Chronic lymphocytic leukemias with trisomy 12 show a distinct DNA methylation profile linked to altered chromatin activation

Chronic lymphocytic leukemia (CLL) is a neoplasm derived from mature B cells showing a broad spectrum of clinico-biological features.[1] The landscape of genetic alterations of CLL is well characterized[2] and found to be extremely heterogeneous, with multiple chromosomal aberrations and dozens of driver genes mutated in relatively small proportions of the cases.[3,4] In spite of this heterogeneity, four cytogenetic alterations, i.e., del(13q) (>50% of the patients), del(11q) (18%), +12 (16%), and less frequently del(17p) (7%), are collectively detected in at least 80% of patients.[1] These copy number changes are part of the routine risk assessment of CLL, as they are robustly associated with treatment choices and the clinical course of the patients. At one end of the prognostic spectrum, the isolated del(13q) is related to favorable prognosis, +12 with intermediate prognosis, del(11q) with poor prognosis and del(17p) with the worst prognosis of all groups. This latter subgroup identifies patients with particular resistance to chemoimmunotherapy who, instead, benefit considerably from biological agents.[1]

Although the targets of del(11q), del(13q) and del(17p) have been narrowed down to *ATM*/mir-34b-c, the *DLEU2*–mir-15-16 cluster, and *TP53*, respectively, how +12 contributes to CLL pathogenesis remains elusive.[1,5] This group of CLL is characterized by high rates of cell proliferation as well as clinical and biological heterogeneity, since it is linked to additional genomic aberrations such us trisomy 19, high frequency of *NOTCH1* mutations[6-8] and enrichment in the aggressive stereotyped subset 8[9] and IgG-switched heavy chains.[10] Moreover, +12 has recently been identified as an important modulator of response to kinase inhibitors in CLL.[11]

In addition to genetic changes, various layers of the CLL epigenome have also been characterized and these analyses have revealed novel insights into the cellular origin and molecular mechanisms underlying disease pathophysiology.[12-14] However, the relationship between the genetic and epigenetic changes in CLL is still unclear. Therefore, the goal of the present study is to identify the epigenetic correlates of the three most frequent cytogenetic subgroups of CLL and to shed light onto the molecular pathogenesis of +12 cases.

We used 450k arrays to analyze the DNA methylomes from 255 CLL cases in whom del(13q), del(11q) and +12 were mutually exclusive. In particular, we investigated 29 CLL cases with del(11q) (3 IGHV-mutated CLL [M-CLL] and 26 IGHV-unmutated CLL [U-CLL]), 45 with +12 (17 M-CLL and 28 U-CLL) and 181 with del(13q) (131 M-CLL and 50 U-CLL), as well as five biological replicates each of normal naive (NBC) and memory (MBC) B cells sorted from healthy donors (*Online Supplementary Table S1*).

We initially applied a series of unsupervised methods to characterize the DNA methylome of the cases. We found that the mean methylation levels of the three CLL subgroups was lower than in normal B cells, and that +12 cases globally had a significantly hypomethylated genome compared to the other CLL cases (+12 *vs.* del(11q), *P*=0.006; +12 *vs.* del(13q), *P*=7.2e-08; del(13q) *vs.* del(11q), *P*=0.01) (Figure 1A). Next, we analyzed the CpG with most variable methylation levels (Standard deviation [SD]≥0.1, n=185,936) by Principal Component Analysis (PCA) and studied the information provided by each of the five principal components (*Online Supplementary Figure S1*). As previously reported,[13] the first two components revealed different fractions of the variability to reflect different concepts: (i) overall CLL are epigenetically more similar to memory B cells (PC1, 7.9%); (ii) U-CLL and M-CLL are epigenetically different (PC1 and PC2); and, (iii) CLL as a whole are different from naive and memory B cells (PC2, 3.7%). The third and fourth components (1.9% and 1.3%) were less clear and showed subtle differences between naive and memory B cells as well as between U-CLL and M-CLL. Remarkably, the fifth component (1%) was unrelated to the *IGHV* gene somatic hypermutation status and revealed clear differences between +12 cases and CLL cases from the other cytogenetic subgroups (Figure 1B and *Online Supplementary Figure S1*), suggesting that +12 cases may indeed show a specific epigenetic configuration.

We then aimed at identifying the +12-specific CpG sites by two complementary approaches. First, we correlated the eigenvalues of the PC5 and the methylation β-values and found that 1,760 CpG were significantly associated with this source of variability and confirmed to be related with +12 by hierarchical clustering (*Online Supplementary Figure S2*). In order to more specifically detect +12-related CpG, we performed a supervised differential methylation (DM) analysis in cases with and without +12 considering U-CLL and M-CLL separately (absolute mean β-value difference of at least 0.25 and a false discovery rate [FDR]<0.05). In +12 U-CLL (n=28), we observed a signature of 646 DMCpG, which was mostly composed of hypomethylated sites in the +12 cases (80.1%) (Figure 1C-D and *Online Supplementary Table S2*). This hypomethylation signature was present in +12 CLL lacking and showing *NOTCH1* mutations, and therefore, was specifically associated with +12. The +12 M-CLL (n=17) were more heterogeneous and a lower number of DMCpG were detected (Figure 1C). However, the 646 DMCpG in +12 U-CLL showed the same trend in M-CLL cases (*Online Supplementary Figure S3*). Furthermore, analyzing B cells spanning the entire B-ell maturation program[15] revealed that the hypomethylation signature was mostly acquired *de novo* in the +12 cases, as for the great majority of the CpG it was not present in any B-cell subpopulation (Figure 1D and *Online Supplementary Figure S3*). Based on these findings, downstream analyses in the +12 cases were focused on the U-CLL subgroup. We also performed a similar analysis for del(13q) and del(11q) cases, which showed few DMCpG (Figure 1C). The predominantly *de novo* acquired hypomethylation signature in +12 U-CLL was related to gene bodies and untranslated (UTR) regions outside CpG islands, and targeted both transcribed regions and enhancer elements (Figure 1E).

The link between +12-specific hypomethylation and enhancer elements prompted us to evaluate the association between DNA methylation and histone 3 lysine 27 acetylation (H3K27ac), a chromatin mark related to active regulatory elements. In a previous study,[12] we generated H3K27ac ChIP-seq data in CLL patients, which included 21 of the 255 cases whose DNA methylome is studied in the present report. We initially observed that 48.3% of the DMCpG (312 of 646) specific for +12 U-CLL were located within 246 H3K27ac peaks (*Online Supplementary Table S3*). Remarkably, studying the H3K27ac signal within the overlapping peaks, we noticed that the regions losing methylation showed a trend towards higher H3K27ac levels, i.e., more activation, and those regions gaining methylation showed

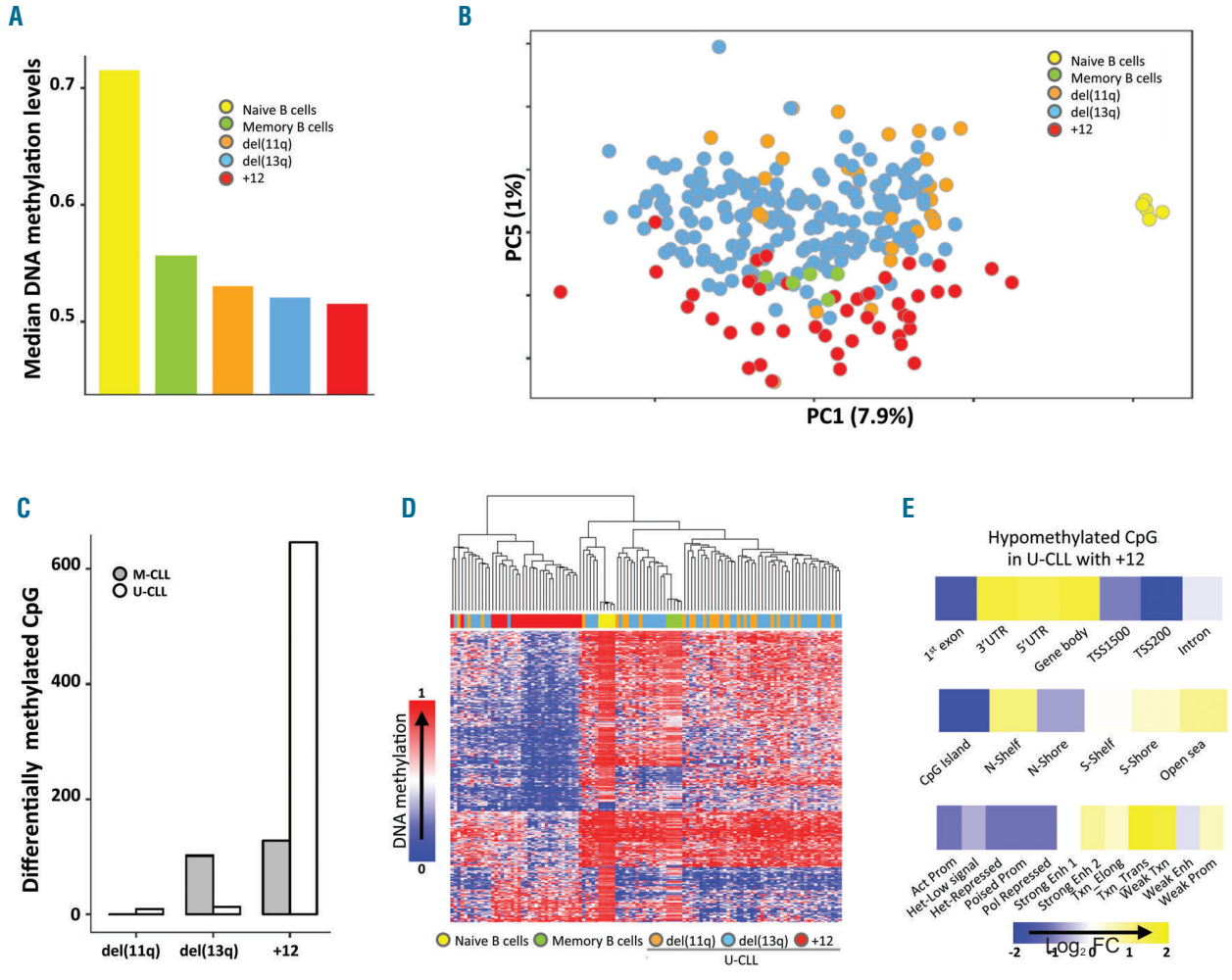

Figure 1. DNA methylation analysis of del(13q), del(11q) and +12 chronic lymphocytic leukemia. (A) Barplot showing the overall DNA methylation levels in the examined cytogenetic subgroups and normal B cells (+12 *vs.* del(11q), *P*=0.006; +12 *vs.* del(13q), *P*=7.166e-08; del(13q) vs. del(11q), *P*= 0.01). (B) Principal component analysis showing components 1 and 5 in the cytogenetic subgroups and normal B cells. (C) Number of differential methylation CpG (DMCpG) resulting from the comparison of each cytogenetic subgroup *versus* all the others chronic lymphocytic leukemia (CLL) cases. The white bar represents the IGHV-unmutated CLL (U-CLL) while the grey bar represents the IGHV-mutated CLL (M-CLL) cases. (D) Hierarchical clustering based on 646 DMCpG of +12 U-CLL cases as compared to other U-CLL and normal B cells. (E) Enrichment analysis of DMCpG hypomethylated in +12 U-CLL on gene location, location related to islands and chromatin states of memory B cells. The density represents the log_2FC of the enriched CpG compared to the background in each condition (hypo- and hyper-methylation). Each color is linked to a particular subgroup: orange for del(11q), blue for del(13q), red for +12, green for memory B cells and yellow for naïve B cells. PC: principal component, FC: fold change.

lower H3K27ac levels and, therefore, less active chromatin (Figure 2A-B). A differential analysis of the 246 H3K27ac peaks in U-CLL with and without +12 (adjusted *P*<0.05) revealed 35 regions with significantly increased H3K27ac levels (Figure 2C). Although a previous report indicated that the overall H3K27ac pattern of +12 cases was similar to normal B cells,[12] the increased statistical power of our targeted analysis of *de novo* hypomethylated sites focused on +12 U-CLL revealed the presence of chromatin activation. As transcription factor (TF) binding has been described to be related to DNA methylation and chromatin activity,[12,16] we postulated that TF could be implicated in inducing the +12-specific epigenomic signature. As TF bind to DNA lacking nucleosomes, we used previously reported chromatin accessibility data generated by ATAC-seq (available data for 20 of 255 examined cases)[12] (Figure 2D). We found that the 35 regions of interest contained 52 sites of accessible chromatin where TF potentially bind (*Online Supplementary Table S4*). The DNA sequences within these 52 accessible sites were significantly enriched (*P*<0.05) in binding sites of the NFIB-C, MYCN, TFCP2 and XBP1 TF (Figure 2E). Interestingly, the *TFCP2* gene is located in chromosomal band 12q13 and show moderate but significant overexpression cases with +12 compared to CLL lacking this genetic change (log_2FC= 0.284, FDR=0.024) (*Online Supplementary Figure S4*). Although detailed functional studies are needed to establish the role of TFCP2 in CLL, our data suggest that its overexpression may be one of the potential mechanisms through which +12 exerts its pathogenic effect in CLL.

In order to identify the potential target genes of the regulatory elements targeted by hypomethylation in +12 cases, we analyzed the gene expression arrays of U-CLL with del(11q), del(13q) or +12. As regulatory elements can affect the expression of distant genes within topologically associating domains (TAD),[12] the +12-specific hypomethylated regions associated with increased acetylation (n=35) were integrated with the TAD from GM12878, a lymphoblastoid B-cell line widely used to

characterize epigenetic patterns in CLL.[12-14] Using gene expression arrays from an extended series of 20 U-CLL with +12 and 54 U-CLL lacking this genetic change, we identified 25 genes with increased expression in +12 compared to del(11q) and del(13q) U-CLLs (Figure 2F and *Online Supplementary Table S5*). Amongst them, *RUNX3* was a remarkable example. This gene is a master regulator of gene expression during development that has been reported to act as tumor suppressor or oncogene in cancer.[17] We identified that +12 U-CLL show a cluster of seven hypomethylated CpG plus increased H3K27ac levels in a regulatory region 35 Kb upstream the *RUNX3*

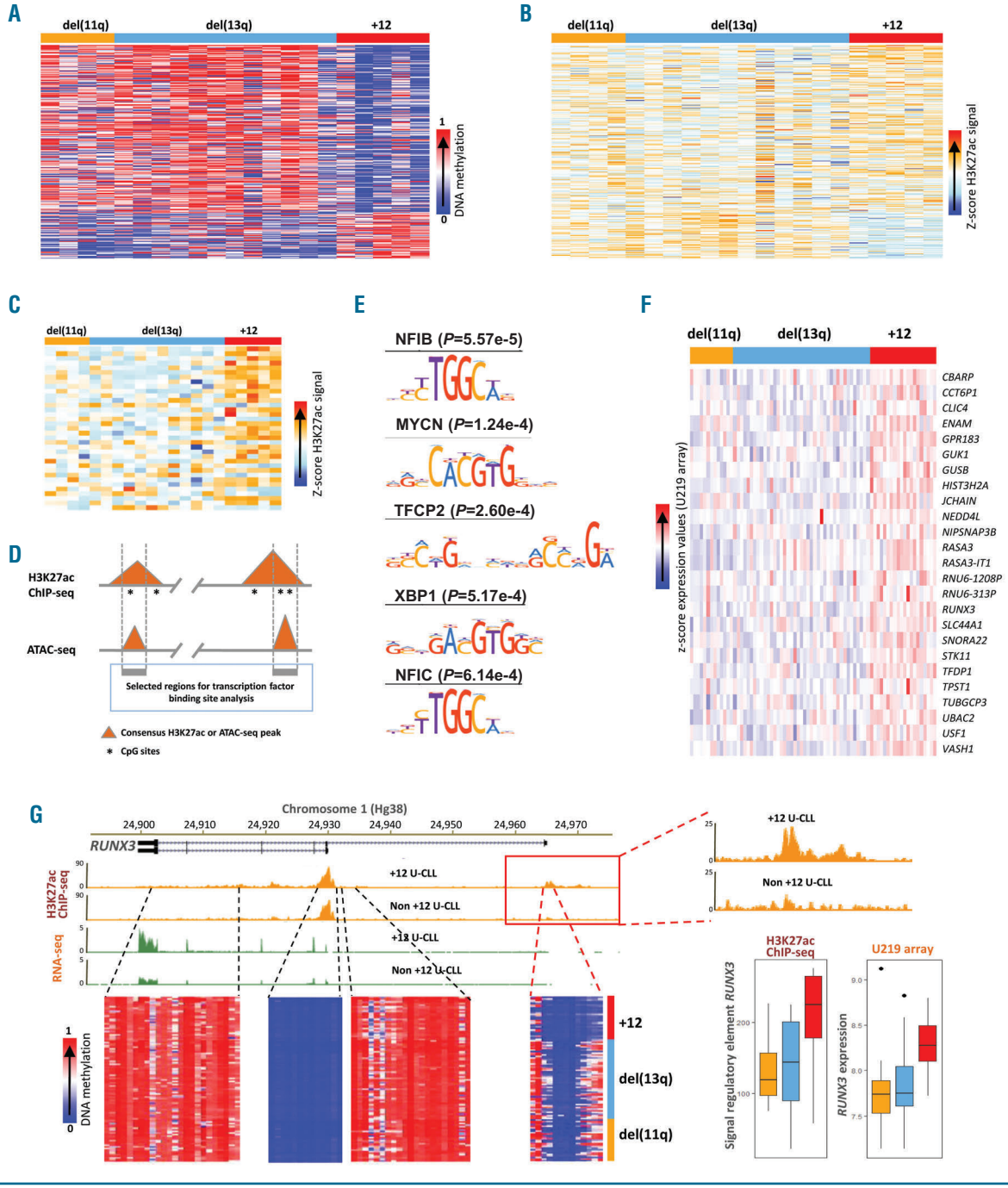

Figure 2. Characterization of the chromatin and transcriptional features of regions *de novo* hypomethylated in +12 cases. (A) DNA methylation levels of 312 CpG that overlap with H3K27ac-containing regions. (B) Heatmap of signal intensities of H3K27ac regions associated with CpG from panel (A). (C) Heatmap of the 35 hypomethylated regions showing significantly higher H3K27ac levels in IGHV-unmutated CLL (U-CLL) compared with cases without +12. (D) Scheme of the strategy used to detect accessible sites within regions with differential methylation CpG (DMCpG) and increased chromatin activity. (E) Transcription factors whose binding sites are significantly enriched in accessible regions within hypomethylated and active regulatory elements. (F) Heatmap showing the relative expression levels of the 25 genes epigenetically-upregulated in +12 U-CLL compared with cases lacking +12. (G) On the left, a genome browser display of the *RUNX3* region including H3K27ac and gene expression of two representative +12 and non +12 U-CLL, as well as DNA methylation data from all U-CLL studied. The red square points to the distant regulatory region that becomes hypomethylated and active in +12 U-CLL. On the right, a zoom-in panel of the H3K27ac levels of the distant enhancer region as well as box plots of H3K27ac and gene expression in all U-CLL studied. Each color is linked to a particular subgroup: orange for del(11q), blue for del(13q), red for +12.

promoter, a gene that was found to be significantly overexpressed (Figure 2G). No other region across the *RUNX3* gene changed DNA methylation levels and the gene promoter showed similar H3K27ac levels in cases showing and lacking +12 (Figure 2G). These results suggest that epigenetic activation of dozens of genes, and in particular the activation of a distant *RUNX3* regulatory element leading to gene overexpression, may account for the distinct biological background of this CLL subtype.

Taken together, our findings further support the unique biological features of +12 CLL from the epigenetic perspective. This group of CLL is associated with a subset of epigenetically-upregulated genes that may account for its distinct biological background. These novel insights into +12 CLL may provide a biological rationale to identify specific therapies to treat this unique subtype of CLL.

Maria Tsagiopoulou,[1,2] Vicente Chapaprieta,[3] Martí Duran-Ferrer,[3] Theodoros Moysiadi,[1,4] Fotis Psomopoulos,[1,4] Panagoula Kollia,[2] Nikos Papakonstantinou,[1] Elias Campo,[3,5,6] Kiostas Stamatopoulos[1,4] and Jose I. Martin-Subero[3,5,6,7]

[1]Institute of Applied Biosciences, Center for Research and Technology Hellas, Thessaloniki, Greece; [2]Department of Biology, National and Kapodistrian University of Athens, Athens, Greece; [3]Institut d'Investigacions Biomèdiques August Pi i Sunyer (IDIBAPS), Universitat de Barcelona, Barcelona, Spain; [4]Department of Molecular Medicine and Surgery, Karolinska Institute, Stockholm, Sweden; [5]Departamento de Fundamentos Clínicos, Universitat de Barcelona, Barcelona, Spain; [6]Centro de Investigación Biomédica en Red de Cáncer (CIBERONC), Madrid, Spain and [7]Institució Catalana de Recerca I Estudis Avançats (ICREA), Barcelona, Spain.

Correspondence:
JOSE I. MARTIN-SUBERO - imartins@clinic.cat

doi:10.3324/haematol.2019.240721

Acknowledgements: this work was partially developed at the Centro Esther Koplowitz (CEK, Barcelona, Spain).

Funding: data used in the present study have been generated through funding from the European Union's Seventh Framework Programme through the Blueprint Consortium (grant agreement 282510), the International Cancer Genome Consortium (Chronic Lymphocytic Leukemia Genome consortium to EC), the World Wide Cancer Research Foundation Grant No. 16–1285 (to JIM-S), the Generalitat de Catalunya Suport Grups de Recerca AGAUR 2017-SGR-736 (to JIM-S) and 2017-SGR-1142 (to EC), and CIBERONC (CB16/12/00225). EC is an Academia Researcher of the Institució Catalana de Recerca I Estudis Avançats (ICREA) of the Generalitat de Catalunya. MT is a recipient of a fellowship from the State Scholarships Foundation of Greece, and a short term collaboration award winner by the European Hematology Association.

References

1. Kipps TJ, Stevenson FK, Wu CJ, et al. Chronic lymphocytic leukaemia. Nat Rev Dis Primers. 2017;3:16096.
2. Bosch F, Dalla-Favera R. Chronic lymphocytic leukaemia: from genetics to treatment. Nat Rev Clin Oncol. 2019;16(11):684-701.
3. Landau DA, Tausch E, Taylor-Weiner AN, et al. Mutations driving CLL and their evolution in progression and relapse. Nature. 2015;526(7574):525-530.
4. Puente XS, Bea S, Valdes-Mas R, et al. Non-coding recurrent mutations in chronic lymphocytic leukaemia. Nature. 2015; 526(7574):519-524.
5. Deneberg S, Kanduri M, Ali D, et al. microRNA-34b/c on chromosome 11q23 is aberrantly methylated in chronic lymphocytic leukemia. Epigenetics. 2014;9(6):910-917.
6. Baliakas P, Iskas M, Gardiner A, et al. Chromosomal translocations and karyotype complexity in chronic lymphocytic leukemia: a systematic reappraisal of classic cytogenetic data. Am J Hematol. 2014; 89(3):249-255.
7. Del Giudice I, Rossi D, Chiaretti S, et al. NOTCH1 mutations in +12 chronic lymphocytic leukemia (CLL) confer an unfavorable prognosis, induce a distinctive transcriptional profiling and refine the intermediate prognosis of +12 CLL. Haematologica. 2012;97(3):437-441.
8. Ibbotson R, Athanasiadou A, Sutton LA, et al. Coexistence of trisomies of chromosomes 12 and 19 in chronic lymphocytic leukemia occurs exclusively in the rare IgG-positive variant. Leukemia. 2012; 26(1):170-172.
9. Baliakas P, Hadzidimitriou A, Sutton LA, et al. Clinical effect of stereotyped B-cell receptor immunoglobulins in chronic lymphocytic leukaemia: a retrospective multicentre study. Lancet Haematol. 2014; 1(2):e74-84.
10. Vardi A, Agathangelidis A, Sutton LA, et al. IgG-switched CLL has a distinct immunogenetic signature from the common MD variant: ontogenetic implications. Clin Cnacer Res. 2014;20(2):323-330.
11. Dietrich S, Oles M, Lu J, et al. Drug-perturbation-based stratification of blood cancer. J Clin Invest. 2018;128(1):427-445.
12. Beekman R, Chapaprieta V, Russinol N, et al. The reference epigenome and regulatory chromatin landscape of chronic lymphocytic leukemia. Nat Med. 2018;24(6):868-880.
13. 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.
14. 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.
15. 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.
16. Stadler MB, Murr R, Burger L, et al. DNA-binding factors shape the mouse methylome at distal regulatory regions. Nature. 2011;480(7378):490-495.
17. Selvarajan V, Osato M, Nah GSS, et al. RUNX3 is oncogenic in natural killer/T-cell lymphoma and is transcriptionally regulated by MYC. Leukemia. 2017;31(10):2219-2227.