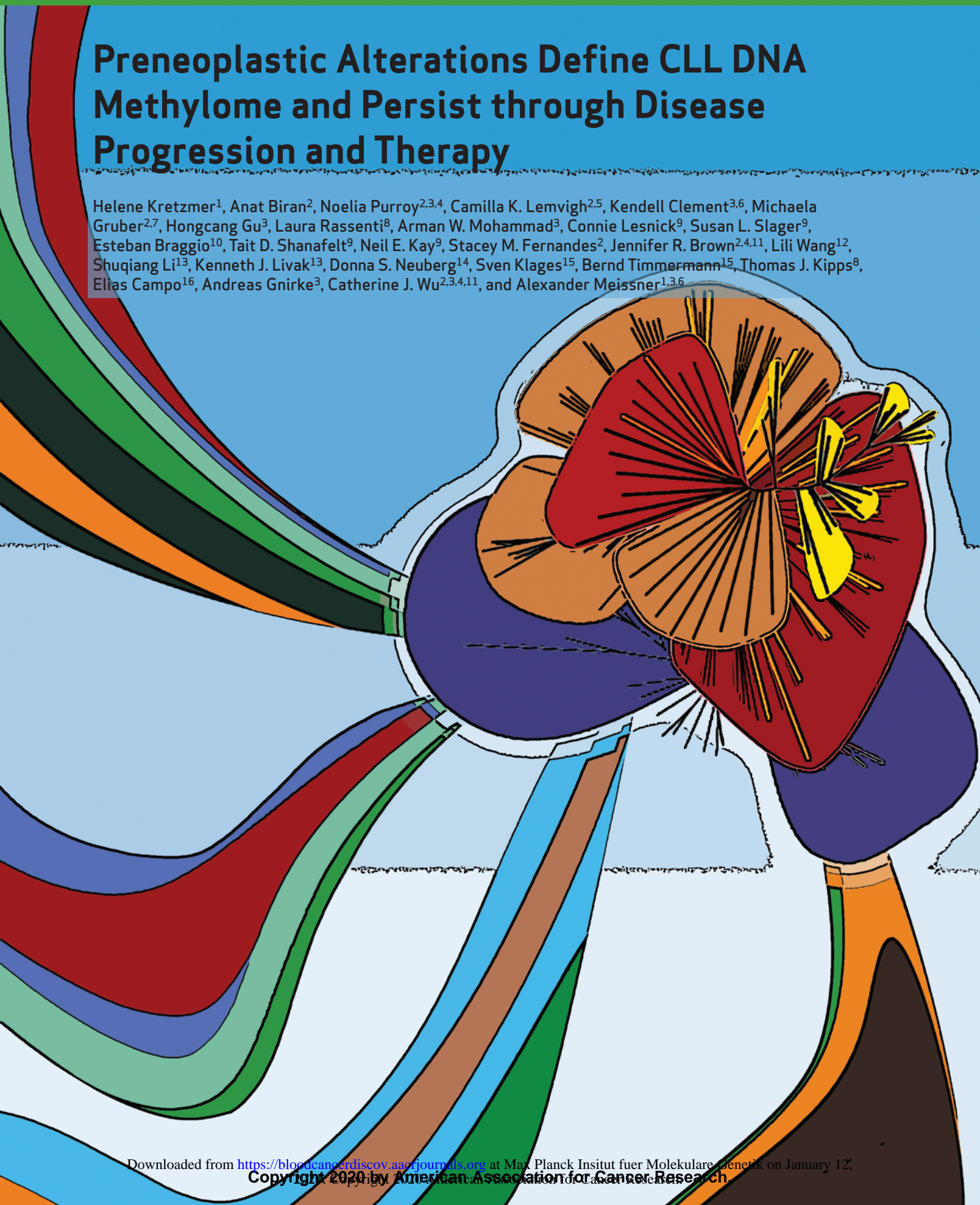


Preneoplastic Alterations Define CLL DNA Methylome and Persist through Disease Progression and Therapy

Helene Kretzmer¹, Anat Biran², Noelia Purroy^{2,3,4}, Camilla K. Lemvigh^{2,5}, Kendell Clement^{3,6}, Michaela Gruber^{2,7}, Hongcang Gu³, Laura Rassenti⁸, Arman W. Mohammad³, Connie Lesnick⁹, Susan L. Slager⁹, Esteban Braggio¹⁰, Tait D. Shanafelt⁹, Neil E. Kay⁹, Stacey M. Fernandes², Jennifer R. Brown^{2,4,11}, Lili Wang¹², Shuqiang Li¹³, Kenneth J. Livak¹³, Donna S. Neuberg¹⁴, Sven Klages¹⁵, Bernd Timmermann¹⁵, Thomas J. Kipps⁸, Elias Campo¹⁶, Andreas Gnirke³, Catherine J. Wu^{2,3,4,11}, and Alexander Meissner^{1,3,6}



ABSTRACT

Most human cancers converge to a deregulated methylome with reduced global levels and elevated methylation at select CpG islands. To investigate the emergence and dynamics of the cancer methylome, we characterized genome-wide DNA methylation in preneoplastic monoclonal B-cell lymphocytosis (MBL) and chronic lymphocytic leukemia (CLL), including serial samples collected across disease course. We detected the aberrant tumor-associated methylation landscape at CLL diagnosis and found no significant differentially methylated regions in the high-count MBL-to-CLL transition. Patient methylomes showed remarkable stability with natural disease and posttherapy progression. Single CLL cells were consistently aberrantly methylated, indicating a homogeneous transition to the altered epigenetic state and a distinct expression profile together with MBL cells compared with normal B cells. Our longitudinal analysis reveals the cancer methylome to emerge early, which may provide a platform for subsequent genetically driven growth dynamics, and, together with its persistent presence, suggests a central role in disease onset.

SIGNIFICANCE: DNA methylation data from a large cohort of patients with MBL and CLL show that epigenetic transformation emerges early and persists throughout disease stages with limited subsequent changes. Our results indicate an early role for this aberrant landscape in the normal-to-preneoplastic transition that may reflect a pan-cancer mechanism.

See related commentary by Rossi.

INTRODUCTION

In normal adult tissues, cell identity is associated with accurate maintenance of a distinct DNA methylation land-

scape (1, 2). By contrast, cells profiled from virtually every human cancer type display local hypermethylation at typically lowly methylated CpG-rich regions and simultaneously global hypomethylation at highly methylated domains (3–6).

¹Department of Genome Regulation, Max Planck Institute for Molecular Genetics, Berlin, Germany. ²Department of Medical Oncology, Dana-Farber Cancer Institute, Boston, Massachusetts. ³Broad Institute of MIT and Harvard, Cambridge, Massachusetts. ⁴Harvard Medical School, Boston, Massachusetts. ⁵Department of Health Technology, Technical University of Denmark, Lyngby, Denmark. ⁶Department of Stem Cell and Regenerative Biology, Harvard University, Cambridge, Massachusetts. ⁷Division of Haematology and Haemostaseology, Department of Internal Medicine I, Medical University of Vienna, Vienna, Austria. ⁸Division of Hematology-Oncology, Department of Medicine, Moores Cancer Center, University of California, San Diego, La Jolla, California. ⁹Mayo Clinic, Division of Hematology, Rochester, Minnesota. ¹⁰Mayo Clinic Arizona, Scottsdale, Arizona. ¹¹Department of Medicine, Brigham and Women's Hospital, Boston, Massachusetts. ¹²Department of Systems Biology, Beckman Research Institute, City of Hope, Monrovia, California. ¹³Translational Immunogenomics Laboratory, Dana-Farber Cancer Institute, Boston, Massachusetts. ¹⁴Data Science, Dana-Farber Cancer Institute, Boston, Massachusetts. ¹⁵Sequencing Core Facility, Max Planck Institute for Molecular Genetics, Berlin, Germany. ¹⁶Lymphoid Neoplasm Program, Institut d'Investigacions Biomèdiques August Pi i Sunyer (IDIBAPS), Hematopathology Section, Hospital Clínic; Departament d'Anatomia Patològica, Universitat de Barcelona, Barcelona, Spain.

Note: Supplementary data for this article are available at Blood Cancer Discovery Online (<https://bloodcancerdiscov.aacrjournals.org/>).

A. Biran, N. Purroy, C. K. Lemvigh, K. Clement, and M. Gruber contributed equally to this article.

C.J. Wu and A. Meissner jointly supervised this work.

Current address for M. Gruber: CeMM Research Center for Molecular Medicine of the Austrian Academy of Sciences, Vienna, Austria; and current address for T.D. Shanafelt, Stanford University School of Medicine, Stanford, California.

Corresponding Authors: Alexander Meissner, Max Planck Institute for Molecular Genetics, Ihnestr. 63-73, 14195 Berlin, Germany. Phone: 49-30-8413-1880; E-mail: meissner@molgen.mpg.de; and Catherine J. Wu, Dana-Farber Cancer Institute, 44 Binney Street, Boston, MA 02115. Phone: 617-632-5943; E-mail: cwu@partners.org

Blood Cancer Discov 2021;2:1–16

doi: 10.1158/2643-3230.BCD-19-0058

©2020 American Association for Cancer Research.

The striking universality of this phenomenon across cancer types raises the fundamental question of whether a cell first becomes cancerous and then acquires an aberrant methylome or if the aberrant methylome is a prerequisite. Methylation dynamics of similar proportions have otherwise only been observed during early embryonic development or the germline specification. At the same time, the generation and propagation of most other benign adult cell types show relatively stable global methylation patterns (1–7). One notable exception to the epigenetic stability of adult cell types is the maturation of B cells from hematopoietic stem cells through several intermediate stages to mature B cells, which is a critical process for the establishment of a highly effective, dynamic immune system (8). This maturation process involves genetic modulation such as somatic hypermutation of the immunoglobulin heavy-chain variable (*IGHV*) region and immunoglobulin class switch recombination (9), as well as a modulation of the methylome (10, 11). Interestingly, the methylation dynamics observed in B-cell maturation share many features with the cancer methylome (10, 11).

Chronic lymphocytic leukemia (CLL) is a malignancy of aberrant clonal mature B cells in the blood, bone marrow, and lymphoid organs that provides an ideal model setting to gain insight into the emergence of the altered methylome. Its typically indolent course enables longitudinal studies within individual patients from a pretreatment “watch and wait” phase—the duration of which is highly variable among patients, lasting months to years (12)—to the posttreatment setting and even onto progression (13, 14). A precursor stage termed monoclonal B-cell lymphocytosis (MBL) has also been described, defined as elevated white blood cell (WBC) counts with clonal B cells of a CLL immunophenotype. High-count MBL on average progresses to CLL that requires treatment in

1% to 2% of patients per year (15). A well-established prognostic factor in CLL is the mutational status of the *IGHV* region genes, with mutated *IGHV* showing a much better prognosis than CLL with unmutated *IGHV* (16, 17). The *IGHV* mutational status has been thought to reflect differences in the cell of origin, with a similarity in methylation profiles of unmutated CLLs and pregerminal center B cells, and of mutated CLL with mature, postgerminal center memory B cells, suggesting that CLL emerges from a spectrum of B cells undergoing broad DNA methylation alterations (11, 16, 18, 19). In addition to these characteristic global changes, we previously identified a pervasive local disorder of methylation across genomic features in CLL, not present in normal tissues (20). Although general changes in methylation profiles during B-cell development and cancer have been described (6, 10, 11, 20–24), little is currently known about: (i) if and which additional methylation changes are necessary to transition from normal into a preneoplastic state and further into cancer, (ii) how this altered cancer methylome is affected by therapy, and (iii) why it is found so ubiquitously across different types and stages of cancer. Furthermore, the chronologic origin of altered methylation with respect to cancer initiation and progression is not well understood but would be of relevance for early detection and could lead to novel therapeutic strategies.

To approach these questions, we used bulk and single-cell reduced representation bisulfite sequencing (RRBS; refs. 25–27) to profile normal mature B cells, as well as cells from patients in the preneoplastic MBL phase and during CLL progression, including after treatment. We characterized the methylation status of samples collected from 53 patients supplemented with WBC counts as a measure of tumor burden, and hence the effect of treatment (average sampling period of 5.7 years). Further, we used single-cell transcriptomics to complement the DNA methylation results in the patients transitioning from MBL to CLL. Our analyses reveal that changes in methylome and transcriptome are established early on, already at the precursor stage, and remain remarkably stable throughout the disease and even after therapy.

RESULTS

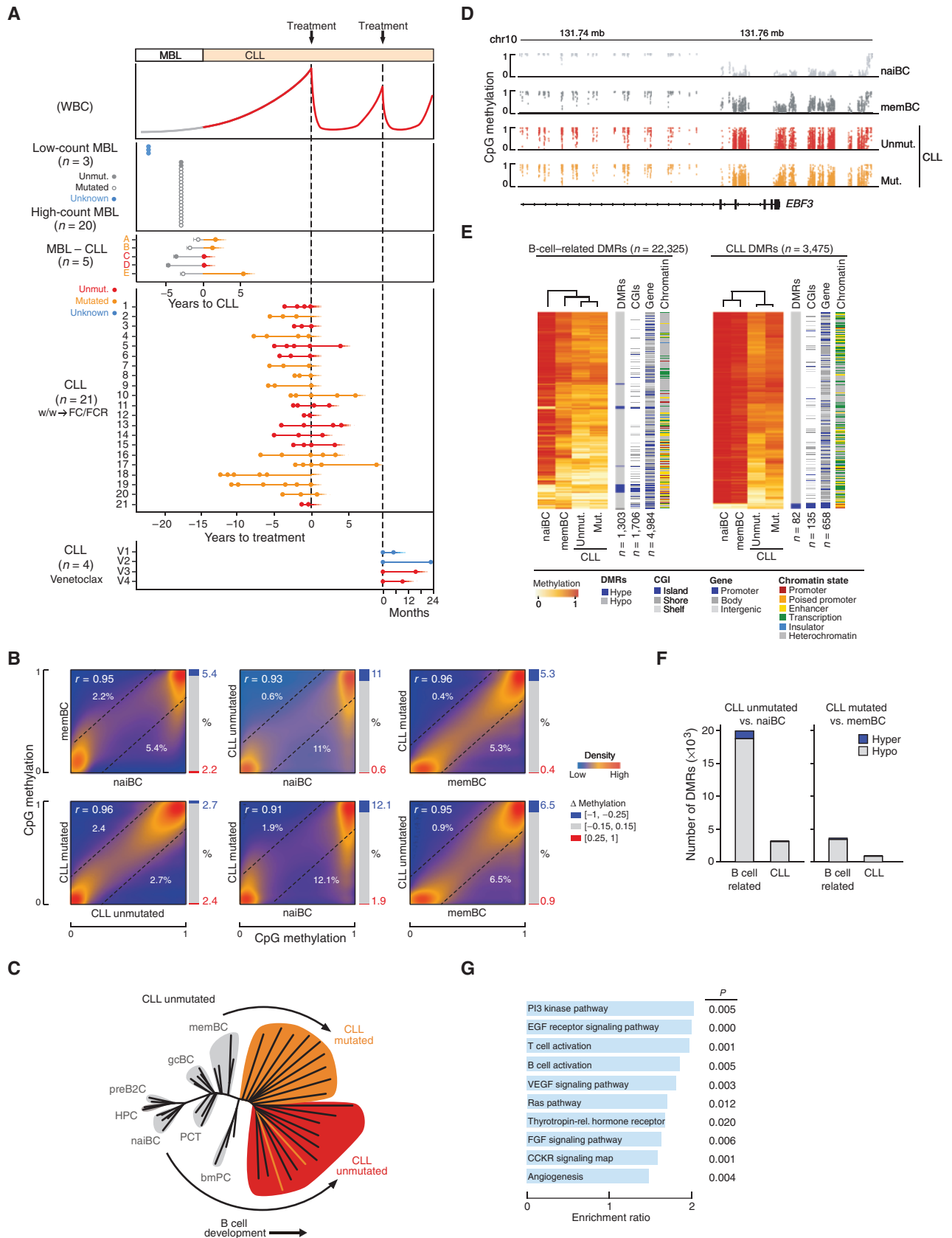
Unmutated and Mutated CLLs Converge to a Similar Methylome

To systematically study the DNA methylation dynamics across the disease course of CLL, we generated RRBS datasets from CD19⁺ CD5⁺ cells collected from 23 individuals with MBL, matched samples for 5 patients capturing both the MBL and their transition to CLL, and serial pre- and posttreatment samples from 25 patients collected following the diagnosis of CLL (28, 29) and compared these with published B-cell-lineage subpopulations (refs. 10, 30; Fig. 1A; Supplementary Table S1).

Genome-wide correlation of single CpG methylation showed a substantial similarity of unmutated and mutated methylation profiles ($r = 0.96$); however, compared with their putative cell of origin, the CLL *IGHV* subtypes showed different degrees of abnormality. Although the unmutated CLL showed more changes compared with naïve B cells, the mutated CLL exhibited a methylation landscape more similar to memory B cells than naïve to memory B cells (Fig. 1B). As noted above, CLLs originate from a range of developmental stages with pregerminal center B cells thought to give rise to unmutated CLL and mature, postgerminal center memory B cells to mutated CLLs (10, 11, 16, 18, 19, 31). Evaluation of single samples in a phylogenetic tree analysis revealed that the unmutated and mutated CLL samples are characterized by a methylome that consistently differs from normal naïve and memory B cells, suggesting a convergent disease-associated methylome, irrespective of *IGHV* mutation status (Fig. 1C; Supplementary Fig. S1A). Together, these results suggest that both *IGHV* subtypes of CLL undergo methylation changes specific to CLL. However, some of these changes also appear to be normally acquired during B-cell maturation, as observed in the example of the *EBF3* locus (Fig. 1D).

To more systematically evaluate regions that are consistently altered in CLL, we identified differentially methylated regions (DMR) between (i) unmutated CLL versus naïve B cells ($n = 23,206$ DMRs) and (ii) mutated CLL versus memory B cells

Figure 1. CLL methylation signatures distinguish CLL from normal B cells. **A**, Schematic representation of progression from the precursor state of MBL to CLL, depicting the extended period of “watch and wait” (w/w) until first treatment and an overview of the patient cohort from which 109 samples were collected to generate RRBS data. Combination chemoimmunotherapy (CIT; FC/FCR) typically leads to a rapid decrement in WBC counts. Our cohort included samples from 23 MBL, 5 paired samples of MBL and CLL, and 25 CLL patients. More specifically, the MBL samples are $n = 20$ high and $n = 3$ low count [2 with unmutated *IGHV* ($\geq 98\%$ homology with germline sequence), 18 with mutated *IGHV* ($< 98\%$ homology with germline sequence), 3 with unknown status], and the CLL are $n = 21$ CIT treated (red: *IGHV* unmutated status; orange: *IGHV* mutated status) and $n = 4$ venetoclax treated, after having already progressed from first-line (fludarabine-based) regimens. Each circle indicates a sample collected. **B**, Correlation of CpG methylation levels between naïve B cell, memory B cells, and unmutated and mutated CLL at CpG-level resolution (purple, low density; orange, high density). The methylome of memory B and CLL cells, in contrast to naïve B cells, is strongly hypomethylated and shows hypermethylation in otherwise lowly methylated regions. Bar charts give fraction of hypermethylated (> 0.25) and hypomethylated (< 0.25) CpGs. $N = 3,490,971$ (mem-naï); 3,202,573 (unmut-naï); 3,034,005 (mut-naï); 3,202,573 (unmut-mem); 3,034,005 (mut-mem); 2,974,458 (mut-unmut). **C**, Phylogenetic CpG methylation tree of normal B cells (gray shading) and first time point samples of patients with CLL [*IGHV*: unmutated (red) and mutated (orange)] in the context of normal B-cell differentiation states. All CLLs cluster to the more mature end of the tree and separate by the mutational status of the *IGHV* chain genes except for two cases (orange lines). Each line represents a sample. Arrows indicate presumed cell of origin to CLL transition. bmPC, bone marrow plasma cell; gcBC, germinal center B cell; HPC, hematopoietic progenitor cell; PCT, plasma cell from tonsil; preB2C, pre-B-II cell. **D**, Average CpG methylation from bulk in naïve B cells (light gray), memory B cells (dark gray), and unmutated (red) and mutated CLL (orange) across the *EBF3* locus. Both CLL samples exhibit similar levels of modulation across the entire region. Specifically, both CLL samples reveal stronger hypermethylation in the promoter region than the normal B cells and loss of methylation in the usually highly methylated gene body. Although to a much lesser degree, this effect can also be found between naïve and memory B cells. **E**, Average CpG methylation levels for CLL DMRs of unmutated CLL versus naïve B cells and mutated versus memory B-cell comparisons. The rows represent overlapping and mutated or unmutated CLL-specific DMRs. Samples are merged into a mean methylation representation per group (columns) and DMR (rows). Rows were ordered using unsupervised hierarchical clustering. Side annotations for DMR location and chromatin state: hyper- or hypomethylated DMR; CpG density; CpG island, shore, or shelf; location: promoter, gene body, or intergenic; chromatin state: promoter, poised promoter, enhancer, transcription related, insulator, or heterochromatin. **F**, Numbers of hypomethylated (gray) and hypermethylated (blue) DMRs (minimum difference of 0.25, minimum 3 CpG in length) in unmutated CLL versus naïve B-cell and mutated versus memory B-cell comparisons. DMRs are classified as B-cell related or CLL based on their overlap with CpGs that are differentially methylated during normal B-cell development. **G**, Overrepresentation enrichment analysis for genes with CLL DMRs compared with the background, i.e., all DMRs. Enriched pathways (Panther) include PI3K, EGFR, Ras, FGF, and CCKR signaling. A common characteristic of these pathways is their implication in cell survival, gene expression regulation, growth factors, activation of proliferation, and cell invasion. Shown are the top 10 pathways based on P value.



($n = 4,653$ DMRs; Supplementary Table S2; ref. 32). To disentangle methylation changes associated with normal cell lineage-specific differentiation from potentially cancer-related changes, we classified the aggregate of these two sets of DMRs as B-cell related ($n = 22,325$) or CLL ($n = 3,475$), depending on whether they were classified as dynamically changing during normal B-cell development (Fig. 1E and F; Methods; ref. 30). The majority (85%) of the DMRs overlapped with developmental regions, whereas 15% were classified as CLL DMRs (Fig. 1E and F; Supplementary Fig. S1B). Based on the clustering, B-cell lineage-related DMRs showed a gradual shift, mostly toward hypomethylation, from naïve to memory and both CLL subtypes, reflecting the normal B-cell developmental changes that are retained in CLL. In contrast, as expected, the set of CLL DMRs readily distinguished normal B cells from CLL (Fig. 1E). Moreover, genes that were associated with CLL DMRs were found to be overrepresented among pathways related to cell growth and survival, proliferation, and neoplastic transformation, suggesting possible regulatory relevance (Fig. 1G).

We additionally confirmed the DMRs to be a distinctive feature between normal B and CLL cells by analyzing replicates of CD5-positive and -negative naïve and memory B cells from a set of three healthy donors. Genome-wide phylogenetic tree clustering and the correlation of methylation rates revealed two major clusters, separating the samples by naïve and memory B cells but not by CD5 status (Supplementary Fig. S2A and S2B). Based on the presence of the CLL-specific DMRs, CD5-sorted healthy donor samples were found to cluster into the group of previously published reference B cells, hence demonstrating similar methylation in DMRs independent of CD5 status of naïve or memory B-cell state (Supplementary Fig. S2C).

CLL Methylome Remains Mostly Unchanged after Treatment

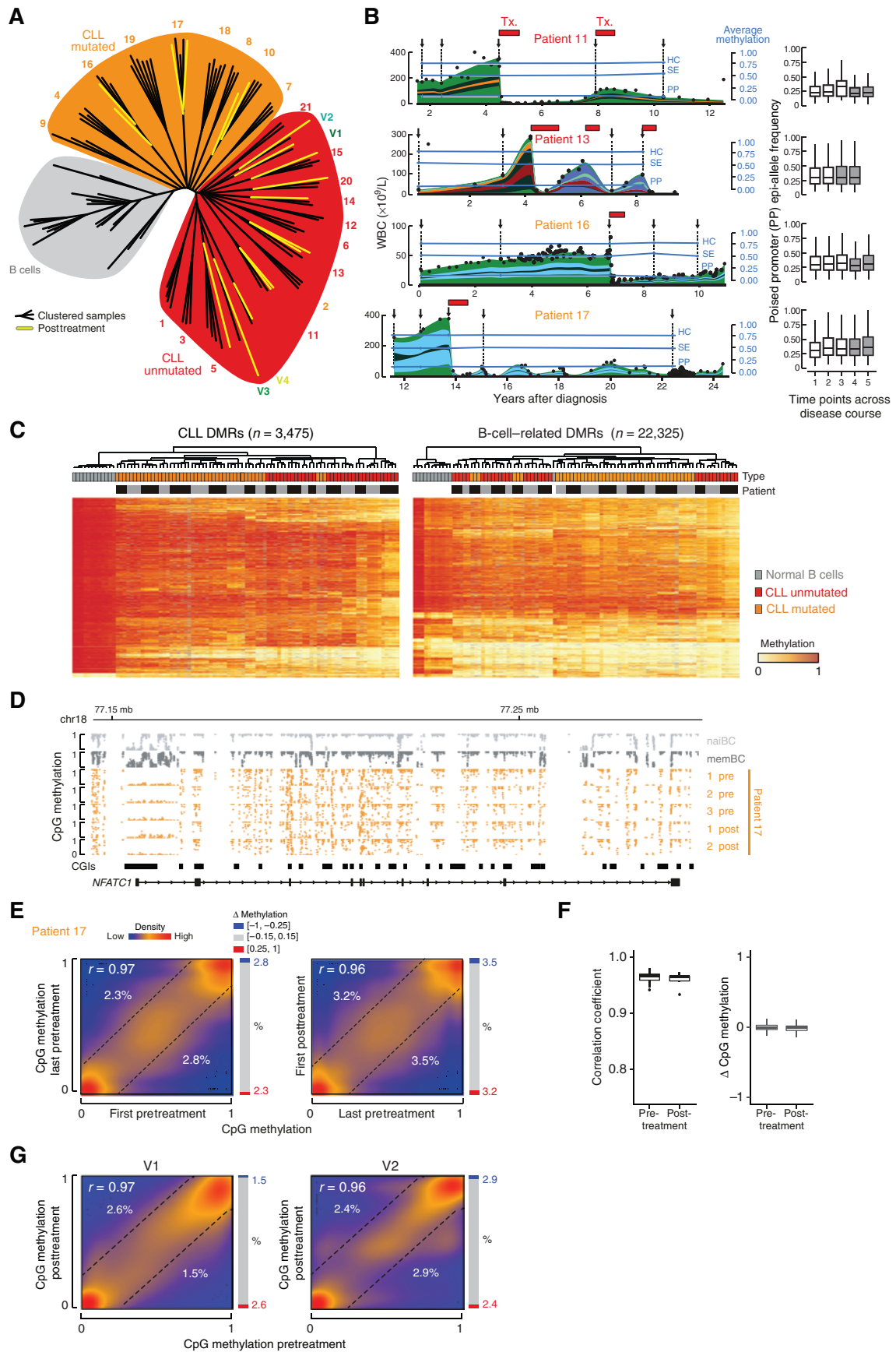
To evaluate the stability of these DMRs and the dynamics of the CLL methylome over time, we analyzed longitudinal

samples collected during natural CLL progression. CLL allows that leukemic burden can be approximately estimated by measuring the WBC count over time since it is, for many patients, primarily a circulating malignancy. To study the CLL methylome before and after the first treatment, we performed unsupervised phylogenetic clustering of the pre- and posttreatment (fludarabine, cyclophosphamide, and rituximab, FCR) samples of patients. Interestingly, we found no consistent methylation differences that separate pre- and posttreatment samples, and also no WBC-related effects could be seen in the clustering (Fig. 2A). Next, we compared methylation levels across patient time points for selected chromatin states derived from published data of the lymphoblastoid cell line GM12878 (33). Despite vastly different growth patterns and subclonal dynamics (defined by prior genetic characterization; ref. 28), global methylation levels of various genomic features, such as heterochromatin, strong enhancers, and poised promoters, for serial samples from all 21 patients remained stable and were consequently independent of the dynamic changes in WBC counts (Fig. 2B, left; Supplementary Fig. S3A and S3B).

We observed substantial posttreatment reduction in WBCs creating population bottlenecks for the nine patients following treatment with FCR (Fig. 2B; Supplementary Fig. S3A). However, this was not associated with any notable DNA methylation changes over the three representative genomic features or the number of distinct epialleles present (Fig. 2B; Supplementary Fig. S3; ref. 34). Indeed, methylation levels were mostly independent of detected subclonal genetic evolution and from patterns of growth. Similar stability was also found at the previously identified DMRs, with once acquired changes appearing to persist during CLL progression (Fig. 2C). These deregulated regions, including the example of *NFATC1* (Fig. 2D), further highlight the remarkable stability of methylation patterns.

We also compared the nine patients by focusing on the clinically most divergent time points, i.e., first and last

Figure 2. The CLL methylome remains mostly unchanged over disease progression, including after treatment. **A**, Phylogenetic tree of normal B cells and all measured time points of patients with CLL ($n = 83$) using global CpG methylation levels. Each line represents a sample; subtrees are multiple samples of the same patient. Yellow lines represent posttreatment samples (chemoimmunotherapy and venetoclax V1–V4). Mutated and unmutated are colored as before, and all patient numbers are shown next to the respective branches. All samples from the same patient clustered together, whereas normal samples are distinct from the CLL cohort. **B**, WBC counts and methylation dynamics for selected genomic features to represent global hypomethylation (HC, heterochromatin; SE, strong enhancers) and hypermethylation (PP, poised promoters) across disease progression for patients 11, 13, 16, and 17 (for all others, see Supplementary Fig. S3A). The methylation levels remain constant over time and after treatment. Black dots: WBC counts (left axis). Blue lines and dots: measurements of CpG methylation levels (right axis). Black arrows and dashed lines indicate collected time points for DNA methylation analysis. Boxplots to the right display the coverage normalized epiallele fraction in poised promoter regions. Treatment exposure at time points is indicated as shaded boxplots. In addition to the methylation level, the epiallelic fractions' distribution stays stable over time and after treatment. WBC plots are taken from the same patients studied in ref. 28 and have been overlaid with our DNA methylation data. Distinct genetically defined subclones are indicated with the different colors. Tx., treatment. **C**, Left, average methylation levels per sample (columns) for CLL DMRs (rows) of unmutated CLL versus naïve B cells and mutated versus memory B-cell comparisons. Rows were ordered using unsupervised hierarchical clustering. Right, average methylation levels per sample (columns) for B-cell-related DMRs (rows) of unmutated CLL versus naïve B cells and mutated versus memory B-cell comparisons. Rows were ordered using unsupervised hierarchical clustering. **D**, Average CpG methylation in naïve B cells (light gray), memory B cells (dark gray), and five serial samples collected from patient 17 (orange, top to bottom: three pretreatment and two posttreatment samples) across the *NFATC1* locus. Dots represent CpG-level methylation of each sample. **E**, Correlation of CpG methylation levels in the first pre- and last pretreatment sample as well as last pre- and first posttreatment of patient 17 at CpG-level resolution ($n = 1,912,382$). For all other samples, see Supplementary Table S3. Bar charts give fraction of hypermethylated (>0.25) and hypomethylated (<0.25) CpGs. Numbers are given within the scatter. **F**, Boxplot of correlation coefficients of genome-wide CpG-level correlation between the first pre- and last pretreatment samples of all patients with CLL as well as last pre- and first posttreatment samples for the posttreatment CLL (left). Corresponding boxplot of genome-wide CpG-level difference between the first pre- and last pretreatment samples as well as last pretreatment to first posttreatment samples (right). $n = 21$ pretreatment data points and $n = 9$ posttreatment data points. In the boxplots, the centerline is median; boxes, first and third quartiles; whiskers, $1.5\times$ interquartile range; data beyond the end of the whiskers are omitted. **G**, Correlation of CpG methylation levels in the pre- and posttreatment samples of the venetoclax-treated patients V1 and V2 at CpG-level resolution. For all other samples, see Supplementary Fig. S4. Bar charts give fraction of hypermethylated (>0.25) and hypomethylated (<0.25) CpGs. Numbers are provided within the scatter.



pretreatment and last pre- and first posttreatment. No joint DMRs between all first pretreatment versus last pretreatment time points could be detected. The variability between samples of the same patient and the lack of shared events appear to be more in line with patient-specific evolution than a common path across patients. Moreover, correlation analysis on the CpG level also confirmed a largely stable methylome across CLL evolution and even after treatment (Fig. 2E and F).

To explore if the observed methylation stability is therapy specific, we next analyzed four patients treated with the BCL2 inhibitor venetoclax (35). As with the FCR chemoimmunotherapy-treated patients, these CLL samples collected before and after venetoclax exposure clustered tightly together, within the group of chemoimmunotherapy-treated patients (Fig. 2A and G). We further confirmed the stability of the B-cell-related and CLL DMRs in this treatment cohort, in which we could only detect globally on average less than 6% of CpGs to vary between pre- and post-venetoclax treatment (Supplementary Fig. S4A–S4D).

Combined with the early emergence of the altered methylation landscape, our posttreatment results highlight the striking stability of the CLL methylome, with minimal changes over disease progression, including after treatment.

Variations in the CLL Methylome Appear Stochastic among Patients

Because only a few patient-specific methylation dynamics were observed, we assessed if their occurrence exceeded random dynamics present among normal B-cell subtypes. Focusing on the chemoimmunotherapy-treated patient samples, we first compared the number of dynamic CpGs between first and last pretreatment, and last pre- and first posttreatment CLL samples with differences between biological replicates of naïve and memory B cells (Supplementary Fig. S5A; Supplementary Table S3). Although we did not detect any correlation with the time to treatment, we observed the fraction of dynamic CpGs to be slightly higher in posttreatment samples. Overall, only 1 of 21 patients stood out with higher variability ($r = 0.94$ pretreatment and $r = 0.93$ for pre- to posttreatment comparison); however, this is very similar to the variation observed at the transition of naïve to memory B cells ($r = 0.95$). Moreover, the relative number of CpGs that exhibited substantial differences was less than 5% of all considered CpGs for most CLL cases, and those CpGs were frequently located in heterochromatin regions and outside of CpG islands (Supplementary Fig. S5B). Most of the dynamic CpGs were restricted to individual patients, and 99.9% of CpGs were shared within a maximum of six and four patients for pre- and posttreatment comparisons, respectively (Supplementary Fig. S5C), thus again confirming limited variation and a high degree of stability of CLL methylome over time.

To further separate random single dynamic position events from consistently altered regions during disease progression and following treatment, we focused on patient-specific DMRs identified either between first and last pretreatment for all 21 patients (Supplementary Fig. S5D, left), or between the last pre- and first posttreatment for the nine patients (Supplementary Fig. S5D, right). We found a median of 106 pre- and 143 posttreatment DMRs per sample. Of these, the

vast majority (89% and 86%) appeared in regions shared with normal B-cell development. The remaining DMRs comprised only 8% of the aforementioned dynamic CpGs. However, with 50% of the dynamic CpGs still localized to CLL-specific regions, our data suggest the presence of randomly modulated individual CpGs rather than stretches of adjacent CpGs. Furthermore, about half of the CLL DMRs were located in heterochromatin, supporting the assumption that these may represent a secondary effect (Supplementary Fig. S5E). Finally, gene set enrichment analysis revealed pathways supported by only a few genes (i.e., B-cell and T-cell pathways supported by three genes and p53 by two genes) or pathways with no apparent link to CLL (Supplementary Fig. S5E).

In sum, although a low number of patient-specific methylation changes accompany the individual tumor evolution, we observed remarkable stability and similarity of the acquired CLL methylome across patients.

The Altered Methylome Is Already Detectable at the MBL Stage

Because we observed an altered methylome already present within the first time points of the characterized CLL specimens, we next turned to specimens collected from our patients with MBL to evaluate the cancer precursor methylome. We again performed unsupervised phylogenetic clustering but now including the high-count MBL samples. Despite clinical classification as a precursor state, all of the MBL cases were found to cluster directly among the group of CLL samples and not to branch earlier from the trunk of this tree (Fig. 3A). Most strikingly, all five matched MBL–CLL cases appeared to be as similar to each other as the biological replicates of different healthy B-cell types and much more similar to each other than to the other CLL cases. Thus, patient-specific methylation signatures appeared to be stronger than any preleukemic versus leukemic methylation signature, which would have otherwise resulted in the separate clustering of the MBL samples from the CLL samples.

Strikingly, the identified CLL-associated DMRs were also present in our patients with MBL (Supplementary Fig. S6A and S6B). We extended the analysis to search for additional consistently occurring methylation changes that could potentially drive the MBL-to-CLL transition. However, no statistically significant DMRs could be detected between the methylomes of the individuals that transitioned from MBL to CLL. As a representative example of this shared landscape between MBL and CLL, we show the methylation patterns for the gene *NFATC1*, which has been reported as overexpressed in CLL due to loss of epigenetic repression (36) and is an upstream effector of *BCL2*, which itself is frequently deregulated due to chromosomal translocations in B-cell malignancies (Fig. 3B). Through a correlation analysis at single CpG resolution of the methylomes of the matched MBL and CLL pairs, we further observed the striking similarity between MBL and CLL methylomes. These results revealed only minor, if any, targeted remodeling of the methylome between the precursor and CLL stages within a given patient (Fig. 3C; Supplementary Fig. S6C). To appreciate this high similarity, we note that comparably high correlations are otherwise found between biological replicates of flow cytometrically isolated normal B-cell subpopulations.

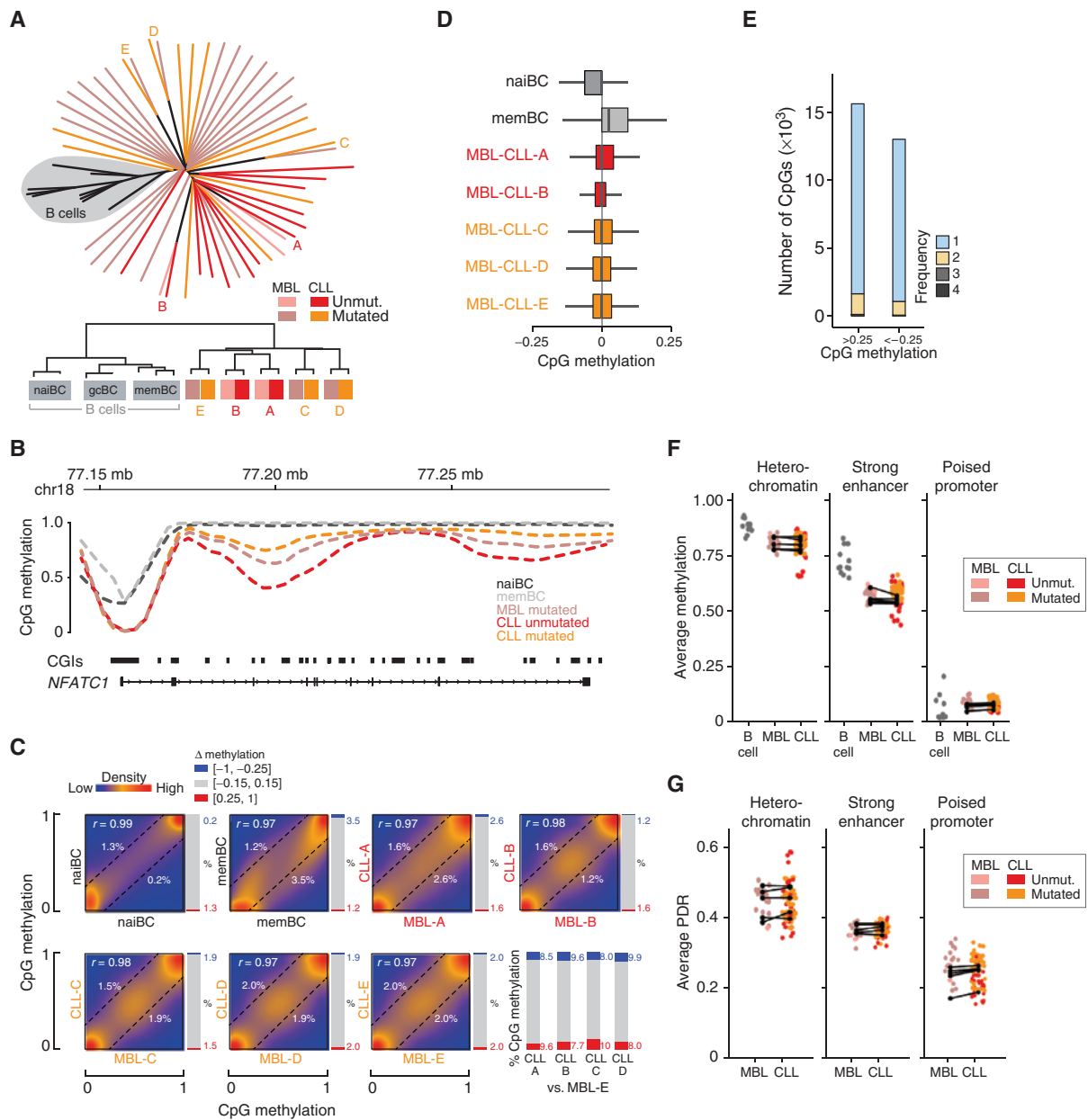


Figure 3. CLL methylation signatures are already present in MBL. **A**, Phylogenetic tree of normal B cells, MBL, CLL (first pretreatment sample per case), and matched MBL-CLL pairs using all CpGs ($n = 3.5$ million). Left, all MBL cases clustered together with CLL cases. Each line represents a sample. Right, unsupervised hierarchical clustering of normal B cells and matched MBL-CLL pairs. Matched MBL-CLL cases clustered as closely as biological replicates of normal B cells; mean joint fork distance ($\log_2 FC$) of matched versus mismatched MBL-CLL versus normal B cells: -0.03 versus -1.06 . gcBC, germinal center B cell. **B**, RBS-based genome browser tracks. Average methylation is shown as a smoothed line across the *NFATC1* locus, grouped by naïve B cells (light gray, $n = 3$ samples), memory B cells (dark gray), unmutated (red) and mutated CLL (orange), and MBL (brown). Naïve and memory B cells are hypermethylated in the gene body, with only a small drop at the transcription start site. Unmutated and mutated CLL are hypomethylated across the promoter and gene body. **C**, Correlation of CpG methylation levels in matched MBL-CLL pairs ($n = 5$), in biological replicates of normal B cells (naïve B cells and memory B cells), at CpG-level resolution ($n = 1,801,907$ CpGs). Plots show a high correlation between biological replicates of normal B cells and a similarly high correlation between MBL-CLL pairs. Numbers give quantification of hypermethylated (>0.25) and hypomethylated (<0.25) CpGs. For comparison of CpG differences with mismatched samples, the quantification of all CLL samples versus one MBL is shown. $r = 0.97-0.98$ for matched pairs, $r = 0.85-0.9$ for mismatched pairs, $r = 0.99$ for naïve B cells, and $r = 0.97$ for memory B cells. **D**, CpG-level resolution differences between matched MBL-CLL pairs and among biological replicates of normal B cells. For all MBL-CLL comparisons, a maximum of 4.2% of positions show a difference of >0.25 ($n = 1,801,907$ CpGs). Methylation differences among biological replicates were less than 2% (naïve B cells) and 3.6% to 9.5% (memory B cells) of CpGs with a difference of >0.25 . Per boxplot, the median value is indicated by the centerline, with first and third quartiles as outlines of boxes, and $1.5\times$ interquartile range as whiskers; data beyond the end of the whiskers are omitted. **E**, Number of CpGs with a difference of >0.25 in MBL-CLL comparisons and frequency of recurrent observations across the five pairs (light gray, unique for one pair, to dark grey, observed in four of five pairs, no five of five detected). The minority of CpGs are recurrently differentially methylated. **F**, Comparison of average chromatin state methylation among DNA from normal B cells, MBL, and CLLs. Black horizontal lines, matched MBL-CLL pairs ($n = 1,801,907$ CpGs). **G**, Comparison of chromatin state proportions of discordant methylation rates (proportion of discordant reads, PDR) between MBL and CLLs. Black horizontal lines, matched MBL-CLL pairs.

Also, at single CpG resolution, we found that at most 4% of all covered CpGs showed a difference of 0.25 or greater between MBL and CLL samples of the same patient (Fig. 3D). An expanded analysis to examine whether individual CpGs were conserved targets across patients revealed this is not the case (Fig. 3E). Finally, we compared methylation and read-discordance levels for different chromatin states. This showed that differences and variability compared with normal B cells affect MBL and CLL cases to the same degree, again highlighting that the similarity of MBL and CLL is not merely based on patient identity (Fig. 3F and G).

Lastly, as most of our patients with MBL had already relatively elevated WBC counts, we further extended our investigation to include CD5⁺ sorted cells from three patients with low-count MBL (Supplementary Fig. S7A). Although the signal is expectedly not as strong due to the rare proportion of cells sequenced and the potential contamination with CD5⁺ normal B cells, we do detect evidence of the same characteristic epigenetic alterations as observed in high-count MBL and CLL (Supplementary Fig. S7B–S7D).

Taken together, our comprehensive analysis of 53 patients and 101 pretreatment RRBS datasets suggests that the transition to the cancer methylome occurs early in the disease. Analysis of patient-matched MBL and CLL shows that no consistent additional DNA methylation changes are seemingly associated with disease progression.

Heterogeneous Expression Patterns Are Present Per Patient but Are Stable across Natural Disease Progression

Complementing our methylation analysis, we profiled the transcriptomes of approximately 60k single cells isolated from healthy donors and the five matched MBL–CLL specimens (Fig. 4A). Unsupervised clustering revealed nine distinct clusters: four clusters representing peripheral blood mononuclear cells of the two healthy donors and the remaining five, with each representing one of five patient B cells (Fig. 4B). From the healthy donors, cell clusters of myeloid and lymphoid origin were readily identifiable based on their marker gene expression. In contrast, the five clusters from patients were identified as CLL/MBL-mixed clusters that showed expression of some B-cell marker genes, although less pronounced (Fig. 4C). When looking more specifically at differentially expressed genes, we found lower and heterogeneous expression for some characteristic B-cell markers and similarly heterogeneous upregulation of genes such as KLF2 and CD27 in the patient samples (Fig. 4D). Of note, the MBL and CLL cells per patient were transcriptionally indistinguishable.

Because the transcription-based clustering could not distinguish the MBL and CLL cells, we instead used barcode information per cell for annotation (Fig. 4E). We observed a remarkable overlap for most clusters, supporting the striking similarity of the MBL and CLL transcriptomes. Only the MBL and CLL cells from patient A were slightly separated in the UMAP visualization. However, upon evaluation of the highest-ranked marker genes for each MBL–CLL set, we found a surprisingly high concordance of expression of even the most differential expressed genes between MBL and CLL cells (Fig. 4F).

Based on the single-cell expression profiles, the differences between MBL and CLL appear to be marginal, which agrees with the lack of separation by clustering on single-cell transcriptomes. Combined with the lack of DNA methylation changes in the MBL-to-CLL transition, it points to an earlier molecular event that sets the normal cells already on the path to tumorigenesis.

Individual MBL and CLL Cells Show Little DNA Methylation Heterogeneity

Our bulk data indicated an early conserved switch in the DNA methylation landscape across patients with CLL, and our single-cell transcriptome data demonstrate the transcriptional similarity between matched MBL and CLL. However, bulk measurements cannot completely distinguish the contributions of diverse cellular subpopulations to the overall picture. Subclonal evolution and genetic heterogeneity are common in CLL (37–39). This understanding motivated us to investigate single-cell methylation maps from two patients with CLL, two patients with MBL, and age-matched B cells collected from two healthy adult volunteers, uncovering a stable level of mean methylation per cell on a global scale (Fig. 5A; Supplementary Table S1; ref. 40).

Analysis of our previously defined DMRs showed the presence of aberrant methylation levels in all MBL/CLL cells with sufficient coverage (Fig. 5B). When comparing naïve to memory with MBL and CLL cells, a gradual gain of methylation in B-cell-related and CLL hyper-DMRs was observed. Conversely, hypomethylated B-cell-related and CLL DMRs appeared slightly stronger in separating normal from diseased (MBL and CLL) cells (Fig. 5B). Phylogenetic clustering separated CLL, memory B, and naïve B cells, with no differences between the B-cell subpopulations with or without the presence of CD5 (Fig. 5C; ref. 12). Moreover, we observed a clear separation between MBL and CLL versus normal, with each forming a tight cluster in line within the observed stability of the methylome per patient. Of note, memory B cells, despite many shared features with the CLL methylome, cluster distinctly next to the naïve B cells and apart from the MBL and CLL cells.

Our genome-wide single-cell methylation analysis thus complements our bulk data by further showing the clear methylation difference between MBL and CLL compared with sorted B-cell subtypes.

DISCUSSION

We show that the aberrant cancer methylome in CLL is already established at the preneoplastic MBL stage and is consistently present at the time of diagnosis across samples collected from 3 low-count and 20 high-count MBL, 5 matched MBL–CLL pairs, and 25 patients with CLL. Although normal B-cell maturation shows some similarities with the CLL methylome, these normal developmental changes are likely insufficient to transform cells into proliferative MBL and CLL. Nonetheless, the shared targets make a better understanding of the underlying mechanism and biological reason highly relevant. We also find a limited set of cancer-specific targets that can be readily applied to distinguish all normal B-cell subtypes from MBL and CLL. These CLL DMRs are

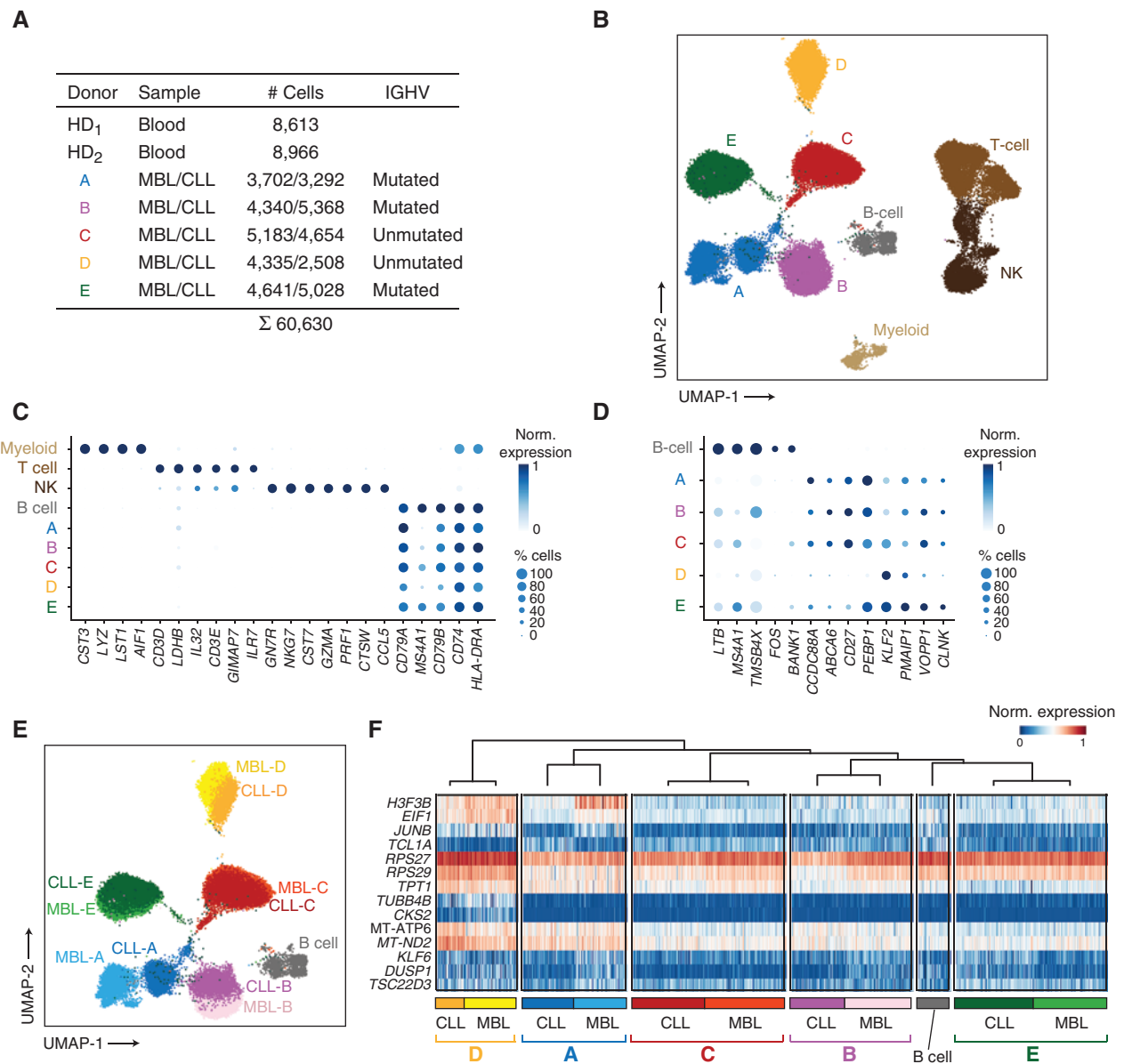


Figure 4. Single-cell transcriptome analysis. **A**, Summary table with details of donors, tissue source, number of cells per sample, and *IGHV* status. Patients were profiled in MBL as well as the CLL state using the 10x Genomics Chromium droplet single-cell RNA sequencing. HD, healthy donor. **B**, UMAP displaying the groups found using the Louvain algorithm. The healthy donor cells split into B cells, T cells, myeloid cells, and natural killer (NK) cells. The MBL and CLL cells of patients build distinct groups but are within a patient not distinguishable. **C**, Normalized gene expression level and the number of positive cells of marker genes used to identify normal cell types in **B**. B-cell-specific genes show an aberrant expression profile in clusters derived from patient cells. **D**, Normalized gene expression level and the number of positive cells of genes identified as marker genes between B cells and all patient cells (Wilcoxon rank-sum test). **E**, UMAP of MBL, CLL, and B cells with artificially introduced identification of MBL and CLL cells. MBL and CLL cells cluster farther apart from the B cells than from other cells of the same patient and are highly overlapping for almost all patients. **F**, Heatmap displaying single-cell gene expression of highest-ranking marker genes between MBL and CLL cells of the same patient (Wilcoxon rank-sum test). Expression levels are very similar among cells of the same patient as compared with other patients or B cells, a parameter that is also supported by hierarchical clustering.

overrepresented among pathways involved in proliferation, cell survival, and growth. Although it remains technically challenging to experimentally explore if, for instance, the addition of just these CLL DMRs alone is sufficient to drive the tumorigenic transition or facilitate extended and rapid proliferation, we anticipate that these targets are certainly worthy of future exploration, including with emerging epigenome editing tools.

Our results provide a comprehensive picture of the DNA methylation alterations in MBL and CLL and demonstrate that the switch to an abnormal landscape has consistently occurred before any of our measured time points. This notably expands findings from prior array-based studies (41, 42) and complements recent work on the genetic evolution across the 21 CLL samples (28). Similar early alterations of the methylome have also been noted in colorectal

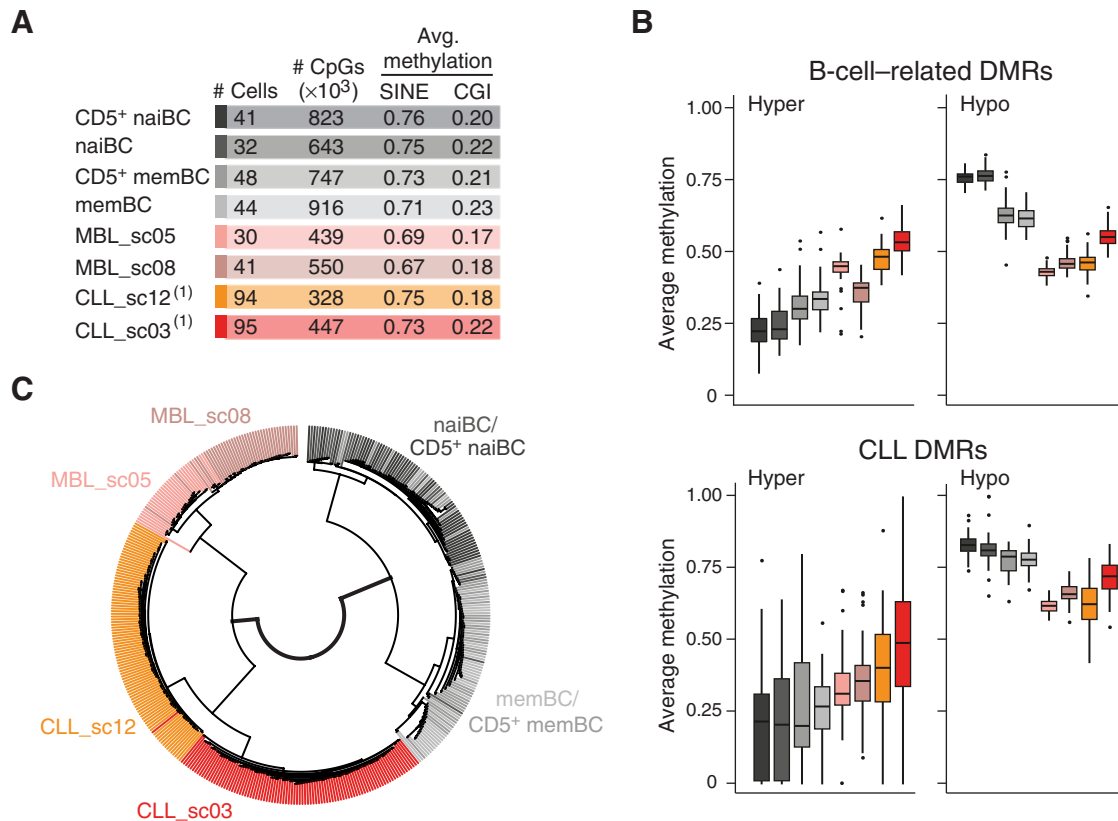


Figure 5. Single-cell DNA methylation analysis of MBL, CLL, and normal B cells. **A**, Summary table of methylation data generated from MBL (rose, brown), CLL (red, orange), and flow cytometrically isolated normal B-cell subpopulations [naïve B cells, CD19⁺CD27⁻ (lightest gray), CD5⁺ naïve B cells (CD19⁺CD27⁻, light gray), memory B cells (CD19⁺CD27⁺, dark gray), CD5⁺ memory B cells (CD19⁺CD27⁺, darkest gray)]. $n = 611,452$ CpGs covered on average. ⁽¹⁾Data for CLL were taken from Gaiti and colleagues (40). **B**, Average methylation levels for B-cell-related and CLL DMRs per cell confirm aberrant methylation of these regions being consistently observed across all cells. CLL and MBL cells show strong patterns of hypomethylation, whereas the small number of hypermethylated DMRs ($n = 888$ and 70, respectively) is already present in MBL but also seems to be slightly more prevalent in CLL cells. **C**, Unsupervised hierarchical clustering of MBL, CLL, and normal B cells. MBL and CLL cluster together into one clade. Although the clonal MBL and CLL separate by donor, naïve and memory B cells are intermingled with their CD5-positive counterparts.

cancer where the aberrant methylation landscape is already detectable in premalignant colorectal adenomas and amplifies upon the colorectal cancer state (43). Taken together with the near universality of the altered cancer methylome (5), the possibly conserved early emergence points to an important role for the epigenetic change in a tumorigenic transition. Although it is difficult to establish causality, we speculate that the altered landscape may provide a receptive platform for the disease progression. Alternatively, the cancer methylome may simply be a consequence of a developmental program that may regulate numerous cellular attributes, including methylation (5). In the latter case, it may well be other features driving the tumorigenic transition, and the methylome is only one biological consequence of the entire program. Although this seems possible, it should be noted that this altered DNA methylation landscape is maintained across patients sometimes for decades and found in nearly all cancer types, raising the question of why it is not diverging if it has no functional role. Another possibility that we can consider here is that the altered methylome presents an optimized epigenetic state to maintain viability with maximum proliferation and the minimal energy requirement for DNA methylation maintenance.

Aside from these considerations, we note that the MBL and CLL methylome and transcriptome are extremely stable once acquired. In contrast to the dramatic fluctuations in tumor burden (estimated by the changes in the level of WBC counts) across disease course, methylation levels are not consistently affected by clonal expansion or treatment-induced bottlenecks. The latter may reflect that cells surviving treatment represent either the average of all subclones or that limited methylation heterogeneity is present across all subclones. From a practical standpoint, the stability of the methylome in patients with CLL limits its utility to track disease progression. Still, it may be valuable for early detection and helpful to assess the efficiency of treatments. We observed neither any notable consistency of dynamic CpGs along with the MBL to CLL nor CLL progression and treatment, indicating that the few observed dynamics over the disease progression are possibly an accumulative secondary effect. During disease progression, considerable increases in WBC counts are only juxtaposed with subtle methylome changes. These largely constant methylation levels within each patient indicate that increased clonal expansion occurred without substantial additional departure from the preexisting, already aberrant landscape. The stability of the altered state is further supported by our single-cell

transcriptomes from the five patients that transition from MBL to CLL without any major expression dynamics. Finally, our finding that the cancer methylome also remains mostly unaffected by conventional chemoimmunotherapy or the BCL2 inhibitor venetoclax may keep patients at an elevated risk for relapse, in line with the fact that CLL is rarely cured, although this treatment landscape is continuously evolving (17).

Genetic and epigenetic diversity of normal tissues, tumors, and even clonally amplified cell populations have been most broadly assessed in-depth so far using bulk sequencing. To date, the degree of heterogeneity in methylation levels among distinct genomic regions within a single cell has not been investigated to our knowledge. Here, we applied single-cell methylome analysis and could show that aberrant methylation affects single cells to a surprisingly similar extent. Despite these convincing findings, it needs to be stated that due to technical limitations, such as stochastically missing values caused by unequal coverage, parts of the genome have not been investigated. Nevertheless, we could confirm that hypomethylation is more pronounced across individual MBL and CLL cells, suggesting that the methylation machinery targets a consistent and specific set of regions, though single CpGs are generally affected discordantly (20). As a result, single-cell analysis can help classify tumors and their micro-environment, which provides a more holistic disease picture and may guide more precise treatments in the future.

In sum, our comprehensive exploration over the disease course of CLL, including its precursor stage MBL, highlights several important lessons toward a better mechanistic understanding of the cancer methylome. First, the transition to the altered methylome occurs very early, possibly as a nongenetic precursor lesion that is not yet tumorigenic. Second, it should play at least some facilitating, if not central, role as it was present in all 53 patients evaluated and at all stages with remarkable stability. And finally, its persistence after treatment, though currently limited to the 13 patients that we investigated, suggests that the current chemoimmunotherapy and BCL2 inhibition approach eradicates only some but not all diseased cells. Although it remains to be investigated, the general nature of the epigenetic transformation extends to other cancer types, including many solid tumors, suggesting that this landscape may reflect convergence toward a commonly utilized regulatory mechanism.

METHODS

Human Samples

Heparinized blood samples were obtained from normal donors and patients enrolled in clinical research protocols approved by the Human Subjects Protection Committee of the Dana-Farber Cancer Institute (DFCI), at University of California, San Diego and the Mayo Clinic (CLL Research Consortium), and through the International Cancer Genome Consortium (42) after obtaining written-informed consent. Treatment indication for all 21 patients in the discovery cohort was determined based on International Workshop on Chronic Lymphocytic Leukemia criteria (12, 44). Peripheral blood mononuclear cells (PBMC) from normal donors and patients were isolated by Ficoll/Hypaque density gradient centrifugation. Mononuclear cells were used fresh or cryopreserved with 10% DMSO FBS and stored in vapor-phase liquid nitrogen until the time of analysis. CD19⁺ B cells from normal volunteers and CLL samples with WBC $\leq 50 \times 10^9/L$ were isolated by immunomagnetic selection (Miltenyi Biotec

and stained with anti-CD19-phycoerythrin (PE; BioLegend) prior to FACS sorting for live single cells in the presence of DAPI. MBL cells and naïve and memory B cells from age-matched healthy donors were isolated as follows: Cryopreserved PBMCs were thawed and stained with anti-CD19-PE, CD5-FITC, and CD27-Allophycocyanin (APC; BioLegend). Cells were gated for naïve B cells (CD19⁺, CD27⁻, and CD5⁻), memory B cells (CD19⁺, CD27⁺, and CD5⁻), or MBL (CD19⁺ and CD5⁺; Supplementary Fig. S7A).

Bulk RRBS Library Generation and Data Processing

RRBS libraries were generated from 25 to 100 ng of input DNA using the Ovation Methyl-Seq System (NuGen) following the manufacturer's recommendation. We used NuGen unique molecular identifier (UMI) technology to measure the rate of PCR duplicates on one patient (four samples) and found the duplicate rate to be below 2%, even at an input of only 25 ng of DNA. On average, 15.7M fragments, resulting in 31.4M paired-end 101-base pair (bp) reads, were sequenced per sample on an Illumina HiSeq2500. These reads were aligned to the human hg19 genome using BSMAP (45) with flags `-v 0.05 -s 16 -w 100 -S 1 -p 8 -u`. An average of 21.1M reads per sample was aligned correctly. Custom scripts written in Perl were used to count the number of times a CpG was observed to be methylated. The methylation percentage for each CpG was calculated as the number of times the CpG appeared methylated divided by the total times the CpG was covered in sequencing reads. Finally, we converted the resulting CpG level files to bigWig files, filtering out all CpGs covered with less than five reads. An average of 3.4M CpGs was covered per sample at an average depth of 14x.

Multiplexed Single-Cell RRBS Library Generation and Data Processing

Single-cell RRBS libraries were prepared by combining the first steps (cell lysis and physical separation of DNA and mRNA) of the single-cell methylome and transcriptome sequencing protocol (46) with multiplexed single-cell RRBS (26) using double MspI+HaeIII digestion. Single cells were sorted into 5 μ L RLT plus buffer (QIAGEN) containing 1 U SUPERase in RNase inhibitor (Invitrogen) in 96-well PCR plates, flash-frozen on dry ice, and stored at -80°C . Upon thawing, 5 μ L of QIAGEN RLT plus buffer and 10 μ L M-280 streptavidin beads conjugated to a biotinylated oligo-dT primer were added to each well. After 30 minutes at 25°C , the plates were transferred to a magnet to capture bead-bound mRNA, and the DNA-containing supernatant was transferred to a new 96-well plate. Beads in the original wells were washed twice with 15 μ L of washing buffer (50 mmol/L Tris-HCl, pH 8, 75 mmol/L KCl, 3 mmol/L MgCl₂, 10 mmol/L DTT, and 0.5% Tween-20), and each wash was added to the DNA plate. To clean up the DNA, 1 volume of a 1:5 dilution of AMPure XT SPRI beads (Beckman Coulter) in 20% PEG/2.5 mol/L NaCl and 0.5 μ L Proteinase K (0.8 U/ μ L, NEB) were added. After 30 minutes at 25°C with mixing, the beads were washed with 80% ethanol and genomic DNA eluted with 8 μ L H₂O, with the beads remaining in the well during library prep. After addition of 2 μ L 1x CutSmart buffer (NEB) containing 10 U of MspI (NEB), or 5 U of MspI plus 5 U of HaeIII (NEB), DNA was digested for 2 hours at 37°C , followed by heat inactivation for 15 minutes at 65°C . MspI sites were filled in and fragment ends adenylated by adding 2 μ L 1x CutSmart containing 2.5 U Klenow fragment (3'-5'exo-, NEB), 0.4 μ L of dNTP mixture (10 mmol/L dATP, 1 mmol/L dCTP, and 1 mmol/L dGTP) followed by a two-step incubation for 25 minutes at 30°C and 30 minutes at 37°C and heat inactivation at 70°C for 10 minutes. After addition of 3 μ L 1x CutSmart containing 800 U T4 DNA ligase (NEB), 0.1 μ L of 100 mmol/L ATP (Roche), 1.5 μ L of 0.1 μ mol/L custom 5mC-substituted and indexed (inline barcode) adapter, overnight ligation at 16°C , and heat inactivation (20 minutes at 65°C), 24 separately indexed ligation reactions were pooled. After addition of 3 μ L sheared and dephosphorylated *Escherichia coli*

carrier DNA (27), DNA was cleaned up with 1.8 volumes of AMPure XP beads (Beckman Coulter), eluted off the beads, and bisulfite converted (EpiTect Fast Bisulfite kit, QIAGEN) following the manufacturer's recommendations with extended conversion time (20 minutes each cycle). Each pool of RRBS libraries from 24 single cells was PCR amplified using KAPA HiFi Uracil+ DNA Polymerase, a universal P5, and a pool-specific indexed P7 primer for a total of 17 cycles. The thermoprofile was 98°C denaturation for 45 seconds, 6 cycles of 98°C for 20 seconds, 58°C annealing for 30 seconds, and 72°C extension for 1 minute, followed by 11 cycles of 98°C for 20 seconds, 65°C annealing for 30 seconds, and 72°C extension for 1 minute, and a final extension at 72°C for 5 minutes. To minimize size bias during sequencing, multiple PCR products, each representing 24 single cells, were pooled together and size selected on a 2% NuSieve agarose gel into two fractions (150–400 bp and 400–800 bp) that were sequenced in separate lanes with a 10% spike-in of a library with a balanced base composition, which is typically 2 lanes (1.5 plus 0.5 lanes for the low and high size cut, respectively) for 96 cells. On average, 4.5M fragments, resulting in 9M paired-end 75-bp reads, were generated per sample on an Illumina HiSeq4000.

Sequencing reads were demultiplexed using the inline barcode, adapters were trimmed, and reads were trimmed for quality. These reads were aligned to the human hg19 genome using BSMAP with flags `-v 0.1 -s 12 -w 100 -S 1 -q 20 -u -R`. An average of 8.4M reads (4.2M pairs) per sample was aligned. To determine the methylation state of all CpGs captured and assess the bisulfite conversion rate, we used the mcall module in the MOABS software suite with standard parameter settings (47). Finally, we converted the resulting CpG level files to bigwig files, filtering out all CpGs covered with more than 250 reads resulting in an average of 1.1M CpGs covered per sample.

10x Single-Cell RNA Library Generation and Data Processing

PBMCs were thawed in Roswell Park Memorial Institute 1640 medium supplemented with 10% FBS and centrifuged at 1,500 rpm for 5 minutes. Each sample was filtered through a 70- μ m filter. Cells were resuspended in PBS–0.04% BSA and stained with anti-human CD5 (FITC), CD19 (PE), CD27 (APC), and 7-aminoactinomycin D for 15 minutes on ice (BioLegend). The samples were washed and resuspended in PBS–0.04% BSA at a concentration of 10×10^6 cells/mL. Samples from the same patient were processed and sorted in parallel on the same day using two FacsAria II cytometers (Becton Dickinson). Cells were sorted through a 70- μ m nozzle into 1.5-mL Eppendorf tubes with 10- μ L PBS–0.04% BSA and immediately stored on ice. Cellular suspensions were loaded on a 10x Genomics Chromium Controller platform (10x Genomics, Inc.) to generate a single-cell Gel bead in Emulsion (GEM). Single-cell RNA sequencing (scRNA-seq) libraries were prepared as previously described (48).

The Cell Ranger pipeline (10x Genomics, Inc.) was used for each scRNA-seq dataset to demultiplex the raw base call files, generate the fastq files, perform the alignment against the mouse reference genome hg19, filter the alignment, and count barcodes and UMIs. Outputs from multiple sequencing runs were also combined using Cell Ranger functions.

Data Analysis

If not stated otherwise, all statistics and plots are generated using R version 3.5.1 “Feather Spray.” In all boxplots, the centerline is median; boxes, first and third quartiles; whiskers, 1.5 \times interquartile range; and data beyond the whiskers' end are omitted.

Bed files were processed using UCSCtools and bedtools (v2.25.0).

Additional Data

Whole genome bisulfite sequencing (WGBS) data of normal B cells were obtained from the European Genome-phenome Archive (EGA) under accession EGAS00001001196 for comparison in the phylogenetic and average methylation analysis. Methylation data were filtered for minimum coverage of 10 \times read coverage, and coordinates were converted to hg19 using bedtools liftOver (49). Additional WGBS data of normal B cells were obtained from Beekman and colleagues (30) and downloaded from <http://resources.idibaps.org/paper/the-reference-epigenome-and-regulatory-chromatin-landscape-of-chronic-lymphocytic-leukemia>. Methylation data were filtered for a minimum of 10 \times read coverage, and coordinates were converted to hg19 using bedtools liftOver. Data were used in phylogenetic comparisons, average methylation analysis, and all comparative analysis, e.g., detecting differential methylated regions and genomic region visualizations. To ensure accurate comparison among samples, all published WGBS data were reduced to positions covered in any of our patient RRBS data, resulting in a set of approximately 5 million comparable CpGs.

Single-cell RRBS and RNA-seq data of the two CLL samples were obtained from Gaiti and colleagues (40).

Feature Annotations

Chromatin states were defined by the standard 15-state model using the ChromHMM algorithm (33) and were downloaded from the UCSC Genome Browser (50). The average methylation rate for each sample and per chromatin state was calculated as the mean of all methylation rates overlapping with a particular chromatin state. CpG islands were downloaded from the UCSC Genome Browser; shores were defined as adjacent 2 kb regions and shelves as the next adjacent 2 kb regions. Gene annotations were downloaded from the UCSC Genome Browser (encode v19), and promoter regions were defined as 5,000 nt upstream to 2,500 nt downstream of annotated transcription start site. DMRs were assigned to genes if overlapping with the promoter or gene body with at least one shared base. For unique annotation of DMRs and if a DMR overlapped more than one feature, the ranking was: promoter, gene body, last intergenic; or CpG island, shelf, and last shore. For chromatin state annotation of DMRs, the 15 chromatin states were collapsed into the 6 main categories (Active Promoter, Poised Promoter, Enhancer, Transcription, Insulator, and Heterochromatin), and each DMR was assigned to the region with its maximal overlap.

Mutation and subclone information was taken from ref. 28.

Phylogenetic Tree

The phylogenetic analysis of DNA methylation was performed as previously described (51). In brief, the phylogenetic trees were inferred using the fastme.bal function in the R package ape, which is based on the minimal evolution method. Trees were computed by applying the fastme.bal function on the Euclidean distance matrices of the methylation rates of all samples in the tree. To always capture the maximal information, the subset of CpGs considered was adapted to the sample shows, resulting in $n =$ (i) 28,343,743; (ii) 5,227,401; and (iii) 3,490,971 CpGs for (i) normal B cells, (ii) normal B cells + first time point CLL, and (iii) normal B cells + first time point CLL + MBL; normal B cells + all CLL time points, respectively.

Scatter Plots and Correlation

Scatter plots were created using the smoothScatter function of R, and correlations were calculated using the cor function of R. For the first figure, the average methylation per group was used ($n = 3$ for naïve and memory B cells, $n = 20$ and 21 for unmutated and mutated CLL, respectively). Missing values were removed from the mean calculation. For the matched MBL–CLL correlation, missing values were

not omitted from the naïve, memory, and the five matched MBL and CLL samples, resulting in 1,801,907 CpGs that were used for correlation analysis and scatter plots. Statistics for the full CLL cohort are given in Supplementary Fig. S5A.

Genome Region Visualization

Visualization of methylation levels per CpG at genomic regions was done using the plotTracks function of the R package Gviz (52). Track data were grouped by cell type, and for the curve, representation was plotted as a smoothed line.

Differential Methylation Analysis

DMRs were called using metilene version 0.2–7 (32). DMRs were defined to have an absolute minimum difference in methylation of 0.25 with a maximum distance of 100 nt between CpGs within a DMR and a minimum of 3 CpGs per DMR (parameter $-M\ 100 -m\ 3 -d\ 0.25$). DMRs were calculated between (i) the normal B cells from Beekmann and colleagues (30), (ii) CLL samples as well as between the first and last pretreatment and the last pre- and first posttreatment time points, and (iii) for each sample individually between the first and last pretreatment and the last pre- and first posttreatment time points. More specifically, DMRs were calculated for normal B cells to CLL between the naïve B cells ($n = 3$) and the first time point of the CLL samples with unmutated *IGHV* as well as between the memory B cells ($n = 3$) and the first time point of the CLL samples with mutated *IGHV*. Only positions on the autosomes (chr1–22) were taken into account that were covered by all three normal B-cell samples (naïve or memory) and 90% of the CLL samples ($n = 18$ unmutated/ $n = 19$ mutated CLL), respectively. For within-patient sample versus sample DMRs, all positions that were covered by both samples were taken into account. After DMR calling, all *P* values were adjusted for multiple testing using the R function *p.adj*, and regions with an adjusted *P* value < 0.05 were considered DMRs. As previously described, DMRs were separated into DMRs that overlap CpGs that are dynamic during B-cell differentiation (difference between all normal B cells > 0.25) and subsequently called B-cell-related DMRs, and those that do not overlap any dynamic position are called CLL DMRs. No DMRs ($FDR < 0.05$) were found when comparing the five matched MBL samples with the respective five CLL samples, when comparing the unmutated and mutated CLL, or when comparing the first pretreatment time points to the last pretreatment time points.

For heatmap visualization, methylation levels of DMRs were calculated as the mean methylation of all CpGs with a DMR for all samples and plotted using the heatmap function of the R package ComplexHeatmap (53). Row annotations were based on overlap with features; see Feature Annotations section above. The enrichment analysis of genes affected by DMRs was done using the online Web tool WebGestalt (54). An overrepresentation enrichment analysis (ORA) was calculated for the CLL DMRs with all DMRs as background for Panther pathways and the patient-specific DMRs by comparing recurrently hit genes among more than two patients with all DMR genes.

Single-Cell Analysis. For comparison of the RRBS single-cell experiments, only positions of the double-digest data (naïve and memory B cells) were considered that were also covered by the single-digest data (CLL).

10x scRNA-seq. The single-cell RNA was analyzed using the python toolkit “Scanpy” with default parameters for clustering and UMAP generation (55). Gene expression profiles were generated using parameters for normalized gene expression representation for dotplot and heatmap representations.

Data Accessibility

Raw methylation sequence data from patients are deposited in the database of Genotypes and Phenotypes (dbGAP) record # phs001431.v1.p to allow controlled access and maintain patient privacy.

scRNA-seq and processed methylation data are available under Gene Expression Omnibus (GEO) accession GSE125499. Go to <https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE125499>.

Authors' Disclosures

A. Biran reports grants from Leukemia & Lymphoma Society and Lymphoma Research Foundation during the conduct of the study. N. Purroy reports other from AstraZeneca outside the submitted work. K. Clement reports a patent for US10801070B2 issued to Harvard College General Hospital Corp., Dana-Farber Cancer Institute Inc, and Broad Institute Inc. E. Braggio reports personal fees from DASA outside the submitted work. T.D. Shanafelt reports grants from Genentech, Pharmacyclics, and AbbVie outside the submitted work, and his institution, Mayo Clinic, has a patent for green tea extract issued and based on the research of T.D. Shanafelt and N.E. Kay. N.E. Kay reports grants and personal fees from AbbVie (data and safety monitoring committee) and Pharmacyclics (advisory board); personal fees from Agios (data and safety monitoring committee), AstraZeneca (data and safety monitoring committee/advisory board), Cytomx Therapy (data and safety monitoring committee/advisory board), Dava Oncology (advisory board), Juno Therapeutics (advisory board), Oncotracker (advisory board), and Rigel (data and safety monitoring committee); and grants from Bristol-Myers Squibb, MEI Pharma, MorphoSys, TG Therapeutics, and Tolero Pharmaceuticals outside the submitted work. J.R. Brown reports personal fees from AbbVie (consulting), Acerta/AstraZeneca (consulting), BeiGene (consulting), Catapult Therapeutics (consulting), Dynamo Therapeutics (consulting), Eli Lilly and Company (consulting), Genentech/Roche (consulting), Juno/Celgene (consulting), Kite (consulting), Loxo (consulting), MEI Pharma (consulting), Nextcea (consulting), Novartis (consulting), Octapharma (consulting), Pfizer (consulting), Pharmacyclics (consulting), Rigel (consulting), Sunesis (consulting), TG Therapeutics (consulting), Janssen (honoraria), and Teva (honoraria); grants and personal fees from Gilead (consulting and research funding) and Verastem (consulting and research funding); personal fees and other from MorphoSys (consulting and data safety monitoring committee service); grants from Loxo/Lilly and Sun Pharmaceuticals; and other from Invectys (data safety monitoring committee service) outside the submitted work. S. Li reports grants from NCI during the conduct of the study. K.J. Livak reports grants from NIH/NCI during the conduct of the study. D.S. Neuberg reports grants from NIH P01 CA206978 during the conduct of the study, as well as other from Pharmacyclics (research support) and Madrigal Pharmaceuticals (stock ownership) outside the submitted work. E. Campo reports grants from Spanish Ministry of Science during the conduct of the study. A. Gnrice reports grants from NIH/NCI during the conduct of the study. C.J. Wu reports other from BioNTech (equity holder) outside the submitted work. No disclosures were reported by the other authors.

Authors' Contributions

H. Kretzmer: Conceptualization, resources, supervision, funding acquisition, investigation, visualization, writing—original draft, project administration, writing—review and editing. **A. Biran:** Conceptualization, resources, investigation, visualization, methodology, writing—original draft, writing—review and editing. **N. Purroy:** Resources, investigation, methodology. **C.K. Lemvigh:** Resources, data curation, investigation. **K. Clement:** Resources, data curation, investigation. **M. Gruber:** Resources, data curation, investigation. **H. Gu:** Resources, data curation, investigation, methodology. **L. Rassenti:** Resources, data curation, methodology. **A.W. Mohammad:** Resources, data curation, methodology. **C. Lesnick:** Resources, data

curation, methodology. **S.L. Slager:** Resources, data curation. **E. Braggio:** Resources, data curation. **T.D. Shanafelt:** Resources, data curation. **N.E. Kay:** Resources, data curation. **S.M. Fernandes:** Resources, data curation, methodology. **J.R. Brown:** Resources, methodology. **L. Wang:** Resources, methodology. **S. Li:** Resources, methodology. **K.J. Livak:** Resources, methodology. **D.S. Neuberg:** Resources, methodology. **S. Klages:** Resources, methodology, project administration. **B. Timmermann:** Supervision, funding acquisition, project administration. **T.J. Kipps:** Supervision, funding acquisition. **E. Campo:** Supervision, funding acquisition, methodology, project administration. **A. Gnirke:** Conceptualization, resources, supervision, funding acquisition, investigation, methodology, writing—original draft, project administration, writing—review and editing. **C.J. Wu:** Conceptualization, resources, supervision, funding acquisition, investigation, methodology, writing—original draft, project administration, writing—review and editing. **A. Meissner:** Conceptualization, resources, supervision, funding acquisition, visualization, methodology, writing—original draft, project administration, writing—review and editing.

Acknowledgments

The authors thank members of the Meissner lab, in particular Jocelyn Charlton, and both Erin M. Parry and Inaki Subero-Martin for critical reading of the article. The authors thank Jerome Ritz and the DFCI Pasquarello Tissue Bank in Hematologic Malignancies for the prospective collection and processing of blood samples from healthy donors. This work was supported in part by the NCI (SP01CA081534-14). A. Biran was supported by the Lymphoma Research Foundation. C.K. Lemvigh acknowledges support from the Fishman Family Fund. M. Gruber received a Marie-Curie International Outgoing Fellowship from the European Union (PIOF-2013-624924). T.D. Shanafelt and N.E. Kay were supported by the NIH (R01CA197120). E. Campo is supported by Instituto de Salud Carlos III, Madrid Spain (PMP15/00007), “La Caixa” Foundation (grant CLLEvolution-HR17-00221), Health Research 2017, and European Research Council (ERC) BCLLatlas - 810287 and is a Researcher of the “Institució Catalana de Recerca i Estudis Avançats” (ICREA) of the Generalitat de Catalunya. S.L. Slager and E. Braggio were supported by R01 CA235026. J.R. Brown is supported by NCI RO1 CA 213442 and the Rosenbach Fund for Lymphoma Research. S. Li is supported by the NCI Research Specialist Award (R50CA251956-01). The single-cell analysis was supported by a SPARC grant of the Broad Institute (to A. Gnirke). C.J. Wu acknowledges support from the CLL Global Research Foundation, NHLBI (1R01HL103532-01), and NCI (1R01CA155010-01A1 and UG1 CA233338), and is a Scholar of the Leukemia & Lymphoma Society. A. Meissner acknowledges support from the Starr Foundation, the New York Stem Cell Foundation, and the Max Planck Society.

Received November 11, 2019; revised October 28, 2020; accepted November 18, 2020; published first December 3, 2020.

REFERENCES

- Bird A. DNA methylation patterns and epigenetic memory. *Genes Dev* 2002;16:6–21.
- Smith ZD, Meissner A. DNA methylation: roles in mammalian development. *Nat Rev Genet* 2013;14:204–20.
- Feinberg AP, Ohlsson R, Henikoff S. The epigenetic progenitor origin of human cancer. *Nat Rev Genet* 2006;7:21–33.
- Easwaran H, Tsai HC, Baylin SB. Cancer epigenetics: tumor heterogeneity, plasticity of stem-like states, and drug resistance. *Mol Cell* 2014;54:716–27.
- Smith ZD, Shi J, Gu H, Donaghey J, Clement K, Cacchiarelli D, et al. Epigenetic restriction of extraembryonic lineages mirrors the somatic transition to cancer. *Nature* 2017;549:543–7.
- 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;50:591–602.
- Ziller MJ, Gu H, Muller F, Donaghey J, Tsai LT, Kohlbacher O, et al. Charting a dynamic DNA methylation landscape of the human genome. *Nature* 2013;500:477–81.
- Rajewsky K. Clonal selection and learning in the antibody system. *Nature* 1996;381:751–8.
- Natkunam Y. The biology of the germinal center. *Hematology Am Soc Hematol Educ Program* 2007;210–5.
- Kulis M, Merkel A, Heath S, Queiros AC, Schuyler RP, Castellano G, et al. Whole-genome fingerprint of the DNA methylome during human B cell differentiation. *Nat Genet* 2015;47:746–56.
- Oakes CC, Seifert M, Assenov Y, Gu L, Przekopowicz 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;48:253–64.
- Hallek M, Cheson BD, Catovsky D, Caligaris-Cappio F, Dighiero G, Dohner H, et al. Guidelines for the diagnosis and treatment of chronic lymphocytic leukemia: a report from the International Workshop on Chronic Lymphocytic Leukemia updating the National Cancer Institute-Working Group 1996 guidelines. *Blood* 2008;111:5446–56.
- Mowery YM, Lanasa MC. Clinical aspects of monoclonal B-cell lymphocytosis. *Cancer Control* 2012;19:8–17.
- Landgren O, Albitar M, Ma W, Abbasi F, Hayes RB, Ghia P, et al. B-cell clones as early markers for chronic lymphocytic leukemia. *N Engl J Med* 2009;360:659–67.
- Rawstron AC, Bennett FL, O'Connor SJ, Kwok M, Fenton JA, Plummer M, et al. Monoclonal B-cell lymphocytosis and chronic lymphocytic leukemia. *N Engl J Med* 2008;359:575–83.
- Zenz T, Mertens D, Kuppers R, Dohner H, Stilgenbauer S. From pathogenesis to treatment of chronic lymphocytic leukaemia. *Nat Rev Cancer* 2010;10:37–50.
- Kipps TJ, Stevenson FK, Wu CJ, Croce CM, Packham G, Wierda WG, et al. Chronic lymphocytic leukaemia. *Nat Rev Dis Primers* 2017;3:16096.
- Hamblyn 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;94:1848–54.
- Kulis M, Heath S, Bibikova M, Queiros AC, Navarro A, Clot G, et al. Epigenomic analysis detects widespread gene-body DNA hypomethylation in chronic lymphocytic leukemia. *Nat Genet* 2012;44:1236–42.
- 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;26:813–25.
- Kretzmer H, Bernhart SH, Wang W, Haake A, Weniger MA, Bergmann AK, et al. DNA methylome analysis in Burkitt and follicular lymphomas identifies differentially methylated regions linked to somatic mutation and transcriptional control. *Nat Genet* 2015;47:1316–25.
- Gama-Sosa MA, Slagel VA, Trewyn RW, Oxenhandler R, Kuo KC, Gehrke CW, et al. The 5-methylcytosine content of DNA from human tumors. *Nucleic Acids Res* 1983;11:6883–94.
- Feinberg AP, Vogelstein B. Hypomethylation distinguishes genes of some human cancers from their normal counterparts. *Nature* 1983;301:89–92.
- Hansen KD, Timp W, Bravo HC, Sabuncian S, Langmead B, McDonald OG, et al. Increased methylation variation in epigenetic domains across cancer types. *Nat Genet* 2011;43:768–75.
- Meissner A, Mikkelsen TS, Gu H, Wernig M, Hanna J, Sivachenko A, et al. Genome-scale DNA methylation maps of pluripotent and differentiated cells. *Nature* 2008;454:766–70.
- Charlton J, Downing TL, Smith ZD, Gu H, Clement K, Pop R, et al. Global delay in nascent strand DNA methylation. *Nat Struct Mol Biol* 2018;25:327–32.
- Gu H, Smith ZD, Bock C, Boyle P, Gnirke A, Meissner A. Preparation of reduced representation bisulfite sequencing libraries for genome-scale DNA methylation profiling. *Nat Protoc* 2011;6:468–81.
- Gruber M, Bozic I, Leshchiner I, Livitz D, Stevenson K, Rassenti L, et al. Growth dynamics in naturally progressing chronic lymphocytic leukaemia. *Nature* 2019;570:474–9.

29. Guièze R, Liu VM, Rosebrock D. Mitochondrial reprogramming underlies resistance to BCL-2 inhibition in lymphoid malignancies. *Cancer Cell* 2019;36:369–84.
30. Beekman R, Chapaprieta V, Russinol N, Vilarrasa-Blasi R, Verdaguero-Dot N, Martens JHA, et al. The reference epigenome and regulatory chromatin landscape of chronic lymphocytic leukemia. *Nat Med* 2018;24:868–80.
31. Bock C, Beeraman I, Lien WH, Smith ZD, Gu H, Boyle P, et al. DNA methylation dynamics during in vivo differentiation of blood and skin stem cells. *Mol Cell* 2012;47:633–47.
32. Juhling F, Kretzmer H, Bernhart SH, Otto C, Stadler PF, Hoffmann S. Metilene: fast and sensitive calling of differentially methylated regions from bisulfite sequencing data. *Genome Res* 2016;26:256–62.
33. Ernst J, Kheradpour P, Mikkelsen TS, Shores N, Ward LD, Epstein CB, et al. Mapping and analysis of chromatin state dynamics in nine human cell types. *Nature* 2011;473:43–9.
34. Landan G, Cohen NM, Mukamel Z, Bar A, Molchadsky A, Brosh R, et al. Epigenetic polymorphism and the stochastic formation of differentially methylated regions in normal and cancerous tissues. *Nat Genet* 2012;44:1207–14.
35. Souers AJ, Levenson JD. ABT-199, a potent and selective BCL-2 inhibitor, achieves antitumor activity while sparing platelets. *Nat Med* 2013;19:202–8.
36. Wolf C, Garding A, Filarsky K, Bahlo J, Robrecht S, Becker N, et al. NFATC1 activation by DNA hypomethylation in chronic lymphocytic leukemia correlates with clinical staging and can be inhibited by ibrutinib. *Int J Cancer* 2018;142:322–33.
37. Landau DA, Carter SL, Stojanov P, McKenna A, Stevenson K, Lawrence MS, et al. Evolution and impact of subclonal mutations in chronic lymphocytic leukemia. *Cell* 2013;152:714–26.
38. 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;526:525–30.
39. Nadeu F, Clot G, Delgado J, Martin-Garcia D, Baumann T, Salaverria I, et al. Clinical impact of the subclonal architecture and mutational complexity in chronic lymphocytic leukemia. *Leukemia* 2018;32:645–53.
40. 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;569:576–80.
41. Cahill N, Bergh AC, Kanduri M, Goransson-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;27:150–8.
42. Puente XS, Bea S, Valdes-Mas R, Villamor N, Gutierrez-Abril J, Martin-Subero JJ, et al. Non-coding recurrent mutations in chronic lymphocytic leukaemia. *Nature* 2015;526:519–24.
43. Bormann F, Rodriguez-Paredes M, Lasitschka F, Edelmann D, Musch T, Benner A, et al. Cell-of-origin DNA methylation signatures are maintained during colorectal carcinogenesis. *Cell Rep* 2018;23:3407–18.
44. Hallek M, Cheson BD, Catovsky D, Caligaris-Cappio F, Dighiero G, Dohner H, et al. iwCLL guidelines for diagnosis, indications for treatment, response assessment, and supportive management of CLL. *Blood* 2018;131:2745–60.
45. Xi Y, Li W. BSMAP: whole genome bisulfite sequence MAPPING program. *BMC Bioinformatics* 2009;10:232.
46. Angermueller C, Clark SJ, Lee HJ, Macaulay IC, Teng MJ, Hu TX, et al. Parallel single-cell sequencing links transcriptional and epigenetic heterogeneity. *Nat Methods* 2016;13:229–32.
47. Sun D, Xi Y, Rodriguez B, Park HJ, Tong P, Meong M, et al. MOABS: model based analysis of bisulfite sequencing data. *Genome Biol* 2014;15:R38.
48. Zheng GX, Terry JM, Belgrader P, Ryvkin P, Bent ZW, Wilson R, et al. Massively parallel digital transcriptional profiling of single cells. *Nat Commun* 2017;8:14049.
49. Quinlan AR, Hall IM. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* 2010;26:841–2.
50. Karolchik D, Baertsch R, Diekhans M, Furey TS, Hinrichs A, Lu YT, et al. The UCSC Genome Browser Database. *Nucleic Acids Res* 2003;31:51–4.
51. Brocks D, Assenov Y, Minner S, Bogatyrova O, Simon R, Koop C, et al. Intratumor DNA methylation heterogeneity reflects clonal evolution in aggressive prostate cancer. *Cell Rep* 2014;8:798–806.
52. Hahne F, Ivanek R. Visualizing genomic data using Gviz and Bioconductor. *Methods Mol Biol* 2016;1418:335–51.
53. Gu Z, Eils R, Schlesner M. Complex heatmaps reveal patterns and correlations in multidimensional genomic data. *Bioinformatics* 2016;32:2847–9.
54. Wang J, Vasaiakar S, Shi Z, Greer M, Zhang B. WebGestalt 2017: a more comprehensive, powerful, flexible and interactive gene set enrichment analysis toolkit. *Nucleic Acids Res* 2017;45:W130–W7.
55. Wolf FA, Angerer P, Theis FJ. SCANPY: large-scale single-cell gene expression data analysis. *Genome Biol* 2018;19:15.

BLOOD CANCER DISCOVERY

Preneoplastic Alterations Define CLL DNA Methylome and Persist through Disease Progression and Therapy

Helene Kretzmer, Anat Biran, Noelia Purroy, et al.

Blood Cancer Discov Published OnlineFirst December 3, 2020.

Updated version	Access the most recent version of this article at: doi: 10.1158/2643-3230.BCD-19-0058
Supplementary Material	Access the most recent supplemental material at: http://bloodcancerdiscov.aacrjournals.org/content/suppl/2020/11/26/2643-3230.BCD-19-0058.DC1

E-mail alerts [Sign up to receive free email-alerts](#) related to this article or journal.

Reprints and Subscriptions To order reprints of this article or to subscribe to the journal, contact the AACR Publications Department at pubs@aacr.org.

Permissions To request permission to re-use all or part of this article, use this link <http://bloodcancerdiscov.aacrjournals.org/content/early/2020/12/29/2643-3230.BCD-19-0058>. Click on "Request Permissions" which will take you to the Copyright Clearance Center's (CCC) Rightslink site.