Transcriptional regulation of HSCs in Aging and MDS reveals DDIT3 as a Potential Driver of Transformation

Short Title: Transcriptional lesions of HSCs in aging and MDS

Nerea Berastegui,[1,2,*] Marina Ainciburu,[1,2,*] Juan P. Romero,[1,2] Ana Alfonso-Pierola,[2,3] Céline Philippe,[4] Amaia Vilas-Zornoza,[1,2] Patxi San Martin,[1,2] Raquel Ordoñez,[5] Diego Alignani,[1,2] Sarai Sarvide,[1,2] Laura Castro,[1,2] José M. Lamo-Espinosa,[3] Mikel San-Julian,[3] Tamara Jimenez,[6] Félix López,[6] Sandra Muntion,[6] Fermin Sanchez-Guijo,[2,6] Antonieta Molero,[7] Julia Montoro,[7] Bárbara Tazón,[7] Guillermo Serrano,[8] Aintzane Diaz-Mazkiaran,[1,8] Mikel Hernaez,[2,8] Sofía Huerga,[3] Findlay Copley,[4] Ana Rio-Machin,[4] Matthew T. Maurano,[5,9] María Díez-Campelo,[2,6] David Valcarcel,[7] Kevin Rouault-Pierre,[4] David Lara-Astiaso,[1,§] Teresa Ezponda,[1,2,§] and Felipe Prosper[1,2,3,§]

*These authors contributed equally to this work
[§]These authors jointly supervised this work

[1]Área de Hemato-Oncología, CIMA Universidad de Navarra, Instituto de Investigación Sanitaria de Navarra (IDISNA), Pamplona, Spain; [2]Centro de Investigación Biomédica en Red de Cáncer, CIBERONC; [3]Clínica Universidad de Navarra, Universidad de Navarra, 31008 Pamplona, Spain; [4]Department of Haemato-Oncology, Barts Cancer Institute, Queen Mary University of London, London, England, UK; [5]Institute for Systems Genetics, NYU School of Medicine, New York, USA; [6]Hematology, Hospital Universitario de Salamanca-IBSAL, Salamanca, Spain; [7]Department of Hematology, University Hospital Vall d'Hebron, Barcelona, Spain; [8]Computational Biology Program, CIMA Universidad de Navarra, Instituto de Investigación Sanitaria de Navarra (IDISNA), Spain; [9]Department of Pathology, NYU School of Medicine, New York, USA;

Corresponding authors:
Felipe Prosper
Clínica Universidad de Navarra. Pio XII 36, 31008, Pamplona, Spain.
E-mail address: fprosper@unav.es
Phone number: +34948255400 Ext: 5807
Fax number: +34948296676

Teresa Ezponda
CIMA Universidad de Navarra. Pio XII 55, 31008, Pamplona, Spain.
E-mail address: tezponda@unav.es
Phone number: +34948194700 Ext:1005
Fax number: +34948194718

Text word count: 3,997
Abstract word count: 221
Number of figures: 6
Number of tables: 0
Number of references: 65

KEY POINTS:

- Human HSCs undergo a complex transcriptional rewiring in aging and MDS that may contribute to myeloid transformation.
- *DDIT3* overexpression induces a failure in the activation of erythroid transcriptional programs, leading to inefficient erythropoiesis.

ABSTRACT

Myelodysplastic syndromes (MDS) are clonal hematopoietic stem cell (HSC) malignancies characterized by ineffective hematopoiesis with increased incidence in elderly individuals. Genetic alterations do not fully explain the molecular pathogenesis of the disease, indicating that other types of lesions may play a role in its development. In this work, we analyzed the transcriptional lesions of human HSCs, demonstrating how aging and MDS are characterized by a complex transcriptional rewiring that manifests as diverse linear and non-linear transcriptional dynamisms. While aging-associated lesions seemed to predispose elderly HSCs to myeloid transformation, disease-specific alterations may be involved in triggering MDS development. Among MDS-specific lesions, we detected the overexpression of the transcription factor *DDIT3*. Exogenous upregulation of *DDIT3* in human healthy HSCs induced an MDS-like transcriptional state, and a delay in erythropoiesis, with an accumulation of cells in early stages of erythroid differentiation, as determined by single-cell RNA-sequencing. Increased *DDIT3* expression was associated with downregulation of transcription factors required for normal erythropoiesis, such as *KLF1*, *TAL1* or *SOX6*, and with a failure in the activation of their erythroid transcriptional programs. Finally, *DDIT3* knockdown in $CD34^+$ cells from MDS patients was able to restore erythropoiesis, as demonstrated by immunophenotypic and transcriptional profiling. These results demonstrate that *DDIT3* may be a driver of MDS transformation, and a potential therapeutic target to restore the inefficient erythropoiesis characterizing these patients.

INTRODUCTION

Deregulation of the molecular mechanisms that control hematopoiesis[1,2] can lead to the development of various hematological disorders, including myelodysplastic syndromes (MDS), which result from alterations in the first steps of hematopoietic differentiation[3-7]. Thus, MDS are characterized by ineffective hematopoiesis, and clinically manifested as cytopenias and increased risk of transformation to acute myeloid leukemia (AML)[3,7-10].

MDS prevalence is almost exclusive to elderly individuals[11], suggesting that alterations associated with aging predispose hematopoietic progenitors to abnormal differentiation. In fact, aging is accompanied by a decline in the hematopoietic system that includes defects in the adaptive and innate immune systems[12-17], and augmented risk of developing myeloid diseases[18]. Increasing evidence suggests that mechanisms intrinsic to HSCs, including changes in the epigenome[19-22] and the transcriptome[23-26] of these cells, are critical for the adverse hematopoietic consequences seen with age. Moreover, in mice, aging causes the expression of genes involved in leukemic transformation in HSCs[25], while in humans, developmental and cancer pathways are epigenetically reprogrammed in elderly $CD34^+$ cells[19], suggesting an aged-mediated predisposition towards the development of myeloid neoplasms. These findings suggest that gene expression alterations in aging may evolve towards more pathological profiles associated with MDS, although such lesions still need to be elucidated (**Figure 1A**).

Most studies aiming to characterize the molecular pathogenesis of MDS have mainly focused on mutational and epigenetic profiles associated with the disease[27-31], but the fundamental molecular bases of this pathology are still very incomplete. It is possible that transcriptional mechanisms, which can act as integrators of different alterations, also play a relevant role. A handful of studies have recently started to identify specific transcriptional alterations with prognostic value or a potential role in the pathogenesis and phenotype of MDS[32-41]. Although different cell types are phenotypically altered in MDS, previous studies suggest that HSCs are key to understanding this disease due to their privileged position at the apex of the differentiation process[3-7]. The low abundance of HSCs in the bone marrow (BM) makes its characterization challenging, implying that their expression profile could be partially masked by other cells when HSCs-enriched populations (i.e: $CD34^+$) are studied. We aimed to overcome such limitation by studying the transcriptional profile of purified HSCs. Furthermore, we integrated MDS-associated transcriptional alterations in

the context of aging, in order to gain a more comprehensive view of the molecular pathogenesis that ultimately triggers pathological hematopoietic differentiation.

We describe how human HSCs undergo a complex transcriptional rewiring in aging and MDS that manifests as linear and non-linear transcriptional dynamisms. While aging-associated lesions seem to confer human HSCs with a transformation-prone state, alterations characterizing MDS could be directly involved in promoting inefficient hematopoiesis. Among MDS-specific lesions, overexpression of the transcription factor (TF) *DDIT3* promotes a failure in the activation of erythroid differentiation programs and a delay in differentiation, whereas *DDIT3* knockdown in primary $CD34^+$ cells from MDS patients restores erythropoiesis. These results demonstrate that *DDIT3* may be a driver of MDS transformation and a potential therapeutic target for the disease.

METHODS

Sample collection and cell isolation

BM aspirates were obtained from healthy young, or elderly donors, and untreated MDS patients after the study was approved by the local ethics committee, and informed consent was obtained. Sample information and cell isolation strategy can be found in supplemental methods and Tables S1, S2 and S3.

Bulk RNA-sequencing (RNA-seq)

Bulk RNA-seq was performed following MARS-seq protocol adapted for bulk RNA-seq[42,43]. Information about MARS-seq and computational analyses is provided in supplemental methods.

DDIT3 overexpression and knockdown systems

DDIT3 cDNA was cloned into the pCDH-MCS-T2A-copGFP-MSCV lentiviral vector. For *DDIT3* knockdown, two different short hairpin RNAs (shRNAs) and a scramble shRNA were cloned into pSIH1-H1-copGFP shRNA lentiviral vector (System Biosciences #SI501A-A). Generation of lentiviruses and further specifications are described in supplemental methods.

Transcriptome profiling of primary healthy HSCs upon *DDIT3* overexpression

Healthy HSCs were transduced with the overexpression system (supplemental methods). Two days after transduction, GFP^+ cells were FACS-sorted using a BD FACSAria™ IIu in Lysis/Binding Buffer for Dynabeads™ mRNA Purification Kits (Invitrogen), and samples were snap-frozen and kept at -80 C until MARS-seq was performed.

Colony assays and ex vivo lineage differentiation assays

$CD34^+$ cells were infected[44], sorted after 4 days based on GFP, and subjected to colony, ex vivo lineage differentiation assays and transcriptomic analyses (details in supplemental methods).

Single-cell RNA-seq (scRNA-seq) of ex-vivo differentiated cells upon *DDIT3* overexpression

$CD34^+$ cells were transduced with lentiviruses as described above and subjected to the ex vivo differentiation system. After 14 days, transduced cells (GFP^+) were sorted and processed using the 10X Genomics Chromium Single Cell platform. The procedure and the computational analyses are explained in supplemental methods.

Transcription factor binding site analysis

The factors driving regulons with decreased activity in *DDIT3*-overexpressing cells were analyzed using CiiiDER, a TF binding site analysis software[45] (details in supplemental methods).

RESULTS

Profiling of HSCs reveals diverse transcriptional dynamics in aging and MDS

FACS-sorted HSCs ($CD34^+CD38^-CD90^+CD45RA^-$) obtained from young and elderly healthy donors, and from untreated MDS patients were analyzed using MARS-seq (**Figure 1B and S1A).** Principal component analysis showed that the largest transcriptional differences were observed between young and elderly samples (**Figure 1C**), with smaller differences between elderly and MDS patients. These observations were confirmed in an unsupervised hierarchical clustering (**Figure S1B**). Differential expression analysis demonstrated 733 genes deregulated between young and elderly healthy samples, and 907 genes between healthy elderly and MDS HSCs (|FC|>2; FDR<0.05) (**Figure S1C, Table S4**). A limited overlap was detected between differentially expressed genes (DEGs) in both comparisons (**Figure 1D**), indicating that specific transcriptional profiles, distinct from those present in aging HSCs, occur in MDS. Gene ontology (GO) analyses demonstrated that, among other processes, aging was characterized by an enrichment in inflammatory signaling, and a decrease in cell proliferation and DNA repair signatures; while MDS transcriptional lesions were enriched in RNA metabolism and showed underrepresentation of processes such as migration, or extracellular matrix organization (**Figure 1E**). Thus, not only deregulated expression of different genes but also distinct biological processes characterize HSCs in aging and the transition to MDS, suggesting that transcriptional alterations taking place in the development of the disease are not a continuous evolution of those found in aging.

Using MaSigPro[46], we identified 8 different patterns of expression in the transition of HSCs from healthy to MDS: two patterns with genes specifically upregulated or downregulated in aging, (clusters C1, C2); two patterns containing genes with augmented or decreased expression specifically in MDS compared to healthy HSCs (C3, C4); two linear trends harboring genes with either increased or decreased expression in aging and with an exacerbation of such changes in MDS (C5, C6); and finally, two patterns in which genes showed deregulation in aging, and alteration in the opposite direction between healthy elderly and MDS HSCs (C7, C8) (**Figure 1F**, **Table S5**). Using GO analyses, we identified the main functional categories and their associated biological processes enriched for each cluster (**Figure 1G, Table S6**). Although some processes such as gene regulation and metabolism were common to different clusters, others were cluster-specific, including DNA replication, or cell adhesion, among others. These results suggested a sequential transcriptional

rewiring in HSCs, and supported the existence of distinct transcriptional dynamisms regulating specific biological process across aging and MDS development.

Specific gene functions are associated with distinct patterns of expression in aging and MDS

To gain insight into the potential functional significance of the transcriptional rewiring taking place in aging and MDS, we analyzed the processes and pathways enriched for each cluster. Genes from C1 (upregulated in aging) were enriched in processes such as inflammation response, or response to bacteria, and in signaling pathways such as NF-Kß, VEGF, or GTPase-mediated transduction (**Figure 1G, 2A**), indicating a physiological response of HSCs to the more inflammatory microenvironment of elderly individuals[47]. Some examples included *RIPK2*, *IL1R1*, *NOD2, NFKßL1* or *RELB*, factors involved in the modulation of immune response and activation of the NF-Kß pathway (**Figure 2B**). Functional analysis of C2 (downregulated in aging) indicated a reduction of DNA replication, cell cycle transition, mitosis and cell proliferation (**Figure S2A**), suggesting a decreased proliferative state of HSCs in elderly individuals. Furthermore, this cluster showed an enrichment of genes involved in telomere maintenance, error-free translation synthesis, detection of DNA damage and nucleotide-excision repair (i.e: *BRCA1, BRCA2*, *FEN1*, *UNG*, *CHEK1*) (**Figure 1G, 2A, 2C**), suggesting that HSCs from elderly individuals lose their ability to overcome DNA insults and have increased genomic instability.

Transcriptional changes in C5 and C6 (exacerbation in MDS of aging-associated lesions) indicated a progressively more quiescent state of HSCs in aging and in MDS, with increased expression of negative cell cycle regulators, such as *TSPYL2* and *TGB1* (**Figure 2D**, **S2B**), and decreased levels of genes involved in DNA replication, G1/S transition or mitotic nuclear division. In addition, increased expression of factors involved in extracellular matrix organization, such as *COL9A3* and *ELN* (**Figure S2B**) and decreased levels of proteins regulating cell adhesion (**Figure S2C**), suggested abnormal interactions of HSCs with the microenvironment. DNA repair genes such as *UHRF*, *FANCA*, *TOP3A*, *CKAP5* or *USP10* were included in C6 (**Figure 2D, 2E**), suggesting an enhancement of the inability to overcome DNA damage in MDS HSCs. Notably, processes related to cell proliferation and DNA repair were mainly enriched in C2 but also in C6 (**Figure S2D**), suggesting that loss of proliferation and DNA repair abilities takes place during aging, and it is further amplified in MDS. Finally, these analyses indicated a progressive repression of genes involved in the maintenance of DNA methylation and of

negative regulators of transcription (**Figure S2C**), suggesting a progressive chromatin rewiring leading to gene activation in MDS.

Interestingly, GSEA demonstrated that genes altered in aging, with or without exacerbation of expression in MDS (C1, C2, C5, C6), were enriched in cancer-related signatures (**Figure 2F**). Furthermore, we identified in these clusters factors with known roles in the development or maintenance of myeloid malignancies (**Figure 2G**), such as *TRIB1*, a transforming gene for myeloid cells[48]; the matrix metallopeptidase *MMP2,* which shows high secretion levels in AML[49]; the adenosine deaminase *ADA2;* and the mini-chromosome maintenance proteins *MCM7* and *MCM4*[50], whose repression has been involved in myeloid leukemias (**Figure 2H**). These results suggest that age-derived transcriptional changes predispose HSCs to malignant transformation.

When analyzing C7 and C8 (reversal of the aging transcriptome), we observed that C7 was enriched in processes related to apoptosis, suggesting an increased regulation in aging and a loss of such physiological control in MDS (**Figure 1G**). Other biological processes such as response to stress, protein or RNA metabolism were enriched in these clusters (**Fig, S2E**), although the biological significance of their transitory deregulation in aging remains to be deciphered.

Clusters showing exclusive deregulation in MDS (C3, C4) represented lesions with a potential direct role in MDS development. GSEA and GO analyses (**Figure 2I, 2J**) demonstrated diminished expression of genes controlling cell division and DNA repair (**Figure S2F**). Decreased levels of genes regulating cell-substrate adhesion and B-cell differentiation as well as increased expression of genes involved in processes such as miRNA processing, exocytosis, type I IFN production, or response to TGF-ß were also detected in MDS-associated HSCs (**Figure S2G**). Finally, increased expression of chromatin and transcriptional regulators, such as *DDIT3*, *SERTAD2*, *RELA*, *EBF4* or *ZNF219* (**Figure 2K**), suggested an epigenetic and transcriptomic rewiring in MDS. Collectively, these transcriptional changes are consistent with an aged-mediated predisposition towards myeloid transformation, and with MDS-specific alterations that may contribute to the development of the disease.

DDIT3 overexpression produces an MDS-like transcriptional state and alters erythroid differentiation

Next, we focused on the transcriptional regulators specifically altered in MDS, and identified the TF *DDIT3* (CHOP/C/EBPζ), which was upregulated in MDS-associated HSCs (range of 2.7-10.5 FC over healthy elderly samples) (**Figure 2K, S3A**). *DDIT3* regulates hematopoietic differentiation in mice, where its overexpression alters lineage commitment, promoting an erythroid bias and a myeloid maturation defect at the GMP state[51]. To determine the impact of such upregulation, *DDIT3* was overexpressed in primary human healthy HSCs for 2 days in the absence of differentiation stimuli, and MARS-seq was performed (**Figure 3A**). *DDIT3*-overexpression induced the up- and downregulation of 427 and 128 genes, respectively (|FC|>2, FDR<0.05) (**Figure 3B**). Consistent with lesions observed in MDS patients, GSEA demonstrated activation of chromatin remodelers and decrease in DNA repair pathways and cell-substrate adhesion signatures upon *DDIT3* overexpression (**Figure 3C, S3B, S3C**). In line with the *DDIT3*-induced erythroid bias previously described in mice, we observed an upregulation of genes associated with heme-metabolism or erythroid differentiation, such as *FN3K*, *GCLM* or *HBB* (**Figure S3D**). Furthermore, *DDIT3* upregulation promoted an enrichment in cancer-related signatures, indicating a potential oncogenic role of this TF (**Figure 3D**). More importantly, using our own gene signatures generated from the DEGs detected in our cohort of MDS samples, we observed that *DDIT3*-overexpression induced a "MDS-like" transcriptional state, activating genes upregulated in MDS patients and repressing factors downregulated in the disease (**Figure 3E-F**). These data suggested that *DDIT3* upregulation plays a key role in inducing a pathological transcriptional state observed in MDS patients.

We next examined the biological effect of *DDIT3* upregulation using an ex vivo myeloid differentiation system starting from primary $CD34^+$ cells (**Figure 3A**). We observed a statistically significant decrease in the number of both erythroid burst-forming units (BFU-E) and granulo-monocytic colony forming units (CFU-G/M/GM) after *DDIT3* upregulation (**Figure 4A, S4A**). Liquid culture differentiation assays demonstrated a delay in the differentiation of *DDIT3*-overexpressing cells, with a statistically significant decrease in later stages of erythroid differentiation (stage IV, $CD71^-CD235a^+$) at days 10 and 14, and increased stage II and III erythroid progenitors ($CD71^+CD235a^-$ and $CD71^+CD235a^+$ erythroblasts) (**Figure 4B-C**). No significant alterations were detected in liquid culture myeloid differentiation assays upon *DDIT3* overexpression (**Figure S4B**).

To further characterize the effect of *DDIT3* overexpression in erythropoiesis, we performed scRNA-seq in human healthy CD34$^+$ cells overexpressing *DDIT3* or a mock control after 14 days of ex vivo differentiation (**Figure 3A**). We focused on clusters representing erythroid progenitors (**Figure S4C**). Cluster 4 was annotated as reticulocytes, due to its low counts, high ribosomal RNA content and the expression of genes characterizing this stage, such as those from the ATG family (**Figure S4C, S4D**). The remaining clusters were annotated as different stages of erythroid differentiation according to published transcriptomic data[52], and included CD34$^+$ cells, burst forming unit (BFU), colony formation unit (CFU), proerythroblasts, and different stages of erythroblasts (early-basophilic, late-basophilic, polychromatic and orthochromatic) (**Figure 4D**, **S4E**). Expression of *DDIT3* was upregulated in every stage compared to control cells (**Figure S4F**), and such overexpression was associated with an increase in the percentage of early erythroid progenitors, and a 34% and 46% decrease in the orthochromatic and reticulocyte stages, respectively (**Figure 4E**). Analysis of differentiation trajectories showed differences in RNA velocity between control and *DDIT3*-overexpressing cells, that were consistent with the decrease in late erythroid progenitors (**Figure 4F**). *DDIT3*-upregulated cells showed decreased RNA velocity from the most undifferentiated cells to the polychromatic state, with shorter, thinner and less dense streamlines being present in these cells as compared to control cells, suggesting an impairment of the correct differentiation to poly- and orthochromatic erythroblasts. Altogether, our data indicate that *DDIT3* overexpression in normal progenitor cells leads to inefficient erythropoiesis.

***DDIT3*-overexpression leads to a failure in the activation of key erythroid transcriptional programs**

To define the molecular mechanisms underlying the erythropoiesis defect induced by *DDIT3* overexpression, we analyzed the transcriptional profile of healthy CD34$^+$ cells transduced with this TF and subjected to erythroid differentiation (**Figure 3A**). GSEA of DEGs (**Figure S5A**) demonstrated an enrichment in transcriptional signatures of stem and early progenitor cells and decreased features of erythroid differentiated cells, such as oxygen transport or erythrocyte development, upon *DDIT3* upregulation (**Figure 5A**). Expression of hemoglobin genes and enzymes involved in heme biosynthesis (i.e: *PPOX*, *FECH*, *ALAS2, HMBS*) showed increased expression during differentiation in control cells but not in *DDIT3*-transduced cells (**Figure 5B, S5B**). On the contrary, factors characteristic of stem cells (*HOXB* genes, *NDN, TNIK)*, which were progressively repressed in control cells, showed aberrant high expression in *DDIT3*-upregulated cells (**Figure 5B, S5C**). We next used the generated scRNA-seq data to analyze *DDIT3*-induced transcriptional lesions at different

stages of erythroid differentiation, from CFU to orthochromatic erythroblast. *DDIT3*-overexpression induced an enrichment in hematopoietic stem and progenitor cell signatures in comparison with control cells (**Figure 5C**), as well as a decrease in the expression of genes related to heme metabolism or oxygen transport at every stage of differentiation. This was also supported by pseudotime analyses showing early hematopoietic progenitor genes downregulated during erythroid differentiation (*SEC61A1*, *CBFB*, *WDR18)* to be upregulated in *DDIT3*-overexpressing cells (**Figure 5D, S5D**); while demonstrating decreased expression of genes involved in erythroid differentiation and heme synthesis (*FAM210B*, *FOXO3*, *ATP5IF1,* hemoglobin genes) upon *DDIT3* upregulation (**Figure 5D, S5E**). These data suggest that *DDIT3* overexpression impairs the physiological transcriptional rewiring of erythroid differentiation, leading to a failure in the repression of immature genes, and in the activation of factors that are required for proper erythrocyte formation.

To dwell into the transcriptional programs involved in the inefficient erythroid differentiation induced by *DDIT3* overexpression, we applied a recently described algorithm, SimiC, that infers the differential activity regulons (TFs and their target genes) between two different conditions[53]. Whereas most regulons behaved similarly between control and *DDIT3*-upregulated cells (**Figure 5E,** low-dissimilarity score), a small number of regulons showed differential activity (high-dissimilarity score). Regulons with higher activity in control cells were guided by TFs with known roles in the promotion of erythroid differentiation, such as *KLF1*, *ARID4B*, *SOX6*, *TAL1* or *HES6* (**Figure 5F**), whereas regulons that gained activity upon *DDIT3* overexpression were driven by TFs that negatively impact erythropoiesis, such as *ARID3A*, *ZFPM1*, *JUND*, *ARID1A*, or *HMGA1* (**Figure S5F**). Differential expression and pseudotime analyses showed that *DDIT3* upregulation was associated with diminished expression of TFs with a role in erythropoiesis promotion, such as *SOX6*, *KLF1*, *TAL1* or *ARID4B* (**Figure 5G, 5H**), potentially leading to lower activity of such regulons. Furthermore, the promoters of these TFs were enriched in binding sites for *CEBPB* and *CEBPG* (**Figure 5I**), which are known to be sequestered and inhibited by *DDIT3*. Thus, these results suggested that abnormally high *DDIT3* expression in MDS could sequester *CEBPB* and *CEBPG*, leading to decreased expression of their target genes, including TFs with key roles in erythropoiesis, ultimately hampering the activity of transcriptional programs that are necessary for proper terminal erythroid differentiation.

DDIT3 knockdown restores erythroid differentiation of primary MDS samples

Based on our results, we hypothesized that inhibition of *DDIT3* in MDS hematopoietic progenitor cells could restore normal erythropoiesis. *DDIT3* was knocked-down in $CD34^+$ cells from patients with MDS using shRNAs, and liquid culture differentiation assays were performed (**Figure 6A**). *DDIT3* knockdown in patient 13, a male with MDS-MLD and anemia (hemoglobin (Hb) 11.8 g/dL), led to an increase in the percentage of cells in stage IV ($CD235^+$ $CD71^-$) at day 7 of differentiation, that was even more notable at day 13 (**Figure 6B, 6C**), where cells expressing the control shRNA were partially blocked at the more immature stage III. In an additional MDS-MLD case showing anemia (male, Hb 7.9 g/dL), we also detected improved erythroid differentiation upon *DDIT3* knockdown, with higher levels of CD71 expression and improved transition to stage III at day 7 of differentiation (**Figure S6A, S6B**). The promotion of erythroid differentiation by *DDIT3* knockdown was validated by transcriptional profiling of these samples at day 7 of differentiation. This analysis revealed that *DDIT3* knockdown induced increased expression of hemoglobins and heme biosynthetic enzymes, and diminished levels of genes characteristic of immature progenitors, such as HOX genes, compared to cells transduced with a control shRNA (**Figure 6D, S6C**). Moreover, using published data of gene expression profiling at different erythroid differentiation stages[54], we observed that *DDIT3* knockdown decreased the expression of genes that are mainly expressed in proerythroblasts, and early and late basophilic erythroblasts, while promoting the activation of genes defining poly- and orthochromatic stages (**Figure 6E, S6D**). All together, these results suggest that inhibition of *DDIT3* in patients with MDS presenting anemia restores proper terminal erythroid differentiation.

DISCUSSION

In this work, we have taken two novel approaches in the study of MDS that we believe are key to better comprehend its molecular pathogenesis. To our knowledge, possibly due to the technical complexity of studying such scarce subpopulation, this is the first time that purified human HSCs have been analyzed in MDS. Notably, many of the genes and processes found deregulated in MDS were different from those previously described[32-41], which may be due to the fact that preceding studies analyzed HSCs-enriched populations, from which HSCs are only a small fraction. Secondly, we have analyzed the molecular lesions of these cells in the context of aging, allowing for the identification of complex alterations in aging and MDS development that include not only specific lesions of aging and MDS, but also continuous alterations, and even reversal of the aging transcriptome, which are findings not previously described. These alterations suggest a profound transcriptional reprogramming of HSCs in aging and MDS development.

The detected age-derived alterations indicated decreased DNA damage sensing and repair abilities of elderly HSCs, suggesting an augmented rate of mutations in these cells. Accordingly, human elderly $CD34^{+}CD38^{-}$ cells exhibit an accumulation of DNA damage, which was linked to decreased self-renewal capacity[55]. Furthermore, elderly HSCs showed an enrichment in cancer-related signatures and factors with known roles in the development of myeloid malignancies. In line with these data, murine aged HSCs overexpress genes involved in leukemic transformation[25], while an epigenetic reprogramming suggested an age-driven leukemic transformation-prone state of human HSC-enriched populations[19]. Collectively, these data suggest that age-dependent transcriptional alterations at the HSC level increase the risk of malignant transformation into myeloid neoplasms.

Our data also identified MDS-associated transcriptional alterations that may contribute to the development of the disease, including decreased expression of DNA repair genes, and of cell adhesion molecules. Loss of attachment in the HSC niche is a known leukemic mechanism of dissemination from the BM[56], and our data suggest that it could start at very early stages of myeloid malignancies. Among genes specifically altered in MDS, we also detected the overexpression of chromatin and transcriptional regulators, which could act as master factors guiding the transcriptional reprogramming of HSCs. Among them, we characterized the overexpression of *DDIT3*, a member of the C/EBP family of TFs.

DDIT3 is altered in numerous tumors, showing a variety of lesions (altered expression, structural abnormalities[57,58]) that lead to increased or diminished activity, suggesting that its role in tumorigenesis may be context-specific. *DDIT3* functions as a dominant-negative inhibitor by forming heterodimers and preventing the DNA binding activity of C/EBP family members and other factors, and also acts as an activator of gene expression[59]. Thus, depending on its partners and its own activity, its functional output may vary, ultimately promoting or inhibiting tumor development. Furthermore, *DDIT3* inhibits adipogenesis by interacting with *CEBPA* and *CEBPB*[60,61], while it enhances osteoblastic differentiation by interplaying with the same factors[62], suggesting that, apart from its partners, other cell-type specific factors may dictate its biological role. Intriguingly, previous works reported downregulation of *DDIT3*[63,64] and hypermethylation of its promoter[65] in MDS. These studies analyzed BM mononuclear cells, from which HSCs represent a very small percentage. Thus, *DDIT3* may be specifically overexpressed in early phases of progenitor cell differentiation and commitment, where its upregulation promotes inefficient hematopoiesis. Future studies characterizing specific subpopulations will determine its role at more mature stages, the nature of its binding partners in those cells, and the phenotypic effect that such lesions imply.

Finally, we demonstrate that knockdown of *DDIT3* in $CD34^+$ cells from MDS cases restores erythroid differentiation. Interestingly, the patients used in these assays did not harbor extremely high levels of *DDIT3* compared to elderly samples, suggesting that the benefit of *DDIT3* inhibition may exceed patients showing upregulation of this factor. Moreover, these data suggest that other mechanisms besides its overexpression, which may include MDS-specific binding partners or novel targeting to specific loci, may lead to aberrant *DDIT3* activity.

Collectively, these results demonstrate the transcriptional rewiring that takes place in aging HSCs and MDS, and highlight the relevance of its analysis for the identification of factors that promote inefficient differentiation, and that could represent therapeutic targets to restore hematopoiesis in these patients.

Acknowledgments

This work was supported by the Instituto de Salud Carlos III and co-financed by FEDER funds (PI17/00701, and PI20/01308), CIBERONC (CB16/12/00489); Gobierno de Navarra (AGATA 0011-1411-2020-000010/0011-1411-2020-000011 and DIANA 0011-1411-2017-000028/0011-1411-2017-000029/0011-1411-2017-000030); Fundación La Caixa (GR-NET NORMAL-HIT HR20-00871); and Cancer Research UK [C355/A26819] and FC AECC and AIRC under the Accelerator Award Program. NB was supported by a PhD fellowship from Gobierno de Navarra (0011-0537-2019-000001); MA was supported by a PhD fellowship from Ministerio de Ciencia, Innovación y Universidades (FPU18/05488); TE was supported by an Investigador AECC award from the Fundación AECC, MH was supported by H2020 Marie S. Curie IF Action, European Commission, Grant Agreement No. 898356. We particularly acknowledge the patients and healthy donors for their participation in this study.

Authorship Contributions

T.E. and F.P. conceived and supervised the study. A.A, J.L, M. S, T.J., F. L., S.M, F.S, A.M, J.M, B.T, S.H, M.D.C, and D.V collected clinical samples and data. A.V, P.S.M, and S.S. generated sequencing data. N.B, D.A, C.P, K.R.P, and T.E performed ex vivo and in vitro experiments. N.B, M.A, J.P.R, R.O, L.C, G.S, A.D, M.H, A.R, D.F, M.T.M, and D.L. performed the data analysis. T.E, and F.P. wrote the manuscript.

Disclosure of Conflicts of interest

The authors have no conflict of interest to disclose.

References

1. Chen L, Kostadima M, Martens JHA, et al. Transcriptional diversity during lineage commitment of human blood progenitors. *Science.* 2014. 345I(6204): 1251033.
2. Wilson NK, Calero-Nieto FJ, Ferreira R, and Gottgens B. Transcriptional regulation of haematopoietic transcription factors. *Stem Cell Res Ther.* 2011. 2I(1): 6.
3. Heaney ML and Golde DW. Myelodysplasia. *N Engl J Med.* 1999. 340I(21): 1649-60.
4. Pellagatti A and Boultwood J. The molecular pathogenesis of the myelodysplastic syndromes. *Eur J Haematol.* 2015. 95I(1): 3-15.
5. Sperling AS, Gibson CJ, and Ebert BL. The genetics of myelodysplastic syndrome: from clonal haematopoiesis to secondary leukaemia. *Nat Rev Cancer.* 2017. 17I(1): 5-19.
6. Woll PS, Kjallquist U, Chowdhury O, et al. Myelodysplastic syndromes are propagated by rare and distinct human cancer stem cells in vivo. *Cancer Cell.* 2014. 25I(6): 794-808.
7. Elias HK, Schinke C, Bhattacharyya S, Will B, Verma A, and Steidl U. Stem cell origin of myelodysplastic syndromes. *Oncogene.* 2014. 33I(44): 5139-50.
8. Greenberg PL, Tuechler H, Schanz J, et al. Revised international prognostic scoring system for myelodysplastic syndromes. *Blood.* 2012. 120I(12): 2454-65.
9. Greenberg PL, Young NS, and Gattermann N. Myelodysplastic syndromes. *Hematology Am Soc Hematol Educ Program.* 2002: 136-61.
10. Raza A and Galili N. The genetic basis of phenotypic heterogeneity in myelodysplastic syndromes. *Nat Rev Cancer.* 2012. 12I(12): 849-59.
11. Ma X, Does M, Raza A, and Mayne ST. Myelodysplastic syndromes: incidence and survival in the United States. *Cancer.* 2007. 109I(8): 1536-42.
12. Kogut I, Scholz JL, Cancro MP, and Cambier JC. B cell maintenance and function in aging. *Semin Immunol.* 2012. 24I(5): 342-9.
13. Montecino-Rodriguez E, Berent-Maoz B, and Dorshkind K. Causes, consequences, and reversal of immune system aging. *J Clin Invest.* 2013. 123I(3): 958-65.
14. Nikolich-Zugich J. Aging of the T cell compartment in mice and humans: from no naive expectations to foggy memories. *J Immunol.* 2014. 193I(6): 2622-9.
15. Pang WW and Schrier SL. Anemia in the elderly. *Curr Opin Hematol.* 2012. 19I(3): 133-40.
16. Rossi DJ, Jamieson CH, and Weissman IL. Stems cells and the pathways to aging and cancer. *Cell.* 2008. 132I(4): 681-96.

17. Shaw AC, Goldstein DR, and Montgomery RR. Age-dependent dysregulation of innate immunity. *Nat Rev Immunol.* 2013. 13I(12): 875-87.
18. Lichtman MA and Rowe JM. The relationship of patient age to the pathobiology of the clonal myeloid diseases. *Semin Oncol.* 2004. 31I(2): 185-97.
19. Adelman ER, Huang HT, Roisman A, et al. Aging Human Hematopoietic Stem Cells Manifest Profound Epigenetic Reprogramming of Enhancers That May Predispose to Leukemia. *Cancer Discov.* 2019. 9I(8): 1080-1101.
20. Challen GA, Sun D, Mayle A, et al. Dnmt3a and Dnmt3b have overlapping and distinct functions in hematopoietic stem cells. *Cell Stem Cell.* 2014. 15I(3): 350-364.
21. Moran-Crusio K, Reavie L, Shih A, et al. Tet2 loss leads to increased hematopoietic stem cell self-renewal and myeloid transformation. *Cancer Cell.* 2011. 20I(1): 11-24.
22. Sun D, Luo M, Jeong M, et al. Epigenomic profiling of young and aged HSCs reveals concerted changes during aging that reinforce self-renewal. *Cell Stem Cell.* 2014. 14I(5): 673-88.
23. Khoo ML, Carlin SM, Lutherborrow MA, Jayaswal V, Ma DD, and Moore JJ. Gene profiling reveals association between altered Wnt signaling and loss of T-cell potential with age in human hematopoietic stem cells. *Aging Cell.* 2014. 13I(4): 744-54.
24. Pang WW, Price EA, Sahoo D, et al. Human bone marrow hematopoietic stem cells are increased in frequency and myeloid-biased with age. *Proc Natl Acad Sci U S A.* 2011. 108I(50): 20012-7.
25. Rossi DJ, Bryder D, Zahn JM, et al. Cell intrinsic alterations underlie hematopoietic stem cell aging. *Proc Natl Acad Sci U S A.* 2005. 102I(26): 9194-9.
26. Rundberg Nilsson A, Soneji S, Adolfsson S, Bryder D, and Pronk CJ. Human and Murine Hematopoietic Stem Cell Aging Is Associated with Functional Impairments and Intrinsic Megakaryocytic/Erythroid Bias. *PLoS One.* 2016. 11I(7): e0158369.
27. Ganguly BB, Banerjee D, and Agarwal MB. Impact of chromosome alterations, genetic mutations and clonal hematopoiesis of indeterminate potential (CHIP) on the classification and risk stratification of MDS. *Blood Cells Mol Dis.* 2018. 69I(90-100.
28. Makishima H, Yoshizato T, Yoshida K, et al. Dynamics of clonal evolution in myelodysplastic syndromes. *Nat Genet.* 2017. 49I(2): 204-212.
29. Ogawa S. Genetics of MDS. *Blood.* 2019. 133I(10): 1049-1059.
30. Bond DR, Lee HJ, and Enjeti AK. Unravelling the Epigenome of Myelodysplastic Syndrome: Diagnosis, Prognosis, and Response to Therapy. *Cancers (Basel).* 2020. 12I(11).

31. Mahmud M and Stebbing J. Epigenetic modifications in AML and MDS. *Leuk Res.* 2010. 34I(2): 139-40.
32. Boultwood J, Pellagatti A, Cattan H, et al. Gene expression profiling of CD34+ cells in patients with the 5q- syndrome. *Br J Haematol.* 2007. 139I(4): 578-89.
33. Hofmann WK, de Vos S, Komor M, Hoelzer D, Wachsman W, and Koeffler HP. Characterization of gene expression of CD34+ cells from normal and myelodysplastic bone marrow. *Blood.* 2002. 100I(10): 3553-60.
34. Im H, Rao V, Sridhar K, et al. Distinct transcriptomic and exomic abnormalities within myelodysplastic syndrome marrow cells. *Leuk Lymphoma.* 2018. 59I(12): 2952-2962.
35. Kazachenka A, Young GR, Attig J, et al. Epigenetic therapy of myelodysplastic syndromes connects to cellular differentiation independently of endogenous retroelement derepression. *Genome Med.* 2019. 11I(1): 86.
36. Miyazato A, Ueno S, Ohmine K, et al. Identification of myelodysplastic syndrome-specific genes by DNA microarray analysis with purified hematopoietic stem cell fraction. *Blood.* 2001. 98I(2): 422-7.
37. Montalban-Bravo G, Class CA, Ganan-Gomez I, et al. Transcriptomic analysis implicates necroptosis in disease progression and prognosis in myelodysplastic syndromes. *Leukemia.* 2020. 34I(3): 872-881.
38. Pellagatti A, Armstrong RN, Steeples V, et al. Impact of spliceosome mutations on RNA splicing in myelodysplasia: dysregulated genes/pathways and clinical associations. *Blood.* 2018. 132I(12): 1225-1240.
39. Pellagatti A, Cazzola M, Giagounidis A, et al. Deregulated gene expression pathways in myelodysplastic syndrome hematopoietic stem cells. *Leukemia.* 2010. 24I(4): 756-64.
40. Pellagatti A, Cazzola M, Giagounidis AA, et al. Gene expression profiles of CD34+ cells in myelodysplastic syndromes: involvement of interferon-stimulated genes and correlation to FAB subtype and karyotype. *Blood.* 2006. 108I(1): 337-45.
41. Ueda M, Ota J, Yamashita Y, et al. DNA microarray analysis of stage progression mechanism in myelodysplastic syndrome. *Br J Haematol.* 2003. 123I(2): 288-96.
42. Jaitin DA, Kenigsberg E, Keren-Shaul H, et al. Massively parallel single-cell RNA-seq for marker-free decomposition of tissues into cell types. *Science.* 2014. 343I(6172): 776-9.
43. Lavin Y, Kobayashi S, Leader A, et al. Innate Immune Landscape in Early Lung Adenocarcinoma by Paired Single-Cell Analyses. *Cell.* 2017. 169I(4): 750-765 e17.

44. Rouault-Pierre K, Lopez-Onieva L, Foster K, et al. HIF-2alpha protects human hematopoietic stem/progenitors and acute myeloid leukemic cells from apoptosis induced by endoplasmic reticulum stress. *Cell Stem Cell.* 2013. 13I(5): 549-63.
45. Gearing LJ, Cumming HE, Chapman R, et al. CiiiDER: A tool for predicting and analysing transcription factor binding sites. *PLoS One.* 2019. 14I(9): e0215495.
46. Conesa A, Nueda MJ, Ferrer A, and Talon M. maSigPro: a method to identify significantly differential expression profiles in time-course microarray experiments. *Bioinformatics.* 2006. 22I(9): 1096-102.
47. Leimkuhler NB and Schneider RK. Inflammatory bone marrow microenvironment. *Hematology Am Soc Hematol Educ Program.* 2019. 2019I(1): 294-302.
48. Nakamura T. The role of Trib1 in myeloid leukaemogenesis and differentiation. *Biochem Soc Trans.* 2015. 43I(5): 1104-7.
49. Chaudhary AK, Chaudhary S, Ghosh K, Shanmukaiah C, and Nadkarni AH. Secretion and Expression of Matrix Metalloproteinase-2 and 9 from Bone Marrow Mononuclear Cells in Myelodysplastic Syndrome and Acute Myeloid Leukemia. *Asian Pac J Cancer Prev.* 2016. 17I(3): 1519-29.
50. Van Montfrans JM, Hartman EA, Braun KP, et al. Phenotypic variability in patients with ADA2 deficiency due to identical homozygous R169Q mutations. *Rheumatology (Oxford).* 2016. 55I(5): 902-10.
51. Pina C, Teles J, Fugazza C, et al. Single-Cell Network Analysis Identifies DDIT3 as a Nodal Lineage Regulator in Hematopoiesis. *Cell Rep.* 2015. 11I(10): 1503-10.
52. Yan H, Hale J, Jaffray J, et al. Developmental differences between neonatal and adult human erythropoiesis. *Am J Hematol.* 2018. 93I(4): 494-503.
53. Peng J, Serrano G, Traniello IA, et al. A single-cell gene regulatory network inference method for identifying complex regulatory dynamics across cell phenotypes. *BioRxiv.*
54. An X, Schulz VP, Li J, et al. Global transcriptome analyses of human and murine terminal erythroid differentiation. *Blood.* 2014. 123I(22): 3466-77.
55. Yahata T, Takanashi T, Muguruma Y, et al. Accumulation of oxidative DNA damage restricts the self-renewal capacity of human hematopoietic stem cells. *Blood.* 2011. 118I(11): 2941-50.
56. Gruszka AM, Valli D, Restelli C, and Alcalay M. Adhesion Deregulation in Acute Myeloid Leukaemia. *Cells.* 2019. 8I(1).

57. Dal Cin P, Sciot R, Panagopoulos I, et al. Additional evidence of a variant translocation t(12;22) with EWS/CHOP fusion in myxoid liposarcoma: clinicopathological features. *J Pathol.* 1997. 182I(4): 437-41.
58. Perez-Losada J, Pintado B, Gutierrez-Adan A, et al. The chimeric FUS/TLS-CHOP fusion protein specifically induces liposarcomas in transgenic mice. *Oncogene.* 2000. 19I(20): 2413-22.
59. Ubeda M, Wang XZ, Zinszner H, Wu I, Habener JF, and Ron D. Stress-induced binding of the transcriptional factor CHOP to a novel DNA control element. *Mol Cell Biol.* 1996. 16I(4): 1479-89.
60. Batchvarova N, Wang XZ, and Ron D. Inhibition of adipogenesis by the stress-induced protein CHOP (Gadd153). *EMBO J.* 1995. 14I(19): 4654-61.
61. Tang QQ and Lane MD. Role of C/EBP homologous protein (CHOP-10) in the programmed activation of CCAAT/enhancer-binding protein-beta during adipogenesis. *Proc Natl Acad Sci U S A.* 2000. 97I(23): 12446-50.
62. Pereira RC, Delany AM, and Canalis E. CCAAT/enhancer binding protein homologous protein (DDIT3) induces osteoblastic cell differentiation. *Endocrinology.* 2004. 145I(4): 1952-60.
63. Qian J, Chen Z, Lin J, Wang W, and Cen J. Decreased expression of CCAAT/enhancer binding protein zeta (C/EBPzeta) in patients with different myeloid diseases. *Leuk Res.* 2005. 29I(12): 1435-41.
64. Qian J, Chen Z, Wang W, Cen J, and Xue Y. Gene expression profiling of the bone marrow mononuclear cells from patients with myelodysplastic syndrome. *Oncol Rep.* 2005. 14I(5): 1189-97.
65. Lin J, Wang YL, Qian J, et al. Aberrant methylation of DNA-damage-inducible transcript 3 promoter is a common event in patients with myelodysplastic syndrome. *Leuk Res.* 2010. 34I(8): 991-4.

Figure legends

Figure 1. Altered transcriptional profiles of HSCs in aging and MDS. (A) Schematic representation of lesions taking place in aging and in MDS. Aging is characterized by transcriptional lesions of HSCs and can also present mutations in specific genes. MDS HSCs usually present genetic lesions but the transcriptional profile of these cells remains unexplored. (B) Schematic representation of the experimental approach: BM specimens from young and elderly healthy donors as well as from MDS patients were obtained, and HSCs ($CD34^+$ $CD38^-$ $CD90^+$ $CD45RA^-$) were isolated using FACS. The transcriptome of these cells was characterized using MARS-seq. (C) Principal component analysis (PCA) of the transcriptome of cells isolated in B. Healthy young (blue), elderly (yellow) and MDS cells (red) are plotted. The percentage of variance explained by PC1 and PC2 principal component is indicated in each axis. (D) Venn diagram representing a partial overlap of genes deregulated in aging (young vs elderly healthy), in the transition to the disease (healthy elderly vs MDS) or between the more distant states (healthy young vs MDS). The number of common or exclusive DEGs is indicated in each area. (E) Barplot showing enriched biological processes determined by GO analysis of genes deregulated in aging or between healthy elderly and MDS HSCs. The -log(p-value) for several statistically significant processes is depicted. (F) Transcriptional dynamisms identified in HSCs in the aging-MDS axis (C1-C8). Left: dot plot depicting the median expression of genes of each cluster in each sample. Each color represents the different states: healthy young: blue, healthy elderly: yellow, and MDS: red. The trend of each cluster is indicated with a line linking the median of each group. Center: heatmap showing z-scores for the expression profile of each cluster of genes in healthy young, elderly and MDS samples. Right: Bar-plot indicating the number of genes per cluster. (G) Heatmap showing the statistical significance (log_{10}p-value) of enrichment of the genes of each cluster in different biological processes, as determined by GO analysis. The different processes have been manually grouped into more general biological functions (right).

Figure 2. Different transcriptional dynamisms of HSCs in the aging-MDS axis associate with specific biological functions. (A) Bubble plot representing statistically significant biological processes and pathways enriched in genes being specifically altered in aging (C1: red dots; C2: blue dots). Bubble size depicts -log10(p-value) and x axis represents GeneRatio (number of genes present in the list/number of total genes in the process). (B-C) Normalized expression of several genes from C1 involved in inflammation response signaling (C), and of genes from C2 related to DNA damage (D),

in healthy young (blue), healthy elderly (orange) and MDS (red) samples. Each point represents a donor or patient and the mean +/- standard deviation (SD) is shown. (D) Bubble plot representing statistically significant biological processes and pathways enriched in genes showing a continuous deregulation in the axis (C5: red dots; C6: blue dots). Bubble size depicts -$\log_{10}$(p-value) and x axis represents GeneRatio. (E) Normalized expression of genes from C6 involved in DNA damage, in healthy young (blue), healthy elderly (orange) and MDS (red) samples. Each point represents a donor or patient and the mean +/- standard deviation (SD) is shown. (F) GSEA analysis of genes from C1, C2, C5 and C6. The normalized enrichment score (NES) for several cancer-related signatures is depicted. Only signatures in which age-matched healthy controls were used as normal counterparts of tumor cells were considered. (G) Heatmap of z-scores of genes with known roles in the development of myeloid malignancies. The cluster to which they correspond, and the gene names are indicated on the left and right side of the heatmap, respectively. (H) Normalized expression of several genes from G in healthy young (blue), healthy elderly (orange) and MDS (red) samples. Each point represents an individual and the mean +/- standard deviation (SD) is depicted. (I) Bubble plot representing statistically significant biological processes and pathways enriched in genes being specifically altered in MDS (C3: red dots; C4: blue dots). Bubble size depicts -log10(p-value) and x axis represents GeneRatio. (J) GSEA analysis of genes from C3 and C4. The NES for several enriched signatures is shown. (K) Normalized expression of several genes from C3 involved in transcriptional regulation in healthy young (blue), healthy elderly (orange) and MDS (red) samples. Each point represents a donor or patient and the mean +/- standard deviation (SD) is shown.

Figure 3. *DDIT3* overexpression promotes and MDS-like transcriptional state. (A) Schematic representation of the experimental procedure: primary HSCs/CD34$^+$ cells from healthy young donors were FACS-sorted and infected with a lentiviral plasmid harboring *DDIT3*, or a control plasmid. Two days after infection, transduced cells were sorted and their transcriptome analyzed using MARS-seq. After 4 days of infection, transduced cells were also isolated and used to perform colony and liquid culture differentiation assays. The later were evaluated by flow cytometry and MARS-seq at different time points and by scRNAseq at day 14 of differentiation. (B) Volcano plot of statistical significance (-$\log_{10}$ (p-value)) against fold-change ($\log_2$ Fold-change) of gene expression between cells overexpressing *DDIT3* and control HSCs. Green points represent genes with |FC|>2, red points depict genes with |FC|>2 and FDR>0.05 and grey points indicate genes with no relevant changes in expression. Several genes with significant up- and down-regulation are indicated. (D-E)

GSEA of cells overexpressing *DDIT3* and control HSCs. The NES for several signatures related to general biological processes (C) and cancer-related signatures (D) are depicted. (E-F) GSEA plots depicting the enrichment upon *DDIT3* overexpression in gene signatures representing genes up- and down-regulated in MDS. The NES and adjusted p-values are indicated for each signature.

Figure 4. *DDIT3* overexpression promotes a defect in erythroid differentiation. (A) Graph indicating the number of colonies (BFU-E and G/M/GM) obtained for control cells or cells overexpressing *DDIT3*. Statistically significant differences are indicated (*p-value < 0.05; **p-value < 0.01). Biological replicates for this experiment can be found in Figure S4A. (B) Flow cytometry charts representing advanced erythroid differentiation (CD71 and CD235a markers; stages I-IV) for control cells and cell overexpressing *DDIT3*, at the indicated time points. (C) Bar-plots representing the percentage of cells observed in C in stages I-IV at 10 and 14 days of differentiation. Three biological replicates +/- SEM is depicted, and statistically significant differences are indicated (*p-value<0.05; **p-value<0.01; ***p-value>0.001). (D) UMAP plot of the transcriptome of cells subjected to ex vivo differentiation for 14 days. Cells have been clustered in different groups representing erythroid differentiation: $CD34^+$ cells, burst forming unit/primitive erythroid progenitor cells (BFU), colony formation unit/later-stage erythroid progenitor cells (CFU), proerythroblast, early-basophilic erythroblast, late-basophilic erythroblast, polychromatic erythroblast, orthochromatic erythroblast, and reticulocyte. (E) Barplot representing the proportions of cells in each of the clusters described in F for control cells or cells overexpressing *DDIT3* upon ex vivo differentiation for 14 days. (F) RNA velocity plotted in UMAP space for control or *DDIT3*-overexpressing cells in clusters defined in E. Streamlines and arrows indicate the location of the estimated future cell state.

Figure 5. *DDIT3* upregulation leads to a failure in the activation of erythroid differentiation programs. (A) Barplot depicting GSEA analysis upon *DDIT3* overexpression at day 14 of cells overexpressing *DDIT3* and control HSCs at day 14 of erythroid differentiation. The NES for several signatures related to erythropoiesis and stem and progenitor profiles are shown. (B) Normalized expression of erythroid differentiation (top) and stem cell genes (bottom) of $CD34^+$ cells at different time points of erythroid differentiation. The mean +/- standard deviation (SD) of 2 biological replicates is depicted. (C) Plot representing GSEA of cells overexpressing *DDIT3* and control cells at different stages of erythroid differentiation detected by scRNAseq. The size of the dots represent NES absolute value, the color indicates the group in which processes are enriched (blue in control

cells, red in *DDIT3*-overexpressing cells), and the intensity of the color depicts the p-value obtained for each geneset. (D) Gene expression trends of early hematopoietic progenitor genes (top) and erythroid differentiation factors (bottom) calculated by pseudotime are represented as a smooth fit with the standard deviation of the fit shown in a lighter shade for control (blue) and *DDIT3*-overexpressing cells (red). (E) Heatmap showing the regulatory dissimilarity score between control and *DDIT3*-overexpressing cells at day 14 of ex vivo liquid culture differentiation in different stages of erythroid differentiation defined by scRNAseq. (F) Ridge plot showing AUC scores for regulons of *KLF1*, *ARID4B* and *SOX6* in control (blue) and *DDIT3*-overexpressing cells (pink) at different stages of erythroid differentiation. (G) Gene expression trends calculated by pseudotime of TFs guiding regulons showing decrease activity in *DDIT3*-overexpressing cells, are represented as a smooth fit with the standard deviation of the fit shown in a lighter shade for control (blue) and *DDIT3*-overexpressing cells (red). (H) Normalized expression of TFs guiding regulons showing decrease activity in *DDIT3*-overexpressing cells, at different time points of ex vivo erythroid differentiation. The mean +/- standard deviation (SD) of 2 biological replicates is depicted. (I) Analysis of putative transcription factor site enrichment in TFs guiding regulons with decreased activity in *DDIT3*-overexpressing cells using CiiiDER. Color and size of circles reflect p-value of enrichment. Over-represented transcription factors of potential interest are depicted.

Figure 6. *DDIT3* knockdown in MDS patients with anemia restores erythroid differentiation. (A) Schematic representation of *DDIT3* knockdown experiments in $CD34^+$ cells from patients with MDS showing anemia. Cells were transduced with a control or *DDIT3*-targeting shRNA, after two days of infection, cells were subjected to ex vivo liquid culture differentiation over OP9 stromal cells. The differentiation state was evaluated by flow cytometry and MARS-seq analyses. (B) Flow cytometry charts representing advanced erythroid differentiation (CD71 and CD235a markers; stages I-IV) for cells from patient #13 harboring a control shRNA (shCtrl) or shRNAs targeting *DDIT3*, at the indicated time points. (C) Bar-plots representing the percentage of cells observed in B in stages I-IV at 7 and 13 days of differentiation. (D) Normalized expression of hemoglobins (top) and stem cell genes (bottom) of cells transduced with a shRNA control or and shRNA targeting *DDIT3* and subjected to 7 days of ex vivo differentiation. (E) Heatmap of z-scores of genes characteristic of proerythroblasts, early and late basophilic erythroblasts (left), and of genes expressed in poly- and orthochromatic stages (right), for cells transduced with a shRNA control, or an shRNA targeting *DDIT3* and subjected to 7 of ex vivo differentiation.

Figure 1

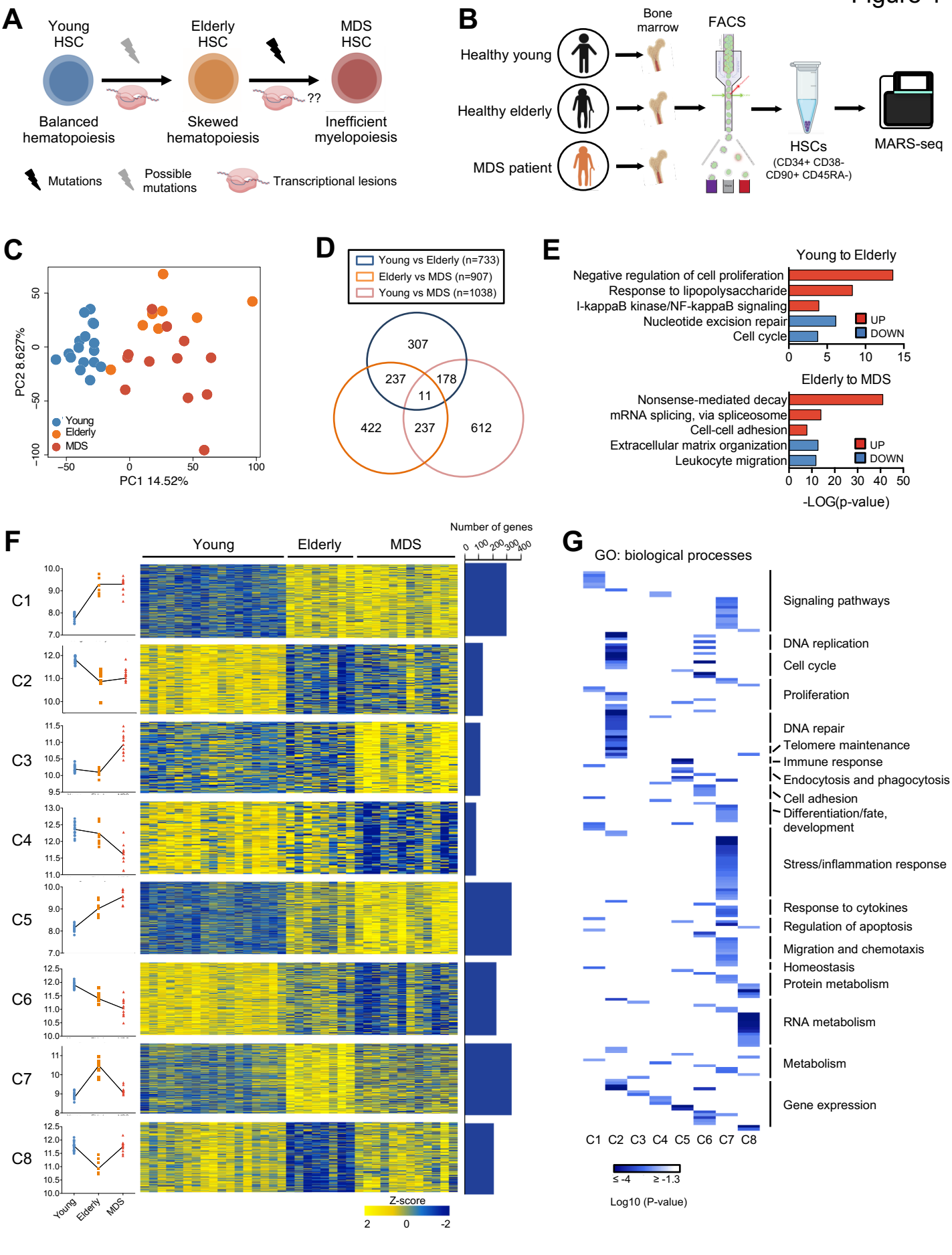

A
Young HSC
Elderly HSC
MDS HSC
??
Balanced hematopoiesis
Skewed hematopoiesis
Inefficient myelopoiesis
Mutations
Possible mutations
Transcriptional lesions
B
Bone marrow
FACS
Healthy young
Healthy elderly
MDS patient
HSCs
(CD34+ CD38- CD90+ CD45RA-)
MARS-seq
C
PC2 8.627%
PC1 14.52%
Young
Elderly
MDS
D
Young vs Elderly (n=733)
Elderly vs MDS (n=907)
Young vs MDS (n=1038)
307
237
178
11
422
237
612
E
Young to Elderly
Negative regulation of cell proliferation
Response to lipopolysaccharide
I-kappaB kinase/NF-kappaB signaling
Nucleotide excision repair
Cell cycle
UP
DOWN
Elderly to MDS
Nonsense-mediated decay
mRNA splicing, via spliceosome
Cell-cell adhesion
Extracellular matrix organization
Leukocyte migration
-LOG(p-value)
F
Young
Elderly
MDS
Number of genes
C1
C2
C3
C4
C5
C6
C7
C8
Z-score
G
GO: biological processes
Signaling pathways
DNA replication
Cell cycle
Proliferation
DNA repair
Telomere maintenance
Immune response
Endocytosis and phagocytosis
Cell adhesion
Differentiation/fate, development
Stress/inflammation response
Response to cytokines
Regulation of apoptosis
Migration and chemotaxis
Homeostasis
Protein metabolism
RNA metabolism
Metabolism
Gene expression
Log10 (P-value)

Figure 2

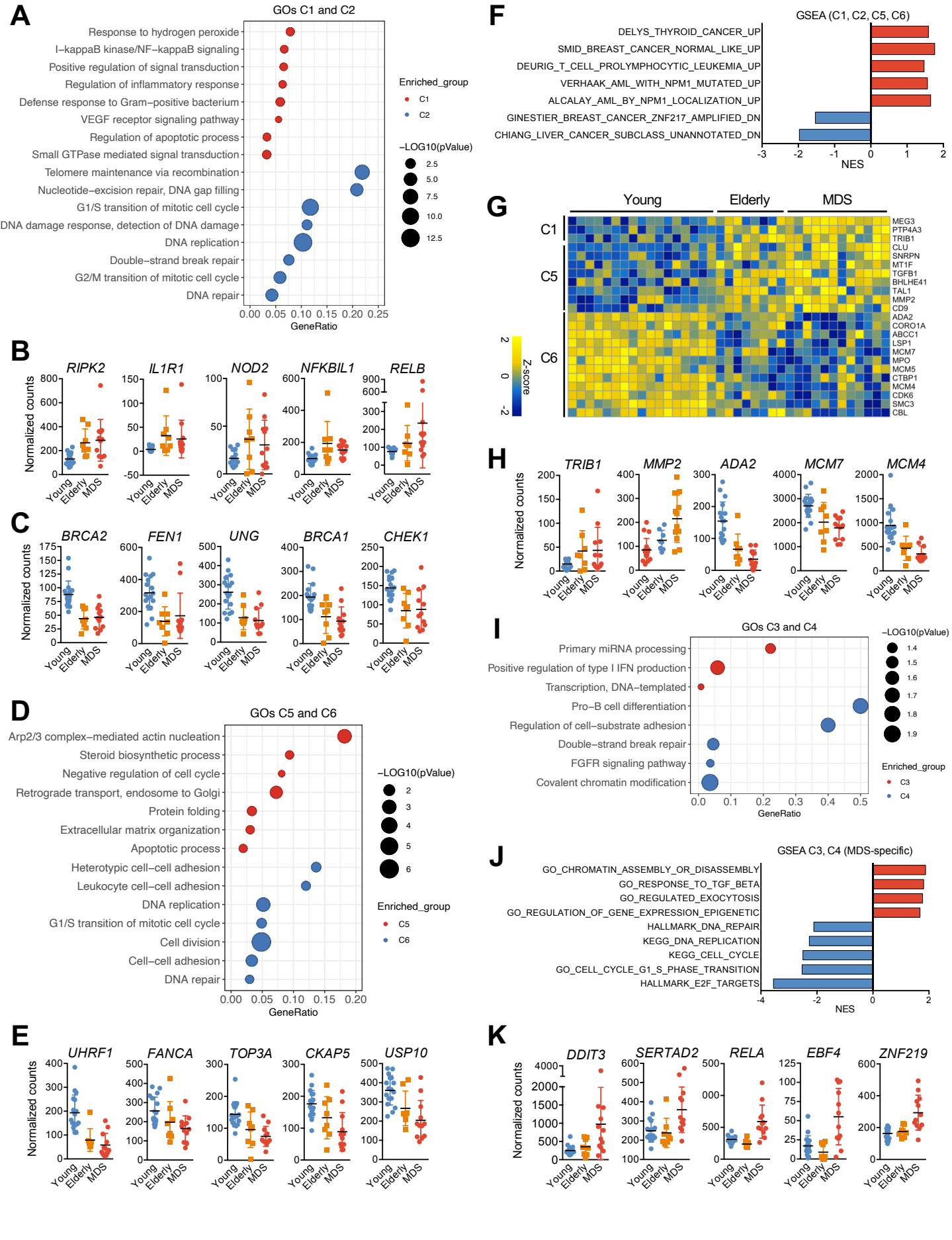
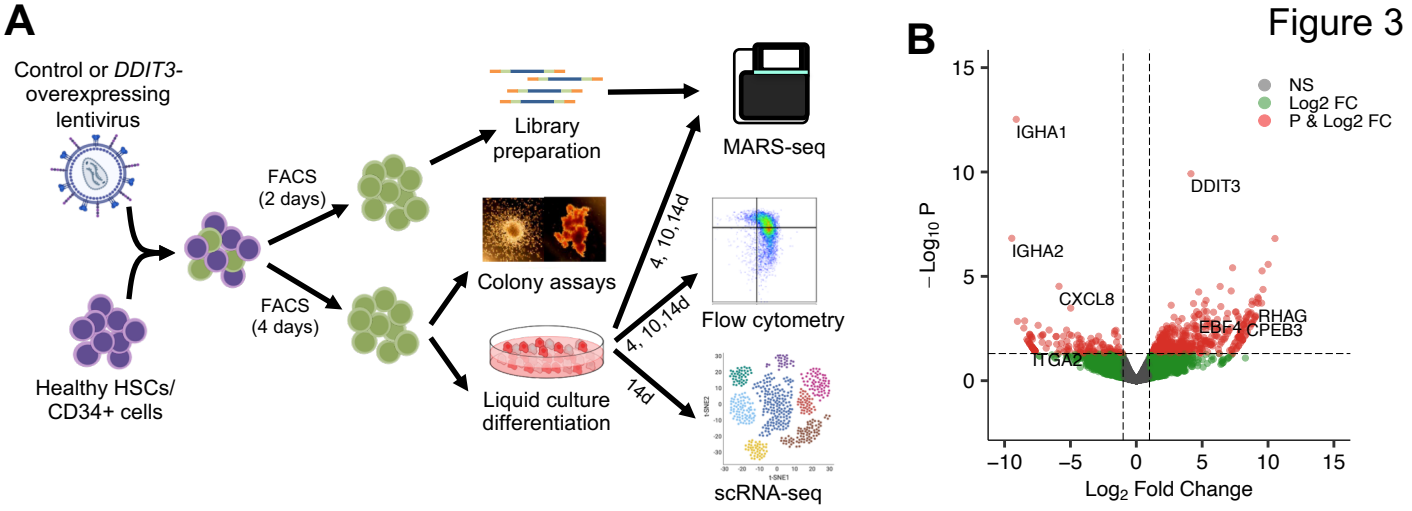
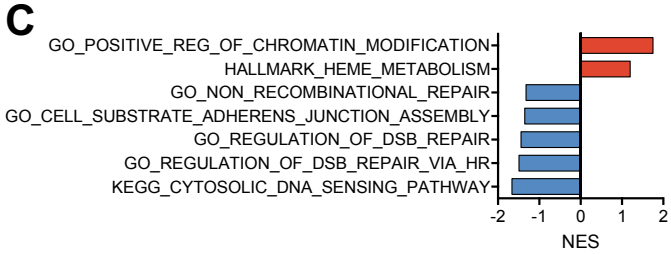
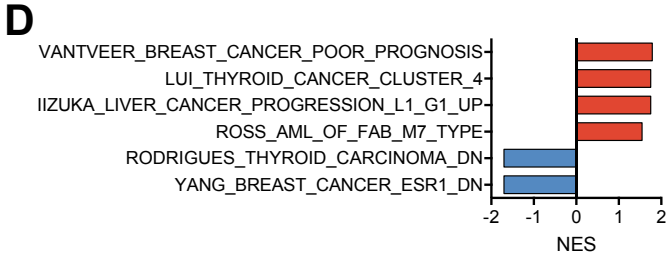
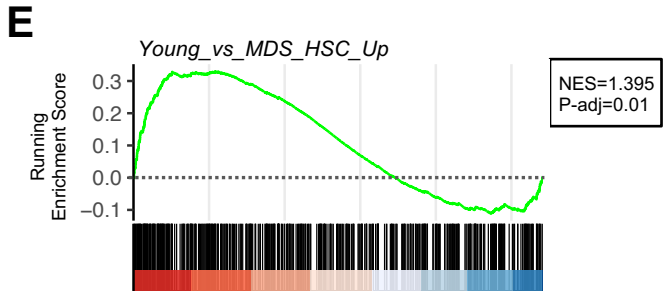

A
GOs C1 and C2
Response to hydrogen peroxide
I−kappaB kinase/NF−kappaB signaling
Positive regulation of signal transduction
Regulation of inflammatory response
Defense response to Gram−positive bacterium
VEGF receptor signaling pathway
Regulation of apoptotic process
Small GTPase mediated signal transduction
Telomere maintenance via recombination
Nucleotide−excision repair, DNA gap filling
G1/S transition of mitotic cell cycle
DNA damage response, detection of DNA damage
DNA replication
Double−strand break repair
G2/M transition of mitotic cell cycle
DNA repair
0.00 0.05 0.10 0.15 0.20 0.25
GeneRatio
Enriched_group
C1
C2
−LOG10(pValue)
2.5
5.0
7.5
10.0
12.5
B
RIPK2
IL1R1
NOD2
NFKBIL1
RELB
Normalized counts
Young
Elderly
MDS
C
BRCA2
FEN1
UNG
BRCA1
CHEK1
Normalized counts
Young
Elderly
MDS
D
GOs C5 and C6
Arp2/3 complex−mediated actin nucleation
Steroid biosynthetic process
Negative regulation of cell cycle
Retrograde transport, endosome to Golgi
Protein folding
Extracellular matrix organization
Apoptotic process
Heterotypic cell−cell adhesion
Leukocyte cell−cell adhesion
DNA replication
G1/S transition of mitotic cell cycle
Cell division
Cell−cell adhesion
DNA repair
0.00 0.05 0.10 0.15 0.20
GeneRatio
−LOG10(pValue)
2
3
4
5
6
Enriched_group
C5
C6
E
UHRF1
FANCA
TOP3A
CKAP5
USP10
Normalized counts
Young
Elderly
MDS
F
GSEA (C1, C2, C5, C6)
DELYS_THYROID_CANCER_UP
SMID_BREAST_CANCER_NORMAL_LIKE_UP
DEURIG_T_CELL_PROLYMPHOCYTIC_LEUKEMIA_UP
VERHAAK_AML_WITH_NPM1_MUTATED_UP
ALCALAY_AML_BY_NPM1_LOCALIZATION_UP
GINESTIER_BREAST_CANCER_ZNF217_AMPLIFIED_DN
CHIANG_LIVER_CANCER_SUBCLASS_UNANNOTATED_DN
−3 −2 −1 0 1 2
NES
G
Young
Elderly
MDS
C1
C5
C6
Z-score
2
0
-2
MEG3
PTP4A3
TRIB1
CLU
SNRPN
MT1F
TGFB1
BHLHE41
TAL1
MMP2
CD9
ADA2
CORO1A
ABCC1
LSP1
MCM7
MPO
MCM5
CTBP1
MCM4
CDK6
SMC3
CBL
H
TRIB1
MMP2
ADA2
MCM7
MCM4
Normalized counts
Young
Elderly
MDS
I
GOs C3 and C4
Primary miRNA processing
Positive regulation of type I IFN production
Transcription, DNA−templated
Pro−B cell differentiation
Regulation of cell−substrate adhesion
Double−strand break repair
FGFR signaling pathway
Covalent chromatin modification
0.0 0.1 0.2 0.3 0.4 0.5
GeneRatio
−LOG10(pValue)
1.4
1.5
1.6
1.7
1.8
1.9
Enriched_group
C3
C4
J
GSEA C3, C4 (MDS-specific)
GO_CHROMATIN_ASSEMBLY_OR_DISASSEMBLY
GO_RESPONSE_TO_TGF_BETA
GO_REGULATED_EXOCYTOSIS
GO_REGULATION_OF_GENE_EXPRESSION_EPIGENETIC
HALLMARK_DNA_REPAIR
KEGG_DNA_REPLICATION
KEGG_CELL_CYCLE
GO_CELL_CYCLE_G1_S_PHASE_TRANSITION
HALLMARK_E2F_TARGETS
−4 −2 0 2
NES
K
DDIT3
SERTAD2
RELA
EBF4
ZNF219
Normalized counts
Young
Elderly
MDS

Figure 3


A
Control or DDIT3-overexpressing lentivirus
Healthy HSCs/ CD34+ cells
FACS (2 days)
FACS (4 days)
Library preparation
Colony assays
Liquid culture differentiation
MARS-seq
Flow cytometry
scRNA-seq
4, 10,14d
4, 10,14d
14d
B
NS
Log2 FC
P & Log2 FC
IGHA1
IGHA2
CXCL8
ITGA2
DDIT3
EBF4
CPEB3
RHAG
−Log10 P
Log2 Fold Change


C
GO_POSITIVE_REG_OF_CHROMATIN_MODIFICATION
HALLMARK_HEME_METABOLISM
GO_NON_RECOMBINATIONAL_REPAIR
GO_CELL_SUBSTRATE_ADHERENS_JUNCTION_ASSEMBLY
GO_REGULATION_OF_DSB_REPAIR
GO_REGULATION_OF_DSB_REPAIR_VIA_HR
KEGG_CYTOSOLIC_DNA_SENSING_PATHWAY
NES


D
VANTVEER_BREAST_CANCER_POOR_PROGNOSIS
LUI_THYROID_CANCER_CLUSTER_4
IIZUKA_LIVER_CANCER_PROGRESSION_L1_G1_UP
ROSS_AML_OF_FAB_M7_TYPE
RODRIGUES_THYROID_CARCINOMA_DN
YANG_BREAST_CANCER_ESR1_DN
NES


E
Young_vs_MDS_HSC_Up
Running Enrichment Score
NES=1.395
P-adj=0.01

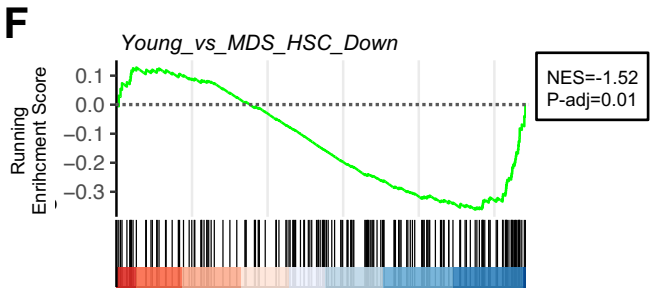

F
Young_vs_MDS_HSC_Down
Running Enrichment Score
NES=-1.52
P-adj=0.01

Figure 4

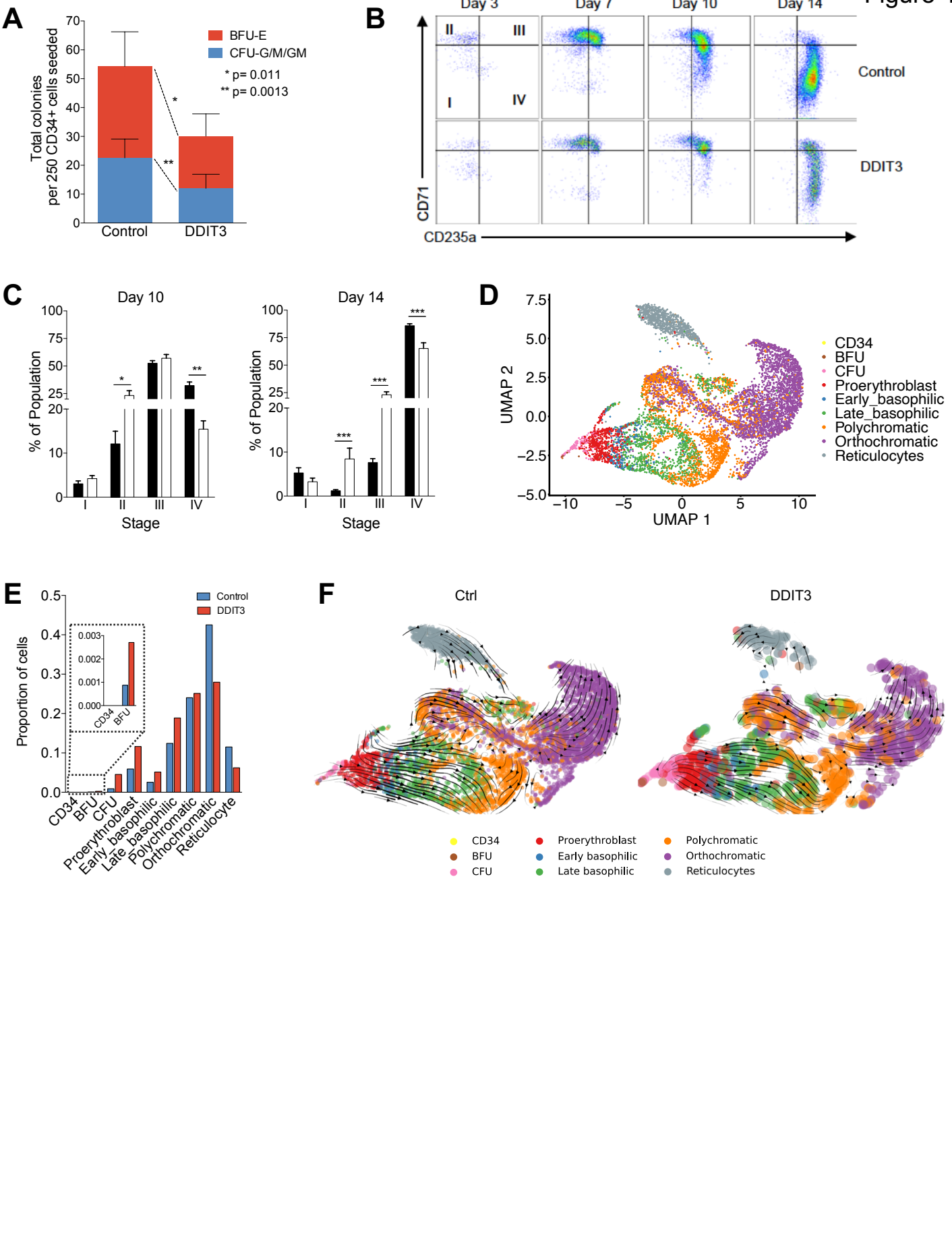

A
Total colonies per 250 CD34+ cells seeded
BFU-E
CFU-G/M/GM
* p= 0.011
** p= 0.0013
Control
DDIT3
B
Day 3
Day 7
Day 10
Day 14
I
II
III
IV
Control
DDIT3
CD71
CD235a
C
Day 10
Day 14
% of Population
Stage
D
UMAP 1
UMAP 2
CD34
BFU
CFU
Proerythroblast
Early_basophilic
Late_basophilic
Polychromatic
Orthochromatic
Reticulocytes
E
Proportion of cells
Control
DDIT3
CD34
BFU
CFU
Proerythroblast
Early_basophilic
Late_basophilic
Polychromatic
Orthochromatic
Reticulocyte
F
Ctrl
DDIT3
CD34
BFU
CFU
Proerythroblast
Early basophilic
Late basophilic
Polychromatic
Orthochromatic
Reticulocytes

Figure 5

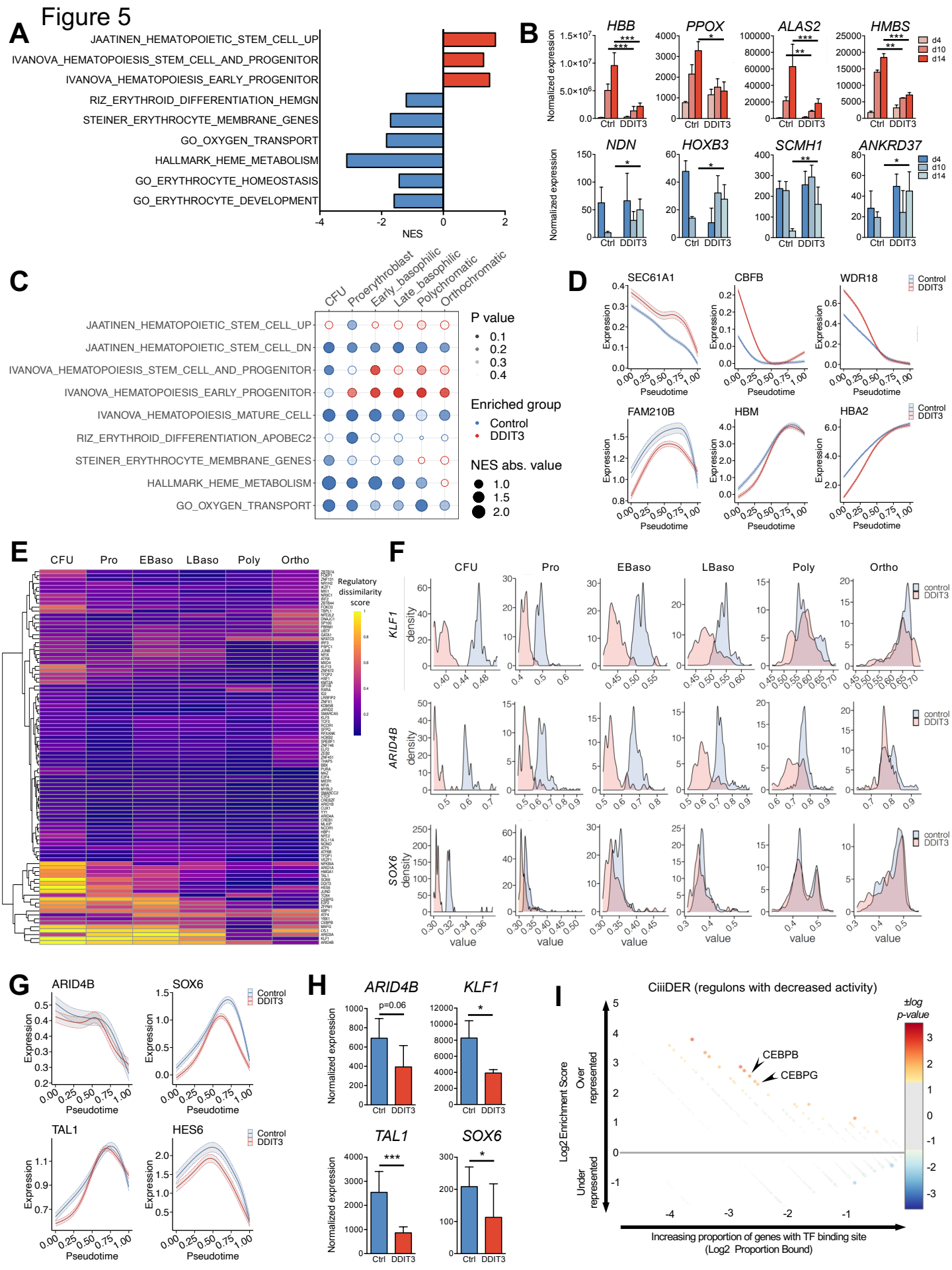
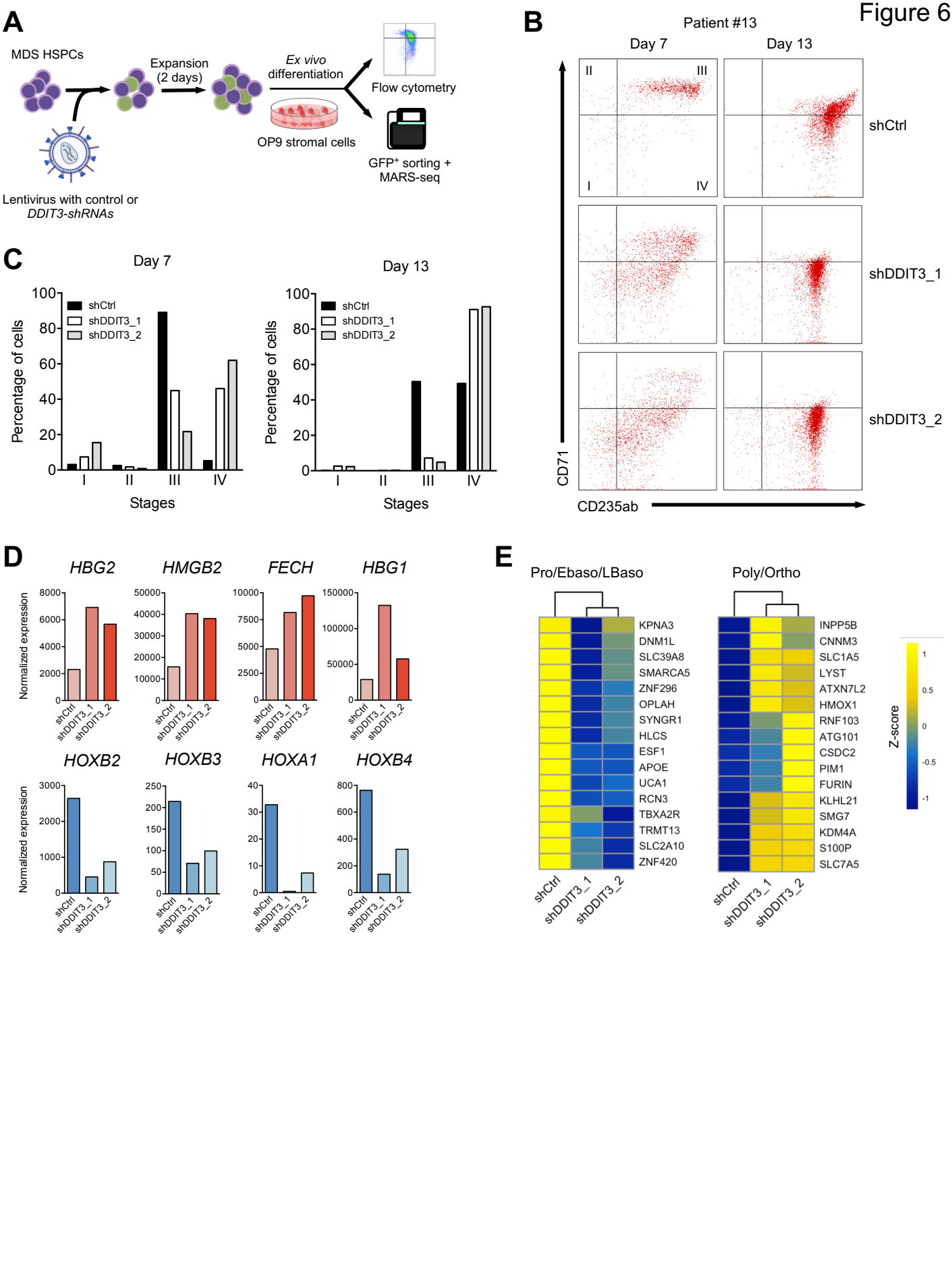

A
JAATINEN_HEMATOPOIETIC_STEM_CELL_UP
IVANOVA_HEMATOPOIESIS_STEM_CELL_AND_PROGENITOR
IVANOVA_HEMATOPOIESIS_EARLY_PROGENITOR
RIZ_ERYTHROID_DIFFERENTIATION_HEMGN
STEINER_ERYTHROCYTE_MEMBRANE_GENES
GO_OXYGEN_TRANSPORT
HALLMARK_HEME_METABOLISM
GO_ERYTHROCYTE_HOMEOSTASIS
GO_ERYTHROCYTE_DEVELOPMENT
-4
-2
0
2
NES
B
HBB
PPOX
ALAS2
HMBS
NDN
HOXB3
SCMH1
ANKRD37
Normalized expression
Ctrl
DDIT3
d4
d10
d14
C
CFU
Proerythroblast
Early_basophilic
Late_basophilic
Polychromatic
Orthochromatic
JAATINEN_HEMATOPOIETIC_STEM_CELL_UP
JAATINEN_HEMATOPOIETIC_STEM_CELL_DN
IVANOVA_HEMATOPOIESIS_STEM_CELL_AND_PROGENITOR
IVANOVA_HEMATOPOIESIS_EARLY_PROGENITOR
IVANOVA_HEMATOPOIESIS_MATURE_CELL
RIZ_ERYTHROID_DIFFERENTIATION_APOBEC2
STEINER_ERYTHROCYTE_MEMBRANE_GENES
HALLMARK_HEME_METABOLISM
GO_OXYGEN_TRANSPORT
P value
0.1
0.2
0.3
0.4
Enriched group
Control
DDIT3
NES abs. value
1.0
1.5
2.0
D
SEC61A1
CBFB
WDR18
FAM210B
HBM
HBA2
Expression
Pseudotime
Control
DDIT3
E
CFU
Pro
EBaso
LBaso
Poly
Ortho
Regulatory dissimilarity score
F
CFU
Pro
EBaso
LBaso
Poly
Ortho
KLF1
ARID4B
SOX6
density
value
control
DDIT3
G
ARID4B
SOX6
TAL1
HES6
Expression
Pseudotime
Control
DDIT3
H
ARID4B
KLF1
TAL1
SOX6
Normalized expression
p=0.06
Ctrl
DDIT3
I
CiiiDER (regulons with decreased activity)
±log p-value
CEBPB
CEBPG
Log2 Enrichment Score
Over represented
Under represented
Increasing proportion of genes with TF binding site
(Log2 Proportion Bound)

Figure 6


A
MDS HSPCs
Expansion (2 days)
Ex vivo differentiation
Flow cytometry
OP9 stromal cells
GFP+ sorting + MARS-seq
Lentivirus with control or DDIT3-shRNAs
B
Patient #13
Day 7
Day 13
shCtrl
shDDIT3_1
shDDIT3_2
CD71
CD235ab
I
II
III
IV
C
Day 7
Day 13
shCtrl
shDDIT3_1
shDDIT3_2
Percentage of cells
Stages
D
HBG2
HMGB2
FECH
HBG1
HOXB2
HOXB3
HOXA1
HOXB4
Normalized expression
shCtrl
shDDIT3_1
shDDIT3_2
E
Pro/Ebaso/LBaso
Poly/Ortho
KPNA3
DNM1L
SLC39A8
SMARCA5
ZNF296
OPLAH
SYNGR1
HLCS
ESF1
APOE
UCA1
RCN3
TBXA2R
TRMT13
SLC2A10
ZNF420
INPP5B
CNNM3
SLC1A5
LYST
ATXN7L2
HMOX1
RNF103
ATG101
CSDC2
PIM1
FURIN
KLHL21
SMG7
KDM4A
S100P
SLC7A5
Z-score