



USP44 Overexpression Drives a MYC-Like Gene Expression Program in Neuroblastoma through Epigenetic Reprogramming

Thomas L. Ekstrom^{1,2}, Sajjad Hussain^{3,4}, Tibor Bedekovics³, Asma Ali³, Lucia Paolini^{3,5}, Hina Mahmood³, Raya M. Rosok², Jan Koster⁶, Steven A. Johnsen², and Paul J. Galardy^{3,7}

ABSTRACT

Neuroblastoma is an embryonic cancer that contributes disproportionately to death in young children. Sequencing data have uncovered few recurrently mutated genes in this cancer, although epigenetic pathways have been implicated in disease pathogenesis. We used an expression-based computational screen that examined the impact of deubiquitinating enzymes on patient survival to identify potential new targets. We identified the histone H2B deubiquitinating enzyme USP44 as the enzyme with the greatest impact on survival in patients with neuroblastoma. High levels of *USP44* significantly correlate with metastatic disease, unfavorable histology, advanced patient age, and *MYCN* amplification. The subset of patients with tumors expressing high levels of *USP44* had significantly worse survival, including those with tumors lacking *MYCN* amplification. We showed experimentally that USP44 regulates neuroblastoma cell proliferation,

migration, invasion, and neuronal development. Depletion of the histone H2B ubiquitin ligase subunit RNF20 resulted in similar findings, strongly implicating this histone mark as the target of USP44 activity in this disease. Integration of transcriptome and epigenome in analyses demonstrates a distinct set of genes that are regulated by USP44, including those in Hallmark MYC target genes in both murine embryonic fibroblasts and the SH-SY5Y neuroblastoma cell line. We conclude that USP44 is a novel epigenetic regulator that promotes aggressive features and may be a novel target in neuroblastoma.

Implications: This study identifies a new genetic marker of aggressive neuroblastoma and identifies the mechanisms by which its overactivity contributes to the pathophysiology of this disease.

Introduction

Neuroblastoma is the most common solid tumor outside the brain that affects children. The disease occurs almost exclusively in young children, with more than 90% of cases occurring within the first decade of life. Neuroblastoma most commonly develops in the adrenal medulla, the largest sympathetic ganglion in humans. Based on studies in humans and animal models, neuroblastoma is known to result from defective development of the peripheral nervous system with cells failing to undergo apoptosis and encountering a differentiation block after ganglion formation—leading to nascent tumors (1). In most cases, tumors found in the neonatal period resolve spontaneously or require comparatively low-intensity chemotherapy or surgical resection to result in cure (2). In a distinct

subset, however, the disease will metastasize and/or harbor aggressive biological features such as *MYCN* amplification, or segmental chromosome alterations that lead to an aggressive cancer that is very difficult to cure. Therefore, there is a tremendous need for more insight into the pathogenesis of this disease in the hope of developing novel and less toxic therapies. High-throughput sequencing of exomes, or whole genomes, has uncovered relatively few recurrent single-gene variants (3–7). In addition to amplification of the *MYCN* oncogene in approximately 20% of tumors, which is one of the most potent negative prognostic findings in this disease, the genomic alterations most associated with aggressive disease are segmental chromosome alterations (8, 9). Alterations in portions of several chromosomes such as losses of chromosome 1p, or 11q, or gain of chromosome 17q are common and associated with poor outcomes (8, 9). Rare cases of germline predisposition to neuroblastoma have uncovered genes such as *ALK* and *PHOX2B*, which, when altered, drive changes in peripheral nervous system development and lead to tumors (10–15). A better understanding of the mediators of the early steps in neuronal development may lead to an improved understanding of tumor pathogenesis and novel therapies.

Over the last several years, several studies have implicated epigenetic pathways in the pathogenesis of neuroblastoma (16–19). Based on the observation of two distinct morphologies of the neuroblastoma cell line SK-N-SH, a super-enhancer-defined gene expression network was uncovered—defining two differentiation states that resemble adrenergic or mesenchymal cells (17, 20). Most tumors are thought to be comprised of cells in both states—though this is challenging to demonstrate using histologic techniques. Recent work has shown that the relative chemoresistance of cells in the mesenchymal state leads to the enrichment of this cell type in

¹Mayo Clinic Graduate School of Biomedical Sciences, Rochester, Minnesota.

²Robert Bosch Center for Tumor Diseases, Stuttgart, Germany. ³Department of Pediatric and Adolescent Medicine, Mayo Clinic, Rochester, Minnesota.

⁴Department of Family Medicine, Mayo Clinic, Rochester, Minnesota. ⁵Department of Pediatrics, University of Milano-Bicocca, San Gerardo Hospital, Monza, Italy.

⁶Department of CEMM, Amsterdam UMC, University of Amsterdam, Amsterdam, the Netherlands. ⁷Division of Pediatric Hematology-Oncology, Mayo Clinic, Rochester, Minnesota.

T.L. Ekstrom and S. Hussain contributed equally to this article.

Corresponding Author: Paul J. Galardy, Mayo Clinic, 200 First Street SW, Rochester, MN 55905. E-mail: Galardy.paul@mayo.edu

Mol Cancer Res 2024;22:812–25

doi: 10.1158/1541-7786.MCR-23-0454

This open access article is distributed under the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International (CC BY-NC-ND 4.0) license.

©2024 The Authors; Published by the American Association for Cancer Research

recurrent or refractory tumors (17, 21–23). Recent evidence indicates that cells may transit between the adrenergic state and the mesenchymal *in vivo* through epigenetic reprogramming (24). Epigenetic regulators have also been implicated in the differentiation blockade in neuroblastoma—with histone deacetylase inhibitors leading to enhanced retinoic acid–induced differentiation (25, 26).

Deubiquitinating enzymes (DUB) are encoded by a family of nearly 100 genes in mammals that reverse selective protein ubiquitination, thereby regulating myriad processes from protein degradation to DNA damage repair. There are several examples (21–23) of DUBs acting as oncogenes (*UCHL1*, *USP7*, and others), tumor suppressors (*A20* (*TNFAIP3*), *BAP1*, *USP24*, and others), and an extensive list of DUBs that under certain contexts may both promote or protect from cancer (27). The increasing body of evidence linking these enzymes with cancer has led to growing interest in this family as a new therapeutic target in cancer. To identify novel potential drug targets in neuroblastoma, we utilized the PREdiction of Clinical Outcomes from Genomic profiles tool (PRECOG; <https://precog.stanford.edu/>), which correlates gene expression with patient survival across multiple human cancers. Using this tool, we profiled the impact of 95 known or predicted DUBs on outcomes in neuroblastoma (28, 29). The results indicated that high expression of the DUB *USP44* had the greatest impact on patient survival. Here, we show that *USP44* is a powerful predictor of poor outcomes in neuroblastoma independently of other established risk factors and provide mechanistic data implicating it as an epigenetic regulator and potential drug target in this aggressive childhood disease.

Materials and Methods

Data analysis

The analysis of the impact of DUBs on cancer outcomes was performed using the PRECOG web tool (<https://precog.stanford.edu/>) as previously published (19). The datasets used in the study are listed in Supplementary Table S1. All survival analyses were initially performed using the R2: Genomics Analysis and Visualization platform (<http://r2.amc.nl>). Gene expression and clinical data from the datasets were downloaded from the R2 platform using the datagraber tool and analyzed using embedded tools in GraphPad Prism or using the BlueSky statistics program. Where indicated, gene expression was stratified by the median expression, or using a percentile ranking derived from the R2 platform's KaplanScan tool to find the cutoff with the most significant *P* value. The CIN25 signature score (21) was generated by determining the percentile rank expression for each of the 25 genes within the signature score using the sample Ranked GeneSetScores tool within R2. A high CIN25 score was defined as a greater than median CIN25 score. The chr. 4p score was generated by identifying the top 25 genes from chromosome 4p that were upregulated in MYCN amplified versus nonamplified tumors (see Supplementary Table S3). Tumors from the SEQC dataset were then scored using the sample Ranked GeneSetScores as described above. A high chr. 4p score was defined as a greater than median score. Unless otherwise indicated, graphs were reproduced using GraphPad Prism by exporting data from R2. Single-cell RNA-seq data were analyzed using the R2 webtool.

Cell culture, proliferation, migration, invasion, immunoblotting

Murine embryonic fibroblasts (MEF) were generated in our laboratory from day 13.5 embryos using standard techniques as reported and characterized (10). The genotype of MEFs was assessed

by PCR. All MEFs were cultured as described previously (30). All experiments using primary MEFs were performed between passages 2 and 5. Mycoplasma testing was not routinely performed on MEF cultures. immortalization of MEFs was performed by lentiviral transduction with the SV40 large T antigen as described (31). The passage number of immortalized MEFs is not monitored. Neuroblastoma cell lines were obtained from ATCC and identity verified every 3 years using the ATCC verification service. Proliferation was monitored for three independent cell cultures for each genotype/condition in replicate (minimum = 3; range 3–9) using an IncuCyte (Essen Bioscience) by monitoring five regions of each well every 2 hours. The device was maintained in a water-jacketed incubator with 5% carbon dioxide. Retinoic acid (RA)-induced differentiation was performed as described (32). Briefly, RA was added to SH-SY5Y cells, and proliferation, confluence, nuclei numbers, and neurite length were monitored using ImageJ or the IncuCyte with built-in tools where indicated. Cell migration and invasion were measured using the transwell system with the migrated/invaded cells quantitated using Cultrex cell migration and invasion kits according to the manufacturer's instructions (R&D systems). Immunoblotting was performed using standard sodium dodecyl sulfate–polyacrylamide gel electrophoresis, semidry transfer, and the following Cell Signaling Technology antibodies: histone H2B (8135), mono-Ub-H2B (5546), RNF20 (9425), histone H3 (4499), and MYCN (9405). The *USP44* antibody is from Santa Cruz (#sc-377203). All blots were developed using film and blots were scanned.

CRISPR editing and RNAi

IMR-32 cells with *USP44* gene deletion were generated using the pCas-Guide system (Origene) using predesigned guide RNAs targeting *USP44*. Following drug selection, cells were plated using limiting dilution, and single-cell clones were selected. The deletion of the *USP44* allele was verified by PCR. Although determining whether *USP44* was essential was not part of the experimental design, only single-allele deletions were identified. Where indicated, the expression of *USP44* was further reduced using previously characterized lentiviral shRNA (33). In experiments targeting *RNF20*, IMR-32 cells were stably transduced to express Cas9 using lentiCas9-Blast (Addgene, #52962). IMR-32^{Cas9} cells were transduced with *RNF20* targeting guide RNAs (gRNA) cloned into the LentiArray vector (Invitrogen). Pooled puromycin-selected transductions were used for experiments. Lentiviral transductions of MEFs or neuroblastoma cell lines were conducted as described (11).

ChIP-seq and qPCR

Chromatin immunoprecipitation was performed by crosslinking cells with 1% formaldehyde for 20 minutes and quenched by glycine for 5 minutes. Cells were lysed with lysis buffer (150-mmol/L NaCl, 20-mmol/L EDTA, 50-mmol/L Tris-HCl, 0.5% v/v NP-40, 1% v/v Triton X-100, and 10-mmol/L NaF). Nuclear pellets were resuspended and sonicated in sonication buffer (150-mmol/L NaCl, 20-mmol/L EDTA, 50-mmol/L Tris-HCl, 1% v/v NP-40, 0.5% v/v sodium deoxycholate, 10-mmol/L NaF, and 0.1% SDS) using a Bioruptor Pico (Diagenode) and a cycle setting of 30 seconds on/off until fragment sizes around 300 to 600 bp was reached. Protease inhibitor (Roche, 11836153001) and DUB inhibitors iodoacetamide and *N*-ethylmaleimide were added to both lysis and sonication buffers. Samples were precleared by 50% slurry of Sepharose 4B (GE Healthcare) resuspended in sonication buffer. Antibodies were added and incubated rotating overnight at 4°C. Antibodies include

1- μ g H3K4me3 (Diagenode, C15410003), 1- μ g H3K27ac (Diagenode, C15410196), 2- μ g H3K27me3 (Diagenode, C15410195), and 20- μ L Ubiquityl-Histone H2B (Lys120; H2Bub1; 5546, Cell Signaling). Protein A-sepharose beads (Cytiva) were added to samples and incubated for 2 hours, washed, decrosslinked, and RNase A and proteinase K treated, and DNA was extracted. Studies were performed in triplicate for each condition. Extracted DNA was sent to the Genome Analysis Core at Mayo Clinic to be library prepped and sequenced paired-end 50 bp on a HiSeq4000 (Illumina). Library preparation was carried out with ThruPLEX DNA-seq Kit V2 (Rubicon Genomics).

ChIP-seq bioinformatic analysis

Paired-end 50-bp sequencing reads were mapped to the reference genome assembly mm10 using bowtie2/2.3.3.1 (34). Triplicates were merged using samtools merge (samtools/1.9; ref. 35). Bigwig files for individual and merged bam files were generated ignoring duplicates and $1 \times$ depth (reads per genome coverage, RPGC) using bamCoverage (deeptools/3.3.2; ref. 36). Localization profiles were viewed using the Integrated Genomics Viewer (2.12.2; ref. 37). MACS2 (38) callpeak (python/3.9.2) was used to call the significant peaks with—broad cutoff 0.01 for H3K4me3 and H3K27ac using input files from respective conditions as background. ChIP occupancy was evaluated by the computeMatrix (deeptools/3.3.2) tool, and the average profiles and heatmaps were generated based on computeMatrix values with plotHeatmap (deeptools/3.3.2). mm10 transcriptional start site (TSS) coordinates were retrieved from the UCSC table browser [group: Genes and Gene Predictions; track: NCBI RefSeq; table: RefSeq All (ncbiRefSeq)]. To ascertain TSS-specific regions in our system, we intersected TSS coordinates with our H3K4me3 bed file (MACS2 output) to retrieve active TSS regions using bedtools/2.27.1 (39). H3K27ac differential binding analysis was performed using DiffBind (Galaxy Version 2.10.0 + galaxy0). Default settings were used for Genomic Regions Enrichment of Annotations Tool (GREAT) and ChIP-Atlas.

RNA extraction coupled with NGS sequencing (RNA-seq)

RNA was extracted using the RNeasy Plus Mini Kit (Qiagen, 74134) and cDNA conversion using SuperScript III First-Strand Synthesis System (ThermoFisher, 18080051) per manufacturer's protocol. Before RNA-sequencing, RNA integrity was validated by gel electrophoresis, and 1 μ g of purified RNA was sent to the Genome Analysis Core at Mayo Clinic to be library prepped and sequenced stranded paired-end 100 bp on a HiSeq4000 (Illumina; MEF data) or sent to the Robert Bosch Center for Tumor Diseases Sequencing Core and sequenced stranded paired-end 59 bp on a NextSeq 2000 (SH-SY5Y data). Library preparation was carried out with TruSeq Stranded mRNA Sample Prep Kit (Illumina). Studies were performed in triplicate.

RNA-seq bioinformatic analysis

Paired-end sequencing reads were mapped to the reference genome assembly hg38 or mm10 using STAR (star/2.7.3a; ref. 40). Reverse (sense strand) reads were quantified using htseq-count (htseq/0.9.1; ref. 41) and used for differential gene expression analysis via DESeq2 (Galaxy Version 2.11.40.7 + galaxy2; ref. 42). Gene set enrichment analysis (GSEA 4.2.2; ref. 43) was performed with default settings using normalized counts from DESeq2 for expressed genes. ChIP platform Mouse_Gene_Symbol_Remaping_Human_Orthologs_MSigDB.v2022.1. Hs.chip was used to convert mouse gene symbols to humans. Z-scores were calculated

with the following formula: $[(\text{sample}_{nc} - \text{row median}_{nc}) / \text{row StDev}_{nc}]$ ($nc = \text{normalized count}$). Subsequent Z-scores were plotted using GraphPad Prism software.

Data availability

Data are available in a public, open-access repository: GSE262506 (SH-SY5Y RNA-seq), GSE228130 (MEF RNA-seq), GSE228131 (MEF ChIP-seq).

Statistics

Quantitative *in vitro* experiments involving MEFs were performed with a minimum of three independently derived MEF lines that were nonimmortalized or immortalized as indicated. Graphs of these data represent the mean \pm SEM for the results obtained for each line—not the individual values. The test used is listed in each respective figure legend. Unless otherwise noted in each figure, a *P* value of less than 0.05 is used to denote significance. Unless otherwise noted in the figure legend, statistics were performed using the functions embedded in the software GraphPad Prism. Where indicated, calculations were performed using the statistical functions embedded in web tools including the z-scores in the PRECOG tool (<https://precog.stanford.edu/>) and the genomic correlations on the R2 website (<http://r2.amc.nl>) which were corrected for multiple comparisons using the FDR. The scan function to find the ideal cutoff values for survival based on gene expression uses *P* values corrected by the Bonferroni method.

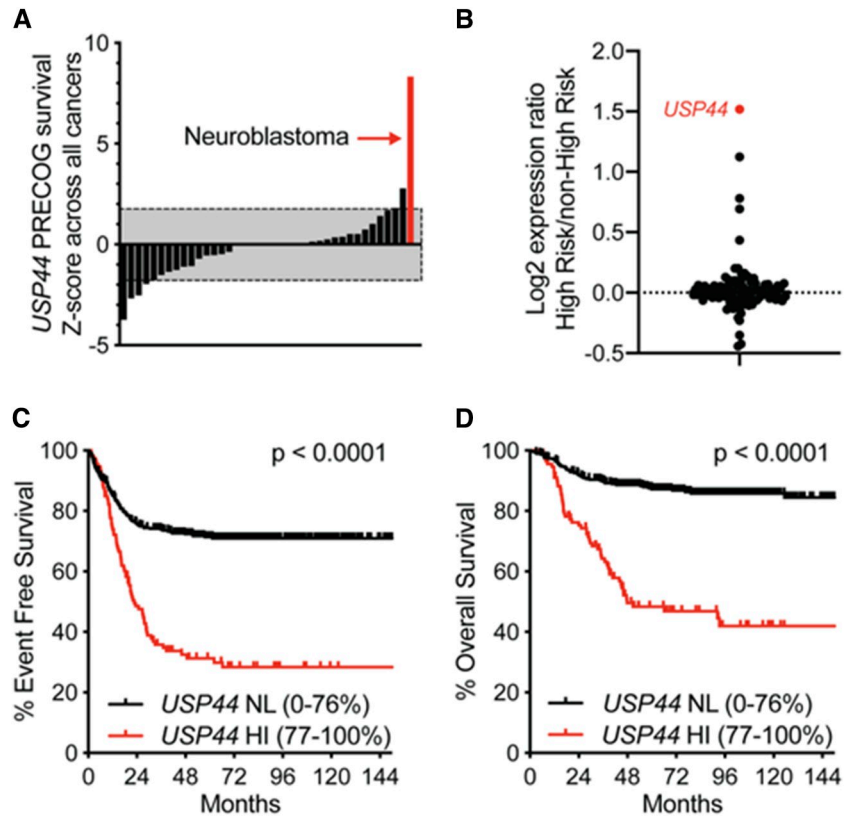
Results

High levels of *USP44* identify an aggressive subset of neuroblastoma tumors in children

We recently published the results of a computational analysis utilizing the PRECOG profiles tool (<https://precog.stanford.edu/>) in which we profiled the impact of 95 known or predicted DUBs on outcomes in neuroblastoma (28, 29). The results revealed that the enzyme *USP44* has the highest positive z-score value of 8.33, which is similar to that of the canonical neuroblastoma oncogene *MYCN* (z-score = 8.18; Fig. 1A). To further investigate the relationship between *USP44* and survival in neuroblastoma, we used public gene expression datasets that have associated clinical outcomes (Supplementary Table S1). We examined the expression of 94 DUBs in tumors classified as high-risk compared with non-high-risk tumors in the SEQC dataset ($n = 498$; Supplementary Table S1). *USP44* was the most overexpressed DUB in high-risk tumors (Fig. 1B). Next, we stratified the cohort based on *USP44* expression to examine its impact on survival using the R2: Genomics Analysis and Visualization Platform (<http://r2.amc.nl>). Patients whose tumors had high *USP44* expression ($>$ median) had significantly worse event-free and overall survival (EFS and OS, respectively) compared with those with lower expression (Supplementary Fig. S1A and S1B). We observed a similar adverse survival impact when we analyzed two distinct datasets [Primary NRC $n = 283$ (Supplementary Fig. S1C and S1D), TARGET $n = 249$ (Supplementary Fig. S1E and S1F)] despite substantial differences in the rates of *MYCN* amplification and metastatic disease in these cohorts (Supplementary Table S1). We next used the KaplanScan function of the R2 platform and identified a cutoff of 76 percentile *USP44* expression that produced the most significant separation of both EFS and OS (76th and 77th percentile, respectively; Fig. 1C and D). Unless otherwise stated, *USP44* normal (NL) includes tumors with transcript between 0 and

Figure 1.

USP44 expression identifies patients with aggressive neuroblastoma. **A**, PRECOG Z-scores for *USP44* across 39 types of human cancer. The gray box indicates nonsignificant Z-scores. **B**, The graph represents the Log₂ expression ratio (high-risk/non-high-risk) for 95 deubiquitinating enzymes in tumors from the SEQC cohort. **C** and **D**, Kaplan-Meier survival curves displaying the EFS and OS for the entire cohort of patients in the SEQC cohort stratified using the optimal cutoff identified using the Kaplan-scan function of *R*². *P*-values were calculated using the Log-rank test.



76th percentiles, whereas *USP44* high (HI) includes tumors between the 77th and 100th percentiles.

We examined the relationship between *USP44* expression and other established prognostic factors. There was a significant and stepwise increase in the expression of *USP44* with increasing clinical stage and was highest in those tumors from patients with metastases (International Neuroblastoma Staging System stage 4; **Fig. 2A**). There was also a significant correlation between *USP44* expression and older patient age at diagnosis, *MYCN* amplification, and unfavorable histology (**Fig. 2B–D**). We confirmed the impact of *USP44* as an independent prognostic factor through univariable and multivariable analyses in which *MYCN* status, histology, age, and stage were included (**Table 1**). Despite the association of high *USP44* expression with *MYCN* amplification, a significant impact of *USP44* on survival was observed in patients with *MYCN* nonamplified tumors and those with non-high-risk tumors (**Fig. 2E and F**). These data lead us to conclude that increased *USP44* expression is an independent prognostic factor for poor survival in children with neuroblastoma.

Because *USP44* regulates H2BK120ub1 (44, 45), we speculated that it may play a role in regulating the transition between super-enhancer-driven adrenergic and mesenchymal cell states (17). However, we found no correlation between *USP44* expression and the mesenchymal or adrenal scores in the neuroblastoma datasets used in our study (Supplementary Fig. S2). We previously demonstrated that *USP44* is important to ensure accurate mitosis, as both *USP44* loss and its overexpression lead to chromosome instability (CIN; refs. 30, 46). We have also shown that neuroblastoma tumors with a higher CIN gene expression signature (CIN25) have a worse survival (28, 47). Looking at *USP44*, there is a significant correlation

between increasing *USP44* expression and the CIN25 score (Supplementary Fig. S3A and S3B). Next, we plotted the EFS in the SEQC dataset stratified on both CIN and *USP44*. There was an additive effect of both parameters on survival in this analysis, with the worst outcome seen in tumors with a positive CIN25 signature and high *USP44* (Supplementary Fig. S3C). The absence of either of these features resulted in an improved outcome, and patients whose tumors lacked both features did generally well. This suggests that there are likely CIN-dependent and CIN-independent effects of *USP44* on neuroblastoma biology.

USP44 expression promotes features of aggressive neuroblastoma

To establish a mechanistic connection between *USP44* and disease aggressiveness in neuroblastoma, we identified cell lines with low or high levels of *USP44* in gene expression datasets for further experiments. The dichotomous expression of the *USP44* protein was confirmed in these cells by Western blotting (Supplementary Fig. S4A and S4B). We then generated cell-line pairs by introducing *USP44* into cells where it was low and depleting *USP44* in cells where it was expressed (Supplementary Fig. S4C). We depleted *USP44* using a two-step process. First, cells stably expressing Cas9 were transduced with a lentivirus encoding *USP44* targeting gRNA and drug resistance. We isolated several single-cell clones from these transductions—but did not observe a biallelic deletion of *USP44*. As cells lacking one allele expressed intermediate amounts of *USP44* protein, in some experiments, we additionally transduced the cells with *USP44* shRNA to reduce expression further. We monitored the impact of these manipulations on cell proliferation using the IncuCyte live cell imaging system. We observed that depletion of

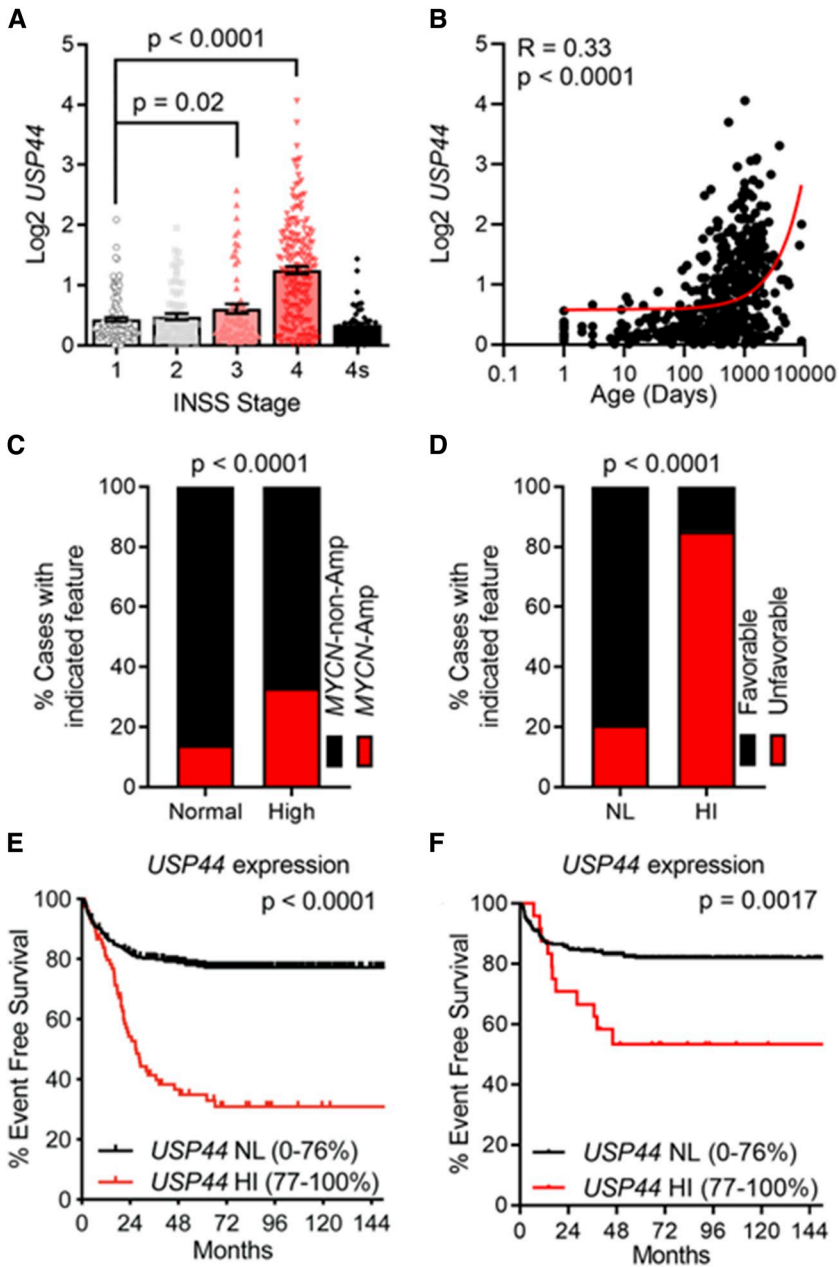


Figure 2. High levels of *USP44* correlate with established poor prognostic features in neuroblastoma. **A**, The graph represents the mean ± SEM for patients in the SEQC dataset classified by the International Neuroblastoma Staging System (INSS) stage. The *P*-values shown were calculated using an unpaired two-tailed *t* test. **B**, The graph shows the level of *USP44* plotted against age at diagnosis for the SEQC cohort. The Pearson correlation (*R*-value) and *P*-value were calculated in GraphPad Prism. The (red) line indicates the trend calculated using a simple linear regression function in Prism. **C** and **D**, Histograms showing the proportions of patients with tumors having *MYCN* amplification (**C**) or unfavorable histology (**D**) stratified by *USP44* expression. **E** and **F**, Event-free survival for patients with normal or high levels of *USP44* in the SEQC dataset whose tumors lacked *MYCN* amplification (**E**) or those stratified as non-high-risk (**F**). High *USP44* defined as >76 percentile.

USP44 slows proliferation in a dose-dependent fashion whereas overexpression promotes it (Fig. 3A and D). As the transcript levels are increased in metastatic tumors, we examined the impact of *USP44* overexpression or depletion on *in vitro* surrogates of metastasis, including assays for cell migration and invasion. Again, we observed that overexpression of *USP44* increased both features whereas *USP44* depletion suppressed them (Fig. 3B, C, E, and F). As *USP44* is both a regulator of histone ubiquitination with likely impacts on global gene expression (44, 45) and as it plays an important role in ensuring accurate mitosis (30, 46), we manipulated H2BK120ub1 independently by depleting an essential component of the ubiquitin E3 ligase complex responsible for this modification. Using stable Cas9-expressing IMR-32 cells, we introduced a control

gRNA or one that targeted *RNF20*—an essential subunit of the RNF20/40 complex (Supplementary Fig. S4D; ref. 48). In results very similar to those seen when *USP44* is overexpressed and consistent with previous work (48), we observed that *RNF20* depletion enhanced the proliferation, migration, and invasion of IMR-32 cells (Fig. 3G–I). These data strongly indicate that *USP44* drives malignant behavior in neuroblastoma through the suppression of H2BK120ub1.

Consistent with the reported role of *USP44* and nuclear co-repressor (N-CoR) in neuronal differentiation (49), we found that *USP44* expression decreased in embryonic stem cells differentiating along the neural lineage in both bulk and single-cell RNA-seq datasets (Fig. 4A and B). The drop of *USP44*

Table 1. Multivariable analysis of USP44 along with other established prognostic factors in neuroblastoma.

Covariate	HR	95% CI	P value
Overall survival			
USP44 HI	7	4.58-10.69	<0.0001
Age at diagnosis (>18 months)	10.04	6.03-16.73	<0.0001
INSS stage 4	22.82	13.83-37.65	<0.0001
MYCN amplified	21.14	13.08-34.15	<0.0001
Event-Free survival			
USP44 HI	7.05	4.62-10.75	<0.0001
Age at diagnosis (>18 months)	10.47	6.28-17.46	<0.0001
INSS stage 4	25.03	14.95-41.93	<0.0001
MYCN amplified	17.09	10.76-27.15	<0.0001

Data from SEQC (GSE62564) were analyzed using the Cox proportional hazard, multiple models. Cases that lacked all variables were omitted (final $n = 270$ patients).

Abbreviations: HR, hazard ratio; CI, confidence interval; INSS, international neuroblastoma staging system.

expression in these experiments occurred concurrently with other pluripotency markers (*POU5F1* and *LIN28A*) and the acquisition of neural markers (*NCAM1* and *NTRK2/3*). We also examined the expression of *USP44* in single cells derived from developing mouse embryos focusing on the developing neural crest (50). The cells with the highest levels of *USP44* were the premigratory neural tube cells, with expression low in most other cell types (Supplementary Fig. S5A). Interestingly, *USP44* is not expressed, or is expressed at very low levels, in all cell types analyzed in a recently published analysis of the developing adrenal medulla in humans (30, 51; Supplementary Fig. S5B). A similar pattern of *USP44* expression was also observed in bulk transcript data from the brain across the developmental spectrum (31; Supplementary Fig. S6A) as well as early human embryogenesis (52; Supplementary Fig. S6B). We conclude from these data that *USP44* is expressed at very early stages of development within the sympathetic nervous system and declines rapidly prior to the differentiation of the adrenal medulla.

The reduction in *USP44* observed during development led us to hypothesize that its continued expression in neuroblastoma tumors may contribute to the differentiation block seen in this cancer. To test this hypothesis, we introduced *USP44* into SH-SY5Y cells and examined its effect on RA-induced differentiation. When exposed to RA, SH-SY5Y cells elongated and developed neuritic projections (Fig. 4C). Concurrent with this there is an increase in H2BK120ub1, as described during neural differentiation from pluripotent cells (44). We observed a significant *USP44*-induced reduction in RA-induced neuritic development and H2BK120ub1 (Fig. 4D and E; Supplementary Fig. S7). Taken together, these data lead us to conclude that *USP44* expression declines during embryonic development of the peripheral nervous system and that aberrant continued expression, such as that observed in neuroblastoma tumors and cell lines, contributes to the differentiation block seen in this disease.

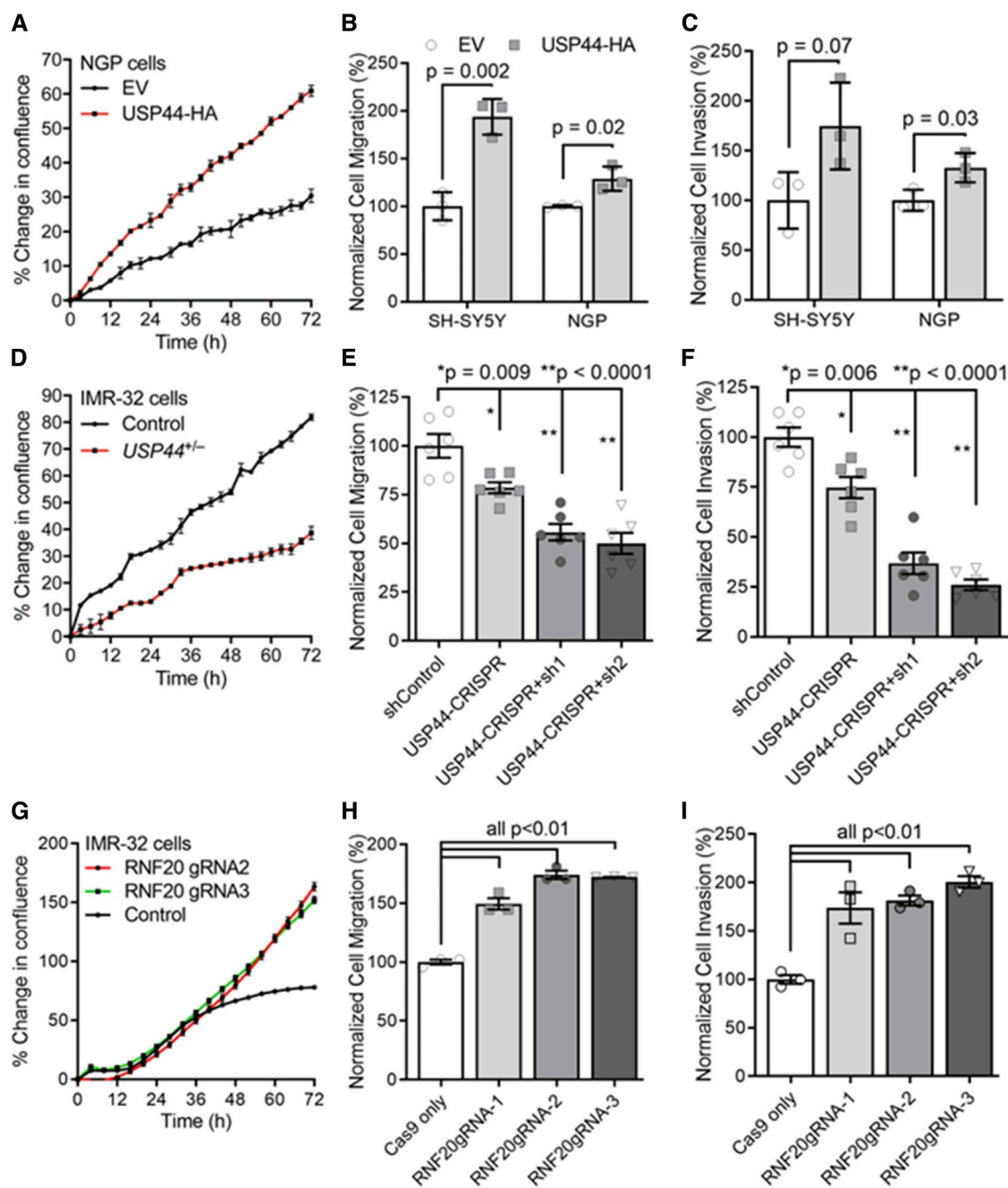
USP44 expression is correlated with copy number of chromosome 4p

Although classified by COSMIC as a tier 2 cancer gene and a likely cancer-causing gene in mice, systematic genomic sequencing projects have not demonstrated recurrent sequence

variants in *USP44* in neuroblastoma. Therefore, we hypothesized that *USP44* expression may be driven by another genomic event. As mentioned previously, high *USP44* levels were associated with *MYCN*-amplification. This led us to investigate whether *USP44* expression was induced by *MYCN*. At the transcript level, there was a slight but significant negative correlation between *USP44* and *MYCN* in tumors (Supplementary Fig. S8A and S8B). Further evidence was observed in cells with manipulated *MYCN* expression, in which no effect on *USP44* is seen (Supplementary Fig. S8C and S8D). These data strongly indicate that the high levels of *USP44* seen in *MYCN*-amplified tumors are not due to a direct impact of *MYCN* on *USP44*.

The long arm of chromosome 12, which includes the *USP44* locus, is gained in up to 20% of tumors, with the highest proportion with gains at 12q-ter (53, 54). In contrast, copy number loss at any portion of chromosome 12 is rare (53, 54). In a distinct dataset in which copy number data were accompanied by gene expression (4), there was no correlation between copy number at the *USP44* locus and its transcript level (Supplementary Fig. S9). When we broadened the analysis and looked for copy number variations that correlated with *USP44*, there was a significant positive correlation (FDR < 0.01) between *USP44* transcript levels and copy number alterations of genes on chromosomes 4p and 14p and a significant negative correlation with genes on chromosomes 1p and 17p (Fig. 5A). When *USP44* expression was correlated with conventionally defined segmental chromosomal alterations, the correlation with chr. 4p remained, with loss correlated with reduced expression (Fig. 5B), whereas there was no significant correlation between *USP44* and chr. 1p status (Fig. 5C). Chromosome 4p is commonly lost in *MYCN* nonamplified tumors observed in *MYCN*-amplified tumors, whereas this region is frequently retained in *MYCN*-amplified tumors (53, 54). In *MYCN*-nonamplified tumors, there was a significant reduction in *USP44* in those with chr. 4p loss (Fig. 5D). Although the number of *MYCN* amplified tumors with chr. 4p loss is small, there remains a trend ($P = 0.08$) toward lower *USP44* expression in *MYCN* amplified cases in which chr. 4p was lost (Fig. 5E). We conclude that the expression of *USP44* is not a target of *MYCN* activity. Instead, *USP44* expression is best accounted for based on the copy number of chr. 4p and that the correlation of *MYCN* amplification with retention of chr. 4p is the cause of the apparent correlation between *USP44* and *MYCN* status.

As there is currently no widely available marker for cases with high *USP44* expression, we next wondered whether the disomy of 4p itself may be of useful prognostic value. Using a list of the top 25 genes from chr. 4p, which are overexpressed in *MYCN* amplified tumors compared with nonamplified tumors, we generated a chr. 4p gene set score using the R2 embedded tool. When we applied this score across all cases within the SEQC dataset, there was a highly significant survival separation between those cases with high (>median) or low chr. 4p scores (Supplementary Fig. S10A). The negative prognostic impact of the chr. 4p score was substantially reduced in tumors where *USP44* was low (<median; Supplementary Fig. S10A). As the negative prognostic impact of this score may be influenced by the correlation with *MYCN* amplification, we examined the survival in *MYCN* nonamplified cases and again found those with a high chr. 4p score to have significantly worse survival (Supplementary Fig. S10B). The score is itself not an accurate test for *USP44* status, as it would have a high false positive and negative rate (Supplementary Fig. S10C). These data indicate that retention of both copies of chr. 4p may be a poor prognostic biomarker for *MYCN* nonamplified neuroblastoma.

**Figure 3.**

USP44 expression drives proliferation, migration, and invasion in neuroblastoma cell lines. **A**, NGP neuroblastoma cells were transduced with lentivirus encoding control EV or *USP44* with a c-terminal hemagglutinin (HA) epitope tag (*USP44*-HA). Proliferation was monitored with live video microscopy using an IncuCyte device. The mean \pm SEM of triplicate cultures is shown at each time point. **B** and **C**, The indicated cell lines were transduced as in **(A)**. Migration **(B)** and invasion **(C)** were measured using transwell devices. The graph represents the mean \pm SEM for three independent experiments. **D** and **G**, IMR-32 cells were transduced with CRISPR-CAS9 and control or *USP44* targeting guide RNA as shown in Supplementary Fig. S4C. Proliferation was monitored as in **(A)**. **E**, **F**, **H**, **I**, IMR-32 cells were transduced with lentivirus as shown. Migration and invasion were monitored as in **(B** and **C)**.

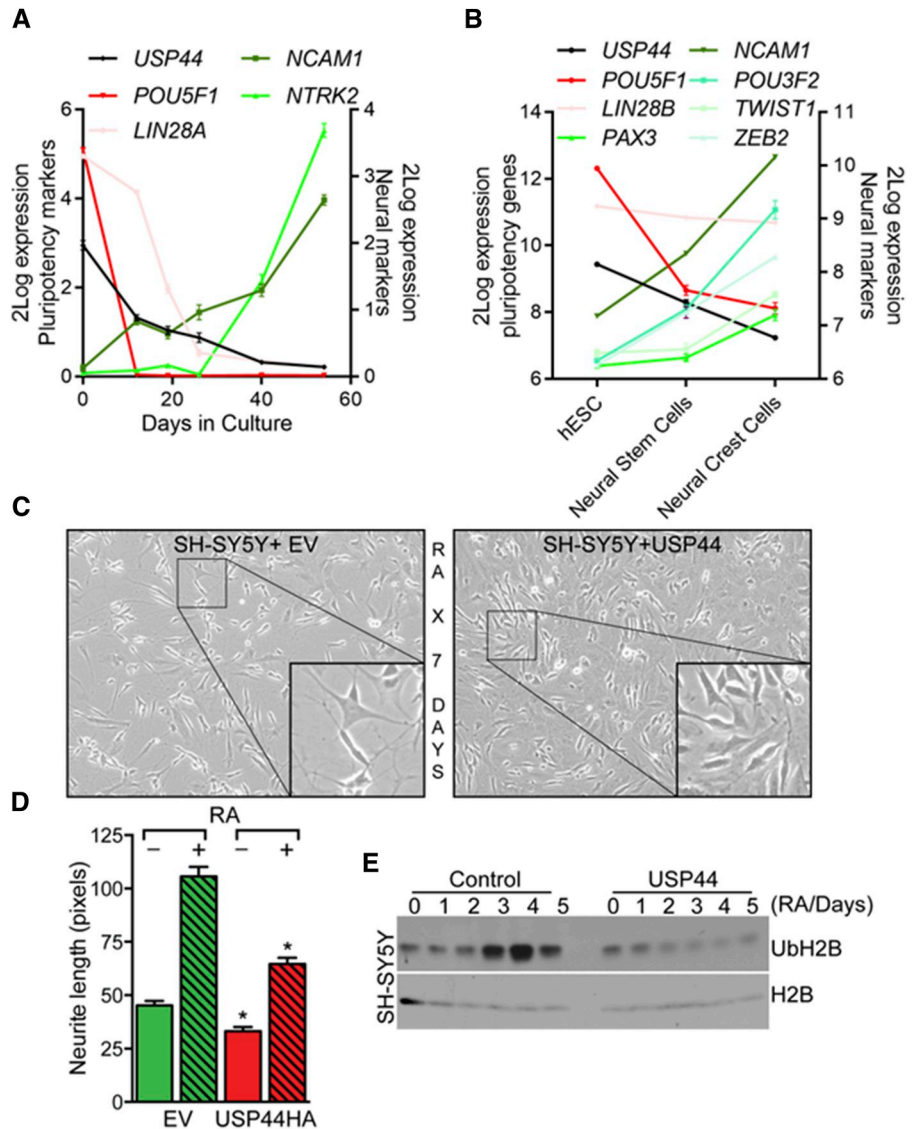
Histone H2B monoubiquitination correlates with gene expression patterns after loss of *USP44*

To directly examine the impact of *USP44* on gene expression, we performed mRNA sequencing (RNA-seq) in SH-SY5Y cells over-expressing (OE) *USP44*. SH-SY5Y cells have low *USP44* expression

and high *MYC* expression. Heatmap analysis shows there are 303 genes ($\text{Log}_2\text{FC} > 0.5$; base mean > 10 ; $P\text{-Adj} < 0.05$) and 476 genes ($\text{Log}_2\text{FC} < -0.5$; base mean > 10 ; $P\text{-Adj} < 0.05$) that are upregulated and downregulated, respectively, in *USP44* OE SH-SY5Y cells when compared with an empty vector (EV) control (Supplementary Fig.

Figure 4.

USP44 restrains differentiation in neuroblastoma. **A**, Data from GSE86977 was reanalyzed to show the expression of the indicated genes. The graph represents the mean ± SEM for the expression of each gene measured by single-cell RNA-seq at each time indicated. **B**, Data from GSE53203 were reanalyzed to show the expression of the indicated genes. The graph represents the mean ± SEM for bulk RNA-seq performed on three replicates of each cell state. **C-E**, SH-SY5Y cells were transduced as indicated and incubated with or without retinoic acid to induce neuritic differentiation. The lengths of neurites were measured using image J (**D**) and the mean ± SEMs of the three experiments are shown. Aliquots of cells were harvested at the indicated times and immunoblotted for the indicated proteins (**E**).



S11A). Interestingly, GSEA demonstrated significant enrichment of MYC target genes (Fig. 6A) in USP44 OE SH-SY5Y cells. To understand if this is a conserved function of *USP44* across species, we next sought to validate this data in an orthogonal approach in MEFs. We compared wild-type (WT) MEFs to those lacking the *Usp44* gene (*Usp44*^{-/-}), as well as *Usp44*^{-/-} MEFs into which *USP44* was reintroduced (*Usp44*^{res}). There were 4,694 genes that changed at least 1-fold (±0.5 Log₂FC; base mean > 10; *P*-Adj < 0.05) between WT and *Usp44*^{-/-} MEFs (2,587 up and 2,107 down in *Usp44*^{-/-}). With the exception of nine genes, all genes were restored to within 0.5 log₂FC of WT levels by the reintroduction of *USP44* (Supplementary Fig. S11B). GSEA in *Usp44*^{-/-} MEFs revealed MYC target genes are downregulated (Fig. 6B) upon loss of *USP44* in MEFs, thus showing the same directionality as the *USP44* OE SH-SY5Y cells in which MYC target genes are enriched. In patient neuroblastoma tumors, we also found these gene sets to be significantly enriched in *USP44* HI compared to *USP44* NL tumors, highlighting the patient relevance of these findings (Fig. 6C). To

further cross-validate these datasets, we performed GSEA using up- and downregulated genes in SH-SY5Y *USP44* OE cells (gene set) and observed how *Usp44*^{-/-} MEFs regulated these genes. We found genes downregulated in *USP44* OE SH-SY5Y cells were slightly enriched (FDR = 0.082) in *Usp44*^{-/-} MEFs showing similar directionality of gene regulation. Genes upregulated in SH-SY5Y *USP44* OE cells were also significantly upregulated in *Usp44*^{-/-} MEFs showing incorrect directionality (i.e., opposite trend of enrichment) of gene regulation (Supplementary Fig. S11C).

To understand the epigenetic impact on *USP44*-dependent gene regulation, we performed chromatin immunoprecipitation followed by high-throughput sequencing (ChIP-seq) for H3K4me3, H2BK120ub1, H3K27me3, and H3K27ac in WT, *Usp44*^{-/-}, and *Usp44*^{res} MEFs. Because we observed similar regulation of two independent MYC target gene sets, we intersected core enrichment genes from the GSEA in both SH-SY5Y *USP44* OE and WT MEF (i.e., downregulated upon *USP44* loss in MEFs) conditions, which resulted in 47 commonly regulated MYC target genes (Fig. 6D).

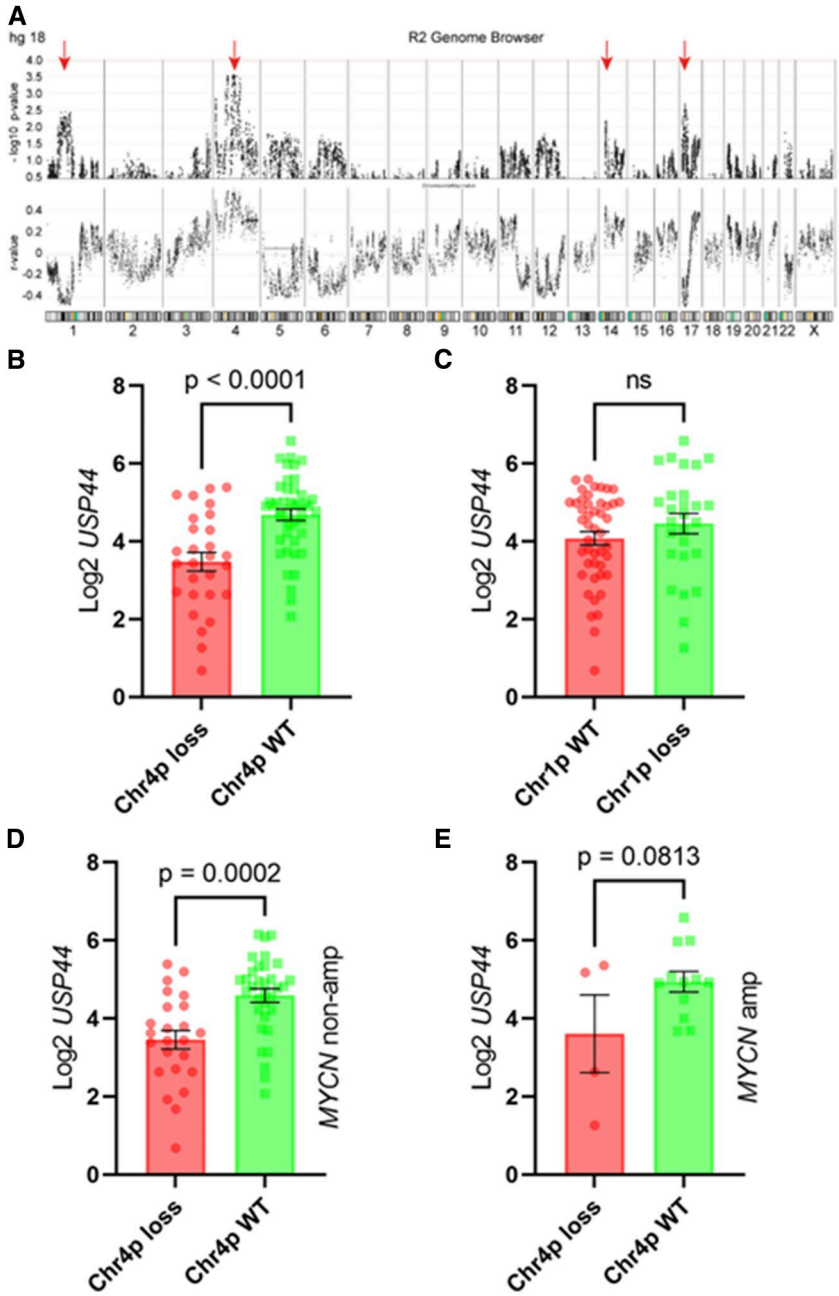


Figure 5. *USP44* expression correlates with the disomy of chromosome 4. **A**, Genome-wide map showing correlation between *USP44* expression and copy number in a dataset containing 108 tumors. Chromosomes are depicted horizontally as indicated with *P*-values for the correlation in the (top) graph and the Pearson correlation coefficient (*R*) in the (bottom). Negative *R* values indicate an inverse correlation. Arrows indicate genomic regions with peaks of significant correlation with *USP44* expression. **B–E**, The expression of *USP44* is shown stratified by chromosome 4p status (**B**), Chromosome 1p loss (**C**). The correlation of *USP44* with chr. 4p status is shown for tumors without (**D**) or with (**E**) *MYCN* amplification.

Interestingly, we did not observe a change in H3K4me3 levels at these loci, however, there was a decrease in H2BK120ub1 occupancy at the TSS and along the gene bodies (GB) downstream of the TSS (Fig. 6E) suggesting ubiquitin-mediated gene regulation. This is in contrast to the tight crosstalk between H3K4me3 and H2BK120ub1 as previously described (55), although we observed a change in H3K4me3 occupancy at downregulated and upregulated genes (Supplementary Fig. S12A–S12B). Integrated Genomics Viewer tracks at the top two commonly regulated *MYC* target genes with the highest H2BK120ub1 occupancy, *Pprc1* and *Nop56*, are shown in Fig. 6F.

Upon further global analysis, there was a significant alteration in H2BK120ub1 occupancy of genes that were altered in

Usp44^{-/-} cells (Supplementary Fig. S12C). Consistent with prior reports about H2BK120ub1, there was a substantial increase along GB downstream of the TSS in genes upregulated by *USP44* loss. This change in H2BK120ub1 was accompanied by an increased occupancy of H3K4me3 near the TSS of these genes (Supplementary Fig. S12B). These changes were again restored by the reintroduction of *USP44* into *Usp44*^{-/-} cells. For unregulated genes, the loss of *USP44* did not alter H2BK120ub1 or H3K4me3 in the gene body (Supplementary Fig. S12B and S12C). Interestingly, downregulated, upregulated, and unregulated genes had lower occupancy of H2BK120ub1 at the TSS in *Usp44*^{-/-} MEFs compared with WT MEFs with only changes of H2BK120ub1 in the gene body correlating with gene expression

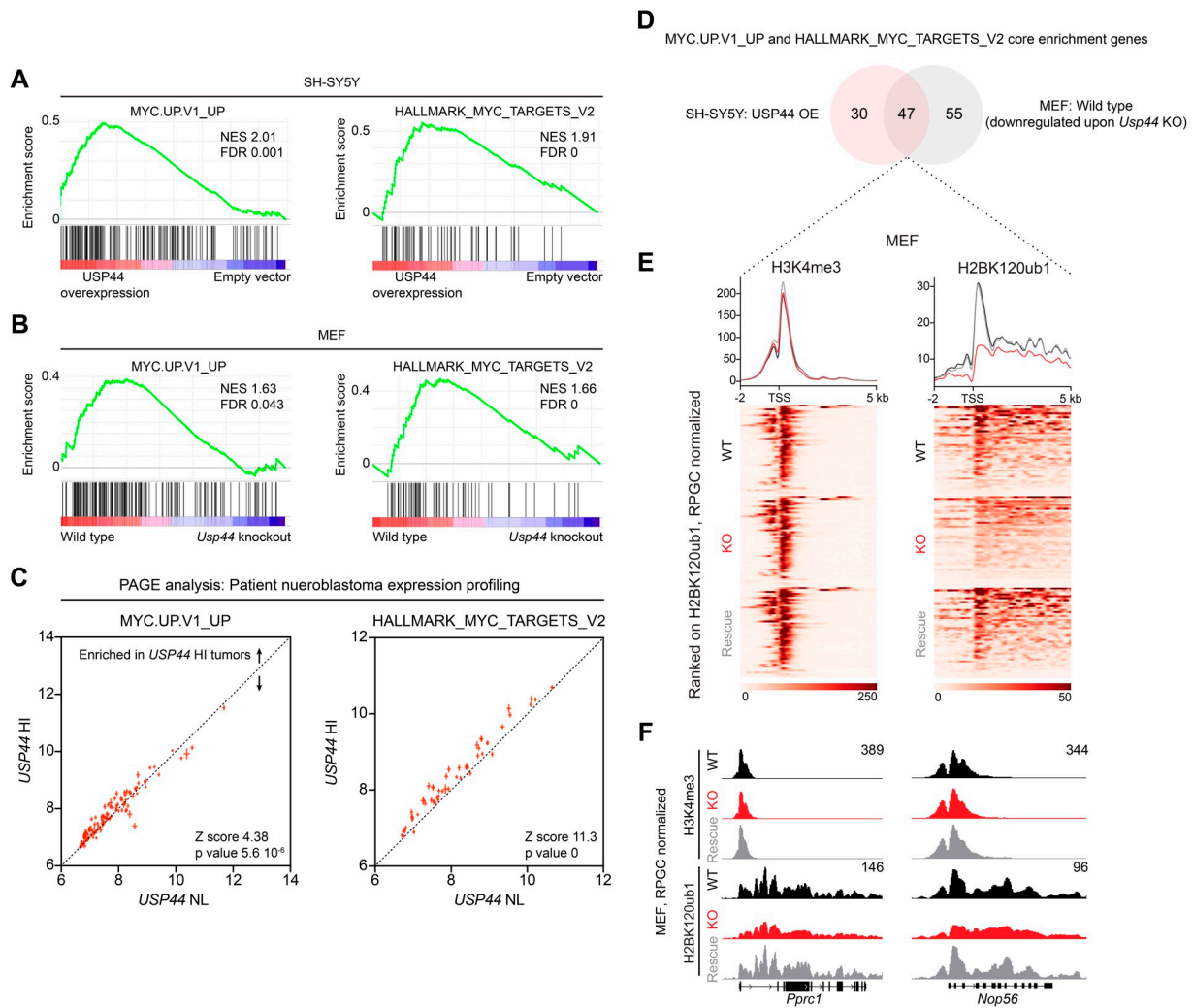


Figure 6.

USP44 alters the epigenome and regulates gene expression. **A** and **B**, GSEA was performed on expressed genes against the Oncogenic (left) and Hallmark (right) gene sets in SH-SY5Y (top) and MEF (bottom) cell lines with either overexpression or genetic depletion of USP44, respectively. **C**, Parametric analysis of gene set enrichment (PAGE) was performed against the Oncogenic (left) and Hallmark (right) gene sets from patient neuroblastoma tumor data expression profiling data. Error bars represent SEM. **D**, Venn diagram intersecting core enriched genes from USP44 overexpression in SH-SY5Y cells (**A**) and WT MEFs (**B**). **E**, Aggregate plots of H3K4me3 (left) and H2BK120ub1 (right) at the TSS (−2 kb/+5 kb) of common enriched genes shown in (**D**). **F**, IGV tracks of H3K4me3 and H2BK120ub1 at the (top) two genes ranked on H2BK120ub1 levels from (**E**): *Pprc1* and *Nop56*. All bigwigs are normalized to 1× depth (reads per genome coverage, RPGC). Bigwigs from each group are represented as follows based on *Usp44* status: WT (black), KO (red), and Rescue (gray).

change. We hypothesize that H2BK120ub1 spreading into the gene body promotes transcriptional elongation to control gene expression mediated by USP44. Lastly, given the finding that H2BK120ub1 facilitates the removal of the repressive mark H3K27me3 during differentiation (56), a histone mark deposited by the Polycomb Repressive Complex-2 that negatively regulates gene expression, we examined the effects of *Usp44* expression on this mark as well. As expected, we observed an increase in genes downregulated and a decrease in genes upregulated in *Usp44*^{−/−} cells, which were restored upon USP44 reintroduction (Supplementary Fig. S12D).

USP44 alters H3K27ac levels at distal elements that are bound by differentially regulated transcription factors after USP44 loss

As enhancers have been shown to drive cell states and mediate key pathways in cancer, we sought to explore differential occupancy of H3K27ac at intergenic enhancer regions as a result of USP44 loss. We performed differential H3K27ac binding analysis in WT and *Usp44*^{−/−} MEFs and found 4,200 (fold change < −2; FDR < 0.01) and 1,428 (fold change > 2; FDR < 0.01) regions that were reduced or enriched in *Usp44*^{−/−} MEFs, respectively. These regions were also restored upon USP44 rescue (Supplementary Fig. S13A), indicating

that they are USP44-dependent. As enhancers are distal regulatory elements, we sought to determine where the differentially bound H3K27ac regions were localized using the Genomic Regions Enrichment of Annotations Tool (57). Interestingly, we observed regions that gain and lose H3K27ac localize distal to the TSS, suggesting that USP44 may influence enhancer activity to alter target gene expression (Supplementary Fig. S13B). We next utilized the ChIP-Atlas tool (58) to identify transcription factors that may localize to these regions. This tool utilizes publicly available datasets for transcription factor occupancy to identify statistically significant enrichment within specific genomic regions of interest. In regions that lost H3K27ac, we observed enrichment of transcription factors that are upregulated (*Prdm16*, *Fos*, *Rnf2*, *Smarca2*, and *Fosl2*) and downregulated (*Fosl1*, *Myc*, *Rela*, and *Smarca4*) in *Usp44*^{-/-} MEFs. In contrast, in regions that gained H3K27ac, we observed enrichment of transcription factors that are upregulated (*Vdr*, *Prdm16*, *Pparg*, *Cebpa*, *Rnf2*, *Smarca2*, and *Cebpb*) and downregulated (*Smarca4*) in *Usp44*^{-/-} MEFs (Supplementary Fig. S13C; Supplementary Table S2). Interestingly, we observed an enrichment of *Myc* at regions that lost H3K27ac and is downregulated in *Usp44*^{-/-} MEFs, suggesting a possible feedback loop. Future studies depleting USP44 with subsequent ChIP-seq for MYC and H2BK120ub1 and HiChIP studies will help dissect the USP44-MYC enhancer-promoter interaction axis to mediate H2BK120ub1 at target genes.

Taken together, these results show that *USP44* upregulation is a feature of aggressive neuroblastoma that regulates key oncogenic pathways such as MYC target genes. USP44 may regulate these pathways by influencing chromatin interactions by controlling H3K27ac levels at intergenic enhancer regions that interact with target genes to promote H2BK120ub1-mediated gene regulation. The transcription factors bound at these regions from ChIP-Atlas analyses may be interesting for future follow-up studies.

Discussion

USP44 is a DUB that was first identified in a screen designed to uncover novel regulators of the metaphase-to-anaphase transition (59). We and others have shown that this enzyme plays a complex role in genomic instability, with its activity required for accurate mitosis and DNA repair (33, 46, 59–61). These data, combined with the observed increased incidence of spontaneous and carcinogen-induced tumors in mice lacking *Usp44*, lead to it being considered as a tumor suppressor gene (46, 61). USP44 has also been identified as a DUB that acts on histone H2B monoubiquitinated at lysine 120 (H2BK120ub1), as part of the N-CoR complex (44, 45). H2BK120ub1 occurs predominantly in the body of genes undergoing active transcription and is catalyzed by the ubiquitin (Ub) ligase complex formed by RNF20/RNF40 (62). H2BK120ub1, USP44, and N-CoR have been implicated in differentiation (44, 49, 56, 63, 64). The role of histone ubiquitination and its regulation in the pathogenesis of neuroblastoma have not been investigated. Here, we report a novel role of USP44 as a prognostic biomarker and oncogene in neuroblastoma through its ability to regulate histone ubiquitination—and promote a MYC-like gene expression program.

Neuroblastoma is an aggressive form of childhood cancer for which novel therapies are urgently needed to improve outcomes and reduce the toxic burden of therapy. In this report, we describe a previously undescribed role of the DUB USP44 in neuroblastoma biology. Risk stratification in this disease helps to guide patient therapy, reserving more intense therapy for those with higher-risk

diseases. Unfortunately, the stratification is imperfect, with many patients in the high-risk group who likely do not need intense therapy and conversely those in the low- or intermediate-risk groups who need more therapy (65). Our data show that tumors with elevated *USP44* levels are at high risk for relapse and that *USP44* is an independent prognostic factor in this disease (Figs. 1 and 2). This is despite the significant coexistence of established prognostic factors, including age >18 months, unfavorable histology, and *MYCN* amplification (Table 1). We found that *USP44* mRNA and protein were correlated in immunoblotting (Supplementary Fig. S4), but there is no clear way to identify *USP44* HI cases using standard pathologic techniques. However, our analysis of genomic alterations revealed that increased *USP44* expression correlated with disomy at chr. 4p (Fig. 5). An expression signature score derived from genes in this region was used as a surrogate to identify tumors with both copies of chr. 4p, we found that chr. 4 disomy is a powerful predictor of poor survival in *MYCN* nonamplified cases (Supplementary Fig. S10). However, only 70% of *USP44* HI cases can be identified with this method. Further work is therefore required to better identify cases within this new high-risk group.

Our previous work on this enzyme showed that USP44 is a non-essential enzyme, with mice lacking both copies of the gene born at Mendelian frequencies. With age, these mice develop tumors at an accelerated rate indicating that USP44 has tumor suppressor activity. Similar tumor suppressor and tumor-promoting effects have been reported for the H2B ubiquitin ligases RNF20 and RNF40 (66). The long latency, however, indicates that the short-term inhibition of USP44 that would be encountered during anticancer therapy may be well tolerated. Additionally, it is not clear whether the loss of USP44 after birth, or during the juvenile period, would have a similar tumor suppressive effect. Embryos and cells lacking USP44 have an increased rate of CIN and aneuploidy due to the mal-positioning of the centrosomes in the early phases of mitosis. This molecular phenotype is rescued by the replacement of WT USP44, but not a mutant enzyme that lacks a critical motif needed for binding to the centriole protein, centrin (46). Proteomics work has shown that this USP44-centrin complex is distinct from the complex that includes USP44 and the N-CoR subunits TBL1X and TBL1XR1, which regulates the ubiquitination status of histone H2B. This separation of function may help to explain—at least in part—how the same enzyme acts as a tumor suppressor and oncogene. Although appealing to think of these activities of mitosis regulator and epigenetic regulator as separate, *in vivo*, it is likely that both high and low levels of USP44 impact both processes. In support of this, we have also observed that elevated levels of USP44 lead to chromosome segregation errors and aneuploidy (34) and that overexpression of USP44 has opposite effects on the transcription of genes affected upon *Usp44* knockout (Fig. 6; Supplementary Fig. S10). Our analysis of clinical data here shows that many tumors with high USP44 also have a high CIN25 signature. The additive effect of high *USP44* and CIN producing worse outcomes than that seen with a high CIN25 signature alone (Supplementary Fig. S3), however, indicates that USP44 has CIN-independent effects on the biology of neuroblastoma.

In addition to its reported tumor suppressor activity, high levels of *USP44* have also been observed in certain cancers. We and others have reported overexpression in T-cell leukemia (33, 67). Others have found USP44 to promote malignant glioma, as well as gastric and prostate cancers (68–71). USP44 activity toward H2BK120ub1 is required for full transcription repression mediated by N-CoR, as depletion of USP44 reduced the repressive activity of N-CoR toward known targets (45). Histone H2B monoubiquitination regulates

processes including gene transcription and DNA damage repair (72). The addition of Ub to histone H2B at K120 is primarily catalyzed by the RNF20/40 E3 ligase that is recruited to chromatin through an interaction between the RNF20/40 adaptor protein WAC (WW domain-containing adapter with coiled-coil) and RNA polymerase II and coordinates transcription-coupled H2B ubiquitination (56, 62). Deregulation of H2BK120ub1 has been implicated in the regulation of early steps in organism development and in the pathogenesis of cancer. Alterations in the coding or copy numbers of the genes encoding the RNF20/40 complex, as well as H2B DUBs, have been commonly observed in several human cancers (72–74). Unlike many histone marks that cluster near the transcription start site of target genes, H2BK120ub1 is characteristically enriched downstream of the TSS in the transcribed regions of GB—a pattern that we also observed (48, 75). Increased H2BK120ub1 occupancy is associated with increased H3K4me3 and increased gene transcription (76, 77). Our RNA-seq and ChIP-seq data reinforce this, as transcripts that were upregulated in response to USP44 had increased H2BK120ub1 and H3K4me3 occupancy in *Usp44* null MEFs (Supplementary Fig. S12). Both the height and width of H3K4me3 peaks were increased for upregulated genes. Increased breadth of H3K4me3 occupancy has been linked with enhanced transcriptional elongation (78) downstream of H2BK120ub1 (55, 64, 79). Genes with particularly broad H3K4me3 peaks are enriched for tumor suppressor genes and additionally may mark genes involved in establishing cell identity (78, 80). In tumors with high USP44, the loss of H2BK120ub1 and the corresponding decrease in the peak and breadth of H3K4me3 may therefore lead to reduced expression of such genes.

In back-to-back publications, the Johnsen and Oren groups also reported that, during the differentiation of stem cells, levels of H2BK120ub1 increase with time and that deregulation of H2BK120ub1 homeostasis during this process disrupts differentiation (44, 56). The association of high USP44 with unfavorable histology in neuroblastoma (Fig. 2D) is consistent with this notion, as these tumors are poorly differentiated. Our finding that USP44 overexpression suppresses H2BK120ub1 and neuritic differentiation response to RA (Fig. 4) further bolsters this conclusion. Further supporting a critical role for reduced H2BK120ub1 in those tumors with high USP44 is our data showing that RNF20 depletion also promotes proliferation, migration, and invasion (Fig. 3). As recently summarized, H2BK120ub1 protein (81) or genomic alterations in RNF20/40 (72) have been observed in several cancers suggesting that similar mechanisms may occur broadly in malignancy. USP22 is also frequently altered in human cancer with deletions and overexpression seen (72). Given the distinct gene expression patterns observed between various H2BK120ub1 DUBs, it is not surprising that they may have different impacts on cancer pathogenesis.

We provide additional compelling data that it is the suppression of histone H2B ubiquitination that drives malignant behavior, as depletion of the ubiquitin ligase component RNF20 has a very similar phenotype as that seen upon USP44 overexpression (Fig. 3). Furthermore, we observed additional factors that are enriched at differentially bound H3K27ac regions in *Usp44* null MEFs (Supplementary Fig. S13C) that USP44 may directly recruit or set the epigenetic landscape for additional factors to bind. For example, MYC-associated feedback loops have been described in neuroblastoma (82–84). In *Usp44* null

MEFs, *Myc* expression is downregulated and was significantly enriched at H3K27ac regions that are lost upon USP44 loss. Thus, there may be a negative feedback loop in which downregulation of *Myc* leads to the loss of enhancer–promoter interactions, which normally drives a USP44–MYC-driven transcriptional program. Although USP44 OE in SH-SY5Y cells did not increase *MYC* expression, possibly due to the already high expression in these cells, USP44 may help promote MYC binding to enhancer regions to drive its respective transcriptional program and promote aggressiveness in neuroblastoma. Future chromatin interaction and association studies with USP44 and MYC will be needed to address these questions as well as ChIP-seq studies for USP44 to relate transcriptional changes to genome occupancy will be critical to complement our novel finding of USP44 mediating neuroblastoma aggressiveness.

We report that the histone H2B DUB USP44 drives malignant behavior and is an independent predictor of poor outcomes in childhood neuroblastoma. High levels of *USP44* drive neuroblastoma cell proliferation, migration, and invasion indicating that the poor outcomes are a direct effect of USP44 activity in this disease. The driving factor(s) leading to high levels of *USP44* are unclear though we have identified that tumors with two intact copies of chromosome 4p have higher expression of USP44 than those where this region is lost. USP44-driven regulation of histone H2B ubiquitination is an important mediator of these effects. Our data support the need for further investigation of the effects of USP44 inhibition as a strategy in neuroblastoma and into the safety of USP44 inhibition in animals. The data validate the need for the development of small-molecule inhibitors of this enzyme to facilitate these studies.

Authors' Disclosures

P.J. Galardy reports grants from Hyundai Hope on Wheels, Fraternal Order of Eagles, and Mayo-Sanford Collaborative Research Fund during the conduct of the study. No disclosures were reported by the other authors.

Authors' Contributions

T.L. Ekstrom: Investigation, writing–review and editing. **S. Hussain:** Data curation, investigation, writing–review and editing. **T. Bedekovics:** Data curation, investigation, writing–review and editing. **A. Ali:** Data curation, investigation, writing–review and editing. **L. Paolini:** Investigation, writing–review and editing. **H. Mahmood:** Investigation, writing–review and editing. **R.M. Rosok:** Resources, data curation, formal analysis, investigation. **J. Koster:** Conceptualization, resources, data curation, formal analysis, writing–review and editing. **S.A. Johnsen:** Conceptualization, data curation, supervision, methodology, writing–review and editing. **P.J. Galardy:** Conceptualization, resources, data curation, supervision, funding acquisition, writing–original draft, project administration, writing–review and editing.

Acknowledgments

This work was supported by grants from the Hyundai Hope on Wheels Foundation, the Mayo Clinic-Sanford Health Collaborative Research Fund, The Fraternal Order of Eagles, and the Mayo Clinic.

Note

Supplementary data for this article are available at Molecular Cancer Research Online (<http://mcr.aacrjournals.org/>).

Received June 5, 2023; revised April 5, 2024; accepted May 20, 2024; published first May 22, 2024.

References

- Marshall GM, Carter DR, Cheung BB, Liu T, Mateos MK, Meyerowitz JG, et al. The prenatal origins of cancer. *Nat Rev Cancer* 2014;14:277–89.
- Matthay KK, Maris JM, Schleiermacher G, Nakagawara A, Mackall CL, Diller L, et al. Neuroblastoma. *Nat Rev Dis Primers* 2016;2:16078.

3. Grobner SN, Worst BC, Weischenfeldt J, Buchhalter I, Kleinheinz K, Rudneva VA, et al. The landscape of genomic alterations across childhood cancers. *Nature* 2018;555:321–7.
4. Valentijn LJ, Koster J, Zwijnenburg DA, Hasselt NE, van Sluis P, Volckmann R, et al. TERT rearrangements are frequent in neuroblastoma and identify aggressive tumors. *Nat Genet* 2015;47:1411–4.
5. Schramm A, Koster J, Assenov Y, Althoff K, Peifer M, Mahlow E, et al. Mutational dynamics between primary and relapse neuroblastomas. *Nat Genet* 2015;47:872–7.
6. Pugh TJ, Morozova O, Attiyeh EF, Asgharzadeh S, Wei JS, Auclair D, et al. The genetic landscape of high-risk neuroblastoma. *Nat Genet* 2013;45:279–84.
7. Molenaar JJ, Koster J, Zwijnenburg DA, van Sluis P, Valentijn LJ, van der Ploeg I, et al. Sequencing of neuroblastoma identifies chromothripsis and defects in neurogenesis genes. *Nature* 2012;483:589–93.
8. Schleiermacher G, Janoueix-Lerosey I, Ribeiro A, Klijjanienko J, Couturier J, Pierron G, et al. Accumulation of segmental alterations determines progression in neuroblastoma. *J Clin Oncol* 2010;28:3122–30.
9. Janoueix-Lerosey I, Schleiermacher G, Michels E, Mosseri V, Ribeiro A, Lequin D, et al. Overall genomic pattern is a predictor of outcome in neuroblastoma. *J Clin Oncol* 2009;27:1026–33.
10. Mosse YP, Laudenslager M, Longo L, Cole KA, Wood A, Attiyeh EF, et al. Identification of ALK as a major familial neuroblastoma predisposition gene. *Nature* 2008;455:930–5.
11. Janoueix-Lerosey I, Lequin D, Brugieres L, Ribeiro A, de Pontual L, Combaret V, et al. Somatic and germline activating mutations of the ALK kinase receptor in neuroblastoma. *Nature* 2008;455:967–70.
12. George RE, Sanda T, Hanna M, Frohling S, Luther W 2nd, Zhang J, et al. Activating mutations in ALK provide a therapeutic target in neuroblastoma. *Nature* 2008;455:975–8.
13. Chen Y, Takita J, Choi YL, Kato M, Ohira M, Sanada M, et al. Oncogenic mutations of ALK kinase in neuroblastoma. *Nature* 2008;455:971–4.
14. Trochet D, Bourdeaut F, Janoueix-Lerosey I, Deville A, de Pontual L, Schleiermacher G, et al. Germline mutations of the paired-like homeobox 2B (PHOX2B) gene in neuroblastoma. *Am J Hum Genet* 2004;74:761–4.
15. Mosse YP, Laudenslager M, Khazi D, Carlisle AJ, Winter CL, Rappaport E, et al. Germline PHOX2B mutation in hereditary neuroblastoma. *Am J Hum Genet* 2004;75:727–30.
16. van Groningen T, Akogul N, Westerhout EM, Chan A, Hasselt NE, Zwijnenburg DA, et al. A NOTCH feed-forward loop drives reprogramming from adrenergic to mesenchymal state in neuroblastoma. *Nat Commun* 2019;10:1530.
17. van Groningen T, Koster J, Valentijn LJ, Zwijnenburg DA, Akogul N, Hasselt NE, et al. Neuroblastoma is composed of two super-enhancer-associated differentiation states. *Nat Genet* 2017;49:1261–6.
18. Henrich KO, Bender S, Saadati M, Dreidax D, Gartlgruber M, Shao C, et al. Integrative genome-scale analysis identifies epigenetic mechanisms of transcriptional deregulation in unfavorable neuroblastomas. *Cancer Res* 2016;76:5523–37.
19. Yang J, Altahan AM, Hu D, Wang Y, Cheng PH, Morton CL, et al. The role of histone demethylase KDM4B in Myc signaling in neuroblastoma. *J Natl Cancer Inst* 2015;107:djv080.
20. Gartlgruber M, Sharma AK, Quintero A, Dreidax D, Jansky S, Park YG, et al. Super enhancers define regulatory subtypes and cell identity in neuroblastoma. *Nat Cancer* 2021;2:114–28.
21. Westerhout EM, Hamdi M, Stroeken P, Nowakowska NE, Lakeman A, van Arkel J, et al. Mesenchymal-type neuroblastoma cells escape ALK inhibitors. *Cancer Res* 2022;82:484–96.
22. Sengupta S, Das S, Crespo AC, Cornel AM, Patel AG, Mahadevan NR, et al. Mesenchymal and adrenergic cell lineage states in neuroblastoma possess distinct immunogenic phenotypes. *Nat Cancer* 2022;3:1228–46.
23. Mabe NW, Huang M, Dalton GN, Alexe G, Schaefer DA, Geraghty AC, et al. Transition to a mesenchymal state in neuroblastoma confers resistance to anti-GD2 antibody via reduced expression of ST8SIA1. *Nat Cancer* 2022;3:976–93.
24. Thirant C, Peltier A, Durand S, Kramdi A, Louis-Brennetot C, Pierre-Eugene C, et al. Reversible transitions between noradrenergic and mesenchymal tumor identities define cell plasticity in neuroblastoma. *Nat Commun* 2023;14:2575.
25. Frumm SM, Fan ZP, Ross KN, Duval JR, Gupta S, VerPlank L, et al. Selective HDAC1/HDAC2 inhibitors induce neuroblastoma differentiation. *Chem Biol* 2013;20:713–25.
26. Hahn CK, Ross KN, Warrington IM, Mazitschek R, Kanegai CM, Wright RD, et al. Expression-based screening identifies the combination of histone deacetylase inhibitors and retinoids for neuroblastoma differentiation. *Proc Natl Acad Sci U S A* 2008;105:9751–6.
27. Gregoire-Mitha S, Gray DA. What deubiquitinating enzymes, oncogenes, and tumor suppressors actually do: are current assumptions supported by patient outcomes? *Bioessays* 2021;43:e2000269.
28. Bedekovics T, Hussain S, Zhang Y, Ali A, Jeon YJ, Galardy PJ. USP24 is a cancer-associated ubiquitin hydrolase, novel tumor suppressor, and chromosome instability gene deleted in neuroblastoma. *Cancer Res* 2021;81:1321–31.
29. Gentles AJ, Newman AM, Liu CL, Bratman SV, Feng W, Kim D, et al. The prognostic landscape of genes and infiltrating immune cells across human cancers. *Nat Med* 2015;21:938–45.
30. Hanemaaijer ES, Margaritis T, Sanders K, Bos FL, Candelli T, Al-Saati H, et al. Single-cell atlas of developing murine adrenal gland reveals relation of Schwann cell precursor signature to neuroblastoma phenotype. *Proc Natl Acad Sci U S A* 2021;118:e2022350118.
31. Kang HJ, Kawasaki YI, Cheng F, Zhu Y, Xu X, Li M, et al. Spatio-temporal transcriptome of the human brain. *Nature* 2011;478:483–9.
32. Attoff K, Johansson Y, Cediel-Ulloa A, Lundqvist J, Gupta R, Caiment F, et al. Acrylamide alters CREB and retinoic acid signalling pathways during differentiation of the human neuroblastoma SH-SY5Y cell line. *Sci Rep* 2020;10:16714.
33. Zhang Y, van Deursen J, Galardy PJ. Overexpression of ubiquitin specific protease 44 (USP44) induces chromosomal instability and is frequently observed in human T-cell leukemia. *PLoS One* 2011;6:e23389.
34. Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. *Nat Methods* 2012;9:357–9.
35. Danecek P, Bonfield JK, Liddle J, Marshall J, Ohan V, Pollard MO, et al. Twelve years of SAMtools and BCFtools. *Gigascience* 2021;10:giab008.
36. Ramirez F, Ryan DP, Gruning B, Bhardwaj V, Kilpert F, Richter AS, et al. deepTools2: a next generation web server for deep-sequencing data analysis. *Nucleic Acids Res* 2016;44:W160–5.
37. Robinson JT, Thorvaldsdottir H, Winckler W, Guttman M, Lander ES, Getz G, et al. Integrative genomics viewer. *Nat Biotechnol* 2011;29:24–6.
38. Feng J, Liu T, Qin B, Zhang Y, Liu XS. Identifying ChIP-seq enrichment using MACS. *Nat Protoc* 2012;7:1728–40.
39. Quinlan AR, Hall IM. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* 2010;26:841–2.
40. Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, et al. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* 2013;29:15–21.
41. Anders S, Pyl PT, Huber W. HTSeq—a Python framework to work with high-throughput sequencing data. *Bioinformatics* 2015;31:166–9.
42. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* 2014;15:550.
43. Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci U S A* 2005;102:15545–50.
44. Fuchs G, Shema E, Vesterman R, Kotler E, Wolchinsky Z, Wilder S, et al. RNF20 and USP44 regulate stem cell differentiation by modulating H2B monoubiquitylation. *Mol Cell* 2012;46:662–73.
45. Lan X, Atanassov BS, Li W, Zhang Y, Florens L, Mohan RD, et al. USP44 is an integral component of N-CoR that contributes to gene repression by deubiquitinating histone H2B. *Cell Rep* 2016;17:2382–93.
46. Zhang Y, Foreman O, Wigle DA, Kosari F, Vasmatzis G, Salisbury JL, et al. USP44 regulates centrosome positioning to prevent aneuploidy and suppress tumorigenesis. *J Clin Invest* 2012;122:4362–74.
47. Carter SL, Eklund AC, Kohane IS, Harris LN, Szallasi Z. A signature of chromosomal instability inferred from gene expression profiles predicts clinical outcome in multiple human cancers. *Nat Genet* 2006;38:1043–8.
48. Shema E, Tirosh I, Aylon Y, Huang J, Ye C, Moskovits N, et al. The histone H2B-specific ubiquitin ligase RNF20/hBRE1 acts as a putative tumor suppressor through selective regulation of gene expression. *Genes Dev* 2008;22:2664–76.
49. Hermanson O, Jepsen K, Rosenfeld MG. N-CoR controls differentiation of neural stem cells into astrocytes. *Nature* 2002;419:934–9.
50. Soldatov R, Kaucaka M, Kastriti ME, Petersen J, Chontorotzea T, Englmaier L, et al. Spatiotemporal structure of cell fate decisions in murine neural crest. *Science* 2019;364:eaas9536.
51. Jansky S, Sharma AK, Korber V, Quintero A, Toprak UH, Wecht EM, et al. Single-cell transcriptomic analyses provide insights into the developmental origins of neuroblastoma. *Nat Genet* 2021;53:683–93.
52. Yi H, Xue L, Guo MX, Ma J, Zeng Y, Wang W, et al. Gene expression atlas for human embryogenesis. *FASEB J* 2010;24:3341–50.

53. Depuydt P, Boeva V, Hocking TD, Cannoodt R, Ambros IM, Ambros PF, et al. Genomic amplifications and distal 6q loss: novel markers for poor survival in high-risk neuroblastoma patients. *J Natl Cancer Inst* 2018;110:1084–93.
54. Depuydt P, Koster J, Boeva V, Hocking TD, Speleman F, Schleiermacher G, et al. Meta-mining of copy number profiles of high-risk neuroblastoma tumors. *Sci Data* 2018;5:180240.
55. Xie W, Nagarajan S, Baumgart SJ, Kosinsky RL, Najafova Z, Kari V, et al. RNF40 regulates gene expression in an epigenetic context-dependent manner. *Genome Biol* 2017;18:32.
56. Karpiuk O, Najafova Z, Kramer F, Hennion M, Galonska C, Konig A, et al. The histone H2B monoubiquitination regulatory pathway is required for differentiation of multipotent stem cells. *Mol Cell* 2012;46:705–13.
57. McLean CY, Bristol D, Hiller M, Clarke SL, Schaar BT, Lowe CB, et al. GREAT improves functional interpretation of cis-regulatory regions. *Nat Biotechnol* 2010;28:495–501.
58. Zou Z, Ohta T, Miura F, Oki S. ChIP-Atlas 2021 update: a data-mining suite for exploring epigenomic landscapes by fully integrating ChIP-seq, ATAC-seq and Bisulfite-seq data. *Nucleic Acids Res* 2022;50:W175–82.
59. Stegmeier F, Rape M, Draviam VM, Nalepa G, Sowa ME, Ang XL, et al. Anaphase initiation is regulated by antagonistic ubiquitination and deubiquitination activities. *Nature* 2007;446:876–81.
60. Mosbech A, Lukas C, Bekker-Jensen S, Mailand N. The deubiquitylating enzyme USP44 counteracts the DNA double-strand break response mediated by the RNF8 and RNF168 ubiquitin ligases. *J Biol Chem* 2013;288:16579–87.
61. Zhang Y, Mandemaker IK, Matsumoto S, Foreman O, Holland CP, Lloyd WR, et al. USP44 stabilizes DDB2 to facilitate nucleotide excision repair and prevent tumors. *Front Cell Dev Biol* 2021;9:663411.
62. Zhu B, Zheng Y, Pham AD, Mandal SS, Erdjument-Bromage H, Tempst P, et al. Monoubiquitination of human histone H2B: the factors involved and their roles in HOX gene regulation. *Mol Cell* 2005;20:601–11.
63. Jepsen K, Hermanson O, Onami TM, Gleiberman AS, Lunyak V, McEvilly RJ, et al. Combinatorial roles of the nuclear receptor corepressor in transcription and development. *Cell* 2000;102:753–63.
64. Najafova Z, Liu P, Wegwitz F, Ahmad M, Tamon L, Kosinsky RL, et al. RNF40 exerts stage-dependent functions in differentiating osteoblasts and is essential for bone cell crosstalk. *Cell Death Differ* 2021;28:700–14.
65. Irwin MS, Naranjo A, Zhang FF, Cohn SL, London WB, Gastier-Foster JM, et al. Revised neuroblastoma risk classification system: a report from the children's oncology group. *J Clin Oncol* 2021;39:3229–41.
66. Wegwitz F, Prokakis E, Pejkovska A, Kosinsky RL, Glatzel M, Pantel K, et al. The histone H2B ubiquitin ligase RNF40 is required for HER2-driven mammary tumorigenesis. *Cell Death Dis* 2020;11:873.
67. Chi Z, Zhang B, Sun R, Wang Y, Zhang L, Xu G. USP44 accelerates the growth of T-cell acute lymphoblastic leukemia through interacting with WDR5 and repressing its ubiquitination. *Int J Med Sci* 2022;19:2022–32.
68. Nishimura S, Oki E, Ando K, Iimori M, Nakaji Y, Nakashima Y, et al. High ubiquitin-specific protease 44 expression induces DNA aneuploidy and provides independent prognostic information in gastric cancer. *Cancer Med* 2017; 6:1453–64.
69. Park JM, Lee JE, Park CM, Kim JH. USP44 promotes the tumorigenesis of prostate cancer cells through EZH2 protein stabilization. *Mol Cells* 2019;42: 17–27.
70. Xiang T, Jiang HS, Zhang BT, Liu G. CircFOXO3 functions as a molecular sponge for miR-143-3p to promote the progression of gastric carcinoma via upregulating USP44. *Gene* 2020;753:144798.
71. Zou Y, Qiu G, Jiang L, Cai Z, Sun W, Hu H, et al. Overexpression of ubiquitin specific proteases 44 promotes the malignancy of glioma by stabilizing tumor-promoter securin. *Oncotarget* 2017;8:58231–46.
72. Jeusset LM, McManus KJ. Characterizing and exploiting the many roles of aberrant H2B monoubiquitination in cancer pathogenesis. *Semin Cancer Biol* 2022;86:782–98.
73. 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.
74. Hoadley KA, Yau C, Hinoue T, Wolf DM, Lazar AJ, Drill E, et al. Cell-of-origin patterns dominate the molecular classification of 10,000 tumors from 33 types of cancer. *Cell* 2018;173:291–304 e6.
75. Minsky N, Shema E, Field Y, Schuster M, Segal E, Oren M. Monoubiquitinated H2B is associated with the transcribed region of highly expressed genes in human cells. *Nat Cell Biol* 2008;10:483–8.
76. Sun ZW, Allis CD. Ubiquitination of histone H2B regulates H3 methylation and gene silencing in yeast. *Nature* 2002;418:104–8.
77. Kim J, Hake SB, Roeder RG. The human homolog of yeast BRE1 functions as a transcriptional coactivator through direct activator interactions. *Mol Cell* 2005;20:759–70.
78. Chen K, Chen Z, Wu D, Zhang L, Lin X, Su J, et al. Broad H3K4me3 is associated with increased transcription elongation and enhancer activity at tumor-suppressor genes. *Nat Genet* 2015;47:1149–57.
79. Kosinsky RL, Zerche M, Kutschat AP, Nair A, Ye Z, Saul D, et al. RNF20 and RNF40 regulate vitamin D receptor-dependent signaling in inflammatory bowel disease. *Cell Death Differ* 2021;28:3161–75.
80. Benayoun BA, Pollina EA, Ucar D, Mahmoudi S, Karra K, Wong ED, et al. H3K4me3 breadth is linked to cell identity and transcriptional consistency. *Cell* 2014;158:673–88.
81. Marsh DJ, Ma Y, Dickson KA. Histone monoubiquitination in chromatin remodelling: focus on the histone H2B interactome and cancer. *Cancers (Basel)* 2020;12:3462.
82. Tan K, Mo J, Li M, Dong Y, Han Y, Sun X, et al. SMAD9-MYCN positive feedback loop represents a unique dependency for MYCN-amplified neuroblastoma. *J Exp Clin Cancer Res* 2022;41:352.
83. Nagy Z, Seneviratne JA, Kanikevich M, Chang W, Mayoh C, Venkat P, et al. An ALYREF-MYCN coactivator complex drives neuroblastoma tumorigenesis through effects on USP3 and MYCN stability. *Nat Commun* 2021;12:1881.
84. Marshall GM, Liu PY, Gherardi S, Scarlett CJ, Bedalov A, Xu N, et al. SIRT1 promotes N-Myc oncogenesis through a positive feedback loop involving the effects of MKP3 and ERK on N-Myc protein stability. *PLoS Genet* 2011;7: e1002135.