

RESEARCH PAPER



MYCN and HIF-1 directly regulate *TET1* expression to control 5-hmC gains and enhance neuroblastoma cell migration in hypoxia

Anastasia E. Hains^a, Sakshi Uppal^a, John Z. Cao^a, Helen R. Salwen^b, Mark A. Applebaum^b, Susan L. Cohn^b, and Lucy A. Godley^{a,c}

^aSection of Hematology/Oncology, Department of Medicine, and the University of Chicago Comprehensive Cancer Center, The University of Chicago, Chicago, IL, USA; ^bDepartment of Pediatrics, The University of Chicago, Chicago, IL, USA; ^cDepartment of Human Genetics, The University of Chicago, Chicago, IL, USA

ABSTRACT

Ten-Eleven-Translocation 5-methylcytosine dioxygenases 1–3 (*TET1*–3) convert 5-methylcytosine to 5-hydroxymethylcytosine (5-hmC), using oxygen as a co-substrate. Contrary to expectations, hypoxia induces 5-hmC gains in *MYCN*-amplified neuroblastoma (NB) cells via upregulation of *TET1*. Here, we show that *MYCN* directly controls *TET1* expression in normoxia, and in hypoxia, HIF-1 augments *TET1* expression and *TET1* protein stability. Through gene-editing, we identify two *MYCN* and HIF-1 binding sites within *TET1* that regulate gene expression. Bioinformatic analyses of 5-hmC distribution and RNA-sequencing data from hypoxic cells implicate hypoxia-regulated genes important for cell migration, including *CXCR4*. We show that hypoxic cells lacking the two *MYCN*/HIF-1 binding sites within *TET1* migrate slower than controls. Treatment of *MYCN*-amplified NB cells with a *CXCR4* antagonist results in slower migration under hypoxic conditions, suggesting that inclusion of a *CXCR4* antagonist into NB treatment regimens could be beneficial for children with *MYCN*-amplified NBs.

KEY POLICY HIGHLIGHTS

- In *MYCN*-amplified neuroblastoma cell lines, *MYCN* directly controls *TET1* expression in normoxia.
- In *MYCN*-amplified neuroblastoma cell lines exposed to hypoxia, HIF-1 augments *TET1* expression and *TET1* protein stability.
- Hypoxic *MYCN*-amplified neuroblastoma cell lines have increased cell migration, mediated by genes including *CXCR4* that gain 5-hydroxymethylcytosine density.
- Treatment of *MYCN*-amplified NB cells with a *CXCR4* antagonist slows hypoxia-associated migration, suggesting a *CXCR4* antagonist could be beneficial in treatment regimens for children with *MYCN*-amplified neuroblastomas.

ARTICLE HISTORY

Received 19 March 2022
Accepted 14 July 2022

KEYWORDS

MYCN; hypoxia; *TET1*; neuroblastoma

Introduction

Most commonly diagnosed in children 5 years and younger, neuroblastoma (NB) tumours originate from the developing sympathetic nervous system, with a predilection for the adrenal gland. NBs exhibit a broad range of clinical behaviours, with some tumours resolving spontaneously, whereas others progress despite intensive medical and surgical intervention [1]. High-risk patients with aggressive tumours are identified based on several diagnostic criteria, including *MYCN* status, age, stage, and grade [1]. Currently, high-risk patients receive a combination of systemic chemotherapy, surgical resection, tandem autologous transplant,

radiation, and immunotherapy as part of established treatment protocols [1]. Although improvements in clinical outcomes have been observed during the past decades, the 5-year survival of high-risk patients remains at less than 60% [2], emphasizing a need for more effective therapies.

Next-generation sequencing has allowed researchers to analyse large datasets of NB genomes and transcriptomes, with the aim of identifying novel genetic biomarkers and therapeutic targets [3,4]. However, results from sequencing studies have shown that NBs have a relatively low mutational burden compared to other cancers [5,6], with only a few recurrent mutations [6]. The

paucity of recurrent somatic genetic alterations has driven a search for epigenetic modifications that contribute to NB pathogenesis [5,7].

NBs arise from neural crest tissue, which is known to have high levels of 5-hydroxymethylcytosine (5-hmC) epigenetic marks [8]. When neural crest cells undergo neuronal differentiation, 5-hmC increases and is enriched in the gene bodies of activated genes, implying 5-hmC is an important epigenetic mark associated with transcription [8,9]. The 5-hmC base functions both as a stable and distinct epigenetic modification associated with open chromatin and active gene transcription [10–12] as well as an intermediate in cytosine demethylation [13,14]. 5-hmC is generated from the oxidation of 5-methylcytosine (5-mC) by three TET enzymes, which use oxygen and α -ketoglutarate in the catalytic reaction and Fe(II) as a co-factor [9].

Solid tumours, including NB, are often characterized by a lack of organized vasculature that generates hypoxic regions in the tumour [15]. The hypoxia-inducible factor (HIF) family of transcription factors (TFs), HIF1/2/3, are the master regulators of the hypoxic response [16]. All the family members are alpha-beta heterodimers with both subunits constitutively transcribed and translated [15]. In a normal oxygen environment, oxygen-dependent prolyl hydroxylases bind and hydroxylate two proline residues in the HIF alpha subunit, marking it for ubiquitination and proteasomal degradation [15,16]. In contrast, in a hypoxic environment, the lack of free oxygen lowers prolyl hydroxylase activity, resulting in accumulation of the alpha subunit. Once the alpha-beta heterodimer forms, it localizes to the nucleus and promotes target gene expression that aids in metabolism, proliferation, angiogenesis, and invasion that characterizes the hypoxic tumour phenotype [15]. When NB cell lines are exposed to hypoxia, they develop an aggressive phenotype, consisting of enhanced proliferation, migration, and invasion, all of which are specifically mediated via HIF-1 [17].

Previously, we described how global 5-hmC increases are essential for induction of the hypoxic transcriptional response specifically in *MYCN*-amplified and tumorigenic NB cell lines [12]. In hypoxia, HIF-1 upregulates *TET1* expression,

which drives an increase in 5-hmC levels along the gene bodies of hypoxic response genes, augmenting their expression. Here, we show that *MYCN* and HIF-1 TFs directly upregulate *TET1* in normoxia and hypoxia, respectively, describing a novel interaction between HIF-1 α and *TET1* proteins, and identify a 5-hmC target, *CXCR4*, as a mediator of the increased migratory phenotype characteristic of hypoxia.

Materials and methods

Cell culture and CRISPR-Cas9 genome editing

NB cell lines were provided by the Cohn Laboratory and were authenticated by STR profiling and tested negative for *Mycoplasma*. SK-N-BE [2], NBL-WN, LA1-55 n, NBL-S, SH-SY5Y, LA1-5S, NBL-WS, and SHEP cells were cultured in RPMI with 10% foetal bovine serum (FBS). Tet-21/N cells [18] were cultured in RPMI with 10% Tet system FBS (Clontech 631,106). Normoxic culture was done at 37°C under 21% O₂ and 5% CO₂ in a humidified incubator. For hypoxic exposure, the cells were incubated under 1% O₂ and 5% CO₂ in a humidified chamber. Tet-21/N cells were maintained with 1 μ g/mL doxycycline, and *MYCN* induction was induced by the removal of doxycycline from the media.

A CRISPR-Cas9 system was used to perform genome editing. First, gRNAs were designed and evaluated with the programme CRISPOR [19]. For NTC, a guide was used that does not target anywhere in the human genome (Table S1). gRNAs were ligated into plasmid Lenticrispr v2 (Addgene plasmid #52961) following the Zhang laboratory protocol [20]. To produce lentiviral particles, 293 T cells were transfected according to the Addgene protocol adapted from [21]. The same protocol was used to transduce target cells with the lentivirus, but to accommodate the low tolerance of NB cells, 4 μ g/mL of polybrene was added to the media [21]. To genotype each clone, DNA was extracted using phenol:chloroform, and PCR using GoTaq (Promega PRM3005) was performed with primers targeting the region surrounding the edited site (Table S1). Sanger sequencing of the amplicon was then used to define the mutation within each clone. For each target, two clones

that lacked the TF binding site but had different deletions were selected for subsequent experiments.

Immunoprecipitation and Western blotting

Nuclear protein extraction was performed via high salt fractionation [22], and proteins were separated via SDS-PAGE on 6% acrylamide gels. TET1 immunoprecipitation was performed using the α -TET1 antibody (Santa Cruz sc-293,186). PVDF membranes were probed with primary antibody: α -TET1 (Genetex GT1462) at 1:1000 dilution overnight at 4°C; α -TOP1 (Abcam ab109374) at 1:1000 dilution for 1 hour at room temperature; or α -MYCN (Abcam ab16898) at 1:1000 dilution for 1 hour at room temperature. After primary antibody incubation, the membranes were incubated with secondary antibodies at 1:5000 dilution against the appropriate species: α -Rabbit IgG (Millipore 401,393–2ML) or α -Mouse IgG (Cell Signaling Technology 70,765) for 1 hour at room temperature. Protein degradation assays in SK-N-BE [2] cells were performed following the protocol outlined in [23]. Protein quantification analysis was done with ImageJ [24].

Chromatin-Immunoprecipitation (ChIP)

ChIP was performed following the Covaris truChIP Chromatin Shearing Kit protocol (PN 520154 and PN 520127). Briefly, cells were cross-linked with 1% formaldehyde for 10 minutes at room temperature and quenched with 175 mM Tris pH 7. Extracted nuclei were sonicated with a Covaris S220 Sonolab 7.2 1.0. Sonicated DNA was immunocleared with 2 μ g/mL sheared salmon sperm DNA (Thermo Fisher Scientific AM9680), 6 μ g/mL normal mouse IgG (Santa Cruz Biotechnology sc-2025), and 30 μ L Protein G Dynabeads (50% vol/vol) (Fisher Scientific 10-004-D). Next, antibody pulldown was performed with 5 μ g α -MYCN IgG (Santa Cruz, B8.4.B) or 5 μ g α -HIF-1 α (BD Biosciences, 610,958) overnight at 4°C. As an antibody control, 5 μ g of normal mouse IgG was used (Santa Cruz, sc-2025). After immunoprecipitation, 30 μ L Protein G Dynabeads and 2 μ g/mL sheared salmon sperm DNA were added for an additional

2 hours. Captured protein-DNA complexes were washed, eluted, and reverse-crosslinked. DNA was purified for qPCR or sequencing with Genelute PCR Clean Up Kit (Sigma-Aldrich).

ChIP-qPCR

DNA was amplified for ChIP-qPCR with Power SYBR Green PCR Master Mix. Serially diluted input DNA was used as the standard curve. ChIP-qPCR data were normalized using the percent input method ($\% \text{ input} = 100 \times 2^{-(\text{adjusted input} - \text{Ct})}$). Primers for target regions can be found in Table S1.

ChIP-sequencing

ChIP samples were sequenced on an Illumina HiSeq 4000, from a single end for a length of 50 bp. Sequenced reads were aligned to the hg19 genome build with the Burrows-Wheeler Aligner [25]. Peaks were called with MACS2 [26]. Files for visualization on IGV were generated with the Count function of IGVtools [27]. Data from HIF-1 α ChIP-sequencing have been deposited in the Gene Expression Omnibus (GEO) under the GEO accession number GSE167477.

Bisulphite Sequencing

Bisulphite sequencing was performed as previously described [11,12], and bisulphite primers can be found in Table S1.

Quantification of 5-hmC by ultra-high performance liquid chromatography coupled with tandem mass spectrometry (UHPLC-MS/MS)

5-hmC quantification was performed as described [28].

Determination of 5-hmC distribution by hMe-SEAL and sequencing

5-hmC selective chemical labelling (hMe-Seal) was performed according to the reported methods [12,29]. DNA libraries were sequenced on an Illumina HiSeq 4000, from a single end for a length of 50 bp. Sequenced reads were aligned

to the hg19 genome build with the Burrows-Wheeler Aligner [25]. Peaks were called with MACS2 [26]. Peak files from all time points were merged. HTSeq-count was used to count the number of reads per peak, which was then converted to FPKM [30]. Peaks that were classified as ‘early’ have a FPKM increase greater than 1.2-fold from 0 to 6 hours and remain stable throughout the rest of the time course. Peaks that were classified as ‘late’ were stable from 0 to 24 hours and increased to greater than 1.1-fold between 24 and 48 hours. Following the reported methods, we conducted genomic element analyses by intersecting 5-hmC peaks with files containing the location and annotation of genomic elements [11,12]. Predicted enhancers were generated from data collected by EnhancerAtlas, downloaded in a bedfile format [31]. The HIF-1 α binding site regions were generated from HIF-1 α ChIP-sequencing peak files, called with MACS2 [26]. Pathway analysis was performed with PANTHER pathway analysis, using adjusted P-values [32]. Data from hMe-Seal sequencing have been deposited in GEO under the accession number GSE167475.

RNA extraction and sequencing

Total RNA was extracted with RNAzol reagent (Sigma-Aldrich, R4533) according to the manufacturer’s protocol. RNA was converted to cDNA with Life Technologies High-Capacity cDNA Reverse Transcription Kit (4,368,814). Quantitative PCR was performed with Power SYBR Green PCR Master Mix (Life Technologies 4,368,706). Primers for qPCR can be found in Table S1.

RNA was sequenced on an Illumina HiSeq 4000, from a single end for a length of 50 bp. Reads from RNA-sequencing in fastq format were aligned to the hg19 reference genome built with Tophat2 [33]. Gene expression was quantified and compared with tools in the Cufflinks package [33]. To detect *TET1*s, the Cufflinks reference annotation gtf file was modified to include the Gnomon [34] predicted transcript: XM_011540207.2. RNA-sequencing experiment data from our experiments can be accessed under GSE167476.

Transwell and wound healing assays

Transwell assays were performed following reported methods, with PBS control and 10 μ g/mL plerixafor added to the media [35]. Wound healing assays began with cells grown to confluency in a 96-well plate. Wells were scratched with an Essen Woundmaker and then placed in IncuCyte and photographed every 4 hours. Alternatively, the wound was created with the Abcam Wound Healing Assay kit (ab242285), and cells were photographed at 0, 6, 12, 24, and 48 hours on an Evos FL Auto 2 microscope. To achieve pseudo-hypoxia, NB cells were treated with FG-4592 at 30 μ M and with 10 μ g/mL plerixafor at hour 0 and every 12 hours thereafter. Drug optimization experiments were performed with both drugs to determine the optimal concentration for NB cells.

Statistical analysis

Statistical significance was calculated using two-tailed t-tests for most biological experiments between the two groups. When more than one group was compared, a one-way ANOVA test was used. For large datasets for which a normal distribution could not be assumed, a Wilcoxon rank-sum test was used. P-values from all tests were considered significant when less than 0.05. Significance tests were carried out on R, GraphPad Prism, and Microsoft Excel. Graphs were generated in R and GraphPad Prism. Statistical thresholds and exact values for n are given in the Figure Legends. Statistical analysis associated with Figure S9 is described in Supplementary Material.

Results

MYCN binds three potential regulatory regions of TET1

We observed that hypoxia induces 5-hmC gains through up-regulation of *TET1* specifically in *MYCN*-amplified NB cells, but not in NB cells with low *MYCN* protein levels [12,36]. This implies *MYCN* regulates *TET1* transcription either directly or indirectly. To address this, we examined the correlation between *MYCN* and *TET1* expression using published RNA-sequencing

(RNA-seq) data from NB cell lines and primary tumours [37,38]. *MYCN*-amplified NB lines have a higher baseline level of *TET1* expression compared to non-*MYCN* amplified lines (Figure 1a), and there is a positive correlation between *MYCN* and *TET1* expression levels ($r = 0.56$, $P = 0.01$, Figure 1b). We observed a similar positive correlation between *TET1* and *MYCN* expression in NB tumours [37] ($r = 0.48$, $P = 1.50 \times 10^{-31}$, Figure 1c). To test if this observation was specific for *TET1*, we performed the same analysis with *TET2* and *TET3* [37,38], and found no correlation between *MYCN* and that of *TET2* or *TET3* in NB cell lines (Fig. S1A,C). However, in tumours, there was a modest correlation between *MYCN* and *TET3* ($r = 0.31$, $P = 5.71 \times 10^{-12}$, Fig. S1D), but not between *MYCN* and *TET2* (Fig. S1B).

To determine whether reduction of *MYCN* expression or protein level also results in lowering of *TET1* expression, we examined publicly available RNA-seq data from NB cells that had been treated with JQ1, a BET-bromodomain inhibitor that binds and inhibits the *MYCN* transcriptional regulator BRD4 [39]. Expression data from SK-N-BE [2]-C cells that had been treated with DMSO or JQ1 at 4, 8, or 24 hours [39] demonstrated that *MYCN* expression decreases by hour 4 and stays at that level for up to 24 hours with a similar trajectory for *TET1* (Fig. S2A). To confirm that this decrease was due to loss of *MYCN* and not a direct result of BRD4 inhibition, we also examined publicly available array data from a mouse NB tumour cell line that had been treated with an Aurora A kinase (AURKA) inhibitor [40]. AURKA stabilizes *MYCN* protein by preventing its degradation [40]. Upon AURKA inhibition, *Tet1* expression decreased (Fig. S2B). Together, these results suggest that *MYCN* is necessary for maintaining baseline *TET1* expression in NB.

We then asked if *TET1* was a direct transcriptional target of *MYCN*. *MYCN*, like HIF-1, is a TF that recognizes and binds an E-box element that consists of a 5'-CANNTG-3' sequence [39]. Using HOMER [41], we identified numerous potential *MYCN* binding motifs within and around *TET1* (Figure 1d), suggesting that *MYCN* could bind and regulate *TET1* from several sites. To determine which of these sites is most likely bound by *MYCN*, we used EnhancerAtlas [42], an enhancer

annotation tool based on compiled Chromatin-immunoprecipitation-sequencing (ChIP-seq) and chromatin accessibility data to visualize enhancer sites that could potentially regulate *TET1* expression (Figure 1d). EnhancerAtlas provides a database of annotated enhancers from a variety of cell lines, including non-*MYCN* amplified NB line SK-N-SH [42]. Using data from this cell line, EnhancerAtlas identified two potential enhancers in *TET1*, in the first and second introns [42] (Figure 1d). The existence of these enhancers has been confirmed experimentally in mouse and human systems. The enhancer located in the first intron is encompassed by a superenhancer originally identified in an intergenic region in mice and remapped to the hg19 human genome [43,44] (Figure 1d). The enhancer in the second intron was identified as a c-MYC target in human embryonic stem cells [45]. Next, we examined the public *MYCN* ChIP-seq data [39,46] at the *TET1* locus from all available *MYCN*-amplified NB lines to determine the experimentally validated *MYCN* binding sites (Figure 1d). From data from the six available *MYCN*-amplified lines, we observed *MYCN* binding to two sites in *TET1*, which we will refer to hereafter as *Site 1* (*S1*) and *Site 2* (*S2*) (Figure 1d). *S1* is located centrally within the predicted enhancer in the first *TET1* intron, and *S2* is located within the predicted enhancer in the second *TET1* intron. In addition, we considered the possibility that *TET1* could be regulated by *MYCN* binding at distal enhancers. EnhancerAtlas identified a potential *TET1*-associated distal enhancer located ~88 kb upstream in the first intron of the gene *DNA2*, which we confirmed was bound by *MYCN* in *MYCN*-amplified NB lines (Figure 1d). For comparison, we also visualized *MYCN* ChIP-seq data from non-*MYCN*-amplified NB cell lines [39,46]. In NB69 and SHEP cells, there is very little *MYCN* present and no binding to *DNA2*, *S1*, or *S2* (Figure 1d). However, one additional non-*MYCN*-amplified cell line, Tet-21/N, derived from the SHEP line, is engineered to express *MYCN* under a tetracycline-off system [18]. When we visualized publicly available *MYCN* ChIP-seq [39] from this cell line in both 'MYCN on' and 'MYCN off' conditions, we found *MYCN* bound *DNA2*, *S1*, and *S2* only in the 'MYCN on' condition. We tested the

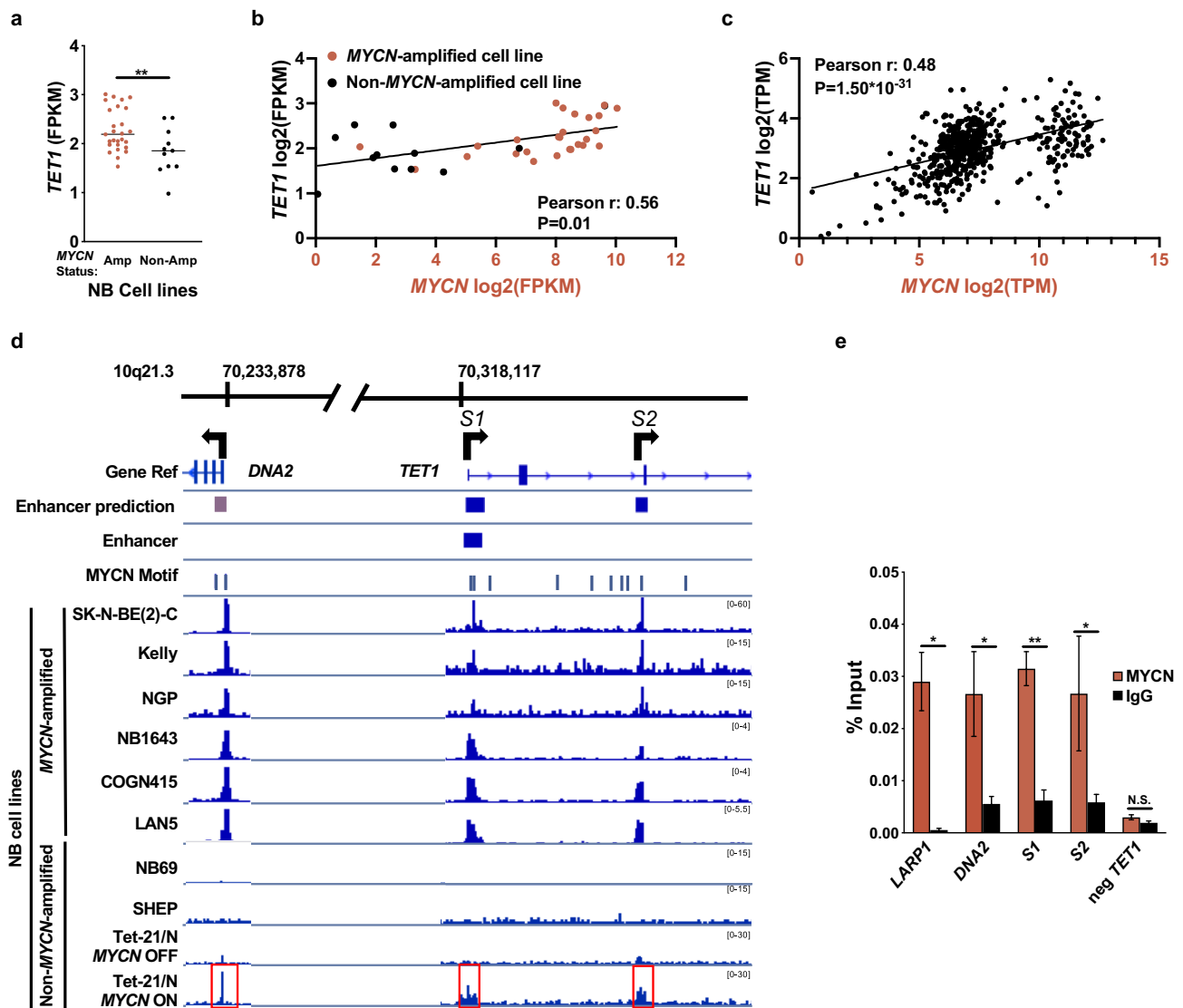


Figure 1. MYCN binds a superenhancer located in the first intron of *TET1* (S1), the second intron of *TET1* (S2), and a predicted upstream enhancer site near *DNA2*. (a) *TET1* expression data (FPKM) from RNA-seq of 38 NB cell lines [38] graphed according to *MYCN*-amplified (Amp, red) or non-amplified (Non-Amp, black) status. GEO accession numbers are given in Accession Numbers. (b) *TET1* (y-axis) and *MYCN* (x-axis) expression data from RNA-seq of 38 NB cell lines [38] are expressed as log₂ of FPKM. *MYCN*-amplified cell lines are denoted by red circles, and non-*MYCN*-amplified lines are denoted by black circles. *r* represents Pearson correlation coefficient, and *P*-values have been corrected with the Bonferroni correction method [67]. Cell line names and expression values can be found in Table S2. (c) *TET1* (y-axis) and *MYCN* (x-axis) expression data from RNA-seq of 579 NB tumours [37]. *TET1* vs *MYCN* expression data are plotted as log₂ of FPKM. *r* represents Pearson correlation coefficient, and *P*-values have been corrected with the Bonferroni correction method [67]. (d) Enhancer regions, MYCN binding motifs, and MYCN ChIP-seencing data visualized around the *TET1* gene. The genomic loci on chromosome 10q21.3 are located at the top according to the hg19 genome build. The Gene Ref track shows the positions of the *TET1* and *DNA2* genes, and arrows show the directions of gene transcription. S1 and S2 are also indicated. Next, enhancers predicted by 'EnhancerAtlas.org' [42] are shown as purple rectangles. An enhancer within *TET1* generated from the UCSC LiftOver tool [41] is shown as a purple rectangle. Below, the 'MYCN Motif' track, generated with HOMER [41], shows the positions of MYCN binding motifs as vertical bars. The bottom rows display MYCN ChIP sequencing data from SK-N-BE [2]-C, Kelly, NGP, NB1643, COGN415, LAN5, NB69, SHEP, and Tet-21/N cells [39,46] at locus 10q21 visualized with the count function of IGVtools [27]. The scale for each track is listed on the right in brackets. The shaded region corresponding to the first peak, from left to right, is within *DNA2*. The second shaded peak contains S1, within *TET1*-intron 1, and the last shaded peak contains S2, within *TET1*-intron 2. (e) Real time qPCR validation of MYCN ChIP-seencing performed in SK-N-BE [2]. From left to right, positive control (LARP1), *DNA2*, S1, S2, and negative control (neg *TET1*) binding sites are plotted on the x-axis. MYCN pulldown values are plotted in red, and IgG controls are plotted in black. The y-axis corresponds to the amount of DNA pulled down normalized to the amount of input DNA. * and ** represent *P* < 0.05 and *P* < 0.01, respectively.

binding of MYCN to each of these three sites in SK-N-BE [2] NB cells with MYCN ChIP-quantitative PCR (qPCR) and found that all three sites were enriched with MYCN binding (Figure 1e). Bisulphite sequencing of the S1 and S2 regions determined that the cytosines are unmodified in normoxia and accessible for TF binding (Fig. S3). These results suggest that MYCN potentially regulates baseline *TET1* through any or all of these three sites.

MYCN is associated with full-length *TET1* transcript expression

Recent studies have described an alternative ‘short’ *TET1* transcript (*TET1s*) in embryonic development, neural, and cancer model systems [45,47–49]. *TET1s* is translated into an isoform of TET1 that lacks the CXXC domain and has an intact catalytic domain (Fig. S4A,B). Although the role of TET1s has yet to be fully understood, it is present in differentiated tissues after an isoform switch from full-length TET1 (TET1fl), which is dominant in embryonic tissues [45,47–49]. Because NB is a tumour of embryonic origin, we investigated whether *TET1s* was present in NB cell lines [38] and whether MYCN preferentially targeted expression of a specific *TET1* transcript. We found that both *TET1fl* and *TET1s* are expressed in NB lines (Figure 2a). However, MYCN-amplified NB cells expressed higher levels and a higher percentage of *TET1fl* (Figure 2a,b). Further analysis revealed that *TET1fl* levels correlate with MYCN expression, but *TET1s* levels do not (Figure 2c,d). Moreover, we found that *TET1fl* levels were reduced by about 60% at 4 hours, whereas *TET1s* was only reduced by about 35% in JQ1-treated SK-N-BE [2]-C cells [39] (Fig. S4C). Together, these results suggest that MYCN preferentially drives *TET1fl* expression.

MYCN is sufficient to induce *TET1* expression in an inducible MYCN expression system

To test if MYCN is sufficient to modulate the baseline *TET1* expression, we used the inducible-MYCN NB line Tet-21/N cells [18], in which we previously showed that, in ‘MYCN on’ condition, MYCN binds S1 and S2 (Figure 1d). MYCN levels

were measured daily after induction, with maximal expression reached on Day 4 (Fig. S5A). RNA-seq and qPCR from Day 4 samples showed *TET1* and *TET3* were induced in MYCN-expressing cells (Fig. S5B,C,D). Analysis of *TET1fl* and *TET1s* transcripts revealed that only *TET1fl* was induced, providing additional support that MYCN targets *TET1fl* preferentially (Fig. S5E). Using an antibody that specifically detects TET1fl, we found TET1fl protein levels were also elevated on Days 4, 5, and 6 (Fig. S5F,G). Despite the induction of *TET1/3*, we observed no increase in 5-hmC levels (Fig. S5H), likely due to the very low baseline levels of *TET1* expression in comparison to MYCN-amplified lines [50].

MYCN controls *TET1* expression directly in normoxia through two binding sites

We hypothesized that one or more of the MYCN binding sites in/near *TET1* (Figure 1 and Figure 3a) directly regulate *TET1* transcription. In order to test the functional role of each site, we used CRISPR-Cas9 gene editing in SK-N-BE [2] cells to introduce double-strand breaks in the E-box motif and remove each potential binding site from each allele (Figure 3b). From the mixed population of cells, we generated single cell clones and used PCR amplification followed by Sanger sequencing to confirm that each binding site of interest was eliminated (Figure 3b). Deletions generated by this process are often unique in their placement and length, but all lack the CpG contained within the E-box motifs of interest (Figure 3b). Two gene-edited single-cell clones per binding site were used in further experiments. Additionally, we generated SK-N-BE [2] lines lacking both S1 and S2 sites (Figure 3c) through CRISPR-Cas9 targeting of the S2 site in one of the $\Delta S1$ cell lines. Clones were named after their binding site deletion ($\Delta DNA2$, $\Delta S1$, $\Delta S2$, and $\Delta S1/2$) followed by ‘1’ or ‘2’ to distinguish individual clones targeting the same site (Fig. S6A,B and Figure 3a,b,c). We also generated a clone that did not feature any genetic alterations at the binding sites to serve as a non-targeting control (NTC). Using qPCR primers that detected expression of both *TET1* transcripts, we found that there was no change in *TET1* expression in SK-N-BE [2] cells

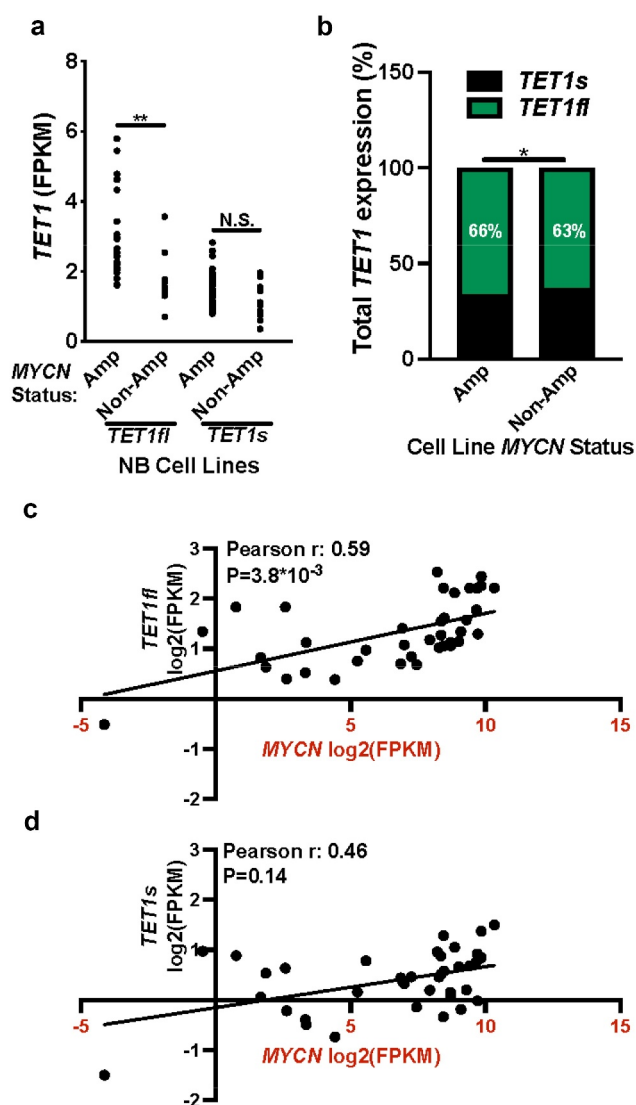


Figure 2. Two distinct *TET1* transcripts are expressed in NB cell lines. (a) FPKM of *TET1fl* and *TET1s* in 38 *MYCN*-amplified (Amp) and non-*MYCN*-amplified (Non-Amp) NB cell lines [38]. ** represents $P < 0.01$; N.S., not significant. (b) Percent of total *TET1* expression in *MYCN*-amplified (Amp) vs non-*MYCN*-amplified (non-Amp) cells. Percent of *TET1fl* is plotted in green, and percent of *TET1s* is plotted in black, with percent calculations displayed in white. * represents $P < 0.05$. (c) *TET1fl* (y-axis) and *MYCN* (x-axis) expression data from RNA-seq of 38 NB cell lines [38] are expressed as log₂ of FPKM. r represents Pearson correlation coefficient, and P -values have been corrected with the Bonferroni correction method [67]. Cell line names can be found in Table S2. (d) *TET1s* (y-axis) and *MYCN* (x-axis) expression data from RNA-seq of 38 NB cell lines [38] are expressed as log₂ of FPKM. r represents Pearson correlation coefficient, and P -values have been corrected with the Bonferroni correction method [67]. Cell line names can be found in Table S2.

when either the potential binding site near *DNA2* or *S2* was deleted, suggesting that *MYCN* binding at these sites is not essential for maintaining baseline *TET1* (Fig. S6C and Figure 3d).

In contrast, $\Delta S1$ and $\Delta S1/2$ cells exhibited reductions in *TET1* expression, of 0.7-fold and 0.25-fold respectively, compared to parental cells (Figure 3d). This indicates that *S1* partially moderates baseline *TET1* levels, but both sites together maintain a high baseline *TET1* expression. To confirm *TET1* protein levels reflected *TET1* mRNA levels, we used Western blotting to detect the *TET1fl* protein, which, for simplicity, we will refer to as *TET1* in subsequent experiments. Unexpectedly, *TET1* protein levels in $\Delta S1$ and $\Delta S1/2$ cells remained comparable to parental and NTC cells (Figure 3e), suggesting that *TET1* protein is relatively stable.

***S1* and *S2* modulate *TET1* expression together in hypoxia**

We previously reported hypoxic *TET1* induction was dependent on the presence of HIF-1 α [12] but had not determined the mechanism by which HIF-1 regulates *TET1*. There are numerous HIF-1 binding motifs in both *DNA2* and *TET1*, including the two that serve as *MYCN* binding sites in *TET1* (Figures 3, 4a). To determine if HIF-1 binds to any of these sites, we performed HIF-1 α ChIP-seq to measure HIF-1 α binding throughout the genome and compared our results with publicly available HIF-1 β ChIP-seq [51]. We found that in hypoxia, both HIF-1 subunits bind *S1* and *S2* in *TET1*, but not at the binding motif in *DNA2* that is bound by *MYCN* in normoxia (Figure 4a). To test if HIF-1 alters the binding of *MYCN* to *S1* or *S2*, we performed *MYCN* ChIP-qPCR in hypoxia and found that *MYCN* still binds to *S1* and *S2* in hypoxic cells (Fig. S7A,B).

To assess the function of HIF-1 binding at *S1* and *S2* on *TET1* expression, we measured *TET1* in $\Delta S1$, $\Delta S2$, and $\Delta S1/2$ cells that were exposed to hypoxia for 48 hours with qPCR (Figure 4b). Although *TET1* expression was low in both normoxic and hypoxic $\Delta S1$ cells (Figure 4b), *TET1* protein levels were elevated in hypoxic cells (Figure 4c). In contrast, *TET1* expression increased as expected in hypoxic $\Delta S2$ cells (Figure 4b) along with its corresponding protein levels (Figure 4d). Despite the very low *TET1* mRNA levels in hypoxic $\Delta S1/2$ cells (Figure 4b), *TET1* protein was readily detectable (Figure 4e), suggesting that

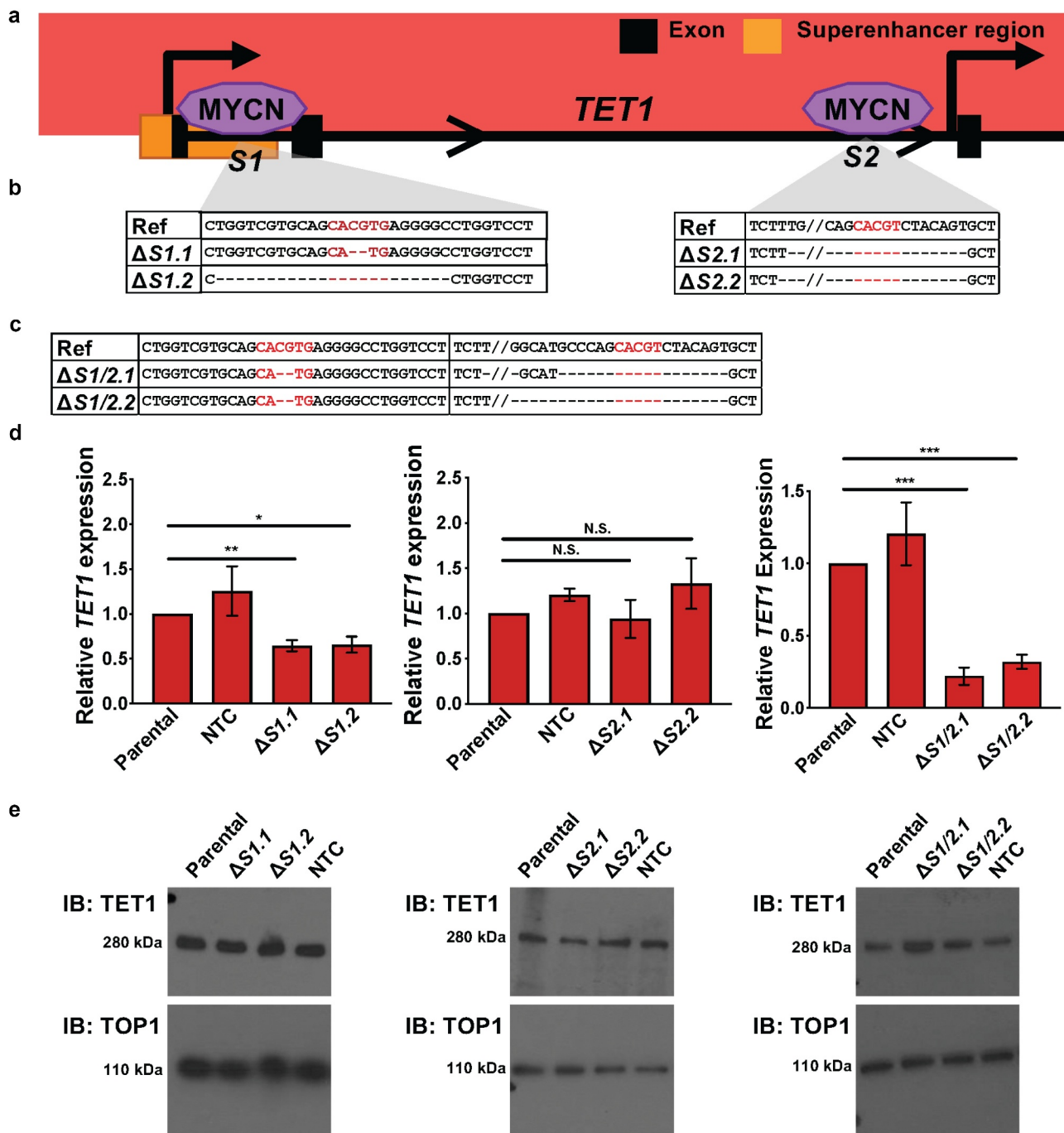


Figure 3. Binding of S1 and S2 by MYCN is important for maintaining normoxic TET1 mRNA expression. (a) Model of MYCN binding around 10q21.3. The two sites are: S1, in a superenhancer (Orange box) in the first intron of TET1, and S2, in the second intron of TET1. (b) Sanger sequencing of the deletions generated in SK-N-BE [2] cells by gRNAs targeting the MYCN binding motifs in TET1: at S1 ($\Delta S1.1$ and $\Delta S1.2$), and at S2 ($\Delta S2.1$ and $\Delta S2.2$). Ref, reference human sequence at this locus. Red font indicates the E-box motif. Slash marks in sequencing from $\Delta S2$ cells represents the loss of 27 base pairs (bps). All clones utilized by this study lacked the CpG that makes up the core of the E-box motif. (c) Sanger sequencing of the deletions generated by gRNAs targeting both MYCN binding motifs at S1 and S2 in TET1 ($\Delta S1/2.1$ and $\Delta S1/2.2$). Ref, reference human sequence at this locus. Red font indicates the E-box motif. Slash marks in sequencing from $\Delta S1/2$ cells represent the loss of 21 bps. (d) Real time qPCR of TET1 mRNA across $\Delta S1$, $\Delta S2$, and $\Delta S1/2$ CRISPR-edited SK-N-BE [2] cell lines ($n = 3$). P-values were determined with one-tailed t-tests between the value of interest and the parental value; *, **, and *** represent $P < 0.05$, $P < 0.01$, and $P < 0.001$, respectively; N.S., not significant. (e) Western blots for TET1 protein (top panels, cropped at ~280 kDa) in normoxic $\Delta S1$ (left panel); $\Delta S2$ (middle panel), and $\Delta S1/2$ (right panel) CRISPR-edited SK-N-BE [2] cell lines. These blots were stripped and re-probed for TOP1 (bottom panels, cropped at ~110 kDa) as loading controls. Parental cells are non-targeted, and NTC, are cells targeted with a non-specific gRNA to control for off-targets effects.

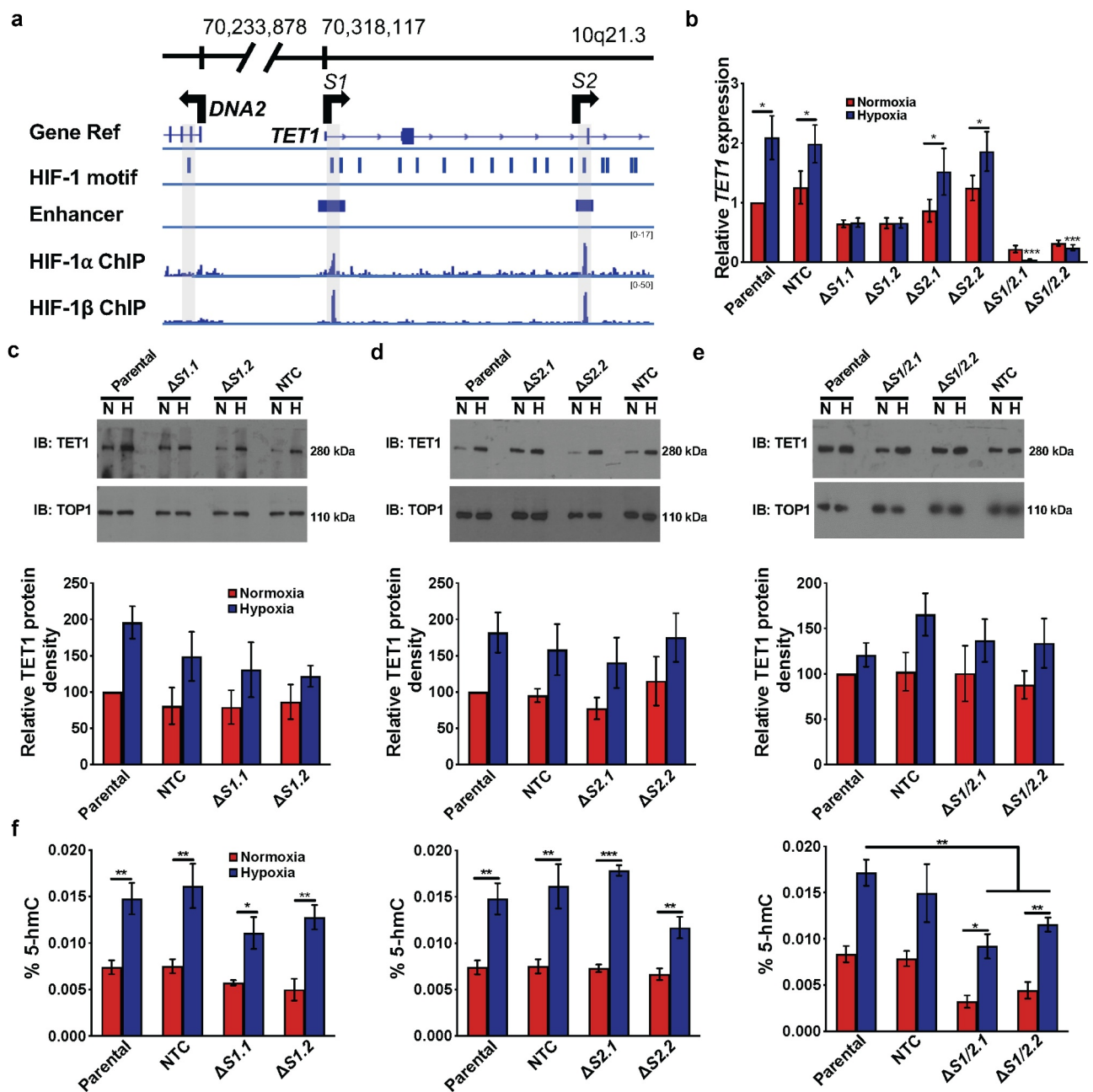


Figure 4. Hypoxia affects *TET1* mRNA transcription and protein levels. (a) HIF-1 motif and sequencing data aligned with the *TET1* promoter, transcriptional start site, and 5' end of the gene. Top, genomic positions from hg19 at 10q21.3. Gene Ref track indicates locations of the *TET1* and *DNA2* genes, with arrows representing the direction of transcription of each gene. *S1* and *S2* are also indicated via shading. HIF-1 motifs were identified by HOMER [41]. Next, an enhancer track marks the location of the enhancers found within *TET1*. HIF-1 α and HIF-1 β [51] ChIP-seq data from SK-N-BE [2] cells were visualized with IGVtools [27]. The scale for each track is listed on the right in brackets. (b) Real time qPCR quantification of *TET1* transcription in normoxia (red) versus hypoxia (dark blue) across Δ *S1*, Δ *S2*, and Δ *S1/2* clones, normalized to normoxic parental SK-N-BE [2] *TET1* expression. P-values were determined with one-tailed t-tests between the parental cell line and the cell line of interest ($n = 3$). * represents $P < 0.05$. (c) Representative Western blots for TET1 protein (upper panels) in Δ *S1* SK-N-BE [2] cells in normoxia (n) versus hypoxia (h), cropped at ~ 280 kDa. TOP1 loading control (lower panels), cropped at ~ 110 kDa. Bottom, mean TET1 protein quantified by ImageJ and normalized to parental normoxic controls in Δ *S1* SK-N-BE [2] cells ($n = 3$). (d) Representative Western blots for TET1 protein (upper panels) in Δ *S2* SK-N-BE [2] cells in normoxia (N) versus hypoxia (H), cropped at ~ 280 kDa. TOP1 loading control (lower panels), cropped at ~ 110 kDa. Bottom, mean TET1 protein quantified by ImageJ and normalized to parental normoxic controls in Δ *S2* SK-N-BE [2] cells ($n = 3$). (e) Representative Western blots for TET1 protein (upper panels) in Δ *S1/2* SK-N-BE [2] cells in normoxia (N) versus hypoxia (H), cropped at ~ 280 kDa. TOP1 loading control (lower panels), cropped at ~ 110 kDa. Bottom, mean TET1 protein quantified by ImageJ and normalized to parental normoxic controls in Δ *S1/2* SK-N-BE [2] cells ($n = 3$). (f) Percent 5-hmC levels in Parental, non-targeted control (NTC), Δ *S1*, Δ *S2*, and Δ *S1/2* CRISPR-edited SK-N-BE [2] cells grown in normoxia (red) versus hypoxia (dark blue). *, **, and *** represent $P < 0.05$, $P < 0.01$, and $P < 0.001$, respectively.

the TET1 protein is stabilized in hypoxic conditions (see below). These results were recapitulated in a second *MYCN*-amplified NB cell line, NBL-WN, in which both *S1* and *S2* were deleted (Fig. S8A-E).

To determine the effects of TET1 catalytic activity under normoxic versus hypoxic conditions, we used mass spectrometry to measure global 5-hmC levels [12]. As expected, exposure of *MYCN*-amplified NB cell lines to hypoxia drove accumulation of 5-hmC levels over cells grown in normoxia (figure 4f and Fig. S8F), consistent with persistent TET activity even under hypoxic conditions [52]. However, global 5-hmC levels in hypoxic $\Delta S1/2$ cells did not reach those measured in parental or NTC cells (figure 4f).

HIF-1 promotes TET1 stability in hypoxia

To determine if elevated TET1 protein levels in hypoxia were a consequence of increased protein

stability, we performed protein degradation assays with cycloheximide in normoxic and hypoxic parental SK-N-BE [2] cells. These degradation assays showed that TET1 had a half-life of 20 hours in normoxia and was still detectable even at 30 hours (Figure 5a). In hypoxia, TET1 had an even longer half-life of 40 hours (Figure 5a,c). To determine if HIF-1 is necessary for the increased stability of TET1 in hypoxia, we performed the same degradation assays in normoxia and hypoxia with cells in which *HIF1A* had been deleted via gene editing ($\Delta HIF1A$) (Figure 5b). Unlike the parental line, the TET1 stability in hypoxia was similar to TET1 stability in normoxia, with TET1 in both conditions having a half-life about 20 hours (Figure 5b, c). This implies HIF-1 α is necessary to stabilize TET1 protein in hypoxia.

Past studies have described TF-TET complexes in which the complex works synergistically to promote TET activity and enhance gene expression from the TF targets [53,54]. We hypothesized that

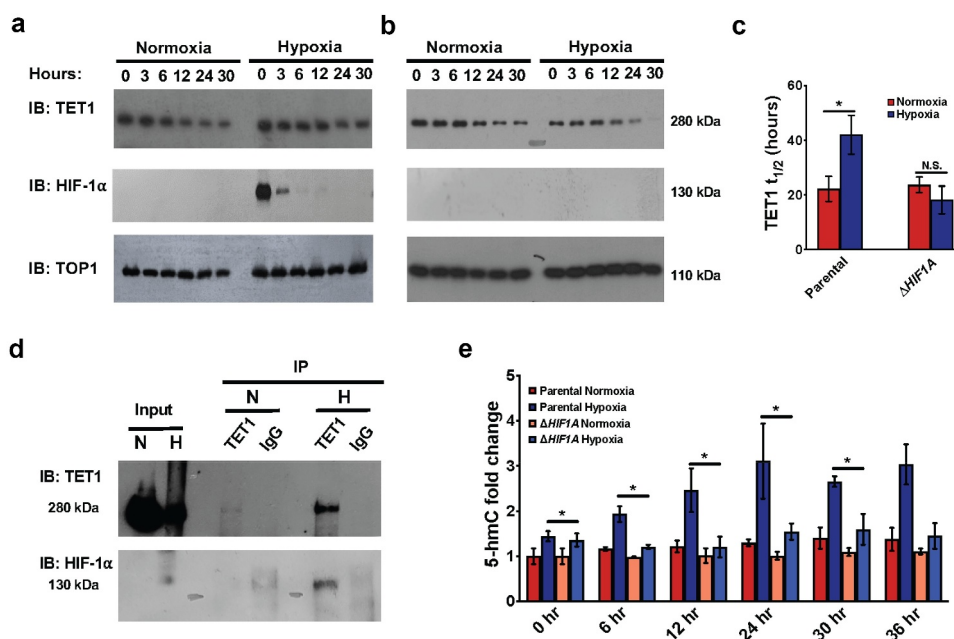


Figure 5. HIF-1 α promotes TET1 protein stability in hypoxia. (a) Western blot of TET1 degradation assay (top) in parental SK-N-BE [2] cells with detection of HIF-1 α (middle) and TOP1 (bottom), cropped at \sim 280 kDa, \sim 130 kDa, and \sim 110 kDa, respectively. Samples from normoxia and hypoxia were collected at 0, 3, 6, 12, 24, and 30 hours post treatment with cycloheximide. (b) Western blot of TET1 degradation assay (top) in gene edited $\Delta HIF1A$ SK-N-BE [2] cells with detection of HIF-1 α (middle) and TOP1 (bottom). Samples from normoxia and hypoxia were collected at 0, 3, 6, 12, 24, and 30 hours post treatment with cycloheximide. (c) Half-life calculations of TET1 protein in normoxia and hypoxia in both parental and $\Delta HIF1A$ SK-N-BE [2] cells. * represents $P < 0.05$; N.S., not significant. (d) Immunoprecipitation (IP) of TET1 protein with detection of HIF-1 α under Normoxic (n) and Hypoxic (h) conditions. IP antibodies were a-TET1 (TET1) and IgG (IgG). (e) Fold change of hypoxic 5-hmC levels over normoxic hour 0 5-hmC levels in parental and $\Delta HIF1A$ SK-N-BE [2] cells 0 to 36 hours after treatment with cycloheximide. Parental SK-N-BE [2] cells are plotted in red (normoxia) and dark blue (hypoxia). $\Delta HIF1A$ SK-N-BE [2] are plotted in coral (normoxia) and sky blue (hypoxia). Percent 5-hmC measured via UHPLC-MS/MS and is calculated relative to guanine. * represents $P < 0.05$.

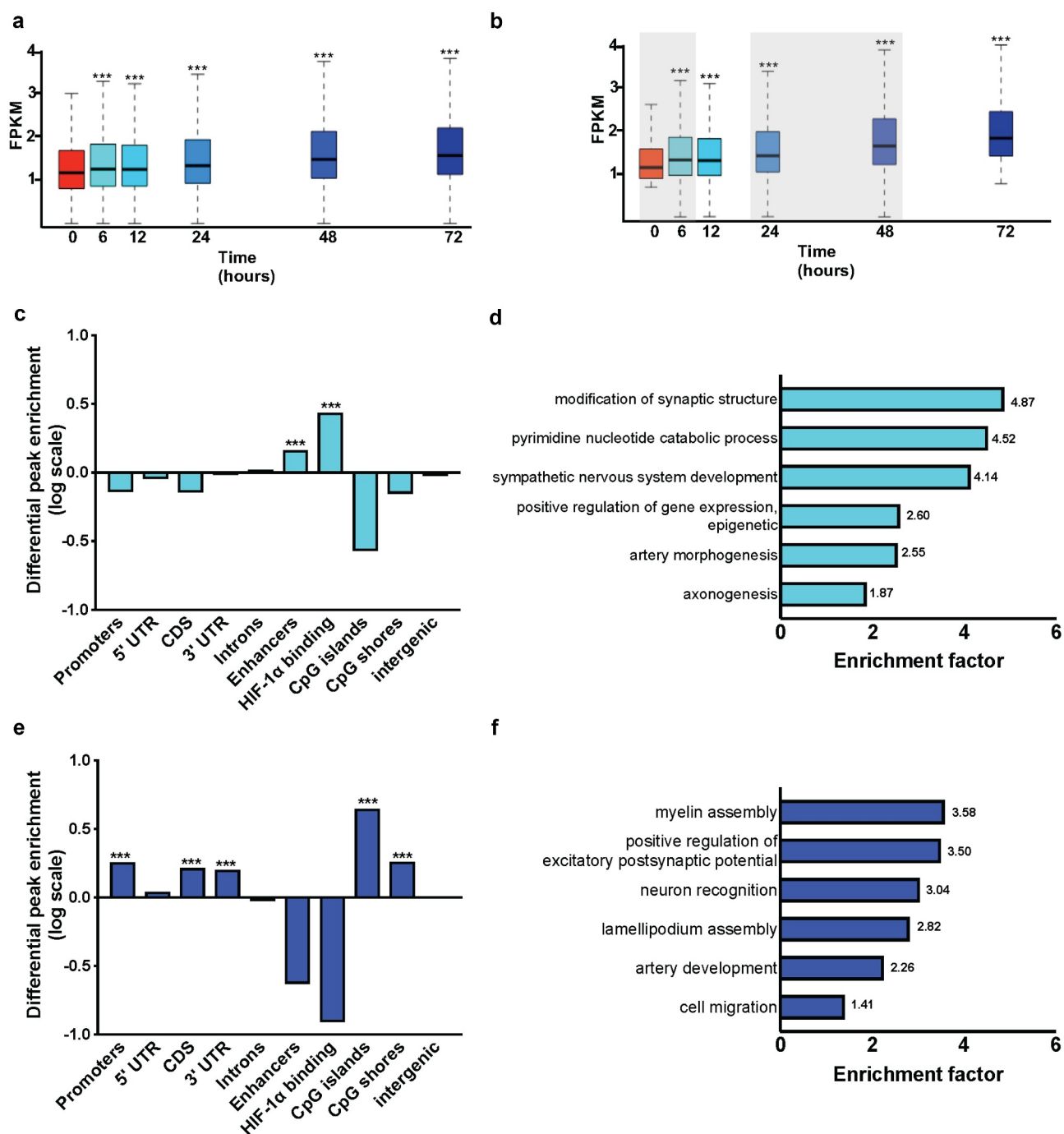


Figure 6. Accumulation of 5-hmC in hypoxic SK-N-BE [2] NB cells occurs in genes important for neuronal morphology, hypoxia adaptation, and cell migration. (a) FPKM values of all 5-hmC peaks ($n = 415,416$) plotted versus time in hypoxic culture: 0, 6, 12, 24, 48, and 72 hours. *** represents $P < 2.2 \times 10^{-16}$. (b) 5-hmC peaks ($n = 189,949$) that show a positive log fold change from 0 to 72 hours. Shaded regions indicate high fold-change between two timepoints. *** represents $P < 2.2 \times 10^{-16}$. (c) Log₂ enrichment (y-axis) of peaks ($n = 26,061$) that increase from 0 to 6 hours plotted by genomic element (x-axis): promoters, 5' untranslated regions (5' UTR), coding domain sequences (CDS), 3' untranslated regions (3' UTR), introns, enhancers, HIF-1 α binding regions (generated from our HIF-1 α ChIP-sequencing data), CpG islands, CpG shores, and intergenic regions. *** represents $P < 2.2 \times 10^{-16}$. (d) Biological processes associated with the genes associated with the 5-hmC-enriched enhancers identified in (C) determined with PANTHER statistical overrepresentation test [32,68]. The biological process is shown on the y-axis, and the enrichment factor is plotted on the x-axis. The number at the right of each process gives the enrichment factor. (e) Log₂ enrichment (y-axis) of peaks ($n = 56,190$) that increase from 0 to 6 hours plotted by genomic element (x-axis): promoters, 5' untranslated regions (5' UTR), coding domain sequences (CDS), 3' untranslated regions (3' UTR), introns, enhancers, HIF-1 α binding regions (generated from our HIF-1 α ChIP-sequencing data), CpG islands, CpG shores, and intergenic regions. *** represents $P < 2.2 \times 10^{-16}$. (f) Biological processes associated with 5-hmC-enriched CDS identified in (E) determined with PANTHER statistical overrepresentation test [32,68]. The biological process is shown on the y-axis, and the enrichment factor is plotted on the x-axis. The number at the right of each process gives the enrichment factor.

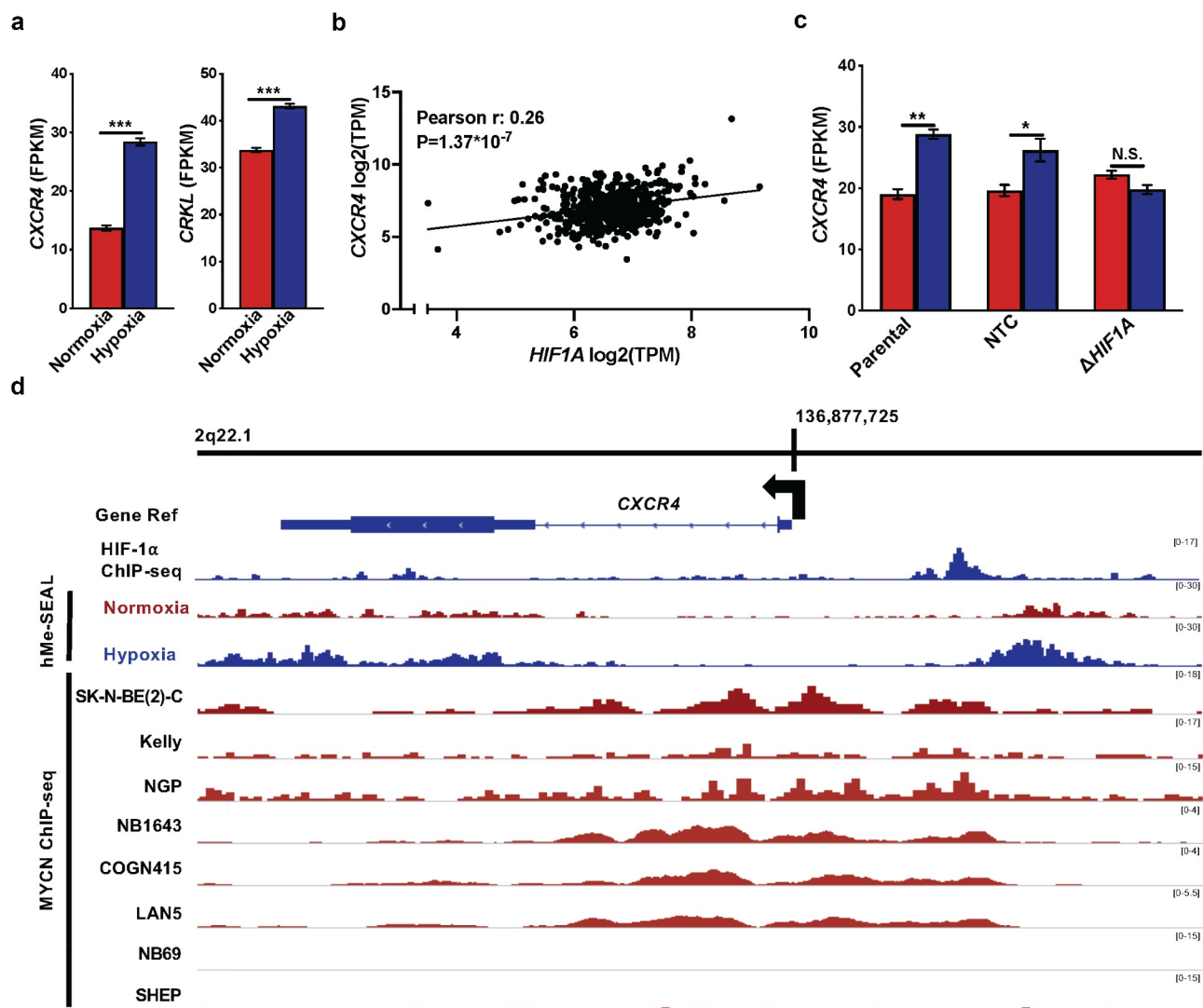


Figure 7. MYCN and HIF-1 likely bind at the *CXCR4* gene. (a) *CXCR4* (left) and *CRKL* (right) expression plotted as FPKM (y-axis) analysed from RNA-seq data from SK-N-BE [2] cells grown under normoxia (red, x-axis) versus hypoxia (dark blue, x-axis). (b) *CXCR4* versus *HIF1A* expression data plotted on the y- and x- axis respectively. r represents Pearson correlation coefficient, and P-values have been corrected with the Bonferroni correction method [67]. (c) *CXCR4* RNA-seq expression data plotted as FPKM (y-axis) from Parental, non-targeted control (NTC), and $\Delta HIF1A$ SK-N-BE [2] cells (x-axis). Red, normoxia; dark blue, hypoxia. * and ** represent $P < 0.05$ and $P < 0.01$, respectively; N.S., not significant. (d) Aligned reads along the *CXCR4* gene at locus 2q22.1 visualized with IGVtools [27]. Top track, the Gene Ref track of the *CXCR4* gene. Next, HIF-1 α ChIP-seq data from SK-N-BE [2] cells at 48 hours hypoxia; followed by tracks presenting the hMe-SEAL data obtained at 0 and 48 hours in hypoxia. The bottom eight tracks display MYCN ChIP-seq data from normoxic NB cell lines: SK-N-BE [2]-C, Kelly, NGP, NB1643, COGN415, LAN5, NB69, and SHEP [39,46]. The scale for each track is listed on the right in brackets.

a similar interaction might occur in NB cells between HIF-1 and TET1, which could promote TET1 stability. To determine if HIF-1 and TET1 bound to each other, we performed co-immunoprecipitation experiments in normoxic and hypoxic SK-N-BE [2] cells.

Immunoprecipitation of TET1 co-immunoprecipitated HIF-1 α , indicating that HIF-1 α and TET1 are part of the same protein complex (Figure 5d). Finally, we examined whether the prolonged half-life of TET1, independent of *TET1* mRNA regulation, was sufficient to result

in higher levels of total 5-hmC in hypoxic cells. We treated $\Delta HIF1A$ SK-N-BE [2] cells with cyclophosphamide to inhibit protein synthesis, cultured them under normoxic versus hypoxic conditions, and measured global 5-hmC levels by mass spectrometry. We found that hypoxic 5-hmC levels continued to increase post-cycloheximide introduction in parental cells (Figure 5e), consistent with the increased half-life of TET1 protein demonstrated in hypoxia (Figure 5c), but 5-hmC levels in hypoxic $\Delta HIF1A$ cells were lower than those in parental cells (Figure 5e), consistent with our data that a HIF-1/TET1 protein complex

increases the stability of TET1 protein (Figure 5a,c).

Hypoxic 5-hmC gains are enriched in and around genes that are important for neuronal morphology, hypoxia adaptation, and cell migration

To further our understanding of the connection between 5-hmC enrichment and initiation of the hypoxic response, we examined the genomic regions where 5-hmC density changes, either increases or decreases, in hypoxic cells over time. To characterize which loci experience altered 5-hmC density over time, we measured 5-hmC

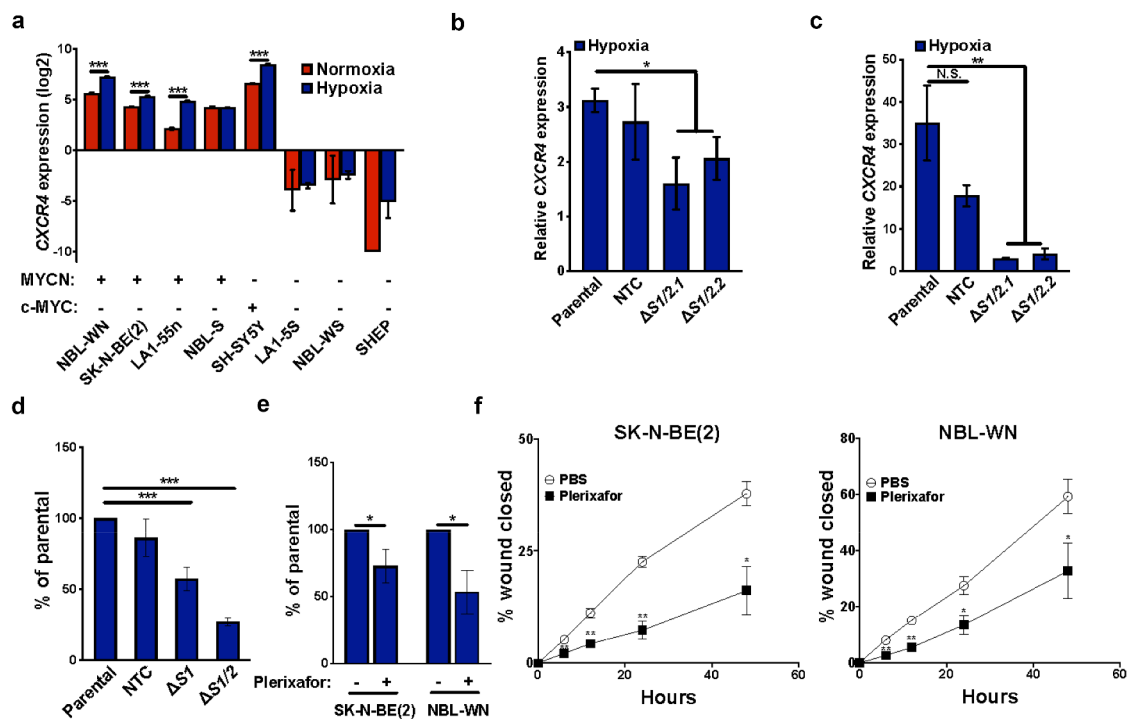


Figure 8. CXCR4 controls migration in hypoxic NB cells with high MYCN protein levels. (a) *CXCR4* expression plotted as log₂ of TPM (y-axis) from RNA-seq data from multiple NB cell lines according to MYCN protein levels (x-axis): NBL-WN, SK-N-BE [2], and LA1-55 n are *MYCN*-amplified with abundant MYCN protein [36], given as '+' in the MYCN track. SH-SY5Y is a cell line with elevated levels of c-MYC, indicated as a '+' in the c-MYC track. RNA levels were measured from cells grown under normoxic (red) or hypoxic (dark blue) conditions. *** represents $P < 0.001$. (b) Real time qPCR of *CXCR4* expression (y-axis) in Parental, non-targeted control (NTC), and $\Delta S1/2$ SK-N-BE [2] cells (x-axis) after 48 hours of hypoxia exposure ($n = 3$). P-values were determined with one-tailed t-tests, * represents $P < 0.05$. (c) Real time qPCR of *CXCR4* expression (y-axis) in Parental, non-targeted control (NTC), or $\Delta S1/2$ NBL-WN cells (x-axis) after 48 hours of hypoxia exposure ($n = 3$). ** represents $P < 0.01$; N.S., not significant. (d) Transwell migration assays were performed with SK-N-BE [2] $\Delta S1$ versus $\Delta S1/2$ cells in hypoxia. Values of $\Delta S1$, $\Delta S1/2$, and NTC were normalized to Parental cells as a percent. P-values were determined with one-tailed t-tests ($n = 3$); * and *** represent $P < 0.05$ and $P < 0.001$, respectively. (e) Transwell assays performed in hypoxia with parental SK-N-BE [2] (left) and NBL-WN (right) cells and treated with (+) or without (-) 10 mg/mL plerixafor ($n = 4$). P-values were determined with one-tailed t-tests, * represents $P < 0.05$. (f) Wound healing assays performed with parental SK-N-BE [2] (left) and NBL-WN (right) cells and treated with 10 mg/mL plerixafor (black squares) or phosphate buffered saline (PBS, open circles). Both control and experimental cells were treated with 30 μ M FG-4592 to induce pseudo-hypoxia. P-values were determined with one-tailed t-tests; * and ** represent $P < 0.05$ and $P < 0.01$, respectively.

distribution changes using hMe-SEAL [29] in *MYCN*-amplified SK-N-BE [2] cells exposed to hypoxia for 0, 6, 12, 24, 48, or 72 hours. We found that mean FPKM at each time point increased over time (Figure 6a), consistent with our prior observations that global 5-hmC levels increase with time in hypoxia (figure 4f and Figure 5e) and our prior work in these cells [12]. To focus on hypoxic induction of genes that mediate the hypoxic programme, we examined genomic loci that experience 5-hmC gains in hypoxia. We removed all FPKM values that did not increase in hypoxia over time and identified ‘early’ peaks that gained 5-hmC density in the first 6 hours and ‘late’ peaks that gained 5-hmC density after 24 hours (Figure 6b). Analysis of the early gains showed that they were enriched in enhancer and HIF-1 binding regions (Figure 6c). We extracted the gene targets of the enriched enhancers in the early gains, and found over-representation GO-terms [32] in biological processes related to neuronal morphology and development (Figure 6d), consistent with the observation that NB cells de-differentiate and undergo morphological changes in hypoxia [55]. Also, over-represented were processes related to the hypoxic responses, such as artery morphogenesis (Figure 6d). Analysis of the late gains showed that they were enriched in different genomic regions than those of early gains. The late gains were enriched in promoters, 5’ untranslated regions, coding domain sequences, 3’ untranslated regions, CpG islands, and CpG shores (Figure 6e). Among the late gains, we found that many of these over-represented targets were part of neuronal development, hypoxia biology, and cell migration (figure 6f).

CXCR4 is critical for the migration of NB cells in hypoxia

We focused on determining which genes are controlled by changes in 5-hmC density and contribute to cell migration as a function of hypoxia, since this phenotype is associated with aggressive cancer cell behaviour and metastasis [56]. To identify candidate genes involved in cell migration, we merged the list of hypoxic 5-hmC-enriched genes from our pathway analysis with our RNA-seq data from normoxic and hypoxic parental SK-N-BE [2]

cells, and identified two 5-hmC-enriched genes with hypoxia-augmented expression: *CXCR4* and *CRKL* (Figure 7a).

We focused on *CXCR4*, since it showed the highest fold induction in hypoxia, 2.1-fold, and prior work had identified *CXCR4* as an important hypoxic target in NB cells [17]. *CXCR4* encodes a surface receptor that functions in the retention of haematopoietic stem cells in the bone marrow and chemotactic guidance in neural progenitor cells [57]. In NB, *CXCR4* expression is correlated with metastatic spread and worse outcome [58]. Consistent with prior work showing that hypoxia induces *CXCR4* expression [17,59], we compared *CXCR4* and *HIF1A* expression in NB tumours [37] and found a positive correlation between them ($r = 0.26$, $P = 1.37 \times 10^{-7}$, Figure 7b). However, because HIF-1 α is regulated largely by post-translational modifications rather than at the transcriptional level, we tested whether *CXCR4* was induced in Δ *HIF1A* cells that lack HIF-1 α protein (Figure 5b). We found that HIF-1 is required for induction of *CXCR4* expression in hypoxia (Figure 7c). We examined our HIF-1 α ChIP-seq and hMe-SEAL data as well as *MYCN* ChIP-seq data [39,46] at the *CXCR4* locus and identified a HIF-1 α peak near the promoter of *CXCR4*, suggesting that HIF-1 regulates *CXCR4* directly (Figure 7d). In addition, we noted that *CXCR4* gains 5-hmC density throughout the gene body, consistent with its increased transcription (Figure 7a,c). Additionally, ChIP-seq for *MYCN* in multiple *MYCN*-amplified NB cell lines [39,46] shows multiple binding peaks across the gene promoter and first intron, suggesting that *MYCN* co-regulates *CXCR4* in normoxia (Figure 7d).

We extended our analysis of *CXCR4* expression to include RNA-seq data from several NB cell lines with varying levels of *MYCN* gene amplification and corresponding protein levels. *CXCR4* was induced in hypoxia in NB lines with abundant *MYCN* protein and one (SH-SY5Y), with c-MYC [36] (Figure 8a). In NBL-S cells, which express *MYCN* protein at levels lower than in *MYCN*-amplified SK-N-BE [2]-C cells [60], *CXCR4* was not induced in hypoxia but had relatively high baseline expression (Figure 8a). To test whether *CXCR4* expression depends on an intact *MYCN*/TET1/HIF-1 axis, we measured *CXCR4* expression

in $\Delta S1/2$ SK-N-BE [2] and NBL-WN cells and observed that *CXCR4* expression was virtually eliminated in $\Delta S1/2$ cell lines grown in hypoxia (Figure 8b,c).

Because prior work has demonstrated that *CXCR4* contributes to NB cell migration [58,61,62], we performed transwell assays in hypoxic SK-N-BE [2] $\Delta S1/2$ cells. We observed that $\Delta S1/2$ cells migrated slower in hypoxia compared to controls (Figure 8d). $\Delta S1$ cells, with an intermediate level of *CXCR4* induction (Figure 8b) also demonstrated decreased migration in hypoxia, albeit not as much as $\Delta S1/2$ cells (Figure 8d) as well as defective wound healing, another means of measuring cell migration (Fig. S9A,B). The cell migration phenotype was specific to hypoxic cells and was not observed in normoxia (Fig. S9C), suggesting that *CXCR4* is involved specifically in hypoxia-directed migration of NB cells. To confirm that this was not the result of altered proliferation, we compared the growth rates of $\Delta S1$ versus control cells and found no difference in growth rate in either normoxia or hypoxia (Fig. S9D,E).

To test if *CXCR4* promotes cell migration directly, we treated hypoxic SK-N-BE [2] and NBL-WN cell lines with plerixafor, a *CXCR4* antagonist [63]. To measure cell migration, we used two complementary approaches: transwell assays performed in hypoxia and wound healing studies carried out in pseudo-hypoxia. Pseudo-hypoxia was achieved by treating cells with FG-4592, a HIF- α prolyl hydroxylase inhibitor, and confirming HIF-1 α stabilization. We found that plerixafor-treated cells migrated slower than their control counterparts in hypoxia and pseudo-hypoxia (Figures 8E,f and Fig. S10), indicating that *CXCR4* directs *MYCN*-amplified NB cell migration in hypoxia.

Discussion

In this study, we describe how the *TET1* gene is transcriptionally regulated by *MYCN* and HIF-1 in *MYCN*-amplified NBs and identify *CXCR4* as a gene induced by hypoxia via *MYCN*/*TET1*-induced 5-hmC deposition, which functions as a driver of NB cell migration in hypoxia (Fig. S11). In normoxia, *TET1* is bound by *MYCN* at

two sites in the first and second introns, *S1* and *S2*. Moreover, *MYCN* drives the expression of the *TET1fl* transcript preferentially. Single deletions of either *S1* or *S2* lower transcription by 30%, but when both sites are deleted, *TET1* expression is reduced by 80%, suggesting that both *S1* and *S2* sites are needed for full induction of the gene. Similarly, in hypoxia, *S1* and *S2* serve as docking sites for HIF-1, which further drives *TET1* transcription. The *MYCN*/*TET1*/HIF-1 axis is further enforced by complex formation by *TET1* and HIF-1, which extends the half-life of *TET1*, a means of maintaining its catalytic activity, which is dependent on oxygen [52].

Several questions remain concerning the role of *TET1* isoforms in NB. Studies that have compared the roles of *TET1fl* versus *TET1s* have identified *TET1fl* as the more 'efficient' isoform and hypothesized its presence is necessary for large-scale epigenetic reprogramming, such as during cell differentiation [47–49]. Future experiments will determine if *TET1* isoforms have distinct roles in NB by examining their impact on the 5-hmC landscape and their relationship with prognosis.

TET1 protein stability is augmented in hypoxia due its binding interaction with HIF-1 α . Our previous findings that *TET1* modifies DNA at HIF-1 binding sites in canonical HIF-1 targets [12] also suggest that the HIF-1/*TET1* interaction may be required to recruit *TET1* to specific genomic sites. Future experiments will determine whether the loss of this interaction changes the 5-hmC landscape and whether other proteins participate in the HIF-1 α /*TET1* complex. Other possible mechanisms of *TET1* protein regulation could also be the subject of future work, including identifying whether any post-translational modifications (PTMs) [64] contribute to the stability and/or activity of *TET1*. Although multiple PTMs of *TET* proteins have been described [65,66], their functional roles could be defined further in future work. Our prior observation that 5-hmC augments expression from HIF-1 targets in *MYCN*-amplified NBs prompted us to investigate 5-hmC profiles from both NB tumour and cell-free DNA. We determined 5-hmC profiles can serve as DNA-based biomarkers that are predictive of metastatic burden and prognostic of overall outcome [7].

Here, we determined that 5-hmC accumulation in hypoxic NB cell lines occurs in genes involved in cell migration, identifying *CXCR4* as a specific hypoxia target gene that mediates this phenotype and may serve as a useful biomarker for aggressive disease and a therapeutic target in *MYCN*-amplified NBs. *CXCR4* has been previously identified in gene expression studies of primary NB tumour samples, correlating with metastatic spread and worse outcome [58,62], but in other solid tumour cell lines, *CXCR4* has been shown to promote or activate MAPK/ERK, JAK/STAT, and PI3K/AKT pathways, resulting in cell migration and survival [63]. Future studies will elucidate the mechanism by which *CXCR4* and its downstream targets augment cell migration in hypoxia in NB. Additionally, future studies could assess whether the addition of a *CXCR4* antagonist to the NB treatment regimen is for the purpose of slowing cell migration/metastasis, rather than as a haematopoietic stem cell mobilizing agent as it is traditionally given, to improve patient outcomes.

Author contributions

A.E.H. designed and performed experiments, analysed data, and wrote the manuscript. S.U. performed experiments. J.Z. C. analysed the data and helped perform experiments. M.A. A. analysed and contributed sequencing data. H.R.S. helped design experiments. L.A.G. and S.L.C. conceived the study, analysed the data, and edited the manuscript.

Acknowledgments

We thank Dr. Paul Geeleher for performing statistical analysis of wound-healing assay data sets. We would also like to thank Dr. Marsha Rosner and Dr. Chilong Nguyen for their invaluable instruction of several protein biochemistry techniques. Finally, we would like to thank Dr. Julie Losman for her CRISPR plasmids targeting *HIF1A*.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

This work was supported by Alex's Lemonade Stand and a Northwestern Mutual Young Investigator Grant. A.E. H. was supported by The Molecular and Cellular Biology

Training Grant, and the HHMI Med-into-Grad Translational Training Program. M.A.A. was supported by the National Institutes of Health, K08CA226237.

Data availability statement

Datasets generated for this study can be found at the Gene Expression Omnibus under the SuperSeries accession number GSE167478.

GEO accession numbers for publicly available ChIP-seq datasets used in this study are GSE80151 (39), GSE138315 (46), and GSE71399 (51). Sequencing data for NB cell lines can be found on GEO databank (GSE89413) and sequencing for NB tumours can be accessed through R2: Genomics Analysis and Visualization Platform (<https://r2.amc.nl/>) (37,38). NB cell line names can be found in Table S2. RNA-seq data from JQ1 experiments can also be accessed under GSE80154 (39). RNA-seq data from MLN8237 studies can be available under GSE57810 (40).

References

- [1] Van Arendonk KJ, Chung DH. Neuroblastoma: tumor biology and its implications for staging and treatment. *Children* (Basel). 2019;6.
- [2] Moreno L, Guo D, Irwin MS, et al. A nomogram of clinical and biologic factors to predict survival in children newly diagnosed with high-risk neuroblastoma: an International Neuroblastoma Risk Group project. *Pediatr Blood Cancer*. 2021;68:e28794.
- [3] Bellini A, Bernard V, Leroy Q, et al. Deep sequencing reveals occurrence of subclonal ALK mutations in neuroblastoma at diagnosis. *Clin Cancer Res*. 2015;21:4913–4921.
- [4] Pinto N, Mayfield JR, Raca G, et al. Segmental chromosomal aberrations in localized neuroblastoma can be detected in formalin-fixed paraffin-embedded tissue samples and are associated with recurrence. *Pediatr Blood Cancer*. 2016;63:1019–1023.
- [5] Maleki Vareki S. High and low mutational burden tumors versus immunologically hot and cold tumors and response to immune checkpoint inhibitors. *J Immunother Cancer*. 2018;6:157.
- [6] Pugh TJ, Morozova O, Attiyeh EF, et al. The genetic landscape of high-risk neuroblastoma. *Nat Genet*. 2013;45:279–284.
- [7] Applebaum MA, Barr EK, Karpus J, et al. 5-Hydroxymethylcytosine profiles in circulating cell-free DNA associate with disease burden in children with neuroblastoma. *Clin Cancer Res*. 2020;26:1309–1317.
- [8] Kriaucionis S, Heintz N. The nuclear DNA base 5-hydroxymethylcytosine is present in Purkinje neurons and the brain. *Science*. 2009;324:929–930.
- [9] Cao JZ, Hains AE, and Godley LA. Regulation of 5-hydroxymethylcytosine distribution by the TET

- enzymes. In: Jurga S, Barciszewski J, editors. *The DNA, RNA, and histone methylomes*. Cham, Switzerland: Springer International Publishing; 2019. p. 229–263.
- [10] Ficiz G, Branco MR, Seisenberger S, et al. Dynamic regulation of 5-hydroxymethylcytosine in mouse ES cells and during differentiation. *Nature*. 2011;473:398–402.
- [11] Madzo J, Liu H, Rodriguez A, et al. Hydroxymethylation at gene regulatory regions directs stem/early progenitor cell commitment during erythropoiesis. *Cell Rep*. 2014;6:231–244.
- [12] Mariani CJ, Vasanthakumar A, Madzo J, et al. TET1-mediated hydroxymethylation facilitates hypoxic gene induction in neuroblastoma. *Cell Rep*. 2014;7:1343–1352.
- [13] He YF, Li BZ, Li Z, et al. Tet-mediated formation of 5-carboxylcytosine and its excision by TDG in mammalian DNA. *Science*. 2011;333:1303–1307.
- [14] Ito S, D'Alessio AC, Taranova OV, et al. Role of Tet proteins in 5mC to 5hmC conversion, ES-cell self-renewal and inner cell mass specification. *Nature*. 2010;466:1129–1133.
- [15] Rankin EB, Giaccia AJ. The role of hypoxia-inducible factors in tumorigenesis. *Cell Death Differ*. 2008;15:678–685.
- [16] Iliopoulos O, Levy AP, Jiang C, et al. Negative regulation of hypoxia-inducible genes by the von Hippel-Lindau protein. *Proc Natl Acad Sci U S A*. 1996;93:10595–10599.
- [17] Huertas-Castaño C, Gómez-Muñoz MA, Pardal R, et al. Hypoxia in the initiation and progression of neuroblastoma tumours. *Int J Mol Sci*. 2019;21:39.
- [18] Lutz W, Stöhr M, Schürmann J, et al. Conditional expression of N-myc in human neuroblastoma cells increases expression of alpha-prothymosin and ornithine decarboxylase and accelerates progression into S-phase early after mitogenic stimulation of quiescent cells. *Oncogene*. 1996;13:803–812.
- [19] Concordet JP, Haeussler M. CRISPOR: intuitive guide selection for CRISPR/Cas9 genome editing experiments and screens. *Nucleic Acids Res*. 2018;46:W242–w245.
- [20] Sanjana NE, Shalem O, Zhang F. Improved vectors and genome-wide libraries for CRISPR screening. *Nat Methods*. 2014;11:783–784.
- [21] Moffat J, Grueneberg DA, Yang X, et al. A lentiviral RNAi library for human and mouse genes applied to an arrayed viral high-content screen. *Cell*. 2006;124:1283–1298.
- [22] Camenisch G, Wenger RH, and Gassmann M. DNA binding activity of hypoxia-inducible factors. In: Armstrong D, editor. *Oxidants and Antioxidants: ultrastructure and molecular biology protocols*. Totowa, New Jersey: Humana Press; 2002. p. 117–129.
- [23] Lignitto L, LeBoeuf SE, Homer H, et al. Nrf2 activation promotes lung cancer metastasis by inhibiting the degradation of Bach1. *Cell*. 2019;178:316–329. e318.
- [24] Schneider CA, Rasband WS, Eliceiri KW. NIH image to imageJ: 25 years of image analysis. *Nat Methods*. 2012;9:671–675.
- [25] Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics*. 2009;25:1754–1760.
- [26] Zhang Y, Liu T, Meyer CA, et al. Model-based analysis of ChIP-Seq (MACS). *Genome Biol*. 2008;9:R137.
- [27] Robinson JT, Thorvaldsdóttir H, Winckler W, et al. Integrative genomics viewer. *Nat Biotechnol*. 2011;29:24–26.
- [28] Song L, James SR, Kazim L, et al. Specific method for the determination of genomic DNA methylation by liquid chromatography-electrospray ionization tandem mass spectrometry. *Anal Chem*. 2005;77:504–510.
- [29] Song CX, Szulwach KE, Fu Y, et al. Selective chemical labeling reveals the genome-wide distribution of 5-hydroxymethylcytosine. *Nat Biotechnol*. 2011;29:68–72.
- [30] Anders S, Pyl PT, Huber W. HTSeq—a Python framework to work with high-throughput sequencing data. *Bioinformatics*. 2015;31:166–169.
- [31] Quinlan AR, Hall IM. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics*. 2010;26:841–842.
- [32] Thomas PD, Campbell MJ, Kejariwal A, et al. PANTHER: a library of protein families and subfamilies indexed by function. *Genome Res*. 2003;13:2129–2141.
- [33] Trapnell C, Roberts A, Goff L, et al. Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. *Nat Protoc*. 2012;7:562–578.
- [34] Souvorov A, Kapustin Y, Kiryutin B, et al. (2010) Gnomon—NCBI eukaryotic gene prediction tool. *National Center for Biotechnology Information*, 1–24.
- [35] Justus CR, Leffler N, Ruiz-Echevarria M, et al. In vitro cell migration and invasion assays. *J Vis Exp*. 2014. DOI:10.3791/51046
- [36] Chlenski A, Park C, Dobratic M, et al. Maternal embryonic leucine zipper kinase (MELK), a potential therapeutic target for neuroblastoma. *Mol Cancer Ther*. 2019;18:507–516.
- [37] Gartlgruber M, Sharma AK, Quintero A, et al. Super enhancers define regulatory subtypes and cell identity in neuroblastoma. *Nat Cancer*. 2021;2:114–128.
- [38] Harenza JL, Diamond MA, Adams RN, et al. Transcriptomic profiling of 39 commonly-used neuroblastoma cell lines. *Sci Data*. 2017;4:170033.
- [39] Zeid R, Lawlor MA, Poon E, et al. Enhancer invasion shapes MYCN-dependent transcriptional amplification in neuroblastoma. *Nat Genet*. 2018;50:515–523.
- [40] Althoff K, Beckers A, Bell E, et al. A Cre-conditional MYCN-driven neuroblastoma mouse model as an

- improved tool for preclinical studies. *Oncogene*. 2015;34:3357–3368.
- [41] Heinz S, Benner C, Spann N, et al. Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. *Mol Cell*. 2010;38:576–589.
- [42] Gao T, He B, Liu S, et al. EnhancerAtlas: a resource for enhancer annotation and analysis in 105 human cell/tissue types. *Bioinformatics*. 2016;32:3543–3551.
- [43] Whyte WA, Orlando DA, Hnisz D, et al. Master transcription factors and mediator establish super-enhancers at key cell identity genes. *Cell*. 2013;153:307–319.
- [44] Kent WJ, Sugnet CW, Furey TS, et al. The human genome browser at UCSC. *Genome Res*. 2002;12:996–1006.
- [45] Neri F, Incarnato D, Krepelova A, et al. TET1 is controlled by pluripotency-associated factors in ESCs and downmodulated by PRC2 in differentiated cells and tissues. *Nucleic Acids Res*. 2015;43:6814–6826.
- [46] Upton K, Modi A, Patel K, et al. Epigenomic profiling of neuroblastoma cell lines. *Sci Data*. 2020;7:116.
- [47] Zhang W, Xia W, Wang Q, et al. Isoform switch of TET1 regulates DNA demethylation and mouse development. *Mol Cell*. 2016;64:1062–1073.
- [48] Good CR, Madzo J, Patel B, et al. A novel isoform of TET1 that lacks a CXXC domain is overexpressed in cancer. *Nucleic Acids Res*. 2017;45:8269–8281.
- [49] Greer CB, Wright J, Weiss JD, et al. Tet1 isoforms differentially regulate gene expression, synaptic transmission, and memory in the mammalian brain. *J Neurosci*. 2021;41:578–593.
- [50] Van groningen T, Koster J, Valentijn LJ, et al. Neuroblastoma is composed of two super-enhancer-associated differentiation states. *Nat Genet*. 2017;49:1261–1266.
- [51] Thienpont B, Steinbacher J, Zhao H, et al. Tumour hypoxia causes DNA hypermethylation by reducing TET activity. *Nature*. 2016;537:63–68.
- [52] Laukka T, Mariani CJ, Ihantola T, et al. Fumarate and succinate regulate expression of hypoxia-inducible genes via TET enzymes. *J Biol Chem*. 2016;291:4256–4265.
- [53] Costa Y, Ding J, Theunissen TW, et al. NANOG-dependent function of TET1 and TET2 in establishment of pluripotency. *Nature*. 2013;495:370–374.
- [54] Lio CW, Zhang J, González-Avalos E, et al. Tet2 and Tet3 cooperate with B-lineage transcription factors to regulate DNA modification and chromatin accessibility. *Elife*. 2016;5.
- [55] Bhaskara VK, Mohanam I, Rao JS, et al. Intermittent hypoxia regulates stem-like characteristics and differentiation of neuroblastoma cells. *PLoS One*. 2012;7:e30905.
- [56] Jögi A, Vallon-Christersson J, Holmquist L, et al. Human neuroblastoma cells exposed to hypoxia: induction of genes associated with growth, survival, and aggressive behavior. *Exp Cell Res*. 2004;295:469–487.
- [57] Domanska UM, Kruizinga RC, Nagengast WB, et al. A review on CXCR4/CXCL12 axis in oncology: no place to hide. *Eur J Cancer*. 2013;49:219–230.
- [58] Russell HV, Hicks J, Okcu MF, et al. CXCR4 expression in neuroblastoma primary tumors is associated with clinical presentation of bone and bone marrow metastases. *J Pediatr Surg*. 2004;39:1506–1511.
- [59] Korbecki J, Kojder K, Kapczuk P, et al. The effect of hypoxia on the expression of CXC chemokines and CXC chemokine receptors-A review of literature. *Int J Mol Sci*. 2021;22:843.
- [60] Cohn SL, Salwen H, Quasney MW, et al. Prolonged N-myc protein half-life in a neuroblastoma cell line lacking N-myc amplification. *Oncogene*. 1990;5:1821–1827.
- [61] Meier R, Mühlethaler-Mottet A, Flahaut M, et al. The chemokine receptor CXCR4 strongly promotes neuroblastoma primary tumour and metastatic growth, but not invasion. *PLoS One*. 2007;2:e1016.
- [62] Klein S, Abraham M, Bulvik B, et al. CXCR4 promotes neuroblastoma growth and therapeutic resistance through miR-15a/16-1-mediated ERK and BCL2/Cyclin D1 pathways. *Cancer Res*. 2018;78:1471–1483.
- [63] Chatterjee S, Behnam Azad B, Nimmagadda S. The intricate role of CXCR4 in cancer. *Adv Cancer Res*. 2014;124:31–82.
- [64] Hornbeck PV, Zhang B, Murray B, et al. PhosphoSitePlus, 2014: mutations, PTMs and recalibrations. *Nucleic Acids Res*. 2015;43:D512–520.
- [65] Bauer C, Göbel K, Nagaraj N, et al. Phosphorylation of TET proteins is regulated via O-GlcNAcylation by the O-linked N-acetylglucosamine transferase (OGT). *J Biol Chem*. 2015;290:4801–4812.
- [66] Jeong JJ, Gu X, Nie J, et al. Cytokine-regulated phosphorylation and activation of TET2 by JAK2 in hematopoiesis. *Cancer Discov*. 2019;9:778–795.
- [67] Bland, JM , and Altman, DG. Multiple significance tests: the Bonferroni method. *BMJ*. 1995;310 (6973):170. doi:10.1136/bmj.310.6973.170.
- [68] Mi H, Muruganujan A, Casagrande JT, et al. Large-scale gene function analysis with the PANTHER classification system. *Nat Protoc*. 2013;8:1551–1566.