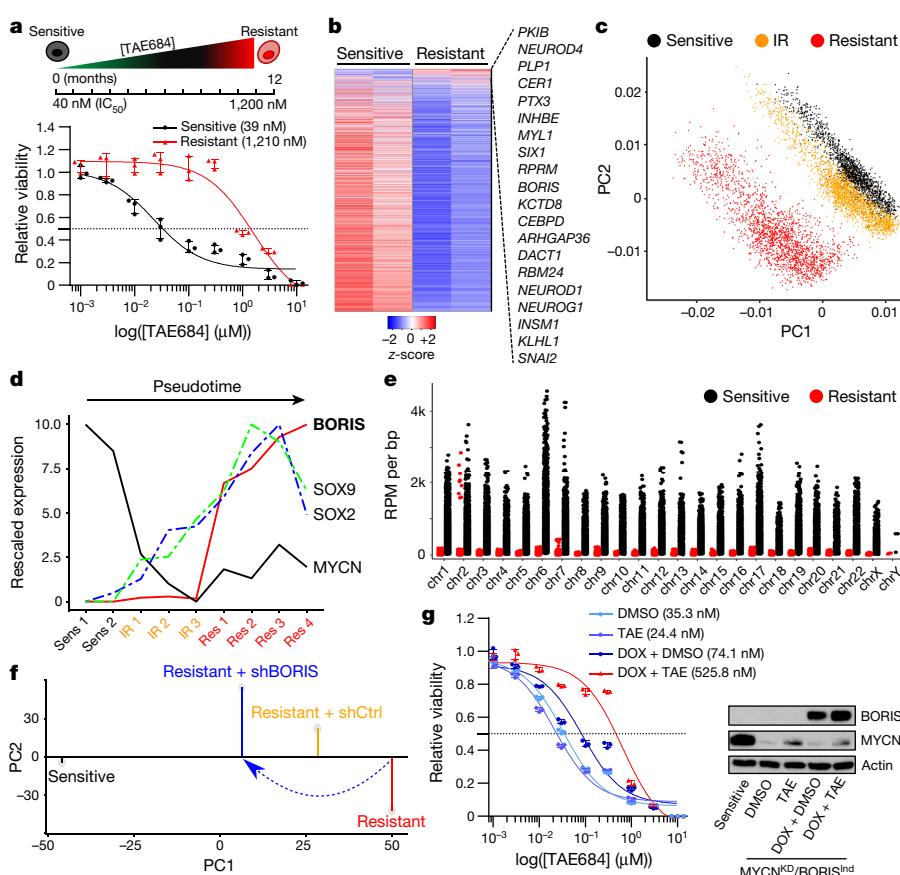


# BORIS promotes chromatin regulatory interactions in treatment-resistant cancer cells

David N. Debruyne<sup>1,2,15</sup>, Ruben Dries<sup>1,2,3,15</sup>, Satyaki Sengupta<sup>1,2</sup>, Davide Seruggia<sup>1,4,5,6</sup>, Yang Gao<sup>1,2</sup>, Bandana Sharma<sup>1,2</sup>, Hao Huang<sup>1,2</sup>, Lisa Moreau<sup>7</sup>, Michael McLane<sup>1,2</sup>, Daniel S. Day<sup>8,9</sup>, Eugenio Marco<sup>3,10</sup>, Ting Chen<sup>11</sup>, Nathanael S. Gray<sup>12,13</sup>, Kwok-Kin Wong<sup>14</sup>, Stuart H. Orkin<sup>1,4,5,6</sup>, Guo-Cheng Yuan<sup>3,10</sup>, Richard A. Young<sup>8,9</sup> & Rani E. George<sup>1,2,4\*</sup>

The CCCTC-binding factor (CTCF), which anchors DNA loops that organize the genome into structural domains, has a central role in gene control by facilitating or constraining interactions between genes and their regulatory elements<sup>1,2</sup>. In cancer cells, the disruption of CTCF binding at specific loci by somatic mutation<sup>3,4</sup> or DNA hypermethylation<sup>5</sup> results in the loss of loop anchors and consequent activation of oncogenes. By contrast, the germ-cell-specific parologue of CTCF, BORIS (brother of the regulator of imprinted sites, also known as CTCFL)<sup>6</sup>, is overexpressed in several cancers<sup>7–9</sup>, but its contributions to the malignant phenotype remain unclear. Here we show that aberrant upregulation of BORIS promotes chromatin interactions in ALK-mutated, MYCN-amplified

neuroblastoma<sup>10</sup> cells that develop resistance 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 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 promote regulatory chromatin interactions that support specific cancer phenotypes.

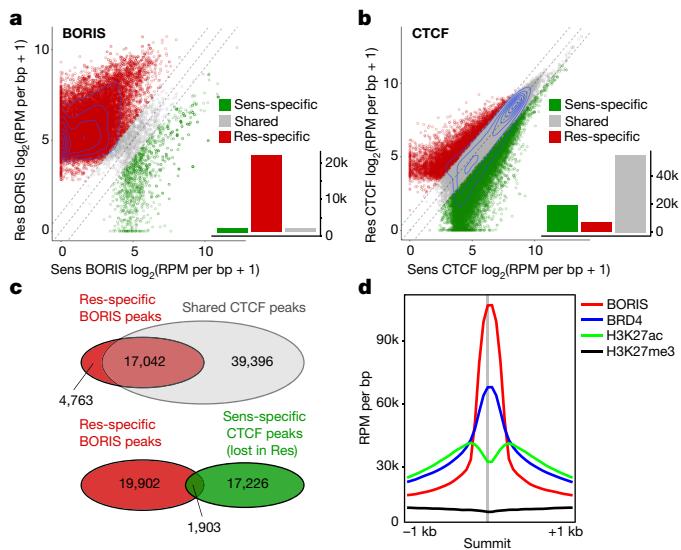


<sup>1</sup>Department of Pediatric Oncology, Dana-Farber Cancer Institute, Boston, MA, USA. <sup>2</sup>Department of Pediatrics, Harvard Medical School, Boston, MA, USA. <sup>3</sup>Department of Biostatistics and Computational Biology, Dana-Farber Cancer Institute, Boston, MA, USA. <sup>4</sup>Division of Hematology/Oncology, Boston Children's Hospital, Boston, MA, USA. <sup>5</sup>Harvard Stem Cell Institute, Harvard Medical School, Boston, MA, USA. <sup>6</sup>Howard Hughes Medical Institute, Boston, MA, USA. <sup>7</sup>Department of Radiation Oncology, Dana-Farber Cancer Institute, Boston, MA, USA. <sup>8</sup>Whitehead Institute for Biomedical Research, Cambridge, MA, USA. <sup>9</sup>MIT Department of Biology, Cambridge, MA, USA. <sup>10</sup>Department of Biostatistics, Harvard TC Chan School of Public Health, Boston, MA, USA. <sup>11</sup>Department of Medical Oncology, Dana-Farber Cancer Institute, Boston, MA, USA. <sup>12</sup>Department of Cancer Biology, Dana-Farber Cancer Institute, Boston, MA, USA. <sup>13</sup>Department of Biological Chemistry and Molecular Pharmacology, Harvard Medical School, Boston, MA, USA. <sup>14</sup>Division of Hematology and Medical Oncology, Laura and Isaac Perlmutter Cancer Center, New York University Langone Medical Center, New York, NY, USA. <sup>15</sup>These authors contributed equally: David N. Debruyne, Ruben Dries.. \*e-mail: rani\_george@dfci.harvard.edu

Unlike CTCF, which is uniformly expressed in healthy tissues and cancer cells, the expression of *BORIS* is typically restricted to the testis<sup>6</sup> and embryonic stem cells<sup>11</sup> (Extended Data Fig. 1a). However, when aberrantly expressed in cancer<sup>7–9</sup>, it is associated with high-risk features that include resistance to treatment (Extended Data Fig. 1b, c). We identified *BORIS* as one of the most differentially expressed genes in neuroblastoma cells driven by amplified *MYCN*<sup>12</sup> and ALK(F1174L)<sup>13</sup> and rendered resistant to ALK inhibition. Kelly human neuroblastoma cells were exposed to increasing concentrations of the ALK inhibitor TAE684<sup>14</sup> 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—which indicates 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 not typically associated with neuroblastoma cells<sup>15,16</sup> (Fig. 1b).

We therefore proposed that the resistant cells had probably undergone transcriptional reprogramming during the development of resistance. To determine the dynamics of resistance development, we performed single-cell RNA sequencing (scRNA-seq) analysis on sensitive, intermediate and fully resistant cell states (Extended Data Fig. 3a). Principal component analysis (PCA) indicated a stepwise transition as cells progressed from the sensitive to the fully resistant state (Fig. 1c). This transition was confirmed by distributed stochastic neighbour embedding (*t*-SNE)<sup>17</sup>, which clustered the cells into three non-overlapping categories (Extended Data Fig. 3b, c). Pseudotime analysis based on the transcription factors that were differentially expressed throughout the development of resistance 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 analysed the status of *MYCN* in these cells, and found that although 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<sup>15,18,19</sup> (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 after depletion of *MYCN* (Extended Data Fig. 3i). Subsequent resistance stages were defined by a gradual increase in the expression of the neural developmental markers *SOX2* and *SOX9*<sup>20</sup>, 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). Overexpression of *BORIS*, which coincided with promoter hypomethylation (Extended Data Fig. 4a, b), was also observed in additional neuroblastoma cell lines rendered resistant to TAE684 (SK-N-SH) or the CDK12 inhibitor E9<sup>21</sup> (SK-N-BE(2)) (Extended Data Fig. 4c, d), which suggests that our findings are not restricted to a single cell line or kinase inhibitor. Indeed, overexpression of *BORIS* in tumours was significantly associated with high-risk disease and a poor outcome in patients with neuroblastoma 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, and observed 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 depletion of *BORIS* in resistant cells ultimately decreased cell viability (Extended Data Fig. 5d, e), which indicates 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 overexpression of

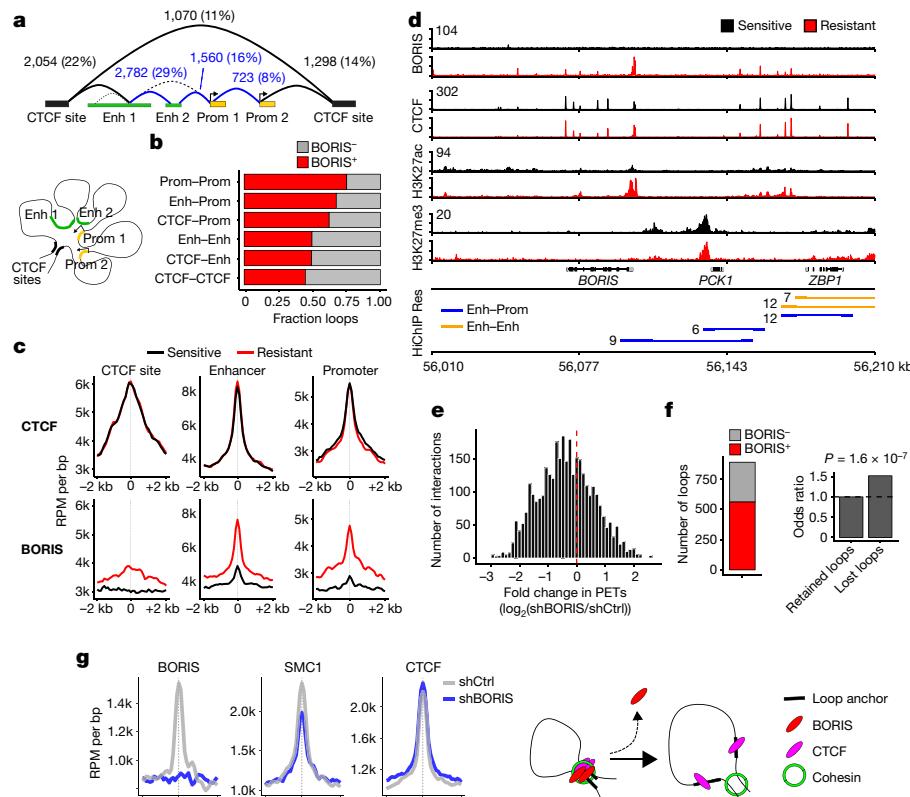


**Fig. 2 | *BORIS* overexpression is associated with its increased chromatin occupancy in resistant cells, whereas CTCF binding is unchanged.**

**a**, Scatter plot of *BORIS* binding in sensitive (Sens) and resistant (Res) cells for all detected *BORIS*-binding sites. *BORIS* peaks unique to resistant cells ( $n = 21,805$ ; 91%), sensitive cells ( $n = 1,125$ ; 4.7%) and shared between the two cell types ( $n = 1,086$ ; 4.5%) are shown. **b**, Scatter plot of CTCF binding in sensitive and resistant cells for all detected CTCF-binding sites. CTCF peaks unique to resistant cells ( $n = 6,808$ ; 8.3%), sensitive cells ( $n = 19,129$ ; 23.2%) and shared between the two cell types ( $n = 56,438$ ; 68.5%) are shown. **c**, Overlap between *BORIS* peaks that are unique to resistant cells and CTCF peaks shared between resistant and sensitive cells (top), and between resistant cell-specific *BORIS* peaks and sensitive cell-specific CTCF peaks (bottom). **d**, Meta-analysis of average ChIP-seq signals at resistant cell-specific *BORIS*-binding sites. All panels,  $n = 2$  biological replicates.

*BORIS* (Extended Data Fig. 5f–h). Given the many sequential steps involved in the evolution of resistance, overexpression of *BORIS* 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 transcription factors that were upregulated in the original TAE684-resistant cells, including *SOX2* and *SOX9* (Extended Data Figs. 3d, 5i). Thus, resistance to inhibition of ALK in neuroblastoma 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 overexpression of *BORIS*. In this context, the resistant cells ultimately become dependent on *BORIS* for survival, which supports a key role for this protein in maintenance of the resistance state.

We next asked whether the aberrant expression of *BORIS*, a DNA-binding protein<sup>6</sup>, affected its genome-wide occupancy in resistant cells. We observed a large (tenfold) gain in *BORIS*-bound peaks after chromatin immunoprecipitation followed by high-throughput sequencing (ChIP-seq) analysis in resistant cells: 22,891 versus 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 versus 63,246 peaks) (Fig. 2b). A considerable 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<sup>22</sup> (Extended Data Fig. 6c). However, only a small proportion ( $n = 1,903$ ; 8.7%) overlapped with CTCF peaks unique to sensitive cells, which suggests 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), which is consistent with its propensity to bind to open chromatin regions<sup>23</sup> (Fig. 2d). Such differential chromatin



**Fig. 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 (Prom) regions (transcription start site (TSS)  $\pm$  2 kb), active enhancer (Enh) regions, or CTCF sites only, in that order). Absolute numbers and percentages for each loop type (structural (black), regulatory (blue)) are shown. Cartoon illustrates 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 centred and extended in both directions ( $\pm 2$  kb). **d**, ChIP-seq tracks of the indicated proteins in sensitive and resistant cells at the BORIS locus

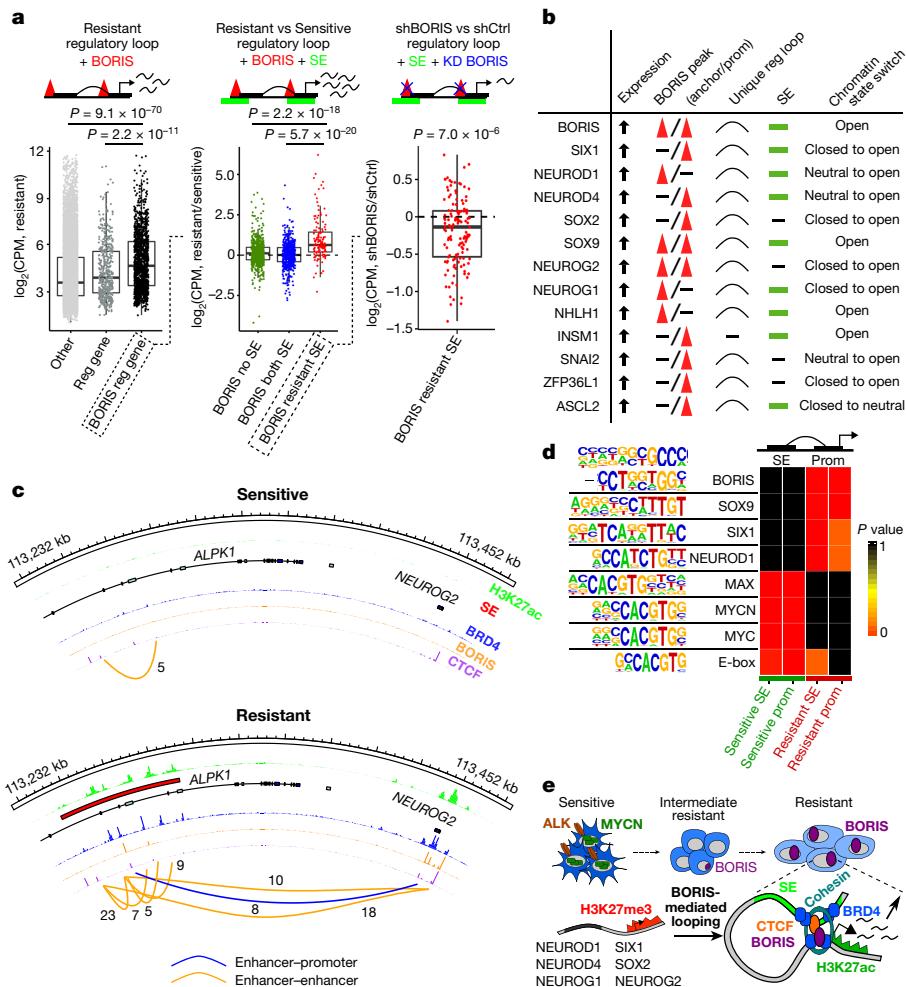
binding at distinct highly expressed genes in resistant versus sensitive cells was commensurate with the MYCN-to-BORIS dependency switch (Extended Data Fig. 6f, g).

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 high-throughput chromosome conformation capture followed by chromatin immunoprecipitation (HiChIP)<sup>24</sup> (Extended Data Fig. 7a). On the basis of the genomic locations of the associated loop anchors, six classes of interactions were identified<sup>25</sup>: 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 most (56%) of the 9,487 interactions gained in resistant cells were positive for BORIS ( $\log_2$ -transformed fold change  $> 1$ ; false discovery rate (FDR)  $< 0.01$ ) (Fig. 3b, Extended Data Fig. 7c). Notably, BORIS was enriched at anchors that were associated with regulatory regions, whereas CTCF binding remained constant, as seen at the BORIS locus itself (Fig. 3c, d). In fact, BORIS binding alone at CTCF-negative loop anchors was sufficient to generate new interactions in resistant cells (Extended Data Fig. 7d).

(representative of two independent experiments), with resistant cell-specific regulatory interactions shown below (HiChIP resistant: paired-end tag (PET) numbers, next to each interaction). Signal intensity is given in the top left corner for each track. **e**, PET interactions in BORIS-depleted (shBORIS) versus control (shCtrl) cells. **f**, Resistant cell-specific loops lost after depletion of BORIS 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 (right).  $P$  value determined by two-sided Fisher's exact test. **g**, Meta-analysis of average BORIS, SMC1A and CTCF ChIP-seq signals at resistant cell-specific loop anchors that were lost after depletion of BORIS ( $n = 2$  biological replicates). BORIS depletion at loop anchors inhibits retention of the cohesin complex, and thus prevents the formation of new loops (loop extrusion model). In **a**, **b**, **e** and **f**,  $n = 3$  biological replicates.

To test whether the newly formed interactions in resistant cells were mediated by BORIS binding, we analysed the consequences of BORIS depletion on loop architecture (Extended Data Fig. 7e). Regulatory interactions specific to resistant cells displayed a global shift towards loss after knockdown of BORIS (Fig. 3e), with more than one-quarter of the total interactions lost, of which 63% were positive for BORIS at their anchors (Fig. 3f). Interactions in which 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 BORIS-bound (Fig. 3f, Extended Data Fig. 7f, g). These results agree with the loop extrusion model<sup>26</sup>, 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 crucial factor in the looping landscape of resistant cells.

Genes associated with new BORIS-positive regulatory interactions were 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 in both healthy and cancer cells<sup>15,27,28</sup>, we characterized the super-enhancer landscape of our cells, observing that the super-enhancers unique to resistant cells were enriched at BORIS-positive regulatory loops (Extended Data Fig. 8a–c).



**Fig. 4 | BORIS-regulated chromatin remodelling supports a phenotypic switch that maintains the resistant state.** **a**, Left, fold change in expression in counts per million (CPM) of genes involved in resistant cell-specific regulatory interactions that are positive for BORIS binding ( $n = 1,368$ ) versus those involved in regulatory interactions that are negative for BORIS binding ( $n = 519$ ) or not associated with a new regulatory interaction (other) ( $n = 16,151$ ). Centre, fold change in expression of genes involved in resistant cell-specific regulatory interactions positive for BORIS binding and associated with super-enhancers (SEs) specific to resistant cells ( $n = 134$ ) versus those with super-enhancers shared by both cell types ( $n = 514$ ) or not associated with super-enhancers ( $n = 720$ ). Right, fold change in expression of genes involved in resistant cell-specific regulatory interactions positive for BORIS binding and associated with resistant cell-specific super-enhancers before and after BORIS knockdown (KD) ( $n = 134$ ) ( $P$  values determined by two-sided Wilcoxon rank-sum test). For all box plots, centre lines denote medians; box limits denote twenty-fifth and seventy-fifth percentiles; whiskers denote minima and maxima ( $1.5 \times$  the interquartile range). **b**, Highest-ranked transcription factors associated with the resistance phenotype selected based on the presence of at least four of the five 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**, Transcription factor recognition motifs at super-enhancers and promoters ( $\pm 2$  kb) of the 1,000 highest-expressed genes in resistant and sensitive cells ( $n = 2$  biological replicates) ( $P$  values determined by 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.

The presence of such super-enhancers correlated significantly with higher expression of their associated genes in resistant versus sensitive cells (Fig. 4a). These BORIS-positive super-enhancer-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 super-enhancers (Fig. 4a, Extended Data Fig. 8f). These observations suggest that BORIS-mediated alterations in chromatin looping lead to interactions of newly formed super-enhancers with their target genes, which results 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, super-enhancer landscape and chromatin state. This analysis revealed 89 genes (Supplementary Table), including

by two-sided Wilcoxon rank-sum test). For all box plots, centre lines denote medians; box limits denote twenty-fifth and seventy-fifth percentiles; whiskers denote minima and maxima ( $1.5 \times$  the interquartile range). **b**, Highest-ranked transcription factors associated with the resistance phenotype selected based on the presence of at least four of the five 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**, Transcription factor recognition motifs at super-enhancers and promoters ( $\pm 2$  kb) of the 1,000 highest-expressed genes in resistant and sensitive cells ( $n = 2$  biological replicates) ( $P$  values determined by 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.

13 transcription factors, that are highly expressed during early neural development and are crucial to cell fate decisions<sup>20,29,30</sup> (Fig. 4b, c, Extended Data Fig. 8g). The expression of these proneural transcription factors 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, analysis of transcription factor binding sites revealed enrichment of BORIS and several of these proneural transcription factors at the regulatory regions of the highest-expressed genes in resistant cells, whereas sensitive cells were dominated by MYC, MYCN and MAX E-box and E-box-like motifs (Fig. 4d). Similar increased expression of proneural transcription factors with increased BORIS occupancy at their promoters was seen in BORIS-overexpressing E9-resistant SK-N-BE(2) neuroblastoma 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, which rendered the resistant cells more sensitive to BET inhibition (Extended Data Fig. 9; Supplementary Note 2). Together, these results indicate the establishment of an alternative transcription factor regulatory network controlled by BORIS-induced chromatin remodelling to support the resistant cell state.

Thus, using a pair of isogenic ALK-inhibitor sensitive and resistant neuroblastoma cell lines, we show that the CTCF parologue BORIS can promote regulatory DNA interactions that support a phenotypic switch in the context of treatment resistance (Fig. 4e). This mechanism appears relevant to different neuroblastoma cell lines and kinase inhibitors and may extend to other cancers. In Ewing sarcoma, in which overexpression of *BORIS* 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 alteration of chromatin looping is a general mechanism by which tumour cells co-opt developmental networks to sustain alternative cell states in response to targeted or conventional therapies.

## Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at <https://doi.org/10.1038/s41586-019-1472-0>.

Received: 21 July 2017; Accepted: 11 July 2019;

Published online: 07 August 2019

- Dixon, J. R. et al. Topological domains in mammalian genomes identified by analysis of chromatin interactions. *Nature* **485**, 376–380 (2012).
- Phillips-Cremins, J. E. et al. Architectural protein subclasses shape 3D organization of genomes during lineage commitment. *Cell* **153**, 1281–1295 (2013).
- Hnisz, D. et al. Activation of proto-oncogenes by disruption of chromosome neighborhoods. *Science* **351**, 1454–1458 (2016).
- Katainen, R. et al. CTCF/cohesin-binding sites are frequently mutated in cancer. *Nat. Genet.* **47**, 818–821 (2015).
- Flavahan, W. A. et al. Insulator dysfunction and oncogene activation in *IDH* mutant gliomas. *Nature* **529**, 110–114 (2016).
- 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. *Proc. Natl. Acad. Sci. USA* **99**, 6806–6811 (2002).
- Klenova, E. M., Morse, H. C. III, Ohlsson, R. & Lobanenkov, V. V. The novel *BORIS* + CTCF gene family is uniquely involved in the epigenetics of normal biology and cancer. *Semin. Cancer Biol.* **12**, 399–414 (2002).
- Martin-Kleiner, I. BORIS in human cancers — a review. *Eur. J. Cancer* **48**, 929–935 (2012).
- Garikapati, K. R. et al. Down-regulation of BORIS/CTCF efficiently regulates cancer stemness and metastasis in *MYCN* amplified neuroblastoma cell line by modulating Wnt/β-catenin signaling pathway. *Biochem. Biophys. Res. Commun.* **484**, 93–99 (2017).
- Cheung, N. K. & Dyer, M. A. Neuroblastoma: developmental biology, cancer genomics and immunotherapy. *Nat. Rev. Cancer* **13**, 397–411 (2013).
- 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. *Mol. Hum. Reprod.* **14**, 347–355 (2008).
- Brodeur, G. M., 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).
- George, R. E. et al. Activating mutations in ALK provide a therapeutic target in neuroblastoma. *Nature* **455**, 975–978 (2008).
- Galkin, A. V. et al. Identification of NVP-TAE684, a potent, selective, and efficacious inhibitor of NPM-ALK. *Proc. Natl. Acad. Sci. USA* **104**, 270–275 (2007).
- Chipumuro, E. et al. CDK7 inhibition suppresses super-enhancer-linked oncogenic transcription in *MYCN*-driven cancer. *Cell* **159**, 1126–1139 (2014).
- Durbin, A. D. et al. Selective gene dependencies in *MYCN*-amplified neuroblastoma include the core transcriptional regulatory circuitry. *Nat. Genet.* **50**, 1240–1246 (2018).
- van der Maaten, L. & Hinton, G. Visualizing data using t-SNE. *J. Mach. Learn. Res.* **9**, 2579–2605 (2008).
- Zeid, R. et al. Enhancer invasion shapes *MYCN*-dependent transcriptional amplification in neuroblastoma. *Nat. Genet.* **50**, 515–523 (2018).
- Muhar, M. et al. SLAM-seq defines direct gene-regulatory functions of the BRD4-MYC axis. *Science* **360**, 800–805 (2018).
- Wegner, M. & Stolt, C. C. From stem cells to neurons and glia: a Soxist's view of neural development. *Trends Neurosci.* **28**, 583–588 (2005).
- Gao, Y. et al. Overcoming resistance to the THZ series of covalent transcriptional CDK inhibitors. *Cell Chem. Biol.* **25**, 135–142 (2018).
- Pugacheva, E. M. et al. Comparative analyses of CTCF and BORIS occupancies uncover two distinct classes of CTCF binding genomic regions. *Genome Biol.* **16**, 161 (2015).
- Bergmaier, P. et al. Choice of binding sites for CTCFL compared to CTCF is driven by chromatin and by sequence preference. *Nucleic Acids Res.* **46**, 7097–7107 (2018).
- Mumbach, M. R. et al. HiChIP: efficient and sensitive analysis of protein-directed genome architecture. *Nat. Methods* **13**, 919–922 (2016).
- Dowen, J. M. et al. Control of cell identity genes occurs in insulated neighborhoods in mammalian chromosomes. *Cell* **159**, 374–387 (2014).
- Sanborn, A. L. et al. Chromatin extrusion explains key features of loop and domain formation in wild-type and engineered genomes. *Proc. Natl. Acad. Sci. USA* **112**, E6456–E6465 (2015).
- Whyte, W. A. et al. Master transcription factors and mediator establish super-enhancers at key cell identity genes. *Cell* **153**, 307–319 (2013).
- Hnisz, D. et al. Super-enhancers in the control of cell identity and disease. *Cell* **155**, 934–947 (2013).
- Kumar, J. P. The sine oculis homeobox (SIX) family of transcription factors as regulators of development and disease. *Cell. Mol. Life Sci.* **66**, 565–583 (2009).
- Dennis, D. J., Han, S. & Schuurmans, C. bHLH transcription factors in neural development, disease, and reprogramming. *Brain Res.* **1705**, 48–65 (2019).

**Publisher's note:** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

© The Author(s), under exclusive licence to Springer Nature Limited 2019

## METHODS

**Cell lines.** Human neuroblastoma cell lines Kelly and SK-N-BE(2) and human Ewing sarcoma cell lines TC-32, TC-71 and CHLA-10<sup>31,32</sup> were obtained from the Children's Oncology Group cell line bank. Human neuroblastoma cell line SK-N-SH and human embryonic kidney cell line HEK293T were obtained from the American Type Culture Collection. 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 medium supplemented with 10% fetal bovine serum (FBS) and 1% penicillin/streptomycin (Life Technologies). HEK293T cells were grown in DMEM medium supplemented with 10% FBS and 1% penicillin/streptomycin (Life Technologies). Resistant cells were grown in the presence of either the ALK inhibitor, TAE684<sup>14</sup> (Kelly and SK-N-SH) or the CDK12 inhibitor, E9<sup>21</sup> (SK-N-BE(2)).

**Compounds.** TAE684 and E9 were synthesized in-house in the Gray laboratory and JQ1<sup>33</sup> was obtained from J.Qi's laboratory at the Dana-Farber Cancer Institute (DFCI). Ceritinib<sup>34</sup>, lorlatinib<sup>35</sup> and I-BET726<sup>36</sup> were purchased from Selleck Chemicals.

**Synthetic RNA spike-in and microarray analysis.** Total RNA and sample preparation was performed as previously described<sup>37</sup>. In brief, cells were either incubated in medium 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) before lysis and RNA extraction. Biological duplicates (equivalent to 5 × 10<sup>6</sup> cells per replicate) were collected and homogenized in 1 ml of TRIzol Reagent (Ambion), 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 analysed on an Agilent 2100 Bioanalyzer (Agilent Technologies) for integrity. RNA with the RNA Integrity Number above 9.8 was hybridized to Affymetrix GeneChip PrimeView Human Gene Expression arrays (Affymetrix).

**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), β-actin (4967), CTCF (3417), normal rabbit IgG (2729) and HRP anti-mouse IgG (7076) from Cell Signaling Technology; HRP anti-rabbit IgG (sc-2357) from Santa Cruz Biotechnology; BRD4 (A301-985A100) and SMC1A (A300-055A) from Bethyl Laboratories; CTCF (07-729), SOX9 (AB5535) and H3K27me3 (07-449) from Millipore; pALK-Y1507 (ab73996), BORIS (ab187163) and H3K27ac (ab2729) from Abcam; BORIS (NBP2-52405) from NOVUS Biologicals; BORIS (39851) from Active Motif; SIX1 (HPA001893) from Sigma-Aldrich; and Vysis LSI N-MYC (2p24) SpectrumGreen/Vysis CEP 2 SpectrumOrange Probe (07J72-001) from Abbott.

**Cell viability and growth curve assays.** Viability and growth experiments were performed using the CellTiter-Glo Luminescent Cell Viability Assay (Promega) according to the manufacturer's instructions, as previously described<sup>38</sup>. Cells were plated in 96-well plates at a seeding density of 4 × 10<sup>3</sup> cells per well. For growth assays, the cells were analysed 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: 2 nM to 20 µM). DMSO without drug served as a negative control. After 72 h of incubation, cells were analysed for cell viability and IC<sub>50</sub> values were determined using a nonlinear regression curve fit with GraphPad Prism 6 software.

**Cell-cycle analysis.** Cell-cycle analysis was performed 24 h after cell plating using propidium iodide staining, as previously described<sup>15</sup>. 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<sup>-1</sup> propidium iodide (BD Biosciences) and 0.2 mg ml<sup>-1</sup> 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).

**Western blotting.** Cell or tumour tissue was lysed in NP-40 buffer (Invitrogen) containing a 1× complete protease inhibitor tablet (Roche) per 10 ml buffer and a cocktail of phosphatase inhibitors (Roche). Protein concentration was measured using the DC Protein Assay (Bio-Rad); 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). The actin loading controls for the protein samples shown in the immunoblots of the following panels (two independent mouse tumour 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.** Cells were collected in immunoprecipitation lysis buffer (50 mM Tris-HCl buffer (pH 7.4), 100 mM NaCl, 1% Triton-100, 1 mM PMSF), containing a 1× complete protease inhibitor tablet (Roche) per 10 ml buffer and a cocktail of phosphatase inhibitors (Roche). Homogenates were centrifuged at 20,000g for 10 min at 4°C to obtain supernatants. DNase I (approximately 1 U ml<sup>-1</sup>) was used to degrade DNA in supernatants by incubation for 1 h at room temperature. Co-immunoprecipitation of endogenously expressed proteins was performed using protein A Dynabeads (Invitrogen), according to the manufacturer's instructions. In brief, 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 (9851, Active Motif).

**Plasmids, shRNA knockdown and overexpression systems.** pLKO.1 shRNA constructs (control: SHC007; MYCN: 1-TRCN000020694 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 D. Sabatini (Addgene plasmid 30323)<sup>39</sup>. Overexpression constructs were generated by cloning BORIS cDNA into the Tet-inducible pInducer20 vector, provided by S. Elledge (Addgene plasmid 44012)<sup>40</sup>. Production of lentiviral particles and subsequent infection were performed as previously described<sup>38</sup>. 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). Virus-containing supernatants were collected 48 h after transfection. Cells were infected with 8 µg ml<sup>-1</sup> polybrene (Sigma-Aldrich) and 24–48 h later selected with puromycin (pLKO.1) (Sigma-Aldrich) and then collected at appropriate time points. When using the Tet-inducible system for BORIS overexpression, induction of gene expression was achieved by treating cells every 2–3 days with doxycycline (0.2 µg ml<sup>-1</sup>) for a total duration of 37 days.

**qRT-PCR.** RNA isolation and PCR amplification were performed as previously described<sup>38</sup>, 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). 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). All experiments were performed in biological triplicates unless stated otherwise. Each individual biological sample was amplified by qPCR 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 on 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) was used as previously described<sup>38</sup>. 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 phosphorylated RTKs.

**Fluorescent in situ hybridization.** Fluorescent in situ hybridization (FISH) analyses were performed using a Vysis LSI N-MYC (2p24) SpectrumGreen/Vysis CEP 2 SpectrumOrange Probe (Vysis), in accordance with the manufacturer's instructions.

**Immunohistochemistry.** All human tumour specimens (formalin-fixed paraffin-embedded slides) were obtained under an Institutional Review Board (IRB)-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 using the ImmPRESS Excel Amplified HRP Polymer Staining Kit (MP-7601, Vector Laboratories) on a Dako Autostainer (Agilent Technologies). Sections were deparaffinized, rehydrated, and subjected to antigen retrieval in citrate-based buffer on a steamer for 25 min. Slides were blocked with BLOXALL blocking solution and 2.5% horse serum sequentially before 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 haematoxylin, followed by dehydration and the addition of coverslips.

**Bisulfite sequencing.** 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 bisulfite sequencing assay. Genomic DNA (500 ng) was treated with the EZ DNA Methylation-Lightning Kit (Zymo Research), 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 on request). PCR products were then sequenced for the 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  $4 \times 10^5$  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).

**Mouse experiments.** All mouse 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  $2 \times 10^6$  sensitive or resistant Kelly neuroblastoma cells into the flanks of NU/NU (Crl:NU-*Foxn1<sup>m</sup>*) (Charles River Laboratories) or NU/NU (CrTac:NCr-*Foxn1<sup>m</sup>*) (Taconic) 7-week-old female mice. Mice were randomized into groups of equal average volumes, and investigators were not blinded to group allocation during data collection. (i) To assess the tumorigenic potential of resistant cells in absence of treatment, mice with established disease (mean tumour volume of  $200 \text{ mm}^3$ ) were monitored for up to 23 days ( $n = 4$  per group). Tumours were obtained, dissociated and used to establish cell lines and for assessment of mRNA levels, protein expression and sensitivity to TAE684. (ii) To ensure that the in vitro resistance to TAE684 was maintained in vivo, mice with established disease were divided into two cohorts and were treated with either TAE684 (10 mg kg<sup>-1</sup>) 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, mice with established disease were divided into two cohorts and treated with either JQ1 (50 mg kg<sup>-1</sup>) 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 (2–3 times a week) and mouse weights were monitored at least twice a week. Tumour volumes were calculated using the modified ellipsoid formula<sup>41</sup>:  $\frac{1}{2}(\text{length} \times \text{width}^2)$ . Animals were euthanized when tumour volumes reached 1,500–2,000 mm<sup>3</sup> based on institutional IACUC criteria for maximum tumour volumes. In none of the experiments were the institutional limits for tumour volumes (<2,000 mm<sup>3</sup> measurement preceding the day of euthanization) exceeded.

**ChIP-seq.** ChIP was carried out as previously described<sup>15</sup> with minor changes as described. Approximately  $1 \times 10^7$  cells were crosslinked for 10 min at room temperature with 1% formaldehyde (Thermo Scientific) in PBS followed by quenching with 0.125 M glycine for 5 min. The cells were then washed twice in ice-cold PBS, and the cell pellets flash frozen and stored at  $-80^\circ\text{C}$ . Fifty microlitres 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  $4^\circ\text{C}$ . 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) at the following settings: 10 cycles, each for 30 s on, followed by 1 min off, at a power of approximately 20 W. The lysates were then centrifuged for 10 min at  $4^\circ\text{C}$ , 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  $4^\circ\text{C}$  with the antibody-bound magnetic beads, washed with low-salt buffer (50 mM HEPES-KOH (pH 7.5), 0.1% SDS, 1% Triton X-100, 0.1% sodium deoxycholate, 1 mM EGTA, 1 mM EDTA, 140 mM NaCl and 1× complete protease inhibitor), high-salt buffer (50 mM HEPES-KOH (pH 7.5), 0.1% SDS, 1% Triton X-100, 0.1% sodium deoxycholate, 1 mM EGTA, 1 mM EDTA, 500 mM NaCl and 1× complete protease inhibitor), LiCl buffer (20 mM Tris-HCl (pH 8), 0.5% NP-40, 0.5% sodium deoxycholate, 1 mM EDTA, 250 mM LiCl and 1× 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 was performed to pellet the magnetic beads and collect the supernatants. The crosslinking was reversed overnight at  $65^\circ\text{C}$ . 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) 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 previously described<sup>24</sup> with a few modifications. Approximately  $1 \times 10^7$  cells were crosslinked for 10 min at room temperature with 1% formaldehyde in growth medium and quenched in 0.125 M glycine. After washing twice with ice-cold PBS, the supernatant was aspirated and the cell pellet flash frozen in liquid nitrogen. Crosslinked cell pellets were thawed on ice, resuspended in 1 ml of ice-cold Hi-C lysis buffer (10 mM Tris-HCl (pH 8.0), 10 mM NaCl, 0.2% NP-40 and 1× complete protease inhibitor) and incubated at  $4^\circ\text{C}$  for 30 min with rotation. Nuclei were pelleted by centrifugation for 5 min at  $4^\circ\text{C}$  and washed once with 500 µl of ice-cold Hi-C lysis buffer. After removing the supernatant, nuclei were resuspended in 100 µl of 0.5% SDS and incubated at  $62^\circ\text{C}$  for 10 min. SDS was quenched by adding 335 µl of 1.5% Triton X-100 and incubating for 15 min at  $37^\circ\text{C}$ . After the addition of 50 µl of 10× NEB Buffer 2 (New England Biolabs, B7002) and 375 U of MboI restriction enzyme (New England Biolabs, R0147), chromatin was digested at  $37^\circ\text{C}$  for 2 h with rotation. After digestion, MboI enzyme was heat-inactivated by incubating the nuclei at  $62^\circ\text{C}$  for 20 min. To fill in the restriction fragment overhangs and mark the DNA ends with biotin, 52 µl of fill-in master mix, containing 37.5 µl of 0.4 mM biotin-dATP (Invitrogen, 19524016), 1.5 µl of 10 mM dCTP (Invitrogen, 18253013), 1.5 µl of 10 mM dGTP (Invitrogen, 18254011), 1.5 µl of 10 mM dTTP (Invitrogen, 18255018), and 10 µl of 5 U µl<sup>-1</sup> DNA Polymerase I, Large (Klenow) Fragment (New England Biolabs, M0210), were added and the tubes were incubated at  $37^\circ\text{C}$  for 1 h with rotation. Proximity ligation was performed by the addition of 948 µl of ligation master mix, containing 150 µl of 10× NEB T4 DNA ligase buffer (New England Biolabs, B0202), 125 µl of 10% Triton X-100, 7.5 µl of 20 mg ml<sup>-1</sup> BSA (New England Biolabs, B9000), 10 µl of 400 U µl<sup>-1</sup> T4 DNA ligase (New England Biolabs, M0202), and 655.5 µl of water, and incubation at room temperature for 4 h with rotation. After proximity ligation, nuclei were pelleted by centrifugation for 5 min and resuspended in 1 ml of ChIP sonication buffer (50 mM HEPES-KOH (pH 7.5), 140 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 and 1× complete protease inhibitor). Nuclei were sonicated using a Misonix 3000 sonicator (Misonix) at the following settings: 12 cycles, each for 30 s on, followed by 1 min off, at a power of approximately 20 W. Sonicated chromatin was clarified by centrifugation for 15 min at  $4^\circ\text{C}$  and the supernatant was transferred to a tube. Sixty microlitres 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 h at  $4^\circ\text{C}$  with rotation. Beads were then separated on a magnetic stand and the supernatant was transferred to a new tube. Seventy-five microlitres of protein G Dynabeads pre-incubated overnight at  $4^\circ\text{C}$  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  $4^\circ\text{C}$  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 buffer 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  $65^\circ\text{C}$  for 15 min. To purify the eluted DNA, RNA was degraded by the addition of 8.5 µl of 10 mg ml<sup>-1</sup> RNase A and incubation at  $37^\circ\text{C}$  for 2 h. Protein was degraded by the addition of 20 µl of 10 mg ml<sup>-1</sup> proteinase K and incubation at  $55^\circ\text{C}$  for 45 min. Samples were then incubated at  $65^\circ\text{C}$  overnight to reverse crosslink protein-DNA complexes. DNA was then purified using Zymo ChIP DNA Clean and Concentrator columns (Zymo, D5205) according to the 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 2× biotin binding buffer (10 mM Tris-HCl (pH 7.5), 1 mM EDTA pH 8.0, 2 M NaCl). Then, 25 ng of purified DNA was added in a total volume of 10 µl water to the beads and incubated at room temperature for 15 min with agitation every 5 min. After capture, beads were separated with a magnet and the supernatant was discarded. Beads were then washed twice with 500 µl of Tween wash buffer, incubating at  $55^\circ\text{C}$  for 2 min with shaking for each wash. Beads were resuspended in 25 µl of Nextera Tagment DNA buffer. To tagment the captured DNA, 1 µl of Nextera Tagment DNA Enzyme 1 was added with 24 µl of Nextera Resuspension Buffer and samples were incubated at  $55^\circ\text{C}$  for 10 min with shaking. Beads were separated on a magnet and supernatant was discarded. Beads were

washed twice with 500 µl of 50 mM EDTA at 50 °C for 30 min, washed twice with 500 µl of Tween wash buffer at 55 °C for 2 min each, and finally washed once with 500 µl of 10 mM Tris-HCl (pH 7.5) for 1 min at room temperature. Beads were separated on a magnet and supernatant was discarded. To generate the sequencing library, PCR amplification of the fragmented DNA was performed while the DNA was still bound to the beads. Beads were resuspended in 15 µl of Nextera PCR Master Mix, 5 µl of Nextera PCR Primer Cocktail, 5 µl of Nextera Index Primer 1, 5 µl of Nextera Index Primer 2 and 20 µl water. DNA was amplified with 9–10 cycles of PCR. After PCR, beads were separated on a magnet and the supernatant containing the PCR-amplified library was transferred to a new tube, purified using Zymo DNA Clean and Concentrator columns (Zymo, D5205) according to the manufacturer's protocol, and eluted in 14 µl water. Purified HiChIP libraries were size-selected to 300–700 bp using a Sage Science Pippin Prep instrument according to the manufacturer's protocol and subjected to 2 × 100 paired-end sequencing using an Illumina HiSeq 2500 system (SY-401-2501, Illumina).

**scRNA-seq.** Kelly cells (sensitive, intermediate and resistant states) were grown to 70% confluence in T75 culture flasks. In brief, growth medium was aspirated and cells were treated with 0.25% Trypsin/EDTA for 3 min at 37 °C, after which cells were washed twice with 1 × PBS. Cells were then resuspended into single cells at a concentration of  $1 \times 10^6$  per ml in 1 × 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) as previously described<sup>42</sup>, 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 × 150 paired-end reads, one full lane per sample, for approximately >90% sequencing saturation.

**Genomics analysis: direct comparison of CTCF and BORIS expression in healthy and tumour samples.** To assess the expression levels and range of *BORIS* and *CTCF* in healthy and tumour cells all GTEx, TCGA and TARGET datasets were downloaded and converted to FPKM values and displayed as [ $\log_2(\text{FPKM} + 1)$ ] (Extended Data Fig. 1a, b).

**Association of BORIS with prognostic features.** For each dataset, processed values were extracted from the Gene Expression Omnibus (GEO) and scaled values were created by normalizing the expression levels by the minimum mean value of the conditions that were compared,  $E_{i,j} = E_{i,j}/\min(\text{average}(E_j))$ . 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 analysed 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<sup>37</sup>. In brief, all microarray chip data were imported in R (<https://www.r-project.org/>, v.3.1.3) using the affy package<sup>43</sup> (v.1.44.0), 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<sup>44</sup> (v.3.22.7) and a FDR value of 0.05. Spike-in normalized absolute expression values (counts) were normalized to 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 *BORIS* shRNA sample was first normalized to the control shRNA 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 (v.0.11.5) and samples were aligned to the human genome (build hg19, GRCh37.75) with STAR (v.2.5.1b\_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 (v.1.3.1) and MarkDuplicates (v.2.1.1) from Picard tools. For each experimental replicate, antibody enrichment was assessed using the plotFingerprint command from deepTools (v.2.2.4). Peaks were identified with MACS2 (v.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 per bp to account for differences in library size. Bigwig files were created with bedGraphToBigWig (v.4). 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/>) (v.3.5.1) using the data.table (v.1.12.2) 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. In brief, 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 the TSS, transcript length and alphabet rank, in that order, until only one 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 [ $\log_2(\text{resistant/sensitive})$ ] was calculated. Peak regions with a twofold density increase or decrease were classified as resistant or sensitive cell-specific peaks, respectively, whereas 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 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 ( $\pm 2$  kb TSS); (ii) BRD4<sup>+</sup> H3K27ac<sup>+</sup> (active) enhancers; (iii) BRD4<sup>-</sup> H3K27ac<sup>+</sup> 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 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 TSSs were extended in both directions by 2 kb and binned in 50-bp bins, and 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 per bp). An aggregated summary profile was created for each cell type. To account for different numbers of identified enhancers in both cell types we calculated a normalization factor ( $N_{\text{resistant enhancers}}/N_{\text{sensitive enhancers}}$ ) 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 trimomatic<sup>45</sup> (v.0.36) and were then processed using the HiC-Pro (v.2.10.0) pipeline<sup>46</sup> 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 one-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<sup>47</sup> (v.0.7.3), with parameter ‘peaks = combined, all’, and subsequently diffloop<sup>47</sup> (v.1.10.0) with default settings. Only loops that were detected in all three biological replicates of a sample (sensitive, resistant, shCtrl or shBORIS) with a minimum of five paired-end tags 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  $\log_2$ -transformed fold change  $> 1$ ) or lost (mango.FDR  $< 0.01$  and  $\log_2$ -transformed fold change  $< -1$ ).

**Classification of HiChIP interactions.** SMC1A-based HiChIP interactions (loops) were classified as previously described<sup>48</sup> with minor adaptations. Associated anchors of loops were overlapped with our ChIP-seq peaks (CTCF, BORIS, H3K27ac, BRD4) and promoter regions (TSS  $\pm 2$  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 four 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 (sensitive versus resistant) and BORIS depletion (shBORIS versus shCtrl) 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 were identified using the ROSE algorithm (v.1) ([https://bitbucket.org/young\\_computation/rose](https://bitbucket.org/young_computation/rose)). In short, H3K27ac-enriched regions were identified with MACS2 and termed typical enhancers. 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 super-enhancers were identified by merging both sensitive- and resistant-cell super-enhancers and determining cell-type specificity based on the differential normalized read density of both H3K27ac and BRD4. In brief, 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 super-enhancer with at least a minimum 0.75 change for each individual mark. Super-enhancers 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 super-enhancers 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.** 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 per 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, and 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 reactivation in resistant cells. For each gene, five features were examined: (i) creation of a unique regulatory interaction; (ii) deposition of BORIS on its promoter or looped enhancer; (iii) association with a resistant cell-specific super-enhancer through overlap with either its promoter or looped anchor; (iv) increased mRNA expression; and (v) transition from a closed or neutral state to an open chromatin state. A unique set of 89 genes (Supplementary Table) that exhibited four out of five 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.cbrckau.edu.sa/tcof/>) (Fig. 4b).

**Allen Brain atlas gene signature.** Expression data and metadata for human brain development were downloaded from the Allen Brain atlas (<http://www.brainspan.org>). Row-normalized z-scores of [ $\log_2(\text{RPKM} + 1)$ ] values were used to create a heat map. 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 ( $\pm 2$  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 ( $\pm 2$  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 motif enrichment analysis.** Statistically overrepresented motifs were identified with HOMER<sup>49</sup> (v.2) using the command findMotifs.pl providing both target and background fasta sequences for regions of interest. For promoter regions we selected the top 1,000 up- and downregulated 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 (v.2.50.1) 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 super-enhancers. Target and background fasta sequences were extracted from the summits of BRD4 peaks located on cell-type-specific super-enhancers and extended by 500 bp in both directions. For a selection of enriched sequences, the associated transcription factor motif and significance level (*P*) was visualized using a heat map (Fig. 4d).

**scRNA-seq analysis.** The Cell Ranger Single Cell Software Suite, v.1.3 was used to perform sample de-multiplexing, barcode and unique molecular identifier (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 fewer than 1,000 genes were detected. To remove low-quality cells, we calculated five technical indicators (ratio of detected genes/UMI, percentage of mitochondrial genes, percentage of ribosomal genes, average GC content of library and library complexity measured by Shannon Entropy) and performed 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 neighbour 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’ (v.1.10.0) to confirm that there were no technical or experimental confounding effects and the R package ‘Seurat’ (v.2.3.4) to analyse and visualize the data. In brief, UMI counts were log-normalized with a scale factor of 10,000 and subsequently centre-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 heat map and elbow plot observation. To identify homogeneous subpopulations, we performed iterative clustering using the network-based clustering algorithm (shared nearest neighbour) with different resolutions as input until each sample was at least separated in two groups. A simple pseudotime analysis was performed by calculating an average expression profile for each identified subpopulation and ordering them according to the summarized expression of transcription factors that displayed variable expression between sensitive and intermediate or intermediate and resistant cells. Variable expression was defined as showing at least a 33% change in the rank of expression between two samples with a minimal normalized expression level  $> 0.2$ . For each sample comparison, at least the top 10 most variable transcription factors were included. In total this resulted in 32 transcription factors. Gene expression values were then linearly rescaled between 0 and 10 to jointly visualize relative expression changes during this pseudotime. To examine co-detection or mutual exclusivity between genes of interest, a two-sided Fisher’s exact test was performed for all cells in a given sample. A score combining both the odds ratio and the  $-\log_{10}(P \text{ value})$  was calculated to visualize both the strength and direction between genes in pairwise co-expression tests.

**Statistical analysis.** Analysis for each plot is listed in the figure legend and/or in the corresponding Methods. In brief, all grouped data are presented as mean  $\pm$  s.d. unless stated otherwise. All box and whisker plots of expression data are presented as: centre lines, medians; box limits, twenty-fifth and seventy-fifth percentiles; whiskers, minima and maxima ( $1.5 \times$  the interquartile range). Statistical significance for pairwise comparisons was determined using the two-sided Wilcoxon rank-sum test or two-sided unpaired *t*-test, unless stated otherwise. Survival analysis was performed using the Kaplan–Meier method and differences between groups calculated by the two-sided log-rank test and the Bonferroni correction method. Tumour volume comparisons for the xenograft studies were analysed by Mann–Whitney *U* test. \**P*  $< 0.05$ ; \*\**P*  $< 0.01$ . Statistical comparisons of distributions of

fold changes for the expression microarrays were done using the Mann–Whitney *U* test. All quantitative analyses are expressed as the mean  $\pm$  s.d. of three biological replicates, unless stated otherwise. Microarray and ChIP-seq data are based on at least two independent experiments. For all experiments, no statistical methods were used to predetermine sample size. Unless stated otherwise, experiments were not randomized and 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](https://github.com/RubD/GeTrackViz2)) or with the circlize package (v.0.4.5) in R.

**Retrospective analysis of gene expression in human samples.** Gene expression levels or correlations across primary tumours, 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 tumour types with poor prognosis feature annotations (GSE49710 (Neuroblastoma)<sup>50</sup>, GSE17679 (Mixed Ewing Sarcoma)<sup>51</sup>, GSE63074 (Non-small cell lung carcinoma)<sup>52</sup>, GSE15709 (ovarian cancer)<sup>53</sup>, GSE16179 (breast cancer)<sup>54</sup> and GSE7181 (Glioblastoma)<sup>55</sup>).

**Reporting summary.** Further information on research design is available in the Nature Research Reporting Summary linked to this paper.

## Data availability

The microarray, ChIP-seq, HiChIP and scRNA-seq datasets generated and analysed during the current study are available in the Gene Expression Omnibus (GEO) 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.

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 Res.* **64**, 5415–5424 (2004).
32. Whang-Peng, J. et al. Cytogenetic characterization of selected small round cell tumors of childhood. *Cancer Genet. Cytogenet.* **21**, 185–208 (1986).
33. Filippakopoulos, P. et al. Selective inhibition of BET bromodomains. *Nature* **468**, 1067–1073 (2010).
34. 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-(isopropylsulfonyl)phenyl)pyrimidine-2,4-diamine (LDK378) currently in phase 1 and phase 2 clinical trials. *J. Med. Chem.* **56**, 5675–5690 (2013).
35. Johnson, T. W. et al. Discovery of (10*R*)-7-amino-12-fluoro-2,10,16-trimethyl-15-oxo-10,15,16,17-tetrahydro-2*H*-8,4-(metheno)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. *J. Med. Chem.* **57**, 4720–4744 (2014).
36. Gosmini, R. et al. The discovery of I-BET726 (GSK1324726A), a potent tetrahydroquinoline ApoA1 up-regulator and selective BET bromodomain inhibitor. *J. Med. Chem.* **57**, 8111–8131 (2014).
37. Lovén, J. et al. Revisiting global gene expression analysis. *Cell* **151**, 476–482 (2012).
38. 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 (2016).
39. Sancak, Y. et al. The Rag GTPases bind Raptor and mediate amino acid signaling to mTORC1. *Science* **320**, 1496–1501 (2008).
40. Meerbrey, K. L. et al. The pINDUCER lentiviral toolkit for inducible RNA interference in vitro and in vivo. *Proc. Natl. Acad. Sci. USA* **108**, 3665–3670 (2011).
41. Tomayko, M. M. & Reynolds, C. P. Determination of subcutaneous tumor size in athymic (nude) mice. *Cancer Chemother. Pharmacol.* **24**, 148–154 (1989).
42. Zheng, G. X. et al. Massively parallel digital transcriptional profiling of single cells. *Nat. Commun.* **8**, 14049 (2017).
43. Gautier, L., Cope, L., Bolstad, B. M. & Irizarry, R. A. affy—analysis of Affymetrix GeneChip data at the probe level. *Bioinformatics* **20**, 307–315 (2004).
44. Smyth, G. K., Yang, Y. H. & Speed, T. Statistical issues in cDNA microarray data analysis. *Methods Mol. Biol.* **224**, 111–136 (2003).
45. Bolger, A. M., Lohse, M. & Usadel, B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* **30**, 2114–2120 (2014).
46. Servant, N. et al. HiC-Pro: an optimized and flexible pipeline for Hi-C data processing. *Genome Biol.* **16**, 259 (2015).
47. Lareau, C. A. & Aryee, M. J. diffloop: a computational framework for identifying and analyzing differential DNA loops from sequencing data. *Bioinformatics* **34**, 672–674 (2018).
48. Ji, X. et al. 3D chromosome regulatory landscape of human pluripotent cells. *Cell Stem Cell* **18**, 262–275 (2016).
49. Heinz, S. et al. Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. *Mol. Cell* **38**, 576–589 (2010).
50. SEQC/MAQC-III Consortium. A comprehensive assessment of RNA-seq accuracy, reproducibility and information content by the Sequencing Quality Control Consortium. *Nat. Biotechnol.* **32**, 903–914 (2014).
51. 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 Oncol.* **2011**, 168712 (2011).
52. Huang, S. et al. Analytical performance of a 15-gene prognostic assay for early-stage non-small-cell lung carcinoma using RNA-stabilized tissue. *JMD* **17**, 438–445 (2015).
53. Li, M. et al. Integrated analysis of DNA methylation and gene expression reveals specific signaling pathways associated with platinum resistance in ovarian cancer. *BMC Med. Genomics* **2**, 34 (2009).
54. Liu, L. et al. Novel mechanism of lapatinib resistance in HER2-positive breast tumor cells: activation of AXL. *Cancer Res.* **69**, 6871–6878 (2009).
55. Beier, D. et al. CD133<sup>+</sup> and CD133<sup>−</sup> glioblastoma-derived cancer stem cells show differential growth characteristics and molecular profiles. *Cancer Res.* **67**, 4010–4015 (2007).

**Acknowledgements** We thank the George, Young and Gray laboratories and J. R. Gilbert for discussions, and C. Li for assistance with the HiChIP experiments. We thank J. Qi for providing JQ1, and D. Sabatini and S. Elledge 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 the human tumour samples. We thank Applied Pathology Systems, the DFCI Molecular Biology Core, the Whitehead Genome Technology Core, D. Adeegbe and the NYULH Genome Technology Center for technical support. The results shown here are in part based on 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.

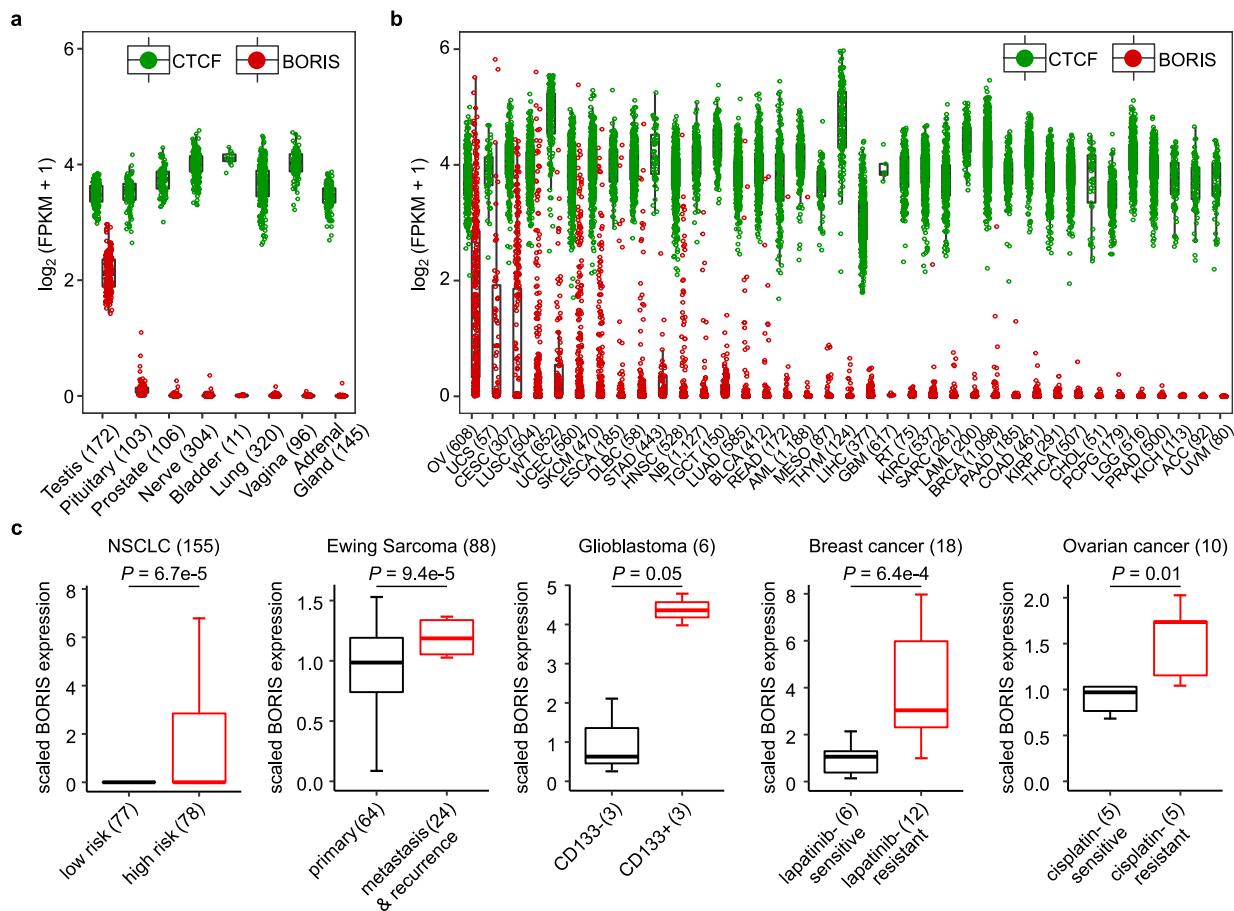
**Author contributions** D.N.D. and R.E.G. conceived the project and designed the experiments. D.N.D. planned and performed the molecular, cellular and genomic studies. R.D. performed computational analyses with input from D.S.D. and E.M. S.S. contributed to the ChIP-seq and HiChIP experiments. D.S. and S.H.O. contributed to the HiChIP experiment. D.N.D., Y.G. and T.C. performed the mouse experiments. B.S. and M.M. provided technical assistance. H.H. performed the co-immunoprecipitation experiments. 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 mouse 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 laboratory receives or has received research funding from Novartis, Takeda, Astellas, Taiho, Janssen, Kinogen, Voronoi, Her2IIC, 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.

## Additional information

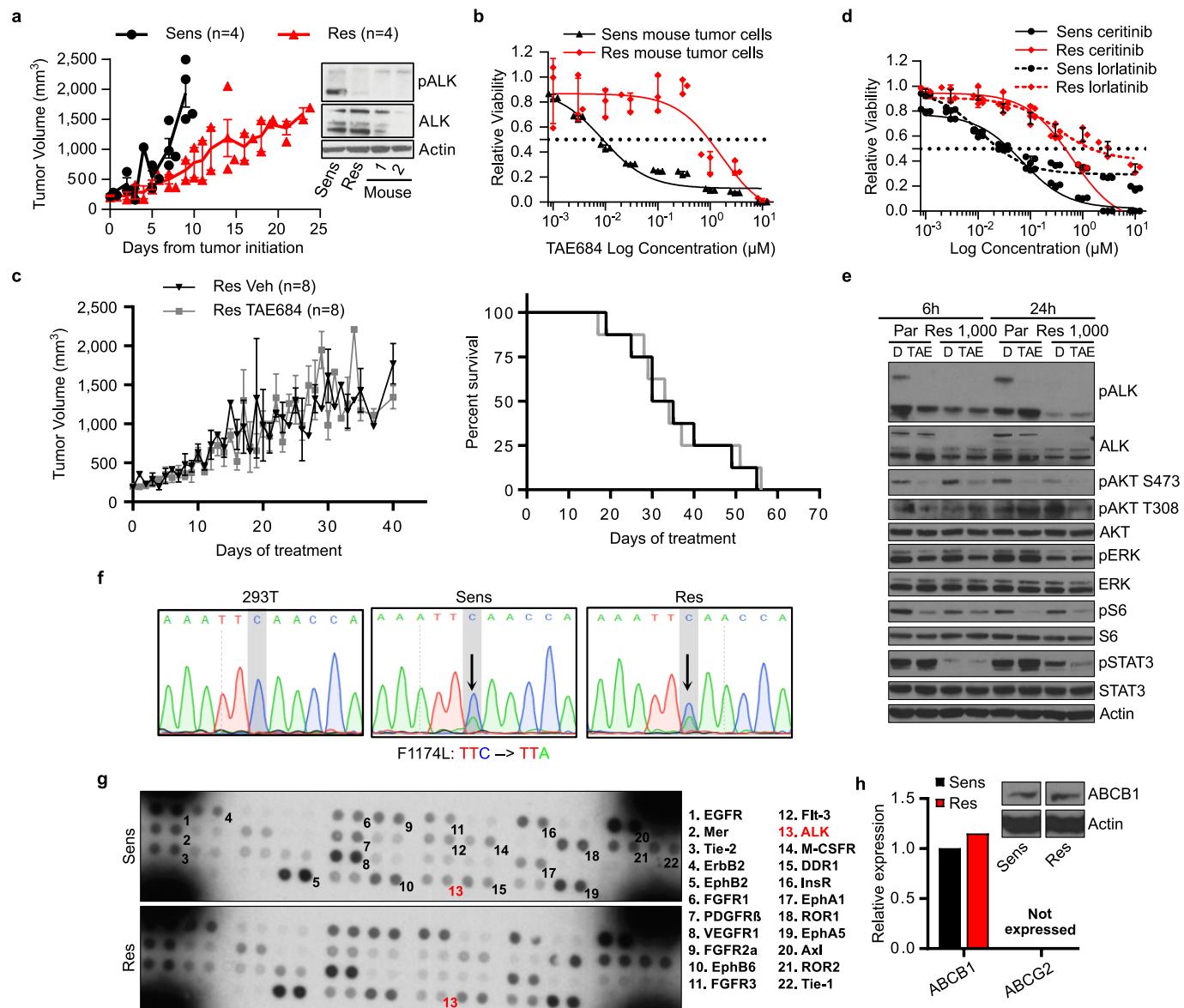
**Supplementary information** is available for this paper at <https://doi.org/10.1038/s41586-019-1472-0>.

**Correspondence and requests for materials** should be addressed to R.E.G. **Reprints and permissions information** is available at <http://www.nature.com/reprints>.



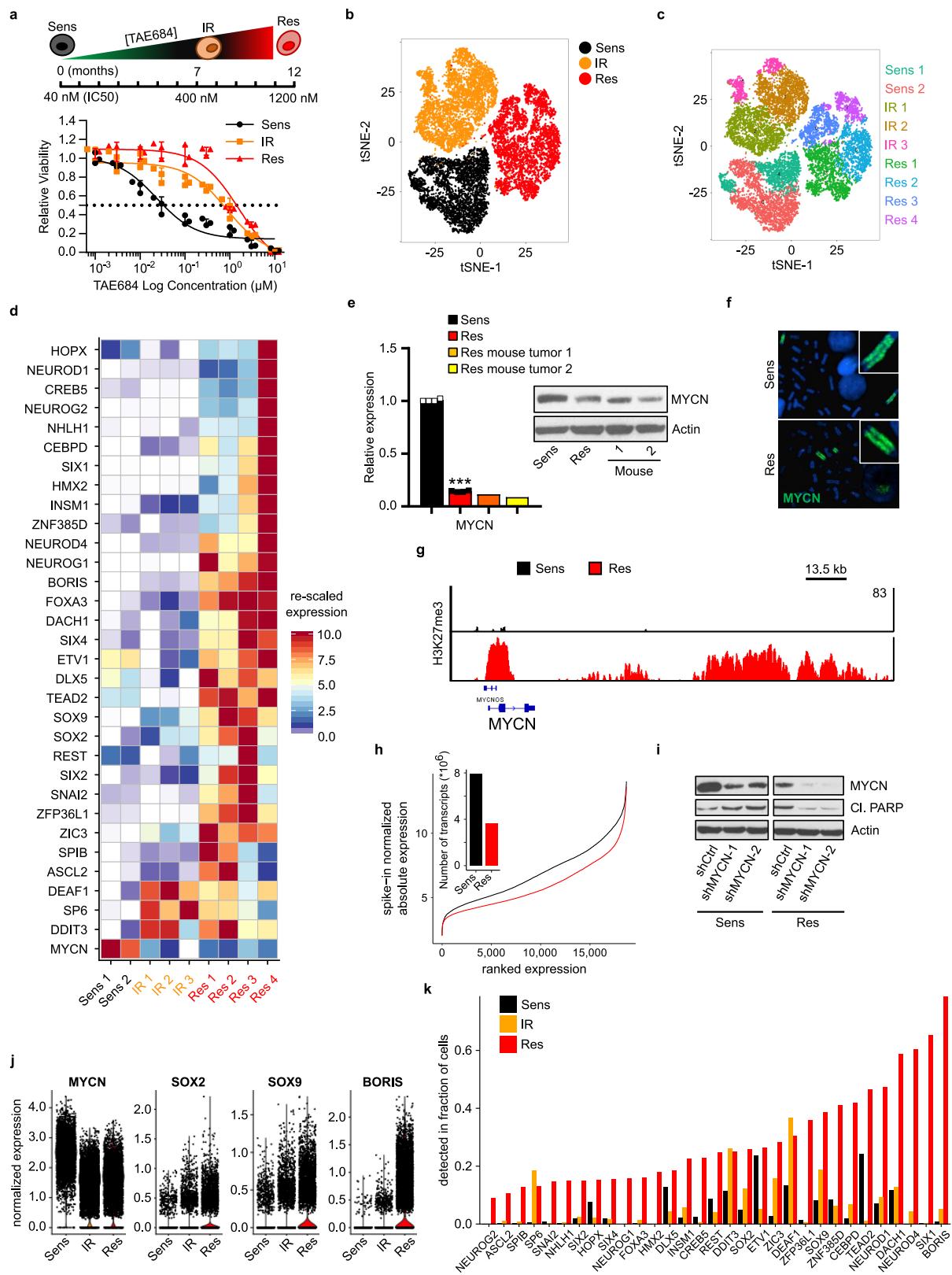
**Extended Data Fig. 1 | BORIS is expressed in several cancers and associated with high-risk features.** **a, b**, Relative mRNA expression [ $\log_2(\text{FPKM} + 1)$ ] of CTCF and BORIS in normal tissues (a) and in various cancer types based on TCGA datasets (b). FPKM, fragments per kilobase of transcript per million mapped reads. Keys to cancer types: ACC, adrenocortical carcinoma; AML, acute myeloid leukaemia; BLCA, bladder urothelial carcinoma; BRCA, breast invasive carcinoma; CESC, cervical squamous cell carcinoma and endocervical adenocarcinoma; CHOL, cholangiocarcinoma; COAD, colon adenocarcinoma; DLBC, diffuse large B-cell lymphoma; ESCA, oesophageal carcinoma; GBM, glioblastoma multiforme; HNSC, head and neck squamous cell carcinoma; LGG, low-grade glioma; KICH, kidney chromophobe; KIRC, renal clear cell carcinoma; KIRP, kidney renal papillary cell carcinoma; LAML, acute myeloid leukaemia; LIHC, hepatocellular carcinoma; LUAD, lung adenocarcinoma; LUSC, lung squamous cell carcinoma; MESO, mesothelioma; NB, neuroblastoma; OV, serous ovarian cystadenocarcinoma; PAAD, pancreatic adenocarcinoma; PCPG,

pheochromocytoma and paraganglioma; PRAD, prostate adenocarcinoma; READ, rectum adenocarcinoma; RT, rhabdoid tumour; SARC, sarcoma; SKCM, skin cutaneous melanoma; STAD, stomach adenocarcinoma; TGCT, testicular germ cell tumour; THCA, thyroid carcinoma; THYM, thymoma; UCEC, uterine corpus endometrial carcinoma; UCS, uterine carcinosarcoma; UVM, uveal melanoma; WT, Wilms tumour. **c**, Box plots showing the correlation of BORIS expression with risk status, tumour stage (primary versus metastasis/recurrence), presence of cancer stem cells (CD133 positivity) and response to targeted (lapatinib) or cytotoxic (cisplatin) therapy in the tumour 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).  $P$  values determined by two-sided Wilcoxon rank-sum test. For all panels, sample sizes ( $n$ ) are depicted in parenthesis and box plots are as defined in Fig. 4.



**Extended Data Fig. 2 | ALK inhibitor-resistant cells exhibit stable resistance in vivo and no longer rely on ALK signalling.** **a**, Left, tumour volumes of sensitive and resistant cell xenografts in untreated NU/NU (*Crl:NU-Foxn1<sup>mu</sup>*) mice established by subcutaneous injection of  $2 \times 10^6$  cells into both flanks. Animals were euthanized when tumours reached  $1,500$ – $2,000 \text{ mm}^3$ . Data are mean  $\pm$  s.e.m.,  $n = 4$  per arm. Right, immunoblot analysis of total and phosphorylated ALK in TAE-resistant xenograft tumours (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 tumour xenografts as in **a** ( $\text{IC}_{50}$  values: sensitive,  $7.9 \text{ nM}$ ; resistant,  $878.6 \text{ nM}$ ). Data are mean  $\pm$  s.d.,  $n = 3$  biological replicates. **c**, Tumour volumes (left) and Kaplan–Meier survival curves (right) of resistant cell xenografts in NU/NU (*CrTac:NCr-Foxn1<sup>mu</sup>*) mice treated with TAE684 ( $10 \text{ mg kg}^{-1}$  by oral gavage once daily) or vehicle control for up to 56 days. Data are mean  $\pm$  s.e.m.,  $n = 8$  per arm.  $P$  values determined by Mann–Whitney  $U$  test for tumour volumes ( $P = 0.8404$ ) and by log-rank test for Kaplan–Meier survival analysis ( $P = 0.8076$ ), both two-sided. **d**, Dose–response curves for TAE684-sensitive and -resistant

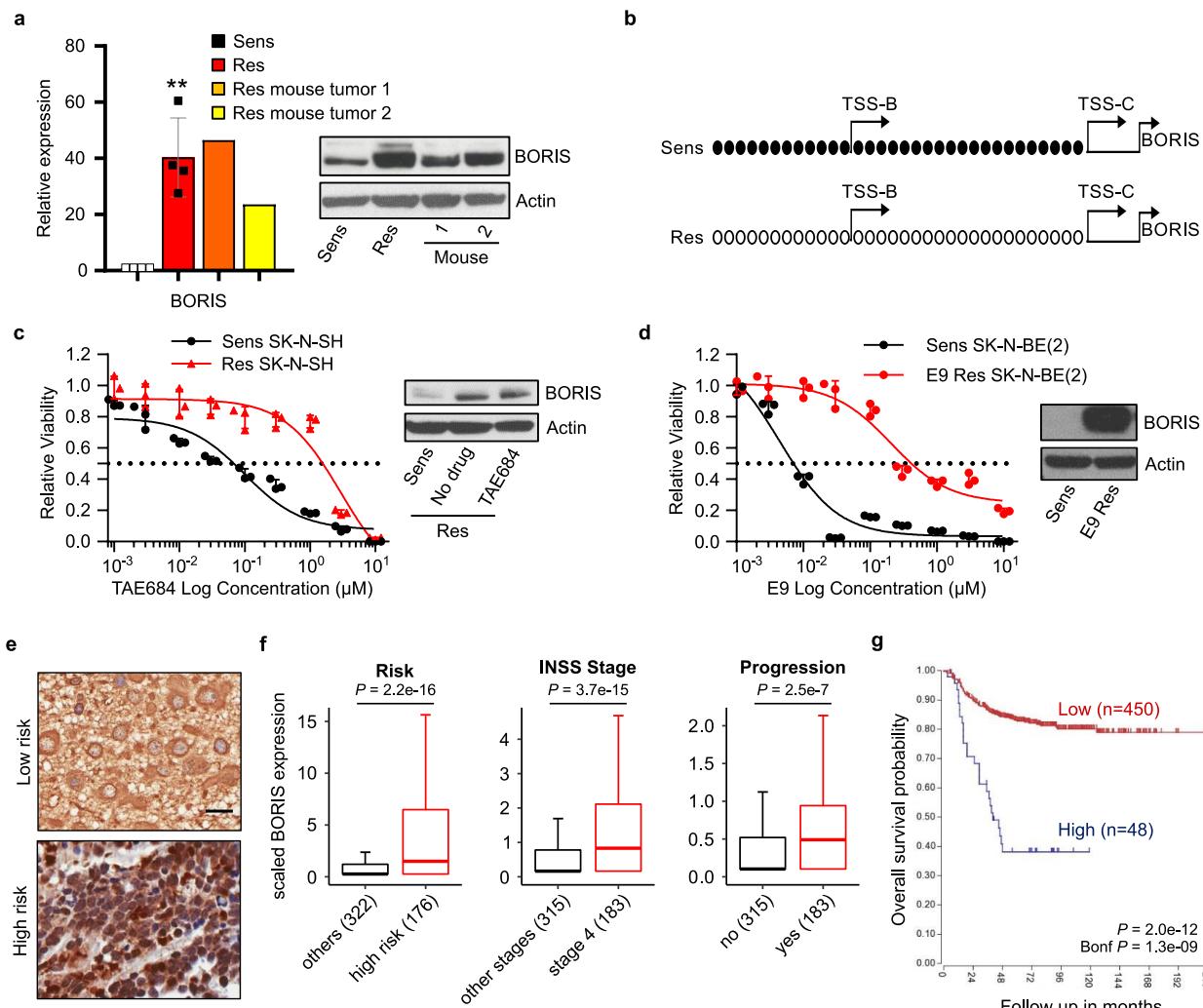
cells treated with ceritinib ( $\text{IC}_{50}$  values: sensitive,  $33.8 \text{ nM}$ ; resistant,  $446.5 \text{ nM}$ ) or lorlatinib ( $\text{IC}_{50}$  values: sensitive,  $47.5 \text{ nM}$ ; resistant,  $2,318 \text{ nM}$ ). Data are mean  $\pm$  s.d.,  $n = 3$  biological replicates. **e**, Immunoblot analysis of the indicated proteins in sensitive and resistant cells treated with DMSO or  $1 \mu\text{M}$  TAE684 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 wild-type 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**, Quantitative reverse transcription PCR (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. In **a** (immunoblot), **d**, **f** and **g**, data are representative of two independent experiments (see Supplementary Note 1 for details; for gel source data, see Supplementary Fig. 1).



Extended Data Fig. 3 | See next page for caption.

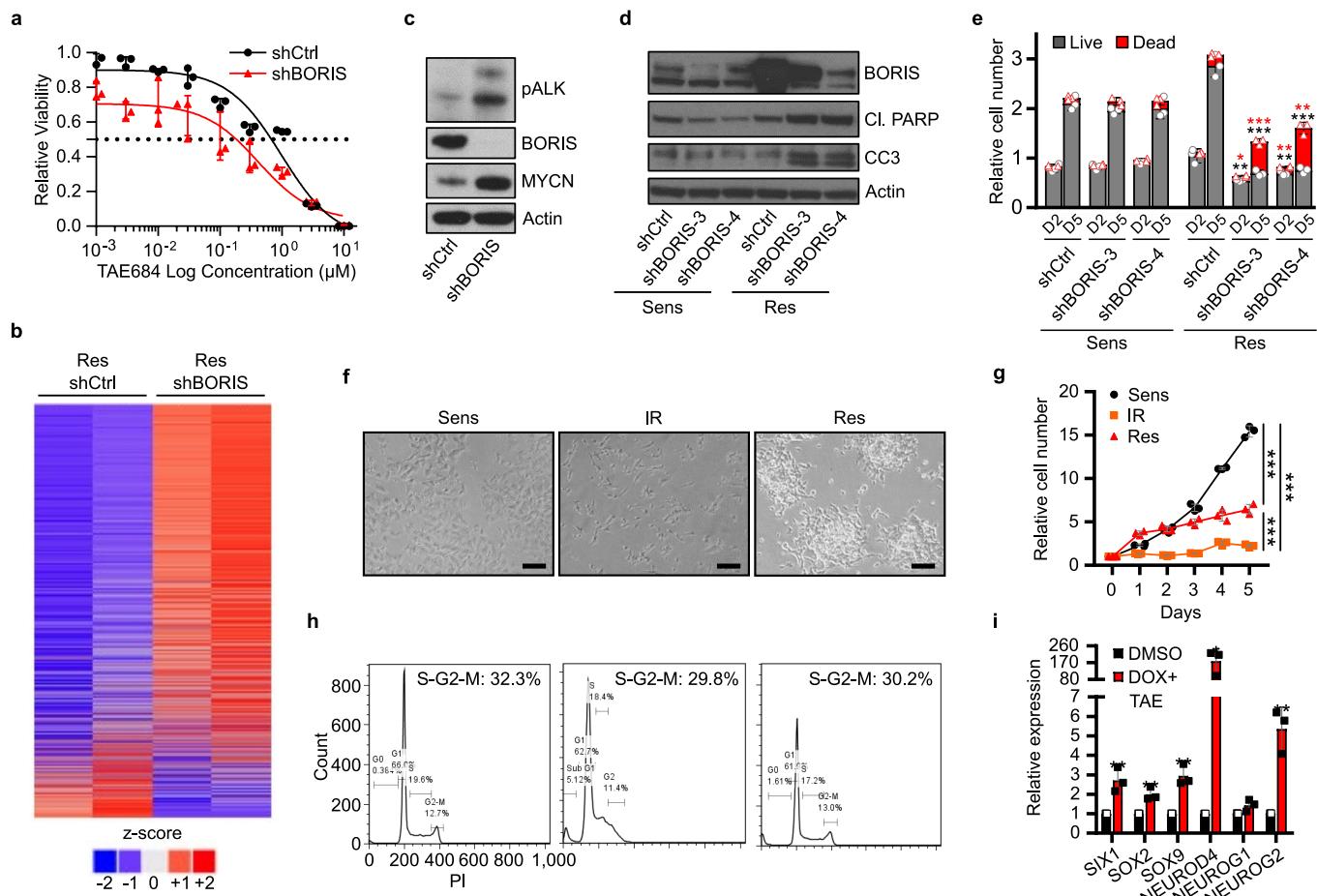
**Extended Data Fig. 3 | Development of resistance is associated with loss of MYCN followed by gradual induction of proneural transcription factors.** **a**, TAE684 dose-response curves of Kelly neuroblastoma cells during resistance establishment ( $IC_{50}$  values: sensitive, 39.4 nM; intermediate, 618 nM; resistant, 1,739 nM). Data are mean  $\pm$  s.d.,  $n = 3$  biological replicates. Schematic representation of development of resistance 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**, Heat map of rescaled gene expression values of the most variable ranked transcription factors in the three cell states. **e**, qRT–PCR and immunoblot analysis of MYCN expression in TAE684-resistant xenograft tumours (1 and 2) and sensitive and resistant cells in culture. The qRT–PCR data are mean  $\pm$  s.d.,  $n = 4$  biological replicates for sensitive and resistant cells

( $***P = 1.396 \times 10^{-11}$ ; unpaired two-sided *t*-test) and  $n = 3$  technical replicates for each tumour. **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 top 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 versus resistant cells. Bar plot, total transcriptional yield in sensitive or 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**, Bar plot showing the fractions of cells with detectable mRNA levels of the same genes as in **d**. In **e** (immunoblot) and **f–i**, data are representative of two independent experiments (for gel source data, see Supplementary Fig. 1).



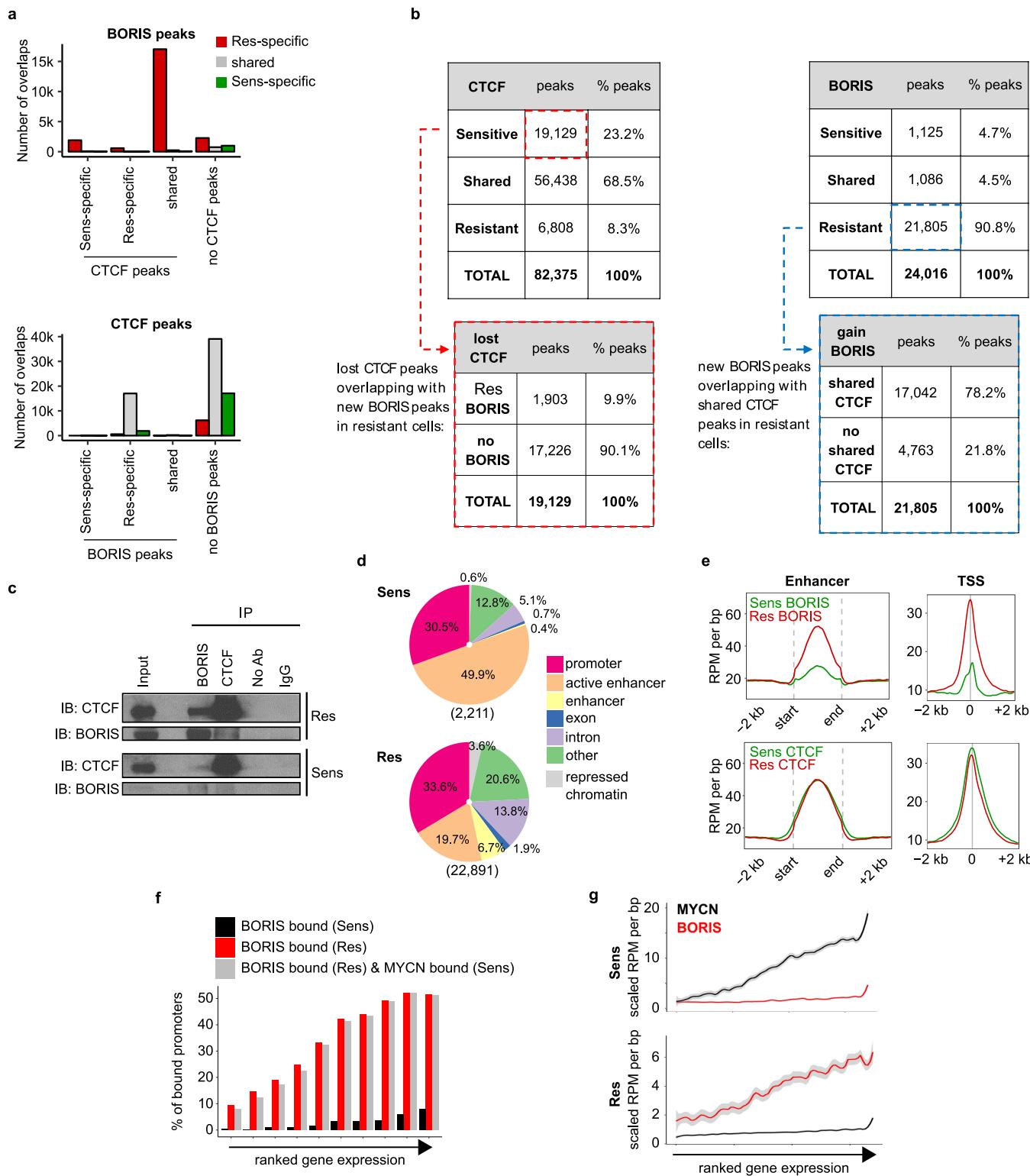
**Extended Data Fig. 4 | Overexpression of BORIS is seen in resistance models of neuroblastoma and correlates with high-risk disease and a poor outcome.** **a**, qRT-PCR and immunoblot analysis of BORIS expression in TAE684-resistant Kelly cell xenograft tumours (1 and 2) and sensitive and resistant cells in culture. The qRT-PCR data are mean  $\pm$  s.d.,  $n = 4$  biological replicates for sensitive and resistant cells ( $**P = 0.0014$ ; unpaired two-sided *t*-test) and  $n = 3$  technical replicates for each tumour. **b**, Bisulfite sequencing of the BORIS promoter in sensitive and resistant cells. Black circles represent methylated cytosine residues in a CpG dinucleotide, empty circles are unmethylated cytosines. The B and C TSSs 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 neuroblastoma cells (IC<sub>50</sub> values: sensitive, 47.9 nM; resistant, 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) neuroblastoma cells (IC<sub>50</sub> values: sensitive, 9.5 nM; resistant, 638 nM). Data are mean  $\pm$  s.d.,  $n = 3$  biological replicates for **c** (left) and **d** (left). **e**, Immunohistochemical staining of BORIS expression in primary neuroblastoma tumour samples (representative of four independent experiments). Scale bar, 20  $\mu$ m. **f**, Box plots showing correlation of BORIS expression with the indicated parameters in a human neuroblastoma dataset ( $n = 498$ ; Tumour Neuroblastoma-SEQC-498; R2: Genomics Analysis and Visualization Platform (<http://r2.amc.nl>)). Box plots are as defined in Fig. 4. *P* values were determined by two-sided Wilcoxon rank-sum test. **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). In **a**, **c**, **d** (immunoblots) and **b**, data are representative of two independent experiments. Sample sizes (*n*) are depicted in parenthesis for **f** and **g** (for gel source data, see Supplementary Fig. 1).



**Extended Data Fig. 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 ( $\text{IC}_{50}$  values: shCtrl, 537.7 nM; shBORIS, 141.2 nM). Data are mean  $\pm$  s.d.,  $n = 3$  biological replicates. **b**, Heat map 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**, **e**, Immunoblot analysis of the indicated proteins (Cl., cleaved; CC3, cleaved caspase 3) (**d**), and quantification of trypan blue staining (**e**) in sensitive and resistant cells expressing control (shCtrl) or BORIS (shBORIS-3 and -4) shRNAs. Data are mean  $\pm$  s.d.,  $n = 3$

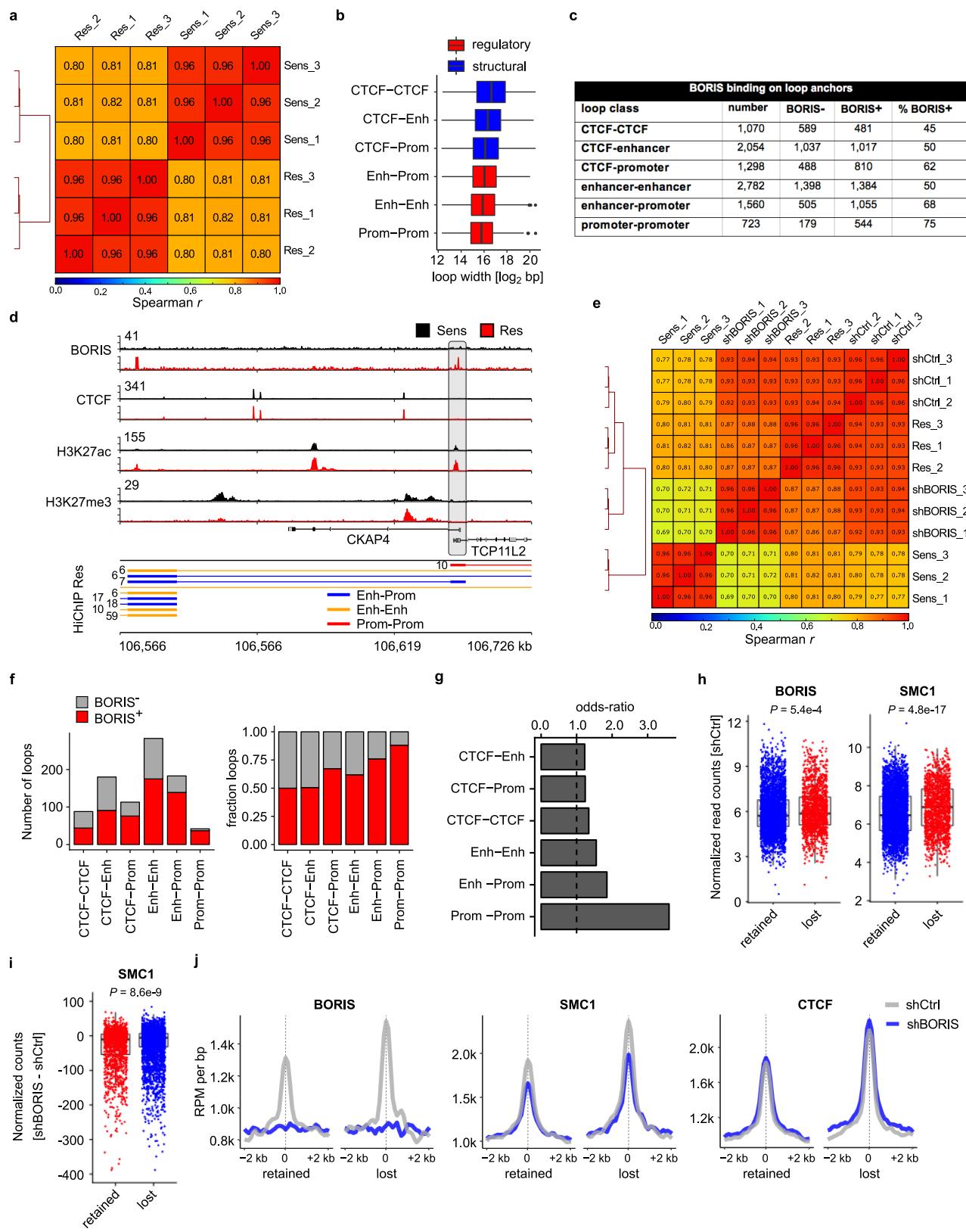
biological replicates (\* $P < 0.05$ ; \*\* $P < 0.01$ ; \*\*\* $P < 0.001$ ; unpaired two-sided *t*-tests). **f–h**, Phase-contrast microscopy images (scale bars, 150  $\mu\text{m}$ ) (**f**), growth curves (**g**) and flow cytometry analyses (**h**) of propidium iodide (PI) staining in sensitive, intermediate and resistant cells. Data are mean  $\pm$  s.d.,  $n = 3$  biological replicates (\*\* $P < 0.0001$  for all comparisons; two-way ANOVA). **i**, qRT-PCR analysis of the expression of the indicated proneural transcription factors in the same sensitive (DMSO) versus MYCN<sup>KD</sup> and BORIS<sup>Ind</sup> (DOX + TAE) cells as in Fig. 1g. Data are mean  $\pm$  SD,  $n = 3$  biological replicates (\* $P < 0.05$ ; \*\* $P < 0.01$ ; unpaired two-sided *t*-tests). In **c**, **d**, **f** and **h**, data are representative of two independent experiments (for gel source data, see Supplementary Fig. 1).



Extended Data Fig. 6 | See next page for caption.

**Extended Data Fig. 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. Most resistant cell-specific BORIS peaks (red) colocalize with CTCF peaks that are shared between the two cell types. The markedly lower number of BORIS peaks that are unique to sensitive cells (green) or shared between sensitive and resistant cells (grey) typically do not overlap with CTCF peaks that are shared or specific to any cell type (top). Most CTCF peaks are shared (grey) between sensitive and resistant cells and either do not overlap with BORIS peaks, or overlap only with those restricted to resistant cells (bottom). **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 (Ab) serve as controls. **d**, Pie charts depicting the percentages of genomic regions bound by BORIS in sensitive (top) and resistant (bottom) cells. Numbers

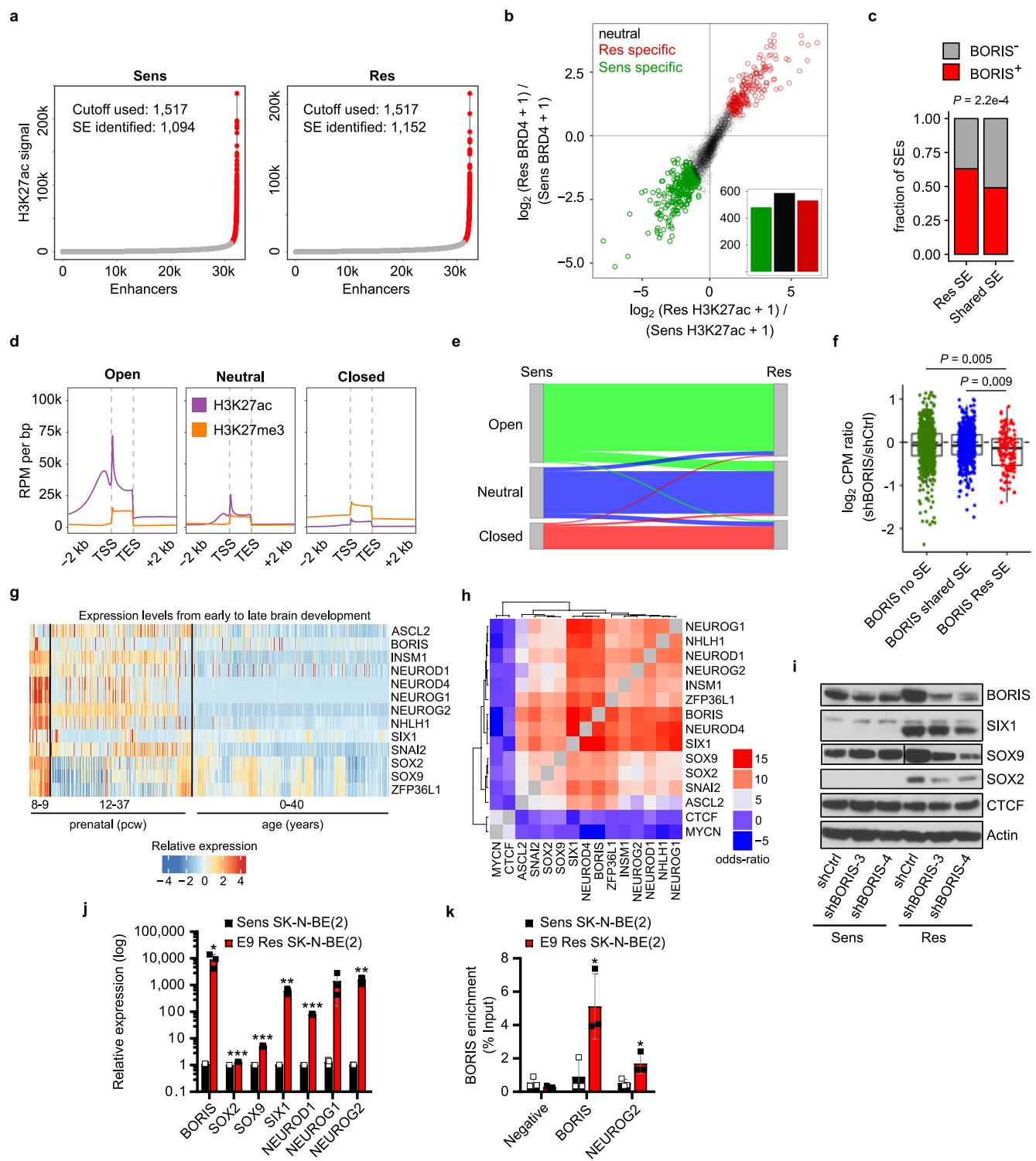
of BORIS-binding peaks in each cell type are given below each pie chart. The regions shown are promoters ( $\text{TSS} \pm 2 \text{ kb}$ ), typical enhancers (H3K27ac), active enhancers (H3K27ac + BRD4), repressed chromatin (H3K27me3), exons, introns, and other (peaks not assigned to any of the previous categories). **e**, Meta-analysis of average CTCF and BORIS ChIP-seq signals in RPM per bp at enhancer and 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 is shown in grey. **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 Fig. 1).



Extended Data Fig. 7 | See next page for caption.

**Extended Data Fig. 7 | Regulatory loops in resistant cells are more vulnerable to BORIS depletion.** **a**, Heat map 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  $\log_2(\text{bp})$ ) 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 top left corner for each track. **e**, Heat map 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**, Bar plots showing the number and

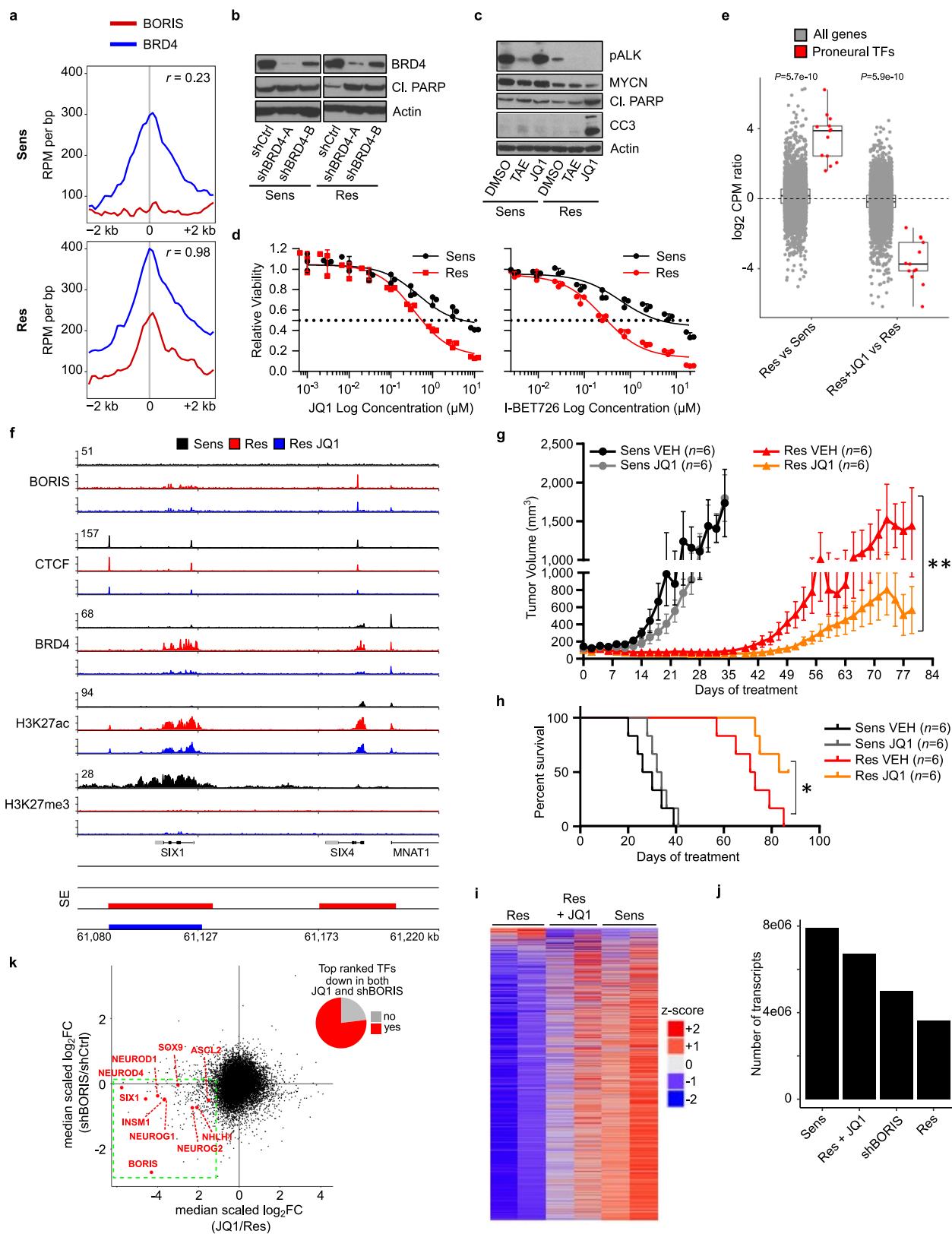
fraction of resistant cell-specific loops for all interaction classes that were BORIS negative and positive in resistant cells, and that were lost after BORIS depletion. **g**, Bar plots showing the odds ratio (two-sided Fisher's exact test) of losing loops that were previously bound by BORIS for all interaction classes. **h**, Box plots showing the initial intensities (in normalized read counts) of BORIS and SMC1A binding in the shRNA control cells at the anchors of the resistant cell-specific loops that were significantly lost versus those that were retained in shBORIS cells (two-sided Wilcoxon rank-sum test). **i**, Box plot showing the difference in SMC1A loss (shBORIS versus shCtrl) on the same anchors as in **h**. **j**, P value determined by two-sided Wilcoxon rank-sum test. All box plots are as defined in Fig. 4. **j**, Metaplots depicting BORIS, SMC1A and CTCF binding at the anchors of the resistant cell-specific loops that were lost or retained after BORIS depletion. In **a–c** and **e–g**, data are from  $n = 3$  biological replicates. In **h–j**, data are from  $n = 2$  biological replicates.



Extended Data Fig. 8 | See next page for caption.

**Extended Data Fig. 8 | Redistribution of the super-enhancer landscape with subsequent expression of a BORIS-dependent proneural network in resistant cells.** **a**, Accumulation of H3K27ac signal at enhancer regions. Typical enhancers (grey) are plotted according to increasing levels of normalized H3K27ac signal (length  $\times$  density) in sensitive and resistant cells. The highest cut-off based on the inclination point in both sensitive and resistant cells was used to delineate super-enhancers (red). **b**, Scatter plot showing differential binding of H3K27ac [ $\log_2(\text{RPM per bp} + 1)$ ] and BRD4 [ $\log_2(\text{RPM per bp} + 1)$ ] for all detected super-enhancers in both sensitive and resistant cells. Cell-specific super-enhancers were identified based on the combined increase in H3K27ac and BRD4 binding. For each individual histone mark, a 0.75  $\log_2$ -transformed fold change threshold was applied and a minimum summed 2.5  $\log_2$ -transformed fold change was used as the final cut-off. **c**, Bar plot depicting the enrichment (two-sided Fisher's exact test) and fractions of resistant cell-specific and shared super-enhancers 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

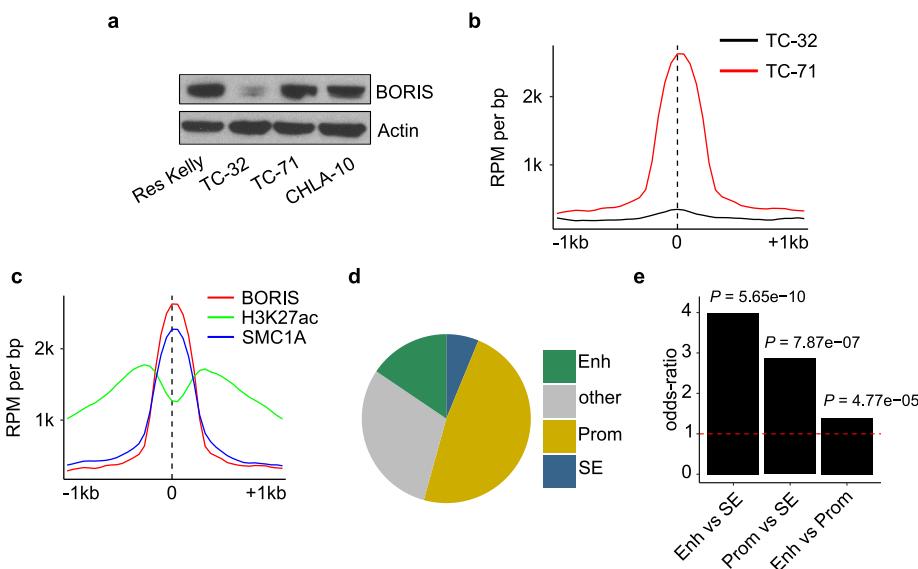
distribution of genes in distinct chromatin states and the switches between sensitive and resistant cells. **f**, Box plots showing the expression level changes upon BORIS depletion for genes that had a resistant cell-specific and BORIS-positive regulatory interaction and were not associated with a super-enhancer ( $n = 720$ ), associated with a super-enhancer in both cell types ( $n = 514$ ) or associated with a super-enhancer seen only in the resistant cells ( $n = 134$ ) (two-sided Wilcoxon rank-sum test). Box plots are as defined in Fig. 4. **g**, Heat map of the expression levels of the indicated proneural transcription factor 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, post-conceptual weeks). **h**, Heat map showing the odds ratios (two-sided Fisher's exact test) for co-detection of the indicated transcription factors 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**, **k**, qRT-PCR analysis of the indicated genes (**j**) and ChIP-qPCR analysis of BORIS binding at the promoter regions of *BORIS* and *NEUROG2* (**k**) in sensitive and resistant SK-N-BE(2) neuroblastoma cells. Data are mean  $\pm$  s.d.,  $n = 3$  biological replicates in **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 Fig. 1).



Extended Data Fig. 9 | See next page for caption.

**Extended Data Fig. 9 | The proneural transcription factor network in resistant cells is sensitive to BRD4 inhibition.** **a**, Metaplots showing the correlation between BRD4 and BORIS co-occupancies at the promoter regions ( $\pm 2$  kb) of the 89 top-ranked genes in resistant versus sensitive cells based on the features in Fig. 4b ( $r$ , Spearman correlation coefficient). **b**, Immunoblot analysis of BRD4 and cleaved PARP expression in sensitive and resistant cells expressing control (shCtrl) or BRD4 (shBRD4-A and -B) shRNAs. **c**, Immunoblot analysis of the indicated proteins in sensitive and resistant cells treated with DMSO, TAE684 (1  $\mu$ M) or JQ1 (2.5  $\mu$ M) for 48 h. **d**, Dose–response curves for sensitive and resistant cells treated with JQ1 or I-BET726 (JQ1 ( $IC_{50}$  values: sensitive, 4,798 nM; resistant, 645 nM); I-BET726 ( $IC_{50}$  values: sensitive, 6,203 nM; resistant, 347 nM)). Data are mean  $\pm$  s.d.,  $n = 3$  biological replicates. **e**, Box plots comparing the expression of the transcription factors listed in Fig. 4b ( $n = 13$ ) with that of all genes ( $n = 18,038$ ) in sensitive versus resistant cells (left), and between DMSO and JQ1-treated resistant cells (right) ( $P$  values determined by two-sided Wilcoxon rank-sum test). **f**, ChIP-seq tracks of the indicated proteins at the *SIX1* or *SIX4* locus in sensitive, resistant and JQ1-treated resistant cells (2.5  $\mu$ M for 48 h). Super-enhancers are depicted as coloured rectangles below the tracks. Signal intensity is shown in the top left corner for each track. **g**, **h**, Tumour volumes (g) and survival curves

(**h**) in sensitive- and resistant-cell xenografts in NU/NU (Crl:NU-*Foxn1<sup>nu</sup>*) mice treated with JQ1 (50 mg kg $^{-1}$  i.p. once daily) and vehicle control for up to 87 days. Data are mean  $\pm$  s.e.m.,  $n = 6$  per arm. Significance was calculated by Mann–Whitney  $U$  test for tumour volumes (sensitive:  $P = 0.3231$ ; resistant:  $P = 0.0023$ ) and by log-rank test for Kaplan–Meier survival analysis (sensitive:  $P = 0.3047$ ; resistant: 0.0348), both two-sided. **i**, Heat map of gene expression values in sensitive, resistant and JQ1-treated resistant cells. Rows are  $z$ -scores calculated for each gene in each condition. **j**, Number of transcripts in sensitive, JQ1-treated resistant, shBORIS-expressing resistant and resistant cells based on expression array data after spike-in normalization. **k**, Scatter plot displaying the median-scaled fold-change gene expression values for shBORIS and JQ1-treated resistant cells. The top-ranked transcription factors that show decreased expression levels after both *BORIS* knockdown and JQ1 treatment are listed in red (bottom left quadrant). The pie chart represents the fraction of all top-ranked transcription factors that are located in the left lower quadrant of the scatter plot. All box plots are as defined in Fig. 4. In **b**, **c** and **f**, data are representative of two independent experiments. In **a**, **e** and **i–k**, data are from  $n = 2$  biological replicates (see Supplementary Note 2 for further details; for gel source data, see Supplementary Fig. 1).



**Extended Data Fig. 10 | Aberrantly expressed BORIS binds to regulatory regions and is associated with new super-enhancers 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 (Kelly) neuroblastoma cells. **b**, Meta-analysis of average BORIS ChIP-seq signals in RPM per 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 per 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 ( $TSS \pm 2$  kb), typical and super-enhancers (H3K27ac), and other (if peaks were not assigned to any of the previous categories). **e**, Bar plot 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 Fig. 1).

Corresponding author(s): Rani E. George

Last updated by author(s): Jun 13, 2019

# Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see [Authors & Referees](#) and the [Editorial Policy Checklist](#).

## Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a Confirmed

- The exact sample size ( $n$ ) for each experimental group/condition, given as a discrete number and unit of measurement
- A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
- The statistical test(s) used AND whether they are one- or two-sided  
*Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
- A description of all covariates tested
- A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
- A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
- For null hypothesis testing, the test statistic (e.g.  $F$ ,  $t$ ,  $r$ ) with confidence intervals, effect sizes, degrees of freedom and  $P$  value noted  
*Give  $P$  values as exact values whenever suitable.*
- For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
- For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
- Estimates of effect sizes (e.g. Cohen's  $d$ , Pearson's  $r$ ), indicating how they were calculated

*Our web collection on [statistics for biologists](#) contains articles on many of the points above.*

## Software and code

Policy information about [availability of computer code](#)

### Data collection

Data collection was done with BD CellQuest Pro software (BD Biosciences) for the flow cytometry experiment. The Cell Ranger Single Cell Software Suite (v1.3) was used to perform sample de-multiplexing, barcode and UMI processing, and singlecell 3' gene counting for the single cell RNA-sequencing experiment.

### Data analysis

R (v3.1.3), affy (v1.44.0), limma (v3.22.7), Fastqc (v0.11.5), STAR (v2.5.1b\_modified), Samtools (v1.3.1), MarkDuplicates (v2.1.1), deepTools (v2.2.4), MACS2 (v2.1.1), bedGraphToBigWig (v4), R (v3.5.1), Rstudio (v1.1.463), data.table (v1.12.2), trimmomatic (v0.36), HiC-Pro (v2.10.0), hicchipper (v0.7.3), diffloop (v1.10.0), ROSE (v1), HOMER (v2) (<http://homer.ucsd.edu/homer/motif/>), Biostrings (v2.50.1), Cell Ranger (v1.3), Scater (v1.10.0), Seurat (v2.3.4), Circlize (v0.4.5), IGV (v2.3.74), FlowJo (v10.0.5) and GraphPad Prism (v7.02). Custom code is available upon reasonable request.

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research [guidelines for submitting code & software](#) for further information.

## Data

Policy information about [availability of data](#)

All manuscripts must include a [data availability statement](#). This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on 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.

# Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

Life sciences       Behavioural & social sciences       Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see [nature.com/documents/nr-reporting-summary-flat.pdf](http://nature.com/documents/nr-reporting-summary-flat.pdf)

## Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	No statistical methods were used to predetermine sample sizes. Sample sizes were chosen in order to be able to perform statistical analyses, as is standard in the field
Data exclusions	No data were excluded from the analyses.
Replication	To verify the reproducibility of our findings, experiments were performed using at least three biological replicates, unless clearly stated otherwise in the figure legends. All attempts at replication were successful.
Randomization	For <i>in vivo</i> mice experiments, mice were randomized into groups of equal average volumes.
Blinding	For experiments involving human research participants, investigators were blinded to group allocation during data collection. For <i>in vivo</i> mice experiments, investigators were not blinded to group allocation during data collection. For experiments involving cell culture, blinding did not apply.

## Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

### Materials & experimental systems

- |                                     |   |
|-------------------------------------|---|
| n/a                                 | Involved in the study   |
| <input type="checkbox"/>            | <input checked="" type="checkbox"/> Antibodies                  |
| <input type="checkbox"/>            | <input checked="" type="checkbox"/> Eukaryotic cell lines       |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Palaeontology                          |
| <input type="checkbox"/>            | <input checked="" type="checkbox"/> Animals and other organisms |
| <input type="checkbox"/>            | <input checked="" type="checkbox"/> Human research participants |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Clinical data                          |

### Methods

- |                                     |  |
|-------------------------------------|--|
| n/a                                 | Involved in the study                              |
| <input type="checkbox"/>            | <input checked="" type="checkbox"/> ChIP-seq       |
| <input type="checkbox"/>            | <input checked="" type="checkbox"/> Flow cytometry |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> MRI-based neuroimaging    |

## Antibodies

### Antibodies used

The following antibodies were used (all commercially available):

- From Cell Signaling Technology (Danvers, MA, USA):
1. N-Myc (#9405, lot 2, 1:1,000);
  2. N-Myc (#51705, Clone D4B2Y, lot 1);
  3. Cleaved PARP Asp214 (#9541, lot 15, 1:4,000);
  4. Cleaved caspase 3 Asp175 (#9661, lot 45, 1:500);
  5. ALK (#3333, Clone C26G7, lot 7, 1:1,000);
  6. Akt (pan) (#4691, Clone C67E7, lot 20, 1:4,000);
  7. Phospho-Akt Thr308 (#9275, lot 26, 1:1,000);
  8. Phospho-Akt Ser473 (#9271, lot 14, 1:2,000);
  9. p44/42 MAPK (ERK1/2) (#4695, Clone 137F5, lot 21, 1:4,000);
  10. Phospho-p44/42 MAPK (ERK1/2) (Thr202/Tyr204) (#4377, Clone 197G2, lot 10, 1:4,000);
  11. S6 Ribosomal Protein (#2217, Clone 5G10, lot 5, 1:8,000);
  12. Phospho-S6 Ribosomal Protein (Ser235/236) (#4857, Clone 91B2, lot 2, 1:8,000);
  13. Stat3 (#4904, Clone 79D7, lot 7, 1:2,000);
  14. Phospho-Stat3 (Tyr705) (#9131, lot 30, 1:1,000);
  15. MDR1/ABCB1 (#12683, Clone D3H1Q, lot 2, 1:5,000);
  16. Sox2 (#3579, Clone D6D9, lot 8, 1:2,000);
  17. β-Actin (#4967, multiple lots, 1:5,000);
  18. CTCF (#3417, Clone D1A7, lot 1);
  19. Normal Rabbit IgG (#2729, lot 7)

20. Anti-mouse IgG, HRP-linked (#7076, lot 33, 1:2,000).
- From Santa Cruz Biotechnology (Santa Cruz, CA, USA):  
 21. Anti-rabbit IgG-HRP (sc-2357, multiple lots, 1:5,000).
- From Bethyl Laboratories (Montgomery, TX, USA):  
 22. BRD4 (A301-985A100, lot 6, 1:30,000);  
 23. SMC1A (A300-055A, lot 6).
- From Millipore (Billerica, MA, USA):  
 24. CTCF (#07-729, lot 2887267, 1:5,000);  
 25. SOX9 (#AB5535, lot 2847051, 1:10,000);  
 26. H3K27me3 (#07-449, lot 2972864).
- From Abcam (Cambridge, MA, USA):  
 27. ALK (phospho Y1507) (ab73996, lot GR57100-14, 1:1,000);  
 28. BORIS (ab187163, Clone EP12204, lot GR228943-6);  
 29. H3K27ac (ab2729, lot GR3198866-1).
- From NOVUS Biologicals (Littleton, CO, USA):  
 30. BORIS (NBP2-52405, Clone 20B11, lot CRT/17/124).
- From Active Motif (Carlsbad, CA, USA):  
 31. BORIS (#39851, lot 18916002, 1:3,000).
- From Sigma-Aldrich (Saint Louis, MO, USA):  
 32. SIX1 (HPA001893, lot I114662, 1:4,000).
- From Abbott (Abbott Park, IL, USA):  
 33. Vysis LSI N-MYC (2p24) SpectrumGreen/Vysis CEP 2 SpectrumOrange Probe (07J72-001).

## Validation

Validation for commercially available antibodies can be found using the links below:

1. <https://www.cellsignal.com/products/primary-antibodies/n-myc-antibody/9405>
2. <https://www.cellsignal.com/products/primary-antibodies/n-myc-d4b2y-rabbit-mab/51705>
3. <https://www.cellsignal.com/products/primary-antibodies/cleaved-parp-asp214-antibody-human-specific/9541>
4. <https://www.cellsignal.com/products/primary-antibodies/cleaved-caspase-3-asp175-antibody/9661>
5. <https://www.cellsignal.com/products/primary-antibodies/alk-c26g7-rabbit-mab/3333>
6. <https://www.cellsignal.com/products/primary-antibodies/akt-pan-c67e7-rabbit-mab/4691>
7. <https://www.cellsignal.com/products/primary-antibodies/phospho-akt-thr308-antibody/9275>
8. <https://www.cellsignal.com/products/primary-antibodies/phospho-akt-ser473-antibody/9271>
9. <https://www.cellsignal.com/products/primary-antibodies/p44-42-mapk-erk1-2-137f5-rabbit-mab/4695>
10. <https://www.cellsignal.com/products/primary-antibodies/phospho-p44-42-mapk-erk1-2-thr202-tyr204-197g2-rabbit-mab/4377>
11. <https://www.cellsignal.com/products/primary-antibodies/s6-ribosomal-protein-5g10-rabbit-mab/2217>
12. <https://www.cellsignal.com/products/primary-antibodies/phospho-s6-ribosomal-protein-ser235-236-91b2-rabbit-mab/4857>
13. <https://www.cellsignal.com/products/primary-antibodies/stat3-79d7-rabbit-mab/4904>
14. <https://www.cellsignal.com/products/primary-antibodies/phospho-stat3-tyr705-antibody/9131>
15. <https://www.cellsignal.com/products/primary-antibodies/mdr1-abcb1-d3h1q-rabbit-mab/12683>
16. <https://www.cellsignal.com/products/primary-antibodies/sox2-d6d9-xp-rabbit-mab/3579>
17. <https://www.cellsignal.com/products/primary-antibodies/b-actin-antibody/4967>
18. <https://www.cellsignal.com/products/primary-antibodies/ctcf-d1a7-xp-rabbit-mab/3417>
19. <https://www.cellsignal.com/products/primary-antibodies/normal-rabbit-igg/2729>
20. <https://www.cellsignal.com/products/secondary-antibodies/anti-mouse-igg-hrp-linked-antibody/7076>
21. <https://www.scbt.com/scbt/product/mouse-anti-rabbit-igg-hrp>
22. <https://www.bethyl.com/product/A301-985A100/BRD4+Antibody>
23. <https://www.bethyl.com/product/A300-055A/SMC1+Antibody>
24. [http://www.emdmillipore.com/US/en/product/Anti-CTCF-Antibody,MM\\_NF-07-729](http://www.emdmillipore.com/US/en/product/Anti-CTCF-Antibody,MM_NF-07-729)
25. [http://www.emdmillipore.com/US/en/product/Anti-Sox9-Antibody,MM\\_NF-AB5535](http://www.emdmillipore.com/US/en/product/Anti-Sox9-Antibody,MM_NF-AB5535)
26. [http://www.emdmillipore.com/US/en/product/Anti-trimethyl-Histone-H3-Lys27-Antibody,MM\\_NF-07-449](http://www.emdmillipore.com/US/en/product/Anti-trimethyl-Histone-H3-Lys27-Antibody,MM_NF-07-449)
27. <https://www.abcam.com/alk-phospho-y1507-antibody-ab73996.html>
28. <https://www.abcam.com/boris-antibody-ep12204-ab187163.html>
29. <https://www.abcam.com/histone-h3-acetyl-k27-antibody-chip-grade-ab4729.html>
30. [https://www.novusbio.com/products/boris-antibody-20b11\\_nb2-52405](https://www.novusbio.com/products/boris-antibody-20b11_nb2-52405)
31. <https://www.activemotif.com/catalog/details/39851/boris-ctcf-antibody-pab>
32. <https://www.sigmal Aldrich.com/catalog/product/sigma/hpa001893?lang=en&region=US>
33. <https://www.molecular.abbott/sal/en-us/staticAssets/AMD-US-Oncology-and-Genetics-Catalog.pdf>

## Eukaryotic cell lines

Policy information about [cell lines](#)

Cell line source(s)	Human NB cell lines Kelly and SK-N-BE(2) and human Ewing sarcoma cell lines TC-32, TC-71 and CHLA-10 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).
Authentication	All cell lines have been authenticated by short tandem repeat analysis.
Mycoplasma contamination	All cell lines tested negative for mycoplasma.
Commonly misidentified lines (See <a href="#">ICLAC</a> register)	No commonly misidentified cell lines were used.

## Animals and other organisms

Policy information about [studies involving animals; ARRIVE guidelines](#) recommended for reporting animal research

Laboratory animals	NU/NU (Crl:NU-Foxn1nu) mice (6-8 weeks old female) were purchased from Charles River Laboratories. NU/NU (CrTac:NCr-Foxn1nu) mice (6-8 weeks old female) were purchased from Taconic.
Wild animals	The study did not involve wild animals.
Field-collected samples	The study did not involve samples collected from the field.
Ethics oversight	All animal experiments were performed with approval from the Institutional Animal Care and Use Committee of the DFCI.

Note that full information on the approval of the study protocol must also be provided in the manuscript.

## Human research participants

Policy information about [studies involving human research participants](#)

Population characteristics	Data involving human research participants pertain to immunohistochemistry of formalin-fixed paraffin-embedded tumor tissue slides, and were de-identified prior to analysis.
Recruitment	Data involving human research participants pertain to immunohistochemistry of formalin-fixed paraffin-embedded tumor tissue slides, and were de-identified prior to analysis.
Ethics oversight	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.

Note that full information on the approval of the study protocol must also be provided in the manuscript.

## ChIP-seq

### Data deposition

- Confirm that both raw and final processed data have been deposited in a public database such as [GEO](#).
- Confirm that you have deposited or provided access to graph files (e.g. BED files) for the called peaks.

Data access links	<a href="https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE103084">https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE103084</a> <small>May remain private before publication.</small>
Files in database submission	<p>ChIPSeq_Original_Submission_sens_res Raw files:</p> <p>Par_BORIS_1.fastq.gz Par_BORIS_2.fastq.gz Par_BRD4_1.fastq.gz Par_BRD4_2.fastq.gz Par_CTCF_1.fastq.gz Par_CTCF_2.fastq.gz Par_H3K27ac_1.fastq.gz Par_H3K27ac_2.fastq.gz Par_H3K27me3_1.fastq.gz Par_H3K27me3_2.fastq.gz Par_input_1.fastq.gz Par_input_2.fastq.gz Par_Pol2_1.fastq.gz Par_Pol2_2.fastq.gz</p>

Res\_BORIS\_1.fastq.gz  
Res\_BORIS\_2.fastq.gz  
Res\_BRD4\_1.fastq.gz  
Res\_BRD4\_2.fastq.gz  
Res\_CTCF\_1.fastq.gz  
Res\_CTCF\_2.fastq.gz  
Res\_H3K27ac\_1.fastq.gz  
Res\_H3K27ac\_2.fastq.gz  
Res\_H3K27me3\_1.fastq.gz  
Res\_H3K27me3\_2.fastq.gz  
Res\_input\_1.fastq.gz  
Res\_input\_2.fastq.gz  
Res\_Pol2\_1.fastq.gz  
Res\_Pol2\_2.fastq.gz  
Res\_JQ1\_BORIS\_1.fastq.gz  
Res\_JQ1\_BORIS\_2.fastq.gz  
Res\_JQ1\_BRD4\_1.fastq.gz  
Res\_JQ1\_BRD4\_2.fastq.gz  
Res\_JQ1\_CTCF\_1.fastq.gz  
Hi2.fastq.gz  
Res\_JQ1\_H3K27ac\_1.fastq.gz  
Res\_JQ1\_H3K27ac\_2.fastq.gz  
Res\_JQ1\_H3K27me3\_1.fastq.gz  
Res\_JQ1\_H3K27me3\_2.fastq.gz  
Res\_JQ1\_input\_1.fastq.gz  
Res\_JQ1\_input\_2.fastq.gz  
Processed files:  
norm\_Par\_BORIS\_merged.noNeg.sorted.bw  
norm\_Par\_BRD4\_merged.noNeg.sorted.bw  
norm\_Par\_CTCF\_merged.noNeg.sorted.bw  
norm\_Par\_H3K27ac\_merged.noNeg.sorted.bw  
norm\_Par\_H3K27me3\_merged.noNeg.sorted.bw  
norm\_Par\_Pol2\_2.noNeg.sorted.bw  
norm\_Res\_BORIS\_merged.noNeg.sorted.bw  
norm\_Res\_BRD4\_merged.noNeg.sorted.bw  
norm\_Res\_CTCF\_merged.noNeg.sorted.bw  
norm\_Res\_H3K27ac\_merged.noNeg.sorted.bw  
norm\_Res\_H3K27me3\_merged.noNeg.sorted.bw  
norm\_Res\_Pol2\_2.noNeg.sorted.bw  
norm\_Res\_JQ1\_BORIS\_merged.noNeg.sorted.bw  
norm\_Res\_JQ1\_BRD4\_merged.noNeg.sorted.bw  
norm\_Res\_JQ1\_CTCF\_merged.noNeg.sorted.bw  
norm\_Res\_JQ1\_H3K27ac\_merged.noNeg.sorted.bw  
norm\_Res\_JQ1\_H3K27me3\_merged.noNeg.sorted.bw  
Par\_BORIS\_vs\_Par\_input\_peaks.narrowPeak  
Par\_BRD4\_vs\_Par\_input\_peaks.narrowPeak  
Par\_CTCF\_vs\_Par\_input\_peaks.narrowPeak  
Par\_narrow\_Pol2\_vs\_Par\_input\_peaks.narrowPeak  
Res\_BORIS\_vs\_Res\_input\_peaks.narrowPeak  
Res\_BRD4\_vs\_Par\_input\_peaks.narrowPeak  
Res\_CTCF\_vs\_Res\_input\_peaks.narrowPeak  
Res\_narrow\_Pol2\_vs\_Res\_input\_peaks.narrowPeak  
Res\_JQ1\_BORIS\_vs\_Res\_JQ1\_input\_peaks.narrowPeak  
Res\_JQ1\_BRD4\_vs\_Res\_JQ1\_input\_peaks.narrowPeak  
Res\_JQ1\_CTCF\_vs\_Res\_JQ1\_input\_peaks.narrowPeak  
Par\_H3K27ac\_vs\_Par\_input\_peaks.broadPeak  
Par\_H3K27me3\_vs\_Par\_input\_peaks.broadPeak  
Res\_H3K27ac\_vs\_Res\_input\_peaks.broadPeak  
Res\_H3K27me3\_vs\_Res\_input\_peaks.broadPeak  
Res\_JQ1\_H3K27ac\_vs\_Res\_JQ1\_input\_peaks.broadPeak  
Res\_JQ1\_H3K27me3\_vs\_Res\_JQ1\_input\_peaks.broadPeak  
ChIPseq\_MYCN  
Raw files:  
Kelly\_input\_MYCN\_1.fastq.gz  
Kelly\_MYCN\_CS\_1.fastq.gz  
Kelly\_MYCN\_CS\_2.fastq.gz  
Res\_Kelly\_input\_MYCN\_CS\_1.fastq.gz  
Res\_Kelly\_input\_MYCN\_CS\_2.fastq.gz  
Res\_Kelly\_MYCN\_CS\_1.fastq.gz  
Res\_Kelly\_MYCN\_CS\_2.fastq.gz  
Processed files:  
Par\_Kelly\_MYCN.noNeg.sorted.bw  
Res\_Kelly\_MYCN.noNeg.sorted.bw  
Par\_Kelly\_MYCN\_vs\_Par\_Kelly\_MYCN\_INPUT\_peaks.narrowPeak  
Res\_Kelly\_MYCN\_vs\_Res\_Kelly\_MYCN\_INPUT\_peaks.narrowPeak

## ChIPseq\_KD\_Boris

Raw files:

shB4\_Res\_Kelly\_BORIS\_1.fastq.gz  
shB4\_Res\_Kelly\_BORIS\_2.fastq.gz  
shB4\_Res\_Kelly\_CTCF\_1.fastq.gz  
shB4\_Res\_Kelly\_CTCF\_2.fastq.gz  
shB4\_Res\_Kelly\_H3K27ac\_1.fastq.gz  
shB4\_Res\_Kelly\_H3K27ac\_2.fastq.gz  
shB4\_Res\_Kelly\_INPUT\_1.fastq.gz  
shB4\_Res\_Kelly\_INPUT\_2.fastq.gz  
shB4\_Res\_Kelly\_INPUT\_BORIS\_1.fastq.gz  
shB4\_Res\_Kelly\_INPUT\_BORIS\_2.fastq.gz  
shB4\_Res\_Kelly\_SMC1\_1.fastq.gz  
shB4\_Res\_Kelly\_SMC1\_2.fastq.gz  
shLUC\_Res\_Kelly\_BORIS\_1.fastq.gz  
shLUC\_Res\_Kelly\_BORIS\_2.fastq.gz  
shLUC\_Res\_Kelly\_CTCF\_1.fastq.gz  
shLUC\_Res\_Kelly\_CTCF\_2.fastq.gz  
shLUC\_Res\_Kelly\_H3K27ac\_1.fastq.gz  
shLUC\_Res\_Kelly\_H3K27ac\_2.fastq.gz  
shLUC\_Res\_Kelly\_INPUT\_1.fastq.gz  
shLUC\_Res\_Kelly\_INPUT\_2.fastq.gz  
shLUC\_Res\_Kelly\_SMC1\_1.fastq.gz  
shLUC\_Res\_Kelly\_SMC1\_2.fastq.gz

Processed files:

shB4\_Res\_Kelly\_BORIS.noNeg.sorted.bw  
shB4\_Res\_Kelly\_CTCF.noNeg.sorted.bw  
shB4\_Res\_Kelly\_H3K27ac.noNeg.sorted.bw  
shB4\_Res\_Kelly\_SMC1.noNeg.sorted.bw  
shLUC\_Res\_Kelly\_BORIS.noNeg.sorted.bw  
shLUC\_Res\_Kelly\_CTCF.noNeg.sorted.bw  
shLUC\_Res\_Kelly\_H3K27ac.noNeg.sorted.bw  
shLUC\_Res\_Kelly\_SMC1.noNeg.sorted.bw  
shB4\_Res\_Kelly\_BORIS\_vs\_shB4\_Res\_Kelly\_INPUT\_peaks.narrowPeak  
shB4\_Res\_Kelly\_CTCF\_vs\_shB4\_Res\_Kelly\_INPUT\_peaks.narrowPeak  
shB4\_Res\_Kelly\_H3K27ac\_vs\_shB4\_Res\_Kelly\_INPUT\_peaks.broadPeak  
shB4\_Res\_Kelly\_SMC1\_vs\_shB4\_Res\_Kelly\_INPUT\_peaks.narrowPeak  
shLUC\_Res\_Kelly\_BORIS\_vs\_shLUC\_Res\_Kelly\_INPUT\_peaks.narrowPeak  
shLUC\_Res\_Kelly\_CTCF\_vs\_shLUC\_Res\_Kelly\_INPUT\_peaks.narrowPeak  
shLUC\_Res\_Kelly\_H3K27ac\_vs\_shLUC\_Res\_Kelly\_INPUT\_peaks.broadPeak  
shLUC\_Res\_Kelly\_SMC1\_vs\_shLUC\_Res\_Kelly\_INPUT\_peaks.narrowPeak

## ChIPseq\_Ewing\_Sarcoma

Raw files:

TC32\_BORIS\_1.fastq.gz  
TC32\_BORIS\_2.fastq.gz  
TC32\_H3K27ac\_1.fastq.gz  
TC32\_H3K27ac\_2.fastq.gz  
TC32\_input\_1.fastq.gz  
TC32\_input\_2.fastq.gz  
TC32\_SMC1\_1.fastq.gz  
TC32\_SMC1\_2.fastq.gz  
TC71\_BORIS\_1.fastq.gz  
TC71\_BORIS\_2.fastq.gz  
TC71\_H3K27ac\_1.fastq.gz  
TC71\_H3K27ac\_2.fastq.gz  
TC71\_input\_1.fastq.gz  
TC71\_input\_2.fastq.gz  
TC71\_SMC1\_1.fastq.gz  
TC71\_SMC1\_2.fastq.gz

Processed files:

Ewings\_TC32\_BORIS.noNeg.sorted.bw  
Ewings\_TC32\_H3K27ac.noNeg.sorted.bw  
Ewings\_TC32\_SMC1.noNeg.sorted.bw  
Ewings\_TC71\_BORIS.noNeg.sorted.bw  
Ewings\_TC71\_H3K27ac.noNeg.sorted.bw  
Ewings\_TC71\_SMC1.noNeg.sorted.bw  
Ewings\_TC32\_BORIS\_vs\_David\_Ewings\_TC32\_INPUT\_peaks.narrowPeak  
Ewings\_TC32\_H3K27ac\_vs\_David\_Ewings\_TC32\_INPUT\_peaks.broadPeak  
Ewings\_TC32\_SMC1\_vs\_David\_Ewings\_TC32\_INPUT\_peaks.narrowPeak  
Ewings\_TC71\_BORIS\_vs\_David\_Ewings\_TC71\_INPUT\_peaks.narrowPeak  
Ewings\_TC71\_H3K27ac\_vs\_David\_Ewings\_TC71\_INPUT\_peaks.broadPeak  
Ewings\_TC71\_SMC1\_vs\_David\_Ewings\_TC71\_INPUT\_peaks.narrowPeak  
Microarray\_Original\_Submission\_sens\_res\_JQ1

Raw files:

DD\_1\_PrimeView\_par\_Kelly\_DMSO.CEL

DD\_2\_PrimeView\_par\_Kelly\_DMSO.CEL  
 DD\_5\_PrimeView\_Res\_Kelly\_TAE.CEL  
 DD\_6\_PrimeView\_Res\_Kelly\_TAE.CEL  
 DD\_7\_PrimeView\_Res\_Kelly\_JQ1.CEL  
 DD\_8\_PrimeView\_Res\_Kelly\_JQ1.CEL  
 Microarray\_KD\_Boris  
 Raw files:  
 Res\_Kelly\_shBORIS\_1.CEL  
 Res\_Kelly\_shBORIS\_2.CEL  
 Res\_Kelly\_shLUC\_1.CEL  
 Res\_Kelly\_shLUC\_2.CEL  
 10X\_sens\_IR\_Res  
 Raw files:  
 Sens.possorted\_genome\_bam.bam  
 Intermediate\_Res.possorted\_genome\_bam.bam  
 Full\_Res.possorted\_genome\_bam.bam  
 Processed files:  
 Sens\_matrix mtx  
 Intermediate\_Res\_matrix.mtx  
 Full\_Res\_matrix.mtx  
 barcodes.tsv  
 genes.tsv  
 HiChIP\_sens\_res  
 Raw files:  
 KellyRes1\_1.fastq.tar.gz  
 KellyRes1\_2.fastq.tar.gz  
 KellyRes2\_1.fastq.tar.gz  
 KellyRes2\_2.fastq.tar.gz  
 KellyRes3\_1.fastq.tar.gz  
 KellyRes3\_2.fastq.tar.gz  
 KellyWTrep1\_1.fastq.bz2  
 KellyWTrep1\_2.fastq.bz2  
 KellyWTrep2\_1.fastq.bz2  
 KellyWTrep2\_2.fastq.bz2  
 KellyWTrep3\_1.fastq.bz2  
 KellyWTrep3\_2.fastq.bz2  
 Processed files:  
 Kelly-Res-SMC1.bedpe  
 Kelly-Res-SMC1-petcount.bedpe  
 Kelly-SMC1.bedpe  
 Kelly-SMC1-petcount.bedpe  
 HiChIP\_KD\_Boris  
 Raw files:  
 Kelly\_shBORIS1\_1.fastq.tar.gz  
 Kelly\_shBORIS1\_2.fastq.tar.gz  
 Kelly\_shBORIS2\_1.fastq.tar.gz  
 Kelly\_shBORIS2\_2.fastq.tar.gz  
 Kelly\_shBORIS3\_1.fastq.tar.gz  
 Kelly\_shBORIS3\_2.fastq.tar.gz  
 Kelly\_shLUC1\_1.fastq.tar.gz  
 Kelly\_shLUC1\_2.fastq.tar.gz  
 Kelly\_shLUC2\_1.fastq.tar.gz  
 Kelly\_shLUC2\_2.fastq.tar.gz  
 Kelly\_shLUC3\_1.fastq.tar.gz  
 Kelly\_shLUC3\_2.fastq.tar.gz  
 Processed files:  
 shBORIS-results.csv  
 shLUC-results.csv

Genome browser session  
 (e.g. [UCSC](#))

No longer applicable.

## Methodology

### Replicates

For each mark assessed, two biological replicates were performed. All ChIP-seq data are derived from the analysis of both replicates.

### Sequencing depth

ChIP-seq libraries were generated using the NEBNext Ultra II DNA Library Prep kit (E7645), following the manufacturer's instructions. Starting DNA material ranged from 2.5 to 10ng, and PCR amplification (8 to 10 cycles) was performed using NEBNext Multiplex Oligos for Illumina (E7335 and E7500) with distinct indices to allow for multiplexing of up to 12 samples to be run on the Illumina NextSeq 500 for 75 bases in single-read mode.

Sample Total\_reads Uniquely\_mapped\_reads  
 Par\_BORIS\_1.fastq.gz 52592735 39434783  
 Par\_BORIS\_2.fastq.gz 54031826 40040869

```

Par_BRD4_1.fastq.gz 32913499 24145172
Par_BRD4_2.fastq.gz 34899996 25050068
Par_CTCF_1.fastq.gz 32357383 21742074
Par_CTCF_2.fastq.gz 38048050 25570483
Par_H3K27ac_1.fastq.gz 34652785 31481286
Par_H3K27ac_2.fastq.gz 36652655 33723955
Par_H3K27me3_1.fastq.gz 35375006 28384362
Par_H3K27me3_2.fastq.gz 33791944 27655558
Par_Pol2_1.fastq.gz 24373861 6139176
Par_Pol2_2.fastq.gz 41084548 2685056
Kelly_MYCN_CS_1.fastq.gz 22544426 19447997
Kelly_MYCN_CS_2.fastq.gz 43975115 38012488
Par_input_1.fastq.gz 74638885 66630061
Par_input_2.fastq.gz 40772400 36361135
Kelly_input_MYCN_1.fastq.gz 43184072 34957425
Res_BORIS_1.fastq.gz 43137736 32555678
Res_BORIS_2.fastq.gz 47143635 35144692
Res_BRD4_1.fastq.gz 32159271 24332339
Res_BRD4_2.fastq.gz 36755542 30051197
Res_CTCF_1.fastq.gz 40911369 28372298
Res_CTCF_2.fastq.gz 36878692 25628082
Res_H3K27ac_1.fastq.gz 33542051 30727706
Res_H3K27ac_2.fastq.gz 37596270 34409319
Res_H3K27me3_1.fastq.gz 35238601 29098617
Res_H3K27me3_2.fastq.gz 35026938 28826263
Res_Pol2_1.fastq.gz 35817388 6953374
Res_Pol2_2.fastq.gz 48429363 2674818
Res_Kelly_MYCN_CS_1.fastq.gz 18821385 14198653
Res_Kelly_MYCN_CS_2.fastq.gz 30543884 24497611
Res_Kelly_input_MYCN_CS_1.fastq.gz 77322913 68513199
Res_Kelly_input_MYCN_CS_2.fastq.gz 53883106 47951287
Res_input_1.fastq.gz 45072526 40107804
Res_input_2.fastq.gz 49535965 44608902
Res_JQ1_BORIS_1.fastq.gz 81479030 61882578
Res_JQ1_BORIS_2.fastq.gz 64057634 49039377
Res_JQ1_BRD4_1.fastq.gz 32714245 25430455
Res_JQ1_BRD4_2.fastq.gz 33031805 25335103
Res_JQ1_CTCF_1.fastq.gz 33287924 24508642
Res_JQ1_CTCF_2.fastq.gz 39674176 29154183
Res_JQ1_H3K27ac_1.fastq.gz 40677082 37149902
Res_JQ1_H3K27ac_2.fastq.gz 32078957 29293906
Res_JQ1_H3K27me3_1.fastq.gz 36342163 30007069
Res_JQ1_H3K27me3_2.fastq.gz 37930668 31338450
Res_JQ1_input_1.fastq.gz 49313027 43993392
Res_JQ1_input_2.fastq.gz 43284562 38433644

```

## Antibodies

The following antibodies were used: MYCN (#51705, Cell Signaling Technology), BRD4 (A301-985A100, lot 6, Bethyl Laboratories), CTCF (#07-729, lot 2887267, Millipore), H3K27me3 (#07-449, lot 2972864, Millipore), H3K27ac (ab2729, lot GR3198866-1, Abcam), and BORIS (#39851, lot 18916002, Active Motif). All antibodies were previously validated by their manufacturers.

## Peak calling parameters

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. Peaks were identified with MACS2 (2.1.1) for narrow peaks with the parameters “--q 0.01--call-summits” and for broad peaks with the parameters “--broad-cutoff 0.01”.

## Data quality

- 1) Peaks overlapping regions with known artefact regions (<http://mitra.stanford.edu/kundaje/akundaje/release/blacklists/>) were blacklisted out.
- 2) Antibody enrichment was assessed using the plotFingerprint command from deepTools (v2.2.4).
- 3) Correlation of replicates was assessed with the deepTools command “multiBigwigSummary BED-file” using all bigwigs and identified peaks as input.
- 4) broad peaks with 5-fold enrichment, 5% FDR and after blacklisting
  - Par\_H3K27ac\_vs\_Par\_input\_peaks.broadPeak.blacklisted 15626
  - Par\_H3K27me3\_vs\_Par\_input\_peaks.broadPeak.blacklisted 3946
  - Res\_H3K27ac\_vs\_Res\_input\_peaks.broadPeak.blacklisted 16725
  - Res\_H3K27me3\_vs\_Res\_input\_peaks.broadPeak.blacklisted 2293
  - Res\_JQ1\_H3K27ac\_vs\_Res\_JQ1\_input\_peaks.broadPeak.blacklisted 15942
  - Res\_JQ1\_H3K27me3\_vs\_Res\_JQ1\_input\_peaks.broadPeak.blacklisted 1819
- 5) narrow peaks with 5-fold enrichment, 5% FDR and after blacklisting
  - Par\_BORIS\_vs\_Par\_input\_peaks\_blacklisted.narrowPeak 1055
  - Par\_BRD4\_vs\_Par\_input\_peaks\_blacklisted.narrowPeak 52158
  - Par\_CTCF\_vs\_Par\_input\_peaks\_blacklisted.narrowPeak 69566
  - Par\_narrow\_Pol2\_vs\_Par\_input\_peaks\_blacklisted.narrowPeak 19193
  - Par\_MYCN\_vs\_Par\_input\_peaks\_blacklisted.narrowPeak 36047

Res\_BORIS\_vs\_Res\_input\_peaks\_blacklisted.narrowPeak 17035  
 Res\_BRD4\_vs\_Res\_input\_peaks\_blacklisted.narrowPeak 26297  
 Res\_CTCF\_vs\_Res\_input\_peaks\_blacklisted.narrowPeak 58724  
 Res\_JQ1\_BORIS\_vs\_Res\_JQ1\_input\_peaks\_blacklisted.narrowPeak 12674  
 Res\_JQ1\_BRD4\_vs\_Res\_JQ1\_input\_peaks\_blacklisted.narrowPeak 9702  
 Res\_JQ1\_CTCF\_vs\_Res\_JQ1\_input\_peaks\_blacklisted.narrowPeak 37791  
 Res\_narrow\_Pol2\_vs\_Res\_input\_peaks\_blacklisted.narrowPeak 54174  
 Res\_MYCN\_vs\_Res\_input\_peaks\_blacklisted.narrowPeak 6500

## Software

Fastqc (v0.11.5), STAR (v2.5.1b\_modified), Samtools (v1.3.1), MarkDuplicates (v2.1.1), deepTools (v2.2.4), MACS2 (v2.1.1), bedGraphToBigWig (v4), R (v3.5.1), Rstudio (v1.1.463), data.table (v1.12.2), trimmomatic (v0.36), HiC-Pro (v2.10.0), hichipper (v0.7.3), diffloop (v1.10.0), ROSE (v1), Circlize (v0.4.5), IGV (v2.3.74). Custom code is available upon reasonable request.

## Flow Cytometry

### Plots

Confirm that:

- The axis labels state the marker and fluorochrome used (e.g. CD4-FITC).
- The axis scales are clearly visible. Include numbers along axes only for bottom left plot of group (a 'group' is an analysis of identical markers).
- All plots are contour plots with outliers or pseudocolor plots.
- A numerical value for number of cells or percentage (with statistics) is provided.

### Methodology

## Sample preparation

Flow cytometry was used for cell cycle analysis in both parental and resistant cells using propidium iodide (PI) staining. After fixation of 1x10<sup>6</sup> cells overnight at 4°C with 80% ethanol, cells were resuspended in PBS supplemented with 0.1% Triton X-100, 25 mg/ml PI, and 0.2 mg/ml RNase A, and incubated for 45 min at 37°C in the dark before analysis.

## Instrument

Samples were run on a FACSCalibur flow cytometer (BD Biosciences).

## Software

Data collection was done with BD CellQuest Pro software (BD Biosciences), and analysis was performed using FlowJo software (v10.0.5).

## Cell population abundance

No FACS sorting was performed for this work.

## Gating strategy

Cell debris as well as non-singlets were excluded by a gate based on FSC-H and SSC-H.

- Tick this box to confirm that a figure exemplifying the gating strategy is provided in the Supplementary Information.