Received: 24 February 2025 | Accepted: 17 August 2025

DOI: 10.1002/hem3.70211

HemaSphere eha

LETTER

Identification of two epitypes of IGHV unmutated chronic lymphocytic leukemia with distinct B-cell-related epigenetic imprinting and genetic alterations

Stella Charalampopoulou[1,2,3] | Silvia Ramos-Campoy[4] | Marti Duran-Ferrer[1,3] | Anna Puiggros[4] | Joanna Kamaso[4] | Florence Nguyen-Khac[5,6,7] | Gian Matteo Rigolin[8] | Antonio Cuneo[8] | Rosa Collado[9] | Rocío García-Serra[9] | Claudia Haferlach[10] | Margarita Ortega[11] | Pau Abrisqueta[11] | Francesc Bosch[11] | Thorsten Zenz[12] | María Laura Blanco[13] | Rocío Salgado[14] | Mª Dolores García-Malo[15] | Eva Gimeno[16] | Armando Lopez-Guillermo[1,3,17] | Elias Campo[1,3,17,18] | Laurence Etter[19] | Jacqueline Schoumans[19] | Blanca Espinet[4,^] | Jose I. Martin-Subero[1,3,18,20,^]

Correspondence: Blanca Espinet (bespinet@psmar.cat) and Jose I. Martin-Subero (imartins@recerca.clinic.cat)

Chronic lymphocytic leukemia (CLL), the most frequent leukemia in the Western world, is characterized by extensive biological and clinical heterogeneity.[1] Immunogenetic, genetic, and epigenetic features have been described as major molecular traits strongly associated with clinical impact in CLL.[2] The levels of somatic hypermutation (SHM) of the immunoglobulin heavy chain variable (IGHV) region segregate cases into IGHV unmutated and mutated CLLs (U-CLL and M-CLL, respectively), being the former clinically more aggressive.[1] This dichotomous classification was further refined using DNA methylation imprints of normal B-cell maturation, leading to the identification of three distinct epigenetic subtypes or epitypes, including naive-like/low-programmed CLL (n-CLL), mostly corresponding to U-CLL, intermediate CLL and memory-like/high-programmed CLL, both of which belonging to M-CLL.[3,4] In spite of this classification into three categories, a model in which CLLs derive from a continuum of mature B-cell maturation stages was proposed, suggesting that CLL could be potentially stratified into additional epitypes.[4] Apart from this link with normal B-cell maturation, the DNA methylome of CLL also reflects the past proliferative history and presents specific DNA methylation signatures related to its pathogenesis.[5,6]

In addition to IGHV subtypes and epitypes, specific genetic lesions and the presence of complex karyotypes (CKs) carry prognostic[1,2,7,8] and predictive value.[9] Such CK has been frequently but not always associated with U-CLL and alterations affecting the

[1]Institut d'Investigacions Biomèdiques August Pi i Sunyer (IDIBAPS), Barcelona, Spain
[2]Programa de doctorat en Biomedicina, Universitat de Barcelona, Barcelona, Spain
[3]Centro de Investigación Biomédica en Red de Cáncer (CIBERONC), Madrid, Spain
[4]Molecular Cytogenetics Laboratory, Pathology Department, Hospital del Mar and Translational Research on Hematological Neoplasms Group, Cancer Programme, Hospital del Mar Research Institute, Barcelona, Spain
[5]Service d'Hématologie Biologique, Hôpital Pitié-Salpêtrière, AP-HP, Sorbonne Université, Paris, France
[6]Centre de Recherche des Cordeliers, Sorbonne Université, Paris, France
[7]Inserm UMRS 1138, Drug Resistance in Hematological Malignancies Team, Université Paris Cité, Paris, France
[8]Hematology Section, St. Anna University Hospital, Ferrara, Italy
[9]Consorcio Hospital General Universitario de Valencia, Valencia, Spain
[10]MLL Munich Leukemia Laboratory, Munich, Germany
[11]Department of Hematology, University Hospital Vall d'Hebron, Barcelona, Spain
[12]Department of Medical Oncology and Hematology, University Hospital and University of Zürich, Zurich, Switzerland
[13]Department of Hematology, Hospital de la Santa Creu i Sant Pau, Barcelona, Spain
[14]Cytogenetics Laboratory, Hematology Department, Fundación Jiménez Díaz, Madrid, Spain
[15]Hematology and Medical Oncology Services, Hospital Universitario Morales Meseguer, Murcia, Spain
[16]Department of Hematology, Hospital del Mar, Barcelona, Spain
[17]Hospital Clínic de Barcelona, Barcelona, Spain
[18]Departament de Fonaments Clinics, Facultat de Medicina, Universitat de Barcelona, Barcelona, Spain
[19]Oncogenomic Laboratory, Hematology Service, Lausanne University Hospital, Lausanne, Switzerland
[20]Institució Catalana de Recerca i Estudis Avançats (ICREA), Barcelona, Spain
^These authors share senior authorship.

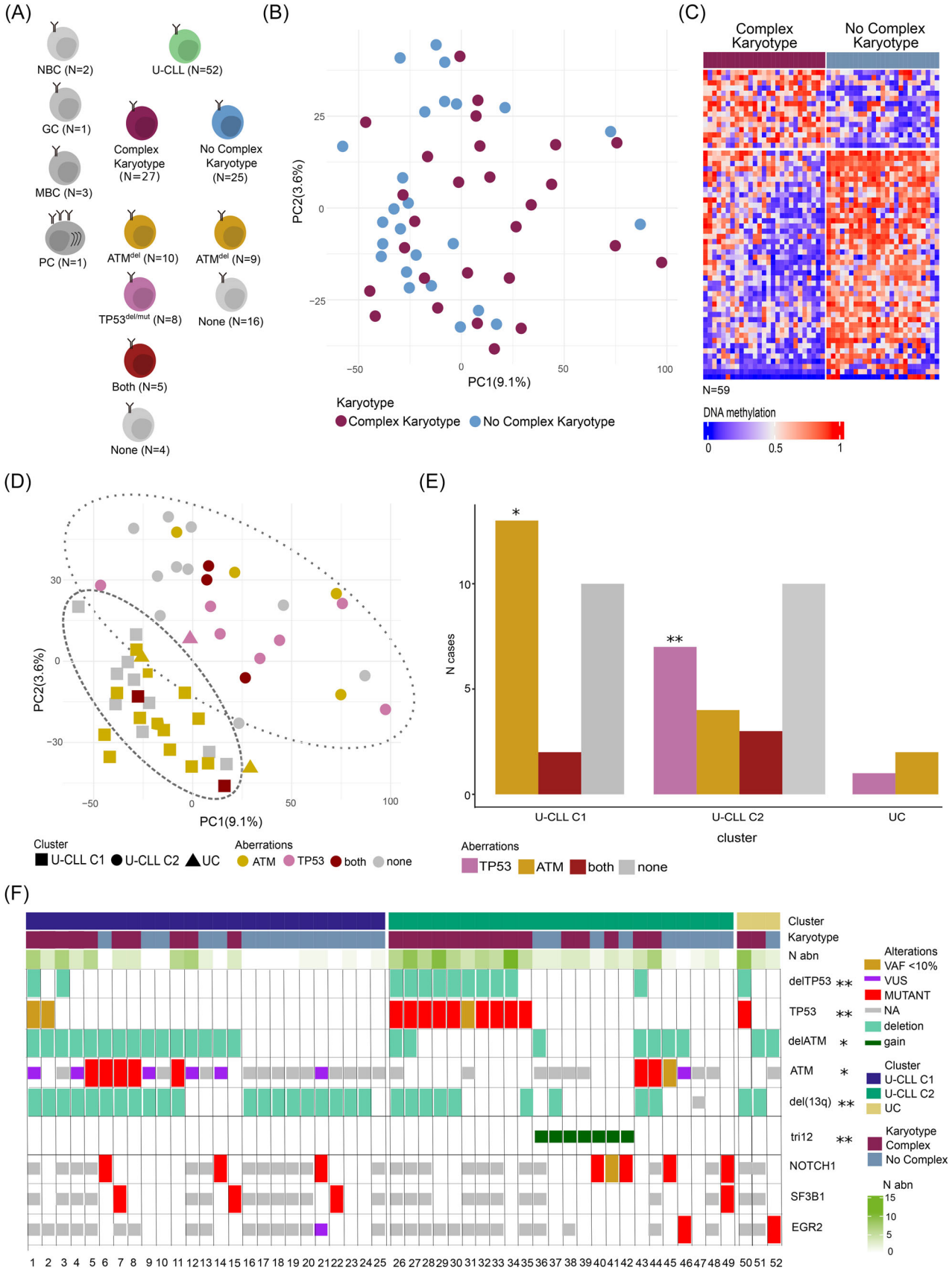

FIGURE 1 (See caption on next page).

FIGURE 1 **Exploration of the relationship between complex karyotype and DNA methylation, and identification of two epitypes. (A)** Experimental design. **(B)** Principal component analysis (PCA) of the whole dataset with annotation of the karyotypic complexity. **(C)** Heatmap of the differentially methylated (DM) CpGs between unmutated chronic lymphocytic leukemia (U-CLL) C1 and U-CLL C2 ($n = 59$). **(D)** PCA of the whole dataset with annotation of the two identified epitypes and the status of *ATM*/*TP53* aberrations. **(E)** Barplots representing the distribution of the *ATM*/*TP53* aberrations in the identified epitypes. **(F)** Genetic characterization of the two epitypes. GC, germinal center B cell; MBC, memory B cell; NA, not available/not tested; NBC, naive B cell; PC, plasma B cell; UC, unclassified; VAF, variant allele frequency; VUS, variant of unknown significance.

TP53 and *ATM* genes.[10,11] To our knowledge, studies analyzing the potential link between CK and epigenetic features have not been reported. Therefore, we have here analyzed the DNA methylome of a U-CLL cohort ($n = 52$, Supporting Information S2: Table 1, all belonging to the n-CLL epitype according to a previously reported classifier[5]) that has been carefully selected based on distinct CK levels and the presence of *ATM* deletion (ATM^{del}) and/or *TP53* deletion or mutation ($TP53^{del/mut}$) status (Figure 1A). We used Illumina EPIC V1 arrays to generate DNA methylation profiles, which were analyzed as previously described[5] (Supporting Information S9: Supplementary Material).

An unsupervised principal component analysis (PCA) showed that CK and non-CK U-CLLs are intermingled (Figure 1B), and the same finding was made when analyzing additional principal components (Supporting Information S1: Figure 1A). In line with this observation, a differential methylation analysis revealed a small and heterogeneous signature of 59 differentially methylated (DM) CpGs between CK and non-CK U-CLLs (Figure 1C, Supporting Information S3: Table 2). Comparing non-CK only with cases with high CK (≥5 aberrations), we obtained a still relatively small signature of 315 CpGs (Supporting Information S1: Figure 1B, Supporting Information S4: Table 3). Overall, these results suggest that despite the biological and clinical relevance of CK and DNA methylation in CLL, they do not seem to be strongly related to each other in our series.

Beyond the poor association between CK and DNA methylation, a careful inspection of the results pointed to an unexpected finding. We observed that the DNA methylome segregated U-CLLs into two groups that were statistically validated using a consensus clustering approach (Supporting Information S1: Figure 2A,B and Supporting Information S9: Supplementary Methods). Most interestingly, one cluster (U-CLL C1, $n = 25$) was strongly enriched in ATM^{del} cases ($P = 0.015$), whereas the other (U-CLL C2, $n = 24$) contained all the $TP53^{del/mut}$ cases ($P = 0.004$) (Figure 1D,E). Three cases could not be unequivocally classified due to low clustering probability (unclassified, UC) and were excluded from all downstream analyses. Further genetic characterization of the two clusters showed that U-CLL C1 was also enriched in *ATM* mutations ($P = 0.017$) and del(13q) ($P = 0.007$), whereas U-CLL C2 was also related to trisomy 12 (tri12) ($P = 0.004$) (Figure 1F), suggesting that the clusters were associated with particular genetic alterations beyond those involving *ATM* and *TP53*. Of note, the two clusters were unrelated to CK ($P = 0.15$, Figure 1F).

We next compared the DNA methylome of U-CLL C1 versus C2 and identified 4333 DM CpGs (Figure 2A,B, Supporting Information S5: Table 4, and Supporting Information S9: Supplementary Methods). To interpret this differential signature, we divided the CpGs into those whose methylation levels change or remain stable during B-cell differentiation, which we termed B-cell-related and B-cell-independent, respectively (Supporting Information S9: Supplementary Methods). A total of 2183 CpGs were B-cell independent (50.4%, Figure 2A), whereas 2150 were B-cell-related (49.6%, Figure 2B and Supporting Information S5: Table 4). The B-cell-independent signature was mainly associated with a de novo methylation loss in U-CLL C2 ($n = 1679$ CpGs, Figure 2C, Supporting Information S6: Table 5, and Supporting Information S1: Figure 3A). In contrast, U-CLL C1 only showed 118 de novo hypomethylated and 53 de novo hypermethylated CpGs (Figure 2C, Supporting Information S7: Table 6, and Supporting Information S1: Figure 3B). The hypomethylation signatures in both clusters were located in gene bodies and in regions with low CpG content, and were enriched in regulatory elements such as enhancers (Supporting Information S1: Figure 3C–H). On the other hand, the U-CLL C1 de novo hypermethylation was preferentially located in CpG-rich regions and poised promoters (Supporting Information S1: Figure 3I–K). A transcription factor binding site analysis of both B-cell-related and independent differential CpGs did not reveal clear enrichments, with the exception of a highly significant association between the de novo hypomethylation signature of U-CLLs C2 and the NFAT TF (Supporting Information S1: Figure 3L), which has been commonly linked to epigenetic signatures in CLL.[4,5,12]

Remarkably, when we analyzed B-cell-related differential CpGs, we noticed that the higher methylation levels in U-CLL C1 resembled pre-germinal center (GC) B cells, while methylation loss in U-CLL C2 was shared with GC-experienced B cells (Figure 2B). Congruent with this observation, Component 1 of a PCA using these differential CpGs clustered U-CLL C1 cases together with naïve B cells and U-CLL C2 cases together with GC B, memory B, and plasma cell samples (Figure 2D). In fact, a more stringent analysis with a methylation difference of 0.25 revealed 1243 CpGs with similar methylation levels between U-CLL C1 and naïve B cells and between U-CLL C2 and post-GC B cells (Supporting Information S1: Figure 4A and Supporting Information S8: Table 7). Of note, IGHV SHM was remarkably low in both groups, as they mostly contained cases with 100% IGHV homology (Supporting Information S1: Figure 4B). Overall, these findings suggest that these two U-CLL epitypes may derive from GC-independent B cells lacking detectable SHM but showing markedly different levels of epigenetic programming related to B-cell maturation. These results are aligned with the model proposed by Oakes and colleagues, in which CLLs may derive from a continuum of B-cell maturation stages that can be segregated into discrete epitypes.[4] Supporting this notion, a phylogenetic analysis of our U-CLL cohort together with other CLL samples[13] showed a continuum that can be segregated into four instead of the three previously reported epitypes[2–4] (Figure 2E). We show that low-programmed CLLs (i.e., U-CLLs) can be separated into two epitypes that are, respectively, enriched in *ATM* alterations/del(13q), and *TP53* alterations/tri12 (Figure 1F). Furthermore, U-CLL C1 presents a significantly lower proliferative history measured by the epiCMIT clock ($P = 0.012$; Figure 2F), which further strengthens the concept that they are derived from B cells with lower levels of epigenetic programming as compared to U-CLL C2. From the clinical perspective, these two epitypes did not show significant differences in overall survival (OS) time (Supporting Information S1: Figure 4C) or time to first treatment (TTT) (Supporting Information S1: Figure 4D). However, in the group with more variable epiCMIT levels (i.e., U-CLL C2), a higher epiCMIT score was significantly associated with a shorter TTT (Supporting Information S1: Figure 4E,F), as previously reported.[5]

We recognize that the initial U-CLL series had a selection bias towards CK with *TP53* and *ATM* alterations. Therefore, to evaluate whether the identification of two new U-CLL/n-CLL epitypes could be generalized beyond this initial cohort, we used two previously published population-based U-CLL/n-CLL cohorts with 159 (ICGC,[14]

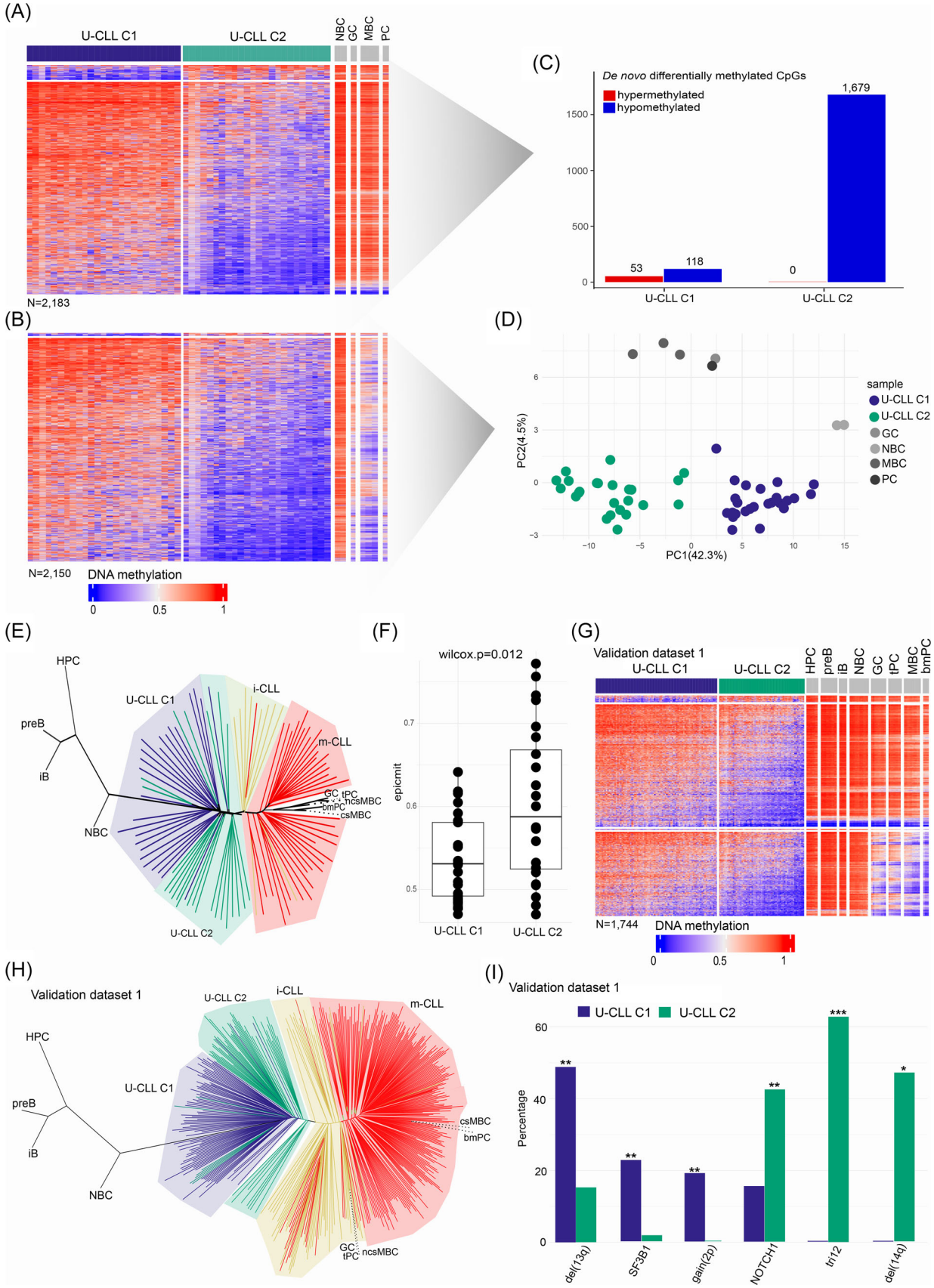

FIGURE 2 (See caption on next page).

FIGURE 2 **Epigenetic characterization of the two epitypes and validation of the results in an independent unmutated chronic lymphocytic leukemia (U-CLL) cohort. (A)** Heatmap of the B-cell independent differentially methylated (DM) CpGs between U-CLL C1 and U-CLL C2 ($n = 2183$). **(B)** Heatmap of the B-cell-related DM CpGs between U-CLL C1 and U-CLL C2 ($n = 2150$). **(C)** Barplots representing the number of the de novo specific CpGs of the two epitypes. **(D)** Principal component analysis (PCA) of the B-cell-related DM CpGs between U-CLL C1 and U-CLL C2. **(E)** Phylogenetic analysis of a previously published set of CpGs significant for B-cell differentiation ($n = 1137$). **(F)** Boxplot of the epiCMIT clock in the two epitypes. **(G)** Heatmap of the DM CpGs between U-CLL C1 and U-CLL C2 that are present in the 450k validation dataset ($n = 1744$). **(H)** Phylogenetic analysis of a previously published set of CpGs significant for B-cell differentiation in the validation dataset ($n = 1220$). **(I)** Barplots representing the significantly enriched genetic aberrations in U-CLL C1 and U-CLL C2 of the validation dataset. bmPC, bone marrow plasma cell; GC, germinal center B cell; HPC, hematopoietic progenitor cell; iB, immature B cell; MBC, memory B cell; m-CLL, mutated chronic lymphocytic leukemia; NBC, naive B cell; PC, plasma cell; preB, pre-B cell; tPC, tonsilar plasma cell; UC, unclassified.

profiled with 450k arrays) and 87 cases (59 profiled with 450k arrays and 28 with EPIC V1 arrays[13]). For the first cohort, we could also identify two epitypes using 1744 CpGs present in the 450k array that overlap with the 4333 differential CpGs detected with the EPIC V1 array in the initial series (Supporting Information S1: Figure 5A,B), and 13 samples had lower cluster probabilities and were left UC. The two identified clusters (U-CLL C1, $n = 86$; U-CLL C2, $n = 60$) showed similar methylation patterns to those previously identified (Figure 2G). In addition, a phylogenetic analysis widely recapitulated our previous findings, where U-CLL C1 cases overall showed less B-cell epigenetic programming than U-CLL C2 (Figure 2H). We also found that U-CLL C1 was enriched in del(13q) ($P = 0.004$) and U-CLL C2 in tri12 ($P < 0.001$) as well as other particular genetic alterations in both clusters, including *SF3B1* mutations ($P = 0.004$) and gain(2p) ($P = 0.004$) in U-CLL C1, and *NOTCH1* mutations ($P = 0.004$) and del(14q) ($P = 0.01$) in U-CLL C2 (Figure 2I, Supporting Information S1: Figure 5C). As in the discovery series, IGHV SHM was low in both groups, and the noncanonical activation-induced cytidine deaminase (AID) mutational signature (SBS9) identified by whole-genome sequencing (WGS) was likewise low (Supporting Information S1: Figure 5D,E). The epiCMIT score was also found to be significantly lower in U-CLL C1 ($P = 0.005$; Supporting Information S1: Figure 5F). Finally, the two clusters did not show differences in OS or TTT in this validation series either (Supporting Information S1: Figure 5G,H).

For the second cohort, we analyzed the 1749/4333 DM CpGs that were present in the normalized matrix of 450k array data and also identified two clusters with methylation patterns similar to those in the discovery cohort (63 C1 and 24 C2 U-CLLs, Supporting Information S1: Figure 6A). The corresponding phylogenetic analysis also pointed to a lower maturation level in U-CLL C1 (Supporting Information S1: Figure 6B) and a lower epiCMIT score ($P = 0.003$; Supporting Information S1: Figure 6C). Finally, U-CLL C1 was enriched in ATM^{del} ($P = 0.01$) and del(13q) ($P = 0.006$) whereas U-CLL C2 was linked to *NOTCH1* mutation ($P = 0.02$) and tri12 ($P < 0.001$) (Supporting Information S1: Figure 6D). Overall, these findings corroborate our observations in the initial and first validation cohort.

In summary, although this project was initiated with the intention of exploring the link between CK and DNA methylation, their association was minor in the studied U-CLL cohort. Instead, we serendipitously detected two U-CLL epitypes, which were validated in two additional U-CLL cohorts. Although these two epitypes did not further stratify U-CLL into distinct clinical behaviors, they showed markedly different biological features. These include different levels of B-cell-related epigenetic imprinting and enrichment in distinct genetic alterations. Such enrichment suggests that specific B-cell maturation stages are more prone to acquire particular genetic alterations, such as, for example, ATM^{del} in U-CLL C1, $TP53^{del/mut}$/tri12 in U-CLL C2, *SF3B1*/IGLV3-21R110 mutations in M-CLL of the intermediate-programmed epitype,[15] and *MYD88* in the M-CLL/high-programmed.[14] This study further highlights the power of integrative (epi)genetic profiling to identify, within a continuum of maturation stages, CLL epitypes linked to specific genetic alterations.

AUTHOR CONTRIBUTIONS

Stella Charalampopoulou: Investigation; methodology; visualization; software; data curation; validation; writing—original draft. **Silvia Ramos-Campoy**: Investigation; methodology; writing—review and editing; data curation; resources. **Marti Duran-Ferrer**: Investigation; writing—review and editing; methodology; validation; data curation; software; visualization. **Anna Puiggros**: Investigation; writing—review and editing; methodology; data curation. **Joanna Kamaso**: Resources; writing—review and editing; data curation. **Florence Nguyen-Khac**: Data curation; resources; writing—review and editing. **Gian Matteo Rigolin**: Resources; data curation; writing—review and editing. **Antonio Cuneo**: Data curation; resources; writing—review and editing. **Rosa Collado**: Data curation; resources; writing—review and editing. **Rocío García-Serra**: Data curation; resources; writing—review and editing. **Claudia Haferlach**: Data curation; resources; writing—review and editing. **Margarita Ortega**: Data curation; resources; writing—review and editing. **Pau Abrisqueta**: Data curation; resources; writing—review and editing. **Francesc Bosch**: Data curation; resources; writing—review and editing. **Thorsten Zenz**: Resources; writing—review and editing; validation. **María Laura Blanco**: Data curation; resources; writing—review and editing. **Rocío Salgado**: Data curation; resources; writing—review and editing. **M^{a} Dolores García-Malo**: Data curation; Resources; Writing—review and editing. **Eva Gimeno**: Data curation; resources; writing—review and editing. **Armando Lopez-Guillermo**: Resources; validation; writing—review and editing. **Elias Campo**: Resources; validation; writing—review and editing. **Laurence Etter**: Data curation; resources; writing—review and editing. **Jacqueline Schoumans**: Data curation; resources; writing—review and editing. **Blanca Espinet**: Conceptualization; investigation; funding acquisition; writing—original draft; methodology; project administration; data curation; supervision; resources. **Jose I. Martin-Subero**: Conceptualization; investigation; funding acquisition; writing—original draft; methodology; validation; visualization; project administration; data curation; supervision; resources.

CONFLICT OF INTEREST STATEMENT

The authors declare no conflicts of interest.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are openly available in Gene Expression Omnibus at https://www.ncbi.nlm.nih.gov/geo/, reference number GSE289415.

FUNDING

This work was partly supported by grants from Generalitat de Catalunya (2021-SGR-00025), Gilead Sciences Fellowship (GLD17/00282), Ministerio de Universidades of Spain (FPU17/00361), Fundación Española de Hematología y Hemoterapia (FEHH)-AstraZeneca, European Research Council (ERC) under the European Union's Horizon 2020 Research and Innovation Programme (grant 810287), Generalitat de Catalunya Suport Grups

de Recerca AGAUR (2021-SGR-01343 and 2021-SGR-01172), and CIBERONC (CB16/12/00225, CB16/12/00334, and CB24/12/00009). This work was partially developed at the Centre Esther Koplowitz (Barcelona, Spain).

SUPPORTING INFORMATION

Additional supporting information can be found in the online version of this article.

ORCID

Gian Matteo Rigolin https://orcid.org/0000-0002-8370-5190
Rocío García-Serra https://orcid.org/0000-0003-4857-1016
Jose I. Martin-Subero https://orcid.org/0000-0001-8809-5195

REFERENCES

1. Kipps TJ, Stevenson FK, Wu CJ, et al. Chronic lymphocytic leukaemia. *Nat Rev Dis Primers*. 2017;3(1):16096. doi:10.1038/nrdp.2016.96
2. Nadeu F, Diaz-Navarro A, Delgado J, Puente XS, Campo E. Genomic and epigenomic alterations in chronic lymphocytic leukemia. *Annu Rev Pathol Mech Dis*. 2020;15(1):149-177. doi:10.1146/annurev-pathmechdis-012419-032810
3. Kulis M, Heath S, Bibikova M, et al. Epigenomic analysis detects widespread gene-body DNA hypomethylation in chronic lymphocytic leukemia. *Nat Genet*. 2012;44(11):1236-1242. doi:10.1038/ng.2443
4. 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. doi:10.1038/ng.3488
5. 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. doi:10.1038/s43018-020-00131-2
6. Duran-Ferrer M, Martín-Subero JI. Epigenomic characterization of lymphoid neoplasms. *Annu Rev Pathol Mech Dis*. 2024;19(1):371-396. doi:10.1146/annurev-pathmechdis-051122-100856
7. Baliakas P, Jeromin S, Iskas M, et al. Cytogenetic complexity in chronic lymphocytic leukemia: definitions, associations, and clinical impact. *Blood*. 2019;133(11):1205-1216. doi:10.1182/blood-2018-09-873083
8. 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. doi:10.1038/s41588-022-01140-w
9. Kittai AS, Miller C, Goldstein D, et al. The impact of increasing karyotypic complexity and evolution on survival in patients with CLL treated with ibrutinib. *Blood*. 2021;138(23):2372-2382. doi:10.1182/blood.2020010536
10. Puiggros A, Collado R, Calasanz MJ, et al. Patients with chronic lymphocytic leukemia and complex karyotype show an adverse outcome even in absence of *TP53/ATM FISH* deletions. *Oncotarget*. 2017;8(33):54297-54303. doi:10.18632/oncotarget.17350
11. Visentin A, Bonaldi L, Rigolin GM, et al. The combination of complex karyotype subtypes and IGHV mutational status identifies new prognostic and predictive groups in chronic lymphocytic leukaemia. *Br J Cancer*. 2019;121(2):150-156. doi:10.1038/s41416-019-0502-x
12. Beekman R, Chapaprieta V, Russiñol N, et al. The reference epigenome and regulatory chromatin landscape of chronic lymphocytic leukemia. *Nat Med*. 2018;24(6):868-880. doi:10.1038/s41591-018-0028-4
13. Dietrich S, Oleś M, Lu J, et al. Drug-perturbation-based stratification of blood cancer. *J Clin Invest*. 2017;128(1):427-445. doi:10.1172/jci93801
14. Puente XS, Beà S, Valdés-Mas R, et al. Non-coding recurrent mutations in chronic lymphocytic leukaemia. *Nature*. 2015;526(7574):519-524. doi:10.1038/nature14666
15. Nadeu F, Royo R, Clot G, et al. Iglv3-21^{R110} identifies an aggressive biological subtype of chronic lymphocytic leukemia with intermediate epigenetics. *Blood*. 2021;137(21):2935-2946. doi:10.1182/blood.2020008311