BORIS Promotes Chromatin Regulatory Interactions in Treatment-Resistant Cancer Cells

BORIS promotes novel chromatin regulatory interactions in treatment-resistant cancer cells

David N. Debruyne[1,2,8], Ruben Dries[1,2,3,8], Satyaki Sengupta[1,2], Davide Seruggia[4], Yang Gao[1,2], Bandana Sharma[1,2], Hao Huang[1,2], Lisa Moreau[5], Michael McLane[1,2], Daniel S. Day[6,7], Eugenio Marco[3,12], Ting Chen[8], Nathanael S. Gray[9,10], Kwok-Kin Wong[11], Stuart H. Orkin[4], Guo-Cheng Yuan[3,12], Richard A. Young[6,7], Rani E. George[1,2]

Department of [1]Pediatric Hematology/Oncology, Dana-Farber Cancer Institute and Boston Children's Hospital, Boston, MA.

Departments of [3]Biostatistics and Computational Biology, [5]Radiation Oncology, [8]Medical Oncology, and [9]Cancer Biology, Dana-Farber Cancer Institute, Boston, MA.

Departments of [2]Pediatrics, [10]Biological Chemistry and Molecular Pharmacology, Harvard Medical School, Boston, MA.

[4]Division of Hematology/Oncology, Boston Children's Hospital and Department of Pediatric Oncology, Dana-Farber Cancer Institute (DFCI), Harvard Stem Cell Institute, Harvard Medical School, Boston, MA 02115, USA; Howard Hughes Medical Institute, Boston, MA 02115, USA.

[6]Whitehead Institute for Biomedical Research and [7]MIT Department of Biology, Cambridge, MA.

[11]Division of Hematology & Medical Oncology, Laura and Isaac Perlmutter Cancer Center, New York University Langone Medical Center, New York, NY.

[12]Department of Biostatistics, Harvard TC Chan School of Public Health, Boston, MA.

[8]These authors contributed equally: David N. Debruyne, Ruben Dries

*e-mail: rani_george@dfci.harvard.edu

The CCCTC-binding factor (CTCF), which anchors DNA loops that organize the genome into structural domains, plays a central role in gene control by facilitating or constraining interactions between genes and their regulatory elements[1,2]. In cancer cells the disruption of CTCF binding at specific loci through somatic mutation[3,4] or DNA hypermethylation[5] results in the loss of loop anchors and consequent activation of oncogenes. By contrast, the germ cell-specific paralog of *CTCF*, *BORIS* (Brother of the Regulator of Imprinted Sites)[6], is overexpressed in multiple cancers[7-9], but its contributions to the malignant phenotype remain unclear. Here we show that aberrant upregulation of *BORIS* promotes novel chromatin interactions in *ALK*-mutated, *MYCN*-amplified neuroblastoma[10] cells rendered resistant to ALK inhibition. These cells are reprogrammed to a distinct phenotypic state during the acquisition of resistance, a process defined by the initial loss of *MYCN* expression followed by subsequent overexpression of *BORIS* and a concomitant switch in cellular dependence from MYCN to BORIS. The resultant BORIS-regulated alterations in chromatin looping lead to the formation of new super-enhancers that drive the ectopic expression of a subset of proneural transcription factors that ultimately define the resistance phenotype. These results identify a previously unrecognized role of BORIS – to engender regulatory chromatin interactions that support specific cancer phenotypes.

Unlike *CTCF*, which is uniformly expressed in normal tissues and cancer cells, *BORIS* expression is typically restricted to the testis[6] and embryonic stem cells[11] (**Extended Data Fig. 1a**). However, when aberrantly expressed in cancer[7-9], it appears to be associated with high-risk features including treatment resistance (**Extended Data Fig. 1b, c**). We identified *BORIS* as one of the most differentially expressed genes in ALK inhibitor-resistant neuroblastoma (NB) cells driven by amplified *MYCN*[12] and *ALK(F1174L)*[13] . Kelly human NB cells were exposed to increasing concentrations of the ALK inhibitor TAE684[14] until stable resistance was achieved (**Fig. 1a, Extended Data Fig. 2a-d**). The acquisition of stable resistance coincided not only with the loss of ALK phosphorylation, indicating that the cells no longer required activation of this receptor tyrosine kinase to maintain their oncogenic properties, but also with the absence of other common instigators of resistance (**Extended Data Fig. 2a, e-h**; **Supplementary Note 1**). However, comparison of the gene expression profiles of the TAE684-sensitive and resistant cells showed generalized downregulation of transcription in the resistant cells, but with marked upregulation of a subset of transcription factors (TFs) not typically associated with NB cells[15,16] (**Fig. 1b**).

We therefore hypothesized that the resistant cells had likely undergone transcriptional reprogramming during the development of resistance. To elucidate the dynamics of resistance development, we performed single-cell RNA sequencing (scRNA-seq) on sensitive, intermediate and fully resistant cell states (**Extended Data Fig. 3a**). Principal component analysis indicated a stepwise transition as cells progressed from the sensitive to the fully resistant state (**Fig. 1c**). This transition was confirmed by distributed stochastic neighbor embedding (t-SNE)[17], which clustered the cells into three nonoverlapping categories (**Extended Data Fig. 3b, c**). Pseudotime analysis based on the TFs that were differentially expressed throughout resistance development revealed that the initial major alteration was loss of *MYCN* expression, which persisted in stably resistant cells (**Fig. 1d**, **Extended Data Fig. 3d, e**). To understand this

unexpected result, we analyzed *MYCN* status in these cells, observing that while genomic amplification was retained, the *MYCN* locus was epigenetically repressed (**Extended Data Fig. 3f, g**). This state was accompanied by a genome-wide reduction of MYCN binding to DNA and a consequent revision of associated downstream transcription outcomes[15,18,19] (**Fig. 1e, Extended Data Fig. 3h**). Coincident with this loss of transcriptional activity, the resistant cells were no longer dependent on MYCN for survival, unlike their sensitive controls, which underwent apoptosis upon MYCN depletion (**Extended Data Fig. 3i**). Subsequent resistance stages were defined by a gradual increase in the expression of the neural developmental markers *SOX2* and *SOX9*[20], followed by upregulation of *BORIS*, ultimately leading to a fully resistant state in which *BORIS* expression was highest and detectable in essentially all cells (**Fig. 1d, Extended Data Fig. 3j, k**). *BORIS* overexpression, which coincided with promoter hypomethylation (**Extended Data Fig. 4a, b**), was also observed in additional NB cell lines rendered resistant to TAE684 (SK-N-SH) or the CDK12 inhibitor E9[21] [SK-N-BE(2)] (**Extended Data Fig. 4c, d**), suggesting that our findings are not restricted to a single cell line or kinase inhibitor. Indeed, tumor *BORIS* overexpression was significantly associated with high-risk disease and a poor outcome in NB patients treated with a variety of regimens (**Extended Data Fig. 4e-g**).

To clarify the role of BORIS in the resistance phenotype, we depleted its expression in resistant cells, observing a partial reversal to the sensitive-cell state with re-emergence of MYCN and ALK expression (**Fig. 1f, Extended Data Fig. 5a-c**). However, this outcome was insufficient to maintain cell growth, as BORIS depletion in resistant cells ultimately decreased cell viability (**Extended Data Fig. 5d, e**), indicating a switch from MYCN to BORIS dependency with stable resistance. This transition was associated with changes in cellular growth kinetics - from a highly proliferative, MYCN-overexpressing sensitive state to an intermediate, slow-cycling phenotype that was partially reversed in fully resistant cells, coincident with BORIS overexpression (**Extended Data Fig. 5f-h**). Given the multiple sequential steps involved in the

evolution of resistance, BORIS overexpression alone was not adequate to induce this phenotype (data not shown). Instead, concomitant downregulation of MYCN expression and BORIS overexpression in the presence of ALK inhibition were required to generate resistance in sensitive cells (**Fig. 1g**). This combination of factors also led to increased expression of the TFs that were upregulated in the original TAE684-resistant cells, including *SOX2* and *SOX9* (**Extended Data Figs. 3d, 5i**). Thus, resistance to ALK inhibition in NB cells evolves through a multistep process that promotes a dependency switch from a dominant oncogenic stimulus - amplified *MYCN* - to a phenotypically distinct state characterized by *BORIS* overexpression. In this context, the resistant cells ultimately become dependent on BORIS for survival, supporting a key role for this protein in maintenance of the resistance state.

We next asked if the aberrant expression of BORIS, a DNA-binding protein[6], affected its genome-wide occupancy in resistant cells. We observed a large (10-fold) gain in BORIS-bound ChIP-seq peaks in resistant cells: 22,891 vs. 2,211 in sensitive cells (**Fig. 2a**, **Extended Data Fig. 6a, b**). By contrast, CTCF binding did not change substantially between sensitive and resistant cells (75,567 vs. 63,246 peaks) (**Fig. 2b**). A significant portion (*n* = 17,042; 78%) of the BORIS peaks unique to resistant cells overlapped with CTCF peaks shared by both cell types (**Fig. 2c**), consistent with their heterodimerization[22] (**Extended Data Fig. 6c**). However, only a small proportion (*n* = 1,903; 8.7%) overlapped with CTCF peaks unique to sensitive cells, suggesting that BORIS does not replace CTCF in resistant cells. BORIS preferentially occupied gene regulatory regions - enhancers and promoters (60%) in resistant cells (**Extended Data Fig. 6d, e**), consistent with its propensity to bind to open chromatin regions (**Fig. 2d**)[23]. Such differential chromatin binding at distinct highly expressed genes in resistant versus sensitive cells was commensurate with the MYCN-to-BORIS dependency switch (**Extended Data Fig. 6f, g**).

The proclivity of aberrantly expressed BORIS for genomic regions associated with active chromatin features in resistant cells suggested that it may, like CTCF and cohesin, regulate gene expression through chromatin looping. Thus, we examined the chromatin looping profiles of sensitive and resistant cells, using cohesin (SMC1A)-based HiChIP[24] (**Extended Data Fig. 7a**). Based on the genomic locations of the associated loop anchors, six classes of interactions were identified[25]: three longer average interaction loops with a CTCF site on at least one anchor and three smaller connecting regulatory regions (**Fig. 3a, Extended Data Fig. 7b**). The overlap of BORIS binding with loop anchors revealed that the majority (56%) of the 9,487 interactions gained in resistant cells were positive for BORIS ($log_2FC > 1$; FDR < 0.01) (**Fig. 3b**, **Extended Data Fig. 7c**). Importantly, BORIS was strikingly enriched at anchors that were associated with regulatory regions, while CTCF binding remained constant, as seen at the BORIS locus itself (**Fig. 3c, d**). Notably, BORIS binding alone at CTCF-negative loop anchors was sufficient to generate new interactions in resistant cells (**Extended Data Fig. 7d**).

To test whether the newly formed interactions in resistant cells were mediated by BORIS binding, we analyzed the consequences of BORIS depletion on loop architecture (**Extended data Fig. 7e**). Novel regulatory interactions specific to the resistant cells displayed a global shift towards loss upon BORIS knockdown (**Fig. 3e**), with more than a quarter of the total interactions lost, of which 63% were positive for BORIS at their anchors (**Fig. 3f**). Interactions whose anchors were bound by BORIS (especially enhancer-promoter and promoter-promoter interactions) were more likely to be lost after BORIS depletion than those that were not (**Fig. 3f, Extended Data Fig. 7f, g**). These results agree with the loop extrusion model[26], as BORIS loss resulted in decreased SMC1A binding, preferentially at lost interactions, whereas CTCF binding did not change significantly (**Fig. 3g, Extended Data Fig. 7h-j**). These data confirm that BORIS is a critical factor in the looping landscape of resistant cells.

Genes associated with new BORIS-positive regulatory interactions were observed to be expressed at higher levels than those associated with BORIS-negative regulatory interactions or genes not associated with new regulatory interactions (**Fig. 4a**). Because genes that define cell identity are often regulated by super-enhancers (SEs) in both normal and cancer cells[15,27,28], we characterized the SE landscape of our cells, observing that the SEs unique to resistant cells were enriched at BORIS-positive regulatory loops (**Extended Data Fig. 8a-c**). The presence of such SEs correlated significantly with higher expression of their associated genes in resistant versus sensitive cells (**Fig. 4a**). These BORIS-positive SE-associated genes were also enriched for genes that underwent a chromatin state switch from a closed or neutral to an open configuration in resistant cells (**Extended Data Fig. 8d, e**). Depletion of BORIS resulted in the decreased expression of genes associated with BORIS-positive interactions, especially genes associated with resistant cell-specific SEs (**Fig. 4a, Extended Data Fig. 8f**). These observations suggest that BORIS-mediated alterations in chromatin looping lead to interactions of newly-formed SEs with their target genes, resulting in their increased expression.

We next sought to identify BORIS-regulated genes that are functionally linked to the resistance phenotype by integrating gene expression, BORIS-mediated looping, SE landscape and chromatin state. This analysis revealed 89 genes (**Supplementary Information Table**), including 13 TFs, that are highly expressed during early neural development and are critical to cell fate decisions[20,29,30] (**Fig. 4b, c, Extended Data Fig. 8g**). The expression of these proneural TFs paralleled that of BORIS in resistant cells, and was dependent on BORIS-mediated looping, as BORIS depletion led to their downregulation (**Extended Data Fig. 8h, i**). Moreover, TF binding site analysis revealed enrichment of BORIS and several of these proneural TFs at the regulatory regions of the highest expressed genes in resistant cells, while sensitive cells were dominated by MYC/MYCN/MAX E-box and E-box-like motifs (**Fig. 4d**). Similar increased expression of proneural TFs with increased BORIS occupancy at their promoters was seen in

BORIS-overexpressing E9-resistant SK-N-BE(2) NB cells compared with their sensitive counterparts (**Extended Data Fig. 8j, k**). The high transcriptional activity of these BORIS-regulated genes was also associated with increased binding of the transcriptional activator BRD4, rendering the resistant cells more sensitive to BET inhibition (**Extended Data Fig. 9**; **Supplementary Note 2**). Together, these results indicate the establishment of an alternative TF regulatory network controlled by BORIS-induced chromatin remodeling to support the resistant cell state.

Thus, using a pair of isogenic ALK-inhibitor sensitive and resistant NB cell lines, we show that the CTCF paralog BORIS is capable of promoting novel regulatory DNA interactions that support a phenotypic switch in the context of treatment resistance (**Fig. 4e**). This mechanism appears relevant to different NB cell lines and kinase inhibitors and may extend to other cancers. In Ewing sarcoma, where BORIS overexpression is associated with metastasis and relapse (**Extended Data Fig. 1c**), we observed increased BORIS occupancy at regulatory regions in chemotherapy-resistant cell lines (**Extended Data Fig. 10**; **Supplementary Note 3**). Further work will establish whether BORIS-mediated altered chromatin looping is a general mechanism whereby tumor cells co-opt developmental networks to sustain alternative cell states in response to targeted or conventional therapies.

References

1 Dixon, J. R. *et al.* Topological domains in mammalian genomes identified by analysis of chromatin interactions. *Nature* **485**, 376-380, doi:10.1038/nature11082 (2012).

2 Phillips-Cremins, J. E. *et al.* Architectural protein subclasses shape 3D organization of genomes during lineage commitment. *Cell* **153**, 1281-1295, doi:10.1016/j.cell.2013.04.053 (2013).

3 Hnisz, D. *et al.* Activation of proto-oncogenes by disruption of chromosome neighborhoods. *Science* **351**, 1454-1458, doi:10.1126/science.aad9024 (2016).

4 Katainen, R. *et al.* CTCF/cohesin-binding sites are frequently mutated in cancer. *Nature genetics* **47**, 818-821, doi:10.1038/ng.3335 (2015).

5 Flavahan, W. A. *et al.* Insulator dysfunction and oncogene activation in IDH mutant gliomas. *Nature* **529**, 110-114, doi:10.1038/nature16490 (2016).

6 Loukinov, D. I. *et al.* BORIS, a novel male germ-line-specific protein associated with epigenetic reprogramming events, shares the same 11-zinc-finger domain with CTCF, the insulator protein involved in reading imprinting marks in the soma. *Proceedings of the National Academy of Sciences of the United States of America* **99**, 6806-6811, doi:10.1073/pnas.092123699 (2002).

7 Klenova, E. M., Morse, H. C., 3rd, Ohlsson, R. & Lobanenkov, V. V. The novel BORIS + CTCF gene family is uniquely involved in the epigenetics of normal biology and cancer. *Seminars in cancer biology* **12**, 399-414 (2002).

8 Martin-Kleiner, I. BORIS in human cancers -- a review. *European journal of cancer* **48**, 929-935, doi:10.1016/j.ejca.2011.09.009 (2012).

9 Garikapati, K. R. *et al.* Down-regulation of BORIS/CTCFL efficiently regulates cancer stemness and metastasis in MYCN amplified neuroblastoma cell line by modulating Wnt/beta-catenin signaling pathway. *Biochem Biophys Res Commun* **484**, 93-99, doi:10.1016/j.bbrc.2017.01.066 (2017).

10 Cheung, N. K. & Dyer, M. A. Neuroblastoma: developmental biology, cancer genomics and immunotherapy. *Nat Rev Cancer* **13**, 397-411, doi:10.1038/nrc3526 (2013).

11 Monk, M., Hitchins, M. & Hawes, S. Differential expression of the embryo/cancer gene ECSA(DPPA2), the cancer/testis gene BORIS and the pluripotency structural gene OCT4, in human preimplantation development. *Molecular human reproduction* **14**, 347-355, doi:10.1093/molehr/gan025 (2008).

12 Brodeur, G., Seeger, R. C., Schwab, M., Varmus, H. E. & Bishop, J. M. Amplification of N-*myc* in untreated human neuroblastomas correlates with advanced disease stage. *Science* **224**, 1121-1124 (1984).

13 George, R. E. *et al.* Activating mutations in ALK provide a therapeutic target in neuroblastoma. *Nature* **455**, 975-978 (2008).

14 Galkin, A. V. *et al.* Identification of NVP-TAE684, a potent, selective, and efficacious inhibitor of NPM-ALK. *Proc Natl Acad Sci U S A* **104**, 270-275 (2007).

15 Chipumuro, E. *et al.* CDK7 inhibition suppresses super-enhancer-linked oncogenic transcription in MYCN-driven cancer. *Cell* **159**, 1126-1139, doi:10.1016/j.cell.2014.10.024 (2014).

16 Durbin, A. D. *et al.* Selective gene dependencies in MYCN-amplified neuroblastoma include the core transcriptional regulatory circuitry. *Nature genetics* **50**, 1240-1246, doi:10.1038/s41588-018-0191-z (2018).

17 van der Maaten, L. & Hinton, G. Visualizing Data using t-SNE. *Journal of Machine Learning Research* **9**, 2579-2605 (2008).

18 Zeid, R. *et al.* Enhancer invasion shapes MYCN-dependent transcriptional amplification in neuroblastoma. *Nature genetics* **50**, 515-523, doi:10.1038/s41588-018-0044-9 (2018).

19 Muhar, M. *et al.* SLAM-seq defines direct gene-regulatory functions of the BRD4-MYC axis. *Science* **360**, 800-805, doi:10.1126/science.aao2793 (2018).
20 Wegner, M. & Stolt, C. C. From stem cells to neurons and glia: a Soxist's view of neural development. *Trends in neurosciences* **28**, 583-588, doi:10.1016/j.tins.2005.08.008 (2005).
21 Gao, Y. *et al.* Overcoming Resistance to the THZ Series of Covalent Transcriptional CDK Inhibitors. *Cell chemical biology* **25**, 135-142 e135, doi:10.1016/j.chembiol.2017.11.007 (2018).
22 Pugacheva, E. M. *et al.* Comparative analyses of CTCF and BORIS occupancies uncover two distinct classes of CTCF binding genomic regions. *Genome biology* **16**, 161, doi:10.1186/s13059-015-0736-8 (2015).
23 Bergmaier, P. *et al.* Choice of binding sites for CTCFL compared to CTCF is driven by chromatin and by sequence preference. *Nucleic acids research* **46**, 7097-7107, doi:10.1093/nar/gky483 (2018).
24 Mumbach, M. R. *et al.* HiChIP: efficient and sensitive analysis of protein-directed genome architecture. *Nature methods* **13**, 919-922, doi:10.1038/nmeth.3999 (2016).
25 Dowen, J. M. *et al.* Control of cell identity genes occurs in insulated neighborhoods in mammalian chromosomes. *Cell* **159**, 374-387, doi:10.1016/j.cell.2014.09.030 (2014).
26 Sanborn, A. L. *et al.* Chromatin extrusion explains key features of loop and domain formation in wild-type and engineered genomes. *Proceedings of the National Academy of Sciences of the United States of America* **112**, E6456-6465, doi:10.1073/pnas.1518552112 (2015).
27 Whyte, W. A. *et al.* Master transcription factors and mediator establish super-enhancers at key cell identity genes. *Cell* **153**, 307-319, doi:10.1016/j.cell.2013.03.035 (2013).
28 Hnisz, D. *et al.* Super-enhancers in the control of cell identity and disease. *Cell* **155**, 934-947, doi:10.1016/j.cell.2013.09.053 (2013).
29 Kumar, J. P. The sine oculis homeobox (SIX) family of transcription factors as regulators of development and disease. *Cellular and molecular life sciences : CMLS* **66**, 565-583, doi:10.1007/s00018-008-8335-4 (2009).
30 Dennis, D. J., Han, S. & Schuurmans, C. bHLH transcription factors in neural development, disease, and reprogramming. *Brain research* **1705**, 48-65, doi:10.1016/j.brainres.2018.03.013 (2019).

Main Figure Legends

Figure 1. Targeted therapy resistance in NB is associated with transcriptional reprogramming and a switch in dependency from amplified MYCN to BORIS. a, Dose–response curves of TAE684-sensitive and resistant Kelly NB cells incubated in increasing concentrations of TAE684 for 72 h. Data are means ± SD, *n* = 3 biological replicates. Schematic representation of resistance development is shown above. **b,** Heatmap of gene expression values in sensitive (Sens) vs. resistant (Res) cells (*n* = 2 biological replicates). Rows are z-scores calculated for each gene in both cell types. **c,** Principal component analysis (PCA) of scRNA-seq data of sensitive (*n* = 5,432), intermediate (IR, *n* = 6,376) and resistant (*n* = 6,379) cells showing the first two principal components (PCs). **d,** Pseudotime analysis of TF expression during resistance development. **e,** ChIP-seq signals of genome-wide MYCN binding in sensitive and resistant cells, reported as rpm/bp for each chromosome (rpm/bp, reads per million/per base pair). **f,** PCA of gene expression profiles showing the first two PCs (*n* = 2 biological replicates). **g,** Dose-response curves for TAE684 (IC_{50} values in parenthesis) and immunoblot analysis (representative of two independent experiments) of BORIS and MYCN expression in sensitive cells expressing MYCN shRNA and doxycycline-inducible BORIS ($MYCN^{KD}$/$BORIS^{Ind}$), treated with DMSO or TAE684, 1 µM, with or without doxycycline. Data are means ± SD, *n* = 3 biological replicates.

Figure 2. BORIS overexpression is associated with its increased chromatin occupancy in resistant cells, while CTCF binding is unchanged. a, Scatterplot of BORIS binding in sensitive and resistant cells. BORIS peaks unique to resistant cells (*n* = 21,805; 91%), sensitive cells (*n* = 1,125; 4.7%) and BORIS peaks shared between the two cell types (*n* = 1,086; 4.5%). **b,** Scatterplot of CTCF binding in sensitive and resistant cells for all detected CTCF binding sites. Unique CTCF peaks in resistant cells (*n* = 6,808; 8.3%), sensitive cells (*n* = 19,129;

23.2%) and CTCF peaks shared between the two cell types (*n* = 56,438; 68.5%). **c,** Overlap between BORIS peaks that are unique to resistant cells and CTCF peaks shared between resistant and sensitive cells (upper), and between sensitive cell-specific CTCF peaks (lower). **d,** Meta-analysis of average ChIP-seq signals at resistant cell-specific BORIS binding sites. All panels, *n* = 2 biological replicates.

Figure 3. **BORIS promotes new chromatin interactions in resistant cells. a,** DNA interactions gained in resistant cells based on SMC1A HiChIP analysis. Interaction classes were determined from the genomic locations of the associated anchors [overlapping promoter regions (TSS ± 2 kb), active enhancer regions, or CTCF sites only, in that order]. Absolute numbers and percentages for each loop type [structural **(black)**, regulatory **(blue)**] are shown. Cartoon illustrating the spatial proximity induced by DNA looping between these regions. **b,** Fractions of loops bound by BORIS within each interaction class. **c,** Meta-analysis of average CTCF and BORIS ChIP-seq signals in sensitive and resistant cells at the three main anchor types normalized by the number of interactions (*n* = 2 biological replicates). Anchor sites were centered and extended in both directions (± 2 kb). **d,** ChIP-seq tracks of the indicated proteins in sensitive and resistant cells at the *BORIS* locus (representative of two independent experiments), with resistant cell-specific regulatory interactions shown below [HiChIP Res: paired-end tag (PET) numbers, next to each interaction]. Signal intensity is given in the upper left corner for each track. **e,** PET interactions in BORIS-depleted (shBORIS) vs. control (shCtrl) cells. **f,** Resistant cell-specific loops lost upon BORIS depletion based on loops negative or positive for BORIS binding in shCtrl cells **(left)**, and the odds-ratio of losing a loop previously bound by BORIS (two-sided Fisher's exact test) **(right)**. **g,** Meta-analysis of average BORIS, SMC1A and CTCF ChIP-seq signals at resistant cell-specific loop anchors that were lost upon BORIS depletion (*n* = 2 biological replicates). BORIS depletion at loop anchors inhibits retention

of the cohesin complex, thus preventing new loop formation (loop extrusion model). Panels **a**, **b**, **e** and **f**, *n* = 3 biological replicates.

Figure 4. BORIS-regulated chromatin remodeling supports a phenotypic switch that maintains the resistant state. a, Box plots comparing: **Left,** Fold-change expression in counts per million (CPM) of genes involved in resistant cell-specific regulatory interactions *positive* for BORIS binding (BORIS reg gene) (*n* = 1,368) vs. those involved in regulatory interactions *negative* for BORIS (reg gene) (*n* = 519) or not associated with a novel regulatory interaction (other) (*n* = 16,151). **Centre,** Fold-change expression of genes involved in resistant cell-specific regulatory interactions positive for BORIS binding *and* associated with SEs specific to resistant cells (BORIS Res SE) (*n* = 134) vs. those with SEs shared by both cell types (BORIS both SE) (*n* = 514) or not associated with SEs (BORIS no SE) (*n* = 720). **Right,** Fold-change expression of genes involved in resistant cell-specific regulatory interactions positive for BORIS binding and associated with resistant cell-specific SEs before and after BORIS knockdown (*n* = 134) (*P*; two-sided Wilcoxon rank-sum test). Centre lines, medians; box limits, 25^{th} and 75^{th} percentiles; whiskers, minima and maxima (1.5X the interquartile range). **b,** Highest ranked TFs associated with the resistance phenotype selected based on presence of at least 4 of 5 of the indicated features. **c,** ChIP-seq tracks of the indicated proteins in sensitive and resistant cells at the *NEUROG2* locus; regulatory interactions with PET numbers indicated below. **d,** TF recognition motifs at SEs and promoters (± 2 kb) of the 1,000 highest expressed genes in resistant and sensitive cells (*n* = 2 biological replicates) (*P*; hypergeometric enrichment test). Panels **a**-**c** integrate data of biological replicates from expression microarrays (*n* = 2), ChIP-seq (*n* = 2) and HiChIP (*n* = 3). **e,** Proposed role of BORIS in resistant cells.

METHODS

Cell lines. Human NB cell lines Kelly and SK-N-BE(2) and human Ewing sarcoma cell lines TC-32, TC-71 and CHLA-10[31,32] were obtained from the Children's Oncology Group cell line bank (Lubbock, TX, USA). Human NB cell line SK-N-SH and human embryonic kidney cell line HEK293T were obtained from the American Type Culture Collection (Manassas, VA, USA). Cell line authenticity was confirmed by genotyping, and cells were tested negative for mycoplasma contamination every 3 months. All cells except HEK293T were grown in RPMI-1640 media supplemented with 10% fetal bovine serum (FBS) and 1% penicillin/streptomycin (Life technologies, Carlsbad, CA). HEK293T cells were grown in DMEM media supplemented with 10% FBS and 1% penicillin/streptomycin (Life technologies). Resistant cells were grown in the presence of either the ALK inhibitor, TAE684[33] (Kelly and SK-N-SH) or the CDK12 inhibitor, E9[34] (SK-N-BE(2)).

Compounds. TAE684 and E9 were synthesized in-house in Dr Nathanael Gray's laboratory and JQ1[35] was obtained from Dr Jun Qi's laboratory at the Dana-Farber Cancer Institute (DFCI) (Boston, MA, USA). Ceritinib[36], lorlatinib[37] and I-BET726[38] were purchased from Selleck Chemicals (Houston, TX, USA).

Synthetic RNA Spike-In and Microarray Analysis. Total RNA and sample preparation was performed as previously described[39]. Briefly, cells were either incubated in media containing DMSO, TAE684 (1 μM) or JQ1 (2.5 μM), or infected with shRNA (Ctrl or BORIS) for 24 h. Cell numbers were determined using a Countess II cell counter (Life Technologies) prior to lysis and RNA extraction. Biological duplicates (equivalent to 5 x 10e6 cells per replicate) were collected and homogenized in 1 ml of TRIzol Reagent (Ambion, Carlsbad, CA, USA), purified using the mirVANA miRNA isolation kit (Ambion) following the manufacturer's instructions and re-suspended in 50 μl nuclease-free water (Ambion). Total RNA was spiked-in with ERCC RNA Spike-In Mix (Ambion), treated with DNA-free™ DNase I (Ambion) and analyzed on an Agilent

2100 Bioanalyzer (Agilent Technologies, Santa Clara, CA, USA) for integrity. RNA with the RNA Integrity Number above 9.8 was hybridized to Affymetrix GeneChip PrimeView Human Gene Expression arrays (Affymetrix, Santa Clara, CA, USA).

Antibodies. The following antibodies were used: N-Myc (#9405), N-Myc (#51705), cleaved PARP (#9541), cleaved caspase 3 (#9661), ALK (#3333), AKT (#4691), pAKT^{T308} (#9275), pAKT^{S473} (#9271), ERK (#4695), pERK (#4377), S6 (#2217), pS6 (#4857), STAT3 (#4904), pSTAT3 (#9131), ABCB1 (#12683), SOX2 (#3579), beta-actin (#4967), CTCF (#3417), normal rabbit IgG (#2729) and HRP anti-mouse IgG (#7076) from Cell Signaling Technology (Danvers, MA, USA); HRP anti-rabbit IgG (sc-2357) from Santa Cruz Biotechnology (Santa Cruz, CA, USA); BRD4 (A301-985A100) and SMC1A (A300-055A) from Bethyl Laboratories (Montgomery, TX, USA); CTCF (#07-729) , SOX9 (#AB5535) and H3K27me3 (#07-449) from Millipore (Billerica, MA, USA); pALK^{Y1507} (ab73996), BORIS (ab187163) and H3K27ac (ab2729) from Abcam (Cambridge, MA, USA); BORIS (NBP2-52405) from NOVUS Biologicals (Littleton, CO, USA); BORIS (#39851) from Active Motif (Carlsbad, CA, USA); and SIX1 (HPA001893) from Sigma-Aldrich (Saint Louis, MO, USA); and Vysis LSI N-MYC (2p24) SpectrumGreen/Vysis CEP 2 SpectrumOrange Probe (07J72-001) from Abbott (Abbott Park, IL, USA).

Cell Viability and Growth Curve Assays. Viability and growth experiments were performed using the CellTiter-Glo® Luminescent Cell Viability Assay (Promega, Madison, WI, USA) according to the manufacturer's instructions, as previously described[40]. Cells were plated in 96-well plates at a seeding density of 4 x 10e3 cells per well. For growth assays, the cells were analyzed each day until Day 5. For viability, after 24 h, the cells were treated with various concentrations of the indicated drug (ranging from 1 nM to 10 µM except for I-BET726: 2nM to 20 µM). DMSO without drug served as a negative control. After 72 h of incubation, cells were analyzed for cell viability and IC_{50} values were determined using a nonlinear regression curve fit with GraphPad Prism 6 software (La Jolla, CA, USA).

Cell cycle analysis. Cell cycle analysis was performed 24 h after cell plating using propidium iodide (PI) staining, as described previously[41]. Cells fixed with 80% ethanol overnight at 4°C were resuspended in PBS supplemented with 0.1% Triton X-100 (Sigma-Aldrich), 25 mg/ml PI (BD Biosciences, San Jose, CA, USA) and 0.2 mg/ml RNase A (Sigma-Aldrich). After 45 min at 37°C in the dark, analysis was performed on a FACSCalibur™ flow cytometer (BD Biosciences). Cell-cycle profiles were plotted as histograms generated using FlowJo software (FLOWJO, Ashland, OR, USA).

Western blotting. Cell or tumor tissue was lysed in NP-40 buffer (Invitrogen, Carlsbad, CA, USA) containing a 1X complete protease inhibitor tablet (Roche, Basel, Switzerland) per 10 ml buffer and a cocktail of phosphatase inhibitors (Roche). Protein concentration was measured using the DC Protein Assay (Bio-Rad, Hercules, CA, USA); protein (50 µg) was denatured in LDS sample buffer with reducing agent (Invitrogen), separated on precast 4%–12% Bis-Tris gels (Life Technologies) and transferred to nitrocellulose membranes (Bio-Rad). Membranes were incubated in blocking buffer (5% dry milk in TBS with 0.2% Tween-20) for 1 h, and then incubated in the primary antibody in blocking buffer overnight at 4°C. Chemiluminescent detection was performed with the appropriate secondary antibodies and developed using Genemate Blue ultra-autoradiography film (VWR, Radnor, PA, USA). The actin loading controls for the protein samples shown in the immunoblots of the following panels (two independent mouse tumor samples, and cell lines representative of two independent experiments) are the same because the samples were run on a single gel but probed for pALK, ALK (**Extended Data Fig. 2a**), MYCN (**Extended Data Fig. 3e**) and BORIS (**Extended Data Fig. 4a**) respectively.

Co-immunoprecipitation (co-IP). Cells were harvested in IP lysis buffer (50 mM Tris-HCl buffer (pH 7.4), 100 mM Nacl, 1% Triton-100, 1 mM PMSF), containing a 1X complete protease inhibitor tablet (Roche) per 10 ml buffer and a cocktail of phosphatase inhibitors (Roche). Homogenates were centrifuged at 20,000 g for 10 min at 4°C to obtain supernatants. DNase I (~

1U/ml) was used to degrade DNA in supernatants by incubation for 1 h at room temperature (RT). Co-immunoprecipitation (co-IP) of endogenously expressed proteins was performed using Protein A Dynabeads® (Invitrogen), according to the manufacturer's instructions. Briefly, antibody-conjugated Dynabeads® were incubated with purified cell lysates to immunoprecipitate the target antigen. Antibodies used for immunoprecipitation were CTCF (#3417, Cell Signaling Technology) and BORIS (NBP2-52405, NOVUS Biologicals). The elution step was conducted by heating the beads for 10 min at 95°C in Lithium dodecyl sulfate (LDS) sample buffer with reducing agent (Invitrogen), after which, western blotting was performed using the following antibodies: CTCF (#3417, Cell Signaling Technology) and BORIS (#39851, Active Motif).

Plasmids, shRNA Knockdown and overexpression systems. pLKO.1 shRNA constructs (Ctrl: SHC007; MYCN: #1-TRCN0000020694 and #2-TRCN0000363425; BORIS: #3-TRCN0000370229 and #4-TRCN0000365141; BRD4: #A-TRCN0000318771 and #B-TRCN0000196576) were purchased from Sigma-Aldrich and pLKO.1 GFP shRNA was a gift from Dr D. Sabatini (Addgene plasmid # 30323)[42]. Overexpression constructs were generated by cloning BORIS cDNA into the Tet-inducible pInducer20 vector, provided by Dr S. Elledge (Addgene plasmid #44012)[43]. Production of lentiviral particles and subsequent infection were performed as previously described[40]. The lentivirus was packaged by co-transfection of either pLKO.1 or pInducer20 plasmid with the helper plasmids, pCMV-deltaR8.91 and pMD2.G-VSV-G into HEK293T cells using TransIT-LT1 Transfection Reagent (Mirus Bio LLC, Madison, WI, USA). Virus-containing supernatants were collected 48 h after transfection. Cells were infected with 8 µg/ml polybrene (Sigma-Aldrich) and 24 to 48 h later selected with puromycin (pLKO.1) (Sigma-Aldrich) and then harvested at appropriate time points. When using the Tet-inducible system for BORIS overexpression, induction of gene expression was achieved by treating cells every 2 to 3 days with doxycycline (0.2 µg/ml) for a total duration of 37 days.

Quantitative RT-PCR. RNA isolation and PCR amplification were performed as previously described[40], except that the RT-PCR was performed using the SuperScript III First-Strand system (Life technologies). Total RNA was isolated from cell lines with the RNeasy kit (Qiagen, Carlsbad, CA, USA). One microgram of purified RNA was reverse transcribed using Superscript III First-Strand (Invitrogen) according to the manufacturer's protocol, and quantitative PCR was performed using SYBR Green on a Viia7 Real-Time PCR system (Thermo Fisher Scientific, Waltham, MA, USA). All experiments were performed in biological triplicates unless stated otherwise. Each individual biological sample was qPCR-amplified in technical replicates and normalized to actin as an internal control. Amplification was carried out with primers specific to the genes to be quantified (sequences available upon request).

Sequence analysis. The kinase domain of ALK was amplified from cDNA extracted from sensitive and resistant cells using the HotStar HiFidelity Polymerase Kit (Qiagen). The PCR products were cloned into the pGEM-T vector (Promega) and confirmed by sequencing.

RTK Array. The Human Phospho-RTK Array Kit (R&D Systems, Minneapolis, MN, USA) was used as previously described[40]. Cell lysate (500 µg) was incubated on a phospho-RTK membrane array (ARY001B) according to the manufacturer's instructions. Target proteins were captured with their respective antibodies. After washing, the proteins were incubated with a phosphotyrosine antibody conjugated to horseradish peroxidase to allow the detection of captured phospho-RTKs.

Fluorescent in situ hybridization (FISH). Fluorescent in situ hybridization (FISH) analyses were performed using a Vysis LSI N-MYC (2p24) SpectrumOrange Probe (Vysis, Downers Grove, IL, USA), in accordance with the manufacturer's instructions.

Immunohistochemistry (IHC). All human tumor specimens (formalin-fixed paraffin-embedded slides) were obtained under an Institutional Review Board-approved protocol of the Dana-

Farber/Boston Children's Cancer and Blood Disorders Center, and informed consent was obtained from all subjects. Staining was performed by Applied Pathology Systems (Worcester, MA, USA) using the ImmPRESS™ Excel Amplified HRP Polymer Staining Kit (MP-7601, Vector Laboratories, Burlingame, CA, USA) on a Dako Autostainer (Agilent Technologies). Sections were deparaffinized, rehydrated, and subjected to antigen retrieval in citrate-based buffer on a steamer for 25 minutes. Slides were blocked with BLOXALL™ blocking solution and 2.5% horse serum sequentially prior to a 1-h incubation with BORIS antibody at 1:50 dilution (ab187163, Abcam). Sections were then incubated with anti-rabbit Amplifier™ antibody and ImmPRESS™ Excel Amplified HRP Polymer Reagent sequentially before incubation with ImmPACT™ DAB EqV Substrate. Finally, slides were counterstained with hematoxylin, followed by dehydration and coverslipping.

Bisulfite Sequencing (Kit and sequencing primers). Methylation analysis of BORIS (NCBI RefSeq NC_000020.11, spanning nucleotides chr20: 57,524,203- 57,525,234 on GRCh38.p7 assembly) was performed using a bisulphite sequencing assay. 500 ng of genomic DNA was treated with the EZ DNA Methylation-Lightning™ Kit (Zymo Research, Irvine, CA, USA), followed by PCR using ZymoTaq Polymerase premix (Zymo Research) and specific primers designed using the Zymo bisulfite primer seeker (http://www.zymoresearch.com/tools/bisulfite-primer-seeker/ - sequences available upon request). PCR products were then sequenced for assessment of CpG site-specific DNA methylation in the BORIS promoter region.

Growth assay. After shRNA-mediated knockdown of BORIS, cells were reseeded at a density of 4 x 10e5 cells per well in 6-well plates. At 48 and 120 h of incubation, cells were stained with trypan blue (Sigma-Aldrich) and counted on a Countess II cell counter (Life technologies).

Animal experiments. All animal experiments were performed with approval from the Institutional Animal Care and Use Committee (IACUC) of the DFCI. Three mouse experiments were performed: (i) To assess the tumorigenic potential of resistant cells *in vivo*, (ii) to assess

that resistance to TAE684 was maintained *in vivo* and (iii) to assess the effect of JQ1 on resistant cells *in vivo*. All experiments were performed using subcutaneous cell xenograft models generated by injecting 2 x 10e6 sensitive or resistant Kelly NB cells into the flanks of NU/NU (Crl:NU-Foxn1^{nu}) (Charles River Laboratories, Wilmington, MA) or NU/NU (CrTac:NCr-Foxn1^{nu}) (Taconic, Rensselaer, NY) 7 weeks female mice. (i) To assess the tumorigenic potential of resistant cells in absence of treatment, animals with established disease (mean tumor volume of 200 mm^3) were monitored for up to 23 days (*n* = 4 per group). Tumors were harvested, dissociated and used to establish cell lines and assessment of mRNA levels, protein expression and sensitivity to TAE684. (ii) To ensure that the *in vitro* resistance to TAE684 was maintained *in vivo*, animals with established disease were divided into two cohorts and were treated with either TAE684 (10 mg/kg) or vehicle control by oral gavage once daily (*n* = 8 per group), and were monitored for up to 56 days from start of treatment. (iii) To assess the sensitivity of resistant cells to BRD4 inhibition, animals with established disease were divided into two cohorts and treated with either JQ1 (50 mg/kg) or vehicle control intraperitoneally (i.p.) once daily (*n* = 6 per group), and were monitored for up to 87 days from start of treatment. For all experiments, disease burden was quantified using electronic caliper measurements (2 to 3 times a week) and mouse weights were monitored at least twice a week. Tumor volumes were calculated using the modified ellipsoid formula (½ (length x width^2)[44]. Animals were euthanized when tumor volumes reached 1500-2000 mm^3 based on institutional IACUC criteria for maximum tumor volumes. In none of the experiments were the institutional limits for tumor volumes (< 2000 mm^3 measurement preceding the day of sacrifice) exceeded.

ChIP-sequencing. ChIP was carried out as previously described[41] with minor changes as described below. 1 x 10e7 cells were crosslinked for 10 minutes at RT with 1% formaldehyde (Thermo Scientific) in PBS followed by quenching with 0.125 M glycine for 5 minutes. The cells were then washed twice in ice-cold PBS, and the cell pellets flash frozen and stored at -80^{o}C.

50 µl of Protein G Dynabeads® per sample (Invitrogen) were blocked with 0.02% Tween20 (w/v) in PBS. Magnetic beads were loaded with 10 µg each of antibody and incubated overnight at 4°C. Crosslinked cells were lysed, placed in sonication buffer with 0.2% SDS, placed on ice and chromatin was sheared using a Misonix 3000 sonicator (Misonix, Farmingdale, NY, USA) and at the following settings: 10 cycles, each for 30 seconds on, followed by 1 minute off, at a power of approximately 20 Watts. The lysates were then centrifuged for 10 min at 4°C, supernatants collected and diluted with an equal amount of sonication buffer to reach a final concentration of 0.1% SDS. The sonicated lysates were incubated overnight at 4°C with the antibody-bound magnetic beads, washed with low-salt buffer (50mM Hepes-KOH pH 7.5, 0.1% SDS, 1% Triton X-100, 0.1% sodium deoxycholate, 1mM EGTA, 1 mM EDTA, 140 mM NaCl and 1X complete protease inhibitor), high-salt buffer (50mM Hepes-KOH pH 7.5, 0.1% SDS, 1% Triton X-100, 0.1% sodium deoxycholate, 1mM EGTA, 1 mM EDTA, 500 mM NaCl and 1X complete protease inhibitor), LiCl buffer (20mM Tris-HCl pH 8, 0.5% NP-40, 0.5% sodium deoxycholate, 1mM EDTA, 250 mM LiCl and 1X complete protease inhibitor) and Tris-EDTA buffer. DNA was then eluted in elution buffer (50 mM Tris-HCl pH 8.0,10 mM EDTA, 1% SDS), and high-speed centrifugation performed to pellet the magnetic beads and collect the supernatants. The crosslinking was reversed overnight at 65°C. RNA and protein were digested using RNase A and Proteinase K, respectively, and DNA was purified with phenol chloroform extraction and ethanol precipitation. Purified ChIP DNA was used to prepare Illumina multiplexed sequencing libraries using the NEBNext® Ultra™ II DNA Library Prep kit and the NEBNext® Multiplex Oligos for Illumina (New England Biolabs, Ipswich, MA, USA) according to the manufacturer's protocol. Libraries with distinct indexes were multiplexed and run together on the Illumina NextSeq 500 (SY-415-1001, Illumina) for 75 bases in single-read mode.

HiChIP. HiChIP was performed as described in Mumbach et al[45] with a few modifications. 1 x 10e7 cells were cross-linked for 10 min at room temperature with 1% formaldehyde in growth

539 medium and quenched in 0.125 M glycine. After washing twice with ice-cold PBS, the
540 supernatant was aspirated and the cell pellet flash frozen in liquid nitrogen. Cross-linked cell
541 pellets were thawed on ice, resuspended in 1 mL of ice-cold Hi-C lysis buffer (10 mM Tris-HCl
542 pH 8.0, 10 mM NaCl, 0.2% NP-40 and 1X complete protease inhibitor), and incubated at 4^{o}C for
543 30 minutes with rotation. Nuclei were pelleted by centrifugation for 5 min at 4^{o}C and washed
544 once with 500 µL of ice-cold Hi-C lysis buffer. After removing the supernatant, nuclei were
545 resuspended in 100 µL of 0.5% SDS and incubated at 62^{o}C for 10 minutes. SDS was quenched
546 by adding 335 µL of 1.5% Triton X-100 and incubating for 15 minutes at 37^{o}C. After the addition
547 of 50 µL of 10X NEB Buffer 2 (New England Biolabs, B7002) and 375 U of MboI restriction
548 enzyme (New England Biolabs, R0147), chromatin was digested at 37^{o}C for 2 hours with
549 rotation. Following digestion, MboI enzyme was heat-inactivated by incubating the nuclei at
550 62^{o}C for 20 min. To fill in the restriction fragment overhangs and mark the DNA ends with biotin,
551 52 µL of fill-in master mix, containing 37.5 µL of 0.4 mM biotin-dATP (Invitrogen, 19524016), 1.5
552 µL of 10 mM dCTP (Invitrogen, 18253013), 1.5 µL of 10 mM dGTP (Invitrogen, 18254011), 1.5
553 µL of 10 mM dTTP (Invitrogen, 18255018), and 10 µL of 5 U/µL DNA Polymerase I, Large
554 (Klenow) Fragment (New England Biolabs, M0210), were added and the tubes incubated at
555 37^{o}C for 1 hour with rotation. Proximity ligation was performed by the addition of 948 µL of
556 ligation master mix, containing 150 µL of 10X NEB T4 DNA ligase buffer (New England Biolabs,
557 B0202), 125 µL of 10% Triton X-100, 7.5 µL of 20 mg/mL BSA (New England Biolabs, B9000),
558 10 µL of 400 U/µL T4 DNA ligase (New England Biolabs, M0202), and 655.5 µL of water, and
559 incubation at room temperature for 4 hours with rotation. After proximity ligation, nuclei were
560 pelleted by centrifugation for 5 minutes and resuspended in 1 mL of ChIP sonication buffer (50
561 mM HEPES-KOH pH 7.5, 140 mM NaCl, 1 mM EDTA pH 8.0, 1 mM EGTA pH 8.0, 1% Triton X-
562 100, 0.1% sodium deoxycholate, 0.1% SDS and 1X complete protease inhibitor). Nuclei were
563 sonicated using a Misonix 3000 sonicator (Misonix) and at the following settings: 12 cycles,
564 each for 30 seconds on, followed by 1 minute off, at a power of approximately 20 Watts.

Sonicated chromatin was clarified by centrifugation for 15 min at 4°C and the supernatant was transferred to a tube. 60 µL of protein G Dynabeads (Invitrogen) were washed three times and resuspended in 50 µL sonication buffer. Washed beads were then added to the sonicated chromatin and incubated for 1 hour at 4°C with rotation. Beads were then separated on a magnetic stand and the supernatant was transferred to a new tube. 75 µL of protein G Dynabeads pre-incubated overnight at 4°C with 10 µg of anti-SMC1A antibody (Bethyl A300-055A) or 10 µg of BORIS antibody (Abcam, ab187163) were added to the tube and incubated overnight at 4°C with rotation. Beads were then separated on a magnetic stand and washed twice with 1 mL of sonication buffer, followed by once with 1 mL high salt sonication buffer (50 mM HEPES-KOH pH 7.5, 500 mM NaCl, 1 mM EDTA pH 8.0, 1 mM EGTA pH 8.0, 1% Triton X-100, 0.1% sodium deoxycholate, 0.1% SDS), once with 1 mL of LiCl wash buffer (20 mM Tris-HCl pH 8.0, 1 mM EDTA pH 8.0, 250 mM LiCl, 0.5% NP-40, 0.5% sodium deoxycholate, 0.1% SDS) and once with 1 mL of TE with salt (10 mM Tris-HCl pH 8.0, 1 mM EDTA pH 8.0, 50 mM NaCl). Beads were then resuspended in 200 µL of elution buffer (50 mM Tris-HCl pH 8.0, 10 mM EDTA pH 8.0, 1% SDS) and incubated at 65°C for 15 minutes. To purify the eluted DNA, RNA was degraded by the addition of 8.5 µL of 10 mg/mL RNase A and incubation at 37°C for 2 hours. Protein was degraded by the addition of 20 µL of 10 mg/mL proteinase K and incubation at 55°C for 45 minutes. Samples were then incubated at 65°C overnight to reverse crosslink protein-DNA complexes. DNA was then purified using Zymo ChIP DNA Clean and Concentrator™ columns (Zymo, D5205) according to manufacturer's protocol and eluted in 14 µL water. The amount of eluted DNA was quantified by Qubit dsDNA HS kit (Invitrogen, Q32854). Tagmentation of ChIP DNA was performed using the Illumina Nextera DNA Library Prep Kit (Illumina, FC-121-1030). First, 5 µL of MyOne™ Streptavidin C1 Dynabeads (Invitrogen, 65001) was washed with 1 mL of Tween wash buffer (5 mM Tris-HCl pH 7.5, 0.5 mM EDTA pH 8.0, 1 M NaCl, 0.05% Tween-20) and resuspended in 10 µL of 2X biotin binding buffer (10 mM Tris-HCl pH 7.5, 1 mM EDTA pH 8.0, 2 M NaCl). 25 ng of purified DNA was

591 added in a total volume of 10 µL water to the beads and incubated at room temperature for 15
592 minutes with agitation every 5 minutes. After capture, beads were separated with a magnet and
593 the supernatant was discarded. Beads were then washed twice with 500 µL of Tween wash
594 buffer, incubating at 55^{o}C for 2 minutes with shaking for each wash. Beads were resuspended in
595 25 µL of Nextera Tagment DNA buffer. To tagment the captured DNA, 1 µL of Nextera Tagment
596 DNA Enzyme 1 was added with 24 µL of Nextera Resuspension Buffer and samples were
597 incubated at 55^{o}C for 10 minutes with shaking. Beads were separated on a magnet and
598 supernatant was discarded. Beads were washed twice with 500 µL of 50 mM EDTA at 50^{o}C for
599 30 minutes, washed twice with 500 µL of Tween wash buffer at 55^{o}C for 2 minutes each, and
600 finally washed once with 500 µL of 10 mM Tris-HCl pH 7.5 for 1 minute at room temperature.
601 Beads were separated on a magnet and supernatant was discarded. To generate the
602 sequencing library, PCR amplification of the tagmented DNA was performed while the DNA was
603 still bound to the beads. Beads were resuspended in 15 µL of Nextera PCR Master Mix, 5 µL of
604 Nextera PCR Primer Cocktail, 5 µL of Nextera Index Primer 1, 5 µL of Nextera Index Primer 2,
605 and 20 µL water. DNA was amplified with 9-10 cycles of PCR. After PCR, beads were
606 separated on a magnet and the supernatant containing the PCR-amplified library was
607 transferred to a new tube, purified using Zymo DNA Clean and Concentrator^{TM} columns (Zymo,
608 D5205) according to the manufacturer's protocol, and eluted in 14 µL water. Purified HiChIP
609 libraries were size selected to 300-700 bp using a Sage Science Pippin Prep instrument
610 according to the manufacturer's protocol and subjected to 2 x 100 paired-end sequencing using
611 an Illumina HiSeq 2500 system (SY–401–2501, Illumina).

612 **Single-cell RNA sequencing (scRNA-seq).** Kelly cells (sensitive, intermediate and resistant
613 states) were grown to 70% confluence in T75 culture flasks. Briefly, growth medium was
614 aspirated and cells were treated with 0.25% Trypsin/EDTA for 3 minutes at 37^{o}C, after which
615 cells were washed twice with 1X PBS. Cells were then resuspended into single cells at a

concentration of 1 x 10e6 per ml in 1X PBS with 0.4% BSA for 10x genomics processing. The sorted cell suspensions were loaded onto a 10x Genomics Chromium instrument to generate single-cell gel beads in emulsion (GEMs). Approximately 5,000 cells were loaded per channel. scRNA-seq libraries were prepared using the following Single Cell 3′ Reagent Kits: Chromium™ Single Cell 3′ Library & Gel Bead Kit v2 (PN-120237), Single Cell 3′ Chip Kit v2 (PN-120236) and i7 Multiplex Kit (PN-120262) (10x Genomics, Pleasanton, CA, USA) as previously described[46], and following the Single Cell 3′ Reagent Kits v2 User Guide (Manual Part # CG00052 Rev A). Libraries were run on an Illumina HiSeq 4000 system (SY-401-4001, Illumina) as 2 x 150 paired-end reads, one full lane per sample, for approximately > 90% sequencing saturation.

Genomics analysis.

Direct comparison of CTCF and BORIS expression in normal and tumor samples. To assess the expression levels and range of BORIS and CTCF in normal and tumor cells all GTEx, TCGA and TARGET datasets were downloaded and converted to FPKM values and displayed as [$\log_2$ (FPKM + 1)] (**Extended Data Fig. 1a, b**).

Association of BORIS with prognostic features. For each dataset processed values were extracted from GEO and scaled values were created by normalizing the expression levels by the minimum mean value of the conditions that were compared, Esi,j = Ei,j / min(average(Ej). The two-sided Wilcoxon rank-sum test on the original values was used to determine statistical differences between the compared conditions (**Extended Data Fig. 1c** and **Extended Data Fig. 4f**).

Microarray Data Analysis. Microarray data were analyzed using a custom CDF file (GPL16043) that contained the mapping information of the ERCC probes used in the spike-in

RNAs. The arrays were normalized as previously described[39]. Briefly, all chip data were imported in R (https://www.r-project.org/, v3.1.3) using the affy package[47] (v1.44.0), converted into expression values using the expresso command, normalized to take into account the different numbers of cells and spike-ins used in the different experiments and renormalized using loess regression fitted to the spike-in probes. Sets of differentially expressed genes were obtained using the limma package[48] (v3.22.7) and a False Discovery Rate (FDR) of 0.05. Spike-in normalized absolute expression values (counts) were normalized to counts per million (CPM) as a measurement of relative gene expression concentrations per condition. Total number of transcripts per sample was determined as the total number of counts after spike-in normalization and the shBORIS sample was first normalized to the control shCtrl sample to account for technical effects that originated from the transfection protocol.

ChIP-seq Analysis. For all ChIP-seq samples high-quality data were confirmed using the Fastqc tool (v0.11.5) and samples were aligned to the human genome (build hg19, GRCh37.75) with STAR (v2.5.1b_modified) and the parameters "--alignIntronMax 1 --alignEndsType EndToEnd--outFilterMultimapNmax 1 --outFilterMismatchMax 5". Next, non-duplicate reads that mapped to the reference chromosomes were retained using Samtools (v1.3.1) and MarkDuplicates (v2.1.1) from Picard tools. For each experimental replicate, antibody enrichment was assessed using the plotFingerprint command from deepTools (v2.2.4). Peaks were identified with MACS2 (2.1.1) for narrow peaks (BORIS, CTCF, BRD4, Pol2, MYCN) with the parameters "--q 0.01--call-summits" and for broad peaks (H3K27ac, H3K27me3) with the parameters "--broad-cutoff 0.01". Peaks overlapping regions with known artefact regions (http://mitra.stanford.edu/kundaje/akundaje/release/blacklists/) were blacklisted out. Input normalized bedgraph tracks were created with the deepTools command bamCompare and the parameters "--scaleFactorsMethod=readCount --ratio=subtract --binSize=50 --numberOfProcessors=4 --extendReads=200". Subsequently, negative values were set to zero

and counts were scaled to rpm/bp to account for differences in library size. Bigwig files were created with bedGraphToBigWig (v4). ChIP-seq replicates (*n* = 2) were merged at the BAM level after assessment of strong correlation with the deepTools command "multiBigwigSummary BED-file" using all replicate bigwigs and identified peaks as input. Identification of peaks and generation of tracks were then repeated for these merged files and used for further analyses. Downstream analyses for ChIP-seq and other genomic interval data was performed in R (https://www.r-project.org/, (v3.5.1) using the data.table (v1.12.2) package.

Gencode annotation and isoform selection. Gencode (http://www.gencodegenes.org/, release 19) annotation was used and for each gene the most likely isoform was selected based on data-driven criteria. Briefly, only genes that were part of the Refseq transcriptome annotation and with a minimum length of 1 kb were considered. Next, isoforms were prioritized according to increased deposition of Pol2 and H3K27ac reads on transcript start sites (TSS), transcript length and alphabet rank, in that order, until only 1 transcript was selected for each gene.

Cell-type specific binding patterns. To determine the cell-specificities of BORIS and CTCF peaks we first combined all peaks identified by MACS2 and merged the peak regions that overlapped by at least 50%. A 50% threshold was empirically selected to avoid merging peaks that had clear and distinct summits. Next, normalized BORIS or CTCF read densities were calculated for each region and a ratio [$\log_2$ (resistant/sensitive)] was calculated. Peak regions with a 2-fold density increase or decrease were classified as resistant- or sensitive-cell-specific peaks respectively, while other regions were denoted as 'shared' to indicate that these peaks had similar BORIS or CTCF deposition in both cell types (**Fig. 2a, b** and **Extended Data Fig. 6a**). To further explore the proximity of BORIS and CTCF peaks and how they were altered during the transition from sensitive to resistant cells, we overlapped all shared and cell type-specific peaks from both cell types in the least stringent way (minimum 1 bp overlap) (**Fig. 2c** and **Extended Data Fig. 6a**).

Genomic enrichment of peak binding sites. To identify genomic locations with BORIS or CTCF binding we determined the number of peaks that overlapped with at least 25% of known functional regions in the following order: (i) broad promoter (± 2 kb TSS), (ii) BRD4+ H3K27ac+ (active) enhancers, (iii) $BRD4^{-}$ $H3K27ac^{+}$ enhancers, (iv) exons, (v) introns, (vi) repressed chromatin represented by H3K27me3 broad peaks or (vii) other if the peak was outside the aforementioned regions (**Extended Data Fig. 6d**). Enrichment of ChIP-seq binding at resistant cell BORIS peaks was performed by extending BORIS summits by 1 kb in both directions and calculating the normalized read densities in 50 bp bins (**Fig. 2d**).

Genomic enrichment of regulatory regions. To further visualize the enrichment of CTCF and BORIS at regulatory regions (enhancers and promoters) and the differences between sensitive and resistant cells, a metagene analysis for CTCF and BORIS occupancies was performed for all H3K27ac enhancer regions and gene promoters. All TSS were extended in both directions by 2 kb and binned in 50 bp bins, while each enhancer (start – end) was divided into 40 equally sized bins and extended with 2 kb in both directions and these extended regions were binned in 50 bp bins. Normalized bedgraph files were used to calculate read density (rpm/bp). An aggregated summary profile was created for each cell type. To account for different numbers of identified enhancers in both cells types we calculated a normalization factor (= N Res enhancers/N Sens enhancer) to divide each aggregated read density (**Extended Data Fig. 6e**).

HiChIP processing and quality control. For all SMC1A-based HiChIP datasets, raw reads were first trimmed to a uniform length of 50 bp using *trimmomatic*[49](v0.36) and were then processed using the *HiC-Pro* (v2.10.0) pipeline[50] with default settings for the human genome (build hg19, GRCh37.75) and corresponding MboI cut sites. To perform intra and inter correlation analysis for biological replicates, forward and reverse reads from the HiC-Pro output were merged together to generate 1-Dimensional SMC1A BAM profiles. Genome-wide Spearman correlation in 5 kb bins was computed for all merged genomic anchor regions on

those merged BAMs for all replicates using the '*multiBamSummary BED-file*' command from deepTools (**Extended Data Fig. 7a, e**).

HiChIP loop calling and differential looping analysis. Loops were directly called from the HiC-Pro output using hichipper[51] (v0.7.3), with parameter 'peaks = combined, all', and subsequently diffloop[51] (v1.10.0) with default settings. Only loops that were detected in all 3 biological replicates of a sample (sensitive, resistant, shGFP or shBORIS) with a minimum of 5 paired-end tags (PETs) in total and an FDR $\leq$ 0.01 were retained for further analysis. To call differential loops between samples, the quickAssocVoom function was used and significantly different loops were either considered reinforced (mango.FDR < 0.01 and $\log_2$FC > 1) or lost (mango.FDR < 0.01 and $\log_2$FC < -1).

Classification of HiChIP interactions. SMC1A based HiChIP interactions (loops) were classified as previously described[52] with minor adaptations. Associated anchors of loops were overlapped with our ChIP-seq peaks (CTCF, BORIS, H3K27ac, BRD4) and promoter regions (TSS ± 2 kb), requiring a minimum 1 bp overlap. Each anchor was then independently classified according to its overlap profile, following a hierarchical tree. If an anchor overlapped a promoter, an enhancer (BRD4 + H3K27ac), or a CTCF peak, it was classified as promoter-, enhancer- or CTCF-anchor, in that order. If there was no overlap, the anchor was considered 'other'. By combining these 4 anchor classes we discriminated 10 different interaction classes. We excluded from further analyses any interaction that contained an anchor classified as other, which also represented on average much shorter interactions (data not shown), and which were hence more likely to have occurred due to linear proximity on the DNA. This resulted in the identification of 6 main interaction classes (**Fig. 3a** and **Extended Data Fig. 7b**).

Association of BORIS with lost loops. Only loops that were detected in both the original (Sens vs. Res) and BORIS depletion (shBORIS vs. shGFP) samples were used for this analysis. First, loops were divided into lost and retained loops upon BORIS depletion, and an

odds-ratio (two-sided Fisher's exact test) was calculated for the initial presence of BORIS binding on the anchors of these two groups (**Fig. 3f**). An analogous strategy was followed after first stratifying loops according to the different identified loop classes (**Extended Data Fig. 7f, g**).

Identification of super-enhancer regions. Super-enhancers (SEs) were identified employing the ROSE algorithm (v1) (https://bitbucket.org/young_computation/rose). In short, H3K27ac enriched regions were identified with MACS2 and termed typical enhancers (TEs). These regions were stitched together if they were within 12.5 kb of each other. Stitched regions were ranked by H3K27ac signal therein and the inclination point at which the two classes of enhancers separated was determined by ROSE. Stitched enhancers above this threshold were considered super-enhancers and the others, typical enhancers. To compare different samples, we used the same maximum threshold between the conditions considered (**Extended Data Fig. 8a**).

Identification of cell-type specific super-enhancers. Cell-type specific and active SEs were identified by merging both sensitive- and resistant-cell SEs and determining cell-type specificity based on the differential normalized read density of both H3K27ac and BRD4. Briefly, ratios [$\log_2$ (resistant/sensitive)] were calculated for H3K27ac and BRD4. A combined threshold of 2.5 was required to identify a cell type-specific SE with at least a minimum 0.75 change for each individual mark. SEs that did not meet these criteria were classed as shared (neutral) between cell-types (**Extended Data Fig. 8b**).

Correlation analysis of looping with gene expression and enhancer landscape. Regulatory interactions were associated to target genes and SEs based on proximity to the TSS and minimal overlap (1 bp) with its anchors respectively (**Fig. 4a** and **Extended Data Fig. 8f**).

Chromatin-based gene classification. Unsupervised metagene clustering on H3K27ac and H3K27me3. Genes were classified as having an "open", "neutral" or "closed" chromatin state based on unsupervised clustering of a metagene representation of ChIP-seq occupancy of H3K27ac and H3K27me3. Each gene (from TSS to TES, and 2 kb up- and downstream of this region) was divided into 20 equally sized bins; the extended regions were binned in regions of 50 bp. Normalized bedgraph files were used to calculate read density (rpm/bp) and k-means clustering was applied to group each extended gene region in one of three clusters (**Extended Data Fig. 8d, e**). An aggregated summary profile was created for each group of genes. The "open" and "closed" clusters were classified based on predominantly H3K27ac and H3K27me3 accumulation, respectively, while the "neutral" cluster displayed on average equal levels of both.

Integrated genomic data analysis. An ensemble analysis was performed to identify the set of genes that showed characteristics of re-activation in resistant cells. For each gene, five features were examined: 1) creation of a unique regulatory interaction, 2) deposition of BORIS on its promoter or looped enhancer, 3) association with a resistant cell-specific SE through overlap with either its promoter or looped anchor, 4) increased mRNA expression, and 5) transition from a closed or neutral state to an open chromatin state. A unique set of 89 genes (**Supplementary Information Table**) that exhibited 4 out of 5 features were identified as the top reactivated genes in resistant cells. Within these 89 genes, 13 were identified as transcription factors by the TcoF database (http://www.cbrc.kaust.edu.sa/tcof/) (**Fig. 4b**).

Allen Brain atlas gene signature. Expression data and metadata for human brain development was downloaded from the Allen Brain atlas (http://www.brainspan.org). Row normalized z-scores of [$\log_2$ (RPKM + 1)] values were used to create a heatmap. Values greater than 3.5 were set to 3.5 to reduce the effect of extreme outliers on the visualization. Samples were ordered according to developmental time points (**Extended Data Fig. 8g**).

BORIS and BRD4 correlation at promoter regions. BORIS and BRD4 colocalization and correlation were assessed for the promoter regions of the 89 top ranked genes. The TSS was extended in both directions by 2 kb and binned in 100 bp regions. Normalized read densities for BORIS and BRD4 were calculated and a Spearman's rank correlation coefficient calculated for sensitive and resistant cells. An aggregated density plot of all 89 genes was created to visualize the increased deposition and correlation of BRD4 and BORIS in resistant cells (**Extended Data Fig. 9a**).

Gene expression and DNA binding analysis. To examine the association between gene expression and overlapping targets of MYCN and BORIS in sensitive and resistant cells respectively, the percentage of gene promoters (± 2 kb TSS) that overlapped with ChIP-seq peaks in 10 equally sized bins based on the expression distribution was calculated (**Extended Data Fig. 6f**). To visualize and correlate gene expression with DNA binding of MYCN or BORIS, genes were ranked based on expression and plotted against the total rescaled (0-100) binding intensities calculated for each gene promoter (± 2 kb TSS). For each ChIP-seq mark a loess regression curve was computed using a span of 0.1 (**Extended Data Fig. 6g**).

Transcription factor (TF) motif enrichment analysis. Statistically overrepresented motifs were identified with HOMER[53] (v2) using the command findMotifs.pl providing both target and background fasta sequences for regions of interest. For promoter regions we selected the top 2,000 up- and down-regulated genes in resistant cells and extended the TSS of each gene by 2 kb in both directions. The genomic coordinates were used to extract fasta sequences with the Biostrings package (v2.50.1) in R and used as target or background to identify motifs associated with promoter regions of genes within each cell type. A similar strategy was followed to identify overrepresented motifs associated with cell-type specific SEs. Target and background fasta sequences were extracted from the summits of BRD4 peaks located on cell-type specific SEs

and extended by 500 bp in both directions. For a selection of enriched sequences, the associated TF motif and significance level (*P*) was visualized using a heatmap (**Fig. 4d**).

Single-cell RNA-seq Analysis. The Cell Ranger Single Cell Software Suite, version 1.3 was used to perform sample de-multiplexing, barcode and UMI processing, and single-cell 3′ gene counting. A detailed description of the pipeline and specific instructions to run it can be found at: https://support.10xgenomics.com/single-cell-gene-expression/software/pipelines/latest/what-is-cell-ranger)**.** A high quality gene expression matrix was created in sequential preprocessing steps. First UMI-based counts were converted to relative expression concentrations by rescaling each cell to a library size of 10,000. Genes were considered detected if rescaled count > $\log_2$ (0.1 + 1) and retained for further analysis if present in at least 0.5% of the cells from the sample with the lowest cell count. Cells were removed if less than 1,000 genes were detected. To remove low quality cells, we calculated five technical indicators (ratio of detected genes/UMI, percent of mitochondrial genes, percent of ribosomal genes, average GC-content of library and library complexity measured by Shannon Entropy) and performed principal component analysis (PCA) on indicators with a coefficient of variation > 5%. Next, density-based clustering was performed on the first and second principal component using an epsilon determined by a k-nearest neighbor plot. All cells that were located outside the main cluster were considered low quality and removed from further analysis. Next, we used the R package 'scater' (v1.10.0) to confirm that there were no technical or experimental confounding effects and the R package 'Seurat' (v2.3.4) to analyze and visualize the data. In brief, UMI-counts were log-normalized with a scale factor of 10,000 and subsequently center-scaled. To visualize cells in a reduced dimensionality, PCA was performed on the most variable genes, which were identified as genes with higher-than-expected variability in consecutive ranked expression bins. Higher complexity clustering was performed with t-SNE using the first 10 principal components, which were deemed most informative based on heatmap and elbow plot observation. To identify

836 homogeneous subpopulations, we performed iterative clustering using the network-based
837 clustering algorithm (shared nearest neighbor) with different resolutions as input until each
838 sample was at least separated in 2 groups. A simple pseudotime analysis was performed by
839 calculating an average expression profile for each identified subpopulation and ordering them
840 according to the summarized expression of TFs that displayed variable expression between
841 sensitive and intermediate or intermediate and resistant cells. Variable expression was defined
842 as showing at least a 33% change in the rank of expression between two samples with a
843 minimal normalized expression level > 0.2. For each sample comparison, at least the top 10
844 most variable TFs were included. In total this resulted in 32 TFs. Gene expression values were
845 then linearly rescaled between 0 and 10 to jointly visualize relative expression changes during
846 this pseudotime. To examine co-detection or mutual exclusivity between genes of interest, a
847 two-sided Fisher's exact test was performed for all cells in a given sample. A score combining
848 both the odds-ratio and the $-\log_{10}$ (*P*-value) was calculated to visualize both the strength and
849 direction between genes in pair-wise co-expression tests.

850 **Statistical analysis.** Analysis for each plot is listed in the figure legend and/or in the
851 corresponding Methods. Briefly, all grouped data are presented as mean ± SD unless stated
852 otherwise. All box and whisker plots of expression data are presented as: centre lines, medians;
853 box limits, 25^{th} and 75^{th} percentiles; whiskers, minima and maxima (1.5X the interquartile
854 range). Statistical significance for pair-wise comparisons was determined using the two-sided
855 Wilcoxon rank-sum test or two-sided unpaired *t*-test, unless stated otherwise. Survival analysis
856 was performed using the Kaplan-Meier method and differences between groups calculated by
857 the two-sided log-rank test and the Bonferroni correction method. Tumor volume comparisons
858 for the xenograft studies were analyzed by Mann-Whitney U test. (*, $P < 0.05$; **, $P < 0.01$).
859 Statistical comparisons of distributions of fold changes for the expression microarrays were
860 done using the Mann-Whitney U test. All quantitative analyses are expressed as the mean ± SD

of three biological replicates, unless stated otherwise. Microarray and ChIP-seq data are based on at least 2 independent experiments. For all experiments, no statistical methods were used to predetermine sample size, and the experiments were not randomized. The investigators were not blinded to allocation during experiments and outcome assessment.

Track visualizations. Peaks, (super-) enhancers and HiChIP interactions were visualized with a custom build tool (github.com/RubD/GeTrackViz2) or with the circlize package (v0.4.5) in R.

Retrospective analysis of gene expression in human samples. Gene expression levels or correlations across primary tumors, healthy tissues or experimental data and patient survival were determined through analysis of the TCGA and TARGET (https://cancergenome.nih.gov/), GTEx (https://www.gtexportal.org/home/), R2 (https://hgserver1.amc.nl/cgi-bin/r2/main.cgi), Allen Brain atlas (http://www.brain-map.org/) and selected datasets representing distinct tumor types with poor prognosis feature annotations [GSE49710 (Neuroblastoma)[54], GSE17679 (Mixed Ewing Sarcoma)[55], GSE63074 (Non-small cell lung carcinoma)[56], GSE15709 (ovarian cancer)[57], GSE16179 (breast cancer)[58] and GSE7181 (Glioblastoma)[59]].

31 Batra, S., Reynolds, C. P. & Maurer, B. J. Fenretinide cytotoxicity for Ewing's sarcoma and primitive neuroectodermal tumor cell lines is decreased by hypoxia and synergistically enhanced by ceramide modulators. *Cancer research* **64**, 5415-5424, doi:10.1158/0008-5472.CAN-04-0377 (2004).

32 Whang-Peng, J. *et al.* Cytogenetic characterization of selected small round cell tumors of childhood. *Cancer genetics and cytogenetics* **21**, 185-208 (1986).

33 Galkin, A. V. *et al.* Identification of NVP-TAE684, a potent, selective, and efficacious inhibitor of NPM-ALK. *Proceedings of the National Academy of Sciences of the United States of America* **104**, 270-275, doi:10.1073/pnas.0609412103 (2007).

34 Gao, Y. *et al.* Overcoming Resistance to the THZ Series of Covalent Transcriptional CDK Inhibitors. *Cell chemical biology* **25**, 135-142 e135, doi:10.1016/j.chembiol.2017.11.007 (2018).

35 Filippakopoulos, P. *et al.* Selective inhibition of BET bromodomains. *Nature* **468**, 1067-1073, doi:10.1038/nature09504 (2010).

36 Marsilje, T. H. *et al.* Synthesis, structure-activity relationships, and in vivo efficacy of the novel potent and selective anaplastic lymphoma kinase (ALK) inhibitor 5-chloro-N2-(2-isopropoxy-5-methyl-4-(piperidin-4-yl)phenyl)-N4-(2-(isopropylsulf onyl)phenyl)pyrimidine-2,4-diamine (LDK378) currently in phase 1 and phase 2 clinical trials. *Journal of medicinal chemistry* **56**, 5675-5690, doi:10.1021/jm400402q (2013).

37 Johnson, T. W. *et al.* Discovery of (10R)-7-amino-12-fluoro-2,10,16-trimethyl-15-oxo-10,15,16,17-tetrahydro-2H-8,4-(m etheno)pyrazolo[4,3-h][2,5,11]-benzoxadiazacyclotetradecine-3-carbonitrile (PF-06463922), a macrocyclic inhibitor of anaplastic lymphoma kinase (ALK) and c-ros oncogene 1 (ROS1) with preclinical brain exposure and broad-spectrum potency against ALK-resistant mutations. *Journal of medicinal chemistry* **57**, 4720-4744, doi:10.1021/jm500261q (2014).

38 Gosmini, R. *et al.* The discovery of I-BET726 (GSK1324726A), a potent tetrahydroquinoline ApoA1 up-regulator and selective BET bromodomain inhibitor. *Journal of medicinal chemistry* **57**, 8111-8131, doi:10.1021/jm5010539 (2014).

39 Loven, J. *et al.* Revisiting global gene expression analysis. *Cell* **151**, 476-482, doi:10.1016/j.cell.2012.10.012 (2012).

40 Debruyne, D. N. *et al.* ALK inhibitor resistance in ALK(F1174L)-driven neuroblastoma is associated with AXL activation and induction of EMT. *Oncogene* **35**, 3681-3691, doi:10.1038/onc.2015.434 (2016).

41 Chipumuro, E. *et al.* CDK7 inhibition suppresses super-enhancer-linked oncogenic transcription in MYCN-driven cancer. *Cell* **159**, 1126-1139, doi:10.1016/j.cell.2014.10.024 (2014).

42 Sancak, Y. *et al.* The Rag GTPases bind raptor and mediate amino acid signaling to mTORC1. *Science* **320**, 1496-1501, doi:10.1126/science.1157535 (2008).

43 Meerbrey, K. L. *et al.* The pINDUCER lentiviral toolkit for inducible RNA interference in vitro and in vivo. *Proceedings of the National Academy of Sciences of the United States of America* **108**, 3665-3670, doi:10.1073/pnas.1019736108 (2011).

44 Tomayko, M. M. & Reynolds, C. P. Determination of subcutaneous tumor size in athymic (nude) mice. *Cancer chemotherapy and pharmacology* **24**, 148-154 (1989).

45 Mumbach, M. R. *et al.* HiChIP: efficient and sensitive analysis of protein-directed genome architecture. *Nature methods* **13**, 919-922, doi:10.1038/nmeth.3999 (2016).

46 Zheng, G. X. *et al.* Massively parallel digital transcriptional profiling of single cells. *Nature communications* **8**, 14049, doi:10.1038/ncomms14049 (2017).

47 Gautier, L., Cope, L., Bolstad, B. M. & Irizarry, R. A. affy--analysis of Affymetrix GeneChip data at the probe level. *Bioinformatics* **20**, 307-315, doi:10.1093/bioinformatics/btg405 (2004).

48 Smyth, G. K., Yang, Y. H. & Speed, T. Statistical issues in cDNA microarray data analysis. *Methods in molecular biology* **224**, 111-136, doi:10.1385/1-59259-364-X:111 (2003).
49 Bolger, A. M., Lohse, M. & Usadel, B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* **30**, 2114-2120, doi:10.1093/bioinformatics/btu170 (2014).
50 Servant, N. *et al.* HiC-Pro: an optimized and flexible pipeline for Hi-C data processing. *Genome biology* **16**, 259, doi:10.1186/s13059-015-0831-x (2015).
51 Lareau, C. A. & Aryee, M. J. diffloop: a computational framework for identifying and analyzing differential DNA loops from sequencing data. *Bioinformatics* **34**, 672-674, doi:10.1093/bioinformatics/btx623 (2018).
52 Ji, X. *et al.* 3D Chromosome Regulatory Landscape of Human Pluripotent Cells. *Cell stem cell* **18**, 262-275, doi:10.1016/j.stem.2015.11.007 (2016).
53 Heinz, S. *et al.* Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. *Molecular cell* **38**, 576-589, doi:10.1016/j.molcel.2010.05.004 (2010).
54 Consortium, S. M.-I. A comprehensive assessment of RNA-seq accuracy, reproducibility and information content by the Sequencing Quality Control Consortium. *Nature biotechnology* **32**, 903-914, doi:10.1038/nbt.2957 (2014).
55 Savola, S. *et al.* High Expression of Complement Component 5 (C5) at Tumor Site Associates with Superior Survival in Ewing's Sarcoma Family of Tumour Patients. *ISRN oncology* **2011**, 168712, doi:10.5402/2011/168712 (2011).
56 Huang, S. *et al.* Analytical Performance of a 15-Gene Prognostic Assay for Early-Stage Non-Small-Cell Lung Carcinoma Using RNA-Stabilized Tissue. *The Journal of molecular diagnostics : JMD* **17**, 438-445, doi:10.1016/j.jmoldx.2015.03.005 (2015).
57 Li, M. *et al.* Integrated analysis of DNA methylation and gene expression reveals specific signaling pathways associated with platinum resistance in ovarian cancer. *BMC medical genomics* **2**, 34, doi:10.1186/1755-8794-2-34 (2009).
58 Liu, L. *et al.* Novel mechanism of lapatinib resistance in HER2-positive breast tumor cells: activation of AXL. *Cancer research* **69**, 6871-6878, doi:10.1158/0008-5472.CAN-08-4490 (2009).
59 Beier, D. *et al.* CD133(+) and CD133(-) glioblastoma-derived cancer stem cells show differential growth characteristics and molecular profiles. *Cancer research* **67**, 4010-4015, doi:10.1158/0008-5472.CAN-06-4180 (2007).

Data availability.

The microarray, ChIP-seq, HiChIP and scRNA-seq datasets generated and analyzed during the current study are available in the Gene Expression Omnibus repository under accession number GSE103084. The authors declare that all other data supporting the findings of this study are available within the paper and its Supplementary Information files.

Code availability

Custom code is available upon reasonable request.

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgments

We thank members of the George, Young and Gray laboratories and J. R. Gilbert for insightful discussions, C. Li for assistance with the HiChIP experiments. We thank J. Qi for providing JQ1, and D. Sabatini (MIT) and S. Elledge (HMS) for sharing plasmids (pLKO.1 GFP shRNA and pInducer20, respectively). We thank A. Ward and C. Clinton at DFCI Pediatric Oncology and J. Chan and the Clark Smith Tumor Bank, Charbonneau Cancer Institute, Calgary, Canada, for help with human tumor samples. We thank Applied Pathology for performing the IHC experiments, the DFCI Molecular Biology Core Facilities and the Whitehead Genome Technology Core for providing genomics services, and D. Adeegbe and the NYULH Genome Technology Center for performing the scRNA-sequencing. The results shown here are in part based upon data generated by the TCGA Research Network: http://cancergenome.nih.gov/, the R2: Genomics Analysis and Visualization Platform: http://r2.amc.nl/ and the Allan Brain Map Data Portal: http://www.brain-map.org/. This work was supported by NIH grants R01CA197336 (R.E.G. and R.A.Y.), R01CA148688 (R.E.G. and N.S.G.) and a Hyundai Hope on Wheels Scholar Grant (R.E.G.). D.N.D. is a recipient of a Young Investigator Grant from the Alex's Lemonade Stand Foundation/Northwestern Mutual Foundation. D.S.D is supported by an American Cancer Society fellowship PF-16-146-01-DMC. The NYULH Genome Technology Center is partially supported by the Cancer Center Support Grant P30CA016087 at the Laura and Isaac Perlmutter Cancer Center.

Author Contributions

D.N.D and R.E.G. conceived the project and designed the experiments. D.N.D. planned and performed the molecular, cellular and genomic studies. R.D. performed computational analyses with input from D.S.D. and E.M. S.S. contributed to the ChIP-seq and HiChIP experiments. D.S. and S.H.O. contributed to the HiChIP experiment. D.N.D., Y.G. and T.C. performed the animal experiments. B.S. and M.M. provided technical assistance. H.H. performed the co-IP experiment. L.M. performed the FISH analysis. N.S.G. provided TAE684 and E9. G-C.Y. supervised the bioinformatics analyses. K.K.W. enabled the animal and scRNA-seq studies. D.N.D, R.D., R.A.Y and R.E.G. interpreted the data. D.N.D., R.D. and R.E.G wrote the manuscript with input from R. A. Y. R.E.G. supervised the research. All authors edited the manuscript.

Competing interests

N.S.G. is a founder, SAB member and equity holder of Gatekeeper, Syros Pharmaceuticals. Petra, C4, B2S and Soltego. The Gray lab receives or has received research funding from Novartis, Takeda, Astellas, Taiho, Janssen, Kinogen, Voronoi, Her2llc, Deerfield and Sanofi. S.H.O. is a SAB member of Syros. R.A.Y. is a founder and shareholder of Syros, Camp4

Therapeutics, Omega Therapeutics and Dewpoint Therapeutics. R.E.G. is a SAB member of Global Gene Corp.

Correspondence and requests for materials should be addressed to rani_george@dfci.harvard.edu.

Extended Data Figure Legends

Extended Data Figure 1. BORIS is expressed in several cancers and associated with high-risk features. a, Relative mRNA expression [log2 (FPKM + 1)] of CTCF and BORIS in normal tissues, and **b,** in various cancer types based on TCGA datasets (see **Supplementary Information** for keys to cancer types). FPKM, fragments per kilo base of transcript per million mapped reads. Keys to cancer types, OV, serous ovarian cystadenocarcinoma; UCS, uterine carcinosarcoma; CESC, cervical squamous cell carcinoma and endocervical adenocarcinoma; LUSC, lung squamous cell carcinoma; WT, Wilms tumor; UCEC, uterine corpus endometrial carcinoma; SKCM, skin cutaneous melanoma; ESCA, esophageal carcinoma; STAD, stomach adenocarcinoma; DLBC, diffuse large B-cell lymphoma; HNSC, head and neck squamous cell carcinoma; NB, neuroblastoma; TGCT, testicular germ cell tumor; LUAD, lung adenocarcinoma; BLCA, bladder urothelial carcinoma; AML, acute myeloid leukemia; READ, rectum adenocarcinoma; MESO, mesothelioma; THYM, thymoma; LIHC, hepatocellular carcinoma; RT, rhabdoid tumor; GBM, glioblastoma multiforme; KIRC, renal clear cell carcinoma; SARC, sarcoma; LAML, acute myeloid leukemia; BRCA, breast invasive carcinoma; PAAD, pancreatic adenocarcinoma; COAD, colon adenocarcinoma; KIRP, kidney renal papillary cell carcinoma; THCA, thyroid carcinoma; CHOL, cholangiocarcinoma; PCPG, pheochromocytoma and paraganglioma; LGG, low-grade glioma; PRAD, prostate adenocarcinoma; KICH, kidney chromophobe; ACC, adrenocortical carcinoma; UVM, uveal melanoma. **c,** Box plots showing the correlation of BORIS expression with risk status, tumor stage (primary vs. metastasis/recurrence), presence of cancer stem cells (CD133 positivity) and response to targeted (lapatinib) or cytotoxic (cisplatin) therapy in the tumor types depicted. NSCLC, non-small cell lung cancer. Datasets (Mixed Ewing Sarcoma-Savola-117 and NSCLC-Plamadeala-410) were extracted from the R2: Genomics Analysis and Visualization Platform (http://r2.amc.nl). GSE7181 (glioblastoma); GSE16179 (breast cancer); GSE15372 (ovarian

cancer). The two-sided Wilcoxon rank-sum test was used for all pairwise comparisons. For all panels, sample sizes (*n*) are depicted in parenthesis and box plots are defined by centre lines, medians; box limits, 25^{th} and 75^{th} percentiles; whiskers, minima and maxima (1.5X the interquartile range of the box).

Extended Data Figure 2. ALK inhibitor-resistant cells exhibit stable resistance *in vivo* and no longer rely on ALK signaling. a, Left, Tumor volumes of sensitive and resistant cell xenografts in untreated NU/NU (Crl:NU-Foxn1nu) mice established by subcutaneous injection of 2 x 10^6 cells into both flanks. Animals were euthanized when tumors reached 1,500-2,000 mm^3. Data are means ± SEM, *n* = 4 per arm. **Right,** Immunoblot analysis of total and phosphorylated ALK in TAE-resistant xenograft tumors (1 and 2) and sensitive and resistant cells in culture. **b,** Dose-response curves for TAE684 in sensitive and resistant cell lines established from the same tumor xenografts as in **a** (IC_{50}: Sens, 7.9 nM; Res, 878.6 nM). Data are means ± SD, *n* = 3 biological replicates. **c, Upper,** Tumor volumes and **Lower,** Kaplan-Meier survival curves of resistant cell xenografts in NU/NU (CrTac:NCr-Foxn1nu) mice treated with TAE684 (10mg/kg by oral gavage once daily) or vehicle control for up to 56 days. Data are means ± SEM, *n* = 8 per arm. Significance was calculated by the Mann-Whitney U test for tumor volumes (*P* = 0.8404) and by the log-rank test for Kaplan-Meier survival analysis (*P* = 0.8076), both two-sided. **d,** Dose-response curves for TAE684-sensitive and -resistant cells treated with ceritinib (IC_{50}: Sens, 33.8 nM; Res, 446.5 nM) or lorlatinib (IC_{50}: Sens, 47.5 nM; Res, 2,318 nM). Data are means ± SD, *n* = 3 biological replicates. **e,** Immunoblot analysis of the indicated proteins in sensitive and resistant cells treated with DMSO (D) or TAE684 (TAE),1 μM for 6 or 24 h. **f,** Electropherograms of ALK kinase domain sequencing in sensitive and resistant cells. Arrows show the F1174L mutation characteristic of Kelly cells. HEK293T cells were used as a control for sequencing WT ALK. **g,** Phosphoproteomic analysis of a panel of receptor tyrosine kinases (RTKs) in sensitive and resistant cells. Each RTK is shown in duplicate and the pairs in the

corners of each array are positive controls. Numbered RTKs with corresponding names listed on the right, represent the highest phosphorylated proteins. ALK is depicted in red. **h,** qRT-PCR and immunoblot analysis of ABCB1 and ABCG2 multidrug transporter expression in sensitive and resistant cells. The qRT-PCR data are means of *n* = 2 biological replicates. Panels **a** (immunoblot), **d**, **f** and **g** are representative of two independent experiments (See Supplementary Note 1 for details; for gel source data, see Supplementary Figure 1).

Extended Data Figure 3. Development of resistance is associated with loss of MYCN followed by gradual induction of proneural TFs. a, TAE684 dose-response curves of Kelly NB cells during resistance establishment (IC_{50}: sensitive (Sens), 39.4 nM; intermediate (IR), 618 nM; resistant (Res), 1,739 nM). Data are means ± SD, *n* = 3 biological replicates. Schematic representation of resistance development is shown above. **b,** t-SNE plot of scRNA-seq data showing the segregation of sensitive (*n* = 5,432), intermediate (*n* = 6,376) and resistant (*n* = 6,379) cells, **c,** t-SNE plot depicting unsupervised clusters for the individual subpopulations that underlie the pseudotime analysis. **d,** Heatmap of re-scaled gene expression values of the most variable ranked TFs in the three cell states. **e,** qRT-PCR and immunoblot analysis of MYCN expression in TAE684-resistant xenograft tumors (1 and 2) and sensitive and resistant cells in culture (Sens, Res). The qRT-PCR data are means ± SD, *n* = 4 biological replicates for sensitive and resistant cells (***, *P* = 1.396e-11; unpaired two-sided *t*-test) and *n* = 3 technical replicates for each tumor. **f,** Fluorescence in situ hybridization of MYCN in sensitive and resistant cells (representative of 20 nuclei per condition). **g,** ChIP-seq track of H3K27me3 binding at the *MYCN* locus in sensitive and resistant cells. Signal intensity is given in the upper right corner. **h,** Line plot showing the association between genes ordered by expression (x-axis) and changes in absolute gene expression levels (y-axis) in sensitive vs. resistant cells. Barplot, total transcriptional yield in sensitive/resistant cells. **i,** Immunoblot analysis of the indicated proteins in sensitive and resistant cells expressing control (shCtrl) or MYCN (shMYCN-1 and -2)

shRNAs. **j,** Violin plots representing the expression distribution of selected genes in the same cells as in **a** (centre line, median). **k,** Barplot showing the fractions of cells with detectable mRNA levels of the same genes as in **d**. Panels **e** (immunoblot) and **f**-**i** are representative of two independent experiments (for gel source data, see Supplementary Figure 1).

Extended Data Figure 4. BORIS overexpression is seen in resistance models of NB and correlates with high-risk disease and a poor outcome. a, qRT-PCR and immunoblot analysis of BORIS expression in TAE684-resistant Kelly cell xenograft tumors (1 and 2) and sensitive and resistant cells in culture (Sens, Res). The qRT-PCR data are means ± SD, *n* = 4 biological replicates for sensitive and resistant cells (**, *P* = 0.0014; unpaired two-sided *t*-test) and *n* = 3 technical replicates for each tumor. **b,** Bisulfite sequencing of the BORIS promoter in sensitive and resistant cells. Black circles represent methylated cytosines in a CpG dinucleotide, empty circles are unmethylated cytosines. Transcription start sites (TSS) B and C are indicated by arrows. **c,** Dose-response curves to TAE684 **(left)** and immunoblot analysis of BORIS expression **(right)** in TAE684-sensitive and -resistant SK-N-SH NB cells (IC_{50}: Sens, 47.9 nM; Res, 1,739 nM). **d,** Dose-response curves to the CDK12 inhibitor, E9 **(left)** and immunoblot analysis of BORIS expression **(right)** in sensitive and resistant SK-N-BE(2) NB cells (IC_{50}: Sens, 9.5 nM; Res, 638 nM). Data are means ± SD, *n* = 3 biological replicates for **c (left)** and **d (left)**. **e,** Immunohistochemical (IHC) staining of BORIS expression in primary NB tumor samples (representative of 4 independent experiments). Scale bars, 20 µm. **f,** Box plots showing correlation of BORIS expression with the indicated parameters in a human NB dataset [*n* = 498; Tumor Neuroblastoma-SEQC-498; R2: Genomics Analysis and Visualization Platform (http://r2.amc.nl)]. Centre lines, medians; box limits, 25th and 75th percentiles; whiskers, minima and maxima (1.5X the interquartile range). The two-sided Wilcoxon rank-sum test was used for all pairwise comparisons. **g,** Kaplan-Meier analysis of overall survival based on BORIS expression in the same dataset as in **f** [*n* = 498; two-sided log-rank test with Bonferroni

correction]. Panels **a**, **c**, **d** (immunoblots) and **b** are representative of two independent experiments. Sample sizes (*n*) are depicted in parenthesis for panels **f** and **g** (for gel source data, see Supplementary Figure 1).

Extended Data Figure 5. Resistant cells are dependent on BORIS for survival. a, Dose-response curves to TAE684 in resistant cells expressing control (shCtrl) or BORIS (shBORIS) shRNAs (IC_{50}: shCtrl, 537.7 nM; shBORIS, 141.2 nM). Data are means ± SD, *n* = 3 biological replicates. **b,** Heatmap of gene expression values in the same cells as in **a** (*n* = 2 biological replicates). Rows are z-scores calculated for each gene in both conditions**. c,** Immunoblot analysis of the indicated proteins in the same cells as in **a**. **d,** Immunoblot analysis of the indicated proteins (Cl., cleaved; CC3, cleaved caspase 3), and **e,** quantification of trypan blue staining in sensitive and resistant cells expressing control (shCtrl) or BORIS (shBORIS-3 and -4) shRNAs. Data are means ± SD, *n* = 3 biological replicates (*, $P < 0.05$; **, $P < 0.01$; ***, $P < 0.001$; unpaired two-sided *t*-tests). **f,** Phase contrast microscopy images (scale bars, 150 µm), **g,** growth curves and. **h,** flow cytometry analyses of propidium iodide (PI) staining in sensitive, intermediate and resistant cells. Data are means ± SD, *n* = 3 biological replicates (***, $P < 0.0001$ for all comparisons; two-way ANOVA). **i,** qRT-PCR analysis of the expression of the indicated proneural TFs in the same sensitive (DMSO) vs. $MYCN^{KD}$/$BORIS^{Ind}$ (DOX+TAE) cells as in **Fig. 1g**. Data are means ± SD, *n* = 3 biological replicates (*, $P < 0.05$; **, $P < 0.01$; unpaired two-sided *t*-tests). Panels **c**, **d**, **f** and **h** are representative of two independent experiments (for gel source data, see Supplementary Figure 1).

Extended Data Figure 6. BORIS colocalizes with CTCF and open chromatin. a, Bar graphs illustrating the overlap of shared and specific BORIS and CTCF binding sites in sensitive and resistant cells. The large majority of resistant cell-specific BORIS peaks **(red)** colocalize with CTCF peaks that are shared between the two cell types. The significantly lower number of BORIS peaks that are unique to sensitive cells **(green)** or shared between sensitive and

resistant cells **(gray)** typically do not overlap with CTCF peaks that are shared or specific to any cell type **(upper)**. Most CTCF peaks are shared **(gray)** between sensitive and resistant cells and either do not overlap with BORIS peaks, or overlap only with those restricted to resistant cells **(lower)**. **b,** Comparison of CTCF and BORIS peaks identified in sensitive and resistant cells. **c,** Co-immunoprecipitation of BORIS with CTCF in sensitive and resistant cells (representative of two independent experiments). IgG and sample without antibody (No Ab) serve as controls. **d,** Pie charts depicting the percentages of genomic regions bound by BORIS in sensitive **(upper)** and resistant **(lower)** cells. Numbers of BORIS binding peaks in each cell type are given below each pie chart. The regions shown are promoters (TSS ± 2 kb), typical enhancers (H3K27ac), active enhancers (H3K27ac + BRD4), repressed chromatin (H3K27me3), exons, introns, and other (peaks not assigned to any of the previous categories). **e,** Meta-analysis of average CTCF and BORIS ChIP-seq signals in rpm/bp at enhancer and transcription start site (TSS) regions in sensitive and resistant cells. **f,** Percentage of gene promoters bound by BORIS in sensitive **(black)** and resistant **(red)** cells for 10 equal-sized groups ordered based on absolute gene expression levels in resistant cells. Percentage of promoters bound by BORIS in resistant cells that were also originally bound by MYCN in sensitive cells are shown in **gray**. **g,** Loess regression analysis of ranked gene expression against BORIS and MYCN occupancies at gene promoters in sensitive and resistant cells. Shaded regions represent 95% confidence intervals. All panels except **c** depict data from $n = 2$ biological replicates (for gel source data, see Supplementary Figure 1).

Extended Data Figure 7. Regulatory loops in resistant cells are more vulnerable to BORIS depletion. a, Heatmap depicting the Spearman correlation between HiChIP biological replicates of sensitive and resistant cells in genome-wide bins of 5 kb for all merged anchor regions. **b,** Box plots showing the genomic length distribution (in log2 bp) for interaction classes that are specific to resistant cells. **c,** Table depicting HiChIP loop class statistics in resistant cells,

including their association with BORIS binding. **d,** ChIP-seq tracks of the indicated proteins in sensitive and resistant cells at the *TCP11L2* locus (representative of two independent experiments), with resistant cell-specific regulatory interactions shown below [HiChIP Res: PET numbers, next to each interaction] Signal intensity is given in the upper left corner for each track. **e,** Heatmap depicting the Spearman correlation between HiChIP biological replicates of sensitive, resistant, shCtrl and shBORIS cells in genome-wide bins of 5 kb for all merged anchor regions. **f,** Barplots showing the number and fraction of resistant cell-specific loops for all interaction classes that were BORIS negative and positive in resistant cells, and that were lost upon BORIS depletion. **g,** Barplots showing the odds-ratio (two-sided Fisher's exact test) of losing loops that were previously bound by BORIS for all interaction classes. **h,** Box plots showing the initial intensities (in normalized read counts) of BORIS and SMC1A binding in the shRNA control cells (shCtrl) at the anchors of the resistant cell-specific loops that were significantly lost vs. those that were retained in shBORIS cells (two-sided Wilcoxon rank-sum test). **i,** Box plot showing the difference in SMC1A loss (shBORIS vs. shCtrl) on the same anchors as in **h** (two-sided Wilcoxon rank-sum test). All box plots are defined by centre lines, medians; box limits, 25^{th} and 75^{th} percentiles; whiskers, minima and maxima (1.5X the interquartile range). **j,** Metaplots depicting BORIS, SMC1A and CTCF binding at the anchors of the resistant cell-specific loops that were lost or retained upon BORIS depletion. Panels **a**-**c** and **e**-**g** depict data from *n* = 3 biological replicates. Panels **h**-**j** depict data from *n* = 2 biological replicates.

Extended Data Figure 8. Redistribution of the SE landscape with subsequent expression of a BORIS-dependent proneural network in resistant cells. a, Accumulation of H3K27ac signal at enhancer regions. Typical enhancers **(gray)** are plotted according to increasing levels of normalized H3K27ac signal (length x density) in sensitive and resistant cells. The highest cutoff based on the inclination point in both sensitive and resistant cells was used to delineate

SEs **(red)**. **b,** Scatterplot showing differential binding of H3K27ac [(log2 (rpm/bp + 1)] and BRD4 [log2 (rpm/bp + 1)] for all detected SEs in both sensitive and resistant cells. Cell-specific SEs were identified based on the combined increase in H3K27ac and BRD4 binding. For each individual histone mark, a 0.75 log2FC threshold was applied and a minimum summed 2.5 log2FC was used as the final cutoff. **c**, Barplot depicting the enrichment (two-sided Fisher's exact test) and fractions of resistant cell-specific and shared SEs that were located at resistant cell-specific regulatory loop anchors in resistant cells. **d,** Density plots showing the aggregated accumulation of H3K27ac and H3K27me3 at gene regions, defined as 2 kb upstream of the TSS and 2 kb downstream of the transcription end site (TES). K-means clustering (k = 3) analysis resulted in the separation of genes associated with 'open', 'neutral' or 'closed' chromatin in both sensitive and resistant cells. **e,** Sankey diagram of the distribution of genes in distinct chromatin states and the switches between sensitive and resistant cells. **f,** Box plots showing the expression level changes upon BORIS depletion for genes that had a resistant cell-specific and BORIS-positive regulatory interaction and were not associated with a SE (BORIS no SE, *n* = 720), associated with a SE in both cell types (BORIS shared SE, *n* = 514), or associated with a SE seen only in the resistant cells (BORIS Res SE, *n* = 134) (two-sided Wilcoxon rank-sum test). Centre lines, medians; box limits, 25th and 75th percentiles; whiskers, minima and maxima (1.5X the interquartile range). **g,** Heatmap of the expression levels of the indicated proneural TF genes during brain development (http://www.brain-map.org/). Gene expression levels are represented as z-scores for different developmental time points (*n* = 413; pcw, postconceptional weeks). **h,** Heatmap showing the odds-ratios (two-sided Fisher's exact test) for co-detection of the indicated TFs based on the scRNA-seq data in resistant cells (*n* = 6,379). **i,** Immunoblot analysis of the indicated proteins in sensitive and resistant cells expressing control (shCtrl) or BORIS (shBORIS-3 and -4) shRNAs. **j**, qRT-PCR analysis of the indicated genes, and **k,** ChIP-qPCR analysis of BORIS binding at the promoter regions of *BORIS* and *NEUROG2* in sensitive and resistant SK-N-BE(2) NB cells. Data are means ± SD, *n*

= 3 biological replicates for panels **j** and **k** (*, $P < 0.05$; **, $P < 0.01$; ***, $P < 0.001$; unpaired two-sided *t*-tests). All other panels except **g** and **h** depict data from $n = 2$ biological replicates (for gel source data, see Supplementary Figure 1).

Extended Data Figure 9. The proneural TF network in resistant cells is sensitive to BRD4 inhibition. a, Metaplots showing the correlation between BRD4 and BORIS co-occupancies at the promoter regions (± 2 kb) of the 89 top-ranked genes in resistant vs. sensitive cells based on the features in **Fig. 4b** (*r*, Spearman correlation coefficient). **b,** Immunoblot analysis of BRD4 and cleaved PARP (Cl. PARP) expression in sensitive and resistant cells expressing control (shCtrl) or BRD4 (shBRD4-A and -B) shRNAs. **c,** Immunoblot analysis of the indicated proteins (CC3, cleaved caspase 3) in sensitive and resistant cells treated with DMSO, TAE684 (1 μM) or JQ1 (2.5 μM) for 48 h. **d,** Dose-response curves for sensitive and resistant cells treated with JQ1 or I-BET726 [JQ1 (IC_{50}: Sens, 4,798 nM; Res, 645 nM); I-BET726 (IC_{50}: Sens, 6,203 nM; Res, 347 nM)]. Data are means ± SD, n = 3 biological replicates. **e,** Box plots comparing the expression of the TFs listed in **Fig. 4b** (n = 13) with that of all genes (n = 18,038) in sensitive vs. resistant cells **(left)**, and between DMSO and JQ1-treated resistant cells **(right)** (*P*; two-sided Wilcoxon rank-sum test). **f,** ChIP-seq tracks of the indicated proteins at the *SIX1/SIX4* locus in sensitive, resistant and JQ1-treated resistant cells (2.5 μM for 48 h). SEs are depicted as colored rectangles below the tracks. Signal intensity is shown in the upper left corner for each track. **g,** Tumor volumes and **h,** survival curves in sensitive- and resistant-cell xenografts in NU/NU (Crl:NU-Foxn1nu) mice treated with JQ1 (50 mg/kg i.p. once daily) and vehicle control for up to 87 days. Data are means ± SEM, n = 6 per arm. Significance was calculated by the Mann-Whitney U test for tumor volumes (Sens: $P = 0.3231$; Res: $P = 0.0023$) and by the log-rank test for Kaplan-Meier survival analysis (Sens: $P = 0.3047$; Res: 0.0348), both two-sided. **i,** Heatmap of gene expression values in sensitive (Sens), resistant (Res) and JQ1-treated resistant (Res + JQ1) cells. Rows are z-scores calculated for each gene in each condition. **j,**

Number of transcripts in sensitive (Sens), JQ1-treated resistant (Res + JQ1), shBORIS-expressing resistant (shBORIS), and resistant (Res) cells based on expression array data after spike-in normalization. **k,** Scatterplot displaying the median-scaled fold-change gene expression values for shBORIS and JQ1-treated resistant cells. The top ranked TFs that show decreased expression levels upon both *BORIS* knockdown and JQ1 treatment are listed in red **(left lower quadrant)**. The pie chart represents the fraction of all top-ranked TFs that are located in the left lower quadrant of the scatterplot. All box plots are defined by centre lines, medians; box limits, 25^{th} and 75^{th} percentiles; whiskers, minima and maxima (1.5X the interquartile range). Panels **b**, **c**, **f** are representative of two independent experiments. Panels **a**, **e** and **i**-**k** depict data from n = 2 biological replicates (See Supplementary Note 2 for further details; for gel source data, see Supplementary Figure 1).

Extended Data Figure 10. Aberrantly expressed BORIS binds to regulatory regions and is associated with new SEs in Ewing sarcoma cells. a, Immunoblot analysis of BORIS expression in TC-32 (pre-chemotherapy), TC-71 and CHLA-10 (relapsed, post-chemotherapy) Ewing sarcoma cells, compared with BORIS expression in resistant (Res Kelly) NB cells. **b,** Meta-analysis of average BORIS ChIP-seq signals in rpm/bp at all combined BORIS binding sites for TC-32 and TC-71 cells. **c,** Meta-analysis of average BORIS, H3K27ac and SMC1A ChIP-seq signals in rpm/bp at TC-71-specific BORIS binding sites. **d,** Pie chart depicting the proportions of genomic regions bound by BORIS in TC-71 cells. The regions shown are promoters (TSS ± 2 kb), typical and super-enhancers (H3K27ac), and other if peaks were not assigned to any of the previous categories. **e,** Barplot showing the odds-ratios (two-sided Fisher's exact test) of BORIS localization to regulatory genomic regions in TC-71 cells. All panels are representative of two independent experiments (See Supplementary Note 3 for further details; for gel source data, see Supplementary Figure 1).

Fig. 1

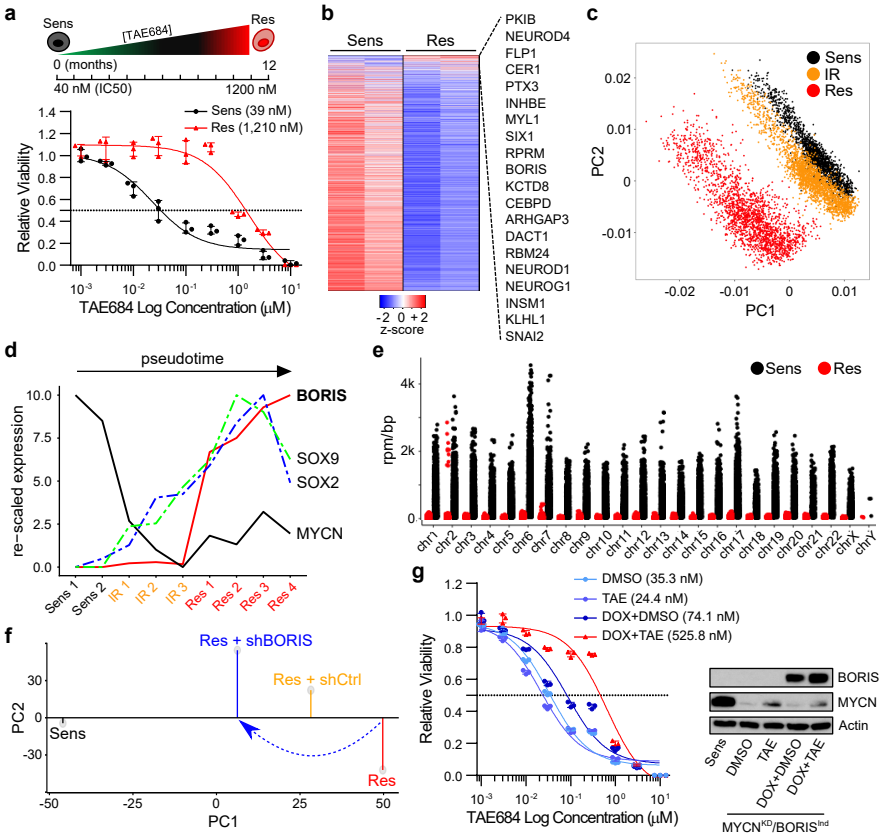

a
Sens
[TAE684]
Res
0 (months)
12
40 nM (IC50)
1200 nM
Sens (39 nM)
Res (1,210 nM)
Relative Viability
TAE684 Log Concentration (μM)
b
Sens
Res
PKIB
NEUROD4
FLP1
CER1
PTX3
INHBE
MYL1
SIX1
RPRM
BORIS
KCTD8
CEBPD
ARHGAP3
DACT1
RBM24
NEUROD1
NEUROG1
INSM1
KLHL1
SNAI2
-2 0 +2
z-score
c
Sens
IR
Res
PC2
PC1
d
pseudotime
re−scaled expression
BORIS
SOX9
SOX2
MYCN
Sens 1
Sens 2
IR 1
IR 2
IR 3
Res 1
Res 2
Res 3
Res 4
e
Sens
Res
rpm/bp
f
Res + shBORIS
Res + shCtrl
Sens
Res
PC2
PC1
g
DMSO (35.3 nM)
TAE (24.4 nM)
DOX+DMSO (74.1 nM)
DOX+TAE (525.8 nM)
Relative Viability
TAE684 Log Concentration (μM)
BORIS
MYCN
Actin
Sens
DMSO
TAE
DOX+DMSO
DOX+TAE
MYCN^KD/BORIS^Ind

Fig. 2

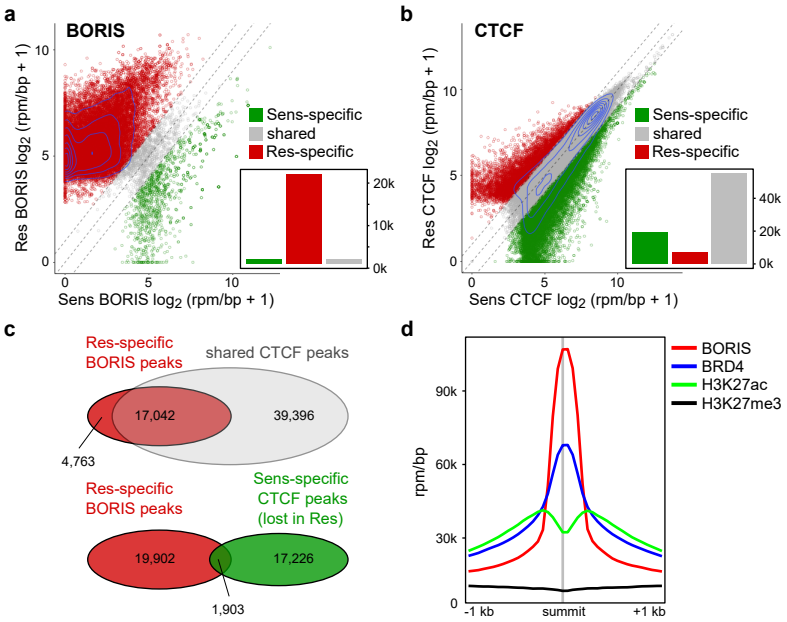

a
BORIS
Res BORIS log2 (rpm/bp + 1)
Sens BORIS log2 (rpm/bp + 1)
Sens-specific
shared
Res-specific
20k
10k
0k
b
CTCF
Res CTCF log2 (rpm/bp + 1)
Sens CTCF log2 (rpm/bp + 1)
40k
20k
0k
c
Res-specific BORIS peaks
shared CTCF peaks
17,042
39,396
4,763
Res-specific BORIS peaks
Sens-specific CTCF peaks (lost in Res)
19,902
17,226
1,903
d
rpm/bp
BORIS
BRD4
H3K27ac
H3K27me3
-1 kb
summit
+1 kb

Fig. 3

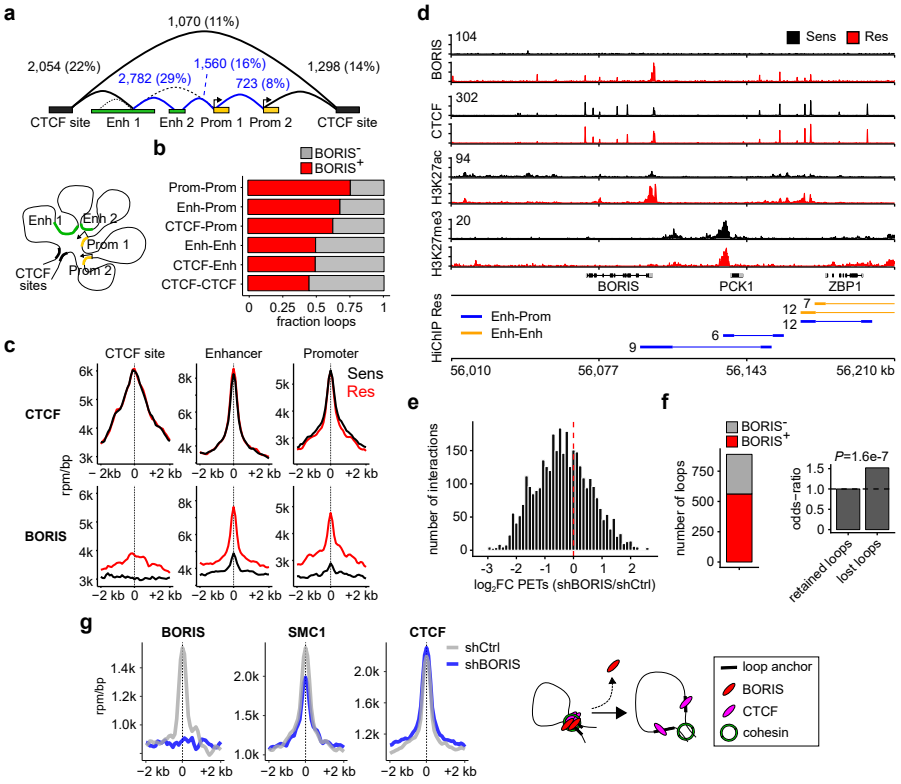
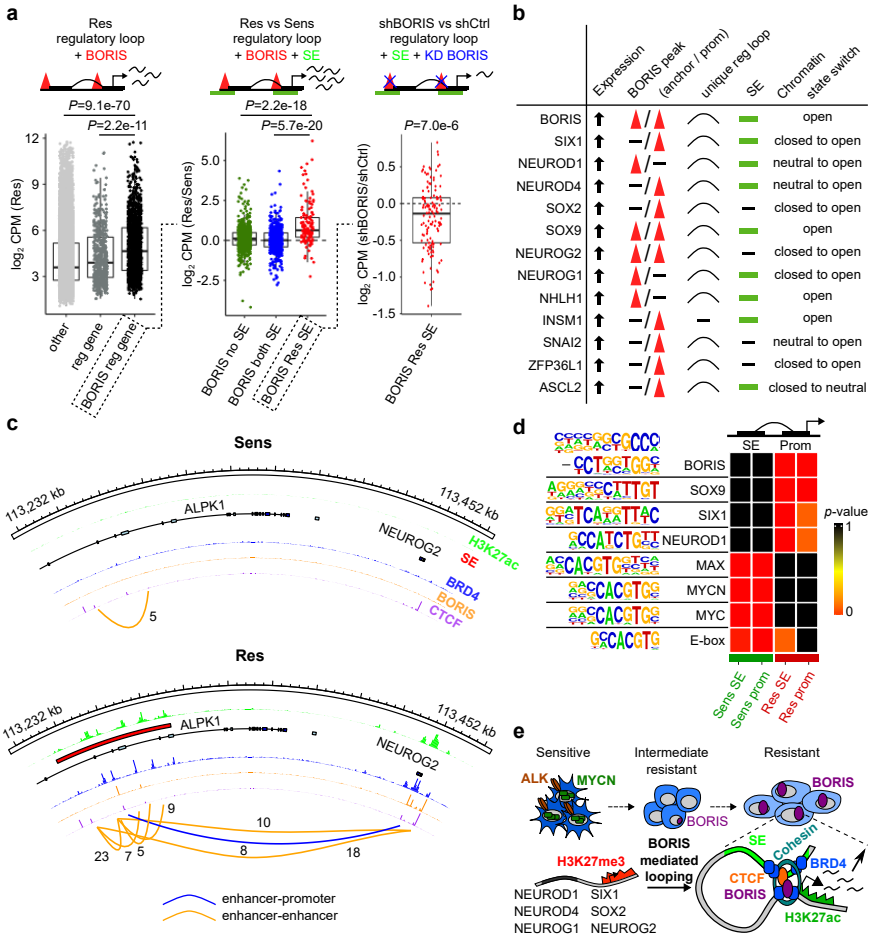

a
1,070 (11%)
2,054 (22%)
1,560 (16%)
2,782 (29%)
723 (8%)
1,298 (14%)
CTCF site
Enh 1
Enh 2
Prom 1
Prom 2
CTCF site
b
BORIS-
BORIS+
Enh 1
Enh 2
Prom 1
Prom 2
CTCF sites
Prom-Prom
Enh-Prom
CTCF-Prom
Enh-Enh
CTCF-Enh
CTCF-CTCF
0 0.25 0.5 0.75 1
fraction loops
c
CTCF site
Enhancer
Promoter
Sens
Res
CTCF
BORIS
rpm/bp
−2 kb 0 +2 kb
d
Sens
Res
BORIS
CTCF
H3K27ac
H3K27me3
HiChIP Res
104
302
94
20
BORIS
PCK1
ZBP1
Enh-Prom
Enh-Enh
12
12
7
6
9
56,010
56,077
56,143
56,210 kb
e
number of interactions
log2FC PETs (shBORIS/shCtrl)
f
BORIS-
BORIS+
number of loops
odds−ratio
P=1.6e-7
retained loops
lost loops
g
BORIS
SMC1
CTCF
shCtrl
shBORIS
rpm/bp
loop anchor
BORIS
CTCF
cohesin

Fig. 4


a
Res regulatory loop + BORIS
Res vs Sens regulatory loop + BORIS + SE
shBORIS vs shCtrl regulatory loop + SE + KD BORIS
P=9.1e-70
P=2.2e-11
P=2.2e-18
P=5.7e-20
P=7.0e-6
log2 CPM (Res)
log2 CPM (Res/Sens)
log2 CPM (shBORIS/shCtrl)
other
reg gene
BORIS reg gene
BORIS no SE
BORIS both SE
BORIS Res SE
BORIS Res SE
b
Expression
BORIS peak (anchor / prom)
unique reg loop
SE
Chromatin state switch
BORIS open
SIX1 closed to open
NEUROD1 neutral to open
NEUROD4 neutral to open
SOX2 closed to open
SOX9 open
NEUROG2 closed to open
NEUROG1 closed to open
NHLH1 open
INSM1 open
SNAI2 neutral to open
ZFP36L1 closed to open
ASCL2 closed to neutral
c
Sens
Res
113,232 kb
113,452 kb
ALPK1
NEUROG2
H3K27ac
SE
BRD4
BORIS
CTCF
enhancer-promoter
enhancer-enhancer
d
SE
Prom
BORIS
SOX9
SIX1
NEUROD1
MAX
MYCN
MYC
E-box
p-value
Sens SE
Sens prom
Res SE
Res prom
e
Sensitive
Intermediate resistant
Resistant
ALK
MYCN
BORIS
H3K27me3
BORIS mediated looping
SE
Cohesin
CTCF
BRD4
H3K27ac
NEUROD1 SIX1
NEUROD4 SOX2
NEUROG1 NEUROG2