

UC San Diego

UC San Diego Previously Published Works

Title

Identification of a novel gene signature for neuroblastoma differentiation using a Boolean implication network

Permalink

<https://escholarship.org/uc/item/4mh5f5hb>

Journal

Genes Chromosomes and Cancer, 62(6)

ISSN

1045-2257

Authors

Zage, Peter E

Huo, Yuchen

Subramonian, Divya

et al.

Publication Date

2023-06-01

DOI

10.1002/gcc.23124

Peer reviewed



Published in final edited form as:

Genes Chromosomes Cancer. 2023 June ; 62(6): 313–331. doi:10.1002/gcc.23124.

Identification of a Novel Gene Signature for Neuroblastoma Differentiation using a Boolean Implication Network

Peter E. Zage^{1,*}, Yuchen Huo¹, Divya Subramonian¹, Christophe Le Cloennec¹, Pradipta Ghosh^{2,3,4,*}, Debashis Sahoo^{1,5,*}

¹Department of Pediatrics, Division of Hematology-Oncology, University of California San Diego (UCSD), La Jolla, CA

²Department of Medicine, UCSD, La Jolla, CA

³Department of Cellular and Molecular Medicine, UCSD, La Jolla, CA

⁴Veterans Affairs Medical Center, La Jolla, CA

⁵Department of Computer Science and Engineering, Jacobs School of Engineering, UCSD, La Jolla, CA

Abstract

Although induction of differentiation represents an effective strategy for neuroblastoma treatment, the mechanisms underlying neuroblastoma differentiation are poorly understood. We generated a computational model of neuroblastoma differentiation consisting of interconnected gene clusters identified based on symmetric and asymmetric gene expression relationships. We identified a differentiation signature consisting of series of gene clusters comprised of 1251 independent genes that predicted neuroblastoma differentiation in independent datasets and in neuroblastoma cell lines treated with agents known to induce differentiation. This differentiation signature was associated with patient outcomes in multiple independent patient cohorts and validated the role

Name/contact information for corresponding authors: Peter Zage, MD, PhD; Associate Professor, Department of Pediatrics, University of California San Diego School of Medicine; Moores Cancer Center, Room 5311; 3855 Health Sciences Dr, MC 0815; La Jolla, CA 92093-0815. **Phone:** 858-534-6494; **Fax:** 858-822-0022; pzage@ucsd.edu ; Pradipta Ghosh, M.D.; Professor, Departments of Medicine and Cell and Molecular Medicine, University of California San Diego; 9500 Gilman Drive (MC 0651), George E. Palade Bldg, Rm 331-333; La Jolla, CA 92093. **Phone:** 858-822-7633; **Fax:** 858-822-7636; prghosh@ucsd.edu ; Debashis Sahoo, Ph.D.; Assistant Professor, Department of Pediatrics, University of California San Diego; 9500 Gilman Drive, MC 0703, BRF II - 2119; La Jolla, CA 92093-0703. **Phone:** 858-246-1803; **Fax:** 858-246-0019; dsahoo@ucsd.edu.

*These authors contributed equally to this work

Author Contribution Statement

Peter E. Zage: Conceptualization, data curation, formal analysis, funding acquisition, investigation, methodology, project administration, resources, supervision, validation, writing - original draft, writing - review and editing

Yuchen Huo: data curation, formal analysis, investigation, validation, writing - review and editing

Divya Subramonian: data curation, formal analysis, investigation, validation, writing - review and editing

Christophe Le Cloennec: data curation, formal analysis, investigation, validation, writing - review and editing

Pradipta Ghosh: data curation, formal analysis, investigation, validation, resources, supervision, writing - review and editing

Debashis Sahoo: Conceptualization, data curation, formal analysis, funding acquisition, investigation, methodology, project administration, resources, supervision, validation, writing - original draft, writing - review and editing

Declaration of Interest statement: The authors declare no potential conflicts of interest with respect to the research, authorship, and/or publication of this article. P.E.Z. has received research funding from Exelixis, Inc.; Aptose Biosciences, Y-mAbs Therapeutics, and QED Therapeutics, Inc.

Ethics Statement

This study (Protocol 161445) was reviewed and approved by the University of California San Diego Institutional Review Board initially on November 10th, 2016 and renewed on October 12th, 2019

of *MYCN* expression as a marker of neuroblastoma differentiation. Our results further identified novel genes associated with *MYCN* via asymmetric Boolean implication relationships that would not have been identified using symmetric computational approaches and that were associated with both neuroblastoma differentiation and patient outcomes. Our differentiation signature included a cluster of genes involved in intracellular signaling and growth factor receptor trafficking pathways that is strongly associated with neuroblastoma differentiation, and we validated the associations of *UBE4B*, a gene within this cluster, with neuroblastoma cell and tumor differentiation. Our findings demonstrate that Boolean network analyses of symmetric and asymmetric gene expression relationships can identify novel genes and pathways relevant for neuroblastoma tumor differentiation that could represent potential therapeutic targets.

Keywords

Neuroblastoma; differentiation; Boolean; MYCN; UBE4B

Introduction

Children with high-risk neuroblastoma suffer from frequent relapses and treatment-resistant tumors that respond poorly to salvage therapy (1-5), and new treatment strategies are needed. Neuroblastoma tumors arise from primordial neural crest cells and are composed of immature, undifferentiated neuroblasts, and tumor differentiation is strongly associated with patient outcomes (6-9). While induction of differentiation represents a potentially curative therapeutic strategy, the signaling networks underlying neuroblastoma differentiation are not well understood, and improved understanding of the mechanisms of differentiation will identify important contributors that could serve as targets for novel therapies.

Pathway and network analyses have been applied to cancer datasets to find novel oncogenic and regulatory pathways (10-13), but gene associations in these analyses were established using symmetric computational frameworks such as analyses of correlation (14-19), linear regression (20), dimension reduction (21), and clustering (21,22). Gene expression relationships are frequently asymmetric, however, and strategies for exploration of asymmetric relationships are likely to uncover novel pathogenic and regulatory mechanisms. Boolean mathematical tools employ unbiased analyses to identify both symmetric and asymmetric gene expression relationships as well as gene expression changes conserved during biological processes despite disease and tissue heterogeneity that are likely to be fundamentally important for any given process (23).

Boolean analyses assign a parameter (such as the RNA level of a gene) with one of two values, i.e., high or low, and relationships between the expression levels of any pair of genes can then be determined (23,24). The Boolean principle dictates six different possible Boolean implication relationships between expression patterns of two genes: two are symmetric (equivalent or opposite) and four are asymmetric. Because the general relationships among pairs of genes across all samples irrespective of their heterogeneous origin (whether normal tissue or cancer) are likely consistent, these conserved relationships are termed "invariants." Using these invariant Boolean relationships from a given dataset,

a Boolean network can be created (23). Prior studies employed Boolean algorithms and MiDReG (Mining Developmentally Regulated Genes) software to identify novel genes expressed in stem and progenitor cells in normal and malignant tissues and to identify novel therapeutic targets in adult cancers (25-31), demonstrating the potential of this strategy for understanding cancer pathogenesis.

To increase our understanding of the molecular steps and pathways involved in the process of neuroblastoma differentiation and to identify candidate therapeutic targets, we evaluated gene expression profiles of neuroblastoma tumors using novel Boolean Network Explorer (BoNE) software (32,33) that incorporates MiDReG algorithms. Our analyses identified a gene signature comprised of a series of gene clusters significantly associated with neuroblastoma differentiation based on identified invariant relationships. We further validated the associations of these genes with patient survival and identified novel candidate genes associated with differentiation and with asymmetric implication relationships with *MYCN* that were only detectable using Boolean approaches. Our data demonstrate that Boolean approaches for analysis of gene expression profiles can identify novel genes and pathways associated with neuroblastoma tumor differentiation and identified a role for growth factor receptor trafficking pathways in neuroblastoma differentiation and pathogenesis.

Materials and Methods

Data collection and annotation

Publicly available microarray and gene expression databases were downloaded from the National Center for Biotechnology Information (NCBI) Gene Expression Omnibus (GEO) website (34,35) and the European Molecular Biology Laboratory European Bioinformatics Institutes (EMBL-EBI) ArrayExpress website (36,37). NCBI GEO and EMBL-EBI ArrayExpress were searched for transcriptomic studies of neuroblastoma patient primary tumor samples (38-43), of transcriptomic studies of neuroblastoma cell lines before and after differentiation-inducing treatment (44-51), and of single-cell RNA-sequencing of neuroblastoma cell lines and tumor samples (52,53). Accession numbers for these crowdsourced datasets are provided (Supplemental Table 1). The National Cancer Institute (NCI) Therapeutically Applicable Research to Generate Effective Treatments dataset (TARGET; phs000467.v1.p1) and the KidsFirst: Neuroblastoma dataset (KidsFirstNB; phs001436.v1.p1) were downloaded from the NIH database for genotypes and phenotypes (NIH dbGaP; https://www.ncbi.nlm.nih.gov/projects/gap/cgi-bin/study.cgi?study_id=phs000218.v4.p1) (54,55).

Boolean Network Explorer (BoNE) analysis

Previously developed and validated computational tools, including StepMiner (24); MiDReG (25), and BooleanNet statistics for the identification of Boolean Implication relationships (23) were integrated into a single platform, the Boolean Network Explorer (BoNE) (see Supplementary Methods; workflow summarized in Supplementary Figures 1,2; 32,33), that analyzes both symmetric and asymmetric properties of Boolean implication relationships to discover natural progressive time-series changes in major cellular processes.

Gene expression summarization from neuroblastoma patient tumors and cell lines was performed by normalizing Affymetrix platforms by RMA (Robust Multichip Average) (56,57) and RNA-Sequencing platforms by computing TPM (Transcripts Per Millions) (58,59) values whenever normalized data were not available in GEO. Boolean logic (23-25,60) is a simple mathematic relationship of two values, i.e., high/low, 1/0, or positive/negative, and expression levels of all genes in each dataset were converted into Boolean values (0/1) using the StepMiner algorithm (24). A noise margin of 2-fold change is applied around the threshold to determine intermediate values, and these values are ignored during Boolean analysis. Pairwise comparison of relative gene expression levels was then carried out to capture fundamental symmetric and asymmetric relationships between genes. BooleanNet statistics (S , p) were used to assess the significance of the Boolean implication relationships (23) and binarization error is controlled using a noise margin of ± 0.5 around the determined StepMiner threshold. $S > 3$ and $p < 0.1$ were used as thresholds, where $S = (\text{expected} - \text{observed}) / (\text{expected})$, for a false discovery rate $< 10^{-3}$ (23,25).

Invariant Boolean relationships were then used to make a universal model. These relationships were simplified by first clustering genes with equivalent relationships to each other. Internal logically consistent clusters were identified by following the equivalent relationships from genes that share at least half of the equivalences within a cluster. Subsequently, a graph was built connecting the individual clusters to each other using Boolean relationships. To identify cluster connections, the most prevalent Boolean relationships were used from ten statistically top ranked members from each cluster. To identify the global continuum of cellular states along the stemness-differentiation axis in neuroblastoma tumors, pathway discovery was performed using the MiDReG algorithm (25). The largest cluster was used as a starting point and other clusters were discovered using specific Boolean implication relationships. A greedy algorithm (61) was employed for choosing clusters during depth-first traversal and was used to identify the pathway involving sequential gene clusters with the strongest association with differentiation (clusters 3-4-5-6) that was designated the differentiation signature. For further validation, multiple pathways were derived from connected clusters in the network as previously described (32), and machine learning and model training were performed by ranking Boolean cluster pathways including up to 4 connected clusters based on their performance (ROC-AUC values) in distinguishing neuroblastoma differentiation as defined by *MYCN* expression. Reactome pathway analysis of each cluster was then performed to identify and organize molecular relationships into biological pathways and processes and to identify the enriched pathways (62).

For comparison studies, a gene signature score from the differentiation signature was computed by using scaled linear combinations of gene expression values which were used to classify sample categories, and the performance of the multi-class classification was measured by ROC-AUC (Receiver Operating Characteristics Area Under the Curve) values (see Supplemental Methods). Comparisons of the differentiation signature with gene expression datasets from independent patient cohorts were analyzed separately using the Boolean analysis framework. A color-coded bar plot was combined with a density plot to visualize the gene signature-based classifications. Boxplots and scatterplots of gene

expression values were computed using the Hegemon (**H**ierarchical **e**xploration of **g**ene **e**xpression **m**icroarrays **o**nline) computational algorithm (28,29,31).

For pathway-based signature analysis, the differentiation signature score was compared to gene signatures from four different pathways: MES (mesenchymal; 63), ADRN (adrenergic; 63), cell cycle (KEGG_CELL_CYCLE, hsa04110; 64), and senescence (FRIDMAN_Senescence_UP, FRIDMAN_Senescence_DN; 65). The gene lists for the cell cycle and senescence pathways were downloaded from the Molecular Signatures Database (MSigDB), a joint project between UC San Diego and the Broad Institute (66). Genes from these pathways were converted to a gene signature score by using the average of the normalized gene expression values based on the Z-score approach centered around the StepMiner threshold (formula = $(\text{expr} - \text{SThr})/3 * \text{stddev}$; see Supplemental Methods).

For selection of candidate genes, genes from clusters 3, 4, 5, and 6 were ranked based on T-test scores between gene expression levels from cell lines with less differentiation (IMR-32 with and without retinoic acid treatment, SK-N-AS with and without retinoic acid treatment, and untreated SK-N-BE(2) cells) and gene expression levels from cell lines with evidence of differentiation (SK-N-BE(2) cells after retinoic acid treatment and SK-N-SH cells with and without retinoic acid treatment) (see Figure 2; Table 1). Analogous comparisons were made comparing T-test scores between gene expression levels from cell lines with and without retinoic acid treatment (see Table 1). Additional genes whose expression was significantly associated with expression of *MYCN* through one of the four asymmetric Boolean implication relationships (high->high, high->low, low->high, low->low) without direct symmetric correlations with *MYCN* expression were selected and prioritized based on BooleanNet statistics and absence of prior published associations with neuroblastoma differentiation or patient outcomes.

Standard t-tests were performed using python `scipy.stats.ttest_ind` package (version 0.19.0) with Welch's Two Sample t-test (unpaired, unequal variance (`equal_var=False`), and unequal sample size) parameters. Multiple hypothesis correction was performed by adjusting *p* values with `statsmodels.stats.multitest.multipletests` (`fdr_bh`: Benjamini/Hochberg principles). The results were independently validated with R statistical software (R version 3.6.1; 2019-07-05). Patient subgroups were compared with respect to survival outcomes with the use of Kaplan–Meier curves, log-rank tests, and multivariate analyses based on the Cox proportional-hazards method. Kaplan-Meier analyses were performed using `lifelines` python package version 0.14.6 and results were validated with R statistical software. Statistical significance of the difference between patient groups is computed using log-rank test. Univariate and multivariate survival analysis is performed using R survival package.

Cell culture

The neuroblastoma cell lines SK-N-AS, SK-N-SH, SK-N-BE(2), CHP-134, IMR-32, NGP, LAN1, SH-SY5Y, and SH-EP have been previously described (67,68) and were generously provided by the Children's Oncology Group (COG) Childhood Cancer Repository (www.cccells.org), Susan Cohn (The University of Chicago Children's Hospital, Chicago, IL), John Maris (Children's Hospital of Philadelphia, Philadelphia, PA), or were purchased

from the American Type Culture Collection (ATCC; Rockville, MD). Cell lines were maintained in RPMI-1640 (Mediatech Inc, Manassas, VA) with 10% fetal bovine serum (FBS; Omega, Tarzana, CA), L-glutamine, non-essential amino acids (Mediatech), sodium pyruvate (Mediatech), and Hyclone antibiotic/antimycotic (Fisher Scientific, Hampton, NH) at 37°C in 5% CO₂. All cell lines were authenticated by DNA profiling prior to use.

Therapeutic agents

Vorinostat (Suberoylanilide Hydroxamic Acid, SAHA) was purchased from Selleck Chemicals (Houston, TX) and 13-*cis*-retinoic acid from Sigma-Aldrich (St. Louis, MO). All compound preparations were stored at -20°C, with dilutions maintained at 4°C for experimental use. 13-*cis*-retinoic acid was diluted directly into media prior to use.

RNA-sequencing

RNA-sequencing of SK-N-AS, SK-N-SH, SK-N-BE(2), CHP-134, IMR-32, NGP, and LAN1 neuroblastoma cells was performed as previously described (69); see Supplemental Methods for details.

Quantitative PCR

SK-N-AS and SK-N-BE(2) neuroblastoma cells were plated and allowed to proliferate until approximately 70% confluent. Cells were harvested, and RNA was isolated from cell populations using a Qiagen RNeasy Kit (Qiagen). RNA was reverse transcribed to cDNA and qPCR was subsequently performed using primers for *PCDHA12*, *DCLK1*, and *UBE4B* with *GAPDH* used as a control. Fold change in gene expression was calculated by comparing levels of the gene of interest against *GAPDH*.

Western blots

SK-N-AS and SK-N-BE(2) neuroblastoma cells were plated in 6-well plates at approximately 70–80% confluence and allowed to adhere overnight. Wells were treated with 5µM 13-*cis*-retinoic acid for 2, 5, and 7 days. Cells were then washed with PBS and lysed with RIPA buffer supplemented with Protease inhibitor and phosphatase inhibitor (Life Technologies, Carlsbad, CA). Lysates were centrifuged and the supernatants were collected. Protein concentration in cell lysates was measured using a protein assay Dye Reagent (Bio-Rad, Hercules, CA). 20–30µg of denatured total protein from each sample was separated by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) using 4–12% Bolt gels (Invitrogen, Carlsbad, CA) and transferred to PVDF membranes. Membranes were blocked with 3% BSA made in 1× TBS + 0.1% Tween-20 and then incubated with primary antibodies to UBE4B (Abcam, Cambridge, United Kingdom; ab97697) and β-actin (Sigma-Aldrich; AS316). All antibodies were diluted in 5% BSA in 1× TBS + 0.01% Tween-20. Bound primary antibodies were incubated in anti-rabbit or anti-mouse HRP-conjugated secondary antibodies (1:5000, Sigma-Aldrich) at room temperature for 1 hour and the signal was visualized using Amersham ECL (GE Healthcare Bio-Sciences, Pittsburgh, PA).

***In vitro* growth and differentiation assays**

Neuroblastoma cells with depleted UBE4B were generated using lentiviral particles produced in HEK 293T cells using 3rd Generation Packaging Mix and UBE4B sgRNA CRISPR/Cas9 All-in-One Lentivector set (Human) containing plasmid pLenti-U6-sgRNA-SFFV-Cas9-2A-Puro (Applied Biological Materials Inc., Richmond, BC, Canada). Lentivirus containing culture media was collected from 293T cells, centrifuged at 300g for 10 minutes to remove cells, filtered through a 0.45 μ M sterile filter and added to target cells with 6 μ g/ml Polybrene. Target cells were incubated at 37°C overnight, and then culture media was replaced with complete media (RPMI supplemented with 10% FBS, 2mM L-glutamine, 1X Antibiotic-Antimycotic solution (Corning) and incubated at 37°C. Infected cells were selected for stable expression with 1 μ g/ml puromycin and UBE4B depletion was validated by quantitative PCR and Western blot.

Parental and UBE4B depleted neuroblastoma cells were plated in 96-well plates at seeding densities between 5,000–10,000 cells/well for 24 hours. Plates were placed in the Incucyte Zoom™ continuous live cell imaging system (Essen Bioscience, Ann Arbor, MI) and phase contrast images were taken every 6 hours at 10 \times magnification for up to 192 hours. Cell growth curves were generated from percent cell confluence acquired from the Incucyte Zoom™ analyzer as previously described (67,68). The change in percent confluence over time was normalized to the percent confluence at time zero. All experiments were performed in triplicate, and differences in normalized confluence were compared using Student's T-tests. For neurite growth experiments, cells were treated with 13-*cis*-retinoic acid or vorinostat, and neurite length per field was quantified using NeuroTrack™ processing software on the Incucyte Zoom™ as previously described (67,68). The growth rate of neurites in each well was obtained by measuring the total neurite length per high-power field at each time point and expressed as mm/mm².

Results

Boolean Network Explorer (BoNE) Analysis Reveals an Invariant Signature for Neuroblastoma Differentiation

Because neuroblastoma tumors arise from primordial neural crest cells and are composed of immature, undifferentiated neuroblasts, and tumor differentiation is strongly associated with improved patient outcomes (6-9), we sought to create a mathematical model for neuroblastoma differentiation by extracting the fundamental gene expression relationships from transcriptomic profiles of neuroblastoma tumors to investigate progressive changes in gene expression along the stemness-differentiation continuum. Amplification of the chromosome 2p24 region, including the oncogene *MYCN*, is found in ~25% of all tumors and is a marker of high-risk disease and poor prognosis (71,72). Elevated *MYCN* expression is also known to drive stemness (73-75) and reduced *MYCN* is often observed during neuroblastoma differentiation (76-79), and *MYCN* expression, but not expression of the related *MYC* gene, is significantly correlated with neuroblastoma tumor differentiation (Supplemental Figure 3). *MYCN* expression is also reduced in both *MYCN*-amplified and nonamplified cell lines and tumors (Supplemental Figure 3), suggesting that

MYCN expression represents a marker of neuroblastoma phenotypes along the stemness-differentiation continuum (Figure 1A, B).

Gene expression levels from 498 neuroblastoma patient tumors (38) including a representative sample of patients and tumors from all risk groups with representative distributions of all known prognostic features were converted to Boolean values, and Boolean cluster relationships were used to chart individual gene expression changes along a Boolean path within the stemness-differentiation continuum using Boolean Network Explorer (BoNE) software, which clusters genes using Boolean equivalent relationships. Pairwise Boolean implication relationships were used to create a Boolean network (23,24) with the edges between clusters defined using the identified implication relationships, and the path direction was derived from the connections established from high *MYCN* expression (representing undifferentiated neuroblastoma) to low *MYCN* expression (representing differentiated neuroblastoma), with the arrow colors reflecting the implication relationships (Figure 1C). We then generated a map using BoNE software with identified Boolean relationships representing the sequential progression of gene clusters in neuroblastoma tumors from undifferentiated to differentiated tumors (Figure 1D).

Using a greedy algorithm (61) and a classical machine learning framework, we performed model selection to identify gene clusters associated with neuroblastoma differentiation based on *MYCN* expression. The pathway involving four sequential gene clusters (clusters 3, 4, 5, and 6) consisting of 1440 total genes (with 1251 independent genes) emerged as one of the best models that spanned multiple clusters with the strongest association with neuroblastoma differentiation (Figure 1D,E,F). Reactome pathway analyses (62) of each of the clusters in the continuum between undifferentiated and differentiated neuroblastoma were then performed, with significant pathways for each cluster identified (Figure 1D). These four gene clusters demonstrated enrichment of gene transcription and protein translation pathways in clusters 3 and 4 and enrichment of intracellular signaling, receptor trafficking, and cell-cell communication and signaling pathways in clusters 5 and 6 (Supplemental Table 2), with a progression of cellular functions likely required for neuroblastoma differentiation. As the 3-4-5-6 gene clusters demonstrated the strongest association with neuroblastoma differentiation, they will henceforth be referred to as a differentiation signature.

The Differentiation Signature Predicts Neuroblastoma Cell and Tumor Differentiation

To validate the association of our differentiation signature with neuroblastoma cell differentiation, we projected our Boolean network onto gene expression profiles from neuroblastoma cell lines with both low and high *MYCN* gene expression levels before and after treatment with the differentiating agent 13-*cis*-retinoic acid (CRA) using BoNE software (Figure 2A). Comparing Boolean continuum scores calculated from expression levels of genes in the differentiation signature, the differentiation signature was able to perfectly separate gene expression profiles of CRA-treated neuroblastoma cell lines from untreated cells (ROC-AUC = 1.00; Figure 2B). CRA treatment also resulted in increased differentiation signature scores in neuroblastoma cells sensitive to CRA (SK-N-SH, $p=3.86e-06$; SK-N-BE(2), $p=0.000107$), while CRA treatment resulted in smaller fold changes in differentiation signature scores in neuroblastoma cells resistant to CRA

(IMR32, $p=2.51e-05$; SK-N-AS, $p=0.0479$) (80) (Figure 2C). SK-N-SH neuroblastoma cells, which consist of a mixture of both more differentiated N-type SH-SY5Y cells and less differentiated S-type SH-EP cells (80), demonstrated an increased differentiation signature score at baseline that was further increased with CRA treatment (Figure 2C).

Our Boolean continuum scores were also able to perfectly separate gene expression profiles of SK-N-SH and SH-SY5Y neuroblastoma cells after 5 days of treatment with 10 μ M all-trans-retinoic acid (ATRA)(48; ROC-AUC=1.00, Figure 2D). The differentiation signature was able to perfectly separate undifferentiated LAN1 neuroblastoma cells from LAN1 cells differentiated with the combination of 10 μ M ATRA and 5-aza-deoxycytidine, of undifferentiated SK-N-BE(2)C neuroblastoma cells from SK-N-BE(2)C cells differentiated with 10 μ M ATRA, and of undifferentiated SH-SY5Y cells from SH-SY5Y cells differentiated with 10 μ M ATRA and neurobasal medium (44-46,50; ROC-AUC=1.00 for each; Figure 2E). The differentiation signature was also able to predict responses of neuroblastoma cells to the histone deacetylase (HDAC) inhibitor vorinostat (suberoylanilide hydroxamic acid/SAHA) (47; Figure 2E) and was able to separate neuroblastoma cells with overexpressed *TFAP2B* and depleted *MYCN* (50,51; Figure 2E).

We then projected our Boolean network onto gene expression profiles from neuroblastoma tumors in the TARGET database (54). Comparing Boolean continuum scores, the differentiation signature was able to separate differentiating from undifferentiated neuroblastoma tumors in the TARGET neuroblastoma tumor gene expression dataset ($p=2.923-06$; Figure 2F). The differentiation signature was also able to separate nodular from intermixed ganglioneuroblastoma tumors from the KidsFirst neuroblastoma dataset (49; $p=0.0397$; Figure 2G). Using single-cell RNA-sequencing data from human fetal adrenal samples, the differentiation signature was able to distinguish neuroblastoma tumor samples from Schwann cell precursors (SCPs), Sympathoblasts (SBs), and Chromaffin (CF) cells (52; Figure 2H). In single-cell RNA-sequencing data from neuroblastoma tumor samples from untreated and previously treated patients, the differentiation signature distinguished neuroblastoma cells from SCPs and CF cells, but there was no significant difference from SBs (53; Figure 2I), further validating the significance of these clusters in neuroblastoma tumor differentiation.

The Differentiation Signature Predicts Neuroblastoma Patient Outcomes

To evaluate the associations of the differentiation signature with patient outcomes, we projected our Boolean network onto gene expression profiles of neuroblastoma tumors from databases that included patient survival information. Elevated differentiation signature scores were also found in patients with neuroblastoma who were alive at the completion of standard therapy, compared to scores in patients who did not survive (43; $p=0.00995$; Figure 3A). To further evaluate the association of the differentiation signature score with neuroblastoma patient outcomes, available datasets of gene expression profiles of neuroblastoma tumors with information about patient survival were divided into subsets with high and low *MYCN* expression. Among tumors with low *MYCN* expression levels, datasets were further subdivided by differentiation signature gene expression levels into low and high expression groups, and Kaplan-Meier curves were generated. Overall survival rates in

patients with tumors with low *MYCN* expression with low cluster differentiation signature gene expression levels were lower in all tested neuroblastoma tumor gene expression profiles (38-42,55); Figure 3B,C,D).

In both univariate and multivariate Cox regression analyses for overall survival (OS) using gene expression patterns and prognostic features from 498 patients (38), including the differentiation signature score, *MYCN* amplification status, tumor stage, patient age at diagnosis, and up or down regulated gene set expression from LAN1 neuroblastoma cells treated with 10 μ M ATRA and 5-aza-deoxycytidine (44) as independent variables, the differentiation signature score was independently associated with overall patient survival (** p < 0.01, *** p < 0.001; Figure 3E).

Association of the Differentiation Signature with Established Neuroblastoma Gene Signatures

Since *MYCN* expression has been shown to be associated with cell fates other than differentiation, including cell cycle progression (81), mesenchymal transition (63) and cellular senescence (49), the differentiation score was compared to gene signatures from four different pathways: MES (mesenchymal; 63); ADRN (adrenergic; 63), cell cycle (KEGG_CELL_CYCLE, hsa04110; 64) and senescence (FRIDMAN_SENESCENCE_UP, FRIDMAN_SENESCENCE_DN; 65). While the differentiation signature score was consistently able to separate undifferentiated from differentiated neuroblastoma cells and tumors (Figure 2B-F), gene signature scores for mesenchymal and adrenergic states and for cell cycle and cellular senescence pathways were unable to consistently separate undifferentiated neuroblastoma cell lines from cell lines differentiated with 13-*cis*-retinoic acid (Figure 4A). Comparisons of the signature-based classifications across all datasets further demonstrated that the differentiation signature score had the best performance and highest ROC-AUC value in separating undifferentiated from differentiated cell lines using data compiled from all previously tested datasets, while other tested gene signatures demonstrated lower ROC-AUC values in at least one dataset with reduced performance in separating undifferentiated from differentiated cell lines (Figure 4B). Additionally, recently identified pediatric cancer gene modules were unable to significantly separate undifferentiated neuroblastoma cell lines from cell lines differentiated with 13-*cis*-retinoic acid (82; Figure 4E, Supplemental Figure 4).

The differentiation signature score was also able to separate neuroblastoma cell lines treated with all-trans retinoic acid (ATRA) to induce differentiation from those treated with topotecan (49), although topotecan treatment appears to also induce neuroblastoma cell line differentiation, as demonstrated by increasing differentiation signature score (Figure 4C, **left**). The senescence gene signature (65) was also able to segregate neuroblastoma cells treated with topotecan from those treated with ATRA (Figure 4D, **left**). However, while the differentiation signature was able to perfectly segregate undifferentiated neuroblastoma cell lines from those differentiated with 13-*cis*-retinoic acid (Figure 4C, **right**), the senescence signature was not able to consistently and correctly segregate undifferentiated from differentiated neuroblastoma cells (Figure 4D, **right**), suggesting that the differentiation score reflects the association of *MYCN* expression with differentiation.

Boolean Analyses Identifies Novel Genes Associated with *MYCN* Expression and Neuroblastoma Differentiation

In order to identify novel genes from our differentiation signature associated with neuroblastoma differentiation, genes from clusters 3, 4, 5, and 6 were ranked based on significance between gene expression levels from cell lines with less differentiation (IMR-32 with and without retinoic acid treatment, SK-N-AS with and without retinoic acid treatment, and untreated SK-N-BE(2) cells) and gene expression levels from cell lines with evidence of differentiation (SK-N-BE(2) cells after retinoic acid treatment and SK-N-SH cells with and without retinoic acid treatment) (Figure 2; Supplemental Table 3A). Analogous comparisons were made comparing T-test scores between gene expression levels from cell lines with and without retinoic acid treatment (Supplemental Table 3B).

The gene lists were then screened for genes whose expression was significantly associated with expression of *MYCN* through one of the four asymmetric Boolean implication relationships (high->high, high->low, low->high, low->low) without direct symmetric correlations with *MYCN* expression and without prior published associations with neuroblastoma differentiation or patient outcomes, and five novel genes were identified (Table 1, Figure 5). These genes were located in clusters 3/4 (*PSMC3*, *SLC25A39*, *PFDN2*) and clusters 5/6 (*PCDHA12*, *DCLK1*) and were associated with *MYCN* expression via high->high (*PSMC3*, *SLC25A39*, *PFDN2*) or high->low (*PCDHA12*, *DCLK1*) implication relationships (Figure 5). In neuroblastoma cell lines, the *SLC25A39* gene from clusters 3/4 demonstrated significantly reduced expression after CRA treatment, while *DCLK1* and *PCDHA12* from clusters 5/6 demonstrated significantly increased expression after CRA treatment (Figure 5, Supplemental Figure 5), consistent with their Boolean relationships with *MYCN* expression. Furthermore, increased expression of *PCDHA12* ($p=1.8e-38$) and *DCLK1* ($p=7.5e-40$) were associated with increased rates of survival in *MYCN*-nonamplified tumors, while increased expression of *PSMC3* ($p=2.0e-39$), *SLC25A39* ($p=3.4e-40$), *PFDN2* ($p=9.9e-40$) were each associated with lower survival rates in *MYCN*-nonamplified tumors (Figure 5).

UBE4B is a Network-predicted Modulator of Neuroblastoma Differentiation

In our Boolean network, The *UBE4B* gene was identified in cluster 5, one of the gene clusters within the differentiation signature and was found to be equivalent to other genes involved in intracellular signaling and receptor trafficking pathways (Figures 1D, 6A). *UBE4B* is an E3/E4 ubiquitin ligase involved in growth factor receptor trafficking in neuroblastoma tumors that has been shown to be associated with neuroblastoma tumor differentiation (69,83,84), and to validate the results of our Boolean analyses, we evaluated neuroblastoma cells before and after treatment with 13-*cis*-retinoic acid (CRA) for *UBE4B* expression. Both *UBE4B* gene and protein expression increased in response to CRA-treatment in CRA-sensitive SK-N-BE(2) neuroblastoma cells while *UBE4B* gene and protein expression decreased in CRA-resistant SK-N-AS neuroblastoma cells ($*p<0.05$; Figure 6B, C), although the more significant observed changes in protein expression compared to gene expression suggest that other post-translational mechanisms may be involved.

Elevated expression of *UBE4B* and the previously established neuroblastoma differentiation marker *GAP43* (85) were also found in differentiating neuroblastoma tumors, compared to undifferentiated tumors, while *MYCN* expression demonstrated lower expression in differentiating tumors, consistent with the established association of *UBE4B* and *MYCN* expression with neuroblastoma differentiation (69,74,76; Figure 6D). We further analyzed the associations of the expression patterns of *UBE4B* and other genes associated with neuroblastoma differentiation (*GAP43*, *SYP*), and *UBE4B* expression in patient tumor samples demonstrated expected Boolean relationships with neuroblastoma differentiation markers *GAP43* and *SYP*, but was not associated with expression of *NTRK2*, *RPFOX3*, or *NPY*, genes previously shown to have increased expression after retinoic acid treatment (86; Figure 6E). Lastly, *UBE4B* depletion in neuroblastoma cell lines also resulted in both in increased growth rate and inhibition of neurite outgrowth in response to either CRA or to histone deacetylase inhibition ($p=5.9e-08$ and $p=4.8e-13$, respectively; Figure 6F,G; Supplemental Figure 6), consistent with reduced neuroblastoma differentiation as predicted by our Boolean network and differentiation signature.

Discussion

Neuroblastoma tumors are childhood tumors that arise from primordial neural crest cells that normally differentiate to form sympathetic ganglion and adrenal chromaffin cells (87-89). Children with high-risk, undifferentiated neuroblastoma suffer from frequent relapses and treatment-resistant tumors that respond poorly to salvage therapy (1-5), while children with well-differentiated tumors have improved survival rates (6-9). Neuroblastoma is also characterized by the capacity for spontaneous tumor regression in a subset of patients (90), resulting in part through neuronal differentiation, and aggressive neuroblastoma tumors can be induced to differentiate along sympathoadrenal lineage pathways, suggesting that a block in the differentiation process likely contributes to neuroblastoma pathogenesis. While induction of differentiation is an effective strategy for neuroblastoma treatment, our understanding of the signaling pathways and networks in the process of neuroblastoma differentiation is incomplete and does not fully account for the complex interplay of intracellular signaling pathways with transcriptional, post-transcriptional, and other regulatory events that underlie the differentiation process. Improved understanding of the mechanisms regulating neuroblastoma differentiation may highlight events that contribute to neuroblastoma tumorigenesis and identify important contributors that could serve as targets for novel therapies.

The multiple, overlapping roles of protein kinases, transcription factors, and other regulatory molecules in neuroblastoma differentiation are best analyzed using a systems-based approach that can more effectively model complex intracellular signaling networks and their role in cellular behaviors. Boolean networks are models of biological networks used to study complex behavior in biological systems, and Boolean mathematical models employ unbiased analyses of transcriptomic datasets to identify general relationships among pairs of genes across all samples irrespective of their origin or prior manipulation, and these relationships are conserved despite disease and tissue heterogeneity and are likely to be fundamentally important for any given process. The Boolean principle dictates only six different possible relationships between the expression levels of any pair of genes:

two are symmetric (equivalent or opposite) and four are asymmetric (low => low, high => low, low => high, and high => high)(Figure 1). A Boolean implication relationship between expression levels of two genes exists when one or more quadrants of a gene expression profile are sparsely populated (23). Analyses of Boolean relationships from a given dataset can be used to generate a Boolean network, which can then be used to identify invariant gene expression changes conserved during biological processes despite disease and tissue heterogeneity and disease evolution, and these Boolean networks can then be easily extrapolated to independent cohorts for validation.

Boolean analyses involve the conversion of individual gene expression patterns into binary values. While use of continuous expression data would incorporate more information, the use of continuous data for the analyses would lead to increased "noise" that would increase the difficulty in interpretation of the results. Prior studies have employed Boolean algorithms and the MiDReG (Mining Developmentally Regulated Genes) software to successfully identify novel genes expressed in stem and progenitor cells in both normal and malignant tissues (25-31) and to identify novel therapeutic targets in bladder (26,27,31) and colorectal cancers (28,29). Using recently developed Boolean Network Explorer (BoNE) software that uses MiDReG principles to sort developmentally regulated genes along the stemness-differentiation axis, we have generated a map with identified Boolean relationships representing the sequential progression of gene clusters in neuroblastoma tumors from undifferentiated to differentiated tumors. Using our Boolean network, we have identified a differentiation signature comprised of 1251 independent genes in 4 gene clusters, and this signature predicted neuroblastoma differentiation in independent datasets and in neuroblastoma cell lines treated with agents known to induce differentiation. This differentiation signature was also associated with patient outcomes in multiple independent patient cohorts and validated the role of *MYCN* expression as a marker of neuroblastoma differentiation. Our results identified novel genes whose expression demonstrated asymmetric Boolean implication relationships with *MYCN* expression and that were associated with both neuroblastoma differentiation and patient outcomes, suggesting that our Boolean network was able to identify novel genes associated with the differentiation process and with *MYCN* expression that would not have been detected using symmetric computational analyses. Our differentiation signature included a cluster of genes involved in intracellular signaling and growth factor receptor intracellular trafficking pathways that is strongly associated with neuroblastoma differentiation, and *UBE4B*, a gene within this cluster associated with growth factor receptor trafficking and intracellular signaling, was associated with neuroblastoma cell and tumor differentiation. Our results have shown that Boolean analyses are able to identify gene expression patterns that more accurately predict neuroblastoma cell and tumor differentiation status, demonstrating the advantages of our approach.

Since prior studies have shown that elevated *MYCN* expression drives neuroblastoma stemness (73-75) and that reduced *MYCN* expression is observed during neuroblastoma differentiation (76-79), we evaluated neuroblastoma differentiation using gene expression relationships from transcriptomic profiles of neuroblastoma tumors to investigate progressive gene expression changes along the stemness-differentiation continuum defined by *MYCN* expression. *MYCN* expression has previously been shown to be associated

with cell fates other than neuronal-like differentiation, such as mesenchymal transition (63) and senescence (49), but our results demonstrate that gene signatures associated with mesenchymal and adrenergic phenotypes and those associated with cell cycle progression and cellular senescence were unable to consistently distinguish undifferentiated from differentiated neuroblastoma cells, while our differentiation signature was consistently able to separate undifferentiated from differentiated neuroblastoma cells and tumors. Furthermore, pathways involved in cell cycle progression and cellular mitosis that might be linked to *MYCN* expression were located in an alternate gene cluster (cluster #10; Figure 1D) that was identified in unbiased fashion by our analysis and that was not included in the differentiation signature, and gene cluster #10 was further located in a completely independent path of gene clusters along the stemness-differentiation continuum (Figure 1D,F). Furthermore, some neuroblastoma tumors are not marked by expression of *MYCN*, but rather, the related c-MYC protein (91). However, the role of c-MYC in neuroblastoma differentiation is not clear, and *MYC* expression is not associated with patient tumor differentiation (Supplemental Figure 3), and our results are generally identical in both *MYCN*-amplified and –non-amplified tumors and cell lines (for example, see Figures 2B and 3E), demonstrating the likely significant role of *MYCN* expression in neuroblastoma differentiation.

The four gene clusters identified in our differentiation signature demonstrated enrichment of gene transcription and protein translation pathways in clusters 3 and 4 and enrichment of intracellular signaling, receptor trafficking, and cell-cell communication and signaling pathways in clusters 5 and 6. While the relative roles of these individual intracellular and extracellular pathways and of individual genes within each cluster remain to be fully delineated, this progression of cellular functions appears to be consistent with a process required for a complex cellular behavior such as differentiation. Further studies are clearly needed to identify the most critical genes and pathways within these clusters that drive the differentiation process. However, we have demonstrated associations between our identified differentiation signature and neuroblastoma differentiation induced *in vitro* by a range of agents, including retinoids and HDAC inhibitors, which have been shown to induce neuroblastoma differentiation in previous studies (45,92-94), as well as associations with neuroblastoma tumor differentiation in multiple independent patient tumor gene expression datasets (54,55), further establishing the validity of our Boolean network. Diverse neuroblastoma prognostic features, such as tumor stage and patient age at diagnosis, are all included within the Boolean analysis framework, and the analyzed datasets include patients with the typical distribution of risk groups, tumor stages, patient ages, and other biologic and prognostic features (38-43), suggesting that our Boolean network and differentiation signature represent invariant gene expression patterns conserved during differentiation despite disease and tissue heterogeneity and disease evolution.

Our Boolean network identified five novel genes, 3 in clusters 3/4 (*PSMC3*, *SLC25A39*, *PFDN2*) and 2 in clusters 5/6 (*PCDHA12*, *DCLK1*), whose expression was associated with *MYCN* expression via asymmetric Boolean implication relationships and with neuroblastoma cell differentiation and patient outcomes. None of the five genes has previously been associated with neuroblastoma differentiation or survival in published literature to our knowledge. While the genes in clusters 3/4 have known or suggested

functions that are not clearly linked to neural differentiation, the identified genes in clusters 5/6 have more established roles in neural development. The *PSMC3* gene encodes one of the ATPase subunits of the 26S proteasome and has been associated with T-cell receptor signaling and with cell cycle checkpoints. *SLC25A39* is a member of the SLC25 transporter or mitochondrial carrier family of proteins that are typically embedded in the inner mitochondrial membrane, although its specific function is unknown. The *PFDN2* gene encodes a subunit of prefoldin, a molecular chaperone complex that binds and stabilizes newly synthesized polypeptides to allow for appropriate folding. Of the genes in clusters 5/6, however, *PCDHA12* encodes a neural protocadherin cell adhesion protein that has been suggested to be involved in the establishment and maintenance of specific neuronal connections in the brain, and *DCLK1* encodes a kinase that has been shown to also have microtubule-polymerizing activity and is involved in neuronal migration, retrograde transport, neuronal apoptosis and neurogenesis and is up-regulated by brain-derived neurotrophic factor. The functional roles of these genes in neuroblastoma differentiation and pathogenesis are currently under further investigation.

To further validate our differentiation signature, and in particular the identified cluster of genes involved in receptor endocytosis and trafficking and intracellular signaling, we evaluated the association of UBE4B, an E3/E4 ubiquitin ligase (95) that has been linked to neuroblastoma differentiation (69), in neuroblastoma cell lines and tumors. We have identified increases in *UBE4B* expression in differentiated neuroblastoma cells and tumors, associations with differentiation marker genes, and increased growth and reduced neurite outgrowth in neuroblastoma cells with depleted UBE4B, suggesting potential roles for UBE4B ubiquitin ligase activity and UBE4B-mediated GFR trafficking and degradation in the regulation of downstream signaling required for neuroblastoma differentiation. However, the molecular mechanisms underlying the associations of *UBE4B* expression and UBE4B-mediated signaling with differentiation are not known. Aberrant expression and activity of several GFRs, including RET, EGFR, TrkA, and TrkB, have been associated with neuroblastoma differentiation (96-108). While further studies are clearly needed to determine the relative roles of individual intracellular and extracellular pathways and of individual genes within each cluster, our data demonstrating the association of *UBE4B* with neuroblastoma differentiation suggests that UBE4B-mediated signaling may play a key role in the differentiation process.

Our results demonstrating that our differentiation signature identified from a Boolean network derived from neuroblastoma tumor gene expression profiles sorted along a stemness-differentiation continuum based on *MYCN* expression is associated with neuroblastoma cell and tumor differentiation in multiple independent datasets further validate our novel approach to identifying critical genes and pathways in the neuroblastoma differentiation process. Our analyses have identified a pathway of 4 gene clusters which predicted neuroblastoma differentiation in neuroblastoma cell lines and patient tumors, and which also predicted neuroblastoma patient outcomes in a series of diverse patient cohorts. We identified a cluster of genes involved in growth factor receptor trafficking and intracellular signaling that was strongly associated with differentiation, and we identified novel candidate genes associated with *MYCN* expression and neuroblastoma differentiation via asymmetric relationships. Our findings therefore demonstrate that Boolean approaches

for gene expression analysis can identify novel pathways relevant for neuroblastoma tumor differentiation that could represent potential therapeutic targets.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

RNA-sequencing for this project was supported in part by the Genomic and RNA Profiling Core (G.A.R.P.) at Baylor College of Medicine with funding from an NIH NCI grant (P30CA125123) and with the assistance of the Core Director Lisa D. White, PhD.

Funding:

Hyundai Hope on Wheels Hope Grant to P.E.Z. and D.S.; UCSD Pediatrics Department Pilot Grant to P.E.Z. and D.S.; NIGMS grant GM138385 to D.S.; NIH grants AI141630, CA238042, CA100768, and CA160911 to P.G. AI155696, UG3TR003355 and UG3TR002968 (to D.S and P.G). Funding sources had no role in study design; data collection, analysis, or interpretation; or in writing or submitting the manuscript.

Data Sharing Statement:

In agreement with the NIH public access policy of open sharing of scientific resources generated with the support of public funding, we will make all reagents and discoveries available free of charge to the scientific community at large upon publication of the findings, and any new software will be made available online for use. The data generated by our studies will be presented on a continual basis at international, national, and local scientific meetings and published in peer-reviewed journals in a timely manner. RNA-sequencing data will be deposited to and made publicly and freely available through Gene Expression Omnibus (GEO) and Sequence Read Archive (SRA) no later than the acceptance date of initial publication. Deposited data will be referenced by repository accession numbers in manuscripts and reports.

References

1. Whittle SB, Smith V, Doherty E, Zhao S, McCarty S, Zage PE. Overview and Recent Advances in the Treatment of Neuroblastoma. *Exp Rev Anticancer Ther* 2017; 17: 369–86.
2. Lau L, Tai D, Weitzman S, Grant R, Baruchel S, Malkin D. Factors influencing survival in children with recurrent neuroblastoma. *J Pediatr Hematol Oncol* 2004; 26: 227–32. [PubMed: 15087949]
3. London WB, Castel V, Monclair T, Ambros PF, Pearson ADJ, Cohn SL, et al. Clinical and Biologic Features Predictive of Survival After Relapse of Neuroblastoma: A Report From the International Neuroblastoma Risk Group Project. *J Clin Oncol* 2011; 29: 3286–92. [PubMed: 21768459]
4. London WB, Bagatell R, Weigel BJ, Fox E, Guo D, Van Ryn C, Naranjo A, Park JR. Historical Time to Disease Progression and Progression-Free Survival in Patients With Recurrent/Refractory Neuroblastoma Treated in the Modern Era on Children's Oncology Group Early-Phase Trials. *Cancer* 2017; 123: 4914–23. [PubMed: 28885700]
5. Zage PE. Novel Therapies for Relapsed and Refractory Neuroblastoma. *Children* 2018; 5(11).
6. Shimada H, Ambros IM, Dehner LP, Hata J-I, Joshi VV, Roald B, et al. The International Neuroblastoma Pathology Classification (the Shimada System). *Cancer* 1999; 86: 364–72. [PubMed: 10421273]
7. Mohlin SA, Wigerup C, Pahlman S. Neuroblastoma aggressiveness in relation to sympathetic neuronal differentiation stage. *Semin Cancer Biology* 2011; 21: 276–282.

8. Fredlund E, Ringner M, Maris JM, Pahlman S. High MYC pathway activity and low stage of neuronal differentiation associate with poor outcome in neuroblastoma. *Proc Natl Acad Sci USA* 2008; 105: 14094–14099. [PubMed: 18780787]
9. Kimura S, Sekiguchi M, Watanabe K, Hiwatarai M, Seki M, Yoshida K, et al. Association of high-risk neuroblastoma classification based on expression profiles with differentiation and metabolism. *PLoS ONE* 2021; 16: e0245526. [PubMed: 33465163]
10. Akavia UD, Litvin O, Kim J, Sanchez-Garcia F, Kotliar D, Causton HC, Pochanard P, Mozes E, Garraway LA, Pe'er D. An integrated approach to uncover drivers of cancer. *Cell* 2010; 143: 1005–1017. [PubMed: 21129771]
11. Hoadley KA, Yau C, Wolf DM, Cherniack AD, Tamborero D, Ng S, et al. Multiplatform analysis of 12 cancer types reveals molecular classification within and across tissues of origin. *Cell* 2014; 158: 929–944. [PubMed: 25109877]
12. Sonabend AM, Bansal M, Guarnieri P, Lei L, Amendolara B, Soderquist C, et al. The transcriptional regulatory network of proneural glioma determines the genetic alterations selected during tumor progression. *Cancer Res* 2014; 74: 1440–1451. [PubMed: 24390738]
13. Carro MS, Lim WK, Alvarez MJ, Bollo RJ, Zhao X, Snyder EY, et al. The transcriptional network for mesenchymal transformation of brain tumours. *Nature* 2010; 463: 318–325. [PubMed: 20032975]
14. Allocco DJ, Kohane IS, Butte AJ. Quantifying the relationship between co-expression, co-regulation and gene function. *BMC Bioinformatics* 2004; 5: 18. [PubMed: 15053845]
15. Jordan IK, Marino-Ramirez L, Wolf YI, Koonin EV. Conservation and coevolution in the scale-free human gene coexpression network. *Mol Biol Evol* 2004; 21: 2058–2070. [PubMed: 15282333]
16. Lee HK, Hsu AK, Sajdak J, Qin J, Pavlidis P. Coexpression analysis of human genes across many microarray data sets. *Genome Res* 2004; 14: 1085–1094. [PubMed: 15173114]
17. Tavazoie S, Hughes JD, Campbell MJ, Cho RJ, Church GM. Systematic determination of genetic network architecture. *Nat Genet* 1999; 22: 281–285. [PubMed: 10391217]
18. Butte AJ, Tamayo P, Slonim D, Golub TR, Kohane IS. Discovering functional relationships between RNA expression and chemotherapeutic susceptibility using relevance networks. *Proc Natl Acad Sci USA* 2000; 97: 12182–12186. [PubMed: 11027309]
19. Paik S, Shak S, Tang G, Kim C, Baker J, Cronin M, et al. A multigene assay to predict recurrence of tamoxifen-treated, node-negative breast cancer. *N Engl J Med* 2004; 351: 2817–2826. [PubMed: 15591335]
20. Witten DM, Tibshirani R. Survival analysis with high-dimensional covariates. *Stat Methods Med Res* 2010; 19: 29–51. [PubMed: 19654171]
21. Alizadeh AA, Eisen MB, Davis RE, Ma C, Lossos IS, Rosenwald A, et al. Distinct types of diffuse large B-cell lymphoma identified by gene expression profiling. *Nature* 2000; 403: 503–511. [PubMed: 10676951]
22. Zhao H, Ljungberg B, Grankvist K, Rasmuson T, Tibshirani R, Brooks JD. Gene expression profiling predicts survival in conventional renal cell carcinoma. *PLoS Med* 2006; 3: e13. [PubMed: 16318415]
23. Sahoo D, Dill DL, Gentles AJ, Tibshirani R, Plevritis SK. Boolean implication networks derived from large scale, whole genome microarray datasets. *Genome Biol* 2008; 9: R157 [PubMed: 18973690]
24. Sahoo D, Dill DL, Tibshirani R, Plevritis SK. Extracting binary signals from microarray time-course data. *Nucleic Acids Res* 2007; 35: 3705–12. [PubMed: 17517782]
25. Sahoo D, Seita J, Bhattacharya D, Inlay MA, Weissman IL, Plevritis SK, Dill DL. MiDReG: a method of mining developmentally regulated genes using Boolean implications. *Proc Natl Acad Sci USA* 2010; 107: 5732–7. [PubMed: 20231483]
26. Cheah MT, Chen JY, Sahoo D, Contreras-Trujillo, Volkmer AK, Scheeren FA, et al. CD14-expressing cancer cells establish the inflammatory and proliferative tumor microenvironment in bladder cancer. *Proc Natl Acad Sci USA* 2015; 112: 4725–30. [PubMed: 25825750]
27. Shin K, Lim A, Zhao C, Sahoo D, Pan Y, Spiekerkoetter E, et al. Hedgehog signaling restrains bladder cancer progression by eliciting stromal production of urothelial differentiation factors. *Cancer Cell* 2014; 26: 521–33. [PubMed: 25314078]

28. Dalerba P, Kalisky T, Sahoo D, Rajendran PS, Rothenberg ME, Leyrat AA, et al. Single-cell dissection of transcriptional heterogeneity in human colon tumors. *Nat Biotechnol* 2011; 29: 1120–7. [PubMed: 22081019]
29. Dalerba P, Sahoo D, Paik S, Guo X, Yothers G, Song N, et al. CDX2 as a Prognostic Biomarker in Stage II and Stage III Colon Cancer. *N Engl J Med* 2016; 374: 211–22 [PubMed: 26789870]
30. Inlay MA, Bhattacharya D, Sahoo D, Serwold T, Seita J, Karsunky H, et al. Ly6d marks the earliest stage of B-cell specification and identifies the branchpoint between B-cell and T-cell development. *Genes Dev* 2009; 23: 2376–81. [PubMed: 19833765]
31. Volkmer JP, Sahoo D, Chin RK, Ho PL, Tang C, Kurtova AV, et al. Three differentiation states risk-stratify bladder cancer into distinct subtypes. *Proc Natl Acad Sci USA* 2012; 109: 2078–83. [PubMed: 22308455]
32. Sahoo D, Swanson L, Sayed IM, Katkar GD, Ibeawuchi SR, Mittal Y, Pranadinata RF, Tindle C, Fuller M, Stec DL, Chang JT, Sandborn WJ, Das S, Ghosh P. Artificial intelligence guided discovery of a barrier-protective therapy in inflammatory bowel disease. *Nat Commun* 2021; 12: 4246. [PubMed: 34253728]
33. Sahoo D, Katkar GD, Khandelwal S, Behroozikhah M, Claire A, Castillo V, Tindle C, Fuller M, Taheri S, Rogers TF, Beutler N, Ramirez SI, Rawlings SA, Pretorius V, Smith DM, Burton DR, Alexander LEC, Duran J, Crotty S, Dan JM, Das S, Ghosh P. AI-guided discovery of the invariant host response to viral pandemics. *EBioMedicine* 2021; 68: 103390. [PubMed: 34127431]
34. Barrett T, Suzek TO, Troup DB, Wilhite SE, Ngau WC, Ledoux P, Rudnev D, Lash AE, Fujibuchi W, Edgar R. NCBI GEO: mining millions of expression profiles--database and tools. *Nucleic Acids Res* 2005; 33: D562–6. [PubMed: 15608262]
35. Barrett T, Wilhite SE, Ledoux P, Evangelista C, Kim IF, Tomashevsky M, Marshall KA, Phillippy KH, Sherman PM, Holko M, Yefanov A, Lee H, Zhang N, Robertson CL, Serova N, Davis S, Soboleva A. NCBI GEO: archive for functional genomics data sets--update. *Nucleic Acids Res* 2013; 41: D991–5. [PubMed: 23193258]
36. Brazma A, Parkinson H, Sarkans U, Shojatalab M, Vilo J, Abeygunawardena N, Holloway E, Kapushesky M, Kemmeren P, Lara GG, et al. ArrayExpress—a public repository for microarray gene expression data at the EBI. *Nucleic Acids Res* 2003; 31: 68–71. [PubMed: 12519949]
37. Parkinson H, Kapushesky M, Kolesnikov N, Rustici G, Shojatalab M, Abeygunawardena N, Berube H, Dylag M, Emam I, Farne A, Holloway E, Lukk M, Malone J, Mani R, Pilicheva E, Rayner TF, Rezwan F, Sharma A, Williams E, Bradley XZ, et al. ArrayExpress update--from an archive of functional genomics experiments to the atlas of gene expression. *Nucleic Acids Res* 2008; 37: D868–72. [PubMed: 19015125]
38. Zhang W, Yu Y, Hertwig F, Thierry-Mieg J, Zhang W, Thierry-Mieg D, et al. Comparison of RNA-seq and Microarray-Based Models of Clinical Endpoint Prediction. *Genome Biol* 2015; 16: 133. [PubMed: 26109056]
39. Rosswog C, Schmidt R, Oberthuer A, Juraeva D, Brors B, Engesser A, et al. Molecular Classification Substitutes for the Prognostic Variables Stage, Age, and MYCN Status in Neuroblastoma Risk Assessment. *Neoplasia* 2017; 19: 982–90. [PubMed: 29091799]
40. Oberthuer A, Berthold F, Warnat P, Hero B, Kahlert Y, Spitz R, et al. Customized Oligonucleotide Microarray Gene Expression–Based Classification of Neuroblastoma Patients Outperforms Current Clinical Risk Stratification. *J Clin Oncol* 2006; 24: 5070–8. [PubMed: 17075126]
41. Oberthuer A, Juraeva D, Li L, Kahlert Y, Westermann F, Eils R, et al. Comparison of performance of one-color and two-color gene-expression analyses in predicting clinical endpoints of neuroblastoma patients. *Pharmacogenomics J* 2010; 10: 258–66. [PubMed: 20676065]
42. Rajbhandari P, Lopez G, Capdevila C, Salvatori B, Yu J, Rodriguez-Barrueco R, et al. Cross-Cohort Analysis Identifies a TEAD4-MYCN Positive Feedback Loop as the Core Regulatory Element of High-Risk Neuroblastoma. *Cancer Discov* 2018; 8: 582–599. [PubMed: 29510988]
43. Ohtaki M, Otani K, Hiyama K, Kamei N, Satoh K, Hiyama E. A robust method for estimating gene expression states using Affymetrix microarray probe level data. *BMC Bioinformatics* 2010; 11: 183. [PubMed: 20380745]

44. Westerlund I, Shi Y, Toskas K, Fell SM, Li S, Surova O, et al. Combined epigenetic and differentiation-based treatment inhibits neuroblastoma tumor growth and links HIF2 α to tumor suppression. *Proc Natl Acad Sci USA* 2017; 114: E6137–E6146. [PubMed: 28696319]
45. Frumm SM, Fan ZP, Ross KN, Duvall JR, Gupta S, VerPlank L, et al. Selective HDAC1/HDAC2 inhibitors induce neuroblastoma differentiation. *Chem Biol* 2013; 20: 713–25. [PubMed: 23706636]
46. Pezzini F, Bettinetti L, Di Leva F, Bianchi M, Zoratti E, Carrozzo R, et al. Transcriptomic Profiling Discloses Molecular and Cellular Events Related to Neuronal Differentiation in SH-SY5Y Neuroblastoma Cells. *Cell Mol Neurobiol* 2017; 37: 665–682. [PubMed: 27422411]
47. Claeys S, Denecker G, Durinck K, Decaestecker B, Mus LM, Loontjens S, et al. ALK positively regulates MYCN activity through repression of HBP1 expression. *Oncogene* 2019; 38: 2690–2705. [PubMed: 30538293]
48. Nishida Y, Adati N, Ozawa R, Maeda A, Sakaki Y, Takeda T. Identification and classification of genes regulated by phosphatidylinositol 3-kinase- and TRKB-mediated signalling pathways during neuronal differentiation in two subtypes of the human neuroblastoma cell line SH-SY5Y. *BMC Res Notes* 2008; 1: 95. [PubMed: 18957096]
49. Taschner-Mandl S, Schwarz M, Blaha J, Kauer M, Kromp F, Frank N, et al. Metronomic topotecan impedes tumor growth of MYCN-amplified neuroblastoma cells in vitro and in vivo by therapy induced senescence. *Oncotarget* 2016; 7: 3571–86. [PubMed: 26657295]
50. Jansky S, Sharma AK, Körber V, Quintero A, Toprak UH, Wecht EM, et al. Single-cell transcriptomic analyses provide insights into the developmental origins of neuroblastoma. *Nat Genet* 2021; 53: 683–693. [PubMed: 33767450]
51. Ikram F, Ackermann S, Kahlert Y, Volland R, Roels F, Engesser A, et al. Transcription factor activating protein 2 beta (TFAP2B) mediates noradrenergic neuronal differentiation in neuroblastoma. *Mol Oncol* 2016; 10: 344–59. [PubMed: 26598443]
52. Dong R, Yang R, Zhan Y, Lai HD, Ye CJ, Yao XY, et al. Single-Cell Characterization of Malignant Phenotypes and Developmental Trajectories of Adrenal Neuroblastoma. *Cancer Cell* 2020; 38: 716–733. [PubMed: 32946775]
53. Kildisiute G, Kholosy WM, Young MD, Roberts K, Elmentaite R, van Hooff SR, et al. Tumor to normal single-cell mRNA comparisons reveal a pan-neuroblastoma cancer cell. *Sci Adv* 2021; 7: eabd3311. [PubMed: 33547074]
54. Pugh TJ, Morozova O, Attiyeh EF, Asgharzadeh S, Wei JS, Auclair D, et al. The genetic landscape of high-risk neuroblastoma. *Nat Genet* 2013; 45: 279–84. [PubMed: 23334666]
55. Egolf LE, Vaksman Z, Lopez G, Rokita JL, Modi A, Basta PV, et al. Germline 16p11.2 Microdeletion Predisposes to Neuroblastoma. *Am J Hum Genet* 2019; 105: 658–668. [PubMed: 31474320]
56. Edgar R, Domrachev M, Lash AE. Gene Expression Omnibus: NCBI gene expression and hybridization array data repository. *Nucleic Acids Res* 2002; 30: 207–10. [PubMed: 11752295]
57. Irizarry RA, Bolstad BM, Collin F, Cope LM, Hobbs B, Speed TP. Summaries of Affymetrix GeneChip probe level data. *Nucleic Acids Res* 2003; 31: e15. [PubMed: 12582260]
58. Irizarry RA, Hobbs B, Collin F, Beazer-Barclay YD, Antonellis KJ, Scherf U, Speed TP. Exploration, normalization, and summaries of high density oligonucleotide array probe level data. *Biostatistics* 2003; 4: 249–64. [PubMed: 12925520]
59. Li B, Dewey CN. RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. *BMC Bioinformatics* 2011; 12: 323. [PubMed: 21816040]
60. Sahoo D The power of boolean implication networks. *Front Physiol* 2012; 3: 276. [PubMed: 22934030]
61. Curtis SA. The Classification of Greedy Algorithms. *Sci Comput Programming* 2003; 49: 125–57.
62. Fabregat A, Jupe S, Matthews L, Sidiropoulos K, Gillespie M, Garapati P, et al. The Reactome Pathway Knowledgebase. *Nucleic Acids Res* 2018; 46: D649–D655. [PubMed: 29145629]
63. van Groningen T, Koster J, Valentijn LJ, Zwijnenburg DA, Akogul N, Hasselt NE, et al. Neuroblastoma is composed of two super-enhancer-associated differentiation states. *Nat Genet* 2017; 49: 1261–6. [PubMed: 28650485]

64. Kanehisa M, Goto S. KEGG: Kyoto encyclopedia of genes and genomes. *Nucleic Acids Res* 2000; 28: 27–30. [PubMed: 10592173]
65. Fridman AL, Tainsky MA. Critical pathways in cellular senescence and immortalization revealed by gene expression profiling. *Oncogene* 2008; 27: 5975–87. [PubMed: 18711403]
66. Liberzon A, Birger C, Thorvaldsdottir H, Ghandi M, Mesirov JP, Tamayo P. The Molecular Signatures Database (MSigDB) Hallmark Gene Set Collection. *Cell Syst* 2015; 1: 417–25 [PubMed: 26771021]
67. Flynn S, Lesperance J, Macias A, Phanthilath N, Paul MR, Kim JW, Tamayo P, Zage PE. The Multikinase Inhibitor RXDX-105 is Effective Against Neuroblastoma In Vitro and In Vivo. *Oncotarget* 2019; 10: 6323–33. [PubMed: 31695841]
68. Subramonian D, Phanthilath N, Huo Y, Flynn S, Rinehardt H, Zhang J, Messer K, Mo Q, Huang S, Lesperance J, Zage PE. Regorafenib Is Effective Against Neuroblastoma in Vitro and in Vivo and Inhibits the RAS-MAPK, PI3K/Akt/mTOR, and Fos/Jun Pathways. *Br J Cancer* 2020; 123: 568–579. [PubMed: 32457362]
69. Woodfield SE, Guo R, Liu Y, Major AM, Hollingsworth EF, Indiviglio S, Whittle SB, Mo Q, Bean AJ, Ittmann M, Lopez-Terrada D, Zage PE. Neuroblastoma patient outcomes, tumor differentiation, and ERK activation are correlated with expression levels of the ubiquitin ligase UBE4B. *Genes Cancer* 2016; 7: 13–26. [PubMed: 27014418]
70. Trapnell C, Pachter L, Salzberg SL. TopHat: discovering splice junctions with RNA-Seq. *Bioinformatics* 2009; 25: 1105–11. [PubMed: 19289445]
71. Brodeur GM, Seeger RC, Schwab M, Varmus HE, Bishop JM. Amplification of N-myc in Untreated Human Neuroblastomas Correlates with Advanced Disease Stage. *Science* 1984; 224: 1121–4. [PubMed: 6719137]
72. Seeger RC, Brodeur GM, Sather H, Dalton A, Siegel SE, Wong KY, Hammond D. Association of Multiple Copies of the N-myc Oncogene with Rapid Progression of Neuroblastomas. *N Engl J Med* 1985; 313: 1111–6. [PubMed: 4047115]
73. Varlakhanova NV, Cotterman RF, deVries WN, Morgan J, Donahue LR, Murray S, Knowles BB, Knoepfler PS. Myc maintains embryonic stem cell pluripotency and self-renewal. *Differentiation* 2010; 80: 9–19. [PubMed: 20537458]
74. Westermark UK, Wilhelm M, Frenzel A, Henriksson MA. The MYCN oncogene and differentiation in neuroblastoma. *Semin Cancer Biol* 2011; 21: 256–66. [PubMed: 21849159]
75. Zhang JT, Weng ZH, Tsang KS, Tsang LL, Chan HC, Jiang XH. MycN Is Critical for the Maintenance of Human Embryonic Stem Cell-Derived Neural Crest Stem Cells. *PLoS One* 2016; 11: e0148062. [PubMed: 26815535]
76. Thiele CJ, Reynolds CP, Israel MA. Decreased expression of N-myc precedes retinoic acid-induced morphological differentiation of human neuroblastoma. *Nature* 1985; 313: 404–406. [PubMed: 3855502]
77. Reynolds CP, Kane DJ, Einhorn PA, Matthy KK, Crouse VL, Wilbur JR, et al. Response of neuroblastoma to retinoic acid in vitro and in vivo. *Prog Clin Biol Res* 1991; 366: 203–211. [PubMed: 2068138]
78. Reynolds CP, Matthy KK, Villablanca JG, Maurer BJ. Retinoid therapy of high-risk neuroblastoma. *Cancer Lett* 2003; 197: 185–192. [PubMed: 12880980]
79. Wei SJ, Nguyen TH, Yang IH, Mook DG, Makena MR, Verlekar D, et al. MYC transcription activation mediated by OCT4 as a mechanism of resistance to 13-cisRA-mediated differentiation in neuroblastoma. *Cell Death Dis* 2020; 11: 368. [PubMed: 32409685]
80. Thiele CJ. Neuroblastoma: In (Ed.) *Masters J Human Cell Culture*. Lancaster, UK: Kluwer Academic Publishers. 1998; 1: 21–53.
81. Petrov I, Suntsova M, Ilnitskaya E, Roumiantsev S, Sorokin M, Garazha A, et al. Gene Expression and Molecular Pathway Activation Signatures of MYCN-amplified Neuroblastomas. *Oncotarget* 2017; 8: 83768–80. [PubMed: 29137381]
82. Savary C, Kim A, Lespagnol A, Gandemer V, Pellier I, Andrieu C, et al. Depicting the genetic architecture of pediatric cancers through an integrative gene network approach. *Sci Rep* 2020; 10: 1224. [PubMed: 31988326]

83. Zage PE, Sirisaengtaksin N, Liu Y, Gireud M, Brown BS, Palla S, et al. UBE4B Levels are Correlated with Clinical Outcomes in Neuroblastoma Patients and with Altered Neuroblastoma Cell Proliferation and Sensitivity to EGFR Inhibitors. *Cancer* 2013; 119: 915–23. [PubMed: 22990745]
84. Sirisaengtaksin N, Gireud M, Yan Q, Kubota Y, Meza D, Waymire J, et al. UBE4B Couples Ubiquitination and Sorting Machineries to Enable EGFR Degradation. *J Biol Chem* 2014; 289: 3026–39. [PubMed: 24344129]
85. Hill DP, Robertson KA. Characterization of the cholinergic neuronal differentiation of the human neuroblastoma cell line LA-N-5 after treatment with retinoic acid. *Dev Brain Res* 1997; 102: 53–67. [PubMed: 9298234]
86. Edsjo A, Lavenius E, Nilsson H, Hoehner JC, Simonsson P, Culp LA, et al. Expression of *trkB* in Human Neuroblastoma in Relation to *MYCN* Expression and Retinoic Acid Treatment. *Lab Invest* 2003; 83: 813–23. [PubMed: 12808116]
87. Jiang M, Stanke J, Lahti JM. The connections between neural crest development and neuroblastoma. *Curr Top Dev Biol* 2011; 94: 77–127. [PubMed: 21295685]
88. Zage PE, Whittle SB, Shohet JM. CD114: A New Member of the Neural Crest-Derived Cancer Stem Cell Marker Family. *J Cell Biochem* 2017; 118: 221–231. [PubMed: 27428599]
89. Tomolonis JA, Agarwal S, Shohet JM. Neuroblastoma pathogenesis: deregulation of embryonic neural crest development. *Cell Tissue Res* 2018; 372: 245–262. [PubMed: 29222693]
90. Tas ML, Nagetegaal M, Kraal KCJM, Tytgat GAM, Abeling NGGM, Koster J, et al. Neuroblastoma Stage 4S: Tumor Regression Rate and Risk Factors of Progressive Disease. *Pediatr Blood Cancer* 2020; 67: e28061. [PubMed: 31736229]
91. Zimmerman MW, Liu Y, He S, Durbin AD, Abraham BJ, Easton J, Shao Y, Xu B, Zhu S, Zhang X, Li Z, Weichert-Leahey N, Young RA, Zhang J, Look AT. *MYC* Drives a Subset of High-Risk Pediatric Neuroblastomas and Is Activated through Mechanisms Including Enhancer Hijacking and Focal Enhancer Amplification. *Cancer Discov* 2018; 8: 320–335. [PubMed: 29284669]
92. de Ruijter AJ, Meinsma RJ, Bosma P, Kemp S, Caron HN, van Kuilenburg AB. Gene expression profiling in response to the histone deacetylase inhibitor BL1521 in neuroblastoma. *Exp Cell Res* 2005; 309: 451–67. [PubMed: 16084510]
93. Hahn CK, Ross KN, Warrington IM, Mazitschek R, Kanegai CM, Wright RD, et al. Expression-based screening identifies the combination of histone deacetylase inhibitors and retinoids for neuroblastoma differentiation. *Proc Natl Acad Sci USA* 2008; 105: 9751–6. [PubMed: 18607002]
94. Politis PK, Akrivou S, Hurel C, Papadodima O, Matsas R. BM88/Cend1 is involved in histone deacetylase inhibition-mediated growth arrest and differentiation of neuroblastoma cells. *FEBS Lett* 2008; 582: 741–8. [PubMed: 18258204]
95. Koegl M, Hoppe T, Schlenker S, Ulrich HD, Mayer TU, Jentsch S. A Novel Ubiquitination Factor, E4, Is Involved in Multiubiquitin Chain Assembly. *Cell* 1999; 96: 635–44. [PubMed: 10089879]
96. Hishiki T, Nimura Y, Isogai E, et al. Glial Cell Line-derived Neurotrophic Factor/Neurturin-induced Differentiation and Its Enhancement by Retinoic Acid in Primary Human Neuroblastomas Expressing *c-Ret*, *GFRa-1*, and *GFRa-2*. *Cancer Res* 1998 58: 2158–65. [PubMed: 9605760]
97. Oppenheimer O, Cheung N-K, Gerald WL. The *RET* Oncogene is a Critical Component of Transcriptional Programs Associated with Retinoic Acid-induced Differentiation in Neuroblastoma. *Mol Cancer Ther* 2007; 6: 1300–9. [PubMed: 17431108]
98. Bunone G, Borrello MG, Picetti R, et al. Induction of *RET* Proto-Oncogene Expression in Neuroblastoma Cells Precedes Neuronal Differentiation and Is Not Mediated by Protein Synthesis. *Exp Cell Res* 1995; 217: 92–9. [PubMed: 7867726]
99. D'Aleesio A, De Vita G, Cali G, et al. Expression of the *RET* Oncogene Induces Differentiation of SK-N-BE Neuroblastoma Cells. *Cell Growth Diff* 1995; 6: 1387–94. [PubMed: 8562477]
100. Zage PE, Zeng L, Palla S, et al. A Novel Therapeutic Combination for Neuroblastoma: the VEGFR/EGFR/*RET* Inhibitor Vandetanib with 13-*cis*-Retinoic Acid. *Cancer* 2010; 116: 2465–75. [PubMed: 20225331]
101. Kaplan DR, Matsumoto K, Lucarelli E, Thiele CJ. Induction of *TrkB* by Retinoic Acid Mediates Biologic Responsiveness to BDNF and Differentiation of Human Neuroblastoma Cells. *Neuron* 1993; 11: 321–31. [PubMed: 8394722]

102. Matsushima H, Bogenmann E. Expression of trkA cDNA in neuroblastomas mediates differentiation in vitro and in vivo. *Mol Cell Biol* 1993; 13: 7447–56. [PubMed: 8246962]
103. Eggert A, Ikegaki N, Liu X-G, et al. Molecular dissection of TrkA signal transduction pathways mediating differentiation in human neuroblastoma cells. *Oncogene* 2000; 19: 2043–51. [PubMed: 10803465]
104. Nakagawara A, Azar CG, Scavarda NJ, Brodeur GM. Expression and function of TRK-B and BDNF in Human Neuroblastomas. *Mol Cell Biol* 1994; 14: 759–67. [PubMed: 8264643]
105. Guo L, Lin W, Zhang Y, Wang J. A Systematic Analysis Revealed the Potential Gene Regulatory Processes of ATRA-Triggered Neuroblastoma Differentiation and Identified a Novel RA Response Sequence in the NTRK2 Gene. *Biomed Res Int* 2020; 2020: 6734048. [PubMed: 32149119]
106. Esposito CL, D'Alessio A, de Franciscis V, Cerchia L. A Cross-Talk between TrkB and Ret Tyrosine Kinases Receptors Mediates Neuroblastoma Cell Differentiation. *PLoS ONE* 2008; 3: e1643. [PubMed: 18286198]
107. Peterson S, Bogenmann E. The RET and TRKA Pathways Collaborate to Regulate Neuroblastoma Differentiation. *Oncogene* 2004; 23: 213–225. [PubMed: 14712226]
108. Gaviglio AL, Knelson EH, Blobe GC. Heparin-binding Epidermal Growth Factor-like Growth Factor Promotes Neuroblastoma Differentiation. *FASEB J* 2017; 31: 1903–15. [PubMed: 28174207]
109. Dabydeen SA, Desai A, Sahoo D. Unbiased Boolean Analysis of Public Gene Expression Data for Cell Cycle Gene Identification. *Mol Biol Cell* 2019; 30: 1770–9. [PubMed: 31091168]
110. Pandey S & Sahoo D. Identification of Gene Expression Logical Invariants in Arabidopsis. *Plant Direct* 2019; 3: e00123. [PubMed: 31245766]

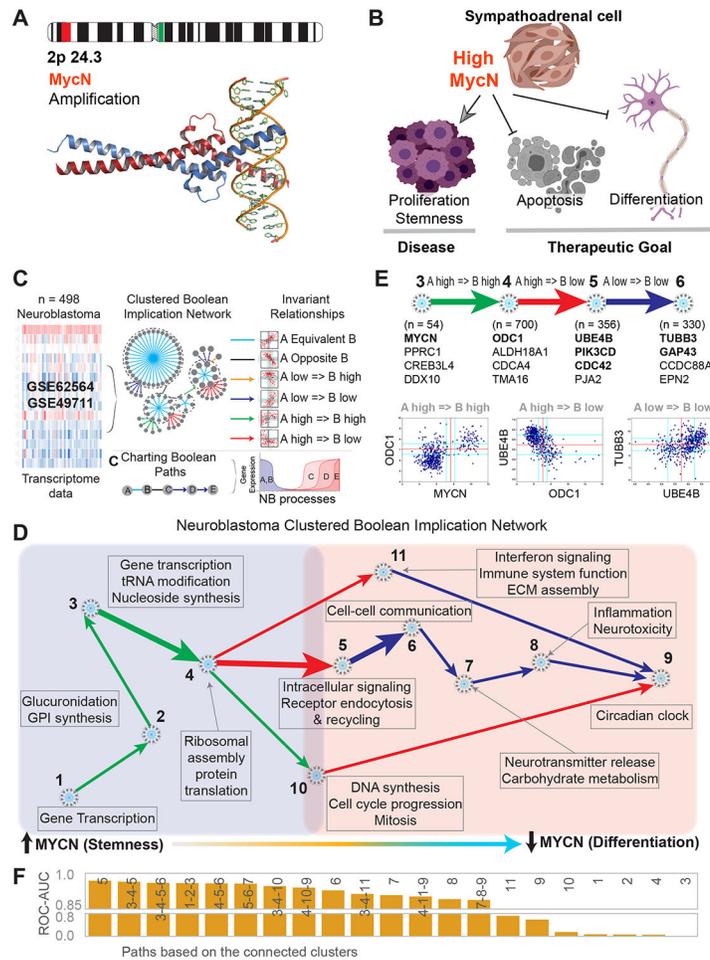


Figure 1. A Boolean Network Model of the Stemness-Differentiation Continuum in Neuroblastoma Tumors

(A,B) Schematics summarizing the significance of *MYCN* amplification in neuroblastoma tumors (A) and its association with disease phenotypes and its antagonistic impact on the desired therapeutic goals (B).

(C) Boolean network analysis identifies clusters of genes that are enriched in a continuum of cellular states. The Boolean network utilizes the six possible invariant Boolean relationships between genes, where the two symmetric relationships each have two diagonally opposite sparse quadrants, while the four asymmetric Boolean relationships each have one sparse quadrant in their gene expression profile. Genes with similar expression profiles are organized into clusters, and their relationships serve as the basis for a directed Boolean Network with nodes as genes and edges corresponding to the Boolean relationships (23,24). Boolean cluster relationships are used to chart individual gene expression changes along a Boolean path within the stemness-differentiation continuum using Boolean Network Explorer (BoNE) software, which clusters genes using Boolean equivalent relationships, while the edges between clusters are defined using the identified implication relationships. (D) Boolean Network Explorer (BoNE) software was used to analyze RNA expression profiles of 498 neuroblastoma patient tumors (38; GSE62564/GSE49711). The path direction was derived from the connections established from high *MYCN* expression

(representing undifferentiated neuroblastoma) to low *MYCN* expression (representing differentiated neuroblastoma), with the arrow colors reflecting the implication relationships. Reactome pathway analyses (62) of each of the clusters in the continuum between undifferentiated and differentiated neuroblastoma were performed, with significant pathways for each cluster listed.

(E) The Boolean relationships and candidate genes within each gene cluster in the identified cluster 3-4-5-6 pathway are shown. Representative Boolean gene expression relationships are shown.

(F) Machine learning and model training was performed by ranking Boolean cluster pathways including up to 4 connected clusters based on their performance (ROC-AUC values) in distinguishing neuroblastoma differentiation as defined by *MYCN* expression.

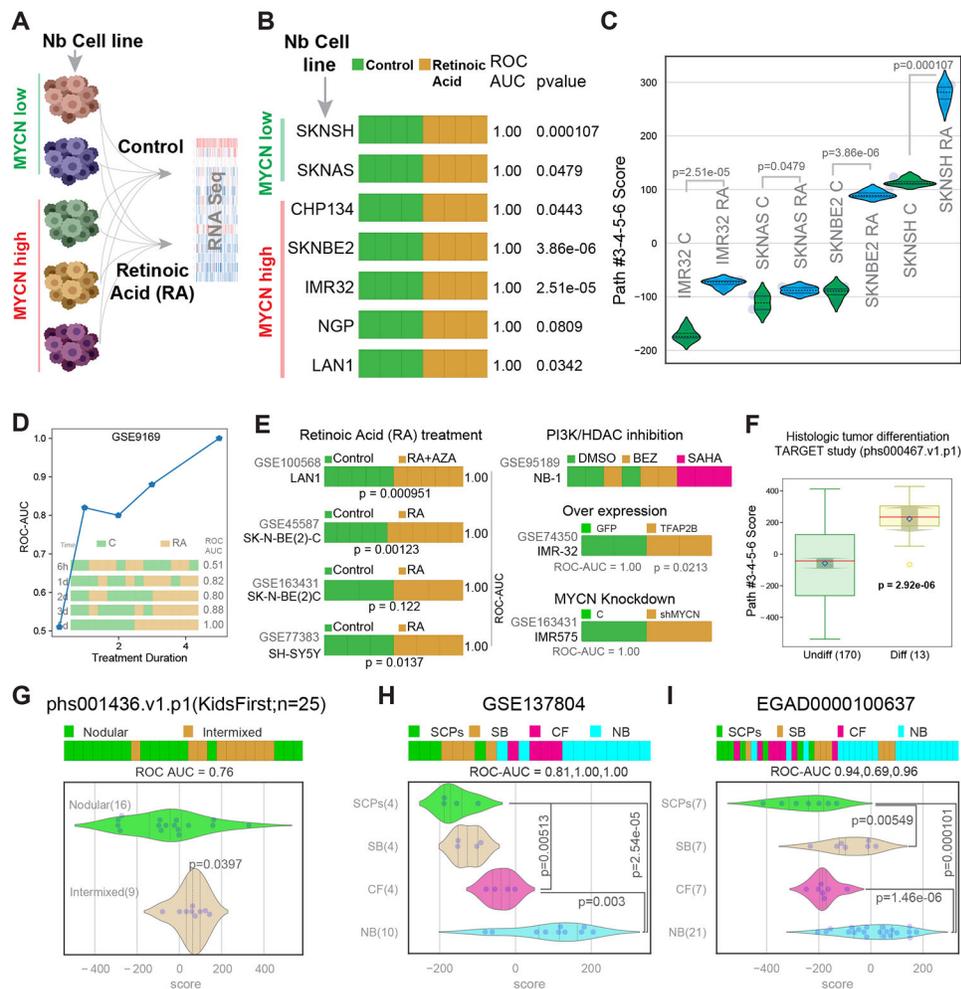


Figure 2. Validation of the Boolean Model of Neuroblastoma Tumor Differentiation.

BoNE software identified a pathway that included gene clusters #3-4-5-6 as the pathway most significantly associated with neuroblastoma differentiation, which we have termed a differentiation signature. The expression of genes within these clusters was converted to a differentiation signature score by using linear combinations of gene expression values for individual genes in each cluster.

(A) Neuroblastoma cell lines with both low and high *MYCN* gene expression levels before and after treatment with the differentiating agent 13-*cis*-retinoic acid (CRA) were analyzed by RNA-sequencing for gene expression profiles.

(B) Sample ordering based on the differentiation signature score using BoNE software of gene expression profiles of untreated control neuroblastoma cell lines and neuroblastoma cell lines after 13-*cis*-retinoic acid treatment (ROC-AUC=1.00 for each). P-values were computed using two-tailed, two sample Welch's T-tests.

(C) Sample ordering based on the differentiation signature using BoNE software of gene expression profiles of control, untreated, neuroblastoma cell lines (C) and neuroblastoma cell lines after 13-*cis*-retinoic acid treatment (RA), with p-values comparing scores in RA treated and untreated cell lines shown.

(D) Time course analysis of differentiation induced by treatment of SK-N-SH, SH-SY5Y-A cells (ATCC: CRL-2266), and SH-SY5Y-E cells (ECACC: EC94030304) (48; GSE9169) with 10 μ M all-trans-retinoic acid (RA). ROC-AUC of the untreated compared to treated samples is plotted against the treatment duration.

(E) (Left) Sample ordering based on the differentiation signature score generated from BoNE software of gene expression profiles of undifferentiated LAN1 neuroblastoma cells and LAN1 cells differentiated with the combination of 10 μ M all-trans-retinoic acid (ATRA) and 5-aza-deoxycytidine (RA+AZA; 44; GSE100568; ROC-AUC=1), of undifferentiated SK-N-BE(2)C neuroblastoma cells and SK-N-BE(2)C cells differentiated with 10 μ M ATRA (RA; 45,50; GSE45587, GSE163431; ROC-AUC=1.00), and of undifferentiated SH-SY5Y cells and SH-SY5Y cells differentiated with 10 μ M ATRA and neurobasal medium (RA; 46; GSE77383; ROC-AUC=1.00). (Right) Sample ordering based on the differentiation signature score of gene expression profiles of untreated control NB-1 neuroblastoma cells and NB-1 cells treated with the PI3K/mTOR inhibitor BEZ-235 (BEZ) and with the histone deacetylase inhibitor vorinostat (suberoylanilide hydroxamic acid/SAHA) (47; GSE95189; ROC-AUC of the cells treated with BEZ-235 compared to control = 0.89; ROC-AUC of the cells treated with SAHA compared to control = 1.0), and sample ordering based on the differentiation signature score of gene expression profiles of IMR-32 neuroblastoma cells expressing either GFP or *TFAP2B* (51; GSE74350; ROC-AUC-1.00) and of gene expression profiles of IMR5 neuroblastoma cells with depleted *MYCN* (shMYCN; 50; GSE163431; ROC-AUC=1.00).

(F) Sample ordering based on the differentiation signature using BoNE software of gene expression profiles from a mixed cohort of undifferentiated and differentiating neuroblastoma tumors (from phs000467.v1.p1; 54; p=2.92e-06).

(G) Sample ordering based on the differentiation signature using BoNE software of gene expression profiles from a mixed cohort of nodular and intermixed ganglioneuroblastoma tumors (from phs001436.v1.p1; 55; p=0.04, ROC-AUC=0.76).

(H,I) Sample ordering based on the differentiation signature using BoNE software of single-cell RNA-sequencing gene expression profiles of human fetal adrenal samples and human neuroblastoma tumor samples (52,53; GSE137804, EGAD0000100637). SCPs = Schwann cell precursors, SB = sympathoblastas, CF = Chromaffin cells, NB = neuroblastoma tumor samples.

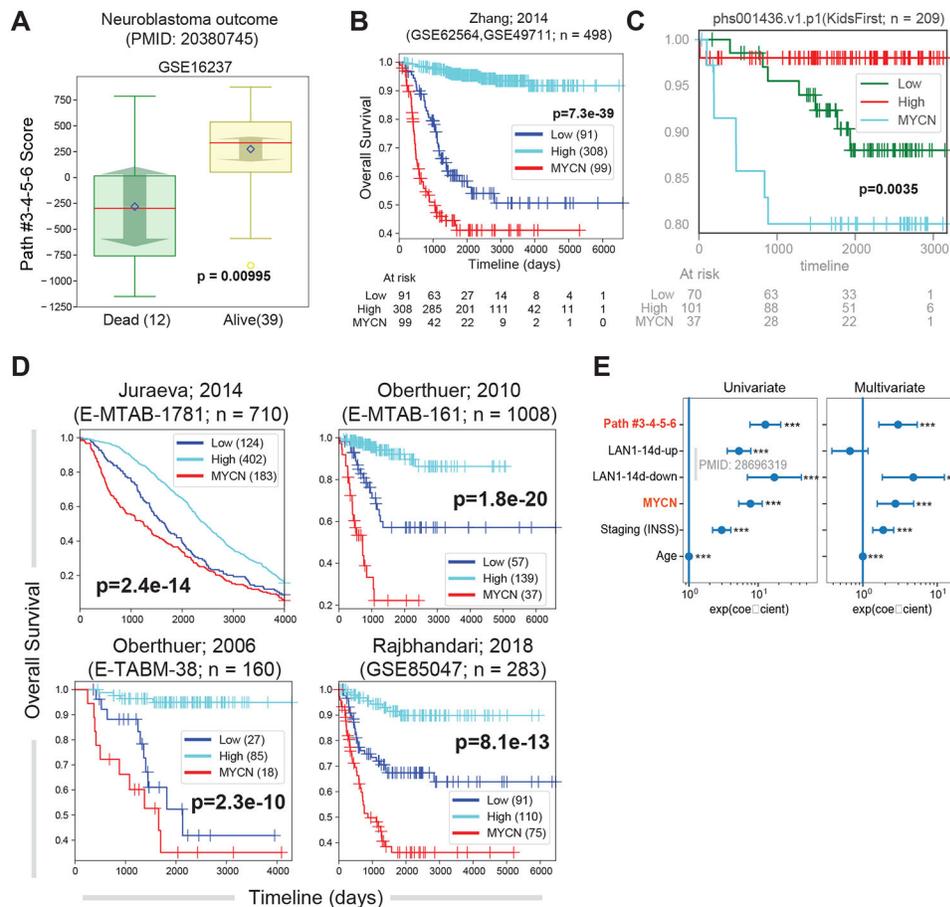


Figure 3. Association of the differentiation signature with neuroblastoma patient outcomes. (A) Relative differentiation signature scores (Path #3-4-5-6 Score) were plotted in patients treated for neuroblastoma who were categorized as alive or dead (43; GSE16237), with patient numbers shown in parentheses ($p=0.00995$). (B) Gene expression values from 498 patient tumors (38; GSE62564/GSE49711) were sorted by *MYCN* expression levels into high *MYCN* (MYCN) and low *MYCN* expression groups, and patient tumors with low *MYCN* expression were further divided into groups with low and high expression of genes in the differentiation signature (Low, High) using BoNE software, with patient numbers in each group shown in parentheses. Kaplan-Meier analyses were performed on each of the three patient groups in each dataset. (C) Gene expression values from neuroblastoma patient tumors (55; phs001436.v1.p1) were sorted by *MYCN* expression levels into high (MYCN) and low expression groups, and patient tumors with low *MYCN* expression were further divided into groups with low and high expression of genes in the differentiation signature (Low, High) using BoNE software, with patient numbers in each group shown in parentheses. Kaplan-Meier analyses were performed on each of the three patient groups in each dataset. (D) Gene expression datasets from four groups of neuroblastoma patient tumors (39-42; E-MTAB-1781, E-MTAB-161, E-TABM-38, GSE85047) were sorted by *MYCN* expression levels into high (MYCN) and low expression groups, and patient tumors with low *MYCN* expression were further divided into groups with low and high expression of genes in the

differentiation signature (Low, High) using BoNE software, with patient numbers in each group shown in parentheses. Kaplan-Meier analyses were performed on each of the three patient groups in each dataset.

(E) Univariate and multivariate Cox regression analyses for overall survival (OS) were performed on gene expression patterns and prognostic features in 498 patients (38; GSE62564/GSE49711) including the differentiation signature score, *MYCN* amplification status, tumor stage, patient age at diagnosis, and up or down regulated gene set expression from LAN1 neuroblastoma cells treated with 10 μ M ATRA and 5-aza-deoxycytidine (LAN1-14d-up, LAN1-14d-down; 44) as independent variables.

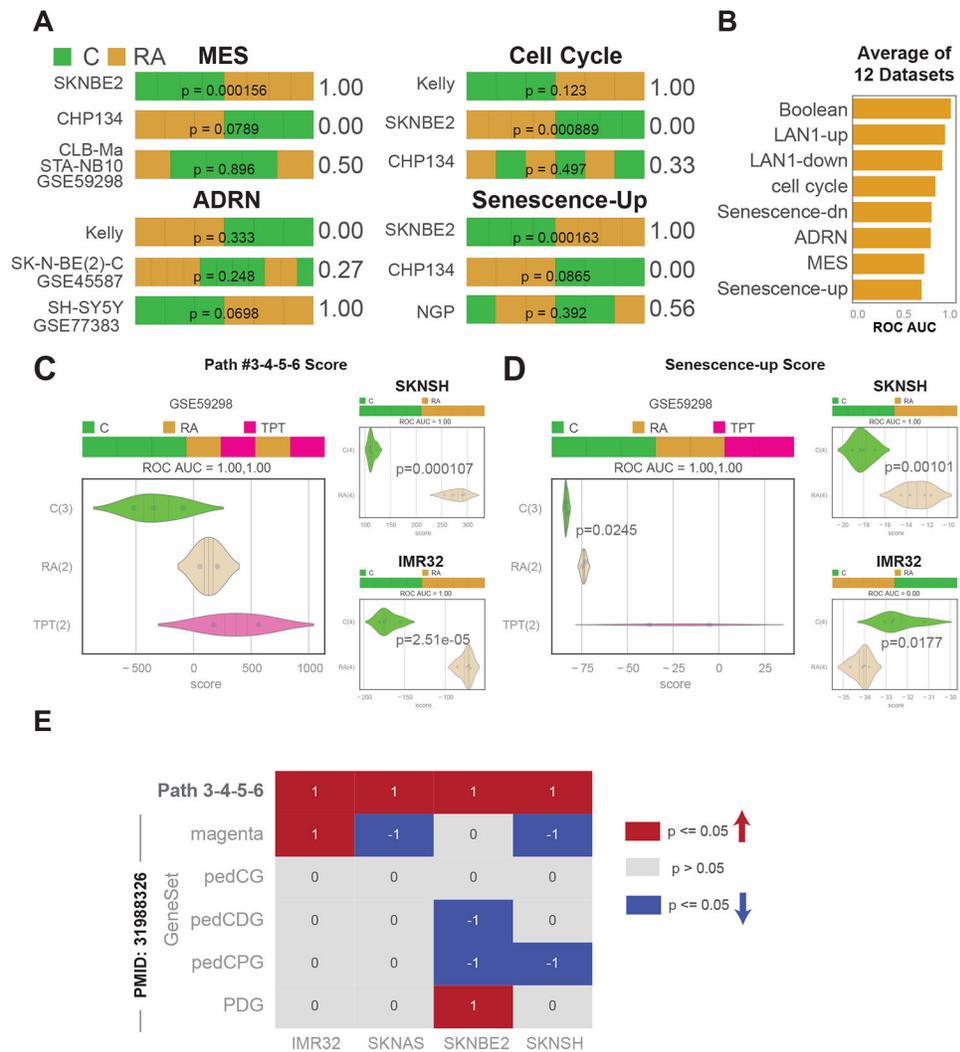


Figure 4. Association of the Differentiation Signature with MYCN Expression

(A) Sample ordering based on the gene expression scores from four pathways (mesenchymal (MES), 63; adrenergic (ADRN), 63; cell cycle, KEGG_CELL_CYCLE, 64; and senescence, FRIDMAN SENESCENCE_UP, 65) are shown in three representative cell line datasets to demonstrate heterogeneity and inconsistency. P-values were computed using two-tailed, two sample Welch's T-tests.

(B) Direct comparison of the performance of the differentiation signature score (using BoNE software) with up or down regulated gene set expression from LAN1 neuroblastoma cells treated with 10 μ M ATRA and 5-aza-deoxycytidine (LAN1-14d-up, LAN1-14d-down; 44), and with other pathways (mesenchymal (MES), 63; adrenergic (ADRN), 63; cell cycle, KEGG_CELL_CYCLE, 64; and senescence, FRIDMAN SENESCENCE_UP, FRIDMAN_SENESCENCE_DN, 65) by computing the average of ROC-AUC values in 12 independent datasets (8 datasets from Figure 2B; GSE100568, GSE45587, GSE77383 and GSE59298; 44-46,49)

(C) (Left) Sample ordering based on the differentiation signature score using BoNE software of gene expression profiles of untreated control neuroblastoma cell lines and

cell lines treated with either 10 μ M ATRA (RA) or topotecan (TPT) (ROC-AUC = 1.00 each)(49; GSE59298), and (Right) of gene expression profiles of SK-N-SH and IMR-32 neuroblastoma cell lines with and without 13-*cis*-retinoic acid treatment (Figure 2B; ROC-AUC=1.00 for each).

(D) (Left) Sample ordering based on the senescence gene signature score (65) of gene expression profiles of untreated control neuroblastoma cell lines and cell lines treated with either 10 μ M ATRA (RA) or topotecan (TPT) (ATRA or TPT, ROC-AUC = 1.00 each), and (Right) of gene expression profiles of SK-N-SH and IMR-32 neuroblastoma cell lines with and without 13-*cis*-retinoic acid treatment (Figure 2B; ROC-AUC=1.00 for SK-N-SH, ROC-AUC=0.00 for IMR-32), demonstrating heterogeneity and inconsistency.

(E) Comparisons of the differentiation score and published pediatric cancer gene modules (82) in distinguishing differentiation induced by 13-*cis*-retinoic acid in 4 neuroblastoma cell lines (IMR-32, SK-N-AS, SK-N-BE(2), and SK-N-SH). Significant changes in the combined gene signatures were color coded by significance and direction of expression change (Red = increased, Blue = decreased, Grey = not significant) based on T-test scores calculated by comparisons of gene signature expression between control and retinoic acid-treated cell lines (Raw p-values in Supplemental Figure 4)

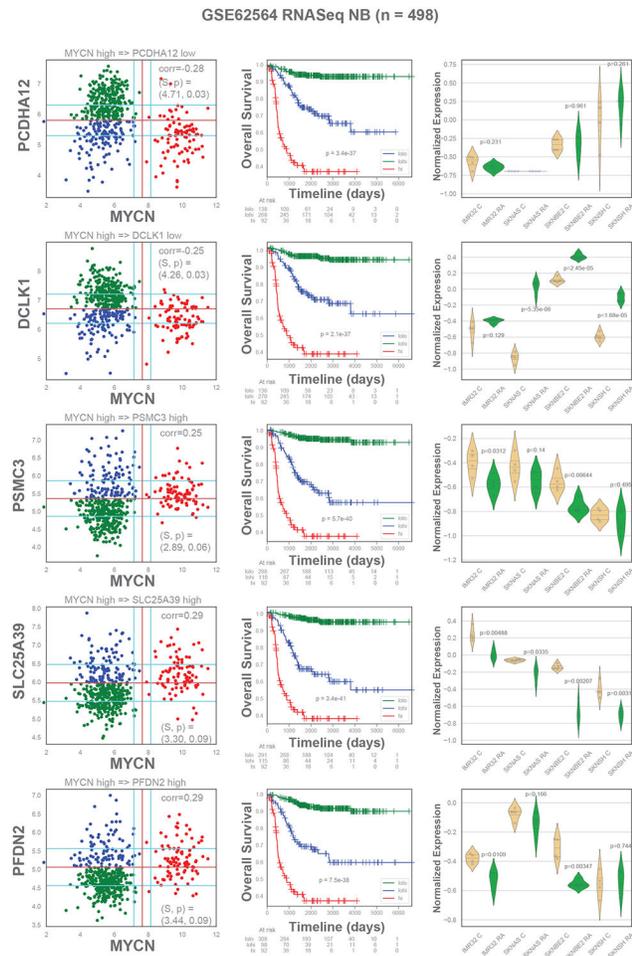


Figure 5. Novel Genes Associated with MYCN Expression and Neuroblastoma Differentiation. (Left) Using BoNE software analysis of RNA expression profiles of 498 neuroblastoma patient tumors (38; GSE62564/GSE49711), the gene expression relationships between *MYCN* and *PCDHA12*, *DCLK1*, *PSMC3*, *SLC25A39*, and *PFDN2* were evaluated using Boolean implication analysis (Center) Gene expression values from neuroblastoma patient tumors (38; GSE62564/GSE49711) were sorted by *MYCN* expression levels into high and low expression groups and were further divided into groups with low and high expression of individual genes using BoNE software, with line colors matching the patient groups from adjacent plots. Kaplan-Meier analyses were performed on each of the four patient groups in each dataset. (Right) Sample ordering based on the differentiation signature using BoNE software of gene expression profiles of control, untreated, neuroblastoma cell lines (C) and neuroblastoma cell lines after 13-cis-retinoic acid treatment (RA). p-values comparing individual untreated (C) and treated (RA) cell lines are shown.

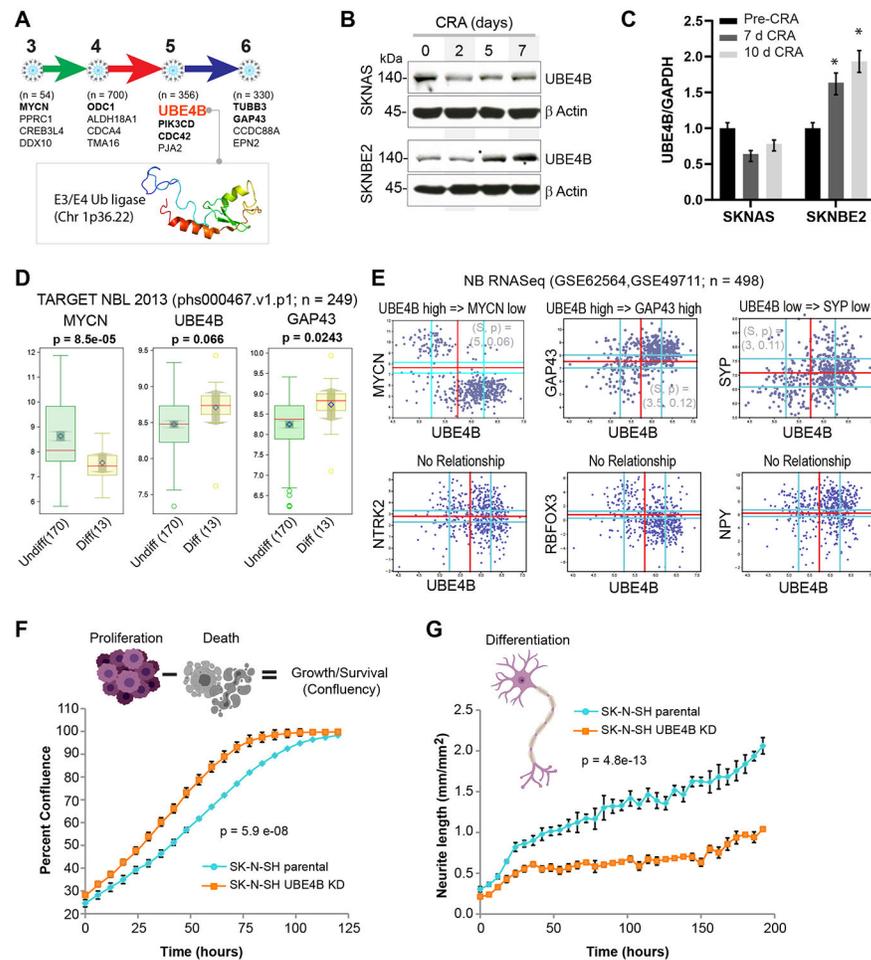


Figure 6. Validation of UBE4B, a Network-predicted Modulator of Neuroblastoma Differentiation.

(A) The *UBE4B* gene is located in cluster 5 in our Boolean network, one of the gene clusters within our differentiation signature, and cluster 5 is comprised of genes that are involved in intracellular signaling and receptor trafficking pathways.

(B,C) Retinoic acid-resistant SK-N-AS and retinoic acid-sensitive SK-N-BE(2) neuroblastoma cells were treated with 5 μ M 13-*cis*-retinoic acid (CRA) for 10 days and analyzed by Western blot for UBE4B protein expression at days 2, 5, and 7 (B) and by qPCR for *UBE4B* gene expression at days 7 and 10 (C).

(D) Using the TARGET dataset (54; phs000467.v1.p1), *MYCN*, *UBE4B*, and *GAP43* expression were compared in undifferentiated (Undiff) and differentiating (Diff) patient tumors ($p=8.5e-05$, $=0.066$, and $=0.0243$, respectively)

(E) Using BoNE software analysis of RNA expression profiles of 498 neuroblastoma patient tumors (38; GSE62564/GSE49711), the gene expression relationships between *UBE4B* and *MYCN* (*UBE4B* high \Rightarrow *MYCN* low; $S>3$, $p=0.06$) and between *UBE4B* and differentiation markers *GAP43* (*UBE4B* high \Rightarrow *GAP43* high; $S>3$, $p=0.12$), *SYP* (*UBE4B* low \Rightarrow *SYP* low; $S>3$, $p=0.11$), *NTRK2* ($p=NS$), *RBFOX3* ($p=NS$), and *NPY* ($p=NS$), were evaluated using Boolean implication analysis.

(F,G) Parental and UBE4B-depleted (UBE4B KD) neuroblastoma cells were treated with 10 μ M 13-*cis*-retinoic acid and were monitored for cell confluence using continuous live cell imaging (F; p= 5.8777e-08) and for differentiation determined by total neurite length per high-power field using NeuroTrack™ software (G; p=4.772763e-13).

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table 1.

Genes with Asymmetric Boolean relationships with MYCN in Neuroblastoma Tumors.

Gene Name	Expression	p-value	Source
cluster 3-4			
PSMC3	-9.26074543	5.05E-10	Suppl Table 2B
SLC25A39	-9.10994675	2.65E-09	Suppl Table 2B
PFDN2	-6.03631336	1.99E-06	Suppl Table 2B
cluster 5-6			
PCDHA12	5.5122718	7.93E-05	Suppl Table 2B
DCLK1	4.3965	5.20E-05	Suppl Table 2A

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript