Pure Appl. Chem., Vol. 80, No. 12, pp. 2609–2626, 2008.
doi:10.1351/pac200880122609

Integrated application of transcriptomics, proteomics, and metallomics in environmental studies*

Macarena González-Fernández[1], Tamara García-Barrera[1], Juan Jurado[2], María J. Prieto-Álamo[2], Carmen Pueyo[2], Juan López-Barea[2], and José Luis Gómez-Ariza[1,‡]

[1]Department of Chemistry and Materials Science, Faculty of Experimental Science, University of Huelva, El Carmen Campus, 21007-Huelva, Spain; [2]Department of Biochemistry and Molecular Biology, Severo Ochoa Building, University of Córdoba, Rabanales Campus, 14071 Córdoba, Spain

Abstract: Here we report a preliminary working scheme for the integrative application of transcriptomic, proteomic, and metallomic methodologies in environmental monitoring, by using as sentinel the wildlife species *Mus spretus* and as reference the gene/protein sequence databases from the key model species *Mus musculus*. We have demonstrated that the absolute transcript expression signatures quantified by reverse transcription (RT) and real-time polymerase chain reaction (PCR) of selected key genes (e.g., those coding for biotransformation enzymes) in *M. spretus* is a useful and reliable novel biomonitoring end-point. The suitability of commercial *M. musculus* oligonucleotide arrays for genome-wide transcriptional profiling in *M. spretus* has been also shown. Transcriptomic studies indicate considerable gene sequence similarities between both mouse species. Based on these similarities, we have demonstrated the applicability in free-living *M. spretus* of high-throughput proteomic methods, based on matrix-assisted laser desorption/ionization with time-of-flight mass spectrometry (MALDI-TOFMS) analysis of tryptic 2D electrophoresis (2-DE) spot digest and peptide matching with *M. musculus* database. A metallomic approach based on size exclusion chromatography inductively coupled plasma-mass spectrometry (SEC-ICP-MS) was applied to trace metal-biomolecule profiles. A preliminary integration of these three -omics has been addressed to *M. musculus*/*M. spretus* couple, two rodent species that separated 3 million years ago. The integrated application of transcriptomic and proteomic data and the bidirectional use of metallomics and proteomics for selective isolation of metal-biomolecules are covered in the working scheme MEPROTRANS-triple-OMIC reported in this study.

Keywords: transcriptomics; proteomics; metallomics; *Mus musculus*; *Mus spretus*; environment; Doñana.

INTRODUCTION

Monitoring environmental issues requires integrating approaches to sum up the multiple variables and factors that contribute to ecosystem behavior. Living organisms are usually the most unambiguous in-

*Paper based on a presentation at the International Symposium on Metallomics 2007 (ISM 2007), 28 November–1 December 2007, Nagoya, Japan. Other presentations are published in this issue, pp. 2565–2750.

‡Corresponding author

dicators of environmental concern (bioindicators), since they can reflect the effect of contaminants on cellular metabolism and global homeostasis. Generally, environmental stress situations can be assessed at molecular level using different parameters (biomarkers). They include hundreds of cytochromes P450 (CYPs), which add polar groups into an organic substrate, e.g., introducing one O atom that promotes its hydroxylation (RH → ROH) plus other reactions, or dozens of glutathione-*S*-transferases (GSTs), that conjugate electrophiles to reduced glutathione (GSH). In addition, heat-shock proteins refold denatured proteins and metallothioneins (MTs) protect organisms from toxic metals [1]. Besides, primary antioxidant enzymes, such as superoxide dismutases (SODs), glutathione peroxidases (GPXs), and peroxiredoxins (PRXs), secondary antioxidant enzymes, including glucose-6-phosphate dehydrogenase and glutathione reductase (GSSGrase), and antioxidant proteins, such as thioredoxins (TRXs) and glutaredoxins (GRXs), are also induced by exposure to many pollutants. Finally, low Mr antioxidants, including GSH and vitamins C or E, are also greatly enhanced in exposed organisms [2,3]. Other biomarkers arise from the damage of key biomolecules by toxic chemicals when they exceed pollutant-elicited defenses, this includes malondialdehyde and 4-hydroxy-2-nonenal, product of lipid peroxidation, and oxidized bases, abasic sites, and chain breaks, generated following nucleic acid damage. In proteins, disulfide bridges are formed and Met or His are oxidized to methionine sulfoxide (MetSO) and 2-oxohistidine, respectively. GSH is oxidized to GSSG, which forms mixed disulfides with protein thiols. These molecular damages trigger repair enzymes, such as phospholipases, GSHPx and PRXs in lipids injury, glycosidases for nucleic acids, MetSO reductase and TRXs for proteins, and GRXs and GSSGrase for GGSG and mixed disulfides [2,3].

These numerous conventional biomarkers are a good tool in pollution assessment, but require a deep knowledge of their toxicity mechanisms and assess a limited number of well-known proteins, excluding others also altered but whose relationship with pollution is unknown. Therefore, massive identification of proteins by proteomic approaches can provide a more general appraisal of proteins altered under pollutant exposure [1]. Environmental proteomics provides a more comprehensive assessment of the toxic and defensive mechanisms triggered by pollutants without requiring any previous knowledge. However, this approach has problems in environmental studies that generally use nonmodel sentinel organisms, whose genetic sequences are not included in databases. This fact makes difficult the identification of proteins with high throughput proteomics methods, generally based on MALDI-TOF (matrix-assisted laser desorption/ionization time-of-flight) analysis of tryptic 2D electrophoresis (2-DE) spot digests, that requires highly expensive and cumbersome *de novo* sequencing of limited applicability to the vast number of proteins involved in these studies [1].

While genes exert their functions typically at the level of proteins, genetic responses to stress conditions are often regulated at transcriptional level. Nowadays, microarray technology is used to generate genome-wide transcriptional profiles. Unquestionably, this methodology is making a huge contribution to functional genomics as a whole, but it also has a number of pitfalls and shortcomings. One critical issue in microarray-based transcript quantification is sensitivity. Most robust methodology for investigations aimed at quantitative analysis is reverse transcription followed by PCR (RT-PCR; reverse transcription-polymerase chain reaction). In fact, changes detected by microarrays have to be confirmed (partially at least) by RT-PCR. Quantitative RT-PCR can be carried out in a relative or absolute manner. While relative quantification gives us the number of times (fold-variation) that the level of a particular mRNA increases or decreases in a problem relative to a reference sample, absolute quantification provides the actual mRNA molecule numbers.

Absolute quantification relates the PCR signal to input copy number using a calibration curve, therefore, its reliability depends on the condition of identical PCR amplification efficiencies for both the target and the calibrator (see Experimental and Results and Discussion sections). Relative quantification is easier to perform since a calibration curve is not necessary, and the expression level of the target gene is normalized, preferentially by using an internal standard (also referred to as control or reference gene). Hence, most RT-PCR studies make (as microarray analysis) relative comparisons of mRNA levels between samples. Besides that modulation of transcription is much better understood in absolute

Table 1 Use of metallomic techniques for traditional biomarkers analysis.

Biomarkers/element-tag		
Cytochrome P450/Fe	Superoxide dismutase/Zn,Cu,Mn	GPX/Se
Cytochrome P450 was induced in trout by intraperitoneal treatment with β-naphthoflavone. Microsomal suspensions containing 7.4 ± 0.1 nmol mL^{-1} P450 enzymes were further separated by AE-FPLC, with UV detection, or coupled to ICP-MS with an octapole reaction system, ICP-(ORS)MS (monitoring specifically Fe signals). An accurate information on the prosthetic group of the protein is obtained, which constitutes an advantageous alternative to classical methods for detection of these hemoproteins [27].	ICP-MS combined to spectrophotometry used to study the change of CuZnSOD catalytic action in relation to the SOD activity against the GSNO-reductase activity, this later inhibited by polyaminocarboxylate metal ion chelators, EDTA and DTPA [28].	Stability of Se-containing proteins in human serum studied by SEC-ICP-MS in non-denaturing conditions. GSHPx peak (tetramer) was the most affected; peaks of the GSHPx monomer and of selenite were found to appear with time [32].
	CE-ICP-MS and cHPLC-ICP-MS compared for metallothionein (MT-I and -II mix) separation from SOD and for analysis of Cu-, Zn-species in erythrocyte extracts to explore more adequate separation methods for metalloprotein biomarker analysis in complex biological matrices [29].	HPLC-ICP-MS chromatography using tandem HPLC columns with ICP-MS detection was used to detect the major selenium-containing proteins in human plasma such as GPX, albumin, and selenoprotein P [33].
	HPLC-ICP-MS used for Zn-binding protein carbonic anhydrase III (CA-III) identification in liver of male adult rats. A comparable amount of Zn to that of Cu, Zn-SOD was bound to CA-III at 8 weeks of age. CA-III concentration was decreased specifically by repeated injections of Cu(II) ions without Cu, Zn-SOD concentration being affected [30].	GMO Brassica juncea extracts studied by SEC-ICP-MS to monitor As, Cd, and S. Further purification of As fractions was performed by RP-HPLC. Structural elucidation of PCs and other thiols, and their As and Cd complexes made by ESI-Q-TOF. In both Cd- and As-exposed plants, oxidized PC2 was detected, GS-PC2 (-Glu) as well as GSH and GSSG. As $(GS)_3$ species identified for the first time in *B. juncea* [34].
	SEC and RP-HPLC/ICP-MS used for metal species quantification in pig liver with off-line identification by electrospray ionization (ESI)-MS. Transferrin, MT, and SOD were identified. Cu- and Zn-species were detected mainly in the cytosolic fraction and Fe in the microsomal fraction [31].	

quantitative terms, absolute quantification prevents the inaccuracy of most internal standards as quantitative references (see, e.g., [4–8]).

Finally, metallomics [9,10] considers that biomolecules which bind metals and metalloids constitute a substantial portion of molecules involved in cell metabolism and behavior, and that identification of a metal cofactor into a protein can greatly assist its functional assignment and help place it in the context of known cellular pathways [11–15]. The use of chromatography and highly sensitive atomic detectors (also referred to as inductively coupled plasma-mass spectrometer, ICP-MS) for heteroatom-tags species are the central components in the metallomic analytical approach [11], generally combined with MS for species identification. The ICP-MS detector allows higher sensitivity, selectivity, and reliability in molecule tracing. For this reason, metallomics is being used successfully in traditional biomarker analysis (Table 1) and provides simplified and quantitative analytical approaches in proteomics (metalloproteomics), which converts this relatively recent -omic tecnology to a valuable alternative for massive biomolecule analysis [10–12,14,16].

The aim of the present paper is to provide a preliminary working scheme for the integration of transcriptomics, proteomics, and metallomics in environmental monitoring, using *M. musculus* as reference species and *M. spretus* as bioindicator (Fig. 1), two rodent species that separated around 3 million years ago. Our studies were focused in Doñana Natural Park and its surrounding areas (SW Spain, Fig. 2). The complementary character of these three -omics would allow a more comprehensive and unequivocal assessment of environmental issues.

Fig. 1 Comparison of *M. musculus* and *M. spretus* mice.

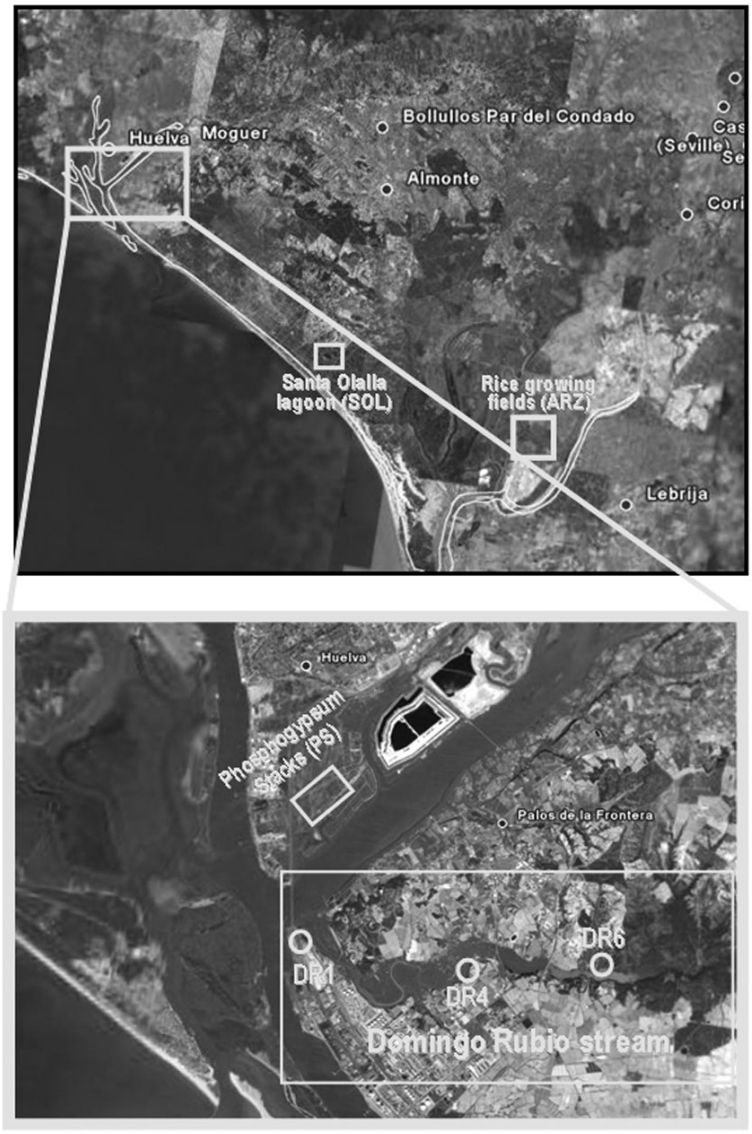

Fig. 2 Sampling areas in Doñana surroundings and the Domingo Rubio stream (Huelva province, SW Spain). Upper figure: localization of DRS, Santa Olalla lagoon (SOL), and "Matochal" rice-growing fields (ARZ) in Doñana surroundings. Lower figure: localization of sampling points in the lower (DR1), medium (DR4), and upper course (DR6) of Domingo Rubio stream, and of the phosphogypsum disposal stalks (PS) [8,17] .

EXPERIMENTAL

Sampling areas and animal trapping

Mice were collected in February 2004 at six sites from Doñana surroundings and the Domingo Rubio stream, both at Huelva province (see later for a further description). Animals were captured with live traps and taken alive to the nearest laboratory (Huelva University or Doñana Biological Reserve-CSIC). Their sex and weight were determined and those of 11–12 g were killed by cervical dislocation and dissected. Individual livers and kidneys were frozen in liquid nitrogen and stored at –80 °C [8,17].

2-DE analysis and protein identification

Around 50 mg of livers from four male mice/site were pooled and homogenized in 20 mM Tris-HCl, pH 7.6, with 0.5 M sucrose, 0.15 M KCl, 20 mM DTT (dithiothreitol), 1 mM PMSF (phenylmethanesulfonyl fluoride), and protease inhibitors, at a ratio of 3 mL/g as described [17]. Cell debris was cleared by centrifugation, and the supernatant was treated with benzonase and ultracentrifuged as described [18]. Protein extract (115 μg) was incubated 30 min in 450 μL rehydration buffer (7 M urea, 2 % CHAPS, 20 mM DTT, 0.5 % Pharmalyte 3–10, bromophenol blue traces), spun and loaded on 24 cm (pH 4–7) Amersham Immobiline Dry-Strips®. After 6 h passive and 6 h active (50 V) rehydration in a BioRad Protean IEF cell (20 °C, 50 mA/strip), the voltage was raised following published conditions [18]. After freezing at –80 °C, the strips were soaked 20 min in equilibration mix (50 mM Tris-HCl, pH 8.8, 6 M urea, 30 % glycerol, 2 % sodium dodecyl sulfate (SDS), bromophenol blue traces) with 65 mM DTT, drained and again soaked 20 min in this mix with 25 mM iodoacetamide. SDA-polyacrylamide gel electrophoresis (SDS-PAGE) was done in 12.5 % gels using the BioRad Protean® Plus Dodeca cell (20 °C) at 2.5 W/gel, 10 min, and 10 W/gel until separation was finished. Gels were silver-stained following a protocol compatible with MS analysis [17]. Analytical-quality chemicals and Milli-Q water (Millipore®) were used throughout.

Gel images of 3 replicates/sample were obtained with a BioRad GS-800 densitometer. Spot volumes were quantitated using the PDQuest software (v7.1, BioRad). Initially, only spots exhibiting in "Santa Olalla" lagoon (SOL) an over/underexpression ratio of at least threefold with respect to any other sampling site was considered. One-way analysis of variance followed by the Student–Newman–Keuls post-test was then used for a definitive selection of the spots showing altered expression patterns between the different animal groups. Differentially expressed spots were manually excised, reduced (10 mM DTT), alkylated (55 mM iodoacetamide), digested overnight at 30 °C with trypsin (Promega) and the peptides extracted with ACN/TFA (acetonitrile/trifluoroacetic acid) as described [17]. Aliquots of 0.5 μL were analyzed by MALDI-TOF peptide mass fingerprint (PMF) in a Voyager DE-PRO instrument (Applied Biosystems) in reflectron mode. PMF data were contrasted against mammalian sequences included at Swiss-Prot (EBI, Heidelberg, Germany) and nonredundant NCBI (Bethesda, MD, USA) databases using ProteinProspector (California University, San Francisco, CA, USA) and MASCOT (Matrix Science, London, UK) software [17].

Transcript quantifications

Primer design, RNA preparation, RT, and absolute quantification by real-time PCR was as detailed in [5,19]. Briefly, PCR reactions were performed in quadruplicate. No primer dimers were detected. Primers showed optimal (~100 %) PCR efficiencies in the range of 20 to 2×10^5 pg of total RNA input with high linearity ($r > 0.99$) (see Fig. 3). An absolute calibration curve was constructed with an external standard in the range of 10^2–10^9 RNA molecules. The number of mRNA molecules was calculated from the linear regression of the calibration curve ($y = -3.326x + 39.693$; $r = 0.998$) as described [5,19].

Microarray-based transcript quantification was performed by using the Whole Mouse Genome Oligo Microarray Kit (Agilent), which includes 60-mer oligonucleotide probes representing all known genes and transcripts (~41.000) of the model *M. musculus* species. Approximately, 20 μg of total RNA were converted to fluorescently labeled cDNA (with Cy3-dCTP or Cy5-dCTP) following the Agilent Fluorescent Direct Label Kit instructions. Hybridization was carried out in Agilent's SureHyb Hybridization Chambers at 65 °C for 17 h using Agilent's Gene Expression Hybridization Kit. The hybridized microarrays were then disassembled and washed at room temperature, as described in the Agilent Microarray Based Gene Expression Analysis protocol. To eliminate dye-bias, dye swap replicates were performed. Microarrays were scanned at 532 and 635 nm using a confocal scanner (Axon 4000B). The ratio of Cy5 to Cy3 was adjusted to 1 varying PMT gain as a global normalization of each

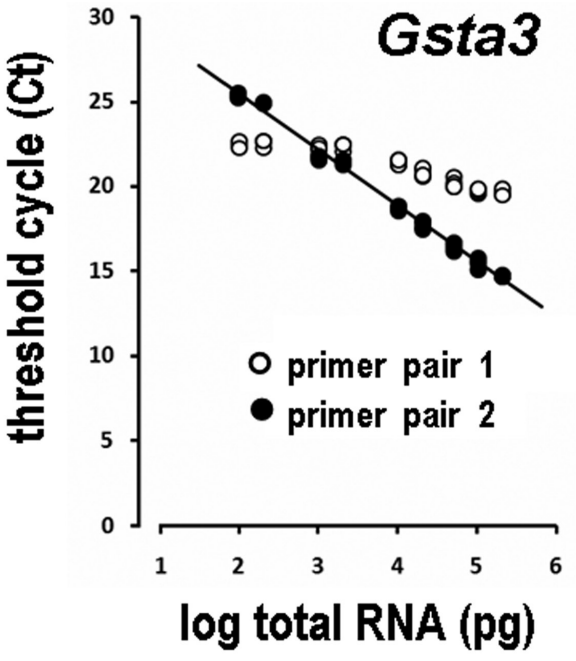

Fig. 3 Real-time PCR efficiency of *M. spretus Gsta3* mRNA amplification. Ten-fold serial dilutions from 2×10^5 to 2×10^1 pg of total RNA input were prepared, retrotranscribed, and amplified by real-time PCR. RNA sample was from *M spretus* liver. A plot of the log RNA dilution vs. the Ct value was made for each pair of primers. Nucleotide sequence results revealed a C/A mismatch at the 3'-end of one primer of pair 1. Primers of pair 2 were exactly complementary to the desired *M. spretus* template location and showed optimal 100 % efficiency with a very high linearity ($y = -3.317x + 32.13$; $r = 0.990$) over 4 orders of magnitude. The efficiency (E) value is calculated from the slope of the efficiency curve equation, as $E = 10^{[-1/\text{slope}]} - 1$.

array. The images were analyzed using GenePix Pro v4.1 software (Axon), and data were subsequently input to Genespring v7.3 software (Agilent) for further analysis.

Total elements determination

Mineralization of extracts from different organs of mice was performed in a microwave-accelerated reaction system. For this purpose, the extracts were weighed exactly (0.5000 g) in 55-ml microwave vessels and 10 ml of a mixture containing nitric acid and hydrogen peroxide (4:1 v/v) was added. After 10 min, the poly(tetrafluoroethylene) (PTFE) vessels were closed and introduced into the microwave oven. The mineralization was carried out at 1200 W from room temperature ramped to 180 °C in 15 min and with holding for 40 min at this temperature. Then, solutions were made up to 25 ml and analyzed by ICP-MS. Finally, Mn, Ni, Cu, Zn, As, Pb, Cr, Fe, Co, Se, and Cd were measured in the extracts from the different organs.

Analysis of extracts by size exclusion chromatography (SEC) coupled with ICP-MS

Extracts were twofold diluted with the mobile phase and centrifuged at 15 557 x *g* for 1 h at 4 °C, and latterly filtered through Iso-Disc poly(vinylidene difluoride) filters (25-mm diameter, 0.2-µm pore size) to avoid column overloading or clogging. Elemental fractionation profiles were obtained by SEC coupled to ICP-MS as detector. Two columns were used in the experiment: Hiload 26/60 Superdex 30 Prep column for a separation range below 10 kDa (low molecular mass, LMM) and a Superdex 75 Prep col-

umn for a separation range of 3–70 kDa (high molecular mass, HMM), both from Amersham Biosciences (Uppsala, Sweden). These columns were calibrated using standards of known molecular mass, such as bovine serum albumin (67 kDa), MT-I (7 kDa), gastrin rat I (2126 Da) and Gly6 (360 Da) for low-molecular-weight (LMW) column, and bovine serum albumin (67 kDa), chymotrypsinogen A (25 kDa), ribonuclease A (13.7 kDa), and MT-I (7 kDa) for high-molecular-weight (HMW) column. The void retention time was estimated with bovine serum albumin (67 kDa) and blue dextran (2000 kDa), for LMW and HMW, respectively.

RESULTS AND DISCUSSION

A combined application of transcriptomic, proteomic, and metallomic approaches has been performed in relation to Doñana Natural Park, one of the most important European biological reserves, in which millions of migrating birds land each year on their way to/from Africa. This uncontaminated area was used as negative reference (SOL) in comparison to the neighboring "Domingo Rubio" stream (DR1 to DR6) and the positive references (PS and ARZ). These areas are contaminated by mining, agricultural, and industrial effluents (DR1 to DR6 and PS) and by pesticides and fertilizers (ARZ) (Fig. 2).

M. spretus is an unprotected rodent that attains high population densities, typically inhabits marshlands, and feeds on plants, seeds, and insects around its burrow. This aboriginal species has proved to be a useful sentinel organism in several monitoring programs, through the use of conventional biometric, cytogenetic, and biochemical biomarkers [20–22]. The use of modern molecular biology methodologies in ecotoxicological studies is limited by the fact that most popular bioindicators are poorly represented in gene/protein sequence databases. This problem can be circumvented by using as sentinels, species close to model organisms that are consequently well represented in public databases. Here, we applied new -omic technologies in the wildlife species *M. spretus* based on the gene/protein sequence databases from the key model species *M. musculus.*

Transcript expression signatures

We first evaluated whether the absolute measurement of mRNA levels from selected key genes is a reliable novel biomonitoring end-point to assess the exposure and biological effects of pollutants on free-living nonmodel *M. spretus* [8]. To this end, the mRNA molecules of genes coding for different CYPs and GSTs were quantified in mice dwelling at SOL and PS.

A key feature in quantitative RT-PCR is primer design. For absolute transcript quantification we always design primers that amplify the targets and the calibrator with optimal (100 %) PCR efficiencies (see, e.g., [4,5,19]). Primers for RT-PCR quantification of *M. spretus Cyp* and *Gst* mRNAs were designed based on known gene sequences from *M. musculus*. Remarkably, these primers, when amplified in *M. spretus*, gave single products exhibiting in most cases 100 % nucleotide sequence identity. Therefore, most designed primers were exactly complementary to the desired *M. spretus* templates and amplified them with 100 % efficiency. In a few cases, however, primers should be redesigned based on nucleotide sequences of PCR fragments from *M. spretus* (see example in Fig. 3) [8].

As an example, Fig. 4 shows the concomitant up-regulation of some *Cyp* transcripts in *M. spretus* PS population, as compared to that at SOL in the Doñana Biological Reserve. From a quantitative perspective, the relevance of the absolute measurements reported in Fig. 4 is highlighted when comparing the increments in transcript molecules with the conventional fold variations (in parentheses). For instance, although a relative 2.7-fold increase in *Cyp2e1* transcript levels might look modest compared to the 6.8-fold increment in *Cyp1a2* mRNA, the actual scenario is that *Cyp2e1* (highly abundant mRNA in liver) exhibited a much higher increase in copy number than *Cyp1a2* (low abundant mRNA). The hepatic levels of the *Cyp* transcripts that were increased in PS population are known to be regulated by a myriad of structurally unrelated compounds through a wide variety of mechanisms (including the generation of reactive oxygen species). For instance, *Cyp1a2* is typically activated by dioxins and poly-

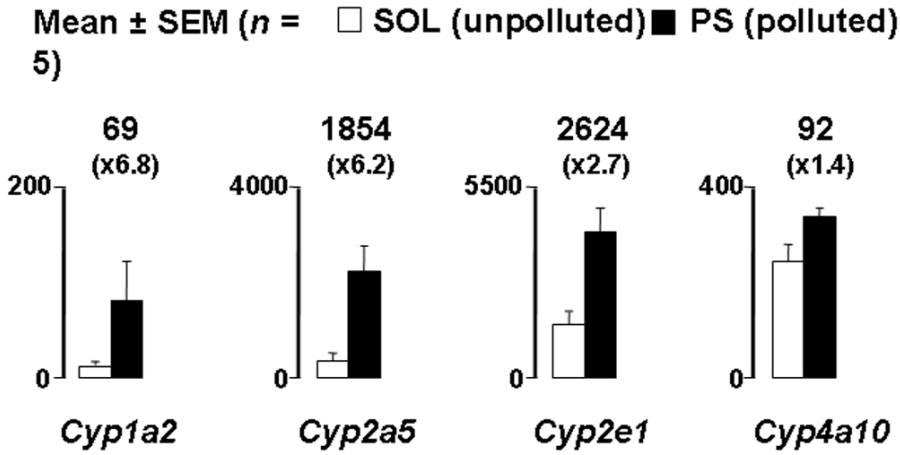

Fig. 4 Area-associated differences in *Cyp* transcript levels. Data are means ± SEM (n = 5) of hepatic mRNA molecules/pg RNA from males captured at SOL and PS. The increments in mRNA copy numbers and the conventional fold-variations (in parentheses) are given for comparison.

cyclic aromatic hydrocarbons, *Cyp2a5* by nitrogen heterocycles and chlorinated hydrocarbons, *Cyp2E1* by low Mr solvents and industrial monomers, and *Cyp4a* subfamily by a diverse group of chemicals, known as peroxisome proliferators. Therefore, the absolute *Cyp* transcript expression signature of mice living at PS seems the logical result of exposure to a complex set of organic pollutants, in line with the industrial activity at this area of concern [8].

Genome-wide transcriptional profiling using microarrays is an increasingly important molecular tool for identifying responsive genes that may serve as biomarkers. In the postgenomic era, there is extensive interest in the application of this technology to the study of nonmodel organisms in the environment. The considerable gene sequence similarity between the aboriginal and the model mouse species, suggested by the above-referred study, prompted us to investigate the suitability of commercial *M. musculus* oligonucleotide microarray for application in the sentinel species *M. spretus*. To this end, the transcriptomes of mice dwelling at SOL and DR1 were compared. As exemplified in Fig. 5, the considerable high level of cross-hybridization to the heterologous microarray allows detecting even small-fold changes in mRNA expression, and so has the potential to greatly enhance the utility of this -omic to environmental studies.

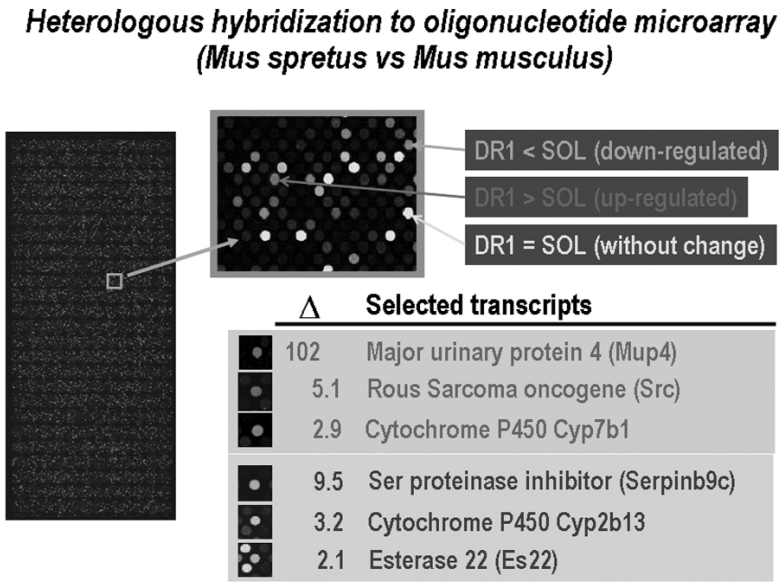

Fig. 5 Heterologous microarray analysis for area-associated differences at the transcriptome level. Total RNA from the livers of six male *M. spretus* mice per sampling site was grouped. Each pooled mixture was then converted to fluorescently labeled cDNA as described in the Experimental section. We swapped dye usage between SOL and DR1 samples in order to account for potential dye bias. A total of four microarrays were hybridized. In this microarray image, DR1 and SOL samples were labeled with red Cy5 and green Cy3 fluorescent dyes, respectively. The numbers of times that three of the up-regulated and three of the down-regulated transcripts were increased or decreased in DR1 population as compared to SOL reference are indicated as example.

Proteomics study of cytosolic liver extract from the mouse *M. spretus*

In view of the similarity of both mouse species, *M. musculus* and *M. spretus*, at DNA sequence level [7,8], a comparison was also made at proteomic level. Figure 6 compares the proteomes from the laboratory mouse, *M. musculus* (BALB/c), and the aboriginal species, *M. spretus* (SPRET/EiJ), both inbred strains from commercial sources (Jackson Laboratory, USA), focusing on hepatic soluble proteins of Mr 5–100 kDa and pI 4.0–7.0. Although both proteomes were similar, clear differences were also visible, both at quantitative and qualitative levels. For instance, the spot of pI 6.5 and 6 kDa was quite intense in *M. musculus* but fairly invisible in *M. spretus*. Oppositely, the set of pI ~4.5 and ~16 kDa was quite intense in *M. spretus*, where spots were clearly detected, while in *M. musculus* only 7 of these spots were barely visible. In principle, these differences could be attributed to any of the multiple steps at which gene expression can be regulated in mammals, e.g., synthesis, splicing, and stability of transcripts, and synthesis, modification, and degradation of proteins. Although the nature of the proteomic differences between the two mouse species has not been further studied, at least some of them most probably arise from differences in the cellular machinery responsible for post-translational modifications [23], as suggested by the clear differences shown by a "spot train" of pI 4.0–4.5 and ~40 kDa.

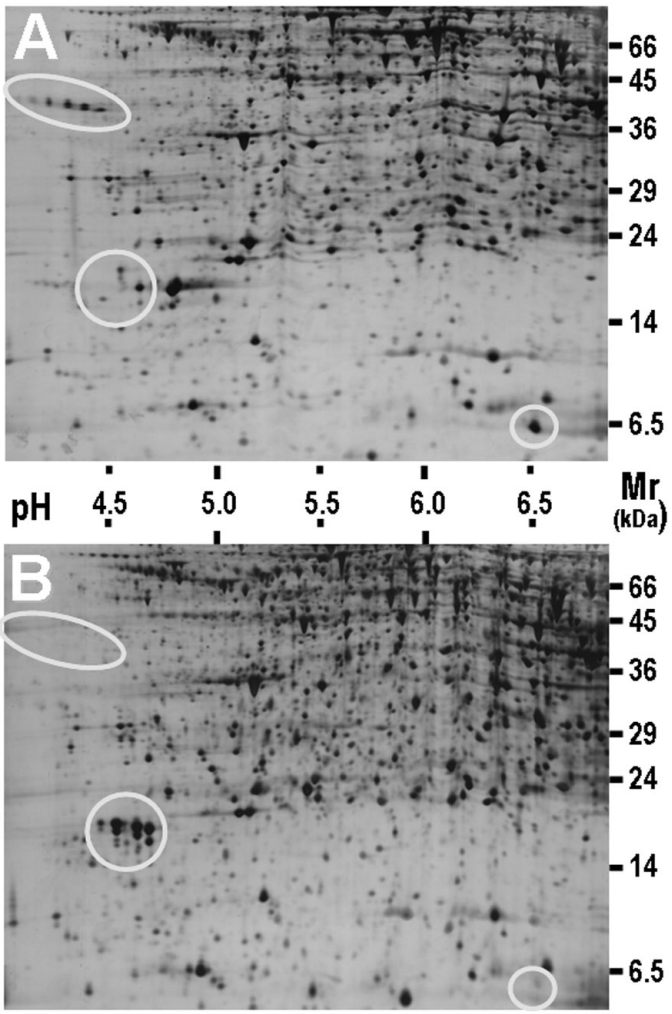

Fig. 6 Comparative 2-DE proteomes of hepatic cytosolic proteins from *M. musculus* (A) and *M. spretus* (B) mice. (See Experimental section for further details.)

Cytosolic fractions from a liver pool (4 male *M. spretus* mice) per sampling site of Doñana and Domingo Rubio streams were analyzed by 2-DE in search of protein expression differences. Over 2500 spots were resolved in the pH range 4–7 and 14–70 kDa Mr. Image analysis of the gels yielded 36 spots with significantly altered expression. Of them, 16 proteins were identified by MALDI-TOF-PMF and heterologous search against *M. musculus* databases, as summarized in Table 2. They showed three expression patterns: (i) high in all/some of the DRS sites, as compared to SOL reference site; (ii) lower in all/some of the DRS sites than in SOL; and (iii) with the highest intensity at PS. The identified proteins were sorted into different cellular functions: axonal transport and cell division, proteolysis, central pathways of glucid, fatty acid, amino acid, methyl, and urea metabolism, biotransformation, and adaptation to oxidative stress [17]. Animals from different polluted environments showed contrasting differences in their proteomes, with specific increases and decreases in selected groups of proteins that seemed to be coordinately regulated. Some protein expression changes protected mice from the toxic effects of pollutants, such as the increase of GSTω1 that protected from As toxicity or the increase of glycine *N*-methyltransferase protective against polycyclic aromatic hydrocarbon (PAH)-derived hepa-

totoxicity. Other changes rendered the mice more susceptible, such as the decrease of putative L-aspartate dehydrogenase and the increase of cysteine protease ATG4B. Proteomic data were consistent with metal biomonitoring and conventional biomarker responses, indicating that DRS (and PS/ARZ) mice sustained a heavier pollutant burden than SOL specimens and suffered a chronic oxidative stress. In fact, 10 of the 16 identified proteins protected polluted mice from oxidative stress or were targets of oxidative damages, directly or due to increased proteolytic susceptibility [17].

Table 2 Proteins identified by MALDI-TOF-PMF and database search [17].

Protein name	Expression pattern	Environmental relevance
Tubulin beta-3 chain	High in DRS	Abundant cytosqueletal protein that acts as a redox buffer to protect other biomolecules from oxidative damages
Triosephosphate isomerase	High in DRS	Redox-sensitive glycolytic enzyme that reprograms glucose metabolism under oxidative stress conditions
Fructose-1,6-bisphosphatase 1	High in DRS	Redox-sensitive gluconeogenic enzyme activated to help pensose-P cycle to form NADPH under oxidative stress
GPX1	High in DRS	Major Se-containing antioxidant enzyme for GSH-dependent removal of peroxides under oxidative stress
GPX1*	High in DRS	Acidic GPX1 isoform (1.4 pH U lower pI) formed by oxidative modification under oxidative stress
EnoylCoA hydratase mitochond. [prec]	Low in DRS	β-oxidation enzyme that diminishes in oxidatively stressed animals by inhibition of its mitochondrial import
Glutathione transferase omega-1	Low in DRS	Reduces various As species to highly toxic +3As forms. Low activity would be protective against As toxicity.
Putative L-aspartate dehydrogenase	Low in DRS	Putatively involved in NAD synthesis. Its decrease would be detrimental for oxidatively stressed animals.
PRX-1	Low in DRS	2-Cys-enzyme for peroxide removal. It almost completely disappears by inactivation under oxidative stress.
PRX-2	Low in DRS	Thiol-specific enzyme that removes peroxides and disappears by irreversible inactivation under oxidative stress
Cysteine protease ATG4B	PS the highest	Proteolytic product of a Cys-protease involved in autophagy. It would negatively affect many biological processes.
Methionine adenosyltransferase 1, alpha	PS the highest	Involved in SAM synthesis. Its overexpression increases GSH synthesis from SAhCys to combat oxidative stress
Glycine *N*-methyltransferase	PS the highest	Methylates many acceptors using SAM as donor. It also protects from B(a)P and other PAH hepatocarcinogens
Ornitine transcarbamylase mitochond. [prec]	PS the highest	For urea synthesis. In fasting conditions it increases to adapt to glucose shortage and higher protein degradation
HMG-CoA synthase 2 mitochond. [prec]	PS the highest	Involved in ketogenesis. It would increase after fast to adapt to glucose shortage and higher protein degradation
PRX-6	PS the highest	2-Cys-enzyme for peroxide removal. Its up-regulation could compensate the low GST1 levels in PS and ARZ mice

Different sources of protein alteration can be envisaged to explain the intensity changes observed when organisms with different pollutant loads are subjected to proteomic analysis. Although increased

protein expression has been demonstrated in some cases, several post-translational modifications may shift protein position in 2-DE gels. In our proteomic analysis with *M. spretus*, the presence of two isospots corresponding to GPX1 and another isoform with the same Mr (~22 kDa) but 1.5 pH units lower pI (6.8 and 5.3, respectively) was remarkable. Comparison with *M. musculus* values suggested that the neutral spot (GPX1) was the native protein and the acidic (GPX1*) the modified form [17]. Exposure to oxidative stress introduces a wide range of reversible or irreversible alterations to amino acid side chains. They include carbonylation, glutathionylation, formation of mixed disulfides, effects of disulfide bridge patterns, ubiquitinylation, and racemization [24]. Other modifications could also explain spot intensity changes, including phosphorylation of Ser, Tyr, or Thr residues, Cys nitrosation, Tyr nitration, and limited proteolysis, as shown in *Chamaelea gallina* actin [25], and in *M. spretus* with Cys-protease ATG4B [17] that could be explained by polypeptide fragmentation promoted by chemical modifications of some residues [26].

Comparison between proteomic and transcriptomic data

The changes observed in protein profiles may not reflect alterations in gene expression at the transcript level. The consistency between protein and mRNA levels was investigated by quantifying the copy numbers of the transcripts encoding four of the proteins with altered expression profiles, i.e., GSTω1, GPX1, PRX1, and PRX6. No significant alterations were found in the absolute amounts of three out of the four investigated transcripts, in agreement with the idea of lack of correlation between mRNA and protein levels due to post-transcriptional regulations. Interestingly, however, significant area-associated differences were found in the absolute amounts of the *gpx1* transcript [17]. Of note also is the additional observation of a positive quantitative correlation ($p = 0.04$, $r = 0.96$) between the GPX1 spot intensities and the *gpx1* transcript levels. Contrarily, no quantitative relationship was observed between the GPX1* spot intensities and the *gpx1* mRNA levels, supporting further the idea that this acidic isoform most likely results from post-transductional modifications.

Metallomics approximation to the mouse *M. musculus*

In this research, the presence of unknown metallobiomolecules in *M. musculus* is studied for the first time. Metals bound to cytosolic molecules extracted from lung, liver, spleen, kidney, brain, testicle, heart, and muscle from inbred *M. musculus* mice were analyzed by molecular mass using the coupling of SEC with ICP-MS. Previous estimation of total metal levels in these samples reveals that Fe is the most abundant with an averaged level in the different organ extracts of 7500 $\mu g\ l^{-1}$, followed by Zn 2375 $\mu g\ l^{-1}$, Cu 555 $\mu g\ l^{-1}$, Cr 402 $\mu g\ l^{-1}$, Ni 165 $\mu g\ l^{-1}$, Mn 160 $\mu g\ l^{-1}$, and Se 145 $\mu g\ l^{-1}$. A toxic element as Pb is present at low levels (26 $\mu g\ l^{-1}$). Other nonessential elements such as Cd and As were not detected when they were evaluated as total element in the different tissues, and other metals were absent in some organs, such as Fe (in brain and testicle) and Se (in spleen). Cobalt was only detected in the lung sample. Most metals (Cu, Zn, Ni, and Se) were accumulated in liver, although Cr and Mn presented higher concentration in lung. The very high content of Zn and Cu in brain, related to the presence of a biomolecule of about 7 kDa (Fig. 7), is also remarkable

Molecular size distribution patterns of elements by SEC-ICP-MS

SEC-ICP-MS profiles were obtained for Mn, Ni, Cu, Zn, As, Pb, and Fe. In contrast, Cr, Cd, Co, and Se were not detected in any sample possibly because of the low intensity and high noise of the signal.

The combined use of two SEC columns, LMM and HMM, allowed successful separations of molecules from 360 to 70 000 Da. All extracts showed peaks in the range of 300 and 2000 Da whatever metal or organ considered. In addition, a peak was observed in the void volume of the LMM column (Mr < 10 kDa) for all profiles, although they showed low intensities except for Cu that was studied in more detail (Fig 7). A Cu-containing fraction (7–10 kDa) was found in the brain, and was absent in all

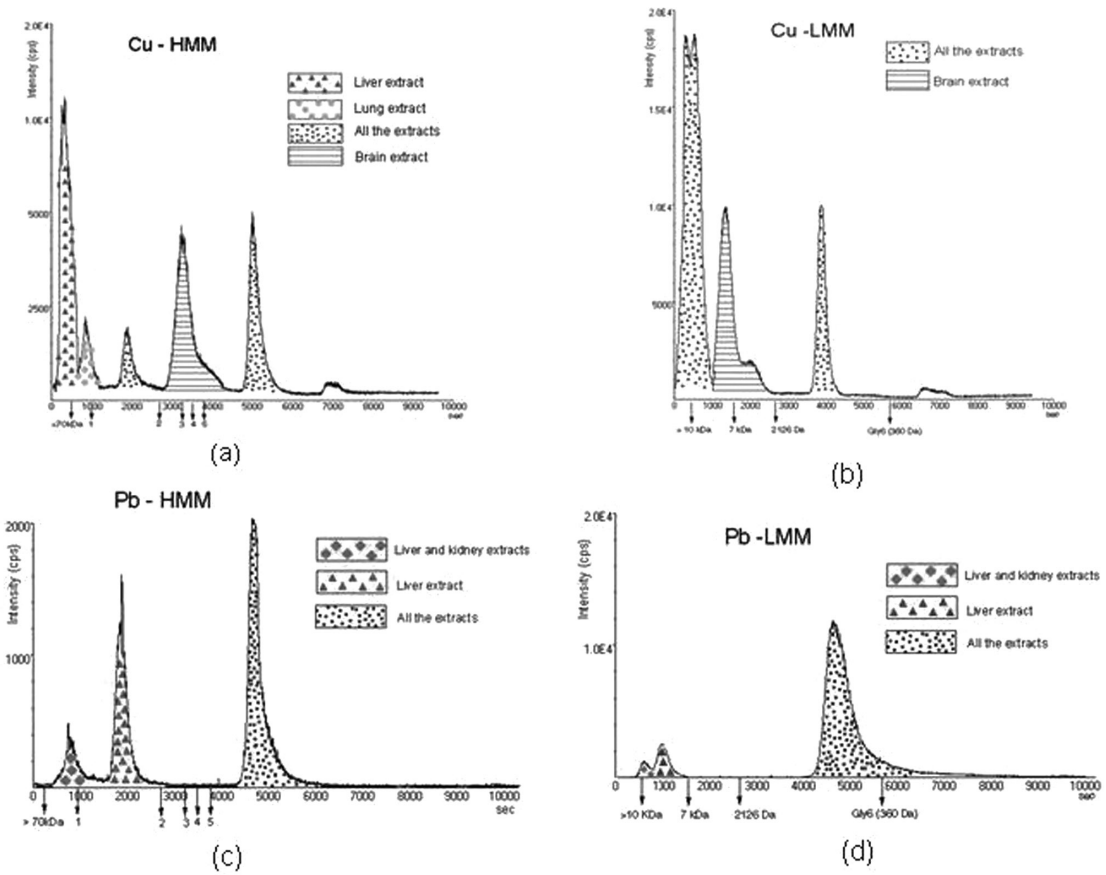

Fig. 7 Cu and Pb biomolecules profiles from *M. musculus* extracts. Numbers in *x*-axis at the HMW column correspond to the standards of known molecular weight: (1) 67 kDa, (2) 25 kDa, (3) 13.7 kDa, (4) 7 kDa, and (5) <3 kDa.

other M. musculus organs. Other elements, e.g. Pb (Fig. 7), nor have peaks equivalent in that range of masses. Therefore, further studies involving subsequent purification by another chromatography technique and isolation of this Cu-bound biomolecule for MS identification are necessary. The Cu and Pb fractions resolved in the HMM column were mainly associated with molecules with molecular masses from 25 to 67 kDa, which are present in all the organs for Cu and only in liver for Pb. A singular chromatographic peak associated to Cu was detected in the molecular mass range 67–70 kDa for the lung extract, and in liver and kidneys for Pb. In addition, a Cu peak can be observed in liver extract at the void volume of the HMM column.

Combination of metallomics and proteomics

Crude extract of *M. musculus* liver was analyzed by 2-DE (Fig. 8a). A similar study was performed on the SEC-ICP-MS Cu fraction isolated in the mass range 60–70 kDa (Fig. 8b). About 1300 spots were detected in the crude extract (Fig. 8a), while purification using SEC and ICP-MS for Cu tracing clearly diminished the number of molecules, and therefore of spots to about 400. Actually, this combination of metallomics and proteomics approaches proves that Cu-associated SEC peak at 60–70 kDa corresponds to a numerous set of molecules. In fact, combination of 2-DE and SEC-ICP-MS approaches is a useful tool to assess the metal-biomolecule isolation performed by SEC-ICP-MS. In addition, this integration

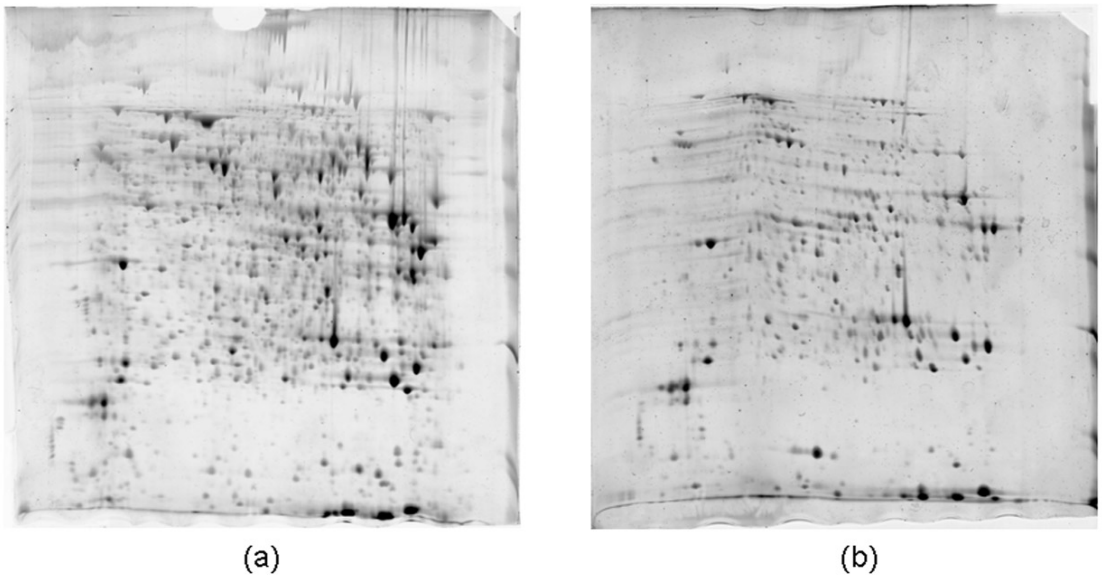

Fig. 8 2-DE gels from *M. musculus* liver extract: (A) crude extract (100 µg); (B) Cu 60–70 kDa fraction after SEC-ICP-MS. 2-DE separations were carried out in 18-cm gels using a nonlinear pH gradient ranging from 3.0 to 11 pH units.

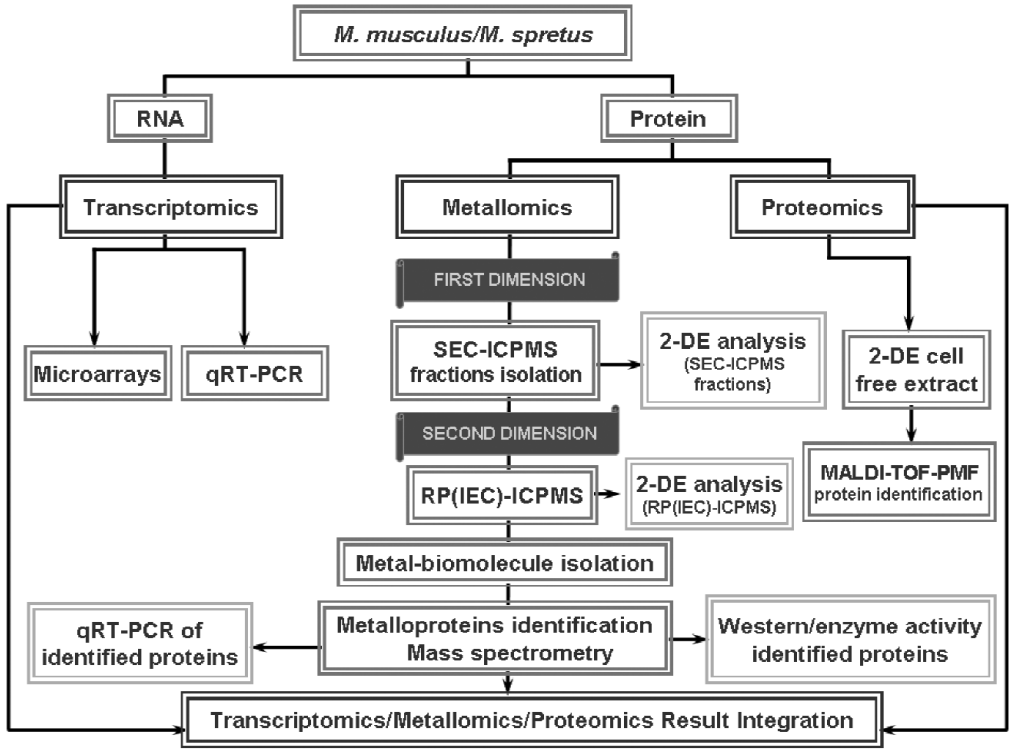

Fig. 9 Workflow for integration of metallomics with proteomics and transcriptomics.

can assure the final presence in the extract of few or even only one biomolecule for MS/MS identification in the *Second Dimension* step proposed in the integrated metallomic-proteomic-transcriptomic scheme that follows in the following section (Fig. 9) [11].

Integration of metallomics with proteomics and transcriptomics

The general aim in current studies of environmental pollution claims for the integration of the three -omics focusing in this field, metallomics, proteomics, and transcriptomics. The availability of the complete sequence of an increasing number of genomes allows the developments in proteomics, which can be achieved by electrospray and MALDI-TOFMS. However, many proteins can be sensitively detected by ICP-MS due to the presence of ICP-ionizable elements (heteroelements). Therefore, the three -omics are complementary, and in this context the combination of validated metallomic data with the transcriptome and proteome knowledge of a cell is one of the largest challenges for future research in this topic.

A crucial question is that better established -omics (transcriptomics and proteomics) support their studies on model organisms whose gene/protein sequence is well collected in databases. On the contrary, bioindicators currently used in environmental studies are nonmodel organisms absent from public databases. The design of primers that allows the confirmation of homology in gene expression at the transcript level between bioindicators used in environmental studies and sequenced model specimens is a very good tool to perform environmental metallomics and proteomics.

This new triple-omics approach (MEPROTRANS-*triple*-OMICS) is being applied to the *M. musculus–M. spretus* couple in Doñana Natural Park for environmental assessment, following the steps shown in the Fig. 9 workflow:

i. Extraction of RNA for transcriptomic and of proteins for proteomic and metallomic analyses.
ii. Quantification of changes at the transcriptome level by means of commercially available microarrays and further validation of selected microarray results by absolute quantification using real-time RT-PCR.
iii. Quantification of changes at the proteome level by 2-DE analysis and subsequent protein identification by MALDI-TOF-PMF.
iv. Integration of transcriptomic and proteomic results to distinguish transcriptional from post-transcriptional changes.
v. Isolation of heteroatom-tagged molecules by SEC fractionation of tissue extracts guided by ICP-MS (***First Dimension***) and further purification by reverse-phase and/or ion exchange chromatography with ICP-MS on the previous selected fractions (***Second Dimension***). This metallomic scheme will be assisted by 2-DE analysis to reveal the complexity of the successive subproteomes. The purification steps involving metallomics and proteomics will be repeated until the final isolation of target metal-biomolecules for identification by tandem MS.
vi. Validation of metallomic results by Western hybridization and/or enzymatic assays, and by absolute quantification of transcripts coding the identified metalloproteins.
vii. Integrative multidisplinary decision on the optimal assay (high sensitivity, low cost, and easy to perform) to be recommended for routine assessment of differentially expressed metalloproteins as novel biomarker in ecotoxicological studies.

CONCLUDING REMARKS

The study of complex environmental systems, where many variables are involved, requires multidisciplinary tools for a comprehensive assessment of contaminant effects on living organisms. The use of bioindicators and biomarkers has been a good alternative widely used for this purpose, although these approaches have limitations since the use of a short number of biomolecules as biomarkers under-

estimates actual serious situations. Several -omics, such as transcriptomics, proteomics, and metallomics, offer a valuable alternative in this field since they provide massive information about biomolecules in cells and organisms that overcome these problems. Nevertheless, an overall evaluation of changes that contaminants induce in cells is only possible by integration of -omics, since the transcripts induced by pollutants (transcriptomics) encode proteins with altered expression profiles, which undergo post-traductional modifications (proteomics). In addition, many proteins related to enviromental issues are bound to metals (i.e., CYPs, SODs, GPXs, MTs) that make advisable the use of metal-tagged techniques (metallomics) as a preliminary step to simplify proteomic and transcriptomic approaches. Reversely, proteomics can guide metallomics by checking the presence of molecules in extracts obtained with the chromatography-ICP-MS couplings, which assists metal biomolecule isolation for MS identification since 2D analysis provides an image of the number of molecules in the extract and their molecular mass. Thus, this triple -omic approach (MEPROTRAN-*triple*-OMICS) is a very useful and comprehensive alternative in the study of environmental issues and the diagnosis of contamination threats. Finally, the triple -omic approach can assist in the absolute validation of traditional biomarkers with the purpose of obtaining cheap, fast, and reliable methodologies to assess environmental issues, which can be realistically used in routine laboratory procedures for environmental monitoring.

ACKNOWLEDGMENTS

This work was supported by the projects CTM2006-08960-C01/02 and BFU2005-02896 from the Ministerio de Educación y Ciencia and by projects FQM-348 and RNM-523 from the Consejería de Innovación, Ciencia y Empresa (Junta de Andalucía). M.G.-F. thanks the Ministerio de Educación y Ciencia for a predoctoral scholarship. J.L.G.-A. thanks Prof. H. Haraguchi for the invitation to "International Metallomics Symposium" held in Nagoya (Japan) in November–December 2007.

REFERENCES

1. J. López-Barea, J. L. Gómez-Ariza. *Proteomics* **6**, Suppl. 1, S51 (2006).
2. J. López-Barea. *Arch. Toxicol.* Suppl. 17, 57 (1995).
3. J. López-Barea, C. Pueyo. *Mutat. Res.* **399**, 3 (1998).
4. J. Jurado, M. J. Prieto-Álamo, J. Madrid-Risquez, C. Pueyo. *J. Biol. Chem.* **278**, 45546 (2003).
5. M. J. Prieto-Álamo, J. M. Cabrera-Luque, C. Pueyo. *Gene. Expr.* **11**, 23 (2003).
6. A. Jiménez, M. J. Prieto-Álamo, C. A. Fuentes-Almagro, J. Jurado, J. A. Gustafsson, C. Pueyo, A. Miranda-Vizuete. *Biochem. Biophys. Res. Commun.* **330**, 65 (2005).
7. J. Ruiz-Laguna, N. Abril, M. J. Prieto-Álamo, J. López-Barea, C. Pueyo. *Gene Exp.* **12**, 165 (2005).
8. J. Ruiz-Laguna, N. Abril, T. García-Barrera, J. L. Gómez-Ariza, J. López-Barea, C. Pueyo. *Environ. Sci. Technol.* **40**, 3646 (2006).
9. H. Haraguchi, H. Matsura. In *Bitrel, Wako (Saitama)*, S. Enomoto (Ed.), Fujiyoshida (Yamanashi), Japan (2003).
10. J. Szpunar. *Anal. Bioanal. Chem.* **378**, 54 (2004).
11. J. L. Gómez-Ariza, T. García-Barrera, F. Lorenzo, V. Bernal, M. J. Villegas, V. Oliveira. *Anal. Chim. Acta* **524**, 15 (2004).
12. J. Szpunar, R. Lobinski, A. Prange. *Appl. Spectrosc.* **57**, 102A (2003).
13. N. Jakubowski, R. Lobinski, L. Moens. *J. Anal. At. Spectrom.* **19**, 1 (2004).
14. A. Sanz-Medel. *Anal. Bioanal. Chem.* **381**, 1 (2005).
15. D. W. Koppenaal, G. M. Hieftje. *J. Anal. At. Spectrom.* **22**, 111 (2007).
16. M. González-Fernández, T. García-Barrera, A. Arias-Borrego, D. Bonilla-Valverde, J. López-Barea, C. Pueyo, J. L. Gómez-Ariza. *Anal. Bioanal. Chem.* **390**, 17 (2008).

17. R. Montes-Nieto, C. A. Fuentes-Almagro, D. Bonilla-Valverde, M. J. Prieto-Álamo, J. Jurado, M. Carrascal, J. L. Gómez-Ariza, J. López-Barea, C. Pueyo. *Proteomics* **7**, 4376 (2007).
18. A. Romero-Ruiz, M. Carrascal, J. Alhama, J. L. Gómez-Ariza, J. Abian, J. López-Barea. *Proteomics.* **6**, Suppl 1, S245 (2006).
19. J. Jurado, C. A. Fuentes-Almagro, M. J. Prieto-Álamo, C. Pueyo. *BMC Mol. Biol.* **8**, 83 (2007).
20. A. C. Nunes, M. L. Mathias, A. M. Crespo. *Environ. Pollut.* **113**, 87 (2001).
21. J. Ruiz-Laguna, C. Garcia-Alfonso, J. Peinado, S. Moreno, L. Leradi, M. Cristaldi, J. López-Barea. *Biomarkers* **6**, 146 (2001).
22. D. Bonilla-Valverde, J. Ruiz-Laguna, A. Muñoz, J. Ballesteros, F. Lorenzo, J. L. Gómez-Ariza, J. López-Barea. *Toxicology* **197**, 123 (2004).
23. J. Klose, C. Nock, M. Herrmann, K. Stuhler, K. Marcus, M. Bluggel, E. Krause, L. C. Schalkwyk, S. Rastan, S. D. Brown, K. Bussow, H. Himmelbauer, H. Lehrach. *Nat. Genet.* **30**, 385 (2002).
24. D. Sheehan. *Biochem. Biophys. Res. Commun.* **349**, 455 (2006).
25. M. J. Rodríguez-Ortega, B. E. Grosvik, A. Rodríguez-Ariza, A. Goksoyr, J. López-Barea. *Proteomics* **3**, 1535 (2003).
26. B. J. Kim, B. L. Hood, R. A. Aragon, J. P. Hardwick, T. P. Conrads, T. D. Veenstra, B. J. Song. *Proteomics* **6**, 1250 (2006).
27. A. Rodríguez-Cea, M. R. Fernández de la Campa, A. Sanz-Medel. *Anal. Bioanal. Chem*. **381**, 388 (2005).
28. M. Ye, A. M. English. *Biochemistry* **45**, 12723 (2006).
29. M. Montes-Bayón, D. Profrock, A. Sanz-Medel, A. Prange. *J. Chromatogr., A* **1114**, 138 (2006).
30. K. T. Suzuki, J. Takenaka, Y. Ogra. *Chem.-Biol. Interact.* **122**, 185 (1999).
31. V. Nischwitz, B. Michalke, A. Kettrup. *J. Anal. At Spectrom.* **18**, 444 (2003).
32. O. Palacios, R. Lobinski. *Talanta* **71**, 1813 (2007).
33. G. A. Jacobson, A. M. Featherstone, A. T. Townsend, R. Lord, G. M. Peterson. *Biol. Trace Element Res.* **107**, 213 (2005).
34. A. P. Navaza, M. Montes-Bayón, D. L. LeDuc, N. Terry, A. Sanz-Medel. *J. Mass Spectrom.* **41**, 323 (2006).