haematologica
Journal of the Ferrata Storti Foundation

Single-cell proteo-transcriptomic profiling reveals altered characteristics of stem and progenitor cells in patients receiving cytoreductive hydroxyurea in early-phase chronic myeloid leukemia

by Hana Komic, Malin S. Nilsson, Lovisa Wennström, Tagore Sanketh Bandaru, Pekka Jaako, Kristoffer Hellstrand, Fredrik B. Thorén, and Anna Martner

Received: January 15, 2024.
Accepted: August 6, 2024.

Citation: Hana Komic, Malin S. Nilsson, Lovisa Wennström, Tagore Sanketh Bandaru, Pekka Jaako, Kristoffer Hellstrand, Fredrik B. Thorén, and Anna Martner. Single-cell proteo-transcriptomic profiling reveals altered characteristics of stem and progenitor cells in patients receiving cytoreductive hydroxyurea in early-phase chronic myeloid leukemia.
Haematologica. 2024 Aug 15. doi: 10.3324/haematol.2024.285071 [Epub ahead of print]

Single-cell proteo-transcriptomic profiling reveals altered characteristics of stem and progenitor cells in patients receiving cytoreductive hydroxyurea in early-phase chronic myeloid leukemia

Hana Komic[1,2*], Malin S. Nilsson[1,3*], Lovisa Wennström[4], Tagore Sanketh Bandaru[2], Pekka Jaako[5], Kristoffer Hellstrand[1,6], Fredrik B. Thorén[1,2#], Anna Martner[1,3#]

[1] TIMM Laboratory at Sahlgrenska Center for Cancer Research, University of Gothenburg, Gothenburg, Sweden

[2] Department of Medical Biochemistry and Cell Biology, Institute of Biomedicine, Sahlgrenska Academy, University of Gothenburg, Gothenburg, Sweden

[3] Department of Microbiology and Immunology, Institute of Biomedicine, Sahlgrenska Academy, University of Gothenburg, Gothenburg, Sweden

[4] Department of Hematology, Sahlgrenska University Hospital, Gothenburg, Sweden

[5] Sahlgrenska Center for Cancer Research, Department of Microbiology and Immunology, Institute of Biomedicine, Sahlgrenska Academy, University of Gothenburg, Gothenburg, Sweden

[6] Department of Infectious Diseases, Institute of Biomedicine, Sahlgrenska Academy, University of Gothenburg, Gothenburg, Sweden

[*], [#] Equal contribution

Running head: CITE-seq analysis of HU effects on CML cells

Corresponding author:

Anna Martner

ORCID ID: 0000-0002-6598-5221

Medicinaregatan 1F, 41390 Gothenburg

email: anna.martner@gu.se

Data-sharing statement:

CITE-seq expression data are available online in a Zenodo repository (10.5281/zenodo.12580559). R codes are available upon request. For fastq files, please contact anna.martner@gu.se.

Word count:

- Abstract: 242
- Main text: 3,497
- Number of tables: 0
- Number of figures: 8
- Number of supplementary files: 2

Authors' contributions

HK and MSN contributed equally as co-first authors, and FBT and AM as co-last authors. HK and MSN designed, performed, and analyzed experiments. MSN designed and performed the bioinformatic analysis with input from HK. TSB performed the SCENIC analysis. LW provided the CML patient material. HK and MSN isolated mononuclear cells from patient samples. AM, FBT, KH, PJ and LW conceived, designed, and supervised the study. HK,

MSN, KH, FBT and AM wrote the manuscript. All authors approved the manuscript prior to submission.

Disclosures

The authors received research funding in the form of free reagents from BD Biosciences within the BD Multiomics Alliance, but declare no other competing financial interests.

Acknowledgements

The authors would like to acknowledge all patients and caregivers involved in the study. Figures were created with BioRender.com.

Funding

This work was supported by grants from the Swedish Research Council (2020-02783 (FBT) and 2023-02193 (AM)), the Swedish Cancer Society (21 1864 Pj (KH), 22 2388 Pj (FBT) and 22 2128 Pj (AM)), the Swedish state via the ALF agreement (ALFGBG-718421 (KH), ALFGBG-963642 (FBT) and ALFGBG-724881 (AM)), the Swedish Childhood Cancer Fund (PR2021-0015 (PJ)), the Assar Gabrielsson Foundation and the Sahlgrenska Academy at the University of Gothenburg.

Abstract

Hydroxyurea (HU) is frequently used in the early phase of chronic myeloid leukemia (CML) to achieve cytoreduction prior to tyrosine kinase inhibitor (TKI) therapy. However, its impact on CML stem and progenitor cells (SPC) remains largely unknown. This study utilized targeted proteo-transcriptomic expression data on 596 genes and 51 surface proteins in 60,000 $CD14^{-}CD34^{+}$ cells from chronic phase CML patients to determine effects of short-term HU treatment (4-19 days) on CML SPC. Peripheral blood and bone marrow samples were obtained from 17 CML patients eligible for short-term HU treatment (three patients before and after HU; seven patients before HU; and seven patients after HU) and subjected to single-cell CITE-seq and/or flow cytometry analysis. The analysis revealed enhanced frequencies of hemoglobin-expressing (*HBA1, HBA2, HBB*) erythroid progenitor cells in blood and bone marrow following HU treatment. In addition, there was an accumulation of cell subsets with S/G2/M phase-related gene and protein expression, likely representing cells arrested in, or progressing slowly through, the cell cycle. The increased frequency of cells in S/G2/M phase after HU was observed already among the most immature leukemic stem cells (LSC), and patients with a high fraction of LSC in the S/G2/M phase showed poor responsiveness to TKI treatment. We conclude that short-term HU treatment entails differentiation of erythroid progenitor cells and alters the characteristics of LSC in CML. The results imply that studies of LSC and progenitor populations in CML should take effects of initial HU therapy into account.

KEYWORDS: hydroxyurea, chronic myelogenous leukemia, leukemic stem cell, progenitor cells, CITE-seq

Introduction

Chronic myeloid leukemia (CML) is a clonal myeloproliferative neoplasm characterized by a balanced t(9;22) translocation leading to formation of the Philadelphia chromosome and the *BCR::ABL1* fusion oncogene.[1,2] BCR::ABL1 conveys constitutive tyrosine kinase activity that translates into expansion and accumulation of malignant cells of different maturation stages in blood and bone marrow.[2] Imatinib as well as second and third generation tyrosine kinase inhibitors (TKI) competitively block BCR::ABL1 kinase activity and have drastically improved survival in CML.[3,4]

Prior to the introduction of TKI therapy, CML patients received treatment with the chemotherapeutics hydroxyurea (HU) and busulfan, the cytokine interferon-alpha or allogeneic stem cell transplants.[5,6] While these treatments are no longer first line options, many CML patients receive a short-term course of HU to reduce hyperleukocytosis while awaiting definitive diagnosis, as advised by European LeukemiaNet (ELN) recommendations.[7] Although frequently used, recently published data suggest that HU pretreatment does not add value in terms of meeting ELN CML treatment response milestones in the TKI era, and that early-phase HU should be restricted to patients with symptomatic hyperleukocytosis or splenomegaly.[8]

HU inhibits ribonucleoside diphosphate reductase, thus limiting the available deoxyribonucleotides to target rapidly proliferating cells.[9] In addition to direct effects on cell cycle proliferation, HU has been shown to generate nitric oxide (NO) radicals.[10–12] While the role of NO for the antileukemic effects of HU is unclear, NO-induced elevation of cyclic GMP reportedly promotes gamma-globin and fetal hemoglobin production and contributes to the favorable effects of HU in sickle-cell anemia.[10,13] *In vitro* studies suggest that NO may

exert pro-proliferative, anti-proliferative, or differentiating effects on hematopoietic stem and progenitor cells (HSPC) depending on concentration.[14]

CML is maintained and propagated by a rare population of *BCR::ABL1*-expressing leukemic stem cells (LSC).[15,16] These cells are considered partly resistant to TKI, which is why many patients require lifelong treatment with ensuing toxicity and risk of TKI resistance development.[17] Considerable efforts have been devoted to developing strategies to target the therapy-resistant LSC population.[18] In many studies, CML patients have received HU prior to TKI, but to our knowledge no studies have evaluated to what extent HU affects the phenotype of the LSC and progenitors. For the present study, we aimed to address this question through targeted proteo-transcriptomic single-cell profiling of >60,000 CML stem and progenitor cells (SPC) from peripheral blood and bone marrow samples obtained before and after HU treatment. Our results revealed (i) enhanced erythroid maturation and (ii) increased fractions of cells in S/G2/M phase among CML SPC after HU treatment.

Methods

CML patient samples

Peripheral blood and/or bone marrow (BM; from the posterior iliac crest) diagnosis samples were collected from 17 chronic phase CML patients eligible for short-term treatment with hydroxyurea (HU, 4-19 days) at the Sahlgrenska University Hospital (Gothenburg, Sweden) and Uddevalla Hospital (Uddevalla, Sweden). From seven patients only pre-HU samples were collected, from seven patients only post-HU samples were collected, and from three patients samples were collected both before and after HU treatment. Patient and sample characteristics, including HU dosage information, are summarized in Table S1. The study was approved by the Regional Ethics Review Board in Gothenburg (approval no. 011-17) and

conducted according to the principles of the Declaration of Helsinki. All patients gave written informed consent prior to sample collection.

Isolation and phenotypic analyses of peripheral blood and BM mononuclear cells

Patient peripheral blood and BM mononuclear cells (PBMC and BM MNC) were isolated by density gradient centrifugation on Lymphoprep (STEMCELL Technologies). Samples were enriched for stem and progenitor cells using FACS sorting of MNC, CD15 depletion, or $CD34^+$ cell selection using MACS Microbeads (Miltenyi Biotec) and were thereafter cryopreserved. Cells were subjected to immunophenotypic analysis and cell cycle analysis as detailed in the Supplementary Methods.

CITE-seq library preparation

PBMC obtained before HU and PBMC and BM MNC obtained after HU from two patients (sCML14 and sCML23) were subjected to single-cell proteo-transcriptomic CITE-seq analysis where the expression of 596 genes and 51 surface proteins was assessed (Tables S2 and S3). Samples were barcoded and pooled, after which cells were labeled with AbSeq, fluorescent and unconjugated antibodies (Tables S3 and S4), and $CD14^-CD34^+$ MNC FACS-sorted. CITE-seq sequencing libraries were generated from the isolated $CD14^-CD34^+$ cells using the BD Rhapsody Single-Cell Analysis System (BD Biosciences) as described in Nilsson *et al*.[19] Further details regarding the CITE-seq library preparation and sequencing are provided in the Supplementary Methods.

Data analysis

Fastq files were processed using the BD Rhapsody Targeted Analysis Pipeline (v.1.10.1; BD Biosciences) on the Seven Bridges Genomics platform (https://www.sevenbridges.com).

Low-quality cells were removed from the UMI count files using genes expressed vs library size dot plots in SeqGeq (v. 1.8.0; BD Biosciences).[19] The mRNA and AbSeq RSEC UMI count files were analyzed in Rstudio (v. 2022.07.1+554; R (v. 4.2.1)) using the Seurat package (v. 4.2.0)[20] and in Python (v. 3.1.0) using the pySCENIC package (v. 0.12.1) as described in the Supplementary Methods.

Analyses were also performed on a previously published dataset of $CD14^{-}CD34^{+}$ CML BM cells that were subjected to CITE-seq using the same set of genes and surface proteins. Details for the generation of the UMAPs in this analysis can be found in the original publication (Nilsson *et al.*[19]).

Statistics

All statistical analyses were performed in GraphPad Prism (v. 9.5.1) and Rstudio (v. 2022.07.1+554; R (v. 4.2.1)). Differences in cell type proportions were evaluated in GraphPad Prism using the unpaired Mann-Whitney test. Differentially expressed genes and proteins were assessed by the Wilcoxon Rank Sum test in R using Seurat's FindMarkers and FindAllMarkers functions.

Results

Hydroxyurea alters the proportions of stem and progenitor cells in CML blood and BM

Peripheral blood and bone marrow (BM) samples from 17 newly diagnosed CML patients obtained before and after hydroxyurea (HU) treatment were analyzed for stem and progenitor cell (SPC) frequencies using flow cytometry (Figure 1A). The analysis revealed a non-significant reduction of the frequency of $CD34^{+}$ cells following HU (Figure 1B), possibly reflecting reduced proliferation of CML SPC. In addition, the frequency of immature

$CD34^+CD38^-$ cells among $CD34^+$ SPC in BM was significantly lower after treatment, with a similar but non-significant trend in blood (Figure 1C). Results from paired samples obtained before and after HU treatment from three patients supported these findings (Figure S1).

Proteo-transcriptomic CITE-seq analysis of CML stem and progenitor cells

To further elucidate HU effects on the $CD34^+$ CML SPC compartment, paired blood and BM samples from two patients (obtained before and after 7 or 9 days of HU) were subjected to proteo-transcriptomic CITE-seq analysis of 596 genes and 51 proteins (Figure 2). 26,000 single $CD14^-CD34^+$ cells were successfully sequenced with sequencing saturations of 93 and 97% for mRNA and AbSeq libraries, respectively. After quality control, 21,044 cells remained for analysis.

HU treatment induces erythroid maturation among CML SPC

Following integration of mRNA expression data by patient, dimensionality reduction and clustering of cells from all six analyzed samples (blood before and blood/BM after HU) was performed (Figure 3A). The obtained UMAP did not contain patient-specific clusters (Figure S2A) and the confined protein expression of known cell type markers (Figure S2B) supported the validity of the mRNA-based clustering. Clusters were manually annotated based on expression patterns of known marker genes (Figure 3B, and Table S5), and proteins (Figure 3C, Table S6). Cells annotated as leukemic stem cells (LSC) were $CD34^+CD38^-CD45RA^-CD90^+$ and expressed the CML LSC markers CD25 and CD26. Myeloid progenitors (MP) were characterized by expression of *SPINK2*, *CSF3R*, *CEBPA*, CD45RA and CD371. While all megakaryocytic/erythroid progenitor clusters (MEP, MKP, EP) expressed the lineage-specific transcription factor *GATA1*, megakaryocytic progenitors (MKP) were further defined by expression of *MPL*, *VWF* and CD9, and erythroid progenitor (EP) clusters by CD35.

Megakaryocytic/erythroid progenitors (MEP) expressed *GATA1*, but lacked concurrent committed MKP/EP marker expression. The most mature EP cluster (EP-II) showed a distinct expression of *HBB, AHSP* and CD235a, and eosinophil/basophil/mast cell progenitors (EBMP) were characterized by expression of *HDC* and *GATA2*.

To assess proportional shifts within the SPC compartment as a result of HU treatment, cells from blood and BM samples obtained before and after treatment were separately highlighted on the UMAP (Figure 3D). Although all nine clusters were represented in all samples, post-treatment samples showed apparent cell density increases in EP-II and EP-Cy clusters and concomitant decreases in the EP-I cluster. BM samples obtained after HU showed a similar pattern of cell cluster distribution as the paired blood samples, though the fraction of cells in the EP-II cluster was even more pronounced in the BM samples. The EP-Cy and EP-II clusters displayed increased expression of an array of checkpoint-related genes such as *CCNE1* (EP-Cy, EP-II), *CHEK2* (EP-Cy), *CDKN2C* (EP-Cy), *FANCI* (EP-Cy) and *TRIP13* (EP-Cy) (Table S5), with cells in the EP-II cluster additionally defined by upregulation of the hemoglobin subunits *HBA1, HBA2,* and *HBB* (Figure 3E). Based on this finding, we additionally employed SCENIC analysis focusing on *GATA1*; a key transcription factor for erythropoiesis.[21] The *GATA1* regulon was identified based on expressional patterns in the paired blood samples from both patients. Many genes within this regulon showed significant upregulation following HU treatment, including *HBB*, *HBA1, E2F2* and *TFRC* (Table S7).

By utilizing Seurat's CellCycleScoring function, cycling erythroid progenitors (EP-Cy) were defined by the high proportion of S/G2/M phase cells within the cluster (Figure 4A). Their cycling nature was also supported by the cluster differential expression analysis (Table S5), where many of the upregulated genes were associated with cell division. Further exploration

of the clusters displaying consistent proportional changes across patients revealed that the proportionally decreased EP-I population primarily comprised cells in G0/G1 phase, whereas the increasing EP-Cy and EP-II clusters mainly held S/G2/M phase cells (Figure 4A). The shift towards S/G2/M phase following HU treatment was also seen among all SPC, disregarding cluster identity (Figure 4B). Furthermore, the expression pattern of the *CHEK2* gene (known to be upregulated following cell cycle arrest[22]), suggested that cells in EP-Cy and EP-II clusters may be arrested in the cell cycle (Figure 4C).

To further address aspects of cell cycling in blood samples obtained before and after HU treatment, we performed flow cytometry analysis using DAPI and Ki67 staining of $CD14^-$ $CD34^+$ cells (Figure 4D). In agreement with the CITE-seq findings, a larger fraction of $CD14^-CD34^+$ cells appeared in the S/G2/M cell cycle phases after HU treatment (Figure 4E). Furthermore, an increased frequency of 2n $Ki67^{hi}$ cells was observed in samples following HU (Figure 4F). Cells with this phenotype are often denoted G1 cells, but at the transcriptional level there was no increase in G1 cells following HU. We therefore speculate that this cluster represents cells with S phase transcriptional profile that fail to synthesize new DNA due to HU. In combination, CITE-seq and flow cytometry findings thus implied HU-induced effects, including erythroid maturation and a shift towards S/G2/M phase, among CML SPC.

To validate these findings, we utilized a previously published dataset of CITE-seq-subjected CML SPC (Nilsson *et al.*[19]). This cohort included samples from five patients prior to HU administration and seven patients who had previously received HU (Figure 5A). CML patients who never received HU in this dataset (likely harboring less proliferative malignant clones) were excluded from the analysis to avoid bias. Clusters were annotated using a

similar approach as described above (Figure S3A-B). Also in this cohort, SPC from HU-treated patients had a significantly higher fraction of cells in the hemoglobin subunit-expressing erythroid cluster (EP-IV; Figure 5A-C). The treated patients additionally showed a higher proportion of cells in the EP-Cy-II cluster that displayed increased expression of S/G2/M and checkpoint-related markers, including *CHEK2* (Figure 5B-D).

The shift towards S/G2/M-related gene expression after HU is present already among the most immature CML cells

To achieve higher resolution for analysis of the effects of HU on the most immature CML stem cell population in the paired CITE-seq data, CD34 and CD38 protein expression data were utilized for flow cytometry-like gating (Figure S4A). This identified 4,222 immature $CD34^{+}CD38^{-}$ cells that were included in subsequent analysis. The positions of the gated cells within the larger $CD14^{-}CD34^{+}$ UMAP confirmed that the selected cells mainly derived from the immature LSC and MP clusters, with additional contribution from MEP and MKP clusters (Figure S4B). In line with the flow cytometry results in Figure S1B, the CITE-seq data showed marginally reduced frequencies of $CD38^{-}$ out of $CD34^{+}$ cells following HU treatment (Figure S4C).

Reclustering of the selected cells yielded nine clusters (Figure 6A), without major patient-specific clustering (Figure S4D). Based on expressional patterns (Figure S4E-F, Tables S8 and S9), the UMAP largely divided into stem cells/multipotent progenitors (SC/MPP; e.g. characterized by CD90, *HLF*, CD25, CD26), cells with myeloid lineage commitment (myeloid progenitors, MP; e.g. CD371, CD45RA, *SPINK2*, *CSF3R*, *CEBPA*) and cells showing signs of megakaryocytic/erythroid/eosinophil/basophil/mast cell lineage bias (MEP/EBMP; e.g. *GATA1*, *MPL*, CD35, *HDC*).

Assessment of the relative proportions of cells in G0/G1 vs S/G2/M phase across samples revealed an HU-induced shift towards S/G2/M phase-related gene expression also among the immature $CD34^{+}CD38^{-}$ cells (Figure 6B). The proportional increase in cells with S/G2/M phase-related gene expression observed after HU treatment thus seemingly occurs across differentiation stages and lineages, ranging from the most immature $CD14^{-}CD34^{+}CD38^{-}$ cells to more mature erythroid-committed progenitors.

HU reduces the frequency of CML LSC with quiescent characteristics

In the previous CITE-seq study (Nilsson *et al.*[19]), the CML LSC compartment (defined by high expression of CD90, CD26 and CD25 along with low expression of CD38 and CD45RA) separated into two clusters in a gene expression-based UMAP. One of the LSC subsets, denoted 'LSC-I', appeared quiescent with an expression pattern resembling that of TKI-resistant LSC.[23] The other LSC subset, denoted 'LSC-II', showed higher expression of S/G2/M phase markers.[19]

To assess whether HU impacted on the most immature LSC population in CML, we used the aforementioned dataset as a reference and performed Seurat-based cell label transfer to annotate the $CD34^{+}CD38^{-}$ cells in the current analysis of paired samples (Figure 7A). The identities of cells annotated as hematopoietic (HSC-I and -II) and leukemic (LSC-I and -II) stem cells (Figure 7B) as well as other cell types (Figure S5A) largely corresponded to the previous lineage annotations (Figure 6A), supporting the validity of the cell label transfer. Expressional patterns for each annotated cell type are provided in Figure S5B-C.

When assessing the impact of HU on stem cell subtype distribution, a proportional shift from the LSC-I to the LSC-II subpopulation was noted within the stem cell compartment after HU treatment (Figure 7C). The decreasing LSC-I population was dominated by G0/G1 cells, while the increasing LSC-II population predominantly consisted of cells in S/G2/M phase (Figure 7D). The actively cycling phenotype of the LSC-II population compared to that of LSC-I was supported by upregulation of a wide range of cell cycle-related genes (e.g. *CCNA2*, *BIRC5*, *E2F2*, *ZWINT*, *NDC80*, *TUBB*) in differential expression analysis (Figure S5D, Table S10).

To validate these results, the effect of HU on stem cell distribution was assessed also in the previously published dataset (Nilsson *et al*.[19]), where BM samples were obtained from five patients prior to HU and from seven other patients after HU. In line with the results above, the stem cell compartment of patients who had received HU contained a significantly lower proportion of G0/G1-dominated LSC-I and a higher proportion of actively cycling (S/G2/M-dominated) LSC-II, while the proportion of HSC was unaltered (Figures 8A and S6A). Of note, the frequency of LSC-I and LSC-II among all $CD34^+CD38^-$ immature cells in this analysis was also significantly reduced and increased, respectively, after HU treatment (Figure S6B). Among HU-treated patients, the proportion of LSC-II did not correlate with time on HU treatment (R^2=0.18, P=0.35).

The majority of patients in the combined study cohort responded well to TKI treatment, but two patients did not attain a complete cytogenetic response (CCyR, 0% Ph^+ cells)[24] after three months of TKI treatment and eventually required second or third line ponatinib treatment to achieve CCyR. These insufficient responders, which were analyzed after HU, showed the highest proportion of LSC confined to the LSC-II subtype among all analyzed

patients (Figure 8B). Mutational analysis, according to clinical practice in Sweden, did not show presence of *BCR::ABL1* mutations in the poorly responding patients.

Discussion

Although many CML patients currently receive a course of cytoreductive hydroxyurea (HU) prior to starting TKI, this is to our knowledge the first study to report effects of this treatment on the stem and progenitor cells (SPC). As HU pre-treatment is indicated only in a subgroup of patients (those with high white blood cell counts or high platelet counts[7,25]), HU-related effects on the SPC compartment may impact conclusions of CML studies utilizing diagnosis samples with and without prior HU treatment. In this study, detailed proteo-transcriptomic CITE-seq analysis of SPC obtained from CP-CML patients before or after HU treatment indicated multiple HU-induced transcriptional changes within this compartment, including (i) an increased erythroid maturation, and (ii) an increased proportion of cells in S/G2/M phase.

Paired and unpaired comparisons of SPC from peripheral blood and BM samples implied increased frequencies of cells appearing in the most mature hemoglobin subunit-expressing erythroid clusters (EP-II/EP-IV) following HU, which is in line with previous studies linking HU treatment to erythroid differentiation, in part via nitric oxide-induced *GATA1* activation, in other contexts.[13,26] A recent study by Krishnan *et al.* suggested that CML patients with a high degree of erythroid differentiation prior to TKI were more likely to respond favorably to TKI treatment; however, it did not take prior HU treatment into account.[27] In our limited cohort of patients, we did not observe any impact of erythroid progenitor cell frequencies on TKI responses. However, post-HU SPC comprised a higher proportion of erythroid progenitors than the paired samples obtained before treatment (blood pre | post HU: sCML14,

56.5% | 63.4%; sCML23, 53.0% | 66.9%), highlighting the need to take HU status into account in such analyses.

SPC from post-HU samples contained an increased fraction of cells in S/G2/M phase as compared to those obtained before HU treatment. This shift was observed among relatively mature erythroid progenitors as well among the most immature CML cells in both paired and unpaired samples. While an increased proportion of cells in S/G2/M phase is typically reflective of cells in active proliferation, it may in this instance instead derive from HU-induced reduction in the availability of deoxyribonucleotides and consequent accumulation of cells in S phase.[28,29] Accordingly, following HU treatment we observed a larger fraction of 2n $Ki67^{hi}$ cells suggesting transcriptional commitment to S phase but failure to replicate DNA. These assumptions are also in line with a study by Behbehani *et al.*[30] showing that though incorporation of 5-iodo-2´-deoxyuridine in primary acute myeloid leukemia cells was strongly reduced after *in vivo* HU treatment, the fraction of cells with signs of active cycling was increased rather than decreased. The reduced proportions of SPC observed in CML blood and BM following HU treatment, as well as the elevated expression of checkpoint-related genes (*CCNE1, CDKN2C, FANCI* and *TRIP13)* in the expanding cell populations, speak in favor of this mechanism.

At the LSC level, a consistent and significant shift towards stem cells expressing S/G2/M phase markers (LSC-II) was observed after HU, while the fraction of the more quiescent LSC-I subset was reduced. Interestingly, the patients with the highest proportion of LSC-II cells after HU treatment did not respond optimally to initial TKI treatment and required treatment with second or third-line ponatinib treatment to attain CCyR. As the total amount of LSC (LSC-I + LSC-II) was unaltered by HU treatment (data not shown), we speculate that

LSC-I and LSC-II are subgroups within the LSC population that differ in cell cycle stage. Upon HU treatment, the LSC may remain in the S/G2/M phase to a higher extent due to slower or halted progression through the cell cycle, reflected by enhanced proportions of LSC-II cells in the post-HU samples. Having a larger proportion of LSC in S/G2/M phase upon HU treatment may thus identify patients with more proliferative LSC clones, which may explain their slower and/or less straightforward normalization of hematopoiesis upon TKI treatment.

While no randomized trials assessing long-term outcomes of early-phase HU treatment have been performed, previous studies suggest that HU treatment for more than 6 months prior to TKI therapy is associated with a significantly lower rate of MMR[31] and that patients with HU pre-treatment may have poorer TKI responses[8]. Also, combined HU/TKI treatment reportedly does not provide additional clinical benefit compared with TKI treatment alone.[32,33] However, since HU is mainly given to patients with larger disease burden, it is difficult to draw conclusions regarding its effects on disease outcome in the absence of randomized comparisons. Our results showed an increased proportion of LSC-II following HU treatment, and a non-optimal TKI response in the two patients with the highest LSC-II content. Although the small sample size precludes conclusions regarding the role of LSC-II for TKI responsiveness, these findings point towards analysis of LSC-II/LSC-I ratios in a larger number of HU-treated patients to elucidate long-term effects on CML outcome.

In conclusion, these results imply that short-term HU therapy enhances erythroid progenitor cell differentiation and alters the characteristics of CML LSC. Hence, HU status should be considered in future analyses addressing the SPC compartment in CML.

References

1. Hungerford DA. Chromosome studies in human leukemia. 1. Acute leukemia in children. J Natl Cancer Inst. 1961;27:983-1011.

2. Quintás-Cardama A, Cortes JE. Chronic myeloid leukemia: diagnosis and treatment. Mayo Clin Proc. 2006;81(7):973-988.

3. Cortes JE, Saglio G, Kantarjian HM, et al. Final 5-year study results of DASISION: the dasatinib versus imatinib study in treatment-naïve chronic myeloid leukemia patients trial. J Clin Oncol. 2016;34(20):2333-2340.

4. Hochhaus A, Saglio G, Hughes TP, et al. Long-term benefits and risks of frontline nilotinib vs imatinib for chronic myeloid leukemia in chronic phase: 5-year update of the randomized ENESTnd trial. Leukemia. 2016;30(5):1044-1054.

5. Hehlmann R, Heimpel H, Hasford J, et al. Randomized comparison of interferon-alpha with busulfan and hydroxyurea in chronic myelogenous leukemia. The German CML Study Group. Blood. 1994;84(12):4064-4077.

6. Wolfe HR, Rein LAM. The Evolving Landscape of Frontline Therapy in Chronic Phase Chronic Myeloid Leukemia (CML). Curr Hematol Malig Rep. 2021;16(5):448-454.

7. Hochhaus A, Baccarani M, Silver RT, et al. European LeukemiaNet 2020 recommendations for treating chronic myeloid leukemia. Leukemia. 2020;34(4):966-984.

8. Kockerols CCB, Geelen I, Levin MD, et al. The use of hydroxyurea pretreatment in chronic myeloid leukemia in the current tyrosine kinase inhibitor era. Haematologica. 2022;107(8):1940-1943.

9. Kennedy BJ. Hydroxyurea therapy in chronic myelogenous leukemia. Cancer. 1972;29(4):1052-1056.

10. Cokic VP, Smith RD, Beleslin-Cokic BB, et al. Hydroxyurea induces fetal hemoglobin by the nitric oxide-dependent activation of soluble guanylyl cyclase. J Clin Invest. 2003;111(2):231-239.

11. Lou TF, Singh M, Mackie A, Li W, Pace BS. Hydroxyurea generates nitric oxide in human erythroid cells: mechanisms for gamma-globin gene activation. Exp Biol Med. 2009;234(11):1374-1382.

12. Subotički T, Mitrović Ajtić O, Djikić D, et al. Nitric Oxide Mediation in Hydroxyurea and Nitric Oxide Metabolites' Inhibition of Erythroid Progenitor Growth. Biomolecules. 2021;11(11):1562.

13. Ikuta T, Sellak H, Odo N, Adekile AD, Gaensler KML. Nitric Oxide-cGMP Signaling Stimulates Erythropoiesis through Multiple Lineage-Specific Transcription Factors: Clinical Implications and a Novel Target for Erythropoiesis. PLoS One. 2016;11(1):e0144561.

14. Hümmer J, Kraus S, Brändle K, Lee-Thedieck C. Nitric Oxide in the Control of the in vitro Proliferation and Differentiation of Human Hematopoietic Stem and Progenitor Cells. Front Cell Dev Biol. 2020;8:610369.

15. Holyoake T, Jiang X, Eaves C, Eaves A. Isolation of a highly quiescent subpopulation of primitive leukemic cells in chronic myeloid leukemia. Blood. 1999;94(6):2056-2064.

16. Vetrie D, Helgason GV, Copland M. The leukaemia stem cell: similarities, differences and clinical prospects in CML and AML. Nat Rev Cancer. 2020;20(3):158-173.

17. Houshmand M, Simonetti G, Circosta P, et al. Chronic myeloid leukemia stem cells. Leukemia. 2019;33(7):1543-1556.

18. Mojtahedi H, Yazdanpanah N, Rezaei N. Chronic myeloid leukemia stem cells: targeting therapeutic implications. Stem Cell Res Ther. 2021;12(1):603.

19. Nilsson MS, Komic H, Gustafsson J, et al. Multiomic single-cell analysis identifies von Willebrand factor and TIM3-expressing BCR-ABL1+ CML stem cells. bioRxiv. 2023 Sep 17. doi:10.1101/2023.09.14.557507 [preprint, not peer-reviewed].

20. Hao Y, Hao S, Andersen-Nissen E, et al. Integrated analysis of multimodal single-cell data. Cell. 2021;184(13):3573-3587.

21. Gutiérrez L, Caballero N, Fernández-Calleja L, Karkoulia E, Strouboulis J. Regulation of GATA1 levels in erythropoiesis. IUBMB Life. 2020;72(1):89-105.

22. Buscemi G, Carlessi L, Zannini L, et al. DNA damage-induced cell cycle regulation and function of novel Chk2 phosphoresidues. Mol Cell Biol. 2006;26(21):7832-7845.

23. Warfvinge R, Geironson L, Sommarin MNE, et al. Single-cell molecular analysis defines therapy response and immunophenotype of stem cell subpopulations in CML. Blood. 2017;129(17):2384-2394.

24. Cortes J, Quintás-Cardama A, Kantarjian HM. Monitoring molecular response in chronic myeloid leukemia. Cancer. 2011;117(6):1113-1122.

25. Cortes J, Kantarjian H. How I treat newly diagnosed chronic phase CML. Blood. 2012;120(7):1390-1397.

26. Gui CY, Jiang C, Qian RL. [The role of hydroxyurea in globin gene expression and cell differentiation of human erythroleukemia cell line]. Shi Yan Sheng Wu Xue Bao. 1997;30(4):375-382.

27. Krishnan V, Schmidt F, Nawaz Z, et al. A single-cell atlas identifies pretreatment features of primary imatinib resistance in chronic myeloid leukemia. Blood. 2023;141(22):2738-2755.

28. Kapor S, Vukotić M, Subotički T, et al. Hydroxyurea Induces Bone Marrow Mesenchymal Stromal Cells Senescence and Modifies Cell Functionality In Vitro. J Pers Med. 2021;11(11):1048.

29. Apraiz A, Mitxelena J, Zubiaga A. Studying Cell Cycle-regulated Gene Expression by Two Complementary Cell Synchronization Protocols. J Vis Exp. 2017;(124):55745.

30. Behbehani GK, Samusik N, Bjornson ZB, Fantl WJ, Medeiros BC, Nolan GP. Mass Cytometric Functional Profiling of Acute Myeloid Leukemia Defines Cell-Cycle and Immunophenotypic Properties That Correlate with Known Responses to Therapy. Cancer Discov. 2015;5(9):988-1003.

31. Rinaldi I, Reksodiputro AH, Jusman SW, et al. Longer Hydroxyurea Administration Prior to Imatinib Mesylate is Risk Factor for Unsuccessful Major Molecular Response in Chronic-Phase Chronic Myeloid Leukemia: Possibility of P-Glycoprotein Role. Asian Pac J Cancer Prev. 2019;20(12):3689-3695.

32. Lange T, Niederwieser C, Gil A, et al. No advantage of Imatinib in combination with hydroxyurea over Imatinib monotherapy: a study of the East German Study Group (OSHO) and the German CML study group. Leuk Lymphoma. 2020;61(12):2821-2830.

33. Thiesing JT, Ohno-Jones S, Kolibaba KS, Druker BJ. Efficacy of STI571, an abl tyrosine kinase inhibitor, in conjunction with other antileukemic agents against bcr-abl-positive cells. Blood. 2000;96(9):3195-3199.

Figure legends

Figure 1. Hydroxyurea alters CML stem and progenitor cell frequencies in blood and bone marrow. (A) Samples and representative gating strategy used for flow cytometry analysis of the effects of hydroxyurea (HU) on stem and progenitor cells in chronic phase CML (CP-CML) patient blood and bone marrow (BM). **(B-C)** Effect of HU treatment on the frequency of **(B)** $CD34^+$ and **(C)** $CD34^+CD38^-$ cells in CML patient blood (PBMC) and BM (BM MNC) samples obtained before (n=7) or after 4-19 (median 9) days of HU (n=10). Statistics by Mann-Whitney. Error bars represent SEM.

Figure 2. Workflow for proteo-transcriptomic CITE-seq analysis of $CD14^-CD34^+$ stem and progenitor cells from chronic phase CML patients. CP-CML, chronic phase CML; HU, hydroxyurea; FACS, fluorescence-activated cell sorting: CITE-seq, Cellular Indexing of Transcriptomes and Epitopes by Sequencing.

Figure 3. Hydroxyurea stimulates erythroid maturation among CML stem and progenitor cells. Analysis of $CD14^-CD34^+$ CITE-seq data from two chronic phase CML patients (sCML14 and sCML23) with peripheral blood samples collected before and after hydroxyurea (HU) treatment and paired bone marrow (BM) samples obtained after treatment. After HU samples were collected after 7 and 9 days of treatment, respectively. **(A)** Uniform Manifold Approximation and Projection (UMAP) representation of an mRNA-based clustering analysis of cells from all six samples (21,044 cells) after integration by patient. EBMP, eosinophil/basophil/mast cell progenitors; EP, erythroid progenitors; LSC, leukemic stem cells; MP, myeloid progenitors; MEP, megakaryocytic/erythroid progenitors; MKP, megakaryocytic progenitors. **(B-C)** Scaled expression of selected marker **(B)** genes and **(C)**

proteins used for annotation of each UMAP cluster. Colors represent average expression level and dot size percentage of cells expressing the specific marker within the cluster. **(D)** Distribution of cells from each analyzed sample across clusters. In the left panel, all cell populations are subsampled to 1,286 cells. **(E)** Feature plots of expression of the genes *HBA1***,** *HBA2* and *HBB* among cells in the UMAP. Red indicates maximum expression and blue minimum expression.

Figure 4. Hydroxyurea induces transcriptional and proteomic features associated with S/G2/M phases of the cell cycle among CML stem and progenitor cells. Graphs **(A-C)** show results from analysis of aspects of cell cycling in the paired CITE-seq dataset (n=2 patients; blood samples obtained before and after, and bone marrow (BM) samples obtained after, 7-9 days of hydroxyurea (HU) treatment). **(A-B)** Proportion of cells in G0/G1 vs S/G2/M phase **(A)** in each UMAP cluster and **(B)** among $CD14^{-}CD34^{+}$ cells in each sample. **(C)** Feature plot of *CHEK2* gene expression among cells in the UMAP. Red indicates maximum expression and blue minimum expression. Graphs **(D-F)** show results from flow cytometry cell cycle analysis of paired blood samples collected before and after HU treatment (n=3 patients). **(D)** Representative gating strategy. **(E-F)** Effect of HU treatment on frequencies of **(E)** cells in S/G2/M phase and **(F)** 2n $Ki67^{hi}$ cells among CML stem and progenitor cells. MNC, mononuclear cells.

Figure 5. CITE-seq analysis of unpaired CML bone marrow samples confirms enhanced frequencies of stem and progenitor cells with features of erythroid maturation and S/G2/M cell cycle phases after hydroxyurea treatment. Analysis of CITE-seq data from a previously published dataset of $CD14^{-}CD34^{+}$ bone marrow (BM) cells collected from chronic phase CML (CP-CML) patients before (n=5; 17,968 cells) or after (n=7; 29,736 cells)

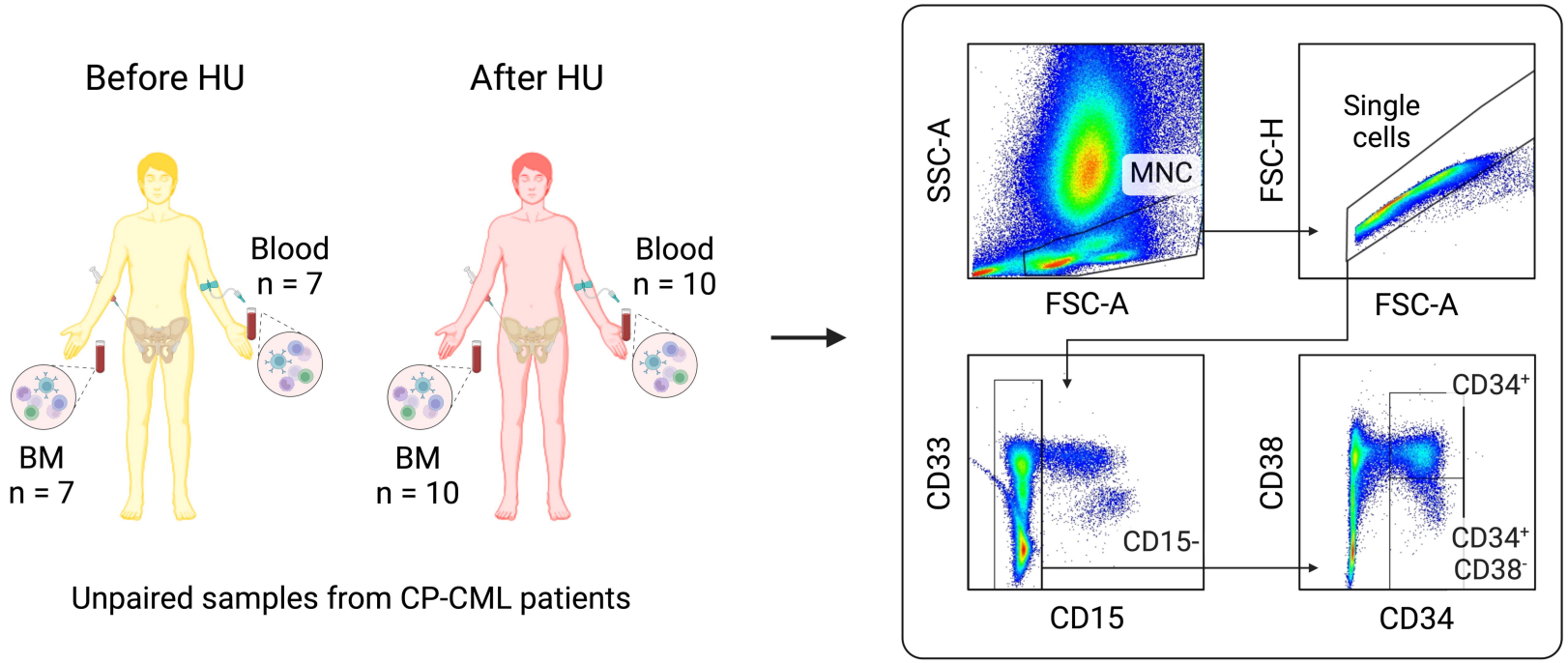
4-19 (median 9) days of hydroxyurea (HU) treatment (unpaired samples; Nilsson *et al.*[19]). **(A)** Visualization of distribution of cells obtained before and after HU across the Nilsson *et al.*[19] $CD14^{-}CD34^{+}$ UMAP. BCP, B cell progenitors; EBMP, eosinophil/basophil/mast cell progenitors; EP, erythroid progenitors; HSC, hematopoietic stem cells; LMP, lympho-myeloid progenitors; LSC, leukemic stem cells; MDP, monocyte/dendritic cell progenitors; MEP, megakaryocytic/erythroid progenitors; MKP, megakaryocytic progenitors; NP, neutrophil progenitors. **(B)** Percentage of cells in each erythroid UMAP cluster in samples obtained before and after HU. Statistics by Mann-Whitney. * $p<0.05$. **(C)** Feature plots of *HBA1***,** *HBA2*, *HBB* and *CHEK2* gene expression. Red indicates maximum expression and blue minimum expression. **(D)** Proportion of cells in G0/G1 vs S/G2/M phase in each erythroid cluster.

Figure 6. Hydroxyurea increases the fraction of cells in S/G2/M phase within the $CD14^{-}CD34^{+}CD38^{-}$ compartment in CML. Analysis of the $CD14^{-}CD34^{+}CD38^{-}$ compartment in the paired CITE-seq dataset (n=2 patients; 4,222 cells; blood and bone marrow (BM) samples obtained before and/or after 7-9 days of hydroxyurea (HU) treatment). **(A)** UMAP visualization of an mRNA expression-based clustering analysis of $CD14^{-}CD34^{+}CD38^{-}$ cells (left panel) and large-scale cluster annotation based on expressional patterns (right panel). MEP/EBMP, megakaryocytic/erythroid and eosinophil/basophil/mast cell progenitors; MP, myeloid progenitors; SC/MPP, stem cells and multipotent progenitors. **(B)** Proportion of $CD14^{-}CD34^{+}CD38^{-}$ cells in G0/G1 vs S/G2/M phase in each analyzed sample.

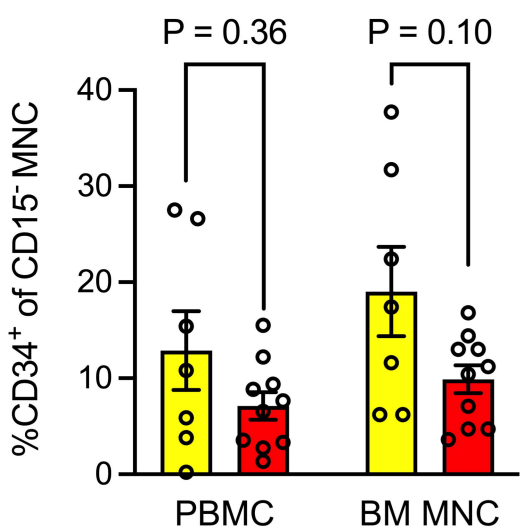
Figure 7. Hydroxyurea increases the fraction of cells in S/G2/M phase among the most immature leukemic stem cells. (A) Annotation of $CD14^{-}CD34^{+}CD38^{-}$ cells in the paired CITE-seq dataset based on cell label transfer from a previously published dataset of $CD14^{-}$

$CD34^+CD38^-$ healthy/chronic phase CML (CP-CML) bone marrow (BM) cells (Nilsson *et al.*[19]). EP, erythroid progenitors; EP/EBMP, erythroid and eosinophil/basophil/mast cell progenitors; HSC, hematopoietic stem cells; LMP, lympho-myeloid progenitors; LSC, leukemic stem cells; MDP, monocyte/dendritic cell progenitors; MEP, megakaryocytic/erythroid progenitors; MKP, megakaryocytic progenitors; NP, neutrophil progenitors. **(B)** Distribution of cells annotated as immature HSC-I, HSC-II, LSC-I, and LSC-II across the UMAP. **(C-D)** Proportion of **(C)** stem cells annotated as HSC (HSC-I and HSC-II), LSC-I and LSC-II in each analyzed sample (after HU samples were collected after 7-9 days of treatment), and **(D)** cells in G0/G1 vs S/G2/M phase in each stem cell subset. HU, hydroxyurea.

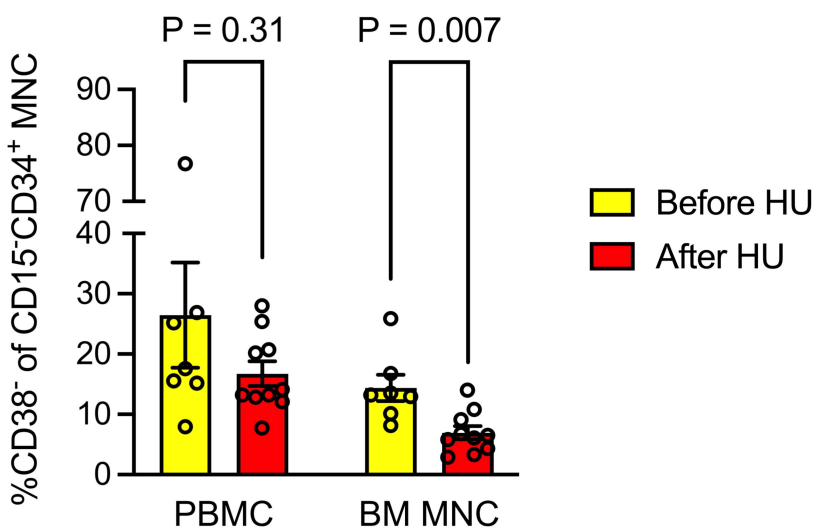
Figure 8. Induction of LSC-II may associate with poor response to tyrosine kinase inhibitor therapy. (A) Proportions of HSC, LSC-I and LSC-II in the previously published dataset of unpaired $CD14^-CD34^+CD38^-$ bone marrow (BM) cells obtained before (n=5) or after 4-19 (median 9) days of hydroxyurea (HU) treatment (n=7) (Nilsson *et al*[19]). HSC, hematopoietic stem cells; LSC, leukemic stem cells. **(B)** Frequency of LSC-II within the diagnosis LSC compartment for chronic phase CML (CP-CML) patients who did (poor responders, n=2) or did not (good responders, n=10) require second or third line ponatinib treatment to achieve a complete cytogenetic response (CCyR). Statistics by Mann-Whitney. * $p < 0.05$, ** $p < 0.01$.

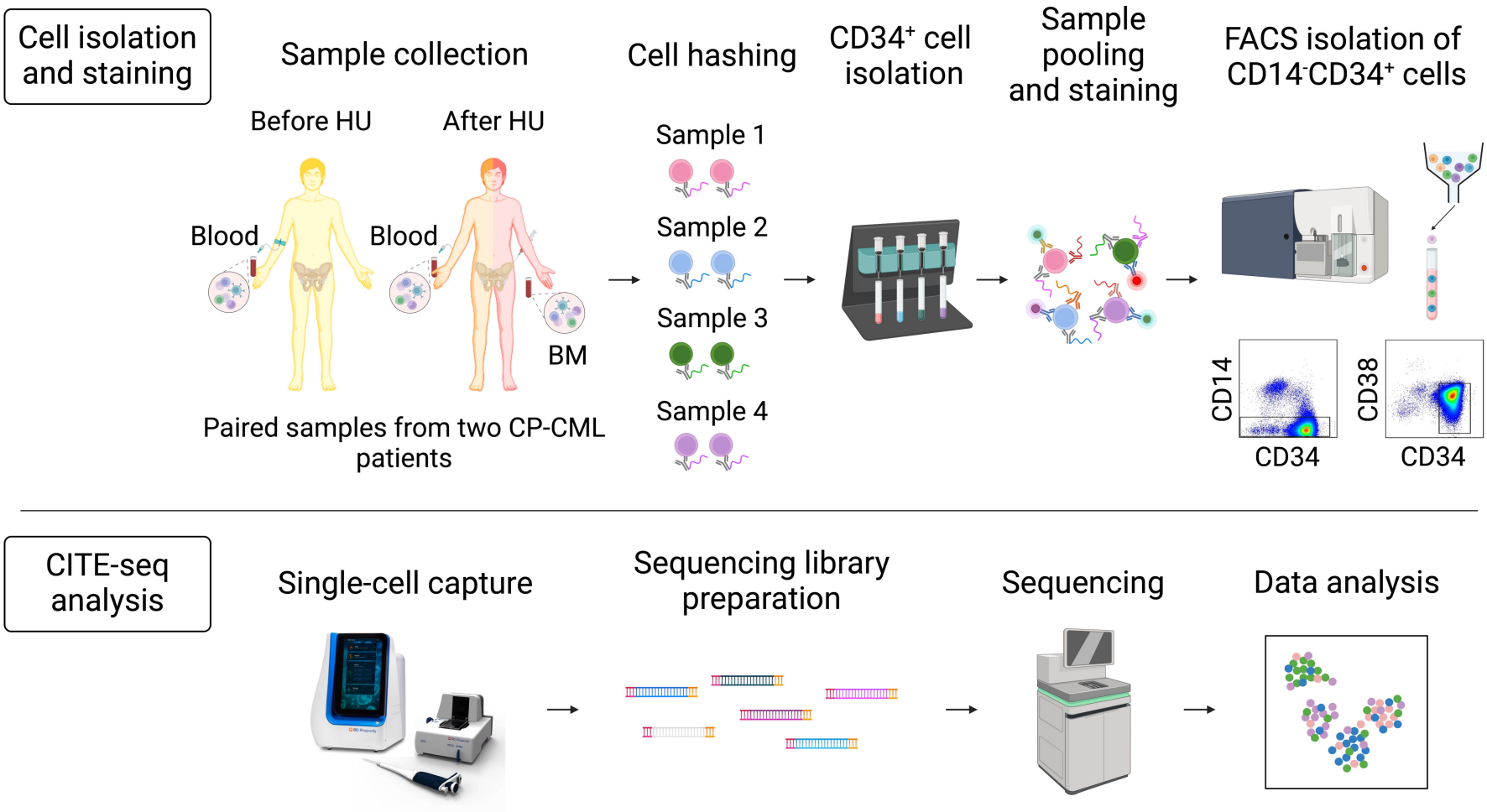

Before HU
After HU
Blood
n = 7
BM
n = 7
Blood
n = 10
BM
n = 10
Unpaired samples from CP-CML patients
SSC-A
FSC-A
MNC
FSC-H
FSC-A
Single
cells
CD33
CD15
CD15-
CD38
CD34
CD34⁺
CD34⁺
CD38⁻

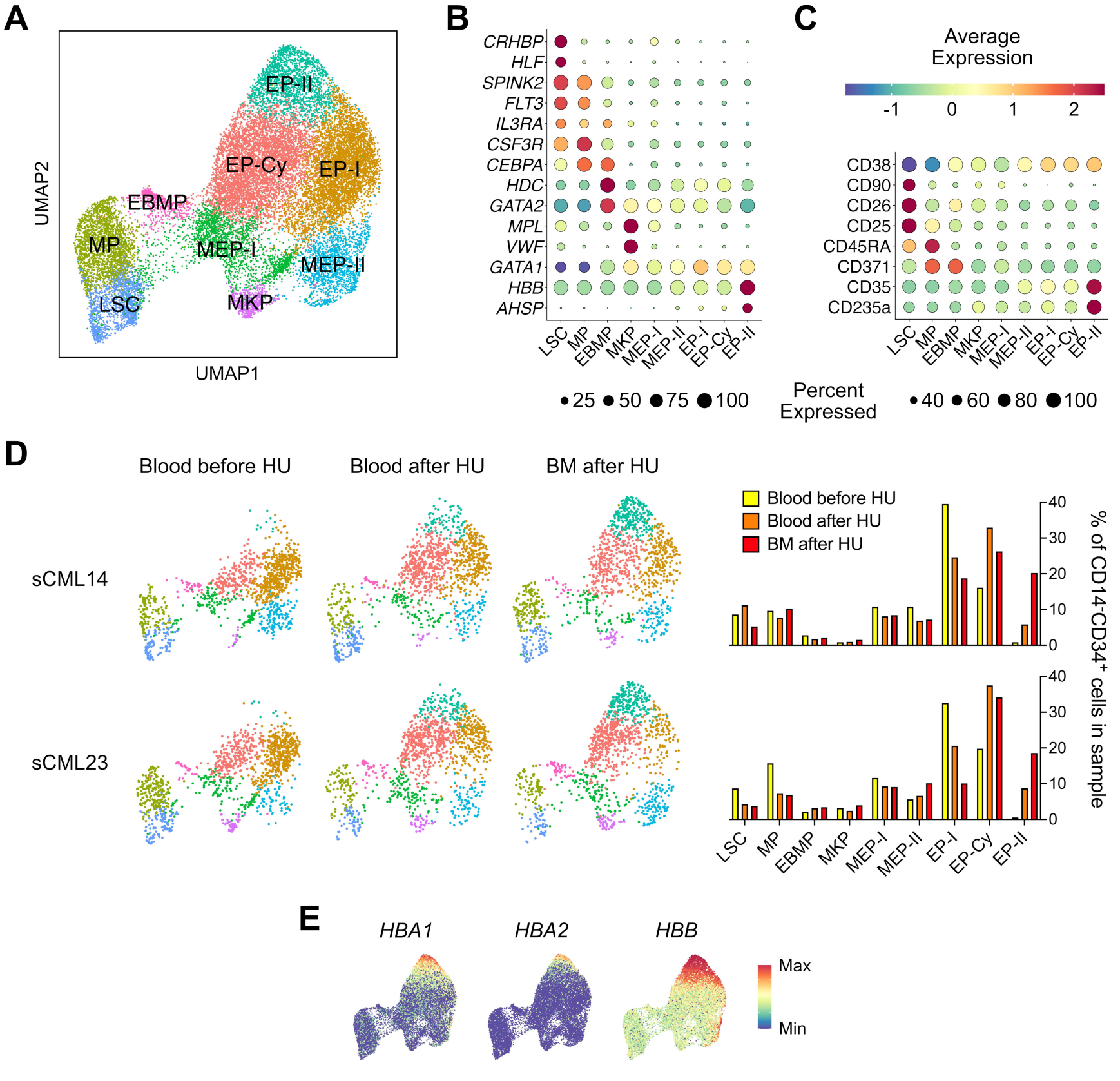
B

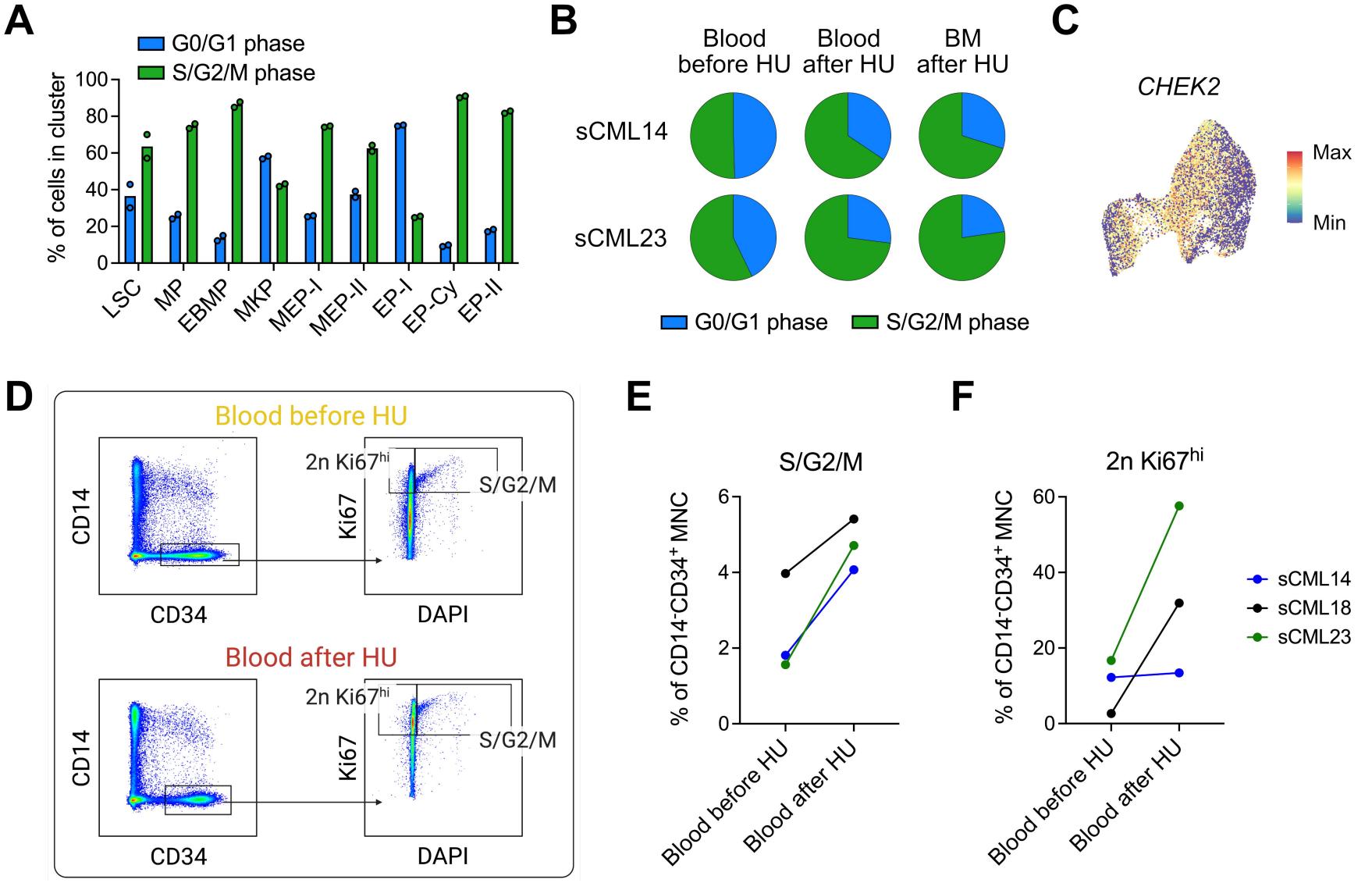

P = 0.36
P = 0.10
%CD34⁺ of CD15⁻ MNC
0
10
20
30
40
PBMC
BM MNC

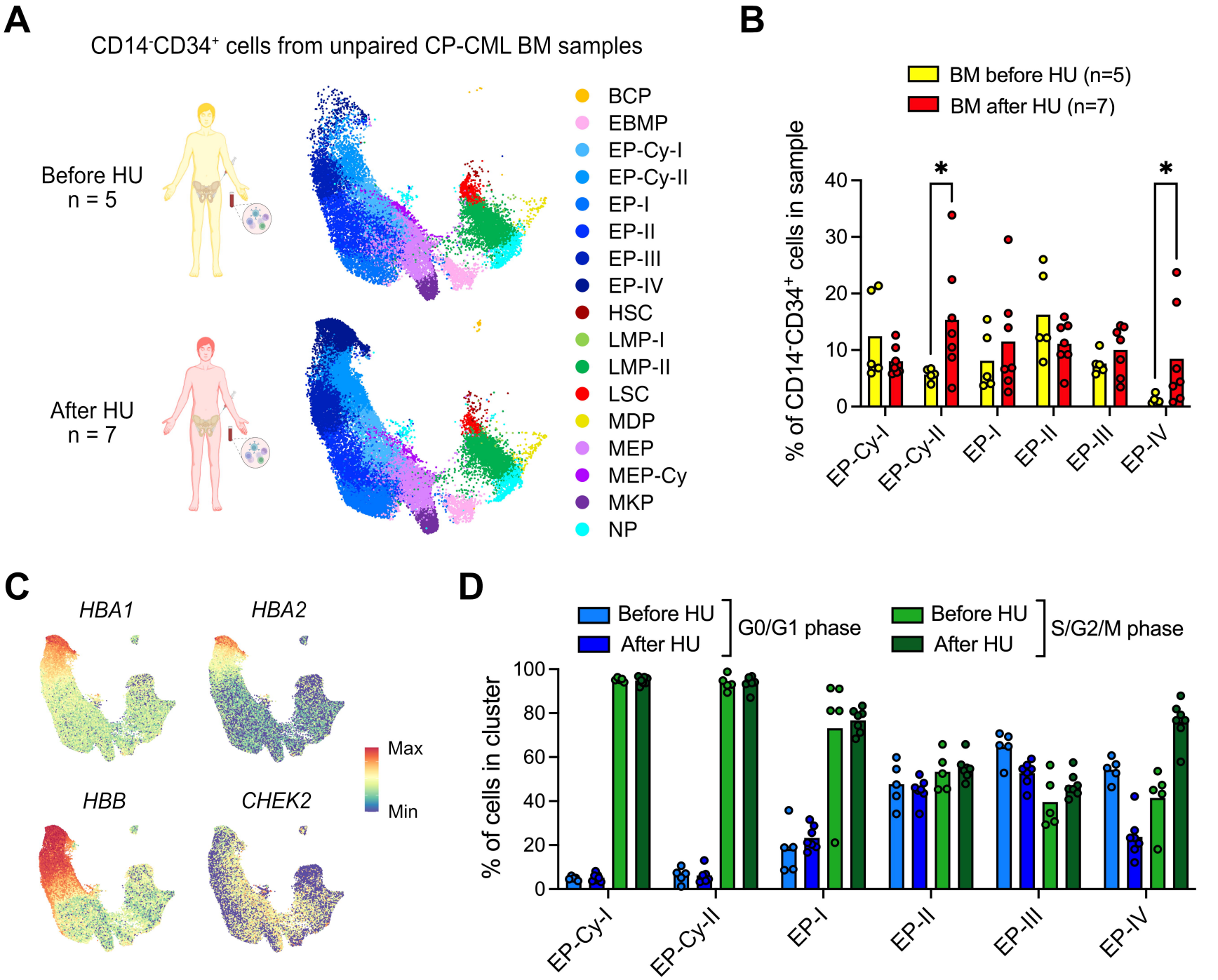
C

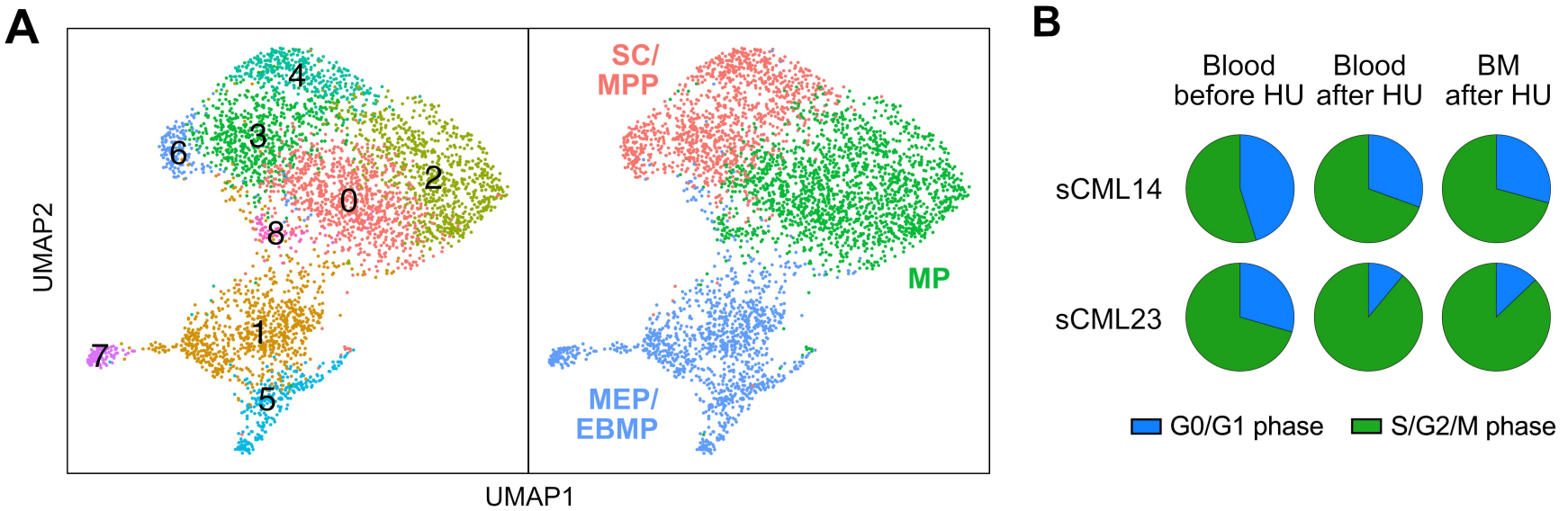

P = 0.31
P = 0.007
%CD38⁻ of CD15⁻CD34⁺ MNC
0
10
20
30
40
70
80
90
PBMC
BM MNC
Before HU
After HU

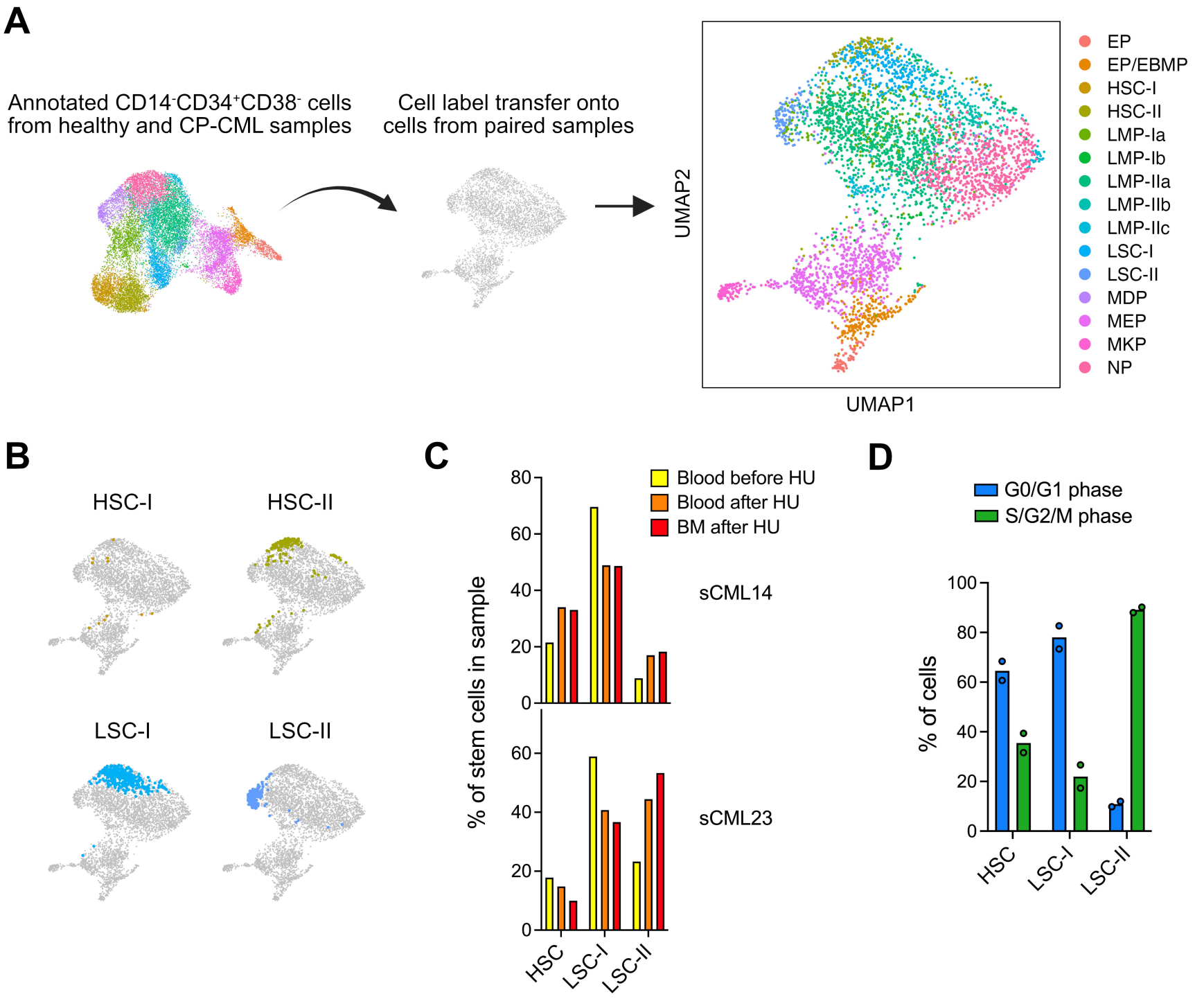

Cell isolation and staining
Sample collection
Before HU
After HU
Blood
Blood
BM
Paired samples from two CP-CML patients
Cell hashing
Sample 1
Sample 2
Sample 3
Sample 4
CD34+ cell isolation
Sample pooling and staining
FACS isolation of CD14-CD34+ cells
CD14
CD34
CD38
CD34
CITE-seq analysis
Single-cell capture
Sequencing library preparation
Sequencing
Data analysis

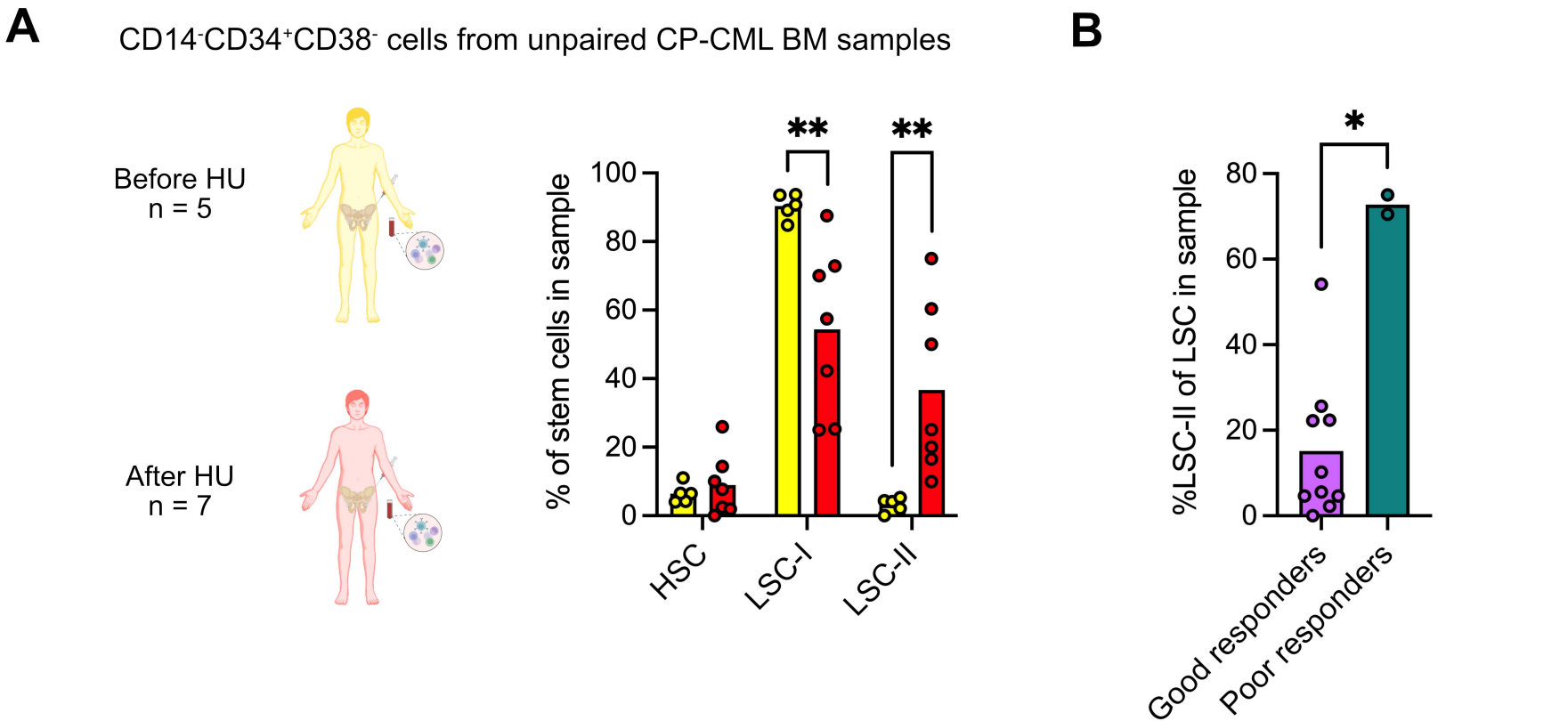

A
UMAP2
UMAP1
EP-II
EP-Cy
EP-I
EBMP
MP
MEP-I
MEP-II
LSC
MKP
B
CRHBP
HLF
SPINK2
FLT3
IL3RA
CSF3R
CEBPA
HDC
GATA2
MPL
VWF
GATA1
HBB
AHSP
LSC MP EBMP MKP MEP-I MEP-II EP-I EP-Cy EP-II
25 50 75 100
C
Average Expression
-1 0 1 2
CD38
CD90
CD26
CD25
CD45RA
CD371
CD35
CD235a
LSC MP EBMP MKP MEP-I MEP-II EP-I EP-Cy EP-II
Percent Expressed
40 60 80 100
D
Blood before HU
Blood after HU
BM after HU
sCML14
sCML23
Blood before HU
Blood after HU
BM after HU
% of CD14-CD34+ cells in sample
0 10 20 30 40
LSC MP EBMP MKP MEP-I MEP-II EP-I EP-Cy EP-II
E
HBA1
HBA2
HBB
Max
Min


A
G0/G1 phase
S/G2/M phase
% of cells in cluster
100
80
60
40
20
0
LSC
MP
EBMP
MKP
MEP-I
MEP-II
EP-I
EP-Cy
EP-II
B
Blood before HU
Blood after HU
BM after HU
sCML14
sCML23
G0/G1 phase
S/G2/M phase
C
CHEK2
Max
Min
D
Blood before HU
CD14
CD34
2n Ki67hi
Ki67
S/G2/M
DAPI
Blood after HU
CD14
CD34
2n Ki67hi
Ki67
S/G2/M
DAPI
E
S/G2/M
% of CD14-CD34+ MNC
6
4
2
0
Blood before HU
Blood after HU
F
2n Ki67hi
% of CD14-CD34+ MNC
60
40
20
0
Blood before HU
Blood after HU
sCML14
sCML18
sCML23


A
CD14-CD34+ cells from unpaired CP-CML BM samples
Before HU
n = 5
After HU
n = 7
BCP
EBMP
EP-Cy-I
EP-Cy-II
EP-I
EP-II
EP-III
EP-IV
HSC
LMP-I
LMP-II
LSC
MDP
MEP
MEP-Cy
MKP
NP
B
BM before HU (n=5)
BM after HU (n=7)
% of CD14-CD34+ cells in sample
0
10
20
30
40
*
*
EP-Cy-I
EP-Cy-II
EP-I
EP-II
EP-III
EP-IV
C
HBA1
HBA2
HBB
CHEK2
Max
Min
D
Before HU
After HU
G0/G1 phase
Before HU
After HU
S/G2/M phase
% of cells in cluster
0
20
40
60
80
100
EP-Cy-I
EP-Cy-II
EP-I
EP-II
EP-III
EP-IV


A
UMAP2
UMAP1
0
1
2
3
4
5
6
7
8
SC/
MPP
MP
MEP/
EBMP
B
Blood
before HU
Blood
after HU
BM
after HU
sCML14
sCML23
G0/G1 phase
S/G2/M phase


A
Annotated CD14-CD34+CD38- cells from healthy and CP-CML samples
Cell label transfer onto cells from paired samples
UMAP2
UMAP1
EP
EP/EBMP
HSC-I
HSC-II
LMP-Ia
LMP-Ib
LMP-IIa
LMP-IIb
LMP-IIc
LSC-I
LSC-II
MDP
MEP
MKP
NP
B
HSC-I
HSC-II
LSC-I
LSC-II
C
Blood before HU
Blood after HU
BM after HU
% of stem cells in sample
sCML14
sCML23
HSC
LSC-I
LSC-II
D
G0/G1 phase
S/G2/M phase
% of cells
HSC
LSC-I
LSC-II


A
CD14-CD34+CD38- cells from unpaired CP-CML BM samples
Before HU
n = 5
After HU
n = 7
% of stem cells in sample
100
80
60
40
20
0
**
**
HSC
LSC-I
LSC-II
B
%LSC-II of LSC in sample
80
60
40
20
0
*
Good responders
Poor responders

Supplementary Material

Supplementary Methods

Immunophenotypic analysis of blood and BM samples

Peripheral blood and bone marrow mononuclear cells (PBMC and BM MNC) were stained with a panel consisting of anti-CD15-BV786 (HI98), anti-CD33-BV711 (WM53), anti-CD34-BV421 (581) and anti-CD38-BV510 (HIT2) antibodies (BD Biosciences) and acquired on a five laser BD LSR Fortessa or a BD FACS Aria Fusion (BD Biosciences). Analyses were performed using FlowJo (v10 or later, BD Biosciences).

Cell cycle analysis

Blood samples obtained from three patients before and after HU treatment were thawed and stained for cell surface markers using anti-CD14-PE-Cy7 and anti-CD34-PE (clones M5E2 and 8G12, respectively; BD Biosciences). Cells were washed and fixed with 2% paraformaldehyde (ThermoFisher), and permeabilized with permeabilization buffer (PBS with 0.1% EDTA, 0.5% BSA, 0.1% saponin and 3% FCS). Thereafter, cells were stained with 1µg/ml DAPI (Invitrogen) and anti-Ki67-Alexa Fluor 647 (B56, BD Biosciences) for 30 minutes at room temperature. Cells were acquired on a five laser BD LSR Fortessa (BD Biosciences) at a low flow rate (100-200 events/second) and DAPI fluorescence was analyzed on the linear scale. Analyses were performed using FlowJo (v10.8.1, BD Biosciences).

CITE-seq library preparation

PBMC and BM MNC obtained before or after HU were thawed and labeled with sample-specific barcoded cell hashing antibodies ('Sample Tags', BD Biosciences). MACS MicroBead (Miltenyi Biotec) $CD34^+$ cell selection was performed on samples that had not undergone this enrichment prior to cryopreservation. Cells were pooled to achieve 1:1 ratio between cells from

different samples, and thereafter co-labeled with unconjugated, fluorescent, and oligo-conjugated BD AbSeq antibodies (BD Biosciences) (Tables S3 and S4). $CD14^{-}CD34^{+}$ MNC were isolated from the sample pool using a BD FACSAria Fusion Cell sorter (BD Biosciences). CITE-seq sequencing libraries were generated from the isolated $CD14^{-}CD34^{+}$ cells using the BD Rhapsody Single-Cell Analysis System (BD Biosciences) as described in Nilsson *et al*.[1]

Sequencing

mRNA, AbSeq and Sample tag libraries were sequenced at SNP&SEQ Technology Platform (Uppsala, Sweden; part of the National Genomics Infrastructure (NGI) Sweden and Science for Life Laboratory). Sequencing was performed on a NovaSeq 6000 instrument (Illumina) using a S4-200 v1.5 flow cell (Illumina), a custom read setup of 64-8-0-74 (R1-i7-i5-R2) and 20% PhiX control v3 (Illumina).

Data analysis

After initial quality control, mRNA expression data were normalized using the SCTransform function[2], and differences between proliferating cells (i.e. G2/M vs S phase) based on Seurat's CellCycleScoring function were regressed out of the data. To remove patient-specific clustering from the analysis presented in Figure 3, mRNA data were integrated by patient using reciprocal PCA (rPCA) on the top 22 PCs (k.anchor = 5). SCT or integrated assays were used for principal component analysis (PCA) followed by K-nearest neighbor (KNN) graph construction using the top 22-23 PCs, and Louvain algorithm-based clustering. Uniform Manifold Approximation and Projection (UMAP) was employed for clustering visualization.

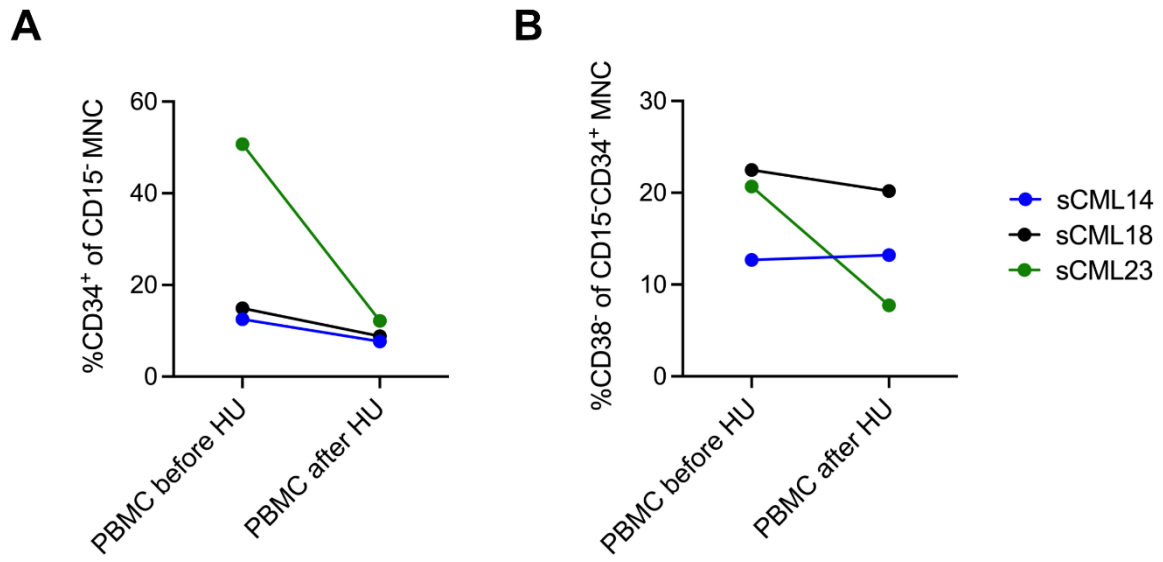
Selection of $CD38^{-}$ cells for reclustering was performed based on centered log ratio (CLR) normalized CD38 protein expression data using Seurat's CellSelector function. A previously

annotated $CD14^{-}CD34^{+}CD38^{-}$ CML dataset (Nilsson *et al.*[1]) was used as a reference to annotate $CD14^{-}CD34^{+}CD38^{-}$ cells by cell label transfer using Seurat's FindTransferAnchors and TransferData functions.

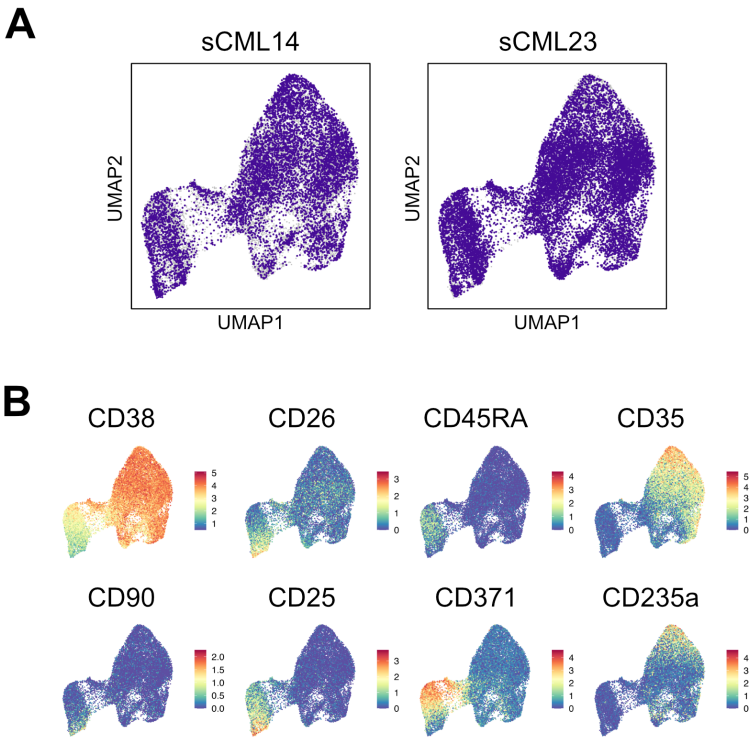
Visualization and differential expression analyses utilized log-normalized RNA or CLR normalized AbSeq counts. UMAP cluster annotations were manually assigned based on differentially expressed genes and proteins from Seurat's FindAllMarkers function as well as feature and dot plot visualizations of expression patterns across the UMAPs. Visualizations were created using ggplot2 (v. 3.4.0), SCPubr (v. 1.0.1), RColorBrewer (v. 1.1.3), GraphPad Prism (v. 9.4.1) and BioRender.com.

The pySCENIC (v. 0.12.1) package in Python (v. 3.1.0) was utilized to address genes regulated by the GATA1 transcription factor. The org was set to hgnc and linked to the RcisTarget databases. A database containing gene promoters (hg38_500bp_up_100bp_down_full_tx_v10_clust.genes_vs_motifs.rankings.feather) was used and only motifs with a Normalized Enrichment Score (NES) > 3.0 were considered as significantly enriched in the transcription factors (TF) module. Downstream analyses were performed using *Grnboost2*, followed by using *pyscenic ctx* and *pyscenic aucell*, using mask dropouts (cells in which expression of either TF or target gene is 0). As input data matrices had been subjected to quality control (QC) in the R analysis pipeline, this aspect was not addressed during SCENIC analysis.

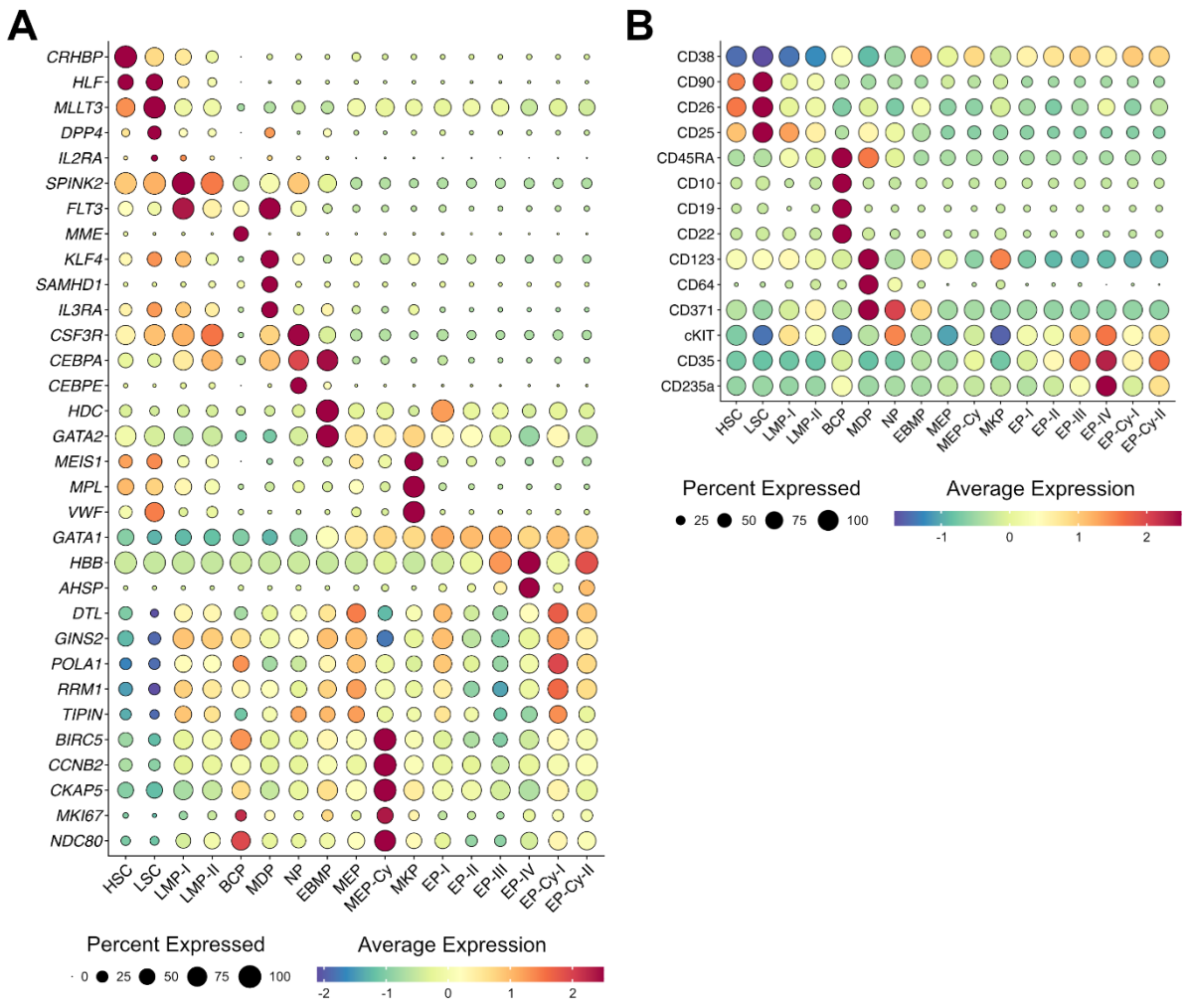
Supplementary Figures



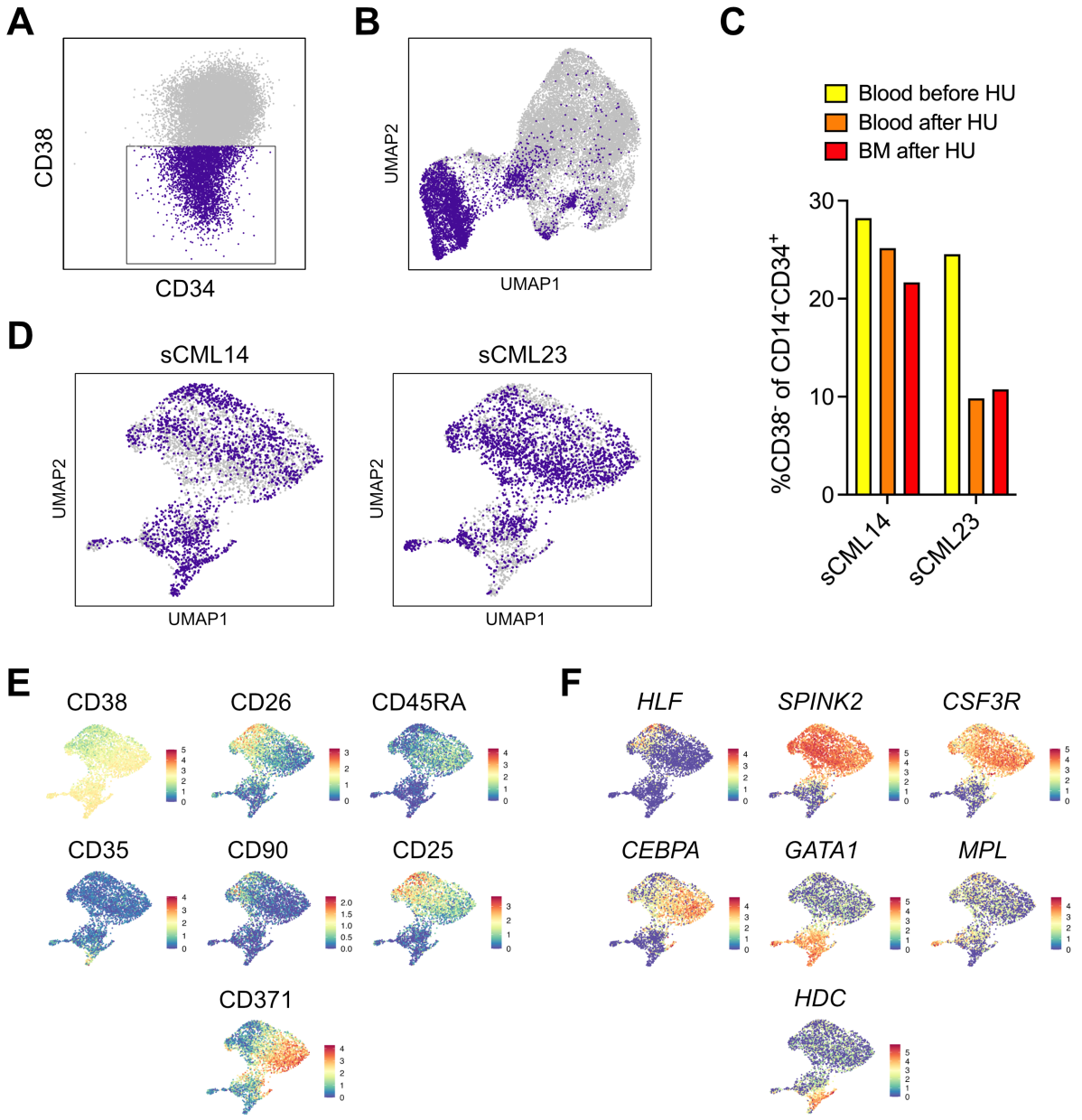
Supplementary Figure 1. Effect of hydroxyurea (HU) treatment on the frequency of **(A)** $CD34^+$ and **(B)** $CD34^+CD38^-$ cells in paired CML patient peripheral blood (PBMC) samples obtained before and after HU treatment. Patients sCML14, sCML18 and sCML23 had received HU treatment for 7, 4, and 9 days respectively, prior to 'after HU' sampling. Samples highlighted in blue (sCML14) and green (sCML23) were subsequently subjected to CITE-seq analysis. MNC, mononuclear cells.



Supplementary Figure 2. Characteristics of cells in the UMAP comprising $CD14^{-}CD34^{+}$ cells from blood and BM samples obtained from patients sCML14 and sCML23 before and after 7-9 days of hydroxyurea (HU) treatment. **(A)** Distribution of cells from each patient across the UMAP. Cells from samples obtained from patient sCML14 are highlighted in purple in the left panel, and those from patient sCML23 in the right panel. **(B)** Feature plots of expression of indicated marker proteins among cells in the UMAP. Red indicates maximum expression and blue minimum expression.

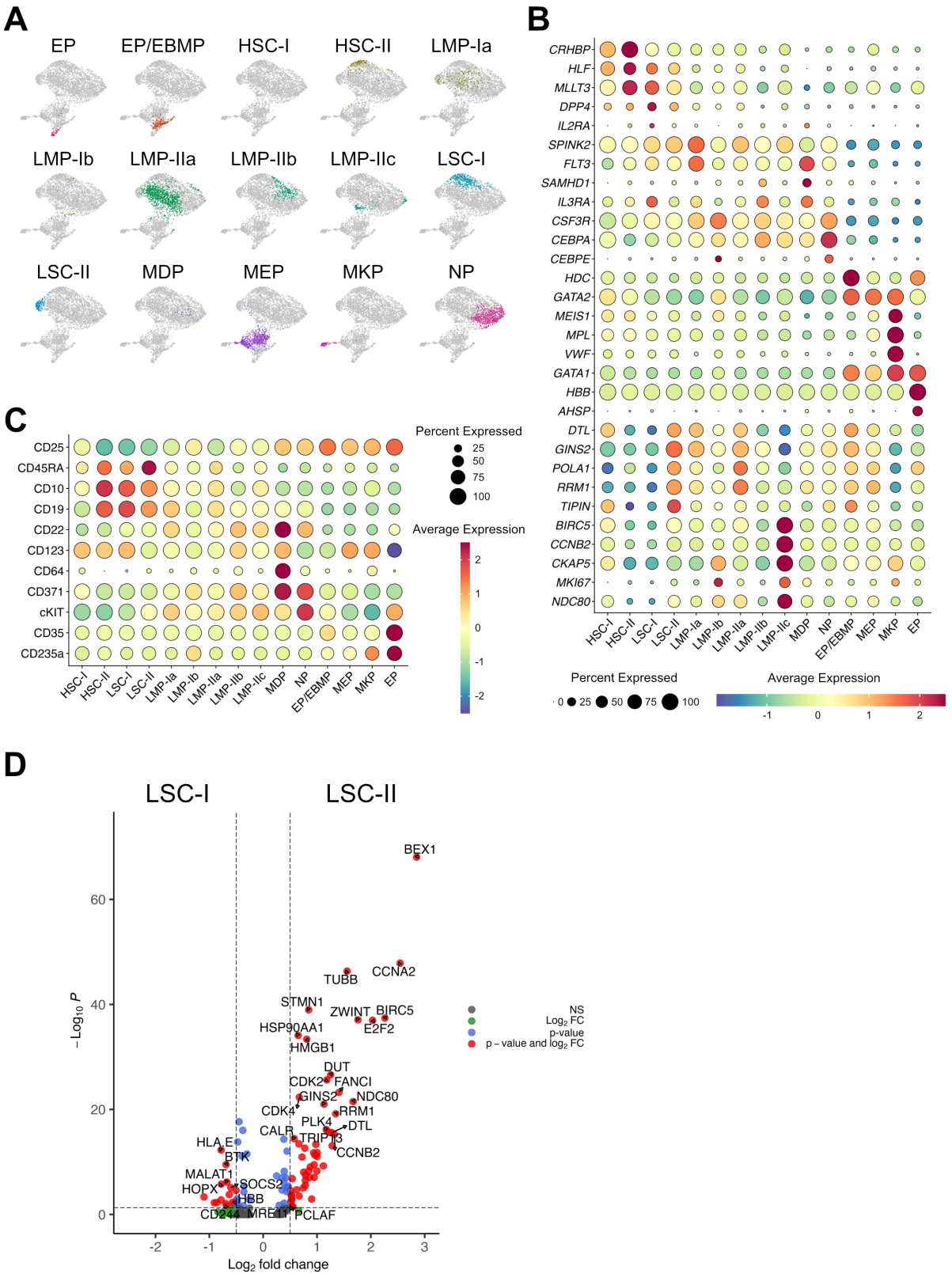


Supplementary Figure 3. Scaled expression of **(A)** genes and **(B)** proteins used to confirm the cell type annotation in Figure 5. Cells in the analysis derive from a previously published dataset of $CD14^{-}CD34^{+}$ bone marrow cells obtained before (n=5) or after (n=7) hydroxyurea (HU) treatment (unpaired samples). Colors represent average expression level and dot size percentage of cells expressing the specific marker. For further detail on the original annotation of these cell types, the reader is referred to Nilsson *et al.*[1] BCP, B cell progenitors; EBMP, eosinophil/basophil/mast cell progenitors; EP, erythroid progenitors; HSC, hematopoietic stem cells; LMP, lympho-myeloid progenitors; LSC, leukemic stem cells; MDP, monocyte/dendritic cell progenitors; MEP, megakaryocytic/erythroid progenitors; MKP, megakaryocytic progenitors; NP, neutrophil progenitors.



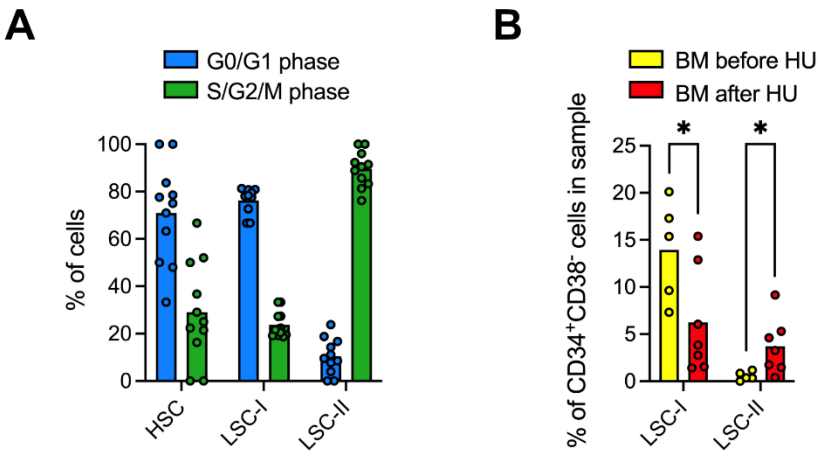
Supplementary Figure 4. Generation and characterization of the UMAP comprising $CD34^{+}CD38^{-}$ cells from blood and bone marrow (BM) samples obtained from patients sCML14 and sCML23 before and after 7-9 days of hydroxyurea (HU) treatment. **(A)** Gating strategy for CITE-seq protein expression-based subsetting of immature $CD34^{+}CD38^{-}$ cells. **(B)** Distribution of subsetted $CD34^{+}CD38^{-}$ cells across the $CD14^{-}CD34^{+}$ UMAP. Subsetted cells are highlighted in purple. **(C)** Frequency of $CD38^{-}$ among $CD14^{-}CD34^{+}$ cells based on CITE-seq protein expression-based gating in blood samples obtained before HU treatment and blood and BM samples obtained after HU treatment for each patient. **(D)** Distribution of cells from

each patient in the reclustered $CD34^{+}CD38^{-}$ UMAP. Cells from samples obtained from patient sCML14 are highlighted in purple in the left panel, and those from patient sCML23 in the right panel. **(E-F)** Feature plots highlighting the expression of selected marker **(E)** proteins and **(F)** genes across the UMAP. Red indicates maximum expression and blue minimum expression.



Supplementary Figure 5. Characteristics of CD34^{+}CD38^{-} cell subset annotations generated through cell label transfer from a previously published reference dataset comprising CD14^{-} CD34^{+}CD38^{-} CML/healthy bone marrow cells (Nilsson *et al.*[1]). **(A)** Plots highlighting cells

given each individual cell label in the transfer. EBMP, eosinophil/basophil/mast cell progenitors; EP, erythroid progenitors; HSC, hematopoietic stem cells; LMP, lympho-myeloid progenitors; LSC, leukemic stem cells; MDP, monocyte/dendritic cell progenitors; MEP, megakaryocytic/erythroid progenitors; MKP, megakaryocytic progenitors; NP, neutrophil progenitors. **(B-C)** Scaled expression of **(B)** genes and **(C)** proteins used to confirm the validity of the cell label transfer annotation in Figure 7. Colors represent average expression level and dot size percentage of cells expressing the specific marker. For further detail on the original reference dataset annotation, see Nilsson *et al*.[1] **(D)** Volcano plot of differentially expressed genes between cells annotated as LSC-I and LSC-II. Green dots mark non-significantly regulated genes with $|Log_2FC| > 0.5$, blue dots significantly regulated genes (adjusted p values < 0.05) with $|Log_2FC| \leq 0.5$ and red dots significantly regulated genes with $|Log_2FC| > 0.5$. Statistics by Wilcoxon Rank Sum test with Bonferroni correction.



Supplementary Figure 6. Characterization of stem cell subsets in the confirmatory unpaired cohort, comprising bone marrow (BM) cells obtained from CML patients before (n=5) or after 4-19 (median 9) days of hydroxyurea (HU) treatment (n=7). **(A)** Proportion of cells in G0/G1 vs S/G2/M phase in each stem cell subset in the 12 samples. HSC, hematopoietic stem cells; LSC, leukemic stem cells. As HSC and LSC-II subsets were each missing in one patient, 11 data points are displayed for these subsets. **(B)** Frequency of LSC-I and LSC-II out of all $CD34^{+}CD38^{-}$ cells in samples obtained before or after HU treatment. Statistics by Mann-Whitney. * $p<0.05$.

Supplementary Tables

Supplementary Tables are available in a separate excel file.

Supplementary Table 1. Chronic phase CML patient samples included in flow cytometry and CITE-seq analyses. Data include patient ID, age, gender, and specification of peripheral blood/bone marrow samples included in flow cytometry and CITE-seq analyses (with cell numbers analyzed). In addition, hydroxyurea (HU) treatment duration and dosing is indicated for samples obtained after HU treatment.

Supplementary Table 2. Genes included in the CITE-seq analysis.

Supplementary Table 3. Oligo-conjugated AbSeq antibodies used in the CITE-seq analysis.

Supplementary table 4. Antibodies used for fluorescence-activated cell sorting of $CD14^-$ $CD34^+$ cells and AbSeq signal muting. The table additionally specifies the purpose of inclusion of each antibody.

Supplementary Table 5. Seurat FindAllMarkers cluster differential gene expression analysis in the $CD14^-CD34^+$ UMAP. Differential gene expression analysis of each cluster in Figure 3A (specified under "cluster") against cells in all other clusters. Columns specify Wilcoxon Rank Sum test p values (p_val), average $\log_2$ fold change (avg_log2FC), fraction of cells expressing the gene in each population (pct.1 and pct.2), and adjusted p values based on Bonferroni correction (p_val_adj). The analysis was run only considering genes expressed in

at least 25% of cells in either population and the table filtered to only contain upregulated genes with $log_2FC > 0.25$.

Supplementary Table 6. Seurat FindAllMarkers cluster differential protein expression analysis in the CD14$^-$CD34$^+$ UMAP. Differential protein expression analysis of each cluster in Figure 3A (specified under "cluster") against cells in all other clusters. Columns specify Wilcoxon Rank Sum test p values (p_val), average log_2 fold change (avg_log2FC), fraction of cells expressing the protein in each population (pct.1 and pct.2), and adjusted p values based on Bonferroni correction (p_val_adj). The analysis was run only considering proteins expressed in at least 25% of cells in either population and the table filtered to only contain upregulated proteins with $log_2FC > 0.25$.

Supplementary Table 7. Differential expression analysis shows that target genes of *GATA1* are upregulated among CD14$^-$CD34$^+$ cells following HU treatment. Genes analyzed were identified as belonging to the *GATA1* regulon in SCENIC analyses of paired blood samples from patients sCML14 and sCML23 obtained before and after (7 and 9 days, respectively) HU. Columns **(B-F)** show results from differential gene expression analysis in blood samples obtained after vs before HU for patient sCML14, and columns **(G-K)** corresponding results for patient sCML23. Subcolumns specify Wilcoxon Rank Sum test p values (p_val), average log_2 fold change (avg_log2FC), fraction of cells expressing the gene in each population (pct.1 and pct.2), and adjusted p values based on Bonferroni correction (p_val_adj). The table was filtered to only include genes significantly differentially expressed (adjusted p value < 0.05) and with $|log2FC| > 0.25$ in at least one of the comparisons, and cells colored according to direction of regulation (upregulated genes in red and downregulated genes in blue).

Supplementary Table 8. Seurat FindAllMarkers cluster differential gene expression analysis in the CD14$^-$CD34$^+$CD38$^-$ UMAP. Differential gene expression analysis of each cluster in Figure 6A (specified under "cluster") against cells in all other clusters. Columns specify Wilcoxon Rank Sum test p values (p_val), average $\log_2$ fold change (avg_log2FC), fraction of cells expressing the gene in each population (pct.1 and pct.2), and adjusted p values based on Bonferroni correction (p_val_adj). The analysis was run only considering genes expressed in at least 25% of cells in either population and the table filtered to only contain upregulated genes with $\log_2$FC > 0.25.

Supplementary Table 9. Seurat FindAllMarkers cluster differential protein expression analysis in the CD14$^-$CD34$^+$CD38$^-$ UMAP. Differential protein expression analysis of each cluster in Figure 6A (specified under "cluster") against cells in all other clusters. Columns specify Wilcoxon Rank Sum test p values (p_val), average $\log_2$ fold change (avg_log2FC), fraction of cells expressing the protein in each population (pct.1 and pct.2), and adjusted p values based on Bonferroni correction (p_val_adj). The analysis was run only considering proteins expressed in at least 25% of cells in either population and the table filtered to only contain upregulated proteins with $\log_2$FC > 0.25.

Supplementary Table 10. Differential gene expression analysis between LSC-II and LSC-I from patients sCML14 and sCML23. Columns specify Wilcoxon Rank Sum test p values (p_val), average $\log_2$ fold change (avg_log2FC), fraction of cells expressing the gene in each population (pct.1 and pct.2), and adjusted p values based on Bonferroni correction (p_val_adj). The analysis was run only considering genes expressed in at least 10% of cells in either population and the table filtered to only contain genes with |log2FC| > 0.25 and p_val_adj < 0.05.

Supplementary References

1. Nilsson MS, Komic H, Gustafsson J, et al. Multiomic single-cell analysis identifies von Willebrand factor and TIM3-expressing BCR-ABL1+ CML stem cells. bioRxiv. 2023 Sep 17. doi:10.1101/2023.09.14.557507 [preprint, not peer-reviewed].

2. Hafemeister C, Satija R. Normalization and variance stabilization of single-cell RNA-seq data using regularized negative binomial regression. Genome Biol. 2019;20(1):296.