# Nucleosome repositioning in chronic lymphocytic leukemia

Kristan V. Piroeva,[1] Charlotte McDonald,[2] Charalampos Xanthopoulos,[2] Chelsea Fox,[1] Christopher T. Clarkson,[1,9] Jan-Philipp Mallm,[3,4,5] Yevhen Vainshtein,[6] Luminita Ruje,[1] Lara C. Klett,[4,5] Stephan Stilgenbauer,[7] Daniel Mertens,[7,8] Efterpi Kostareli,[2] Karsten Rippe,[4,5] and Vladimir B. Teif[1]

*[1]School of Life Sciences, University of Essex, Wivenhoe Park, Colchester CO4 3SQ, United Kingdom; [2]Wellcome-Wolfson Institute for Experimental Medicine, Queen's University Belfast, Belfast BT9 7BL, United Kingdom; [3]German Cancer Research Center (DKFZ) Heidelberg, Single Cell Open Lab, 69120 Heidelberg, Germany; [4]German Cancer Research Center (DKFZ) Heidelberg, Division of Chromatin Networks, 69120 Heidelberg, Germany; [5]Center for Quantitative Analysis of Molecular and Cellular Biosystems (BioQuant), Heidelberg University, 69120 Heidelberg, Germany; [6]Fraunhofer-Institut für Grenzflächen- und Bioverfahrenstechnik IGB, 70569 Stuttgart, Germany; [7]Division of CLL, University Hospital Ulm, Department of Internal Medicine III, 89081 Ulm, Germany; [8]German Cancer Research Center (DKFZ) Heidelberg, Cooperation Unit Mechanisms of Leukemogenesis, 69120 Heidelberg, Germany*

The location of nucleosomes in the human genome determines the primary chromatin structure and regulates access to regulatory regions. However, genome-wide information on deregulated nucleosome occupancy and its implications in primary cancer cells is scarce. Here, we conducted a genome-wide comparison of high-resolution nucleosome maps in peripheral blood B cells from patients with chronic lymphocytic leukemia (CLL) and healthy individuals at single-base-pair resolution. Our investigation uncovered significant changes of nucleosome positioning in CLL. Globally, the spacing between nucleosomes—the nucleosome repeat length (NRL)—is shortened in CLL. This effect is stronger in the more aggressive IGHV-unmutated CLL subtype than in the IGHV-mutated CLL subtype. Changes in nucleosome occupancy at specific sites are linked to active chromatin remodeling and reduced DNA methylation. Nucleosomes lost or gained in CLL marks differential binding of 3D chromatin organizers such as CTCF as well as immune response–related transcription factors and delineated mechanisms of epigenetic deregulation. The principal component analysis of nucleosome occupancy in cancer-specific regions allowed the classification of samples between cancer subtypes and normal controls. Furthermore, patients could be better assigned to CLL subtypes according to differential nucleosome occupancy than based on DNA methylation or gene expression. Thus, nucleosome positioning constitutes a novel readout to dissect molecular mechanisms of disease progression and to stratify patients. Furthermore, we anticipate that the global nucleosome repositioning detected in our study, such as changes in the NRL, can be exploited for liquid biopsy applications based on cell-free DNA to stratify patients and monitor disease progression.

[Supplemental material is available for this article.]

Chronic lymphocytic leukemia (CLL) is the most common blood cancer in adults in the Western world. Over the past decade, novel therapies against specific targets, like Bruton tyrosine kinase, have emerged in parallel with an increased understanding of its molecular pathogenesis (Quesada et al. 2013; Ferreira et al. 2014; Nabhan and Rosen 2014; Bosch and Dalla-Favera 2019). Previous genome-wide studies of CLL (epi)genomics and transcriptomics have focused on DNA mutations (Landau et al. 2015; Puente et al. 2015) and the interplay between gene expression, deregulated chromatin features like DNA methylation, histone modifications, chromatin accessibility, and transcription factor (TF) binding, as well as long-range chromatin interactions (Ferreira et al. 2014; Kulis et al. 2015; Queirós et al. 2015; Oakes et al. 2016; Rendeiro et al. 2016; Beekman et al. 2018; Mallm et al. 2019; Vilarrasa-Blasi et al. 2021). However, the primary chromatin structure of CLL with respect to the location of nucleosomes within the genome has not been systematically characterized in patient samples, which is also the case for almost all other tumor entities. Nucleosome positioning affects gene expression by modulating accessibility of TFs to their DNA binding sites as an important part of gene regulation in eukaryotes. Thus, nucleosome maps provide insight into gene regulatory mechanisms in disease and can be used for diagnostics in liquid biopsies (Shtumpf et al. 2022). Genome-wide maps of nucleosome positions are usually obtained by digesting linker DNA with nucleases, such as MNase, followed by sequencing the DNA associated with the histone octamer core (Teif and Clarkson 2019; Shtumpf et al. 2022). Most previous studies of this kind investigated cell lines or nonmalignant cells (Schones et al. 2008; Valouev et al. 2011; Gaffney et al. 2012;

**[9]Present address: University College London, London WC1E 6BT, UK**
**Corresponding authors: vteif@essex.ac.uk, karsten.rippe@dkfz.de, e.kostareli@qub.ac.uk**

Article published online before print. Article, supplemental material, and publication date are at https://www.genome.org/cgi/doi/10.1101/gr.277298.122. Freely available online through the *Genome Research* Open Access option.

Kundaje et al. 2012; Diermeier et al. 2014; Ho et al. 2014). Related methods include single-cell-based Strand-seq (Jeong et al. 2023), which is inherently more stochastic, as well as ATAC-seq, which is frequently used for cancer cells and tissues (Grandi et al. 2022). The latter method is well suited to detect repositioning of one to two nucleosomes near TF binding sites (Beekman et al. 2018; Mallm et al. 2019) but does not provide genome-wide nucleosome occupancy maps. Precisely mapped nucleosome locations across the whole genome of malignant and nonmalignant cells are becoming increasingly important for clinical diagnostics of cell-free DNA (cfDNA), which contains nucleosome-protected genomic regions (Snyder et al. 2016; Shtumpf et al. 2022). This need goes beyond blood cancers, because even in the case of solid tumors, most of the pieces of cfDNA come from blood cells and only a small fraction originates from tumor tissues.

Here, we exploited our high-throughput sequencing experiments with MNase-assisted histone H3 ChIP-seq (Mallm et al. 2019) at a sequencing depth of more than 4 billion DNA sequence reads to address the following aims: (1) derive complete genome-wide nucleosome occupancy maps at base-pair resolution and detect individual nucleosome position changes in human primary malignant versus nonmalignant B cells (NBCs), from patients with CLL and healthy donors; (2) dissect both the repositioning of individual nucleosomes and the collective behavior of nucleosomes in CLL at the global chromatin reorganization level, for example, with respect to the average spacing between nucleosomes; and (3) compare nucleosome occupancy in two clinically important CLL subtypes that have the immunoglobulin heavy-chain variable region (IGHV) mutated (M-CLL) or unmutated (U-CLL), associated with favorable or poor prognosis, respectively.

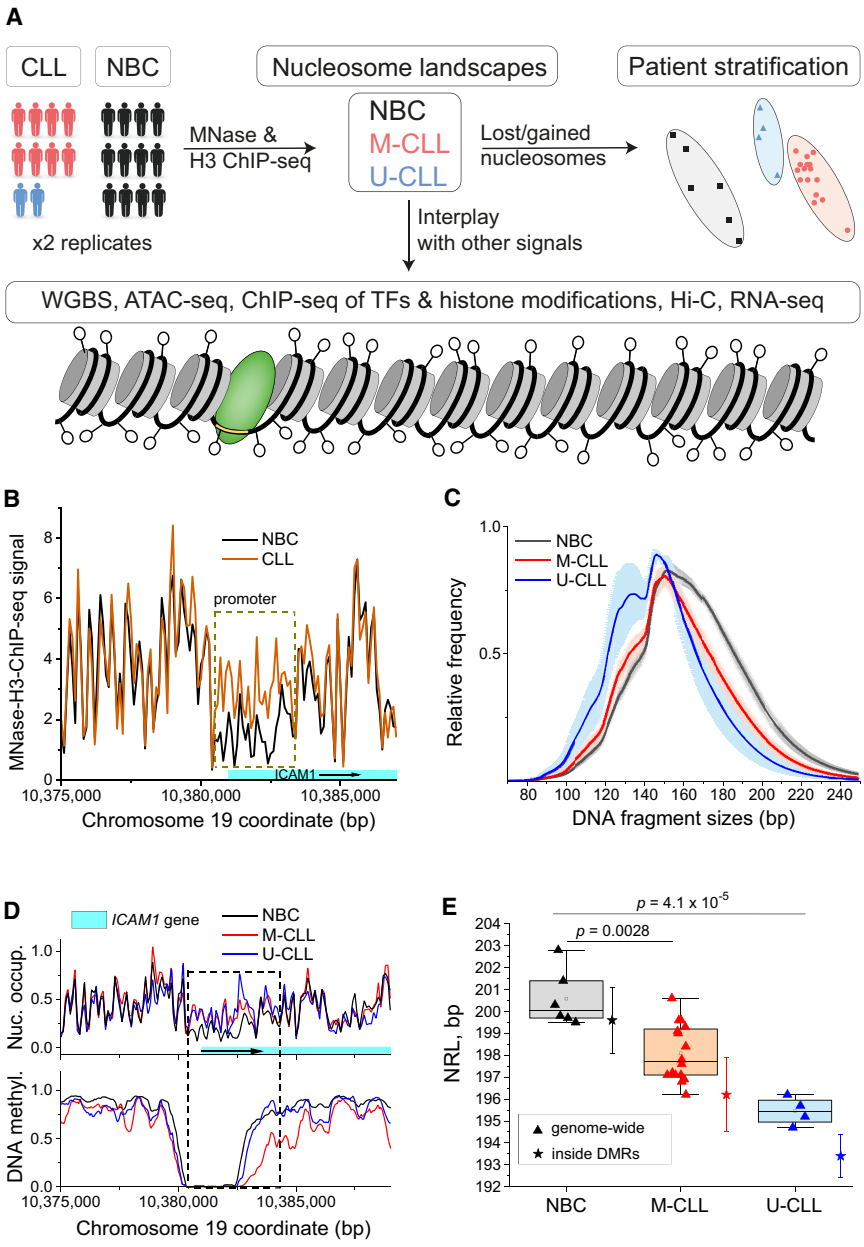


**Figure 1.** Changes of nucleosome positioning between NBCs, M-CLL, and U-CLL. (*A*) Data sets and readouts. Nucleosome landscapes were derived from MNase-assisted H3 ChIP-seq for nonmalignant B cells (NBCs) from healthy donors or CLL patients stratified into IGHV-mutated (M-CLL) or unmutated (U-CLL). These maps were integrated with data from WGBS, ATAC-seq, ChIP-seq of transcription factors (TFs) and six histone modifications, Hi-C, and RNA-seq to dissect molecular mechanisms of nucleosome repositioning. (*B*) Nucleosome occupancy maps in NBCs and CLL at a genomic region enclosing the promoter of the gene *ICAM1*. Gene expression of *ICAM1* was 10-fold reduced in CLL. The total signal of MNase-assisted H3 ChIP-seq is used without size selection of DNA fragments. (*C*) Distributions of DNA fragment sizes from MNase-assisted H3 ChIP-seq in NBCs, M-CLL, and U-CLL. (*D*) Nucleosome occupancy and DNA methylation for the same region as in panel *B* but using only 120- to 180-bp-sized DNA fragments. (*E*) Genome-wide nucleosome repeat length (NRL). A decrease from ~200 bp in NBCs (black) to ~198 bp in M-CLL (red) to ~195 bp in U-CLL (blue) is apparent. Each triangle symbol corresponds to one biological sample. Colored boxes indicate 25%–75% confidence interval; whiskers, range within 1.5 IQR; horizontal line, median; and open squares, mean values. Values inside DMRs shown by asterisk symbols correspond to cohort averages for NBCs (black), M-CLL (red), and U-CLL (blue), with vertical bars depicting the standard deviation.

## Results

### CLL is characterized by global changes of nucleosome positioning

We determined nucleosome positions in malignant B cells from CLL patients and NBCs pooled from healthy individuals using high-coverage MNase-assisted histone H3 ChIP-seq (Fig. 1A). About 4 billion paired-end reads obtained in our experiments were combined to compare the average nucleosome occupancies in 100-bp bins across the genome in CLL patients versus healthy individuals. The Pearson's correlation between samples was high ($r = 0.95$), indicating that only a small fraction of the genome

undergoes significant changes of nucleosome occupancy in CLL versus NBCs, and random nonspecific regions displayed very similar profiles (Supplemental Fig. S2A). However, the averages across groups showed clear differences between nonmalignant and malignant cells (Supplemental Fig. S1) particularly at regulatory regions (Fig. 1B; Supplemental Figs. S2B–D, S3). Such regions typically displayed distinct nucleosome occupancy changes in nanodomains of ~1–5 kb in size. In addition, we observed two global genome-wide effects. First, the length of nucleosome-protected DNA fragments was shorter in both CLL subtypes than in the nonmalignant controls (Fig. 1C), with higher enrichment of subnucleosomal fragments in U-CLL. Studies on cfDNA of cancer patients have shown that it is also enriched with shorter fragments in comparison to cfDNA of healthy individuals (Lo et al. 2021). To study nucleosome repositioning separately from this effect, in the following analysis we applied filtering to consider only DNA fragment sizes between 120 and 180 bp (Fig. 1D; Supplemental Figs. S1G, S3). A second global, genome-wide difference was the shortening of the average distance between the centers of neighboring nucleosomes, which is given by the nucleosome repeat length (NRL) (Fig. 1E). The average NRL decreased from 200.6 ± 0.6 bp in NBCs to 198.1 ± 0.5 bp in M-CLL to 195.5 ± 0.6 bp in U-CLL. These differences were statistically significant (two-sample *t*-test; $P = 0.0028$ for M-CLL vs. NBCs, $P = 4.1 \times 10^{-5}$ for U-CLL vs. NBCs) and comparable to NRL changes previously reported between different cell types during stem cell differentiation (Teif et al. 2012). Furthermore, this effect was even more pronounced inside differentially methylated regions (DMRs), where average NRL decreased from 199.6 bp in NBCs to 196.2 bp in M-CLL to 193.4 bp in U-CLL.

## Nucleosome repositioning occurs preferentially in DNA methylation–depleted regions

Next, we asked what distinguishes genomic loci with (un)changed nucleosome occupancy in CLL. DNA methylation appeared to stabilize nucleosomes, with nucleosomes in highly methylated regions keeping their locations across NBCs, U-CLL, and M-CLL (Fig. 2A). In contrast, nucleosomes that shifted by >20% between NBCs and CLL were mostly depleted of DNA methylation in all conditions (Fig. 2B). In addition, we also detected significant differences of nucleosome occupancy associated with CLL-specific changes of DNA methylation (Fig. 2C). The difference in nucleosome occupancy was clearly detectable in NBCs versus M-CLL and was even more

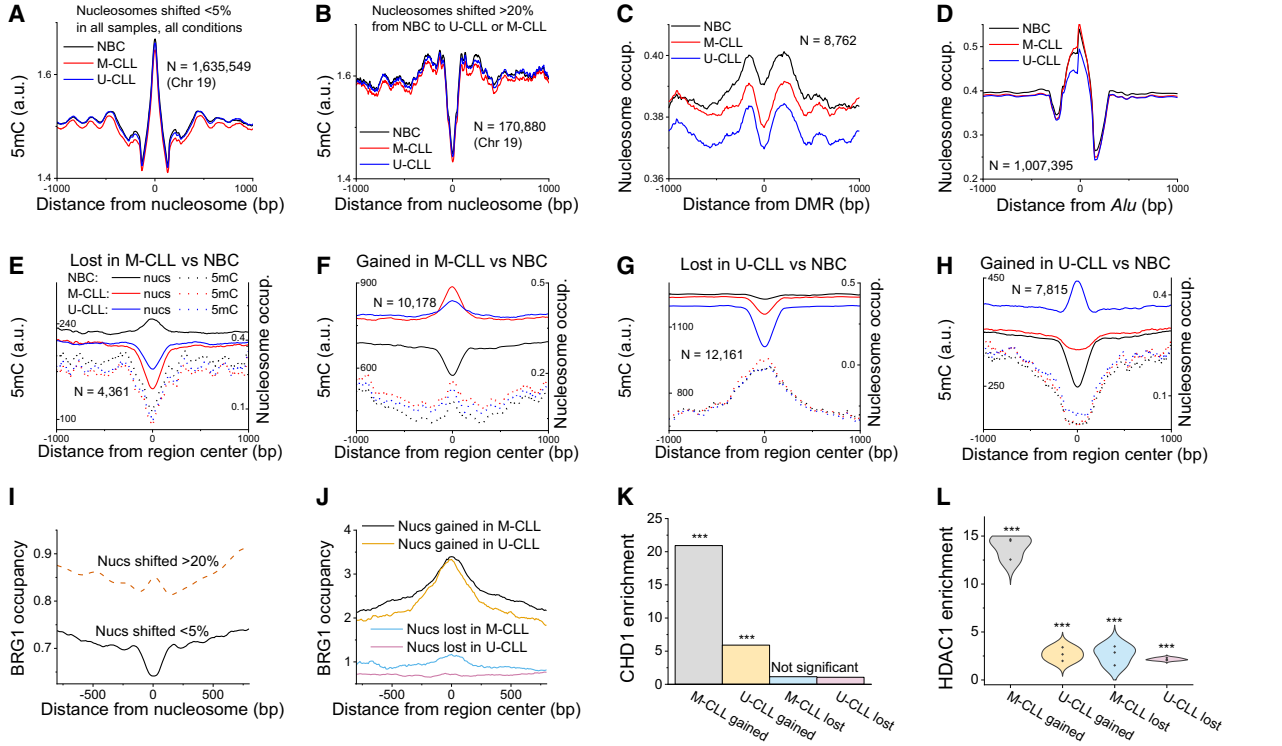


**Figure 2.** Mechanistic determinants of nucleosome repositioning in CLL. (*A*) DNA methylation occupancy profiles around centers of nucleosome-protected DNA fragments on Chromosome 19 for stable nucleosomes with an overlap between different sample types of >95%. (*B*) Same as panel *A* but for shifted nucleosomes. The overlap of positions between NBCs and CLL was <80%. (*C*) Nucleosome occupancy around DMRs with decreased DNA methylation in CLL versus NBCs. (*D*) Nucleosome occupancy around *Alu* repeats. (*E*) Averaged nucleosome occupancy (solid lines) and DNA methylation profiles (dashed lines) around centers of 100-bp regions with decreased nucleosome occupancy (lost nucleosomes) in M-CLL versus NBCs. The profiles of nucleosome occupancy (solid lines) and DNA methylation (dashed lines) were averaged across all NBCs (black), M-CLL (red), and U-CLL (blue) samples. The number of regions used in the calculation (N) is indicated on the graph. (*F*) Same as panel *E* but for regions with increased nucleosome occupancy (gained nucleosomes). (*G*) Same as panel *E* but for regions with decreased nucleosome occupancy in U-CLL versus NBCs. (*H*) Same as panel *G* but for regions with increased nucleosome occupancy. (*I*) BRG1 occupancy in NBCs around centers of stable and shifted nucleosomes at the same regions shown in panel *A*. (*J*) BRG1 occupancy in NBCs around regions that lost or gained nucleosomes in CLL versus NBCs for the same regions shown in panels *E*–*H*. (*K*) Enrichment of CHD1 ChIP-seq peaks determined in the lymphoblastoid cell line GM12878 inside regions that undergo nucleosome loss/gain in CLL versus NBCs. (*L*) Enrichment of HDAC1 ChIP-seq peaks determined in peripheral blood mononuclear cells from CLL patients with a 11q deletion genotype (GSE216287) inside regions that undergo nucleosome loss/gain in CLL versus NBCs. The violin plots correspond to the distribution based on three ChIP-seq replicates. Fisher's test, $P < 10 \times 10^{-3}$ for all points.

pronounced in NBCs versus U-CLL. For comparison, we calculated nucleosome profiles around *Alu* repeats, which are known to stabilize nucleosomes (Fig. 2D; Teif et al. 2017). Nucleosome occupancy around centers of *Alu* repeats showed a distinct, stable profile across all conditions, underlining that the variability we detected in CLL cells is coupled to DNA methylation.

### Nucleosome relocation correlates with chromatin remodeling activity

To characterize regions with increased or decreased nucleosome occupancy in CLL in further detail, we segmented the whole genome into nonoverlapping 100-bp bins and calculated the normalized nucleosome occupancy for each bin. We then defined regions with stable nucleosome occupancy across all samples from the same condition and conducted a pairwise comparison between conditions. Regions where nucleosome occupancy was significantly different between two conditions were selected for further analysis and annotated as lost-nucleosome or gained-nucleosome regions correspondingly. This analysis identified thousands of regions with differential nucleosome occupancy in the comparison of M-CLL versus NBCs (Fig. 2E,F) and U-CLL versus NBCs (Fig. 2G,H). The nucleosome occupancy profiles of these regions were more similar between the two CLL subtypes and distinct from NBCs. DNA methylation profiles also showed differences between the M-CLL and U-CLL subtypes (Fig. 2E–H): Regions that lost nucleosomes in U-CLL versus NBCs were enriched with DNA methylation, whereas those that lost nucleosomes in M-CLL versus NBCs were depleted of DNA methylation. Similarly, regions that gained nucleosomes in U-CLL versus NBCs (but not in M-CLL vs. NBCs) were depleted of DNA methylation. These DNA methylation profiles were highly reproducible across all samples of the same condition (Supplemental Fig. S4). To assess whether nucleosome repositioning in CLL can be linked to the activity of chromatin remodeling complexes, we integrated three previously determined ChIP-seq data sets of chromatin remodelers into our analysis. (1) We compared the nucleosome maps in NBCs with the binding sites of BRG1, the ATPase of the SWI/SNF chromatin remodeling complex (Abraham et al. 2013). Nucleosomes that did not change locations between conditions were depleted of BRG1 in NBCs, whereas shifted nucleosomes (>20% change of their start/end coordinates) were enriched with BRG1 (Fig. 2I). The regions that gained nucleosomes in both M-CLL and U-CLL versus NBCs had a more than threefold enrichment of BRG1 in comparison with neighboring regions (Fig. 2J). (2) The loci bound by another chromatin remodeler, CHD1, in lymphoblastoid B cells (The ENCODE Project Consortium 2012) were enriched at sites that gained nucleosomes (Fig. 2K). M-CLL samples were characterized by a more than zerofold enrichment. (3) The regions that gained nucleosomes in M-CLL displayed an approximately 15-fold enrichment at binding sites of the histone deacetylase HDAC1, a subunit of the NuRD chromatin remodeling complex (Fig. 2L; Supplemental Fig. S5; Lai and Wade 2011). This value was determined using a recently published ChIP-seq data set from peripheral blood mononuclear cells from CLL patients (Lai et al. 2023). These consistent findings for three different complexes, SWI/SNF, CHD1, and NuRD, suggest that nucleosome relocation in CLL is subtype specific and is likely to involve active chromatin remodeling.

### Nucleosome repositioning marks CLL-specific gene regulation

Regions with gained nucleosome occupancy in one of the CLL subtypes versus NBCs were enriched with promoters, active enhancers, CpG islands, and CTCF binding sites, whereas the most pronounced nucleosome loss was at enhancers active in CLL (Fig. 3A,B). For CTCF sites, the fold enrichment of regions with increased nucleosome occupancy was larger for those with high similarity to the consensus CTCF motif. Regions with gained nucleosome occupancy in CLL were enriched with cancer-related pathways including both generic pathways typically deregulated in cancer as well as CLL-specific ones, such as B cell receptor signaling (BCR) (Fig. 3C, D). The largest group of genes marked by regions that gained nucleosomes in CLL was related to the immune system (more than 350 genes, $P = 0.01$). A similar pairwise comparison of regions with differential nucleosome occupancy between U-CLL and M-CLL is shown in Supplemental Figure S6, C and D. In addition, we repeated this analysis including all DNA fragments from MNase-assisted H3 ChIP-seq without size selection (Supplemental Fig. S7). It revealed that the largest fraction of nucleosomes repositioned in CLL versus NBCs resides in 700,000 regions of 100-bp size. These loci were termed "variable" because their nucleosome occupancy significantly varied across CLL patients while being stable across all NBC samples (Supplemental Fig. S7A). The most informative group of nucleosome changes that distinguished CLL was the fraction of gained nucleosomes in regions where nucleosome occupancy in CLL increased in comparison with NBCs (Supplemental Fig. S7D). The intersection of such gained-nucleosome regions with gene promoters marked the B cell receptor signaling pathway (BCR) as the top hit ($P = 1.6 \times 10^{-5}$), followed by the T cell receptor signaling pathway ($P = 3.2 \times 10^{-4}$) (Supplemental Tables S1–S3).

The correlated differences of nucleosome occupancy and gene activity could also be related to a differential activity of *cis*-regulatory enhancer regions in NBCs and CLL. Active enhancers, as defined previously from multiomics profiling in this patient cohort (Mallm et al. 2019), were characterized by nucleosome depletion. In contrast, all enhancers, which contain many inactive ones, were characterized on average by increased nucleosome occupancy (Fig. 4A–D). Of note, active NBC enhancers had very similar DNA methylation profiles between U-CLL and M-CLL, whereas their nucleosome occupancy profiles distinguished between the two CLL subtypes (Fig. 4C). Nucleosome occupancy differed between CLL and NBCs even at the scale of ~100 kb, based on A and B chromatin compartments defined with Hi-C (Vilarrasa-Blasi et al. 2021). The more open and active A compartments displayed nucleosome depletion in NBCs but not in U-CLL (Fig. 4E). In contrast, the inactive and closed B compartments displayed a flat profile of nucleosome occupancy in all conditions (Fig. 4F). Thus, functionally relevant differences in CLL nucleosome occupancy occurred both at the level of single nucleosomes at promoters and enhancers and on larger scale of active chromatin subcompartments.

### CLL-specific changes in nucleosome positioning, CTCF binding, and histone modifications are correlated

Next, we focused on regions with dysregulated CTCF binding (Fig. 5). Sites that were bound by CTCF both in NBCs and CLL were defined as "common." Nucleosome occupancy profiles around common CTCF sites had the same shape for NBCs and CLL. They showed a pronounced depletion in the center surrounded by characteristic oscillations, reflecting the regular nucleosome array positioned by CTCF. The average DNA methylation profiles were depleted in a wider area of ±200 bp surrounding common CTCF sites both in NBCs and CLL (Fig. 5A). In the case of sites that lost CTCF binding in CLL (termed "lost"), NBCs were characterized by nucleosome depletion around CTCF binding sites, as opposed to a

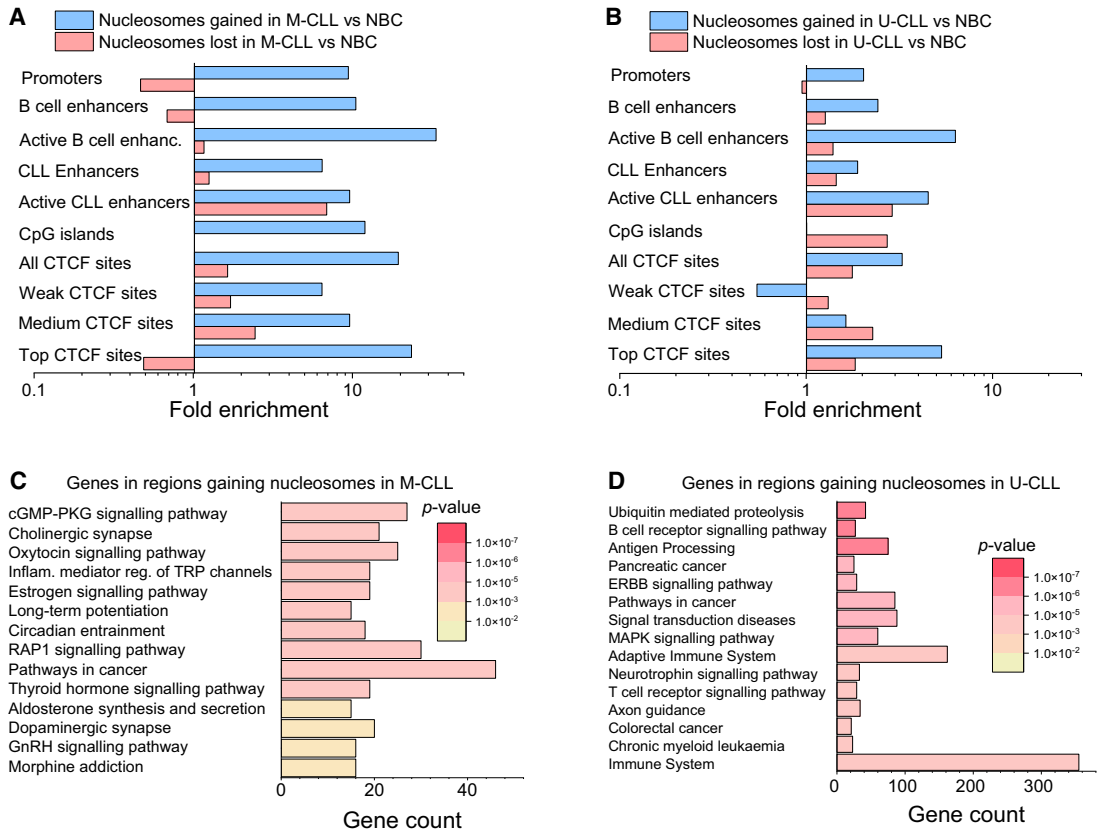


**Figure 3.** Annotation of sites with differential nucleosome occupancy. (*A*) Genomic regions that lost (red) or gained (blue) nucleosomes in M-CLL versus NBC. (*B*) Same as panel *A* but for U-CLL versus NBC. (*C*) Gene Ontology analysis of genes that overlap with regions that gained nucleosomes in M-CLL versus NBC. (*D*) Same as panel *C* but for U-CLL versus NBC.

peak of nucleosome occupancy in U-CLL and M-CLL (Fig. 5B). DNA methylation profiles also gained peaks near lost CTCF sites in CLL. Moreover, M-CLL and U-CLL displayed different trends. M-CLL had higher nucleosome occupancy around both common and lost CTCF sites, whereas the increase in DNA methylation near lost CTCF sites was higher in U-CLL. We also investigated the effect of CTCF loss in CLL on different histone modifications (Fig. 5C). The largest changes were observed for a reduction of the active enhancer mark H3K4me1 and increase of the repressive heterochromatin marks H3K9me3 and H3K27me3 in CLL. In addition, we evaluated a recently reported pan-cancer set of CTCF sites undergoing consistent changes across several cancer types (Supplemental Fig. S8A,B; Fang et al. 2020). Nucleosome occupancy at these CTCF sites was not significantly different between NBCs and CLL, suggesting that the pan-cancer data set of Fang et al. (2020) is less specific for CLL than the variable CTCF sites defined in our work.

## Nucleosome repositioning reveals differential TF binding in CLL

To study effects of TF-nucleosome interaction, we determined enriched TF motifs inside regions with differential nucleosome occupancy in CLL (Fig. 6A), calculated nucleosome profiles around their binding sites (Fig. 6B,C), and visualized gene networks affected by nucleosome repositioning at the corresponding TF binding sites in M-CLL and U-CLL (Fig. 6D). This analysis revealed three major cases. (1) For one set of TF binding sites, nucleosome profiles in NBCs, M-CLL, and U-CLL had a similar shape but differed in average nucleosome occupancy (e.g., ZNF263, PAX9, and PAX5 in Fig. 6B). For these motifs, U-CLL and M-CLL were more similar to each other than to NBCs. (2) In NBCs, M-CLL, and U-CLL, the shape of nucleosome profiles was similar, but NBCs and M-CLL profiles were located closer to each other and further away from U-CLL (e.g., TCLF5, HES5, and E2F3 in Fig. 6B). In this group, the different average nucleosome occupancy level likely reflects the change in local nucleosome landscape not related directly to the binding of a given TF but rather determined by the binding of other factors nearby. (3) The shape of nucleosome occupancy profiles changed between CLL and NBCs. These included the transition from low nucleosome occupancy in NBCs to a peak in CLL, for example, JUND, PKNOX2, and FOSL1 (Fig. 6C). Such nucleosome profile changes were accompanied by gene expression differences in CLL versus NBCs, for example, an approximately twofold reduction for FOSL1. Supplemental Figures S9 and S10 show nucleosome occupancy profiles of TFs enriched in regions with differential nucleosome occupancy defined above but limited to binding sites inside CLL-specific ATAC-seq peaks.

In addition to the analysis based on nucleosome-size DNA fragments above, we also performed a similar analysis for all DNA fragments without size-filtering. The analysis of 2508 promoters with gained nucleosomes revealed B cell receptor signaling as the top enriched pathway (Supplemental Tables S1, S2) and the so-called CG-box as the top enriched motif ($P = 1.1 \times 10^{-37}$, 591 out of 2508 sites) (Supplemental Fig. S11B). Thus, TFs recognizing this

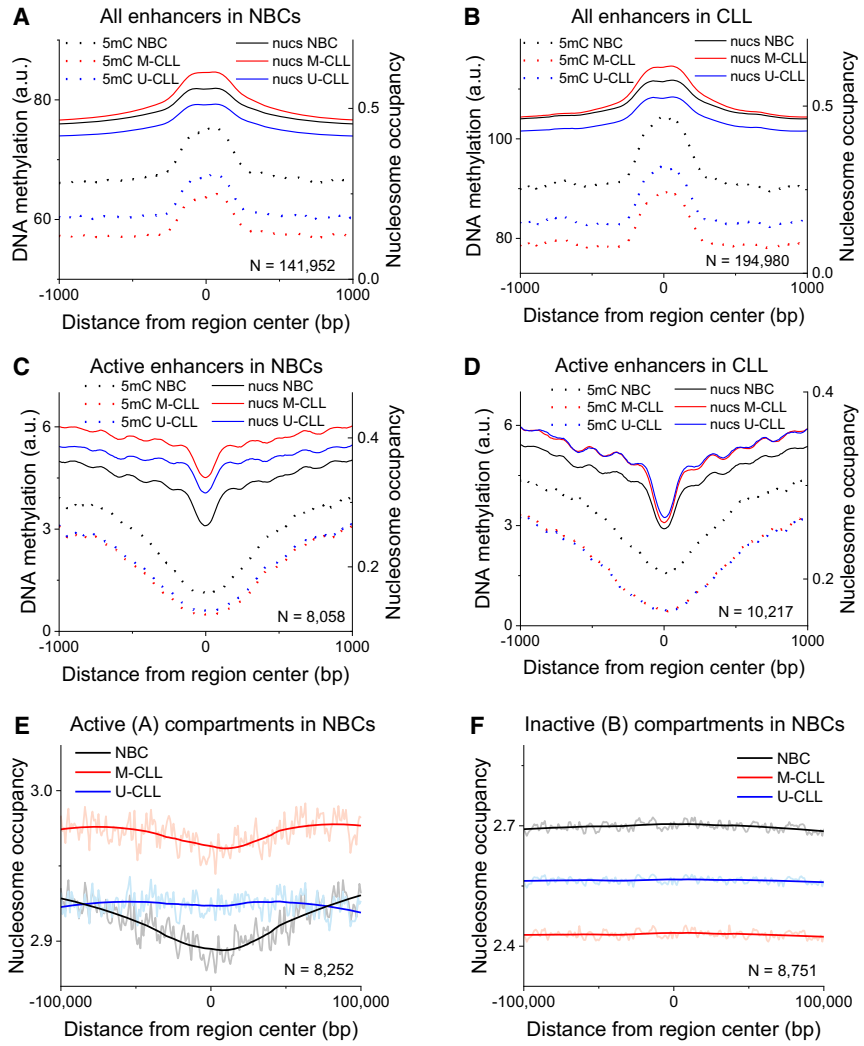


**Figure 4.** Average nucleosome occupancy and DNA methylation occupancy profiles at *cis*-regulatory elements. (*A*) Averaged nucleosome and DNA methylation occupancy profiles at enhancers in NBCs. (*B*) Same as panel *A* but for CLL. (*C*) Same as panel *A* but for active NBC enhancers as determined from ATAC-seq. (*D*) Same as panel *C* but for active CLL enhancer. (*E*) Averaged nucleosome occupancy profiles within the transcriptionally active A chromatin compartments determined from Hi-C in NBCs by Vilarrasa-Blasi et al. (2021). The profiles are averaged over all NBCs (black), M-CLL (red), and U-CLL (blue) samples. The number of regions (N) is indicated on the graphs. (*F*) Same as panel *E* but for the transcriptionally inactive B compartment.

motif such as SP1/2 and EGR1/2 are likely to be involved in CLL deregulation. The expression of genes encoding these TFs was only moderately changed in comparison with nonmalignant controls ($\log_2$ fold change −0.65 for SP1, −2.4 for SP2, −3.0 for EGR1, −0.26 for EGR2), indicating that nucleosome repositioning can be more informative than TF gene expression levels in peripheral blood B cells for assessing TF activity. Supplemental Figure S12 shows that nucleosome occupancy profiles around SP1 change from nucleosome depletion across all SP1 sites to nucleosome occupancy peak for a subset of these sites covered by "variable" nucleosomes in CLL. A less-frequent motif present in promoters of these genes can be recognized by SOX11 and SOX4 among other TFs (Supplemental Fig. S11). SOX11 was proposed as a prognostic marker in CLL (Roisman et al. 2015), whereas SOX4 plays a role in other B cell malignancies (Sarkar and Hochedlinger 2013). Based on RNA-seq in this patient cohort, SOX4 was 128-fold down-regulated in CLL versus nonmalignant controls. An SP1 binding site inside the SOX4 promoter was associated with a distinct nucleosome peak appearing at the place of the nucleosome depletion in nonmalignant controls (Supplemental Fig. S2D), suggesting that differential SP1/2 or EGR1/2 binding might occur upstream of SOX4 dysregulation.

### CLL patients can be stratified based on nucleosome occupancy

We examined the classification of samples based on different epigenetic parameters. First, we asked whether it is possible to distinguish NBCs, M-CLL, and U-CLL based on gene expression. Figure 7A shows that principal component analysis (PCA) of gene expression values clearly distinguished CLL from NBCs but not the CLL subtypes from each other. Second, we attempted the same classification using DNA methylation of gene promoters. Figure 7B depicts the clear differences in DNA methylation between sample types. However, promoter DNA methylation did not segregate the three groups very effectively, and there was a partial overlap among the three clouds/sample types. Finally, we asked whether the nucleosome repositioning as identified in this study can serve as a novel biomarker. For example, nucleosome occupancy at the 100-bp regions that lost nucleosomes in U-CLL versus NBCs in each of the 26 samples (including six NBCs, four U-CLLs, and 16 M-CLLs) was informative in PCA. As shown in Figure 7, C–E, PCA based on nucleosome occupancy allows discerning not only CLL from NBCs but also M-CLL from U-CLL. Notably, regions with differential nucleosome occupancy defined using only one CLL subtype worked to distinguish the other CLL subtype from NBCs as well (Fig. 7C–F). Regions that gained or lost nucleosomes in U-CLL versus NBCs or lost nucleosomes in M-CLL versus NBCs allowed good stratification of all studied subtypes (Fig. 7C–E), whereas regions that gained nucleosomes in M-CLL were common between CLL subtypes and less effective in stratifying patients (Fig. 7F). Our additional calculations showed that unlike gained-nucleosome and lost-nucleosome regions, nucleosome occupancy or DNA methylation inside other genomic features such as promoters, enhancers, or CTCF sites is not as effective in distinguishing between M-CLL, U-CLL, and NBCs (Supplemental Figs. S13, S14).

## Discussion

Large progress in understanding the contribution of chromatin features to CLL pathogenesis has been made in recent years

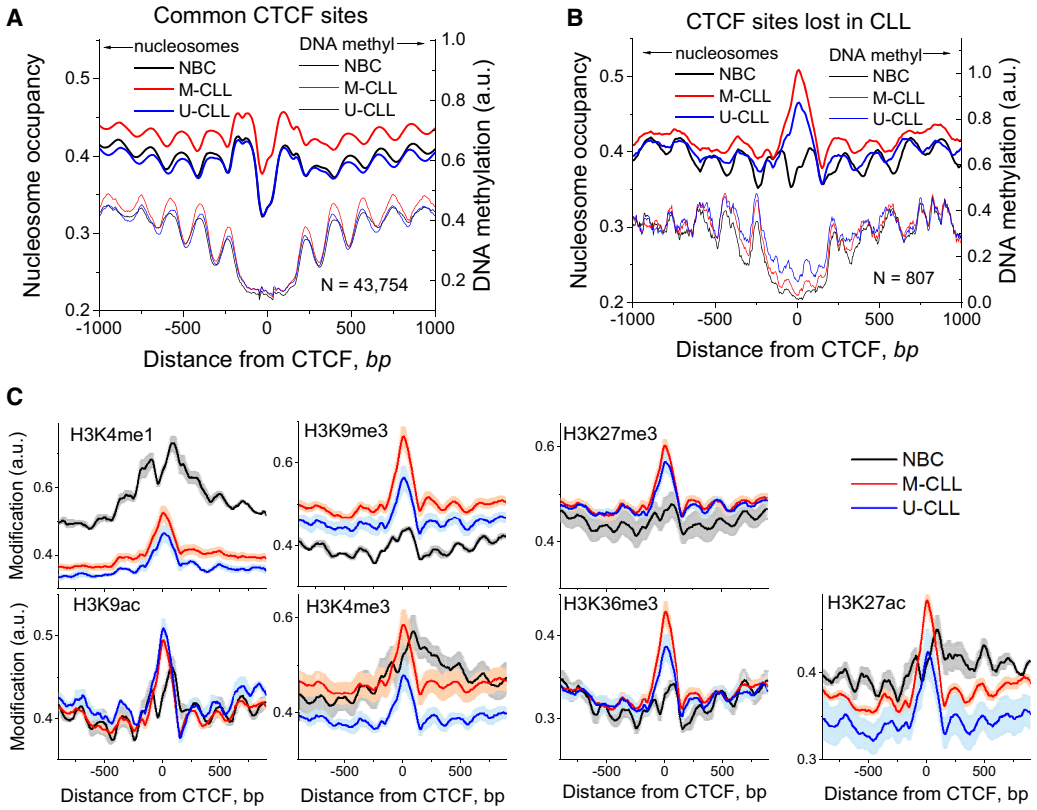


**Figure 5.** Nucleosome occupancy, histone modifications, and DNA methylation around CTCF sites. (*A*) Averaged DNA methylation and nucleosome occupancy profiles around CTCF motifs are bound by CTCF both in NBC and CLL. (*B*) Same as panel *A* but for sites bound by CTCF in NBC but not in CLL. (*C*) Histone modifications around the same set of lost CTCF sites depicted in panel *B*. The profiles were averaged over all NBCs (black), M-CLL (red), and U-CLL (blue) samples. Standard errors are shown for each line in light colors. The number of regions (N) is indicated on the graph

(Beekman et al. 2018; Gaiti et al. 2019; Mallm et al. 2019; Pastore et al. 2019; Rendeiro et al. 2020; Vilarrasa-Blasi et al. 2021). However, the contribution of nucleosome positioning has not been integrated. This essential knowledge gap was addressed here by compiling occupancy maps at single-nucleosome resolution in NBCs, U-CLL, and M-CLL. Based on these data, our study introduces nucleosome positioning as a powerful biomarker and provides novel mechanistic insights of chromatin-mediated deregulation in CLL (Fig. 7G). We detected changes in the nucleosome landscape at two levels. First, our analysis revealed a genome-wide decrease of the NRL in CLL (Fig. 1E), which was most pronounced in the U-CLL subtype. This finding has implications for understanding nucleosome positioning patterns in malignant and nonmalignant B cells and could be applied to cancer diagnostics based on liquid biopsy approaches (Shtumpf et al. 2022). NRL changes at the scale of a few base pairs are known to occur during cell differentiation (Teif et al. 2012) but have not been evaluated previously in primary tumor cells against the same nonmalignant cell type. Furthermore, we observed differences between the two CLL subtypes, with the more aggressive U-CLL subtype having the smallest NRL. The effect of NRL shortening in cancer cells reported here may not be limited to CLL. In a separate study, we found a similar NRL shortening effect in paired tumor versus normal samples from breast cancer patient tissue samples (Jacob et al. 2023). Thus, NRL changes may be generally informative to identify tumor subtypes with distinct disease course. It is worth noting that the cancer-specific decrease in NRL represents a genome-wide feature, integrating information from the large part of the genome, including regions outside of annotated genes.

Our analysis of local changes of nucleosome occupancy showed that DNA methylation correlated with more stably bound nucleosomes, whereas nucleosomes that shifted their position in CLL were at less methylated sites (Fig. 2A,B). This relation was also observed in our previous studies in mouse embryonic stem cells (Teif et al. 2014; Wiehle et al. 2019), but it can be overwritten in specific regions by other mechanisms. For example, regions characterized by differential nucleosome occupancy in CLL can be either depleted or enriched with DNA methylation depending on the CLL subtype (Fig. 2E–H). Thus, DNA methylation can act in two ways: (1) as a targeted mechanism changing the cell fate and (2) as a general protection mechanism from stochastic cancer-associated epigenetic changes. Such nucleosome repositioning is likely happening as an active process, which is reflected by the enrichment of the chromatin remodelers SWI/SNF (BRG1 subunit), CHD1, and NuRD (HDAC1 subunit) at these sites (Fig. 2I–L). SWI/SNF, NuRD, and other remodelers play an important role in normal hematopoiesis and are linked to malignant transformation in leukemia (Andrades et al. 2023; Najm et al. 2023). Mutations in the CHD1 homolog CHD2 are among the most frequent mutations in CLL and are characteristic of the M-CLL subtype (Rodríguez et al. 2015). The concept of preferential nucleosome repositioning in active regions is also consistent with large-scale changes of nucleosome occupancy that we observed at the active A compartment (Fig. 4). It is worth mentioning

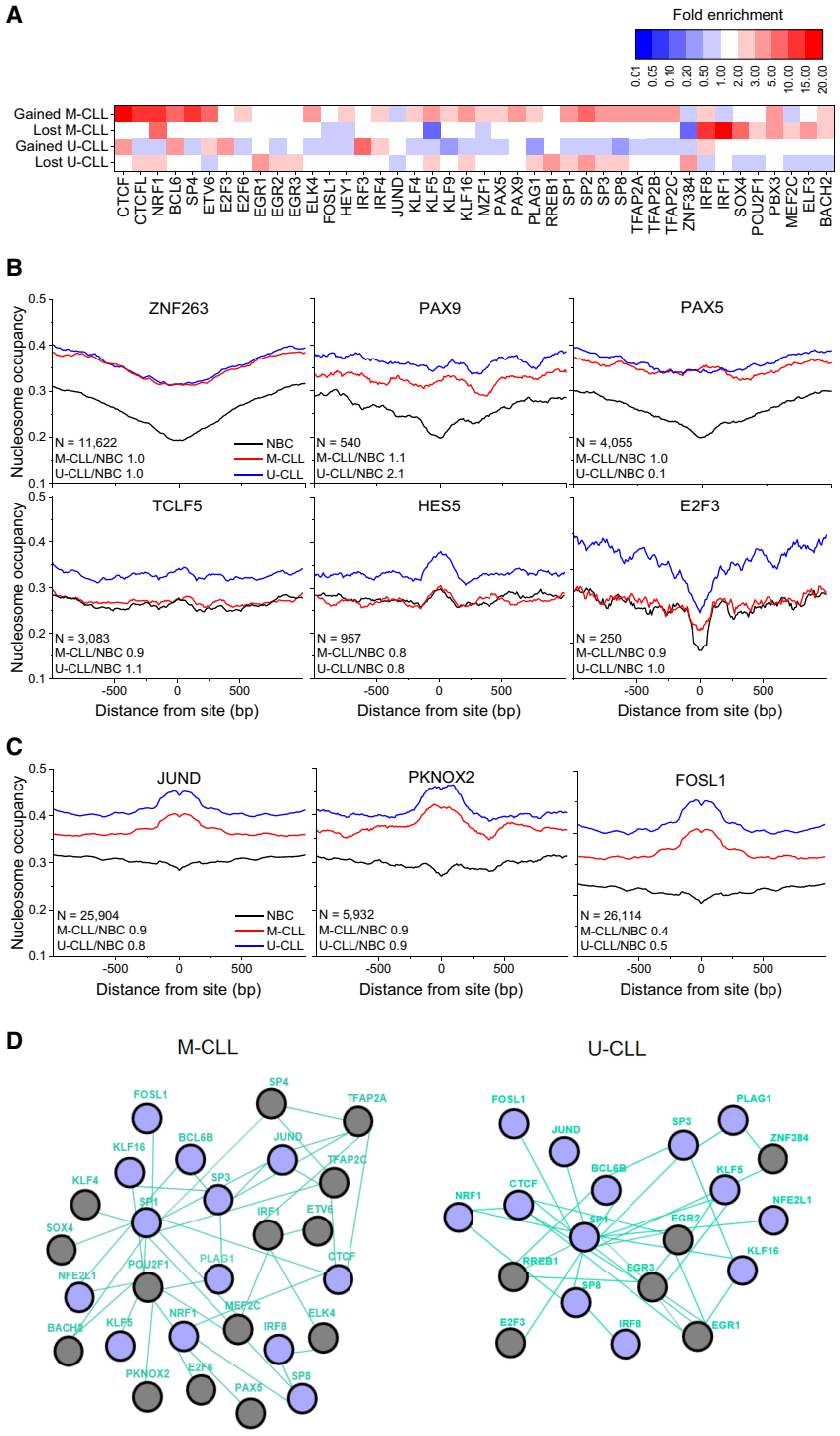


**Figure 6.** Analysis of TF binding in regions with differential nucleosome occupancy. (*A*) Heatmap showing fold-enrichment levels for TFs with more than twofold enrichment of their binding sites in regions that lost or gained nucleosomes in M-CLL or U-CLL versus NBCs or significantly changed the shape of their binding profile in CLL. (*B*) Nucleosome occupancy profiles around binding sites of TFs inside regions that lost ATAC-seq signal in CLL versus NBCs. (*C*) Nucleosome occupancy profiles around binding sites of TFs inside regions that gained ATAC-seq signal in CLL versus NBCs. The profiles are averaged over all NBC (black), M-CLL (red), and U-CLL (blue) samples. The number of regions (N) and fold expression change of the corresponding TF in two conditions, M-CLL versus NBC and U-CLL versus NBC, are indicated on the graph. (*D*) Nucleosome repositioning marks distinct TF pathways, which are different between M-CLL and U-CLL, whereas some network features remain common. Green TF connecting lines indicate known relationships of type "controls expression" based on Pathway Commons (http://www.pathwaycommons.org); purple circles, TFs affected both in U-CLL and M-CLL; and gray circles, TFs unique for CLL subtypes.

that the interplay between DNA methylation and nucleosome remodeling could also be one of the mechanisms leading to NRL changes, because nucleosome spacing changes were more pronounced in DMRs (Fig. 1E).

In addition to the global deregulation of nucleosome positioning, distinct changes of nucleosome occupancy at the scale of single nucleosomes marked individual regulatory regions critical for CLL. These regions were particularly enriched for CTCF binding sites, enhancers, and promoters. Sites with CTCF binding loss became covered by nucleosomes and gained methylation in a CLL subtype–dependent manner (Fig. 5A,B). The changes at promoters were associated with CLL-specific and immunity-related pathways (Fig. 3). The analysis of the interplay between TF binding and nucleosome positioning allowed us to reconstruct CLL-specific TF networks (Fig. 6D). The most enriched TF binding sites inside regions with differential nucleosome occupancy were related to innate immunity (such as IRF8 and IRF1), the germinal center establishment (BCL6) (Capello et al. 2000; Mlynarczyk et al. 2019), 3D genome organization (CTCF and its interaction partners, including the cancer-specific competitor BORIS/CTCFL), and NRF1, a TF that displays DNA methylation–dependent binding (Domcke et al. 2015). It is worth noting that several deregulated TFs detected in this study have not been previously associated with CLL, although some are involved in other cancers (Fig. 6). Overall, three major groups of CLL-specific TFs were identified: (1) TFs for which nucleosome occupancy profiles around binding sites had quantitatively similar changes in the average occupancy level in U-CLL and M-CLL versus NBCs, including PAX9 and PAX5 (Fig. 6B). Higher expression of PAX9 was reported for U-CLL and linked to a poorer clinical outcome (Rani et al. 2017), whereas PAX5 is critical for both NBC development and CLL pathogenesis and evolution (Puente et al. 2015; Ott et al. 2018; Klintman et al. 2021); (2) TFs for which nucleosome profiles in M-CLL were closer to NBC, such as TCLF5, HES5, and E2F3 (Fig. 6B). E2F3 is a miR-34a target and a repressor of the ARF/TP53 pathway. miR-34a has been linked to chemotherapy resistance in CLL, and its low expression is associated with TP53 aberrations and Richter transformation (Balatti et al. 2018); (3) TFs with different shapes of nucleosome

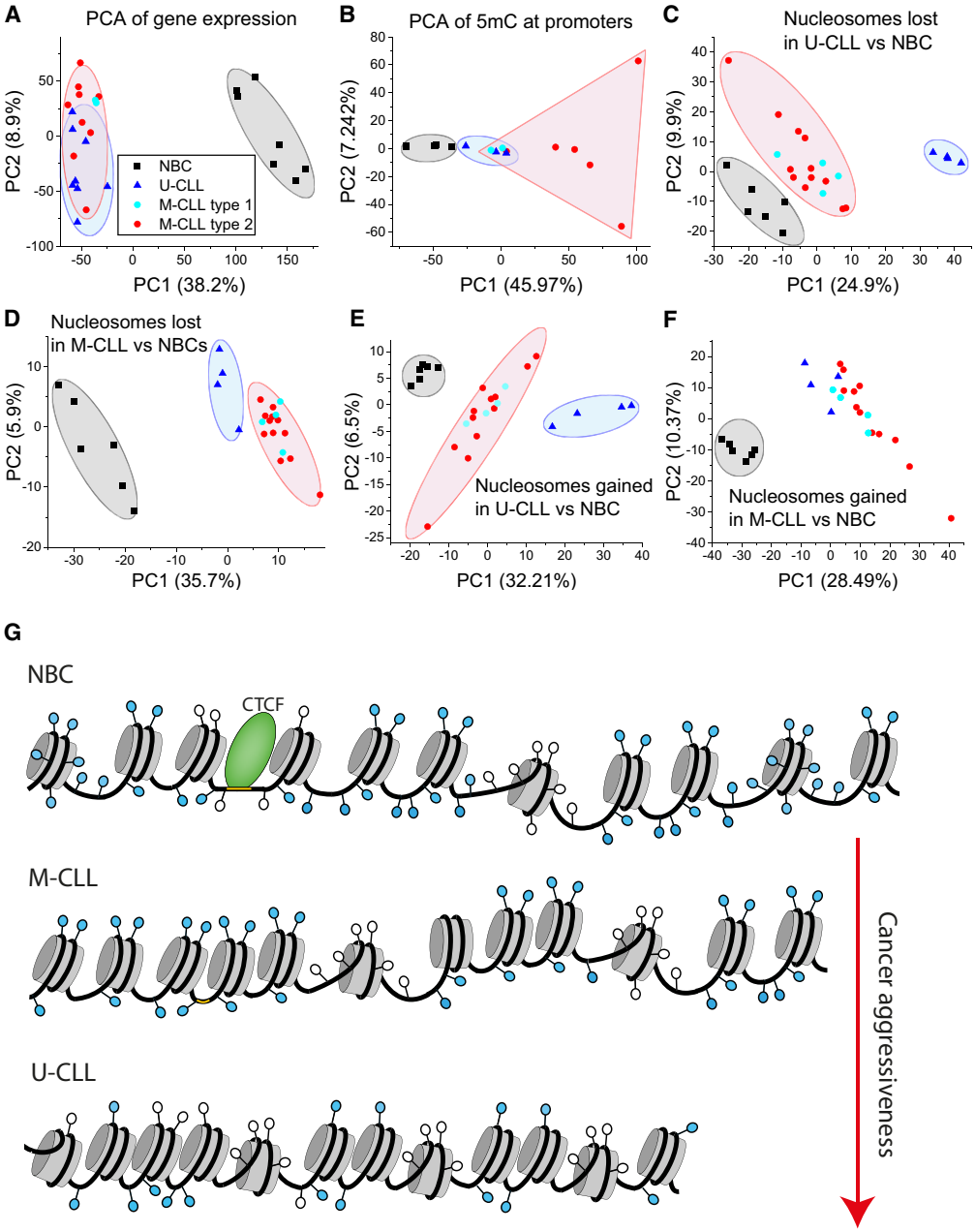


**Figure 7.** Nucleosome positioning as a new marker to stratify CLL patients. Panels *A* to *F* show principal component analysis (PCA) based on gene expression, DNA methylation, and nucleosome occupancy. Each dot represents one sample from NBC (black squares), M-CLL (red circles), and U-CLL (blue triangles). (*A*) PCA based on gene expression. (*B*) PCA based on 5mC at promoters. (*C*) PCA based on nucleosome occupancy at regions with reduced nucleosome occupancy in U-CLL versus NBC. (*D*) Same as in panel *C* but for regions with reduced nucleosome occupancy in M-CLL versus NBC. (*E*) Same as panel *C* but for regions with increased nucleosome occupancy. (*F*) Same as panel *D* but for regions with increased nucleosome occupancy. (*G*) A scheme of molecular mechanisms of nucleosome repositioning in CLL. CLL B cells have shorter NRL and more partially unwrapped nucleosomes than NBCs. These features are linked to differential DNA methylation, rearrangement of CTCF and other TFs, and active chromatin remodeling. The aberrant nucleosome positioning in CLL is more pronounced for the more aggressive U-CLL subtype compared to M-CLL.

occupancy profiles between CLL and NBCs, such as a depletion of nucleosome occupancy in NBCs changed to a peak in CLL (e.g., JUND, PKNOX2, and FOSL1 in Fig. 6C). PKNOX2 acts as tumor suppressor in gastric cancer via activation of IGFBP5 and TP53 (Zhang et al. 2019). JUND is involved in diffuse large B cell lymphoma (Papoudou-Bai et al. 2016), whereas FOSL1 was reported to promote carcinogenesis and metastasis in various cancer types (Jiang et al. 2020; Zhang et al. 2021), but not in CLL. The *FOSL1* gene encodes the FRA1 antigen (Jiang et al. 2020), which forms an AP-1 complex with protein of the JUN family to exert its oncogenic effect. Thus, it is likely no coincidence that both FOSL1 and JUND changed the shape of nucleosome occupancy profiles at their binding sites. Gene expression of *FOSL1* decreased approximately twofold in CLL versus NBCs, which may explain why

many of the binding sites occupied by FOSL1 in NBCs become occupied by a nucleosome in CLL. It is also worth noting that AP-1 is known to interact with SWI/SNF chromatin remodelers (Vierbuchen et al. 2017), thus providing a possible connection to the involvement of BRG1 in active nucleosome repositioning (Fig. 2I,J).

Finally, our study suggests that nucleosome positioning represents a novel biomarker for CLL patient stratification. Nucleosome occupancies in regions that lost or gained nucleosomes in U-CLL versus NBCs allowed it to distinguish U-CLL, M-CLL, and NBCs (Fig. 7C–F), whereas a similar analysis based on gene expression or DNA methylation was less effective (Fig. 7A,B). In part, this can be explained by the fact that CLL-specific gene expression occurring in lymph nodes as the primary site of CLL proliferation is less stably retained in peripheral than the nucleosome positioning associated with CLL in the lymph node microenvironment (Herishanu et al. 2011). In addition, using DNA methylation as a marker for the CLL disease state is confounded by the developmental trajectory of individual CLL cases (Fig. 7B), whereas nucleosome occupancy–based analysis was able to resolve this issue, for example, when using subsets of regions that lost nucleosomes in CLL. Thus, given the extensive heterogeneity in CLL, the analysis of nucleosome positioning could enhance our understanding of critical malignant transformation events that are indicative of CLL subtype, cell of origin, and microenvironmental cues.

In summary, nucleosome positioning undergoes large-scale changes in malignant B cells in CLL, as reflected by the NRL decrease, in addition to specific local changes at functional genomic elements. This makes the analysis of aberrant nucleosome positions in primary tumor cells a powerful approach for distinguishing malignant subtypes for patient stratification that, at the same time, informs about associated deregulation mechanisms. One particular area of application of the nucleosome positioning framework described here is patient diagnostics and monitoring based on liquid biopsies of peripheral blood samples. Identifying nucleosome positions from the analysis of cfDNA and comparing them with cancer-specific nucleosome signatures in cells of origin, as derived here for CLL or for breast cancer in our recent study (Jacob et al. 2023), could provide a novel approach to dissect the disease course from liquid biopsies. Thus, mapping changes in nucleosome positioning in tumor versus nonmalignant cells provides a wealth of information that could be exploited for various applications in personalized precision oncology.

## Methods

### Generation of normalized nucleosome occupancy profiles

The MNase-assisted histone H3 ChIP-seq reads in NBCs, M-CLL, and U-CLL were from our previous study and accessed from the European Genome-phenome Archive (EGA; https://ega-archive.org) under accession number EGAS00001002518 (Mallm et al. 2019). The CLL samples studied here were from asymptomatic patients under a watch-and-wait treatment strategy with high leukocyte counts of $101 \times 10^9$ $L^{-1}$ (median) and with no significant biases between U-CLL and M-CLL samples. The leukocyte count of the healthy donors was in the range of $4.5 \times 10^9$ $L^{-1}$ to $11 \times 10^9$ $L^{-1}$. Thus, for the purpose of our analysis, the potential contribution of NBCs in the CLL samples can be neglected. Paired-end sequenced reads were mapped with Bowtie (Langmead et al. 2009) to human genome hg19, allowing up to one mismatch and retaining only uniquely mappable reads. This resulted in about 150 million paired reads/sample for a total of 26 samples (six NBCs, 16 M-CLL, and four U-CLL) (Supplemental Table S3). The cfDNAtools script extract_nuc_sizes.pl (see Data access) was used to extract fragments of 120–180 bp in size, after which we obtained, on average, 105 million paired reads/sample. NucTools (Vainshtein et al. 2017) was used to calculate genome-wide nucleosome occupancies per chromosome with single-base-pair resolution. This analysis workflow is described in further detail below and includes a normalization of nucleosome occupancy per sample per chromosome for sequencing depth. Averaged profiles for a given sample were calculated based on the normalized profiles of individual patients using NucTools with a 100-bp running window. The resulting genome-wide nucleosome occupancy maps for the complete sample set derived here are available at the NCBI Gene Expression Omnibus (GEO; https://www.ncbi.nlm.nih.gov/geo/) under accession number GSE158745. Because we limited our analysis to uniquely mappable regions, realigning these data to a newer human genome assembly would not affect the major conclusions of this study. As described previously, our experimental procedures included counting the numbers of cells in each sample to have the same amount of input material for chromatin digestion and to reach the same degree of MNase digestion for the sample set studied (Mallm et al. 2019).

### Identification of nucleosomes with changed location in CLL

The annotation as a stable/shifted nucleosome refers to DNA fragments of 120–180 bp size that had genomic coordinates overlapping >95% between all replicates of all conditions ("stable") or <80% overlapping between NBCs and CLL ("shifted"). Only nucleosomes with stable positions across all NBC samples (overlapped >95% between all NBC replicates) were included in this analysis. This set of nucleosomes was determined by performing a pairwise intersection between the corresponding samples using BEDTools with the parameters -u -f 0.95. Coordinates of nucleosomes that were shifted >20% were obtained by intersection with the parameters -f 0.80 -r -v. Aggregate methylation profiles at such nucleosomes were calculated using NucTools considering only CpGs with beta-values of 0.8 or more, which were selected using the script methylationThresholds.pl (see Data access).

### Identification of genomic regions of differential nucleosome occupancy in CLL

Differences in nucleosome occupancy between NBCs, M-CLL, and U-CLL were identified with NucTools. First, we averaged nucleosome occupancy values genome-wide with a sliding window of 100 bp or 1000 bp (as detailed below) separately for each sample. Next, we defined regions that have stable nucleosome occupancy within each condition. This step included the normalization of nucleosome occupancy profiles by the sequencing depth per chromosome per sample. Normalized profiles in each 100-bp window were averaged across all samples in a given condition (six NBCs, 16 M-CLL, four U-CLL). Genomic regions with a relative error of the averaged nucleosome occupancy <0.5 were annotated as "stable-nucleosome" regions. The stable-nucleosome regions were used to perform pairwise occupancy comparisons between M-CLL and NBCs, U-CLL and NBCs, and U-CLL and M-CLL with the NucTools script compare_two_conditions.pl. This script calculates the relative occupancy change ($O_{diff}$) as $O_{diff} = 2 \times (\langle O_{N1} \rangle - \langle O_{N2} \rangle)/(\langle O_{N1} \rangle + \langle O_{N2} \rangle)$. The parameter $\langle O_{N1} \rangle$ corresponds to the averaged nucleosome occupancy in data set 1, and $\langle O_{N2} \rangle$ is the average occupancy in data set 2. Regions that lost or gained nucleosomes between conditions are defined as those in which the relative occupancy change exceeds a given threshold. A 1000-bp sliding window and $O_{diff} = 0.70$ were used to calculate relative

occupancy changes for the comparison of U-CLL versus NBC and M-CLL versus NBC reported in Figure 7 and Supplemental Figures S13 and S14. For all other calculations, a 100-bp sliding window size was used with $O_{diff} = 0.99$ for U-CLL versus NBC and M-CLL versus U-CLL and $O_{diff} = 0.70$ for M-CLL versus NBC due to the lower number of regions reported.

### Enrichment analysis

Fold enrichment of genomic features in lost/gain-nucleosome regions was calculated using BEDTools "fisher" function (Quinlan 2014). All results were statistically significant based on the two-tailed Fisher's test, $P < 0.05$. Pathway enrichment analysis was performed in DAVID 6.8 (Jiao et al. 2012) and gprofiler2 (Raudvere et al. 2019) and visualized with Origin Pro (OriginLab). The visualization of TF networks based on binding sites with differential nucleosome occupancy in M-CLL and U-CLL was performed with Pathway Commons Network Visualizer (Rodchenkov et al. 2020).

### NRL analysis

NRL was calculated with NucTools following our protocol described previously (Vainshtein et al. 2017) but discarding the first peak of the nucleosome–nucleosome distance distributions to make the calculation more robust against the level of MNase digestion. The standard error of the NRL estimation was <1 bp for all samples.

### CTCF binding analysis

CTCF binding sites inside CTCF ChIP-seq peaks were predicted with GimmeMotifs (van Heeringen and Veenstra 2011) using the weight matrix MA0139.1 from JASPAR (Khan et al. 2018) and accepting motifs with a similarity score >80%. These sites were intersected using BEDTools with the experimentally determined CTCF ChIP-seq peaks for each sample from GEO accession number GSE113336 to obtain CTCF binding sites specific to NBCs, M-CLL, and U-CLL. CTCF sites present in both NBCs and B cells from all CLL patients were termed "common," and sites present only in NBCs but not in CLL were termed "lost"; 41,202 CTCF-bound ChIP-seq peaks observed both in NBCs and CLL were defined as common, which contained 43,754 unique CTCF motifs. In addition, 718 CTCF-bound ChIP-seq peaks that were lost in CLL versus NBCs contained 708 CTCF motifs; 430 peaks contained 124 motifs that gained CTCF binding in M-CLL versus NBCs; and 369 peaks contained 255 motifs that gained CTCF binding in U-CLL versus NBCs. Furthermore, the computationally predicted CTCF motifs were split into three quantiles based on the motif similarity determined with TFBStools (Tan and Lenhard 2016). Sites with a similarity score of 80%–81% were defined as quantile 1 CTCF motifs; a score of 81%–83% as quantile 2 CTCF motifs; and a score >83% as quantile 3.

### Calculation of aggregate profiles of nucleosome occupancy and histone modifications

Calculation of aggregate nucleosome profiles was performed with HOMER (Heinz et al. 2010), using only DNA fragments with sizes of 120–180 bp unless specified otherwise in the text. Aggregate profiles of seven histone modifications measured with ChIP-seq for this cohort (Mallm et al. 2019) were calculated using HOMER considering all ChIP-seq reads uniquely mapped with Bowtie to hg19 genome with up to one mismatch. Calculation of the aggregate methylation profiles was performed using custom Perl scripts as detailed below.

### Calculation of aggregate DNA methylation profiles

We calculated two types of DNA methylation profiles. In the first type of analysis (Figs. 2A,B, 2E–H, 5A,B, 7B; Supplemental Figs. S4, S6, S8), CpGs with beta-values higher than 0.8 for each sample were extracted using the script methylationThresholds.pl (see Data access) and split into chromosomes using NucTools script extract_chr_bed.pl followed by NucTools script bed2occupancy_average.pl to calculate DNA methylation density arising from such CpGs genome-wide with a 1-bp sliding window. This was then used for calculating occupancy profiles around genomic features with NucTools script aggregate_profile.pl. In this type of calculation, each CpG with a beta-value greater than 0.8 contributes equally. In the second type of calculation (Fig. 4A–D), DNA methylation data reported previously for this cohort (Mallm et al. 2019) were processed with the Perl script bed2occupancy.v3d.methyl.pl (see Data access) as described previously (Wiehle et al. 2019). The actual methylation beta-values with values from zero to one were added up for each individual CpG located at a given distance from the genomic feature of interest. The value of "DNA methylation (a.u.)" in the plots corresponds to the sum of all corresponding beta-values without further normalization. DNA methylation profiles of individual example regions were calculated considering all methylation beta-values (Fig. 1D; Supplemental Fig. S3).

### Principal component analysis

DNA methylation and nucleosome occupancy values averaged over a 100-bp window for each sample were intersected with coordinates of genomic features of interest (e.g., promoters or regions that lost/gained nucleosome occupancy in CLL) to create a matrix of occupancy values at the specified genomic regions, for each sample. PCA was performed in R (R Core Team 2023), and the relationship of principal components 1 and 2 was visualized in Origin Pro (OriginLab).

### Annotation of genomic features

The coordinates of the 100,000-bp genomic regions, annotated as A and B chromatin compartments in NBCs from peripheral blood, were kindly provided by Vilarrasa-Blasi et al. (2021) as BED files in hg38 genome assembly. This included 8300 A and 8786 B compartments in NBCs. The liftOver utility from the UCSC Genome Browser (Kent et al. 2002) was used to convert the coordinates from hg38 to hg19 with default settings. Conversion failed on 48 and 35 A and B compartment records, respectively. Promoters were defined as regions ±1000 bp around the TSS, based on RefSeq annotation for hg19. Enhancers specific to NBCs and CLL were defined based on the same cohort following the method of our previous report (Mallm et al. 2019). Binding sites of SP1 and CHD1 were determined based on ChIP-seq peaks in lymphoblastoid cell line GM12878 reported by the ENCODE Consortium (GEO accession numbers GSM803363 and GSM935301 correspondingly) (The ENCODE Project Consortium 2012). Binding sites of HDAC1 in peripheral blood mononuclear cells from CLL patients with the 11q deletion genotype were determined based on ChIP-seq peaks downloaded from GEO accession number GSE216287 (Lai et al. 2023). Enrichment of BRG1 was calculated using ChIP-seq of BRG1 binding in naïve B cells J1 (GSM971343) (Abraham et al. 2013). For the latter, we mapped the raw data to the hg19 genome, calculated aggregate profiles with HOMER, and then normalized these by dividing the BRG1 ChIP-seq signal by the corresponding Input reported by the investigators.

## Data access

The processed sequencing data generated in this study have been submitted to the NCBI Gene Expression Omnibus (GEO; https://www.ncbi.nlm.nih.gov/geo/) under accession number GSE158745. The raw sequencing data, including MNase-assisted H3 ChIP-seq, have been submitted to the European Genome-phenome Archive (EGA; https://ega-archive.org) under accession number EGAS00001002518. All custom scripts generated in this study are available at GitHub (https://github.com/TeifLab/cfDNAtools) and as Supplemental Scripts.

## Competing interest statement

The authors declare no competing interests.

## Acknowledgments

This work was supported by Wellcome Trust grant 200733/Z/16/Z and Cancer Research UK grant EDDPMA-Nov21\100044 to V.B.T.; Genetics Society grants to K.V.P. and L.R.; DFG grants for subprojects B1, B2, and Z1 within SFB1074 to D.M., S.S., and K.R.; and the Leukaemia & Lymphoma NI R2740CEM and the Academy of Medical Sciences SBF005\1113 grants to E.K. V.B.T. thanks Victor Zhurkin for fruitful discussions and Stuart Newman for the computer cluster support.

*Author contributions:* K.V.P. performed analyses related to CTCF, regulatory regions, remodelers, A/B compartments, histone modifications, and DNA methylation, as well as PCA analysis for patient stratification. C.M.D., C.X., C.T.C., and C.F. performed analyses related to TF binding. J.P.M. contributed to the data preprocessing. L.R. performed data visualization. L.C.K. performed patient classification based on B cell developmental status. Y.V. maintained and upgraded the NucTools software. V.B.T. performed initial data analyses leading to this manuscript and NRL calculations. S.S. and D.M. advised on medical aspects. E.K., K.R., and V.B.T. supervised the study. V.B.T. and K.R. wrote the manuscript draft. All authors read and approved the final manuscript.

## References

Abraham BJ, Cui K, Tang Q, Zhao K. 2013. Dynamic regulation of epigenomic landscapes during hematopoiesis. *BMC Genomics* **14:** 193. doi:10.1186/1471-2164-14-193

Andrades A, Peinado P, Alvarez-Perez JC, Sanjuan-Hidalgo J, García DJ, Arenas AM, Matia-González AM, Medina PP. 2023. SWI/SNF complexes in hematological malignancies: biological implications and therapeutic opportunities. *Mol Cancer* **22:** 39. doi:10.1186/s12943-023-01736-8

Balatti V, Tomasello L, Rassenti LZ, Veneziano D, Nigita G, Wang HY, Thorson JA, Kipps TJ, Pekarsky Y, Croce CM. 2018. *miR-125a* and *miR-34a* expression predicts Richter syndrome in chronic lymphocytic leukemia patients. *Blood* **132:** 2179–2182. doi:10.1182/blood-2018-04-845115

Beekman R, Chapaprieta V, Russiñol N, Vilarrasa-Blasi R, Verdaguer-Dot N, Martens JHA, Duran-Ferrer M, Kulis M, Serra F, Javierre BM, et al. 2018. The reference epigenome and regulatory chromatin landscape of chronic lymphocytic leukemia. *Nat Med* **24:** 868–880. doi:10.1038/s41591-018-0028-4

Bosch F, Dalla-Favera R. 2019. Chronic lymphocytic leukaemia: from genetics to treatment. *Nat Rev Clin Oncol* **16:** 684–701. doi:10.1038/s41571-019-0239-8

Capello D, Fais F, Vivenza D, Migliaretti G, Chiorazzi N, Gaidano G, Ferrarini M. 2000. Identification of three subgroups of B cell chronic lymphocytic leukemia based upon mutations of *BCL-6* and IgV genes. *Leukemia* **14:** 811–815. doi:10.1038/sj.leu.2401766

Diermeier S, Kolovos P, Heizinger L, Schwartz U, Georgomanolis T, Zirkel A, Wedemann G, Grosveld F, Knoch TA, Merkl R, et al. 2014. TNFα signalling primes chromatin for NF-κB binding and induces rapid and widespread nucleosome repositioning. *Genome Biol* **15:** 536. doi:10.1186/s13059-014-0536-6

Domcke S, Bardet AF, Adrian Ginno P, Hartl D, Burger L, Schübeler D. 2015. Competition between DNA methylation and transcription factors determines binding of NRF1. *Nature* **528:** 575–579. doi:10.1038/nature16462

The ENCODE Project Consortium. 2012. An integrated encyclopedia of DNA elements in the human genome. *Nature* **489:** 57–74. doi:10.1038/nature11247

Fang C, Wang Z, Han C, Safgren SL, Helmin KA, Adelman ER, Serafin V, Basso G, Eagen KP, Gaspar-Maia A, et al. 2020. Cancer-specific CTCF binding facilitates oncogenic transcriptional dysregulation. *Genome Biol* **21:** 247. doi:10.1186/s13059-020-02152-7

Ferreira PG, Jares P, Rico D, Gómez-López G, Martínez-Trillos A, Villamor N, Ecker S, González-Pérez A, Knowles DG, Monlong J, et al. 2014. Transcriptome characterization by RNA sequencing identifies a major molecular and clinical subdivision in chronic lymphocytic leukemia. *Genome Res* **24:** 212–226. doi:10.1101/gr.152132.112

Gaffney DJ, McVicker G, Pai AA, Fondufe-Mittendorf YN, Lewellen N, Michelini K, Widom J, Gilad Y, Pritchard JK. 2012. Controls of nucleosome positioning in the human genome. *PLoS Genet* **8:** e1003036. doi:10.1371/journal.pgen.1003036

Gaiti F, Chaligne R, Gu H, Brand RM, Kothen-Hill S, Schulman RC, Grigorev K, Risso D, Kim KT, Pastore A, et al. 2019. Epigenetic evolution and lineage histories of chronic lymphocytic leukaemia. *Nature* **569:** 576–580. doi:10.1038/s41586-019-1198-z

Grandi FC, Modi H, Kampman L, Corces MR. 2022. Chromatin accessibility profiling by ATAC-seq. *Nat Protoc* **17:** 1518–1552. doi:10.1038/s41596-022-00692-9

Heinz S, Benner C, Spann N, Bertolino E, Lin YC, Laslo P, Cheng JX, Murre C, Singh H, Glass CK. 2010. Simple combinations of lineage-determining transcription factors prime *cis*-regulatory elements required for macrophage and B cell identities. *Mol Cell* **38:** 576–589. doi:10.1016/j.molcel.2010.05.004

Herishanu Y, Pérez-Galán P, Liu D, Biancotto A, Pittaluga S, Vire B, Gibellini F, Njuguna N, Lee E, Stennett L, et al. 2011. The lymph node microenvironment promotes B-cell receptor signaling, NF-κB activation, and tumor proliferation in chronic lymphocytic leukemia. *Blood* **117:** 563–574. doi:10.1182/blood-2010-05-284984

Ho JW, Jung YL, Liu T, Alver BH, Lee S, Ikegami K, Sohn KA, Minoda A, Tolstorukov MY, Appert A, et al. 2014. Comparative analysis of metazoan chromatin organization. *Nature* **512:** 449–452. doi:10.1038/nature13415

Jacob DR, Guiblet WM, Mamayusupova H, Shtumpf M, Ciuta I, Ruje L, Gretton S, Bikova M, Correa C, Dellow E, et al. 2023. Nucleosome reorganisation in breast cancer tissues. bioRxiv doi:10.1101/2023.04.17.537031

Jeong H, Grimes K, Rauwolf KK, Bruch PM, Rausch T, Hasenfeld P, Benito E, Roider T, Sabarinathan R, Porubsky D, et al. 2023. Functional analysis of structural variants in single cells using strand-seq. *Nat Biotechnol* **41:** 832–844. doi:10.1038/s41587-022-01551-4

Jiang X, Xie H, Dou Y, Yuan J, Zeng D, Xiao S. 2020. Expression and function of FRA1 protein in tumors. *Mol Biol Rep* **47:** 737–752. doi:10.1007/s11033-019-05123-9

Jiao X, Sherman BT, Huang da W, Stephens R, Baseler MW, Lane HC, Lempicki RA. 2012. DAVID-WS: a stateful web service to facilitate gene/protein list analysis. *Bioinformatics* **28:** 1805–1806. doi:10.1093/bioinformatics/bts251

Kent WJ, Sugnet CW, Furey TS, Roskin KM, Pringle TH, Zahler AM, Haussler D. 2002. The human genome browser at UCSC. *Genome Res* **12:** 996–1006. doi:10.1101/gr.229102

Khan A, Fornes O, Stigliani A, Gheorghe M, Castro-Mondragon JA, van der Lee R, Bessy A, Chèneby J, Kulkarni SR, Tan G, et al. 2018. JASPAR 2018: update of the open-access database of transcription factor binding profiles and its web framework. *Nucleic Acids Res* **46:** D1284. doi:10.1093/nar/gkx1188

Klintman J, Appleby N, Stamatopoulos B, Ridout K, Eyre TA, Robbe P, Pascua LL, Knight SJL, Dreau H, Cabes M, et al. 2021. Genomic and transcriptomic correlates of Richter transformation in chronic lymphocytic leukemia. *Blood* **137:** 2800–2816. doi:10.1182/blood.2020005650

Kulis M, Merkel A, Heath S, Queirós AC, Schuyler RP, Castellano G, Beekman R, Raineri E, Esteve A, Clot G, et al. 2015. Whole-genome fingerprint of the DNA methylome during human B cell differentiation. *Nat Genet* **47:** 746–756. doi:10.1038/ng.3291

Kundaje A, Kyriazopoulou-Panagiotopoulou S, Libbrecht M, Smith CL, Raha D, Winters EE, Johnson SM, Snyder M, Batzoglou S, Sidow A. 2012. Ubiquitous heterogeneity and asymmetry of the chromatin environment at regulatory elements. *Genome Res* **22:** 1735–1747. doi:10.1101/gr.136366.111

Lai AY, Wade PA. 2011. Cancer biology and NuRD: a multifaceted chromatin remodelling complex. *Nat Rev Cancer* **11:** 588–596. doi:10.1038/nrc3091

Lai TH, Ozer HG, Gasparini P, Nigita G, Distefano R, Yu L, Ravikrishnan J, Yilmaz S, Gallegos JJ, Shukla SA, et al. 2023. HDAC1 regulates the chromatin landscape to control transcriptional dependencies in chronic lymphocytic leukemia. *Blood Adv* **7:** 2897–2911. doi:10.1182/bloodadvances.2022007998

Landau DA, Tausch E, Taylor-Weiner AN, Stewart C, Reiter JG, Bahlo J, Kluth S, Bozic I, Lawrence M, Böttcher S, et al. 2015. Mutations driving CLL and their evolution in progression and relapse. *Nature* **526:** 525–530. doi:10.1038/nature15395

Langmead B, Trapnell C, Pop M, Salzberg SL. 2009. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol* **10:** R25. doi:10.1186/gb-2009-10-3-r25

Lo YMD, Han DSC, Jiang P, Chiu RWK. 2021. Epigenetics, fragmentomics, and topology of cell-free DNA in liquid biopsies. *Science* **372:** eaaw3616. doi:10.1126/science.aaw3616

Mallm JP, Iskar M, Ishaque N, Klett LC, Kugler SJ, Muino JM, Teif VB, Poos AM, Großmann S, Erdel F, et al. 2019. Linking aberrant chromatin features in chronic lymphocytic leukemia to transcription factor networks. *Mol Syst Biol* **15:** e8339. doi:10.15252/msb.20188339

Mlynarczyk C, Fontán L, Melnick A. 2019. Germinal center-derived lymphomas: the darkest side of humoral immunity. *Immunol Rev* **288:** 214–239. doi:10.1111/imr.12755

Nabhan C, Rosen ST. 2014. Chronic lymphocytic leukemia: a clinical review. *JAMA* **312:** 2265–2276. doi:10.1001/jama.2014.14553

Najm FJ, DeWeirdt P, Moore MM, Bevill SM, El Farran CA, Macias KA, Hegde M, Waterbury AL, Liau BB, van Galen P, et al. 2023. Chromatin complex dependencies reveal targeting opportunities in leukemia. *Nat Commun* **14:** 448. doi:10.1038/s41467-023-36150-7

Oakes CC, Seifert M, Assenov Y, Gu L, Przekopowitz M, Ruppert AS, Wang Q, Imbusch CD, Serva A, Koser SD, et al. 2016. DNA methylation dynamics during B cell maturation underlie a continuum of disease phenotypes in chronic lymphocytic leukemia. *Nat Genet* **48:** 253–264. doi:10.1038/ng.3488

Ott CJ, Federation AJ, Schwartz LS, Kasar S, Klitgaard JL, Lenci R, Li Q, Lawlor M, Fernandes SM, Souza A, et al. 2018. Enhancer architecture and essential core regulatory circuitry of chronic lymphocytic leukemia. *Cancer Cell* **34:** 982–995.e7. doi:10.1016/j.ccell.2018.11.001

Papoudou-Bai A, Goussia A, Batistatou A, Stefanou D, Malamou-Mitsi V, Kanavaros P. 2016. The expression levels of JunB, JunD and p-c-Jun are positively correlated with tumor cell proliferation in diffuse large B-cell lymphomas. *Leuk Lymphoma* **57:** 143–150. doi:10.3109/10428194.2015.1034704

Pastore A, Gaiti F, Lu SX, Brand RM, Kulm S, Chaligne R, Gu H, Huang KY, Stamenova EK, Béguelin W, et al. 2019. Corrupted coordination of epigenetic modifications leads to diverging chromatin states and transcriptional heterogeneity in CLL. *Nat Commun* **10:** 1874. doi:10.1038/s41467-019-09645-5

Puente XS, Beà S, Valdés-Mas R, Villamor N, Gutiérrez-Abril J, Martín-Subero JI, Munar M, Rubio-Pérez C, Jares P, Aymerich M, et al. 2015. Non-coding recurrent mutations in chronic lymphocytic leukaemia. *Nature* **526:** 519–524. doi:10.1038/nature14666

Queirós AC, Villamor N, Clot G, Martinez-Trillos A, Kulis M, Navarro A, Penas EM, Jayne S, Majid A, Richter J, et al. 2015. A B-cell epigenetic signature defines three biologic subgroups of chronic lymphocytic leukemia with clinical impact. *Leukemia* **29:** 598–605. doi:10.1038/leu.2014.252

Quesada V, Ramsay AJ, Rodríguez D, Puente XS, Campo E, López-Otín C. 2013. The genomic landscape of chronic lymphocytic leukemia: clinical implications. *BMC Med* **11:** 124. doi:10.1186/1741-7015-11-124

Quinlan AR. 2014. BEDTools: the Swiss-Army tool for genome feature analysis. *Curr Protoc Bioinformatics* **47:** 1–34. doi:10.1002/0471250953.bi1112s47

Rani L, Mathur N, Gupta R, Gogia A, Kaur G, Dhanjal JK, Sundar D, Kumar L, Sharma A. 2017. Genome-wide DNA methylation profiling integrated with gene expression profiling identifies PAX9 as a novel prognostic marker in chronic lymphocytic leukemia. *Clin Epigenetics* **9:** 57. doi:10.1186/s13148-017-0356-0

Raudvere U, Kolberg L, Kuzmin I, Arak T, Adler P, Peterson H, Vilo J. 2019. g:Profiler: a web server for functional enrichment analysis and conversions of gene lists (2019 update). *Nucleic Acids Res* **47:** W191–W198. doi:10.1093/nar/gkz369

R Core Team. 2023. *R: a language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna. https://www.R-project.org/.

Rendeiro AF, Schmidl C, Strefford JC, Walewska R, Davis Z, Farlik M, Oscier D, Bock C. 2016. Chromatin accessibility maps of chronic lymphocytic leukaemia identify subtype-specific epigenome signatures and transcription regulatory networks. *Nat Commun* **7:** 11938. doi:10.1038/ncomms11938

Rendeiro AF, Krausgruber T, Fortelny N, Zhao F, Penz T, Farlik M, Schuster LC, Nemc A, Tasnády S, Réti M, et al. 2020. Chromatin mapping and single-cell immune profiling define the temporal dynamics of ibrutinib response in CLL. *Nat Commun* **11:** 577. doi:10.1038/s41467-019-14081-6

Rodchenkov I, Babur O, Luna A, Aksoy BA, Wong JV, Fong D, Franz M, Siper MC, Cheung M, Wrana M, et al. 2020. Pathway Commons 2019 update: integration, analysis and exploration of pathway data. *Nucleic Acids Res* **48:** D489–D497.

Rodríguez D, Bretones G, Quesada V, Villamor N, Arango JR, López-Guillermo A, Ramsay AJ, Baumann T, Quirós PM, Navarro A, et al. 2015. Mutations in CHD2 cause defective association with active chromatin in chronic lymphocytic leukemia. *Blood* **126:** 195–202. doi:10.1182/blood-2014-10-604959

Roisman A, Stanganelli C, Nagore VP, Richardson GV, Scassa ME, Bezares RF, Cabrejo M, Slavutsky I. 2015. *SOX11* expression in chronic lymphocytic leukemia correlates with adverse prognostic markers. *Tumour Biol* **36:** 4433–4440. doi:10.1007/s13277-015-3083-1

Sarkar A, Hochedlinger K. 2013. The Sox family of transcription factors: versatile regulators of stem and progenitor cell fate. *Cell Stem Cell* **12:** 15–30. doi:10.1016/j.stem.2012.12.007

Schones DE, Cui K, Cuddapah S, Roh TY, Barski A, Wang Z, Wei G, Zhao K. 2008. Dynamic regulation of nucleosome positioning in the human genome. *Cell* **132:** 887–898. doi:10.1016/j.cell.2008.02.022

Shtumpf M, Piroeva KV, Agrawal SP, Jacob DR, Teif VB. 2022. NucPosDB: a database of nucleosome positioning in vivo and nucleosomics of cell-free DNA. *Chromosoma* **131:** 19–28. doi:10.1007/s00412-021-00766-9

Snyder MW, Kircher M, Hill AJ, Daza RM, Shendure J. 2016. Cell-free DNA comprises an in vivo nucleosome footprint that informs its tissues-of-origin. *Cell* **164:** 57–68. doi:10.1016/j.cell.2015.11.050

Tan G, Lenhard B. 2016. TFBSTools: an R/Bioconductor package for transcription factor binding site analysis. *Bioinformatics* **32:** 1555–1556. doi:10.1093/bioinformatics/btw024

Teif VB, Clarkson CT. 2019. Nucleosome positioning. In *Encyclopedia of bioinformatics and computational biology* (ed. Ranganathan S, et al.), pp. 308–317. Academic Press, Oxford.

Teif VB, Vainshtein Y, Caudron-Herger M, Mallm JP, Marth C, Höfer T, Rippe K. 2012. Genome-wide nucleosome positioning during embryonic stem cell development. *Nat Struct Mol Biol* **19:** 1185–1192. doi:10.1038/nsmb.2419

Teif VB, Beshnova DA, Vainshtein Y, Marth C, Mallm JP, Höfer T, Rippe K. 2014. Nucleosome repositioning links DNA (de)methylation and differential CTCF binding during stem cell development. *Genome Res* **24:** 1285–1295. doi:10.1101/gr.164418.113

Teif VB, Mallm JP, Sharma T, Mark Welch DB, Rippe K, Eils R, Langowski J, Olins AL, Olins DE. 2017. Nucleosome repositioning during differentiation of a human myeloid leukemia cell line. *Nucleus* **8:** 188–204. doi:10.1080/19491034.2017.1295201

Vainshtein Y, Rippe K, Teif VB. 2017. NucTools: analysis of chromatin feature occupancy profiles from high-throughput sequencing data. *BMC Genomics* **18:** 158. doi:10.1186/s12864-017-3580-2

Valouev A, Johnson SM, Boyd SD, Smith CL, Fire AZ, Sidow A. 2011. Determinants of nucleosome organization in primary human cells. *Nature* **474:** 516–520. doi:10.1038/nature10002

van Heeringen SJ, Veenstra GJ. 2011. GimmeMotifs: a *de novo* motif prediction pipeline for ChIP-sequencing experiments. *Bioinformatics* **27:** 270–271. doi:10.1093/bioinformatics/btq636

Vierbuchen T, Ling E, Cowley CJ, Couch CH, Wang X, Harmin DA, Roberts CWM, Greenberg ME. 2017. AP-1 transcription factors and the BAF complex mediate signal-dependent enhancer selection. *Mol Cell* **68:** 1067–1082.e12. doi:10.1016/j.molcel.2017.11.026

Vilarrasa-Blasi R, Soler-Vila P, Verdaguer-Dot N, Russiñol N, Di Stefano M, Chapaprieta V, Clot G, Farabella I, Cuscó P, Kulis M, et al. 2021. Dynamics of genome architecture and chromatin function during human B cell differentiation and neoplastic transformation. *Nat Commun* **12:** 651. doi:10.1038/s41467-020-20849-y

Wiehle L, Thorn GJ, Raddatz G, Clarkson CT, Rippe K, Lyko F, Breiling A, Teif VB. 2019. DNA (de)methylation in embryonic stem cells controls CTCF-dependent chromatin boundaries. *Genome Res* **29:** 750–761. doi:10.1101/gr.239707.118

Zhang L, Li W, Cao L, Xu J, Qian Y, Chen H, Zhang Y, Kang W, Gou H, Wong CC, et al. 2019. PKNOX2 suppresses gastric cancer through the transcriptional activation of IGFBP5 and p53. *Oncogene* **38:** 4590–4604. doi:10.1038/s41388-019-0743-4

Zhang M, Hoyle RG, Ma Z, Sun B, Cai W, Cai H, Xie N, Zhang Y, Hou J, Liu X, et al. 2021. FOSL1 promotes metastasis of head and neck squamous cell carcinoma through super-enhancer-driven transcription program. *Mol Ther* **29:** 2583–2600. doi:10.1016/j.ymthe.2021.03.024

*Received September 8, 2022; accepted in revised form September 7, 2023.*

# Nucleosome repositioning in chronic lymphocytic leukemia

Kristan V. Piroeva, Charlotte McDonald, Charalampos Xanthopoulos, et al.

*Genome Res.* 2023 33: 1649-1661 originally published online September 12, 2023
Access the most recent version at doi:10.1101/gr.277298.122