

## **Arsenite exposure inhibits histone acetyltransferase p300 for attenuating H3K27ac at enhancers in low-dose exposed mouse embryonic fibroblast cells**

Yan Zhu\* a, Yanqiang Li\* a, Dan Lou\*, Yang Gao\*, Jing Yu\*, Dehui Kong $\beta$ , Qiang Zhang $\Sigma$ , Yankai Jia $\varphi$ , b, Haimou Zhang $\beta$ , b, and Zhibin Wang $\beta$ , $\delta$ , b, \*

\* Laboratory of Human Environmental Epigenome, Department of Environmental Health & Engineering, Bloomberg School of Public Health, Johns Hopkins University, 615 N. Wolfe St., Baltimore, MD, USA 21205

$\beta$  School of Life Sciences, Hubei University, Wuhan, Hubei Province, China.

$\Sigma$  Department of Occupational and Environmental Health, School of Public Health, Tianjin Medical University. No. 22, Qixiangtai Road, Tianjin, China 300070

$\varphi$  GENEWIZ Suzhou, 218 Xinghu Road, Suzhou Industrial Park, Suzhou, China, 215123;

$\delta$  The Sidney Kimmel Comprehensive Cancer Center and Department of Oncology, School of Medicine, Johns Hopkins University, Baltimore, MD, USA 21205

a Contribute equally

b Correspondence should be addressed to

Dr. Zhibin Wang zwang47@jhu.edu Phone (410)-955-7840

Dr. Haimou Zhang haimou@hubu.edu.cn

Dr. Yankai Jia yankai.jia@genewiz.com

## Abstract

Epidemiological investigations have validated the positive relationships between arsenic in drinking water and several cancers, including skin, liver and lung cancers. Besides genotoxicity, arsenic exposure-related pathogenesis of disease is widely considered through epigenetic mechanisms; however, the underlying mechanistic insight remains elusive. Herein we explore the initial epigenetic changes via acute low-dose arsenite exposures of mouse embryonic fibroblast (MEF) cells and Dot1L knockout MEF (Dot1L-/- for abbreviation) cells. Our RNA-seq and Western blot data demonstrated that, in both cell lines, acute low-dose arsenite exposure abolished histone acetyltransferase p300 at the RNA level and subsequent protein level. Consequently, p300-specific main target histone H3K27ac, a marker separating active from poised enhancers, decreased dramatically as validated by both Western blot and ChIP-seq analyses. Concomitantly, H3K4me1 as another well-known marker for enhancers also showed significant decreases, suggesting an underappreciated crosstalk between H3K4me1 and H3K27ac involved in arsenite exposure. Significantly, arsenite exposure-reduced H3K27ac and H3K4me1 inhibit the expression of genes including *EP300* itself and *Kruppel Like Factor 4 (Klf4)*, a tumor suppressor gene. Collectively, our investigations identify p300 as an internal bridging factor within cells to sense external environmental arsenite exposure to alter chromatin, thereby changing gene transcription for disease pathogenesis.

## Keywords:

Arsenite exposure, histone acetyltransferase p300, H3K27ac, enhancers, epigenetic mechanism

## 1. Introduction

Arsenic, as a ubiquitous element, ranks the 20th most abundant element in the earth's crust (Mandal and Suzuki, 2002). Human exposure to environmental arsenic occurs primarily via ingestion of inorganic arsenic contaminated water. It is estimated that over 200 million individuals worldwide are exposed to drinking water containing arsenic above 10 µg/L, the safety standard set by the World Health Organization (Argos *et al.*, 2014). Through numerous epidemiological investigations, accumulating evidence demonstrate the positive relationship between inorganic arsenic and carcinomas, including both skin and internal cancers. In Wisconsin, USA, residents

who were over 35 years old and consumed arsenic-contaminated water for more than a decade suffered a higher prevalence of skin cancer than control group (Knobeloch *et al.*, 2006). A significant dose-response relationship between arsenic concentrations and other tumors was found in Taiwan, including bladder, kidney, and lung cancer in both sexes, and tumors of prostate and liver in males (Wu *et al.*, 1989). However, the exact pathogenic mechanisms underlying inorganic arsenic exposure remain to be determined. Inorganic arsenic shows genotoxicity through several aspects, including chromosomal aberration (Oya-Ohta *et al.*, 1996), genomic instability (Bhattacharjee *et al.*, 2013) and p53 dysfunction (Hamadeh *et al.*, 1999). Besides, the role of inorganic arsenic in epigenome has also been initially explored, including DNA methylation, microRNA, and histone modification. The methylation of CpG dinucleotides at the 5' position on the pyrimidine ring to form 5-methylcytosine (5-mC), which can disrupt transcription by blocking the binding of transcription factors and attracting methyl-binding proteins that initiate chromatin compaction and bring about gene silencing (Lunnon and Mill, 2013). So methylation results in gene silencing, while de-methylation causes gene activation (Tellez-Plaza *et al.*, 2014). S-adenosyl-methionine (SAM) is essential for methylation of inorganic arsenic to detoxication, and it is also the methyl-donor required by DNA methyltransferases. And it was validated that arsenic exposure was associated with hypo-methylation of DNA and facilitates oncogenes expression (Zhao *et al.*, 1997). Besides, hypermethylation of *p53* and *p16* promoters were found in people exposed to arsenic. Histone modification (including histone acetylation and methylation), a covalent post-translational modification, plays an integral role in gene transcription (Allfrey and Mirsky, 1964; Dai H, 2014). For instance, chronic arsenite exposure decreased H4K16ac in UROtsa cells (Jo *et al.*, 2009). However, systematic characterizations to identify effector histone acetyltransferases and altered genomic loci with reduced histone acetylations remain to be done.

For histone acetylation, it is regulated by two group of enzymes, histone acetyltransferases (HATs) and deacetylases (HDACs), with antagonizing functions (Wang *et al.*, 2009). Among members of HATs, p300 and CBP are homologous. They are also global transcriptional co-activators (Ogryzko *et al.*, 1996), and both of them play integral roles in regulation of gene transcription via at least two (not necessarily mutual exclusive) mechanisms. In scenario 1, p300 acts as associated partners of transcription factors (TFs). p300 is recruited by these TFs including p53 to acetylate

nucleosomes at promoters and/or enhancers for opening chromatin, thereby facilitating the recruitment of RNA polymerase II and other general transcription factors (Guermah *et al.*, 2006; Wang *et al.*, 2009). For example, our data demonstrate that the increases of histone H3K9ac and H4K16ac at transcription start sites, via inhibition of HDACs with HDAC inhibitor drugs, facilitate the RNA polymerase II recruitment. In scenario 2, p300 functions through acetylation of TFs themselves, thereby impacting the function of TFs. For example, p300 directly acetylates exposure-relevant TF Nrf2. Acetylations of multiple lysine residues of Nrf2 increase the binding of Nrf2 to promoters (Sun *et al.*, 2009). This increased Nrf2 binding presumably contributes to the regulation of Nrf2-regulated gene pathways. In addition, the loss of p300 functions or simple reduction of p300 protein levels has linked to human diseases. For example, DNA mutations of p300 are linked to a small percentage of patients of Rubinstein-Taybi syndrome (Roelfsema *et al.*, 2005). Such mutations lead to either truncated non-functional p300 molecules, or the loss of one copy of *EP300* gene in each cell, thereby reducing the p300 amount by half. While the reduction of p300 amount is linked to development of Rubinstein-Taybi syndrome, the exact underlying mechanism is not clear. Lastly, it seems that p300/CBP specifically mediate H3K18ac and H3K27ac, whereas two other homologous HATs, GCN5 and PCAF, mediate H3K9ac (Jin *et al.*, 2011).

Herein our analyses reveal that sodium arsenite exposure-related health problems may be linked to the reduction of the amount of p300 in exposed cells. We found that acute arsenite exposure in low doses reduced the expression of p300 at both the mRNA levels via RNA-seq and qRT-PCR analyses and the protein levels via Western blotting. The consequence of reduced p300 (i.e., reduction of H3K27ac, an active marker for enhancers) was further identified. In addition, we examined the expression of tumor suppressor genes. Collectively, our data reveal an intriguing epigenetic mechanism that arsenic exposure may impact human health through altering functions of p300 and following histone acetylation modifications it mediated.

## 2. Materials and methods

### 2.1. Cell culture and arsenite exposure

MEF and Dot1L<sup>-/-</sup> are diploid and nontumorigenic. Both cells were cultured in high glucose Dulbecco's Modified Eagle Medium (DMEM) with 10% (v/v) fetal bovine serum (FBS) in an

incubator at 37 °C with 5% CO<sub>2</sub>. At approximately 80% confluent, cells were treated with fresh medium containing 0, 2, 5, 10 µM sodium arsenite for 24 h. After exposure, cells were collected for following experiments.

## 2.2. *Western Blot*

Total protein extracts were prepared from cells and loaded into each lane of a 4-12% SDS-PAGE gel as described before (Wang *et al.*, 2008). After electrophoresis, total proteins were transferred onto a nitrocellulose membrane. Membranes were blocked in TBS-T buffer (20 mM Tris base, 150 mM NaCl, 0.05% Tween 20, pH 7.4) containing 5% nonfat milk at room temperature for 2-4 h prior to the addition of primary antibody and incubation at 4°C overnight. Nrf2 antibody (Santa Cruz, sc-13032, Lot # C1416), p300 antibody (Santa Cruz, sc-584, Lot # C0805), H3K4me1 antibody (Abcam, ab8895, Lot # GR243233-1) and H3K27ac antibody (Abcam, ab4729, Lot # GR103379-1) were used to detect each protein. TBS-T-washed membranes were incubated with goat anti-rabbit (Santa Cruz, sc-2004, Lot # A1416) or goat anti-mouse (Santa Cruz, sc-2005, Lot # J2215) secondary antibodies in TBS-T-5% nonfat milk and washed with TBS-T, and membrane-bound antibody was visualized in a chemiluminescence assay. Beta-actin (Sigma, A5441, Lot # 030M4788) was used for normalization.

## 2.3. *mRNA extraction and characterization*

Total RNAs from one exposed and three control groups of two cell lines were isolated and quality controlled with protocols from Illumina. A total amount of 2 µg RNA per sample was used as input material for library construction. Ribosomal RNA (rRNA) was removed through Epicentre Ribo-zero rRNA Removal Kit (Epicentre, USA) according to the manufacturer's instructions. Strand-specific sequencing libraries were generated through the dUTP method by using the resulting RNA from NEBNext Ultra Directional RNA Library Prep Kit for Illumina (NEB, USA) following manufacturer's recommendations. RNA-seq was performed on an Illumina Hiseq 2000 platform and 100 bp paired-end reads were generated according to Illumina's protocol.

## 2.4. *Whole-genome gene expression analysis*

The adapter sequences were removed from the raw sequencing data and the individual libraries were converted to the FASTQ format. Sequence reads were aligned to the mouse genome (mm10) with TopHat2 (v2.0.9) and the resulting alignment files were reconstructed with Cufflinks (v2.1.1) and Scripture (beta2). For mRNA analyses, the RefSeq database (Build 37.3) was chosen as the annotation references. The read counts of each transcript were normalized to the length of the individual transcript and to the total mapped fragment counts in each sample and expressed as fragments per kilobase of exon per million fragments mapped (FPKM) of mRNAs in each sample. The mRNA differential expression analyses were applied for exposed and control groups. An adjusted P value <0.05 (Student's t-test with Benjamini-Hochberg false discovery rate (FDR) adjustment) was used as the cut-off for significantly differentially expressed genes. Differentially expressed genes (DEGs) were analyzed by enrichment analyses to detect over-represented functional terms present in the genomic background. Gene ontology (GO) analysis was performed using the GO-seq R package, in which gene length bias was corrected.

## 2.5. Gene expression validation by quantitative Reverse Transcription-PCR (qRT-PCR)

Total RNAs were isolated and quality controlled using protocols from Illumina, Inc. All cDNAs were synthesized using Superscript III (Invitrogen; Thermo Fisher Scientific, Inc., Waltham, MA, USA). Five over- and under-expressed genes were selected for qRT-PCR assay. All qRT-PCR primers are presented in Table 1. All qPCR reactions were performed on a Roche Lightcycler 480 PCR system (Roche Applied Science, Penzberg, Germany) using Toyobo Thunderbird SYBR RT-qPCR Mix (Toyobo Life Science, Osaka, Japan), with three technical repeats. The amplification procedure was as follows: 95 °C for 5 min, followed by 40 cycles of 95 °C for 10 sec and 60 °C for 20 sec. Relative quantification of target genes was performed using the 2- $\Delta\Delta CT$  method with GAPDH as a reference gene (Livak and Schmittgen, 2001). Pearson correlation was used to calculate the association between RNA sequence and qRT-PCR (Hane *et al.*, 1993).

Table 1 Primers used for qRT-PCR

GAPDH	F: TGCACCACCAACTGCTTAGC R: GGCATGGACTGTGGTCATGAG
Ndufa5	F: GGTGTGCTGAAGAAGACAC R: GGTTCCGCTTAAACCATAGCC
Hsd17b10	F: CGGTAAATAACCGGAGGAGCC R: TGGTCATCACCCGGATACCT
Ndufa11	F: GTCACACTCAATCCTCCGGG

	R: CGCTATGCCAAAGTACACGC
Cox8a	F: CGCCAAGATCCATTGTTGC R: TCACGAAGCAGGAGGTAAGC
Ndufa3	F: CCCCATTGAGCCCCTACTTC R: CGGGCACTGGTAGTTGTAG
EP300	F: CCCCAGATGGGAGGACAAAC R: ACTGGCTCCAATCTGCTGTC
Clock	F: GGCCACTATGTGAGAACCCC R: GACAGTCGTCCACGTTCACT
Postn	F: CAACGCAGCGCTATTCTGAC R: TCGGAAGCCACTTGTCTCC
Notch1	F: GAATGGCGGGAAAGTGTGAAG R: CACAGCTGCAGGCATAGTCT
Klf4	F: GGGAGAAGACACTGCGTCAA R: GGAAGTCGCTTCATGTGGGA
Crebbp	F: TGAGAACTTGCTGGACGGAC R: CACTGAGGCTGGCCATGTTA
UTX	F: AGCTTTGTCGAGCCAAGGA R: GCATTGGACAAAGTGCAGGG
Sirt1	F: GACTCCAAGGCCACGGATAG R: TGTCGAGGATCTGTGCCAA
Hdac4	F: GCAGATCCAGAGGCAGATCC R: TTTGGCGTCGTACATCCCCA

## 2.6. Chromatin immunoprecipitation sequencing (ChIP-Seq)

Chromatin immunoprecipitation (ChIP) was performed as described previously (Wang *et al.*, 2009). In brief, cells were crosslinked with 1% formaldehyde for 10 min at room temperature and the reaction was quenched with 125 mM glycine. After chromatin fragmented to 200 to 500 bp by sonication, chromatin templates from 20 million cells were used for ChIP experiment. Samples were immunoprecipitated with 2-4 µg of H3K4me1 antibody or H3K27ac antibody overnight at 4 °C followed by washing steps. After reverse crosslinking, the ChIP DNA fragment were purified and repaired using End-It™ DNA End-Repair Kit (Epicentre) followed by treatment with Taq polymerase to generate single-base 3' overhangs used for adaptor ligation. Following ligation of a pair of Illumina adaptors to the repaired ends, the ChIP DNA was amplified using the adaptor primers. Amplified DNA fragments around 300 bp were isolated from agarose gel prior to sequencing.

## 2.7. Identification of binding regions and peaks

For ChIP-seq of H3K4me1 and H3K27ac, reads were mapped to the mm10 genome using Bowtie aligner allowing up to two mismatches (Langmead, 2010). Only the uniquely mapping reads were used for further analysis. Macs2 call peak program was used to call the peaks and the macs2 bdgdiff was used to identify the difference between MEF and Dot1L (Zhang *et al.*, 2008).

### 2.8. ChIP-qPCR validation

Five genes' promoters were selected for qRT-PCR assay and presented in Table 2. All qPCR reactions were performed on a Roche Lightcycler 480 PCR system (Roche Applied Science, Penzberg, Germany) using Toyobo Thunderbird SYBR RT-qPCR Mix (Toyobo Life Science, Osaka, Japan), with three technical repeats. The amplification procedure was as follows: 95 °C for 5 min, followed by 40 cycles of 95 °C for 10 sec and 60 °C for 20 sec. The cycle threshold (Ct) of samples was used for calculation. Each value was normalized to the percentage of Input DNA by using  $IP/INPUT = 2^{CtINPUT - CtIPDNA}$  (Alamdar *et al.*, 2017).

Table2 Primers used for ChIP-qPCR validation

Gene	Location (5'-3')	Sequence	Size (bp)
<i>Klf4</i> promoter	-1000 to -500	F:CTAACCGTTGGCGTGAGGAA R:CGGGTTGTTACTGCTGCAAG	137
<i>EP300</i> promoter	-500 to 0	F:TGGCACAGATTGGTTCACTG R: TCGCCACCATTGGTTAGTCC	125
<i>Nqo1</i> promoter	-500 to 0	F: GCGAGAAGAGCCCTGATTGT R: CAATATCTGGCTCAGGCGT	151
<i>Notch2</i> promoter	-1000 to -500	F: CAGCACTGTGACAGCCCTTA R: ATAGCCTCCGTTCGGTTGG	221
<i>Igf2r</i> promoter	-1500 to -1000	F: ACCAGACACTCCGGTACTCA R: TGACGAGCCAACACAGACAG	277

### 2.9. Statistical analysis

Results were presented in Means  $\pm$  standard error of the mean (SEM). Student's t-test was used to examine difference between two groups and ANOVA was used for multi-groups. \* means  $p < 0.05$ , \*\*means  $p < 0.01$ .

### 3. Results

#### 3.1. Identification of a low dose exposure condition

Toward a low dose exposure condition for epigenetic insights of arsenite exposure, we set up exposure experiments and used Nrf2 expression as a readout to characterize. Based on previous reports, we started our exposure experiments with two time points (6 and 24 h) and four concentrations (0, 2, 5, and 10  $\mu$ M). Without arsenite exposure, there was limited amount of Nrf2 expressed in both MEF and Dot1L-/- cells (Fig. 1A and 1B). After 6 h treatment, the amount of Nrf2 showed dose-effect relationships with the concentrations of sodium arsenite in both cell lines (Fig. 1C). In addition, there was a significant difference between 2  $\mu$ M and 5  $\mu$ M groups. Results after 24 h exposure (Fig. 1D and 1 E) were similar to 6 h treatment. However, compared with 6 h groups, which showed same results in both cells, 24 h groups (Fig. 1F) showed different pattern of Nrf2 concentrations, which might be caused by *Dot1L* lacking. Collectively, we consider 5  $\mu$ M and 24 h appropriate for our mechanistic exploration and choose this exposure condition for further experimental characterizations.

#### 3.2. Acute low dose exposure of mouse MEF cells for mechanistic insights via RNA-seq analyses

To identify an intermediate factor bridging (external) environmental exposure and (internal) cellular pathway changes, we focused on transcriptomic changes. In other word, we would like to identify the bridging factor(s) based on transcript changes. To our goal, we purified mRNAs from exposed cells and followed our established protocols for RNA-seq library construction (Li *et al.*, 2015). Sequenced reads were aligned to the mouse genome, and expression changes were analyzed. Results demonstrated that arsenite treatment upregulated 518 genes in MEF and 696 genes in Dot1L-/- MEF cells and the overlapping genes were 339 in two cell lines (Fig. 2A). Arsenite treatment also downregulated genes, with 158 genes in MEF and 413 genes in Dot1L-/- MEF cells and the overlapping number was 101 (Fig. 2B). To further validate our RNA-seq results,

we selected 10 genes (Up-regulated genes were chosen among genes that played roles in response to oxidative stress and down-regulated genes were chosen among gene that played roles in epigenetics or tumor initiation) for qRT-PCR analyses in both cell lines and confirmed their expression changes (Fig. 2C and 2D). R squares from Pearson correlation analyses were 0.85 for MEF cell and 0.90 for Dot1L-/- cell, suggesting the results from our RNA-seq analyses reliable (Fig. 2E and 2F).

### 3.3. Arsenite exposure inhibited histone acetyltransferase p300

Among these 101 overlapping genes that were inhibited by arsenite treatment in two cell lines, we noticed an intriguing candidate, *EP300* (Log2 fold, -3.5 in MEF cell and -4.6 in Dot1L-/-) (Fig. 2C and 2D). *EP300* encodes histone acetyltransferase p300 (also known as E1A-associated protein p300), essential for regulating cell growth and differentiation and preventing the growth of tumors. For the latter, many have reported the role of p300 as a tumor suppressor (Iyer *et al.*, 2004)

We first did qRT-PCR (Fig. 2C and 2D) and validated the observed changes from RNA-seq analyses. Because reduced RNA level may not reflect the actual protein level and the protein executes enzymatic activity, we next did Western blot to reveal the exact p300 protein level within cells. Western blot from three biological replicates demonstrated that arsenite exposure indeed reduced p300 protein level and showed dose-effect relationships with arsenite concentrations (Fig. 3A and 3B). In combination with RNA-seq and qRT-PCR data, we conclude that arsenite exposure reduces the expression of *EP300* initially at the transcription level and eventually at the protein level.

### 3.4. Altered histone acetylation H3K27ac, an enhancer marker

Next, we determined the consequence of the abolishment of p300 due to arsenite exposure. As a HAT, p300 is expected to acetylate nucleosomal tails. Though p300 seems to acetylate many residues of histone H3 or H4 especially from in vitro assays, there are also reports suggesting H3K27ac as p300-specific mark (Li *et al.*, 2011; Tang *et al.*, 2013). Therefore, we examined the changes of H3K27ac after arsenite exposure. With the increasing concentration of arsenite, our data demonstrated that the levels of H3K27ac decreased accordingly (Fig. 3A and 3C) from three

replicates. These data demonstrate that arsenite exposure caused decreases of histone H3K27ac, as a consequence of diminished p300.

### 3.5. Arsenite exposure decreased H3K4me1, also a well-known enhancer marker

In epigenetic community, both H3K27ac and H3K4me1 are two well-known markers for enhancers. With H3K27ac down, we asked the extent to which H3K4me1 changes accordingly. To our aim, we checked the signals of H3K4me1 in the same samples. Intriguingly, our Western blot results from three replicates reveal the concomitant decreases of H3K4me1 (Fig. 3A and 3D). These data collectively suggest that arsenite exposure involves underappreciated crosstalk between two enhancer markers, H3K27ac and H3K4me1.

### 3.6. Identification of vulnerable genomic loci with arsenite exposure-reduced H3K27ac

Data of our Western blotting assays reveal the changes of histone acetylation at the global level. To identify the vulnerable genomic loci with altered chromatin after arsenite exposure, we did ChIP-seq analyses using a previously characterized anti-H3K27ac (Wang *et al.*, 2008; Wang *et al.*, 2009). Following our established pipeline, we analyzed H3K27ac ChIP-seq signals genomewide. As expected, arsenite exposed sample has much less H3K27ac peaks compared to control samples. There are 4,772 peaks specific to exposed sample, whereas 18,302 peaks specific to control sample. Two samples shared 21,548 peaks (Fig. 4). These data are consistent with our Western blotting results (Fig. 3).

### 3.7 Decreased H3K27ac and H3K4me1 inhibited gene expression

Having established that arsenite exposure reduced H3K27ac and H3K4me1 (Fig. 3 and 4), we next determined the reductions on transcription of selected genes. Snapshots of ChIP-seq results were presented for genes including *EP300*, *Klf4*, *Notch2*, and *Igf2r* (Fig.5). These data demonstrate that reduction of H3K27ac repressed the transcripts of these genes. Notice the different scales used for presenting RNA-seq data.

To provide independent confirmation of H3K27ac at these loci and provide additional information of H3K4me1, we did ChIP-qPCR analyses. Results from Fig. 6 showed the lower enrichment of five gene promoters after arsenic exposure in MEF cell, which were in accordance

with ChIP-seq results. *Klf4*, a tumor suppressor gene, showed de-acetylation of H3K27ac and down-regulation (Log2 fold, -5.0), which was validated through ChIP-qPCR (Fig. 3C) after arsenite exposure. Accompany to the decreases of H3K27ac at these loci, our ChIP-qPCR demonstrated the decreases of H3K4me1. These data convinced us that arsenite exposure inhibited p300 protein level within cells, thereby impacting histone markers for enhancers and transcription start sites for gene transcription.

#### 4. Discussion

Considering the expensive nature of deep sequencing and our focus of mechanistic insights to initiate the understanding of epigenomic mechanisms during arsenic exposure, we set out experiments for identification of one low dose condition for future acute exposure. One more reasonable assumption is that exposure at high doses and low doses highly likely shares similar epigenetic mechanism. To our goal, we choose 0, 2, 5, and 10  $\mu$ M sodium arsenite based on published papers (Shi *et al.*, 2004; Chervona *et al.*, 2012a; Wang *et al.*, 2013; Eckstein *et al.*, 2017) for treatment.

We choose to use MEF cells for a few reasons: First, MEF cell line is a common model used in toxicology. Second, though still very preliminary, H3K79 methylation may be associated with arsenic exposure (Howe *et al.*, 2017). Meanwhile, we have Dot1L knockout MEF cells with depleted histone mark H3K79 dimethylataion (H3K79me2). Therefore, treatments to both MEF and Dot1L-/- help to determine the crosstalk between arsenic exposure and histone H3K79me2 marks in cells. Third, treatments of two independent cell lines would serve as replicates for us. Currently, it is still not conclusive about inorganic arsenic exposure changing histone acetylations for altering gene expression. For example, inorganic arsenic exposure showed H3K27ac and H3K18ac in negative association in females, but positive association in males (Chervona *et al.*, 2012b). Sodium arsenite exposure increased H3K9ac after 24 h exposure in HepG2 cells (Ramirez *et al.*, 2008), which is in dramatic contrast to the decrease of H3K9ac in arsenic exposed mice (Cronican *et al.*, 2013; Pournara *et al.*, 2016). Previous investigations also reported that the chronic arsenic exposure decreased H4K16ac in human UROtsa cells (Jo *et al.*, 2009), via affecting one of the HATs, MYST1 (also known as MOF). MOF forms two distinct multiprotein complexes (MSL and NSL) that both acetylate H4K16ac. Herein, through three biological

replicates, we demonstrated that acute arsenic exposure decreased p300 in both mRNA and protein levels (Fig. 2C, 2D, 3A). Consequently, p300-specific H3K27ac was reduced as demonstrated by both ChIP-seq and ChIP-qPCR analyses in our exposed mouse MEF cells. In our system, RNA-seq data did not show the altered expression of MYST1, different from the case in human UROtsa cells. To summarize, three independent investigations including ours showed the decreases of histone acetylations (H3K9ac, H3K27ac, or H4K16ac) (Jo *et al.*, 2009; Cronican *et al.*, 2013), whereas one report showed a contradictory increase of H3K9ac (Ramirez *et al.*, 2008).

Taken together, it remains to be finalized the changes of histone acetylations and affected HAT(s) responsible for the changed histone acetylation upon arsenic exposure. From our observation, all three doses showed similar inhibition of p300 (Figure 3A), maybe ruling out of different doses for different observations (Ramirez *et al.*, 2008). We also noticed the different cell types/lines used or differences in model system used (animal or cell line). Whether these differences explain the controversies of increased or decreased histone acetylation requests further investigation. Lastly, our convincing data suggest that the loss of p300 (and resulting H3K27ac) is related to inorganic arsenic related diseases. Therefore, it may open a new avenue for alleviating the consequence of H3K27ac, for example, via boosting CBP enzymatic activity for at least partially compensating p300 functions (Iyer *et al.*, 2004; Kasper *et al.*, 2006). Both p300 and its regulated gene *Klf4* (silenced after arsenic exposure; Figure 2C) are tumor suppressors(Iyer *et al.*, 2004). For example, *EP300* works as a tumor suppressor in hematological tumors (Oike *et al.*, 1999; Kung *et al.*, 2000). Down-expression of p53 and p300 were also found in azoxymethane treated mice, which is a colon-specific carcinogen(Aizu *et al.*, 2003). The role of *Klf4* as a tumor suppressor has been validated in colorectal cancer (Zhao *et al.*, 2004), non-Hodgkin lymphoma (Guan *et al.*, 2010), classic Hodgkin lymphoma (Guan *et al.*, 2010), and pancreatic ductal carcinoma (Zammarchi *et al.*, 2011). According to our results, it is reasonable to speculate that arsenite induced down-regulation of p300, and following decreased H3K27ac in *Klf4*. This change led to down expression of *Klf4*, which promoted the initiation of tumors. Other researchers also reported this result (Evans *et al.*, 2007).

## Figure Legends.

Fig.1 Identification of a low dose exposure condition. 1A and 1B showed results of Western blot of Nrf2 protein in MEF and Dot1L-/- cells after 6 h exposure of arsenite. 1C showed the quantitative results of western blot in 1A and 1B. 1D and 1E showed results of Western blot of Nrf2 protein in MEF and Dot1L-/- cells after 24 h exposure of arsenite. 1F showed the quantitative results of Western blot in 1D and 1E

Fig. 2 Results from mRNA-seq and qRT-PCR validation. 2A and 2B showed up-regulation and down-regulation genes after arsenite exposure in MEF and Dot1L-/- cells. 2C and 2D showed qRT-PCR results from two cells. 2E and 2F showed Pearson correlation of two cells.

Fig. 3 Protein expression and histone modification signals. 3A showed expression of p300, and relative signals of H3K27ac and H3K4me1 in MEF and Dot1L-/- cells in different concentrations of arsenite. 3B, 3C and 3D showed the quantitative results of Western blot.

Fig. 4 Results from ChIP-seq, which showed peaks in ChIP-seq results using H3K27ac antibody.

Fig. 5 Results from WashU EpiGenome Browser. 6A, 6B, 6C, 6D showed ChIP-seq and mRNA-seq results of *EP300*, *Klf4*, *Notch2*, and *Igf2r* gene.

Fig.6 Results from ChIP-qPCR. 7A showed ChIP-qPCR results using H3K27ac antibody and 7B showed ChIP-qPCR results using H3K4me1 antibody.

## Author Contribution:

Z.W conceived this project (with inputs from H.Z, Y.J, and Q.Z) and supervised the experiments. Y.Z, D.L performed the experiments; Y.L did the bioinformatic analyses; D.K, H.Z, Y.J. and Q.Z also conceptually contributed significantly; Y.Z and Z.W wrote the manuscript.

## Acknowledgment:

The laboratory of Z.W was supported by U.S. National Institutes of Health (R01ES25761, U01ES026721 opportunity fund, and R21ES028351) and Johns Hopkins Catalyst Award. Q.Z

thanks China Scholarship Council for support. This work was also made possible by the ChuTian Professorship from Hubei University, and support from GENEWIZ Suzhou. Lastly, extensive bioinformatic analyses were possible only with computational resources (and/or scientific computing services) at the Maryland Advanced Research Computing Center (MARCC). We are therefore greatly indebted to the support from MARCC.

### **Conflict of Interest:**

ZW serves as a consultant for CUSABIO (CusAb) Company. GENEWIZ Suzhou also contributes to the sequencing in this research. However, they had no influence on the experimental design and data presentation here.

### **References**

Aizu, W., Guda, K., Nambiar, P., Xin, T., Thibodeau, M., Rosenberg, D.W., Giardina, C., 2003. p53 and its co-activator p300 are inversely regulated in the mouse colon in response to carcinogen. *Toxicology letters* **144**, 213-224.

Alamdar, A., Xi, G., Huang, Q., Tian, M., Eqani, S., Shen, H., 2017. Arsenic activates the expression of 3beta-HSD in mouse Leydig cells through repression of histone H3K9 methylation. *Toxicology and applied pharmacology* **326**, 7-14.

Allfrey, V.G., Mirsky, A.E., 1964. Structural Modifications of Histones and their Possible Role in the Regulation of RNA Synthesis. *Science* **144**, 559.

Argos, M., Parvez, F., Rahman, M., Rakibuz-Zaman, M., Ahmed, A., Hore, S.K., Islam, T., Chen, Y., Pierce, B.L., Slavkovich, V., Olopade, C., Yunus, M., Baron, J.A., Graziano, J.H., Ahsan, H., 2014. Arsenic and lung disease mortality in Bangladeshi adults. *Epidemiology* **25**, 536-543.

Bhattacharjee, P., Banerjee, M., Giri, A.K., 2013. Role of genomic instability in arsenic-induced carcinogenicity. A review. *Environment international* **53**, 29-40.

Chervona, Y., Arita, A., Costa, M., 2012a. Carcinogenic metals and the epigenome: understanding the effect of nickel, arsenic, and chromium. *Metallomics : integrated biometal science* **4**, 619-627.

Chervona, Y., Hall, M.N., Arita, A., Wu, F., Sun, H., Tseng, H.C., Ali, E., Uddin, M.N., Liu, X., Zoroddu, M.A., Gamble, M.V., Costa, M., 2012b. Associations between arsenic exposure and global posttranslational histone

modifications among adults in Bangladesh. *Cancer epidemiology, biomarkers & prevention* : a publication of the American Association for Cancer Research, cosponsored by the American Society of Preventive Oncology **21**, 2252-2260.

Cronican, A.A., Fitz, N.F., Carter, A., Saleem, M., Shiva, S., Barchowsky, A., Koldamova, R., Schug, J., Lefterov, I., 2013. Genome-wide alteration of histone H3K9 acetylation pattern in mouse offspring prenatally exposed to arsenic. *PLoS one* **8**, e53478.

Dai H, W.Z., 2014. Histone Modification Patterns and Their Responses to Environment. *Current Environmental Health Reports* **1**, 11.

Eckstein, M., Eleazer, R., Rea, M., Fondufe-Mittendorf, Y., 2017. Epigenomic reprogramming in inorganic arsenic-mediated gene expression patterns during carcinogenesis. *Reviews on environmental health* **32**, 93-103.

Evans, P.M., Zhang, W., Chen, X., Yang, J., Bhakat, K.K., Liu, C., 2007. Kruppel-like factor 4 is acetylated by p300 and regulates gene transcription via modulation of histone acetylation. *The Journal of biological chemistry* **282**, 33994-34002.

Guan, H., Xie, L., Leithauser, F., Flossbach, L., Moller, P., Wirth, T., Ushmorov, A., 2010. KLF4 is a tumor suppressor in B-cell non-Hodgkin lymphoma and in classic Hodgkin lymphoma. *Blood* **116**, 1469-1478.

Guermah, M., Palhan, V.B., Tackett, A.J., Chait, B.T., Roeder, R.G., 2006. Synergistic functions of SII and p300 in productive activator-dependent transcription of chromatin templates. *Cell* **125**, 275-286.

Hamadeh, H.K., Vargas, M., Lee, E., Menzel, D.B., 1999. Arsenic disrupts cellular levels of p53 and mdm2: a potential mechanism of carcinogenesis. *Biochemical and biophysical research communications* **263**, 446-449.

Hane, B.G., Jager, K., Drexler, H.G., 1993. The Pearson product-moment correlation coefficient is better suited for identification of DNA fingerprint profiles than band matching algorithms. *Electrophoresis* **14**, 967-972.

Howe, C.G., Liu, X., Hall, M.N., Ilievski, V., Caudill, M.A., Malyshева, O., Lomax-Luu, A.M., Parvez, F., Siddique, A.B., Shahriar, H., Uddin, M.N., Islam, T., Graziano, J.H., Costa, M., Gamble, M.V., 2017. Sex-Specific Associations between One-Carbon Metabolism Indices and Posttranslational Histone Modifications in Arsenic-Exposed Bangladeshi Adults. *Cancer epidemiology, biomarkers & prevention* : a publication of the American Association for Cancer Research, cosponsored by the American Society of Preventive Oncology **26**, 261-269.

Iyer, N.G., Ozdag, H., Caldas, C., 2004. p300/CBP and cancer. *Oncogene* **23**, 4225-4231.

Jin, Q., Yu, L.R., Wang, L., Zhang, Z., Kasper, L.H., Lee, J.E., Wang, C., Brindle, P.K., Dent, S.Y., Ge, K., 2011. Distinct roles of GCN5/PCAF-mediated H3K9ac and CBP/p300-mediated H3K18/27ac in nuclear receptor transactivation. *The EMBO journal* **30**, 249-262.

Jo, W.J., Ren, X., Chu, F., Aleshin, M., Wintz, H., Burlingame, A., Smith, M.T., Vulpe, C.D., Zhang, L., 2009. Acetylated H4K16 by MYST1 protects UROtsa cells from arsenic toxicity and is decreased following chronic arsenic exposure. *Toxicology and applied pharmacology* **241**, 294-302.

Kasper, L.H., Fukuyama, T., Biesen, M.A., Boussouar, F., Tong, C., de Pauw, A., Murray, P.J., van Deursen, J.M., Brindle, P.K., 2006. Conditional knockout mice reveal distinct functions for the global transcriptional coactivators CBP and p300 in T-cell development. *Molecular and cellular biology* **26**, 789-809.

Knobeloch, L.M., Zierold, K.M., Anderson, H.A., 2006. Association of arsenic-contaminated drinking-water with prevalence of skin cancer in Wisconsin's Fox River Valley. *Journal of health, population, and nutrition* **24**, 206-213.

Kung, A.L., Rebel, V.I., Bronson, R.T., Ch'ng, L.E., Sieff, C.A., Livingston, D.M., Yao, T.P., 2000. Gene dose-dependent control of hematopoiesis and hematologic tumor suppression by CBP. *Genes & development* **14**, 272-277.

Langmead, B., 2010. Aligning short sequencing reads with Bowtie. *Current protocols in bioinformatics Chapter 11*, Unit 11 17.

Li, Y., Yang, H.X., Luo, R.Z., Zhang, Y., Li, M., Wang, X., Jia, W.H., 2011. High expression of p300 has an unfavorable impact on survival in resectable esophageal squamous cell carcinoma. *The Annals of thoracic surgery* **91**, 1531-1538.

Li, Z., Dai, H., Martos, S.N., Xu, B., Gao, Y., Li, T., Zhu, G., Schones, D.E., Wang, Z., 2015. Distinct roles of DNMT1-dependent and DNMT1-independent methylation patterns in the genome of mouse embryonic stem cells. *Genome biology* **16**, 115.

Livak, K.J., Schmittgen, T.D., 2001. Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. *Methods* **25**, 402-408.

Lunnon, K., Mill, J., 2013. Epigenetic studies in Alzheimer's disease: current findings, caveats, and considerations for future studies. *American journal of medical genetics. Part B, Neuropsychiatric genetics : the official publication of the International Society of Psychiatric Genetics* **162B**, 789-799.

Mandal, B.K., Suzuki, K.T., 2002. Arsenic round the world: a review. *Talanta* **58**, 201-235.

Ogryzko, V.V., Schiltz, R.L., Russanova, V., Howard, B.H., Nakatani, Y., 1996. The transcriptional coactivators p300 and CBP are histone acetyltransferases. *Cell* **87**, 953-959.

Oike, Y., Takakura, N., Hata, A., Kaname, T., Akizuki, M., Yamaguchi, Y., Yasue, H., Araki, K., Yamamura, K., Suda, T., 1999. Mice homozygous for a truncated form of CREB-binding protein exhibit defects in hematopoiesis and vasculo-angiogenesis. *Blood* **93**, 2771-2779.

Oya-Ohta, Y., Kaise, T., Ochi, T., 1996. Induction of chromosomal aberrations in cultured human fibroblasts by inorganic and organic arsenic compounds and the different roles of glutathione in such induction. *Mutation research* **357**, 123-129.

Pournara, A., Kippler, M., Holmlund, T., Ceder, R., Grafstrom, R., Vahter, M., Broberg, K., Wallberg, A.E., 2016. Arsenic alters global histone modifications in lymphocytes in vitro and in vivo. *Cell biology and toxicology* **32**, 275-284.

Ramirez, T., Brocher, J., Stopper, H., Hock, R., 2008. Sodium arsenite modulates histone acetylation, histone deacetylase activity and HMGN protein dynamics in human cells. *Chromosoma* **117**, 147-157.

Roelfsema, J.H., White, S.J., Ariyurek, Y., Bartholdi, D., Niedrist, D., Papadia, F., Bacino, C.A., den Dunnen, J.T., van Ommen, G.J., Breuning, M.H., Hennekam, R.C., Peters, D.J., 2005. Genetic heterogeneity in Rubinstein-Taybi syndrome: mutations in both the CBP and EP300 genes cause disease. *American journal of human genetics* **76**, 572-580.

Shi, H., Shi, X., Liu, K.J., 2004. Oxidative mechanism of arsenic toxicity and carcinogenesis. *Molecular and cellular biochemistry* **255**, 67-78.

Sun, Z., Chin, Y.E., Zhang, D.D., 2009. Acetylation of Nrf2 by p300/CBP augments promoter-specific DNA binding of Nrf2 during the antioxidant response. *Molecular and cellular biology* **29**, 2658-2672.

Tang, Z., Chen, W.Y., Shimada, M., Nguyen, U.T., Kim, J., Sun, X.J., Sengoku, T., McGinty, R.K., Fernandez, J.P., Muir, T.W., Roeder, R.G., 2013. SET1 and p300 act synergistically, through coupled histone modifications, in transcriptional activation by p53. *Cell* **154**, 297-310.

Tellez-Plaza, M., Tang, W.Y., Shang, Y., Umans, J.G., Francesconi, K.A., Goessler, W., Ledesma, M., Leon, M., Laclaustra, M., Pollak, J., Guallar, E., Cole, S.A., Fallin, M.D., Navas-Acien, A., 2014. Association of global DNA methylation and global DNA hydroxymethylation with metals and other exposures in human blood DNA samples. *Environmental health perspectives* **122**, 946-954.

Wang, Z., Humphries, B., Xiao, H., Jiang, Y., Yang, C., 2013. Epithelial to mesenchymal transition in arsenic-transformed cells promotes angiogenesis through activating beta-catenin-vascular endothelial growth factor pathway. *Toxicology and applied pharmacology* **271**, 20-29.

Wang, Z., Zang, C., Cui, K., Schones, D.E., Barski, A., Peng, W., Zhao, K., 2009. Genome-wide mapping of HATs and HDACs reveals distinct functions in active and inactive genes. *Cell* **138**, 1019-1031.

Wang, Z., Zang, C., Rosenfeld, J.A., Schones, D.E., Barski, A., Cuddapah, S., Cui, K., Roh, T.Y., Peng, W., Zhang, M.Q., Zhao, K., 2008. Combinatorial patterns of histone acetylations and methylations in the human genome. *Nature genetics* **40**, 897-903.

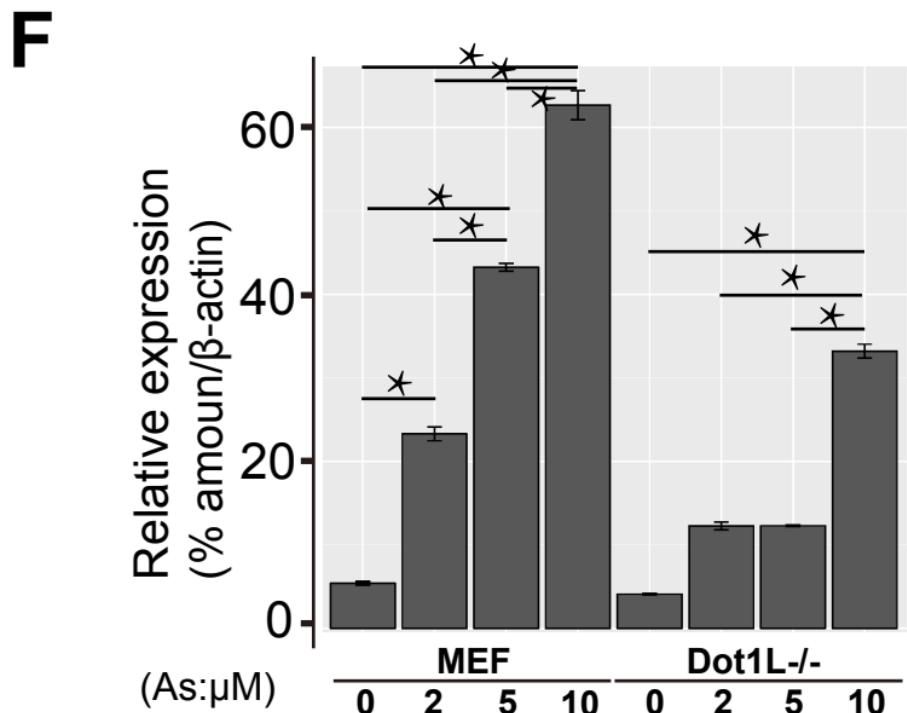
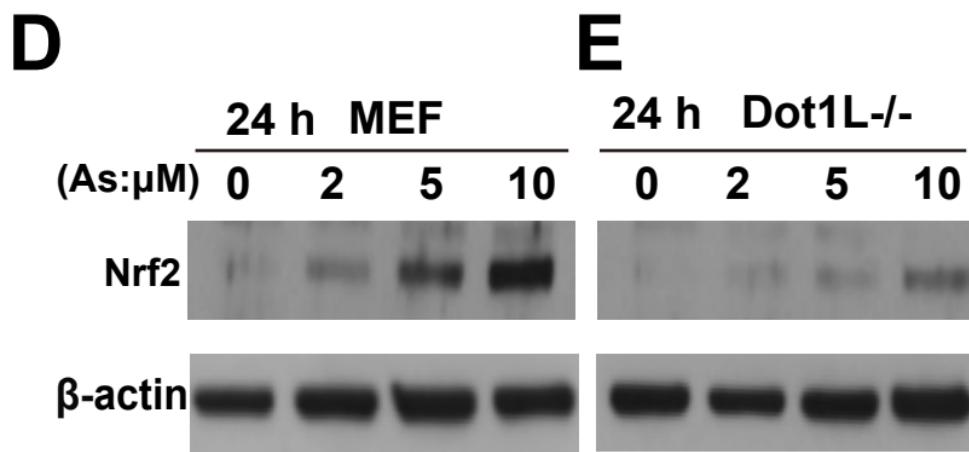
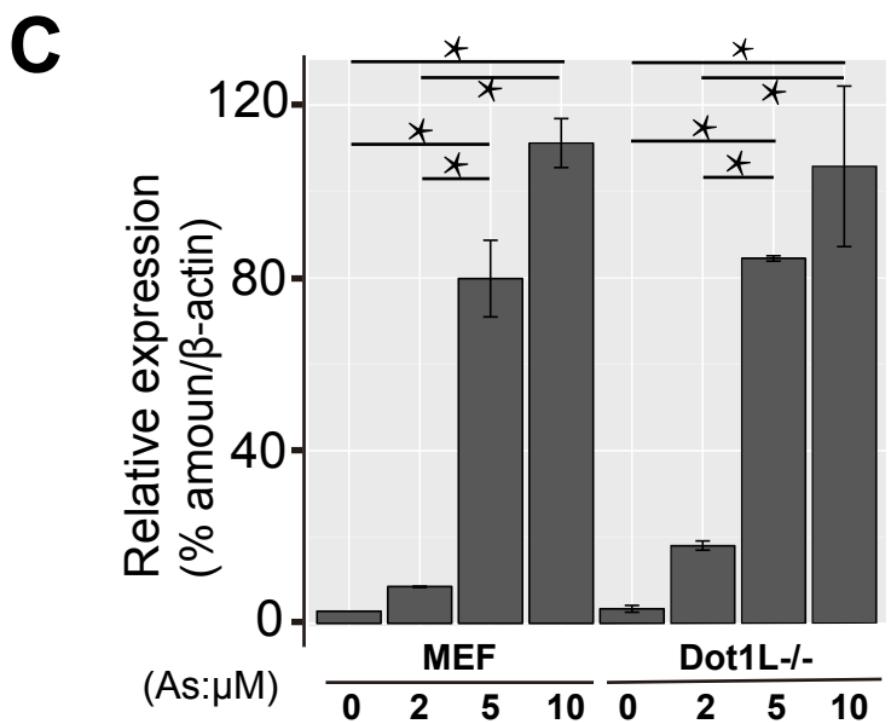
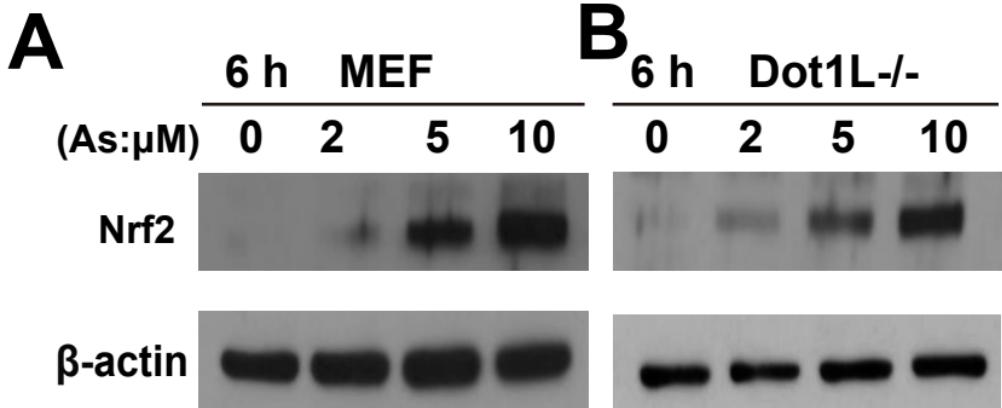
Wu, M.M., Kuo, T.L., Hwang, Y.H., Chen, C.J., 1989. Dose-response relation between arsenic concentration in well water and mortality from cancers and vascular diseases. *American journal of epidemiology* **130**, 1123-1132.

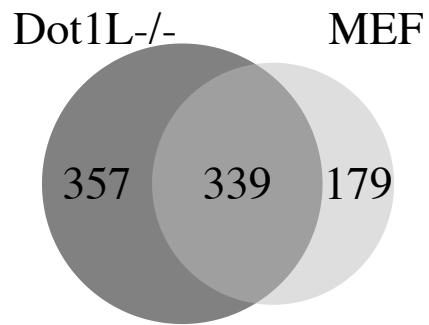
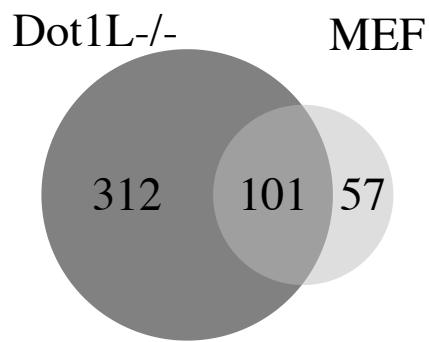
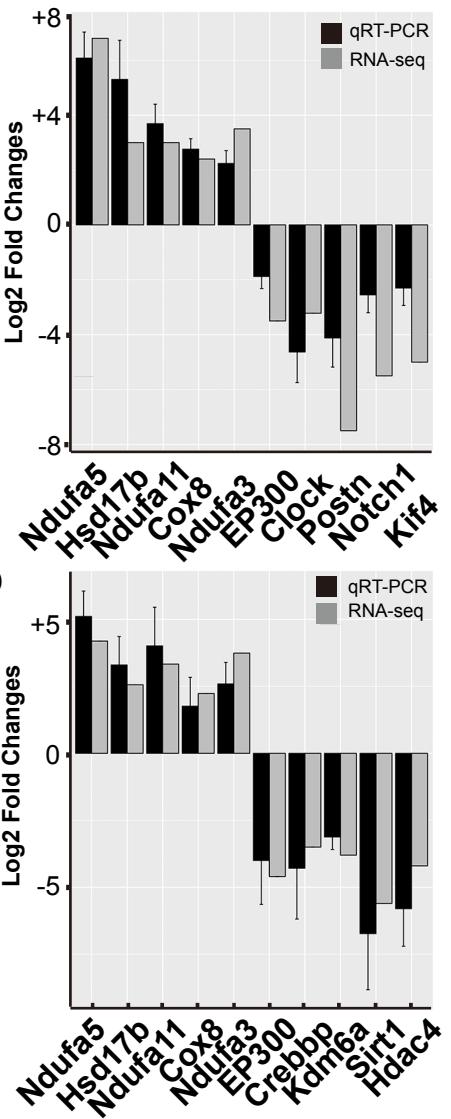
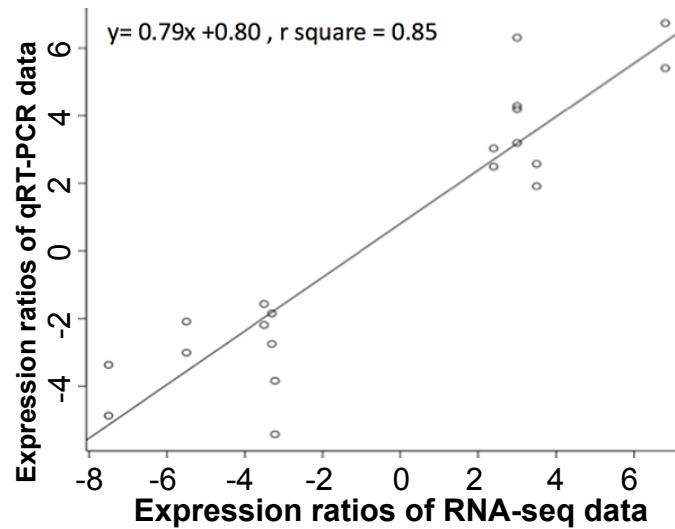
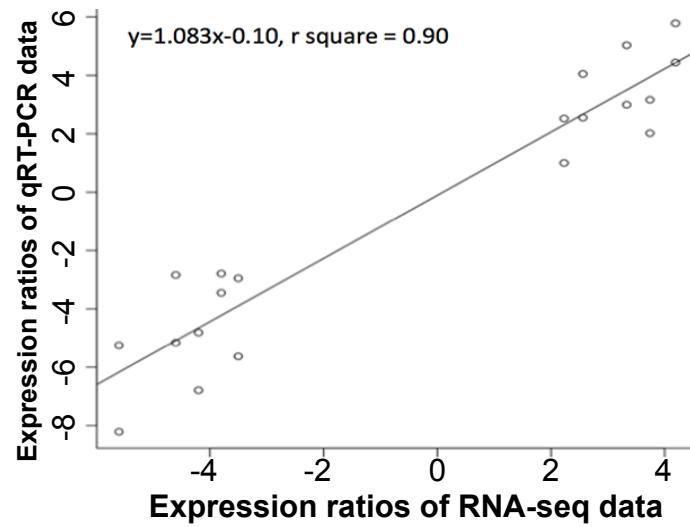
Zammarchi, F., Morelli, M., Menicagli, M., Di Cristofano, C., Zavaglia, K., Paolucci, A., Campani, D., Aretini, P., Boggio, U., Mosca, F., Cavazzana, A., Cartegni, L., Bevilacqua, G., Mazzanti, C.M., 2011. KLF4 is a novel candidate tumor suppressor gene in pancreatic ductal carcinoma. *The American journal of pathology* **178**, 361-372.

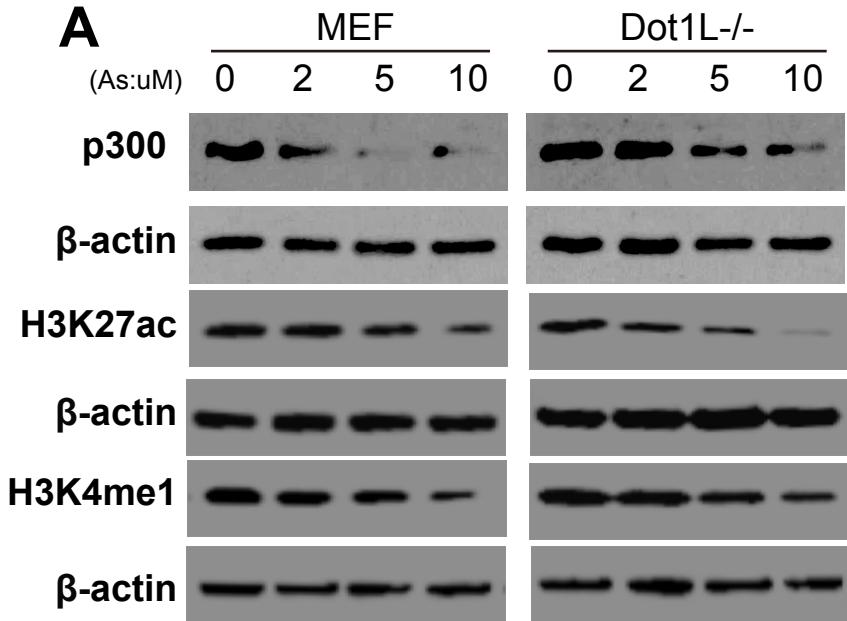
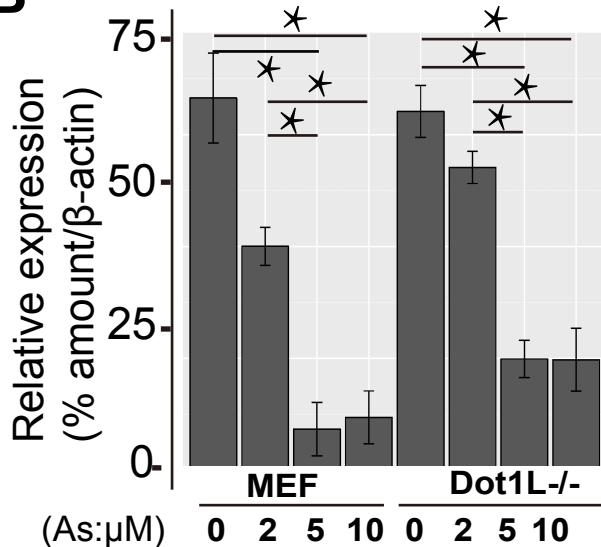
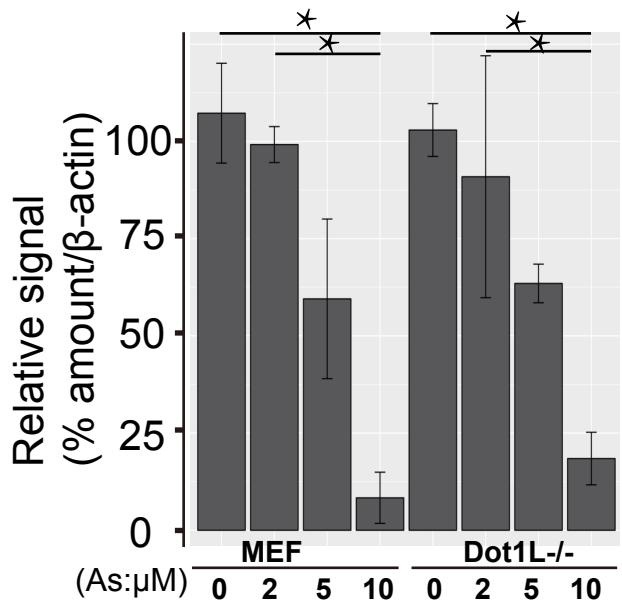
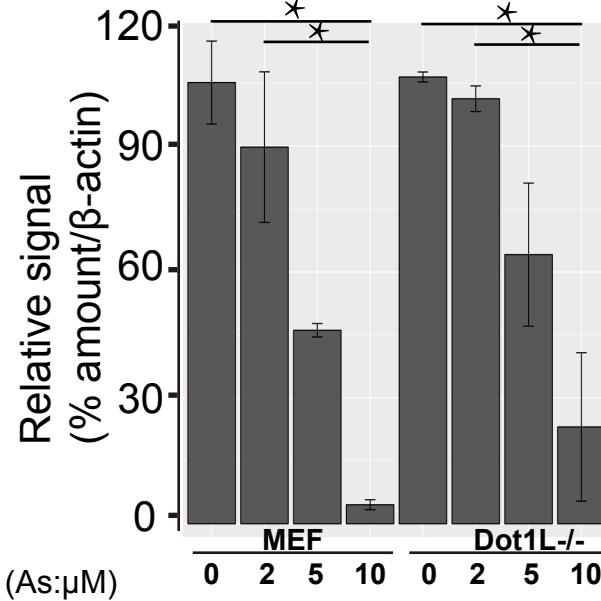
Zhang, Y., Liu, T., Meyer, C.A., Eeckhoute, J., Johnson, D.S., Bernstein, B.E., Nusbaum, C., Myers, R.M., Brown, M., Li, W., Liu, X.S., 2008. Model-based analysis of ChIP-Seq (MACS). *Genome biology* **9**, R137.

Zhao, C.Q., Young, M.R., Diwan, B.A., Coogan, T.P., Waalkes, M.P., 1997. Association of arsenic-induced malignant transformation with DNA hypomethylation and aberrant gene expression. *Proceedings of the National Academy of Sciences of the United States of America* **94**, 10907-10912.

Zhao, W., Hisamuddin, I.M., Nandan, M.O., Babbin, B.A., Lamb, N.E., Yang, V.W., 2004. Identification of Kruppel-like factor 4 as a potential tumor suppressor gene in colorectal cancer. *Oncogene* **23**, 395-402.



**A****B****C****E****F**

**A****B****C****D**

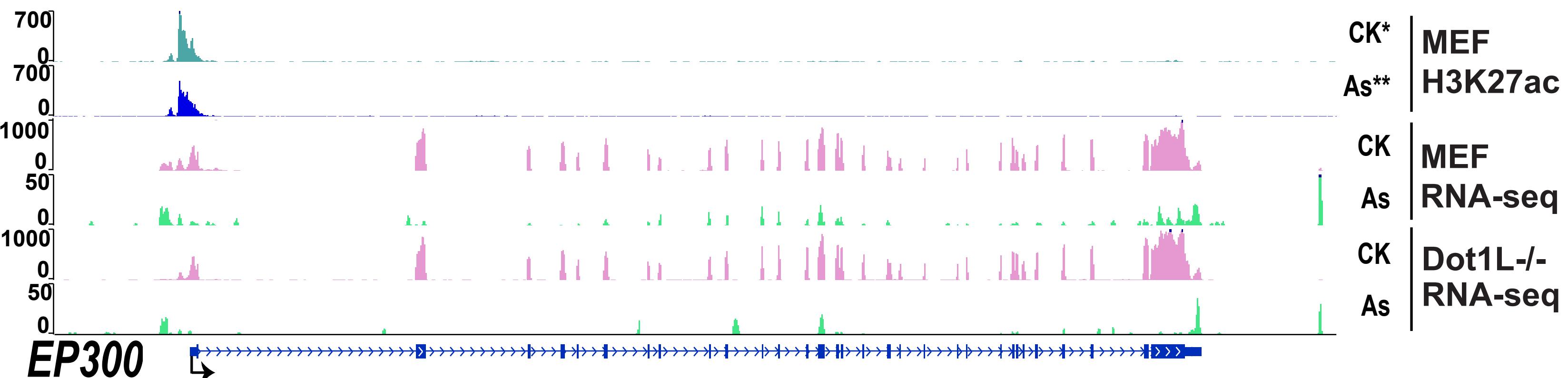
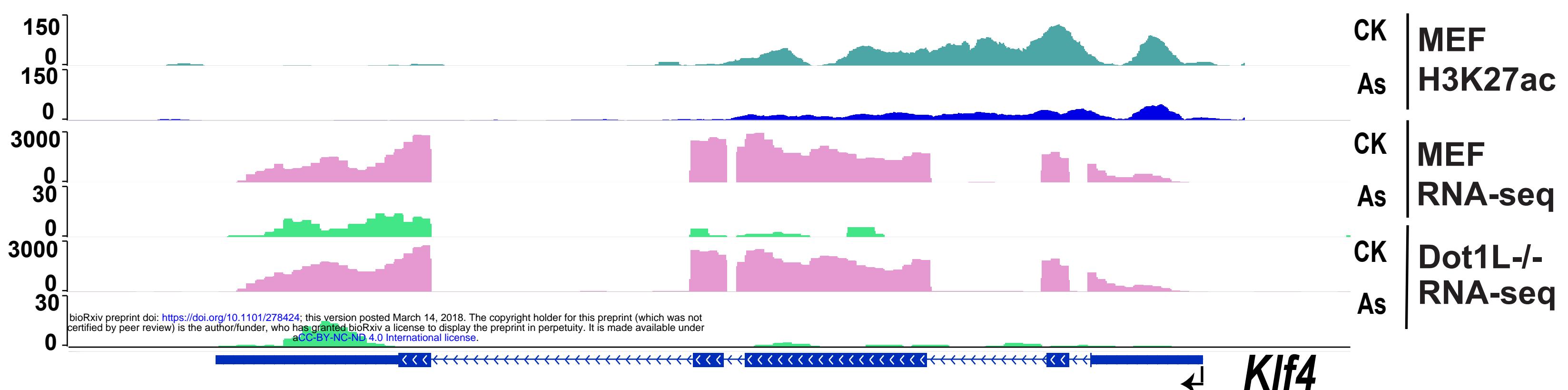
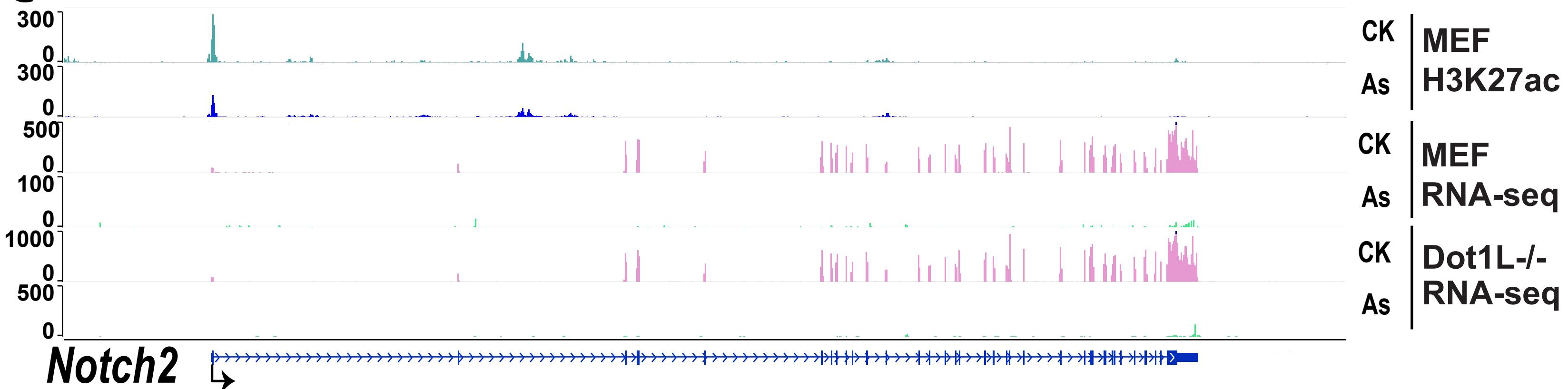
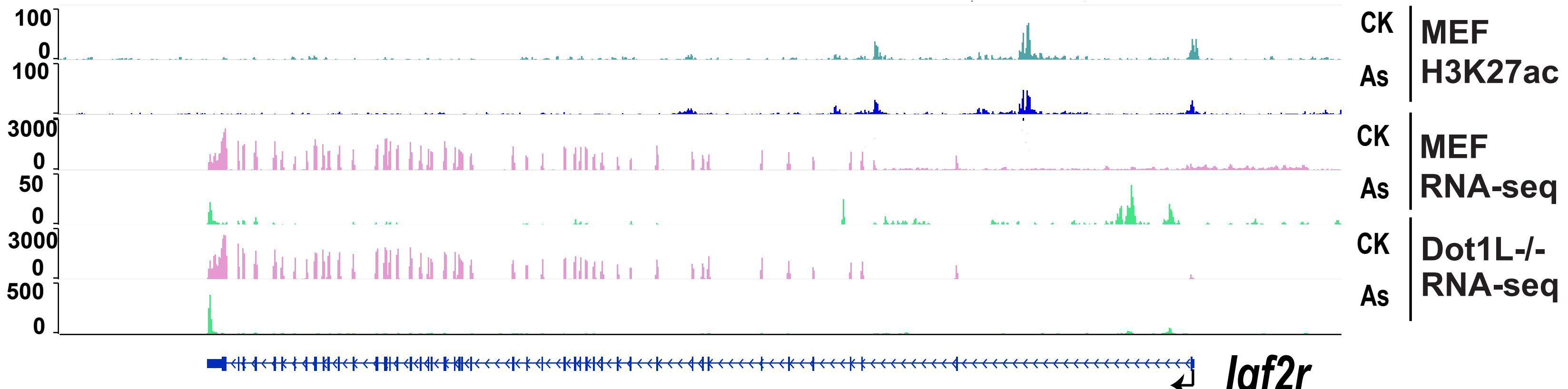
Control

As

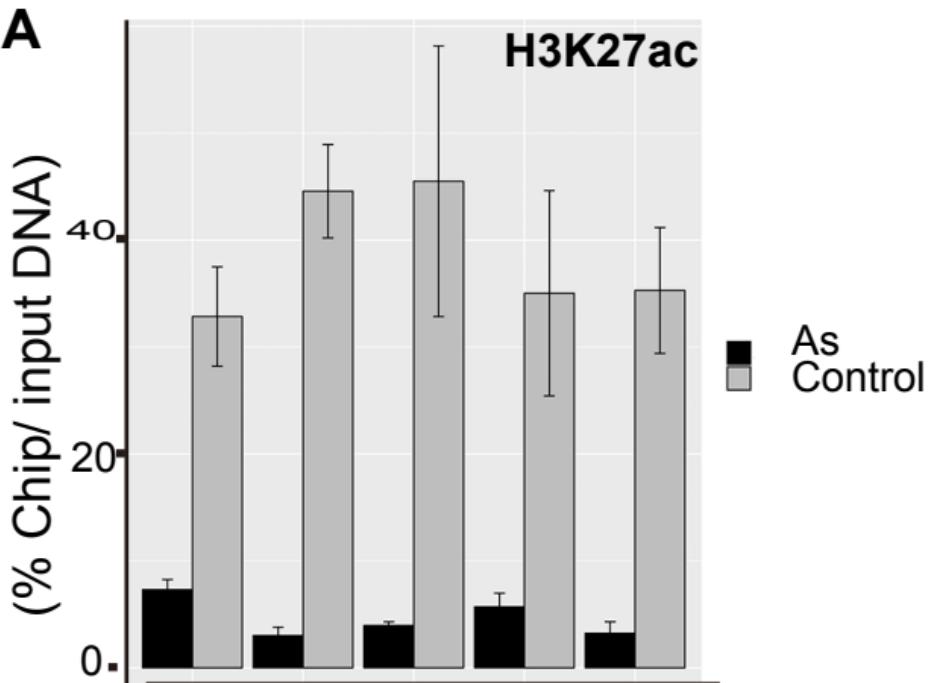
18302

21548

4772

**A****B****C****D**

\*CK, Control; \*\*As, Arsenite

**A****B**