RESEARCH ARTICLE

Mutant EZH2 alters the epigenetic network and increases epigenetic heterogeneity in B cell lymphoma

Ofir Griess[1], Noa Furth[1], Nofar Harpaz[1], Nicoletta Di Bernardo[1], Tomer-Meir Salame[2], Bareket Dassa[3], Ioannis Karagiannidis[4], Yusuke Isshiki[4], Menachem Gross[5], Ari M. Melnick[4], Wendy Béguelin[4], Guy Ron[6]*, Efrat Shema[1]*

1 The Department of Immunology and Regenerative Biology, Weizmann Institute of Science, Rehovot, Israel, 2 Mass Cytometry Unit, Life Sciences Core Facilities, Weizmann Institute of Science, Rehovot, Israel, 3 Bioinformatics Unit, Department of Life Sciences Core Facilities, Faculty of Biochemistry, Weizmann Institute of Science, Rehovot, Israel, 4 Division of Hematology, Weill Cornell Medical College, New York, New York, United States of America, 5 Department of Otolaryngology-Head and Neck Surgery, Hadassah Medical Center, Jerusalem, Israel, 6 Racah Institute of Physics, Hebrew University of Jerusalem, Jerusalem, Israel

* guy.ron2@mail.huji.ac.il (GR), Efrat.shema@weizmann.ac.il (ES)

Abstract

Diffuse large B cell lymphomas and follicular lymphomas show recurrent mutations in epigenetic regulators; among these are loss-of-function mutations in KMT2D and gain-of-function mutations in EZH2. To systematically explore the effects of these mutations on the wiring of the epigenetic network, we applied a single-cell approach to probe a wide array of histone modifications. We show that mutant-EZH2 elicits extensive effects on the epigenome of lymphomas, beyond alterations to H3K27 methylations, and is epistatic over KMT2D mutations. Utilizing the single-cell data, we present computational methods to measure epigenetic heterogeneity. We identify an unexpected characteristic of mutant-EZH2, but not KMT2D, in increasing heterogeneity, shedding light on a novel oncogenic mechanism mediated by this mutation. Finally, we present tools to reconstruct known interactions within the epigenetic network, as well as reveal potential novel cross talk between various modifications, supported by functional perturbations. Our work highlights novel roles for mutant-EZH2 in lymphomagenesis and establishes new concepts for measuring epigenetic heterogeneity and intra-chromatin connectivity in cancer cells.

Introduction

Diffuse large B cell lymphoma (DLBCL) is the most common type of non-Hodgkin's lymphoma (NHL) [1,2]. Most of these cancers arise from germinal centers (GC), an immune compartment governed and maintained by the transcriptional repressor BCL6 [3–5]. GCs are transient structures in which B cells are rapidly dividing and

undergoing somatic hypermutation and class switch recombination of their immuno-globulin loci, rendering them inherently susceptible to malignant transformation [6,7]. Under normal conditions, the GC reaction is tightly controlled to ensure that GC B cells stop proliferating to differentiate and form high-affinity antibody-secreting cells. Genetic lesions (mainly 3q27 translocations) at the master regulator BCL6 gene locus result in constitutive expression of this gene [8], ultimately leading to B cells hyperplasia and lymphomagenesis [5,9–12].

Epigenetic deregulation is one of the hallmarks of cancer, and specifically DLBCL [13–15]. The chromatin landscape of cells, consisting of nucleosomes structures composed of DNA wrapped around an octamer of core histones, is decorated by an extensive array of post-transcriptional modifications (acetylation, methylation, phos-phorylation etc.) [16,17]. Precise regulation of these epigenetic modifications, which is governed by cell-type specific transcription factors (TFs) and chromatin regulators, is critical in driving differentiation programs and maintaining cell identity [18,19]. In DLBCL, 85% of patients carry at least one mutation in prominent regulators of the chromatin landscape [6]. The two genes that show the highest mutation frequency, as a single or double mutations, are EZH2 (20%–30% of patients), which methylates histone 3 lysine 27 (H3K27me1/2/3), and KMT2D (30% of patients), which deposits mono-methylation of H3K4 (H3K4me1) [6,14,20–24].

EZH2 is the catalytic unit of the Polycomb Repressive Complex 2 (PRC2), asso-ciated with transcriptional repression [25,26]. Normally, upon GC entry, B cells exhibit upregulation of EZH2, and it was shown to cooperate with BCL6 in GC formation [15,27,28]. Germinal center B cell subtype of DLBCL (GCB–DLBCL), as well as follicular lymphomas (FL) which are the second most common type of NHL, exhibit recurrent heterozygous somatic point mutations in EZH2. These mutations converge to substitutions of amino acids that are positioned in the catalytic methyl-transferase SET domain (mainly at position Y646, yet mutations at additional sites have been reported), thus altering enzymatic activity [29,30]. These gain-of-function (GOF) mutations alters EZH2 substrate preference from unmethylated lysine 27 to mono- and mainly di-methylated lysine 27 (H3K27me1/2), leading to accumula-tion of H3K27me3 at the expense of H3K27me1/2 [30,31]. EZH2 GOF mutations were shown to promote lymphoid transformation and contribute to lymphoma cells' proliferation and survival [27,28,32]. These mutations alter epigenetic regulation, by changing the topology and function of chromatin domains, inducing spreading of H3K27me3, and silencing of Polycomb target genes [27,28,33,34]. Moreover, mutant-EZH2 was shown to maintain the GC phenotype in DLBCL, potentially by mediating a significant increase in bivalent nucleosomes, marked by both the repres-sive H3K27me3 and the active tri-methylation of lysine 4 on histone H3 (H3K4me3) [27,28,35]. These bivalent chromatin domains mark critical GC B cell promoters, perhaps maintaining them in a repressed state.

KMT2D mutations, like EZH2 GOF, were also shown to enhance B cell proliferation and promote lymphomagenesis in mice. While loss of KMT2D was suggested to affect locus-specific H3K4 methylation and dampen the expression of tumor-suppressor genes, whether it affects global levels of H3K4me1 remains controversial [23,24]. A

recent study identified interaction between KMT2D and CREBBP, providing explanation for the frequent co-selection of inactivating both chromatin regulators in lymphoma [36]. Intriguingly, DLBCL patients also exhibit recurrent loss of KMT2D in combination with EZH2 GOF [20,37], yet the composite effects of these mutations on the epigenetic network and the cancer phenotype are not well understood.

DLBCL cells often carry mutations in several epigenetic regulators, rendering it challenging to isolate the effects of each mutation on the epigenome. In addition, accumulative evidence suggests extensive cross-talk and inter-relationships between various histone modifications [16,17,38–40]. This underscores the need to evaluate mutations in epigenetic regulators as potentially affecting the entire epigenetic network rather than a single histone modification. Here, we apply a single-cell approach, based on Cytometry by Time of Flight (CyTOF), to systematically explore the effects of EZH2 GOF and KMT2D mutations, alone or in combination, on a wide panel of epigenetic modifications. Our data reveal extensive and epistatic effects of mutant-EZH2 on the epigenetic and transcriptional network, and an uncharacterized oncogenic function for this mutation in increasing epigenetic heterogeneity. Interestingly, we show that the single-cell data can be leveraged to reconstruct and expose new putative interactions within the epigenetic network. Finally, we demonstrate a non-linear relationship between H3K27me3 and the lymphoma master regulator BCL6. Overall, this work highlights the power of single-cell epigenetic analysis to reveal novel principles of epigenetic deregulation in cancer.

Results

Single-cell epigenetic analysis reveals systematic alterations to various histone modifications mediated by EZH2 and KMT2D mutations

To systematically explore the effects of EZH2 Y646N GOF and KMT2D mutations on the epigenetic network, we employed a single-cell technology we recently developed that uses CyTOF to profile an extensive array of histone modifications in single cells [38] (Fig 1A). The epigenetic panel contained metal-conjugated antibodies targeting 16 histone modifications, three core histone proteins, cell cycle indicators, the proliferation marker Ki67 and the germinal center master regulator BCL6. We first analyzed two patient-derived xenografts (PDXs) and cell lines of GCB-DLBCL tumors, which either carry the Y646N mutant EZH2 allele in combination with KMT2D biallelic loss-of-function mutations (the PDX^MUT and the cell line OCI-Ly1) or carry two WT alleles for both genes (PDX^WT and the cell line OCI-Ly7). For each cell, the measurements of histone modifications were normalized to the levels of the core histones (see "Methods"), followed by dimensionality reduction using uniform manifold approximation and projection (UMAP) analysis. Both EZH2-mutant PDX and cell line clustered separately from the cells carrying WT EZH2, indicating differential epigenetic patterns in each line (Figs 1B, 1C, S1A, and S1B). Consistent with the known functions of EZH2 mutation, cells carrying it exhibited elevated levels of H3K27me3 concomitant with a reduction in H3K27me2 (Figs 1C, 1D, S1B, and S1C). While KMT2D mutations had no effect on global H3K4me1 levels in the PDXs, the OCI-Ly1 cells exhibited reduced levels of this modification in comparison to OCI-ly7 cells (S1D–S1E Fig). Finally, in both KMT2D-EZH2-mutant PDX and cell line we observed reduced levels of histone acetylation, perhaps due to an additional mutation in the histone acetyltransferase CREBBP (Figs 1C, 1D, and S1B–S1E).

To isolate the direct effects of EZH2 and KMT2D mutations, we utilized the patient-derived lymphoma cell line OCI-Ly7, which carry WT alleles of all the main epigenetic regulators that are known to be recurrently mutated in DLBCL [20,21,37]. We then generated in these OCI-Ly7 cells the EZH2 Y646N mutation, KMT2D biallelic loss-of-function mutation, or the combination of both mutations ('dMUT' for double-mutants), using CRISPR/Cas9 genome editing (S2A–S2C Fig). CyTOF analysis of the parental OCI-Ly7 cells and the different mutants revealed wide-spread epigenetic alterations mediated by EZH2 GOF mutation; while KMT2D-null cells clustered with the WT cells, most EZH2-mutant cells clustered separately (Fig 1E and 1F). Two independent clones of EZH2 GOF showed highly similar epigenetic phenotypes (Fig 1G). Interestingly, EZH2 GOF mutation showed a robust and epistatic phenotype over the KMT2D mutation, as visualized by the
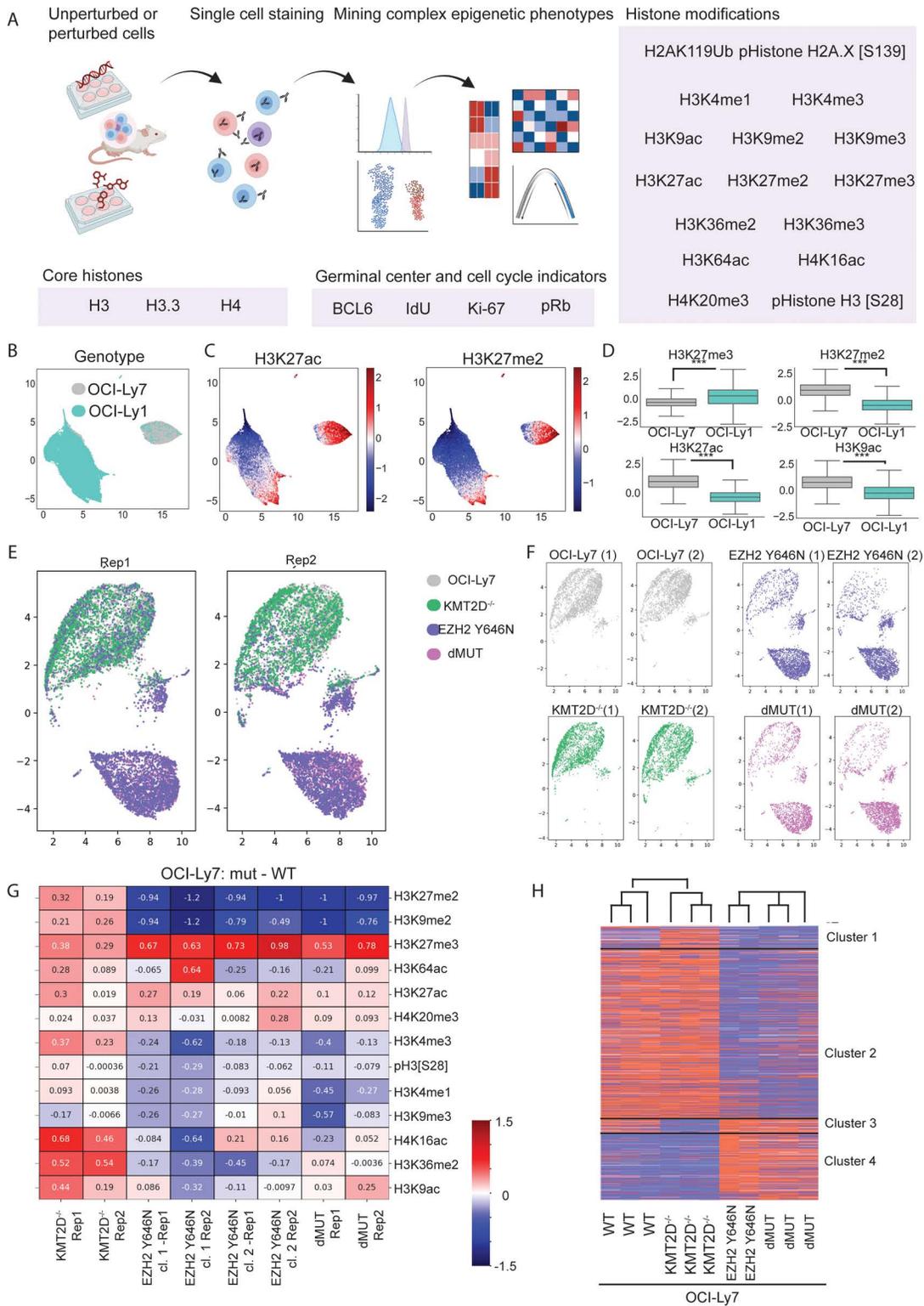
Fig 1. Mutant-EZH2 elicits a wide array of epigenetic alterations and shows an epistatic phenotype over KMT2D mutations. A. Scheme of the CyTOF experimental system. Human lymphoma patient-derived xenografts, or cell lines, unperturbed or perturbed genetically or chemically, were

stained with an epigenetic-oriented metal-conjugated antibody panel (see also S1 Table for list of antibodies and metals). From this single cell data, we mined complex epigenetic phenotypes exhibited by these cells. Drawn with Biorender. **B.** CyTOF analysis of the lymphoma patient-derived cell lines OCI-Ly7, carrying WT copies of EZH2 and KMT2D, and OCI-Ly1, carrying the EZH2 GOF mutation and biallelic loss of KMT2D. Colors indicate the sample index. **C.** Scaled, normalized levels of the indicated histone modifications on the UMAP of the cell lines that is shown in **B**. **D.** Expression levels of the indicated modifications in OCI-Ly1 and OCI-Ly7 cell lines, as measured by the CyTOF. P values were calculated by Welch's t test. ***p value < 0.001. **E–F.** OCI-Ly7 cells (WT), as well as their isogenic counterparts carrying mutant-EZH2 (EZH2 Y646N), biallelic knockout of KMT2D (KMT2D$^{-/-}$), or a combination of both mutations in EZH2 and KMT2D (dMUT), were analyzed by CyTOF. UMAP was performed based on all epigenetic marks measured in two independent biological repeats, following scaling and normalization. **E.** plotting of the indicated mutant cells on the joined UMAP in each of the repeats. **F.** Each line is shown separately. EZH2-mutant cells (alone or in combination with KMT2D loss) cluster together and separately from the WT cells, indicating an epistatic epigenetic phenotype for EZH2 GOF mutation. **G.** Heatmap showing the differences in the means of the indicated histone modifications in each of the indicated mutant lines compared to the WT OCI-Ly7 cells. Two biological CyTOF replicates are shown. Red: higher in the mutant cells compared to WT. Blue: lower in mutant cells. **H.** RNA-sequencing analysis of OCI-Ly7 cells and the indicated isogenic mutant lines. Heatmap shows the rld values of differentially expressed genes between the indicated mutant lines versus OCI-Ly7. Rows are standardized and clustered; red and blue denotes high and low levels of the modifications, respectively. The list of differentially expressed genes was obtained using DESeq2 and filtered for >1 or <−1 LogFC and p.adj ≤ 0.05. K-means clustering = 2. The data underlying this figure can be found in Raw data 1–2 at 10.17605/OSF. IO/NTGUX, under CyTOF and MARS-seq folders.

https://doi.org/10.1371/journal.pbio.3003191.g001

highly similar epigenetic patterns of the single EZH2 Y646N mutation and the double mutant containing both EZH2 Y646N and KMT2D loss. Importantly, in the double mutant the GOF mutation was generated on the background of the KMT2D null cells, thus it represents an independent transfection and a third mutant clone. This further highlights the robust and epistatic effects of EZH2 GOF on mediating these epigenetic alterations. Of note, while loss of KMT2D did not lead to prominent global epigenetic alterations, as these caused by EZH2 GOF mutation, and did not affect H3K4me1 levels, we did observe mild changes to various histone marks (Fig 1G). These results are in agreement with previous works, showing that KMT2D mutations mainly affect H3K4me1 and H3K4me3 genomic distribution rather than their global levels [24,41]. All mutants were analyzed by CyTOF in two biological repeats; cells from both repeats overlapped in the UMAP analysis and showed similar epigenetic patterns, confirming the quantitative nature of CyTOF in providing accurate measurements of histone modifications in hundreds of thousands of cells (Figs 1E, 1G, and S2D).

As expected, introducing mutant-EZH2 in this isogenic system resulted in elevated levels of H3K27me3 and reduction in H3K27me2, comparable to the levels observed in the EZH2-mutant line OCI-Ly1 (Figs 1G, S2C, and S2D). Interestingly, all cells expressing mutant-EZH2 showed robust loss of H3K9me2 (also seen in OCI-Ly1 cells, S1E Fig), similar in extent to the loss of H3K27me2, which is the direct substrate of mutant-EZH2 enzymatic activity. These cells also showed reduced H3K9me3 levels, yet to a lesser extent, and a slight increase in H3K27ac. Indeed, SHAP analysis revealed that multiple histone modifications, in addition to H3K27me2/3, contributed to the separation of the OCI-Ly7 WT and EZH2 Y646N cells into distinct clusters (S2E and S2F Fig). Interestingly, while loss of KMT2D had no effect on H3K4me1, it led to elevated levels of histone acetylation and a mild gain of H3K27me3, reminiscent of the EZH2 GOF mutants. These results suggest that the loss of histone acetylation observed in the PDX and OCI-Ly1 cells is likely attributed to loss of CREBBP, rather than reflecting epigenetic alterations directly mediated by EZH2 or KMT2D mutations. Importantly, the epigenetic phenotypes were validated using different methods (single-molecule imaging, FACS, western blot and image stream), and with different antibodies, confirming our results (S3 Fig). Taken together, the data reveal an epistatic role for mutant EZH2 in altering a wide array of histone modifications, beyond lysine 27 methylations, and indicate a potential cross talk between various epigenetic marks in this network.

We next aimed to explore whether these epigenetic phenotypes were reflected in gene expression changes. To that end, we performed RNA sequencing of the WT OCI-Ly7 cells as well as the isogenic mutant lines (Figs 1H and S4A–S4D). In agreement with the mild epigenetic changes, loss of KMT2D resulted in minimal transcriptional changes (cluster 1 and 3, corresponding to genes upregulated or downregulated in KMT2D$^{-/-}$ cells, respectively). Expression of EZH2 GOF mutation, however, led to robust transcriptional changes, with 702 genes downregulated (cluster 2) and 346 upregulated (cluster 4); the higher number of downregulated genes agrees with the function of H3K27me3 in transcriptional

repression. The independent clone of the double mutant of KMT2D$^{-/-}$ and EZH2 Y646N showed highly similar transcriptional phenotypes to that of the single EZH2 Y646N mutation (Figs 1H and S4A). These results indicate an epistatic role for mutant-EZH2 in mediating both transcriptional changes as well as epigenetic alterations.

Finally, exploring the list of up- and down-regulated genes revealed genes with interesting functions related to lymphomagenesis. For example, EZH2 mutation suppressed the expression of several genes that encode the linker histone H1 (S4E Fig). Recent studies showed that recurrent mutations in H1 genes, leading to reduced H1 levels, result in disruption of the 3D chromatin architecture and promote lymphomagenesis [42,43]. Thus, suppression of H1 expression by mutant EZH2 may benefit lymphoma cells and lead to a more aggressive phenotype. Furthermore, both EZH2 Y646N and the double mutant cells showed upregulation of B-Cell Receptor (BCR) and CDKN1A, reported to promote survival in germinal-center B cells [44,45]. We also observed downregulation of IL7R, which is associated with poorer prognosis in DLBCL [46].

Epigenetic CyTOF reveals a role for mutant-EZH2 in inducing epigenetic variability between cells to generate higher heterogeneity

The power of the CyTOF approach is that not only does it reveal global epigenetic changes in the population due to perturbations (as seen in Fig 1), but it also enables single-cell analysis of parameters such as the variability between cells and the heterogeneity in the population. To explore these concepts, we first calculated for each epigenetic mark the Gini coefficient, commonly used to measure the inequality among the values of a frequency distribution [47]. Interestingly, across most modifications the Gini index was higher in the EZH2-mutant cells compared to the WT, suggesting increased epigenetic heterogeneity (S5A Fig).

We next sought to define additional parameters that would reflect epigenetic heterogeneity in a population, leveraging all measurements rather than examining each modification separately. Thus, we generated a joint UMAP, based on all epigenetic modifications, of the OCI-Ly7 WT cells versus each of the isogenic mutants (Fig 2A). Heterogeneity following dimensionality reduction could potentially be manifested in different aspects (Fig 2B). We could consider global heterogeneity, as measured by (i) the global area each population of cells occupies (i.e., the area enclosed by the bounding convex hull of all points in the distribution, Fig 2Bi); and (ii) the global distance between each two cells in the population, defined by the 95th quantile of the distribution of distances between any two points in the distribution (Fig 2Bii). Another meaningful value is the local heterogeneity, reflecting whether cells in each cluster are relatively homogenous and form a tight cluster, or show greater variability to form a dispersed cluster. This can be measured by the area of the latent space occupied by the cells in a population (Fig 2Biii). Strikingly, in all these parameters, EZH2-mutant cells exhibited higher local and global epigenetic heterogeneity compared to their WT counterparts, reproducible in all experimental repeats, and irrespectively of the UMAP parameters that were used (Figs 2C and S5B). This higher heterogeneity was observed in both the EZH2 Y646N single mutant clones, as well as in the third independent clone of cells comprising the double mutant (EZH2 Y646N+KMT2D$^{-/-}$), but was not seen in cells carrying only KMT2D biallelic loss. Finally, we also calculated the nearest-neighbor distance (NND), which refers to the distance between each point and its nearest neighbor (Fig 2Biv). If a population is more "clumped", the distribution function for the NND will peak at shorter distances. Plotting the ratio of the cumulative distribution functions (CDFs) of EZH2-mutant cells versus WT cells revealed values below 1, indicating higher heterogeneity (corresponding to higher NND values) in EZH2-mutant cells (Fig 2D). Of note, the increase in local and global epigenetic heterogeneity was also observed in EZH2-mutant cells when excluding H3K27me2/3 from the UMAP (S5C Fig). This indicates that the effect of EZH2 GOF on epigenetic heterogeneity is mediated via multiple modifications, and not solely through H3K27 methylations.

To further explore this phenomenon of increased heterogeneity, we examined whether it could be recapitulated in the additional EZH2-mutant systems- the PDXs and the endogenously mutant cell line OCI-Ly1. Indeed, OCI-Ly1 carrying the mutant allele EZH2 Y646N showed higher heterogeneity in all parameters compared to OCI-Ly7 (Figs 1B and 2E). Similarly, the PDX^{MUT} was more locally heterogenous than the PDX^{WT}, and occupied a larger global area, although the maximal distance between cells was smaller (S5D Fig).
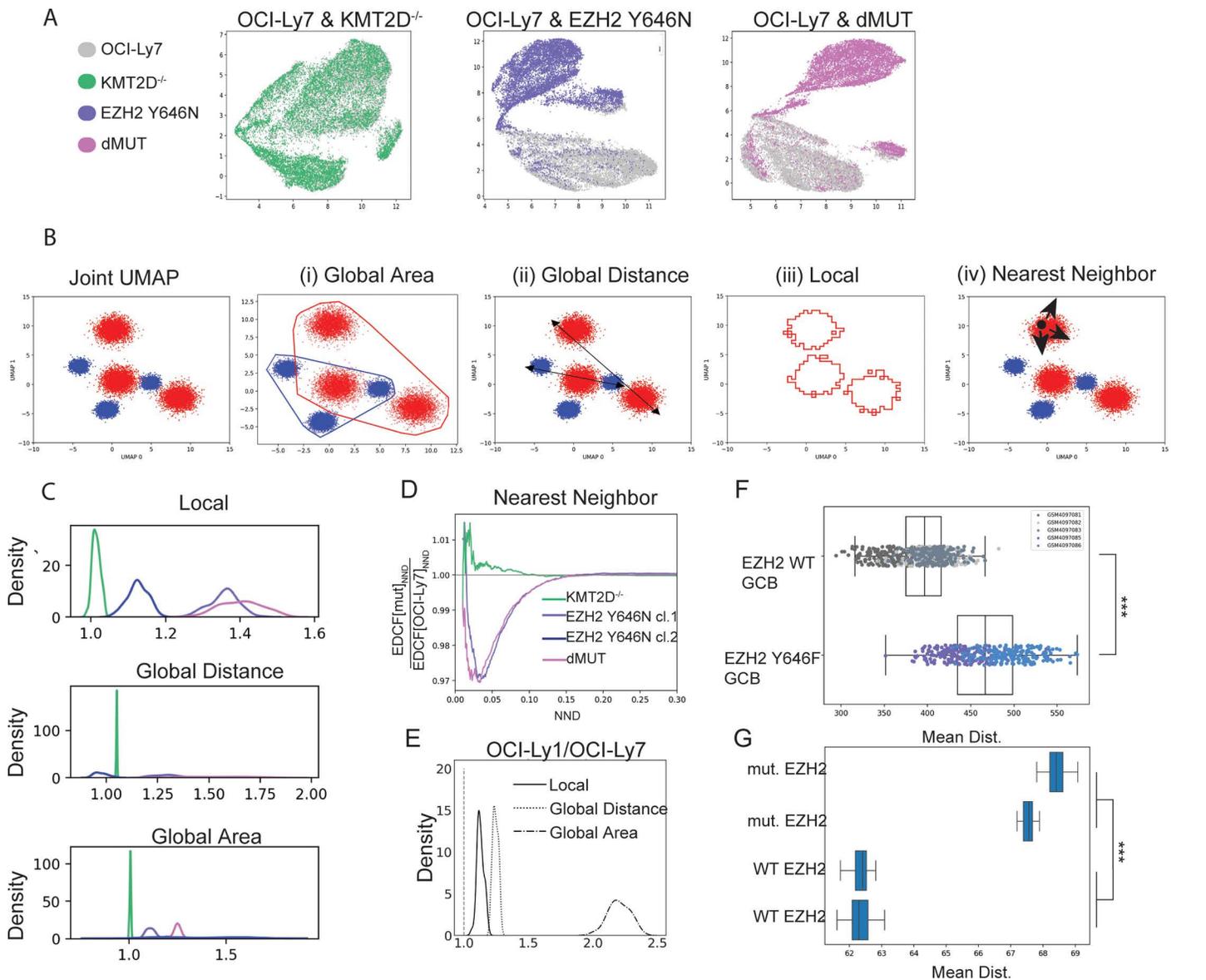
Fig 2. EZH2 Y646N mutation contributes to higher epigenetic heterogeneity. A. OCI-Ly7 cells and the indicated isogenic mutants (KMT2D$^{-/-}$, EZH2 Y646N, and the double mutant 'dMUT') were analyzed by CyTOF. Joint UMAP was performed for the WT OCI-Ly7 cells versus each of the three mutants, based on all epigenetic marks measured, following scaling and normalization. Colors indicate the sample index. **B.** Illustration of the various measurements of epigenetic heterogeneity. Shown is a hypothetical UMAP of two populations, marked in red and blue. The global heterogeneity of each population is reflected by: (i) Global area, as measured by the area enclosed by the bounding convex hull of all points in the distribution, and (ii) Global distance, measured by the distance between each two cells in the population, defined by the 95th quantile of the distribution of distances between any two cells. The local heterogeneity (iii) is measured by the net area of the clusters. If clusters are more dispersed, they would occupy a larger area. (iv) Nearest-neighbor distance (NND), which refers to the distance between each point and its nearest neighbor. For a population that is more homogenous, the distribution function for the NND will peak at shorter distances. **C.** The indicated heterogeneity measurements that were defined in panel **C** were calculated for OCI-Ly7 cells versus each of the mutant lines, across a wide range of UMAP parameters. Plotted are the ratios of the values of each mutant line versus OCI-Ly7 (WT). Thus, values over 1 indicate excess heterogeneity in the mutant over that of the WT. EZH2 mutant cells showed higher heterogeneity in all measurements. **D.** Ratio of the cumulative distribution functions (CDFs) of the NND values in the indicated mutant lines versus the WT cells. Values below 1 indicate higher heterogeneity of EZH2-mutant cells. **E.** Heterogeneity measurements calculated as in panel C for OCI-Ly7 cells carrying WT EZH2 versus OCI-Ly1 cells expressing the mutant form. Overall, OCI-Ly1 cells carrying EZH2 GOF show higher heterogeneity. **F.** Transcriptional heterogeneity of GCB cells from three mice expressing WT EZH2, versus two mice expressing the EZH2 Y646F GOF mutation. Sc-RNA-seq data is from Béguelin and colleagues [34] Heterogeneity was measured as the distribution of the mean distance between cells in each genotype (EZH2

GOF and WT), each color indicates a different mouse. To calculate the distribution, 300 cells were sampled from each sample and the mean Euclidean distance (based on the gene expression matrix) was calculated. The sampling was repeated 200 times for each sample. P value = 9e−132, as calculated by Welch's t test between the two distributions (from the EZH2 GOF and the WT mice). **G.** Transcriptional heterogeneity for two WT EZH2 and two mutant-EZH2 human FL tumors, showing increased heterogeneity in cells carrying EZH2 mutation. Heterogeneity was measured by the Euclidean pairwise distances, using the gene expression matrix, between each pair of cells in the sample. EZH2 WT and mutant tumors were matched according to grade (grades 2–3). All tumors selected for this analysis contain WT KMT2D and CREBBP. Sc-RNA-seq data is from Han and colleagues [48] P value was calculated by Welch's t test as in **F**. The data underlying this figure can be found in Raw data 1 at 10.17605/OSF.IO/NTGUX, under CyTOF folder.

Finally, we aimed to explore whether the observed increase in epigenetic heterogeneity is reflected in higher transcriptional heterogeneity. To that end, we leveraged two published datasets of single-cell RNA sequencing: (1) An isogenic mouse model with GC-specific expression of $Ezh2^{Y641F}$ from the endogenous Ezh2 locus, versus GCB cells carrying the WT Ezh2 gene [34]; and (2) Human FL tumors, with or without mutant-EZH2, that were profiled as part of a large single-cell RNA-seq study by Han and colleagues [48]. Strikingly, in both datasets we observed an increase in transcriptional heterogeneity, as measured by the mean distance between cells in each sample, in the mouse or human tumor samples expressing mutant-EZH2 compared to cells expressing the WT gene (Fig 2F and 2G). These results suggest that the increase in epigenetic heterogeneity might be linked to an increase in transcriptional heterogeneity, can be observed in diverse systems including clinical samples, and is likely relevant to human cancer. This increased heterogeneity is unique to EZH2 mutations and is not observed in the cells containing KMT2D knockout alone, suggesting a specific characteristic of EZH2 GOF mutations rather than a general effect of mutations in any chromatin regulator.

Mutant-EZH2 cells comprise distinct epigenetic subpopulations

Across all experiments of the isogeneic mutant-EZH2 and WT OCI-Ly7 cells, we observed that a fraction of cells expressing EZH2 Y646N clustered with the WT cells (Figs 3A, S6A and S6B). The percentage of these 'WT-like' cells was dynamic and varied between biological repeats yet was robustly detected in the two clones of the single EZH2 mutation and the double mutant with KMT2D$^{-/-}$ (S6C Fig). These 'WT-like' cells seemed to reflect an intermediate epigenetic phenotype, with some modifications showing similar levels to WT cells, while others were like the EZH2-mutant cells that strongly manifest the mutation and cluster separately (EZH2 Y646N 'Extreme', Figs 3B and S6D). For example, while 'WT-like' cells had intermediate levels of H3K27me2, indicating some activity of the mutant enzyme, their H3K27me3 levels were identical to cells that did not express mutant-EZH2.

To further explore these two subpopulations within EZH2-mutant cells, we compared their epigenetic patterns across different CyTOF experiments (Fig 3C). These data revealed that 'WT-like' cells consistently had elevated levels of the heterochromatin marks di- and tri-methylation of histone H3 on lysine 9 (H3K9me2/3) and lower levels of H3K27ac, perhaps indicating a more compact chromatin state. While speculative, it might suggest that mutant EZH2 activity could be influenced by a pre-existing chromatin state, thus generating heterogeneity. Alternatively, these two subpopulations may reflect transitions on a differentiation axis. Importantly, these 'WT-like' cells were observed throughout the cell cycle, indicating they do not correspond to cells in a specific cell cycle phase (S6E Fig).

The two subpopulations comprising EZH2-mutant cells can also be seen in the bimodal distribution of H3K27me3 levels; most mutant cells exhibited elevated H3K27me3 levels, and a minor fraction of cells exhibited 'normal' H3K27me3 levels (Fig 3D). Similar result was obtained by image stream analysis of fluorescent H3K27me3 signals, validating the CyTOF data (S6F Fig). Of note, the H3K27me3 distribution of WT OCI-Ly7 cells or cells knocked out for KMT2D had a single peak. To examine whether this phenomenon is also present in additional patient-derived cells that endogenously carry mutant-EZH2, we analyzed the PDX^{MUT}, as well as OCI-Ly1. Indeed, both models showed bimodal distribution of H3K27me3 (S6G and S6H Fig). These data suggest that, across different systems, EZH2-mutant cells comprise distinct subpopulations which differ in H3K27me3 levels.
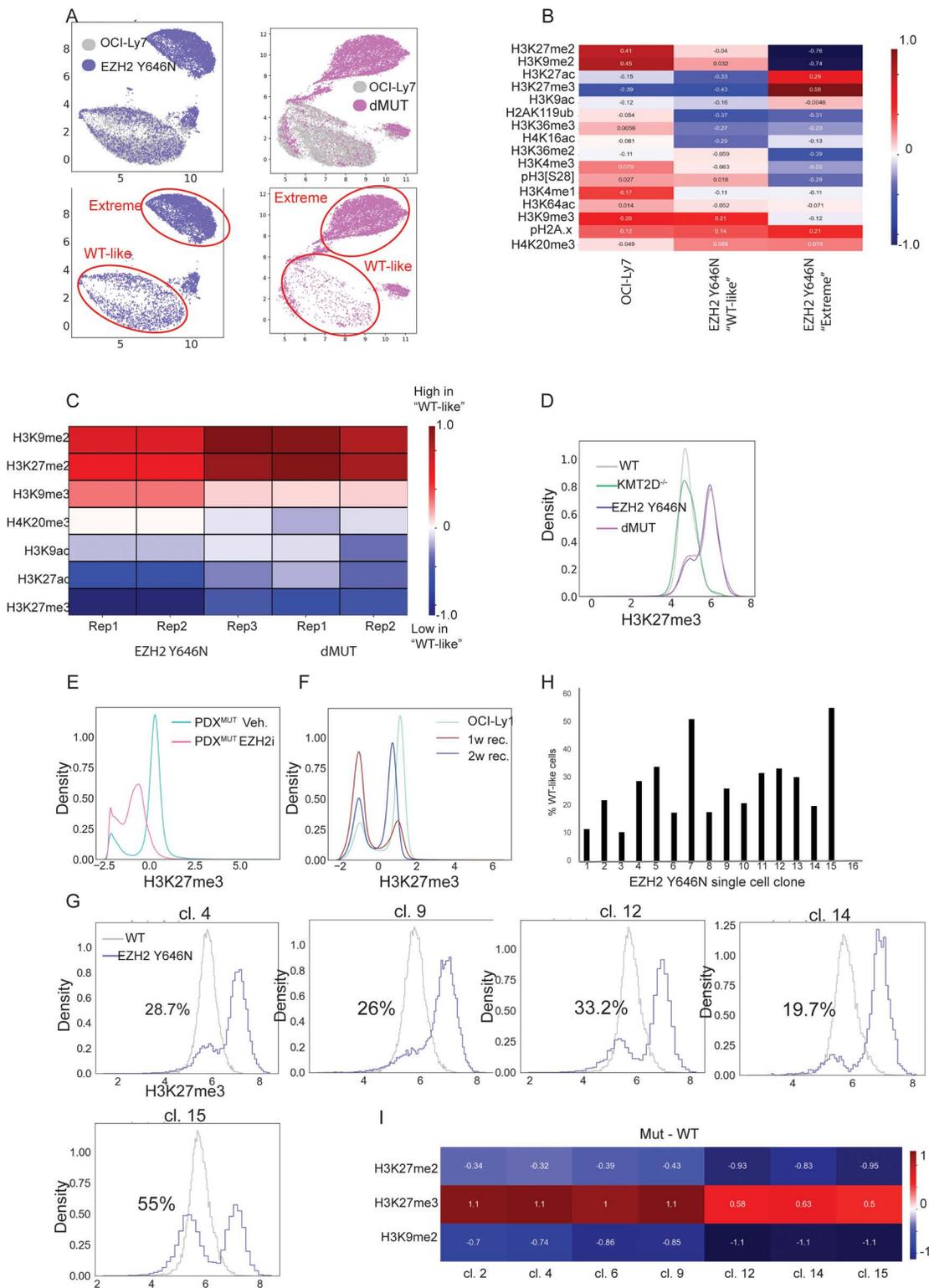
Fig 3. EZH2-mutants exhibit two distinct epigenetic populations. A. Top: Joint UMAP of OCI-Ly7 WT cells and the isogenic EZH2 Y646N cells (Left) or dMUT cells (Right). Colors indicate the sample index. Bottom: Only mutant-EZH2 cells are plotted, to highlight the subpopulation of mutant cells that cluster with WT cells, referred to as 'WT-like'. The subpopulation of cells that cluster separately from WT is referred to as EZH2 Y646N 'extreme' or

dMUT 'extreme', respectively. **B.** Heatmap showing standardized mean values measured by CyTOF for the indicated histone modifications in OCI-Ly7 cells expressing WT EZH2 ('WT'), and in the two populations identified in the EZH2 Y646N cells (marked in panel **F**), which either manifest robustly the GOF phenotype and form a distinct cluster (EZH2 Y646N 'Extreme'), or clusters with OCI-Ly7 WT cells (EZH2 Y646N 'WT-like'). **C.** Heatmap showing the differences in the means of the indicated histone modifications of the 'extreme' population compared to the 'WT-like' sub-population'. Shown are three biological CyTOF replicates for the single EZH2 Y646N mutant line, and two biological replicates for the double mutant. 'WT-like' cells show higher levels of the heterochromatin marks H3K9me2 and H3K9me3, and reduced levels of histone acetylations. **D.** Histogram of scaled and normalized H3K27me3 levels, as measured by CyTOF, in OCI-Ly7 cells ('WT') or the indicated isogenic mutant lines. **E.** Histogram of scaled and normalized H3K27me3 levels, as measured by CyTOF, in the PDX^MUT from mice treated with EZH2 inhibitor for two weeks or with vehicle as control. The two H3K27me3 subpopulations are present in both control and treated mice. EZH2 inhibition shifted the fraction of cells in each population and reduced H3K27me3 levels in the 'high' population. **F.** OCI-Ly1 cells expressing mutant-EZH2 were treated with EZH2 inhibitor at a concentration of 10 μM for one week to completely deplete H3K27me3 (see S4K Fig). Next, the inhibitor was washed, and the cells were allowed to recover for one or two weeks. Shown is a histogram of H3K27me3 levels in untreated OCI-Ly1 cells, as well as the cells after one or two weeks of recovery. The two subpopulations of H3K27me3 are maintained throughout the experiment, yet the fraction of cells in each subpopulation shifts over time. **G.** EZH2 Y646N were sorted into single cells and allowed to proliferate to generate clones. Histogram of scaled and normalized H3K27me3 levels, as measured by CyTOF, in OCI-Ly7 cells (WT) or the indicated EZH2 Y646N single-cell-derived clones. The percentage of 'WT-like' cells is shown per clone. **H.** Bar plot depicting percentages of 'WT-like' cells in 16 EZH2 Y646N clones as in **G**. The 'WT-like' percentage was calculated based on H3K27me3 values of EZH2 Y646N overlapping with OCI-Ly7 H3K27me3 values. **I.** Heatmap showing the differences in the means of the indicated histone modifications in each of the indicated clones compared to the WT OCI-Ly7 cells. The data underlying this figure can be found in Raw data 1 at 10.17605/OSF.IO/NTGUX, under CyTOF folder.

We next explored whether manipulating H3K27me3 levels via inhibition of EZH2, or the H3K27me3 demethylase KDM6A/B, would affect the distribution of cells between the two subpopulations (S6I Fig). Indeed, KDM6A/B inhibition resulted in upregulation of H3K27me3, concomitant with a reduction in the fraction of 'WT-like' cells, from 7.1% of the population to 4.8% (S6J and S6K Fig). Next, we analyzed H3K27me3 single-cell distribution in the PDX^MUT treated in vivo for 14 days with the FDA-approved EZH2 inhibitor Tazemetostat (250 mg/kg PO, BID) [49,50], versus control (vehicle). Cells from vehicle-treated mice showed bimodal distribution of H3K27me3, as expected (Fig 3E). EZH2 inhibition mainly affected the subpopulation expressing the robust GOF phenotype, reducing H3K27me3 levels in these cells. In addition, the fraction of cells comprising each of the H3K27me3 subpopulations was altered towards 'WT-like' cells. This observation was reproducible in the OCI-Ly7 cells expressing mutatnt-EZH2; its inhibition showed greater effect on the EZH2 'extreme' cells compared to the 'WT-like' cells (S6L Fig). Taken together, these results imply higher sensitivity to EZH2 inhibitors of cells expressing robustly the GOF phenotype. Importantly, both subpopulations were maintained in control and EZH2-inhibitor treated cells, suggesting this heterogeneity is an inherent feature of EZH2-mutant cancer cells. Furthermore, we robustly depleted H3K27me3 levels in OCI-Ly1 cells (carrying mutant-EZH2) by prolonged EZH2 inhibition, and then allowed the cells to recover for one or two weeks (Figs 3F and S6M). Interestingly, both H3K27me3 subpopulations were reestablished following a week of recovery, with time-dependent increase in the proportion of cells comprising the H3K27me3-high subpopulation (Fig 3F).

The experiments above imply that cells expressing mutant-EZH2 can dynamically shift between the two states of 'WT-like' and 'EZH2 extreme' phenotypes. To prove that such shifts indeed occur, we isolated single cells from the EZH2 Y646N population and allowed them to recover and form clones, followed by analysis of their H3K27me3 levels. In 15 out of the 16 clones that were analyzed, we identified a bimodal distribution of H3K27me3, corresponding to the two subpopulations (Figs 3G, 3H and S7). Interestingly, the percentage of 'WT-like' cells varied between clones, ranging from 10.3% to 55%, indicating dynamic fluctuations. Globally, the clones exhibited reduced levels of H3K27me3 and H3K9me2, similar to the parental population (Fig 3B and 3I). These results indicate that the two subpopulations observed in cells expressing mutant-EZH2 are not stable; rather, cells can transition between the 'WT-like' and 'EZH2-extreme' phenotypes. Even following a harsh perturbation, EZH2-mutant cells regenerate the two H3K27me3 subpopulations, perhaps revealing an inherent and stable feature of EZH2-mutant tumors. It highlights the power of single-cell epigenetic analysis to uncover such phenomena.

Single-cell analysis reconstructs and reveals novel interactions within the epigenetic network

Our CyTOF data provide further support to the notion that various histone modifications are inter-connected to form a network, leading to a 'domino effect' of EZH2/H3K27me2/3 alterations on various additional histone marks. For each

modification, cells in the population show a distribution of values, from high to low levels. We hypothesized that this information can be leveraged to gleam potential interactions between modifications, by examining coupled behaviors of epigenetic marks across cells in the population. To that end, we applied three different computational models: (1) Gradient boosting algorithm (XGBoost) that allows us to determine the effect on each modification stemming from all other modifications, calculated using the SHAP algorithm [51]. In this analysis, we employed a Gradient Boosting regression; the value for each modification was fit using the values of all other modifications, allowing for complex (and potentially not linear) relationship between the various modifications (see "Methods"). This operation was repeated many times (>10,000) on a random subsample of the data, and the results are averaged to obtain an adjacency matrix (Fig 4A). (2) A metric we term "Probability of link": After discretizing the data and subsampling it, we constructed a Bayesian Network for the data as a directed acyclic graph (DAG, using BNLearn – https://github.com/erdogant/bnlearn/). In a DAG, each node represents a variable, and directed edges indicate the direction of influence from one variable to another, without forming any cycles. This acyclic property ensures that there are no feedback loops, allowing for clear interpretation of causal pathways. The operation was repeated >10,000 and the mean of each link across the adjacency matrices is calculated. Thus, this metric denotes the probability for each specific direct link to be extracted from the data (Fig 4B). (3) Partial correlations, calculated in the standard fashion; the partial correlation of modifications X and Y is the correlation of the residuals of X and Y, when regressed against all other modifications (Fig 4C).

Comparing the results of the different models, we observed high overlap in reconstructing several of the well-established cross-talks between histone modifications in the network (Fig 3A–3C and a biological repeat shown in S8A–S8C Fig). For example, all models predicted a connection between histone H3 acetylations on lysine residues 9, 27 and 64, in agreement with their mutual deposition and removal by CREBBP/p300 and histone deacetylases, respectively. Interestingly, we found stronger connections between H3K9 and H3K64 acetylations, compared to their connections with H3K27ac; this is perhaps due to additional regulatory pathways affecting lysine 27 acetylation, such as the Polycomb complex and H3K27me2, especially in DLBCL. Another prominent example was the predicted strong effect of H3K27me2 on H3K27me3; both are deposited by EZH2, the former serves as the substrate for the latter. Interestingly, this interaction is not symmetrical; while H3K27me2 strongly affects H3K27me3, H3K27me3 only mildly affects H3K27me2. This result is in line with the notion that the enzymatic deposition of histone methylations ('writing'), including H3K27me3, is more robust and dominant compared to their removal ('erasure'). In that case, the model would predict a stronger effect of H3K27me2 on H3K27me3 compared to the reverse interaction. To illustrate that point, we performed a Markov chain simulation of the H3K27 methylation and acetylation dynamics (S8D and S8E Fig). We picked transition probabilities for the stochastic matrix that roughly reproduced the methylation and acetylation populations in the cellular steady state, in particular, such that the demethylation probability is lower than the methylation probability. Indeed, we found that under these conditions, our model accurately predicted a stronger effect of H3K27me2 on H3K27me3 compared to the reverse direction.

The network analysis also predicted an effect of H3K27ac and H4K16ac on γH2Ax; this is in-line with single-molecule studies that showed enrichment of this DNA-damage mark on nucleosomes containing H3K27ac, as well as an increase in γH2Ax in cells treated with HDAC inhibitors [35]. Moreover, a recent study reported elevated γH2Ax levels in cells carrying a deletion of MOF, the enzyme that deposits H4K16ac [52]. An additional cross-talk reconstructed by this analysis was between H3K9 and H3K27 methylations. Previous studies reported physical and functional interactions between G9a, one of the enzymes that deposits H3K9me2, and EZH2 [53–55]. While these studies focused on the output of these interactions on H3K27me3 deposition, our analysis predicted a connection between H3K9me2 and H3K27me2 rather than H3K27me3 (Fig 4A–4C). While EZH2 is known to deposit both di- and tri-methylations, the majority of literature is focused solely on H3K27me3, due to its known regulatory function in silencing of gene expression. Surprisingly, our analyses suggest H3K27me2 as being a central node in the epigenetic network, affecting many more modifications in comparison to H3K27me3. All models predicted multiple marks affected by H3K27me2, such as H3K27ac, H2Aub, H3K36me3 and H3K4me1, while only H3K9me3 seemed to be preferentially affected by H3K27me3. H3K27ac is of particular interest,

Fig 4. Single-cell analysis uncovers potential interactions within the epigenetic network. A–C. Models to decipher interactions within the epigenetic network, applied to unperturbed OCI-Ly7 single-cell CyTOF data. **A.** XGBoost analysis. **B.** 'Probability of link' analysis. For both **A** and **B**, the *Y* axis indicates 'source' modification and *X* axis indicates its 'target' modification affected by the *Y* axis. **C.** Partial correlations between histone modifications.

Yellow blocks and arrows highlight connections that are discussed in the text. **D.** Left – Heatmap showing standardized mean values measured by CyTOF, for the indicated histone modifications, of unperturbed OCI-Ly7 cells compared to each of the following perturbations: EZH2 inhibition ('EZH2i'), over expression of WT EZH2 ('EZH2 OE'), or expression of EZH2 GOF ('EZH2 Y646N'). For EZH2i, cells were incubated with DMSO (control) or 10 μM of the EZH2 inhibitor Tazemetostat for 48 h. For EZH2 OE, cells were nucleofected with PBS (control) or 1 μg of plasmid expressing WT EZH2. For EZH2 Y646N, cells underwent mutagenesis for EZH2 using CRISPR/Cas9. Right – heatmap scoring the effect of the perturbation by showing the mean of the absolute change per each mark. Effect shown in descending order from high (orange) to low (white). **E.** Average coverage of Cut&Run reads of the indicated histone modifications, over H3K27me2 and H3K27me3 peaks, in OCI-Ly7 cells and the isogenic EZH2-mutant cells (EZH2 Y646N). **F–G.** XGBoost analysis done on CyTOF data of unperturbed germinal-center B cells (CD20⁺, CD38⁺), derived from tonsils, for two independent samples derived from different patients. Tonsils were dissociated to single cells followed by staining with the panel of metal-conjugated antibodies and CyTOF analysis. Yellow blocks and arrows highlight connections that are discussed in the text. The data underlying this figure can be found in Raw data 1, 3 at 10.17605/OSF.IO/NTGUX, under CyTOF and Cut and Run folders.

https://doi.org/10.1371/journal.pbio.3003191.g004

as there are multiple publications, spanning diverse biological systems, reporting a cross talk with H3K27me3 [56–58]. Yet, these studies manipulated components of the Polycomb complex or lysine 27 itself, and thus cannot rule out that the effect is mediated by H3K27me2, as predicted by our data. This central role for H3K27me2 in affecting various modifications might stem from the wide distribution of this mark in the genome, as opposed to H3K27me3, which is mainly restricted to gene promoters. Finally, in addition to reconstructing known interactions, our data revealed new connections yet to be explored. For example, H4K16ac, which is also known to be widespread across large genomic regions [59], especially in gene bodies, is predicted to affect H3K27me2 levels.

In order to experimentally validate the predicted connections, and to decouple the effects of H3K27me2 and H3K27me3, we manipulated EZH2 levels and function by three complementary means: (1) EZH2 enzymatic inhibition by Tazemetostat, leading to reduced levels of both H3K27me2 and H3K27me3; (2) Over expression of wild-type EZH2, leading to an increase in H3K27me2/3, and; (3) Expression of the gain-of-function EZH2 mutation, leading to opposing effects on these modifications: upregulation of H3K27me3 concomitant with downregulation of H3K27me2. We then scored the combined effect on all modifications in the three perturbations and ranked them in a descending order (Fig 4D). As expected, the modifications showing the largest effect were H3K27me2/3, the direct targets of EZH2. Next in ranking we observed H3K9me2, which positively correlated with the change in H3K27me2 levels rather than H3K27me3: it decreased upon EZH2 inhibition and expression of mutant-EZH2 yet increased in cells over expressing WT-EZH2. This result validates the known interaction between G9a and EZH2, in addition to providing strong support for the predicted cross-talk specifically with H3K27me2 rather than H3K27me3. H3K27ac levels also followed the change in H3K27me2, yet in an opposite direction: in both perturbations that led to loss of H3K27me2, we observed gain of H3K27ac, while in the cells gaining H3K27me2, H3K27ac levels were reduced. Finally, we also observed changes in H3K36me3 and H3K4me1, predicted to be affected by H3K27me2 in the various models. Of note, the perturbation experiments did not fully recapitulate the predicted interactions. For example, H2Aub was not affected by any of the perturbations, despite the known crosstalk between the PRC1 (which deposits H2Aub) and PRC2 complexes [60,61], as well as the predicted link with H3K27me2 observed in our models. Taken together, these functional data support an upstream position for H3K27me2 in the epigenetic network, and validate the utility of our single-cell data and computational approaches to reveal potential coupling between modifications.

To further explore the relationship between H3 lysine 27 modifications, we profiled the genomic distributions of H3K27me3, H3K27me2 and H3K27ac in the isogenic cells expressing WT or mutant-EZH2. We first examined the full correlation matrix between all marks, to assess the extent of genomic alterations mediated by the introduction of mutant-EZH2 in the isogenic cells (S8F Fig). H3K27ac showed the highest correlation between Ly7 WT and EZH2 Y646N, as expected. Interestingly, H3K27me2 showed the highest distance between the genotypes (even compared to H3K27me3); H3K27me2 genomic distribution in EZH2-mutant cells was more like H3K27me3 compared to its patterns in the WT cells. We next analyzed the peak distribution and shape of H3K27 methylations in the isogenic cells. H3K27me3 peaks in EZH2-mutant cells were wider and showed higher basal signal across the whole genome, in agreement with spreading of this mark by the mutant enzyme (Fig 4E). Concomitantly, H3K27me2 signal was significantly reduced over these regions.

Finally, we aimed to test whether the connections observed in B cell lymphoma represent the wiring of the epigenetic network in normal germinal center B cells, or the re-wiring of this network during the tumorigenic process. To that end, we repeated the single-cell epigenetic CyTOF analysis on tonsils from two individuals, removed during tonsillectomy surgery and stained immediately without any cell culturing. The CyTOF included antibodies marking the GCB cells (CD20+, CD38+), allowing us to identify and reconstruct the network analysis on normal B cells (Fig 4F and 4G show XGBoost analysis of the two individuals, and SG8 and S8H Fig show the 'probability of link'). Most of the connections observed in lymphoma were also seen in the normal GCB cells, including the links between histone acetylations, and the robust effect of H3K27me2 on various modifications, including H3K9me2, H3K36me3 and H3K4me1. Nevertheless, some connections were cancer-specific, such as the predicted effects of H3K27ac and H4K16ac on γH2Ax; thus, these connections in lymphoma may reflect deregulation of the DNA damage response in cancer rather than inherent wiring of the network. In addition, some new connections were observed in the normal cells and lost in lymphoma, such as a predicted effect of H3K36me2 on H3K9me3. Interestingly, all the connections were stronger in the normal cells compared to the lymphoma cells, as can be seen by comparing the heatmaps, noting the different scale (Fig 4A: scale between 0 and 0.1, versus Fig 4G and 4H: scale between 0 and 0.3). We speculate that this may reflect lowering of epigenetic barriers and decreased stability of the epigenetic network in cancer cells, perhaps contributing to increased heterogeneity.

Non-linear relationship between H3K27me3 and the GC master regulator BCL6

EZH2 and BCL6 are pivotal for the formation of the germinal-center response, and deregulation of these key factors by genetic lesions in BCL6 locus and EZH2 GOF mutations promote lymphomagenesis [9,27]. Upon perturbations of EZH2 (Fig 4D), we noticed that not only the various histone modifications were affected, but also the expression of BCL6. Counterintuitively to the expected cooperation of these two master regulators in lymphomagenesis, in our isogenic OCI-Ly7 cells, introducing EZH2 mutation led to downregulation of BCL6. This phenomenon was observed both at the protein level in our single-cell CyTOF data and by western blot (Figs 5A, S9A, and S9B), and at the mRNA level (S9C Fig). In agreement with a potential role for H3K27me3 in suppressing BCL6 expression, Cut&Run analysis for H3K27me3 revealed gain of H3K27me3 at the BCL6 locus in cells expressing mutant-EZH2 (Fig 5B). Moreover, in the EZH2-mutant cells, H3K27me3 and BCL6 showed mild negative correlation ($R = -0.16$, Pearson pairwise correlation measured by CyTOF). Interestingly, in the OCI-Ly7 cells expressing WT EZH2 we observed a positive correlation between these marks ($R = 0.15$). These seemingly contradicting observations, and the wealth of literature pointing to the importance of these factors in lymphoma, prompted us to further explore their relationship.

We first leveraged all the data from the multiple CyTOF runs for OCI-Ly7 WT cells, binning cells according to their H3K27me3 levels and plotting the associated BCL6 levels. Interestingly, we observed a non-linear relationship: low H3K27me3 levels are positively correlated with BCL6, yet high levels are negatively correlated (Fig 5C). This non-linear 'bell shape' relationship was more evident in the EZH2-mutant cells, presumably because they express a wider range of H3K27me3 levels; cells expressing low H3K27me3 (such as the 'WT-like' cells) showed positive correlation with BCL6, while most cells expressing higher H3K27me3 showed negative correlation (Fig 5D and 5E). This phenomenon was recapitulated in the PDX^MUT and PDX^WT models (S6D Fig).

To further validate these non-linear dynamics, we treated the OCI-Ly7 WT and EZH2-mutant cells with EZH2 inhibitor. Confirming our single cell data, in cells expressing WT EZH2, reducing H3K27me3 by EZH2 inhibition led to downregulation of BCL6 expression, at both protein (measured by CyTOF and western blot) and RNA levels (Figs 4F, S9E and S9F). On the contrary, EZH2 inhibition in cells expressing mutant-EZH2 led to upregulation of BCL6, indicating a negative correlation (Figs 4F, S9G and S9H). This upregulation was also observed in OCI-Ly1 cells carrying the EZH2 mutation (S9I–S9J Fig).
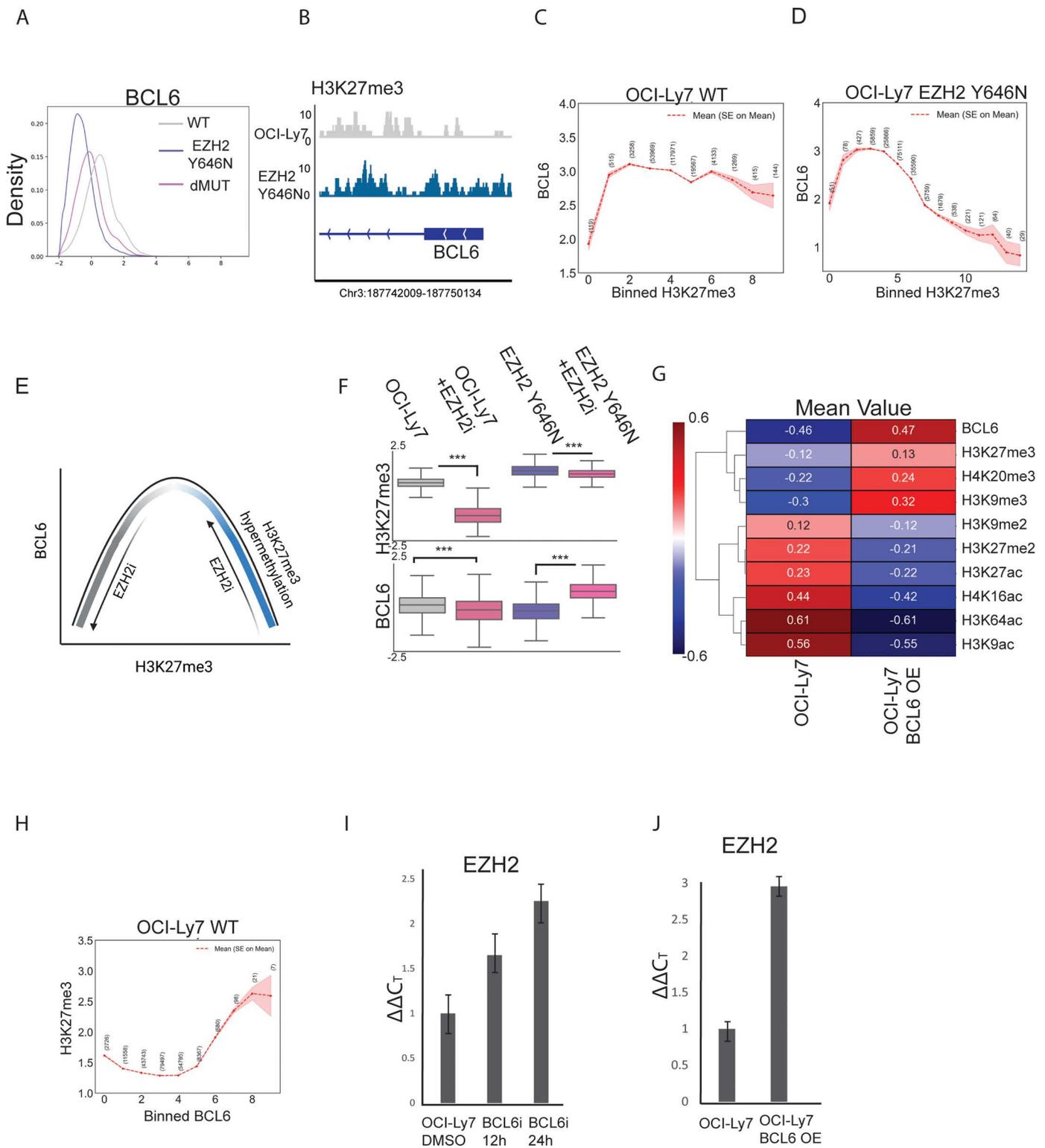
Fig 5. Single-cell analysis reveals non-linear relationship between H3K27me3 and BCL6. A. Histogram of BCL6 levels in OCI-Ly7 WT, EZH2 Y646N cells, and the double mutant of EZH2 Y646N and KMT2D$^{-/-}$, as measured by CyTOF. **B.** Normalized coverage values of H3K27me3 on the promoter of BCL6 gene, as measured by Cut&Run, in OCI-Ly7 and EZH2 Y646N. Chr3:187742009-187750134. **C–D.** BCL6 mean levels in cells binned

according to H3K27me3 levels. Red hue represents standard error of the mean. Number of cells in each bin is shown. **C.** OCI-Ly7 cells with WT EZH2. **D.** OCI-Ly7 cells with mutant-EZH2 (isogenic cell line). For this analysis measurements from four CyTOF experiments were pooled. H3K27me3 and BCL6 show a non-linear relationship. **E.** Model for the non-linear 'bell-shaped' relationship between H3K27me3 and BCL6. For low H3K27me3 levels, as in most OCI-Ly7 cells, it is positively correlated with BCL6. For high H3K27me3 levels, as in most EZH2-mutant cells, it is negatively correlated with BCL6. Thus, inhibiting EZH2 in the context of WT or mutant-EZH2 generates opposite effects on BCL6 levels. Generated with Biorender.com. **F.** Expression levels of H3K27me3 and BCL6, as measured by CyTOF, in the isogenic OCI-Ly7 WT and EZH2-mutant cells treated with EZH2 inhibitor for 48 h at a concentration of 10 µM. In OCI-Ly7 with low H3K27me3 levels, EZH2 inhibition downregulated BCL6. In the same cells expressing mutant-EZH2 and high H3K27me3, EZH2 inhibition led to upregulation of BCL6, validating a non-linear relationship. P values were calculated by Welch's t test. ***p value < 0.001. **G.** Heatmap showing mean values of OCI-Ly7 control and BCL6 overexpressing cells, for the indicated histone modifications. BCL6 overexpression results in an increase in H3K27me3 as well as additional marks of heterochromatin, and a decrease in histone acetylation levels. **H.** H3K27me3 mean levels in cells binned according to BCL6 levels. For this analysis measurements from four CyTOF experiments were pooled. Red hue represents standard error of the mean. Number of cells in each bin is shown. **I–J.** Quantitative RT-PCR analysis of EHZ2 expression. $\Delta\Delta C_T$ values relative to OCI-Ly7 ± s.d (n = 3) are shown. **I.** OCI-Ly7 cells were treated with 50 µM of the BCL6 inhibitor (BCL6i) FX1 for the indicated time points. **J.** OCI-Ly7 control and BCL6 over expression. Both treatments show an increase in EZH2 expression. The data underlying this figure can be found in Raw data 1, 6 at 10.17605/OSF.IO/NTGUX, under CyTOF and RT-PCR folders.

https://doi.org/10.1371/journal.pbio.3003191.g005

Finally, we tested whether manipulating BCL6 levels would affect epigenetic patterns, and specifically H3K27me3. Over expression of BCL6 in OCI-Ly7 cells led to global gain of the heterochromatin marks H3K9me3 and H4K20me3, concomitant with loss of histone acetylation, compatible with BCL6's role as a transcriptional repressor which recruits HDAC3 to silence genes [62] (Fig 5G). Examining specifically the effect of BCL6 on H3K27me3 in the single-cell data, we see the expected non-liner relationship (Fig 4H). In agreement with this data, both inhibition and overexpression of BCL6 led to upregulation of EZH2 and H3K27me3 levels (Fig 5G, 5I, and 5J). To conclude, our data highlights a complex non-linear relationship between BCL6 and EZH2-mediated deposition of H3K27me3. This relationship, which could only be revealed by the application of the single-cell approach, suggests a fine balance between these oncogenic pathways.

Discussion

In this study, we present the application of CyTOF to reveal novel aspects of EZH2 gain-of-function and KMT2D loss-of-function mutations in lymphoma. EZH2 mutation was shown to alter substrate preference of EZH2 from H3K27me0 to mainly H3K27me2, manifesting in robust upregulation of H3K27me3 and loss of H3K27me2 [31,63]. By establishing and analyzing isogenic lymphoma cells expressing WT or mutant-EZH2, we demonstrated the broad epigenetic phenotype mediated by this mutation, affecting a wide array of histone modifications other than lysine 27 methylations. For example, we observed a robust loss of H3K9me2 and gain of various histone acetylation marks in cells expressing mutant-EZH2. The isogenic system revealed that the EZH2 GOF mutation, in addition to generating a more repressed chromatin by increasing H3K27me3, also promoted loss of the 'close' histone H3K9 methylations, perhaps as a counter mechanism for the general repression. Moreover, the EZH2 GOF phenotype was epistatic over the biallelic loss of the epigenetic regulator KMT2D; cells carrying both KMT2D knockout and mutant-EZH2 showed epigenetic phenotypes highly similar to the cells carrying only the EZH2 mutation, and were epigenetically different from KMT2D$^{-/-}$ cells. Transcriptomic analysis revealed a similar epistatic role for EZH2 Y646N mutation, indicating that the global epigenetic alterations observed by CyTOF are functionally reflected in transcriptional changes. It is noteworthy that while KMT2D mutations are found across all DLBCL genetic subtypes, EZH2 mutations are unique to the GCB genetic subtypes, where they contribute to a unique disease phenotype, alluding to the epistatic nature of this mutation [21,22].

The cross talk between the trimethylation and acetylation of lysine 27 on histone H3 (H3K27me3 and H3K27ac) is of particular interest, and was studied in diverse biological systems, from stem cells to cancer [56,64,65]. In the OCI-Ly7 cells, introducing the genetic perturbation of EZH2 GOF mutation led to an increase in H3K27me3, which was accompanied with a mild gain of H3K27ac, suggesting a positive link between these marks. Yet, when OCI-Ly7 cells were treated with EZH2 inhibitor which reduced H3K27me3 levels, we also observed an elevation in H3K27ac levels. A similar negative correlation between H3K27me3 and H3K27ac was reported in stem cells with EZH2 knock-out, or cancer cells expressing

the H3-K27M mutation, in which loss of H3K27me3 led to elevated histone acetylations [56,64,65]. Interestingly, our network analysis, aimed at revealing direct links between histone marks, revealed that H3K27ac is tightly connected to H3K27me2, rather than H3K27me3. It is important to note that, while expression of mutant-EZH2 and application of EZH2 inhibitor exert opposite effects on H3K27me3 levels, both mediate a strong depletion of H3K27me2. Based on this data, and specifically the perturbation experiment in Fig 4D, it is tempting to speculate that the loss of H3K27me2 may be more prominent in mediating the gain of histone acetylation, rather than alterations in H3K27me3 levels. Of note, while our CyTOF panel did not include antibodies targeting the mono-methylation of histone H3 lysine 27 (H3K27me1), it was previously reported to also be downregulated in cells expressing mutant-EZH2 [34]. This downregulation is due to the loss of mono-methylation activity by mutant-EZH2, thus generating a dependency on the WT EZH2 protein that is co-expressed in these cells, yet only from one allele. It is plausible that mono- and di-methylations of lysine 27 may function in blocking acetylation of this residue in certain genomic regions. Thus, upon H3K27me1/2 downregulation, we see accumulation of H3K27ac signal.

Epigenetic heterogeneity and plasticity have recently emerged as inherent features of cancer cells, contributing to their persistence and aggressiveness, and allowing them to cope with hostile environments such as those encountered during treatment with anti-cancer drugs [66–68]. Yet, it remains elusive how this heterogeneity can be measured and quantified. In this work, we present several conceptual advances in leveraging single-cell epigenetic data to measure epigenetic heterogeneity. We devised computational methods to measure both local heterogeneity (i.e., within a subpopulation) and global heterogeneity (within the entire population). These methods identified a novel feature of EZH2 mutation in increasing epigenetic heterogeneity in diverse lymphoma models. Of note, loss of KMT2D did not promote heterogeneity, suggesting this is a unique feature of the EZH2 GOF mutation rather than a consequence of disruption of any epigenetic regulator. Importantly, analysis of single-cell RNA-sequencing data of human tumors, as well as an isogenic mouse model, revealed increased transcriptional heterogeneity in cells carrying EZH2 GOF. Thus, it is likely that epigenetic heterogeneity is linked to transcriptional heterogeneity, highlighting the relevance of our fundings to human tumors. This increased heterogeneity might contribute to the oncogenic functions of mutant-EZH2, providing yet another explanation for the selective advantage it confers to the cancer cells.

One manifestation of this heterogeneity in EZH2-mutant cells is the existence of a subpopulation that, despite the GOF mutation, exhibited an attenuated phenotype and lower levels of H3K27me3. This 'WT-like' subpopulation was robust and maintained across different lymphoma lines, as well as in-vivo in the xenograft model. In agreement with previous works [66,69–71], we provide evidence that the epigenetic subpopulations are not stable; rather, cells can transition between states, so that clones derived from single cells recapitulated the heterogeneity of the parental population (Figs 3G–3I and S7). Given that these 'WT-like' cells are characterized by hypoacetylation and elevated H3K9me2/3 levels, it is possible that in these cells there is lower genomic accessibility for the mutant enzyme to act on, resulting in a milder epigenetic phenotype. It may also be that these two subpopulations of EZH2-mutant cells, expressing low/high H3K27me3, reflect transitions on a differentiation axis or between light and dark zones programs. Yet another alternative possibility would be that these two subpopulations represent a dynamic balance between EZH2 and BCL6. Our single-cell data revealed a non-liner relationship between these key factors. In cells that express high BCL6, our data indicates that it represses EZH2 expression, and generates a more compact chromatin state; compatible with the phenotype of the 'WT-like' cells. On the contrary, cells with lower BCL6 would express higher levels of EZH2, manifesting a robust GOF phenotype. Further studies exploring the transcriptomics and probing for differentiation markers in these two subpopulations would be key in deciphering this heterogeneity. An interesting finding is that the two subpopulations show differential sensitivity to the FDA-approved EZH2 inhibitor (Figs 3E and S6L), and thus would likely have implications to cancer treatment.

Finally, our data strongly emphasize that perturbing a single epigenetic regulator, either by drug inhibition or by genetic manipulation, affects multiple downstream histone modifications. As various epigenetic drugs are in research use and in clinical trials for multiple cancers including lymphoma, these results highlight the need for systematic studies that view

epigenetic regulation as a network, rather than focusing on a single histone mark. We demonstrate the application of several models to reveal, and potentially predict, the interactions within this network. It would be of high interest to apply these methods to single-cell epigenetic data generated in diverse biological systems, elucidating whether these interactions are conserved and can be used for tailoring better anti-cancer approaches.

Methods

Ethics statement

For PDX samples. Ethics Board of the University of British Columbia. Six- to 8-week-old male NOD SCID IL2Rγnull (NSG) mice (The Jackson Laboratory) were used for the establishment of PDX models. All related experimental protocols were approved by the University of British Columbia Animal Care Committee and the Dana-Farber Cancer Institute Animal Care and Use Committee.

All clinical studies were approved by the appropriate national research ethics committee (Helsinki committee of the Hadassah Medical Center under protocol 0058-19-HMO),and have been performed in accordance with the ethical standards. Consent was obtained verbally by guardians.

Patient derived xenografts

Six to eight-week-old NSG mice were implanted subcutaneously in the flank with tumor fragments from a previously characterized DLBCL PDX [72]. Tumor formation was monitored weekly by inspection and calipers. Treatment began when tumors reached 600−700 mm^3. Mice were randomized to treatment with vehicle and Tazemetostat with 5 mice in each group. Tazemetostat was administered PO BID at 250 mg/kg in 0.5% NaCMC, 0.1% Tween-80 in dH2O by oral gavage. Mice were treated for 14 days total. Tumors were measured by caliper three times weekly. Mice were sacrificed when 14 days-treatment finished. All procedures performed were in accordance with Cornell University IACUC. For experiments in Fig 1, the untreated DLBCL PDX used were mutant EZH2^{Y646N} [72] and wild type EZH2 "DB-CJ" [73].

Cell culture conditions and generation of transgenic lines

Cell lines (OCI-Ly1 and OCI-Ly7) were a generous gift by Dr. Mark Minden (Ontario Cancer Institute, Toronto, Ontario, Canada) and were grown in Iscove's Modified Dulbeco Medium (IMDM) (Thermofisher: 21980032) supplemented with 20% Fetal Bovine Serum (Thermofisher: 10270106) and 1% Penicillin–Streptomycin (Biological Industries: 03-031-1B) at 37°C and 5% CO$_2$. Cells were maintained at density of 0.5–2 million cells/ml. Cells were routinely checked for Mycoplasma contamination.

EZH2 Y646N mutant lines were generated through nucleofection with Lonza 4D SF kit (V4XC-2032) of 0.5 million OCI-Ly7 cells with Alt-R CRISPR-Cas9, sgRNA and HDR sequence (IDT). Total nucleofected population was divided into single cell colonies. DNA extraction was done with boiling of pellets in 50 μL of 50 mM NaOH for 10 min at 99°C, then supplemented with 5 μL of Tris-HCl pH = 8. Template was used for PCR using Hy-Tiger mix (2×) (Hylabs: EZ2031), followed by cleanup using Nucleospin Gel Extract II/PCR (Macherey-Nagel: 740609.5). An amount of 20 ng clean template was analyzed by Sanger sequencing. Heterozygous clones were isolated.

BCL6 or EZH2 overexpressing cells were obtained by nucleofecting OCI-Ly7 with BCL6 overexpression plasmid pCXN2-BCL6 (Addgene: #40346) or EZH2 WT overexpression plasmid (Addgene: #173717), and compared to blank controls. Overexpression was tested using real-time PCR and Immunoblotting.

Human subjects

All clinical studies were approved by the appropriate national research ethics committee (Helsinki committee of the Hadassah Medical Center under protocol 0058-19-HMO), and have been performed in accordance with the ethical standards.

Tonsil processing and epigenetic profiling

Human inflamed tonsils were harvested (Helsinki committee of the Hadassah Medical Center under protocol 0058-19-HMO) and crushed into 70 μm mesh nylon for single cell separation. Fresh cells underwent cisplatin staining and extracellular staining as described in Cytometry by Time of Flight chapter. Cell identity profiles were done according to the tonsil atlas [74]. To avoid batch effect of intracellular staining of histone modifications, samples were fixed in 1.6% PFA and frozen. Then all samples were thawed and stained simultaneously.

Drug assays

A total of 1.5 million cells of OCI-Ly7 or EZH2 Y646N were incubated with DMSO, EPZ-6438 10 μM (Enco: 16174-5), GSK-J4 1 μM (Merck: 420205-10MG), FX1 50 μM (Databiotech: S8591 – 5 mg) for 12, 24 or 48 h. Inhibitor effects were validated by immunoblotting for H3K27me3 (EPZ-6438 or GSK-J4) or using real-time PCR for CXCR4 (FX1).

Immunoblotting

Whole cell lysates were prepared by suspension with 4× Laemmli sample buffer (Bio-rad: 1610747) and 50 μM DTT and repeated boiling at 95°C and vortexing. Diluted 1× sample was loaded onto 4%–20% Novex Tris-Glycine gel (Thermofisher: XP04205BOX) at 115 V for 1 h. Transfer was executed using semi dry Bio-rad Trans turbo blotter with 0.2 μm nitrocellulose membranes (Bio-rad: 1704158). Then, membrane was blocked with 5% non-fat milk and washed with TBS-T. Incubation of primary antibody was conducted overnight in 5% BSA solution supplied with 0.04% sodium azide. Incubation with secondary antibody was conducted for 1 h in 5% non-fat milk. Membranes were read with Clarity Western ECL Substrate (Bio-rad: 1705061). Antibodies were diluted according to the manufacturers' instructions.

RNA extraction and real-time PCR

RNA was extracted from 2 million cells using Nucleospin RNA isolation kit (Macherey-Nagel: 740955.50). Quality of RNA was assessed by running in agarose gel and measuring concentration in nanodrop. A total of 500 ng of RNA was used for first strand synthesis using M-MLV Reverse transcriptase (Promega: M1701). Real-time PCR was done by using KAPA SYBR FAST ABI Prism 2X qPCR Master Mix (KAPA Biosystems: KK4603). We normalized with the housekeeping gene HPRT and calculated $\Delta\Delta C_T$ values to evaluate differential expression.

MARS-seq and RNA expression analysis

Total 20 ng of RNA were used for library preparation according to a published protocol [75]. Three replicates per each genotype were used in the preparation. Libraries were then sequenced using a NextSeq 500/550 High Output Kit v2.5 on NextSeq500. Analysis was conducted using the previously published User-friendly Transcriptome Analysis Pipeline (UTAP) [76]. Differential expression was obtained by using DESeq2 [77]. Genes showing 2-fold change expression and p. adj values of ≤0.05 were considered differentially expressed.

Cut and Run sample preparation and data analysis

Cut and Run assay was executed as described in previous works [78]. OCI-Ly7 cells were harvested, with 200,000 cells taken per reaction. Permeabilized cells, bound to Concanavalin A-coated beads, were mixed with individual primary antibody (H3K27me3: CSTC36B11, H3K27me2: abcam24684, H3K27ac: CSTD5E4) and incubated overnight at 4°C. Secondary antibody, anti-mouse HRP (JIR 115-035-003), was used as a negative control. pAG-MNase enzyme (generated in the Department of Life Sciences Core Facilities, WIS, using Addgene plasmid 123461) was added to each sample followed by incubation step of 1 h at 4°C. Targeted digestion was done by 15 min incubation on ice block (0°C) under low salt conditions. DNA purification was done using Nucleospin gel and PCR clean-up kit (Machery-Nagel, 740609).

Sequencing was done on a Next-Seq 500 instrument (Illumina) using a V2 150 cycles mid output kit, allocating 10 M reads per sample (paired end sequencing).

Paired-end reads of each sample (approximately 15.7 M average reads per sample) were preprocessed with cut adapt [79], to remove adapters and low-quality bases (parameters: --times 2, -q 30, -m 20), following evaluation of quality with FastQC. Reads were mapped to human genome (hg38, UCSC) using Bowtie [80] version 2.4.5 (--local, --very-sensitive-local, --no-unal, --no-mixed, --no-discordant, --dovetail, -I 10, -X 700). Nucleosome fragments at the length >120 bp were selected from the remaining unique reads using picard-tools. Alignments were normalized based on spike-in scaling factor calculated by *E. coli* "carry-over" DNA from the pAG-Mnase, as reported by Meers and Henikoff [78,81]. Bigwig files were constructed from BAM alignments using deepTools2 suite [82], by 'bamCoverage' command, in 10 bp bins. Heatmaps and reads coverage around TSS was visualized using missing Data As Zero parameter. Peaks were called using SEACR [81] algorithm, with Empirical false discovery rate of 0.01.

Antibody conjugation

Antibodies (BSA and azide free) were conjugated to metals using the MIBItag Conjugation Kit (IONpath) according to the manufacturer's instructions.

Cytometry by time of flight (CyTOF)

Total of 6 million cells were harvested and washed with Maxpar CyTOF PBS (Fluidigm: 201058). Live/dead staining was conducted by resuspension of the cells with 1.25 μM Cisplatin (Fluidigm: 201064) in CyTOF PBS for 1 min. Cisplatin was neutralized with warm IMDM supplied with 10% FBS (Invitrogen: 10270106). Fixation and permeabilization was conducted by incubation of cells with Maxpar nuclear antigen staining buffer set (Fluidigm: 201063) according to the manufacturer's instructions. Samples were multiplexed by taking 3 million cells/sample and using the Cell-ID 20-Plex Pd Barcoding Kit (Fluidigm: 201060). After barcoding, 1 million cells from each sample were pooled to prevent batch effect and blocked using FBS for 10 min. Then, the pooled sample was directly incubated for 30 min at room temperature with the mixture of antibodies and washed with Maxpar Cell Staining Buffer (Fluidigm: 201068). Cells were fixed with fresh 4% PFA overnight. On acquisition day, samples were suspended with 4% PFA supplied with Cell-ID Intercalator-Iridium (Fluidigm: 201192A) at 125 nM. Cells were then washed with Maxpar Cell acquisition solution (Fluidigm: 201244) and suspended again in Maxpar Cell acquisition solution supplied with 10% EQ Four Element Calibration Beads (Fluidigm: 201078). Prior to the run on the CyTOF machine, cells were filtered through a 35 μM mesh cell strainer. Data was acquired on a Fluidigm Helios CyTOF system. The list of antibodies and conjugated metals is provided in S1 Table.

Single-cell RNA sequencing analysis

Processed aligned reads of scRNA-seq data of mice GCB cells expressing either WT EZH2 or EZH2 Y646F (SRP223373) were obtained from Béguelin and colleagues [34]. We omitted one EZH2 Y646F sample from our analysis, as the reads quantity was not comparable to the other samples [34]. In addition, we analyzed processed aligned reads of scRNA-deq data of FL tumors and cohort matched grades 2–3 tumors (FL-13, FL-30, FL-32, FL-51) which either carry WT or mutant EZH2. KMT2D and CREBBP profiles were WT (GSE203610) [48]. To measure heterogeneity, we calculated the Euclidean pairwise distances using the gene expression matrix between each pair of cells in the sample.

Single molecule profiling

A total of 2 million cells were harvested and subjected to nucleosome extraction followed by Mnase digestion as depicted in Fedyuk and colleagues [83]. Biotin and fluorophore-labelled antibodies were loaded onto a biotin-streptavidin surface. Nucleosomes were incubated with fluorophore-conjugated primary antibodies targeting histone modifications (depicted

in). Samples were incubated with secondary antibody for 30 min (abcam: ab150114 and ab150077). Overlap of nucleosome and modification was quantified using Cell Profiler.

Immunofluorescence

A total of 0.5 million cells were overlaid on a surface using Poly-L-Lysine 0.01% (SIGMA: P4707) for 30 min, then fixed using 4% PFA, and permeabilized using 0.1% Triton X-100 (SIGMA: T8787). Twenty percent BSA was used for blocking. Cells were stained with antibodies labelling histone modifications (S2 Table) or DAPI (Thermo: 62248). Quantification was done using Fiji ImageJ.

FACS analysis

Six million cells were harvested, fixed and permeabilized using pure chilled Methanol for 30 min. Blocking was done by incubating in BSA 20% for 30 min. Zombie Violet was used for live/dead staining (Biolegend: 423113). Cells were stained for histone modifications using fluorescent primary antibodies targeting histone modifications (S2 Table).

CyTOF data analysis

Data processing

CyTOF data underwent the following pre-processing prior to analyses: First, the CyTOF software by Fluidigm was used for normalization and concatenation of the acquired data. Then, several gates were applied using the Cytobank platform (Beckman Coulter): First, the normalization beads were gated out using the 140Ce channel. Then, live single cells were gated using the cisplatin 195Pt, iridium DNA label in 193Ir, event length, and the Gaussian parameters of width, center, offset and residual channels. CyTOF software was then used for samples de-barcoding [84].

Data manipulation and scaling

The data analysis pipeline follows the same general scheme presented in our previous work [38], with some modifications. Here we outline the steps taken.

Before beginning the analysis procedure, the cells were gated using the core histones (H3, H3.3, and H4). The gate allowed only cells with a minimum raw value of 5 for all core histones into the next steps of the analysis. For the PDX lines gating was also applied on the CD45 human versus mouse markers to remove all mouse cells.

For the various markers, a hyperbolic arcsine transforms (with a scale factor of 5) was first applied to the data:

$$M^{ij} = arcsinh\frac{M^{ij}}{5}$$

The data were then normalized to remove systematic effects using the measured values of the core histones:

$$M^{ij} = \frac{M^{ij}}{\left(1 + \alpha\, H^i_{3.3} + \beta\, H^i_4 + \gamma H^i_3\right)}$$

Where as in [34] the coefficients were selected to minimize the sum of the variances of the core histones, and from normalization $\gamma = (1 - \alpha - \beta)$. A reduction of a few dozen percent in the sum of variances was observed after the subtraction, indicating the removal of systematic effects.

Finally, the data were z-transformed:

$$M_{ij} = \frac{M^{ij} - \mu^i}{\sigma^i},$$

where in all the previous *i* denotes the observation, *j* denotes the column, μ^i and σ^i are the column mean and standard deviation, respectively. Z-transform scaling was done at the same time on samples that were acquired together.

Gradient boosting with SHAP values

For this analysis, the value for each modification was fit using the values of all other modifications, as following: to allow for the complex (and potentially not linear) relationship between the various modifications, we employed a Gradient Boosting regression, rather than, for example, a simple linear fit. Briefly, Gradient Boosting Regression is an ensemble learning technique that builds a predictive model in a stage-wise fashion by sequentially adding weak learners to minimize the loss function (here the root means square error). At each iteration, the algorithm fits a new decision tree to the negative gradient of the loss function computed from the residuals of the previous ensemble. This process effectively focuses on correcting the errors of prior models by emphasizing observations that were previously mispredicted. The ensemble of trees is combined to produce a strong predictive model that optimizes performance on the training data. Specifically, we used the XGBoost package, described by Chen and colleagues, 2016 [85]. To interpret the complex models generated by XGBoost, we utilized the SHAP (SHapley Additive exPlanations) algorithm [51], which quantifies the contribution of each feature (modification) to the model's predictions, assigning an importance value based on game-theoretic principles. By repeating this modeling process over 10,000 times on random subsets of the data, we ensured the stability and robustness of our results. Averaging the SHAP values across all iterations allowed us to construct an adjacency matrix that represents the average influence of each modification on every other modification.

Probability of link using Bayesian networks

In addition to the gradient boosting analysis, we also introduced a metric we term "Probability of Link". We construct a Bayesian Network for the data, which allows to model the data as a Directed Acyclical Graph – DAG. In a DAG, each node represents a variable, and directed edges indicate the direction of influence from one variable to another, without forming any cycles. This acyclic property ensures that there are no feedback loops, allowing for clear interpretation of causal pathways. Since Bayesian networks are designed to work with discrete data, we first discretize our continuous data set into 20 uniformly size bins for each feature, spanning the range from the maximal to minimal value of that feature. The DAG is then constructed using the python bnlearn package [86], designed to infer the graphical structure of Bayesian networks from a dataset, each edge in the graph indicates a possible direct influence of one modification on another. Since the DAG is acyclical, which is not necessarily the case for our data, and to account for the variability in our data, we repeatedly subsampled data and reconstructed the DAG over 10,000 iterations, which allowed each iteration to construct a different graph. We then calculated the probability of each link across all networks, resulting in a probability score for each possible direct interaction between modifications.

Markov chain simulation

We simulated the evolution of nucleosome modification populations over discrete Markovian time steps. The steady-state populations in our simulation represent the fraction of nucleosomes modified within a cell. To conduct the analysis, we modeled 5,000 cells, each containing 30 million nucleosomes, with initial modification fractions chosen randomly.

The cellular populations were propagated through 1,000 times steps using a stochastic matrix. Subsequently, XGBoost analysis was applied to the simulated cellular populations, focusing specifically on the H3K27 modifications. This subset was selected to reflect a key feature of the XGBoost analysis, as described. The entire simulation and analysis were repeated 500 times to ensure statistical robustness.

S8E Fig presents a histogram of the inferred interactions between H3K27me2 and H3K27me3 modifications, based on the XGBoost analysis. The results clearly indicate that the influence of H3K27me2 on H3K27me3 is substantially stronger than the reverse influence of H3K27me3 on H3K27me2.

Supporting information

S1 Fig. A. Two patient-derived xenografts; PDXWT which carries WT copies of EZH2 and KMT2D, and PDXMUT which carries the EZH2 Y646N mutation and biallelic loss of KMT2D, were analyzed by CyTOF. UMAP was performed based on all epigenetic marks measured, following scaling and normalization. Colors indicate the sample index. B. Scaled, normalized levels of the indicated histone modifications on the UMAP of the PDXs that is shown in A. C. Expression levels of the indicated modifications in PDXWT and PDXMUT, as measured by the CyTOF. P values were calculated by Welch's t test. ***p value < 0.001. D. Mean differences between PDXWT and PDXMUT for the indicated histone modifications. The mean values for PDXWT were subtracted from PDXMUT. Uncertainties were estimated using bootstrapping. E. Mean differences between OCI-Ly7 and OCI-Ly1 for the indicated histone modifications. The mean values for OCI-Ly7 (WT) were subtracted from OCI-Ly1 (EZH2 Y646N). Uncertainties were estimated using bootstrapping. The data underlying this figure can be found in Raw data 1 at 10.17605/OSF.IO/NTGUX, under CyTOF folder.
(TIF)

S2 Fig. **A.** Left: Sequencing traces of Exon 16 of EZH2, indicating the heterozygous gain-of-function EZH2 Y646N mutation, generated by CRISPR-Cas9 genome editing, in OCI-Ly7 cells. Right: Sequencing traces of Exon 2 of KMT2D, indicating a biallelic insertion of A/C by CRISPR-Cas9 genome editing resulting in a premature stop codon, in OCI-Ly7 cells. **B–C.** Western blot analysis of the indicated modifications in the isogenic OCI-Ly7 WT cells and their counterparts carrying the EZH2 GOF mutation, KMT2D knockout, or a combination of both. Also shown are OCI-Ly1 cells expressing mutant-EZH2 and biallelic loss of KMT2D. Histone H3 represents loading control. **D.** OCI-Ly7 cells (WT), as well as their isogenic counterparts carrying mutant-EZH2 (EZH2 Y646N), biallelic knockout of KMT2D (KMT2D$^{-/-}$), or a combination of both mutations in EZH2 and KMT2D (dMUT), were analyzed by CyTOF. UMAP was performed based on all epigenetic marks measured in two independent biological repeats, following scaling and normalization (see Fig 1E). Shown are the indicated modifications. **E.** OCI-Ly7 and EZH2 Y646N were analyzed by CyTOF, all epigenetic marks were scaled and normalized. UMAP was performed based on all epigenetic marks measured, all marks without H3K27me2/3, or all marks without either H3K27me3 or H3K27me2, as indicated. Colors indicate the sample index. **F.** SHAP analysis between OCI-Ly7 and EZH2 Y646N, indicating the epigenetic marks that contribute to the separation between the two cell lines. H3K27me2, H3K27me3, H3K9me2 and H3K27ac contribute most to the separation of the two cell lines. The data underlying this figure can be found in Raw data 1 at 10.17605/OSF.IO/NTGUX, under CyTOF folder and S1 Raw images.
(TIF)

S3 Fig. A. Single molecule analysis, as described in Furth and colleagues [87], of the percentage of modified nucleosomes for the indicated modification in the indicated sample. Nucleosomes were extracted using MNase digestion and tagged with a fluorophore-biotin conjugate, then captured on a streptavidin-coated surface. Next, nucleosomes were incubated with fluorescently-labeled antibodies targeting the indicated modifications. Values represent colocalization events of antibody signal with nucleosome signal, measured via TIRF microscopy. Full list of all antibodies is depicted in S2 Table. Results validate the CyTOF analysis in Fig 1G. B. Western blot analysis of the indicated modifications in the isogenic OCI-Ly7 EZH2 WT and EZH2 Y646N cells. Histone H3 represents loading control. Results validate the CyTOF analysis in Fig 1G. C–D. Immunofluorescence of OCI-Ly7 and EZH2 Y646N cells for the indicated histone modifications (H3K9me2/3 AF-555 and H3K27me3 AF-488) and DAPI. EZH2 inhibitor was applied to OCI-Ly7 cells at a concentration of 10 μM for 48 h. EZH2 Y646N cells show robust reduction in H3K9me2/3 levels, compatible with the CyTOF analysis. Of note,

H3K9me2/3 levels are detected here by a different antibody (clone and vendor) than the one used for CyTOF, providing an additional validation of the data. D. Mean values of the fluorescent signal (40–60 cells per sample) of the indicated histone modifications, as seen in S3C Fig. P values were calculated by T test. E. Two independent repeats of FACS analysis of OCI-Ly7 cells, stained with fluorescent H3K27me3 and H3K27ac antibodies. Scatterplots represent antibody signal against Zombie Violet dye of samples stained with one antibody only, or both antibodies together. The results indicate that there is no apparent interference between the antibodies, as the signals of staining the cells with a single antibody or a combination of both are comparable. F. Western blot analysis of H3K9me2 in OCI-Ly7 or OCI-Ly18 cells. Cells were incubated with DMSO, or the G9a inhibitor UNC-0642 at concentrations of 10 µM or 20 µM for 48 h. As expected, cells treated with the inhibitor showed reduced H3K9me2 levels, validating the antibody's detection of this mark. Histone H3 represents loading control. G. Box plot of H3K9me2, H3K27me2 and H3K27me3, as measured by CyTOF, in OCI-Ly7 and OCI-Ly18 cells. Cells were incubated with DMSO, or the G9a inhibitor UNC-0642 at a concentration of 10 µM, for 24 or 48 h. In OCI-Ly7 cells, the reduction of H3K9me2 was accompanied by a reduction in H3K27me2, validating the link between these marks that is also observed in Fig 4. In OCI-Ly18 cells, however, while H3K9me2 levels were reduced by the inhibitor treatment, we observed an increase in H3K27me2. This data strongly suggests that there is no significant cross-reactivity of the H3K27me2/3 antibodies with H3K9me2. P values were calculated by Welch's t test. ***p value < 0.001. The data underlying this figure can be found in Raw data 1, 4–5 at 10.17605/OSF.IO/NTGUX, under CyTOF, Single molecule and IF folders and S1 Raw images.
(TIF)

S4 Fig. A. RNA-sequencing analysis of OCI-Ly7 cells and the indicated isogenic mutant lines. Principal component analysis of count number for all genes. Showing PC1 (40.3%) and PC2 (12.05%). B–D. Volcano plots showing Log fold change and p.adj values of differentially expressed genes between the indicated mutant lines versus OCI-Ly7. The list of differentially expressed genes was obtained using DESeq2. Upregulated genes marked in blue were filtered by >1 LogFC and p.adj ≤ 0.05. Downregulated genes marked in red were filtered by <−1 LogFC and p.adj ≤ 0.05. E. Visualization of OCI-Ly7 H3K27me3 read coverage around the TSS, as measured by Cut&Run, of differential genes from clusters 2 and 4 (Fig 1H). F. Bar plot showing Log2 change values of the indicated genes in the mutant cell lines versus OCI-Ly7. Colors indicate sample index. The data underlying this figure can be found in Raw data 1, 3 at 10.17605/OSF.IO/NTGUX, under CyTOF and Cut and Run folders.
(TIF)

S5 Fig. A. The Gini coefficient, commonly used to measure the inequality among the values of a frequency distribution, was used to determine the relative heterogeneity of each modification between the different cell lines. The ratio of the Gini coefficient between the individual lines and the WT was used as a measure of the excess heterogeneity of the mutant over the WT for each modification. B. The indicated heterogeneity measurements that were defined in Fig 2C were calculated for OCI-Ly7 cells versus each of the mutant lines, across a wide range of UMAP parameters. This experiment is a biological repeat of the experiment presented in Fig 2C. Plotted are the ratios of the values of each mutant line versus OCI-Ly7. Thus, values over 1 indicate excess heterogeneity in the mutant over that of the WT. EZH2 mutant cells showed higher heterogeneity in all measurements. C. Heterogeneity measurements calculated for OCI-Ly7 versus the GOF cells or the double mutant with KMT2D. Values were calculated based on the UMAPs presented in S2E Fig, without H3K27me2/3. EZH2-mutant cells show higher heterogeneity locally and globally, even when removing H3K27me2/3 from the analysis. D. Heterogeneity measurements calculated for the two patient derived xenografts with WT EZH2 (PDXWT) versus mutant-EZH2 (PDXMUT). Overall, EZH2-mutant cells show higher heterogeneity locally, and for global area. The data underlying this figure can be found in Raw data 1 at 10.17605/OSF.IO/NTGUX, under CyTOF folder and S1 Raw images.
(TIF)

S6 Fig. A–B. Top: OCI-Ly7 cells expressing WT EZH2, mutant-EZH2 (EZH2 Y646N), or a combination of mutant-EZH2 with KMT2D biallelic knockout (dMUT) were analyzed by CyTOF. Shown are joint UMAPs of OCI-Ly7 cells (WT) and the indicated lines. A represents a repeat of the experiment shown in Fig 2F. B shows two repeats for the double mutant (KMT2D−/− EZH2 Y646N). Colors indicate the sample index. Bottom: Only mutant-EZH2 cells are plotted, to highlight the subpopulation of mutant cells that cluster with WT cells, referred to as 'WT-like'. C. The fraction of WT-like cells in three biological CyTOF replicates for the single EZH2 Y6464 mutant line, and two biological replicates for the double mutant. The percentage of cells consisting of this subpopulation is dynamic and varies between experiments. D. The mean of distribution of the indicated histone modifications for OCI-Ly7 cells expressing WT EZH2 ('WT'), and for the double mutant of EZH2 Y646N with KMT2D knockout (dMUT) cells that either express robustly the GOF phenotype and form a distinct cluster (EZH2 Y646N dMUT 'Extreme'), or the 'WT-like' subpopulation that clusters with OCI-Ly7 WT cells. See also Fig 2G for the single EZH2 mutant line. E. OCI-Ly7 cells and the indicated isogenic mutants were analyzed by CyTOF, that included all epigenetic modifications as well as the Maxpar cell cycle panel kit. The cell cycle markers, included in the CyTOF panel, were used to determine the cell cycle phase of each cell: G1, S and G2. For each phase, a joint UMAP of the WT and mutant cells was generated, based on all epigenetic modifications. 'WT-like' cells are observed for each of the indicated cell cycle phases. F. Histogram showing image stream values of H3K27me3 in OCI-Ly7 and EZH2 Y646N. In this complementary technology to CyTOF, we also observed a bimodal distribution for H3K27me3, indicating a sub-population of WT-like cells (8% of the EZH2 Y646N cells). G–H. Histogram of H3K27me3 levels in: G. Patient-derived xenograft expressing mutant-EZH2 (PDXMUT). H. OCI-Ly1 cells expressing mutant-EZH2. Cells expressing the mutant enzyme show bimodal distribution of H3K27me3, indicating two distinct subpopulations. I. Illustration of the effects of EZH2 inhibitor (EZH2i) Tazemetostat that blocks EZH2-mediated deposition of H3K27me2/3, and GSK-J4 (KDM6i) that inhibits the H3K27 demethylase KDM6A/B, resulting in elevated H3K27me3 levels. J. OCI-LY7 with mutant-EZH2 were treated with the histone lysine 27 demethylase KDM6A/B inhibitor GSK-J4 at a concentration of 1 µM for 24 h, or left untreated as control. Shown are the expression levels of H3K27me3, as measured by CyTOF. Treatment with the inhibitor led to the expected upregulation of H3K27me3 levels. P values were calculated by Welch's t test. ***p value < 0.001. K. OCI-Ly7 cells were treated with KDM6A/B inhibitor at 1 µM concentration for 24 h, followed by CyTOF. The percentage of 'WT-like' cells in control and EZH2i-treated cells is shown. EZH2 inhibition reduced the fraction of 'WT-like' cells. L. EZH2 Y646N cells were treated with EZH2 inhibitor at a concentration of 10 µM for 48 h, followed by CyTOF analysis. While the 'EZH2-extreme' population was highly affected by the drug and was reduced by 15-fold, the 'WT-like' cells were less affected (9-fold reduction). M. OCI-Ly1 cells expressing mutant-EZH2 were treated with EZH2 inhibitor at a concentration of 10 µM for one week to completely deplete H3K27me3. Shown is a western blot for H3K27me3 in the control and two repeats of the treated cells. Histone H3 is used as a loading control. The data underlying this figure can be found in Raw data 1 at 10.17605/OSF.IO/NTGUX, under CyTOF folder.
(TIF)

S7 Fig. **EZH2 Y646N were sorted into single cells and allowed to proliferate to generate clones, similar to** Fig 3G. Histogram of scaled and normalized H3K27me3 levels, as measured by CyTOF, in OCI-Ly7 cells (WT) or the indicated EZH2 Y646N single-cell clones. The percentage of 'WT-like' cells is shown per clone. The data underlying this figure can be found in Raw data 1 at 10.17605/OSF.IO/NTGUX, under CyTOF folder.
(TIF)

S8 Fig. A–C. Models to decipher interactions within the epigenetic network, applied to unperturbed OCI-Ly7 single-cell CyTOF data, on a biological repeat of the experiment shown in Fig 3A–C. A. XGBoost analysis. B. 'Probability of link' analysis. For both A and B, the Y axis indicates 'source' modification and X axis indicates its 'target' modification affected by the Y axis. C. Partial correlations between histone modifications. D–E. Markov chain simulation of the H3K27 methylation and acetylation dynamics (See "Methods"). D. The table shows the steady-state populations in our simulation,

representing the fraction of nucleosomes modified within a cell. E. Shown are the distribution of the SHAP values, associated with the effect of H3K27me2 on H3K27me3 (blue) or H3K27me3 on H3K27me2 (red). The distribution shown is for 500 runs over a subsample of the data in order to allow for the statistical variation. As our model predicts, there is a stronger effect of H3K27me2 on H3K27me3 compared to the reverse direction. F. Dendrogram of the hierarchical clustering of spearman correlations of Cut&Run normalized reads between all samples, calculated on 10kbp genomic bins. G–H. Probability of link analysis done on CyTOF data of unperturbed germinal-center B cells (CD20+, CD38+), derived from tonsils, for two independent samples derived from different patients. Tonsils were dissociated to single cells followed by staining with the panel of metal-conjugated antibodies and CyTOF analysis. The data underlying this figure can be found in Raw data 1, 3 at 10.17605/OSF.IO/NTGUX, under CyTOF and Cut and Run folders and S1 Raw images.
(TIF)

S9 Fig. A. Scaled, normalized levels of BCL6 on the joint UMAP of OCI-Ly7 and EZH2 Y646N cells, corresponding to the UMAP shown in Fig 2A. Cells expressing mutant-EZH2 show downregulation of BCL6 levels. B. Western blot analysis of H3K27me3 and BCL6 in the indicated samples. Histone H3 is used as a loading control. C. Quantitative RT-PCR analysis of BCL6 expression in the isogenic OCI-Ly7 WT and EZH2-mutant cells. ΔΔCT values relative to OCI-Ly7 ±s.d (n = 3) are shown. HPRT was used for normalization. D. BCL6 mean levels in cells binned according to H3K27me3 levels. Red hue represents standard error of the mean. Number of cells in each bin is shown. Top: PDX with WT-EZH2 (PDXWT). Bottom: PDX with mutant-EZH2 (PDXMUT). H3K27me3 and BCL6 show non-linear relationship. E–F. Western blot analysis of H3K27me3 and BCL6. Histone H3 is used as a loading control. E. OCI-Ly7 cells were treated with EZH2 inhibitor at a concentration of 10 μM for the indicated times. F. EZH2-mutant cells were treated with EZH2 inhibitor at a concentration of 10 μM for 48 h. G–H. Quantitative RT-PCR analysis of BCL6 expression in the isogenic OCI-Ly7 WT and EZH2-mutant cells, treated with EZH2i 10 μM at the indicated times. ΔΔCT values relative to OCI-Ly7 ±s.d (n = 3) are shown. HPRT was used for normalization. I. Western blot analysis of H3K27me3 and BCL6. OCI-Ly1 cells, carrying mutant-EZH2, were treated with EZH2 inhibitor at a concentration of 10 μM for 48 h. J. Quantitative RT-PCR analysis of BCL6 expression in OCI-Ly1 cells treated with EZH2 inhibitor at a concentration of 10 μM for 48 h versus DMSO control. ΔΔCT values relative to DMSO treated sample ±s.d (n = 3) are shown. The data underlying this figure can be found in Raw data 1, 6 at 10.17605/OSF.IO/NTGUX, under CyTOF and RT-PCR folders and S1 Raw images.
(TIF)

S1 Table. CyTOF antibodies used in this study.
(XLSX)

S2 Table. All antibodies used in this study and their usages.
(XLSX)

S1 Raw images. Raw uncropped images of Western Blotting in this study for S2B, S2C; S3B, S3F; S6M, S9B, S9E, S9G and S9I Figs. Doi: 10.6084/m9.figshare.28737236.
(PDF)

Author contributions

Conceptualization: Ofir Griess, Noa Furth, Nofar Harpaz, Guy Ron, Efrat Shema.

Data curation: Ofir Griess, Nicoletta Di Bernardo, Tomer-Meir Salame.

Formal analysis: Ofir Griess, Nicoletta Di Bernardo, Bareket Dassa, Guy Ron.

Funding acquisition: Efrat Shema.

Investigation: Ofir Griess, Guy Ron, Efrat Shema.

Methodology: Ofir Griess, Guy Ron, Efrat Shema.

Project administration: Efrat Shema.

Resources: Ofir Griess, Ioannis Karagiannidis, Yusuke Isshiki, Menachem Gross, Ari M. Melnick, Wendy Béguelin, Efrat Shema.

Software: Efrat Shema.

Supervision: Efrat Shema.

Validation: Ofir Griess.

Visualization: Ofir Griess, Noa Furth, Efrat Shema.

Writing – original draft: Ofir Griess, Efrat Shema.

Writing – review & editing: Ofir Griess, Noa Furth, Guy Ron, Efrat Shema.

References

1. Alizadeh AA, Eisen MB, Davis RE, Ma C, Lossos IS, Rosenwald A, et al. Distinct types of diffuse large B-cell lymphoma identified by gene expression profiling. Nature. 2000;403(6769):503–11. https://doi.org/10.1038/35000501 PMID: 10676951.

2. Swerdlow SH, Campo E, Pileri SA, Harris NL, Stein H, Siebert R, et al. The 2016 revision of the World Health Organization classification of lymphoid neoplasms. Blood. 2016;127(20):2375–90. https://doi.org/10.1182/blood-2016-01-643569 PMID: 26980727.

3. Crotty S, Johnston RJ, Schoenberger SP. Effectors and memories: Bcl-6 and Blimp-1 in T and B lymphocyte differentiation. Nat Immunol. 2010;11(2):114–20. https://doi.org/10.1038/ni.1837 PMID: 20084069.

4. Song S, Matthias PD. The transcriptional regulation of germinal center formation. Front Immunol. 2018;9:2026. https://doi.org/10.3389/fimmu.2018.02026 PMID: 30233601.

5. Gatto D, Brink R. The germinal center reaction. J Allergy Clin Immunol. 2010;126(5):898–907; quiz 908–9. https://doi.org/10.1016/j.jaci.2010.09.007 PMID: 21050940.

6. Pasqualucci L, Dalla-Favera R. Genetics of diffuse large B-cell lymphoma. Blood. 2018;131(21):2307–19. https://doi.org/10.1182/blood-2017-11-764332 PMID: 29666115.

7. Pasqualucci L, Neumeister P, Goossens T, Nanjangud G, Chaganti RS, Küppers R, et al. Hypermutation of multiple proto-oncogenes in B-cell diffuse large-cell lymphomas. Nature. 2001;412(6844):341–6. https://doi.org/10.1038/35085588 PMID: 11460166.

8. Ye BH, Lista F, Lo Coco F, Knowles DM, Offit K, Chaganti RS, et al. Alterations of a zinc finger-encoding gene, BCL-6, in diffuse large-cell lymphoma. Science. 1993;262(5134):747–50. https://doi.org/10.1126/science.8235596 PMID: 8235596.

9. Yang H, Green MR. Epigenetic programing of B-Cell lymphoma by BCL6 and its genetic deregulation. Front Cell Dev Biol. 2019;7:272. https://doi.org/10.3389/fcell.2019.00272 PMID: 31788471.

10. Stavnezer J, Guikema JEJ, Schrader CE. Mechanism and regulation of class switch recombination. Annu Rev Immunol. 2008;26:261–92. https://doi.org/10.1146/annurev.immunol.26.021607.090248 PMID: 18370922.

11. Klein U, Dalla-Favera R. Germinal centres: role in B-cell physiology and malignancy. Nat Rev Immunol. 2008;8(1):22–33. https://doi.org/10.1038/nri2217 PMID: 18097447.

12. Bunting KL, Melnick AM. New effector functions and regulatory mechanisms of BCL6 in normal and malignant lymphocytes. Curr Opin Immunol. 2013;25(3):339–46. https://doi.org/10.1016/j.coi.2013.05.003 PMID: 23725655.

13. Hanahan D. Hallmarks of cancer: new dimensions. Cancer Discov. 2022;12(1):31–46. https://doi.org/10.1158/2159-8290.CD-21-1059 PMID: 35022204.

14. Jiang Y, Melnick A. The epigenetic basis of diffuse large B-cell lymphoma. Semin Hematol. 2015;52(2):86–96. https://doi.org/10.1053/j.seminhematol.2015.01.003 PMID: 25805588.

15. Caganova M, Carrisi C, Varano G, Mainoldi F, Zanardi F, Germain P-L, et al. Germinal center dysregulation by histone methyltransferase EZH2 promotes lymphomagenesis. J Clin Invest. 2013;123(12):5009–22. https://doi.org/10.1172/JCI70626 PMID: 24200695.

16. Millán-Zambrano G, Burton A, Bannister AJ, Schneider R. Histone post-translational modifications - cause and consequence of genome function. Nat Rev Genet. 2022;23(9):563–80. https://doi.org/10.1038/s41576-022-00468-7 PMID: 35338361.

17. Furth N, Shema E. It's all in the combination: decoding the epigenome for cancer research and diagnostics. Curr Opin Genet Dev. 2022;73:101899. https://doi.org/10.1016/j.gde.2022.101899 PMID: 35091256.

18. Shema E, Bernstein BE, Buenrostro JD. Single-cell and single-molecule epigenomics to uncover genome regulation at unprecedented resolution. Nat Genet. 2019;51(1):19–25. https://doi.org/10.1038/s41588-018-0290-x PMID: 30559489.

19. Atlasi Y, Stunnenberg HG. The interplay of epigenetic marks during stem cell differentiation and development. Nat Rev Genet. 2017;18(11):643–58. https://doi.org/10.1038/nrg.2017.57 PMID: 28804139.

20. Schmitz R, Wright GW, Huang DW, Johnson CA, Phelan JD, Wang JQ, et al. Genetics and pathogenesis of diffuse large B-cell lymphoma. N Engl J Med. 2018;378(15):1396–407. https://doi.org/10.1056/NEJMoa1801445 PMID: 29641966.

21. Wright GW, Huang DW, Phelan JD, Coulibaly ZA, Roulland S, Young RM, et al. A probabilistic classification tool for genetic subtypes of diffuse large B cell lymphoma with therapeutic implications. Cancer Cell. 2020;37(4):551-568.e14. https://doi.org/10.1016/j.ccell.2020.03.015 PMID: 32289277.

22. Pasqualucci L, Trifonov V, Fabbri G, Ma J, Rossi D, Chiarenza A, et al. Analysis of the coding genome of diffuse large B-cell lymphoma. Nat Genet. 2011;43(9):830–7. https://doi.org/10.1038/ng.892 PMID: 21804550.

23. Zhang J, Dominguez-Sola D, Hussein S, Lee J-E, Holmes AB, Bansal M, et al. Disruption of KMT2D perturbs germinal center B cell development and promotes lymphomagenesis. Nat Med. 2015;21(10):1190–8. https://doi.org/10.1038/nm.3940 PMID: 26366712.

24. Ortega-Molina A, Boss IW, Canela A, Pan H, Jiang Y, Zhao C, et al. The histone lysine methyltransferase KMT2D sustains a gene expression program that represses B cell lymphoma development. Nat Med. 2015;21(10):1199–208. https://doi.org/10.1038/nm.3943 PMID: 26366710.

25. Heyn H, Esteller M. EZH2: an epigenetic gatekeeper promoting lymphomagenesis. Cancer Cell. 2013;23(5):563–5. https://doi.org/10.1016/j.ccr.2013.04.028 PMID: 23680141.

26. Comet I, Riising EM, Leblanc B, Helin K. Maintaining cell identity: PRC2-mediated regulation of transcription and cancer. Nat Rev Cancer. 2016;16(12):803–10. https://doi.org/10.1038/nrc.2016.83 PMID: 27658528.

27. Béguelin W, Popovic R, Teater M, Jiang Y, Bunting KL, Rosen M, et al. EZH2 is required for germinal center formation and somatic EZH2 mutations promote lymphoid transformation. Cancer Cell. 2013;23(5):677–92. https://doi.org/10.1016/j.ccr.2013.04.011 PMID: 23680150.

28. Béguelin W, Teater M, Gearhart MD, Calvo Fernández MT, Goldstein RL, Cárdenas MG, et al. EZH2 and BCL6 cooperate to assemble CBX8-BCOR complex to repress bivalent promoters, mediate germinal center formation and lymphomagenesis. Cancer Cell. 2016;30(2):197–213. https://doi.org/10.1016/j.ccell.2016.07.006 PMID: 27505670.

29. Morin RD, Johnson NA, Severson TM, Mungall AJ, An J, Goya R, et al. Somatic mutations altering EZH2 (Tyr641) in follicular and diffuse large B-cell lymphomas of germinal-center origin. Nat Genet. 2010;42(2):181–5. https://doi.org/10.1038/ng.518 PMID: 20081860.

30. McCabe MT, Graves AP, Ganji G, Diaz E, Halsey WS, Jiang Y, et al. Mutation of A677 in histone methyltransferase EZH2 in human B-cell lymphoma promotes hypertrimethylation of histone H3 on lysine 27 (H3K27). Proc Natl Acad Sci U S A. 2012;109(8):2989–94. https://doi.org/10.1073/pnas.1116418109 PMID: 22323599.

31. McCabe MT, Ott HM, Ganji G, Korenchuk S, Thompson C, Van Aller GS, et al. EZH2 inhibition as a therapeutic strategy for lymphoma with EZH2-activating mutations. Nature. 2012;492(7427):108–12. https://doi.org/10.1038/nature11606 PMID: 23051747.

32. Velichutina I, Shaknovich R, Geng H, Johnson NA, Gascoyne RD, Melnick AM, et al. EZH2-mediated epigenetic silencing in germinal center B cells contributes to proliferation and lymphomagenesis. Blood. 2010;116(24):5247–55. https://doi.org/10.1182/blood-2010-04-280149 PMID: 20736451.

33. Donaldson-Collier MC, Sungalee S, Zufferey M, Tavernari D, Katanayeva N, Battistello E, et al. EZH2 oncogenic mutations drive epigenetic, transcriptional, and structural changes within chromatin domains. Nat Genet. 2019;51(3):517–28. https://doi.org/10.1038/s41588-018-0338-y PMID: 30692681.

34. Béguelin W, Teater M, Meydan C, Hoehn KB, Phillip JM, Soshnev AA, et al. Mutant EZH2 induces a pre-malignant lymphoma niche by reprogramming the immune response. Cancer Cell. 2020;37(5):655-673.e11. https://doi.org/10.1016/j.ccell.2020.04.004 PMID: 32396861.

35. Shema E, Jones D, Shoresh N, Donohue L, Ram O, Bernstein BE. Single-molecule decoding of combinatorially modified nucleosomes. Science. 2016;352(6286):717–21. https://doi.org/10.1126/science.aad7701 PMID: 27151869.

36. Vlasevska S, Garcia-Ibanez L, Duval R, Holmes AB, Jahan R, Cai B, et al. KMT2D acetylation by CREBBP reveals a cooperative functional interaction at enhancers in normal and malignant germinal center B cells. Proc Natl Acad Sci U S A. 2023;120(11):e2218330120. https://doi.org/10.1073/pnas.2218330120 PMID: 36893259.

37. Chapuy B, Stewart C, Dunford AJ, Kim J, Kamburov A, Redd RA, et al. Molecular subtypes of diffuse large B cell lymphoma are associated with distinct pathogenic mechanisms and outcomes. Nat Med. 2018;24(5):679–90. https://doi.org/10.1038/s41591-018-0016-8 PMID: 29713087.

38. Harpaz N, Mittelman T, Beresh O, Griess O, Furth N, Salame T-M, et al. Single-cell epigenetic analysis reveals principles of chromatin states in H3.3-K27M gliomas. Mol Cell. 2022;82(14):2696-2713.e9. https://doi.org/10.1016/j.molcel.2022.05.023 PMID: 35716669.

39. Lasserre J, Chung H-R, Vingron M. Finding associations among histone modifications using sparse partial correlation networks. PLoS Comput Biol. 2013;9(9):e1003168. https://doi.org/10.1371/journal.pcbi.1003168 PMID: 24039558.

40. Perner J, Lasserre J, Kinkley S, Vingron M, Chung H-R. Inference of interactions between chromatin modifiers and histone modifications: from ChIP-Seq data to chromatin-signaling. Nucleic Acids Res. 2014;42(22):13689–95. https://doi.org/10.1093/nar/gku1234 PMID: 25414326.

41. Li J, Chin CR, Ying H-Y, Meydan C, Teater MR, Xia M, et al. Loss of CREBBP and KMT2D cooperate to accelerate lymphomagenesis and shape the lymphoma immune microenvironment. Nat Commun. 2024;15(1):2879. https://doi.org/10.1038/s41467-024-47012-1 PMID: 38570506.

42. Yusufova N, Kloetgen A, Teater M, Osunsade A, Camarillo JM, Chin CR, et al. Histone H1 loss drives lymphoma by disrupting 3D chromatin architecture. Nature. 2021;589(7841):299–305. https://doi.org/10.1038/s41586-020-3017-y PMID: 33299181.

43. Willcockson MA, Healton SE, Weiss CN, Bartholdy BA, Botbol Y, Mishra LN, et al. H1 histones control the epigenetic landscape by local chromatin compaction. Nature. 2021;589(7841):293–8. https://doi.org/10.1038/s41586-020-3032-z PMID: 33299182.

44. Béguelin W, Rivas MA, Calvo Fernández MT, Teater M, Purwada A, Redmond D, et al. EZH2 enables germinal centre formation through epigenetic silencing of CDKN1A and an Rb-E2F1 feedback loop. Nat Commun. 2017;8(1):877. https://doi.org/10.1038/s41467-017-01029-x PMID: 29026085.

45. Young RM, Shaffer AL 3rd, Phelan JD, Staudt LM. B-cell receptor signaling in diffuse large B-cell lymphoma. Semin Hematol. 2015;52(2):77–85. https://doi.org/10.1053/j.seminhematol.2015.01.008 PMID: 25805587.

46. Zhao Y, Cui W-L, Feng Z-Y, Xue J, Gulinaer A, Zhang W. Expression of Foxp3 and interleukin-7 receptor and clinicopathological characteristics of patients with diffuse large B-cell lymphoma. Oncol Lett. 2020;19(4):2755–64. https://doi.org/10.3892/ol.2020.11374 PMID: 32218828.

47. Gini C. On the Measure of Concentration with Espacial Reference to Income and Wealth. Cowles Commission. 1936;2(3).

48. Han G, Deng Q, Marques-Piubelli ML, Dai E, Dang M, Ma MCJ, et al. Follicular lymphoma microenvironment characteristics associated with tumor cell mutations and MHC class II expression. Blood Cancer Discov. 2022;3(5):428–43. https://doi.org/10.1158/2643-3230.BCD-21-0075 PMID: 35687817.

49. Study of tazemetostat in participants with relapsed or refractory B-cell non-Hodgkin's lymphoma with EZH2 gene mutation. Available from: https://clinicaltrials.gov/ct2/show/NCT03456726

50. Morschhauser F, Tilly H, Chaidos A, McKay P, Phillips T, Assouline S, et al. Tazemetostat for patients with relapsed or refractory follicular lymphoma: an open-label, single-arm, multicentre, phase 2 trial. Lancet Oncol. 2020;21(11):1433–42. https://doi.org/10.1016/S1470-2045(20)30441-1 PMID: 33035457.

51. Scott M, Lundberg S-IL. A Unified Approach to Interpreting Model Predictions. 2017.

52. Valerio DG, Xu H, Chen C-W, Hoshii T, Eisold ME, Delaney C, et al. Histone acetyltransferase activity of MOF is required for MLL-AF9 leukemogenesis. Cancer Res. 2017;77(7):1753–62. https://doi.org/10.1158/0008-5472.CAN-16-2374 PMID: 28202522.

53. Mozzetta C, Pontis J, Fritsch L, Robin P, Portoso M, Proux C, et al. The histone H3 lysine 9 methyltransferases G9a and GLP regulate polycomb repressive complex 2-mediated gene silencing. Mol Cell. 2014;53(2):277–89. https://doi.org/10.1016/j.molcel.2013.12.005 PMID: 24389103.

54. Mozzetta C, Pontis J, Ait-Si-Ali S. Functional crosstalk between lysine methyltransferases on histone substrates: the case of G9A/GLP and polycomb repressive complex 2. Antioxid Redox Signal. 2015;22(16):1365–81. https://doi.org/10.1089/ars.2014.6116 PMID: 25365549.

55. Casciello F, Kelly GM, Ramarao-Milne P, Kamal N, Stewart TA, Mukhopadhyay P, et al. Combined inhibition of G9a and EZH2 suppresses tumor growth via synergistic induction of IL24-mediated apoptosis. Cancer Res. 2022;82(7):1208–21. https://doi.org/10.1158/0008-5472.CAN-21-2218 PMID: 35149587.

56. Lavarone E, Barbieri CM, Pasini D. Dissecting the role of H3K27 acetylation and methylation in PRC2 mediated control of cellular identity. Nat Commun. 2019;10(1):1679. https://doi.org/10.1038/s41467-019-09624-w PMID: 30976011.

57. Pasini D, Malatesta M, Jung HR, Walfridsson J, Willer A, Olsson L, et al. Characterization of an antagonistic switch between histone H3 lysine 27 methylation and acetylation in the transcriptional regulation of Polycomb group target genes. Nucleic Acids Res. 2010;38(15):4958–69. https://doi.org/10.1093/nar/gkq244 PMID: 20385584.

58. Tie F, Banerjee R, Stratton CA, Prasad-Sinha J, Stepanik V, Zlobin A, et al. CBP-mediated acetylation of histone H3 lysine 27 antagonizes *Drosophila* Polycomb silencing. Development. 2009;136(18):3131–41. https://doi.org/10.1242/dev.037127 PMID: 19700617.

59. Radzisheuskaya A, Shliaha PV, Grinev VV, Shlyueva D, Damhofer H, Koche R, et al. Complex-dependent histone acetyltransferase activity of KAT8 determines its role in transcription and cellular homeostasis. Mol Cell. 2021;81(8):1749-1765.e8. https://doi.org/10.1016/j.molcel.2021.02.012 PMID: 33657400.

60. Laugesen A, Højfeldt JW, Helin K. Molecular mechanisms directing PRC2 recruitment and H3K27 methylation. Mol Cell. 2019;74(1):8–18. https://doi.org/10.1016/j.molcel.2019.03.011 PMID: 30951652.

61. Kim JJ, Kingston RE. Context-specific Polycomb mechanisms in development. Nat Rev Genet. 2022;23(11):680–95. https://doi.org/10.1038/s41576-022-00499-0 PMID: 35681061.

62. Hatzi K, Jiang Y, Huang C, Garrett-Bakelman F, Gearhart MD, Giannopoulou EG, et al. A hybrid mechanism of action for BCL6 in B cells defined by formation of functionally distinct complexes at enhancers and promoters. Cell Rep. 2013;4(3):578–88. https://doi.org/10.1016/j.celrep.2013.06.016 PMID: 23911289.

63. Sneeringer CJ, Scott MP, Kuntz KW, Knutson SK, Pollock RM, Richon VM, et al. Coordinated activities of wild-type plus mutant EZH2 drive tumor-associated hypertrimethylation of lysine 27 on histone H3 (H3K27) in human B-cell lymphomas. Proc Natl Acad Sci U S A. 2010;107(49):20980–5. https://doi.org/10.1073/pnas.1012525107 PMID: 21078963.

64. Mohammad F, Weissmann S, Leblanc B, Pandey DP, Højfeldt JW, Comet I, et al. EZH2 is a potential therapeutic target for H3K27M-mutant pediatric gliomas. Nat Med. 2017;23(4):483–92. https://doi.org/10.1038/nm.4293 PMID: 28263309.

65. Piunti A, Hashizume R, Morgan MA, Bartom ET, Horbinski CM, Marshall SA, et al. Therapeutic targeting of polycomb and BET bromodomain proteins in diffuse intrinsic pontine gliomas. Nat Med. 2017;23(4):493–500. https://doi.org/10.1038/nm.4296 PMID: 28263307.

66. Flavahan WA, Gaskell E, Bernstein BE. Epigenetic plasticity and the hallmarks of cancer. Science. 2017;357(6348):eaal2380. https://doi.org/10.1126/science.aal2380 PMID: 28729483.

67. Feinberg AP, Levchenko A. Epigenetics as a mediator of plasticity in cancer. Science. 2023;379(6632):eaaw3835. https://doi.org/10.1126/science.aaw3835 PMID: 36758093.

68. De S, Shaknovich R, Riester M, Elemento O, Geng H, Kormaksson M, et al. Aberration in DNA methylation in B-cell lymphomas has a complex origin and increases with disease severity. PLoS Genet. 2013;9(1):e1003137. https://doi.org/10.1371/journal.pgen.1003137 PMID: 23326238.

69. Aylon Y, Furth N, Mallel G, Friedlander G, Nataraj NB, Dong M, et al. Breast cancer plasticity is restricted by a LATS1-NCOR1 repressive axis. Nat Commun. 2022;13(1):7199. https://doi.org/10.1038/s41467-022-34863-9 PMID: 36443319.

70. Pitter KL, Grbovic-Huezo O, Joost S, Singhal A, Blum M, Wu K, et al. Systematic comparison of pancreatic ductal adenocarcinoma models identifies a conserved highly plastic basal cell state. Cancer Res. 2022;82(19):3549–60. https://doi.org/10.1158/0008-5472.CAN-22-1742 PMID: 35952360.

71. Laisné M, Lupien M, Vallot C. Epigenomic heterogeneity as a source of tumour evolution. Nat Rev Cancer. 2025;25(1):7–26. https://doi.org/10.1038/s41568-024-00757-9 PMID: 39414948.

72. Chapuy B, Cheng H, Watahiki A, Ducar MD, Tan Y, Chen L, et al. Diffuse large B-cell lymphoma patient-derived xenograft models capture the molecular and biological heterogeneity of the disease. Blood. 2016;127(18):2203–13. https://doi.org/10.1182/blood-2015-09-672352 PMID: 26773040.

73. Kotlov N, Bagaev A, Revuelta MV, Phillip JM, Cacciapuoti MT, Antysheva Z, et al. Clinical and biological subtypes of b-cell lymphoma revealed by microenvironmental signatures. Cancer Discov. 2021;11(6):1468–89. https://doi.org/10.1158/2159-8290.CD-20-0839 PMID: 33541860.

74. Massoni-Badosa R, Aguilar-Fernández S, Nieto JC, Soler-Vila P, Elosua-Bayes M, Marchese D, et al. An atlas of cells in the human tonsil. Immunity. 2024;57(2):379-399.e18. https://doi.org/10.1016/j.immuni.2024.01.006 PMID: 38301653.

75. Jaitin DA, Kenigsberg E, Keren-Shaul H, Elefant N, Paul F, Zaretsky I, et al. Massively parallel single-cell RNA-seq for marker-free decomposition of tissues into cell types. Science. 2014;343(6172):776–9. https://doi.org/10.1126/science.1247651 PMID: 24531970.

76. Kohen R, Barlev J, Hornung G, Stelzer G, Feldmesser E, Kogan K, et al. UTAP: user-friendly transcriptome analysis pipeline. BMC Bioinformatics. 2019;20(1):154. https://doi.org/10.1186/s12859-019-2728-2 PMID: 30909881.

77. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. 2014;15(12):550. https://doi.org/10.1186/s13059-014-0550-8 PMID: 25516281.

78. Meers MP, Bryson TD, Henikoff JG, Henikoff S. Improved CUT&RUN chromatin profiling tools. Elife. 2019;8:e46314. https://doi.org/10.7554/eLife.46314 PMID: 31232687.

79. Martin M. Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet J. 2011;17(1):10–2.

80. Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. Nat Methods. 2012;9(4):357–9. https://doi.org/10.1038/nmeth.1923 PMID: 22388286.

81. Meers MP, Tenenbaum D, Henikoff S. Peak calling by sparse enrichment analysis for CUT&RUN chromatin profiling. Epigenetics Chromatin. 2019;12(1):42. https://doi.org/10.1186/s13072-019-0287-4 PMID: 31300027.

82. Ramírez F, Ryan DP, Grüning B, Bhardwaj V, Kilpert F, Richter AS, et al. deepTools2: a next generation web server for deep-sequencing data analysis. Nucleic Acids Res. 2016;44(W1):W160-5. https://doi.org/10.1093/nar/gkw257 PMID: 27079975.

83. Fedyuk V, Erez N, Furth N, Beresh O, Andreishcheva E, Shinde A, et al. Multiplexed, single-molecule, epigenetic analysis of plasma-isolated nucleosomes for cancer diagnostics. Nat Biotechnol. 2023;41(2):212–21. https://doi.org/10.1038/s41587-022-01447-3 PMID: 36076083.

84. Bagwell CB, Inokuma M, Hunsberger B, Herbert D, Bray C, Hill B, et al. Automated data cleanup for mass cytometry. Cytometry A. 2020;97(2):184–98. https://doi.org/10.1002/cyto.a.23926 PMID: 31737997.

85. Chen T, Guestrin C. Xgboost: A scalable tree boosting system. Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. 2016.

86. Erdogan E. Bnlearn: Bayesian Network Structure Learning in Python. GitHub repository. Github; 2020. https://github.com/erdogant/bnlearn

87. Furth N, Algranati D, Dassa B, Beresh O, Fedyuk V, Morris N, et al. H3-K27M-mutant nucleosomes interact with MLL1 to shape the glioma epigenetic landscape. Cell Rep. 2022;39(7):110836. https://doi.org/10.1016/j.celrep.2022.110836 PMID: 35584667.