

RESEARCH ARTICLE

Open Access

# Comparative genomic analyses identify common molecular pathways modulated upon exposure to low doses of arsenic and cadmium

Margaret Ann Benton<sup>†</sup>, Julia E Rager<sup>†</sup>, Lisa Smeester and Rebecca C Fry<sup>\*</sup>

## Abstract

**Background:** Exposure to the toxic metals arsenic and cadmium is associated with detrimental health effects including cancers of various organs. While arsenic and cadmium are well known to cause adverse health effects at high doses, the molecular impact resulting from exposure to environmentally relevant doses of these metals remains largely unexplored.

**Results:** In this study, we examined the effects of *in vitro* exposure to either arsenic or cadmium in human TK6 lymphoblastoid cells using genomics and systems level pathway mapping approaches. A total of 167 genes with differential expression were identified following exposure to either metal with surprisingly no overlap between the two. Real-time PCR was used to confirm target gene expression changes. The gene sets were overlaid onto protein-protein interaction maps to identify metal-induced transcriptional networks. Interestingly, both metal-induced networks were significantly enriched for proteins involved in common biological processes such as tumorigenesis, inflammation, and cell signaling. These findings were further supported by gene set enrichment analysis.

**Conclusions:** This study is the first to compare the transcriptional responses induced by low dose exposure to cadmium and arsenic in human lymphoblastoid cells. These results highlight that even at low levels of exposure both metals can dramatically influence the expression of important cellular pathways.

## Background

Arsenic and cadmium are ranked among the top ten priority hazardous substances by the Agency for Toxic Substances and Disease Registry (ATSDR) [1]. Exposure to arsenic and cadmium can lead to adverse health outcomes such as lung and kidney cancers as well as cardiovascular disease and diabetes [2,3]. Further, exposure to these two toxic and well-classified chemicals is of particular interest because of their extensive global impact [4-6]. For example, it is estimated that more than 40 million people worldwide drink water containing arsenic at concentrations that exceed the World Health Organization (WHO) and Environmental Protection Agency (EPA) drinking water guideline of 10 ppb [7]. Also, humans are exposed to low levels of cadmium

through food consumption, typically ranging between 8 and 25 ug per day [3]. Smoking populations experience higher levels of cadmium exposure, as one cigarette may contain 1-2 ug cadmium [3].

Both arsenic and cadmium are classified as Group 1 carcinogens by the International Agency for Research on Cancer (IARC) [4,5]. Arsenic exposure has been linked to several types of cancer, including skin, lung, liver, and bladder [5,8]. Proposed mechanisms of arsenic-induced disease include oxidative stress, DNA repair inhibition, and epigenetic events [8]. Likewise, cadmium exposure has been associated with various cancers, such as prostate, kidney, pancreas, and lung [3]. The probable mechanisms of cadmium carcinogenesis are similar to those of arsenic and include aberrant gene expression, inhibition of DNA damage repair and apoptosis, and oxidative stress [9]. Studies suggest that even at low levels, chronic exposure to arsenic and cadmium is associated with increased risk for disease including

\* Correspondence: rfry@unc.edu

† Contributed equally

Department of Environmental Sciences and Engineering, Gillings School of Global Public Health, University of North Carolina, Chapel Hill, NC, USA

cancer [10,11]. However, the exact mechanisms that associate low level exposures of arsenic and cadmium with many of these negative health outcomes remain largely unknown.

In this study, we set out to contrast the cellular responses of human TK6 lymphoblastoid cells upon exposure to low, environmentally relevant doses of arsenic and cadmium. The rationale for this research was based on studies suggesting that differential gene expression occurs in human cells following exposure to low levels of either arsenic [12-15] or cadmium [16]. Here we characterized the effects of low dose exposure to either arsenic or cadmium by examining changes in the expression of genes and their associated biological pathways and functions. We found that while the modified gene sets were distinct for each of the two metals, similar biological pathways were modulated between the two. Future research will extend these findings to identify modulated protein activity associated with such low dose exposure.

## Results

### Metal-induced gene sets identified

In this study, human TK6 lymphoblastoid cells were exposed to either arsenic (sodium arsenite) or cadmium (cadmium chloride) at low, equimolar doses (0.1  $\mu\text{M}$ ) for 24 hours (see Methods). These metal concentrations were minimally cytotoxic with 99% cell survival for both arsenic and cadmium (Table 1). RNA was extracted from metal-exposed or mock-treated control cells and hybridized to Affymetrix GeneChip<sup>®</sup> Human Gene 1.0 ST arrays (see Methods). Differentially expressed genes were identified using an ANOVA model (see Methods). A total of 167 genes (209 probesets) were significantly differentially expressed in the metal-exposed cells; 62 of these genes were dysregulated in the arsenic-exposed

cells and 105 in the cadmium exposed cells (Table 1 see Additional File 1: Arsenic-modulated genes, Additional File 2: Cadmium-modulated genes). It should be noted that the two metal-induced gene sets were distinct with no overlap between them.

### Metal-induced molecular networks enriched for numerous biological processes

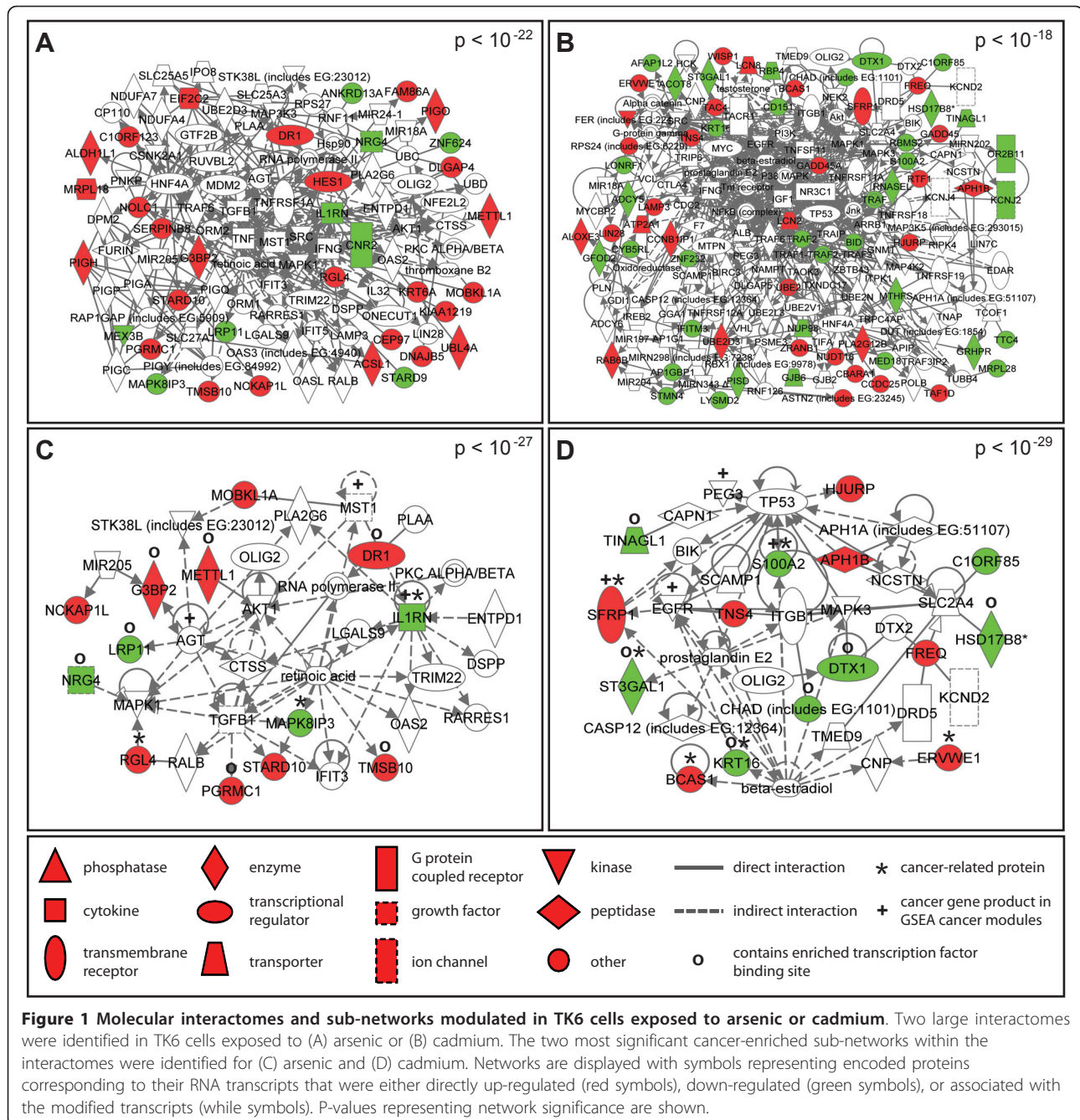
To identify molecular networks associated with low dose arsenic or cadmium exposure, we analyzed the metal-modulated gene sets for known interactions among their encoded proteins (see Methods). For the arsenic gene set, 41 of the 62 differentially expressed genes were eligible for network generation (see Additional File 3: Gene products in networks). For the cadmium gene set, a total of 71 of the 105 differentially expressed genes were eligible for network generation (see Additional File 3: Gene products in networks).

Through network analysis we identified that exposure to either arsenic or cadmium exposure results in the modulation of large interacting protein networks (e.g. interactomes). The arsenic-induced interactome was comprised of a total of 103 proteins, 35 of which were encoded by transcripts that were differentially expressed following exposure ( $p < 10^{-22}$ , Figure 1A, see Additional File 3: Gene products in networks). Within this arsenic-modulated interactome, 26 genes showed increased expression and 9 genes showed decreased expression as a result of arsenic exposure (Figure 1A, see Additional File 3: Gene products in networks).

The cadmium-modulated interactome contained a total of 167 proteins with significant interactions ( $p < 10^{-18}$ ); 64 of these proteins were encoded by transcripts that were differentially expressed following exposure (Figure 1B, see Additional File 3: Gene products in networks). Within this large cadmium-modulated interactome, the

**Table 1 Summary of biological responses of TK6 cells exposed to arsenic or cadmium**

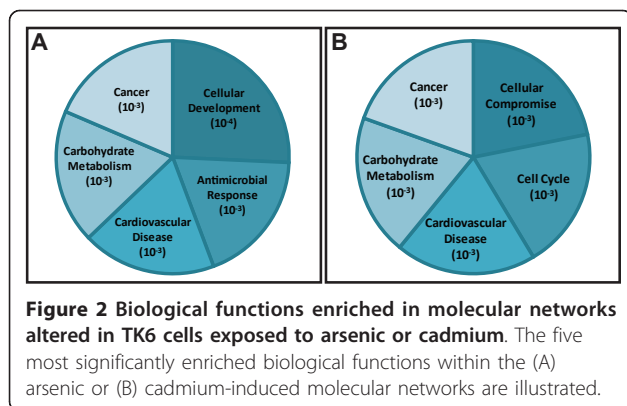
	<b>Arsenic (0.1 <math>\mu\text{M}</math>)</b>	<b>Cadmium (0.1 <math>\mu\text{M}</math>)</b>
Percent Survival	99%	99%
Genes Differentially Expressed (probesets)	62 (74)	105 (135)
Gene Set Enrichment (FWER)	Cancer Module 88 (Liver, heart, lung) (0.33)	Cancer Module 88 (Liver, heart, lung) (0.49)
	Cancer Module 55 (Liver, lung, breast) (0.33)	Cancer Module 55 (Liver, lung, breast) (0.49)
		Cancer Module 23 (Liver, lung, prostate) (0.49)
		Cancer Module 6 (Lung, prostate, liver) (0.34)
Transcription Factor Binding Site Enrichment (p-value)	RBP-Jkappa (0.001)	ZF5 (0.002)
	Oct-1 (0.002)	Sp1 (0.011)
	E2F (0.015)	SREBP-1 (0.016)



expression of 29 genes was up-regulated and the expression of 35 genes was down-regulated (Figure 1B, see Additional File 3: Gene products in networks).

Within the large cadmium and arsenic-induced interactomes were smaller, more focused sub-networks. The arsenic interactome was comprised of three sub-networks with p values ranging from  $10^{-27}$  to  $10^{-22}$  (Figure 1C, see Additional File 4: Metal modulated sub-networks). These arsenic-modulated sub-networks were enriched for biological processes such as cellular

development ( $p < 10^{-4}$ ), antimicrobial response ( $p < 10^{-3}$ ), carbohydrate metabolism ( $p < 10^{-3}$ ), cardiovascular disease ( $p < 10^{-3}$ ), and cancer ( $p < 10^{-3}$ ) (Figure 2A). Contained within the large interactome were genes with known associations with cancer development and progression. These included interleukin 1 receptor antagonist (*IL1RN*), mitogen-activated protein kinase 8 interacting protein 3 (*MAPK8IP3*), ral guanine nucleotide dissociation stimulator-like 4 (*RGL4*), cannabinoid receptor 2 (*CNR2*) and lin-28 (*LIN28*).



As was seen in response to arsenic, we identified significant sub-networks in the cadmium-modulated interactome with p values ranging from  $10^{-31}$  to  $10^{-18}$  (Figure 1D, see Additional File 4: Metal modulated sub-networks). The five sub-networks are enriched for processes involving cellular compromise ( $p < 10^{-3}$ ), carbohydrate metabolism ( $p < 10^{-3}$ ), cardiovascular disease ( $p < 10^{-3}$ ), cell cycle functions ( $p < 10^{-3}$ ), and cancer ( $p < 10^{-3}$ ) (Figure 2B). In the cadmium-modulated interactome and within the significant sub-networks, many proteins with known roles in tumorigenesis were identified. Such proteins include breast carcinoma amplified sequence 1 (BCAS1), endogenous retroviral family W, env(C7), member 1 (ERVWE1), keratin 16 (KRT16), S100 calcium binding protein A2 (S100A2), secreted frizzled-related protein 1 (SFRP1), and ST3 beta-galactoside alpha-2,3-sialyltransferase 1 (ST3GAL1), and actin filament associated protein 1-like 2 (AFAP1L2). These analyses also show the enrichment of transcripts that encode proteins involved in pathways regulated through nuclear factor kappa-light-chain-enhancer of activated B cells (NF- $\kappa$ B) and p38 mitogen-activated protein kinase (p38 MAPK) (see Additional File 4: Metal modulated sub-networks, Figure E).

While the arsenic and cadmium-induced gene sets shared no genes in common, we set out to establish whether the metal-modulated genes map to common pathways. For this, we combined the arsenic and cadmium gene lists and analyzed the resulting metal-modulated molecular networks. Interestingly, the networks showed overlap with both arsenic and cadmium-associated gene products mapping to common pathways including the p38 mitogen-activated protein kinase (p38 MAPK) and hepatocyte nuclear factor 4 (HNF-4) pathways (see Additional File 4: Metal modulated sub-networks).

#### Cancer modules enriched in the metal-modulated gene sets

As an alternative approach to identify biological pathways that are modulated upon exposure to arsenic or

cadmium, we performed gene set enrichment analysis (see Methods). We compared the metal-modulated gene sets to genes known to be dysregulated under certain physiological and environmental conditions (e.g. cancer modules) [17]. These results showed significant enrichment for two cancer modules common to both arsenic and cadmium: cancer modules 88 and 55 which include gene sets associated with liver, lung, heart, breast, and prostate cancer (Table 1). In addition, the cadmium-modulated gene set was uniquely enriched for cancer modules 6 and 23 (Table 1).

#### Transcription factors identified that may mediate metal-induced transcriptional responses

To identify regulatory mechanisms that potentially underlie the metal-modulated transcript levels, we investigated whether binding sites for specific transcription factors were computationally enriched in the promoter regions of the metal-modulated gene sets (see Methods). In the arsenic-induced gene set, computational analysis of the promoter regions of the differentially expressed genes identified a significant enrichment for the following two transcription factors: EF2 transcription factor (E2F) ( $p < 0.015$ ), octamer-1 transcription factor (Oct-1) ( $p < 0.002$ ) and RBP-Jkappa transcription factor ( $p < 0.001$ ) (Table 1). In the cadmium-modulated gene set, the enriched transcription factors included sp1 transcription factor (Sp1) ( $p < 0.011$ ), sterol regulatory element binding transcription factor 1 (SREBP-1) ( $p < 0.016$ ), and zinc finger protein 161 homolog (ZF5) ( $p < 0.002$ ) (Table 1). A comparison of the predicted gene targets of these transcription factors shows the enrichment for their putative activity in the cancer-associated proteins in the cadmium sub-network. Specifically, ST3 beta-galactoside alpha-2,3-sialyltransferase 1 (ST3GAL1) is predicted to be regulated by ZF5 and Sp1. Likewise, keratin 16 (KRT16) is predicted to be regulated by Sp1 (see Additional File 5 Figure 1).

#### Microarray gene expression results validated using qRT-PCR

Within each of the metal-modulated interactomes, target genes were selected and analyzed for their expression levels using quantitative real-time PCR (qRT-PCR). Selected from the arsenic-modulated networks, qRT-PCR analysis showed that StAR-related lipid transfer (START) domain containing 9 (STARD9) was down-regulated by arsenic exposure (-1.31 fold change (FC)), and methyltransferase like 1 (METTL1) was up-regulated (+1.46 FC). Likewise, qRT-PCR analysis showed that cannabinoid receptor 2 (CNR2) was down-regulated (-1.26 FC). A comparison of the fold changes obtained with qRT-PCR with those obtained by microarray shows a high correlation (0.87) (see Additional File 6).

Selected from the cadmium interactome, qRT-PCR showed that actin filament associated protein 1-like 2 (*AFAPIL2*) was down-regulated (-0.76 FC), and lin-28 homolog (*LIN28*) was up-regulated (+4.0 FC) following low dose cadmium exposure. qRT-PCR analysis showed that S100 calcium binding protein A2 (*S100A2*) was down-regulated (-1.27 FC). In order to confirm that metallothionein genes were not differentially expressed upon exposure to low-dose arsenic or cadmium, metallothionein 1A (*MT1A*) was assessed and found to be insignificant ( $p$  value  $\geq 0.3$ ). A comparison of the fold changes obtained with qRT-PCR with those obtained by microarray shows a high correlation (0.92) (see Additional File 6).

## Discussion

In this study we set out to compare the transcriptional responses and associated biological pathways modulated in TK6 cells exposed to low levels of arsenic or cadmium. Our aim was to contrast the genomic responses of TK6 cells exposed to either of these metals at environmentally relevant doses. Our rationale for the selection of TK6 lymphoblastoid cells is that they are well characterized at the molecular level and have been used to analyze cellular responses upon exposure to a host of environmental agents, including arsenic [18-22]. Using the most recently developed Affymetrix ST arrays that more fully cover open reading frames across the genome, we established the differential gene expression between untreated and metal-exposed TK6 cells. The transcriptional effect of exposure was greater for the cells exposed to cadmium than for arsenic with 105 and 62 genes, respectively. It was interesting to find that there were no genes in common between the two gene sets identified. The microarray gene expression results were validated using real time PCR and a high correlation (0.87-0.92) was found.

Using network analysis, we evaluated the effects of exposure to the two metals by overlaying the differentially expressed genes onto known molecular interaction maps and identified two large interactomes of inter-connected proteins. The most significantly enriched cellular functions within these networks included processes related to cancer, carbohydrate metabolism, and cardiovascular disease. Not surprisingly as their gene sets were distinct, there were also biological functions that were unique to TK6 cells exposed to either arsenic or cadmium. For example, TK6 cells exposed to arsenic modulated biological pathways associated with cellular development and antimicrobial response, whereas TK6 cells exposed to cadmium modulated cellular compromise and cell cycle functions. It is important to note that while the arsenic and cadmium-induced gene sets were distinct with no overlap in genes, the metal-modulated

genes do map to common pathways. Specifically, the arsenic and cadmium-associated gene products both map to the p38 MAPK and HNF-4 pathways.

## Arsenic exposure results in the modulation of transcripts associated with cancer, cell proliferation, cell death, and inflammation

Within the arsenic-modulated interactome, there were gene products with known associations to cancer, cell proliferation, cell death, and inflammation. These include the protein products CNR2, IL1RN, LIN28, MAPK8IP3, and RGL4. *CNR2*, found to be down-regulated through both qRT-PCR and microarray analysis, has several anticancer functions. For example, it inhibits cell proliferation, reduces angiogenesis, reduces cell migration and metastasis, and attenuates inflammatory processes [23]. *IL1RN*, here down-regulated by arsenic exposure, is known to inhibit the activity of IL-1, a major human cytokine implicated in inflammatory responses that contribute to tumor development [24]. Evidence also suggests that *IL1RN* activity may play an important role in the development of cervical squamous cell carcinoma [25]. Further, in a chronically arsenic exposed population, Argos *et al.* found that *IL1RN* was down-regulated among individuals with skin lesions compared to those without lesions [26]. Our microarray results also highlight the up-regulation of *LIN28* further validated through qRT-PCR. *LIN28* has known associations with cancer, as decreased transcription decreases cancer metastasis [27].

## Cadmium exposure results in the alteration of the expression of genes involved in tumorigenesis, cell cycle, apoptosis, and oxidative stress

The cadmium-modulated interactome contains gene products with known involvement in tumorigenesis, cell cycle, apoptosis, and oxidative stress. For example, *AFAPIL2* showed decreased expression levels as assessed through both qRT-PCR and microarray analysis. *AFAPIL2* has been identified as one of six potential gene markers of tumors exhibiting local aggressiveness [28]. Also, *S100A2* showed decreased transcript levels in TK6 cells exposed to cadmium assessed through both qRT-PCR and microarray analysis. *S100A2* is a known tumor suppressor that controls tumor progression and growth [29,30]. *S100A2* differential expression has been observed in cancers such as laryngeal squamous cell carcinoma [31] and non-small cell lung cancer [32]. Another cancer-associated gene, *SFRP1*, is up-regulated in TK6 cells exposed to cadmium. *SFRP1* is an antagonist of Wnt signaling and is involved in apoptosis, cell differentiation, and signal transduction [33,34]. A previous association between the Wnt signaling pathway and cellular responses to cadmium exposure has been

found [35]. *ST3GAL1*, up-regulated by cadmium exposure, is differentially expressed in several types of cancer including bladder [36], ovarian [37], breast [38], and colorectal [39,40]. We note that these genes are involved in cellular mechanisms that control cell growth and proliferation and here they are altered in TK6 cells exposed to low doses of cadmium.

Cadmium exposure has been associated with the activation of oxidative stress pathways [9]. Our findings support that oxidative stress responsive genes and potential regulatory pathways are altered in TK6 cells exposed to low doses of cadmium. Here we find the *S100A2* gene is modulated by exposure to cadmium. *S100A2* is associated with oxidative stress in keratinocytes [41], and is down-regulated by ultraviolet B exposure, possibly in response to free radical activity [42]. In addition, our pathway analysis shows the enrichment for NF- $\kappa$ B and p38 MAP kinase-associated proteins altered by cadmium exposure. Both of these molecules have been shown to regulate genes associated with oxidative stress [43].

#### **Gene Set Enrichment Analysis shows that common cancer gene sets are modulated by environmentally relevant levels of arsenic and cadmium**

We performed Gene Set Enrichment Analysis (GSEA) to determine whether transcripts with common biological functions were differentially expressed in the metal-treated cells. Four cancer-related gene expression signatures (modules) altered by arsenic or cadmium were identified. Two of the signatures were common for both the arsenic and cadmium-induced gene sets. Specifically, cancer modules 88 and 55 were enriched in both gene sets. Cancer module 55 resembles a gene expression profile present in various tumors including liver, lung, B lymphoma, and breast cancer. Cancer module 88 resembles a gene expression profile observed in various tumors including those of the liver, lung and breast. The enrichment of cancer module 6 in the cadmium-induced gene set represents a cancer signature identified in liver, lung, prostate, and breast cancers and also leukemia. The cadmium-modulated genes were similar to those that control liver functions, such as metabolism (module 23), that are associated with liver, lung, prostate, and breast cancer and leukemia and B lymphoma [44]. It is interesting to note again that while no genes were commonly dysregulated between the two metal exposures, here the gene set enrichment analysis identified common cancer-associated modules within both. Thus, this analysis suggests that even low dose environmentally relevant exposures to arsenic or cadmium can induce similar gene expression signatures, reflective of those seen in several types of cancer.

#### **Promoter sites of metal-modulated genes are enriched for transcription factors involved in cell cycle regulation**

To identify transcription factors that may mediate the transcriptional responses in TK6 cells exposed to arsenic or cadmium, we employed computational prediction analysis of transcription factor binding sites for both metal-induced gene sets. Computational analysis was used to identify transcription factors predicted as regulators of the observed molecular response. Analysis of the promoter regions of the arsenic-modulated genes showed enrichment for E2F, Oct-1, and RBP-JKappa. E2F regulates the expression of genes that generate proteins required for cell cycle progression, cell proliferation, apoptosis, and cell differentiation [45,46]. Oct-1 is induced by DNA damaging therapeutic agents through a posttranscriptional mechanism that does not require the tumor suppressor p53, suggesting that Oct-1 may play an important role in p53-independent gene activation to regulate cellular responses to DNA damage [47]. It has been shown that RBP-JKappa is a repressor of NF- $\kappa$ B transcription, which regulates the expression of genes involved in cell cycle regulation and inflammation [48].

Transcription factor binding site enrichment analysis of the promoter regions of the cadmium-modulated genes showed significant enrichment for transcription factors Sp1, SREBP-1, and ZP5. Sp1, an important mediator of the cell cycle [49], may activate IL-10 in response to inflammatory signals [50] and play a role in the cellular responses to DNA damage and tumor metastasis [51,52]. Sp1 also has known involvement in oxidative stress response pathways [43]. Interestingly, Sp1 is a known regulator of two cancer-associated genes in the cadmium-associated networks. Specifically, ST3 beta-galactoside alpha-2,3-sialyltransferase 1 (*ST3GAL1*) and keratin 16 (*KRT16*) are regulated by Sp1. The Sterol regulatory element-binding protein-1 (SREBP-1) regulates the expression of lipogenic genes, such as Fatty acid synthase (*FAS*) [53]. SREBP-1 activity has been seen in both colorectal neoplasia and breast cancer [54]. Although no common transcription factors were identified between the arsenic and cadmium-modulated gene sets, the induced transcriptional changes of the target genes are associated with similar biological processes.

#### **Comparison of metal-induced genomic response of TK6 cells to existing genomic studies**

In an effort to gain further understanding of the relevance of these findings, we compared the genes and pathways identified here to existing genomic studies. Specifically, we compared our findings to the gene expression changes that were observed in a human population in Thailand exposed to arsenic at varying levels [55]. In that study, a total of 447 genes with altered expression were associated with arsenic

exposure. Two of the genes, namely *IL1RN* and *G3BP2*, were also found to be differentially expressed here in TK6 cells exposed to low-dose arsenic. As discussed previously, *IL1RN* is a human cytokine involved in inflammation and tumorigenesis [24]. Also, *G3BP2* is often over-expressed in human cancers, and has been suggested to act as a negative regulator of the p53 tumor suppressor pathway [56]. In the previous study, enrichment for NF- $\kappa$ B binding sites in the arsenic-modulated genes was identified. Interestingly, here we find enrichment for NF- $\kappa$ B-associated genes altered in response to low-dose cadmium exposure, but not to arsenic.

We further compared the results obtained here with the metal-induced transcriptional response in yeast exposed to various transition metals [57]. In the study by Jin *et al.* the investigators exposed yeast to arsenic and cadmium and compared the transcriptional responses at various concentrations [57]. Two genes modified here in response to arsenic exposure, namely *C1ORF123* and *DNAJB5*, were also found to have altered expression in yeast exposed to arsenic. There were no genes with common transcriptional response between the TK6 cells and yeast exposed to cadmium. These genomic comparisons highlight some commonalities in the transcriptional changes induced by arsenic and cadmium exposure, but the larger lack of overlap suggests that different transcriptional responses are likely influenced by cell type, organism, and dose.

## Conclusions

In this study, we identified genes with expression levels that were significantly changed in TK6 cells exposed to low doses of either arsenic or cadmium. Our selection of these two metals was based on their environmental prevalence and our interest in identifying biological processes that may be modulated by low level exposure. Using a systems level approach to examine genes and their associated networks, we show that even low-level exposure to these metals impacts gene expression. In addition, both metals were found to modulate genes that encode proteins with similar biological functions in the cell. These results highlight that even at low doses, exposure to either of these metals has an impact on the expression of molecular networks associated with important biological functions. These results suggest that more research is needed to examine the biological effects of low dose exposure to toxic metals as such pathways are potentially modulated not only *in vitro*, but modulated *in vivo* as well. Future research will include the examination of protein activity associated with the metal-modulated genes and predicted transcription factors. In addition, these findings suggest that studies examining samples from humans with known

low-level exposures to these metals would be warranted and results compared to our *in vitro* studies.

## Methods

### Cell Type

The human TK6 lymphoblastoid cell line was obtained from ATCC (Manassas, VA) and cultured according to ATCC's guidelines (RPMI 1640 with L-glutamine, Penicillin-Streptomycin, FetalPlex™ Animal Serum Complex).

### Metal exposures

Sodium arsenite and cadmium chloride were purchased from Sigma (St. Louis, MO). Human TK6 cells were plated at  $1 \times 10^6$  cells/ml and treated with 0.1  $\mu$ M of either sodium arsenite or cadmium chloride for 24 hours. These equimolar, nonlethal doses were equivalent to 7.5 ppb of arsenic and 11.2 ppb of cadmium. Cell survival was assessed with the Trypan Blue Exclusion Assay using a hemocytometer with a 1:1 cell dilution and a 1:5 Trypan Blue dilution. For the arsenic exposures, two separate biological treatment replicates and two mock-treated controls were performed for subsequent microarray analysis. Likewise, for the cadmium exposures, two separate biological treatment replicates and two mock-treated controls were performed for subsequent microarray analysis.

### Microarray Processing

Total RNA was isolated from either arsenic or cadmium treated and mock treated human TK6 cells using the Qiagen's RNeasy® Kit (Qiagen, Valencia CA). RNA was quantified with the NanoDrop™ 1000 Spectrophotometer (Thermo Scientific, Waltham MA) and its integrity was verified by the Agilent 2100 Bioanalyzer. RNA was biotin-labelled according to the Affymetrix protocol and hybridized to Affymetrix GeneChip® Human Gene 1.0 ST arrays.

### Differential Gene Expression Analysis

Data were first normalized using Robust Multi-Chip Average (RMA) [58] and filtered for background noise ( $> \text{abs}[30]$ ), which resulted in a reduction of probesets from 28869 to 20695 for arsenic and 21450 for cadmium. Differential gene expression was determined as a significant difference in the expression of a gene (exposed versus unexposed) where  $p < 0.01$  (ANOVA model, Partek). Network analyses were performed using the Ingenuity software <http://www.ingenuity.com>. Statistical analyses were carried out within Ingenuity on both the arsenic and cadmium-modulated gene lists to identify the most significantly altered biological functions. Gene set enrichment analysis (GSEA) [17] was performed using the GSEA desktop software [59].

Significance is defined by the family wise error rate (FWER) p-value, which describes the probability that results are falsely positive. Microarray data have been submitted to the Gene Expression Omnibus (GEO) repository (GSE20320).

### Transcription Factor-binding Site Analysis

Transcription factor binding site analysis was performed using the EXPANDER 2.0 Software <http://acgt.cs.tau.ac.il/expander/>. For the analysis, Affymetrix probesets were linked to sequence data for regions 1,000 base pairs upstream and 200 base pairs downstream of the transcription start sites, and these were analyzed for significant enrichment of transcription factor binding sites. Significance ( $p \leq 0.016$ ) was calculated where significance is the probability of obtaining an equal or greater number of sequences with a model match in a randomly drawn sample of the same size as the input sequence set.

### Quantitative Real Time PCR Verification of Expression Data

Expression levels of the selected genes within the interactomes and sub-networks were validated with quantitative real time PCR (qRT-PCR) using RNA obtained from separate biological exposures (e.g. independent experiments) as those used for the microarray analyses. The forward and reverse primers were designed using Integrated DNA Technologies' Real Time PCR SciTool <http://www.IDTdna.com/SciTools>. The Qiagen QuantiTect SYBR Green PCR kit and Roche Lightcycler 480 were used for all qRT-PCR experiments. Fold changes between treated and untreated samples were calculated based on delta-delta cycle threshold (ddCt) values and normalized with  $\beta$ -actin as a housekeeping gene. Significance was calculated using a t-test.

### Additional material

**Additional file 1: Arsenic-modulated genes.** Additional File 1 lists all 74 probesets (62 genes) that showed significantly altered transcript levels resulting from arsenic exposure (0.1  $\mu$ M for 24 hours). Corresponding gene symbols, reference sequences IDs, and fold changes are included.

**Additional file 2: Cadmium-modulated genes.** Additional File 2 lists all 135 probesets (105 genes) that showed significantly altered transcript levels resulting from cadmium exposure (0.1  $\mu$ M for 24 hours). Corresponding gene symbols, reference sequences, and fold changes are included.

**Additional file 3: Gene products in networks.** Additional File 3 lists all of the proteins contained within the arsenic and cadmium-induced interactomes. Each protein is listed as either directly altered resulting from arsenic or cadmium exposure or associated with the directly altered transcripts. Also, proteins that are cancer-related are identified.

**Additional file 4: Metal-modulated sub-networks.** Additional File 4 shows eight additional metal-modulated sub-networks. Arsenic sub-networks include (A) TNF-alpha associated sub-network 2 and (B) HNF-4 associated sub-network 3. Cadmium sub-networks include (C) p53

associated sub-network 1, (D) MYC-associated sub-network 3, (E) NF- $\kappa$ B associated sub-network 4, and (F) PI3K associated sub-network 5. Networks are displayed with symbols representing encoded proteins corresponding to their RNA transcripts that were either directly up-regulated (red symbols), down-regulated (green symbols), or associated with the modified transcripts (while symbols). Arsenic and cadmium gene sets were combined and mapped to (G) p38 MAPK associated sub-network 1, and (H) HNF-4 associated sub-network 3. Networks are displayed with symbols representing gene products of arsenic-modulated genes (red symbols) and cadmium-modulated genes (green symbols).

**Additional file 5: Genes with enriched transcription factor binding sites.** Metal-modulated genes and their predicted transcription factors are listed.

**Additional file 6: Comparison of gene expression levels of target genes assessed with qRT-PCR and microarray.** Fold changes in transcript levels for (A) arsenic and (B) cadmium-exposed samples are plotted for selected target genes. Correlations (R) between microarray and qRT-PCR values are displayed.

### Acknowledgements

We would like to acknowledge the staff at the microarray facility at UNC and for their assistance in the processing of the arrays. We would also like to acknowledge Dr. Elyse Lee for her help with the Real-time PCR experiments. This research was supported in part by a grant from the National Institute of Environmental Health Sciences (P30ES10126).

### Authors' contributions

RCF conceived and designed experiments. LS and MAB performed exposures and determined cell killing. LS, JR, and MAB processed samples for microarray experiments. RCF, JR, and MAB analyzed the data. Manuscript was written by MAB, RCF, JR, and LS. All authors read and approved the final manuscript.

### Authors' Information

RCF is Assistant Professor at Gillings School of Global Public Health in the Department of Environmental Sciences and Engineering, University of North Carolina-Chapel Hill, MHRC 1213, CB#7431, Phone: 919-843-6864, Fax: 919-966-7911.

Received: 8 April 2010 Accepted: 1 April 2011 Published: 1 April 2011

### References

1. Agency for Toxic Substances and Disease Registry (ATSDR), *Priority List of Hazardous Substances* 2007.
2. Kapaj S, Peterson H, Liber K, Bhattacharya P: **Human health effects from chronic arsenic poisoning—a review.** *J Environ Sci Health A Tox Hazard Subst Environ Eng* 2006, **41(10)**:2399-2428.
3. Jarup L, Akesson A: **Current status of cadmium as an environmental health problem.** *Toxicology and Applied Pharmacology* 2009, **238**:201-208.
4. IARC: **Beryllium, Cadmium, Mercury and Exposures in the Glass Manufacturing Industry.** IARC Monographs on the Evaluation of Carcinogenic Risk of Chemicals to Humans. 1993, **58**:444.
5. IARC: **IARC Monographs on the Evaluation of the Carcinogenic Risks to Humans, Suppl. 7, Overall Evaluations of Carcinogenicity: An Updating of IARC Monographs Volumes 1 to 42.** 1987.
6. IARC: **IARC Monographs on the Evaluation of the Carcinogenic Risk of Chemicals to Humans, Some Metals and Metallic Compounds.** 1980, **23**.
7. IARC: **Some drinking-water disinfectants and contaminants, including arsenic.** 2004, **84**:1-477.
8. Huang C, Qingdong K, Costa M, Shi X: **Molecular mechanisms of arsenic carcinogenesis.** *Molecular and cellular biochemistry* 2004, **255(1-2)**:57.
9. Joseph P: **Mechanisms of cadmium carcinogenesis.** *New Insights into the Mechanisms of Cadmium Toxicity - Advances in Cadmium Research* 2009, **238(3)**:272-279.
10. Liao CM, Shen HH, Chen CL, Hsu LI, Lin TL, Chen SC, Chen CJ: **Risk assessment of arsenic-induced internal cancer at long-term low dose exposure.** *J Hazard Mater* 2009, **165(1-3)**:652-663.



11. Satarug S, Moore MR: Adverse health effects of chronic exposure to low-level cadmium in foodstuffs and cigarette smoke. *Environ Health Perspect* 2004, **112**(10):1099-1103.
12. Andrew AS, Warren AJ, Barchowsky A, Temple KA, Klei L, Soucy NV, O'Hara KA, Hamilton JW: Genomic and proteomic profiling of responses to toxic metals in human lung cells. *Environ Health Perspect* 2003, **111**(6):825-835.
13. Hester SD, Drobná Z, Andrews DMK, Liu J, Waalkes MP, Thomas DJ, Styblo M: Expression of AS3MT alters transcriptional profiles in human urothelial cells exposed to arsenite. *Human and Experimental Toxicology* 2009, **28**(1):49-61.
14. Hu Y, Jin X, Snow ET: Effect of arsenic on transcription factor AP-1 and NF-kappaB DNA binding activity and related gene expression. *Toxicol Lett* 2002, **133**(1):33-45.
15. Zheng XH, Watts GS, Vought S, Gandolfi AJ: Low-level arsenite induced gene expression in HEK293 cells. *Toxicology* 2003, **187**(1):39-48.
16. Yamada H, Koizumi S: DNA microarray analysis of human gene expression induced by a non-lethal dose of cadmium. *Industrial health* 2002, **40**(2):159.
17. Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, Paulovich A, Pomeroy SL, Golub TR, Lander ES, et al: Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. *Proceedings of the National Academy of Sciences of the United States of America* 2005, **102**(43):15545-15550.
18. Danaee H, Nelson HH, Liber H, Little JB, Kelsey KT: Low dose exposure to sodium arsenite synergistically interacts with UV radiation to induce mutations and alter DNA repair in human cells. *Mutagenesis* 2004, **19**(2):143-148.
19. Guillamet E, Creus A, Ponti J, Sabbioni E, Fortaner S, Marcos R: In vitro DNA damage by arsenic compounds in a human lymphoblastoid cell line (TK6) assessed by the alkaline Comet assay. *Mutagenesis* 2004, **19**(2):129-135.
20. Hickman MJ, Samson LD: Apoptotic signaling in response to a single type of DNA lesion, O(6)-methylguanine. *Mol Cell* 2004, **14**(1):105-116.
21. Hornhardt S, Gomolka M, Walsh L, Jung T: Comparative investigations of sodium arsenite, arsenic trioxide and cadmium sulphate in combination with gamma-radiation on apoptosis, micronuclei induction and DNA damage in a human lymphoblastoid cell line. *Mutat Res* 2006, **600**(1-2):165-176.
22. Tapio S, Danescu-Mayer J, Assmus M, Posch A, Gomolka M, Hornhardt S: Combined effects of gamma radiation and arsenite on the proteome of human TK6 lymphoblastoid cells. *Mutat Res* 2005, **581**(1-2):141-152.
23. Oesch S, Gertsch J: Cannabinoid receptor ligands as potential anticancer agents-high hopes for new therapies? *J Pharm Pharmacol* 2009, **61**(7):839-853.
24. Balkwill F, Mantovani A: Inflammation and cancer: back to Virchow? *The Lancet* 2001, **357**(9255):539-545.
25. Fujiwaki R, Iida K, Nakayama K, Kanasaki H, Hata K, Katabuchi H, Okamura H, Miyazaki K: Clinical significance of interleukin-1 receptor antagonist in patients with cervical carcinoma. *Gynecologic oncology* 2003, **89**(1):77-83.
26. Argos M, Kibriya MG, Parvez F, Jasmine F, Rakibuz-Zaman M, Ahsan H: Gene Expression Profiles in Peripheral Lymphocytes by Arsenic Exposure and Skin Lesion Status in a Bangladeshi Population. *Cancer Epidemiology Biomarkers & Prevention* 2006, **15**(7):1367-1375.
27. Dangi-Garimella S, Yun J, Eves EM, Newman M, Erkeland SJ, Hammond SM, Minn AJ, Rosner MR: Raf kinase inhibitory protein suppresses a metastasis signalling cascade involving LIN28 and let-7. *EMBO J* 2009, **28**(4):347-358.
28. Cunha IW, Carvalho KC, Martins WK, Marques SM, Muto NH, Falzoni R, Rocha RM, Jr SA, Simoes ACQ, Fahham L, et al: Identification of Genes Associated with Local Aggressiveness and Metastatic Behavior in Soft Tissue Tumors. *Translational Oncology* 2010, **3**:23-32.
29. Wicki R, Franz C, Scholl FA, Heizmann CW, Schäfer BW: Repression of the candidate tumor suppressor gene S100A2 in breast cancer is mediated by site-specific hypermethylation. *Cell calcium* 1997, **22**(4):243-254.
30. Lee SW, Tomasetto C, Swisshelm K, Keyomarsi K, Sager R: Down-regulation of a member of the S 100 gene family in mammary carcinoma cells and reexpression by azadeoxycytidine treatment. *Proceedings of the National Academy of Sciences of the United States of America* 1992, **89**(6):2504-2508.
31. Almadori G, Bussu F, Galli J, Rigante M, Lauriola L, Michetti F, Maggiano N, Schäfer BW, Heizmann CW, Ranelletti FO, et al: Diminished expression of S100A2, a putative tumour suppressor, is an independent predictive factor of neck node relapse in laryngeal squamous cell carcinoma. *J Otolaryngol Head Neck Surg* 2009, **38**(1):16-22.
32. Bulk E, Sargin B, Krug U, Hascher A, Jun Y, Knop M, Kerkhoff C, Gerke V, Liersch R, Mesters RM, et al: S100A2 induces metastasis in non-small cell lung cancer. *Clin Cancer Res* 2009, **15**(1):22-29.
33. Hu J, Dong A, Fernandez-Ruiz V, Shan J, Kawa MM-AE, Prieto J, Qian C: Blockade of Wnt Signaling Inhibits Angiogenesis and Tumor Growth in Hepatocellular Carcinoma. *Cancer Research* 2009, **69**(17):6951-6959.
34. Han XXAS: Secreted Frizzled-related Protein 1 (SFRP1) Protects Fibroblasts from Ceramide-induced Apoptosis. *The Journal of Biological Chemistry* 2004, **279**(4):2832-2840.
35. Hanson ML, Brundage KM, Schafer R, Tou JC, Barnett JB: Prenatal cadmium exposure dysregulates sonic hedgehog and Wnt/beta-catenin signaling in the thymus resulting in altered thymocyte development. *Toxicol Appl Pharmacol* 2002, **182**(2):136-145.
36. Videira PA, Ligeiro D, Correia M, Trindade H: Gene expression analysis in superficial bladder cancer: comparison of two suitable endogenous reference genes. *Curr Urol* 2007, **1**:145-150.
37. Wang P-H, Lee W-L, Juang C-M, Yang Y-H, Lo W-H, Lai C-R, Hsieh S-L, Yuan C-C: Altered mRNA expressions of sialyltransferases in ovarian cancers. *Gynecologic oncology* 2005, **99**(3):631-639.
38. Burchell J, Poulson R, Hanby A, Whitehouse C, Cooper L, Clausen H, Miles D, Taylor-Papadimitriou J: An alpha2,3 sialyltransferase (ST3Gal II) is elevated in primary breast carcinomas. *Glycobiology* 1999, **9**:1307-1311.
39. Kudo T, Ikehara Y, Togayachi A, Morozumi K, Watanabe M, Nakamura M, Nishihara S, Narimatsu H: Up-regulation of a set of glycosyltransferase genes in human colorectal cancer. *Laboratory investigation; a journal of technical methods and pathology* 1998, **78**(7):797-811.
40. Schneider F, Kemmer W, Haensch W, Franke G, Gretschel S, Karsten U, Schlag PM: Overexpression of sialyltransferase CMP-sialic acid:Galbeta1, 3GalNAc-R alpha6-Sialyltransferase is related to poor patient survival in human colorectal carcinomas. *Cancer Res* 2001, **61**:4605-4611.
41. Zhang T, Woods TL, Elder JT: Differential Responses of S100A2 to Oxidative Stress and Increased Intracellular Calcium in Normal, Immortalized, and Malignant Human Keratinocytes. 2002, **119**(5):1196-1201.
42. Enk CD, Jacob-Hirsch J, Gal H, Verbovetski I, Amariglio N, Mevorach D, Ingber A, Givol D, Rechavi G, Hochberg M: The UVB-induced gene expression profile of human epidermis in vivo is different from that of cultured keratinocytes. *Oncogene* 2006, **25**(18):2601-2614.
43. Allen RG, Tresini M: Oxidative stress and gene regulation. *Free Radical Biology and Medicine* 2000, **28**(3):463-499.
44. MSigDB Molecular Signatures Database. [http://www.broadinstitute.org/gsea/msigdb/index.jsp].
45. laquinta PJ, Lees JA: Life and death decisions by the E2F transcription factors. *Cell differentiation/Cell division, growth and death* 2007, **19**(6):649-657.
46. Muller H, Bracken AP, Vernell R, Moroni MC, Christians F, Grassilli E, Prosperini E, Vigo E, Oliner JD, Helin K: E2Fs regulate the expression of genes involved in differentiation, development, proliferation, and apoptosis. *Genes Dev* 2001, **15**(3):267-285.
47. Zhao H, Jin S, Fan F, Fan W, Tong T, Zhan Q: Activation of the transcription factor Oct-1 in response to DNA damage. *Cancer Res* 2000, **60**(22):6276-6280.
48. Oswald F, Liptay S, Adler G, Schmid RM: NF-kappa B2 Is a Putative Target Gene of Activated Notch-1 via RBP-Jkappa. *Molecular and cellular biology* 1998, **18**(4):2077-2088.
49. Tapias A, Ciudad CJ, Roninson IB, Noe V: Regulation of Sp1 by cell cycle related proteins. *Cell cycle (Georgetown, Tex)* 2008, **7**(18):2856-2867.
50. Larsson L, Johansson P, Jansson A, Donati M, Rymo L, Berglund T: The Sp1 transcription factor binds to the G-allele of the -1087 IL-10 gene polymorphism and enhances transcriptional activation. *Genes Immun* 2009, **10**(3):280-284.
51. Olofsson BA, Kelly CM, Kim J, Hornsby SM, Azizkhan-Clifford J: Phosphorylation of Sp1 in Response to DNA Damage by Ataxia Telangiectasia-Mutated Kinase. *Molecular Cancer Research* 2007, **5**(12):1319-1330.
52. Takami Y, Russell MB, Gao C, Mi Z, Guo H, Mantyh CR, Kuo PC: Sp1 regulates osteopontin expression in SW480 human colon adenocarcinoma cells. *Surgery* 2007, **142**(2):163-169.

53. Li JN, Mahmoud MA, Han WF, Ripple M, Pizer ES: **Sterol Regulatory Element-Binding Protein-1 Participates in the Regulation of Fatty Acid Synthase Expression in Colorectal Neoplasia.** *Experimental cell research* 2000, **261**(1):159-165.
54. Yang Yu-An, Morin PJ, Han WF, Chen T, Bornman DM, Gabrielson EW, Pizer ES: **Regulation of fatty acid synthase expression in breast cancer by sterol regulatory element binding protein-1c.** *Experimental cell research* 2003, **282**(2):132-137.
55. Fry RC, Navasumrit P, Valiathan C, Svensson JP, Hogan BJ, Luo M, Bhattacharya S, Kandjanapa K, Soontararuks S, Nookabkaew S, *et al*: **Activation of Inflammation/NF- $\kappa$ B Signaling in Infants Born to Arsenic-Exposed Mothers.** *PLoS Genet* 2007, **3**(11):e207.
56. Kim MM, Wiederschain D, Kennedy D, Hansen E, Yuan ZM: **Modulation of p53 and MDM2 activity by novel interaction with Ras-GAP binding proteins (G3BP).** *Oncogene* 2007, **26**(29):4209-4215.
57. Jin YH, Dunlap PE, McBride SJ, Al-Refai H, Bushel PR, Freedman JH: **Global Transcriptome and Deletome Profiles of Yeast Exposed to Transition Metals.** *PLoS Genet* 2008, **4**(4):e1000053.
58. Irizarry RA, Hobbs B, Collin F, Beazer-Barclay YD, Antonellis KJ, Scherf U, Speed TP: **Exploration, normalization, and summaries of high density oligonucleotide array probe level data.** *Biostatistics* 2003, **4**(2):249-264.
59. Subramanian A, Kuehn H, Gould J, Tamayo P, Mesirov JP: **GSEA-P: a desktop application for Gene Set Enrichment Analysis.** *Bioinformatics* 2007, **23**(23):3251-3253.

doi:10.1186/1471-2164-12-173

**Cite this article as:** Benton *et al.*: Comparative genomic analyses identify common molecular pathways modulated upon exposure to low doses of arsenic and cadmium. *BMC Genomics* 2011 **12**:173.

**Submit your next manuscript to BioMed Central  
and take full advantage of:**

- Convenient online submission
- Thorough peer review
- No space constraints or color figure charges
- Immediate publication on acceptance
- Inclusion in PubMed, CAS, Scopus and Google Scholar
- Research which is freely available for redistribution

Submit your manuscript at  
[www.biomedcentral.com/submit](http://www.biomedcentral.com/submit)

