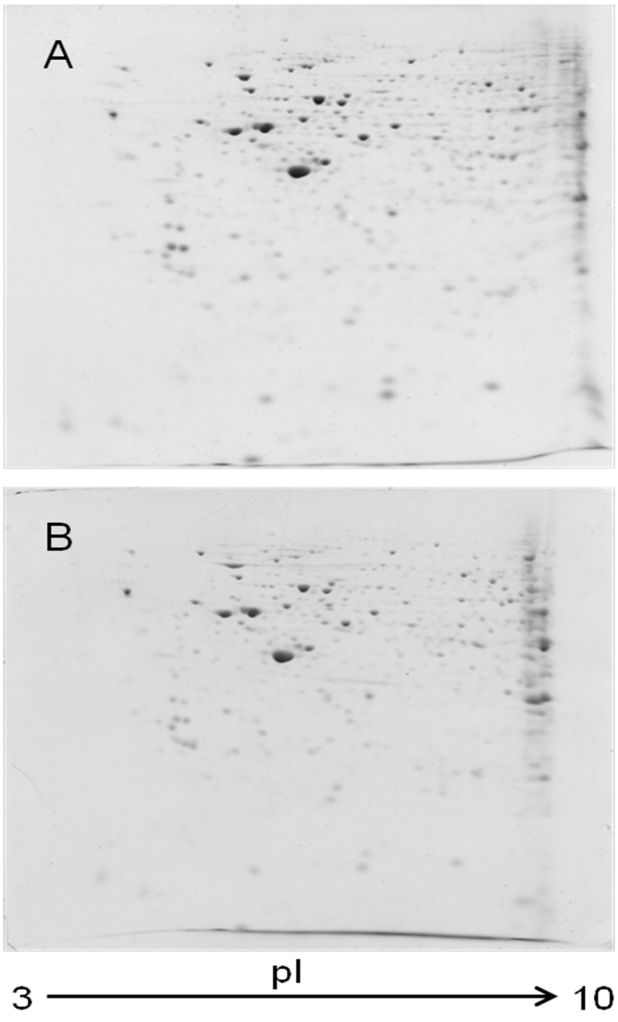

Fig. 3 Representative 2D-PAGE gel images of the proteins extracted using (A) Trizol and (B) UT which show similar protein patterns, good resolution of proteins, and more spots, with a higher intensity, especially in the low MW range for the Trizol-extracted proteins.

by more efficient protein staining.

Since an artificial Trizol-triggered modification of proteins that might lead to a masking of epitopes has previously been discussed [19], the suitability of Trizol-prepared protein extracts for immunoblotting was exemplarily tested for α-GAPDH. Fig. 1B shows α-GAPDH identified at approximately 37kDa, demonstrating preservation of epitope recognition of the proteins extracted and the lack of influence of the extraction method on α-GAPDH signal intensity. Other studies have also investigated the stability of the proteins extracted with Trizol by using PAGE and Western blot analysis, with similar results [8,9,20].

Representative gels of 2D-gel electrophoresis performed with Trizol and UT protein extracts from S9 cell line are presented in Fig. 3. Our modified Trizol extraction method revealed good resolution of proteins, presenting a similar protein pattern as it was detected on the gels in which proteins of the UT extracts were separated. Thus, our data contradict those of Xiong and colleagues, who reported spot chains, smears or diffuse spots or even potential protein degradation as possible consequences of artificial protein modifications in the presence of Trizol [19]. In agreement with our data, comparable good results obtained after conventional urea extraction were reported for 2D-PAGE experiments with Trizol derived extracts [20]. Again, 2D gels of Trizol extracts yielded more spots, with a higher intensity, especially in the low MW range. This effect was hypothesized to occur as a result of the high efficiency of removal of nucleic acids, lipids, carbohydrates and salts [18]. Such a high purity of the protein extracts can be a significant advantage for different protein enrichment methods such as immobilized metal affinity chromatography (IMAC) for phosphopeptides, potentially providing an increased sensitivity [24,25].

3.3 Protein identification by ESI nano-HPLC-MS

In order to analyze the samples by bottom-up proteomics we first trypsinized the protein extracts and interpreted the raw data obtained from nano-HPLC-MS/MS using Proteome Discoverer (Supplemental Table S2). Overall, the number of peptide spectrum matches (PSMs) and peptides found in the Trizol samples was similar to the ones detected in the UT samples (PSMs: T1=4339, T2=4432; T3=3389; UT1=4295, UT2=4432, UT3=4327p=0.46, and peptides: T1=3117, T2=3108, T3=2292; UT1=3074, UT2=3156, UT3=3090, p=0.43). From the identified peptides, we inferred a similar number of proteins in the Trizol and the UT samples (proteins identified based on at least one proteospecific peptide: Trizol=814; UT=798; proteins identified based on at least two proteospecific peptides: Trizol=459; UT=507). However, except for sample T3, Trizol extraction revealed a higher number of protein identifications compared to UT. Improved protein identification was also reported by others, recommending the use of Trizol to the mapping of whole proteomes [21,26], with the advantage that Trizol can simultaneously extract RNA, DNA, and protein from the same sample leading to sample economy, which is especially useful when dealing with small and precious samples (biopsies, sorted cells etc.) [8,14].

Overall, there was a 77.49 % overlap between the proteins identified using UT or Trizol (Supplemental Figure 2). For both protein extraction methods, we found an 80% overlap between the proteins identified in the three bioreplicates used which is in the range of the technical variance across the replicates.

To assess the similarity of the protein profiles between the two extraction methods, proteins were classified using Protein Center (Figure 3). Top cellular components covered by the extracted proteins were very similar for the two extraction methods - cytoplasm (T=23.63% /UT=23.88%),

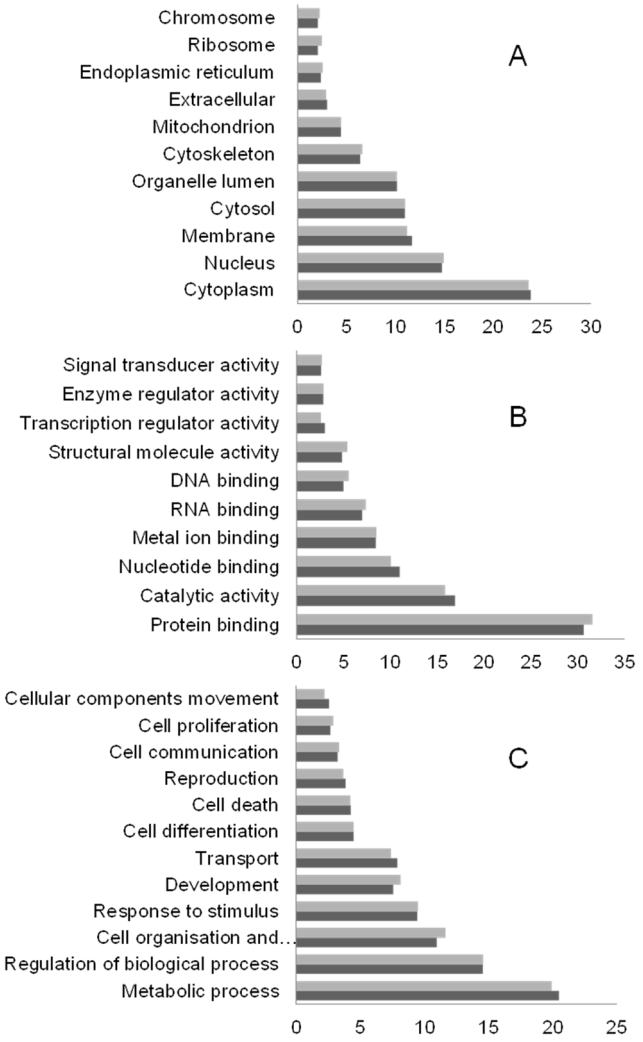

Fig. 4 ProteinCenter assignment of proteins to the corresponding (A) cellular components, (B) molecular functions and (C) biological processes covered by the proteins extracted either by UT (dark bars) or by Trizol (grey bars).

nucleus (14.94% / 14.78%) and membrane (11.19% / 11.73%). Likhite et al. identified the inability to analyze nuclear proteins as a major limitation of the Trizol method [9], but it seems that our modification to the Trizol protocol also improved coverage of the nuclear protein fraction. The presence of membrane proteins in the Trizol samples shows that the modification to the Trizol manufacturer's protocol led to the improvement of the reconstitution of membrane proteins, which sometimes poses difficulty during extraction and dissolution. Top molecular functions of the identified proteins were protein binding (T=31.59%, UT=30.67%), catalytic activity (15.85% / 16.92%) and nucleotide binding (10.04% / 11.01%) and top biological processes covered by the identified proteins were metabolic (19.92% / 20.5%), regulation of biological processes (14.58% / 14.56%) and cell organization and biogenesis (11.64% / 10.97%) and all were very similarly covered by both extraction methods. Molecular functions and biological processes were also similarly covered by the proteins extracted by both methods.

4. Concluding Remarks

In the present study we have performed a Trizol protein extraction from a cell culture and compared it to traditional urea/ thiourea lysis buffer extraction resulting in comparable quality of the proteins fractions recovered by the two methods, with the major advantage that the Trizol protocol also enables simultaneous RNA and DNA extraction from the same sample.

We have demonstrated that sufficient amounts of protein for further analysis can be extracted from S9 cells for further proteomics applications using a modified Trizol protocol. Our study highlights that the modified Trizol extraction allows rapid protein extraction with minimal protein degradation by proteolysis and yields highly pure protein extracts, compatible with many types of protein analysis techniques such as 1D-, 2D-PAGE, Western blot analysis and most importantly LC-tandem mass spectrometry.

5. Supplementary material

Supplementary data and information is available at: http://www.jiomics.com/index.php/jio/rt/suppFiles/185/0

Supplementary Table S1 LC-MS/MS parameters and presentation of proteomic data

Supplementary Table S2 List of all proteins extracted from human airway epithelial cell line S9 using Trizol and urea/ thiourea

Supplementary Fig. S1 Venn diagram depicting the protein overlap for the merged results from 3 biological replicates for UT and Trizol extraction methods.

Acknowledgements

The present study has partially been conducted with the support of the POSDRU grant 107/1.5/S/78702 from the University of Medicine and Pharmacy "Iuliu Hatieganu", Cluj-Napoca, Romania.

References

[1] H. Tanaka, S. Ogishima, Methods Mol Biol, 719 (2011) 499-509. doi: 10.1007/978-1-61779-027-0_23

[2] P. Mallick, B. Kuster, Nat Biotechnol, 28 (2010) 695-709. doi: 10.1038/nbt.1658

[3] I. Gonzalez, K.A. Cao, M.J. Davis, S. Dejean, BioData Min, 5 (2012) 19. doi: 10.1186/1756-0381-5-19

[4] P. Chomczynski, N. Sacchi, Anal Biochem, 162 (1987) 156-159. doi: 10.1006/abio.1987.9999

[5] P. Chomczynski, N. Sacchi, Nat Protoc, 1 (2006) 581-585. doi: 10.1038/nprot.2006.83

[6] C. Yu, S. Young, V. Russo, B.G. Amsden, L.E. Flynn, Tissue Eng Part C Methods, 19 (2013) 829-838. doi: 10.1089/ten.TEC.2012.0693

[7] M. Remáková, M. Škoda, M. Faustová,.J. Vencovský, P. Novotá, Folia Biol(Praha) 59 (2013) 47–50

[8] A.B.Hummon, .S.R.Lim, M.J. Difilippantonio, T. Ried, Bio Techniques 42 (2007) 467–470 doi: 10.2144/000112401

[9] N.Likhite, U.M. Warawdekar, J Biomol Tech 22 (2011) 37-44

[10] A.E. Simoes, D.M. Pereira, J.D. Amaral, A.F. Nunes, S.E. Gomes, P.M. Rodrigues, A.C. Lo, R. D'Hooge, C.J. Steer, S.N. Thibodeau, P.M. Borralho, C.M. Rodrigues, BMC Genomics, 14 (2013) 181. doi: 10.1186/1471-2164-14-181

[11] T.K. Man, Y. Li, T.A. Dang, J. Shen, L. Perlaky, C.C. Lau, Proteome Sci, 4 (2006) 3. doi: 10.1186/1477-5956-4-3

[12] F. Schmidt, S.S. Scharf, P. Hildebrandt, M. Burian, J. Bernhardt, V. Dhople, J. Kalinka, M. Gutjahr, E. Hammer, U. Volker, Proteomics, 10 (2010) 2801-2811. doi: 10.1002/pmic.201000045

[13] E. Hammer, S. Bien, M.G. Salazar, L. Steil, C. Scharf, P. Hildebrandt, H.W. Schroeder, H.K. Kroemer, U. Volker, C.A. Ritter, Proteomics, 10 (2010) 99-114. doi: 10.1002/pmic.200800626

[14] C. Pop, S. Ameling, K. Empen, J. Rudebusch, V.M. Dhople, S.B. Felix, F. Loghin, U. Volker, E. Hammer, Clin Chim Acta, 438 (2015) 246-247. doi: 10.1016/j.cca.2014.08.033

[15] M.M. Bradford, Analytical Biochemistry, 72 (1976) 248-254. doi: 10.1016/0003-2697(76)90527-3

[16] K. Nishtala, T.Q. Phong, L. Steil, M. Sauter, M.G. Salazar, R. Kandolf, H.K. Kroemer, S.B. Felix, U. Volker, K. Klingel, E. Hammer, Proteomics, 11 (2011) 4310-4320. doi: 10.1002/pmic.201100229

[17] T. Thiele, L. Steil, S. Gebhard, C. Scharf, E. Hammer, M. Brigulla, N. Lubenow, K.J. Clemetson, U. Volker, A. Greinacher, Transfusion, 47 (2007) 1221-1233. doi: 10.1111/j.1537-2995.2007.01255.x

[18] P.A. Kirkland, J. Busby, S. Stevens, Jr., J.A. Maupin-Furlow, Anal Biochem, 351 (2006) 254-259. doi: 10.1016/j.ab.2006.01.017

[19] J. Xiong, Q. Yang, J. Kang, Y. Sun, T. Zhang, G. Margaret, W. Ding, Electrophoresis, 32 (2011) 321-330. doi: 10.1002/elps.201000425

[20] H. Yamaguchi, K. Hasegawa, M. Esumi, Exp Mol Pathol, 95 (2013) 46-50. doi: 10.1016/j.yexmp.2013.05.002

[21] B.M. Ham, F. Yang, H. Jayachandran, N. Jaitly, M.E. Monroe, M.A. Gritsenko, E.A. Livesay, R. Zhao, S.O. Purvine, D. Orton, J.N. Adkins, D.G. Camp, 2nd, S. Rossie, R.D. Smith, J Proteome Res, 7 (2008) 2215-2221. doi: 10.1021/pr700575m

[22] T. Rabilloud, Electrophoresis, 19 (1998) 758-760. doi: 10.1002/elps.1150190526

[23] D.A. MacIntyre, R. Smith, E.C. Chan, Anal Biochem, 359 (2006) 274-276. doi: 10.1016/j.ab.2006.07.044

[24] K. Moser, F.M. White, J Proteome Res, 5 (2006) 98-104. doi: 10.1021/pr0503073

[25] F. Yang, D.L. Stenoien, E.F. Strittmatter, J. Wang, L. Ding, M.S. Lipton, M.E. Monroe, C.D. Nicora, M.A. Gristenko, K. Tang, R. Fang, J.N. Adkins, D.G. Camp, 2nd, D.J. Chen, R.D. Smith, J Proteome Res, 5 (2006) 1252-1260. doi: 10.1021/pr060028v

[26] K.G. Kline, B. Frewen, M.R. Bristow, M.J. Maccoss, C.C. Wu, J Proteome Res, 7 (2008) 5055-5061. doi: 10.1021/pr800239e