Enhancer-promoter hubs organize transcriptional networks promoting oncogenesis and drug resistance

Brent S. Perlman [1,2,3,7], Noah Burget[1,2,3,7], Yeqiao Zhou[1,2,3], Gregory W. Schwartz [4,5], Jelena Petrovic[6], Zora Modrusan[6] & Robert B. Faryabi [1,2,3] ✉

Recent advances in high-resolution mapping of spatial interactions among regulatory elements support the existence of complex topological assemblies of enhancers and promoters known as enhancer-promoter hubs or cliques. Yet, organization principles of these multi-interacting enhancer-promoter hubs and their potential role in regulating gene expression in cancer remain unclear. Here, we systematically identify enhancer-promoter hubs in breast cancer, lymphoma, and leukemia. We find that highly interacting enhancer-promoter hubs form at key oncogenes and lineage-associated transcription factors potentially promoting oncogenesis of these diverse cancer types. Genomic and optical mapping of interactions among enhancer and promoter elements further show that topological alterations in hubs coincide with transcriptional changes underlying acquired resistance to targeted therapy in T cell leukemia and B cell lymphoma. Together, our findings suggest that enhancer-promoter hubs are dynamic and heterogeneous topological assemblies with the potential to control gene expression circuits promoting oncogenesis and drug resistance.

Genome spatial organization facilitates enhancer-promoter communication, which is crucial for control of oncogenic transcriptional programs[1,2]. Emerging evidence from studies of cancer genome topology supports that multiple enhancers and promoters can spatially coalesce, forming topological assemblies that are variably referred to as enhancer-promoter hubs or cliques[3,4]. Nevertheless, the fundamental properties of these topological assemblies and their potential role in promoting oncogenesis remain unclear.

Investigation of oncogenic enhancer-promoter hubs has unique potential to advance our understanding of cancer given that enhancer dysregulation is a key hallmark of oncogenesis[5]. Furthermore, current models have yet to fully grasp how distal enhancers exert their regulatory functions across large genomic distances. Genome topology, which is organized at various length scales from megabase-scale compartments and topologically associating domains (TADs) to fine-scale chromatin loops, contributes to spatial positioning of enhancers and their target promoters, influencing their activity and specificity[6–9]. Given that the number of active enhancers is 2–3 times more than active genes[10], it is often possible that multiple enhancers control the expression of a single gene, giving rise to complex enhancer regulatory circuits[11–13]. Although chromatin interaction data alone cannot capture the complexity of potential multi-enhancer regulation, its integration with chromatin activity datasets at a few oncogenes revealed that multiple distal enhancers can spatially cluster with promoters to form

[1]Department of Pathology and Laboratory Medicine, University of Pennsylvania, Philadelphia, USA. [2]Penn Epigenetics Institute, University of Pennsylvania, Philadelphia, USA. [3]Abramson Family Cancer Research Institute, University of Pennsylvania, Philadelphia, USA. [4]Princess Margaret Cancer Centre, University Health Network, Toronto, Canada. [5]Department of Medical Biophysics, University of Toronto, Toronto, Canada. [6]Department of Proteomic and Genomic Technologies, Genentech, South San Francisco, USA. [7]These authors contributed equally: Brent S. Perlman, Noah Burget. ✉e-mail: faryabi@pennmedicine.upenn.edu

enhancer-promoter hubs in cancer genomes[3,14–18]. More recent studies have demonstrated that enhancer-promoter hubs facilitate enhancer cooperativity and target specificity to control gene expression dosage[12,14,19]. Despite these advances, a systematic understanding of enhancer-promoter hub prevalence, organization principles, and regulatory importance in mediating oncogenic enhancer function is lacking.

In this work, we systematically identify enhancer-promoter hubs in T cell acute lymphoblastic leukemia (T-ALL), mantle cell lymphoma (MCL), and triple negative breast cancer (TNBC) to elucidate prevalence and organization principles of these topological assemblies across diverse cancer types. Examination of enhancer-promoter hubs reveals that they are ubiquitous and different from TADs and super-enhancers. Study of T-ALL, MCL and TNBC enhancer-promoter interactions further shows that hubs are heterogeneous with asymmetric distribution of interactions among enhancers and promoters. Notably, a small subset of enhancer-promoter hubs is hyperinteracting, exhibiting exceptionally high spatial interactivity among constituent enhancer and promoter elements. We demonstrate that hyperinteracting hubs are uniquely enriched for transcription, predominantly form around transcription factors and coregulators, and are more lineage associated than regular (i.e. non-hyperinteracting) hubs. To further substantiate the structure-function relationship of enhancer-promoter hubs, we examine their reorganization in Notch inhibitor resistant T-ALL and Bruton's tyrosine kinase (BTK) inhibitor resistant MCL. Our population-based and single-cell resolution chromatin mapping studies reveal the role of enhancer-promoter hub reorganization in setting gene expression programs permissive to Notch inhibitor and BTK inhibitor resistance in T-ALL and MCL, respectively. Together, our data suggest that enhancer-promoter hub formation is an epigenetic mechanism which is potentially hijacked by cancer cells to set gene expression programs promoting oncogenesis and drug resistance.

Results

Interactions among enhancers and promoters are asymmetrically distributed in T leukemic cells

Complex interactions among enhancers and promoters measured by high-resolution chromatin conformation capture assays such as in-situ Hi-C or HiChIP can be conceptualized as a network of connected nodes within nuclear space and modeled using undirected graph mathematical abstraction[14,20]. To detect groups of highly interacting enhancers and promoters, known as enhancer-promoter hubs or cliques, from the graph of frequently interacting enhancers and promoters, we leveraged an efficient implementation of divisive hierarchical spectral clustering (see Methods)[21]. Using global information about the enhancer-enhancer, enhancer-promoter, and promoter-promoter interactions embedded in the interactivity graph, our clustering approach identifies a hierarchy of densely interacting enhancer and promoter groups with high intra-group and sparse inter-group interactions (Fig. 1a). Notably, our implementation of divisive hierarchical spectral clustering has tunable parameters (see Methods), enabling identification of hubs with granularity that matches user preferences. Given that enhancer-promoter hubs are dually defined by regulatory element composition and spatial organization, we hypothesized that these topological assemblies organize transcription across cancer genomes.

To evaluate this hypothesis, we first identified and characterized enhancer-promoter hubs using multi-omic data from DND41 T-ALL[22]. From Hi-C data, 1377 hubs with an average size of roughly 414 Kb were identified across the T-ALL genome (Supplementary Data 1). On a genomic organization perspective, these size estimates put hubs on a similar order of magnitude as TADs[23]. To determine if T-ALL hubs were simply a subset of TADs, we identified TAD boundaries from DND41 Hi-C data and evaluated their overlap with hub boundaries. This analysis revealed that hubs were distinct from TADs with only 5.3% of the 2754 hub boundaries overlapping with the more strongly insulated TAD boundaries (Supplementary Fig. 1a, b), a conclusion that was corroborated by analysis of cohesin subunit SMC1 HiChIP (Supplementary Fig. 1c). Given that hubs appeared to be unique from TADs, we characterized them by the count of spatial interactions between their constituent enhancer and promoter elements (Fig. 1a), where an interaction represents the presence of a contact between two regulatory elements (see Methods). As discussed in subsequent sections, we then leveraged Hi-C or HiChIP-measured enhancer and promoter interactivity within hubs, instead of individual algorithmically defined enhancer-promoter loops, as the basis for identifying differential hubs between two conditions (Fig. 1b).

To elucidate the basic organizational principles of enhancer-promoter hubs, we first examined within enhancer-promoter hub interaction counts. This analysis revealed that T-ALL hubs identified from either Hi-C or SMC1 HiChIP distributed asymmetrically, with only a small number of hubs harboring substantial numbers of spatial interactions (Fig. 2a and Supplementary Fig. 1d). While 50.5% of Hi-C enhancer-promoter hubs contained less than 20 interactions, 11.5% or 158 hubs demonstrated high interactivity with more than 83 interactions in T-ALL (Supplementary Data 1). Inspection of the relationship between number of promoters and enhancers participating in each hub and the extent of hub interactivity indicated that there was a positive correlation between interaction and regulatory element counts (Fig. 2b and Supplementary Fig. 1e) such that hubs with the most interactions also tended to have the highest interaction to promoter/enhancer ratios (Fig. 2c and Supplementary Fig. 1f). These observations suggest that the largest T-ALL hubs contain regulatory elements that are highly interacting, leading us to term this subset of enhancer-promoter assemblies as hyperinteracting hubs and refer to non-hyperinteracting hubs as regular hubs.

Given that both regular and hyperinteracting hubs are defined by the presence of active enhancers, we considered the possibility that they could be akin to super-enhancers. To examine the relationship between super-enhancers and enhancer-promoter hubs, we compared their location on the linear genome, observing that 60.5% of regular hubs did not coincide with super-enhancers. Importantly, 71.1% of hyperinteracting hubs contained two or more super-enhancers (Supplementary Fig. 1g). Among super-enhancer-containing hyperinteracting hubs, 92.6% included additional enhancers that were not part of super-enhancers. Together, these findings suggest that hyperinteracting hubs may have a potentially enhanced regulatory status compared to regular spatial hubs and further highlight their difference from super-enhancers, which are defined based on the linear genome clustering of enhancers.

Hyperinteracting hubs organize control of transcription factor expression in T leukemic cells

Postulating that enhancer-promoter hubs function to tightly direct transcription of certain genes, we sought to examine the transcriptional status of T-ALL hubs. To this end, we evaluated RNA enrichment over hub loci in comparison to random representative loci of identical genomic length using total transcript RNA-seq in DND41 T-ALL cells. Our analysis showed that median RNA enrichment over the observed hubs was significantly greater than median enrichment in all of the 5,000 sets of size-matched random representative loci that served as comparators for Hi-C and SMC1 HiChIP hubs (Fig. 2d, $p = 0.0002$; and Supplementary Fig. 1h, $p = 0.0002$). Together, these data suggest that T-ALL hubs are uniquely associated with transcriptional activation.

The observation that enhancer-promoter hubs are enriched for transcription in T-ALL led us to test if hyperinteracting hubs organize transcription on a broader scale than regular hubs and whether these two hub types are functionally distinct. Compared to regular hubs, SMC1 HiChIP hyperinteracting hubs on average spanned

~3.2 times more base pairs (Supplementary Fig. 1i, $p = 1E\text{-}53$). Examining transcription over hyperinteracting hubs revealed that median expression over these hubs was significantly higher than median expression over 10,000 sets of size-matched random representative loci from across the DND41 T-ALL genome (Fig. 2e, $p = 0.0001$; and

Supplementary Fig. 1j, $p = 0.0001$). Given that both regular and hyperinteracting hubs were distinctly enriched for transcription in comparison to random representative loci (Fig. 2d, e, and Supplementary Fig. 1h,j), we aimed to clarify their transcriptional state relative to one another. Consequently, we observed that hyperinteracting hubs
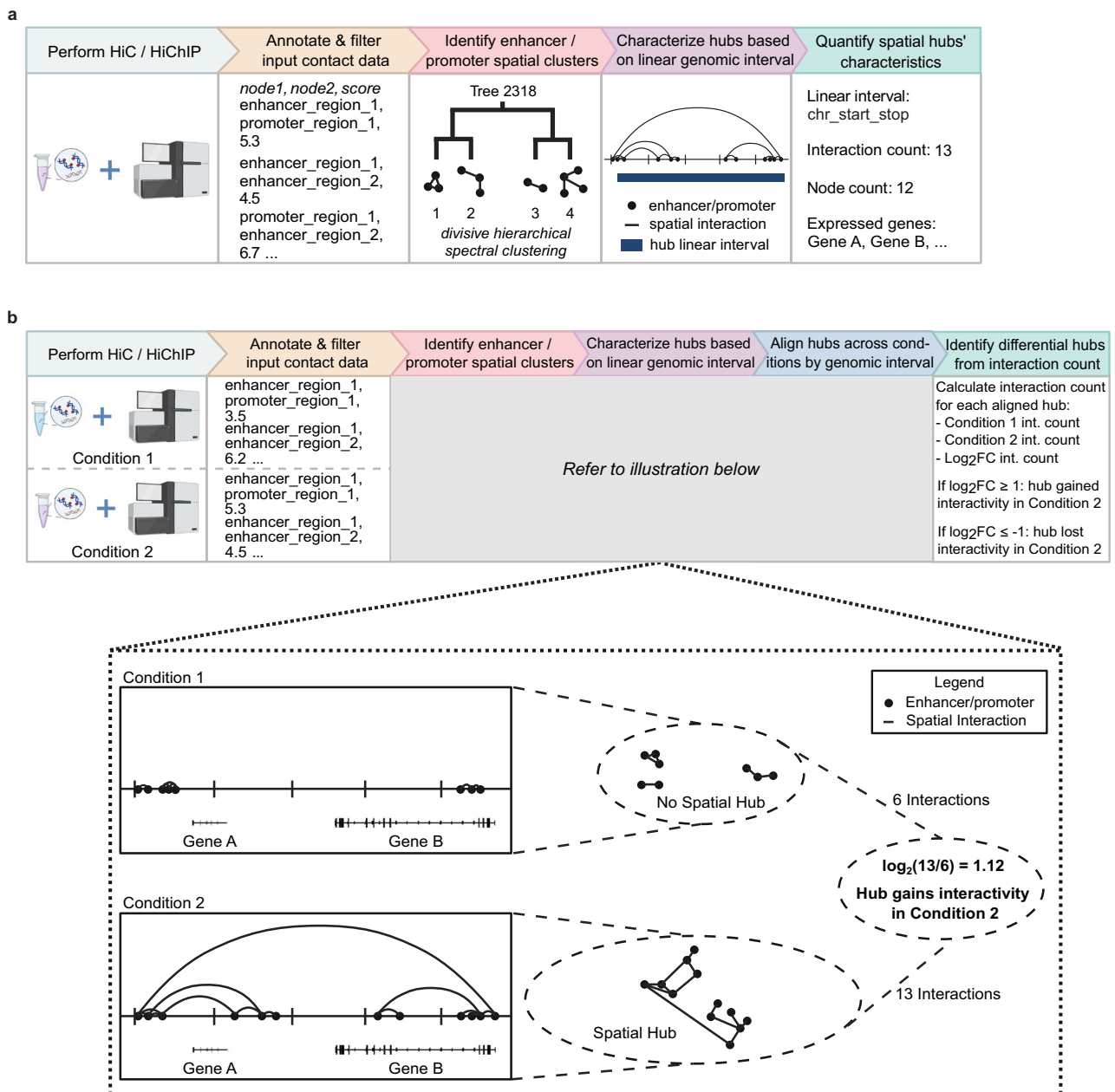

Fig. 1 | Enhancer-promoter hubs are systematically identified by divisive hierarchical spectral clustering of an undirected interactivity graph of enhancers and promoters. a Process of detecting enhancer-promoter hubs from raw chromatin conformation capture data. First, the interactivity graph of enhancer and promoter elements is created from regulatory element nodes connected by pairwise Hi-C or SMC1 HiChIP spatial interactions. Next, an efficient matrix-free divisive hierarchical spectral clustering algorithm is used to partition the enhancer-promoter interactivity graph into spatial clusters. Clusters are then characterized by their contiguous linear genomic intervals from their most upstream spatially interacting regulatory element to their most downstream spatially interacting regulatory element to form enhancer-promoter hubs. These hubs can be ranked based on their interaction count, enhancer/promoter number, and the expressed genes contained within their linear genomic intervals. Input interaction data depicted is for illustrative purposes only. Created with BioRender.com.

b Procedure for identifying differential enhancer-promoter hubs between two conditions based on within-hub total interactivity. Top: enhancer-promoter hubs separately identified in each condition are combined into a union set of hubs based on their linear genomic coordinates. The input interaction data depicted is for illustrative purposes only. Spatial enhancer hubs with markedly differential interactivity are identified based on $\log_2$ fold change of interaction count between the two conditions. Bottom: diagram of an illustrative differential hub. Bottom left depicts this hub on the linear genome where Hi-C valid interactions (arcs) connect enhancers and promoters (circle nodes) in condition 1 (upper) and condition 2 (lower). Bottom right illustrates a simplified, potential 3D rendering of this hub in each condition, demonstrating that cells in condition 2 gain more than two-fold interactivity at this locus to form a (differential) spatial hub. Created with BioRender.com.
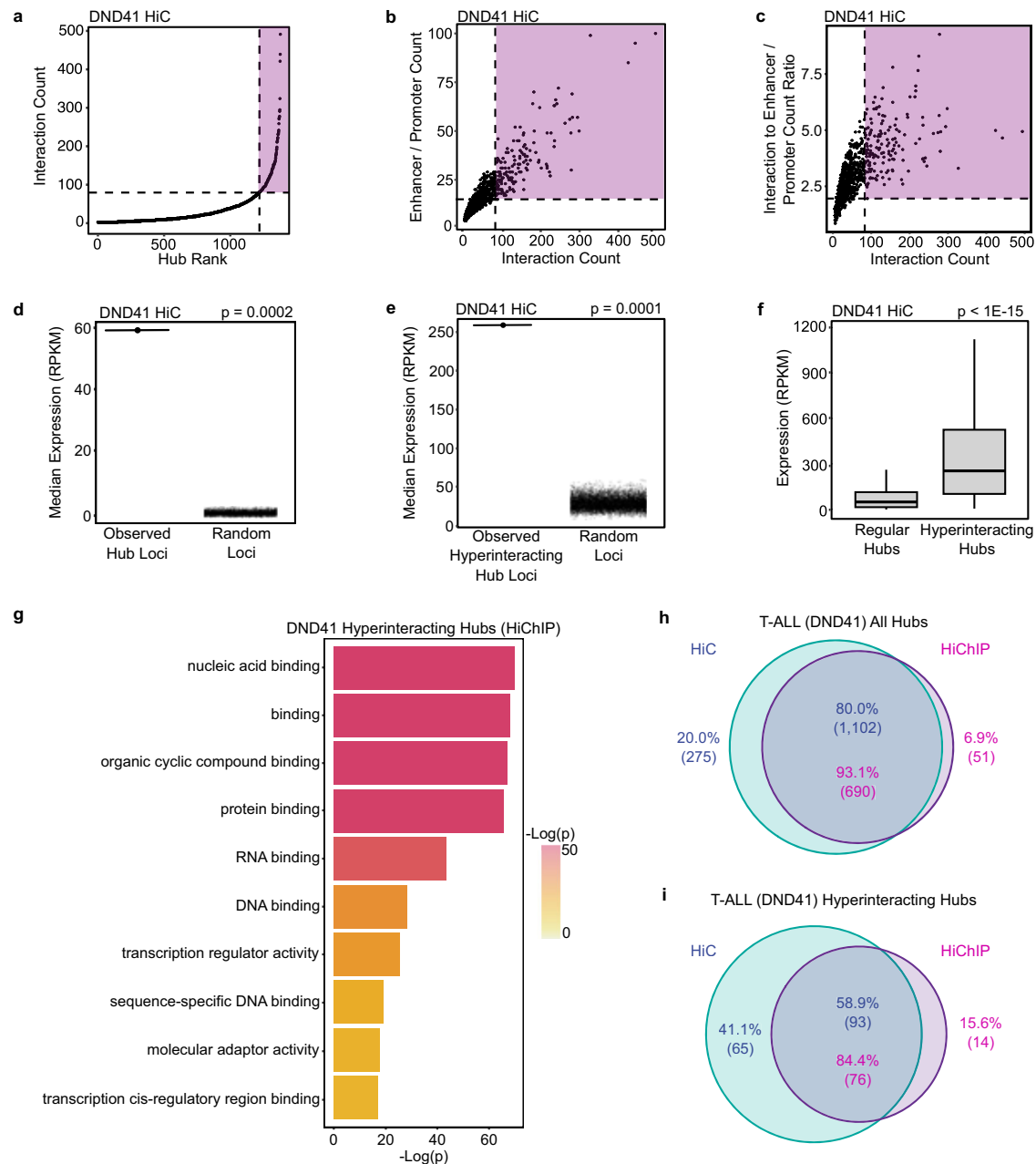
Fig. 2 | T-ALL hyperinteracting hubs are markedly transcribed and organize expression of genes encoding transcription factors and cofactors. a Enhancer-promoter hubs detected from T-ALL DND41 Hi-C data are plotted in ascending order of their total interactivity. The purple region marks hyperinteracting hubs, defined as hubs above the elbow of the total interactivity ranking curve. **b** Plot of Hi-C interaction count vs. enhancer/promoter element count in each T-ALL DND41 hub with hyperinteracting hubs marked in purple. **c** Plot of Hi-C interaction count vs. ratio of Hi-C interaction to enhancer/promoter counts in each T-ALL DND41 hub with hyperinteracting hubs marked in purple. **d** Median gene expression of T-ALL DND41 Hi-C hubs vs. 5000 sets of matched, randomly selected regions (empirical permutation *p*-value). **e** Median gene expression of T-ALL DND41 Hi-C hyperinteracting hubs vs. 10,000 sets of matched, randomly selected regions (empirical permutation *p*-value). **f** Box-and-whisker plots comparing transcription levels of T-ALL DND41 Hi-C regular (*n* = 1219) and hyperinteracting (*n* = 158) hubs. Box-and-whisker plots: center line, median; box limits, upper (75th) and lower (25th) percentiles; whiskers, 1.53 interquartile range. *P*-value: two-tailed Wilcoxon rank sum test. **g** Molecular function GO enrichment analysis of DND41 SMC1 HiChIP hyperinteracting hubs. *P*-value: Fisher's exact test. **h** Overlap of T-ALL DND41 hubs identified from Hi-C and SMC1 HiChIP data. Percentages and counts of overlapping and unique hubs are shown separately for Hi-C and SMC1 HiChIP. **i** Overlap of T-ALL DND41 hyperinteracting hubs identified from Hi-C and SMC1 HiChIP data. Percentages and counts of overlapping and unique hyperinteracting hubs are shown separately for Hi-C and SMC1 HiChIP.

contained more highly expressed genes (Supplementary Fig. 1k, *p* = 1E-25) and were significantly enriched for gene expression (Fig. 2f, *p* < 1E-15) compared to regular hubs detected from Hi-C data, an observation that was corroborated by SMC1 HiChIP-measured interactions among enhancers and promoters (Supplementary Fig. 1l, *p* < 1E-15). These findings support that hyperinteracting hubs coalesce enhancers and promoters to potentially establish transcriptionally permissive environments.

In light of these observations, we sought to determine the molecular functions of the genes contained within these highly transcribed T-ALL hyperinteracting hubs. Gene ontology (GO) enrichment analysis revealed that a significant fraction of expressed genes located within

these topological assemblies encoded proteins functioning as transcription factors and cofactors (Fig. 2g, *p* < 1E-15; and Supplementary Data 2), including genes encoding DNA binding factors *MYC*, *TP53*, and *YY1* as well as genes encoding chromatin and transcriptional coregulators *HDAC4*, *KAT5*, *DNMT3B*, and *EZH1* (Supplementary Data 1) with demonstrated role in leukemia[24–28]. As expected from the high degree of concordance between Hi-C and SMC1 HiChIP hubs (Fig. 2h, i), GO enrichment analysis with hyperinteracting hubs identified from Hi-C corroborated these hubs' distinct enrichment for transcription factors (Supplementary Data 2). Collectively, these data suggest that hyperinteracting hubs may serve to not only organize local transcription, but also control transcription across the genome by regulating transcription factors that act in *trans* upon distal loci.

To assess whether our observations in DND41 were generalizable to other T-ALL models, we performed hub analysis in CUTLL1 cells[29–31]. Similar to DND41, we observed asymmetrical distribution of interactions among enhancers and promoters and formation of hyperinteracting hubs in CUTLL1 (Supplementary Fig. 2a–c). Moreover, CUTLL1 hubs were not identical to super-enhancers (Supplementary Fig. 2d) and were significantly transcribed (Supplementary Fig. 2e, *p* = 0.0002; and Supplementary Fig. 2f, *p* = 0.0001) with hyperinteracting hubs distinctly enriched for gene expression (Supplementary Fig. 2g, *p* < 1E-15) and encoding molecules involved in nucleic acid/protein binding, among other functions characteristic of transcriptional regulators (Supplementary Fig. 2h, *p* < 1E-15; and Supplementary Data 2).

To better understand the potential regulatory environment formed by hyperinteracting hubs in T-ALL, we next examined these hubs at histone methyltransferase *DOT1L* and DNA methyltransferase *DNMT3B* (Supplementary Fig. 2i,j), two well-known genes involved in leukemogenesis with prognostic significance[25,32,33]. Both *DOT1L* and *DNMT3B* hyperinteracting hubs contained several long-range interactions among highly accessible genomic elements with elevated levels of active histone mark H3K27ac and/or transcription. The *DOT1L* locus was among the top 10 most interacting hubs in both DND41 and CUTLL1 and exhibited similar local interaction organization. On the other hand, in the hyperinteracting hub containing *DNMT3B*, a number of active elements adjacent to the *DNMT3B* promoter demonstrated more variable spatial interactivity between the two T-ALL cell lines (Supplementary Fig. 2i,j). Taken in conjunction, these data support the ability of hub-based analysis to detect key topological assemblies with varying spatial interaction structure across different models of a given cancer.

Organizational principles of enhancer-promoter hubs are shared between T cell leukemia and B cell lymphoma

Observing organizational principles of highly interacting enhancer-promoter hubs in T-ALL led us to investigate whether these topological assemblies create complex networks of regulatory elements interacting with genes encoding transcriptional regulators in other cancer types. For this reason, we identified hubs from Rec-1 MCL cells using both Hi-C and SMC1 HiChIP data[14]. Similar to T-ALL, Rec-1 hubs were largely different from both TADs and super-enhancers (Supplementary Fig. 3a–d), and demonstrated significant asymmetry in interactivity and enhancer/promoter count distributions (Fig. 3a, b and Supplementary Fig. 3e, f). The existence of a positive correlation between interaction counts and interaction to enhancer/promoter ratios further supports the presence of hyperinteracting hubs in MCL (Fig. 3c and Supplementary Fig. 3g). Rec-1 hubs in general (Fig. 3d, *p* = 0.0002; and Supplementary Fig. 3h, *p* = 0.0002) and hyperinteracting hubs in particular (Fig. 3e, *p* = 0.0001; and Supplementary Fig. 3i, *p* = 0.0001) both demonstrated significant transcriptional activity. In comparison to regular hubs, Rec-1 hyperinteracting hubs spanned significantly more base pairs on the linear genome (Supplementary Fig. 3j, *p* = 1E-31), contained increased numbers of highly expressed genes

(Supplementary Fig. 3k, *p* = 1E-32), and exhibited markedly more transcriptional activity than regular hubs (Fig. 3f, *p* < 1E-15; and Supplementary Fig. 3l, *p* < 1E-15), similar to T-ALL.

To examine the potential role of genes located in MCL hyperinteracting hubs, we performed molecular function GO enrichment analysis. This analysis revealed that, in agreement with T-ALL, a significant fraction of genes located within MCL hyperinteracting hubs encoded transcription factors and cofactors (Fig. 3g, *p* < 1E-9; and Supplementary Data 2), including *MYC*, *CTCF*, *ETS1*, *KAT5*, and *DOT1L*. These genes were located in hyperinteracting hubs in both Rec-1 MCL and DND41 T-ALL cells (Supplementary Data 1); yet, the structure of these hubs appeared different between the two cancer types, as exemplified by the *DOT1L* locus (Supplementary Fig. 3m). On the other hand, several transcription regulators with roles in oncogenesis, such as *DNMT3B*[25] (Supplementary Fig. 3n,o; Supplementary Data 1) and B lymphocyte lineage transcription factor *PAX5*[34] (Supplementary Fig. 3p,q; Supplementary Data 1) were only expressed and positioned at hyperinteracting hubs in either T-ALL or MCL and not both. As such, these data corroborate observations in T-ALL (Fig. 2g) and further support the role of hyperinteracting hubs as potential regulatory assemblies that orchestrate gene expression in hematological malignancies.

Enhancer-promoter hub identification is robust to chromatin conformation capture assay resolution

In-situ Hi-C provides unbiased chromatin conformation maps at the expense of short-range enhancer-promoter loop resolution. In contrast, protein-centric assays, including HiChIP, are biased to increase resolution and support the identification of looping interactions mediated by a particular protein[35]. To assess the impact of chromatin conformation capture technology on the detection of enhancer-promoter hubs, we first compared hubs identified with Hi-C and SMC1 HiChIP in DND41 T-ALL. Hubs detected with SMC1 HiChIP in DND41 were fewer and on average larger than hubs detected with Hi-C. 741 SMC1 HiChIP hubs were identified with an average span of 1.09 Mb compared to 1377 Hi-C hubs spanning an average of 414 Kb. A similar trend was observed for hyperinteracting hubs, where SMC1 HiChIP and Hi-C identified 90 and 158 hyperinteracting hubs with an average span of 2.73 Mb and 1.00 Mb on the linear DND41 genome, respectively (Supplementary Data 1). Despite these differences, the vast majority of hubs identified with SMC1 HiChIP were also detectable with Hi-C, where 93.1% of total SMC1 HiChIP hubs and 84.4% of hyperinteracting SMC1 HiChIP hubs specifically coincided with Hi-C total and hyperinteracting hubs, respectively (Fig. 2h, i). Hence, DND41 SMC1 HiChIP hyperinteracting hubs were largely a subset of Hi-C hubs. While 84.4% of SMC1 HiChIP hyperinteracting hubs overlapped with Hi-C hyperinteracting hubs, only 58.9% of Hi-C hyperinteracting hubs overlapped with their SMC1 HiChIP counterparts (Fig. 2i). As shown in Fig. 2a–f and Supplementary Fig. 1d–f, h, j, l, both hubs detected from SMC1 HiChIP and Hi-C exhibited similar transcriptional activity as well as interaction and enhancer/promoter count distributions. In sum, these data exhibit the high fidelity of our analysis to identify enhancer-promoter hubs in T-ALL from both Hi-C and SMC1 HiChIP experiments.

In order to corroborate these observations, we repeated comparative analyzes with SMC1 HiChIP and Hi-C data from Rec-1 MCL cells. Similar to DND41 T-ALL, Hi-C hubs coincided with more than 75% of HiChIP hubs (Fig. 3h). However, the percentage of Rec-1 SMC1 HiChIP hyperinteracting hubs overlapping with Hi-C hyperinteracting hubs was only 32.0% (Fig. 3i), which was likely due to two Hi-C outlier hubs with disproportionately large interaction counts (Fig. 3a and Supplementary Fig. 3r) that resulted in a more stringent cutoff for categorizing a hub as hyperinteracting from Hi-C compared to HiChIP measurements. Nevertheless, Rec-1 Hi-C and SMC1 HiChIP hyperinteracting hubs exhibited similar
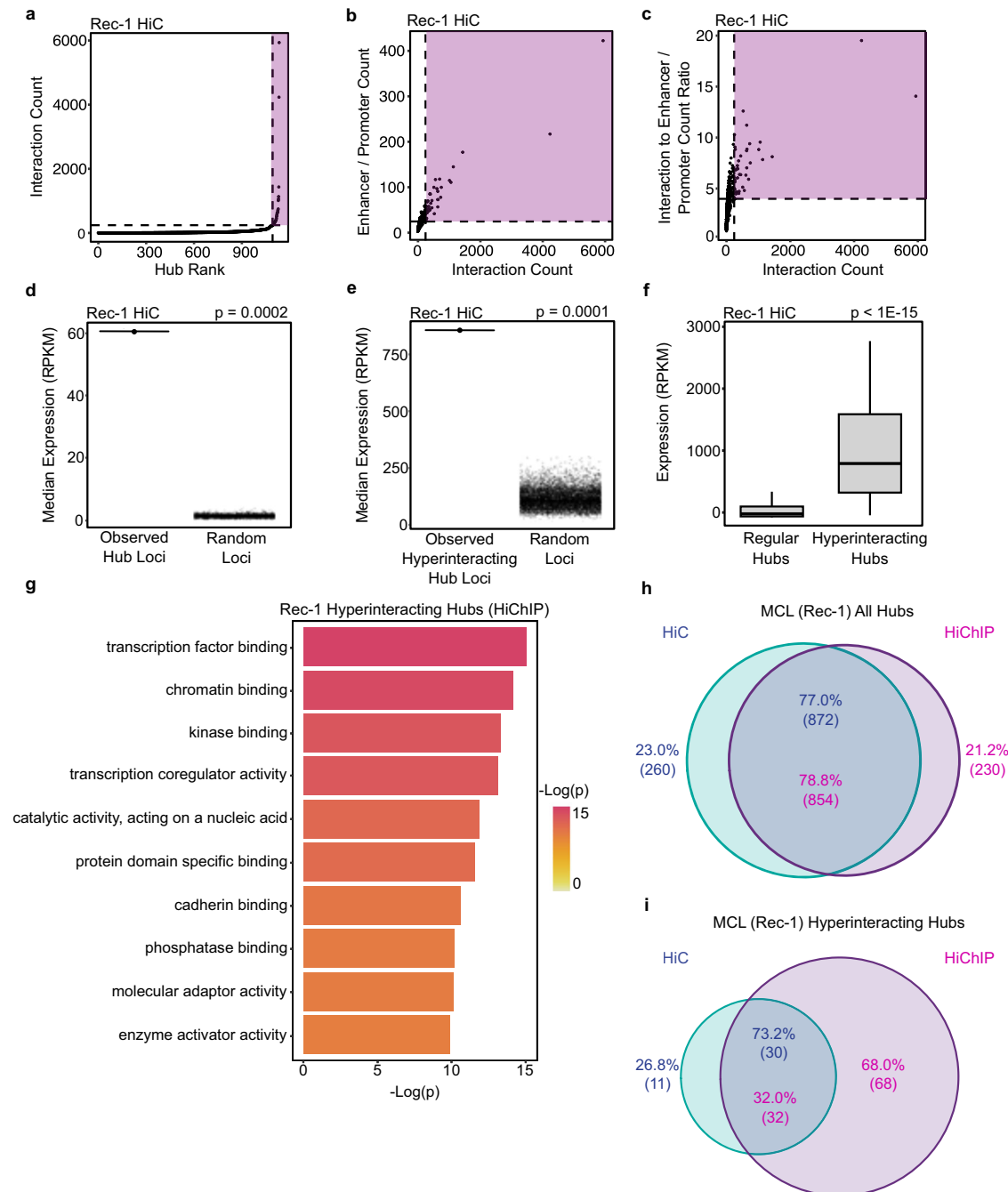
Fig. 3 | MCL hyperinteracting hubs are markedly transcribed and organize expression of genes encoding transcription factors and cofactors. a Enhancer-promoter hubs detected from MCL Rec-1 Hi-C data are plotted in ascending order of their total interactivity. The purple region marks hyperinteracting hubs, defined as hubs above the elbow of the total interactivity ranking curve. **b** Plot of Hi-C interaction count vs. enhancer/promoter element count in each MCL Rec-1 hub with hyperinteracting hubs marked in purple. **c** Plot of Hi-C interaction count vs. ratio of Hi-C interaction to enhancer/promoter counts in each MCL Rec-1 hub with hyperinteracting hubs marked in purple. **d** Median gene expression of MCL Rec-1 Hi-C hubs vs. 5000 sets of matched, randomly selected regions (empirical permutation *p*-value). **e** Median gene expression of MCL Rec-1 Hi-C hyperinteracting hubs vs. 10,000 sets of matched, randomly selected regions (empirical permutation *p*-value). **f** Box-and-whisker plots comparing transcription levels of MCL Rec-1 Hi-C regular ($n = 1091$) and hyperinteracting ($n = 41$) hubs. Box-and-whisker plots: center line, median; box limits, upper (75th) and lower (25th) percentiles; whiskers, 1.53 interquartile range. *P*-value: two-tailed Wilcoxon rank sum test. **g** Molecular function GO enrichment analysis of MCL Rec-1 SMC1 HiChIP hyperinteracting hubs. *P*-value: hypergeometric test. **h** Overlap of MCL Rec-1 hubs identified from Hi-C and SMC1 HiChIP data. Percentages and counts of overlapping and unique hubs are shown separately for Hi-C and SMC1 HiChIP. **i** Overlap of MCL Rec-1 hyperinteracting hubs identified from Hi-C and SMC1 HiChIP data. Percentages and counts of overlapping and unique hyperinteracting hubs are shown separately for Hi-C and SMC1 HiChIP.

structural and transcriptional characteristics (Fig. 3a–f and Supplementary Fig. 3e–i, l). In tandem with the results of our comparative analyzes in T-ALL, these data suggest that hub analysis is robust across a spectrum of chromatin conformation capture assays with varying resolution of enhancer-promoter looping. More importantly, our analysis supports the role of enhancer-promoter hubs as potentially important units of genome organization rather than artifacts of a particular chromatin capture assay.

Organizational principles of enhancer-promoter hubs in hematological cancers are generalizable to non-hematological cancers

Intrigued by the commonality of hub organizational principles in leukemia and lymphoma, we sought to evaluate structural and transcriptional characteristics of enhancer-promoter hubs in non-hematological cancers. To this end, we identified hubs in two triple-negative breast cancer (TNBC) cell lines, HCC1599 and MB157, using SMC1 HiChIP data[14]. Analysis of TNBC enhancer-promoter hubs confirmed that they could be stratified on the basis of interaction count into two distinct groups of regular and hyperinteracting hubs (Fig. 4a–c and Supplementary Fig. 4a–c), both of which were markedly transcribed (Fig. 4d, $p = 0.0002$; Fig. 4e, $p = 0.0001$; Supplementary Fig. 4d, $p = 0.0002$; and Supplementary Fig. 4e, $p = 0.0001$) and different from super-enhancers (Supplementary Fig. 4f, g). Similar to T-ALL and MCL, hyperinteracting hubs in TNBC contained a greater number of highly expressed genes (Supplementary Fig. 4h, $p = 1E-23$; and Supplementary Fig. 4i, $p = 1E-32$). Furthermore, TNBC hyperinteracting hubs were significantly more transcribed (Fig. 4f, $p = 1E-11$; and Supplementary Fig. 4j, $p < 1E-15$) and spanned larger genomic distances than regular hubs (Supplementary Fig. 4k, $p = 1E-23$; and Supplementary Fig. 4l, $p = 1E-14$), again mirroring T-ALL and MCL hubs. TNBC hyperinteracting hubs also predominantly formed at genes encoding transcription factors and cofactors (Fig. 4g and Supplementary Fig. 4m; Supplementary Data 2), some of which were only present in TNBC. For example, TNBC-associated hyperinteracting hubs formed around *SOX9* (Fig. 4h, i), which encodes a transcription factor with demonstrated role in TNBC oncogenesis[36–38], and *TRPS1* (Supplementary Fig. 4n,o), which encodes a transcription factor that serves as a highly specific marker for breast carcinoma including TNBC[39,40]. In contrast, some of the most prominent hyperinteracting hubs in T-ALL and/or MCL, including *DOT1L*, *DNMT3B*, and *PAX5* (Supplementary Fig. 3m,n,p), were not hyperinteracting or not present in TNBC (Supplementary Fig. 4p–r). Taken together, our characterization of TNBC hubs corroborates T-ALL and MCL observations, and suggests that hyperinteracting hubs may organize transcriptional regulation of *trans*-acting factors in a lineage-associated manner to inform broader gene expression programs.

Hyperinteracting hubs are more lineage associated than regular hubs

The lineage association of certain key hyperinteracting hubs (Supplementary Figs. 2i,j 3m–q, 4n–r, and Fig. 4h, i) led us to systematically assess the similarity of hubs identified from T-ALL, MCL, and TNBC SMC1 HiChIP data. By defining hub similarity as the percentage of hubs with overlapping genomic loci between two cancer types, we observed that hubs are relatively conserved, with at least 51% of hubs being shared between any two cancer types (Fig. 5a), and at least 23% of hubs being shared across all four cancer types. Some of the common hubs in T-ALL, MCL, and TNBC contained genes encoding key transcription and DNA replication regulators including TET3, CTCF, KLF10, AKT1, E2F1/4/6, PARP enzymes, polymerases, and NFkB proteins (Supplementary Data 1).

Observing that hubs are generally shared across T-ALL, MCL, and TNBC, we went on to compare hyperinteracting hubs in these cancers. In contrast to all hubs, hyperinteracting hubs were more lineage associated such that less than ~50% of these hubs were shared between any two cancer types (Fig. 5b). The observation that hyperinteracting hubs were more lineage associated than regular hubs was further supported by comparison of hubs detected from Hi-C data in DND41 and CUTLL1 T-ALL as well as Rec-1 MCL (Supplementary Fig. 5a, b).

To evaluate lineage correlation of hyperinteracting hubs in greater depth, we closely scrutinized their organization across T-ALL, MCL, and TNBC. Lineage-associated hyperinteracting hubs were separated into two groups based on their presence and interactivity.

The first group consisted of hyperinteracting hubs that only existed in one cancer type, as exemplified by the *SOX9*, *PAX5*, and *DNMT3B* genes, which were only coalesced within enhancer-promoter hubs in TNBC, MCL, and T-ALL, respectively (Fig. 4h, i, and Supplementary Figs. 3n, 3p, and 4q–r). The second group consisted of hubs that were highly interacting in some cancer types and less interacting in others, as exemplified by the *DOT1L*, *KAT5* (also known as *TIP60*), and *CEBPB* genes. Similar to the hub containing *DOT1L* (Supplementary Figs. 3m and 4p), the hub containing *KAT5*, which encodes a histone acetyltransferase with known role in driving *HOXA* gene expression in leukemia[24,41], was hyperinteracting in T cell leukemia and B cell lymphoma but regularly interacting in TNBC (Fig. 5c, d). In contrast, the hub containing *CEBPB*, which encodes a transcription factor implicated in normal and malignant breast epithelium[42], was regularly interacting in leukemia and lymphoma, but was hyperinteracting in TNBC (Fig. 5e, f). Other notable hyperinteracting hub genes that were spatially proximate to multiple regulatory elements in T-ALL and MCL but to only a few in TNBC included the retinoic acid receptor *RXRB*, the apoptosis regulator *BAK1*, and the heat shock protein *HSPA9* (also known as mortalin; Supplementary Data 1). On the other hand, as previously discussed, the transcription factor *TRPS1* was located within a hyperinteracting hub only in TNBC and not T-ALL or MCL (Supplementary Fig. 4n,o). Together, these data suggest that hyperinteracting hubs are a distinct subset of enhancer-promoter hubs that demonstrate notable lineage association and are potentially involved in transcriptional control.

Despite observing that most hyperinteracting hubs were lineage associated, we identified ten hyperinteracting hubs that were present in T-ALL, MCL, and TNBC (Supplementary Fig. 5c). Some of these topological assemblies, which were conserved on the basis of genomic overlap, formed at genes implicated in oncogenesis. Notable expressed genes from these hubs include the proto-oncogene *MYC*, the *MYC*-interacting chromatin effector *PYGO2*[43], the cell proliferation driver *EFNA4*[44,45], the well-studied protein deacetylase *SIRT2*[46], and the highly expressed transcription factor and cancer biomarker *ZNF217*[47] (Supplementary Fig. 5c–e). The intriguing commonality of these few hyperinteracting hubs highlights the potential importance of enhancer-promoter hubs in oncogenesis.

Enhancer-promoter hubs are reorganized in GSI-resistant T-ALL

Given the correlation between enhancer-promoter hubs' interactivity and transcriptional level in T-ALL, MCL, and TNBC, we aimed to examine the potential structure-function relationship of hubs by studying their changes during anticancer drug resistance acquisition. To this end, we first screened for differential hubs in T leukemic cells that were either sensitive or resistant to gamma secretase inhibitor (GSI), an antagonist of NOTCH1 signaling, with *NOTCH1* being the most frequently mutated gene in T-ALL[48,49]. Differential hubs were identified based on within-hub interaction count changes between GSI-sensitive and GSI-resistant DND41 T-ALL cells (Fig. 1b and Supplementary Data 3) (see Methods). We postulated that identifying differential hubs should reveal genes with key roles in GSI resistance without the need for differential loop calling analysis, should hubs function as topological assemblies of gene expression control.

Differential hub analysis in GSI-sensitive and GSI-resistant DND41 T-ALL identified 217 differential hubs that had at least 2-fold change in interactivity or were de novo gained/lost in GSI resistance (Fig. 6a). Notably, these differential hubs were distinct from differential compartments and TADs identified in GSI-sensitive and -resistant cells (Supplementary Fig. 6a–d) in accordance with enhancer-promoter hubs' broader separation from TADs and super-enhancers (Supplementary Figs. 1a–c, g, and 3a–d). To determine whether these spatially differential hubs were also epigenetically or transcriptionally altered in GSI resistance, we examined their differential chromatin accessibility, chromatin activity, and gene expression using ATAC-seq, H3K27ac
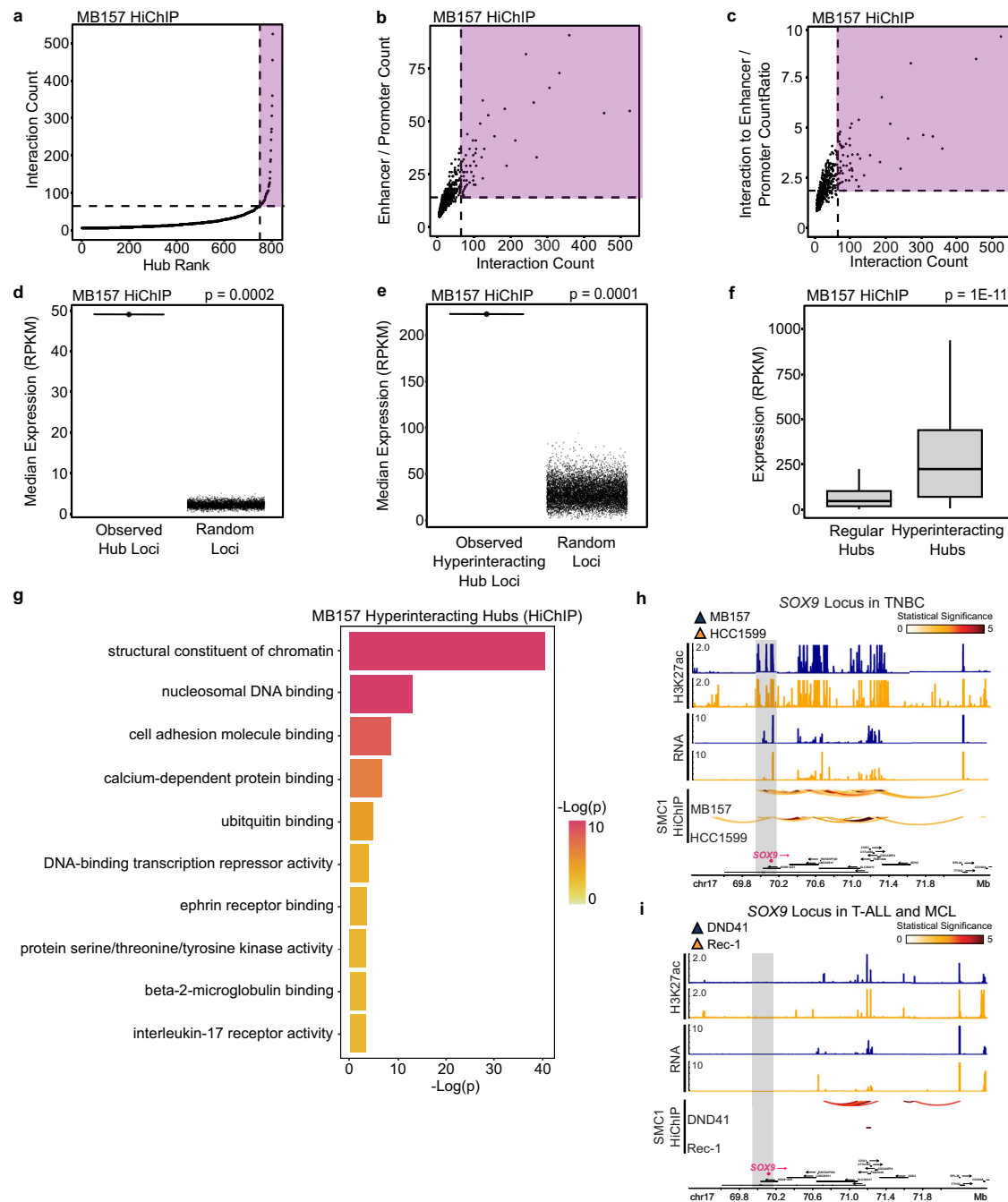
Fig. 4 | TNBC hyperinteracting hubs are markedly transcribed and organize expression of genes encoding transcription factors and cofactors. a Enhancer-promoter hubs detected from TNBC MB157 SMC1 HiChIP data are plotted in ascending order of their total interactivity. The purple region marks hyperinteracting hubs, defined as hubs above the elbow of the total interactivity ranking curve. **b** Plot of SMC1 HiChIP interaction count vs. enhancer/promoter element count in each TNBC MB157 hub with hyperinteracting hubs marked in purple. **c** Plot of SMC1 HiChIP interaction count vs. ratio of interaction to enhancer/promoter counts in each TNBC MB157 enhancer-promoter hub with hyperinteracting hubs marked in purple. **d** Median gene expression of TNBC MB157 SMC1 HiChIP hubs vs. 5000 sets of matched, randomly selected regions (empirical permutation *p*-value).

e Median gene expression of TNBC MB157 SMC1 HiChIP hyperinteracting hubs vs. 10,000 sets of matched, randomly selected regions (empirical permutation *p*-value). **f** Box-and-whisker plots comparing transcription levels of TNBC MB157 SMC1 HiChIP regular (*n* = 752) and hyperinteracting (*n* = 54) hubs. Box-and-whisker plots: center line, median; box limits, upper (75th) and lower (25th) percentiles; whiskers, 1.53 interquartile range. *P*-value: two-tailed Wilcoxon rank sum test. **g** Molecular function GO enrichment analysis of TNBC MB157 SMC1 HiChIP hyperinteracting hubs. *P*-value: hypergeometric test. **h,i** H3K27ac ChIP-seq, RNA-seq, and SMC1 HiChIP interactions surrounding gray box-marked *SOX9* markedly differ between TNBC (**h**) and T-ALL/MCL (**i**). Bottom tracks indicate the Ensembl gene position.

ChIP-seq, and RNA-seq, respectively. Interestingly, hubs with significant loss or gain of interactivity in GSI resistance demonstrated markedly decreased or increased chromatin opening (Fig. 6b, *p* = 1E-5; and Fig. 6c, *p* = 1E-8) and chromatin activity (Fig. 6b, *p* = 1E-8; and Fig. 6c, *p* = 1E-12), respectively, with a similar trend for gene expression

(Fig. 6b, *p* = 0.14; and Fig. 6c, *p* = 0.20). The absence of statistically significant changes in hubs' transcriptional activity could be in part attributed to a lack of concordant changes in all the genes located within these topological assemblies such that aggregate measurements of hub transcription appear less variant. Nonetheless, close
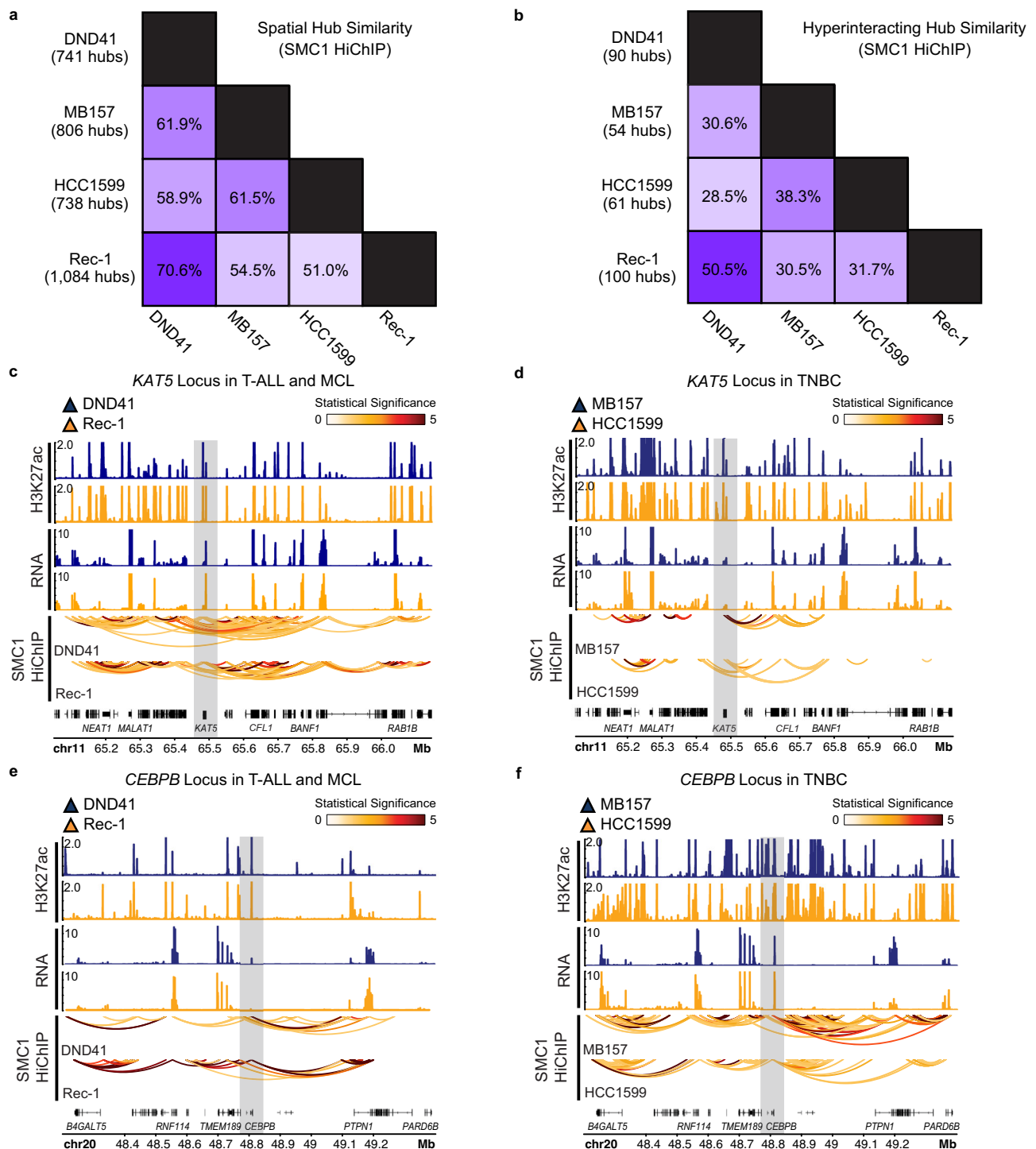
Fig. 5 | Hyperinteracting hubs are more lineage associated compared to regular hubs in T-ALL, MCL, and TNBC. a Matrix of pairwise SMC1 HiChIP enhancer-promoter hub similarity across T-ALL, MCL, and TNBC cells. Total hub count for each cell line is listed on the left. **b** Matrix of pairwise SMC1 HiChIP hyperinteracting hub similarity across T-ALL, MCL, and TNBC cells. Total hyperinteracting hub count for each cell line is listed on the left. **c**, **d** Gray box-marked *KAT5* forms a hyperinteracting enhancer-promoter hub with SMC1 HiChIP arcs connecting active regulatory elements and genes marked with H3K27ac ChIP-seq and RNA-seq, respectively, in T-ALL DND41 and MCL Rec-1 (**c**) but not TNBC MB157 and HCC1599 (**d**). Bottom tracks indicate Ensembl gene position. **e**, **f** Gray box-marked *CEBPB* forms a hyperinteracting enhancer-promoter hub with SMC1 HiChIP arcs connecting active regulatory elements and genes marked with H3K27ac ChIP-seq and RNA-seq, respectively, in TNBC MB157 and HCC1599 (**f**) but not T-ALL DND41 and MCL Rec-1 (**e**). Bottom tracks indicate Ensembl gene position.

examination of hubs with a gain of interactivity in GSI-resistant T leukemic cells showed that some of these topological assemblies formed at B cell-related genes, including *Early B-cell factor 1* (*EBF1*; Supplementary Data 3), as well as genes involved in glucocorticoid signaling, including glucocorticoid receptor *NR3C1* (Supplementary Data 2).

EBF1, which encodes a transcription factor promoting development of GSI resistance in T-ALL[22], is both more interacting (Fig. 6a) and expressed (Supplementary Fig. 6e, $p = 0.003$) in GSI-resistant compared to GSI-sensitive cells. Similarly, *NR3C1*, a key driver of T-ALL steroid resistance[50], participated in an enhancer-promoter hub with
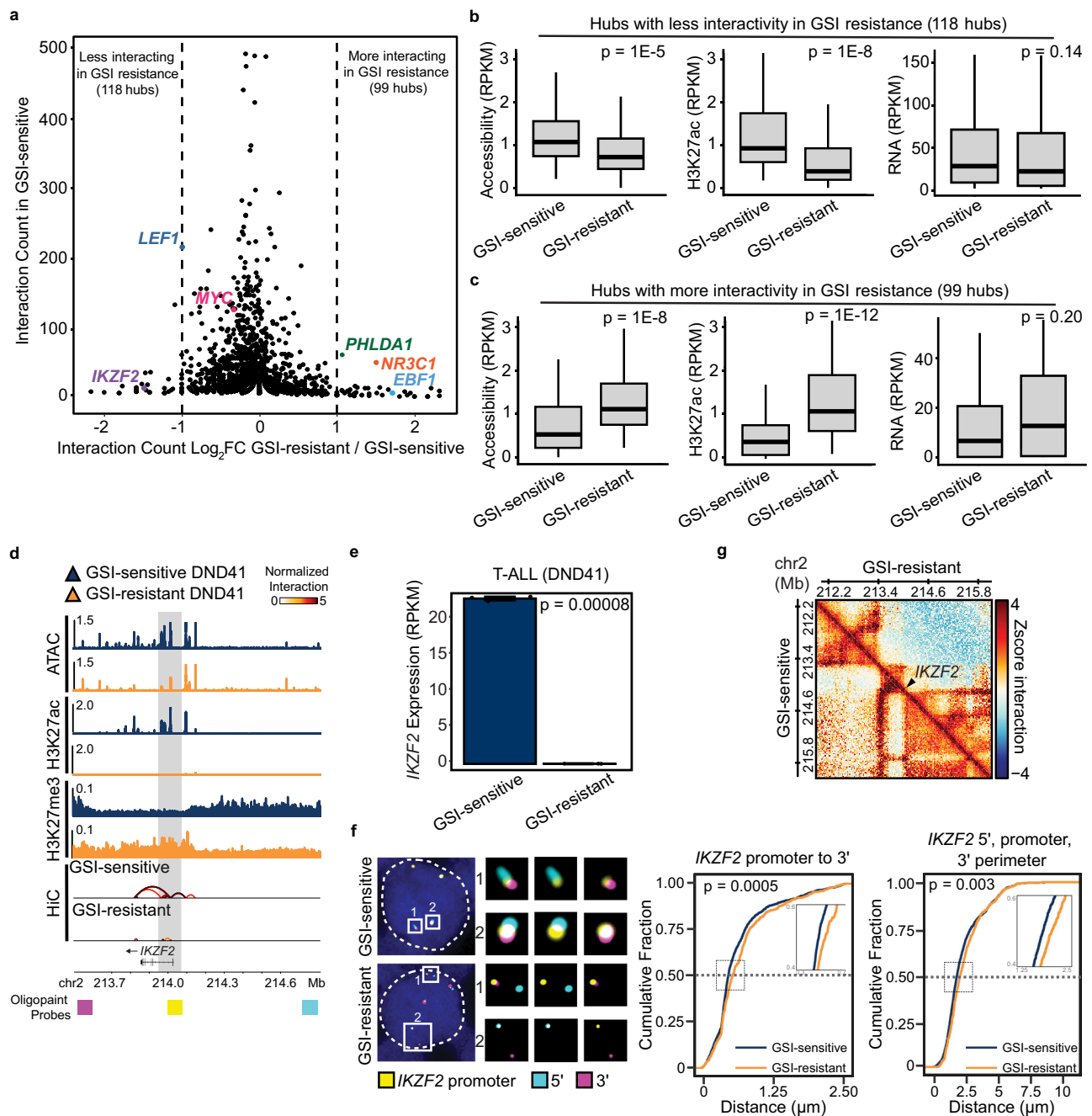
Fig. 6 | Loss of *IKZF2* hyperinteracting hub coincides with decrease in chromatin activity, gene expression, and architectural stripe in GSI-resistant T-ALL. **a** Scatter plot showing log$_2$ fold change of hub interactivity in GSI-resistant DND41 vs. hub interaction counts in GSI-sensitive cells. Dotted lines mark hubs with ≥ 2-fold decrease ("less interacting in GSI resistance") or increase ("more interacting in GSI resistance") in interaction counts in GSI-resistant vs. GSI-sensitive DND41 cells. Selected hubs are labeled. **b** Box-and-whisker plots comparing ATAC-seq (left), H3K27ac ChIP-seq (center), and RNA-seq (right) in hubs that lost interactivity in GSI-resistant cells. Box-and-whisker plots: center line, median; box limits, upper (75th) and lower (25th) percentiles; whiskers, 1.53 interquartile range (n = 118). *P*-value: two-tailed Wilcoxon rank sum test. **c** Box-and-whisker plots comparing ATAC-seq (left), H3K27ac ChIP-seq (center), and RNA-seq (right) in hubs that gained interactivity in GSI-resistant cells. Box-and-whisker plots: center line, median; box limits, upper (75th) and lower (25th) percentiles; whiskers, 1.53 interquartile range (n = 99). *P*-value: two-tailed Wilcoxon rank sum test. **d** Concordant changes in ATAC-seq, H3K27ac ChIP-seq, H3K27me3 CUT&RUN, and Hi-C over the *IKZF2* hub between GSI-sensitive and GSI-resistant DND41 cells. Oligopaint DNA FISH probes are marked with pseudo-color magenta (*IKZF2* 3'), yellow (gray box-marked *IKZF2* promoter), and cyan (*IKZF2* 5') below the Ensembl gene track. **e** *IKZF2* normalized RNA-seq reads in GSI-sensitive and GSI-resistant DND41 cells. Each dot represents a biological replicate (n = 3). *P*-value: two-sided *t*-test; error bars: ± SEM. **f** Center and right: cumulative distribution plots of the closest distance between the noted elements in the same cell between 1548 GSI-sensitive and 533 GSI-resistant allelic interactions (Kolmogorov–Smirnov *p*-value). Inserts show magnified curves at the gray boxes (not drawn to scale). Center: mean (± S.D.) distance in GSI-sensitive and GSI-resistant cells is 0.642 (± 0.50) μm, and 0.704 (± 0.52) μm, respectively. Right: mean (± S.D.) in GSI-sensitive and GSI-resistant cells is 2.26 (± 1.31) μm, and 2.41 (± 1.27) μm, respectively. Left: representative cells. **g** Normalized Hi-C contact maps in GSI-resistant (upper triangle) and GSI-sensitive (lower triangle) DND41 cells at the *IKZF2* locus.

more than a 2.5-fold gain of interactivity (Fig. 6a) and expression (Supplementary Fig. 6f, $p = 0.003$) in the GSI-resistant state. On the other hand, scrutinization of hubs with loss of interactivity in GSI-resistant cells revealed disruption of highly interacting enhancer-promoter assemblies at genes involved in T cell biology and T cell receptor signaling (Supplementary Data 2), including *LEF1* and *IKZF2* (Fig. 6a). These data are supported by earlier findings showing that T-ALL GSI resistance is partially mediated by downregulation of a T cell-related transcription program in favor of a B cell-related one[22], hence reinforcing the potential functional importance of enhancer-promoter hub restructuring during GSI resistance.

Optical mapping confirms reorganization of enhancer-promoter hubs at individual GSI-resistant T-ALL cells

Given that downregulation of a T cell-associated transcription program spurs acquisition of GSI resistance in T-ALL[22], we closely examined hubs at genes encoding T cell lineage-restricted transcription factors *LEF1* and *IKZF2*. The *LEF1* and *IKZF2* loci experienced nearly 2- and 3-fold reductions in interactivity in GSI-resistant T-ALL, respectively (Supplementary Fig. 6g and Fig. 6d, Supplementary Data 3). In line with decreased interactivity at these two loci, we observed marked reduction in chromatin accessibility, substantial depletion of active enhancer mark H3K27ac, and deposition of repressive mark H3K27me3 (Supplementary Fig. 6g and Fig. 6d), which was concomitant with significant repression of *LEF1* (Supplementary Fig. 6h, $p = 0.0003$) and *IKZF2* (Fig. 6e, $p = 0.00008$).

We next sought to establish how reduction of interaction frequency at *LEF1* and *IKZF2* enhancer-promoter hubs detected from Hi-C relates to physical separation of regulatory elements within these two loci in individual GSI-resistant cells. To this end, we used high-resolution Oligopaint DNA fluorescence in situ hybridization (FISH) and 3D confocal imaging to visualize physical perimeters of three genomic elements at each locus in GSI-sensitive and GSI-resistant DND41 cells.

To detect the disruption of *LEF1* at a single-cell resolution, we designed Oligopaint DNA FISH probes hybridizing to three regulatory elements surrounding *LEF1*: the *LEF1* promoter (5' end of the locus), a lineage-restricted *LEF1* enhancer (center of the locus), and the *RPL34* promoter (3' end of the locus). Measuring pairwise distances between the *LEF1* promoter, the *RPL34* promoter, and the lineage-restricted enhancer probes across 1319 GSI-sensitive and 640 GSI-resistant allelic interactions showed a significant increase in their spatial perimeter (Supplementary Fig. 6i, $p = 0.002$), suggesting expansion of the *LEF1* hyperinteracting hub in GSI resistance. This observation was in agreement with genomic data, which indicated the loss of an architectural stripe connecting *LEF1* with its flanking region (Supplementary Fig. 6j).

Similar to the *LEF1* hub, optical mapping of the *IKZF2* locus using three-color Oligopaint DNA FISH with 3D confocal microscopy showed significant separation of the genomic elements at this hub in GSI-resistant DND41 (Fig. 6f, $p = 0.003$), in line with the visible loss of an architectural stripe on the *IKZF2* Hi-C contact frequency map (Fig. 6g). Although additional optical mapping is required to demonstrate multiway interactions within hubs in various contexts, our single-cell resolution studies affirm the dynamic structure of enhancer-promoter hubs and further support their potential role in organizing regulation of genes involved in anticancer drug resistance.

Ibrutinib resistance reorganizes enhancer-promoter hubs in MCL

The observation that enhancer-promoter hub spatial changes coincide with transcriptional changes associated with GSI resistance in T-ALL led us to examine whether the same relationship holds in MCL upon resistance to the BTK inhibitor ibrutinib, which is approved for the treatment of various hematological malignancies[51]. We followed the same methodology as our T-ALL studies to identify differential hubs between ibrutinib-sensitive and ibrutinib-resistant Rec-1 MCL. Similar to GSI-resistant T-ALL, this analysis revealed that a small fraction of hubs (159 or ~15%) were differential, notably gaining or losing interactivity in ibrutinib-resistant MCL (Fig. 7a). Similar to GSI-resistant T-ALL, differential hubs were separate from differential compartments and TADs identified in ibrutinib-resistant MCL (Supplementary Fig. 7a–d), further supporting enhancer-promoter hubs' status as distinct topological features (Supplementary Figs. 1a–c and 3a–c). Investigation of epigenetic and transcriptional changes showed that significant loss or gain of interactivity in ibrutinib-resistant hubs coincided with marked decreases or increases in chromatin opening (Fig. 7b, $p = 0.006$; and Fig. 7c, $p = 0.0009$) and chromatin activity (Fig. 7b, $p = 0.04$; and Fig. 7c, $p = 0.0004$), respectively, with a similar trend for gene expression levels (Fig. 7b, $p = 0.53$; and Fig. 7c, $p = 0.12$).

Close inspection of hubs with loss of interactivity in ibrutinib-resistant MCL showed the dismantling of enhancer-promoter assemblies at genes involved in regulating GTPase activity, suggestive of dysregulation of GTPase signaling downstream of BTK inhibition within the BTK signaling cascade (Supplementary Data 2)[52,53]. The hub containing tumor suppressor *PTPRG*[54] was the most differentially interacting hub with loss of interactivity in ibrutinib-resistant MCL (Fig. 7d). In line with loss of the *PTPRG* hub, we observed marked reductions in chromatin activity (Fig. 7d), Hi-C map contacts (Fig. 7f), and *PTPRG* expression levels at the *PTPRG* locus (Fig. 7e, $p = 0.007$).

In contrast, examination of hubs with gain of interactivity in ibrutinib-resistant MCL indicated that these topological assemblies predominantly formed at genes involved in processes promoting drug resistance, including cell cycle and apoptosis regulation, chromatin organization, stem cell maintenance, as well as regulation of protein transport (Supplementary Data 2). Indeed, the hub containing *BCL2L1* (also known as *BCL-xL*), a gene with proven anti-apoptotic function[55], gained over 700 interactions in ibrutinib resistance to become hyperinteracting (Fig. 7g, Supplementary Data 3). Increases in chromatin activity over this hyperinteracting hub coincided with a significant increase in *BCL2L1* transcriptional activity (Fig. 7h, $p = 0.00009$) and visible changes in the Hi-C contact map (Fig. 7i), suggestive of a structural variation (such as a translocation) at this locus in ibrutinib-resistant MCL (Supplementary Data 4). Other genes located in this hyperinteracting hub, including DNA damage response gene *TPX2* and DNA methyltransferase *DNMT3B*, were also markedly upregulated in ibrutinib resistance (Supplementary Fig. 7e, $p = 0.0004$; and Supplementary Fig. 7f, $p = 0.0001$). In line with observations in GSI-resistant T-ALL, this data further supports the role of enhancer-promoter hub reorganization in setting gene expression programs permissive to anticancer drug resistance.

Discussion

As an emerging unit of chromatin architecture, the spatial enhancer-promoter hub has remained a topological feature with cryptic structural and functional relevance. Here, we used a graph-based approach to systematically identify enhancer-promoter hubs from multi-omic data and examine their organizational principles and potential function in oncogenesis and drug resistance. By studying enhancer-promoter hubs in T-ALL, MCL, and TNBC, we found that these topological assemblies were distinct from TADs and ubiquitously enriched for transcription, with the most highly interacting hubs spatially coalescing genes encoding lineage-associated and oncogenic transcription factors and coregulators including *MYC*, *DOT1L*, *KAT5*, and *SOX9*, observations that may extend to other cancers.

Upon acquisition of Notch inhibitor resistance, a subset of hubs was reorganized as supported by optical mapping of the *LEF1* and *IKZF2* hubs in individual T-ALL cells. Our data further showed that differential hubs contained genes that were recognized as key promoters of anticancer drug resistance in T-ALL[22,56]. Similarly, application
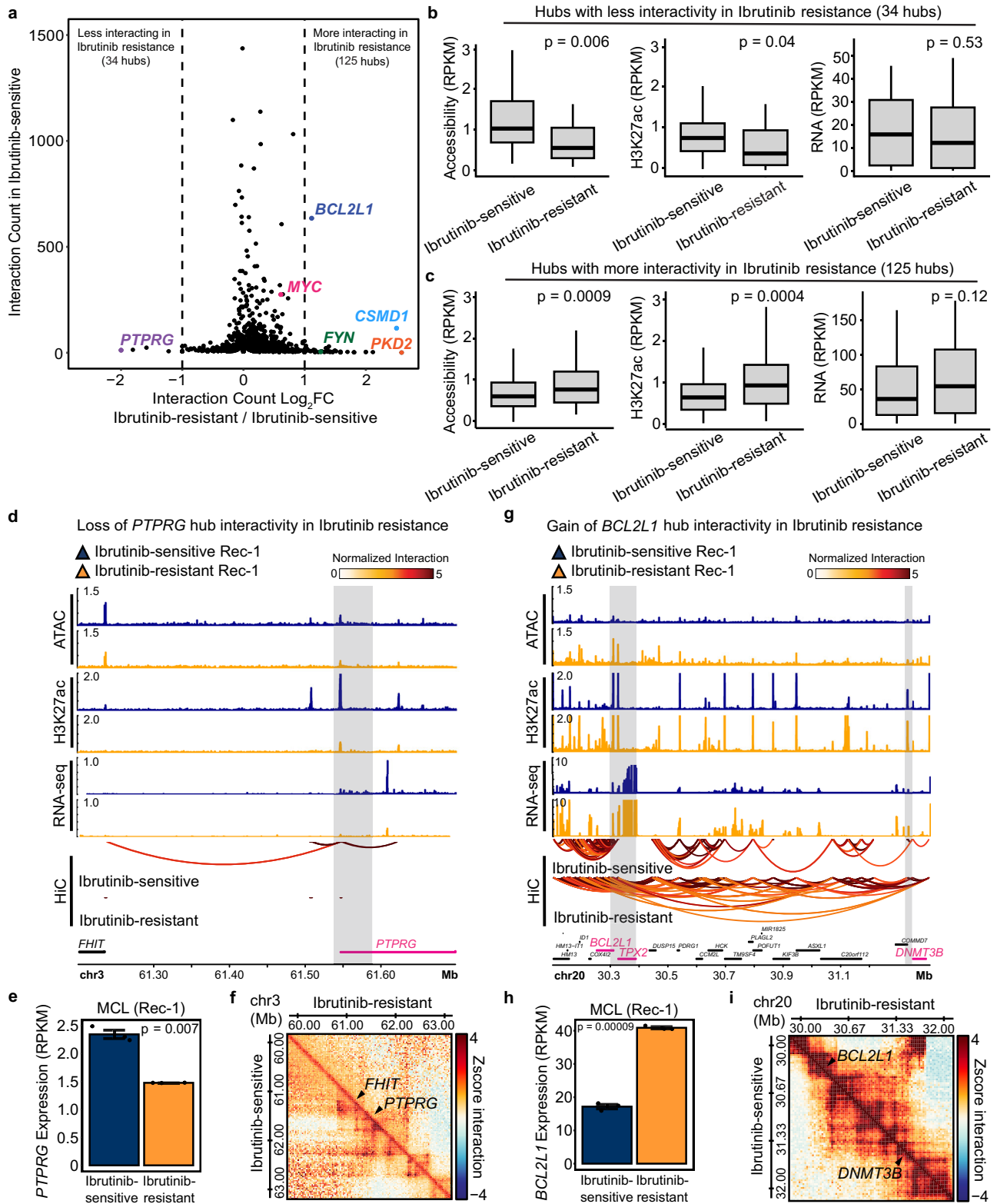
of our differential hub screen to ibrutinib-sensitive and -resistant MCL revealed a variety of differential hubs containing genes with potentially important roles in supporting the drug-resistant phenotype, including *BCL2L1*. Indeed, it has been shown that BCL2L1 inhibitors were able to induce potent cytotoxicity in MCL cells resistant to ibrutinib and Bcl-2 inhibitor venetoclax[57]. Although further investigation is required, our findings suggest that systematic examination of clusters of enhancers and promoters converging through space may guide identification of

therapeutic targets by revealing key regulatory elements and genes promoting drug resistance in cancer.

Our data also suggests that enhancer-promoter hubs uniquely straddle the structure-function axis. While the chain of causality remains elusive, our observations support a model of enhancer-promoter hubs in which the number of spatial interactions within a hub coincides with its relative transcriptional state. This model is reinforced by the correlation between transcriptional activity and

Fig. 7 | Genome-wide differential hub screen identifies loci that are aberrantly folded and expressed in ibrutinib-resistant MCL. a Scatter plot showing $\log_2$ fold change of hub interactivity in ibrutinib-resistant Rec-1 vs. hub interaction counts in ibrutinib-sensitive cells. Dotted lines mark enhancer-promoter hubs with ≥ 2-fold decrease ("less interacting in ibrutinib resistance") or increase ("more interacting in ibrutinib resistance") in interaction counts in ibrutinib-resistant vs. ibrutinib-sensitive Rec-1 cells. Selected hubs are labeled. **b** Box-and-whisker plots comparing ATAC-seq (left), H3K27ac ChIP-seq (center), and RNA-seq (right) of hubs with marked loss of interactivity in ibrutinib-resistant cells. Box-and-whisker plots: center line, median; box limits, upper (75th) and lower (25th) percentiles; whiskers, 1.53 interquartile range ($n = 34$). P-value: two-tailed Wilcoxon rank sum test. **c** Box-and-whisker plots comparing ATAC-seq (left), H3K27ac ChIP-seq (center), and RNA-seq (right) in hubs that gained interactivity in ibrutinib-resistant cells. Box-and-whisker plots: center line, median; box limits, upper (75th) and lower (25th) percentiles; whiskers, 1.53 interquartile range ($n = 125$). P-value: two-tailed Wilcoxon rank sum test. **d** Concordant changes in ATAC-seq, H3K27ac ChIP-seq, RNA-seq, and Hi-C over the *PTPRG* locus between ibrutinib-sensitive and ibrutinib-resistant Rec-1 cells. **e** *PTPRG* normalized RNA-seq reads in ibrutinib-sensitive and ibrutinib-resistant cells. Each dot represents a biological replicate ($n = 3$). P-value: two-sided t-test; error bars: ± SEM. **f** Normalized Hi-C contact maps in ibrutinib-resistant (upper triangle) and ibrutinib-sensitive (lower triangle) Rec-1 cells at the *PTPRG* hub locus. **g** Coordinated changes in ATAC-seq, H3K27ac ChIP-seq, RNA-seq, and Hi-C over the *BCL2L1* locus between ibrutinib-sensitive and ibrutinib-resistant Rec-1 cells. **h** *BCL2L1* normalized RNA-seq reads in ibrutinib-sensitive and ibrutinib-resistant cells. Each dot represents a biological replicate ($n = 3$). P-value: two-sided t-test; error bars: ± SEM. **i** Normalized Hi-C contact maps in ibrutinib-resistant (upper triangle) and -sensitive (lower triangle) Rec-1 cells at the *BCL2L1* hub.

enhancer positioning that has been documented at individual loci across various cancers[14,15,58,59]. It is likely that the mechanisms underlying enhancer-promoter hub formation also contribute to the distinct relationship between hub interactivity and transcription. Given the concordance of enhancer-promoter hubs identified from SMC1 HiChIP and Hi-C data, it is possible that cohesin-mediated loop extrusion may regulate hub spatial interactions and thus hub transcriptional activity, similar to cohesin's role in organizing multi-way contact 'hubs' in single cells[60]. In line with optical mapping performed in this study, further work is needed to concretely demonstrate the existence of multi-way interactions within hubs beyond correlative analysis of multi-omic data. Future studies leveraging advanced microscopy will refine our current understanding of hubs from population-based sequencing experiments and may shed light on hub landscapes within single cells. Taken together, our investigation suggests that enhancer-promoter hubs in cancer spatially organize transcriptional programs, which in turn may promote oncogenesis and drug resistance, parallel to enhancer-promoter hubs' broader role in directing gene expression circuits in other diseases[12,13,61,62].

Methods
Contact for reagent and resource sharing
Further information and request for reagents may be directed to and will be fulfilled by the corresponding author, Robert B. Faryabi (faryabi@pennmedicine.upenn.edu).

Experimental procedures
Cell culture. All of the data analyzed in this study for TNBC cell lines MB157 (ATCC, Cat# CRL-7721) and HCC1599 (ATCC, Cat# CRL-2331) and for T-ALL cell line CUTLL1[63] were taken from previous investigations (see subsequent section). For the purpose of the FISH experiments conducted in this study, DND41 (DSMZ, Cat# ACC525) GSI-sensitive and GSI-resistant cells were cultured and maintained as previously described[22]. Briefly, the DND41 cell line used for FISH analysis in this study was purchased from the Leibniz-Institute DSMZ-German Collection of Microorganisms and Cell Lines, and was grown in suspension with RPMI 1640 (Corning, cat# 10-040-CM) supplemented with 10% fetal bovine serum (Thermo Fisher Scientific, cat# SH30070.03), 2 mM L-glutamine (Corning, cat# 25-005-CI), 100 U/mL and 100 µg/mL penicillin/streptomycin (Corning, cat# 30-002-CI), 100 mM nonessential amino acids (GIBCO, cat# 11140-050), 1 mM sodium pyruvate (GIBCO, cat# 11360-070) and 0.1 mM of 2-mercaptoethanol (Sigma, cat# M6250). GSI-resistant DND41 cells were constantly cultured with GSI compound E (125 nM, Calbiochem, cat# 565790) and were periodically validated to have maintained the drug-resistant state by Western blotting for Notch intracellular domain 1 (NICD1), which is not present in cells constantly treated with GSI. Parental and resistant DND41 cells were used at a low passage number (<12) and subjected to regular (approximately every 6 months) mycoplasma testing and short tandem repeat (STR) profiling.

For experiments involving ibrutinib-sensitive and ibrutinib-resistant Rec-1 MCL, Rec-1 cells from the Genentech cell bank were cultured in RPMI 1640 (Corning, cat# 10-040-CM) supplemented with 10% fetal bovine serum (Thermo Fisher Scientific, cat# SH30070.03), 2 mM L-glutamine (Corning, cat# 25-005-CI), 100 U/mL and 100 µg/mL penicillin/streptomycin (Corning, cat# 30-002-CI), 100 mM nonessential amino acids (GIBCO, cat# 11140-050), 1 mM sodium pyruvate (GIBCO, cat# 11360-070) and 0.1 mM of 2-mercaptoethanol (Sigma, cat# M6250). Ibrutinib-resistant cells were generated over prolonged period of time using ibrutinib (Selleckchem, cat# S2680) dose escalation until they were stable in cell culture media supplemented with 100 nM ibrutinib. Ibrutinib resistance was confirmed by 100-fold shift for ibrutinib IC50 between parental and resistant Rec1 lines. Parental and resistant Rec-1 cells were used at a low passage number (< 12) and subjected to regular (approximately every 6 months) mycoplasma testing and short tandem repeat (STR) profiling. Parental and resistant cells, when used for RNA-seq, ATAC-seq, Hi-C, and ChIP-seq experiments, were treated with DMSO (parental) or ibrutinib (resistant) for 24 h following BCR crosslinking with IgM. Cells used for ChIP-seq and Hi-C assays were subsequently fixed with 1% or 2% formaldehyde, respectively. Fixed frozen cell pellets were stored at −80C and used when needed.

Multi-omic assays. In this study, we performed in-situ Hi-C, RNA-seq, H3K27ac ChIP-seq, and ATAC-seq on ibrutinib-sensitive and ibrutininb-resistant Rec-1 cells. Refer to the subsequent sections for descriptions on how these assays were performed. Excluding ibrutinib-sensitive and -resistant Rec-1 cells, multi-omic data for DND41, MB157, HCC1599, (untreated) Rec-1 cells, and CUTLL1 were obtained from previous investigations[14,22,29-31]. Specifically, in situ Hi-C, SMC1 HiChIP, ATAC-seq, RNA-seq, H3K27ac ChIP-seq, and H3K27me3 CUT&RUN data from DND41 GSI-sensitive and GSI-resistant cells were downloaded from GSE173872. SMC1 HiChIP data for MB157, HCC1599, and Rec-1 were downloaded from GSE116876 along with H3K27ac ChIP-seq data for GSI-washout MB157, Rec-1, and HCC1599 and RNA-seq data for GSI-washout MB157 and HCC1599. RNA-seq data for GSI-washout Rec-1 was downloaded from GSE59810. GSI-washout cells were treated with GSI (1 uM, Calbiochem) for 72 h before being washed and cultured in media containing DMSO for 5 h, at which time RNA-seq and ChIP-seq assays were run. For the purposes of this study, data collected from GSI-washout cells was considered equivalent to data collected from untreated cells for hub identification and analysis. For CUTLL1, in situ Hi-C and H3K27ac ChIP-seq data were downloaded from GSE115896, RNA-seq data was downloaded from GSE59810, and ATAC-seq data was downloaded from GSE216430.

Oligopaint FISH probe synthesis. DNA FISH Oligopaint probe libraries targeting three 50 Kb elements within the *LEF1* hub and three 75 Kb elements in the *IKZF2* hub were designed using OligoMiner[64]. Each of the six distinct probe sublibraries was amplified and isolated from a

larger pooled library using several short oligonucleotide primers (*RPL34* F primer: CTCGAATCGGTGTCGCATTC, R primer: TTGAC GTTTGCGCCGAATAC; *LEF1* promoter F primer: TCCGCCGTGTTATCG ATTTG, R primer: ATTCAACGGCCCTCGATTTG; *LEF1* enhancer F primer: TCATAATTCGGCGCTTGGTG, R primer: TGTATCGCGCGGTCA ATTTC; *IKZF2* promoter F primer: TCGCTACGCCGGTTGTAATG, R primer: ATTACCGCGACCGGTTGAAG; *IKZF2* 5′ F primer: CAGTTAC CGGTCCGTCGATG, R primer: ACGTATCGTCCCGCAACATG; *IKZF2* 3′ F primer: TTGTCGCGATGCCATAGACG, R primer: AGCTCAATCGTCGC ACGATC), as previously described[22]. Briefly, the pooled library was amplified via low cycle PCR and the T7 promoter was separately cloned into oligos within the six probe sublibraries of interest using primers specific to each probe sublibrary, which were purchased from IDT. Each probe was transcribed to RNA using a T7 RNA polymerase, and then RNA was reverse transcribed back into DNA oligos, which were purified and isolated for subsequent nuclear DNA hybridization. In order to visualize primary probes with fluorescence microscopy, secondary DNA probes conjugated with Alexa-488 (sequence: 5Alex488N/ CACACGCTCTTCCGTTCTATGCGACG

TCGGTG/3AlexF488N), Atto-565 (sequence: 5ATTO565N/ACACC CTTGCACGTCGTGGACCT

CCTGCGCTA/3ATTO565N), and Alexa-647 (sequence: 5Alex647N/ TGATCGACCACGGCCAA

GACGGAGAGCGTGTG/3AlexF647N) reporters, which were also purchased from IDT, were used.

3D Oligopaint DNA FISH on slides. Cells were prepared for DNA FISH as previously described[22]. Briefly, GSI-sensitive and GSI-resistant DND41 cells were first incubated on poly-*L*-lysine-coated glass slides (ThermoScientific, cat# P4981-001) and fixed in a solution of 4% formaldehyde in PBS for 10 min. Cell membranes were permeabilized by submerging slides in 0.5% Triton X-100 in PBS for 15 min and then denatured by immersion in increasing concentrations of ethanol (70%, 90%, and 100%). Cells were further permeabilized with cycles of immersion in heated 2× SSCT/50% formamide before incubation with a hybridization buffer containing 100 pmol of each of the three Oligopaint probes targeting the locus of interest. Slides were then sealed with rubber cement and a coverslip. After incubating in a 37 °C humidified chamber for 16 h, the coverslip and probe hybridization solution were removed from the slides, which were once again cyclically submerged in 2× SSCT and 0.2× SSCT in heated water baths to re-permeabilize cell and nuclear membranes. Another hybridization mix containing secondary probes conjugated to fluorophores was aliquoted onto each slide before slides were sealed with rubber cement and coverslips. After incubation in the 37 °C humidified chamber for 2 h, slides were submerged in 2× SSCT. DAPI dye was then added to stain nuclei, and slides were submerged in 2× SSCT for a final time. Finally, mounting media (Invitrogen, Ref# 336936) was added to each slide and coverslips were sealed onto the slides using transparent nail polish. Slides were then imaged using a Leica SP8 confocal microscope (*IKZF2* locus) with a 40× oil immersion objective or the Vutara VXL microscope (*LEF1* locus) (Bruker) in the widefield, epi-fluorescence microscope setting.

Rec-1 ibrutinib in-situ Hi-C. 10^6 parental or ibrutinib-resistant Rec-1 cells fixed with 2% formaldehyde were used as input for Hi-C assay performed with Arima Hi-C kit (Arima Genomics, cat#A510008) per the manufacturer's instructions. Both samples passed Arima QC1 and were subsequently used for library generation with Accel-NGS 2S Plus DNA Library Kit (Swift, cat# 21024) and 2S Set A Indexing Kit (Swift, cat# 26148). After passing Arima QC2, samples were PCR amplified for 6 cycles and quality was inspected using D5000 on Agilent 4200 Tapestation. Libraries were paired-end sequenced (2x150bp) using NovaSeq.

Rec-1 ibrutinib H3K27ac chromatin immunoprecipitation sequencing (ChIP-seq). ChIP-seq was performed using a previously published protocol[14]. 10×10^6 parental or ibrutinib-resistant Rec-1 cells previously fixed with 1% formaldehyde (Pierce, cat#28908) and quenched with 0.125 M Glycine (Fisher scientific, cat#AAJ1640736), were sonicated for 5.5 min on Covaris L220 with the following settings: PIP 350, DF 15%, CPB 200. Chromatin was cleared with recombinant protein G–conjugated agarose beads (Invitrogen, cat# 15920-010) and subsequently immunoprecipitated with H3K27ac antibody at 1:500 dilution (Active Motif, cat# 39133). Antibody-chromatin complexes were captured with recombinant protein G–conjugated agarose beads, washed with Low Salt Wash Buffer, High Salt Wash Buffer, LiCl Wash Buffer and TE buffer with 50 mM NaCl and eluted. Input sample was prepared by the same approach without immunoprecipitation. After reversal of crosslinking, RNase (Roche, cat# 10109169001) and Proteinase K (Invitrogen, cat# 25530-049) treatments were performed, and DNA was purified with QIAquick PCR Purification Kit (QIAGEN, cat# 28106). Libraries were then prepared using the NEBNext Ultra II DNA library Prep Kit for Illumina (NEB, cat# E7645S). Two replicates were performed for each condition. Indexed libraries were validated for quality and size distribution using a TapeStation 4200 (Agilent). Libraries were paired-end sequenced (2 × 50 bp) on Illumina NextSeq instrument.

Rec-1 ibrutinib assay for transposase accessible chromatin (ATAC-seq). ATAC-seq assay was performed as previously described[14]. Briefly, 60,000 parental or ibrutinib-resistant Rec-1 cells were washed with 50 μL of ice cold 1 × PBS (Corning, cat# 21031CV), followed by 2 min treatment with 50 μL lysis buffer containing 10 mM Tris-HCl, pH 7.4, 3 mM $MgCl_2$, 10 mM NaCl, and 0.1% NP-40 (Igepal CA-630). After pelleting nuclei, nuclei were resuspended in 50 μL of transposition buffer consisting of 25 μL of 2× TD buffer, 22.5 μL of molecular biology grade water, and 2.5 μL Tn5 transposase (Illumina, cat# FC-121-1030) to tag the accessible chromatin for 45 min at 37 °C. Tagmented DNA was purified with MinElute Reaction Cleanup Kit (QIAGEN, cat# 28204) and amplified with 5 cycles. Additional number of PCR cycles was determined from the side reaction and ranged from 8–9 total cycles of PCR. Two replicates were performed for each condition. Libraries were purified using QiaQuick PCR Purification Kit (QIAGEN, cat# 28106) and eluted in 20 μL EB buffer. Indexed libraries were assessed for nucleosome patterning on the TapeStation 4200 (Agilent) and paired-end sequenced (50 bp+50 bp) on HiSeq (Illumina).

Rec-1 ibrutinib RNA sequencing (RNA-seq). Strand-specific RNA-seq was performed on parental and ibrutinib-resistant Rec-1 cells using SMARTer Stranded Total RNA Sample Prep Kit (Takara, cat# 634873) per the manufacturer's instructions. Briefly, 24 h post treatment with DMSO and 100 nM ibrutinib followed by IgM stimulation for 5 min, parental and resistant cells were lysed with 350 μL RLT Plus buffer (QIAGEN) supplemented with 2-mercaptoethanol (Sigma, cat# M6250) and total RNA was isolated using the RNeasy Plus Micro Kit (QIAGEN, cat# 74034). RNA integrity numbers were determined using TapeStation 2200 (Agilent). 800 ng of total RNA was used, and libraries were prepared using the SMARTer Standard Total RNA Sample Prep Kit - HI Mammalian. Libraries were paired-end sequenced (38 + 38 bp) on a HiSeq. Three biological replicates were performed for each condition.

Quantification and statistical analysis
Definition of regulatory elements. For the purposes of this study, regions within 2.5 Kb of transcription start sites (TSSes) of expressed genes were considered to be promoters. H3K27ac ChIP-seq peaks that did not overlap with promoter regions were considered to be enhancers. For identification of hubs from Hi-C data, only enhancers and promoters that overlapped with ATAC-seq peaks were considered valid loop anchors.

Gene annotation. All genes presented in this study were annotated in accordance with the Human Genome version 19 (hg19) GRCh37.75 assembly. 2,828,317 Ensembl transcripts from this assembly were downloaded, and the longest transcript from each Ensembl gene id (ENSG) was used to generate a list of transcription start sites (for promoter annotation) and gene position. 57,209 gene annotations were used for RNA-seq analysis after excluding rRNA and chrM genes.

Rec-1 ibrutinib H3K27ac ChIP-seq and ATAC-seq data analysis. Rec-1 ibrutinib-sensitive and -resistant H3K27ac ChIP-seq reads were trimmed with Trim Galore (version 0.6.6) using parameters -q 5 --phred33 --fastqc --gzip --stringency 5 -e 0.1 --length 20 –paired. Ensembl GRCh37.75 primary assembly, which included chromosome 1-22, chrX, chrY, chrM and contigs was used for alignment of trimmed reads with BWA (version 0.7.13)[65]. BWA was run using the command bwa aln -q 5 -l 32 -k 2 -t 12 and paired-end reads were grouped using bwa sampe -P -o 1000000 -r. After grouping, reads which were considered duplicates from Picard (version 2.1.0) as well as reads that were matched with ENCODE blacklist regions or contigs were filtered out so that valid reads were kept and used for all subsequent analysis. This procedure was repeated for alignment of Rec-1 ibrutinib-sensitive and -resistant ATAC-seq reads. See the following section for ChIP-seq and ATAC-seq peak-calling protocols.

Rec-1 ibrutinib Hi-C data analysis. HiC-Pro (version v2.8.1)[66] was used to process Rec-1 ibrutinib-sensitive and -resistant Hi-C raw reads for each sample using default parameters except LIGATION_SITE and GENOME_FRAGMENT were provided by Arima. Putative interactions identified from HiC-Pro were used as the basis for hub calling. Similar to Rec-1, CUTLL1 Hi-C data was processed using HiC-Pro (version v2.8.1) with default parameters.

H3K27ac ChIP-seq and ATAC-seq peak calling. Peak calling of all H3K27ac ChIP-seq and ATAC-seq data was performed similar to as previously described[14]. Briefly, fragment length of H3K27ac ChIP-seq reads was estimated with HOMER (version 4.8)[67], and MACS was used to identify peaks with the parameters -q 1E-5 –shiftsize = 0.5∗fragment_length –format = BAM –bw = 300 –keep-dup = 1. After peak calling, H3K27ac signal over peaks was quantified and normalized to reads per kilobase per million mapped reads (RPKM). ATAC-seq reads were processed similarly to H3K27ac reads except MACS peak calling was performed with parameters -p 1E-5 –nomodel –nolambda –format = BAM –bw = 300 –keep-dup = 1. After peaks were called for each condition, BED files containing the union of peaks across relevant conditions (e.g. drug-sensitive, drug-treated, and drug-resistant) were created using the bedtools merge command. Note that for CUTLL1 H3K27ac ChIP-seq and ATAC-seq data, a slightly less stringent *q*-value cutoff of 1E-4 was used for peak calling to obtain a comparative number of peaks between cell lines. Exact steps for each workflow can be found at https://github.com/faryabiLab/dockerize-workflows/tree/master/workflows.

RNA-seq analysis. Bulk total transcript RNA-seq data from T-ALL DND41 and CUTLL1, MCL Rec-1 (including ibrutinib-sensitive and -resistant), and TNBC MB157 and HCC1599 cells was used to annotate expressed genes and analyze RNA enrichment over hub intervals. Alignment to Ensembl GRCh37.75 primary assembly and normalization with RPKM was performed as previously described[22]. RNA-seq for DND41, Rec-1, MB157, and HCC1599 was originally performed on at least three biological replicates for each cell line, and expressed genes were determined using a cutoff of RPKM > 1 in at least two of three replicates. For CUTLL1, a single RNA-seq experiment was analyzed with expressed genes determined using the same cutoff of RPKM > 1.

Topologically associating domain (TAD) boundary and differential TAD boundary identification. TAD boundaries were identified in both DND41 and Rec-1 drug-sensitive and drug-resistant cells from SMC1 HiChIP and Hi-C data as previously described[22]. Briefly, the cooltools insulation function (https://cooltools.readthedocs.io/en/latest/) was applied to a .cool file converted from HiC-Pro[66] valid pairs output using the hicpro2higlass.sh function followed by the hic2cool command-line tool with default options. Using a window size of 100 Kb and a bin size of 5 Kb, insulation scores were calculated. Valid, adjacent boundaries were defined as those with total reads between them exceeding the 75th percentile. Differential Hi-C TAD boundaries between drug-sensitive and drug-resistant conditions were categorized as boundaries with absolute log2 fold change in insulation score greater than 0.75.

Differential compartment identification. Hi-C data was used to detect differential compartments between drug-sensitive and drug-resistant DND41 and Rec-1. Compartments were initially identified in each condition from the first principle component (PC1) of Hi-C data, which was calculated with the Homer v4.11[67] runHiCpca.pl method using H3K27ac ChIP-seq data to avoid arbitrary sign assignment. Next, the getHiCcorrDiff command was used to calculate each compartment's correlation difference between drug-sensitive and drug-resistant conditions. To identify differential compartments switching from A to B in the drug-resistant state, the findHiCCompartments command was used with a correlation difference threshold (-corr) and a PC1 threshold of at least 0.4 for DND41 and 0.65 for Rec-1. For detection of compartments switching from B to A in the drug-resistant state, the same findHiCCompartments command was used with the addition of the -opp flag.

Enhancer-promoter hub identification and analysis
Spatial interaction pre-processing overview. Enhancer-promoter hubs in T-ALL DND41, T-ALL CUTLL1, MCL Rec-1, TNBC MB157, and TNBC HCC1599 were identified using Hi-C or SMC1 HiChIP contact frequency data annotated for enhancers and promoters. Specifically, SMC1 HiChIP data was used for detection of hubs in DND41, HCC1599, MB157, and Rec-1 cells whereas Hi-C data was used for detection of enhancer-promoter hubs in CUTLL1, GSI-sensitive and -resistant DND41 cells and ibrutinib-sensitive and -resistant Rec-1 cells. A high-level description of the data filtering process to create the inputs for our hub identification program is as follows: in order to generate data tables of valid enhancer-enhancer (EE), enhancer-promoter (EP), or promoter-promoter (PP) interactions in each condition, Hi-C/HiChIP loop anchors were first filtered by intersecting them with BED files of H3K27ac ChIP-seq peaks (i.e. enhancers) and active transcription start sites (TSS) from RNA-seq (i.e. promoters). For this analysis, peaks were defined as the 5 Kb or 10 Kb centered around the summit of the original protein enrichment signal. Anchors that were annotated as active TSSes with H3K27ac peaks were considered to be promoters. Once spatial interactions between putative promoter and/or enhancer elements were isolated, they were assigned normalized contact frequency scores and further filtered (using a user-defined score cutoff for Hi-C) to create the final data table for input into the hub-calling algorithm. This process of interaction filtering was slightly different depending upon the source assay of contact frequency data (i.e. Hi-C vs. SMC1 HiChIP) and upon other genomic data (e.g. ATAC-seq) that was readily available for analysis.

SMC1 HiChIP spatial interaction pre-processing. For SMC1 HiChIP data, valid EE/EP/PP input interactions into the hub-calling pipeline were detected from raw reads by filtering significant interactions identified from FitHiChIP v11.0[68] to only include those between enhancers and active promoters. Briefly, SMC1 HiChIP reads were

processed with Hi-C Pro version v2.5.0[66]. High-confidence loop anchors were identified from FitHiChIP using Hi-C Pro's allValidPairs file as input, a significance cutoff of $q = 0.05$ or $p = 0.05$ (see next paragraph), coverage bias regression for normalization, an interaction type of all to all / IntType=4, and default values for the remaining options. Anchors of significant FitHiChIP interactions were divided into two BED files and separately intersected with a BED file containing the union of H3K27ac ChIP-seq peaks and actively transcribed TSSes, where active genes were defined as those with >1 RNA-seq RPKM in at least two thirds of the replicates across the conditions of interest. After valid promoter and enhancer anchors were identified, the original list of FitHiChIP interactions was filtered to keep only interactions between valid enhancer and/or promoter elements for input to the hub-calling pipeline.

For filtering interactions from Rec-1, MB157, and HCC1599 with FitHiChIP, a significance threshold of $q < 0.05$ was used to yield over 100,000 significant interactions in each cell type. However, since FitHiChIP identified less than 25,000 significant interactions from DND41 SMC1 HiChIP data with a significance cutoff of $q < 0.05$, we decided that a significance threshold of $p < 0.05$ was more appropriate for filtering interactions in this cell line. To ensure stringent filtering of DND41 SMC1 HiChIP data given this lower threshold, FitHiChIP significant interactions ($p < 0.05$) for DND41 were subject to further filtering. Specifically, BED files containing anchors of significant FitHiChIP interactions, ATAC-seq peaks, SMC1 ChIP-seq peaks, and the union of H3K27ac ChIP-seq peaks and actively transcribed TSSes were intersected to create a list of high-confidence, accessible regulatory element anchors. Only significant FitHiChIP interactions between two of these anchors were considered valid and used for DND41 hub calling. Since enhancers were not filtered by accessibility in Rec-1, MB157, and HCC1599 cells, they were defined more stringently as H3K27ac ChIP-seq peaks of > 500 bp.

In-situ Hi-C spatial interaction pre-processing. For Hi-C data, valid EE/EP/PP interactions were filtered from Hi-C reads by quantifying the number of contacts between accessible enhancers and accessible, actively expressed promoters. First, a data table of all possible *cis* interactions between accessible enhancers and accessible, active promoters was generated for the conditions of interest. BED files containing active, accessible promoters were created by intersecting a BED file of actively transcribed TSSes, where active genes were defined as those with >1 RNA-seq RPKM in 2 out of 3 replicates for either drug-sensitive or drug-resistant cells, with a BED file containing the union of ATAC-seq peaks from drug-sensitive and drug-resistant conditions. BED files of accessible enhancers were created by intersecting a BED file containing the union of H3K27ac ChIP-seq peaks from drug-sensitive and drug-resistant cells with a BED file containing the union of ATAC-seq peaks from drug-sensitive and drug-resistant cells. A matrix containing all possible combinations of *cis* linkages (maximum length of 2 Mb) between these accessible enhancer and promoter elements was then constructed and used to normalize Hi-C reads for a given condition. Specifically, Hi-C reads were processed with Hi-C Pro[66], rearranged into BEDPE file format, and intersected with the aforementioned BED file of ATAC-seq union peaks such that only interactions between pairs of accessible anchors were kept. These contacts were then mapped onto the matrix containing all possible enhancer and promoter interactions, and the number of ATAC-filtered Hi-C reads overlapping with each possible EE/EP/PP interaction was counted. Next, these linkage counts were summed and normalized to contacts per 100 million, and interactions with a normalized contact frequency of >3 (for DND41 and Rec-1) or >5 (for CUTLL1) were considered valid interactions for hub calling.

Hub-calling pipeline. We adapted an implementation of matrix-free divisive hierarchical spectral clustering (https://github.com/ faryabiLab/hierarchical-spectral-clustering), which was originally developed for single cell RNA-seq[21], to identify enhancer-promoter hubs from the enhancer-promoter connectivity graph. We reasoned that this approach can overcome limitations associated with heuristic global optimization-based community detection methods such as Louvain-based algorithms to efficiently identify enhancer-promoter hubs, which we define as groups of densely connected enhancers and promoters with high intra-group and sparse inter-group interactions. Briefly, our matrix-free divisive hierarchical spectral clustering uses all of the information embedded in the enhancer-promoter connectivity graph at each partitioning to create a tree of enhancer and/or promoter clusters by recursively bi-partitioning the input spatial interactions between genomic elements. Time and memory efficiency is achieved by replacing factorization of the normalized Laplacian matrix at each iteration with direct calculation of the second left singular vector corresponding to the second largest singular value of a new matrix derived from the sparse connectivity matrix[21]. To enable simultaneous detection of large and small interacting enhancer/promoter clusters and to avoid creating arbitrary small clusters, our approach uses Newman-Girvan modularity[69] as a stopping criterion for recursive cluster bi-partitioning. Using modularity as a stopping criterion instead of an optimization parameter also bypasses limitations associated with heuristic global optimization-based clustering such as Louvain-based algorithms[70,71]. To this end, our approach produces a hierarchy of nested enhancer-promoter clusters where each inner node is a cluster at a given scale and each leaf node is the finest-grain cluster such that any additional partitioning would be as good as randomly splitting connected enhancers/promoters. Importantly, this divisive hierarchical spectral clustering approach maintains relationships among clusters at various levels. As a result, the nested cluster structure can be used for clustering tree pruning, which provides flexibility in analysis and interpretation of enhancer-promoter hubs when hub topologies are a priori unclear.

To prevent the possibility that a functionally relevant cluster is divided into multiple clusters for downstream analysis, we used parameters cluster-tree -c dense -n 2 -s for all the analysis, and defined hubs as self-contained networks of connected regulatory elements. We next categorized hubs by the largest contiguous genomic interval covered by their enhancer/promoter anchors, and computed their within-hub spatial interaction counts and enhancer/promoter counts. We also annotated hubs with the expressed genes contained within their contiguous intervals (Supplementary Data 1). To ensure that we identified hubs rather than just a few interacting elements, we removed clusters containing fewer than 6 spatial interactions among less than 4 enhancer/promoters for downstream analysis. For each cell type, distributions of hubs were plotted on the basis of their interaction count and enhancer/promoter count. The interaction count cutoff for hyperinteracting hubs was determined by calculating the point of tangency on the elbow of the curve comparing hub rank to interaction count as previously described[14].

Hub vs. random loci RNA enrichment analysis. Scatterplots of median RNA enrichment over hub loci and random representative hub loci were created for each cancer type to analyze transcriptional activity of hubs. Lists of random representative hub loci were generated by iteratively calling the bedtools shuffle command with parameters -chrom -noOverlapping -g hg19.genome on input BED files containing observed hubs. This bedtools command generates a list of random loci that contains an identical number of loci as the input list of observed hubs, where each random locus on the output list has identical length and chromosomal localization as each observed hub in the input list. As noted in the bedtools documentation, the process of generating random representative loci using this command could very infrequently skip loci violating the noOverlapping criteria. Median RNA enrichment was then quantified over each list of random

representative loci to create a single data point on the output scatterplot, and this process was repeated several thousand times (5000 times for all total hub analyzes and 10,000 times for all hyperinteracting hub analyzes) to enhance the accuracy of the simulated comparison. Additionally, comparisons of transcription levels between regular and hyperinteracting hubs were performed without genomic span normalization.

SMC1 HiChIP vs. Hi-C hub similarity Venn diagrams. In order to compare DND41 and Rec-1 hubs identified from Hi-C with DND41 and Rec-1 hubs identified from SMC1 HiChIP (respectively), Venn diagrams were used to depict the similarity of hubs on the basis of their genomic overlap. These diagrams indicate the percentage of distinct, non-overlapping HiChIP hubs (relative to the total number of HiChIP hubs) and Hi-C hubs (relative to the total number of Hi-C hubs) as well as the percentages of overlapping HiChIP and Hi-C hubs reported both in terms of the total number of HiChIP hubs and the total number of Hi-C hubs. Note that the sizes of the Venn diagram circles are proportional to the number of hubs between those conditions.

All hub and hyperinteracting hub similarity matrices. Similarity matrices of all hubs and hyperinteracting hubs identified from Rec-1, HCC1599, MB157, and DND41 SMC1 HiChIP were created by comparing the genomic positions of hubs between each pair of cancer types. For example, to compare SMC1 HiChIP hyperinteracting hubs between DND41 and Rec-1 cells, the number of non-overlapping (where overlap is defined as $\geq 1\,bp$) hyperinteracting hubs in each cancer was calculated by using bedtools intersect $-v$ on two input BED files of hyperinteracting hubs in each cancer type. The similarity score for the comparison was then calculated using the following formula: similarity = 100 − (average percentage of distinct hubs). Since no more than 5% of within-condition hubs or hyperinteracting hubs for SMC1 HiChIP were overlapping, this methodology of similarity analysis was considered to be minimally biased.

The ten shared (i.e. overlapping) hyperinteracting hubs across T-ALL DND41, MCL Rec-1, and TNBC MB157 and HCC1599 were identified by finding the intersection of the four SMC1 HiChIP hyperinteracting hub lists using bedtools intersect, and manually filtering the resulting overlapping intervals based upon their connectivity structure and size. Specifically, overlapping regions were only considered to be conserved hyperinteracting hubs if there were valid SMC1 HiChIP contacts over the overlapping region in all four cell lines and if the region spanned more than ~75 Kb (i.e. size of the smallest hyperinteracting hub across the four cell types).

TAD and hub boundary similarity analysis. Hub boundaries were defined as the 5 Kb upstream and downstream from the start position and end position of hubs (respectively) resulting in 2 h intervals, where h is the number of hubs identified. Valid TAD boundaries from DND41 and Rec-1 SMC1 HiChIP or Hi-C data were identified as described in the previous section and intersected with the aforementioned HiChIP or Hi-C hub boundaries (respectively) to produce Venn diagrams depicting hub/TAD boundary overlap. Additionally, pile-up plots of Hi-C hub and Hi-C TAD boundaries were separately generated using coolpup.py with options −pad 250000 −local and graphed with plotpup.py[72]. Note that few boundaries in each condition were automatically excluded from pile-up plots by coolpup.py.

Hyperinteracting vs. regular hub overlap with super-enhancers. Super-enhancers were identified in each cell line by applying previously described methods[73] to H3K27ac ChIP-seq peaks for CUTLL1, DND41, Rec-1, MB157, and HCC1599 in R. Super-enhancers were intersected with hub genomic locations and stratified by type (hyperinteracting or regular). Super-enhancer and hub similarity was

determined by calculating the proportion of regular and hyperinteracting hubs overlapping with 0, 1, or 2+ super-enhancers.

Hyperinteracting vs. regular hub genomic length violin plots. Genomic lengths of hubs were calculated as the difference between the start coordinate of the farthest upstream regulatory element in the hub and the stop coordinate of the farthest downstream regulatory element in the hub for DND41, Rec-1, MB157, and HCC1599 hubs identified from SMC1 HiChIP data. These genomic distances were stratified by hub type (i.e. hyperinteracting vs. regular) and plotted in R using ggplot2's geom_violin function.

Hyperinteracting vs. regular hub highly expressed gene plots. Highly expressed genes were considered to be genes within the top 2.5% quantile of gene expression, as defined from RNA RPKM in DND41, Rec-1, MB157, and HCC1599. The genomic coordinates of these highly expressed genes were intersected with hyperinteracting and regular hubs, and the number of highly expressed genes overlapping each hyperinteracting or regular hub was plotted in R using ggplot2's geom_violin function.

Identification of structural variants in MCL. Structural variations in ibrutinib-sensitive and ibrutinib-resistant Rec-1 were identified from ICE-balanced Hi-C data using the predicts command of EagleC v0.1.9 with default options[74]. Hi-C matrices were prepared as .cool files with resolutions 10 Kb, 50 Kb, and 100 Kb, as required by the EagleC pipeline.

Differential Enhancer-Promoter Hub Identification and Analysis
Differential hub-calling pipeline. Differential enhancer-promoter hubs were identified from the hub-calling output for two different conditions (i.e. two BED-like files containing hub genomic intervals and interaction counts). In order to compare hubs in each condition on the basis of interaction count, the bedtools merge command was used to find the union of hubs across the two conditions. For union hubs that were created from ≥ 3 overlapping precursor hubs, new interaction counts were calculated by summing the number of interactions of the individual precursor hubs in each condition. The union hub lists also contained de novo lost/gained non-overlapping hubs. Finally, the $\log_2$ fold change in interaction count over union hubs was calculated, where $\log_2$ fold change = $\log_2$(case interaction count) − $\log_2$(control interaction count), and union hubs were annotated by the expressed genes contained within their contiguous genomic intervals.

Differential hub analysis. For each set of union hubs, a volcano-like plot comparing the drug-sensitive interaction counts and $\log_2$ fold change in interaction count of each union hub was graphed. These plots were used to illustrate genome-wide changes in hub interactivity, and to highlight large differential hubs of interest for downstream analysis. Hubs were considered to become more interacting in the drug-resistant state either if $\log_2$FC(interaction count) ≥ 1 or if they increased from <6 interactions to ≥ 6 in the drug-resistant condition (i.e. de novo gained hubs). Similarly, hubs were considered to become less interacting in the drug-resistant state either if $\log_2$FC(interaction count) ≤ -1 or if they decreased from ≥ 6 interactions to <6 interactions in the drug-resistant condition (i.e. lost hubs). In order to best visualize all of the hubs in the volcano-like plot for DND41 GSI-sensitive vs. GSI-resistant hubs, the $\log_2$ fold change of all hubs in the plot was adjusted using a pseudocount such that $\log_2$ fold change = $\log_2$[(GSI-resistant interaction count + 1)/(GSI-sensitive interaction count + 1)]. This pseudocount adjustment was not necessary to create the volcano-like plot of differential hubs in ibrutinib-sensitive vs. ibrutinib-resistant Rec-1. Instead, for visual clarity, the two outlier hubs with >1500 interaction counts in ibrutinib-sensitive and -resistant cells were excluded from the plot. RNA enrichment, H3K27ac enrichment, and ATAC-seq enrichment were then examined

over these differential hubs. Gene ontology and pathway enrichment analysis over the expressed genes contained within differential hubs was also performed.

Differential compartment and differential hub similarity Venn diagrams. Differential A to B and B to A compartments were identified as described in the previous section and intersected with differential hubs that lost or gained interactivity in the drug-resistant condition, respectively. For similarity analysis, the overlap between the full genomic length of differential hubs and differential compartments was calculated and plotted with Venn diagrams.

Differential TAD and differential hub boundary similarity Venn diagrams. Differential TAD boundaries losing insulation or gaining insulation were identified as described in the previous section and intersected with boundaries of differential hubs that lost or gained interactivity in the drug-resistant condition, respectively. For similarity analysis, the overlap between the differential hub and differential TAD boundaries was calculated and plotted with Venn diagrams.

Genomic feature intersection/similarity analyzes

Intersection analysis to determine the linear genomic overlap between distinct populations of enhancer-promoter hubs and between hubs and other genomic features (e.g. TADs, compartments, super-enhancers, etc.) was performed using either the intersect command of bedtools[75] or the findOverlaps command of R library GenomicRanges. Unless otherwise noted, for hub populations and/or genomic features that overlapped such that one hub in list A intersected more than one hub/feature in list B (or vice versa), bedtools was used to determine the degree of overlap between the two lists. For genomic features and hubs that overlapped such that at most one hub in list A intersected one feature in list B (and vice versa), GenomicRanges was used to determine the degree of overlap between the two lists.

All Venn diagrams illustrating the output of feature intersection analysis were created in R using the eulerr package such that each segment's size was proportional to the number of features that it represented. For Venn diagrams illustrating intersection analyzes that had at least one feature in one list intersecting more than one feature in the other, the sizes of the diagram segments were set such that they reflected the number of exclusive features in each set and the average degree of overlap, where average overlap = (total number features - total number exclusive features)*0.5. All Venn diagram segments were manually labeled with the actual number and percentage of exclusive and overlapping features in each list.

Gene ontology (GO) and pathway enrichment analysis

GO enrichment analysis was conducted using the enrichment function of Metascape[76] or the PANTHER enrichment function of the Gene Ontology Knowledgebase[77–79]. For hyperinteracting hubs in each condition, the list of expressed genes contained within the intervals of hyperinteracting hubs was used as input for analysis of GO molecular function, biological process, and/or cellular component annotation enrichment. Metascape was used for GO enrichment analysis of SMC1 HiChIP hyperinteracting hubs in Rec-1, HCC1599, and MB157 while PANTHER was used for GO enrichment analysis of SMC1 HiChIP hyperinteracting hubs in DND41 and Hi-C hyperinteracting hubs in CUTLL1 because the number of expressed genes in T-ALL hyperinteracting hubs exceeded the gene limit for Metascape. Metascape was also used for GO enrichment analysis of DND41 and Rec-1 hyperinteracting hubs identified from Hi-C data. For GO enrichment analysis of differential DND41 and Rec-1 hubs, expressed genes in drug-resistant cells that were located within hubs that gained interactivity or expressed genes in drug-sensitive cells that were located within hubs that lost interactivity were used as inputs to Metascape. In addition to GO analysis, pathway enrichment analysis was also performed on these

expressed genes within differential hubs in Metascape using the Hallmark Gene Set, Reactome Gene Set, Biocarta Gene Set, Canonical Pathway, WikiPathway, and KEGG Pathway annotations.

For discussion of Metascape GO/pathway enrichment output in this paper, significant 'summary' annotations ($p < 0.01$) were presented (as opposed to significant 'non-summary' annotations) given that the 'summary' annotations represented overarching groups of GO terms and therefore limited redundancy in annotation reporting. Reported p-values from Metascape analyzes were computed using hypergeometric tests. For all Metascape hyperinteracting hub GO plots (Figs. 3g, 4g, and S4m), the top 10 most significantly overrepresented (i.e. enriched) 'summary' GO molecular function annotations from expressed genes in SMC1 HiChIP hyperinteracting hubs were ranked by p-value and plotted. For presentation of PANTHER GO enrichment output, we plotted the top 10 most significantly overrepresented (i.e. enriched) GO molecular function annotations ranked by p value, with a comprehensive list of significant GO annotations (including both underrepresented and overrepresented GO terms) contained in Supplementary Data 2. We also excluded the significantly underrepresented 'molecular function' GO term from Figs. 2g and S4h because this annotation could lead to confusion between underrepresented and overrepresented terms. Also note that the 'binding' GO term in these figures is not immediately relevant for direct gene annotation given its breadth as a parent GO annotation. Reported p-values from PANTHER analyzes were computed using Fisher's exact test.

DNA FISH analysis

DNA FISH image analysis was performed similar to a previously described protocol[22]. Briefly, DAPI signal was used for manual nuclei segmentation, with 1548 GSI-sensitive allelic interactions and 533 GSI-resistant allelic interactions analyzed for the *IKZF2* locus and 1319 GSI-sensitive allelic interactions and 640 GSI-resistant allelic interactions analyzed for the *LEF1* locus. For each manually segmented nucleus, spots indicative of probe signal were manually thresholded. Centroid positions for each spot in xy were found by fitting a Gaussian. X, Y, and Z coordinates were extracted, and pairwise Euclidean distances between nearest neighbors were calculated. Representative FISH cell images were taken using a Leica SP8 63× oil immersion objective and photo brightness, contrast, and smoothing was adjusted in ImageJ to facilitate probe visualization.

Genomic data visualization

Normalized reads from ATAC-seq, H3K27ac ChIP-seq, H3K27me3 CUT&RUN, RNA-seq, SMC1 HiChIP, and/or Hi-C over selected hubs were visualized using the R package Sushi (version 1.18.0)[80]. Specifically, bedgraph (.bg) files of reads normalized using reads per million from ATAC-seq, H3K27ac ChIP-seq, H3K27me3 CUT&RUN, and RNA-seq data were created with command genomecov, or by converting RPM-normalized BigWig (.bw) files into.bg format using the UCSC tools (version 329) BigWigToBedGraph[81]. These .bg files were visualized using the Sushi command plotBedgraph. For Hi-C data, normalized EE/EP/PP interactions exceeding the contact frequency cutoff of 3 (for DND41 and Rec-1) or 5 (for CUTLL1) were plotted using the Sushi command plotBedpe with intensity equivalent to normalized contact frequency. For SMC1 HiChIP data, EE/EP/PP FitHiChIP significant interactions ($p < 0.05$ for DND41 and $q < 0.05$ for Rec-1, MB157, and HCC1599) were plotted with intensity equivalent to $-\log_{10}(X)$, where X is the minimum FitHiChIP p-value (for DND41) or q-value (for Rec-1, MB157, and HCC1599) for the interaction. Z-scored contact maps were created to visualize Hi-C interaction heatmaps over selected hubs. These plots were generated by first applying z-score transformation to 25Kb resolution VC_SQRT normalized Hi-C contact maps for each chromosome as previously described[82] using the R command loessFit with parameters iter = 100, span = 0.02. The Sushi command plotHiC was used to plot these transformed maps.

Data presentation & statistical analysis

All analysis and quantification of ATAC-seq and ChIP-seq output was performed using peak-called data with normalized RPKM measurements. Median RNA enrichment over observed spatial hubs and hyperinteracting hubs was compared to median RNA enrichment over random representative loci using empirical p-values conservatively calculated as $(n+1)/(r+1)$, where r is the total number of replicates of random representative loci lists (i.e. 5000 or 10,000) and n is the number of replicates in which theoretical median RNA enrichment exceeded observed median RNA enrichment over hubs/hyperinteracting hubs. Two-tailed Wilcoxon Rank Sum tests were used for comparisons of regular hubs and hyperinteracting hubs and for differential hub RNA-seq/ATAC-seq/H3K27ac ChIP-seq enrichment analyzes with the wilcox.test function in R (version 4.2.1), and plot axes were abridged for the purpose of visualization. Comparisons of RNA-seq differential gene enrichment between DND41/Rec-1 drug-sensitive/drug-resistant cells for selected genes were evaluated with two-sided Student's t-tests ($n = 3$). Finally, statistical values for the comparison of probe distances in cumulative distribution plots were calculated using a Kolmogorov–Smirnov test. All p values less than 1E-5 were rounded up one decimal place (e.g. $p = 8E-7$ becomes $p = 1E-6$) for reporting in the text and figures. Unless otherwise noted, all box and whisker plots are shown without outliers to facilitate data visualization.

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

The RNA-seq, ChIP-seq, ATAC-seq, and HiC data generated in this study have been deposited in the GEO database under accession code GSE268228. The RNA-seq, ChIP-seq, The RNA-seq, ChIP-seq, CUT&RUN, ATAC-seq, HiC, and SMC1 HiChIP publicly available data used in this study are available in the GEO database under accession codes GSE173872[22], GSE116876[14], GSE59810[30], GSE115896[29], GSE216430[31]. The remaining processed and / or raw data are available within the Article, Supplementary Information or Source Data file. Source data are provided with this paper.

Code availability

All the code needed to evaluate the conclusions in the paper is publicly available at https://github.com/faryabiLab/hierarchical-spectral-clustering, https://github.com/faryabiLab/dockerize-workflows, and https://github.com/faryabiLab/enhancer-promoter-hub.

References

1. Finn, E. H. et al. Extensive heterogeneity and intrinsic variation in spatial genome organization. *Cell* **176**, 1502–1515.e1510 (2019).
2. Roukos, V. et al. Spatial dynamics of chromosome translocations in living cells. *Science* **341**, 660–664 (2013).
3. Zhao, J. & Faryabi, R. B. Spatial promoter-enhancer hubs in cancer: organization, regulation, and function. *Trends Cancer*, (2023).
4. Uyehara C. M., Apostolou E. 3D enhancer-promoter interactions and multi-connected hubs: Organizational principles and functional roles. *Cell Rep*, 112068 (2023).
5. Suva, M. L., Riggi, N. & Bernstein, B. E. Epigenetic reprogramming in cancer. *Science* **339**, 1567–1570 (2013).
6. Schoenfelder, S. & Fraser, P. Long-range enhancer-promoter contacts in gene expression control. *Nat. Rev. Genet* **20**, 437–455 (2019).
7. Rao, S. S. et al. A 3D map of the human genome at kilobase resolution reveals principles of chromatin looping. *Cell* **159**, 1665–1680 (2014).
8. Nuebler, J., Fudenberg, G., Imakaev, M., Abdennur, N. & Mirny, L. A. Chromatin organization by an interplay of loop extrusion and compartmental segregation. *Proc. Natl Acad. Sci. USA* **115**, E6697–E6706 (2018).
9. Fulco, C. P. et al. Activity-by-contact model of enhancer-promoter regulation from thousands of CRISPR perturbations. *Nat. Genet* **51**, 1664–1669 (2019).
10. Grubert, F. et al. Genetic control of chromatin states in humans involves local and distal chromosomal interactions. *Cell* **162**, 1051–1065 (2015).
11. Frankel, N. et al. Phenotypic robustness conferred by apparently redundant transcriptional enhancers. *Nature* **466**, 490–493 (2010).
12. Chandra A. et al. Quantitative control of Ets1 dosage by a multi-enhancer hub promotes Th1 cell differentiation and protects from allergic inflammation. *Immunity*, (2023).
13. Fasolino, M. et al. Genetic variation in type 1 diabetes reconfigures the 3d chromatin organization of T cells and alters gene expression. *Immunity* **52**, 257–274.e211 (2020).
14. Petrovic, J. et al. Oncogenic notch promotes long-range regulatory interactions within hyperconnected 3D cliques. *Mol. Cell* **73**, 1174–1190e1112 (2019).
15. Rhie, S. K. et al. A high-resolution 3D epigenomic map reveals insights into the creation of the prostate cancer transcriptome. *Nat. Commun.* **10**, 4154 (2019).
16. Liu, Y. et al. Chromatin looping shapes KLF5-dependent transcriptional programs in human epithelial cancers. *Cancer Res* **80**, 5464–5477 (2020).
17. Yang, L. et al. 3D genome alterations associated with dysregulated HOXA13 expression in high-risk T-lineage acute lymphoblastic leukemia. *Nat. Commun.* **12**, 3708 (2021).
18. Zhao, J. et al. Oncogenic transcription factors instruct promoter-enhancer hubs in individual triple negative breast cancer cells. *Sci. Adv.* **10**, eadl4043 (2024).
19. Chen, L. F. et al. Structural elements promote architectural stripe formation and facilitate ultra-long-range gene regulation at a human disease locus. *Mol. Cell* **83**, 1446–1461.e1446 (2023).
20. Phanstiel, D. H. et al. Static and dynamic DNA loops form AP-1-bound activation hubs during macrophage development. *Mol. Cell* **67**, 1037–1048.e1036 (2017).
21. Schwartz, G. W. et al. TooManyCells identifies and visualizes relationships of single-cell clades. *Nat. Methods* **17**, 405–413 (2020).
22. Zhou, Y. et al. EBF1 nuclear repositioning instructs chromatin refolding to promote therapy resistance in T leukemic cells. *Mol. Cell* **82**, 1003–1020.e1015 (2022).
23. Dixon, J. R. et al. Topological domains in mammalian genomes identified by analysis of chromatin interactions. *Nature* **485**, 376–380 (2012).
24. Yamagata, K., Shino, M., Aikawa, Y., Fujita, S. & Kitabayashi, I. Tip60 activates Hoxa9 and Meis1 expression through acetylation of H2A.Z, promoting MLL-AF10 and MLL-ENL acute myeloid leukemia. *Leukemia* **35**, 2840–2853 (2021).
25. Hayette, S. et al. High DNA methyltransferase DNMT3B levels: a poor prognostic marker in acute myeloid leukemia. *PLoS One* **7**, e51527 (2012).
26. Wang, Z. Y. & Chen, Z. Acute promyelocytic leukemia: from highly fatal to highly curable. *Blood* **111**, 2505–2515 (2008).
27. Izutsu, K. et al. An open-label, single-arm phase 2 trial of valemetostat for relapsed or refractory adult T-cell leukemia/lymphoma. *Blood* **141**, 1159–1168 (2023).
28. Gruhn, B. et al. The expression of histone deacetylase 4 is associated with prednisone poor-response in childhood acute lymphoblastic leukemia. *Leuk. Res* **37**, 1200–1207 (2013).
29. Kloetgen, A. et al. Three-dimensional chromatin landscapes in T cell acute lymphoblastic leukemia. *Nat. Genet* **52**, 388–400 (2020).
30. Stoeck, A. et al. Discovery of biomarkers predictive of GSI response in triple-negative breast cancer and adenoid cystic carcinoma. *Cancer Discov.* **4**, 1154–1167 (2014).

31. Tan, J. et al. Cell-type-specific prediction of 3D chromatin organization enables high-throughput in silico genetic screening. *Nat. Biotechnol.* **41**, 1140–1150 (2023).

32. Bernt, K. M. et al. MLL-rearranged leukemia is dependent on aberrant H3K79 methylation by DOT1L. *Cancer Cell* **20**, 66–78 (2011).

33. Bernt, K. M. & Armstrong, S. A. Targeting epigenetic programs in MLL-rearranged leukemias. *Hematol. Am. Soc. Hematol. Educ. Program* **2011**, 354–360 (2011).

34. Cobaleda, C., Schebesta, A., Delogu, A. & Busslinger, M. Pax5: the guardian of B cell identity and function. *Nat. Immunol.* **8**, 463–470 (2007).

35. Mumbach, M. R. et al. HiChIP: efficient and sensitive analysis of protein-directed genome architecture. *Nat. Methods* **13**, 919–922 (2016).

36. Christin, J. R. et al. Stem cell determinant SOX9 promotes lineage plasticity and progression in basal-like breast cancer. *Cell Rep.* **31**, 107742 (2020).

37. Guo, W. et al. Slug and Sox9 cooperatively determine the mammary stem cell state. *Cell* **148**, 1015–1028 (2012).

38. Ma, Y. et al. SOX9 is essential for triple-negative breast cancer cell survival and metastasis. *Mol. Cancer Res.* **18**, 1825–1838 (2020).

39. Witwicki, R. M. et al. TRPS1 is a lineage-specific transcriptional dependency in breast cancer. *Cell Rep.* **25**, 1255–1267.e1255 (2018).

40. Ai, D. et al. TRPS1: a highly sensitive and specific marker for breast carcinoma, especially for triple-negative breast cancer. *Mod. Pathol.* **34**, 710–719 (2021).

41. Tan, J. et al. CBX8, a polycomb group protein, is essential for MLL-AF9-induced leukemogenesis. *Cancer Cell* **20**, 563–575 (2011).

42. Sterken, B. A. et al. C/EBPbeta isoform-specific regulation of migration and invasion in triple-negative breast cancer cells. *NPJ Breast Cancer* **8**, 11 (2022).

43. Andrews, P. G. P., Popadiuk, C., Belbin, T. J. & Kao, K. R. Augmentation of Myc-dependent mitotic gene expression by the pygopus2 chromatin effector. *Cell Rep.* **23**, 1516–1529 (2018).

44. Lin, J. et al. EFNA4 promotes cell proliferation and tumor metastasis in hepatocellular carcinoma through a PIK3R2/GSK3beta/beta-catenin positive feedback loop. *Mol. Ther. Nucleic Acids* **25**, 328–341 (2021).

45. Kuang, S. Q. et al. Aberrant DNA methylation and epigenetic inactivation of Eph receptor tyrosine kinases and ephrin ligands in acute lymphoblastic leukemia. *Blood* **115**, 2412–2419 (2010).

46. Zhang, L., Kim, S. & Ren, X. The clinical significance of SIRT2 in malignancies: a tumor suppressor or an oncogene? *Front Oncol.* **10**, 1721 (2020).

47. Cohen, P. A., Donini, C. F., Nguyen, N. T., Lincet, H. & Vendrell, J. A. The dark side of ZNF217, a key regulator of tumorigenesis with powerful biomarker value. *Oncotarget* **6**, 41566–41581 (2015).

48. Aster, J. C., Pear, W. S. & Blacklow, S. C. The varied roles of notch in cancer. *Annu Rev. Pathol.* **12**, 245–275 (2017).

49. Liu, Y. et al. The genomic landscape of pediatric and young adult T-lineage acute lymphoblastic leukemia. *Nat. Genet* **49**, 1211–1218 (2017).

50. Autry, R. J. et al. Integrative genomic analyses reveal mechanisms of glucocorticoid resistance in acute lymphoblastic leukemia. *Nat. Cancer* **1**, 329–344 (2020).

51. Timofeeva, N. & Gandhi, V. Ibrutinib combinations in CLL therapy: scientific rationale and clinical results. *Blood Cancer J.* **11**, 79 (2021).

52. Pal Singh, S., Dammeijer, F. & Hendriks, R. W. Role of Bruton's tyrosine kinase in B cells and malignancies. *Mol. Cancer* **17**, 57 (2018).

53. Qiu, Y. & Kung, H. J. Signaling network of the Btk family kinases. *Oncogene* **19**, 5651–5661 (2000).

54. Boni, C. & Sorio, C. The role of the tumor suppressor gene protein tyrosine phosphatase gamma in cancer. *Front Cell Dev. Biol.* **9**, 768969 (2021).

55. Loo, L. S. W. et al. BCL-xL/BCL2L1 is a critical anti-apoptotic protein that promotes the survival of differentiating pancreatic cells from human pluripotent stem cells. *Cell Death Dis.* **11**, 378 (2020).

56. Pui, C. H. & Evans, W. E. Treatment of acute lymphoblastic leukemia. *N. Engl. J. Med* **354**, 166–178 (2006).

57. Jayappa, K. D. et al. Microenvironmental agonists generate de novo phenotypic resistance to combined ibrutinib plus venetoclax in CLL and MCL. *Blood Adv.* **1**, 933–946 (2017).

58. Liu, Y. et al. Systematic inference and comparison of multi-scale chromatin sub-compartments connects spatial organization to cell phenotypes. *Nat. Commun.* **12**, 2439 (2021).

59. Flavahan, W. A. et al. Altered chromosomal topology drives oncogenic programs in SDH-deficient GISTs. *Nature* **575**, 229–233 (2019).

60. Hafner, A. et al. Loop stacking organizes genome folding from TADs to chromosomes. *Mol. Cell* **83**, 1377–1392.e1376 (2023).

61. Miguel-Escalada, I. et al. Human pancreatic islet three-dimensional chromatin architecture provides insights into the genetics of type 2 diabetes. *Nat. Genet* **51**, 1137–1148 (2019).

62. Sobreira, D. R. et al. Extensive pleiotropism and allelic heterogeneity mediate metabolic effects of IRX3 and IRX5. *Science* **372**, 1085–1091 (2021).

63. Palomero, T. et al. CUTLL1, a novel human T-cell lymphoma cell line with t(7;9) rearrangement, aberrant NOTCH1 activation and high sensitivity to gamma-secretase inhibitors. *Leukemia* **20**, 1279–1287 (2006).

64. Beliveau, B. J. et al. OligoMiner provides a rapid, flexible environment for the design of genome-scale oligonucleotide in situ hybridization probes. *Proc. Natl Acad. Sci. USA* **115**, E2183–E2192 (2018).

65. Li, H. & Durbin, R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* **25**, 1754–1760 (2009).

66. Servant, N. et al. HiC-Pro: an optimized and flexible pipeline for Hi-C data processing. *Genome Biol.* **16**, 259 (2015).

67. 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).

68. Bhattacharyya, S., Chandra, V., Vijayanand, P. & Ay, F. Identification of significant chromatin contacts from HiChIP data by FitHiChIP. *Nat. Commun.* **10**, 4221 (2019).

69. Newman M. E. J., Girvan M. Finding and evaluating community structure in networks. *Phys. Rev. E* **69**, (2004).

70. Blondel V. D., Guillaume J. L., Lambiotte R., Lefebvre E. Fast unfolding of communities in large networks. *J. Stat. Mech.-Theory E*, (2008).

71. Lancichinetti A., Fortunato S. Limits of modularity maximization in community detection. *Phys Rev E* **84**, (2011).

72. Flyamer, I. M., Illingworth, R. S. & Bickmore, W. A. Coolpup.py: versatile pile-up analysis of Hi-C data. *Bioinformatics* **36**, 2980–2985 (2020).

73. Whyte, W. A. et al. Master transcription factors and mediator establish super-enhancers at key cell identity genes. *Cell* **153**, 307–319 (2013).

74. Wang, X., Luan, Y. & Yue, F. EagleC: A deep-learning framework for detecting a full range of structural variations from bulk and single-cell contact maps. *Sci. Adv.* **8**, eabn9215 (2022).

75. Quinlan, A. R. & Hall, I. M. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* **26**, 841–842 (2010).

76. Zhou, Y. et al. Metascape provides a biologist-oriented resource for the analysis of systems-level datasets. *Nat. Commun.* **10**, 1523 (2019).

77. Ashburner, M. et al. Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. *Nat. Genet.* **25**, 25–29 (2000).

78.   Thomas, P. D. et al. PANTHER: Making genome-scale phylogenetics accessible to all. *Protein Sci.* **31**, 8–22 (2022).
79.   Gene Ontology C., et al. The Gene Ontology knowledgebase in 2023. *Genetics* **224**, (2023).
80.   Phanstiel, D. H., Boyle, A. P., Araya, C. L. & Snyder, M. P. Sushi.R: flexible, quantitative and integrative genomic visualizations for publication-quality multi-panel figures. *Bioinformatics* **30**, 2808–2810 (2014).
81.   Kent, W. J., Zweig, A. S., Barber, G., Hinrichs, A. S. & Karolchik, D. BigWig and BigBed: enabling browsing of large distributed datasets. *Bioinformatics* **26**, 2204–2207 (2010).
82.   Crane, E. et al. Condensin-driven remodelling of X chromosome topology during dosage compensation. *Nature* **523**, 240–244 (2015).

Author contributions
Conceptualization: R.B.F., B.P.; Methodology: B.P., N.B., R.B.F., G.W.S., Y.Z.; Investigation: B.P., N.B., Y.Z., J.P., R.B.F.; Formal Analysis: B.P., N.B., Y.Z., J.P.; Resources and Reagents: R.B.F., Z.M.; Writing-Original Draft: B.P., R.B.F.; Writing-Review & Editing: Y.Z., N.B., J.P., G.W.S.; Z.M.; Funding Acquisition: R.B.F., Supervision: R.B.F.

Competing interests
The authors declare no competing interests.

Additional information
Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41467-024-52375-6.

Correspondence and requests for materials should be addressed to Robert B. Faryabi.

Peer review information *Nature Communications* thanks the anonymous reviewers for their contribution to the peer review of this work. A peer review file is available.

Reprints and permissions information is available at http://www.nature.com/reprints

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