

# Clonal evolution of the 3D chromatin landscape in patients with relapsed pediatric B-cell acute lymphoblastic leukemia

Received: 28 June 2023

Accepted: 9 August 2024

Published online: 28 August 2024



Sonali Narang<sup>1,8</sup>, Yohana Ghebrechristos<sup>1,2,8</sup>, Nikki A. Evensen<sup>1</sup>, Nina Murrell<sup>1,2</sup>, Sylwia Jasinski<sup>1,3</sup>, Talia H. Ostrow<sup>1</sup>, David T. Teachey<sup>4</sup>, Elizabeth A. Raetz<sup>1,3</sup>, Timothee Lionnet<sup>5,6</sup>, Matthew Witkowski<sup>7</sup>, Iannis Aifantis<sup>1,2,9</sup>✉, Aristotelis Tsirigos<sup>1,2,9</sup>✉ & William L. Carroll<sup>1,3,9</sup>✉

Relapsed pediatric B-cell acute lymphoblastic leukemia (B-ALL) remains one of the leading causes of cancer mortality in children. We performed Hi-C, ATAC-seq, and RNA-seq on 12 matched diagnosis/relapse pediatric leukemia specimens to uncover dynamic structural variants (SVs) and 3D chromatin rewiring that may contribute to relapse. While translocations are assumed to occur early in leukemogenesis and be maintained throughout progression, we discovered novel, dynamic translocations and confirmed several fusion transcripts, suggesting functional and therapeutic relevance. Genome-wide chromatin remodeling was observed at all organizational levels: A/B compartments, TAD interactivity, and chromatin loops, including some loci shared by 25% of patients. Shared changes were found to drive the expression of genes/pathways previously implicated in resistance as well as novel therapeutic candidates, two of which (*ATXN1* and *MNI*) we functionally validated. Overall, these results demonstrate chromatin reorganization under the selective pressure of therapy and offer the potential for discovery of novel therapeutic interventions.

Relapsed pediatric B-cell acute lymphoblastic leukemia (B-ALL) remains one of the leading causes of cancer mortality in children<sup>1,2</sup>. While five-year survival rates for newly diagnosed pediatric B-ALL have improved and now approach 90%, up to 20% of patients will suffer relapse and face a dismal prognosis<sup>3</sup>. Although recent therapeutic approaches to treating relapsed B-ALL seem encouraging, leukemia sub-clones continue to emerge through the selective pressures of therapy<sup>4</sup>. Therefore, investigating the mechanisms responsible for cancer recurrence and therapy resistance is crucial for preventing and treating B-ALL relapse.

Previous studies have focused on the discovery of relapse-enriched somatic mutations and copy number changes, which proved to be fruitful in the identification of mutations involved in resistance to one or more drugs used in treatment<sup>5–8</sup>. Our recent work characterized the evolution of the epigenetic landscape of B-ALL from diagnosis to relapse. While changes in the epigenome varied widely across patients, we were able to identify shared relapse-specific superenhancers along with concordant gene expression changes. We also found convergent evolution of histone modifications on pathways

<sup>1</sup>Perlmutter Cancer Center, NYU Langone Health, New York, NY, USA. <sup>2</sup>Department of Pathology, NYU Langone Health, New York, NY, USA. <sup>3</sup>Division of Pediatric Hematology/Oncology, Department of Pediatrics, NYU Langone Health, New York, NY, USA. <sup>4</sup>Department of Pediatrics and the Center for Childhood Cancer Research, Children's Hospital of Philadelphia, Philadelphia, PA, USA. <sup>5</sup>Institute for Systems Genetics and Department of Cell Biology, NYU Langone Health, New York, NY, USA. <sup>6</sup>Department of Biomedical Engineering, NYU Tandon School of Engineering, Brooklyn, NY, USA. <sup>7</sup>Department of Pediatrics, University of Colorado Anschutz Medical Campus, Aurora, CO, USA. <sup>8</sup>These authors contributed equally: Sonali Narang, Yohana Ghebrechristos. <sup>9</sup>These authors jointly supervised this work: Iannis Aifantis, Aristotelis Tsirigos, William L. Carroll. ✉e-mail: [ioannis.Aifantis@nyulangone.org](mailto:ioannis.Aifantis@nyulangone.org); [aristotelis.tsirigos@nyulangone.org](mailto:aristotelis.tsirigos@nyulangone.org); [william.carroll@nyulangone.org](mailto:william.carroll@nyulangone.org)

known to be associated with resistance, such as p53, JAK/STAT, and Ras, thus highlighting the importance of chromatin changes in disease progression<sup>9</sup>. Although studies of the linear genome have been enormously productive in yielding fundamental insights into B-ALL biology and therapeutic options, recent studies suggest that an additional critical level of control of the epigenome is connected to the 3D structure of chromosomes.

Thus far, no one has explored B-ALL progression through the lens of 3D chromatin organization. The advent of chromosome conformation capture (3C) techniques, such as Hi-C, revealed that the folding of large complex chromosomes is nonrandom and involves a hierarchy of biologically relevant structures including chromosome territories, A and B compartments, topologically associating domains (TADs), and enhancer-promoter chromatin loops<sup>10–14</sup>. These organizational layers have been shown to be tightly linked to epigenetic and transcriptional regulation. Additionally, a growing number of cancers have been shown to harbor disruptions in various layers of chromosome topology including gliomas, non-Hodgkin lymphoma, and T-cell ALL<sup>15–17</sup>.

In B-cell lymphomas, a recurrent histone H1 (*HIST1H1B-E*) linker protein mutation was shown to cause large scale chromatin decompaction coupled with a gain of H3K36me2 and/or loss of H3K27me3, which created a permissive landscape for the expression of stem cell genes normally silenced during development<sup>18</sup>. Loss of *Hic* and *Hle* in mice enhanced fitness and self-renewal properties of GC B-cells, causing aggressive lymphomas. Similarly, recurrent alterations in TAD boundaries and TADs in T-ALL were shown to be linked to changes in gene expression and affect downstream oncogenic effectors including NOTCH1 signaling<sup>17</sup>. In B-ALL, we showed that the relapse-specific NSD2 E1099K mutation drives relapse by disrupting 3D chromatin organization at the compartment level<sup>19</sup>. More recently, recurrent and subtype-specific alterations in A/B compartments, TAD activity, and chromatin loops were identified in AML<sup>20</sup>. Additionally, inhibition of DNA methylation with a hypomethylating agent restored 3D genome organization and gene regulation in AML cells. We now aim to capitalize on chromatin conformation capture techniques to discover how each layer of 3D chromatin organization evolves over the course of B-ALL progression and contributes to therapy resistance and relapse.

To uncover 3D chromatin architecture-related mechanisms underlying disease progression, we performed Hi-C, ATAC-seq, and RNA-seq on 12 matched primary pediatric leukemia specimens at diagnosis and relapse. Mapping of structural variations using Hi-C data revealed previously unidentified stable, diagnosis-specific, and relapse-specific SVs as well as relapse-specific fusion transcripts indicating a possible role in clonal evolution (e.g. MBTD1::WDR89). Moreover, Hi-C analysis revealed genome wide chromatin remodeling specifically in terms of A/B compartments, TAD interactivity, and chromatin loops. We showed that 12.17% of changes were shared by at least 25% of the patients, including at the *JAK2* locus. Integration with ATAC-seq and RNA-seq datasets revealed strong correlation with both gene expression and chromatin accessibility. Importantly, integration of looping data with ATAC-seq and Hi-ChIP for H3K27ac demonstrated rewiring events can occur at stable chromatin accessibility sites. Additionally, we identified recurrent A/B compartments and TAD interactivity changes across the patient cohort supporting a crucial role in the clonal evolution of B-ALL. Recurrent TAD interactivity changes were identified in genes previously found dysregulated in cancer, such as *PCDH9*, as well as pathways previously implicated in drug resistance such as Wnt-beta catenin signaling, highlighting the activation of these vital pathways through 3D chromatin architecture alterations. Lastly, we discovered novel genes dysregulated by recurrent 3D changes, such as *MNI* and *ATXN1*, which were found to play a role in drug response validating the functional significance of our findings. These results significantly extend our knowledge of shifts in the genetic and epigenetic landscape that drive therapy resistance in childhood B-ALL and identify novel therapeutic targets.

## Results

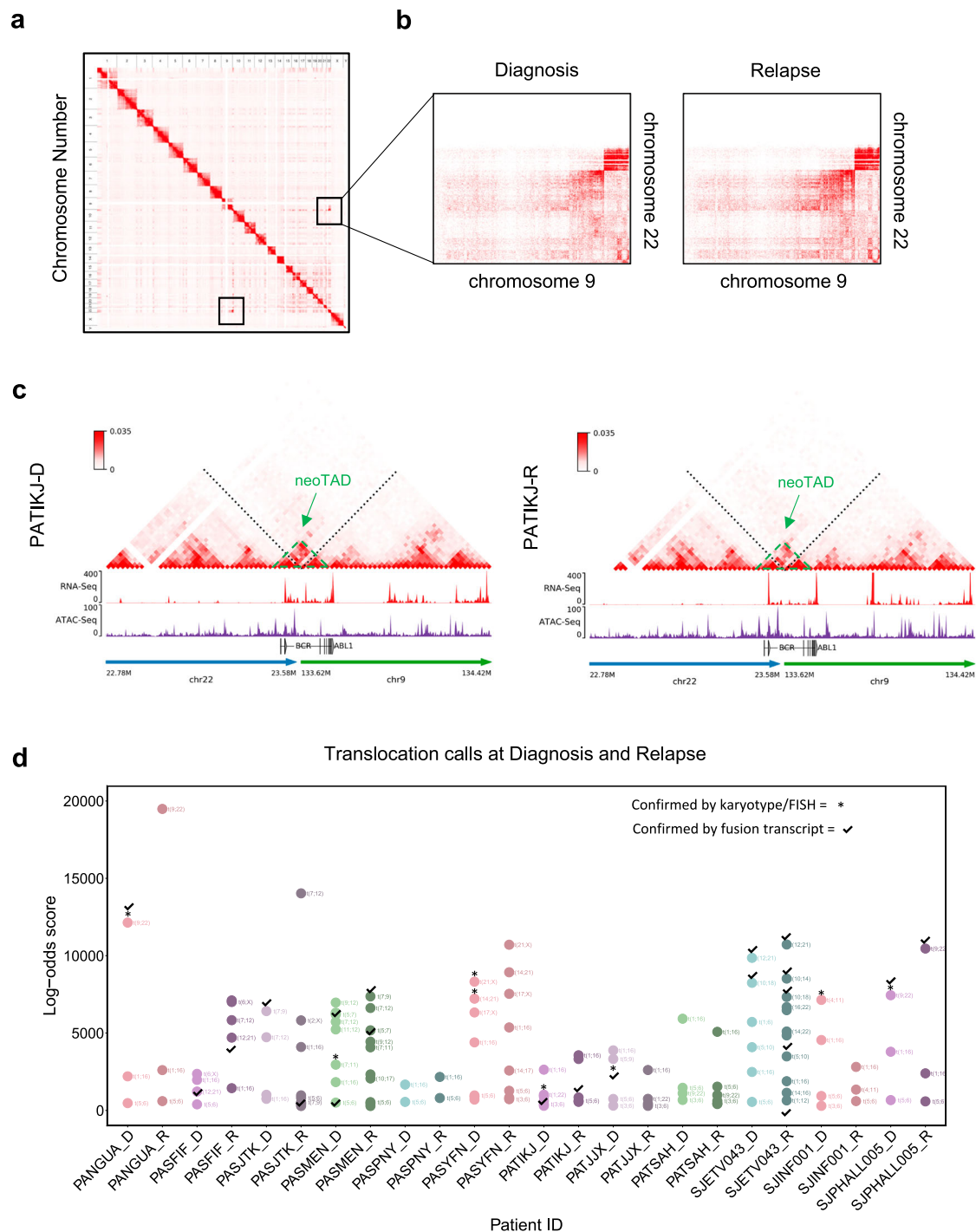
### Mapping structural variations in pediatric ALL using Hi-C

To broadly assess global chromatin architecture throughout the progression of pediatric B-ALL, we performed Hi-C, RNA-seq, and ATAC-seq on sorted blasts (CD45dim, CD19+) from 12 matched primary bone marrow B-ALL samples at diagnosis and at relapse from patients enrolled in Children's Oncology Group (COG) protocols or St. Jude Children's Research Hospital Total XV protocol<sup>21</sup> (Supplementary Fig. 1a; Supplementary Table 1). Recurrent cytogenetic abnormalities in pediatric B-ALL serve to identify subgroups important for risk stratification of therapy<sup>22,23</sup>, therefore we first examined the Hi-C data for structural variants. Patient Hi-C data was processed by our in-house HiC-bench platform<sup>24</sup>. Alignment rates showed a high percentage of usable long-range read-pairs across our patient cohort (~100 million; Supplementary Fig. 1b, c).

Visual inspection of entire patient Hi-C contact matrices revealed inter-chromosomal contacts (Fig. 1a). Closer inspection of the Hi-C contact matrices of patient PATIKJ demonstrated a canonical ALL rearrangement, the *BCR-ABL1* translocation between chromosomes 9 and 22 at both diagnosis and relapse (Fig. 1b). Reconstruction of Hi-C contact matrices revealed the presence of inter-chromosomal contacts across the breakpoint and the formation of a new topologically associating domain, referred to as a neoTAD<sup>25</sup> (Fig. 1c). Fusion transcript analysis<sup>26</sup> with RNA-seq data confirmed the presence of the *BCR::ABL1* fusion transcript in this patient (Supplementary Data 1). This example of a well-known B-ALL translocation screened for at diagnosis by conventional methods provided evidence that our methods in detecting these chromosomal abnormalities using Hi-C was successful.

To further our investigation, we utilized Hi-C-Breakfinder<sup>27</sup> and HiNT<sup>28</sup> to call inter-chromosomal translocations from each Hi-C dataset in our cohort. We overlapped translocations predicted by the methods with cytogenetics provided for each patient, and found Hi-C-Breakfinder performed better and revealed previously identified as well as novel translocations (Supplementary Fig. 2a, b; Supplementary Data 2). Additionally, EagleC was used to call all SVs and was found to produce approximately 50% overlapping translocation calls with Breakfinder. While 50 stable translocations were called, for the first time, we report diagnosis-specific (5) and relapse-specific (21) translocations providing further evidence for clonal evolution as a mechanism for drug resistance (Fig. 1d). To further investigate these 26 dynamic translocations, we visually inspected each of them using Juicebox web application<sup>29</sup> and confirmed one of the diagnosis-specific and 11 of the relapse-specific translocations (Supplementary Data 3). Interestingly, two of the relapse-specific translocations had faint signal at diagnosis upon visual inspection but were below the level of detection for Breakfinder, supporting clonal selection of a minor subclone at diagnosis. For example, while a translocation was called between chromosomes 16 and 22 in SJETV043 at relapse but not at diagnosis, upon visual inspection we were able to detect faint signal at the breakpoint at diagnosis (Fig. 2a). Another example in the same patient revealed a faint signal at diagnosis at the breakpoint of the t(chr1,12) called at relapse (Supplementary Fig. 2c).

While the presence of inter-chromosomal contacts across breakpoints were evidenced by the reconstruction of Hi-C contact matrices<sup>25</sup>, we wanted to determine whether these contacts resulted in any fusions at the RNA level. Overall the fusion transcript analysis using STAR-fusion called a variable number of fusion transcripts across patients (Supplementary Fig. 2d; Supplementary Data 1). Of the 50 translocations shared from diagnosis to relapse, nine were found to have fusion transcripts. The one diagnosis-specific and three of the 11 relapse-specific translocations also had fusion transcripts detected, none of which were shared across patients. An example of a diagnosis-specific translocation in patient PATJXX showed the presence of a fusion between chromosomes 5 and 9 at diagnosis only (Fig. 2b). Fusion transcript analysis revealed the existence of a



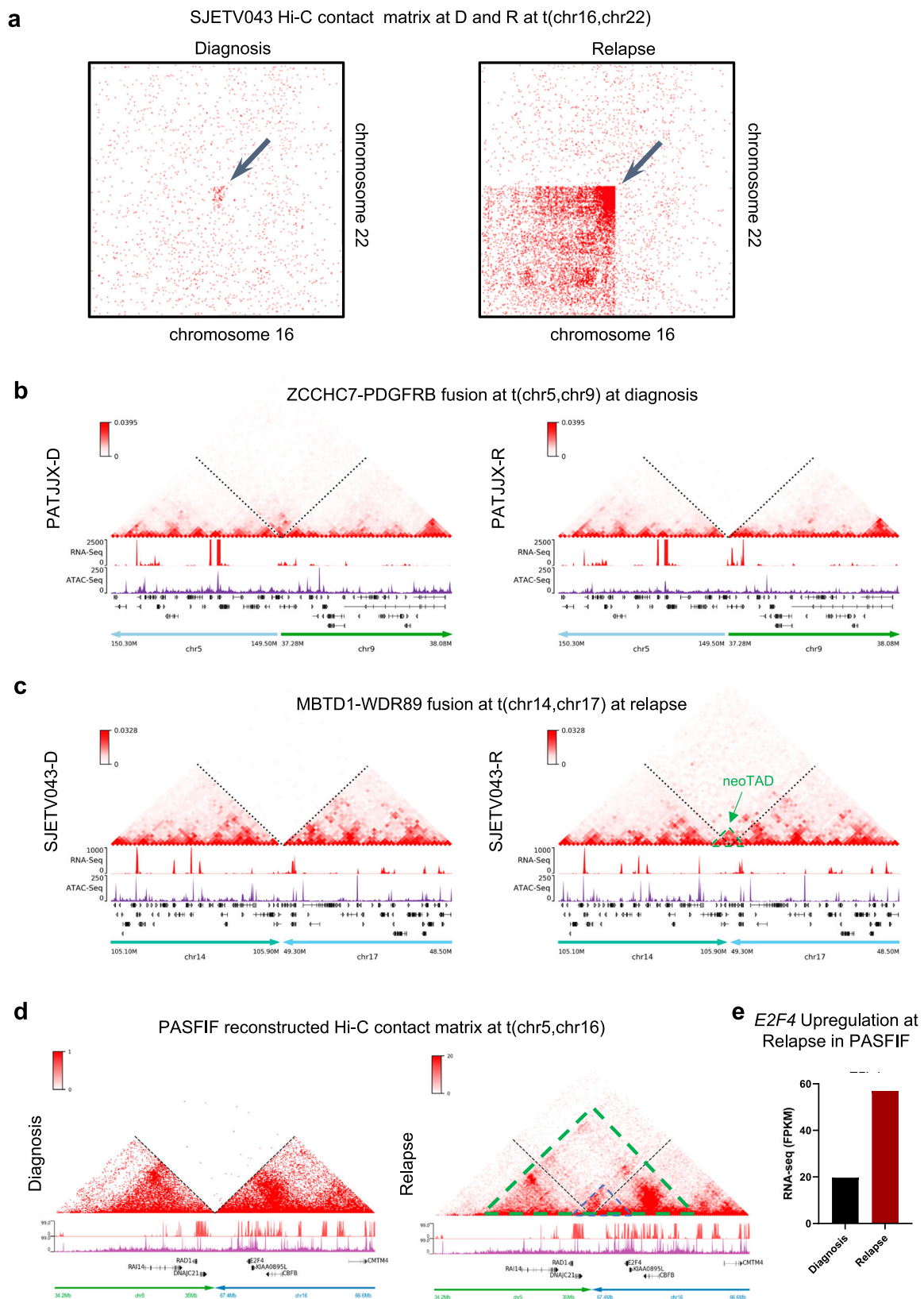
**Fig. 1 | Mapping of structural variations at Diagnosis and Relapse using Hi-C.** **a** Entire genome Hi-C contact matrix for patient PATIKJ. **b** Zoomed-in Hi-C contact matrix demonstrating BCR-ABL1 translocation between chromosome 9 and chromosome 22 in patient PATIKJ at diagnosis and at relapse (left and right respectively). **c** Reconstructed Hi-C contact matrix presenting an example of a shared

BCR-ABL1 translocation in patient PATIKJ between chromosomes 9 and 22 at diagnosis and relapse (left and right respectively). Green dashed triangle highlights a NeoTAD. **d** Translocations and their log-odds score identified in diagnosis and relapse per patient. Those also found by cytogenetics or fusion transcripts are marked.

previously unreported fusion transcript in the diagnosis sample, ZCCHC7::PDGFRB (Supplementary Data 1 and 3). *PDGFRB* alterations are commonly found in Ph-like B-ALL cases, which has clinical implications as these patients can be responsive to tyrosine kinase inhibition<sup>30,31</sup>. Lastly, an example of a relapse-specific translocation in patient SJETV043 showed the presence of inter-chromosomal contacts between chromosomes 14 and 17 only in the relapse sample (Fig. 2c). Fusion transcript analysis confirmed the presence of the known ETV6::RUNX1 at both diagnosis and relapse, and demonstrated the

existence of a previously unreported fusion transcript MBTD1::WDR89 in the relapse sample (Supplementary Data 1 and 3). *MBTD1* is found as a recurrent translocation partner with *ZMYND11* in acute myeloid leukemia and has been demonstrated to play a role in transcriptional activation of oncogenic genes such as *myc*<sup>32</sup>.

For the remaining relapse-specific, visually validated translocations that did not have fusion transcripts detectable by STAR-fusion, we sought to determine if there were examples of dysregulated gene expression surrounding breakpoints. One example was found for the



**Fig. 2 | Mapping of structural variations using Hi-C reveals Diagnosis and Relapse-specific translocations.** **a** Hi-C contact matrix demonstrating relapse-specific translocation between chromosomes 16 and 22 in patient SJETV043 with faint signal visually detected at diagnosis (arrow). **b** Reconstructed Hi-C contact matrix presenting a diagnosis-specific translocation in patient PATJX between chromosomes 5 and 9 at diagnosis and relapse (left and right respectively). **c** Reconstructed Hi-C contact matrix presenting a relapse-specific translocation in

patient SJETV043 between chromosomes 14 and 17 at diagnosis and relapse (left and right respectively). Green dashed triangle highlights a NeoTAD. **d** Reconstructed Hi-C contact matrix demonstrating a relapse-specific NeoTAD (dashed green triangle). New contacts overlap region of ATAC peaks (purple tracks) and increased transcription at relapse (red tracks) overlapping the *E2F4* locus (dashed blue triangle). **e** Bar plot showing upregulation of *E2F4* transcripts at relapse in PASFIF.



relapse-specific translocation between chromosomes 5 and 16 in patient PASIF (Fig. 2d; Supplementary Fig. 2e). Upon reconstruction around the breakpoint, new contacts were observed across the two chromosomes in the relapse sample compared to the diagnosis sample. Interestingly, several genes within the neoTAD show differential gene expression. Two ATAC peaks near a transcriptionally active region are evident on chromosome 5 brought into contact with genes near the breakpoint on chromosome 16 (Fig. 2d), suggestive of potential promoter/enhancer hijacking as Xu et al. previously demonstrated<sup>20</sup>. One of the genes, E2F transcription factor 4 (*E2F4*), that is upregulated at relapse (Fig. 2e) has been implicated in acute myeloid leukemia (AML). AML patients with high expression of *E2F4* had significantly lower overall survival and in vitro models its overexpression lead to increased proliferation and differentiation blockage<sup>33</sup>. These results extend the landscape of cytogenetic abnormalities in pediatric B-ALL that could play a significant role in the identification of subgroups that have significant clinical implications for risk stratification and therapeutic strategy.

### B-ALL relapse is associated with recurrent changes in A/B compartments

Next, we sought to determine changes in the various layers of 3D chromatin architecture throughout the progression of pediatric B-ALL. We first assessed compartmentalization of the genome between B-ALL samples at diagnosis and relapse. A/B compartment status was analyzed in individual patient pairs using Cscore<sup>34</sup>. Principal Component Analysis (PCA) of Cscore compartment scores revealed two distinct groups of patient samples in which the matched diagnosis and relapse pairs cluster together and the patient with a *KMT2A* rearrangement, SJETV043, clusters separately (Fig. 3a). Overall, the number of A and B compartment calls was comparable between diagnosis and relapse samples (Fig. 3b). We next assessed A/B compartment switching in individual diagnosis-relapse patient pairs. We observed 4.88% compartment switching between diagnosis and relapse samples across the 12 patient pairs (Fig. 3c; Supplementary Fig. 3a). In order to confirm that 3D changes were not driven by copy number variations (CNVs), we called CNVs using EagleC. Less than 1.3% of compartment bins that underwent a switch were found to overlap with a CNV (Supplementary Data 4). B-ALL is a heterogeneous disease classified by various chromosomal abnormalities, such as hyperdiploidy, hypodiploidy, translocations, and intra-chromosomal amplification. Given this heterogeneity, B-ALL relapse expectedly has various mechanisms of clonal selection. While the small size of our cohort is limiting, we were particularly interested in identifying recurrent compartment switches across the patient cohort and/or switches that were concordant with gene expression and chromatin accessibility as they could be indicative of a role in clonal evolution. Overall, we identified 12.17% recurrent switches shared by at least 3 out of 12 patients (25% of patients). Specifically, 5.73% (420/7335) of A to B switches and 5.99% (483/8059) of B to A switches were shared by at least 3 patient samples (Fig. 3d).

We then integrated the Hi-C data with RNA-seq and ATAC-seq data to investigate whether compartment switches are associated with changes in gene expression and chromatin accessibility (Supplementary Fig. 3b–e). We performed differential expression analysis of all expressed genes (FPKM > 1) and differential chromatin accessibility analysis of all ATAC-seq peaks ( $\text{abs}(\log_2(\text{fold change})) > 1.0$ , false-discovery rate (FDR) < 0.01) within compartment switches at the individual patient pair level and found that B to A compartment switches upon relapse correlated with an increase in gene expression, whereas A to B compartment switches upon relapse were associated with a decrease in gene expression relative to expression changes within stable compartments (Fig. 3e, left panel). We witnessed a similar trend with chromatin accessibility changes (Fig. 3e, right panel). Interestingly, we found a recurrent B to A compartment switch at the *JAK2* locus in 3 of the 12 patient pairs. This B to A compartment switch

was also concordant with increased gene expression (Fig. 3f). The *JAK2* gene is a member of the non-receptor tyrosine kinase family and mediates intracellular signaling upon activation of cytokine receptors<sup>35</sup>. Several studies have identified gain of function mutations and translocations affecting *JAK2* in pediatric B-ALL<sup>36,37</sup>.

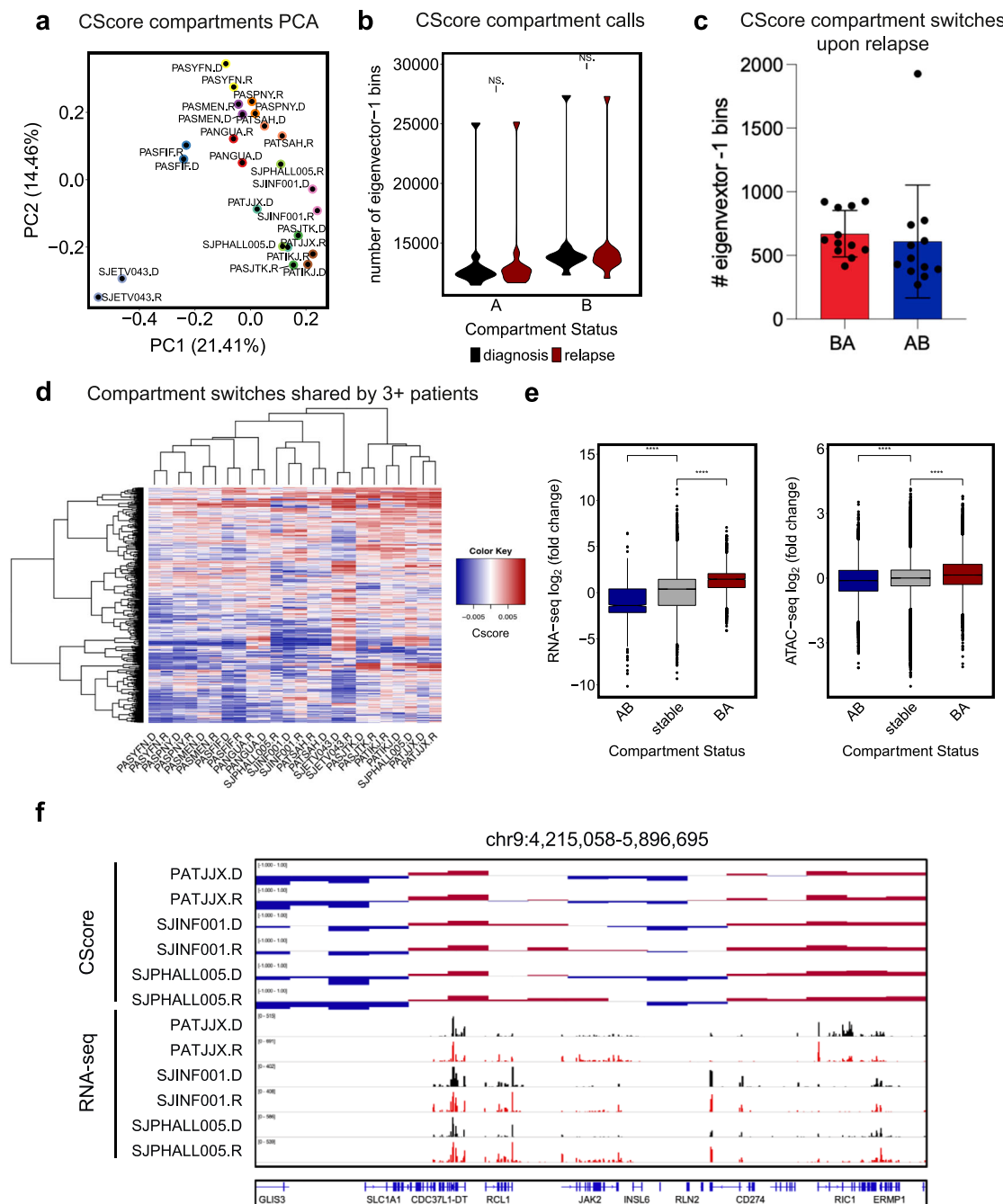
### B-ALL relapse drives recurrent changes in TAD interactivity

Following compartment analysis, we investigated TAD interactivity changes (i.e. changes in intra-TAD chromatin interactions) between B-ALL diagnosis and relapse patient pairs. We first called TADs at diagnosis and relapse and identified ~3000 TADs in each patient (Supplementary Fig. 4a). Comparison of TAD interactivity between B-ALL samples at diagnosis and relapse identified a range of significantly differential TADs across patients, with 8 out of 12 patients having more gains in TAD activity at relapse (Fig. 4a, b; (FDR) < 0.1 and  $\text{abs}(\log_2(\text{fold change})) > 0.25$ ; Supplementary Fig. 4b). Less than 2.4% of differential TADs overlapped with CNVs (Supplementary Data 5). Next, to investigate whether TAD interactivity associated with changes in gene expression and chromatin accessibility, we performed differential expression analysis of all expressed genes (FPKM > 1) and differential chromatin accessibility analysis of all ATAC-seq peaks ( $\text{abs}(\log_2(\text{fold change})) > 1.0$ , FDR < 0.01) within differentially active TADs at the individual patient pair level. Increased TAD interactivity at relapse significantly associated with positive fold changes in gene expression, whereas decreased TAD interactivity at relapse associated with negative fold changes in gene expression relative to expression changes within stable TADs (Fig. 4c, left). We saw a similar trend with chromatin accessibility changes (Fig. 4c, right).

Of note, patient SJPHALLO05 had more than 800 changes in TAD interactivity, when applying the same thresholds, with the majority being loss in activity (Supplementary Fig. 4c). We observed a similar trend with patient SJPHALLO05 when integrating the TAD interactivity data with either gene expression or chromatin accessibility data (Supplementary Fig. 4d). This patient also showed the greatest amount of compartment changes at 10.85%, with the majority being A to B (Supplementary Fig. 3a). Interestingly, we saw increased *CEBPA* expression and decreased *PAX5* at relapse (Supplementary Fig. 4e), suggesting a more myeloid-like profile upon relapse. These observations are consistent with patient SJPHALLO05 having a therapy related or lineage switch secondary malignancy as previously described<sup>38</sup>. Furthermore, we also observed myeloid lineage architectural changes, including a gain of TAD activity, a switch to A compartment, and increased expression of *DACH1* at relapse (Supplementary Fig. 4f, g). *DACH1* has been shown to regulate cell cycle progression in myeloid cells<sup>39</sup>.

As with the compartment analysis, we next aimed to identify TAD interactivity changes that were shared amongst our patient cohort possibly indicating major modes of tumor escape from therapy. Notably, we found a recurrent decrease in TAD interactivity at the *PCDH9* locus in 4 of the 12 patient pairs (Fig. 4d). All 4 patients demonstrated decreased mean TAD interactivity, gene expression, Cscore, and accessibility (Fig. 4e). Closer examination of the patient with the most dramatic change in TAD interactivity, patient PASPNY, revealed concordant changes in compartments, chromatin accessibility, and gene expression (Fig. 4f). The *PCDH9* (Protocadherin 9) gene encodes a transmembrane protein involved in cell adhesion and calcium ion binding and has been identified as one of the top down-regulated genes in pediatric B-ALL<sup>40–44</sup>.

Lastly, we performed pathway enrichment analysis<sup>45–47</sup> to identify pathways associated with changes in 3D architecture. Pathway enrichment analysis on genes that were concordant in expression with 3D changes such as A/B compartment switches or TAD interactivity changes revealed enrichment of previously identified key pathways including, mTORC1, Wnt-beta catenin, and p53 signaling (Supplementary Fig. 4h)<sup>9,48</sup>. These data demonstrate that the activation of



**Fig. 3 | Compartments undergo discrete changes upon B-ALL relapse.** **a** PCA of A/B compartment calls with Cscore. **b** Violin plot showing A and B compartment calls for diagnosis and relapse samples. Significant differences were calculated with a paired one-sided *t* test comparing number of A or B eigenvector-1 bins between diagnosis and relapse. **c** Bar plot showing number of compartment switches for each patient pair upon relapse. Bar plot shows mean, bars represent standard deviation, and each point represents a patient ( $n = 12$ ). **d** Heatmap representation of compartment switches shared by at least 3 patients upon relapse. **e** Correlation boxplots between gene expression and compartment changes (left) and chromatin

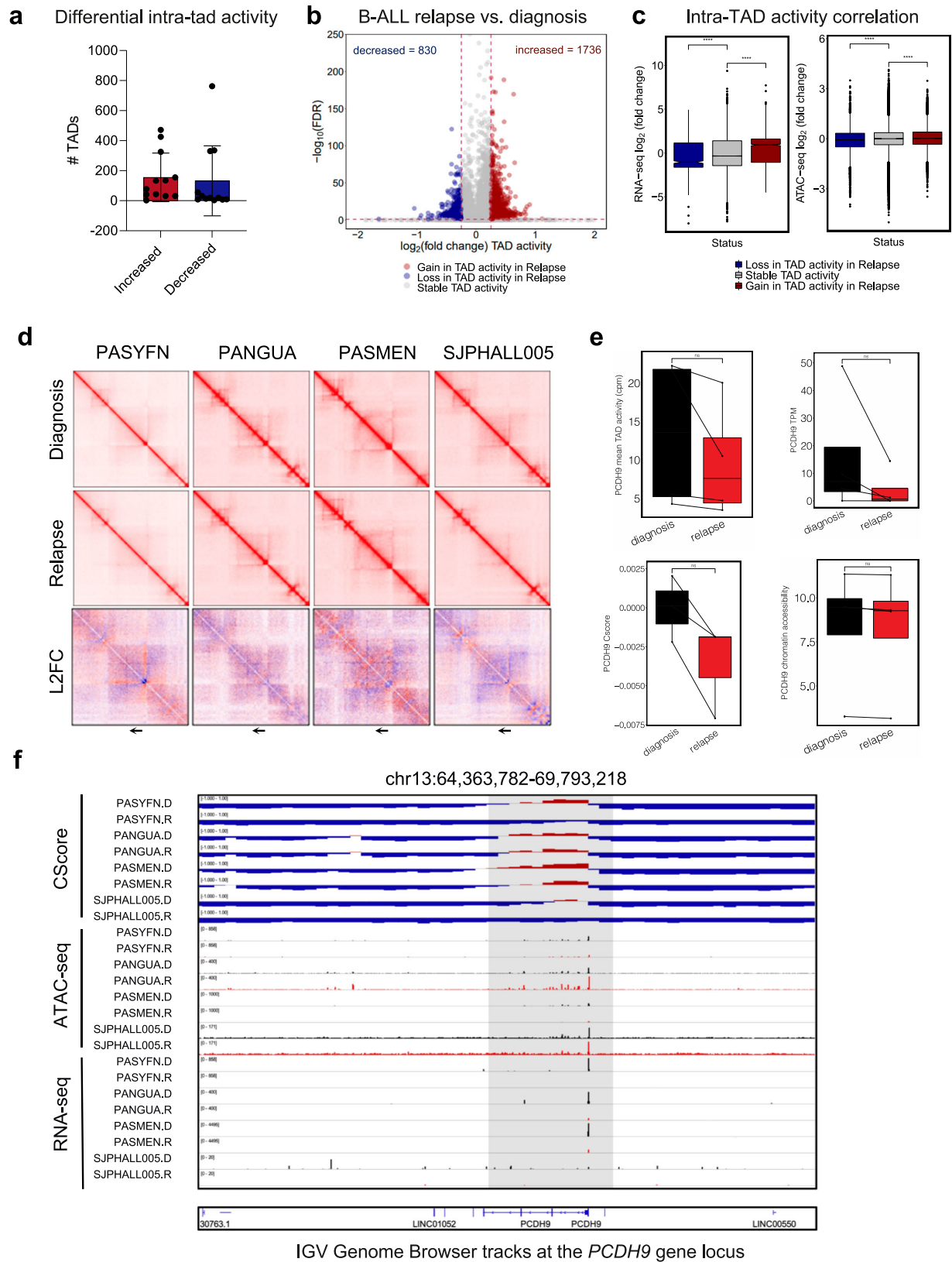
accessibility and compartment changes (right) from relapse to diagnosis of 12 individual patient pairs. Significant differences were calculated by unpaired one-sided *t* test comparing genes or peaks from compartment changes AB or BA to genes or peaks from stable compartments (\*\*\*\* $p < .00001$ ). Boxplots show medians (horizontal line in each box), interquartile ranges (boxes), 1.5 interquartile (whiskers) and each point represents a gene or peak. **f** IGV tracks demonstrating concordant B to A compartment switches and increased gene expression in 3 patient pairs at the *JAK2* gene locus.

these important pathways goes beyond genetic alterations and validate these pathways as novel targets to prevent and treat relapse.

### Chromatin loop data reveals enriched factors at relapse

We next wanted to investigate the rewiring of chromatin loops and the factors responsible for mediating such changes during selection

from therapy. We performed a looping analysis with FitHiC2<sup>49,50</sup>. Looping calls varied based on sequencing depth across the patient cohort (Fig. 5a). We performed differential looping analysis to identify loops that increased, decreased, or remained the same upon relapse (Fig. 5b) and visually confirmed them using Aggregate Peak Analysis (APA) (Supplementary Fig. 5a). We observed

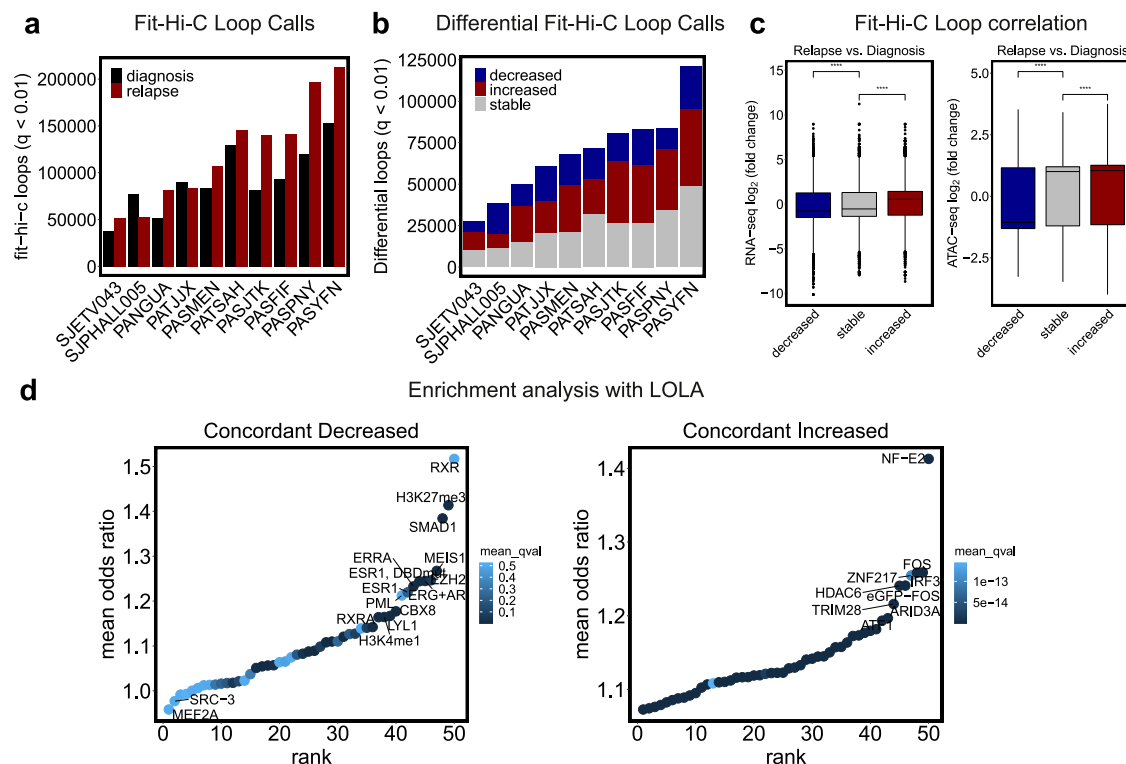


a strong correlation of loop changes with both gene expression changes and chromatin accessibility changes (Fig. 5c). Additionally, we identified a positive association between looping changes and both compartment and TAD interactivity changes (Supplementary Fig. 5b)

To identify key regulators of transcription attributed to loop changes, we performed motif enrichment analysis with Locus Overlap Analysis (LOLA)<sup>51</sup> with ATAC-seq peaks overlapping concordant chromatin loops. We identified several chromatin binding factors enriched in regions with decreased chromatin accessibility and decreased

**Fig. 4 | B-ALL patients share concordant changes in TAD activity.** **a** Bar plot presenting number of intra-TAD activity changes per patient pair upon relapse ( $\text{abs}(\text{L2FC}) > 0.25$ ,  $\text{FDR} < 0.01$ ). Bar plot shows means, bars represent standard deviation, and each point represents a patient ( $n = 12$ ). **b** Volcano plot presenting increased, decreased, and stable TADs between diagnosis and relapse samples (red, blue, and gray respectively). **c** Correlation boxplots showing gene expression and chromatin accessibility changes within differential TADs (paired  $t$  test,  $p < 0.05$ ) from relapse to diagnosis of 12 individual patient pairs. Significant differences were calculated by unpaired one-sided  $t$  test comparing gene expression or ATAC peaks within TADs with interactivity changes, losses, or gains, to gene expression or ATAC peaks within stable TADs ( $****p < .00001$ ). Boxplots show medians (horizontal line in each box), interquartile ranges (boxes), 1.5 interquartile (whiskers) and each point represents a gene or peak. **d** Hi-C contact matrices presenting a TAD

decreased in activity at relapse in 4 of 12 patient pairs at the *PCDH9* locus (Diagnosis, Relapse, and L2FC (Relapse/Diagnosis) from top to bottom). **e** Box plots demonstrating decreased *PCDH9* mean TAD activity in cpm (top left), decreased gene expression in tpm (top right), decreased *PCDH9* compartment Cscore (bottom left), and decreased chromatin accessibility in cpm (bottom right) in the 4 patient pairs at diagnosis and relapse. Significant differences were calculated with a paired one-sided  $t$  test comparing mean TAD activity, TPM, Cscore, and cpm between diagnosis and relapse. Boxplots show medians (horizontal line in each box), interquartile ranges (boxes), 1.5 interquartile (whiskers) and each point represents a patient with pairs connected by lines. **f** IGV genome browser tracks demonstrating A to B compartment switch, decreased chromatin accessibility, and decreased gene expression in the 4 patients with decreased TAD interactivity at the *PCDH9* gene locus.



**Fig. 5 | Chromatin loop data reveals enriched factors at relapse.** **a** Bar plot presenting number of loops called with Fit-Hi-C2 at diagnosis and relapse ( $q < 0.01$ ). **b** Bar plot presenting number of loops increased, decreased or stable upon relapse (blue, red, gray respectively). **c** Fit-Hi-C loop correlation with gene expression and chromatin accessibility (left and right respectively). Significant differences were calculated by unpaired one-sided  $t$  test comparing genes or peaks from loops that

increased or decreased to genes/peaks from stable loops ( $****p < 0.00001$ ) from relapse to diagnosis of 12 individual patient pairs. Boxplots show medians (horizontal line in each box), interquartile ranges (boxes), 1.5 interquartile (whiskers) and each point represents a gene or peak. **d** Enrichment analysis with LOLA of decreased and increased ATAC-seq peaks (left and right) concordant with Fit-Hi-C loops and gene expression.

chromatin looping including RXR and EZH2 (Fig. 5d, left panel). We also identified several enriched factors in regions with increased chromatin accessibility and increased chromatin looping including FOS and NF-E2 (Fig. 5d, right panel).

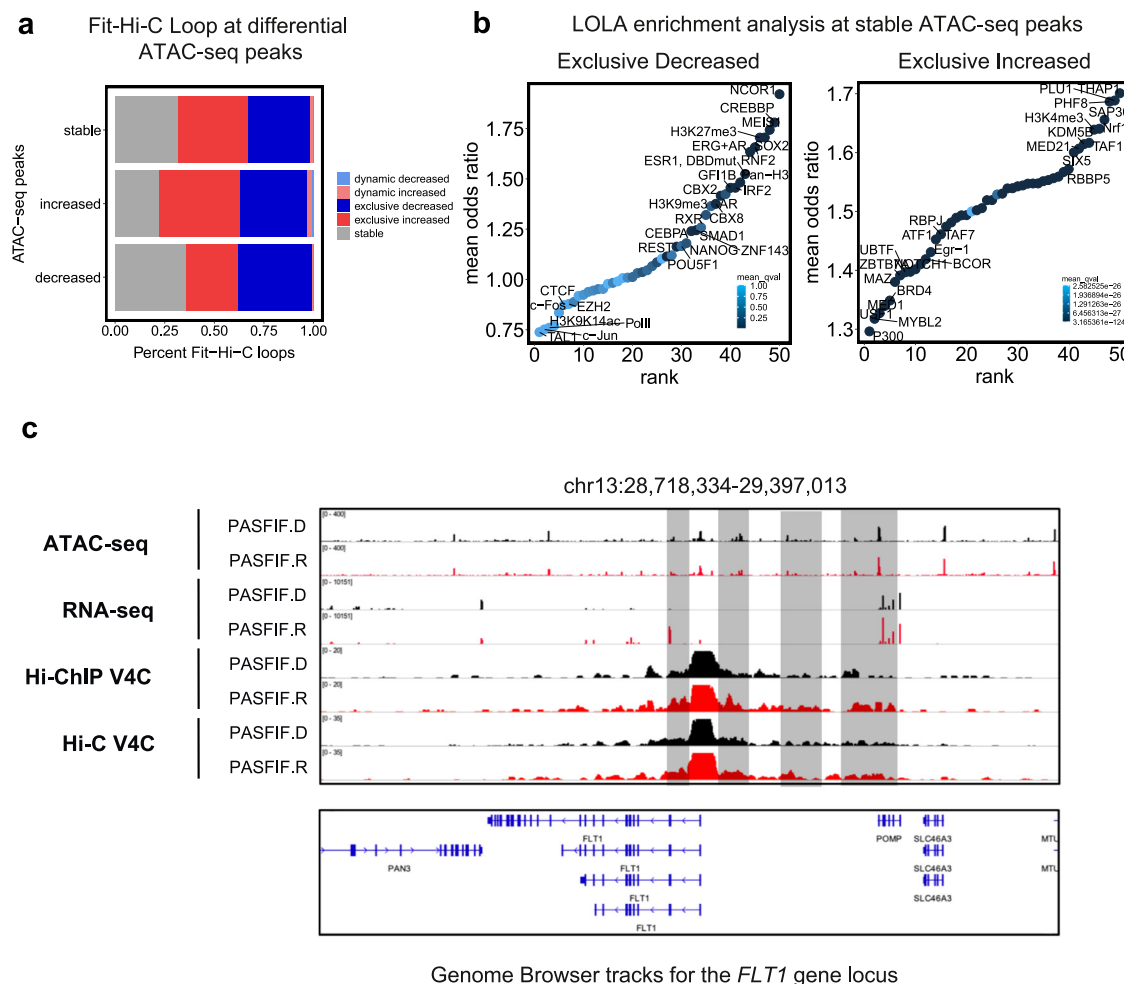
### H3K27ac HiChIP data reveals exclusive looping at stable chromatin accessibility sites

We were also interested in the behavior of enhancer anchored loops upon B-ALL relapse based on our previous work demonstrating the importance of enhancers and superenhancers in relapse<sup>9</sup>. We performed H3K27ac HiChIP with one patient pair, PASFIF, to evaluate H3K27ac anchored loops. Looping analysis was performed with FitHiC2<sup>49,50</sup>. To examine the relationship between alterations in the H3K27ac chromatin loops and chromatin accessibility, we categorized the H3K27ac chromatin loops into those that were exclusively increased (exclusive to relapse), exclusively decreased (exclusive to

diagnosis), dynamically increased (increased), dynamically decreased (decreased), or stable and assessed chromatin accessibility changes. Interestingly, integration of H3K27ac HiChIP with ATAC-seq data revealed significantly increased and decreased exclusive loops at regions with stable chromatin accessibility sites (Fig. 6a). This data suggests that significant 3D rewiring events can occur at stable chromatin accessibility sites that would otherwise be entirely ignored if analysis is focused only on sites with changes in chromatin accessibility using ATAC-seq.

To expand on this finding, we performed enrichment analysis using LOLA<sup>51</sup> with specifically stable ATAC-seq peaks overlapping concordant and exclusive H2K27ac anchored loops. We identified several interesting chromatin binding factors enriched in regions with stable chromatin accessibility and decreased enhancer-promoter looping including NCOR1, CREBBP, and MEIS (Fig. 6b, left panel). We also identified several factors enriched in regions with stable





**Fig. 6 | PASFIF H3K27ac HiChIP data reveals exclusive loops at stable ATAC-seq peaks. a** ATAC-seq and Fit-Hi-C loop association bar plot. **b** Enrichment analysis with LOLA of increased, decreased, or stable ATAC-seq peaks with exclusively decreased, dynamic decreased, stable, exclusively increased, or dynamic increased

loops. **c** IGV snapshot of the Fms Related Receptor Tyrosine Kinase 1 (*FLT1*) gene locus presenting ATAC-seq, RNA-seq, and H3K27ac Hi-ChIP and Hi-C virtual 4C tracks at the TSS as a viewpoint. Gray boxes highlight regions of differential contact frequency.

chromatin accessibility and increased enhancer-promoter looping including *PLU1*, *THAP1*, and *PHF8* (Fig. 6b, right panel).

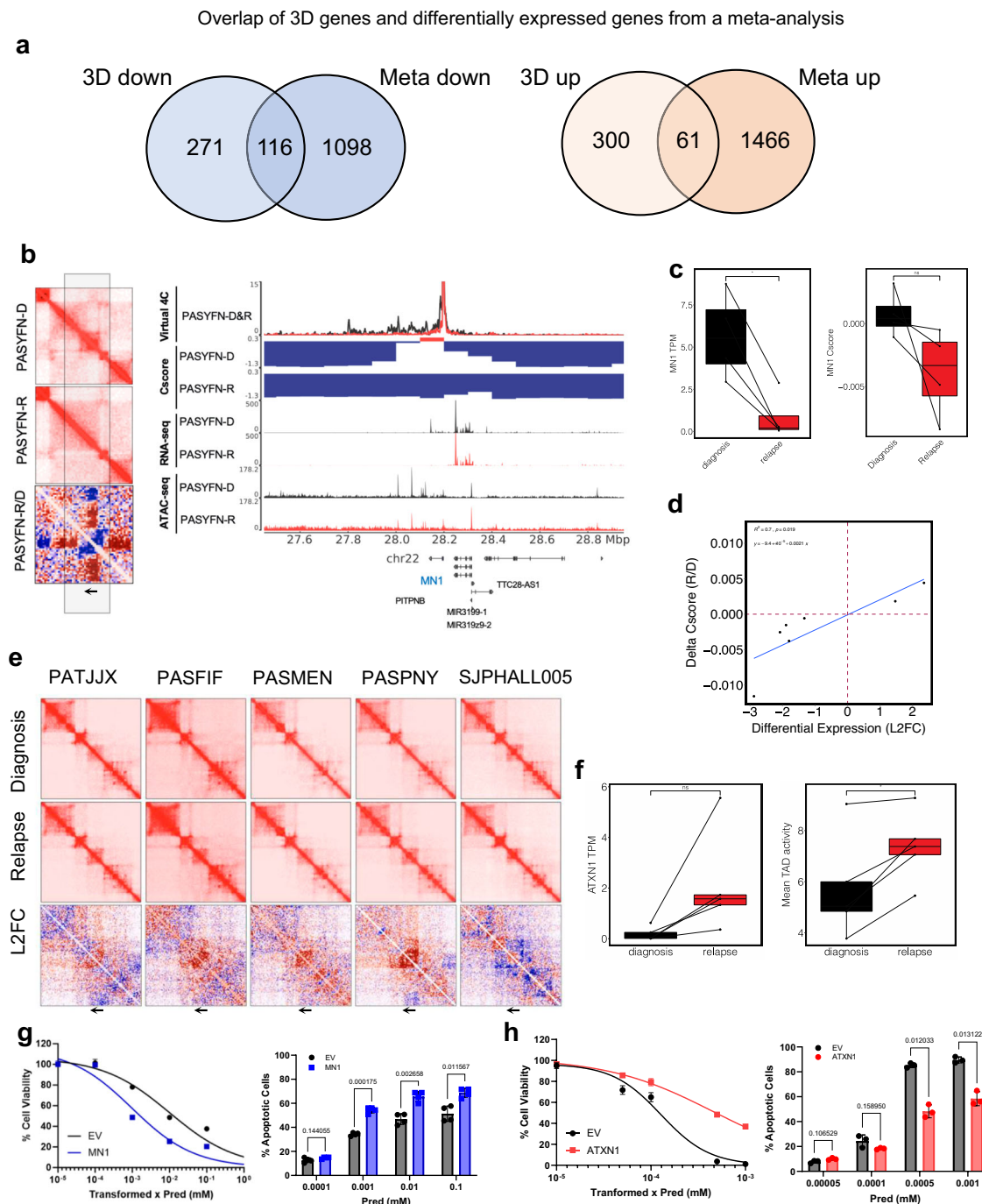
Lastly, we present an example of stable chromatin accessibility and exclusively increased H3K27ac chromatin looping at the vascular endothelial growth factor receptor 1 (*FLT1*) gene locus<sup>52</sup> in patient PASFIF (Fig. 6c). Genome browser tracks show stable chromatin accessibility and increased gene expression. H3K27ac HiChIP and Hi-C virtual 4C tracks show differential contact frequencies between diagnosis and relapse samples. *FLT1* activation in ALL has been shown to increase survival of blasts by influencing their localization into the bone marrow and exit back into circulation<sup>52</sup>.

### Functional validation of candidate genes suggests a role in drug resistance

In an effort to discover potential gene targets involved in drug resistance or relapse that are downstream of 3D chromatin changes, we identified genes that demonstrated 3D change from diagnosis to relapse (A/B compartment, TAD interactivity, or chromatin loop), that were concordant with gene expression, and were shared by at least 3 patients, indicating a possible major role in clonal evolution. This resulted in a list of 361 upregulated and 387 downregulated genes. To validate the expression changes, we first overlapped our 3D gene-based lists with a list of the most significantly differentially expressed genes from a meta-analysis of a large cohort of three microarray datasets generated from pediatric B-ALL diagnosis/relapse pairs<sup>53</sup>.

Thirty-two percent (116 of 387) of the 3D mediated downregulated and 17% (61 of 361) of the upregulated genes were also differentially regulated in this much larger cohort of samples (Fig. 7a; Supplementary Table 2) indicating a prominent role of 3D reorganization in clonal evolution. Furthermore, using published data from a genome wide CRISPR screen performed with the B-ALL cell line REH and various chemotherapies used in pediatric B-ALL treatment performed by Oshima et al.<sup>54</sup>, we found that 42% (153 of 361) of the upregulated and 24% (93 of 387) of the downregulated genes were within the top 10% of guides negatively (for up) or positively (for down) selected from at least one of the drugs tested, suggesting a cell intrinsic role in drug resistance (Supplementary Data 6).

Interestingly, at the *MNI* locus we identified a recurrent A to B compartment switch in 4 of the 12 patient pairs that was concordant with decreased gene expression, chromatin accessibility, and virtual 4C signal (Fig. 7b–d). Linear regression model revealed a positive relationship between changes in compartment Cscore and gene expression ( $R^2 = 0.7$ ,  $p < 0.019$ , Fig. 7d). *MNI* encodes a transcriptional co-regulator involved in the co-activation of various transcription factors<sup>55,56</sup>. Notably, we also found a recurrent increase in TAD interactivity with a concomitant increase in gene expression at the *ATXN1* locus in 5 of the 12 patient pairs (Fig. 7e, f). Closer examination of the patient with the most dramatic change in TAD interactivity, PASPNY, revealed concordant changes in TAD interactivity, compartments, gene expression, chromatin accessibility, and virtual 4C signal



**Fig. 7 | Functional validation of *MN1* and *ATXN1* suggests a role in drug resistance.** **a** Venn diagrams presenting overlap of 3D gene-based lists with a list of the most significantly differential genes from a meta-analysis of three microarray datasets. **b** Example shown for PASYFN of an A to B compartment switch shared by 4 patient pairs at the *MN1* gene locus including TAD activity, V4C signal, A/B compartments, RNA-seq, and ATAC-seq tracks. **c** Box plot demonstrating a decrease in *MN1* gene expression in tpm (left) and decrease in Cscore (right) in the 4 patient pairs with the A to B compartment switch. Significant differences were calculated using paired t test ( $p < 0.05$ ). Boxplots show medians (horizontal line in each box), interquartile ranges (boxes), 1.5 interquartile (whiskers) and each point represents a patient with pairs connected by lines. **d** Linear regression model demonstrating positive correlation between compartment and gene expression changes at the *MN1* locus for patients with differential expression  $\text{absL2FC} > 1.0$  ( $n = 7$ ). **e** Hi-C contact matrices presenting a TAD increased at relapse in 5 of 12 patient pairs at the *ATXN1* locus (Diagnosis, Relapse, and L2FC (Relapse/Diagnosis) from top to bottom). **f** Box plots demonstrating increased *ATXN1* gene expression

(top) and increased intra-TAD activity (bottom) in the 5 patient pairs upon relapse. Significant differences were calculated using paired t test ( $p < 0.05$ ). Boxplots show medians (horizontal line in each box), interquartile ranges (boxes), 1.5 interquartile (whiskers) and each point represents a patient with pairs connected by lines. **g** Cell viability (left) and apoptosis assays (right) demonstrating increased sensitivity to treatment with Prednisolone in SEM cells overexpressing *MN1* compared to empty vector control (EV). **h** Cell viability (left) and apoptosis assays (right) demonstrating increased resistance to treatment with Prednisolone in 697 cells overexpressing *ATXN1* compared to EV. Cell viability graphs shown (**g**, **h**) are one experiment representative of three independent experiments, each with technical triplicates. The dots represent the mean of technical triplicates for that experiment with standard deviation shown. Curve represents the nonlinear fit of transformed X. The null hypothesis of one curve fits all datasets was rejected with  $p < 0.001$  (**g**, **h**). Apoptosis graphs shown represents the mean and standard deviation of 4 (**g**) and 3 (**h**) individual experiments. Statistical significance shown based on unpaired t test with Welch correction with  $p$  value indicated.

(Supplementary Fig. 6a). This was further validated by increased H3K27ac peak changes in many patients from our previous study<sup>9</sup>. The Ataxin-1 (*ATXN1*) gene is a chromatin binding factor that represses Notch signaling<sup>57,58</sup>.

We validated a functional role for both *MNI* and *ATXN1* in B-ALL cell lines. The B-ALL cell line SEM was engineered to overexpress *MNI* or EV as control and verified via western blot (Supplementary Fig. 6b). We observed a significant increase in sensitivity to prednisolone upon overexpression of *MNI* (Fig. 7g) but no difference was observed with thiopurines, suggesting that loss of *MNI* would provide an advantage to cells undergoing steroid treatment. The B-ALL cell line 697 was engineered to overexpress *ATXN1* or EV as control and verified via western blot (Supplementary Fig. 6c). Overexpression of *ATXN1* resulted in increased resistance to prednisolone compared to EV (Fig. 7h), suggesting upregulation of *ATXN1* leads to drug resistance. There was no difference in IC50s for other chemotherapies, including thiopurines or doxorubicin (Supplementary Fig. 6d). Note when *ATXN1* or *MNI* expression was modified in other cell lines (REH, RS4;11, SEM, 697), we did not observe changes in chemosensitivity indicating cell context plays an important role, a finding consistent with the fact that these changes were shared in a subset of patients.

## Discussion

Relapsed pediatric B-ALL remains one of the leading causes of cancer mortality in children. Previous studies have characterized the genetic, epigenetic, and transcriptional landscapes of B-ALL. However, no one has investigated B-ALL progression through the lens of 3D chromatin organization. In this work, we investigate the dynamics of 3D chromatin architecture in B-ALL progression by assessing 12 matched primary pediatric leukemia specimens at diagnosis and relapse. We demonstrate that 3D genome organization plays an additional critical level of control in B-ALL progression.

The current clinical standard for the detection of genetic abnormalities in B-ALL is fluorescence in-situ hybridization (FISH). Despite its sensitivity, FISH can only detect known genetic aberrations with the provision of specific probes. Hi-C in contrast is a powerful tool that gives insight into previously unknown structural variants. We sought to leverage Hi-C and investigate the stability of structural variants between diagnosis and relapsed B-ALL. Robust SV analysis with Hi-C data from diagnosis and relapse samples confirmed translocations provided by conventional cytogenetics but also some previously unidentified stable, diagnosis-specific, and relapse-specific translocations. This data suggests that Hi-C data can be used as a computationally less rigorous alternative to WGS methods to discover cytogenetic abnormalities in ALL and other disease models. We confirmed the presence of fusion transcripts such as BCR::ABL1 and ETV6::RUNX1 as well as identified novel translocation-driven fusion transcripts offering possible new avenues for therapeutic intervention especially those that appear to be enriched at relapse.

In general, such translocations are thought to be an early event in tumorigenesis and are maintained at diagnosis and relapse. We show a striking number of additional translocations indicating a fundamental role in tumorigenesis. The loss of translocations at relapse indicates a dispensable role in clonal evolution while relapse-specific translocations may provide additional fitness to the leukemic clone. Interestingly, both *PDGFRB* and *ZCCHC7* have been previously observed in rearrangements with different loci in B-ALL. The detection of *ZCCHC7*::*PDGFRB* at diagnosis would support the use of a tyrosine kinase inhibitor, although in this case the clone appeared to be extinguished with chemotherapy. MBTD1, malignant brain tumor domain containing 1, has been identified as a part of a chimeric protein with ZMYND11 (zinc finger MYND-type containing 11) that leads to active chromatin states at stemness related genes in AML<sup>59</sup>. Herein, we show a novel fusion transcript with MBTD1 and WDR89, a member of the WD40-repeat proteins whose function is largely unknown but may

play a role in immunotherapy response<sup>60,61</sup>. While the functional relevance is beyond the scope of the present study, the fact that this novel transcript was relapse-specific suggests a role in disease progression. Performing Hi-C on a larger cohort of samples could enable the rapid benchmarking of cytogenetic abnormalities associated with B-ALL progression and would allow for an unbiased view of the genome to identify unknown fusions that would otherwise be missed using current methods. Lastly, while our studies show a number of diagnosis or relapse specific translocations, all pairs had shared translocations indicating that the origin of the relapsed clone was from a common ancestral precursor including case SJPHALL005 where the initial diagnostic clone was lymphoid and the patient relapsed with AML. Thus, the subsequent relapse likely originated from a common ancestral stem-like clone, a feature known to be associated with Ph+ ALL<sup>62</sup>. Furthermore, there were several relapse-specific translocations that were visually detected (but not called by Breakfinder) at diagnosis, supporting the idea of clonal selection of a minor subclone at diagnosis<sup>5,7</sup>.

In addition to structural changes, we demonstrate significant A/B compartment, TAD interactivity, and looping changes between diagnosis and relapse that strongly correlated with expression and accessibility. These findings support a prominent role of 3D architecture in modulating transcription. We saw that the magnitude of these architectural changes varied among pairs. This is consistent with our previous findings showing variable degrees of epigenetic changes from diagnosis to relapse, highlighting the plasticity of the chromatin landscape under the selective pressures of therapy<sup>9</sup>. Analysis of 3D-regulated genes revealed previously discovered pathways associated with drivers of clonal evolution. For example, activation of JAK/STAT signaling has been implicated in leukemia progression and we observed a recurrent B to A compartment switch at the *JAK2* locus. Previous studies have shown mutations as well as translocations lead to activated *JAK2* in pediatric B-ALL, which could be targeted through JAK inhibitors<sup>53,54</sup>. Our Hi-C analysis demonstrates an additional mechanism by which blasts may increase *JAK2* thus extending the use of JAK inhibitors at relapse.

Importantly our work identified additional novel pathways involved in drug resistance specifically to prednisolone, a cardinal feature of ALL at relapse. A meta-analysis of gene expression in B-ALL diagnosis and relapse pairs confirms that *MNI* expression is down-regulated while *ATXN1* is upregulated at relapse consistent with our findings<sup>53</sup>. *MNI* is a transcriptional co-activator and indirectly activates retinoic acid receptor and vitamin D receptor (VDR) target genes<sup>63</sup>. Many studies have confirmed an association between polymorphisms of VDR with osteonecrosis, a side effect of glucocorticoids in ALL therapy<sup>64,65</sup>. Furthermore, retinoids have been found to promote glucocorticoid-induced apoptosis of various T-cell lines. Thus, down-regulation of *MNI* may impair apoptosis induced by retinoid signaling under steroid treatment<sup>66</sup>. Conversely, overexpression of *ATXN1* resulted in increased resistance to prednisolone. *ATXN1*, thought to function in extracellular matrix remodeling, has been shown to regulate Notch and Capicua (CIC) controlled developmental processes<sup>67</sup>. We were unable to replicate phenotypic changes in multiple cell lines; however, we have previously demonstrated cell context specific changes in drug response, which can be explained in part due to 3D chromatin differences<sup>56</sup>.

By integrating looping changes with gene expression and chromatin accessibility, we were able to identify potential regulators of these changes. We found enrichment of sites for chromatin binding factors shown to be key players in cancer. One such example is EZH2, which is part of the polycomb repressive complex (PRC2) that is crucial for proliferation and differentiation and plays a complex role in cancer<sup>68</sup>. We previously demonstrated DNA hypermethylation at PRC2 target genes at relapse<sup>9</sup>, which would be consistent with decreased expression, accessibility, and 3D looping interactions where there are



EZH2 motifs. Studies related to *EZH2* and *NSD2* mutations in leukemia and myeloma have also revealed a complex interplay between these two epigenetic modifiers<sup>69,70</sup> and perhaps understanding the 3D chromatin changes associated with EZH2 could provide more insight for the best method to target epigenetic modifiers. The importance of epigenetic modifiers is strengthened further by the enrichment of H3K27me3 and the role this mark has been demonstrated to play in repressor hijacking events<sup>20</sup>. The highest-ranked factor for decreased looping and expression was retinoid X receptor (RXR), which is a nuclear receptor that interacts with retinoic acid receptor alpha and recruits corepressors to chromatin to regulate target genes, including genes involved in differentiation and apoptosis<sup>71</sup>. Targeting these receptors in *IKZF1* mutated BCR-ABL1 ALL has also been shown to reverse stem cell like phenotype<sup>72</sup>. The highest-ranking factor motif for activated regions was nuclear factor erythroid-2 (NF-E2). Overexpression in mouse models has been shown to lead to increased stem and progenitor cells and spontaneous development of AML<sup>73</sup>.

Interestingly, the H3K27ac HiChIP with patient PASFIF demonstrated a significant amount of de novo loops appearing and disappearing at regions of stable chromatin upon relapse. This finding indicates that sites of stable chromatin accessibility provide invaluable information on a significant number of 3D rewiring events that would otherwise be entirely ignored if analyzes are directed to changes in chromatin accessibility exclusively (e.g., ATAC-seq). For example, CREBBP motifs were enriched in regions of decreased looping at relapse. *CREBBP* loss of function mutations have previously been observed at relapse in a significant number of patients and is associated with resistance to glucocorticoids<sup>74</sup>. Likewise, PHF8 motifs were enriched in regions of increased looping and PHF8 expression is associated with ALL progression through activation of the MEK/ERK pathway<sup>75</sup>.

Overall, this study demonstrates 3D genome organization as a critical factor in reshaping the transcriptional output of leukemia cells under the selective pressures of therapy. These processes appear to drive pathways shown to be involved in drug resistance but also previously undiscovered genes/pathways that may play a role in clonal evolution. Our findings extend indications for inhibition of kinase and cytokine receptor signaling in B-ALL using available agents, as well as, highlight the potential of novel approaches to modulate *MNI*, *ATXN1* and *FLT1* at relapse. The identification of common downstream targets among patients indicates the convergence of drug resistance mechanisms and has important clinical implications for possible future therapeutic interventions.

## Methods

Cell culture, drug preparation, viral preparation, immunoblotting, and phenotypic assays were performed according to methods published previously<sup>5–7</sup>.

### Cells and reagents

The B-lineage leukemia cell lines Reh (ATCC), RS4;1I (ATCC), RCH-ACV (DSMZ), and SEM (kindly gifted by Jun Yang, St. Jude Children's Hospital) were grown in RPMI1640 medium. All media were supplemented with 10% FBS, 1% penicillin/streptomycin under 5% CO<sub>2</sub> at 37 °C. No cell lines were used beyond passage 20. Each leukemia line was validated by short tandem repeat analysis through IDEXX cellcheck9+ service. Cell lines were routinely monitored for mycoplasma contamination by PCR using ATCC Universal Mycoplasma Detection Kit (20-1012 K).

### Experimental Procedures

**Primary B-ALL patient samples.** Cryopreserved paired diagnosis/relapse primary patient bone marrow samples from individual patients were obtained from the Children's Oncology group (COG) ALL biorepository or St. Jude Children's Research Hospital<sup>21</sup>. All subjects provided consent for banking and future research use of these specimens in accordance with the regulations of the institutional review boards of

all participating institutions. Samples were thawed and stained for CD45 (PerCP-Cy<sup>TM</sup>5.5 Mouse Anti-Human CD45 clone HI30, BD Bioscience Cat# 564105) and CD19 (APC Mouse Anti-Human CD19 Clone HIB19, BD Bioscience Cat# 555415). Leukemic blasts (CD45dimCD19+) were sorted using SY3200 cell sorter (Sony Technologies) in the NYU Cytometry and Cell Sorting Laboratory and then immediately processed for downstream steps.

### RNA-seq

Following FACS, 50,000–200,000 leukemic cells were immediately resuspended in Buffer RLT plus BME and frozen on dry ice. RNA was extracted using the QIAGEN RNeasy Micro Kit and quality was verified by an Agilent Bioanalyzer 2100 (PICO chip). RNA-seq libraries were generated by the NYU Genome Technology Center using the Low input Clontech SMART-Seq kit and sequenced on the Illumina NovaSeq 6000.

### ATAC-seq

Following FACS, 50,000 leukemic cells were immediately processed for ATAC-seq by the NYU Genome Technology Center. ATAC libraries were generated based on the protocol by Buenrostro et al.<sup>76</sup> Briefly, cells are resuspended in cold lysis buffer (10 mM Tris Cl, 10 mM NaCl, 3 mM MgCl<sub>2</sub>, 0.1% (v/v) Igepal CA-630, pH 7.4) and centrifuged for 1 min at 500 × g. Nuclei were tagged using Nextera (Illumina) Tagmentation DNA buffer and enzyme. PCR amplification was performed as described by Buenrostro et al.<sup>76</sup> Libraries were sequenced on the Illumina NovaSeq 6000.

### Hi-C

Following FACS, leukemic cells were immediately fixed and cross-linked with 2% formaldehyde at room temperature for 10 min. Glycine was added to stop the reaction. Crosslinked cells were washed with PBS and then frozen on dry ice. Crosslinked cell pellets were stored at –80 until ready for processing. Hi-C libraries were generated using the Arima-HiC kit (A410110) with Swift Biosciences® Accel-NGS® 2 S Plus DNA Library Kit (Cat # 21024) and Indexing kit (Cat #) according to manufacturer's protocol. Libraries were amplified and quantified using KAPA Library Amplification kit (Cat # KK2620). Final libraries were sequenced at the NYU Genome Technology Center on the Illumina NovaSeq 6000.

### H3K27ac HiChIP

HiChIP libraries were generated using the Arima-HiC<sup>+</sup> kit (A410232) with Swift Biosciences® Accel-NGS® 2 S Plus DNA Library Kit (Cat # 21024) and Indexing kit (Cat # 26148) according to manufacturer's protocol with minor modifications. After in-situ proximity ligation, samples were sonicated on Covaris LE 220 at 4 °C, 300 W, 15% duty factor, 200 cycles per burst, for 300 s. Samples are then precleared with protein G beads for an hour, and subsequently incubated with 5 µg of H3K27AC antibody (ab4729, Lot:GR3231988) at 4 °C overnight. After IP, libraries were amplified and quantified using KAPA Library Amplification kit (Cat # KK2620). Final libraries were sequenced at the NYU Genome Technology Center on the Illumina NovaSeq 6000.

### Drug preparation

Stock solutions of Doxorubicin (Dox) (Sigma-Aldrich, St. Louis, MO) were prepared in double-distilled water, Prednisolone (Pred) (Pharmacia, St. Paul, MN) in 0.9% NaCl (Saline), 6-thioguanine (6-TG) and 6-mercaptopurine (6-MP) in 0.1 M NaOH. Drugs were diluted and added to cell culture media at indicated concentrations. Vehicle toxicity was tested at highest drug dose used to ensure no impact on viability.

### Generation of overexpression cell lines

The open reading frames of *MNI* or *ATXN1* were PCR amplified from the pcDNA3.1 plasmids purchased from Genscript (OHu18815D and



OHu26181D, respectively) and cloned into pLenti-R1R2-DEST (kindly gifted by Kenneth Scott, Baylor College of Medicine) using Invitrogen's Gateway(R) Technology. HEK293T cells were then transfected with these constructs (empty vector or gene of interest) along with packaging plasmids coding for VSV-G, Gag-pol, and Rev using lipofectamine 2000 (Invitrogen, Carlsbad, CA). Viral supernatant was used to spin infect B-ALL cell lines with 8 µg/ml Polybrene (Millipore). After 72 h, infected cell lines were selected with 0.75–3 µg/ml of Puromycin.

### Cytotoxicity assays

Cells were plated at a density of 60,000–80,000 cells/well and exposed to chemotherapy agents commonly used in the treatment of pediatric B-ALL for 4 days. Cell viability was measured using CellTiter-Glo® Luminescent Cell viability Assay (Promega) according to the manufacturer's protocol. The luminescence was read using a PerkinElmer EnSpire multimode plate reader. Luminescent readings were normalized to the untreated control for each cell line to account for any differences in baseline growth kinetics or survival. Cytotoxicity percentages were plotted using nonlinear regression with a four parametric variable slope with log transformed x concentrations with a bottom constraint equal to zero on GraphPad Prism 7.2 (GraphPad Prism Software Inc., La Jolla, CA). For apoptosis, following treatment, cells were stained with Annexin V and DAPI and analyzed via flow cytometry using the BD LSR II cytometer within the NYU Cytometry and Cell Sorting core. The percent Annexin V+ cells were determined for percent apoptotic using FlowJo Software. The percentage of apoptotic cells was compared between empty vector control and MN1 or ATXN1 overexpressing cells using an unpaired t test. Each experiment was plated in triplicate and repeated at least three times.

### Computational analysis

**RNA-seq analysis.** Patient RNA-seq fastq files were processed in single replicates using the route “rna-star” and “rna-star-groups-dge” from the Slide-n-Seq (sns) pipeline: <https://igordot.github.io/sns/>. Processing steps include alignment of paired-end reads to the human reference genome (hg19) using the STAR aligner with default parameters<sup>77</sup>. Counts were obtained using featureCounts<sup>78</sup>. Bigwig tracks were obtained for visualization on individual samples using deeptools (v3.1.0)<sup>79</sup>. Downstream analysis including normalization and differential expression analysis was performed using DESeq2<sup>80</sup>. Genes were categorized as differentially expressed if  $\text{abs}(\text{L2FC}) > 0.58$ ,  $p$  value  $< .05$ . Pathway analysis was performed using enrichR<sup>46</sup>.

**ATAC-seq analysis.** Patient ATAC-seq fastq files were processed in two replicates using the route “atac” from the Slide-n-Seq (sns) pipeline: <https://igordot.github.io/sns/>. Processing steps included aligning paired-end reads to the human reference genome (hg19) with Bowtie2(v2.3.4.1)<sup>81</sup>. Reads with a mapping quality  $< 30$  were removed. Duplicated reads were removed using Sambamba (v0.6.8)<sup>82</sup>. Remaining reads were analyzed by applying the peak-calling algorithm MACS2(v2.1.1)<sup>83</sup>. Bigwig tracks were obtained for visualization on individual samples using deeptools (v3.1.0)<sup>79</sup>. Differential ATAC-seq peak analysis was performed using DiffBind<sup>84</sup>. Nearest genes were annotated using ChIPseeker<sup>85</sup>. Enrichment analysis of genomic regions sets was performed using Bioconductor package LOLA (Locus overlap analysis or enrichment of genomic ranges; R package version 1.24.0) with RStudio (v3.6.1) with the hg19 LOLA core database<sup>51</sup>. LOLA core is curated from many sources including TF binding sites from Encode and epigenome databases from Cistrome.

**Hi-C and H3K27ac HiChIP analysis.** Raw Hi-C sequencing data was processed with the hic-bench platform<sup>24</sup>. Cell line Hi-C data was

processed as single replicates. Data was aligned against the human reference genome (GRCh37/hg19) with bowtie2(version 2.3.1)<sup>81</sup>. The reads used for downstream analyzes were filtered for by the GenomicTools tools-hic filter command in the hic-bench platform using default parameters. The GenomicTools tools-hic filter command discards reads including multi-mapped reads (“multihit”), read-pairs with only one mappable read (“single sided”), duplicated read-pairs (“ds.duplicate”), read-pairs with a low mapping quality of MAPQ  $< 20$ , read-pairs resulting from self-ligated fragments (together called “ds.filtered”), and short range interactions resulting from read-pairs aligning within 25 kb (“ds.too.short”). Downstream analysis was performed with the accepted intra-chromosomal read-pairs (“ds.accepted intra”). The number of accepted intra-chromosomal read-pairs varied between ~40 and ~140 million for all samples (Chapter 1; Fig. 2). Hi-C interaction matrices were generated for each chromosome separately using the hic-bench platform at 40 kb resolution. Filtered read counts were normalized by iterative correction and eigenvector decomposition (ICE)<sup>86</sup>. To account for variances in read counts of more distant loci, distance normalization for each chromosome matrix was performed.

### Translocation calling from patient Hi-C data

Hi-C breakfinder<sup>27</sup> and HiNT<sup>28</sup> were used to call previously identified, as well as, novel intra-chromosomal and inter-chromosomal translocations from the matched B-ALL diagnosis/relapse patient Hi-C data. Hi-C breakfinder required 3 input files including a bam file, an inter-chromosomal expectation file, and an intra-chromosomal expectation file. The bam file was generated by hic-bench and the two expectations files for hg19 were provided on the Dixon lab GitHub page: [https://github.com/dixonlab/hic\\_breakfinder](https://github.com/dixonlab/hic_breakfinder). Hi-C breakfinder generated lists of structural variant predictions at different resolutions including 1 Mb, 100 kb, and 10 kb. For the purposes of comparing translocations between Hi-C breakfinder and HiNT, the 100 kb resolution was used. Translocations were called at 100 kb resolution using HiNT through the “translocations” step within the hic-bench platform<sup>24</sup>. EagleC<sup>87</sup> (version 0.1.9) was used to call copy number inferences (deletions and duplications) at three resolutions 5 kb, 10 kb, 50 kb. High-confidence calls were determined by setting the probability cutoff to 0.95.

### Fusion transcript analysis

Fusion transcript analysis was performed using STAR-fusion: <https://github.com/STAR-Fusion/STAR-Fusion/wiki><sup>88</sup>. STAR-fusion required paired-end RNA-seq fastq files and a reference genome (hg19) in order to identify candidate fusion transcripts.

### Hi-C contact matrix visualization

To visualize Hi-C contact matrices, ICE normalized Hi-C contact matrices for the corresponding chromosome were loaded and normalized by the total number of intra-chromosomal interactions for Diagnosis and Relapse samples. The log2FC Hi-C contact matrix was produced by applying the log2 function on the division product of the Relapse Hi-C table by the Diagnosis Hi-C table.

To visualize translocations, Hi-C contact matrices were generated in the hic format using the “tracks” step within the hic-bench platform<sup>24</sup>. Hi-C contact matrices were visualized using Juicebox<sup>29</sup>.

### Hi-C contact matrix reconstruction

Hi-C contact matrices were reconstructed to demonstrate inter-chromosomal translocations at 20 kb resolution using NeoLoopFinder<sup>25</sup>. NeoLoopFinder required cool files which were generated using the “tracks” step within the hic-bench platform<sup>24</sup>. NeoLoopFinder also required files containing a list of structural variations that were generated at 10 kb resolution using Hi-C breakfinder<sup>27</sup> as described above.

## A/B compartments analysis

A/B compartments analysis was performed using the “compartments” and “compartments-stats” steps within the hic-bench platform<sup>24</sup>. The Cscore tool algorithm was used to assign active (A) and inactive (B) compartments<sup>34</sup>. For a bin to be considered a switch from A to B or from B to A, the compartment score sign had to flip in sample 2 (Relapse) when compared with the reference sample (sample 1 or Diagnosis). The absolute difference between the compartment scores had to be higher than the cutoff (default: 1.2). The difference was computed as a relative delta:

$$\text{delta} = (Y - X) / \text{abs}(Y) \# \text{delta value calculation}$$

$X$  = compartment score of bin in sample 1 (reference)

$Y$  = compartment score of bin in sample 2

## TAD interactivity analysis

TAD interactivity was assessed using the “domains” and “domains-diff” steps within the hic-bench platform<sup>24</sup>. The “domains” step uses the hic-ratio algorithm for TAD calling developed within hic-bench by previous Tsirigos lab member Haris Lazaris in which the average of the normalized interaction scores is calculated for all interactions taking place within a particular TAD. The “domains-diff” step assesses TAD interactivity alterations and was developed by previous lab members Sofia Nomikou and Andreas Kloetgen<sup>17,24</sup>. To identify TADs with differential interactivity between Relapse and Diagnosis, we used the TADs identified in the Diagnosis sample as a reference to identify common TADs. Once mean TAD interactivity values were obtained from the Hi-C data, a wilcoxon two-sided rank sum non-parametric test was performed to determine the p-values for each TAD. Multiple testing was used to correct these p-values by adjusting to the total number of TADs. Lastly, the log2 fold change (log2FC) of intra-TAD activity value was calculated between the samples. TAD interactivity alterations were categorized as significant if  $\text{abs}(\text{L2FC}) > 0.25$  and  $\text{FDR} < 0.01$ .

## Looping analysis

Looping analysis was performed using the “loops” and “loops-diff” steps within the hic-bench platform<sup>24</sup>. Chromatin loops from Hi-C and H2K27ac HiChIP data were called using FitHiC2<sup>49,50</sup>. Loops were classified as diagnosis-specific, relapse-specific, or common. The common loops were also classified as increased, decreased, or stable using a log2 fold change cutoff ( $\text{abs}(\text{L2FC}) > 1$ ). A loop was considered sample-specific if it was found to be significant only in sample ‘x’ but not in sample ‘y’, when using a  $q$ -value cutoff of ‘qcut1’ (default: 0.01) as well as the loop shouldn’t be found significant in sample ‘y’ when using a more relaxed cutoff of ‘qcut2’ (default: 0.1). This is to avoid considering a loop sample-specific when the significance was relatively close in both samples but in one sample it was lower than the cutoff (significant) and in the other sample it was slightly higher than the cutoff (not significant). These cases were classified as common loops.

## Integration of patient Hi-C data with other datasets

To show the correlation between the Hi-C data and the various sequencing datasets, the peaks obtained from ATAC-seq, and the genes obtained from RNA-seq were mapped to the AB, BA, and stable regions or the increased, decreased, or stable TADs using the “bed-tools intersect” command<sup>89,90</sup>. We calculated the peak intensity fold change or gene expression fold change for peaks or genes assigned to a compartment region or TAD activity region between diagnosis and relapse and showed the correlation with boxplots or bar plots. Statistical significance was assessed using a paired two-sample  $t$  test.

## Reporting summary

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

## Data availability

All raw and processed data files generated through high-throughput sequencing for this publication (Hi-C, H3K27ac HiChIP, ATAC-seq, RNA-seq) has been deposited in the National Center for Biotechnology Information Gene Expression Omnibus (GEO; <http://www.ncbi.nlm.nih.gov/geo>) and are accessible through GEO Series accession number GSE222687. Previously generated high-throughput sequencing data (ChIP-seq) used in this publication is accessible through GEO Series accession number GSE156563.

## References

1. Nguyen, K. et al. Factors influencing survival after relapse from acute lymphoblastic leukemia: a Children’s Oncology Group study. *Leukemia* **22**, 2142–2150 (2008).
2. Carroll, W. L. & Raetz, E. A. Clinical and laboratory biology of childhood acute lymphoblastic leukemia. *J. Pediatr.* **160**, 10–18 (2012).
3. Hunger, S. P. et al. Improved survival for children and adolescents with acute lymphoblastic leukemia between 1990 and 2005: a report from the children’s oncology group. *J. Clin. Oncol.* **30**, 1663–1669 (2012).
4. Park, J. H. et al. Long-term follow-up of CD19 CAR therapy in acute lymphoblastic leukemia. *N. Engl. J. Med.* **378**, 449–459 (2018).
5. Meyer, J. A. et al. Relapse-specific mutations in NT5C2 in childhood acute lymphoblastic leukemia. *Nat. Genet.* **45**, 290–294 (2013).
6. Evensen, N. A. et al. MSH6 haploinsufficiency at relapse contributes to the development of thiopurine resistance in pediatric B-lymphoblastic leukemia. *Haematologica* **103**, 830–839 (2018).
7. Jones, C. L. et al. MAPK signaling cascades mediate distinct glucocorticoid resistance mechanisms in pediatric leukemia. *Blood* **126**, 2202–2212 (2015).
8. Yang, J. J. et al. Genome-wide copy number profiling reveals molecular evolution from diagnosis to relapse in childhood acute lymphoblastic leukemia. *Blood* **112**, 4178–4183 (2008).
9. Saint Fleur-Lominy, S. et al. Evolution of the epigenetic landscape in childhood B acute lymphoblastic leukemia and its role in drug resistance. *Cancer Res.* **80**, 5189–5202 (2020).
10. Aronson, B. E. et al. A bipartite element with allele-specific functions safeguards DNA methylation imprints at the Dlk1-Dio3 locus. *Dev. Cell* **56**, 3052–3065.e5 (2021).
11. Fortin, J. P. & Hansen, K. D. Reconstructing A/B compartments as revealed by Hi-C using long-range correlations in epigenetic data. *Genome Biol.* **16**, 180 (2015).
12. Dixon, J. R. et al. Topological domains in mammalian genomes identified by analysis of chromatin interactions. *Nature* **485**, 376–380 (2012).
13. Nora, E. P. et al. Spatial partitioning of the regulatory landscape of the X-inactivation centre. *Nature* **485**, 381–385 (2012).
14. Spielmann, M., Lupianez, D. G. & Mundlos, S. Structural variation in the 3D genome. *Nat. Rev. Genet.* **19**, 453–467 (2018).
15. Flavahan, W. A. et al. Insulator dysfunction and oncogene activation in IDH mutant gliomas. *Nature* **529**, 110–114 (2016).
16. Lupianez, D. G. et al. Disruptions of topological chromatin domains cause pathogenic rewiring of gene-enhancer interactions. *Cell* **161**, 1012–1025 (2015).
17. Kloetgen, A. et al. Three-dimensional chromatin landscapes in T cell acute lymphoblastic leukemia. *Nat. Genet.* **52**, 388–400 (2020).
18. Yusufova, N. et al. Histone H1 loss drives lymphoma by disrupting 3D chromatin architecture. *Nature* **589**, 299–305 (2021).

19. Narang, S. et al. NSD2 E1099K drives relapse in pediatric acute lymphoblastic leukemia by disrupting 3D chromatin organization. *Genome Biol.* **24**, 64 (2023).
20. Xu, J. et al. Subtype-specific 3D genome alteration in acute myeloid leukaemia. *Nature* **611**, 387–398 (2022).
21. Pui, C. H. et al. Treating childhood acute lymphoblastic leukemia without cranial irradiation. *N. Engl. J. Med.* **360**, 2730–2741 (2009).
22. Inaba, H. & Pui, C. H. Advances in the diagnosis and treatment of pediatric acute lymphoblastic leukemia. *J. Clin. Med.* **10**, 1926 (2021).
23. Inaba, H. & Mullighan, C. G. Pediatric acute lymphoblastic leukemia. *Haematologica* **105**, 2524–2539 (2020).
24. Lazaris, C., Kelly, S., Ntziachristos, P., Aifantis, I. & Tsirigos, A. HiC-bench: comprehensive and reproducible Hi-C data analysis designed for parameter exploration and benchmarking. *BMC Genomics* **18**, 22 (2017).
25. Wang, X. et al. Genome-wide detection of enhancer-hijacking events from chromatin interaction data in rearranged genomes. *Nat. Methods* **18**, 661–668 (2021).
26. Akers, N. K., Schadt, E. E. & Lolic, B. STAR Chimeric Post for rapid detection of circular RNA and fusion transcripts. *Bioinformatics* **34**, 2364–2370 (2018).
27. Dixon, J. R. et al. Integrative detection and analysis of structural variation in cancer genomes. *Nat. Genet.* **50**, 1388–1398 (2018).
28. Wang, S. et al. HiNT: a computational method for detecting copy number variations and translocations from Hi-C data. *Genome Biol.* **21**, 73 (2020).
29. Robinson, J. T. et al. Juicebox.js provides a cloud-based visualization system for Hi-C data. *Cell Syst.* **6**, 256–258.e1 (2018).
30. Zhang, Y. et al. PDGFRB mutation and tyrosine kinase inhibitor resistance in Ph-like acute lymphoblastic leukemia. *Blood* **131**, 2256–2261 (2018).
31. Pui, C. H., Roberts, K. G., Yang, J. J. & Mullighan, C. G. Philadelphia chromosome-like acute lymphoblastic leukemia. *Clin. Lymphoma Myeloma Leuk.* **17**, 464–470 (2017).
32. Devoucoux, M. et al. Oncogenic ZMYND11-MBT1 fusion protein anchors the NuA4/TIP60 histone acetyltransferase complex to the coding region of active genes. *Cell Rep.* **39**, 110947 (2022).
33. Feng, Y., Li, L., Du, Y., Peng, X. & Chen, F. E2F4 functions as a tumour suppressor in acute myeloid leukaemia via inhibition of the MAPK signalling pathway by binding to EZH2. *J. Cell Mol. Med.* **24**, 2157–2168 (2020).
34. Ouyang, X., Handoko, S. D. & Kwok, C. K. CScore: a simple yet effective scoring function for protein-ligand binding affinity prediction using modified CMAC learning architecture. *J. Bioinform Comput Biol.* **9**, 1–14 (2011).
35. Smith, C. A. & Fan, G. The saga of JAK2 mutations and translocations in hematologic disorders: pathogenesis, diagnostic and therapeutic prospects, and revised World Health Organization diagnostic criteria for myeloproliferative neoplasms. *Hum. Pathol.* **39**, 795–810 (2008).
36. Mullighan, C. G. et al. JAK mutations in high-risk childhood acute lymphoblastic leukemia. *Proc. Natl Acad. Sci. USA* **106**, 9414–9418 (2009).
37. Tirado, C. A. et al. Novel JAK2 rearrangement resulting from a t(9;22)(p24;q11.2) in B-acute lymphoblastic leukemia. *Leuk. Res.* **34**, 1674–1676 (2010).
38. Waanders, E. et al. Mutational landscape and patterns of clonal evolution in relapsed pediatric acute lymphoblastic leukemia. *Blood Cancer Discov.* **1**, 96–111 (2020).
39. Lee, J. W. et al. DACH1 regulates cell cycle progression of myeloid cells through the control of cyclin D, Cdk 4/6 and p21Cip1. *Biochem Biophys. Res Commun.* **420**, 91–95 (2012).
40. Chen, Y., Xiang, H., Zhang, Y., Wang, J. & Yu, G. Loss of PCDH9 is associated with the differentiation of tumor cells and metastasis and predicts poor survival in gastric cancer. *Clin. Exp. Metastasis* **32**, 417–428 (2015).
41. Lv, J. et al. PCDH9 acts as a tumor suppressor inducing tumor cell arrest at G0/G1 phase and is frequently methylated in hepatocellular carcinoma. *Mol. Med Rep.* **16**, 4475–4482 (2017).
42. Wang, C. et al. Dual inhibition of PCDH9 expression by miR-215-5p up-regulation in gliomas. *Oncotarget* **8**, 10287–10297 (2017).
43. Wang, C. et al. Downregulation of PCDH9 predicts prognosis for patients with glioma. *J. Clin. Neurosci.* **19**, 541–545 (2012).
44. Xiao, X. et al. The Gene Encoding Protocadherin 9 (PCDH9), a Novel Risk Factor for Major Depressive Disorder. *Neuropsychopharmacology* **43**, 1128–1137 (2018).
45. Chen, E. Y. et al. Enrichr: interactive and collaborative HTML5 gene list enrichment analysis tool. *BMC Bioinforma.* **14**, 128 (2013).
46. Kuleshov, M. V. et al. Enrichr: a comprehensive gene set enrichment analysis web server 2016 update. *Nucleic Acids Res.* **44**, W90–W97 (2016).
47. Xie, Z. et al. Gene Set Knowledge Discovery with Enrichr. *Curr. Protoc.* **1**, e90 (2021).
48. Hogan, L. E. et al. Integrated genomic analysis of relapsed childhood acute lymphoblastic leukemia reveals therapeutic strategies. *Blood* **118**, 5218–5226 (2011).
49. Kaul, A., Bhattacharyya, S. & Ay, F. Identifying statistically significant chromatin contacts from Hi-C data with FitHiC2. *Nat. Protoc.* **15**, 991–1012 (2020).
50. Ay, F., Bailey, T. L. & Noble, W. S. Statistical confidence estimation for Hi-C data reveals regulatory chromatin contacts. *Genome Res.* **24**, 999–1011 (2014).
51. Sheffield, N. C. & Bock, C. LOLA: enrichment analysis for genomic region sets and regulatory elements in R and Bioconductor. *Bioinformatics* **32**, 587–589 (2016).
52. Frago, R. et al. VEGFR-1 (FLT-1) activation modulates acute lymphoblastic leukemia localization and survival within the bone marrow, determining the onset of extramedullary disease. *Blood* **107**, 1608–1616 (2006).
53. Chow, Y. P., Alias, H. & Jamal, R. Meta-analysis of gene expression in relapsed childhood B-acute lymphoblastic leukemia. *BMC Cancer* **17**, 120 (2017).
54. Oshima, K. et al. Mutational and functional genetics mapping of chemotherapy resistance mechanisms in relapsed acute lymphoblastic leukemia. *Nat. Cancer* **1**, 1113–1127 (2020).
55. MN1 Linked to syndrome characterized by craniofacial abnormalities and severe developmental delay. *Am J Med Genet A* **182**, 615–616 (2020).
56. Carella, C. et al. MN1 overexpression is an important step in the development of inv(16) AML. *Leukemia* **21**, 1679–1690 (2007).
57. Kang, S. & Hong, S. Molecular pathogenesis of spinocerebellar ataxia type 1 disease. *Mol. Cells* **27**, 621–627 (2009).
58. Alluri, R. V. et al. Molecular analysis of CAG repeats at five different spinocerebellar ataxia loci: correlation and alternative explanations for disease pathogenesis. *Mol. Cells* **24**, 338–342 (2007).
59. Li, J. et al. ZMYND11-MBT1 induces leukemogenesis through hijacking NuA4/TIP60 acetyltransferase complex and a PWWP-mediated chromatin association mechanism. *Nat. Commun.* **12**, 1045 (2021).
60. Yuan, F. et al. A Novel Immune Classification for Predicting Immunotherapy Responsiveness in Patients With Adamantinomatous Craniopharyngioma. *Front Neurol.* **12**, 704130 (2021).
61. Jiang, Y. et al. [Study on expression and mechanism of serum differential proteins after rush immunotherapy of allergic rhinitis]. *Lin. Chung Er Bi Yan Hou Tou Jing Wai Ke Za Zhi* **34**, 683–689 (2020).



62. Reardon, D. A., Hanson, C. A., Roth, M. S. & Castle, V. P. Lineage switch in Philadelphia chromosome-positive acute lymphoblastic leukemia. *Cancer* **73**, 1526–1532 (1994).
63. Grosveld, G. C. MN1, a novel player in human AML. *Blood Cells Mol. Dis.* **39**, 336–339 (2007).
64. Relling, M. V. et al. Pharmacogenetic risk factors for osteonecrosis of the hip among children with leukemia. *J. Clin. Oncol.* **22**, 3930–3936 (2004).
65. Sherief, L. M. et al. Genetic polymorphism of vitamin D receptors and plasminogen activator inhibitor-1 and osteonecrosis risk in childhood acute lymphoblastic leukemia. *Mol. Genet. Genom. Med.* **9**, e1700 (2021).
66. Toth, K. et al. Retinoids enhance glucocorticoid-induced apoptosis of T cells by facilitating glucocorticoid receptor-mediated transcription. *Cell Death Differ.* **18**, 783–792 (2011).
67. Lee, Y. et al. ATXN1 protein family and CIC regulate extracellular matrix remodeling and lung alveolarization. *Dev. Cell* **21**, 746–757 (2011).
68. Safaei, S. et al. Double sword role of EZH2 in leukemia. *Biomed. Pharmacother.* **98**, 626–635 (2018).
69. Pierro, J. et al. The NSD2 p.E1099K Mutation Is Enriched at Relapse and Confers Drug Resistance in a Cell Context-Dependent Manner in Pediatric Acute Lymphoblastic Leukemia. *Mol. Cancer Res* **18**, 1153–1165 (2020).
70. Li, J. et al. PRC2 Inhibitors Overcome Glucocorticoid Resistance Driven by NSD2 Mutation in Pediatric Acute Lymphoblastic Leukemia. *Cancer Discov.* **12**, 186–203 (2022).
71. Martens, J. H. et al. PML-RAR $\alpha$ /RXR Alters the Epigenetic Landscape in Acute Promyelocytic Leukemia. *Cancer Cell* **17**, 173–185 (2010).
72. Churchman, M. L. et al. Efficacy of Retinoids in IKZF1-Mutated BCR-ABL1 Acute Lymphoblastic Leukemia. *Cancer Cell* **28**, 343–356 (2015).
73. Kaufmann, K. B. et al. A novel murine model of myeloproliferative disorders generated by overexpression of the transcription factor NF-E2. *J. Exp. Med.* **209**, 35–50 (2012).
74. Mullighan, C. G. et al. CREBBP mutations in relapsed acute lymphoblastic leukaemia. *Nature* **471**, 235–239 (2011).
75. Fu, Y. et al. The histone demethylase PHF8 promotes adult acute lymphoblastic leukemia through interaction with the MEK/ERK signaling pathway. *Biochem Biophys. Res. Commun.* **496**, 981–987 (2018).
76. Buenrostro, J. D., Wu, B., Chang, H. Y. & Greenleaf, W. J. ATAC-seq: A Method for Assaying Chromatin Accessibility Genome-Wide. *Curr. Protoc. Mol. Biol.* **109**, 21 29 1–21 29 9 (2015).
77. Dobin, A. et al. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15–21 (2013).
78. Liao, Y., Smyth, G. K. & Shi, W. featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics* **30**, 923–930 (2014).
79. Ramirez, F., Dundar, F., Diehl, S., Gruning, B. A. & Manke, T. deepTools: a flexible platform for exploring deep-sequencing data. *Nucleic Acids Res* **42**, W187–W191 (2014).
80. Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15**, 550 (2014).
81. Langmead, B. & Salzberg, S. L. Fast gapped-read alignment with Bowtie 2. *Nat. Methods* **9**, 357–359 (2012).
82. Tarasov, A., Vilella, A. J., Cuppen, E., Nijman, I. J. & Prins, P. Sambamba: fast processing of NGS alignment formats. *Bioinformatics* **31**, 2032–2034 (2015).
83. Zhang, Y. et al. Model-based analysis of ChIP-Seq (MACS). *Genome Biol.* **9**, R137 (2008).
84. Ross-Innes, C. S. et al. Differential oestrogen receptor binding is associated with clinical outcome in breast cancer. *Nature* **481**, 389–393 (2012).
85. Yu, G., Wang, L. G. & He, Q. Y. ChIPseeker: an R/Bioconductor package for ChIP peak annotation, comparison and visualization. *Bioinformatics* **31**, 2382–2383 (2015).
86. Imakaev, M. et al. Iterative correction of Hi-C data reveals hallmarks of chromosome organization. *Nat. Methods* **9**, 999–1003 (2012).
87. 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).
88. Haas, B. J. et al. Accuracy assessment of fusion transcript detection via read-mapping and de novo fusion transcript assembly-based methods. *Genome Biol.* **20**, 213 (2019).
89. Quinlan, A. R. & Hall, I. M. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* **26**, 841–842 (2010).
90. Quinlan, A. R. BEDTools: The Swiss-Army Tool for Genome Feature Analysis. *Curr. Protoc. Bioinforma.* **47**, 11 12 1–11 12 34 (2014).

## Acknowledgements

W.L.C has received grants from the National Cancer Institute of Health (R01CA140729 and R01CA260028), The Leukemia and Lymphoma Society Specialized Center for Research (7010-14), Perlmutter Cancer Center Arline and Norman M. Feinberg Pilot Grant for Lymphoid Malignancies, HHOW Scholar Hope Grant, and the Perlmutter Cancer Center (P30 CA016087). A.T. is supported by the NCI/NIH P01CA229086, NCI/NIH R01CA252239, NCI/NIH R01CA260028 and NIH/NCI R01CA140729. IA is supported by the NIH/NCI and the awards 1R01CA228135 and P01CA229086. We would like to thank the Genome Technology Center (GTC) for expert library preparation and sequencing, and the Applied Bioinformatics Laboratories (ABL) for providing bioinformatics support and helping with the analysis and interpretation of the data. GTC and ABL are shared resources partially supported by the Cancer Center Support Grant P30CA016087 at the Laura and Isaac Perlmutter Cancer Center. This work has used computing resources at the NYU School of Medicine High Performance Computing (HPC) Facility. We would also like to acknowledge the Children's Oncology group (COG) ALL biorepository which is supported by the NCTN Operations Center Grant (U10CA180886) and the COG Biospecimen Bank Grant (U24 CA196173).

## Author contributions

Conceptualization & study design, W.L.C., A.T., T.L., and I.A.; Investigation, S.N., Y.G. and N.A.E.; Formal analysis, S.N., N.M.; Writing, S.N., N.A.E., and Y.G.; Editing, D.T.T., E.A.R.; COG sample acquisition, D.T.T., E.A.R.; Experimental assays, N.A.E., Y.G., S.J., T.H.O. and M.W.; Supervision, N.A.E., W.L.C., A.T., T.L. and I.A.; Funding acquisition, S.N., N.A.E., D.T.T., E.A.R., W.L.C., A.T. and I.A.

## 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/s41467-024-51492-6>.

**Correspondence** and requests for materials should be addressed to Iannis Aifantis, Aristotelis Tsirigios or William L. Carroll.

**Peer review information** *Nature Communications* thanks Argyris Papanonis and the other, anonymous, reviewer(s) 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.



**Open Access** This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

© The Author(s) 2024