Am J Transl Res 2023;15(10):5984-5996
www.ajtr.org /ISSN:1943-8141/AJTR0150430

Original Article

Gene expression profiles to analyze the anticancer and carcinogenic effects of arsenic in bladder cancer

Xiaomin Han[1,2,4*], Tengteng Zhang[4*], Qiang Ma[4*], Ruyi Chang[4], Siyuan Xin[4], Qingzhen Yu[3,5], Guojun Zhang[2], Yukun Wang[2,3]

[1]Department of Oncology and Chemistry, The First Affiliated Hospital, Jinan University, Guangzhou 510632, Guangdong, China; [2]Department of Pharmacy, Southern University of Science and Technology Hospital, Shenzhen 518055, Guangdong, China; [3]Department of Pharmacology, School of Medicine, Southern University of Science and Technology, Shenzhen 518055, Guangdong, China; [4]School of Basic and Forensic Medicine, Baotou Medical College, Baotou 014060, Inner Mongolia, China; [5]Medical Research Center, Southern University of Science and Technology Hospital, Shenzhen 518055, Guangdong, China. [*]Equal contributors.

Received March 29, 2023; Accepted October 11, 2023; Epub October 15, 2023; Published October 30, 2023

Abstract: Objectives: Arsenic is one of the greatest hazards as an environmental carcinogen. At the same time it is also a promising anticancer agent, that can be used to treat acute promyelocytic leukemia (APL) and some other tumors. Arsenic trioxide (ATO) exerts its therapeutic effect by promoting degradation of an oncogenic protein that drives the growth of APL cells. However, the molecular mechanisms that govern these paradoxical effects of arsenic in bladder cancer remain unclear. We speculate that they share the common mechanism that arsenic binds to the target proteins and subsequently impacts the expression of downstream genes. Methods: To address this issue, three Gene Set Enrichments (GSE) were loaded from the Gene Expression Omnibus (GEO) database with four expression matrices. Three of them were mice samples at exposure times of 1, 2, and 12 weeks, and the last was a human urothelial cell (HUC1) sample. Differentially expressed genes (DEGs) from 4 expression groups were identified at iDEP and analyzed at Metascape and Cytoscape for signaling pathway analysis and protein-protein interaction (PPI) analysis. The web-portals UALCAN and GEPIA were used to analyze the role of DEGs in the crosstalk between carcinogenic and anticancer effects. The putative downstream genes of arsenic binding proteins were retrieved using the Cistrome Data Browser. Real-time PCR was used to validate the expression of DEGs. Results: The signaling pathways referred to lipid metabolism. Responses to various stimuli or hormones were overrepresented in 4 expression matrices. The PPI network emphasized the role of KRAS and TNF signaling in different groups. Furthermore, BDKRB2, FOS, NR4A1, PLAU, SH3BGRL, and F10 played an important role in the crosstalk between carcinogenic and anticancer effects in bladder cancer. Arsenic may impact the activity of ACTB, BACH1, NME2, RBBP4, PARP1, and PML by direct binding, and thus influence the expression of downstream genes such as PAX6, MLLT11, LTBP1, PCSK5, ZFP36, COL8A2, and IL1R2. Conclusion: Arsenic exerted carcinogenic and anticancer functions by altering the expression of crosstalk genes such as BDKRB2, FOS, NR4A1, PLAU, SH3BGRL, and F10, and these were due to arsenic binding proteins.

Keywords: Arsenic, bladder cancer, gene expression profile, carcinogenesis, anticancer

Introduction

Arsenic species as traditional and modern medicines have been often used and studied, even though paradoxically it is known for its toxicity. Much effort has been put into interpreting its contradictory effects of carcinogenic and anticancer activity. A strong association between chronic arsenic exposure and the development of various tumors including cancers of skin, bladder, lung, and liver was found [1]. There is evidence of increases in lung, bladder, and kidney cancer even 40 years after high arsenic exposure [2]. The cancer risk attributable to arsenic exposure is reported in Northern New England [3], Iran [4], Hungary [5], China [6], and others. Chronic As^{3+} exposure can convert a non-invasive papillary bladder cancer to an invasive form that acquires some basal characteristics as shown by immunohistochemical analysis [7]. One possible mechanism involves epigenetic reprogramming through arsenic

Table 1. Primers for real-time PCR assay

Gene	Forward	Reverse
GAPDH	CCTCTGACTTCAACAGCGACAC	TGGTCCAGGGGTCTTACTCC
BDKRB2	TGGAACGCCAGATTCACAAA	TCAACTCATTTCCCAAGGTCG
FOS	ACCGCCACGATGATGTTCTC	TGCGGGTGAGTGGTAGTAAGAG
NR4A1	TGCCAATCTCCTCACTTCCC	ATCCCCAGCATCTTCCTTCC
PLAU	CAACGACATTGCCTTGCTGA	CTTGTGCCAAACTGGGGATC
SH3BGRL	CTGCTCCACAGCCCTTTTCA	GCATCATAGTCCCCGCGATA
IL1R2	GCTTTCCTGCCGTTCATCTC	CAGCGGTAATAGCCAGCATCTT
PAX6	AGCCCTCACAAACACCTACAGC	CATAACTCCGCCCATTCACC
PCSK5	CGCAGGAATCCAGAGTCACAA	CCAGGCGAAAAGTACCCAGA

exposure, which causes changes in cell signaling and dysfunctions of different epigenetic elements, including DNA methylation, histone post-translational modifications, histone variants, and microRNAs [8, 9]. Furthermore, 218-5p/EGFR signaling pathway may be a therapeutic target for the treatment of lung cancer induced by chronic arsenic exposure [10]. The connections among ER stress-associated unfolded protein response, mitochondrial dysfunction, and autophagy in arsenic malignancies need further investigation [11].

Despite its carcinogenicity, arsenic and its compounds have also been identified as promising anticancer drugs. The ability of arsenic, specifically arsenic trioxide (ATO), to treat acute promyelocytic leukemia has played an important role in the re-emergence of this candidate to western medicine for the treatment of various cancers. ATO exerts its therapeutic effect by promoting degradation of an oncogenic protein that drives the growth of acute promyelocytic leukemia cells, PML-RARα (a fusion protein containing sequences from the PML zinc finger protein and retinoic acid receptor alpha) [12]. Chen and colleagues proposed that arsenic binding stabilizes the DNA-binding loop-sheet-helix motif, endowing TP53 mutants with thermostability and transcriptional activity, and then mediates tumor suppression in cellular and mouse xenograft models [13]. Also, ATO directly bound to GLI1 protein, inhibited its transcriptional activity, and decreased expression of endogenous GLI target genes [14]. It is well known that arsenic binds to a variety of proteins containing dithiols formed by vicinal cysteine structures, such as ring finger domain and zinc finger domain [15]. This specific interaction can be used to elucidate the function mechanisms of arsenic. A variety of labeled arsenics, for example, BIO-PAO (III) and azide-based AS-AC, are used to identify the arsenic binding protein in HepG2 cells and NB4 cells [16, 17].

We hypothesized that arsenic binding proteins govern paradoxical effects of arsenic in bladder cancer and subsequently impact the expression of downstream genes. The biological consequence of the binding between arsenic and target proteins may be reflected by their gene expression profiles. In this study, 3 GEO datasets were grouped into 4 expression matrices designated 1-week, 2-week, 12-week, and HUC1. Signaling pathway and PPI network analysis of DEGs were performed. DEGs were also integrated with altered expression genes in BLCA to illustrate the arsenic impact. To explore how arsenic induced the gene expression differentiation, the arsenic binding proteins and their downstream genes were analyzed as well.

Methods

Data information

The Gene Expression Omnibus Datasets (https://www.ncbi.nlm.nih.gov/gds/) were searched with the terms arsenic and bladder cancer. Arsenate, realgar, orpiment and AsS were also used but produced no search results. Three mRNA expression data GSE116554, GSE32102 and GSE90023 were obtained. The GSE116554 had 4 C57BL/6 female mice samples which were provided drinking water with or without 50 ppm arsenic (0.0867 mg/ml sodium arsenite) for 2 weeks. The GSE32102 had 49 C57BL/6J female mice samples, 20 of which were provided with 50 mg As/L drinking water and were grouped as 1-week and 12-week. Two HUC1 cell samples of the GSE90023 were chronically exposed to 1 μM arsenic trioxide for 10 months constituted the HUC1 group. These 4 expression matrices were used for differential gene expression analyses.

DEGs identification

Differential expression genes of 4 expression groups were analyzed at iDEP (version 0.96)

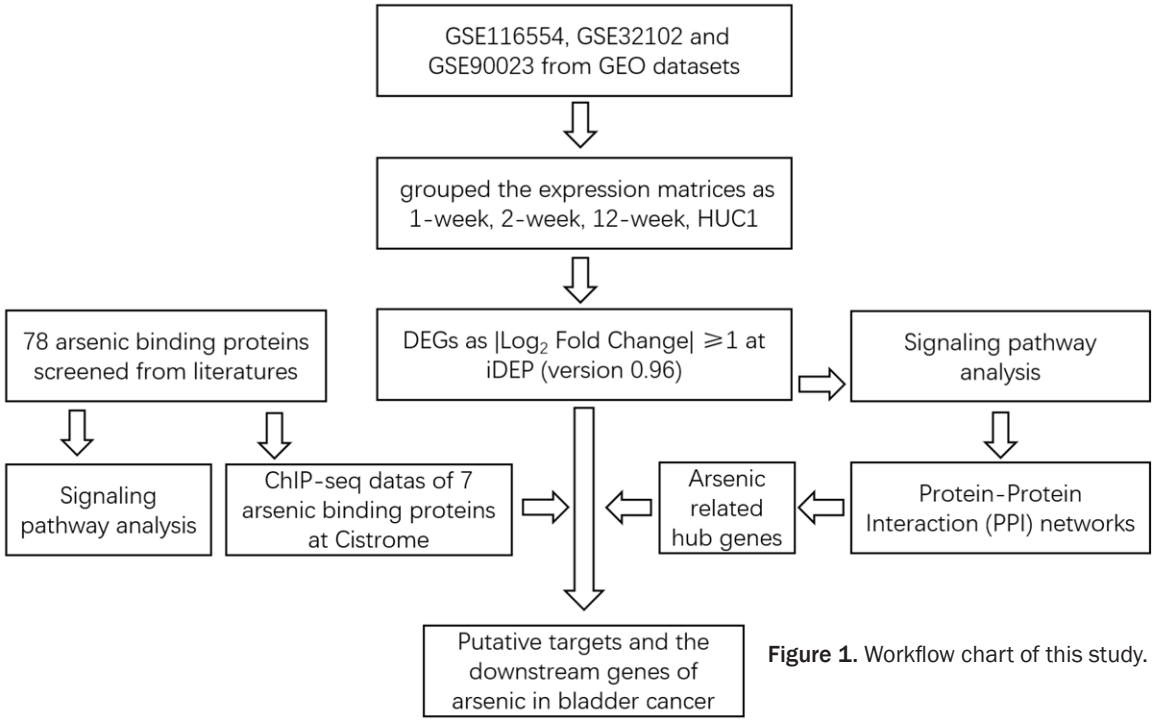

Figure 1. Workflow chart of this study.

Table 2. Characteristics of GEO datasets

Accession	Object	Treatment	Platform	Samples	Ref
GSE116554	C57BL/6 female mice	Fed with 50 ppm As-contained water for 2-week	GPL15523	GSM3242725, GSM3242726, GSM3242727, GSM3242728	[29]
GSE32102	C57BL/6J female mice	Fed with 50 mg/L As-contained water for 1-week or 12-week	GPL1261	GSM795662, GSM795663, GSM795664, GSM795665, GSM795666, GSM795667, GSM795668, GSM795669, GSM795670, GSM795671, GSM795692, GSM795693, GSM795694, GSM795695, GSM795696, GSM795697, GSM795698, GSM795699, GSM795700, GSM795701	[30]
GSE90023	HUC1	Cultured with 1 μM ATO in F-12K medium for 10-month	GPL10558	GSM2395808, GSM2395810	[31]

(http://bioinformatics.sdstate.edu/idep964/) [18]. The criteria for identifying the DEGs was as |Log2 Fold Change| > 1. *P*-value was not considered owing to HUC1 group having no replicate data. Non-annotated genes from GSE116554 and GSE32102 in particular were blast in NCBI for manual annotation. TBtools (version 1.108) was used to identify the common DEGs among 4 expression matrices and construct a Venn diagram [19].

Signaling pathway analysis and PPI network construct

All DEGs were used for signaling pathway analysis at Metascape (https://metascape.org/) [20]. Gene ontology biological processes, KEGG pathway, Wiki pathway, and Reactome gene set were included (*P*-value < 0.05). DEGs were further used for PPI analysis using string (https://cn.string-db.org/) (version 11.5) [21] and cytoscape (https://cytoscape.org/) (version 3.9.1) [22]. CytoHubba was employed to discover the highest linkage hub genes in the networks. MCODE plugin was performed to identify significant modules from PPI network with default settings (degree cutoff = 2, node score cutoff = 0.2, k-core = 2, and max. depth = 100).

UALCAN database analysis

UALCAN (http://ualcan.path.uab.edu) is an interactive web-portal to perform in-depth analyses of TCGA gene expression data [23]. It enables the analysis of mRNA expression differences to compare normal samples with primary tumor tissue samples, as well as clinico-

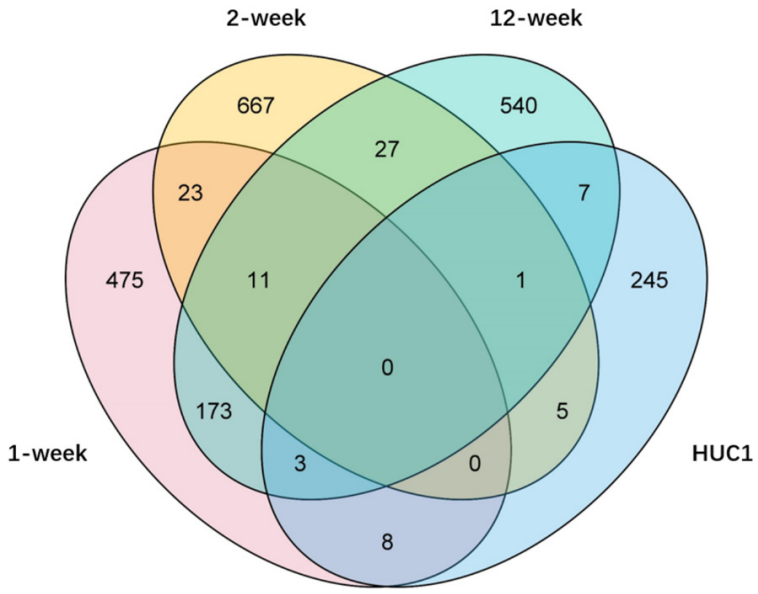

Figure 2. Venn diagram of DEGs from 4 expression matrices.

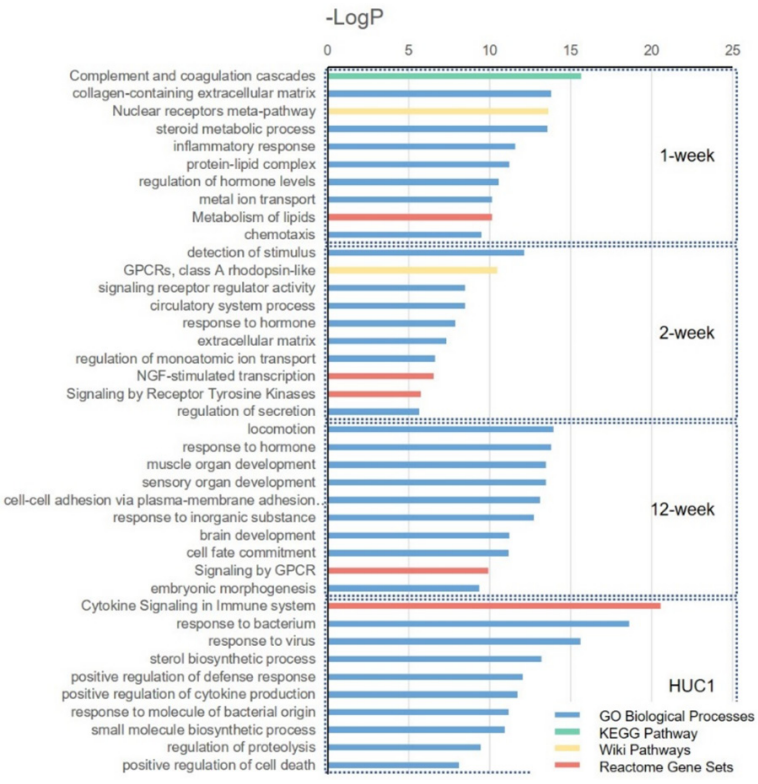

Figure 3. Signaling pathway analysis of DEGs. Top 10 pathways of each expression group are shown. GO biological processes, KEGG pathway, Wiki pathway, and Reactome gene set are included and shown in different colors.

pathologic features. The top 250 BLCA-UP and top 250 BLCA-DOWN genes were downloaded from this portal.

GEPIA survival analyses

GEPIA (http://gepia.cancer-pku.cn) was applied to conduct the overall survival analyses of hub genes and putative downstream genes using the Kaplan-Meier method. Patients were divided into 2 groups using median values as the cutoff points. Hazards ratio was also calculated.

Cistrome Data Browser analysis

Cistrome Data Browser (http://cistrome.org/db/) is a resource of human and mouse cis-regulatory information derived from ChIP-seq, DNase-seq and ATAC-seq chromatin profiling assays [24]. Seventy-eight arsenic binding proteins were analyzed. Downstream genes of ChIP-seq data were loaded for further analysis.

Cell culture and treatment

The T24 human bladder cancer cell line was purchased from Procell Co. Ltd., and mycoplasma testing was conducted. Reagents were as follows: Roswell Park Memorial Institute (RPMI)-1640 cell culture medium, trypsin, fetal bovine serum, 1% streptomycin (Thermo Fisher Scientific, Inc.), ATO (Sigma, Cat# 2026-73). The T24 cells were cultured in RPMI-1640 medium containing 10% fetal bovine serum and 1% streptomycin, and grown at 37°C and 5% carbon dioxide (CO_2). The cells were routinely subcultured when they reached ~80% confluence. ATO was dissolved in 1 M NaOH and further diluted to 10 mM in culture medium as a stock solution. The T24 cells were treated with 1 μM ATO for 1 h or 6 h, twice washed with PBS and were collected for the further use.

Table 3. The common signaling pathways and their shared genes

Signaling pathway	Group	Genes
hsa04610 Complement and coagulation cascades	1-week and 12-week	FGB, BDKRB2, F12, CPB2, GM4788, SERPINA5, F10, F5
GO:0006935 Chemotaxis	1-week and 2-week	MED1, ANKRD1, NODAL, PAX6, PTPRZ1
GO:0009410 Response to xenobiotic stimulus	1-week and 12-week	TGFBR2, ARG1, CRKL, CYP3A25, FOSL1, CES1B, CYP2C29, MTHFR, CPB2, CYP2A22, DNMT3A, SLCO1B2
GO:0009725 Response to hormone	1-week, 2-week and 12-week	PLCB1, TGFBR2, FGB, ARG1, NTRK3, CRKL, FOSL1, AGTR2, NR4A1, IGFBP1, PTPN2, ACHE, CTSJ, MUP3, DEFA-RS7, DNMT3A, GABRB3, LY6H, NSCME3L, ANGPTL3
GO:0007167 Enzyme-linked receptor protein signaling pathway	1-week and 12-week	PLCB1, TGFBR2, HGF, ARG1, NTRK3, CRKL, PIK3CD, BDKRB2, NOG, AFP, PTN, NR4A1, IGFBP1, PTPN2, FGFR4, ANGPTL1, FGF4, IBSP, AKAP4, PAX9
GO:0007420 Brain development	1-week and 12-week	PLCB1, TGFBR2, KRAS, CRKL, AGTR2, NOG, PAX6, CPHX3, CPHX1, CPHX2, NEUROD2, POU6F1
GO:0032102 Negative regulation of response to external stimulus	1-week and 2-week	MED1, FGB, KNG1, SAMHD1, CLEC4E
GO:0009617 Response to bacterium	2-week and HUC1	S100A9

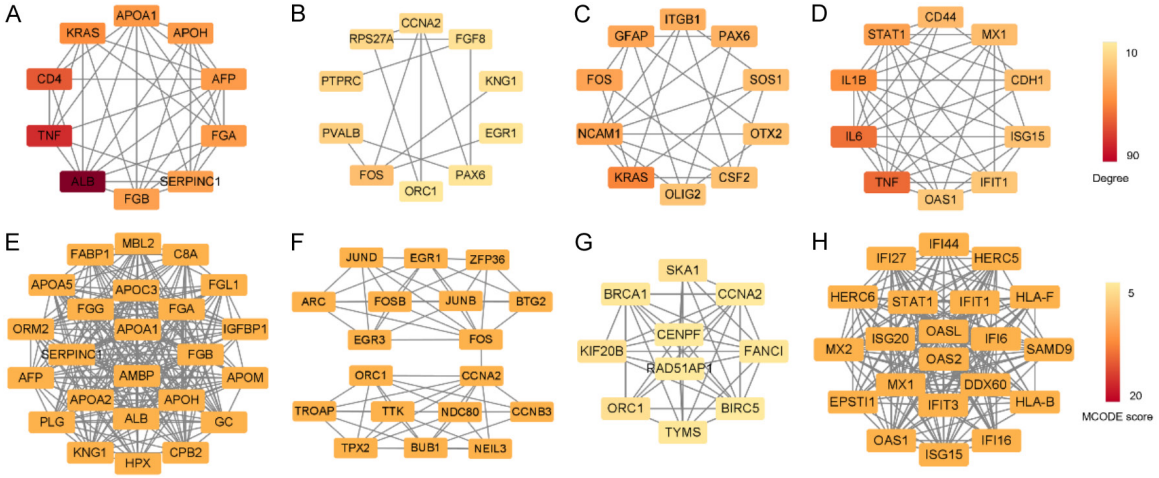

Figure 4. PPI network of the DEGs identified through the STRING database and Cytoscape. A-D. Top 10 hub genes with the highest degrees of 4 expression matrices. E-H. The key module of MCODE analysis of 4 expression group. Squares indicate DEGs, and lines indicate interactions. The color of squares means the degrees of connectivity and the MCODE score.

Quantitative real-time PCR assay

Total RNA was isolated according to manufacturer's instructions using Trizol reagent (TaKaRa, Cat# 9108Q). One thousand nanograms of RNA was used for gDNA eraser and reverse transcription (TaKaRa, Cat# RR047A). The cDNA was diluted 5 times, and 4 μL was used as a template in a 20-μL PCR reaction. Real-time PCR analysis was performed using TB Green Premix Ex TaqII (TaKaRa, Cat# RR820A) on an Applied Biosystems 7500 Real-Time PCR System, with 40 cycles under the following conditions: 95°C for 5 s, and 60°C for 45 s. GAPDH (GenBank: NM_002046.7) was used to normalize the samples. The primers are listed in **Table 1**.

Results

Acquisition of datasets and identification of DEGs

The gene expression data were retrieved from the GEO database using search terms "arsenic", "arsenite" and "bladder cancer". The workflow is illustrated in **Figure 1**. Three gene expression datasets (accession no. GSE116554, GSE32102 and GSE90023) from various platforms were obtained (**Table 2**). The expression

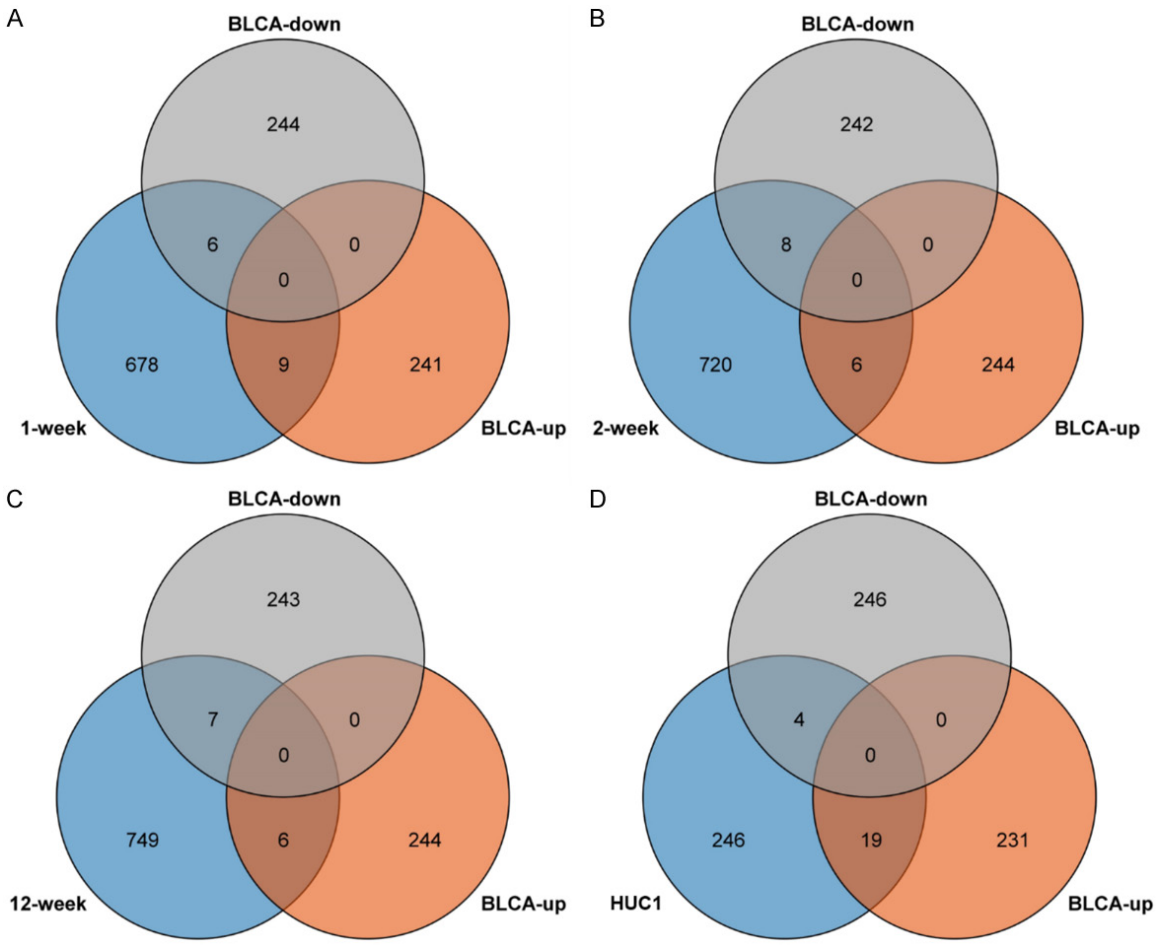

Figure 5. Venn diagram of DEGs and top 250 up-regulated and down-regulated genes in BLCA. (A) 1-week, (B) 2-weeks, (C) 12-weeks, (D) HUC1.

matrices were grouped as 1-week, 2-week, 12-week, and HUC1 from these datasets. Differential expression genes were analyzed at iDEP (version 0.96) [18]. Six hundred and eighty-five DEGs were identified from the 1-week group, 734 DEGs from the 2-week group, 762 DEGs from the 12-week group, and 269 DEGs from HUC1. Venn diagram of these DEGs were constructed to illustrate the overlapped genes (**Figure 2**). Common DEGs were found, especially between 1-week and 12-weeks, which were both from GSE32102. However, no common genes were found among the four groups.

Signaling pathway enrichment of DEGs

Signaling pathway analysis of DEGs, including GO biological processes, KEGG pathway, Wiki pathway, and Reactome gene set, were performed by Metascape [20]. The results of the top 10 signalings are presented in **Figure 3**. Some signaling pathways were overrepresented, for example lipid metabolism, and response to various stimuli or hormones. The strongest signaling of each expression group was complemented and coagulation cascades, detection of stimulus, locomotion, and cytokine signaling in the immune system. The common signaling pathways of the top 20 were complement and coagulation cascades (hsa04610), chemotaxis (GO:0006935), response to xenobiotic stimulus (GO:0009410), response to hormones (GO:0009725), enzyme-linked receptor protein signaling pathway (GO:0007167), brain development (GO:0007420), negative regulation of response to external stimulus (GO:0032102), and response to bacterium (GO:0009617) (**Table 3**). More common signaling pathways were found between 1-week and 12-weeks which were both from GSE32102.

Table 4. DEGs intersecting with 250 BLCA-UP and 250 BLCA-DOWN genes

Intersection	Expression group	Genes
Intersection 1 (BLCA-UP and AS-UP)	1-week	UCK2, NUSAP1, **FANCI**, OAS3, FCHO1, PLA2G2F
	2-week	TPX2, TROAP, NDC80, LAMC2, CCNA2, **BUB1**
	12-week	CENPF, TUBB3, BIRC5, TYMS
	HUC1	IFIT3, EPSTI1, IFI27, IFI44, HIST2H2AA3, IFI6, UCA1, RARRES1, ISG15, HIST1H2BK, MUC1
Intersection 2 (BLCA-DOWN and AS-DOWN)	1-week	**F10**, NFIA, **NR4A1**
	2-week	JUNB, BTG2, **FOS**, ZFP36, ADAMTS1
	12-week	**FOS**, **F10**, **NR4A1**, **BDKRB2**
	HUC1	SPON1, TMOD1, OLFM1
Intersection 3 (BLCA-UP and AS-DOWN)	1-week	**PLAU**, CCNB1, **BUB1**
	2-week	NF
	12-week	**FANCI**, BRCA1
	HUC1	TRIB3, SQLE, APOBEC3B, **PLAU**, DHCR7, MDK, SCD, PSAT1
Intersection 4 (BLCA-DOWN and AS-UP)	1-week	BIN1, BOC, **BDKRB2**
	2-week	ZEB2, AHNAK, MFAP4
	12-week	**SH3BGRL**, ITGA8, FGL2
	HUC1	**SH3BGRL**

NF: not found. Genes in bold belonged to different intersections.

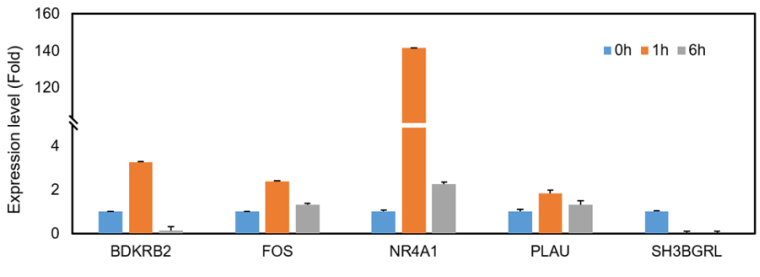

Figure 6. The expression of DEGs belonging to different intersections. Human bladder cancer cells T24 were treated with 1 μM ATO for 1 h and 6 h. Expression values were calculated using $2^{-\Delta\Delta CT}$ method and GAPDH as the endogenous reference gene.

PPI network construction and identification of hub genes

DEGs were further used for PPI analysis using String and Cytoscape. The top 10 hub genes were identified by degree in CytoHubba (**Figure 4A**). Among these, KRAS, TNF, FOS, and PAX6 appeared in different groups. Utilizing the MCODE plugin, the interconnected clusters were retrieved and the module with the highest score was selected for further analysis (**Figure 4B**). 24, 18, 10, and 21 nodes were identified from 1-week, 2-week, 12-week, and HUC1 expression groups individually. The common proteins were CCNA2 and ORC1. These results imply that RAS, TNF, FOS, PAX6, CCNA2, and ORC1 signaling may be disturbed by arsenic.

DEGs integrated with altered expression genes in bladder cancer (BLCA) describe the paradoxical effects of arsenic

Arsenic as a potent carcinogen, is causally related to various tumors. At the same time, As_2O_3 has been identified as an effective anticancer agent for the treatment of acute promyelocytic leukemia. To show these multiple effects in bladder cancer, the top 250 BLCA-UP and 250 BLCA-DOWN genes were loaded from UALCAN, and intersected with our DEGs (**Figure 5** and **Table 4**). DEGs of intersection 1 (BLCA-UP and AS-UP) and intersection 2 (BLCA-DOWN and AS-DOWN) had a coordinative expression, meaning that these genes may indicate a carcinogenic effect. Among them, FOS, F10, and NR4A1 arose more than once in different expression groups. The other two intersections contained genes expressed inversely, which may illustrate arsenic's anticancer effect. PLAU and SH3BGRL appeared twice in the converse intersections. However, FANCI, BUB1, and BDKRB2 may facilitate the crosstalk between carcinogenic and anticancer effects during different treatments because they exist in both coordinative and converse expression intersections.

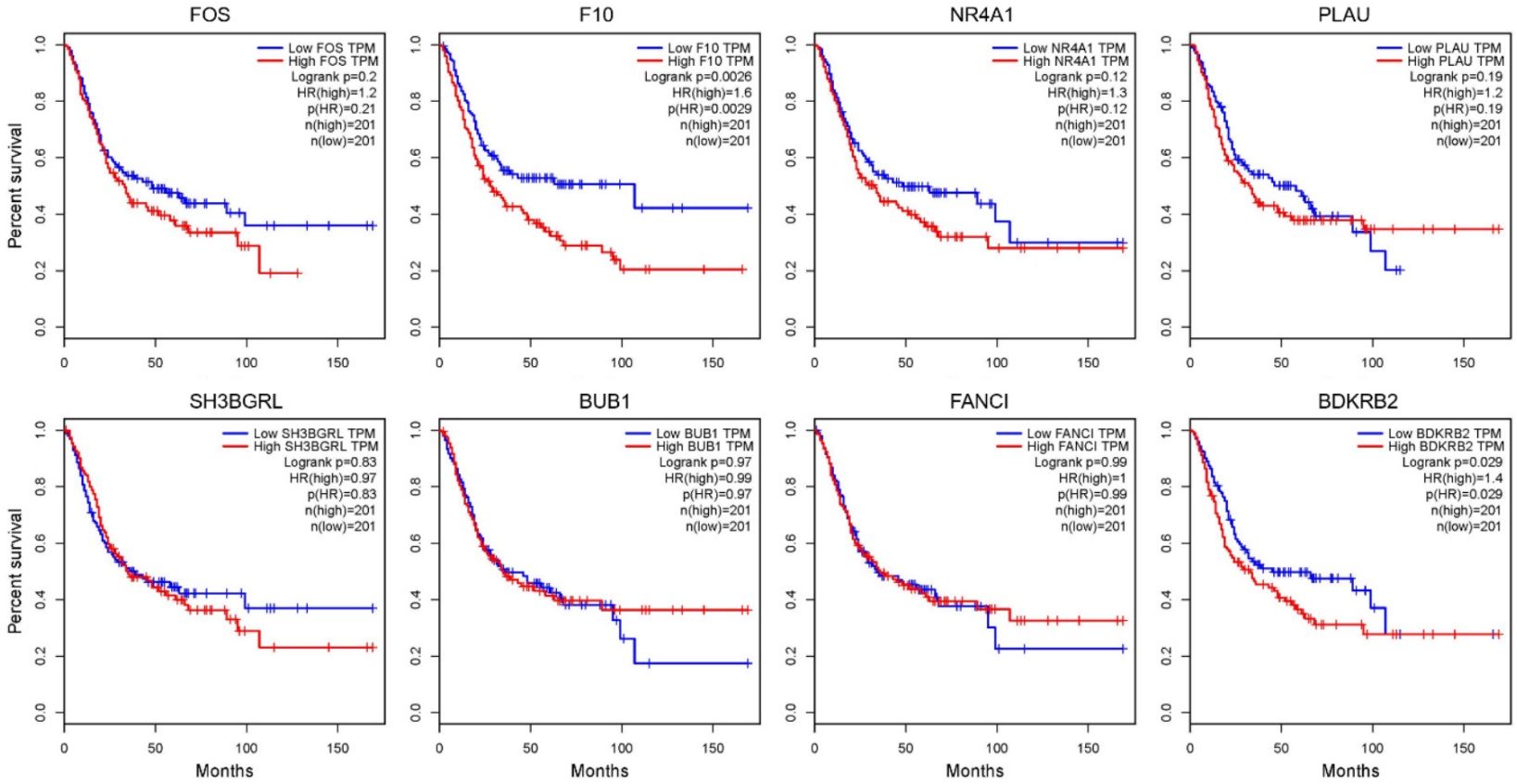

Figure 7. The overall survival of common genes from different intersections using Kaplan-Meier plotter at GEPIA. F10 and BDKRB2 were significantly correlated with decreased overall survival. $P < 0.05$ was considered statistically significant.

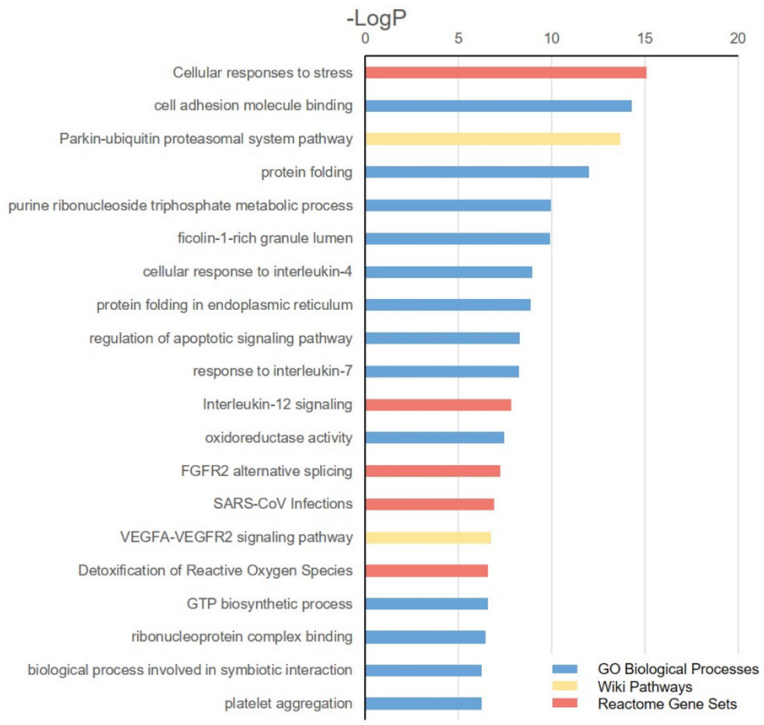

Figure 8. Signaling pathway analysis of 78 arsenic binding proteins. The top 20 pathways are shown. GO biological processes, KEGG pathway, Wiki pathway, and Reactome gene set are included and shown in different colors.

Real-time PCR assay was performed to explore the specific functions of DEGs belonging to different intersections. Human bladder cancer cell line T24 was treated with 1 μM ATO. Results show that BDKRB2, FOS, NR4A1, PLAU, and SH3BGRL were responsive to arsenic (**Figure 6**). Interestingly, the expression of BDKRB2, which may facilitate the crosstalk between carcinogenic and anticancer effects was increased at 1 h, and decreased when the treatment extended to 6 h. The survival curves between the expression of these DEGs and overall survival rate were also examined by the Kaplan-Meier method at GEPIA (**Figure 7**). Results show that high expressions of F10 and BDKRB2 were significantly correlated with decreased overall survival (P = 0.0026 and 0.0029, Hazard Ratio = 1.6 and 1.4), while the expressions of other genes were mildly correlated. Taken together, BDKRB2, FOS, NR4A1, PLAU, SH3BGRL, and F10 play an important role in the crosstalk between carcinogenic and anticancer effects in the bladder.

Putative downstream gene analysis of arsenic binding proteins

To predict direct targets of arsenic and to resolve how arsenic functions in its double-edge effect, 78 arsenic binding proteins that were confirmed by mass spectrometry [16, 17] were selected for further analysis. Signaling pathway enrichment of these proteins was performed by Metascape (**Figure 8**). The most significantly enriched clusters were cellular responses to stress (R-HSA-2262752), cell adhesion molecule binding (GO:0050839), and the parkin-ubiquitin proteasomal system pathway (WP2359). Cistrome Data Browser was used to analyze the downstream genes of these 78 arsenic binding proteins. Among them, ACTB, BACH1, NME2, RBBP4, PARP1, and PML underwent a ChIP-seq assay. Some of the putative downstream genes of these 6 arsenic binding proteins had an altered expression level and are listed in **Table 5**. Kaplan-Meier plotter analysis showed that high expressions of PAX6, MLLT11, LTBP1, PCSK5, ZFP36, and COL8A2 were significantly associated with shortened survival (**Figure 9**). However, their expressions in 4 expression groups were inconsistent (**Figure 10A**). Quantitative real-time PCR was conducted to detect the expressions of these putative downstream genes in bladder cancer cell line T24 under 1 μM ATO treatment. The expressions of IL1R2, PAX6, and PCSK5 were regulated by arsenic in T24 cells (**Figure 10B**).

Among the putative downstream genes, PAX6, LY6H, TNFRSF8, FZD9, IL1R2, and ZFP36 were also found as nodes in MCODE (**Figure 4**). Pearson correlation tests were performed in LinkedOmics. The expressions of C15ORF48 and UGCG showed stronger positive correlation with their own regulatory genes in the 408 BLCA patient dataset (**Table 5**). These results indicate that arsenic may impact the activity of ACTB, BACH1, NME2, RBBP4, PARP1, and PML by direct binding, further influencing the expression of downstream genes such as PAX6, MLLT11, LTBP1, PCSK5, ZFP36, COL8A2, LY6H, TNFRSF8, FZD9, and IL1R2, facilitating the carcinogenic and anticancer functions of arsenic.

Table 5. Arsenic binding proteins and their putative downstream genes based on Cistrome analysis

Arsenic binding proteins	1-week	2-week	12-week	HUC1
ACTB	**PAX6**, LY6H	**PAX6**, GNAT3	**PAX6**, L2HGDH, LY6H	C15ORF48 (0.31)
BACH1	MLLT11	NF	ERF, SYNGR3, TNFRSF8	**MLLT11**
NME2	NPY2R, GNMT	SLC15A1, **COL8A2**, TTR	FZD9	BASP1, IL1R2
PARP1	**LTBP1**	**PCSK5**, OXSR1	NF	NF
RBBP4	NF	NF	UGCG (0.38)	NF
PML	NF	ERBIN, **ZFP36**	NF	H2BC6

NF: not found. Genes in bold were corelated with a poor overall survival. The values in brackets represent Pearson correlation between arsenic binding proteins and their putative downstream genes.

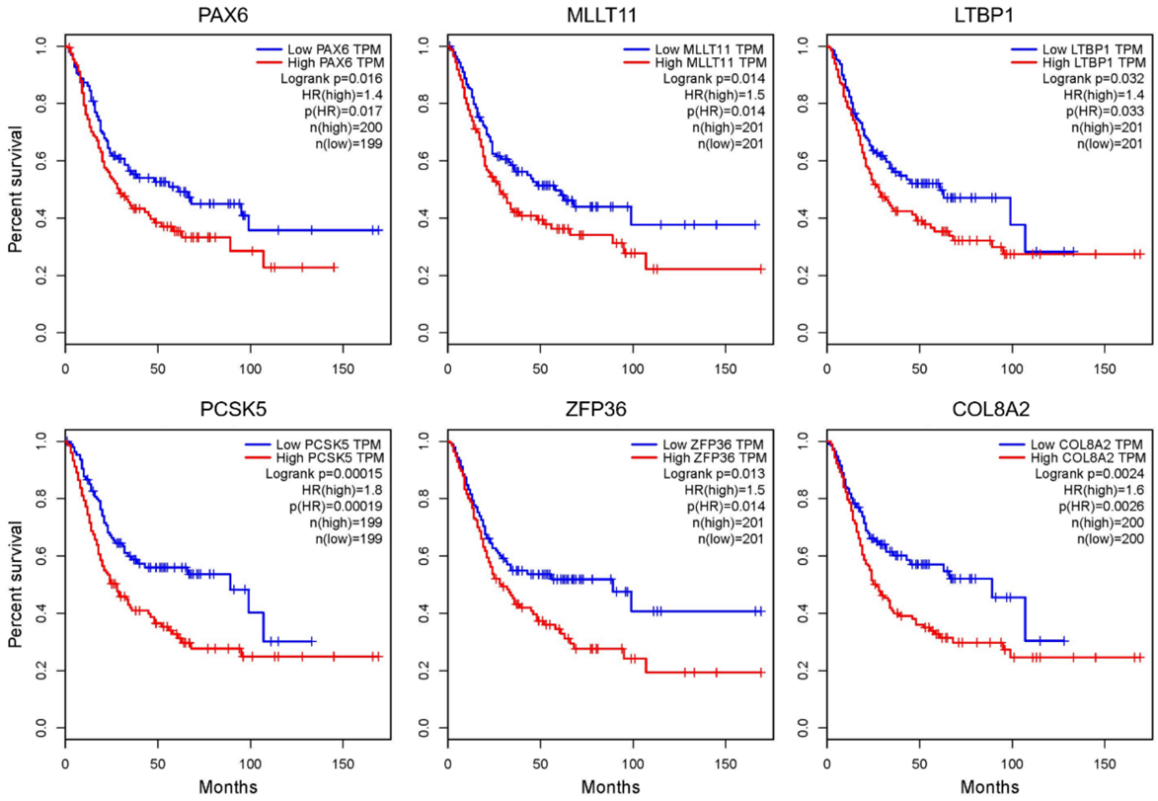

Figure 9. Overall survival of putative downstream genes using Kaplan-Meier plotter at GEPIA. $P < 0.05$ was considered significant.

Discussion

The paradoxical roles of arsenic in carcinogenic and anticancer activities have puzzled many researchers. Three gene expression datasets (accession no. GSE116554, GSE32102 and GSE90023) were obtained and were grouped into four expression matrices named 1-week, 2-week, 12-week and HUC1 (**Table 2**). Differentially expressed genes were analyzed in order to show the effect of arsenic. No common genes were found among the four groups, though *P*-value was not considered because the HUC1 group had no replicate data. More common DEGs were found between 1-week and 12-weeks which were from the same experimental dataset, GSE32102 (**Figure 2**). Less common genes were found among the other expression groups may partially result from weaker comparability of different experimental conditions.

To resolve the DEGs, the signaling pathway and PPI analysis were conducted. The highest link-

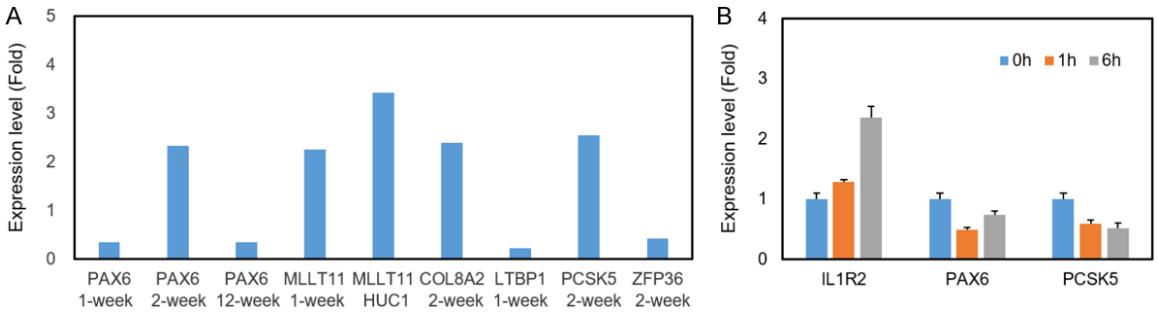

Figure 10. Expression of putative downstream genes of arsenic binding proteins in each expression matrix (A) or T24 cells examined by real-time PCR (B). Human T24 bladder cancer cells were treated with 1 μM ATO for 1 h and 6 h. Expression values were calculated using the $2^{-\Delta\Delta CT}$ method and GAPDH as the endogenous reference gene.

age hub gene at 1-week was albumin (ALB). This protein acts as a carrier protein for a wide range of endogenous molecules including hormones, fatty acids, and metabolites, as well as exogenous drugs [25]. KRAS and TNF were found in different groups (**Figure 4**). It is no doubt that these two pathways play the key roles in carcinogenesis. A previous study identified recurrent mutations in FGFR and HRAS in urothelial bladder carcinoma [26]. A transcriptomic analysis of arsenic-induced bladder carcinogenesis emphasized the cell proliferation and survival pathways, such as the MAPK, PI3K/AKT, and Hippo signaling pathways [27].

We know that arsenic impacts the target protein structure by immediately binding to Cys residues such as in PML-RAR, TP53, and Gli1. Therefore it became possible to deduce a model of action for arsenic [28]. Putative downstream genes of 6 possible arsenic binding proteins from a previous study [16, 17] and the web-portal were analyzed. Some of the downstream genes such as PAX6, MLLT11, LTBP1, PCSK5, ZFP36, and COL8A2 were significantly associated with shortened survival (**Figure 9**). Among them, PAX6, MLLT11, and ZFP36 have been reported to be associated with poor prognosis in various cancers. These may be possible mechanisms of arsenic carcinogenesis and anticancer effects.

Conclusion

Three Gene Set Enrichments (GSE) were loaded from the Gene Expression Omnibus (GEO) database and grouped into four expression matrices named 1-week, 2-weeks, 12-weeks, and HUC1 according to exposure time. The PPI network emphasized the role of KRAS and TNF signaling in different groups. Furthermore, BDKRB2, FOS, NR4A1, PLAU, SH3BGRL, and F10 played an important role in the crosstalk between carcinogenic and anticancer effects in bladder. Arsenic may impact the activity of ACTB, BACH1, NME2, RBBP4, PARP1, and PML by direct binding, further influencing the expression of downstream genes such as PAX6, MLLT11, LTBP1, PCSK5, ZFP36, COL8A2, and IL1R2.

Acknowledgements

We are grateful to Mark Goettel (Former Editor-in-Chief of the journal, Biocontrol Science and Technology) for scientific editing. This work was supported by the Stability Support Plan for Higher Education Institutions in Shenzhen (20200925160201001); Shenzhen Science and Technology Plan Project (JCYJ20220530-115203008); Open Project of Putian University Key Laboratory of Tumor Transformation Medicine (2022KF004); Dean Project of Southern University of Science and Technology Hospital (2020-A2); and Technology Major Project of Nanshan District Health System (NSZD2023-063).

Disclosure of conflict of interest

None.

Abbreviations

APL, Acute promyelocytic leukemia; ATO, Arsenic trioxide; GSE, Gene Set Enrichments; GEO, Gene Expression Omnibus; HUC1, Human urothelial cell 1; DEGs, Differential Expression Genes; iDEP, integrated Differential Expression and Pathway analysis; UALCAN, The University

of Alabama at Birmingham Cancer Data Analysis Portal; GEPIA, Gene Expression Profiling Interactive Analysis; PPI, Protein-Protein interaction; TCGA, The Cancer Genome Atlas; GO, Gene Ontology; KEGG, Kyoto Encyclopedia of Genes and Genomes; MCODE, Molecular Complex Detection; KRAS, Kirsten Rat Sarcoma Viral Oncogene; TNF, Tumor Necrosis Factor; ACTB, β-Actin; BACH1, BTB Domain and CNC Homolog 1; NME2, NME/NM23 Nucleoside Diphosphate Kinase 2; RBBP4, RB Binding Protein 4; PARP1, Poly(ADP-ribose) Polymerase 1.

Address correspondence to: Yukun Wang, Department of Pharmacology, School of Medicine, Southern University of Science and Technology, Shenzhen 518055, Guangdong, China. E-mail: wangyk@sustech.edu.cn

References

[1] Chen QY and Costa M. Arsenic: a global environmental challenge. Annu Rev Pharmacol Toxicol 2021; 61: 47-63.

[2] Smith AH, Marshall G, Roh T, Ferreccio C, Liaw J and Steinmaus C. Lung, bladder, and kidney cancer mortality 40 years after arsenic exposure reduction. J Natl Cancer Inst 2018; 110: 241-249.

[3] Koutros S, Lenz P, Hewitt SM, Kida M, Jones M, Schned AR, Baris D, Pfeiffer R, Schwenn M, Johnson A, Karagas MR, Garcia-Closas M, Rothman N, Moore LE and Silverman DT. RE: elevated bladder cancer in Northern New England: the role of drinking water and arsenic. J Natl Cancer Inst 2018; 110: 1273-1274.

[4] Abtahi M, Dobaradaran S, Koolivand A, Jorfi S and Saeedi R. Assessment of cause-specific mortality and disability-adjusted life years (DALYs) induced by exposure to inorganic arsenic through drinking water and foodstuffs in Iran. Sci Total Environ 2023; 856: 159118.

[5] Pál L, Jenei T, McKee M, Kovács N, Vargha M, Bufa-Dőrr Z, Muhollari T, Bujdosó MO, Sándor J and Szűcs S. Health and economic gain attributable to the introduction of the World Health Organization's drinking water standard on arsenic level in Hungary: a nationwide retrospective study on cancer occurrence and ischemic heart disease mortality. Sci Total Environ 2022; 851: 158305.

[6] Lamm SH, Boroje IJ, Ferdosi H and Ahn J. A review of low-dose arsenic risks and human cancers. Toxicology 2021; 456: 152768.

[7] Mehus AA, Bergum N, Knutson P, Shrestha S, Kalonick M, Zhou X, Garrett SH, Sens DA, Sens MA and Somji S. Chronic arsenic exposure upregulates the expression of basal transcriptional factors and increases invasiveness of the non-muscle invasive papillary bladder cancer line RT4. Int J Mol Sci 2022; 23: 12313.

[8] Islam R, Zhao L, Wang Y, Lu-Yao G and Liu LZ. Epigenetic dysregulations in arsenic-induced carcinogenesis. Cancers (Basel) 2022; 14: 4502.

[9] Saintilnord WN and Fondufe-Mittendorf Y. Arsenic-induced epigenetic changes in cancer development. Semin Cancer Biol 2021; 76: 195-205.

[10] Islam R, Zhao L, Zhang X and Liu LZ. MiR-218-5p/EGFR signaling in arsenic-induced carcinogenesis. Cancers (Basel) 2023; 15: 1204.

[11] Wadgaonkar P and Chen F. Connections between endoplasmic reticulum stress-associated unfolded protein response, mitochondria, and autophagy in arsenic-induced carcinogenesis. Semin Cancer Biol 2021; 76: 258-266.

[12] Zhang XW, Yan XJ, Zhou ZR, Yang FF, Wu ZY, Sun HB, Liang WX, Song AX, Lallemand-Breitenbach V, Jeanne M, Zhang QY, Yang HY, Huang QH, Zhou GB, Tong JH, Zhang Y, Wu JH, Hu HY, de Thé H, Chen SJ and Chen Z. Arsenic trioxide controls the fate of the PML-RARalpha oncoprotein by directly binding PML. Science 2010; 328: 240-243.

[13] Chen S, Wu JL, Liang Y, Tang YG, Song HX, Wu LL, Xing YF, Yan N, Li YT, Wang ZY, Xiao SJ, Lu X, Chen SJ and Lu M. Arsenic trioxide rescues structural p53 mutations through a cryptic allosteric site. Cancer Cell 2021; 39: 225-239, e8.

[14] Beauchamp EM, Ringer L, Bulut G, Sajwan KP, Hall MD, Lee YC, Peaceman D, Ozdemirli M, Rodriguez O, Macdonald TJ, Albanese C, Toretsky JA and Uren A. Arsenic trioxide inhibits human cancer cell growth and tumor development in mice by blocking Hedgehog/GLI pathway. J Clin Invest 2011; 121: 148-160.

[15] Kaiming C, Sheng Y, Zheng S, Yuan S, Huang G and Liu Y. Arsenic trioxide preferentially binds to the ring finger protein PML: understanding target selection of the drug. Metallomics 2018; 10: 1564-1569.

[16] Nan K, He M, Chen B and Hu B. Analysis of arsenic binding proteins in HepG2 cells based on a biotinylated phenylarsenite probe. Anal Chim Acta 2021; 1183: 339007.

[17] Hu X, Li H, Ip TK, Cheung YF, Koohi-Moghadam M, Wang H, Yang X, Tritton DN, Wang Y, Wang Y, Wang R, Ng KM, Naranmandura H, Tse EW and Sun H. Arsenic trioxide targets Hsp60, triggering degradation of p53 and survivin. Chem Sci 2021; 12: 10893-10900.

[18] Ge X. iDEP web application for RNA-Seq data analysis. Methods Mol Biol 2021; 2284: 417-443.

[19] Chen C, Chen H, Zhang Y, Thomas HR, Frank MH, He Y and Xia R. TBtools: an integrative toolkit developed for interactive analyses of big biological data. Mol Plant 2020; 13: 1194-1202.
[20] Zhou Y, Zhou B, Pache L, Chang M, Khodabakhshi AH, Tanaseichuk O, Benner C and Chanda SK. Metascape provides a biologist-oriented resource for the analysis of systems-level datasets. Nat Commun 2019; 10: 1523.
[21] Szklarczyk D, Kirsch R, Koutrouli M, Nastou K, Mehryary F, Hachilif R, Gable AL, Fang T, Doncheva NT, Pyysalo S, Bork P, Jensen LJ and von Mering C. The STRING database in 2023: protein-protein association networks and functional enrichment analyses for any sequenced genome of interest. Nucleic Acids Res 2023; 51: D638-D646.
[22] Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, Amin N, Schwikowski B and Ideker T. Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome Res 2003; 13: 2498-2504.
[23] Chandrashekar DS, Bashel B, Balasubramanya SAH, Creighton CJ, Ponce-Rodriguez I, Chakravarthi BVSK and Varambally S. UALCAN: a portal for facilitating tumor subgroup gene expression and survival analyses. Neoplasia 2017; 19: 649-658.
[24] Zheng R, Wan C, Mei S, Qin Q, Wu Q, Sun H, Chen CH, Brown M, Zhang X, Meyer CA and Liu XS. Cistrome Data Browser: expanded datasets and new tools for gene regulatory analysis. Nucleic Acids Res 2019; 47: D729-D735.
[25] De Simone G, di Masi A and Ascenzi P. Serum albumin: a multifaced enzyme. Int J Mol Sci 2021; 22: 10086.
[26] Cancer Genome Atlas Research Network. Comprehensive molecular characterization of urothelial bladder carcinoma. Nature 2014; 507: 315-322.
[27] Shukla V, Chandrasekaran B, Tyagi A, Navin AK, Saran U, Adam RM and Damodaran C. A comprehensive transcriptomic analysis of arsenic-induced bladder carcinogenesis. Cells 2022; 11: 2435.
[28] Chang KJ, Yang MH, Zheng JC, Li B and Nie W. Arsenic trioxide inhibits cancer stem-like cells via down-regulation of Gli1 in lung cancer. Am J Transl Res 2016; 8: 1133-1143.
[29] Jou YC, Wang SC, Dai YC, Chen SY, Shen CH, Lee YR, Chen LC and Liu YW. Gene expression and DNA methylation regulation of arsenic in mouse bladder tissues and in human urothelial cells. Oncol Rep 2019; 42: 1005-1016.
[30] Clewell HJ, Thomas RS, Kenyon EM, Hughes MF, Adair BM, Gentry PR and Yager JW. Concentration- and time-dependent genomic changes in the mouse urinary bladder following exposure to arsenate in drinking water for up to 12 weeks. Toxicol Sci 2011; 123: 421-432.
[31] Ooki A, Del Carmen Rodriguez Pena M, Marchionni L, Dinalankara W, Begum A, Hahn NM, VandenBussche CJ, Rasheed ZA, Mao S, Netto GJ, Sidransky D and Hoque MO. YAP1 and COX2 coordinately regulate urothelial cancer stem-like cells. Cancer Res 2018; 78: 168-181.