Published in final edited form as:

Biochem Biophys Res Commun. 2019 June 25; 514(2): 475–481. doi:10.1016/j.bbrc.2019.04.172.

Quantitative proteomic study of arsenic treated mouse Liver sinusoidal endothelial cells using a reverse super-SILAC method

Wenbo Li^{1,2}, Jiyang Zhang³, Yongzhuang Lv¹, Nader Sheibani^{4,*}

¹State Key Laboratory of Proteomics, Beijing Proteome Research Center, 100850, China

²National Engineering Research Center for Protein Drugs, Beijing 100850, China

³College of Mechatronic Engineering and Automatic Control, National University of Defense Technology, Changsha 410073, China

⁴Departments of Ophthalmology and Visual Sciences, University of Wisconsin School of Medicine and Public Health, Madison, WI USA

Abstract

Liver sinusoidal endothelial cells are the border patrol in the liver. Their open transcellular fenestrations allow the transfer of small and dissolved substances from the blood into the liver parenchymal cells. Fenestrations are dynamic structures, and many drugs and diseases alter their size and number, thus making them an important target for modulation. There is an urgent need to understand how various diseases, toxic substances, and physiological conditions influence liver endothelial cell fenestrations, and how these changes affects liver function. This work represents a straightforward quantitative proteomics study of the in vivo arsenic-stressed liver sinusoidal endothelial cells using a reverse super-SILAC based method. The aim of this study was to identify proteins, which are up- or down- regulated in response to arsenic. This knowledge will aid in identification of potential targets and mechanisms of arsenic toxicity and novel ways to reverse these changes.

Keywords

Mouse liver sinusoidal endothelial cell; Arsenic; Mass spectrometry; Proteomics; Super-SILAC

1. Introduction

Endothelial cells, which line the liver sinusoids are the most abundant non-parenchymal cells in the liver and control the transfer of nutrients, lipids, and lipoproteins. Under physiological conditions, these cells are perforated by fenestrations and lack a basal lamina.

*Correspondence: nsheibanikar@wisc.edu.

Conflict of interests

The authors have no conflicts of interest to declare.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

However, under some pathological conditions, they lose their fenestrations and form a continuous basal lamina [1]. This phenomenon is referred to as “capillarization”. Sinusoidal capillarization precedes liver fibrosis in various liver diseases [2]. Miyao et al studied a cholangiopathy model and demonstrated that capillarization appears before liver fibrosis [3]. Xie et al reported that capillarization has a pivotal role in hepatic stellate cell (HSC) activation and fibrogenesis during the late stage in a rat liver fibrosis model [4]. Liver sinusoidal endothelial cells (LSEC) are speculated to have an anti-inflammatory role in cooperation with Kupffer cells, and to have a substantial role in fibrogenesis by promoting HSC activation [5].

Arsenic is a toxic metalloid. Drinking arsenic-contaminated water increases risk of many diseases. Even at 10 ppb, arsenic promotes angiogenesis and vascular remodeling in mice. Hedgehog (Hh) signaling is a critical component of maintaining the LSEC phenotype. The loss of this phenotype induces capillarization, which is regulated by Hh signaling [6]. LSEC also serve as the body’s scavenger system. However, during capillarization defenestration and the formation of an organized basal lamella occur, creating an unhealthy condition.

Stable Isotopic Labeling of Amino acids in Cell Culture (SILAC) is used to quantify differences in protein abundance between two cell culture conditions by means of incorporating stable isotopically labeled or “heavy” amino acid(s) (AA) [7]. $^{13}\text{C}_6$ lysine and $^{13}\text{C}_6$ arginine are commonly used to quantify every peptide during tandem mass spectrometry since most digests are performed with trypsin [8]. This concept has been recently extended to mice allowing the quantitative comparison of tissue samples in *in vivo* experiments [9].

Mann’s group previously developed the super-SILAC mix, which is a mixture of several cell lines that serve as an internal spike-in standard for the study of human tumor tissue [10]. The super-SILAC mix greatly improves the quantification accuracy while lowering error rates, and it is a simple, economic, and robust technique. The super-SILAC mix is an integration of cell cultures that differ in origin and cell states and encompass the complexity of the tissue sample. The design and development of the super-SILAC mix is crucial to the quality of the internal standard. The main parameters that determine standard quality are the ratios toward the target tissue, the coverage of the tissue proteome, and the number of orphan peptides that do not have a heavy SILAC partner. Low SILAC ratios (<5-fold) ensures accurate quantification, and high coverage ensures representation of biologically relevant proteins.

The design of a typical super-SILAC for tumor quantification involves the selection of 3 to 7 cell lines representing the tissue type, preferably cell lines that are diverse but are similar to the tissue. Selecting fewer cell lines may lead to under-representation of tissue proteins and too many cell lines tend to dilute out one another. Here this work represents a straightforward quantitative proteomics study of *in vivo* arsenic-stressed LSEC by a reverse super-SILAC based method. The aim was to determine proteins that are up- or down-regulated in response to arsenic treatment.

2. Material and methods

Animals

Animal exposures were performed in agreement with institutional guidelines for animal safety and welfare. C57BL/6N male mice, ages 6–8 weeks weighing approximately 20 g were obtained from the institute of experimental animals, Chinese academy of medical sciences. Standard mouse chow and drinking water were fed *ad libitum* for 5 weeks to mice housed in cages of 3 animals. Fresh drinking water solutions containing 250 ppb sodium arsenite (Sodium (meta)arsenite; Sigma, St. Louis, MO) were prepared every third week in drinking water. The mice were fasted for ~16 h before sacrificed. All procedures were conducted as previously described [11]. For SILAC studies mice were fed ^{13}C -Lys-specific labelled diet (Silantes; Munich, Germany). The corresponding livers were obtained from two 12-week-old females with 99.48% incorporation rate of heavy lysine. SILAC tissues were pooled before processing. All tissue samples were stored at -80°C until use.

Mouse liver perfusion and fixation

Livers were cleared of blood by perfusion with 30 mL of PBS (1 mL/min) through the inferior vena cava, and perfusion fixed with 3 mL 2.5% glutaraldehyde in PBS. Livers were removed and immersed in 2.5% glutaraldehyde for 2 days at 4°C . Samples were processed for transmission electron microscopy [1]. Briefly, several 1 mm^3 tissue blocks were washed in PBS, and post-fixed in aqueous 1% OsO_4 and 1% $\text{K}_3\text{Fe}(\text{CN})_6$ for 1 h. After several PBS washes, the blocks were dehydrated through a graded series of 70% to 100% ethanol, 100% propylene oxide, and infiltrated for 1 hour in a 1:1 mixture of propylene oxide: Polybed 812 epoxy resin (Polysciences, Warrington, PA). After several changes of 100% resin over 24 hours, the blocks were embedded in molds, cured at 37°C overnight, and then hardened for 2 days at 65°C .

Digestion

Tissues were disrupted by either grinding in a liquid nitrogen-cooled mortar, sonication, shearing, or homogenization. Proteins were then solubilized with repeated sonication in SDS sample buffer (1% SDS, 100 mM Tris-HCl (pH 9.5)). After protein solubilization, extracts were centrifuged at $40,000 \times g$ for 1 h at 15°C . Protein extracts were used or stored in aliquots at -80°C . For improved cell lysis, the samples were sonicated on ice, and then were incubated for 30 min at room temperature with repeated vortexing. Protein content was determined using the BCA Protein Assay Kit (ThermoFisher, Rockford, IL). For normal and SILAC labeled mouse liver, 30 μg protein from each were mixed with an identical amount of the corresponding SILAC labeled standard. The DTT was then added to a final concentration of 0.1 M followed by incubation at 75°C for 5 min. Proteins were separated using a 5–12% SDS-PAGE. Each gel slice was diced into small pieces (1 mm^2) and placed into a 0.65 mL siliconized tube. Next, $\sim 100\text{ }\mu\text{L}$ of 25mM NH_4HCO_3 /50% ACN was added to each tube and vortexed for 10 min. Using gel loading pipet tip, the supernatant was removed and discarded. The gel pieces were completely dried ($\sim 20\text{ min}$) using a Speed Vac. DTT (25 μL ; 10 mM prepared in 25 mM NH_4HCO_3) was added to dried gels, vortexed, spun briefly, and allowed the reaction to proceed at 56°C for 30 min. The supernatant was discarded and 25 μL of 55 mM iodoacetamide was added to the gel pieces, vortexed, and spun briefly. The

reaction was allowed to proceed in the dark for 45 min at room temperature, and supernatant was discarded. The gels were washed with ~100 μ L NH_4HCO_3 , vortexed 10 min, spun, and discard the supernatant. Gels were dehydrated with ~100 μ L of 25 mM NH_4HCO_3 in 50% ACN, vortexed 5 min, spun and repeated once. The samples were incubated with a trypsin solution for 15 h. The digestion was stopped by addition of formic acid (3%), and the organic solvent was removed in a SpeedVac concentrator. The obtained peptides were acidified with trifluoroacetic acid. Peptide mixtures were measured both directly and after fractionation into nine fractions via anion exchange (Anion Exchange Columns; 1–12 ml) chromatography.

LC-MS/MS

Dissolved peptides were separated on a 50-cm column (75 μ m id) packed with PepMap R RSLC C18 reverse-phase material. Samples were loaded using a UPLC system (EASYnLC 1000, Fisher Scientific). Eluted peptides were introduced into an Orbitrap Elite Mass Spectrometer (Fisher Scientific) using a nanoelectrospray ion source (EASY-SPRAY, Fisher Scientific). A linear gradient of 0–30% ACN in 0.1% formic acid over 240 min was used at a flow rate of 250 nL/min. The mass spectra were obtained using one full MS scan (400–1600 m/z) in the Orbitrap mass analyzer, followed by 20 data-dependent MS/MS scans. For MS scans, the automatic gain control target was 500,000 with a maximum ion injection time of 120 ms, and the resolution was 240,000 (full width half-maximum) at m/z 400. For MS/MS spectra, up to 20 most intense ions were selected and fragmented by CID with normalized collision energies of 30. Automatic gain control target for MS/MS was 10,000 with a maximum injection time of 50 ms. A minimum ion intensity of 1,000 was required to trigger an MS/MS spectrum. Dynamic exclusion duration was set at 30 s.

Data analysis

The raw files were analyzed by MaxQuant [12] (version 1.3.0.5). Peaks were searched against the UniProt human database (released July 2017; <http://www.uniprot.org>) using the Andromeda search engine included in MaxQuant [13]. Two miscleavages were allowed and at least six amino acids per identified peptide were required. The false discovery rate (FDR) for peptides and proteins was set at 0.01. “Match between runs” option was selected within a time window of 2 min. A minimum ratio count of 2 was used for quantification of SILAC pairs. The MS proteomics data is deposited to the Proteome Xchange Consortium (<http://proteomecentral.proteomexchange.org>) via the PRIDE partner repository with the dataset identifier PXD000438. Further data analysis was performed using Perseus software (version 1.3.0.4; <http://www.perseus-framework.org>). Unsupervised hierarchical clustering was performed based on Euclidean distance with average linkage. DAVID Bioinformatics Resources 6.7 (<http://david.abcc.ncifcrf.gov/>) was used for the GO analysis. For the analysis of the set of 12 xenografts by spectrum counts, peptide separations were performed using a modified five-step Multidimensional Protein Identification Technology [14]. Raw data were converted to m/z XML by using Re AdW and searched using X! Tandem (v.2008.02.01.1) against the human UniPort protein sequence database and employing a target-decoy strategy to establish a FDR [15]. Fragment ion mass tolerance was 0.4 Da, and 10 ppm for parent ions. A peptide quality control strategy was applied to minimize false-positive identifications

and protein inferences. Protein quantification by spectrum counts and maintaining a FDR <0.05 [16].

Circular representation

The chromosomal distributions of liver proteins and the interactions among those proteins were visualized using Circos software [17], which is widely used to visualize genomic data and their features.

Protein-protein interaction network

Functional connections among the 661 differentially expressed proteins were analyzed using the multiple sequences module from the STRING-DB facilities (<http://string-db.org/>) with the high confidence score (0.700) used as edges to construct the network. Unsupervised hierarchical clustering (UHC) within 86 significant proteins filtered based on p-value as determined by One-way ANOVA ($P < 0.05$). UHC was performed using Euclidean distance and average linkage. Cluster analysis was performed using Cluster software, version 3.0, and data was visualized using TreeView software, version 1.1.6. Clustering of 86 significant proteins was generated by unsupervised hierarchical clustering using Euclidean distance and average linkage.

3. Results and discussion

3.1. Schematic of the workflow

Fig. 1 provides a schematic outline of the experimental approach used to compare normal and arsenic treated LSEC in mice aged 11 weeks, using stable isotope labeling of whole animals (SILAC mouse) and high resolution mass spectrometry. Primary LSEC lysates from each of four mice aged 11 weeks were analyzed in separate experiments and mixed each time with an identical standard derived from SILAC mice aged 12 weeks. Proteins were digested using trypsin and fractionated before analysis by high resolution mass spectrometry. Proteins were quantified by dividing the individual SILAC peptide ratios of the proteomes to be compared (“ratio of ratios”).

3.2. Examination of SILAC mouse labeling efficiency

For incorporation testing SILAC-mouse labeling efficiency, SILAC-mouse liver were lysed and digested using standard protocols. Protein amount (10–50 μg) is sufficient for the test. After digestion, peptides were purified on StageTips and analyzed with a single LC-MS/MS run. MaxQuant analysis of the raw file determined the ratio of heavy labeled peptides to the remaining non-labeled ones. The incorporation efficiency can be calculated as $(1 - 1/\text{Ratio (H/L)})$ on the peptide level. Since arginine and lysine may have different labeling efficiencies, it is recommended to perform the calculations separately for each peptide. For proper SILAC experiments, it is necessary to obtain labeling $>95\%$. To examine arginine to proline conversion, heavy proline can be added as a variable modification in the MaxQuant analysis (Fig. 2).

3.3. Protein identification and quantification

All raw files were converted to mzXML and MGF files using the Msconvert module [12] in Trans-Proteomic Pipeline (TPP v4.5.2) [13]. The MS/MS peak lists were searched using the local Mascot v2.3.2 server [14] against the database containing sequences of all mouse proteins from Uniprot database (50,517 protein entries, release 2012_12, and download from <ftp.uniprot.org/pub/databases/uniprot/>) and common contaminant protein sequences (115 proteins, download from <ftp.thegpm.org/fasta/cRAP>). The Decoy checkbox were chosen to perform an automatic decoy database search by Mascot. The monoisotopic mass was used for both peptide and fragment ions with fixed modification (Carbamidomethylation, +57.0214 Da) on cysteine and variable modifications (Oxidation, +15.9949 Da) on methionine. Specific cleavages were selected for trypsin and lysC precedence trypsin (Lys and Arg) as well as for LysC (Lys), and up to 2 missed cleavage sites were allowed. The precursor and fragment ion mass tolerance is 20 ppm and 0.5 Da for this hybrid linear trap high-accuracy data. PepDistiller [9] was used for quality control to facilitate the validation of Mascot search results. Peptides length shorter than seven amino acids were removed and all peptide-spectra matches were filtered by the probability value of PepDistiller to keep the FDR measured by the decoy hits <1%.

The peptide and protein quantification was assessed using SILVER [18], a program developed by Beijing Proteome Research Center (BPRC), which can be downloaded from <http://bioinfo.hupo.org.cn/silver/download.php>. The proteins with significant differences in expression were selected by a normal distribution estimation on $\log_2$ (ratio) of proteins. This method is based on the hypothesis that the majority of the protein expressions are not statistically significant events, and the $\log_2$ (ratio) is centered on zeros with a normal distribution. The 3-sigma rule of thumb was used to select the proteins with significant different expression, which means the risk of the proteins with no biological significance was treated as ones with different expression that was <0.0027 (Fig. 3).

3.4. Gene Ontology, protein-protein interaction and disease association analysis of differentially expressed proteins upon arsenic treatment

DAVID Bioinformatics Resources 6.7 (<http://david.abcc.ncifcrf.gov/>) was used for GO analysis. For analysis of the set of 12 xenografts by spectrum counts, peptide separations were performed using a modified five-step Multidimensional Protein Identification Technology [12]. Cellular component distribution of the identified proteins and/or differentially expressed proteins is shown in Fig. 4A. The biological process distribution of the identified proteins/differentially expressed proteins are presented in Fig. 4B. The molecular function distribution of the identified proteins/differentially expressed proteins is listed in Fig. 4C. Top five biological processes, cell components, and molecular function of up-regulated proteins are shown in Fig. 4D. DAVID v6.7 [13] was used to assemble the functional annotation of the proteins, including the biological processes, cellular components and molecular functions (Gene Ontology; [17]). The biological processes, cellular components, molecular functional distribution of the identified proteins and the network distribution of identified proteins was also computed using David based on Gene Ontology. The enrichment p-values in tables were calculated based on EASE Score, a

modified Fisher Exact Test. The terms with p-value <0.01 were considered as significantly enriched.

To confirm the physiological relevance of protein expression changes, we characterized the proteins with the largest increase or decrease following treatment with arsenic (Fig. 4E). Gene ontology analysis of the biological processes, enriched in the proteins increased by more than 50% using DAVID16, identified processes mainly associated with metabolic pathways such as oxydoreduction (p-value of $29\text{e-}19$) and carbohydrate catabolic processes (p-value of $6.60\text{e-}10$) as well as several processes implicated in the metabolism. Circular representation of the genome features including the genome sequences (ring 1) and gene symbols of interacting proteins (ring 2) are shown (Fig. 4F). Gene symbols with red and bold font are liver fibrosis and liver cirrhosis-related genes. Interacting proteins are linked by edges, red edges indicate interacting proteins on the same chromosome and green edges indicate interacting proteins on different chromosomes. The image was created using the Circos software.

We next determined the interactions among the 661 candidate proteins. Interaction maps generated using STRING database (Fig. 4G). Circular representation of the genome features of 661 candidate proteins were also determined. Genome sequences (ring1), protein symbols (ring2) and interacting proteins (ring 3) are shown. Interacting proteins are linked by edges; blue edges indicate interacting proteins on different chromosomes, and red edges indicate interacting proteins on the same chromosome. Protein symbols with red color also indicate interacting proteins on same chromosome. This image was created using the Circos software (Fig. 4H).

The super-SILAC approach is scalable and flexible, allowing the generation of reference libraries of SIL peptides that can be applied over the duration of a lengthy biomarker discovery campaign, spanning different tissue types and sample sources. Improved quantification of complex tissue proteomic samples in the discovery phase could substantially improve confidence in the identification of differentially expressed proteins, effectively triaging the long lists of candidate biomarkers requiring validation.

Spiking in a whole proteome's worth of SIL peptides brings new analytical challenges. The combined super-SILAC and liver sinusoidal cell's proteome mixture has at least doubled in complexity, and the dynamic range of accurate peptide quantification may not span the full range of analytes of interest. Indeed, the whole-proteome SIL standard is unlikely to be useful in the validation phase of biomarker discovery. Greater quantitative accuracy, afforded by the use of a super-SILAC proteome standard or other means, will undoubtedly improve the quality of tissue protein expression profiles and our ability to confidently identify subtle changes in protein expression. Widespread use of whole-proteome SIL standards may provide a framework, similar to approaches commonly used in gene expression profiling, to standardize quantitative analyses of complex tissue samples in clinical proteomics.

The ability to robustly compare different clinical proteomics datasets would facilitate the integration of datasets from proteomics and genomics, and transform the field of clinical proteomics. At the same time, we are seeing a general shift in metabolism with arsenic that

relates to more lipid utilization and possible insulin resistant metabolism. This is not unusual for arsenic, but the proteins are interesting for their close connection with lipid conjugation reactions by the change in fenestrations. Among the novel capillarization related proteins, which we found, the N-acetyltransferase is interesting and it would be very interesting to determine whether we observe a change in lysine acetylation, as well as the change in protein abundance. For both HMGCoA synthase and the N-acetyltransferase, it might be that oxidative stress is affecting activity through post-translational modification, and that results in an increased need for more enzyme. A change in lysine acetylation status might indicate that arsenic affecting sirtuin activity that regulates de-acetylation in response to metabolic shifts.

Acknowledgments

Funding

The work in NS laboratory is supported by an unrestricted departmental award from Research to Prevent Blindness, RRF, P30 EY016665, P30 CA014520, EPA 83573701, R24 EY022883, and R01 EY026078. NS is a recipient of Stein Innovation Award.

References

- [1]. Li W, Ding W, Ji G, Wang L, Zhang J, Sun F, Three-dimensional visualization of arsenic stimulated mouse liver sinusoidal by FIB-SEM approach, *Protein Cell*, 7 (2016) 227–232. [PubMed: 26856874]
- [2]. Lee YA, Wallace MC, Friedman SL, Pathobiology of liver fibrosis: a translational success story, *Gut*, 64 (2015) 830–841. [PubMed: 25681399]
- [3]. Miyao M, Ozeki M, Abiru H, Manabe S, Kotani H, Tsuruyama T, Tamaki K, Bile canalicular abnormalities in the early phase of a mouse model of sclerosing cholangitis, *Digestive and liver disease : official journal of the Italian Society of Gastroenterology and the Italian Association for the Study of the Liver*, 45 (2013) 216–225.
- [4]. Xie G, Wang X, Wang L, Wang L, Atkinson RD, Kanel GC, Gaarde WA, Deleve LD, Role of differentiation of liver sinusoidal endothelial cells in progression and regression of hepatic fibrosis in rats, *Gastroenterology*, 142 (2012) 918–927.e916. [PubMed: 22178212]
- [5]. Deleve LD, Wang X, Guo Y, Sinusoidal endothelial cells prevent rat stellate cell activation and promote reversion to quiescence, *Hepatology*, 48 (2008) 920–930. [PubMed: 18613151]
- [6]. Omenetti A, Diehl AM, Hedgehog signaling in cholangiocytes, *Current opinion in gastroenterology*, 27 (2011) 268–275. [PubMed: 21423008]
- [7]. Ong SE, Blagoev B, Kratchmarova I, Kristensen DB, Steen H, Pandey A, Mann M, Stable isotope labeling by amino acids in cell culture, SILAC, as a simple and accurate approach to expression proteomics, *Mol Cell Proteomics*, 1 (2002) 376–386. [PubMed: 12118079]
- [8]. Kruger M, Moser M, Ussar S, Thievensen I, Luber CA, Forner F, Schmidt S, Zanivan S, Fassler R, Mann M, SILAC mouse for quantitative proteomics uncovers kindlin-3 as an essential factor for red blood cell function, *Cell*, 134 (2008) 353–364. [PubMed: 18662549]
- [9]. Ma J, Li W, Lv Y, Chang C, Wu S, Song L, Ding C, Wei H, He F, Jiang Y, Zhu Y, A new insight into the impact of different proteases on SILAC quantitative proteome of the mouse liver, *Proteomics*, 13 (2013) 2238–2242. [PubMed: 23703833]
- [10]. Geiger T, Cox J, Ostasiewicz P, Wisniewski JR, Mann M, Super-SILAC mix for quantitative proteomics of human tumor tissue, *Nature methods*, 7 (2010) 383–385. [PubMed: 20364148]
- [11]. Straub AC, Stolz DB, Ross MA, Hernandez-Zavala A, Soucy NV, Klei LR, Barchowsky A, Arsenic stimulates sinusoidal endothelial cell capillarization and vessel remodeling in mouse liver, *Hepatology*, 45 (2007) 205–212. [PubMed: 17187425]

- [12]. Park SK, Venable JD, Xu T, Yates JR 3rd, A quantitative analysis software tool for mass spectrometry-based proteomics, *Nature methods*, 5 (2008) 319–322. [PubMed: 18345006]
- [13]. Ma J, Zhang J, Wu S, Li D, Zhu Y, He F, Improving the sensitivity of MASCOT search results validation by combining new features with Bayesian nonparametric model, *Proteomics*, 10 (2010) 4293–4300. [PubMed: 21086516]
- [14]. Huang da W, Sherman BT, Lempicki RA, Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources, *Nat Protoc*, 4 (2009) 44–57. [PubMed: 19131956]
- [15]. Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, Davis AP, Dolinski K, Dwight SS, Eppig JT, Harris MA, Hill DP, Issel-Tarver L, Kasarskis A, Lewis S, Matese JC, Richardson JE, Ringwald M, Rubin GM, Sherlock G, Gene ontology: tool for the unification of biology. The Gene Ontology Consortium, *Nat. Genet*, 25 (2000) 25–29. [PubMed: 10802651]
- [16]. Picotti P, Rinner O, Stallmach R, Dautel F, Farrah T, Domon B, Wenschuh H, Aebersold R, High-throughput generation of selected reaction-monitoring assays for proteins and proteomes, *Nature methods*, 7 (2010) 43–46. [PubMed: 19966807]
- [17]. Krzywinski M, Schein J, Birol I, Connors J, Gascoyne R, Horsman D, Jones SJ, Marra MA, Circos: an information aesthetic for comparative genomics, *Genome Res*, 19 (2009) 1639–1645. [PubMed: 19541911]
- [18]. Chang C, Zhang J, Han M, Ma J, Zhang W, Wu S, Liu K, Xie H, He F, Zhu Y, SILVER: an efficient tool for stable isotope labeling LC-MS data quantitative analysis with quality control methods, *Bioinformatics*, 30 (2014) 586–587. [PubMed: 24344194]
- [19]. Ma J, Zhang J, Wu S, Li D, Zhu Y, He F, Improving the sensitivity of MASCOT search results validation by combining new features with Bayesian nonparametric model, *PROTEOMICS*, 10 (2010) 4293–4300. [PubMed: 21086516]

High lights for review

- There is an urgent need to understand how various diseases, toxic substances, and physiological conditions influence liver endothelial cell fenestrations, and how these changes affects liver function.
- This work represents a straightforward quantitative proteomics study of the in vivo arsenic-stressed liver sinusoidal endothelial cells using a reverse super-SILAC based method.
- The results presented will aid in identification of potential targets and mechanisms of arsenic toxicity and novel ways to reverse these changes.

Fig. 1. Schematic of the workflow.

Primary LSEC lysates from each of six mice aged 11 weeks were analyzed in separate experiments, and mixed each time with an identical standard derived from SILAC mice aged 12 weeks. Proteins were digested using trypsin and fractionated before analysis by high-resolution mass spectrometry. Proteins were quantified by dividing the individual SILAC peptide ratios of the proteomes to be compared (“ratio of ratios”).

Fig. 2. The probability density function (PDF) of parent (M+H) ion + error (ppm) and scatter plot for score and delta score.

(A) The probability density function (PDF) of parent (M+H) ion + error (ppm) was plotted using the peptide spectrum match (PSM) passing through the initial filtration (FDR < 0.01, without parent error requirement). The distribution is asymmetrical (around 0). Thus, the lower bound used in our database search result filtration is -7.066 (the mean -3*stand deviation) and the upper bound is 9.543 (the mean +3*stand deviation). (B) The scatter plot of Mascot score and delta score for target/decoy matches. The decoy matches located on the left bottom corner and the positive matches located on the right and top corner. As shown in the figure, we cannot give a simple boundary to distinguish the target and decoy PSMs. In order to improve the sensitivity of the positive PSMs selection, both the database search score and the delta score were used to select the positive PSMs and reject the false matches given an FDR [19].

Fig. 3. The distribution of ratio values for all proteins identified.

Proteins with a ratio larger than 4.7425 are up-regulated while those with a ratio smaller than -5.2213 are down-regulated.

Fig. 4. Gene Ontology, protein-protein interaction and disease association analysis of differentially expressed proteins upon arsenic treatment.

DAVID Bioinformatics Resources 6.7 (<http://david.abcc.ncifcrf.gov/>) was used for the GO analysis. For the analysis of the set of 12 xenografts by spectrum counts, peptide separations were performed using a modified five-step Multidimensional Protein Identification Technology as previously reported²⁴ Cellular component distribution of the identified proteins / differential expressed proteins (**A**); Biological process distribution of the identified proteins / differential expressed proteins (**B**); Molecular function distribution of the

identified proteins / differential expressed proteins (**C, D**). Top five biological processes, cell components and molecular functions of up-regulated proteins (**E**). (**F**) Circular representation of the genome features. Genome sequences (ring 1) and gene symbols of interacting proteins (ring 2) are shown. Gene symbols with red and bold font are liver fibrosis and liver cirrhosis-related genes. Interacting proteins are linked by edges, red edges indicate interacting proteins on the same chromosome and green edges indicate interacting proteins on different chromosomes. The image was created by using the software Circos. (**G**) Interactions between proteins among 661 candidate proteins. Interaction map generated using STRING database. (**H**) Circular representation of the genome features of 661 candidate proteins. Genome sequences (ring1), protein symbols (ring2) and interacting proteins (ring3) are shown. Interacting proteins are linked by edges, blue edges indicate interact proteins on different chromosomes and red edges indicate interacting proteins on the same chromosome. Protein symbols with red color indicate interact proteins on the same chromosomes. The image was created using the software Circos.