



**HAL**  
open science

## Epigenetic features support the diagnosis of B-cell prolymphocytic leukemia and identify two clinico-biological subtypes

Stella Charalampopoulou, Elise Chapiro, Ferran Nadeu, Thorsten Zenz, Silvia Beà, Marta Aymerich, Ares Martinez-Farran, Maria Rozman, Damien Roos-Weil, Olivier Bernard, et al.

### ► To cite this version:

Stella Charalampopoulou, Elise Chapiro, Ferran Nadeu, Thorsten Zenz, Silvia Beà, et al.. Epigenetic features support the diagnosis of B-cell prolymphocytic leukemia and identify two clinico-biological subtypes. *Blood Advances*, 2024, 10.1182/bloodadvances.2024013327 . hal-04796777

HAL Id: hal-04796777

<https://hal.science/hal-04796777v1>

Submitted on 22 Nov 2024

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution - NonCommercial - NoDerivatives 4.0  
International License

# **Epigenetic features support the diagnosis of B-cell prolymphocytic leukemia and identify two clinico-biological subtypes**

Stella Charalampopoulou<sup>1</sup>, Elise Chapiro<sup>2,3</sup>, Ferran Nadeu<sup>1,4</sup>, Thorsten Zenz<sup>5</sup>, Sílvia Beà<sup>1,4,6,7</sup>, Ares Martínez-Farran<sup>1</sup>, Marta Aymerich<sup>6</sup>, Maria Rozman<sup>6</sup>, Damien Roos-Weil<sup>2,3</sup>, Olivier Bernard<sup>8</sup>, Santos A. Susin<sup>3</sup>, Helen Parker<sup>9</sup>, Renata Walewska<sup>10</sup>, Christopher C. Oakes<sup>11</sup>, Jonathan C. Strefford<sup>9</sup>, Elías Campo<sup>1,4,6,7</sup>, Estela Matutes<sup>6</sup>, Martí Duran-Ferrer<sup>1,\*</sup>, Florence Nguyen-Khac<sup>2,3,\*</sup>, José I. Martín-Subero<sup>1,7,12,\*</sup>.

- 1 Institut d'Investigacions Biomèdiques August Pi i Sunyer (IDIBAPS), 08036 Barcelona, Spain.
- 2 Service d'Hématologie Biologique, Hôpital Pitié-Salpêtrière, Sorbonne Université, AP-HP, Paris, France.
- 3 Centre de Recherche des Cordeliers, Sorbonne Université, Université Paris Cité, Inserm UMRS 1138, Drug Resistance in Hematological Malignancies Team, F-75006 Paris, France.
- 4 Centro de Investigación Biomédica en Red de Cáncer (CIBERONC), Madrid, Spain.
- 5 Department of Medical Oncology and Hematology, University Hospital and University of Zürich, Zurich, Switzerland.
- 6 Hematopathology Section, Pathology Department, Hospital Clínic de Barcelona, Barcelona, Spain.
- 7 Departament de Fonaments Clínics, Facultat de Medicina, Universitat de Barcelona, Barcelona, Spain.
- 8 Gustave Roussy, INSERN U1170, Rue Edouard Vaillant, Villejuif, France.
- 9 Cancer Genomics, School of Cancer Sciences, Faculty of Medicine, University of Southampton, Southampton, UK.
- 10 Department of Molecular Pathology, University Hospitals Dorset, Bournemouth, UK
- 11 The Ohio State University, Columbus, Ohio, United States.
- 12 Institució Catalana de Recerca i Estudis Avançats (ICREA), Barcelona, Spain.

\* These authors contributed equally to this work as senior authors.

**Corresponding authors:** Florence [Nguyen-Khac \(florence.nguyen-khac@aphp.fr\)](mailto:florence.nguyen-khac@aphp.fr) and José I. Martín-Subero ([imartins@recerca.clinic.cat](mailto:imartins@recerca.clinic.cat)).

**Data Sharing:** For data sharing please contact [imartins@recerca.clinic.cat](mailto:imartins@recerca.clinic.cat). DNA methylation data will be deposited in GEO.

## **Abstract**

The recognition of B-cell prolymphocytic leukemia (B-PLL) as a separate entity is controversial based on the current classification systems. Here, we analyze the DNA methylome of a cohort of 20 B-PLL cases diagnosed according to the ICC/WHO-HAEM4R guidelines, and compare them with chronic lymphocytic leukemia (CLL), mantle cell lymphoma (MCL), splenic marginal zone lymphoma (SMZL), and normal B cell subpopulations. Unsupervised principal component analyses suggest that B-PLL is epigenetically distinct from CLL, MCL and SMZL, which is further supported by robust differential methylation signatures in B-PLL. We also observe that B-PLL can be segregated into two epitypes with differential clinico-biological characteristics. B-PLL epitype 1 carries lower IGHV somatic hypermutation and a less profound germinal center-related DNA methylation imprint than epitype 2. Furthermore, epitype 1 is significantly enriched in mutations affecting *MYC* and *SF3B1*, and displays DNA hypomethylation and gene upregulation signatures enriched in *MYC* targets. Despite the low sample size, patients from epitype 1 have an inferior overall survival than those of epitype 2. This study provides relevant insights into the biology and differential diagnosis of B-PLL, and potentially identifies two subgroups with distinct biological and clinical features.

## **Key points**

- B-PLL has a differential DNA methylation signature compared to CLL, MCL, and SMZL.
- Identification of two potential B-PLL subtypes with different biological and clinical features.

## Introduction

B-cell prolymphocytic leukemia (B-PLL) is a very rare and clinically aggressive mature B-cell disorder characterized by a percentage of prolymphocytes exceeding 55% of the lymphoid cells in peripheral blood and presenting with lymphocytosis, cytopenia(s) and splenomegaly<sup>1</sup>. Although it was first described in the 70s<sup>2,3</sup>, it was not recognized as a separate entity by the World Health Organization Classification (WHO) of Tumors until 2001<sup>4</sup>. From the genetic perspective, B-PLL frequently shows *MYC* rearrangements and gains, *TP53* deletions and mutations as well as complex karyotypes<sup>1,5</sup>. However, whether the clinico-biological features of B-PLL are sufficiently different from other B-cell malignancies such as chronic lymphocytic leukemia (CLL), mantle cell lymphoma (MCL) and splenic marginal zone lymphoma (SMZL) to justify its designation as an independent entity is a matter of international debate<sup>6,7</sup>. Indeed, the 5<sup>th</sup> WHO classification, whose online version was recently released, does not anymore recognize B-PLL as a separate entity and propose to reclassify cases within the umbrella of other lymphoid tumors<sup>8</sup>. According to this classification, B-PLLs expressing CD5 are classified either as MCL variants or as a prolymphocytic progression of CLL/SLL when there is more than 15% prolymphocytes in peripheral blood/bone marrow. CD5 negative cases should be classified in the new category of “splenic B-cell lymphoma/leukaemia with prominent nucleoli” which encompasses the previously designated hairy cell leukemia variant (HCL-v) and CD5- B-PLL. On the other hand, B-PLL is still recognized in the International Consensus Classification (ICC)<sup>9</sup> which is based on the previous criteria of WHO-HAEM4R<sup>10</sup> and emphasizes the need to perform a careful differential diagnosis assessment to exclude other lymphoid neoplasms, especially CLL, MCL and SMZL.

Complementary to genetic changes, DNA methylation features have greatly contributed to our understanding of the cellular origin, pathogenetic mechanism and clinical behavior of B-cell neoplasms<sup>11</sup>. For instance, three epigenetic subtypes (i.e. epitopes) with different clinico-biological features have been identified in CLL, which show imprints of germinal center (GC)-inexperienced (low-programmed or naïve-like CLL, n-CLL) and various degrees of GC-experienced mature B cells (moderate GC imprint in intermediate CLL, i-CLL, and strong GC imprint in high-programmed/memory-like CLL, m-CLL)<sup>11–14</sup>. Likewise, two epitopes have been identified in MCL based on their resemblance to GC-inexperienced (C1-MCL) and GC-experienced B cells (C2-MCL), which show a major overlap with conventional and leukemic non-nodal MCLs<sup>15</sup>, respectively. Moreover, two epigenetic subtypes have been identified in Waldenström’s macroglobulinemia based on similarities to memory B cells and plasma cells<sup>16</sup>. Finally, a study on promoter DNA

methylation in SMZL reported the presence of two epitypes with distinct clinico-biological characteristics<sup>17</sup>. In addition to epigenetics-based disease subtyping, the DNA methylome of B cell tumors also reflects their proliferative history and shows *de novo* changes affecting key pathways<sup>18</sup>.

Here, we have analyzed the DNA methylome of a B-PLL cohort diagnosed according to the ICC/WHO-HAEM4R guidelines, and compared it to CLL, MCL, SMZL, and normal B-cell subpopulations, aiming to shed light on its biological nature as well as on the differential diagnosis with other mature B-cell tumors.

## **Methods**

### Selection of B-PLL cases

We gathered a total of 20 B-PLL cases which were diagnosed according to the criteria established in the WHO-HAEM4R classification. Genetic and immunophenotypic data is available from 17 cases previously published<sup>5,19</sup>. Detailed immunophenotypic and clinical features of all the B-PLL cases are shown in Supplementary Table 1. None of the cases had a previous history of any other B-cell neoplasm nor showed any molecular hallmark of HCL-v and SMZL such as *MAP2K1* and *KLF2* mutations, respectively, or MCL such as *CCND1*, *CCND2* and *CCND3* chromosomal rearrangements. Careful cytological examination showed lymphoid cells with prominent nucleoli and lack of cytoplasmic projections in all cases<sup>1</sup>. To further support the differential diagnosis between B-PLL and MCL or CLL, we assessed in B-PLL the expression of previously reported genes specific for MCL and CLL<sup>20</sup> using RNA-seq data from 11 B-PLL cases<sup>5</sup>, 5 MCL<sup>21</sup> (2 conventional and 3 leukemic non-nodal), 294 CLL<sup>22</sup> cases (Supplementary Fig. 1A). In addition, we compared the mutational landscape of B-PLL with that of previously described MCL<sup>23</sup> and CLL<sup>24</sup> cohorts (Supplementary Fig. 1B). A summary of the biological and clinical characteristics of B-PLL as compared to CLL, MCL and SMZL is provided in Supplementary Table 2.

### DNA methylation profiling

We used Illumina EPIC V1 arrays to generate DNA methylation profiles of the 20 B-PLLs as well as from splenic tumor material of nine SMZLs. These SMZLs were diagnosed according to current guidelines<sup>25</sup> of the University Hospital Dorset (IGHV sequencing analysis [IGHV1-02\*04, n=6/9], deletion 7q status [n=4/9 positive], and *KLF2* [n=3/9 mutated] and *NOTCH2* [n=3/9 mutated]). Moreover, we used previous EPIC array data from 78 CLL<sup>26</sup> (27 n-CLL, 10 i-CLLs, 39 m-CLLs and 2 unclassified), 70 MCL<sup>23</sup> (49 C1-

MCLs (39 with classical morphology, 10 with pleomorphic/blastoid), 19 C2 MCLs, and 2 unclassified) (Supplementary Table 3). Although the Illumina arrays are robust and show little slide number batch effects, we compared the methylation profiles of cases from the same disease epitype located in different slides. We found a minimal number of differentially methylated CpGs indicating that the batch effect in our series is negligible as compared to the biological differences (data not shown). We additionally used 450k data of 476 CLL cases and 5 CLL cases containing from 16 to 44% of prolymphocytic cells (CLL/PL)<sup>22</sup>(Supplementary Table 4), as well as EPIC V1 data of two sequential samples of a CLL case that underwent a clonally-related prolymphocytic progression (CLL-pPL)<sup>27</sup>. Finally, we also included EPIC V1 data of normal B-cell subpopulations including 2 Naive B-cells (NBC), 1 Germinal Center (GC), 3 Memory B-cells (MBC) and 1 Plasma cell (PC)<sup>27</sup> (Fig. 1A) as well as complementary 450k data of 6 Hematopoietic Precursor Cells, 12 pre-B cells, 4 intermediate B cells, 15 NBC, 9 GC, 8 tonsil PC, 10 MBC and 3 bone marrow PC<sup>28</sup> to extend our normal B cell dataset for phylogenetic analyses.

#### DNA methylation data analyses

DNA methylation data was analyzed as previously described<sup>18</sup>. Briefly, we removed CpGs with a pvalue  $\leq 0.01$  in more than 10% of the samples as well as SNP associated CpGs, sexual chromosome CpGs and individual-specific CpGs. After all the filtering steps, the data was normalized using the SWAN algorithm. CpGs were annotated using the annotation package *IlluminaHumanMethylationEPICanno.ilm10b4.hg19*. CpGs whose methylation are modulated during B cell differentiation were selected from a previous publication using 450k Illumina data<sup>28</sup> or calculated for EPIC array-specific CpGs using an absolute DNA methylation difference of at least 0.25 between NBC and MBC. Previously generated chromatin states of the GM12878 B-cell cell line<sup>29</sup> were used to annotate CpGs to regulatory regions. CLL and MCL cases were further subdivided into epitypes using a previous DNA methylation classifier of B cell tumours<sup>18</sup> available at <https://durant-ferrerm.github.io/Pan-B-cell-methylome/B.cell.tumor.classifier.html>. Unsupervised analyses of all the DNA methylation data were performed using Principal Component Analyses (PCA). Differential methylation (DM) analyses were performed using *limma* with FDR < 0.05 and methylation difference of at least 0.25 as thresholds. Gometh was used to identify the functions of genes showing differential methylation.

We used a permutation analysis to test the robustness of the differential methylation signatures between B-PLL and CLL or MCL. The permutation analysis was carried out by performing 1,000 DM analyses with distinct sets of randomly selected CLL or MCL cases

in each permutation. The number of randomly selected CLL or MCL cases in each permutation was the same as the number of B-PLL cases. In that way, we obtained the distribution of the expected number of DM CpGs assuming that our B-PLL cases were in reality misdiagnosed CLL or MCL cases, and compared it with the observed number of DM CpGs in B-PLL vs CLL or MCL. The p-value was calculated as the number of permutations resulting in lower number of DM CpGs than the observed divided by the total number of permutations.

#### Phylogenetic analyses with DNA methylation

Phylogenetic analyses were performed as in Oakes 2016<sup>13</sup>. Briefly, phylogenetic trees were inferred by the minimal evolution method using the *fastme.bal* function in the R package ape. Phylogenies were generated by applying the minimal evolution algorithm on Euclidean distance matrices based on continuous methylation values of 1,112 CpGs present in the normalized data matrix.

#### Gene expression analyses

We used previous RNA-seq data of 11 B-PLL cases<sup>5</sup> to perform differential expression analyses with *DESeq2* package considering as differentially expressed (DE) the genes with FDR < 0.05. Gene set enrichment analysis was done with the *fgsea* package using the Hallmark gene sets of the Broad Institute. For the gene expression comparisons performed with B-PLL, MCL and CLL data together, we transformed the normalized expression values to percentiles to minimize the possible batch effects from RNA-seq data of distinct studies and data modalities.

#### Transcription factor binding analysis

Transcription factor (TF) binding analysis was performed with PWMEnrich package in windows of +/- 100 bp around the target CpGs as described previously<sup>18</sup>. TFs were considered significant if the binding site was present in at least 10% of the sequences and the p value was below 0.05.

#### Additional statistical analyses

Genetic and clinical associations were tested using two-tailed Fisher test and the log-rank test, respectively. All analyses were performed in R software.

## **Results**

### B-PLL shows a DNA methylation signature different from CLL, MCL and SMZL

We initially performed an unsupervised PCA including cases from all four diseases and analyzed separately each of the first eight principal components to systematically evaluate similarities and differences among them (Fig. 1B, Supplementary Fig. 2). We performed pair-wise entity comparisons of the values of each component, and we detected significant differences between B-PLL vs. CLL, MCL and SMZL in five, four and five components, respectively. We next performed differential methylation analyses and we detected that B-PLL shows 5,058 differentially methylated (DM) CpGs vs. CLL, 2,755 vs. MCL, and 60,841 vs. SMZL (Fig. 1C, Supplementary Tables 5-7). In the case of B-PLL vs. CLL and B-PLL vs. MCL, we further validated the robustness of the signatures by permutation analyses which aimed at calculating the number of DM CpGs that would be obtained by chance if the B-PLL were misdiagnosed CLL or MCL cases (Methods). These analyses indicated that the number of observed DM CpGs between B-PLL and CLL or MCL are very unlikely to be obtained by chance ( $P < 0.001$ ) (Supplementary Fig. 3). **Despite the fact that most DNA methylation changes in lymphoid tumors do not have a direct impact on the expression of underlying genes<sup>11</sup> we performed a Gene Ontology analysis of the DM CpGs between B-PLL and each of the other diseases and normal B cells. Although we did not obtain highly significant enrichments, hypomethylated CpGs in B-PLL as compared to CLL, MCL and SMZL were related to genes involved in functions associated with different pathways and the immune system (see Supplementary Table 8).** We also identified a signature of 31 CpGs that can differentiate B-PLL vs CLL, MCL and SMZL (Fig. 1D, Supplementary Table 9). Collectively, these analyses suggest that B-PLL -diagnosed on the basis of the ICC/WHO-HAEM4R criteria- has a different epigenetic make-up from CLL, MCL, and SMZL, and the identified methylation patterns may set the basis for an epigenetic-based differential diagnosis approach.

### B-PLLs are epigenetically distinct from CLLs with prolymphocytoid cells, a prolymphocytic progression of CLL, and blastoid/pleomorphic MCLs

As previously mentioned, the WHO-HAEM5 classification states that some B-PLL cases should be considered prolymphocytic progressions of CLL or blastoid MCL<sup>8</sup>. Therefore, we next performed DM analysis between B-PLL and 5 CLL cases with percentage of prolymphocytes between 16% and 44% (CLL/PL) and we identified 3,197 DM CpGs (Supplementary Table 10, Fig. 2A). Moreover, we identified 2,906 CpGs comparing 476 CLLs with a lower percentage of prolymphocytes (<10 %) vs B-PLL (Supplementary Table 11). We then performed a PCA with those CpGs including the 5 CLL/PL cases to evaluate



whether they cluster with CLL or B-CLL. We observed that the CLL/PL cases are separated from B-PLL cases and cluster with their respective CLL epitype regardless of the percentage of prolymphocytes (Fig 2B). Finally, within the same PCA we additionally represented two sequential samples from a highly informative CLL case, which represents a *bona fide* clonally-related prolymphocytic progression of a previous CLL with classic morphology (CLL-pPL)<sup>27</sup>. In this case, neither the progressed nor the initial sample cluster with the B-PLLs but show a CLL-related DNA methylome (Fig 2B). Overall, these analyses indicate that *de novo* B-PLLs (i.e., lacking evidence of a previous B-cell tumor) are epigenetically distinct from CLLs with prolymphocytoid cells or even a prolymphocytic progression of a previous CLL. These observations seem to question the assumption that B-PLLs are prolymphocytic progressions of CLL, as proposed by the WHO-HAEM5 classification.

Similarly, the comparison of B-PLL vs 10 pleomorphic/blastoid MCL cases with EPIC data led to a clear signature of 10,922 DM CpGs (Supplementary Table 12, Fig. 2C). Considering that all pleomorphic/blastoid cases belong to the C1 epitype, we then compared 39 classical morphology C1 MCLs with B-PLL and identified 8,809 DM CpGs (Supplementary Table 13). The PCA of those CpGs clearly indicated that the 10 blastoid C1-MCL cases do cluster with C1-MCLs with classical morphology rather than with B-PLLs (Fig. 2D). These data indicate that the B-PLLs included in this study show a different DNA methylome than blastoid MCLs, which in turn are epigenetically similar to MCLs with classical morphology.

#### Identification of two B-PLL epitypes with different levels GC-related epigenetic imprinting

As epigenetic variability in CLL and MCL is in part related to different epitypes with methylation imprints of pre- and post-germinal center B cells<sup>11–15</sup>, we further explored whether this was also the case for B-PLL. We then performed a PCA with B-PLL and normal B cell subsets (Fig. 3A, Supplementary Fig. 4) and showed that PC1 separated B-PLL and normal B cells, whereas PC2 suggested the presence of two epitypes of B-PLL aligned with variability in normal B cell subsets. One cluster (n=6, epitype 1) was epigenetically closer to naive B cells whereas the other (n=14, epitype 2) was closer to germinal center-experienced B cells (Fig. 3A). This finding suggests that B-PLL could be segregated into two subgroups related to B cells at different maturation stages. Next, we compared their DNA methylome and identified 5,465 DM CpGs (Fig. 3B, Supplementary Table 14). Out of those, 5,036 and 429 CpGs were hypomethylated and hypermethylated in epitype 1 as compared to epitype 2, respectively. Additionally, when we examined those

CpGs in the context of normal B-cell differentiation, we noticed 9.5% of CpGs hypomethylated and 81% of CpG hypermethylated in epitype 1 were B cell-related (Supplementary Table 14), suggesting that the hypermethylation signature mostly contains a B-cell differentiation epigenetic imprint. Evaluating those B cell related CpGs in the context of NBC and MBC, we identified that 249/429 (58%) CpGs hypermethylated in epitype 1 and 130/5036 (2.6%) CpGs hypomethylated in epitype 1 show a clear signature linking epitype 1 to NBC (i.e. GC-inexperienced cells) and epitype 2 to MBC (i.e. GC-experienced) (Fig. 3C, Supplementary Table 15). As B-PLL almost always shows IGHV somatic hypermutation (Supplementary Table 1)<sup>5</sup>, it is unlikely that some cases arise from GC-inexperienced cells. Therefore, to evaluate the overall cellular origin of B-PLL, we used CpGs distinguishing GC-experienced vs GC-inexperienced B cells that were previously applied to trace the origin of CLL<sup>13</sup>. A phylogenetic analysis (Methods) with those CpGs revealed that B-PLL as a whole shows an epigenetic imprint associated with the GC-experienced cells (Supplementary Fig. 5), consistent with our initial PCA analysis (PC1). Thus, the most plausible explanation is that although all B-PLLs derive from GC experienced B-cells, the epitypes may be subjected to different levels of GC exposure. Epitype 1 shows a partial GC epigenetic programming, maintaining thereby some similarities with GC-inexperienced cells, whereas epitype 2 shows a more extensive GC programming. This interpretation is concordant with the fact that epitype 1 has a tendency towards lower IGHV somatic hypermutation (mean germline identity 96.45% vs 94.32% respectively, P=0.08) (Fig. 3D) and a significantly lower proliferative history measured by the epiCMIT mitotic clock (P=0.008)<sup>18</sup> (Fig. 3E), similar to n-CLL and C1/cMCL epitypes which show earlier cells of origin<sup>11,18</sup>.

#### (Epi)genetic characterization, CD5 expression and clinical relevance of the B-PLL epitypes

We further characterized the DM CpGs between B-PLL epitypes. Epitype 2 did not show any specific *de novo* methylation signature. However, epitype 1 hypomethylation was mostly occurring *de novo*, as both normal B cells and epitype 2 were methylated at those CpGs. Moreover, 799/5,036 of the hypomethylated CpGs in B-PLL epitype 1 (Fig. 4A, Supplementary Table 16) were also *de novo* compared to CLL and MCL epitypes. These CpGs were mainly located in low CpG-content regions and enriched in regulatory regions and transcription-related chromatin states (Supplementary Fig. 6A-B). A TF analysis using the genomic regions associated to these CpGs (Methods) showed a significant enrichment in binding motifs of several TF families (Supplementary Fig 6C and Supplementary Table 17). Out of all the TFs identified, *BHLHA15/MIST1* was also found to be significantly

overexpressed in epitype 1 B-PLL ( $P=0.016$ , Supplementary Fig 6D), suggesting that it could be involved in the pathogenesis of this epitype. At the genetic level, epitype 1 was enriched in *MYC* ( $P=0.034$ ) and *SF3B1* ( $P=0.027$ ) mutations and 1q gain ( $P=0.024$ ) (Fig. 4B). Further supporting the existence of these 2 B-PLL epitypes, we identified 95 differentially expressed genes (Fig. 4C, Supplementary Table 18). Overexpressed genes in epitype 1 were found to be enriched in *MYC* targets (Fig. 4D), although *MYC* expression was consistently high in all B-PLL cases (Fig. 4E). In addition, since the WHO-HAEM5 classification highlights the importance of CD5 expression to reclassify B-PLL cases, we also explored CD5 expression in the two B-PLL epitypes. CD5 was expressed in all six epitype 1 and 6/14 epitype 2 B-PLLs ( $P=0.04$ ), which could suggest a relationship between CD5 and methylation patterns. Nonetheless, when comparing the presence or absence of CD5 within epitype 2 cases, we only detect 5 DM CpGs. In contrast, 2,014 CpGs were differential comparing CD5+ cases between epitypes (Fig. 4F, Supplementary Table 19). These data strongly suggest that CD5, although enriched in epitype 1, is unrelated to the differential methylation pattern observed between B-PLL epitypes. Furthermore, they suggest that the decision of the WHO classification of segregating B-PLLs into other entities based on the expression of CD5 is not supported by a differential DNA methylation signature. Finally, despite the small sample size inherent to this rare leukemia, epitype 1 showed a shorter overall survival (OS) than epitype 2 (Fig. 4G), although the proportion of cases with *TP53* abnormalities was similar in both epitypes (Fig. 4B).

## Discussion

This project on the DNA methylome of B-PLL was initiated when the disease was still recognized as a separate entity in the WHO-HAEM4R classification. However, when the new WHO-HAEM5 and ICC classifications were published in 2022<sup>8,9</sup>, an international debate started on whether B-PLL truly exists or not. Whereas the ICC maintained B-PLL as a separate entity, the new WHO classification eliminated this leukemia and cases were fragmented to become variants or progressive forms of MCL, CLL, or SMZL<sup>8</sup>. Therefore, we started our analysis by evaluating, in the first place, whether B-PLL, as diagnosed following the ICC/WHO-HAEM4R criteria, was epigenetically distinct from MCL, CLL, and SMZL. Following exhaustive unsupervised and supervised statistical analyses, our results suggest that the 20 B-PLL cases studied herein show an epigenetic configuration that is overall different from of other three entities. The differential features of B-PLL versus each of the three entities is discussed below.

In the case of MCL, initial reports on B-PLL included some cases that in reality were misdiagnosed MCLs, as they had the *CCND1::IGH* translocation<sup>30</sup>. In our series, all the B-PLL lack any of the *CCND1*, 2 or 3 rearrangements, which is a key criterion to differentiate B-PLL from MCL in the ICC/WHO-HAEM4R classifications. Also, they don't show overexpression of any of the CCN family genes excluding the presence of cryptic *CCND* translocations that have been reported previously in MCL<sup>31,32</sup>. Moreover, B-PLLs and MCL show differential mutational patterns and gene expression profiles, and present 2,755 DM CpGs. We could also show that MCL cases with blastoid/pleomorphic morphology present a clear differential methylation signature from B-PLL, further ruling out that, as the WHO-HAEM5 classification states, some B-PLL cases are actually MCLs with blastoid morphology<sup>8</sup>. Therefore, it seems that none of the B-PLLs included in our study are misdiagnosed MCLs.

Regarding CLL, the WHO-HAEM5 classification asserts that CD5+ B-PLL cases represent prolymphocytic progressions of CLL. We included in our analyses CLL cases with >15% and <55% of cells with prolymphocytoid morphology and observed that these cases were epigenetically distinct from B-PLL and similar to CLLs with typical morphology. Therefore, prolymphocytic progressions of CLL maintain a CLL-related methylome and do not resemble B-PLLs. Moreover, we further analyzed a highly informative case in which a classic CLL progressed into a clonally-related CLL with prolymphocytic morphology<sup>27</sup>. Both CLL and CLL-pPL cells showed a DNA methylation pattern consistent with CLL and not with B-PLL. This scenario is similar to the Richter transformation of CLL into DLBCL. Such DLBCLs arising from a previous CLL seem to maintain a CLL epigenetic imprint and can be differentiated from *de novo* DLBCL<sup>33</sup>. Additionally, the B-PLLs studied herein and CLLs show differential mutational landscapes and expression of key CLL genes. Overall, we could not find any molecular evidence that B-PLLs merely represent progressed forms of CLL.

As compared to SMZL, B-PLLs were epigenetically very different. However, we recognize that the SMZL sample size was low, and therefore, a more detailed comparison of B-PLL versus a larger SMZL series is needed to properly characterize the potential overlap between these two entities.

Overall, this initial part of our analyses seems to indicate that B-PLL, diagnosed on the basis of the ICC/WHO-HAEM4R classification, is epigenetically different from MCL, CLL and SMZL. Additionally, we identified a DNA methylation signature that segregates B-PLL from the other three entities, which can potentially be used to develop molecular biomarkers that may contribute to the differential diagnosis of B-PLL.

In addition to identifying an epigenetic signature of B-PLL as a whole, we detected that the disease can potentially be divided into two epigenetic subtypes. This finding is aligned with findings in MCL, CLL and WM, in which different epitypes linked to distinct B-cell maturation stages are described<sup>11-16</sup>. As compared to epitype 2, B-PLL epitype 1 showed a less profound GC-related epigenetic imprint, a tendency towards lower IGHV somatic hypermutation and a lower proliferative history. These findings may suggest that B-PLL epitype 1 derive from post-GC B cells that underwent a moderate exposure to the (epi)mutagenic environment of the GC whereas epitype 2 shows a stronger GC imprint (i.e. various rounds of recirculation between light and dark zone). Additionally, epitype 1 was the only of the two that presented a *de novo* specific hypomethylation signature enriched for *BHLHA15/MIST1*, a TF also overexpressed in this epitype and previously reported to be related to lower OS in cervical cancer<sup>34</sup>. At the genetic level, epitype 1 was enriched in 1q gains as well as *MYC* and *SF3B1* mutations, and showed an upregulation of genes related to *MYC* pathways. *MYC* is a *bona fide* poor prognostic factor in cancer<sup>35</sup>, while *SF3B1* is a known unfavorable marker in CLL<sup>36</sup> associated with the intermediate epitype<sup>12,22</sup> and stereotyped subset #2<sup>22,37</sup>. All these molecular features of epitype 1 may account for the poorer clinical outcome of this epitype as compared to epitype 2.

In conclusion, our study provides fresh insights into the molecular features of B-PLL carefully diagnosed based on the ICC/WHO-HAEM4R criteria. We did not find epigenetic evidence that B-PLLs, or a fraction of them, resemble MCL, CLL or SMZL. Instead, it seems that B-PLL is a *de novo* disease independent from the other three entities. We are aware that the relatively small sample size imposes a careful interpretation of the data, and in particular the presence of two epitypes, needs to be validated in the future. Moreover, a comprehensive analysis of the B-PLL methylome in the context of a broader range of B-cell lymphoproliferative disorders is needed. We expect that the data provided by this study will significantly contribute to the current discussion on B-PLL, with the hope that an international consensus will be reached.

## **Acknowledgements**

This research was funded by the Accelerator award CRUK/AIRC/AECC joint funder-partnership (to J.I.M.S.), Generalitat de Catalunya Suport Grups de Recerca AGAUR (2021-SGR-1343 to J.I.M.S., 2021-SGR-01172 to E.C, and 2021-SGR-01293 to S.B.), la Caixa" Foundation (CLLEvolution - LCF/PR/HR17/52150017 [HR17-00221LCF] and CLLSYSTEMS - LCF/PR/HR22/52420015 [HR22-00172] Health Research 2017 and 2022 Programs, to E.C.), Force Hemato (grant reference: 03-2022), French Innovative

Leukemia Organization (FILO) group, ACLF (grant 2022) and SIRIC-CURAMUS (Cancer United research Associating Medecine, University and Society; grant reference: INCa-DGOS-INSERM\_12560 and INCa-DGOS-INSERM-ITMO Cancer\_18010). M.D-F is supported by a postdoctoral grant from the Spanish Association Against Cancer (AECC). J.C.S was supported by research grants from the Kay Kendall Leukemia Fund (grant number), and Cancer Research UK (ECRIN-M3 accelerator award C42023/A29370). This work was partially developed at the Centro Esther Koplowitz (CEK, Barcelona, Spain).

### **Authorship Contributions**

S.C. analyzed and interpreted the data, and wrote the manuscript; E.Ch. provided samples and data, and assisted in data interpretation, F.N., A.M.-F., M.A., M.R., H.P., R.W., C.C.O., J.C.S., and T.Z. provided essential data and tools, S.B., E.C. and E. M. provided samples and assisted in data interpretation; M.D-F, F.N.K and J.I-M.S. jointly designed the research, supervised the study and wrote the manuscript.

### **Disclosure of Conflicts of Interest**

The authors declare no financial conflicts of interest with regard to the contents of the present manuscript.

## References

1. Algrin C, Pérol L, Chapiro E, et al. Retrospective analysis of a cohort of 41 de novo B-cell prolymphocytic leukemia patients: impact of genetics and targeted therapies (a FILO study). *Haematologica*. 2023;108(6):1691–1696.
2. Catovsky D, Galetto J, Okos A, et al. Prolymphocytic leukaemia of B and T cell type. *Lancet*. 1973;(2):233–234.
3. Galton DAG, Goldman JM, Wiltshaw E, et al. Prolymphocytic Leukaemia. 1974.
4. Catovsky D, Montserrat E, Harris N, Muller-Hermelink H. Tumours of Hematopoietic and Lymphoid Tissues. World Health Organization Classification of Tumours. IARC Press: Lyon, France; 2001.
5. Chapiro E, Pramil E, Diop boyba, et al. Genetic characterization of B-cell prolymphocytic leukemia: a prognostic model involving MYC and TP53. *Blood*. 2019 Nov 21;134(21):1821-1831.
6. Nguyen-Khac F, Bidet A, Troadec MB, et al. The 5th edition of the WHO classification of haematolymphoid tumors: comments from the Groupe Francophone de Cytogénétique Hématologique (GFCH). *Leukemia*. 2023;37(4):946–947.
7. Siebert R, Schuh A, Ott G. Response to the Comments from the Groupe Francophone de Cytogénétique Hématologique (GFCH) on the 5th edition of the World Health Organization classification of haematolymphoid tumors. *Leukemia*. 2023;
8. Alaggio R, Amador C, Anagnostopoulos I, et al. The 5th edition of the World Health Organization Classification of Haematolymphoid Tumours: Lymphoid Neoplasms. *Leukemia*. 2022;36(7):1720–1748.
9. Campo E, Jaffe ES, Cook JR, et al. The International Consensus Classification of Mature Lymphoid Neoplasms: a report from the Clinical Advisory Committee. *Blood*. 2022;140(11):1229–1253.
10. Swerdlow SH, Campo E, Harris NL, et al. WHO Classification of Tumours of Haematopoietic and Lymphoid Tissues. Revised 4th Edition. Lyon, France: International Agency for Research on Cancer; 2017.
11. Duran-Ferrer M, Martín-Subero JI. Epigenomic Characterization of Lymphoid Neoplasms. *Annu Rev Pathol*. 2024;19:371–396.
12. Kulis M, Heath S, Bibikova M, et al. Epigenomic analysis detects widespread genome-wide DNA hypomethylation in chronic lymphocytic leukemia. *Nat Genet*. 2012;44(11):1236–1242.
13. Oakes CC, Seifert M, Assenov Y, et al. DNA methylation dynamics during B cell maturation underlie a continuum of disease phenotypes in chronic lymphocytic leukemia. *Nat Genet*. 2016;48(3):253–264.
14. Kulis M, Martín-Subero JI. Integrative epigenomics in chronic lymphocytic leukaemia: Biological insights and clinical applications. *Br J Haematol*. 2023;200(3):280–290.
15. Queirós AC, Beekman R, Vilarrasa-Blasi R, et al. Decoding the DNA Methylome of Mantle Cell Lymphoma in the Light of the Entire B Cell Lineage. *Cancer Cell*. 2016;30(5):806–821.
16. Roos-Weil D, Giacomelli B, Armand M, et al. Identification of 2 DNA methylation subtypes of waldenström macroglobulinemia with plasma and memory b-cell features. *Blood*. 2020;136(5):585–595.
17. Arribas AJ, Rinaldi A, Mensah AA, et al. DNA methylation profiling identifies two splenic marginal zone lymphoma subgroups with different clinical and genetic features. *Blood*. 2015;125(12):1922–1931.

18. Duran-Ferrer M, Clot G, Nadeu F, et al. The proliferative history shapes the DNA methylome of B-cell tumors and predicts clinical outcome. *Nat Cancer*. 2020;1(11):1066–1081.
19. Magnano L, Navarro A, López-Guerra M, et al. Chronic lymphocytic leukemia and prolymphocytic leukemia. Two coins or two sides of the same coin? *Haematologica*. 2020 Sep 1;105(9):e484.
20. Navarro A, Clot G, Martínez-Trillos A, et al. Improved classification of leukemic B-cell lymphoproliferative disorders using a transcriptional and genetic classifier. *Haematologica*. 2017;102(9):360–363.
21. Vilarrasa-Blasi R, Soler-Vila P, Verdaguer-Dot N, et al. Dynamics of genome architecture and chromatin function during human B cell differentiation and neoplastic transformation. *Nat Commun*. 2021;12:651.
22. Puente XS, Beà S, Valdés-Mas R, et al. Non-coding recurrent mutations in chronic lymphocytic leukaemia. *Nature*. 2015;526(7574):519–524.
23. Nadeu F, Martín-García D, Clot G, et al. Genomic and epigenomic insights into the origin, pathogenesis, and clinical behavior of mantle cell lymphoma subtypes. *Blood*. 2020;136(12):1419–1432.
24. Knisbacher BA, Lin Z, Hahn CK, et al. Molecular map of chronic lymphocytic leukemia and its impact on outcome. *Nat Genet*. 2022;54(11):1664–1674.
25. Walewska R, Eyre TA, Barrington S, et al. Guideline for the diagnosis and management of marginal zone lymphomas: A British Society of Haematology Guideline. *Br J Haematol*. 2024;204(1):86–107.
26. Dietrich S, Oleś M, Lu J, et al. Drug-perturbation-based stratification of blood cancer. *Journal of Clinical Investigation*. 2018;128(1):427–445.
27. Nadeu F, Royo R, Massoni-Badosa R, et al. Detection of early seeding of Richter transformation in chronic lymphocytic leukemia. *Nat Med*. 2022;28(8):1662–1671.
28. Kulis M, Merkel A, Heath S, et al. Whole-genome fingerprint of the DNA methylome during human B cell differentiation. *Nat Genet*. 2015;47(7):746–756.
29. Ernst J, Kheradpour P, Mikkelsen TS, et al. Mapping and analysis of chromatin state dynamics in nine human cell types. *Nature*. 2011;473(7345):43–49.
30. Ruchlemer R, Parry-Jones N, Brito-Babapulle V, et al. B-prolymphocytic leukaemia with t(11;14) revisited: A splenomegalic form of mantle cell lymphoma evolving with leukaemia. *Br J Haematol*. 2004;125(3):330–336.
31. Martín-García D, Navarro A, Valdés-Mas R, et al. CCND2 and CCND3 hijack immunoglobulin light-chain enhancers in cyclin D12 mantle cell lymphoma. *Blood*. 2019;133(9):940–951.
32. Fuster C, et al. Cryptic insertions of the immunoglobulin light chain enhancer region near *CCND1* in t(11;14)-negative mantle cell lymphoma. *Haematologica*. 2020 Aug;105(8):e408-e411.
33. Broséus J, Hergalant S, Vogt J, et al. Molecular characterization of Richter syndrome identifies de novo diffuse large B-cell lymphomas with poor prognosis. *Nat Commun*. 2023 Jan 19;14(1):309.
34. Wang W, Xie X, Zhou Z, Zhang H. Expression analysis of MIST1 and EMT markers in primary tumor samples points to MIST1 as a biomarker of cervical cancer. *Int J Gen Med*. 2021;14:1293–1300.
35. Dang C V. MYC on the path to cancer. *Cell*. 2012;149(1):22–35.
36. Nadeu F, Delgado J, Royo C, et al. Clinical impact of clonal and subclonal TP53, SF3B1, BIRC3, NOTCH1, and ATM mutations in chronic lymphocytic leukemia. *Blood*. 2016;127(17):2122–2130.
37. Strefford JC, Sutton LA, Baliakas P, et al. Distinct patterns of novel gene mutations in poor-prognostic stereotyped subsets of chronic lymphocytic leukemia: The case of SF3B1 and subset #2. *Leukemia*. 2013;27(11):2196–2199.



## Figure Legends

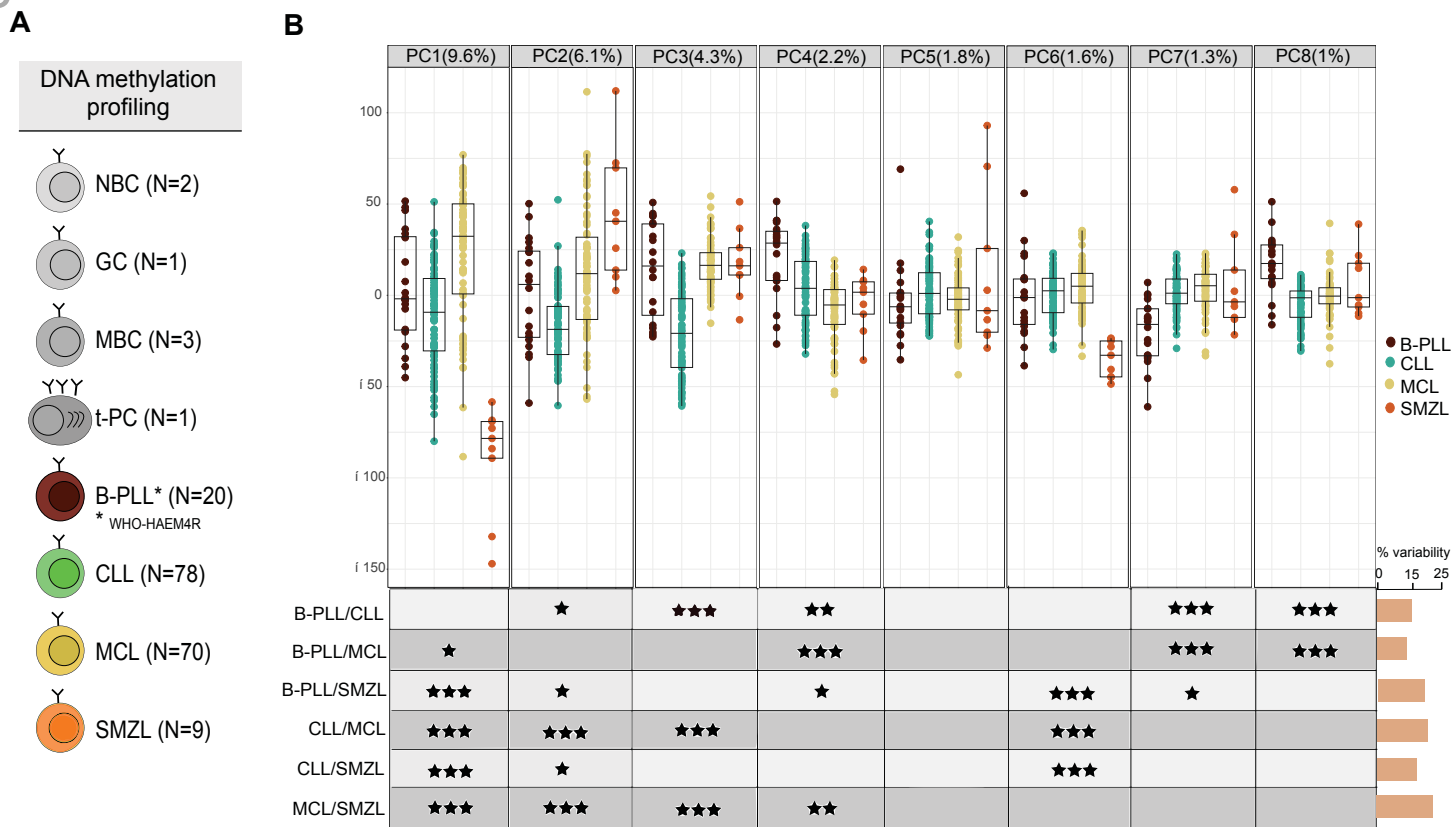
**Figure 1. Differential DNA methylation patterns of B-PLL compared with CLL, MCL SMZL.** **A.** Number of samples with DNA methylation data available. **B.** Boxplots of the first eight principal components (PC) including B-PLL, CLL, MCL and SMZL cases. At bottom, the table shows the FDR corrected p-values of Wilcoxon test of pair-wise comparisons across entities. On the right, the bars represent the total percentage of the variability showing significant differences for each PC comparison. **C.** Heatmaps of the differentially methylated CpGs between B-PLL and CLL (n=5,058), B-PLL and MCL (n=2,755) and B-PLL and SMZL (n = 60,841). In the case of SMZL, the top 10,000 CpGs with the highest standard deviation are represented. **D.** Heatmap of the 31 simultaneous differentially methylated CpGs comparing B-PLL to CLL, MCL and SMZL. FDR levels of significance: FDR<0.001=\*\*\*, FDR <0.01=\*\* and FDR<0.05=\*.

**Figure 2. Differential DNA methylation patterns of B-PLL compared to CLL/PL and MCL P/B.** **A.** Heatmap of the differentially methylated CpGs between B-PLL and CLL/PL (n=3,197). **B.** PCA of the differentially methylated CpGs between B-PLL and CLL with less than 10% prolymphocytoid cells (n=2,906), including also the CLL/PL (showing the percentage of cells with prolymphocytoid morphology) and the paired CLL/pPL case. The percentage of prolymphocytes of each of the CLL/PL samples is indicated. **C.** Heatmap of the differentially methylated CpGs between B-PLL and MCL P/B (n=10,922). **D.** PCA of the differentially methylated CpGs between B-PLL and C1 MCL with classic morphology (n=8.809) including in the plot the C1 MCL P/B.

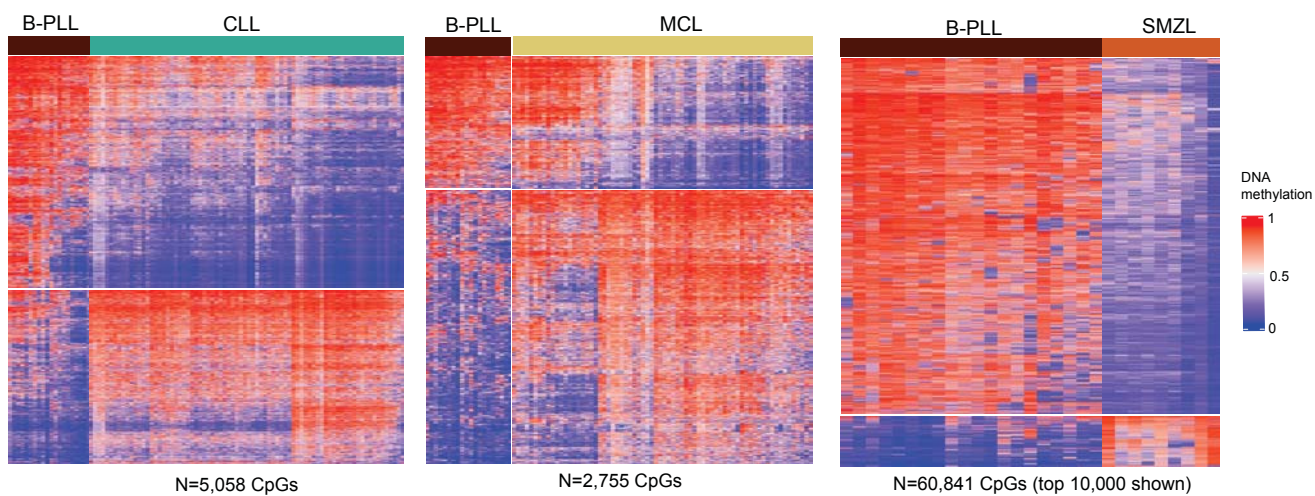
**Figure 3. Identification of two B-PLL epitypes with signatures related to normal B-cell differentiation.** **A.** PCA analysis showing PC1 and PC2 with B-PLL and normal B cell samples. On its side, the boxplots display differences between B-PLL epitypes and normal B cells. NBC: naive B cells; GC: germinal center; MBC: mature B cells, PC: plasma cells. **B.** Heatmap of the differentially methylated CpGs (n=5,465) between B-PLL epitypes. **C.** Heatmap of the differentially methylated CpGs between B-PLL epitypes related to methylation patterns of NBC and MBC (n=379, out of which 249 are hypermethylated and 130 hypomethylated in epitype 1). **D.** Percentage of IGHV gene identity B-PLL epitypes. **E.** Magnitude of the epiCMIT scores between B-PLL epitypes.

**Figure 4. Epigenetic and genetic characterization of the differential signature of the B-PLL epitypes.** **A.** Heatmap of the *de novo* epitype 1 specific methylation signature (n=799 CpGs) **B.** Oncoplot showing genetic alterations of B-PLL cases. **C.** Heatmap of 97 differentially expressed genes between B-PLL epitypes. **D.** Gene set enrichment analysis (GSEA) showing enrichment of MYC targets in epitype 1 B-PLL **E.** MYC expression percentile of B-PLL, CLL and MCL cases **F.** Barplot of the numbers of differentially methylated CpGs between CD5+ epitype 2 and CD5- epitype 2 (top) and heatmap of DM CpGs between CD5+ epitype 1 and CD5+ epitype 2 (n=2,014, bottom) **G.** Kaplan-Meier curves for overall survival between B-PLL epitypes.

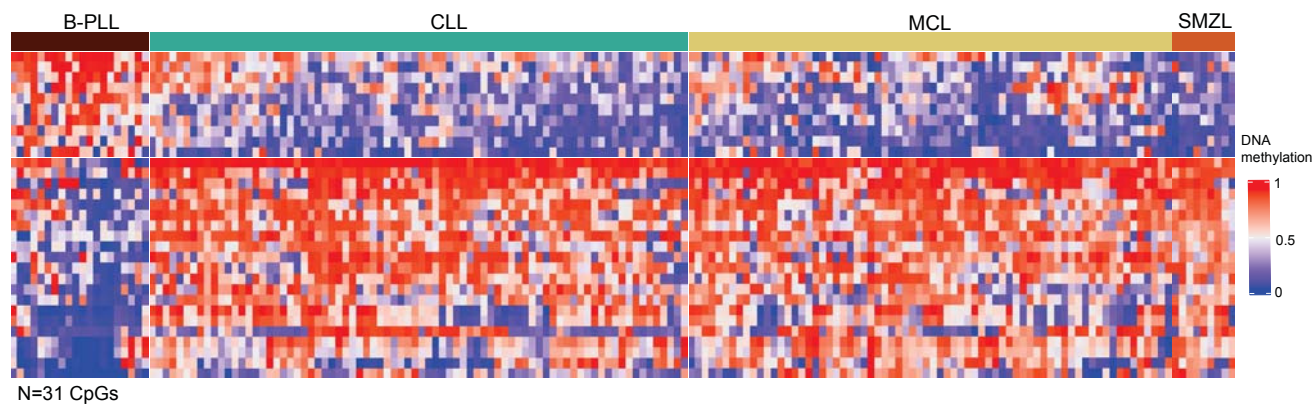
# Figure 1



**C**



**D**



# Figure 2

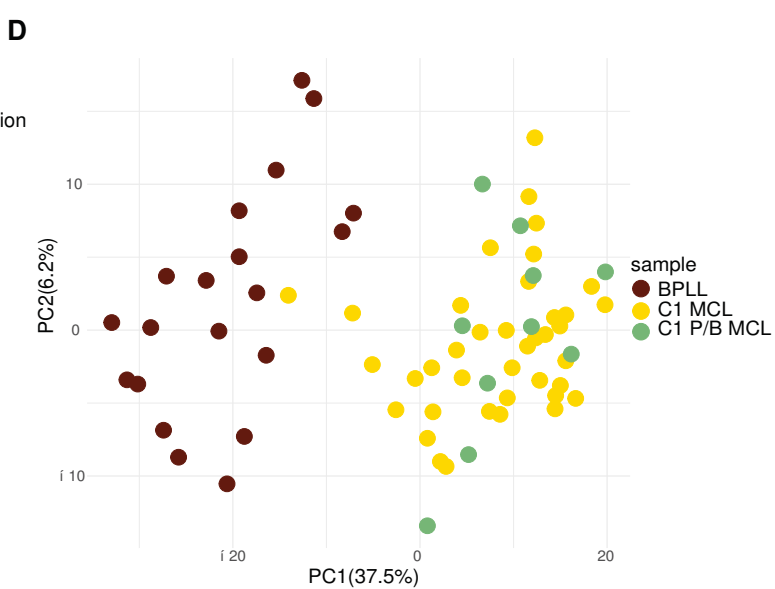
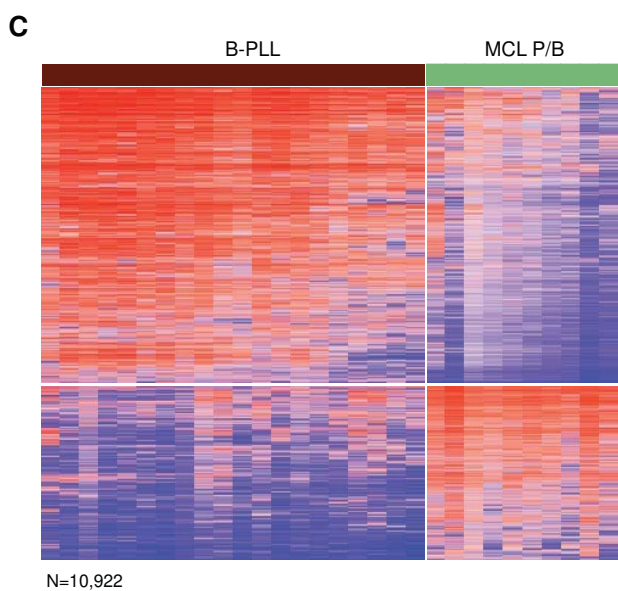
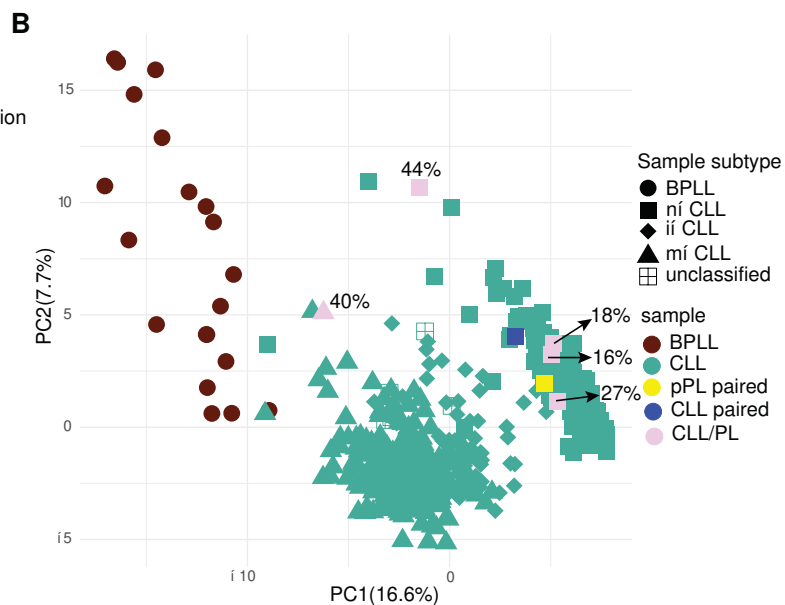
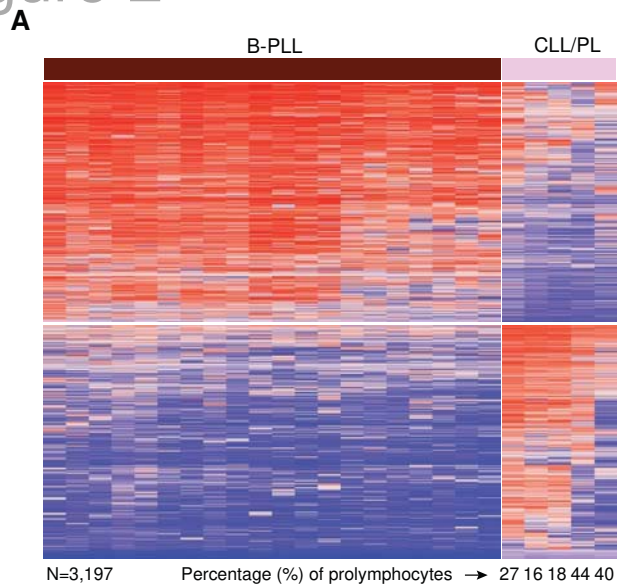
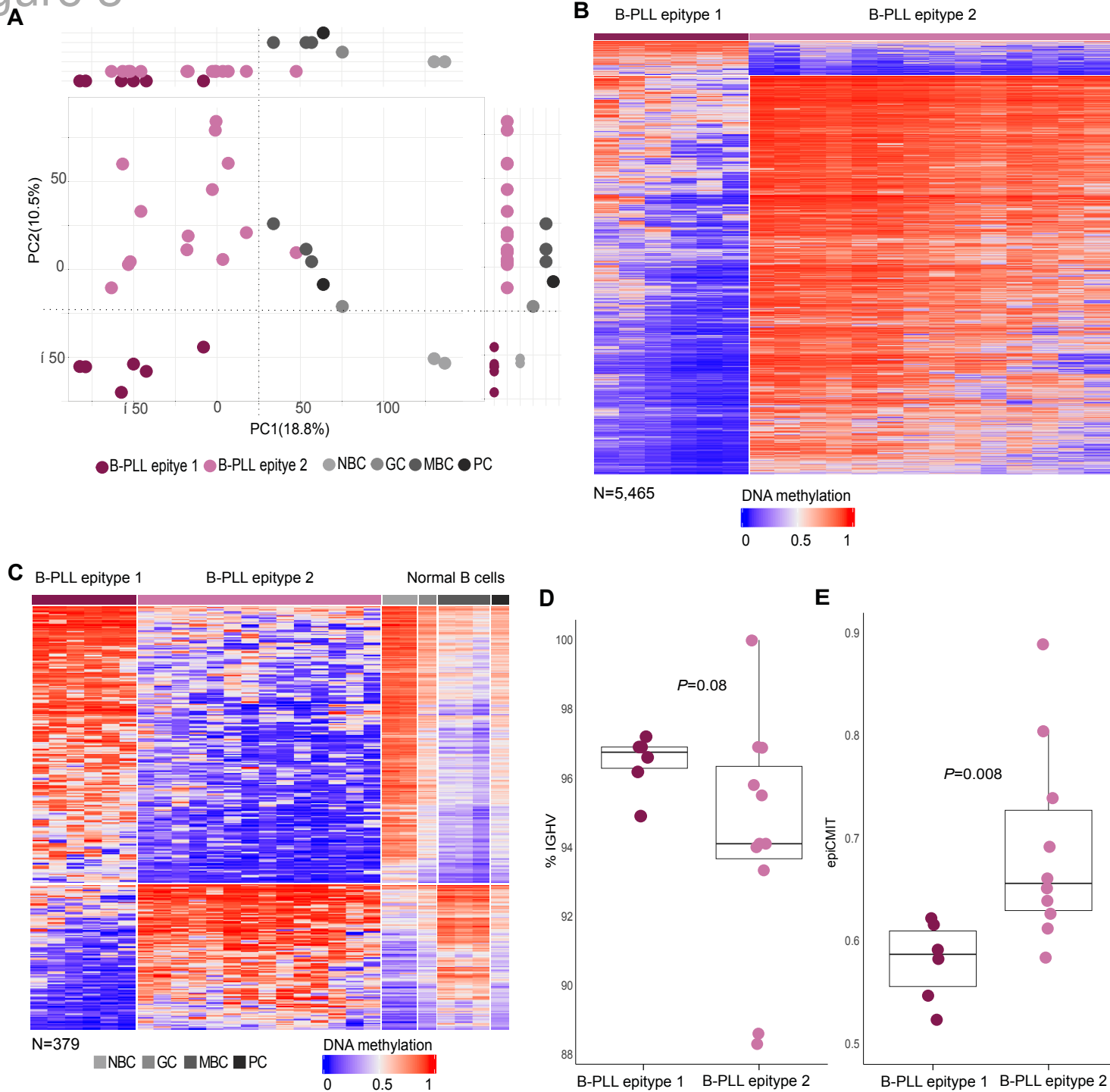


Figure 3



# Figure 4

