Volume 3 | Issue 1 | 21 June 2013 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

Food Research Division, Bureau of Chemical Safety, Health Canada, Ottawa, Canada

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 3 | ISSUE 1 | JUNE 2013

REVIEW ARTICLES

ORIGINAL ARTICLES

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

REVIEW ARTICLE | DOI: 10.5584/jiomics.v3i1.120

-Omics fields of study related to plant-parasitic nematodes

Juan E. Palomares-Rius[1,2,*] and Taisei Kikuchi[1,3]

[1]Department of Forest Pathology, Forestry and Forest Products Research Institute (FFPRI) Tsukuba 305-8687, Japan; [2]Institute for Sustainable Agriculture (IAS), Spanish National Research Council (CSIC), Apdo. 4084, 14080 Córdoba, Campus de Excelencia Internacional Agroalimentario, ceiA3, Spain; [3]Division of Parasitology, Fuculty of Medicine, University of Miyazaki, Miyazaki 889-1692, Japan.

Received: 09 January 2013 **Accepted:** 07 March 2013 **Available Online**: 12 March 2013

ABSTRACT

Plant-parasitic nematodes (PPN) cause significant losses and these pathogens must be addressed amid the growing demand for food, global warming, and the discarded use of inorganic pesticides. For these reasons, acquiring deeper knowledge about PPN and devising new management strategies are important in order to meet future food demand. This review focuses on PPN and their applicable and diverse –omics fields of study. While most efforts have been centered on transcriptomics, other –omics studies have recently begun to expand. The few genomes sequenced (Meloidogyne incognita, M. hapla, and Bursaphelenchus xylophilus) have shown high diversity in PPN. This review also discusses the future prospects and uses of –omics relative to PPN.

Keywords: Plant-parasitic nematodes; Disease; Control; Management; -Omics .

Abbreviations:

PPN, Plant-Parasitic Nematodes; **PCN**, Potato Cyst Nematodes; **NGS**, Next Generation Sequencing; **RNAi**, RNA interference; **dsRNA**, Double-stranded RNA; **EST**, Expressed Sequenced Tag; **HGT**, Horizontal Gene Transfer.

1. Importance of plant-parasitic nematodes in agriculture

Plant-parasitic nematodes (PPN) cause an estimated crop yield loss of 14.6% in tropical and sub-tropical climates and losses of 8.8% in developing countries (1). These losses are estimated to exceed 100 billion US dollars (2; 3). Thus, PPN have a significant impact on food security and our ability to feed a growing human population (35% increase by 2050, 4) in the coming years. Other estimates project a 75% increase in food demand between 2010 and 2050, including changes in diet (toward consuming more protein) and steady population growth (5).

Plant-parasitic nematodes are considered the "unseen enemies" of plants because the symptoms seen in the aerial parts of plants are generally associated with forms of abiotic stress (e.g., lack of nitrogen, water stress). Diseased plants are usually found in patched patterns in the field (Fig. 1). Aside from the lack of specific symptoms, it is difficult to detect PPN, which are small, soil dwelling organisms. Nematodes can affect crops by directly feeding on plants through the stylet (a protrusible, syringe-needle-like structure), disrupting plant physiology through the growth of plant-specific structures, enabling secondary infection by opportunistic pathogens (bacteria and fungi) or, in some cases, transmitting plant viruses. The damage caused by PPN largely depends on the type of crop, its stage of development, and edaphic/climatic conditions. Table 1 lists the specific damage caused to crops in detail. Some genera are of major importance as plant parasites (e.g. *Meloidogyne*, *Globodera*, *Heterodera*, *Pratylenchus*, *Ditylenchus*, *Aphelenchoides*, *Bursaphelenchus*, *Xiphinema*, *Trichodorus*), while others are of minor importance (given their limited crop damage or low numbers in the soil). Some PPN are highly polyphagous (e.g.. *Meloidogyne* spp., *Pratylenchus* spp.) and can infect many species of plants, while others are species or genus specific (e.g. *Globodera* spp., *Heterodera* spp.).

***Corresponding author:** Juan E. Palomares-Rius, Department of Forest Pathology, FFPRI1 Matsunosato, Tsukuba, Ibaraki, 305-8687, Japan. Phone number: +81-29-873-3211; Fax number: +81-29 874 3720; E-mail address: palomaresje@ias.csic.es

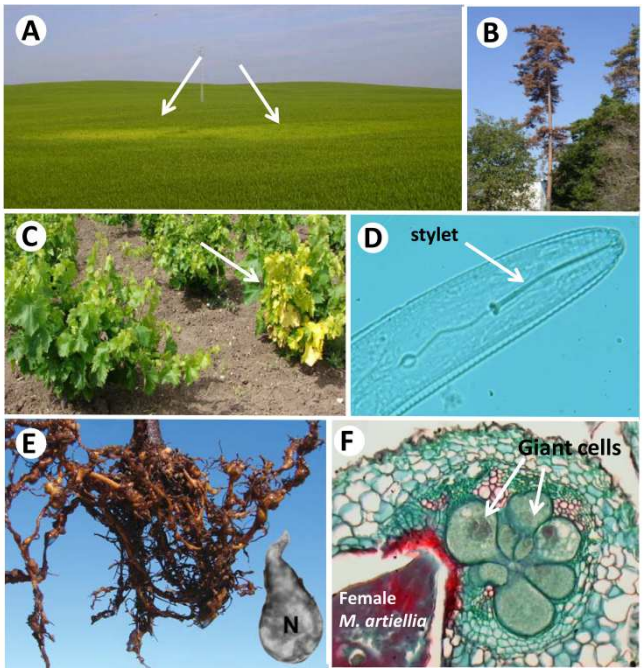

Fig. 1. Plant-parasitic nematodes and their damage to plants. A) Patchy distribution of symptomatic plants (yellow and reduced growth) in a wheat field caused by Heterodera avenae; B) pine tree death by Bursaphelenchus xylophilus; C) Grapevine Fan Leaf virus in grapevine transmited by Xiphinema index; D) Anterior part of Hemicycliophora sp. showing the stylet; E) Grapevine rootstock (41 B Millardet et De Grasset) infected by the root-knot nematodes Meloidogyne incognita race 1, in detail: a female extracted from roots; F) Cross section of M. artiellia chickpea infected root showing the typical feeding site with giant cells.

Plant-parasitic nematodes can be separated based on their life strategy into several groups:

i) ectoparasites: nematodes remaining in the soil and not entering plant tissue; ii) endoparasites: nematodes that fully penetrate root tissue; and iii) semi-endoparasites: nematodes that penetrate roots with only their anterior part, with their posterior part remaining in the soil phase. This division becomes more complicated with the presence of migratory or sedentary life cycles. Nematodes also serve important roles in the flow of energy in the soil and are important in other ecological niches (such as bacterial and fungal feeders, predatory and omnivorous). Some PPN, primarily the sedentary endoparasites (*Meloidogyne* spp., *Heterodera* spp., *Globodera* spp.), have a complicated relationship with their host. They induce the re-differentiation or fusion of root cells into "specialized feeding sites" (Fig 1) from which they feed continuously. Feeding sites are thought to be produced as a result of the action of proteins and/or small molecules (*e.g.* hormones) secreted into the host via the stylet. Other secretions from the nematodes (via cuticles, amphids and rectal glands) may also play an important role in plant parasitism. The current trend toward discontinuing the use of inorganic pesticides due to environmental and human health concerns, and the effects of climate change which may increase the geographical range of some nematodes, are predicted to increase the damage caused by PPN (1). For these reasons, greater and more detailed knowledge about the causal agents of disease and the life cycle (e.g., development, survival, interaction with host, virulence, pest resistance) of PPN are essential for devising new management strategies. Omics studies will offer new ways of obtaining the knowledge necessary for achieving these objectives.

2. –Omics fields of study related to plant-parasitic nematodes

Plant-parasitic nematodes are intimately associated with their hosts and are highly adapted to this lifestyle. Studies conducted on the model nematode (*Cænorhabditis elegans*) have contributed enormously towards advancing our understanding of the basic biology of PPN. However, studies of PPN are far more difficult than those on *C. elegans* as several features of their life cycle make them extremely difficult to work with. These include: i) different specific development stages that exist only inside the roots (endoparasitic nematodes); ii) difficulty to culture the nematodes as most PPN are biotrophic organisms; and iii) high diversity in some species. Moreover, many PPN studies are dictated by the importance of certain groups of nematodes over the rest. Studies tend to be focused on the most damaging PPN (i.e., sedentary endoparasites *Meloidogyne* spp., *Globodera* spp., *Heterodera* spp.). However, studies on other nematode groups will expand as more genomes and sets of genes are deposited in public databases.

2.1. *Genomics*

Whole genome sequencing yields access to all the genes, avoiding the difficulty of finding genes expressed at low levels by using other strategies (such as ESTs). However, the gene catalogue from whole genome sequencing may still be incomplete, as some regions remain difficult to sequence and certain transcripts could be difficult to predict based on the genomic sequence alone (6). Second generation of sequencing technologies collectively known as "Next Generation Sequencing" (NGS) has emerged in recent years including: 454 sequencing (Roche, Branford, CT, USA), Illumina sequencing (Illumina Inc., San Diego, CA, USA), and sequencing oligonucleotide ligation and detection (SOLiD) (Applied Biosystems, Carlsbad, CA, USA) (7). These new technologies offer advantages over Sanger sequencing in terms of i) reduced cost per DNA base sequenced; ii) speed at which large volumes of data are generated; iii) an ability to work with less starting material, and iv) no need to go through a cloning vector or host organism (7). However, challenges remain when using NGS including issues with the assembly process and data storage. In addition, specific problems are associated with some technologies, such as the unreliable determination of homopolymer regions and large repeats in 454 sequencing technology (6). Nevertheless, several programs are available for the assembly process and data management (see reviews by 6 and 7).

Table 1. International survey of crop losses due to nematodes based on 371 responses (Source: 1).

Life-sustaining Crops	Annual Loss (%)	Economically-important Crops	Annual Loss (%)
Banana	19.7	Cacao	10.5
Barley	6.3	Citrus	4.2
Cassava	8.4	Coffee	15.0
Chickpea	13.7	Cotton	10.7
Coconut	17.1	Cowpea	15.1
Corn	10.2	Eggplant	16.9
Field bean	10.9	Forages	8.2
Millet	11.8	Grape	12.5
Oat	4.2	Guava	10.8
Peanut	12.0	Melons	13.8
Pigeon pea	13.2	Misc. other	17.3
Potato	12.2	Okra	20.4
Rice	10.0	Ornamentals	11.1
Rye	3.3	Papaya	15.1
Sorghum	6.9	Pepper	12.2
Soybean	10.6	Pineapple	14.9
Sugar beet	10.9	Tea	8.2
Sugar cane	15.3	Tobacco	14.7
Sweet potato	10.2	Tomato	20.6
Wheat	7.0	Yam	17.7
Average	10.7%	**Average**	14.0%
	Overall Average 12.3%		

Caenorhabditis elegans was the first genome of a multicellular organism to be completely sequenced, thereby providing a platform for further nematode genomics (*C. elegans* Sequencing Consortium, 1998). Nematode genomes analysed to date range from 15 megabases (*Pratylenchus* spp.) up to 0.5 gigabase (in some estimates of the *Ascaris suum* genome), and were organized from as few as one chromosome (*Parascaris univalens*) to many tens (in certain *Meloidogyne* spp.) (6). Representative species across the diversity of the phylum Nematoda is very important for understanding the molecular and ecological aspects of nematodes. To date, three species of PPN including two from the genus *Meloidogyne* (sedentary endoparasites), *M. incognita* (8), *M. hapla* (9), and *Bursaphelenchus xylophilus* (migratory endoparasites) (10) have been sequenced and published. Further plant nematode genome sequencing programmes are in progress including, *Bursaphelenchus mucronatus*, *Ditylenchus destructor*, *Globodera pallida*, *Globodera rostochiensis*, *Meloidogyne arenaria*, *Heterodera glycines*, *Meloidogyne floridensis*, and *Meloidogyne javanica* (www.nematodes.or). Information on with the progress of some of these sequencing projects is now available on various websites. However, other important PPN not related to the order Rhabditida are not yet being sequenced. These include nematodes in the orders Enoplea (i.e., genera *Xiphinema*, *Longidorus*, *Paralongidorus*) and Triplonchida (i.e., genera *Trichodorus*, *Paratrichodorus*).

The genome sequences obtained to date have showed important differences between them, even in species from the same genus (e.g., *M. incognita*, *M. hapla*). These differences may reflect the important biological differences that exist between these species. *Meloidogyne incognita* is a parthenogenetic and polyphagous species, while *M. hapla* is capable of sexual reproduction and has a narrower host range (11). Many isolates of *M. incognita* are subject to extensive polyploidy and/or aneuploidy, while *M. hapla* has a meiotic reproduction lifestyle, which helps make controlled crossings possible (11). The main differences between both genomes are the very small size of *M. hapla*; in contrast, the *M. incognita* genome has shown the genetic consequences of reproduction by asexual mitosis, as demonstrated by the high sequence divergence between aligned regions (11). The num-

ber of proteins predicted in the two RKN genomes is also different (14,454 and 19,212 for *M. hapla* and *M. incognita*, respectively). This is due to the duplications that have occurred within the *M. incognita* genome (11).

The pine wood nematode, *B. xylophilus* is a migratory endoparasite that causes severe damage to forest ecosystems. Several species of pine trees (mainly outside the area of co-evolution with pines in North America) are susceptible to *B. xylophilus*. *Bursaphelenchus xylophilus* has a complex ecology that combines fungal feeding and the plant-parasitic/insect-associated stages. The genome of this species contains 18,074 genes distributed across six chromosomes. The assembled genome showed a G+ C content of 40.4%, which is higher than that of both *M. incognita* (31.4%) and *M. hapla* (27.4%). This nematode showed great expansion in terms of digestive and detoxification proteins, which may reflect an unusual diversity in the foods consumed and the environments encountered during its life cycle (10). In addition, *B. xylophilus* has the largest number of digestive proteases known for any nematode, and shows expanded families of lysosome pathway genes, ABC transporters, and cytochrome P450 pathway genes (10). *Bursaphelenchus xylophilus* sequences that matched those of parasitism genes (except cell wall degrading enzymes) either did not have the predicted signal peptide or, if one was predicted, homologues were also present in a wide range of other species including *C. elegans* and animal PPN (10). Two exceptions were found—venom allergen proteins and a putative cystein protease inhibitor. These findings are consistent with the independent evolution of plant parasitism within *Bursaphelenchus* compared to other nematodes (e.g., *Meloidogyne* and *Heterodera*, *Globodera* spp.,). In addition, *B. xylophilus* is a migratory endoparasite that is not biotrophic (10).

These analysis of genomes from PPN have led to the identification of novel secreted proteins that could be important parasitism genes (e.g., cell wall-degrading enzymes, those involved in modulation of the plant's defense system, those important for establishment of nematode feeding sites, and those required for synthesis or processing of nutrients). Some of these identified parasitism genes were acquired by horizontal gene transfer (HGT) from bacteria or fungi (12). In addition, these genomes have given researchers important tools for understanding parasitism and have led to the identification of genes that could be good targets for nematode control and basic biological research (reviewed by 13).

The sequenced genomes of other nematodes having different ways of life (e.g., free moving bacterial feeder, bacterial-fungus-nematode feeder, necromenic and animal parasitic) and/or different life-related vectors take different genomic approaches to deal with their environments. In some cases, such unexpected results as cellulases or diapausin have been found in the genome of *Pristionchus pacificus*. These cellulases exhibited a probable and different functional role (probably used for biofilm degradation of certain bacteria), and origin by HGT (amoebozoa) than cellulases in PPN (14). Moreover, diapausin and other genes (509 genes in total) have insect-like codon usage more akin to insects than nematodes (15). For this reason, they are considered to be acquired via HGT from insects (14; 15). *Ascaris suum* (an animal parasite) presents a large genome with a similar number and size of genes. However, it has a low repeat content (4.4%) and a greater intron size compared to the other species sequenced (Table 2; 16). Conversely, the genome or *Dirofilaria immitis* (an animal parasite) harbors neither DNA transposons nor active retrotransposons (17). Some animal-parasitic nematodes with sequenced genomes (e.g., *B. malayi*, *D. immitis*) harbor intracellular symbiotic bacteria of the genus *Wolbachia* (17), and the genomes reveal the genetic basis of this interrelationship. As explained before, the genomes of nematodes are very diverse between and within the same genus employing the same parasite strategy; thus, new genomes will show more diversity in this sense.

Phylogenomics, which uses genomes (or large portions thereof) to reconstruct evolutionary relationships between species, is dependent on the development of genomic resources. This approach is not yet applicable to PPN as only three species have been sequenced to date. However, Kikuchi *et al.* (2011) conducted this analysis on the seven nematodes studied. The study of PPN families with difficult evolutionary relationships (such as Meloidogynidae or Pratylenchidae) will become more defined as more genomes become available.

2.2. *Transcriptomics*

The most rapid and cost-effective approach to gene discovery in eukaryotic genomes (including PPN) has been the generation of expressed sequence tags (ESTs) (18). However, EST data has such shortcomings as: i) base-calling errors when dataset redundancy is not sufficiently considered; ii) ESTs are generally not complete and do not cover the gene's entire coding sequence; and iii) short inserts tending to clone more efficiently than longer ones (18). Nevertheless, ESTs are considered among the most important tools for studying PPN at the molecular level. Analysis of the GenBank division housing ESTs (http://www.ncbi.nlm.nih.gov/dbEST) showed 73 datasets for nematodes, 21 of which were derived from PPN (28.8%). However, PPN only accounted for 210,422 (17.7%) of the total ESTs (1,190,246), with sedentary endoparasites accounting for the majority of these sequences. Some web tools provide more useful data for researchers, such as wormbase (www.wormbase.or; 19), Nembase4 (http://www.nematodes.org/nembase4; 20), and Nematode.net (www.nematode.ne; 20). The use of such new techniques as NGS or microarrays (based on EST datasets, complete nematode genomes, or joined nematode/plant genes) will expand as the techniques become less expensive and more available at a greater number of research centers.

Sedentary endoparasites account for the majority of transcriptomic data at all levels. These species include *M. incognita*, *M. hapla*, *M. javanica*, *M. paranaensis*, *M. chitwoodi*., *Globodera pallida*, *G. rostochiensis*, *G. mexicana*, *Heterodera*

Table 2. General features of nematode genomes sequenced to date (adapted 10).

Features	*B. xylophilus*	*M. incognita*[a]	*M. hapla*[b]	*C. elegans*[b]	*P. pacificus*[b]	*B. malayi*[b]	*A. suum*[c]	*D. immitis*[d]
Overall/life-vector[e]	PPN, insect	PPN; -	PPN; -	BF; -	BF, FF, NP; insect	AP; insect	AP, -	AP, insect
Estimated size of genome (Mb)	61-73[f]	47-51	54	100*	169	90-95	309	84,2
Chromosome number	6	Variable	16	6	6	6	24	10
Total size of assembled sequence (Mb)	74,6	86	53	100	172,5	95,8	273	84,2
Number of scaffolds / chromosomes	1.231	2.817	1,523[g]	6 chr.	2,894[g]	8,180[g]	29.831	1.618
N50 of scaffolds (kb)	1.158	83	84[g]	-	1.244	94[g]	408	10.474
Maximum length of scaffold (kb)	3.612	593	360	-	5.268	6.534	-	168
G + C content (%)	40,4	31,4	27,4	35,4	42,8	30,5	37,9	28,3
Completeness[h]								
Cegma completeness (%): (complete/partial)	97/98	73/77	95/96	100/100	95/98	95/96	96/98	97/98
Average CEG gene number: (complete/partial)	1.08/1.09	1.53/1.61	1.07/1.12	1.05/1.06	1.20/1.23	1.07/1.11	1.18/1.22	1.18/1.22
Protein-coding regions								
Number of gene models	18.074	19.212	14.420	20.416	21.416	18.348	18.542	11.375
Number of proteins	18.074	20.365	13.072	24.890	24.217	21.252	18.542	12.344
Average protein length	345	354	309,7	439,6	331,5	311,9	327,7	-
Gene density (genes per Mb)	242,3	223,4	270	249	140,4	221,8	-	135
Mean exon size (bp)	288,9	169	171,5	201,6	96,7	159,8	153	-
Mean number of exons per gene	4,5	6,6	6,1	6,5	10,3	5,9	6,4	-
Mean intron size (bp)	153	230	154	320	309	280	1081	226

a 8, b WormBase release 221, c 16, d 17.

e PPN: Plant-parasitic nematode, BF: bacterial feeder, FF: fungal feeder, NP: nematode predator, AP: animal parasite.

f Genome size was estimated using real time PCR.

g These values are from 9, 80 and 81 respectively.

h Assembly completeness was estimated by CEGs (Core Eukaryotic Genes) with CEGMA software. Assemblies in WormBase release 221 was used for *M. incognita*, *M. hapla*, *C. elegans* and *P. pacificus* genomes. *Ascaris suum* and *D. immitis* were downloaded form http://www.wormbase.org/species/all#01--10 and http://nematodes.org/genomes/dirofilaria_immitis/, respectively.

The genome features of *B. xylophilus*, *M. incognita*, *M. hapla*, *P. pacificus*, *B. malayi*, *A. suum* and *D. immitis* are based on incomplete genome drafts and represent statistical estimates while *C. elegans* genome has been completely sequenced and well annotated.

glycines, *H. schachtii*, and *H. avenae* (6; 9, 22-31). The migratory endoparasite species studied are *Pratylenchus coffeae*, *P. thornei*, *P. penetrans*, *Radopholus similis*, *Ditylenchus africanus*, and *B. xylophilus* (30; 32-37). The only semi-endoparasitic species studied is *Rotylenchulus reniformis* (38), and the only ESTs from an ectoparasite are from *Xiphinema index* (39). The majority of these studies have been centered on understanding the mechanisms (mainly protein effectors) involved in plant-parasitism caused by PPN. Few studies have been directed at a comparison of different environmental studies, different specific nematode stages, or different nematodes strains/species (35; 40- 45). The study of avirulence factors has been achieved by employing a cDNA-amplification fragment length polymorphism (AFLP)-based strategy in *M. incognita* for the gene *Mi* (39; 46). However, the avirulence factor (*map-1*) (46) was not found for *M. javanica*, and when nematode juveniles of the *Mi-1*-avirulent strain were soaked in dsRNA from a different potential avirulence factor *Cg-1*, they produced progeny that were virulent on tomato carrying the *Mi-1* gene (41). Other example of avirulence factor is the cyst nematode SPRYSEC protein RBP-1 in *G. pallida*, which elicits Gpa2 and RanGAP2 dependent plant cell death (47). This recognition of Gp-RBP-1 correlated to a single amino acid polymorphism at position 187 in the Gp-RBP-1 SPRY domain (47). The effects of different plant resistant genotypes on the same nematode (*G. pallida*) have also been studied using microarrays in different breeding potato lines, revealing similarities in the mode of action against the nematode (45). Different *H. glycines* populations (virulent or avirulent) exposed to the resistant *Glycine max* genotype (Peking) were studied at pre-parasitic stages and at different times of post-infection (43), with numerous putative parasitism genes being found expressed differentially, as well as numerous genes (1668) being suppressed in the avirulent population, and induced in the virulent population (43).

Next Generation Sequencing has been used for several nematodes, the majority of which are migratory endoparasites. For example, *P. thornei* and *P. coffeae* transcriptomes have been studied using a 454 strategy (36; 37). Both *Pratylenchus* spp. showed similarities in terms of effectors relative to other species of PPN with different lifestyles. Illumina NGS has been used for a comparison between *B. xylophilus* (PPN) and *B. mucronatus* (only PPN under controlled conditions) that showed a similar adaptation to life on pine hosts (35). NGS has also been used in the genome annotation analysis of *B. xylophilus* (10). New web integration platforms (such as Nematode.net) enabling the use of NGS data and the development of new programs or web tools for large-scale genome analyses will help promote the development of these studies (48).

2.3. *Proteomics*

Proteomics is the study of proteins or proteomes. A specific proteome can be defined as all the proteins in an organism, organ, tissue or cell under specific environmental and temporal conditions. The use of proteomics in PPN has been hampered due to a lack of sequenced genomes or ESTs datasets for some species and due to the problems in obtaining the required sample amounts from species that have long life cycles or which are obligate endoparasites. The few studies using proteomics that have been conducted on PPN have been conducted with the aim of i) identifying effectors secreted by the nematode in order to understand their role in plant-nematode interaction; ii) map the entire proteome, and iii) confirm annotated genomes.

The use of proteomics in identifying effectors was made possible by the study of specific subproteomes. The secretome of *M. incognita* was investigated using 2-DE gels (49), followed by a more complete analysis using nano-LC-ESIMS/MS to find effectors (50). Rehman *et al*., (2009) used a monoclonal antibody to immunopurify the most abundant cellulases in the stylet secretions of pre-parasitic juveniles of *G. rostochiensis* (51). Different effectors have also been detected using these techniques. However, the detection capabilities are limited due to the amounts of protein required for their correct detection and identification.

Different proteomic maps using 2-DE gels have been made for *Heterodera glycines* (a soybean cyst nematode), *B. xylophilus*, *Ditylenchus dipsaci,* and different species of *Meloidogyne*. *Heterodera glycines* showed 803 proteins using 2-D gels, with 426 spots being identified by LC-MS/MS, indicating that those showing metabolic, developmental, and biological regulation processes were the most abundant (52). Navas *et al*. (2002) compared different isolates from the species of *Meloidogyne* (i.e., *M. incognita*, *M. javanica*, *M. arenaria*), in order to find proteomic markers for identification (53). The identification of several differential proteins in gel suggested their possible use as putative diagnostic markers for the species (54). A similar study was conducted on the discovery of possible protein biomarkers in *D. dipsaci* races (55). The peptides of galactose-binding lectin-1 of *B. xylophilus* were demonstrated as being the antigen target of MAb-D9-F10, as based several types of proteomic analyses (e.g., SDS-PAGE, 2 -DE, anion exchange chromatography, immunoprecipitation) (56), while other studies more centered on subproteomes showed important differences when the nematode is inside the host or at various stages of development (57; 58). On the other hand, the complete proteome of *M. hapla* (59) is the only one that has been used thus far to improve genome annotation and provide experimental confirmation of the computational predictions of intron/exon structures.

2.4. *Metabolomics and other –omics fields of study*

Technical advances made in high-resolution NMR spectroscopy and mass spectrometry, in an attempt to capture the complexity of metabolic networks (60), have increased the possibilities of metabolomics, which could be subdivided into other subdivisions such as lipidomics and glycomics. In a widest sense of the word, metabolomics has been applied

to *C. elegans* in order to study the regulation of signals and aggregating behavior (61), the effects of pesticides and heavy metals or a combination of both (62), changes in diet (63), and the detection of differences between mutants (64). Lipidomics characterizes the composition of intact lipid molecular species in biological systems (65). The discovery of new antiparasitic drug targets using membrane lipidomics and the changes in lipid balance in the membranes of parasites could provide clues to the dynamics of drugs and some mechanisms of drug resistance; work in this area has focusedmainly on vertebrate nematode parasites (66). However, few studies have been conducted using lipidomics in PPN. Only a few approaches to the variation in lipid reserves of second-stage juveniles of *Meloidogyne exigua* in a coffee field and its relation with infectivity have been studied (67). While in cyst nematodes, several studies showed the effect of long-term storage on the lipid reserves and fatty acid composition of cyst and hatched juveniles of *G. rostochiensis* and *G. pallida* (68); during rehydratation, exposure to the hatching stimulus and hatch in *G. rostochiensis* (69) or the influence of the host plant on lipid reserves of *G. rostochiensis* (70).Some studies attempted to identify fatty patterns in the soil for biodiversity assessment (71; 72). The glycome consists of all glycans (or carbohydrates) within a biological system, and modulates a wide range of important biological activities from protein folding to cellular communication (73). Arrays and mass spectrometry techniques are used to study the glycome. Glycomics is still in the development stage in nematodes, and only partial studies on certain glycans related to this matter have been reported (74; 75).

3. Use of –omics studies in plant-parasitic nematode population management

The best management of PNN is to avoid its introduction to the field, because once PPN is introduced and the population stabilized, they are almost impossible to eradicate given their long survival stages, deep soil penetration, and the ineffectiveness of most pesticides now permitted to be used in soil. The management of PPN in general, or that of a specific species, should be centered on maintaining population levels below the economic damage threshold for the crops being grown. The cost of control measures must also be adjusted relative to the cost of the expected yield reduction, as compared to the yield in a situation where there is no need for control (76). For these reasons, the nematodes in certain low -cash crops could significantly impact their management. The most efficient and economically viable strategy is usually host-plant resistance. However, "Integrated Pest Management", which integrates several control measures (agronomical, chemical, and genetic) is often the best solution for PPN management.

The impact of the several –omics studies discussed above is or will be significant in terms of PPN management. Correctly and quickly identifying species is the first point in the correct integrated management of PPN. Plant-parasitic nematodes have a conserved morphology and few characters are useful for experienced nematologists in delimitating species, but morphologically very similar species may have markedly different pathogenic features. As mentioned before, some species are highly host-specific and genomic or proteomic approaches could help to identify species. More importantly, both pathogenicity genes and survival genes play an important role in PPN management. –Omics studies offer a faster way to study these genes, but the sequences of interest will be pioneers that do not have orthologues in model organisms, such as *C. elegans*. Pathogenicity genes, specifically those implied in the formation of "feeding sites," are of great interest for their biological significance and putative use as control measures targeting their expression using RNAi. RNAi is a useful research tool for determining gene function and has potential applications in commercial nematode control through transgenic plant-derived dsRNA (11). This tool has been used extensively to knock down gene expression in *C. elegans* by introducing dsRNA using microinjection, soaking, and feeding (77). The soaking of second-stage juveniles (J2s) in dsRNA solution in combination with drugs (e.g., resorcinol) has been applied to several species of PPN (reviewed in 78), but more recent studies have highlighted the challenges of non-specific phenotypic effects from long dsRNAs (79).

4. Concluding Remarks

Plant-parasitic nematology will generally be assisted in finding solutions to minimize crop loss caused by PPN. In this sense, the next step in using these genome sequences will be comparative and functional analysis in order to find out genes involved in their pathogenicity/virulence which will be achieved by analyzing different strains or different species from the same genus. The future availability of "third-generation" sequencing approaches based on single DNA molecule sequencing, with longer read capabilities and greater parallel sequencing densities (i. e., Pacific Biosciences technology or Oxford Nanopore) may be helpful in this regard. Such new technologies will help reduce the costs of genome sequencing. On the other hand, the use of other –omics fields of study, such as metabolomics, lipidomics and glycomics in plant-parasitic nematology, will expand as more genomes become available to researchers. However, all these techniques should have an important applied role in crop protection as much as possible. Otherwise, we run the risk of creating science for scientists, and not for farmers.

Acknowledgements

The authors wish to thank the Japanese Society for Promotion of Science for the fellowship provided to Dr. Juan E. Palomares-Rius and Dr. Pablo Castillo (Institute of Sustainable Agriculture-CSIC), and Dr. Blanca B. Landa (Institute of Sustainable Agriculture-CSIC) for revision of the article.

References

1. J.M. Nicol, S.J. Turner, D.L. Coyne, L. den Nijs, S. Hockland, Z. Thana Maafi, Current nematodes threats to world agriculture, in: J. Jones, G. Gheysen, C. Fenoll, (Eds.), Genomics and molecular genetics of plant-nematode interactions, Springer, Dordrecht, 2011, pp. 21-43. DOI: 10.1007/978-94-007-0434-3.
2. J.N. Sasser, D.W. Freckman, A world perspective on Nematology: the role of the society, in: J.A. Veech, D.W. Dickson, (Eds.), Vistas on Nematology, Society of Nematologists, Hyattsville, Maryland, 1987, pp. 7-14.
3. S.R. Koenning, C. Overstreet, J.W. Noling, P.A. Donald, J.O. Becker, B.A. Fortnum, J. Nematol. 31 (1999) 587-618.
4. World Bank, World Development Report 2008: Agriculture for Development. The World Bank, Washington DC. (2008).
5. B.A. Keating, P.S. Carberry, P.S. Bindraban, S. Asseng, H. Meinke, J. Dixon, Crop. Sci. 50 (2010) S-109-S-119. DOI: 10.2135/cropsci2009.10.0594.
6. M. Blaxter, Kumar, S., Kaur, G., Koutsovoulos, G., Elsworth, B., Parasite Immunol. 34 (2012) 108-120. DOI: 10.1111/j.1365-3024.2011.01342.x.
7. N. Holroyd, A. Sanchez-Flores, Parasite Immunol. 34 (2012) 100-107. DOI: 10.1111/j.1365-3024.2011.01311.x.
8. P. Abad, J. Gouzy, J. Aury, P. Castagnone-Sereno, E. Danchin, E. Deleury, L. Perfus-Barbeoch, V. Anthouard, F. Artiguenave, V. Blok, M.-C. Caillaud, P.M. Coutinho, C. Dasilva, F. De Luca, F. Deau, M. Esquibet, T. Flutre, J.V. Goldstone, N. Hamamouch, T. Hewezi, O. Jaillon, C. Jubin, P. Leonetti, M. Magliano, T.R. Maier, G.V. Markov, P. McVeigh, G. Pesole, J. Poulain, M. Robinson-Rechavi, E. Sallet, B. Ségurens, D. Steinbach, T. Tytgat, E. Ugarte, C. van Ghelder, P. Veronico, T.J. Baum, M. Blaxter, T. Bleve-Zacheo, E.L. Davis, J.J. Ewbank, B. Favery, E. Grenier, B. Henrissat, J.T. Jones, V. Laudet, A.G. Maule, H. Quesneville, M.-N. Rosso, T. Schiex, G. Smant, J. Weissenbach, P. Wincker, Nat. Biotechnol. 26 (2008) 909-915. DOI:10.1038/nbt.1482.
9. C.H. Opperman, D.M. Bird, V.M. Williamson, D.S. Rokhsar, M. Burke, J. Cohn, J. Cromer, S. Diener, J. Gajan, S. Graham, T.D. Houfek, Q. Liu, T. Mitros, J. Schaff, R. Schaffer, E. Scholl, B.R.Sosinski, V.P. Thomas, E. Windham, P. Nat. Acad. Sci. USA 105 (2008) 14802-14807. DOI: 10.1073/pnas.0805946105.
10. T. Kikuchi, J.A. Cotton, J.J. Dalzell, K. Hasegawa, N. Kanzaki, P. McVeigh, T. Takanashi, I.J. Tsai, S.A. Assefa, P.J.A. Cock, T. Dan Otto, M. Hunt, A.J. Reid, A. Sanchez-Flores, K. Tsuchihara, T. Yokoi, M.C. Larsson, J. Miwa, A.G. Maule, N. Sahashi, J.T. Jones, M. Berriman, PLoS Pathog. 7 (2011) e1002219. DOI: 10.1371/journal.ppat.1002219.
11. P. Abad, J.P. McCarter, Genome analysis of plant parasitic nematodes, in: J.T. Jones, Gheysen, G., Fenoll, C., (Ed.), Genomics and molecular genetics of plant-nematode interactions, Springer, Dordrecht, 2011, pp. 103-117. DOI: 10.1007/978-94-007-0434-3.
12. A. Haegeman, J.T. Jones, E.G.J. Danchin, Mol. Plant Microbe Interact. 24 (2011a) 879-887. DOI: 10.1094/MPMI-03-11-0055.
13. A. Haegeman, S. Mantellin, J.T. Jones, G. Gheysen, Gene 492 (2012) 19-31. DOI: 10.1016/j.gene.2011.10.040.
14. E.G.J. Danchin, Mobile Genetic Elements 1 (2011) 269-273. DOI: 10.4161/mge.1877.
15. C. Rödelsperger, R.J. Sommer, BMC Evol. Biol. 11 (2011) 239. DOI: 10.1186/1471-2148-11-239.
16. A.R. Jex, S. Liu, B. Li, N.D. Young, R.S. Hall, Y. Li, L. Yang, N. Zeng, X. Xu, Z. Xiong, F. Chen, X. Wu, G. Zhang, X. Fang, Y. Kang, G.A. Anderson, T.D. Harris, B.E. Campbell, J. Vlaminck, T. Wang, C. Cantacessi, E.M. Schwarz, S. Ranganathan, P. Geldhof, P. Nejsum, P.W. Sternberg, H. Yang, J. Wang, J. Wang, R.B. Gasser, Nature 479 (2011) 529-536. DOI: 10.1038/nature10553.
17. C. Godel, S. Kumar, G. Koutsovoulos, P. Ludin, D. Nilsson, F. Comandatore, N. Wrobel, M. Thompson, C.D. Schmid, S. Goto, F. Bringaud, A. Wolstenholme, C. Bandi, C. Epe, R. Kaminsky, M. Blaxter, P. Mäser, FASEB J. 26 (2012) 4650-4661. DOI: 10.1096/fj.12-205096.
18. J. Jacob, M. Mitreva, Transcriptomes of plant-parasitic nematodes, in: J. Jones, Gheysen, G., Fenoll, C., (Ed.), Genomics and molecular genetics of plant-nematode interactions, Springer, Dordrecht, 2011, pp. 119-138. DOI: 10.1007/978-94-007-0434-3.
19. T.W. Harris, I. Antoshechkin, T. Bieri, Blasiar D., J.C. Chan, W.J. Chen, N. De La Cruz, P. Davis, M. Duesbury, R. Fang, J. Fernandes, M. Han, R. Kishore, R. Lee, H.-S. Müller, C. Nakamura, P. Ozersky, P. Petcherski, A. Rangarajan, A. Rogers, G. Schindelman, E.M. Schwarz, M.A. Tuli, K. Van Auken, D. Wang, X. Wang, G. Williams, K. Yook, R. Durbin, L.D. Stein, J. Spieth, P.W. Sternberg, Nucleic Acids Res. 38 (2010) D463-D467. DOI: 10.1093/nar/gkp952.
20. B. Elsworth, J. Wasmuth, M. Blaxter, Int. J. Parasitol. 41 (2011) 881-894. DOI: 10.1016/j.ijpara.2011.03.009.
21. J. Martin, S. Abubucker, E. Heizer, C.M. Taylor, M. Mitreva, Nucleic Acids Res. 40 (2012) D720–D728. DOI: 10.1093/nar/gkr1194.
22. M. Popeijus, V.C. Blok, L. Cardle, E. Bakker, M.S. Phillips, J. Helder, G. Smant, J.T. Jones, Nematology 2 (2000) 567-574.
23. J.M. De Boer, J.P. McDermott, X.H. Wang, T. Maier, F. Qu, R.S. Hussey, E.L. Davis, T.J. Baum, Mol. Plant Pathol. 3 (2002) 261-270. DOI: 10.1046/j.1364-3703.2002.00122.x.
24. B.L. Gao, R. Allen, T. Maier, E.L. Davis, T.J. Baum, R.S. Hussey, Mol. Plant Microbe Interact. 16 (2003) 720-726.
25. G.Z. Huang, B.L. Gao, T. Maier, R. Allen, E.L. Davis, T.J. Baum, R.S. Hussey, Mol. Plant Microbe Interact. 16 (2003) 376-381. DOI: 10.1094/MPMI.2003.16.5.376
26. J.P. McCarter, M.D. Mitreva, J. Martin, M. Dante, T. Wylie, U. Rao, D. Pape, Y. Bowers, B. Theising, C.V. Murphy, A.P. Kloek, B.J. Chiapelli, S.W. Clifton, D.M. Bird, R.H. Waterston, Genome Biol. 4 (2003) R26. DOI: 10.1186/gb-2003-4-4-r2.
27. G.Z. Huang, R.H. Dong, T. Maier, R. Allen, E.L. Davis, T.J. Baum, R.S. Hussey, Mol. Plant Pathol. 5 (2004) 217-222. DOI: 10.1111/j.1364-3703.2004.00220.x.
28. B. Vanholme, M. Mitreva, W. Van Criekinge, M. Logghe, D. Bird, J.P. McCarter, G. Gheysen, Parasitology Res. 98 (2006) 414-424. DOI: 10.1007/s00436-005-0029-3.
29. S. Bekal, J.P. Craig, M.E. Hudson, T.L. Niblack, L.L. Domier, K.N. Lambert, Mol. Genet. Genomics 279 (2008) 535-543. DOI: 10.1007/s00438-008-0331-8.
30. E. Roze, B. Hanse, M. Mitreva, B. Vanholme, J. Bakker, G. Smant, Mol. Plant Pathol. 9 (2008) 1-10. DOI: 10.1111/j.1364-3703.2007.00435.x.
31. M.J. Kang, Y.H. Kim, B.S. Hahn, Genes Genom. 32 (2010) 553-562. DOI: 10.1007/s13258-010-0065-y.
32. T. Kikuchi, T. Aikawa, H. Kosaka, L. Pritchard, N. Ogura, J.T. Jones, Mol. Biochem. Parasit. 155 (2007) 9-17. DOI: 10.1016/j.molbiopara.2007.05.00.
33. J. Jacob, M. Mitreva, B. Vanholme, G. Gheysen, Mol. Genet. Genomics 280 (2008) 1-17. DOI: 10.1007/s00438-008-0340-

7
34. A. Haegeman, J. Jacob, B. Vanholme, T. Kyndt, M. Mitreva, G. Gheysen, Mol. Biochem. Parasit. 167 (2009) 32-40. DOI: 10.1016/j.molbiopara.2009.04.004.
35. X. Yan, X.-Y. Chen, Y.-S. Wang, J. Luo, Z.-C. Mao, V.R. Ferris, B.-Y. Xie, Gene 505 (2012) 81-90. DOI: 10.1016/j.gene.2012.05.041.
36. A. Haegeman, S. Joseph, G. Gheysen, Mol. Biochem. Parasit. 178 (2011b) 7-14. DOI: 10.1016/j.molbiopara.2011.04.00.
37. P. Nicol, R. Gill, J. Fosu-Nyarko, M.G.K. Jones, Int. J. Parasitol. 42 (2012) 225-237. DOI: 10.1016/j.ijpara.2011.11.010.
38. M.J. Wubben, F.E. Callahan, B.S. Scheffler, Mol. Biochem. Parasit. 172 (2010) 31-40. DOI: 10.1016/j.molbiopara.2010.03.011.
39. C. Furlanetto, L. Cardle, D.J.F. Brown, J.T. Jones, Nematology 7 (2005) 95-104.
40. C. Neveu, S. Jaubert, P. Abad, P. Castagnone-Sereno, Mol. Plant Microbe Interact. 16 (2003) 1077-1084. DOI: 10.1094/MPMI.2003.16.12.1077.
41. C.A. Gleason, Q.L. Liu, V.M. Williamson, Mol. Plant Microbe Interact. 21 (2008) 576-585. DOI: 10.1094/MPMI-21-5-0576.
42. J.T. Jones, A. Kumar, L.A. Pylypenko, A. Thirugnanasambandam, L. Castelli, S. Chapman, P.J.A. Cock, E. Grenier, C.J. Lilley, M.S. Phillips, V.C. Blok, Mol. Plant Pathol. 10 (2009) 815-828. DOI: 10.1111/j.1364-3703.2009.00585.x.
43. V.P. Klink, P. Hosseini, M.H. MacDonald, N.W. Alkharouf, B.F. Matthews, BMC Genomics 10 (2009) 111. DOI: 10.1186/1471-2164-10-111.
44. P. Castagnone-Sereno, E. Deleury, E.G.J. Danchin, L. Perfus-Barbeoch, P. Abad, Genomics 97 (2011) 29-36. DOI: 10.1016/j.ygeno.2010.10.002.
45. J.E. Palomares-Rius, P.E. Hedley, P.J.A. Cock, J.A. Morris, J.T. Jones, N. Vovlas, V. Blok, Mol. Plant Pathol. 13 (2012) 1120-1134. DOI: 10.1111/j.1364-3703.2012.00821.x.
46. J.P. Semblat, M.N. Rosso, R.S. Hussey, P. Abad, P. Castagnone-Sereno, Mol. Plant Microbe Interact. 14 (2001) 72-79. DOI: 10.1094/MPMI.2001.14.1.72.
47. M.A. Sacco, K. Koropacka, E. Grenier, M.J. Jaubert, A. Blanchard, A. Goverse, G. Smant, P., PLoS Pathog. 5 (2009) e1000564. DOI: 10.1371/journal.ppat.1000564.
48. J. Goecks, A. Nekrutenko, J. Taylor, Genome Biol. 11 (2010) R86. DOI: 10.1186/gb-2010-11-8-r86.
49. S. Jaubert, T.N. Ledger, J.B. Laffaire, C. Piotte, P. Abad, M.-N. Rosso, Mol. Biochem. Parasit. 121 (2002) 205-211.
50. S. Bellafiore, Z. Shen, M-N. Rosso, P. Abad, P. Shih, S.P. Briggs, PLoS Pathog. 4 (2008) e1000192. DOI: 10.1371/journal.ppat.1000192.
51. S. Rehman, P. Butterbach, H. Popeijus, H. Overmars, E.L. Davis, J.T. Jones, A. Goverse, J. Bakker, G. Smant, Phytopathology 99 (2009) 194-202. DOI: 10.1094/PHYTO-99-2-0194.
52. X. Chen, MacDonald, M. H., Khan, F., Garrett, W. M., Matthews, B. F., Natarajan, S. S., Two-dimensional proteome reference maps for the soybean cyst nematode Heterodera glycines. Proteomics 11 (2011) 4742-4746.
53. A. Navas, J.A. López, G. Espárrago, E. Camafeita, J.P. Albar, J. Proteome Res. 1 (2002) 421-427. DOI: 10.1021/pr0255194.
54. E. Calvo, P. Flores-Romero, J.A. López, A. Navas, J. Proteome Res. 4 (2005) 1017–1021. DOI: 0.1021/pr0500298.
55. M.R. Perera, S.P. Taylor, V.A. Vanstone, M.G.K. Jones, Nematology 11 (2009) 555-563. DOI: 10.1111/j.1364-3703.2012.00821.x.
56. D.W.P. Lee, J.B.P. Seo, M.H. Nam, J.S. Kang, S.Y. Kim, A.Y. Kim, W.T. Kim, J.K. Choi, Y. Um, Y. Lee, I.S. Moon, H.R. Han, S.H. Koh, Y.H. Je, K.J. Lim, S.H. Lee, Y.H. Koh, Mol. Cell. Proteomics 10 (2011) M900521-MCP900200. DOI: 10.1074/mcp.M900521-MCP200.
57. R. Shinya, Y. Takeuchi, N. Miura, K. Kuroda, M. Ueda, K. Futai, Nematology 11 (2009) 429-438.
58. R. Shinya, H. Morisaka, Y. Takeuchi, M. Ueda, K. Futai, Phytopathology 100 (2010) 1289-1297. DOI: 10.1094/PHYTO-04-10-0109.
59. F. Mbeunkui, E.H. Scholl, C.H. Opperman, M.B. Goshe, D. McK. Bird, J. Proteome Res. 9 (2010) 5370–5381. DOI: 10.1021/pr1006069.
60. J.L. Griffin, H. Atherton, J. Shockcor, L. Atzori, Nature reviews cardiology 8 (2011) 630-643. DOI: 10.1038/nrcardio.2011.138.
61. J. Srinivasan, S.H. von Reuss, N. Bose, A. Zaslaver, P. Mahanti, M.C. Ho, O.G. O'Doherty, A.S. Edison, P.W. Sternberg, F.C. Schroeder, Plos Biol. 10 (2012) e1001237. DOI: 10.1371/journal.pbio.1001237.
62. O.A.H. Jones, S.C. Swain, C. Svendsen, J.L. Griffin, S.R. Sturzenbaum, D.J. Spurgeon, J Proteome Res. 11 (2012) 1446-1453. DOI: 10.1021/pr201142c.
63. S.N. Reinke, X. Hu, B.D. Sykes, B.D. Lemire, Mol. Genet. Metab. 100 (2010) 274-282. DOI: 10.1016/j.ymgme.2010.03.013.
64. J.A. Butler, N. Ventura, T.E. Johnson, S.L. Rea, FASEB J. 24 (2010) 4977-4988. DOI: 10.1096/fj.10-162941.
65. A.D. Postle, Curr. Opin. Clin. Nutr. 15 (2012) 127-133. DOI: 10.1093/jncimonographs/lgs019.
66. E. Marechal, M. Riou, D. Kerboeuf, F. Beugnet, P. Chaminade, P.M. Loiseau, Trends in parasitol. 27 (2011) 496-504. DOI: 10.1016/j.pt.2011.07.002.
67. F.P. Rocha, V.P. Campos, J.T. de Souza, Nematology 12 (2010) 365-371.
68. R.A. Holz, D.J. Wright, R.N. Perry, J. Helminthol. 72 (1998) 133-141. DOI: 10.1017/S0022149X0001631.
69. R.A. Holz, Wright, R.N. Perry, Parasitology 116 (1998) 183-190.
70. R.A. Holz, Wright, R.N. Perry, Nematologica 44 (1998) 153-169. DOI: 10.1163/005325998X0004.
71. L. Ruess, M.M. Haggblom, E.J.G. Zapata, J. Dighton, Soil Biol. & Biochem. 34 (2002) 745-756.
72. E. Yergeau, T.M. Bezemer, K. Hedlund, S.R. Mortimer, G.A. Kowalchuk, W.H. van der Putten, Environ. Microbiol. 12 (2010) 2096-2106. DOI: 10.1111/j.1462-2920.2009.02053.x.
73. S. Hua, H.J. An, Bmb Rep. 45 (2012) 323-330. DOI: 10.5483/BMBRep.2012.45.6.132
74. A.J. Hanneman, J.C. Rosa, D. Ashline, V.N. Reinhold, Glycobiology 16 (2006) 874-890. DOI: 10.1093/glycob/cwl011.
75. D. Rendic, I.B.H. Wilson, G. Lubec, M. Gutternigg, F. Altmann, R. Leonard, Electrophoresis 28 (2007) 4484-4492. DOI: 10.1002/elps.200700098.
76. C.H. Schomaker, T.H. Been, Plant growth and population dynamics, in: R.N. Perry, M. Moens, (Eds.), Plant Nematology, CABI, Wallingford, Oxfordshire, UK, 2006.
77. A.G. Fraser, R.S. Kamath, P. Zipperlen, M. Martinez-Campos, M. Sohrmann, J. Ahringer, Nature 408 (2000) 325-330. DOI: 10.1038/35042517.
78. M.N. Rosso, J.T. Jones, P. Abad, Annu. Rev. Phytopathol. 47 (2009) 207-232. DOI: 10.1146/annurev.phyto.112408.132605.
79. J.J. Dalzell, S. McMaster, M.J. Johnston, R. Kerr, C.C. Fleming, A.G. Maule, Int. J. Parasitol. 39 (2009) 1503-1516. DOI:

10.1016/j.ijpara.2009.05.006.

80. C. Dieterich, S.W. Clifton, L.N. Schuster, Chinwalla. A., K. Delehaunty, I. Dinkelacker, L. Fulton, R. Fulton, J. Godfrey, P. Minx, M. Mitreva, W. Roeseler, H. Tian, H. Witte, S.P. Yang, R.K. Wilson, R.J. Sommer, Nat. Genet. 40 (2008) 1193-1198. DOI: 10.1038/ng.227.

81. E. Ghedin, S. Wang, D. Spiro, E. Caler, Q. Zhao, J. Crabtree, J.E. Allen, A.L. Delcher, D.B. Guiliano, D. Miranda-Saavedra, S.V. Angiuoli, T. Creasy, P. Amedeo, B. Haas, N.M. El-Sayed, J.R. Wortman, T. Feldblyum, L. Tallon, M. Schatz, M. Shumway, H. Koo, S.L. Salzberg, S. Schobel, M. Pertea, M. Pop, O. White, G.J. Barton, C.K.S. Carlow, M.J. Crawford, J. Daub, M.W. Dimmic, C.F. Estes, J.M. Foster, M. Ganatra, W.F. Gregory, J. N.M., J. Jin, R. Komuniecki, I. Korf, S. Kumar, S. Laney, B.-W. Li, W. Li, T. H., T.H. Lindblom, S. Lustigman, D. Ma, C.V. Maina, D.M.A. Martin, J.P. McCarter, L. McReynolds, M. Mitreva, T.B. Nutman, J. John Parkinson, J.M. Peregrín-Alvarez, C. Poole, Q. Ren, L. Saunders, A.E. Ann E. Sluder, K. Smith, M. Stanke, T.R. Unnasch, J. Ware, A.D. Wei, G. Weil, D.J. Williams, Y. Zhang, S.A. Williams, C. Fraser-Liggett, B. Slatko, M.L. Blaxter, A.L. Scott, Science 317 (2007) 1756. DOI: 10.1126/science.1145406.

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

ORIGINAL ARTICLE | DOI: 10.5584/jiomics.v3i1.123

Red Blood Cell Lipidomics analysis through HPLC-ESI-qTOF: application to red blood cell storage

Anna Maria Timperio[1], Cristiana Mirasole[1], Angelo D'Alessandro[1], Lello Zolla[1,*]

[1]Department of Ecological and Biological Sciences, University of Tuscia, Largo dell'Università, snc, 01100 Viterbo, Italy

Received: 11 January 2013 **Accepted:** 02 February 2013 **Available Online**: 11 March 2013

Abstract

Recent developments in mass spectrometry (MS) have enabled fast and sensitive detection of lipid species in different biological matrices. In the present study we performed an on-line HPLC-microTOF-Q MS approach to the red blood cell (RBC) lipidome. We thus exploited bioinformatic tools for the interrogation of novel databases, such as LIPID MAPS. By means of *ad hoc* software suites for mass spectrometry-based metabolomics analyses, we could address the key biological issue of the RBC lipidome, within the framework of RBC storage for transfusion purposes. Samples were collected from subjects living in the province of Viterbo, where olive oil consumption represents a central aspect of the diet. On this ground, we could postulate a diet specific effect on the accumulation of lipid-specific storage lesions. The analyses yielded the tentative identification of a huge number of lipid molecules on the basis of accurate intact mass values and retention times, and MS/MS validation. This analytical workflow was exploited to consolidate existing knowledge on the RBC lipid composition and individuate statistically significant fluctuations of lipids throughout storage duration of RBC concentrates under blood bank conditions. Our analysis indicated ceramides, glycerophospholipids and sterols as key targets of RBC storage lesions to the lipidome, that will deserve further targeted investigations in the future. It also emerged how compositional analyses of the RBC lipidome might end up yielding different results on the basis of the background of the blood donor (i.e. diet), which might translate into region-specific lipidomic alterations over storage progression of RBC concentrates.

Keywords: Red Blood Cells; Mass Spectrometry; Blood Storage; Lipidomics.

1. Introduction

Lipidomics is the systematic identification of the lipid molecular species of a biological matrix (either a cell, organelle, globule, or whole organism) with emphasis on the relative quantitation of composition changes in response to a perturbation, such as ageing or drug treatments [1].While the term "lipidomics" dates back to a decade ago, investigations of the lipid content of specific biological matrices was an already consolidated field of research over the last fifty years [2]. In particular, this holds true for those matrices that are largely available and display limited biological complexity, such as anucleated cells and, in particular red blood cells (RBCs) [2-14].

Indeed, RBCs are also devoid of organelles and of any *de novo* lipid synthesis capacity, which makes their lipidome rather stable in comparison to other cell types. Indeed, phospholipid synthesis is known to be active in reticulocytes and suppressed in mature RBCs [15]. Nonetheless, alteration of lipid homeostasis is strictly tied to membrane reorganization during RBC ageing *in vivo* and *in vitro* (RBC storage), mainly owing to lipid peroxidation phenomena which promote membrane shape alterations through the progressive loss of lipids (and membrane-associated proteins) via vesiculation [16-18]. Therefore, it is small wonder that the RBC lipidome has long attracted a great deal of interest over the last five

***Corresponding author:** Lello Zolla, E-mail Address: zolla@unitus.it; Phone Number: +39 0761 357 100; Fax Number: +39 0761 357 630

decades.

Yet in 1959, Phillips and Roome provided a preliminary portrait of the human RBC phospholipidome [2]. However, it was only in 1960 that Hanahan and colleagues described a more complex scenario, also by including species-specific differences between human and bovine RBCs [3]. Four years later, Ways and Hanahan reported a detailed lipid class composition of normal human RBCs, indicating the following percentages: cholesterol 25%, choline glycerophosphatides 30%, sphingomyelin 24%, ethanolamine glycerophosphatides 26%, and serine glycerophosphatides 15% [5]. Meanwhile, Farquhar and Ahrens [4] had showed that 67% of the PE, 8% of the PS, and 10% of the lecithin of human RBCs are in the plasmalogen form, with a vinyl ether linkage at the sn-1 and an ester linkage at the sn-2 position. In 1967, Dodge and Philips described a silicic acid thin-layer chromatography strategy to investigate the phospholipid and phospholipid fatty acids and aldehydes in human RBCs [6]. Thirty-three fatty acids and five aldehydes were separated and tentatively classified into lipid classes, including phosphatidyl ethanolamine (PE), phosphatidyl serine (PS), lecithin, and sphingomyelin (SM) 24:0 and 24:1, while fatty acid moieties were tentatively attributed. Of note, the values reported by Dodge and Philips [6] were consistent with those by Ways and Hanahan [5]. Interesting results were obtained also as far as it concerned the composition in fatty acid moieties of the different lipid classes. About 37% of the total fatty acid in PS was 18:0, while only about 3% was 16:0; in PE and lecithin, 16:0 was the major saturated fatty acid, with the level in lecithin being over twice that in PE. The relative amount of 18:1 was also much lower in PS than in PE and lecithin. The fatty acid distribution of sphingomyelin differed markedly from that of the glycerophospholipids (GP), in particular in the greater degree of saturation [6]. Only about 33% of the fatty acids were unsaturated; in addition, less than 6% of the fatty acids appeared to have more than one double bond and less than 3% more than two double bonds. The 16:0, 24:0, and 24:1 made up almost 7% of the total fatty acids. Essentially all of the 24:0 and most of the 24:1 of the human RBC phospholipids appeared to reside in sphingomyelin.

Different instrumentations and techniques have been tested for the improvement of lipid analysis. During the last two decades, big technological strides have prompted the dissemination of chromatography separation and mass spectrometry-based lipidomics studies of RBCs [9-14]. At the dawn of the mass spectrometry-based lipidomics era the complexity of the lipidome did not enable comprehensive studies like the ones performed with thin layer chromatography (TLC) or gas chromatography (GC) described in the previous paragraphs [2-8]. The expensive instrumentation and the lack of bioinformatic tools to handle the high-throughput amount of data collected via the mass spectrometry-based workflow hampered at first its diffusion in the field [1]. More recently, the introduction of highly accurate and less expensive instruments (in comparison to the ones available decades ago) was also paralleled by consistent improvements in the field of bioinformatic elaboration of the raw mass spectra [1]. The acquired expertise have helped laboratories worldwide to cope with the intrinsic difficulties related to lipid mass attribution and fueled new efforts to bring about the systematic classification of lipid species and structures [19, 20]. The current burgeoning of OMICS disciplines has thus given new verve to the field of lipidomics research, while enabling further steps forward.

Regarding RBC lipid homeostasis, as premised by Farquhar and Ahrens [4], lipid composition of human RBCs is largely influenced by the diet. In this view, Dougherty and colleagues performed an extensive investigation to relate region specific diets to the lipid content of plasma, platelets and RBCs [8]. By comparing RBCs of individuals from rural areas in Finland, Italy (province of Viterbo) and the United States, the Authors demonstrated how diets largely relying on fish and olive oil consumption (in Finland and Italy, respectively), resulted in a significant decrease (in comparison to the US counterparts) in the levels of polyunsaturated fatty acids (PUFA), which they relate to the potential production of unhealthy prostaglandins (thromboxane and prostacyclins) byproducts [8]. Finally, the Authors also noted that in all plasma and RBC glycerolphospholipids, the monounsaturated fatty acids (especially oleic acid 16:1 and palmitic acid, 16:0) were highest in the Italian and the saturated fatty acids were highest in the Finnish samples. In this frame, we exploit novel databases such as LIPID MAPS and *ad hoc* software suites for mass spectrometry-based metabolomics analyses (such as MAVEN [21]) to address the key biological issue of the RBC lipidome. Our investigation shares some features with the study by Dougherty and colleagues [8], for it was performed on RBCs collected from subjects living in the province of Viterbo, where olive oil consumption represents a central aspect of the diet. We further address the RBC storage issue (from a lipidomic standpoint) as to conclude that wider transfusion medicine-relevant studies should be carried out to investigate whether inter-regional donor differences might lie upon peculiar RBC lipidomic profiles, which in turn are likely to reflect the heterogeneity of local alimentation regimes across Italy.

2. Materials and Method

2.1. *Sample collection*

Red blood cell units were drawn from healthy donor volunteers according to the policy of the Italian National Blood Centre guidelines (Blood Transfusion Service for donated blood) and upon informed consent in accordance with the declaration of Helsinki. We studied RBC units collected from 10 healthy donor volunteers [male=5, female=5, age 39.4 ± 7.5 (mean ± S.D.)]. RBC units were stored for up to 42 days under standard conditions (CDP-SAGM, 4°C), while samples were removed aseptically for the analysis on a weekly basis (at 0, 7, 14, 21, 28, 35 and 42 days of storage).

2.2. *Untargeted Metabolomics Analyses*

2.2.1. Metabolite extraction

For each sample, 0.5mL from the pooled erythrocyte stock was transferred into a microcentrifuge tube (Eppendorf ® Germany). Erythrocyte samples were then centrifuged at 1000g for 2 minutes at 4°C. Tubes were then placed on ice while supernatants were carefully aspirated, paying attention not to remove any erythrocyte at the interface. Samples were extracted following the protocol by D'Alessandro et al. [22]. The erythrocytes were resuspended in 0.15mL of ice cold ultra-pure water (18 MΩ) to lyse cell, then the tubes were plunged into a water bath at 37°C for 0.5 min. Samples were mixed with 0.6mL of -20°C methanol and then with 0.45mL chloroform. Subsequently, 0.15mL of ice cold ultra-pure water were added to each tube and they were transferred to -20°C freezer for 2-8 h. An equivalent volume of acetonitrile was added to the tube and transferred to refrigerator (4°C) for 20 min. Samples with precipitated proteins were thus centrifuged for 10000 x g for 10 min at 4 °C .

Finally, samples were dried in a rotational vacuum concentrator (RVC 2-18 - Christ Gmbh; Osterode am Harz, Germany) and re-suspended in 200 µl of water, 5% formic acid and transferred to glass auto-sampler vials for LC/MS analysis.

2.3. *Rapid Resolution Reversed-Phase HPLC*

An Ultimate 3000 Rapid Resolution HPLC system (LC Packings, DIONEX, Sunnyvale, USA) was used to perform metabolite separation. The system featured a binary pump and vacuum degasser, well-plate autosampler with a six-port micro-switching valve, a thermostated column compartment. Samples were loaded onto a Reprosil C18 column (2.0mm×150mm, 2.5 µm - Dr Maisch, Germany) for metabolite separation.

For lipids multi-step gradient program was used. It started with 8% solvent A (ddH_20, 20 mmol L^{-1} ammonium formiate; pH 5) to 6% solvent A for 3 min than to 2% solvent A for 35 min and finally to 100% solvent B (methanol) in 30 minutes. At the end of gradient, the column was reconditioned with 8% solvent A for 10 min. The overall run time was 68 min. Column oven was set to 50°C and the flow rate was 0.2 mL/min.

2.4. *Mass spectrometry analysis through microTOF-Q*

Due to the use of linear ion counting for direct comparisons against naturally expected isotopic ratios, time-of-flight instruments are most often the best choice for molecular formula determination. Thus, mass spectrometry analysis was carried out on an electrospray hybrid quadrupole time-of flight mass spectrometer MicroTOF-Q (Bruker-Daltonik, Bremen, Germany) equipped with an ESI-ion source.

MS analysis was carried out in negative ion mode capillary voltage 2800V, nebulizer 45 psi and dry gas of 9 l/min, scan mode 100-1500 m/z. For sample injection, solutions were evaporated to dryness and reconstituted in an adequate volume of methanol:ethanol 1:1. Lipids extracts were prepared by dilution to a concentration of 5 pmol*L^{-1} (where total phospholipids concentration was 2.5 pmol*L^{-1}). Tandem mass spectrometry (MS/MS) is used for glycerophospholipid species structural characterization. Unambiguous species identification is done by analysis of the retention time and fragmentation pattern and through direct comparison against the same parameters, as acquired from chemically defined standards (Avanti Polar Lipids, Inc., Alabaster, Al.), in agreement with Ivanova et al. [23].

Automatic isolation and fragmentation (AutoMS^n mode) was performed on the 4 most intense ions simultaneously throughout the whole scanning period (30 min per run). Calibration of the mass analyzer is essential in order to maintain an high level of mass accuracy. Instrument calibration was performed externally every day with a sodium formate solution consisting of 10 mM sodium hydroxide in 50% isopropanol, water, 0.1 % formic acid. Automated internal mass scale calibration was performed through direct automated injection of the calibration solution at the beginning and at the end of each run by a 6-port divert-valve.

2.5. *Data elaboration and statistical analysis*

In order to reduce the number of possible hits in molecular formula generation, we exploited the in house SmartFormula application of MAVEN [21], which directly calculates molecular formulae based upon the MS spectrum (isotopic patterns) and transition fingerprints (fragmentation patterns). This software generates a confidence-based list of chemical formulae on the basis of the precursor ions and all fragment ions, and the significance of their deviations to the predicted intact mass and fragmentation pattern (within a predefined window range of 5 ppm). Triplicate runs for each one of the 10 biological replicate at day 0 and over storage duration were exported as mzXML files and processed through METLIN/XCMS data analysis software (Scripps Centre for Metabolomics) [24] and MAVEN [21]. Mass spectrometry chromatograms were elaborated for peak alignment (m/zwidth = 0.025, minfrac = 0.5, bw = 5), matching and comparison of parent and fragment ions, and tentative metabolite identification (within a 20 ppm mass-deviation range between observed and expected results against the imported LIPID MAPS database [20] – annotations included adduct ions in positive ion mode). XCMS and MAVEN are open-source software that could be freely used or downloaded from their websites (http://metlin.scripps.edu/download/ and http://genomics-pubs.princeton.edu/mzroll/index.php?show=download).
Quantitative variations were determined against day 0 controls and only statistically significant results were considered (fold change > 2.5 and ANOVA *p-values* < 0.01). Data were further refined and plotted with GraphPad Prism 5.0

(GraphPad Software Inc.).

3. Results and Discussion

HPLC-MS analysis of the RBC lipidome yielded the tentative identification of a huge number of lipid molecules on the basis of accurate intact mass values and retention times (RT) (Supplementary Tables 1-5). Results were further validated against MS/MS feature transitions (fragmentation patterns) for RBC storage time course analyses, where we reported statistically significant variations ($p < 0.01$ ANOVA) of specific lipid molecules over storage duration on a weekly basis in comparison to day 0 controls (Table 1). This helped coping with the difficulties related to the attribution of fatty acid moieties in detected lipids, a problem that hampered major translational applications of early MS-based approaches to the RBC lipidome [9,10].

A 2D map overview of the lipid features identified in a single run is provided in Figure 1, where compound class specific separations are indicated according to the established nomenclature (fatty acids – FA; glycerolipids – GL; glycerophospholipids – GP; sphingolipids – SP; sterols – ST; prenols – PR and polyketides – PK). While FA, GL, GP and SP eluted rather early (within the first six minutes of RT), ST first and PR or PK displayed higher RTs, consistently with their more hydrophobic nature.

In the following paragraphs, we will detail the major findings of the currently proposed investigation through the description of the main distinct lipid classes. Results will be discussed in the light of existing literature in the field.

3.1. *Fatty acids*

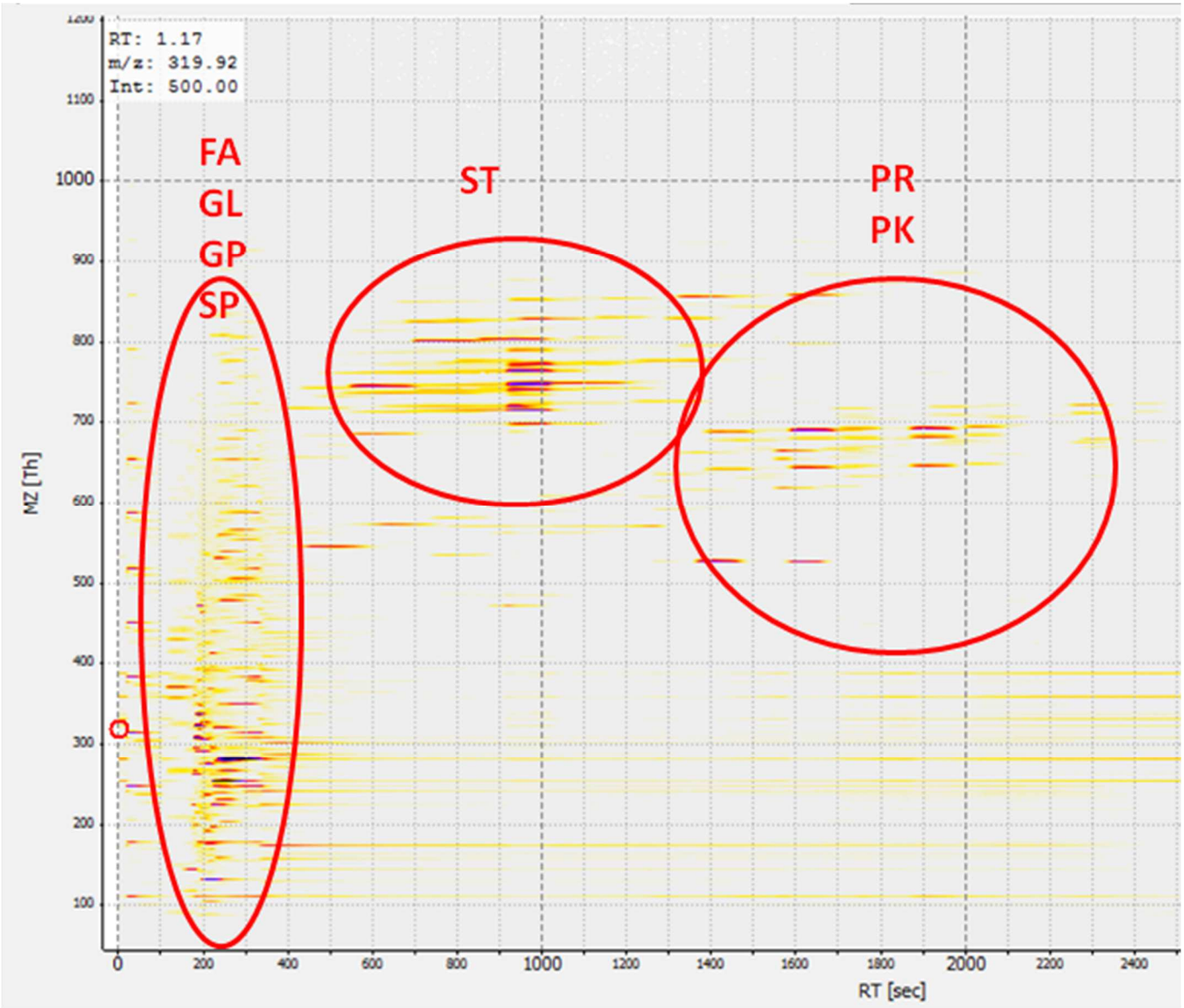

Figure 1. 2D map overview of the lipid features identified in a single run compound. Class specific separations are indicated according to the established nomenclature (fatty acids – FA; glycerolipids – GL; glycerophospholipids – GP; sphingolipids – SP; sterols – ST; prenols – PR and polyketides – PK).

Table 1. Time course alterations of lipids in long stored Red Blood Cells (RBCs). The table reports a weekly comparison of the lipidomes of stored RBCs. Mass spectra chromatograms of day 0, 7, 14, 21, 28, 35 and 42 lipid extract samples were aligned and matched with XCMS software. Statistically significant variations (fold change > 2.5 and p-value < 0.01 ANOVA) against day 0 controls were reported in the table, depending on the specific lipid class (through a direct reference to the LIPIDMAPS database). Red=Increase;

Lipid class		0 vs 7	0 vs 14	0 vs 21	0 vs 28	0 vs 35	0 vs 42
FA		Anandamide (20:l, n-9)	Clupanodonyl carnitine	Cervonyl carnitine		Mayolene-18	11,12-dihydroxy arachidic acid
		4-fumarylacetoacetic acid		Dodecanoylcarnitine		N-(5-hydroxy-pentyl) arachidonoyl amine	omega-hydroxy hendecanoic acid
		3-O-L-rhamnosyl-3-hydroxydecanoyl-3-hydroxydecanoic acid		triacontan-1,14-diol		Triacontan-1,14-diol	15-oxo-18Z-tetracosenoic acid
		1-O-[(6'-O-hexadecanoyl)-aD-glucopyranosyl]-(2-hexadecanoyloxy)-eicosan-1-ol		Hentriacontan-16-one		Hentriacontan-16-one	
				15(S)-15-methyl PGF2α isopropyl ester		(Z)-N-(2-hydroxyethyl) hexacos-17-enamide	
				PGF2α dimethyl amine			
		C24:1-OH Sulfatide		Prostaglandin E2		Prostaglandin E2	
				(1R,2R)-3-oxo-2-pentyl-cyclopentanehexanoic acid		3-methyl-tetradecanedioic acid	
				7-oxo-11E-Tetradecenoic acid		5,13-docosadienoic acid	
DG		DG(17:1(9Z)/17:1(9Z)/0:0)					
		54:3 SLBPA					
TG		TG(20:5(5Z,8Z,11Z,14Z,17Z)/22:3(10Z,13Z,16Z)/22:5(7Z,10Z,13Z,16Z,19Z))[iso6]					
		TG(17:0/18:2(9Z,12Z)/20:0)[iso6]					
GP	PC				PC(O-1:0/O-18:0)		PC(10:0/18:0)
	PI	PIP2(16:0/18:0)	D-myo-Inositol-1,2,4,5,6-pentaphosphate			PI(13:0/0:0)	D-myo-Inositol-1,3,4-triphosphate (sodium salt)
		D-myo-Inositol-1,2,4,5,6-pentaphosphate					
		Inosine 5'-tetraphosphate					
	PE			LysoPE(0:0/22:2(13Z,16Z))		LysoPE(0:0/22:2(13Z,16Z))	

Table 1 Continuation) Red=Increase; Blue=decrease

Lipid class	0 vs 7	0 vs 14	0 vs 21	0 vs 28	0 vs 35	0 vs 42
ST	1-Hydroxyvitamin D3 3-D-glucopyranoside	Vitamin D3 / Cholecalciferol (Vitamin D3) skeleton	1 α -hydroxy-25-methoxyvitamin D3 / 1 α -hydroxy-25-methoxycholecalciferol	Cucurbitacin H	Cucurbitacin H	(6R)-vitamin D2 6,19-sulfur dioxide adduct / (6R)-ergocalciferol 6,19-sulfur dioxide adduct
					1-Hydroxyvitamin D3 cellobioside	
					19-norandrosterone	
	1α,25-dihydroxy-11-(4-hydroxymethylphenyl)-9,11-didehydrovitamin D3 / 1α,25-dihydroxy-11-(4-hydroxymethylphenyl)-9,11-didehydrocholecalciferol	5β-Cholestane-3α,7α-diol	Campesterol	Desmosterol(d6)	Campesterol	
					5a-Tetrahydrocorticosterone	
			1 α -hydroxy-23-[3-(1-hydroxy-1-methylethyl) phenyl]-22,22,23,23-tetradehydro-24,25,26,27-tetranorvitamin D3	(22R)-1 α,25-dihydroxy-22-methoxy-26,27-dimethyl-23,23,24,24-tetradehydrovitamin D3 / (22R)-1 α,25-dihydroxy-22-methoxy-26,27-dimethyl-23,23,24,24-tetradehydrocholecalciferol	Desmosterol(d6)	
		1α-hydroxy-2β-(5-hydroxypentoxy)vitamin D3 / 1α-hydroxy-2β-(5-hydroxypentoxy) cholecalciferol	Ecalcidene		(2-fluoroethyl)-25-hydroxy-26,27-dimethyl-24a-homo-22-oxavitamin D3 / 1α-(2-fluoroethyl)-25-hydroxy-26,27-dimethyl-24a-homo-22-oxacholecalciferol	
			1 α,25-dihydroxy-11-(4-hydroxymethylphenyl)-9,11-didehydrovitamin D3 / 1α,25-dihydroxy-11-(4-hydroxymethylphenyl)-9,11-didehydrocholecalciferol		3α,7α,12α,24-tetrahydroxy-24-methyl-5α-cholestan-26-oic acid	
		1α-hydroxy-25-methoxyvitamin D3 / 1α-hydroxy-25-methoxycholecalciferol	3α,12α,25-trihydroxy-5β-cholestan-7-one	3 α,12 α,25-trihydroxy-5 α -cholestan-7-one	1α,25-dihydroxy-11-(4-hydroxymethylphenyl)-9,11-didehydrovitamin D3 / 1α,25-dihydroxy-11-(4-hydroxymethylphenyl)-9,11-didehydrocholecalciferol	
					1α -hydroxy-22-(3-hydroxyphenyl)-23,24,25,26,27-pentanorvitamin D3 / 1α-hydroxy-22-(3-hydroxyphenyl)-23,24,25,26,27-pentanorcholecalciferol	
			(22R)-1α,25-dihydroxy-22-methoxy-26,27-dimethyl-23,23,24,24-tetradehydrovitamin D3 / (22R)-1 α,25-dihydroxy-22-methoxy-26,27-dimethyl-23,23,24,24-tetradehydrocholecalciferol		Gorgosterol	
					5?-Cholestane-3?,7?-diol	
					Ecalcidene	

Table 1 Continuation) Red=Increase; Blue=decrease

Lipid class	0 vs 7	0 vs 14	0 vs 21	0 vs 28	0 vs 35	0 vs 42
SAC	1-(O-alpha-D-glucopyranosyl)-3-keto-(1,25R,27R)-octacosanetriol			1-O-alpha-D-glucopyranosyl-1,2-nonadecandiol		
PR			30,32-dihydroxy-2b-methyl-bishomohopane	4,4'-Diapo-zeta-carotene	4,4'-Diapo-zeta-carotene	
			(+)-24-Dammarene-3alpha,12beta,20S-triol		2-Hexaprenyl-3-methyl-6-methoxy-1,4 benzoquinone	
			2-Hexaprenyl-3-methyl-5-hydroxy-6-methoxy-1,4-benzoquinone		3-Hexaprenyl-4-hydroxy-5-methoxybenzoic acid	
					4,4'-Diapophytoene	
PK			Delphinidin 3-(6-p-coumaroylglucoside)-5-[6-(malonyl)-4-(rhamnosyl)glucoside)]			
SP Cer			C-8 Ceramide		C-8 Ceramide	C-2 Ceramide
			Ceramide (d18:1/12:0)		Ceramide (d18:1/12:0)	
SP SP			N,N,N-trimethyl-sphingosine	Sphingosine	Phytosphingosine	
				Phytosphingosine	D-erythro-Sphingosine C-15	

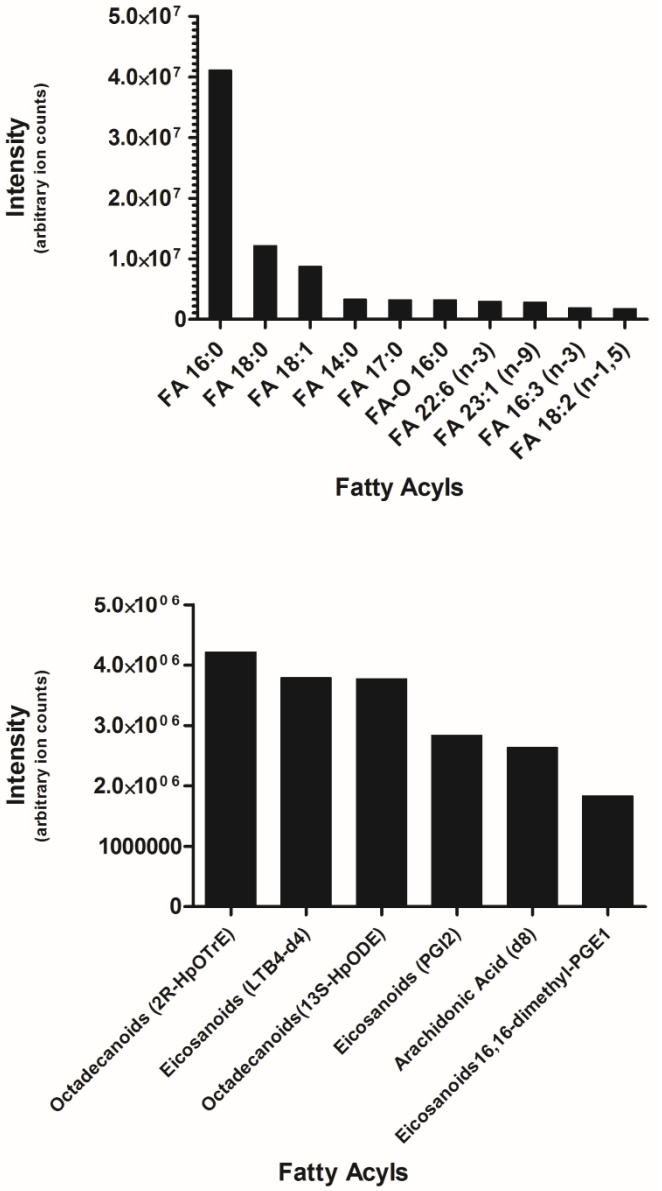

Figure 2. Fatty acid distribution obtained by exporting data from microQtof as mzXML files and processed through MAVEN by interrogating LIPID MAPS database [20].

Fatty acid distribution indicated that palmitic acid (FA C16:0 – Figure 2) was the most abundant free fatty acid (extended results are reported in Supplementary Table 1). This is also consistent with oldest reports on the RBC fatty acid composition available from the literature [5], despite the extreme differences between the TLC and the HPLC-MS analytical approaches. Palmitic acid might be tied to the modulation of calcium signaling in RBCs by mediating Ca^{2+} fluxes via specific membrane pores [25], thereby modulating RBC survival.

Among the most abundant individual fatty acids we could detect 16:0 (palmitic), 18:0 (stearic), 18:1 (oleic), and 22:6 n-3 (docosahexaenoic acid), in agreement with previous studies on fatty acids of erythrocytes obtained from healthy Italian subjects [26]. Furthermore, octadecadienoic acid (18:2 n-1,5) had already been reported among the most abundant ten fatty acids of RBCs [5]. The abundance of oleic acid in particular was an expected result, since olive oil holds a key role in the frame of the Mediterranean diet and, in particular, in the province of Viterbo (Italy) where blood samples were collected from healthy donor volunteers. The intertwinement between oleic acid relative concentrations and high olive oil consumption rates had already been postulated and demonstrated through TLC approaches [8], and hereby confirmed through MS.

On the other hand, no previous investigation indicated myristic acid (14:0) as one of the most abundant fatty acid in RBCs, except for those studies suggesting a role for myristic acid supplementation as a substitute of oleic acid in the diet, which results in the relative increase of α-linolenic and doxosahexaenoic acid levels [27] and alterations of RBC membrane fluidity [28]. Analogously, heptadecanoic acid (17:0) has been proposed as a controversial biomarker for the assessment of energy and macronutrient composition in response to specific diets [29].

Eicosanoids and octadecanoids and their peroxidation products (relative abundances are reported in Figure 1) are thought to play a role in mediating RBC maturation from reticulocytes by promoting the degradation of mitochondrial membranes and thus elimination of these organelles [30]. Also, eicosanoids serve as substrates for cycloxygenase, lipoxygenase and epoxygenase activities, which result in the production of pro-inflammatory factors that are associated with increased cardiovascular risk and cancer [31].

3.2. *Glycerolipids and glycerophospholipds*

Relative abundances of RBCs glycerolipids (GL) and glycerophospholipids (GP) are reported in Figure 3 and 4 (extended results are reported in Supplementary Table 2 and 3), respectively, whereas the latter class has been further subdivided into phosphatidic acid (PA), phosphatidylcholines (PC), phosphatidylethanolamines (PE) and phosphatidylserine (PS), in the light of the observed elevated concentrations in RBCs.

Fatty acid incorporation stages into RBC membrane GLs and GPs has been long investigated [32-34], indicating higher rates for reticulocytes in comparison to adult RBCs [35,36]. Indeed, RBC, GL and GP metabolism is a key aspect in RBC survival [37], since during their 120 days approximate lifespan in the circulatory system RBCs shed approximately 1 microvesicle/h, thus continuously remodeling their membrane and its lipid composition. Also, early approaches to GP composition of RBCs have been
purported via TLC [38] indicated a relation of GP composition with RBC membrane anomalies, such as in the case of spheorcytosis [39]. These considerations are relevant in the light of the incomplete long chain fatty acid synthesizing system which characterizes RBCs [40]. The introduction of high-resolution capillary gas chromatography approaches recently shed new light on this delicate issue [41], further evidencing compositional anomalies of GL and GP in cancer

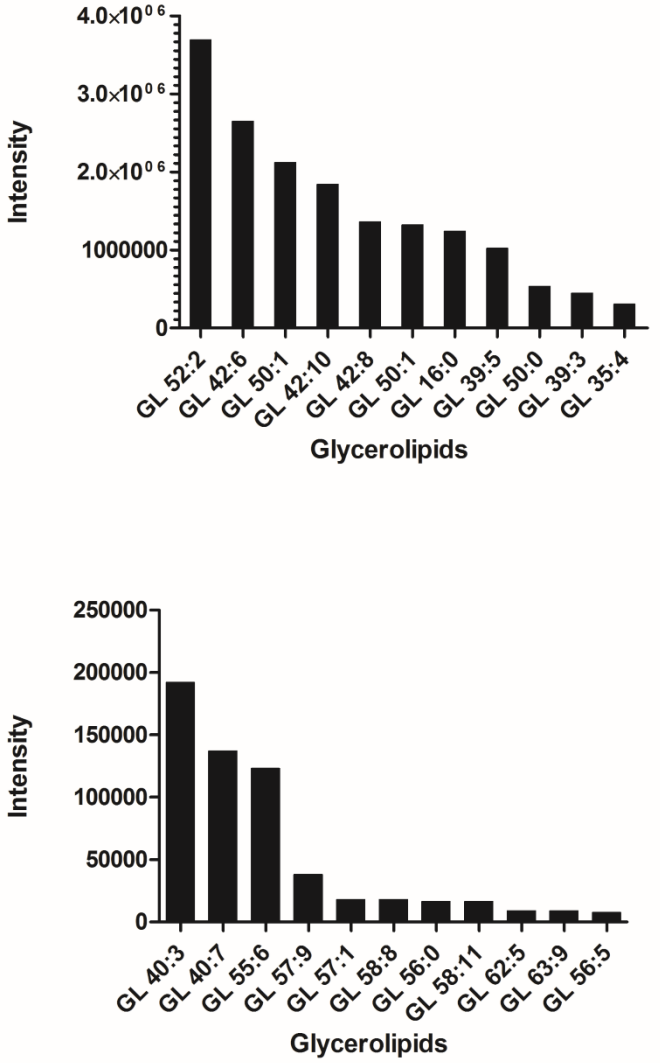

Figure 3. Glycerolipids distribution obtained by exporting data from microQtof as mzXML files and processed through MAVEN by interrogating LIPID MAPS database [20].

evidencing compositional anomalies of GL and GP in cancer patients [42].

Our results provide further supporting evidence about increased levels of elevated levels of C16:0 and C18:0 fatty acids in lyso-PCs from adult RBCs (Figure 4), as previously reported with different approaches [8,41,43]. Analogous considerations can be made for PE 38:4, PE 40:6 and lyso-PE 18:0, as well as for PS 38:4 (Figure 4), in agreement with the literature [8,43]. In particular, PS 38:4 had been recently indicated as the most abundant RBC-specific PS, in comparison to other blood cell types [43]. Distribution of saturated and unsaturated fatty acids in GL and GP was also consistent with the literature [5,8,43].

As expected, compositional differences were observed as well, which are probably attributable to the different diets of the subjects enrolled in the present study in comparison to data available from the literature.

3.3. *Sphingolipids, sterols and prenol lipids*

Sphingolipids (SP), among which ceramides (Cer), have been recently associated with *in vivo* and *in vitro* ageing of RBCs [18]. Though the mechanisms have not yet been fully elucidated, sphinogosines and ceramides seem to affect RBC survival by serving as signaling molecules upon acid sphingomyelinase hydrolysis of sphingomyelin into ceramide [44] or rather they directly affect RBC membrane stability by forming specific pores and thus altering membrane permeability and potential [45]. Ceramide-enriched membrane domains have been indeed associated with hot-cold hemolysis [46]. Besides, SP metabolites, including ceramides, sphingosine and sphingosine 1-phosphate have recently emerged as a new class of lipid biomodulators also in the extracellular space [47].

In Figures 5 and 6 we report the relative abundances of the top SP, with a focus on ceramides, respectively. Of note, C16 sphingosine (Figure 5) has been recently reported to be the most abundant RBC-specific SP [43]. On the other hand, while we expected elevated levels of Cer 24:1 and 24:0, in agreement with Leidl et al. [43], we could instead observe elevated levels of Cer 18:0 and 18:1 in all ceramide subclass (Figure 6), which might reflect the relative composition of free fatty acid, as mentioned in the previous paragraphs.

Recent studies have also demonstrated that sphingolipids dynamically cluster with sterols to form lipid microdomains or rafts, which function as platforms for effective signal transduction and protein sorting [48]. Sterol profiling of RBCs is also a powerful diagnostic tool to investigate the effects of total parental nutrition diet supplementation to the newborn, a lifesaving therapy in children with intestinal failure [49].

An overview of the most abundant sterol lipids is provided in Figure 7, where sterol lipids are reported with their relative name from the Lipid Maps database [20], owing to the impossibility to adapt graphic limitations to the lengthy extended names of each sterol lipid. However, full details are provided in the Supplementary Table 4.

Prenol lipids are often under-investigated class of lipids, which are synthesized from five carbon isoprene units. Recent lipidomics studies focused on plasma levels of dolichols (a group of α-saturated polyprenols characterized by 14 to 24 isoprene subunits) and ubiquinones (a group of 1,4-benzoquinones modified with 9-10 repeated isoprene units) [50]. However, to the best of the Authors' knowledge, little is known about the composition of prenol lipids in adult RBCs.

While in Figure 8 we graphed the relative abundances of prenol lipids on the basis of their relative molecular formula, in Supplementary Table 5 we also provided extended details about their common name and abbreviations, according to the Lipid Maps database nomenclature [20]. However, further dedicated studies are mandatory to shed further lights on the relative concentrations and biological functions of these molecules within the framework of RBC biology.

3.4. *Red blood cell Lipidomics: application to the storage of erythrocyte concentrates*

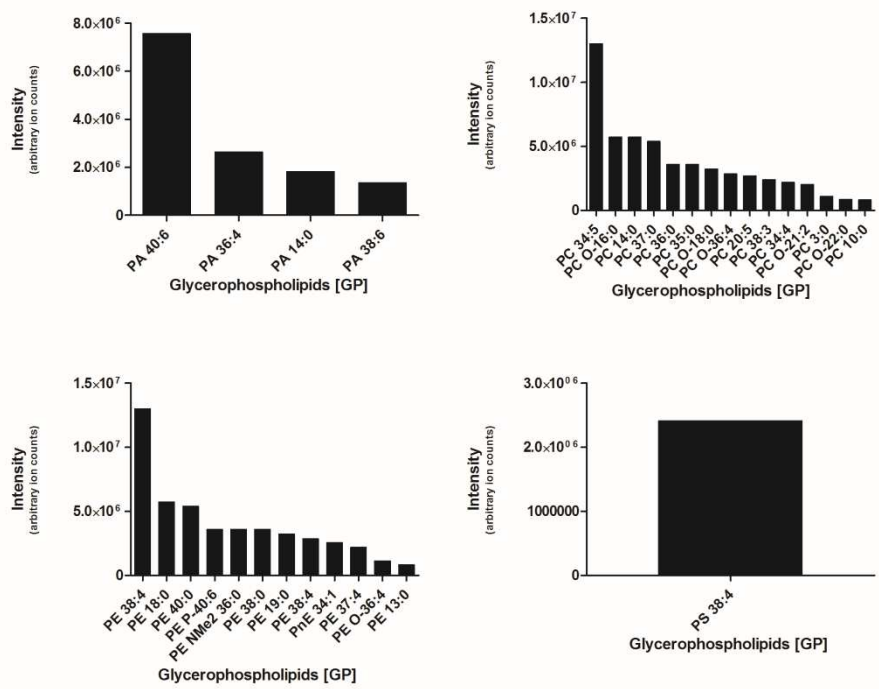

Figure 4. Glycerophospholipids distribution obtained by exporting data from microQtof as mzXML files and processed through MAVEN by interrogating LIPID MAPS database [20].

RBC concentrates for transfusion purposes are routinely stored at 2-6°C for up to 42 days, according to international standard guidelines [51].

Despite decades of substantial improvements in the field of RBC storage [52], concerns still arise and persist about the quality of longer stored RBCs, since it is clearly emerging –at least from a biochemical standpoint– that storage progression corresponds to the accumulation of a wide series of RBCs storage lesions [52], as, among others, we could recently document at the morphologic, metabolomics and proteomics level [53-55]. On the other hand, lipidomic aspects of RBC storage in the blood bank still lie undisclosed. Recently, Bosman's group demonstrated that specific treatment with exogenous sphingomyelinases resulted in the accumulation of ceramides RBC morphological lesions, thereby mimicking the effects of long-stored RBCs [18]. Indeed, accumulation of ceramides and their metabolites (sphingosine and sphingosine 1-phosphate – S1P) might promote intrinsic stimuli leading to the exacerbation of ageing phenomena in RBCs [18], by altering membrane conformation [45] or mediating specific intra- or extra-cellular signaling cascades [44]. In this view, it is worthwhile noting that plasma S1P mainly originates from erythrocytes, since RBCs display alkaline (but not acid or neutral) ceramidase activity on D-e-C(18:1)-ceramide [56]. First of all, we wish to stress that the most abundant ceramides we could detect in control adult RBCs could be catalogued as C18:0 or 18:1 (Figure 6). Moreover, in the present study prolonged storage of RBC was apparently associated with statistically significant decrease (p < 0.01 ANOVA) of ceramides (C-8, Ceramide d18:1/12:0 and ceramide C-2 – Table 1) after three weeks of storage, which is a critical timespan threshold for the accumulation of storage lesions at the biochemical level, as we could previously report at the proteomics and metabolomics level [54,55]. These results are suggestive of a likely ceramidase-mediated digestion of ceramides, or rather of an alteration of the lipid composition of long-stored RBCs probably reflecting the membrane remodeling occurring over RBC storage duration [53]. However, we could also observe a decrease in the levels of several sphingosines (N,N,N-trimethyl sphingosine, sphingosine, phytosphingosine and D-erythro-Sphingosine C-15 – Table 1), which did not help us ruling out any definitive scenario to explain the observed phenomena.

After an initial increase (attributable to phospholipase activities [57]), prolonged storage of RBC concentrates hereby resulted in the progressive statistically significant (p<0.01 ANOVA) decrease of a wide series of fatty acids, prostaglandins (such as PGF2α and prostaglandin E2) and fatty acid oxidation products ((1R,2R)-3-oxo-2-pentyl-cyclopentanehexanoic acid) (Table 1). This is consistent with the reported progressive accumulation of lipid oxidation byproducts in the supernatants of long-stored erythrocyte concentrate units [17,57].

The initial increase in the levels of diacyl-glycerols (DG) and triacyl-glycerols (TG) (Table 1) is difficult to interpret, if not in the light of the need for RBCs to cope with the initial free fatty acid accumulation through their sequestering and accumulation in the form of DGs and TGs. This is consistent with the hypothesis that, whether a Save or Sacrifice mechanism is innate in RBCs, as suggested by *in silico* elaborations [58,59], this mechanism is active within the first two weeks of storage [53-54].

Recently, Bicalho et al. investigated the alterations to the RBCs phospholipidome by performing a direct comparison of fresh RBC phospholipids against the phospholipid composition of RBC-shed microvesicles [60]. As a result, the Authors could point out the alterations of PS 38:4 and PS 38:1 composition in fresh controls and RBC-derived microvesicles [60]. In the present study, while we could confirm previous evidences about PS 38:4 being preponderant in RBCs (Figure 4), also in agreement with Leidl et al. [43], we could not detect any statistically significant variation as far as PS are concerned. On the other hand, we could detect significant decrease in the levels of two PCs (O-1:0/O-18:0 and 10:0/18:0 – Table 1), PEs (lyso-PE(0:0/22:2(13Z,16Z)) and lyso-PE(0:0/22:2(13Z,16Z)) – Table 1), while PIs followed a controversial trend, especially within the first two weeks of storage.

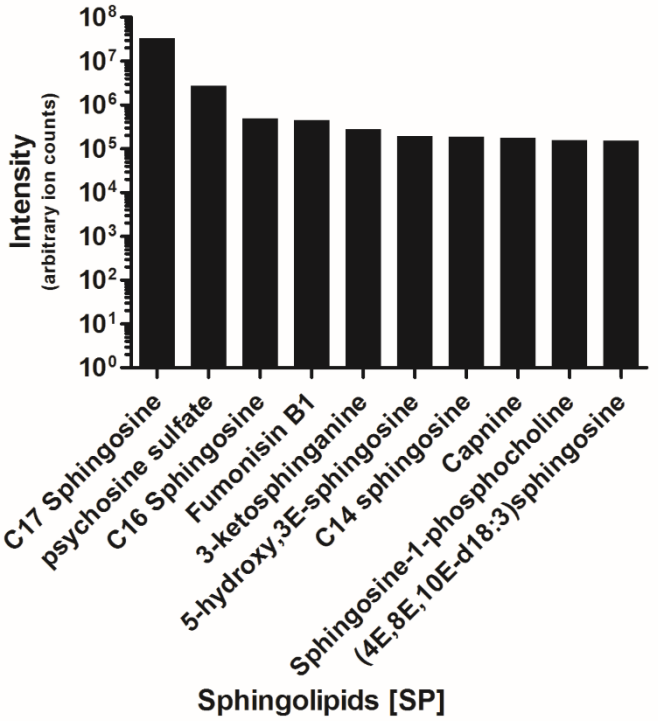

Figure 5. Sphingolipids distribution obtained by exporting data from microQtof as mzXML files and processed through MAVEN by interrogating LIPID MAPS database [20].

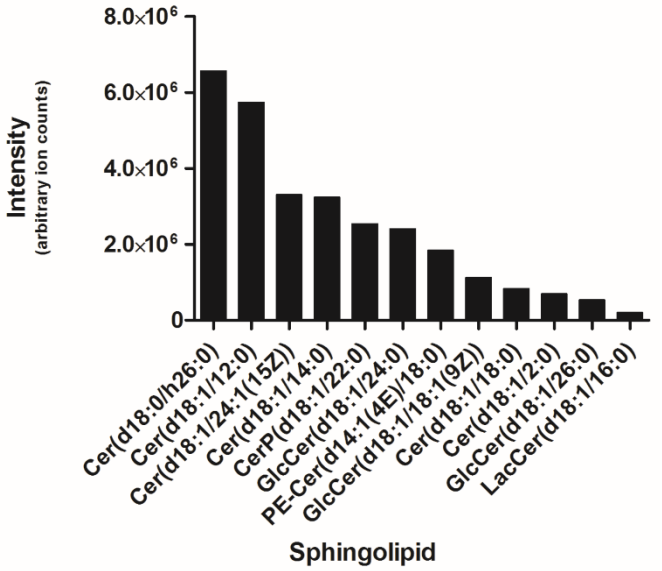

Figure 6. Ceramides distribution obtained by exporting data from microQtof as mzXML files and processed through MAVEN by interrogating LIPID MAPS database [20].

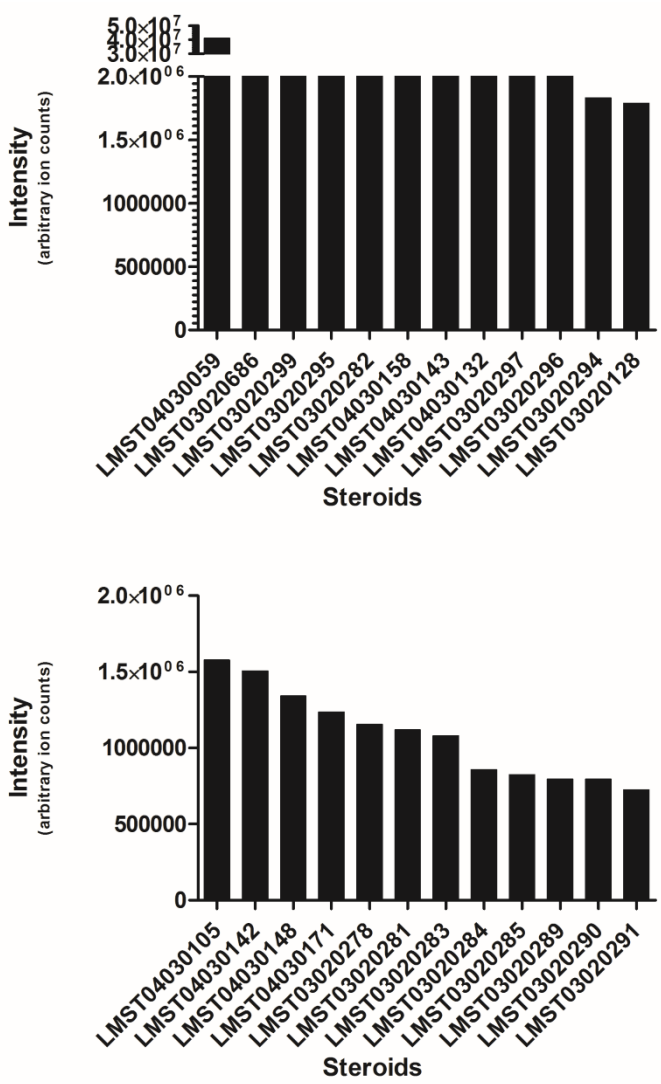

Figure 7. Steroids distribution obtained by exporting data from microQtof as mzXML files and processed through MAVEN by interrogating LIPID MAPS database [20].

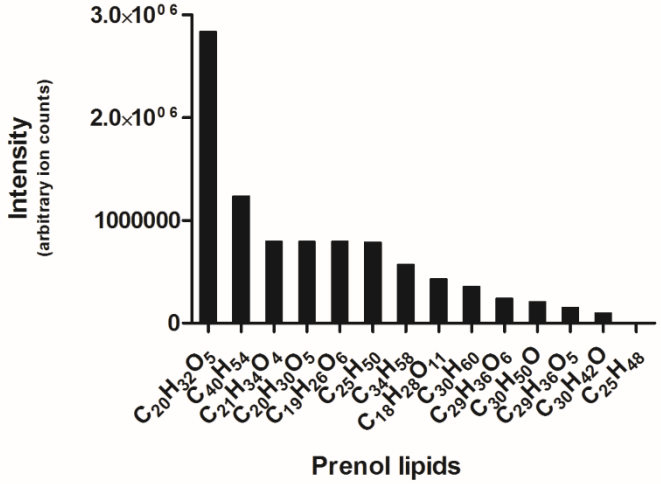

Figure 8. Prenols distribution obtained by exporting data from microQtof as mzXML files and processed through MAVEN by interrogating LIPID MAPS database [20].

Finally, sterols, prenols, saccharolipids and polyketides were hereby investigated for the first time within the framework of RBC storage. Intriguingly, all these classes of lipids statistically significant decreases throughout storage duration (Table 1). Most of the observed decreases account for sterols (e.g. desmosterol, gorgosterol), and in particular for vitamin D3-related metabolites (Table 1). This is relevant in the light of the well-established role for Vitamin D in modulating RBC survival [61], also by influencing anti-oxidant potential and Ca^{2+} permeability [62], a phenomenon which is strictly tied to erythrocyte-specific apoptosis, also known as eryptosis [63].

On the other hand, earliest studies on the likely long term effect of RBC storage on the lipidome suggested that cholesterol loss is limited in comparison to the loss of phospholipids and phosphoinositides [64]. Finally, our results about a generalized decrease in lipid contents of the major lipid classes in long stored RBCs also confirm and expand/complement recent evidences by Acker's group [60,65].

4. Conclusion

Despite decades of investigations, the field of lipidomics recently drained new lymph from the introduction of recent technical innovations. From TLC to gas chromatography and MS, consolidated lipidomics expertise in the field of RBC biology has paved the way for a deeper understanding of the functioning of this pivotal cell and, in parallel, to the accumulation of a wealth of knowledge that will be soon transferred to the clinical setting. Indeed, owing to their relative abundances and widespread biological activities, lipids are well suited to play the role of biological markers and will soon serve this purpose.

In this study, we presented an HPLC-microTOF-Q approach to investigate the RBC lipidome. We could exploit this analytical workflow to consolidate existing knowledge on the RBC lipid composition and individuate statistically significant fluctuations of lipids throughout storage duration of RBC concentrates under blood bank conditions. While this field of research still warrants future investigations, which could be exploiting for example recently introduced imaging mass spectrometry approaches [66], we could indicate ceramides, glycerophospholipids and sterols as key targets of RBC storage lesions to the lipidome, that will deserve further targeted investigations in the future.

Finally, in the light of minor differences with other reports available from the literature, we posited how compositional analyses of the RBC lipidome might end up yielding different results on the basis of the background of the blood donor (above all, the diet), which might translate into region-specific lipidomic alterations over storage progression of RBC concentrates. This is relevant in the light of the constant efforts pursued by transfusion services to improve the quality of blood-derived therapeutics [67], by shifting the focus of attention from the end-product (RBC concentrates) to their providers (the donors).

Acknowledgments

The authors are supported by funds from the Italian National Blood Centre (Centro Nazionale Sangue - CNS - Istituto Superiore Sanita` - Rome, Italy). Angelo D'Alessandro was supported by mobility studentship funds by the Interuniversity Consortium for Biotechnology (CIB). **Funding** TAM, CM, ADA and LZ are supported by funds from the Italian National Blood Centre (Centro Nazionale Sangue - CNS - Istituto Superiore Sanita` - Rome, Italy).

Conflict of interest statement. All the authors disclose any conflict of interest.

Authorship contributions

Timperio Anna Maria – Designed the experiment and performed mass spectrometry analyses.

Cristiana Mirasole – elaborated the results, prepared figures and tables.

Angelo D'Alessandro – wrote the paper.

Lello Zolla – Supervised the experimental design and methods, provided laboratory equipment and contributed with his expertise in the field of biochemistry to the revision of the draft of the manuscript..

References

1. H. Alex Brown, Curr Opin Chem Biol 16 (2012) 221-226. DOI: 10.1016/j.cbpa.2012.02.003.
2. G.B. Philips, N.S. Roome, Proc Soc Exp Biol Med 100 (1959) 489-492. DOI: 10.3181/00379727-100-24670.
3. D.J. Hanahan, R.M. Watts, D. Papajohn, J Lipid Res 1 (1960) 421-432. DOI:jlr.org/content/1/5/421.short.
4. J.W. Farquhar, E.H. Jr Ahrens, J Clin Invest 42 (1963) 675-685. DOI: 10.1172/JCI104759.
5. P. Ways, D.J. Hanahan, Lipid Res 5 (1964) 318-328. DOI:jlr.org/content/5/3/318.short.
6. J.T. Dodge, G.B. Phillips, J Lipid Res 8 (1967) 667-675. DOI:jlr.org/content/8/6/667.short.
7. J.S. Owen, K.R. Bruckdorfer, R.C. Day, N. McIntyre, J Lipid Res. 23 (1982) 124-132. DOI: jlr.org/content/23/1/124.short.
8. R.M. Dougherty, C. Galli, A. Ferro-Luzzi, J.M. Iacono, Am J Clin Nutr 45(1987) 443-455. DOI: http://ajcn.nutrition.org/content/45/2/443.short.
9. X. Han, R.W. Gross, Proc Natl Acad Sci U S A 91 (1994) 10635-10639. DOI: 22/10635/10639.
10. C. Beermann, M. Mobius, N. Winterling, et al. Lipids 40 (2005) 211-218. DOI: 40/L9604/2005.
11. C.M. Skeaff, L. Hodson, J.E. McKenzie, J Nutr 136 (2006) 565-569. DOI: jn.nutrition.org/content/136/3/565.short.
12. P. Rise, S. Eligini, S. Ghezzi, et al. Prostaglandins Leukot Essent Fatty Acids 76 (2007) 363-369. DOI: 10.1016/j.plefa.2007.05.003.
13. E.K. Kabagambe, M.Y. Tsai, P.N. Hopkins, et al. Clin Chem 54 (2008) 154-162. DOI: 10.1373/clinchem.2007.095059.
14. T.P. Novgorodtseva, Y.K. Karaman, N.V. Zhukova, E.G. Lobanova, M.V. Antonyuk, T.A. Kantur, Lipids Health Dis 10 (2011) 82-92. DOI: 10.1186/1476-511X-10-82.

15. A.K. Percy, E. Schmell, B.J. Earles, W.J. Lennarz, Biochemistry 12 (1973) 2456–2461. DOI: 10.1021/bi00737a014.
16. A. D'Alessandro, G.M. D'Amici, S. Vaglio, L. Zolla, Hematologica 97 (2012) 107–115. DOI: 10.3324/haematol.2011.051789.
17. F. Gevi, A. D'Alessandro, S. Rinalducci, L. Zolla, J Proteomics. 76 (2012) 168-180. DOI: 10.1016/j.jprot.2012.03.012.
18. S. Dinkla, K. Wessels, W.P. Verdurmen, C. Tomelleri, J.C. Cluitmans, J. Fransen, B. Fuchs, J. Schiller, I. Joosten, R. Brock, G.J. Bosman, Cell Death Dis 18 (2012) 410-418. DOI: 10.1038/cddis.2012.143.
19. E. Fahy, S. Subramaniam, H.A. Brown, et al. J Lipid Res 46 (2005) 839-861. DOI: 10.1194/jlr.E400004-JLR200.
20. M. Sud, E. Fahy, D. Cotter, A. Brown, E.A. Dennis, C.K. Glass, A.H. Merrill Jr, R.C. Murphy, C.R. Raetz, D.W. Russell, S. Subramaniam. Nucleic Acids Res. 35 (Database issue) (2007) D527-532. DOI: 10.1093/nar/gkl838.
21. M.F. Clasquin, E. Melamud, J.D. Rabinowitz, Current Protocols in Bioinformatics. 14 (2012) 14-11. DOI: 10.1002/0471250953.bi1411s37.
22. A. D'Alessandro, F. Gevi, L. Zolla, Mol Biosyst. 27 (2011) 1024-1032. DOI: 10.1039/c0mb00274g.
23. P.T. Ivanova, S.B. Milne, H.A. Brown, J Lipid Res 51 (2010) 1581-1590. DOI: 10.1194/jlr.D003715.
24. R. Tautenhahn, G.J. Patti, D, Rinehart, G, Siuzdak, Anal Chem 84 (2012) 5035-5039. DOI: 10.1021/ac300698c.
25. K.N. Belosludtsev, A.S. Trudovishnikov, N.V. Belosludtseva, A.V. Agafonov, G.D. Mironova, J Membr Biol 237 (2010) 13-19. DOI: 10.1007/s00232-010-9302-1.
26. V. Pala, V. Krogh, P. Muti, V. Chajès, E. Riboli, A. Micheli, M. Saadatian, S. Sieri, F. Berrino, J Natl Cancer Inst 93 (2001) 1088-1095. jnci.oxfordjournals.org/content/93/14/1088.short.
27. V. Rioux, D. Catheline, E. Beauchamp, J. Le Bloc'h, F. Pédrono, P. Legrand. Animal. 2 (2008) 636-644. DOI: 10.1017/S1751731108001705.
28. H. Dabadie, C. Motta, E. Peuchant, P. LeRuyet, F. Mendy, Brit J Nutr 96 (2006) 283-289. DOI:10.1079/BJN20061813.
29. S.D. Poppitt, P. Kilmartin, P. Butler, G.F. Keogh, Lipids Health Dis 4 (2005) 30-36. DOI: 10.1186/1476-511X-4-30.
30. C. Grüllich, R.M. Duvoisin, M. K. Wiedmann, van Leyen. FEBS Lett. 489 (2001) 51-54. DOI: 10.1016/S0014-5793(01)02080-4.
31. W.L. Smith, R.C. Murphy, Chapter 13 In Biochemistry of lipids, lipoproteins and membranes (4th Edition) 2002; Vance DE and Vance JE Eds. Elsevier. DOI: books.it/booksid=LZprTsjcvIMCpg=PA331d.
32. M.M. Oliveira, M. Vaughn, Incorporation of fatty acids into phospholipids of erythrocyte membranes. J Lipid Res 5 (1964)156-162. DOI: jlr.org/content/5/2/156.short.
33. R.K. Donabedian, A. Karmen, J Clin Invest 46 (1967) 1017-1027. DOI:10.1172/JCI105591.
34. E. Mulder, J. de Gier, L.L.M. Van Deenen. Biochim Biophls Acta 70 (1962) 94-96. DOI:11.1572/BBA133211
35. S.B. Shohet, D.G. Nathan, M.L. Karnovsky, J Clin Invest 47 (1968) 1096-108. DOI: 10.1172/JCI105799.
36. C.Van Gastel, D. Van Den Berg, J. de Gier, L.L.M. Van Deenen, Brit. J Haematol 11 (1965) 193-199. DOI: 10.1111/j.1365-2141.1965.tb06577x.
37. E. Mulder, J.W.O. Van Den Berg, L.L.M. Van Deenen, Biochim. Biophys. Acta. 106 (1965) 118-125. DOI:10.1016/0005-2760(65)90100-1.
38. V.P. Skipski, R.F. Peterson, M. Barclay, Biochem. J 90 (1964) 374-378. DOI: 12.1202627/0003(64)-2.
39. M. Kates, A.C. Allison, A.T. James, 1961. Biochim Biophys Acta 48 (1961) 571-582. DOI: 10.1016/0006-3002(61)90055-5.
40. J.G. Pittman, D.B.J. Martin, Clin Invest 45 (1066) 165-172. DOI: 10.1172/JCI105328
41. V. Jakobik, I. Burus, T. Decsi, Eur J Pediatr 168 (2009) 141-147. DOI: 10.1007/s00431-008-0719-9.
42. N. Mikirova, H.D. Riordan, J.A. Jackson, K. Wong, J.R. Miranda-Massari, M.J. Gonzalez, PR Health. Sci J. 23 (2004) 107-113. DOI: 10.1002/ijc.2910450502.
43. K. Leidl, G. Liebisch, D. Richter, G. Schmitz, Biochim Biophys Acta 1781 (2008) 655-664. DOI:10.1016/j.bbalip.2008.07.008.
44. E.L. Smith, E.H. Schuchman, FASEB J. 22 (2008) 3419-3431.DOI: 10.1096/fj.08-108043
45. L.J. Siskind, S. Fluss, M. Bui, M. Colombini, J Bioenerg Biomembr 37 (2005) 227-236. DOI: 10.1007/s10863-005-6632-2.
46. L.R. Montes, D.J. López, J. Sot, L.A. Bagatolli, M.J. Stonehouse, M.L. Vasil, B.X. Wu, Y.A. Hannun, F.M. Goñi, A. Alonso, Biochemistry. 47 (2008) 11222-11230. DOI: 10.1021/bi801139z.
47. M. Tani, M. Ito, Y. Igarashi. Cell Signal 19 (2007) 229-237. DOI:/10.1016/j.cellsig.2006.07.001.
48. K. Simons, E. Ikonen, Nature 387 (1997) 569-572. DOI: 10.1038/42408.
49. P. Pianese, G. Salvia, A. Campanozzi, O. D'Apolito, A. Dello M. Russo, G. Pettoello-Mantovani, J. Corso, Pediatr Gastroenterol Nutr 47 (2008) 645-651. DOI: 10.1097/MPG.0b013e318170956a.
50. O. Quehenberger, A.M. Armando, A.H. Brown, S.B. Milne, D.S. Myers, A.H. Merrill, S. Bandyopadhyay, K.N. Jones, S. Kelly, et al., J Lipid Res 51 (2010):3299-3305. DOI: 10.1194/jlr.M009449.
51. European Directorate for the Quality of Medicines & HealthCare. Guide to the Preparation,Use and Quality Assurance of Blood Components. Recommendation No. R (95) 15.16th Edition. Council of Europe (2011) DOI:10.1111/j.1423-0410.2002.tb05271.x.
52. A. D'Alessandro. G. Liumbruno, G. Grazzini, L. Zolla, Blood Transfus 8 (2010) 82-88. DOI: 10.2450/2009.0122-09.
53. B. Blasi, A. D'Alessandro, N. Ramundo, L. Zolla, Transfus Med 22 (2012) 90-96. DOI: 10.1111/j.1365-3148.2012.01139.x.
54. G.M. D'Amici, C. Mirasole, A. D'Alessandro, T. Yoshida, L.J. Dumont, L. Zolla. Blood Transfus 10 (2012) 46-54. DOI: 10.2450/2012.008S.
55. R. Xu, W. Sun, J. Jin, L.M. Obeid, C. Mao, FASEB J 24 (2010) 2507-2515. DOI: 10.1096/fj.09-153635.
56. N.A. Neidlinger, S.K. Larkin, A. Bhagat, G.P. Victorino, F.A. Kuypers, J Biol Chem 281 (2006) 775-781 DOI:10.1074/jbc.M50579020.
57. R. Chaudhary, R. Katharia, Blood Transfus 10 (2012) 59-62 . DOI: 10.2450/2011.0107-10.
58. S.R. Goodman, A. Kurdia, L. Ammann, D. Kakhniashvili, O. Daescu, Exp Biol Med (Maywood) 232 (2007) 1391-408. DOI: 10.3181/0706-MR-156.
59. A. D'Alessandro, P.G. Righetti, L. Zolla, J Proteome Res 9 (2010) 144-163. DOI: 10.1021/pr900831f.
60. B. Bicalho, J.L. Holovati, J.P. Acker, Biochim Biophys Acta 1828 (2013) 317-326. DOI: 10.1016/j.bbamem.2012.10.026.
61. M. Alexander, Biomedicine (1977) 108-110. DOI: 10.890015/k.1965-1765.1977.
62. R.P. Holmes, M. Mahfouz, B.D. Travis, N.L. Yoss, M.J.

Keenan, Ann N Y Acad Sci 414 (1983) 44-56. DOI: 10.1111/j.1749-6632.1983.tb31673.

63. G. Pompeo, M. Girasole, A. Cricenti, G. Boumis, A. Bellelli, S. Amiconi, Biochim Biophys Acta 1798 (2010) 1047-1055. DOI: 10.1016/j.bbamem.2010.02.002.
64. T.J. Greenwalt, C. Zehner Sostok, U.J. Dumaswala. Vox Sang 58 (1990) 90-3. DOI: 10.1111/j.1423-0410.1990.tb02068.x.
65. R. Almizraq, J.D. Tchir, J.L. Holovati, J.P. Acker. Transfusion. 2013 DOI: 10.1111/trf.12080.
66. Mainini V, Bovo G, Chinello C, et al. Mol Biosyst. 2013 DOI: 10.1039/C2MB25296A.
67. G. Liumbruno, A. D'Alessandro, G. Grazzini, L. Zolla, J Proteomics 73 (2010) 483-507. DOI: 10.1016/j.jprot.2009.06.010.

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

ORIGINAL ARTICLE | DOI: 10.5584/jiomics.v3i1.115

Comparison of preservation methods for bacterial cells in cytomics and proteomics

Michael Jahn[1], Jana Seifert[2], Thomas Hübschmann[1], Martin von Bergen[2,3,4], Hauke Harms[1], Susann Müller[1]

[1]Department of Environmental Microbiology, UFZ - Helmholtz Centre for Environmental Research, Permoserstr. 15, D-04318 Leipzig, Germany; [2]Department of Proteomics, UFZ - Helmholtz Centre for Environmental Research, Permoserstr. 15, D-04318 Leipzig, Germany; [3]Department of Metabolomics, UFZ - Helmholtz Centre for Environmental Research, Permoserstr. 15, D-04318 Leipzig, Germany; [4]Department of Biotechnology, Chemistry and Environmental Engineering, Aalborg University, Sohngaardsholmsvej 49, DK-9000 Aalborg, Denmark

Received: 19 December 2012 **Accepted:** 20 March 2013 **Available Online**: 29 March 2013

Abstract

Cell sampling during long-term experiments usually requires reliable storage of cells for later analysis. In this study, we evaluated three different preservation strategies (sodium azide fixation, deep freezing and vacuum drying) with regard to their effects on bacterial cells. *Pseudomonas putida* was used as a model organism and stored for shorter (2 d) and longer periods (28 d). The impact of the treatments (preservation method, duration) was evaluated on the level of single cells using flow cytometry and on the population level using protein mass spectrometry. On the single cell level, the effect of sodium azide fixation was found to be small ($1.01 \leq sd \leq 1.76$) for short term and larger for long term storage ($1.59 \leq sd \leq 2.33$), as determined by FlowFP fingerprinting. In contrast, deep frozen and vacuum dried cells showed properties highly similar to fresh reference cells, even after extended storage ($0.5 \leq sd \leq 1.2$). On the population level, the mass spectrometric analysis revealed about 800 proteins for each sample and storage condition. The proteome profiles evaluated by label-free quantification showed that variation within functional groups was least for deep frozen and vacuum dried cell samples after 2 d (sd log2 relative protein quantity < 1) and marginally increased after 28 d. In contrast, sodium azide fixation caused higher variations between functional groups although the number of detected proteins and the respective peptide coverage excluded protein degradation. In conclusion, deep freezing was found to be the method of choice, but simple vacuum drying of cells with storage at 4°C can be a convenient alternative.

Keywords: Fixation; Cryopreservation; Vacuum drying; Bacteria; Flow cytometry; Mass spectrometry.

Abbreviations:

F, fresh**; VD,** vacuum drying**; DF,** deep freezing**; SAF,** sodium azide fixation**; FCM,** flow cytometry**; MS,** mass spectrometry

1. Introduction

The number of samples that can be analyzed by flow cytometry is steadily increasing and recently reached 1536-well plate format as a platform for parallel analysis [1]. The advent of high-throughput analyses raises the question of efficient sample preservation, both for plain flow cytometry and for cell sorting in combination with other *Omics* techniques in microbiology [2]. With such applications in mind, the recovery of whole intact cells after storage is essential for flow cytometric analysis and the reliability of this recovery is particularly important for long-term experiments. The process of sample preservation consists of two parts, sample preparation and sample storage. Here, we define sample preparation as a measure to stop cellular activity and prevent cells from alteration or decay. Usually relatively harsh methods derived from histology were used for preparation, commonly known as fixation. For instance, alcohols and alde-

***Corresponding author:** Susann Müller, UFZ - Helmholtz Centre for Environmental Research, Permoserstr. 15, D-04318 Leipzig, Germany; Phone number: +49 341 235 1318; E-mail address: susann.mueller@ufz.de

hydes have the advantage of inactivating biohazardous specimens [3]. In general, different fixation and preparation techniques have different benefits and drawbacks, as they change the sample in characteristic ways. The risk of further changes caused by subsequent storage until sample analysis is mainly determined by the storage time and temperature [4]. The diversity of preparation techniques and storage conditions is large, and the choice depends on the kind of analysis samples are subjected to.

The most widely-used fixation methods for microbial cells are alcohol- or aldehyde-based methods, which mainly work via protein cross-linking (aldehydes), cell permeabilization and water removal (alcohol) [3, 5]. Although aldehydes permit easy fixation and storage even at room temperature (RT), major drawbacks include alterations of cell morphology, cell aggregation, loss of biomolecules and increased autofluorescence [3, 6, 7]. Another chemical preservation method, which can be also regarded as fixation, is the use of (sodium-) azide, which was proven useful for microorganisms [8, 9]. Azide inhibits the terminal enzyme of the respiratory chain and thereby prevents energy production of aerobic cells [10-12]. Metal salts like barium, nickel and molybdenum in combination with 10% sodium azide proved especially useful for anaerobic bacteria and provided highly reproducible cytometric patterns for at least nine days [13, 14]. These chemical preservation techniques are usually sufficient for flow cytometry, but destructive effects like protein cross-linking in case of aldehydes disqualify a method for protein mass spectrometry. A method widely used in biobanking is cryopreservation by deep freezing (DF) at temperatures below -60°C with addition of a cryoprotective agent, such as glycerol or sugars [15-17]. Deep freezing was suitable for preservation of methane oxidizing bacteria as tested by flow cytometry [18], but performed poorly for natural microbial communities [13]. In the food and pharmaceutical industry, drying of bacterial cells by different techniques is used for product conservation [19, 20]. Freeze drying, spray drying or low temperature (~0°C) vacuum drying were shown to preserve different bacterial strains [21-23], but have not been applied in single cell analysis.

In this study, we compared three different preparation and storage procedures suitable for single cell analysis and simple enough for standard laboratory application: Sodium azide fixation (SAF) followed by storage at 4°C, vacuum drying (VD) followed by storage at 4°C, and deep freezing (DF) with liquid nitrogen and subsequent storage at -80°C. The bacterial strain *Pseudomonas putida* was used as model organism for analysis by flow cytometry (FCM) and by protein mass spectrometry (MS). The combination of both techniques is promising, as cells with different characteristics can be identified and sorted by FCM, and the subpopulation's proteome can be analyzed by MS. Subpopulation proteomics was already performed for an artificial mixture of bacteria [24] and for pure cultures of *P. putida* [2]. Besides MS, FCM can also be coupled with other downstream applications like RNA isolation [25, 26]. In the presented study, FCM and MS were used to evaluate the effect of the selected preservation methods (VD, DF, SAF) on cell characteristics (DNA content, light scatter) and protein profile, respectively, in comparison to fresh samples (F). A gate-free similarity fingerprinting method was applied for analyzing cytometry data [27] and a novel functional clustering algorithm was used for comparison of proteome profiles.

2. Materials and Methods

2.1. *Bacterial strains and cultivation conditions*

Pseudomonas putida KT2440 (source: DSMZ – German Collection of Microorganisms and Cell Cultures) was cultivated overnight at 30°C on a rotary shaker (180 rpm) in minimal medium (Na_2HPO_4 6 g/L, KH_2PO_4 3 g/L, NaCl 0.5 g/L, NH_4Cl 1 g/L, $MgSO_4$ 0.5 g/L, $CaCl_2$ 15 mg/L, $ZnSO_4$ x $7H_2O$ 3.6 mg/L, $CuSO_4$ x $5H_2O$ 0.625 mg/L, H_3BO_3 0.15 mg/L, $FeSO_4$ x $7H_2O$ 6 mg/L, $CaCO_3$ 5 mg/L, $MnSO_4$ x $7H_2O$ 3 mg/L, $CoSO_4$ x $7H_2O$ 0.7 mg/L) with 2 g/L glucose as sole carbon and energy source. A 500-ml shaking flask with 100 ml of minimal medium was inoculated with a volume of overnight culture corresponding to an initial optical density of 0.05 at 600 nm (OD_{600nm}, $d_{cuvette}$ = 0.5 cm). For comparison, *Escherichia coli* DH5α (source: DSMZ) was cultivated overnight at 37°C on a rotary shaker (180 rpm) in LB medium. A 500 ml shaking flask with 100 ml of LB medium was inoculated to an initial OD_{600nm} of 0.05. The growth of the cells was monitored by measurement of OD_{600nm} up to a maximum incubation time of 12 h.

2.2. *Cell preparation and storage*

Cells of *P. putida* were collected every two hours for a total of 12 hours after inoculation covering all cell cycle stages (Figure 1B). For *E. coli,* cells from the lag and exponential growth phases at 0, 1 and 2 hours were used. Samples were taken at various time points by centrifugation of up to 2 ml cell suspension in a microcentrifuge (Heraeus Fresco 21) for 5 min at room temperature and 5,000 x g, and the supernatant was discarded. For DF, the cells were resuspended in 1 ml phosphate buffered saline (PBS; 6 mM Na_2HPO_4, 1.8 mM NaH_2PO_4, 145 mM NaCl, pH 7.2) containing 15 % (v/v) glycerol as cryoprotective agent, incubated for 10 min on ice and shock frozen in liquid nitrogen for subsequent storage at -80°C. For VD, any residual medium was removed and the cell pellet was vacuum dried for 30 min at 30°C using a vacuum concentrator (N-Biotek Micro-Cenvac). The dried cell pellet was stored at 4°C. For SAF, the cells were resuspended in 2 ml 10 % (v/v) sodium azide (NaN_3) and stored at 4°C.

2.3. Cell staining

Fresh (only *P. putida*), deep frozen and sodium azide fixed cells were centrifuged for 5 min at RT and 5,000 x g and the supernatant was discarded. The resulting cell pellets as well

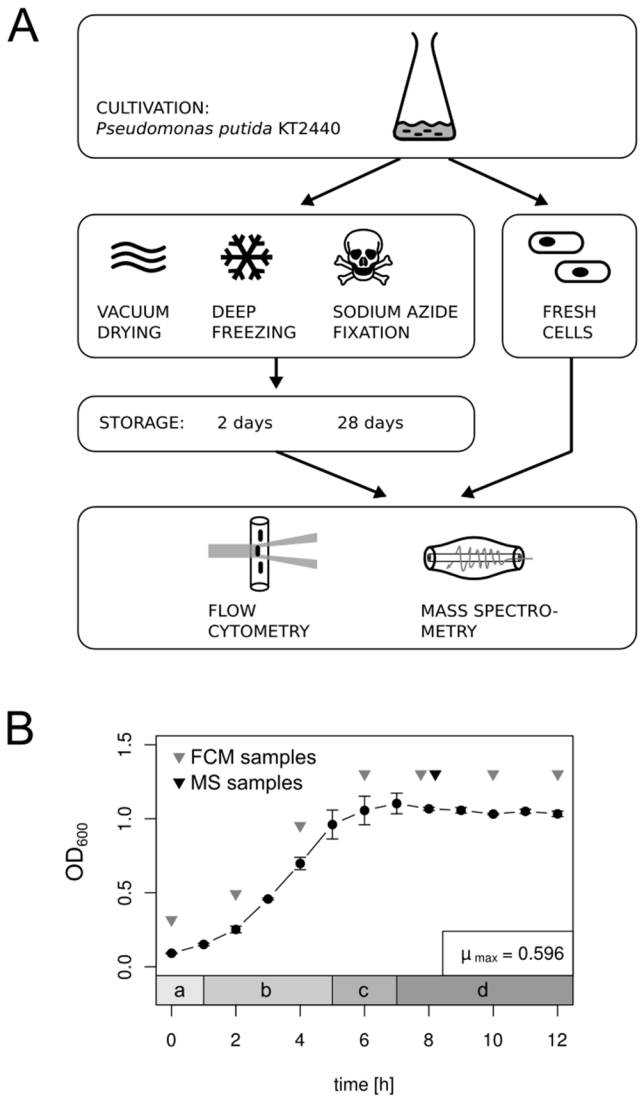

Figure 1. Experimental design of the study. (**A**) Bacterial cultures were grown in shaking flasks, samples were taken at different time points and prepared for storage by vacuum drying (VD), deep freezing (DF) and sodium azide fixation (SAF). The samples were stored for 2 d or 28 d and then analyzed by flow cytometry and mass spectrometry. Fresh samples (F) were directly analyzed and served as reference. (**B**) Growth of *P. putida* KT2440 was monitored by optical density measurement ($OD_{600nm,}$ $d_{cuvette}$ = 0.5 cm) and covered lag phase (a), exponential (b), early stationary (c) and stationary (d) growth phase. Samples for flow cytometry were taken every 2 h (▼) and for mass spectrometry at 8 h of growth (▼).

as VD pellets were resuspended in ice cold PBS by repeated pipetting, and the OD_{600nm} was adjusted to 0.05. For DNA staining, 1 ml of the cells was harvested by centrifugation, and the pellet was resuspended in 1 ml of permeabilization buffer (0.3 M citric acid, 5 g/L Tween 20) and incubated for 10 min on ice. After centrifugation the supernatant was removed and the cells were resuspended in 1 ml DNA staining solution (0.68 µM 4',6-diamidino-2'-phenylindole (DAPI), 0.4 M Na_2HPO_4) [28] and incubated for at least 15 min at RT. Prior to flow cytometry, cell clusters were removed by filtration using a membrane of 50 µm pore diameter (CellTrics, Partec) to prevent clogging of the cytometer nozzle. A pumping of DAPI by fresh cells, as reported for several nucleic acid stains [29], was not observed.

2.4. *Flow cytometry and data analysis*

Flow cytometry was performed using a MoFlo cell sorter (Beckman-Coulter, USA) as described in [24]. The DAPI fluorescence was determined using a multi-line UV laser (333–365 nm, 100 mW) for excitation, and emission was detected in the FL4 channel (450±30 nm). Prior to all measurements, the instrument was adjusted using fluorescent beads and a biological standard. Data were recorded with the Summit v4.3 software (Beckman-Coulter) and further analyzed using the Bioconductor framework for R [30]. The electronic noise was removed (forward scatter, threshold of 25) to prevent bias of similarity analyses. For better comparability, the dominant peaks of the forward scatter (FSC) and DAPI (FL4) channels were normalized (warpSet function, variable peak number, grouping on parameter time) using the Bioconductor *flowStats* package [31]. Similarity fingerprinting with the FSC and FL4 channels (standard deviation method, five recursions) was performed using the *FlowFP* package [27].

2.5. *Identification of proteins by LC-MS-MS*

For proteomics, cells of *P. putida* KT2440 (three technical replicates) were harvested after 8 h of growth and either stored as described above or directly prepared for MS by resuspension in 2 ml PBS. The cell number was determined using flow cytometry and a sample volume corresponding to 1 x 10^8 cells was harvested by centrifugation and resuspended in 25 µl 25 mM ammonium bicarbonate buffer (NH_4HCO_3, pH 7.8) with 1 µl acetonitrile and 5 µl of trypsin (0.25 µg/µl, Promega, Madison, USA) for proteolytic digestion. The samples were incubated over night at 37°C with continuous shaking (180 rpm) and the digestion was stopped by addition of formic acid (FA, 0.1 % (v/v) final concentration). Cell debris was removed by centrifugation for 10 min at 13,000 x g and RT and the supernatant was transferred to a fresh 0.5 ml-tube. Samples were stored at -20°C until analysis. The peptide solution was then purified using the ZipTip protocol (Millipore, Bedford, USA), dried using a vacuum concentrator and the remaining peptides resuspended in FA. The solution was sonicated in a water bath for 5 min prior to injection. Peptides were separated and measured by a high-pressure liquid chromatography (nano-UPLC) system (nano -Acquity, Waters) coupled to an LTQ Orbitrap XL mass spectrometer (Thermo Fisher Scientific, Bremen, Germany) as described in [32]. Continuous scanning of eluted peptide ions was carried out between 300-1600 m/z, automatically switching to MS/MS CID mode on ions exceeding an intensity of 3000.

The retrieved raw data were analyzed by MaxQuant (version 1.2.2.5) [33] with the genome sequence of *P. putida*

KT2440 as the database. The MaxQuant settings can be found in more detail in supplementary Table S1. The label-free quantification (LFQ) values were used for protein quantification and can be found in supplementary Table S3. For peptide mapping, the original mass spectra were further analyzed by Thermo Discoverer (v.1.2.0.208), Mascot (v 2.3) and the NCBInr database (as of February 2013) with a restriction to sequence entries of *P. putida* KT2440 (available in supplementary Table S4).

2.6. *Proteome mapping*

The mean and standard deviation of the obtained LFQ values were calculated for each triplicate. The relative protein quantity was calculated as ratio of the protein quantity of the respective stored sample to the fresh sample (F). To add biological information to the detected proteins, their subcellular localization was predicted *in silico* by PSORTb v3.0 [34]. Furthermore, proteins were annotated using the KEGG BRITE functional hierarchy for *P. putida* KT2440 (http://www.genome.jp/kegg-bin/get_htext?ppu00001.keg, [35]), adding four hierarchical levels (here called 'system', 'process', 'pathway' and 'protein'). Sunburst treemaps were created using a custom recursive function in R (supplementary information S2).

3. Results

3.1. *Cell preparation and storage*

The influence of preservation method and storage time on the stability of bacterial cells was investigated. To this end, three different cell parameters (FSC, SSC and DNA content) were selected as indicators and analyzed by FCM. Likewise, the protein profile of the same samples was analyzed by MS. Samples of *P. putida* KT2440 were taken at various time points during batch cultivation. Hence, the cells display a different morphology, DNA content and possibly also sensitivity to preparation methods due to various growth stages. At every time point, cells were either directly analyzed by flow cytometry (fresh, F) or after preservation by VD, DF or SAF as depicted in Figure 1A. Since VD has not been used for cell preservation in cytometry before, the method was optimized regarding drying temperature and duration. Out of three different drying durations (10, 30, 60 min) and two temperatures (30°C, 60°C), the most distinct distribution (FSC, DNA) was obtained at 30°C for 30 min (supplementary Figure S1). Likewise, DF was tested with 15 and 50 % (v/v) glycerol in PBS as a cryoprotective agent, but no difference was observed (supplementary Figure S2).

3.2. *Analysis of flow cytometric pattern similarity*

Samples of *P. putida* KT2440 from seven time points were prepared by four different methods (F, VD, DF, SAF) and either directly analyzed (F) or stored for 2 d and 28 d (VD, DF, SAF). The light scattering and DNA content of the cells were analyzed by flow cytometry (Figure 2A). The side scatter signal showed no remarkable differences between samples and was disregarded for similarity analysis. Predominantly, the FSC signal showed a unimodal distribution, which shifted according to the growth of the cells. The cellular DNA content changed during cultivation, ranging from a single chromosome equivalent (C1n) to two or more copies

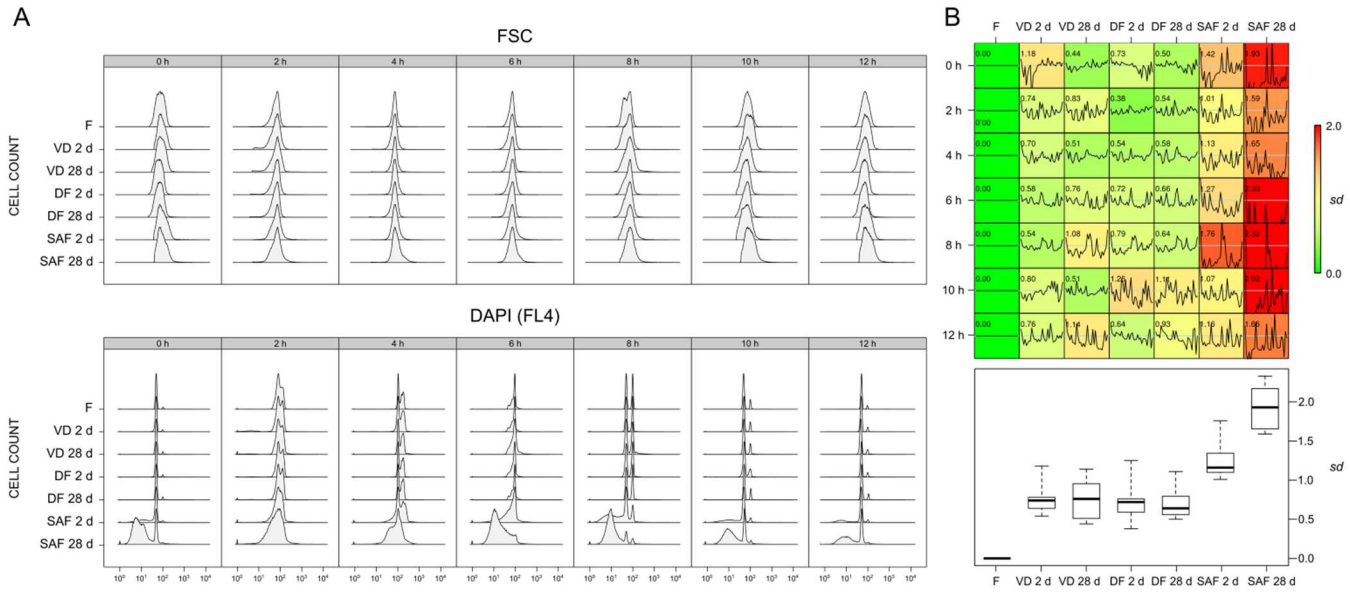

Figure 2. Similarity analysis of stored *P. putida* KT2440 cells using flow cytometry. The cells were either directly analyzed (F) or prepared by vacuum drying (VD), deep freezing (DF) and sodium azide fixation (SAF) and stored for 2 d or 28 d. (**A**) Histograms of forward scatter (FSC, upper panel) and DAPI fluorescence (FL4, lower panel) after 0-12 h of incubation. (**B**) Similarity fingerprint of stored samples in comparison to the fresh sample (F). Given is the standard deviation *sd* as index for bin differences, which increases with increasing dissimilarity. The retrieved *sd* values for each method are summarized in a box plot.

(≥ C2n). The DNA pattern of the stored samples was highly similar to the fresh samples for VD and DF, but SAF treated cells showed a certain proportion of cells below the C1n peak after 2 d, which further increased after 28 d of storage. To quantify similarity of flow cytometric patterns a fingerprinting method (*FlowFP*) was employed, involving FSC and DNA (Figure 2B). It is based on a probability binning algorithm, which constructs a model from a reference sample based on the distribution of events in it. Here, the respective fresh sample (F) of each time point served as the reference sample. This is then compared to the stored samples yielding an index for similarity, the standard deviation of bin differences (*sd*). It is 0 for identical samples and increases with increasing dissimilarity. The sensitivity limit of the method was determined by comparison of two virtual halves of the same sample F, yielding an $sd \leq 0.25$ (supplementary Figure S3). An equal similarity to the fresh reference sample was found for VD and DF ($0.5 \leq sd \leq 1.2$, Figure 2B) independent of the storage duration. Samples stored by SAF were less similar after 2 d ($1.01 \leq sd \leq 1.76$) and particularly after prolonged storage of 28 d ($1.59 \leq sd \leq 2.33$). The growth phase of *P. putida* had only a small influence on the similarity to the reference sample. For comparison, a similar experiment was conducted using another bacterial strain, *E. coli* DH5α. In contrast to *P. putida*, cells of the exponential growth phase were more amenable to alteration than lag phase cells (supplementary Figure S4). Furthermore, the deleterious effect of SAF was much less pronounced for *E. coli* than for *P. putida*.

3.3. *Identification of proteins by MS*

Flow cytometry is a powerful application for analysis and sorting of cells, and can be readily combined with other downstream applications such as proteomics. However, the storage of cells until sorting may influence their protein profile. Therefore, we tested the ability of the selected methods (DF, VD, SAF) to preserve the protein composition of *P. putida* cells for two time periods (2 d, 28 d). Cell samples were acquired at the early stationary growth phase (8 h) and either stored, or instantly analyzed serving as the reference (F). The cellular proteins were analyzed by shotgun mass spectrometry with label-free quantification. A total number of 971 different proteins was detected across all samples, with a range of 793 to 858 different proteins present in at least one replicate per single sample (Figure 3A). The number of proteins common to all samples was 751 (intersection, IS), indicating comparably good protein recovery for all preservation techniques, although the number of recovered proteins across replicates varied by each method (supplementary Figure S5). Particularly, samples stored for 28 d with VD and SAF showed a reduced number of detected proteins per replicate as low as 611 and 713, respectively. Overall, the obtained coverage was in the range of other MS based studies for *P. putida*, which found 604 to 2383 different proteins [36 -39]. Here, the 971 identified proteins showed a different distribution in theoretical subcellular localization (as determined *in silico* by PSORTb) compared to the 5350 proteins annotated at www.pseudomonas.com (Figure 3B). For instance, the proportion of cytoplasmic proteins was higher for

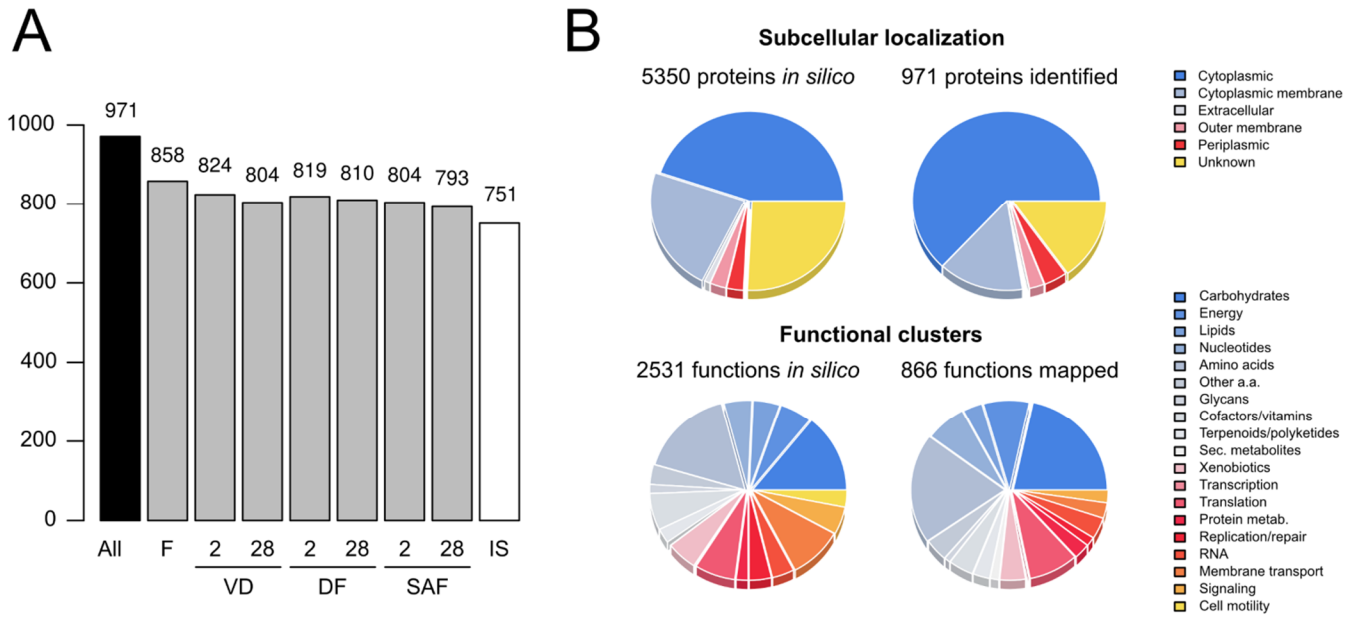

Figure 3. Proteins detected by mass spectrometry after storage of *P. putida* KT2440 cells. (**A**) Number of unique proteins present in at least one replicate per sample for fresh (F), vacuum dried (VD), deep frozen (DF) and sodium azide fixed cells (SAF). The intersection (IS) denotes the number of proteins common in all samples. (**B**) Subcellular localization of all 971 detected proteins in comparison to 5350 protein encoding sequences in the *P. putida* database (www.pseudomonas.com). A subset of the identified proteins was mapped to 866 functions by KEGG BRITE, compared to 2531 functions of all annotated proteins.

the experimental data (63 % instead of 45 %), whereas the proportion of inner membrane localized proteins was smaller (15 % instead of 23 %). To gain a deeper insight into the nature of the identified proteins, 463 out of 971 unique proteins were assigned to 866 functional annotations using the KEGG BRITE database (Figure 3B), which comprised a total of 1494 unique proteins assigned to 2531 functional annotations for *P. putida* KT2440. The clustering of proteins according to the second highest hierarchy level (here called 'process') with 19 subgroups illustrates the over-representation of detected proteins in carbohydrate, energy and amino acid metabolism. An under-representation was observed for proteins involved in membrane transport, signal transduction and cell motility.

3.4. *Protein profile similarity of stored samples*

Besides the number of proteins identified per sample, the relative quantity of the detected proteins was taken into account to test the impact of the three preservation methods on cell protein composition. Therefore, we investigated if certain functional clusters of proteins were more prone to alteration than others. The hierarchical annotation by KEGG BRITE was used to arrange proteins according to the top three out of four hierarchical levels ('system', 'process' and 'pathway'). Protein clusters were drawn as a sunburst treemap (Figure 4), where each layer of the treemap represents one hierarchical level, starting with the broadest level as innermost layer ('system') and ending with the most detailed layer at the surface ('pathway'). The width of a sector corresponds to the number of proteins within the functional group, and the color represents the standard deviation of the $\log_2$ relative protein quantity (*sd*), with the fresh sample (F) as reference. For example, a yellow or red color indicates a high *sd* and therefore a higher or lower quantity of (some) proteins in a group, compared to the proteins of the fresh sample. Thus, storage-induced variations affecting only certain functional groups can be easily spotted. The variation was small within the most groups for VD and DF samples ($sd < 1$), although the pathways for methane, glycine, pyruvate and glyoxylate metabolism as well as xenobiotics biodegradation showed increased variations ($1.2 \leq sd \leq 2.0$). The storage time of 2 or 28 d had no significant influence on the protein profile. However, a stronger overall variation was displayed by SAF treated samples already after 2 d of storage. After 28 d the variation was further increased for proteins involved in oxidative phosphorylation, protein export, ABC transporter and sulfur and lysine metabolism ($1.8 \leq sd \leq 2.2$). The variation in protein quantity within functional groups is an indicator for the impact of the preservation method. To further elucidate the cause for this variation, we analyzed the peptide coverage obtained by MS for 60 selected proteins of different chain lengths across samples (supplementary Figure S6). For these 20 largest, 20 smallest and 20 proteins of medium polypeptide chain length, the peptide coverage was very similar for all samples and no decay at the termini was observed. This finding suggests, that degradation of proteins was not the cause for the deviating protein profile of SAF cells in comparison to DF and VD cells. Furthermore, even a slightly increased number of peptides was detected for the SAF treated cells after 28 d compared to the 2 d stored cells. For VD, the effect was inverse with fewer peptides detected after 28 d compared to 2 d.

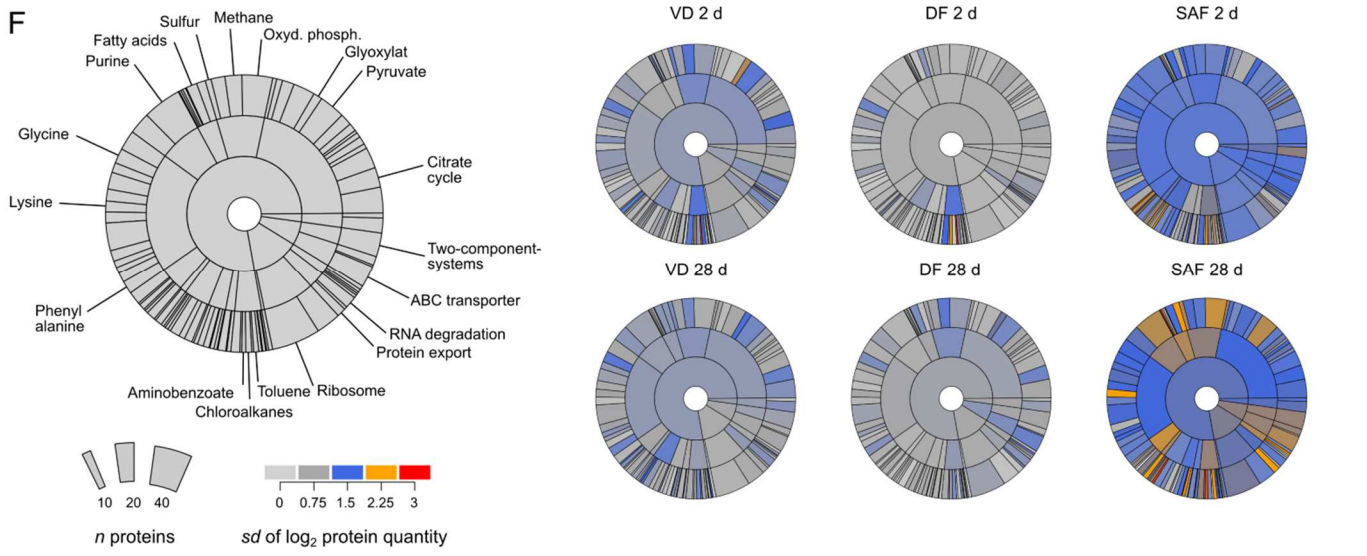

Figure 4. Functional clustering of proteins derived from cell samples prepared by VD, DF and SAF and stored for 2 d and 28 d. The sunburst treemaps represent 463 proteins mapped to functional groups using KEGG BRITE for visualization of sample variation. Treemaps consist of three hierarchically ordered layers ('system', 'process', 'pathway') with the most general ('system') being in the center. Each sector represents a group, with sector width encoding the number of proteins *n* and color encoding standard deviation of the $\log_2$ relative protein quantity (*sd*), compared to the reference (fresh cells, F). Thus, a gray color represents low variation in protein quantity and a yellow or red color a high variation.

4. Discussion

In this study, we tested three different preservation techniques (VD, DF, SAF) for bacterial cells with respect to their influence on sample integrity. These methods were selected for speed, simplicity, and omission of organic solvents. The effect of short-term and long-term storage was compared using *P. putida* as model organism. The two techniques used here –FCM and MS– are state-of-the-art for the analysis of microbes by addressing one specific *Omics* level each, namely cytomics and proteomics. The combination of both techniques can be used to obtain a detailed picture of the cellular protein interior for different subpopulations, as was recently shown for microorganisms [2, 24]. However, this kind of experiments requires the reliable preservation of cell samples before FCM and MS are applied. Using cytometry, we found VD and DF to preserve scatter characteristics and DNA content equally well, with high similarity to fresh cells even after long term storage (28 d). Based on the results for *P. putida*, a threshold of $sd = 1.5$ can be considered appropriate for indicating high sample similarity, as measured by *FlowFP* fingerprinting (Figure 2). The similarity of SAF samples was generally lower, mainly due to differing DNA patterns of the cells.

Likewise, the protein profile of VD and DF samples for *P. putida* showed higher similarity to the fresh sample than that of SAF samples as determined by MS. Interestingly, the peptide coverage of 60 selected proteins showed no decreased numbers of detected peptides for SAF, but even a slight increase after prolonged storage (28 d). A possible cause for this may be that the three dimensional structure of proteins is disturbed or unfolded over time due to the high concentration of sodium azide salt. Depending on the nature of the used salt, cellular proteins may be more amenable for whole cell tryptic digestion under such conditions [40]. Nevertheless, the increase of peptide recovery over time is not a desirable effect, since it might influence the comparability between samples. Regarding VD and DF, the higher similarity to fresh cells in proteome profiles was also reflected on the level of peptide coverage. Most likely, the absence of an aqueous environment (VD) and the very low temperature (DF) preserves protein structure more effectively.

However, functional clustering revealed that proteins of specific pathways are more prone to alteration than others, and these pathways coincided for VD and DF (e. g. glycine and pyruvate metabolism). Similar results were found in a proteomic study with human cells, where protein degradation during cold storage affected not all proteins in a sample and not all sample types equally [41]. Furthermore, the number of detected proteins in two of three replicates of VD and SAF was considerably lower after 28 days of storage, pointing towards the superiority of low storage temperatures (-80° C) over moderate ones (4°C). The loss of culturability of stored bacteria in relation to elevated storage temperature or time is a known phenomenon, both for freeze-dried and deep-frozen cells [15, 17, 20]. But little is known about the integrity of biomolecules in whole cells, when stored at different temperatures. What was shown at least, is the beneficial effect of deep storage temperatures (≤ -80°C) on protein stability for already isolated protein extracts [41, 42].

Most storage procedures for whole cells are optimized for biobanking, aiming at resuscitation of cells after storage. For this purpose, cryopreservation procedures like deep freezing and freeze drying may be the most important preservation methods [4]. But not every specimen is equally cryotolerant. Some microbial genera like *Helicobacter* or *Neisseria* are notoriously difficult to freeze or to recover [4] and complex microbial communities may require a completely different treatment (anaerobic sampling, metal ion treatment) for stabilization [13]. Moreover, the preservation of cells for single-cell analysis or for resuscitation are two different objectives, and the chosen technique is not necessarily suitable for both. However, cryopreservation is a preferred option, as the majority of cells is cryotolerant and the required equipment is affordable even for small laboratories [4]. The storage temperature for cryopreservation should preferably be lower than -20°C, which was reported to result in degradation of serum proteins compared to -80°C [42]. If storage capacities at -80°C are limited, alternatives like freeze drying and low temperature vacuum drying may be considered [22, 23], The vacuum drying procedure applied here preserved the cellular DNA content, light scattering properties, and protein profile with similar efficiency as deep freezing. It requires neither chemical treatment nor other technical effort than a generic vacuum concentrator, and storage of dried cells can take place at 4°C.

Whichever technique is chosen, it is necessary to test the applicability of the desired work flow to the target organism. The bacterial species covered here are not representative for all bacterial genera, but are commonly used in biotechnology. The presented preservation methods are intended for the use in sub-population proteomics, a combination of flow cytometric cell sorting and protein mass spectrometry recently applied for microbes [2, 24, 43]. As the process of cell sorting may impose further stress on recovered cells and change protein abundances independent of storage, it was omitted here. However, the results of this study may be of interest for other analytical disciplines as well. Regarding flow cytometry for instance, other markers such as fluorescent proteins are often used and their function should be conserved by DF [2]. Proteins, however, belong to the more stable biomolecules. And although more 'delicate' biomolecules such as RNA could be stronger affected by unfavorable storage conditions, the findings of this study may very well apply to these biomolecules as well.

5. Conclusions

Three different methods for the preservation of bacterial cells –vacuum drying (VD), deep freezing (DF), sodium azide fixation (SAF)– were evaluated using flow cytometry and protein mass spectrometry. Cells of *P. putida* were stored for 2 d and 28 d and the similarity to a fresh reference sample

quantified by cytometry fingerprinting and proteome profiling. Both DF and VD ensured high agreement between stored and fresh samples as analyzed by flow cytometry, whereas SAF samples showed reduced similarity. Furthermore, 971 different proteins were identified across all samples, and 463 of these proteins were functionally clustered and revealed susceptibility of certain protein groups to alteration, depending on the preservation method. Overall, most of the functional groups displayed low variation in protein quantity for VD and DF samples, but high variation in case of SAF, particularly after 28 d. Interestingly, no peptide decay — for example at the protein termini — was found for SAF but rather an increase in peptide coverage. We assume, that the protein structure is made more amenable for trypsin digestion by the action of sodium azide. Nevertheless, DF and VD are recommended for use in flow cytometry and further downstream applications like protein mass spectrometry, whereas SAF should be avoided for *P. putida*.

6. Supplementary material

Supplementary data and information is available at: http://www.jiomics.com/index.php/jio/rt/suppFiles/115/0.
Supplementary Material includes Figures S1 (Evaluation of vacuum drying (VD) conditions for *P. putida* KT2440 using flow cytometry), S2 (Evaluation of deep freezing (DF) conditions for *P. putida* KT2440 cells using flow cytometry), S3 (Internal variation of identical samples when using FlowFP fingerprinting [27]), S4 (Test of different storage methods for *E. coli* DH5α using flow cytometry), S5 (Variability of protein number and quantity across replicates as identified by mass spectrometry) and S6 (Peptide coverage of selected proteins). Supplementary material further includes S1 for MaxQuant settings, S2 for the R treemap function, and tables S3 and S4 for the list of detected proteins by MaxQuant and by Thermo Discoverer/Mascot, respectively.

Acknowledgements

The work was integrated in the internal research and development program of the UFZ and the CITE program (Chemicals in the environment). The support of the European Regional Development Fund (ERDF/EFRE) and the Sächsische Aufbaubank (Free State of Saxony) is gratefully acknowledged. Jana Seifert and Martin von Bergen were partially funded by DFG SPP1319 and DFG FG1530.

References

1. B.S. Edwards, J. Zhu, J. Chen, M.B. Carter, D.M. Thal, J.J.G. Tesmer, S.W. Graves, L.A. Sklar, Cytometry A 81A (2012) 419–429. http://view.ncbi.nlm.nih.gov/pubmed/22438314
2. M. Jahn, J. Seifert, M.V. Bergen, A. Schmid, B. Bühler, S. Müller, Curr Opin Biotechnol (2012) in press. http://www.sciencedirect.com/science/article/pii/S0958166912001723
3. H.M. Shapiro, Practical flow cytometry, 4th edition, John Wiley & Sons, 2003.
4. P. De Paoli, FEMS Microbiol Rev 29 (2005) 897-910. http://view.ncbi.nlm.nih.gov/pubmed/16219511
5. D. Hopwood, Histochem J 1 (1969) 323-360. http://view.ncbi.nlm.nih.gov/pubmed/4113286
6. S. Müller, G. Nebe-von-Caron, FEMS Microbiol Rev 34 (2010) 554-587. http://view.ncbi.nlm.nih.gov/pubmed/20337722
7. M.A. Perlmutter, C.J.M. Best, J.W. Gillespie, Y. Gathright, S. González, A. Velasco, W.M. Linehan, M.R. Emmert-Buck, R.F. Chuaqui, J Mol Diagn 6 (2004) 371-377. http://view.ncbi.nlm.nih.gov/pubmed/15507677
8. S. Müller, A. Lösche, T. Bley, Acta Biotechnologica 13 (1993) 289-297. http://dx.doi.org/10.1002/abio.370130311
9. S. Müller, Cell Prolif 40 (2007) 621-639. http://view.ncbi.nlm.nih.gov/pubmed/17877606
10. F. Palmieri, M. Klingenberg, Eur J Biochem 1 (1967) 439-446. http://view.ncbi.nlm.nih.gov/pubmed/6061963
11. T. Noumi, M. Maeda, M. Futai, FEBS Lett 213 (1987) 381-384. http://view.ncbi.nlm.nih.gov/pubmed/2881810
12. J. Weber, A.E. Senior, J Biol Chem 273 (1998) 33210-33215. http://view.ncbi.nlm.nih.gov/pubmed/9837890
13. S. Günther, T. Hübschmann, M. Rudolf, M. Eschenhagen, I. Röske, H. Harms, S. Müller, J Microbiol Methods 75 (2008) 127-134. http://view.ncbi.nlm.nih.gov/pubmed/18584902
14. C. Vogt, A. Lösche, S. Kleinsteuber, S. Müller, Cytometry A 66 (2005) 91-102. http://view.ncbi.nlm.nih.gov/pubmed/16003722
15. M.Y. Lin, S.H. Kleven, Avian Dis 26 (1982) 426-430. http://view.ncbi.nlm.nih.gov/pubmed/7049151
16. T. Ahn, S.S. Kang, C. Yun, Biotechnol Lett 26 (2004) 1593-1594. http://view.ncbi.nlm.nih.gov/pubmed/15604803
17. F. Fonseca, M. Marin, G.J. Morris, Appl Environ Microbiol 72 (2006) 6474-6482. http://view.ncbi.nlm.nih.gov/pubmed/17021195
18. S. Hoefman, K. Van Hoorde, N. Boon, P. Vandamme, P. De Vos, K. Heylen, PLoS One 7 (2012) e34196. http://view.ncbi.nlm.nih.gov/pubmed/22539945
19. Y. Wang, R. Yu, C. Chou, Int J Food Microbiol 93 (2004) 209-217. http://view.ncbi.nlm.nih.gov/pubmed/15135959
20. Y. Wong, S. Sampson, W.A. Germishuizen, S. Goonesekera, G. Caponetti, J. Sadoff, B.R. Bloom, D. Edwards, Proc Natl Acad Sci U S A 104 (2007) 2591-2595. http://view.ncbi.nlm.nih.gov/pubmed/17299039
21. A. Spengler, A. Gross, H. Kaltwasser, J Clin Pathol 45 (1992) 737. http://view.ncbi.nlm.nih.gov/pubmed/1401192
22. G. Siberry, K.N. Brahmadathan, R. Pandian, M.K. Lalitha, M.C. Steinhoff, T.J. John, Bull World Health Organ 79 (2001) 43-47. http://view.ncbi.nlm.nih.gov/pubmed/11217666
23. S.A.W. Bauer, S. Schneider, J. Behr, U. Kulozik, P. Foerst, J Biotechnol 159 (2011) 351-357. http://view.ncbi.nlm.nih.gov/pubmed/21723344
24. N. Jehmlich, T. Hübschmann, M. Gesell Salazar, U. Völker, D. Benndorf, S. Müller, M. von Bergen, F. Schmidt, Appl Microbiol Biotechnol 88 (2010) 575-584. http://view.ncbi.nlm.nih.gov/pubmed/20676634
25. J. Achilles, F. Stahl, H. Harms, S. Müller, Nat Protoc 2 (2007) 2203-2211. http://view.ncbi.nlm.nih.gov/pubmed/17853877
26. A. Lemme, L. Gröbe, M. Reck, J. Tomasch, I. Wagner-

Döbler, J Bacteriol 193(2011) 1863-1877. http://view.ncbi.nlm.nih.gov/pubmed/21317319

27. W.T. Rogers, H.A. Holyst, Adv Bioinformatics (2009) 193947. http://view.ncbi.nlm.nih.gov/pubmed/19956416
28. M.L. Meistrich, W. Göhde, R.A. White, J. Schumann, Nature 274 (1978) 821-823. http://view.ncbi.nlm.nih.gov/pubmed/567280
29. M. Walberg, P. Gaustad, H.B. Steen, J Microbiol Methods 35 (1999) 167-176. http://view.ncbi.nlm.nih.gov/pubmed/10192050
30. R.C. Gentleman, V.J. Carey, D.M. Bates, B. Bolstad, M. Dettling, S. Dudoit, B. Ellis, L. Gautier, Y. Ge, J. Gentry, K. Hornik, T. Hothorn, W. Huber, S. Iacus, R. Irizarry, F. Leisch, C. Li, M. Maechler, A.J. Rossini, G. Sawitzki, C. Smith, G. Smyth, L. Tierney, J.Y.H. Yang, J. Zhang, Genome Biol 5 (2004) R80. http://view.ncbi.nlm.nih.gov/pubmed/15461798
31. F. Hahne, A.H. Khodabakhshi, A. Bashashati, C. Wong, R.D. Gascoyne, A.P. Weng, V. Seyfert-Margolis, K. Bourcier, A. Asare, T. Lumley, R. Gentleman, R.R. Brinkman, Cytometry A 77 (2010) 121-131. http://view.ncbi.nlm.nih.gov/pubmed/19899135
32. E. Marco-Urrea, S. Paul, V. Khodaverdi, J. Seifert, M. von Bergen, U. Kretzschmar, L. Adrian, J Bacteriol 193 (2011) 5171-5178. http://view.ncbi.nlm.nih.gov/pubmed/21784924
33. J. Cox, M. Mann, Nat Biotechnol 26 (2008) 1367-1372. http://view.ncbi.nlm.nih.gov/pubmed/19029910
34. N.Y. Yu, J.R. Wagner, M.R. Laird, G. Melli, S. Rey, R. Lo, P. Dao, S.C. Sahinalp, M. Ester, L.J. Foster, F.S.L. Brinkman, Bioinformatics 26 (2010) 1608-1615. http://view.ncbi.nlm.nih.gov/pubmed/20472543
35. H. Ogata, S. Goto, K. Sato, W. Fujibuchi, H. Bono, M. Kanehisa, Nucleic Acids Res 27 (1999) 29-34. http://view.ncbi.nlm.nih.gov/pubmed/9847135
36. Y. Kasahara, H. Morimoto, M. Kuwano, R. Kadoya, J Microbiol Methods 91 (2012) 434-442. http://view.ncbi.nlm.nih.gov/pubmed/23022446
37. S. Yun, G.W. Park, J.Y. Kim, S.O. Kwon, C. Choi, S. Leem, K. Kwon, J.S. Yoo, C. Lee, S. Kim, S.I. Kim, J Proteomics 74 (2011) 620-628. http://view.ncbi.nlm.nih.gov/pubmed/21315195
38. D. Wijte, B.L.M. van Baar, A.J.R. Heck, A.F.M. Altelaar, J Proteome Res 10 (2011) 394-403. http://view.ncbi.nlm.nih.gov/pubmed/20979388
39. D.K. Thompson, K. Chourey, G.S. Wickham, S.B. Thieman, N.C. VerBerkmoes, B. Zhang, A.T. McCarthy, M.A. Rudisill, M. Shah, R.L. Hettich, BMC Genomics 11 (2010) 311. http://view.ncbi.nlm.nih.gov/pubmed/20482812
40. D.L. Beauchamp, M. Khajehpour, Biophys Chem 161 (2012) 29-38. http://view.ncbi.nlm.nih.gov/pubmed/22197350
41. D. Pieragostino, F. Petrucci, P. Del Boccio, D. Mantini, A. Lugaresi, S. Tiberio, M. Onofrj, D. Gambi, P. Sacchetta, C. Di Ilio, G. Federici, A. Urbani, J Proteomics 73(2010) 579-592. http://view.ncbi.nlm.nih.gov/pubmed/19666151
42. D.H. Lee, J.W. Kim, S.Y. Jeon, B.K. Park, B.G. Han, Ann Clin Lab Sci 40(2010) 61-70. http://view.ncbi.nlm.nih.gov/pubmed/20124332
43. F. Schmidt, U. Völker, Proteomics 11 (2011) 3203-3211. http://view.ncbi.nlm.nih.gov/pubmed/21710565

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

ORIGINAL ARTICLE | DOI: 10.5584/jiomics.v3i1.125

Proteome response to heat stress in the Antarctic clam *Laternula elliptica*

Manuela Truebano[a,b], Angel P. Diz[b,c], Michael A.S. Thorne[a], Melody S. Clark[a], David O.F. Skibinski[b]

[a]British Antarctic Survey, Natural Environment Research Council, High Cross, Madingley Road, Cambridge CB3 0ET, UK; [b]Institute of Life Sciences, College of Medicine, Swansea University, Swansea SA2 8PP, Wales, UK; [c]Department of Biochemistry, Genetics and Immunology, Faculty of Biology, University of Vigo, 36310, Vigo, Spain

Received: 13 January 2013 **Accepted:** 20 March 2013 **Available Online**: 30 March 2013

Abstract

The proteome can be regarded as a molecular phenotype, as changes in protein expression patterns have a direct effect on organismal physiology and fitness. The analysis of the proteome can therefore be an invaluable tool for our understanding of the mechanisms underlying phenotypic changes in response to environmental change. However, proteomic studies on thermal stress in marine species have mainly focused on heat shock protein expression, and little information is available for other components of the cellular stress response. This is particularly limiting for Antarctic species, which can lack the ability to induce heat shock protein expression in response to experimentally induced heat stress. The present study analysed changes in protein expression patterns in the Antarctic clam *Laternula elliptica* after exposure to elevated temperatures using two dimensional gel electrophoresis and mass spectrometry. Acute exposure to elevated temperatures had an effect in global protein expression patterns, suggesting that *L. elliptica* has the capacity to alter protein expression in response to heat stress. Changes in the expression of 14 proteins out of 264 analysed were observed in response to different levels of heat stress. Four of the 14 proteins had database matches and were identified as the cytoskeletal protein tubulin and associated chaperone TCP-1, and the enzymes enolase and aldehyde dehydrogenase, part of the minimal stress proteome and involved in redox regulation.

Keywords: Thermal stress; Proteomics; Two-dimensional gel electrophoresis; Antarctic invertebrates.

Abbreviations:

CSR, Cellular stress response; **HSPs**, heat shock proteins; **ANOVA**, analysis of variance; **AIC**, Akaike information criterion; **REML**, Restricted Maximum Likelihood; **FDR**, false discovery rate; **TCP-1**, T-complex polypeptide-1; **GST**, Glutathione s-transferase.

1. Introduction

Genomics has provided an invaluable toolkit to aid in the understanding of organism responses to environmental changes at all levels of biological organization, from molecular to organismal. In this respect, proteomic and transcriptomic analyses are complementary approaches. Whilst much variation in mRNA expression is biologically meaningful, protein expression is more likely to determine the phenotype of an organism [1]. In fact the proteome, the full complement of proteins expressed by the genome of a cell, a tissue or an organism at a specific time point [2], can be regarded as a molecular phenotype [3] as changes in protein expression patterns will have a direct effect on organismal physiology and fitness [4]. The analysis of the proteome can therefore be an invaluable method to advance our understanding of the mechanisms underlying phenotypic responses to environmental change in ecologically relevant species.

***Corresponding author:** Manuela Truebano, 609, Davy Building, Marine Biology and Ecology Research Centre, Plymouth University, Drake Circus, Plymouth, PL4 8AA, UK; E-mail address: mtruebano@gmail.com

However, the main thrust of discovery-driven genomics and particularly proteomics work is oriented towards well-known model species. In ecologically relevant non-model species, a greater number of investigations on the mechanisms underlying phenotypic change have focused on the transcriptome. Recent work has demonstrated that new hypotheses can be tested in proteomics studies in non-model organisms, opening the possibility of "addressing old problems, from a new perspective" (reviewed in [5]). The present paper demonstrates the application of an exploratory proteomics approach in the study of the physiologically important stress response in a non-model Antarctic species.

Non-lethal heat shock results in the activation of the cellular stress response (CSR), a proposed universal response which comprises a series of biochemical changes aimed at maintaining homeostasis [6]. This mechanism protects cells from sudden fluctuations in the environment, and ultimately reacts to the threat of macromolecular change [7]. Some of the best understood components of the CSR are heat shock proteins (HSPs), which play a housekeeping role in the normal functioning of the cell, but are best known for being important molecular mechanisms of stress tolerance [8]. Proteomic studies of thermal stress in marine organisms have thus mainly focused on the expression of HSPs, using 1 -DE or 2-DE (for a review of proteomic studies in the marine environment see [9]). The expression of the inducible HSPs in the presence of a thermal challenge has been identified in most species studied to date, leading to the proposal of a universal response. However, some Antarctic species have been found to be an exception (reviewed in [10]) and while others are able to induce HSP expression in response to elevated temperatures [11], the experimental induction temperature of the inducible form of HSP70 in these animals is in excess of +8°C, a temperature never experienced in the Antarctic marine environment. Whilst some other components of the CSR may be present in Antarctic species, information is currently very limited, and it is possible that other mechanisms differ. Exploratory proteomics approaches, which look at a large number of proteins at any one time of which no *a priori* knowledge is needed, will likely aid in the identification of such mechanisms.

In the Antarctic marine environment, much work has been conducted on the clam *Laternula elliptica* (King and Broderip), a keystone species of the Antarctic marine ecosystem [12,13]. As a highly abundant infaunal filter-feeder, it plays a significant role in benthopelagic coupling [13,14]. It is a relatively large mollusc, growing over 100 mm in shell length [15] and this, allied to its abundance and experimental tractability, makes it an attractive candidate for understanding Antarctic marine ectotherm responses to environmental perturbation. *L. elliptica* is also one of the most thermally sensitive species studied [16-18]. It suffers 50% failure in its ability to burrow, an essential biological activity, under acute exposure at 2–3 °C, and complete loss at 5°C. *L. elliptica* moves vertically in the sediment during normal activity cycles, and needs to rebury when ploughed from the sediment by ice disturbance, hence the ability to bury is considered an essential biological function [19]. The upper lethal temperature was estimated around 9°C when animals were warmed at a rate of 1–2°C per week [16]. Much of the work carried out to study the response to elevated temperatures in *L. elliptica* initially concentrated on physiological responses. This has identified the thermal limits and changes in metabolism in response to temperature and oxygen availability [16,17,20]. Other temperature related studies included investigations on burrowing capacity [19], seasonal energetics [21], lipid radical [22] and reactive oxygen species generation [23]. Molecular analyses included the study of HSP expression [11,24], and the study of antioxidant systems with the characterization of a glutathione s-transferase (GST) [25] and two peroxiredoxin genes [26]. More recently, a microarray was developed to study gene expression changes on individuals exposed acutely to elevated temperatures [27]. This, together with a 454 sequencing project [28] have significantly increased the genomic resources available for this species. At the protein level, an investigation of the effect of thermal stress on antioxidant defense systems using enzyme assays has been carried out [29]. To our knowledge, there are currently no studies on global patterns of protein expression in this species and the techniques used in the present study have not been applied to this, or any other Antarctic marine invertebrate.

In the present study, the proteome of the Antarctic clam *Laternula elliptica* was analysed by two dimensional gel electrophoresis (2-DE) and mass spectrometry (MS) after exposure to elevated temperatures *in vivo*. Changes in protein expression patterns in response to different levels of heat stress were analysed with the aim of i) determining whether exposure to elevated temperatures has an effect of protein expression and ii) identifying proteins associated with the stress response, which may be candidates for stress biomarkers in this species, a representative example of Antarctic marine ectotherms.

2.Material and Methods

2.1. *Sample collection*

L. elliptica specimens were collected by scuba divers at a depth of 10-18m in February 2006 at North Cove, Rothera Point, Adelaide Island, Antarctic Peninsula (67°34'07"S, 68° 07'30"W) and shipped to the British Antarctic Survey facilities in Cambridge, UK. The animals were allowed to acclimate for four weeks in a closed water system at water temperature and salinity of 0±0.5°C and 34±1 p.p.t. respectively. A marine microalgae concentrate (Nannochloropsis, Reed Mariculture) was added to the water on a weekly basis. All necessary permits were obtained. This work was approved by the British Antarctic Survey ethics review committee and meets the requirements of UK legislation. All work on living organisms was conducted under the Antarctic Act (1994) Section 3 permit issued by the UK Foreign and Common-

wealth Office. No endangered or protected species were used.

2.2. *Heat shock experiments*

At the end of the acclimation period, a control and two heat shock experiments were performed at 0, 3 and 9°C as described in an earlier transcriptomic study [27]. Briefly, seawater temperature was raised gradually from 0˚C to 3˚C±0.5˚C or 9˚C±0.5˚C over a 12 h period, after which the animals were left in contact with a sand surface to allow them to burrow. The temperature was maintained at a constant value for 12 h. Treatment temperatures were selected based on previous studies [19] where *L. elliptica* was shown to suffer 50% failure in essential biological activities (burrowing) at 3˚C and survive only a few days at 9˚C. The ability to burrow is therefore a measure of the physiological status in this species, thus animals at 3°C were allowed to bury to determine whether differences observed at the organismal level could be detected at the molecular level. Immediately after the treatment, the number of buried animals was noted, individual animals were dissected and samples from mantle tissue were taken, snap frozen and stored at -80ºC. The procedure was repeated for a control group, for which seawater temperature was maintained constant at 0˚C±0.5˚C. Approximately 90% of the animals reburied at 0°C, 25% at 3°C and none at 9°C (n = 10, 20 and 10 for 0, 3 and 9°C respectively). Animals exposed to 3°C were divided into two groups named 3°C buried and 3°C not buried. No mortalities were recorded. The choice of mantle material was based on previous work which showed that this tissue was one of the most responsive to heat challenge [11] and it also allowed for direct comparisons with a previous gene expression analysis carried out in the same population, tissue and experimental condition [27].

2.3. *Experimental design for 2-DE*

Six 2-DE gels could be run at the same time in the electrophoresis chamber. This influenced the design of the experiment. Six independent biological replicates (i.e. six individual animals) were used per treatment, with the exception of the 3°C buried treatment where only five replicates were available (i.e. only five animals buried during the 3°C exposure period). A subset of all biological replicates was run more than once to generate technical replicates, which allows estimation of experimental variation due to the technique itself. Eight out of the total 23 animals across treatments were run on more than one gel to generate technical replicates. In order to determine whether there was a batch effect, replicates were run on the same day and on different days (i.e. same or different batches of six gels, the maximum allowed by the chamber). Animals 13, 15 and 19 were run on two different days, and animals 1, 7, 9, 17 and 23 were run twice within a day and again on a separate day, permitting one within day and two between day estimates.

2.4. *Preparation of protein mixtures*

Approximately 30 mg of mantle tissue was crushed in N_2 (l) with pestle and mortar, and the powder homogenised in 700 µl lysis buffer (7 M urea, 2 M thiourea, 4% CHAPS, 3% DTT and 2% IPG ampholytes) with a sonicator (Branson digital sonicator 250). After centrifugation at 4ºC for 30 min (15000g), the pellet was discarded and the supernatant stored at -80ºC. Protein concentrations were measured with a 2-D Quant Kit (GE Healthcare) and a total of 90 µg (for analytical gels) or 300 µg (for preparative gels) of protein from each individual was cleaned using a 2-D Clean-Up Kit (GE Healthcare) to remove interfering substances (salt or charged detergents) for the first dimension isoelectric focusing (IEF).

2.5. *2-DE electrophoresis, silver staining and image analysis*

The first dimension, isoelectric focusing (IEF), was carried out on immobilized pH gradient strips (pH 3-10NL/24cm, GE Healthcare) with a horizontal electrophoresis apparatus (Ettan IPGphor, GE Healthcare). Strips were equilibrated in two steps in 10 ml equilibration buffer (6M urea, 2% (w/v) SDS, 75 mM Tris-HCl (pH 8.8), 30% glycerol and 0.001% (w/v) bromophenol blue). In the first step, the strips were washed in equilibration buffer containing 100 mg DTT for 15 min with light shaking. In the second step, the first equilibration buffer was decanted and the wash procedure repeated with an extra 10 ml equilibration buffer with 250 mg iodoacetamide replacing the DTT. After equilibration, strip were loaded along the top of the 12.5% DALT pre-cast polyacrylamide gels and transferred to an Ettan Dalt six Electrophoresis System (GE Healthcare) according to manufacturer instructions (for further details see [30]). Silver staining of the analytical gels was used to visualize protein spots using a modification of the protocol developed by Heukeshoven and Dernick [31]. For mass spectrometric analysis of candidate spots, gels were stained with a modified protocol [32], compatible with mass spectrometry.

Silver staining has a linear range up to two orders of magnitude in terms of quantitative response [33,34]. This limited dynamic range compared to fluorescently labeling methods is in part due to the well known "bleaching" effect of silver staining on spots for highly abundant proteins. After spot filtering, when saturated and negatively stained spots were eliminated, the difference between the weakest and strongest spots fell within this range, as observed in a previous study carried out under similar conditions [30]. Silver-stained gels were scanned to TIFF files using an Image Scanner (Amersham Pharmacia Biotech). Progenesis Samespots v2.0 software (Nonlinear Dynamics Ltd) was used for gel alignment, spot detection and spot volume measurement. Visual inspection was incorporated into the filtering process as detailed in [35]. The "lowest on boundary" method for background subtraction was applied, as recommended by Nonlinear Dynamics Ltd. For analysis, the total number of fea-

tures on the gel (including spots and artifacts) was reduced from over 1000 to 264 spots. In each treatment, only spots present in >50% of the animals, and all technical replicates in available, were selected. Absolute spot volumes were normalised using a modification of a method previously described for one channel microarray data [36]. The gel with the lowest median spot volume was selected as a baseline. A normalisation factor was calculated for each gel by dividing the mean spot volume in the baseline gel (i.e. for the total number of spots in the gel retained after filtering) by the mean spot volume of any gel to be normalised. Individual spot volumes within a gel were multiplied by the normalisation factor calculated for that gel. Normalised spot volumes were $\log_2$ transformed for further analysis. Normalisation removes variation arising from different overall staining intensities between gels and allows comparison of individual spot volumes between gels.

2.6. *Statistical analysis*

Data were analysed to identify individual spots for which volume showed statistically significant between treatments, as well as to test for an overall treatment effect on global protein expression. In addition, sources of variation were analysed. Statistical analysis was carried out using R version 2.8.1 [37]. Normalised spot volumes were used throughout the analysis. The $\log_2$ transformed normalised spot volumes gave acceptable fit to normality and homogeneous variance values.

2.7. *Treatment comparisons*

Comparisons between treatments were performed on a spot by spot basis and also using a global protein expression approach. A spot by spot analysis was performed to determine those spots with a statistically significant treatment effect. The characteristics of the dataset (unbalanced design with some animals replicated in multiple days) determined the statistical method applied. Two linear mixed effects (lme) models were fitted to the data for each spot and implemented with the lme package in R [38]. As multiple animals were sampled per treatment, and a number of them replicated over several days, *Animal* and *Day* were included as random factors. For the first model, *Treatment* was not included as a factor, and for the second model it was specified as a fixed factor whilst *Animal* and *Day* remained in the model as before. To test the null hypothesis that for each individual spot there are no differences in spot volume between treatments, the fit of the two models was compared for each spot using a likelihood ratio test implemented with the ANOVA (analysis of variance) function in R. The Akaike information criterion (AIC) and the log likelihood ratio and associated probability were used to compare the models. A p-value of less than 0.05 was considered statistically significant, revealing differences between treatments in the normalised spot volume of that particular spot. For those spots showing statistically significant differences between treatments, further comparisons between treatments were made in a pairwise manner by fitting an lme model to each spot using Restricted Maximum Likelihood (REML). A single model was considered in which *Treatment* was specified as fixed and *Day* and *Animal* as random factors. Since multiple tests were carried out, false discovery rate (FDR) corrections were applied to account for potential type I errors after Benjamini and Hochberg [39]. Fisher´s method for combining probabilities was also applied to the 264 p-values from each pairwise treatment comparison [40]. This method combines the *a priori* p-values of all the spots into a single p-value for the data as a whole, which is then compared with the chosen significance level α value. If the combined p-value is statistically significant, it can be concluded that for at least one of the spots in the list, the null hypothesis that all individual spots show no difference between treatments is false; the best candidates for showing a treatment effect being those with the lowest p-values.

An additional complementary approach to provide evidence of treatment effects on global expression patterns was to determine whether there is a correlation between treatments in patterns of protein expression compared with the control. A 2x2 contingency table was created with the number of observations for which spots were 1) up-regulated in both treatments compared to the control, 2) down-regulated in both treatments , 3) up-regulated at 3°C, then down-regulated at 9°C and 4) down-regulated at 3°C, then up-regulated at 9°C. Fisher's exact test was used to test for association in the 2x2 table.

2.8. *Analysis of technical variation*

Analysis of technical variation was carried in two ways. Firstly, to test for differences in between and within day technical variation, factorial two-way ANOVAs were carried out on normalised spot volume with the variables *Spot* and *Gel* as random factors for each animal for which technical replicates were available. *Spot* represents all selected spots within a gel and *Gel* represents technical replicates for a single animal. The interaction *Spot*Gel* thus measures the technical error. Separate ANOVAs were carried out to extract the within day and between day technical errors (i.e. the same animal run on different gels on the same day or on different gels on different days). The ratio of the mean square of the two measures of the technical error (within and between days) and associated probabilities were calculated. *P*-values less than 0.05 in a variance ratio test indicate statistically significant differences in the technical variation measured between and within days. Secondly, for all animals replicated within or between days, the coefficient of determination (r^2) of normalised spot volume was calculated for each replicated animal across spots between and within days. This measures the technical error between and within days. Values can be compared with those reported in previous studies.

2.9. *Mass Spectrometry*

Gel preparation and mass spectrometric analysis were carried out at the mass spectrometry facility at the Cambridge Center for Proteomics at Cambridge University. All protein spots showing statistically significant changes between treatments in *a priori* tests were manually excised from the gel and subsequent sample preparation was performed in a MassPrep Station (Micromass). Briefly, proteins were reduced and alkylated with dithiothreitol (DTT) and iodoacetamide (IAA) and subjected to enzymatic digestion with porcine trypsin (Promega). After digestion, 10 µl of supernatant was pipetted into a sample vial and analysed by LC-MS/MS. All LC-MS/MS experiments were performed using an Eksigent NanoLC-1D Plus (Eksigent Technologies, Dublin) HPLC system and an LTQ Orbitrap Classic mass spectrometer (ThermoFisher). Separation of peptides was performed by reverse phase chromatography with the conditions described in [41]. The Orbitrap was set to perform in data dependent acquisition (DDA, Top3) mode where all m/z values of eluting ions were subjected to a survey scan in the Orbitrap mass analyzer, set at a resolution of 7500 over an m/z range 360-1500. Peptide ions with charge states of 2^+ and 3^+ were then automatically isolated and fragmented in the LTQ linear ion trap by collision-induced dissociation and MS/MS spectra were acquired. Resulting MS/MS spectra were processed using Bioworks Browser (version 3.3.1 SP1, ThermoFisher) and resulting mgf files were submitted to the Mascot search algorithm (version 2.2, Matrix Science) and searched against the NCBI (all entries) database with the following parameter settings: a fixed modification of carbamidomethyl cysteine, a variable methionine oxidation and two miscleavages. Average atomic masses were used in the searches and a tolerance of 0.8 Da for fragment ions and 2.0 Da for precursor ions was allowed. Putatively identified proteins were examined and only protein matches with two or more unique and statistically significant peptide matches were accepted. When different isoforms/homologies were identified, these were ranked according to their protein score and the match with the highest score was accepted. Observed molecular weights (MW) based on the spot position on the gel were used for confirmation with a 10 kDa tolerance. To further validate the identifications, all peptide sequences from each spot were used to search against a local *L. elliptica* database previously generated using 454 pyrosequencing (peptide against nucleotide database) [28]. Those contigs that had a match to one or more peptides were then identified using sequence similarity searching [42] against the NCBI nr database.

3. Results

3.1. *Analysis of technical variation*

Results of the analysis of within and between day technical variation are shown in Table 1. Apart from animal A1 there is no clear evidence that the mean square for between day replicates is significantly greater than that within days. For both of A9 and A17, one comparison is statistically significant and one is not. Unexpectedly, for A7, technical variation measured within days is greater than that measured between days, though this is not statistically significant. There is therefore no indication that technical variation is consistently higher between days, but appears to be sample dependent (i.e. not batch or treatment dependent). The coefficient of determination (r^2) for normalised spot volume for each animal for which technical replicates were run ranges from 0.97 to 0.92 for between day technical error, and from 0.97 to 0.95 for within day technical error.

Table 1. Comparisons of between and within technical error using ANOVA analyses of normalized spot volume with Spot and Gel as random factors. The Spot*Gel interaction for within day technical replicates (MSwithin) and between day technical replicates (MSbetween) is indicated for each animal. Degrees of freedom (df), F ratios for between divided by within Spot*Gel mean square values and associated probabilities (P) are also shown.

Animal	MS_{within}	$MS_{between}$	df	F ratio	*p-value*
A1	0.15	0.35	263	2.26	<0.001
		0.40	263	2.59	<0.001
A7	0.53	0.36	263	0.68	0.999
		0.42	263	0.79	0.972
A9	0.42	0.56	263	1.32	0.012
		0.39	263	0.93	0.722
A17	0.28	0.27	263	0.97	0.597
		0.39	263	1.38	0.005
A23	0.22	0.23	263	1.06	0.319
		0.23	263	1.06	0.319

3.2. *Comparisons of spot volumes between treatments*

Mean protein expression levels in treatment groups containing six animals each (five in the case of 3˚C buried) were compared. Of the 264 spots that passed quality screening, 14 (Figure 1) showed statistically significant differences in expression between at least two of the treatments at a p-value less than 0.05 in *a priori* tests (Table 2). The AIC for both models, log likelihood ratios and associated p-values are indicated for each spot. The AIC [43] balances the number of parameters in the model, against the maximised likelihood for the model, and can be used to determine the simplest model that best explains the data. Given a dataset and several competing models to explain it, the model with the lowest AIC can be considered the best. For the spots shown, AIC values are lower when the treatment effect is included in the model. Thus the model including treatment effect can be considered the preferred explanation for the data. After FDR corrections, none of the 14 spots showing statistically significant differences between treatments in *a priori* tests remain statistically significant. However, Fisher´s method for combining probabilities is statistically significant (p=0.04) for

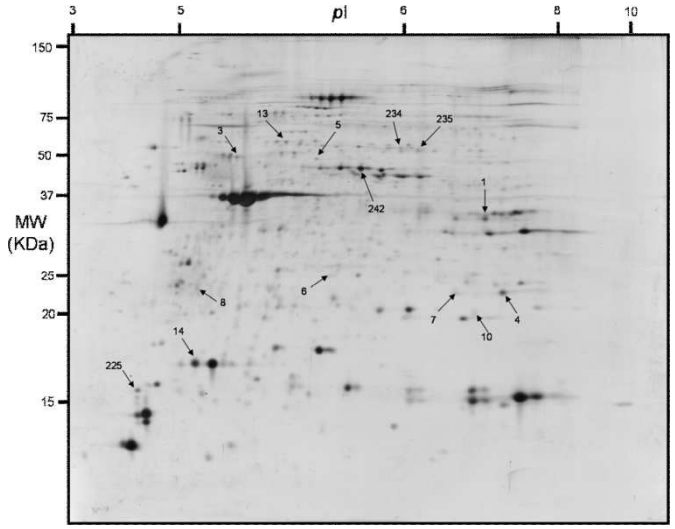

Figure 1. Representative 2DE gel from *L. elliptica* mantle tissue exposed to control conditions (0°C). The isoelectric point (pI) and molecular weight (MW) in kilodaltons are indicated on the horizontal and vertical axes respectively. Arrows indicate the 14 spots that showed significant differences between at least two of the temperature treatments (3°C and 9°C) or the treatments and the control.

Table 2. Results from the lme tests for model comparisons showing spot numbers for spots with expression levels significantly different between treatments are given, followed by the Akaike information criterion (AIC) for models 1 (no fixed effect) and 2 (treatment as a fixed effect), log likelihood ratios (L. Ratio) and associated *a priori* p-values.

Spot	AIC 1	AIC 2	L. Ratio	*p-value*
1	68.56	66.27	8.28	0.041
3	85.14	79.50	11.64	0.009
4	118.06	116.02	8.04	0.045
5	71.65	67.44	10.21	0.017
6	86.79	83.49	9.29	0.026
7	91.37	87.70	9.67	0.022
8	91.04	88.61	8.43	0.038
10	81.29	78.46	8.82	0.032
13	74.00	71.68	8.32	0.040
14	63.74	60.32	9.42	0.024
225	115.29	112.91	8.38	0.039
234	83.74	81.03	8.71	0.034
235	43.73	40.95	8.78	0.032
242	46.87	42.22	10.65	0.014

one of the pairwise treatment comparisons suggesting that at least one of the 264 null hypotheses of no treatment effect is false. The best candidates for further analysis would be the spots with the lowest *p*-values. Some of these will be false positives but this should not be of great concern in exploratory research, as they would be eliminated by repeat experimentation [44].

REML statistics (*t*-value) and associated probabilities from comparisons between treatments are shown in table 2. Mean fold changes in protein expression of each treatment in relation to the control, and to each other, are also shown for all spots showing statistically significant changes in expression between any of the treatments in *a priori* tests. More spots show statistically significant changes in their level of expression between 9˚C and all other temperature treatments than between any of the treatments and the control (Table 3). The two 3˚C treatments (buried and not buried) show the lowest number of proteins with statistically significant changes in expression between them.

In addition to differences in the expression of individual spots between treatments, the analysis using Fisher's exact test suggests that there is a correlation between the treatments across spots in the direction of changes in expression in relation to the control (P<0.001 one tailed Fisher´s exact test). That is, spots are up-regulated (or down-regulated) compared to the control in all treatments more often than this is expected by chance, assuming random variation in direction of expression.

3.3. *Protein identifications by MS/MS*

Protein identifications by mass spectrometry were attempted only on the 14 spots showing statistically significant differences between treatments prior to corrections, as the identification of all spots in the gel was beyond the scope of this study. Of these, four were identified by MS/MS (28.6%). Results are shown in table 4. All spots are found approximately within their theoretical MW (±10 kDa). The four spots are identified as alpha tubulin, aldehyde dehydrogenase, T-complex polypeptide-1 (TCP-1) alpha subunit and enolase.

4. Discussion

In the present exploratory study, changes in the expression of 14 proteins out of 264 analysed in *L. elliptica* exposed to different levels of heat stress were observed to be statistically significant in *a priori* tests. Although these do not remain statistically significant after application of the FDR method, Fisher's combined probability test and application of the Akaike information criterion provide evidence of a treatment effect. In addition, there is a statistically significant correlation between the treatments in the up- and down-regulation of spots in the treated animals compared to the control, which also provides evidence of a treatment effect. Environmental changes are likely to cause coordinated changes in the up- or down-regulation of different proteins, and may be a means of discovering protein networks underlying physiological or adaptive response as observed in a 2-DE study of the effect of mutational change in yeast [45]. Four of the 14 proteins were identified by mass spectrometry. The potential significance of the observed changes is explained and candidates for further analysis are proposed. The coefficient of determination (r^2) for normalised spot volume for each animal for which technical replicates were run (0.92 to 0.97) was within the range reported in previous studies in bivalves [30,46], indicating that technical variation was consistent with expectations from 2-DE analysis.

Table 3. Mean fold changes (FC) in protein expression of each treatment in relation to the control, and to each other for *L. elliptica* mantle tissue. Results are shown for significantly up- or down- regulated spots between at least two of the treatments in *a priori* tests. Positive and negative values indicate up- and down-regulation respectively. Restricted Maximum Likelihood statistics (t-value) and associated probability (P) are also shown. Bold probabilities indicate significant differences between the treatments in the level of expression of a given spot (P<0.05). "B", buried; "NB", not buried.

	FC			FC			FC		
Spot	**3°CB/0°C**	**t-value**	**P**	**3°CNB/0°C**	**t-value**	**P**	**9°C/0°C**	**t-value**	**P**
1	-1.35	1.34	0.195	-1.12	0.54	0.594	1.35	1.41	0.176
3	1.31	1.26	0.223	1.36	1.50	0.151	-1.39	1.54	0.141
4	-2.10	2.12	**0.047**	-2.03	2.09	0.050	-1.06	0.18	0.862
5	1.41	2.02	0.058	1.56	2.69	**0.015**	1.06	0.32	0.751
6	1.16	0.72	0.481	1.59	2.24	**0.037**	1.75	2.61	**0.017**
7	-1.19	0.77	0.452	-1.15	0.65	0.522	1.54	1.90	0.072
8	1.94	2.87	**0.010**	1.59	2.04	0.055	1.59	1.96	0.065
10	1.59	2.36	**0.029**	1.70	2.77	**0.012**	1.67	2.55	**0.020**
13	-1.46	2.05	0.054	-1.17	0.89	0.383	-1.61	2.56	**0.019**
14	9.94	0.03	0.976	10.31	0.21	0.836	6.02	2.40	**0.027**
225	-1.89	1.88	0.075	-1.84	1.86	0.079	1.10	0.28	0.786
234	1.52	1.91	0.071	1.78	2.71	**0.014**	1.77	2.58	**0.018**
235	1.03	0.28	0.780	1.20	1.56	0.135	1.37	2.56	**0.019**
242	-1.19	1.28	0.216	1.17	1.21	0.242	1.24	1.60	0.127
	FC			**FC**			**FC**		
Spot	**3°CNB/3°CB**	**t-value**	**P**	**9°C/3°CB**	**t-value**	**P**	**9°C/3°CNB**	**t-value**	**P**
1	1.20	0.85	0.408	1.82	2.73	**0.013**	1.52	1.99	0.061
3	1.04	0.20	0.846	-1.82	2.89	**0.009**	-1.90	3.20	**0.005**
4	1.04	0.11	0.913	1.98	2.00	0.060	1.91	1.96	0.065
5	1.11	0.67	0.513	-1.33	1.75	0.097	-1.48	2.43	**0.025**
6	1.36	1.60	0.125	1.51	2.02	0.058	1.11	0.51	0.618
7	1.03	0.14	0.891	1.83	2.80	**0.011**	1.78	2.73	**0.013**
8	-1.23	0.94	0.357	-1.23	0.90	0.377	1.00	0.00	1.000
10	1.07	0.38	0.709	1.05	0.26	0.801	-1.02	0.11	0.916
13	1.24	1.29	0.213	-1.11	0.57	0.576	-1.38	1.84	0.081
14	1.04	0.17	0.865	-1.65	2.36	**0.029**	-1.71	2.66	**0.016**
225	1.03	0.10	0.925	2.08	2.21	**0.040**	2.02	2.20	**0.041**
234	1.17	0.76	0.457	1.16	0.71	0.486	1.00	0.01	0.989
235	1.16	1.35	0.192	1.32	2.42	**0.026**	1.14	1.16	0.261
242	1.39	2.59	**0.018**	1.47	2.94	**0.008**	1.06	0.45	0.656

In a previous study of transcriptomic responses to heat stress in *L. elliptica* exposed to the same 3°C treatment described herein, a total of 160 unique transcripts, 7.5% of the total number of clones retained for analysis, showed statistically significant changes in expression between treatments [27]. The reasons for the disparity between the transcriptome and proteome studies in the number of statistically significant changes in expression could be related to the subset of transcripts and proteins analysed. Only a subset of the transcriptome was printed on the array (8500 clones) and likewise, only a subset of the total proteome expressed at any one time can be separated under the same running and preparation conditions. It is possible that the set of transcripts on the arrays by chance showed relatively more differences between treatments than did the set of proteins visualised by 2-DE. In addition, it is likely that some changes occurring in mRNA expression do not result in changes at the corresponding protein level. Several studies have shown that predictions on protein expression levels based on changes in the corresponding mRNA abundance are not always met (reviewed in [5]). It is also possible that there is a time lag between the generation of mRNAs and subsequent protein production in response to changes in temperature. However, when taken as a proportion of the total number of proteins changing in expression between the treatments, the percentage of identified proteins (approximately 29%) was higher than that of identified transcripts (17% in the microarray and 454 sequencing work previously referred to).

A higher number of spots showing changes in expression in the 9˚C treated animals when compared to other treatments could be expected. At 9˚C, the critical temperature has been exceeded and the transfer to anaerobic metabolism has occurred [16], potentially resulting in different protein expression profiles and a larger set of proteins showing sta-

Table 4. Protein spots identified by MS/MS in *L. elliptica* mantle. Mascot protein score, number of peptides, % sequence cover, theoretical molecular weight (MW), sequences for retained peptides and allocated function are shown.

Spot number	Protein	Uniprot ID	Protein score	Peptides	Sequence cover (%)	Theoretical MW	Peptides	GO annotation
3	Tubulin, alpha 1c	Q6P8G7	686	17	31.4	50532	K.DVNAAIATIK.T	Microtubule based movement
							R.QLFHPEQLITGK.E	
							R.TIQFVDWCPTGFK.V	
							K.VGINYQPPTVVPGGDLAK.V	
							R.AVCMLSNTTAIAEAWAR.L	
							R.AFVHWYVGEGMEEGEFSEAR.E	
							R.FDGALNVDLTEFQTNLVPYPR.I	
5	Aldehyde dehydrogenase	A7XZK3	153	3	6.0	53544	K.ILGLIESGK.K	Oxidation reduction
							R.EEIFGPVQQILK.Y	
235	T-complex protein 1, alpha subunit	Q4AE76	269	5	7.6	59962	R.LLEVEHPAAK.V	Protein folding
242	Enolase	A8DU76	233	9	10.8	40012	R.ICDDELIIIK.G	Glycolysis
							R.TQNVMAASSIANIVK.S	
							K.IQIGMDVAASEFCK.D	
							R.AAVPSGASTGIYEALEMR.D	

tistically significant changes in expression. However, the differences in physiological performance (buried and not buried) reflected at the organismal level was not detected at the protein level (only one protein showed statistically significant differences between the 3˚C buried and not buried groups, table 3). Of the proteins identified, aldehyde dehydrogenase and the glycolytic enzyme enolase changed in their level of expression in *L. elliptica* exposed to elevated temperatures. Both proteins are part of the minimal stress proteome of cellular organisms, a set of proteins that participate in different aspects of the cellular stress response and are ubiquitously conserved in all three super kingdoms [6]. In the absence of stress, aldehyde dehydrogenase is an enzyme in the glycolytic pathway. However, under acute stress, it acts as an oxidoreductase, involved in redox regulation by detoxifying aldehydes. The induction of this protein may be necessary for generating reducing equivalents (NADH, NADPH) that are needed for cellular antioxidant systems, or to respond to the energetic requirements of protein degradation, protein chaperoning, and DNA repair [6]. Both are in accordance with the observed decrease in aerobic capacities [20] and the induction of antioxidant defenses at 3˚C [27].

Tubulin is a cytoskeletal protein and main constituent of microtubules, which consists of many isoforms [47,48]. The upregulation of tubulin in clams exposed to 3˚C is consistent with the induction of cytoskeletal protein transcripts observed in heat stressed *L. elliptica* [27]. It is also likely to be associated with the changes in the expression of the fourth protein identified, T-complex polypeptide-1 (TCP-1), a member of the TRiC family which functions as a chaperone in eukaryotes directing folding of cytoskeletal proteins (reviewed in [49]) and is key in the biogenesis of tubulin in the cytosol [50]. It is thus possible that the upregulation of TCP-1 responds to the demand for tubulin and, based on the gene expression data previously cited, other cytoskeletal proteins. The induction of TCP-1 is likely linked with the induction of a prefoldin transcript observed at the level of gene expression [27]. Prefoldin is a molecular chaperone complex which transfers target proteins to chaperonins, promoting folding of newly synthesised or denatured proteins. The biogenesis of the cytoskeletal proteins actin and tubulin involves the interaction of nascent chains of each of the two proteins with prefoldin and their subsequent transfer to the cytosolic chaperonin containing TCP-1 [50,51]. TCP-1 was highly upregulated during heat stress in the coral *Montastraea faveolata* [52]. Owing to the absence of inducible HSPs at environmentally relevant temperatures, TCP-1 deserves further attention as a potential heat stress induced chaperone in *L. elliptica*. Its role in other Antarctic species and potential use as a marker for thermal stress in the Antarctic marine environment awaits examination.

The number of identified proteins in this study using protein databases (4 out of 14 spots) reflects the amount of information available in public databases for *L. elliptica* (18 proteins at the time of writing), which has also been a feature in previous work in this species [27,28]. Identifications in the

present study were limited to highly conserved sequences. Searches for the 14 proteins showing *a priori* statistically significant differences between treatments against the 454 *L. elliptica* transcriptome did not result in further proteins being identified but confirmed the MASCOT results, increasing confidence in the identifications. Since the individuals used for 454 transcriptome analysis were not exposed to stress, it is possible that genes involved in the stress response are underrepresented. This would emphasise the desirability in transcriptomics studies of constructing databases derived from material exposed to contrasting experimental treatments. Identifying proteins associated with specific environmental stressors, may thus require the generation of further transcriptomic resources.

5. Conclusions

In summary, this exploratory study provides evidence for a treatment effect due to experimentally induced heat stress at the protein level in *L. elliptica*. Four candidate proteins have been identified, providing evidence of the activation of a cellular stress response in *L. elliptica* at 3°C, and a candidate spot (TCP-1) that deserves further attention as a potential stress marker in this species. Successful identifications will likely increase as genomic resources rapidly increase through the development of next generation sequencing projects. A strong statistically significant correlation in the pattern of expression of spots in the treated animals compared to the control was observed. This is an interesting approach, which might be investigated more generally where *a priori* values are not statistically significant, as a means of identifying treatment effects and coordinated regulation of protein networks underlying the physiological response.

Acknowledgements

Financial support was provided by the Marine Genomics Europe (MGE) Network (EU-FP6 contract No. GOCE-CT-2004-505403) and the BAS Q4 BIOREACH/ BIOFLAME core programme. MGE also funded a PhD fellowship for MT, and postdoctoral fellowship for APD. MAST and MSC were supported by the British Antarctic Survey Polar Science for Planet Earth Programme, Adaptation and Physiology Work Package. APD is currently supported by "I. Parga Pondal" fellowship and "Grupos de Referencia Competitiva" (code 2010/80) from Xunta de Galicia (Spain) and Fondos FEDER "*Unha maneira de facer Europa*". The authors would like to thank all members of the Rothera dive team for collecting animals.

References

1. B.B. Rees, T. Andacht, E. Skripnikova, D.L. Crawford , Mol Biol Evol 28 (2011) 1271-1279.
2. M.R. Wilkins, J.C. Sanchez, A.A. Gooley, R.D. Appel, I. Humphery-Smith, D.F. Hochstrasser, K.L. Williams, Biotechnol Genet Eng 13 (1995) 19-50.
3. D.G. Biron, C. Brun, T. Lefevre, C. Lebarbenchon, H.D. Loxdale, F. Chevenet, J.P. Brizard, F. Thomas, Proteomics 6 (2006) 5577-5596.
4. M.E. Feder, J. Walser, J Evol Biol 18 (2005) 901-910.
5. A.P. Diz, M. Martínez-Fernández, E. Rolán Alvarez, Mol Ecol 21 (2012) 1060-1080.
6. D. Kültz, Annu Rev Physiol 67 (2005) 225-257.
7. D. Kültz, J Exp Biol 206 (2003) 3119-3124.
8. M.E. Feder, Am Zool 39 (1999) 857-864.
9. L. Tomanek, Annu Rev Mar Sci 3 (2011) 14.1-14.27.
10. M.S. Clark, L.S. Peck, Marine Genomics 2 (2009) 11-18.
11. M.S. Clark, K.P.P. Fraser, L.S. Peck, Cell Stress Chaperon 13 (2008) 39-49.
12. I-Y. Ahn, Mem Natl Inst Polar Res Spec Issue 50 (1994) 1-10.
13. I-Y. Ahn, S.H. Lee, K.T. Kim, J.H. Shim, D.Y. Kim, Mar Pollut Bull 32 (1996) 592-598.
14. R. Ralph, J.G.H. Maxwell, Mar Biol 42 (1977) 171-175.
15. T. Brey, A. Mackensen, Polar Biol 17 (1997) 465-468.
16. L.S. Peck, H.O. Pörtner, I. Hardewig, Physiol Biochem Zool 75 (2002) 123-133.
17. L.S. Peck, S.A. Morley, H.O. Pörtner, M.S. Clark. Oecologia 154 (2007) 479-484.
18. L.S. Peck, M.S. Clark, S.A. Morley, A. Massey, H. Rossetti, Func Ecol 23 (2009) 248-256.
19. L.S. Peck, K.E. Webb, D.M. Bailey, Func Ecol 18 (2004) 625 -630.
20. H.O. Pörtner, L.S. Peck, T. Hirse, Polar Biol 29 (2006) 688-693.
21. S.A. Morley, L.S. Peck, A.J. Miller, H.O. Pörtner, Oecologia 153 (2007) 29-36.
22. M.S. Estevez, D. Abele, S. Puntarulo, Comp Biochem Phys B 132 (2002) 729-737.
23. K. Heise, S. Puntarulo, H.O. Pörtner, D. Abele, Comp Biochem Physiol C 134 (2003) 79-90.
24. H. Park, I-Y. Ahn, H.E. Lee, Cell Stress Chaperon 12 (2007) 275-282.
25. M. Kim, I-Y. Ahn, J. Cheon, H. Park, Comp Biochem Physiol A 152 (2009) 207-213.
26. H. Park, I-Y. Ahn, H. Kim, J. Cheon, M. Kim, Shellfish Immun 25 (2008) 550-559.
27. M. Truebano, G. Burns, M.A.S. Thorne, G. Hillyard, L.S. Peck, D.O.F. Skibinski. M.S. Clark, J Exp Mar Biol Ecol 391 (2010) 65-72.
28. M. Clark, M.A.S. Thorne, F.A. Viera, J.C.R. Cardoso, D.M. Power, L.S. Peck, BMC Genomics 11 (2010) 362-376.
29. H. Park, I-Y. Ahn, K.I. Park, S. Hyun Antarct Sci 20 (2008) 521-526.
30. A.P. Diz, D.O.F. Skibinski, Proteomics 7 (2007) 2111-2120.
31. J. Heukeshoven, R. Dernick (1985), Electrophoresis 6 (1985) 103-112.
32. S.C. Carpentier, E. Witters, K. Laukens, P. Deckers, R. Swennen, B. Panis, Proteomics 5 (2005) 2497-2507.
33. I. Miller, J. Crawford, E. Gianazza E, Proteomics 6 (2006) 5385–5408.
34. T. Rabilloud, Silver staining of 2D electrophoresis gels. Methods in Molecular Biology, Clifton, N.J., 2012, pp. 61-73.
35. A.P. Diz, M. Truebano, D.O.F. Skibinski DOF, Electrophoresis 30 (2009) 2967-2975.
36. B.M. Bolstad, R.A. Irizarry, M. Astrand, T.P. Speed, Bioin-

formatics 19 (2003) 185-193.
37. R Development Core Team, R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, 2008. URL http://www.R-project.org.
38. J.C. Pinheiro, D.M. Bates, eds, Mixed-effects models in S and S-PLUS. Springer, New York, 2000
39. Y. Benjamini, Y. Hochberg, J R Stat Soc Ser B 57 (1995) 289 –300.
40. R.A. Fisher, Statistical Methods for Research Workers. Oliverb and Boyd, Edinburgh, 1932.
41. M.A.S. Thorne, M.R. Worland, R. Feret, R.J. Deery, K.S. Lilley, M.S. Clark, Insect Mol Biol 20 (2011) 303–310.
42. S.F. Altschul, W. Gish, W. Miller, E.W. Myers, D. Lipman, J Mol Biol 215 (1990) 403-10.
43. H. Akaike, IEEE Transactions on Automatic Control 19 (1974) 716-723.
44. A.P. Diz, A. Carvajal-Rodriguez, D.O.F. Skibinski, Mol Cell Proteomics 10 (2011) M110.004374.
45. C.G. Knight, N. Zitzmann, S. Prabhakar, R. Antrobus, R. Dwek, H. Hebestreit, P.B. Rainey, Nat Genet 38 (2006) 1015 -1022.
46. J.L. Lopez, E. Mosquera, J. Fuentes, A. Marina, J. Vazquez, G. Alvarez, Mar Ecol Prog Ser 224 (2001) 149-156.
47. M.B. Yaffe, G.W. Farr, D. Miklos, A.L. Horwich, M.L. Sternlicht, H. Sternlicht, Nature 358 (1992) 245-248.
48. A. Forer, J Cell Sci 121 (2008) 7-9.
49. P. Liang, T.H. MacRae, J Cell Sci 110 (1997) 1431-1440.
50. O. Llorca, E. McCormack, G. Hynes, J. Grantham, J. Cordell, J. Carrascosa, K. Willison, J. Fernandez, J. Valpuesta, Nature 402 (1999) 693-696.
51. J. Martín-Benito, J. Boskovic, P. Gómez-Puertas, J.L. Carrascosa, C.T. Simons, S.A. Lewis, F. Bartolini, N.J. Cowan, J.M. Valpuesta JM, EMBO 21 (2002) 6377-6386.
52. M.K. Desalvo, C.R. Voolstra, S. Sunagawa S, J.A. Schwarz, J.H. Stillman, M.A. Coffroth, A.M. Szmant, M. Medina, Mol Ecol 17 (2008) 3952-3971.

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

ORIGINAL ARTICLE | DOI: 10.5584/jiomics.v3i1.131

Enzymatic protein digests do not assist in *E. coli* discrimination at the strain level using mass spectrometry

Ricardo J. Carreira[a], J. D. Nunes-Miranda[b,c], Alexandre Gonçalves[b,c,d,e], Gilberto Igrejas[b,c], Patrícia Poeta[d,e], Silvana Gómez-Meire[f], Miguel Reboiro-Jato[f], Florentino Fdez-Riverola[f], Carlos Lodeiro[g], José-Luis Capelo-Martínez[g]

[a]Biomolecular Mass Spectrometry Unit, Department of Parasitology, Leiden University Medical Center, Leiden, the Netherlands; [b]Department of Genetics and Biotechnology, University of Trás-os-Montes and Alto Douro, Vila Real, Portugal; [c]Institute for Biotechnology and Bioengineering, Centre of Genomics and Biotechnology, University of Trás-os-Montes and Alto Douro, Vila Real, Portugal; [d]Veterinary Science Department, University of Trás-os-Montes and Alto Douro, Vila Real, Portugal; [e]Center of Studies of Animal and Veterinary Sciences, Vila Real, Portugal; [f]SING Group, Informatics Department, Higher Technical School of Computer Engineering, University of Vigo, Ourense, Spain; [g]REQUIMTE, Departamento de Química, Faculdade de Ciências e Tecnologia, FCT, Universidade Nova de Lisboa, 2829-516 Caparica, Portugal.

Received: 13 February 2013 **Accepted:** 31 March 2013 **Available Online**: 03 April 2013

ABSTRACT

Different procedures for matrix assisted laser desorption ionization time of flight mass spectrometry-based *E. coli* classification at the strain level using the enzymatic digestion of proteins from the cell lysate have been studied. The effects of ultrasonic energy as well as the effects of protein reduction and protein alkylation in the sample treatment and in the subsequent classification were assessed. The final optimal method for classification was then compared with an intact cell-based approach in a different set of samples. Our results show that *E. coli* classification at the strain level is possible as 12 different strains were correctly classified using intact cell analysis. Overall, the confidence level in classification was higher when the analysis was performed with the intact cell approach.

Keywords: Bacterial; *E. Coli*; Fingerprinting; MALDI; Classification.

1. Introduction

Bacterial classification, referred as the assignment of an isolate to a species formally described [1], has been generally accomplished based on conventional techniques that rely on biochemical reactions and phenotypic characteristics, such as Gram staining and colony morphology [2]. The complete identification with these methods is usually time-consuming and, in some cases, correct identification is not possible [3]. Thereby, molecular biology methods, mostly based on nucleic acid tests such as polymerase chain reaction (PCR), have been used as complementary or alternative approaches to overcome these disadvantages due to their rapidity and sensitivity [4]. However, these techniques require a high level of technical expertise and remain expensive [5].

In recent years, besides its wide field of application in conventional proteomics, matrix-assisted laser desorption ionization time-of-flight mass spectrometry (MALDI-TOF MS) has been increasingly applied with success for microbial identification, including bacteria, fungi and viruses [6-12]. A unique mass spectral fingerprint of the microorganism is acquired and used for identification at the species and strain-level. The rapidity, accuracy and cost-effectiveness make it a promising tool to replace current methods in clinical microbial diagnostics within the next few years [1, 13].

Nevertheless, the increment in published work led to an increase in the number of sample processing methods that differ in particular steps [14]. Nowadays, among all treatment procedures published regarding MALDI-TOF MS microbial profiling, two main approaches can be discerned. The direct analysis of intact cells, without any disruption process, has been referred as an easy way to achieve microbial identification [15]. On the other hand, cell disruption and protein extraction analysis or whole cell lysate analysis represent an

***Corresponding author: Jose Luis Capelo Martínez,** REQUIMTE, Departamento de Química, Faculdade de Ciencias e Tecnologia, FCT, Universidade Nova de Lisboa, 2829-516 Caparica, Portugal; Phone number: +351 919 404 933; E-mail address: jlcm@fct.unl.pt

alternative method for microorganism identification [14]. Whereas few biomarkers are required for identification at the species-level, much more are needed for strain differentiation [16]. Protein fraction analysis enhances the discriminatory power of the technique by increasing the number of reproducible peaks detected that are useful in intra-species classification. For instance, Bizzini et al. improved the identification yield of strains of different species of bacteria and yeasts, from 70,3% to 93,2% after a protein extraction step [17].

In this work, the usefulness of digesting the whole bacterial lysate as well as digesting the whole bacterial protein content for bacterial discrimination at the strain level through MALDI-TOF MS has been assessed. The effectiveness of ultrasonic assisted enzymatic digestion was compared with the conventional overnight digestion, in order to reduce the identification time to a few minutes. Finally, the optimized method was applied to a distinct set of 12 *E. coli* isolates in order to compare it with an intact cell procedure for classification.

2. Material and methods

2.1. *Apparatus*

A Branson Sonifier® SLPe 150W equipped with a 3 mm microtip from Branson Ultrasonics (Danbury, USA) and a sonicator model SONOPULS HD 2200 with a Cup Horn BB6 accessory, from Bandelin (Berlin, Germany), were used for cell lysis and for accelerating protein reduction, alkylation and digestion steps. A mini-centrifuge CM-50 from ELMI (Riga, Latvia) and a refrigerated centrifuge MPW-65R from MPW Med. Instruments (Warsaw, Poland) were used throughout the experiments when necessary. Deionized ultrapure water was obtained from a Simplicity™ system from Millipore (Milan, Italy). Mass spectrometry analyses were performed with a MALDI-TOF/TOF UltraFlex model from Bruker Daltonics (Bremen, Germany).

2.2. *Bacterial samples*

Escherichia coli T1 Mach (Invitrogen) (A), NovaBlue (Merck-Millipore) (B), DH5α (C) and BL21 (D) were used for method development. For method validation, a set of 12 previously characterized *E. coli* isolates recovered from wild animals was used (E-P): GV5A (E) [18], L196 (F) [19], WA57 (G) [20], L65 (H) [19], W4B[a] (I) [20], W151A (J) [20], W184 (K) [20], L63[†] (L) [19], L18 (M) [21], L40 (N) [21], L115 (O) [21] and L185 (P) [22].

2.3. *Standards and reagents*

Luria-Broth (Sigma-Aldrich, Steinheim, Germany) and Brain Heart Infusion (Oxoid) liquid mediums were used for *E. coli* growth. Phosphate saline buffer (PBS, pH = 7.4), sodium deoxycholate (NaDOC), trichloroacetic acid (TCA), trifluoroacetic acid (TFA), acetonitrile, DL-dithiothreitol (DTT), iodoacetamide (IAA) and trypsin from porcine (proteomics grade) were also from Sigma-Aldrich. Ammonium bicarbonate (AmBic), formic acid, and α-cyano-4-hydroxycinnamic acid (α-CHCA) were purchased from Fluka (Buchs, Switzerland). Peptide and protein standards for mass calibration were purchased from Bruker Daltonics (Germany).

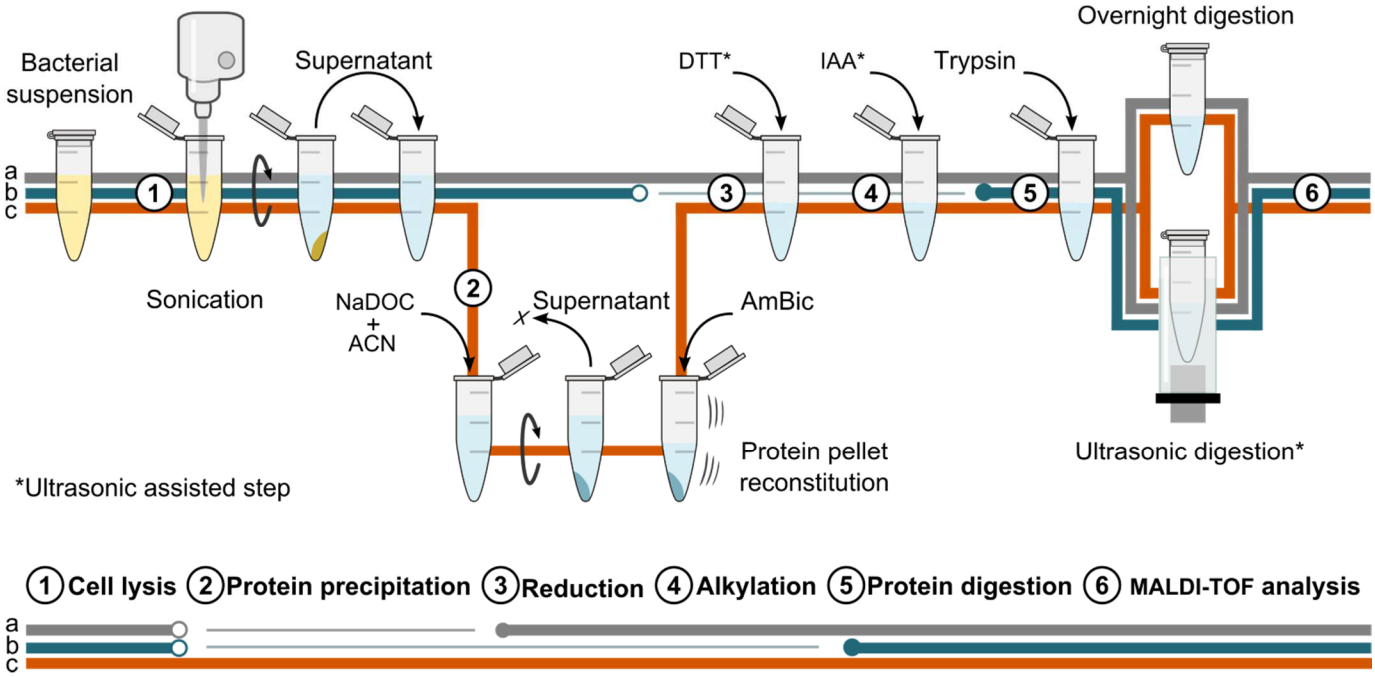

Figure 1. Overview of the different sample treatments applied in protein digestion. **(a)** Direct digestion of proteins from supernatant with the reduction and alkylation steps. **(b)** Direct digestion of proteins from supernatant without the reduction and alkylation steps. (c) Protein digestion after protein precipitation, resuspension and reduction and alkylation steps.

2.4. *Bacterial growth and protein extraction*

E. coli strains A, B, C and D were cultured in Luria-Broth (LB) medium and incubated overnight at 37 ºC, 190 rpm in an orbital shaker. After incubation, cells were centrifuged at 8000 rpm for 5 min, decanted and washed twice with PBS-buffer. Cell disruption was performed with an ultrasonic probe (3 mm probe tip diameter; 35 % ultrasonic amplitude) during 6 minutes with intercalated cycles of ultrasonication (10 s) and rest periods without sonication (10 s). This procedure was performed in an ice-bath to prevent sample overheating. Cell debris was removed by centrifugation at 13000 rpm at 4 ºC for 10 min. The supernatant was collected to new vials and the pellet was discarded. Three procedures were then applied: direct enzymatic digestion of supernatant with (i) and without (ii) reduction and alkylation steps and enzymatic digestion of supernatant's protein content after precipitation (iii). The procedures are described in detail below and depicted in Fig.1

2.5. *Enzymatic digestion of supernatant*

Two aliquots of 50 μL of supernatant were made for each method tested. Direct digestion of proteins was performed with (Fig. 1a) and without (Fig. 1b) reduction and alkylation steps. The protein content of each aliquot was processed as follows. For protein reduction, 2 μL of DTT (10 mM final concentration in sample mixture) were added and the sample was ultrasonicated with the sonoreactor, SR, SONOPULS (50 % amplitude) for 2 min. Then, 2 μL of IAA (50 mM final concentration in sample mixture) were added and protein alkylation was enhanced with the SR (50 % amplitude) for 2 min. Finally, 44 μL of trypsin (0,01 μg/μL final concentration in sample mixture) were added to the sample and protein digestion was performed during 5 min with the SR (50 % amplitude), or during overnight (ca. 12 h) at 37ºC. Protein digestion was stopped with the addition of 2 μL of formic acid (1 % v/v final concentration in sample mixture).

2.6. *Enzymatic digestion of supernatant's protein content*

Two aliquots of 500 μL of supernatant were taken for each method tested and proteins were precipitated. For protein precipitation samples were made up to 0.02 % with NaDOC (2% m/v). Samples were incubated during 30 min. Then, TCA was added until solution was made up to 10 % v/v in this reagent, and the samples were incubated at 4 ºC for 1 h. After centrifugation at 13000 rpm, 4 ºC, during 15 min, the supernatant was discarded and the pellet was washed with 400 μL of cold acetone (-20 ºC), incubated during 15 min at -20 ºC and centrifuged at 13000 rpm, 4 ºC, for 15 min. The supernatant was discarded and the pellet was air dried to remove acetone residues. Finally, 50 μL of AmBic 25 mM were added to dissolve the proteins from the pellet. Proteins were reduced, alkylated and digested as described above (Fig. 1c).

2.7. *Sample treatment for whole bacterial analysis*

Twelve isolates of *E. coli* were used. The isolates were taken from wild animals and they were collected during a biological survey done within a program of the UTAD University. Bacteria were cultured in BHI (brain heart infusion) medium. After incubation, cells were centrifuged at 8000 rpm for 5 min, decanted and washed twice with MilliQ water. The resulting pellet was then resuspended in 250 μL of 0.1 % TFA and vortexed for 60 s. Three biological replicates for each isolate were done (Fig. SM-1).

2.8. *MALDI-TOF-MS analysis*

2.8.1. *Peptide digests*

The peptide digests were mixed (1:1) with α-CHCA matrix at a concentration of 10 mg/mL in acetonitrile/water (50 % v/v) and TFA (0.1%). Then, each sample was applied onto the MALDI sample plate (Bruker Daltonics parts no. 209519) in five different spots (1 μL per spot), air dried and analyzed without any further purification steps, in a Ultraflex II MALDI-TOF/TOF instrument from Bruker Daltonics equipped with a 200 Hz Smartbeam laser system. Data was acquired using the FlexControl software (version 3.3.108.0). Measurements were performed in the reflectron positive-ion mode. A total of 1500 spectra were acquired at each spot position with a laser frequency of 100 Hz. Close external calibration was performed with the $[M + H]^+$ monoisotopic peaks of bradykinin 1-7 (757.3992), angiotensin II (1046.5418), angiotensin I (1296.6848), substance P (1347.7345), bombesin (1619.8223), renin substrate (1758.9326), ACTH clip 1-17 (2093.0862), ACTH 18-39 (2465.1983), and somatostatin 28 (3147.4710). The peak lists were generated from the mass spectra using the peak detection algorithm SNAP (sophisticated numerical annotation procedure).

2.8.2. *Intact bacterial analysis*

A volume of 10 μL of the bacterial suspension was mixed with 10 μL of a sinapinic acid matrix solution (20 mg/mL in 50 % acetonitrile/ 2 % TFA. Then, 1 μL of each sample was applied in quintuplicate onto a MALDI target plate (MTP 384 target plate ground steel T F). The same Ultraflex II MALDI-TOF/TOF instrument described above was used. Measurements were performed in the linear positive-ion mode from 2000 to 20000 m/z. Close external calibration was performed with the protein calibration standards I. Peak lists were generated with the FlexAnalysis software (version 3.3) using the peak detection algorithm Centroid.

3. Results and discussion

Currently, the identification of bacteria by MALDI-TOF-MS relies mostly on the analysis of intact bacteria due to the

simplicity of the procedures and the rapidity of the analysis. Yet, these procedures may not be suitable for the analysis of pathogenic bacteria due to the risk of contamination and exposure of the operator to the infectious agent. Alternatives relying on the analysis of inactivated bacterial lysates and enzymatic digests of proteins from the bacterial lysate have been recently developed [14]. New approaches to such alternatives are presented below.

3.1. *The effect of ultrasonic energy in the digestion of whole bacterial lysates.*

Ultrasound has been reported as a useful and reliable tool to accelerate not only protein digestion with trypsin, but also the entire protocol for protein identification by MS including protein reduction with DTT and alkylation with IAA [23 -26]. The mechanism responsible for the enhancement of protein enzymatic digestion with ultrasound is not fully understood yet, but it is thought to rely on cavitation [27]. The cavitation phenomena occurs when a ultrasound wave travels through a liquid media generating micro bubbles of gas, known as cavitation bubbles, which grow during successive rarefaction and compression cycles. Eventually, these bubbles reach an unstable size and undergo violent collapse generating localized spots of high pressure and high temperature, which can be used as micro reactors to enhance chemical and physical reactions. In this work, ultrasonic energy from a sonoreactor apparatus, SR, was used to accelerate the digestion of proteins from crude bacterial lysates and proteins obtained after precipitation of the bacterial lysate, as well as to enhance protein reduction and alkylation steps. The SR works as a small ultrasonic bath. However, the ultrasound intensity delivered is much higher than the one obtained from a normal ultrasonic bath. This it allows to speed up protein digestion with great efficiency [28]. Furthermore, ultrasonication with the SR is performed in closed vials. This prevents sample contamination and allows hazardous samples to be safely handled.

In the first part of this work, the ultrasonic (5 min) and overnight (12 h) enzymatic digestion of crude bacterial lysates were compared. As a general role, the mass spectra obtained after ultrasonication had lower intensity than the spectra obtained after overnight digestion (Fig. SM-2). Table SM-1 presents the number of mass peaks with relative intensity higher than 5 % or 10 % obtained after protein enzymatic digestion with the overnight (12 h) procedure or with the ultrasonic (5 min) procedure. As this table shows, the number of peaks with a relative intensity higher than 5 % obtained with the overnight procedure varied from 128 to 564, which represents a difference of more than 400 peaks between spectra of D and B strains, respectively. Regarding the ultrasound-assisted digestion, the number of mass peaks with a relative intensity higher than 5 % varied from 279 to 478 peaks. The same pattern was observed when only the mass peaks with a relative intensity higher than 10 % were considered. As a trend, the number of significant mass peaks

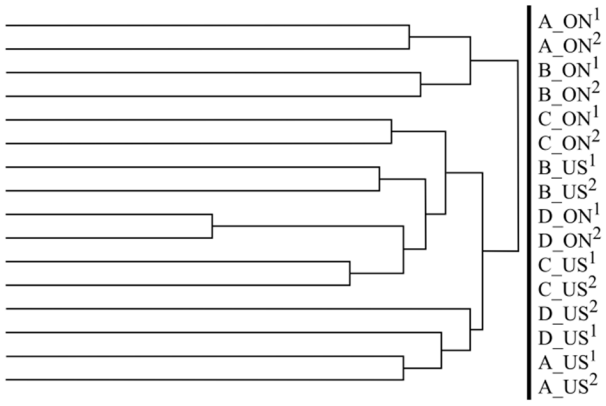

Figure 2. Hierarchical clustering of peak lists obtained by MALDI-TOF MS from overnight (ON) and ultrasound (US) assisted enzymatic digestion of crude bacterial lysates of E. coli strains A, B, C and D. Two replicates were performed in each procedure. Cluster analysis was accomplished with BACTERIAL IDENTIFICATION software (Euclidean distance, minimum variance 0.1%).

is higher in spectra from overnight digested samples than in spectra from samples digested with ultrasound., with the exception of strain D. In this case, one of the main reasons for the discrepancy in the size of the list of peptides can be related with ion suppression effect, as can be seen by the mass spectrum obtained after the overnight digestion of proteins from strain D_ON (Fig. SM-2). This spectrum clearly shows a high intense mass peak at 1480 m/z. The major effect is the decrease of the relative intensities of the remaining mass peaks, which in turn are not selected for analysis, since they do not meet the relative intensity threshold. This is very important because the higher the number of peaks selected for analysis the higher the possibility to find differences between strains. The mass spectra obtained were analysed by hierarchical clustering to study the classification achieved by each sample treatment. As shown in Fig. 2, most replicates of the different samples are grouped in pairs, according to the strain and to the type of sample treatment used. This shows that both, the overnight and ultrasonic sample treatments; render reproducible mass spectra, which can be used to distinguish bacterial samples at the strain level. Furthermore, it is clear from this result that the sample treatment is the variable that most influences the classification of the strains, as they are grouped as a function of the sample treatment used. This result is of great importance because it reveals that sets of data obtained from two different sample treatments should not be used to classify bacteria.

3.2. *Influence of protein reduction and alkylation in the sample treatment for bacterial classification.*

Protein reduction with DTT followed by alkylation with IAA is used in proteomics to break disulphide bonds between cysteine residues and to prevent their formation, respectively. These reactions introduce a chemical modification which destroys the native structure of the protein and

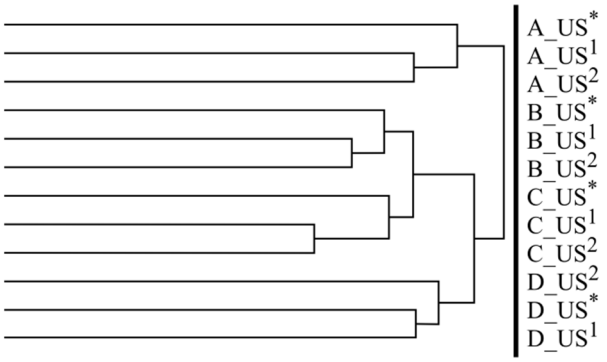

Figure 3. Hierarchical clustering of peak lists obtained by MALDI-TOF MS from ultrasound assisted enzymatic digestion of crude bacterial lysates of E. coli strains A, B, C and D, with (US) and without (US*) previous reduction and alkylation treatments. Two replicates were performed in the US procedure. Cluster analysis was accomplished with BACTERIAL IDENTIFICATION software (Euclidean distance, minimum variance 0.1%).

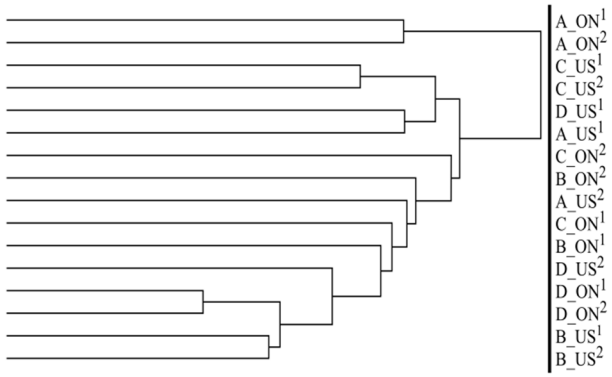

Figure 4. Hierarchical clustering of peak lists obtained by MALDI-TOF MS from overnight (ON) and ultrasound (US) assisted enzymatic digestion of precipitated proteins of E. coli strains A, B, C and D. Two replicates were performed in each procedure. Cluster analysis was accomplished with BACTERIAL IDENTIFICATION software (Euclidean distance, minimum variance 0.1 %).

helps enzymatic digestion by exposing cleavage sites that otherwise would remain hindered by the native structure. In the previous section, the analysis was done after protein reduction and alkylation. To assess the influence of protein reduction and alkylation on bacteria classification, direct enzymatic digestion assisted with ultrasound energy was performed on a sample set without reduction and alkylation steps. The number of peptides obtained with a relative intensity higher than 5% in the non-reduced and non-alkylated samples was 536, 403, 333 and 481 for strains A, B, C and D, respectively (Table SM-2). These values are of the same magnitude than the ones obtained in samples where protein reduction and alkylation was performed (Table SM-1). Fig. 3 presents the hierarchical clustering analysis of the spectra obtained for ultrasonic treated samples, both non-reduced and non-alkylated together with reduced and alkylated. As may be seen, 4 main groups were formed, each one corresponding to a different strain. These results suggest that performing protein reduction and alkylation on proteins from crude bacterial lysates before the enzymatic digestion has a minor effect in the classification obtained by MALDI-TOF-MS analysis.

3.3. *Evaluation of the discriminatory power of the protein content in the bacterial classification.*

When working with bacterial lysates, protein precipitation is routinely used to concentrate and separate the protein content from other components present in the lysate. Yet, one of the main disadvantages of this approach is the incomplete protein recovery from the precipitate [29]. In this work, proteins were precipitated from bacterial lysates with TCA/NaDOC. After precipitation, the pellet was resuspended in ammonium bicarbonate buffer (25 mM), then the proteins were reduced and alkylated, and finally digested with trypsin by the overnight (12 h) or the ultrasonic assisted (5 min) procedures. Results are shown in Fig. SM-3. As a general rule, the intensity of the spectra was lower than for the analysis of crude bacterial lysates. Even though the number of mass peaks obtained per spectra was similar to the ones obtained for the crude bacterial lysates, a higher variance between replicates was observed (Table SM-3). Fig. 4 presents the hierarchical clustering analysis of the spectra obtained after overnight and ultrasound assisted enzymatic digestion. Once again, the same strains are grouped in pairs according to the procedure used for protein digestion. For instance, replicates of the ultrasound assisted digestion of proteins from strain D were found closely related, but distant from spectra of the same strain obtained after overnight enzymatic digestion. This confirms that bacterial identification by MS seems highly dependent on the sample treatment used.

3.4. *Evaluation of the discriminatory power of the digested protein content analysis in a set of E.* coli *isolates*

According to the results presented in the sections above, the protein digestion of the crude bacterial lysate assisted with ultrasonic energy appears to be the most favourable method for strain discrimination. In order to evaluate the sensitivity and robustness of the method, 12 *E. coli* isolates recovered from wild animals previously classified according to the antimicrobial resistance gene profile were tested. Results presented in Fig. 5, show that nearly half of the replicated isolates were clustered grouped together after ultrasonic assisted enzymatic digestion. However, others were clustered distantlydistantly related, as shown for example in the replicates of isolate K. Therefore, it is possible to verify that this approach may have limitations. On the other hand, as may be observed in Fig. 6, the direct analysis of whole bacterial cells allows the classification of almost all strains. This suggest that this method is highly reproducible and also that the MALDI-TOF-based mass spectrometry can be used to

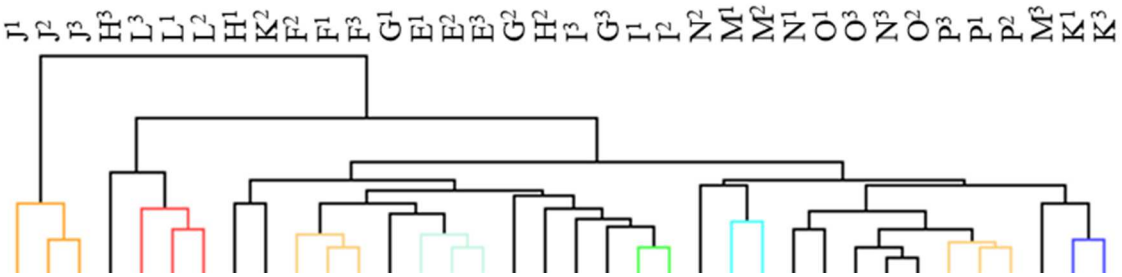

Figure 5. Hierarchical clustering of peak lists obtained by MALDI-TOF MS from ultrasound assisted enzymatic digestion of crude bacterial lysates of twelve (E-P) E. coli isolates characterized by the antimicrobial resistance gene profile. Three biological replicates were performed for each isolate.

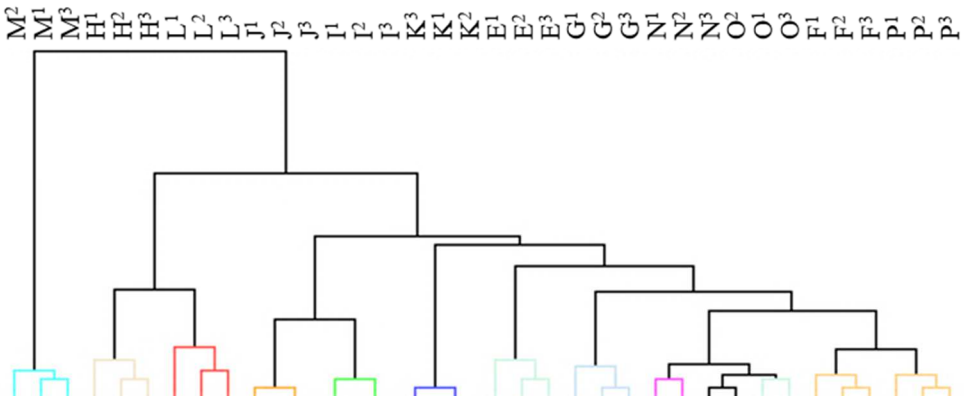

Figure 6. Hierarchical clustering of peak lists obtained by MALDI-TOF MS from whole cell analysis of twelve (E-P) E. coli isolates characterized by the antimicrobial resistance gene profile. Three biological replicates were performed for each isolate.

differentiate bacteria of the same species at the strain level. This approach, whole cell mass spectrometry (WC-MS), has been a widely adopted routine procedure for bacterial differentiation.

4. Conclusions

Different sample treatments for bacterial identification relying on the analysis of digested bacterial lysates have been compared. Crude bacterial lysates or proteins obtained after precipitation were enzymatically digested during overnight or with the aid of ultrasonic energy. In addition, the effect of protein reduction and alkylation on the classification was also studied. The results showed that reduction and alkylation have little effect in the classification of bacteria when ultrasonic energy was used to enhance enzymatic digestion of crude bacterial extracts. It was also verified that samples were grouped according to the sample treatment used rather than the corresponding bacterial strain. Even though, samples corresponding to the same strain and with the same sample treatment were grouped together in the majority of the studied cases. Nevertheless, the use of enzymatic digests failed in classifying 12 wild *E. coli* strains whilst such classification was achieved with the analysis of the intact bacteria. Therefore the use of digested samples is not recommended for classification of *E. coli* strains. Also bacterial classification through MALDI-based mass spectrometry is highly dependent on the sample treatment used and therefore classification studies must be always done using the same sample treatment. In fact this study seems to suggest that there is an urgent need to method standardization for MALDI-TOF sample preparation for microbial identification. It must be highlighted that this study was performed with gram negative species. Several known difficulties regarding mass spectrometry are inherent to the peptidoglycan layer of the gram positive bacterial cell wall, wherein a previous disruption step may be required to improve signal record. Therefore, extrapolation of these outcomes to other cases should be done carefully.

5. Supplementary material

Supplementary data and information is available at: http://www.jiomics.com/index.php/jio/rt/suppFiles/131/0. Supplementary Material includes Figures SM_1, SM_2 and SM_3.

Acknowledgments

The authors thank Dr. Pablo Gonzalez, Biochemistry Laboratory, FCT-UNL at Portugal, for gently providing the *E. coli* samples, DH5α, T1 Mach, NovaBlue and BL21, to this study. European Interreg Project POCTEC 47 Tubercontrol_1_E is also acknowledge for financial support. Alexan-

dre Gonçalves has a PhD fellowship granted by FCT-Fundação para a Ciência e a Tecnologia and ESF-European Social Fund (SFRH/BD/47833/2008).

References

1. Welker, M., Proteomics for routine identification of microorganisms. *Proteomics* 2011, *11*, 3143-3153.
2. Bizzini, A., Greub, G., Matrix-assisted laser desorption ionization time-of-flight mass spectrometry, a revolution in clinical microbial identification. *Clin Microbiol Infect* 2010, *16*, 1614-1619.
3. Lotz, A., Ferroni, A., Beretti, J. L., Dauphin, B., *et al.*, Rapid identification of mycobacterial whole cells in solid and liquid culture media by matrix-assisted laser desorption ionization-time of flight mass spectrometry. *J Clin Microbiol* 2010, *48*, 4481-4486.
4. Carbonnelle, E., Mesquita, C., Bille, E., Day, N., *et al.*, MALDI-TOF mass spectrometry tools for bacterial identification in clinical microbiology laboratory. *Clin Biochem* 2011, *44*, 104-109.
5. Stevenson, L. G., Drake, S. K., Shea, Y. R., Zelazny, A. M., Murray, P. R., Evaluation of matrix-assisted laser desorption ionization-time of flight mass spectrometry for identification of clinically important yeast species. *J Clin Microbiol* 2010, *48*, 3482-3486.
6. Angelakis, E., Million, M., Henry, M., Raoult, D., Rapid and Accurate Bacterial Identification in Probiotics and Yoghurts by MALDI-TOF Mass Spectrometry. *Journal of Food Science* 2011, *76*, M568-M572.
7. El Khechine, A., Couderc, C., Flaudrops, C., Raoult, D., Drancourt, M., Matrix-assisted laser desorption/ionization time-of-flight mass spectrometry identification of mycobacteria in routine clinical practice. *PloS one* 2011, *6*, e24720.
8. Hettick, J. M., Kashon, M. L., Simpson, J. P., Siegel, P. D., *et al.*, Proteomic Profiling of Intact Mycobacteria by Matrix-Assisted Laser Desorption / Ionization Time-of-Flight Mass Spectrometry culture are time-consuming , requiring as long as 12. *October* 2004, *193*, 5769-5776.
9. Lamy, B., Kodjo, A., Laurent, F., Identification of Aeromonas isolates by matrix-assisted laser desorption ionization time-of-flight mass spectrometry. *Diagn Microbiol Infect Dis* 2011, *71*, 1-5.
10. Pignone, M., Greth, K. M., Cooper, J., Emerson, D., Tang, J., Identification of mycobacteria by matrix-assisted laser desorption ionization-time-of-flight mass spectrometry. *J Clin Microbiol* 2006, *44*, 1963-1970.
11. Vargha, M., Takats, Z., Konopka, A., Nakatsu, C. H., Optimization of MALDI-TOF MS for strain level differentiation of Arthrobacter isolates. *J Microbiol Methods* 2006, *66*, 399-409.
12. Giebel, R., Worden, C., Rust, S. M., Kleinheinz, G. T., *et al.*, Microbial fingerprinting using matrix-assisted laser desorption ionization time-of-flight mass spectrometry (MALDI-TOF MS) applications and challenges. *Advances in applied microbiology* 2010, *71*, 149-184.
13. Croxatto, A., Prod'hom, G., Greub, G., Applications of MALDI-TOF mass spectrometry in clinical diagnostic microbiology. *FEMS Microbiol Rev* 2012, *36*, 380-407.
14. Šedo, O., Sedláček, I., Zdráhal, Z., Sample preparation methods for MALDI-MS profiling of bacteria. *Mass Spectrometry Reviews* 2011, *30*, 417-434.
15. van Veen, S. Q., Claas, E. C., Kuijper, E. J., High-throughput identification of bacteria and yeast by matrix-assisted laser desorption ionization-time of flight mass spectrometry in conventional medical microbiology laboratories. *J Clin Microbiol* 2010, *48*, 900-907.
16. Dieckmann, R., Helmuth, R., Erhard, M., Malorny, B., Rapid classification and identification of salmonellae at the species and subspecies levels by whole-cell matrix-assisted laser desorption ionization-time of flight mass spectrometry. *Appl Environ Microbiol* 2008, *74*, 7767-7778.
17. Bizzini, A., Durussel, C., Bille, J., Greub, G., Prod'hom, G., Performance of matrix-assisted laser desorption ionization-time of flight mass spectrometry for identification of bacterial strains routinely isolated in a clinical microbiology laboratory. *J Clin Microbiol* 2010, *48*, 1549-1554.
18. Poeta, P., Radhouani, H., Igrejas, G., Goncalves, A., *et al.*, Seagulls of the Berlengas natural reserve of Portugal as carriers of fecal Escherichia coli harboring CTX-M and TEM extended-spectrum beta-lactamases. *Appl Environ Microbiol* 2008, *74*, 7439-7441.
19. Goncalves, A., Igrejas, G., Radhouani, H., Estepa, V., *et al.*, Detection of extended-spectrum beta-lactamase-producing Escherichia coli isolates in faecal samples of Iberian lynx. *Letters in applied microbiology* 2012, *54*, 73-77.
20. Goncalves, A., Igrejas, G., Radhouani, H., Estepa, V., *et al.*, Iberian wolf as a reservoir of extended-spectrum beta-lactamase-producing Escherichia coli of the TEM, SHV, and CTX-M groups. *Microb Drug Resist* 2012, *18*, 215-219.
21. Goncalves, A., Igrejas, G., Radhouani, H., Santos, T., *et al.*, Iberian Lynx (Lynx pardinus) as reservoir of antimicrobial resistant enterococci and Escherichia coli isolates. *Submitted* 2012.
22. Goncalves, A., Igrejas, G., Radhouani, H., Santos, T., *et al.*, Detection of antibiotic resistant enterococci and Escherichia coli in free range Iberian Lynx (Lynx pardinus). *Science of the Total Environment* 2013, *In Press*.
23. Lopez-Ferrer, D., Capelo, J. L., Vazquez, J., Ultra fast trypsin digestion of proteins by high intensity focused ultrasound. *J Proteome Res* 2005, *4*, 1569-1574.
24. Capelo, J. L., Carreira, R., Diniz, M., Fernandes, L., *et al.*, Overview on modern approaches to speed up protein identification workflows relying on enzymatic cleavage and mass spectrometry-based techniques. *Analytica chimica acta* 2009, *650*, 151-159.
25. Santos, H. M., Carreira, R., Diniz, M. S., Rivas, M. G., *et al.*, Ultrasonic multiprobe as a new tool to overcome the bottleneck of throughput in workflows for protein identification relaying on ultrasonic energy. *Talanta* 2010, *81*, 55-62.
26. Santos, H. M., Mota, C., Lodeiro, C., Moura, I., *et al.*, An improved clean sonoreactor-based method for protein identification by mass spectrometry-based techniques. *Talanta* 2008, *77*, 870-875.
27. Santos, H. M., Capelo, J. L., Trends in ultrasonic-based equipment for analytical sample treatment. *Talanta* 2007, *73*, 795-802.
28. Vale, G., Santos, H. M., Carreira, R. J., Fonseca, L., *et al.*, An assessment of the ultrasonic probe-based enhancement of protein cleavage with immobilized trypsin. *Proteomics* 2011, *11*, 3866-3876.
29. Jiang, L., He, L., Fountoulakis, M., Comparison of protein precipitation methods for sample preparation prior to proteomic analysis. *Journal of chromatography. A* 2004, *1023*, 317-320.

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

ORIGINAL ARTICLE | DOI: 10.5584/jiomics.v3i1.126

Vascular Smooth Muscle Cells activation revealed by quantitative phosphoproteomics analysis

Claudia Boccardi[1,2,*,§], Vittoria Matafora[3,4,§], Giovanni Signore[2], Silvia Rocchiccioli[1], Maria Giovanna Trivella[1], Emanuele Alpi[3], Lorenzo Citti[1], Angela Bachi[3], Antonella Cecchettini[1,5]

[§]These authors contributed equally to this work; [1]Institute of Clinical Physiology, CNR, Pisa, Italy; [2]Center for Nanotechnology Innovation @NEST, Istituto Italiano di Tecnologia, Pisa, Italy; [3]San Raffaele Scientific Institute, Division of Genetics & Cell Biology, Milan, Italy; [4]Department of Internal Medicine, Second University of Naples, Naples, Italy; [5]Clinical and Experimental Medicine, University of Pisa, Pisa, Italy.

Received: 22 January 2013 **Accepted:** 14 April 2013 **Available Online**: 10 May 2013

Abstract

Vascular smooth-muscle cells (VSMCs) are the main components of the artery medial layer and if activated by growth factors as a consequence of vessel injuries, acquire the ability to proliferate and migrate contributing to the formation of neointima. In the early phase/events of VSMC stimulation a cascade of kinases and phosphatases initiates phospho-events that are decisive for VSMC activation. In this work, a SILAC approach in a multi-strategy combined method for phosphopeptides enrichment was used, in order to gain insights into the phosphomodulation of the proteome of VSMCs stimulated by platelet derived growth factor BB (PDGF-BB).. The quantitative SILAC phosphoproteome analysis allowed the identification of 1300 phosphopeptides among which 47 resulted novel phosphosites. Some important factors involved in cytoskeleton remodeling, focal adhesions, gap junction assembly and cell activation have been found differentially phosphorylated, highlighting their pivotal role in early VSMC reorganization and suggesting a novel starting point to assess the precise actors of VSMC activation and their roles.

Keywords: quantitative phosphoproteomics; SILAC; VSMC activation; PDGF-BB.

1. Introduction

Vascular smooth muscle cells (VSMCs), although highly specialized, are not terminally differentiated cells. They maintain a great plasticity and, when stimulated by external inputs, they are able to switch to an activated phenotype characterized by increased cell proliferation capability and triggering of cell proliferation and migration programs [5,6]. These properties are helpful for vascular healing following injury and damage, but can also be detrimental since they exert, a pivotal role in onset and progression of vascular proliferative disorders contributing to the thickening of the tunica intima, named neointima, and the beginning of the atherosclerotic plaque [7]. A wide range of elements contributes to VSMC activation and, among them, growth factors are the most critical ones. Platelet derived growth factor –BB (PDGF-BB), produced by activated platelets and macrophages, has so far been the only factor demonstrated to selectively and directly promote VSMC phenotype changes [5,8].

Following growth factor binding to its receptor, a cascade of kinases and phosphatases are activated and, within a few minutes, a series of events and specific responses are triggered while the phosphorylation state of the proteins changes dramatically [9]. These early events, even if phenotypically imperceptible, are decisive for the activation process in VSMCs [6, 10], and result with a complete reorganization of the cell structure.

Indeed, several stimuli regulate both VSMC differentiation and activation through phosphorylation [11]. Chen and colleagues in 2010 investigated the signal transduction pathways of aeat shock protein 27 (HSP27) phosphorylation in VSMCs after angiotensin II (ANG-II) and PDGF-BB stimu-

***Corresponding author:** Claudia Boccardi, Institute of Clinical Physiology, CNR, Pisa, Italy; Phone number: +39050509769; Fax number: +39050509417; E-mail address: claudia.boccardi@iit.it

lation [12]. Similarly, Liu et al observed the promotion of VSMC proliferation via the phosphoinositol-3-kinase/ Protein Kinase B (PI3K/Akt) pathway induced by apelin-13 [13]. Recently, in an *in vivo* rat carotid injury model, overexpression of Smad3 produced an increase in phosphorylated extracellular signal-regulated kinases (ERK) mitogen-activated protein kinases (MAPK) as well as increased VSMC proliferation [14].

Several MS-based quantification methods have been exploited for the study of the phosphoproteome [15, 16, 17], highlighting how accuracy and reliability of phosphorylation sites identification and quantitation is of key importance. Among others, stable-isotope labeling of amino acids in cell culture (SILAC) is a simple and straightforward approach [18], which allows the metabolic labeling in living cells and minimizes variations that could be introduced during sample preparation. However, some limitations still remain in a global phosphoproteomics approach, due to the low amount of phosphorylated proteins [19], and these are overcome by the use of phosphopeptide enrichment methods [20,21,22].

In this work, we used SILAC in a multi-strategy combined method for phosphopeptides enrichment in order to gain insights into the phospho-modulation of VSMC proteome stimulated by PDGF-BB. In particular, as the phosphorylation events are really fast and dynamic, we analyzed two different time points after stimulation: 10 minutes and 2 hours. 1300 phosphopeptides, clustering into 380 protein groups were identified. Among them, 21 phosphopeptides (corresponding to 16 proteins) resulted phospho-modulated at different times upon PDGF-BB stimulation. These proteins are involved in cytoskeleton remodeling, focal adhesions, gap junction assembly and cell activation indicating that the events related to PDGF-BB phosphorylation pathway are the triggering phenomena in VSMC activation.

2. Material and methods

2.1 *Chemicals, standards and consumables*

ACN, water, formic acid used for LC-MS analysis were from Fluka, Sigma Aldrich St. Louis, MO; sequencing grade trypsin was purchased from Promega (Madison, WI). SILACTM DMEM medium, fetal bovine serum (FBS), L-lysine, L-arginine, [13C6]-L-Lysine and [13C6]-L Arginine were purchased from Invitrogen (USA, CA). [U-13C6, U-15N2] Lysine and [U-13C6,U-15N4] Arginine were purchased at CIL (USA). Recombinant Human PDGF-BB was from R&D (USA). Coomassie Plus protein assay was from Pierce, Bio-Spin Column from Biorad and PHOS-Select Iron Affinity Gel (IMAC resin) was purchased from Sigma Aldrich. All other chemicals were purchased from Sigma Aldrich.

2.2 *Cell Culture and SILAC-Labeling*

SILAC™ D-MEM medium containing 2 mM L-Glutamine, 10% dialyzed FBS, and 100 U/mL penicillin and streptomycin was supplemented with 100 mg/L L-lysine and 100 mg/L L-arginine; 100 mg/L [^{13}C6]-L-Lysine and 100 mg/L [^{13}C6]-L Arginine or 100 mg/L [U-^{13}C6, U-^{15}N2] Lysine and 100 mg/L [U-^{13}C6,U-^{15}N4] Arginine to prepare the "Light" (L), "Medium" (M)" or "Heavy" (H) SILAC culture media, respectively. VSMCs were obtained from coronaries of 8-month-old domestic crossbred pigs (*Sus scrofa domestica*) and the investigation was conformed to Ministerial Decree to authorize testing n.06/2009-b. Cells isolation was performed according to the method described by Christen et al [23] and VSMCs were expanded in SILAC medium for 6 doublings to ensure nearly 100% incorporation of labeled amino acids. For the experiments three biological replicates (cells isolated from at least three different pieces of coronary coming from three different animals) and then pooled before analysis. Prior to cell treatment, all "heavy", "medium" and "light" cells were cultured overnight in their corresponding SILAC media without FBS. For PDGF-treatment, medium labeled cells (M) were treated for 10 minutes (t10min) with 100 ng/mL of PDGF-BB; heavy labeled cells (H) were stimulated for 2 hours (t2h) with PDGF-BB while Light labeled cells (L) were not treated with the growth factor (t0).

2.3 *Protein lysate preparation and in-solution digestion*

VSMCs were harvested with trypsin and lysed in 50mM Tris-HCl pH 8, 0.25% sodium deoxycholate, 0.5% Triton X100, 10mM sodium orthovanadate, 5mM sodium fluoride. For each of three specimens (L, M and H cells), three million cells were re-suspended in 400 µl of lysis buffer. After 30 min in ice, samples were sonicated and then centrifuged at 10000xg for 30 min at 4 °C to produce a clarified lysate. L, M and H VSMC lysates were mixed at 1:1:1 [w/w] ratio, the protein amount was determined by Coomassie Plus protein assay. In total, 9 mg of the 1:1:1 mixed protein lysate was precipitated with a standard chloroform/methanol precipitation protocol. Precipitated proteins were re-suspended in 450 µl of 100 mM Tris-HCl pH 8.

Disulfide bonds were reduced with 5 mM DTT for 60 min at room temperature, and then the free sulfhydryl groups were alkylated with 10 mM iodoacetamide at room temperature in the dark for 30 min. Protein digestion was performed using a total of 18 µg of sequencing grade trypsin with 15 h incubation at 37°C.

2.4 *Strong Cation Exchange (SCX) Chromatography*

Preparative separations were carried out with SCX resin, using Bio-Spin Microcolumns. The first equilibration resin step was carried out using 4.5 mL of solution A (5 mM KH_2PO_4, pH 2.65; 30% ACN; 0,1% TFA) for 5 min. Then 4.5 mL of regeneration buffer were used (5mM KH_2PO_4, 30% ACN, 0.1% TFA, 1M KCl, pH 2.7). The first step was repeated. The peptide mixtures were diluted with 4.5 mL of SCX

solution A and placed on the resin at room temperature for 40 minutes. Increasing concentrations of solution B (5mM KH_2PO_4, 30% ACN, 0,1% TFA, 350mM KCl, pH 2.7) from 5% to 100% in 7 steps (5, 10, 15, 20 ,25, 50, 100) were used to elute peptides.

2.5 *Immobilized Metal-Affinity Chromatography (IMAC)*

PHOS-Select Iron Affinity Gel was used for IMAC fractionation using the same microcolumn. As first step, resin was washed three times with 200 µL of binding solution (0.1% TFA, 30% ACN). 50µL of PHOS-Select slurry were added to each SCX fraction. After 60-min incubation at room temperature with vigorous shaking, the supernatant was removed (IMAC flow-through). The resin was washed with 200 µL of binding solution and added to IMAC flow-through. The mono-phosphorylated peptides and non-phosphorylated peptides were eluted from IMAC resin using acidic conditions (200 µL of 1% TFA, 20% ACN), while the poly-phosphorylated peptides were subsequently eluted from the same IMAC resin using ammonium hydroxide (200 µL of NH_4OH pH 11.2). After concentration (from 1mL to 0.1mL) in dry vacuum the pooled flow-through was further analyzed by TiO_2 chromatography. The mono-phosphorylated peptide fraction was also subjected to TiO_2 chromatography as described below. Basic elution samples were submitted to LC-MS/MS analysis directly.

2.6 *Titanium Dioxide (TiO2) Chromatography*

Samples were diluted nine times with TiO_2 binding solution (5mg/mL DHB, 80% ACN and 3% TFA). 160 mg of TiO_2 resin were re-suspended in 1.4 mL of TiO_2 binding solution. 200µl of TiO_2 resin slurry were added to each sample for 30 min at room temperature. The supernatant was removed by centrifugation in a microcolumn. Non- specific binders were removed by two washes with 500 µl of 50% ACN, 0.2% TFA solution. Phosphopeptides bound were eluted from TiO_2 micro-columns using 200 µL of ammonium hydroxide, 40% ACN pH 12.5 followed by 100 µL of 100% ACN.

2.7 *Mass Spectrometry*

Mass spectrometry analysis was performed on a LTQ-Orbitrap mass spectrometer (Thermo Scientific, Bremen, Germany). Tryptic digests for each sample were cleaned using Stage Tips as described previously [24] and then injected in a capillary chromatographic system (EasyLC, Proxeon Biosystems, Denmark); Peptide separation occurred on a homemade column obtained with a 10-cm fused silica capillary (75 µm inner diameter and 360 µm outer diameter; Proxeon Biosystems) filled with 3 µm resin Reprosil-Pur C18 (Dr. Maisch GmbH, Ammerbuch-Entringen, Germany) using a pressurized "packing bomb". A gradient of eluents A (distilled water with 2% (v/v) ACN, 0.1% (v/v) formic acid) and B (2% (v/v) ACN, distilled water with 0.1% (v/v) formic acid) was used to achieve separation from 8% B (at 0 min, 0.2 µL-min flow rate) to 50% B (at 80 min, 0.2 µL-/min flow rate). The LC system was connected to the Orbitrap equipped with a nanoelectrospray ion source (Proxeon Biosystems). Full-scan mass spectra were acquired in the mass range m/z 350 to 1500 Th and with the multistage activation enabled for neutral loss of phosphoric acid [32.66, 48.99, and 97.97 atomic mass units (amu)][25]. All full-scan spectra were recalibrated in real time with the lock-mass option [26]. The 5 most intense doubly and triply charged ions were automatically selected and fragmented in the ion trap.

2.8 *Data analysis*

The raw MS data were processed and analyzed using MaxQuant software version 1.0.13.13. The required FDR was set to 0.01 at both the peptide and protein level and the minimum required peptide length was set six amino acids. The MS/MS spectra were searched against Uniprot mammalia 2010_07 database supplemented by frequently observed contaminants, concatenated with reversed versions of all sequences. Mascot (version 2.2.07) was used for the database search. Enzyme specificity was set to trypsin, additionally allowing up to two missed cleavages. The search included cysteine carbamidomethylation as fixed modification and N-acetylation of protein, oxidation of methionine, and phosphorylation of serine, threonine and tyrosine as variable modifications. Protein SILAC ratios are reported as the median of the ratios derived from SILAC triplets assigned to the protein. For phosphopeptides, the phosphorylation site(s) were assigned with a modified version of the PTM score [27] given in MaxQuant. MaxQuant automatically quantified SILAC peptides and proteins. SILAC protein ratios were calculated as the median of all peptide ratios assigned to the protein. Ratio Variability was less than 30%. Systematic deviations, such as mixing errors, were corrected by the quantitation algorithm in the MaxQuant software by normalizing all peptide ratios such that the mean of all log-transformed ratios were zero. A posterior error probability for each MS/MS spectrum below or equal to 0.1 was required. In case the identified peptides of two proteins were the same or the identified peptides of one protein included all peptides of another protein, these proteins (e.g. isoforms and homologs) were combined and reported as one protein group. Phosphorylation sites were made non-redundant with regards to their surrounding peptide sequence. Nevertheless all alternative proteins that matched a particular phosphosite were reported as one group. The PTM score was used for assignment of the phosphorylation site(s) as described [27]. To match identifications across different replicates and adjacent fractions, the "match between runs" option in MaxQuant was enabled within a time window of 5 min.

2.9 *Bioinformatics analysis*

We based the annotation of proteins on Uniprot identifiers. Pathway membership information was obtained from the KEGG database (http://www.genome.jp/kegg/pathway.html). Phospho_ELM BLAST Search PhosphoSite (http://phospho.elm.eu.org/pELMBlastSearch.html) database was used to retrieve known phosphorylated sites and relative kinases.

2.10 *SILAC incorporation*

For incorporation test, a small aliquot of cells was lysed using the above mentioned protocol. Briefly, 10µg of total lysate were loaded on SDS-PAGE. A section of the gel was digested with trypsin, and peptides were purified on StageTips and analyzed with a single LC-MS/MS run. Raw files were processed by MaxQuant analysis to determine the ratio of heavy labeled peptides to the remaining unlabeled ones. The calculation for lysine- and arginine-containing peptides was performed separately. The probability density function analysis was obtained by kernel density estimation (KDE) using the R software package (ver. 2.10.1) (Figure 1–Supplementary data).

3. Results and Discussion

3.1 *Workflow: description and assessment*

In this study, we performed SILAC experiments to quantify changes in phosphorylation upon PDGF-BB stimulus. The schematic experimental workflow is depicted in Figure 1. One advantage of SILAC is that it allows the combination of differentially labeled protein samples early during sample preparation. This is especially important for the experiments related to this study because several sample fractionation and enrichment steps were involved. Our method combines SILAC for phosphopeptide quantitation, SCX fractionation,

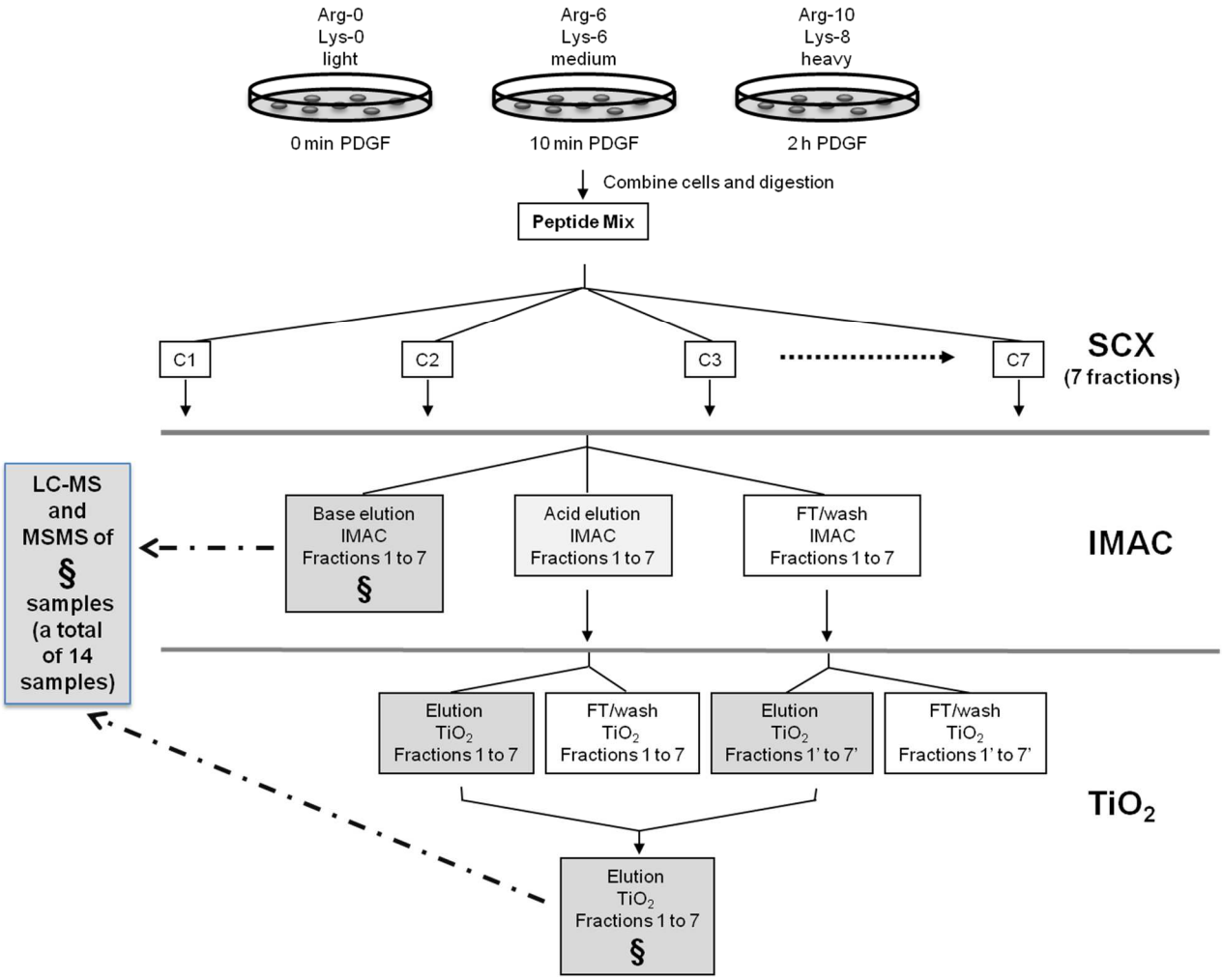

Figure 1. Schematic workflow representation. Primary VSMCs, were cultured and stimulated with PDGF-BB at different time and different SILAC medium (L for t0; M for t10minand H for t2h). Protein extracts were mixed 1:1:1 and enzymatically in solution digested. The resulting peptide mixtures were separated into 7 fractions by SCX fractionation. Every fraction was enriched for phosphopeptides using first IMAC and then, TiO2 beads. Only the 14 different sample fractions obtained after IMAC (7 fractions) and after TiO2 (7 fractions) were analyzed by LC-MS/ MS (fractions marked with §). Two technical LC-MS/MS replicates were examined for each fraction.

sequential elution from IMAC (SIMAC) [28] and multi-stage LC-MS to obtain a highly accurate phospho selection and identification. Primary VSMCs, obtained from pig coronary arteries, were cultured and stimulated with PDGF-BB (100 ng/mL) for 10 minutes and 2 hours and labeled with both arginine and lysine using three distinct isotope forms (L for t0, M for t10minand H for t2h). Protein extracts were mixed 1:1:1 and enzymatically digested in solution. The resulting peptide mixtures were separated into 7 fractions by SCX [29] and phosphopeptides enriched by IMAC [30, 31]. Every fraction was enriched in phosphopeptides using first IMAC and then, TiO2 beads.

Only the 14 different fractions obtained after IMAC (7 fractions, see figure 1) and after TiO2 (7 fractions, see figure 1) were analyzed by LC-MS/ MS, to specifically identify phosphor-peptides and proteins. Two replicates of each fraction were examined by online liquid chromatography on a linear ion trap/Orbitrap mass spectrometer. Each peptide was eluted as characteristic triplets: one peptide unlabeled for t0 and the other two peptides differently labeled on arginine or lysine for the two different PDGF-BB stimulation times (t10min and t2h). The intensity of each component of triplet reflects the relative amounts at the three time points chosen. Technical reproducibility of SILAC quantification for each peptide was proved to be really excellent (R2 > 0.93) (Figure 2-Supplementary data).

Comparing results in terms of phosphopeptides identification, TiO2 selection enable the identification of 86% of the total phosphopeptides, with only 12% overlapping with IMAC selection (Figure 2A).

Considering that the SILAC approach was applied for the first time to VSMC model to quantitatively analyze time-course modifications, we evaluated the level of labeled amino acid incorporation in these primary cells and the possible cellular morphological changes induced by the treatments. Labeling efficiency, calculated separately for lysine- and for arginine-containing peptides, was very high, above 96% for both labeled samples (M and H) (Figure 1, panel ASupplementary data) demonstrating both the feasibility of SILAC quantitation in primary VSMCs and the precision of the experiment (Figure 1, panel B_Supplementary data). Moreover, confocal microscope images obtained staining VSMCs with the specific marker a-SM-actin showed no differences in cell morphology and/or marker distribution in cells grown in the three SILAC media compared to those cultured in standard conditions (data not shown).

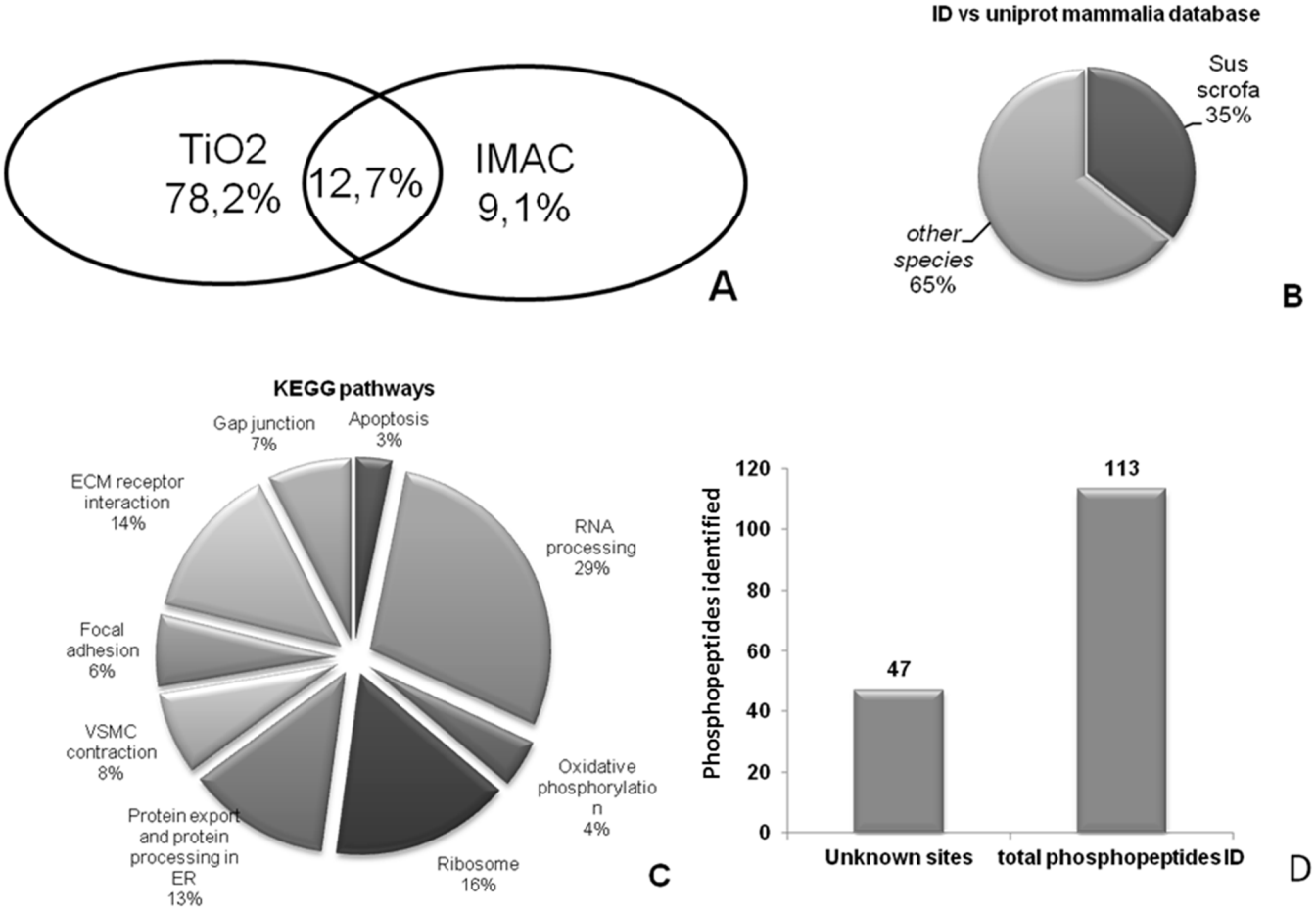

Figure 2. Characterization of identified proteins. A:Venn diagram compares the number of phosphorylated peptides identified in TiO_2 and IMAC fractions. B: percentage of peptides identified in *Sus scrofa* dataset. C: KEGG pathway analysis. D: analysis of total non-redundant identified phosphopetides.

Table 1. Differentially phosphorylated peptides. Proteins corresponding to the peptides are reported in the first column. The numbers of phosphorylation sites, the putative amino-acidic phosphorylated residues, the enzymes likely involved in the phosphorylations are indicated. The columns tagged as 10min and 2h quote the phosphorylation trend at 10 minutes or 2 hours after stimulation (▲; up-regulated, ▼; down-regulated) . The last two columns report the information on the known phosphorylation site.

Protein Name	KEGG Pathway	N°P-sites (STY)	AA	Probable Phospho (STY)	10'	2h	Known site	Best Motif
Calreticulin	Protein processing in ER	1	Y	FY(1)GDQEKDK	▲			
Gap junction alpha-1 protein (Connexin 43)	Gap junction	1	S	KLDAGHELQPLAIVDQRPSS(0.926)R	▲			**CAMK2**
Gap junction alpha-1 protein (Connexin 43)	Gap junction	1	S	ASS(0.988)RPRPDDLEI	▲		**Known site**	**CAMK2**
Gap junction alpha-1 protein (Connexin 43)	Gap junction	1	S	AS(0.5)SRPRPDDLEI	▲			**PKA**
60S acidic ribosomal protein P2	Ribosome	1	S	KEESEES(1)DDDMGFGLFD	▲	▲	**Known site**	**CK1**
60S acidic ribosomal protein P2	Ribosome	1	S	KEES(1)EESDDDMGFGLFD	▲	▲		
High mobility group protein HMG-I/HMG-Y		1	S	KLEKEEEEGISQES(0.9)SEEEQ	▲	▲	**Known site**	**CK2**
High mobility group protein HMG-I/HMG-Y		2	S	KLEKEEEEGISQES(0.998)S(0.998)EEEQ	▲	▲	**Known site**	**CK2**
HSP27		1	S	QLSSGVS(0.73)EIQQTADR		▲	**Known site**	**CK1**
Caldesmon	VSMC contraction	1	S	QSVDKVTS(0.967)PTKV		▲		
Filamin A		1	S	APS(1)VANVGSHCDLSLK	▼			**PKA/AKT**
PDZ and LIM domain protein 5		1	S	SPS(0.746)WQRPNQAAPSTGR	▼			
Vinculin	Focal adhesion	1	S	DPTAS(1)PGDAGEQAIR	▼			
MCG7378;Sec61 beta subunit; translocation complex beta variant	Protein processing in endoplasmic ER	1	S	PGPTPSGTNVGSSGRS(0.949)PSK	▼		**Known site**	**CDK2**
Nucleophosmin		1	S	LLSIS(1)GK	▼		**Known site**	**NEK6**
Aa2-041;BCL2-associated transcription factor 1, isoform CRA_a		1	Y	Y(0.414)SPSQNSPIHHIPSR	▼	▼	**Known site**	
Coiled-coil domain-containing protein 67		1	S	VLYQTHLVSLDAQQKLLS(0.99)EK	▼	▼		**GSK3**
Collapsin response mediator protein 4 long variant		1	S	GS(0.986)PTRPNPPVR	▼	▼	**Known site**	**DYRK2**
Collapsin response mediator protein 4 long variant		1	T	GSPT(0.777)RPNPPVR	▼	▼		**CDK2**
EPLIN-b		1	S	LENEETLERPAQLPNAAEIPQS(1)PGVEDAPIAK	▼	▼		**WGroupIV**
Ybx1 protein;Nuclease-sensitive element-binding protein 1		1	S	NYQQNYQNSES(0.598)GEKNEGSESAPEGQAQQR	▼	▼	**Known site**	

3.2 *Phosphoproteome analysis of PBGF-BB stimulated VSMCs*

Quantitation of phosphorylation sites was obtained via MaxQuant analysis [32]. A total of 42 LC- MS/MS analyses allowed the identification of a total of 1300 phosphopeptides distributed on 380 protein groups (see Supplementary Table 1 and 2). Although satisfactory, these data are underestimated owing to the partial completion of the *Sus scrofa* database [33]. In order to compensate this shortcoming we used the Uniprot cp mammalia database to identify not only phosphopeptides belonging to *Sus scrofa* but also phosphorylated peptides highly homologous in other mammalian species (Figure 2B).

Phospho_ELM BLAST Search PhosphoSite was used to look for novel, still unidentified, phosphorylated sites. Among the non-redundant 113 phosphorylated sites identified, 42% (corresponding to 47 new phosphorylation sites) were not yet previously reported in Mouse-Human database (identified through sequence homology between Pig and Mouse-Human) (see Supplementary Table 1 and Figure 2D). For the phospho-modulated proteins, more than 50% of the sites were identified as new in the Mouse-Human database.

It is worth noting that all identified phosphorylation sites are not present in the Sus Scrofa database highlighting the great drawback of using pig as a model organism. Notwith-

standing, pig remains an important animal model for cardiovascular disease experimentations and thus it could be valuable to annotate the database with our findings.

Using 1.7 fold as SILAC ratio cutoff (corresponding to $P < 0.05$ as calculated by MaxQuant significance B and using the ratio M/L and H/L normalized by MaxQuant on median intensity), we found 21 phosphopeptides statistically modulated at different times of stimulation (Table 1, Figure 3). Ten of them (belonging to 6 different proteins) were up-regulated, while the remaining eleven (belonging to 10 different proteins) resulted to be down-regulated upon PGDF-BB stimulation at both time points.

One potential concern for quantitative phosphoproteomics is that changes in protein expression can affect phosphopeptide/phosphoprotein ratios. In our SILAC analyses, for each replicate, we used the same cell population under the same growing conditions. At the last passage, cells were subdivided in three groups, all of them cultured for 24h without serum and afterwards two parts were treated with PDGF-BB as indicated while the thirdset of cells was left untreated. The time of treatment was short and thus we reasoned that it would not lead to significant changes in protein expression. This hypothesis was also confirmed from the analysis of the ratio M/L and H/L of the entire proteome identified (Figure 3 and 4-Supplementary data.) Before harvesting, cells were counted in order to use in the successive analyses the same number of them for each treatment.

KEGG pathway analysis of the identified phosphoproteins revealed that phosphorylation events in the activation process induced by PDGF-BB are involved in several functions including RNA and protein processing, extra cellular matrix (ECM)-receptor interaction, VSMC contraction, gap junction regulation and oxidative phosphorylation (Figure 2C). The most represented functions (RNA and protein processing) are related to transcription and translation activities. This last observation is interesting since the early events (from 10 min to 2h of growth factor stimulation) are initial stages towards the acquisition of the activated, synthetic phenotype which extensively engaged in new proteins synthesis.

Finally, it is interesting , the analysis of the kinases potentially involved in the activation process induced by PDGF-BB . Figure 4 shows the results of phosphopeptides that were predicted to be associated with kinase binding motif by SCANSITE. The computer algorithm SCANSITE was used to predict the kinase substrate relationship from dataset. The panel A shows the known motif distribution for all identified phosphorylation sites. The panel B shows instead the same analysis restricted for the sites of modulated proteins in our "in vitro" cardiovascular pathogenesis model. The analysis of the most frequent protein kinase related with the phospho-modulated proteins were normalized in order to better analyze the frequency of phosphorylation motif in our system. (Figure 4 panel C)

It is interesting to note that Ca2+/calmadulin-dependent protein kinases 2 (CAMK2), cyclin-dependent kinase 2 (CDK2) and casein kinase 1 (CK1) appear to be mainly involved in the signaling evoked during VMSC activation induced by PGDF-BB. In particular, the modulation of HSP27 and gap junction a-1 protein (connexin 43), seems to be relevant for VSMC phenotype changes. HSP27 has been recently demonstrated to be phosphorylated after PDGF-BB stimulation and to exert a role in VSMC de-differentiation through cytoskeleton remodeling [34].

Connexin 43 is the major protein of gap junctions and it is highly expressed in the neointima during the early stage of arteriosclerosis [35,36]. It has been shown that connexin 43 phosphorylation can regulate the early assembly of functional gap junctional channels [37].

It is also worth discussing the modulation of other proteins such as calreticulin, filamin A (FLNA), PDZ LIM protein 5 and vinculin, since also the modulation of these proteins might be involved in VSMC activation.

Vinculin is a membrane protein that links adhesion molecules to actin cytoskeleton and has a key role in modulating focal adhesion structures and functions thus controlling cell migration. The phosphorylation of Vinculin on tyrosine and serine is well-known and seems to regulate actin filament assembly and cell spreading [38,39,40].

Also the PDZ and LIM domain protein family is involved in focal adhesion assembly [41]. LIM and PDZ domains

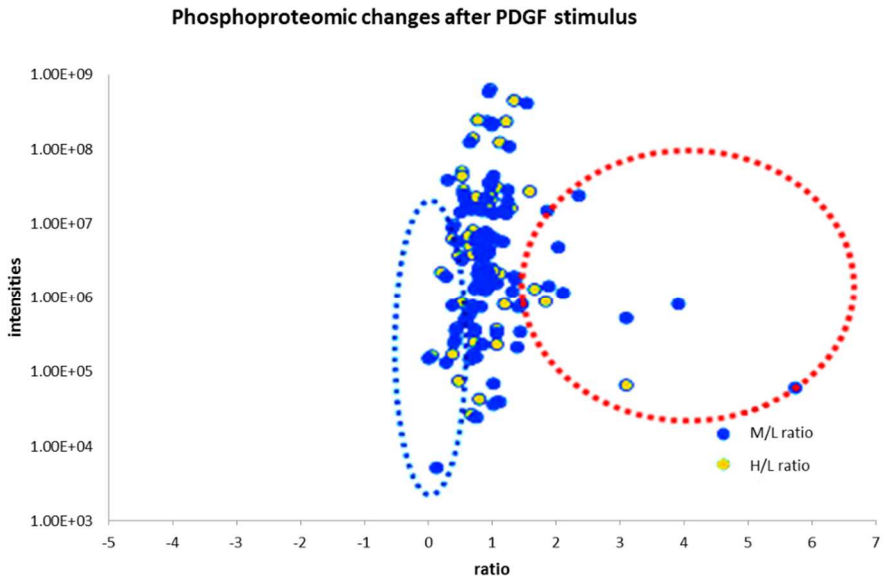

Figure 3. Quantitative phosphoproteomic analysis of PBGF-BB stimulated VSMCs . Scatter plot of the ratios M/L or H/L versus intensities of the identified phosphopeptides. Up-regulated phosphopeptides are grouped in the red circle, while in the blue circle the down-regulated ones are highlighted.

may have diverse functions but essentially they provide recognition sites for assembling multiprotein complexes and cytoskeleton. Noteworthy, PDZ-domain-containing proteins serve as scaffolds for assembling complexes involved in cell-cell junctions and signal transduction [42,43]. The extent and role of phosphorylation in modulating the biological functions of these proteins has been poorly investigated.

Caldesmon is an actin-binding protein that plays an essential role in the regulation of cytoskeleton organization modulating VSMC contraction and stabilizing stress fibers and focal adhesions. It has been demonstrated that phosphorylation decreases its actin-binding capacity thereby allowing dynamic changes of actin cytoskeleton that lead to motility and likely activation [44].

Also filamin A is an actin-binding protein and it regulates re-organization of cytoskeleton. It is essential for cell locomotion, interacting with integrins and membrane receptor complexes in VSMCs [45, 46]. Recently, it has been shown that phosphorylated filamin A regulates actin-linked caveolae dynamics [47], moreover, phosphorylation seems to facilitate filamin A binding to integrin suggesting a role, modulated by the degree of phosphorylation, as a mechano-sensor [48].

Calreticulin is an endoplasmic reticulum-resident chaperone that binds misfolded proteins and prevents them from being exported, but it is also found in the nucleus and thus may have a role in transcription regulation. More recently, an important role of serine de-phosphorylation and tyrosine phosphorylation on mRNA stability of angiotensin II receptor type 1 (AT1) receptor was investigated [49]. This receptor mediates important biological effects in VSMCs, such as proliferation, vasoconstriction and reactive oxygen species release [50].

It is interesting to underline that up- and down-phosphorylation events are equally represented at early times after stimulation. Probably, once PDGF-BB growth factor binds to its receptor, within a few minutes a cascade of kinases and phosphatases activate a series of biological events and the phosphorylation state of several proteins changes dramatically. As discussed previously, these early events, even if phenotypically imperceptible, are decisive for the activation process in VSMCs [6,10], and the entire reorganization of the cell structure.

4. Concluding Remarks

In summary, we performed for the first time a quantitative phosphoproteome analysis on primary VSMCs in order to investigate biological events induced by PDGF-BB stimulation. Moreover, the SILAC approach applied to VSMC model allowed evaluation of time-course modifications in a quantitative manner. Our approach allowed the identification of 1300 phosphopeptides, one hundred of them resulted novel phospho-sites in pig model, 47 of which are new even in Mouse-Human databases. Furthermore, in VSMC signaling events, some important factors involved in cytoskeleton remodeling, focal adhesions, gap junction assembly and cell activation were found differentially phosphorylated these early phosphorylations, pointed out here for the first time can be considered the starting point for the activation and the reorganization of VSMC.

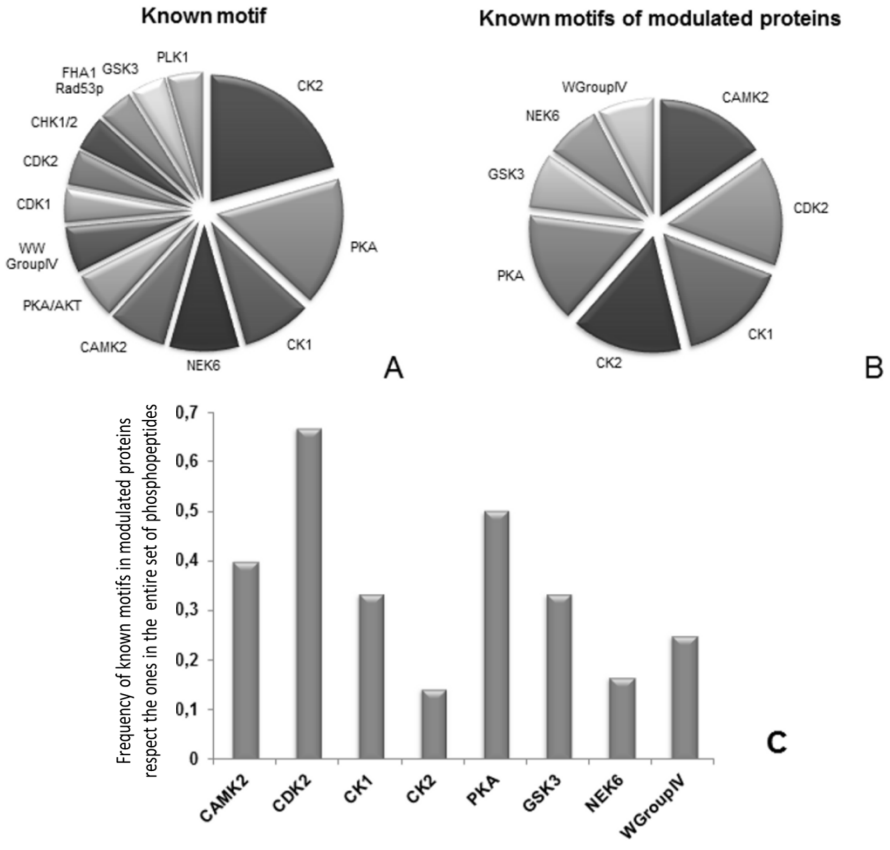

Figure 4. Analysis of the identified phosphopeptides: A: analysis of all identified phosphopetides. B: analysis of modulated phosphopeptides . C: frequency of known motifs in modulated proteins with respect to the entire list.

5. Supplementary material

Supplementary data and information is available at: http://www.jiomics.com/index.php/jio/rt/suppFiles/126/0.

Supplementary Material includes Supplementary Tables 1 (List of total identified peptide), 2 (List of

total identified groups of proteins) and Supplementary Data (Supplementary Figure 1, 2,3 and Figure 4).

References

1. JD Graves, EG Krebs in: Pharmacology and Therapeutics 1999; pp. 111–121.
2. T Hunter, Cell (2000) 113– 127.
3. G Manning, DB Whyte, R Martinez, T Hunter, S Sudarsanam, Science (2002) 1912–1934.
4. Adams MD, Myers EW, Li PW, Mural RJ, Sutton GG, et. al. Science (2001) 1304 –1351.
5. GK Owens, MS Kumar, BR Wamhoff, Physiological Reviews (2004) 767-801.
6. A Cecchettini, S Rocchiccioli, C Boccardi, L Citti, International review of cell and molecular biology (2011) 43-99.
7. R Ross, Nature (1993) 801-809.
8. WT Gerthoffer , Circulation Research (2007) 607-621.
9. C Lande, CBoccardi, L Citti , A Mercatanti , M Rizzo , S Rocchiccioli , L Tedeschi , MG Trivella , A Cecchettini , BMC Res Notes (2012) 268 doi:10.1186/1756-0500-5-268.
10. C Boccardi, A Cecchettini, A Caselli, G Camici, M Evangelista, A Mercatanti, G Rainaldi, L Citti, Cell and tissue research (2007) 185-95.
11. K Kawai-Kowase, GK Owens, Am. J. Physiol. Cell Physiol. (2007) C59–C69 doi: 10.1152/ajpcell.00394.2006.
12. HF Chen, LD Xie, CS Xu, Mol. Cell. Biochem (2010) 49–56.
13. C Liu , T Su , F Li , L Li , X Qin , W Pan , F Feng , F Chen , D Liao , L Chen , Acta Biochim. Biophys. Sin. (2010) 396–402.
14. PA Suwanabol, SM Seedial, X Shi, F Zhang, D Yamanouchi, D Roenneburg, B Liu, KC Kent, J Vasc Surg. (2012) 446-54.
15. Mann M, Ong SE, Gronborg M, Steen H, Jensen ON, Pandey A. Analysis of protein phosphorylation using mass spectrometry: deciphering the phosphoproteome. Trends Biotechnology 2002;20: 261–8.
16. LB Areces, V Matafora, A Bachi, Eur J Mass Spectrom (Chichester, Eng) (2004) 383-92.
17. B Bodenmiller, R Aebersold, Methods Enzymol. (2010) 317 -34.
18. SE Ong, B Blagoev, I Kratchmarova, DB Kristensen, H Steen, A Pandey, M Mann, Molecular and Cellular Proteomics (2002) 376–86.
19. F Delom, E Chevet, Proteome Science (2006) 4:15 doi:10.1186/1477-5956-4-15.
20. GH Han, Ml Ye, HF Zou, Analyst (2008) 1128-38.
21. A Leitner, Trends in Analytical Chemistry (2010) 177-85.
22. F Torta, M Fusi, CS Casari, CE Bottani, A Bachi, J Proteome Res. (2009) 1932-42.
23. T Christen, ML Bochaton-Piallat, P Neuville, S Renzen, Circ Res (1999) 99-107.
24. V Matafora, A D'Amato, S Mori, F Blasi, A Bachi, Mol Cell Proteomics (2009) 2243-55.
25. M. J. Schroeder, J. Shabanowitz, J. C. Schwartz, D. F. Hunt, J. J. Coon, Anal. Chem (2004) 3590-8.
26. 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 (2005) 2010–21.
27. J. V. Olsen, B. Blagoev, F. Gnad, B. Macek, C. Kumar, P. Mortensen, M. Mann, Cell (2006) 635–48.
28. S Tine, E Thingholm, ON Jensen, JP Robinson, MR Larsen, Molecular & Cellular Proteomics (2008) 661–71.
29. B Boris Macek, I Mijakovic, JV Olsen, F Gnad, C Kumar, PR Jensen, M Mann, Molecular & Cellular Proteomics (2007) 697–707.
30. GT Cantin, W Yi, B Lu, SK Park, T Xu, JD Lee, JR Yates III, Journal of Proteome Research (2008) 1346–51.
31. B Zhai, J Villén, SA Beausoleil, J Mintseris, SP Gygi, Journal of Proteome Research (2008) 1675–82.
32. WX Schulze , M Mann, Journal of Biological Chemistry (2004) 10756-64.
33. S Rocchiccioli, L Citti, C Boccardi, N Ucciferri, L Tedeschi, C Lande, MG Trivella, A Cecchettini. Proteome Science (2010) 8:15 doi:10.1186/1477-5956-8-15.
34. HF Chen, LD Xie, CS Xu, Molecular and Cellular Biochemistry (2010) 49–56.
35. BR Kwak, N Veillard, G Pelli, F Mulhaupt, RW James, Circulation (2003) 1033-39.
36. CE Chadjichristos, CM Matter, I Roth, E Sutter, G Pelli, TF Lüscher, M Chanson, BR Kwak, Circulation (2006) 2835-43.
37. JL Solan, PD Lampe, Biochimica et Biophysica Acta (2005) 154-63.
38. Z Zhang, G Izaguirre, SY Lin, HY Lee, E Schaefer, B Haimovich. Molecular Biology of the Cell (2004) 4234–47.
39. K Küpper, N Lang, C Möhl, N Kirchgeßner, S Born, WH Goldmann, R Merkel, B Hoffmann, Nature (2010) 171–5.
40. K Küpper , N Lang , C Möhl , N Kirchgessner , S Born , WH Goldmann , R Merkel , B , Biochemical and Biophysical Research Communications (2010) 560-4.
41. N Tamura, K Ohno, T Katayama, N Kanayama, Ko Sat, Biochemical and Biophysical Research Communications (2007) 589–94.
42. T Khurana, B Khurana, AA Noegel, Protoplasma (2002) 1–12.
43. ZH Baruch, AL Wendell, Journal of Cell Science (2001) 3219-31.
44. Q Jiang, R Huang, S Cai, ALA Wang, Journal of Biomedical Science (2010) 17:6 doi: 10.1186/1423-0127-17-6
45. N Yu, L Erb, R Shivaji, GA Weisman, CI Seye, Circulation Research (2008) 581-8.
46. Y Xu, TA Bismar, J Su, B Xu, G Kristiansen, Z Varga, L Teng, DE Ingber, A Mammoto, R Kumar, MA Alaoui-Jamali, Journal Expert Medicine (2010) 2421-37.
47. O Murie, A Echarri, C Hellriege, DM Pavón, L Beccari, MA Del Pozo, Journal of Cell Science (2011) 2763-76.
48. HS Chen, KS Kolahi, MRK Mofrad, Biophysical Journal (2009) 3095–04.
49. CFH Mueller, K Wassmann, A Berger, S Holz, S Wassmann, G Nickenig, Biochemical and Biophysical Research Communications (2008) 669–74.
50. G Nickenig, DG Harrison, Circulation (2002) 393–406.

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

ORIGINAL ARTICLE | DOI: 10.5584/jiomics.v3i1.130

Values for evaluating the nutritional status of water-soluble vitamins in humans

Katsumi Shibata, Tsutomu Fukuwatari

Department of Nutrition, School of Human Cultures, University of Shiga Prefecture, Hikone, Shiga 522-8533, Japan

Received: 12 February 2013 **Accepted:** 21 April 2013 **Available Online**: 9 May 2013

ABSTRACT

Previously, we clarified that the amount of urinary excretion of water-soluble vitamins closely reflects the surplus amount of water-soluble vitamins in the body stores of rats and humans. We tried to set a tentative amount of urinary excretion of eight water-soluble vitamins of nine water-soluble vitamins (except vitamin B_{12}) for maintaining health based on experiments in healthy young females administered a semi-chemically defined diet according to Japanese Dietary Reference Intakes and related data. We proposed a tentative value for the amount of urinary excretion of water-soluble vitamins for maintaining health. The values were: 200–2000 nmol/d for vitamin B_1; 200–2000 nmol/d for vitamin B_2; 2–15 μmol/d for 4-pyridoxic acid (a catabolite of vitamin B_6); 50–300 μmol/d for the sum of the nicotinamide catabolites N^1-methylnicotinamide, N^1-methyl-2-pyridone-5-carboxamide, and N^1-methyl-4-pyridone-3-carboxamide; 10–30 μmol/d for pantothenic acid; 15–100 nmol/d for folate; 50–200 nmol/d for biotin; and 100–2000 μmol/d for vitamin C. By using these values, we attempted to evaluate the nutritional status of water-soluble vitamins for 709 young Japanese females. The percentage within the tentative value of urinary excretion of water-soluble vitamin for maintaining health was 73.6% for vitamin B_1, 63.5% for vitamin B_2, 90.0% for vitamin B_6, 85.6% for niacin, 58.1% for folate, 85.6% for pantothenic acid, 70.2% for biotin, and 65.4% for vitamin C. The percentage beyond the lower limit of detection was 22.4% for vitamin B_1, 31.3% for vitamin B_2, 6.2% for vitamin B_6, 14.0% for niacin, 40.9% for folate, 12.4% for pantothenic acid, 26.2% for biotin, and 33.0% for vitamin C. The percentage over the upper limit of detection was 4.1% for vitamin B_1, 5.2% for vitamin B_2, 3.8% for vitamin B_6, 0.4% for niacin, 1.0% for folate, 2.0% for pantothenic acid, 3.6% for biotin, and 1.6% for vitamin C. Nutritional assessment using urinary excretion amounts of water-soluble vitamins is persuasive, and leads to the transformation of habitual dietary intakes.

Keywords: Urine; Water-soluble vitamins; Reference value; Human; Evaluation.

Abbreviations:

DRIs, dietary reference intakes**; MNA**, N^1-methylnicotinamide; 4-**PIC**, 4-pyridoxic acid; 2-**Py**, N^1-methyl-2-pyridone-5-carboxamide; **4-Py**, N^1-methyl-4-pyridone-3-carboxamide; **PteGlu**, pteroylmonoglutamic acid.

1. Introduction

Nutritional evaluation for individual persons is important because the metabolic ability is not the same. The method that has been used frequently involves recording dietary intake and calculating nutrient intake. Another approach to evaluate nutritional status involves using biomarkers such as blood and urine samples.

We reported that the urinary excretory amounts of water-soluble vitamins closely reflect the surplus amount of water-soluble vitamins in the bodies of rats [1–15] and humans [16–27]. The nutritional assessment using biomarkers is persuasive, and leads readily to the transformation of habitual dietary intakes. Thus, we tried to set tentative amounts of urinary excretion for eight water-soluble vitamins of nine water-soluble vitamins (except vitamin B_{12})

***Corresponding author:** Katsumi Shibata, Department of Nutrition, School of Human Cultures, University of Shiga Prefecture, Hikone, Shiga 522-8533, Japan. E-mail: kshibata@shc.usp.ac.jp. Tel: +81-749-28-8449. Fax: +81-749-28-8499.

for preventing water-soluble vitamin deficiency and its toxicity of excess intake for maintaining health in adults. The method was applied for evaluation of the status of water-soluble vitamins for young Japanese females.

2. Material and Methods

The study protocol was reviewed and approved by the Ethical Committee of the University of Shiga Prefecture (Shiga, Japan).

2.1. *A semi-chemically defined diet experiment to calculate tentative urinary excretion amounts of eight water-soluble vitamins of nine water-soluble vitamins (except vitamin B_{12}) for maintaining health*

2.1.1. Subjects

We enrolled 20 healthy female Japanese college students. Their ages, body weights and heights were 21.2 ± 0.7 y (mean ± SD), 163.9 ± 3.7 cm, and 54.2 ± 1.1 kg, respectively. Before experimentation, they underwent a physical examination. Their hematological and blood biochemical analyses showed normal values.

2.1.2. Diet and experimental design

All subjects were housed in the same facility for 9 days. The experimental design is the same as that shown in a previous report [17]. The experimental design is the same as in a preceding article [17]. The subjects took a semi-purified diet including vitamins based on Japanese Dietary Reference Intakes (DRIs) [28] during the experiment. The diet consisted of wheat flour, gluten, gelatinized-cornstarch, sucrose, soybean oil, rapeseed oil, lard, soluble dietary fiber, insoluble dietary fiber and mineral mixture. The diet contained 1,800 kcal/d of energy, 55 g/d of protein, 40 g/d of fat and 292 g/d of carbohydrate. The vitamin mixture contained 0.9 mg/d (2.7 μmol/d) of thiamin hydrochloride, 1.0 mg/d (2.7 μmol/d) of riboflavin, 1.5 mg/d (5.7 μmol/d) of pyridoxine hydrochloride, 2.4 μg/d (1.8 nmol/d) of cyanocobalamin, 2.8 mg/d (23 μmol/d) of nicotinamide and 9.2 mg of nicotinamide equivalent derived from tryptophan in protein, 5.5 mg/d (23 μmol/d) of calcium pantothenate, 200 μg/d (453 nmol/d) of pteroylmonoglutamic acid, 30 μg/d (123 nmol/d) of D(+)-biotin, and 100 mg/d (568 μmol/d) of L(+)-ascorbic acid.

When subjects are administered the same diet for several days, the urinary excretion amounts of water-soluble vitamins become constant after 3 d [22]. Twenty-four-hour urine samples were collected from the second urinary excretion on d 4 to the first one on d 5 (referred to as “urine sample d 4”). Urine samples were also collected on d 5, 6, and 7. After the volumes of the urine samples had been measured, the collected urine samples were treated immediately as described in the section "Analyses" to avoid destruction of water-soluble vitamins. They were then stored at –20°C until needed.

2.2. *Collection of urine samples for evaluating levels of water-soluble vitamins*

2.2.1. Subjects

The subjects were female Japanese college students (n = 709). They consumed the diet and had a lifestyle that did not involve restrictions.

2.2.2. Twenty-four-hour urine collection

A single 24-h urine sample was collected each day. Subjects were instructed in writing and verbally on the methods of urine collection and the necessity of obtaining a complete 24-h urine collection. Subjects were requested to eat and drink normally during the collection and to follow their usual pattern of activity. Subjects were then provided with a bag, three or four 1-L plastic bottles (containing no additives) and ten 400-mL cups. A recording sheet was provided. In the morning, subjects were asked to discard the first specimen and to record the time (usually 06:00–09:00 h) on the sheet (the start of the collection period). Subjects were asked to collect all specimens by the time of the start of the collection period the following morning. If a specimen was missed, subjects were asked to record the estimated volume of missing urine and the time. The following morning, subjects were asked to collect the last specimen at the time when the specimen was discarded the previous morning, and record the time on the collection sheet (the end of the collection period). The collection sheet was reviewed by the research staff when the samples were returned, and missing information was obtained from the subjects. The height of urine in each bottle was measured and later converted into volume using an empirical formula based on repeated measurements of volume in identical bottles. All urine from the 24-h collection period was then combined and mixed thoroughly by vigorous stirring. Urinary aliquots were taken and used for determination of vitamins and metabolites.

2.3. *Chemicals*

Wheat flour (soft flour, first grade) was obtained Nisshin Flour Milling Inc. (Tokyo, Japan). Wheat gluten, gelatinized cornstarch, soybean oil, 13 types of vitamins, and minerals were purchased from Wako Pure Chemical Industries (Osaka, Japan). Rapeseed oil was from Ajinomoto Co., Ltd. (Tokyo, Japan). Coconut oil and lard were obtained from Clea Japan (Tokyo, Japan). "Fibersol", used as a soluble dietary fiber, was purchased from Matsutani Chemical Industries (Osaka, Japan), and “Ramie powder”, used as an insoluble dietary fiber, was from Tosco (Tokyo, Japan).

Thiamin hydrochloride ($C_{12}H_{17}C_lN_4OS$-HCl = 337.27), riboflavin ($C_{17}H_{20}N_4O_6$ = 376.37), cyanocobalamin

($C_{63}H_{88}CoN_{14}O_{14}P$ = 1355.40), nicotinamide ($C_6H_6N_2O$ = 122.13), calcium pantothenate ($C_{18}H_{32}N_2O_{10}$-Ca = 476.54), folic acid ($C_{19}H_{19}N_7O_6$ = 441.40), D(+)-biotin ($C_{10}H_{16}N_2O_3S$ = 244.31), and L(+)-ascorbic acid ($C_6H_8O_6$ = 176.13) were purchased from Wako Pure Chemical Industries. N^1-Methylnicotinamide (MNA) chloride ($C_7H_9N_2O$-HCl = 159.61) was obtained from Tokyo Kasei Kogyo (Tokyo, Japan). N^1-Methyl-2-pyridone-5-carboxamide (2-Py, $C_7H_8N_2O_2$ = 152.15) and N^1-methyl-4-pyridone-3-carboxamide (4-Py, $C_7H_8N_2O_2$ = 152.15) were synthesized using the methods of Pullman and Colowick [29] and Shibata et al. [30], respectively. All other chemicals used were of the highest purity available from commercial sources.

2.4. *Determination of vitamins and metabolites in urine*

For the analysis of thiamin, 1 mL of 1 mol/L HCl was added to 9 mL of urine. Urinary content of thiamin was determined by the high-performance liquid chromatography (HPLC)-post labeled fluorescence method [31].

For the analysis of riboflavin, 1 mL of 1 mol/L HCl was added to 9 mL of urine. Urinary content of riboflavin was determined by HPLC [32].

For the analysis of 4-PIC (a metabolite of pyridoxal), 1 mL of 1 mol/L HCl was added to 9 mL of urine. Urinary content of 4-PIC was determined by HPLC [33].

For the analysis of vitamin B_{12}, acetate buffer and potassium cyanide were added to urine, and vitamin B_{12} in urine was converted to cyanocobalamin by autoclave [34]. Urinary content of cyanocobalamin was determined by the microbioassay method using *Lactobacillus leichmannii*, ATCC 7830 [34].

For the analysis of MNA, 2-Py, 4-Py, and nicotinamide metabolites, 1 mL of 1 mol/L HCl was added to 9 mL of urine. Urinary content of 2-Py, 4-Py and MNA was determined by HPLC [30, 35].

For the analysis of pantothenic acid, a urine sample was injected directly into a HPLC system [36].

For the analysis of folate, 1 mL of 1 mol/L L(+)-ascorbic acid was added to 9 mL of urine. Urinary content of folate was determined by the microbioassay method using *Lactobacillus rhamnosus*, ATCC 27773 [37].

For the analysis of ascorbic acid, 4 mL of 10% metaphosphate was added to 4 mL of urine. Urinary ontent of reduced and oxidized ascorbic acid and of 2,3-diketogulonic acid was determined by HPLC [38].

2.5. *Statistical analyses*

Spearman correlation coefficients were calculated to determine the association between a urinary vitamin and another urinary vitamin. Graph Pad Prism 5.0 (Graph Pad Software, San Diego, CA, USA) was used for all statistical analyses.

3. Results and Discussion

3.1. *Selection of surrogate indicators for evaluating water-soluble vitamins in urine*

A potential approach for calculating the tentative urinary excretion amounts of eight water-soluble vitamins of nine water-soluble vitamins (except vitamin B_{12}) for maintaining health in apparently healthy subjects is based on the observation that a water-soluble vitamin or its catabolites can be detected in the urine. For example, the nicotinamide catabolite MNA is traditionally used for a surrogate indicator of niacin nutritional status [39].

The catabolites of thiamin in human urine such as 4-methylthiazole-5-acetic acid [40], and 2-methyl-4-amino-5-pyrimidine carboxylic acid [41] have been reported. However, the synthetic methods for standard chemicals and the practical methods of measurement have not been successful. Thus, we chose thiamin for evaluating the nutritional status of vitamin B_1.

The catabolites of riboflavin in human urine (7 alpha-hydroxyriboflavin and 8 alpha-hydroxyriboflavin) have been reported [31], but the catabolites have not been synthesized. Thus, we chose riboflavin for evaluating the nutritional status of vitamin B_2.

4-PIC is well known as a major catabolite of vitamin B_6 [33] and is from available commercial resources. Thus, we chose 4-PIC for evaluating the nutritional status of vitamin B_6.

As shown in the present study, vitamin B_{12} is also eliminated into urine. However, the major elimination pathway is bile via the enterohepatic circulation [42], and not urine via the kidney. Furthermore, the amount of urinary excretion does not reflect the intake of vitamin B_{12} [23]. Therefore, the data on vitamin B_{12} are informative. We already clarified that the amount of urinary excretion of vitamin B_{12} reflects urine volume [23].

The metabolism of niacin in humans is well known (Fig. 1). Nicotinamide is biosynthesized from the essential amino acid L-tryptophan. The conversion amount is reported to be ≈1 mg of nicotinamide from 60 mg of L-tryptophan [18,43]. Nicotinic acid cannot coin in humans under a physiological concentration of nicotinamide (≈0.1 mmol/L) because the Km value of nicotinamide deamidase (which catalyzes the reaction of nicotinamide → nicotinic acid) is very high (≈0.1 mol/L) [44]. Indeed, we could not detect nicotinic acid in human urine as well as in rat urine [45]. Ingested nicotinic acid is a good precursor of NAD^+ and therefore of nicotinamide (Fig. 1). In rat experiments, even when rats are administered a diet containing nicotinic acid, we could not detect nicotinic acid in their urine [46, 47]. Detection of nicotinic acid was limited only when a large amount of nicotinic acid (for example, a 1.0% nicotinic acid diet) [48] or nicotinamide (1.0% nicotinamide diet) [49] was fed to rats. In human experiments, even when 150 mg of nicotinamide was administered for 44 weeks, we could not

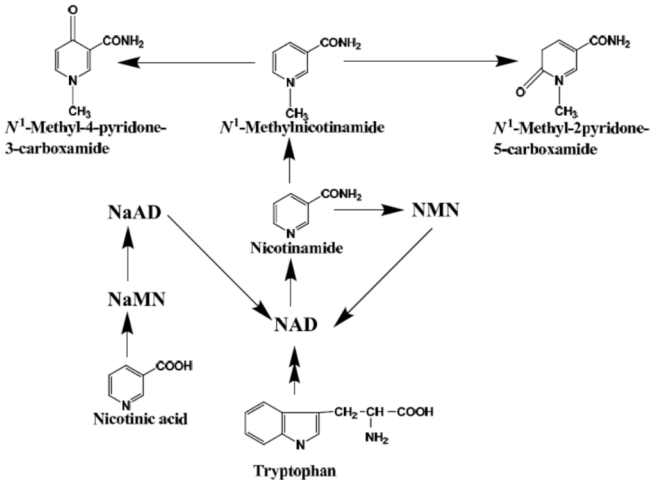

Figure 1. Metabolism of nicotinamide in humans. NaMN, nicotinic acid mononucleotide; NaAD, nicotinic acid adenine dinucleotide; NMN, nicotinamide mononucleotide.

find any amount of nicotinic acid in urine [50]. Therefore, in our opinion, nicotinic acid itself cannot be used for evaluating niacin nutritional status. Nicotinamide can be detected in human urine [50,51], but the amount is approximately the limit of detection even when 150 mg of nicotinamide is administered [50]. Therefore, nicotinamide cannot be used for evaluating niacin nutritional status. Catabolites of nicotinamide in humans are known: MNA, 2-Py, and 4-Py (Fig. 1) [30]. MNA can be obtained from commercial resources, but 2-Py and 4-Py cannot. Shibata et al. successfully synthesized 2-Py and 4-Py [30] and showed that MNA, 2-Py, and 4-Py are major urinary catabolites in humans [30]. Thus, we chose the sum of MNA, 2-Py, and 4-Py for evaluating niacin nutritional status.

The catabolism of folate in humans is known. These catabolites are the cleavage products of the C-9—N-10 bond of folate, including *p*-aminobenzoylglutamate and *p*-acetamidobenzoylglutamate [52,53]. In general, the measurement methods for the catabolites are not practical and straightforward. Therefore, they are difficult to carry out as routine work. Thus, we chose folate for evaluating the nutritional status of folate. In the present study, the urinary excretory levels of folate was measured by microbiological assay using *L. rhamnosus*, ATCC 27773 [37]. This assay can use many types of folate compounds, such as pteroylmonoglutamic acid (PteGlu), dihydroPteGlu, tetrahydroPteGlu, 5-formyltetrahydroPteGlu, 10-formyltetrahydroPteGlu, 5,10-methylenetetrahydroPteGlu, and 5-methyltetrahydroPteGlu, as growth factors [54]. Therefore, the urinary amounts in the present study did not contain the catabolites of folate.

The catabolites of pantothenic acid in human urine are not known. Pantoic acid and β-alanine are not the enzymatic degradation products of pantothenic acid but are non-enzymatic degradation products, the reaction of which proceeds under acidic conditions. Thus, we chose pantothenic acid for evaluating the nutritional status of pantothenic acid.

The catabolites of biotin in human urine are known. Mock et al. [55] reported that biotin is catabolized to bisnorbiotin and biotin sulfoxide in humans, and that the bioassay organism grows equally well on the biotin as well as the biotin metabolites present in urine. The standard catabolites are not purchased from commercial sources but, in the present study, *Lactobacillus plantarum* was used as the bioassay organism to assess biotin. In the present study, the urinary excretion levels for bisnorbiotin and biotin sulfoxide as well as biotin are detailed. Thus, we chose the sum of biotin and its catabolite for evaluating the nutritional status of biotin.

The catabolism of vitamin C is not well known. However, we know that reduced ascorbic acid → oxidized ascorbic acid → 2,3-diketogulonic acid → oxaloacetic acid [38]. In the present study, we measured three major compounds which occur in urine: reduced ascorbic acid, oxidized ascorbic acid, and 2,3-diketogulonic acid. In the present study, the measurement method used can detect total osazones (into which the three compounds were converted). Thus, we chose total osazones for evaluating the nutritional status of vitamin C.

3.2. *Tentative urinary excretion amounts of eight water-soluble vitamins of nine water-soluble vitamins (except vitamin B_{12}) for maintaining health*

Apparently healthy subjects were administered a chemically defined vitamin diet followed by Japanese DRIs for 7 d. Daily urine samples were collected from d 4 to d 7. The frequency distribution of urinary excretion of each of eight vitamins is shown in Fig. 2. The minimum and fifth percentile urine excretion of each vitamin is described in Table 1. The maximum and 95th percentile urinary excretion of each vitamin is described in Table 2. We have shown the average urinary excretion when subjects are administered a diet containing 3-fold levels of vitamins compared with Japanese DRIs [20]. Such doses are safe for maintaining health because no accumulation occurs in blood [20]. We tried to set a tentative value of urinary excretion for each water-soluble vitamin for maintaining health from these results (Fig. 1, Table 1, and Table 2). Proposed values are shown in Table 3.

3.3. *Evaluation of the nutritional status of water-soluble vitamin in Japanese females*

Twenty-four-hour urine samples (n = 709) from females who lived and consumed a diet in an uninhibited manner were collected each day. Table 4 shows the minimum, maximum, median, and mean ± SD of each water-soluble vitamin. Figure 3 shows a fragmentary view of the histograms (with a focus on omitting extremely high excretion) of daily urinary excretion amounts of vitamin B_1, vitamin B_2, vitamin B_6, vitamin B_{12}, niacin, folic acid,

Table 1. Minimum urinary excretion of water-soluble vitamins in apparently healthy young Japanese females fed a chemically defined vitamin diet followed by Japanese DRIs.

Water-soluble vitamin	Minimum urinary excretion when fed a chemically defined vitamin diet followed by Japanese DRIs	Urinary excretion at 5th percentile when fed a chemically defined vitamin diet followed by Japanese DRIs	Proposed lower limit of excretion for maintaining health
Vitamin B_1 (as thiamin)	181 nmol/d	259 nmol/d	200 nmol/d
Vitamin B_2 (as riboflavin)	205 nmol/d	284 nmol/d	200 nmol/d
Vitamin B_6 (as 4-pyridoxic acid)	1.5 μmol/d	1.8 μmol/d	2.0 μmol/d
Vitamin B_{12} (as cyanocobalamin)	70 pmol/d	87 pmol/d	not determined
Niacin (as sum of MNA + 2-Py + 4-Py)	46 μmol/d	52 μmol/d	50 μmol/d
Folate (as sum of pteroylmonoglutamic acid (PteGlu) + dihydroPteGlu + tetrahydroPteGlu + 5-formyltetrahydroPteGlu + 10-formyltetrahydroPteGlu + 5,10-methylenetetrahydroPteGlu + 5-methyltetrahydroPteGlu)	14 nmol/d	16 nmol/d	15 nmol/d
Pantothenic acid	10.2 μmol/d	12.4 μmol/d	10 μmol/d
Biotin (as sum of biotin + bisnorbiotin + biotin sulfoxide)	53 nmol/d	58 nmol/d	50 nmol/d
Vitamin C (sum of reduced and oxidized ascorbic acid + 2,3-diketogulonic acid)	80 μmol/d	94 μmol/d	100 μmol/d

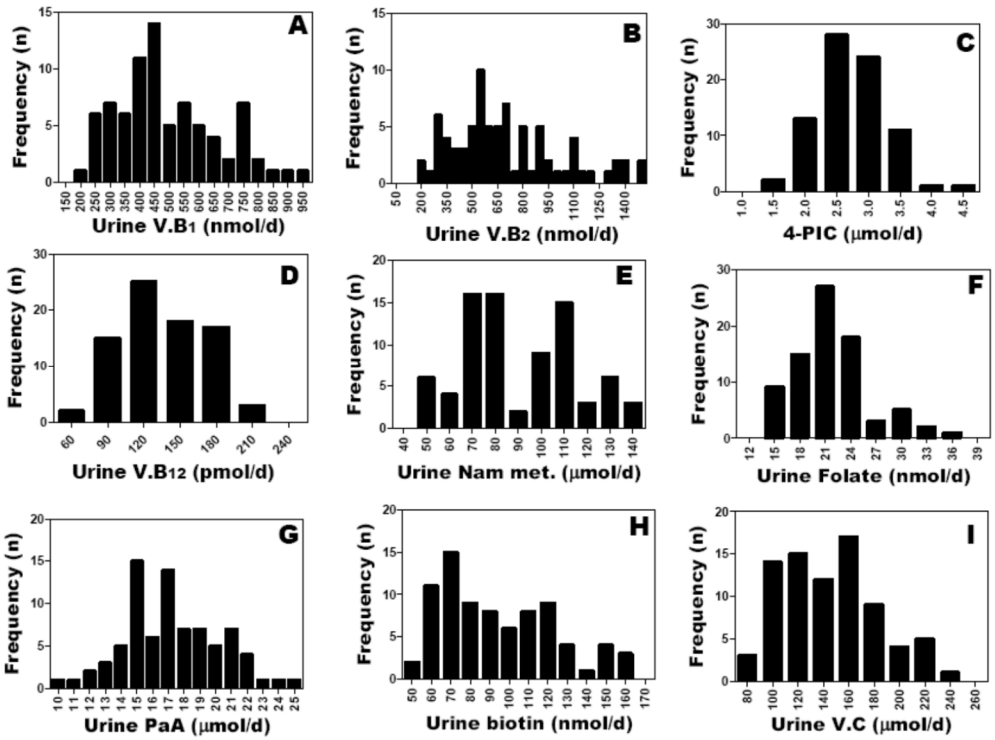

Figure 2. Histograms of daily urinary excretion amounts of (A) vitamin B_1, (B) vitamin B_2, (C) vitamin B_6*1, (D) vitamin B_{12}, (E) niacin*2, (F) folates*3, (G) pantothenic acid, (H) biotin*4, and (I) vitamin C*5 in apparently healthy young Japanese females administered a chemically defined vitamin diet followed Japanese DRIs.
*1 Vitamin B_6; the value is 4-PIC, a major catabolite of V.B_6.
*2 Niacin; the value is the sum of MNA, 2-Py, and 4-Py, which are major catabolites of nicotinamide.
*3 Folates; the value is the sum of pteroylmonoglutamic acid (PteGlu) + dihydroPteGlu + tetrahydroPteGlu + 5-formyltetrahydroPteGlu + 10-formyltetrahydroPteGlu + 5,10-methylenetetrahydroPteGlu + 5-methyltetrahydroPteGlu.
*4 Biotin; the value is the sum of biotin + bisnorbiotin + biotin sulfoxide
*5 Vitamin C; the value is the sum of the reduced and oxidized ascorbic acid and its catabolite 2,3-diketogulonic acid.

Table 2. Maximum urinary excretion of water-soluble vitamins in apparently healthy young Japanese females fed a chemically defined vitamin diet followed by Japanese DRIs.

Water-soluble vitamin	Maximum urinary excretion when fed a chemically defined vitamin diet followed by Japanese	Urinary excretion at 95th percentile when fed a chemically defined vitamin diet followed by Japanese	Average urinary excretion when administered a diet containing 3-fold vitamins compared with DRIs*1	Proposed upper limit of excretion for maintaining health
Vitamin B_1 (as thiamin)	930 nmol/d	782 nmol/d	2042 nmol/d	2000 nmol/d
Vitamin B_2 (as riboflavin)	1492 nmol/d	1370 nmol/d	1654 nmol/d	2000 nmol/d
Vitamin B_6 (as 4-pyridoxic acid)	4.4 μmol/d	3.6 μmol/d	12.6 μmol/d	15 μmol/d
Vitamin B_{12} (as cyanocobalamin)	215 pmol/d	184 pmol/d	no data	not determined
Niacin (as sum of MNA + 2-Py + 4-Py)	139 μmol/d	131 μmol/d	280 μmol/d	300 μmol/d
Folate (as sum of pteroylmonoglutamic acid (PteGlu) + dihydroPteGlu + tetrahydroPteGlu + 5-formyl-tetrahydroPteGlu + 10-formyl-tetrahydroPteGlu + 5,10-methylene-tetrahydroPteGlu + 5-methyltetrahydroPteGlu)	37 nmol/d	30 nmol/d	98 nmol/d	100 nmol/d
Pantothenic acid	24.6 μmol/d	22.3 μmol/d	39.4 μmol/d	40 μmol/d
Biotin (as sum of biotin + bisnorbiotin + biotin sulfoxide)	163 nmol/d	154 nmol/d	207 nmol/d	200 nmol/d
Vitamin C (sum of reduced and oxidized ascorbic acid + 2,3-diketogulonic acid)	244 μmol/d	215 μmol/d	1956 μmol/d	2000 μmol/d

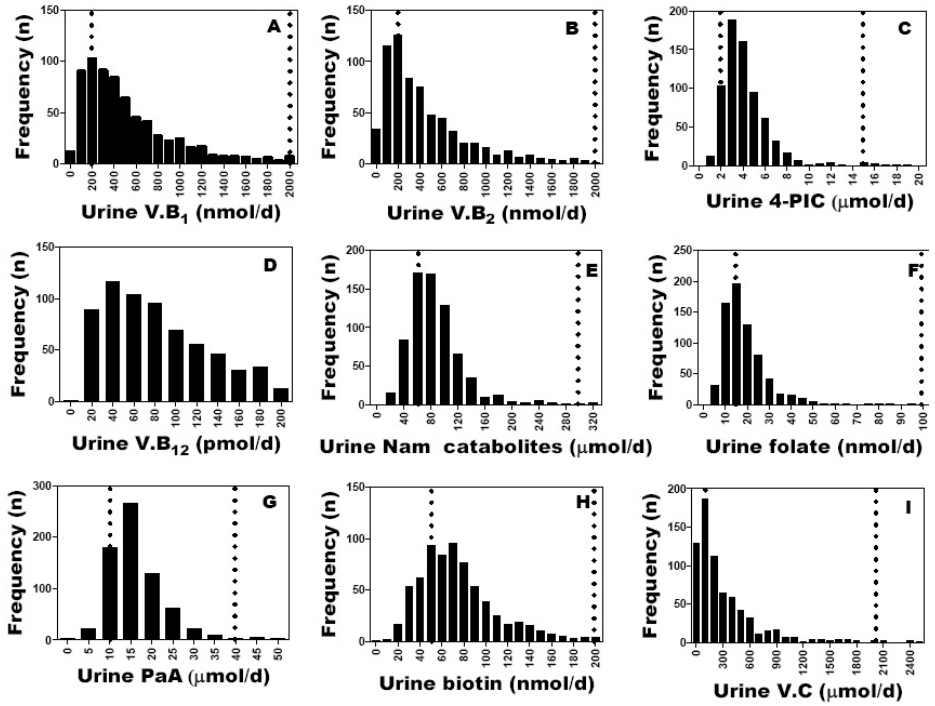

Figure 3. Fragmentary view of histograms (with a focus on omitting extremely high excretion) of daily urinary excretion amounts of (A) vitamin B_1, (B) vitamin B_2, (C) vitamin B_6*1, (D) vitamin B_{12}, (E) niacin*2, (F) folates*3, (G) pantothenic acid, (H) biotin*4, and (I) vitamin C*5 in apparently healthy young Japanese females.

*1 Vitamin B_6; the value is 4-PIC, a major catabolite of $V.B_6$.

*2 Niacin; the value is the sum of MNA, 2-Py, and 4-Py, which are major catabolites of nicotinamide.

*3 Folates; the value is the sum of pteroylmonoglutamic acid (PteGlu) + dihydroPteGlu + tetrahydroPteGlu + 5-formyltetrahydroPteGlu + 10-formyltetrahydroPteGlu + 5,10-methylenetetrahydroPteGlu + 5-methyltetrahydroPteGlu.

*4 Biotin; the value is the sum of biotin + bisnorbiotin + biotin sulfoxide

*5 Vitamin C; the value is the sum of the reduced and oxidized ascorbic acid and its catabolite 2,3-diketogulonic acid.

Table 3. Tentative urinary excretion of water-soluble vitamins for maintaining health.

	Tentative values for maintaining health
Vitamin B_1 (nmol/d)	200–2000
Vitamin B_2 (nmol/d)	200–2000
Vitamin B_6*1 (μmol/d)	2–15
Vitamin B_{12} (pmol/d)	–*6
Niacin*2 (mmol/d)	50–300
Folate*3 (nmol/ d)	15–100
PaA (μmol/d)	10–40
Biotin*4 (nmol/d)	50–200
V.C*5 (μmol/d)	100–2000

*1 Vitamin B_6; the value is 4-PIC, a major catabolite of V.B_6.
*2 Niacin; the value is the sum of MNA, 2-Py, and 4-Py, which are major catabolites of nicotinamide.
*3 Folates; the value is the sum of pteroylmonoglutamic acid (PteGlu) + dihydroPteGlu + tetrahydroPteGlu + 5-formyltetrahydroPteGlu + 10 -formyltetrahydroPteGlu + 5,10-methylenetetrahydroPteGlu + 5-methyltetrahydroPteGlu.
*4 Biotin; the value is the sum of biotin + bisnorbiotin + biotin sulfoxide
*5 Vitamin C; the value is the sum of the reduced and oxidized ascorbic acid and its catabolite 2,3-diketogulonic acid.
*6 Urinary elimination is not a main pathway, so we did not try to calculate the urinary excretion of vitamin B_{12}.

Table 4. Several parameters (minimum, maximum, median, average ± SD, and tentative urinary excretion of water-soluble vitamins for maintaining health) in apparently healthy young Japanese females.

	Minimum	**Maximum**	**Median**	**Average ± SD**
Vitamin B_1 (nmol/d)	10	54933	426	710 ± 2208
Vitamin B_2 (nmol/d)	8	20969	348	650 ± 1300
Vitamin B_6*1 (μmol/d)	1	126.2	3.8	5.4 ± 8.5
Vitamin B_{12} (pmol/d)	8	1134	79	99 ± 81
Niacin*2 (mmol/d)	13	325	80	86 ± 42
Folate*3 (nmol/d)	2.6	718.8	16.6	21.3 ± 34
PaA (μmol/d)	0.3	97	15	16.7 ± 8.6
Biotin*4 (nmol/d)	3.5	884.9	68.6	80.9 ± 59.5
Vitamin C*5 (μmol/d)	6	3971	178	327 ± 442

*1 Vitamin B_6; the value is 4-PIC, a major catabolite of V.B_6.
*2 Niacin; the value is the sum of MNA, 2-Py, and 4-Py, which are major catabolites of nicotinamide.
*3 Folates; the value is the sum of pteroylmonoglutamic acid (PteGlu) + dihydroPteGlu + tetrahydroPteGlu + 5-formyltetrahydroPteGlu + 10 -formyltetrahydroPteGlu + 5,10-methylenetetrahydroPteGlu + 5-methyltetrahydroPteGlu.
*4 Biotin; the value is the sum of biotin + bisnorbiotin + biotin sulfoxide
*5 Vitamin C; the value is the sum of the reduced and oxidized ascorbic acid and its catabolite 2,3-diketogulonic acid.

Table 5. Trial evaluation of water-soluble vitamins in in 709 apparently healthy young Japanese females.

	Below the lower limit (%)	**Within range (%)**	**Over the upper limit (%)**
Vitamin B_1	22.4	73.6	4.1
Vitamin B_2	31.3	63.5	5.2
Vitamin B_6	6.2	90	3.8
Vitamin B_{12}	–	–	–
Niacin	14	85.6	0.4
Folate	40.9	58.1	1
Pantothenic acid	12.4	85.6	2
Biotin	26.2	70.2	3.6
Vitamin C	33	65.4	1.6

Table 6. Spearman correlation between the amounts of urinary excretion of water-soluble vitamins in 709 apparently healthy young Japanese females.

	Spearman r								
	Vitamin B_1	**Vitamin B_2**	**Vitamin B_6**	**Vitamin B_{12}**	**Niacin**	**Folate**	**PaA**	**Biotin**	**Vitamin C**
Vitamin B_1		0.203	0.400	-0.036	0.213	0.134	0.184	0.165	0.089
Vitamin B_2	< 0.0001		0.254	0.138	0.208	0.233	0.329	0.237	0.160
Vitamin B_6	< 0.0001	< 0.0001		0.231	0.452	0.180	0.362	0.218	0.218
Vitamin B_{12}	0.365	0.0003	< 0.0001		0.259	- 0.123	0.128	0.159	0.292
Niacin	< 0.0001	< 0.0001	< 0.0001	< 0.0001		0.168	0.332	0.285	0.219
Folate	0.008	< 0.0001	< 0.0001	0.002	< 0.0001		0.270	0.177	0.107
PaA	< 0.0001	< 0.0001	< 0.0001	0.001	< 0.0001	< 0.0001		0.281	0.160
Biotin	< 0.0001	< 0.0001	< 0.0001	< 0.0001	< 0.0001	< 0.0001	< 0.0001		0.134
Vitamin C	0.025	< 0.0001	< 0.0001	< 0.0001	< 0.0001	0.0001	< 0.0001	0.001	

The values above the diagonal line are "Spearman r", and those below the line are P values (two-tailed).
The number of data used here is 635 because subjects who obviously took vitamin supplements were omitted.

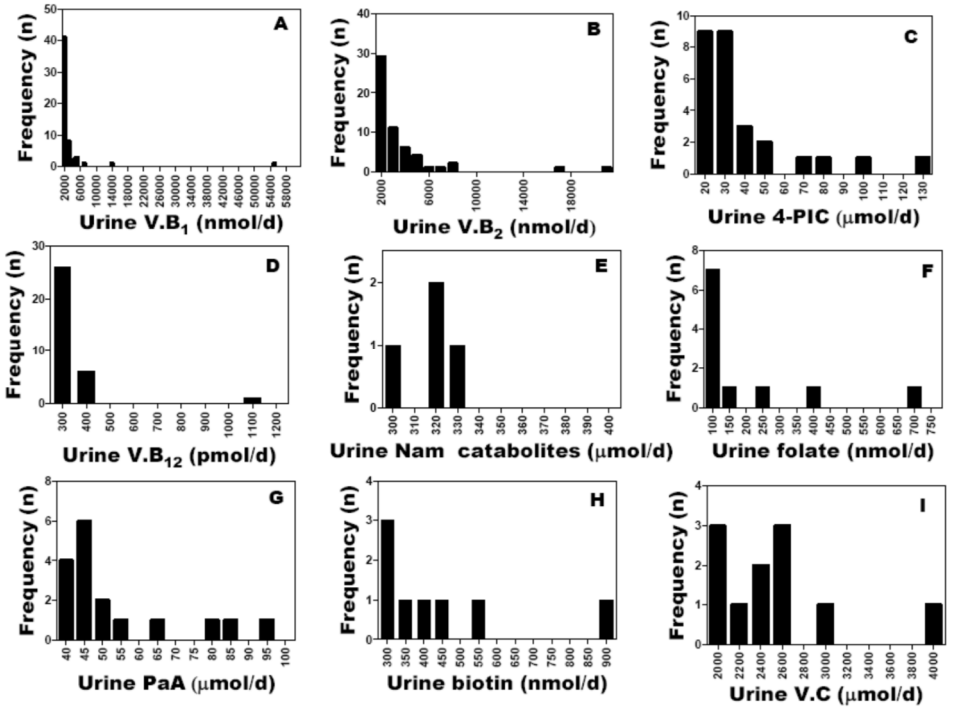

Figure 4. Fragmentary view of histograms (with a focus on extremely high excretion) of daily urinary excretion amounts of (A) vitamin B_1, (B) vitamin B_2, (C) vitamin B_6*1, (D) vitamin B_{12}, (E) niacin*2, (F) folates*3, (G) pantothenic acid, (H) biotin*4, and (I) vitamin C*5 in apparently healthy young Japanese females.

*1 Vitamin B_6; the value is 4-PIC, a major catabolite of V.B_6.

*2 Niacin; the value is the sum of MNA, 2-Py, and 4-Py, which are major catabolites of nicotinamide.

*3 Folates; the value is the sum of pteroylmonoglutamic acid (PteGlu) + dihydroPteGlu + tetrahydroPteGlu + 5-formyltetrahydroPteGlu + 10-formyltetrahydroPteGlu + 5,10-methylenetetrahydroPteGlu + 5-methyltetrahydroPteGlu.

*4 Biotin; the value is the sum of biotin + bisnorbiotin + biotin sulfoxide

*5 Vitamin C; the value is the sum of the reduced and oxidized ascorbic acid and its catabolite 2,3-diketogulonic acid.

pantothenic acid, biotin, and vitamin C in the study cohort. Figure 4 shows the fragmentary view of histograms (with a focus on extremely high excretion) of daily urinary excretion amounts of vitamin B_1, vitamin B_2, vitamin B_6, vitamin B_{12}, niacin, folic acid, pantothenic acid, biotin, and vitamin C. Table 5 shows a trial evaluation of water-soluble vitamins in the study cohort. Vitamin B_2, folate, and vitamin C occupied >30% beyond the lower limit. The percentage over the upper limit was for each water-soluble vitamin was very low. Table 6 shows the correlation between the amounts of urinary excretion of water-soluble vitamins. A relatively close relationship between vitamin B_1 and vitamin B_6, as well as between vitamin B_6 and niacin, was observed.

Of 709 subjects, the urinary excretion of eight water-soluble vitamins was within the tentative values for maintaining health (Table 3) in 181 subjects (25.5%). The number of subjects in whom only one vitamin was below the lower limit of detection was 185 (26.1%), for two vitamins was 157 (22.1%), for three vitamins was 97 (13.7%), for four vitamins was 148 (20.9%), for five vitamins was (22) 3.1%, for six vitamins was 12 (1.7%), for 7 vitamins was 2 (0.28%), and for 8 vitamins was 1 (0.14%).

The urinary excretion of water-soluble vitamins reflects that for the most recent week [25–27]. Hence, the lower urinary excretion does not immediately mean vitamin insufficiency. However, if a subject maintains a diet for a long time, he/she will be in an appropriate state to reveal a deficiency syndrome.

K.S. and T.F. designed the study. K.S. drafted the manuscript and T.F. reviewed the manuscript. The authors express their sincere appreciation to the many subjects who gave urine samples and to our students for measuring the levels of water-soluble vitamins. All authors approved the final manuscript.

4. Concluding Remarks

We calculated a tentative urinary excretion values for water-soluble vitamins for maintaining health based on data obtained from intervention studies. We evaluated the nutritional status of water-soluble vitamin in 709 Japanese females.

Acknowledgements

This investigation was part of the project “Studies on the Dietary Reference Intakes for Japanese” (principal investigator, Sinkan Tokudome), which was supported by a Research Grant for Comprehensive Research on Cardiovascular and Life-Style Related Diseases from the Ministry of Health, Labour and Welfare of Japan. This work is a publication of the Department of Nutrition, School of Human Cultures, University of Shiga Prefecture. None of the authors have any conflicts of interest.

References

1. K. Shibata, T. Fukuwatari, A. Enomoto, E. Sugimoto, J Nutr Sci Vitaminol 47 (2001) 263-266.
2. K. Shibata, T. Fukuwatari, E. Sugimoto, Biosci Biotechnol Biochem 65 (2001) 1339-1346.
3. T. Fukuwatari, Y. Morikawa, F. Hayakawa, E. Sugimoto, K. Shibata K, Biosci Biotechnol Biochem 65 (2001) 2154-2161..
4. T. Fukuwatari, Y. Suzuki, E. Sugimoto, K. Shibata, Biosci Biotechnol Biochem 66 (2002) 705-710.
5. T. Fukuwatari, E. Sugimoto, K. Shibata, Biosci Biotechnol Biochem 66 (2002) 1435-1441.
6. T. Fukuwatari, Y. Suzuki, E. Sugimoto, K. Shibata, Biosci Biotechnol Biochem 66 (2002) 2687-2691.
7. T. Fukuwatari, M. Ohta, E. Sugimoto, R. Sasaki, K. Shibata, Biochim Biophys Acta 1672 (2004) 67-75.
8. T. Fukuwatari, S. Ohsaki, S. Fukuoka, R. Sasaki, K. Shibata, Toxicol Sci 81 (2004) 302-308.
9. A. Okuno, T. Fukuwatari, K. Shibata, Biosci Biotechnol Biochem 72 (2008) 1667-1672.
10. K. Shibata, C. Takahashi, T. Fukuwatari, R. Sasaki, J Nutr Sci Vitaminol 51 (2005) 385-391
11. H. Sawamura, T. Fukuwatari, K. Shibata, Biosci Biotechnol Biochem 71 (2007) 2977-2984.
12. A. Moriya, T. Fukuwatari, M. Sano, K. Shibata, Br J Nutr 107 (2012) 52-60.
13. K. Shibata, E. Imai, M. Sano, T. Fukuwatari, Biosci Biotechnol Biochem 76 (2012) 186-188.
14. Imai E, Sano M, Fukuwatari T, Shibata K, J Nutr Sci Vitaminol 58 (2012) 54-58.
15. K. Shibata, C. Sugita, M. Sano, T. Fukuwatari, J Nutr Sci (2013) e12. doi: 10.1017/jns.2013.3.
16. H. Okamoto, A. Ishikawa, H. Yoshiike, N. Kodama, T. Nishimuta, T. Fukuwatari, K. Shibata, Am J Clin Nutr 72 (2003) 406-410.
17. K. Shibata, T. Fukuwatari, M. Ohta, H. Okamoto, T. Watanabe, T. Fukui, M. Nishimuta, M. Totani, M. Kimura, N. Ohishi, M. Nakashima, F. Watanabe, E. Miyamoto, S. Shigeoka, T. Takeda, M. Murakami, H. Ihara, N. Hashizume, J Nut Sci Vitaminol 51 (2005) 319-328.
18. T. Fukuwatari, M. Ohta, N. Kimura, R. Sasaki, K. Shibata, J Nutr Sci Vitaminol 50 (2004) 385-391.
19. T. Fukuwatari, K. Shibata, Int J Vitam Nutr Res 77 (2007) 255-262.
20. T. Fukuwatari, K. Shibata, J Nutr Sci Vitaminol 54 (2008) 223-229.
21. T. Fukuwatari, K. Shibata, J Nutr Sci Vitaminol 55 (2009) 279-281.
22. K. Shibata, T. Fukuwatari, T. Watanabe, M. Nishimuta, J Nutr Sci Vitaminol 55 (2009) 459-470.
23. T. Fukuwatari, E. Sugimoto, T. Tsuji, J. Hirose, T. Fukui, K. Shibata, Nutr Res 29 (2009) 839-845.
24. E. Imai, T. Tsuji, M. Sano, T. Fukuwatari, K. Shibata, Asian Pac J Clin Nutr 20 (2011) 507-513.
25. T. Tsuji, T. Fukuwatari, S. Sasaki, K. Shibata, Eur J Clin Nutr 64 (2010) 800-807.
26. T. Tsuji, T. Fukuwatari, S. Sasaki, K. Shibata, Nutr Res 30 (2010) 171-178.
27. Tsuji T, Fukuwatari T, Sasaki S, Shibata K. (2011) Twenty-four-hour urinary water-soluble vitamin levels correlate with their intakes in free-living Japanese school children. Public Health Nutr 14:327-33.
28. Dietary reference Intakes for Japanese, 2010, Tokyo, Japan. Ministry of Health, Labor, and Welfare of Japan. Ministry of Health, Labor, and Welfare of Japan, 2009.
29. M.E. Pullman, S.P. Colowick, J Biol Chem 206 (1954) 121-

127.
30. K. Shibata, T. Kawada, K. Iwai, J Chromatogr 424 (1988) 23 -28.
31. T. Fukuwatari, C. Suzuura, R. Sasaki, K. Shibata, Shokuhin Eiseigaku Zasshi (Japanese) 45 (2004) 231-238.
32. H. Ohkawa, N. Ohishi, K. Yagi, J Biol Chem 258 (1983) 5623-5628.
33. J.F. Gregory 3rd, J.R. Kirk, Am J Clin Nutr 32 (1979) 879-883.
34. F. Watanabe, K. Abe, H. Katsura, S. Takenaka, Z.H. Mazumder, R. Yamaji, S. Ebara, T. Fujita, S. Tanimori, M. Kirihata, Y. Nakano, J Agric Food Chem 46 (1998) 5177-5180.
35. K. Shibata K, Vitamins (Japan) 61 (1987) 599-604.
36. K. Takahashi, T. Fukuwatari, K. Shibata, J Chromatogr 877 (2009) 2168-2172.
37. H. Taisun, T. Tamura, Exp Biol Med 230 (2005) 444-454.
38. K. Kishida, Y. Nishimoto, S. Kojo, Anal Chem 64 (1992) 1505-1507.
39. G.A. Goldsmith, H.P. Sarett, U.D. Register, J. Gibbens, J Clin Invest 31 (1952) 533-542.
40. T. Matsuo, Z. Suzuoki, J Biochem 65 (1969) 953-960.
41. W.H Amo Jr, R.A. Neal, J Biol Chem 245 (1970) 5643-5648.
42. N.K. Shinton, Br Med J 1 (1972) 556-559.
43. M.K. Horwitt, A. E. Harper, L.M. Hendeson, Am J Clin Nutr 34 (1981) 423-427.
44. J. Kirchner, J.G. Watson, S. Chaykin, J Biol Chem 241 (1966) 953-960.
45. K. Shibata, Agric Biol Chem 52 (1988) 2973-2976.
46. K. Shibata, J Nutr 119 (1989) 892-895.
47. K. Shibata, H. Matsuo Agric Biol Chem 54 (1990) 223-224.
48. T. Fukuwatari, K. Kurata, K. Shibata, Shokuhin Eiseigaku Zasshi (Japanese)50 (2009) 80-84.
49. K. Shibata, H. Kakehi, H. Matsuo, J Nutr Sci Vitaminol 36 (1990) 87-98.
50. K. Shibata, M. Onodera, S. Shimada, K. Yasuda, Vitamins (Japanese) 66 (1992) 309-314..
51. K. Shibata, H. Matsuo, Am J Clin Nutr 50 (1989) 114-119.
52. M.L. Willoughby, F.G. Jewell, Br Med J 9 (1968) 356-360.
53. J. McPartlin, A. Halligan, J.M. Scott, M. Darling, Lancet 341 (1993) 148-149.
54. D.W. Home, D. Patterson, D. Clin Chem 34 (1988) 2357-2359.
55. D.M. Mock, G.L. Lankford, J. Cazin Jr, J Nutr 123 (1993) 1844-1851.

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

ORIGINAL ARTICLE | DOI: 10.5584/jiomics.v3i1.129

Starvation of Jurkat T cells causes metabolic switch from glycolysis to lipolysis as revealed by comprehensive GC-qMS

Janina M. Tomm[1], Wolfgang Otto[1], Ulrich Servos[2], Bernd Bertram[2], Martin von Bergen[1,3,4], Sven Baumann[3]

[1]Department of Proteomics, UFZ – Helmholtz Centre for Environmental Research, Permoserstraße 15, 04318 Leipzig, Germany; [2]Shimadzu Germany, Albert-Hahn-Straße 6-10, 47269 Duisburg, Germany, [3]Department of Metabolomics, UFZ – Helmholtz Centre for Environmental Research, Permoserstraße 15, 04318 Leipzig, Germany; [4]Department of Biotechnology, Chemistry and Environmental Engineering, Aalborg University, Sohngaardsholmsvej 49, 9000 Aalborg, Denmark

Received: 11 February 2013 **Accepted:** 25 April 2013 **Available Online**: 13 May 2013

ABSTRACT

T cells play a central role in the cellular part of the immune system and their metabolomic activity is strictly bound to the status of activation. Little is known about the metabolomic resilience of T helper cells in the resting status. Therefore we analyzed the metabolomic profile of non-activated Jurkat T cells under normal and starvation conditions by a two-dimensional GC-MS approach. We detected 52 organic and amino acids in the biological replicates covering the majority of central metabolic pathways. Under starvation conditions 21 analytes representing major metabolic pathways showed a significant down-regulation. For palmitoleic acid a significant up-regulation was detected. The annotation of differentially abundant metabolites to the pathways of TCA-cycle, amino acid metabolism and fatty acid biosynthesis indicating a metabolic switch from glycolytic to lipolytic energy generation upon starvation.

Keywords: Comprehensive GC-qMS; Metabolite profiling; Starvation; Jurkat T cells.

1. Introduction

In the postgenome era the focus shifts from the static blueprint to the factors that actually dominate the phenotype of a cell, proteins and metabolites. Both are linked to the genotype but are also regulated independently and therefore provide targets for the analysis of factors that can be both indicative and causative to a great variety of cellular processes [1]. Due to their central role for the cellular life most processes eventually influence the metabolome. In addition, the perspective on the metabolome as a mere consequence of catabolic and anabolic reactions is changing with the discovery of the different levels of crosstalk between metabolites and signalling pathways.

Furthermore, functions exerted by metabolites can be highly tissue specific. For example, glutamine is the most abundant amino acid in plasma and in skeletal muscle [2]. In contrast, glutamine serves in liver as a precursor for glucose and urea synthesis, whereas it is used in the brain in synthesis of neurotransmitters. Hence glutamine is reported to play a role in maintenance of skeletal muscle, immune systems function and glucose and glycogen metabolism [3]. As a consequence, the metabolome has to be analyzed in a cell-type specific manner. This is particularly true for the immune system in which the differentiation of cells leads to a great variety of cellular subtypes. In the present study Jurkat cells are used as a widely accepted model for the study of activation of T cells and the resulting effect on the metabolome [4].

T cells control the adaptive immune response to pathogens either by destruction of the antigen presenting cell directly (cytotoxic T cells) or by secretion of cytokines (T helper cells) which activate effector cells. T cells are activated by interaction of their T cell receptor (TCR) with an antigen-

***Corresponding author:** Dr. Sven Baumann. Helmholtz Centre for Environmental Research – UFZ, Department of Metabolomics , Permoserstraße 15, 04318 Leipzig. Phone: +49-341/2351099; E-Mail address: sven.baumann@ufz.de

bound major histocompatibility complex containing an antigen-presenting cell [5, 6]. In addition to the TCR induction, T cells require additional activation signals from one or more costimulatory receptors, such as CD3 and CD28, to become fully active [7, 8].

Peripheral T cells which expand upon signal responses to foreign antigens spend the majority of their life span in a resting state. Therefore their metabolic resilience to suboptimal conditions is crucial for their survival. Besides changes in protein expression, changes in concentration of central metabolites during activation and differentiation of T cells are of particular interest. Since metabolites are the end products of cellular regulatory processes, their levels can be regarded as the ultimate response of biological systems to genetic or environmental changes. In order to tackle the question of metabolomic resilience of resting T cells we established a metabolomic analysis of Jurkat T cell grown under normal and starvation conditions as a model system to elucidate the cellular response towards suboptimal conditions.

For the analysis of the central carbon and nitrogen metabolites that are part of the glycolysis, pyruvate metabolism as well as the tricarboxylic acid cycle (TCA-cycle) and the amino acid synthesis methods are required that cope with the high polarity of the metabolites. GC-MS is a suitable technique for this purpose. Several GC-MS based metabolomics methods have been reported [9-13]. Most of these methods are based on one or more derivatization steps by oximation and/or silylation analysis prior to conversion of the polar functional groups. These methods are suitable for the analysis of cell extracts, body fluids and tissues and allow the measurement of a broad range of small, medium-polar to polar metabolites.

However, profiling analyses and quantification of target substances from complex biological samples is often hampered by insufficient chromatographic separation and/or matrix effects resulting from difficult sample matrices. Comprehensive two-dimensional gas chromatography (GCxGC) offers several advantages over one-dimensional gas chromatography, i.e. higher peak capacity, a broader dynamic range and lower detection limits. Metabolite profiling of differentially grown Jurkat T cells was therefore carried out on a comprehensive GCxGC quadrupole MS to characterize molecular changes in cellular systems by quantitative signatures of primary metabolites.

2. Material and Methods

2.1 *Materials and reagents*

Methanol and Chloroform were purchased from Merck (Darmstadt, Germany). Water was deionized and filtered through a 0.22 µm filter using a Millipore water generation system (Millipore, Schwalbach, Germany). Methoxyamine hydrochloride (MOX) and *N*-Methyl-*N*-(trimethylsilyl) trifluoroacetamide (MSTFA) were purchased from Sigma-Aldrich (Seelze, Germany).

2.2 *Cell culture*

Jurkat T cells clone E6-1 (TIB-152, ATCC, Manassas, VA, USA) were routinely cultured in RPMI-1640 medium (Biochrom AG., Berlin, Germany) containing 10% fetal bovine serum (Biochrom AG, Berlin, Germany), 1% L- Glutamine (Biochrom AG, Berlin, Germany) and 1% penicillin (100 U/mL)/ streptomycin (100 mg/mL) (PAA, Pasching, Austria) at an atmosphere of 5% CO_2, 95% humidity at 37°C in a CO_2 incubator (MCO-18AIC, Sanyo Electric Co Ltd, Gunma-ken, Japan). Jurkat cells were maintained between 1 x 10^5 and 1 x 10^6 cells/mL of medium in flasks with a surface area of 25 cm^2 (Greiner Bio-One, Solingen, Germany). For the experiment control cells were cultured in 75 cm^2 flasks as described earlier [14]. Starving cells were incubated in RPMI-1640 medium containing 1% L-Glutamine, 1% penicillin/streptomycin for three to four days in 75 cm^2 flasks under the conditions stated above.

Before and after starvation, cell numbers were recorded by cell counting after Trypan blue staining in a Neugebauer-chamber. After cultivation cells were spun down at 300*g* for 5 min, the medium was removed and the cell pellets were immediately shock-frozen in liquid nitrogen and stored at –80°C until further processing.

2.3 *Metabolite extraction*

In order to separate hydrophilic target analytes from lipophilic metabolites, phases of different polarity were obtained by performing a liquid-liquid extraction with ice cold methanol, chloroform and water. For comprehensive GCxGCqMS analysis frozen cell pellets were extracted using 0.5 mL methanol, 1 mL chloroform and 0.375 mL water. Cell disruption was performed for 5 min on ice using an ultrasonic homogenization device. The mixture was then centrifuged at 4000 *g* for 5 min at 4°C. 500 µL of the upper hydrophilic methanol/water phase was transferred into 2 mL reaction glass vials. The remaining lipophilic phase as well as protein precipitate was discarded. All samples were dried under a stream of nitrogen at 40 °C for a maximum of 2 h. First derivatization step of hydrophilic compounds was a methoximation using 200 µL MOX for 2 h at 80°C. After first derivatization step the fractions were dried under a stream of nitrogen at 40 °C and silylated using 100 µL MSTFA for 20 min at 90°C.

2.4 *GCxGC-qMS analysis and data processing*

Analyses were done in biological and technical duplicates by comprehensive gas chromatography in combination with a fast quadrupole mass spectrometer, GCxGCqMS. The dimension of the first column (Rtx-1; Restek) were, 30 m length, 0.25 mm i.D. with 0.25 µm film, of the second column (BPX50 (SGE)) 1 m length, 0.15 mm i.D. and 0.15 µm film. The oven was heated from 60°C to 290°C with a ramp of 5°C/mi. The modulation time was 6 sec and the hot pulse time 0.375 sec, respectively. The detection was performed in

Analyte	Fold change Starvation vs Normal	p-Value
4-Hydroxyproline	0.234	0.004
Glutamic acid	0.267	0.006
Fumaric acid	0.169	0.008
Glycolic acid	0.594	0.008
Methylsuccinic acid	0.380	0.009
Lactic acid	0.247	0.010
Succinic acid	0.235	0.010
Malic acid	0.267	0.011
Palmitoleic acid	2.738	0.012
Phenylalanine	0.391	0.012
Serine	0.518	0.013
Proline	0.188	0.014
Tyrosine	0.500	0.016
3-Hydroxybutyric acid	0.118	0.019
Urea	0.416	0.020
L-Glycine	0.238	0.026
a-Aminobutyric acid	0.343	0.031
2-Hydroxyglutaric acid	0.498	0.034
Isoluecine	0.440	0.036
5-Oxoproline	0.382	0.039
Aspartic acid	0.548	0.040
L-Alanine	0.242	0.048
Threonin	0.372	0.057
Asaparagin	0.064	0.067
Tryptophan	0.060	0.069
Glycerol	0.710	0.071
Methionine	0.314	0.073
Benzoic acid	0.612	0.113
Valine	0.446	0.121
Citric acid	0.728	0.133
Phenol	0.871	0.163
Leucine	0.486	0.168
Glutaric acid	0.330	0.199
3-Hydroxy-3-methylglutaric acid	0.249	0.204
Thymine	0.071	0.269
Oxalic acid	0.450	0.333
Palmitic acid	1.115	0.336
Caproic acid	0.733	0.366
Stearic acid	1.056	0.392
Myristic acid	0.690	0.397
Azelaic acid	0.871	0.427
Sebacic acid	0.879	0.458

Table 1. Summary of detected organic acids in Jurkat T cells after GCxGC-qMS analyses sorted according to ascending p-values. Italicized analyte names represent significant regaulated organic and amino acids.

Not detected in starvation cells
2-Hydroxybutyric acid
Sarcosine
2-Hydroxyisovaleric acid
3-Hydroxyisovaleric acid
Ethylhydracrylic acid
Phenylacetic acid
Maleic acid
3-Methylglutaconic acid
L-Glutamine
Hippuric acid
Indolelactic acid

full scan mode from 35 to 280 amu with 20.000 amu/sec. The chromatograms were recorded and processed by GCMSsolution software (Shimadzu). The identification of metabolites was performed by Excel based macro, taking spectra similarity, linear retention index and the relation between modulated peaks into account. Fold changes were calculated between the accumulated peak areas of modulated signals of starvation cultivation versus the normal cultivated Jurakt T cells. An unpaired, heteroscedastic, two-tailed Student's t-test was conducted on the peak areas for the comparison of starved and normal cultered cells. Metabolites with a p-value less than 0.05 and an average linear fold change higher than 1.5 were considered as significantly regulated.

3. Results and Discussion

Hydrophilic extracts gave reproducible characteristic chromatograms. In total, 42 metabolites from the classes of organic and amino acids could be detected and verified in the biological duplicates in normal and starvation cultivated cells. Identification of metabolites with a substance specific database was supported by use of linear retention time indices and the time-dependent relation between modulated GC -peaks. As a result of high retention time stability of the GCxGCqMS system, precise determination of linear retention time indices was possible. A typical chromatogram of an extract from Jurkat T cells grown under normal conditions is shown in Figure 1. In addition, eleven metabolites comprising among others sarcosine, phenylacetic and hippuric acid could be observed only in T cells grown under normal conditions. Obviously due to decelerated metabolism of T cells under starving conditions and generally lower signal intensities in the total ion chromatogram, these analytes were not detectable in samples derived from starving T cells (Table 1). Above all, also glutamine could only be observed in normal cultivated cells, even it was also supplied during cultivation,

which suggests an almost complete consumption cells under starvation.

Due to sample complexity several compounds were coeluting within the first separation dimension. Therefore, comprehensive GC-qMS was applied to increase the separation power of coeluting substances in the second dimension (Figure 2). Besides this, GCxGC separation leads to a better matrix separation and provides a basis for better spectra quality and library based identification of target analytes.

In metabolite extracts of control and starving Jurkat T cells clear differences in analyte intensities became evident. Due to decelerated metabolism of T cells under starving conditions, reduced intensities for several metabolites could be observed. Wilcoxon signed-rank test was employed to detect metabolites that are significantly different between control and starving Jurkat T cells. Using a threshold of $p < 0.05$ and a fold change >1.5, in total 11 metabolites were detected that showed statistical relevant differences (Figure 3, Table 1).

The most affected amino and organic acids were assigned showing varying dynamics during different growth conditions. Cellular concentrations of all identified amino acids decreased during growth under starving conditions.

For serine, phenylalanine, alanine, glycine, proline, isoleucine and tyrosine a significant decrease of cellular concentrations (fold changes 0.2 - 0.5) from normal to starvation conditions could be observed (Table 1). The effect of increased amino acid consumption during starvation could also be shown in mouse macrophages -like RAW cells [15]. Certain amino acids are now known to play important nutrient-sensing roles involving the (mTOR)-mediated signaling pathway [16]. mTOR is a component of a signaling pathway that couples insulin receptor stimulation and nutrient availability with protein synthesis via activation and phosphorylation of the ribosomal protein S6. Energy levels and nutrients regulate mTOR effectors such as S6K1 (S6 kinase 1) through the interaction with the mTOR complex. In cells growing in nutrient-rich conditions, the mTOR kinase activity is high, whereas under nutrient-poor conditions the mTOR kinase activity is decreased. It is not known how amino acids activate the mTOR complex, but it is probable that stimulation of a kinase or inhibition of a phosphatase that act upon mTOR as a substrate is involved [17-19].

In addition, various organic acids revealed the same effect of decreased signal intensities compared to Jurkat T cells grown under normal conditions (Table 1). A 0.2-fold change of signal intensities could be observed for fumaric and succinic acid after starvation cultivation ($p<0.05$). The intracellular decrease of TCA-cycle metabolites could be explained with consumption of these metabolites for energy generation of the cells during starvation period.

Significant differences could also be detected in pyruvate metabolism, resulting in lowered amounts of lactate as well as malate (fold changes 0.2 and 0.3, respectively).

For stearic acid, palmitic acid and especially its oxidation product palmitoleic acid higher signal intensities in the starvation culture could be observed (Fold changes 1.1, 1.1 and

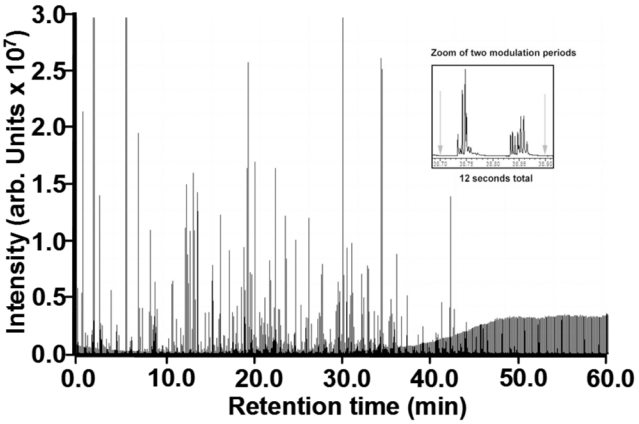

Figure 1. Comprehensive GC-qMS chromatogram of a hydrophilic extract from Jurkat T cells grown under normal conditions. Due to modulation by continuous freezing and hot pulse evaporation of analytes eluting from 1st dimension column peaks were cut into multiple fractions. Thermally desorbed analytes will be further separated on 2nd dimension column and reveal chromatographic peaks with a temporal distance according to the chosen modulation time (see insert; modulation time of 6s).

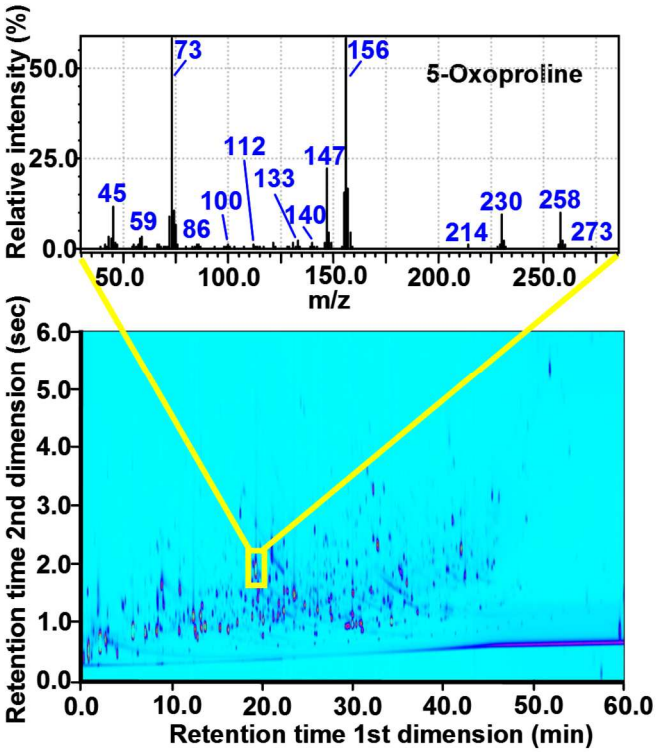

Figure 2. Comprehensive two-dimensional GC-qMS plot of a hydrophilic extract from Jurkat T cells grown under normal conditions. Additionally, the corresponding mass spectrum for 5-Oxoproline is shown as an example. Separation of analytes into two chromatographic dimensions by preferably orthogonal separation characteristics. Resulting 'blobs' in horizontal direction illustrating non-separated substances coeluting under "classical" chromatographic conditions. Referring mass spectra can be obtained from the modulated single peaks and used for compound identification.

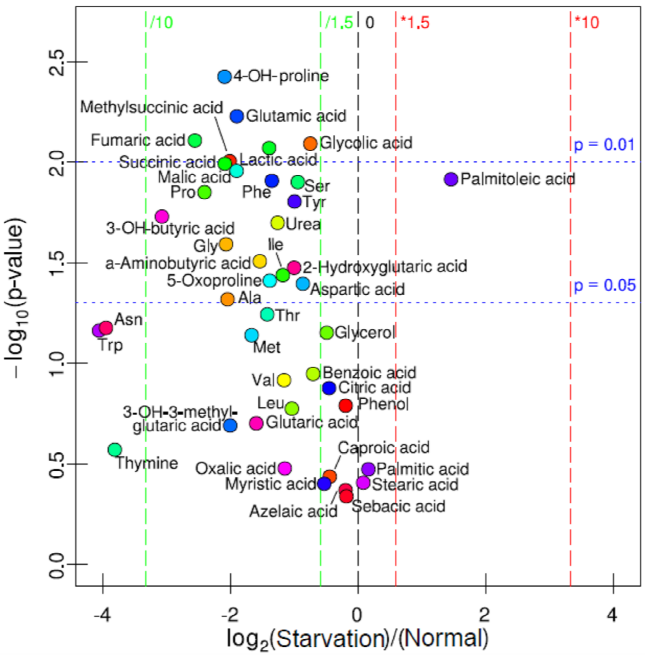

Figure 3. Volcano plot of for comparison of Jurkat T cell extracts after normal and starvation cultivation. As a reaction to cultivation of cells under starvation conditions and a resulting decelerated metabolism reduced signal intensities for several amino and organic acids could be detected. For three fatty acids an increase of signal intensities due to lipolytic shift could be observed.

2.7, respectively). This could be explained by a shift to lipolysis as one possible alternative for energy generation as a result of the energy deficit of starving Jurkat T cells. Thereby, triglycerides were hydrolyzed to provide free fatty acids as energy source [20, 21].

4. Concluding Remarks

The application of GCxGC-qMS allowed the reproducible detection of metabolites from the central carbon and amino acid metabolism of T cells. Starvation conditions induced significant changes in resting T cells metabolism, predominantly in the metabolic pathways of amino acid synthesis, the TCA cycle and fatty acid biosynthesis. These results will help to complement the puzzle of the metabotype of the variant status of immune cells. Since the involved pathways are swiftly affected the next step will be to elucidate if consumption of amino and organic acids as well as the lipolytic switch during starvation is a reversible process or not and indicating for the final fate of these cells.

References

1. I. Kouskoumvekaki, G. Panagiotou, J Biomed Biotechnol 29 (2011) 94.
2. J.M. Lacey, D.W. Wilmore, Nutr Rev 48 (1990) 297.
3. J. Antonio, C. Street, Can J Appl Physiol 24 (1999) 1.
4. R.T. Abraham, A. Weiss, Nat Rev Immunol 4 (2004) 301.
5. A.E. Nel, J Allergy Clin Immunol 109 (2002) 758.
6. E.J. Sundberg, L. Deng, R.A. Mariuzza, Semin Immunol 19 (2007) 262.
7. A.E. Nel, J Allergy Clin Immunol 109 (2002) 758.
8. I.A. Wilson, K.C. Garcia, Curr Opin Struct Biol 7 (1997) 839.
9. O. Fiehn, J. Kopka, R.N. Trethewey, L. Willmitzer, Anal Chem 72 (2000) 3573.
10. P. Jonsson, J. Gullberg, A. Nordstrom, M. Kusano, M. Kowalczyk, M. Sjostrom, T. Moritz, Anal Chem 76 (2004) 1738.
11. M.M. Koek, B. Muilwijk, M.J. van der Werf, T. Hankemeier, Anal Chem 78 (2006) 1272.
12. U. Roessner, C. Wagner, J. Kopka, R.N. Trethewey, L. Willmitzer, Plant J 23 (2000) 131.
13. R.A. Shellie, W. Welthagen, J. Zrostlikova, J. Spranger, M. Ristow, O. Fiehn, R. Zimmermann, J Chromatogr A 1086 (2005) 83.
14. M. Rockstroh, S.A. Müller, C. Jende, A. Kerzhner, M. von Bergen, J.M. Tomm, JIOMICS 1 (2011) 135.
15. H. Sakagami, K. Kishino, O. Amano, Y. Kanda, S. Kunii, Y. Yokote, H. Oizumi, T. Oizumi, Anticancer Res 29 (2009) 343.
16. M.L. McDaniel, C.A. Marshall, K.L. Pappan, G. Kwon, Diabetes 51 (2002) 2877.
17. I. Briaud, M.K. Lingohr, L.M. Dickson, C.E. Wrede, C.J. Rhodes, Diabetes 52 (2003) 974.
18. G. Kwon, C.A. Marshall, K.L. Pappan, M.S. Remedi, M.L. McDaniel, Diabetes 53 Suppl 3 (2004) S225.
19. M.L. McDaniel, C.A. Marshall, K.L. Pappan, G. Kwon, Diabetes 51 (2002) 2877.
20. J.S. Bruce, A.M. Salter, Biochem J 316 (Pt 3) (1996) 847.
21. M. Dieuaide, I. Couee, A. Pradet, P. Raymond, Biochem J 296 (Pt 1) (1993) 199.