Epigenomic profiling of neuroblastoma cell lines

Kristen Upton[1], Apexa Modi[1,2], Khushbu Patel[3], Karina L. Conkrite[1], Robyn T. Sussman[1], Gregory P. Way[2], Rebecca N. Adams[4], Gregory I. Sacks[1], Paolo Fortina[4], John M. Maris[1,5], Jo Lynne Rokita[1,3,6]

1. Division of Oncology, Children's Hospital of Philadelphia, Philadelphia, Pennsylvania 19104

2. Genomics and Computational Biology Graduate Group, University of Pennsylvania, Philadelphia, PA, 19104, USA

3. Department of Bioinformatics and Health Informatics, Children's Hospital of Philadelphia, Philadelphia, Pennsylvania 19104, USA

4. Cancer Genomics and Bioinformatics Laboratory, Sidney Kimmel Cancer Center, Philadelphia, Pennsylvania 19107, USA

5. Department of Pediatrics, Perelman School of Medicine at the University of Pennsylvania, Philadelphia, Pennsylvania 19104, USA

6. Center for Data-Driven Discovery in Biomedicine, Children's Hospital of Philadelphia, Philadelphia, Pennsylvania 19104, USA

Corresponding author(s): Jo Lynne Rokita (rokita@email.chop.edu), John Maris (maris@email.chop.edu)

Abstract

Understanding the aberrant transcriptional landscape of neuroblastoma is necessary to provide insight to the underlying influences of the initiation, progression and persistence of this developmental cancer. Here, we present chromatin immunoprecipitation sequencing (ChIP-Seq) data for the oncogenic transcription factors, MYCN and MYC, as well as regulatory histone marks H3K4me1, H3K4me3, H3K27Ac, and H3K27me3 in ten commonly used human neuroblastoma-

derived cell line models. In addition, for all of the profiled cell lines we provide ATAC-Seq as a measure of open chromatin. We validate specificity of global MYCN occupancy in MYCN amplified cell lines and functional redundancy of MYC occupancy in MYCN non-amplified cell lines. Finally, we show with H3K27Ac ChIP-Seq that these cell lines retain expression of key neuroblastoma super-enhancers (SE). We anticipate this dataset, coupled with available transcriptomic profiling on the same cell lines, will enable the discovery of novel gene regulatory mechanisms in neuroblastoma.

Background & Summary

An estimated 15,780 children in the United States will be diagnosed with cancer in 2019 (ACCO 2016). While 80% of pediatric cancer patients overcome this disease, 20% of children do not survive, and survivors often have multiple side effects of therapy (ACCO 2016). Neuroblastoma accounts for more than 7% of malignancies in patients under 15 years of age and approximately 12% of all pediatric cancer-related deaths (for review (Matthay et al. 2016)). Neuroblastoma shows wide phenotypic variability, with tumors arising in children diagnosed under the age of 18 months often spontaneously regressing with little or no treatment, but patients diagnosed at an older age or with unfavorable genomic features often showing a relentlessly progressive and widely metastatic disease pattern despite intensive, multimodal therapy (for review see (Matthay et al. 2016; Maris 2010)). Ninety-eight percent of low-risk neuroblastoma disease are currently cured (Twist et al. 2019), however, the survival rate for patients with high-risk neuroblastoma remains less than 50% (Pinto et al. 2015). Relapsed high-risk neuroblastoma is typically incurable (Simon et al. 2011; Ambros et al. 2009), and thus these children require improved therapeutic options.

A major prognostic factor predicting the severity, risk, and inferior outcome for neuroblastoma patients is amplification of the proto-oncogene *MYCN*. *MYCN* amplification occurs in nearly 20% of all neuroblastomas, and approximately 50% of patients with high-risk disease

(Gherardi et al. 2013; Rickman, Schulte, and Eilers 2018). It is a truncal genomic event, and typical stable across the spectrum of therapy and disease recurrence. MYCN, along with structural and binding homologues MYC and MYCL, are members of the MYC transcription factor family (Pistoia et al. 2012) and have been implicated in transcriptional regulation of proteins involved in cell growth (Coller et al. 2000), proliferation (Ji et al. 2011), and ribosome biogenesis (Ji et al. 2011). Mounting evidence has also indicated that MYCN and MYC are functionally redundant (He et al. 2013; Malynn et al. 2000; Chappell and Dalton 2013). However, the protein expression of MYC and MYCN appears to be mutually exclusive. For example, neuroblastoma tumors with *MYCN* amplification typically lack or have low *MYC* mRNA expression (Rickman, Schulte, and Eilers 2018). The strong influence of MYCN on the progression and metastasis of neuroblastoma makes it a key target for therapy, but due to its global transcriptional activity, it is necessary to develop a better understanding of which of its gene targets directly influence oncogenesis.

To better understand the regulatory effects of MYC family proteins in neuroblastoma, we performed ChIP-Seq data for MYCN in six neuroblastoma cell lines with *MYCN* amplification, MYC in four neuroblastoma cell lines without *MYCN* amplification, and H3K27Ac, H3K27me3, H3K4me1, and H3K4me3 histone modifications along with ATAC-Seq in all ten neuroblastoma cell lines (with ATAC data in four additional lines also reported here). All of the cell lines here also have RNA sequencing data freely available (Harenza et al. 2017).

Methods

Online Table 1 summarizes which assays were performed for each cell line, and an overview of the workflow is shown in Figure 1.

Cell Growth and Expansion

The cell lines used to collect this data were obtained from multiple sources: the Children's Oncology Group (COG) Cell Culture and Xenograft Repository at Texas Tech University Health Sciences Center (www.cccells.org), the American Type Culture Collection (Manassas, MA), or the Children's Hospital of Philadelphia (CHOP) cell line bank. All of the cell growth and preparations were done at CHOP. The neuroblastoma cell lines were cultured using media (Metadata Table 1) and methods as previously described (Harenza et al. 2017). Briefly, cells were thawed by floating in a 37°C water bath for 2-3 minutes. Cells were then added to a 15 mL conical tube, containing 5 mL of the appropriate growth media, and centrifuged at 300 *xg* for 5 minutes at room temperature (RT). Media was then carefully aspirated off, and the pelleted cells were resuspended in 1 mL of media before being transferred to a 75 mm^2 flask containing 10 mL of growth media. Cells were incubated at 37°C with a 5.0% CO_2 concentration. When cells reach 70-80% confluency, media was aspirated off and cells were gently washed with 1X PBS. Following aspiration of the PBS, 3 mL of the appropriate detachment solution (noted in Metadata Table 1) was added and the flask was incubated at 37°C for 2-5 minutes. Cells were then gathered by tilting the plate at a 45° angle and washing with at least 4 mL of the appropriate growth media, and transferred to a 15 mL conical. After centrifugation for 5 minutes at 300 *x g*. Media was aspirated off, and the pellet was resuspended in 1 mL of growth media and transferred to a 150 mm cell culture dish containing 19 mL of growth media. Cells were incubated at 37°C with a 5.0% CO_2 concentration until reaching 70-80% confluency. Necessary materials and reagents are listed in Online Table 2.

ChIP-Seq Protocol

The ChIP-Seq Protocol is separated into four sections: Cell Fixation, Chromatin Immunoprecipitation (ChIP) and Library Preparation, Library Sequencing, and ChIP-Seq Analysis. Of note, the MYCN ChIP-Seq for Kelly and NGP cell lines was performed using a varied

procedure and is noted in a separate section within this protocol. Necessary materials and reagents are listed in Online Table 2.

Cell Fixation

Cells were grown as described in Cell Growth and Expansion section of protocol to 70-80% confluence in 150 mm tissue culture plates in 20 mL of media. The Formaldehyde solution (Online Table 2) was freshly prepared. Cells were removed from incubation and 1/10th of the growth media volume of the Formaldehyde Solution was added to the existing media in the plate (i.e. if the current volume of the plate is 20 mL of media, 2 mL of Formaldehyde Solution would be added). The solution was gently swirled, and then rocked at RT for 15 minutes. To stop the fixation, 1/20th the current volume of the Glycine Solution (Online Table 2) was added to the plate (i.e. if the current volume in the plate is 22 mL then 1.1 mL of Glycine Solution should be added). The plate was gently swirled to mix, and then allowed to sit at RT for 5 minutes. Following this incubation, a cell scraper was used to collect the cells, and then all cells and solution were transferred to a 50 mL conical on ice. From this point forward, all samples were kept on ice. The 50 mL conical was centrifuged at 800 *x g* at 4°C for 10 minutes to pellet the cells. Supernatant was removed and discarded, and the cells were resuspended with 10 mL of chilled, sterile PBS. Centrifugation of the tube at 800 *x g* at 4°C for 10 minutes was repeated. The supernatant was removed and discarded, and the cells were resuspended with 10 mL of chilled, sterile PBS with 100 uL of PMSF. The tube was centrifuged at 800 *x g* at 4°C for 10 minutes, the supernatant was removed, and then the cells were snap frozen on dry ice and stored at -80°C. The cells were then shipped to Active Motif on dry ice following the instructions listed at on the Sample Submission Form, downloaded from www.activemotif.com/sample-submission.

ChIP and Library Preparation by Active Motif

Chromatin immunoprecipitation was completed by Active Motif. Full methods are proprietary. Chromatin was isolated using a lysis buffer and membranes were disrupted with a dounce homogenizer. The lysates were then sonicated with Active Motif's EpiShear probe sonicator (#53051) and cooled sonication platform (#53080) to an average fragment length 300-500 bp. A portion of the sample was collected as the Input DNA, treated with RNase, proteinase K, and incubated to reverse crosslinking. The DNA was then collected by ethanol precipitation. The Input DNA was resuspended and concentration was quantified by a NanoDrop spectrophotometer. Extrapolation of this concentration to the original chromatin volume allowed for quantitation of the total chromatin yield. Aliquots of the fixed chromatin were used in the immunoprecipitation were precleared with protein A agarose beads (Invitrogen, #15918014). Genomic DNA regions of interest were isolated using specific ChIP antibodies (Online Table 2). Antibody DNA complexes were isolated using additional protein A agarose beads, and the crosslinked DNA, antibody, and bead complexes were washed. The cross-linked DNA was eluted from the beads with SDS buffer, and subjected to RNase and proteinase K treatment. Reverse crosslinking was done in an overnight incubation at 65°C, and ChIP DNA was purified with a phenol-chloroform extraction and ethanol precipitation.

Illumina sequencing libraries were prepared from the ChIP and Input DNAs using the standard consecutive enzymatic steps of end-polishing, dA-addition, and adaptor ligation using Active Motif's custom liquid handling robotics pipeline. Samplers were amplified with a 15 cycle PCR amplification and then quantified before being shipped to the Jefferson Cancer Genomics Laboratory at the Kimmel Cancer Center for sequencing.

MYCN ChIP-Seq: Kelly and NGP Cell Lines

Chromatin immunoprecipitation was performed on adherent cells as described in Bosse et al., 2017(Bosse et al. 2017). Of note, a different MYCN antibody was used than listed in Bosse

et al., 2017 (Santa Cruz B8.4B, sc-53993). Cells were grown as described in Cell Growth and Expansion section of protocol to 70-80% confluence in 150 mm tissue culture plates in 20 mL of media. To the existing media, 415 mL of 37% formaldehyde (final concentration of 0.75%) was added, and rocked for 10 minutes at RT. To this, 1.5 mL of 2.5 M glycine (Online Table 2) (final concentration of 0.18 M) was added to inactivate the formaldehyde, and the plate was rocked for an additional 5 min. Cells were lysed with a volume of FA Lysis Buffer (Online Table 2) equivalent to 5 pellet volumes. Beads were washed 3 times in ChIP Wash Buffer (Online Table 2) and one time with Final Wash Buffer (Online Table 2). Libraries were constructed using NEB Ultra Kit following manufacturer's instructions. Libraries were sequenced as single-end, 50 bp reads on a MiSeq to a depth of ~50 M reads by the Children's Hospital of Philadelphia Nucleic Acid and PCR Core.

ChIP Library Sequencing for ChIP

Sequencing was conducted by the Jefferson Cancer Genomics Laboratory at the Kimmel Cancer Center. Samples were quality control tested using an Agilent High Sensitivity Screen Tape to determine average fragment length. The concentration of each library was measured using a High Sensitivity Qubit Quantification kit, and samples were diluted to an appropriate amount for the loading protocol (4 nM or less). Samples were normalized to the same nanomolar concentration, and libraries were pooled together in equal amounts. Samples were diluted to 1.51 pM in Low EDTA TE Buffer. Samples were then sequenced as single-end, 70 bp reads to an average depth of ~30 M reads on a NextSeq 500.

ATAC-Seq Protocol

The following ATAC-Seq protocol was adapted from Buenrostro, et. al, 2015(Buenrostro et al. 2015). This protocol consists of four parts: Cell Preparation, Transposition Reaction and Purification, PCR Amplification, qPCR, and Library Preparation. Primer 1 and Primer 2 were

custom synthesized by Integrated DNA Technologies (IDT), using sequences provided in Buenrostro, et. al, 2015. Note: ATAC-Seq for NB-69 and NGP was performed using a slightly varied procedure and is noted in a separate section.

Cell Preparation

Cells were grown as described in Cell Growth and Expansion section of protocol to 70-80% confluence in a 75 mm^2 tissue culture flasks in 10 mL of media. Following detachment and pelleting, cells were resuspended in 1.0 mL of the appropriate growth media. Cells were triturated until they were in a homogenous single-cell suspension. Using an automated cell counter, the volume for 500,000 cells was determined and aliquoted into a sterile 1.5 mL Eppendorf tube containing 500 µL of sterile 1X PBS. Cells were centrifuged at 500 *x g* for 5 minutes at 4°C. The supernatant was carefully aspirated, and the cells were resuspended in 500 mL of sterile 1X PBS. Centrifugation was repeated and cells were resuspended in 500 mL of cold lysis buffer by gently pipetting up and down, and then immediately centrifuged at 500 *x g* for 10 minutes at 4°C. The supernatant was carefully removed and discarded. The pellet was immediately resuspend in 50 µL of nuclease free water by gently pipetting up and down, and the protocol immediately continued on to Transposition Reaction and Purification section.

Transposition Reaction and Purification

The pellet was placed on ice. The following reagents were prepared and combined: transposition reaction mix (25 µL TD (2X reaction buffer from Nextera Kit), 2.5 µL TDE1 (Nextera Tn5 Transposase from Nextera Kit), 17.5 µL nuclease-free water, and 5.0 µL of resuspended DNA/protein from the final step in *Cell Preparation* (resuspended pellet in 50 µL of nuclease free water). The transposition reaction was incubated in a thermocycler at 37°C for 30-35 minutes. The reaction was immediately purified using Qiagen MinElute PCR Purification Kit, and the transposed DNA was eluted in 10.5 µL of elution buffer (Buffer EB from the MinElute Kit consisting

of 10 mM Tris-Cl (pH 8)). The eppendorf tube containing purified DNA was parafilmed, and stored at -20°C. *NOTE:* This can act as a good stopping point, however these DNA fragments are not PCR amplifiable if melted at this point.

PCR Amplification

Primer sequences are shown in Online Table 3. To amplify the Transposed DNA, the following were combined into a 0.2 mL PCR tube: 10 µL transposed DNA, 10 µL nuclease-free H_2O, 2.5 µL 25 mM PCR Primer 1 (Ad1), 2.5 µL 25 mM Barcoded PCR Primer 2 (Ad2.X, X being the unique number of samples), and 25 µL NEBNext High-Fidelity 2X PCR Master Mix. The thermal cycle was as follows:

1 cycle:	5 min 72°C
	30 sec 98°C
5 cycles:	10 sec 98°C
	30 sec 63°C
	1 min 72°C

The five minute extension in the first cycle is critical to allow extension on both ends of the primer after transposition, thereby generating amplifiable fragments. This ensures that downstream quantitative PCR (qPCR) quantitation will not change the complexity of the original library.

qPCR

To reduce the GC and size bias in PCR, the appropriate number of PCR cycles (N) was determined using qPCR, allowing us to stop prior to saturation. The samples were kept in the thermocycler following the PCR Amplification reaction, and the qPCR side reaction was run. In a 0.2 mL PCR tube the following were added: 5 µL of DNA PCR amplified DNA, 2 µL of nuclease free H--$_2$O, 1 µL of 6.25 mM Custom Nextera PCR Primer (Ad1), 1 µL of 6.25 mM Custom Nextera

PCR Primer 2 (Ad2.X), 1 µL 9X SYBR Green I, and 5 µL NEBNext High-Fidelity 2X PCR Master Mix. This sample was run in the qPCR instrument with the following cycles:

1 cycle:	30 sec 98°C
20 cycles:	10 sec 98°C
	30 sec 63°C
	1 min 72°C

To calculate the additional number of cycles needed, a linear plot of Rn versus cycle was generated. This determined the cycle number (N) that corresponds to one-third of the maximum fluorescent intensity.

The remaining 45 mL PCR reaction was run to the cycle number (N) determined by qPCR. Cycles are as follows:

1 cycle:	30 sec 98°C
N cycles:	10 sec 98°C
	30 sec 63°C
	1 min 72°C

The amplified library was purified using Qiagen MinElute PCR Purification Kit after the additional PCR. The purified library was eluted in 20 µL of elution buffer (Buffer EB from the MinElute Kit consisting of 10 mM Tris-Cl (pH 8)). It is important to make sure that the column is dry prior to adding elution buffer to avoid ethanol contamination of final library. The amplified library was purified using AMPure XP beads at a 1.8x ratio to get rid of adapter dimers, using 80% ethanol for the wash steps. Sample was eluted in 50 µL of nuclease free H_2O. The concentration of the DNA eluted from the column should be about 30 nM.

Library Preparation

The quality of the purified libraries was assessed using a Bioanalyzer High-Sensitivity DNA Analysis kit (Agilent). If libraries contained predominant peaks around 1000 bp, SPRI beads were used to remove these fragments. This was accomplished by first, with a new vial of SPRI beads, performing size selection with various ratios to ensure larger peaks are removed. For example, ratios could include 0.4X, 0.45X, 0.5X. Choose the ratio that removes 1000 bp fragments, but leaves 800 bp fragments. Libraries were eluted in 20 µL of nuclease-free water, and sequenced as described below.

Sequencing for ATAC-Seq by Beijing Genomics International (BGI):

Sequencing was conducted by Beijing Genomics International at the Children's Hospital of Philadelphia. Samples were quality control tested using an Agilent High Sensitivity Screen Tape to confirm average fragment sizes were ~180, 380, 580, 780, and 980 bp. The concentration of each library was measured using a High Sensitivity Qubit Quantification kit, to ensure they were 5.5 nM. Samples were normalized and libraries were pooled together in equal amounts. Samples were then sequenced as paired-ends, 100 bp to an average depth of 80M reads on a HiSeq 2500.

ATAC-Seq NB-69 and NGP cell lines via Active Motif:

Cells were grown as described in Cell Growth and Expansion section of protocol to 70-80% confluence in a 75 mm^2 tissue culture flasks in 10 mL of media. Following detachment and pelleting, cells were resuspended in 1.0 mL of the appropriate growth media. Cells were triturated into a homogenous single-cell suspension. Using an automated cell counter, the volume for 100,000 cells was determined, and aliquoted into a sterile 1.5 mL eppendorf tube containing 500 µL of sterile 1X PBS. Cells were then centrifuged at 500 *x g* for 5 minutes at 4°C. The supernatant was carefully aspirated off, and the cells were resuspended in 500 µL of growth media with 5% DMSO. The sample was transferred to a 1.7 mL microfuge tube on ice. Cells were frozen with a

slow cooling to minimize cell lysis. Samples were shipped on dry ice to Active Motif (1914 Palomar Oaks Way, Ste 150, Carlsbad, CA 92008) following the instructions listed at on the Sample Submission Form, downloaded from www.activemotif.com/sample-submission. Samples were prepared and sequenced following Active Motif's ATAC-Seq proprietary protocol. Cells were thawed in a 37°C water bath, pelleted, washed with cold PBS, and tagmented as previously described (Buenrostro et al. 2015), with some modifications based on (Corces et al. 2017). Cell pellets were resuspended in lysis buffer, pelleted, and tagmented using the enzyme buffer provided in the Nextera Library Prep Kit (Illumina). Tagmented DNA was then purified using the MinElute PCR purification kit (Qiagen), amplified with 10 cycles of PCR, and purified using Agencourt AMPure SPRI beads (Beckman Coulter). The resulting material was quantified using the KAPA Library Quantification Kit for Illumina platforms (KAPA Biosystems) and sequenced with PE42 sequencing on the NextSeq 500 sequencer (Illumina).

ChIP-Seq Data Analysis

FASTQ quality was assessed using FastQC v0.11.4 (http://www.bioinformatics.babraham.ac.uk/projects/fastqc/) and sequences were adapter- and quality-trimmed using Trim Galore v.0.4.0 (https://github.com/FelixKrueger/TrimGalore) and CutAdapt v.1.12 (Krueger n.d.; Lindgreen 2012). Since multiple sequencers were used, FASTQ phred sequencing scores (Ewing et al. 1998) were calculated using a perl script (https://raw.githubusercontent.com/douglasgscofield/bioinfo/master/scripts/phredDetector.pl). This value was used as input into the alignment algorithm. The bwa v.0.7.12 samse (Li and Durbin 2009) was used to align the reads to hg19 reference genome and Picard tools v.2.17.9-SNAPSHOT (Alec Wysoker KT, Mike McCowan, Nils Homer, Tim Fennell, n.d.) was used to remove duplicates. Fragment sizes were estimated using MaSC 1.2.1 (Ramachandran et al. 2013) and these values were used as input into MACS2 v.2.1.1 (Zhang et al. 2008) for narrow peak calling (transcription factors) or broad peak calling (histone marks). Broad peaks were called

significant using a q-value (minimum False Discovery Rate) cut off of 0.05 and narrow peaks at a q-value cutoff of 0.1. Results were returned in units of signal per million reads to get normalized peak values. Finally, repetitive centromeric, telomeric and satellite regions known to have low sequencing confidence were removed using the blacklisted regions defined by UCSC.

ATAC-Seq Data Analysis

Samples were quality-controlled and trimmed as described in Chip-Seq Analysis. FASTQ files were aligned using bwa aln for BGI samples (100 bp reads) and bwa mem for Active Motif samples (42 bp reads). Reads with mapping quality < 10 were discarded. Biological duplicate BAMs were merged using Picard v.2.17.9-SNAPSHOT. Broad peaks were called using --extsize 200, --shift 100, and using a q-value (minimum False Discovery Rate) cut off of 0.1. Results were returned in units of signal per million reads to get normalized peak values. Finally, repetitive centromeric, telomeric and satellite regions known to have low sequencing confidence were removed using the blacklisted regions defined by UCSC.

ChIP-Seq Quality Control Metrics

We investigated three metrics to assess ChIP-seq quality. To calculate enrichment of reads within peaks we determined the FRiP score using deeptools2 (Ramírez et al. 2016). The FRiP score is defined as the fraction of reads that fall within a peak divided by the total number of reads. To measure read enrichment independent of peak calling we calculated the NSC (normalized strand cross-correlation) and the RSC (relative strand cross-correlation) using phantompeakqualtools (Landt et al. 2012; Kharchenko, Tolstorukov, and Park 2008)as part of the ENCODE ChIP-seq processing pipeline. All ChIP-Seq data passed quality control and results are reported in Online Table 4.

Super-enhancer Calling and Comparison

Super-enhancers (SEs) were called from H3K27Ac BAM files using the default parameters of LILY (https://github.com/BoevaLab/LILY), which includes correction for copy number variation inherently present in cancer samples. Enhancers were classified into SEs, enhancers, and promoters and annotated using Homer v4.10.4. Scripts to run LILY can be found on Github (https://github.com/marislab/epigenomics-data-descriptor). SEs were also called from H3K27Ac MACS2 peaks using ROSE v.0.1 (https://bitbucket.org/young_computation/rose/src/master/) using default parameters and annotated using Homer v4.10.4. SEs which overlapped with the MYCN locus (hg19, chr2:16080683-16087129) were removed from the analysis. SE genes which we annotated as transcription factors (Lambert et al. 2018) were used for comparison to two literature studies (Boeva et al. 2017; van Groningen et al. 2017).

Heatmap Preparation

The 5,000 most significant (sorted by highest $-\log_{10}$(p-value) and $-\log_{10}$(q-value)) MYCN peaks for each of the five MYCN amplified cell line were intersected using bedtools. Heatmaps were generated for regions +/- 4 kb from the transcription start site (TSS) for the 5,046 peaks common to at least four MYCN amplified cell lines. Heatmaps were created for LA-N-5 and NB-69 at loci annotated as enhancers, SEs, and promoters-TSS by LILY. All ChIP-seq heatmaps were created using deepTools 3.2.0 package plotHeatmap tool (Ramírez et al. 2016). The code and parameters used to generate heatmaps can be found on GitHub (https://github.com/marislab/epigenomics-data-descriptor).

Cell Line Authentication

All cell lines were STR-authenticated by Guardian Forensic Sciences (Abington, PA) using the GenePrint 24 (Promega, #B1870).

Data Records

Raw, concatenated FASTQ files and processed BIGWIG files for all sequencing data were deposited into the Gene Expression Omnibus (GEO) under SuperSeries Accession Number GSE138315. MYCN and MYC ChIP-Seq data for the Kelly and NGP cell lines were deposited into GEO under Accession Number GSE94782, all other MYCN and MYC ChIP-Seq were deposited under Accession Number GSE138295, histone ChIP-Seq data were deposited under Accession Number GSE138314, and ATAC-Seq data were deposited under Accession Number GSE138293.

Technical Validation

Prior to selecting cell lines for MYCN and MYC profiling, we assessed RNA expression (Figure 2A-B) and protein expression (Figure 2C-D) across a subset of neuroblastoma cell lines. NB-LS, while *MYCN* non-amplified, has substantial MYCN RNA and protein expression (Cohn et al. 1990), but was not chosen, as we restricted MYCN ChIP-Seq to *MYCN* amplified cell lines plus one negative control. SK-N-BE(2)-C, a *MYCN* amplified cell line, showed high *MYCN* mRNA expression, but surprisingly low protein expression, and thus was excluded. The remaining cell lines had concordant *MYCN* and *MYC* mRNA and protein expression, thus, COG-N-415, KELLY, NB-1643, LA-N-5, and NGP were chosen for MYCN ChIP-Seq while NB-69, SK-N-AS, and SK-N-SH were chosen for MYC ChIP-Seq. As additional controls, we performed MYCN ChIP-Seq in the *MYCN* non-amplified line NB-69, and MYC ChIP-Seq on the *MYCN* amplified cell line KELLY. To validate the MYCN and MYC ChIP-Seq antibodies, we first intersected loci bound by MYCN in two or more cell lines and of the 157 MYCN transcriptional targets previously reported using ChIP-on-ChIP (Valentijn et al. 2012), found 139 loci occupied by the MYCN via ChIP-Seq (Figure 2E). Next, we integrated the top 5,000 MYCN peaks from each *MYCN* amplified cell line. We generated heatmaps for the peaks (1,330) which overlapped in all five cell lines (as defined in ***Heatmap Preparation***) and depict occupancy of MYCN (Figure 2F) and MYC (Figure 2G) at

these sites. As expected, the MYCN amplified cell lines COG-N-415, KELLY, NB-1643, LA-N-5, and NGP show similar binding profiles, while the negative control MYCN non-amplified line NB-69 depicted an absence of binding for MYCN at the same loci. We observed MYC bound to the same loci in the MYCN non-amplified cell lines, SK-N-AS, SK-N-SH, and NB-69 as well as the *MYCN* amplified and low MYC-expressing line KELLY (Figure 2G), supporting the notion of redundant functionality of MYC family protein members.

Next, we evaluated genome-wide binding densities of the histone antibodies and assessed open chromatin by plotting binding of one *MYCN* amplified cell line LA-N-5 (Figure 3A), and one *MYCN* non-amplified cell line NB-69 (Figure 3B). Of note, cell-line specific promoters are located in regions of open chromatin and strongly occupied by narrow regions of H3K4me3 and devoid of H3K27me3 and H3K4me1, as expected. The majority of promoters are also occupied by MYCN in LA-N-5 and MYC in NB-69. Enhancers have bivalent marking of MYCN, H3K4me3, H3K27Ac, open chromatin, and absence of H3K27me3. SEs are broadly marked by MYCN, H3K4me3, H3K27Ac, H3K4me1, and open chromatin.

Finally, we used our H3K27Ac ChIP-Seq data to compare SEs to those reported in two other publications describing the SE landscape in neuroblastoma (Figure 4). Boeva and colleagues identified 5,205 SE-associated genes in Table S3 to identify core regulatory transcriptional circuitry in neuroblastoma using 25 cell lines (Boeva et al. 2017). Four cell lines were common to our study: SK-N-BE(2)C, SK-N-FI, SK-N-AS, and NB-69. Therefore, to validate our H3K27Ac ChIP-Seq, we utilized the same algorithm (LILY, see **Methods**) to call SEs from our H3K27Ac data, and restricted comparison analyses to genes defined as transcription factors (TFs), as defined by core regulatory circuitry (Lambert et al. 2018). We annotated 396 of the SEs reported by Boeva and colleagues as transcription factors and found 59-85% concordance of our TF SE calls (Figure 4A). While a majority of SEs called in each of our cell lines was concordant with Boeva and colleagues, the variance likely stems from the diversity of cell lines in both studies. Thus, we additionally compared our TF SE calls to those from an independent neuroblastoma

study (van Groningen et al. 2017) which used the ROSE algorithm (see **Methods**) and reported smaller SE genesets (found in Figure 3) driving the lineage-specific mesenchymal (MES, N = 20 TFs) and adrenergic (ARDN, N = 18 TFs) subtypes. To mimic the analysis performed by van Gronigen and colleagues, we ran ROSE on our H3K27Ac ChIP-Seq data and removed any peaks which overlapped the *MYCN* locus (see **Methods**) to account for false SE calls due to *MYCN* amplification. There were no common neuroblastoma cell lines between van Gronigen and colleagues study and the lines used in our study. We assessed the number of MES or ADRN SE-associated TFs detected in each of our study and found between five and eight ADRN SEs were detected using ROSE (Figure 4B) and between five and 11 ADRN SEs were detected using LILY (Figure 4C). SK-N-SH has a known MES subtype; its subclone, SH-SY-5Y, was profiled as MES by van Gronigen and colleagues. Combining the calls, we were able to significantly (Fisher's exact test, $p < 0.05$) validate ADRN subtypes in eight of the ten cell lines we profiled (Figure 4D). Interestingly, SK-N-AS contains SEs from both subtypes and thus may reflect a heterogeneous cell line. Specific SEs are reported per algorithm per cell line in Online Table 5. As further validation, we re-analyzed publicly-available SK-N-SH H3K27Ac (Biosample SAMN05733860, Run SRR5338927) and SK-N-SH Input (Biosample SAMN05733844, Run SRR5471111) ChIP-Seq data (GEO accession GSM2534162) using the same peak-calling and SE pipelines used on our data (see **Methods**). We observed enhancer binding (H3K27Ac) and open chromatin (ATAC) at the same loci we observe strong MYC occupancy (Supplemental Figure 1A). Further, we assessed concordance of SEs called in SK-N-SH with those previously reported and found 76% of TF SEs called in SK-N-SH in common with those from Boeva, et. al, similar to our findings.

Together, we have validated both MYCN and MYC ChIP-Seq antibodies for use in ChIP-Seq, as well as genome-wide occupancy profiles for histone markers and open chromatin across a cohort of neuroblastoma cell lines. We ran two algorithms (LILY and ROSE) and compared our data to two independent datasets to validate reproducibility of lineage-specific SEs in neuroblastoma cell lines. Finally, we demonstrate integration of publicly-available H3K27Ac data

from SK-N-SH with our MYC ChIP-Seq and ATAC-Seq data, and show reproducibility of SE calls between the publicly-available data and two independent reports. These data should be a valuable resource to the childhood cancer and MYC research communities.

Usage Notes

Here, we provide a comprehensive, validated ChIP-Seq (MYCN, MYC, H3K27Ac, H3K27me3, H3K4me3, and H3K4me1) and ATAC-Seq neuroblastoma cell line dataset which can be coupled with our previous RNA-Seq profiling dataset (Harenza et al. 2017) to interrogate novel transcriptional regulation in this disease. For example, the H3K27me3 ChIP-Seq can be used to identify genes being repressed via the PRC2 complex, while H3K27Ac and H3K4me1 ChIP-Seq can be used to interrogate promoter-enhancer mechanisms. CSI-ANN can be used to integrate histone ChIP-Seq data to predict regulatory DNA segments (Firpi, Ucar, and Tan 2010), and IM-PET can use the results from CSI-ANN to predict enhancer-promoter interactions without the need for Hi-C data (He et al. 2013). Additionally, chromatin states can be inferred (Ernst and Kellis 2012; Sohn et al. 2015), and these data can be later integrated with whole exome or genome sequencing data or genome-wide association studies to identify molecular alterations driving transcriptional regulatory marked by histone marks or open chromatin.

All data are openly-available from GEO as described in the Data Records section.

Code Availability

Code for SE calling, filtering, and heatmap generation is available on Github (https://github.com/marislab/epigenomics-data-descriptor).

Acknowledgements

Funding for this research was provided by NIH grants R01 CA180692 (JMM), R35 CA220500 (JMM), and an Alex's Lemonade Stand Young Investigator Award (JLR). We thank the neuroblastoma patients and families for donating tumor tissue from which cell lines used in this study were derived.

Author contributions

Conceptualization: JLR, KU

Methodology: JLR, KU, KP, RTS, RNA, PF, GPW

Validation: JLR, KU, KP, KLC, AM

Formal Analysis: JLR, KU, KP, AM

Investigation: JLR, KU, KP

Resources: JMM

Data Curation: JLR, KP, GIS

Writing - Original Draft: JLR, KU, KP, AM

Writing - Review & Editing: JLR, KLC, KU, GPW, JMM

Visualization: JLR, KU, KP

Supervision: JLR

Funding Acquisition: JLR, JMM

Competing interests

The authors declare no competing interests.

Figure Legends

Figure 1. Experimental workflow. (**A**) Cells were thawed, grown, and expanded until 70-80% confluency in a 150 mm dish. (**B**) For ChIP-Seq, cells were fixed, collected, and frozen (N = 1 biological replicate per cell line). Libraries were prepared, sequenced, and data analyzed. (**C**) For ATAC-Seq (n = 14 samples with n = 2 biological replicates), cells were incubated in a transposition reaction, DNA was purified, and amplified with limited PCR. Libraries were prepared, sequenced, and analyzed. Diagram was created using Servier Medical ART (https://smart.servier.com/).

Figure 2. Comparison of MYCN and MYC binding based on MYCN amplification status. The $\log_2$ FPKM mRNA expression of *MYCN* (**A**) and *MYC* (**B**) in all neuroblastoma cell lines assayed herein. Protein expression for MYCN (**C**) and MYC (**D**) in a subset of cell lines. (**E**) Comparison of unique MYCN peaks called across all MYCN ChIP-Seq samples to known MYCN regulated genes (Valentijn et al. 2012) demonstrates concordance of MYCN ChIP-Seq to MYCN ChIP-ChIP. (**F**) The top 1,330 MYCN peaks ($p < 0.05$, $q < 0.05$) are plotted as ChIP-Seq heatmaps for five MYCN amplified cell lines (COG-N-415, Kelly, NB-1643, LAN-5, and NGP) and one MYCN non-amplified cell line (NB-69). All heat map densities ranges from +/- 4.0 Kb from the TSS, with average signal plots shown above. (**G**) MYC ChIP-Seq heat maps for the same peaks in two MYCN non-amplified lines (SK-N-AS, and NB-69) and one MYCN amplified cell line (Kelly) showing redundant binding of MYC in non-amplified cell lines.

Figure 3. Validation of ChIP-Seq promoter, enhancer, and open chromatin occupancy. (**A**) Binding densities of MYCN, MYC, histone antibodies, and open chromatin for promoter regions of the MYCN-amplified cell line LA-N-5 (**A**) and the MYCN non-amplified cell line NB-69 (**B**) are depicted (+/- 4.0 Kb from gene TSS) and distinct profiles are shown for promoters (+/- 4.0 Kb from gene TSS), enhancers (+/- 10.0 Kb from gene TSS), and SEs (+/- 10.0 Kb from gene TSS). For LA-N-5: $N_{promoters\text{-}TSS} = 4{,}662$, $N_{enhancers} = 25{,}601$, $N_{SE} = 826$, and for NB-69: $N_{promoters\text{-}TSS} = 4{,}718$, $N_{enhancers} = 31{,}769$, $N_{SE} = 667$.

Figure 4. Comparison of super-enhancers called to literature reports. SEs were called using the H3K27Ac ChIP-Seq data collected in ten neuroblastoma cell lines plus SK-N-SH H3K27Ac ChIP-Seq data from ENCODE. Results were compared to two other literature sources which examined SE signatures in neuroblastoma. (**A**) LILY TF-specific SEs were 59-85% concordant with (Boeva et al. 2017). (**B**) Total number of lineage-specific SEs called from ROSE and (**C**) LILY per cell line (ADRN = adrenergic, MES = mesenchymal subtype). (**D**) Total number of lineage-specific SEs called when combining ROSE and LILY results. One-tailed fisher's exact test p-values are listed (*denotes significance at an α level of 0.05).

Table Legends

Online Table 1. Neuroblastoma cell lines profiled in this study.

Identification of which cell lines were used to perform ChIP-Seq and ATAC-Seq, along with their MYCN amplification status.

Online Table 2. Materials and reagents used in this study.

Information regarding the resources used for obtaining the ChIP-Seq and ATAC-Seq data. This includes manufacturer, item number or identification code, and the volume or concentration used.

Online Table 3. ATAC-Seq primer sequences.

Primer sequences used in ATAC-Seq to amplify transposed DNA.

Online Table 4. ChIP-Seq Quality Control Metrics.

Total number of reads, FRiP score, RSC, and NSC per sample.

Online Table 5. Super-enhancer-associated transcription factor comparisons.

Manuscript metadata tables

Metadata Table 1. Neuroblastoma cell line metadata.

Listed are the cell lines used in this study, MYCN amplification status, and culturing media information.

References

ACCO. 2016. "Childhood Cancer Statistics." February 16, 2016. https://www.acco.org/childhood-cancer-statistics/.

Alec Wysoker KT, Mike McCowan, Nils Homer, Tim Fennell. n.d. "A Set of Tools (in Java) for Working with next Generation Sequencing Data in the BAM." https://broadinstitute.github.io/picard/.

Ambros, P. F., I. M. Ambros, G. M. Brodeur, M. Haber, J. Khan, A. Nakagawara, G. Schleiermacher, et al. 2009. "International Consensus for Neuroblastoma Molecular Diagnostics: Report from the International Neuroblastoma Risk Group (INRG) Biology Committee." *British Journal of Cancer* 100 (9): 1471–82.

Boeva, Valentina, Caroline Louis-Brennetot, Agathe Peltier, Simon Durand, Cécile Pierre-Eugène, Virginie Raynal, Heather C. Etchevers, et al. 2017. "Heterogeneity of Neuroblastoma Cell Identity Defined by Transcriptional Circuitries." *Nature Genetics* 49 (9): 1408–13.

Bosse, Kristopher R., Pichai Raman, Zhongyu Zhu, Maria Lane, Daniel Martinez, Sabine Heitzeneder, Komal S. Rathi, et al. 2017. "Identification of GPC2 as an Oncoprotein and Candidate Immunotherapeutic Target in High-Risk Neuroblastoma." *Cancer Cell* 32 (3): 295–309.e12.

Buenrostro, Jason D., Beijing Wu, Howard Y. Chang, and William J. Greenleaf. 2015. "ATAC-Seq: A Method for Assaying Chromatin Accessibility Genome-Wide." *Current Protocols in Molecular Biology / Edited by Frederick M. Ausubel ... [et Al.]* 109 (January): 21.29.1–9.

Chappell, James, and Stephen Dalton. 2013. "Roles for MYC in the Establishment and Maintenance of Pluripotency." *Cold Spring Harbor Perspectives in Medicine* 3 (12): a014381.

Cohn, S. L., H. Salwen, M. W. Quasney, N. Ikegaki, J. M. Cowan, C. V. Herst, R. H. Kennett, S. T. Rosen, J. A. DiGiuseppe, and G. M. Brodeur. 1990. "Prolonged N-Myc Protein Half-Life in a Neuroblastoma Cell Line Lacking N-Myc Amplification." *Oncogene* 5 (12): 1821–27.

Coller, H. A., C. Grandori, P. Tamayo, T. Colbert, E. S. Lander, R. N. Eisenman, and T. R. Golub. 2000. "Expression Analysis with Oligonucleotide Microarrays Reveals That MYC Regulates Genes Involved in Growth, Cell Cycle, Signaling, and Adhesion." *Proceedings of the National Academy of Sciences of the United States of America* 97 (7): 3260–65.

Corces, M. Ryan, Alexandro E. Trevino, Emily G. Hamilton, Peyton G. Greenside, Nicholas A. Sinnott-Armstrong, Sam Vesuna, Ansuman T. Satpathy, et al. 2017. "An Improved ATAC-Seq Protocol Reduces Background and Enables Interrogation of Frozen Tissues." *Nature Methods* 14 (10): 959–62.

Ernst, Jason, and Manolis Kellis. 2012. "ChromHMM: Automating Chromatin-State Discovery and Characterization." *Nature Methods* 9 (3): 215–16.

Ewing, B., L. Hillier, M. C. Wendl, and P. Green. 1998. "Base-Calling of Automated Sequencer Traces Using Phred. I. Accuracy Assessment." *Genome Research* 8 (3): 175–85.

Firpi, Hiram A., Duygu Ucar, and Kai Tan. 2010. "Discover Regulatory DNA Elements Using Chromatin Signatures and Artificial Neural Network." *Bioinformatics* 26 (13): 1579–86.

Gherardi, Samuele, Emanuele Valli, Daniela Erriquez, and Giovanni Perini. 2013. "MYCN-Mediated Transcriptional Repression in Neuroblastoma: The Other Side of the Coin." *Frontiers in Oncology* 3 (March): 42.

Groningen, Tim van, Jan Koster, Linda J. Valentijn, Danny A. Zwijnenburg, Nurdan Akogul, Nancy E. Hasselt, Marloes Broekmans, et al. 2017. “Neuroblastoma Is Composed of Two Super-Enhancer-Associated Differentiation States.” *Nature Genetics* 49 (8): 1261–66.

Harenza, Jo Lynne, Maura A. Diamond, Rebecca N. Adams, Michael M. Song, Heather L. Davidson, Lori S. Hart, Maiah H. Dent, Paolo Fortina, C. Patrick Reynolds, and John M. Maris. 2017. “Transcriptomic Profiling of 39 Commonly-Used Neuroblastoma Cell Lines.” *Scientific Data* 4 (March): 170033.

He, Stanley, Zhihui Liu, Doo-Yi Oh, and Carol J. Thiele. 2013. “MYCN and the Epigenome.” *Frontiers in Oncology* 3 (January): 1.

Ji, Hongkai, George Wu, Xiangcan Zhan, Alexandra Nolan, Cheryl Koh, Angelo De Marzo, Hoang Mai Doan, et al. 2011. “Cell-Type Independent MYC Target Genes Reveal a Primordial Signature Involved in Biomass Accumulation.” *PloS One* 6 (10): e26057.

Kharchenko, Peter V., Michael Y. Tolstorukov, and Peter J. Park. 2008. “Design and Analysis of ChIP-Seq Experiments for DNA-Binding Proteins.” *Nature Biotechnology* 26 (12): 1351–59.

Krueger, Felix. n.d. *TrimGalore*. Github. Accessed October 7, 2019. https://github.com/FelixKrueger/TrimGalore.

Lambert, Samuel A., Arttu Jolma, Laura F. Campitelli, Pratyush K. Das, Yimeng Yin, Mihai Albu, Xiaoting Chen, Jussi Taipale, Timothy R. Hughes, and Matthew T. Weirauch. 2018. “The Human Transcription Factors.” *Cell* 172 (4): 650–65.

Landt, S. G., G. K. Marinov, A. Kundaje, P. Kheradpour, F. Pauli, S. Batzoglou, B. E. Bernstein, et al. 2012. “ChIP-Seq Guidelines and Practices of the ENCODE and modENCODE Consortia.” *Genome Research* 22 (9): 1813–31.

Li, Heng, and Richard Durbin. 2009. “Fast and Accurate Short Read Alignment with Burrows-Wheeler Transform.” *Bioinformatics* 25 (14): 1754–60.

Lindgreen, Stinus. 2012. “AdapterRemoval: Easy Cleaning of next-Generation Sequencing Reads.” *BMC Research Notes* 5 (July): 337.

Malynn, B. A., I. M. de Alboran, R. C. O’Hagan, R. Bronson, L. Davidson, R. A. DePinho, and F. W. Alt. 2000. “N-Myc Can Functionally Replace c-Myc in Murine Development, Cellular Growth, and Differentiation.” *Genes & Development* 14 (11): 1390–99.

Maris, John M. 2010. “Recent Advances in Neuroblastoma.” *The New England Journal of Medicine* 362 (23): 2202–11.

Matthay, Katherine K., John M. Maris, Gudrun Schleiermacher, Akira Nakagawara, Crystal L. Mackall, Lisa Diller, and William A. Weiss. 2016. “Neuroblastoma.” *Nature Reviews. Disease Primers* 2 (November): 16078.

Pinto, Navin R., Mark A. Applebaum, Samuel L. Volchenboum, Katherine K. Matthay, Wendy B. London, Peter F. Ambros, Akira Nakagawara, et al. 2015. “Advances in Risk Classification and Treatment Strategies for Neuroblastoma.” *Journal of Clinical Oncology: Official Journal of the American Society of Clinical Oncology* 33 (27): 3008–17.

Pistoia, Vito, Fabio Morandi, Annalisa Pezzolo, Lizzia Raffaghello, and Ignazia Prigione. 2012. “MYCN: From Oncoprotein to Tumor-Associated Antigen.” *Frontiers in Oncology* 2 (November): 174.

Ramachandran, Parameswaran, Gareth A. Palidwor, Christopher J. Porter, and Theodore J. Perkins. 2013. “MaSC: Mappability-Sensitive Cross-Correlation for Estimating Mean Fragment Length of Single-End Short-Read Sequencing Data.” *Bioinformatics* 29 (4): 444–50.

Ramírez, Fidel, Devon P. Ryan, Björn Grüning, Vivek Bhardwaj, Fabian Kilpert, Andreas S. Richter, Steffen Heyne, Friederike Dündar, and Thomas Manke. 2016. “deepTools2: A next Generation Web Server for Deep-Sequencing Data Analysis.” *Nucleic Acids Research* 44 (W1): W160–65.

Rickman, David S., Johannes H. Schulte, and Martin Eilers. 2018. “The Expanding World of N-MYC-Driven Tumors.” *Cancer Discovery* 8 (2): 150–63.

Simon, Thorsten, Frank Berthold, Arndt Borkhardt, Bernhard Kremens, Boris De Carolis, and Barbara Hero. 2011. “Treatment and Outcomes of Patients with Relapsed, High-Risk Neuroblastoma: Results of German Trials.” *Pediatric Blood & Cancer* 56 (4): 578–83.

Sohn, Kyung-Ah, Joshua W. K. Ho, Djordje Djordjevic, Hyun-Hwan Jeong, Peter J. Park, and Ju Han Kim. 2015. “hiHMM: Bayesian Non-Parametric Joint Inference of Chromatin State Maps.” *Bioinformatics* 31 (13): 2066–74.

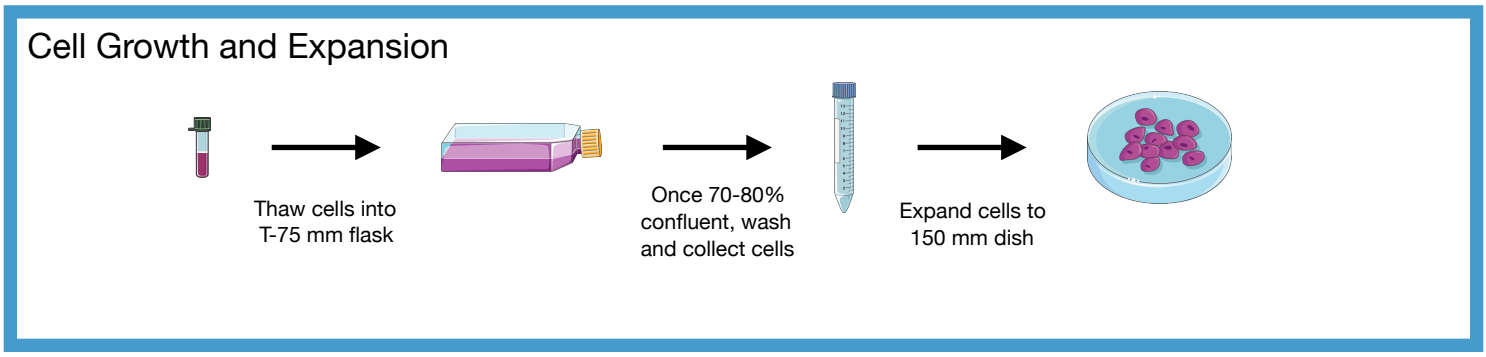
Twist, Clare J., Mary Lou Schmidt, Arlene Naranjo, Wendy B. London, Sheena C. Tenney, Araz Marachelian, Hiroyuki Shimada, et al. 2019. “Maintaining Outstanding Outcomes Using Response- and Biology-Based Therapy for Intermediate-Risk Neuroblastoma: A Report From the Children’s Oncology Group Study ANBL0531.” *Journal of Clinical Oncology: Official Journal of the American Society of Clinical Oncology*, August, JCO1900919.

Valentijn, Linda J., Jan Koster, Franciska Haneveld, Rachida Ait Aissa, Peter van Sluis, Marloes E. C. Broekmans, Jan J. Molenaar, Johan van Nes, and Rogier Versteeg. 2012. “Functional MYCN Signature Predicts Outcome of Neuroblastoma Irrespective of MYCN Amplification.” *Proceedings of the National Academy of Sciences of the United States of America* 109 (47): 19190–95.

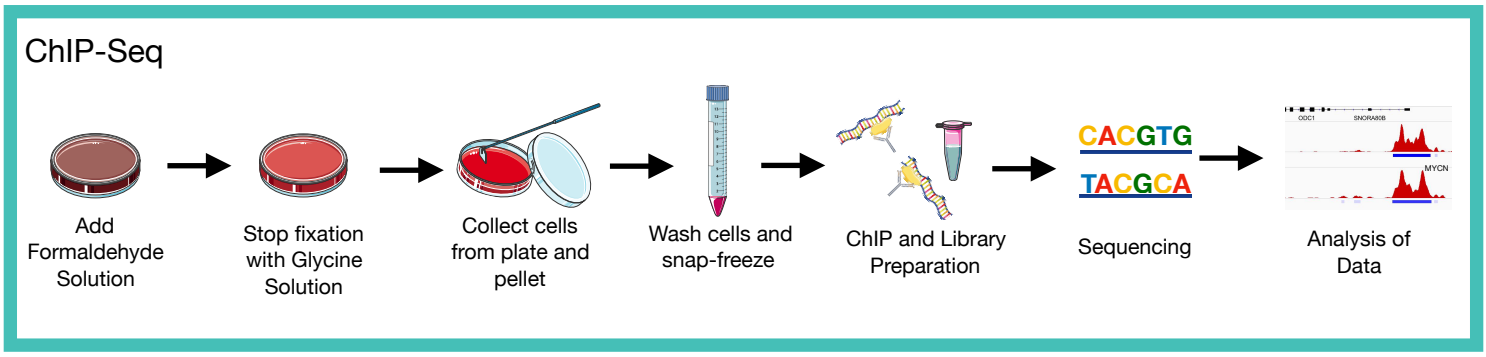
Zhang, Yong, Tao Liu, Clifford A. Meyer, Jérôme Eeckhoute, David S. Johnson, Bradley E. Bernstein, Chad Nusbaum, et al. 2008. “Model-Based Analysis of ChIP-Seq (MACS).” *Genome Biology* 9 (9): R137.

Figure 1

a


Cell Growth and Expansion
Thaw cells into T-75 mm flask
Once 70-80% confluent, wash and collect cells
Expand cells to 150 mm dish

b


ChIP-Seq
Add Formaldehyde Solution
Stop fixation with Glycine Solution
Collect cells from plate and pellet
Wash cells and snap-freeze
ChIP and Library Preparation
CACGTG
TACGCA
Sequencing
Analysis of Data

c

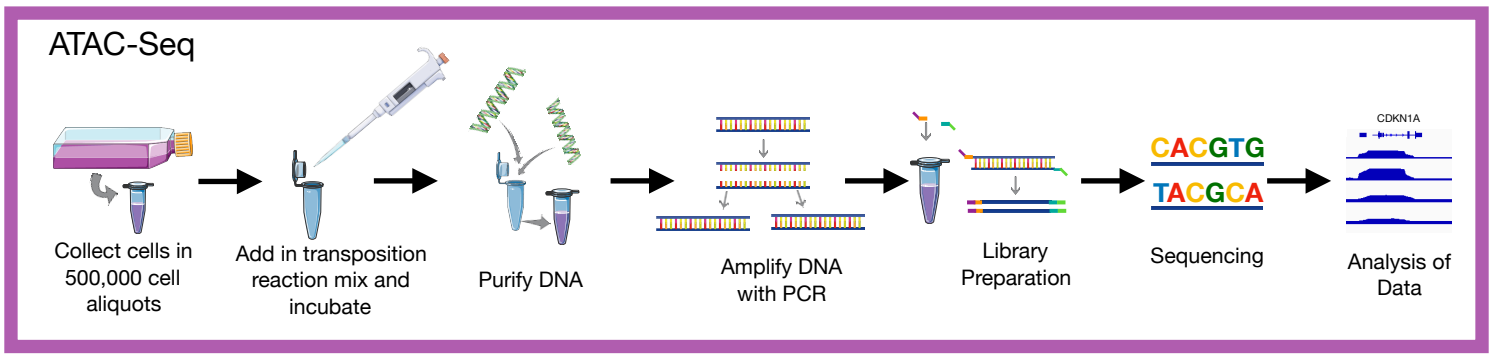

ATAC-Seq
Collect cells in 500,000 cell aliquots
Add in transposition reaction mix and incubate
Purify DNA
Amplify DNA with PCR
Library Preparation
CACGTG
TACGCA
Sequencing
CDKN1A
Analysis of Data

Figure 2

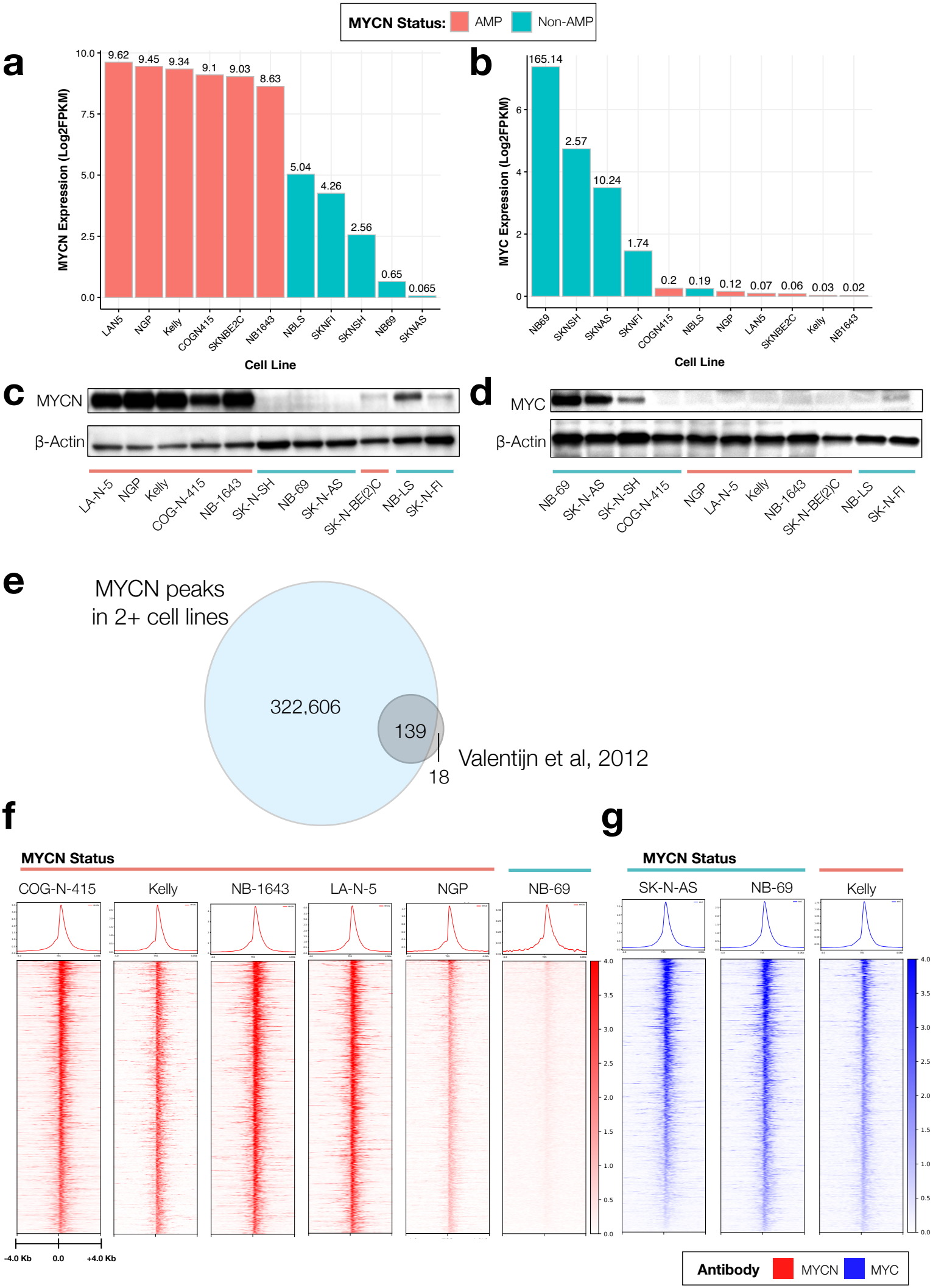
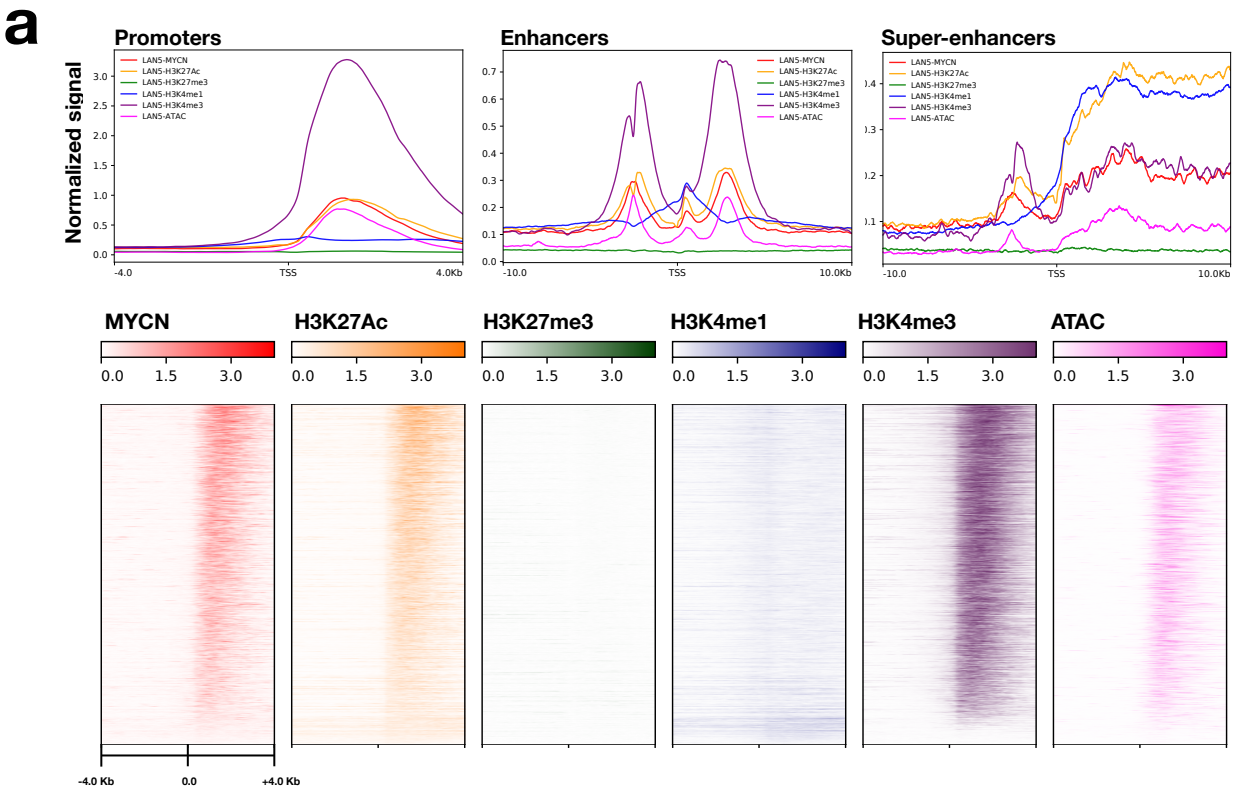

MYCN Status: AMP Non-AMP
a
MYCN Expression (Log2FPKM)
10.0
7.5
5.0
2.5
0.0
9.62
9.45
9.34
9.1
9.03
8.63
5.04
4.26
2.56
0.65
0.065
LAN5
NGP
Kelly
COGN415
SKNBE2C
NB1643
NBLS
SKNFI
SKNSH
NB69
SKNAS
Cell Line
b
MYC Expression (Log2FPKM)
6
4
2
0
165.14
2.57
10.24
1.74
0.2
0.19
0.12
0.07
0.06
0.03
0.02
NB69
SKNSH
SKNAS
SKNFI
COGN415
NBLS
NGP
LAN5
SKNBE2C
Kelly
NB1643
Cell Line
c
MYCN
β-Actin
LA-N-5
NGP
Kelly
COG-N-415
NB-1643
SK-N-SH
NB-69
SK-N-AS
SK-N-BE(2)C
NB-LS
SK-N-FI
d
MYC
β-Actin
NB-69
SK-N-AS
SK-N-SH
COG-N-415
NGP
LA-N-5
Kelly
NB-1643
SK-N-BE(2)C
NB-LS
SK-N-FI
e
MYCN peaks
in 2+ cell lines
322,606
139
18
Valentijn et al, 2012
f
MYCN Status
COG-N-415
Kelly
NB-1643
LA-N-5
NGP
NB-69
-4.0 Kb
0.0
+4.0 Kb
g
MYCN Status
SK-N-AS
NB-69
Kelly
Antibody
MYCN
MYC

Figure 3


a
Promoters
Enhancers
Super-enhancers
Normalized signal
LAN5-MYCN
LAN5-H3K27Ac
LAN5-H3K27me3
LAN5-H3K4me1
LAN5-H3K4me3
LAN5-ATAC
-4.0
TSS
4.0Kb
-10.0
10.0Kb
MYCN
H3K27Ac
H3K27me3
H3K4me1
H3K4me3
ATAC
0.0
1.5
3.0
-4.0 Kb
+4.0 Kb

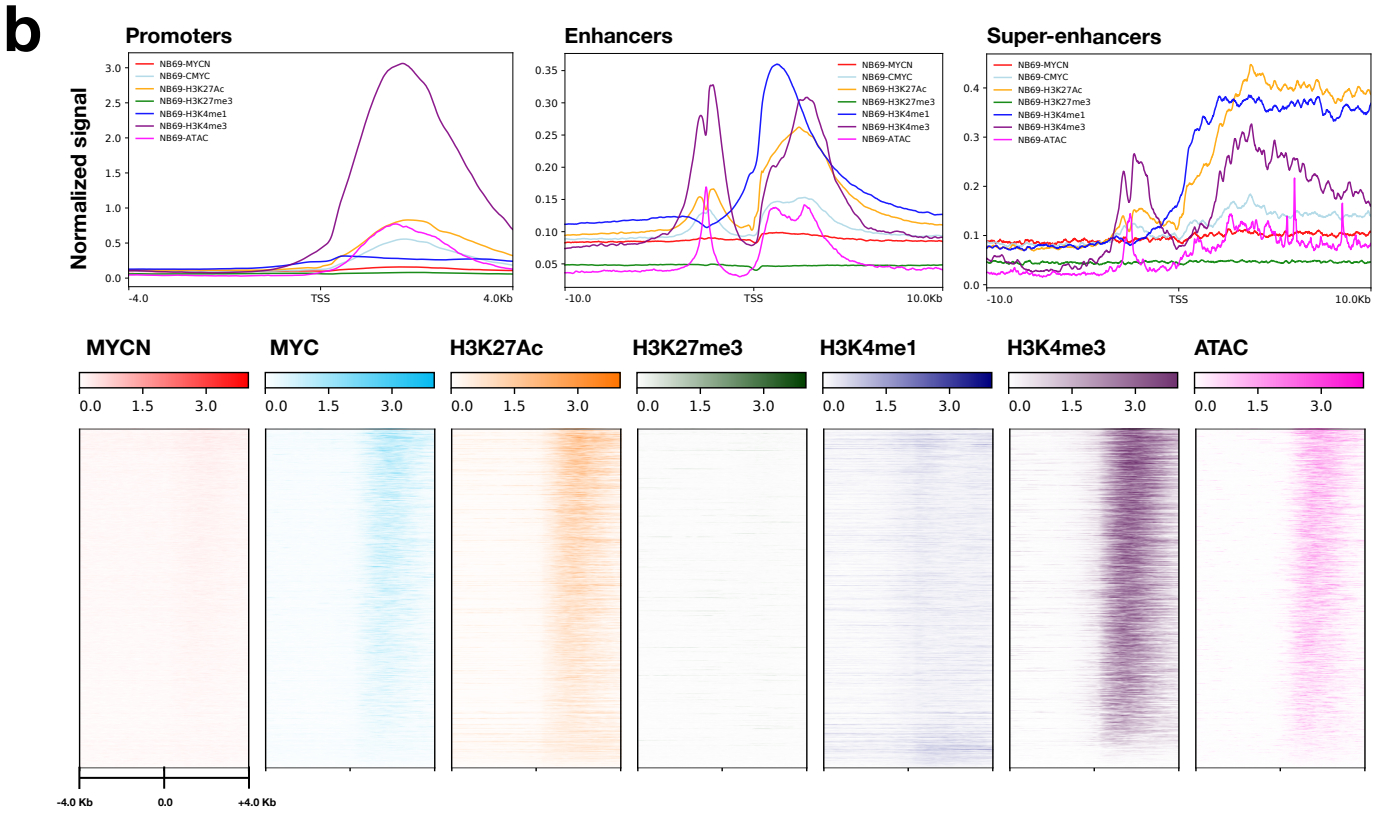
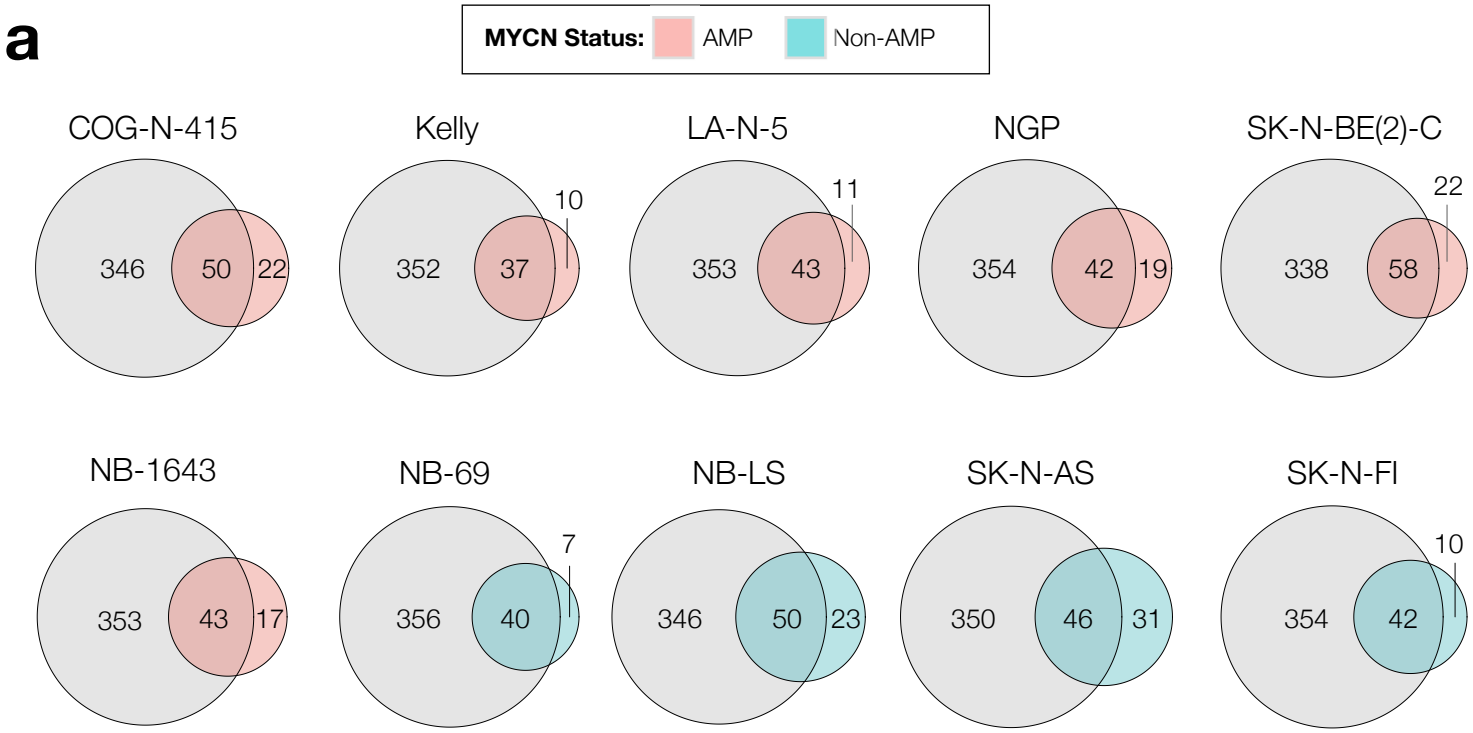
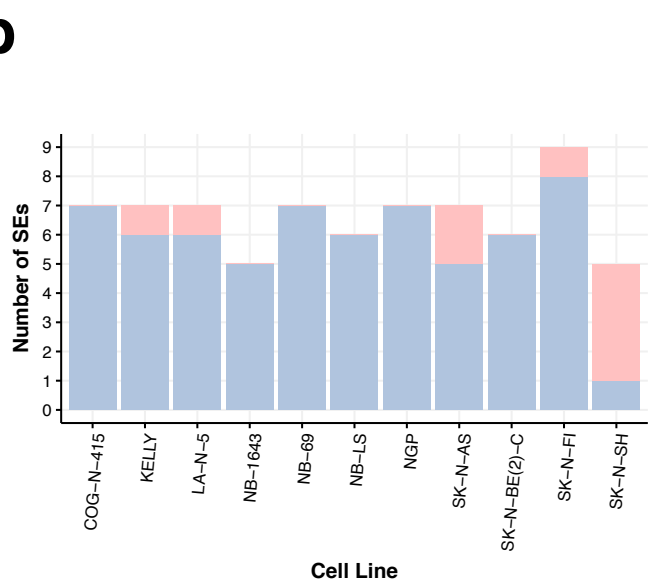
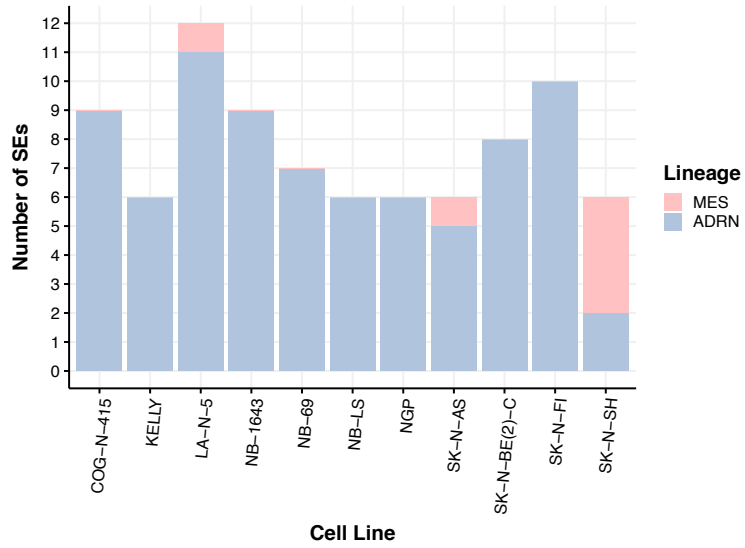

b
Promoters
Enhancers
Super-enhancers
Normalized signal
NB69-MYCN
NB69-CMYC
NB69-H3K27Ac
NB69-H3K27me3
NB69-H3K4me1
NB69-H3K4me3
NB69-ATAC
-4.0
TSS
4.0Kb
-10.0
10.0Kb
MYCN
MYC
H3K27Ac
H3K27me3
H3K4me1
H3K4me3
ATAC
0.0
1.5
3.0
-4.0 Kb
+4.0 Kb

Figure 4



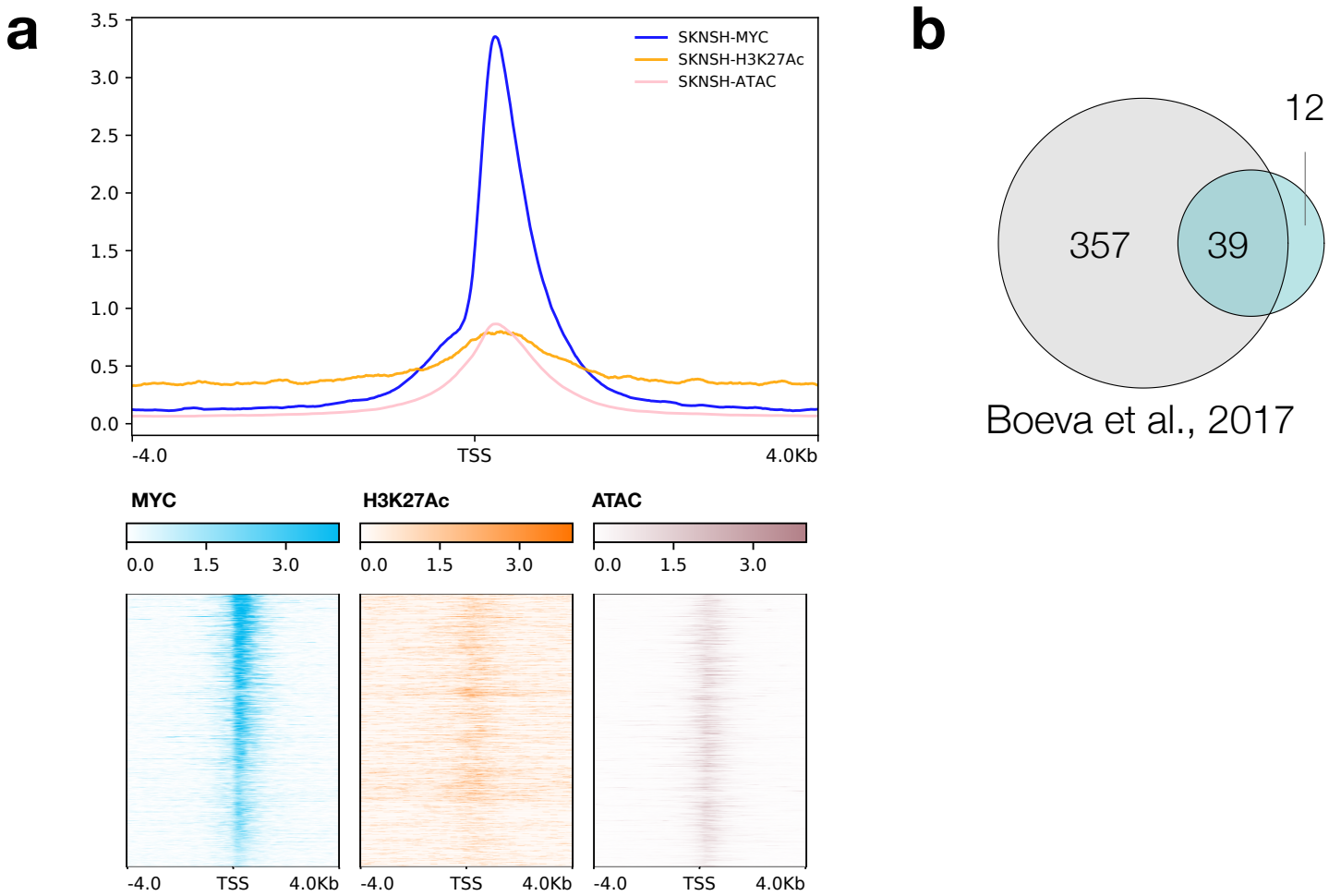




d

Cell Line	MES SEs	ADRN Ses	p-value
LA-N-5	1	11	0.0064*
SK-N-FI	1	10	0.0102*
COG-N-415	0	9	0.0034*
NB-1643	0	9	0.0034*
NGP	0	9	0.0034*
SK-N-BE(2)-C	0	9	0.0034*
NB-69	0	7	0.0106*
KELLY	1	6	0.0699
NB-LS	0	6	0.0191*
SK-N-AS	2	6	0.1512
SK-N-SH	5	2	0.9114

Supplemental Figure 1



Supplemental Figure 1. Integration of ENCODE SK-N-SH H3K27Ac ChIP-Seq with our dataset. Depicted are heatmaps and the binding profiles (**a**) for SK-N-SH MYC, H3K27Ac (ENCODE), and ATAC-Seq for the top 5,000 MYC peaks. (**b**) LILY super-enhancer calls were 76% concordant with Boeva, et. al, 2017 (N = 39/51).