Published in final edited form as:

Oncogene. 2019 January ; 38(5): 671–686. doi:10.1038/s41388-018-0474-y.

An Activating Mutation of the NSD2 Histone Methyltransferase Drives Oncogenic Reprogramming in Acute Lymphocytic Leukemia

Alok Swaroop[#1,2], **Jon A. Oyer**[#1], **Christine M. Will**[1], **Xiaoxiao Huang**[1,3], **Wenbo Yu**[4], **Catalina Troche**[2], **Marinka Bulic**[1], **Benjamin H. Durham**[5], **Qiang Jeremy Wen**[1], **John D. Crispino**[1], **Alexander D. MacKerell Jr**[4], **Richard L. Bennett**[2], **Neil L. Kelleher**[3], and **Jonathan D. Licht**[1,2]

[1]Division of Hematology/Oncology, University of Florida Health Cancer Center, Gainesville, FL

[2]Division of Hematology/Oncology, Robert H. Lurie Comprehensive Cancer Center, Northwestern University Feinberg School of Medicine, Chicago, IL

[3]Department of Chemistry, Department of Molecular Biosciences, and the Chemistry of Life Processes Institute, Northwestern University, Evanston, IL

[4]Computer-Aided Drug Design Center, Department of Pharmaceutical Sciences, School of Pharmacy, University of Maryland, Baltimore

[5]Department of Pathology, Memorial Sloan Kettering Cancer Center

[#] These authors contributed equally to this work.

Abstract

NSD2, a histone methyltransferase specific for methylation of histone 3 lysine 36 (H3K36), exhibits a glutamic acid to lysine mutation at residue 1099 (E1099K) in childhood acute lymphocytic leukemia (ALL), and cells harboring this mutation can become the predominant clone in relapsing disease. We studied the effects of this mutant enzyme *in silico*, *in vitro*, and *in vivo* using gene edited cell lines. The E1099K mutation altered enzyme/substrate binding and enhanced the rate of H3K36 methylation. As a result, cell lines harboring E1099K exhibit increased H3K36 dimethylation and reduced H3K27 trimethylation, particularly on nucleosomes containing histone H3.1. Mutant NSD2 cells exhibit reduced apoptosis and enhanced proliferation, clonogenicity, adhesion, and migration. In mouse xenografts, mutant NSD2 cells are more lethal and brain invasive than wildtype cells. Transcriptional profiling demonstrates that mutant NSD2 aberrantly activates factors commonly associated with neural and stromal lineages in addition to signaling and adhesion genes. Identification of these pathways provides new avenues for therapeutic interventions in NSD2 dysregulated malignancies.

Corresponding author: Jonathan D. Licht, MD, The University of Florida Health Cancer Center, Cancer/Genetics Research Complex, 2033 Mowry Road, Suite 145, Gainesville, FL 32610, 352-273-8143, jdlicht@ufl.edu.

Conflict of Interest: The authors declare no conflict of interest

Keywords

leukemia; relapse; epigenetics; histone; methylation

INTRODUCTION

Acute lymphocytic leukemia (ALL) is the most common pediatric cancer in the United States (1). A disease of immature lymphocytes that overgrow their niche, ALL impedes normal hematopoiesis and infiltrates vital organs (2). While decades of laboratory and clinical studies have led to the development of multi-agent chemotherapy regimens that yield 5-year survival rates of nearly 90%, relapse is a serious concern, and ALL is one the commonest causes of cancer death in children (3). ALL pathogenesis is frequently associated with disruption of transcription factors that guide lymphocyte development, aberrant activation of key signaling pathways, and loss of tumor suppressor genes required for cell cycle control (4). Additionally, mutations in the genes encoding epigenetic regulators are often detected in ALL (4–7). Included among epigenetic regulators are histone methyltransferases (HMT), demethylases, acetyltransferases (HAT), and deacetylases (HDAC) that modify histone tails, and chromatin remodeling enzymes that reposition nucleosomes, altering accessibility of genes for transcription. Deregulated expression of such genes can have wide-ranging consequences, affecting expression of multiple genes and pathways (8).

Most of the mutations of chromatin regulators, including HMTs, in cancer yield loss of protein function (9). Gain-of-function mutations are rare and include recurrent point mutations in EZH2, the enzymatic component of the polycomb repressive complex 2 (PRC2), that enhance its ability to trimethylate histone 3 lysine 27 (H3K27me3) (10, 11). The other gain-of-function mutation of an HMT in malignancy occurs in NSD2 (MMSET, WHSC1), an epigenetic regulator with a role in hematopoietic development (12). Translocation (4;14), present in 10–20% of cases of multiple myeloma (MM), places the *NSD2* gene under the control of the immunoglobulin heavy chain promoter (13). The resulting *NSD2* overexpression leads to a genome-wide increase in the transcription-activating mark histone 3 lysine 36 dimethyl (H3K36me2), with a concurrent loss of the repressive mark, H3K27me3. The spread of H3K36me2 leads to aberrant gene activation and the promotion of cancer phenotypes (14–16).

A heterozygous mutation resulting in a glutamic acid to lysine change at position 1099 (E1099K) within the enzymatic SET domain of NSD2 was recently described in ALL cell lines (17, 18). Analysis revealed increased H3K36me2 and decreased H3K27me3 compared to lines without NSD2 mutation, and exogenous expression of the mutant protein increased the colony-forming ability of MM cells (18). Subsequently, recurrent E1099K mutations were identified in 5–10% of pediatric ALL patients (18–20). In one study the mutant allele was detected in a subclonal population at diagnosis but present in a majority of cells at relapse, suggesting that the mutation contributes to therapeutic resistance (20, 21). As such, the mutation may be even more prevalent in relapsed ALL. The E1099K mutation was also identified in chronic lymphocytic leukemia (CLL) (22), and, together with another mutation,

T1150A, in 12% of mantle cell lymphomas (MCL) (5). Collectively these findings suggest that the NSD2 mutations are oncogenic. However, the mechanism for increased activity of NSD2 E1099K as well as how this mutation contributes to the development or progression of lymphoid malignancy beyond simply increasing proliferation remains unclear.

Here, we characterize the enhanced enzymatic activity of mutant NSD2 through *in silico* and *in vitro* analysis and using CRISPR-edited cell lines study the biological, epigenetic, and transcriptional consequences of the E1099K mutation *in vivo*. We find that the NSD2 mutation activates a distinctively abnormal gene expression program that stimulates growth and other aggressive biological properties of ALL cells that may contribute to relapse. Activation of oncogenic transcriptional pathways in NSD2 mutant cells suggests numerous therapeutic approaches to counteract the mutant enzyme.

RESULTS

The E1099K mutation enhances NSD2 interaction with chromatin

To investigate how the E1099K mutation may alter the structure and function of NSD2, we used the structure of NSD1 (PDB ID:3OOI) (23) and the nucleosome core particle (NCP, PDB ID: 1KX5) to construct a model in which NSD2 interacts with the H3 N-terminus and nucleosome DNA, allowing for a molecular understanding of the stabilization of the complex during H3K36 methylation (Figure 1A-B). Our model shows significant similarity to a recently published crystal structure of NSD2 that was solved with multiple point mutations, substantiating the results from our analysis (24).

Molecular dynamics (MD) simulation predicted that the E1099K mutation makes the DNA-NSD2 interaction energy more favorable (Figure 1C, first row). The interaction of residue 1099 with other parts of the system was significantly augmented when glutamic acid is changed to lysine (Figure 1C, first row). The interaction energy of K1099 with the rest of the system was even more favorable when H3.3 was replaced with replication-associated histone variant 3.1 (Figure 1C, second row). H3.1 had a more stable interaction with DNA and NSD2 due to the presence of hydrophobic alanine instead of serine at residue 31 (Figure 1C, second row). The solvent-accessible surface area (SASA) of the NSD2 histone binding site surrounding H3K36 was predicted to be larger in the E1099K mutant allowing for more contact between H3K36 and the S-adenosylmethionine (SAM) binding site required for methylation (Figure 1D). The SASA was further increased in the H3.1 model (Figure 1D). These results predict increased potential for histone methylation by mutant NSD2 relative to wildtype and suggest that nucleosomes containing H3.1 might be an even better target for methylation by NSD2 E1099K than those containing H3.3.

To determine whether mutant NSD2 has altered affinity for chromatin *in vivo*, fluorescence recovery after photobleaching (FRAP) was performed on HEK293 cells transfected with GFP-tagged NSD2. Following photobleaching, cells expressing $NSD2^{E1099K}$ recovered fluorescence intensity at a slower rate than cells expressing $NSD2^{WT}$ (Figure 1E). Furthermore, the amount of unbound $NSD2^{E1099K}$ in the nucleus, as measured by mobile fraction, was less than that of $NSD2^{WT}$ (Figure 1E, inset). To further determine whether $NSD2^{E1099K}$ exhibits increased binding to nucleosomes, direct measurement of this

interaction was performed by quantifying wildtype and mutant NSD2 bound to nucleosomes in reactions containing increasing concentrations of salt (Figure 1F). While both wildtype NSD2 and NSD2^{E1099K} were bound to nucleosomes at 100mM NaCl, at 200mM NaCl wildtype NSD2 did not remain bound while mutant NSD2 maintained substantial interaction with nucleosomes (Figure 1F). Together these data indicate that mutant NSD2 has increased affinity for chromatin, which may underlie its enhanced activity.

Gene editing to remove the E1099K mutation restores the balance of H3K36me2 to H3K27me3

To elucidate the oncogenic effects of the E1099K mutation, we used CRISPR/Cas9 to generate isogenic cell lines differing only in the mutation status of NSD2. A single guide RNA (sgRNA) was designed to target the mutant allele of NSD2 (Supplementary Fig. S1A-B). Three genotypes were isolated from two B-ALL cell lines (SEM and RCH-ACV) and one T-ALL cell line (RPMI-8402). Control, non-targeted, cell lines were heterozygous for mutant NSD2 (NSD2$^{WT/E1099K}$) (Figure 2A, left). Two hemizygous wildtype (NSD2$^{WT/-}$) cell lines with frameshift mutations in the NSD2^{E1099K} allele were isolated from both SEM and RPMI-8402 cells (Figure 2A, right). NSD2$^{WT/-}$ RCH-ACV cells were isolated but difficult to maintain in culture and eventually lost. Finally, we isolated homozygous wildtype (NSD2$^{WT/WT}$) cells from RCH-ACV, which were presumably the result of interchromosomal gene conversion, as no donor oligonucleotide was included in the gene editing procedure (Figure 2A, middle).

NSD2$^{WT/E1099K}$ cells had globally high H3K36me2 and low H3K27me3, while NSD2$^{WT/-}$ or NSD2$^{WT/WT}$ had low H3K36me2 and high H3K27me3 (Figure 2B). The increase in H3K27me3 was not due to changes in EZH2 expression (Supplementary Fig. S1C). The increased activity of NSD2 E1099K was further characterized by mass spectrometry and modeling of histone methylation kinetics (M4K), which quantifies the rate of histone modification *in vivo*. Relative to RCH-ACV NSD2$^{WT/WT}$ cells, NSD2$^{WT/E1099K}$ cells formed H3K36me1 at a six-fold higher rate on histones devoid of H3K27 modification and an eight-fold greater rate on histones bearing a monomethylated H3K27 (Figure 2C). By contrast, the rate of conversion of H3K36me1 to H3K36me2 was similar in NSD2$^{WT/E1099K}$ RCH-ACV cells and in cells where the mutation was removed (Figure 2C; Supplementary Fig. S2A). In SEM NSD2$^{WT/E1099K}$ cells, the monomethylation rate was two-fold higher compared with NSD2$^{WT/-}$ cells on histones with unmodified H3K27 and three-fold increased when K27 was monomethylated (Supplementary Fig. S2B). The dimethylation rate was unchanged between mutant and wildtype cells. The overall result was a 2.5-fold increase in global levels of H3K36me2 and 5-fold decreased level of H3K27me3 in NSD2$^{WT/E1099K}$ cells as compared to NSD2$^{WT/WT}$ and NSD2$^{WT/-}$ (Fig. 2B). Of note, upon elimination of the mutant NSD2 allele there was an accumulation of unmethylated H3K36 without an increase of H3K36me1, consistent with *in vitro* data indicating NSD2 catalyzes the mono- and dimethylation of H3K36 (25). In analyzing histone variants H3.3 and H3.1, we detect high H3K36 dimethylation on histone H3.3 in NSD2$^{WT/E1099K}$ cells that remains elevated even upon elimination of the mutant allele (Supplementary Fig. S3A). In contrast, H3K36 dimethylation on histone H3.1 is high in NSD2$^{WT/E1099K}$ cells and significantly reduced in cells without the E1099K mutation (Supplementary Fig. S3B). This relationship

also holds true for NSD2-overexpressing cells compared to NSD2-underexpressing cells derived from the MM cell line, KMS11 (Supplementary Fig. S3C). In both ALL and MM, mutant NSD2 dimethylates K36 of H3.1 more than H3.3, thus facilitating the spread of H3K36me2 and inhibition of H3K27me3 across the epigenome.

E1099K mutation promotes cancer-associated phenotypes

Next, we used the edited cell lines to explore the role of the E1099K mutation in ALL biology. Removal of the E1099K mutation reduced the proliferation of RCH-ACV, SEM, and RPMI-8402 cells. After 8 days in culture, RCH-ACV NSD2$^{WT/E1099K}$ cells were 8.5 times more numerous than cells without the mutation, while there were 1.4 times as many NSD2$^{WT/E1099K}$ SEM and 1.7 times as many RPMI-8402 cells than wildtype (Figure 3A). Similarly, there were 9-, 2-, and 3-fold more colonies formed by RCH-ACV, SEM, and RPMI-8402 NSD2$^{WT/E1099K}$ cells as compared to cells without the mutation (Figure 3B). Correlating with the decreased growth in culture, RCH-ACV and SEM cells in which NSD2^{E1099K} was removed demonstrated a modest increase of cells in G1, fewer cells in S and G2/M, and an increase in spontaneous apoptosis demonstrated by annexin V staining and a sub-G1 population (Figure 3C, Supplementary Fig. S4A-C). The decreased growth of NSD2$^{WT/-}$ RPMI-8402 cells was attributable to a marked, 2-fold increase in apoptosis (Figure 3C, Supplementary Fig. S4C). Strikingly, the ability of the leukemia cells to adhere to stromal cells was significantly reduced upon removal of the E1099K mutation (Figure 3D), and RCH-ACV and SEM NSD2$^{WT/E1099K}$ cells migrated through Boyden chambers twice as effectively as cells without the mutation (Figure 3E). Thus, removal of the E1099K mutation resulted in the loss of many phenotypes associated with aggressive malignancy and disease progression.

NSD2-mutant cells behave aggressively *in vivo*

To examine the role of the E1099K mutation on tumor growth *in vivo*, luciferase-tagged mutant and CRISPR-edited ALL cell lines were injected by tail vein into NOD-SCID mice and monitored by *in vivo* imaging. Mice xenografted with RCH-ACV NSD2$^{WT/E1099K}$ cells had a median survival of 22 days, requiring sacrifice due to neurological impairment, while mice injected with NSD2$^{WT/WT}$ cells had a median survival of 27.5 days (Figure 4A, left panel). Similarly, mice injected with SEM NSD2$^{WT/E1099K}$ cells survived for 30 days compared to mice injected with two different clones of NSD2$^{WT/-}$ cells that had median survivals of 50 days and 59 days (Figure 4A, right panel). The total tumor burden detected by photon flux in mice injected with NSD2$^{WT/WT}$ or NSD2$^{WT/-}$ cells was 3-fold higher at time of death than those injected with NSD2$^{WT/E1099K}$ cells (Figure 4B), indicating that mice were able to tolerate wildtype cells better than mutant. Furthermore, while all cells were widely disseminated late in the disease course, NSD2$^{WT/E1099K}$ cells were detected in the head earlier and with more intensity than NSD2$^{WT/WT}$ or NSD2$^{WT/-}$ cells (Figure 4C-D), suggesting altered migration and response to brain microenvironment. Histological analysis of brain sections indicated leptomeningeal involvement of NSD2$^{WT/E1099K}$and NSD2$^{WT/-}$ SEM cells, but only NSD2$^{WT/E1099K}$ cells invaded the brain parenchyma (Figure 4E). These observations indicate that the E1099K mutation promotes cancer phenotypes *in vitro* and dictates aggressive biological behavior *in vivo*.

Widespread gene activation in E1099K cells drives oncogenic programs

In order to characterize the pathways that drive aggressive phenotypes associated with the E1099K mutation, we performed global transcriptome analysis of CRISPR-edited cell lines. Most of the genes differentially expressed in $NSD2^{WT/E1099K}$ cells relative to wildtype cells were upregulated (Figure 5A-B). Accordingly, there were 84 genes commonly over-expressed in $NSD2^{WT/E1099K}$ RCH-ACV, SEM, and RPMI-8402 cells, but only one gene commonly repressed (Figure 5C), suggesting that the primary function of the E1099K mutation in NSD2 is to activate transcription. Additionally, transcriptional activation in $NSD2^{WT/E1099K}$ cells converged on similar pathways among all three cell lines, while there were many fewer common downregulated pathways (Supplementary Fig. S5A-B). Gene set enrichment analysis (GSEA, Enrichr) revealed that E1099K upregulated genes overlapped with those regulated by repressive polycomb proteins EZH2 and CBX8 and displaying the H3K27me3 modification in mouse spleen and thymus (Supplementary Fig. S6). In fact, the E1099K upregulated genes tended to be under-expressed in normal B and T cells and over-expressed in stromal cells (Supplementary Fig. S7;www.immgen.org; (26)), suggesting that E1099K activated a gene program distinctly ectopic to the hematopoietic lineage. GSEA showed that genes upregulated in $NSD2^{E1099K/WT}$ cells were part of oncogenic pathways including cell adhesion, cytokine-receptor interaction, and signaling (Figure 6A). In order to assess how relevant the data from our CRISPR-edited cell lines are to ALL patients we compared it to datasets of genes upregulated in patient samples with NSD2 mutations (27). Analysis of microarray data from E2A-PBX1 ALL patients revealed 261 genes significantly upregulated in $NSD2^{WT/E1099K}$ samples (18, 28). These genes also tended to be activated in all three $NSD2^{WT/E1099K}$ cell lines (Figure 6B, row 1). From this same study we compared ETV6-RUNX1 ALL samples and identified 144 E1099K upregulated genes. The ETV6-RUNX1 gene-set was enriched in RCH-ACV and RPMI-8402 $NSD2^{WT/E1099K}$ cells (Figure 6B, row 2). Intriguingly, the data generated from the CRISPR-edited cell lines may be applicable beyond ALL. Genes differentially upregulated in mutant NSD2 MCL or t(4;14) MM were also enriched in $NSD2^{WT/E1099K}$ cell lines (Figure 6B, row 3–4; (5, 29)). The strong correlation between these gene expression data and patient data indicate our cell line models can recapitulate human disease and that similar pathways may be functioning downstream of dysregulated NSD2 in various malignancies.

DISCUSSION

Next generation sequencing has revealed the molecular basis of early relapse of ALL (19) and offers potential clues to prevent this outcome. The discovery of the activating E1099K mutation in NSD2 presents a potential therapeutic opportunity, as targeting EZH2, which harbors the only other activating mutation in a histone methyltransferase, has already shown promise in clinical trials (30, 31). Our finding that NSD2 mutation activates the enzymatic activity of the protein *in vivo* and activates transcriptional pathways that promote ALL progression nominates the development of NSD2 inhibitors as a priority. Furthermore, targeting pathways downstream of mutant NDS2 may offer additional therapeutic strategies.

Our model of the structure of NSD2 suggested that the E1099K mutation changes the binding of NSD2 to histones and DNA, providing insight that could be useful for the design

of specific drugs that attenuate the action of the mutant protein, while sparing the function of normal NSD2. The predicted increased stability of the NSD2-nucleosome complex and the larger SASA of the NSD2 histone-binding pocket that occur due to the E1099K mutation provide a mechanistic basis for increased methyltransferase activity. A pulldown assay with purified wildtype and mutant NSD2 confirmed improved binding of the mutant protein with nucleosomes. FRAP analysis suggested that this was also the case *in vivo*, as mutant NSD2 was not as mobile as wildtype. Interestingly, we found superior stability and a larger SASA of the mutant NSD2-nucleosome complex containing canonical, replication-dependent histone H3.1 rather than replacement histone H3.3. H3.1 is more associated with repressed genes and is commonly enriched in H3K27me3, while H3.3 is associated with active transcription and H3K36me2 (32, 33). Although further study is required to confirm the increased functionality of mutant NSD2 on nucleosomes containing H3.1 versus H3.3, mass spectroscopic analysis revealed that H3K36me2 was increased more on H3.1 rather than H3.3 when comparing $NSD2^{WT/E1099K}$ cells to cells without the mutation. This suggests that mutant NSD2 may be outcompeting the H3K27me3-depositing PRC2 complex at newly synthesized DNA, marking H3.1 with K36me2 and activating transcription of genes that were designated for silencing after DNA replication.

Overall, mass spectroscopic analysis of NSD2-mutant and isogenic wildtype cells showed that NSD2 mutation caused a 2.5-fold increase in H3K36me2 levels. Measurement of *in vivo* rates of histone methylation in NSD2-mutant versus wildtype cells showed that the rate of H3K36me0 to H3K36me1 was particularly accelerated, perhaps reflecting the greater *in vivo* binding affinity of mutant NSD2 for chromatin that lends it the ability to overcome a rate limiting step for methylation. By contrast, in the case of t(4;14) MM, sheer mass action of the overexpressed NSD2 may increase the interaction of NSD2 with chromatin, increasing the rates of conversion of H3K36me0 to H3K36me1 and also H3K36me1 to H3K36me2 (34). Both E1099K mutation and NSD2 overexpression result in a global change in chromatin: increasing H3K36me2 while decreasing H3K27me3, particularly at H3.1. H3K36me2, normally found along with the H3K4me3 modification at the 5’ end of genes near transcriptional start sites, is deposited through the action of NSD family HMTs, including NSD1, NSD2, and NSD3, as well as ASH1L (16, 23, 25, 35–37). The H3K36me2 and H3K4me3 modifications are antagonistic to PRC2 and EZH2 activity while EZH2 activity and H3K27me3 inhibits the action of NSD2 and related enzymes (34, 38–40). Thus, high-level NSD2 activity inhibits EZH2/polycomb action on specific loci, resulting mainly in activation of genes and pathways that change the properties of cancer cells. This was reflected by our finding that the genes consistently upregulated by NSD2 were enriched for previously identified PRC2 targets. Neither NSD2 overexpression nor E1099K mutation has the effect of increasing the rate of H3K36 trimethylation, a modification found within gene bodies and towards the 3’ end of genes, which is deposited by the SETD2 enzyme and serves to prevent adventitious initiation of transcription within genes (41).

This work presents the clearest picture of biological and transcriptional changes that occur due to the oncogenic point mutation $NSD2^{E1099K}$. Past work compared cell lines harboring the mutation versus those without, or reduced NSD2 expression less specifically with shRNAs or sgRNAs (17–19). Here, we generated wildtype NSD2 isogenic cell lines derived from 3 different parent cell lines harboring the E1099K mutation. In these cells, the

NSD2^{E1099K} mutation reduced apoptosis and enhanced proliferation, clonogenicity, adhesion, and migration, all hallmarks of cancer that are accentuated with leukemia progression (42). Knockdown of NSD2 in MM also results in impaired adhesion and proliferation (15, 43). Pathway analysis of RNA-seq from NSD2$^{WT/E1099K}$ and NSD2 wildtype cells indicated that genes involved in cell adhesion were prominently upregulated by NSD2^{E1099K}. Adhesion proteins, including neural cell adhesion molecule 1 (*NCAM1, CD56*) and desmoglein 2 (*DSG2*), often activate downstream signaling pathways (44, 45). NCAM1 is expressed on neurons, epithelial cells, muscle cells, and pancreatic β cells, but its role in neuronal development and plasticity is most well defined (44). In MM, NCAM1 and DSG2 are both upregulated by NSD2, and their expression is implicated in adhesion and proliferation (46, 47). Disrupting the activity of these genes could have cascading effects on NSD2^{E1099K} ALL and MM beyond simply reducing adhesion.

The aggressive biology of NSD2$^{WT/E1099K}$ cells was evident *in vivo,* as mice xenografted with mutant cells died significantly faster than those grafted with wildtype cells. Enhanced proliferation and clonogenicity of mutant cells were likely not primary reasons for increased lethality, however. NSD2$^{WT/E1099K}$ cells also exhibited earlier invasion of the central nervous system, a common site of relapse in pediatric ALL (48). Interestingly, the mutant cells also invade the brain parenchyma, a clinically rare but significant finding, which is likely the major reason for their increased lethality because general tumor burden was not predictive of death. Children with ALL get CNS disease in 10% of cases, but prior to widespread CNS prophylaxis, it was between 50% and 75% (49). Expression of *NCAM1* in ALL correlates with increased likelihood of CNS invasion, and was suggested as a biomarker to identify patients needing aggressive CNS therapy (50). As mentioned, *NCAM1* was upregulated in all three E1099K cell lines along with other genes involved in neural processes. The activation of such a transcriptional program might contribute to the ability of mutant NSD2 cells to adhere and proliferate within the CNS microenvironment. Studying the role of *NCAM1* and other downstream effectors of NSD2^{E1099K} may provide additional possibilities for therapy.

Mutant NSD2 plays a crucial role in progression of ALL and has significant potential for therapeutic inhibition. Our modeling and analysis of the differential binding between wildtype and mutant NSD2 with nucleosomes and DNA suggests a mechanism by which the E1099K mutation hyperactivates NSD2 and lays out a path for inhibition of mutant NSD2 that involves disrupting its altered binding to chromatin. Specific targeting of mutant NSD2 binding to H3.1-containing nucleosomes may prove to be particularly effective. The significant biological effects of NSD2^{E1099K} are explained by spurious activation of numerous genes and pathways, which reveal more avenues for therapy. Further study of NSD2 target genes should yield other insights into leukemia progression, relapse, and therapy.

Methods

Modeling

The SWISS-MODEL server (51) was employed to construct a NSD2 homology model based on the NSD1 crystal structure. Hydrogen atoms were added to the structure using REDUCE

software (52). The system was minimized for 1,000 steepest descent steps followed by 500 adopted basis Newton-Raphson steps with harmonic restraints of 1 kcal/(mol Å). We overlaid the structure of H3K9 peptide from the structure of the homologous GLP-H3K9 complex. The H3-NSD2 complex was modeled using Langevin MD simulation in combination with the ZDOCK server (53). Three MD runs were conducted using the CHARMM program (54) with CHARMM C36 protein (55, 56) and nucleic acid (57) force fields. The Generalized Born-Molecular Volume method (58) accounted for solvation effects. Leapfrog (59) integrated the equations of motion. The SHAKE algorithm (60) was applied to constrain length of covalent bonds involving hydrogen to their equilibrium values.

Cell culture

SEM cells (DSMZ, ACC-546) were cultured in IMDM/10% HI-FBS/1% penicillin-streptomycin. RPMI-8402 (ATCC, CRL-1994) and RCH-ACV cells (DSMZ, ACC-548) were cultured in RPMI-1640/10% HI-FBS/1% penicillin-streptomycin. Cells and subclones were authenticated by STR profiling (Genetica). HS-5 (ATCC, CRL-11882) and 293T cells were cultured in DMEM/10% HI-FBS/1% penicillin-streptomycin at 37°C with 5% CO_2. Routine mycoplasma testing was performed using the MycoAlert Detection Kit (Lonza).

Plasmids, mutagenesis, and FRAP

NSD2 cDNA was PCR-amplified and inserted into Mfe-I digested pCAG-GFP (Addgene, #11150, (61)) using HiFi DNA Assembly (NEB). HEK 293T were transfected using FuGENE 6 (Promega). Cells in heated and CO_2-regulated chamber were imaged using the 63x objective on Nikon A1R+ confocal system. 8–12 bright, uniform nuclei were imaged per experiment, for a total of 6 experiments, and analyzed in GraphPad v6. Imaging and mobile fraction calculation performed as described previously (62). Imaging work was performed at the Northwestern University Center for Advanced Microscopy supported by NCI CCSG P30 CA060553. Mobile fraction plotted as mean with error bars equal representing SEM. Two-tailed t-test with unequal variance p=0.0006. In this experiment, as in those described below, the more stringent two-tailed t-test with unequal variance was chosen to preclude assumptions regarding the data and due to smaller sample sizes than would allow for normal distribution of data.

NSD2-nucleosome association assay

Recombinant NSD2 (Cat #HMT-21-122, lot #1963), E1099K-NSD2 (Cat #HMT-21-159, lot #1520) and biotinylated oligonucleosomes purified from HeLa cells (Cat #HMT-35–160, lot#1542) were purchased from Reaction Biology Corp. (Malvern, PA). Briefly, 0.25μM of recombinant enzymes and biotinylated nucleosome were incubated for 15 minutes at 30C in a 30ul reaction containing 50mM TrisCl pH 8.5, 5mM $MgCl_2$, 5uM SAM, 2μM TCEP, 1% BSA, and either 100mM, 200mM, 300mM, or 400mM NaCl in siliconized tubes. Following incubation, 10% of the reaction was kept as the input fraction. Reactions were diluted to 300ul in reaction buffer with 0.5% NP-40 and 10μl of streptavidin magnetic beads (Pierce Cat#88817). Reactions were incubated at 4C for 1 hour and nucleosome associated protein was isolated by magnetic bead capture. Complexes were washed three times in reaction buffer with 0.5% NP-40 and the appropriate salt concentration. Input and bound fractions were resolved by SDS-PAGE, and immunoblotting with antibodies to NSD2 (Abcam, Cat#

ab75359) and H3 (Cell signaling Cat# 14269S) was performed to evaluate the fraction of NSD2 bound. Immunoblot was quantified by ImageJ densitometry analysis.

CRISPR Generation of NSD2-Edited Cells

A double-stranded DNA fragment (IDT gBlock) was used to generate a customized sgRNA expression construct composed of a U6 promoter, sequence targeting the K1099 allele of *NSD2* (TGGGGAGCTGATCGACAAGG), and an sgRNA scaffold described previously (63). This construct was PCR-amplified, column-purified (Qiagen), and transfected with Cas9-GFP (Addgene #42234, (64)) using the Neon Transfection System (Life Technologies). GFP-positive cells were sorted (BD FACSAria) 20 hours post-transfection, plated in methylcellulose (MethoCult H4100, Stem Cell Technologies), and single-cell derived colonies isolated. DNA extracted from colonies (QuickExtract, Epicentre) was amplified and sequenced to detect CRISPR-editing events. PCR products were cloned and sequenced to assess mutant allele frequency (PCR Cloning Kit, NEB).

Immunoblot

Nuclear extracts were prepared using the Nuclear Complex Co-IP kit (Active Motif). Proteins were separated using SDS-PAGE, transferred to PVDF membranes, and probed with antibodies to EZH2 (5246, Millipore), HDAC2 (05–814, Millipore), NSD2 (39), and histone H4 (2935, Cell Signaling Tech.). Secondary antibodies were HRP-conjugated anti-mouse IgG (95017–332, VWR) or donkey anti-rabbit IgG (95017–556, VWR).

Quantitation of histone modifications by targeted mass spectrometry

Acid extracted histones were digested with trypsin and chemically derivatized using propionic anhydride as described (65). Histone peptides were analyzed by nanoLC-QqQ mass spectrometer (Dionex UltiMate 3000 and ThermoFisher TSQ Quantum) using a selected reaction monitoring method (34). Data were analyzed using Skyline software (66). The relative levels of specific modified peptides were calculated by dividing peak area against the total of all modified peptides sharing the same sequence. Individual H3K27 and H3K36 methylation levels were aggregated from levels of 16 combinatorial methylation species found in histone H3 peptide from residue 27 to 40 (KSAPATGGVKKPHR). 2 biological experiments performed per cell line with mean and error bars representing SD.

M4K

Cells were cultured as described above, then switched to media with [$^{13}C_6$]Arg and [$^{13}C_1$,$^{2}H_3$]Met (Cambridge Isotope Laboratories) and 10% dialyzed FBS (Sigma F0392). Cells were harvested at 0, 5, 10 and 25 hours after the switch. Acid-extracted histones were derivatized, and digested with trypsin prior to LC-MS. Selected reaction monitoring transitions were developed and data were analyzed using Skyline (66). The measured relative abundance of each species was used in a kinetic model (67).

Lentivirus production and infection of leukemia cell lines

293T cells were transfected with luciferase vector (pFU-L2T; gift of Dr. Marcus Peter, Northwestern University) and lentiviral packaging plasmids (psPAX2 and VSVG) using

FuGENE 6 (Promega). Viral supernatant was collected and passed through 0.45um filter. Cells were infected with spinoculation at 2000 RPM for 90 minutes and sorted by flow cytometry using the fluorescent Tomato protein.

In Vivo Tumorigenicity

Power analysis showed that having 19 mice per experimental group would give an 80% chance to detect an extension in survival from median of 1 month to 2 months with $P<0.05$ (http://stattools.crab.org). However, experiments were ended early as a statistically significant difference in survival was identified before reaching this number. 12–18-week-old female NOD-SCID mice were obtained from Jackson Laboratory (Bar Harbor, ME) and randomly numbered by an independent researcher. The numbers were then separated into experimental groups and 5×10^6 luciferase-expressing ALL cells into the tail veins of these mice. A total of 24 mice were injected with RCH-ACV cells and 51 mice were injected with SEM cells. Tumor growth was monitored by bioluminescent imaging (IVIS Spectrum, Xenogen). 15 RCH-ACV-injected mice were not included in bioluminescence analysis due to rapid progression of disease requiring euthanasia before imaging could be performed. Mice were euthanized when tumor burden became excessive. No blinding was performed. The log-rank test was used for analysis of survival curves. Total bioluminescence was plotted as log-transformed mean flux for surviving mice with error bars representing SEM. Head/body bioluminescence ratio plotted as mean with SEM and two-tailed t-test with unequal variance for statistical analysis. Experiments were conducted in ethical, pathogen-free conditions according to protocols approved by Northwestern University (#IS00000557) and University of Florida (#201509176) Animal Care and Use Committees. Organs were fixed in neutral buffered formalin embedded in paraffin, sectioned, and stained with hematoxylin and eosin or hHLA-ABC by the University of Florida Molecular Pathology Core.

Proliferation

ALL cells were plated at 1000 cells/well in 96-well black plates (BD Falcon) in sextuplicate. Cell mass was measured using CellTiter-Glo (Promega) and a BMG plate-reader at each time point. Results shown are the mean of 3 independent experiments for time points between 0–6 days and 2 experiments for day 8 with error bars representing SEM. Two-tailed t-tests with equal variance were performed and the raw p-values from the t-test were adjusted for multiplicity using Holm's method (68). Analysis was performed by the Biostatistics & Quantitative Sciences Shared Resource at the University of Florida.

Colony formation

Methylcellulose-based medium (Methocult H4100, StemCell Technologies) was prepared using appropriate media for each cell line. Cells were resuspended in methylcellulose medium at 4000/ml and 1ml media/well was added to a 12-well plate. After 7 days colonies were counted in 5 fields and averaged. Results shown are the mean of 3 independent experiments with error bars representing SEM and statistical analysis by two-tailed t-test with unequal variance.

Flow Cytometry

ALL cells were stained using a FITC-Annexin V kit (BioLegend) or were fixed in 70% ethanol and stained with FxCycle PI/RNase Staining Solution (ThermoFisher). Analysis was on an Accuri C6. Results shown are the mean of 3 independent experiments with error bars representing SD and statistical analysis by two-tailed t-test with unequal variance.

Adhesion

HS-5 cells were plated at 1500/well of a 96-well black plate for 24 hours. ALL cells were stained with 10uM CellTracker Green CMFDA (Life Technologies) and plated in triplicate on HS-5 cells at 5000/well, spun at 1200 rpm for 5 minutes and allowed to adhere for 24 hours. Wells were gently washed twice with PBS, and a plate reader (BMG) was used to measure fluorescence. Results shown are the mean of 4 independent experiments with error bars representing SEM and statistical analysis by two-tailed t-test with unequal variance.

Migration

Cells were suspended in serum-free media and added to a Boyden chamber (VWR 29442–118) at 2×10^5 per well, while serum-containing media was added to bottom chamber. After 24 hours cells in bottom chamber were quantified by flow cytometry. Results shown are the mean of 3 (RCH-ACV) or 4 (SEM) independent experiments with error bars representing SEM and statistical analysis by two-tailed t-test with unequal variance.

RNA-seq

RNA was extracted with the RNeasy Plus Mini kit (Qiagen). Library preparation and sequencing were performed by the Epigenomics Core Facility of Weill Cornell Medicine using TruSeq RNA Library preparation kit (Illumina) and HiSeq 2500 system (Illumina). Short reads were filtered and trimmed using Trimmomatic (v.0.36) and aligned to the hg38 reference genome using STAR (v.2.5.2a). RNA-seq data was deposited in the Gene Expression Omnibus database, reference number SRP135822. Expression quantification and differential expression analysis were performed using RSEM (v.1.2.31), using log2 (fold change) limit of 1 and p-value<5%. GSEA was performed using the Enrichr web server (69, 70) or Broad GSEA software (27). The significance of the intersections between sets of genes was determined by random sampling. 1000 sets of the same size as the input ones containing random genes were generated and the fraction of cases in which the intersection was equal to or larger than the observed one was computed. This fraction represents the P value of the observed intersection. For RNA-seq validation, cDNA was synthesized with iScript cDNA synthesis kit (BioRad). Quantitative real time PCR was performed using SYBR green (Roche) on a Lightcycler 480 II (Roche). Data plotted as mean of two independent experiments with SEM.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

Financial Support: Supported by T32 CA009560 (AS), F30 CA203292 (AS), R01 CA195732 (JDL), R01 GM051501 (ADM), a Leukemia and Lymphoma Society Specialized Center of Excellence grant (JDL), the Samuel Waxman Cancer Research Foundation (ADM and JDL) and Celgene (JDL), P41 GM108569 (NLK) and the Lauri Strauss Leukemia Foundation (JAO, CT).

References

1. Society AC. Cancer Facts & Figures 2014. Atlanta: American Cancer Society 2014.
2. Golub TR. Genomics: global views of leukaemia. Nature. 2007;446:739–40.
3. Society AC. Cancer Facts & Figures 2016. Atlanta: American Cancer Society 2016.
4. Mullighan CG. Genome sequencing of lymphoid malignancies. Blood. 2013;122:3899–907. [PubMed: 24041576]
5. Beà S, Valdés-Mas R, Navarro A, Salaverria I, Martín-Garcia D, Jares P, et al. Landscape of somatic mutations and clonal evolution in mantle cell lymphoma. Proceedings of the National Academy of Sciences of the United States of America. 2013;110:18250–5. [PubMed: 24145436]
6. Lafave LM, Levine RL. Mining the epigenetic landscape in ALL. Nature genetics. 2013;45:1269–70. [PubMed: 24165727]
7. Lawrence MS, Stojanov P, Mermel CH, Robinson JT, Garraway LA, Golub TR, et al. Discovery and saturation analysis of cancer genes across 21 tumour types. Nature. 2014;505(7484):495–501. [PubMed: 24390350]
8. Berdasco M, Esteller M. Aberrant epigenetic landscape in cancer: how cellular identity goes awry. Developmental cell. 2010;19:698–711. [PubMed: 21074720]
9. Van Rechem C, Whetstine JR. Examining the impact of gene variants on histone lysine methylation. Biochim Biophys Acta. 2014;1839(12):1463–76. [PubMed: 24859469]
10. Sneeringer CJ, Scott MP, Kuntz KW, Knutson SK, Pollock RM, Richon VM, et al. Coordinated activities of wild-type plus mutant EZH2 drive tumor-associated hypertrimethylation of lysine 27 on histone H3 (H3K27) in human B-cell lymphomas. Proc Natl Acad Sci U S A. 2010;107(49): 20980–5. [PubMed: 21078963]
11. Morin RD, Johnson NA, Severson TM, Mungall AJ, An J, Goya R, et al. Somatic mutations altering EZH2 (Tyr641) in follicular and diffuse large B-cell lymphomas of germinal-center origin. Nat Genet. 2010;42(2):181–5. [PubMed: 20081860]
12. Campos-Sanchez E, Deleyto-Seldas N, Dominguez V, Carrillo-de-Santa-Pau E, Ura K, Rocha PP, et al. Wolf-Hirschhorn Syndrome Candidate 1 Is Necessary for Correct Hematopoietic and B Cell Development. Cell Rep. 2017;19(8):1586–601. [PubMed: 28538178]
13. Keats JJ, Maxwell Ca, Taylor BJ, Hendzel MJ, Chesi M, Bergsagel PL, et al. Overexpression of transcripts originating from the MMSET locus characterizes all t(4;14)(p16;q32)-positive multiple myeloma patients. Blood. 2005;105:4060–9. [PubMed: 15677557]
14. Popovic R, Martinez-Garcia E, Giannopoulou EG, Zhang Q, Zhang Q, Ezponda T, et al. Histone methyltransferase MMSET/NSD2 alters EZH2 binding and reprograms the myeloma epigenome through global and focal changes in H3K36 and H3K27 methylation. PLoS genetics. 2014;10:e1004566. [PubMed: 25188243]
15. Lauring J, Abukhdeir AM, Konishi H, Garay JP, Gustin JP, Wang Q, et al. The multiple myeloma associated MMSET gene contributes to cellular adhesion, clonogenic growth, and tumorigenicity. Blood. 2008;111:856–64. [PubMed: 17942756]
16. Kuo AJ, Cheung P, Chen K, Zee BM, Kioi M, Lauring J, et al. NSD2 Links Dimethylation of Histone H3 at Lysine 36 to Oncogenic Programming. Molecular Cell. 2011;44:609–20. [PubMed: 22099308]
17. Ja Oyer, Huang X, Zheng Y, Shim J, Ezponda T, Carpenter Z, et al. Point mutation E1099K in MMSET/NSD2 enhances its methyltranferase activity and leads to altered global chromatin methylation in lymphoid malignancies. Leukemia. 2014;28:198–201. [PubMed: 23823660]

18. Jaffe JD, Wang Y, Chan HM, Zhang J, Huether R, Kryukov GV, et al. Global chromatin profiling reveals NSD2 mutations in pediatric acute lymphoblastic leukemia. Nature genetics. 2013;45:1386–91. [PubMed: 24076604]

19. Ding LW, Sun QY, Tan KT, Chien W, Thippeswamy AM, Eng Juh Yeoh A, et al. Mutational Landscape of Pediatric Acute Lymphoblastic Leukemia. Cancer Res. 2017;77(2):390–400. [PubMed: 27872090]

20. Loh ML, Ma X, Rusch M, Wu G, Harvey RC, Wheeler DA, et al. Comparison Of Mutational Profiles Of Diagnosis and Relapsed Pediatric B-Acute Lymphoblastic Leukemia: A Report From The COG ALL Target Project. Blood. 2013;122:824.

21. Ma X, Edmonson M, Yergeau D, Muzny DM, Hampton Oa, Rusch M, et al. Rise and fall of subclones from diagnosis to relapse in pediatric B-acute lymphoblastic leukaemia. Nature communications. 2015;6:6604.

22. Fabbri G, Rasi S, Rossi D, Trifonov V, Khiabanian H, Ma J, et al. Analysis of the chronic lymphocytic leukemia coding genome: role of NOTCH1 mutational activation. J Exp Med. 2011;208(7):1389–401. [PubMed: 21670202]

23. Qiao Q, Li Y, Chen Z, Wang M, Reinberg D, Xu RM. The structure of NSD1 reveals an autoregulatory mechanism underlying histone H3K36 methylation. J Biol Chem. 2011;286(10): 8361–8. [PubMed: 21196496]

24. Tisi D, Chiarparin E, Tamanini E, Pathuri P, Coyle JE, Hold A, et al. Structure of the Epigenetic Oncogene MMSET and Inhibition by N-Alkyl Sinefungin Derivatives. ACS Chem Biol. 2016;11(11):3093–105. [PubMed: 27571355]

25. Li Y, Trojer P, Xu C-F, Cheung P, Kuo A, Drury WJ, et al. The target of the NSD family of histone lysine methyltransferases depends on the nature of the substrate. The Journal of biological chemistry. 2009;284:34283–95. [PubMed: 19808676]

26. Heng TS, Painter MW, Immunological Genome Project C. The Immunological Genome Project: networks of gene expression in immune cells. Nat Immunol. 2008;9(10):1091–4. [PubMed: 18800157]

27. Subramanian A, Kuehn H, Gould J, Tamayo P, Mesirov JP. GSEA-P: a desktop application for Gene Set Enrichment Analysis. Bioinformatics. 2007;23(23):3251–3. [PubMed: 17644558]

28. Zhang J, Ding L, Holmfeldt L, Wu G, Heatley SL, Payne-Turner D, et al. The genetic basis of early T-cell precursor acute lymphoblastic leukaemia. Nature. 2012;481(7380):157–63. [PubMed: 22237106]

29. Broyl A, Hose D, Lokhorst H, de Knegt Y, Peeters J, Jauch A, et al. Gene expression profiling for molecular classification of multiple myeloma in newly diagnosed patients. Blood. 2010;116(14): 2543–53. [PubMed: 20574050]

30. McCabe MT, Ott HM, Ganji G, Korenchuk S, Thompson C, Van Aller GS, et al. EZH2 inhibition as a therapeutic strategy for lymphoma with EZH2-activating mutations. Nature. 2012;492(7427): 108–12. [PubMed: 23051747]

31. Kim KH, Roberts CW. Targeting EZH2 in cancer. Nat Med. 2016;22(2):128–34. [PubMed: 26845405]

32. Szenker E, Ray-Gallet D, Almouzni G. The double face of the histone variant H3.3. Cell Research. 2011;21(3):421–34. [PubMed: 21263457]

33. Hake SB, Allis CD. Histone H3 variants and their potential role in indexing mammalian genomes: the “H3 barcode hypothesis”. Proc Natl Acad Sci U S A. 2006;103(17):6428–35. [PubMed: 16571659]

34. Zheng Y, Sweet SMM, Popovic R, Martinez-Garcia E, Tipton JD, Thomas PM, et al. Total kinetic analysis reveals how combinatorial methylation patterns are established on lysines 27 and 36 of histone H3. Proceedings of the National Academy of Sciences of the United States of America. 2012;109:13549–54. [PubMed: 22869745]

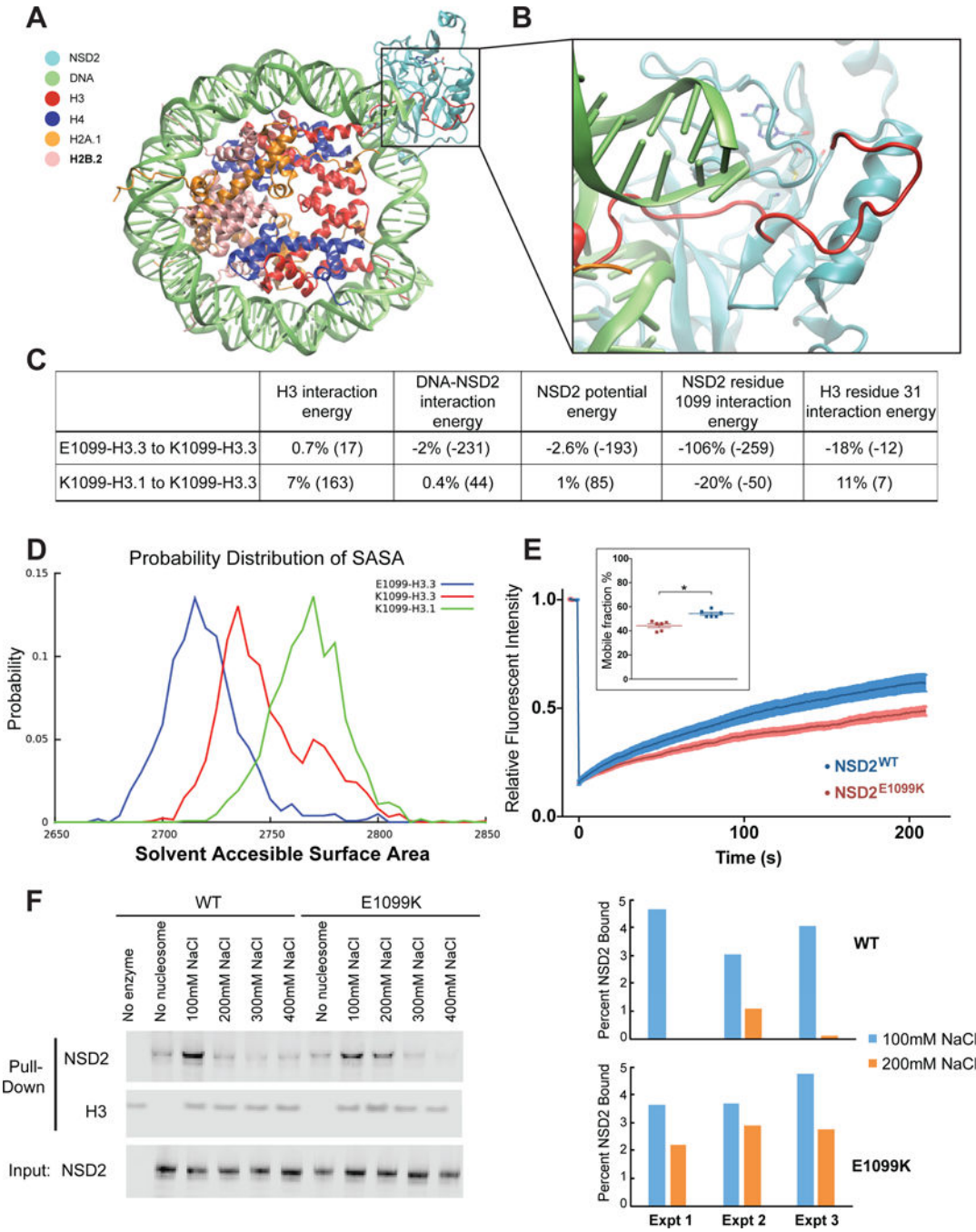
35. Kim SM, Kee HJ, Eom GH, Choe NW, Kim JY, Kim YS, et al. Characterization of a novel WHSC1-associated SET domain protein with H3K4 and H3K27 methyltransferase activity. Biochemical and Biophysical Research Communications. 2006;345:318–23. [PubMed: 16682010]

36. Lucio-Eterovic AK, Singh MM, Gardner JE, Veerappan CS, Rice JC, Carpenter PB. Role for the nuclear receptor-binding SET domain protein 1 (NSD1) methyltransferase in coordinating lysine

36 methylation at histone 3 with RNA polymerase II function. Proceedings of the National Academy of Sciences of the United States of America. 2010;107:16952–7. [PubMed: 20837538]

37. Zhu L, Li Q, Wong SH, Huang M, Klein BJ, Shen J, et al. ASH1L Links Histone H3 Lysine 36 Dimethylation to MLL Leukemia. Cancer Discov. 2016;6(7):770–83. [PubMed: 27154821]
38. Schmitges FW, Prusty AB, Faty M, Stutzer A, Lingaraju GM, Aiwazian J, et al. Histone methylation by PRC2 is inhibited by active chromatin marks. Mol Cell. 2011;42(3):330–41. [PubMed: 21549310]
39. Martinez-Garcia E, Popovic R, Min DJ, Sweet SMM, Thomas PM, Zamdborg L, et al. The MMSET histone methyl transferase switches global histone methylation and alters gene expression in t(4;14) multiple myeloma cells. Blood. 2011;117:211–20. [PubMed: 20974671]
40. Ezponda T, Popovic R, Shah MY, Martinez-Garcia E, Zheng Y, Min D-J, et al. The histone methyltransferase MMSET/WHSC1 activates TWIST1 to promote an epithelial-mesenchymal transition and invasive properties of prostate cancer. Oncogene. 2013;32:2882–90. [PubMed: 22797064]
41. Wagner EJ, Carpenter PB. Understanding the language of Lys36 methylation at histone H3. Nature Reviews Molecular Cell Biology. 2012;13:115–26. [PubMed: 22266761]
42. Hanahan D, Weinberg Ra. Hallmarks of cancer: the next generation. Cell. 2011;144:646–74. [PubMed: 21376230]
43. Huang Z, Wu H, Chuai S, Xu F, Yan F, Englund N, et al. NSD2 Is recruited through Its PHD domain to oncogenic gene loci to drive multiple myeloma. Cancer Research. 2013;73:6277–88. [PubMed: 23980095]
44. Cavallaro U, Christofori G. Cell adhesion and signalling by cadherins and Ig-CAMs in cancer. Nat Rev Cancer. 2004;4(2):118–32. [PubMed: 14964308]
45. Brennan D, Hu Y, Joubeh S, Choi YW, Whitaker-Menezes D, O'Brien T, et al. Suprabasal Dsg2 expression in transgenic mouse skin confers a hyperproliferative and apoptosis-resistant phenotype to keratinocytes. J Cell Sci. 2007;120(Pt 5):758–71. [PubMed: 17284515]
46. Brito JLR, Walker B, Jenner M, Dickens NJ, Brown NJM, Ross FM, et al. MMSET deregulation affects cell cycle progression and adhesion regulons in t(4;14) myeloma plasma cells. Haematologica. 2009;94:78–86. [PubMed: 19059936]
47. Damgaard T, Knudsen LM, Dahl IMS, Gimsing P, Lodahl M, Rasmussen T. Regulation of the CD56 promoter and its association with proliferation, anti-apoptosis and clinical factors in multiple myeloma. Leukemia & lymphoma. 2009;50:236–46. [PubMed: 19235015]
48. Wynn RF. Acute Lymphoblastic Leukemia. Pediatric Hematology and Oncology: Wiley-Blackwell; 2010 p. 75–94.
49. Estey EH, Faderl SH, Kantarjian H. Hematologic malignancies : acute leukemias. Berlin ; New York: Springer; 2008 x, 294 p. p.
50. Ravandi F, Cortes J, Estrov Z, Thomas D, Giles FJ, Huh YO, et al. CD56 expression predicts occurrence of CNS disease in acute lymphoblastic leukemia. Leuk Res. 2002;26(7):643–9. [PubMed: 12008081]
51. Schwede T, Kopp J, Guex N, Peitsch MC. SWISS-MODEL: An automated protein homology-modeling server. Nucleic Acids Res. 2003;31(13):3381–5. [PubMed: 12824332]
52. Word JM, Lovell SC, Richardson JS, Richardson DC. Asparagine and glutamine: using hydrogen atom contacts in the choice of side-chain amide orientation. J Mol Biol. 1999;285(4):1735–47. [PubMed: 9917408]
53. Pierce BG, Wiehe K, Hwang H, Kim BH, Vreven T, Weng Z. ZDOCK server: interactive docking prediction of protein-protein complexes and symmetric multimers. Bioinformatics. 2014;30(12): 1771–3. [PubMed: 24532726]
54. Brooks BR, Brooks CL, 3rd, Mackerell AD, Jr, Nilsson L, Petrella RJ, Roux B, et al. CHARMM: the biomolecular simulation program. J Comput Chem. 2009;30(10):1545–614. [PubMed: 19444816]
55. MacKerell AD, Bashford D, Bellott M, Dunbrack RL, Evanseck JD, Field MJ, et al. All-atom empirical potential for molecular modeling and dynamics studies of proteins. J Phys Chem B. 1998;102(18):3586–616. [PubMed: 24889800]

56. Best RB, Zhu X, Shim J, Lopes PE, Mittal J, Feig M, et al. Optimization of the additive CHARMM all-atom protein force field targeting improved sampling of the backbone phi, psi and side-chain chi(1) and chi(2) dihedral angles. J Chem Theory Comput. 2012;8(9):3257–73. [PubMed: 23341755]
57. Hart K, Foloppe N, Baker CM, Denning EJ, Nilsson L, Mackerell AD, Jr. Optimization of the CHARMM additive force field for DNA: Improved treatment of the BI/BII conformational equilibrium. J Chem Theory Comput. 2012;8(1):348–62. [PubMed: 22368531]
58. Lee MSS FR, Jr.; Brooks III CL. Novel generalized Born methods. J Chem Phys. 2002;116(10606).
59. Hockney RW. Methods in Computational Physics. New York: Academic Press; 1970.
60. Ryckaert J-PC G; Berendsen HJC Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes. J Comput Phys. 1977;23:327–41.
61. Matsuda T, Cepko CL. Electroporation and RNA interference in the rodent retina in vivo and in vitro. Proc Natl Acad Sci U S A. 2004;101(1):16–22. [PubMed: 14603031]
62. Mueller F, Karpova TS, Mazza D, McNally JG. Monitoring dynamic binding of chromatin proteins in vivo by fluorescence recovery after photobleaching. Methods Mol Biol. 2012;833:153–76. [PubMed: 22183594]
63. Chen B, Gilbert LA, Cimini BA, Schnitzbauer J, Zhang W, Li GW, et al. Dynamic imaging of genomic loci in living human cells by an optimized CRISPR/Cas system. Cell. 2013;155(7):1479–91. [PubMed: 24360272]
64. Jinek M, East A, Cheng A, Lin S, Ma E, Doudna J. RNA-programmed genome editing in human cells. Elife. 2013;2:e00471. [PubMed: 23386978]
65. Garcia BA, Mollah S, Ueberheide BM, Busby SA, Muratore TL, Shabanowitz J, et al. Chemical derivatization of histones for facilitated analysis by mass spectrometry. Nat Protoc. 2007;2(4):933–8. [PubMed: 17446892]
66. MacLean B, Tomazela DM, Shulman N, Chambers M, Finney GL, Frewen B, et al. Skyline: an open source document editor for creating and analyzing targeted proteomics experiments. Bioinformatics. 2010;26(7):966–8. [PubMed: 20147306]
67. Zheng Y, Tipton JD, Thomas PM, Kelleher NL, Sweet SM. Site-specific human histone H3 methylation stability: fast K4me3 turnover. Proteomics. 2014;14(19):2190–9. [PubMed: 24826939]
68. Holm S A Simple Sequentially Rejective Multiple Test Procedure. Scandinavian Journal of Statistics. 1979;6(2):65–70.
69. Chen EY, Tan CM, Kou Y, Duan Q, Wang Z, Meirelles GV, et al. Enrichr: interactive and collaborative HTML5 gene list enrichment analysis tool. BMC Bioinformatics. 2013;14:128. [PubMed: 23586463]
70. Kuleshov MV, Jones MR, Rouillard AD, Fernandez NF, Duan Q, Wang Z, et al. Enrichr: a comprehensive gene set enrichment analysis web server 2016 update. Nucleic Acids Res. 2016;44(W1):W90–7. [PubMed: 27141961]



	H3 interaction energy	DNA-NSD2 interaction energy	NSD2 potential energy	NSD2 residue 1099 interaction energy	H3 residue 31 interaction energy
E1099-H3.3 to K1099-H3.3	0.7% (17)	-2% (-231)	-2.6% (-193)	-106% (-259)	-18% (-12)
K1099-H3.1 to K1099-H3.3	7% (163)	0.4% (44)	1% (85)	-20% (-50)	11% (7)

Figure 1. E1099K mutation enhances NSD2 interaction with chromatin.

A,in the NSD2-SAM-NCP model NSD2, DNA, Histone H3, H4, H2A.1 and H2B.2 are colored in cyan, green, red, blue, orange and pink, respectively. **B**, shows a closer view of the interaction interface between NSD2 and NCP. SAM and H3 K36 residue are shown in stick representation. **C**, the percentage difference of averaged interaction energies between three simulated systems were calculated from MD simulations. The energy difference is given in kcal/mol in parentheses such that negative values indicate more favorable interactions. **D**, probability distribution of the SASA ($Å^2$) of the NSD2 histone-binding site

(residues M1119 to A1130 and N1178 to N1186) for the three systems. **E**, representative FRAP curve comparing C terminus GFP-tagged wildtype and E1099K NSD2 transiently expressed in 293 cells. Plotted as average (n=11) ±SEM. E (inset), average mobile fraction ±SEM of GFP-tagged wildtype and E1099K mutant NSD2 measured by FRAP (n=6). Two sample t-test with unequal variance, p=0.0006. **F**, H3 and NSD2 immunoblot of bound and input fractions from *in vitro* pulldown assay showing wildtype and mutant NSD2 bound to nucleosomes following incubation with increasing concentrations of NaCl.

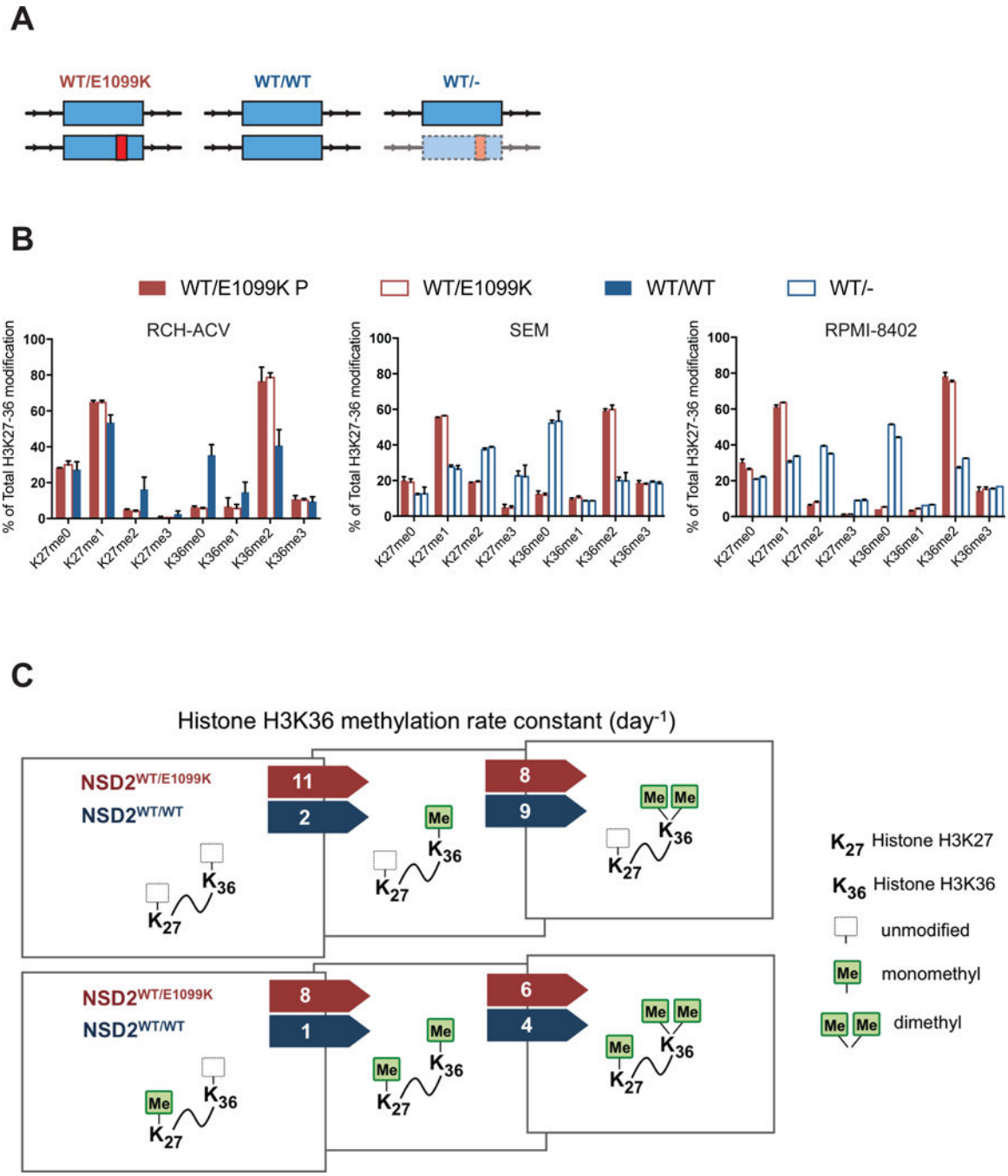


Figure 2. Gene editing to correct the E1099K mutation restores the balance of H3K36me2 to H3K27me3.

A,diagram illustrating the varieties of NSD2-mutant and wildtype subclones isolated from RCH-ACV, SEM, and RPMI-8402 cell lines. **B,** mass spectrometry quantifying histone marks in CRISPR-edited and unedited cell lines RCH-ACV (mean +SD, n=2 each). Bars represent cell lines, with colors and fill-patterns indicating genotype. **C,** diagram representing the comparison of histone H3K36 methylation rate constants between

$NSD2^{WT/E1099K}$ and $NSD2^{WT/WT}$ RCH-ACV cells when neighboring K27 is unmodified (top) or monomethylated (bottom) as determined by M4K.

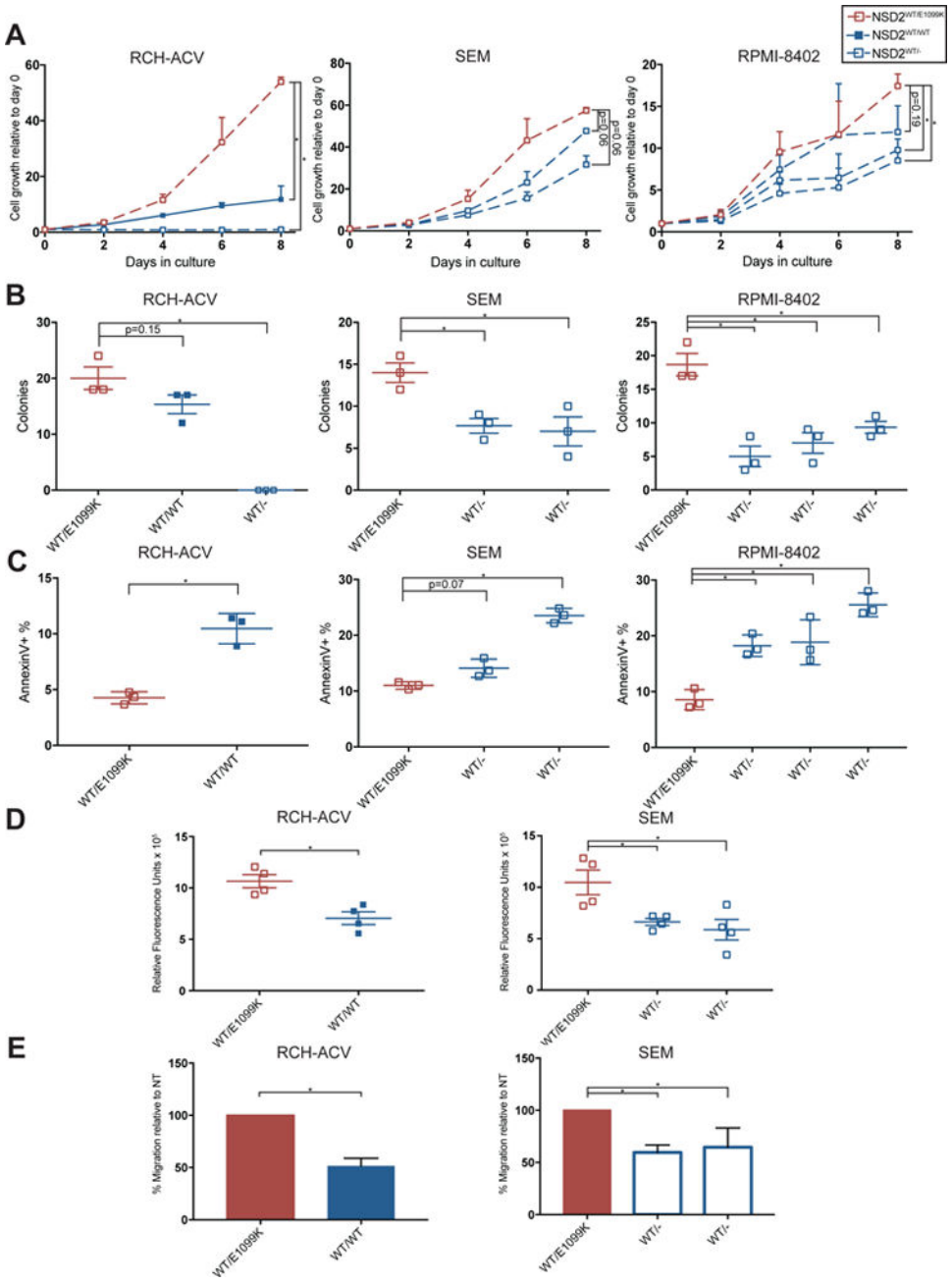


Figure 3. E1099K mutation promotes cancer-associated phenotypes

A, proliferation of RCH-ACV, SEM, and RPMI-8402 cells in liquid culture was monitored by Celltiter-Glo Luminescent Cell Viability Assay every two days (n=2–3; mean +SEM). Day 8 data analyzed by two-tailed t-test with equal variance. **B,** colonies formed in methylcellulose were counted 7 days after plating (n=3). **C,** cells were stained for Annexin V using a FITC conjugated antibody and then quantified by flow cytometry (n=3, mean ±SD). **D,** cells were fluorescently labeled and incubated on a layer of HS-5 bone marrow stromal cells for 24 hours. Following washes, adherent cell numbers were estimated by relative fluorescence (n=4). **E,** cells were seeded in top of Boyden chamber containing no

serum and allowed to migrate for 24 hours to serum-containing bottom chamber. Cells in bottom chamber were quantified by flow cytometry and normalized to non-targeted, $NSD2^{WT/E1099K}$ (RCH-ACV, n=3; SEM, n=4). Data are presented as mean ±SEM and analyzed by two sample t-test with unequal variance unless otherwise indicated; * $p<0.05$.

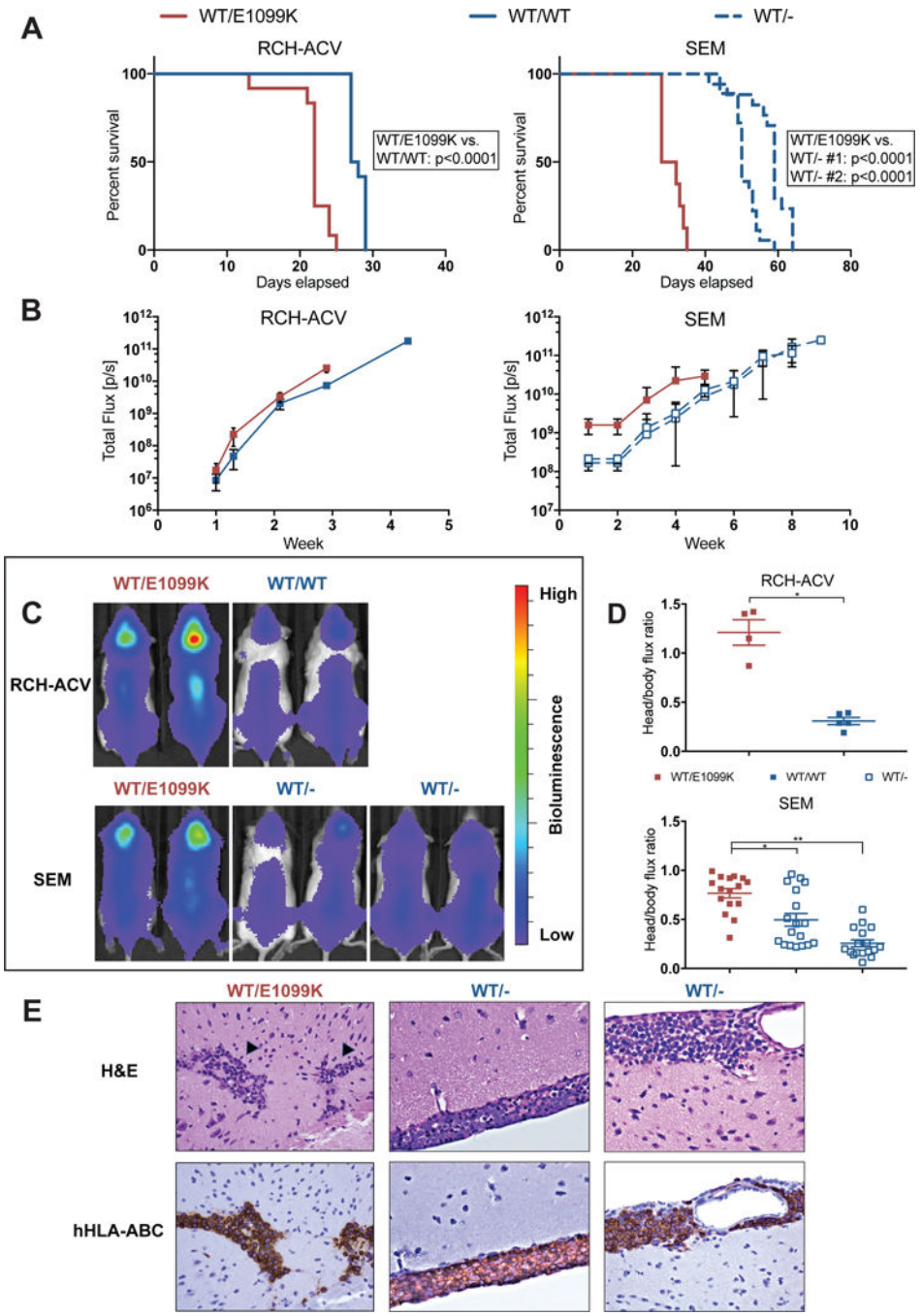


Figure 4. NSD2-mutant cells behave aggressively *in vivo*

A, survival curves following injection of NOD-SCID mice with 5×10^6 RCH-ACV- or SEM-derived cells. Log-rank test was used for statistical analysis to compare mice injected with RCH-ACV $NSD2^{WT/E1099K}$ cells to $NSD2^{WT/WT}$ cells and to compare SEM $NSD2^{WT/E1099K}$ injected mice to those injected with two independently derived $NSD^{WT/-}$ cell lines (RCH-ACV, n=24; SEM, n=51). **B,** the total bioluminescence signal (radiance, photons/second) was quantified weekly after xenograft. Y-axis is logarithmic to highlight differences in data. **C,** representative bioluminescence images from 3 weeks (RCH-ACV) or

4 weeks (SEM) post-injection. **D,** the ratio of bioluminescence signal between head and body was quantified at these time points for RCH-ACV- (n=9) and SEM-injected (n=51) mice. * $p<0.01$, ** $p<0.0001$; two sample t-test with unequal variance. **E**, representative brain sections from mice injected with SEM cells showing CNS infiltration by blasts with surface and cytoplasmic immunoreactivity for hHLA-ABC immunohistochemistry **(400x magnification)**. Black arrowheads indicate parenchymal infiltration.

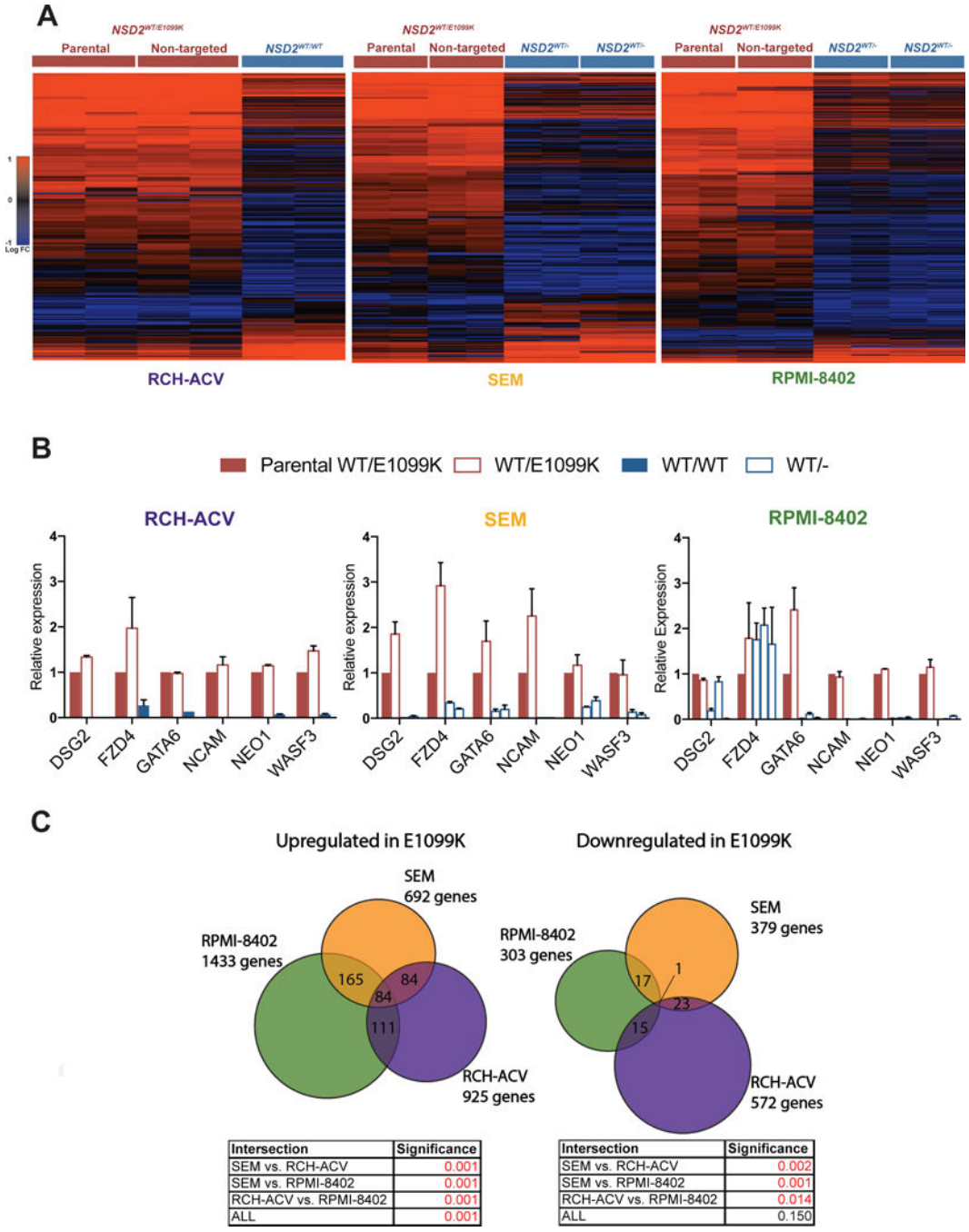


Figure 5. E1099K mutation leads to widespread gene activation

A, RCH-ACV, SEM, and RPMI-8402 heat maps depicting differentially expressed genes between parental NSD2$^{WT/E1099K}$, non-targeted NSD2$^{WT/E1099K}$, and CRISPR-edited NSD2$^{WT/WT}$ or NSD2$^{WT/-}$ cell lines. **B,** Selected genes were tested by qPCR to validate RNA-seq results. Data are presented mean ±SEM (n=2). **C,** Venn diagram displays overlap between different cell lines in genes upregulated or downregulated in NSD2$^{WT/E1099K}$ versus NSD2$^{WT/WT}$ and/or NSD2$^{WT/-}$ cells (fold-change>2, p<0.05).

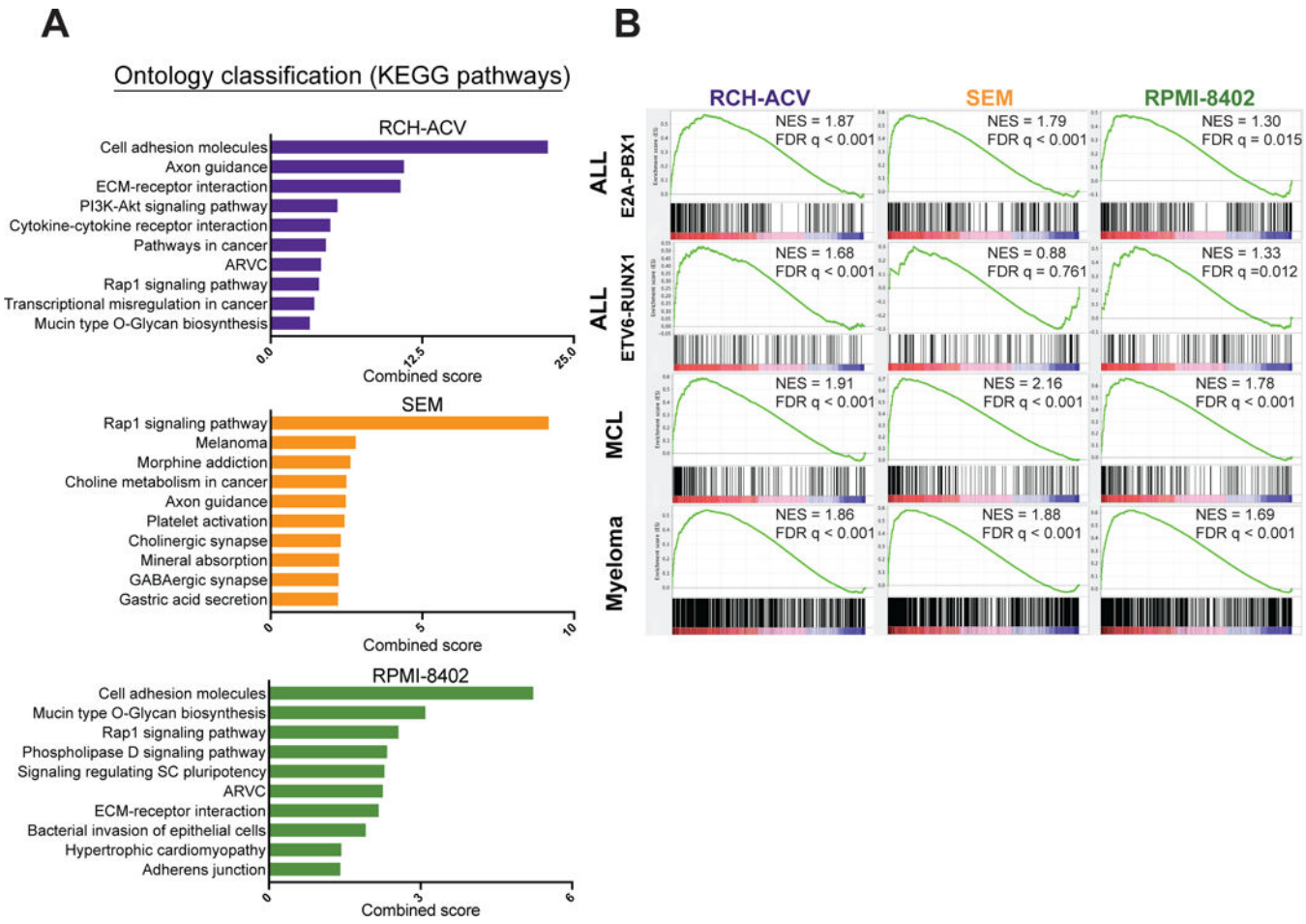


Figure 6. Similar genes and pathways are differentially activated in NSD2-mutant cell lines and patients

A, KEGG pathways significantly enriched in NSD2$^{WT/E1099K}$ cells using gene ontology analysis (p and q combined score). **B,** GSEA analysis comparing genes upregulated in mutant NSD2 ALL, MCL, and MM patients with RNA-seq from CRISPR-edited RCH-ACV, SEM, and RPMI-8402 cell lines; NES = normalized enrichment score; FDR q = false discovery rate q value