The BRD4–NUT Fusion Alone Drives Malignant Transformation of NUT Carcinoma

R. Taylor Durall[1], Julianna Huang[1], Luke Wojenski[2], Yeying Huang[1], Prafulla C. Gokhale[3], Brittaney A. Leeper[3], Joshua O. Nash[4,5], Pedro L. Ballester[4], Scott Davidson[4], Adam Shlien[4,5], Emmanuel Sotirakis[6], Fabien Bertaux[6], Vincent Dubus[6], Jia Luo[7,8], Catherine J. Wu[7,8], Derin B. Keskin[7,9,10], Kyle P. Eagen[11,12,13,14,15], Geoffrey I. Shapiro[7,8], and Christopher A. French[1]

ABSTRACT

NUT carcinoma (NC) is an aggressive squamous carcinoma defined by the BRD4–NUT fusion oncoprotein. Routinely effective systemic treatments are unavailable for most NC patients. The lack of an adequate animal model precludes identifying and leveraging cell-extrinsic factors therapeutically in NC. Here, we created a genetically engineered mouse model (GEMM) of NC that forms a *Brd4::NUTM1* fusion gene upon tamoxifen induction of *Sox2*-driven Cre. The model displayed complete disease penetrance, with tumors arising from the squamous epithelium weeks after induction and all mice succumbing to the disease shortly thereafter. Closely resembling human NC (hNC), GEMM tumors (mNC) were poorly differentiated squamous carcinomas with high expression of MYC that metastasized to solid organs and regional lymph nodes. Two GEMM-derived cell lines were developed whose transcriptomic and epigenetic landscapes harbored key features of primary GEMM tumors. Importantly, GEMM tumor and cell line transcriptomes co-classified with those of human NC. BRD4–NUT also blocked differentiation and maintained the growth of mNC as in hNC. Mechanistically, GEMM primary tumors and cell lines formed large histone H3K27ac-enriched domains, termed megadomains, that were invariably associated with the expression of key NC-defining proto-oncogenes, *Myc* and *Trp63*. Small-molecule BET bromodomain inhibition (BETi) of mNC induced differentiation and growth arrest and prolonged survival of NC GEMMs, as it does in hNC models. Overall, tumor formation in the NC GEMM is definitive evidence that BRD4–NUT alone can potently drive the malignant transformation of squamous progenitor cells into NC.

Significance: The development of an immunocompetent model of NUT carcinoma that closely mimics the human disease provides a valuable global resource for mechanistic and preclinical studies to improve treatment of this incurable disease.

Introduction

NUT carcinoma (NC) is a rare, aggressive subtype of squamous cell carcinoma defined by NUT fusion oncoproteins. NC most often arises from thoracic (50%) and head and neck (40%) sites in predominantly adolescents and young adults (the AYA gap, median age, ∼23, range, 0–83; ref. 1). With a median survival of 6.5 months, NC is the second-most aggressive solid tumor in humans (1, 2), yet is vastly under-diagnosed and under-recognized. The estimated annual U.S. incidence of NC, based on the sequential next-generation sequencing (NGS) of solid tumors (3), is approximately 1,080–3,600 cases. This incidence is more than 4-fold more prevalent than Ewing sarcoma (https://rarediseases.org/). Aside from surgical resection, appropriate for only a minority of patients, there are no routinely effective systemic treatments for NC. Thus, this disease represents an extreme unmet need (2).

Mechanistically, BRD4 is a BET family protein whose dual bromodomains bind acetyl-histones. BRD4, when fused to NUT, recruits the histone acetyltransferase, p300 and the BRD4–NUT–p300 complex forms enormous super-enhancer-like chromatin regions called megadomains (4–6). BRD4–NUT megadomains maintain expression of progrowth, antidifferentiation transcription factors including *MYC*, *SOX2*, and *TP63* (4, 7). BRD4–NUT also acts by reprogramming the chromatin 3D structure, resulting in interchromosomal DNA interactions between these lineage-defining transcription factors. These ultra-long-range contacts occur within a novel nuclear subcompartment, subcompartment M (8, 9). Despite compelling evidence that BRD4–NUT may alone drive this cancer, including the lack of recurrent additional oncogenic mutations (1, 10) and cases occurring in infants, it has not been experimentally proven that BRD4–NUT can alone drive malignant transformation.

Pathologically, NC exhibits poorly differentiated squamous morphology and IHC features (11). However, because NC arises from a wide variety of sites, and is nearly always locally advanced without

[1]Department of Pathology, Brigham and Women's Hospital, Harvard Medical School, Boston, Massachusetts. [2]BasePair, New York, New York. [3]Experimental Therapeutics Core and Belfer Center for Applied Cancer Science, Dana-Farber Cancer Institute, Boston, Massachusetts. [4]Program in Genetics and Genome Biology, The Hospital for Sick Children (SickKids), University of Toronto, Toronto, Ontario, Canada. [5]Laboratory of Medicine and Pathobiology, University of Toronto, Toronto, Ontario, Canada. [6]genOway, Lyon, France. [7]Department of Medical Oncology, Dana-Farber Cancer Institute, Harvard Medical School, Boston, Massachusetts. [8]Department of Medical Oncology, Brigham and Women's Hospital, Harvard Medical School, Boston, Massachusetts. [9]Broad Institute of MIT and Harvard, Cambridge, Massachusetts. [10]Translational Immunogenomics Laboratory, Dana-Farber Cancer Institute, Boston, Massachusetts. [11]Department of Molecular and Cellular Biology, Baylor College of Medicine, Houston, Texas. [12]Center for Precision Environmental Health, Baylor College of Medicine, Houston, Texas. [13]Stem Cells and Regenerative Medicine Center, Baylor College of Medicine, Houston, Texas. [14]Center for Cell and Gene Therapy, Baylor College of Medicine, Houston, Texas. [15]Dan L Duncan Comprehensive Cancer Center, Baylor College of Medicine, Houston, Texas.

Corresponding Author: Christopher A. French, Brigham and Women's Hospital/Harvard Medical School, New Research Building Rm. 630G, 77 Avenue Louis Pasteur, Boston, MA 02115. E-mail: cfrench@bwh.harvard.edu

Cancer Res 2023;83:3846–60

doi: 10.1158/0008-5472.CAN-23-2545

identifiable *in situ* lesions at presentation, the tissue of origin is not known.

Treatment of cancer with BET bromodomain inhibitors (BETi), small molecules that competitively inhibit the binding of BET bromodomains to chromatin (12), was first demonstrated experimentally and in humans by our group in NC (13, 14). The on-target activity of BETi in NC led to a new field investigating the role of BRD4 in cancer (15). However, it has become clear that monotherapy with BETi does not fully address NC biology. Unfortunately, apart from limited case reports, almost nothing is known about cell-extrinsic mechanisms of NC growth and metastasis, including for example the immune response to NC. The lack of an adequate animal model precludes the ability to leverage cell-extrinsic factors therapeutically in this disease. Thus, in response to the need for an immunocompetent animal model in which both NC cell-autonomous and -nonautonomous pathogenic mechanisms can be rigorously investigated, we sought to develop a genetically engineered mouse model (GEMM) of NC.

Herein, we present an immunocompetent genetically engineered animal model of NC. To validate the GEMM as representative of human disease, we compare pathologic, transcriptomic, epigenomic, and molecular pathogenetic features with those of human NC (hNC).

Materials and Methods

Mouse strains

All mouse research was conducted with the approval of the DFCI Institutional Animal Care and Use Committee (IACUC). The *Sox2*-CreER^{T2} [B6;129S-Sox2tm1(cre/ER^{T2})Hoch/J] model has been previously described (16). The methods for the creation of the mouse bearing the *BRD4–NUTM1* allele are presented below. *Sox2*-CreER^{T2}-expressing mice (Jackson Laboratories) were mated with *BRD4–NUTM1* allelic mice to heterozygosity and homozygosity for the *BRD4–NUTM1* allele. Genotyping was performed by Transnetyx using primers for the *Brd4tm1*-WT, *Brd4tm1*-FL, and *Sox2*-CreER^{T2} alleles listed in Supplementary Table S1. Control mice not treated with tamoxifen showed no abnormalities and normal fecundity. Mice were treated with tamoxifen (Sigma 74136) at 7–8 weeks of age (75 mg/kg) via single i.p. injection.

Creation of the transgenic Brd4–NUTM1 mouse allele

To generate an inducible *Brd4–NUTM1* knock-in mouse allele, the FLEx system was used. Generation of this mouse was performed in collaboration with genOway. A targeting vector containing the *NUTM1* exon 3-stop CDS in an inverted orientation was created. A codon-optimized version of the human *NUTM1* gene was used in this model to secure the splicing event (Supplementary Table S2) while conserving the amino acid sequence, thus allowing for detection by IHC, IF, or immunoblotting using a validated anti-NUT antibody (clone C52B1, cat no. 3625). The *NUTM1* exon 3-stop CDS is fused to the hGH polyA terminator sequence. These coding sequences were double-flanked by two wild-type LoxP sites and two modified LoxP sites (Lox2272). This vector further contained a neomycin-negative selection cassette that was flanked by RoxP excision sites. Finally, this insert was surrounded by homology domains to the genomic region spanning intron 13 to intron 16 of endogenous *Brd4* to allow for homologous recombination. The generated targeting vector was then electroporated into mouse embryonic stem cells and treated with neomycin for selection. The generated colonies were screened for and demonstrated homologous recombination with the insertion of the FLEx cassette in the correct orientation by PCR and whole locus sequencing. PCR primers were as follows: wild-type allele (220 bp): forward: TGGCATTTTCCATAGTGTAAGCTCAGGC, reverse: ATCTGGGTAGTGGGTAGACCTTGAGGACTG; inducible allele (334bp): forward: TGGCATTTTCCATAGTGTAAGCTCAGGC, reverse: ATGGTGATACAAGGGACATCTTCCCATTC; induced allele (329bp): forward: CTCCAACTCCTAATCTCAGGTGATC-TACCC, reverse: GTGTAAGCTCAGGCTACAGACAACAGAGC.

These recombined ES cell clones were subsequently injected into mouse E3.5 blastocysts, and the blastocysts were implanted into pseudopregnant females. The resultant chimeric mice were then bred with wild-type C57BL/6N mice to generate progeny containing the transgene allele. The RoxP-flanked neomycin cassette was removed by Dre excision, thus yielding the mature inducible. This line was further validated by whole locus sequencing. The steps during the mouse allele generation are shown graphically in Supplementary Fig. S1.

Mouse studies

All *in vivo* studies were conducted at Dana-Farber Cancer Institute with the approval of the IACUC in an Association for Assessment and Accreditation of Laboratory Animal Care International–accredited vivarium. Bioluminescent imaging was performed to detect tumor growth in mice by injecting D-luciferin subcutaneously at 75 mg/kg (Promega) and imaged with the IVIS Spectrum Imaging System (PerkinElmer). To quantify bioluminescence, identical regions of interest were drawn, and the integrated total flux of photons (the sum of the prone and supine values) using the Living Image software (PerkinElmer) was used for initial randomization of mice into various treatment groups and subsequently weekly imaging for assessing tumor response.

For *in vivo* efficacy studies, tumor induction was performed by treating mice with tamoxifen (Sigma 74136) at 7 to 8 weeks of age (75 mg/kg) via single i.p. injection. Five days after tamoxifen injection, mice were randomized into either vehicle control or ABBV-744 treatment groups. ABBV-744 was formulated in 2% DMSO, 30% PEG 400, and 68% Phosal-50PG, and administered once daily by oral gavage for 28 days. Mice were followed for survival based on morbidity criteria e.g., body weight loss of >15%, ruffled fur, lethargy, and impaired ambulation. Confirmation of tumor growth was made based on necropsy.

Tumor cell lines

mNC-derived cell lines 311E and 317E were created by finely mincing fresh mNC tumors and culturing in Epithelial Cell Medium (Cell Biologics) supplemented with 2% FBS (Sigma-Aldrich; batch number 17H392) + 1% pen–strep (GIBCO/Invitrogen) until passage 15. Both cell lines were then maintained as a monolayer culture in DMEM (Invitrogen) supplemented with 10% FBS, 1% pen–strep, and 1× GlutaMax (Gibco). Human NC line PER-403 was maintained as a monolayer in DMEM supplemented with 20% BGS (HyClone, lot number AG29548253), 1% pen–strep, and 1× GlutaMax. TC-797, 14169, HeLa, and NIH-3T3 cell lines were cultured as a monolayer in DMEM supplemented with 10% FBS, 1% pen–strep, and 1× GlutaMax. Cell lines 311E and 317E were authenticated by IF staining using antibodies to NUT and H3K27ac (Supplementary Tables S3–S4). The PER-403, TC-797, 14169, HeLa, and NIH-3T3 cell lines were authenticated by STR profiling. The TC-797, PER-403, and 14169 cell lines were further authenticated by fluorescent *in situ* hybridization as described (17), demonstrating rearrangement of the *NUTM1* and *BRD4* genes.

IHC

Formalin-fixed, paraffin-embedded tumor samples from mice were prepared using standard methods (18). All IHC was performed on the Leica Bond III automated staining platform. Antibodies and conditions used are listed in Supplementary Table S5.

Hemacolor cytologic staining

Adherent cells were grown on 25-mm glass coverslips (Fisherbrand, Fisher Scientific) in 6-well format, and stained using the MilliporeSigma Hemacolor Staining Kit for Microscopy according to the manufacturer's instructions (MilliporeSigma).

Immunoblotting

Cells were lysed in RIPA Buffer (50 mmol/L Tris-HCl, 250 mmol/L NaCl, 1% NP-40, 0.5% sodium deoxycholate, 0.1% SDS, 5 mmol/L EDTA) containing 250 mmol/L NaCl and Halt Protease and Phosphatase Inhibitor Cocktail (78445, Thermo Fisher Scientific) for 30 minutes at 4°C with rotation. The lysate was centrifuged at 16,100 × *g*, and the supernatant was collected. Immunoblotting was performed as described previously (19). Antibodies and conditions used are listed in Supplementary Table S6.

Immunofluorescence staining

Immunofluorescence was performed as described previously (5). Briefly, 311E, 317E, and NIH-3T3 cells grown on coverslips were fixed in 4% paraformaldehyde in PBS and permeabilized using 0.3% Triton X-100. The cells were then blocked in 5% normal goat serum (CST, 5425) prepared in PBS + 0.3% Triton X-100. The cells were incubated with primary antibody dissolved in 1% BSA + PBS-T for 1 hour at room temperature. Primary and secondary antibodies used are described in Supplementary Tables S3–S4. Samples were washed after primary and secondary antibody incubation. Coverslips were mounted onto glass slides with ProLong Gold Antifade with DAPI (CST) or SlowFade Glass Antifade Mountant (Invitrogen) for confocal analysis.

Cell growth assays (CellTiter-Glo)

Cells were seeded onto 96-well plates at a density of 300–1,500 cells per well in a total volume of 90 μL media. Compounds were delivered to sextuplet wells using an HP D300e digital dispenser (Hewlett Packard) at the ICCB-Longwood Screening Facility at Harvard Medical School. For long-term (>96 hours) treatments, cells were refed every four days using a Thermo Multidrop Reagent Dispenser (Thermo Fisher Scientific). Following a 96 hours or 10-day incubation at 37°C, cells were lysed, and wells were assessed for total ATP content using a commercial proliferation assay (CellTiter-Glo; Promega). Three biological replicates were performed. Estimates of IC_{50} were calculated by logistic regression (GraphPad Prism).

Plasmids, cloning, and viral transduction

pInducer20 blast blast-N-BRD4-NUT-C-HA lentiviral plasmid was created by Gibson assembly of gBlocks (synthesized by Integrated DNA Technologies, Inc.; IDT) of the *BRD4–NUT* coding sequence into pInducer20 blast (Addgene #109334) using Gateway LR Clonase II Enzyme Mix according to the manufacturer's instructions (Thermo Fisher Scientific).

To create lentivirus, 293T cells were transfected with vector DNA, pRev, pTat, pHIV Gag/pol, and pVSVG by transfection using Lipofectamine 2000 (Invitrogen). The virus was collected 48 hours after transfection and purified using a 45-μm pore filter. Polybrene (4 μg/mL, Sigma) was used to transduce HeLa cells, and the infected cells were subsequently selected with blasticidin (Gibco) for at least one week.

Chemicals and compounds

ABBV-744 and ABBV-075 were kindly provided by AbbVie, Inc. Dimethyl sulfoxide (DMSO) was purchased from Sigma-Aldrich.

RNA extraction and library preparation for RNA-seq

Whole RNA was also extracted from tissue or cultured cells using the RNeasy mini kit (Qiagen) with the following replicate structure: fresh treated or untreated cultured mNC cell lines, 311E and 317E, were sampled in biological triplicates; fresh control normal esophagus was sampled from three separate mice; frozen GEMM tumor tissue from 311E and 317E were sampled in singlicate.

Following extraction samples were treated with a TURBO DNA-free DNase kit (Invitrogen) and RNA Clean and Concentrator (Zymo Research). 1 μg of total RNA was used for the construction of ribosome-depleted sequencing libraries using the KAPA RNA HyperPrep Kit with RiboErase kit (HMR, Roche) according to the manufacturers' instructions. ERCC ExFold RNA spike-in controls (Invitrogen; 4456739) were added according to the manufacturer's instructions to 1 μg of purified RNA from each sample to help normalize expression. Libraries were 50-bp paired-end sequenced on a NovaSeq X plus by Admera Health.

RNA-seq data analysis

The raw fastq data were analyzed using Basepair software (https://www.basepairtech.com/) using the following steps. Raw reads were trimmed using fastp (20) to remove both low-quality bases from reads (quality <10) and adapter sequences. Trimmed reads with a minimum length of 15 bp were aligned to the mouse genome assembly GRCm39/mm39 with refseq gene annotations and the ERCC spike-in sequences using STAR (21) with default parameters. Alignments were sorted and indexed using samtools (22) and ERCC spike-in normalized bigwig files were created for visualization using deeptools (23) with a bin size of 5bp. Read counts were collected at gene annotations using featureCounts and data were visualized using the IGV browser (24). Replicate-averaged bedGraph, bigwig, and TDF files were generated using a combination of bedTools (25), kentUtils (https://hgdownload.soe.ucsc.edu/admin/exe/), wiggletools (26), and igvTools (24). Heat maps and average profiles were generated using deepTools (23). Differential expression analyses were performed using DESeq2 (27) on genes with an average spike-in normalized expression level >7. Genes with adjusted *P* values <0.05 and fold change >2 were identified as differentially expressed. Principal component analysis (PCA) was done on all genes after filtering by average spike-in normalized expression level >7. Gene set enrichment analysis (GSEA) of gene ontology (GO) Hallmark gene sets were performed using the R packages clusterProfiler (28) and fgsea (bioRxiv 2016.06.20.060012), before combining results and plotting heat maps of select terms with hierarchical clustering using the gplots R package (https://github.com/talgalili/gplots). Venn diagrams of differentially expressed genes were created using the R package VennDiagram (https://cran.r-project.org/package=VennDiagram).

OTTER classification of human and mouse transcriptomes

Human hNC RNA-seq data were downloaded from the Sequencing Read Archive through the NCBI. FASTQs were processed using the STAR + RSEM Toil RNA-seq pipeline (29) and the resulting $\log_2$ transcripts per million normalized gene-expression counts were inputted into OTTER, an ensemble convolutional neural net classifier for cancer and normal transcriptomes (30). OTTER scores range from 0 to 1, with 0 being no similarity and 1 being full similarity to a cluster. OTTER assignments for all clusters with a score >0.05 were reported.

Mouse FASTQs were processed using the STAR + RSEM Toil RNA-seq pipeline. Mouse and human orthologs from HUGO were used to convert mouse gene counts to their human counterparts (31). We applied the following rules for ortholog conversion: when multiple

mouse genes mapped to a single human gene, the counts were averaged; and when a single mouse gene mapped to multiple human genes, the counts were applied without alteration to each human gene. To classify nonhuman tissues more effectively, OTTER was trained with data augmentation strategies to yield a more generalized model for out-of-distribution samples. The data augmentation pipeline included rescaling and merging samples (arXiv 2018.1710.09412). Human ortholog counts were then run through this augmented OTTER to classify them according to human tumor/normal classes. Due to this optimization, this version of OTTER outputs higher overall scores, so OTTER results for mice are only reported when scores >0.25.

All OTTER clusters and their characteristics mentioned in this publication are described in Supplementary Table S7.

siRNA knockdown

siRNAs to the following target sequences were purchased from IDT and transfected at a concentration of 50 nmol/L by electroporation (Lonza) into 2e6 cells using the program, X-005: negative control DsiRNA (#51-01-14-03), si*NUT* #1 (CCUCUGAGAUCUAAAGAGAAUCAAG), si*NUT* #4 (UGAGAUCUAAAGAGAAUCAAGAGCA).

Cleavage under targets and release using nuclease

Cleavage under targets and release using nuclease (CUT&RUN) reactions were conducted using cultured GEMM-derived cells lines, 311E and 317E, fresh tumor tissue from mouse number 342, and fresh normal stomach tissue collected from two *Sox2-Cre-ER*$^{T2\ -/-}$; *Brd4-NUTM1*$^{+/-}$ control mice. The initial procedure for cell lines and fresh tissue differed as follows.

Cell lines: Cells cultured in biological duplicate were cross-linked *in situ* with 0.1% formaldehyde for 1 minute. Fixation was quenched with 125 mmol/L glycine. Cells were washed twice with HBSS (Gibco) and then incubated with TypLE Express (Gibco) for 10–20 minutes to detach cells. Detached cells were centrifuged at 1,000 × *g* for 3 minutes and washed twice in wash buffer (20 mmol/L HEPES K-OH pH 7.9, 150 mmol/L NaCl, 0.5 mmol/L Spermidine, 1× cOmpleteTM, Mini, EDTA-free Protease Inhibitor; Roche).

Fresh tissue: 100 mg of tissue was finely minced on ice before being suspended in 2 mL fixation buffer (PBS + 1:100 Protease Inhibitor Cocktail (Thermo Scientific) + 0.1% formaldehyde) and vortexed. Tissue was incubated at room temperate for 2 minutes before adding 125 mmol/L glycine to quench fixation. Fixed tissue was centrifuged at 2,000 × *g* for 5 minutes at 4°C to pellet tissue, then washed twice in 2 mL PBS + 1:100 protease inhibitor cocktail (Roche). Following the last wash, tissue was resuspended in 2 mL wash buffer and disaggregated into a single-cell suspension by dounce homogenization.

For reactions with cell lines or fresh tissue, 500,000 cells per reaction were counted using a Cellometer Mini (Nexcelom). Cells were incubated with Concanavalin A-coated magnetic beads (EpiCypher) for 10 minutes at room temperature. Bead-bound cells were permeabilized in antibody binding buffer (wash buffer + 2 mmol/L EDTA + 0.0025% digitonin). Samples were incubated with anti-H3K27ac antibody (1:100, Cell Signaling Technologies; #8173) or rabbit IgG (1:50, EpiCypher 13-0042K) overnight at 4°C with nutation. For samples incubated with rabbit IgG, 1 μL of SNAP-CUTANA K-MetStat Panel (EpiCypher #19-1002) was added to each sample. Bead-bound cells were washed twice with 50 μL digitonin buffer (wash Buffer + 0.0025% digitonin), and then incubated with 2.5× pAG-MNase (EpiCypher, #15-1016) in 50 μL Digitonin Buffer for 1 hour at 4°C with nutation. Bead-bound cells were washed twice in 50 μL digitonin buffer and then incubated with 1 μL, 100 mmol/L $CaCl_2$ in 50 μL digitonin buffer for 30 minutes at 4°C. *E. Coli* spike-in DNA (EpiCypher, #18-1401, Lot #: 22274005-09) was added to Stop buffer (340 mmol/L NaCl, 20 mmol/L EDTA, 4 mmol/L EGTA, 50 μg/mL RNase A, 50 μg/mL glycogen, 0.02% digitonin) so each sample received 0.5 ng of DNA. 50 μL Stop Buffer + *E. Coli* spike-in was added to each sample, and then samples were incubated at 37°C for 10 minutes. Following the stop reaction, 1 μL 10% SDS and 1.2 μL proteinase K were added to samples and then incubated at 55°C overnight to reverse cross-link. DNA was extracted and concentrated using Oligo Clean and Concentrator (Zymo Research, D4060). Sample yield was determined by Qubit Fluorometer 2.0 with the 1× dsDNA HS Assay kit (Invitrogen).

Immunoprecipitated DNA was quantified with Qubit 2.0 DNA HS Assay (Thermo Fisher) and quality was assessed by Tapestation genomic DNA Assay (Agilent Technologies). Library preparation was performed using a KAPA Hyper Prep kit with PCR (Roche) per the manufacturer's recommendations. Library quality and quantity were assessed with Qubit 2.0 DNA HS Assay (Thermo Fisher), Tapestation High Sensitivity D1000 Assay (Agilent Technologies), and QuantStudio 5 System (Applied Biosystems). Illumina 8-nt unique dual-indices were used. Equimolar pooling of libraries was performed based on QC values and sequenced on an Illumina NovaSeq X Plus 10B (Illumina) with a read length configuration of 150 PE for 8M PE reads (4M in each direction).

CUT&RUN data analysis

The raw fastq data were analyzed using Basepair software (https://www.basepairtech.com/) using the following steps. Raw reads were trimmed using fastp (20) to remove both low-quality bases (quality <20) and adapter sequences. Reads less than 15 bp were discarded, and remaining reads were aligned with Bowtie2 (32) to the mouse genome assembly GRCm39/mm39 and the E.coli genome using "very-sensitive" mode. Alignments with MAPQ <30 were removed before sorting and indexing using samtools (22), deduplicated using sambamba (33), and mitochondrial reads were removed. Reads were converted to BED format and genome-wide read coverage was computed using bedTools (5). The insert-size distribution was calculated using the Picard toolkit (http://broadinstitute.github.io/picard/). Replicate-averaged bedGraph, bigwig, and TDF files were generated using a combination of bedTools (25), kentUtils, wiggletools (26), and igvTools (24), and data were visualized using the IGV browser (24). Peaks were called with epic2 (34) using a bin size of 350 bp and a gap number of 3, before identifying domains using the rank-ordering of super-enhancers (ROSE; ref. 35) algorithm to a threshold based on peak signal and size. Domains within 25 kb were merged and a list of high-confidence domains was created per biological condition by retaining only domains sharing overlap between biological replicates. Heat maps and average profiles were generated using deepTools (23). Venn diagrams were generated using the R package ChIPpeakAnno (36). For transcriptional start site (TSS) profiles, deepTools (23) was used to compute the count matrices for the TSS within 500 kb flanking regions of large domains, and for those not (TSS ± 1 Kb, reference-point TSS); R was used to plot the results. Heat maps were generated by deepTools (plotHeatmap) from count matrices (region ± 1 Kb, scale-regions) for all regions for each mark per cell line. The ideogram was created using bedtools to window the genome and count domains in windows, and then the ggplot2 R package (http://public.ebookcentral.proquest.com/choice/publicfullrecord.aspx? $P = 511468$) was used to plot the windows.

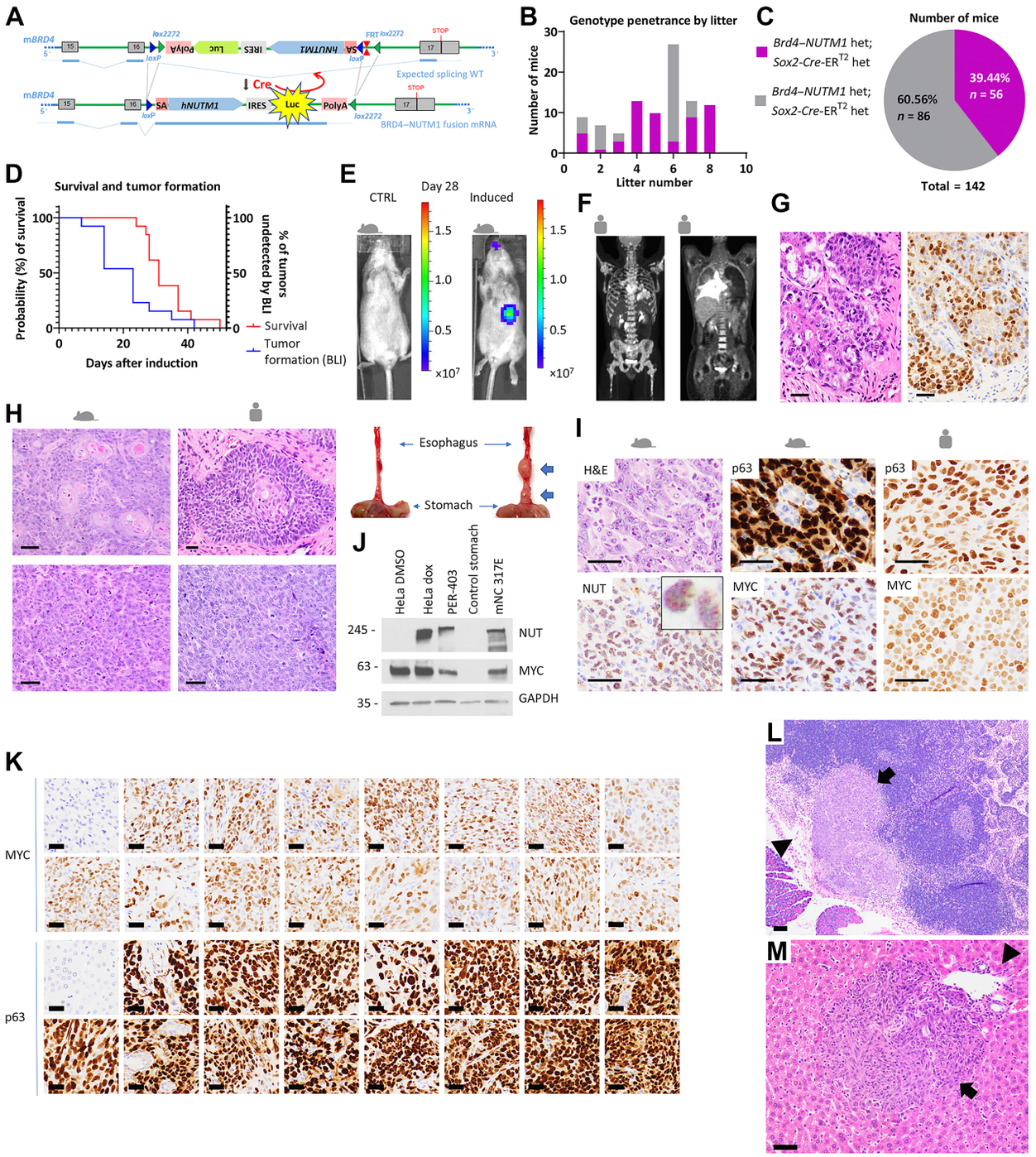


Figure 1.

Creation of an inducible, highly penetrant genetically engineered mouse model of NUT carcinoma. **A,** Schematic of NC GEMM design. **B** and **C,** Genotype penetrance by litter (**B**) and for all mice (**C**). het, heterozygote. **D,** Kaplan–Meier estimated survival and tumor formation ($n = 13$ mice). **E,** Involvement of esophagus and GEJ by tumors in NC GEMM. Bioluminescence images of mice (top) and corresponding gross photos (bottom) of the esophagus and stomach. Control (left) is female *Sox2* (neg); *Brd4–NUTM1* (het) mouse, and induced is *Sox2* (het); *Brd4–NUTM1* (het) GEM with separate tumors in the upper esophagus and GEJ. Thick blue arrows, tumors. Heat map scale units are radiance in p/sec/cm^2/sr. **F,** Images of human thoracic NC for comparison with **E**. Left, coronal slice from 3D reconstructed CT scan with contrast showing NC involving left lung and mediastinum. Right, coronal slice from PET-CT scan showing NC involving mediastinum and right lung. In both cases, it is difficult to determine where the tumor arose due to extensive invasion of multiple structures. **G,** *In situ* carcinoma was identified within the esophageal squamous epithelium at the GEJ one week following induction of *Brd4–NUTM1* fusion in an NC GEM. Left, hematoxylin and eosin; right, NUT IHC corresponding to left. Scale bar, 50 μm. (*Continued on the following page.*)

Confocal microscopy procedures and analysis

311E and 317E cells were prepared for immunofluorescent staining as described above. Images were acquired on a Zeiss Axio Observer Z1 single-point scanning confocal microscope with a linear encoded motorized stage from Zeiss, an LSM 980 scan head, and Toptica iChrome MLE laser CW launch using the Airyscan2 super-resolution external detector. A Zeiss Plan Apo 63x/1.4 DIC III oil immersion objective lens was used for the imaging. Images were acquired using a galvo, unidirectional scan with a pixel dwell time of 0.75 µs, 2× line averaging, and a pixel size of 42 nm × 42 nm (set as optimal for Airyscan detection) by Zen Blue Acquisition Software and exported as CZI files. For all images, the pinhole diameter was 5AU (set as optimal for Airyscan detection). Alexa Fluors were excited with either 493 nm or 590 nm solid state diode lasers with AOTF modulation, and emissions were collected on a GaAsP-PMT detector with collection windows of 495–550 nm and 607–735 nm, respectively. Super-resolution images were generated by Airyscan SuperResolution 2D automatic processing (Weiner noise filtering value range 6.7–7 as optimally determined by the Zen software). Laser power and detector settings for each sample are listed in Supplementary Table S8. Images cropped to isolate individual nuclei and 2D fluorescence intensity profiles for each channel were generated using Fiji ImageJ.

Data availability

RNA-seq and CUT&RUN data generated in this study are publicly available in Gene-Expression Omnibus (GEO) at accession number GSE241477. All other raw data are available upon request from the corresponding author. The data analyzed in this study were obtained from Clinical Sequencing Evidence-Generating Research at https://cser-consortium.org/; and the GEO at GSE142481 (https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSM4230244) and GSE228533 (https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE228533).

Results

Creation of an inducible, conditional expression model of *Brd4–NUTM1* fusion

Both the level of BRD4–NUT expression and the cell type in which it is expressed are critical to avoid toxicity from mis-expression. We thus endeavored to create a conditional inducible model where BRD4–NUT levels are dictated by the native *Brd4* promoter, and BRD4–NUT can be induced using a *Cre*-driver expressed in tissues predicted to give rise to NC. We used *Sox2* for the *Cre*-driver because it is pathogenetically linked to BRD4–NUT function and is expressed highly in NC (37). Moreover, the tissues and cell types in which SOX2 is expressed (adult stem cells of stomach, esophagus, bronchi, and trachea) overlap with those from which NC is presumed to arise, i.e., upper respiratory tract epithelium (1, 16). Our GEMM contains a heterozygous knock-in of human *NUTM1* in inverse orientation within the mouse homolog of the canonical breakpoint intron of murine *Brd4* (intron 16; **Fig. 1A**; Supplementary Fig. S1A and S1B). Intraperitoneal tamoxifen injection-induced *Sox2*-promoter-driven Cre recombinase leads to an inversion of *NUTM1*, encoding an in-frame *Brd4::NUTM1* fusion. A bicistronic luciferase gene allows for the visualization of tumor formation by bioluminescence imaging (BLI).

Brd4::NUTM1 homozygous mice are developmentally normal. When crossed with *Sox2*-CreER^{T2} heterozygotes (het), healthy, developmentally normal *Sox2*-CreER^{T2} (het); *Brd4::NUTM1* (het) offspring are born at an overall frequency of approximately 40% (**Fig. 1B** and **C**).

Tumor formation in the GEMM is highly penetrant and rapidly progressive

In our initial cohort of 13 mice, 100% developed tumors, and all died from the disease within 50 days following induction with tamoxifen (range, 22–50 days, median 31 days; **Fig. 1D**). Tumors are first detected by BLI 1–6 weeks following tamoxifen induction (**Fig. 1D**). In all cases, tumor BLI signal occurs within the thoracic region (**Fig. 1E**; Supplementary Fig. S2A), and at necropsy, these tumors consistently arise within the esophagus, affecting all regions, including gastroesophageal junction (GEJ; Supplementary Fig. S2B). Multiple tumors are often seen (**Fig. 1E**; Supplementary Fig. S2B). Many tumors that arise in the GEJ invade into gastric mucosa mimicking a gastric primary (Supplementary Fig. S2B).

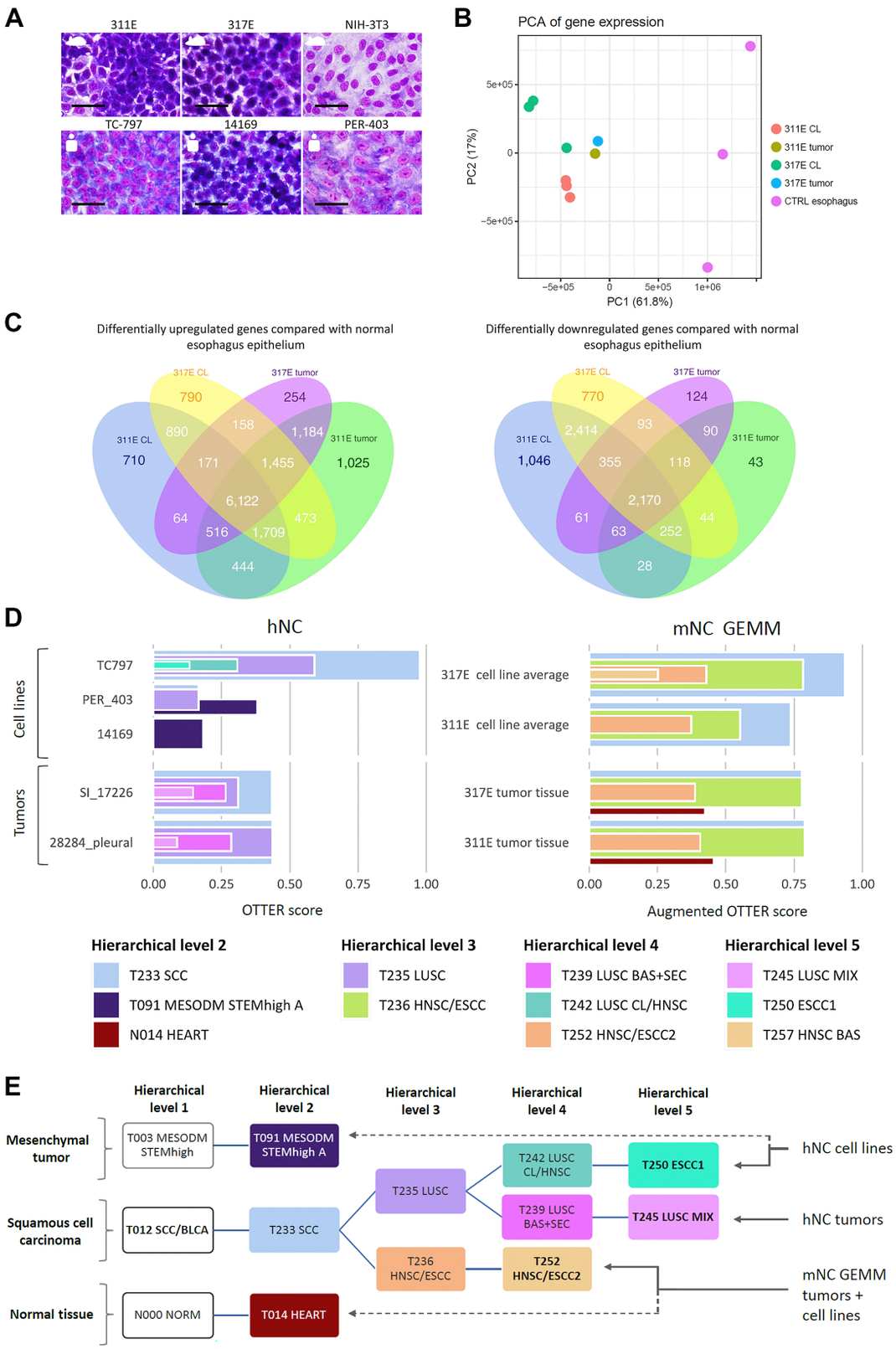
mNC arises from squamous epithelium

Like our GEMM tumors, heretofore referred to as mNC, human NC (hNC) most often arises in the thorax (**Fig. 1F**). In contrast to the GEMM tumors, hNC is believed to derive from the respiratory tract. However, most cases are extensively invasive (**Fig. 1F**) and lack *in situ* precursor lesions; thus, the primary tissue of origin has never been established. To determine the tissue of origin in NC genetically engineered mice (GEM), mice were euthanized seven days following tamoxifen induction, prior to the development of detectable lesions by BLI, and pathologic examination of all tissues and organs was evaluated. In a subset of these mice, we identified noninvasive, microscopic squamous carcinoma *in situ* mNC lesions arising in esophageal squamous epithelium (**Fig. 1G**). This finding demonstrates that mNC is of squamous epithelial origin and suggests that it may share a common cell lineage with hNC, a squamous cell carcinoma.

The histopathology of mNC closely resembles that of hNC

Pathologic examination of mNCs reveals that all are poorly differentiated squamous carcinomas with striking overlap with the histopathologic spectrum of hNCs. Two patterns were evident in mNCs that are classic features of hNC. These include the presence of focal, abrupt squamous differentiation with keratinization (compare mNC with hNC in **Fig. 1H**). Squamous differentiation is recognized by an

(*Continued.*) **H,** Histologic comparison of mNC and hNC on hematoxylin and eosin–stained tissue. Top, focal abrupt squamous differentiation is seen in mNC and hNC. Bottom, sheet-like growth pattern of monomorphic poorly differentiated carcinoma in mNC and hNC. Scale bar, 50 µm. **I,** Top left, Hematoxylin and eosin (H&E) staining of invasive mNC involving GEJ. Anti-MYC, -NUT, and -p63 IHC of mNC, and MYC and p63 IHC of hNC as indicated. Anti-MYC and -p63 antibodies recognize both human and mouse proteins. The NUT antibody recognizes human NUT, which is expressed in the NC GEMM fusion. Inset, cropped and zoomed section of NUT IHC image demonstrating BRD4–NUT nuclear foci. Scale bar, 50 µm. **J,** Immunoblot from indicated lysates. mNC 317E, primary tumor taken from the esophagus (GEJ). Normal stomach control tissue was taken from a SOX2–CRE-ER^{T2} (neg) BRD4–NUT (het) mouse. HeLa cells stably transduced with pInducer 20 Blast-BRD4–NUT treated with DMSO or doxycycline for 48 hours were used as positive control. PER-403, human NC endogenous BRD4–NUT-positive control. **K,** p63 and MYC IHC in 15 different mNC primary tumors. Negative control stomach (for MYC) and liver (for p63) tissue is shown at top left for each antibody. Scale bar, 25 µm. Note, 16 tumors were stained showing the same results, but only 15 are shown for space considerations. **L,** Hematoxylin and eosin staining of mNC metastatic to the subcapsular sinus of peripancreatic lymph node. Arrow, metastatic tumor. Arrowhead, subcapsular sinus. Left lower corner is normal pancreatic parenchyma. Scale bar, 100 µm. **M,** Hematoxylin and eosin staining of mNC metastatic to portal region of liver. Arrow, metastatic tumor. Arrowhead, portal tract. Scale bar, 50 µm.


A
311E
317E
NIH-3T3
TC-797
14169
PER-403
B
PCA of gene expression
PC2 (17%)
PC1 (61.8%)
311E CL
311E tumor
317E CL
317E tumor
CTRL esophagus
C
Differentially upregulated genes compared with normal esophagus epithelium
Differentially downregulated genes compared with normal esophagus epithelium
317E CL
317E tumor
311E CL
311E tumor
790
254
158
890
1,184
710
171
1,455
1,025
6,122
64
473
516
1,709
444
770
124
93
2,414
90
1,046
355
118
43
2,170
61
44
63
252
28
D
hNC
mNC GEMM
Cell lines
Tumors
TC797
PER_403
14169
SI_17226
28284_pleural
317E cell line average
311E cell line average
317E tumor tissue
311E tumor tissue
0.00
0.25
0.50
0.75
1.00
OTTER score
Augmented OTTER score
Hierarchical level 2
T233 SCC
T091 MESODM STEMhigh A
N014 HEART
Hierarchical level 3
T235 LUSC
T236 HNSC/ESCC
Hierarchical level 4
T239 LUSC BAS+SEC
T242 LUSC CL/HNSC
T252 HNSC/ESCC2
Hierarchical level 5
T245 LUSC MIX
T250 ESCC1
T257 HNSC BAS
E
Hierarchical level 1
Hierarchical level 2
Hierarchical level 3
Hierarchical level 4
Hierarchical level 5
Mesenchymal tumor
T003 MESODM STEMhigh
T091 MESODM STEMhigh A
Squamous cell carcinoma
T012 SCC/BLCA
T233 SCC
T235 LUSC
T242 LUSC CL/HNSC
T239 LUSC BAS+SEC
T250 ESCC1
T245 LUSC MIX
T236 HNSC/ESCC
T252 HNSC/ESCC2
Normal tissue
N000 NORM
T014 HEART
hNC cell lines
hNC tumors
mNC GEMM tumors + cell lines

increase in size of tumor cells, which accumulate large amounts of eosinophilic, cytoplasmic keratin intermediate filaments. The second histologic pattern is a sheet-like growth of monotonous appearing, poorly differentiated, yet medium-sized cells that lack the pleomorphism and enlargement seen in smoking-related squamous cancers (**Fig. 1H**, bottom).

Key targets of hBRD4–NUT, *MYC* and *TP63*, are also upregulated in mNC

mNC resembles hNC immunohistochemically, exhibiting diffuse, strong expression of MYC and p63 proteins (**Fig. 1I**). MYC is a canonical target of hBRD4–NUT and hNC is considered an MYC-driven tumor, one reason these tumors are likely poorly differentiated and fast growing (4, 7). p63, also a conserved target of hBRD4–NUT, is a squamous-lineage transcription factor required for hNC viability (4). By western blot analysis, mNC tumors express BRD4–NUT protein (mBRD4–NUT) at the expected size compared with recombinant (expressed in HeLa cells) or endogenous (PER-403 hNC cell line) hBRD4–NUT. Moreover, we found that MYC was upregulated in mNC tissue compared with adjacent normal gastric epithelium (**Fig. 1J**). As in hNC, we found that coexpression of MYC and p63 is a highly conserved feature of mNC; of 16 different mNC tumors tested, all expressed high levels of p63 (>95% of nuclei) and MYC (≥75% of nuclei; **Fig. 1K**). The strong morphologic and IHC similarities between mNC and hNC further suggest common pathogenetic mechanisms.

mNC, like hNC, exhibits a pattern consistent with metastases to lymph nodes and solid organs

hNC frequently metastasizes to lymph nodes and solid organs, such as liver, lung, and ovary. Likewise, we found that mNC exhibited frequent presumptive metastases ($n/N = 5/12$ mice, 42%), either to regional lymph nodes ($n/N = 3/12$ mice, 25%) or solid organs (liver, lung, and omentum, $n/N = 3/12$ mice, 25%; **Fig. 1L** and **M**). Although the pattern of tumors present in the subcapsular sinuses of lymph nodes and portal tracts of liver is consistent with metastatic spread, a molecular analysis would confirm and identify genetic factors that predispose these tumors to metastases. These findings indicate that the aggressive biological behavior of mNC is similar to that of hNC, especially considering that these metastases occurred within less than five weeks of tumor formation.

Two mNC-derived cell lines were created that share transcriptional profiles with tumor tissues, characterized by MYC pathway upregulation

We developed two cell lines derived from mNCs arising from separate mice, termed 311E and 317E. Like hNC cell lines, these are highly proliferative (doubling time ∼36 hours), adherent, poorly differentiated epithelioid cells (**Fig. 2A**). These cell lines can potentially serve as a valuable *in vitro* experimental model to test biological hypotheses and screen therapeutic interventions for *in vivo* investigation in the GEMM. To determine whether these cell lines are representative of the tumor tissue from which they are derived, we performed RNA sequencing (RNA-seq). Replicate reproducibility was confirmed based on PCA, which also revealed that the two cell lines showed more similarity to the tumor tissues from which they are derived than with control esophageal mucosa (**Fig. 2B**). Indeed, roughly one-third of all differentially expressed (DE) genes [upregulated: $n/N = 6{,}122/15{,}965$ (38.3%); downregulated: $n/N = 2{,}170/7{,}761$ (28.3%)] relative to control esophageal mucosa are shared amongst the two cell lines and their parental tumor tissues (**Fig. 2C**). GSEA of these DE genes reveal common transcriptional programs of MYC target gene upregulation and metabolic reprogramming (e.g., decreased fatty acid metabolism, adipogenesis, and oxidative phosphorylation; ref. 38) in both mNC tumor tissue and cell lines (Supplementary Fig. S3). These findings are consistent with the mechanism of BRD4–NUT to upregulate MYC and associated programs to drive the growth of NC (7, 39, 40) and highlight the potential translatability of the mNC cell line biology to that of the NC GEMM and ultimately hNC.

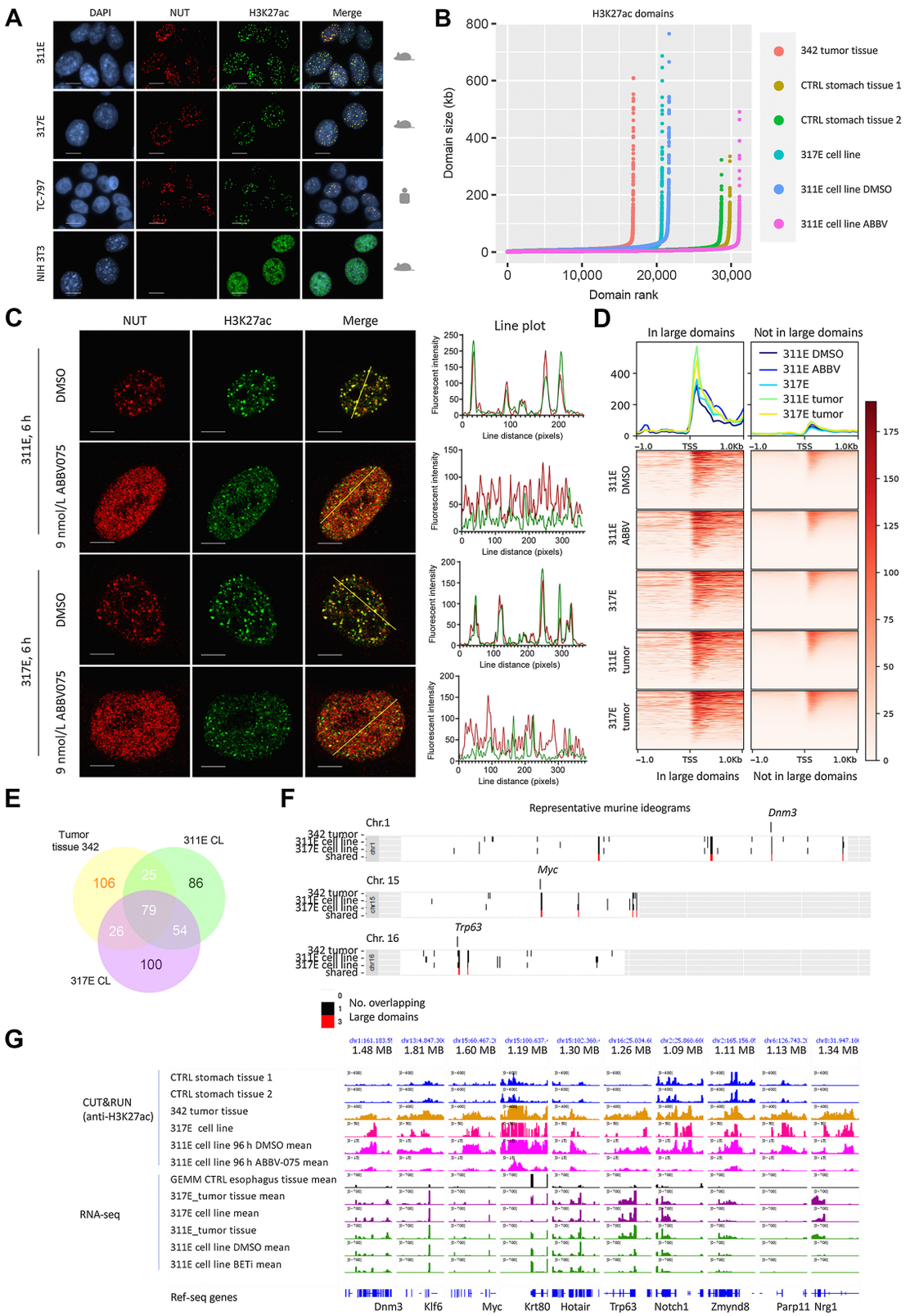
Human and GEMM NC cell lines and tissue co-classify as squamous carcinomas

hNC is considered a squamous cancer based on squamous differentiation observed in tumors and the expression of p63 (**Fig. 1H–K**; ref. 41). However, squamous differentiation can generally be identified in nonsquamous tumor types and thus does not establish a tumor as being squamous origin. Thus, we ran the transcriptomes of two available hNC primary tumors (28284, from our collection, and MCTP_SI_17226) and three hNC cell lines (14169, PER-403, and TC-797) through OTTER, an ensemble convolutional neural net classifier trained on the pan-cancer atlas ($n = 11{,}578$; ref. 30). Indeed, both hNC tissues and two of three hNC cell lines are classified as lung or esophageal squamous cell carcinoma (**Fig. 2D**; see Supplementary Table S9 for the list of relevant cluster characteristics, and Supplementary Tables S10 and S11 for the list of cluster assignments for hNC tissue and cell lines, respectively). The remaining hNC cell line, 14169, is classified wholly as mesodermal, possibly due to its poorly differentiated state. The findings indicate that hNC classifies as a squamous carcinoma of either lung or esophageal origin.

To determine how mNC classifies in comparison with hNC, we analyzed the transcriptomes from our two mNC cell lines, 311E and 317E, and the parental tumor tissues from which they were derived through OTTER. To classify mouse tissue, mouse genes were

Figure 2.

mNC-derived cell lines share transcriptional profiles with tumor tissues that are characterized by MYC pathway upregulation. **A,** Hemacolor-stained mouse and human NC cell lines as indicated grown on coverslips. NIH-3T3 mouse fibroblasts served as benign control. Scale bars, 25 μm. **B,** Principal components analysis of RNA-seq on the indicated samples based on all genes after filtering by average spike-in normalized expression level >7. Three biological replicates were used for each sample type, as shown. CL, cell line; tumor, primary tumor tissue; CTRL esophagus, normal esophageal mucosa. **C,** Venn diagrams of DE genes relative to normal esophagus control mouse tissue identified by RNA-seq. DE genes were identified based on biological triplicates of each sample. 317E tumor, 317E primary tumor tissue; 311E tumor, 311E primary tumor tissue. **D,** OTTER and augmented OTTER scores of hNC and mNC GEMM tumors and tumor-derived cell lines across hierarchical levels 2–5. Assignments for mNC GEMM cell line replicates are averaged for clarity and legibility. Bar width decreases with each subsequent level. See Supplementary Table S9 for explanations of each cluster assignment. **E,** Hierarchical tree of assigned OTTER clusters illustrating major tumor lineages for the assignments of each tumor/cell line group. Arrows of minor assignments are notated using dotted lines. Bolded groups represent the most important clusters for this analysis.


A
DAPI
NUT
H3K27ac
Merge
311E
317E
TC-797
NIH 3T3
B
H3K27ac domains
Domain size (kb)
Domain rank
342 tumor tissue
CTRL stomach tissue 1
CTRL stomach tissue 2
317E cell line
311E cell line DMSO
311E cell line ABBV
C
NUT
H3K27ac
Merge
311E, 6 h
317E, 6 h
DMSO
9 nmol/L ABBV075
Line plot
Fluorescent intensity
Line distance (pixels)
D
In large domains
Not in large domains
311E DMSO
311E ABBV
317E
311E tumor
317E tumor
TSS
E
Tumor tissue 342
311E CL
317E CL
106
25
86
79
26
54
100
F
Representative murine ideograms
Chr.1
Chr. 15
Chr. 16
Dnm3
Myc
Trp63
342 tumor
311E cell line
317E cell line
shared
No. overlapping
Large domains
G
CUT&RUN (anti-H3K27ac)
RNA-seq
CTRL stomach tissue 1
CTRL stomach tissue 2
342 tumor tissue
317E cell line
311E cell line 96 h DMSO mean
311E cell line 96 h ABBV-075 mean
GEMM CTRL esophagus tissue mean
317E_tumor tissue mean
317E cell line mean
311E_tumor tissue
311E cell line DMSO mean
311E cell line BETi mean
Ref-seq genes
Dnm3
Klf6
Myc
Krt80
Hotair
Trp63
Notch1
Zmynd8
Parp11
Nrg1

converted to human orthologs and OTTER was retrained using its original data set along with data augmentation procedures to increase its generalization (Materials and Methods). Using this augmented OTTER, all mNC tissues and cell lines were classified as squamous cell carcinomas of esophageal or head and neck origin (**Fig. 2D**; Supplementary Tables S12 and S13). There is also a low assignment to normal heart in the tumor tissue due to retained musculature in the tumor sample. The findings indicate that mouse and human NCs co-classify transcriptionally as lung or esophageal squamous carcinomas (summarized in **Fig. 2E**) and add key evidence that the cell identity of murine NC is the same as that of human NC. The data also demonstrate that the mouse cell lines resemble their parental tissue transcriptionally, confirming their suitability as *in vitro* models of the *in vivo* disease with respect to cell identity.

The epigenetic landscape of mNC is similar to that of hNC, highlighted by the presence of BRD4–NUT/H3K27ac megadomains

hNC is characterized by the presence of BRD4–NUT and H3K27ac nuclear foci, also known as nuclear condensates due to their liquid-like properties (42, 43). Like hNC, we found that mNC cell lines also form these foci as observed by immunofluorescence (IF) microscopy (**Fig. 3A**). Similar foci can be observed *in vivo* by IHC (**Fig. 1I**, inset). In contrast, NIH 3T3 fibroblasts lack BRD4–NUT expression and thus BRD4–NUT foci such that H3K27ac is dispersed throughout the nucleus.

These foci arise in concert with chromatin 3D structure reprogramming, which positions very large super-enhancer-like chromatin domains, frequently over 1MB in size and referred to as megadomains (4) in close spatial proximity with one another (8). Megadomains are enriched with BRD4–NUT, p300, and multiple histone marks associated with transcriptional activation, including H3K27ac. H3K27ac is a p300/CBP-specific histone mark of active enhancers, and because it colocalizes faithfully with BRD4–NUT, it is used as a surrogate chromatin marker for these domains (4).

We thus performed CUT&RUN on primary mNC tumor tissue and mNC cell lines to compare domain sizes of H3K27ac with those in control murine gastric epithelial tissue. We found that the H3K27ac domain size distribution in mNC tumor tissue is similar to that in mNC cell lines, but markedly differs from that in control tissue (**Fig. 3B**). Indeed, domain sizes appear much larger in mNC samples compared with those in normal. Thus, we sought to compare the size distribution of large, super-enhancer-like domains (here termed "large domains") between the sample types. We used the approach described by Whyte and colleagues (35) defining super-enhancers to identify large domains in our samples, with the presumption that the large domains in these cell types are comprised wholly or in part of super-enhancers and megadomains. To define large domains, we used as cutoff the point at which rank-ordered lists of region size rapidly increase for H3K27ac domains in mNC tumor tissue (342), and applied to all samples (**Fig. 3B**; see also Materials and Methods), as described (35). Indeed, the size distribution of H3K27ac large domains was similar amongst different mNC samples, but was significantly larger than those of control tissue by a vast margin ($P = 4.9\times10^{-15}$ – 6×10^{-16}, **Fig. 3B**; Supplementary Fig. S4A). The same is true for the group of largest H3K27ac domains that are larger than 200 kb (**Fig. 3B**; Supplementary Fig. S4B). These domain size distributions are consistent with the presence of massive large domains, namely, megadomains in mNC cell lines and tissue. The presence of megadomains in mNC establishes that this pathognomic epigenetic hNC structure is also characteristic of mNC and is conserved between cultured mNC cell lines and *in vivo* tumor tissue.

hNC megadomains and large domains are dependent upon the binding of BRD4–NUT to chromatin via the bromodomains of BRD4. Thus, eviction of BRD4–NUT from chromatin using BETi leads to the dissolution of these large domains (4, 12). Likewise, we found that large domain size is significantly reduced in 311E cells treated with the BETi, ABBV-075 (Supplementary Fig. S4A; ref. 44; **Fig. 3B**). Because these observations are based on bulk analyses, we used immunofluorescence confocal microscopy to determine the effects of BETi on BRD4–NUT/H3K27ac distribution in individual mNC cell nuclei. Similar to what has been observed in hNC (12), BETi treatment resulted in rapid dispersion (increased granularity and spreading) of BRD4–NUT/H3K27ac foci (megadomains) within 6 hours (**Fig. 3C**). These findings may be explained by redistribution of BRD4–NUT and p300 to regions previously depleted of these factors (4).

Using the above cutoff for large domains, we found that, as predicted for super-enhancers, they are associated with increased transcription of associated genes compared with regular domains (**Fig. 3D**). Unexpectedly, however, we found that transcription at large domains was globally unaffected by BETi, and even shows the expansion of longer transcripts. Given that the length of the 5′ untranslated region combined with exon 1 typically ranges from 200 to 500 bp, the roughly 50%–75% increase in longer transcripts in BETi-treated cells may suggest perturbed RNA processing. Possibilities include defective splicing or impaired cross-talk between chromatin and RNA polymerase activity mediated by BRD4 (45). Although transcription of key genes such as *Myc* did decrease after 96 hours of BETi (see below), the global RNA-seq profiles suggest a

Figure 3.

The epigenetic landscape of mNC cell lines is similar to that of mNC tumor tissue and hNC. **A,** Immunofluorescence images using indicated antibodies comparing mNC cell lines 317E and 311E with the hNC cell lines TC-797 and NIH-3T3 are made. Scale bars, 3 μm. **B,** Size distribution of H3K27ac domains based on CUT&RUN performed on the samples indicated. For 311E and 317E cell lines, CUT&RUN was performed on duplicate samples and are shown based on mean values. 342 tumor, primary mNC tissue sampled fresh from an esophageal primary tumor invading stomach; CTRL stomach tissue, stomach mucosa sampled fresh from two different *Sox2*-Cre-ER^{T2} (neg); *Brd4–NUTM1* (het) mice; ABBV, ABBV-075–treated 311E cells, 96 hours; DMSO, DMSO-treated 311E cells, 96 hours. **C,** Confocal immunofluorescent images with line-scan profiles as indicated. Scale bars, 3 μm. **D,** Scaled average transcriptional profiles (top) and heat maps (bottom) based on RNA-seq tag density of the indicated samples at the TSS of genes within large domains, and those that are not. Large domains are defined as those whose size is above the point at which rank-ordered lists of region size rapidly increase for H3K27ac domains in mNC tumor tissue. 317E, 317E cell line; ABBV, ABBV-075-treated 311E cells, 96 hours; DMSO, DMSO-treated 311E cells, 96 hours. **E,** Venn diagram of the number of large domains that overlap between different mNC samples. **F,** Representative ideograms of murine chromosomes 1, 15, and 16 showing the chromosomal location of large domains based on CUT&RUN on the three mNC sample types shown. Callouts mark the genomic regions of the indicated genes. **G,** Integrated Genomics Viewer visualization of CUT&RUN and RNA-seq peaks at megadomains shared between all mNC cell lines and tissue. BETi, ABBV-075 BET bromodomain inhibitor administered for 96 hours at a concentration of 9 nmol/L. Control DMSO-treated samples were treated for 96 hours as well. Shown for cell lines are the mean of two (CUT&RUN) or three (RNA-seq) biological replicates. Individual samples are shown for tissues.

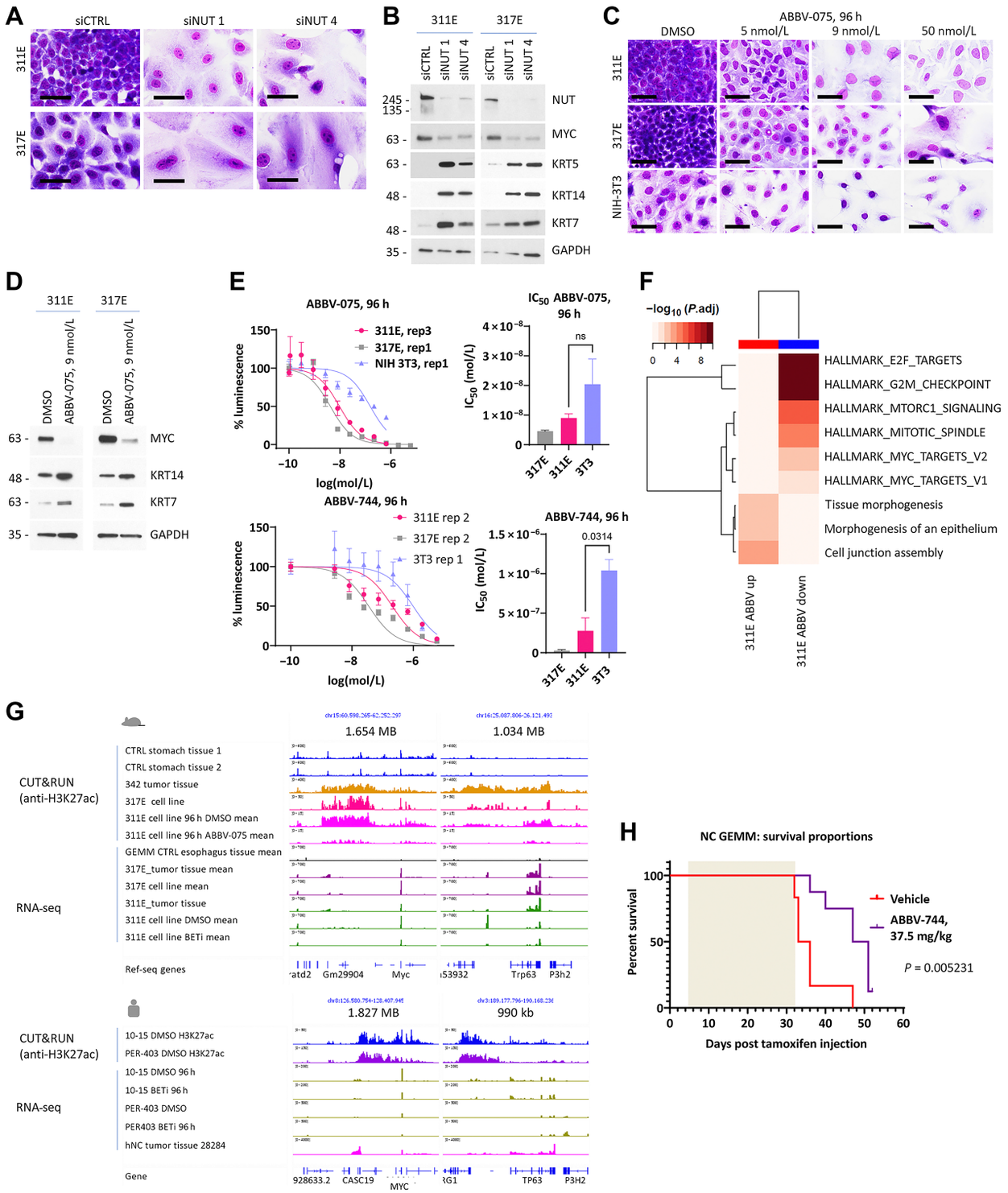


Figure 4.

Like human BRD4–NUT, mBRD4–NUT functions to block differentiation and maintain the growth of mNC in association with the formation of megadomains at the *Myc* locus. **A,** Hemacolor-stained cells grown on coverslips 48 hours following siRNA transfection. Scale bars, 25 μm. **B,** Immunoblot of lysates taken at 48 hours as indicated. **C,** Hemacolor-stained cells grown on coverslips 96 hours following ABBV-075 treatment. Scale bars, 25 μm. **D,** Immunoblot of lysates taken at 96 hours following treatment with DMSO or ABBV-075 (9 nmol/L). **E,** Dose–response curves shown are for one representative biological replicate (three were performed), and the error bars depict the six technical replicates performed for the single biological replicate shown. (*Continued on the following page.*)

potential opportunity to probe the specific effects on RNA processing in future investigations.

Super-enhancers, and likely megadomains, define cell identity (4, 35). Indeed, we found that 79 of 476 large domains (16.6%) were shared amongst all mNC samples across 18 of the 21 chromosomes (311E and 317E cell lines and 342 tumor tissue; **Fig. 3E** and **F**; Supplementary Fig. S4C; Supplementary Table S14). 43% of these domains were shared between the mNC tissue and each cell line, and 51% overlapped between the two cell lines (**Fig. 3E**). This considerable amount of overlap comparing two mNC cell types, including between cell line and tissue, indicates similar epigenetic landscapes and thus related cell types. The RNA-seq and CUT&RUN profiles of 10 such examples are shown in **Fig. 3G**. Every overlapping megadomain in these regions is highly BET bromodomain-dependent, as evidenced by their marked depletion in the presence of BETi. We note that this bulk sequencing data do not account for cell-to-cell variability. Single-cell analysis, in future studies, may reveal subpopulations that could show greater or less overlap with other cell types, depending on the state of differentiation or subclone. Overall, we conclude that the presence of overlapping large and megadomains in mNC cell lines and tissue indicates that mNC cell lines are a powerful *in vitro* model to study NC chromatin biology.

As in hNC, mBRD4–NUT functions to block differentiation and maintain growth of mNC

The general mechanism by which BRD4–NUT drives the growth of hNC is through the blockade of differentiation by upregulating MYC, thereby maintaining cells in an undifferentiated, proliferative state (7). Knockdown of BRD4–NUT leads to the differentiation and growth arrest of hNC cells; this phenotype can be rescued by MYC overexpression (7). To test whether BRD4–NUT functions similarly in mNC cells, we knocked down BRD4–NUT by electroporation of 311E and 317E cells with siRNAs targeting the codon-optimized, GEMM-specific *NUTM1* sequence. Within 24 to 48 hours, both cell lines exhibited morphologic features of squamous differentiation, including marked flattening and increased surface area of cells, enlargement, and increased pallor of nuclei (**Fig. 4A**). The morphologic changes were accompanied by increased expression of the epithelial-specific differentiation marker, KRT7, and the squamous differentiation-specific markers, KRT14 and KRT5 (**Fig. 4B**). In keeping with a role of BRD4–NUT in upregulating MYC, MYC expression was reduced (**Fig. 4B**), recapitulating what occurs in hNC cells (17, 18). The findings confirmed that BRD4–NUT serves the same overall function in mNC to block differentiation in an MYC-dependent manner, as has previously been shown in hNC (7).

We have previously shown that blocking the binding of BET proteins to chromatin using BET bromodomain inhibitors (BETi) phenocopies the effect of BRD4–NUT knockdown to induce differentiation and growth arrest in hNC (7). hNC, the only known malignancy harboring a *BRD4* oncogene, was the first cancer in which the on-target antineoplastic effects of BETi were shown (12, 13). Likewise, we found that treatment with ABBV-075 has the same effect on mNC cells, demonstrating similar morphologic and IHC changes of squamous differentiation induced by BRD4–NUT knockdown, including downregulation of MYC and upregulation of KRT7 and 14 (**Fig. 4C** and **D**; ref. 39). By comparison, although non-mNC NIH-3T3 cells did show slowed growth, morphologic changes of differentiation, including flattening and nuclear-cytoplasmic enlargement observed in mNC cells, were not seen (**Fig. 4C**). As in hNC, the growth of 311E and 317E cells is exquisitely sensitive to ABBV-075, with mean half maximal inhibitory (IC_{50}) concentrations ranging from 4.5 to 9 nmol/L (**Fig. 4E**, top). Although the IC_{50} of 3T3 cells to ABBV-075 was not significantly different from the mNCs (20.4 nmol/L), these cells did not fully growth arrest even at the highest concentration of ABBV-075 (∼1 μmol/L, **Fig. 4E**, top). In separate work (39), we have shown that hNC is exceptionally sensitive to ABBV-744, a BETi selective for the second bromodomain (BD2) that has demonstrated activity in a narrow range of cancers, specifically prostate cancer and acute leukemia, but has a much more favorable toxicity profile compared with pan-BETi, such as ABBV-075 (39, 46). As in hNC, we found that mNC cell lines were sensitive to ABBV-744, with significantly greater sensitivity to this compound compared with 3T3 cells (**Fig. 4E**, bottom). The findings indicate that epigenetic dependencies overlap between hNC and mNC.

To determine how the loss of BRD4-NUT function impacts transcriptional programs in mNC cells, we evaluated the effects of BETi on the 311E mNC cell line by RNA-seq. Unsurprisingly, and similar to what is seen in hNC (40), we observed marked changes in global transcription induced by BETi (**Fig. 4F**; Supplementary Fig. S5A and S5B), and GSEA notably revealed that MYC targets and cell division pathways (**Fig. 4F**) were negatively enriched, similarly to hNC cells treated with BETi (40). Also evident was that BETi treatment leads to the enrichment of epithelial differentiation programs as observed morphologically and by expression of differentiation markers (**Fig. 4C** and **D**). These findings are in keeping with the observation that BRD4-NUT upregulates MYC expression in hNC, which in turn is required and sufficient for the blockade of differentiation in this cancer (4, 7).

Myc and *Trp63* megadomains are defining features of mNC and hNC

The association of MYC with BRD4–NUT expression in mNC cells led us to inquire whether *Myc* is an epigenetic target of BRD4–NUT/H3K27ac megadomains in mNC cells, as it is in hNC (4, 7). In all hNC cell lines and tissue that have been tested, megadomains are invariably adjacent to *MYC* and *TP63* loci (**Fig. 4G**, bottom), which encode proteins that are required for hNC growth and viability, respectively (4). In fact, interchromosomal interaction between these

(*Continued.*) IC_{50}s from biological triplicate dose–response assays corresponding with **D**. Fisher exact *P* values comparing 311E and 3T3 are shown; ns, not significant. **F,** Heat map plotting the adjusted *P* value from GSEA (Hallmark or Genome Ontology gene sets) of ranked log-fold change values comparing RNA-seq from ABBV-075-treated cells (96 hours) relative to DMSO control. Up, differentially upregulated genes; down, differentially downregulated genes. **G,** Integrated Genomics Viewer visualization of CUT&RUN and RNA-seq peaks at *Myc* and *Trp63* loci in mouse cells and tissues (top), and for hNC cells and tissue (bottom). BETi, ABBV-075 BET bromodomain inhibitor administered for 96 hours at a concentration of 9 nmol/L in murine cells and 0.7 and 13.45 nmol/L in 10–15 and PER-403 hNC cell lines, respectively. Control DMSO-treated samples were treated for 96 hours as well. Top, cell lines shown are the mean of two (CUT&RUN) or three (RNA-seq) biological replicates. Individual samples are shown for tissues. Bottom, single representative replicates are shown for both cell lines and tissue. RNA-seq and CUT&RUN data for 10–15 and PER-403 cell lines (39) and are available at GEO with accession numbers GSE142481 and GSE228533, respectively. **H,** Kaplan–Meier survival plot of NC GEM treated with either vehicle control ($n = 6$) or ABBV-744 (37 mg/kg/day, $n = 8$). Beige rectangle denotes the time during treatment, 28 days.

megadomains was observed in hNC cells (8), raising the yet untested possibility that these transcription factors may be components of the cell identity-defining core regulatory circuitry in hNC. Using anti-H3K27ac as a surrogate for BRD4–NUT/H3K27ac megadomains, as previously described (4), we found that in all mNC tissues and cell lines, in contrast with control stomach tissue, a H3K27ac megadomain was present adjacent to *Myc* (range, ∼600 kb–1.2 MB) and *Trp63* (range, ∼190 kb–2.7 MB) loci (**Figs. 3F** and **G**; **4G**, top). Moreover, robust expression of both genes could be seen in all mNC tumor tissues, derivative cell lines, and in hNC cells. In line with the reduced protein expression of MYC, its transcription significantly dropped in the presence of BETi in both mNC and hNC cells (**Fig. 4G**). This is seen to occur in association with megadomain depletion in 311E cells (**Fig. 4G**, top). The consistent presence of megadomains in both hNC and mNC at NC-cell defining transcription factors provides strong evidence that similar epigenetic and oncogenic mechanisms drive growth and block differentiation of mNC and hNC.

Taken together, we find that mNC and hNC share similar epigenetic mechanisms to drive growth by blocking the differentiation of squamous-lineage progenitor cells. In both species, BRD4–NUT forms H3K27ac-enriched megadomains that are associated with expression of the key oncogenic target, *MYC/Myc*, and the cell lineage-defining gene, *TP63/Trp63*. These core similarities, together with the *in vitro* sensitivity of mNC and hNC to BETi, suggest that the NC GEMM may serve as a relevant model of NC. Thus, this work sets the foundational groundwork for further validation of mNC as a preclinical model to study cell-intrinsic and -extrinsic vulnerabilities for therapeutic development.

BET inhibition prolongs survival of NC GEMM *in vivo*

To evaluate the NC GEMM as a preclinical model for targeted inhibition of BRD4–NUT, we tested the *in vivo* activity of the BETi, ABBV-744. In this study, starting five days after tamoxifen-induced tumor initiation, we treated animals with ABBV-744 (37.5 mg/kg/day, one-half maximal tolerated dose (46)) or vehicle control. Treatment was for 28 days, then terminated, and animals were followed for survival based on morbidity criteria. This study revealed a significant prolongation of survival of BETi-treated NC GEMs [$P = 0.005$, log-rank (Mantel–Cox) test] compared with vehicle controls (**Fig. 4H**). However, no mice were cured and all had gross involvement of the esophagus and stomach with tumors at necropsy (Supplementary Fig. S6). The findings demonstrate *in vivo* activity of BETi in mNC and resemble findings reported in xenograft models of hNC treated with BETi monotherapy, where there is suppression of tumor growth and prolonged survival, but not cure of animals (39, 40, 47). The preclinical response to BETi in our GEMM also resembles what has been reported in human NC clinical trials of single-agent BETi, where responses have been seen in a subset of patients, but are short-lived (median duration of response, 2.8 months; range, 1–10.3 months), and all patients eventually succumb to disease (2, 48–51). Thus, the results of our *in vivo* study indicate that the NC GEMM can be used as a rigorous preclinical model of targeted therapy of human NC.

Discussion

The rapid formation of mNC tumors in 100% of our NC GEMs is definitive evidence that BRD4–NUT alone potently drives the malignant transformation of squamous progenitor cells. This observation is of utmost significance: the presence of a single cancer driver affords deep mechanistic insights through a focus on a single oncoprotein. We can now study those mechanisms *in vivo* with the NC GEMM. The first discovery of premalignant *in situ* lesions in this model provides a window in which to investigate the cascade of molecular changes driven by *de novo* BRD4–NUT expression that eventuate into malignant transformation, tumor invasion, and eventual metastasis.

Our NC GEMM has been validated as a faithful model of hNC in several key aspects. First, the rapid tumor growth, metastasis, and death caused by induction of BRD4–NUT (**Fig. 1D**, **L** and **M**) closely mimic the pathophysiology of the human disease, which in thoracic cases has a median survival of 4.4 months (1). Second, and importantly, the poorly differentiated squamous histopathology and IHC staining of mNC is indistinguishable from that of hNC (**Fig. 1H–K**). Third, the overriding function of BRD4–NUT in both mNC and hNC is to block differentiation; its knockdown results in rapid epithelial differentiation of NC cell lines from both species (**Fig. 4A** and **B**; refs. 7, 17, 18). Fourth, the key downstream target of BRD4–NUT in hNC and mNC, *MYC/Myc,* is shared in all hNC and mNC tissues and cell lines (**Fig. 4B**, **D**, **F**, and **G**; ref. 4). Fifth, we found that BRD4–NUT-dependent MYC-driven programs of cell-cycle progression and altered metabolism, characteristic of hNC (39, 40), are also properties of mNC tissues and cell lines (**Fig. 4F**; Supplementary Fig. S3). Sixth, upregulation of MYC expression is associated with the formation of H3K27ac/BRD4–NUT megadomains, and their presence in mNC is an important validation that similar epigenetic mechanisms drive tumor formation in mice and humans (**Figs. 3A–G** and **4G**; Supplementary Fig. S4).

Another powerful resource generated from the GEMM is the creation of two mNC cell lines. We found that many key biological aspects of these cell lines, namely, the blockade of differentiation by BRD4–NUT, the formation of megadomains, and MYC pathway upregulation, match those of primary mNC tissue (**Figs. 3A–G**, **4B**, **D**, **F**, and **G**; Supplementary Fig. S3). Nevertheless, there are also a few differences that may be related to *ex vivo* growth on plastic, including downregulation of epithelial–mesenchymal transition (Supplementary Fig. S3). Although the significance of these differences is undetermined, the mNC cell lines will serve as a powerful experimental model to test hypotheses in high throughput (i.e., CRISPR-Cas9 screens) that can later be validated *in vivo* in the GEMM.

In summary, this novel creation and validation of a highly penetrant and robust NC GEMM will provide the global scientific community with a much-needed immuno-competent animal model that recapitulates the human disease, with the potential to greatly accelerate NC translational and basic research.

Authors' Disclosures

P.C. Gokhale reports grants from Kymera Therapeutics, Moderna, and Pfizer, Inc. outside the submitted work. A. Shlien reports a patent for use of transcriptional analysis to diagnose cancer and predict patient prognosis pending. J. Luo reports other support from Targeted Oncology, Physicians Education Resource, VJ Oncology, Cancer GRACE, Community Cancer Education Inc., AstraZeneca, Astellas, and Amgen; grants from Erasca, Genentech, Kronos Bio, Novartis, Revolution Medicines; personal fees from Erasca, Blueprint Medicines, and Daiichi Sankyo outside the submitted work; and a patent filed by Memorial Sloan Kettering Cancer Center related to multimodal features to predict response to immunotherapy(PCT/US2023/115872) pending. C.J. Wu reports other support from BioNTech and Pharmacyclics outside the submitted work. D.B. Keskin is a scientific advisor for Immunitrack and Breakbio. D.B. Keskin owns equity in Affimed N.V., Agenus, Armata Pharmaceuticals, Breakbio, BioMarin Pharmaceutical, Celldex Therapeutics, Editas Medicine, Gilead Sciences, Immunitybio, ImmunoGen, IMV, Lexicon Pharmaceuticals, and Neoleukin Therapeutics. BeiGene, a Chinese biotech company, supported unrelated SARS

COV-2 research at TIGL. K.P. Eagen reports grants from NIH, Cancer Prevention and Research Institute of Texas, and Alex's Lemonade Stand Foundation for Childhood Cancer during the conduct of the study and personal fees from Atheneum Partners GmbH outside the submitted work. G.I. Shapiro reports grants and personal fees from Merck KGaA/EMD-Serono, grants from Tango Therapeutics, Bristol Myers Squibb, Pfizer, Lilly, personal fees from Bicycle Therapeutics, Cybrexa Therapeutics, Boehringer Ingelheim, Bayer, ImmunoMet, Concarlo Holdings, Syros, Zentalis, CytomX Therapeutics, Blueprint Medicines, Kymera, Janssen, Xinthera, and Artios outside the submitted work; in addition, G.I. Shapiro has a patent for Dosage regimen for sapacitabine and seliciclib issued to Cyclacel Pharmaceuticals and Geoffrey Shapiro and a patent for Compositions and Methods for predicting response to CDK4/6 inhibition issued to Liam Cornell and Geoffrey Shapiro. C.A. French reports other support from AbbVie, Max Vincze Foundation, Ryan Richards Foundation, McDevittStrong, DFCI NUT Carcinoma Fund, Jay Dion Memorial Golf Classic, grants from NIH RO1 award, Bertarelli Cancer Research Fund, and Samuel Waxman Cancer Research Foundation during the conduct of the study; personal fees from Boehringer-Ingelheim and other support from Boehringer-Ingelheim outside the submitted work. No disclosures were reported by the other authors.

Authors' Contributions

R. Durall: Data curation, investigation, methodology, writing–review and editing. **J. Huang:** Investigation. **L. Wojenski:** Formal analysis, investigation, writing–review and editing. **Y. Huang:** Investigation, methodology. **P.C. Gokhale:** Investigation, methodology, writing–review and editing. **B.A. Leeper:** Investigation. **J.O. Nash:** Software, formal analysis. **P.L. Ballester:** Data curation, software. **S. Davidson:** Formal analysis. **A. Shlien:** Supervision, project administration. **E. Sotirakis:** Supervision, Investigation. **F. Bertaux:** Investigation, methodology. **V. Dubus:** Resources, methodology. **J. Luo:** Writing–review and editing. **C.J. Wu:** Conceptualization, writing–review and editing. **D.B. Keskin:** Conceptualization, writing–review and editing. **K.P. Eagen:** Formal analysis, writing–original draft, writing–review and editing. **G.I. Shapiro:** Conceptualization, funding acquisition, writing–original draft, writing–review and editing. **C.A. French:** Conceptualization, resources, data curation, formal analysis, supervision, funding acquisition, investigation, visualization, methodology, writing–original draft, project administration, writing–review and editing.

Acknowledgments

This work was supported by grants from the NIH R01 CA124633 (C.A. French), NIH RO1 CA285308 (C.A. French, D.B. Keskin, G. Shapiro), the Samuel Waxman Cancer Research Foundation (C.A. French), the Bertarelli Rare Cancer Research Fund (C.A. French), the NIH DP5 OD024587 (K.P. Eagen), Dana-Farber/Harvard Cancer Center SPORE in Lung Cancer Career Enhancement Program Grant (J. Luo), the Alex's Lemonade Stand Foundation for Childhood Cancer 22-25794 (K.P. Eagen), philanthropic gifts from the Friends of Jay Dion Memorial Golf Classic, the Max Vincze Foundation, the Ryan Richards Foundation, McDevittStrong, and multiple donations to the DFCI NUT fund (G.I. Shapiro, C.A. French). K.P. Eagen is a CPRIT Scholar in Cancer Research (RR210082). D.B. Keskin is supported by 1R01HL157174-01A1. A. Shlien is partially supported by a Garron Family Chair in Childhood Cancer Research. ABBV-075 and ABBV-744 were generously provided by AbbVie, Inc.

The publication costs of this article were defrayed in part by the payment of publication fees. Therefore, and solely to indicate this fact, this article is hereby marked "advertisement" in accordance with 18 USC section 1734.

Note

Supplementary data for this article are available at Cancer Research Online (http://cancerres.aacrjournals.org/).

Received August 24, 2023; revised September 26, 2023; accepted October 6, 2023; published first October 11, 2023.

References

1. Chau NG, Ma C, Danga K, Al-Sayegh H, Nardi V, Barrette R, et al. An anatomical site and genetic-based prognostic model for patients with Nuclear Protein in Testis (NUT) midline carcinoma: analysis of 124 patients. JNCI Cancer Spectr 2020;4:pkz094.
2. French CA, Cheng ML, Hanna GJ, DuBois SG, Chau NG, Hann CL, et al. Report of the first international symposium on NUT carcinoma. Clin Cancer Res 2022; 28:2493–505.
3. Stevens TM, Morlote D, Xiu J, Swensen J, Brandwein-Weber M, Miettinen MM, et al. NUTM1-rearranged neoplasia: a multi-institution experience yields novel fusion partners and expands the histologic spectrum. Mod Pathol 2019; 32:764–73.
4. Alekseyenko AA, Walsh EM, Wang X, Grayson AR, Hsi PT, Kharchenko PV, et al. The oncogenic BRD4-NUT chromatin regulator drives aberrant transcription within large topological domains. Genes Dev 2015;29: 1507–23.
5. Reynoird N, Schwartz BE, Delvecchio M, Sadoul K, Meyers D, Mukherjee C, et al. Oncogenesis by sequestration of CBP/p300 in transcriptionally inactive hyperacetylated chromatin domains. EMBO J 2010;29:2943–52.
6. Alekseyenko AA, Walsh EM, Zee BM, Pakozdi T, Hsi P, Lemieux ME, et al. Ectopic protein interactions within BRD4-chromatin complexes drive oncogenic megadomain formation in NUT midline carcinoma. Proc Natl Acad Sci USA 2017;114:E4184–E92.
7. Grayson AR, Walsh EM, Cameron MJ, Godec J, Ashworth T, Ambrose JM, et al. MYC, a downstream target of BRD-NUT, is necessary and sufficient for the blockade of differentiation in NUT midline carcinoma. Oncogene 2014; 33:1736–42.
8. Rosencrance CD, Ammouri HN, Yu Q, Ge T, Rendleman EJ, Marshall SA, et al. Chromatin hyperacetylation impacts chromosome folding by forming a nuclear subcompartment. Mol Cell 2020;78:112–26.
9. Eagen KP, French CA. Supercharging BRD4 with NUT in carcinoma. Oncogene 2021;40:1396–408.
10. Lee JK, Louzada S, An Y, Kim SY, Kim S, Youk J, et al. Complex chromosomal rearrangements by single catastrophic pathogenesis in NUT midline carcinoma. Ann Oncol 2017;28:890–7.
11. French CA, Badve S, den Bakker MA, Jain D. NUT carcinoma of the thorax. 5th ed. Chan JKC, Lantuejoul S, Marx A, editors. Lyon, France: International Agency for Research on Cancer; 2021; 4p.
12. Filippakopoulos P, Qi J, Picaud S, Shen Y, Smith WB, Fedorov O, et al. Selective inhibition of BET bromodomains. Nature 2010;468:1067–73.
13. Stathis A, Zucca E, Bekradda M, Gomez-Roca C, Delord JP, de La Motte Rouge T, et al. Clinical response of carcinomas harboring the BRD4-NUT oncoprotein to the targeted bromodomain inhibitor OTX015/MK-8628. Cancer Discov 2016;6: 492–500.
14. Filippakopoulos P, Picaud S, Mangos M, Keates T, Lambert JP, Barsyte-Lovejoy D, et al. Histone recognition and large-scale structural analysis of the human bromodomain family. Cell 2012;149:214–31.
15. Filippakopoulos P, Knapp S. Targeting bromodomains: epigenetic readers of lysine acetylation. Nat Rev Drug Discov 2014;13:337–56.
16. Arnold K, Sarkar A, Yram MA, Polo JM, Bronson R, Sengupta S, et al. Sox2(+) adult stem and progenitor cells are important for tissue regeneration and survival of mice. Cell stem cell 2011;9:317–29.
17. French CA, Rahman S, Walsh EM, Kuhnle S, Grayson AR, Lemieux ME, et al. NSD3-NUT fusion oncoprotein in NUT midline carcinoma: implications for a novel oncogenic mechanism. Cancer Discov 2014;4:928–41.
18. French CA, Ramirez CL, Kolmakova J, Hickman TT, Cameron MJ, Thyne ME, et al. BRD-NUT oncoproteins: a family of closely related nuclear proteins that block epithelial differentiation and maintain the growth of carcinoma cells. Oncogene 2008;27:2237–42.
19. Schwartz BE, Hofer MD, Lemieux ME, Bauer DE, Cameron MJ, West NH, et al. Differentiation of NUT midline carcinoma by epigenomic reprogramming. Cancer Res 2011;71:2686–96.
20. Chen S, Zhou Y, Chen Y, Gu J. fastp: an ultra-fast all-in-one FASTQ preprocessor. Bioinformatics 2018;34:i884–i90.
21. 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.
22. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, et al. The sequence alignment/map format and SAMtools. Bioinformatics 2009;25: 2078–9.

23. 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.
24. Robinson JT, Thorvaldsdottir H, Winckler W, Guttman M, Lander ES, Getz G, et al. Integrative genomics viewer. Nat Biotechnol 2011;29:24–6.
25. Quinlan AR, Hall IM. BEDTools: a flexible suite of utilities for comparing genomic features. Bioinformatics 2010;26:841–2.
26. Zerbino DR, Johnson N, Juettemann T, Wilder SP, Flicek P. WiggleTools: parallel processing of large collections of genome-wide datasets for visualization and statistical analysis. Bioinformatics 2014;30:1008–9.
27. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol 2014;15:550.
28. Yu G, Wang LG, Han Y, He QY. clusterProfiler: an R package for comparing biological themes among gene clusters. OMICS 2012;16:284–7.
29. Vivian J, Rao AA, Nothaft FA, Ketchum C, Armstrong J, Novak A, et al. Toil enables reproducible, open source, big biomedical data analyses. Nat Biotechnol 2017;35:314–6.
30. Comitani F, Nash JO, Cohen-Gogo S, Chang AI, Wen TT, Maheshwari A, et al. Diagnostic classification of childhood cancer using multiscale transcriptomics. Nat Med 2023;29:656–66.
31. Seal RL, Braschi B, Gray K, Jones TEM, Tweedie S, Haim-Vilmovsky L, et al. Genenames.org: the HGNC resources in 2023. Nucleic Acids Res 2023;51: D1003–D9.
32. Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. Nat Methods 2012;9:357–9.
33. Tarasov A, Vilella AJ, Cuppen E, Nijman IJ, Prins P. Sambamba: fast processing of NGS alignment formats. Bioinformatics 2015;31:2032–4.
34. Stovner EB, Saetrom P. epic2 efficiently finds diffuse domains in ChIP-seq data. Bioinformatics 2019;35:4392–3.
35. Whyte WA, Orlando DA, Hnisz D, Abraham BJ, Lin CY, Kagey MH, et al. Master transcription factors and mediator establish super-enhancers at key cell identity genes. Cell 2013;153:307–19.
36. Zhu LJ, Gazin C, Lawson ND, Pages H, Lin SM, Lapointe DS, et al. ChIPpeakAnno: a bioconductor package to annotate ChIP-seq and ChIP-chip data. BMC Bioinf 2010;11:237.
37. Wang R, Liu W, Helfer CM, Bradner JE, Hornick JL, Janicki SM, et al. Activation of SOX2 expression by BRD4-NUT oncogenic fusion drives neoplastic transformation in NUT midline carcinoma. Cancer Res 2014;74:3332–43.
38. Hsieh AL, Walton ZE, Altman BJ, Stine ZE, Dang CV. MYC and metabolism on the path to cancer. Semin Cell Dev Biol 2015;43:11–21.
39. Huang Y, Durall RT, Luong NM, Hertzler HJ, Huang J, Gokhale PC, et al. EZH2 cooperates with BRD4-NUT to drive NUT carcinoma growth by silencing key tumor suppressor genes. Cancer Res 2023:83:3956–73.
40. Morrison-Smith CD, Knox TM, Filic I, Soroko KM, Eschle BK, Wilkens MK, et al. Combined targeting of the BRD4-NUT-p300 axis in NUT midline carcinoma by dual selective bromodomain inhibitor, NEO2734. Mol Cancer Ther 2020;19:1406–14.
41. French CA, den Bakker MA. WHO classification of head and neck tumours. 4th ed. El-Naggar A, Chan JKC, Grandis JR, Takata T, Slootweg P, editors. Lyon: International Agency for Research on Cancer (IARC); 2017. p.20–21.
42. Yu D, Liang Y, Kim C, Jaganathan A, Ji D, Han X, et al. Structural mechanism of BRD4-NUT and p300 bipartite interaction in propagating aberrant gene transcription in chromatin in NUT carcinoma. Nat Commun 2023;14:378.
43. Ibrahim Z, Wang T, Destaing O, Salvi N, Hoghoughi N, Chabert C, et al. Structural insights into p300 regulation and acetylation-dependent genome organisation. Nat Commun 2022;13:7759.
44. Bui MH, Lin X, Albert DH, Li L, Lam LT, Faivre EJ, et al. Preclinical characterization of BET family bromodomain inhibitor ABBV-075 suggests combination therapeutic strategies . Cancer research 2017;77:2976–89.
45. Lam FC, Kong YW, Huang Q, Vu Han TL, Maffa AD, Kasper EM, et al. BRD4 prevents the accumulation of R-loops and protects against transcription-replication collision events and DNA damage. Nat Commun 2020;11:4083.
46. Faivre EJ, McDaniel KF, Albert DH, Mantena SR, Plotnik JP, Wilcox D, et al. Selective inhibition of the BD2 bromodomain of BET proteins in prostate cancer. Nature 2020;578:306–10.
47. Liao S, Maertens O, Cichowski K, Elledge SJ. Genetic modifiers of the BRD4-NUT dependency of NUT midline carcinoma uncovers a synergism between BETis and CDK4/6is. Genes Dev 2018;32:1188–200.
48. Piha-Paul SA, Hann CL, French CA, Cousin S, Braña I, Cassier PA, et al. Phase 1 study of molibresib (GSK525762), a bromodomain and extra-terminal domain protein inhibitor, in NUT carcinoma and other solid tumors. JNCI Cancer Spectr 2020;4:pkz093.
49. Cousin S, Blay JY, Garcia IB, de Bono JS, Le Tourneau C, Moreno V, et al. Safety, pharmacokinetic, pharmacodynamic and clinical activity of molibresib for the treatment of nuclear protein in testis carcinoma and other cancers: results of a Phase I/II open-label, dose escalation study. Int J Cancer 2022; 150:993–1006.
50. Lewin J, Soria JC, Stathis A, Delord JP, Peters S, Awada A, et al. Phase Ib trial with birabresib, a small-molecule inhibitor of bromodomain and extraterminal proteins, in patients with selected advanced solid tumors. J Clin Oncol 2018; 36:3007–14.
51. Shapiro GI, LoRusso P, Dowlati A, K TD, Jacobson CA, Vaishampayan U, et al. A Phase 1 study of RO6870810, a novel bromodomain and extra-terminal protein inhibitor, in patients with NUT carcinoma, other solid tumours, or diffuse large B-cell lymphoma. Br J Cancer 2021;124:744–53.