Article

Histone H1 loss drives lymphoma by disrupting 3D chromatin architecture

Nevin Yusufova[1,2], Andreas Kloetgen[3,4], Matt Teater[1], Adewola Osunsade[5,6], Jeannie M. Camarillo[7,8,9], Christopher R. Chin[10,11], Ashley S. Doane[11,12], Bryan J. Venters[13], Stephanie Portillo-Ledesma[14], Joseph Conway[15], Jude M. Phillip[1], Olivier Elemento[10], David W. Scott[16], Wendy Béguelin[1], Jonathan D. Licht[17], Neil L. Kelleher[7,8,9], Louis M. Staudt[18], Arthur I. Skoultchi[19], Michael-Christopher Keogh[13], Effie Apostolou[20,21], Christopher E. Mason[10,22,23], Marcin Imielinski[15], Tamar Schlick[14,24,25], Yael David[5,6], Aristotelis Tsirigos[3,26], C. David Allis[27], Alexey A. Soshnev[27✉], Ethel Cesarman[15✉] & Ari M. Melnick[1✉]

Linker histone H1 proteins bind to nucleosomes and facilitate chromatin compaction[1], although their biological functions are poorly understood. Mutations in the genes that encode H1 isoforms B–E (*H1B*, *H1C*, *H1D* and *H1E*; also known as *H1-5*, *H1-2*, *H1-3* and *H1-4*, respectively) are highly recurrent in B cell lymphomas, but the pathogenic relevance of these mutations to cancer and the mechanisms that are involved are unknown. Here we show that lymphoma-associated H1 alleles are genetic driver mutations in lymphomas. Disruption of H1 function results in a profound architectural remodelling of the genome, which is characterized by large-scale yet focal shifts of chromatin from a compacted to a relaxed state. This decompaction drives distinct changes in epigenetic states, primarily owing to a gain of histone H3 dimethylation at lysine 36 (H3K36me2) and/or loss of repressive H3 trimethylation at lysine 27 (H3K27me3). These changes unlock the expression of stem cell genes that are normally silenced during early development. In mice, loss of *H1c* and *H1e* (also known as *H1f2* and *H1f4*, respectively) conferred germinal centre B cells with enhanced fitness and self-renewal properties, ultimately leading to aggressive lymphomas with an increased repopulating potential. Collectively, our data indicate that H1 proteins are normally required to sequester early developmental genes into architecturally inaccessible genomic compartments. We also establish H1 as a bona fide tumour suppressor and show that mutations in H1 drive malignant transformation primarily through three-dimensional genome reorganization, which leads to epigenetic reprogramming and derepression of developmentally silenced genes.

Linker histones are encoded in humans by ten different genes, five of which (*H1A* (also known as *H1-1*), *H1B*, *H1C*, *H1D* and *H1E*) are expressed in a replication-dependent manner. Linker histones act as transcriptional repressors by limiting chromatin accessibility[2], and are depleted from actively transcribed domains[3]. Their functions are mediated either directly, through condensation of chromatin fibre; or indirectly, through recruitment of transcriptional repressors or by impairing the access of transcriptional activators to core nucleosomes[1]. A triple knockout of three H1 isoforms (H1C, H1D and H1E) in mouse embryonic stem cells impaired differentiation[4], suggesting that H1 is involved in the epigenetic programming of cellular phenotypes. Although mutations in core nucleosomal histones are well-documented as having a role in cancer, little is known about H1 mutations[5,6]. Recurrent H1 mutations occur in around 30–40% of diffuse large B cell lymphomas (DLBCLs), around 30% of follicular lymphomas and around 50% of Hodgkin lymphomas[7–9]. These diseases originate from germinal centre (GC) B cells, which arise transiently from resting B cells during the

T-cell-dependent humoral immune response[10]. During the GC reaction, the immunoglobulin loci undergo extensive mutagenesis through the activity of activation-induced cytosine deaminase (AICDA), and mutant H1 alleles in lymphoma exhibit AICDA mutation signatures[11]. H1 mutations in lymphoma are often highly clonal and their mutational landscape suggests loss of function; however, their effect at the chromatin or functional level has not been defined, although one study of a single H1 mutant reported impaired binding to mononucleosomes[7]. In mice, homozygous knockout of one or both of *H1c* and *H1e*, the most commonly mutated isoforms of H1 in human lymphoma, did not induce a developmental phenotype[12]. Whether or how the dosage or functionality of H1 isoforms contribute to lymphomagenesis is unknown.

Lymphoma H1 alleles are driver mutations

We examined data from the Pan-Cancer Atlas of The Cancer Genome Atlas (TCGA) project and observed that B cell lymphomas have the

---

A list of affiliations appears at the end of the paper.

Article

highest frequency of mutant H1 alleles as compared to all other cancers. The majority (97%) of mutant H1 alleles encode missense mutations that affect the globular and C-terminal domains, with H1C and H1E being the most commonly affected isoforms (Extended Data Fig. 1a, b). Although H1 mutations occurred across DLBCL subtypes, there was significant enrichment for H1 single-nucleotide variants (SNVs) and focal deletions in the newly defined MCD subtype of DLBCLs[13] (Extended Data Fig. 1c–f). We analysed germline-controlled whole-genome sequencing profiles from 101 patients with DLBCL and observed H1 mutation rates of 8.9% (*H1B*), 24.7% (*H1C*), 11.0% (*H1D*) and 42.6% (*H1E*) (Extended Data Fig. 1g). A rigorous analysis controlling for genomic and epigenomic covariates (Supplementary Methods) found that *H1C* and *H1E* were among the top ten driver mutations (Extended Data Fig. 1h). The variant-allele frequency ranged from 0.2 to 0.4, consistent with clonal heterozygous mutation. There was significant co-occurrence between *H1C* and *H1E* mutations, as well as those affecting other H1 alleles (Extended Data Fig. 1i, j). Furthermore, 85% of mutations in *H1B–H1E* that altered the globular domain of the protein scored as deleterious and affected amino acids within conserved interaction interfaces, including an ASGS motif that directly binds to DNA[14] (Extended Data Fig. 2a).

Expression of wild-type or C-terminal-mutant H1C tagged with monomeric eGFP in mouse 3T3 cells showed patterns consistent with localization to chromatin, whereas globular domain mutants affecting the ASGS loop formed extensive nuclear aggregates. In fluorescence recovery after photobleaching (FRAP) assays, wild-type H1C and C-terminal H1C mutants showed similar in vivo dynamics, with recovery rates in the order of minutes (Extended Data Fig. 2b, c). By contrast, globular domain mutants recovered rapidly, consistent with failure to incorporate into chromatin. In line with this, using biolayer interferometry we observed that globular domain mutants had higher mononucleosome dissociation constants than did wild-type or C-terminal mutants. Furthermore, Mg2+ precipitation and atomic force microscopy of 12-mer nucleosome arrays revealed impaired chromatin compaction upon loading of C-terminal-mutant H1C relative to wild-type H1C (Extended Data Fig. 2d–f). Together, these results show that H1 mutations may result in loss of function through several biochemical mechanisms.

H1 deficiency increases GC B cell fitness

Quantitative PCR with reverse transcription (RT–PCR) analysis showed that the expression of *H1B–H1E* was twofold to fourfold higher in GC B cells than in naive B cells (Extended Data Fig. 3a, b). Given the common co-occurrence of *H1C* and *H1E* mutations, we assessed GC formation in *H1c⁻/⁻H1e⁻/⁻* mice, which were previously reported to have no developmental phenotype[2,12]. Immunized *H1c⁻/⁻H1e⁻/⁻* mice had no splenomegaly (Extended Data Fig. 3c) or disruption of splenic architecture (Fig. 1a). However, they exhibited enlarged and more abundant GCs based on PNA (peanut agglutinin) and Ki67⁺ staining (Fig. 1b–d, Extended Data Fig. 3d). There was no observable effect on apoptosis (active caspase-3) or DNA damage (γ-H2AX) (Extended Data Fig. 3e, f). The increase in GC B cells was confirmed by flow cytometry (Extended Data Fig. 3g–i, Supplementary Fig. 2). The proportions of other mature and immature B cells were similar to wild-type mice, with minor differences in the number of Ki67⁺ cells (Extended Data Fig. 3j–m). After sequential immunization with NP-KLH and NP-CGG to induce formation of NP-specific plasma cells, we found no difference between *H1c⁻/⁻H1e⁻/⁻* and wild-type mice in the ratio of high- to low-affinity NP antibody titres or in the number of plasma cells secreting anti-NP immunoglobulins (Extended Data Fig. 3n–q). GCs are composed of a dark zone that contains proliferative B cells (centroblasts) and a light zone that contains mostly non-dividing B cells (centrocytes). *H1c⁻/⁻H1e⁻/⁻* mice showed a selective increase in the number of centrocytes (Extended Data Fig. 3r). To determine whether a lack of H1C and H1E endows GC B cells with a fitness advantage, we performed mixed bone marrow chimaera experiments and observed a robust competitive advantage

for *H1c⁻/⁻H1e⁻/⁻* GC B cells at both time points (day 10 and day 16) as shown by flow cytometry (Fig. 1e–h) and immunofluorescence analysis (Extended Data Fig. 3s, t). The competitive advantage was specific to centrocytes (Fig. 1i, j). Administering 5-ethynyl-2′-deoxyuridine (EdU) revealed a significant increase in the replication of GC B cells that was specific to centrocytes (Fig. 1k, Extended Data Fig. 3u). Therefore, H1 deficiency leads to increased fitness of GC B cells, which manifests as an increase in the self-renewal of centrocytes.

H1 deficiency induces stem cell genes

RNA sequencing (RNA-seq) in sorted *H1c⁻/⁻H1e⁻/⁻* and wild-type GC B cells revealed distinct transcriptional profiles, with 782 differentially expressed genes that were markedly skewed towards transcriptional activation (Fig. 2a, Extended Data Fig. 4a). Many upregulated genes were linked to stem cell functionality, including *Klf4*, *Klf5*, *Meis1*, *Prdm5*, *Mycn*, *Spry2* and *Hoxa9*. Pathway analysis revealed enrichment for (i) signatures associated with partially reprogrammed induced pluripotent stem (iPS) cells, adult tissue stem cells and haematopoietic stem and progenitor cells; (ii) direct targets of stem cell transcription factors such as SOX2, NANOG and SUZ12 (a subunit of polycomb repressive complex 2; PRC2); and (iii) genes marked by H3K27me3 in stem and mature haematopoietic cells, including monovalent H3K27me3 in GC B cells (Fig. 2b). H3K27me3 is formed by the PRC2 complex and is opposed by H3K36me2, which is mediated by NSD2 and related histone methyltransferases[15]. Notably, NSD2 is induced in wild-type human and mouse GC B cells, and *H1c⁻/⁻H1e⁻/⁻* GC B cells showed a marked upregulation of genes activated by NSD2 gain of function in B and T cells[16] (Fig. 2b, Extended Data Fig. 4b, c).

Among normal immune and haematopoietic cells, only long-term repopulating haematopoietic stem cells were enriched for the *H1c⁻/⁻H1e⁻/⁻* signature, and there was significant enrichment for mesenchymal-like state transitions, and cancer stromal cells (Extended Data Fig. 4d, e). Of note, genes repressed by EZH2 in GC B cells through the formation of bivalent chromatin were not de-repressed by loss of *H1c* and *H1e*[17]. Therefore, H1C and H1E deficiency primarily reversed the silencing of developmental PRC2 targets (Fig. 2b, Extended Data Fig. 4f). When we examined RNA-seq profiles from patients with double-mutant DLCBL (that is, with mutations in both *H1C* and *H1E*; hereafter, *H1C/H1E*-mutant DLCBL) and compared these to RNA-seq profiles from patients with DLCBL with wild-type H1 alleles[18], we identified 453 significantly differentially expressed genes in the patients with H1 mutations. These genes were again skewed towards transcriptional upregulation and enrichment for iPS cell signatures, H3K27me3-marked genes in haematopoietic cells and cistromes for NANOG and PRC2 (Extended Data Fig. 4g–i).

To determine whether the effect of H1 deficiency on gene expression was derived from an aberrant subpopulation of GC B cells, we performed single-cell RNA-seq in GC B cells from *H1c⁻/⁻H1e⁻/⁻* and wild-type mice. When we plotted centroblasts and centrocytes across a pseudotime axis (Fig. 2c, Extended Data Fig. 4j, k), we again observed an increased abundance of centrocytes among *H1c⁻/⁻H1e⁻/⁻* GC B cells (Fig. 2d, e). Yet genes upregulated in *H1c⁻/⁻H1e⁻/⁻* GC B cells or *H1C/H1E*-mutant DLBCLs were uniformly upregulated across centroblasts and centrocytes (Fig. 2f, g) with no new subpopulations among H1-deficient GC B cells (data not shown). Notably, *H1c⁻/⁻H1e⁻/⁻* centrocytes (but not centroblasts) showed a significant upregulation of proliferation-related genes (Extended Data Fig. 4l), consistent with aberrant self-renewal being most evident among centrocytes.

H1 deficiency decompacts stem cell genes

We next performed chromatin conformation capture (Hi-C) analyses in sorted *H1c⁻/⁻H1e⁻/⁻* and wild-type GC B cells. These showed distinct contact profiles, and analyses of chromatin compartment
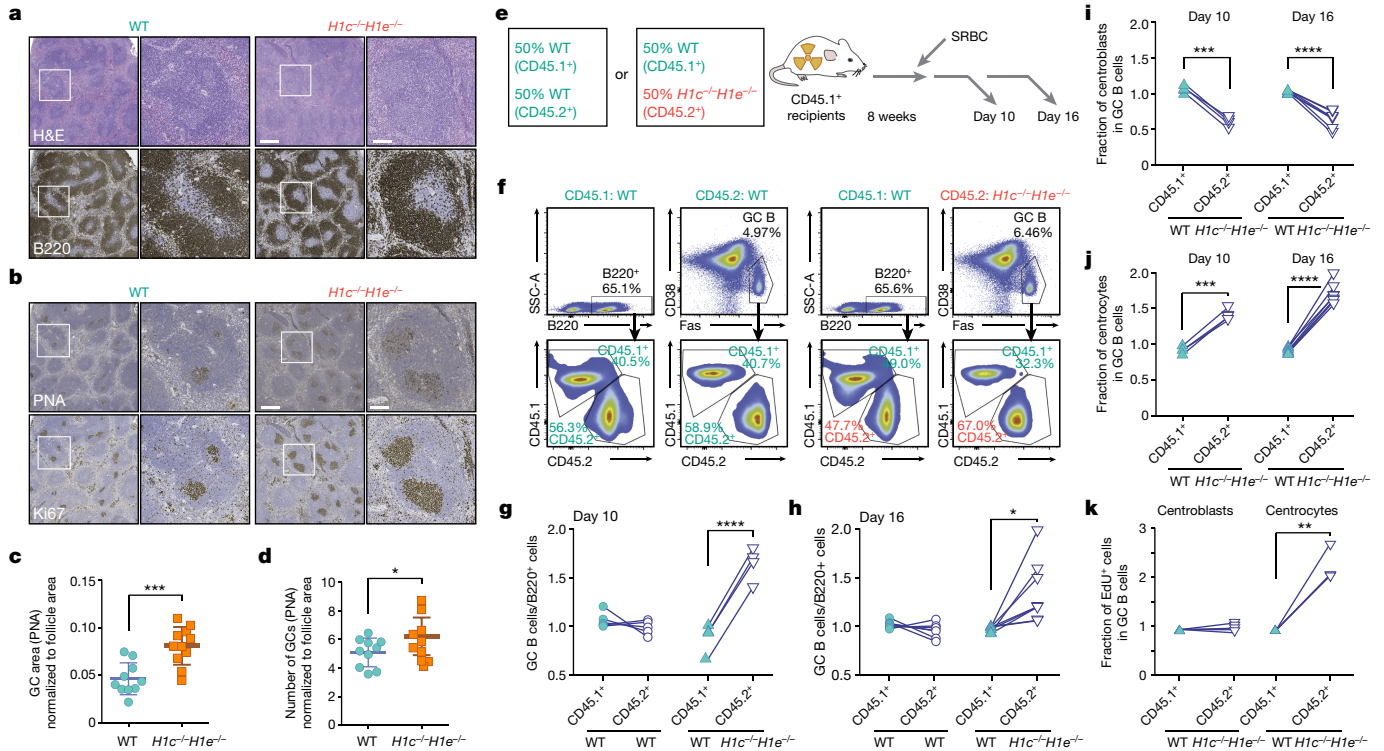
Fig. 1 | Characterization of *H1c[−/−]H1e[−/−]* GC B cells. a, **b**, Spleens of wild-type (WT) and *H1c[−/−]H1e[−/−]* mice at day 9 after immunization with sheep red blood cells (SRBC). Images were stained for haematoxylin and eosin (H&E) or B220 antibody (**a**), or PNA or Ki-67 antibody (**b**). Scale bars, 500 μm (low-magnification images); 100 μm (high-magnification images). Images are representative of three independent experiments. **c**, **d**, Quantification of GC area (**c**; *n* = 11 per genotype, ***P* = 0.0004) and number of GCs (**d**; *n* = 10 per genotype, **P* = 0.044) in the spleens of wild-type and *H1c[−/−]H1e[−/−]* mice. Mean ± s.d., unpaired *t*-test. **e**, Schematic of competitive mixed bone marrow chimaera experiments. **f**, Representative flow cytometry plots at day 10 after immunization for relative fractions of CD45.1[+] and CD45.2[+] cells within naive B220[+] cells and GC B cells. **g**, **h**, Relative ratio of GC B cells to B220[+] cells in chimaeras injected with wild-type CD45.1[+] cells or *H1c[−/−]H1e[−/−]* CD45.2[+] cells, measured at day 10 (**g**; *n* = 4 chimaeras, *****P* < 0.0001) or day 16 (**h**; *n* = 7

chimaeras, **P* = 0.0138) after immunization. Paired *t*-test. GC B/B220[+] ratios in chimaeras injected with wild-type CD45.1[+] cells and wild-type CD45.2[+] cells were unchanged at day 10 (*n* = 5) and day 16 (*n* = 7) after immunization. **i**, Relative fraction of centroblasts (CXCR4[+]CD86[−]) among GC B cells for wild-type (CD45.1[+]) and *H1c[−/−]H1e[−/−]* (CD45.2[+]) fractions in mixed chimaeras at day 10 (*n* = 4, ****P* = 0.0004) or day 16 (*n* = 7, *****P* < 0.0001) after immunization. Paired *t*-test. **j**, Relative fraction of centrocytes (CXCR4[−]CD86[+]) among GC B cells for wild-type (CD45.1[+]) and *H1c[−/−]H1e[−/−]* (CD45.2[+]) fractions at day 10 (*n* = 4, ****P* = 0.0005) or day 16 (*n* = 7, *****P* < 0.0001) after immunization. Paired *t*-test. **k**, Relative fraction of EdU[+] cells among centroblasts (*n* = 4) or centrocytes (*n* = 3, ***P* = 0.0040, two-sided paired *t*-test) for wild-type (CD45.1[+]) and *H1c[−/−]H1e[−/−]* (CD45.2[+]) fractions in mixed chimaeras at day 7 after immunization.

states (compartment scores; c-scores) revealed a strong separation between *H1c[−/−]H1e[−/−]* and wild-type GC B cells (Extended Data Fig. 5a, b). Compartment-B chromatin is highly compacted and transcriptionally silent, whereas compartment A is associated with transcriptionally poised or active chromatin[19]. We observed extensive, yet focal, decompaction that affected 5,320 discrete 100-kb chromatin domains, and increased compaction of only 386 regions (Fig. 3a, b, Extended Data Fig. 5c). Among these, 637 domains shifted entirely from compartment A to compartment A (B to A). The remaining shifts consisted of decompaction within respective compartments (Fig. 3a). Focal compartment shifting occurred across all chromosomes, only excluding regions with the most-extreme compartment-B c-scores, which contain gene deserts and pericentric heterochromatin[20] (Extended Data Fig. 5d, e). Accordingly, regions that underwent B-to-A decompaction had a significantly higher content of genes and CpG islands than did non-shifting B regions (*P* = 0.007 and *P* = 0.0005, respectively). Boundaries of topologically associating domains (TADs) were not significantly affected, but 26 TADs exhibited a significant gain of intra-TAD interactivity (Extended Data Fig. 5f) and were enriched for regions with compartment decompaction (Fig. 3b). Assay for transposase-accessible chromatin using sequencing (ATAC–seq) profiling revealed 488 differentially accessible peaks, 90% of which manifested gain of accessibility (Extended Data Fig. 5g), and significant enrichment within regions that experienced decompaction

(Fig. 3c), whereby more than 99% of differentially accessible sites gained accessibility (Extended Data Fig. 5h).

Chromatin shifts from compartment B to compartment A affected genes that are upregulated in H1-deficient GC B cells, such as *Klf5*, *Meis1*, *Tusc1* and *Spry2* (Fig. 3d, Extended Data Fig. 5i). In line with this, we observed a significant upregulation of genes shifting from B to A (*n* = 224, normalized enrichment score (NES) = 1.73, false discovery rate (FDR) < 0.001), as well as genes decompacting within cognate compartments (Extended Data Fig. 5j, k). Decompacted genes were enriched for iPS cell reprogramming, mesenchymal-transition states, stem cell transcription factor cistromes, H2K27me3-marked genes in haematopoietic cells, and genes that are induced by NSD2 gain of function. GC-B-cell-specific EZH2 targets were unaffected, consistent with a previous report on the absence of H1 from poised gene promoters[21] (Extended Data Fig. 5l).

Notably, we observed a significant association between genomic domains that decompact during early iPS cell reprogramming and those that decompact in H1-deficient GC B cells (Extended Data Fig. 6a–d). Further analysis of the stem-cell-associated *Klf5* locus using virtual 4C (v4C) revealed a gain of promoter interaction with distal elements that shift from B compartments to A compartments in *H1c[−/−]H1e[−/−]* GC B cells, similar to the gains in *Klf5* promoter–enhancer interactions that occur in iPS cells (Extended Data Fig. 6e). Notably, these newly
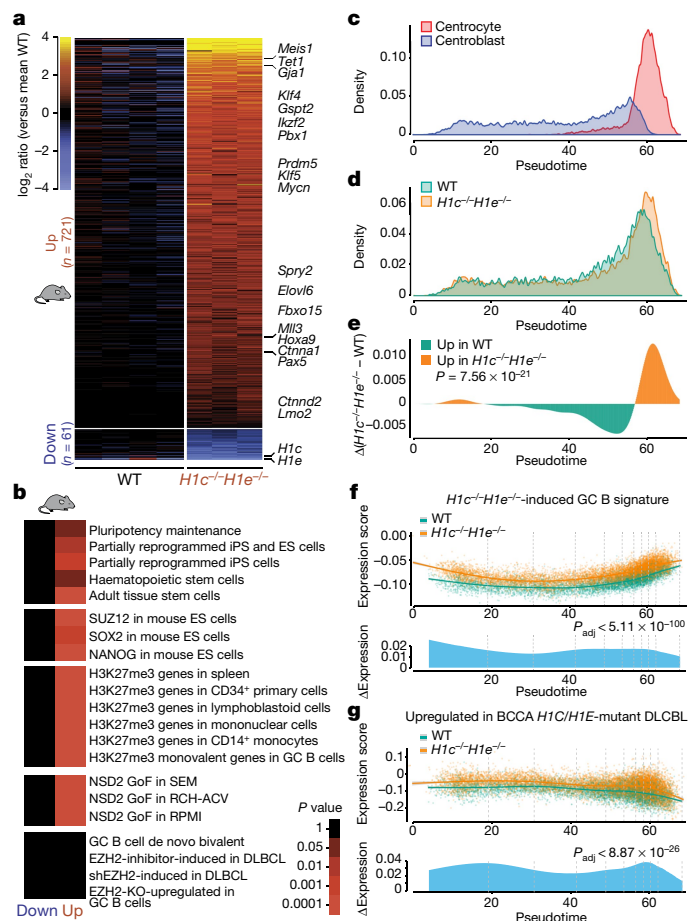
Fig. 2 | Loss of *H1c* and *H1e* induces stem-cell-like transcriptional profiles in GC B cells. a, Heat map of differentially expressed genes as log₂-transformed ratio of every sample to the mean wild-type gene expression (FDR < 0.05, fold change > 1.5) in sorted GC B cells from independent *H1c⁻/⁻H1e⁻/⁻* (*n* = 3) and wild-type (*n* = 4) mice. *Mll3* is also known as *Kmt2c*. **b**, Gene pathway enrichment analysis of upregulated and downregulated genes in *H1c⁻/⁻H1e⁻/⁻* versus wild-type GC B cells (hypergeometric mean test). ES cell, embryonic stem cell; KO, knockout; NSD2 GoF, NSD2 gain of function (in three cell lines: SEM, RCH-ACV and RPMI); shEZH2, short hairpin RNA against EZH2. **c**, Single-cell RNA-seq density plot showing the frequency of centroblasts and centrocytes across the Slingshot pseudotime axis. **d**, Density plot of the frequency of *H1c⁻/⁻H1e⁻/⁻* (*n* = 9,807 cells) and wild-type (*n* = 6,774 cells) GC B cells across the pseudotime axis. Data are pooled from two independent biological replicates. **e**, Differential density plot between *H1c⁻/⁻H1e⁻/⁻* and wild-type GC B cells, with analysis by two-sided Wilcoxon test. **f**, **g**, Top, expression of the upregulated *H1c⁻/⁻H1e⁻/⁻* GC B cell gene signature (**f**) and human *H1C/H1E*-mutant DLBCL gene signature (**g**) was plotted for each cell on the *y* axis, with spline curves representing the average for *H1c⁻/⁻H1e⁻/⁻* and wild-type cells. Bottom, differential expression shown as a delta spline plot across pseudotime, tested by two-sided Wilcoxon test within ten bins of equal cell number (dashed lines). BCCA, British Columbia Cancer Agency.

interactive *Klf5* sites were significantly enriched for two canonical consensus motifs for transcription factors that are associated with GC B cells—OCT2 (*P* = 0.00295) and IRF8 (*P* = 0.00699)—suggesting that decompacted genes could become targets for GC-B-cell associated transcription factors. We wondered whether these iPS-cell-like architectural states would facilitate stem cell reprogramming from differentiated H1C-and-H1E-deficient cells. Indeed, expression of the pluripotency-related transcription factors OCT4, KLF4, SOX2 and MYC (OKSM) in *H1c⁻/⁻H1e⁻/⁻* and wild-type mouse embryonic fibroblasts resulted in a three- to fourfold-increased efficiency in forming H1C-and-H1E-deficient colonies of iPS cells as compared to wild-type

cells (Extended Data Fig. 6f–h). These data suggest that the H1C and H1E isoforms maintain the inactivation of primitive stem cell genes that are silenced during lineage specification and differentiation.

H1 deficiency reprograms epigenetic states

Mass-spectrometry-based quantification of histone post-translational modifications in wild-type and *H1c⁻/⁻H1e⁻/⁻* GC B cells revealed a significant gain of H3K36me1 and H3K36me2—but not H3K36me3—in *H1c⁻/⁻H1e⁻/⁻* cells. As expected given that H3K36me2 can license chromatin for transcription and antagonize PRC2[16], there was a significant reduction in H3K27me2 and H3K27me3, but little change in other histone modifications (Extended Data Fig. 7a, b). Gain of H3K36me2 and loss of H3K27me3 was confirmed by western blotting, whereas the levels of EZH2 and NSD2 were unchanged (Extended Data Fig. 7c). We found little difference in the relative abundance of H3.3 versus H3.1 or H3.2 in *H1c⁻/⁻H1e⁻/⁻* GC B cells, and a similar gain of K36me2 and loss of K27me3 across H3 variants (Extended Data Fig. 7d, e).

Using chromatin immunoprecipitation with reference exogenous genomic DNA (ChIP-RX) analysis for H3K36me2 and H3K27me3, we found that there was a clear difference in the distribution of these marks in *H1c⁻/⁻H1e⁻/⁻* relative to wild-type GC B cells, yielding 7,901 gained versus 33 lost H3K36me2 peaks, and 792 gained versus 4,736 lost H3K27me3 peaks, with a general inverse correlation of these marks (Fig. 4a, Extended Data Fig. 7f, g). Notably, H3K36me2 gain more closely followed B-to-A compartment shifts as compared to H3K27me3 loss (Fig. 4b, Extended Data Fig. 7h). Further analysis of H3K36me2 and H3K27me3 marks in normal GC B cells showed that both were generally absent from compartment-B chromatin, except for H3K27me3 at the least compacted B regions, whereas H3K36me2 increased progressively across compartment-A chromatin, and H3K27me3 was depleted from the most decompacted regions (Extended Data Fig. 7i). Thus, the degree of chromatin interactivity is reflected in alternatively demarcated H3K27me3 or H3K36me2 epigenetic states. In line with this, hierarchical clustering of these histone marks in regions with compartment decompaction revealed five distinct epigenetic states (Fig. 4c, d, Extended Data Fig. 7j): (1) regions that remain in compartment B, with both marks stably absent; (2) regions that start in compartment B, with low H3K36me2 and H3K27me3, and exhibit modest gains of both, with H3K27me3 gains preferentially remaining in compartment B (Extended Data Fig. 7k); (3) regions that start in compartment B or A and undergo moderate gain of H3K26me2 and loss of H3K27me3; (4) regions that start in compartment A and show substantial gain of H3K36me2 and modest loss of H3K27me3; and (5) regions in compartment A that become highly decompacted and mostly show a reduction of H3K27me3.

Genes within the epigenetic state groups 3, 4 and 5 exhibited significant transcriptional activation, which was consistent across centroblasts and centrocytes (Extended Data Fig. 8a, b). Accordingly, upregulated promoters gained H3K4me3 and H3K27ac marks, primarily in compartment A (Extended Data Fig. 8c, d). By contrast, genes in groups 1 and 2 remained mostly silenced. Profiling of the distribution of H3K9me2 and H3K9me3 in wild-type GC B cells revealed that these histone marks were confined to compartment B (Extended Data Fig. 8e). Profiling these marks in *H1c⁻/⁻H1e⁻/⁻* GC B cells showed a reduction in H3K9me3 but not in H3K9me2 peaks (Extended Data Fig. 8f). However, regions undergoing decompaction in *H1c⁻/⁻H1e⁻/⁻* GC B cells exhibited a significant reduction in both H3K9me2 and H3K9me3 (Extended Data Fig. 8g–i). Moreover, H3K9me2 and H3K9me3 were reduced but not completely lost in group 1 and 2 regions, thus marking compacted territories that are more resistant to full decompaction (Extended Data Fig. 8j).

To gain further insight into how the abundance of H1 influences chromatin compaction, we performed in silico modelling of 50-mer nucleosome arrays[22]. At higher concentrations, H1 formed a rigid
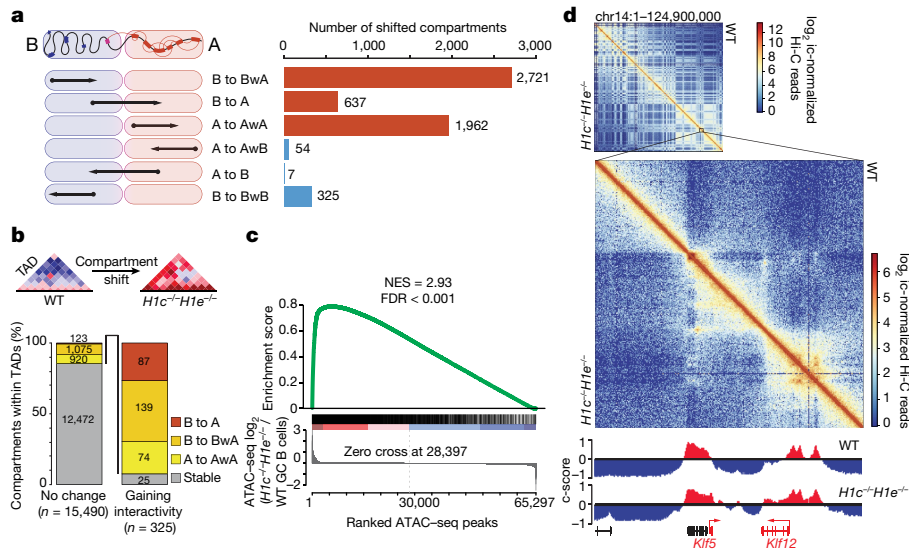
Fig. 3 | Loss of *H1c* and *H1e* induces three-dimensional decompaction in GC B cells. a, Schematic of directionality of shifting within defined compartments (left) and number of shifted compartment regions (100-kb scale) (right), showing decompaction in red bars and compaction in blue bars (BwA, A decompacting shift in compartment B; AwA, A decompacting shift in compartment A). **b**, Bar graph showing the proportion of stable or decompacting 100-kb compartment bins within TADs gaining (*n* = 26 TADs) or showing no change in (*n* = 1,444 TADs) intra-TAD interactivity in *H1c⁻/⁻H1e⁻/⁻* GC B cells. Two sided Fisher's exact test, *P* < 2.2 × 10⁻¹⁶. **c**, Gene set enrichment analysis (GSEA) of ATAC–seq peaks from decompacting compartments against ranked log₂-transformed fold changes in ATAC–seq peaks in *H1c⁻/⁻H1e⁻/⁻* versus wild-type GC B cells. **d**, Top, contact heat map of chromosome 14, showing wild-type GC B cells on the top right and *H1c⁻/⁻H1e⁻/⁻* GC B cells on the bottom left. Bottom, contact heat map in the region containing the *Klf5* locus. Integrated genome viewer (IGV) tracks below represent the eigenvector for compartments A (red) and B (blue) positioned on chr14:98,000,000–101,000,000. ic, iterative correction (normalization method).

stem and straightened chromatin fibres (Extended Data Fig. 9a). These rigid conformations occupied smaller volumes as compared to nucleosome arrays with lower H1 concentrations and demonstrated negligible long-range contacts between nucleosomes. With decreasing H1:nucleosome ratio, fewer DNA linkers were shielded by H1, allowing greater bending and loop formation[22] and resulting in enhanced long-range nucleosome–nucleosome contacts (Extended Data Fig. 9b-f). These results mimic the chromatin state changes that were observed by Hi-C and provide a rationale for the transition states that we captured experimentally.

Collectively, these data suggest that H1 dosage affects gene expression by: (i) sequestering genes within compartment B through dense compaction; (ii) making less-compact compartment-B regions accessible enough to enable PRC2 to form H3K27me3 domains[23]; and (iii) contributing to the regulation of compartment-A genes by establishing chromatin compaction states that are optimal for specific epigenetic complexes (for example, PRC2 and NSD2), with regions that lack H3K27me3 at baseline primarily gaining H3K36me2, and regions that contain H3K27me3 primarily experiencing loss of this mark without substantial gain of H3K36me2. These data suggest that there is a critical functional link between H1 and NSD2, as NSD2 gain of function mimics the H1-deficient transcriptional signature in B cells[16], disrupts genomic compartmentalization[24] and leads to the upregulation of genes associated with epithelial–mesenchymal transition[25].

H1 deficiency induces lymphomagenesis

We crossed *H1c⁻/⁻H1e⁻/⁻* mice with *VavP-Bcl2* mice to model DLBCLs of the MCD subtype (Fig. 5a), as MCD DLBCLs have the highest expression of BCL2 among DLBCLs[13]. At early time points, *VavP-Bcl2* mice had intact lymph node architecture, albeit with hyperplastic follicles containing foci of small lymphocytes with condensed chromatin. By contrast, additional loss of *H1c* and *H1e* resulted in disrupted lymph nodes that were diffusely infiltrated by immunoblastic cells with large nuclei, open chromatin and increased H3K36me2 (Fig. 5b, Extended

Data Fig. 10a, b). *VavP-Bcl2;H1c⁻/⁻H1e⁻/⁻* mice also exhibited a more extensive invasion of the liver and the lungs by neoplastic Ki67⁺ B cells, with T cell infiltrates as often observed in activated B cell-like (ABC) DLBCLs[26], and more evident expansion of monoclonal B cell populations (Fig. 5c, Extended Data Fig. 10c–f). Long-term observation revealed that *VavP-Bcl2;H1c⁺/⁻H1e⁺/⁻* and *VavP-Bcl2;H1c⁻/⁻H1e⁻/⁻* mice had significantly shorter survival times than did *VavP-Bcl2*-only mice (Fig. 5d). *H1c⁻/⁻H1e⁻/⁻* and *H1c⁺/⁻H1e⁺/⁻* mice without *Bcl2* also had shorter survival times compared to controls, and at early time points manifested lymphoproliferative disease invading extranodal tissues (Extended Data Fig. 10g). We noted a trend for greater lethality of the heterozygous condition, with two *VavP-Bcl2;H1c⁺/⁻H1e⁺/⁻* mice already suffering from immunoblastic DLBCL at the early time point, which did not occur in *VavP-Bcl2* or *VavP-Bcl2;H1c⁻/⁻H1e⁻/⁻* mice (Extended Data Fig. 10h). These findings are consistent with H1 mutations being heterozygous in humans, perhaps to avoid loss of isoform-specific functions (for example, mouse H1D uniquely interacts with DNA methyltransferases[8,27]), which might place homozygous mutants at a relative disadvantage.

RNA-seq performed in mouse lymphomas showed that there was significant similarity between the *VavP-Bcl2;H1c⁺/⁻H1e⁺/⁻* and *VavP-Bcl2;H1c⁻/⁻H1e⁻/⁻* transcriptional profiles, as well as *H1c⁻/⁻H1e⁻/⁻* GC B cells (Extended Data Fig. 10i, j). H1-deficient lymphomas in mice exhibited upregulation of many genes—including stem cell factors such as *Klf5* (Extended Data Fig. 10k), which was also induced at the protein level (Fig. 5f). We observed significant enrichment for stem cell and mesenchymal-transition genes, target genes of PRC2 and H3K27me3, and NSD2 gain-of-function signatures (Fig. 5e). The genetic signature of human *H1C/H1E*-mutant DLBCL was significantly enriched in *VavP-Bcl2;H1c⁺/⁻H1e⁺/⁻* lymphomas, with a trend also seen in *VavP-Bcl2;H1c⁻/⁻H1e⁻/⁻* (Extended Data Fig. 10l). The common theme of stem cell signatures among mouse and human H1-deficient lymphomas prompted us to perform secondary transplantation of lymphoma cells from moribund *VavP-Bcl2;H1c⁺/⁻H1e⁺/⁻* or *VavP-Bcl2*-only mice into NOD-SCID recipients (Fig. 5g). All mice were euthanized
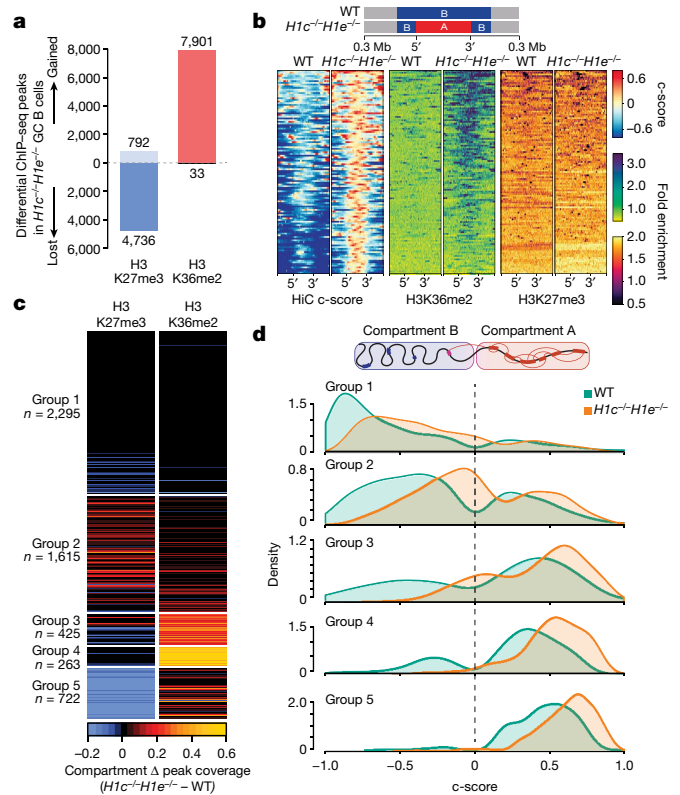
Fig. 4 | Loss of *H1c* and *H1e* reprograms H3K36me2 and H3K27me3 epigenome trajectories. a, Differential chromatin immunoprecipitation followed by sequencing (ChIP–seq) peaks for H3K36me2 and H3K27me3 between *H1c*[−/−]*H1e*[−/−] and wild-type GC B cells (fold change > 1.5). **b**, Heat map of c-score and H3K36me2 and H3K27me3 fold enrichment centred within shifting B-to-A compartment 'islands' (100 kb) and surrounding 300 kb in *H1c*[−/−]*H1e*[−/−] and wild-type GC B cells. **c**, Heat map of changing fraction peak coverage for H3K27me3 and H3K36me2 in *H1c*[−/−]*H1e*[−/−] compared to wild-type GC B cells within decompacting compartments (*n* = 5,320), subdivided into the five groups captured by unsupervised hierarchical clustering. **d**, Density plots showing the distribution of c-scores for shifting compartment groups 1–5 defined in (**c**) for *H1c*[−/−]*H1e*[−/−] versus wild-type GC B cells.

within six weeks, at which point 100% of the *VavP-Bcl2;H1c*[+/−]*H1e*[+/−] mice, but none of the *VavP-Bcl2*-only mice, had developed lymphomas (Fig. 5h, i) Tertiary transfer into recipient NOD-SCID mice again yielded 100% engraftment of *VavP-Bcl2;H1c*[+/−]*H1e*[+/−] cells (Fig. 5h). Hence, loss of H1 endows lymphoma cells with enhanced self-renewal properties, consistent with the highly aggressive nature of H1-mutant DLBCL.

In conclusion, our studies point to decompaction of three-dimensional chromatin as the dominant effect of H1 loss of function in GC B cells. By contrast, however, compartment shifting was not observed in mouse embryonic stem cells with triple knockout of *H1c*, *H1d* and *H1e*[28]. Chromatin compartmentalization by H1 might therefore occur specifically during differentiation, consistent with recent findings in conditional H1 triple-knockout haematopoietic cells[29]. H1 isoforms could thus function as key maintenance factors for compaction and epigenetic memory, particularly for regions that are initially marked for repression by PRC2 during lineage specification. We propose a stoichiometric gradient model to suggest that GC B cells are especially sensitive to H1 dosage (Extended Data Fig. 10m) owing to their extremely rapid proliferative rate[10]. Reduced H1 dosage in this context could impair the maintenance of proper chromatin compartmentalization (and hence 'epigenetic memory') in daughter cells (that is, centrocytes). The greater inefficiency in
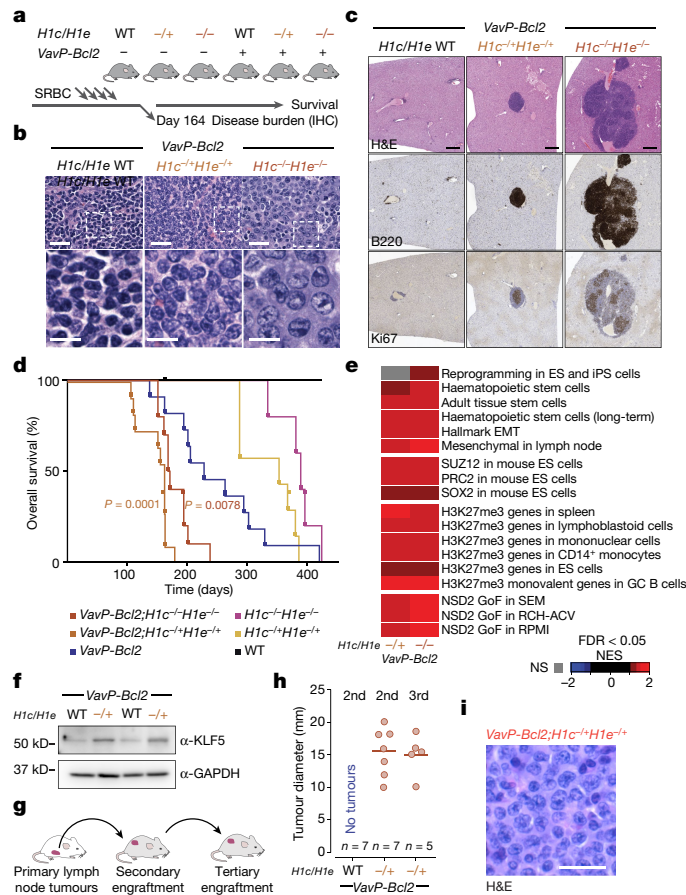
Fig. 5 | Loss of *H1c* and *H1e* leads to aggressive *VavP-Bcl2* lymphomas. a, Experimental scheme for lymphomagenesis. Disease burden was examined by immunohistochemistry (IHC). **b**, IHC images of H&E-stained lymph nodes from mice of indicated genotypes. Scale bars, 25 μm (top); 10 μm (bottom). Images are representative of two independent experiments. **c**, Representative IHC images of liver tissue stained for H&E, B220 and Ki67 from mice of indicated genotypes. Scale bars, 100 μm. **d**, Kaplan–Meier curves depicting overall survival of *VavP-Bcl2;H1c*[−/−]*H1e*[−/−] (*n* = 10), *VavP-Bcl2;H1c*[−/+]*H1e*[−/+] (*n* = 11), *VavP-Bcl2*-only (that is, with wild-type H1 alleles) (*n* = 11), *H1c*[−/+]*H1e*[−/+] (*n* = 7), *H1c*[−/−]*H1e*[−/−] (*n* = 5) and wild-type mice (*n* = 1), assessed by time of death or euthanasia after bone marrow transplantation. *P* values are shown compared to *VavP-Bcl2*-only control (log-rank test). **e**, Normalized enrichment score of GSEA on indicated gene sets, using ranked log[2]-transformed fold change in expression between indicated genotypes versus *VavP-Bcl2*-only control. EMT, epithelial–mesenchymal transition; NS, not significant. **f**, Immunoblot for KLF5 protein from *VavP-Bcl2;H1c*[−/+]*H1e*[−/+] and control *VavP-Bcl2*-only B220-enriched lymphoma cells with GAPDH loading control. Sample pairs collected over two independent experiments are shown. **g**, Schematic of tumour engraftment assay. **h**, Diameter of engrafted tumours after six weeks for secondary (2nd) engraftments (*n* = 7 recipients per genotype) and tertiary (3rd) engraftments of *VavP-Bcl2; H1c*[−/+]*H1e*[−/+] cells into seven recipients. **i**, H&E staining of secondary engrafted *VavP-Bcl2; H1c*[−/+]*H1e*[−/+] lymphomas. Scale bar, 25 μm. Image is representative of *n* = 7 mice.

generating iPS cells from terminally differentiated cells has been attributed to their restrictive epigenetic states[30], an effect that our data suggest is reversed by H1 deficiency. Given that stem-cell-like transcriptional programs are linked to cancer[31], it is possible that silencing of these through H1-mediated compartmentalization represents a major barrier to transformation of mature B cells. The highly recurrent nature of H1 loss-of-function alleles in lymphoma may reflect the fitness advantage conferred to these mature cells by primitive stem cell programs that are normally sequestered within compartment B.

Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41586-020-3017-y.

1. Fyodorov, D. V., Zhou, B. R., Skoultchi, A. I. & Bai, Y. Emerging roles of linker histones in regulating chromatin structure and function. *Nat. Rev. Mol. Cell Biol.* **19**, 192–206 (2018).
2. Fan, Y. et al. Histone H1 depletion in mammals alters global chromatin structure but causes specific changes in gene regulation. *Cell* **123**, 1199–1212 (2005).
3. Cao, K. et al. High-resolution mapping of H1 linker histone variants in embryonic stem cells. *PLoS Genet.* **9**, e1003417 (2013).
4. Zhang, Y. et al. Histone H1 depletion impairs embryonic stem cell differentiation. *PLoS Genet.* **8**, e1002691 (2012).
5. Nacev, B. A. et al. The expanding landscape of 'oncohistone' mutations in human cancers. *Nature* **567**, 473–478 (2019).
6. Bennett, R. L. et al. A mutation in histone H2B represents a new class of oncogenic driver. *Cancer Discov.* **9**, 1438–1451 (2019).
7. Okosun, J. et al. Integrated genomic analysis identifies recurrent mutations and evolution patterns driving the initiation and progression of follicular lymphoma. *Nat. Genet.* **46**, 176–181 (2014).
8. Li, H. et al. Mutations in linker histone genes *HIST1H1 B, C, D*, and *E; OCT2* (*POU2F2*); *IRF8*; and *ARID1A* underlying the pathogenesis of follicular lymphoma. *Blood* **123**, 1487–1498 (2014).
9. Reichel, J. et al. Flow sorting and exome sequencing reveal the oncogenome of primary Hodgkin and Reed-Sternberg cells. *Blood* **125**, 1061–1072 (2015).
10. Mesin, L., Ersching, J. & Victora, G. D. Germinal center B Cell dynamics. *Immunity* **45**, 471–482 (2016).
11. Chapuy, B. et al. Molecular subtypes of diffuse large B cell lymphoma are associated with distinct pathogenic mechanisms and outcomes. *Nat. Med.* **24**, 679–690 (2018).
12. Fan, Y. et al. H1 linker histones are essential for mouse development and affect nucleosome spacing in vivo. *Mol. Cell. Biol.* **23**, 4559–4572 (2003).
13. Wright, G. W. et al. A probabilistic classification tool for genetic subtypes of diffuse large B cell lymphoma with therapeutic implications. *Cancer Cell* **37**, 551–568 (2020).
14. Zhou, B. R. et al. Structural mechanisms of nucleosome recognition by linker histones. *Mol. Cell* **59**, 628–638 (2015).
15. Lu, C. et al. Histone H3K36 mutations promote sarcomagenesis through altered histone methylation landscape. *Science* **352**, 844–849 (2016).
16. Swaroop, A. et al. An activating mutation of the NSD2 histone methyltransferase drives oncogenic reprogramming in acute lymphocytic leukemia. *Oncogene* **38**, 671–686 (2019).
17. Béguelin, W. et al. EZH2 is required for germinal center formation and somatic EZH2 mutations promote lymphoid transformation. *Cancer Cell* **23**, 677–692 (2013).
18. Ortega-Molina, A. et al. The histone lysine methyltransferase KMT2D sustains a gene expression program that represses B cell lymphoma development. *Nat. Med.* **21**, 1199–1208 (2015).
19. Eagen, K. P. Principles of chromosome architecture revealed by Hi-C. *Trends Biochem. Sci.* **43**, 469–478 (2018).
20. Rao, S. S. et al. A 3D map of the human genome at kilobase resolution reveals principles of chromatin looping. *Cell* **159**, 1665–1680 (2014).
21. Izzo, A. et al. The genomic landscape of the somatic linker histone subtypes H1.1 to H1.5 in human cells. *Cell Rep.* **3**, 2142–2154 (2013).
22. Bascom, G. D., Myers, C. G. & Schlick, T. Mesoscale modeling reveals formation of an epigenetically driven HOXC gene hub. *Proc. Natl Acad. Sci. USA* **116**, 4955–4962 (2019).
23. Kim, J. M. et al. Linker histone H1.2 establishes chromatin compaction and gene silencing through recognition of H3K27me3. *Sci. Rep.* **5**, 16714 (2015).
24. Lhoumaud, P. et al. NSD2 overexpression drives clustered chromatin and transcriptional changes in a subset of insulated domains. *Nat. Commun.* **10**, 4843 (2019).
25. Ezponda, T. et al. The histone methyltransferase MMSET/WHSC1 activates TWIST1 to promote an epithelial–mesenchymal transition and invasive properties of prostate cancer. *Oncogene* **32**, 2882–2890 (2013).
26. Mulder, T. A., Wahlin, B. E., Österborg, A. & Palma, M. Targeting the immune microenvironment in lymphomas of B cell origin: from biology to clinical application. *Cancers* **11**, 915 (2019).
27. Yang, S. M., Kim, B. J., Norwood Toro, L. & Skoultchi, A. I. H1 linker histone promotes epigenetic silencing by regulating both DNA methylation and histone H3 methylation. *Proc. Natl Acad. Sci. USA* **110**, 1708–1713 (2013).
28. Geeven, G. et al. Local compartment changes and regulatory landscape alterations in histone H1-depleted cells. *Genome Biol.* **16**, 289 (2015).
29. Wilcockson, M. et al. H1 histones control the epigenetic landscape by local chromatin compaction. *Nature* https://doi.org/10.1038/s41586-020-3032-z (2020).
30. Eminli, S. et al. Differentiation stage determines potential of hematopoietic cells for reprogramming into induced pluripotent stem cells. *Nat. Genet.* **41**, 968–976 (2009).
31. Wong, D. J. et al. Module map of stem cell genes guides creation of epithelial cancer stem cells. *Cell Stem Cell* **2**, 333–344 (2008).

¹Division of Hematology and Medical Oncology, Sanford I. Weill Department of Medicine, Weill Cornell Medicine, New York, NY, USA. ²Cell and Molecular Biology Graduate Program, Weill Cornell Medicine, New York, NY, USA. ³Department of Pathology, NYU School of Medicine, New York, NY, USA. ⁴Department of Computational Biology of Infection Research, Helmholtz Centre for Infection Research, Braunschweig, Germany. ⁵Chemical Biology Program, Memorial Sloan Kettering Cancer Center, New York, NY, USA. ⁶Tri-Institutional PhD Program in Chemical Biology, New York, NY, USA. ⁷Department of Chemistry, Northwestern University, Evanston, IL, USA. ⁸Department of Molecular Biosciences, Northwestern University, Evanston, IL, USA. ⁹Proteomics Center of Excellence, Northwestern University, Evanston, IL, USA. ¹⁰Department of Physiology and Biophysics, Weill Cornell Medicine, New York, NY, USA. ¹¹Tri-Institutional PhD Program in Computational Biomedicine, New York, NY, USA. ¹²Institute for Computational Biomedicine, Department of Physiology and Biophysics, Weill Cornell Medicine, New York, NY, USA. ¹³EpiCypher, Durham, NC, USA. ¹⁴Department of Chemistry, New York University, New York, NY, USA. ¹⁵Department of Pathology and Laboratory Medicine, Weill Cornell Medicine, New York, NY, USA. ¹⁶Centre for Lymphoid Cancer, BC Cancer, Vancouver, British Columbia, Canada. ¹⁷University of Florida Health Cancer Center, The University of Florida Cancer and Genetics Research Complex, Gainesville, FL, USA. ¹⁸Lymphoid Malignancies Branch, National Cancer Institute, National Institutes of Health, Bethesda, MD, USA. ¹⁹Department of Cell Biology, Albert Einstein College of Medicine, New York, NY, USA. ²⁰Sanford I. Weill Department of Medicine, Weill Cornell Medicine, New York, NY, USA. ²¹Sandra and Edward Meyer Cancer Center, Weill Cornell Medicine, New York, NY, USA. ²²The WorldQuant Initiative for Quantitative Prediction, Weill Cornell Medicine, New York, NY, USA. ²³The Feil Family Brain and Mind Research Institute, Weill Cornell Medicine, New York, NY, USA. ²⁴Courant Institute of Mathematical Sciences, New York University, New York, NY, USA. ²⁵New York University–East China Normal University Center for Computational Chemistry at New York University Shanghai, Shanghai, China. ²⁶Institute for Computational Medicine, NYU School of Medicine, New York, NY, USA. ²⁷Laboratory of Chromatin Biology and Epigenetics, The Rockefeller University, New York, NY, USA. ✉e-mail: asoshnev@rockefeller.edu; ecesarm@med.cornell.edu; amm2014@med.cornell.edu

Article

32. Stadhouders, R. et al. Transcription factors orchestrate dynamic interplay between genome topology and gene regulation during cell reprogramming. *Nat. Genet.* **50**, 238–249 (2018).

Extended Data Fig. 1 | Characterization of H1 allele mutations and deletions in DLBCL. a, A PanCancer Atlas cancer mutation survey was performed using the cBioPortal to search the cumulative mutation frequency of *H1A–H1E* across a total of 10,953 non-redundant patient samples across all cancer types. **b**, *H1B–H1E* mutation landscape across non-redundant TCGA and BCCA (British Columbia Cancer Agency) DLBCL samples (number of samples indicated for each H1 isoform); location of mutations is plotted onto protein structure, with bar height corresponding to mutation counts, total frequency of mutations in each H1 isoform is plotted to the right (red bars). **c**, Prevalence of mutant H1 versus wild-type H1 cases in ABC and GC B DLBCLs (Fisher's exact test for enrichment in ABC or GC B-DLBCL, *P* > 0.05) **d**, Prevalence of *H1A–H1E* missense mutations and heterozygous loss in MCD-subtype DLBCL and their enrichment (−log₁₀(*P* value)) over non-MCD DLBCLs. **e**, Summary table of frequency of

mutations and heterozygous loss of *H1A–H1E* in DLBCL subtypes (MCD, ST2, BN2, EZB, N1 and AP53). **f**, Co-occurrence as odds ratio and −log₁₀(*P* value) among H1 alleles in 101 germline-matched WGS DLBCLs. **g**, Frequency and location of missense mutations for *H1A–H1E* in 101 germline-controlled DLBCLs. **h**, Quantile-quantile plot showing the *P* values for SNVs across 101 germline-matched WGS DLBCLs. Driver analysis derived as probability of mutation count greater than or equal to the observed mutation count under *P* values expected from a Gamma-Poisson distribution for SNVs. Shaded grey zone contains mutant genes with FDR < 0.01, depicted as blue or red dots, several of which are listed in the figure. **i**, Oncoprint for *H1A–H1E* in 101 germline-matched WGS DLBCLs. **j**, Co-occurrence as odds ratio (OR) and −log₁₀(*P* value) among H1 alleles in 101 germline-matched WGS DLBCLs.
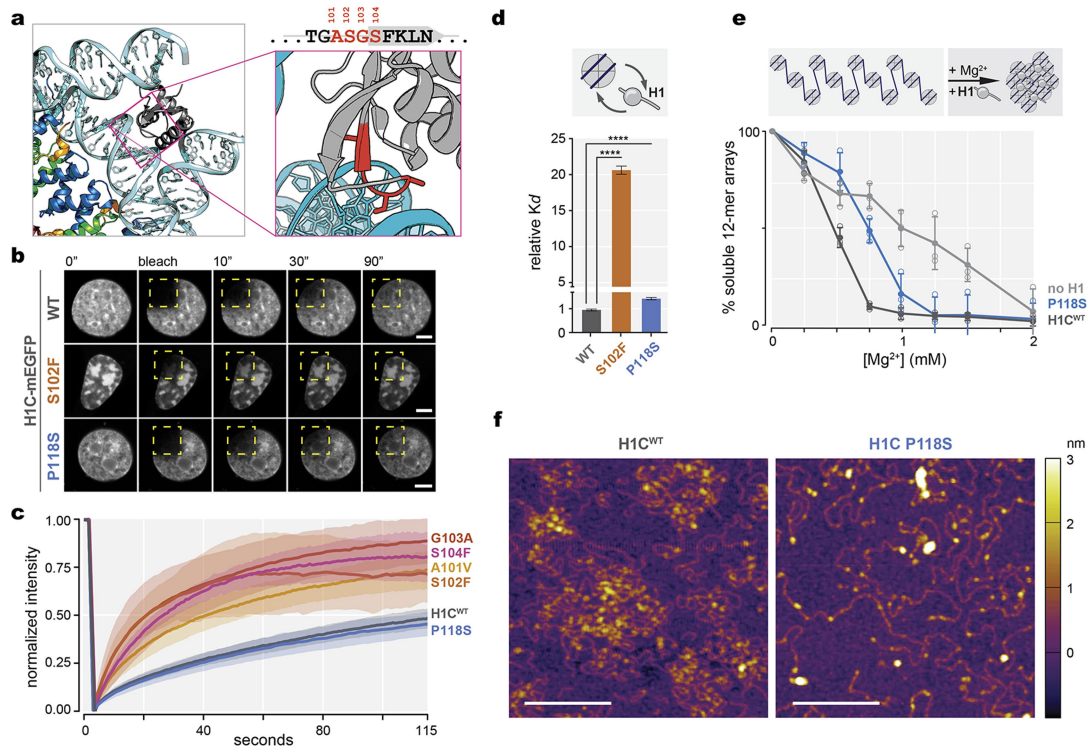
Extended Data Fig. 2 | H1 mutations are genetic drivers in lymphoma and confer loss of function. a, Crystal structure of the linker histone globular domain (grey) bound to nucleosome (Protein Data Bank (PDB), 4QLC), with magnified view of ASGS amino acid residues highlighted in red. **b**, Representative images of FRAP of ectopically expressed, monomeric eGFP-tagged wild-type H1C and H1C mutants S102F and P118S in 3T3 cells before, immediately after, and at 10, 30 and 90 s after bleaching the area (yellow dashed square). Scale bars, 5 μm. **c**, Quantification of normalized intensity as representation of turnover kinetics from **b** for wild-type H1C ($n = 18$) and mutants A101V ($n = 15$), S102F ($n = 9$), S104F ($n = 10$), G103A ($n = 10$) and P118S ($n = 10$); shaded area indicates 95% confidence intervals. Data are pooled from two independent biological experiments. **d**, Dissociation constant (K_d) of the binding to mononucleosomes of recombinant H1C mutants S102F and P118S compared to wild-type H1C, as determined by biolayer interferometry. Data are mean ± s.e.m. (two-sided unpaired t-test, ****$P < 0.0001$). Data are global fits from five concentration measurements. **e**, Chromatin fibre oligomerization upon serial precipitation by Mg2+ as per cent soluble 12-mer arrays was determined for no H1, wild-type H1C and the C-terminal P118S mutant. Data are mean ± s.d. Data are pooled from three independent biological experiments. **f**, Atomic force microscopy imaging of chromatin arrays in the presence of wild-type H1 and the C-terminal P118S mutant. Scale bars, 200 nm. Images are representative of two independent biological experiments.
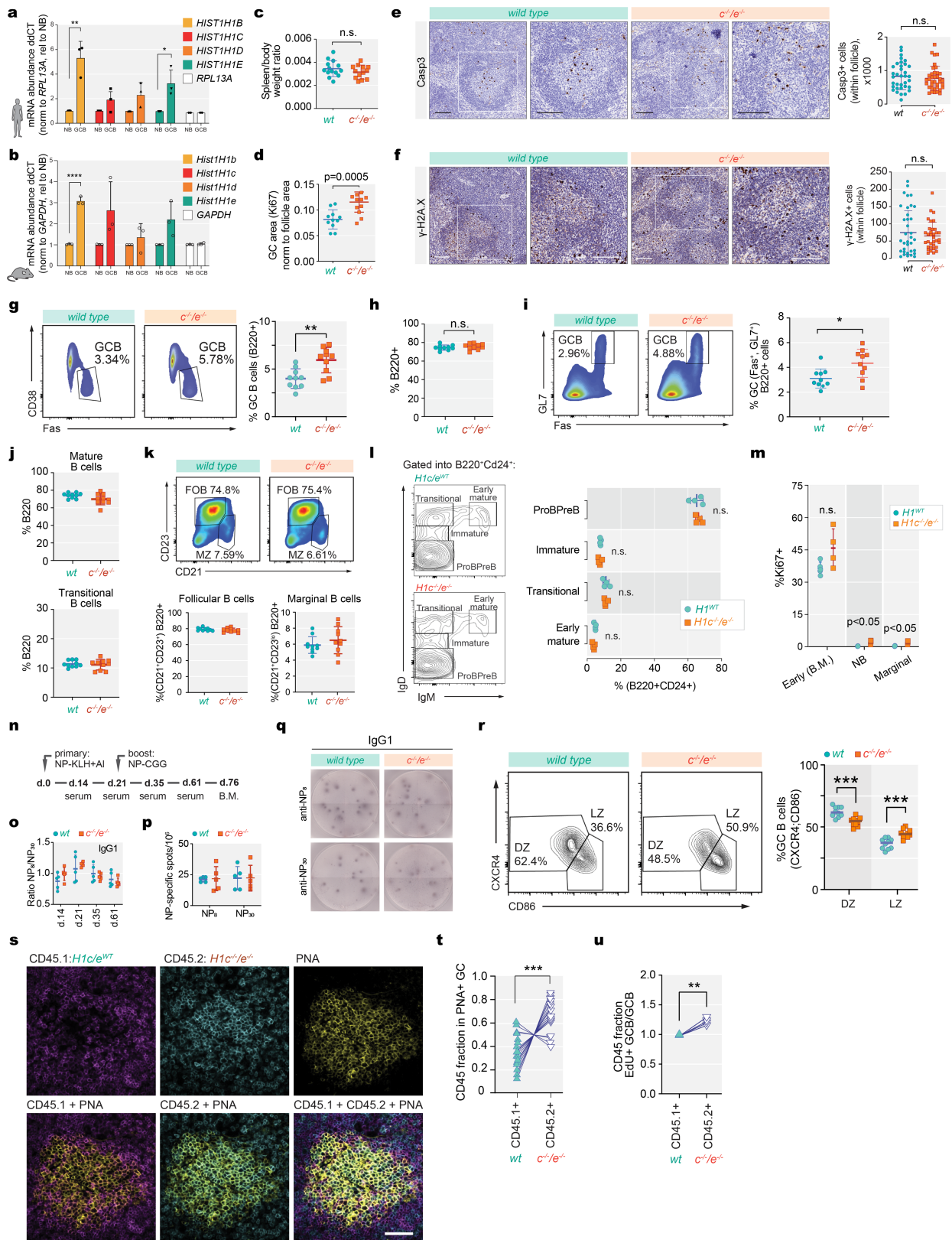
Extended Data Fig. 3 | See next page for caption.

Article

**Extended Data Fig. 3 | $H1c^{-/-}H1e^{-/-}$ mature B cells show normal development in spleen and bone marrow. a**, mRNA of human $H1B-H1E$ normalized to $RPL13A$ in GC B cells relative to naive B cells ($H1B$, **$P = 0.004$; $H1E$, *$P = 0.027$), isolated from three independent specimens of human tonsils. Data are mean ± s.d., two-sided unpaired t-tests. **b**, Mouse $H1b-H1e$ mRNA levels normalized to $Gapdh$ in sorted GC B cells ($n = 3$) relative to naive B cells ($n = 3$) ($H1b$, ****$P < 0.0001$). Data are mean ± s.d., two-sided unpaired t-tests. **c**, Quantification of spleen/body weight ratios of two-month-old $H1c^{-/-}H1e^{-/-}$ ($n = 13$) and wild-type littermate control ($n = 14$) mice. Data are pooled from two independent experiments. $P < 0.05$; NS, not significant; two-sided unpaired t-tests. Data are mean ± s.d. **d**, Quantification of GC area (Ki67 staining) in the spleens of $H1c^{-/-}H1e^{-/-}$ ($n = 10$) and WT ($n = 10$) mice. ***$P = 0.0005$. Data are mean ± s.d., two-sided unpaired t-tests. **e**, **f**, Immunohistochemistry images of spleen sections of cleaved caspase-3 (**e**) and γ-H2AX (**f**) staining (left) and quantification (right) of positively stained follicular cells from $H1c^{-/-}H1e^{-/-}$ ($n = 3$) and wild-type littermate control ($n = 3$) mice immunized with SRBCs and euthanized 10 days after immunization. Scale bars, 100 μm. $P < 0.05$ (not significant; NS), two-sided unpaired t-tests. Data are mean ± s.d. **g**, Flow cytometry analysis and quantification of (Fas$^+$CD38$^-$) GC B cells within total B cells from $H1c^{-/-}H1e^{-/-}$ and wild-type mice ($n = 10$ per genotype). **$P = 0.0018$, two-sided unpaired t-tests. Data are mean ± s.d. **h**, Quantification of the percentage of B220$^+$ splenocytes in $H1c^{-/-}H1e^{-/-}$ ($n = 10$) and wild-type ($n = 10$) mice 9 days after SRBC immunization. $P < 0.05$, two-sided unpaired t-tests. Data are mean ± s.d. **i**, Flow cytometry analysis and quantification of GC B cells (Fas$^+$GL7$^+$) from $H1c^{-/-}H1e^{-/-}$ ($n = 10$) and wild-type ($n = 10$) mice. Two-sided unpaired t-tests, *$P = 0.041$. Data are mean ± s.d. **j**, Flow cytometry analysis and quantification of mature B cells (B220$^+$IgD$^+$IgM$^+$) and transitional B cells (B220$^+$IgD^{int}IgM$^+$) in spleens from $H1c^{-/-}H1e^{-/-}$ ($n = 10$) and wild-type ($n = 10$) mice. $P < 0.05$, two-sided unpaired t-tests. Data are mean ± s.d. **k**, Flow cytometry quantification of follicular B cells (B220$^+$D23$^+$CD21$^+$) and marginal zone B cells (B220$^+$D23^{lo}CD21$^+$) in spleens from $H1c^{-/-}H1e^{-/-}$ ($n = 10$) and

wild-type ($n = 10$) mice. $P < 0.05$, two-sided unpaired t-tests. Data are mean ± s.d. **l**, Flow cytometry analysis gated on B220$^+$CD24$^+$ and quantification of ProBPreB (IgM$^-$IgD$^-$), immature (IgM$^-$IgD^{lo}), transitional (IgD$^+$IgM$^-$) and early mature (IgD$^+$IgM$^+$) B cells in bone marrow of $H1c^{-/-}H1e^{-/-}$ ($n = 4$) and wild-type ($n = 5$) mice. $P < 0.05$, two-sided unpaired t-tests. Data are mean ± s.d. **m**, Percentage of Ki67$^+$ early B cells (B220$^+$CD24$^+$) in bone-marrow of $H1c^{-/-}H1e^{-/-}$ ($n = 4$) and wild-type ($n = 5$) mice, as well as naive B cells (***$P = 0.0004$) and marginal zone B cells (***$P = 0.001$) in the spleens of $H1c^{-/-}H1e^{-/-}$ ($n = 5$) and wild-type ($n = 5$) mice. **n**, Schematic diagram of primary immunization with NP-KLH and secondary immunization 21 days after with NP-CGG. **o**, Ratio between high (NP$_8$) and low (NP$_{30}$) affinity NP-specific IgG1 antibody titres in sera of $H1c^{-/-}H1e^{-/-}$ ($n = 5$) and wild-type ($n = 5$) mice by enzyme-linked immunosorbent assay (ELISA). $P < 0.05$, two-sided unpaired t-tests. Data are mean ± s.d. **p**, Enzyme-linked immunosorbent spot (ELISPOT) quantification of NP-specific (anti-NP$_8$ and anti-NP$_{30}$) IgG1-secreting cells from the bone marrow of $H1c^{-/-}H1e^{-/-}$ ($n = 5$) and wild-type ($n = 5$) mice. $P < 0.05$, two-sided unpaired t-test. Data are mean ± s.d. Data are representative of two independent experiments. **q**, Representative images of anti-NP$_8$ and anti-NP$_{30}$ 96-well ELISPOT. **r**, Flow cytometry analysis and quantification of centroblasts within the dark zone (DZ) (CXCR4$^+$CD86$^-$) (***$P = 0.0002$) and centrocytes within the light zone (LZ) (CXCR4$^-$CD86$^+$) (***$P = 0.0002$) within GC B cells from $H1c^{-/-}H1e^{-/-}$ ($n = 10$) and wild-type ($n = 10$) mice. Two-sided unpaired t-test. Data are representative of three independent experiments. **s**, Immunofluorescence confocal microscopy images of GCs at day 7 after immunization in mixed chimaeras. Scale bar, 50 μm. Images are representative of two independent experiments. **t**, Quantification from **s** of the fraction of PNA$^+$CD45.1 or CD45.2 cells (17 GCs, $n = 3$ mice). Two-sided paired t-test, ***$P = 0.0004$. **u**, Relative EdU$^+$ GC B cell/GC B cell fraction for wild-type CD45.1$^+$ and $H1c^{-/-}H1e^{-/-}$ CD45.2$^+$ cells at day 7 after immunization ($n = 4$ chimaeras). Two-sided paired t-test, **$P = 0.0065$. Data are representative of two independent experiments.
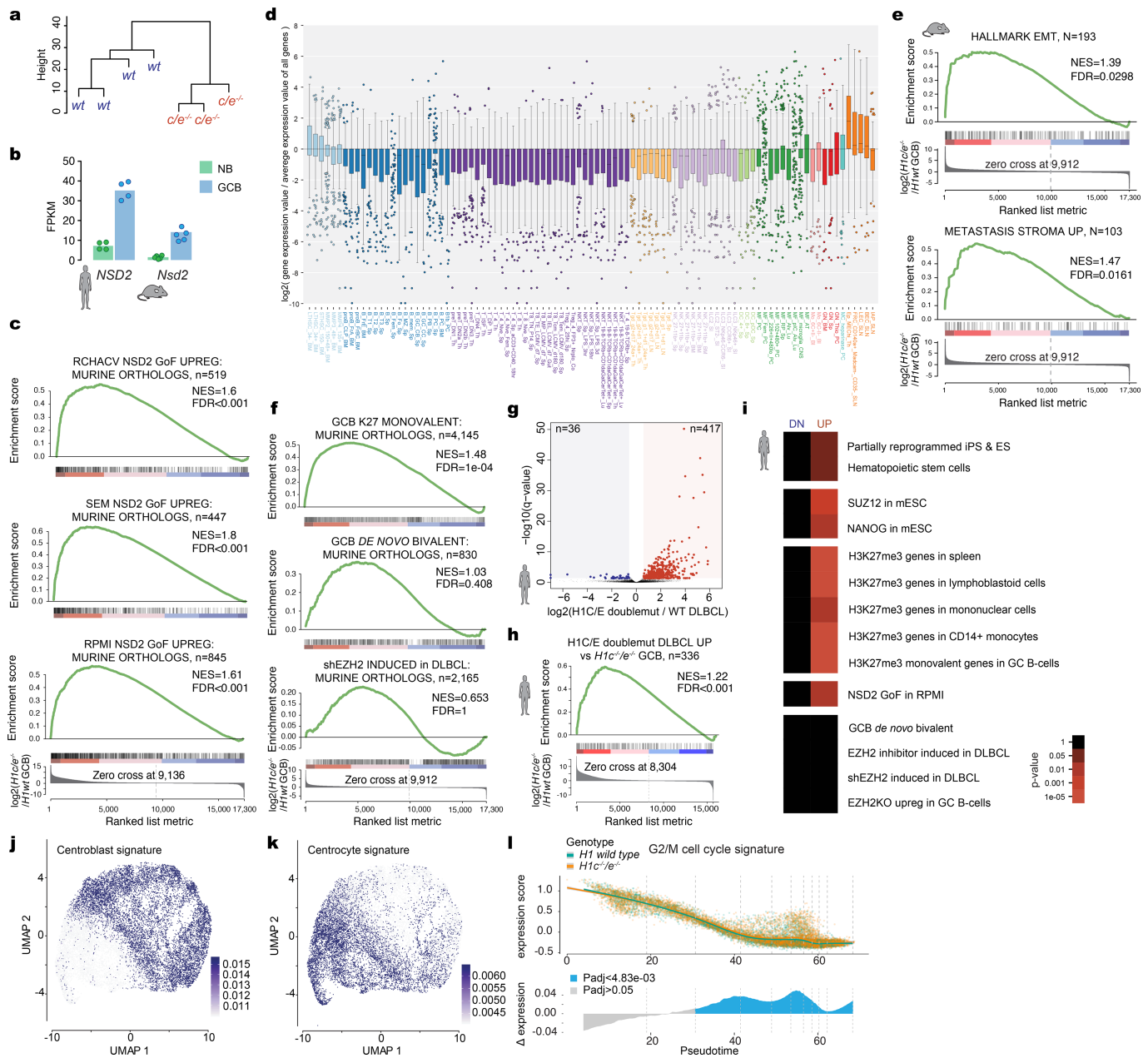
Extended Data Fig. 4 | Loss of *H1c* and *H1e* induces stem-cell-like transcriptional profiles in GC B cells and DLBCLs. a, Unsupervised hierarchical clustering analysis of RNA-seq data from sorted *H1c*−/−*H1e*−/− and wild-type H1 GC B cells, based on genes in the top 90th percentile of variability. **b**, Expression of *NSD2* (measured as fragments per kilobase of exon model per million reads mapped; FPKM) from RNA-seq profiles of human and mouse naive B and GC B cells. **c**, GSEA analysis of genes linked to *NSD2* gain-of-function mutation in three cell lines (RCHACV, SEM, RPMI) against ranked gene expression changes in mouse *H1c*−/−*H1e*−/− GC B cells. **d**, Box plot of log2-transformed relative gene expression normalized to the average expression value of all genes from the top 200 differentially upregulated genes in the *H1c*−/−*H1e*−/− signature against the ImmGen database. Box plot centre represents median, bounds of box are the first and third quartiles and whiskers extend to 1.5 × the interquartile range. **e**, GSEA analysis with indicated gene sets, using ranked log2-transformed fold change in expression between *H1c*−/−*H1e*−/− and wild-type mouse GC B cells. **f**, GSEA analysis of gene sets linked to EZH2 against ranked gene expression changes in mouse *H1c*−/−*H1e*−/− GC B

cells. **g**, Volcano plot showing differentially expressed genes between patients with *H1C/E*-mutant DLCBL (*n* = 18) versus patients with H1-wild-type DLBCL (*n* = 237) (FDR < 0.05, fold change > 1.5). Red field denotes upregulated and blue field downregulated genes. **h**, GSEA analysis of upregulated genes in patients with *H1C/E*-mutant versus H1-wild-type DLBCL, using ranked changes in log2-transformed fold change in mouse *H1c*−/−*H1e*−/− GC B cells. **i**, Gene pathway enrichment analysis of upregulated and downregulated genes from **g**, hypergeometric mean test. **j**, **k**, Sorted *H1c*−/−*H1e*−/− or wild-type GC B cells (*n* = 2 per genotype) were subjected to droplet based (10X) single-cell RNA-seq. Centroblasts (**j**) and centrocytes (**k**) were defined on the basis of enrichment for centroblast and centrocyte signature profiles, respectively projected onto the uniform manifold approximation and projection (UMAP) distribution of cells. **l**, Top, expression of G2M cell cycle proliferation gene signature was plotted for each cell on the *y* axis with spline curves representing the average for *H1c*−/−*H1e*−/− and wild-type GC B cells. Bottom, differential expression is shown as a delta spline plot (blue) across pseudotime and tested by two-sided Wilcoxon test within ten bins of equal cell number (dashed lines).
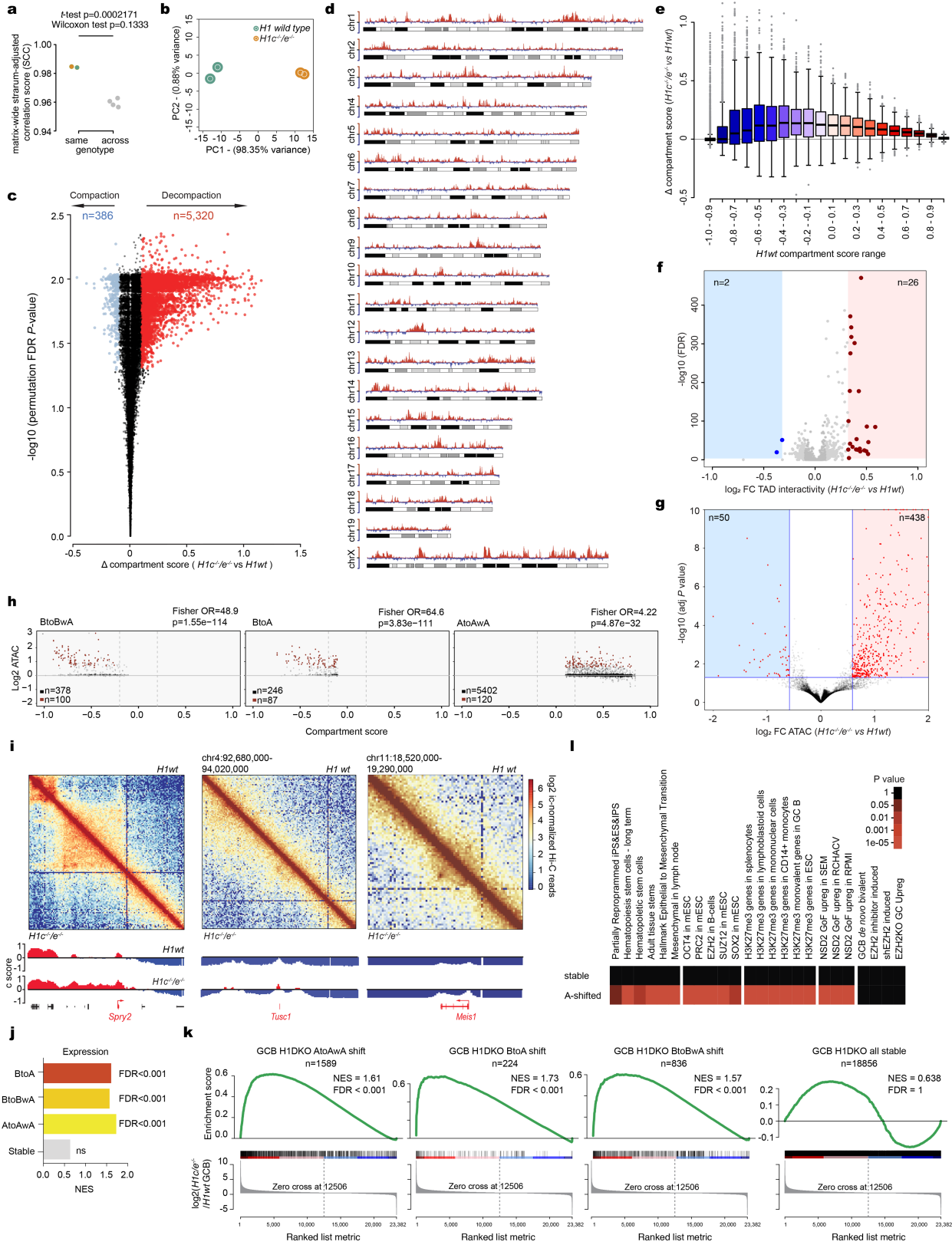
Extended Data Fig. 5 | See next page for caption.

Extended Data Fig. 5 | HiC compartment analysis and integration with ATAC–seq in $H1c^{-/-}H1e^{-/-}$ and wild-type GC B cells. a, Genome-wide correlation score (Stratum adjusted correlation score, SCC) of HiC matrices within same genotype (***P = 0.0002171, two-sided unpaired t-test), and across genotypes (P < 0.05, two-sided Wilcoxon test). **b**, Principal component analysis (PCA) of compartment bins processed with Hi-C bench at a resolution of 100kb from Hi-C replicates for $H1c^{-/-}H1e^{-/-}$ and wild-type GC B cells. **c**, Volcano plot showing significant c-score shifts in $H1c^{-/-}H1e^{-/-}$ mouse GC B cells based on Δ c-score and −log10(permutation FDR P value): decompacted (red) compartment bins (n = 5,320) and compacted (blue) compartment bins (n = 386). **d**, Δ c-score ($H1c^{-/-}H1e^{-/-}$ versus wild type) across mouse chromosomes (positive y axis is in red for decompacting loci and negative y axis is blue for compacting loci) plotted as ideograms. **e**, Box plots of Δ c-score ($H1c^{-/-}H1e^{-/-}$ versus wild type) across c-score range (−1.0 to 1.0) separated into 0.1 bins. No statistical evaluation was derived for this graph. **f**, Volcano plot showing TADs with significant gain of intra-TAD interactivity (n = 26, red) and reduction of intra-TAD interactivity (n = 2, blue; two-sided unpaired t-test, FDR-adjusted P value < 0.05, fold change > 1.5). **g**, Volcano plot showing significant ATAC–seq peaks gaining accessibility (n = 438) and losing accessibility (n = 53; two-sided unpaired t-test FDR-adjusted P value < 0.05, fold-change > 1.5).

h, Scatter plot showing log$_2$-transformed fold change in ATAC–seq peaks between $H1c^{-/-}H1e^{-/-}$ and wild-type GC B cells in decompacting compartments versus wild-type c-score: B to BwA (left), B to A (middle), and A to AwA (right). Peaks gaining accessibility (FC > 1.5, P_{adj} < 0.05) are marked in red. Estimated odds ratios and P values were calculated using Fisher's exact test. Although ATAC–seq peaks are more prevalent in the A-to-AwA compartment (right), increased accessibility is more enriched in B-to-BwA and B-to-A compartments. **i**, HiC contact maps of regions surrounding *Spry* (left), *Tusc1* (middle) and *Meis1* (right) genes. The top of each square shows H1 wild-type contacts and the bottom of each square shows those in $H1c^{-/-}H1e^{-/-}$ GC B cells. Heat maps represent the Pearson correlation of interactions in wild-type and $H1c^{-/-}H1e^{-/-}$ GC B cells. Bottom tracks represent the eigenvector (PC1) for compartments A and B in red and blue, respectively, and show the position of genes within these loci. **j**, **k**, GSEA analysis of genes shifting to A compartments (B to A, B to BwA, A to AwA) or genes contained in stable compartments using ranked log$_2$-transformed gene expression in $H1c^{-/-}H1e^{-/-}$ GC B cells. (NES and FDR values as implemented by GSEA). **l**, Gene pathway enrichment analysis of genes in decompacting and stable compartments in $H1c^{-/-}H1e^{-/-}$ versus wild-type GC B cells (hypergeometric mean test).

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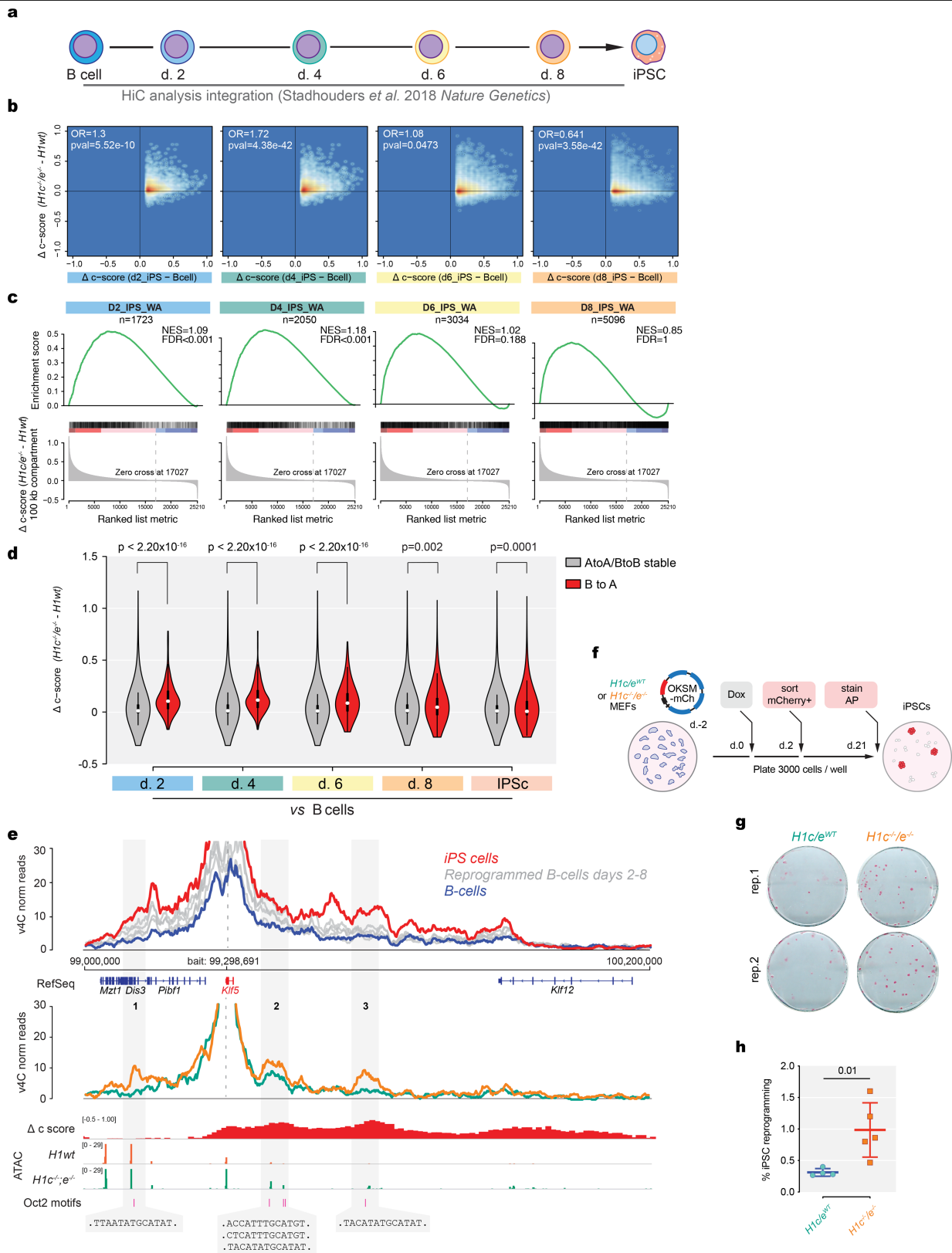
Extended Data Fig. 6 | See next page for caption.

Extended Data Fig. 6 | Three-dimensional changes in chromatin architecture owing to a deficiency of H1C and H1E in GC B cells recapitulate decompaction during iPS cell differentiation. a, Schematic of a previous study of B cell differentiation to iPS cells[32] with time points for HiC analysis. **b**, Correlation plots comparing shifting to A c-scores in iPS cells (day 2, day 4, day 6, day 8) compared to control B cells versus c-score changes in $H1c^{-/-}H1e^{-/-}$ GC B cells. Estimated odds ratios and P values were calculated using Fisher's exact test. **c**, GSEA analysis of shifting to A compartments in iPS cells (day 2, day 4, day 6, day 8) against ranked Δ c-scores derived from $H1c^{-/-}H1e^{-/-}$ minus wild-type mouse GC B cells (NES and FDR values as implemented by GSEA). **d**, Violin plots comparing stable and shifting B to A compartments during iPS cell differentiation (day 2, day 4, day 6, day 8 and fully undifferentiated) to the estimated Δ c-score due to deficiency of H1C and H1E in those same regions in GC B cells. (day 2, $P < 2.2 \times 10^{-16}$; day 4, $P < 2.2 \times 10^{-16}$; day 6, $P = 0.002$; day 8, 0.0001; estimated with two-sided Wilcoxon test). Box plot centre represents median, bounds of box are the first and third quartiles and whiskers extend to $1.5 \times$ the interquartile range. **e**, v4c analyses on the *Klf5* locus (chr14: 99,000,000–100,200,000) anchored on the *Klf5* promoter for (top) B cell reprogramming (blue) to iPS cells (red) with four intermediate time states (grey) from a previous study[32], as well as GC B cells (bottom) ($H1c^{-/-}H1e^{-/-}$ and wild-type H1). IgV tracks below comprise Δ c-score and ATAC–seq signal in $H1c^{-/-}H1e^{-/-}$ versus wild-type GC B cells. Gained HiC interactions in $H1c^{-/-}H1e^{-/-}$ compared to wild-type GC B cells (shaded in grey 1–3: pval = 0.04; pval = 0.059; pval = 0.02, respectively, two-sided unpaired t-test) have OCT2 motif sequences as shown. **f**, Schematic of experimental set-up with $H1c^{-/-}H1e^{-/-}$ or wild-type littermate mouse embryonic fibroblasts. **g**, Representative images of alkaline phosphatase (AP)-stained $H1c^{-/-}H1e^{-/-}$ and wild-type iPS cell colonies at day 21. **h**, Per cent iPS cell reprogramming efficiency of $H1c^{-/-}H1e^{-/-}$ ($n = 5$ transfections on two biological replicates) and wild-type ($n = 4$ transfections on two biological replicates) mouse embryonic fibroblasts determined as the ratio of AP+ colonies to the number of seeded mCherry+ cells. $P = 0.01$, two-sided unpaired t-tests. Data are mean ± s.d. Data are representative of three independent experiments.
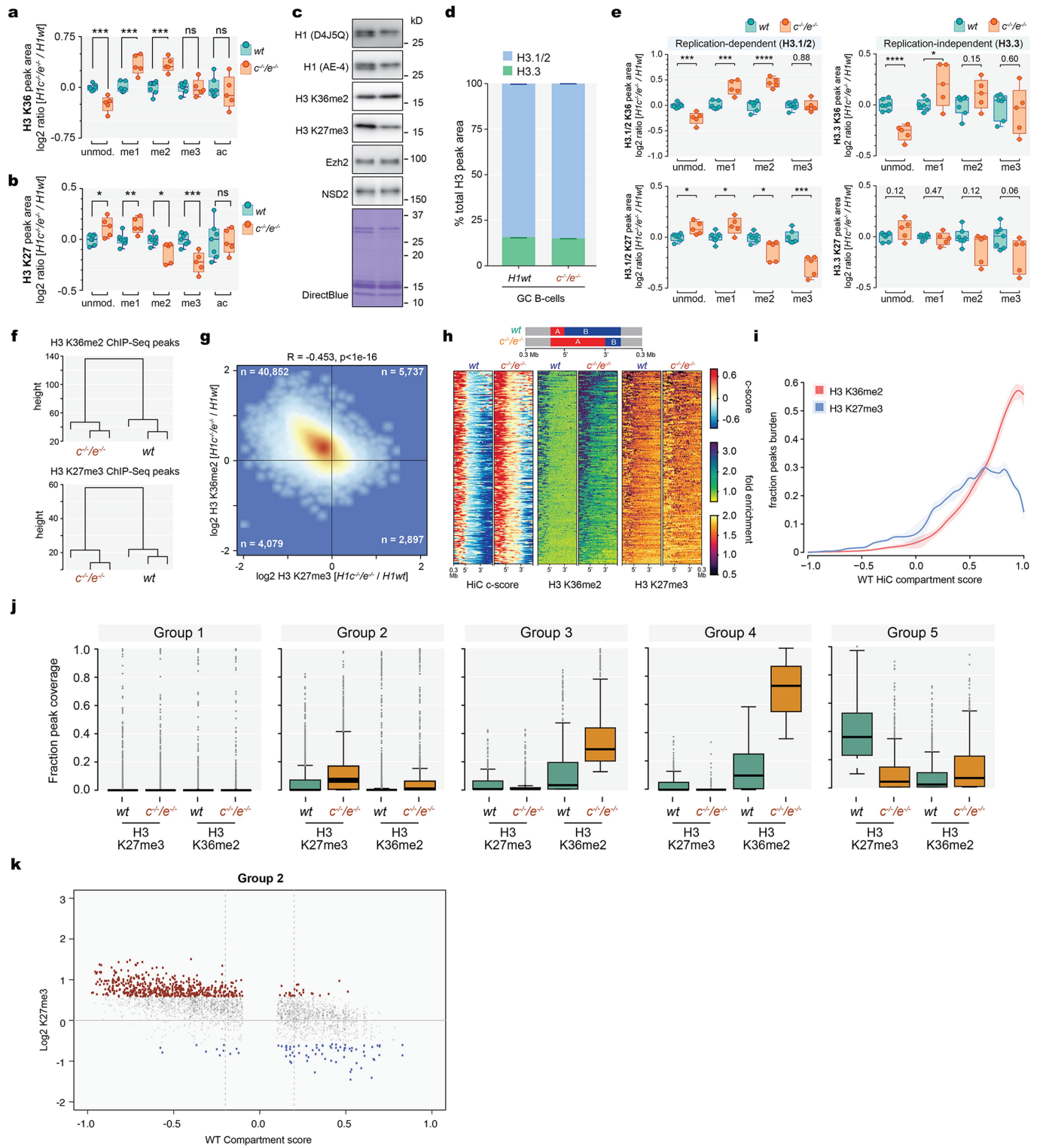
Extended Data Fig. 7 | See next page for caption.

Extended Data Fig. 7 | Altered H3K36me2 and H3K27me3 distribution in *H1c*⁻/⁻*H1e*⁻/⁻ **GC B cells. a, b,** Mass spectrometry of H3K36 (**a**) and H3K27 (**b**) post-translational modifications, $\log_2$-transformed and normalized to average wild-type peak area from *H1c*⁻/⁻*H1e*⁻/⁻ ($n = 5$) and wild-type ($n = 7$) acid-extracted samples from GC B cells; two-sided unpaired t-test: K36 unmodified (unmod.), ***$P = 0.0005$; K36me1, ***$P = 0.0003$; K36me2 ***$P = 0.0002$; K36me3, $P = 0.93$; K36ac, $P = 0.56$; K27 unmod., *$P = 0.0157$; K27me1, **$P = 0.0072$; K27me2, *$P = 0.0175$; K27me3, ***$P = 0.0007$; K27ac, $P = 0.9337$. Box plots show median and 25th to 75th percentiles, whiskers indicate data range. Data are representative of two independent experiments. **c,** Immunoblots for H1 (D4J5Q and AE-4 antibodies), H3K36me2, H3K27me3, EZH2 and NSD2 from sorted wild-type and *H1c*⁻/⁻*H1e*⁻/⁻ GC B cells. Direct blue stain is included as loading control. A representative image of at least three experiments is shown. Uncropped gels are shown in Supplementary Fig. 1. **d,** Mass-spectometry-based relative abundance of H3.1 or H3.2 (replication-dependent) and H3.3 (replication-independent) isoforms, shown as the average percentage of the total peak area of H3K27-K36 peptide containing H3.3-specific Ser31 in acid-extracted histones from wild-type (H3.3, 15.76%, $n = 7$) and *H1c*⁻/⁻*H1e*⁻/⁻ (H3.3, 15.07%, $n = 5$) GC B cells; two-sided unpaired t-test, $P = 0.0004$, Data are mean ± s.d. **e,** Mass spectrometry of H3 K36 (top) and K27 (bottom) post-translational modifications across H3.1/2 (left) and H3.3 (right) isoforms, $\log_2$-transformed and normalized to average wild-type peak area from samples acid-extracted from wild type ($n = 7$) and *H1c*⁻/⁻*H1e*⁻/⁻ ($n = 5$) GC B cells; two-sided unpaired t-test: H3.1/2 K36 unmod., ***$P = 0.0005$; H3.1/2 K36me1, ***$P = 0.0003$; H3.1/2 K36me2, ****$P < 0.0001$; H3.1/2 K36me2, $P = 0.88$; H3.1/2 K27 unmod., *$P = 0.0100$; H3.1/2 K27me1, *$P = 0.0162$; H3.1/2 K27me2, *$P = 0.0129$; H3.1/2 K27me3, ***$P = 0.0002$; H3.3 K36 unmod., ****$P < 0.0001$; H3.3 K36me1, *$P = 0.036$; H3.3 K36me2, $P = 0.15$; H3.3 K36me3, $P = 0.5974$; H3.3

K27 unmod., $P = 0.1187$; H3.3 K27me1, $P = 0.4743$; H3.3 K27me2, $P = 0.1199$; H3.3 K27me3, $P = 0.0628$. Box plots show median and 25th to 75th percentiles, whiskers indicate data range. **f,** Unsupervised hierarchical clustering analysis of ChIP–seq data for H3K27me3 and H36me2 in biological triplicates from sorted *H1c*⁻/⁻*H1e*⁻/⁻ and wild-type GC B cells. **g,** Genome-wide correlation plot of $\log_2$-transformed fold change (*H1c*⁻/⁻*H1e*⁻/⁻ versus wild type) of normalized reads within ChIP–seq peak union for H3K36me2 and H3K27me3 (Pearson correlation coefficient $R = -0.453$, $P < 1 \times 10^{-16}$). **h,** Heat map of HiC c-score, H3K36me2 and H3K27me3 centred within shifting B-to-A compartments (100 kb) and surrounding 300 kb for *H1c*⁻/⁻*H1e*⁻/⁻ and wild-type GC B cells for compartment 'extensions' (top). **i,** Fraction of ChIP–seq peak coverage (H3K36me2 in red and H3K27me3 in blue) within 100-kb compartments across HiC c-score (x axis, −1 to 1) for wild-type GC B cells. Cubic smoothing splines of data are presented with shaded regions indicating 99% confidence intervals. **j,** Fraction peak (H3K27me3 and H3K36me2) coverage of regions within shifting compartment groups 1–5 in *H1c*⁻/⁻*H1e*⁻/⁻ and wild-type GC B cells. Paired Wilcoxon test, group 1: H3K27me3, $P < 1 \times 10^{-16}$; H3K36me2, $P = 0.597$; group 2: H3K27me3, $P < 1 \times 10^{-16}$; H3K36me2, $P < 1 \times 10^{-16}$; group 3: H3K27me3, $P = 3.38 \times 10^{-13}$; H3K36me2, $P < 1 \times 10^{-16}$; group 4: H3K27me3, $P = 8.71 \times 10^{-15}$; H3K36me2, $P < 1 \times 10^{-16}$; group 5: H3K27me3, $P < 1 \times 10^{-16}$; H3K36me2, $P < 1 \times 10^{-16}$. Box plot centre represents median, bounds of box are the first and third quartiles and whiskers extend to 1.5 × the interquartile range. **k,** Scatter plot of H3K27me3 peak $\log_2$-transformed fold change (*H1c*⁻/⁻*H1e*⁻/⁻ versus wild-type GC B cells) versus wild-type c-score score for decompacting group 2. Gain of H3K27me3 (red dots) largely occurred within regions shifting from compartment B whereas loss of H3K72me3 (blue dots) was more prevalent within regions shifting compartments from compartment A.
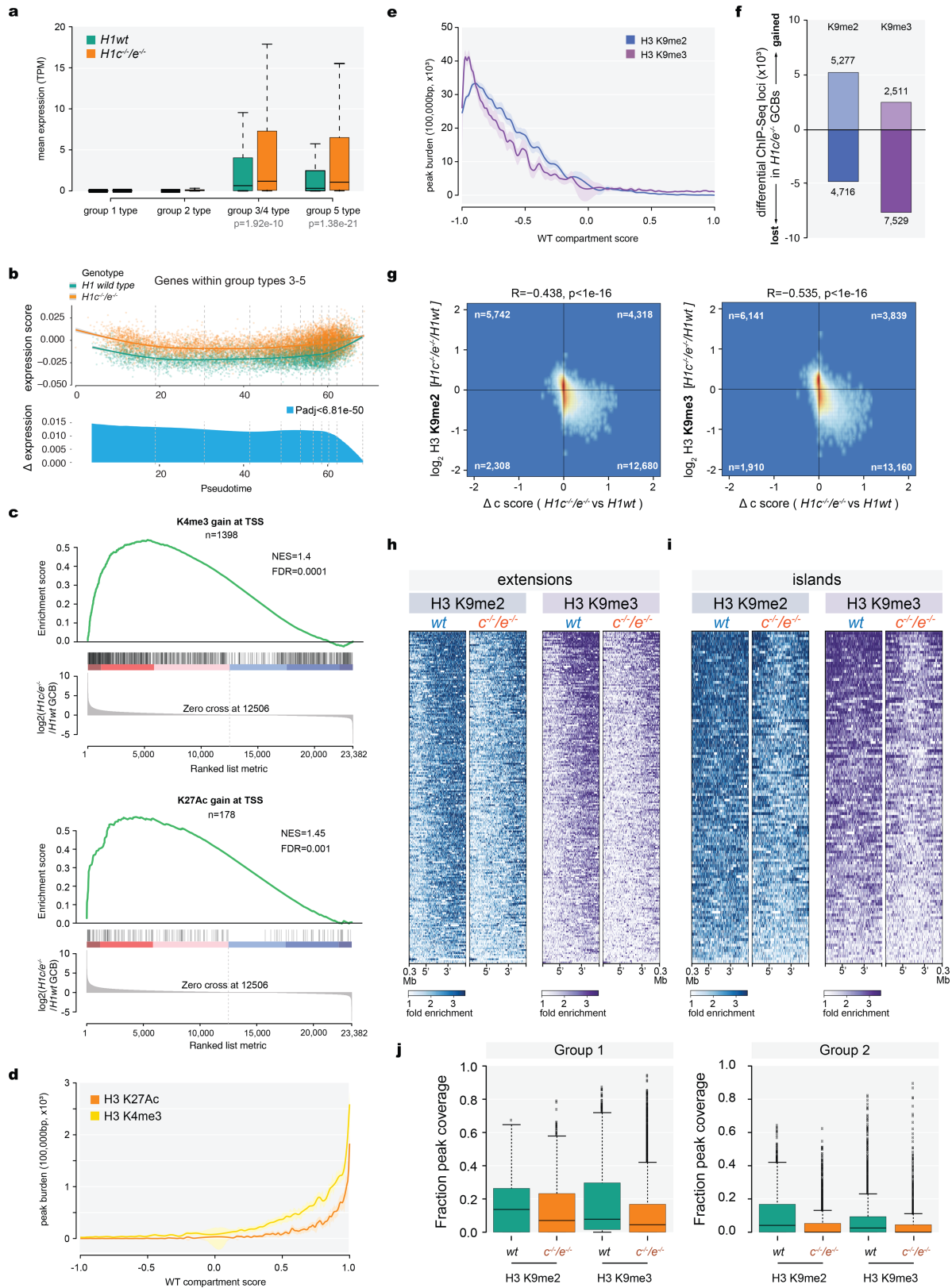
Extended Data Fig. 8 | See next page for caption.

Extended Data Fig. 8 | Changes in activation marks H3K4me3 and H3K27ac and repressive marks H3K9me2 and H3K9me3 in H1-deficient GC B cells are associated with compartment-B decompaction. a, Genes defined as groups 3 and 4 ($n = 108$ genes) and group 5 ($n = 152$ genes) show transcriptional activation and significant upregulation in $H1c^{-/-}H1e^{-/-}$ GC B cells compared to wild-type GC B cells. Paired Wilcoxon test. Box plot centre represents median, bounds of box are the first and third quartiles and whiskers extend to $1.5 \times$ the interquartile range. **b**, Top, expression of the genes defined as groups 3–5 was plotted for each cell on the y axis. Average expression is represented by the different coloured spline curves for each genotype as indicated. Bottom, differential expression between $H1c^{-/-}H1e^{-/-}$ and wild-type cells is represented as the delta spline plot across pseudotime. Cells are divided by pseudotime into bins of equal cell number (grey vertical dashed lines) and tested for signature enrichment compared to wild-type GC B cells by two-sided Wilcoxon test ($P < 6.81 \times 10^{-50}$) **c**, GSEA with genes marked with gain in H3K4me3 (top) or H3K27Ac (bottom) on their promoters (transcription start site (TSS) ± 500 bp) using the ranked $\log_2$-transformed fold change in mouse $H1c^{-/-}H1e^{-/-}$ GC B cells. **d**, Fraction of histone peak coverage for H3K4me3 and H3K27ac within 100-kb compartments across HiC c-score (x axis) for wild-type GC B cells. Cubic smoothing splines of data are presented with shaded regions indicating 99% confidence intervals. **e**, Fraction of histone peak coverage (for H3K9me2 and

H3K9me3) within 100-kb compartments across HiC c-score (x axis) for wild-type GC B cells. Cubic smoothing splines of data are presented with shaded regions indicating 99% confidence intervals. **f**, CUT&RUN peaks for H3K9me2 and H3K9me3 show altered abundance between $H1c^{-/-}H1e^{-/-}$ and wild-type GC B cells (FC > 1.5, $n = 5,277$ gained and $n = 4,717$ lost H3K9me2 and $n = 2,511$ gained and $n = 7,529$ lost H3K9me3 peaks). **g**, Genome-wide correlation plot of $\log_2$-transformed fold change ($H1c^{-/-}H1e^{-/-}$ versus wild-type H1 GC B cells) in H3K9me2 (left) and H3K9me3 (right) peaks versus Δ c-score ($H1c^{-/-}H1e^{-/-}$ minus wild-type H1). Both H3K9me2 and H3K9me3 changes were largely loss and were anti-correlated with compartment decompaction (Pearson correlation coefficient $R = -0.438$, $P < 1 \times 10^{-16}$ and $R = -0.543$, $P < 1 \times 10^{-16}$, respectively). **h**, **i**, Heat maps of H39me2 and H3k9me3 centred within shifting B-to-A compartments (100 kb) and surrounding 300 kb for $H1c^{-/-}H1e^{-/-}$ and wild-type GC B cells for compartment 'extensions' (**h**) and 'islands' (**i**). **j**, Fraction peak (H3K9me2 and H3K9me3) coverage of regions within shifting compartment groups and 2 in $H1c^{-/-}H1e^{-/-}$ and wild-type GC B cells. Paired Wilcoxon test, group 1: H3K9me2, $P < 1 \times 10^{-16}$, H3K9me3, $P < 1 \times 10^{-16}$; group 2: H3K9me2, $P < 1 \times 10^{-16}$; H3K9me3, $P < 1 \times 10^{-16}$. Box plot centre represents median, bounds of box are the first and third quartiles and whiskers extend to $1.5 \times$ the interquartile range.
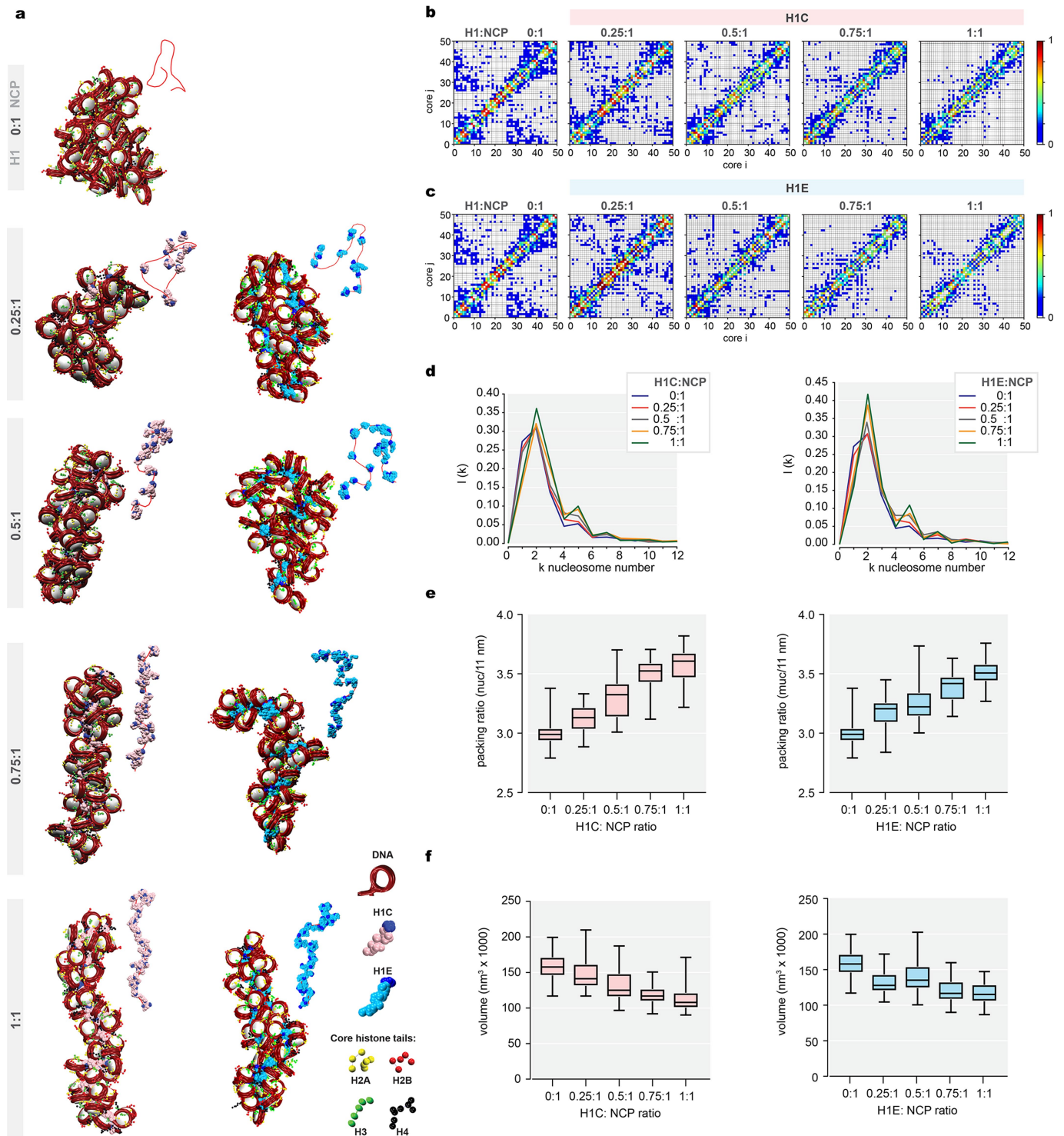
Extended Data Fig. 9 | Linker histone incorporation reduces interactivity of chromatin fibre. a, Representative equilibrated configurations of 50-nucleosome chromatin fibres obtained in silico in the absence of H1 and in the presence of 0.25, 0.5, 0.75 and 1 H1 molecule per nucleosome (H1C (left) and H1E (right)). Fibre contour (in red) on which the H1 mean positions are shown is shown on the top right of each model. Colour key shows DNA, linker histone, H2A, H2B, H3 and H4 tails. **b**, **c**, Contact maps for the 1,000-configuration ensembles obtained from left to right, in the absence of H1 and in the presence of 0.25, 0.50, 0.75 and 1 H1 per nucleosome (**b**, H1C; **c**, H1E). **d**, The nucleosome–nucleosome interaction patterns, or a one-dimensional decomposition of each contact map shown in **b** and **c**, indicate the dominant zigzag pattern of the fibre ($i \pm 2$) and increase of long-range interactions as the density of H1C (left) or H1E

(right) decreases. **e**, Packing ratio calculated as the number of nucleosomes contained in 11 nm of fibre for systems without H1 and in the presence of 0.25, 0.50, 0.75 and 1 H1 molecule per nucleosome (left, H1C; right, H1E); $n = 1,000$ chromatin ensemble configurations for each H1 per nucleosome ratio. Ordinary one-way analysis of variance for multiple comparisons, ****$P < 0.0001$. **f**, Volume of chromatin fibres calculated assuming a cylindrical shape for systems with no H1 and in the presence of 0.25, 0.5, 0.75 and 1 H1 molecule per nucleosome (left, H1C; right, H1E); $n = 1,000$ chromatin ensemble configurations for each H1 per nucleosome ratio. Ordinary one-way analysis of variance for multiple comparisons, ****$P < 0.0001$. Box plot centre represents median, bounds of box are the first and third quartiles and whiskers extend to $1.5 \times$ the interquartile range (**e**, **f**).

a

VavP-Bcl2

H1 wild type | *H1c⁻ᐟ⁺/e⁻ᐟ⁺* | *H1c⁻ᐟ⁻/e⁻ᐟ⁻*

Lymph node

H&E

B220

b

VavP-Bcl2

H3 K36me2

H1 wild type | *H1c⁺ᐟ⁻/e⁺ᐟ⁻* | *H1c⁻ᐟ⁻/e⁻ᐟ⁻*

H3K36me2

High / Mid / Low / Neg

% lymph node tumor cells

c

VavP-Bcl2

H1 wild type | *H1c⁺ᐟ⁻/e⁺ᐟ⁻* | *H1c⁻ᐟ⁻/e⁻ᐟ⁻*

lung

H&E

B220

Ki67

d

CD3

H1c⁺ᐟ⁻/e⁺ᐟ⁻
H1c⁻ᐟ⁻/e⁻ᐟ⁻

VavP-Bcl2

e

Igλ clonal rearrangement

H1 wild type — *c⁺ᐟ⁻/e⁺ᐟ⁻* — *c⁻ᐟ⁻/e⁻ᐟ⁻*
–*Bcl2ʷᵗ*– — *VavP-Bcl2*

f

H1 wild type | *H1c⁺ᐟ⁻/e⁺ᐟ⁻* | *H1c⁻ᐟ⁻/e⁻ᐟ⁻*

liver

H&E

B220

lung

H&E

B220

g

VavP-Bcl2

H1 wild type | *H1c⁺ᐟ⁻/e⁺ᐟ⁻*

lymph node tumors

H&E B220 H&E B220

h

Murine tumor *VavP-Bcl2, H1c⁻ᐟ⁻/e⁻ᐟ⁻* UPREG
n=1054

NES = 2.17
FDR < 0.001

Enrichment score (ES)

log2(*H1c⁻ᐟ⁻/H1wt* GCB)

Zero cross at 10577

Ranked list metric

i

Murine tumor *VavP-Bcl2, H1c⁺ᐟ⁻/e⁺ᐟ⁻* UPREG
n=240

NES=1.39
FDR<0.001

Enrichment score

log2(*H1c⁻ᐟ⁻/H1wt* GCB)

Zero cross at 12506

Ranked list metric

Murine tumor *VavP-Bcl2, H1c⁻ᐟ⁻/e⁻ᐟ⁻* UPREG
n=1054

NES=1.64
FDR<0.001

Enrichment score

log2(*H1c⁻ᐟ⁻/H1wt* GCB)

Zero cross at 12506

Ranked list metric

j

leading edge (n=898)

Klf5
Meis1
Tet1

Gja1
Gspt2
Prdm5

H1c⁺ᐟ⁻/e⁺ᐟ⁻ *H1c⁻ᐟ⁻/e⁻ᐟ⁻*

VavP-Bcl2

log2 ratio vs mean WT

−4 −2 0 2 4

k

BCCA DLBCL H1 doublemut UPREG
murine orthologs
n=78

NES=1.31
FDR=0.0436

Enrichment score (ES)

log2(*H1c⁻ᐟ⁻/H1wt* GCB)

Zero cross at 9063

Ranked list metric

BCCA DLBCL H1 doublemut UPREG
murine orthologs
n=78

NES=1.05
FDR=0.403

log2(*H1c⁻ᐟ⁻/H1wt* GCB)

Zero cross at 9063

Ranked list metric

l

wild type

H3 K27me3
H3 K36me2

B   A

global compartment shift towards A

Hist1h1 loss

H3 K27me3
H3 K36me2

derepression of developmentally regulated genes → clonal expansion of H1-deficient GC B cells → lymphomagenesis

Extended Data Fig. 10 | See next page for caption.

Article

Extended Data Fig. 10 | Loss of *H1c* and *H1e* leads to aggressive *Vav-PBcl2* lymphomas with DLBCL-like morphology. a, Immunohistochemistry images of lymph node stained for H&E and B220 from *VavP-Bcl2;H1c$^{-/-}$H1e$^{-/-}$*, *VavP-Bcl2;H1c$^{-/+}$H1e$^{-/+}$* and *VavP-Bcl2*-only mice at day 164. Scale bar, 1 mm. Images are representative of *n* = 11 mice per genotype examined over two independent experiments. **b**, Representative immunohistochemistry images of lymphomatous *VavP-Bcl2;H1c$^{-/-}$H1e$^{-/-}$*, *VavP-Bcl2;H1c$^{-/+}$H1e$^{-/+}$* and *VavP-Bcl2*-only lymph nodes stained for H3K36me2, and quantification of intensity (binned as high, mid, low and negative). Scale bar, 50 μm. Tissue derived from three mice per genotype with four tumour lymph nodes each. Data are mean ± s.d., two-sided unpaired *t*-test. **c**, Immunohistochemistry images of lung tissue stained for H&E, B220 and Ki67 from *VavP-Bcl2;H1c$^{-/-}$H1e$^{-/-}$*, *VavP-Bcl2;H1c$^{-/+}$H1e$^{-/+}$* and *VavP-Bcl2*-only mice at day 164. Scale bar, 100 μm. Images are representative of *n* = 11 mice per genotype examined over two independent experiments. **d**, Quantification of B220$^+$ lesion areas in liver tissue (Fig. 5c) from *VavP-Bcl2;H1c$^{-/-}$H1e$^{-/-}$* (****P < 0.0001) and *VavP-Bcl2;H1c$^{-/+}$H1e$^{-/+}$* (*P = 0.0308) compared to *VavP-Bcl2*-only (*n* = 7 mice per genotype, mean ± s.d.; two-sided unpaired *t*-tests). **e**, Immunohistochemistry stains for CD3 from *VavP-Bcl2;H1c$^{-/-}$H1e$^{-/-}$* and *VavP-Bcl2;H1c$^{-/+}$H1e$^{-/+}$* lymphomas. Scale bar, 50 μm. Images are representative of *n* = 4 mice per genotype examined over two independent experiments. **f**, PCR for Igλ clonal rearrangement to report on tumour clonality of B220$^+$ cells from *VavP-Bcl2;H1c$^{-/-}$H1e$^{-/-}$*, *VavP-Bcl2;H1c$^{-/+}$H1e$^{-/+}$* and *VavP-Bcl2*-only mice at day 164. **g**, Immunohistochemistry images of liver and lung tissues stained for H&E and B220 in from *H1c$^{-/-}$H1e$^{-/-}$*, *H1c$^{-/+}$H1e$^{-/+}$* and wild-type mice at day 164. Scale bar, 500 μm. Images are representative of *n* = 6 mice per genotype examined over two independent experiments. **h**, Immunohistochemistry images of lymph node tissue stained for H&E and B220 from *VavP-Bcl2;H1c$^{-/+}$H1e$^{-/+}$* and *VavP-Bcl2*-only mice. Images are representative of *n* = 4 mice per genotype examined over two independent experiments. Scale bars, 100 μm. **i**, GSEA with the *VavP-Bcl2;H1c$^{-/-}$H1e$^{-/-}$* versus *VavP-Bcl2* lymphoma upregulated gene set ranked against log$_2$-transformed changes in fold change from mouse *VavP-Bcl2;H1c$^{-/+}$H1e$^{-/+}$* versus *VavP-Bcl2*. **j**, Top, GSEA for genes upregulated in *VavP-Bcl2;H1c$^{-/+}$H1e$^{-/+}$* versus *VavP-Bcl2* lymphomas using the ranked log$_2$-transformed fold change in mouse *H1c$^{-/-}$H1e$^{-/-}$* GC B cells. Bottom, GSEA for genes upregulated in *VavP-Bcl2;H1c$^{-/-}$H1e$^{-/-}$* versus *VavP-Bcl2* mice using the ranked log$_2$-transformed fold change in mouse *H1c$^{-/-}$H1e$^{-/-}$* GC B cells. **k**, Heat map showing differential expression of leading-edge genes (*n* = 898) from *VavP-Bcl2;H1c$^{-/-}$H1e$^{-/-}$* and *VavP-Bcl2;H1c$^{-/+}$H1e$^{-/+}$* lymphomas. **l**, GSEA for genes upregulated in human *H1C/E*-mutant DLBCL, against the ranked log$_2$-transformed fold change gene expression profiles of mouse *VavP-Bcl2;H1c$^{-/+}$H1e$^{-/+}$* (left) and *VavP-Bcl2;H1c$^{-/-}$H1e$^{-/-}$* (right) lymphomas. **m**, Summary model depicting chromatin as contiguous B-to-A space, with H3K27 and H3K36 methylations occupying distinct compartments within. Loss of H1 results in a global shift of compartment interactivity towards A, with both H3K27 and H3K36 methylations shifting into ectopic regions. Most compacted regions devoid of either K27 or K36 methylation appear largely protected from H1 loss. The biological effects of H1 loss in GC B cells are summarized below.

nature research

Corresponding author(s): Ethel Cesarman and Ari M. Melnick

Last updated by author(s): Sep 3, 2020

Reporting Summary

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

Statistics

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

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

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	FACS BD Aria II, Influx, FACS Canto II were used to collect flow cytometry sorting and analysis data. Halo software (v3.0.311.201) was used to quantify data from immunohistochemistry. Skyline software (v4.1) was used for raw mass spectometry files. Asylum Research sotware packager version IX was used for height profiles on atomic force microscopy. ELISPOT plates were evaluated using an automated Zeiss ELISPOT reader system (ZellNet Consulting, Inc.). Biolayer interferometry measurements were taken on an Octen Red96e system (v11).
Data analysis	FlowJo software v10.5.3, R version 3.4.4 (2017-11-30), R package Slingshot version 1.2.0, R package Seurat 3.0.2, R package EdgeR v3.16.5 and v3.20.9, R package EDAseq 2.22.0, R package DeSeq2 1.27.1, Sarek pipeline version 2.5.1 ChIP-seq reads were aligned using bwa-mem function of the BWA suite. Peaks were called using the SICER algorithm. RNA-seq reads were aligned to mm10 using STAR and annotated to RefSeq using the R subread package. Whole-genome sequencing data aligned to human genome hg-19 was retrieved from the European Genome Archive and used to generate unmapped bam files using picard tools RevertSam according to GATK best practice. Bam files were aligned to human genome hg38 using BWA mem. Somatic alterations were called from matched tumor-normal pairs using Manta followed by Strelka2, and annotated by VEP (https://github.com/nf-core/sarek). FishHook (https://github.com/mskilab/fishHook) was used to model background mutational processes. All study and public Hi-C data used have been pre-processed with the hic-bench pipeline. FRAP data were analyzed using ImageJ (v1.51j) and FRAP Profiler plugin (http://worms.zoology.wisc.edu/research/4d/4d.html)

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

Data

The sequencing datasets generated in this study are available at GEO (GSE143293). No restriction of data availability.

Field-specific reporting

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

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

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

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

Sample size	For high throughput NG, mass spec experiments, a minimum of three independent biological replicates were used where possible (ChIP-seq, ATAC-seq, RNA-seq, Cut and Run). Two biological replicates were used for HiC and Single cell RNA-seq. Sample size for in vivo experiments were chosen based on our prior experience for detecting statistically significant differences compared to control animals (Béguelin et al Cancer Cell, Vol37, Issue 5, Pages 655-673). Sample size for in vitro experiments were determined with preliminary experiments with a minumum of 2 samples per group, and full cohort experiments repeated at least twice.
Data exclusions	No samples were excluded.
Replication	All findings were reproducible. Confirmation of reproducibility was assessed with at least three independent experiments for in vivo and at least two independent experiments for in vitro assays, and in the case of high throughput NGS analysis, through the use of sample sized that provided comprehensive genome-wide coverage of chromatin changes.
Randomization	Sex-matched littermates were randomly assigned to experimental arms. Samples were allocated to groups according to genotype.
Blinding	Studies were not blinded and group allocation was done before data collection as only one investigator ran the experiments.

Reporting for specific materials, systems and methods

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

Materials & experimental systems

n/a	Involved in the study
☐	☒ Antibodies
☐	☒ Eukaryotic cell lines
☒	☐ Palaeontology
☐	☒ Animals and other organisms
☒	☐ Human research participants
☒	☐ Clinical data

Methods

n/a	Involved in the study
☐	☒ ChIP-seq
☐	☒ Flow cytometry
☒	☐ MRI-based neuroimaging

Antibodies

Antibodies used	For immunoblotting, the following primary antibodies were used: D4J5Q, anti-H1E, Cell-Signaling 41328; AE-4, anti-H1, Millipore 05-457; C36B11 anti-H3 K27me3, Cell Signaling 9733; C75H12 anti-H3 K36me2, Cell Signaling 2901; AC22 anti-Ezh2, Cell Signaling 3147; 29D1 anti-NSD2, Millipore Mabe191; rabbit polyclonal anti-KLF5, Proteintech 21017-1-AP; 6C5 anti-GAPDH, Ab8245) For ChIP, the following primary antibodies were used: anti-H3 K27me3, C36B11 - Cell Signaling 9733 (20 micrograms), anti-H3 K36me2, C75H12 - Cell Signaling 2901 (20 micrograms). For CutandRun, the following antibodies were used: ant-H3K9me2 (Millipore 05-1249), anti-H3K9me3 (Abcam 176916), and anti-H3K4me3 (Epicypher 13-0041) (0.5ug each).
Validation	anti-H3 K27me3 and anti-H3 K36me2 antibodies have been extensively validated in literature, including Weinberg et al., 2019

Nature; Papillon-Cavanagh et al., Nat Genet 2017; Lu et al., Science 2016. Cut and Run antibodies are validated to SNAP-ChIP nucleosome standards (S1hah et al., Mol Cell 2018).D4J5Q and AE-4 anti-H1 antibodies have been validated in mouse knock-out and human shRNA-treated knock-down cells. C36B11, C75H12, Ac22 and 29D1 have been validated in human and mouse knock-out and transgenic cell lines (using ChIP-qPCR in wild type and loss-of-function (K27M, K36M, knock-ou, Ezh2 knock-out) backgrounds.

Eukaryotic cell lines

Policy information about cell lines

Cell line source(s)	NIH-3T3 (CRL-1658) purchased from ATCC; Kc167 (RRID:CVCL_Z834) cells were purchased from Drosophila Genomics Resource Center
Authentication	Cell lines were not specifically authenticated for these studies
Mycoplasma contamination	Cell line was tested negative for mycoplasma contamination
Commonly misidentified lines (See ICLAC register)	No commonly misidentified lines were used in this study

Animals and other organisms

Policy information about studies involving animals; ARRIVE guidelines recommended for reporting animal research

Laboratory animals	The study involved H1c-/-/e-/- mice and littermate controls (C57BL/6J) that were age- and sex-matched (tested within 8-10-weeks old). Bone-marrow recipient mice were female, 8-weeks old and C57BL/6J strain with donor bone marrow (CD45.1 and CD45.2) pooled from sex-matched, 8-10-week old mice. VavP-Bcl2 (C57BL/6J), crossed with H1c-/-/e-/-, were used sex-matched 8-week old mice as bone-marrow donors for lymphomagenesis and survival studies.
Wild animals	The study did not involve wild animals.
Field-collected samples	The study did not involve samples collected from the field.
Ethics oversight	Animal care was in strict compliance with institutional guidelines established by Weill Cornell Medicine, the Guide for the Care and Use of Laboratory Animals (National Academy of Sciences 1996) (Silverman et al., 2006), and the Association for Assessment and Accreditation of Laboratory Animal Care International. All mice were followed until any one of several criteria for euthanizing were met, including severe lethargy, more than 10% body weight loss, and palpable splenomegaly that extended across the midline, in accordance with our Weill Cornell Medicine Institutional Animal Care and Use Committee (IACUC)–approved animal protocols.

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

ChIP-seq

Data deposition

☒ Confirm that both raw and final processed data have been deposited in a public database such as GEO.

☒ Confirm that you have deposited or provided access to graph files (e.g. BED files) for the called peaks.

Data access links *May remain private before publication.*	https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE143293 Publicly available
Files in database submission	GSM4256478 Sample_1_GCB_WT_repl1_RNAseq GSM4256479 Sample_2_GCB_WT_repl2_RNAseq GSM4256480 Sample_3_GCB_WT_repl3_RNAseq GSM4256481 Sample_4_GCB_WT_repl4_RNAseq GSM4256482 Sample_5_GCB_H1_DKO_repl1_RNAseq GSM4256483 Sample_6_GCB_H1_DKO_repl2_RNAseq GSM4256484 Sample_7_GCB_H1_DKO_repl3_RNAseq GSM4256485 Sample_8_GCB_WT_H3K27me3_repl1_ChIPseq GSM4256486 Sample_9_GCB_WT_H3K27me3_repl2_ChIPseq GSM4256487 Sample_10_GCB_WT_H3K27me3_repl3_ChIPseq GSM4256488 Sample_11_GCB_H1 DKO_H3K27me3_repl1_ChIPseq GSM4256489 Sample_12_GCB_H1 DKO_H3K27me3_repl2_ChIPseq GSM4256490 Sample_13_GCB_H1 DKO_H3K27me3_repl3_ChIPseq GSM4256491 Sample_14_GCB_WT_H3K36me2_repl1_ChIPseq GSM4256492 Sample_15_GCB_WT_H3K36me2_repl2_ChIPseq GSM4256493 Sample_16_GCB_WT_H3K36me2_repl3_ChIPseq GSM4256494 Sample_17_GCB_H1 DKO_H3K36me2_repl1_ChIPseq GSM4256495 Sample_18_GCB_H1 DKO_H3K36me2_repl2_ChIPseq GSM4256496 Sample_19_GCB_H1 DKO_H3K36me2_repl3_ChIPseq GSM4256497 Sample_20_GCB_WT_Input_repl1_ChIPseq

GSM4256498 Sample_21_GCB_WT_Input_repl2_ChIPseq
GSM4256499 Sample_22_GCB_WT_Input_repl3_ChIPseq
GSM4256500 Sample_23_GCB_H1 DKO_Input_repl1_ChIPseq
GSM4256501 Sample_24_GCB_H1 DKO_Input_repl2_ChIPseq
GSM4256502 Sample_25_GCB_H1 DKO_Input_repl3_ChIPseq
GSM4363364 Sample_26_GCB_WT_repl1_HiC
GSM4363365 Sample_27_GCB_WT_repl2_HiC
GSM4363366 Sample_28_GCB_H1_DKO_repl1_HiC
GSM4363367 Sample_29_GCB_H1_DKO_repl2_HiC
GSM4523349 Sample_30_GCB_WT_repl1_ATACseq
GSM4523350 Sample_31_GCB_WT_repl2_ATACseq
GSM4523351 Sample_32_GCB_WT_repl3_ATACseq
GSM4523352 Sample_33_GCB_WT_repl4_ATACseq
GSM4523353 Sample_34_GCB_H1_DKO_repl1_ATACseq
GSM4523354 Sample_35_GCB_H1_DKO_repl2_ATACseq
GSM4523355 Sample_36_GCB_H1_DKO_repl3_ATACseq
GSM4523356 Sample_37_GCB_H1_DKO_repl4_ATACseq
GSM4523357 Sample_38_GCB_WT_H3K4me3_CutAndRun
GSM4523358 Sample_39_GCB_WT_H3K9me2_CutAndRun
GSM4523359 Sample_40_GCB_WT_H3K9me3_CutAndRun
GSM4523360 Sample_41_GCB_H1_DKO_H3K4me3_CutAndRun
GSM4523361 Sample_42_GCB_H1_DKO_H3K9me2_CutAndRun
GSM4523362 Sample_43_GCB_H1_DKO_H3K9me3_CutAndRun
GSM4523363 Sample_44_GCB_WT_H3K27Ac_rep1_ChIPseq
GSM4523364 Sample_45_GCB_WT_H3K27Ac_rep2_ChIPseq
GSM4523365 Sample_46_GCB_WT_H3K27Ac_rep3_ChIPseq
GSM4523366 Sample_47_GCB_H1_DKO_H3K27Ac_rep1_ChIPseq
GSM4523367 Sample_48_GCB_H1_DKO_H3K27Ac_rep2_ChIPseq
GSM4523368 Sample_49_GCB_H1_DKO_H3K27Ac_rep3_ChIPseq
GSM4752544 Sample_DKO2-GCB-3_10X_SingleCell
GSM4752545 Sample_WT4-GCB-1_10X_SingleCell
GSM4752546 Sample_WT5-GCB-2_10X_SingleCell
GSM4752547 Sample_DKO6-GCB-4_10X_SingleCell

Genome browser session
(e.g. UCSC)

Not applicable - visualized data using IGV

Methodology

Replicates

For ChIP-seq experiments with primary mouse GC B-cells, three biological replicates (n=3)per group were tested. For CutandRun experiments with primary mouse GC B-cells, two biological replicates (n=2) per group were tested. For HiC experiments with primary mouse GC B-cells, two biological replicates per group were tested. For ATAC-seq with primary mouse GCB-cells, at least three biological replicates (n=3) per group were tested. For Single Cell RNA-seq with primary mouse GC B-cells, two biological replicates per group were tested.

Sequencing depth

ChIP-seq experiments: 75 bp single-end
1-WT-H3K27me3, total number reads: mm10, 18,627,064; dm6, 23,345,945
2-WT-H3K27me3, total number reads: mm10, 15,809,793; dm6, 21,688,452
3-WT-H3K27me3, total number reads: mm10, 18,663,073; dm6, 21,202,327
1-H1DKO-H3K27me3, total number reads: mm10, 20,042,774; dm6, 26,505,971
2-H1DKO-H3K27me3, total number reads: mm10, 17,020,424; dm6, 25,308,467
3-H1DKO-H3K27me3, total number reads: mm10, 16,444,400; dm6, 23,322,436
1-WT-H3K36me2, total number reads: mm10, 19,010,885; dm6, 2,397,743
2-WT-H3K36me2, total number reads: mm10, 18,804,732; dm6, 2,275,684
3-WT-H3K36me2, total number reads: mm10, 16,474,978; dm6, 1,760,707
1-H1DKO-H3K36me2, total number reads: mm10, 27,599,022; dm6, 2,585,188
2-H1DKO-H3K36me2, total number reads: mm10, 22,605,400; dm6, 2,310,040
3-H1DKO-H3K26me2, total number reads: mm10, 28,325,523; dm6, 2,900,788
1-WT-H3K27Ac, total number reads: mm10, 22,811,922; dm6, 8,933,764
2-WT-H3K27Ac, total number reads: mm10, 23,571,881; dm6, 8,717,785
3-WT-H3K27Ac, total number reads: mm10, 26,129,676; dm6, 9,378,885
1-H1DKO-H3K27Ac, total number reads: mm10, 18,992,351; dm6, 6,909,044
2-H1DKO-H3K27Ac, total number reads: mm10, 16,722,924; dm6, 7,466,262
3-H1DKO-H3K27Ac, total number reads: mm10, 21,355,372; dm6, 8,752,978

CutAndRun experiments:
1-WT-H3K4me3, total number reads: mm10, 3,815,224
2-WT-H3K4me3, total number reads: mm10, 5,541,738
1-H1DKO-H3K4me3, total number reads: mm10, 4,812,484
2-H1DKO-H3K4me3, total number reads: mm10, 5,854,955
1-WT-H3K9me2, total number reads: mm10, 6,981,537
2-WT-H3K9me2, total number reads: mm10, 6,448,278
1-H1DKO-H3K9me2, total number reads: mm10, 7,147,881
2-H1DKO-H3K9me2, total number reads: mm10, 6,496,731

1-WT-H3K9me3, total number reads: mm10, 7,617,674
2-WT-H3K9me3, total number reads: mm10, 9,044,632
1-H1DKO-H3K9me3, total number reads: mm10, 7,585,753
2-H1DKO-H3K9me3, total number reads: mm10, 8,276,380

ATAC-seq experiment:
DKO1-GCB: 75,394,476
DKO2-GCB: 68,819,276
DKO3-GCB: 63,692,712
WT1-GCB: 79,495,180
WT2-GCB: 68,697,976
WT3-GCB: 73,202,966

Hi-C number of intra-chromosomal reads used for downstream reads ('ds.accepted intra'):
WT_H1_1, total number of reads: 216,296,915
WT_H1_2, total number of reads: 222,714,033
H1DKO_1, total number of reads: 246,312,138
H1DKO_2, total number of reads: 237,201,022

Single cell-RNA seq:
WT4-GCB: 283,715,494
WT5-GCB: 237,596,846
DKO2-GCB: 210,177,503
DKO6-GCB: 257,475,313

Antibodies	For ChIP, the following primary antibodies were used: anti-H3 K27me3, C36B11 - Cell Signaling 9733 (20micrograms), anti-H3 K36me2, C75H12 - Cell Signaling 2901 (20micrograms)

Peak calling parameters

Peaks were called using SICER algorithm:
GCB_WT_K27me3_pooledSicer: 24,131
GCB_H1DKO_K27me3_pooledSicer: 13,603
GCB_WT_K36me2_pooledSicer: 29,466
GCB_H1DKO_K36me2_pooledSicer: 35,553
GCB_WT_K27ac_pooledSicer: 10,522
GCB_H1DKO_K27ac_pooledSicer: 16,790
GCB_WT_K4me3_pooledSicer: 13,393
GCB_H1DKO_K4me3_pooledSicer: 14,761
GCB_WT_K9me2_pooledSicer: 77,330
GCB_H1DKO_K9me2_pooledSicer: 90,670
GCB_WT_K9me3_pooledSicer: 103,436
GCB_H1DKO_K9me3_pooledSicer: 73,051

ATAC-seq peaks were called:
Sample_30_GCB_WT_repl1_ATACseq: 141,700
Sample_31_GCB_WT_repl2_ATACseq: 143,301
Sample_32_GCB_WT_repl3_ATACseq: 147,478
Sample_34_GCB_H1_DKO_repl1_ATACseq: 157,608
Sample_35_GCB_H1_DKO_repl2_ATACseq: 150,869
Sample_36_GCB_H1_DKO_repl3_ATACseq: 134,151

Data quality

Resulting FASTQ files were aligned to mouse mm10 and Drosophila dm6 genomes using bwa-mem function of the BWA suite. Individual sequencing experiments were assessed for their percent of mapped reads (to mm10 and dm6) to ensure proper coverage. ChIPseq data was normalized to dm6 spike-in reads using CompChIPseq algorithm. H3K27me3 and H3K36me2 ChIPseq peaks were called using the SICER algorithm on pooled BAM files from all respective replicates. Loci showing differences in ChIPseq abundance were determined by calculating the Comp-ChIPseq normalized read count within the union of peaks from both genotypes using the multiBigwigSummary function of the deepTools package.

For Hi-C, mapped reads were filtered by the GenomicTools tools-hic filter command integrated in HiC-bench for known artifacts of the Hi-C protocol. The filtered reads include multi-mapped reads ('multihit') read-pairs with only one mappable read, duplicated read-pairs, low mapping quality reads, read-pairs resulting from self-ligated fragments, and short-range interactions, resulting from read-pairs aligning within 25kb. For the downstream analyses, all the accepted intra-chromosomal read-pairs ('ds.accepted intra') were used.

Software

ChIP-seq reads were aligned using bwa-mem function of the BWA suite. Peaks were called using the SICER algorithm. Read density tracks were visualized using the Integrative Genomics Viewer (IGV).

Flow Cytometry

Plots

Confirm that:

☒ The axis labels state the marker and fluorochrome used (e.g. CD4-FITC).

☒ The axis scales are clearly visible. Include numbers along axes only for bottom left plot of group (a 'group' is an analysis of identical markers).

☒ All plots are contour plots with outliers or pseudocolor plots.

☒ A numerical value for number of cells or percentage (with statistics) is provided.

Methodology

Sample preparation	Single cell suspensions from mouse spleens were separated by Ficoll gradient centrifucation and stained using fluorescent labeled anti mouse antibodies, incubated on ice in the dark for 30min, then washed 2x with BS with .5% BSA and 5mM EDTA and resuspended in 200ul washing buffer for acquisition
Instrument	BD FACS Canto II
Software	FlowJo v10
Cell population abundance	Sorted human and murine GC B cells were confimed to be >90% following each sorting.
Gating strategy	For selection of single cells, cells were first gated based on FSC-H/FSC-A, then DAPI/FSC-A followed by FSC-A/SSC-A gating. For Sorting: lymphocytes were further gated on APC-B220 positive population on a SSC-A/APC-A dotplot. GC B cells were gated on GL7+FAS+ populations on a BV421-A/PE-Cy7 boxplot (Log axes). For regular flow: live B cells were selected as B220-APC+ cells on APC-A/SSC-A dotplot followed by gating GC B cells either as GL7-FITC+/FAS-PE+ or CD38-APC negative /FAS-PE positive on dotplot with logarithmic axes.

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