

# An immuno-(epi)genomic classification of chronic lymphocytic leukemia refines outcome prediction

Martí Duran-Ferrer, Larry Mansouri, Guillem Clot, Ferran Nadeu, Sujata Bhoi, Lesley Ann Sutton, Panagiotis Baliakas, Sara Ek, Venera Kuci Emruli, Karla Plevova, Zadie Davis, Hanna Goransson-Kultima, Anders Isaksson, Karin E Smedby, Gianluca Gaidano, Anton W Langerak, Frederic Davi, Davide Rossi, David Oscier, Sarka Pospisilova, Maria Karypidou, Andreas Agathangelidis, Junyan Lu, Thorsten Zenz, Julio Delgado, Armando López-Guillermo, Paolo Ghia, Elías Campo, Kostas Stamatopoulos, Richard Rosenquist, José I. Martín-Subero.  
Contact: [maduran@recerca.clinic.cat](mailto:maduran@recerca.clinic.cat), IDIBAPS, Barcelona, Spain.

## INTRODUCTION

Immunogenetic features, such as IGHV and IGLV3-21<sup>R110</sup> (R110) mutational status, and B-cell receptor stereotypy, along with epigenetic features, such as the 3 epitypes (naïve-like, n-CLL; intermediate, i-CLL; and memory-like, m-CLL), contribute to the pathogenesis of chronic lymphocytic leukemia (CLL). However, the interplay between them remains largely unexplored, primarily due to the scarcity of matched data. Here, we report a comprehensive analysis of a cohort of 995 patients enriched in cases harboring the R110 mutation (n=99) and the most frequent stereotyped subsets (#1-8, n=180).

## OBJECTIVES

To gain new biological and clinical insights into the relationship between immunogenetics and epigenetics in the pathogenesis of CLL.

## METHODS

We systematically profiled the DNA methylome of 995 CLL patients (Illumina arrays), the chromatin landscape (H3K27ac-ChIP-seq, n=99; ATAC-seq, n=101), the transcriptome (RNA-seq, n=285) and correlated with time to first treatment (TTFT) and overall survival (OS) (n=885) from diagnosis. Immunogenetic features were available for 893 (IGHV status), 829 (CLL stereotypes), and 816 patients (IGLV3-21<sup>R110</sup> mutation status). The data was divided into a discovery cohort (n=631) and validation cohort (n=364).

## CONCLUSIONS

Collectively, our results identify two new immuno-(epi)genetic clusters, reflecting mainly IGHV and R110 mutational status together with epitype subgroups. We also identify discordant patients exhibiting molecular and clinical features that align more closely with their cluster assignment than with their immunogenetic status. These clusters, together with epitypes and the epiCMIT proliferative history score appear to improve the prognostic power of IGHV and R110 mutational status, and can be detected with a single DNA methylation assay.

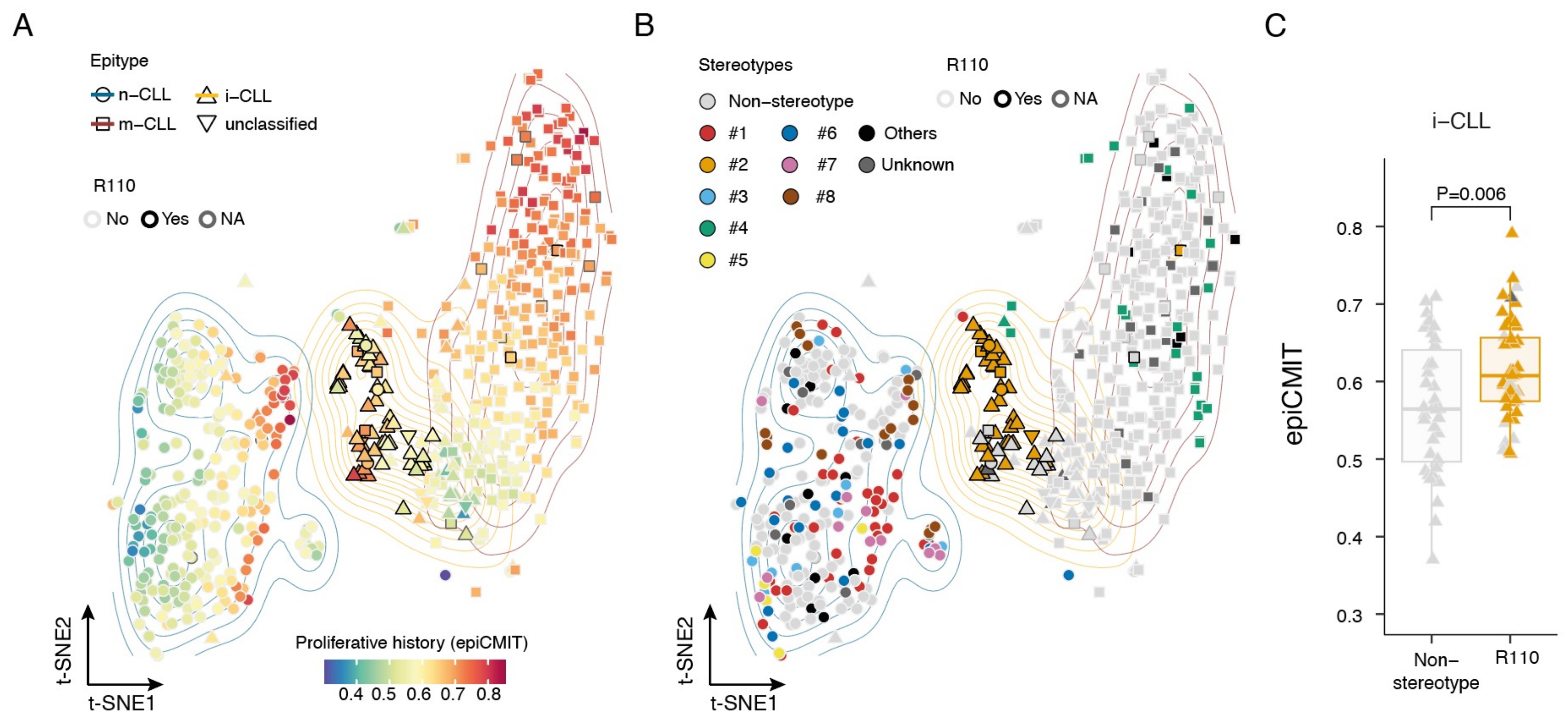
## ACKNOWLEDGEMENTS



## RESULTS

### 1. THE DNA METHYLOME OF IMMUNOGENETIC CLL SUBGROUPS

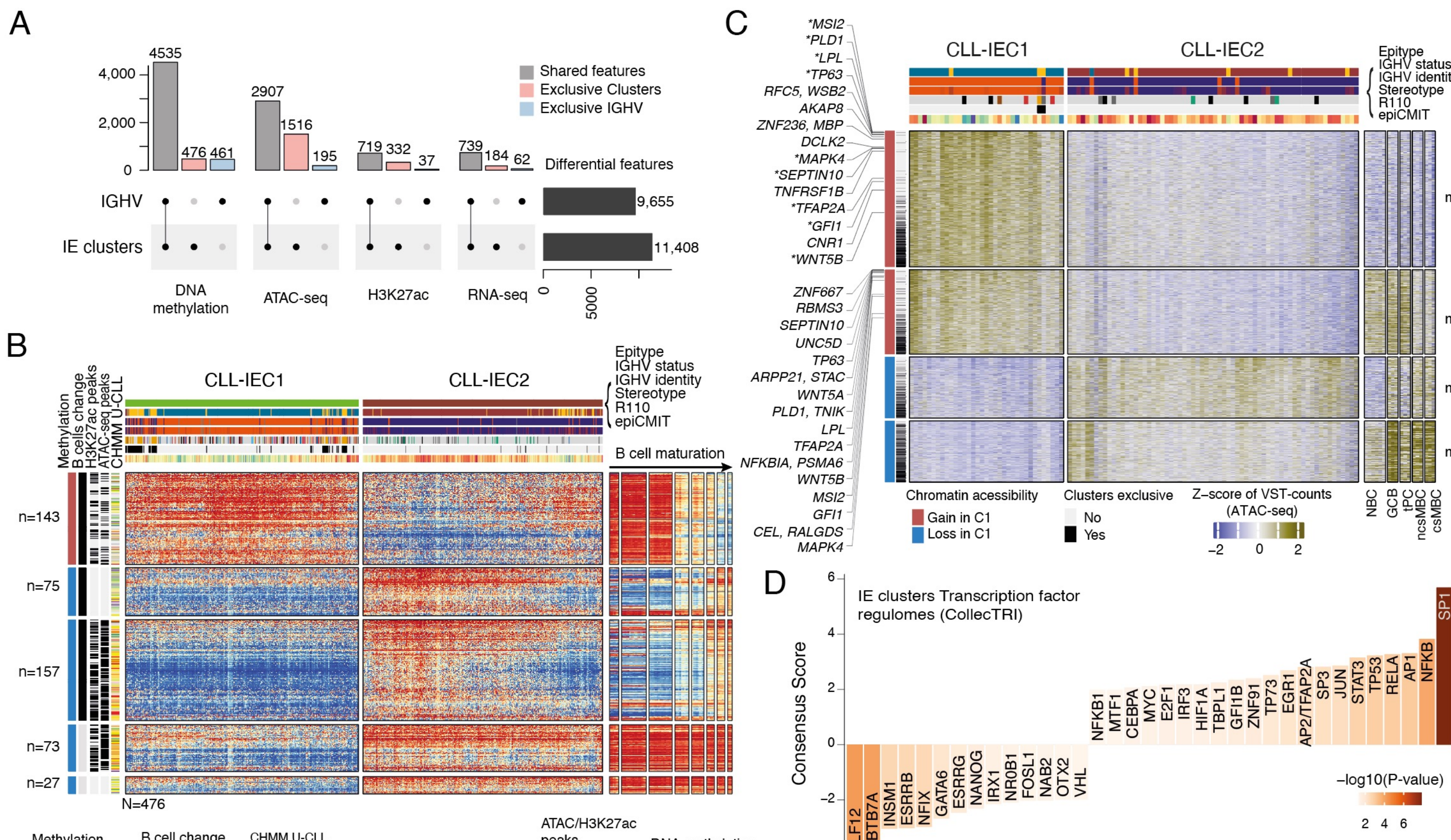
The main sources of DNA methylation variability were related to components of CLL cellular memory, including the cell-of-origin epitypes and the proliferative history score epiCMIT (**Fig. 1A**). CLL stereotypes were not strongly related to DNA methylation variability (**Fig. 1B**). Conversely, the R110 separated the i-CLL epitype into 2 major groups showing significantly different epiCMIT (**Fig. 1C**).



**Fig. 1 | The DNA methylome of CLL immunogenetics.** A, B T-SNE analysis highlighting CLL epitypes, the CLL proliferative history measured by the epiCMIT mitotic clock (A) as well as CLL stereotypes and R110 mutation (B).

### 3. MULTI-OMIC CHARACTERIZATION OF CLL IMMUNO-(EPI)GENOMIC CLUSTERS

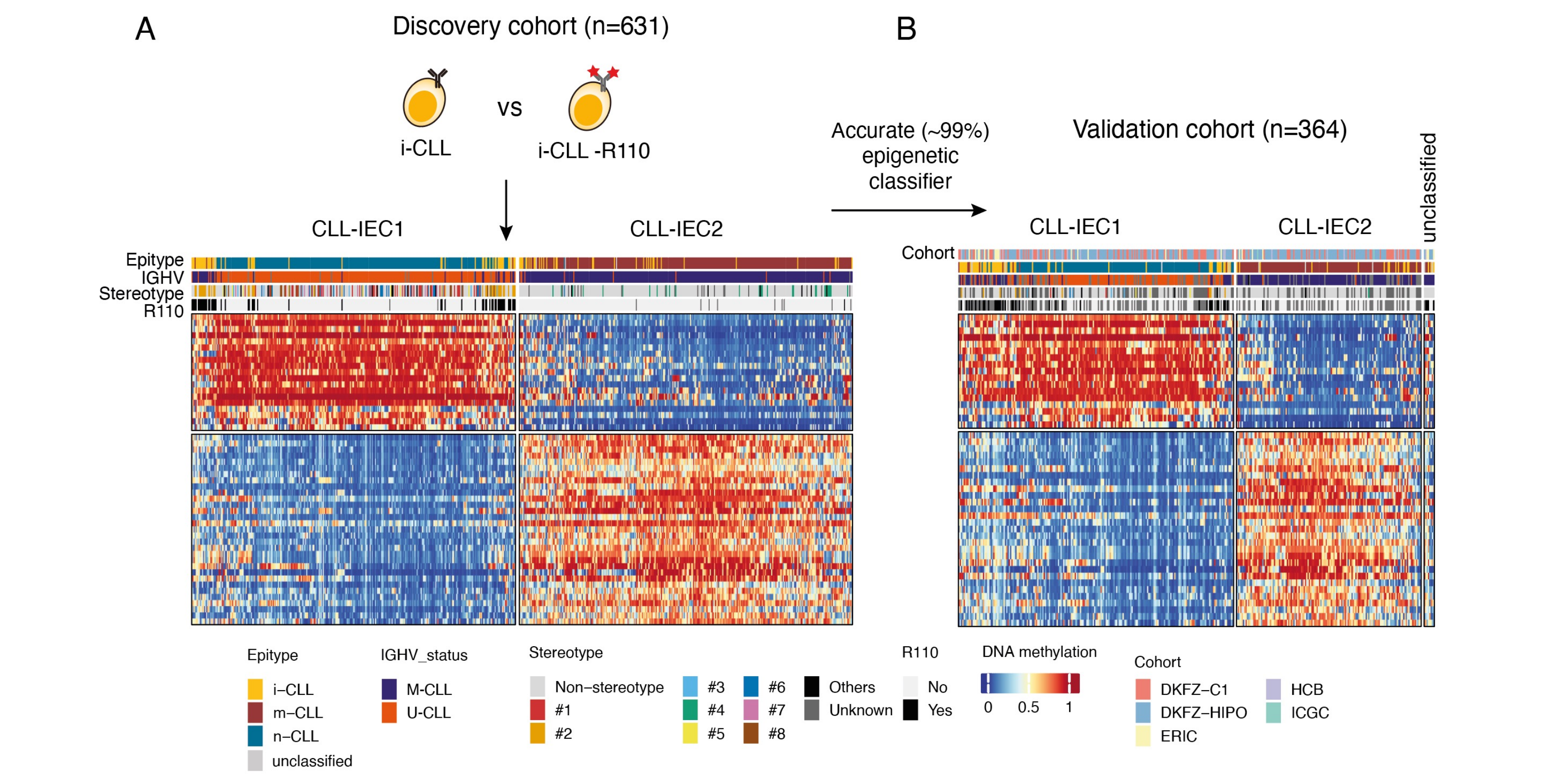
Multi-omic analysis revealed ~20% more differential features between clusters than IGHV groups, with discordant cases aligning better with clusters (**Fig. 3A**). DNA hypomethylation, chromatin accessibility, and activation (**Figs. 3B-C**) affected both known and novel CLL genes (MSI2, MYC, TFAP2A, GF1) linked to pathways such as NF-κB or WNT. RNA-seq supported associated transcription factor regulomes (**Fig. 3D**).



**Fig. 3 | Multi-omic analyses.** A, Differential features across omics layers. B, Differential DNA methylation and chromatin accessibility (C) in clusters. D, Transcription factor regulome analysis in clusters using RNA-seq data.

### 2. THE R110 METHYLATION SIGNATURE DICHOTOMIZE CLL PATIENTS INTO TWO IMMUNO-(EPI)GENOMIC CLUSTERS

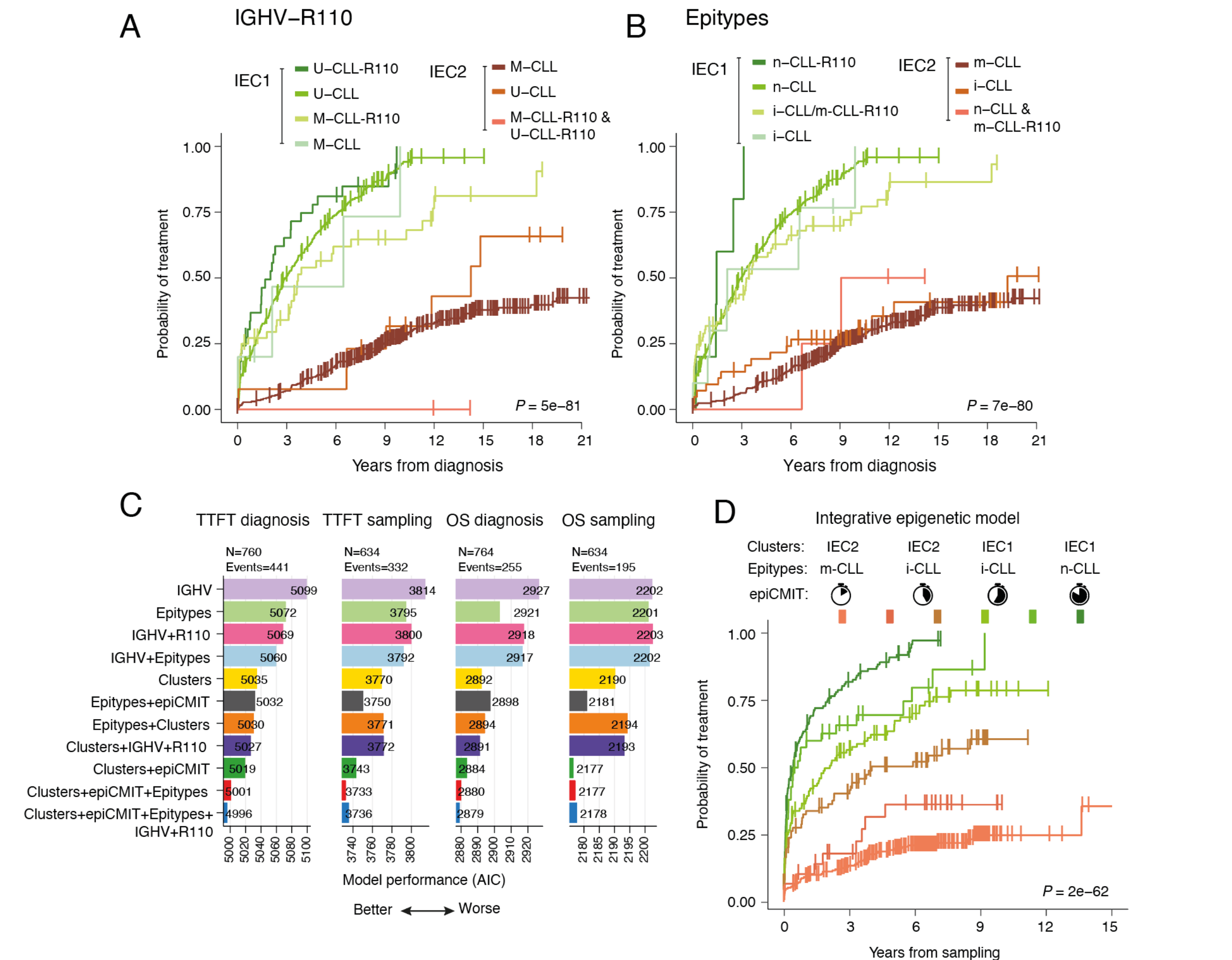
CLL i-CLL and i-CLL-R110 cases showed 316 differentially methylated CpGs, dividing CLL into two immuno-(epi)genomic clusters (**Fig. 2A, B**). IEC1 mainly included U-CLL/n-CLLs and R110-mutated cases, while IEC2 comprised mostly M-CLL and i-CLL/M-CLL without R110 mutations. Remarkably, up to 3% of cases were discordantly classified by IGHV and R110 status but aligned more closely with their cluster assignment in molecular and clinical features.



**Fig. 2 | Identification and validation of CLL immuno-(epi)genomics clusters.** A, Top 50 of 316 CpGs differentiating i-CLL vs i-CLL-R110 used to build an epigenetic classifier. B, Classifier applied to validation cohort.

### 4. CLINICAL IMPLICATIONS OF CLL IMMUNO-(EPI)GENOMIC CLUSTERS

IEC1 showed shorter TTFT and OS regardless of their immunogenetic or epitype assignment (**Fig. 4A, B**). Akaike information criterion suggested that an integrative epigenomic model including CLL epitypes, clusters and epiCMIT significantly improves the prognostic model performance of IGHV and R110 status (**Fig 4C,D**).



**Fig. 5 | Clinical associations of CLL clusters.** A, Kaplan-Meier curves of clusters for TTFT considering IGHV and RR10 (A) or epitypes and R110 (B). C, Akaike information criterion across distinct Cox models and endpoint variables. D, Kaplan-Meier curves of an integrative epigenomic model for TTFT.