

## RESEARCH ARTICLE

# Arsenite malignantly transforms human prostate epithelial cells *in vitro* by gene amplification of mutated KRAS

B. Alex Merrick<sup>1</sup>\*, Dhiral P. Phadke<sup>2</sup>, Meredith A. Bostrom<sup>3</sup>, Ruchir R. Shah<sup>2</sup>, Garron M. Wright<sup>3</sup>, Xinguo Wang<sup>3</sup>, Oksana Gordon<sup>3</sup>, Katherine E. Pelch<sup>1</sup>\*, Scott S. Auerbach<sup>1</sup>, Richard S. Paules<sup>1</sup>, Michael J. DeVito<sup>1</sup>, Michael P. Waalkes<sup>1</sup>, Erik J. Tokar<sup>1</sup>

**1** Division of the National Toxicology Program, National Institute of Environmental Health Sciences, Research Triangle Park, North Carolina, United States of America, **2** Sciome, LLC, Research Triangle Park, North Carolina, United States of America, **3** David H. Murdock Research Institute, Kannapolis, North Carolina, United States of America

\* These authors contributed equally to this work.

✉ Current address: The Endocrine Disruptor Exchange, Eckert, Colorado, United States of America

\* [merrick@niehs.nih.gov](mailto:merrick@niehs.nih.gov)



## OPEN ACCESS

**Citation:** Merrick BA, Phadke DP, Bostrom MA, Shah RR, Wright GM, Wang X, et al. (2019) Arsenite malignantly transforms human prostate epithelial cells *in vitro* by gene amplification of mutated KRAS. PLoS ONE 14(4): e0215504. <https://doi.org/10.1371/journal.pone.0215504>

**Editor:** Kamaleshwar P Singh, Texas Tech University, UNITED STATES

**Received:** February 8, 2019

**Accepted:** April 4, 2019

**Published:** April 22, 2019

**Copyright:** This is an open access article, free of all copyright, and may be freely reproduced, distributed, transmitted, modified, built upon, or otherwise used by anyone for any lawful purpose. The work is made available under the [Creative Commons CC0](https://creativecommons.org/licenses/by/4.0/) public domain dedication.

**Data Availability Statement:** The data underlying the results presented in the study are available from the Sequence Read Archive (SRA) at <https://www.ncbi.nlm.nih.gov/sra> under Study Accession Nos. SUB4913016 and PRJNA514436.

**Funding:** This work was supported by funds from the Division of the National Toxicology Program (DNTP). The funder is the federal government at the National Institute of Environmental Health Sciences (NIEHS) <https://www.niehs.nih.gov/index.cfm>] and has provided support in the form of

## Abstract

Inorganic arsenic is an environmental human carcinogen of several organs including the urinary tract. RWPE-1 cells are immortalized, non-tumorigenic, human prostate epithelia that become malignantly transformed into the CAse-PE line after continuous *in vitro* exposure to 5μM arsenite over a period of months. For insight into *in vitro* arsenite transformation, we performed RNA-seq for differential gene expression and targeted sequencing of KRAS. We report >7,000 differentially expressed transcripts in CAse-PE cells compared to RWPE-1 cells at >2-fold change,  $q < 0.05$  by RNA-seq. Notably, KRAS expression was highly elevated in CAse-PE cells, with pathway analysis supporting increased cell proliferation, cell motility, survival and cancer pathways. Targeted DNA sequencing of KRAS revealed a mutant specific allelic imbalance, 'MASI', frequently found in primary clinical tumors. We found high expression of a mutated KRAS transcript carrying oncogenic mutations at codons 12 and 59 and many silent mutations, accompanied by lower expression of a wild-type allele. Parallel cultures of RWPE-1 cells retained a wild-type KRAS genotype. Copy number analysis and sequencing showed amplification of the mutant KRAS allele. KRAS is expressed as two splice variants, KRAS4a and KRAS4b, where variant 4b is more prevalent in normal cells compared to greater levels of variant 4a seen in tumor cells. 454 Roche sequencing measured KRAS variants in each cell type. We found KRAS4a as the predominant transcript variant in CAse-PE cells compared to KRAS4b, the variant expressed primarily in RWPE-1 cells and in normal prostate, early passage, primary epithelial cells. Overall, gene expression data were consistent with KRAS-driven proliferation pathways found in spontaneous tumors and malignantly transformed cell lines. Arsenite is recognized as an important environmental carcinogen, but it is not a direct mutagen. Further investigations into this *in vitro* transformation model will focus on genomic events that cause arsenite-mediated mutation and overexpression of KRAS in CAse-PE cells.

salaries for DHMRI authors, MAB, GMW, XW, OG, under NIEHS contract HHSN273201100016C; and salaries for Sciome, LLC authors DPP and RRS under NIEHS contract HHSN273201700001C. The website for Sciome, LLC is <https://www.sciome.com> and for DHMRI are <https://transforming-science.com/dhmri/>. The Sciome, LLC and DHMRI (David H. Murdoch Research Institute) authors did not have any additional role in the study design, decision to publish, or preparation of the manuscript. Sciome, LLC provided support in the form of salary for authors DPP, RRS, but did not have any additional role in the study design, data collection and analysis, decision to publish, or preparation of the manuscript. The specific role of these authors is articulated in the 'author contributions' section.

**Competing interests:** We have the following interests: Dhiral P. Phadke and Ruchir R. Shah are employed by Sciome, LLC. The US government at NIEHS is the sole funder of this research through contracts with Sciome, LLC and DHMRI. Neither Sciome, LLC nor DHMRI derive benefits from the publication of this work and had no role in funding of the work. There are no patents, products in development or marketed products to declare. This does not alter our adherence to all the PLOS ONE policies on sharing data and materials.

## Introduction

Environmental exposure to arsenic increases the risks of skin, lung, kidney, liver and urinary-bladder cancers [1, 2]. Although the mode of action for arsenic-induced tumors is unclear, many animal and human studies suggest arsenic can act as a carcinogen [3, 4], co-carcinogen [5, 6], or transplacental carcinogen [7]. Arsenite and arsenate, the inorganic tri- and pentavalent forms of arsenic, are considered non-mutagenic in bacterial and human cells [8, 9]. However, arsenic may indirectly cause DNA damage, chromosomal abnormalities, and generation of reactive oxygen species (ROS) like superoxide or hydrogen peroxide [10, 11]. Other transformational effects of arsenic may involve disruption of signaling pathways, miRNA dysregulation, inhibition of DNA repair, or formation of cancer stem cells or polycomb proteins [12–19]. Arsenite and other trivalent species can be acutely cytotoxic by readily binding to intracellular thiols (e.g. GSH) and sulfhydryl sites on macromolecules to inhibit critical biochemical processes [17]. Persistent cytotoxicity from prolonged arsenic exposure and subsequent regenerative proliferation may contribute to carcinogenesis as well [3]. Biotransformation of arsenic involves S-adenosylmethione (SAM), methyltransferases and sulfur redox metabolism so that arsenic-induced interference of methyl-donor pathways could lead to abnormal DNA methylation and histone modification patterns and epigenetic transformation [14, 15, 17, 20–24].

The prostate gland, as part of the urogenital system, is among the many target organs in arsenic carcinogenesis [25–27]. Epidemiologic studies have shown an association of environmental inorganic arsenic exposure with prostate cancer incidence and mortality in the U.S. and abroad [28–30]. Development of immortalized human prostate epithelial cells have greatly advanced prostate cancer research [31, 32]. *In vitro* transformation assays induced by various metals, including arsenic, have provided an invaluable model for examining the multistep events underlying tumor formation (see reviews [19, 33]). RWPE-1 cells [34] were developed as non-tumorigenic, human prostate epithelia for research. Subsequently, our group was able to demonstrate that RWPE-1 cells can be malignantly transformed into the CAse-PE cell line by continuous exposure after 30 weeks of non-cytotoxic levels of sodium arsenite at 5 $\mu$ M in culture [35]. CAse-PE cells create tumors when injected into nude mice and demonstrate many *in vitro* characteristics of malignant transformation, including loss of contact inhibition, anchorage-independent growth, resistance to apoptosis, and increased activity of secreted MMP-9 [35, 36].

Early work with CAse-PE cells reported an increased KRAS expression by northern and western blot with absence of mutations by PCR-RFLP, along with MAP kinase and ErbB-2 activation and androgen independent growth [25, 37]. KRAS mRNA and protein were highly overexpressed in CAse-PE cells beginning at 12 weeks and increasing over a 30 week period [25]. Initial evaluation showed KRAS overexpression in absence of mutations or proximal DNA methylation changes to the gene body [25]. KRAS knockdown in shRNA experiments can reverse many CAse-PE malignant properties that include a reduced metalloprotease activity, limited anchorage independent growth, reduced cell proliferation, less invasiveness, and decreased MAPK signaling activities [38]. Other studies using these two cell lines reported epigenetic changes such as transcriptional silencing of ZNF (zinc finger) transcriptional repressors, developmental HOX and protocadherin gene families [39, 40] by focal hypermethylation. Carcinogenesis-related genes such as S100P, HYAL1, NES, and NTM have been similarly shown to have an inverse relationship of DNA methylation to gene expression [41]. Further, KRAS activation in arsenic transformed cells has been associated with dysregulated miRNA expression linked to transcriptional regulation of many downstream processes supporting malignancy [42]. Thus, accumulated evidence suggests increased KRAS expression plays a key role in arsenic-induced malignancy in the CAse-PE cell model.

KRAS is well-recognized as a transformative factor in prostate cancer [43–47] and gene amplification at KRAS loci (e.g. increased copy number) has also been detected in many tumors including prostate [47–49]. In this study, we hypothesized genomic and transcriptomic analyses and targeted KRAS sequencing could reveal insights into arsenite transformation of RWPE-1 into CAsE-PE cells by looking for sequence variations in KRAS, alterations in gene and pathway expression, and genomic alterations contributing to development of malignancy.

## Results

### Alignment of reads, transcript assembly and differential expression

RNA-seq analysis was performed after polyA selection from three independently grown RWPE-1 and CAsE-PE cultures. Low quality bases at the ends of reads were removed and trimmed according to quality metrics (Phred Scores) and equivalent nucleotide distributions (S1 Fig) as previously described [50]. The total number of sequenced reads after quality-based trimming (S1 Table) ranged from 176–198 million paired end (PE) reads. Reads were aligned to hg19 using TopHat for a >200X total coverage of the transcriptome. The initial CAsE-PE alignments to hg19 were found to average 56.8% and were less than the average 77% of RWPE hg19 alignments. DNA-seq at low level coverage was performed to provide additional data to improve genomic alignment. Nearly 500 million DNA-seq reads were obtained for each cell line that were combined and ultimately assembled into 226,987 contigs, mapped to hg19 and RefSeq, excluding non-repetitive regions. Comparison with hg19 and contigs improved alignments to a mean of 71.6% for CAsE-PE and 82.0% for RWPE cells as shown in S1 Table.

CuffDiff was used to determine differential expression using RefSeq annotation from 46,090 transcripts and isoforms of which there were 7,265 total differences (q-value < 0.05; >2-fold). From this total number of differentially expressed genes (DEGs) there were 3,261 up-regulated transcripts and 4,004 transcripts which were downregulated compared to RWPE-1 expression (S2 Table).

### Differential expression and pathway activation by arsenic transformation

The top thirty increased and decreased DEGs are shown in Table 1. Notably, the proliferation gene transcript KRAS at 432-fold increase and the maternally imprinted gene transcript, H19, at 154-fold increase were highly upregulated. PSCA (prostate stem cell antigen), a recognized surface marker of prostate cancer [51], was increased 52-fold (Table 1). Altered molecular functions and disease pathways are displayed in Fig 1. Developmental, morphological and cell growth pathways were altered, as well as changes in reproductive and neoplastic disease pathways that are consistent with prostate malignancy. Fig 2 shows 34 downstream gene-pathway interactions annotated for KRAS (see S3 Table for citations) that involved substantially upregulated transcripts at  $\geq 5$ -fold. High levels of KRAS were associated with increased expression of several growth factors such as CSF3, IGF2, HBEGF, VEGFA; sodium channel isoforms like SCNN1A, B, and G; transcriptional regulators including FOS, NFKB1B, PLAGL1; and with upregulation of enzymes such as HMOX1, PTGS1 and DNMT3a.

Ingenuity (IPA) connectivity analysis was performed for the top 1,000 upregulated DEGs (algorithm limit) to determine the complex regulatory interrelationships of KRAS as displayed in a circular plot (Fig 3). In this analysis, gene-to-gene connections encompass changes in expression, gene and pathway activation, protein-protein interactions, phosphorylation and other interactions (see S4 Table). A greater number of connections for specific genes infers a higher level of biochemical and regulatory importance (darkened convergence of lines on edge of circle). The number of up- or downstream relationships among DEGs were tabulated in S4 Table. Fig 3 shows the top 25 upstream and downstream DEGs in large font at the edges of the

**Table 1. Top 30 upregulated and downregulated gene fold changes in CA5E-PE versus RWPE-1 cells.**

Fold Δ	Transcript ID	Gene Symbol	Entrez Gene Name	Function
959.26	NM_003490	SYN3	synapsin III	synaptic vesicles
432.37	NM_033360	KRAS	KRAS proto-oncogene, GTPase	enzyme
320.77	NM_015568	PPP1R16B	protein phosphatase 1 regulatory subunit 16B	phosphatase
255.16	NM_004121	GGT5_Variant2	gamma-glutamyltransferase 5	enzyme
188.98	NM_022369	STRA6	stimulated by retinoic acid 6	transporter
159.35	NM_033159	HYAL1	hyaluronoglucosaminidase 1	enzyme
154.58	NR_002196	H19	H19, imprinted maternally expressed transcript	imprinted gene, lncRNA
151.47	NM_000507	FBP1	fructose-bisphosphatase 1	phosphatase
119.89	NM_019601	SUSD2	sushi domain containing 2	cytokine receptor
98.69	NM_001190202	CES4A	carboxylesterase 4A	enzyme
86.06	NM_020142	NDUFA4L2	NDUFA4, mitochondrial complex associated like 2	enzyme
85.29	NM_002281	KRT81	keratin 81	structural protein
83.44	NR_052017	MED24	mediator complex subunit 24	transcription regulator
83.01	NM_138567	SYT8	synaptotagmin 8	transporter
78.25	NM_001099781	GGT5_Variant1	gamma-glutamyltransferase 5	enzyme
72.57	NM_001270991	EPGN	epithelial mitogen	growth factor
72.39	NM_130842	PTPRN2	protein tyrosine phosphatase, receptor type N2	phosphatase
67.91	NM_001242767	MTHFD1L	methylenetetrahydrofolate dehydrogenase 1 like	enzyme
67.89	NM_007129	ZIC2	Zic family member 2	transcription regulator
65.52	NM_172084	CAMK2B	calcium/calmodulin dependent protein kinase II beta	kinase
61.26	NM_153480	IL17RE	interleukin 17 receptor E	cytokine receptor
57.39	NM_001958	EEF1A2	eukaryotic translation elongation factor 1 alpha 2	translation regulator
56.66	NM_138768	MYEOV	myeloma overexpressed	tissue invasion
55.13	NR_047690	HYAL1_lncRNA	hyaluronoglucosaminidase 1	enzyme
52.87	NM_005672	PSCA	prostate stem cell antigen	cell growth regulator
52.30	NM_000336	SCNN1B	sodium channel epithelial 1 beta subunit	ion channel
48.60	NR_033861	LINC00514	long intergenic non-protein coding RNA 514	lncRNA
47.85	NR_046224	LINC00659	long intergenic non-protein coding RNA 659	lncRNA
47.52	NM_001171946	SUN1	Sad1 and UNC84 domain containing 1	nuclear envelope
45.58	NR_105044	LOC102546229	uncharacterized LOC102546229	unknown function
-162.16	NM_001001668	ZNF470	zinc finger protein 470	transcription regulator
-163.01	NM_018650	MARK1	microtubule affinity regulating kinase 1	kinase
-167.92	NM_024501	HOXD1	homeobox D1	transcription regulator
-177.01	NM_005602	CLDN11	claudin 11	tight junction protein
-178.56	NM_033518	SLC38A5	solute carrier family 38 member 5	transporter
-179.77	NM_001129891	FAM196B	family with sequence similarity 196 member B	cell growth regulator
-181.54	NM_032682	FOXP1	forkhead box P1	transcription regulator
-185.37	NM_007361	NID2	nidogen 2	ECM protein
-188.33	NM_001010924	FAM171A1	family with sequence similarity 171 member A1	unknown function
-188.57	NM_005103	FEZ1	fasciculation and elongation protein zeta 1	cell growth regulator
-195.20	NM_199320	JADE1	jade family PHD finger 1	transcription regulator
-196.28	NM_000922	PDE3B	phosphodiesterase 3B	enzyme
-197.65	NM_005708	GPC6	glypican 6	transmembrane receptor
-227.31	NM_018043	ANO1	anoctamin 1	ion channel
-237.88	NM_001289861	PER3	period circadian regulator 3	circadian regulator
-246.68	NM_020814	MARCH4	membrane associated ring-CH-type finger 4	enzyme
-249.17	NM_001197294	DPYSL3	dihydropyrimidinase like 3	enzyme
-252.33	NM_001190972	C8orf88	chromosome 8 open reading frame 88	negative translation

(Continued)

Table 1. (Continued)

Fold Δ	Transcript ID	Gene Symbol	Entrez Gene Name	Function
-267.80	NM_002403	MFAP2	microfibril associated protein 2	microfibril
-290.47	NM_002523	NPTX2	neuronal pentraxin 2	synaptic vesicles
-305.87	NM_020828	ZFP28	ZFP28 zinc finger protein	transcription regulator
-316.95	NM_152476	ZNF560	zinc finger protein 560	transcription regulator
-337.73	NM_000885	ITGA4	integrin subunit alpha 4	transmembrane receptor
-357.26	NM_170697	ALDH1A2	aldehyde dehydrogenase 1 family member A2	enzyme
-375.46	NR_030299	mir-573	microRNA 573	microRNA
-391.38	NR_036521	ZNF667-AS1	ZNF667 antisense RNA 1 (head to head)	transcription regulator
-513.45	NM_001007026	ATN1	atrophin 1	transcription regulator
-630.89	NM_002535	OAS2	2'-5'-oligoadenylate synthetase 2	enzyme
-683.93	NM_003070	SMARCA2	SWI/SNF related, a2	transcription regulator
-750.77	NM_016608	ARMCX1	armadillo repeat containing, X-linked 1	protein binding

Differentially expressed genes (DEGs) were determined by RNA-seq analysis of arsenic transformed cells (CAsE-PE) compared to the control non-transformed human prostate epithelial cells, RWPE-1. The top 30 upregulated or downregulated gene transcripts are accompanied by their relative fold change (Fold Δ), RefSeq identity, gene symbol, gene description and cellular function.

<https://doi.org/10.1371/journal.pone.0215504.t001>

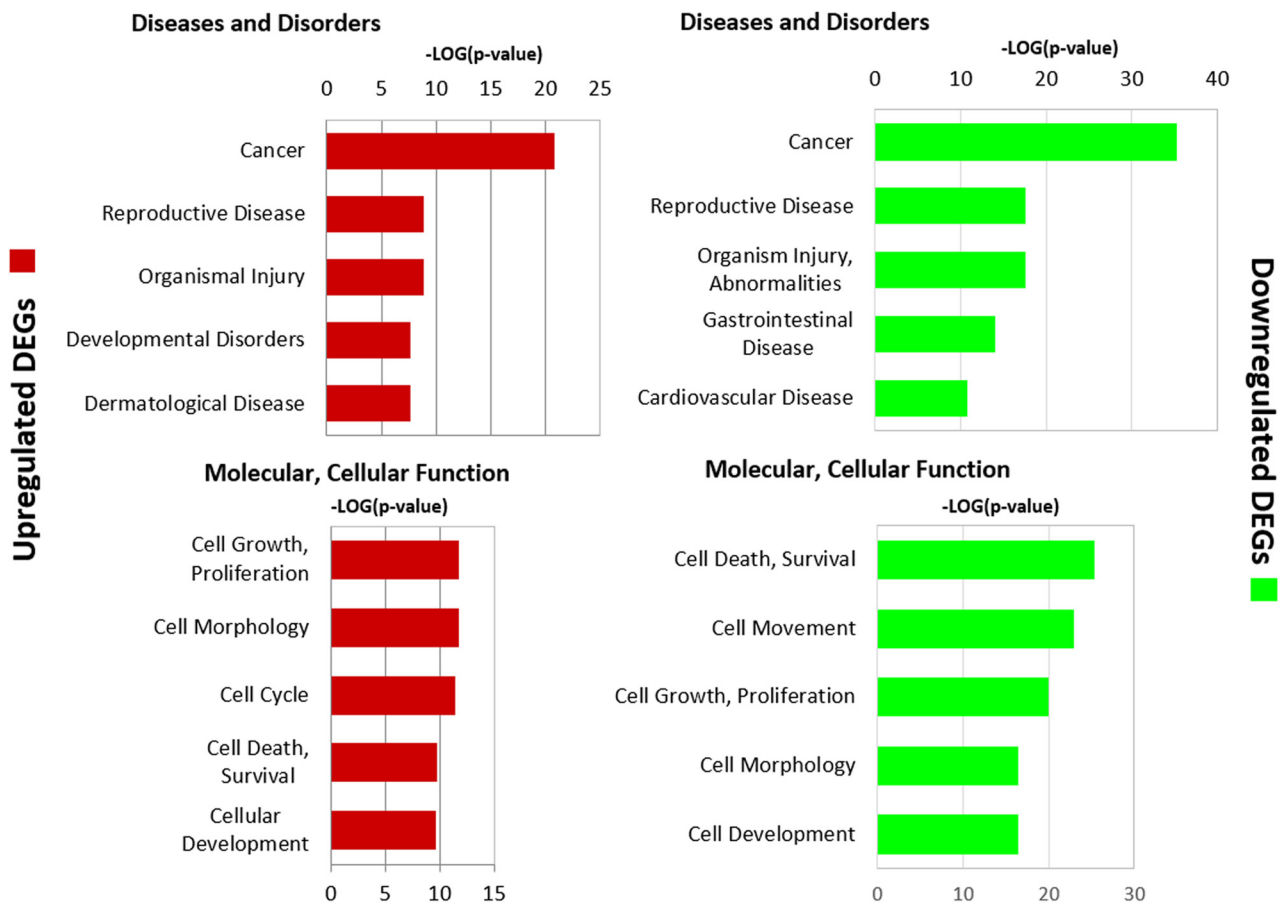
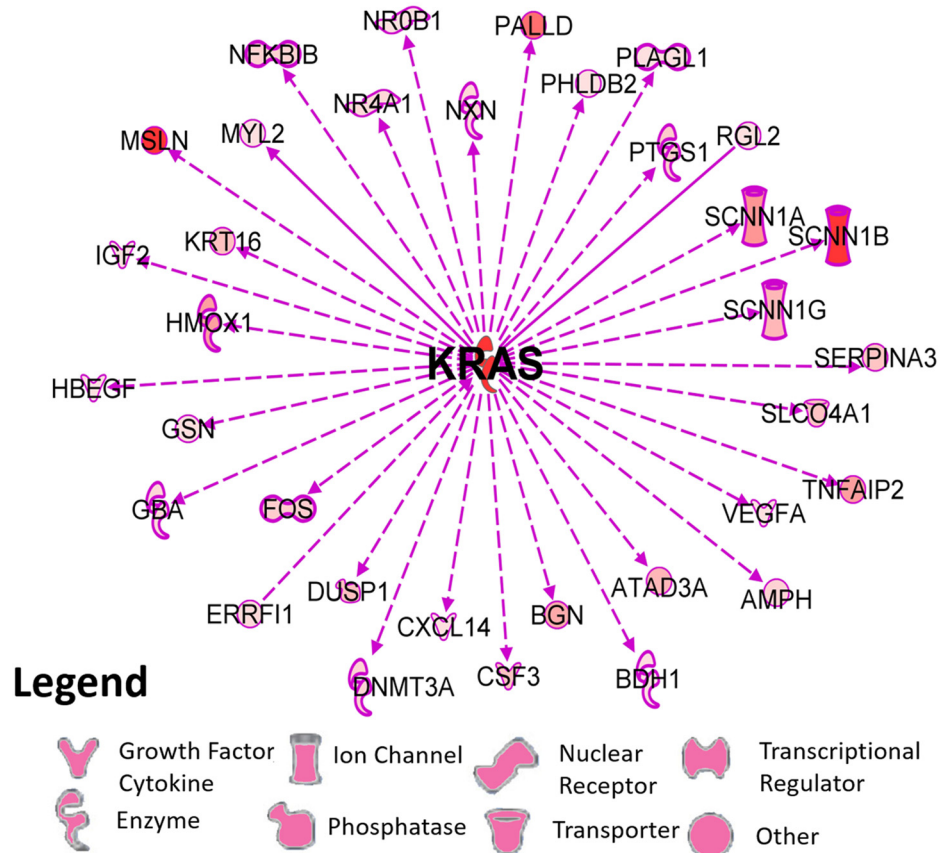


Fig 1. Differentially expressed genes (DEGs) and pathways in arsenite transformed CAsE-PE cells compared to normal RWPE-1 prostate epithelia cells. Expression profiling by RNA-seq showed over 7,000 DEGs at  $q < 0.05$  in CAsE-PE cells compared to RWPE-1 cells. DEGs for upregulated genes (red bar graphs) or downregulated genes (green bar graphs) populated pathways that are associated with diseases and health disorders, and changes in molecular and cellular pathways (IPA, Ingenuity Pathway Analysis).

<https://doi.org/10.1371/journal.pone.0215504.g001>



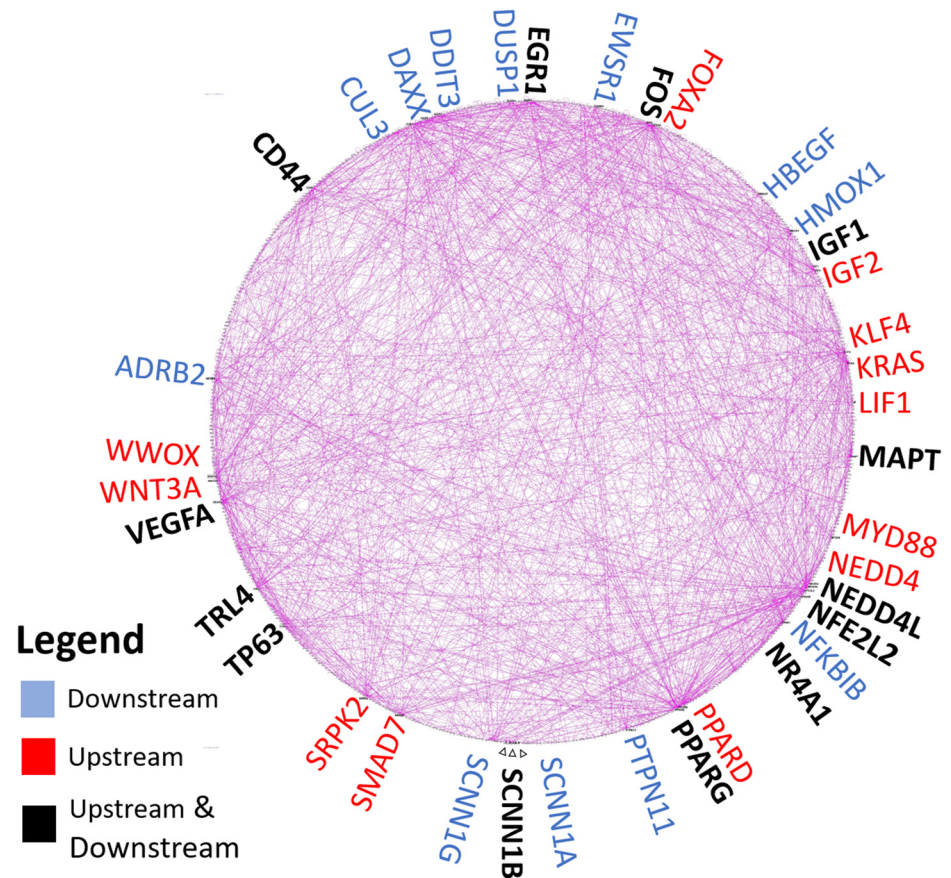
**Fig 2. Upregulated genes in CAse-PE versus RWPE-1 cells that influence KRAS expression.** Up-regulated genes were filtered by IPA annotation as either upstream and regulating KRAS (arrow tip towards KRAS) or downstream as KRAS-controlled genes (arrow tip towards gene). Solid lines indicate direct relationship, while dotted lines indicate an indirect relationship. The legend shows symbols representing primary functions of each gene.

<https://doi.org/10.1371/journal.pone.0215504.g002>

circle plot (14 bolded genes in black were common to both lists of up- and downstream relationships). Among these interconnected genes with KRAS were several growth factors, such as IGF1, IGF2, EGR1, VEGFA, HBEGF, transcription factors FOS, FOXA2, TP63, DAXX, DDIT3, and enzymes like NEDD4, NEDD4L, HMOX1, and CUL3, that are consistent with cell proliferation, stress and turnover. DAVID analysis for functionally annotated genes was also performed and showed clustering for EGF-binding domains, cell growth and developmental factors, cellular adhesion factors, Zn finger proteins and cellular membrane changes (S5 Table).

### KRAS mutations and copy number

KRAS is one of the most frequently mutated genes in human cancers. Prior work [25] had used PCR-RFLP to examine codons 12, 13 and 61 for mutations in parent RWPE-1 and arsenite malignantly transformed CAse-PE cells. These data reported an increase in transcript expression of wild type (wt, as unmutated) KRAS in CAse-PE cells. Since CAse-PE cells are malignantly transformed, we decided to re-examine the KRAS genotype in these cells by Sanger sequencing of PCR products from DNA and RNA extracted from each cell type. In normal tissues, KRAS consists of 5 coding exons, exons 1, 2, 3, 4a and 4b (see S2 Fig for structure). A 450bp invariant region includes all exons, except exon 4a and exon 4b. KRAS gene is



**Fig 3. Circular plot of gene connection analysis in CAsE-PE cells.** Connections analysis was applied to the top 1000 upregulated DEGs (limit of algorithm) in CAsE-PE vs RWPE-1 cells. All 1000 genes were plotted. IPA (Ingenuity Pathways Analysis) Connections analysis finds annotated relationships among DEGs as participating in regulating (upstream genes) or regulated (downstream genes) relationships. The number of connections were scored for each gene and are summarized in [S4 Table](#). The top 25 upstream and downstream genes are shown in large colored font around the circular plot (downstream genes in blue; upstream genes in red) with bolded (black font) genes having connections for both upstream and downstream relationships. The greater number of connections infers a higher level of regulatory and interactive relationships compared to other genes that are visually indicated by the dark convergence of connections on circle's rim for specific genes and quantitatively scored by counting the number of upstream and downstream relationships.

<https://doi.org/10.1371/journal.pone.0215504.g003>

normally expressed as two splice variants, where the variant 4b is more prevalent in normal cells compared to variant 4a [52]. Mutations in either variant renders them oncogenic, and in human cancers, variant 4a often becomes the more widely expressed KRAS transcript [53].

Initially, we designed primers to examine the genomic sequence of KRAS in both cell types and made several key observations, outlined below. First, exon-specific primers were designed from wt normal sequence of KRAS4a (NM033360) and produced amplicons with sequences aligning to wt KRAS in both cell lines. Similarly, we also designed intronic primers surrounding each exon (exons 1, 2, 3 and 4a) from wt KRAS4a and found Sanger sequences aligned to wt KRAS for RWPE-1, CAsE-PE and a normal human prostate DNA sample. Second, we considered the possibility that increased KRAS expression could be related to genomic amplification and designed assays to examine copy number variation (CNV) using wt primers. We found that DNA isolated from RWPE and normal human prostate cells showed two copies of exon 4a and exon 3 as well as two copies of introns 3 and 5 (S3 Fig). These results suggested

the presence of normal diploid KRAS alleles in RWPE-1 cells and normal prostate epithelia. However, CAsE-PE cells showed 24 copies of exon 4a, and only one copy of exon 3, intron 3 and intron 5. We interpreted these results, and the wt KRAS exon sequencing data described above, to mean CAsE-PE had one normal KRAS allele (one copy of exon 3, introns 3 and 5) even though qPCR for exon 4a showed 24 copies. We considered that the primer set for exon4a might be detecting two different KRAS alleles, one normal and one aberrant KRAS allele.

Third, the differing copy number of KRAS exons and introns in CAsE-PE cells made us suspect allelic variation of KRAS. To understand the increase in DNA copy number of exon 4a in CAsE-PE cells, experiments were designed for KRAS genomic amplification. To test this hypothesis, we performed PCR on gDNA using various primers sets spanning multiple exons that included exons 1–2, exons 2–3, exons 1-4a, exons 2-4a, and exons 3-4a. Using gDNA, we found amplicons of continuous exons of the expected size for all these regions in CAsE-PE cells only, and not RWPE-1 cells (S4 Fig).

Fourth, Sanger sequencing on the largest KRAS amplicon (exons 1-4a), spanning codons 27–182 in the preceding experiment, showed multiple sequence variants (Table 2) that

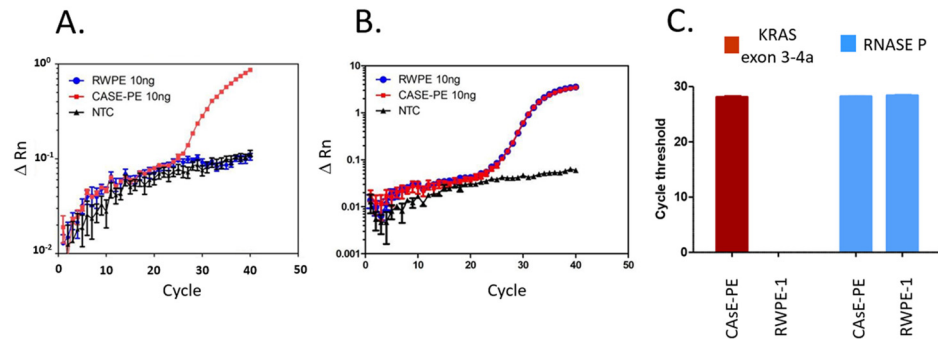
**Table 2. KRAS mutations in CAsE-PE cells.**

Mutation (CDS)	Mutation (AA)	Mutation ID (COSM No.)	Mutation Type (AA)	FATHMM	Mutation Type	A	G	C	T
c.34G>A	p.G12S	COSM1152506	Substitution–Missense	Pathogenic (score 0.98)	Transition		X		
c.90C>T	p.D30D		Substitution–coding silent		Transition			X	
c.102A>T	p.P34P		Substitution–coding silent		Transversion	X			
c.105A>G	p.T35T		Substitution–coding silent		Transition	X			
c.114T>C	p.D38D	COSM2202568	Substitution–coding silent	Pathogenic (score 0.75)	Transition				X
c.126G>A	p.K42K	COSM2202566	Substitution–coding silent	N/A	Transition		X		
c.175G>A	p.A59T	COSM1562187	Substitution–Missense	Pathogenic (score 0.98)	Transition		X		
c.219G>A	p.R73R	COSM2202564	Substitution–coding silent	Pathogenic (score 0.82)	Transition		X		
c.309T>A	p.V103V		Substitution–coding silent		Transversion				X
c.327A>G	p.V109V		Substitution–coding silent		Transition	X			
c.345A>G	p.G115G		Substitution–coding silent		Transition	X			
c.351A>G	p.K117K		Substitution–coding silent		Transition	X			
c.359T>C	p.D119D		Substitution–coding silent		Transition				X
c.381A>G	p.T127T		Substitution–coding silent		Transition	X			
c.393G>A	p.Q131Q		Substitution–coding silent		Transition		X		
c.396C>G	p.D132E		Substitution–Missense		Transversion			X	
c.405A>G	p.R135R		Substitution–coding silent		Transition	X			
c.414A>G	p.G138G		Substitution–coding silent		Transition	X			
c.420T>A	p.P140P		Substitution–coding silent		Transversion				X
c.423T>C	p.F141F		Substitution–coding silent		Transition				X
c.429A>G	p.E143E		Substitution–coding silent		Transition	X			
c.432A>C	p.T144T		Substitution–coding silent		Transversion	X			
c.438A>G	p.A146A		Substitution–coding silent		Transition	X			

Mutations in KRAS for CAsE-PE cells were determined by Sanger sequencing and confirmed by Roche 454 sequencing. The specific nucleotide change from wild type sequence is described from the CDS start site (e.g. c.34G>A) accompanied by the amino acid (AA) change (e.g. p.G12S). COSMIC database annotations were noted if present, including the type of mutation (e.g. substitution; transition) and affected nucleotide (marked by X) in far-right columns. FATHMM is the Functional Analysis through Hidden Markov Models assigning relativity pathogenicity scores to point mutations (See, <http://fathmm.biocompute.org.uk/>). Further details are described in the text.

<https://doi.org/10.1371/journal.pone.0215504.t002>





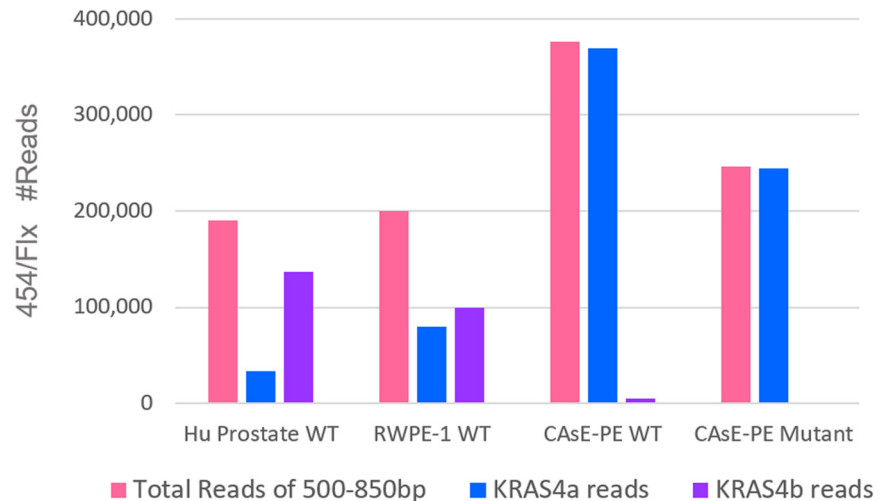
**Fig 4. qPCR analysis of RWPE-1 and CAsE-PE cells for KRAS exon-3 to exon-4a amplicon.** Each point in panel A is the mean  $\pm$  SEM of 4 replicates per sample at 10ng DNA input. RNase P in Panel B was multiplexed to show equivalent DNA input per sample. NTC is the non-template control. Panel C shows the mean  $\pm$  SEM for cycle threshold (Ct) of 4 replicates for CAsE-PE and RWPE-1 cells for KRAS and RNaseP from data in Panels A and B.

<https://doi.org/10.1371/journal.pone.0215504.g004>

suggested presence of a mutated KRAS gene allele. Missense mutations were present at codons 59 and 132 as well as twenty synonymous mutations. A redesign of primers was necessary to obtain a complete sequence of exon 1 (S6 Table) and sequence results showed a missense mutation at codon 12 (p.G12S). A table of all KRAS mutations is shown in Table 2. The Catalogue Of Somatic Mutations in Cancer (COSMIC) database was reviewed for annotated mutations. For the three missense mutations, two had COSMIC identifications as ‘oncogenic’ at codons 12 and 59 while the mutation at codon 132 has not yet been annotated. The majority of KRAS mutations were silent coding mutations for which three were identified in the COSMIC database at codons 38, 42 and 73. For mutation types within KRAS, we noted more transition mutations (18) than transversions (5). Comparison of the CAsE-PE mutated allele sequence with the pseudogene, KRASP1, shows KRAS mutations were almost completely unique from the pseudogene (S7 Table) and suggests results were not confounded by KRASP1.

Fifth, increased copy number of mutant KRAS allele was confirmed by a qPCR assay using primers crossing over exon3 and exon4A. qPCR included RNase P as a housekeeping gene to demonstrate equivalent DNA input. Fig 4 shows plots of the  $\Delta R_n$  (normalized reporter value) against the PCR cycle for the KRAS crossover amplicon (Panel A) and RNase P (Panel B). The Ct (cycle threshold above background) for RWPE-1 and CAsE-PE samples are shown in Panel C. The CAsE-PE samples show clear amplification curves for KRAS exon3-4a amplicon formation while there was no product found in the RWPE-1 samples or the non-template controls. Amplification of RNase P shows equivalent amount of DNA input into the assay as an internal control.

Finally, we wanted to determine the relative proportion of transcripts for KRAS variants 4a and 4b. RNA was isolated from RWPE-1 cells, CAsE-PE cells and normal human prostate cells and cDNA was reverse transcribed with an oligo dT primer. PCR primers were designed from wt KRAS as well as mutant KRAS derived from the mutated allele sequence. Amplicons were produced (S5 Fig) for all three cell types for wt KRAS, but only CAsE-PE produced a mutant amplicon. Indexed libraries were constructed from amplicons from each sample for 454 sequencing. Sequences were filtered for reads between 500 and 850 bp and then aligned to the two splice variants. Those reads containing mutated bases as described in Table 2 were designated as ‘mutant KRAS’ and those reads without those base changes were designated as, ‘wild type KRAS’. Cumulative read counts for 454 sequencing (S8 Table) are shown in Fig 5. The normal human prostate cells had a 4:1 higher proportion of wt KRAS variant 4b compared to variant 4a. Slightly more KRAS4b to 4a reads (55:45) were observed for RWPE cells. Only



**Fig 5. KRAS isoforms, KRAS4a and KRAS4b were measured by 454 sequencing.** Sequences were filtered for reads between 500 and 850 bp and then aligned to the two KRAS splice variants. Those reads containing the mutated bases in Table 2 were described as 'mutant KRAS' and those reads without those base changes were described as 'wild type KRAS'. See Methods for further details.

<https://doi.org/10.1371/journal.pone.0215504.g005>

CAsE-PE produced amplicons from mutant primers that were almost completely mutant variant KRAS4a transcript reads with only a few reads detected for mutant variant KRAS4b transcripts. By comparison, CAsE-PE RNA produced amplicons from wt KRAS primers with a high proportion of mutant KRAS4a variant and lesser amounts of reads aligning to the wt KRAS4b variant. In summary, both RWPE and CAsE-PE expressed wt KRAS4a and KRAS4b but only CAsE-PE cells over-expressed a mutated, oncogenic form of KRAS primarily comprised of variant KRAS4a.

## Discussion

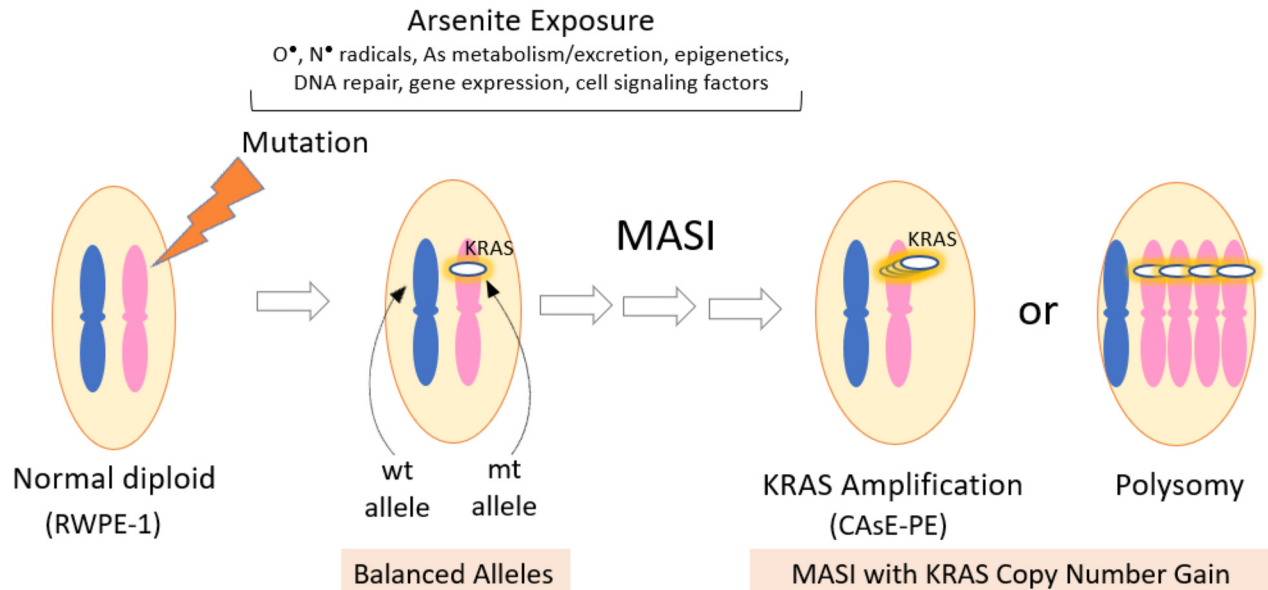
A major finding of this study was the highly mutated nature of a KRAS allele as well as its substantial overexpression. Missense mutations at codon 12 (p.G12S) and codon 59 (p.A59T) are both pathogenic as documented in the COSMIC database of human cancer mutations [54]. Codon 12 and codon 59 KRAS mutations are most commonly found in the large intestine tumors and to a lesser extent in lung, hematopoietic and stomach tumors. The sequence variant at codon p.D132E has not yet been reported for KRAS in human tumors. It is notable the other sequence variants result in silent mutations, most of which have not been annotated. KRAS mutations likely accumulate through clonal selection over months of arsenite exposure. Importantly, the concurrence of wt and amplified mutant KRAS DNA in CAsE-PE cells is consistent with the genomic aberration, 'mutant allele specific imbalance', known as MASI, that is frequently observed for oncogenes in many primary tumors [55]. The incidence of KRAS MASI genotypes range from about 12 to 18% of clinical cases that include colorectal [56, 57] and lung [58], and pancreatic [59] cancers involving mutations in codons 12 or 13. One study that examined over 400 human tumors, cell lines and xenografts of various tumor types reported that KRAS MASI was often present in primary tumors and malignantly transformed cell lines and that the combination of oncogene mutation, copy number gains and MASI may have a greater role in development and maintenance of malignancy than any individual alteration alone [47]. A more recent study involving more than 1,110 KRAS mutant tumors found an incidence of allelic imbalance at 55% [60]. Even though malignantly transformed cells may

contain multiple gene mutations, sequencing studies suggest that mutations in driver oncogenes are often mutually exclusive, meaning the actions of one oncogene may predominate over others [61]. In recent work, knockdown of KRAS expression in CAsE-PE cells greatly reduced several malignant characteristics including anchorage independent growth and metalloprotease activity [38]. Thus, the high level of KRAS gene amplification, transcript expression and activating mutations suggest it is the lead oncogene driving the malignantly transformed CAsE-PE phenotype.

We also measured relative expression of the two splice variants, KRAS4a and KRAS4b, based on long 454 reads. Read sequences containing all detected mutations were called 'mutated KRAS' and if no changes were observed from the KRAS invariant sequence, reads were designated 'wild type KRAS'. KRAS4b is the major splice variant in normal tissues and organs, while the expression of KRAS4a variant is proportionately less and varies from moderate (e.g. colon, pancreas), to low (e.g. lung, prostate) to undetectable (e.g. brain, heart) amounts [62]. Here, normal human prostate early passage epithelia expressed KRAS transcripts (Caucasian donor) had a higher proportion of wt KRAS variants 4b over 4a at a 4:1 ratio. By comparison, RWPE-1 cells had a slightly greater amount of 4b over 4a variants which were both wild type. However, reads for mutant KRAS4a transcripts in CAsE-PE cells were at a much higher ratio at >100:1 ratio compared to either wt KRAS4b reads or the few detectable mutant KRAS4b reads. This is consistent with others, reporting that KRAS4a is more highly expressed in many cancers [53, 63, 64] compared to KRAS4b. The relative variant expression may be important since different biological activities of variants 4a or 4b may uniquely contribute to the malignant transformation process and maintenance of a malignant phenotype. Given the growth-inhibitory actions of wt KRAS in tumors [65, 66], CAsE-PE cells fit an established pattern of increased copy number gains of mutated KRAS resulting in allelic imbalance (MASI) that counteracts the presence of any wt KRAS transcripts. Whether continued exposure of CAsE-PE cells to arsenite might eventually lead to complete loss of wt KRAS has yet to be determined.

How mutations arise in the parent RWPE-1 cells during arsenite exposure and malignant transformation is not clear. Arsenite is recognized as an indirect carcinogen, acting by a combination of redox reactions, DNA damage by oxygen and nitrogen radicals, interference of DNA replication and repair, and epigenetic changes [67–70]. Specifically, arsenite may produce multiple reactive oxygen and nitrogen species that result in DNA damage and mutations [71, 72]. Among various DNA lesions formed by oxidative stress, 8-oxo-guanine is one of the most common and well-studied oxidative species which frequently results in G→T transversions [73, 74]. Here, we primarily observed purine transitional mutations in KRAS and 4 transversions which were generally silent. The p.G12S substitution at codon 12 was a G>A transition mutation as was the p.A59T substitution at codon 59. These data suggest the mechanism for arsenite-induced mutations is likely more complicated than oxidative DNA damage alone in the CAsE-PE malignant transformation model.

The molecular processes for malignant transformation of RWPE-1 cells into CAsE-PE cells after prolonged 5  $\mu$ M arsenite exposure are not completely clear [35]. A 5  $\mu$ M concentration of arsenite was selected because of minimal cytotoxicity [36] and absence of oxidative DNA damage [75] in RWPE-1 cells. In other studies, when RWPE-1 cells were exposed to low concentrations of arsenite at 100 pg/ml for six months, changes in expression of epigenetic regulatory genes were observed along with global changes in DNA methylation and histone modifications, suggesting epigenetic alterations may also contribute to malignant arsenite transformation [76]. What we do report in this study is formation of an arsenite-induced, oncogenic mutated KRAS allele that becomes amplified in the CAsE-PE line. Increased copy number or gene amplification is a common observation in different types of human tumors. For example,



**Fig 6. Overview of arsenic-induced malignant transformation in CAsE-PE cells.** Prolonged, continuous exposure of RWPE-1 prostate epithelial cells to 5  $\mu$ M sodium arsenite results in the malignantly transformed CAsE-PE cell line [35]. Our observations suggest that arsenite-induced changes in various cellular factors produce the CAsE-PE cell line with a mutant (mt) KRAS allele and a wild type (wt) allele, in a process termed, mutant allele specific imbalance (MASI). The mt KRAS allele becomes amplified in what is likely a multi-step process (multiple arrows) for a copy number gain that does not involve polysomy. RNA-seq showed a substantial gene expression differences between RWPE-1 and CAsE-PE cells. Figure adapted from Yu et al [55]. Please see text for further explanation.

<https://doi.org/10.1371/journal.pone.0215504.g006>

high copy numbers of ERBB2, TOP2A, CCND1, EGFR and MYC are observed in many colorectal cancers [77]. Wide regions of DNA (kilobases to megabases) containing these oncogenes can be organized as extrachromosomal copies (double minutes), as tandem repeats within a chromosome or may be distributed at various sites within the genome [78]. KRAS amplifications has been reported in various malignancies, primarily in pancreatic [59, 79], colorectal [80, 81] and lung [58, 82] tumors, but also have been noted in gastric [83], ovarian [84] and endometrial [85] malignancies, but less frequently in prostate cancers [46]. In our study, the data did not suggest either amplification of large chromosomal portions of DNA encapsulating the KRAS gene, or chromosome duplication (e.g. polysomy); rather the data indicate an insertion of a fully processed mutated KRAS transcript into the genome resulting in increased copy number within CAsE-PE cells. The mechanism of how this type of amplification might occur is currently under investigation. Fig 6 provides a summary of arsenite-induced malignant transformation in RWPE-1 cells beginning with DNA damage and mutation resulting in mutant KRAS allele imbalance and then KRAS gene amplification in CAsE-PE cells.

Expression analysis by RNA-seq shows As transformation activates pathways and cellular processes supporting cell proliferation for tumor development in CAsE-PE cells [29]. Prominent upregulation of KRAS transcripts stands out as a leading driver of malignant transformation in CAsE-PE cells because of its known role in GTPase-mediated cell signaling and proliferation [86]. RAS proteins are master regulators of multiple downstream signaling cascades [87]. Some of the most proximal pathways for KRAS activation involve Raf-MEK-ERK (MAPK) and PI3K/Akt signaling [88]. In our study, we filtered elevated CAsE-PE transcripts and found 34 genes that had annotated interactions with KRAS (S3 Table), including growth factor genes like CSF3, HBEGF, IGF2 and VEGFA and genes for transcriptionally active proteins such as FOS, NR0B1, NR4A1 and PLAGL1. Establishing KRAS-signaling effector genes

after global expression screening has often required use of innovative experimental approaches. For example, subtractive suppression hybridization (SSH) was combined with specific signaling inhibitors of known MAPK and PI3K effector pathways to identify >200 differentially expressed genes associated with KRAS-induced malignant transformation [89]. A more recent study involved creation of a transformable, mouse cell line with an inducible, oncogenic Kras<sup>G12D</sup> mutation that was used for exploring downstream signaling genes in pancreatic cancer [90]. These researchers found that KRAS activation induces an EGFR signaling loop to drive proliferation through c-MYC to play an important role for pancreatic carcinogenesis. In our study, the 2.3-fold and 3.0-fold increases in MYC and EGFR, respectively, suggest a contributory but minor role of these pathways in CAsE-PE cells.

We also performed connectivity analysis and results suggest a sizeable number of annotated relationships occur among upregulated CAsE-PE transcripts. Among the twenty-five genes with the highest number of functional interconnections, we found upregulated CAsE-PE genes that strongly support cell proliferation (in addition to KRAS) such as growth factors (EGR1, HBEGF, IGF1, IGF2, LIF, VEGFA, WNT3A), transcription factors (FOS, NFE2L2, NR4A1, FOXA2) and the nuclear receptors, PPARG and PPARD. A related connectivity finding of interest was that SCNN1 $\alpha$ ,  $\beta$  and  $\gamma$  isoforms of the epithelial sodium ion channel, or ENaC, were each highly upregulated at 23-fold, 52-fold and 16-fold, respectively in CAsE-PE cells. The three SCNN1 isoforms comprise a heteromeric complex that forms a nonvoltage-gated sodium channel to regulate fluid, electrolytes and cell mobility in epithelial cells [91] and along with acid-sensing ion channels (ASICs) are part of the ENaC/degenerin family of sodium channels [92]. In prostate cells, this family of sodium channels is responsible for polarized movement of fluids and proteins for acinar lumen formation [93] but these normal ENaC activities may be repurposed for cell proliferation, migration and invasion during tumor development and progression [94]. While further work will be needed to determine if sodium channels are a direct consequence of KRAS overexpression, researchers have proposed ENaC should be targets for therapeutic intervention in various cancers [94–96].

## Conclusion

CAsE-PE cells comprise a well-studied *in vitro* human epithelial malignant transformation model produced by continuous low level arsenite exposure. Expression analysis showed >3,000 upregulated and about 4,000 downregulated transcripts that support increased cell growth, motility, survival and tumorigenic pathways consistent with prostate cancer. KRAS transcript was highly expressed in CAsE-PE cells, consistent with our prior work [25, 38], and is driven in part by an increase in KRAS copy number. Genomic analysis revealed a KRAS allelic imbalance with high expression of a mutated transcript carrying oncogenic mutations at codons 12 and 59 and many silent mutations, accompanied by relatively low expression of a wt allele. KRAS4a is the predominant transcript variant in CAsE-PE cells compared to KRAS4b in parental RWPE-1 epithelia and compared to normal primary prostate epithelium. These data are consistent with KRAS driven proliferation pathways found in spontaneous tumors or cell lines. Future work will focus on how KRAS becomes amplified in the *in vitro* arsenic transformation model to provide further insight into this important environmental carcinogen.

## Material and methods

### Cells and cell culture

RWPE-1 cells were originally isolated from normal human prostate epithelial cells and immortalized with a single copy of human papillomavirus 18 (HPV 18), are diploid, and have been shown to be non-tumorigenic, showing anchorage dependence and no tumors in nude mice

[34, 97]. RWPE-1 cells were originally obtained from the Webber lab [97] and were grown in K-SFM containing 50 µg/ml bovine pituitary extract, 5 ng/ml epidermal growth factor, supplemented with antibiotic–antimycotic mixture. Cells were incubated at 37°C in a humidified atmosphere containing 5% CO<sub>2</sub> and passaged weekly. CAsE-PE cells were originally developed in our lab by Waalkes [35] and maintained by Tokar and colleagues [38] after continuous exposure of parental cells to 5 µM sodium arsenite where media was changed 3 times per week for 29–30 weeks. RWPE-1 (control) and CAsE-PE cells (arsenic transformed) were harvested from three different cultures and frozen at -80°C until DNA and RNA isolation.

### DNA and RNA isolation

Genomic DNA and RNA were isolated from frozen cell pellets using RNeasy or DNeasy spin columns (Qiagen Valencia, CA, USA). For RNA-seq, RNA libraries were created from three independent isolates of RWPE-1 and CAsE-PE cells. Starting with 1 µg total RNA, polyA-tailed mRNA was isolated by oligo(dT) and fragmented by adaptive focused acoustic energy (Covaris Inc., MA, USA). A random hexamer primed, cDNA library of nucleotide sequences (350 bp median fragment size) was created from which millions of short DNA 100bp paired end reads were generated. Sequencing was performed on an Illumina HiSeq2000 instrument (Illumina, San Diego, CA, USA).

DNA was sheared in a Covaris instrument (S220, Woburn, MA) from which a DNA library was created by priming with random hexamers and nucleotide sequences (400 bp median fragment size) were created from 100bp DNA reads in paired-end orientation. Sequencing was performed in two lanes per cell line on an Illumina HiSeq2000 instrument (San Diego, CA) for which each lane produced 240 million raw reads.

### Bioinformatic analysis: Alignment of paired end reads

RNA samples were sequenced by the standard Illumina protocol to create raw sequence files (.fastq files) which underwent quality control analysis using FastQC (<http://www.bioinformatics.babraham.ac.uk/projects/fastqc/>). Quality Control (QC) plots are provided in the supplementary information (S1 Fig). We aligned the quality checked reads to the hg19 of human genome using TopHat version 2.0.11 (parameters: max gap length: 5, max mismatches: 5, max edit distance: 5, read realign edit dist: 0, mate inner distance: 250)[50].

Aligned reads were converted to UCSC genome browser tracks and uploaded to the browser to allow for visual inspection of normalized signal at any genomic location. The UCSC browser tracks contain RPKM (Reads Per Kilobase per Million reads) normalized read counts. Deep sequencing fastq data files for expression analysis and alignment are stored in the Sequence Read Archive (SRA) under Study Accession Nos. SUB4913016 and PRJNA514436.

### Analysis of differential gene expression

Differentially expressed genes (DEGs) were identified using Cuffdiff version 2.2.0 [98, 99]. Differential expression analysis was done on the human RefSeq transcriptome.

### Pathway analysis

The set of differentially expressed genes from RNA-Seq was generated (2-fold change,  $q < 0.05$ ) and used as input for Ingenuity Pathway Analysis (IPA) software (licensed use of Ingenuity Systems, [www.ingenuity.com](http://www.ingenuity.com)). Core Analysis was performed (Fig 1) to determine top canonical and disease pathways populated by differential expression. The significance value

associated with overrepresented pathways measures the likelihood of an association between an experimental gene set and a reference gene set for a specific process or pathway. The p-value is calculated with the right-tailed Fisher's Exact Test. Ingenuity (IPA) uses public databases (e.g. HumanCyc) and performs in-house curation to formulate and update signaling pathways and gene transcript and product interactions.

In our study, we focused on using differentially upregulated transcripts to build IPA 'Grow' pathways and IPA Connectivity analysis since there was more enrichment of annotated interactions for cell-proliferation and transformation processes compared to using down-regulated genes. The IPA Build pathway function used the 'Grow' feature (Fig 2) to show annotated KRAS relationships with other upregulated RefSeq DEGs (filtered to top 500 DEGs to focus on the most critical relationships). The Connect feature displayed Ingenuity-curated relationships among the top upregulated 1,000 DEGs (gene limit of connectivity algorithm), shown in Fig 3 and results were displayed in a circle graph to reduce visual complexity of gene-to-gene relationships. The number of connections to each DEG were counted as either a regulatory ability (upstream arrows 'From' each DEG) or as a regulated gene with a downstream connection (arrows 'To' regulated DEG). The higher the number of connections, the more overall regulatory importance was inferred about connections for a specific differentially expressed gene. The number of regulatory connections for each DEG (either 'To' or 'From' connections to other genes) were rank ordered for the top 25 'To' (Downstream) or 'From' (Upstream) connections and appear in large font in Fig 3. Bolded genes in black font in the Total Scoring tab in S4 Table were those common genes that were shared in the Top 25 'To' (Downstream) or 'From' (Upstream) connections.

### KRAS sequencing and CNV

A series of exon primers were designed from the DNA sequences of KRAS (NM033360.3) that were anchored in exons or surrounding intron-exon boundaries (S6 Table). Each exon was PCR amplified from gDNA (Phusion high fidelity DNA polymerase, NEB Inc, Ipswich ME), gel-purified and Sanger sequenced. Some experiments also involved amplification across multiple exons (exons 1–2, 2–3, 1–4a, 2–4a, and 3–4a) and the resulting amplicons were also gel purified prior to sequencing. Exon 1 was specifically examined for DNA sequence variation (e.g. codons 12 and 13) using primers at the intron-exon junctions, using wt and mutant primers (S6 Table), using sequence data derived from DNA-seq of CAsE-PE samples. CAsE-PE samples had an amplicon of the expected size whereas RWPE had nonspecific amplification (S5 Fig). The amplicons of the target size from the CAsE-PE samples were gel extracted and purified for Sanger Sequencing.

Quantitative differences in KRAS from genomic DNA were determined between the two cell types using exon specific primers. Specific KRAS exons were PCR amplified by Phusion DNA polymerase and quantitated by qPCR (Model 7900HT, ABI Life Technologies, Foster City CA USA) using a FAM labeled TaqMan probe. Three separate assays were also formulated to test if introns of the KRAS gene were part of a duplicated genomic region containing the KRAS gene. Amplicons for each intron were 150–200bp in length with a primer set targeting intron 3 between exons 2 and 3, and a commercial TaqMan primer set for intron 5 (Assay ID Hs05280621\_cn, Life Technologies #4400291) between exons 4a and 4b (S6 Table). Possible interference of a pseudogene, KRASP1 (NC\_000006.12) was accounted by designing primers and probes for amplicons in the sequence specific regions of KRAS that were distinguishable from it. All Taqman targets were tested prior to qPCR with traditional PCR and Sanger sequencing to ensure the generation of a single product, which matches exactly the KRAS gene sequence (and not the pseudogene retro-KRAS). Gel separation was performed to ensure a

single amplicon from PCR reactions. DNA samples from CAsE-PE and RWPE-1 were amplified in quadruplicate using each custom Taqman assay. DNA isolated from normal prostate of a male Caucasian (Coriell Institute, Camden, NJ; Cat# NA17223) was included for copy number comparison.

TaqMan Copy Number Assays were conducted simultaneously with a TaqMan Copy Number Reference Assay in a duplex real-time polymerase chain reaction (PCR). The Copy Number Assay detects the target gene or genomic sequence of interest, and the Reference Assay detects a sequence from a known two-copy gene, human RNase P H1. Copy number is determined by relative quantitation (RQ) from the  $\Delta\Delta C_t$  method. The  $\Delta C_t$  difference is measured between target and reference sequences, and then compared to the  $\Delta C_t$  values of test samples and a calibrator sample, known to have two copies. The copy number of the target is calculated as two times the relative quantity. In a copy number quantitation reaction, purified genomic DNA is combined with the TaqMan Copy Number Assay, containing two primers and a FAM dye labeled MGB probe to detect the genomic DNA target sequence. The TaqMan Copy Number Reference Assay contains two primers and a VIC dye-labeled TAMRA probe to detect the genomic DNA reference sequence. The TaqMan Genotyping Master Mix, contains AmpliTaq Gold DNA Polymerase, UP (Ultra Pure) and dNTPs required for the PCR reactions. The reference assay was run simultaneously in each well. The relative copy number of each exon for the CAsE-PE sample was compared to the RWPE-1 DNA sample using the CopyCaller v2.0 software (Applied Biosystems).

## 454 sequencing

Total RNA from RWPE, CAsE-PE and normal human prostate epithelial cells (Sciencell, Carlsbad, CA; total RNA isolated from de-identified human Caucasian primary, early passage, epithelial cultures, Cat No. 4415) was used to synthesize cDNA with a Roche cDNA Synthesis System (Cat No. 11 117 831 001) and an oligo (dT)<sub>15</sub> primer. Primers for amplification targeted wt KRAS (NM\_033360) and mutant KRAS sequences as shown in [S6 Table](#) and amplicons were visualized by electrophoresis ([S5 Fig](#)). CAsE-PE RNA generated primarily one amplicon using “mutantcDNA3” primers ([S6 Table](#)) while no amplification products were observed with RWPE and normal prostate samples. Two amplicons were generated from the RWPE and prostate samples, while one amplicon was primarily observed with CAsE-PE samples (gel slices were taken and extracted above this amplicon band). The two bands were consistent with the two known KRAS splice variants, KRAS4a and KRAS4b. Indexed libraries were constructed with the Roche Rapid Library Preparation Method Manual using the 454-sequencing adapter. Libraries were pooled together for sequencing on a Roche 454 GS Flx instrument. Sequences were filtered to only visualize reads between 500 and 850 bp in length to target reads that comprised the entire amplicon of each variant. Sequences were aligned to KRAS (variant 4a, NM\_033360 and variant 4b, NM\_004985). 454 reads were described as ‘mutant KRAS’ if the sequence contained the base changes described in [Table 2](#) found by Sanger sequencing and 454 reads were designated as ‘wildtype’ KRAS if mutant base changes were absent upon alignment to NCBI mRNA.

## Supporting information

**S1 Fig. Quality metrics for RNA-seq.**  
(PPTX)

**S2 Fig. Structure of KRAS variants, KRAS4a and KRAS4b.**  
(DOCX)



**S3 Fig. qPCR of CNVs for KRAS exons 4a and exon 3, intron 5 and intron 3.**  
(PPTX)

**S4 Fig. KRAS cross-exon transcripts in CAsE-PE shown by PCR bands in gels.**  
(PPTX)

**S5 Fig. PCR amplification of wt and mutant cDNA KRAS in RWPE-1, CAsE-PE and normal human prostate cells.**  
(PPTX)

**S1 Table. RNA-seq total reads and alignment summary.**  
(XLSX)

**S2 Table. RNA-seq DEGs by CuffDiff for transformed CAsE-PE cells compared to control RWPE-1 cells.**  
(XLSX)

**S3 Table. 34 gene references for KRAS annotations.**  
(XLSX)

**S4 Table. DEG gene interaction tables for data in Fig 3.**  
(XLS)

**S5 Table. DAVID analysis.**  
(XLSX)

**S6 Table. Primers for PCR, Sanger sequencing and 454 sequencing.**  
(XLSX)

**S7 Table. KRASP1 and KRAS alignment.**  
(DOCX)

**S8 Table. 454 sequencing read counts.**  
(XLSX)

## Acknowledgments

Internal manuscript reviews by Dr. Alison H. Harrill and Dr. Amy Wang of the DNTP are greatly appreciated. Internal and logistical support by Dr. John R. Bucher and Dr. Raymond R. Tice, at DNTP is gratefully acknowledged.

## Disclaimer

This article is the work product of employees of the National Institute of Environmental Health Sciences (NIEHS), National Institutes of Health (NIH); however, the statements, opinions or conclusions contained therein do not necessarily represent the statements, opinions or conclusions of NIEHS, NIH or the United States government.

## Author Contributions

**Conceptualization:** B. Alex Merrick, Scott S. Auerbach, Richard S. Paules, Michael J. DeVito, Michael P. Waalkes, Erik J. Tokar.

**Data curation:** Ruchir R. Shah, Garron M. Wright.

**Formal analysis:** B. Alex Merrick, Dhiral P. Phadke, Meredith A. Bostrom, Ruchir R. Shah, Garron M. Wright.

**Investigation:** B. Alex Merrick, Meredith A. Bostrom, Xinguo Wang, Oksana Gordon.

**Methodology:** Dhiral P. Phadke, Meredith A. Bostrom, Xinguo Wang, Katherine E. Pelch, Erik J. Tokar.

**Resources:** Richard S. Paules, Michael J. DeVito, Michael P. Waalkes, Erik J. Tokar.

**Software:** Ruchir R. Shah.

**Supervision:** B. Alex Merrick, Erik J. Tokar.

**Validation:** Meredith A. Bostrom, Katherine E. Pelch.

**Visualization:** Dhiral P. Phadke, Ruchir R. Shah.

**Writing – original draft:** B. Alex Merrick.

**Writing – review & editing:** B. Alex Merrick, Dhiral P. Phadke, Scott S. Auerbach, Richard S. Paules, Michael J. DeVito, Michael P. Waalkes, Erik J. Tokar.

## References

1. Arsenic and inorganic arsenic compounds. *Rep Carcinog*. 2011; 12:50–3. Epub 2011/08/11. PMID: [21829252](https://pubmed.ncbi.nlm.nih.gov/21829252/).
2. Tchounwou PB, Centeno JA, Patolla AK. Arsenic toxicity, mutagenesis, and carcinogenesis—a health risk assessment and management approach. *Mol Cell Biochem*. 2004; 255(1–2):47–55. Epub 2004/02/20. PMID: [14971645](https://pubmed.ncbi.nlm.nih.gov/14971645/).
3. Cohen SM, Chowdhury A, Arnold LL. Inorganic arsenic: A non-genotoxic carcinogen. *J Environ Sci (China)*. 2016; 49:28–37. <https://doi.org/10.1016/j.jes.2016.04.015> PMID: [28007178](https://pubmed.ncbi.nlm.nih.gov/28007178/).
4. Mandal P. Molecular insight of arsenic-induced carcinogenesis and its prevention. *Naunyn-Schmiedeberg's Arch Pharmacol*. 2017; 390(5):443–55. <https://doi.org/10.1007/s00210-017-1351-x> PMID: [28229170](https://pubmed.ncbi.nlm.nih.gov/28229170/).
5. Burns FJ, Uddin AN, Wu F, Nadas A, Rossman TG. Arsenic-induced enhancement of ultraviolet radiation carcinogenesis in mouse skin: a dose-response study. *Environ Health Perspect*. 2004; 112(5):599–603. Epub 2004/04/06. <https://doi.org/10.1289/ehp.6655> PMID: [15064167](https://pubmed.ncbi.nlm.nih.gov/15064167/).
6. Salnikow K, Zhitkovich A. Genetic and epigenetic mechanisms in metal carcinogenesis and cocarcinogenesis: nickel, arsenic, and chromium. *Chem Res Toxicol*. 2008; 21(1):28–44. Epub 2007/11/01. <https://doi.org/10.1021/tx700198a> PMID: [17970581](https://pubmed.ncbi.nlm.nih.gov/17970581/).
7. Waalkes MP, Liu J, Ward JM, Diwan BA. Animal models for arsenic carcinogenesis: inorganic arsenic is a transplacental carcinogen in mice. *Toxicol Appl Pharmacol*. 2004; 198(3):377–84. Epub 2004/07/28. <https://doi.org/10.1016/j.taap.2003.10.028> PMID: [15276417](https://pubmed.ncbi.nlm.nih.gov/15276417/).
8. Rossman TG, Stone D, Molina M, Troll W. Absence of arsenite mutagenicity in E coli and Chinese hamster cells. *Environ Mutagen*. 1980; 2(3):371–9. Epub 1980/01/01. PMID: [6459229](https://pubmed.ncbi.nlm.nih.gov/6459229/).
9. Rossman TG, Uddin AN, Burns FJ, Bosland MC. Arsenite cocarcinogenesis: an animal model derived from genetic toxicology studies. *Environ Health Perspect*. 2002; 110 Suppl 5:749–52. Epub 2002/11/12. <https://doi.org/10.1289/ehp.02110s5749> PMID: [12426125](https://pubmed.ncbi.nlm.nih.gov/12426125/).
10. Guidarelli A, Fiorani M, Cerioni L, Scotti M, Cantoni O. Arsenite induces DNA damage via mitochondrial ROS and induction of mitochondrial permeability transition. *Biofactors*. 2017. <https://doi.org/10.1002/biof.1375> PMID: [28703385](https://pubmed.ncbi.nlm.nih.gov/28703385/).
11. Lee CH, Yu HS. Role of mitochondria, ROS, and DNA damage in arsenic induced carcinogenesis. *Front Biosci (Schol Ed)*. 2016; 8:312–20. PMID: [27100709](https://pubmed.ncbi.nlm.nih.gov/27100709/).
12. Huang C, Ke Q, Costa M, Shi X. Molecular mechanisms of arsenic carcinogenesis. *Mol Cell Biochem*. 2004; 255(1–2):57–66. Epub 2004/02/20. PMID: [14971646](https://pubmed.ncbi.nlm.nih.gov/14971646/).
13. Tokar EJ, Qu W, Liu J, Liu W, Webber MM, Phang JM, et al. Arsenic-specific stem cell selection during malignant transformation. *J Natl Cancer Inst*. 2010; 102(9):638–49. Epub 2010/03/27. <https://doi.org/10.1093/jnci/djq093> PMID: [20339138](https://pubmed.ncbi.nlm.nih.gov/20339138/).

14. Reichard JF, Puga A. Effects of arsenic exposure on DNA methylation and epigenetic gene regulation. *Epigenomics*. 2010; 2(1):87–104. Epub 2010/06/02. <https://doi.org/10.2217/epi.09.45> PMID: 20514360.
15. Brocato J, Costa M. Basic mechanics of DNA methylation and the unique landscape of the DNA methylome in metal-induced carcinogenesis. *Crit Rev Toxicol*. 2013; 43(6):493–514. Epub 2013/07/13. <https://doi.org/10.3109/10408444.2013.794769> PMID: 23844698.
16. Kim HG, Kim DJ, Li S, Lee KY, Li X, Bode AM, et al. Polycomb (PcG) proteins, BMI1 and SUZ12, regulate arsenic-induced cell transformation. *J Biol Chem*. 2012; 287(38):31920–8. Epub 2012/07/31. <https://doi.org/10.1074/jbc.M112.360362> PMID: 22843710.
17. Hughes MF, Beck BD, Chen Y, Lewis AS, Thomas DJ. Arsenic exposure and toxicology: a historical perspective. *Toxicol Sci*. 2011; 123(2):305–32. Epub 2011/07/14. <https://doi.org/10.1093/toxsci/kfr184> PMID: 21750349.
18. Cardoso APF, Al-Eryani L, States JC. Arsenic-Induced Carcinogenesis: The Impact of miRNA Dysregulation. *Toxicol Sci*. 2018; 165(2):284–90. <https://doi.org/10.1093/toxsci/kfy128> PMID: 29846715.
19. Chen QY, DesMarais T, Costa M. Metals and Mechanisms of Carcinogenesis. *Annu Rev Pharmacol Toxicol*. 2019; 59:537–54. <https://doi.org/10.1146/annurev-pharmtox-010818-021031> PMID: 30625284.
20. Pogribny IP, Rusyn I. Environmental toxicants, epigenetics, and cancer. *Adv Exp Med Biol*. 2013; 754:215–32. Epub 2012/09/08. [https://doi.org/10.1007/978-1-4419-9967-2\\_11](https://doi.org/10.1007/978-1-4419-9967-2_11) PMID: 22956504.
21. Pilsner JR, Liu X, Ahsan H, Ilievski V, Slavkovich V, Levy D, et al. Folate deficiency, hyperhomocysteinemia, low urinary creatinine, and hypomethylation of leukocyte DNA are risk factors for arsenic-induced skin lesions. *Environ Health Perspect*. 2009; 117(2):254–60. Epub 2009/03/10. <https://doi.org/10.1289/ehp.11872> PMID: 19270796.
22. Ren X, McHale CM, Skibola CF, Smith AH, Smith MT, Zhang L. An emerging role for epigenetic dysregulation in arsenic toxicity and carcinogenesis. *Environ Health Perspect*. 2011; 119(1):11–9. Epub 2010/08/05. <https://doi.org/10.1289/ehp.1002114> PMID: 20682481.
23. Pournara A, Kippler M, Holmlund T, Ceder R, Grafstrom R, Vahter M, et al. Arsenic alters global histone modifications in lymphocytes in vitro and in vivo. *Cell Biol Toxicol*. 2016; 32(4):275–84. <https://doi.org/10.1007/s10565-016-9334-0> PMID: 27165195.
24. Roy RV, Son YO, Pratheeshkumar P, Wang L, Hitron JA, Divya SP, et al. Epigenetic targets of arsenic: emphasis on epigenetic modifications during carcinogenesis. *J Environ Pathol Toxicol Oncol*. 2015; 34(1):63–84. PMID: 25746832.
25. Benbrahim-Tallaa L, Waterland RA, Styblo M, Achanzar WE, Webber MM, Waalkes MP. Molecular events associated with arsenic-induced malignant transformation of human prostatic epithelial cells: aberrant genomic DNA methylation and K-ras oncogene activation. *Toxicol Appl Pharmacol*. 2005; 206(3):288–98. Epub 2005/07/26. <https://doi.org/10.1016/j.taap.2004.11.017> PMID: 16039940.
26. Brown KG, Ross GL. Arsenic, drinking water, and health: a position paper of the American Council on Science and Health. *Regul Toxicol Pharmacol*. 2002; 36(2):162–74. Epub 2002/12/04. PMID: 12460751.
27. Yang CY, Chang CC, Chiu HF. Does arsenic exposure increase the risk for prostate cancer? *J Toxicol Environ Health A*. 2008; 71(23):1559–63. Epub 2008/10/17. <https://doi.org/10.1080/15287390802392065> PMID: 18923998.
28. Wu MM, Kuo TL, Hwang YH, Chen CJ. Dose-response relation between arsenic concentration in well water and mortality from cancers and vascular diseases. *Am J Epidemiol*. 1989; 130(6):1123–32. Epub 1989/12/01. PMID: 2589305.
29. Benbrahim-Tallaa L, Waalkes MP. Inorganic arsenic and human prostate cancer. *Environ Health Perspect*. 2008; 116(2):158–64. Epub 2008/02/22. <https://doi.org/10.1289/ehp.10423> PMID: 18288312.
30. Garcia-Esquinas E, Pollan M, Umans JG, Francesconi KA, Goessler W, Guallar E, et al. Arsenic Exposure and Cancer Mortality in a US-based Prospective Cohort: the Strong Heart Study. *Cancer Epidemiol Biomarkers Prev*. 2013. Epub 2013/06/27. <https://doi.org/10.1158/1055-9965.EPI-13-0234-T> PMID: 23800676.
31. Rhim JS, Li H, Furusato B. Novel human prostate epithelial cell culture models for the study of carcinogenesis and of normal stem cells and cancer stem cells. *Adv Exp Med Biol*. 2011; 720:71–80. Epub 2011/09/09. [https://doi.org/10.1007/978-1-4614-0254-1\\_6](https://doi.org/10.1007/978-1-4614-0254-1_6) PMID: 21901619.
32. Rhim JS. In vitro human cell culture models for the study of prostate cancer. *Prostate Cancer Prostatic Dis*. 2000; 3(4):229–35. Epub 2002/12/24. <https://doi.org/10.1038/sj.pcan.4500470> PMID: 12497069.
33. Chen QY, Costa M. A comprehensive review of metal-induced cellular transformation studies. *Toxicol Appl Pharmacol*. 2017; 331:33–40. <https://doi.org/10.1016/j.taap.2017.05.004> PMID: 28506836.

34. Bello D, Webber MM, Kleinman HK, Wartinger DD, Rhim JS. Androgen responsive adult human prostatic epithelial cell lines immortalized by human papillomavirus 18. *Carcinogenesis*. 1997; 18(6):1215–23. Epub 1997/06/01. PMID: [9214605](#).
35. Achanzar WE, Brambila EM, Diwan BA, Webber MM, Waalkes MP. Inorganic arsenite-induced malignant transformation of human prostate epithelial cells. *J Natl Cancer Inst*. 2002; 94(24):1888–91. Epub 2002/12/19. PMID: [12488483](#).
36. Brambila EM, Achanzar WE, Qu W, Webber MM, Waalkes MP. Chronic arsenic-exposed human prostate epithelial cells exhibit stable arsenic tolerance: mechanistic implications of altered cellular glutathione and glutathione S-transferase. *Toxicol Appl Pharmacol*. 2002; 183(2):99–107. Epub 2002/10/22. PMID: [12387749](#).
37. Benbrahim-Tallaa L, Webber MM, Waalkes MP. Mechanisms of acquired androgen independence during arsenic-induced malignant transformation of human prostate epithelial cells. *Environ Health Perspect*. 2007; 115(2):243–7. Epub 2007/03/27. <https://doi.org/10.1289/ehp.9630> PMID: [17384772](#).
38. Ngalame NN, Tokar EJ, Person RJ, Waalkes MP. Silencing KRAS overexpression in arsenic-transformed prostate epithelial and stem cells partially mitigates malignant phenotype. *Toxicol Sci*. 2014; 142(2):489–96. Epub 2014/10/03. <https://doi.org/10.1093/toxsci/kfu201> PMID: [25273566](#).
39. Severson PL, Tokar EJ, Vrba L, Waalkes MP, Futscher BW. Agglomerates of aberrant DNA methylation are associated with toxicant-induced malignant transformation. *Epigenetics*. 2012; 7(11):1238–48. Epub 2012/09/15. <https://doi.org/10.4161/epi.22163> PMID: [22976526](#).
40. Severson PL, Tokar EJ, Vrba L, Waalkes MP, Futscher BW. Coordinate H3K9 and DNA methylation silencing of ZNFs in toxicant-induced malignant transformation. *Epigenetics*. 2013; 8(10). Epub 2013/08/27. <https://doi.org/10.4161/epi.25926> PMID: [23974009](#).
41. Pelch KE, Tokar EJ, Merrick BA, Waalkes MP. Differential DNA methylation profile of key genes in malignant prostate epithelial cells transformed by inorganic arsenic or cadmium. *Toxicol Appl Pharmacol*. 2015; 286(3):159–67. Epub 2015/04/30. <https://doi.org/10.1016/j.taap.2015.04.011> PMID: [25922126](#).
42. Ngalame NN, Tokar EJ, Person RJ, Xu Y, Waalkes MP. Aberrant microRNA expression likely controls RAS oncogene activation during malignant transformation of human prostate epithelial and stem cells by arsenic. *Toxicol Sci*. 2014; 138(2):268–77. Epub 2014/01/17. <https://doi.org/10.1093/toxsci/kfu002> PMID: [24431212](#).
43. Salmaninejad A, Ghadami S, Dizaji MZ, Golchehre Z, Estiar MA, Zamani MR, et al. Molecular Characterization of KRAS, BRAF, and EGFR Genes in Cases with Prostatic Adenocarcinoma; Reporting Bioinformatics Description and Recurrent Mutations. *Clin Lab*. 2015; 61(7):749–59. PMID: [26299074](#).
44. Silan F, Gultekin Y, Atik S, Kilinc D, Alan C, Yildiz F, et al. Combined point mutations in codon 12 and 13 of KRAS oncogene in prostate carcinomas. *Mol Biol Rep*. 2012; 39(2):1595–9. <https://doi.org/10.1007/s11033-011-0898-8> PMID: [21607617](#).
45. Nodin B, Zendeherkh N, Sundstrom M, Jirstrom K. Clinicopathological correlates and prognostic significance of KRAS mutation status in a pooled prospective cohort of epithelial ovarian cancer. *Diagn Pathol*. 2013; 8:106. <https://doi.org/10.1186/1746-1596-8-106> PMID: [23800114](#).
46. Wang XS, Shankar S, Dhanasekaran SM, Ateeq B, Sasaki AT, Jing X, et al. Characterization of KRAS rearrangements in metastatic prostate cancer. *Cancer Discov*. 2011; 1(1):35–43. <https://doi.org/10.1158/2159-8274.CD-10-0022> PMID: [22140652](#).
47. Stephens RM, Yi M, Kessing B, Nissley DV, McCormick F. Tumor RAS Gene Expression Levels Are Influenced by the Mutational Status of RAS Genes and Both Upstream and Downstream RAS Pathway Genes. *Cancer Inform*. 2017; 16:1176935117711944. <https://doi.org/10.1177/1176935117711944> PMID: [28634423](#).
48. Murakami R, Matsumura N, Brown JB, Higasa K, Tsutsumi T, Kamada M, et al. Exome Sequencing Landscape Analysis in Ovarian Clear Cell Carcinoma Shed Light on Key Chromosomal Regions and Mutation Gene Networks. *Am J Pathol*. 2017; 187(10):2246–58. <https://doi.org/10.1016/j.ajpath.2017.06.012> PMID: [28888422](#).
49. Hamidi H, Lu M, Chau K, Anderson L, Fejzo M, Ginther C, et al. KRAS mutational subtype and copy number predict in vitro response of human pancreatic cancer cell lines to MEK inhibition. *Br J Cancer*. 2014; 111(9):1788–801. <https://doi.org/10.1038/bjc.2014.475> PMID: [25167228](#).
50. Merrick BA, Phadke DP, Auerbach SS, Mav D, Stiegelmeier SM, Shah RR, et al. RNA-Seq profiling reveals novel hepatic gene expression pattern in aflatoxin B1 treated rats. *PLoS One*. 2013; 8(4): e61768. <https://doi.org/10.1371/journal.pone.0061768> PMID: [23630614](#).
51. Zhigang Z, Wenlv S. Prostate stem cell antigen (PSCA) expression in human prostate cancer tissues: implications for prostate carcinogenesis and progression of prostate cancer. *Jpn J Clin Oncol*. 2004; 34(7):414–9. Epub 2004/09/03. <https://doi.org/10.1093/jjco/hyh073> PMID: [15342669](#).

52. Nussinov R, Tsai CJ, Chakrabarti M, Jang H. A New View of Ras Isoforms in Cancers. *Cancer Res.* 2016; 76(1):18–23. Epub 2015/12/15. <https://doi.org/10.1158/0008-5472.CAN-15-1536> PMID: 26659836.
53. Tsai FD, Lopes MS, Zhou M, Court H, Ponce O, Fiordalisi JJ, et al. K-Ras4A splice variant is widely expressed in cancer and uses a hybrid membrane-targeting motif. *Proc Natl Acad Sci U S A.* 2015; 112(3):779–84. Epub 2015/01/07. <https://doi.org/10.1073/pnas.1412811112> PMID: 25561545.
54. Forbes SA, Beare D, Boutselakis H, Bamford S, Bindal N, Tate J, et al. COSMIC: somatic cancer genetics at high-resolution. *Nucleic Acids Res.* 2017; 45(D1):D777–D83. <https://doi.org/10.1093/nar/gkw1121> PMID: 27899578.
55. Yu CC, Qiu W, Juang CS, Mansukhani MM, Halmos B, Su GH. Mutant allele specific imbalance in oncogenes with copy number alterations: Occurrence, mechanisms, and potential clinical implications. *Cancer Lett.* 2017; 384:86–93. <https://doi.org/10.1016/j.canlet.2016.10.013> PMID: 27725226.
56. Hartman DJ, Davison JM, Foxwell TJ, Nikiforova MN, Chiosea SI. Mutant allele-specific imbalance modulates prognostic impact of KRAS mutations in colorectal adenocarcinoma and is associated with worse overall survival. *Int J Cancer.* 2012; 131(8):1810–7. <https://doi.org/10.1002/ijc.27461> PMID: 22290300.
57. Malapelle U, Sgariglia R, De Stefano A, Bellevisine C, Vigliar E, de Biase D, et al. KRAS mutant allele-specific imbalance (MASI) assessment in routine samples of patients with metastatic colorectal cancer. *J Clin Pathol.* 2015; 68(4):265–9. <https://doi.org/10.1136/jclinpath-2014-202761> PMID: 25609577.
58. Chiosea SI, Sherer CK, Jelic T, Dacic S. KRAS mutant allele-specific imbalance in lung adenocarcinoma. *Mod Pathol.* 2011; 24(12):1571–7. <https://doi.org/10.1038/modpathol.2011.109> PMID: 21743433.
59. Krasinskas AM, Moser AJ, Saka B, Adsay NV, Chiosea SI. KRAS mutant allele-specific imbalance is associated with worse prognosis in pancreatic cancer and progression to undifferentiated carcinoma of the pancreas. *Mod Pathol.* 2013; 26(10):1346–54. <https://doi.org/10.1038/modpathol.2013.71> PMID: 23599154.
60. Burgess MR, Hwang E, Mroue R, Bielski CM, Wandler AM, Huang BJ, et al. KRAS Allelic Imbalance Enhances Fitness and Modulates MAP Kinase Dependence in Cancer. *Cell.* 2017; 168(5):817–29.e15. Epub 2017/02/22. <https://doi.org/10.1016/j.cell.2017.01.020> PMID: 28215705.
61. Cisowski J, Bergo MO. What makes oncogenes mutually exclusive? *Small GTPases.* 2017; 8(3):187–92. <https://doi.org/10.1080/21541248.2016.1212689> PMID: 27416373.
62. Plowman SJ, Berry RL, Bader SA, Luo F, Arends MJ, Harrison DJ, et al. K-ras 4A and 4B are co-expressed widely in human tissues, and their ratio is altered in sporadic colorectal cancer. *J Exp Clin Cancer Res.* 2006; 25(2):259–67. Epub 2006/08/22. PMID: 16918139.
63. Calvert RJ, Gupta M, Maciag A, Shiao YH, Anderson LM. K-ras 4A and 4B mRNA levels correlate with superoxide in lung adenocarcinoma cells, while at the protein level, only mutant K-ras 4A protein correlates with superoxide. *Lung Cancer.* 2013; 80(3):263–9. Epub 2013/03/12. <https://doi.org/10.1016/j.lungcan.2013.01.022> PMID: 23474128.
64. To MD, Wong CE, Karnezis AN, Del Rosario R, Di Lauro R, Balmain A. Kras regulatory elements and exon 4A determine mutation specificity in lung cancer. *Nat Genet.* 2008; 40(10):1240–4. Epub 2008/09/02. <https://doi.org/10.1038/ng.211> PMID: 18758463.
65. Westcott PM, Halliwill KD, To MD, Rashid M, Rust AG, Keane TM, et al. The mutational landscapes of genetic and chemical models of Kras-driven lung cancer. *Nature.* 2015; 517(7535):489–92. Epub 2014/11/05. <https://doi.org/10.1038/nature13898> PMID: 25363767.
66. Zhang Z, Wang Y, Vikis HG, Johnson L, Liu G, Li J, et al. Wildtype Kras2 can inhibit lung carcinogenesis in mice. *Nat Genet.* 2001; 29(1):25–33. Epub 2001/08/31. <https://doi.org/10.1038/ng721> PMID: 11528387.
67. Kligerman AD, Tennant AH. Insights into the carcinogenic mode of action of arsenic. *Toxicol Appl Pharmacol.* 2007; 222(3):281–8. <https://doi.org/10.1016/j.taap.2006.10.006> PMID: 17118416.
68. Rossman TG, Klein CB. Genetic and epigenetic effects of environmental arsenicals. *Metallomics.* 2011; 3(11):1135–41. <https://doi.org/10.1039/c1mt00074h> PMID: 21976018.
69. Sage AP, Minatel BC, Ng KW, Stewart GL, Dummer TJB, Lam WL, et al. Oncogenomic disruptions in arsenic-induced carcinogenesis. *Oncotarget.* 2017; 8(15):25736–55. <https://doi.org/10.18632/oncotarget.15106> PMID: 28179585.
70. Klein CB, Leszczynska J, Hickey C, Rossman TG. Further evidence against a direct genotoxic mode of action for arsenic-induced cancer. *Toxicol Appl Pharmacol.* 2007; 222(3):289–97. <https://doi.org/10.1016/j.taap.2006.12.033> PMID: 17316729.
71. Bustaffa E, Stocco A, Bianchi F, Migliore L. Genotoxic and epigenetic mechanisms in arsenic carcinogenicity. *Arch Toxicol.* 2014; 88(5):1043–67. <https://doi.org/10.1007/s00204-014-1233-7> PMID: 24691704.

72. Faita F, Cori L, Bianchi F, Andreassi MG. Arsenic-induced genotoxicity and genetic susceptibility to arsenic-related pathologies. *Int J Environ Res Public Health*. 2013; 10(4):1527–46. <https://doi.org/10.3390/ijerph10041527> PMID: 23583964.
73. Markkanen E. Not breathing is not an option: How to deal with oxidative DNA damage. *DNA Repair (Amst)*. 2017; 59:82–105. <https://doi.org/10.1016/j.dnarep.2017.09.007> PMID: 28963982.
74. Waris G, Ahsan H. Reactive oxygen species: role in the development of cancer and various chronic conditions. *J Carcinog*. 2006; 5:14. <https://doi.org/10.1186/1477-3163-5-14> PMID: 16689993.
75. Orihuela R, Kojima C, Tokar EJ, Person RJ, Xu Y, Qu W, et al. Oxidative DNA damage after acute exposure to arsenite and monomethylarsonous acid in biomethylation-deficient human cells. *Toxicol Mech Methods*. 2013; 23(6):389–95. Epub 2013/01/11. <https://doi.org/10.3109/15376516.2012.762570> PMID: 23301828.
76. Treas JN, Tyagi T, Singh KP. Effects of chronic exposure to arsenic and estrogen on epigenetic regulatory genes expression and epigenetic code in human prostate epithelial cells. *PLoS One*. 2012; 7(8): e43880. Epub 2012/09/07. <https://doi.org/10.1371/journal.pone.0043880> PMID: 22952798.
77. Al-Kuraya K, Novotny H, Bavi P, Siraj AK, Uddin S, Ezzat A, et al. HER2, TOP2A, CCND1, EGFR and C-MYC oncogene amplification in colorectal cancer. *J Clin Pathol*. 2007; 60(7):768–72. <https://doi.org/10.1136/jcp.2006.038281> PMID: 16882699.
78. Albertson DG. Gene amplification in cancer. *Trends Genet*. 2006; 22(8):447–55. <https://doi.org/10.1016/j.tig.2006.06.007> PMID: 16787682.
79. Shiroma N, Arihiro K, Oda M, Orita M. KRAS fluorescence in situ hybridisation testing for the detection and diagnosis of pancreatic adenocarcinoma. *J Clin Pathol*. 2018. Epub 2018/04/27. <https://doi.org/10.1136/jclinpath-2018-205002> PMID: 29695486.
80. Docs O, Hegyi K, Mokanszky A, Monusne A, Beke L, Andras C, et al. Mutant KRAS Status Is Associated with Increased KRAS Copy Number Imbalance: a Potential Mechanism of Molecular Heterogeneity. *Pathol Oncol Res*. 2017; 23(2):417–23. Epub 2016/10/16. <https://doi.org/10.1007/s12253-016-0126-x> PMID: 27743339.
81. Valtorta E, Misale S, Sartore-Bianchi A, Nagtegaal ID, Paraf F, Lauricella C, et al. KRAS gene amplification in colorectal cancer and impact on response to EGFR-targeted therapy. *Int J Cancer*. 2013; 133(5):1259–65. Epub 2013/02/14. <https://doi.org/10.1002/ijc.28106> PMID: 23404247.
82. Bahcall M, Awad MM, Sholl LM, Wilson FH, Xu M, Wang S, et al. Amplification of wild type KRAS imparts resistance to crizotinib in MET exon 14 mutant non-small cell lung cancer. *Clin Cancer Res*. 2018. Epub 2018/08/04. <https://doi.org/10.1158/1078-0432.CCR-18-0876> PMID: 30072474.
83. Qian Z, Zhu G, Tang L, Wang M, Zhang L, Fu J, et al. Whole genome gene copy number profiling of gastric cancer identifies PAK1 and KRAS gene amplification as therapy targets. *Genes Chromosomes Cancer*. 2014; 53(11):883–94. Epub 2014/06/18. <https://doi.org/10.1002/gcc.22196> PMID: 24935174.
84. Rahman MT, Nakayama K, Rahman M, Katagiri H, Katagiri A, Ishibashi T, et al. KRAS and MAPK1 gene amplification in type II ovarian carcinomas. *Int J Mol Sci*. 2013; 14(7):13748–62. Epub 2013/07/04. <https://doi.org/10.3390/ijms140713748> PMID: 23820584.
85. Birkeland E, Wik E, Mjos S, Hoivik EA, Trovik J, Werner HM, et al. KRAS gene amplification and overexpression but not mutation associates with aggressive and metastatic endometrial cancer. *Br J Cancer*. 2012; 107(12):1997–2004. Epub 2012/10/27. <https://doi.org/10.1038/bjc.2012.477> PMID: 23099803.
86. McCormick F. KRAS as a Therapeutic Target. *Clin Cancer Res*. 2015; 21(8):1797–801. Epub 2015/04/17. <https://doi.org/10.1158/1078-0432.CCR-14-2662> PMID: 25878360.
87. Shrestha G, MacNeil SM, McQuerry JA, Jenkins DF, Sharma S, Bild AH. The value of genomics in dissecting the RAS-network and in guiding therapeutics for RAS-driven cancers. *Semin Cell Dev Biol*. 2016; 58:108–17. Epub 2016/06/25. <https://doi.org/10.1016/j.semcdb.2016.06.012> PMID: 27338857.
88. McCubrey JA, Steelman LS, Chappell WH, Abrams SL, Wong EW, Chang F, et al. Roles of the Raf/MEK/ERK pathway in cell growth, malignant transformation and drug resistance. *Biochim Biophys Acta*. 2007; 1773(8):1263–84. Epub 2006/11/28. <https://doi.org/10.1016/j.bbamcr.2006.10.001> PMID: 17126425.
89. Tchernitsa OI, Sers C, Zuber J, Hinzmann B, Grips M, Schramme A, et al. Transcriptional basis of KRAS oncogene-mediated cellular transformation in ovarian epithelial cells. *Oncogene*. 2004; 23(26):4536–55. Epub 2004/04/06. <https://doi.org/10.1038/sj.onc.1207585> PMID: 15064704.
90. Diersch S, Wirth M, Schneeweis C, Jors S, Geisler F, Siveke JT, et al. Kras(G12D) induces EGFR-MYC cross signaling in murine primary pancreatic ductal epithelial cells. *Oncogene*. 2016; 35(29):3880–6. Epub 2015/11/26. <https://doi.org/10.1038/onc.2015.437> PMID: 26592448.
91. Hanukoglu I, Hanukoglu A. Epithelial sodium channel (ENaC) family: Phylogeny, structure-function, tissue distribution, and associated inherited diseases. *Gene*. 2016; 579(2):95–132. Epub 2016/01/17. <https://doi.org/10.1016/j.gene.2015.12.061> PMID: 26772908.

92. Boscardin E, Alijevic O, Hummler E, Frateschi S, Kellenberger S. The function and regulation of acid-sensing ion channels (ASICs) and the epithelial Na(+) channel (ENaC): IUPHAR Review 19. *Br J Pharmacol.* 2016; 173(18):2671–701. <https://doi.org/10.1111/bph.13533> PMID: 27278329.
93. Pearson JF, Hughes S, Chambers K, Lang SH. Polarized fluid movement and not cell death, creates luminal spaces in adult prostate epithelium. *Cell Death Differ.* 2009; 16(3):475–82. <https://doi.org/10.1038/cdd.2008.181> PMID: 19096393.
94. Xu S, Liu C, Ma Y, Ji HL, Li X. Potential Roles of Amiloride-Sensitive Sodium Channels in Cancer Development. *Biomed Res Int.* 2016; 2016:2190216. <https://doi.org/10.1155/2016/2190216> PMID: 27403419.
95. He M, Liu S, Gallolu Kankanamalage S, Borromeo MD, Girard L, Gazdar AF, et al. The Epithelial Sodium Channel (alphaENaC) Is a Downstream Therapeutic Target of ASCL1 in Pulmonary Neuroendocrine Tumors. *Transl Oncol.* 2018; 11(2):292–9. Epub 2018/02/08. <https://doi.org/10.1016/j.tranon.2018.01.004> PMID: 29413762.
96. Vila-Carriles WH, Kovacs GG, Jovov B, Zhou ZH, Pahwa AK, Colby G, et al. Surface expression of ASIC2 inhibits the amiloride-sensitive current and migration of glioma cells. *J Biol Chem.* 2006; 281(28):19220–32. Epub 2006/05/18. <https://doi.org/10.1074/jbc.M603100200> PMID: 16704974.
97. Webber MM, Bello D, Kleinman HK, Hoffman MP. Acinar differentiation by non-malignant immortalized human prostatic epithelial cells and its loss by malignant cells. *Carcinogenesis.* 1997; 18(6):1225–31. Epub 1997/06/01. PMID: 9214606.
98. Trapnell C, Williams BA, Pertea G, Mortazavi A, Kwan G, van Baren MJ, et al. Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. *Nat Biotechnol.* 2010; 28(5):511–5. <https://doi.org/10.1038/nbt.1621> PMID: 20436464.
99. Roberts A, Pimentel H, Trapnell C, Pachter L. Identification of novel transcripts in annotated genomes using RNA-Seq. *Bioinformatics.* 2011; 27(17):2325–9. <https://doi.org/10.1093/bioinformatics/btr355> PMID: 21697122.