Genome-Wide CRISPR Screening Identifies the Tumor Suppressor Candidate OVCA2 As a Determinant of Tolerance to Acetaldehyde

Amin Sobh,^{*,†} Alex Loguinov,^{*} Alessia Stornetta,[‡] Silvia Balbo,^{‡,§} Abderrahmane Tagmount,^{*} Luoping Zhang,[¶] and Chris D. Vulpe^{*,1}

^{*}Department of Physiological Sciences, College of Veterinary Medicine, University of Florida, Gainesville, Florida; [†]Department of Nutritional Sciences & Toxicology, Comparative Biochemistry Program, University of California, Berkeley, California; [‡]Masonic Cancer Center and [§]Division of Environmental Health Sciences, University of Minnesota, Minneapolis, Minnesota; and [¶]Division of Environmental Health Sciences, School of Public Health, University of California, Berkeley, California

¹To whom correspondence should be addressed at Department of Physiological Sciences, Center for Environmental and Human Toxicology, University of Florida, 2187 Mowry Drive, Building 471, PO Box 110885, Gainesville, FL 32611. Fax: (352)-392-2938. E-mail: cvulpe@ufl.edu.

ABSTRACT

Acetaldehyde, a metabolite of ethanol, is a cellular toxicant and a human carcinogen. A genome-wide CRISPR-based loss-of-function screen in erythroleukemic K562 cells revealed candidate genetic contributors affecting acetaldehyde cytotoxicity. Secondary screening exposing cells to a lower acetaldehyde dose simultaneously validated multiple candidate genes whose loss results in increased sensitivity to acetaldehyde. Disruption of genes encoding components of various DNA repair pathways increased cellular sensitivity to acetaldehyde. Unexpectedly, the tumor suppressor gene OVCA2, whose function is unknown, was identified in our screen as a determinant of acetaldehyde tolerance. Disruption of the OVCA2 gene resulted in increased acetaldehyde sensitivity and higher accumulation of the acetaldehyde-derived DNA adduct N²-ethylidene-dG. Together these results are consistent with a role for OVCA2 in adduct removal and/or DNA repair.

Key words: acetaldehyde; ethanol; OVCA2; CRISPR screening; DNA repair.

Acetaldehyde, a toxic intermediate in the metabolic breakdown of alcohol, is designated as a Group 1 carcinogen by the international agency for research on cancer when associated with alcohol consumption (Secretan *et al.*, 2009). Ethanol consumption is associated with increased risk of several forms of cancer, including esophageal (Brooks *et al.*, 2009; Yokoyama and Omori, 2003), liver (McKillop and Schrum, 2009), and breast cancer (Chen *et al.*, 2011). Acetaldehyde is an electrophile which targets cellular proteins and DNA through adduct formation. Cellular protein targets of acetaldehyde include erythrocyte membrane proteins, tubulin, ketosteroid reductase and the ethanol metabolizing enzyme Cytochrome P450 2E1 (CYP2E1) (Niemelä, 1999). Acetaldehyde-induced DNA adducts include the more abundant N²-ethylidene-2'-deoxyguanosine (N²-ethylidene-dG) and the

mutagenic N²-propano-2'-deoxyguanosine (PdG) (Setshedi *et al.*, 2010). PdG induces the formation of DNA-protein and DNA interstrand crosslinks (ICLs) and is implicated in acetaldehyde-induced carcinogenesis (Brooks and Theruvathu, 2005). Acetaldehyde may induce cancers by damaging the genome directly, through the formation of DNA adducts, and indirectly by the formation of adducts with proteins that maintain genomic and epigenomic stability (Brooks and Zakhari, 2014). The underlying mechanisms of toxicity and carcinogenesis and factors modulating susceptibility to acetaldehyde, and other aldehydes are poorly understood (Brooks and Zakhari, 2014; Seitz and Stickel, 2007).

Functional studies identified key processes in modulating acetaldehyde toxicity (Garaycochea *et al.*, 2012, 2018; Langevin

et al., 2011; Parmar and D'Andrea, 2012; Venton *et al.*, 2016). Individual disruption of several genes in the Fanconi Anemia (FA) DNA repair pathway increased susceptibility of cultured B cells to exogenous acetaldehyde (Langevin *et al.*, 2011). Similarly, mice lacking the acetaldehyde-catabolizing enzyme Aldh2 (Aldehyde dehydrogenase 2) and Fancd2 (Fanconi anemia complementation group D2) are highly susceptible to ethanol toxicity (Langevin *et al.*, 2011). In humans, blood levels of acetaldehyde-derived DNA modifications were significantly higher in alcoholics carrying the inactive ALDH2 allele (ALDH2*2) (Matsuda *et al.*, 2006). In addition, FA patients with ALDH2 deficiency associated with an ALDH2 variant allele showed a faster progression of bone marrow failure compared with control patients (Hira *et al.*, 2013). Inactivation of components of the homologous recombination (HR) DNA repair pathway including BRCA1/2 (BRCA1/2, DNA Repair Associated) or RAD51 (RAD51 Recombinase) increased sensitivity to exogenous acetaldehyde treatment and intracellular acetaldehyde accumulation induced by pharmacological inhibition of ALDH2 (Tacconi *et al.*, 2017). The nucleotide excision repair (NER) pathway repairs PdG adducts (Johnson *et al.*, 1997) and thus likely mitigates acetaldehyde-induced DNA damage. A comprehensive functional screen in fission yeast demonstrated key roles for the FA, HR, and NER pathways in modulating acetaldehyde toxicity and revealed roles for base excision repair (BER), translesion DNA synthesis (TLS) and DNA-protein crosslink (DPC) repair in reversing acetaldehyde-induced DNA damage (Brooks and Schuebel, 2017; Noguchi *et al.*, 2017). To our knowledge, a similar large-scale genomic approach to study mechanisms affecting cellular sensitivity and tolerance to acetaldehyde in mammalian, and particularly human cells, has not been performed.

Gene expression profiling following acetaldehyde exposure has been performed in yeast (Aranda and del Olmo, 2004) and human cells (Cheah *et al.*, 2013) and provided some insight into molecular mechanisms of acetaldehyde toxicity. However, transcriptomic analysis is not sufficient to determine mechanisms affecting susceptibility to acetaldehyde because the expression of cellular determinants of sensitivity or tolerance to acetaldehyde is not necessarily modulated by acetaldehyde exposure. In this study, we used a CRISPR-based functional screening approach to decipher molecular mechanisms influencing cellular vulnerability to acetaldehyde. We performed a genome-wide loss-of-function screen in human erythroleukemic cells exposed to acetaldehyde and identified several genes whose loss enhances susceptibility to the toxic effects of acetaldehyde. We found that disruption of multiple DNA repair pathways, including ICL repair and transcription-coupled NER (TC-NER), increased acetaldehyde sensitivity. Unexpectedly, a tumor suppressor gene, OVCA2 (Ovarian Cancer-Associated gene 2), of unknown function was identified as a primary determinant of acetaldehyde tolerance. Disruption of the OVCA2 gene confers increased acetaldehyde cytotoxicity associated with increased levels of the DNA adduct N²-ethylidene-dG, suggesting a role for OVCA2 in DNA damage response, DNA repair and/or adduct removal.

MATERIALS AND METHODS

Cell culture. Human HEK293T and K562 cells lines were obtained from the biosciences divisional services cell culture facility, UC Berkeley (<https://bds.berkeley.edu/facilities/cell-culture> last accessed February, 27 2019). HEK293T cells were cultured in DMEM (ThermoFisher 11965092) supplemented with 10% fetal

bovine serum (FBS; Corning 35-015-CV) and 1% penicillin/streptomycin (PS; ThermoFisher 15140122). K562 cells were cultured in RPMI 1640 (ThermoFisher 11835030) supplemented with 10% FBS and 1% PS. Cells were cultured in a humidified incubator with 5% CO₂ at 37°C.

Cell viability assays. Evaluation of cell viability was performed by measuring ATP levels using the CellTiter Glo luminescent cell viability assay kit (Promega G7572) following the manufacturer's protocol. Briefly, K562 cells were seeded at a density of 10⁵ cells/ml (10⁴ cells/well) in opaque 96-well cell culture plates and exposed to multiple concentrations (0–50 mM) of acetaldehyde (Sigma 402788) for 72 h. To measure the ATP content in each well, cells were mixed with 100 µl of the CellTiter Glo reagent and lysed on an orbital plate shaker for 2 min. Plates were incubated for 10 min at room temperature in the dark to stabilize the signal. Luminescent signals from all wells were read on a Synergy H1 microplate reader (BioTek Instruments).

Genome-wide and focused CRISPR/Cas9 libraries. The human genome-wide CRISPR knockout (GeCKO) version 2 sgRNA library cloned in LentiCRISPR v2 vector (Addgene No. 1000000048, kindly deposited by Dr Feng Zhang) was used for genome-wide screening. The GeCKO v2 library targets 19 050 protein-coding genes and 1864 miRNAs and contains 1000 nontargeting sgRNAs with a total of 123 411 sgRNAs split into 2 half-libraries A and B (Sanjana *et al.*, 2014). For primary screening, we used half-library A containing 65 383 sgRNAs with an average of 3 sgRNAs targeting each gene. For the focused (validation) library, sgRNA designs targeting each selected gene were picked up from the GeCKO v2 (half-libraries A and B) and the Brunello library (Doench *et al.*, 2016). The validation library targets 307 genes (6–8 sgRNAs/gene) and contains 500 nontargeting sgRNAs, with a total of 2784 sgRNAs. Pooled custom oligonucleotides (79 bp) comprised of the 20 bp sgRNA sequence and the appropriate upstream (5'-cttGTGGAAAGACGAAACACCG-3') and downstream (5'-gttttagagctaGAAAtagcaagtaaataagg ct-3') flanking sequences were synthesized by CustomArray pooled oligo synthesis service (CustomArray, Inc., Bothell, Washington). The obtained full-size oligos were PCR amplified, gel-purified and cloned into the LentiCRISPR v2 vector (Addgene No. 52961) using Gibson assembly as previously described (Shalem *et al.*, 2014). Both GeCKO v2 and validation libraries were transformed into Endura electrocompetent cells (Lucigen 60242) using previously described protocols for library amplification (Sanjana *et al.*, 2014; Shalem *et al.*, 2014). Transformation efficiency for each library ensured sufficient representation of all constructs (approximately 150-fold library size for GeCKO v2 and approximately 500-fold library size for validation library). Plasmid DNA was isolated from amplified colonies using the Maxiprep plasmid DNA purification kit (Qiagen 12362).

Lentiviral production and functional titration. Lentivirus production was performed as previously described (Shalem *et al.*, 2014), with minor modifications. Briefly, HEK293T cells cultured in a T225 flasks were cotransfected with 20 µg of the plasmid library, 15 µg of the packaging plasmid (psPAX2, Addgene No. 12260), and 10 µg of the envelope plasmid (pMD2.G, Addgene No. 12259). Media containing the virus were collected 60 h post-transfection and filtered through a Steriflip-HV 0.45 µm low protein binding PVDF membrane (Millipore SE1M003M00). The lentiviral supernatant was concentrated 50-fold using Lenti-X Concentrator (Takara 631231) following the manufacturer's protocol. Viral solutions were aliquoted and stored at –80°C until use. To perform

functional titration of the prepared viral solutions, K562 cells were suspended in transduction medium (RPMI 1640, 10% FBS, 1% PS + 8 µg/ml polybrene) and seeded at a density of 1.25×10^6 cells/ml in 12-well plates (2.5×10^6 cells per well). Different volumes (0, 2.5, 5, 10, 15, and 20 µl) of the virus were mixed with the cell suspension in each well and the plates were centrifuged at $1000 \times g$ for 2 h at 33°C. Transduced cells from each well were suspended in fresh media and recovered for 48 h. For each transduction volume, cells were seeded in 96-well plate at a density of 10^5 cells/ml (10^4 cells/well; 100 µl) with or without puromycin (2 µg/ml) and maintained for 7 days during which 25 µl of cell suspension from each well were added to 75 µl of fresh media in a new replica plate every 48 h. Following puromycin selection, cell viability in each condition was evaluated by CellTiter Glo and the multiplicity of infection (MOI) corresponding to each transduction volume was calculated by dividing the average luminescence signal from wells with puromycin by the average luminescence signal from wells without puromycin. A transduction volume corresponding to an MOI of 0.25–0.5 was used in the large-scale transduction.

Genome-wide (primary) screening. GeCKO v2 library, packaged in lentiviral particles was transduced into 100×10^6 K562 cells in 12-well plates using the same protocol described for viral titration. Each well, containing 2.5×10^6 cells, was transduced with 10 µl of the GeCKO v2 virus which results in an MOI of approximately 0.4 (determined from titration). Transduced cells from all wells were pooled and the noninfected cells were eliminated by puromycin selection (2 µg/ml) for 7 days during which the infected cells were expanded. The pooled cell library was split into treatment and control groups (in T225 flasks) with 2 replicates for each group. In the treatment group, cells were pulsed with 2.5 mM acetaldehyde for 24 h followed by 48 h recovery, with a total of 3 pulses. Control cells were left untreated and the media were changed following the same pattern used for the treatment group. At least 25×10^6 cells were maintained in each replicate resulting in a representation of approximately 400-fold the library size. Using this strategy, selection was applied for 8 days which corresponds to 8 K562 doublings. At the end of the screen, 25×10^6 cells from each replicate were washed with PBS and the pellets were saved for DNA extraction.

Secondary screening using the validation library. 10×10^6 K562 cells were transduced with the validation library in a 12-well plate using the described lentiviral transduction protocol. Each well, containing 2.5×10^6 cells, was transduced with 2.5 µl of the validation library virus which results in an MOI of approximately 0.3 (determined from titration). Cells were pooled and treated with puromycin for 7 days to remove noninfected cells. Unlike the pulse/recovery treatment protocol applied for the primary screen, selection was performed under continuous exposure to 1 mM acetaldehyde for 8 days whereas control cells were left untreated. For each condition, the medium was changed every 48 h and at least 2.5×10^6 cells were maintained resulting in a representation of approximately 900-fold the library size. Screens were performed in T25 cell culture flasks and each condition was run in triplicate. At the end of the screen, 2.5×10^6 cells from each replicate were washed with PBS, and the pellets were saved for DNA extraction.

DNA extraction, library preparation and next-generation sequencing. Genomic DNA was isolated from 25×10^6 cells (primary screen) using the Blood and Cell Culture DNA Midi kit (Qiagen 13343)

or 2.5×10^6 cells (secondary screen) using the DNeasy Blood and Tissue kit (Qiagen 69504) following the manufacturer's protocols. Library preparation for next-generation sequencing was performed as previously described (Sanjana et al., 2014) with minor modifications. For each sample, the pool of guide sequences was amplified from genomic DNA by high fidelity PCR using the Herculase II Fusion DNA Polymerase kit (Agilent 600679). For the genome-wide screen, 150 µg genomic DNA were amplified for each sample (10 µg genomic DNA/reaction; 15 reactions/sample). For the secondary screen, 15 µg genomic DNA were amplified for each sample (5 µg genomic DNA/reaction; 3 reactions/sample). In addition to the appropriate amount of genomic DNA template, each PCR reaction contained 20 µl of the $5 \times$ reaction buffer, 500 nM of each of the forward and reverse primers, 1 mM dNTPs, 1 µl polymerase and an appropriate volume of water to reach a final volume of 100 µl. PCR was conducted with the following conditions: 95°C/2 min; 18 cycles of 95°C/20 s, 60°C/20 s, 72°C/30 s; followed by 72°C/3 min. Following amplification, reactions corresponding to the same sample were pooled. To prepare the samples for next-generation sequencing, the obtained amplicons were further amplified using primers that include appropriate P5 and P7 Illumina adapter sequences. In order to increase the diversity of the libraries, the forward primer used in the second PCR included a 5N shuffle sequence. To allow multiplexing of samples, multiple reverse primers were used in the second PCR and each primer contained a unique 8 bp index that was used to label each sample. For each PCR2 reaction, 5 µl of the first PCR product was used as a template. Seven PCR2 reactions per sample were performed for the genome-wide screen whereas a single PCR2 reaction was performed for each sample of the validation screen. PCR2 conditions and amplification protocol were similar to those used for PCR1 but 20 amplification cycles were applied instead of 18. For the genome-wide screen, PCR2 reactions corresponding to the same sample were pooled. Primers used in the first and second PCRs are shown in [Supplementary Table 1](#). The quality of the 358 bp PCR2 amplicon was assessed on a 2% agarose gel and then using a 2100 bioanalyzer (Agilent). If necessary, unincorporated primers and nonspecific products were removed from each sample using pippin prep (Sage Science). Following purification, individual samples labeled with different indices were quantified on a Qubit fluorometer (ThermoFisher Scientific) and pooled in equimolar amounts. Pooled libraries were deep sequenced using the Illumina HiSeq2500 platform (single read 50 bp) with a coverage >500-fold the size of each library.

Data processing and computational analysis. Raw FASTQ files were demultiplexed using the FASTX-Toolkit (http://hannonlab.cshl.edu/fastx_toolkit/ last accessed February, 27 2019) and processed to contain only the unique 20-bp guide sequences. To align the processed reads with the reference library, guide sequences from the library were assembled into a Burrows-Wheeler index (Li and Durbin, 2009) using the Bowtie build-index function (Langmead et al., 2009). Reads were aligned using the Bowtie aligner and the number of uniquely aligned reads (perfect match) for each guide was calculated. Individual guide counts were used as an input into edgeR, where the counts were normalized using the upper-quartile method. Differential abundance of each guide between acetaldehyde-treated and control pools was determined using the negative binomial generalized linear model approach implemented in edgeR (Lun et al., 2016). False discovery rates (FDRs) were estimated to correct for multiple comparisons. Primary candidate selection was

based on individual guide sequences displaying differential representation between acetaldehyde treated and control pools with FDR <0.1. Validation of primary candidate genes in the secondary screen involved multiple guide sequences targeting a gene showing differential representation (FDR <0.1) between treatment and control conditions.

Generation of OVCA2 knockout and control pools. sgRNAs targeting OVCA2 or a nontargeting control (NTC) sgRNA (sequences shown in [Supplementary Table 2](#)) were cloned into the CRISPR lentiviral backbone vector LentiCRISPR v2 using the Golden Gate method. To produce lentiviral particles, we cotransfected HEK293T cells in T25 culture flasks with 3.4 µg of the target vector, 2.6 µg of the packaging plasmid (psPAX2) and 1.7 µg of the envelope plasmid (pMD2.G) using Lipofectamine 2000 (ThermoFisher 11668027) and Plus reagent (ThermoFisher 11514015) following the manufacturer's instructions. Viral solutions were collected 60 h post-transfection, filtered through a Steriflip-HV 0.45 µm low protein binding PVDF membrane (Millipore SE1M003M00), and concentrated 50-fold using Lenti-X Concentrator (Takara 631231) following the manufacturer's protocol. Cells were transduced with targeting or NTC vectors at an MOI <0.5 and the transduced cells were enriched by puromycin selection. The obtained KO cellular pools were used in cytotoxicity assays to evaluate their sensitivity to acetaldehyde.

Measuring N²-ethylidene-2'-deoxyguanosine (N²-ethylidene-dG) DNA adduct levels. Evaluation of N²-ethylidene-dG adduct accumulation was performed using previously described methods ([Balbo et al., 2016](#); [Chen et al., 2007](#)) with some modifications. Wild-type (WT), NTC and OVCA2 knockout K562 pools were either treated with 5 mM acetaldehyde or left untreated for 48 h. For each sample, DNA was extracted from 1×10^6 cells using the Gentra Puregene Cell kit (Qiagen 158745) following the manufacturer's protocol. DNA pellets were suspended in 100 µl water. [¹⁵N]₅N²-ethyl-2'-deoxyguanosine ([¹⁵N]₅N²-ethyl-dG) standard was prepared as previously described ([Wang et al., 2006](#)). A total of 400 µl of 10 mM Tris/5 mM MgCl₂ buffer (pH 7) containing [¹⁵N]₅N²-ethyl-dG (50 fmol) and sodium cyanoborohydride (NaBH₃CN, 30 mg) were added to DNA solutions, and the pH was adjusted to 7 with 0.12 N HCl. NaBH₃CN in the buffer reduces N²-ethylidene-dG into the more stable N²-ethyl-dG. For DNA digestion, the mixtures were initially incubated with 626 units of DNase I (Sigma D4527) overnight at room temperature. An additional 626 units of DNase I, 32.5 mU of phosphodiesterase I (Sigma P3243) and 225 units of alkaline phosphatase (Roche 11097075001) were added and the mixtures were incubated at 37°C for 70 min and then overnight at room temperature. Enzymes were removed using Centrifree ultrafiltration device (30K molecular weight, Millipore) following the manufacturer's protocol. A 10 µl aliquot was taken from each sample for 2'-deoxyguanosine (dG) quantification. Samples were desalted and purified using a solid-phase extraction cartridge [Strata-X 33 µm, 30 mg/ml (Phenomenex)]. The 70% CH₃OH fraction (1 ml) was collected, evaporated to dryness, dissolved in 250 µl H₂O and 20 µl 0.2 N NaOH (pH 13.3), and purified using a mixed mode, anion exchange reversed phase extraction cartridge [Oasis MAX, 30 mg/cartridge (Waters)]. Adducts were eluted with 500 µl of 70% CH₃OH, and the solution was evaporated to dryness. Samples were reconstituted in 10 µl of LCMS-grade H₂O and 1 µl was analyzed by liquid chromatography-electrospray ionization-tandem mass spectrometry (LC-ESI-MS/MS). A blank containing all the reagents without DNA was prepared to check for instrument baseline and possible contamination during sample

preparation. Samples containing 50 µg calf thymus DNA (Worthington Biochemical LS002105) were processed similarly and used as controls to determine interday precision and accuracy. The LC-ESI-MS/MS analysis was performed on an UltiMate 3000 Ultra-high-performance liquid chromatography (UHPLC) System (ThermoFisher) with a 250 × 0.5 mm Luna C18 100 A column (Phenomenex) and a TSQ Vantage (ThermoFisher) triple quadrupole mass spectrometer. The solvent elution program was a 10 µl/min gradient from 5 to 35% CH₃OH in 30 min, followed by a wash at 95% CH₃OH for 5 min and re-equilibration at 5% CH₃OH for 8 min. The ESI source was set in the positive ion mode as follows: Voltage, 3.0 kV; heated ion transfer tube, 300°C. The collision energy was 12 eV, and the Ar collision gas pressure was 1.3 mTorr. Adducts were quantified by MS/MS with selected reaction monitoring at m/z 296 → m/z 180 ([M+H]⁺ → [BH]⁺) for N²-ethyl-dG, and at the corresponding transition m/z 301 → m/z 185 for [¹⁵N]₅N²-ethyl-dG. A calibration curve was derived by injecting standard solutions prepared by mixing a 5 fmol of [¹⁵N]₅N²-ethyl-dG with various amounts of N²-ethyl-dG (0, 0.25, 0.5, 1, 2, 3, 4, 20, and 30 fmol). The amount of N²-ethyl-dG (in fmol) is determined from the ratio of the peak area of N²-ethyl-dG to that of [¹⁵N]₅N²-ethyl-dG and extrapolated from the calibration curve. These values are then normalized to dG amounts which were quantified for each sample. Quantitation of dG was carried out on an UltiMate 3000 UHPLC System (ThermoFisher) with a UV detector set at 254 nm. A 250 × 0.5 mm Luna C18 100 A column (Phenomenex) was used with a flow rate of 10 µl/min and a gradient from 5% to 20% CH₃OH in H₂O over the course of 12 min followed by a 10 min hold at 20% CH₃OH. The column was then washed with 95% CH₃OH for 5 min and re-equilibrated to initial conditions for 15 min.

Statistical analysis. Data were presented as the mean ± standard deviation. Statistical significance was determined by 2-tailed, 2-sample t test assuming unequal variance. *p* < .05 was considered statistically significant.

RESULTS

Genome-Wide Screening Identifies Genes Affecting Cellular Sensitivity to Acetaldehyde

We performed a genome-wide CRISPR-based loss-of-function screen in human K562 erythroleukemic cells to define key genes that modulate acetaldehyde cytotoxicity. To generate a genome-wide genetically perturbed cellular pool, we transduced K562 cells with the GeCKO v2 pooled sgRNA library cloned into the LentiCRISPR v2 backbone vector which also expresses Cas9. Prior to screening, we determined acetaldehyde cytotoxicity in K562 cells treated with multiple concentrations of acetaldehyde (2.5–50 mM) for 72 h and observed a dose-dependent reduction in cell viability ([Figure 1B](#)). However, even the lowest acetaldehyde concentration (2.5 mM), that causes 15% reduction in cell viability after 72 h, resulted in complete cell death after 7 days of continuous exposure (data not shown). To apply selective pressure in the screen while maintaining enough viable cells to cover approximately 400-fold the library size, we used a pulse-recovery approach in which the mutant cell pool was treated with 3 pulses of acetaldehyde (2.5 mM) for 24 h followed by 48 h recovery ([Figure 1C](#)), which resulted in approximately 50% reduction in cell viability by the end of the screen (data not shown). This treatment strategy provided both positive and negative selection phenotypes. Gene disruptions leading to altered acetaldehyde sensitivity were revealed by differential

Figure 1. Identification of candidate genes involved in acetaldehyde toxicity by CRISPR-Cas9 genome-wide knockout screening. **A**, Production of acetaldehyde from ethanol and its subsequent metabolism into acetate in mammalian cells. **B**, Cytotoxicity of acetaldehyde in K562 cells. Cells were treated with different concentrations of acetaldehyde for 72 h and cell viability was evaluated by CellTiter Glo assay. Bars represent cell viability as percentage of untreated control. Data are represented as mean $\pm$ standard deviation ($n = 3$). **C**, Overview of the CRISPR screening workflow. **D**, Volcano plot showing differential abundance of each sgRNA sequence between acetaldehyde treated and control pools. Fold changes (FC) were calculated by dividing the average normalized counts ($n = 2$) for each sgRNA sequence in acetaldehyde pools to that in control pools. Differentially abundant guides with log₂ Counts Per Million (CPM) > 2 and FDR < 0.1 ($-\log_{10}$ FDR > 1) are considered significant. Significantly depleted ($\log_2$ FC < 0) and enriched ($\log_2$ FC > 0) sgRNA sequences are shown in different colors.

representation of the corresponding sgRNA sequences between acetaldehyde-treated and control pools determined by next-generation sequencing as described in the Materials and Methods section. Genes were considered for further evaluation if at least 1 well-represented (average log₂ CPM > 2) sgRNA sequence targeting the gene was differentially abundant between acetaldehyde-treated and control pools at an FDR < 0.1 . Using these criteria, we identified 21 genes whose loss potentially increases sensitivity to acetaldehyde and 19 candidate genes whose disruption leads to acetaldehyde tolerance (Figure 1D; Supplementary Table 3). These candidate genes were selected for further validation.

Secondary Screening Validates the Role of Multiple Candidates in Acetaldehyde Tolerance

A focused sgRNA library (validation library; 8 sgRNAs/gene) targeting candidate genes identified from the primary screen was used in secondary screening. The validation library also included multiple sgRNAs targeting noncandidate genes in addition to nontargeting sgRNAs (total approximately 3000 sgRNAs). We focused our efforts on confirming candidate genes whose loss increases sensitivity to acetaldehyde. Thus, we performed

the secondary screen at a lower dose of acetaldehyde (1 mM) which also obviated the need for the pulse-recovery strategy used in the primary screen. K562 cells transduced with the validation library were continuously treated with 1 mM acetaldehyde for 8 days (Figure 2A). This treatment resulted in approximately 20% drop in overall cell viability by the end of the screen (data not shown). Validation of candidate genes required that at least 2 sgRNA sequences targeting a gene display significant differential abundance between acetaldehyde-treated and control pools. Indeed, our secondary screening approach validated genes whose disruption increases sensitivity to acetaldehyde. Only a small proportion (20%) of the primary sensitivity hits were validated in our stringent confirmatory approach. Multiple sgRNA sequences targeting each of these genes were significantly depleted in acetaldehyde-treated pools compared with control pools (Figure 2B; Supplementary Table 4).

Loss of DNA Repair Genes Confers Acetaldehyde Sensitivity

Among the candidate genes validated with multiple sgRNAs in the secondary screen were those involved in DNA repair. The majority of sgRNA sequences targeting the ICL repair gene *HELQ* (Helicase POLQ-like gene) were significantly depleted in acetaldehyde-treated

Figure 2. Simultaneous validation of genetic perturbations increasing acetaldehyde sensitivity by secondary screening. **A**, Secondary screening workflow using the focused confirmatory CRISPR library. **B**, Volcano plot showing differential abundance of each sgRNA sequence of the confirmatory library between acetaldehyde treated and control pools. Fold changes (FC) were calculated by dividing the average normalized counts ($n = 3$) for each sgRNA sequence in acetaldehyde pools to that in control pools. Depleted sgRNA sequences ($\log_2 FC < 0$) with $FDR < 0.1$ ($-\log_{10} FDR > 1$) are considered significant and are shown in different color. **C**, Depletion of multiple sgRNA sequences targeting each of the DNA repair genes *HELQ*, *ERCC8*, *PPP4R2* and *UVSSA* in acetaldehyde relative to control pools. **D**, Depletion of all confirmatory sgRNA sequences targeting *OVCA2* in acetaldehyde relative to control pools. Bars in (**C**) and (**D**) represent relative abundance of each sgRNA sequence relative to untreated control. Data are represented as mean ($n = 3$). FDRs corresponding to the differential abundance of each sgRNA sequence are represented as follows: ^a $FDR < 0.001$, ^b $FDR < 0.01$, ^c $FDR < 0.1$, ^d $FDR < 0.2$.

pools (Figure 2C; Supplementary Table 4). Multiple sgRNA sequences targeting each of 2 genes involved in the TC-NER, *ERCC8* (Excision Repair Cross-Complementation Group 8 gene), and *UVSSA* (UV Stimulated Scaffold Protein A gene) were depleted in acetaldehyde-treated pools (Figure 2C; Supplementary Table 4). sgRNA sequences targeting *PPP4R2* (Protein Phosphatase 4 Regulatory Subunit 2 gene), which encodes another DNA repair component, were also depleted in acetaldehyde-treated pools as revealed by the secondary screen (Figure 2C; Supplementary Table 4), although *PPP4R2* was not initially identified as a candidate in the primary screen.

Inactivation of the Tumor Suppressor Gene *OVCA2* Confers Sensitivity to Acetaldehyde

The tumor suppressor gene, *OVCA2*, was identified as a candidate determinant of acetaldehyde tolerance in the initial screen and confirmed in the secondary screen. All the validation sgRNA sequences targeting *OVCA2* were significantly depleted in acetaldehyde-treated pools compared with untreated controls (Figure 2D). To further confirm the role of *OVCA2* in acetaldehyde resistance, we individually targeted the *OVCA2* gene in K562 cells

with 2 independent sgRNA/Cas9 vectors. We evaluated the pools of *OVCA2* knockouts corresponding to each sgRNA rather than individual isolated clones. Previous work has suggested that the pooled approach provides a robust evaluation of the functional importance of the gene and avoids clone specific effects (Parnas et al., 2015; Potting et al., 2018; Shalem et al., 2014; Xia et al., 2016). A nontargeting sgRNA/Cas9 vector was used to generate a control pool. Cell viability assays showed increased sensitivity to acetaldehyde in the *OVCA2* knockout pools when compared with the control pool (Figure 3A).

OVCA2 Is Potentially Involved in Acetaldehyde-Induced DNA Adduct Repair

Acetaldehyde exposure induces the formation of DNA adducts that likely underlie genotoxicity. One of the major acetaldehyde DNA adducts is *N*²-ethylidene-dG (Brooks and Zakhari, 2014). We measured the levels of *N*²-ethyl-2'-deoxyguanosine (*N*²-ethyl-dG), the stable derivative of *N*²-ethylidene-dG, in *OVCA2* knockout and control pools treated with 5 mM acetaldehyde for 48 h using LC-ESI-MS/MS. Accumulation of *N*²-ethyl-dG was observed in all cellular pools treated with acetaldehyde but *OVCA2*

Figure 3. Increased acetaldehyde-induced genotoxicity upon OVCA2 disruption. K562 cells were transduced with CRISPR/sgrNA vectors targeting OVCA2 or nontargeting control (NTC) and nontransduced cells were depleted by puromycin selection. A, Cells were treated with 2.5 mM acetaldehyde for 72 h and cell viability was evaluated by CellTiter Glo assay. Cell viability is represented as percentage of untreated control in each group. Data are represented as mean of 3 independent experiments $\pm$ standard deviation. B, Cells were treated with 5 mM acetaldehyde for 48 h and untreated cells from each group served as controls. Levels of N²-ethyl-dG in DNA extracted from each sample were determined by LC-MS/MS. N²-ethyl-dG levels were normalized to the 2'-deoxyguanosine (dG) levels in each sample. Data are represented as mean $\pm$ standard deviation (n = 3). Statistical significance was determined by Student's t-test, where *p < .05, **p < .01.

knockout pools exhibited significantly elevated accumulation of N²-ethyl-dG when compared with control pools (Figure 3B).

DISCUSSION

Acetaldehyde is a highly reactive intermediate metabolite of ethanol metabolism (Brooks and Zakhari, 2014). Acetaldehyde-related adverse health effects, including hematological complications and increased cancer risk, can result from ethanol consumption (Ratna and Mandrekar, 2017), tobacco smoking (Salaspuuro and Salaspuuro, 2004), or inadequate cellular detoxification mechanisms (Seitz and Stickel, 2010). Cellular processes previously shown to mitigate acetaldehyde toxicity include metabolism by the aldehyde dehydrogenase (ALDH) (Amanuma et al., 2015) and DNA repair (Langevin et al., 2011; Tacconi et al., 2017). However, our understanding of the cellular mechanisms influencing acetaldehyde toxicity remains incomplete and a more comprehensive identification of determinants of susceptibility to acetaldehyde could help reduce the risks associated with acetaldehyde exposure.

We employed a 2-tier screening system to identify genes modulating susceptibility to acetaldehyde in human K562 erythroleukemic cells. Alcohol toxicity in blood cell precursors can cause several hematological complications including decreased red blood cell production (Ballard, 1997). Accumulation of endogenous acetaldehyde compromises bone marrow HSCs and can lead to aplastic anemia (Garaycochea et al., 2012), and predisposes mice to leukemia (Langevin et al., 2011). Hence, investigating mechanisms influencing susceptibility to acetaldehyde genotoxicity in erythroleukemic cells is relevant for understanding predisposition to acetaldehyde-driven hematological complications. Our primary screen, a whole genome-wide approach with 3 sgRNAs per gene, identified 40 candidates whose loss affected sensitivity to a pulsed acetaldehyde exposure (Figure 1D; Supplementary Table 3). In our secondary screen, we used a custom library with 8 sgRNAs per gene and continuous lower dose of acetaldehyde exposure to selectively validate key genes whose disruption leads to increased sensitivity to acetaldehyde

(Figure 2). The primary screening used relatively less stringent criteria to maximize identification of candidate genes whereas the secondary screening provided a cost-effective and rapid alternative for validation when compared with individual targeted follow up of each candidate. Because we used a less selective acetaldehyde dose in the secondary screen, perturbations identified in the primary screen that result in acetaldehyde resistance were not generally validated but this does not preclude a role for these genes in acetaldehyde toxicity.

A comprehensive functional genomic study of acetaldehyde toxicity in budding yeast (*Saccharomyces cerevisiae*) identified multiple determinants of acetaldehyde tolerance including 2 aldehyde dehydrogenases (ALD3 and ALD6), in addition to components of the oleic acid biosynthesis and pentose phosphate pathways (Matsufuji et al., 2008). Human homologs of the genes relevant to acetaldehyde toxicity in yeast were not identified as candidates in our screen. Despite the known protective role of aldehyde dehydrogenases, especially ALDH2, against acetaldehyde toxicity in humans and mice (Amanuma et al., 2015; Garaycochea et al., 2012; Langevin et al., 2011), our screen did not reveal any role for ALDH2 in acetaldehyde tolerance in K562 cells. K562 cells express low levels of ALDH2 in addition to other aldehyde dehydrogenases that could also breakdown acetaldehyde (Moreb et al., 2012). Similar to our results, disruption of ALDH2 in the chicken bursal lymphoma cell line DT40 did not result in increased sensitivity to acetaldehyde treatment (Langevin et al., 2011). Together, these findings suggest that ALDH2 activity is dispensable for acetaldehyde tolerance in certain cell types.

Consistent with the established role of acetaldehyde in DNA damage (Brooks and Zakhari, 2014), the majority of genes confirmed by the secondary screen as determinants of acetaldehyde tolerance encode components of DNA repair (Figure 2C). Previous studies revealed a major role for the FANCD1-BRCA pathway in mitigating acetaldehyde toxicity (Garaycochea et al., 2012; Langevin et al., 2011; Tacconi et al., 2017). Mouse hematopoietic stem cells (HPCs) lacking *Fancc* (Fanconi anemia complementation group A) or *Fancd2* genes display a considerable increase in sensitivity to acetaldehyde exposure (Garaycochea

et al., 2012). Similarly, chicken lymphoid DT40 cells lacking individual components of the FA pathway are more vulnerable to acetaldehyde toxicity (Langevin et al., 2011). In addition, inactivation of FANCD2, BRCA-1, or BRCA-2 in human non-small cell lung carcinoma H1299 cells, and colorectal adenocarcinoma DLD1 cells, leads to a substantial increase in sensitivity to acetaldehyde exposure (Tacconi et al., 2017). Our primary screen revealed a single component of the FA pathway, PALB2 (Partner And Localizer of BRCA2; also known as FANCN), as a candidate determinant of acetaldehyde tolerance (Figure 1D; Supplementary Table 3). Consistent with this finding, disruption of PALB2 has been shown to increase cellular sensitivity to acetaldehyde (Ghosh et al., 2014). However, PALB2 was not validated by our secondary screening approach where a low selective acetaldehyde dose was used. One plausible explanation could be that the protective role of PALB2 against acetaldehyde toxicity in K562 cells can only be detected at higher acetaldehyde doses where other detoxification systems become saturated.

Our approach identified novel DNA repair components whose disruption enhances cellular sensitivity to acetaldehyde. These include the helicase HELQ, which is involved in DNA ICL repair (Takata et al., 2013). Mice with deleted *Helq* are more predisposed to developing tumors and exhibit a phenotype similar to that observed in mice mutant for different Fanconi Anemia related genes (Adelman et al., 2013; Parmar et al., 2009). Intriguingly, *Helq* deficient cells display increased sensitivity to ICL-inducing agents that is further exacerbated upon deletion of *fancd2* indicating that both *Helq* and *Fancd2* are involved in ICL repair (Adelman et al., 2013). *Fancd2* was previously shown to counteract the genotoxic effects of exogenous and endogenous acetaldehyde (Langevin et al., 2011). Collectively, these observations and our finding that HELQ disruption increases acetaldehyde sensitivity indicate that the role of HELQ in ICL repair is critical for acetaldehyde tolerance. Correspondingly the reaction of acetaldehyde with DNA was shown to induce ICL formation (Liu et al., 2006; Wang et al., 2000) and the FANC-BRCA pathway, a recognized determinant of acetaldehyde tolerance, is primarily involved in repairing ICLs (Kim and D'Andrea, 2012). Another revealed gene with a potential role in mitigating the toxic effects of acetaldehyde encodes PP4R2, the regulatory subunit of protein phosphatase 4 (PP4). PP4R2 is recurrently deleted in leukemia and is involved in DNA repair (Chowdhury et al., 2008; Herzig et al., 2017). Phosphorylation of the histone H2AX (which becomes γ -H2AX) is an early response to DNA ICLs that allows recruitment of multiple repair factors, including FANC-BRCA components, to the damaged loci (Clingen et al., 2008; Wang, 2007) and a subsequent dephosphorylation of γ -H2AX by PP4 is required for efficient DNA repair (Chowdhury et al., 2008). Hence, PP4R2 could contribute to the efficient repair of acetaldehyde-induced ICLs consistent with a role in acetaldehyde tolerance. Two additional validated genes, ERCC8 and UVSSA, encode components of the TC-NER pathway (Fousteri and Mullenders, 2008; Nakazawa et al., 2012). A direct role for TC-NER in the removal of acetaldehyde-derived DNA adducts has not been reported. However, TC-NER has been recognized as an adduct-repairing pathway due to its role in detoxifying adduct-forming agents like cisplatin (Furuta et al., 2002). In addition, polymorphisms in ERCC8 are associated with susceptibility to gastric cancer related to smoking and alcohol consumption, two major sources of exogenous acetaldehyde exposure (Jing et al., 2015). Collectively, these findings suggest that multiple DNA repair pathways can contribute to mitigating acetaldehyde genotoxicity. In line with this suggestion, a comprehensive

study of DNA repair genes in fission yeast (*Schizosaccharomyces pombe*) revealed roles for components belonging to the FA, HR, NER, BER, TLS, and DPC repair pathways in acetaldehyde tolerance (Noguchi et al., 2017).

We revealed a novel acetaldehyde tolerance gene, OVCA2, whose inactivation results in increased sensitivity to acetaldehyde (Figure 2D). Unlike the other acetaldehyde tolerance genes confirmed in our secondary screen, OVCA2 was not previously shown to be involved in DNA repair and has unknown biological function. Intriguingly, human OVCA2 is located in a small region of chromosome 17 that is deleted in the vast majority of esophageal squamous cell carcinoma (ESCC) (Huang et al., 2000). Importantly, ESCC incidence is strongly correlated to alcohol consumption and tobacco smoking, two of the most important acetaldehyde sources (Freedman et al., 2007; Pandeya et al., 2013; Yang et al., 2017). More recently, a genome-wide association study identified multiple ESCC susceptibility loci, including a locus on chromosome 17 that contains OVCA2 (Wu et al., 2012). The same study verified the well-known association of the ALDH2 locus to increased ESCC risk thus highlighting the role of acetaldehyde in this disease. Together, these findings suggest that OVCA2 mitigates the genotoxic effects of acetaldehyde and its loss increases susceptibility to ESCC associated with alcohol exposure. Hence, defining the role of OVCA2 as a determinant of acetaldehyde tolerance is relevant for understanding susceptibility to ESCC and other alcohol-related cancers. Our results demonstrated that disruption of OVCA2 confers a cellular growth disadvantage in the presence of acetaldehyde that is associated with accumulation of the N²-ethylidene-dG DNA adduct (Figure 3). In agreement with our finding, deletion of FSH1, the yeast homolog of OVCA2, results in defective growth in a medium containing ethanol, the precursor of acetaldehyde (Schlecht et al., 2014). These findings suggest that OVCA2 confers acetaldehyde tolerance through its role in removal of DNA adducts or acetaldehyde metabolism.

The acetaldehyde doses used in our study are relatively high and exceed reported physiological or pathological acetaldehyde concentrations. High acetaldehyde concentrations were used in multiple studies *in vitro* to evaluate acetaldehyde cytotoxicity in different cell lines and primary cells (Amanuma et al., 2015; Garaycochea et al., 2012; Langevin et al., 2011; Yan et al., 2016). Lower acetaldehyde concentrations can result in adduct formation (Amanuma et al., 2015; Moeller et al., 2013) but do not necessarily lead to a drop in cell viability (Amanuma et al., 2015). Hence, we utilized higher doses which affect cell viability to investigate the genetic modulators of acetaldehyde toxicity. In a previous study, the increased sensitivity of *Fancd2* deficient mouse B cells to acetaldehyde was demonstrated by short-term exposure to high exogenous acetaldehyde concentrations (1–4 mM) (Langevin et al., 2011). Intriguingly, *Fancd2* deletion in *Aldh2* deficient mice led to acute leukemia in the absence of any external acetaldehyde source. These observations provide a strong evidence that the cytotoxicity resulting from acute exposure to high nonphysiological acetaldehyde concentrations *in vitro* can predict long-term genotoxic effects of chronic real-world acetaldehyde exposure. Hence, the protective role of OVCA2 against short-term exposure to high acetaldehyde doses supports a role for OVCA2 in protecting against the genotoxic effects of chronic exposure to physiological and pathological acetaldehyde concentrations.

Functional domain analysis of OVCA2 predicted a serine hydrolase activity which was supported by activity based serine hydrolase probes (Baxter et al., 2004). OVCA2 expression is downregulated in response to retinoid treatment (Prowse et al.,

2002), but its biological function remains unclear. OVCA1 and OVCA2 are ubiquitously expressed tumor suppressor candidates whose expression is considerably reduced in ovarian cancer as a result of allelic loss in the corresponding locus (Schultz et al., 1996). Unlike OVCA1, ectopic expression of OVCA2 in ovarian cancer cells does not suppress cell proliferation (Bruening et al., 1999). Accordingly, the tumor suppressor role of OVCA1 (DPH1), but not OVCA2, was highlighted in multiple studies (Chen and Behringer, 2004; Kong et al., 2011; Tong et al., 2017). OVCA1 and OVCA2 genes partially overlap but do not share coding regions and OVCA1 was not identified in our screen. OVCA2 has 2 exons, the first is located within the last intron of OVCA1 and the second overlaps with the region encoding the 3' UTR of OVCA1 (Chen and Behringer, 2004). Our confirmatory screen showed that all the sgRNAs targeting OVCA2 exons (5 sgRNAs targeting exon 1 and 3 sgRNAs targeting exon 2) resulted in increased sensitivity to acetaldehyde (Figure 2D). Because these sgRNAs target different noncoding and untranslated regions of OVCA1, we are confident that the observed acetaldehyde sensitivity results from OVCA2 and not OVCA1 disruption.

Our work revealed a role for OVCA2 in mitigating acetaldehyde toxicity and suggests the removal of acetaldehyde-derived DNA adducts as a function of OVCA2 underlying acetaldehyde tolerance. However, a role for OVCA2 in acetaldehyde metabolism is possible, because inactivation of the acetaldehyde-metabolizing enzyme ALDH2 has also been shown to increase acetaldehyde-derived N²-ethylidene-dG levels (Amanuma et al., 2015). Hence, additional work is required to elucidate whether OVCA2-mediated acetaldehyde tolerance arises from a role for OVCA2 in acetaldehyde metabolism, adduct removal or DNA repair.

SUPPLEMENTARY DATA

Supplementary data are available at *Toxicological Sciences*.

FUNDING

This work was supported by grant from the Superfund Hazardous Substance Research and Training Program (P42ES004705 to M.T.S.) as well as support from the University of Florida, Gainesville to C.V. A.S. was a trainee in the Superfund Research Program (University of California, Berkeley). The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institute of Environmental Health Sciences.

DECLARATION OF CONFLICTING INTERESTS

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

REFERENCES

- Adelman, C. A., Lolo, R. L., Birkbak, N. J., Murina, O., Matsuzaki, K., Horejsi, Z., Parmar, K., Borel, V., Skehel, J. M., and Stamp, G. (2013). HELQ promotes RAD51 paralogue-dependent repair to avert germ cell loss and tumorigenesis. *Nature* **502**, 381–384.
- Amanuma, Y., Ohashi, S., Itatani, Y., Tsurumaki, M., Matsuda, S., Kikuchi, O., Nakai, Y., Miyamoto, S., Oyama, T., Kawamoto, T., et al. (2015). Protective role of ALDH2 against acetaldehyde-derived DNA damage in oesophageal squamous epithelium. *Sci. Rep.* **5**, 14142.
- Aranda, A., and del Olmo, M. (2004). Exposure of *Saccharomyces cerevisiae* to acetaldehyde induces sulfur amino acid metabolism and polyamine transporter genes, which depend on Met4p and Haa1p transcription factors, respectively. *Appl. Environ. Microbiol.* **70**, 1913–1922.
- Balbo, S., Juanes, R. C., Khariwala, S., Baker, E. J., Daunais, J. B., and Grant, K. A. (2016). Increased levels of the acetaldehyde-derived DNA adduct N 2-ethyldeoxyguanosine in oral mucosa DNA from Rhesus monkeys exposed to alcohol. *Mutagenesis* **31**, 553–558.
- Ballard, H. S. (1997). The hematological complications of alcoholism. *Alcohol Health Res. World* **21**, 42–52.
- Baxter, S. M., Rosenblum, J. S., Knutson, S., Nelson, M. R., Montimurro, J. S., Gennaro, J. A. D., Speir, J. A., Burbaum, J. J., and Fetrow, J. S. (2004). Synergistic computational and experimental proteomics approaches for more accurate detection of active serine hydrolases in yeast. *Mol. Cell Proteomics* **3**, 209–225.
- Brooks, P. J., Enoch, M.-A., Goldman, D., Li, T.-K., and Yokoyama, A. (2009). The alcohol flushing response: An unrecognized risk factor for esophageal cancer from alcohol consumption. *PLoS Med.* **6**, e1000050.
- Brooks, P. J., and Schuebel, K. (2017). Timeless insights into prevention of acetaldehyde genotoxicity? *Cell Cycle* **16**, 308–309.
- Brooks, P. J., and Theruvathu, J. A. (2005). DNA adducts from acetaldehyde: Implications for alcohol-related carcinogenesis. *Alcohol* **35**, 187–193.
- Brooks, P. J., and Zakhari, S. (2014). Acetaldehyde and the genome: Beyond nuclear DNA adducts and carcinogenesis. *Environ. Mol. Mutagen* **55**, 77–91.
- Bruening, W., Prowse, A. H., Schultz, D. C., Holgado-Madruga, M., Wong, A., and Godwin, A. K. (1999). Expression of OVCA1, a candidate tumor suppressor, is reduced in tumors and inhibits growth of ovarian cancer cells. *Cancer Res.* **59**, 4973–4983.
- Cheah, N. P., Pennings, J. L. A., Vermeulen, J. P., van Schooten, F. J., and Opperhuizen, A. (2013). In vitro effects of aldehydes present in tobacco smoke on gene expression in human lung alveolar epithelial cells. *Toxicol. In Vitro* **27**, 1072–1081.
- Chen, C.-M., and Behringer, R. R. (2004). Ovca1 regulates cell proliferation, embryonic development, and tumorigenesis. *Genes Dev.* **18**, 320–332.
- Chen, L., Wang, M., Villalta, P. W., Luo, X., Feuer, R., Jensen, J., Hatsukami, D. K., and Hecht, S. S. (2007). Quantitation of an acetaldehyde adduct in human leukocyte DNA and the effect of smoking cessation. *Chem. Res. Toxicol.* **20**, 108–113.
- Chen, W. Y., Rosner, B., Hankinson, S. E., Colditz, G. A., and Willett, W. C. (2011). Moderate alcohol consumption during adult life, drinking patterns, and breast cancer risk. *JAMA* **306**, 1884–1890.
- Chowdhury, D., Xu, X., Zhong, X., Ahmed, F., Zhong, J., Liao, J., Dykxhoorn, D. M., Weinstock, D. M., Pfeifer, G. P., and Lieberman, J. (2008). A PP4-phosphatase complex dephosphorylates γ -H2AX generated during DNA replication. *Mol. Cell* **31**, 33–46.
- Clingen, P. H., Wu, J. Y.-H., Miller, J., Mistry, N., Chin, F., Wynne, P., Prise, K. M., and Hartley, J. A. (2008). Histone H2AX phosphorylation as a molecular pharmacological marker for DNA interstrand crosslink cancer chemotherapy. *Biochem. Pharmacol.* **76**, 19–27.
- Doench, J. G., Fusi, N., Sullender, M., Hegde, M., Vaimberg, E. W., Donovan, K. F., Smith, I., Tothova, Z., Wilen, C., Orchard, R.,

- et al. (2016). Optimized sgRNA design to maximize activity and minimize off-target effects of CRISPR-Cas9. *Nat. Biotechnol.* **34**, 184–191.
- Fousteri, M., and Mullenders, L. H. (2008). Transcription-coupled nucleotide excision repair in mammalian cells: Molecular mechanisms and biological effects. *Cell Res.* **18**, 73–84.
- Freedman, N. D., Abnet, C. C., Leitzmann, M. F., Mouw, T., Subar, A. F., Hollenbeck, A. R., and Schatzkin, A. (2007). A prospective study of tobacco, alcohol, and the risk of esophageal and gastric cancer subtypes. *Am. J. Epidemiol.* **165**, 1424–1433.
- Furuta, T., Ueda, T., Aune, G., Sarasin, A., Kraemer, K. H., and Pommier, Y. (2002). Transcription-coupled nucleotide excision repair as a determinant of cisplatin sensitivity of human cells. *Cancer Res.* **62**, 4899–4902.
- Garaycoechea, J. I., Crossan, G. P., Langevin, F., Daly, M., Arends, M. J., and Patel, K. J. (2012). Genotoxic consequences of endogenous aldehydes on mouse haematopoietic stem cell function. *Nature* **489**, 571.
- Garaycoechea, J. I., Crossan, G. P., Langevin, F., Muldergig, L., Louzada, S., Yang, F., Guilbaud, G., Park, N., Roerink, S., Nik-Zainal, S., et al. (2018). Alcohol and endogenous aldehydes damage chromosomes and mutate stem cells. *Nature* **553**, 171.
- Ghosh, S., Sur, S., Yerram, S. R., Rago, C., Bhunia, A. K., Hossain, M. Z., Paun, B. C., Ren, Y. R., Iacobuzio-Donahue, C. A., Azad, N. A., et al. (2014). Hypersensitivities for acetaldehyde and other agents among cancer cells null for clinically relevant Fanconi anemia genes. *Am. J. Pathol.* **184**, 260–270.
- Herzig, J. K., Bullinger, L., Tasdogan, A., Zimmermann, P., Schlegel, M., Teleanu, V., Weber, D., Rücker, F. G., Paschka, P., Dolnik, A., et al. (2017). Protein phosphatase 4 regulatory subunit 2 (PPP4R2) is recurrently deleted in acute myeloid leukemia and required for efficient DNA double strand break repair. *Oncotarget* **8**, 95038–95053.
- Hira, A., Yabe, H., Yoshida, K., Okuno, Y., Shiraishi, Y., Chiba, K., Tanaka, H., Miyano, S., Nakamura, J., Kojima, S., et al. (2013). Variant ALDH2 is associated with accelerated progression of bone marrow failure in Japanese Fanconi anemia patients. *Blood* **122**, 3206–3209.
- Huang, J., Hu, N., Goldstein, A. M., Emmert-Buck, M. R., Tang, Z. Z., Roth, M. J., Wang, Q. H., Dawsey, S. M., Han, X. Y., Ding, T., et al. (2000). High frequency allelic loss on chromosome 17p13.3-p11.1 in esophageal squamous cell carcinomas from a high incidence area in northern China. *Carcinogenesis* **21**, 2019–2026.
- Jing, J., Sun, L., Xu, Q., and Yuan, Y. (2015). Effect of ERCC8 tagSNPs and their association with *H. pylori* infection, smoking, and alcohol consumption on gastric cancer and atrophic gastritis risk. *Tumor Biol.* **36**, 9525–9535.
- Johnson, K. A., Fink, S. P., and Marnett, L. J. (1997). Repair of propenodeoxyguanosine by nucleotide excision repair in vivo and in vitro. *J. Biol. Chem.* **272**, 11434–11438.
- Kim, H., and D'Andrea, A. D. (2012). Regulation of DNA cross-link repair by the Fanconi anemia/BRCA pathway. *Genes Dev.* **26**, 1393–1408.
- Kong, F., Tong, R., Jia, L., Wei, W., Miao, X., Zhao, X., Sun, W., Yang, G., and Zhao, C. (2011). OVCA1 inhibits the proliferation of epithelial ovarian cancer cells by decreasing cyclin D1 and increasing p16. *Mol. Cell Biochem.* **354**, 199–205.
- Langevin, F., Crossan, G. P., Rosado, I. V., Arends, M. J., and Patel, K. J. (2011). Fancd2 counteracts the toxic effects of naturally produced aldehydes in mice. *Nature* **475**, 53.
- Langmead, B., Trapnell, C., Pop, M., and Salzberg, S. L. (2009). Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol.* **10**, R25.
- Li, H., and Durbin, R. (2009). Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* **25**, 1754–1760.
- Liu, X., Lao, Y., Yang, I.-Y., Hecht, S. S., and Moriya, M. (2006). Replication-coupled repair of crotonaldehyde/acetaldehyde-induced guanine-guanine interstrand cross-links and their mutagenicity. *Biochemistry* **45**, 12898–12905.
- Lun, A. T. L., Chen, Y., and Smyth, G. K. (2016). It's DE-licious: A recipe for differential expression analyses of RNA-seq experiments using Quasi-likelihood methods in edgeR. *Methods Mol. Biol.* **1418**, 391–416.
- Matsuda, T., Yabushita, H., Kanaly, R. A., Shibutani, S., and Yokoyama, A. (2006). Increased DNA damage in ALDH2-deficient alcoholics. *Chem. Res. Toxicol.* **19**, 1374–1378.
- Matsufuji, Y., Fujimura, S., Ito, T., Nishizawa, M., Miyaji, T., Nakagawa, J., Ohyama, T., Tomizuka, N., and Nakagawa, T. (2008). Acetaldehyde tolerance in *Saccharomyces cerevisiae* involves the pentose phosphate pathway and oleic acid biosynthesis. *Yeast* **25**, 825–833.
- McKillop, I. H., and Schrum, L. W. (2009). Role of alcohol in liver carcinogenesis. *Semin. Liver Dis.* **29**, 222–232.
- Moeller, B. C., Recio, L., Green, A., Sun, W., Wright, F. A., Bodnar, W. M., and Swenberg, J. A. (2013). Biomarkers of exposure and effect in human lymphoblastoid TK6 cells following [13C2]-acetaldehyde exposure. *Toxicol. Sci.* **133**, 1–12.
- Moreb, J. S., Ucar, D., Han, S., Amory, J. K., Goldstein, A. S., Ostmark, B., and Chang, L.-J. (2012). The enzymatic activity of human aldehyde dehydrogenases 1A2 and 2 (ALDH1A2 and ALDH2) is detected by Aldefluor, inhibited by diethylaminobenzaldehyde and has significant effects on cell proliferation and drug resistance. *Chem. Biol. Interact.* **195**, 52–60.
- Nakazawa, Y., Sasaki, K., Mitsutake, N., Matsuse, M., Shimada, M., Nardo, T., Takahashi, Y., Ohyama, K., Ito, K., Mishima, H., et al. (2012). Mutations in UVSSA cause UV-sensitive syndrome and impair RNA polymerase II processing in transcription-coupled nucleotide-excision repair. *Nat. Genet.* **44**, 586–592.
- Niemelä, O. (1999). Aldehyde-protein adducts in the liver as a result of ethanol-induced oxidative stress. *Front. Biosci. J. Virtual Libr.* **4**, D506–D513.
- Noguchi, C., Grothausen, G., Anandarajan, V., Martínez-Lage García, M., Terlecky, D., Corzo, K., Tanaka, K., Nakagawa, H., and Noguchi, E. (2017). Genetic controls of DNA damage avoidance in response to acetaldehyde in fission yeast. *Cell Cycle* **16**, 45–58.
- Pandeya, N., Olsen, C. M., and Whiteman, D. C. (2013). Sex differences in the proportion of esophageal squamous cell carcinoma cases attributable to tobacco smoking and alcohol consumption. *Cancer Epidemiol.* **37**, 579–584.
- Parmar, K., D'Andrea, A., and Niedernhofer, L. J. (2009). Mouse models of Fanconi anemia. *Mutat. Res.* **668**, 133–140.
- Parmar, K., and D'Andrea, A. D. (2012). Stressed out: Endogenous aldehydes damage hematopoietic Stem Cells. *Cell Stem Cell* **11**, 583–584.
- Parnas, O., Jovanovic, M., Eisenhaure, T. M., Herbst, R. H., Dixit, A., Ye, C. J., Przybylski, D., Platt, R. J., Tirosh, I., Sanjana, N. E., et al. (2015). A genome-wide CRISPR screen in primary immune cells to dissect regulatory networks. *Cell* **162**, 675–686.
- Potting, C., Crochemore, C., Moretti, F., Nigsch, F., Schmidt, I., Manneville, C., Carbone, W., Knehr, J., DeJesus, R., Lindeman, A., et al. (2018). Genome-wide CRISPR screen for PARKIN regulators reveals transcriptional repression as a determinant of mitophagy. *Proc. Natl. Acad. Sci.* **115**, E180–E189.

- Prowse, A. H., Vanderveer, L., Milling, S. W. F., Pan, Z.-Z., Dunbrack, R. L., Xu, X.-X., and Godwin, A. K. (2002). OVCA2 is downregulated and degraded during retinoid-induced apoptosis. *Int. J. Cancer* **99**, 185–192.
- Ratna, A., and Mandrekar, P. (2017). Alcohol and cancer: Mechanisms and therapies. *Biomolecules* **7**, 61.
- Salaspuro, V., and Salaspuro, M. (2004). Synergistic effect of alcohol drinking and smoking on in vivo acetaldehyde concentration in saliva. *Int. J. Cancer* **111**, 480–483.
- Sanjana, N. E., Shalem, O., and Zhang, F. (2014). Improved vectors and genome-wide libraries for CRISPR screening. *Nat. Methods* **11**, 783–784.
- Schlecht, U., Suresh, S., Xu, W., Aparicio, A. M., Chu, A., Proctor, M. J., Davis, R. W., Scharfe, C., and St Onge, R. P. (2014). A functional screen for copper homeostasis genes identifies a pharmacologically tractable cellular system. *BMC Genomics* **15**, 263.
- Schultz, D. C., Vanderveer, L., Berman, D. B., Hamilton, T. C., Wong, A. J., and Godwin, A. K. (1996). Identification of two candidate tumor suppressor genes on chromosome 17p13.3. *Cancer Res.* **56**, 1997–2002.
- Secretan, B., Straif, K., Baan, R., Grosse, Y., El Ghissassi, F., Bouvard, V., Benbrahim-Tallaa, L., Guha, N., Freeman, C., Galichet, L., et al. (2009). A review of human carcinogens—Part E: Tobacco, areca nut, alcohol, coal smoke, and salted fish. *Lancet Oncol.* **10**, 1033–1034.
- Seitz, H. K., and Stickel, F. (2007). Molecular mechanisms of alcohol-mediated carcinogenesis. *Nat. Rev. Cancer* **7**, 599–612.
- Seitz, H. K., and Stickel, F. (2010). Acetaldehyde as an underestimated risk factor for cancer development: Role of genetics in ethanol metabolism. *Genes Nutr.* **5**, 121–128.
- Setshedi, M., Wands, J. R., and de la Monte, S. M. (2010). Acetaldehyde adducts in alcoholic liver disease. *Oxid. Med. Cell Longev.* **3**, 178–185.
- Shalem, O., Sanjana, N. E., Hartenian, E., Shi, X., Scott, D. A., Mikkelsen, T. S., Heckl, D., Ebert, B. L., Root, D. E., Doench, J. G., et al. (2014). Genome-scale CRISPR-Cas9 knockout screening in human cells. *Science* **343**, 84–87.
- Tacconi, E. M., Lai, X., Folio, C., Porru, M., Zonderland, G., Badie, S., Michl, J., Sechi, I., Rogier, M., Matía García, V., et al. (2017). BRCA1 and BRCA2 tumor suppressors protect against endogenous acetaldehyde toxicity. *EMBO Mol. Med.* **9**, 1398–1414.
- Takata, K., Reh, S., Tomida, J., Person, M. D., and Wood, R. D. (2013). Human DNA helicase HELQ participates in DNA interstrand crosslink tolerance with ATR and RAD51 paralogs. *Nat. Commun.* **4**, 2338.
- Tong, R., Yang, Q., Wang, C., Bi, F., and Jiang, B. (2017). OVCA1 expression and its correlation with the expression levels of cyclin D1 and p16 in cervical cancer and intraepithelial neoplasia. *Oncol. Lett.* **13**, 2929–2936.
- Venton, G., Pérez-Alea, M., Baier, C., Fournet, G., Quash, G., Labiad, Y., Martin, G., Sanderson, F., Poullin, P., Suchon, P., et al. (2016). Aldehyde dehydrogenases inhibition eradicates leukemia stem cells while sparing normal progenitors. *Blood Cancer J.* **6**, e469.
- Wang, M., McIntee, E. J., Cheng, G., Shi, Y., Villalta, P. W., and Hecht, S. S. (2000). Identification of DNA adducts of acetaldehyde. *Chem. Res. Toxicol.* **13**, 1149–1157.
- Wang, M., Yu, N., Chen, L., Villalta, P. W., Hochalter, J. B., and Hecht, S. S. (2006). Identification of an acetaldehyde adduct in human liver DNA and quantitation as N2-ethyldeoxyguanosine. *Chem. Res. Toxicol.* **19**, 319–324.
- Wang, W. (2007). Emergence of a DNA-damage response network consisting of Fanconi anaemia and BRCA proteins. *Nat. Rev. Genet.* **8**, 735–748.
- Wu, C., Kraft, P., Zhai, K., Chang, J., Wang, Z., Li, Y., Hu, Z., He, Z., Jia, W., Abnet, C. C., et al. (2012). Genome-wide association analyses of esophageal squamous cell carcinoma in Chinese identify multiple susceptibility loci and gene-environment interactions. *Nat. Genet.* **44**, 1090–1097.
- Xia, P., Zhang, X., Xie, Y., Guan, M., Villeneuve, D. L., and Yu, H. (2016). Functional toxicogenomic assessment of triclosan in human HepG2 cells using genome-wide CRISPR-Cas9 screening. *Environ. Sci. Technol.* **50**, 10682–10692.
- Yan, T., Zhao, Y., and Zhang, X. (2016). Acetaldehyde induces cytotoxicity of SH-SY5Y cells via inhibition of Akt activation and induction of oxidative stress. *Oxid. Med. Cell Longev.* **2016**, 4512309.
- Yang, X., Chen, X., Zhuang, M., Yuan, Z., Nie, S., Lu, M., Jin, L., and Ye, W. (2017). Smoking and alcohol drinking in relation to the risk of esophageal squamous cell carcinoma: A population-based case-control study in China. *Sci. Rep.* **7**, 17249.
- Yokoyama, A., and Omori, T. (2003). Genetic polymorphisms of alcohol and aldehyde dehydrogenases and risk for esophageal and head and neck cancers. *Jpn. J. Clin. Oncol.* **33**, 111–121.