HHS Public Access

Author manuscript

Mol Cancer Ther. Author manuscript; available in PMC 2019 July 01.

Published in final edited form as:

Mol Cancer Ther. 2018 July ; 17(7): 1585–1594. doi:10.1158/1535-7163.MCT-17-0937.

Disruption of *NSD1* in head and neck cancer promotes favorable chemotherapeutic responses linked to hypomethylation

Nam Bui[1,*], **Justin K. Huang**[2,*], **Ana Bojorquez-Gomez**[1], **Katherine Licon**[1], **Kyle S. Sanchez**[1], **Sean N. Tang**[1], **Alex N. Beckett**[1], **Tina Wang**[1], **Wei Zhang**[1], **John Paul Shen**[1,†], **Jason F. Kreisberg**[1,†], and **Trey Ideker**[1,2,3,†]

[1]Department of Medicine, UC San Diego, La Jolla, California, USA

[2]Bioinformatics and Systems Biology Program, UC San Diego, La Jolla, California, USA

[3]Department of Bioengineering, UC San Diego, La Jolla, California, USA

Abstract

Human papillomavirus (HPV) negative head and neck squamous cell carcinoma (HNSCC) represents a distinct classification of cancer with poor expected outcomes. Of the 11 genes recurrently mutated in HNSCC, we identify a singular and substantial survival advantage for mutations in the gene encoding Nuclear Set Domain Containing Protein 1 (*NSD1*), a histone methyltransferase altered in approximately 10% of patients. This effect, a 55% decrease in risk of death in *NSD1*-mutated versus non-mutated patients, can be validated in an independent cohort. *NSD1* alterations are strongly associated with widespread genome hypomethylation in the same tumors, to a degree not observed for any other mutated gene. To address whether *NSD1* plays a causal role in these associations, we use CRISPR-Cas9 to disrupt *NSD1* in HNSCC cell lines and find that this leads to substantial CpG hypomethylation and sensitivity to cisplatin, a standard chemotherapy in head and neck cancer, with a 40 – 50% decrease in IC_{50}. Such results are reinforced by a survey of 1,001 cancer cell lines, in which loss-of-function *NSD1* mutations have an average 23% decrease in cisplatin IC_{50} compared to cell lines with wild type *NSD1*. This study identifies a favorable subtype of head and neck cancer linked to *NSD1* mutation, hypomethylation and cisplatin sensitivity.

Keywords

HPV(−) head and neck cancer; Cancer Epigenetics; Cisplatin Sensitivity

[†]Address correspondences to J.P.S. (jpshen@ucsd.edu), J.F.K. (jkreisberg@ucsd.edu) or T.I. (tideker@ucsd.edu).
[*]Equal contributions

Contact Information

John Paul Shen, M.D., UC San Diego, Department of Medicine, 9500 Gilman Drive MC-0688, La Jolla, CA 92093-0688, (858) 822-4704

Jason F. Kreisberg, Ph.D., UC San Diego, Department of Medicine, 9500 Gilman Drive MC-0688, La Jolla, CA 92093-0688, (858) 534-3578

Trey Ideker, Ph.D., UC San Diego, Department of Medicine, 9500 Gilman Drive MC-0688, La Jolla, CA 92093-0688, (858) 822-4558

Disclosure of Potential Conflicts of Interest

Trey Ideker is co-founder of Data4Cure, Inc. and has an equity interest. Trey Ideker has an equity interest in Ideaya BioSciences, Inc. The terms of this arrangement have been reviewed and approved by the University of California, San Diego, in accordance with its conflict of interest policies. No potential conflicts of interest were disclosed by the other authors.

INTRODUCTION

Head and neck squamous cell carcinoma (HNSCC) is the sixth most common cause of cancer worldwide, with more than 500,000 cases leading to 300,000 deaths each year (1). In the last decade, it has become clear that there are two distinct classes of HNSCC based on the presence or absence of human papillomavirus (HPV). HPV(+) head and neck cancers have a more favorable prognosis than HPV(−) cases (74% versus 30% 5-year overall survival rate) (2). For this reason, HPV(+) and HPV(−) tumors are now regarded as separate diseases with distinct objectives for further research, with a focus on de-intensification of therapy in HPV(+) and novel therapeutic approaches in HPV(−) tumors (3). For both of these diseases, the current standard of care for localized HNSCC involves surgery, radiation and concomitant chemotherapy, typically by the platinum DNA-damaging agent cisplatin. Other therapeutic strategies have been attempted, including combination chemotherapy (4,5) and inhibition of epidermal growth factor (EGFR) with cetuximab (6,7). However, none of these chemotherapy options have resulted in a definitively improved prognosis in HPV(−) cases.

Recently, The Cancer Genome Atlas (TCGA) performed a comprehensive molecular analysis of all HNSCC (HPV(−) and HPV(+)) identifying recurrent mutations in 11 genes including *TP53* (72%), *FAT1* (23%), *CDKN2A* (22%), *NOTCH1* (19%) and *NSD1* (10%) (8). However, this initial study did not attempt to associate these genetic events with clinical outcomes. With this goal in mind, we sought to identify recurrently mutated genes that stratify HNSCC patients into clinically informative subgroups. In what follows, we report that somatic mutations in *NSD1*, a histone methyltransferase (HMT), are strongly correlated with cisplatin sensitivity as well as better patient outcomes, and that these effects can be recapitulated by disrupting *NSD1* in HNSCC cell lines using CRISPR-Cas9.

MATERIALS AND METHODS

Data Acquisition

TCGA data were obtained from the Genome Data Analysis Center Broad Firehose website (https://gdac.broadinstitute.org/) (9) including full clinical information, mutation calls, mRNA sequencing data and methylation CpG (beta) fractions. All data were downloaded from the run on January 28, 2016 (https://doi.org/10.7908/C11G0KM9). Research was conducted in accordance with the U.S. Common Rule. Per institutional guidelines (Common Rule: 45 CFR 46 subpart A), this study was exempt from Institutional Review Board (IRB) review due to the fact that it involved publicly available data from which subjects cannot be identified.

Determining HPV status

HPV calls for the 279 patients analyzed in the original TCGA paper were also obtained (8). For remaining patients, we examined the clinical information as well as MassArray data: patients with p16 or *in situ* hybridization results were noted as HPV(+) if either of those tests were positive. Lacking either test we turned to the MassArray calls (PCR for 16 HPV types) from TCGA to determine HPV status.

Survival analysis

Cox regression models were constructed using the mutation status of *NSD1* and ten other recurrently mutated genes along with the clinical co-variates age, stage, grade, gender, smoking status and anatomical location. Kaplan-Meier methods were used to generate survival curves. The 'survival' package from R was used for this analysis (10).

TCGA methylation analysis

We selected the 1000 most variable CpG probes from HPV(−) HNSCC samples in TCGA, excluding SNP-associated probes and probes located on sex chromosomes. We then performed unsupervised hierarchical clustering of the HPV(−) HNSCC samples using the methylation values of the top 500 of these probes with the highest average methylation value. To determine whether other gene alterations had an effect on the methylome, we took each gene mutated in more than 5% of HPV(−) HNSCC samples in TCGA and calculated whether each CpG site was differentially methylated (between gene mutant versus wild type) using the Wilcoxon rank-sum test. The resulting p-values for each CpG site were Bonferroni corrected and called significant if $q < 0.05$. To determine location of differentially methylated regions for *NSD1* mutated tumors, CpG sites were binned in 200-marker-long sliding windows along the length of the chromosome. The number of differentially methylated CpG sites was summed, indexed against a standard normal distribution and assigned a Z-score with a corresponding p-value.

Cell lines and disruption of *NSD1*

Two of these lines were generated from CAL33, an HPV(−) HNSCC cell line, and one from UM-SCC47, an HPV(+) HNSCC cell line. The UM-SCC47 cell line was obtained from the laboratory of Dr. Silvio Gutkind on April 20, 2016, where the identity and HPV(+) status was authenticated using STR profiling by IDEXX BioResearch on September 01, 2016. The CAL33 cell line was obtained from the German Collection of Microorganisms and Cell Cultures (DSMZ, Catalog# ACC-447), also via the Gutkind lab on April 20, 2016. The identity and HPV(−) status of the CAL33 cell line was originally confirmed with STR profiling by Genetica DNA Laboratories via the Gutkind lab and was reconfirmed by STR profiling at IDEXX BioResearch on February 08, 2018. Both cell lines were tested for mycoplasma using a PCR-based test kit (Applied Biological Materials, Inc.) upon receipt and again each time a new frozen vial was started (the latest test was performed on January 10, 2018). Neither CAL33 nor UM-SCC47 were mutated in *NSD1* prior to our CRISPR experiments, and there is a 0.75 copy number amplification in UM-SCC47 but no copy number alteration in CAL33 (11). To generate *NSD1* disrupted cell lines, two guide RNAs were selected from the GeCKO v2 CRISPR library (12) and synthesized with overhanging regions mapping to the GeCKO v2 backbone sequence. The synthesized oligos (20 bp gRNA sequence is underlined below) were then assembled onto the CRISPR v2 backbone via Gibson assembly (New England Biosciences, #E5510S) and transformed into STBL3 competent cells (Invitrogen, #C7373-03). The synthesized oligos were:

Library ID HGLibA_32744:

GAAAGGACGAAACACCG<u>CTGGCTCGAGATTTAGCGCA</u>GTTTTAGAGCT AGAAATAGCAAGTAAAATAAGGC

Library ID HGLibA_32745:

GGAAAGGACGAAACACCGAATCTGTTCATGCGCTTACGGTTTTAGAGC TAGAAATAGCAAGTTAAAATAAGGC

Transformed cells were grown overnight at 37°C on LB agar with 100 μg/mL ampicillin. Single clones were picked, cultured in liquid, miniprepped, and Sanger sequenced to confirm successful assembly. Successfully assembled vectors were packaged into virus by transfecting 293T cells using lipofectamine 3000 (Invitrogen, L3000-015) with the following plasmid amounts per 10 cm culture dish: 1.2 μg PMD2.G (Addgene, #12259), 4.8 μg of pCMV-dR8.2 dvpr (Addgene, #8455) and 3.6 μg of CRISPRv2-*NSD1* vector. Virus was collected at 48 and 72 hours, filtered (0.45 μm) and concentrated (Millipore, #UFC910024).

The CAL33 and UM-SCC47 cell lines were separately transduced using 0.8 μg/mL polybrene and 10 – 20 μL of CRISPRv2-*NSD1* lentivirus. Previously performed viability assays found that 1 μg/mL of puromycin was sufficient for selecting stable cell lines. To generate monoclonal populations, puromycin selection was started at 48 hours post-transduction, after which cultures were diluted and single clones selected for further study. Disruptions in the *NSD1* gene were identified by extracting genomic DNA, PCR amplifying 100 bp upstream to 100 bp downstream of the guide RNAs and performing Sanger sequencing on these amplicons. NSD1 and TBP (TATA Binding Protein) expression levels were determined by extracting total protein from various cell lines and quantitated using the Wes electropherogram (ProteinSimple) using an anti-NSD1 antibody (EMD Millipore, ABE1009, 1:100 dilution) and an anti-TBP antibody (Abgent, AP6680b, 1:50 dilution for CAL33, 1:500 dilution for UM-SCC47). Experiments using pools of *NSD1* disrupted cells (as opposed to any single clone) were constructed and grown using the same procedure described above without selecting for monoclonal populations.

CpG methylation arrays and analysis

Wild type and *NSD1* disrupted cell lines were trypsinized and counted so that 4×10^6 cells could be pelleted, washed in PBS, pelleted again and then snap-frozen in liquid nitrogen. The DNeasy Blood & Tissue kit (Qiagen, 69506) was used to extract genomic DNA. Genomic DNA was quantified using the Qubit assay (Thermo Fisher). Methylation was assayed using the Infinitum MethylationEPIC BeadChip Kit (Illumina) with 750 ng of genomic DNA per sample. The R package ‘Minfi’ (13) was used to process the raw data. The resulting beta values were quantile normalized using Minfi, and probe biases were normalized using BMIQ (14). The top 10,000 most differentially methylated CpG loci were identified by taking the absolute value of the difference between the methylation beta value each CpG site between the respective parent and *NSD1* disrupted cell lines. Identification of the hypomethylated peak was done by fitting a Gaussian mixture model using the Sci-Kit Learn Package (15) to the density plot of differential methylation values and extracting the peak density value at the smallest Gaussian component mean for each distribution. Shared CpG probes between the parent and *NSD1* disrupted cell line were determined by mapping CpG probes to genes and performing set pairwise intersections.

RT-qPCR

500,000 cells were aliquoted into an Eppendorf tube, washed once with PBS, snap frozen in liquid nitrogen and stored at −80 C until ready for RT-qPCR assay. Cells were lysed and RNA extracted using a Quick-RNA miniprep kit (Zymo Research) and then converted to cDNA using Superscript III First-strand synthesis kit (Invitrogen). RT-qPCR assays were run on a Bio-Rad CFX96 using Sso Advanced Universal SYBR Green (Bio-Rad) using two technical replicates per gene. Differential expression was measured relative to the *LMNB1* probe:

Fwd: CTG GCG AAG ATG TGA AGG TTA T

Rev: TCC TCC TCT TCT TCA GGT ATG G

The probe sequences for the genes tested are as follows:

COL13A

Fwd: GCA GAC ACT TGA AGG GAA AGA

Rev: CGT TCC AAG TCC AGG AAA GTT A

NTM

Fwd: CAT CCT CTA TGC TGG GAA TGA C

Rev: CGT CAT ACA CAT CCA CGT TCT

PDE1A

Fwd: CCA TGA GTG ATG GGT CCT ATT C

Rev: CAG CTA ACT CTT TCC ACC TCT C

Drug sensitivity assay

Cell viability in response to cisplatin (Spectrum Chemical, #C1668) was assayed in 96 well plates with continuous exposure to cisplatin for 72 hours. Cells were plated at 5,000 cells per well, allowed to attach overnight and then treated 24 hours later with cisplatin at doses from 0 to 20 μM. Six technical replicates were performed for each dose. After 72 hours exposure to cisplatin, a 10× stock of resazurin (working concentration 44 μM) was added and incubated for 4 to 6 hours. Fluorescence intensity at 590 nM was measured using a plate-reading spectrophotometer (Tecan). The resulting data were analyzed with GraphPad Prism. For experiments with the HMT inhibitor (HMTi) UNC0379 (Selleckchem, #S7570) (16), dose-response curves in both cell lines were initially performed to select non-toxic doses. The highest dose without a significant toxic effect was 0.5 μM for both CAL33 and UM-SCC47. Prior to plating for the cisplatin assay, cells were pretreated at this dose for 72 hours.

γH2AX immunofluorescence assay

2,000 cells were seeded into clear-bottom 384-well plates (Nunc), allowed to attach overnight, and treated with cisplatin or vehicle the following day. After 48 hours, cells were fixed with 4% formaldehyde, blocked with 2% bovine serum albumin in Tris Buffered

Saline with 0.1% TRITON X-100 (TBST), and stained with Hoechst (1:1000) and FITC-conjugated anti-γH2AX antibody (1:333, Millipore). Plates were imaged with an ImageXpress Micro automated epi-fluorescent microscope (Molecular Devices). Images were scored with MetaExpress analysis software (Molecular Devices), and statistical analysis was performed with Prism 7 (GraphPad Software). The percentage of γH2AX positive cells in cisplatin-treated samples was normalized to untreated controls.

Clonogenic radiation assays

Clonogenic radiation assays were performed with slight modification to a previously published protocol (17). A Canon Rebel T3i digital camera was used to create a digital image of each plate. Colonies were then scored using a custom Matlab script calibrated against manually counted control plates for each cell line. A range of 1,000 – 10,000 cells was used in an initial experiment to determine plating efficiency. For radiation experiments, cells were counted, radiated while in suspension, then immediately plated and allowed to grow for eight days. The percent viability was calculated by normalizing to the number of colonies on plates without radiation treatment. Each cell line was normalized independently. Normalized survival data were then fitted to a weighted, stratified regression according to the following linear–quadratic formula for radiation dose-effect (18):

$$Y = 100 * e^{-(\alpha X + \beta X^2)}$$

where Y is the percentage of surviving cells, X is the radiation dose in Gy, α is the coefficient for linear killing and β is the coefficient for quadratic killing; α and β are constrained to be greater than zero. Curves for parent and knockout cell lines were fit using Prism v7.03 (GraphPad Software). An extra-sum-of-squares F-test with a significance threshold of $p < 0.05$ was used to determine if a single curve or two separate curves for parent and *NSD1* disrupted cell lines best fit the data.

Analysis of drug sensitivity in 1,001 cell lines

Data for cell lines, mutation calls and drug sensitivity were downloaded from the Genomics of Drug Sensitivity in Cancer database, maintained by the Sanger Institute (http://www.cancerrxgene.org/) (19). Cell lines with *NSD1* loss-of-function mutations (nonsense or frameshift mutations) were separated from *NSD1* wild type cell lines. A Volcano plot was constructed by performing Student's t-test on the $\ln(IC_{50})$ for all drugs with sensitivity data on ≥15 *NSD1* loss-of-function cell lines. Effect size was represented by the mean difference in $\ln(IC_{50})$, and *p*-value was derived from the t-test.

RESULTS

NSD1 mutations are associated with significantly improved patient survival

We began by analyzing 421 HPV(−) HNSCC patients from TCGA with complete exome sequencing data. Previous results from MutSig (8) were used to identify 11 distinct genes that are recurrently mutated in this cohort (20). When we compared patients with and without mutations in each of these genes, only patients with mutations in *NSD1* showed a

difference in survival after accounting for clinical covariates (Hazard Ratio (HR) 0.45, p = 0.007, Cox Proportional Hazards) (Fig. 1A). Patients with mutations in *NSD1* had a markedly improved clinical outcome, with an approximately five-year absolute increase in median overall survival time (8.0 versus 3.1 years) (Fig. 1B). Interestingly, patients with *NSD1* mutations were enriched for those with a history of smoking (p = 0.002, Chi-squared). When restricting analysis to only current and former smokers, those with *NSD1* mutations had significantly improved survival relative to wild type (HR 0.46, p = 0.008, Cox Proportional Hazards) (Supplemental Fig. 1A–B). There were too few *NSD1* mutations in non-smokers to evaluate the corresponding survival effects for those patients. This survival advantage was validated in a second, independent cohort of 68 HPV(−) HNSCC patients from the University of Chicago (21). In this second cohort, *NSD1*-mutated patients demonstrated an improvement in both progression free and overall survival (Supplemental Fig. 1C–D).

When *NSD1* was examined across other tissue cohorts in TCGA, we found that HNSCC was the tissue with both the highest percentage of *NSD1* mutations (12.2% of patients) and the highest percentage of deleterious mutations (66% of *NSD1* mutations), reflecting a tissue-specific phenotype (Fig. 1C). In the HPV(−) HNSCC cohort, loss-of-function *NSD1* alterations (*i.e.*, nonsense mutations, frameshift mutations or homozygous copy number deletions) were associated with significantly lower mRNA expression levels (Fig. 1D). Missense mutations did not significantly impact *NSD1* mRNA expression levels but tended to cluster near the SET domain (Fig. 1E). To investigate the pathogenicity of these missense mutations, we separated tumors with truncating loss-of-function *NSD1* alterations into a distinct group from those with missense *NSD1* mutations and tested the association of each group with survival. Strikingly, patients with *NSD1* missense mutations had increased survival compared to *NSD1* wild type patients (p = 0.042 by Log-Rank Test, Supplemental Fig. 1E), with an effect that was indistinguishable from *NSD1* loss-of-function mutations. This evidence suggested that the SET domain in *NSD1* is important to the function of the protein, such that missense mutations in this domain lead to loss-of-function of *NSD1.*

NSD1 is a key regulator of the epigenome

Given the role of *NSD1* as an HMT, we sought to determine if somatic mutations in *NSD1* in HPV(−) head and neck cancer patients might also be associated with CpG hypomethylation. Therefore, we hierarchically clustered the HPV(−) HNSCC samples from TCGA for which CpG methylation data were available (n=421) based on the methylation status of 500 selected CpG sites (**Methods**). We found that most patients with mutations in *NSD1* were placed in the same cluster due to a clear pattern of hypomethylated CpG sites (Fig. 2A). Loss-of-function alterations comprised the majority of this cluster whereas missense mutations were more likely to be outliers.

To determine if disruptions in other genes also correlated with changes in CpG methylation, we examined every gene that was mutated in more than 5% of the HPV(−) HNSCC samples in TCGA (n=132) and determined the percentage of CpG sites that were differentially methylated between wild-type and mutant tumors. Whereas about 14% of CpG sites were differentially mutated between *NSD1* mutant and wild-type tumors, no other gene mutation

impacted more than 2% of CpG sites (Fig. 2B). For the *NSD1*-associated differentially methylated CpG sites, a striking 98.9% were hypomethylated. Therefore, the profound association between genetic alteration and hypomethylation is unique to *NSD1*.

Next, we asked whether CpG hypomethylation in tumors with *NSD1* mutations is preferentially located in any particular region of the genome. Using a sliding window consisting of 200 consecutive CpG sites along each chromosome, we identified a region enriched for hypomethylated CpGs on chromosome 6 (Supplemental Fig. 2). This hypomethylated region includes the MHC I and MHC III loci as well as genes that regulate connective tissue and skin structure (Supplemental Table 1).

Disrupting *NSD1* in HNSCC cell lines leads to CpG hypomethylation

To determine whether disruptions to *NSD1* are sufficient to alter CpG methylation levels, and the dependence of this effect on HPV status, we used CRISPR-Cas9 to generate three monoclonal cell lines with NSD1 truncating mutations. In each case, at least one allele of *NSD1* was disrupted by CRISPR, leading to decreased protein expression levels (Supplemental Fig. 3A–D). Methylation status in the parental or *NSD1* disrupted cell lines was determined using the Illumina MethylationEpic BeadChip, which measures CpG methylation levels at >850,000 CpG sites. For each pair of parental and *NSD1* disrupted cell lines, we examined the methylation levels for the 10,000 most differentially methylated CpG sites (**Methods**). Substantial hypomethylation was also observed in all *NSD1* disrupted cell lines, regardless of HPV status (Fig. 3A–D). The associated differentially methylated regions (DMRs) were consistently enriched in enhancer and intergenic regions, and depleted in promoter regions. This finding is consistent with observations in TCGA patients with *NSD1* mutations and patients with Sotos Syndrome (22), a childhood disease caused by germline mutations in *NSD1* (Supplemental Fig. 4A).

Analysis of the hypomethylated CpG sites revealed eight genes with differentially hypomethylated CpGs in all three *NSD1* disrupted cell lines and across HNSCC tumors (Supplemental Table 2). The expression levels of some of these genes have been associated with chemotherapy response or implicated as tumor suppressors (Supplemental Table 2). We found that four of these genes were expressed at detectable levels in HNSCC TCGA patients, of which three were significantly down-regulated when *NSD1* was mutated (Student's T-Test): *COL13A1* ($p = 4.1\times10^{-3}$), *NTM* ($p = 1.3\times10^{-2}$), and *PDE1A* ($p = 4.7\times10^{-2}$). We performed RT-qPCR on these three genes to determine if disrupting *NSD1* leads to similar expression changes as observed in patients. Indeed, two of these genes were consistently down-regulated by *NSD1* disruption in two distinct cell lines (Supplemental Figs. 4B–C).

NSD1 disruption confers sensitivity to cisplatin

Given reported associations between DNA hypomethylating agents and platinum sensitivity (23–26), we hypothesized that the improved survival of *NSD1*-mutated patients might be due to increased sensitivity to cisplatin, a common chemotherapy used to treat HNSCC patients. In each case, cell lines with *NSD1* disruption were more sensitive to cisplatin than the parental wild-type cell lines (Figs. 4A–C). To mimic the loss of *NSD1*

pharmacologically, we performed a separate experiment in which parental cells were pre-treated with the HMT inhibitor UNC0379, which also rendered the HNSCC cell lines more sensitive to cisplatin with a growth response that was nearly identical to direct *NSD1* disruption (Figs. 4A–B). To investigate whether the sensitivity to cisplatin was related to its DNA damage activity, we performed a high-throughput immunofluorescence assay to measure phosphorylation of histone H2AX at Ser139 (γH2AX), an established marker of DNA damage (27,28). Indeed, *NSD1*-disrupted CAL33 cells had increased γH2AX signal when treated with cisplatin relative to wild-type controls (**Methods**, Supplemental Fig. 5A). Since radiation is also standard treatment for patients with HNSCC, we tested whether *NSD1* disruption caused sensitivity to radiation using clonogenic assays on the CAL33 cell line **(Methods)**. While the disruption of *NSD1* significantly reduced the formation of colonies (i.e. plating efficiency) relative to wild type (Supplemental Fig. 5B), after normalizing for this effect, we did not observe a significant difference in the radiation dose-response curves for CAL33 ($p = 0.15$, Extra-sum-of-squares F-test, Supplemental Fig. 5C).

Finally, we investigated whether this drug sensitivity was specific to HNSCC, by analyzing a collection of 1,001 cancer cell lines representing a diverse set of tumor types (Supplemental Fig. 6) with full genomic profiles and measured responses to 265 anti-cancer drugs (19). Comparing differential drug sensitivity between cell lines containing at least one *NSD1* allele with a truncating mutation (n=17) and those with wild-type *NSD1* (n=774), we found that drugs targeting DNA replication or genome integrity were more likely to be effective in cell lines with *NSD1* disrupted ($p = 0.003$, Wilcoxon rank sum; Fig. 4D). One of the most effective drugs in this category was cisplatin, with a 24% decrease in IC_{50} relative to wild type ($p = 0.02$, Student's t-test; Fig. 4E). Taken together, these data suggest that *NSD1* loss-of-function increases sensitivity to DNA damaging chemotherapies, such as cisplatin, and the effect may generalize beyond HNSCC cell lines.

DISCUSSION

While our study has focused on somatic mutations of a particular gene, *NSD1*, in a particular context, HNSCC, the implications may in fact be broader. *NSD1* is altered in other tumor types, including epigenetic inactivation through promoter hypermethylation in glioma (29) and translocations with a fusion protein in acute myeloid leukemia (*NUP98*/*NSD1*) (30–32). While *NSD1* has been shown as a biomarker for global epigenetic changes in cancer (33,34), we have also shown here that *NSD1* is a prognostic biomarker in patients with HPV(−) HNSCC. Beyond *NSD1* itself, the NSD family of HMTs has been linked to various cancers, with *NSD2* mutations seen in breast cancer, lung cancer and acute myeloid leukemia (35,36).

The connection between *NSD1* loss-of-function mutations and CpG hypomethylation is also seen in the germline setting. Patients with Sotos syndrome have inherited loss-of-function mutations in *NSD1* and present clinically with childhood overgrowth, non-progressive developmental delay and a distinctive facial appearance (37). A recent genomic analysis of Sotos syndrome patients found a genome-wide DNA hypomethylation signature that distinguishes them from normal controls (22). The affected genes function in cellular morphogenesis and neuronal differentiation, consistent with the clinical phenotype. Sotos

Syndrome follows an autosomal dominant inheritance pattern, consistent with our observation that the *NSD1* truncating mutations found in HNSCC are hemizygous, suggesting that loss of a single copy of *NSD1* is sufficient to cause hypomethylation.

An important question is how *NSD1*, an HMT, can impact methylation of not only histones but also DNA. Indeed, histone methylation and DNA methylation are intertwined in a complex relationship (38), and at least two mechanisms are plausible. First, cells deficient in *NSD1* are unable to mono- and di-methylate H3K36 (39–41). In turn, this defect likely affects the ability of these histones to recruit DNA methyltransferases (34), leading to a global DNA hypomethylation signature. Another connection between HMTs and DNA methylation is that some SET-domain containing HMTs physically recruit DNA methyltransferase leading to CpG methylation (42).

A second question relates to how hypomethylation of DNA is connected to cisplatin sensitivity. Indeed, DNA hypomethylation has been previously implicated as a potential sensitizer for several chemotherapeutic agents, including cisplatin and other platinum-based treatments (23,24,43). Treating cisplatin-resistant HNSCC cell lines with decitabine, a cytidine analog that inhibits DNA methylation leading to global DNA hypomethylation, also renders these cells sensitive to cisplatin (25). In diffuse large B-cell lymphoma, treating cells with DNA methyltransferase inhibitors leads to the expression of previously repressed genes and renders these cells sensitive to chemotherapy (26). Based on some of these observations, combinations of hypomethylating agents and cisplatin have been attempted in head and neck cancer in phase I clinical trials (NCT00901537 and NCT00443261), however both trials were terminated early due to accrual problems. Preliminary results (NCT00901537) show encouraging activity with one partial response, one patient with progression free survival for 15 months and another with progression free survival for greater than 6 months (44). Given our finding that cells become more sensitive to cisplatin after *NSD1* disruption or pharmacological inhibition of HMTs, perhaps an HMTi could be used along with platinum-based therapy to more effectively treat HPV(−) HNSCC patients. In addition to platinum sensitivity, we also found that disrupting *NSD1* dramatically reduced the clonogenic growth capacity of the CAL33 cell line. This finding may also be related to the survival advantage seen in patients with *NSD1* mutant tumors, and should be studied in a greater number of cell lines across cancer lineages.

Given that *NSD1* mutation is associated with a dramatic increase in the survival of HPV(−) HNSCC patients in multiple cohorts, we propose that patients with loss-of-function *NSD1* mutations should be considered a distinct clinical subclass of HPV(−) HNSCC. In addition to serving as a prognostic biomarker, the *in vitro* cisplatin sensitivity data suggest that *NSD1* mutation is also predictive of response to cisplatin chemotherapy. Although clinical validation of this finding is still needed, our results suggest that cisplatin should be strongly considered for any HNSCC patient with *NSD1* loss-of-function mutation, especially since platinum chemotherapy is already part of the standard of care. Given the clear influence on survival as well as the distinct molecular features of *NSD1* mutant tumors, future prospective clinical trials of HPV(−) HNSCC should include these tumors as a planned subgroup with expected differences in therapeutic response.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

The results published here are in part based on data generated by TCGA, established by the National Cancer Institute and the National Human Genome Research Institute, and we are grateful to the specimen donors and relevant research groups associated with this project. We would also like to thank investigators at the University of Chicago for clinical and sequencing data from their head and neck cohort. We would also like to thank the laboratory of Dr. Mark Tuszynski for use of their imaging equipment for our γH2AX Immunofluorescence assay and the lab of Dr. Silvio Gutkind for helping in the acquisition of the cell lines used.

Grant Support

This work was generously supported by the following grants from the U.S. National Institutes of Health: R01 ES014811, P41 GM103504 and U24 CA184427 to T. Ideker, U54 CA209891 and P50 GM085764 to J.F. Kreisberg and T. Ideker, L30 CA171000 to J.P. Shen, T32 GM008806 to J.K. Huang and T32 GM008666 to T. Wang. J.P. Shen was also supported by a Career Development Grant from the Tower Cancer Research Foundation.

References

1. Jemal A, Bray F, Center MM, Ferlay J, Ward E, Forman D. Global cancer statistics. CA Cancer J Clin. 2011; 61:69–90. [PubMed: 21296855]
2. Huang SH, Xu W, Waldron J, Siu L, Shen X, Tong L, et al. Refining American Joint Committee on Cancer/Union for International Cancer Control TNM stage and prognostic groups for human papillomavirus-related oropharyngeal carcinomas. J Clin Oncol. 2015; 33:836–45. [PubMed: 25667292]
3. Bhatia A, Burtness B. Human Papillomavirus-Associated Oropharyngeal Cancer: Defining Risk Groups and Clinical Trials. J Clin Oncol. 2015; 33:3243–50. [PubMed: 26351343]
4. Hitt R, Grau JJ, López-Pousa A, Berrocal A, García-Girón C, Irigoyen A, et al. A randomized phase III trial comparing induction chemotherapy followed by chemoradiotherapy versus chemoradiotherapy alone as treatment of unresectable head and neck cancer. Ann Oncol. 2014; 25:216–25. [PubMed: 24256848]
5. Cohen EEW, Karrison TG, Kocherginsky M, Mueller J, Egan R, Huang CH, et al. Phase III randomized trial of induction chemotherapy in patients with N2 or N3 locally advanced head and neck cancer. J Clin Oncol. 2014; 32:2735–43. [PubMed: 25049329]
6. Magrini SM, Buglione M, Corvò R, Pirtoli L, Paiar F, Ponticelli P, et al. Cetuximab and Radiotherapy Versus Cisplatin and Radiotherapy for Locally Advanced Head and Neck Cancer: A Randomized Phase II Trial. J Clin Oncol. 2016; 34:427–35. [PubMed: 26644536]
7. Buglione M, Maddalo M, Corvò R, Pirtoli L, Paiar F, Lastrucci L, et al. Subgroup Analysis According to Human Papillomavirus Status and Tumor Site of a Randomized Phase II Trial Comparing Cetuximab and Cisplatin Combined With Radiation Therapy for Locally Advanced Head and Neck Cancer. Int J Radiat Oncol Biol Phys. 2017; 97:462–72. [PubMed: 27986347]
8. Cancer Genome Atlas Network. Comprehensive genomic characterization of head and neck squamous cell carcinomas. Nature. 2015; 517:576–82. [PubMed: 25631445]
9. Broad Institute TCGA Genome Data Analysis Center. Analysis-ready standardized TCGA data from Broad GDAC Firehose 2016_01_28 run. Broad Institute of MIT and Harvard; 2016. Dataset [Internet]
10. Therneau TM, Grambsch PM. Modeling Survival Data: Extending the Cox Model. 2000
11. Martin D, Abba MC, Molinolo AA, Vitale-Cross L, Wang Z, Zaida M, et al. The head and neck cancer cell oncogenome: a platform for the development of precision molecular therapies. Oncotarget. 2014; 5:8906–23. [PubMed: 25275298]
12. Sanjana NE, Shalem O, Zhang F. Improved vectors and genome-wide libraries for CRISPR screening. Nat Methods. 2014; 11:783–4. [PubMed: 25075903]

13. Aryee MJ, Jaffe AE, Corrada-Bravo H, Ladd-Acosta C, Feinberg AP, Hansen KD, et al. Minfi: a flexible and comprehensive Bioconductor package for the analysis of Infinium DNA methylation microarrays. Bioinformatics. 2014; 30:1363–9. [PubMed: 24478339]
14. Teschendorff AE, Marabita F, Lechner M, Bartlett T, Tegner J, Gomez-Cabrero D, et al. A beta-mixture quantile normalization method for correcting probe design bias in Illumina Infinium 450 k DNA methylation data. Bioinformatics. 2013; 29:189–96. [PubMed: 23175756]
15. Varoquaux G, Buitinck L, Louppe G, Grisel O, Pedregosa F, Mueller A. Scikit-learn. GetMobile: Mobile Computing and Communications. 2015; 19:29–33.
16. Ma A, Yu W, Li F, Bleich RM, Herold JM, Butler KV, et al. Discovery of a selective, substrate-competitive inhibitor of the lysine methyltransferase SETD8. J Med Chem. 2014; 57:6822–33. [PubMed: 25032507]
17. Franken NAP, Rodermond HM, Stap J, Haveman J, van Bree C. Clonogenic assay of cells in vitro. Nat Protoc. 2006; 1:2315–9. [PubMed: 17406473]
18. Barendsen GW. Parameters of linear-quadratic radiation dose-effect relationships: dependence on LET and mechanisms of reproductive cell death. Int J Radiat Biol. 1997; 71:649–55. [PubMed: 9246179]
19. Iorio F, Knijnenburg TA, Vis DJ, Bignell GR, Menden MP, Schubert M, et al. A Landscape of Pharmacogenomic Interactions in Cancer. Cell. 2016; 166:740–54. [PubMed: 27397505]
20. Lawrence MS, Stojanov P, Polak P, Kryukov GV, Cibulskis K, Sivachenko A, et al. Mutational heterogeneity in cancer and the search for new cancer-associated genes. Nature. 2013; 499:214–8. [PubMed: 23770567]
21. Seiwert TY, Zuo Z, Keck MK, Khattri A, Pedamallu CS, Stricker T, et al. Integrative and comparative genomic analysis of HPV-positive and HPV-negative head and neck squamous cell carcinomas. Clin Cancer Res. 2015; 21:632–41. [PubMed: 25056374]
22. Choufani S, Cytrynbaum C, Chung BHY, Turinsky AL, Grafodatskaya D, Chen YA, et al. NSD1 mutations generate a genome-wide DNA methylation signature. Nat Commun. 2015; 6:10207. [PubMed: 26690673]
23. Fu S, Hu W, Iyer R, Kavanagh JJ, Coleman RL, Levenback CF, et al. Phase 1b-2a study to reverse platinum resistance through use of a hypomethylating agent, azacitidine, in patients with platinum-resistant or platinum-refractory epithelial ovarian cancer. Cancer. 2011; 117:1661–9. [PubMed: 21472713]
24. Asadollahi R, Hyde CAC, Zhong XY. Epigenetics of ovarian cancer: From the lab to the clinic. Gynecol Oncol. 2010; 118:81–7. [PubMed: 20421130]
25. Viet CT, Dang D, Achdjian S, Ye Y, Katz SG, Schmidt BL. Decitabine rescues cisplatin resistance in head and neck squamous cell carcinoma. PLoS One. 2014; 9:e112880. [PubMed: 25391133]
26. Clozel T, Yang S, Elstrom RL, Tam W, Martin P, Kormaksson M, et al. Mechanism-based epigenetic chemosensitization therapy of diffuse large B-cell lymphoma. Cancer Discov. 2013; 3:1002–19. [PubMed: 23955273]
27. Sharma A, Singh K, Almasan A. Histone H2AX phosphorylation: a marker for DNA damage. Methods Mol Biol. 2012; 920:613–26. [PubMed: 22941631]
28. Paull TT, Rogakou EP, Yamazaki V, Kirchgessner CU, Gellert M, Bonner WM. A critical role for histone H2AX in recruitment of repair factors to nuclear foci after DNA damage. Curr Biol. 2000; 10:886–95. [PubMed: 10959836]
29. Berdasco M, Ropero S, Setien F, Fraga MF, Lapunzina P, Losson R, et al. Epigenetic inactivation of the Sotos overgrowth syndrome gene histone methyltransferase NSD1 in human neuroblastoma and glioma. Proc Natl Acad Sci U S A. 2009; 106:21830–5. [PubMed: 20018718]
30. Xu H, Valerio DG, Eisold ME, Sinha A, Koche RP, Hu W, et al. NUP98 Fusion Proteins Interact with the NSL and MLL1 Complexes to Drive Leukemogenesis. Cancer Cell. 2016; 30:863–78. [PubMed: 27889185]
31. Struski S, Lagarde S, Bories P, Puiseux C, Prade N, Cuccuini W, et al. NUP98 is rearranged in 3.8% of pediatric AML forming a clinical and molecular homogenous group with a poor prognosis. Leukemia. 2017; 31:565–72. [PubMed: 27694926]
32. Hollink IHIM, van den Heuvel-Eibrink MM, Arentsen-Peters STCJM, Pratcorona M, Abbas S, Kuipers JE, et al. NUP98/NSD1 characterizes a novel poor prognostic group in acute myeloid

leukemia with a distinct HOX gene expression pattern. Blood. 2011; 118:3645–56. [PubMed: 21813447]

33. Lee S-T, Wiemels JL. Genome-wide CpG island methylation and intergenic demethylation propensities vary among different tumor sites. Nucleic Acids Res. 2016; 44:1105–17. [PubMed: 26464434]
34. Papillon-Cavanagh S, Lu C, Gayden T, Mikael LG, Bechet D, Karamboulas C, et al. Impaired H3K36 methylation defines a subset of head and neck squamous cell carcinomas. Nat Genet. 2017; 49:180–5. [PubMed: 28067913]
35. Morishita M, di Luccio E. Cancers and the NSD family of histone lysine methyltransferases. Biochim Biophys Acta. 2011; 1816:158–63. [PubMed: 21664949]
36. He C, Li F, Zhang J, Wu J, Shi Y. The methyltransferase NSD3 has chromatin-binding motifs, PHD5-C5HCH, that are distinct from other NSD (nuclear receptor SET domain) family members in their histone H3 recognition. J Biol Chem. 2013; 288:4692–703. [PubMed: 23269674]
37. McClelland J, Burgess B, Crock P, Goel H. Sotos syndrome: An unusual presentation with intrauterine growth restriction, generalized lymphedema, and intention tremor. Am J Med Genet A. 2016; 170A:1064–9. [PubMed: 26738611]
38. Tamaru H, Selker EU. A histone H3 methyltransferase controls DNA methylation in Neurospora crassa. Nature. 2001; 414:277–83. [PubMed: 11713521]
39. Qiao Q, Li Y, Chen Z, Wang M, Reinberg D, Xu R-M. The structure of NSD1 reveals an autoregulatory mechanism underlying histone H3K36 methylation. J Biol Chem. 2011; 286:8361–8. [PubMed: 21196496]
40. Suzuki S, Murakami Y, Takahata S. H3K36 methylation state and associated silencing mechanisms. Transcription. 2016; 8:26–31. [PubMed: 27723431]
41. Mutations in Histone H3K36 Prevent Methylation and Drive Sarcomagenesis. Cancer Discov. 2016; 6:689–689.
42. Cedar H, Bergman Y. Linking DNA methylation and histone modification: patterns and paradigms. Nat Rev Genet. 2009; 10:295–304. [PubMed: 19308066]
43. Li M, Balch C, Montgomery JS, Jeong M, Chung JH, Yan P, et al. Integrated analysis of DNA methylation and gene expression reveals specific signaling pathways associated with platinum resistance in ovarian cancer. BMC Med Genomics. 2009; 2:34. [PubMed: 19505326]
44. Liao Y-M, Mirshahidi H, Zhang K, Mirshahidi S, Williamson S, Hsueh C-T. Abstract 2663: Phase I study of azacitidine and cisplatin in patients with advanced head and neck or non-small cell lung cancer. Cancer Res. 2012; 72:2663–2663.
45. Mermel CH, Schumacher SE, Hill B, Meyerson ML, Beroukhim R, Getz G. GISTIC2.0 facilitates sensitive and confident localization of the targets of focal somatic copy-number alteration in human cancers. Genome Biol. 2011; 12:R41. [PubMed: 21527027]
46. Gao J, Aksoy BA, Dogrusoz U, Dresdner G, Gross B, Sumer SO, et al. Integrative analysis of complex cancer genomics and clinical profiles using the cBioPortal. Sci Signal. 2013; 6:11.
47. Cerami E, Gao J, Dogrusoz U, Gross BE, Sumer SO, Aksoy BA, et al. The cBio cancer genomics portal: an open platform for exploring multidimensional cancer genomics data. Cancer Discov. 2012; 2:401–4. [PubMed: 22588877]

Precis

Disruptions in the *NSD1* gene in HPV(−) head and neck cancer are associated with improved prognosis, genome-wide CpG hypomethylation and cisplatin sensitivity.

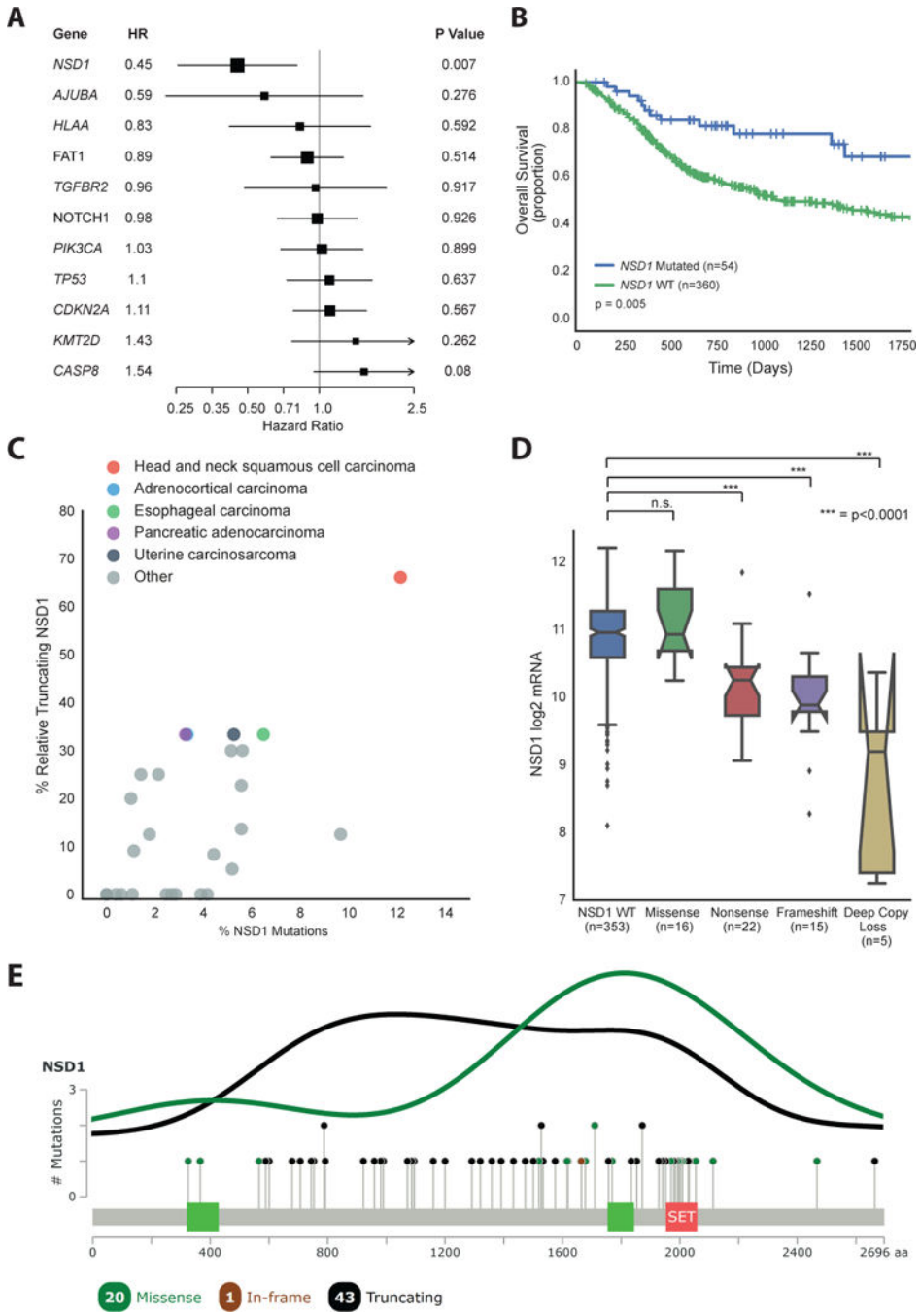

Figure 1.

NSD1 mutations are associated with improved survival in the HPV(−) HNSCC cohort in TCGA. **A**, Forest plot of the prognostic influence of the 11 most recurrently mutated genes in the HPV(−) HNSCC cohort in TCGA. Hazard ratios derived from Cox proportional hazards model incorporating the clinical covariates age, stage, grade, gender, smoking status and anatomical location. **B**, Kaplan-Meier curve showing overall survival from the HPV(−) HNSCC cohort in TCGA. **C**, Head and neck cancer possess a high percentage of *NSD1* mutations and a high percentage of relative truncating mutations. **D**, Loss-of-function *NSD1*

mutations and homozygous deletions, defined as a −2 copy number change by GISTIC (45), have significantly lower gene expression than wild-type or missense mutations. **E**, Lollipop plot of location of *NSD1* mutations as generated by cBioPortal (46,47). The lines represent density plots of truncating (black) and missense (green) mutations.

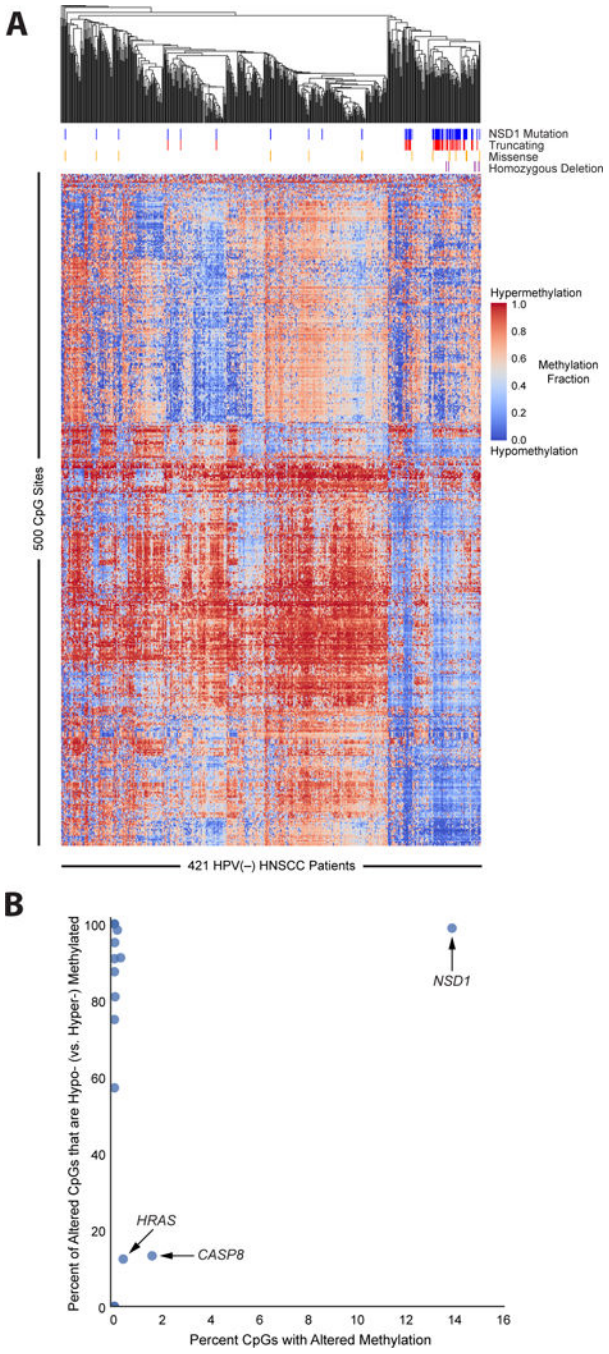

Figure 2.

CpG hypomethylation in patients with *NSD1* loss-of-function mutations in the HPV(−) HNSCC cohort in TCGA. **A**, Unsupervised hierarchical clustering based on the methylation status of 500 selected CpG sites reveals a tight cluster of hypomethylation centered around *NSD1* mutations. Analysis of alteration type (top color bar) reveals that missense mutations (orange) were more likely to be outliers while truncating (red) and homozygous deletions (purple) were associated with the hypomethylation signal. **B**, Gene level methylation analysis reveals that *NSD1* is the only gene where mutations cause a significant change to

the methylome (x-axis: ~13% of all CpG sites) with all other genes at <2%. The direction of methylation changes are strikingly in the hypomethylated direction with a full 98.9% of differentially methylated CpG probes being hypomethylated (y-axis).

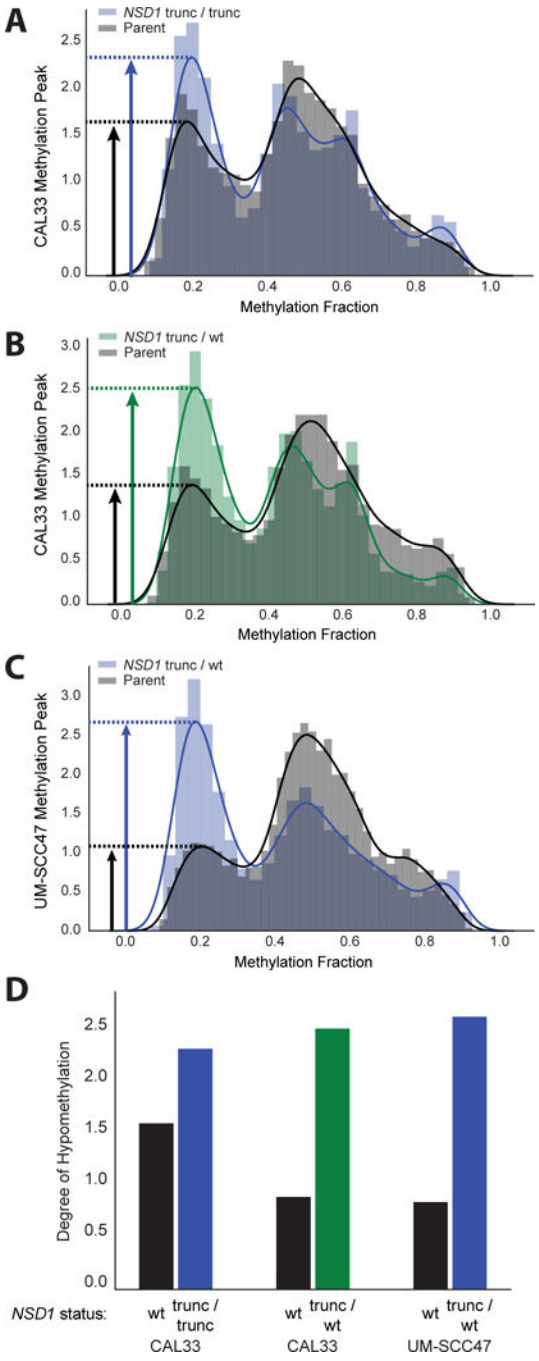

Figure 3.

CpG hypomethylation in cell lines with *NSD1* disrupted. **A–B**, Methylation analysis of top 10,000 most differentially methylated CpG sites in CAL33 with and without *NSD1* disrupted demonstrates that the cell lines with *NSD1* disrupted have a much higher hypomethylation peak than their respective parents. **C**, Same as A and B except for UM-SCC47. **D**, Bar plot of the above three cell lines showing the increase in the hypomethylation peak in the *NSD1* disrupted cell lines. *NSD1* alleles from monoclonal populations are characterized as follows: wt, wild type; trunc, contains a truncating mutation

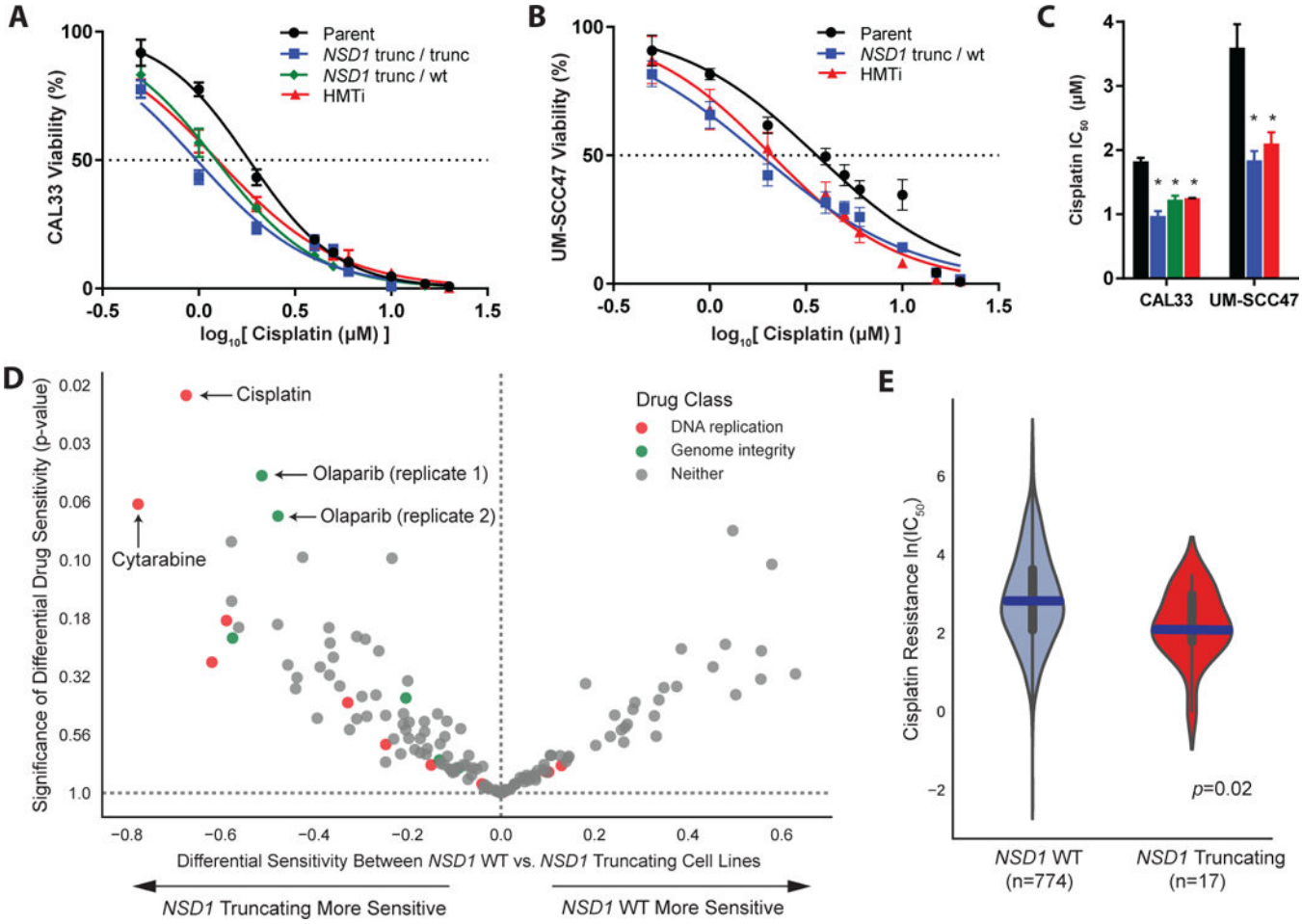

Figure 4.

NSD1 loss-of-function mutations confer increased cisplatin sensitivity. **A–B**, Cisplatin sensitivity curves for cell lines with and without *NSD1* disruption, showing greater sensitivity in the disrupted cell lines (blue and green lines). Pretreatment with the HMT inhibitor UNC0379 (HMTi) also increased sensitivity to cisplatin. *NSD1* alleles from monoclonal populations are characterized as follows: wt, wild type; trunc, contains a truncating mutation. **C**, Barplot of cisplatin IC_{50} in parental cell lines and cell lines with *NSD1* disrupted. Asterisk (*) indicates f sum-of-squares $p < 0.0001$ when compared to parental cell line. **D**, Volcano plot showing differential effect of 143 drugs on *NSD1* mutated versus *NSD1* wild type cell lines. Cisplatin is highly effective (2nd most left point) and the most significant (most upward point). The drug classes “DNA replication” and “Genome integrity” are highly represented on the *NSD1* sensitizing side. **E**, Violin plot showing increased sensitivity of *NSD1* mutated cell lines to cisplatin.