Integrative epigenomics in chronic lymphocytic leukaemia: biological insights and clinical applications

Marta Kulis[1] & Jose Ignacio Martin-Subero[1,2,3,4]

1. Institut d'Investigacions Biomèdiques August Pi I Sunyer (IDIBAPS), Barcelona, Spain.

2. Institució Catalana de Recerca i Estudis Avançats (ICREA), Barcelona, Spain.

3. Departamento de Fundamentos Clínicos, Universitat de Barcelona, Barcelona, Spain.

4. Centro de Investigación Biomédica en Red de Cáncer (CIBERONC), Madrid, Spain.

Corresponding author:

Jose I. Martin-Subero: imartins@clinic.cat

Keywords: chronic lymphocytic leukaemia, genetics, epigenetics, DNA methylation, chromatin

Summary

Chronic lymphocytic leukaemia (CLL) is not only characterised by driver genetic alterations but by extensive epigenetic changes. Over the last decade, epigenomic studies have described the DNA methylome, chromatin accessibility, histone modifications and the 3D genome architecture of CLL. Beyond its regulatory role, the DNA methylome contains imprints of the cellular origin and proliferative history of CLL cells. These two aspects are strong independent prognostic factors. Integrative analyses of chromatin marks have uncovered novel regulatory elements and altered transcription factor networks as non-genetic means mediating gene deregulation in CLL. Additionally, CLL cells display a disease-specific pattern of 3D genome interactions. From the technological perspective, we are currently witnessing a transition from bulk omics to single cell analyses. This review aims at summarizing the major findings from the epigenomics field as well as providing a prospect of the present and future of single cell analyses in CLL.

Introduction

Chronic lymphocytic leukaemia (CLL) is a frequent neoplasm of CD5+ mature B cells displaying a heterogeneous spectrum of biological features and clinical manifestations (1). CLL is preceded by monoclonal B cell lymphocytosis, which is a frequent condition in the elderly. Once CLL is diagnosed, some patients remain stable for years without any treatment, and even a small proportion spontaneously regresses (2). Others, however, require early treatment, become resistant and end up transforming to a high-grade lymphoma, which is a condition with dismal prognosis called Richter transformation. A variety of genetic features have been described to underlie the clinico-biological heterogeneity of CLL, including immunogenetic characteristics and somatic genetic alterations of the neoplastic clone. Importantly, CLLs can be classified into two subtypes with markedly differential clinical evolution. Those patients whose leukaemic cells have somatic hypermutation (SHM) in the variable region of the immunoglobulin heavy chain (IGHV) locus (M-CLL) show a clearly more benign clinical course than those patients with low levels or no SHM (U-CLL) (3,4). The landscape of genomic alterations in CLL has been thoroughly studied by classical molecular cytogenetics as well as whole-exome and whole-genome next generation sequencing (NGS) approaches (5) (6)(7)(8). Beyond the classical clinically-relevant copy number changes del(13q), del(11q), del(17p) and trisomy 12, which are present in roughly 80% of the patients, NGS has revealed a heterogeneous pattern of mutated genes. Few genes are mutated in 10-20% of the patients (e.g. *SF3B1*, *NOTCH1* and *ATM*) and a large tail of genes are mutated at small frequencies. In spite of the large number of low-frequency genes, mutations can be functionally grouped into several pathways that further underline the different pathogenetic mechanisms that can give rise to CLL. These include B-cell receptor (BCR) signalling, cell cycle regulation, apoptosis, DNA damage response, chromatin remodelling, NF-κB signalling, NOTCH1 signalling, and RNA processing/export(6).

In addition to the undisputed importance of genetic alterations in the pathophysiology and clinical management of CLL, leukaemic cells also show widespread epigenetic alterations (9,10). Epigenetics encodes a variety of molecular features that regulate gene expression and provides function to genomic information (11) (Figure 1A).

Cytosine methylation at CpG dinucleotides is perhaps the most widely studied mark, and can be measured genome-wide by high-density microarrays of bisulfite-based NGS. This mark has been classically defined as a repressive mark, although its functions are broader and genomic-context dependent(12)(13). An important layer of epigenetic information is coded in the post-translational modifications of histones, which can be measured at the whole-genome scale by chromatin immunoprecipitation followed by NGS (ChIP-seq) with specific antibodies. Small chemical changes in the histone tail aminoacid residues, such as methylation and acetylation can produce significant changes in the properties of the chromatin. To mention some, monomethylated lysine 4 of histone 3 (H3K4me1) is related to enhancer elements whereas trimethylation of the same aminoacid (H3K4me3) is linked to promoters. If any of these regulatory elements acquires acetylation in e.g. lysine 27 of histone 3 (H3K27ac), they will become active. Trimethylation of lysine 36 of H3 (H3K36me3) is associated with transcriptional elongation, and other histone marks are related to polycomb-based repression (H3K27me3) or long-term repression (H3K9me3). Most importantly, different combinations of histone modifications provide the basis to segment the genome into chromatin states (14) (Figure 1B). For instance, if a region contains both H3K4me3 and H3K27me3, it will be classified as a poised promoter, whereas if a region containing H3K4me3 is combined with H3K27ac, it will be an active promoter. Additionally, a large fraction of the genome is densely packed and lacks any of these histone modifications. This low-signal chromatin state is also a form of heterochromatin. Some regions of the genome lack nucleosomes and are accessible to non-histonic DNA binding proteins, such as transcription factors (TF) (15). The pattern of chromatin accessibility can be measured by specific assays such as DNase I hypersensitive sites sequencing (DNAase-seq) or Assay for Transposase-Accessible Chromatin using sequencing (ATAC-seq). Finally, chromatin creates a complex three-dimensional (3D) network within the interphase cell nucleus, which can be characterised by a variety of methods collectively known as chromosome conformation capture techniques (16,17). This 3D genome architecture can be characterised at different resolution levels, such as chromosome territories, chromosome compartments (A for active and B for inactive), topologically-associating domains, and chromatin loops between for instance distant enhancers and promoters.

Over the last decade, several epigenomic studies have generated whole-genome maps of DNA methylation, various chromatin features and 3D genome interactions in CLL and normal B cell differentiation. The goal of this review will be to distil the major findings derived from these studies and how they have contributed to decipher the complexity of CLL, with fresh insights into the cellular origin, pathogenesis, evolution and clinical behaviour.

Epigenetic insights into the cellular origin of CLL

The cellular origin of CLL has been the subject of intense research and debate, as different approaches can lead to apparently different interpretations (18,19). Immunogenetic studies suggest that M-CLLs are derived from germinal center-experienced B cells whereas U-CLLs have matured independently of the germinal center (GC) reaction. Additionally, the existence of subsets of patients with (quasi)identical BcR IG indicates that antigen selection plays a role in the clonal expansion of individual cells that will give rise to CLL (20). Transcriptional studies suggest that U-CLL derives from unmutated mature CD5+ B cells and M-CLL derives from a low abundant subpopulation of CD5+/CD27+ post-germinal center B cells (21), although a previous study interpreted that CLL cells show an overall memory B cell-like transcriptome (22), which is congruent with the immunogenetic studies indicating that CLLs are antigen experienced.

From an epigenetic perspective, DNA methylation analyses have much to add to this discussion. First, normal B cells extensively modify their DNA methylome during their maturation process and each differentiation stage has a specific epigenetic fingerprint (23)(24); second, DNA methylation has been associated with cellular memory(25); and third, cancer cells maintain an epigenetic imprint of their cellular origin (26). In CLL, an initial genome-wide DNA methylation study identified that the majority of the differences between U-CLL and M-CLLs could be attributed to differences between naïve B cells (NBC) and memory B cells (MBC) (27). Furthermore, a more profound exploration of the data revealed the presence of another epigenetic subtype of CLL with an intermediate profile between NBC and MBC and moderate SHM. These

subtypes were named NBC-like CLL (n-CLL), MBC-like CLL (m-CLL) and intermediate CLL (i-CLL). Interestingly, the new i-CLL subtype was strongly enriched for patients with stereotyped BcR IG subset #2 and *SF3B1* mutations. A subsequent independent study used a different analytic approach and further confirmed the existence of these 3 clinico-biological CLL subtypes, which were named low-programmed, intermediate-programmed and high-programmed to reflect the relationship of the CLL methylome to B cells during the differentiation program (28). With regard to the i-CLL subgroup, no clear link to a normal B cell subpopulation was established (29). However, recent findings shed light into this issue, as the i-CLL group is heavily enriched for IGLV3-21 rearrangements (30), in particular those bearing the R110 mutation, which are also associated with the stereotyped subset #2 cases (31). The fact that the IGLV3-21^{R110} mutation seems to be mediated by AID and thus associated with the mutagenic microenvironment of the GC (31), is an evident argument to conclude that i-CLLs are most likely derived from GC-experienced B cells. The germinal center is not only a mutagenic microenvironment but also a niche where strong DNA methylation programming takes place (32)(24). GC B cells (GCBC) circulate though the dark zone and light zone until they are selected to leave the GC reaction. Thus, the level of SHM and epigenetic programming may be dependent on how many recirculation cycles a B cell has undergone before exiting the GC reaction. In this context, the fact that i-CLLs show both a moderate SHM and an intermediate DNA methylation pattern could reflect that their cell of origin has been positively selected early in the GC reaction. In contrast, m-CLL or high-programmed, which show high SHM load and strong epigenetic programming, could reflect that they derive from a B cell that was positively selected to leave the GC after several maturation affinity cycles. Based on these arguments, we propose that n-CLLs/low-programmed CLLs could be derived from GC-inexperienced MBC cells, i-CLLs/int-programmed CLLs from GC-experienced B cells that have been early selected to leave the GC and m-CLLs/high-programmed CLLs from GC-experienced B cells that have undergone several maturation-affinity cycles after leaving the GC (Figure 2A). For the sake of clarity, we will maintain the n-CLL, i-CLL and m-CLL nomenclature throughout the manuscript.

Remarkably, from a clinical point of view, the 3 subtypes of CLL with different cellular origin identify patients with distinct clinical features (Figure 2B). N-CLL, m-CLL and i-CLL have been associated with a worse, better, and intermediate prognosis, respectively, and this classification seems to be superior to the classical immunogenetic U-CLL/M-CLL categorization in multivariate models. This prognostic impact has been confirmed in several patient series using specific epigenetic biomarkers (30) (29)(33,34). Epigenetic assays to determine these three subtypes are attractive in the clinical setting, as they are stable over time and are not affected by the biological source of patient material (i.e. peripheral blood and lymph node) (29,35). Interestingly, as both i-CLLs and m-CLLs are originated from GC-experienced B cells, the causes underlying their differential prognosis were unknown. A recent study has shed light into this issue, as i-CLLs can be further classified into those showing the IGLV3-21^{R110} mutation and those lacking this feature. Remarkably, those i-CLLs bearing the IGLV3-21^{R110} mutation have a clinical behaviour similar to n-CLLs, and i-CLLs lacking this mutation are more similar to m-CLLs (36). Thus, the overall intermediate clinical behaviour of i-CLLs seems to be influenced by the IGLV3-21^{R110} mutation. Therefore, although the 3 epigenetic subtypes are biologically relevant and provide information regarding the maturation stage of their cellular origin, their clinical importance may be questioned. Additional studies are needed to evaluate whether the epigenetic classification plus the IGLV3-21^{R110} has an added clinical value over a more standard immunogenetics-based approach (i.e. IGHV and IGLV3-21^{R110} mutational status).

The proliferative history of CLL cells leaves DNA methylation imprints and predicts clinical behaviour

Although most studies analysing DNA methylation in cancer have focused on regulatory elements, there is compelling evidence indicating that cancer cells also acquire extensive DNA methylation changes in repressed, late-replicating DNA regions. These changes, although in principle do not have an impact on gene expression, have been used as a faithful strategy to assess accumulated mitotic cell divisions (37,38). As cells undergo subsequent rounds of mitosis, they gradually accumulate

hypomethylation in low-CpG content heterochromatin and hypermethylation in high-CpG content regions marked by the polycomb-repressive complex. This gradual accumulation can therefore be used to trace the proliferative history of the cells. In the context of CLL and other B cell malignancies, a recent study has developed a novel mitotic clock, named epiCMIT (**epi**genetically-determined **c**umulative **mit**oses), which is based both on hyper- and hypomethylation (39). The CpG sites included in this epigenetic clock were selected from repressed regions and their differential methylation did not impact gene expression of nearby genes, which remained silent regardless of the methylation level. In normal B cells, this epiCMIT score gradually increases as cell differentiation progresses, and sharply rises in highly proliferative GCBCs. The highest epiCMIT is observed in terminally-differentiated PCs, which do not proliferate but carry the accumulated cell divisions of the entire B cell maturation process. Therefore, it is not surprising that the 3 epigenetic subtypes of CLL show distinct baseline epiCMIT because they derive from B cells with different proliferative histories, i.e. GC-independent MBCs with low proliferative history, early selected GC-derived MBCs with intermediate proliferative history and late-selected GC-derived MBCs with extensive proliferative history (Figure 3A). If we compare the prognosis of CLLs with different cellular origins, it may look like the lowest proliferative history, the worse the clinical behaviour, as NBC-like CLLs show the poorest outcome. However, if we compare cases sharing the same cellular origin, we can observe variable epiCMIT levels, which presumably are associated with the proliferative history of the leukaemic cells acquired during clonal expansion. In this case, the more profound the proliferative history of the neoplastic clone (i.e. the highest epiCMIT) the worse the clinical behaviour of the patient (Figure 3B). The epiCMIT, analysed as continuous variable, represents an independent prognostic variable in multivariate models with cell of origin epigenetic subtypes and driver genetic alterations (39)(5). It is important to note that the epiCMIT reflects proliferative history and not proliferative status, which can be assessed by proliferation signatures or markers such as Ki-67. The proliferative status of neoplastic cells is in general an important prognostic variable, but this variable depends on time of sampling and type of sample. This is relevant in CLL, as CLL cells recirculate from peripheral blood and lymph nodes, and only undergo cell division in the proliferation centers in lymph nodes. Thus, the epiCMIT, as

measured in peripheral blood, seems to reflect the accumulated mitosis in the proliferative niches.

Although the epiCMIT reflects the overall proliferation history, some studies have also focused on intratumor variability within CLLs methylomes (40,41). Intratumoral DNA methylation showed higher heterogeneity in neoplastic than in normal B cells and it was associated with worst prognosis. Interestingly, it seems that CLL is locally acquiring DNA methylation in a stochastic manner, creating epigenetic "noise" (41), which at least in repressed regions may also be associated with proliferative history. However, such heterogeneity, when targeting regulatory elements may then affect gene expression, likely generating epigenetic plasticity to facilitate tumour evolution.

Another biological variable that may reflect proliferative history is telomere length, which is known to decrease during cell division (42). Studies in CLL have revealed that shorter telomeres are related to a worse clinical behaviour (43,44). Although, to our knowledge, a detailed association between telomere length and epiCMIT has not yet been performed, there seems to be a fundamental difference between them. The epiCMIT is accumulative and gradually increases during B cell maturation (39). In the case of telomere length, it does not linearly decrease as B cells proliferate and differentiate, but sharply increases in GCBCs (due to transient telomerase upregulation, (45)) as compared to naïve and memory B cells, which show similar telomere lengths when sorted from peripheral blood (46). Therefore, in principle, we would not expect U-CLLs and M-CLLs to have different cell of origin-related telomere lengths. Based of these lines of evidence, we believe that the epiCMIT is a more faithful measure of the accumulated proliferative history, including cellular origin and leukaemic expansion, whereas telomere length may be more related to leukaemia-specific proliferation. However, telomere dysfunction may reflect other aspects rather than only proliferative history, as mutations in *POT1* and deregulated expression of *TERT*, both members of the telomere system, have been described in CLL (47).

Chromatin alterations and transcription factor networks

Beyond DNA methylation reports, CLL has been the subject of several genome-wide studies characterising chromatin accessibility and histone modifications, including H3K4me3, H3K4me1, H3K27ac, H3K36me3, H3K9me3 and H3K27me3 (48)(49)(50)(51). Overall, CLL undergoes a genome-wide reconfiguration of each chromatin mark as compared to normal B cell differentiation stages (Figure 4). Although a fraction of the CLL chromatin landscape seems to be largely influenced by distinct dynamics during normal B-cell maturation, an analysis of chromatin states revealed approximately 500 of *de novo* activated regulatory elements in all CLLs regardless of their genetic heterogeneity. These regions targeted regulatory elements of genes known to be involved in CLL pathogenesis, such as *TCF4*, *FMOD*, *CTLA4* or *LEF1*, among others (49,51). Remarkably, the activation of these regulatory elements does not seem to be independent from each other, as their DNA sequence is enriched for binding sites of only few transcription factor families. In particular, NFAT, FOX and TCF/LEF TF families seem to be associated with de novo, CLL-specific chromatin activation (Figure 4A). These TFs have also identified as major modulators of the CLL epigenome in independent studies (28)(49). Further research on these TF families shall determine their role in the initiation and maintenance of chromatin activation in CLL, and whether their inhibition may represent suitable therapeutic path to explore. Complementary to these studies centered on CLL-specific chromatin alterations, an additional study explored chromatin maps of CLL from a different perspective. This study identified that, as compared to normal B cells, CLL cells show an altered pattern of histone modification combinations, as typically mutually exclusive modifications unexpectedly co-localize (50). This finding suggests the presence of intra-tumoral epigenetic diversity generated by an admixture of cells with diverging cellular identities. Thus, as it has been done for DNA methylation (41), chromatin modifications could also be informative on different paths of subclonal evolution.

In addition to chromatin patterns in CLL as a whole, reports have also identified clearly distinct landscapes of chromatin accessibility and histone modifications, as well as TF networks in U-CLL and M-CLL (51,52). If the differential patterns are placed in the context of normal B cell maturation, regions not only can be classified as *de novo* altered in U-CLL or M-CLL, but also were associated with multiple distinct dynamics

during B cell differentiation (51)(Figure 4A). These reflected in part the link between M-CLL and GC-experienced cells, and between U-CLL and GC-inexperienced cells, but also more unexpected patterns. For instance, part of the regions active in U-CLL as compared to M-CLL are also more active in GCBCs than other maturation stages, suggesting that the more proliferative U-CLL subgroup may hijack some of the mechanisms associated with proliferative GCBCs. Intriguingly, the i-CLL group described by DNA methylation analyses does not show a clear differential chromatin signature (51,52). This finding suggests that DNA methylation is a better tracer of cellular memory than chromatin-based marks, which are more associated with the activity of regulatory elements.

Chromatin profiling has also been exploited in the context of therapy in CLL. Rendeiro and coworkers identified that ibrutinib treatment induces a specific chromatin regulatory program, identifying decreased NF-kB binding, B cell TF activity, and CLL identity and acquisition of a quiescent expression pattern (Figure 4B). Interestingly pretreated samples could be used to predict the temporal dynamics in the response to ibrutinib treatment (53). Additionally, it seems that prolonged ibrutinib treatment induces a global reduction of specific histone modifications such as H3K27ac and H3K27me3 (54). Another study applied chromatin profiling and chemosensitivity profiling to identify ibrutinib combination therapies. During ibrutinib treatment of CLL patient samples, different patterns of chromatin accessibility were associated with changes in chemosensitivity during ibrutinib treatment (55). This approach suggessted that proteasome, PLK1, and mTOR inhibitors were the preferencial drug combinations with ibrutinib, although further validation experiments are needed. Furthermore, it seems that not only ibrutinib but also standard chemo- and chemoimmunotherapies can affect the epigenome. In particular, the study of sequential CLL samples has identified post-treatment DNA methylation changes (56,57).

The extensive chromatin remodelling observed in CLL may itself represent a vulnerability axis to be exploited therapeutically. For instance, recent studies have explored the possibility of using inhibitors of Bromodomain and Extra-Terminal domain (BET) proteins, which target histone acetylation, in CLL therapy. These agents exert in vitro and in vivo antiproliferative effects in preclinical models of CLL (58,59). Moreover,

specific blocking of BRD4, a BET protein member overexpressed in CLL, leads to downregulation of multiple important signalling pathways, including the BCR (58). Hence, BCR4 inhibitors may represent attractive epigenetic agents to be combined with standard therapies.

Finally, a recent report analysed the 3D genome architecture of CLL and normal B cell subpopulations using in situ HiC (60), a whole-genome chromosome conformation capture technique (61). This study analysed 3D genome interactions in the context of other epigenomic marks and chromatin states, and identified a third 3D genome compartment, coined as intermediate, between the A (active) and B (inactive) compartments. This intermediate compartment, which was independently described also in colorectal cancer (62), was enriched in regions marked with the polycomb-repressive complex, and represents a transitional 3D genome feature between A and B compartments. As compared to normal B cell maturation, CLL undergoes a widespread reorganization of the 3D genome architecture. At a 100 Kb resolution, several hundred regions changed from one 3D compartment to another, including both activation and inactivation shifts. Interestingly, the chromatin blocks changing compartments contained genes associated with CLL pathogenesis. A remarkable example is the *EBF1* locus, which in normal B cells is open and contains multiple 3D interactions between the gene and its regulatory elements (Figure 4C). In CLL cells, however, the 3D interactions are abolished, the gene becomes heterochromatic and the space that the *EBF1* locus occupies within the interphase nucleus is significantly reduced.

Linking epigenetic and genetic features

The genomic landscape of CLL has been thoroughly investigated through whole-genome and exome sequencing in large series of patients (6–8), and some studies have analysed the association between genetic and epigenetic features. Although B cell neoplasms such as follicular lymphoma and diffuse large B cell lymphoma frequently carry mutations in epigenetic modifiers (63), such mutations in CLL are rather infrequent. Mutations in components of chromatin structure/function such as *CHD2*,

KMT2D, *ARID1A*, *ASXL1* or *SETD2* have been reported at low frequencies (1-5%) in CLL (5–7).

Although genetic changes in epigenetic genes are rare in CLL, the link between genetics and epigenetics in CLL has been studied through several perspectives. First, CLLs derived from the 3 different cellular origins are related to different driver mutations, being e.g. *NOTCH1* enriched in n-CLL, *SF3B1* in i-CLL and *MYD88* in m-CLL (Figure 2B) (6,27). This finding suggests that B cells at different maturation stages provide a fertile soil for differential driver genes to be selected. Second, within each of the 3 epigenetic subtypes, specific driver mutations have been linked to DNA methylation changes derived from a higher proliferative history (i.e. epiCMIT score) (39). Therefore, this association with the epiCMIT can identify driver mutations that confer a proliferative advantage. A striking example is the presence of *SF3B1* mutations, which identifies i-CLLs with higher epiCMIT as compared to i-CLLs lacking *SF3B1*. Third, another aspect related to the genetics-epigenetics link is whether the presence of particular driver alterations leads to a defined epigenetic signature. Mutations in some drivers are related to the expected functional changes such as NOTCH signalling in *NOTCH1* mutations or alternative splicing in *SF3B1* mutations (8). Based on these findings, it is logical to speculate that genetic alterations may induce particular downstream epigenetic signatures. However, the results in CLL are less clear than expected. DNA methylation, chromatin accessibility and chromatin activation measured by H3K27ac in patients with specific driver genetic alterations have revealed that mostly trisomy 12 and *MYD88* mutations are related to specific epigenetic patterns (51,64) (Figure 5A and B). However, this moderate association between genetic drivers and epigenetic signatures may in part be related to the fact that some mutations are subclonally present in CLL cells. Supporting this potential confounder, it is intriguing that trisomy 12 and *MYD88* mutations, which are known to be frequently clonal, are those related to more robust epigenetic signatures. Further studies, considering subclonality, ideally by single cell genetic and epigenetic analyses, shall shed light into the association between specific drivers and downstream epigenetic patterns.

The fourth aspect to highlight regarding the genetics-epigenetics connection is in the identification of potentially functional somatic non-coding mutations (Figure 5C). Chromatin profiling of CLL provides a high-resolution map of regulatory elements across the genome (51), which can be used to map non-coding mutations to enhancers and promoters. Whole-genome studies in CLL have also identified the significant accumulation of somatic mutations in non-coding regions (6,65–67). Most non-coding mutations are located in heterochromatic regions, and few of them overlap with regulatory elements (51). Among these, a distant enhancer in chromosome 9p13 was found to be recurrently mutated in M-CLL cases and the presence of the mutation is related with lower expression of *PAX5*, which is located 330 Kb away. Chromosome conformation capture analyses further confirms that the enhancer and the *PAX5* locus are in close proximity in the 3D space (6) (Figure 5C). Other examples of potentially relevant non-coding mutations affect disease-relevant genes such as *ATM*, *TCL1A*, *IKZF3*, *SAMHD1*, and *BIRC3* (67).

Finally, the genetics-epigenetics connection has been exploited in the field of genetic predisposition (Figure 5D). CLL is a one of the human cancers with highest familial risk, and a total of 43 SNPs have been described to influence CLL risk (68,69). Some risk SNPs have been already linked with regulatory elements, such as the *BMF* super-enhancer polymorphism (70), but a more recent study revealed that 93% of the risk loci are actually located in regulatory regions in CLL (71). Some SNPs were located within transcription factor (TF) binding sites, and the presence of the risk allele was related to altered TF binding affinity. Remarkably, the binding affinity for TFs of the NFAT, TCF/LEF, and FOX families was increased, and these TFs have been associated with *de novo* chromatin activation in CLL as compared to normal B cells (51) (Figure 5D). Interestingly SNPs in regulatory elements of *LEF1* itself are also related to CLL risk.

Overall, genetic and epigenetic features constitute two interconnected layers of the molecular landscape of CLL. On the one hand, epigenetics is useful to determine the potential function of non-coding genetic variants. On the other hand, some genetic alterations can lead to a higher proliferation, which will impact the epigenetic mitotic clock, and for few driver alterations, a specific epigenetic signature is observed. However, it seems that a large fraction of epigenetic changes are not associated with

genetic alterations (e.g. 500 regions become *de novo* active in all CLLs regardless of their genetic drivers). Therefore, other factors such as autonomous BCR signalling (72) or microenvironmental interactions (1) and their downstream signalling cascades also need to be analysed to better understand the causes underlying epigenetic programming in CLL.

Next stop: single cell omics

There is no doubt that bulk epigenomics has provided valuable biological and clinical insights related to the cellular origin, pathogenic mechanisms and evolution of CLL. However, subclonal genetic, epigenetic and transcriptional heterogeneity is definitely present in CLL (73). The relatively recent development of technologies allowing us to characterise the molecular profile of individual cells is revolutionising the field of biology in general and of cancer genomics in particular. In CLL, the first single cell papers are already appearing in the literature, with breakthrough discoveries. For instance, in an integrative bulk and single cell multi-omic analyses of sequential CLL samples undergoing a Richter transformation, Nadeu and coworkers have recently identified that the seeds of Richter transformation are already present, but dormant, in the CLL samples at diagnosis even 19 years prior to Richter transformation (74). Furthermore, Penter and colleagues described a new multiome single cell assay that combines the use of mitochondrial DNA mutations as barcodes of genetic evolution together with chromatin accessibility (75). This assay was applied to sequential samples during CLL progression, and revealed the presence of extensive changes associated with relapse and transformation, as well as an association between CLL subclones to distinct chromatin states, providing evidence for epigenetic modulation during relapse. Finally, Gaiti and colleagues integrated single cell DNA methylation analyses with gene expression and specific driver mutations to trace the past history of CLL cells (76). Their analyses of lineage trees indicated that CLL cells undergo a rapid drift after initial steps of clonal expansion. This may give rise to the early presence of subclones that compete to become expanded, a finding that may explain the early seeds of Richter transformation identified by Nadeu (74). Additionally, this study also identified, using

the example of *SF3B1* mutations, that genetic subclones arising during disease development mapped to distinct DNA methylation branches.

These three articles are just examples of the unprecedented power of single cell technologies to dive into the dynamic evolution and subclonal architecture of CLL cells. However, single cell technologies have the added value of being able to also characterize the non-neoplastic cells accompanying leukaemic cells. Microenvironmental cells such as T cells, nurse-like cells and stromal have a strong impact on CLL biology, as they nourish CLL cells with survival and proliferation signals in the lymph nodes (1). On the other hand, modulation of the microenvironment, in particular the reversal of T-cell exhaustion, may also lead to the rare occurrence of spontaneous CLL regression (2). Therefore, a thorough single cell characterisation of the CLL microenvironment in patients with different disease courses, must complement the one of the neoplastic cells, as CLL pathogenesis and clinical behaviour depends on the continuous dialogue between neoplastic cells and their niche. Ideally, the upcoming integrative single cells studies shall be used to predict the evolution of individual patients for a precision clinical management. In this scenario, the early seeds of future clinical complications, if detected at diagnosis, may be counteracted by an early interception.

ACKNOWLEDGEMENTS

The authors thank all members of the biomedical epigenomics group and the lymphoid neoplasms program of the IDIBAPS for stimulating discussions. The author's research on epigenetics is funded by the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation program (810287, BCLLatlas), Fundació La Marató de TV3, the Accelerator award CRUK/AIRC/AECC joint funder-partnership, Generalitat de Catalunya Suport Grups de Recerca AGAUR, and CIBERONC (CB16/12/00225). This work was developed at the Centre Esther Koplowitz (CEK, Barcelona, Spain).

AUTHOR CONTRIBUTIONS

Both authors have written the manuscript and designed the figures.

CONFLICT OF INTEREST

The authors declare no conflicts of interest

DATA AVAILABILITY STATEMENT

Data sharing is not applicable to this review article as no new data were created or analysed in this study.

REFERENCES

1. Kipps TJ, Stevenson FK, Wu CJ, Croce CM, Packham G, Wierda WG, et al. Chronic lymphocytic leukaemia. Nat Rev Dis Primers. 2017;3:16096.

2. Kwok M, Oldreive C, Rawstron AC, Goel A, Papatzikas G, Jones RE, et al. Integrative analysis of spontaneous CLL regression highlights genetic and microenvironmental interdependency in CLL. Blood. 2020;135(6):411–28.

3. Hamblin TJ, Davis Z, Gardiner A, Oscier DG, Stevenson FK. Unmutated Ig V(H) genes are associated with a more aggressive form of chronic lymphocytic leukemia. Blood. 1999 Sep 15;94(6):1848–54.

4. Damle RN, Wasil T, Fais F, Ghiotto F, Valetto A, Allen SL, et al. Ig V gene mutation status and CD38 expression as novel prognostic indicators in chronic lymphocytic leukemia. Blood. 1999 Sep 15;94(6):1840–7.

5. Knisbacher BA, Lin Z, Hahn CK, Nadeu F, Duran-Ferrer M, Stevenson KE, et al. Molecular map of chronic lymphocytic leukemia and its impact on outcome. Nat Genet. 2022 Aug 4;

6. Puente XS, Beà S, Valdés-Mas R, Villamor N, Gutiérrez-Abril J, Martín-Subero JI, et al. Non-coding recurrent mutations in chronic lymphocytic leukaemia. Nature. 2015 Oct 22;526(7574):519–24.

7. Landau DA, Tausch E, Taylor-Weiner AN, Stewart C, Reiter JG, Bahlo J, et al. Mutations driving CLL and their evolution in progression and relapse. Nature. 2015 Oct 22;526(7574):525–30.

8. Bosch F, Dalla-Favera R. Chronic lymphocytic leukaemia: from genetics to treatment. Nat Rev Clin Oncol. 2019;16(11):684–701.

9. Mansouri L, Wierzbinska JA, Plass C, Rosenquist R. Epigenetic deregulation in chronic lymphocytic leukemia: Clinical and biological impact. Semin Cancer Biol. 2018;51:1–11.

10. Nadeu F, Diaz-Navarro A, Delgado J, Puente XS, Campo E. Genomic and Epigenomic Alterations in Chronic Lymphocytic Leukemia. Annu Rev Pathol. 2020;15:149–77.

11. Bernstein BE, Meissner A, Lander ES. The mammalian epigenome. Cell. 2007 Feb 23;128(4):669–81.

12. Kulis M, Queirós AC, Beekman R, Martín-Subero JI. Intragenic DNA methylation in transcriptional regulation, normal differentiation and cancer. Biochim Biophys Acta. 2013 Nov;1829(11):1161–74.

13. Jones PA. Functions of DNA methylation: islands, start sites, gene bodies and beyond. Nat Rev Genet. 2012 May 29;13(7):484–92.

14. Ernst J, Kheradpour P, Mikkelsen TS, Shoresh N, Ward LD, Epstein CB, et al. Mapping and analysis of chromatin state dynamics in nine human cell types. Nature. 2011 May 5;473(7345):43–9.

15. Thurman RE, Rynes E, Humbert R, Vierstra J, Maurano MT, Haugen E, et al. The accessible chromatin landscape of the human genome. Nature. 2012 Sep 6;489(7414):75–82.

16. Lieberman-Aiden E, van Berkum NL, Williams L, Imakaev M, Ragoczy T, Telling A, et al. Comprehensive mapping of long-range interactions reveals folding principles of the human genome. Science. 2009 Oct 9;326(5950):289–93.

17. Dekker J, Marti-Renom MA, Mirny LA. Exploring the three-dimensional organization of genomes: interpreting chromatin interaction data. Nat Rev Genet. 2013 Jun;14(6):390–403.

18. Chiorazzi N, Ferrarini M. Cellular origin(s) of chronic lymphocytic leukemia: cautionary notes and additional considerations and possibilities. Blood. 2011 Feb 10;117(6):1781–91.

19. Ng A, Chiorazzi N. Potential Relevance of B-cell Maturation Pathways in Defining the Cell(s) of Origin for Chronic Lymphocytic Leukemia. Hematol Oncol Clin North Am. 2021;35(4):665–85.

20. Stamatopoulos K, Agathangelidis A, Rosenquist R, Ghia P. Antigen receptor stereotypy in chronic lymphocytic leukemia. Leukemia. 2017;31(2):282–91.

21. Seifert M, Sellmann L, Bloehdorn J, Wein F, Stilgenbauer S, Dürig J, et al. Cellular origin and pathophysiology of chronic lymphocytic leukemia. J Exp Med. 2012 Nov 19;209(12):2183–98.

22. Klein U, Tu Y, Stolovitzky GA, Mattioli M, Cattoretti G, Husson H, et al. Gene Expression Profiling of B Cell Chronic Lymphocytic Leukemia Reveals a Homogeneous Phenotype Related to Memory B Cells✪. Journal of Experimental Medicine. 2001 Dec 3;194(11):1625–38.

23. Martin-Subero JI, Oakes CC. Charting the dynamic epigenome during B-cell development. Semin Cancer Biol. 2018;51:139–48.

24. Kulis M, Merkel A, Heath S, Queirós AC, Schuyler RP, Castellano G, et al. Whole-genome fingerprint of the DNA methylome during human B cell differentiation. Nat Genet. 2015 Jul;47(7):746–56.

25. Bird A. DNA methylation patterns and epigenetic memory. Genes Dev. 2002 Jan 1;16(1):6–21.

26. Hoadley KA, Yau C, Hinoue T, Wolf DM, Lazar AJ, Drill E, et al. Cell-of-Origin Patterns Dominate the Molecular Classification of 10,000 Tumors from 33 Types of Cancer. Cell. 2018;173(2):291-304.e6.

27. Kulis M, Heath S, Bibikova M, Queirós AC, Navarro A, Clot G, et al. Epigenomic analysis detects widespread gene-body DNA hypomethylation in chronic lymphocytic leukemia. Nat Genet. 2012 Nov;44(11):1236–42.

28. Oakes CC, Seifert M, Assenov Y, Gu L, Przekopowitz M, Ruppert AS, et al. DNA methylation dynamics during B cell maturation underlie a continuum of disease phenotypes in chronic lymphocytic leukemia. Nat Genet. 2016 Mar;48(3):253–64.

29. Queirós AC, Villamor N, Clot G, Martinez-Trillos A, Kulis M, Navarro A, et al. A B-cell epigenetic signature defines three biologic subgroups of chronic lymphocytic leukemia with clinical impact. Leukemia. 2015 Mar;29(3):598–605.

30. Giacopelli B, Zhao Q, Ruppert AS, Agyeman A, Weigel C, Wu YZ, et al. Developmental subtypes assessed by DNA methylation-iPLEX forecast the natural history of chronic lymphocytic leukemia. Blood. 2019;134(8):688–98.

31. Maity PC, Bilal M, Koning MT, Young M, van Bergen CAM, Renna V, et al. IGLV3-21*01 is an inherited risk factor for CLL through the acquisition of a single-point mutation enabling autonomous BCR signaling. Proc Natl Acad Sci U S A. 2020;117(8):4320–7.

32. Shaknovich R, Cerchietti L, Tsikitas L, Kormaksson M, De S, Figueroa ME, et al. DNA methyltransferase 1 and DNA methylation patterning contribute to germinal center B-cell differentiation. Blood. 2011 Sep 29;118(13):3559–69.

33. Bhoi S, Ljungström V, Baliakas P, Mattsson M, Smedby KE, Juliusson G, et al. Prognostic impact of epigenetic classification in chronic lymphocytic leukemia: The case of subset #2. Epigenetics. 2016 Jun 2;11(6):449–55.

34. Wojdacz TK, Amarasinghe HE, Kadalayil L, Beattie A, Forster J, Blakemore SJ, et al. Clinical significance of DNA methylation in chronic lymphocytic leukemia patients: results from 3 UK clinical trials. Blood Adv. 2019;3(16):2474–81.

35. Cahill N, Bergh AC, Kanduri M, Göransson-Kultima H, Mansouri L, Isaksson A, et al. 450K-array analysis of chronic lymphocytic leukemia cells reveals global DNA methylation to be relatively stable over time and similar in resting and proliferative compartments. Leukemia. 2013 Jan;27(1):150–8.

36. Nadeu F, Royo R, Clot G, Duran-Ferrer M, Navarro A, Martín S, et al. IGLV3-21R110 identifies an aggressive biological subtype of chronic lymphocytic leukemia with intermediate epigenetics. Blood. 2021;137(21):2935–46.

37. Zhou W, Dinh HQ, Ramjan Z, Weisenberger DJ, Nicolet CM, Shen H, et al. DNA methylation loss in late-replicating domains is linked to mitotic cell division. Nat Genet. 2018;50(4):591–602.

38. Yang Z, Wong A, Kuh D, Paul DS, Rakyan VK, Leslie RD, et al. Correlation of an epigenetic mitotic clock with cancer risk. Genome Biol. 2016;17(1):205.

39. Duran-Ferrer M, Clot G, Nadeu F, Beekman R, Baumann T, Nordlund J, et al. The proliferative history shapes the DNA methylome of B-cell tumors and predicts clinical outcome. Nat Cancer. 2020;1(11):1066–81.

40. Oakes CC, Claus R, Gu L, Assenov Y, Hüllein J, Zucknick M, et al. Evolution of DNA methylation is linked to genetic aberrations in chronic lymphocytic leukemia. Cancer Discov. 2014 Mar;4(3):348–61.

41. Landau DA, Clement K, Ziller MJ, Boyle P, Fan J, Gu H, et al. Locally disordered methylation forms the basis of intratumor methylome variation in chronic lymphocytic leukemia. Cancer Cell. 2014 Dec 8;26(6):813–25.

42. Allsopp RC, Chang E, Kashefi-Aazam M, Rogaev EI, Piatyszek MA, Shay JW, et al. Telomere shortening is associated with cell division in vitro and in vivo. Exp Cell Res. 1995 Sep;220(1):194–200.

43. Bloehdorn J, Braun A, Taylor-Weiner A, Jebaraj BMC, Robrecht S, Krzykalla J, et al. Multi-platform profiling characterizes molecular subgroups and resistance networks in chronic lymphocytic leukemia. Nat Commun. 2021;12(1):5395.

44. Jebaraj BMC, Tausch E, Landau DA, Bahlo J, Robrecht S, Taylor-Weiner AN, et al. Short telomeres are associated with inferior outcome, genomic complexity, and clonal evolution in chronic lymphocytic leukemia. Leukemia. 2019;33(9):2183–94.

45. Hu BT, Lee SC, Marin E, Ryan DH, Insel RA. Telomerase is up-regulated in human germinal center B cells in vivo and can be re-expressed in memory B cells activated in vitro. J Immunol. 1997 Aug 1;159(3):1068–71.

46. Son NH, Joyce B, Hieatt A, Chrest FJ, Yanovski J, Weng N ping. Stable telomere length and telomerase expression from naïve to memory B-lymphocyte differentiation. Mech Ageing Dev. 2003 Apr;124(4):427–32.

47. Jebaraj BMC, Stilgenbauer S. Telomere Dysfunction in Chronic Lymphocytic Leukemia. Front Oncol. 2020;10:612665.

48. Ott CJ, Federation AJ, Schwartz LS, Kasar S, Klitgaard JL, Lenci R, et al. Enhancer Architecture and Essential Core Regulatory Circuitry of Chronic Lymphocytic Leukemia. Cancer Cell. 2018;34(6):982-995.e7.

49. Mallm JP, Iskar M, Ishaque N, Klett LC, Kugler SJ, Muino JM, et al. Linking aberrant chromatin features in chronic lymphocytic leukemia to transcription factor networks. Mol Syst Biol. 2019;15(5):e8339.

50. Pastore A, Gaiti F, Lu SX, Brand RM, Kulm S, Chaligne R, et al. Corrupted coordination of epigenetic modifications leads to diverging chromatin states and transcriptional heterogeneity in CLL. Nat Commun. 2019;10(1):1874.

51. Beekman R, Chapaprieta V, Russiñol N, Vilarrasa-Blasi R, Verdaguer-Dot N, Martens JHA, et al. The reference epigenome and regulatory chromatin landscape of chronic lymphocytic leukemia. Nat Med. 2018;24(6):868–80.

52. Rendeiro AF, Schmidl C, Strefford JC, Walewska R, Davis Z, Farlik M, et al. Chromatin accessibility maps of chronic lymphocytic leukaemia identify subtype-specific epigenome signatures and transcription regulatory networks. Nat Commun. 2016;7:11938.

53. Rendeiro AF, Krausgruber T, Fortelny N, Zhao F, Penz T, Farlik M, et al. Chromatin mapping and single-cell immune profiling define the temporal dynamics of ibrutinib response in CLL. Nat Commun. 2020 Jan 29;11(1):577.

54. Holmes KB, Sadreev II, Rawstron AC, Munir T, Westhead DR, Hillmen P, et al. Ibrutinib induces chromatin reorganisation of chronic lymphocytic leukaemia cells. Oncogenesis. 2019 May 10;8(5):32.

55. Schmidl C, Vladimer GI, Rendeiro AF, Schnabl S, Krausgruber T, Taubert C, et al. Combined chemosensitivity and chromatin profiling prioritizes drug combinations in CLL. Nat Chem Biol. 2019;15(3):232–40.

56. Barrow TM, Nakjang S, Lafta F, Bilotkach K, Woodhouse L, Junge G, et al. Epigenome-wide analysis reveals functional modulators of drug sensitivity and post-treatment survival in chronic lymphocytic leukaemia. Br J Cancer. 2021;124(2):474–83.

57. Tsagiopoulou M, Papakonstantinou N, Moysiadis T, Mansouri L, Ljungström V, Duran-Ferrer M, et al. DNA methylation profiles in chronic lymphocytic leukemia patients treated with chemoimmunotherapy. Clin Epigenetics. 2019;11(1):177.

58. Ozer HG, El-Gamal D, Powell B, Hing ZA, Blachly JS, Harrington B, et al. BRD4 Profiling Identifies Critical Chronic Lymphocytic Leukemia Oncogenic Circuits and Reveals Sensitivity to PLX51107, a Novel Structurally Distinct BET Inhibitor. Cancer Discov. 2018;8(4):458–77.

59. Ott CJ, Federation AJ, Schwartz LS, Kasar S, Klitgaard JL, Lenci R, et al. Enhancer Architecture and Essential Core Regulatory Circuitry of Chronic Lymphocytic Leukemia. Cancer Cell. 2018;34(6):982-995.e7.

60. Vilarrasa-Blasi R, Soler-Vila P, Verdaguer-Dot N, Russiñol N, di Stefano M, Chapaprieta V, et al. Dynamics of genome architecture and chromatin function during human B cell differentiation and neoplastic transformation. Nat Commun. 2021;12(1):651.

61. Rao SSP, Huntley MH, Durand NC, Stamenova EK, Bochkov ID, Robinson JT, et al. A 3D map of the human genome at kilobase resolution reveals principles of chromatin looping. Cell. 2014 Dec 18;159(7):1665–80.

62. Johnstone SE, Reyes A, Qi Y, Adriaens C, Hegazi E, Pelka K, et al. Large-Scale Topological Changes Restrain Malignant Progression in Colorectal Cancer. Cell. 2020;182(6):1474-1489.e23.

63. Morin RD, Mendez-Lago M, Mungall AJ, Goya R, Mungall KL, Corbett RD, et al. Frequent mutation of histone-modifying genes in non-Hodgkin lymphoma. Nature. 2011 Jul 27;476(7360):298–303.

64. Tsagiopoulou M, Chapaprieta V, Duran-Ferrer M, Moysiadis T, Psomopoulos F, Kollia P, et al. Chronic lymphocytic leukemias with trisomy 12 show a distinct DNA methylation profile linked to altered chromatin activation. Haematologica. 2020;105(12):2864–7.

65. Kasar S, Kim J, Improgo R, Tiao G, Polak P, Haradhvala N, et al. Whole-genome sequencing reveals activation-induced cytidine deaminase signatures during indolent chronic lymphocytic leukaemia evolution. Nat Commun. 2015 Dec 7;6:8866.

66. Spina V, Rossi D. Overview of non-coding mutations in chronic lymphocytic leukemia. Mol Oncol. 2019;13(1):99–106.

67. Burns A, Alsolami R, Becq J, Stamatopoulos B, Timbs A, Bruce D, et al. Whole-genome sequencing of chronic lymphocytic leukaemia reveals distinct differences in the mutational landscape between IgHVmut and IgHVunmut subgroups. Leukemia. 2018;32(2):332–42.

68. Berndt SI, Camp NJ, Skibola CF, Vijai J, Wang Z, Gu J, et al. Meta-analysis of genome-wide association studies discovers multiple loci for chronic lymphocytic leukemia. Nat Commun. 2016 Mar 9;7:10933.

69. Law PJ, Berndt SI, Speedy HE, Camp NJ, Sava GP, Skibola CF, et al. Genome-wide association analysis implicates dysregulation of immunity genes in chronic lymphocytic leukaemia. Nat Commun. 2017;8:14175.

70. Kandaswamy R, Sava GP, Speedy HE, Beà S, Martín-Subero JI, Studd JB, et al. Genetic Predisposition to Chronic Lymphocytic Leukemia Is Mediated by a BMF Super-Enhancer Polymorphism. Cell Rep. 2016;16(8):2061–7.

71. Speedy HE, Beekman R, Chapaprieta V, Orlando G, Law PJ, Martín-García D, et al. Insight into genetic predisposition to chronic lymphocytic leukemia from integrative epigenomics. Nat Commun. 2019;10(1):3615.

72. Dühren-von Minden M, Übelhart R, Schneider D, Wossning T, Bach MP, Buchner M, et al. Chronic lymphocytic leukaemia is driven by antigen-independent cell-autonomous signalling. Nature. 2012 Sep 13;489(7415):309–12.

73. Gohil SH, Wu CJ. Dissecting CLL through high-dimensional single-cell technologies. Blood. 2019;133(13):1446–56.

74. Nadeu F, Royo R, Massoni-Badosa R, Playa-Albinyana H, Garcia-Torre B, Duran-Ferrer M, et al. Detection of early seeding of Richter transformation in chronic lymphocytic leukemia. Nat Med. 2022;28(8):1662–71.

75. Penter L, Gohil SH, Lareau C, Ludwig LS, Parry EM, Huang T, et al. Longitudinal Single-Cell Dynamics of Chromatin Accessibility and Mitochondrial Mutations in Chronic

Lymphocytic Leukemia Mirror Disease History. Cancer Discov. 2021 Dec 1;11(12):3048–63.

76. Gaiti F, Chaligne R, Gu H, Brand RM, Kothen-Hill S, Schulman RC, et al. Epigenetic evolution and lineage histories of chronic lymphocytic leukaemia. Nature. 2019;569(7757):576–80.

FIGURE LEGENDS

Figure 1. Epigenetic regulation. (A) Schematic representation of different epigenetic layers of information whose integrative analysis is associated with genome function and gene expression regulation. (B) Heatmap showing the combination of histone modifications associated with different chromatin states.

Figure 2. Cell of origin of CLL. (A) Tentative model summarizing the different cells of origin of CLL subtypes (i.e. n-CLL, i-CLL and m-CLL). (B) Table showing selected differential features among the distinct CLL subtypes.

Figure 3. Epigenetics and proliferative history in CLL (A) Boxplots representing the concept of the epiCMIT score, made out of two components. Below is the component representing the different proliferative histories of the putative cell of origin and above the component related to the proliferation history of the neoplastic clone. (B) Kaplan-Meier plot exemplifying the association between higher epiCMIT score and worse prognosis within each epigenetic subtype of CLL.

Figure 4. Chromatin alterations in CLL (A) Schematic representation of different patterns of chromatin activations of CLL cells in the context of normal B cells. The active chromatin regions in the figure correspond to chromatin states containing H3K27ac, such as active promoters and enhancers as well as transcriptional transition. (B) Chromatin accessibility changes upon ibrutinib treatment. (C) Visualization of HiC contact maps of a chromosomal region containing the *EBF1* locus in representative NBC and CLL samples. In NBC, a high density of 3D interactions within the EBF1-containing TAD (marked as black triangle) are observed. These interactions are mostly lost in the CLL sample in which the gene is completely inactive. NBC - naïve B cells, MBC – memory B cells, GCBC – germinal center B cells

Figure 5. Links between genetic features and epigenetics. (A) Chromatin changes associated with the presence of coding mutations, exemplified by CLL cases with or without *MYD88* mutation. (B) Chromatin changes associated with chromosome alterations, exemplified by CLL cases with or without trisomy 12. (C) Non-coding mutations in CLL can modulate the activity of regulatory elements, as shown by mutations in a distant enhancer region that is associated with changes in the expression of *PAX5.* (D) SNPs associated with genetic predisposition in CLL are commonly located in regulatory elements.

Figure 1

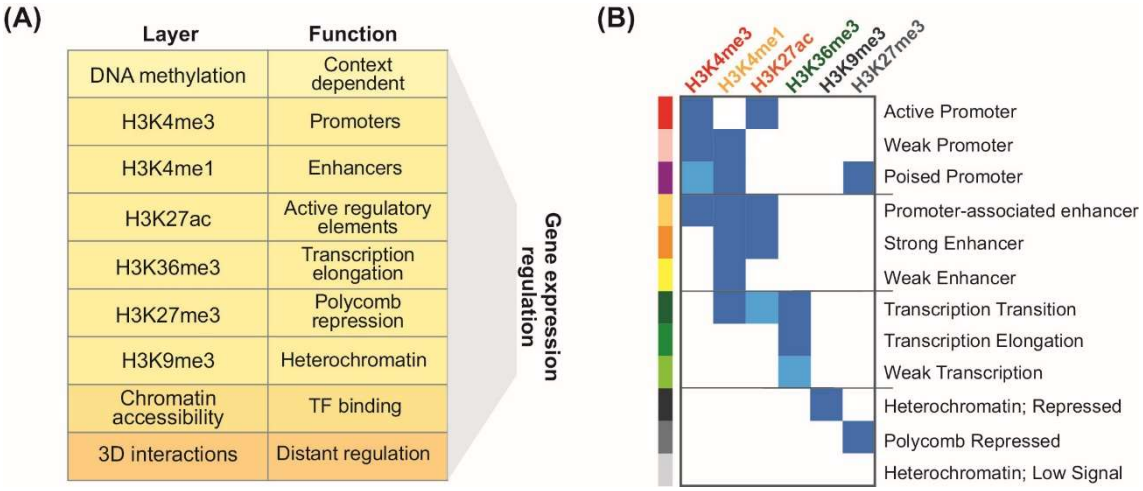
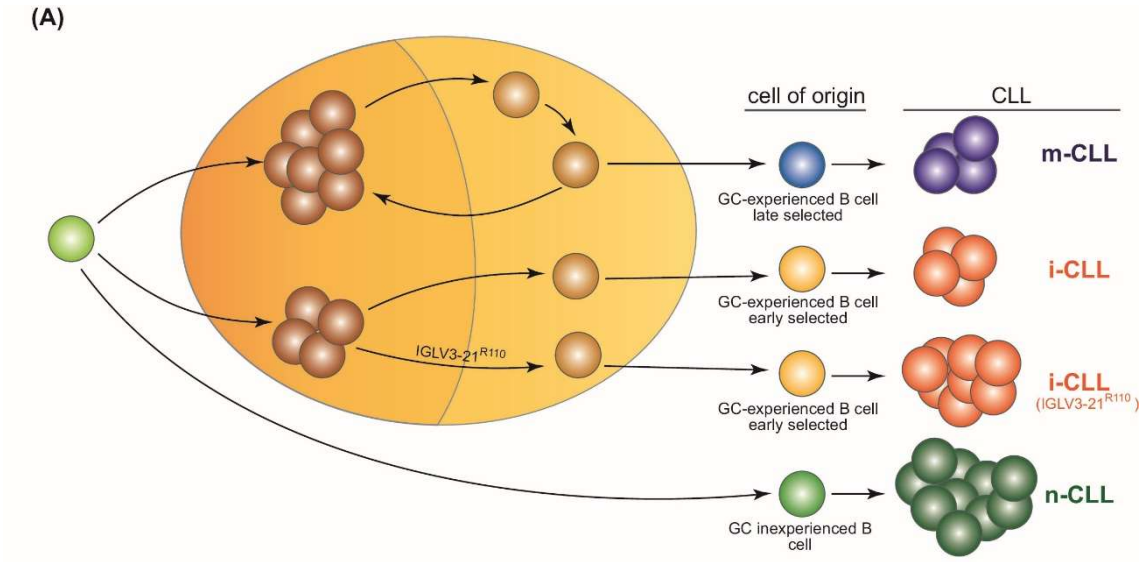


Figure 2



(B)

	DNA methylation	IGHV identity	Stereotyped BCR subset	Genetics	Prognosis
m-CLL	high GC imprint	<95%	#4	*MYD88* mut	better
i-CLL	intermediate GC imprint	95-98%	#2	*SF3B1* mut	intermediate
n-CLL	lack of GC imprint	>98%	#1, #5, #6, #8	del(11q), *NOTCH1* mut	worse

IGLV3-21^{R110}: no → better; yes → worse

Figure 3

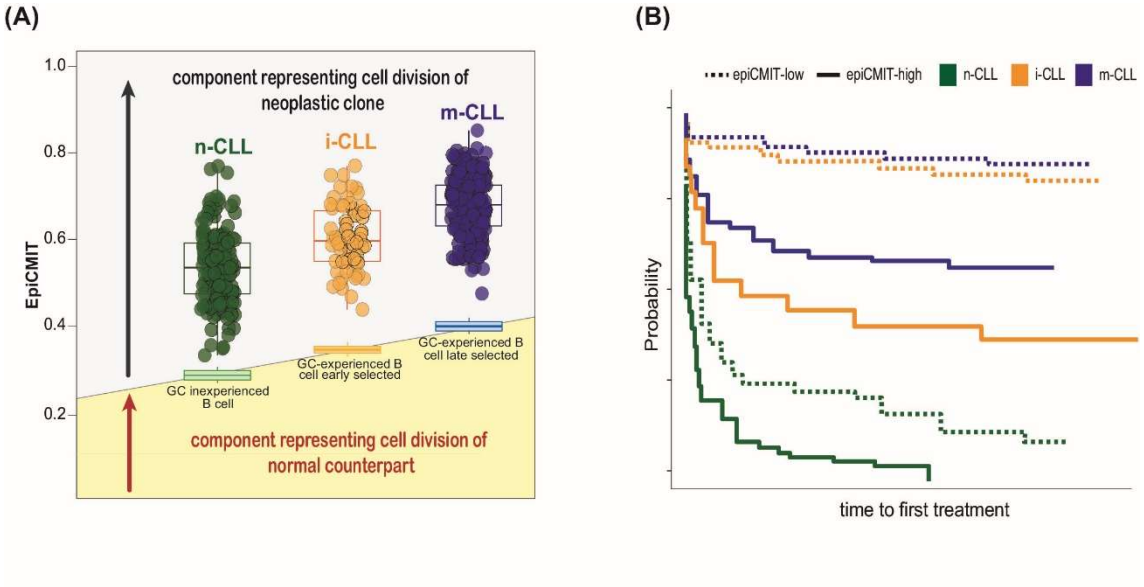


Figure 4

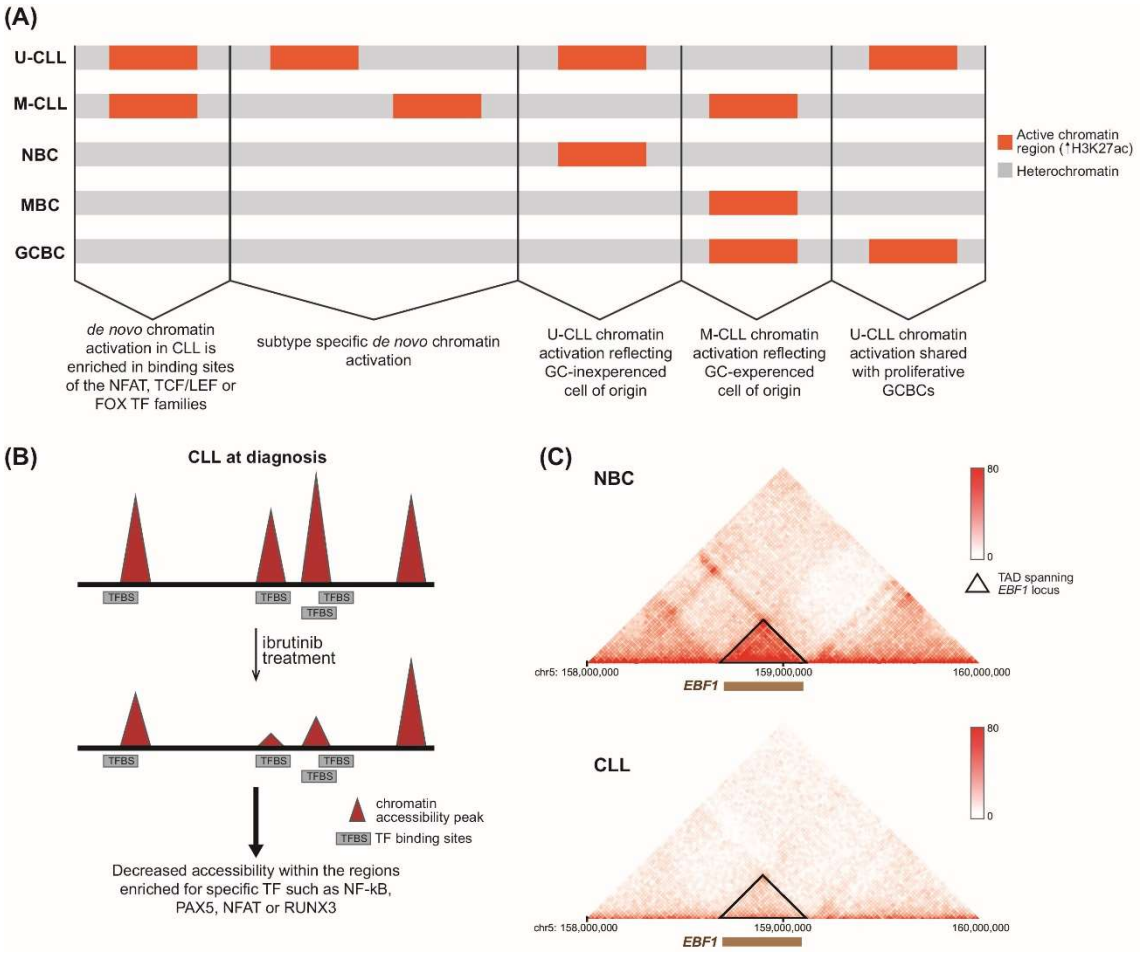
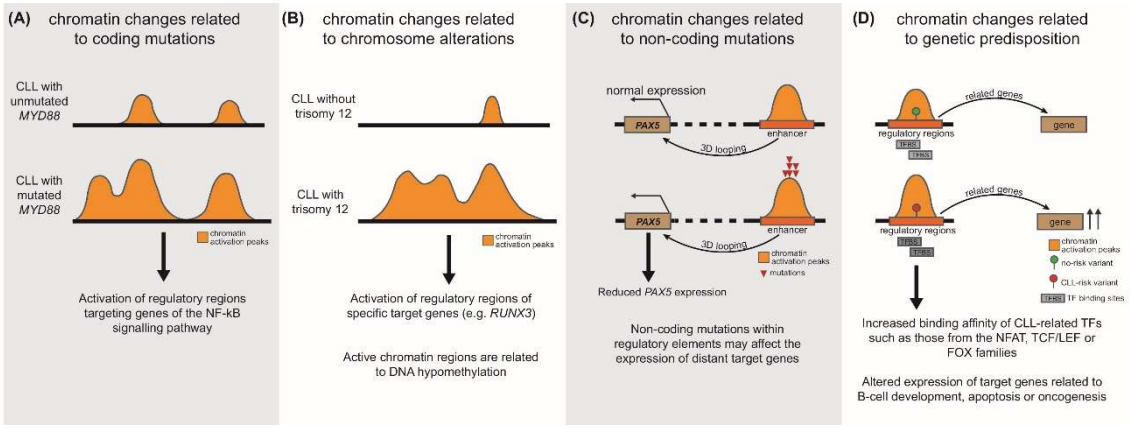


Figure 5


(A) chromatin changes related to coding mutations
CLL with unmutated MYD88
CLL with mutated MYD88
chromatin activation peaks
Activation of regulatory regions targeting genes of the NF-kB signalling pathway
(B) chromatin changes related to chromosome alterations
CLL without trisomy 12
CLL with trisomy 12
chromatin activation peaks
Activation of regulatory regions of specific target genes (e.g. RUNX3)
Active chromatin regions are related to DNA hypomethylation
(C) chromatin changes related to non-coding mutations
normal expression
PAX5
enhancer
3D looping
chromatin activation peaks
mutations
Reduced PAX5 expression
Non-coding mutations within regulatory elements may affect the expression of distant target genes
(D) chromatin changes related to genetic predisposition
related genes
regulatory regions
gene
chromatin activation peaks
no-risk variant
CLL-risk variant
TF binding sites
Increased binding affinity of CLL-related TFs such as those from the NFAT, TCF/LEF or FOX families
Altered expression of target genes related to B-cell development, apoptosis or oncogenesis