Targeting Heterochromatin Eliminates Malignant Hematopoietic Stem and Progenitor Cells in Chronic Myelomonocytic Leukemia Through Reactivation of Retroelements and Innate Immune pathways

Donia Hidaoui[1,2,3,], Rabie Chelbie[1,2,3,4], Mathieu Bohm[1,2,3,4], Audrey Porquet[1,2,3], Aikaterini Polyzou[5], Vincent Alcazer[6], Stéphane Depil[6], Aygun Imanci[1,2,3,4], Margot Morabito[1,2,3,], Aline Renneville[1,2,3,7], Dorothée selimoglu-Buet[1,2,3], Sylvain Thépot[8], Raphael Itzykson[9], Lucie Laplane[1,2,3,10], Nathalie Droin[1,2,3,7], Eirini Trompouki[5], Emilie Elvira-Matelot[1,2,3], Eric Solary[1,2,3,11] and Françoise Porteu[1,2,3,*]

1. *Inserm UMR1287, Gustave Roussy Cancer Center, 94805, Villejuif, France*.
2. *Université Paris-Saclay,*
3. *Equipe labellisée Ligue Nationale Contre le Cancer*
4. *Inovarion, 75005 Paris, France.*
5. *Institute for Research on Cancer and Aging of Nice (IRCAN), Université Côte d'Azur, INSERM, CNRS, Nice, France*.
6. *Cancer Research Center of Lyon, UMR INSERM U1052 CNRS 5286, Centre Léon Bérard, Lyon, France*
7. *INSERM US23, CNRS UMS 3655, Gustave Roussy Cancer Center, Villejuif, France.*
8. *Clinical Hematology Department, University Hospital, Angers, France.*
9. *Hematology Department, Saint-Louis Hospital, Paris, France.*
10. *Institut d'Histoire et Philosophie des Sciences et des Techniques, Université Paris I Panthéon-Sorbonne, Paris, France.*
11. *Clinical Hematology Department, Gustave Roussy Cancer Center, Villejuif, France.*

** to whom correspondence should be sent*

Françoise Porteu (francoise.porteu@gustaveroussy.fr)

INSERM U1287

Cancer Campus Gustave Roussy, PR1

114 rue Édouard-Vaillant

94805 Villejuif Cedex-France

Tel: (33) 1 42 11 23 48

Running title: Eradication of CMML HSPCs through modulation of retroelements expression.

The authors declare no potential conflicts of interest

Complementary list of authors:

Thorsten Braun[1], Céline Berthon[2], Gabriel Etienne[3], Pierre Fenaux[4], Christophe Willekens[5], also provided patient samples.

1. *Clinical Hematology Department, Hôpital Avicenne, Bobigny, France*
2. *Clinical Hematology Department, University Hospital, Lille, France.*
3. *Clinical Hematology Department, University Hospital, Bordeaux, France*
4. *Hematology Department, Saint-Louis Hospital, Paris, France*
5. *Clinical Hematology Department, Gustave Roussy Cancer Center, Villejuif, France.*

Abstract

Chronic myelomonocytic leukemia (CMML) is a severe myeloid malignancy affecting the elderly, for which therapeutic options are limited. DNA hypomethylating agents (HMAs) provide transient responses, failing to eradicate the malignant clone. Hematopoietic stem cell (HSC) aging involves a reorganization of heterochromatin, evidenced by alterations in histone marks H3K9me2 and H3K9me3. These repressive marks together with DNA methylation are essential for maintaining genome stability by suppressing transposable elements (TEs). In solid cancers, the antitumor efficacy of HMAs involves the derepression of TEs, mimicking a state of viral infection. In this study, we demonstrate a significant disorganization of heterochromatin in CMML HSCs and progenitors (HSPCs) characterized by an increase in the repressive mark H3K9me2, mainly at the level of TEs, and a repression of immune and age-associated transcripts. Combining HMAs with G9A/GLP H3K9me2 methyltransferase inhibitors reactivates these pathways, selectively targeting mutated cells while preserving wild-type HSCs, thus offering new therapeutic avenues for this severe myeloid malignancy.

Statement of significance

This study reveals significant heterochromatin disorganization in CMML HSPCs, marked by increased repressive H3K9me2 at transposable elements and repression of immune- and aging-related transcripts. Combining DNA hypomethylating agents with H3K9me2 inhibitors selectively targets mutated cells while preserving wild-type cells, offering promising novel therapeutic avenues for this severe myeloid malignancy linked to aging.

Introduction

Chronic myelomonocytic leukemia (CMML) is a myeloid neoplasm originating from a hematopoietic stem cell (HSC). This disease usually emerges in the elderly (median age at diagnosis is 72 years) and combines dysplastic and proliferative features, therefore is classified as a myelodysplastic/myeloproliferative neoplasm. The World Health Organization distinguishes a dysplastic from a proliferative subtype, as proliferative forms (based on white blood cell count with a cut-off of 13 x 10^9/L) have an inferior outcome. The outcome also correlates with blast cell count in bone marrow (BM) and peripheral blood (PB), separating CMML-1 from CMML-2 (BM and PB blast cells 10-20% and > 5%, respectively) (1). Somatic mutations detected in CMML cells typically and diversely affect genes encoding epigenetic, splicing and signaling molecules. These mutations accumulate almost linearly in the stem cell compartment, with an early dominance of the clone that invades most of this compartment and a growth advantage to the most mutated cells with differentiation (2). Such a relatively simple genotype contrasts with the heterogeneous phenotype of the disease. Consistently, abnormal DNA methylation profiles were identified as an additional feature of the disease (3-5).

CMML is a severe neoplasm with a median survival lower than 3 years in most series. The sole curative therapeutic option is allogeneic stem cell transplantation, which is difficult to implement due to age and comorbidities. Cytoreductive drugs and DNA hypomethylating agents (HMAs) are used in patients with predominantly proliferative and dysplastic CMML variants, respectivey (6). The therapeutic response to these drugs is always transient. The response to HMAs correlates with demethylation of leukemic cell DNA with a restored hematopoiesis that contrasts with the lack of decrease in the fraction of mutated cells (4).

Thus, HMA-induced correction of hematopoietic imbalance is likely attributable to epigenetic rather than cytotoxic effects.

In accordance with the age at diagnosis, two clock-like molecular signatures were identified in CMML cells (4). Cellular aging also involves heterochromatin reorganization (7,8), with alterations in DNA methylation and histone H3K9me2 and H3K9me3 mark patterns. Reduced H3K9me3 in HSCs is observed in aging or premature aging models, correlating with characteristic hematopoietic changes related to aging (9-11). DNA methylation and repressive histone marks are crucial regulators of genome stability by repressing transposable elements (TEs), including DNA transposons and retrotransposable elements (RTEs). RTEs, which are classified into long terminal repeat (LTR) that characterize endogenous retroviruses (ERVs), and non-LTR elements such as long or short interspersed elements (LINE-1/L1; SINE), utilize a copy/paste mechanism to spread in the genome, causing double-strand breaks, and contribute to gene regulatory networks with a potential oncogenic role (12). TEs also serve as a source of endogenous double-stranded (ds) RNA or cytoplasmic cDNA that are sensed as genomic parasites, ultimately triggering a type I interferon (IFN-I) response (13). In HSCs, TE overexpression induced by irradiation and chemotherapy leads to DNA damage accumulation, and creates an intrinsic sterile inflammatory context potentially contributing to functional changes that generate an aging phenotype (14,15).

HMA activity in solid tumors was ascribed to their capacity to derepress TEs that generate a type I interferon (IFN-I) and NFκB response and promote the expression tumor-associated antigens (16,17). The role of these pathways in hematological malignancies treated with HMAs is less known. In the present study, we show a significant chromatin disorganization in HSPCs of CMML patients, which repress TE expression and inflammatory signaling. We also provide

evidence that a combination of HMAs and H3K9me2 methyltranferase inhibitors could reactivate these pathways, specifically targeting the mutated clone while preserving unmutated HSCs.

Results

CMML HSPCs show a decrease in inflammatory response pathways

To explore the characteristics of pathological aged hematopoiesis in CMML patients versus healthy aging, we assembled a cohort of untreated CMML patients, at the time of diagnosis, and age-matched healthy donors (herein referred to as controls) (supplementary Table S1). Bulk RNA-sequencing (RNA-seq) analysis was performed on $CD34^+$ hematopoietic stem and progenitor cells (HSPCs) isolated from the BM of 19 patients (mean age 70.6, range 49-84) and 12 age-matched healthy donors (mean age 68, range 54-91). Principal component analysis (PCA) revealed a distinct separation of control and patient transcriptomes (Fig. 1A). Of 2,664 differentially expressed genes (DEGs) (padj<0.1), 1,430 (54%) were upregulated (Fig. 1B; supplementary Table S2). Gene set enrichment analysis (GSEA) examining HALLMARK and KEGG gene sets indicated that control cell transcriptomes were significantly enriched in various inflammatory, innate immune response, and cytokine signaling, as well as cell cycle, P53, and apoptosis pathways (Fig. 1C; supplementary S1A). Gene Ontology process analysis unveiled that, beyond immune response, terms associated with chromatin remodeling and DNA packaging were notably enriched in controls (Fig. 1D; supplementary S1B). In contrast, patient cells exhibited enrichment in only a few terms related to ribosomes, vascular system development, actin, and adhesion (Fig. 1D and supplementary S1A).

Increased expression of inflammatory genes, P53 signaling and chromatin deregulation are indicative of senescence and aging (8,18-22). The transcriptome of patient HSPCs therefore suggests a younger state compared to control HSPCs. Supporting this possibility, senescence and aging signatures were more pronounced in control cells (Fig. 1E and F; supplementary S1C). Furthermore, a signature composed of genes upregulated in aged versus young human

BM HSCs (22) was depleted in patient cells, while genes downregulated with age were marginally enriched (Fig. 1G).

Chromatin is remodeled in CMML HSPCs

We have previously shown that murine and human HSC aging is associated with decreased expression of the H3K9-trimethyl transferase SUV39H1, leading to a loss of the heterochromatin mark H3K9me3 (9), which was also detected in HSC premature aging induced by sublethal irradiation (11). Decreased expression of the heterochromatin mark H3K9me2, assessed by immunofluorescence, was also detected in HSCs from aged mice and in $CD34^+$ BM cells collected from aged individuals compared with umbilical cord blood (supplementary Fig. S2A and B). These results prompted us to analyze H3K9me3 and H3K9me2 heterochromatin marks in HSPCs from patients and controls. Surprisingly, immunofluorescence analysis suggested a significant decrease in H3K9me3 expression in patient cells (Fig. 2A), which was confirmed by immunoblot analysis in five patients (supplementary Fig. S2C). Conversely, a significant increase in the expression of H3K9me2 was detected by immunofluorescence in $CD34^+$ cells from 28 out of 29 patients (Fig. 2B) and validated by immunoblot analysis in five patients (supplementary Fig. S2D). This correlated with heightened expression of the H3K9-dimethyltransferases G9A and GLP, particularly at the protein level (Fig. 2C and D; supplementary S2E and F).

To explore the impact of these opposite modifications on two heterochromatin marks on chromatin opening genome-wide, we conducted Assay for Transposase-Accessible Chromatin-sequencing (ATAC-seq) experiments in $CD34^+$ cells collected from 4 patients and 3 controls. We detected 40,335 differential peaks, with 31,110 (77%) gained (unique to CMML) and 9,245 (23%) lost (unique to controls) (Fig. 2E; supplementary Table S3). Altered peaks

were significantly enriched in intergenic and intronic regions, while common peaks were predominantly located at the transcription start sites (supplementary Fig. S2G). Assigning differential ATAC-seq peaks to genes (15) demonstrated a substantial overlap between DEGs and ATAC-seq peaks (Fig. 2F). Consistent with the lower expression of TNF-NFκB and IFN signaling genes in CMML cells, numerous processes related to immune/inflammation/NFκB signaling were associated with genes exhibiting loss of chromatin accessibility (Fig. 2G; supplementary S2H). Further, an analysis of transcription factor (TF) motifs in peaks with reduced chromatin accessibility using HOMER demonstrated specific enrichment of NFκB motifs, along with increased enrichment of IFN regulatory factors (IRF1, IRF3) and IFN-stimulated response element (ISRE) motifs (Fig. 2H). This suggests that the attenuated expression of certain immune/inflammatory response genes in patient cells might be related to an altered chromatin accessibility.

Together, these results highlight a profound chromatin remodeling that may contribute to repressing inflammatory genes in CMML compared to healthy control $CD34^+$ cells.

H3K9me2 is enriched at TEs in CMML HSPCs

H3K9me2 is recognized as a suppressor of innate immune genes, notably exerting direct control over the expression of *IFNB1* gene, encoding IFN-β, and certain IFN-stimulated genes (ISGs) (23,24). To ascertain whether increased H3K9me2 in CMML cells could contribute to the observed decrease in inflammatory/immune gene expression, we conducted H3K9me2 CUT&Tag analysis in $CD34^+$ cells, comparing 4 patient with 3 control samples. As expected from immunofluorescence and western blot experiments, patient cells exhibited a greater number of recovered peaks (15,406 ± 1,484) compared to control cells (4,831 ± 1,600) (supplementary Fig. S2I). Focusing on peaks shared by all four patients or three controls, we

identified 1,396 gained peaks exclusive to patients and only 88 lost peaks unique to controls (Fig. 2I; supplementary Table S4). Gene annotation revealed that H3K9me2 was primarily enriched in intergenic regions, while being excluded from exons, in both patient and control cells (Fig. 2J). Biological processes associated with the nearest genes gaining H3K9me2 in patient cells included adhesion and endocytic recycling but did not show enrichment in immune pathway genes or IRF/NFκB transcription factor binding sites (not shown). Moreover, there was no significant difference between patients and controls in the expression of H3K9me2-linked genes (supplementary Fig. S2J), and no correlation was identified between gene expression and H3K9me2 concentration at the promoter level (supplementary Fig. S2K).

Since H3K9me2 is also known to repress repeated sequences (25,26), we analyzed CUT&Tag data by considering all mapping reads, either unique and multiple, that mapped to the genome without mismatch, and we randomly assigned them at one of their best possible positions in the genome (11). This analysis identified a pronounced enrichment of the H3K9me2 mark at TE relative to genes in patients, with 13,364 TEs among H3K9me2-bound peaks in CMML cells, compared to 255 in control cells (Fig. 2K; supplementary S2L). A permutation test evaluating the number of each TE class in single peaks from patients and controls versus random peaks revealed that H3K9me2 peaks gained in patient cells were significantly enriched for all TE classes (Fig. 2L; supplementary S2M). In contrast, ERVs and SINEs were under-represented in lost peaks. The distribution of H3K9me2 among the TE classes demonstrated that ERVs were overrepresented in gained peaks, while LINEs were overrepresented in lost peaks (Fig. 2M). Overall, these results highlight that the profound chromatin remodeling detected in CMML $CD34^+$ cells involves an enrichment of H3K9me2 at TEs.

The HMA and G9A/GLP inhibitor combination reduces CMML HSPC clonogenicity

HMAs are clinically approved for treating high-risk CMML patients. Recent data have shown that HMAs, when used in various pathological contexts, can act through the reactivation of TEs (16,17), ultimately triggering the activation of IFN-I and NF-κB pathways. Our initial results led us to speculate that deregulated G9A/GLP methyltransferases and histone marks could interfere with CMML cell response to HMAs. To test this hypothesis, CMML and healthy donor $CD34^+$ cells were treated *ex vivo* with the G9A/GLP inhibitor UNC0638, the HMA decitabine (DAC) and their combination, then seeded in methylcellulose for 14 days (Fig. 3A). Since H3K9me2 has been shown to bind to demethylated regions where it can then reinstall DNA methylation (26,27), we added UNC0638 2 days after DAC. Dose-response experiments demonstrated that each compound alone had a modest impact on colony formation by patient $CD34^+$ cells, with a 33% and 6% decrease observed at the highest dose of DAC (80 nM) and UNC0638 (2 µM), respectively. In contrast, the combination of DAC and UNC0638, even at low doses, significantly reduced the clonogenicity of CMML $CD34^+$ cells (Fig. 3B). Using the Chou-Tabalay method (28), a synergistic effect (combination index (CI) <1) was observed under all conditions (Fig. 3C and D). Notably, the highest dose of DAC also induced a 20% decrease in the number of colonies formed by control $CD34^+$ cells. However, consistent with the low expression of G9A/GLP in control cells, no synergy between DAC and UNC0638 on their clonogenicity was detected (supplementary Fig. S3A).

Subsequent experiments were performed with 10 nM DAC followed by 1 µM UNC0638. A significant reduction in global DNA methylation level was detected by ELISA after 2 days of DAC, which was maintained at day 4, either alone or combined with UNC0638, while UNC0638 alone had no effect on DNA methylation (supplementary Fig. S3B and C). The level of H3K9me2, measured by immunofluorescence, remained unchanged after 2 days of culture in

the presence of DAC alone (supplementary Fig. S3D), while it was markedly reduced by UNC0638 (supplementary Fig. S3E). These experiments validated the proper functioning of the tested drugs.

In the absence of treatment, the numbers of colonies formed by $CD34^+$ cells were comparable between control and patient cells (on average 82 +/-10, n=12 and 72 +/- 4, n=28, colonies per 600 cells cultured, respectively) (Fig. 3E). Treatment with DAC followed by UNC0638 inhibited the clonogenicity of 26 out of 28 tested CMML $CD34^+$ cells and reduced the colony size while each inhibitor alone had a weak (DAC) or no (UNC0638) effect (Fig. 3F; supplementary Fig. S3F). Again, none of these treatments affected the growth of control $CD34^+$ cells (Fig. 3G). Similar results were obtained with the combination of another HMA, 5-Azacytidine (AZA) combined with UNC0638, and with another G9A/GLP inhibitor, UNC0642, combined with DAC (supplementary Fig. S3G-I). Finally, CM-272, a dual inhibitor of DNMT1 and G9A/GLP (29), specifically reduced the clonogenicity of CMML but not control $CD34^+$ cells (Fig. 3H and I).

To more precisely identify the cell population targeted by the combination, $CD34^+$ cells were enriched in $CD34^+CD38^-CD90^+$ (HSC-enriched, referred to as stem cells) and $CD34^+CD38^-CD90^-$ (progenitor-enriched, referred to as progenitors) cells. We functionally assessed how the inhibitors affected the cell clonogenic potential by testing their capacity to serially replate in methylcellulose over 3 passages after treatment at day 2 and 4 of the culture (Fig. 3A). In the absence of treatment, at all three passages, the number of colonies formed by control and patient stem cells was not significantly different (supplementary Fig. S4A). DAC alone specifically decreased the size of the colonies generated by CMML progenitors and, starting from P2, significantly reduced their number, while having no effect on the stem cell-enriched population (Fig. 4A; supplementary S4B). The DAC/UNC0638 combination triggered a drastic

reduction in the number and size of colonies formed by both CMML stem cells and CMML progenitors. None of the treatments, either alone or in combination, affected the capacity of stem cells and progenitors from controls to form colonies (supplementary Fig. S4C). Similar results were obtained by measuring the cumulative clonogenic potential at P2 (30), which considers both the number of colonies and their size (Fig. 4B).

We further explored drug effects in single-cell liquid culture experiments for 14 days, measuring cell viability (number of wells with at least one cell) and proliferation (total number of cells in wells with at least one cell). In the absence of treatment, vialbility and expansion of both stem cells and progenitors were similar in control and patient samples (supplementary Fig. S4D and E). However, the colony size formed by CMML stem cells and progenitors was larger than those of controls (supplementary Fig. S4F). The combination reduced the viability and expansion of both stem cells and progenitors from patients (Fig. 4C-E). DAC treatment alone targeted patient progenitors only, UNC0638 alone had no effect, and none of the treatments impaired the viability and proliferation of control stem cells and progenitors (supplementary Fig. S4G-I).

We subsequently interrogated the ability of the drug combination to selectively eliminate leukemic cells by performing Sanger sequencing of pre-identified mutated genes (supplementary Table S5) in colonies grown from $CD34^+$ cells in methylcellulose and from $CD34^+CD38^-CD90^+$ cells cultured at one cell per well for 14 days, focusing on somatic mutations with a variant allele frequency (VAF) of 30% or more. Detecting mutations with the highest VAF likely captures the primary events and major clones, as previous studies demonstrated that mutations accumulate almost linearly in CMML stem and progenitor cells, with limited branching events and early clonal dominance (2). In line with previous findings (2), a minority

of wild-type cells (ranging from 0 to 40%) were found in untreated conditions (Fig. 4F). The lack of detection of wild-type cells in patient #3564 could be due to an overrepresentation of mutated clones in methylcellulose (2) and the limited number of clones analyzed. While DAC and UNC0638 showed limited effects when tested alone, the combination drastically reduced the proportion of mutated clones in all patients (Fig. 4F). In patients #3459 and #3562, with mutations in *TET2* and *KRAS*, respectively, the major clones specifically and completely disappeared. Patients #3480 (harboring clones with 0, 1, or 2 mutations) and #3564 (harboring clones with 1, 2, or 3 mutations), showed increased proportions of non-mutated cells, along with reduction in the most mutated clones carrying 2 or 3 mutations. UNC0638 alone slightly increased the proportion of clones or colonies without any mutation (patients #3459 and #3562) or harboring 1 vs. 2 mutations. DAC alone tended to favor the growth of clones harboring mutations (patients #3459 and #3562) or the most mutated clones (increased size of clones with 3 mutations from 59 to 78% in patient #3564). Altogether, these results show that DAC and UNC0638 demonstrate a synergistic and specific impact on leukemic cells while preserving wild-type cells.

The HMA and G9A/GLP inhibitor combination increases TE expression in CMML CD34$^+$ HSPCs

To assess the impact of DAC and UNC0638 on TE expression, we performed RNA-seq analysis on CMML CD34$^+$ cells (n=4 patients) treated *ex vivo* for four days with or without DAC and/or UNC0638, as outlined in Figure 4A. At day 4 of culture, the number of cells was not significantly different between treated and non-treated conditions (supplementary Fig. S5A). Multimapping analysis of RNA-seq data revealed 340 differentially expressed TE families ($p<0.05$), all exhibiting increased expression after treatment with the DAC/UNC0638 combination compared to untreated cells (Fig. 5A; supplementary Table S6). In contrast, only two differential TE families were identified after 4 days of treatment with either DAC or UNC

alone (supplementary Fig. S5B and C). Upon DAC/UNC0638 combination, ERVs and DNA transposons accounted for 43% and 32% of differentially expressed TEs, respectively (Fig. 5B). Notably, ERV retrotransposons showed significant enrichment compared to LINEs and DNA transposons (Fig. 5C). Comparative analysis of these deregulated TEs across all culture conditions revealed that DAC and UNC0638 alone elevated the expression of ERV and DNA transposons compared to untreated cells, with the combination synergistically increasing the expression of ERVs, DNA and LINEs (Fig. 5D; supplementary S5D).

Focusing on ERVs that displayed the most significant differential expression upon DAC/UNC0638 treatment, we conducted a loci-based quantification, as described in our previous work (31), incorporating four additional patient samples. We identified 2,120 differentially expressed ERV loci ($p<0.05$, FC 1.5) between untreated and DAC/UNC0638 conditions, including 2,056 (97%) upregulated and 64 (3%) downregulated loci (Fig. 5E and F; supplementary Table S7). Confirming the results above, the impact of DAC or UNC0638 treatment alone was notably weaker, resulting in 120 (63 upregulated and 57 downregulated) and 122 (59 upregulated and 63 downregulated) differentially regulated ERV loci, respectively, compared to non-treated condition (Fig. 5E and F). A majority of the ERVs induced by the combination (1991/2056, 97%) were specific to this condition, whereas either treatment alone upregulated very few specific ERVs (10% and 20% for UNC0638 and DAC, respectively). Conversely, differentially downregulated loci were largely specific to each condition (48/58, 83%, and 43/43, 100%, of the ERVs downregulated by UNC0638 or DAC alone). Focusing on ERVs differentially expressed in all conditions compared with untreated cells, the vast majority of loci exhibited a synergistically higher expression level in cells treated with the combination than in DAC or UNC0638 treatment alone (Fig. 5G). RT-qPCR experiments targeting ERV family members deregulated by the combination further validated their greater or synergistic

induction after 4 days of treatment with DAC/UNC0648 compared to either inhibitor alone (Fig. 5H). Together, these results indicate that the HMA and G9A/GLP inhibitor combination, but not each inhibitor alone, induces a significant and rapid derepression of TEs, particularly ERVs, in CMML CD34$^+$ cells.

The DAC/UNC0638 combination reactivates viral and inflammatory pathways in CMML cells

ERVs can generate double-stranded RNAs (dsRNAs) that are recognized by antiviral pattern sensors, leading to IFN-I production and activation of the IFN/NFκB pathway—a process known as viral mimicry, which exhibits an anti-tumoral effect (13,16,17). Analysis of coding gene expression revealed 3,227 DEGs ($p<0.05$, FC 1.2), of which 2,914 (90%) are positively and 313 (10%) are negatively expressed in DAC/UNC0638 treated compared with untreated CD34$^+$ cells (Fig. 6A and B; supplementary Table S8). Treatment with DAC or UNC0638 alone induced only 187 and 158 DEGs, respectively, with 104 (56%) and 84 (53%) being overexpressed, and 83 (44%) and 74 (47%) being downregulated. GSEA revealed that genes upregulated by the DAC/UNC0638 combination were enriched for IFN, NFκB, and other inflammatory pathways, TP53, and apoptosis. In contrast, downregulated genes included those associated with DNA and histone methylation, transcription, and cell proliferation pathways (Fig. 6C and D and supplementary S6A-D). No specific gene sets were found to be enriched in cells treated with DAC or UNC0638 alone. Gene Set Variation Analysis (GSVA) revealed that the majority of IFN signaling and IFN production signatures were positively enriched as compared with any individual treatment (Fig. 6E). RT-qPCR experiments showed that the combination significantly increased the expression of *IFNB1* mRNA, encoding IFN-β, and TFs such as *IRF9* and *IRF7* involved in IFN signaling (Fig. 6F), confirming RNA-seq results. Altogether, these data demonstrate that TE derepression induced by the combined treatment correlates with the reactivation of viral and inflammatory pathways in CD34$^+$ cells from CMML patients.

The induction of IFN signaling by the combination is essential for eliminating CMML HSPCs

We further examined whether IFN signaling is dysregulated at the protein level. Immunofluorescence analysis demonstrated increased levels of key proteins involved in IFN-I production (IRF3 and IRF7) and IFN signaling (STAT1) in CD34$^+$ cells from CMML patients after 4 days of culture with DAC/UNC0638, compared to untreated cells (68-fold, 50-fold, and 3-fold increase, respectively) (Fig. 7A-C). While HMA treatment alone significantly increased IRF7 and STAT1 protein levels (Fig. 7B and C), the combination induced a much stronger response. Notably, only the combination led to a substantial increase in the phosphorylated form of STAT1 (pSTAT1) (98-fold compared to untreated), indicating the activation of IFN signaling after 4 days of culture (Fig. 7D). In control cells, only a modest increase in IRF3 was observed upon treatment with the combination (1.5-fold), and this increase was notably lower compared to that observed in patients (68-fold) (supplementary Fig. S7A-D). These findings further underscore the specific and potent activation of IFN signaling achieved by the combination treatment in CMML CD34$^+$ cells.

To investigate the potential involvement of ERVs in activating IFN signaling, we assessed dsRNA formation using J2 antibody that specifically recognizes dsRNAs. The combination treatment markedly increased dsRNA expression in patient CD34$^+$ cells, while having no significant effect on control cells (Fig. 7E; supplementary Fig. S7E). dsRNA induction was already evident at day 2 of treatment, suggesting that it may precede activation of IFN signaling. Consistent with its modest impact on IRF7 and STAT1 expression, DAC alone showed a slight increase in dsRNA expression, whereas UNC0638 alone had no observable effect (Fig. 7E). We then explored whether cell death induced by the combination could be related to IFN-I signaling activation by treating CD34$^+$ cells with the DAC/UNC0638 combination in the presence or absence of an antibody blocking IFNAR2 (IFN-I receptor) before being cultured in

methylcellulose (Fig. 7F). The addition of the IFNAR2 blocking antibody partially restored the clonogenic capacity of CMML $CD34^+$ cells in the presence of DAC/UNC0638 (Fig. 7G and H). This finding highlights the contribution of the antiviral response of malignant cells to the efficacy of the DAC/UNC0638 combination.

Discussion

This study reports the first comprehensive characterization of chromatin and transcriptomic changes in CMML compared to age-matched healthy donor CD34$^+$ cells upon *ex vivo* exposure to HMAs. We demonstrate that CMML CD34$^+$ cells, regardless of their mutation pattern, exhibit repression of immune and age-associated transcripts, coupled with an increase in the repressive H3K9me2 mark, predominantly at TEs. This chromatin context may protect them from both TE activation (9,14) and the highly inflammatory environment that characterize aging, and promote their expansion (18-22,32,33). Reactivation of TEs and TE-induced innate immune response, achieved through a combination of HMAs and G9A/GLP inhibitors, appears to decrease the fitness of mutated cells compared to residual wild-type cells. This discovery opens up new therapeutic avenues for treating this severe myeloid malignancy.

The depicted downregulation of proinflammatory genes in CMML compared to age-matched healthy donor CD34$^+$ cells contrasts with the proinflammatory transcriptome of CMML monocytes (32) and the increased level of various inflammatory cytokines in peripheral blood plasma of CMML patients (33-35). HSC aging is associated with an increased expression of proinflammatory genes (18,19,22) likely indicative of the cumulative impact of infections and various stresses over the lifespan, and chronic inflammation promotes premature HSC aging in mice (36). Recent findings highlight chromatin accessibility changes in aged mouse HSCs, with open regions notably enriched for TF activated in response to cytokine signals and inflammation, such as the JAK-STAT pathway. These open regions exhibit a correlation with gene expression levels (18). Senescent human HSPCs exhibit heightened expression of TEs and inflammatory genes (37). The presence of enriched NFκB and IRF motifs in peaks displaying reduced chromatin accessibility suggests that CMML HSPCs may have a diminished capacity

to respond to this highly inflammatory milieu. This might protect them from the deleterious effects of aging and inflammation when compared to wild-type CD34$^+$ cells, a difference that should contribute to the expansion of the leukemic clone. Supporting this hypothesis, we observed a decline in senescence and aging signatures in CD34$^+$ cells from CMML compared to those from aged healthy donors. Our recent results also indicate that dysplastic granulocytes generated by the CMML clone produce high levels of CXCL8 that inhibits the *ex vivo* growth of wild-type CD34$^+$ cells while having limited effects on mutated CD34$^+$ CMML cells, thereby promoting leukemic clone expansion (Deschamps et al., submitted). Consistent with these findings, studies in zebrafish and humans have reported that the muted inflammatory response of mutant HSCs, compared to wild-type HSCs in the same samples, renders them more resistant to inflammation and aging, facilitating their selection (38,39). These results suggest that epigenetic drugs that enable the reactivation of inflammatory/antiviral responses in CMML HSCs may break out such vicious circle in CMML as well as other age-associated malignancies.

In a previous work, we showed that HMAs could restore a balanced hematopoiesis in CMML patients without decreasing the mutation VAF nor preventing clonal evolution, indicating that these compounds could not eradicate the malignant clone (4). Similar observations were made in MDS patients treated with 5-AZA (40,41). An additional observation is provided by our experiments indicating that HMAs selectively target progenitors while having no detected effect on stem cells in both liquid culture and methylcellulose assays. Even more, DAC appears to slightly promote the growth of the most mutated cells, which could explain why CD34$^+$ cells from HMA-responsive MDS patients were observed to be enriched in progenitor cells as compared with non-responders (40,41). Similarly, HMAs were reported to predominantly target granulo-monocytic progenitors in high-grade MDS and acute myeloid leukemia (AML)

(41). The high expression of cell cycle-associated genes depicted in CMML CD34$^+$ cells who respond to HMAs may promote their incorporation in nucleic acids (42). The lower expansion of stem cells compared to progenitors (mean fold expansion of 20.34 +/- 0.34 and 32.26 +/- 0.98 in 8 days, respectively) may also contribute to the greater sensitivity of progenitors to HMAs. Alternatively, differences in DNA and histone methylation, expression of genes encoding RNA helicases and autophagy genes (43) might exist between CMML progenitor and stem cells.

Our findings indicate that DAC as a single agent is an inefficient inducer of either TE expression or inflammatory / IFN signaling in CMML CD34$^+$ cells *ex vivo*. The role of these events in HMA activity in myeloid malignancies remains controversial. 5-AZA was initially reported to upregulate the expression of some ERVs in primary CD34$^+$ cells collected from MDS patients (44). Overexpression of TEs in DNMT3A-mutant AML cells renders them more sensitive to 5-AZA via viral mimicry response (45). An increased transcription of evolutionary young TEs with an activation of innate immune responses was depicted in some 5-AZA-responders (46) but not validated in others (47). These discrepancies could be related to the heterogeneity of the studied cohorts with diverse diseases (46,47) and genetic backgrounds (48,49) as well as methodological issues.

While TEs are commonly recognized for their tumor-promoting functions, recent studies have suggested that RTEs, including active LINE-1 and ERVs, have been evolutionarily selected for their tumor-suppressing function through their ability to induce viral mimicry and/or p53 activation (50). Accordingly, TE transcription is commonly suppressed along tumor progression, for example from low-grade to high grade MDS, then a full-blown AML (43), where a low repeat/gene expression ratio correlates with a poor outcome (51). This is

associated with the inactivation of innate immune pathways, including IFN and NFκB signatures, potentially contributing to the immune evasion mechanism observed in this pathology. One study also reported that CMML and MDS CD34$^+$ cells expressed lower levels of TEs than healthy donor cells (47). Epigenetic mechanisms that were shown to contribute to TE repression in myeloid malignancies include the H3K9 methyltransferase SETDB1 (52) and the epigenetic repressor complex HUSH whose component MMP8 mediates LINE-1 element repression (53). The decrease in H3K9me3 observed in CMML CD34$^+$ cells suggests this mark may not be involved in the repression of TEs and TE-induced innate immune response in these cells. The sharp increase in H3K9me2 and G9A/GLP methyltransferase expression, which was detected in CMML cells may compensate for H3K9me3 loss. A similar increase in H3K9me2 expression was shown to prevent type I IFN pathway activation and immune gene expression in AML cell lines (24). Our CUT&Tag analysis show that H3K9me2 is mainly located at TEs and did not indicate a direct inhibitory effect of H3K9me2 on immune gene expression. In addition, the G9A/GLP inhibitor UNC0638 alone did not induce IFN/inflammatory signaling in CMML CD34$^+$ cells. G9A interacts with DNMT1 and recruits DNMT3A/B through MPP8 to promote the compaction of chromatin in H3K9me2-marked regions (54). In embryonic stem cells, loss of H3K9me2 was shown insufficient to generate a global DNA demethylation (55). In HSPCs, the H3K9me2 mark is detected preferentially in demethylated regions and participates in global changes in chromatin structure during myeloid differentiation (27). This could explain a synergistic effect of G9A/GLP inhibition with HMAs on TE expression.

Besides controlling TE expression, G9A/GLP has been shown to play a role in AML leukemic cell stemness (56). Altogether, the present study designates G9A/GLP inhibition as a promising strategy to specifically target leukemic stem cells in CMML. Unfortunately, our attempts to explore the ability of HMAs combined to the current G9A/GLP inhibitors to eradicate leukemic

stem cells in patient-derived xenografts established in mice was precluded by the toxic side effects of the combination compounds, suggesting that less toxic direct inhibitors or alternative indirect approaches will be needed to validate the merit of this therapeutic approach.

Materials and methods

Patient and healthy donor samples and processing

The study was carried out on patients diagnosed with CMML, before any treatment according to the 2016 iteration of the WHO classification (supplementary Table S1). Patient BM samples were obtained as part of a BM assessment of the disease carried out in one of the hematology department affiliated to the Groupe Francophone des Myélodysplasies (GFM). Informed consents were obtained in agreement with Ile-de-France ethics committee (MYELOMONO cohort, DC-2014-2091). Mutations were detected in patients' monocyte DNA by next generation sequencing (NGS) using a myeloid panel, as previously described (4). Age-matched healthy donor BM samples were obtained from femoral heads from individuals who had undergone surgery to treat mechanical osteoarthritis, thanks to Dr Juliette RAFFIN (Banque de tissus osseux, Hôpital Cochin). Cord blood was obtained from Unité de Thérapie Cellulaire, CRB-Banque de Sang de Cordon, (AP-HP Hôpital Saint-Louis, Paris) - Authorization no.: AC-2022-5325. Informed consent was obtained from all donors, in accordance with the Declaration of Helsinki. Samples were collected sequentially and cell culture were performed directly on arrival without prior freezing / thawing. Patient samples were collected by BM puncture with EDTA aspiration. BM cells from healthy elderly subjects were obtained by flushing the femoral head in PBS (Gibco). After washing and resuspension in PBS, the cells were filtered on 70 µM nylon mesh. HSPCs were obtained from mononuclear cells sorted on a Ficoll gradient (Pancoll, PAN-Biotech) and isolated by positive magnetic cell sorting performed on CD34 marker expression following the manufacturer's instructions (CD34 MicroBead Kit, AutoMACS ProSeparator, Miltenyi Biotech). $CD34^+$ enrichment was evaluated after staining with anti-CD34-APC (BD Biosciences; clone 581) and FACS analysis. For sorting stem cell-enriched ($CD34^+CD38^-CD90^+$) and progenitor-enriched ($CD34^+CD38^-CD90^-$)

populations, cells were further stained with CD38-PE-Cy7 (BD Biosciences; clone HB-7), CD90-FITC (Sony Biotechnologies; clone HIT2) and sorted on ARIA fusion cell sorter.

Mice and mouse HSC isolation

C57BL/6 mice were from ENVIGO. All the mice were housed in a specific pathogen-free environment. All procedures were reviewed and approved by the Animal Care Committee N°26 (APAFIS #23286-2019121214392327 and APAFIS #35147-2022020315419972). BM was harvested from femur, tibia and hip bones. Total BM was depleted of differentiated hematopoietic cells (lineage-positive cells) using Mouse Hematopoietic Progenitor Stem Cell Enrichment Set (BD Biosciences) and HSCs ($Lin^{-}Sca^{+}c\text{-}Kit^{+}CD34^{low}Flk2^{-}$) were sorted, as described (11) using the following antibodies: CD3ε (Lin) – APC clone 145-2C11 (553066, BD), TER-119 (Lin) – APC clone Ter-119 (557909), CD45R/B220 (Lin) – APC clone RA3-6B2 (553092), Ly6G-6C (Lin)-APC clone RB6-8C5 (553129), Ly-6A/E (Sca-1) - PeCy7 clone d7 (558162), CD117 (c-Kit) – PE or PerCP-Cy5.5, clone 2B8 (553355 or 560557 respectively), CD34– FITC or AF700 clone RAM34 (560238 or 560518), CD135 (Flk2) – BV421 or PE clone A2F10.1 (562898 or 553842 respectively), all from BD Biosciences.

Cell cultures and treatments

Freshly isolated human HSPCs were plated at 200,000 cells/ml and cultured in IMDM medium supplemented with 25% BIT9500 (StemCell Technologies), 1% penicillin/streptomycin (P/S), containing the following cytokines (all from Miltenyi Biotech): FTL3-Ligand (FTL3-L, 100ng/µL), Thrombopoietin (THPO, 100ng/µL), stem cell factor (SCF, 60ng/µL) and interleukin 3 (IL3, 10ng/µL), at 37°C, 20% O2, 5% CO2. For treatments, decitabine (DAC, Sigma Aldrich), or 5-Azacytidine (AZA, Sigma Aldrich) were added at the start of the cultures (day 0) and again on day 1. G9A/GLP inhibitor (UNC0638 or UNC0642, MedChem) was added once on day 2. The

dual DNMT/G9A inhibitor, CM-272 (Axon Medchem) was added on day 0 and again 48h later and the cells were directly seeded in methylcellulose.

Single cell colony and CFU assays

Single human or mouse HSCs/HSPCs were FACS-sorted into individual wells of round-bottom 96 well plates and cultured in the presence or absence of inhibitors. At different times starting day one, each well was checked manually for the the presence of cells. Survival index was calculated as the percent of wells containing at least one cell divided by the total number of seeded wells. The proliferation index was calculated as the total number of cells counted at given time point on the total number of wells containing at least one cell. For CFU assays, 300 $CD34^+$ or 100 $CD34^+CD38^-CD90^{-/+}$ treated or not with the inhibitors were seeded in duplicate in 24-well plates containing 300µl of methylcellulose (H4230, StemCell), supplemented with 1% P/S (ThermoFischer Scientific), FLT3-L (25ng/µL), THPO (25ng/µL), SCF (60ng/µL), IL3 (10ng/mL), EPO (2 units), GM-CSF (10ng/mL). Colonies were counted after 14 days of culture. The mean size of the colonies was estimated by harvesting them and counting the total number of cells in the well. For replating assays, 5000 cells were replated in methylcellulose in the same medium. This process was repeated twice. The cumulative CFU potential that takes into account both the number of CFUs and their size was calculated at passage 2 (P2) as: [number of colonies / number of cells seeded at P2] x total cell number recovered at P1, as described (30).

Gene mutation screening in colonies

Mutations with VAF $\geq$ 30% were analyzed after *in vitro* treatment in 5 patients (supplementary Table S5). Colonies from $CD34^+$ or $CD34^+CD38^-CD90^+$ cells grown in methylcellulose or in single cell liquid cultures, respectively, in the presence or absence of DAC and/or UNC0638, were

picked up individually after 14 days, washed and stored at -80°C as dry pellets. Cells were lysed with a solution containing proteinase K (0.1 mg/mL, Fisher Scientific), 2% Tween-20 (Sigma-Aldrich) and incubated for 1 h at 65°C followed by 15 min at 95°C. DNA was amplified by PCR using the Q5 High-Fidelity DNA polymerase kit (New England BioLabs) following the manufacturer's recommendations and with primers designed around the mutations (supplementary Table S9). Amplicon sizes were verified by migrating the amplified DNA sample onto a 1.8% agarose gel. The amplified DNA fragments were anlayzed by Sanger sequencing (Eurofins Genomics, Ebersberg, Germany). 8-12 clones were tested per condition.

ELISA for DNA methylation

Cell lysis and extraction of genomic DNA from dry pellets were performed using the NucleoSpin Tissue XS kit (Macherey-Nagel). Genomic DNA quality and quantity were measured using the NanoDrop (ND-1000, Marshall Scientific). Global DNA methylation was studied using the global DNA methylation assay kit - LINE-1 (Active Motif) according to the manufacturer’s instructions.

Western Blot

Dry pellets of 10,000 $CD34^+$ cells stored at -80°C were resuspended in Laemmli sample buffer (ref. 1610747, bioRad®), and incubated at 95°C for 10 min. Denatured protein extracts were separated on a 15% polyacrylamide-SDS gel and transferred to a nitrocellulose membrane (ThermoFischer Scientific). Membranes were saturated using a solution of 5% BSA, 0.1% Tween-20 for 30 min under agitation at room temperature and incubated overnight at 4°C with anti-H3, anti-H3K9me2 or anti-H3K9me3 primary antibodies (supplementary Table S10), followed by incubation with goat anti-rabbit HRP-coupled antibodies (ThermoFisher Scientific 31460, 1/2000) for 1h at RT on a shaker. The membranes were revealed by

chemiluminescence using ECL reagent (Enhanced ChemiLuminescence, ThermoFischer Scientific™ SuperSignal™ West Pico) and images visualized with iBright 1500 Western Blot imaging (Invitrogen).

Immunofluorescence

5000 human and murine HSPCs/HSCs were fixed with 4% paraformaldehyde (PFA) for 10 min at RT and centrifuged (Tharmac) on poly-L-Lysine slides (Fisher Scientific). Immunofluorescence was performed as described previously (11,14). The primary antibodies used are listed in supplementary Table S10. Detection was performed using anti-rabbit secondary antibody coupled to Alexa-Fluor 488 or anti-mouse coupled to Alexa-Fluor-555. All slides were viewed under a TCS SPE confocal microscope (Leica). Images were analyzed and fluorescence intensities calculated (in at least 50 cells) using ImageJ software, as described (11,14).

RNA purification and quantitative RT-PCR

50,000 - 150,000 cells were lysed in Tri-Reagent solution (Zymo Research) and stored at -80°C until use. Total RNA was extracted using the Direct-Zol RNA microprep kit (Zymo Research). RNAs were reverse transcribed using the EZ DNAse VILO IV kit (Invitrogen). 1.25 µL of cDNA was pre-amplified for 15-18 cycles in a multiplex reaction using Preamp Master-Mix solution (ref. 100-5580, Fluidigm) and primer mix (200 µM of each primer), as described (11). The primers chosen bind to different exons, so as to minimize amplification of genomic DNA (supplementary Table S9). In addition, primer mix were preanalyzed on MFE3.0 PCR Primer Quality Control to eliminate the risk of primer dimerization or hairpin formation (11). Quantitative PCR (qPCR) was performed using LUNA Universal qPCR Master Mix (New England

Biolabs) on the 7500 real-time PCR machine (Applied Biosystems). Data were normalized with the mean expression of household genes (HPRT, Tubulin, GUS, PPIA and / or RPL32).

Assay for Transposase-Accessible Chromatin with high throughput sequencing (ATAC-seq) and data processing

100,000 freshly isolated $CD34^+$ cells were subjected to lysis, transposition and purification steps using the ATAC-seq kit (Active Motif), according to the manufacturer's recommendations. Transposed DNA fragments were amplified 10-fold by PCR using adapters supplied by the kit. PCR purification was carried out using MinElute PCR Purification Kit (Qiagen, 28004) to remove remaining primers and large fragments. The quantity and quality of the libraries were assessed on Agilent 2100 Bioanalyzer (Agilent Technologies 50567-4626). Sequencing of the libraries was performed on the NovaSeq-6000 at Gustave Roussy (Illumina; 50 bp paired-end reads). ATAC-seq data were processed essentially as described (15) (supplementary methods).

RNA-sequencing

RNA quality was checked on the Agilent Bioanalyzer and quantified. 100-250 pg of total RNA extracted from $CD34^+$ cells were used to prepare cDNA libraries, as described (11), using the SMARTer Ultra Low Input RNA Kit for Sequencing - v3 or -SMARTer Stranded Total RNA-seq kit v3, according to the manufacturer's recommendations (Takara). The fragmentation time was adjusted according to RNA quality. The quantity and quality of the libraries were evaluated on the Agilent Bioanalyzer. Libraries were pooled and sequenced in (2 x 100bp) on the NovaSeq-6000 (Illumina). Gene and ERV expression was quantified as described (11,31) (supplementary methods).

CUT&Tag

CUT&Tag assays were performed using CUT&Tag-IT assay kit (Active Motif) on 50,000 CD34+ cells according to manufacturer's instructions. Cells were incubated O/N with 0.5µg of H3K9me2 (Abcam Ab1220) (supplementary Table S10). Sequencing of the libraries was performed on the NovaSeq-6000 (Illumina; 2x 50 bp). CUT&Tag data were processed essentially as decribed (11) (supplementary methods). TE enrichment was evaluated through a permutation test (x1000) conducted independently for each TE class using peaks that were exclusively found in CMML patient control cells and equal number of peaks randomly chosen from peaks that do not belong to any section of the Venn diagram.

Statistical analyses:

Results were statistically evaluated by one-way or two-way ANOVA tests or a t-test using Prism GraphPad PrismTM version 9.0 software (GraphPad Software Inc., San Diego, CA, USA). Graphs show mean and standard error to the mean (SEM). $p<0.05$ is considered significant. For all figures, *$p<0.05$, **$p<0.01$, ***$p<0.001$, ****$p<0.001$.

Supplemental material

Supplementary material includes 7 supplementary figures and figure legends related to figures 1 to 7, 10 tables, additional methods for RNA-seq, ATAC-seq and CUT&Tag data processing and additional references.

Acknowledgments

We thank animal facility, the Genomic and the Imaging and Cytometry Platforms of Gustave Roussy for sequencing and cell sorting and confocal analysis, respectively and Dr. Louis Paré (Nantes University, École Centrale Nantes) for RNA-seq analysis. This work was supported by

INSERM and grants from Ligue National Contre le Cancer (LNCC, Equipe labellisée EL2020) and Institut National du Cancer (PLBIO N°2020-095) to F.P. D.H. and A.P. are recipients of fellowship from the Ministère de l'Enseignement Supérieur de la Recherche et de l'Innovation.

References

1. Khoury JD, Solary E, Abla O, Akkari Y, Alaggio R, Apperley JF*, et al.* The 5th edition of the World Health Organization Classification of Haematolymphoid Tumours: Myeloid and Histiocytic/Dendritic Neoplasms. Leukemia **2022**;36(7):1703-19 doi 10.1038/s41375-022-01613-1.
2. Itzykson R, Kosmider O, Renneville A, Morabito M, Preudhomme C, Berthon C*, et al.* Clonal architecture of chronic myelomonocytic leukemias. Blood **2013**;121(12):2186-98 doi 10.1182/blood-2012-06-440347.
3. Binder M, Carr RM, Lasho TL, Finke CM, Mangaonkar AA, Pin CL*, et al.* Oncogenic gene expression and epigenetic remodeling of cis-regulatory elements in ASXL1-mutant chronic myelomonocytic leukemia. Nat Commun **2022**;13(1):1434 doi 10.1038/s41467-022-29142-6.
4. Merlevede J, Droin N, Qin T, Meldi K, Yoshida K, Morabito M*, et al.* Mutation allele burden remains unchanged in chronic myelomonocytic leukaemia responding to hypomethylating agents. Nat Commun **2016**;7:10767 doi 10.1038/ncomms10767.
5. Palomo L, Malinverni R, Cabezon M, Xicoy B, Arnan M, Coll R*, et al.* DNA methylation profile in chronic myelomonocytic leukemia associates with distinct clinical, biological and genetic features. Epigenetics **2018**;13(1):8-18 doi 10.1080/15592294.2017.1405199.
6. Solary E, Itzykson R. How I treat chronic myelomonocytic leukemia. Blood **2017**;130(2):126-36 doi 10.1182/blood-2017-04-736421.
7. Tsurumi A, Li WX. Global heterochromatin loss: a unifying theory of aging? Epigenetics **2012**;7(7):680-8 doi 10.4161/epi.20540.
8. Zhang W, Qu J, Liu GH, Belmonte JCI. The ageing epigenome and its rejuvenation. Nat Rev Mol Cell Biol **2020**;21(3):137-50 doi 10.1038/s41580-019-0204-5.
9. Djeghloul D, Kuranda K, Kuzniak I, Barbieri D, Naguibneva I, Choisy C*, et al.* Age-Associated Decrease of the Histone Methyltransferase SUV39H1 in HSC Perturbs Heterochromatin and B Lymphoid Differentiation. Stem Cell Reports **2016**;6(6):970-84 doi 10.1016/j.stemcr.2016.05.007.
10. Keenan CR, Iannarella N, Naselli G, Bediaga NG, Johanson TM, Harrison LC, Allan RS. Extreme disruption of heterochromatin is required for accelerated hematopoietic aging. Blood **2020**;135(23):2049-58 doi 10.1182/blood.2019002990.
11. Pelinski Y, Hidaoui D, Stolz A, Hermetet F, Chelbi R, Diop MK*, et al.* NF-kappaB signaling controls H3K9me3 levels at intronic LINE-1 and hematopoietic stem cell genes in cis. J Exp Med **2022**;219(8) doi 10.1084/jem.20211356.
12. Chuong EB, Elde NC, Feschotte C. Regulatory activities of transposable elements: from conflicts to benefits. Nat Rev Genet **2017**;18(2):71-86 doi 10.1038/nrg.2016.139.
13. Gazquez-Gutierrez A, Witteveldt J, S RH, Macias S. Sensing of transposable elements by the antiviral innate immune system. RNA **2021**;27(7):735-52 doi 10.1261/rna.078721.121.
14. Barbieri D, Elvira-Matelot E, Pelinski Y, Geneve L, de Laval B, Yogarajah G*, et al.* Thrombopoietin protects hematopoietic stem cells from retrotransposon-mediated damage by promoting an antiviral response. J Exp Med **2018**;215(5):1463-80 doi 10.1084/jem.20170997.
15. Clapes T, Polyzou A, Prater P, Sagar, Morales-Hernandez A, Ferrarini MG*, et al.* Chemotherapy-induced transposable elements activate MDA5 to enhance haematopoietic regeneration. Nat Cell Biol **2021**;23(7):704-17 doi 10.1038/s41556-021-00707-9.
16. Chiappinelli KB, Strissel PL, Desrichard A, Li H, Henke C, Akman B*, et al.* Inhibiting DNA Methylation Causes an Interferon Response in Cancer via dsRNA Including Endogenous Retroviruses. Cell **2015**;162(5):974-86 doi 10.1016/j.cell.2015.07.011.
17. Roulois D, Loo Yau H, Singhania R, Wang Y, Danesh A, Shen SY*, et al.* DNA-Demethylating Agents Target Colorectal Cancer Cells by Inducing Viral Mimicry by Endogenous Transcripts. Cell **2015**;162(5):961-73 doi 10.1016/j.cell.2015.07.056.

18. Itokawa N, Oshima M, Koide S, Takayama N, Kuribayashi W, Nakajima-Takagi Y*, et al.* Epigenetic traits inscribed in chromatin accessibility in aged hematopoietic stem cells. Nat Commun **2022**;13(1):2691 doi 10.1038/s41467-022-30440-2.
19. Mann M, Mehta A, de Boer CG, Kowalczyk MS, Lee K, Haldeman P*, et al.* Heterogeneous Responses of Hematopoietic Stem Cells to Inflammatory Stimuli Are Altered with Age. Cell Rep **2018**;25(11):2992-3005 e5 doi 10.1016/j.celrep.2018.11.056.
20. Kowalczyk MS, Tirosh I, Heckl D, Rao TN, Dixit A, Haas BJ*, et al.* Single-cell RNA-seq reveals changes in cell cycle and differentiation programs upon aging of hematopoietic stem cells. Genome Res **2015**;25(12):1860-72 doi 10.1101/gr.192237.115.
21. Kasbekar M, Mitchell CA, Proven MA, Passegue E. Hematopoietic stem cells through the ages: A lifetime of adaptation to organismal demands. Cell Stem Cell **2023**;30(11):1403-20 doi 10.1016/j.stem.2023.09.013.
22. Adelman ER, Huang HT, Roisman A, Olsson A, Colaprico A, Qin T*, et al.* Aging Human Hematopoietic Stem Cells Manifest Profound Epigenetic Reprogramming of Enhancers That May Predispose to Leukemia. Cancer Discov **2019**;9(8):1080-101 doi 10.1158/2159-8290.CD-18-1474.
23. Fang TC, Schaefer U, Mecklenbrauker I, Stienen A, Dewell S, Chen MS*, et al.* Histone H3 lysine 9 di-methylation as an epigenetic signature of the interferon response. J Exp Med **2012**;209(4):661-9 doi 10.1084/jem.20112343.
24. Hansen AM, Ge Y, Schuster MB, Pundhir S, Jakobsen JS, Kalvisa A*, et al.* H3K9 dimethylation safeguards cancer cells against activation of the interferon pathway. Sci Adv **2022**;8(11):eabf8627 doi 10.1126/sciadv.abf8627.
25. Avgustinova A, Laudanna C, Pascual-Garcia M, Rovira Q, Djurec M, Castellanos A*, et al.* Repression of endogenous retroviruses prevents antiviral immune response and is required for mammary gland development. Cell Stem Cell **2021**;28(10):1790-804 e8 doi 10.1016/j.stem.2021.04.030.
26. Liu M, Thomas SL, DeWitt AK, Zhou W, Madaj ZB, Ohtani H*, et al.* Dual Inhibition of DNA and Histone Methyltransferases Increases Viral Mimicry in Ovarian Cancer Cells. Cancer Res **2018**;78(20):5754-66 doi 10.1158/0008-5472.CAN-17-3953.
27. Schones DE, Chen X, Trac C, Setten R, Paddison PJ. G9a/GLP-dependent H3K9me2 patterning alters chromatin structure at CpG islands in hematopoietic progenitors. Epigenetics Chromatin **2014**;7:23 doi 10.1186/1756-8935-7-23.
28. Chou TC. Theoretical basis, experimental design, and computerized simulation of synergism and antagonism in drug combination studies. Pharmacol Rev **2006**;58(3):621-81 doi 10.1124/pr.58.3.10.
29. San Jose-Eneriz E, Agirre X, Rabal O, Vilas-Zornoza A, Sanchez-Arias JA, Miranda E*, et al.* Discovery of first-in-class reversible dual small molecule inhibitors against G9a and DNMTs in hematological malignancies. Nat Commun **2017**;8:15424 doi 10.1038/ncomms15424.
30. Higa KC, Goodspeed A, Chavez JS, De Dominici M, Danis E, Zaberezhnyy V*, et al.* Chronic interleukin-1 exposure triggers selection for Cebpa-knockout multipotent hematopoietic progenitors. J Exp Med **2021**;218(6) doi 10.1084/jem.20200560.
31. Alcazer V, Bonaventura P, Tonon L, Michel E, Mutez V, Fabres C*, et al.* HERVs characterize normal and leukemia stem cells and represent a source of shared epitopes for cancer immunotherapy. Am J Hematol **2022**;97(9):1200-14 doi 10.1002/ajh.26647.
32. Franzini A, Pomicter AD, Yan D, Khorashad JS, Tantravahi SK, Than H*, et al.* The transcriptome of CMML monocytes is highly inflammatory and reflects leukemia-specific and age-related alterations. Blood Adv **2019**;3(20):2949-61 doi 10.1182/bloodadvances.2019000585.
33. Niyongere S, Lucas N, Zhou JM, Sansil S, Pomicter AD, Balasis ME*, et al.* Heterogeneous expression of cytokines accounts for clinical diversity and refines prognostication in CMML. Leukemia **2019**;33(1):205-16 doi 10.1038/s41375-018-0203-0.

34. Pronier E, Imanci A, Selimoglu-Buet D, Badaoui B, Itzykson R, Roger T*, et al.* Macrophage migration inhibitory factor is overproduced through EGR1 in TET2(low) resting monocytes. Commun Biol **2022**;5(1):110 doi 10.1038/s42003-022-03057-w.
35. Sevin M, Debeurme F, Laplane L, Badel S, Morabito M, Newman HL*, et al.* Cytokine-like protein 1-induced survival of monocytes suggests a combined strategy targeting MCL1 and MAPK in CMML. Blood **2021**;137(24):3390-402 doi 10.1182/blood.2020008729.
36. Esplin BL, Shimazu T, Welner RS, Garrett KP, Nie L, Zhang Q*, et al.* Chronic exposure to a TLR ligand injures hematopoietic stem cells. J Immunol **2011**;186(9):5367-75 doi 10.4049/jimmunol.1003438.
37. Capone S, Connor KM, Colombo A, Li X, Triche TJ, Jr., Ramsingh G. Senescent human hematopoietic progenitors show elevated expression of transposable elements and inflammatory genes. Exp Hematol **2018**;62:33-8 e6 doi 10.1016/j.exphem.2018.03.003.
38. Avagyan S, Henninger JE, Mannherz WP, Mistry M, Yoon J, Yang S*, et al.* Resistance to inflammation underlies enhanced fitness in clonal hematopoiesis. Science **2021**;374(6568):768-72 doi 10.1126/science.aba9304.
39. Jakobsen NA, Turkalj S, Zeng AGX, Stoilova B, Metzner M, Nagree MS*, et al.* Selective advantage of mutant stem cells in clonal hematopoiesis occurs by attenuating the deleterious effects of inflammation and aging. bioRxiv **2023**:2023.09.12.557322 doi 10.1101/2023.09.12.557322.
40. Unnikrishnan A, Papaemmanuil E, Beck D, Deshpande NP, Verma A, Kumari A*, et al.* Integrative Genomics Identifies the Molecular Basis of Resistance to Azacitidine Therapy in Myelodysplastic Syndromes. Cell Rep **2017**;20(3):572-85 doi 10.1016/j.celrep.2017.06.067.
41. Ali A, Penneroux J, Dal Bello R, Jr., Masse A, Quentin S, Unnikrishnan A*, et al.* Granulomonocytic progenitors are key target cells of azacytidine in higher risk myelodysplastic syndromes and acute myeloid leukemia. Leukemia **2018**;32(8):1856-60 doi 10.1038/s41375-018-0076-2.
42. Meldi K, Qin T, Buchi F, Droin N, Sotzen J, Micol JB*, et al.* Specific molecular signatures predict decitabine response in chronic myelomonocytic leukemia. J Clin Invest **2015**;125(5):1857-72 doi 10.1172/JCI78752.
43. Colombo AR, Zubair A, Thiagarajan D, Nuzhdin S, Triche TJ, Ramsingh G. Suppression of Transposable Elements in Leukemic Stem Cells. Sci Rep **2017**;7(1):7029 doi 10.1038/s41598-017-07356-9.
44. Tobiasson M, Abdulkadir H, Lennartsson A, Katayama S, Marabita F, De Paepe A*, et al.* Comprehensive mapping of the effects of azacitidine on DNA methylation, repressive/permissive histone marks and gene expression in primary cells from patients with MDS and MDS-related disease. Oncotarget **2017**;8(17):28812-25 doi 10.18632/oncotarget.15807.
45. Scheller M, Ludwig AK, Gollner S, Rohde C, Kramer S, Stable S*, et al.* Hotspot DNMT3A mutations in clonal hematopoiesis and acute myeloid leukemia sensitize cells to azacytidine via viral mimicry response. Nat Cancer **2021**;2(5):527-44 doi 10.1038/s43018-021-00213-9.
46. Ohtani H, Orskov AD, Helbo AS, Gillberg L, Liu M, Zhou W*, et al.* Activation of a Subset of Evolutionarily Young Transposable Elements and Innate Immunity Are Linked to Clinical Responses to 5-Azacytidine. Cancer Res **2020**;80(12):2441-50 doi 10.1158/0008-5472.CAN-19-1696.
47. Kazachenka A, Young GR, Attig J, Kordella C, Lamprianidou E, Zoulia E*, et al.* Epigenetic therapy of myelodysplastic syndromes connects to cellular differentiation independently of endogenous retroelement derepression. Genome Med **2019**;11(1):86 doi 10.1186/s13073-019-0707-x.
48. Bejar R, Lord A, Stevenson K, Bar-Natan M, Perez-Ladaga A, Zaneveld J*, et al.* TET2 mutations predict response to hypomethylating agents in myelodysplastic syndrome patients. Blood **2014**;124(17):2705-12 doi 10.1182/blood-2014-06-582809.
49. Duchmann M, Yalniz FF, Sanna A, Sallman D, Coombs CC, Renneville A*, et al.* Prognostic Role of Gene Mutations in Chronic Myelomonocytic Leukemia Patients Treated With Hypomethylating Agents. EBioMedicine **2018**;31:174-81 doi 10.1016/j.ebiom.2018.04.018.

50. Kelsey MMG. Reconsidering LINE-1's role in cancer: does LINE-1 function as a reporter detecting early cancer-associated epigenetic signatures? Evol Med Public Health **2021**;9(1):78-82 doi 10.1093/emph/eoab004.
51. Onishi-Seebacher M, Erikson G, Sawitzki Z, Ryan D, Greve G, Lubbert M, Jenuwein T. Repeat to gene expression ratios in leukemic blast cells can stratify risk prediction in acute myeloid leukemia. BMC Med Genomics **2021**;14(1):166 doi 10.1186/s12920-021-01003-z.
52. Cuellar TL, Herzner AM, Zhang X, Goyal Y, Watanabe C, Friedman BA*, et al.* Silencing of retrotransposons by SETDB1 inhibits the interferon response in acute myeloid leukemia. J Cell Biol **2017**;216(11):3535-49 doi 10.1083/jcb.201612160.
53. Gu Z, Liu Y, Zhang Y, Cao H, Lyu J, Wang X*, et al.* Silencing of LINE-1 retrotransposons is a selective dependency of myeloid leukemia. Nature Genetics **2021**;53(5):672-82 doi 10.1038/s41588-021-00829-8.
54. Esteve PO, Chin HG, Smallwood A, Feehery GR, Gangisetty O, Karpf AR*, et al.* Direct interaction between DNMT1 and G9a coordinates DNA and histone methylation during replication. Genes Dev **2006**;20(22):3089-103 doi 10.1101/gad.1463706.
55. Jiang Q, Ang JYJ, Lee AY, Cao Q, Li KY, Yip KY, Leung DCY. G9a Plays Distinct Roles in Maintaining DNA Methylation, Retrotransposon Silencing, and Chromatin Looping. Cell Rep **2020**;33(4):108315 doi 10.1016/j.celrep.2020.108315.
56. Lehnertz B, Pabst C, Su L, Miller M, Liu F, Yi L*, et al.* The methyltransferase G9a regulates HoxA9-dependent transcription in AML. Genes Dev **2014**;28(4):317-27 doi 10.1101/gad.236794.113.

Figure legends

Figure 1: CMML HSPCs show a decrease in inflammatory response pathways.

(A) Principal component analysis (PCA) plot of BM $CD34^+$ cells isolated from 19 CMML patients at diagnosis and 12 controls, based on RNA-seq regularized log transformed expression values.

(B) MA-plots showing differentially expressed genes (padj <0.1), upregulated (blue dots) or downregulated (black dots) in CMML cells.

(C-E) Dot plots showing selected enriched (C) HALLMARK and (D) GO and (E) senescence gene sets, upregulated (positive NES, red dots) and downregulated (negative NES, blue dots) in CMML cells. Gene sets are ranked by their NES and gene ratio scores.

(F, G) Enrichment plots for published gene sets upregulated or downregulated with age. FDR, false discover rate; NES, normalized enrichment score. * FDR <0.05.

Figure 2: Chromatin is remodeled in CMML HSPCs.

(A-D) Immunostaining for heterochromatin marks H3K9me3 (A); H3K9me2 (B), and factors G9A (C) and GLP (D) in $CD34^+$ cells from CMML patients and controls. Representative images and quantification of mean global immunofluorescence intensity using ImageJ software are shown. Means +/- SEM from 2-3 independent experiments normalized to the mean intensity of controls. Each dot represents cells from a single individual, with at least 50 cells counted per sample. Bars, 3 µm. t test, **P<0.01; ***p<0.001; ****, P<0.0001.

(E) Heat map of the differentially chromatin accessible regions in CMML patient (n=4) and control (n=3) $CD34^+$ cells +/- 2.5 kb from the center of the peak (CoP).

(F) Venn diagrams depicting the overlap of differentially expressed genes (DEGs) in CMML HSPCs with genes assigned to gained or lost ATAC peaks (-100+/-25 kb from TSS). Hypergeometric test.

(G) KEGG gene set analysis of deregulated genes that also exhibit gained or loss chromatin accessibility in CMML patient cells.

(H) Motifs and p values for transcription factor motif enrichment in ATAC-seq peaks unique to patients (gained) and controls (lost) using HOMER. ND: non detected.

(I) Venn diagram of peak retrieved from H3K9me2 CUT&Tag peak calling on $CD34^+$ from 4 patients (P) and 3 controls (CTA).

(J) Repartition of the CUT&Tag peaks found exclusively in patients and controls compared with hg19 reference genome annotation repartition (upper panel).

(K) Numbers of peaks found exclusively in the 4 patients (red) and 4 controls (blue) and numbers of TEs annotated in these peaks.

(L) Enrichment or depletion of the different classes of TEs in gained or lost H3K9me2-CUT&Tag peaks. The Y-axis representing the enrichment score was calculated as the log2 fold change of the number of specific peaks overlapping with TEs found in patient gained or lost peaks over the median number of TEs found in 1000 times the same number of randomly selected peaks overlapping with TEs. Positive log2 fold change = enrichment, and negative log2 fold change = depletion.

(M) Repartition of the different TE classes in CUT&Tag peaks found exclusively in patients and controls.

Figure 3: the HMA and G9A/GLP inhibitor combination reduces CMML HSPC clonogenicity.

(A) Treatment protocol of $CD34^+$ cells. DAC: decitabine; AZA: 5-azacytidine; UNC0638 and UNC0642: G9A/GLP inhibitors; CM-272: dual DNMT1 and G9A/GLP inhibitor. Colonies are counted after 14 days of culture and replated.

(B) Dose-response effect of DAC (X-axis) and UNC0638 (color scale) on CMML $CD34^+$ colony formation. Mean +/- SEM of duplicate wells from 3 patients.

(C-D) Chou-Talalay model (27) of the effects of combining different doses of DAC and UNC0638. The fraction affected (fa) corresponds to the proportion of colonies eliminated by the treatments. The combination index (CI) is calculated by CompuSyn software (ComboSyn. Inc.) where CI<1, CI=1, CI>1 indicates a synergistic, additive and antagonistic effect respectively. Mean+/- SEM of duplicate wells from 3 patients.

(E) Number of colonies formed from 600 CMML (n=28) and control (n=12) $CD34^+$ cells in the absence of treatment. One dot represents a single individual.

(F, G) Number of colonies formed by $CD34^+$ cells from patients (F) or controls (G) after treatment with 10 nM DAC and 1µM UNC0638 alone or in combination. NT: non-treated. One dot represents a single individual; controls, n=12; patients, n=28. Results are normalized to the number of colonies in the NT condition. 2 or 3 wells per sample. Means +/- SEM. One-way ANOVA Bonferroni's Multiple Comparison.

(H, I) Number of colonies formed by $CD34^+$ cells from patients, n=5 (L) or controls, n=4, (M) after treatment with or without (NT) 200 nM CM-272. Results are normalized to the number of colonies in the NT condition. Means +/- SEM. t.test.

(E-I) One dot represents a single individual.

**p<0.01; ***p<0.001; **** p<0.0001.

Figure 4: Specific targeting of CMML mutant HSCs by DAC and G9A/GLP inhibitor combination.

(A-B) Methylcellulose serial replating assays. (A) Number of colonies formed from 5000 $CD34^+CD38^-CD90^+$ (stem cells) and $CD34^+CD38^-CD90^-$ (progenitor cells) from patients (n=3) and controls (n=2) in methylcellulose after the different passages (P), in the presence or absence of DAC +/- UNC0638. (B) Effects of DAC and/or UNC0638 treatments on the

cumulative clonogenic potential ([number of colonies at P2/number of cells planted at P0] X number cells retrieved at P1) of stem (left panel) and progenitor (right panel) cells from patients, n=3. Mean +/- SEM. One-way ANOVA.

(C-E) Single cell liquid cultures. (C) Viability, (D) proliferation index and (E) cell number distribution per well of CMML patient stem cells (left panel) or progenitors (right panel) treated with either DAC (green), UNC0638 (blue) or both (red), or left non-treated (NT, black). N=3 patients. 48-96 wells counted per sample. Means +/- SEM. Two-way ANOVA Bonferroni's Multiple Comparison. *p<0.05; **p<0.01; ***p<0.001; ****p<0.0001.

(F) Evaluation of the VAF of identified mutated genes after 14 days of culture of $CD34^+$ cells in methylcellulose (left panel) or $CD34^+CD38^-CD90^+$ stem cells in liquid cultures (right panel), in the 4 conditions, as indicated. Percentage of cells with 0 (bright blue), 1 (pattern light blue or light grey), 2 (dark grey, grey pattern) or 3 (red) mutation. Mutations anf VAFs are indicated over each panel. Hz: heterozygous; Ho: homozygous.

Figure 5: Synergistic increase in TE expression in CMML patient $CD34^+$ cells after treatment with the combination.

(A-D) TE family expression in RNA-seq analysis from $CD34^+$ cells from 4 patients, after treatments for 4 days with either DAC, UNC0638 or DAC+UNC0368, or left untreated. (A) Volcano plots showing differential TE expression (pval<0.05) in the combination condition compared to NT cells. (B) TE family (SINE, LINE, ERV, DNA) repartition in differentially expressed TEs. Numbers indicate the percentage represented by each family. (C) Log2 Fold-change increase of differentially expressed TE classes in the DAC+UNC0638 condition vs. NT cells. One-way ANOVA Multiple Comparison. (D) Expression levels of TEs differentially

expressed in the DAC+UNC0638 vs. NT condition in each of the treatment conditions, sorted by TE class.

(E-G) Differential analysis of ERV loci in RNA-seq data from 7 patients after treatments with DAC, UNC0638 or both compared to the untreated condition (NT). (E, F) Venn diagrams of the differentially expressed ERV transcripts up- and down-regulated (pval<0.05, FC 1.5), in each of the treated condition vs. NT, as indicated. (G) Heatmap representation of the median expression level of differentially expressed ERVs in DAC, UNC0638, or DAC+UNC0638 vs. NT. Clusters with fewer than 20 nodes were excluded from the analysis.

(H) RT-qPCR analysis of mRNA expression of chosen ERV families in the different conditions. Ct were normalized to HPRT, Tubulin, Gus, PPIA and/or RPL32 and reported to the NT condition. One line represents a single patient. Means +/- SEM. One-way ANOVA Bonferroni's Multiple Comparison

* p<0.05; **p<0.01

Figure 6: Induction of transcriptional viral mimicry state by the combination of HMA and G9A/GLP inhibitors.

(A-E) RNA-seq-analysis of coding genes from $CD34^+$ cells of 7 patients after 4 days of treatment or not with DAC, UNC0638 or the combination. Each condition is compared to NT cells. (A) Venn diagram showing the number of common or unique DEGs in each condition. (B) Distribution of DEGs overexpressed (red) and downregulated (blue). (C, D) GSE analysis of pathways enriched in upregulated (C) and downregulated (D) genes in DAC+UNC0638 vs. NT. (E) Enrichment analysis of different IFN signatures by single sample gene-set variation analysis (ssGSVA).

(F) mRNA expression level of IFNB1, IRF9 and IRF7 measured by RT-qPCR. Ct were normalized by custom housekeeping genes (HPRT, Tubulin, Gus, PPIA and/or RPL32) and reported to the

NT condition. Each line represents a single patient. Means +/- SEM. One-way ANOVA Bonferroni's Multiple Comparison. * $p<0.05$.

Figure 7: IFN signaling is activated by the HMA and G9A/GLP inhibitor combination and is required for cell death.

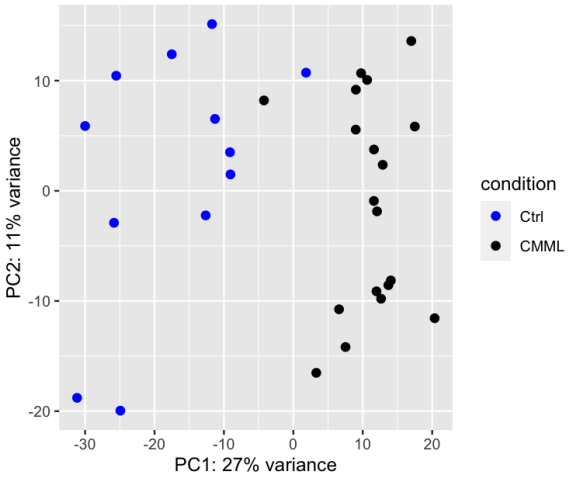
(A-E) Representative images and quantification of the mean overall immunofluorescence (IF) intensity of IRF3 (A), IRF7 (B), STAT1 (C), pSTAT1 (D) and dsRNAs (E) in patient $CD34^+$ cells after 4 days (A-D) or 2 days (E) of culture. Results are reported to the mean IF intensity in the NT condition. N= 2 patients. At least 50 cells counted per sample. Bars, 3 µm. Quantifications were performed using ImageJ software. One-way ANOVA Bonferroni's Multiple Comparison.

(F-H) Blocking IFN receptor inhibits the effect of the combination on colony formation. (F) Protocol scheme. Patient $CD34^+$ cells were incubated in liquid medium with the various treatments as in legend to Fig. 4A. The anti-IFNAR2 antibody (1µg/ml) was added on days 2 and 4 and the cells were seeded in methylcellulose on day 4. (G) Number of colonies at passages 1 and 2 and (H) cumulative clonogenic potential at passage 2. Results are reported to those obtained in the NT condition. Mean +/- SEM; n=6 patients. One-way ANOVA Bonferroni's Multiple Comparison.

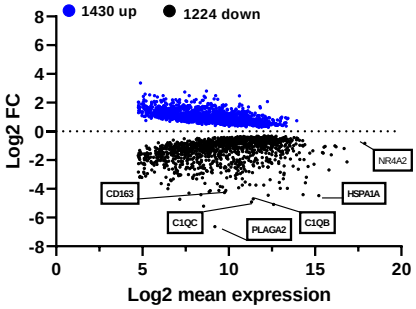
* $p<0.05$; ** $p<0.01$; *** $p<0.001$.

Hidaoui et al. Figure 1

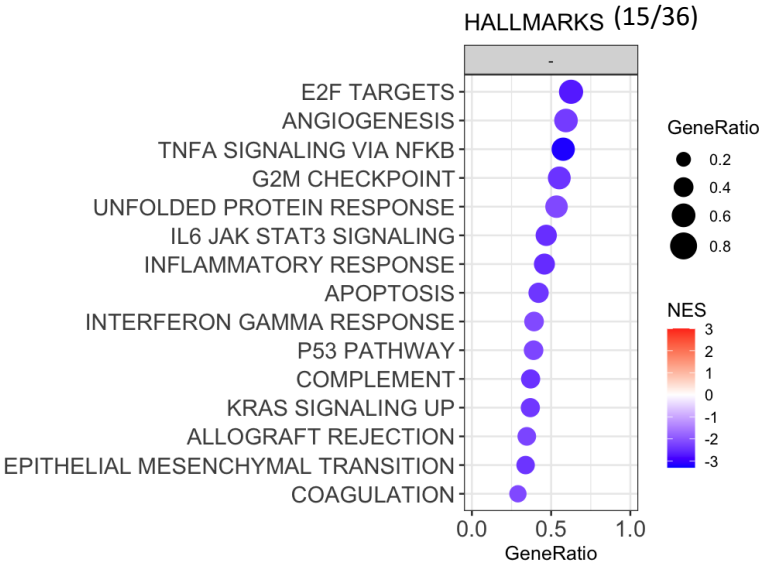
A.



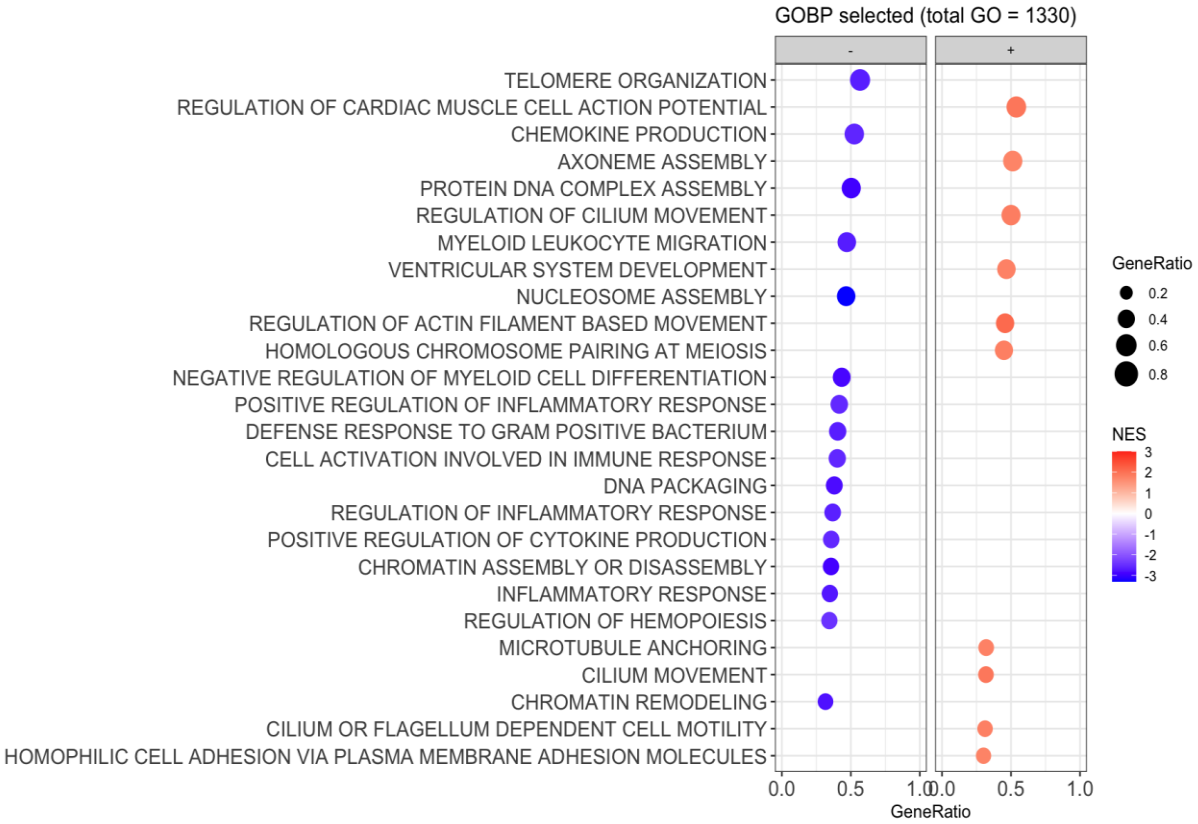
B.



C.

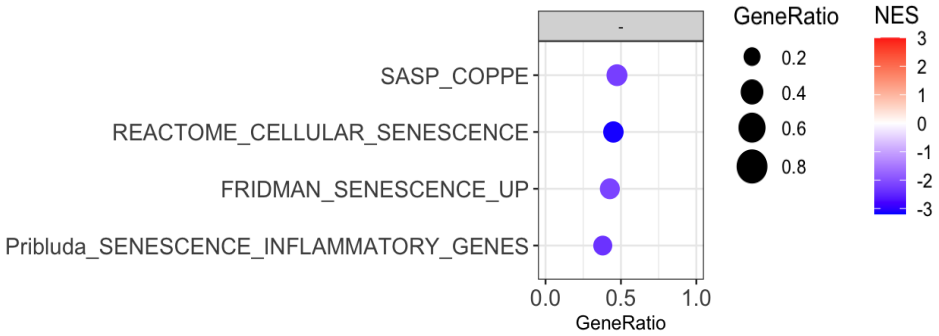


D.

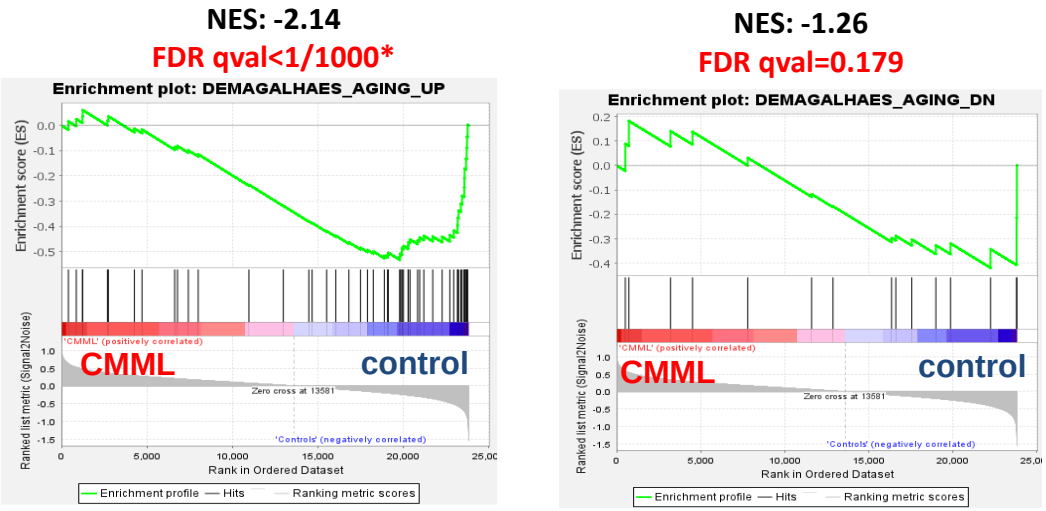


Hidaoui et al. Figure 1, continued

E.



F.



G.

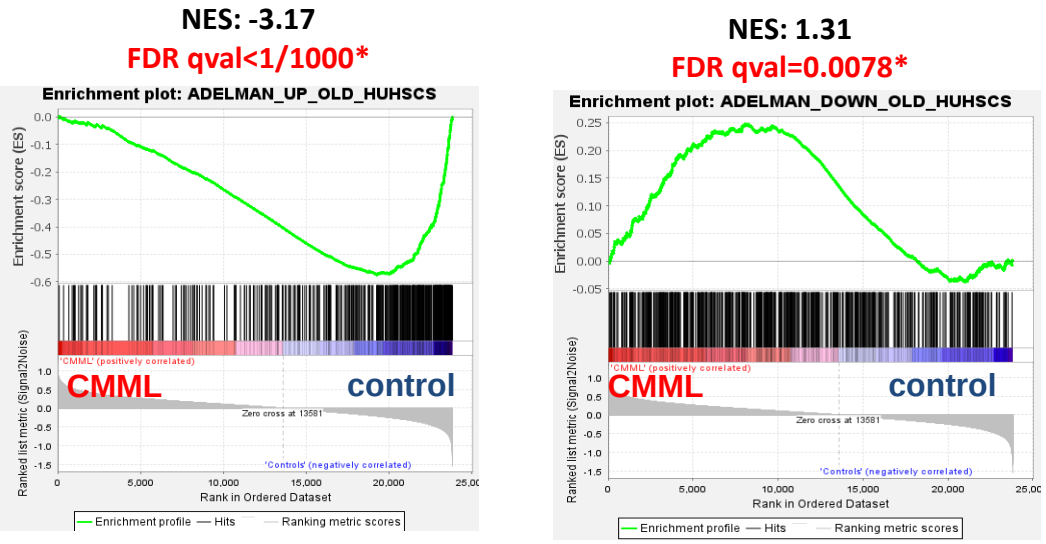


Figure 1: CMML HSPCs show a decrease in inflammatory response pathways

(A) Principal component analysis (PCA) plot of BM CD34$^+$ cells isolated from 19 CMML patients at diagnosis and 12 controls, based on RNA-seq regularized log transformed expression values.

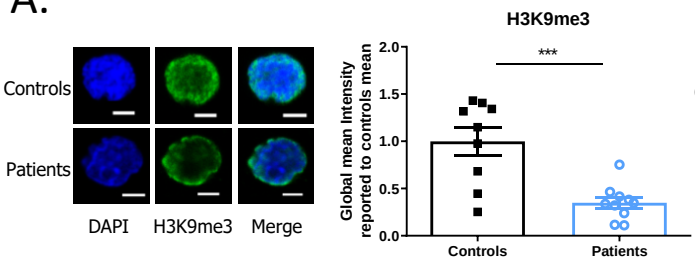
(B) MA-plots showing differentially expressed genes (padj <0.1), upregulated (blue dots) or downregulated (black dots) in CMML cells.

(C-E) Dot plots showing selected enriched (c) HALLMARK and (D) GO and (E) senescence gene sets, upregulated (positive NES, red dots) and downregulated (negative NES, blue dots) in CMML cells. Gene sets are ranked by their NES and gene ratio scores.

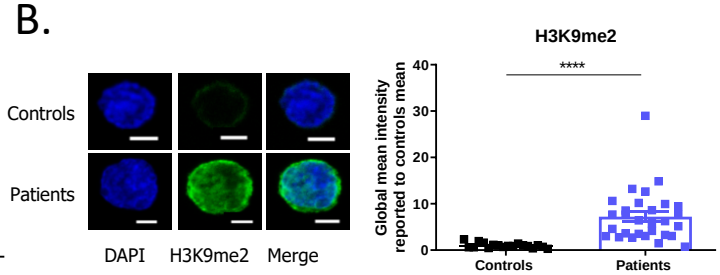
(F, G) Enrichment plots for published gene sets upregulated or downregulated with age. FDR, false discover rate; NES, normalized enrichment score. * FDR <0.05

Hidaoui et al. Figure 2

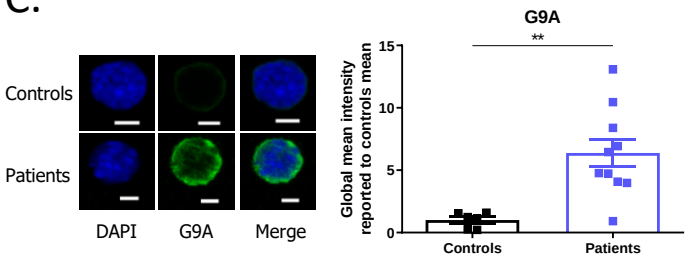
A.



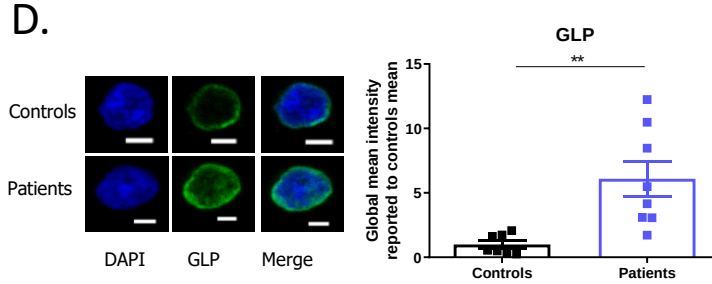
B.



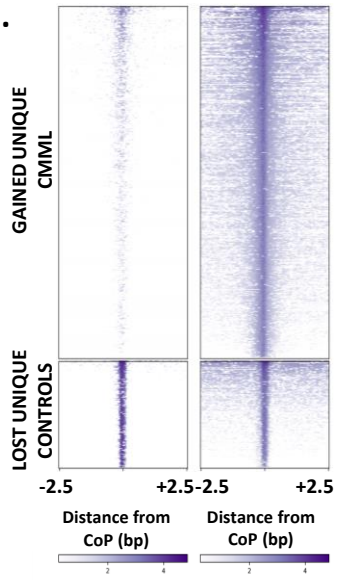
C.



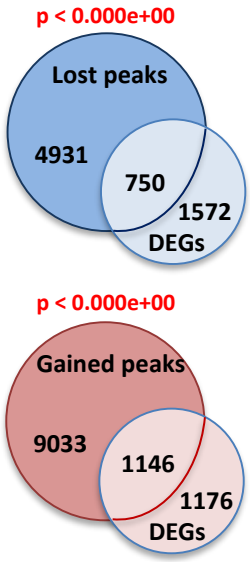
D.



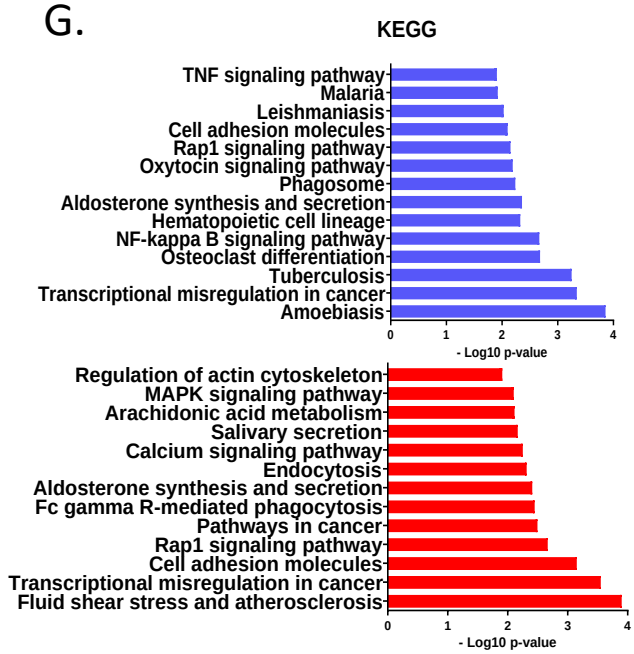
E.



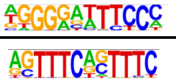
F.



G.

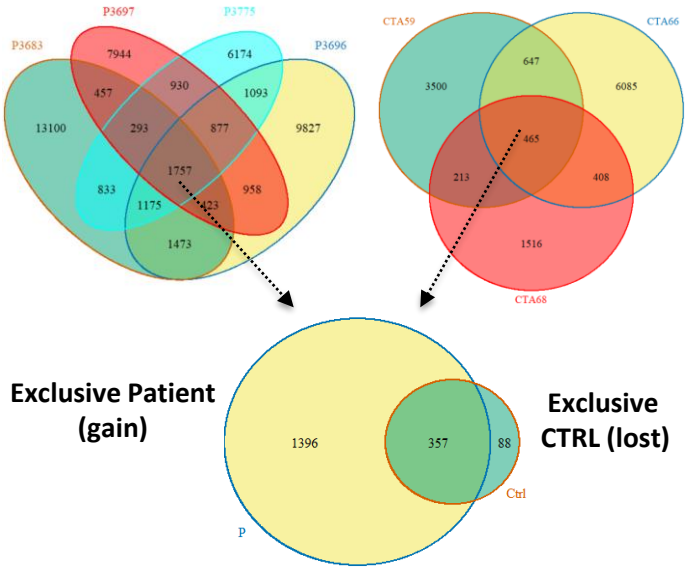


H.

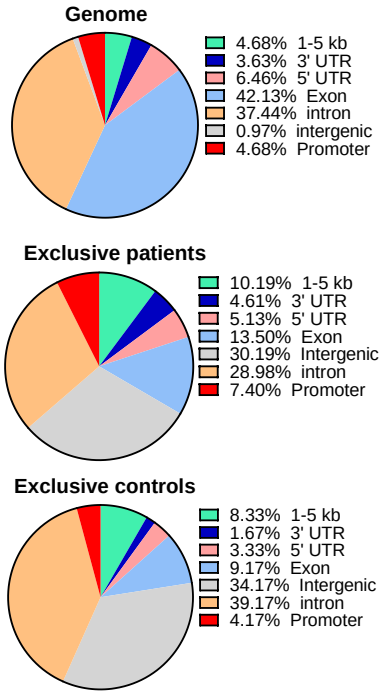
Motif	Transcription factor	Lost	Gained
	PU1-IRF	1^{e-53}	1^{e-5}
	IRF3	1^{e-27}	1^{e-5}
	IRF1	1^{e-25}	1^{e-7}
	NFkB-p65-Rel	1^{e-11}	ND
	NFkB-p65	1^{e-11}	ND
	ISRE	1^{e-10}	1^{e-4}
	NFkB-p50,p52	1^{e-4}	ND

Hidaoui et al. Figure 2, continued

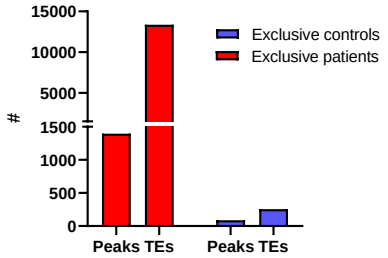
I.



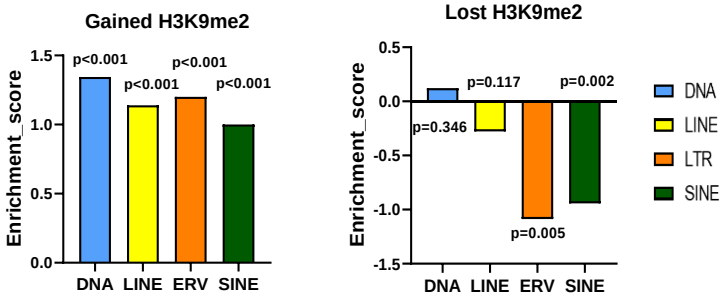
J.



K.



L.



M.

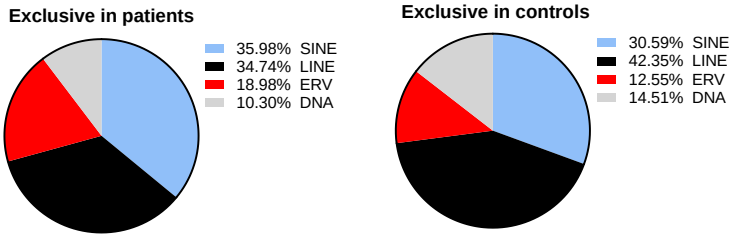


Figure 2: Chromatin is remodeled in CMML HSPCs

(A-D) Immunostaining for heterochromatin marks H3K9me3 (A); H3K9me2 (B), and factors G9A (C) and GLP (D) in $CD34^+$ cells from CMML patients and controls. Representative images and quantification of mean global immunofluorescence intensity using ImageJ software are shown. Means +/- SEM from 2-3 independent experiments normalized to the mean intensity of controls. Each dot represents cells from a single individual, with at least 50 cells counted per sample. Bars, 3 µm. t test, **P<0.01; ***p<0.001; ****, P<0.0001.

(E) Heat map of the differentially chromatin accessible regions in CMML patient (n=4) and control (n=3) $CD34^+$ cells +/- 2.5 kb from the center of the peak (CoP).

(F) Venn diagrams depicting the overlap of differentially expressed genes (DEGs) in CMML HSPCs with genes assigned to gained or lost ATAC peaks (-100+/-25 kb from TSS). Hypergeometric test.

(G) KEGG gene set analysis of deregulated genes that also exhibit gained or loss chromatin accessibility in CMML patient cells.

(H) Motifs and p values for transcription factor motif enrichment in ATAC-seq peaks unique to patients (gained) and controls (lost) using HOMER. ND: non detected.

(I) Venn diagram of peak retrieved from H3K9me2 CUT&Tag peak calling on $CD34^+$ from 4 patients (P) and 3 controls (CTA).

(J) Repartition of the CUT&Tag peaks found exclusively in patients and controls compared with hg19 reference genome annotation repartition (upper panel).

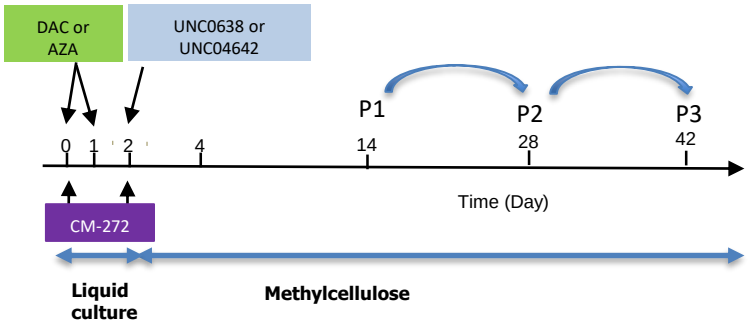
(K) Numbers of peaks found exclusively in the 4 patients (red) and 4 controls (blue) and numbers of TEs annotated in these peaks.

(L) Enrichment or depletion of the different classes of TEs in gained or lost H3K9me2-CUT&Tag peaks. The Y-axis representing the enrichment score was calculated as the log2 fold change of the number of specific peaks overlapping with TEs found in patient gained or lost peaks over the median number of TEs found in 1000 times the same number of randomly selected peaks overlapping with TEs. Positive log2 fold change = enrichment, and negative log2 fold change = depletion.

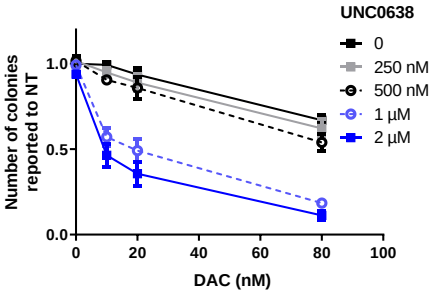
(M) Repartition of the different TE classes in CUT&Tag peaks found exclusively in patients and controls.

Hidaoui et al. Figure 3

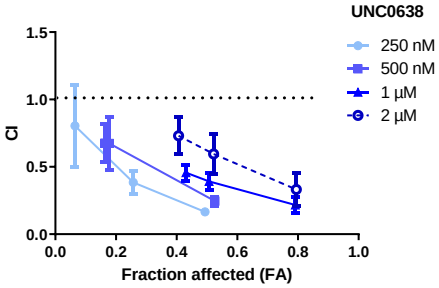
A.



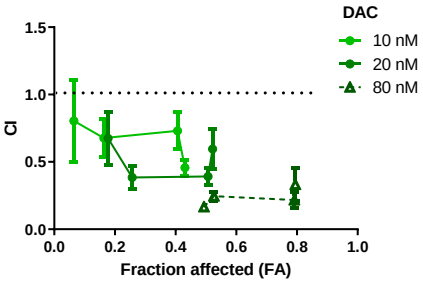
B.



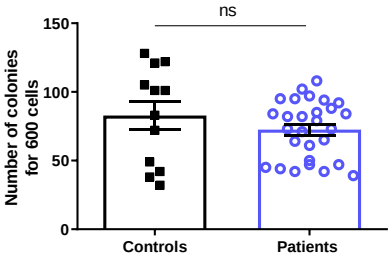
C.



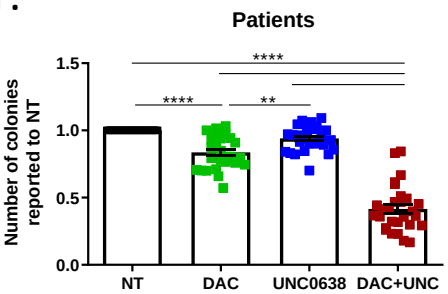
D.



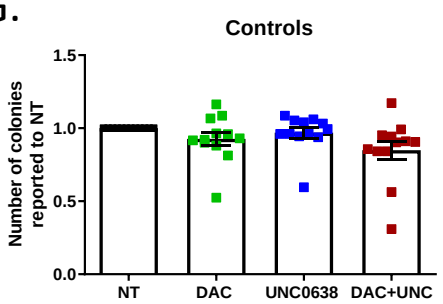
E.



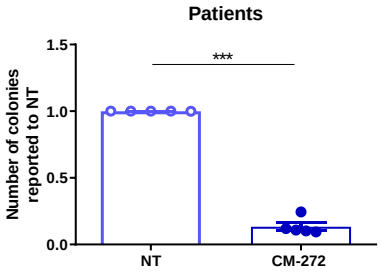
F.



G.



H.



I.

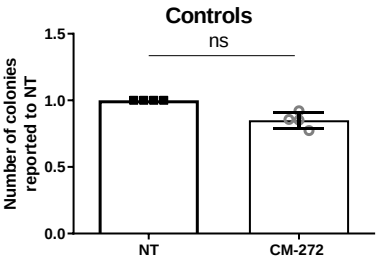


Figure 3: the HMA and G9A/GLP inhibitor combination reduces CMML HSPC clonogenicity

(A) Treatment protocol of $CD34^+$ cells. DAC: decitabine; AZA: 5-azacytidine; UNC0638 and UNC0642: G9A/GLP inhibitors; CM-272: dual DNMT1 and G9A/GLP inhibitor. Colonies are counted after 14 days of culture and replated.

(B) Dose-response effect of DAC (X-axis) and UNC0638 (color scale) on CMML $CD34^+$ colony formation. Mean +/- SEM of duplicate wells from 3 patients.

(C-D) Chou-Talalay model (27) of the effects of combining different doses of DAC and UNC0638. The fraction affected (fa) corresponds to the proportion of colonies eliminated by the treatments. The combination index (CI) is calculated by CompuSyn software (ComboSyn. Inc.) where CI<1, CI=1, CI>1 indicates a synergistic, additive and antagonistic effect respectively. Mean+/- SEM of duplicate wells from 3 patients.

(E) Number of colonies formed from 600 CMML (n=28) and control (n=12) $CD34^+$ cells in the absence of treatment. One dot represents a single individual.

(F, G) Number of colonies formed by CD34+ cells from patients (F) or controls (G) after treatment with 10 nM DAC and 1µM UNC0638 alone or in combination. NT: non-treated. One dot represents a single individual; controls, n=12; patients, n=28. Results are normalized to the number of colonies in the NT condition. Means +/- SEM. 2 or 3 wells per sample. One-way ANOVA Bonferroni's Multiple Comparison.

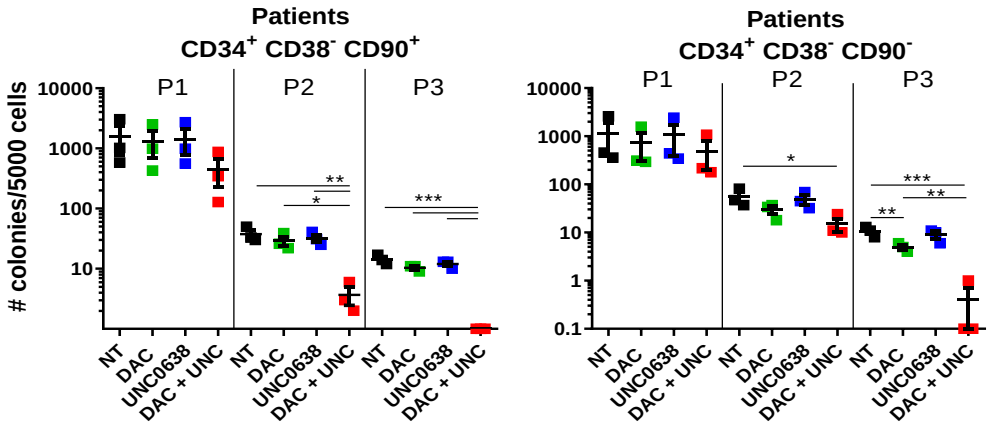
(H, I) Number of colonies formed by $CD34^+$ cells from patients, n=5 (L) or controls, n=4, (M) after treatment with or without (NT) 200 nM CM-272. Results are normalized to the number of colonies in the NT condition. Means +/- SEM. t.test.

(E-I) One dot represents a single individual.

**p<0.01; ***p<0.001; **** p<0.0001.

Hidaoui et al. Figure 4

A.

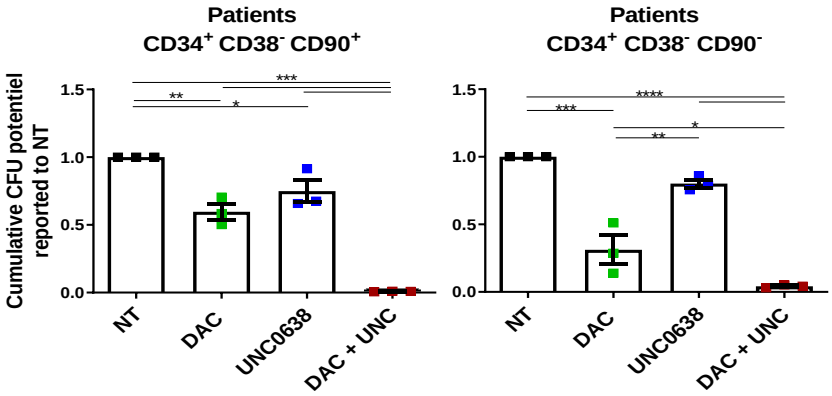


B.

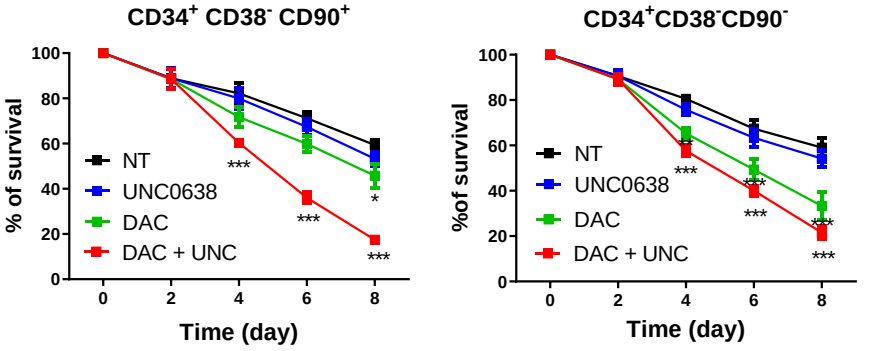
Cumulative CFU potentiel

Taking account of colonies size

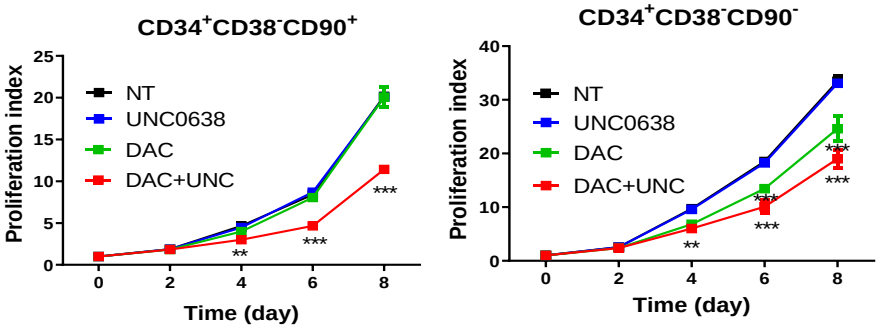
$$\#cells\ at\ P1 \times \frac{\#CFU}{5000}$$



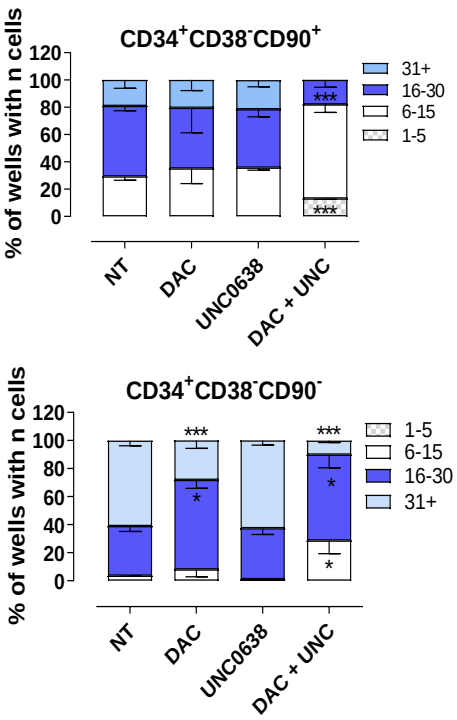
C.



D.



E.

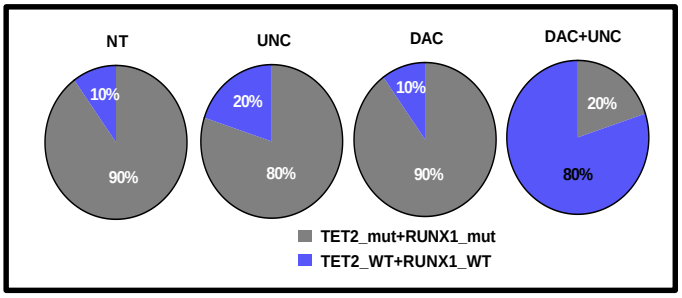


Hidaoui et al. Figure 4, continued

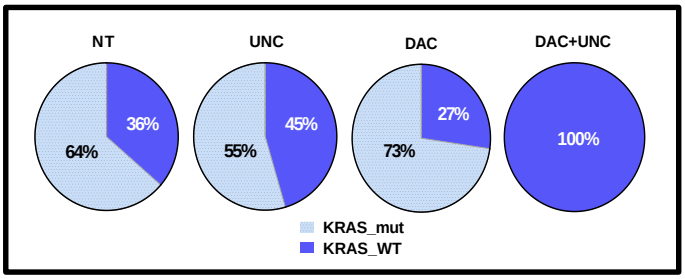
F.

CD34⁺ methylcellulose colonies

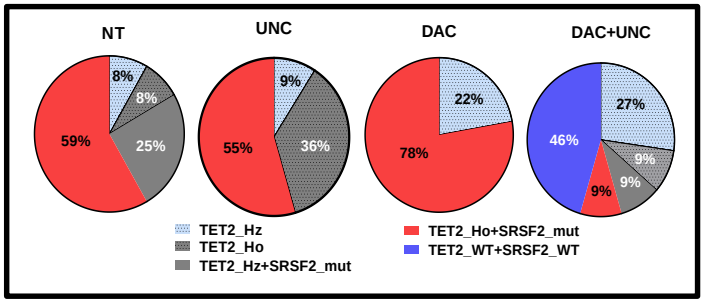
Patient #3440
TET2-K1338X (54%) RUNX1-S114L (61%) SRSF295H (17%)



Patient #3562
KRAS-G12D (30%) TET2-S407fs (9%) RUNX1-Y285F (11%)

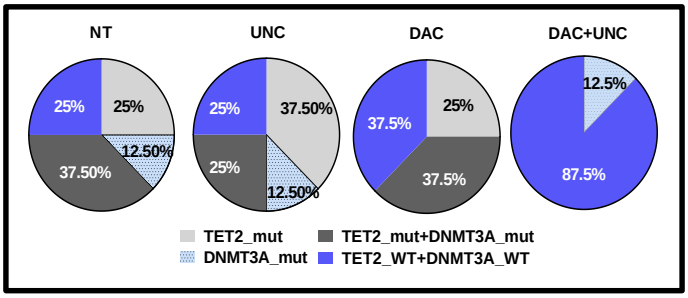


Patient #3564
TET2-L1229P (41%) SRSF2-P95R (37%)

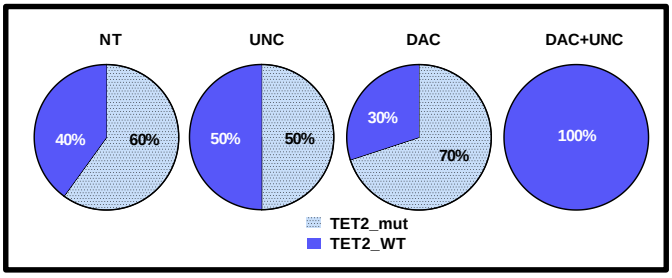


CD34⁺C38⁻CD90⁺ single cell liquid cultures

Patient #3480
DNMT3A-DEL-G (34%) TET2-H1380R (37%)



Patient #3459
TET2-H1480Q Hz (44%) ASXL1(2%)



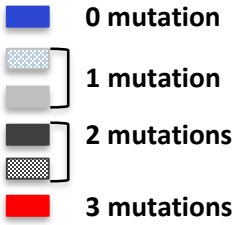


Figure 4: Specific targeting of CMML mutant HSCs by DAC and G9A/GLP inhibitor combination

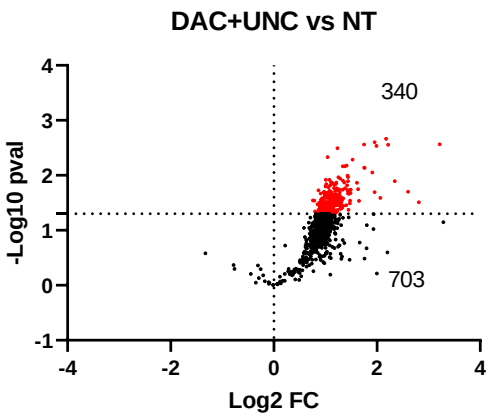
(A-B) Methylcellulose serial replating assays. (A) Number of colonies formed from 5000 CD34⁺CD38⁻CD90⁺ (stem cells) and CD34⁺CD38⁻CD90⁻ (progenitor cells) from patients (n=3) and controls (n=2) in methylcellulose after the different passages (P), in the presence or absence of DAC +/- UNC0638. (B) Effects of DAC and/or UNC0638 treatments on the cumulative clonogenic potential ([number of colonies at P2/number of cells planted at P0] X number cells retrieved at P1) of stem (left panel) and progenitor (right panel) cells from patients, n=3. Mean +/- SEM. One-way ANOVA.

(C-E) Single cell liquid cultures. (C) Viability, (D) proliferation index and (E) cell number distribution per well of CMML patient stem cells (left panel) or progenitors (right panel) treated with either DAC (green), UNC0638 (blue) or both (red), or left non-treated (NT, black). N=3 patients. 48-96 wells counted per sample. Means +/- SEM. Two-way ANOVA Bonferroni's Multiple Comparison. *p<0.05; **p<0.01;***p<0.001;****p<0.0001.

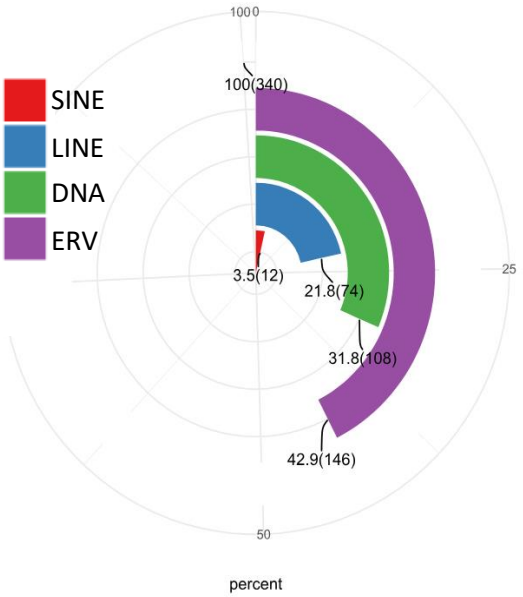
(F) Evaluation of the VAF of identified mutated genes after 14 days of culture of CD34⁺ cells in methylcellulose (left panel) or CD34⁺CD38⁻CD90⁺ stem cells in liquid cultures (right panel), in the 4 conditions, as indicated. Percentage of cells with 0 (bright blue), 1 (light blue pattern, light grey), 2 (dark grey, grey pattern) or 3 (red)mutation. Mutations and VAFs are indicated over each panel. Hz: heterozygous; Ho: homozygous.

Hidaoui et al. Figure 5

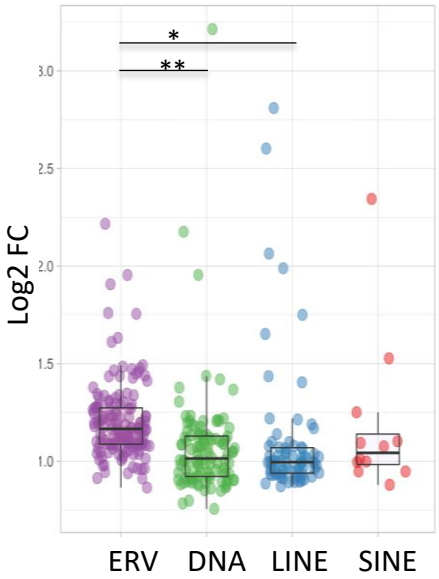
A.



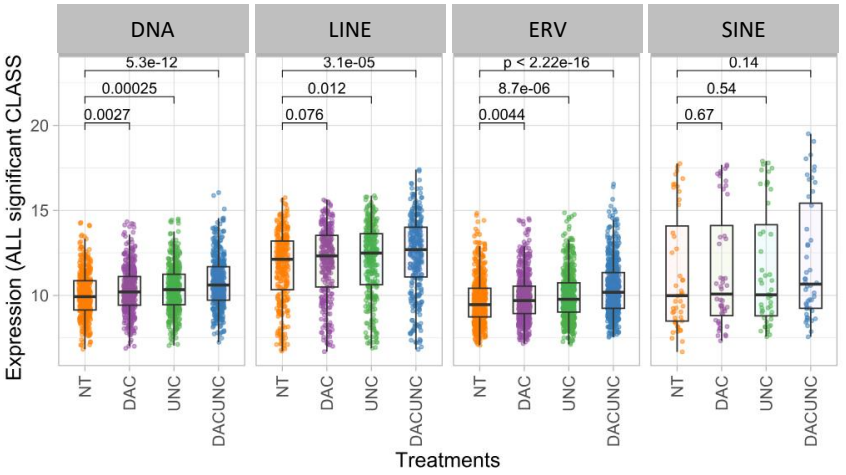
B.



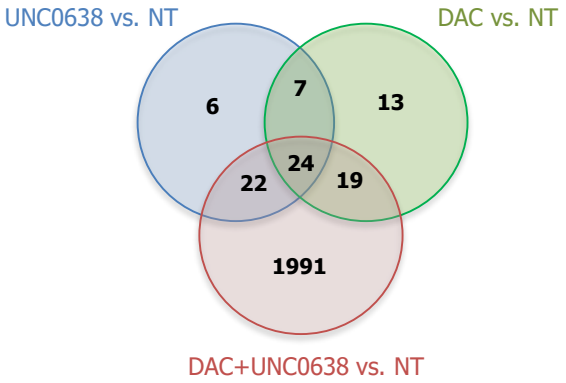
C.



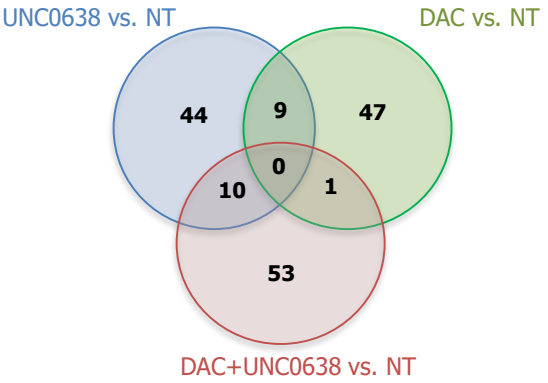
D.



E.

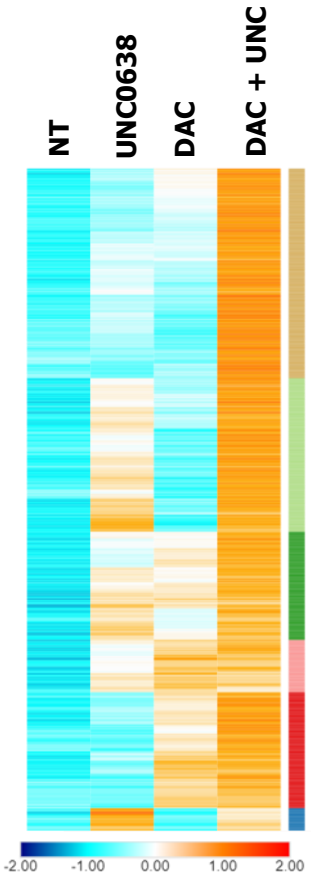


F.



Hidaoui et al. Figure 5 continued

G.



H.

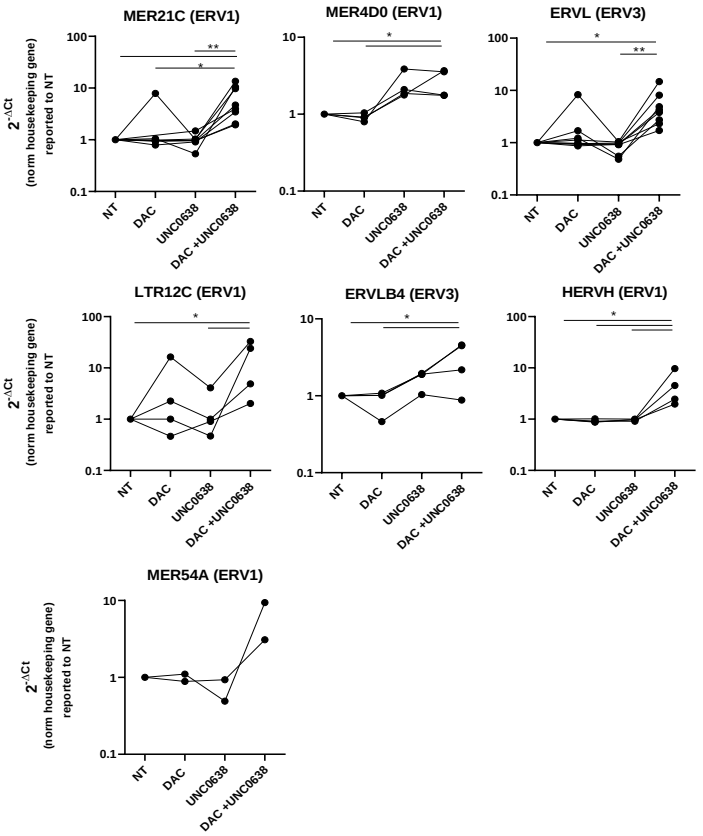


Figure 5: Synergistic increase in TE expression in CMML patient CD34⁺ cells after treatment with the combination

(A-D) TE family expression in RNA-seq analysis from CD34⁺ cells from 4 patients, after treatments for 4 days with either DAC, UNC0638 or DAC+UNC0368, or left untreated. (A) Volcano plots showing differential TE expression (pval<0.05) in the combination condition compared to NT cells. (B) TE family (SINE, LINE, ERV, DNA) repartition in differentially expressed TEs. Numbers indicate the percentage represented by each family. (C) Log2 Fold-change increase of differentially expressed TE classes in the DAC+UNC0638 condition vs. NT cells. One-way ANOVA Multiple Comparison. (D) Expression levels of TEs differentially expressed in the DAC+UNC0638 vs. NT condition in each of the treatment conditions, sorted by TE class.

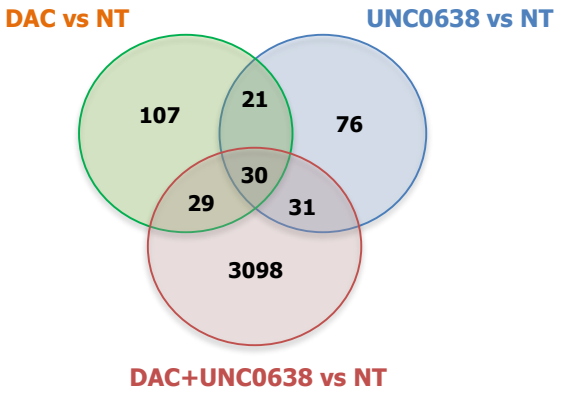
(E-G) Differential analysis of ERV loci in RNA-seq data from 7 patients after treatments with DAC, UNC0638 or both compared to the untreated condition (NT). (E, F) Venn diagrams of the differentially expressed ERV transcripts up- and down-regulated (pval<0.05, FC 1.5), in each of the treated condition vs. NT, as indicated. (G) Heatmap representation of the median expression level of differentially expressed ERVs in DAC, UNC0638, or DAC+UNC0638 vs. NT. Clusters with fewer than 20 nodes were excluded from the analysis.

(H) RT-qPCR analysis of mRNA expression of chosen ERV families in the different conditions. Ct were normalized to HPRT, Tubulin, Gus, PPIA and/or RPL32 and reported to the NT condition. One line represents a single patient. Means +/- SEM. One-way ANOVA Bonferroni's Multiple Comparison

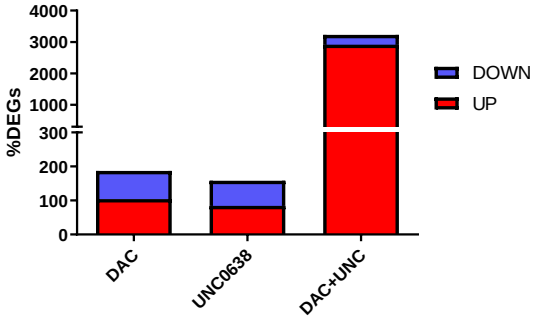
* p<0.05; **p<0.01

Hidaoui et al. Figure 6

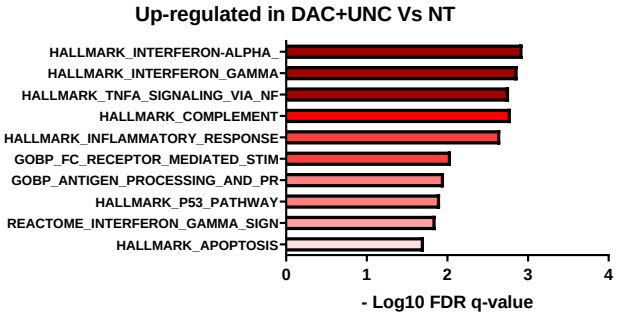
A.



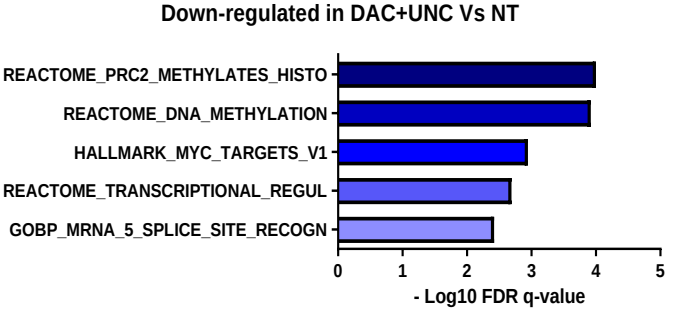
B.



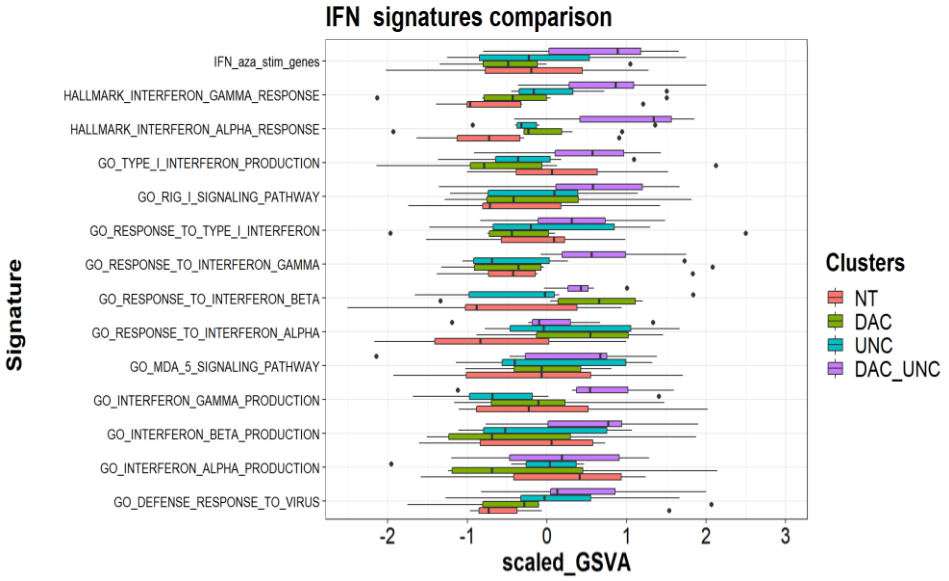
C.



D.



E.



F.

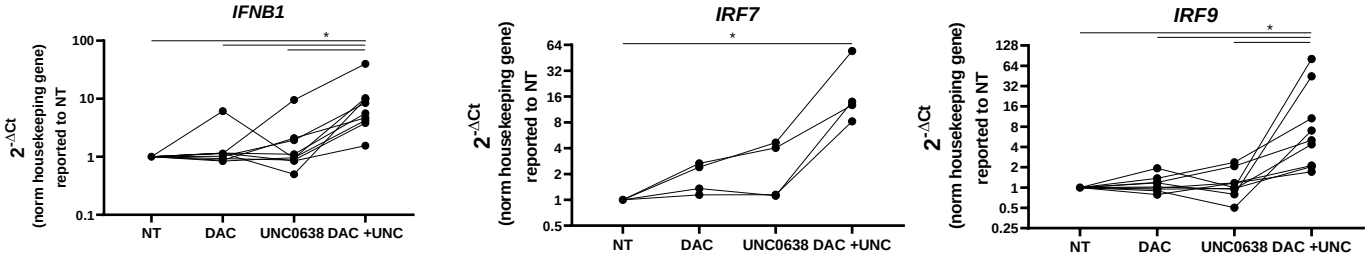


Figure 6: Induction of transcriptional viral mimicry state by the combination of HMA and G9A/GLP inhibitors.

(A-E) RNA-seq-analysis of coding genes from $CD34^+$ cells of 7 patients after 4 days of treatment or not with DAC, UNC0638 or the combination. Each condition is compared to NT cells. (A) Venn diagram showing the number of common or unique DEGs in each condition. (B) Distribution of DEGs overexpressed (red) and downregulated (blue). (C, D) GSE analysis of pathways enriched in upregulated (C) and downregulated (D) genes in DAC+UNC0638 vs. NT. (E) Enrichment analysis of different IFN signatures by single sample gene-set variation analysis (ssGSVA).

(F) mRNA expression level of IFNB1, IRF9 and IRF7 measured by RT-qPCR. Ct were normalized by custom housekeeping genes (HPRT, Tubulin, Gus, PPIA and/or RPL32) and reported to the NT condition. Each line represents a single patient. Means +/- SEM. One-way ANOVA Bonferroni's Multiple Comparison. * $p<0.05$.

Hidaoui et al. Figure 7

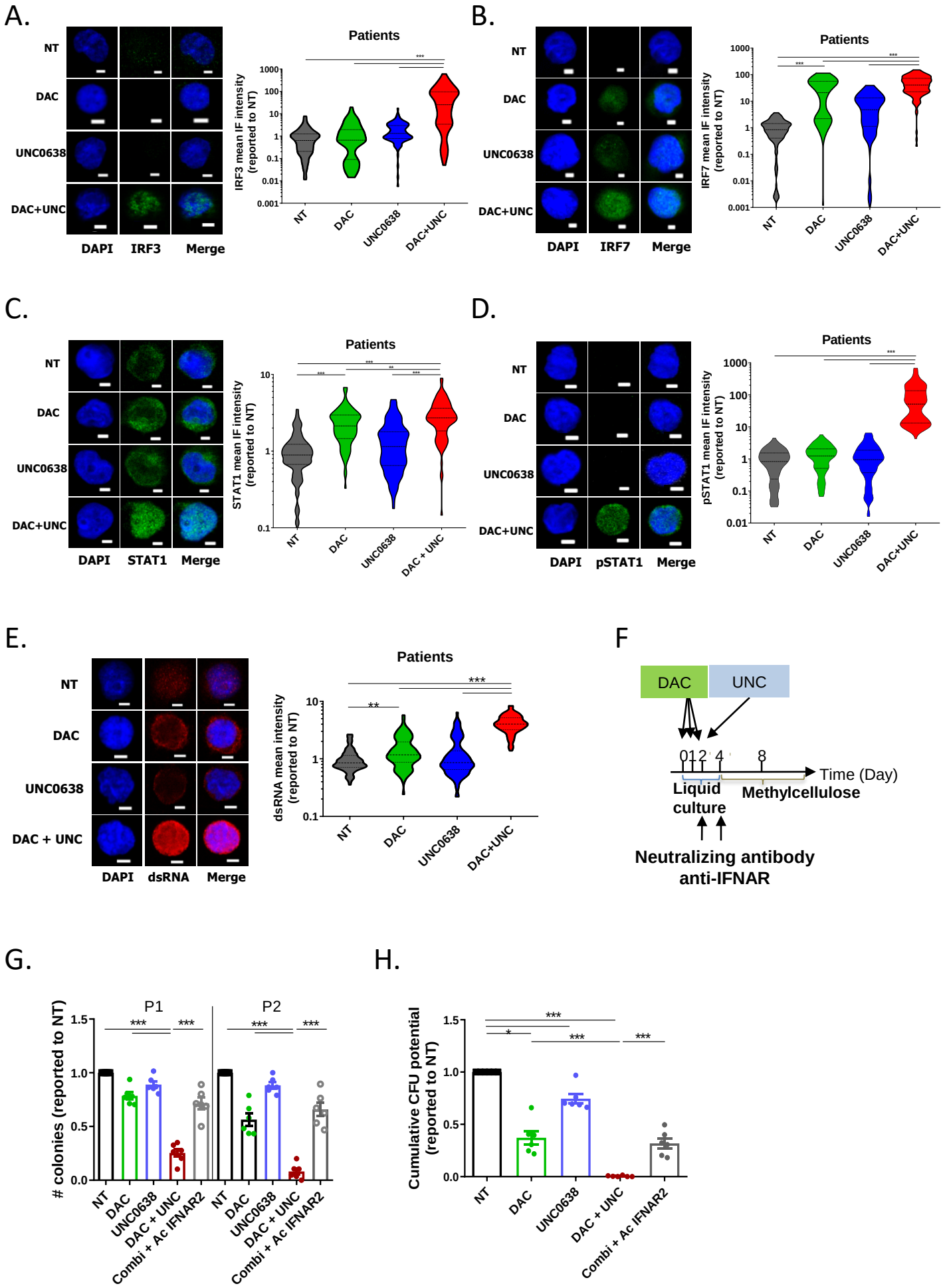


Figure 7: IFN signaling is activated by the HMA and G9A/GLP inhibitor combination and required for cell death.

(A-E) Representative images and quantification of the mean overall immunofluorescence (IF) intensity of IRF3 (A), IRF7 (B), STAT1 (C), pSTAT1 (D) and dsRNAs (E) in patient $CD34^+$ cells after 4 days (A-D) or 2 days (E) of culture. Results are reported to the mean IF intensity in the NT condition. N= 2 patients. At least 50 cells counted per sample. Bars, 3 µm. Quantifications were performed using ImageJ software. One-way ANOVA Bonferroni's Multiple Comparison.

(F-H) Blocking IFN receptor inhibits the effect of the combination on colony formation. (F) Protocol scheme. Patient $CD34^+$ cells were incubated in liquid medium with the various treatments as in legend to Fig. 4A. The anti-IFNAR2 antibody (1µg/ml) was added on days 2 and 4 and the cells were seeded in methylcellulose on day 4. (G) Number of colonies at passages 1 and 2 and (H) cumulative clonogenic potential at passage 2. Results are reported to those obtained in the NT condition. Mean +/- SEM. ; n=6 patients. One-way ANOVA Bonferroni's Multiple Comparison.

* $p<0.05$; ** $p<0.01$; *** $p<0.001$.

Targeting Heterochromatin Eliminates Malignant Hematopoietic Stem and Progenitor Cells in Chronic Myelomonocytic Leukemia Through Reactivation of Retroelements and Innate Immune pathways

Donia Hidaoui[1,2,3,], Rabie Chelbie[1,2,3,4], Mathieu Bohm[1,2,3,4], Audrey Porquet[1,2,3], Aikaterini Polyzou[5], Vincent Alcazer[6], Stéphane Depil[6], Aygun Imanci[1,2,3,4] Margot Morabito[1,2,3,], Aline Renneville[1,2,3,7], Dorothée selimoglu-Buet[1,2,3,], Sylvain Thépot[8], Raphael Itzykson[9], Lucie Laplane[1,2,3,10], Nathalie Droin[1,2,3,7], Eirini Trompouki[5], Emilie Elvira-Matelot[1,2,3], Eric Solary[1,2,3,11] and Françoise Porteu[1,2,3,*]

Supplementary materials

Supplementary figures 1 to 7

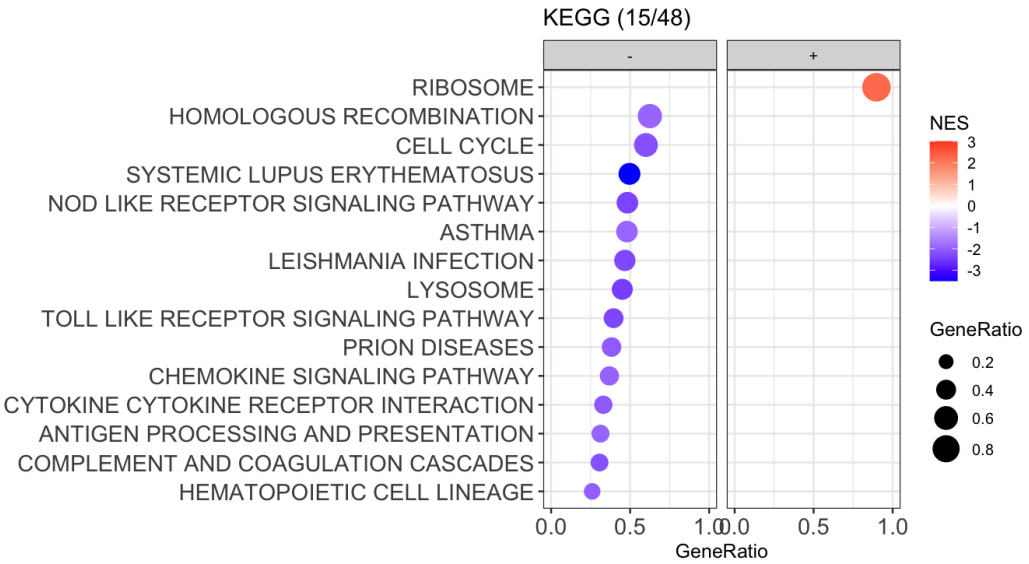
Supplementary Tables 1, 5, 9, 10

Supplementary methods

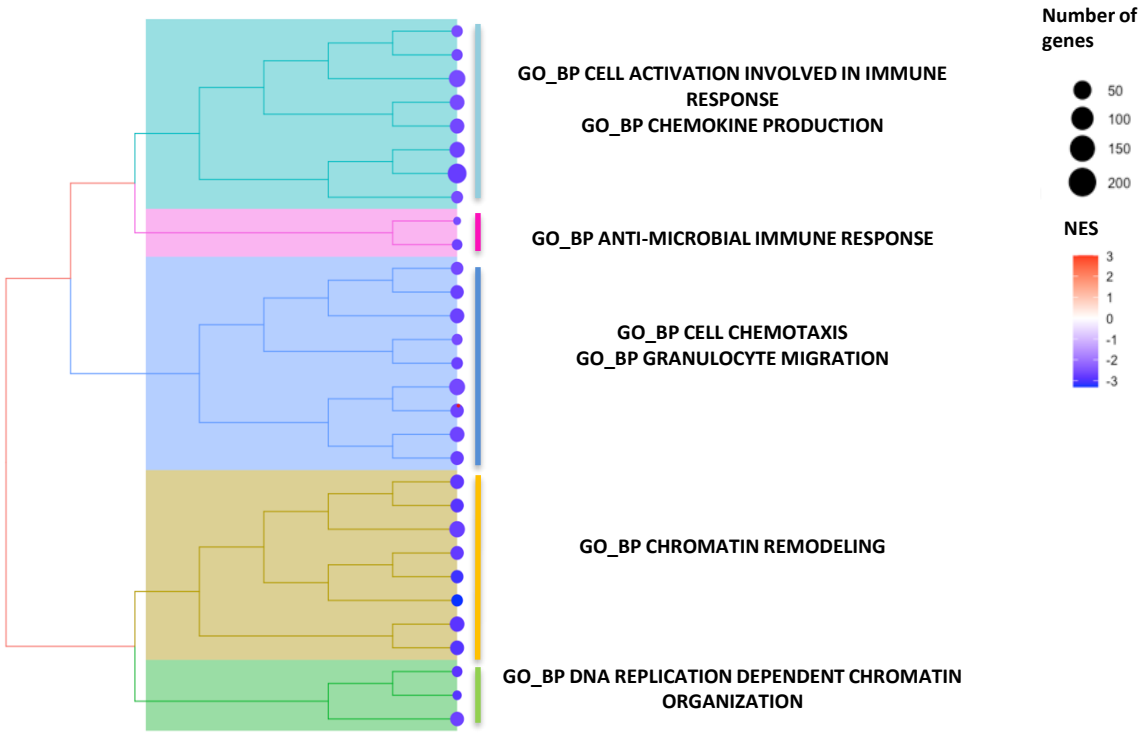
Supplementary references

Hidaoui et al. Figure S1, related to Figure 1

A.



B.



C.

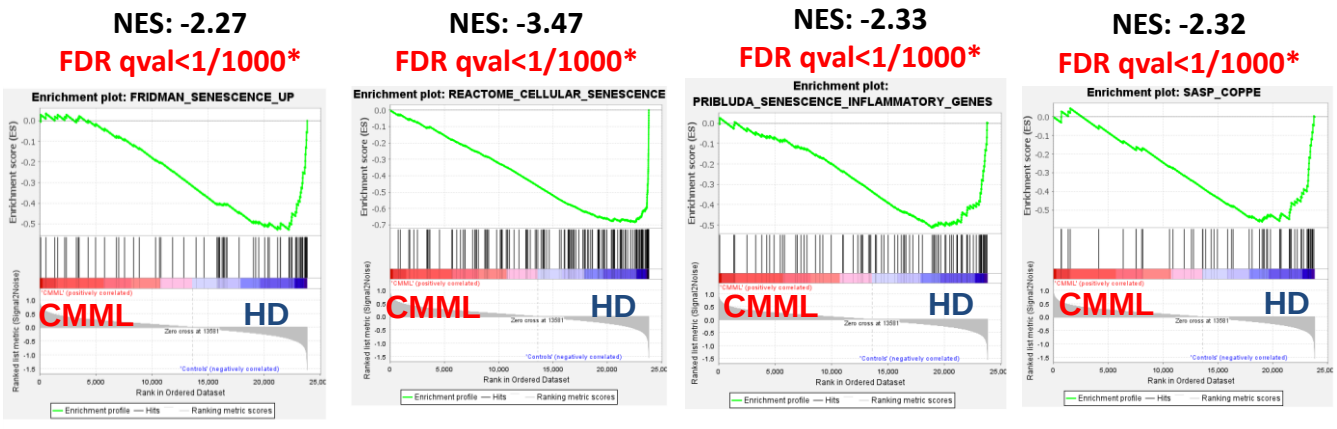


Figure S1, related to Figure 1: Gene score enrichment analysis of differentially expressed genes in CD34$^+$ cells from CMML patients vs. controls.

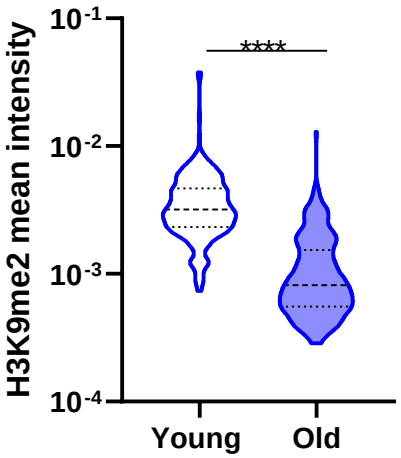
(A) Dot plots showing top 15 enriched Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways (FDR<0.05) upregulated (positive NES, red dots) and downregulated (negative NES, blue dots) in CMML cells compared to controls. Gene sets are ranked by their NES and gene ratio scores.

(B) Top 30 enriched Gene Ontology Biological Processes (GO-BP) organized as a tree plot regrouping redundant pathways. NES scale and gene ratio scores are shown on the right.

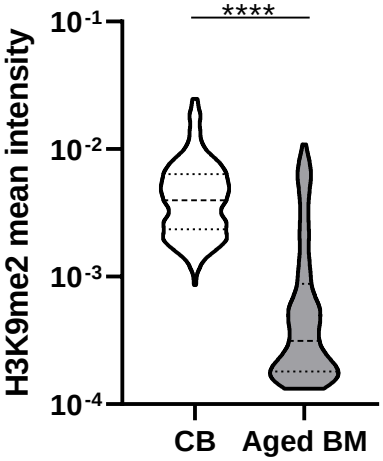
(C) Enrichment plots for published senescence and senescence-associated secretory phenotype (SASP) gene sets. FDR, false discover rate; NES, normalized enrichment score. *, FDR<0.05.

Hidaoui et al. Figure S2, related to Figure 2

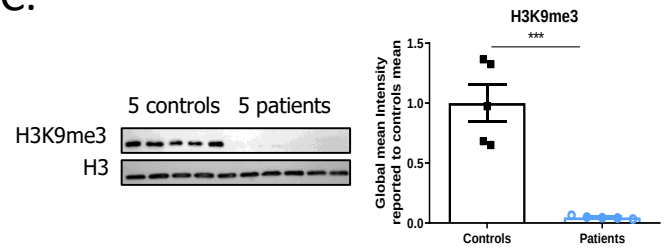
A.



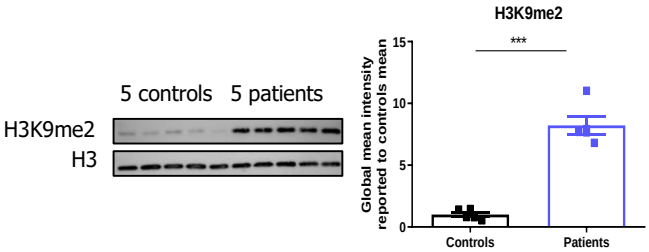
B.



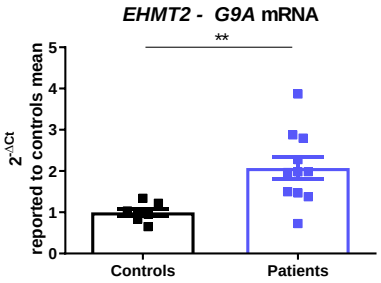
C.



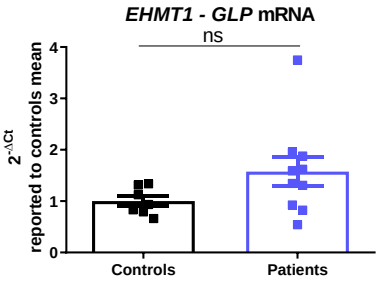
D.



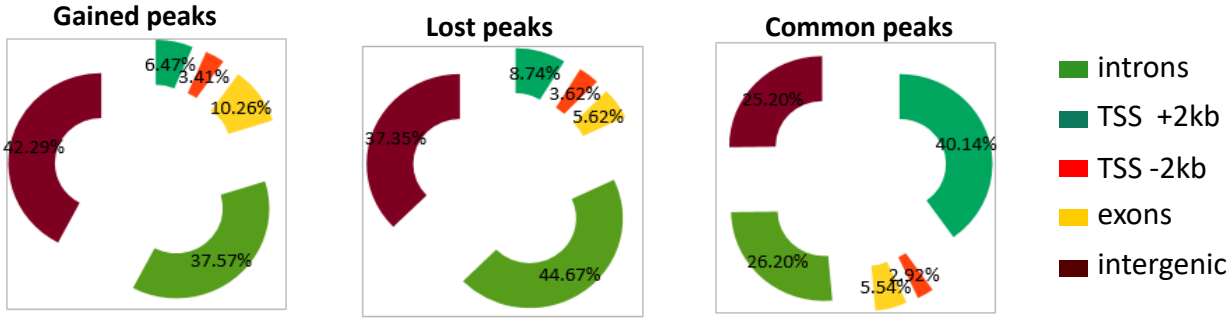
E.



F.

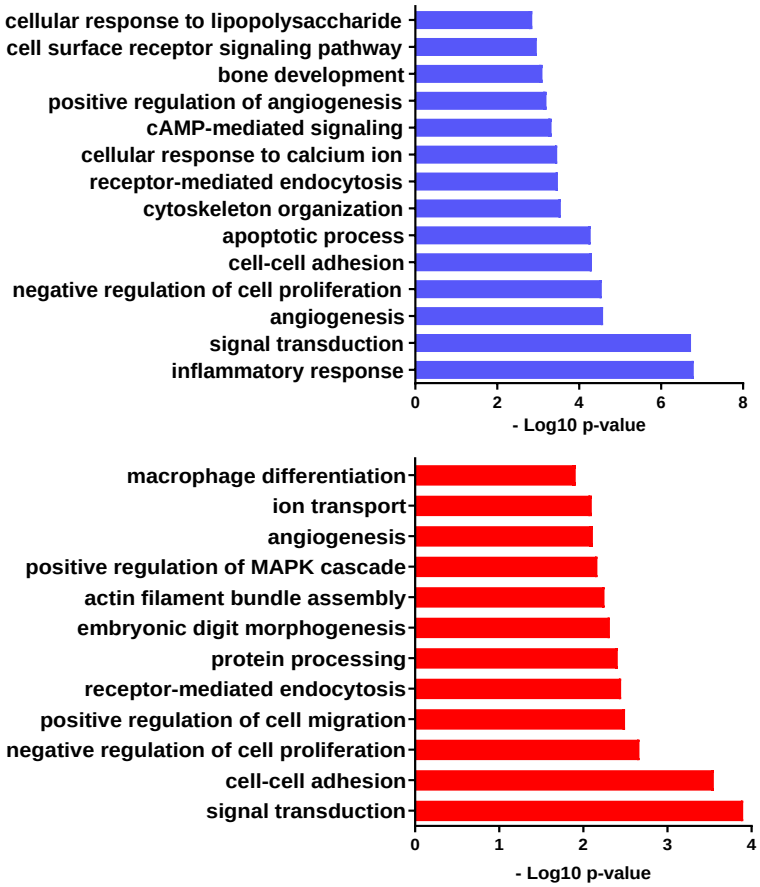


G.

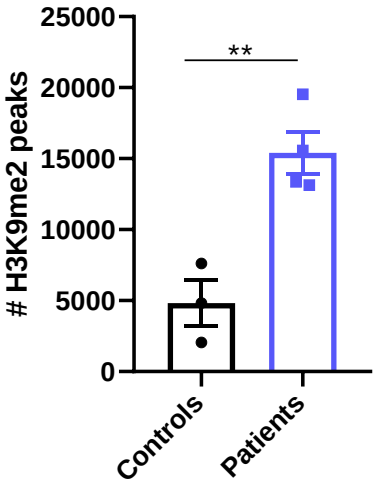


Hidaoui et al. Figure S2, related to Figure 2, continued

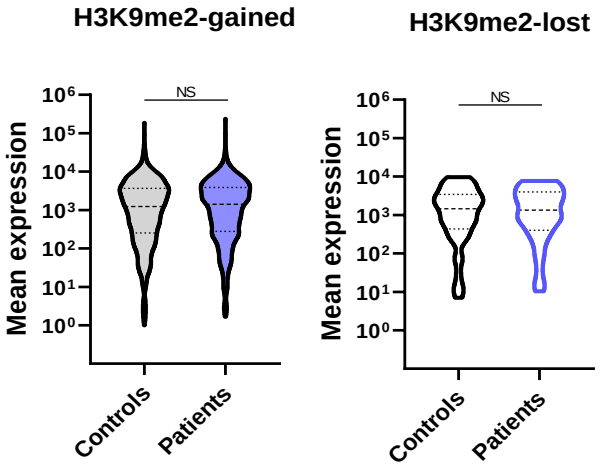
H.



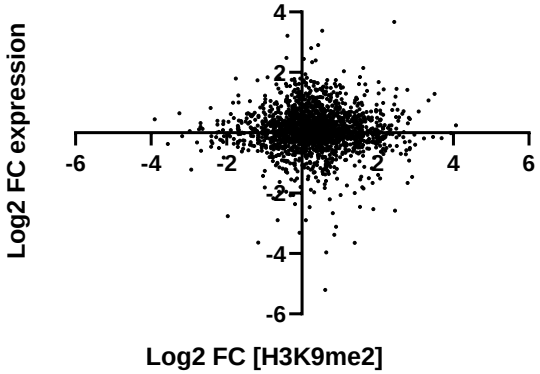
I.



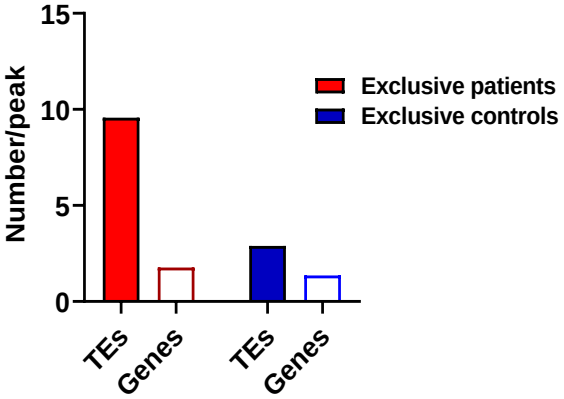
J.



K.



L.



Hidaoui et al. Figure S2, related to Figure 2, continued

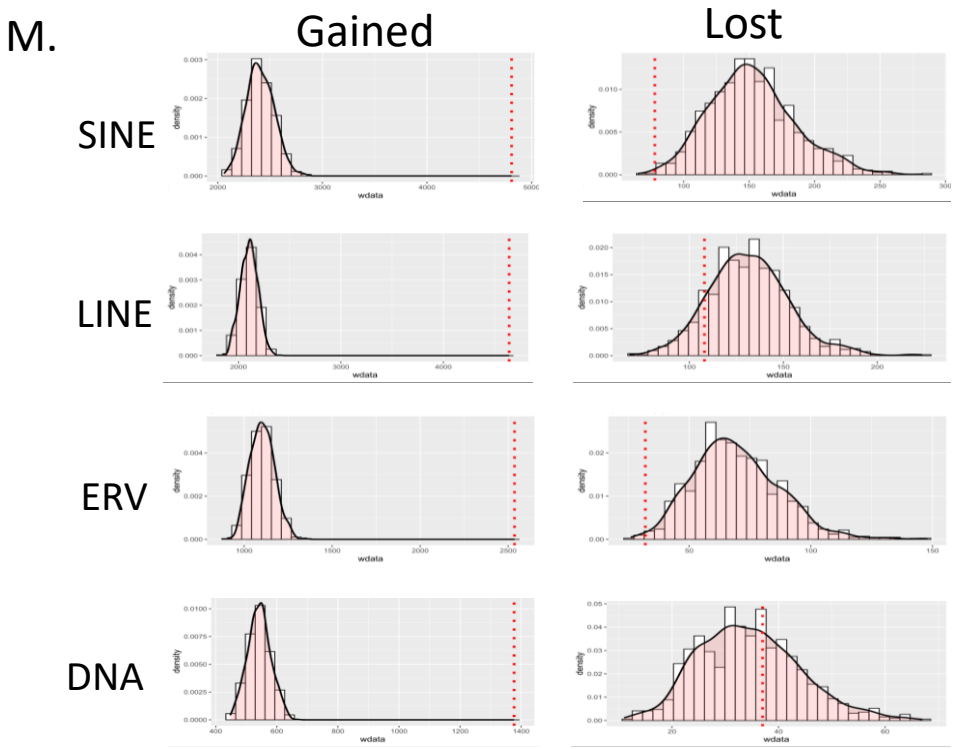
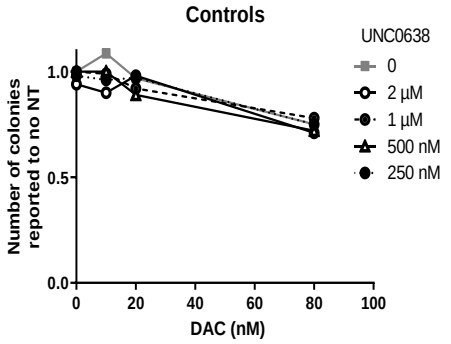


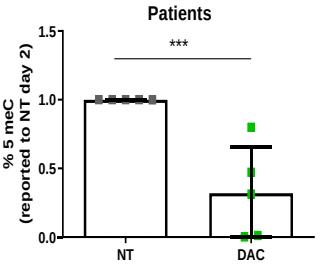
Figure S2, related to figure 2: Chromatin is remodeled in CMML HSPCs.
(A, B) H3K9me2 immunostaining in (A) Lin$^-$Sca$^+$Kit$^+$Flk2$^-$CD34$^-$ HSCs isolated from young (2 mo-old, n=3) and aged (20-24 mo-old, n=4) mice, and (B) CD34$^+$ cells isolated from cord blood (CB, n=4) and aged healthy donor BM (n=3). Violin plot showing the global mean intensity of each cell, from 2 independent experiments, measured by IMAGE J. 50-300 cells counted per sample. Dotted lines show median, and 25 and 75 percentiles. t-tests. ****p<0.0001.
(C, D) Western-blot analysis of H3K9me3 (C) and H3K9me2 (D) expression relative to total H3 in CD34$^+$ cells from 5 controls and 5 CMML patients. t-test. **p<0.01; *** p<0.001
(E, F) mRNA expression level measured by RT-qPCR of EHMT2, encoding G9A (E) and EHMT1, encoding GLP (F) in CD34$^+$ cells from controls and CMML patients. Ct values are normalized to the mean of housekeeping genes (HPRT, PPIA, Tubulin, Gus and/or RPL32) and reported to the mean of controls. Means +/- SEM from 2 independent experiments. Each dot represents cells from a single individual. t-test. **p<0.01
(G) Genomic repartition of ATAC-seq peaks uniquely found in patients (gained chromatin accessibility) or in controls (lost chromatin accessibility) or in common peaks.
(H) KEGG gene set analysis of deregulated genes associated with CMML peaks that are lost (upper panel, blue) or gained (lower panel, red) in CMML patient cells.
(I-N) H3K9me2 CUT&Tag analysis. (I) number of H3K9me2-bound peaks detected in the cells from the 3 controls and the 4 patients. Means +/- SEM; t-test. **p<0.01. (J) Mean expression levels of genes linked to H3K9me2 gained and lost peaks in CMML patients. (K) Correlation plot representing the $\log_2$ fold-change [FC] in H3K9me2 concentration at gene promoters vs. $\log_2$ fold-change in gene expression between patients and controls. (L) Numbers of TEs (filled bars) or genes (open bars) in H3K9me2-bound peaks that are found exclusively in the 4 patients (red) and 3 controls (blue). (M) Permutation test comparing of the number of patient (gained) and control (lost) specific H3K9me2-bound peaks overlapping with TEs specific and in 1000 lists of identical numbers of random peaks. Black curve: Distribution of the number of peaks overlapping TEs found in random peaks. Red vertical dotted line: number of peaks overlapping TEs in the differential peaks.

Hidaoui et al. Figure S3, related to Figure 3

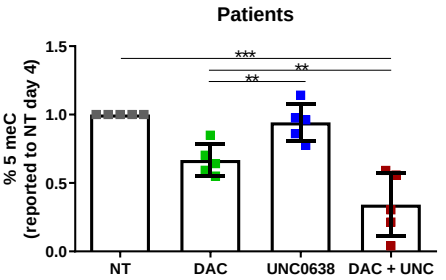
A.



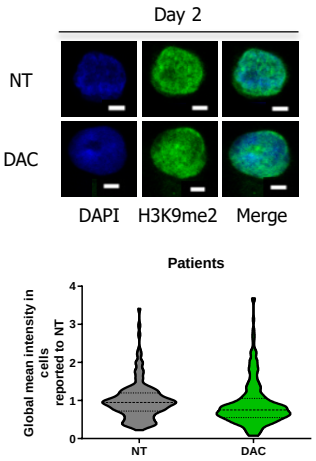
B.



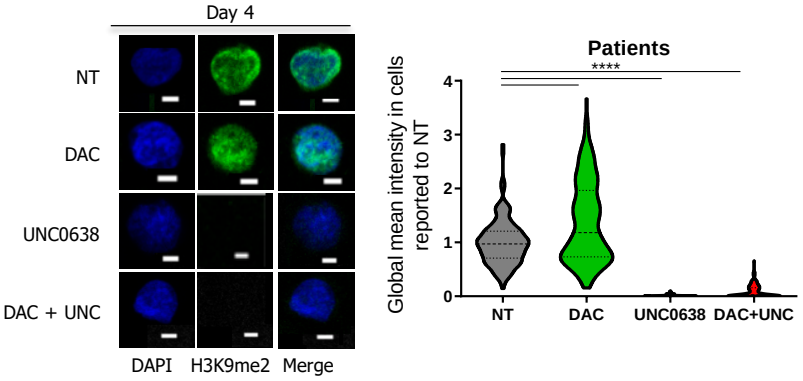
C.



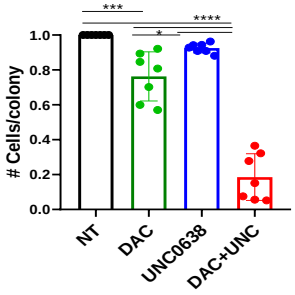
D.



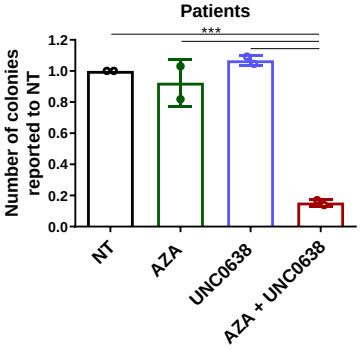
E.



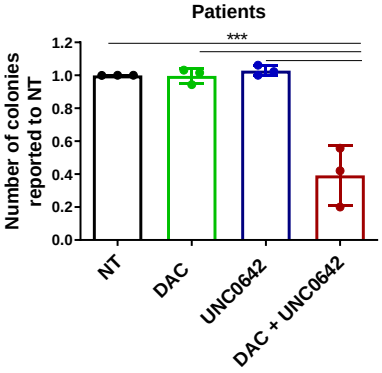
F.



G.



H.



I.

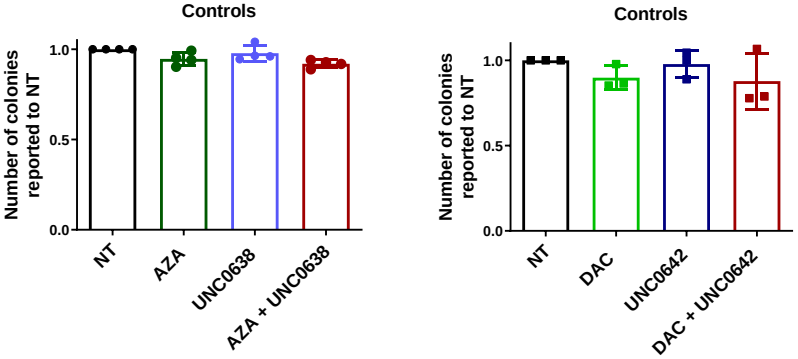


Figure S3, related to figure 3: The combinations of HMAs and G9A/GLP inhibitors induce a specific loss of CMML CD34$^+$ cell clonogenicity.

(A) Dose-response effect of DAC (X-axis) and UNC0638 (dark/grey scale) on CD34+ colonies formed by cells from one control.

(B, C) Measurement of DNA methylation level (% 5meC) after 2 (B) and 4 (C) days of liquid culture of CMML CD34$^+$ cells in the presence or absence of DAC (10 nM) and UNC0638 (1µM), as indicated. Means +/- SEM, n=5 patients. Results are normalized to the percent of 5meC in the NT condition.

(D, E). Representative images and quantification of H3K9me2 IF mean intensity from CMML patient (n=3) CD34$^+$ cells after 2 (D) and 4 (E) days of liquid culture reported to the mean intensity of non-treated cells. Bars, 3 µm. Images were quantified using ImageJ software.

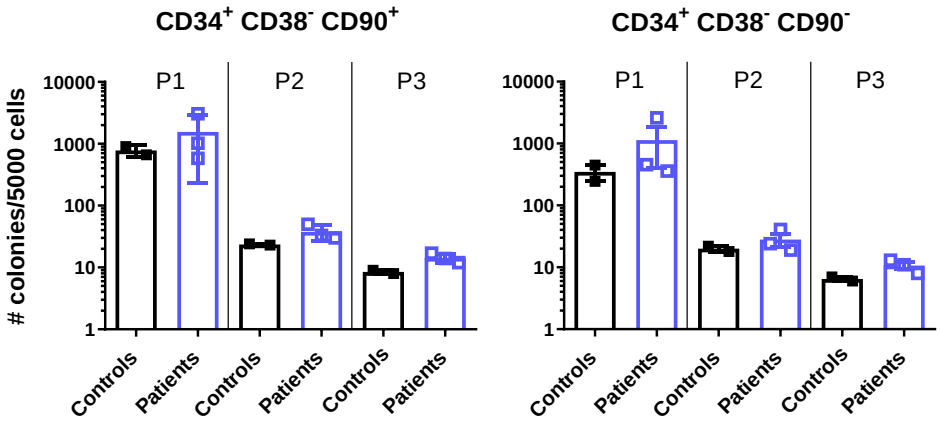
(F) Number of cells per colony formed by CD34$^+$ cells from CMML patients treated with DAC, UNC0638 or both.

(G-I) Number of colonies formed from patient (G, H) or control (I) CD34$^+$ cells after culture with 5-AZA-cytidine (AZA, 500 nM) in the presence of absence of G9A/GLP inhibitor UNC0638 (1 µM) (left panel) or the combination of DAC (10 nM) with the G9A/GLP inhibitor UNC0642 (2 nM) (right panel). One dot represents a single individual. Means +/- SEM.

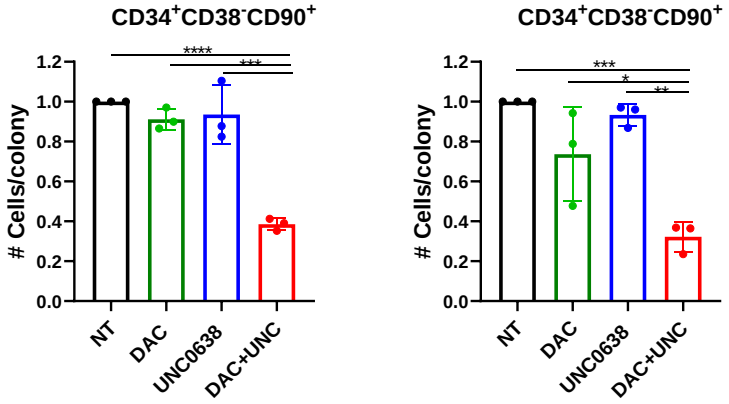
(B), t. test. (C, E-I). One-way ANOVA Bonferroni's Multiple Comparison *$p<0.05$; ***$p<0.001$.

Hidaoui et al. Figure S4, related to Figure 4

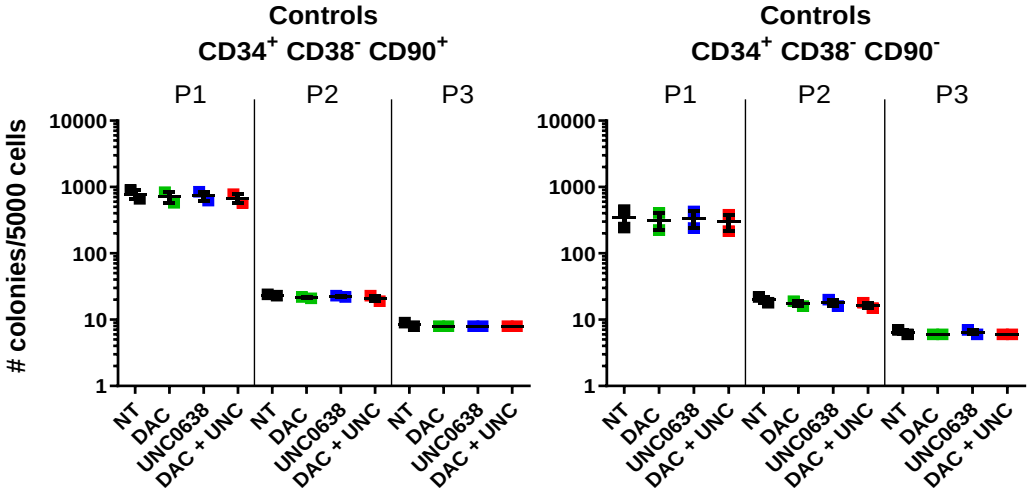
A.



B.

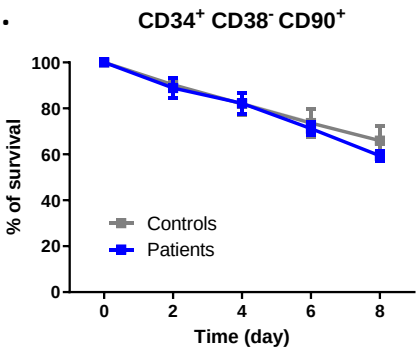


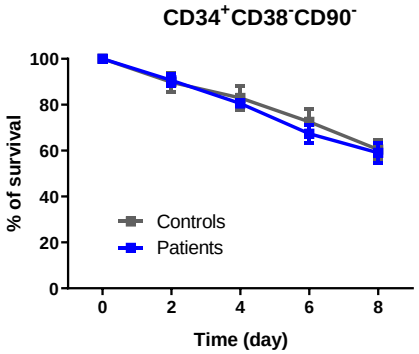
C.



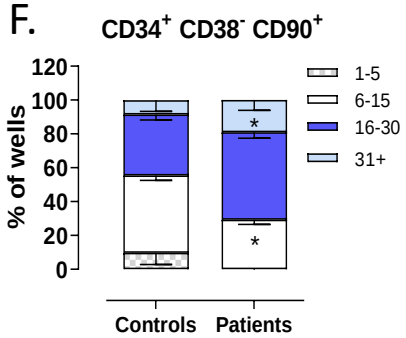
Hidaoui et al. Figure S4 related to Figure 4, continued

D.

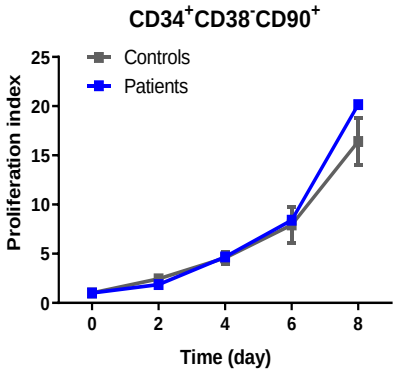


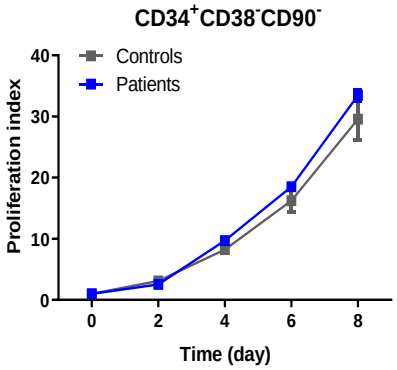


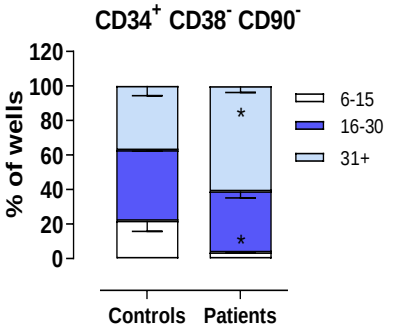
F.



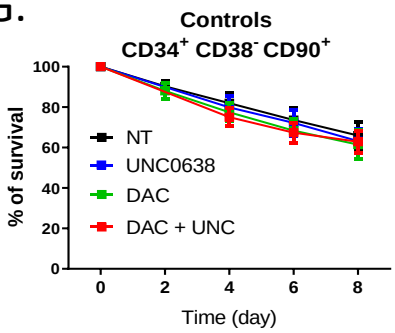
E.

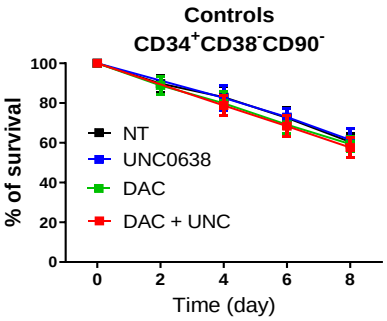




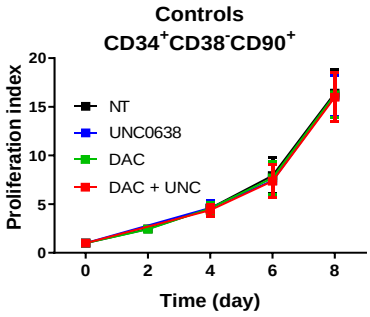


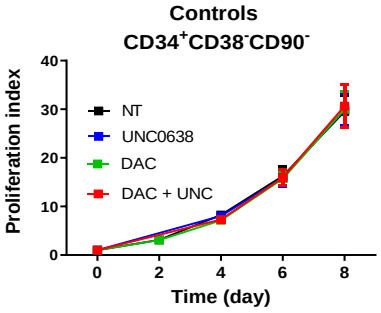
G.



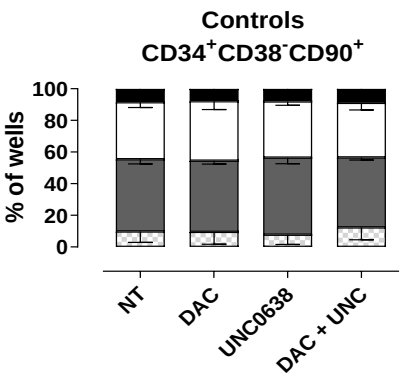


H.





I.



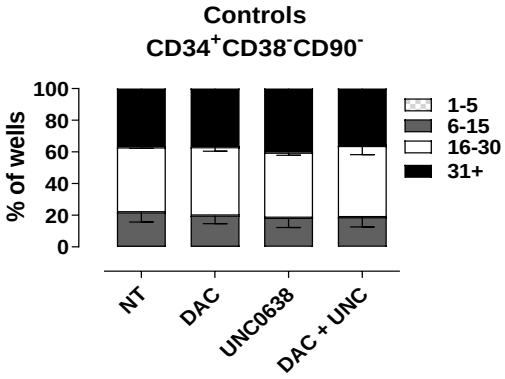


Figure S4, related to figure 4: The HMA and G9A/GLP inhibitor combination reduces CMML mutated stem cells clonogenicity.

(A) Number of colonies formed at repeated passages from stem cells ($CD34^+CD38^-CD90^+$, left panel) or progenitors ($CD34^+CD38^-CD90^-$, right panel) from CMML patients and controls before treatment. One dot represents a single individual

(B) Number of cells per colony formed by CMML patient stem cells (left panel) or progenitors (right panel) treated with DAC, UNC0638 or both, or left untreated. Results are reported to the mean number of cell/colony found in the untreated samples. Means +/- SEM. One dot represents a single individual One-way ANOVA Bonferroni's Multiple Comparison.

(C) Number of colonies formed at passages 1 to 3 by stem cells and progenitors from controls, in the presence or absence of treatment. Means +/- SEM. One dot represents a single individual One-way ANOVA Bonferroni's Multiple Comparison..

(D-I) Single cell liquid cultures comparing stem cells (left panel) and progenitors (right panel) from patients (n=3, blue curves) and control (n=3, grey curves) (D-F) or control cells after treatment with either DAC, UNC0638 or both, or left non-treated (n-=3) (G-I). (D, G) Viability, (E, H) proliferation index, (F, I) Cell number repartition. Means +/- SEM. One dot represents a single individual. Two-way ANOVA Bonferroni’s Multiple Comparison.

*$p<0.05$; **$p<0.01$; ***$p<0.001$.

Hidaoui et al. Figure S5, related to Figures 5

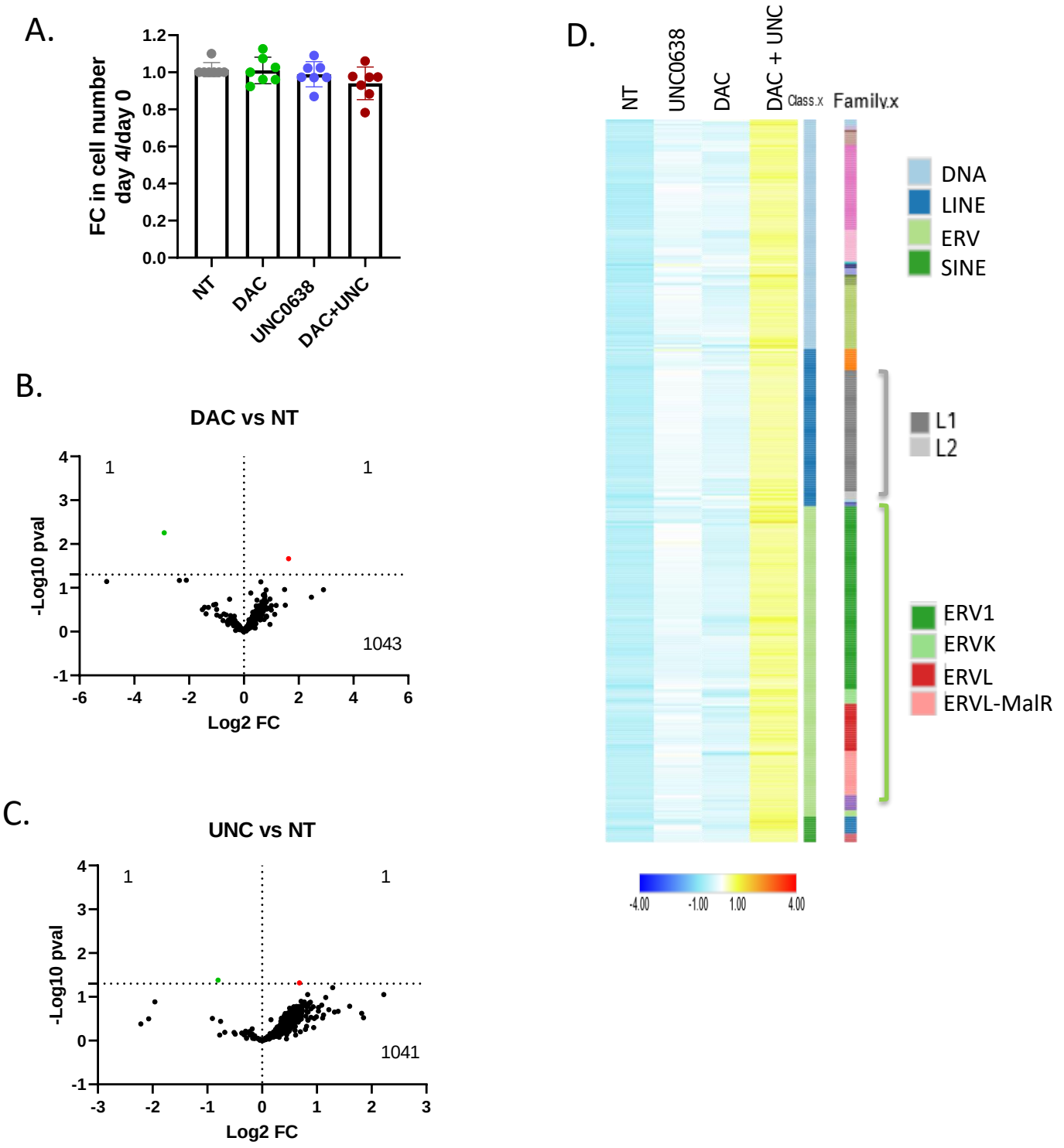


Figure S5, related to figure 5: The HMA and G9A/GLP inhibitor combination increases TE expression in CMML CD34⁺ HSPCs.

(A) Number of cells remaining in the cultures at day 4 in the presence of the various inhibitors. Fold-change in cell numbers reported to the number of cells seeded at day 0.

(B, C) Volcano plots showing differential TE expression in cells treated with DAC (B) or UNC0638 (C), compared to NT. Red and green dots: upregulated and downregulated TEs, respectively.

(D) Heatmap showing the median expression of DAC+UNC0638 vs. NT differentially expressed TEs in each of the treatment conditions, clustered on both TE class and TE family.

Hidaoui et al. Figure S6, related to Figure 6

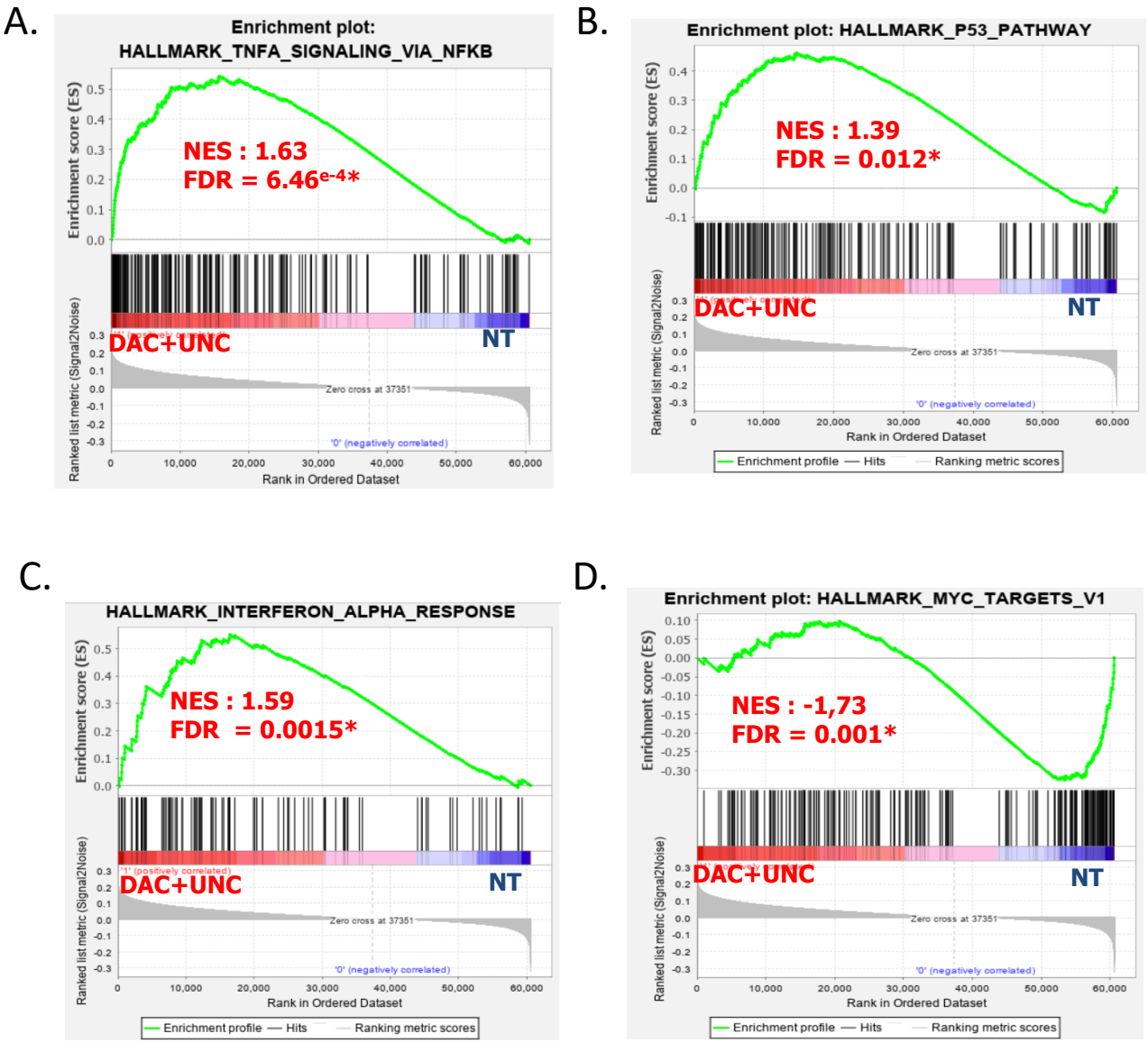


Figure S6, related to figure 6: The DAC/UNC0638 combination reactivates viral and inflammatory pathways in CMML cells.

(A-D) GSEA enrichment plots for selected pathways upregulated (A-C) or downregulated (D) in CMML cells after treatment with DAC+UNC0638. FDR, false discover rate; NES, normalized enrichment score. *, FDR<0.05.

Hidaoui et al. Figure S7, related to Figure 7, continued

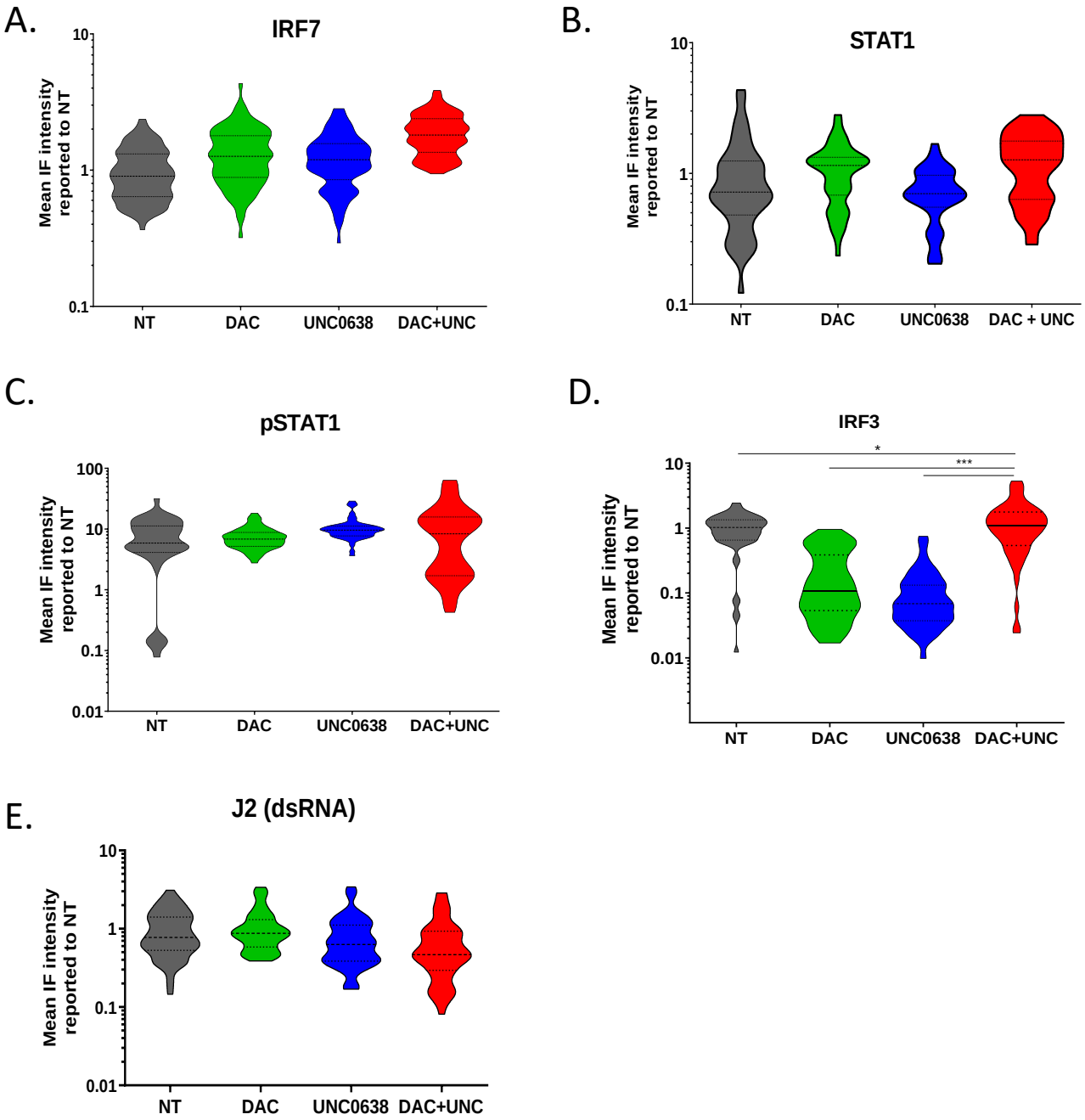


Figure S7, related to figure 7: The DAC/UNC0638 combination does not trigger IFN signaling in control cells(

(A-E) Immunostaining with antibodies against IRF7 (A), STAT1 (B), pSTAT1 (C), IRF3 (D) and dsRNAs (E) in CD34+ cells from one healthy donor after treatment with DAC +/- UNCO638, or untreated. Violin plots showing the global mean intensity of each cell, reported to the NT condition. One-way ANOVA Bonferroni's Multiple comparison. *$p<0.05$; ***$p<0.001$.

Table S1: CMML patient and healthy donor cohorts

	Cohort	
	Healthy controls	**Patients**
Number Males, N (%)	39 16 (42%)	100 54 (54%)
Age, median [range]	68 [45-91]	72 [40-90]
CMML-1 (%) CMML-2 (%)		78% 22%
Dysplastic (%) Proliferative (%)		42% 58%
Leukocytosis G/L; median [range]		15.8 [3.07-157.2]
Hemoglobin g/dL; median [range]		11.4 [4.5-16.6]
Platelets, G/L; median[range]		119 [11.6-986]
Neutrophils, G/L; median [range]		7.9 [0.6-95.9]
Monocytes, G/L; median G/L [range]		3.2 [0.65-53.2]
Monocytes, %; median [range]		20.4 [5-57.7]
Bone marrow blasts, median % [range]		5 |0-19]
Next generation sequencing analysis ; N (%)		100
ASXL1		39
BRCC3		1
CBL		15
CSF3R		3
CTCF		1
CUX1		6
DNMT3A		12
ETNK1		1
ETV6		2
EZH2		7
IDH1		5
IDH2		4
KDM6A		2
KIT		1
KRAS		13
NRAS		14
PHF6		4
PTPN11		1
RIT1		1
RUNX1		15
SF3B1		7
SRSF2		41
TET2		65
TP53		2
U2AF35		3
U2AZF1		1
ZRSR2		3

Table S5: Mutations and variant allele frequencies in 5 patients

Patient #3440	Patient #3480	Patient #3459	Patient #3562	Patient #3564
RUNX1 NM_001001890.2 exon 2 c.341C>T : p.Ser114Leu (61%)	**TET2** NM_001127208.2 exon 9 c.4139A>G : p.His1380Arg (37%)	**TET2** NM_001127208.2 exon 9 c.4140T>A : p.His1380Gln (44%)	**KRAS** c.G35A: p.G12D (30%)	**TET2** NM_001127208.2 exon:6 c.3686T>C p.Leu1229Pro (66%)
TET2 NM_001127208.2 exon 8 c.4012A>T : p.Lys1338*, nonsense mutation, (54%)	**DNMT3A** NM_022552.4 exon 8 c.1014+1delG : splice_donor_variant (34%)			**SRSF2** NM_001195427/1 Exon 1 c.284C>G p.Pro95Arg (37%)

Table S9: Primers

Primers for VAF analysis

Gene	Sense	Anti-sense
RUNX1-S114L (Patient #3340)	GTCCTTTGACTGGTGTTTAGG	GCATAGTTTTGACAGATAACG
TET2-K1338X (Patient #3340)	CTAACTGATAGTCTCTTTTACATAGAG	GTAACAAGTAAGTTGTTACAATTGC
TET2-H1380R (Patient #3380)	CAGATTGAATATGAACACAGAGC	CAGACTCAACAGCTGCTAAGC
DNMT3A c.1014+1 delG (Patient #3380)	GCATTGTGTCTTGGTGGATG	CATCACCCCAATTCCAGACTG
TET2-L1229P (Patient #3564)	GTGGTTCGCAGAAGCAGCAG	CATTCAAGGCACACCGGCG
SRSF2-P95R (Patient #3564)	TGTAAAACGACGGCCAGTGGCCGCCACTCAGAGCTA	GGAAACAGCTATGACCATGCGCGGACCTTTGTGAGGT
TET2-H1380Q (Patient #3459)	CAGATTGAATATGAACACAGAGC	CAGACTCAACAGCTGCTAAGC
KRAS-G12D (Patient #3562)	GTGTGCATGTTCTAATATAGTC	CTATTGTTGGATCATATTCGTCC

Primers for RT-QPCR analysis

Gene	Sense	Anti-sense
EHMT2	GGCCCCCTGAAATACAGCAT	TCGTCAGGGTCACTTCACCT
EHMT1	GTGTGAAAACCGAGCTGCTG	TCCTGCCTGTTTCTCTGCTG
ERVL	ATATCCTGCCTGGATGGGGT	GAGCTTCTTAGTCCTCCTGTGT
ERVLB4	CAGTCAGCCTCTTTCCCCAG	TGTTGCTGAGCCCATGCATA
MER4D0	TCCGGAGAGTCTGACACCTT	CCAGGTAGCAGGCTTCAGAG
ERVH	CCCTGTCCTCCTGCTCTTTG	TGGTGAGATGTTCCTTGGGC
ALUYH9	CGCCTGTAATCCCAGCACTT	GCCAGGATGGTCTCGATCTC
IFNB1	AAACTCATGAGCAGTCTGCA	AGGAGATCTTCAGTTTCGGAGG
IRF7	CGAGACGAAACTTCCCGTCC	GCTGCCTCGGTATGGATCTC
IRF9	GTCCAGCTGTCTGGAAGACTC	GGTACTTTCTGAGTCCCTGGC
GUS	GAAAATATGTGGTTGGAGAGCTCATT	CCGAGTGAAGATCCCCTTTTTA
HPRT	GGACAGGACTGAACGTCTTGC	CTTGAGCACACAGAGGGCTACA
TUBULINE	TCCAGATTGGCAATGCCTG	GGCCATCGGGCTGGAT
PPIA	GTCGACGGCGAGCCC	TCTTTGGGACCTTGTCTGCAA
RPL32	TGTCCTGAATGTGGTCACCTG	CTGCAGTCTCCTTGCACACCT

Table S10: Primary and secondary antibodies used in immunofluorescence and western blots

Specificity	**Référence**	Host	Dilution	Fluorochrome
G9A	GTX129153 *(GeneTex)*	Rabbit	1/200	
GLP	HPA060022 *(Altas antibody)*	Rabbit	1/200	
IRF3	TA311894 *(Origen)*	Rabbit	1/100	
IRF7	HPA052757 *(Sigma-Aldrich)*	Rabbit	1/200	
DsRNA	MABE1134-J2 *(Millipore)*	Rabbit	1/150	
H3K9me2	39239 *(Active Motif)*	Rabbit	1/500	
H3K9me3	8898 *(abcam)*	Rabbit	1/1000	
H3	18521 *(abcam)*	Rabbit	1/5000	
STAT1	14994 *(Cell Signaling)*	Rabbit	1/400	
pSTAT1 (Tyr701)	9167 *(Cell Signaling)*	Rabbit	1/400	
Rabbit	A11008 *(ThermoFisher Scientific)*	Goat	1/600	AlexaFluor 488
Mouse	A21425 *(ThermoFisher Scientific)*	Goat	1/600	AlexaFluor 555

Supplementary methods

RNA-seq data processing

Read quality was assessed with Fastqc v0.11.9, and corrected with Fastqp (1) v0.23.1. Read alignments were performed with STAR (v2.7.10b) on GRCh38 human genome reference. Mapped reads underwent filtration to retain only those associated with regular chromosomes (1-23, XY) and the remaining reads were sorted and indexed. Raw count table were generated with FeatureCounts (package subread, v2.0.1). Transcript expression levels were aggregated in gene expression levels using tximport Bioconductor package v1.13.16. For the CMML versus controls, a batch effect correction was first done using ComBat-seq function of the sva package, before subsequent DESeq2 analysis (2,3). PCA analysis was performed with DESeq2, after regularized log transformation of gene counts using default parameters. The MA plot was done on shrinked log2 fold changes using the apeglm methode (4). Pathway analysis (HALLMARK, GO, KEGG) were performed and plotted on R using msigdbr (7.5.1), and clusterProfiler (4.2.2 or 4.6) (5) or the GSEA app of MSigDB (6,7). GSEA analyses of CMML vs. controls cells were performed and genes pre-ranked on the statistical result of the differential analysis, and pathways were sorted on their NES absolute score before plotting. The tree plot was performed using the pairwise_termsim function of enrichplot (1.18.3).

Multimapping and expression of TE families was analyzed with TEtranscript (v 2.2.1). (8). Paired-end reads were aligned to GRCh38 genome using STAR (v2.7.11a) with these parameters: --winAnchorMultimapNmax 100 --outFilterMismatchNmax 3 --alignEndsType EndToEnd --alignIntronMax 1 --alignMatesGapMax 350. Differentially expressed TEs were identified using the R (v4.1.2) package DESeq2(v1.3.4), with criteria including p-value < 0.05 and an absolute value of log2(FC) > 0. Downstream analysis was conducted using custom R scripts.

ERV expression was quantified using a loci-based approach, as previously described (9). Briefly, RNA-seq reads were aligned to a custom transcriptome using bowtie2 v2.2.1 (10) with custom parameters to retain multimapped reads (-k 100 --very-sensitive-local --score-min "L,0,1.6"). The custom transcriptome consisted in the hg38 reference transcriptome with 14,968 ERV transcriptional units compiled from RepeatMasker annotations (11). SAM outputs were converted to BAM files using SAMtools v1 (12). ERV and gene expression were then quantified using Telescope (11) and HTSeq 0.12.3 (13), respectively. Raw counts were concatenated and normalized independently for each dataset using DESeq2 v1.28.0 with variance stabilizing transformation (VST)(3). Differential expression analysis was performed using DESeq2. Raw counts of ERVs and genes were merged and integrated into the same DESeq object, using patient ID and condition as a covariate in the design formula to account for data pairing (design = ~ patient_id + condition). For each comparison, the untreated (NT) condition was considered as reference. Fold change were shrunk with the ASHR method (14). ERV and coding gene expression is shown in median z-score of log2 normalized counts. Union of differentially expressed coding genes or ERV were selected based on a 1.5-fold change and pvalue < 0.05. Hierarchical clustering was performed with complete Linkage Method and Pearson correlation distance. Data was processed with R version 4.1.2 and visualized using Morpheus (https://software.broadinstitute.org/morpheus). Functional annotations (HALLMARK, GO, KEGG) of each cluster were performed using several R packages including clusterProfiler (4.2.2), org.Hs.eg.db (3.14.0) and msigdbr (7.5.1). IFN signatures were retrieved from The Molecular Signatures Database (MSigDb). Individual enrichment scores were calculated from each patient by single sample gene-set variation analysis (ssGSVA), and center-scaled. Supervised and unsupervised hierarchical clustering were then performed on scaled ssGSVA scores using the ComplexeHeatmap R package.

ATAC-seq data processing

Read quality was assessed using Fastqc (1) (v0.11.9) and MultiQC (15) (1.10.1). Bowtie2 (10) (v.2.3.4.1) using the «--very-sensitive» parameter was used for aligning ATAC-seq reads to the human genome version GCCh38/hg38 and Samtools (12) (v.1.7) for data filtering and file format conversion. Duplicate reads were removed with Picard tool - MarkDuplicates (http://broadinstitute.github.io/picard), blacklist regions, chrM and Y were removed with BEDtools (16) intersect and grep -v respectively before peak calling. All filtered .bam files were converted to bedgraphs using the DeepTools (17) bamCoverage subcomm and, with the reads per kilobase of transcript, per million mapped reads (RPKM) normalization method. MACS2 (v.2.1.1) (18) algorithm was used for peak identification (Q-value cut-off = 0.01). Gained and lost peaks were identified from the narrow peaks in two steps. First, the peak lists from Control and CMML patients were merged separately with the subcommands cat and mergeBed to obtain consensus peaks. Second, with intersectBed and the parameters --a <consensus peaks> --b <narrow peaks> --wa --u and subsequently --v, we compare and subtract them, to take the gained and lost peaks for each group. The reads of these consensus peaks were counted and a statistical model based on edgeR (19) was used to identify the significantly differential peaks. To identify the common peaks between patients or controls, DESeq (3) which uses the Negative Binomial distribution to compute a P value and a fold change for each estimated peak was used. Peaks highly enriched in comparison to the rest were considered gained. Annotation of peaks to genes (100 kb upstream and 25 kb downstream from the TSS) and genomic distribution of accessible regions identified by MACS2 was performed using BEDTools and the -closetBed and -intersectBed subcommands, respectively. Clustering of regions was generated with the ComputeMatrix function of DeepTools, using the reference point --referencePoint center -b 2500 -a 2500 -R <bed files> -S <bigwig files> as parameters. The function plotHeatmap from the same package was used for displaying the average profiles

heat map. Differentially accessible chromatin regions were scanned for enriched short-sequence motifs using HOMER (20) software with “findMotifsGenome.pl” command. Gene ontology and pathway analysis for integrated genes between ATAC and RNA-seq genes was performed with DAVID (21) and Metascape (metascape.org) (22). Only pathways and biological processes with p value less than 0.05 were considered to be significantly enriched.

CUT&Tag data processing

CUT&Tag data were processed and analyzed mostly as described (23). Data integrity was assessed through md5sum (v8.22), followed by the verification of quality control through Fastqc (1) (v0.12.1) and MultiQC (15) (v1.13), adapter and quality trimming with Cutadapt (v0.6.10). Paired-end reads were aligned to GRCh37 genome using Bowtie2 (v2.4.1) with the following parameters: --end-to-end --very-sensitive --no-mixed --no-discordant --phred33 -I 10 -X 700. Duplicate reads were excluded using Picard (v2.26.9). Reads aligned on regular chromosomes were retained and the aligned read quality score was set to «0» using samtools (v1.13) to get both uniquely mapped and multimapped reads. The Y chromosome was removed from the downstream analysis. The aligned reads underwent two processing methods: 1) direct sorting and indexing BAM files using samtools (v1.13), followed by peak calling with MACS2 (v2.2.7.1) with the specified parameters adapted to broad marks: -B --broad --broad-cutoff 0.1 -f BAMPE -g hs --max-gap 2000 --min-length 200. 2) Downsampling BAM files to the minimum read count across BAM files, subsequent sorting, and indexing using samtools (v1.13). These processed BAM files were then utilized for extracting read counts within peaks. Subsequent downstream analysis was carried out using custom scripts in R (v4.1.2). A matrix of counts for union peaks was generated using chromVAR (v1.16.0) R package and differentially expressed peaks between CMML Patients and controls were identified using the R (v4.1.2) package DESeq2(v1.3.4) , with criteria including an p-value < 0.05 and an absolute value of log2(FC) > 0. Peaks consistently identified across all three

biological replicates of controls (n=3) and the four biological replicates of CMML patients (n=4) were kept for further analysis.

Enrichment of transposable elements

Enrichment of TEs in peaks was performed as described (24), using a permutation test conducted independently for each TE class. A peak was considered associated with a TE if there is at least 1 base pair overlap between them. Peaks that intersected with any TE belonging to a given TE class were counted in CMML unique (gained) and control unique (lost) peaks. An equal number of peaks were randomly chosen from peaks that do not belong to any section of the Venn diagram, and the count of peaks overlapping with the designated TE class was determined. This iterative process was repeated 1000 times to establish a null distribution for the number of peaks overlapping with the TEs. Enrichment or depletion of the different classes of TEs was calculated as the log2 fold change of the number of specific peaks overlapping with TEs in any TE class found in patient gained or lost peaks over the median number of TEs of the same class in the randomly selected peaks. The empirical p-value was calculated by dividing by 1000 the number of occurrences where the permutation resulted in values exceeding the number of gained peaks overlapping the TE class.

Supplementary references

1. Wingett SW, Andrews S. FastQ Screen: A tool for multi-genome mapping and quality control. F1000Res **2018**;7:1338 doi 10.12688/f1000research.15931.2.
2. Zhang Y, Parmigiani G, Johnson WE. ComBat-seq: batch effect adjustment for RNA-seq count data. NAR Genom Bioinform **2020**;2(3):lqaa078 doi 10.1093/nargab/lqaa078.
3. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome biology **2014**;15(12):550 doi 10.1186/s13059-014-0550-8.
4. Zhu A, Ibrahim JG, Love MI. Heavy-tailed prior distributions for sequence count data: removing the noise and preserving large differences. Bioinformatics **2019**;35(12):2084-92 doi 10.1093/bioinformatics/bty895.
5. Wu T, Hu E, Xu S, Chen M, Guo P, Dai Z*, et al.* clusterProfiler 4.0: A universal enrichment tool for interpreting omics data. Innovation (Camb) **2021**;2(3):100141 doi 10.1016/j.xinn.2021.100141.
6. Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA*, et al.* Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression

profiles. Proceedings of the National Academy of Sciences of the United States of America **2005**;102(43):15545-50 doi 10.1073/pnas.0506580102.

7. Mootha VK, Lindgren CM, Eriksson KF, Subramanian A, Sihag S, Lehar J*, et al.* PGC-1alpha-responsive genes involved in oxidative phosphorylation are coordinately downregulated in human diabetes. Nature genetics **2003**;34(3):267-73 doi 10.1038/ng1180.
8. Jin Y, Tam OH, Paniagua E, Hammell M. TEtranscripts: a package for including transposable elements in differential expression analysis of RNA-seq datasets. Bioinformatics **2015**;31(22):3593-9 doi 10.1093/bioinformatics/btv422.
9. Alcazer V, Bonaventura P, Tonon L, Michel E, Mutez V, Fabres C*, et al.* HERVs characterize normal and leukemia stem cells and represent a source of shared epitopes for cancer immunotherapy. Am J Hematol **2022**;97(9):1200-14 doi 10.1002/ajh.26647.
10. Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. Nat Methods **2012**;9(4):357-9 doi 10.1038/nmeth.1923.
11. Bendall ML, de Mulder M, Iniguez LP, Lecanda-Sanchez A, Perez-Losada M, Ostrowski MA*, et al.* Telescope: Characterization of the retrotranscriptome by accurate estimation of transposable element expression. PLoS Comput Biol **2019**;15(9):e1006453 doi 10.1371/journal.pcbi.1006453.
12. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N*, et al.* The Sequence Alignment/Map format and SAMtools. Bioinformatics **2009**;25(16):2078-9 doi 10.1093/bioinformatics/btp352.
13. Anders S, Pyl PT, Huber W. HTSeq--a Python framework to work with high-throughput sequencing data. Bioinformatics **2015**;31(2):166-9 doi 10.1093/bioinformatics/btu638.
14. Stephens M. False discovery rates: a new deal. Biostatistics **2017**;18(2):275-94 doi 10.1093/biostatistics/kxw041.
15. Ewels P, Magnusson M, Lundin S, Kaller M. MultiQC: summarize analysis results for multiple tools and samples in a single report. Bioinformatics **2016**;32(19):3047-8 doi 10.1093/bioinformatics/btw354.
16. Quinlan AR, Hall IM. BEDTools: a flexible suite of utilities for comparing genomic features. Bioinformatics **2010**;26(6):841-2 doi 10.1093/bioinformatics/btq033.
17. Ramirez F, Dundar F, Diehl S, Gruning BA, Manke T. deepTools: a flexible platform for exploring deep-sequencing data. Nucleic acids research **2014**;42(Web Server issue):W187-91 doi 10.1093/nar/gku365.
18. Zhang Y, Liu T, Meyer CA, Eeckhoute J, Johnson DS, Bernstein BE*, et al.* Model-based analysis of ChIP-Seq (MACS). Genome biology **2008**;9(9):R137 doi 10.1186/gb-2008-9-9-r137.
19. Robinson MD, McCarthy DJ, Smyth GK. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics **2010**;26(1):139-40 doi 10.1093/bioinformatics/btp616.
20. Heinz S, Benner C, Spann N, Bertolino E, Lin YC, Laslo P*, et al.* Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. Molecular cell **2010**;38(4):576-89 doi 10.1016/j.molcel.2010.05.004.
21. Sherman BT, Hao M, Qiu J, Jiao X, Baseler MW, Lane HC*, et al.* DAVID: a web server for functional enrichment analysis and functional annotation of gene lists (2021 update). Nucleic acids research **2022**;50(W1):W216-W21 doi 10.1093/nar/gkac194.
22. Zhou Y, Zhou B, Pache L, Chang M, Khodabakhshi AH, Tanaseichuk O*, et al.* Metascape provides a biologist-oriented resource for the analysis of systems-level datasets. Nature communications **2019**;10(1):1523 doi 10.1038/s41467-019-09234-6.
23. Pelinski Y, Hidaoui D, Stolz A, Hermetet F, Chelbi R, Diop MK*, et al.* NF-kappaB signaling controls H3K9me3 levels at intronic LINE-1 and hematopoietic stem cell genes in cis. J Exp Med **2022**;219(8) doi 10.1084/jem.20211356.
24. Zezulin AU, Yen D, Ye D, Howell ED, Bresciani E, Diemer J*, et al.* RUNX1 is required in granulocyte-monocyte progenitors to attenuate inflammatory cytokine production by neutrophils. Genes & development **2023**;37(13-14):605-20 doi 10.1101/gad.350418.123.