MODELLING AND TARGETING EPIGENETIC REGULATORS IN ACUTE LEUKEMIA

INAUGURALDISSERTATION

ZUR

ERLANGUNG DER WÜRDE EINES DOKTORS DER PHILOSOPHIE

VORGELEGT

DER PHILOSOPHISCH-NATURWISSENSCHAFTLICHEN FAKULTÄT

DER UNIVERSITÄT BASEL

VON

KATHARINA LUCIE LEONARDS, GEB. DUMRESE

AUS TRIER, DEUTSCHLAND

BASEL, 2016

Genehmigt von der Philosophisch- Naturwissenschaftlichen Fakultät

auf Antrag von
Dissertationsleiter: Prof. MD Jürg Schwaller
Fakultätsvertreter: Prof. dr. Markus Affolter
Korreferent: Prof. dr. Thomas Mercher

Basel, den 18.10.2016

Prof. Dr. Jörg Schibler (Dekan)

To Christopher.

1 TABLE OF CONTENTS

2 SUMMARY

A complex network of transcription factors controls self-renewal and differentiation of hematopoietic stem cells (HSC). Mutations or translocations of these epigenetic regulators may result in malignant transformation leading to acute leukemia. A significant fraction of childhood acute myeloid leukemia (AML) patients carry a translocation involving the nuclear receptor-binding SET domain protein 1 (NSD1) histone methyltransferase. To understand its function we ablated the gene in the hematopoietic system of the mouse. Surprisingly, all *"Nsd1-null"* (*Vav1-iCre;Nsd1*$^{fl/fl}$) mice developed a lethal, malignant accumulation of CD71$^{dim/+}$ TER119^{-} erythroid progenitor cells with aberrant clonogenic activity and impaired terminal maturation of self-renewing erythroblasts *in vitro*, a phenotype that resembles human acute erythroleukemia. The lack of *Nsd1* also reduced the number of HSC starting during fetal liver hematopoiesis. Although gene expression signatures revealed reduced mRNA expression of the erythroid master transcription factor *Gata1*, erythroblasts of *Vav1-iCre;Nsd1*$^{fl/fl}$ mice expressed constitutively high levels of GATA1 protein. Interestingly, the cells were significantly impaired in activation but still able to repress several known GATA1 targets. Strikingly, retroviral overexpression of *Gata1* induced terminal maturation of *Vav1-iCre;Nsd1*$^{fl/fl}$ proerythroblasts which was associated with activation of GATA1 target genes.

Knockdown of *NSD1* in human adult or cord-blood derived CD34^{+} HSC cells also impaired erythroid differentiation associated with increased protein levels of GATA1. In addition, we found high GATA1 protein levels in several human erythroleukemia cell lines suggesting a key role in aberrant erythroid differentiation. Currently ongoing experiments aim to mechanistically understand Nsd1-mediated GATA1 regulation and erythroid differentiation. Preliminary observations with peptide array-based *in vitro* methylation assays suggest the possibility for direct methylation of GATA1 by NSD1. We also found aberrant expression of erythroid- associated transcription factor complex members with increased levels of GATA1 and ETO2 but reduced levels of TAL1, E2A and LBD1. Moreover, significant changes in global histone H3K36 methylation were seen in proerythroblasts lacking Nsd1. Taken together, our data so far revealed Nsd1 as a novel regulator of normal and malignant erythropoiesis. Ongoing studies may not only provide mechanistic insights of aberrant transcriptional regulation leading to erythroleukemia but could also set the

base to develop novel therapies against this rare but very aggressive disease that is currently incurable in most patients.

Next to histone methyltransferases like NSD1, histone acetyltransferases like CBP/p300 are recurrently involved in AML-associated chromosomal translocations and also serve as co-activators of other fusion oncogenes, suggesting therapeutic potential of specific targeting of CBP/p300. We characterized the anti-leukemic potential of I-CBP112, a novel small molecule chemical probe that selectively binds the CBP/p300 bromodomain (BRD). BRDs belong to a diverse family of evolutionary conserved protein-interaction modules recognizing acetylated lysine residues and thereby mediating recruitment of proteins to macromolecular complexes. We found that I-CBP112 significantly impaired the clonogenic activity of a series of murine cell lines immortalized by the MLL-CBP fusion and other leukemic fusion oncogenes (MLL-AF9, MLL-ENL, NUP98-HOXA9) in a dose-dependent manner. Similar to the murine cells, we found that I-CBP112 did not cause immediate cytotoxic effects but impaired colony formation and induced cellular differentiation of a series of 18 human leukemic cell lines. Likewise, I-CBP112 also reduced colony formation of human primary AML blasts but not of normal $CD34^+$ HSC. Importantly, combination of I-CBP112 with another BRD inhibitor targeting BET proteins (JQ1) or with a chemotherapeutic agent (Doxorubicin) revealed clear synergistic effects on cell survival of the AML cell lines. Extreme limited dilution assays in methylcellulose, as well as bone marrow transplantation experiments revealed that I-CBP112 significantly impaired self-renewal of leukemic stem cells *in vitro* and reduced the leukemia-initiating potential *in vivo*. Taken together, we found that selective interference with the CBP/p300 BRD by I-CBP112 has the potential to selectively target leukemic stem cells and opens the way for novel combinatory “BRD inhibitor” therapies for AML. In addition to I-CBP112, we tested a pan- bromodomain inhibitor (“bromosporine”, BSP) broadly targeting BRDs including BET. Evaluation of BSP in BET- inhibitor sensitive and non-sensitive leukemic cell-lines revealed strong anti-proliferative activity in semi-solid medium. Similar to treatment with JQ1 (a selective BET inhibitor) BSP arrested in S- cell cycle phase suggesting BET-mediated effects. Finally, non-selective targeting of BRDs by BSP identified BETs as master regulators of primary transcription response in leukemia.

Collectively the experiments of this thesis investigated the role of epigenetic regulators in normal and malignant hematopoiesis and explored strategies for selective interference as novel anti-leukemic therapies.

3 INTRODUCTION

3.1 Murine blood development

First signs for blood formation, or hematopoiesis, during murine embryogenesis are found right after implantation of the blastocyst into the wall of the uterus. The blastocyst is composed of the inner cell mass that gives rise amongst others to the epiblast consisting of epithelial cells. During development, the epiblast undergoes epithelial- mesenchymal- transition (EMT), which in turn results in formation of the three germ layers endoderm, ectoderm and mesoderm at embryonic day (E) 6.5, a process called gastrulation (**Figure 1A**). Subsequently the mesoderm produces hemangioblasts that migrate to the yolk sac in order to commit towards endothelial or hematopoietic progenitors [1]. At E7.5 the hematopoietic cells arise in so- called „blood islands“ producing large, nucleated, megaloblastic, primitive cells of the erythroid lineage (EryP) [2]. Production of red blood cells is of particular importance to ensure growth and survival of the embryo and later fetus. Moreover, macrophages and megakaryocytes can be detected at this developmental stage [3,4]. At E10.5, the aorta–gonad–mesonephros (AGM) region generates the first hematopoietic stem cells (HSC) [5]. Around E12.5, hematopoiesis is shifted to the developing fetal liver organ where enucleated definitive erythroid (EryD) cells are generated (**Figure 1B**) [6]. In so- called “erythroblastic islands” erythroblasts undergo complete maturation with the help of a macrophage to which the cells attach [7]. Whereas EryP cells predominantly express fetal hemoglobin chains, EryD cells express already adult globin [8]. The fetal liver also provides an environment favoring hematopoietic stem cell (HSC) development and expansion before cells finally migrate to the BM and become quiescent (own observations; see **Figure 1C**) [9,10]. Around the time of birth at E19.5, HSCs can be also detected in spleen tissue and the BM contains hematopoietic cells of different lineages (own observations, see **Figure 1B-C**) [5,9]. In young and adult mice, the BM contributes to steady- state hematopoiesis whereas the spleen is later primarily used for extramedullary stress hematopoiesis [5].

Hematopoiesis functions as a dynamic system continuously trying to build up equilibrium of supply and demand. Shortage or imbalance is sensed by extracellular

signals such as cytokines or by communication with neighboring cells residing with the hematopoietic cells in a niche in the BM.

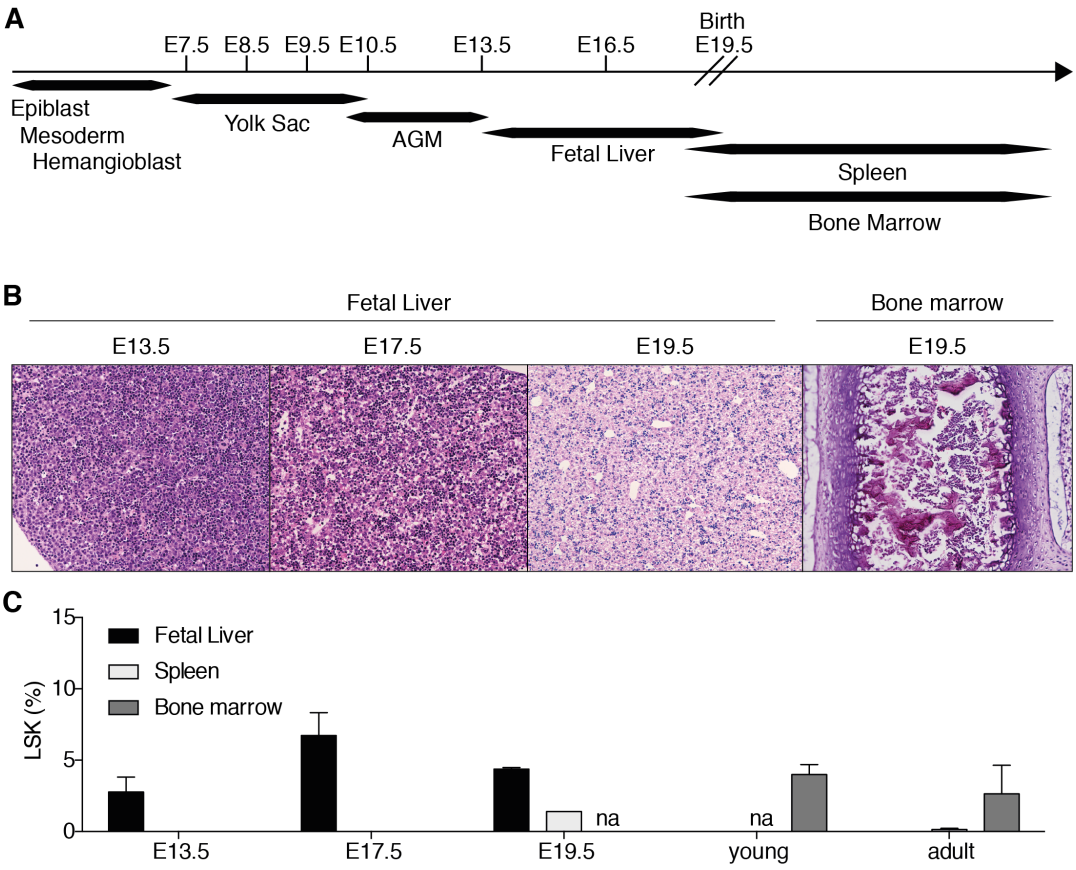


FIGURE 1. Embryonic development of hematopoiesis.

(A) Timeline of embryonic development of murine hematopoiesis starting after implantation of blastocyst. The epiblasts gives rise to mesoderm from which the hemangioblast develops. First hematopoietic maturated cells arise at embryonic (E) day 7.5 from the yolk sac. At E10.5, HSC arise from aorta–gonad–mesonephros (AGM) region before hematopoiesis shifts to fetal liver at E13.5. At E19.5, mice are born and hematopoiesis is shifted to BM and partially to the spleen. (B) Hematoxylin-Eosin staining of fetal liver sections of murine embryos from E13.5 to 19.5 and BM at E19.5 reflecting developmental stages and appearance of hematopoietic cells. (C) Bar graph reflecting percentages of Lineage$^-$, Sca-1$^+$, c-Kit$^+$ (LSK) population by flow cytometry staining hematopoietic stem cells in fetal liver as well as BM and spleen of young (4 weeks after birth) and adult (10- 15 weeks after birth) mice.

For example after bleeding, especially red blood cells (erythrocytes) need to be manufactured immediately to compensate the loss. Erythrocytes contain hemoglobin, an iron- containing molecule able to bind oxygen and transport it through the blood flow to all tissues. If a large amount of erythrocytes is lost, a state of low oxygen, called hypoxia, is sensed in the kidney and liver by stabilization of the hypoxia

inducible factor 1 alpha (HIF1α). HIF1α then forms a heterodimer with HIF1β and together they bind to DNA and induce expression of the erythropoietin (EPO) gene. EPO binds to the EPO receptor (EPOR) and induces a signaling cascade via Janus Kinase 2 (JAK2)- Signal Transducers and Activators of Transcription (STAT) finally resulting in upregulation of a large number of genes regulating erythroid differentiation [11].

Extracellular signals are essential to maintain differentiation and proliferation of hematopoietic cells. Stromal cell-derived factor 1 (SDF-1, also known as CXLX12), stem cell factor (SCF) and thrombopoietin (TPO) have been shown to be necessary for maintenance of HSCs *in vivo* [12]. Moreover regulatory factors such as TGF-beta or other signaling pathways such as Wnt or Notch are involved in regulation of the HSC niche [13,14].

Flow cytometry with antibodies recognizing specific surface marker expression allows to immunophenotype HSCs and their differentiated progeny. In the hematopoietic hierarchy, HSC give rise to progenitor cells that ultimately differentiate into different lineages. Lymphocytes are separated into B-, T-, natural killer and dendritic cells and myelocytes into macrophages, granulocytes, platelets and erythrocytes. A summary of surface markers for each lineage is depicted in **Figure 2**. Restriction to the B-cell lineage can be primarily recognized by staining for B220, to T lymphocytes by staining for cluster of differentiation (CD) 3, 4 and 8, to macrophages for Mac-1, to granulocytes for Gr-1, to platelets for CD41 and to erythrocytes for Ter-119. In order to distinguish stem and progenitor cells from more differentiated cells, lineage marker^{low} cells are additionally stained with antibodies recognizing stem cell antigen 1 (Sca-1) (=LS) and tyrosine kinase kit (c-Kit, also known as CD117) (=LSK). Moreover, CD34 and signaling lymphocytic activation molecule (SLAM) family markers CD150 and CD48 are used to distinguish multi-potent progenitors (MPP) from long- term (LT-HSC) and short- tem repopulating stem cells (ST- HSC) [15–17]. Myeloid progenitors are separated by CD34 and Fc-gamma-receptor (FcγR) into common myeloid, granulocyte- macrophage and megakaryocyte-erythroid progenitors [18,19]. Importantly, the “hematopoietic hierarchy tree” is only a theoretic model. There are constant changes and discussion about origin of cells, e. g. the “myeloid bypass model” involving a direct link of LT-HSC to megakaryocytic, erythroid and myeloid progenitor circumventing the “general” path [20–24].

Taken together, hematopoiesis is a process starting early in development to ensure generation of mature blood cells. It functions in a highly organized, hierarchical manner and needs regulation of signaling pathways within lineages that is reflected in expression of surface markers.

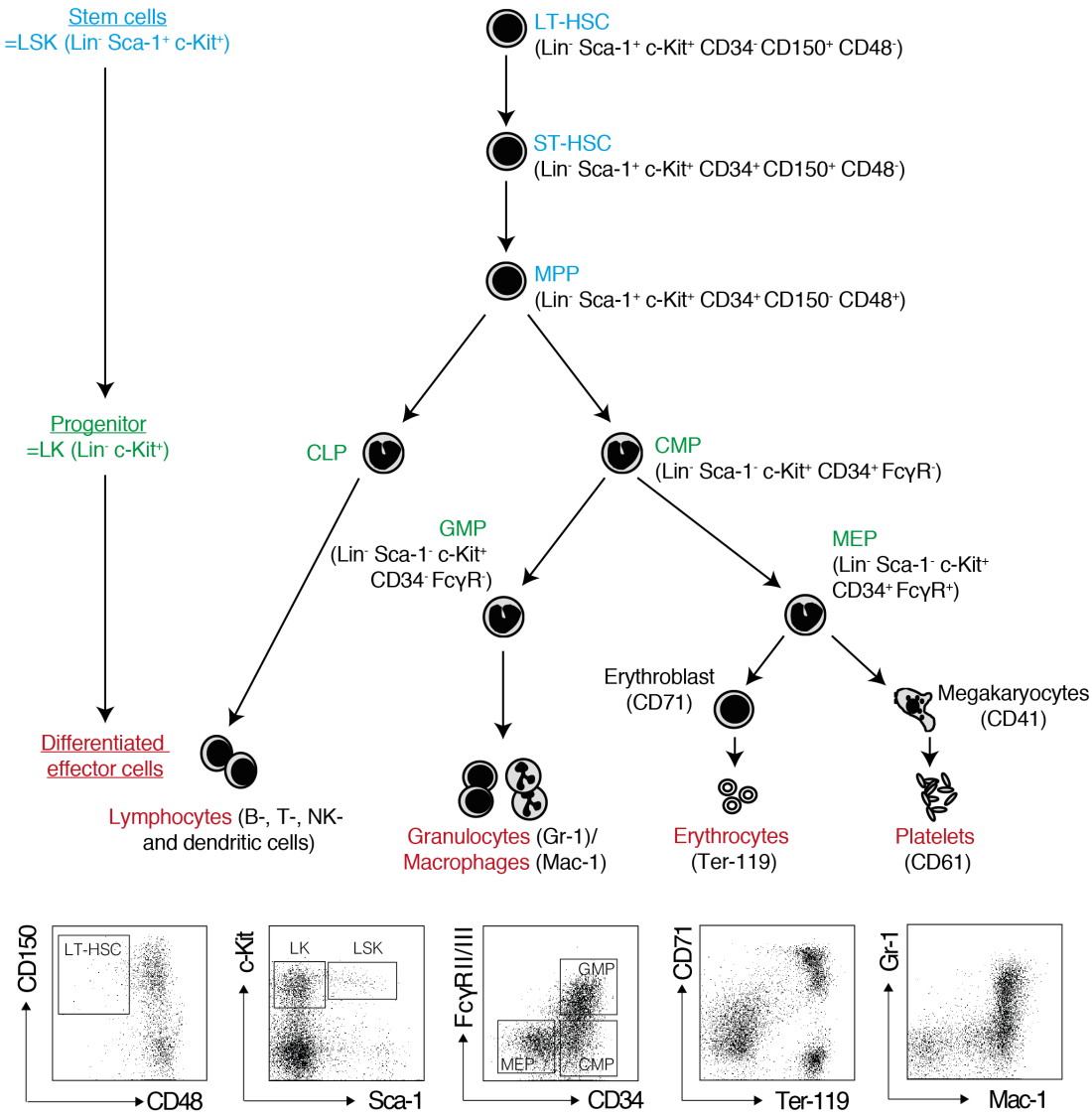


FIGURE 2. Schematic depiction and gating strategy of hematopoietic lineage hierarchy based on flow cytometry staining.

Stem cells are classified as Lin$^-$, Sca-1$^+$, c-Kit$^+$ (LSK) and include long term repopulating stem cells (LT-HSC), short term repopulating stem cells (ST-HSC) and multipotent progenitor cells (MPP) that are distinguished using CD34, CD150 and CD48 surface markers. Myeloid progenitors are within Lin$^-$, c-Kit$^+$ (LK) poulation distinguished with CD34 and FcγR. Differentiated cells include lymphocytes, myelo-granulocytes, erythrocytes and platelets.

3.2 Transcriptional regulation of hematopoiesis

The high cellular demand urges HSCs to balance their output to produce sufficient mature blood cells as well as to maintain their own pool at any time. In order to execute the gene expression programs needed for lineage choice and differentiation, extracellular signaling activates pathways that are regulating transcriptional regulation of hematopoiesis [9]. Interestingly, HSCs constitutively express low levels of multiple cytokine receptors and transcription factors (TF) that get sequentially restricted during lineage fate decision [19]. There is great controversy regarding lineage fate decisions. It has been published that the relative amount of the transcription factors GATA1 and PU.1, which are explained in detail later, account for the earliest lineage decisions at the CMP level, deciding whether cells go into the direction of myeloid (GMP) or megakaryocyte/erythroid (MEP) differentiation [25,26]. Graf & Enver concluded in a "bi-colored marble model" that progenitor cells co-expressing the factors fluctuate between the decisions and finally decide for one direction, which would be a binary decision [27]. To date it is not clear whether lineage choice is controlled by stochastic events, or whether external factors regulate the fate. Most likely, this decision is a combination of both [28].

In general, transcription often works through binding of the basal machinery including RNA polymerase to the TATA (TATAAA consensus sequence) box in the promoter element of the DNA. In order to fine-tune the output, this machinery is also in contact with proteins bound to distal control elements, such as enhancers or repressors [29]. The key feature of transcription factors is to facilitate or to block the access of the RNA polymerase to DNA in order to enhance or suppress mRNA transcription. There are two minimal necessary domains for this action needed. One one hand the DNA- binding domain (DBD) ensures binding to cognate motifs within the promoter or distal enhancer elements. On the other hand, a transactivating domain (TAD), also called activation function (AF), is needed to bind co-regulators. There are different ways how to regulate TFs. First a TF needs to be expressed and each gene has regulatory domains to control the output. Interestingly in this context, TFs can bind the promoter sequence of their own gene. Second, in order to be active the TF needs to be localized in the nucleus. Third, some transcription factors might

need to harbor post- translational modifications determining their activation status. Fourth, cofactors can regulate their function or TFs need to form dimers. Fifth, TFs might compete for the same locus and antagonize each other [30]. **Figure 3** depicts an arbitrary selection of TFs in hematopoietic lineage choices that have been extensively studied using mostly murine knockout models. Our current knowledge about TFs mostly results from studies using genetically engineered embryonic stem (ES) cells and murine models [31–33]. Whether a TF actually binds to chromatin, is studied using electrophoretic mobility shift assay (EMSA) or chromatin immunoprecipitation (ChIP). This technique can identify specific loci that are bound by one or more TF, such as the case for GATA1 and GATA2 [31,34,35] .

Interestingly, single- cell ChIP-Sequencing (ChIP-Seq) has been recently developed in order to study histone marks in embryonic stem cells [36]. Regarding transcription factor binding, small population (500 cells) ChIP-Seq revealed heterogeneity in sorted populations during hematopoietic differentiation [37]. In the following section, TFs are highlighted that are involved in more than one lineage choice.

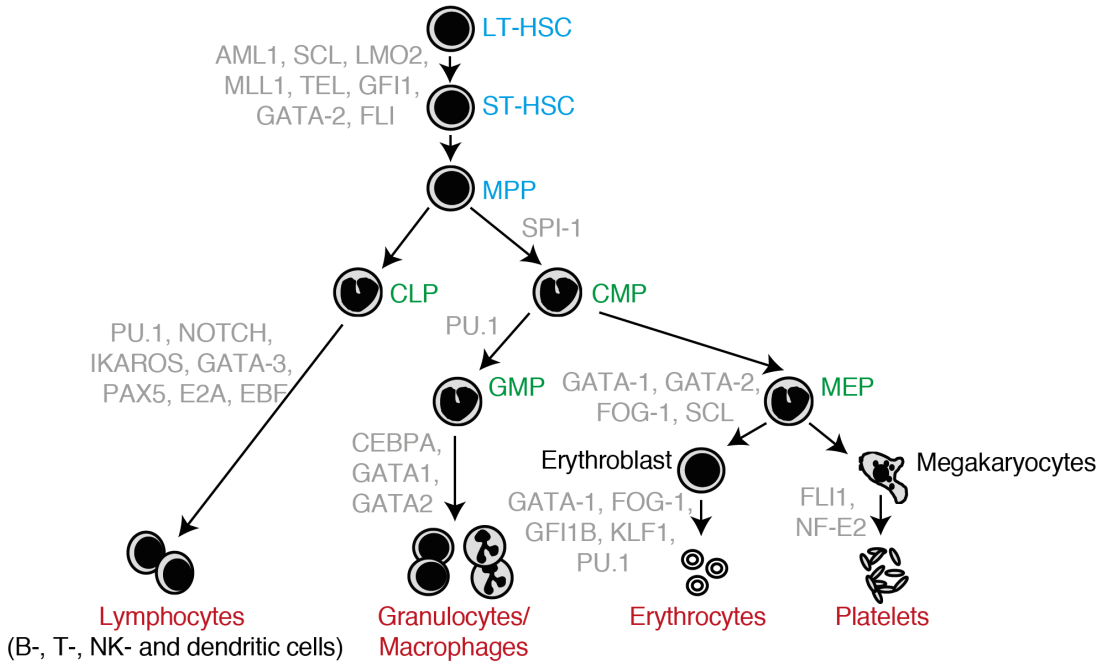


FIGURE 3. Schematic depiction of TFs in hematopoiesis involved in lineage decision execution. (adapted from [33])

3.2.1 GATA1: a master TF regulating erythro- megakaryocytic lineage

Probably one of the best- studied TF in hematopoiesis is GATA1. GATA1 was discovered as the founding member of the GATA family including six other factors of which three (GATA1-3) are implicated in hematopoietic development. The human *GATA1* and murine *Gata1* genes are both located on the X chromosome and are expressed primarily in more mature cell types such as erythroid cells, megakaryocytes, eosinophils, mast cells and in the Sertoli cells of the testis [38–45]. Recently, it has been shown that GATA1 is also expressed at low levels in stem cells [25,46].

The *GATA1* gene locus is composed of different enhancer and promoter elements regulating its expression. It contains two alternative first exons and five more exons contributing to the translation of the protein. The alternative first exons called “IT” and “IE” regulate the expression of the gene in the testis or in erythroid cells respectively [47]. Moreover, a hematopoietic enhancer element (“G1HE”) containing a GATA motif serves as a docking site for the TF itself (**Figure 4**) [48].

The *GATA1* gene translates into a 413aa protein bearing the TAD near the N-terminus and two DNA- binding domains in form of zinc finger motifs, called N- and C- finger due to their location to the adjacent terminus [49,50]. GATA1 binds with its C-finger to (A/T) GATA (A/G) consensus DNA sequences and thereby regulates genes participating in erythroid differentiation and development [49,51–53]. As mentioned, it can bind its own gene locus and is known to negatively regulate its own expression. The N- finger is involved in co- factor binding occurring also in collaboration with the C-finger [32,54]. Interestingly, there is a report demonstrating self- association capability of GATA1 resulting in dimer formation. Critical domains for association have been found to be downstream of the N- and C- finger, in particular aa224-254 and aa278-318. If mutated, transcription activation is impaired [54].

GATA1 has been discovered through its binding to the globin enhancer locus and was shown to be essential for erythropoietic development [32,55]. Interestingly, its cellular localization seems to depend on the activation status of the cell. In undifferentiated erythroid cells, the protein is predominantly located in the cytoplasm

and during activation reshuffles to the nucleus where it remains excluded from nucleoli [56,57]. It has been shown that the protein also undergoes multiple post-translational modifications such as phosphorylation, acetylation or SUMOylation. However, the functional consequences of these modifications are not fully understood [58–62]. This might be partially due to the bi-potential role of GATA1 being a transcriptional activator and repressor depending on the cell type or context in development [63]. It has been shown that GATA1 involves many different co- factors to regulate its target genes and can also dimerize [54,64–66]. Finally GATA1 regulates its target genes e.g. by competition for GATA2 bound gene loci, a process called "GATA switch" or by direct protein interaction with PU.1 (also called SPI1) [67–70].

Regarding blood development, GATA1 has been shown to be a major regulator in primitive and definitive erythroid differentiation in the embryo. Deficient embryonic stem (ES) cells were unable to produce mature erythrocytes and arrested at the pro-erythroblastic stage. The severe anemic phenotype present resulted in embryonic lethality around E10.5 despite attempts to rescue developmental block through overexpression of GATA2 [39,71,72]. In another murine model, GATA1 is knocked down by introduction of a neo cassette into the IE promoter locus. As a functional consequence, male mice ($GATA^{-/Y}$), like $GATA^{null}$ embryos, die at E12.5 of severe defects in primitive erythropoiesis starting at E9.5. Female mice bearing a heterozygous mutation were able to survive and dependent on inactivation of the X chromosome, expressed either wildtype or 5% of wildtype GATA1 ($GATA^{1.05/X}$). The mice developed anemia and thrombocytopenia and were prone to develop an acute erythroleukemia (AEL)- and acute lymphoblastic leukemia (ALL) phenotype in adulthood [73,74]. Another murine model involved introduction of a neomycin antibiotic selection cassette in front of the IE promoter and removal of regulatory sites (called "neoΔHS" & „ΔneoΔHS"). The resulting knockdown of GATA1 expression was milder and impaired erythroid maturation. Part of the mice died *in utero* at E12.5-14.5 and the remaining mice were anemic at birth but recovered during adulthood [75]. These experiments suggested that erythroid differentiation highly depends on the GATA1 expression dosage.

Several GATA1 mutations have been found in human diseases. Mutations of the N- terminal region of GATA1 have been reported in transient myeloproliferative disorder (TMD) and Down's syndrome-related acute megakaryocytic leukemia (DS-AMKL). These mutations lead to translation of a shorter GATA1 isoform, called GATA1-s lacking the TAD (**Figure 4**). Therefore GATA1-s can still bind to DNA, but

changes proliferation of megakaryocytic progenitors in fetal liver. The reason for the increased incidence of these mutations in patients with trisomy 21 is unclear [76–81]. Moreover, there are cases of X- linked thrombocytopenia and anemia that were reported to harbor a GATA mutation reflecting the involvement of this protein in the control of megakaryocytic- erythroid differentiation programs [32].

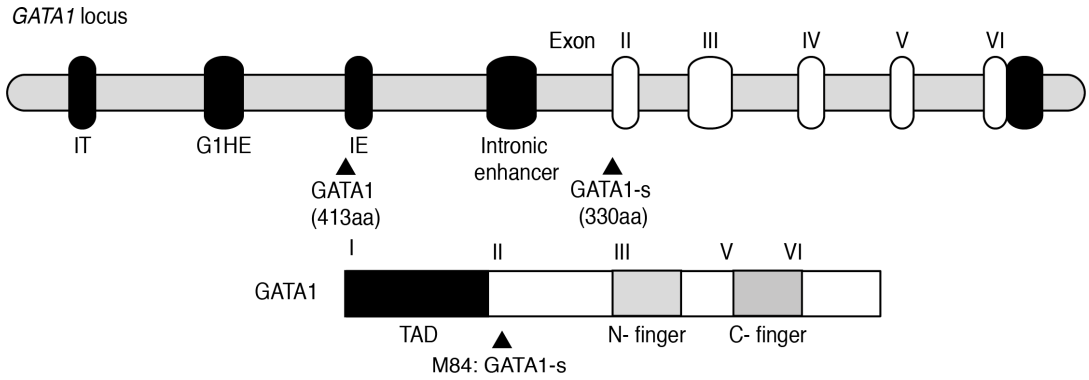


FIGURE 4. Schematic depiction of *GATA1* gene locus and resulting GATA1 protein.
GATA1 gene locus includes two alternative first exons called IT (for testis) and IE (for erythroid). A hematopoietic enhancer (G1HE) element contains a GATA motif and serves as a docking site for autoregulation of the locus. Moreover, the locus contains an intronic enhancer and five more exons (white). GATA1 protein is 413aa long and contains a transactivation domain (TAD) and two zinc fingers (N & C). Above the protein depiction, the exons are marked which contain the respective mRNA later translated into the regions. GATA1-short (s) starts at methionin 84 and translates into a 330aa big protein missing the TAD.

3.2.2 PU.1: key regulator of myeloid fate decisions

The Spleen Focus Forming Virus (SFFV) Proviral Integration Oncogene 1 (=SPI-1, also called PU.1) belongs to the family of E-Twenty-Six (ETS) TF and recognizes the AGAGGAAGTG consensus sequence. It is highly expressed in GMP and also in HSC, CMP and CLP cells at various degrees [25,46].

Inactivation of the SPI-1 gene in mice results in embryonic/ newborn lethality due to blocked maturation of neutrophils, monocytes, B-, T- and NK cell formation whereas megakaryocytic and erythroid development was reported to remain intact. Initially it was assumed that PU.1 is not essential for erythroid development.

However, it was shown that also erythroid progenitors express SPI-1 at very low levels required for proper erythroid homeostasis and that during erythroid differentiation, levels decrease [82–86]. Furthermore, PU.1 was proposed to play a role in HSC maintenance [26].

Several lines of evidence indicate that PU.1 is also a key regulator of leukemogenesis. Erythroid differentiation requires downregulation of PU.1. If disturbed, such as happening in murine erythroleukemia (MEL) cells immortalized by Friend virus or in SPI-1 overexpressing transgenic mice, erythroid differentiation associated targets remain blocked and cause accumulation of immature cells [68,69,87–89]. Furthermore overexpression of SPI-1 in murine erythroblasts inhibits the transition from immature progenitors to pro- erythroblasts due to S- phase blockage and induces an erythroleukemia phenotype in mice [90]. 20% Knockdown of PU.1 in mice due to insertion of a neo cassette in the distal enhancer element resulted in development of myeloblastic AML and lymphoma [91]. Loss of function mutations affecting its DBD have been found in some patients with monocytic AML [92].

As mentioned earlier, there are controversial reports about cross-regulation of PU.1 and GATA1 by mutual inhibition to intrinsically determine the lineage choices or whether they just execute the primary decision initially made by other factors [25–27,46,93]. In order to enable erythroid maturation happen, it is essential to block PU.1 function [89]. GATA1 directly binds to PU.1 with its C- finger and displaces the transcription factor c-jun. Moreover, it uses the N- finger to indirectly regulate the SPI-1 gene locus. Therefore removal of the N- and/or C- finger of GATA1 can both reduce this inhibitory function [68]. On the other hand, PU.1 can displace the activating complex members from GATA1 and replace it with repressing complexes including pRB, SUV39H and HP1 [27]. The interplay between the two TFs is complex and on top, they can also autoregulate themselves and interact with other complexes [32].

3.2.3 GATA2: another GATA TF regulating stem cells

GATA2 is another GATA TF family member that has been identified as an important player in regulation of hematopoiesis. Although its expression is not restricted to hematopoietic cells, it plays a major role in HSC maintenance, mast cell and megakaryocytic-erythroid development [94,95]. GATA2 knockout mice die during embryonic development due to severe anemia and impaired self-renewal capacity of HSC [95]. Interestingly, GATA2 seems to autoregulate its expression by binding to its own promoter. Since GATA1 also bears this consensus sequence in its promoter elements, both factors can replace each other, in a process called “GATA factor switching”. This GATA switch allows reciprocal expression changes and leads to downregulation of the factors if needed [96]. Of note, GATA2 can also interact with other transcription factors such as PU.1 [68]. Recently, mutations in GATA2 were reported in AML patients harboring also C/EBPα mutations [97]. Moreover, GATA2 mutations could de detected in patients suffering from CML [98,99].

There are numerous additional transcriptional regulators involved in hematopoietic development include the CCAAT/enhancer binding protein (C/EBPα), Krüppel- like-factor 1 (KLF1), Runt- related transcription factor 1 (RUNX1, also known as AML1), T-Cell Acute Lymphocytic Leukemia 1 (TAL1, also known as SCL), nuclear factor 1 A (NFIA), Friend of GATA1 (FOG1, also known as ZFPM1), nuclear factor erythroid 2 (NF-E2), CBFA2/RUNX1 Translocation Partner 3 (CBFA2T3, also known as ETO2), Growth Factor Independent 1B Transcription Repressor (1B) and many more [8,31,100]. Interestingly, most genes encoding for these TFs have been found altered in hematological malignancies, as partners in translocations or occur as somatic point mutations [9].

On the basis of the above-mentioned factors, we conclude that fine changes in the stoichiometry result in a deregulation of the dense transcriptional network that affects self-renewal of HSC as well as maturation and proliferation of more committed progenitor cells.

3.3 Epigenetic regulation of hematopoiesis

As mentioned before, it has been recognized that HSCs express cytokine receptors and transcription factors at low level. In order to ensure proper differentiation into lineages, the expression “programs” have to be silenced [19]. Hereby, the genetic code resides in a highly organized and compacted structure and is wrapped around nucleosomes, together called chromatin [101]. The nucleosomes are composed of total eight histone proteins, two of each core histone H2A, H2B, H3 and H4 forming an octamer. Moreover, there are linker histones H1 that bind DNA and provide stability [102]. During the differentiation process, the nucleus gets reorganized and chromatin needs to be made accessible for TFs and the transcriptional machinery at different loci and times. In general, differentiation is associated with condensation of chromatin. This status is called heterochromatin (*greek* hetero=different, chroma=color), in contrast to euchromatin (*greek* eu=real), since the DNA appears more stained when a DNA intercalating fluorescence marker is used [103]. The closure of the chromatin is associated with a decrease in active global transcription [104]. Chromatin needs to be modified in order to control that process. There is a variety of different player involved in this organization, some of which are explained in the following section. Interestingly, all chromatin- modifying enzymes have a role in HSC self- renewal or maintenance. Therefore it is very important to tightly control the organization in a dynamic way to ensure cellular processes such as proliferation, cell death, DNA repair, replication and various others [105]. An overview of some modifications that are discussed in more detail, are depicted in **Figure 5**.

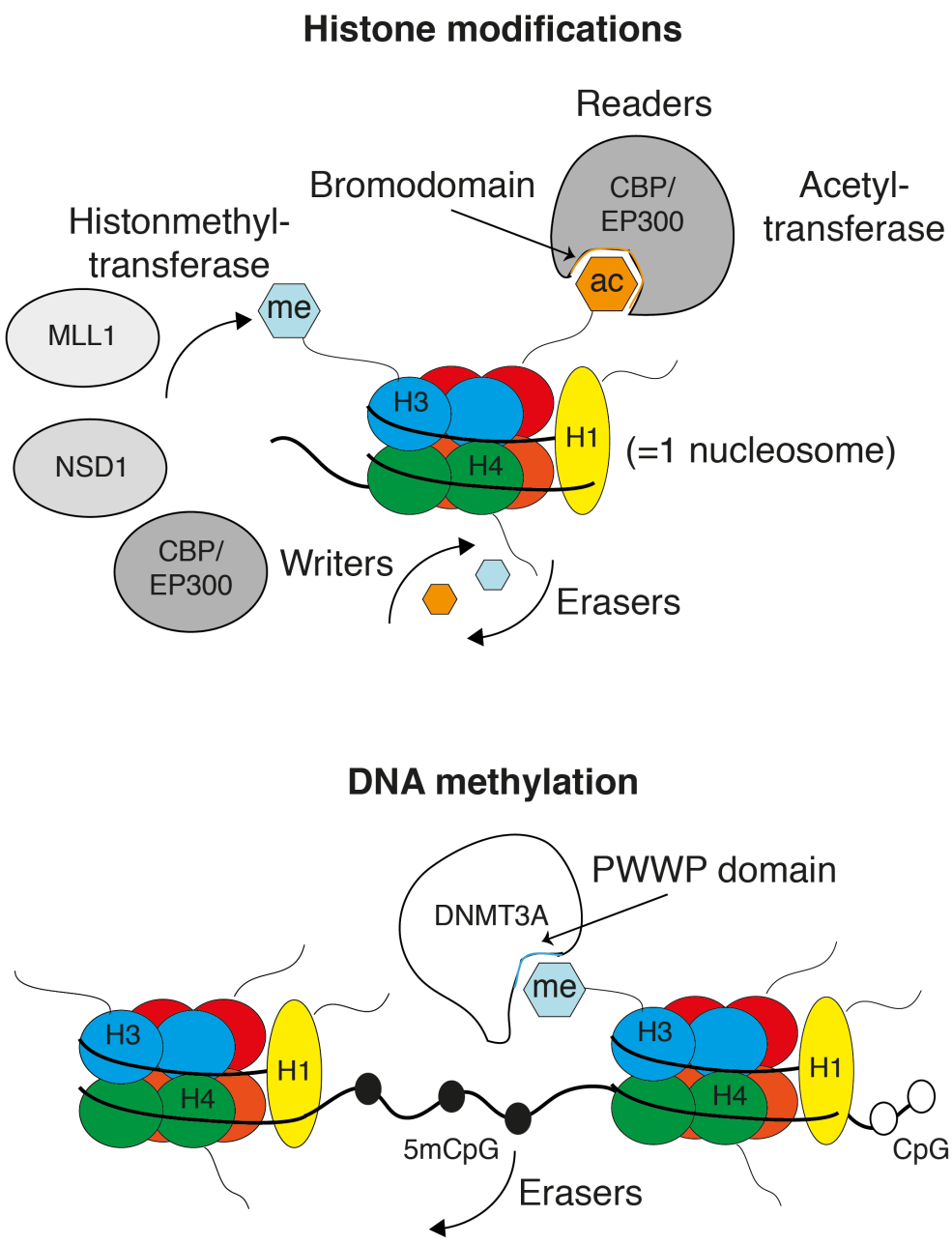


FIGURE 5. Epigenetic regulation of hematopoiesis.

Nucleosomes are composed of total eight histone proteins, two of each core histones H2A, H2B, H3 and H4 are forming an octamer. In addition, there is a linker histone H1 per nucleosome. Histone modifications alter the accessibility of transcription factor and associated machinery to the DNA. The reversible changes are made by so- called “writers”, “readers” and “erasers” of the so-called “histone code”. The figure depicts examples described in the introduction, such as the histone acetyltransferase CBP/EP300 and methyltransferases MLL1 and NSD1. CBP/EP300 can also recognize acetylated histones with their bromodomain. Another epigenetic modification discussed here is DNA methylation. In this process, methyl groups are attached to the nucleotide cytosine (=5mCpG) residing in so-called CpG islands, e.g. by DNMT3A. There are many connections between different epigenetic mechanisms, e.g. by binding of DNMT3A to methylated histones through the PWWP “reader” domain.

3.3.1 Histone modifications

During gene transcription, certain loci are opened or closed more than others depending on a specific code "written" on the histones. The so- called "histone code" refers to an old hypothesis stating that the genetic code is partly regulated through defined combinations of posttranslational modifications that occur at the tail amino acids of histones. Histone modifications alter the accessibility of transcription factors and associated machinery to the DNA. Reversible changes are made by so- called "writer, reader and erasers" [106]. "Writers" can acetylate, methylate, phosphorylate, ubiquitinate, sumoylate, ribosylate or biotinylate histone tails. In general, writers have a specific target residue that they modify [107]. In the following section, two main modifications that are involved in many hematopoietic processes are explained in great detail.

Histone acetylation occurs at lysine residues of H3 and H4. During acetylation the positive charge of the histone tail is neutralized and loosens the affinity of the negatively charged DNA to histones and therefore opens the locus to enhance transcription. Acetylation marks are set e.g. by histone acetyltransferases (HAT). There are three classes, first the Gcn5 N-acetyltransferases (GNAT) comprising its members Gcn5, p300/CBP-associated factor (PCAF), Elp3, Hat1, Hpa2. Second, there are Nut1, MOZ, MORF, Ybf2/Sas3, Sas2, HBO1 and Tip60 making the MYST family. And third the p300/CBP family including the cAMP-responsive element (CREB)- binding protein (CBP) or its homologue adenovirus E1A-associated 300-kD protein (p300) [108]. CBP/p300 function as transcriptional regulators by acetylating histone tails and other nuclear proteins. Moreover, it acts as a scaffold by recruiting transcription factors to active loci through a large diversity of protein interaction domains [109,110]. "Readers" bearing special reading motifs such as bromo- or chromodomain concomitantly recognize acetylated histones. CBP/p300 contain a bromodomain recognizing acetylated lysines finally resulting in a positive feedback loop and maintenance of CBP/p300 enzymatic activity. For this reason, CBP/p300 have two functions: writing and reading [111,112]. In mice, homozygous loss of either *Cbp* or *p300* leads to embryonic lethality in utero due to developmental defects including impaired hematopoiesis [113,114]. Studies in heterozygous mice have characterized Cbp as an essential regulator of HSC self-renewal. Likewise, conditional ablation of Cbp in adult mice altered differentiation, quiescence,

apoptosis, and self- renewal of adult HSCs [115]. CBP/p300 has been functionally linked to the development of multiple human cancers, including solid tumors and hematological malignancies [116]. CBP/p300 has been detected in several oncogenic fusions in leukemia involving either the MOZ acetyltransferase or the mixed linage leukemia (MLL) gene product promoting cell proliferation [117,118]. The most prevalent is the chromosomal translocation t(11;16)(q23;p13) associated with mostly therapy-related acute myeloid leukemia (AML) or myelodysplastic syndromes (MDS) resulting in a fusion protein containing the bromodomain of CBP and parts of MLL [118,119]. CBP/p300 was also proposed to act as transcriptional co-activator of other leukemogenic proteins such as the NUP98-HOXA9 fusion [120]. Recently, it has been shown that p300 interacts with the AML1- ETO fusion protein, present in over 20% of human AML regulating transcription of multiple AML1–ETO target genes that are drivers of self- renewal of hematopoietic stem/progenitor cells [121]. Furthermore, inhibition of p300 abrogates acetylation of AML1-ETO and impaired clonogenic growth *in vitro* and leukemic transformation *in vivo* [122].

Finally, these chromatin marks are reversible and can be removed by erasers such as histone deacetylases (HDAC), which are associated with decreased transcription. Interestingly, (de)acetylation takes not only place on histone protein, but also a large variety of non- histone proteins have been discovered to be modified [123].

Histone methylation occurs at the N-terminal tails of histones. To date, many different histone methyltransferases (HMT) have been identified creating different marks involved in opening and closing of the gene loci (**Table 1**). In general it has been recognized that the methylation mark is bivalent and can stand either for active or repressed transcription, depending where it is deposited. The best described HMT in hematopoiesis is mixed lineage leukemia 1 (MLL1, also known as KMT2A). It belongs to the trithorax group proteins with conserved functionality and more specifically to the subgroup containing a Su(var)3-9, Enhancer-of-zeste, Trithorax (SET) domain. This enzymatic unit is involved in Histone 3 Lysine 4 (H3K4) mono-/di-/tri- methylation at transcriptional start sites (TSS) of genes including HOX genes [124]. HOX genes are essential for embryonic development and encode transcription factors involved in cellular differentiation. MLL1 has been reported to act as an upstream regulator of HOXA9 and the cofactor MEIS1, which are normally expressed in early hematopoietic hierarchy but later downregulated [125].

MLL functions in a large multi-protein complex including WDR5, RBBP5, ASH2L associating at the C- terminus and multiple endocrine neoplasia type 1 (Menin), Lens-epithelial growth factor (LEDGF) and Myb at the N terminus [123,126,127].

Mll1 inactivation in mice results in embryonic lethality due to altered homeobox (HOX) gene expression. Moreover, MLL1 has been shown to be involved in stem cell renewal and differentiation programming. In human leukemia, MLL located on chromosome 11q23 has been found in more than 70 different translocations leading to expression of fusion proteins. Many of these fusions have strong oncogenic activity as shown in cell cultures and in mouse models [128–130]. The majority (>75%) of the N- terminal MLL is fused to transcriptional activators such as ENL, AF9 and AF4. The main task of these transcriptional activators is to recruit other chromatin modifying enzymes. Importantly, all MLL- fusions miss the C terminus and the protein- protein interactions of the N- terminus have been shown to be crucial for normal as well as leukemogenic activity of MLL [131–134].

To conclude, oncogenic fusion proteins often combine a DNA binding part recruiting the complex to the vicinity of transcriptional starts sites and continuous recruitment of the basal transcriptional machinery finally resulting in aberrant transcription [126]. Loss or gain of conserved catalytic subunits of an enzyme, leads to an altered protein function. Concomitantly, protein- protein interactions are destroyed and proteins might be attracted to the “wrong” loci. The chromatin structure cannot be maintained anymore and results in aberrant transcription. If genes are not controlled in a spatial and temporal manner, malignant hematopoietic disorders are the result of faulty target gene expression [105].

On the basis of these two well- studied examples, it can be concluded that active transcription in hematopoietic cells is marked by acetylation and methylation of histones nearby transcription-regulatory elements such as enhancers, promoter or as well gene bodies in a specific “code” combination. Another epigenetic mechanism contributing to the regulation of transcription is DNA methylation.

TABLE 1. Histone lysine methylations, respective enzymes that carry out the reaction and proposed function (+ activates transcription, - diminishes transcription) adapted from [135]. * no independent reproduction of result yet

Histone (lysine)	Modification (methylation)	Enzyme	Function
H1K26	me1	EZH2	
H1K168	me1	NSD1*	
H3K4	me1	MLL1/2/3/4, SETD1A/B, SETD7, ASH1L*	+
	me2	MLL1/2/3/4, SETD1A/B, SMYD3*	+
	me3	MLL1/2/3/4, SETD1A/B, SETD7, ASH1L, PRMD9, SMYD3*	+
H3K9	me1	G9A, SETDB1, EHMT1, PRDM2*	+
	me2	SUV39H1, SUV39H2, G9A, SETDB1, PRDM2*, EHMT1	-
	me3	SUV39H1, SUV39H2, SETDB1, PRDM2*	-
H3K27	me2	EZH1/2	-
	me3	EZH1/2	-
H3K36	me1	SETD2, NSD1/2/3	
	me2	SETD2, NSD1/2/3, SMYD2*	
	me3	SETD2	+
H3K79	me1	DOT1L	+
	me2	DOT1L	+
	me3	DOT1L	+
H4K20	me1	SETD8	
	me2	SUV420H1/2	
	me3	NSD1, SUV420H1/2	

During DNA methylation, methyl groups are attached to the nucleotide cytosine. It almost exclusively occurs in so-called CpG islands, which often are in close proximity of TSSs. In germ line and during pre- implantation of the blastocyst, there is active demethylation of both parental genomes taking place [136]. Upon differentiation, pluripotency- as well as germline- associated genes need to be silenced achieved by hypermethylation of their TSSs. This methylation is carried out by *de novo* methylating enzymes, the DNA methyl transferases (DNMT) 3A and 3B. Moreover, there is another DNA methyltransferase, DNMT1, dedicated to maintain the methylation. In total 60-80% of CpG islands are methylated in the mammalian genome. Methylation of a TSS is associated with repression of gene transcription, preferentially of housekeeping and developmentally important genes [137]. Deletion of DNMT1 and DNMT3B in mice resulted in embryonic lethality, whereas deletion of DNMT3A resulted in production of viable pups [138]. However, these mice did only survive up to four weeks after birth. DNMT3A was identified as a regulator of embryonic development, imprinting and X chromosome inactivation [139]. In hematopoiesis, conditional deletion of DNMT3A leads to expansion of LT-HSCs and changes myeloid differentiation [123,140,141]. In human hematopoiesis, aberrant promoter CpG island methylation of DNA is one of the major causes identified in leukemia. The third most common mutation in AML occurs in the gene of DNMT3A and is associated with poor prognosis [142–146]. The mutations result in a decrease of the catalytic activity and DNA binding properties. The mutated protein inhibits the wild type protein since they normally cooperate in an oligomer to exert their function. Therefore, a haploinsufficiency seems already sufficient to cause a phenotype. However, the consequences are not fully understood [142,147]. Global methylation levels have been reported unchanged, with exception of specific loci, When DNA methylation was decreased, it did not translate into changes in gene expression [144]. Interestingly, HOX-A and HOX-B clusters showed increased expression upon DNMT3A mutations [146]. Moreover, DNMT3A has been reported to be present in MDS patients before acquisition of the driver mutation as well as in healthy subjects [148,149].

There also other proteins involved in regulating DNA methylation being mutated in AML such as DNMT3B, DNMT1, IDH1, IDH2, TET1 and TET2 [142]. Both, histone modifications as well as DNA methylation play an important role in the spatial and temporal control of gene expression during development. Moreover, other proteins such as chromatin remodelers, histone readers, co-activators and kinases

are involved in the orchestration of opening and closing of chromatin at the right time, that are not discussed in detail here. Of note, these pathways do not interact independently e.g. DNMT3A bears a Pro-Trp-Trp-Pro (PWWP) motif binding to H3K36me3 in gene bodies and in turn methylates CpG dinucleotides. The combination of histone and DNA methylation was proposed to influence the splicing of introns [150,151]. It has also been suggested that CpG islands attract MLL fusion proteins and the attached complex containing DOT1L catalyzing methylation of H3K79, a mark associated with active transcription [152–154]. Moreover there has been a direct link between mutated DNMT3A and aberrant HOX gene expression reported [155].

In summary, transcriptional as well as chromatin regulation are essential to maintain the stem cell pool and produce sufficient differentiated cells for functional hematopoiesis (**Figure 6**).

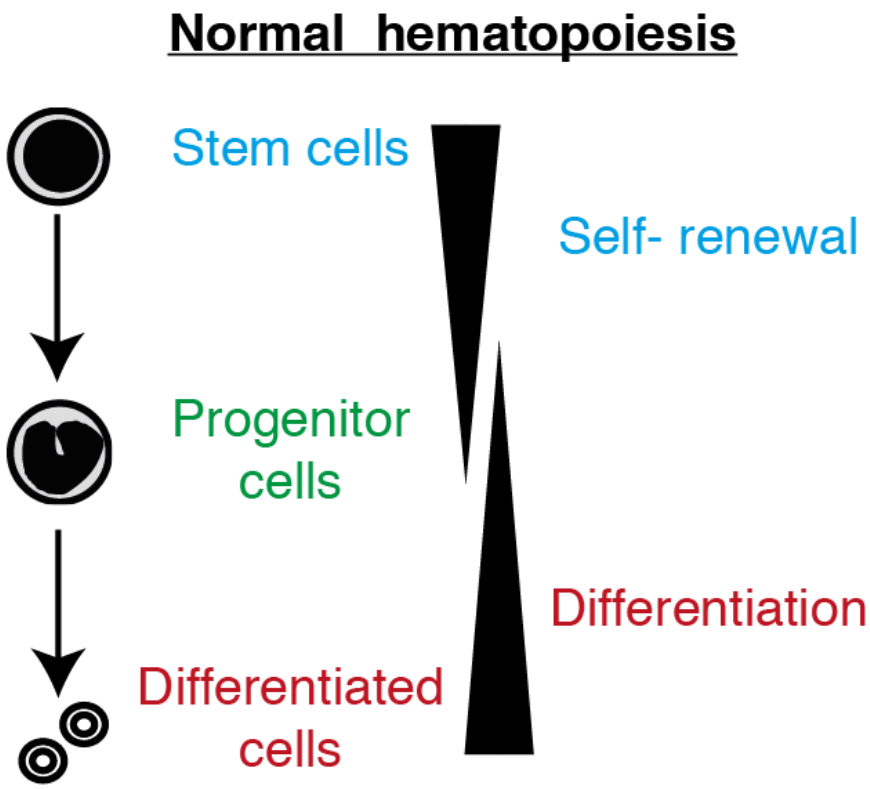


FIGURE 6. Normal hematopoiesis.

Schematic depiction of normal hematopoiesis. Stem cells need to self- renew their pool and at the same time produce progenitors giving rise to differentiated effector cells.

3.4 Acute myeloid leukemia

An aberrant transcriptional regulation caused by mutations or translocations in hematopoietic cells may result in malignant transformation and induction of leukemia. Acute leukemia is classified as either acute lymphoblastic leukemia (ALL) or acute myeloid leukemia (AML). Within the AML classification, there are eight subgroups (M0-7) according to the French- American- British (FAB) system that distinguishes cytomorphological features. Interestingly, acute leukemia is the product of only a limited number of functionally cooperating genetic alterations. A recent analysis of AML genomes using whole- genome as well as whole exome sequencing data suggested an average of thirteen gene mutations per AML genome of which only five were recurrently mutated [142,156]. The prevalence of somatic mutations compared to other cancers was lower than the average [157]. The current concept of leukemia development involves a differentiation block of a lineage through mutations of transcriptional or chromatin regulators combined with mutations affecting cellular signaling pathways conferring growth advantage finally resulting in aberrant differentiation and increased self-renewal capacity of hematopoietic stem and progenitor cells (**Figure 7**) [142,158,159].

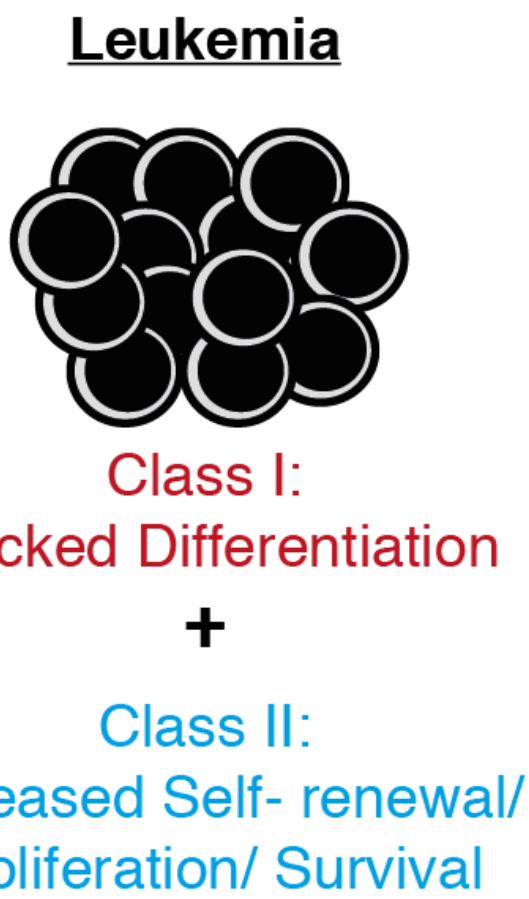


FIGURE 7.

Malignant hematopoiesis. The current concept of leukemia involves so-called class I mutations causing a block in differentiation in combination with class II mutations increasing self-renewal, proliferation and/or survival.

3.4.1 NUP98-NSD1: a molecular hallmark of pediatric AML

Approximately 20% of pediatric AML cases are cytogenetically normal (CN-AML) (**Figure 8**). However, such cases can also harbor chromosomal translocations that often contain telomeric breakpoints not detected by traditional karyotyping. One example is t(5;11)(p35;p15) leading to expression of the NUP98-NSD1 fusion [160,161]. Initially, *NUP98-NSD1* was identified by cloning of a chromosomal translocation t(5;11)(q35;p15) from a pediatric AML patient and revealed expression of a fusion between the genes coding for the nucleoporin 98 (*NUP98*) and the nuclear receptor binding Su(var)3-9, enhancer of Zeste and Trithorax domain protein 1 (*NSD1*) [162]. NUP98-NSD1 is mostly found in AML of children and only rarely found in adult patients and generally associated with a poor prognosis [163–165]. Hollink *et al.* reported that NUP98-NSD1 positive AML are characterized by aberrant expression of the *HOX A-B-C* gene clusters distinguishing them from AML with MLL- rearrangements [163].

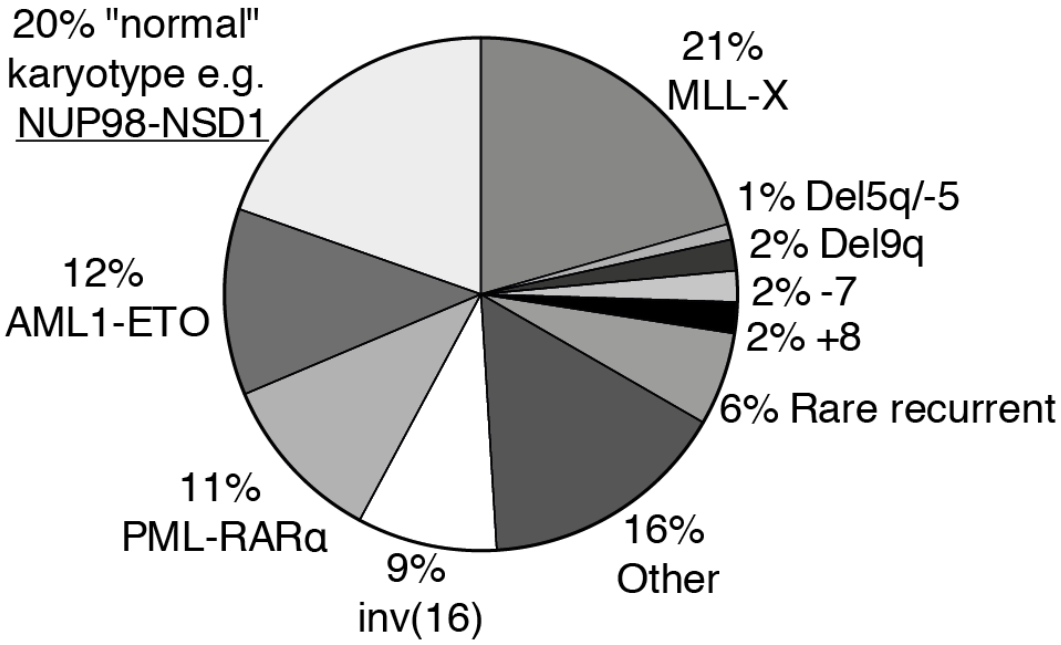


FIGURE 8. NUP98-NSD1 is involved in pediatric AML.

Karyotypic alterations in childhood acute myeloid leukemia (AML). A significant fraction of the patients harbours MLL translocations (21%) or appears as cytogenetically normal (CN) AML including Nucleoporin 98 (NUP98)- Nuclear Receptor-Binding SET Domain-Containing Protein 1 (NSD1) translocations (adapted from [160]) that are not detected by traditional karyotyping.

The *NUP98* gene is located on chromosome 11p15. The NUP98 protein results from cleavage of an 1800aa precursor that is cleaved into 96kDa (NUP96) and 98kDa sized proteins, which form parts of the nuclear pore complex involved in shuffling macromolecules between the cytoplasm and nucleus. In order to exert their function, one subfamily of nucleoporins contain phenylalanine- glycine (FG) repeat domains binding to nuclear receptors and forming a nuclear barrier to decide which macromolecules may pass the pore [166,167]. NUP98 has been also found in other translocations and in all of these fusions, the translocation product still harbors the Phenylalanine-Glycine (FG) repeat of NUP98 [168]. These repeats are a hallmark of nucleoporin proteins and have been shown to interact with CBP/p300 (**Figure 9**) [120].

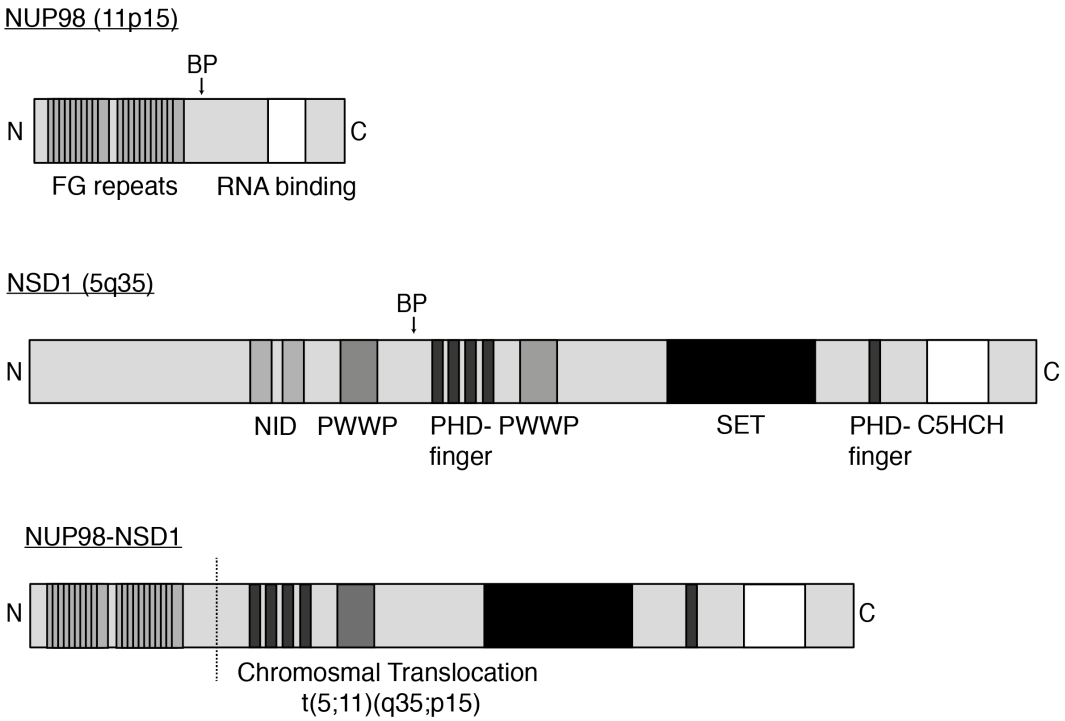


FIGURE 9. Protein structure of NUP98 and NSD1 and NUP98-NSD1 fusion.

Schematic depiction of *NUP98* located in chromosome 11p15 and *NSD1* located on chromosome 5q35. NUP98 harbors FG repeats and RNA binding domain. NSD1 harbors nuclear receptor interacting (NID), PWWP, PHD, SET and C5HCH domain. Upon chromosomal translocation t(5;11)(q35;p15), NUP98 looses its RNA binding domain and NSD1 the NID and first PWWP domain.

Nuclear receptor binding domain protein 1 (*NSD1*) consists of 2696aa and its gene is located on the long arm of chromosome 5. The murine *Nsd1* (2691aa, Chr. 13 B1) was initially identified in a two- hybrid screen investigating potential binding

partners of retinoic acid receptor alpha (RAR). The authors identified Nsd1 as a novel protein interacting with several nuclear receptors such as the retinoid X receptor gamma (RXR), the retinoid acid receptor alpha (RAR), the thyroid hormone receptor (TR) and the androgen receptor (AR). For its binding activity, two distinct nuclear receptor- interacting binding domains (NID) were described. NID^{-L} binds unliganded RAR and TR, whereas NID^{+L} binds ligand-bound TR, RAR, RXR and ER [169]. Moreover, NSD1 is also known as AR-associated (ARA) protein 267, based on observations that it interacts and stimulates AR transactivation [170]. In addition to the N-terminal located NIDs, NSD1 contains two Pro-Trp-Trp-Pro (PWWP) motifs, five plant homeodomain (PHD) fingers and a C-terminal Su(var)3-9, Enhancer-of-zeste, Trithorax (SET) domain with methyltransferase activity (**Figure 9**) [169,171]. Huang et al. reported NSD1 to be present in the nucleus tightly associating with chromatin, excluded from nucleoli and condensed heterochromatin. Deletion of the FG domain of NUP98 as well as mutations in NSD1 inactivating the methyltransferase activity or preventing binding of NUP98–NSD1 to the HOX-A locus decreased HOX-A gene transcription and concomitant immortalization of myeloid progenitor cells (**Figure 10**) [171].

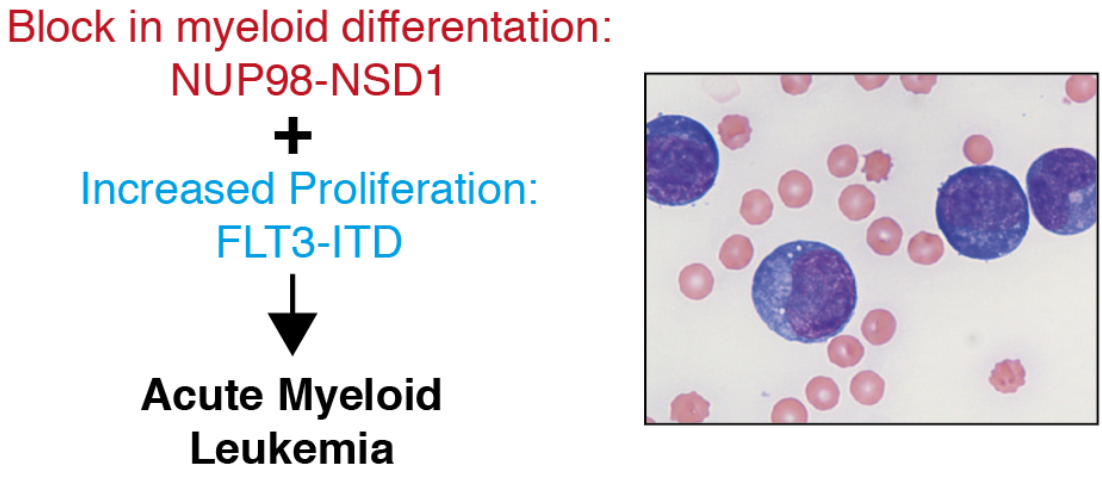


FIGURE 10. NUP98-NSD1 in AML.

NUP98-NSD1 causes blocked differentiation. In combination with FMS-like tyrosine kinase-3 internal tandem duplication (FLT3-ITD), which confers increased proliferation, acute myeloid leukemia develops and so called „blasts“ arise in the peripheral blood. Wright- Giemsa stained blood smear of leukemic mouse (courtesy of A. Thanasopoulou, [172]).

3.4.2 Nuclear receptor binding domain protein 1 (NSD1)

Losson *et al.* generated an Nsd1 knockout model by replacing the largest exon 5 by a floxed *pgk-neomycin* antibiotic selection cassette. Nsd1 knockout mice were embryonically lethal at E10.5. At earlier developmental stages, the authors observed degenerated extra-embryonic tissue lacking any growth structure. At E6.5, knockout embryos were smaller in size and pyknotic nuclei in the ectoderm were detected. Moreover, presence of TUNEL positive cells confirmed occurrence of apoptosis and positional markers of the *anterior-posterior* axis were not expressed. Besides this, mesodermal markers as well as homeobox (HOX) genes were not expressed. The group concluded, that the *Nsd1* gene is required for post- implantation development in mice and mutant embryos did not develop the mesodermal layer. Heterozygous Nsd1 knockout mice were viable and fertile. Furthermore they also suggested that Nsd1 may play a role at later developmental stages as differential expression was observed in multiple regions during embryonic development [173].

Alignment using the protein BLAST program of murine Nsd1 with the human NSD1 revealed 84% identity with domain conservation. A direct comparison of domains revealed 99% homology of the SET domain and 87% of the NID domain between human (NP_071900) and mouse (NP_032765) protein sequences (**Figure 11**) [174].

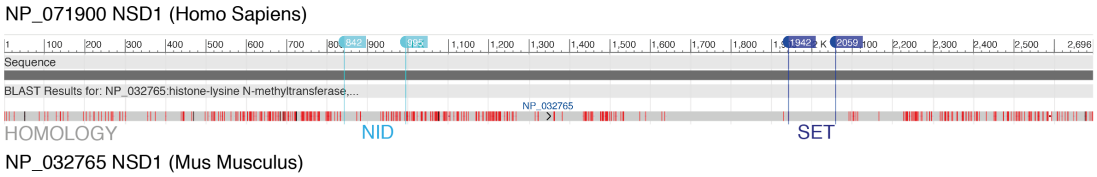


FIGURE 11. Homology of human and murine *NSD1*.

Protein blast homology comparison of human and mouse NSD1. NID and SET domain are highlighted. Red stripes indicate difference in protein sequences.

Upon fusion to NUP98, the NID of NSD1 is lost, while the PHD fingers and SET domains are maintained. Expression of the NUP98-NSD1 fusion has been previously proposed to have oncogenic activity dependent on the integrity of SET-domain-mediated methyltransferase activity resulting in increased H3K36 methylation and expression of the *Hox-A* gene cluster [171,175]. Co-expression of NUP98-NSD1 and FLT3-ITD in BM (BM) cells revealed potent cooperation for the induction of AML in mice [172]. In addition to chromosomal translocations, *NSD1* is the target of genomic

amplifications as well as recurrent point mutations in various solid cancers [162,171,176–182] (**Table 2**).

It has been reported that the major cellular substrates of NSD1 are mono and di-methylation of H3K36 as well as mono methylation of lysine 168 of linker histone 1.5 (H1.5K168) [183,184]. Moreover there are mixed reports regarding methylation of H4K20 as well as non- histone proteins such as nuclear factor kappa B (NF-κB) [173,179,180,184–186]. Interestingly, NSD1 controls methylation of histones at the bone morphogenic protein 4 (BMP4) gene locus in human HCT116 colorectal cancer cells and knockdown of NSD1 in these cells reduced mono-, di- and tri- methylation of H3K36 within the gene body as well as gene expression levels [187]. Whether this regulation holds true in hematopoiesis has not been reported so far. Nevertheless, this data suggests a potential role of Nsd1-mediated mono- and dimethylation marks as a substrate for SET Domain Containing 2 (SETD2, also known as HYPB) that trimethylates H3K36 [183].

In addition, heterozygous *NSD1* point mutations have been proposed to be the molecular correlate for SOTOS, a multiple anomaly syndrome characterized by overgrowth, distinctive craniofacial appearance, and learning disabilities with an increased risk to develop cancer [188,189]. The *NSD1* gene locus was also found to be a target of aberrant CpG island promoter hypermethylation in blood samples of SOTOS patients [179]. A complete list of percentages of pathogenic point mutations of sequenced cancer samples by Sanger Cosmic database can be found in **Table 2**. In hematopoietic and lymphoid tissue, approximately 25% of the mutated cases bore a point mutation that occurred either in the NID, PHD finger or SET domain (**Table 3**). Mutations of PHD domains are believed to reduce recognition of methylated histones whereas mutations of SET domain might interfere with the methyltransferase activity [190,191]. To date, there is no proposal of functional consequences of NID mutations. Collectively, these observations suggested that in addition to its role being involved in the *NUP98-NSD1* fusion oncogene, *NSD1* might also act as a tumor suppressor. As already proposed by Huang *et al*. it contains activation or repression domains and might therefore act in a context- dependent manner [169].

TABLE 2. NSD1 mutation and copy number variant (CNV) distribution in cancer samples according to COSMIC database.

Tissue	Point Mutations (% mutated)	Copy Number Variation (% variant)
Adrenal gland	-	2.78 (gain)
Breast	0.58	0.1 (loss), 0.3 (gain)
Central nervous system	0.14	0.47 (gain)
Cervix	0.31	-
Endometrium	3.75	-
Haematopoietic and lymphoid	0.24	-
Kidney	0.95	0.24 (loss), 0.96 (gain)
Large intestine	2.29	0.14 (loss)
Liver	0.25	0.16 (gain)
Lung	1.34	0.45 (loss), 0.45 (gain)
Oesophagus	0.37	-
Ovary	0.24	0.28 (gain)
Pancreas	0.07	-
Prostate	0.39	-
Skin	1.09	0.41 (gain)
Soft tissue	-	1.41 (gain)
Stomach	2.53	0.28 (loss)
Thyroid	0.18	-
Upper aerodigestive tract	1	0.64 (loss)
Urinary tract	1.05	-

TABLE 3. Pathogenic NSD1 mutations in hematopoietic and lymphoid cancer samples according to COSMIC database.

Mutation (CDS)	Mutation (aa)	Count	Mutation type	Exon	Domain
c.1457C>G	p.S486C	1	Substitution - Missense	5	NID
c.2170G>A	p.E724K	1	Substitution - Missense	5	NID
c.4552G>T	p.G1518C	1	Substitution - Missense	11	PHD
c.5912A>G	p.Y1971C	2	Substitution - Missense	19	SET
c.6085A>G	p.T2029A	2	Substitution - Missense	20	SET

3.4.3 The NSD protein family

The NSD family contains two more members, Nsd2 and 3. Interestingly, both proteins differ from NSD1 by not containing the nuclear receptor interacting domain. NSD2 (also called Wolf-Hirschhorn Syndrome Candidate 1= WHSC1) encodes a protein with a PWWP domain, a high mobility group (HMG) box, a SET domain, and PHD zinc fingers [192]. It is expressed ubiquitously in early development and murine knockout is embryonically lethal. This data proposed a non- redundant function of NSD1 and NSD2. Its second name derives from its involvement in the Wolf-Hirschhorn syndrome (WHS), which is a malformation syndrome, associated with a hemizygous deletion of the distal short arm of chromosome 4. Moreover NSD2 is also involved in multiple myeloma formation through chromosomal translocation t(4;14)(p16.3;q32.3) resulting in increased expression of NSD2 under the control of the immunoglobuline (Ig) enhancer. Continuous expression of NSD2 results in increased H3K36 methylation and proposed oncogene function [192–197]. NSD3 (also known as multiple myelomas =WHSCL1) encodes a protein with a PWWP domain, a SET domain, and PHD zinc fingers. It has been also found to be fused to NUP98 in rare AML cases and found amplified in cases breast cancer [198,199]. There is to date no knockout mouse model that would determine its role in development. Taken together, NSD family members play an important role in human cancer.

4 INVESTIGATING THE ROLE OF THE EPIGENETIC REGULATOR NSD1 IN NORMAL HEMATOPOIESIS

4.1 WORKING HYPOTHESIS

Previous work has suggested a bi-potent role of Nsd1 in human cancer. Our lab and others have shown that when fused to NUP98, NSD1 has leukemogenic activity *in vitro* and *in vivo* [163,171,172]. Important to note is that a chromosomal translocation inherently leads to the loss of one functional allele, suggesting that reduced levels of NSD1 may contribute to the leukemic phenotype. In addition, recurrent presence of putatively loss- of- function mutations of NSD1 in human cancers including AML suggests that NSD1 could also act as tumor suppressor. This idea is also supported by the observation of aberrant DNA methylation of the NSD1 locus in some brain cancers [179]. These observations led us to propose that NSD1 may play an important role in hematopoiesis and reduction of NSD1 could increase the risk to develop a malignant disease.

To experimentally address this hypothesis we established mice in which we constitutively and conditionally inactivated the *Nsd1* gene in the hematopoietic system. The aim of this thesis is to describe the phenotype in Nsd1 knockout mice and to address the underlying molecular mechanisms.

4.2 MATERIAL AND METHODS

4.2.1 Transgenic mice

Mice with a *Nsd1*$^{+/L3}$ allele were obtained from Regine Losson (IGBMC, Strassbourg) [173]. The floxed *pgk-neomycin* selection cassette was removed, leaving two *loxP* sites flanking the largest coding exon 5, here referred as *Nsd1*$^{fl/fl}$. *Nsd1*$^{fl/fl}$ mice were intercrossed with a *Vav1-iCre*$^{tg/+}$ transgenic strain leading to inactivation of the gene in fetal and adult hematopoiesis. We also crossed *Nsd1*$^{fl/fl}$ mice with *Mx1-iCre* and *Scl-Cre-ER*^{T2} (called also *Scl-iCre* in this report) mice for conditional ablation. Both strains were obtained from Radek Skoda (DBM, Basel). For *Mx1-iCre*-mediated deletion, polyinosinic:polycytidylic acid (poly(I:C)) (Cat. P1530, Sigma Aldrich, Buchs, Switzerland) was administered on ten consecutive days with 2 days of break via intraperitoneal (i.p.) injections at 300μg/ml starting 6-8 weeks after birth. For *Scl-Cre-ER*^{T2} mediated deletion, mice were fed with tamoxifen (Tx) impregnated food supplemented with 10% sucrose (1 mg/g; Harlan Laboratories, Venray, The Netherlands) for 8-14 weeks. All mice in this study were kept under specific pathogen-free conditions at the animal facility of the DBM (ZLF, Hebelstrasse 20) with free access to food and water in accordance to Swiss Federal Regulations. All transgenic mice established on a mixed background were backcrossed to C57BL/6J (Ly5.2) for more than 10 generations.

4.2.2 Genotyping

Mice were genotyped using the KAPA Mouse Genotyping Kit HotStart (Cat. KK7352, KapaBiosystems, Wilmington, US) for extraction of genomic DNA and PCR reaction. Briefly, tissue was digested in 100μl extraction master mix (88μl H_2O, 10μl buffer, 2μl enzyme), mixed and heated for 10min at 75°C, followed by inactivation of enzymatic reaction at 95°C for 5min. Mix was vortexed and spun at 14.000rpm for 1min in a Hereaus table top centrifuge. 1μl of reaction was taken for PCR reaction with 12.5μl Kapa Mix, 0.25μl $MgCl_2$ (of 50mM stock), 2.5μl primers (according to **Table 4** at final concentration 10μM) and H_2O. PCR reaction program was: 5min 95°C, 40x cycles

95°C 15sec, 60°C 15sec, 72°C 30sec followed by 5min 72°C and 1min 4°C. PCR-amplicons were visualized on 2% agarose gels containing ethidiumbromide.

TABLE 4. Primer sequences for genoytping PCR

Target	Sequence forward (5'-3')	Sequence reverse (5'-3')
Vav1-iCre	CTCTGACAGATGCCAGGACA	TGATTTCAGGGATGGACACA
Nsd1	GTCTGCATTAAGTAATTGTGCCCTGAAG	ACTGACTCCTCTTCTGGAGATCTGAGTTC
Mx1-iCre	AGGTGTAGAGAAGGCACTTAGC	CTAATCGCCATCTTCCAGCAGG
Scl-Cre-ER^{T2}	TAGTGGGTTCTTTGGGGAAC	GTGAAACAGCATTGCTGTCACTT

4.2.3 *In vivo* experiments

Asymptomatic *Vav1-iCre$^{+/tg}$;Nsd1$^{fl/fl}$* mice were sacrificed at 4- 6 weeks after birth. *Vav1-iCre$^{+/tg}$;Nsd1$^{fl/fl}$* mice became symptomatic 7- 17 weeks after birth: once they reached an “animal welfare score” ≥ 3, based from appearance, natural and provoked behaviour and body weight the mice were sacrificed by CO_2 asphyxia and cervical dislocation. Organs were fixed in 4% formalin solution to preserve for histopathological analysis (Pathology, University Hospital Basel).

4.2.4 Histology & Mircoscopy

Formalin-fixed paraffin-embedded tissue sections were stained with hematoxylin and eosin (H&E). Differential blood counts were analyzed on smears stained using Wright- Giemsa staining (Hematology, University Hospital Basel). Tissue sections and blood smears were evaluated by Prof. J. Schwaller (trained in pathology) with advice of Prof. Dr. A. Tzankov, Institute for Pathology, Basel). Sections were analyzed on an Olympus BX61 microscope (Tokyo, Japan) or Nikon TI (Tokyo, Japan).

4.2.5 Peripheral blood analysis

Blood was collected from the tail vein or by vena cava inferior puncture (terminal) and counts were determined using an Advia120 Hematology Analyzer using Multispecies Version 5.9.0-MS software (Bayer, Leverkusen, Germany).

4.2.6 Timed mating & fetal liver extraction

Nsd1$^{fl/fl}$ and *Vav1-iCre*$^{+/tg}$*;Nsd1*$^{fl/+}$ mice were mated for one night to obtain 1 out of 8 born pups "positive" (=*Vav1-iCre*$^{+/tg}$*;Nsd1*$^{fl/fl}$) for analysis. Plugs were controlled and positive mice marked with E0.5. Pregnant mice were sacrificed at E13.5, 16.5 and 19.5 (day of birth) by CO_2 asphyxia and cervical dislocation. Pups were dissected and kept in PBS (pH 7.2, Cat. 20012068, Gibco, Thermo Fisher Scientific, Reinach, Switzerland). Decaptured whole bodies were fixed in 4% formalin solution to preserve for pathological analysis (Pathology, University Hospital Basel). Fetal livers were dissected with a bended forceps and scissors and kept in PBS. Single cells suspensions were obtained by pressing the liver gently with a pre-moistured plunger through a 40µm cell strainer (Cat. 352340, BD, New Jersey, USA) in RPMI (Cat. 61870, Gibco, Thermo Fisher Scientific, Reinach, Switzerland) containing 10%FCS (Cat. 2-01F10-I, Amimed, Bioconcept, Allschwil, Switzerland) and 1%Penicillin/Streptomycin (P/S) (Cat. 15140, Gibco, Thermo Fisher Scientific, Reinach, Switzerland).

4.2.7 Cell isolation

Total BM was harvested by crushing long bones and spine in RPMI (Cat. 61870, Gibco, Thermo Fisher Scientific, Reinach, Switzerland) containing 10% FCS (Cat. 2-01F10-I, Amimed, Bioconcept, Allschwil, Switzerland) and 1% P/S (Cat. 15140, Gibco, Thermo Fisher Scientific, Reinach, Switzerland) and then filtered through 40µm cell strainer (Cat. 352340, BD, New Jersey, USA). Spleens were dissected

and single cell suspensions obtained by pressing the spleen gently with a pre-moistured plunger through a 70µm cell strainer (Cat. 352350, BD, New Jersey, USA). When needed, red blood cells were lysed with ammonium- chloride potassium (ACK) lysis buffer (150 mM NH_4Cl, 10 mM $KHCO_3$, and 0.1 mM EDTA, pH 8.0) for 10min on ice. After 5min at 1500rpm centrifugation, cells were resuspended in PBS plus 10%FCS and counted. Lineage depletion was achieved according to manufacturers protocol of mouse hematopoietic lineage depletion kit (Cat. 130-090-858, Miltenyi Biotech, Bergisch Gladbach, Germany). Briefly, lineage positive cells were stained with biotin and streptavidin- conjugated beads used for magnetic cell separation using LS columns (Cat. 130-042-401, Miltenyi Biotech, Bergisch Gladbach, Germany). Solutions and antibody concentrations were adjusted to total number of red blood cell depleted cells.

4.2.8 BM transplantation

Transplantations were performed using whole BM or sorted cells of *Nsd1$^{fl/fl}$;Vav-iCre$^{tg/+}$* mice at indicated ages. For competitive transplantation, 1x10^6 total BM cells of symptomatic or asymptomatic *Nsd1$^{fl/fl}$;Vav-iCre$^{tg/+}$* mice (CD45.2) was mixed in a 0:1, 1:0, 1:1 or 1:10 ratio with supporting BM of B6.SJL (CD45.1) donor mice and transplanted into lethally irradiated (2x 600cGy) B6.SJL (CD45.1) recipients via tail vein. For competitive transplantation of distinct cell populations, cells were fluorescence activated cell sorted (FACS) from BM of asymptomatic or diseased *Vav1-iCre;Nsd1$^{fl/fl}$* mice into pre-cooled 2ml Eppendorf tubes filled with RPMI. Cells were spun down for 8min at 2500rpm in a Hereaus table-top centrifuge, counted and dilutions were made in PBS. CD45.1$^+$ wildtype cells were mixed with CD45.2$^+$ sorted cells, 10^6 cells were used as support in LSK/MEP transplantation and 0.5x10^6 cells in Lin$^-$/c-Kit$^+$/CD71$^{low/mid/high}$ transplantation. Cellular chimerism was determined by flow cytometry as described below in peripheral blood at indicated weeks after transplant. Additionally, for each bleeding, whole blood cell counts were measured on a blood analyzer.

4.2.9 Cytospots

Cytospots of approximately 10^5 cells were made by centrifugation for three minutes at 300rpm using a Shandon Cytospin 3 centrifuge using cytofunnel disposable sample chambers (Cat. 5991040, Thermo Fisher Scientific, Reinach, Switzerland) and one-circle, non-coated cytoslides (Cat. 5991051, Thermo Fisher Scientific, Reinach, Switzerland). Cytospots of the BM were stained with Wright-Giemsa in the Hematology Laboratory of the University Hospital Basel.

4.2.10 Flow cytometry

Cells in suspension were washed with FACS buffer (0.5% BSA, 1mM EDTA in PBS) and incubated with indicated antibodies for 45min on ice, washed and stained with 1ug/ml DAPI (Cat. D1306, Life Technologies, Paisley, UK) in PBS. Stained cells were analyzed on CyAn ADP analyzer (Beckman- Coulter) or LSR Fortessa (BD, New Jersey, USA). Data was analyzed with FlowJo software (Tree Star). For CD71/TER119 staining, the preparation still contained red blood cells, for CD71/c-Kit/Sca-1/ FcγRII/III stem and progenitor staining the red blood cells were depleted. For stem and myeloid progenitor staining, lineage positive cells were depleted as described before. All antibodies were used as indicated in **Table 5**.

For calculating number of stem and progenitor cells in BM, lineage-marker depleted cells were counted and absolute numbers of cells adjusted to this number. For example: on average $9.336x10^6$ Lin^- cells were obtained of control $Nsd1^{fl/fl}$ mouse BM and 178 LT-HSCs ($Singlets/DAPI^-/Lin^-/Sca\text{-}1^+/c\text{-}Kit^+/CD34^-/CD150^+/CD48^-$) were counted by flow cytometry where 1'001'239 FSC/SSC preselected events were acquired = =((1000000/1001239)*9.336)*178= 1660 LT-HSC. For flow cytometric analysis of apoptosis, ratios of early and late apoptotic cells were determined with the Annexin V-APC apoptosis detection kit (Cat. 559763, BD, New Jersey, USA). Briefly, cells were collected and washed twice with cold PBS buffer, resuspended in 100μL of 1x binding buffer, incubated with 5μL of Annexin V conjugated to APC and 5μL 7-AAD for 15 min at room temperature, and analyzed by flow cytometry. For differentiation analysis of mouse or human cells *in vitro*, cells were filtered, washed

twice with PBS and stained in 100µL FACS buffer. For cells sorting, the same antibodies were used than for flow cytometry analysis as indicated in the table.

TABLE 5. Staining protocols including antibodies used. Antibodies were obtained from the following companies: eBioscience (Affymetrix, Thermo Fisher Scientific, Reinach, Switzerland), BD Pharmingen (Allschwill, Switzerland), BioLegend (LucernaChem, Luzern. Switzerland)

Staining	Material	Name	Fluorochrome	Concentration	Company	Cat.Nr.
CD71/Ter119	no lysis	CD71	FITC	1:100	eBioscience	11-0711-81
		Ter119	PE	1:100	BD Pharmingen	553673
CD71/c-Kit/Sca-1/FcγRII/III	RBC lysed	CD71	FITC	1:100	eBioscience	11-0711-81
		c-Kit	APC	1:25	eBioscience	17-1171-83
		Sca-1	Pe-Cy7	1:25	BioLegend	122513
		FcγRII/III	PE	1:50	eBioscience	12-0161-83
Stem cells	RBC lysed, lineage depleted	Streptavidin	Pacific Blue	1:200	Invitrogen	S11222
		c-Kit	APC	1:25	eBioscience	17-1171-83
		Sca-1	Pe-Cy7	1:25	BioLegend	122513
		CD34	FITC	1:50	eBioscience	13-0341-82
		CD150	PE	1:50	eBioscience	12-1502-80
		CD48	A700	1:50	BioLegend	103426
Myeloid progenitor I	RBC lysed, lineage depleted	Streptavidin	Pacific Blue	1:200	Invitrogen	S11222
		Il-7R	biotin	1:10	eBioscience	13-1271-85
		c-Kit	APC	1:25	eBioscience	17-1171-83
		Sca-1	Pe-Cy7	1:25	BioLegend	122513
		CD34	FITC	1:50	eBioscience	13-0341-82
		FcγRII/III	PE	1:50	eBioscience	12-0161-83
Myeloid progenitor II	RBC lysed, lineage depleted	Streptavidin	Pacific Blue	1:200	Invitrogen	S11222
		Sca-1	biotin	1:10	eBioscience	13-5981-82
		FcγRII/III	biotin	1:10	BioLegend	101303
		c-Kit	PE	1:50	eBioscience	12-117-82
		CD41	FITC	1:25	BD Pharmingen	553848
		CD105	Pe-Cy7	1:25	BioLegend	120409
		CD150	APC	1:25	eBioscience	17-1502-80
Chimerism	RBC lysed	CD45.1	APC	1:100	eBioscience	17-0453-81
		CD45.2	PercP-Cy5	1:100	BD Pharmingen	552950
Human CD71/GPA	no lysis	CD71	FITC	1:100	BioLegend	334103
		Glycophorin A	PE	1:100	BioLegend	349105

4.2.11 RNA sequencing of *in vivo* mouse samples

We isolated RNA from sorted LSK (n=3/group), MEP (n=3/group) and GMP (n=2/ *Nsd1*$^{fl/fl}$ and n=3/*Vav1;iCre;Nsd1*$^{fl/fl}$ from diseased *Vav1;iCre;Nsd1*$^{fl/fl}$ mice and age- and sex- matched control *Nsd1*$^{fl/fl}$ mice. In addition, LSK of asymptomatic, young *Vav1;iCre;Nsd1*$^{fl/fl}$ mice and age- and sex- matched control *Nsd1*$^{fl/fl}$ mice were sorted

(n=3/group). Cells were sorted into pre- cooled 2ml Eppendorf tubes containing RPMI. Cells were spin down for 8min at 2500rpm and resuspended in RNA extraction buffer of PicoPure RNA Isolation Kit (Cat. Kit0204, Applied Biosystems, Foster City, USA). The RNA sequencing library was prepared using the NuGen v2 RNA-Seq kit and sequenced using an Illumina HiSeq 2000 machine (D-BSSE, ETH Zurich, Basel). Obtained single-end RNA-seq reads (50-mers) were mapped to the mouse genome assembly, version mm9, with SpliceMap [200,201] and included in the R/Bioconductor package QuasR ([202], version 1.8.4) using the command 'qAlign("samples.txt", "BSgenome.Mmusculus.UCSC.mm9", splicedAlignment= TRUE)'. Using RefSeq mRNA coordinates from UCSC (genome.ucsc.edu, downloaded in December 2013) and the qCount function, we quantified gene expression as the number of reads that started within any annotated exon of a gene (exon-union model). The differentially expressed genes were identified using the edgeR package (Robinson et al., version 3.10.5). Genes with FDR smaller than 0.05 and minimum log2 fold change of +/- 1 were used for further analysis.

4.2.12 RNA & RT-PCR

Quantitative RT-PCR: Total RNA was extracted using the RNA Plus extraction kit (Macherey-Nagel, Cat. 740955.50, Düren, Germany) according to the manufacturer's protocol. cDNA synthesis was carried out using the high capacity cDNA reverse transcription kit (Cat. 4368814, Applied Biosystems, Foster City, USA). Quantitative PCR was performed using SYBR Green reagent (Cat. 4368706, Applied Biosystems, Foster City, USA) and an ABI prism 7500 sequence detection system. Ct values were normalized to Gapdh expression and relative expression was quantified using 1/dCt or the $2^{(-ddCt)}$ method. Primers are indicated in **Table 6**.

TABLE 6. Primer sequences used for RT-PCR.

Target	Sequence forward (5'-3')	Sequence reverse (5'-3')
mGapdh	ATGACATCAAGAAGGTGGTG	CATACCAGGAAATGAGCTTG
mNsd1 exon 2	CAACAGCACTTGCTATGAAACAG	GCATCGTCCACACCAGTAAAA
mNsd1 exon 5/6	CAAAGAGCTCCTCCTACAAGTAAACC	CCGAATAGCTGGCTCAGGGA
mNsd1 exon 13/14	TGCTTCTAAAGGTCGTCTGATGCGC	CTAGGGGTGAAGTGATTAGGGCAGA
mGata1	GTGTCCTCACCATCAGATTCCAC	TCCCTCCATACTGTTGAGCAGTG
mGata2	CACCCCTATCCCGTGAATCC	GGCGGCCACTCGCAG
mSpi1 (Pu.1)	CGATTCAGAGCTATACCAACGTCC	ACTCGTTTGTTGTGGACATGGTG
mKit (CD117)	ACGATGTGGGCAAGAGTTCC	GCCTGGATTTGCTCTTTGTTGT
mHbbA	TGATGTAAGCCACGGCTCTG	CAGTGGCTCAGGAGCTTGAA
mHbbB	GTCTCTTGCCTGTGGGGAAA	CAACCAGCAGCCTGCCC
mGpa	CTCCTGTGGTGGCTTCAACT	ACGGCATTCCTCCAATGTGT
mBc2l1	GCCTTTTTCTCCTTTGGCGG	TCCACAAAAGTGTCCCAGCC
mCdkn1a (p21)	CCTGGTGATGTCCGACCTG	CCATGAGCGCATCGCAATC
mCbfa2t3 (Eto2)	GCTGAAGTGAAGACGCAGC	GCCGTTCATCAGTGTGTGAG
Ha-Tag/mGata	ATGACGTGCCTGACTATGCC	TGCATTTGGGGAAGTGGAAGA
hGapdh	GTGGTCTCCCTGACTTTCAACAGC	ATGAGGTCCACCTGCTTGCTG
hNsd1	AGG TAC AGG AGC AGG TGC ACA	AGC ACT AGA TCG ACC TCG GGC
hGata1	AAA CGG GCA GGT ACT CAG TG	CGG TTC ACC TGG TGT AGC TT

4.2.13 Colony forming assay

For whole BM analysis, approximately $4x10^4$ cells were plated in methylcellulose M3434 (Methocult, StemCell Technologies, Vancouver, Canada). Colonies were scored after 8-10 days. Pictures were taken on Olympus IX50 microscope with 2x, 4x and 10x magnification. Cells were washed, resuspended, counted with trypan blue and if applicable replated into new methylcellulose. LSK and MEP cells were flow-sorted according to fluorescence- activated cell sorting section and plated at densities of 1000 respectively 5000 cells into methylcellulose M3434. EPO-dependence was analyzed by replating BM cells from diseased Vav-iCre$^{+/tg}$;*Nsd1*$^{fl/fl}$ mice into M3234 (Methocult, StemCell Technologies, Vancouver, Canada) with 2U/ml EPO (Eprex 4000, Cat.9096976, Pharmacy of University Hospital Basel) or M3534 (Methocult, StemCell Technologies, Vancouver, Canada). For colony formation analysis of human CD34$^+$ cells, $5x10^3$ cells were plated into H4434 (Methocult, StemCell Technologies, Vancouver, Canada).

4.2.14 Benzidine Staining

After scoring of M3434 plates, dishes were incubated with a mix of 2 volumes of 0.3% hydrogen peroxide and 5 volumes of 0.2% di-hydrochloride benzidine (Cat. B3383, Sigma Aldrich, Buchs, Switzerland) in 0.5M acetic acid/1x PBS for 5min at 37°C.

4.2.15 Western Blotting

For protein detection, nuclear lysates were prepared by resuspending cells in hypotonic lysis buffer (10mM HEPES pH7.9, 10mM KCl, 0.1mM EDTA, 0.1mM EGTA, 1mM DTT) for 15min on ice, followed by treatment with 0.1% NP-40 and 15sec vortexing. Nuclei were spun down at 14.000rpm for 2min at 4°C and supernatant containing cytoplasmic fraction kept for analysis. Pellets were resuspended in nuclear lysis buffer (20mM HEPES pH7.9, 0.4M NaCl, 1mM EDTA, 1mM EGTA, 1mM DTT). In addition, pellets were sonicated for 3 cycles (30sec sonication, 30sec pause) on a Bioruptor pico sonicator (Cat. B01060001, Diagenode, Seraing, Belgium) and left for 20min on ice before spinning down at 14.000rpm for min at 4°C. Lysates were kept for analysis of nuclear proteins and remaining pellets used for histone extraction in 0.2N HCl and β-mercaptoethanol. Lysis buffers were supplemented with Complete Mini protease inhibitors (Cat. 11836153001, Roche). Proteins were quantified by Bradford assay (Cat. 500-0006, Biorad, München, Germany) and loading adjusted. Samples were prepared in 4x Laemmli buffer (Cat. 161-0747, Biorad, München, Germany) and boiled for 10min at 95°C before loading on pre-cast (BioRad) or handcasted gels of different percentages. For Nsd1 blot, 50ug of nuclear extract was loaded on a 5%running gel. Wet transfer was done overnight at 4°C in 5%Methanol/0.1%SDS/Tris-Base-Bicine buffer on 0.45μM nitrocellulose membranes. Membranes were blocked in 5% non-fatty milk (NFM)/4%BSA in PBS-1% Tween for 2 hours at room temperature. For blotting GATA1, 10ug nuclear extract was loaded on 10% gels and semi- dry transfer was done for 30min on nitrocellulose 0.2μM (Cat. 170-4158, Biorad, München, Germany). Membranes were blocked for 2hours at room temperature in 5%NFM/PBS-1%Tween. For histone separation, 15% gels were used and general transfer was

used. Blots were probed overnight with antibody in 2.5%NFM/PBS-1%Tween, washed three times for 15minutes in PBS-1% Tween and probed with a secondary antibody in 2.5%NFM/PBS-1%Tween. Again, blots were washed three times in for 15minutes in PBS-1%Tween and then probed with Supersignal West Femto Max substrate (Cat. 11859290, Thermo Scientific, Reinach, Switzerland). Carestream Biomax Kodak films were used for development (Cat. Z-373508-50EA, Sigma, New York, USA). Information regarding antibodies can be found in Table 7.

TABLE 7. Antibodies used for western blot and histone flow cytometry

Antibody	Concentration	Clone	Manufacturer	Cat.Nr.	Host species
GATA1	1:400 (IF), 1:500 (WB)	D52H6 XP	Cell Signaling	CST3535	rabbit
GATA1	1:1000	N6	Santa Cruz	265	rat
LAMIN-A/C	1:2000	N18	Santa Cruz	6215	goat
NSD1	1:1000	kind gift of Antoine Peters, FMI Basel			rabbit
PU.1	1:500	T-21	Santa Cruz	352	rabbit
ACTIN	1:6000	C11	Santa Cruz	1615	goat
H3K36ME1	1:1000 (WB & Flow)		Cell Signaling	5928	rabbit
H3K36ME2	1:1000 (WB & Flow)	C57H12	Cell Signaling	2901	rabbit
H3K36ME3	1:1000		Abcam	9050	rabbit
H3K4ME3	1:1000		Millipore	07-473	rabbit
H3K9ME3	1:1000		Active motif	39161	rabbit
H4K20ME3	1:1000	kind gift of Antoine Peters, FMI Basel (IMP0083)			rabbit
H3	1:1000	D1H2	Cell Signaling	4499	rabbit
ETO-2	1:200-400 (WB)	C-20	Santa Cruz	9739	goat
488 GOAT ANTI- RABBIT	1:1000 (IF), 1:5000 (Flow)		Invitrogen	A31628	

4.2.16 Extensive self- renewing erythroblast cell culture

Fetal liver derived extensive self-renewing erythroblasts (ESRE) cells were obtained as previously described [203]. Briefly, hematopoietic fetal liver cells were cultured for more than one week in “maintenance medium”: StemSpan SFEM (Cat. 9650, StemCell Technologies, Vancouver, Canada), supplemented with 1%Pen/Strep, 0.4%cholesterol (Cat. 12531-018 Gibco, Thermo Fisher Scientific, Reinach, Switzerland), 2U/ml hEpo (Eprex 4000, 9096976, Pharmacy of University Hospital

Basel), 100ng/ml mScf (Cat. 250-03, Peprotech, London, UK), 10^{-6}M dexamethasone (Cat. 265005, Calbiochem, Sigma Aldrich, Buchs, Switzerland) and 40ng/ml hIGF-1 (Cat. 100-11, Peprotech, London, UK). Cells were split every second day and presence of pro- erythroblasts was verified by flow cytometry ($DAPI^-$ /FSC^+/$CD71^+$/$Ter119^-$) and cytospots. Pro- erythroblasts were subjected to differentiation in IMDM (Cat. 31980022, Gibco, Thermo Fisher Scientific, Reinach, Switzerland), 1%P/S, 10%FCS, 10%PFHMII (Cat. 12040077, Gibco, Thermo Fisher Scientific, Reinach, Switzerland), 5%hPDS (0.45μM filtered, Blood donation centre, University Hospital Basel), monothioglycerol (Cat. M6145, Sigma Aldrich, Buchs, Switzerland), 100ng/ml mSCF and 2U/ml hEpo. Erythroblasts cultures from adult mice were established after lineage depletion of BM cells. Cells were kept in maintenance culture for more than six days before flow cytometry analysis. For colony formation assays, cells were used within five days after isolation.

4.2.17 Retroviral Gene Transfer

Retroviral stocks were produced by transient co-transfection of packaging vectors (*pIPAK6*) and respective plasmids using turbofect transfection reagent (Cat. R0531, Life Technologies, Paisley, UK) in HEK293T-LX cells kept in DMEM (Cat. 61965059, Gibco Lubio, Thermo Fisher Scientific, Reinach, Switzerland) with 10%FSC and 1%P/S. Viral supernatant was harvested 48 and 72 hours after transfection, 10x Vivaspin 20 (Cat. 12.8303.10, Sartorius, Göttingen, Germany) concentrated at 4000rpm for 40min at 4°C and snap frozen in liquid nitrogen and stored in -80 until usage. Cells were spin- infected either in StemSpan SFEM, supplemented with 50ng/ml hTPO (Cat. 300-18, Peprotech, London, UK) and 50ng/ml mSCF (="Stem") or in maintenance medium used for erythroblast culture as described above, in presence of 5μg/ml polybrene (Cat. 10.7689, Sigma Aldrich, Buchs, Switzerland) with virus for 90min, 2500rpm at 30°C. Four hours after spin infection the, cells were washed with PBS and plated in maintenance medium. Two days after spin infection the cells were selected with 2μg/ml puromycin (Cat. A11138-03, Gibco, Thermo Fisher Scientific, Reinach, Switzerland) or $EGFP^+$ cells were FACS enriched as previously described. See **Table 8** for more information about the viral constructs used.

TABLE 8. Plasmids

Short name	Full name
Mock	*pMSCV-pgk-puro*
mGata1(-WT)	*pMSCV-mGata1-pgk-puro*
Mock	*pMSCV-pgk-puro-IRES-GFP*
hETO-2(-WT)	*pMSCV-hETO-2-pgk-puro-IRES-GFP*
mNsd1	*pMSCV-mNsd1-pgk-puro-IRES-GFP*
mNsd1-ΔNID	*pMSCV-mNsd1-ΔNID-pgk-puro-IRES-GFP*
mNsd1-ΔNID/Setmut	*pMSCV-mNsd1- ΔNID /Set^{mut}-pgk-puro-IRES-GFP*
mNsd1-Setmut	*pMSCV-mNsd1-Set^{mut}-pgk-puro-IRES-GFP*
shRNA Ctrl	*pLKO.1 mock shRNA (Addgene plasmid 8453)*
shRNA NSD1 353	*pLKO.1 shRNA targeting NSD1 number 353 (TRCN0000061353)*
shRNA NSD1 369	*pLKO.1 shRNA targeting NSD1 number 369 (TRCN0000238369)*
shRNA NSD1 372	*pLKO.1 shRNA targeting NSD1 number 372 (TRCN0000238372)*
pIPAK6	*pIPAK6 enevlope plasmid*
pMD2G	*pMD2G envelope plasmid*
pMLDg/PRE	*pMLDg/PRE Packaging plasmid*
pRSV/Rev	*pRSV/Rev expression plasmid*

4.2.18 shRNA-mediated knockdown

CD34^{+} cells obtained from indirect CD34 MicroBead Kit, human (Cat. 130-046-701, Miltenyi Biotec, Bergisch Gladbach, Germany) of peripheral blood or cord blood of healthy donors were kept in StemLine II medium (Cat. S0192, Sigma Aldrich, Buchs, Switzerland), supplemented with human cytokines such as 50ng/ml hTPO (Cat. 300-18, Peprotech, London, UK), 50ng/ml hFLT3 ligand (Cat. 300-19, Peprotech, London, UK), 50ng/ml hSCF (Cat. 300-07 Peprotech, London, UK) and 1U/ml hEPO (Eprex 4000, Cat.9096976, Pharmacy of University Hospital Basel). shRNAs were expressd from lentiviral vectors. For transduction lentiviral stock was produced by transient co-transfection of packaging vectors (*pMD2G, pMLDg/PRE, pRSV/Rev*) and respective lentiviral shRNA plasmid (*pLKO*) (shRNA Ctrl and shRNA NSD1 353, 369 and 372) using lipofectamine 2000 (Cat. 11668027075, Invitrogen, Thermo Fisher Scientific, Reinach, Switzerland) in HEK293T-LX cells kept in DMEM (Cat. 61965059, Gibco Lubio, Thermo Fisher Scientific, Reinach, Switzerland) with 10% FSC and 1% P/S.. Viral supernatant was harvested 48 and 72 hours after transfection, snap frozen in liquid nitrogen and stored in -80 until usage. Cells were spin- infected in presence of 5ug/ml polybrene (Cat. 10.7689, Sigma Aldrich, Buchs, Switzerland) with virus for 90min, 2500rpm at 30°C. Six hours after spin infection, cells were washed with PBS

and plated in Stem Line II cytokine enriched medium. Two days after spin infection cells were selected with 2μg/ml puromycin (Cat. A11138-03, Gibco, Thermo Fisher Scientific, Reinach, Switzerland). See **Table 8** for more information regarding the constructs used.

4.2.19 Analysis of megakaryocytes

Megakaryocytes were manually counted at 40x magnification on Olympus BX61 microscope on HE- stained BM sections of young and diseased *Nsd1*$^{fl/fl}$ and *Vav1-iCre*$^{+/tg}$*;Nsd1*$^{fl/fl}$ mice. Per mouse, 20 fields were counted and the average per mouse calculated. DNA ploidy of MK was performed on fresh flushed BM cells resuspended in 2mL of 1/1 CATCH (0.1% EDTA/0.1% Albumin/25mM Hepes) and 5%FCS/PBS media. Cells are pre-incubated for 10min with anti-CD16/CD32 Fc(III/II) antibody (Cat. 14-0161-86 eBioscience, clone 93, dilution 1:50, Affymetrix, Thermo Fisher Scientific, Reinach, Switzerland) at 4°C and stained 30min with APC–anti-CD41 (Cat. 17-0411-82, eBioscience, dilution 1:200, Affymetrix, Thermo Fisher Scientific, Reinach, Switzerland) mAb. Then cells were washed and incubated in a 0.1% hypotonic citrate solution containing 50 μg/mL propidium iodide (PI) and 50μg/ml RNAseA for at least 4h at 4°C. Ploidy was measured on a LSR II (BD, Mountain View, CA) cytometer, gating on CD41+ cells. Ploidy of MK in cell culture followed the same protocol except 2%FCS/PBS was used instead of CATCH/PBS medium, and cells were incubated 30min at RT in hypotonic citrate solution.

4.2.20 Immunofluorescent staining

Cytopsots of erythroblasts kept in “maintenance” or “differentiation” medium were obtained by washing cells in 1x PBS. Cytospots of approximately 1.5x10^5 cells were made by centrifugation for three minutes at 300rpm using a Shandon Cytospin 3 centrifuge using Cytofunnel disposable sample chambers (Cat. 5991040, Thermo Fisher Scientific, Reinach, Switzerland) and one-circle, coated cytoslides (Cat. 5991056, Thermo Fisher Scientific, Reinach, Switzerland). Cells were fixed with cold 4% paraformaldehyde/PBS (Cat. 15710, EMS; Hatfield, US) for 5min at room

temperature and then washed with PBS before blocking with 10% goat serum (G9023, Sigma Aldrich, Buchs, Switzerland) in 0.2%triton-X/PBS for 45min. Then cells were incubated with anti-Gata1 antibody (Table 7) diluted in blocking buffer overnight at 4°C. The other day, slides were washed three times for 5min with 0.2%triton X-100/ PBS and incubated with secondary goat- anti rabbit antibody conjugated to AF488 (Table 7) for 1hr. Again, slides were washed three times for 5min with 0.2%triton X-100/ PBS, incubated with 1:10000 DAPI (Cat. D1306, Life Technologies, Paisley, UK) in blocking buffer for 5min and washed again three times before mounted with Fluorsave (Cat. 345789, Merck, Darmstadt, Germany). All steps were carried out in a wet chamber. Pictures were taken on a Nikon TI (Tokyo, Japan) at 100x magnification.

4.2.21 Detecting chromatin marks by flow cytometry

For intracellular flow analysis of histone marks, lineage depleted BM cells of young *Nsd1*$^{fl/fl}$ and *Vav1-iCre*$^{+/tg}$*;Nsd1*$^{fl/fl}$ mice were washed in PBS at 3000rpm for 5min in a table- top centrifuge and then slowly fixed in 70% cold ethanol. Samples were stored at -20°C up to one week before washing in PBS for 5min at 3000rpm. Cells were permeabilized in 0.5%triton-X/PBS for 30min at room temperature and vortexed mildly every 5min before washing with PBS again. Cells were incubated with primary antibody in 1%BSA/0.5%Tween/ PBS (Table 7) for 1hr at room temperature and then washed with PBS. Secondary goat- anti rabbit antibody conjugated to AF488 in 1%BSA/0.5%Tween/ PBS was applied for 30min at room temperature in the dark. Unstained cells, secondary antibody incubated and histone 3 primary antibody cells were used as controls.

4.2.22 Analysis of NSD1- Set domain on GATA1 methylation

Protein expression and purification

For protein expression, E.coli BL21 cells (Novagen) were transformed with the corresponding plasmids and grown in Luria-Bertani media at 37°C, until they reached an optical density of 0.6 to 0.8 at 600 nm. The cells containing GST fused NSD1

were transferred to 30°C and then induced with 1 mM isopropoyl-beta-D-thiogalactopyranoside and grown for 4 h. Whereas the cells containing GST fused GATA full length protein were transferred to 20°C and then induced with 1 mM isopropoyl-beta-D-thiogalactopyranoside and grown for 10-12 h. The cells were harvested by centrifuging at 4500 RCF for 15 minutes at 4°C.

The cell pellet was resuspended in sonication buffer (50 mM Tris/HCl pH 7.5, 150 mM NaCl, 1 mM DTT and 5 % glycerol) and disrupted by sonication. Then the sonicated sample was centrifuged at 18,000 g for 90 minutes and the supernatant was collected. In the meantime glutathione sepharose 4B beads (600 ml) (GE Healthcare) were equilibrated with 60 ml of sonication buffer. The collected supernatant was then loaded on the pre-equilibrated beads and the beads were later washed with wash buffer (120 ml) to remove the weakly bound proteins (50 mM Tris/HCl pH 8.0, 500 mM NaCl, 1 mM DTT and 5 % glycerol). Finally, the bound proteins were eluted with wash buffer containing 40 mM glutathione. Fractions containing the proteins were pooled and dialyzed against low glycerol dialysis 1 buffer (20 mM Tris/HCl pH 7.4, 100 mM KCl, 0.5 mM DTT and 10 % glycerol) for 3 h and then overnight against high glycerol dialysis 2 buffer (20 mM Tris/HCl pH 7.4, 100 mM KCl, 0.5 mM DTT and 60 % glycerol). The quality of the protein purification was analyzed by loading proteins on the SDS-PAGE.

Methylation of peptide arrays

Peptide arrays were washed for 5 min in the methylation buffer containing 50 mM Tris/HCl (pH 9.0), 5 mM $MgCl_2$, and 4 mM DTT and afterwards incubated at room temperature for 60 min in methylation buffer supplemented with 50 nM NSD1 and 0.76 μM labeled [methyl-3H]-AdoMet (Perkin Elmer). Then, the membranes were washed 5 times for 5 minutes with wash buffer (100 mM NH_4HCO_3 and 0.1% SDS). Finally, the peptide arrays were incubated with Amplify NAMP100V solution (GE Healthcare, Munich, Germany) for 5 minutes. The peptide arrays were exposed to Hyperfilm TM high performance autoradiography film (GE Healthcare, Munich, Germany) in the dark at -80°C for 3-7 days. The autoradiography films were developed in Optimus TR developing machine as described in [204]

In vitro protein methylation assays

Protein methylation reactions were performed by incubating substrate proteins (2 μM) in methylation buffer (50 mM Tris/HCl (pH 9.0), 5 mM $MgCl_2$, and 4 mM DTT)

supplemented with 0.76 μM tritium-labeled AdoMet and 100 nM NSD1 for 4 hours at 25° C. Reaction mixtures were halted by boiling with 2 X SDS loading buffer at 95°C for 5 minutes. Afterwards, the samples were separated on 16% SDS PAGE and the methylation signal was detected by autoradiography.

Peptide arrays were synthesized on cellulose membrane by the SPOT synthesis method using a Multipep system (Intavis AG). Each spot contained approximately 9 nmol peptide (Autospot Reference Handbook, Intavis AG). Successful synthesis of peptides on the membrane was confirmed by bromo- phenol blue staining. The membranes containing the peptides were washed for 20min in methylation buffer containing 50mM Tris-HCl (pH 9.0), 5mM MgCl2 and 4mM dithiothreitol (DTT) and incubated at room temperature for 60min in methylation buffer containing 50 nM NSD1 SET domain (constructs were kind gift of kind gifts from Professor Pierre Chambon, Strassbourg, France) and 0.76mM labeled [methyl-3H]-AdoMet (Perkin Elmer).

4.2.23 Human leukemia cell lines

Human cell lines (MOLM-13, OCI-M2, K-562, HEL, MEL, and F-36P) were obtained from the Leibnitz Institute DSMZ-German Collection of Microorganisms and cell cultures (www.dsmz.de). MOLM13, K-562 and HEL cells were cultured in in RPMI (Cat. 61870, Gibco, Thermo Fisher Scientific, Reinach, Switzerland) containing 10%FCS (Cat. 2-01F10-I, Amimed, Bioconcept, Allschwil, Switzerland) and 1%P/S (Cat. 15140, Gibco, Thermo Fisher Scientific, Reinach, Switzerland). MEL cells were cultured in DMEM (Cat. 61965059, Gibco, Thermo Fisher Scientific, Reinach, Switzerland) containing 10%FCS and 1%P/S. OcI-M2 cells were kept in IMDM (Cat. 31980022, Gibco, Thermo Fisher Scientific, Reinach, Switzerland), containing 20% FCS and 1%P/S. F-36P cells were kept in RPMI medium containing 20% FCS and 1%P/S and supplemented with 10ng/ml hGM-CSF (Cat. 300-03, Peprotech, London, UK). See **Table 9** for information regarding ATCC number and origin of the cell lines used.

TABLE 9. Cell lines

Cell line	ATCC Nr.	Origin
MOLM-13	554	AML FAB M5a
OCI-M2	619	AML M6
K-562	10	CML
HEL	11	AML M6
MEL	from I. Vizirianakis,	Murine
(clone 745-PC–4)	Thessaloniki	erythroleukemia
F-36P	543	AML M6

4.2.24 Statistical analysis

Data has been analyzed in excel using F- Test to test the two- sided probability of unchanged variances between datasets. P- values were taken into account for further analysis using two- sided T- Test to analyze differences in the mean between two groups. In case, two or more groups were compared, variances were analyzed by ANOVA followed by Turkey or Sidak multiple comparison tests using Prism 6 software (Graph Pad). P- values <0.05 were considered as statistically significant. Data presented as mean, error bars represent ±SD or SEM as indicated in the figure legend.

4.3 RESULTS

4.3.1 Generation of Nsd1 knockout mice

To inactivate the *Nsd1* gene in the hematopoietic system we used a strain with a previously established floxed $Nsd1^{+/L3}$ allele [173]. We eliminated the floxed *pgk-neomycin* selection cassette leaving two *loxP* sites flanking the largest coding exon 5, here referred as $Nsd1^{fl/fl}$ (**Figure 12A**). We intercrossed $Nsd1^{fl/fl}$ mice with a *Vav1-iCre*$^{tg/+}$ transgenic strain leading to constitutive cleavage in fetal and adult hematopoiesis (**Figure 12B**) [205]. *Vav1-iCre;*$Nsd1^{fl/fl}$ mice were born according to Mendelian rules without any abnormalities. In the BM of *Vav-iCre;*$Nsd1^{fl/fl}$ mice we only detected the excised allele resulting in almost undetectable levels of *Nsd1* exon 5 mRNA and loss of the NSD1 protein. Heterozygous *Vav-iCre;*$Nsd1^{+/fl}$ mice expressed normal *Nsd1* levels and like the $Nsd1^{fl/fl}$ mice never developed any signs of disease during a normal lifespan (**Figure 12C-E**).

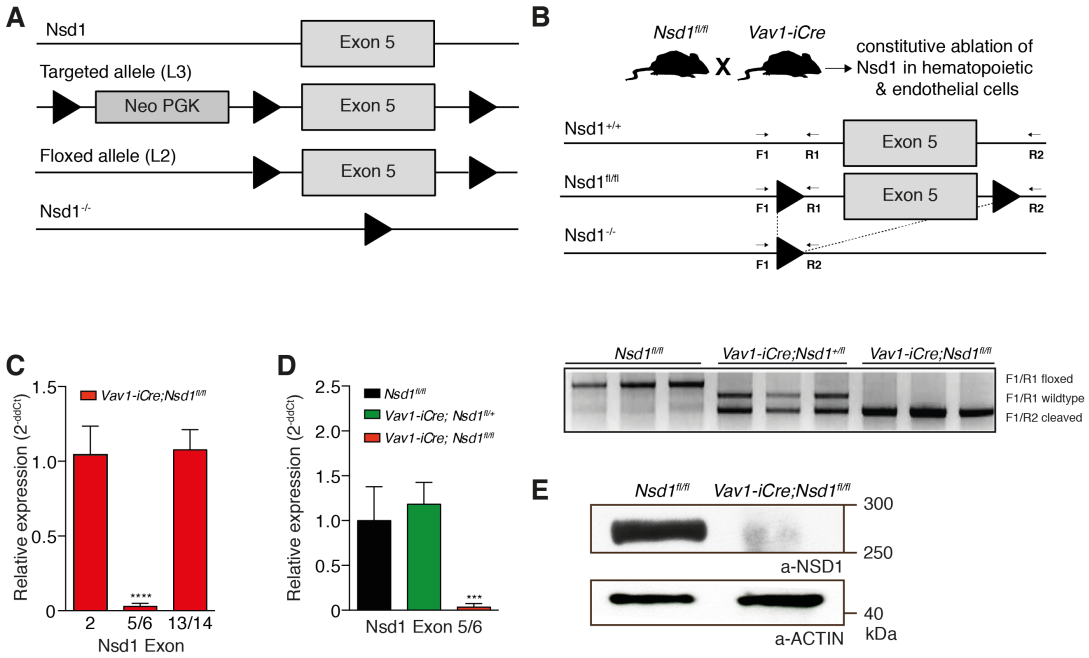


FIGURE 12. Generation of *Nsd1* knockout mice.

(A) Schematic depiction of the targeted *Nsd1* gene allele. The *pgk- neomycin* resistance expression cassette was removed from the targeted *L3* allele to obtain the floxed *L2* allele where exon 5 is flanked by *LoxP* sites. **(B)** Schematic depiction of the model used. *Nsd1*$^{fl/fl}$ mice were intercrossed with the *Vav1-iCre* transgenic strain leading to constitutive ablation of *Nsd1* in hematopoietic and endothelial cells. PCR primer pairs F1 and R1/R2 were used to genotype mice and determine the presence of wildtype and the floxed *Nsd1* alleles in genomic DNA. **(C)** Expression of different exons of *Nsd1* mRNA in BM from *Vav1-iCre;Nsd1*$^{fl/fl}$ mice by quantitative RT-PCR analysis. In contrast to exon 2 (n=12) and the exon 13/14 (n=10) junction, low to undetectable levels of products of the exon5/6 (n=19) junction were detected indicating almost complete excision. Bars represent average relative expression normalized to *Gapdh* and *Nsd1*$^{fl/fl}$ mice (One-way ANOVA with Turkeys multiple comparisons test. **** p<0.0001). **(D)** Expression of *Nsd1* exon 5 mRNA in BM from *Vav1-iCre;Nsd1*$^{fl/fl}$ (red bar, n=4) , *Vav1-iCre;Nsd1*$^{fl/+}$ (green bar, n=4) and *Nsd1*$^{fl/fl}$ (black bar, n=3) mice by quantitative RT-PCR. Bars represent the average relative expression normalized to *Gapdh* and control mice (One-way ANOVA with Turkeys multiple comparisons test. *** p<0.001, **** p<0.0001). **(E)** Expression of Nsd1 protein in nuclear extracts of *Nsd1*$^{fl/fl}$ and *Vav1-iCre;Nsd1*$^{fl/fl}$ spleen tissue. Blots were probed with an anti- NSD1 and anti-ACTIN antibody detecting bands of the appropriate sizes. Data presented as mean, error bars represent ±SD.

4.3.2 Loss of Nsd1 leads to a lethal disease with accumulation of erythroid progenitor cells

At the age of 7-17 weeks (median 85 days) *Vav1-iCre;Nsd1$^{fl/fl}$* mice developed severe signs of distress (**Figure 13A**). All mice had significantly enlarged spleens and livers with abnormal focal infiltrations, also present in the lungs or kidneys suggesting an aggressive malignant disorder (**Figure 13B-E**). Analysis of peripheral blood revealed highly variable white blood counts, with a significant reduction in erythrocytes, severe thrombocytopenia and reticulocytosis (**Figure 13F-I**). Although the white blood counts and myeloid cells were in the normal range and not significantly changed, the counter reported the presence of “unclassified” leukocytes (LUCs) (**Figure 13F,K-L**) and abnormal blast-like cells with a dark-blue cytoplasm were seen on the blood smears (**Figure 14A**).

Flow cytometric analysis of hematopoietic cells in peripheral blood of diseased *Vav1-iCre;Nsd1$^{fl/fl}$* mice revealed a significantly increased number of cells expressing CD71$^+$, CD4$^+$ and TER119$^-$ and a reduction of cells that expressed CD41$^+$ cells. In contrast, cells expressing Mac-1, Gr-1, CD8 or B220 were not significantly changed in peripheral blood (**Figure 14B&C**). The presence of CD71$^+$ and TER119$^-$ expressing cells led us to further analyze the erythroid lineage: hereby we separated early from late progenitors progressing from R0 fraction (CD71$^{-/dim}$, TER119$^-$), CD71^{dim} (CD71^{dim}, TER119$^-$), R1 (CD71$^+$, TER119$^-$), R2 (CD71$^+$, TER119$^+$), R3 (CD71$^{+/dim}$, TER119$^+$) and R4 (CD71$^-$, TER119$^+$) (**Figure 14D**) [16]. We found a significant increase of R1, R2, R3 and CD71^{dim} populations and a significant decrease of the R4 fraction in spleens of diseased mice (**Figure 14E&F**). In BM of diseased mice, we found a significant increase of CD71^{dim} expressing cells (**Figure 14E-G**).

We next asked the question whether accumulated CD71^{dim} cells may co- express other surface markers by investigating BM and spleen cells from diseased *Vav1-iCre;Nsd1$^{fl/fl}$* mice using c-Kit, Sca-1 and FcγRII/III after lysis of the red cells. Spleen cells of diseased *Vav1-iCre;Nsd1$^{fl/fl}$* mice expressed significantly higher levels of CD71$^+$/c-Kit$^-$ and BM cells were significantly enriched for CD71$^+$/c-Kit$^-$ and CD71$^+$/c-Kit$^+$ cells (**Figure 14H-I**). CD71$^+$/c-Kit$^+$ cells expressed variable amounts of FcγRII/III but were mostly negative for CD34 and Sca-1 surface markers (**Figure 14J** & data not shown).

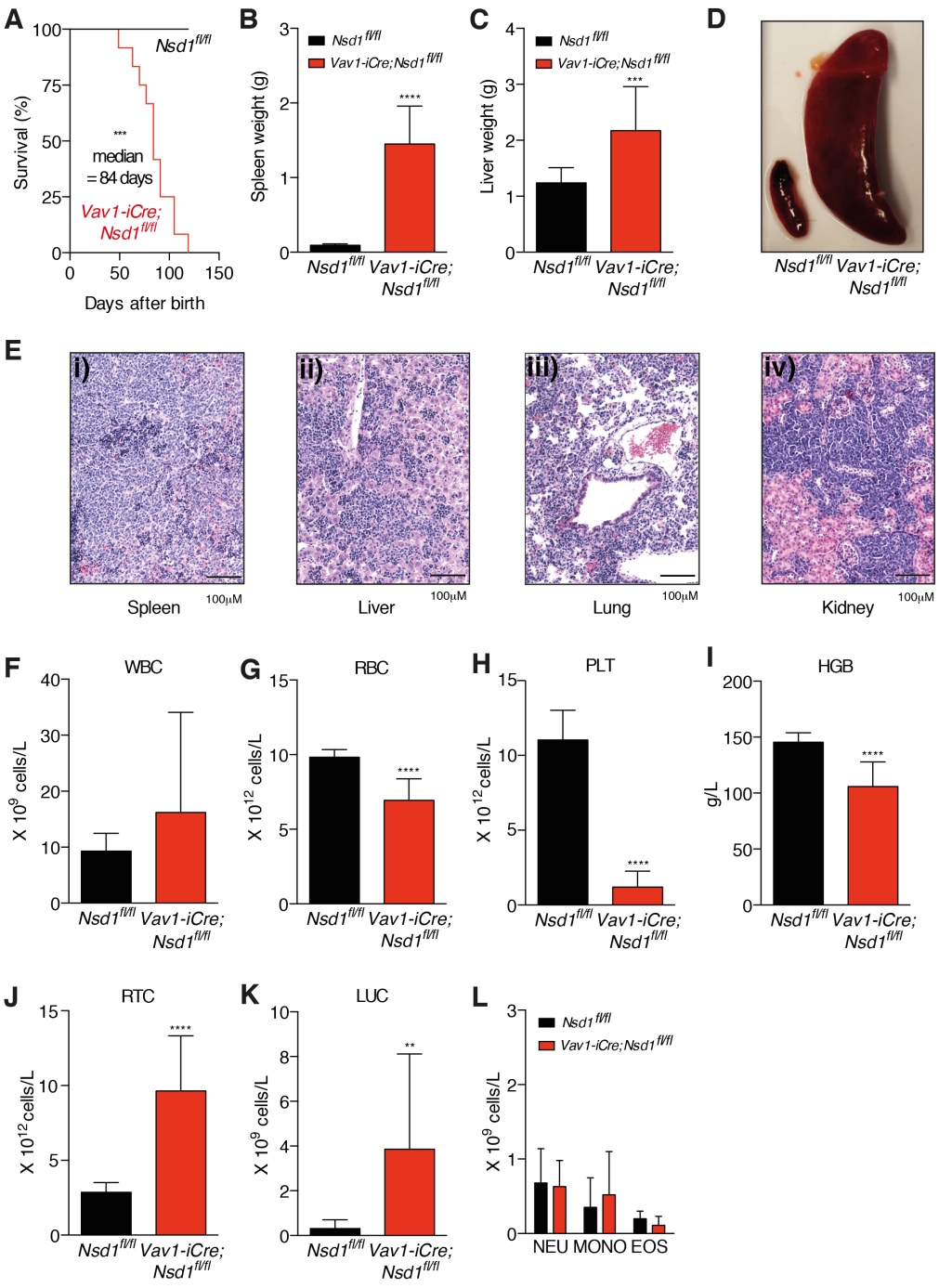

A
Survival (%)
Nsd1fl/fl
***
median
= 84 days
Vav1-iCre;
Nsd1fl/fl
Days after birth
B
Spleen weight (g)
Nsd1fl/fl
Vav1-iCre;Nsd1fl/fl
****
C
Liver weight (g)
***
D
E
i)
ii)
iii)
iv)
Spleen
Liver
Lung
Kidney
100μM
F
WBC
X 10^9 cells/L
G
RBC
X 10^12 cells/L
H
PLT
I
HGB
g/L
J
RTC
K
LUC
**
L
NEU MONO EOS

FIGURE 13. ***Vav1-iCre;Nsd1$^{fl/fl}$*** **mice develop a lethal hematopoietic disease with full penetrance. (A)** Kaplan Meier survival plot of *Nsd1$^{fl/fl}$* (black line) and *Vav1-iCre;Nsd1$^{fl/fl}$* (red line) mice. Survival is shown as percentage at days after birth. Median survival of *Vav1-iCre;Nsd1$^{fl/fl}$* mice is 84 days (n=12 per group). **(B)** Spleen weight of *Nsd1$^{fl/fl}$* (black bar) and symptomatic *Vav1-iCre;Nsd1$^{fl/fl}$* (red bar) mice in gram (n=12 per group. Students t-test with unequal variances. **** p<0.0001). **(C)** Liver weight in grams (n=12 per group. Students t-test with unequal variances. *** p<0.001). **(D)** Representative image of spleens of *Nsd1$^{fl/fl}$* (left) and diseased *Vav1-iCre;Nsd1$^{fl/fl}$* (right) mice. **(E)** Representative images of HE-stained biopsies of the i) spleen, ii) liver, iii) lung and iv) kidney of diseased *Vav1-iCre;Nsd1$^{fl/fl}$* mice. **(F)** Peripheral white cell blood counts (WBC), **(G)** red blood cells (RBC), **(H)** platelets (PLT), **(I)** hemoglobin (HGB), **(J)** reticulocytes (RTC), **(K)** "large unstained cells" (LUC), **(L)** neutrophils (NEU), monocytes (MONO) and eosinophils (EOS) (n=12 per group, Students t-test with unequal variances. ** p<0.01, **** p<0.0001). Data presented as mean, error bars represent ±SD.

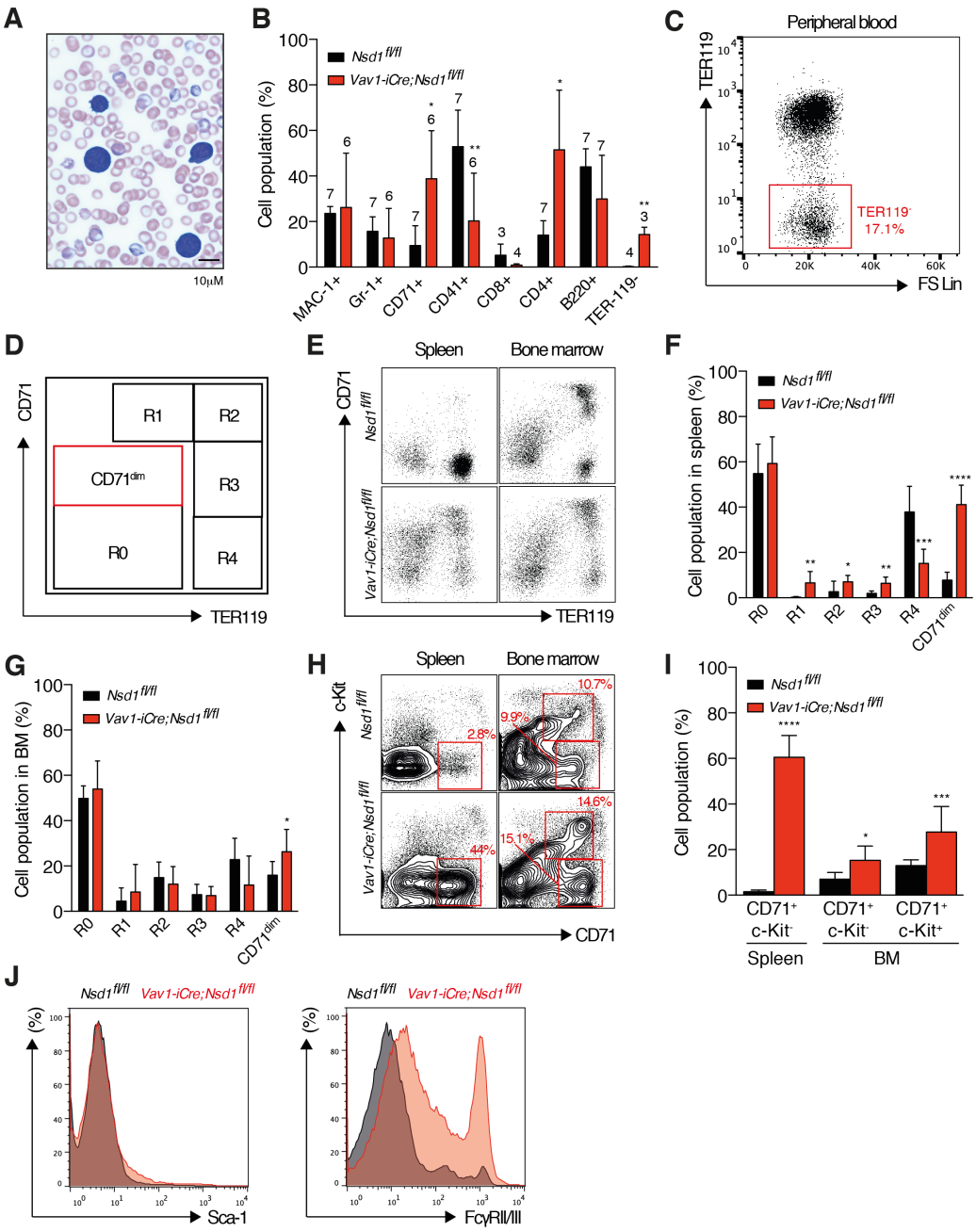

A
10μM
B
Nsd1 fl/fl
Vav1-iCre;Nsd1 fl/fl
Cell population (%)
MAC-1+
Gr-1+
CD71+
CD41+
CD8+
CD4+
B220+
TER-119-
C
Peripheral blood
TER119
TER119-
17.1%
FS Lin
D
CD71
R1
R2
CD71dim
R3
R0
R4
TER119
E
Spleen
Bone marrow
CD71
Nsd1 fl/fl
Vav1-iCre;Nsd1 fl/fl
TER119
F
Nsd1 fl/fl
Vav1-iCre;Nsd1 fl/fl
Cell population in spleen (%)
R0
R1
R2
R3
R4
CD71dim
G
Nsd1 fl/fl
Vav1-iCre;Nsd1 fl/fl
Cell population in BM (%)
R0
R1
R2
R3
R4
CD71dim
H
Spleen
Bone marrow
c-Kit
Nsd1 fl/fl
Vav1-iCre;Nsd1 fl/fl
2.8%
9.9%
10.7%
44%
15.1%
14.6%
CD71
I
Nsd1 fl/fl
Vav1-iCre;Nsd1 fl/fl
Cell population (%)
CD71+ c-Kit-
CD71+ c-Kit-
CD71+ c-Kit+
Spleen
BM
J
Nsd1 fl/fl
Vav1-iCre;Nsd1 fl/fl
(%)
Sca-1
Nsd1 fl/fl
Vav1-iCre;Nsd1 fl/fl
(%)
FcγRII/III

FIGURE 14. Phenotype of hematopoietic organs in diseased *Vav1-iCre;Nsd1$^{fl/fl}$* mice.

(A) Representative image of a Wright-Giemsa stained peripheral blood smear obtained from a symptomatic *Vav1-iCre;Nsd1$^{fl/fl}$* mouse. **(B)** Flow cytometric characterization of cell populations in the peripheral blood of control mice *Nsd1$^{fl/fl}$* mice (black bar) and symptomatic *Vav1-iCre;Nsd1$^{fl/fl}$* (red bar) mice (n as indicated in figure. Students t-test with unequal variances. * $p<0.05$, ** $p<0.01$). **(C)** Ter119-negative red blood cells are found in the peripheral blood of symptomatic *Vav1-iCre;Nsd1$^{fl/fl}$* as shown by a representative flow cytometric analysis of three analyzed mice. **(D)** Flow cytometric gating strategy after CD71 and Ter119 staining to distinguish different maturation steps of erythropoiesis: "R0" fraction ($CD71^{-/dim}$, $Ter119^{-}$), "R1" fraction ($CD71^{+}$, $Ter119^{-}$), "R2" fraction ($CD71^{+}$, $Ter119^{+}$), "R3" fraction ($CD71^{+/dim}$, $Ter119^{+}$), "R4" fraction ($CD71^{-}$, $Ter119^{+}$) and "$CD71^{dim}$" ($CD71^{+/dim}$, $Ter119^{-}$). **(E)** Representative image of flow cytometric analysis of CD71 and Ter119 stained single cell suspensions of the spleens (left column) and bone marrow (right column) of *Nsd1$^{fl/fl}$* control mice (upper row) and symptomatic *Vav1-iCre;Nsd1$^{fl/fl}$* mice (lower row). **(F)** Comparative flow cytometric analysis of erythroid maturation of single cell suspensions (without red cell lysis) of spleens (n=8 per group. left panel) or **(G)** bone marrow (n=9 per group. right panel) of *Nsd1$^{fl/fl}$* control mice (black bar) and symptomatic *Vav1-iCre;Nsd1$^{fl/fl}$* (red bar) mice (Students t-test with unequal variances. * $p<0.05$, ** $p<0.01$, *** $p<0.001$, **** $p<0.0001$). **(H)** Representative image of flow cytometry panel of red blood cell lysed single cell suspensions of spleen (left column) and bone marrow (right column) in *Nsd1$^{fl/fl}$* (upper row) and *Vav1-iCre;Nsd1$^{fl/fl}$* (lower row) mice highlighting CD71 and c-Kit double positive or CD71 single positive staining. **(I)** Percentage flow cytometry marker $CD71^{+}$ and c-Kit^{+} stained cell population in red blood cell lysed single cell suspensions of spleen and bone marrow of *Nsd1$^{fl/fl}$* (n=8/spleen, n=9/BM, black bar) and *Vav1-iCre;Nsd1$^{fl/fl}$* mice (n=8/spleen, n=8/BM, black bar) (Two- way ANOVA with Sidak's multiple comparisons. * $p<0.05$, *** $p<0.001$, **** $p<0.0001$). **(J)** Representative flow cytometry histogram demonstrating Sca-1 and FcγRII/III profile of $CD71^{+}$/c-Kit^{+} in bone marrow of *Nsd1$^{fl/fl}$* (black line) and *Vav1-iCre;Nsd1$^{fl/fl}$* (red line) mice. Data presented as mean, error bars represent ±SD.

4.3.3 Loss of Nsd1 leads to decreased BM cellularity and aberrant colony formation of hematopoietic cells

Despite increased frequencies of CD71-positive cells, the overall BM cellularity was significantly decreased whereas the total number of cells in the spleen was significantly increased (**Figure 15A**). We next investigated whether the loss of cells in the BM correlated with an increase in apoptosis. Flow cytometric analysis revealed no significant changes in early (Annexin-V^+/7-AAD^-) and late apoptotic cells (Annexin-V^+/7-AAD^+) in the BM of diseased *Vav1-iCre;Nsd1$^{fl/fl}$* mice. In contrast, an increase in early apoptotic cells was found in the spleens of diseased mice (**Figure 15B**).

We next explored the functional capacity of the BM and performed colony assays in methylcellulose favoring growth of cells of the myeloid and erythroid lineage. The total number of colonies formed by *Vav1-iCre;Nsd1$^{fl/fl}$* BM cells was slightly reduced (**Figure 15C**). Further classification of the colonies revealed significant reductions in CFU-G/M and BFU-E colonies accompanied by the appearance of sometimes large and abnormally dense, reddish and benzidine-staining positive “BFU-E-like” colonies. These abnormal colonies were composed of myeloid and erythroid progenitors and could be serially plated over three to four rounds in methylcellulose (**Figure 15D-F**). We next investigated cytokine-dependence of these abnormal colonies. Interestingly colony formation (of cells taken from the 4th plating) was significantly reduced in methylcellulose containing only EPO and no colonies were formed in absence of EPO (with SCF, IL-3 and IL-6) (**Figure 15G-I**).

Collectively, our data shows that inactivation of *Nsd1* in the hematopoietic system of the mouse leads to malignant accumulation of erythroid progenitor cells, impaired clonogenic activity and expansion of erythroid progenitors with aberrant self-renewal *in vitro*.

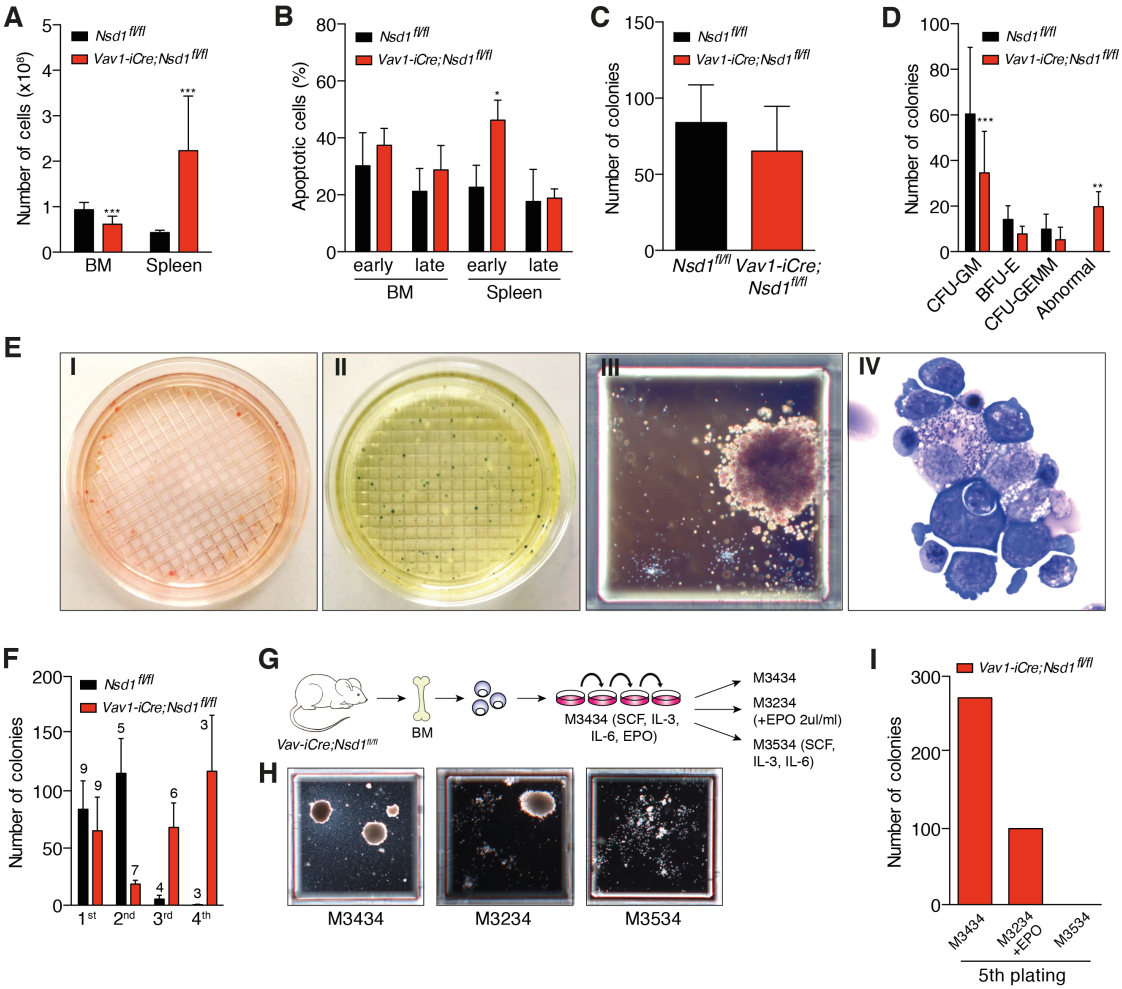

A
Nsd1 fl/fl
Vav1-iCre;Nsd1 fl/fl
Number of cells (x10⁶)
BM
Spleen
B
Apoptotic cells (%)
early
late
BM
Spleen
C
Number of colonies
Nsd1 fl/fl
Vav1-iCre; Nsd1 fl/fl
D
Number of colonies
CFU-GM
BFU-E
CFU-GEMM
Abnormal
E
I
II
III
IV
F
Number of colonies
1st
2nd
3rd
4th
G
Vav-iCre;Nsd1 fl/fl
BM
M3434 (SCF, IL-3, IL-6, EPO)
M3434
M3234 (+EPO 2ul/ml)
M3534 (SCF, IL-3, IL-6)
H
M3434
M3234
M3534
I
Vav1-iCre;Nsd1 fl/fl
Number of colonies
M3434
M3234 +EPO
M3534
5th plating

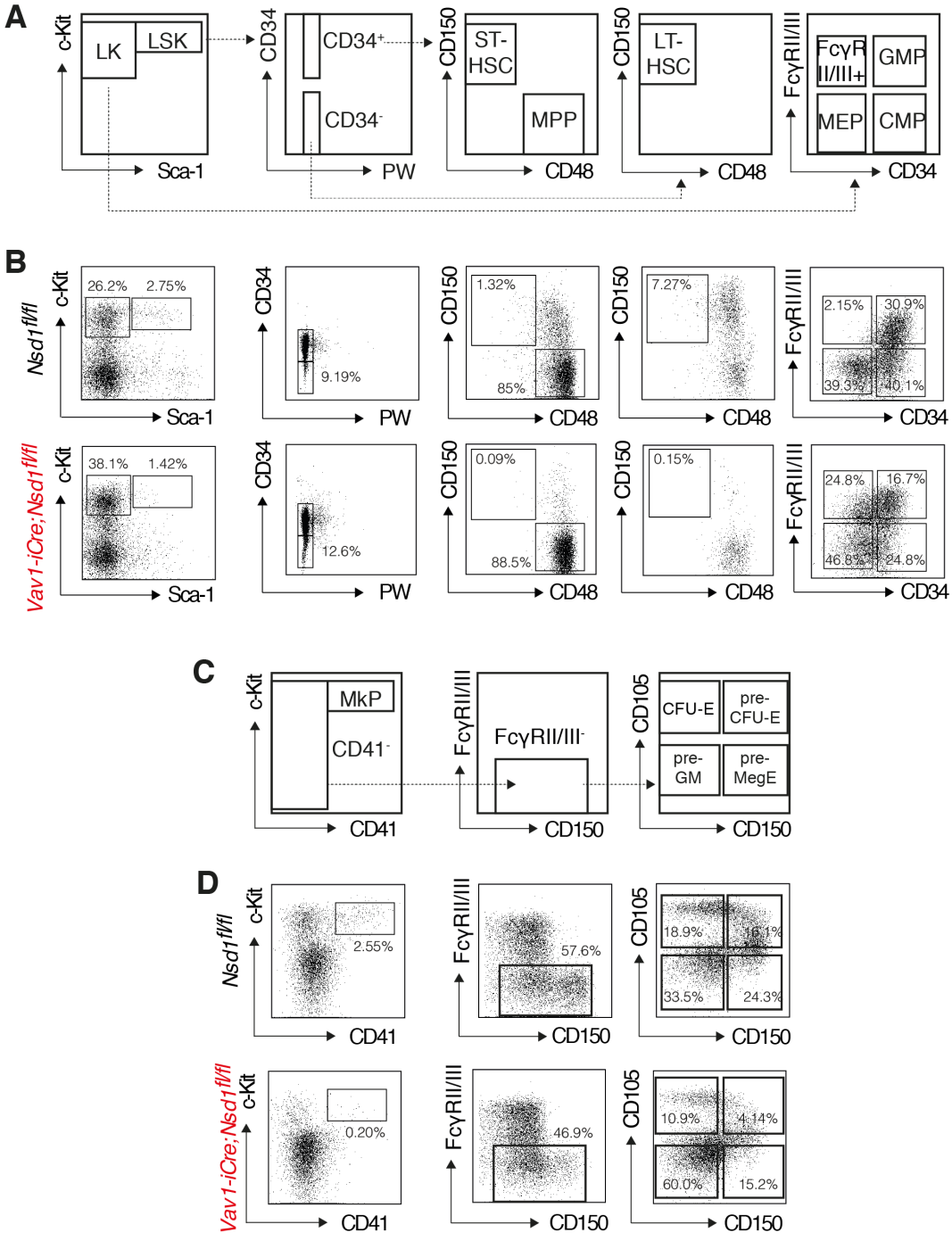
FIGURE 15. Functional analysis of bone marrow cells of *Vav1-iCre;Nsd1$^{fl/fl}$* mice.

(A) Total number of bone marrow and spleen cells after red blood cell lysis (x10^6) in *Nsd1$^{fl/fl}$* (n=7/BM, n=5/spleen, black bar) and *Vav1-iCre;Nsd1$^{fl/fl}$* mice (n=10/BM, n=9/spleen, red bar) (Students t-test with unequal variances. *** p<0.001). **(B)** Percentage of early and late apoptosis using Annexin V and 7-AAD flow cytometry staining in bone marrow and spleen of *Nsd1$^{fl/fl}$* (black bar) and *Vav1-iCre;Nsd1$^{fl/fl}$* mice (red bar) (n=3/group. Two- way ANOVA with Sidak's multiple comparisons; * p<0.05). **(C)** Total number of colonies formed by 40.000 bone marrow cells of *Nsd1$^{fl/fl}$* (black bar) and *Vav1-iCre;Nsd1$^{fl/fl}$* mice (red bar) in M3434 methylcellulose (n=9/group. Students t-test with equal variances). **(D)** Classification of different types of colonies formed by 40.000 bone marrow cells of *Nsd1$^{fl/fl}$* (black bar) and *Vav1-iCre;Nsd1$^{fl/fl}$* mice (red bar) in M3434 methylcellulose (n=9/group. Two-way ANOVA with Sidak's multiple comparisons. ** p<0.05, *** p<0.0001). **(E)** Representative images of colony morphology and content of *Vav1-iCre;Nsd1$^{fl/fl}$* bone marrow cells in M3434 methylcellulose demonstrating appearance of dishes including abnormal red colonies (I), benzidine- staining of dish (II), magnified abnormal colony (III) and cytospot of abnormal colony (IV). **(F)** Total number of colonies formed by 40.000 bone marrow cells of *Nsd1$^{fl/fl}$* (black bar) and *Vav1-iCre;Nsd1$^{fl/fl}$* mice (red bar) in M3434 methylcellulose after replating up to 4th round (n as indicated in figure). **(G)** Schematic depiction of replating experiment to test cells for dependence on erythropoietin (EPO). *Vav1-iCre;Nsd1$^{fl/fl}$* bone marrow cells were plated for four rounds in M3434 methylcellulose containing SCF, IL-3, IL-6 and EPO and then switched for the 5th plating to M3434, M3234 (+ 2ul/ml EPO) or M3534 (SCF, IL-3, IL-6). **(H)** Morphology and **(I)** quantification of total number of colonies formed in M3434, M3234, M3534 methylcellulose in 5th plating of *Vav1-iCre;Nsd1$^{fl/fl}$* bone marrow cells (n=1). Data presented as mean, error bars represent ±SD.

4.3.4 Loss of Nsd1 significantly reduces the HSC pool

To better characterize the hematopoietic system of diseased *Vav1-iCre;Nsd1$^{fl/fl}$* mice we analyzed the cellular hierarchy of the BM and spleen by flow cytometry using standard gating strategies (**Figure 16A-D**) [17]. We observed a striking reduction of lineage marker-negative, Lin$^-$/c-Kit$^+$/Sca-1$^+$ cells (LSKs) with almost a complete lack of long-term (LT-HSC) and short-term repopulating hematopoietic stem cell percentages (ST-HSC) in the BM of diseased *Vav1-iCre;Nsd1$^{fl/fl}$* (**Figure 16B, E-G**). Whereas the percentages of multi-potent progenitors (MPP) were slightly increased, granulocytic-macrophage progenitors (GMP) were reduced, and common myeloid (CMP) and megakaryocyte-erythroid progenitors (MEP) did not significantly differ from littermate controls. However, we found an increase of Lin$^-$/c-Kit$^+$/Sca-1$^-$ /CD34$^-$/FcγRII/III$^+$ cells (**Figure 16B, F-G**). To elucidate the myelo-erythroid hierarchy in more detail, we included endoglin (CD105) into our staining (**Figure 16C&D**) [18]. Lin$^-$ BM cells of diseased *Vav1-iCre;Nsd1$^{fl/fl}$* mice did not differ in CFU-E, pre-Meg-E and MkP cell fractions, whereas pre CFU- E were decreased and pre-GMs were increased (**Figure 16 H**). After the depletion of lineage-marker positive cells, we always obtained decreased Lin$^-$ cell numbers in BM of *Vav1-iCre;Nsd1$^{fl/fl}$* mice (data not shown). Based on this observation and flow cytometry data, we calculated the hypothetic number of stem and progenitor cells present in BM and spleen tissue. An example for the calculation is indicated in the methods section (Flow cytometry chapter). The total number of LSK was significantly decreased in BM and increased in the spleen of diseased *Vav1-iCre;Nsd1$^{fl/fl}$* mice. LT-HSC, ST-HSC, MPP and GMP numbers were significantly decreased whereas MEP, CMP and myelo- erythroid cell numbers remained unchanged (**Figure 16I-L).**

Collectively, these data revealed that aberrant accumulation of erythroid progenitors in diseased *Vav1-iCre;Nsd1$^{fl/fl}$* mice is associated with a striking reduction of the HSC pool and alterations of hematopoietic progenitor cells.


A
c-Kit
LK
LSK
Sca-1
CD34
CD34+
CD34-
PW
CD150
ST-HSC
MPP
CD48
CD150
LT-HSC
CD48
FcγRII/III
FcγRII/III+
GMP
MEP
CMP
CD34
B
Nsd1fl/fl
c-Kit
26.2%
2.75%
Sca-1
CD34
9.19%
PW
CD150
1.32%
85%
CD48
CD150
7.27%
CD48
FcγRII/III
2.15%
30.9%
39.3%
40.1%
CD34
Vav1-iCre;Nsd1fl/fl
c-Kit
38.1%
1.42%
Sca-1
CD34
12.6%
PW
CD150
0.09%
88.5%
CD48
CD150
0.15%
CD48
FcγRII/III
24.8%
16.7%
46.8%
24.8%
CD34
C
c-Kit
MkP
CD41-
CD41
FcγRII/III
FcγRII/III-
CD150
CD105
CFU-E
pre-CFU-E
pre-GM
pre-MegE
CD150
D
Nsd1fl/fl
c-Kit
2.55%
CD41
FcγRII/III
57.6%
CD150
CD105
18.9%
16.1%
33.5%
24.3%
CD150
Vav1-iCre;Nsd1fl/fl
c-Kit
0.20%
CD41
FcγRII/III
46.9%
CD150
CD105
10.9%
4.14%
60.0%
15.2%
CD150

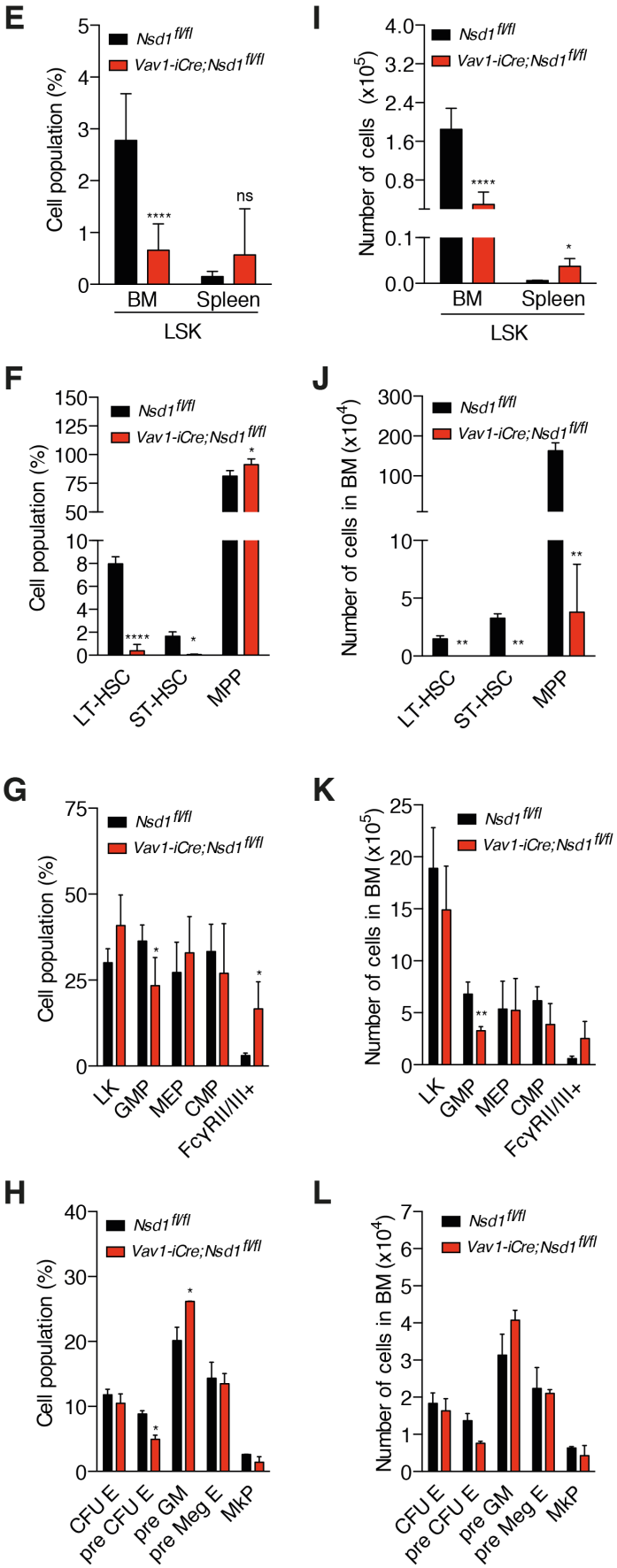

E
Nsd1 fl/fl
Vav1-iCre;Nsd1 fl/fl
Cell population (%)
****
ns
BM
Spleen
LSK
I
Number of cells (x10^5)
****
*
F
Cell population (%)
LT-HSC
ST-HSC
MPP
J
Number of cells in BM (x10^4)
G
Cell population (%)
LK
GMP
MEP
CMP
FcγRII/III+
K
Number of cells in BM (x10^5)
H
Cell population (%)
CFU E
pre CFU E
pre GM
pre Meg E
MkP
L
Number of cells in BM (x10^4)

FIGURE 16. immunephenotype of stem and progenitor cells in *Vav1-iCre;Nsd1$^{fl/fl}$* mice.

(A) Schematic depiction of flow cytometry gating strategy to quantify Lin$^-$/Sca-1$^+$/c-Kit$^+$ (LSK), Lin$^-$/Sca-1$^+$/c-Kit$^+$/CD34$^-$/CD150$^+$/CD48$^-$ (LT-HSC, long- term hematopoietic stem cells), Lin$^-$/Sca-1$^+$/c-Kit$^+$/CD34$^+$/CD150$^+$/CD48$^-$ (ST-HSC, short- term hematopoietic stem cells), Lin$^-$/Sca-1$^+$/c-Kit$^+$/CD34$^+$/CD150$^-$/CD48$^+$ (MPP, multipotent progenitor cells), Lin$^-$/Sca-1$^-$/c-Kit$^+$ (LK), Lin$^-$/Sca-1$^-$/c-Kit$^+$/CD34$^+$/FcγRII/III$^+$ (GMP, granulocyte macrophage progenitor), Lin$^-$/Sca-1$^-$/c-Kit$^+$/CD34$^-$/FcγRII/III$^-$ (MEP, megakaryocyte erythrocyte progenitor), Lin$^-$/Sca-1$^-$/c-Kit$^+$/CD34$^+$/FcγRII/III$^-$ (CMP, common myeloid progenitor) and Lin$^-$/Sca-1$^-$/c-Kit$^+$/CD34$^-$/FcγRII/III$^+$ ("FcγRII/III+") populations. **(B)** Representative image of flow cytometry panel demonstrating stem and progenitor staining in red blood cell lysed and lineage depleted single cell suspensions of bone marrow in *Nsd1$^{fl/fl}$* (upper row) and *Vav1-iCre;Nsd1$^{fl/fl}$* (lower row). **(C)** Schematic depiction of flow cytometry gating strategy to quantify myeloid progenitors C-Kit$^+$/CD41$^-$/FcγRII/III$^-$/CD150$^-$/CD105$^+$ (CFU-E), C-Kit$^+$/CD41$^-$/FcγRII/III$^-$/CD150$^+$/CD105$^+$ (pre CFU-E), C-Kit$^+$/CD41$^-$/FcγRII/III$^-$/CD150$^-$/CD105$^-$ (pre GM), C-Kit$^+$/CD41$^-$/FcγRII/III$^-$/CD150$^+$/CD105$^-$ (pre MegE) and C-Kit$^+$/CD41$^+$/CD150$^+$ (MkP) in bone marrow and spleen of *Nsd1$^{fl/fl}$* (black bar) and *Vav1-iCre;Nsd1$^{fl/fl}$* mice (red bar). **(D)** Representative image of flow cytometry panel demonstrating stem and progenitor staining in red blood cell lysed and lineage depleted single cell suspensions of bone marrow in *Nsd1$^{fl/fl}$* (upper row) and *Vav1-iCre;Nsd1$^{fl/fl}$* (lower row). **(E)** Percentage of LSK cell population in bone marrow and spleen of *Nsd1$^{fl/fl}$* (n=7/BM, n=3/spleen, black bar) and *Vav1-iCre;Nsd1$^{fl/fl}$* mice (n=7/BM, n=4/spleen, red bar) (Students t-test with equal variances. *** p<0.001). **(F)** Percentages of stem cells in bone marrow of *Nsd1$^{fl/fl}$* (n=3, black bar) and *Vav1-iCre;Nsd1$^{fl/fl}$* mice (n=4, red bar) (Students t-test with equal (LT-HSC) or unequal (ST-HSC, MPP) variances, * p<0.05, **** p<0.0001). **(G)** Percentages of progenitor cells in bone marrow of *Nsd1$^{fl/fl}$* (black bar) and *Vav1-iCre;Nsd1$^{fl/fl}$* mice (red bar) (n=4/group. Students t-test with equal (Myeloid, GMP, CMP, MEP) or unequal variances (FcγRII/III$^+$), * p<0.05). **(H)** Percentages of myeloid progenitors in bone marrow of *Nsd1$^{fl/fl}$* (black bar) and *Vav1-iCre;Nsd1$^{fl/fl}$* mice (red bar) (n=2/group. Students t-test with unequal variances. * p<0.05). **(I)** Number of LSK cells (x10^5) in red blood cell lysed and lineage depleted single cell suspensions of bone marrow and spleen of *Nsd1$^{fl/fl}$* (n=7/BM, n=3/spleen, black bar) and *Vav1-iCre;Nsd1$^{fl/fl}$* mice (n=7/BM, n=4/spleen, red bar) relative to total number of lineage depleted cells obtained during each procedure (Students t-test with equal variances, **** p<0.0001). **(J)** Number of stem cells (x10^4) in red blood cell lysed and lineage depleted single cell suspensions of bone marrow in *Nsd1$^{fl/fl}$* (n=3, black bar) and *Vav1-iCre;Nsd1$^{fl/fl}$* (n=4, red bar) relative to total number of lineage depleted cells obtained during each procedure (Students t-test with unequal variances, ** p<0.01). **(K)** Number of progenitor cells (x10^5) in bone marrow of *Nsd1$^{fl/fl}$* (black bar) and *Vav1-iCre;Nsd1$^{fl/fl}$* mice (red bar) (n=4/group. Students t-test with equal variances, ** p<0.01). **(L)** Number of myeloid progenitor cells (x10^4) in bone marrow of *Nsd1$^{fl/fl}$* (black bar) and *Vav1-iCre;Nsd1$^{fl/fl}$* mice (red bar) (n=2/group. Students t-test with unequal variances). Data presented as mean, error bars represent ±SD.

4.3.5 The disease of *Vav1-iCre;Nsd1$^{fl/fl}$* mice is transplantable

To explore the malignant potential of the observed phenotype, we transplanted total BM cells from diseased *Vav1-iCre;Nsd1$^{fl/fl}$* mice (carrying the isogenic CD45.2$^+$ marker) either alone or in competition with normal CD45.1$^+$ cells into lethally irradiated syngeneic recipients (CD45.1$^+$) (**Figure 17A**). All transplanted recipients rapidly developed symptoms of disease within 40 days post transplantation and showed hepato-splenomegaly, increased white blood counts, anemia, thrombocytopenia, mild lymphocytosis, and accumulation of LUC and reticulocytosis (**Figure 17B-I**). In addition, flow cytometry analysis of peripheral blood and BM revealed the accumulation of early erythroid progenitors found in R0 fraction (**Figure 17J-K**). Chimerism analysis of peripheral blood confirmed presence of injected CD45.2$^+$ cells in peripheral blood (**Figure 17L**).

We next assessed whether the disease remains transplantable in presence of increasing number of normal competitor cells, and whether it might be possible to further narrow down the disease-initiating cells based on c-Kit and CD71 expression (**Figure 18A**). First, we checked the colony forming capacity within the Lin$^-$/c-Kit$^+$ compartment and found that only CD71^{low} and CD71^{intermediate} expressing cells of *Vav1-iCre;Nsd1$^{fl/fl}$* mice had colony forming potential (**Figure 18B**). We next transplanted sorted Lin$^-$/c-Kit$^+$ and CD71 low/intermediate/high expressing cells or whole BM alone or in 1:10 competition (CD45.2$^+$) together with equal amounts of competitor whole BM (CD45.1$^+$) into lethally irradiated recipients (CD45.1$^+$) (**Figure 18C&D**). Kaplan- Meier survival curves confirmed development of disease upon transplanting whole knockout BM. Interestingly, all recipients transplanted with BM of diseased mice in 1:10 competition (WT:KO = 10:1) as well as mice receiving CD71^{low} cells (WT:KO CD71^{low} 10:1) developed disease. Moreover, 75% of CD71^{intermediate} transplanted mice (WT:KO CD71mid 10:1) and only 50% of mice that receiving CD71^{high} cells (WT:KO CD71high 10:1) developed disease symptoms with increasing time post- transplantation (**Figure 18E**). All diseased mice showed splenomegaly, increased white blood cell counts, decreased red blood cell counts, thrombocytopenia, variable amounts of reticulocytosis, some lymphocytosis as well as appearance of LUCs (**Figure 18F-K**, data not shown).

Taken together, these data show that disease in *Vav1-iCre;Nsd1$^{fl/fl}$* mice is transplantable into secondary recipients in competition and the disease initiating cell may be within the in the Lin$^-$/c-Kit$^+$ and most probably CD71$^{-/dim}$ compartment.

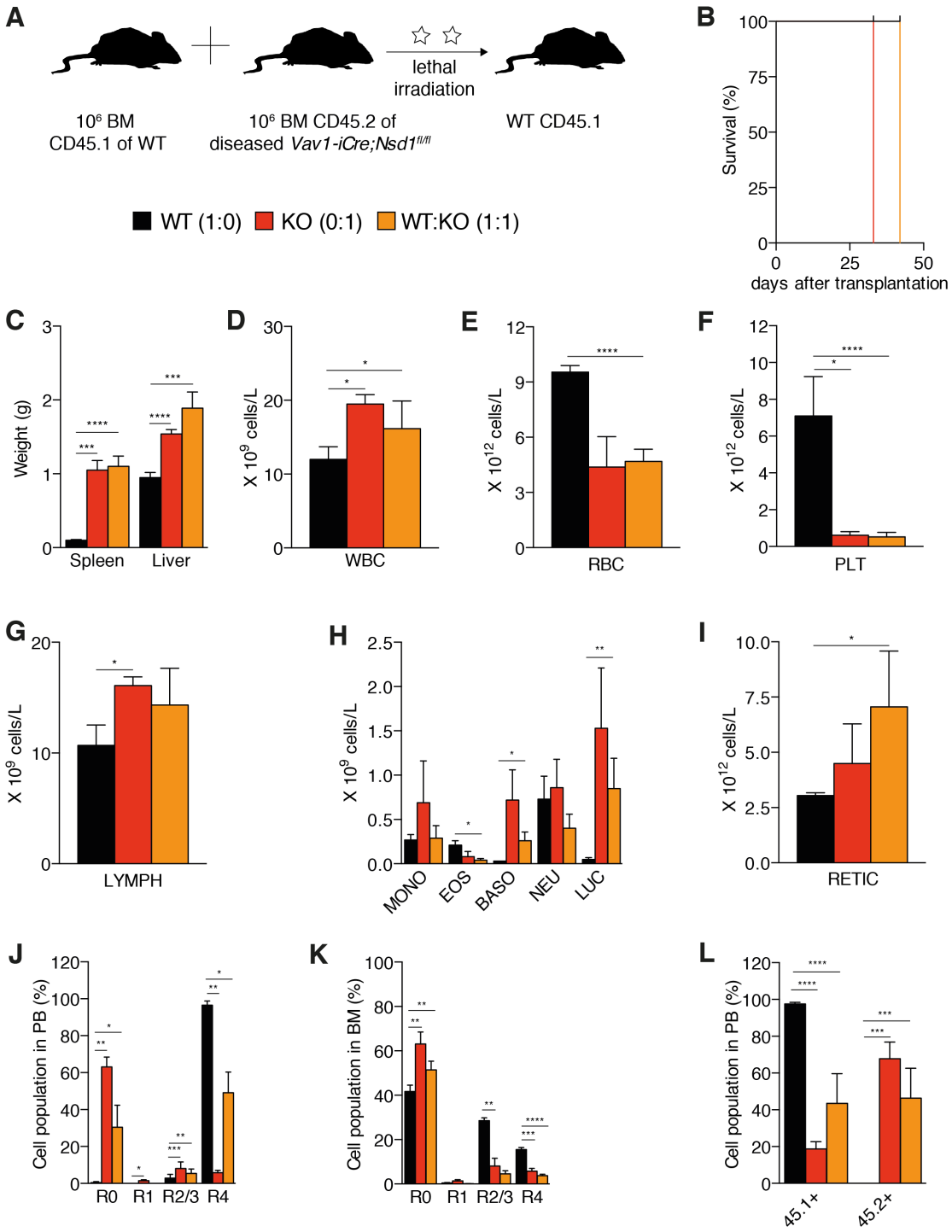
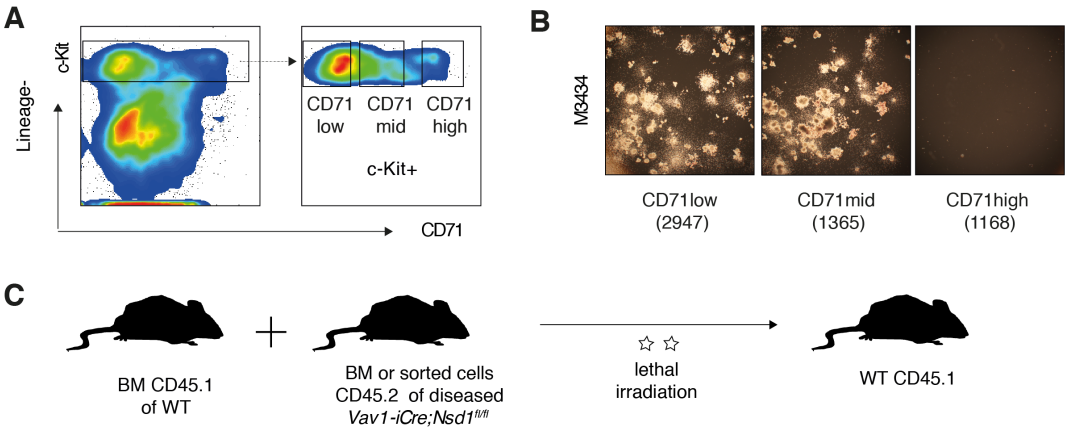

A
10⁶ BM
CD45.1 of WT
10⁶ BM CD45.2 of
diseased Vav1-iCre;Nsd1fl/fl
lethal
irradiation
WT CD45.1
WT (1:0) KO (0:1) WT:KO (1:1)
B
Survival (%)
days after transplantation
C
Weight (g)
Spleen
Liver
D
X 10⁹ cells/L
WBC
E
X 10¹² cells/L
RBC
F
X 10¹² cells/L
PLT
G
X 10⁹ cells/L
LYMPH
H
X 10⁹ cells/L
MONO
EOS
BASO
NEU
LUC
I
X 10¹² cells/L
RETIC
J
Cell population in PB (%)
R0
R1
R2/3
R4
K
Cell population in BM (%)
R0
R1
R2/3
R4
L
Cell population in PB (%)
45.1+
45.2+

FIGURE 17. Competitive BMT of whole BM cells of diseased V*av1-iCre;Nsd1$^{fl/fl}$* mice.

(A) Schematic depiction of competitive transplantation experimental setup. 10^6 CD45.2$^+$ whole bone marrow cells of *Nsd1$^{fl/fl}$* (n=4/"WT 1:0") or diseased *Vav1-iCre;Nsd1$^{fl/fl}$* mice (n=4/"KO 0:1") were transplanted without competition or 10^6 CD45.1$^+$ wildtype cells (n=6/"WT:KO 1:1") into lethally irradiated CD45.1$^+$ recipients. **(B)** Kaplan Meier survival curve of whole bone marrow transplanted mice. Black line represents WT 1:0 group that received bone marrow of *Nsd1$^{fl/fl}$*. Red line represents KO 0:1 group that received bone marrow of diseased *Vav1-iCre;Nsd1$^{fl/fl}$* mice. Orange Line represents competitive transplantation group that received bone marrow from control and diseased *Vav1-iCre;Nsd1$^{fl/fl}$* mice in a 1:1 ratio (WT:KO 1:1). Survival is shown as percentage at days after transplantation. Median survival of KO 0:1 transplanted mice was 33 days and 42 days of WT:KO 1:1 mice. WT 1:0 group was analyzed at 42 days. **(C)** Spleen and liver weight of transplanted mice in gram: WT 1:0 (n=2, black bar), KO 0:1 (n=4, red bar), WT:KO 1:1 (n=6, orange bar) (Students t-test with unequal variances. *** p<0.001, **** p<0.0001). **(D)** Peripheral blood counts of white blood cells, **(E)** red blood cells, **(F)** platelets, **(G)** lymphocytes, **(H)** monocytes (MONO), eosinophils (EOS), basophils (BASO), neutrophils (NEO), "large unstained cells" (LUC) and **(I)** reticulocytes in WT 1:0 (n=3, black bar), KO 0:1 (n=2, red bar), WT:KO 1:1 (n=5, orange bar) groups (Students t-test with unequal variances. * p<0.05, ** p<0.01, *** p<0.001, **** p<0.0001). **(J)** Percentage flow cytometry marker stained cell population in unlysed peripheral blood and **(K)** bone marrow at day of sacrifice in WT 1:0 (n=2, black bar), KO 0:1 (n=4, red bar), WT:KO 1:1 (n=6, orange bar) to distinguish maturation steps in erythropoiesis: R0 fraction (CD71$^{-/dim}$, Ter119$^-$), R1 (CD71$^+$, Ter119$^-$), R2 (CD71$^+$, Ter119$^+$), R3 (CD71$^{+/dim}$, Ter119$^+$) and R4 (CD71$^-$, Ter119$^+$) (Two-way ANOVA with Turkeys multiple comparisons. *** p<0.001, **** p<0.0001). **(L)** Percentage flow cytometry marker stained cell population in lysed single cell suspensions of peripheral blood of WT 1:0 (n=2, black bar), KO 0:1 (n=4, red bar), WT:KO 1:1 (n=6, orange bar) transplanted mice to distinguish donor contribution using CD45.2 surface marker (Two-way ANOVA with Turkeys multiple comparisons. * p<0.05, *** p<0.001, **** p<0.0001). Data presented as mean, error bars represent ±SD.



D

Group	Cells WT	Cells KO	Total cells	CD45.2 (%)	CD45.1 (%)
KO	500000	0	500000	100.0	0.0
WT:KO (10:1)	50000	450000	500000	10.0	90.0
WT:KO CD71low(10:1)	58952	500000	558952	10.5	89.5
WT:KO CD71mid(10:1)	27312	500000	527312	5.2	94.8
WT:KO CD71high(10:1)	23367	500000	523367	4.5	95.5

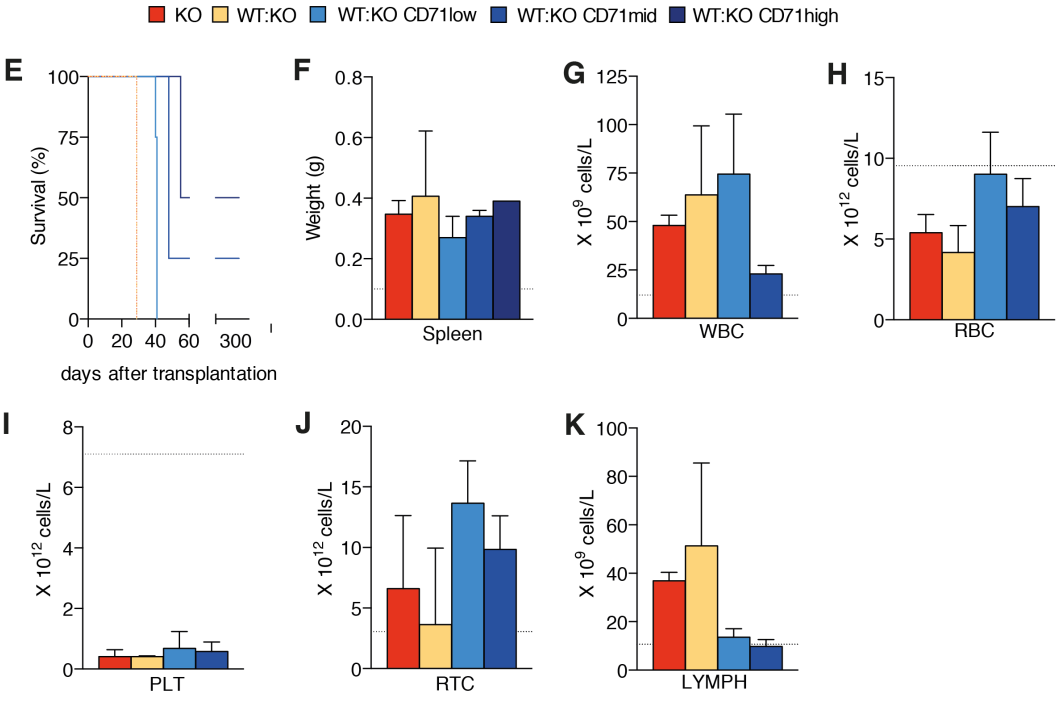


FIGURE 18. Competitive BMT of sorted Lin$^-$/c-Kit$^+$/CD71$^{low/mid/high}$ and whole BM cells of diseased *Vav1-iCre;Nsd1$^{fl/fl}$* mice.

(A) Gating strategy for fluorescence activated cell sorting (FACS). Lineage depleted Lin$^-$/c-Kit$^+$ cells were subfractionated using CD71 surface marker into low, mid and high expressing cells. **(B)** Representative images of sorted Lin$^-$/c-Kit$^+$/CD71$^{low/mid/high}$ cells plated into M3434 methylcellulose. Cell numbers plated are indicated in brackets. **(C)** Schematic depiction of experimental setup. CD45.1$^+$ wildtype cells were mixed with CD45.2$^+$ sorted or whole bone marrow cells of diseased *Vav1-iCre;Nsd1$^{fl/fl}$* mice and transplanted into lethally irradiated CD45.1$^+$ recipients. **(D)** Table containing information about numbers of transplanted cells of control and *Vav1-iCre;Nsd1$^{fl/fl}$* mice and calculated constitution of transplanted material. **(E)** Kaplan Meier survival curve of transplanted mice: KO 0:1 (n=3, red line), WT:KO 10:1 (n=3, yellow line), WT:KO CD71^{low} 10:1 (n=4, light blue line), WT:KO CD71^{mid} 10:1 (n=4, middle blue line) and WT:KO CD71^{high} 10:1 (n=2, dark blue line). Survival is shown as percentage at days after transplantation. Median survival of KO 0:1 and WT:KO 10:1 transplanted mice was 29 days, WT:KO CD71^{low} 10:1 41 days, WT:KO CD71^{mid} 48 days and WT:KO CD71^{high} 180 days. **(F)** Spleen weight of transplanted mice in gram: KO 0:1 (n=3, red bar), WT:KO 10:1 (n=3, yellow bar), WT:KO CD71^{low} 10:1 (n=3, light blue bar), WT:KO CD71^{mid} 10:1 (n=3, middle blue bar) and WT:KO CD71^{high} 10:1 (n=1, dark blue bar) groups. Dashed line indicates average spleen weight in control whole bone marrow transplanted mice from previous experiment. **(G)** Peripheral blood counts of white blood cells, **(H)** red blood cells, **(I)** platelets, **(J)** reticulocytes **(K)** lymphocytes in KO 0:1 (red bar), WT:KO 10:1 (yellow bar), WT:KO CD71^{low} 10:1 (light blue bar), WT:KO CD71^{mid} 10:1 (middle blue bar) and WT:KO CD71^{high} 10:1 (dark blue bar) groups (n=3/group). Dashed lines indicate average values in control whole bone marrow transplanted mice from previous experiment. Data presented as mean, error bars represent ±SD.

4.3.6 Loss of Nsd1 leads to early accumulation of erythroid progenitors

We next wondered whether loss of Nsd1 might affect the hematopoietic system of younger (4-6 weeks old) asymptomatic *Vav1-iCre;Nsd1$^{fl/fl}$* mice, as from now referred as “young” and diseased mice as “old”. Surprisingly, most young animals had already significantly enlarged spleens with abnormal focal cellular infiltrations in the liver and occasionally also in the lungs (**Figure 19A-D**). We also observed reduced white and red blood cell counts, thrombocytopenia, lower levels of lymphocytes and occasional appearance of LUCs in young *Vav1-iCre;Nsd1$^{fl/fl}$* mice (**Figure 19E-K**). Dissecting erythroid maturation revealed no significant changes in the BM, but increases of R1, R2, R3 fractions and CD71^{dim} populations with a significant decrease of the R4 fraction in the spleens of young mice (**Figure 19L&M**). BM cellularity was unchanged and no signs of apoptosis were identified (**Figure 20A&B**). Analysis of whole BM cells (after red cell lysis) revealed a significant expansion of CD71^{+}/c-Kit^{-} erythroid progenitor cells in spleen tissue (**Figure 20C&D**). Whereas the number of colonies formed by whole BM cells in methylcellulose was not altered, the number of CFU-G/M was significantly reduced, and abnormal “BFU-E-like” colonies appeared that could be occasionally replated over 2-3 rounds (**Figure 20E** & data not shown). Flow cytometry analysis of first round platings revealed that abnormal colonies consisted of CD71^{+} cells (**Figure 20F&G**).

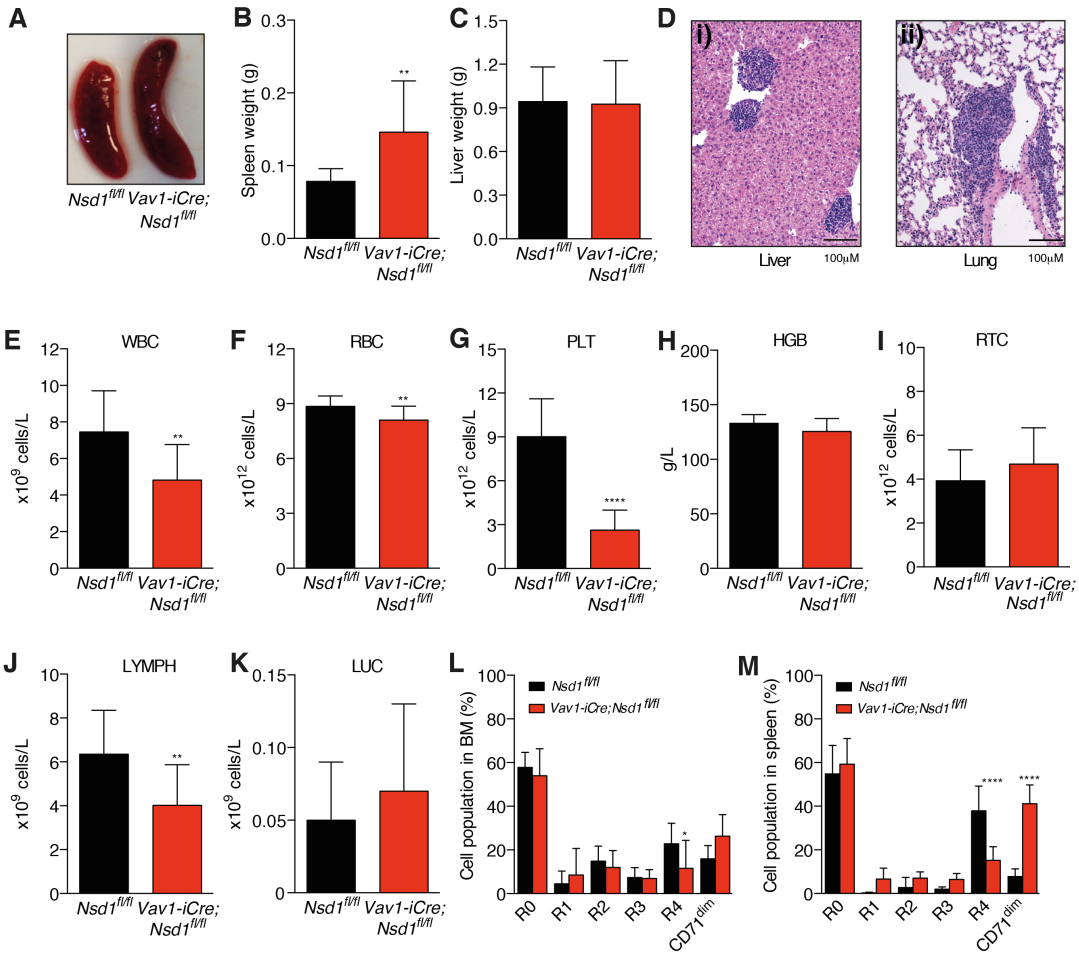

A
Nsd1fl/fl
Vav1-iCre; Nsd1fl/fl
B
Spleen weight (g)
0.0
0.1
0.2
0.3
**
Nsd1fl/fl
Vav1-iCre; Nsd1fl/fl
C
Liver weight (g)
0.0
0.3
0.6
0.9
1.2
1.5
Nsd1fl/fl
Vav1-iCre; Nsd1fl/fl
D
i)
Liver
100µM
ii)
Lung
100µM
E
WBC
x10^9 cells/L
0
2
4
6
8
10
12
**
Nsd1fl/fl
Vav1-iCre; Nsd1fl/fl
F
RBC
x10^12 cells/L
0
3
6
9
12
**
Nsd1fl/fl
Vav1-iCre; Nsd1fl/fl
G
PLT
x10^12 cells/L
0
3
6
9
12
15
****
Nsd1fl/fl
Vav1-iCre; Nsd1fl/fl
H
HGB
g/L
0
50
100
150
200
Nsd1fl/fl
Vav1-iCre; Nsd1fl/fl
I
RTC
x10^12 cells/L
0
2
4
6
8
10
Nsd1fl/fl
Vav1-iCre; Nsd1fl/fl
J
LYMPH
x10^9 cells/L
0
2
4
6
8
10
**
Nsd1fl/fl
Vav1-iCre; Nsd1fl/fl
K
LUC
x10^9 cells/L
0
0.05
0.10
0.15
Nsd1fl/fl
Vav1-iCre; Nsd1fl/fl
L
Cell population in BM (%)
0
20
40
60
80
100
Nsd1fl/fl
Vav1-iCre;Nsd1fl/fl
*
R0
R1
R2
R3
R4
CD71dim
M
Cell population in spleen (%)
0
20
40
60
80
100
Nsd1fl/fl
Vav1-iCre;Nsd1fl/fl
****
****
R0
R1
R2
R3
R4
CD71dim

FIGURE 19. Phenotype in 4-6 week old V*av1-iCre;Nsd1$^{fl/fl}$* mice.

(A) Representative picture of spleen of *Nsd1$^{fl/fl}$* (left) and young *Vav1-iCre;Nsd1$^{fl/fl}$* (right) mice. **(B)** Spleen weight and **(C)** liver weight of *Nsd1$^{fl/fl}$* (n=13, black bar) and young *Vav1-iCre;Nsd1$^{fl/fl}$* (n=17, red bar) mice in gram (Students t-test with unequal variances. ** p<0.01). **(D)** Representative images of i) spleen and ii) lung HE- stained tissue of young *Vav1-iCre;Nsd1$^{fl/fl}$* mice. **(E)** Peripheral blood counts of white blood cells, **(F)** red blood cells, **(G)** platelets, **(H)** hemoglobin, **(I)** reticulocytes, **(J)** lymphocytes and **(K)** "large unstained cells of *Nsd1$^{fl/fl}$* (n=13, black bar) and young *Vav1-iCre;Nsd1$^{fl/fl}$* (n=17, red bar) mice (Students t-test with (un)equal variances. ** p<0.01, **** p<0.0001). **(L)** Percentage flow cytometry marker stained cell population in unlysed single cell suspensions of bone marrow (n=9/group) and **(M)** spleen (n=8/group) of young *Nsd1$^{fl/fl}$* (black bar) and *Vav1-iCre;Nsd1$^{fl/fl}$* (red bar) mice (Two-way ANOVA with Sidak's multiple comparisons. * p<0.05, **** p<0.0001). Data presented as mean, error bars represent ±SD.

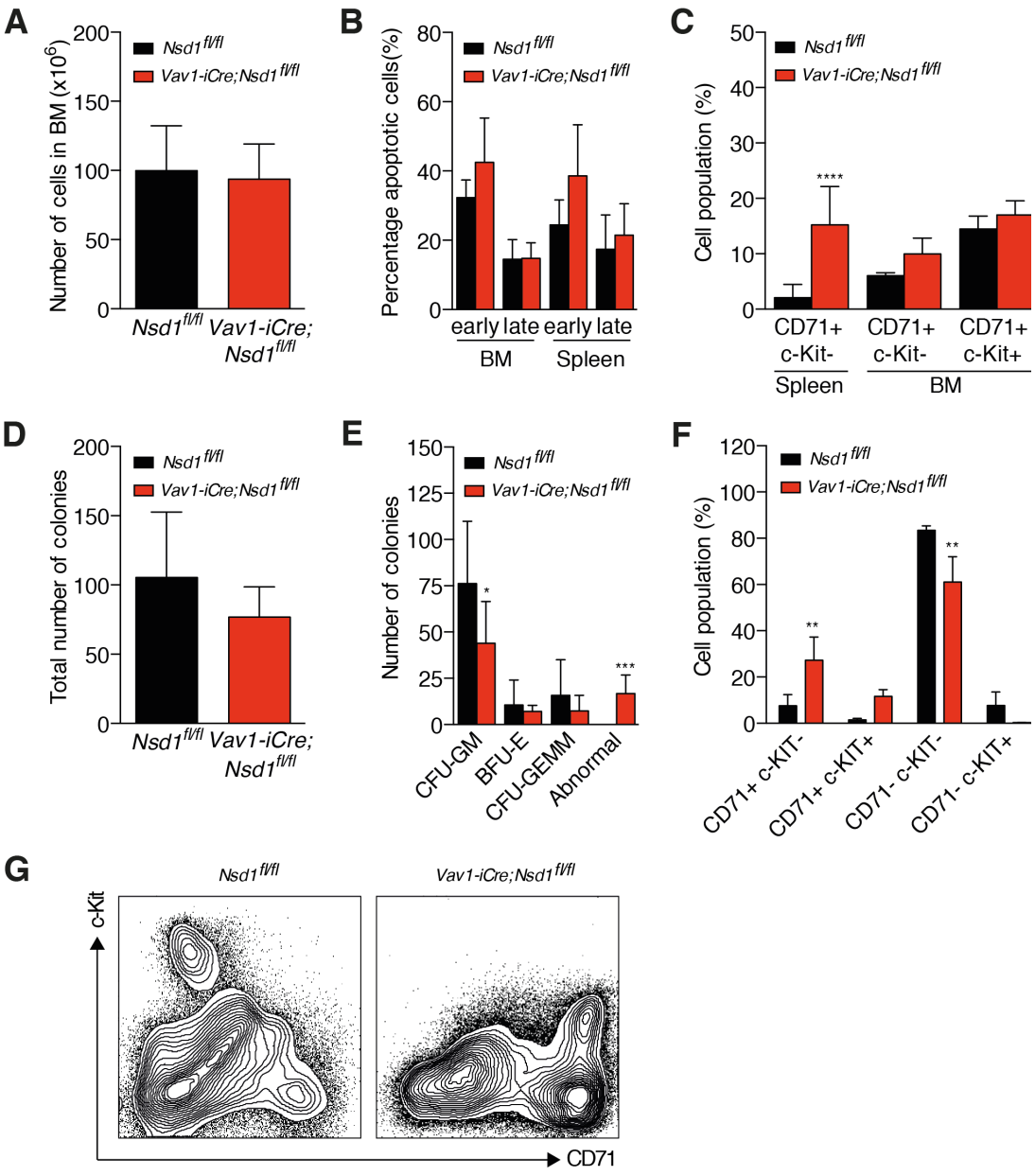
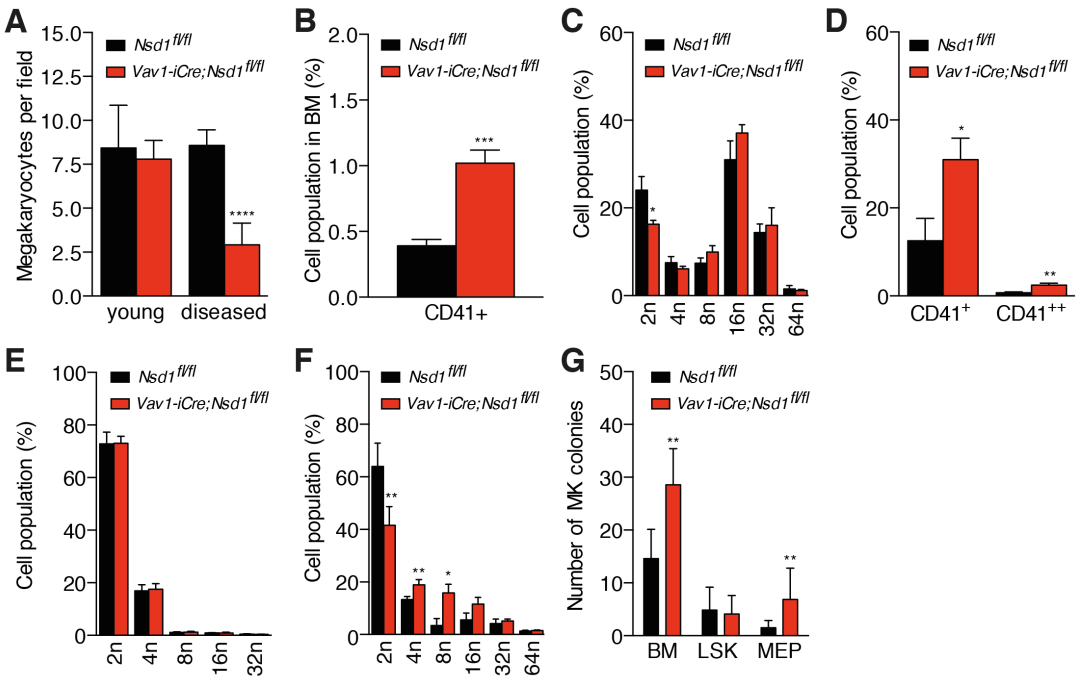

A
Number of cells in BM (x10^6)
200
150
100
50
0
Nsd1fl/fl
Vav1-iCre;Nsd1fl/fl
Nsd1fl/fl Vav1-iCre; Nsd1fl/fl
B
Percentage apoptotic cells(%)
80
60
40
20
0
Nsd1fl/fl
Vav1-iCre;Nsd1fl/fl
early late early late
BM Spleen
C
Cell population (%)
50
40
30
20
10
0
Nsd1fl/fl
Vav1-iCre;Nsd1fl/fl
****
CD71+ c-Kit-
CD71+ c-Kit-
CD71+ c-Kit+
Spleen
BM
D
Total number of colonies
200
150
100
50
0
Nsd1fl/fl
Vav1-iCre;Nsd1fl/fl
Nsd1fl/fl Vav1-iCre; Nsd1fl/fl
E
Number of colonies
150
125
100
75
50
25
0
Nsd1fl/fl
Vav1-iCre;Nsd1fl/fl
*
***
CFU-GM
BFU-E
CFU-GEMM
Abnormal
F
Cell population (%)
120
100
80
60
40
20
0
Nsd1fl/fl
Vav1-iCre;Nsd1fl/fl
**
**
CD71+ c-KIT-
CD71+ c-KIT+
CD71- c-KIT-
CD71- c-KIT+
G
Nsd1fl/fl
Vav1-iCre;Nsd1fl/fl
c-Kit
CD71

FIGURE 20. Immunophenotype of hematopoietic organs in 4-6 week old *Vav1-iCre;Nsd1$^{fl/fl}$* mice. (A) Total number of bone marrow cells (x10^6) after red blood cell lysis in young *Nsd1$^{fl/fl}$* (n=9, black bar) and *Vav1-iCre;Nsd1$^{fl/fl}$* mice (n=10, red bar) (Students t-test with equal variances). **(B)** Percentage of early and late apoptosis using Annexin V and 7-AAD flow cytometry staining in bone marrow and spleen of young *Nsd1$^{fl/fl}$* (n=5, black bar) and *Vav1-iCre;Nsd1$^{fl/fl}$* mice (n=6, red bar) (n=5-6/group. Two-way ANOVA with Sidak's multiple comparisons). **(C)** Percentage flow cytometry CD71 and c-Kit marker stained cell population in red blood cell lysed single cell suspensions of bone marrow and spleen of young *Nsd1$^{fl/fl}$* (n=5, black bar) and *Vav1-iCre;Nsd1$^{fl/fl}$* (n=6, red bar) mice (Two-way ANOVA with Sidak's multiple comparisons. **** p<0.0001). **(D)** Total number of colonies formed by bone marrow cells of young *Nsd1$^{fl/fl}$* (n=8, black bar) and *Vav1-iCre;Nsd1$^{fl/fl}$* mice (n=9, red bar) in M3434 methylcellulose (Students t-test with unequal variances). **(E)** Classification of different types of colonies formed by bone marrow cells of *Nsd1$^{fl/fl}$* (n=8, black bar) and *Vav1-iCre;Nsd1$^{fl/fl}$* mice (n=9, red bar) in M3434 methylcellulose (Two-way ANOVA with Sidak's multiple comparisons. * p<0.05, *** p<0.001). **(F)** Percentage flow cytometry marker CD71 and c-Kit stained population of single cell suspensions obtained after plating 40.000 whole bone marrow in methylcellulose M3434 of young *Nsd1$^{fl/fl}$* (black bar) and *Vav1-iCre;Nsd1$^{fl/fl}$* mice (red bar) (n=3/group. Two-way ANOVA with Sidak's multiple comparisons. ** p<0.01). **(G)** Representative flow cytometry panel of CD71 and c-Kit stained population of single cell suspensions obtained after plating 40.000 whole bone marrow in methylcellulose M3434 of young *Nsd1$^{fl/fl}$* (upper row) and *Vav1-iCre;Nsd1$^{fl/fl}$* mice (lower row). Data presented as mean, error bars represent ±SD.

4.3.7 Loss of Nsd1 leads to altered megakaryopoiesis

Significant thrombocytopenia seen in old and young *Vav1-iCre;Nsd1$^{fl/fl}$* mice led us to characterize the megakaryocytic lineage in more detail. We manually counted megakaryocytes on histological sections of HE stained BM biopsies that did not reveal changes in numbers in young mice, but a significant reduction in diseased mice (**Figure 21A**). Interestingly, we found a significant increase in CD41^{+} expressing cells of the megakaryocytic lineage with a reduced diploid (2N) DNA content in the BM of young *Vav1-iCre;Nsd1$^{fl/fl}$* mice (**Figure 21B-C**). In order to study formation of CD41 expressing cells *in vitro*, we grew BM cells for seven days in liquid medium containing megakaryocyte-favoring cytokines and then quantified the cellular output by flow cytometry. We found a significant increase of CD41^{+} and CD41^{++} cells after culturing young *Vav1-iCre;Nsd1$^{fl/fl}$* BM cells (**Figure 21D**). However CD41^{+} did not show any differences in nuclear ploidy, whereas CD41^{++} cells contained less 2N cells but increased fractions of cells with 4N and 8N (**Figure 21E&F**). To elucidate the colony forming potential, CFU-Mk colonies formed in Megacult^{TM} methylcellulose were quantified. Total BM cells and sorted MEP cells of young *Vav1-iCre;Nsd1$^{fl/fl}$* mice formed significantly more CFU-Mk colonies, whereas LSK cells formed equal numbers (**Figure 21G**). In conclusion, BM of young *Vav1-iCre;Nsd1$^{fl/fl}$* mice shows aberrant megakaryopoiesis that may result in the significant reduction of megakaryocytes in older diseased *Vav1-iCre;Nsd1$^{fl/fl}$* mice.



In collaboration with Cécile Thirant & Thomas Mercher. Gustave Roussy Institute. Paris

FIGURE 21. Megakaryocyte phenotype in *Vav1-iCre;Nsd1$^{fl/fl}$* mice.

(A) Number of manually counted megakaryocytes per field in HE- stained sections of bone marrow tissue of young (n=2) and diseased (n=4) *Nsd1$^{fl/fl}$* (black bar) and *Vav1-iCre;Nsd1$^{fl/fl}$* (red bar) mice. (n=20 fields per mouse. Two-way ANOVA with Sidak's multiple comparisons. *** p<0.001). **(B)** Percentage of flow cytometry marker stained CD41^{+} population in bone marrow cells of young *Nsd1$^{fl/fl}$* (black bar) and *Vav1-iCre;Nsd1$^{fl/fl}$* (red bar) mice. (n=3/group. Students t- test with equal variances. *** p<0.001). **(C)** Ploidy analysis on CD41^{+} population in bone marrow cells of young *Nsd1$^{fl/fl}$* (black bar) and *Vav1-iCre;Nsd1$^{fl/fl}$* (red bar) mice by propidium iodide (PI) staining. (n=3/group. Two-way ANOVA with Sidak's multiple comparisons. * p<0.05, ** p<0.01). **(D)** Percentage of flow cytometry marker stained CD41^{+} and CD41^{++} cells in bone marrow cells of young *Nsd1$^{fl/fl}$* (black bar) and *Vav1-iCre;Nsd1$^{fl/fl}$* (red bar) mice after seven days of megakaryocyte differentiation. (n=3/group. Two-way ANOVA with Sidak's multiple comparisons. *** p<0.001). **(E)** Ploidy analysis on CD41^{+} population in bone marrow cells of young *Nsd1$^{fl/fl}$* (black bar) and *Vav1-iCre;Nsd1$^{fl/fl}$* (red bar) mice by propidium iodide (PI) staining after seven days of megakaryocyte differentiation. (n=3/group. Two-way ANOVA with Sidak's multiple comparisons). **(F)** Ploidy analysis on CD41^{++} population in bone marrow cells of young *Nsd1$^{fl/fl}$* (black bar) and *Vav1-iCre;Nsd1$^{fl/fl}$* (red bar) mice by propidium iodide (PI) staining after seven days of megakaryocyte differentiation (n=3/group. Two-way ANOVA with Sidak's multiple comparisons. ** p<0.01; **** p<0.0001). **(G)** Total numbers of CFU-Mk colonies formed by total bone marrow cells, sorted LSK and MEP population of young *Nsd1$^{fl/fl}$* (black bar) and *Vav1-iCre;Nsd1$^{fl/fl}$* (red bar) mice. Colonies were counted after seven days in megacult assays. (n=3/group. Students t- test with equal variances. ** p<0.01). Data presented as mean, error bars represent ±SD.

4.3.8 Loss of Nsd1 leads to an early reduction and functional defect of hematopoietic stem cells

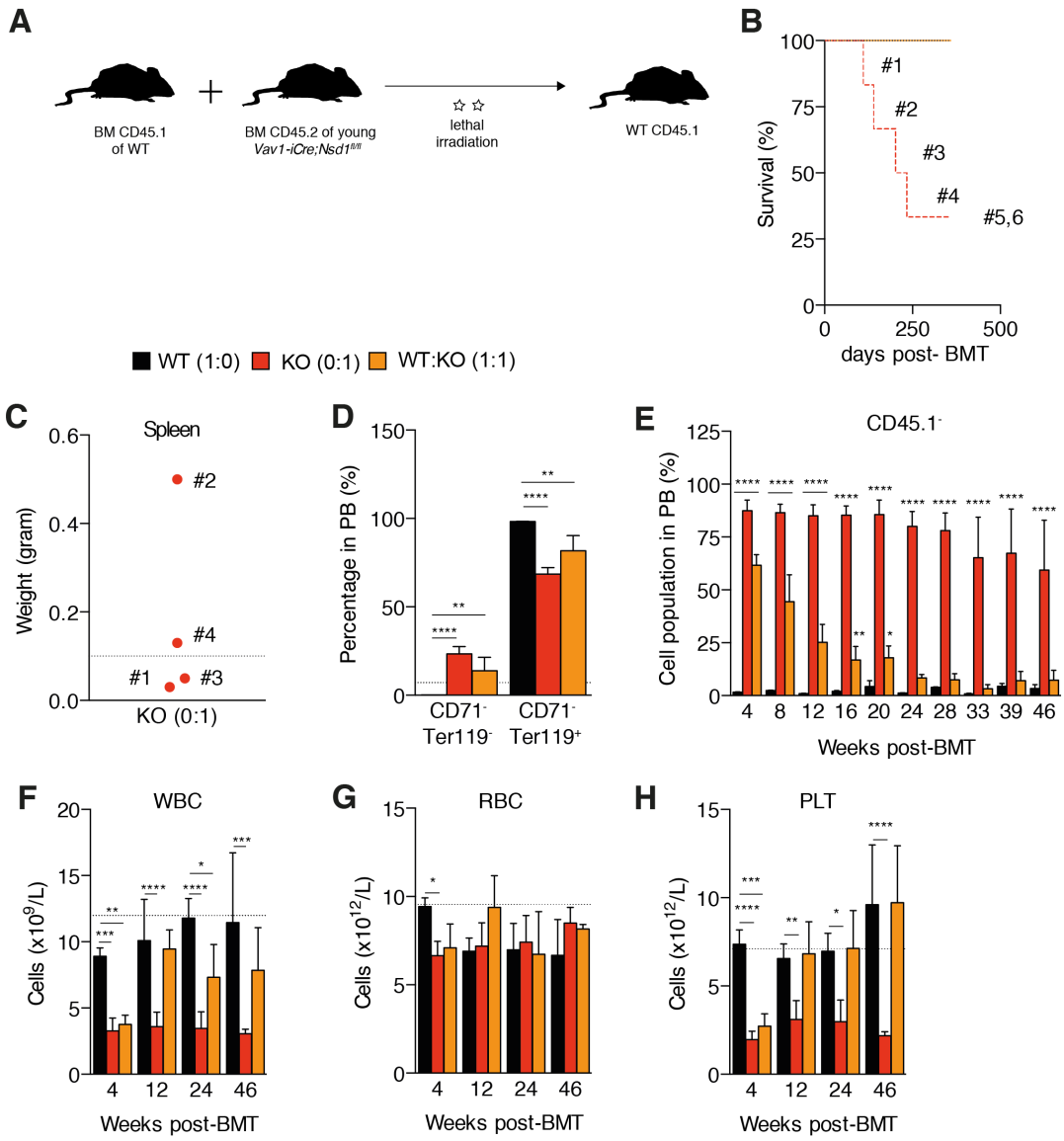
To functionally address the impact of the early changes on the BM function, we performed competitive reconstitution experiments. Hereby we transplanted whole BM cells of young *Vav1-iCre;Nsd1$^{fl/fl}$* mice (CD45.2^{+}) alone or in competition with CD45.1^{+} BM of donor mice into lethally irradiated recipients (CD45.1^{+}) (**Figure 22A**). The Kaplan Meier curve shows that 4 out of 6 mice had disadvantageous survival after more than 100 days post transplantation whereas all mice receiving competitor BM cells did not develop symptoms of disease up to one year post transplantation (**Figure 22B**). When the mice receiving only BM of young *Vav1-iCre;Nsd1$^{fl/fl}$* mice showed signs of distress were sacrificed, only 1 out 4 mice had an enlarged spleen (Mouse #2, **Figure 22C**). Interestingly, both transplanted groups showed significantly increased percentages of CD71^{-}/TER119^{-} and decreased percentages of CD71^{-}/TER119^{+} in peripheral blood 12 weeks post- transplantation compared to control BM transplanted mice (**Figure 22D**). Analysis of peripheral blood obtained by regular tail vein bleeding of transplanted mice revealed almost 100% CD45.1^{-} contribution in “KO (0:1)” mice until the end of the experiment. “WT:KO (1:1)” transplanted mice presented approximately 60% CD45.1^{-} contribution 4 weeks post transplantation, that decreased over time and reached non significant levels 24 weeks post transplantation (**Figure 22E**). White blood cell and platelet counts were significantly decreased in both groups 4 weeks post transplantation, remaining low in the KO (0:1) group and increased in WT:KO (1:1) group (**Figure 22F**) whereas red blood cell counts significantly decreased in both groups 4 weeks post transplantation and normalized to the end of the experiment (**Figure 22F-H**). Further information regarding values obtained from individual mice of KO (0:1) group can be found in **Figure 22I**.

Based on the observation of previous experiments in diseased mice revealing reduction in stem cells (**Figure 16**) as well as above mentioned “outcompetition” of 1:1 transplanted BM cells (**Figure 22E**), we aimed to dissect the hematopoietic hierarchy in BM of young *Vav1-iCre;Nsd1$^{fl/fl}$* mice. Interestingly, we observed a 60% reduction of LSK cell percentages mainly affecting ST-HSC and LT-HSC being more than 60- 90% reduced respectively (**Figure 23A&B**). MPP as well as myeloid progenitor cell percentages were less affected in young *Vav1-iCre;Nsd1$^{fl/fl}$* mice, however, FcγRII/III^{+} cell percentages were significantly increased (**Figure 23B-D**).

Total numbers of lineage depleted cells were not influenced at this time, which is reflected in the calculation- based quantification of stem and progenitor cells (data not shown). Total LSK cell number in BM was significantly reduced, in detail LT-HSC, ST-HSC and MPP cell numbers in BM were significantly reduced up to 90% (**Figure 23E&F**). Except FcγRII/III^{+} cells, myeloid progenitors were not reduced in numbers (**Figure 23G-H**).

FIGURE 22. Competitive BMT of whole BM cells of 4- 6 weeks aged *Vav1-iCre;Nsd1$^{fl/fl}$* mice.

(A) Schematic depiction of competitive transplantation experimental setup. 10^6 CD45.2$^+$ whole bone marrow cells of young *Nsd1$^{fl/fl}$* (WT 1:0) or *Vav1-iCre;Nsd1$^{fl/fl}$* mice (KO 0:1) were transplanted without competition or 10^6 CD45.1$^+$ wildtype cells (WT:KO 1:1) into lethally irradiated CD45.1$^+$ recipients. **(B)** Kaplan Meier survival curve of whole bone marrow transplanted mice. Black line represents WT 1:0 group that received bone marrow of *Nsd1$^{fl/fl}$* (n=3). Red line represents KO 0:1 group that received bone marrow of diseased *Vav1-iCre;Nsd1$^{fl/fl}$* mice (n=6). Orange Line represents competitive transplantation group that received bone marrow from control and diseased *Vav1-iCre;Nsd1$^{fl/fl}$* mice in a 1:1 ratio (n=6, WT:KO 1:1). Survival is shown as percentage at days after transplantation. Median survival of KO 0:1 transplanted mice was 218 days. WT 1:0 group was analyzed at 42 days. Individual mice are marked with numbers for later identification. **(C)** Spleen weight of KO 0:1 transplanted mice in gram (n=4). Individual mice are marked with numbers and normal spleen weight is indicated by dashed line. **(D)** Percentage flow cytometry marker CD71 and Ter119 stained cell population in unlysed single cell suspensions of peripheral blood 12 weeks after transplantation in WT 1:0 (n=3, black bar), KO 0:1 (n=6, red bar), WT:KO 1:1 (n=6, orange bar) (Two- way ANOVA with Turkey's multiple comparison test. ** p<0.01; **** p<0.0001). **(E)** Percentage flow cytometry marker CD45.1$^-$ stained cell population in red blood cell lysed peripheral blood at indicated weeks after transplantation to distinguish donor contribution using CD45.2 surface marker in WT 1:0 (black bar), KO 0:1 (red bar), WT:KO 1:1 (orange bar) (Two- way ANOVA with Turkeys multiple comparison test. * p<0.05; ** p<0.01; **** p<0.0001). **(F)** Peripheral blood counts of white blood cells, **(G)** red blood cells and **(H)** platelets in WT 1:0 (black bar), KO 0:1 (red bar), WT:KO 1:1 (orange bar) groups (n=4-6/group. Two- way ANOVA with Turkey's multiple comparison test. * p<0.05; ** p<0.01; *** p<0.001; **** p<0.0001). **(I)** Table containing information about transplanted mice of KO 0:1 group including survival time, spleen and liver weight, blood values and ability of bone marrow to form abnormal colonies in M3434. Data presented as mean, error bars represent ±SD.



I

Mouse # KO (0:1)	Survival (Days)	Spleen weight (g)	Liver weight (g)	WBC (x10^9 cells/L)	RBC (x10^12 cells/L)	PLT (x10^12 cells/L)	RETIC (x10^12 cells/L)	"abnormal colony formation"	Replating of colonies
1	110	0.03	0.48	45.06	9.00	2.64	6.72	YES	NO
2	140	0.50	2.00	340.53	5.58	2.34	9.76	NA	NA
3	202	0.05	0.80	5.67	9.09	1.35	6.31	YES	NO
4	234	0.13	1.00	19.65	7.98	3.36	4.65	YES	NO
5	359	NA	NA	3.30	9.12	2.04	6.10	NA	NA
6	359	NA	NA	2.82	7.86	2.34	3.89	NA	NA

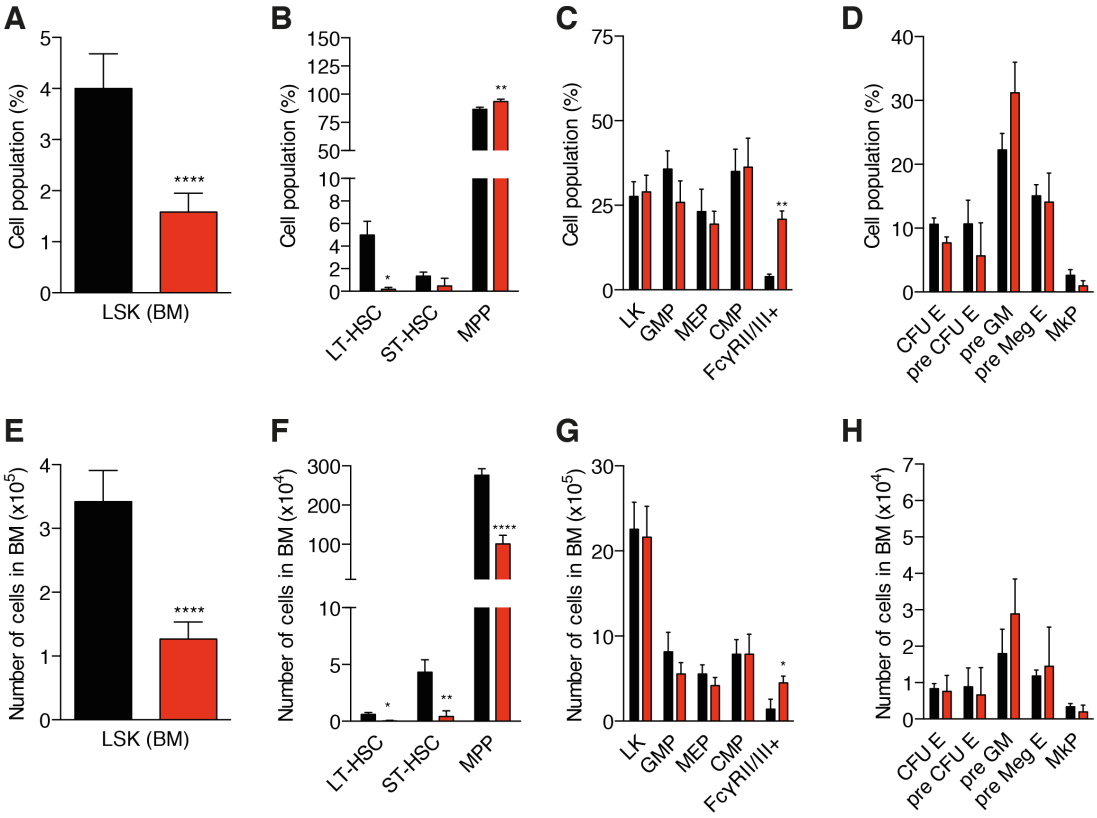

A
Cell population (%)
LSK (BM)
****
B
Cell population (%)
LT-HSC
ST-HSC
MPP
C
Cell population (%)
LK
GMP
MEP
CMP
FcγRII/III+
D
Cell population (%)
CFU E
pre CFU E
pre GM
pre Meg E
MkP
E
Number of cells in BM (x10^5)
LSK (BM)
F
Number of cells in BM (x10^4)
LT-HSC
ST-HSC
MPP
G
Number of cells in BM (x10^5)
LK
GMP
MEP
CMP
FcγRII/III+
H
Number of cells in BM (x10^4)
CFU E
pre CFU E
pre GM
pre Meg E
MkP

FIGURE 23. Immunephenotype of stem and progenitor cells of 4- 6 weeks aged *Vav1-iCre;Nsd1$^{fl/fl}$* mice.

(A) Percentage of Lin$^-$/Sca-1$^+$/c-Kit$^+$ (LSK) cell population in bone marrow of young *Nsd1$^{fl/fl}$* (n=7, black bar) and *Vav1-iCre;Nsd1$^{fl/fl}$* mice (n=8, red bar) (Students t-test with (un)equal variances. **** p<0.0001). **(B)** Percentages of Lin$^-$/Sca-1$^+$/c-Kit$^+$/CD34$^-$/CD150$^+$/CD48$^-$ (LT-HSC, long- term hematopoietic stem cells), Lin$^-$/Sca-1$^+$/c-Kit$^+$/CD34$^+$/CD150$^+$/CD48$^-$ (ST-HSC, short- term hematopoietic stem cells), Lin$^-$/Sca-1$^+$/c-Kit$^+$/CD34$^+$/CD150$^-$/CD48$^+$ (MPP, multipotent progenitor cells) in bone marrow of young *Nsd1$^{fl/fl}$* (n=3, black bar) and *Vav1-iCre;Nsd1$^{fl/fl}$* mice (n=4, red bar) (Unpaired T-test. ** p<0.01, *** p<0.001). **(C)** Percentages of progenitor cells Lin$^-$/Sca-1$^-$/c-Kit$^+$ (LK), Lin$^-$/Sca-1$^+$/c-Kit$^+$/CD34$^+$/FcγRII/III$^+$ (GMP, granulocyte macrophage progenitor), Lin$^-$/Sca-1$^+$/c-Kit$^+$/CD34$^-$/FcγRII/III$^-$ (MEP, megakaryocyte erythrocyte progenitor), Lin$^-$/Sca-1$^+$/c-Kit$^+$/CD34$^+$/FcγRII/III$^-$ (CMP, common myeloid progenitor) and Lin$^-$/Sca-1$^+$/c-Kit$^+$/CD34$^-$/FcγRII/III$^+$ (FcγRII/III+) in bone marrow of young *Nsd1$^{fl/fl}$* (black bar) and *Vav1-iCre;Nsd1$^{fl/fl}$* mice (red bar) (n=4/group. Two-way ANOVA with Sidak's multiple comparisons. ** p<0.01). **(D)** Percentages of myeloid progenitors c-Kit$^+$/CD41$^-$/FcγRII/III$^-$/CD150$^-$/CD105$^+$ (CFU-E), c-Kit$^+$/CD41$^-$/FcγRII/III$^-$/CD150$^+$/CD105$^+$ (pre CFU-E), c-Kit$^+$/CD41$^-$/FcγRII/III$^-$/CD150$^-$/CD105$^-$ (pre GM), c-Kit$^+$/CD41$^-$/FcγRII/III$^-$/CD150$^+$/CD105$^-$ (pre MegE) and c-Kit$^+$/CD41$^+$/CD150$^+$ (MkP) in bone marrow of young *Nsd1$^{fl/fl}$* (black bar) and *Vav1-iCre;Nsd1$^{fl/fl}$* mice (red bar) (n=2/group. Two-way ANOVA with Sidak's multiple comparisons). **(E)** Number of LSK cells (x10^5) in red blood cell lysed and lineage depleted single cell suspensions of bone marrow in young *Nsd1$^{fl/fl}$* (n=7, black bar) and *Vav1-iCre;Nsd1$^{fl/fl}$* (n=8, red bar) relative to total number of lineage depleted cells obtained during each procedure (Students t-test with equal variances. **** p<0.0001). **(F)** Number of stem cells (x10^4) in red blood cell lysed and lineage depleted single cell suspensions of bone marrow in young *Nsd1$^{fl/fl}$* (n=3, black bar) and *Vav1-iCre;Nsd1$^{fl/fl}$* (n=4, red bar) relative to total number of lineage depleted cells obtained during each procedure (Unpaired T-test. ** p<0.01, **** p<0.0001). **(G)** Number of progenitor cells (x10^5) in bone marrow of young *Nsd1$^{fl/fl}$* (black bar) and *Vav1-iCre;Nsd1$^{fl/fl}$* mice (red bar) (n=4/group. Two-way ANOVA with Sidak's multiple comparisons). **(H)** Number of myeloid progenitor cells (x10^4) in bone marrow of *Nsd1$^{fl/fl}$* (black bar) and *Vav1-iCre;Nsd1$^{fl/fl}$* mice (red bar) (n=2/group. Two-way ANOVA with Sidak's multiple comparisons). Data presented as mean, error bars represent ±SD.

We next aimed to study the functional colony forming capacity of flow- sorted LSK and MEP cells in methylcellulose. Total number of colonies did not differ between *Vav1-iCre;Nsd1$^{fl/fl}$* and *Nsd1$^{fl/fl}$* stem and progenitor cells (**Figure 24A**). However benzidine staining of plates revealed difference in colony morphology (**Figure 24B**). LSK and MEP cells formed denser, benzidine- positive colonies compared to control. Flow cytometry analysis of resuspended cells demonstrated shift from CD71$^-$/c-Kit$^-$ to CD71$^+$/c-Kit$^+$ expression (**Figure 24C&D**). Taking the *in vitro* data into account, we aimed to identify *in vivo* whether either LSK or MEP cells can reconstitute (partially) the BM of recipients and/or their presence influences development of differentiated cells in peripheral blood. We therefore injected 1000 ("WT or KO LSK 1000") or 6000 sorted LSK ("WT or KO LSK 6000") or 5000 sorted MEP cells ("WT or KO MEP 5000") of *Vav1-iCre;Nsd1$^{fl/fl}$* and *Nsd1$^{fl/fl}$* mice (CD45.2$^+$) with supporting BM (CD45.1$^+$) into lethally irradiated recipients marrow (CD45.1$^+$) and performed tail vein bleeding to follow their reconstitution ability (**Figure 24E**). Chimerism analysis in peripheral blood four weeks post transplantation revealed around 60- 80% reconstitution ability of *Vav1-iCre;Nsd1$^{fl/fl}$* LSK compared to 20% reconstitution ability of *Nsd1$^{fl/fl}$* LSK cells whereas MEP reconstitution ability did not differ between the groups. CD45.2$^+$ chimerism of *Vav1-iCre;Nsd1$^{fl/fl}$* LSK cells (1000 and 6000) dropped 12 weeks post- transplantation to levels of around 20-40% same as the control and remained stably decreased (**Figure 24F**). Interestingly, levels of TER119$^-$ population in peripheral blood were significantly higher in LSK transplanted recipients four weeks post transplantation and dropped after 12 weeks in KO LSK (1000) group and after 26 weeks in KO LSK (6000) group. KO MEP (6000) group did not show significant changes in formation of TER119$^-$ cells (**Figure 24G**). We next focused on the group receiving 1000 LSK cells and compared their CD45.2$^+$ chimerism in more detail. As mentioned before, *Vav1-iCre;Nsd1$^{fl/fl}$* LSK cells contributed to a significantly higher chimerism and TER119$^-$ cells during the first 12 weeks post transplantation (**Figure 24H-I**). In coincidence with the observed changes, white blood cells, platelets and red blood cells counts were significantly changed during this period (**Figure 24J-L**).

Taken together, our data demonstrate that absence of Nsd1 affects early hematopoiesis by inducing a partially transplantable disease phenotype and reduction in number and functional capacity of stem cells.

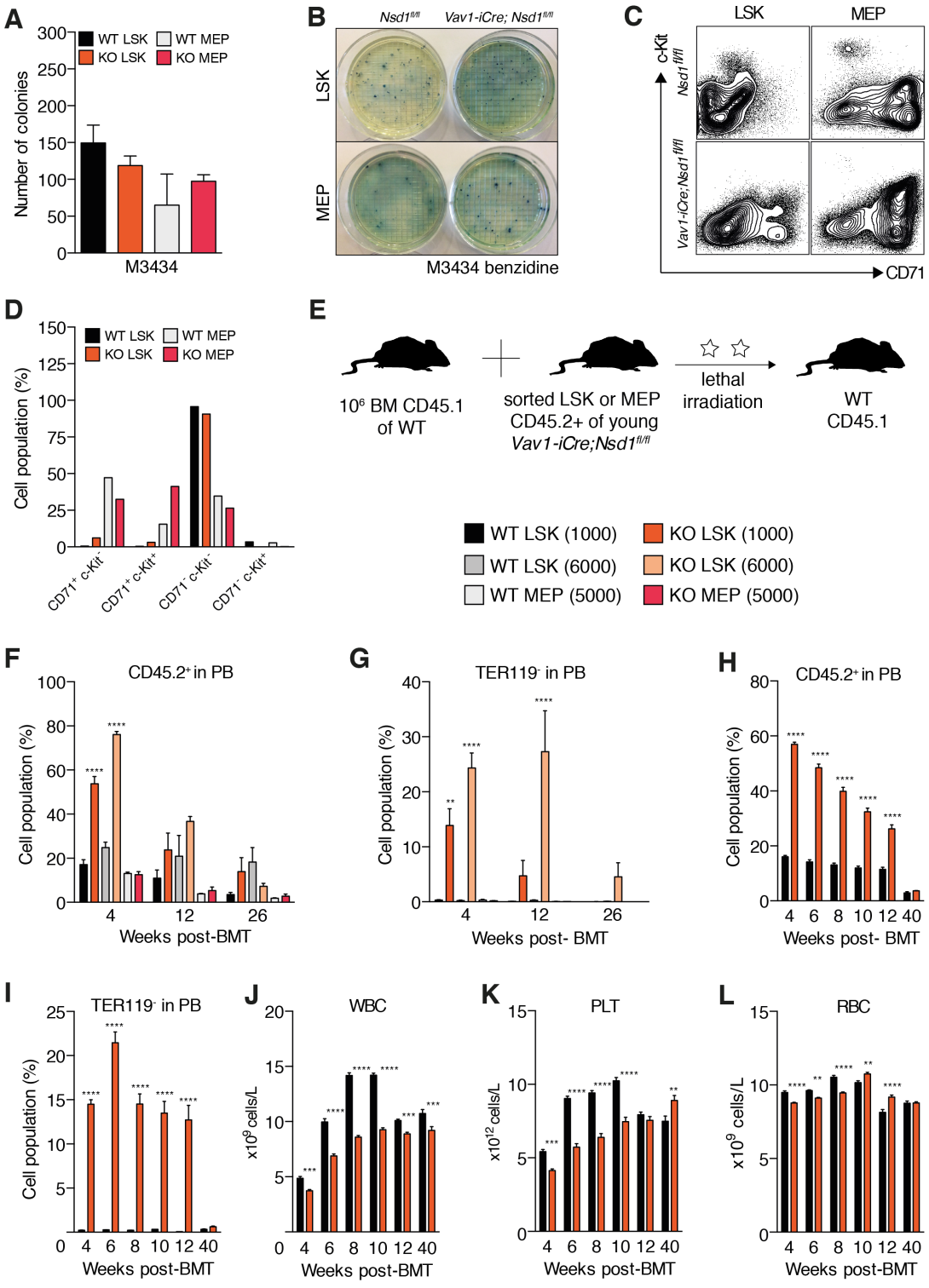

A
WT LSK
WT MEP
KO LSK
KO MEP
Number of colonies
M3434
B
Nsd1fl/fl
Vav1-iCre; Nsd1fl/fl
LSK
MEP
M3434 benzidine
C
LSK
MEP
c-Kit
Nsd1fl/fl
Vav1-iCre;Nsd1fl/fl
CD71
D
Cell population (%)
CD71+ c-Kit-
CD71+ c-Kit+
CD71- c-Kit-
CD71- c-Kit+
E
10⁶ BM CD45.1 of WT
sorted LSK or MEP CD45.2+ of young Vav1-iCre;Nsd1fl/fl
lethal irradiation
WT CD45.1
WT LSK (1000)
WT LSK (6000)
WT MEP (5000)
KO LSK (1000)
KO LSK (6000)
KO MEP (5000)
F
CD45.2+ in PB
Cell population (%)
Weeks post-BMT
G
TER119- in PB
Cell population (%)
Weeks post- BMT
H
CD45.2+ in PB
Cell population (%)
Weeks post- BMT
I
TER119- in PB
Cell population (%)
Weeks post-BMT
J
WBC
x10⁹ cells/L
Weeks post-BMT
K
PLT
x10¹² cells/L
Weeks post-BMT
L
RBC
x10⁹ cells/L
Weeks post-BMT

FIGURE 24. Colony formation and leukemic potential of LSK and MEP of 4- 6 week old *Vav1-iCre;Nsd1$^{fl/fl}$* mice.

(A) Total number of colonies formed by flow cytometry activated sorting of LSK and MEP obtained from red blood cell lysed and lineage depleted bone marrow cells of young *Nsd1$^{fl/fl}$* (black bar) and *Vav1-iCre;Nsd1$^{fl/fl}$* mice (red bar) in M3434 methylcellulose. 1000 LSKs and 5000 MEPs were plated (n=5/LSK, n=3/MEP. Students t-test with equal variances). **(B)** Representative images of colony morphology in M3434 methylcellulose after benzidine staining. 1000 LSKs (upper row) and 5000 MEPs (lower row) of young *Nsd1$^{fl/fl}$* (left column) and *Vav1-iCre;Nsd1$^{fl/fl}$* mice (right column) were plated. **(C)** Flow cytometry panel of CD71 and c-Kit stained population of single cell suspensions obtained after plating sorted LSK (left) or MEP (right) cells of young *Nsd1$^{fl/fl}$* (upper row) and *Vav1-iCre;Nsd1$^{fl/fl}$* mice (lower row) into M3434 methylcellulose. **(D)** Percentage flow cytometry marker CD71 and c-Kit stained population of single cell suspensions obtained after resuspension of cells out of methylcellulose M3434 of young *Nsd1$^{fl/fl}$* (black bar) and *Vav1-iCre;Nsd1$^{fl/fl}$* mice (red bar) (n=1). **(E)** Schematic depiction of experimental setup of competitive stem and progenitor transplantation. 1000 LSK, 6000 LSK or 5000 MEP of young CD45.2$^+$ *Nsd1$^{fl/fl}$* or *Vav1-iCre;Nsd1$^{fl/fl}$* mice were fluorescence activated cell sorted according to previously used markers and were transplanted with competition of 10^6 CD45.1$^+$ wildtype cells into lethally irradiated CD45.1$^+$ recipients. **(F)** Percentage flow cytometry marker stained cell population in lysed single cell suspensions of peripheral blood of WT LSK 1000 (black bar), KO LSK 1000 (orange bar), WT LSK 6000 (grey bar), KO LSK 6000 (light orange bar), WT MEP 5000 (white bar) and KO MEP 5000 (pink bar) transplanted mice at indicated weeks after transplantation to distinguish donor contribution using CD45.2 surface marker (n=3-6/group. Two- way ANOVA with Turkey's multiple comparison test. **** $p<0.0001$). **(G)** Percentage flow cytometry marker Ter119 stained cell population in unlysed single cell suspensions of peripheral blood of WT LSK 1000 (black bar), KO LSK 1000 (orange bar), WT LSK 6000 (grey bar), KO LSK 6000 (light orange bar), WT MEP 5000 (white bar) and KO MEP 5000 (pink bar) transplanted mice at indicated weeks after transplantation (n=3-6/group. Two-way ANOVA with Turkeys multiple comparison test. ** $p<0.01$; **** $p<0.0001$). **(H)** Percentage flow cytometry marker CD45 stained cell population in lysed single cell suspensions of peripheral blood of WT LSK 1000 (black bar), KO LSK 1000 (orange bar) transplanted mice at indicated weeks after transplantation to distinguish donor contribution using CD45.2 surface marker (n=2 independent experiments with n>4 per group. Two- way ANOVA with Turkey's multiple comparisons test. **** $p<0.0001$). **(I)** Percentage flow cytometry marker Ter119 stained cell population in unlysed single cell suspensions of peripheral blood of WT LSK 1000 (black bar), KO LSK 1000 (orange bar) transplanted mice at indicated weeks after transplantation (n=2 independent experiments with n>4 per group. Two-way ANOVA with Turkey's multiple comparisons test. **** $p<0.0001$). **(J)** Peripheral blood counts of white blood cells, **(K)** platelets and **(L)** red blood cells in WT LSK 1000 (black bar), KO LSK 1000 (orange bar) transplanted mice (n=2 independent experiments with n>4 per group. Two- way ANOVA with Turkey's multiple comparisons test. ** $p<0.01$, *** $p<0.001$ **** $p<0.0001$). Data presented as mean, error bars represent ±SD.

4.3.9 Conditional ablation of Nsd1 in adult mice results in mild aberrations of the erythro-megakaryocytic lineage

To address the effects of Nsd1 in adult hematopoiesis, we used two well-established transgenic Cre mouse lines that allow conditional gene ablation in the hematopoietic system. First we crossed *Nsd1*$^{fl/fl}$ mice to *Mx1-iCre* in which *Cre* expression is induced by poly(I:C) triggering an interferon response resulting in ablation of the gene in the liver, intestine and hematopoietic tissue. Adult (6-10 week old) mice obtained 10 poly(I:C) injections on 10 consecutive days with two days of break in between (**Figure 25A**). Nine weeks after induction we bled the mice and determined cre-mediated cleavage of the floxed allele by PCR in the peripheral blood. Interestingly, we only observed very limited cleavage, which seemed not to remain stable over time (**Figure 25B** & data not shown). We analyzed the mice 16- 20 weeks after first induction, a time point where *Vav1-iCre*-induced animals developed an extensive disease. In order to test whether the floxed *Nsd1 exon 5* was still present, we performed quantitative real time PCR and found reduced expression in the BM (**Figure 25C**). Flow cytometry analysis with CD71 and TER119 revealed accumulation of erythroid progenitor cells in peripheral blood (**Figure 25D**). Spleen and liver weights did not significantly differ between induced *Mx1-iCre;Nsd1*$^{fl/fl}$ and *Nsd1*$^{fl/fl}$ mice (**Figure 25E&F**). Slightly decreased white and red blood counts but a significantly reduced number of platelets were seen (**Figure 25G-I**). In addition we observed reduced neutrophils, other parameters like hemoglobin levels, lymphocytes, reticulocytes, LUCs, monocytes and eosinophils remained unchanged (**Figure 25J-N**). Histopathology did not revealed gross changes in the hematopoietic organs of induced mice, however in some mice we found small infiltrations of blast-like cells in liver (data not shown). To conclude, conditional cleavage of Nsd1 by Mx1-iCre in adult mice could be induced but caused only mild changes in the erythroid lineage and platelets. However, it was not sufficient to induce the severe disease observed in *Vav-iCre;Nsd1*$^{fl/fl}$ mice.

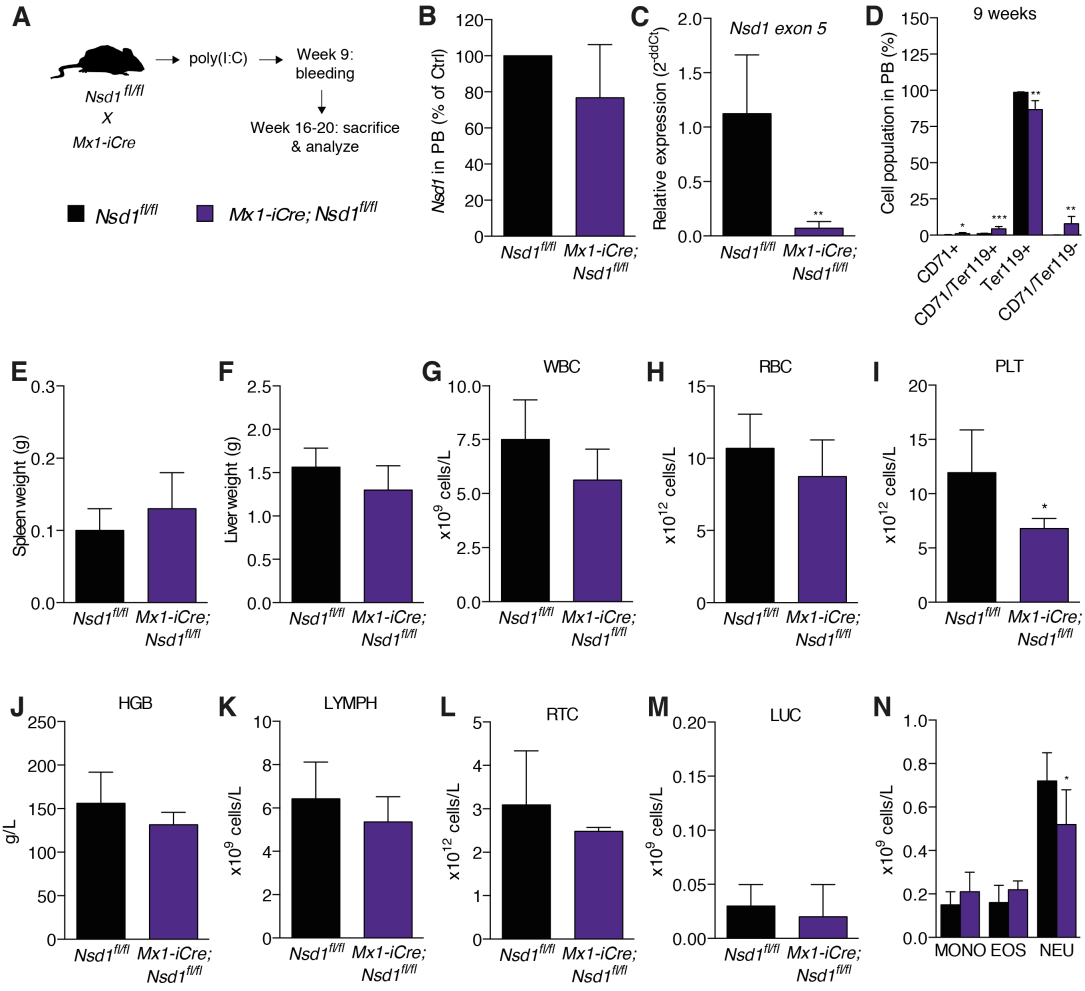

A
Nsd1 fl/fl
X
Mx1-iCre
poly(I:C)
Week 9: bleeding
Week 16-20: sacrifice & analyze
Nsd1fl/fl
Mx1-iCre; Nsd1fl/fl
B
Nsd1 in PB (% of Ctrl)
C
Nsd1 exon 5
Relative expression (2-ddCt)
D
9 weeks
Cell population in PB (%)
CD71+
CD71/Ter119+
Ter119+
CD71/Ter119-
E
Spleen weight (g)
F
Liver weight (g)
G
WBC
x10^9 cells/L
H
RBC
x10^12 cells/L
I
PLT
x10^12 cells/L
J
HGB
g/L
K
LYMPH
x10^9 cells/L
L
RTC
x10^12 cells/L
M
LUC
x10^9 cells/L
N
x10^9 cells/L
MONO
EOS
NEU

FIGURE 25. Phenotype in *Mx1-iCre;Nsd1$^{fl/fl}$* mice.

(A) Schematic depiction of the experimental setup using sex and age- matched *Nsd1$^{fl/fl}$* and *Mx1-iCre;Nsd1$^{fl/fl}$* mice that were injected at six to eight weeks after birth with ten times polyI:C (300ug/ml) intraperitoneal. Nine weeks after first injection, mice were tail vein bleeded and sixteen to twenty weeks later sacrificed and the bone marrow was analyzed. **(B)** Calculation of percentage of *Nsd1* cleavage in red blood cell lysed peripheral blood of polyI:C induced *Mx1-iCre;Nsd1$^{fl/fl}$* compared to their matched controls based on PCR. **(C)** Quantitative RT-PCR of bone marrow sixteen to twenty weeks after first injection of *Nsd1$^{fl/fl}$* (n=6, black bar) and *Mx1-iCre;Nsd1$^{fl/fl}$* mice (n=7, purple bar) indicating mRNA expression of *Nsd1 Exon 5*. Ct values were normalized to *Gapdh* expression and shown as relative expression using 1/dCt method (Students t-test with equal variances. **** p<0.0001). **(D)** Percentage flow cytometry marker CD71 and Ter119 stained cell population in unlysed single cell suspensions of peripheral blood (PB) nine to ten weeks after first injection (n=5 per group, black bar) and *Mx1-iCre;Nsd1$^{fl/fl}$* mice (n=6 per group, purple bar) (Two-way ANOVA with Sidak's multiple comparisons test. *** p<0.001, **** p<0.0001). **(E)** Spleen and **(F)** liver weight in gram of induced *Nsd1$^{fl/fl}$* (n=6, black bar) and *Mx1-iCre;Nsd1$^{fl/fl}$* mice (n=5, purple bar) sixteen to twenty weeks after first injection (Students t-test with equal variances). **(G)** Peripheral blood counts of white blood cells (WBC), **(H)** red blood cells (RBC), **(I)** platelets (PLT), **(J)** hemoglobin (HGB), **(K)** lymphocytes (LYMPH), **(L)** reticulocytes (RTC), **(M)** "large unstained cells" (LUC), **(N)** monocytes (MONO), eosinophils (EOS) and neutrophils (NEU) of induced *Nsd1$^{fl/fl}$* (n=6, black bar) and *Mx1-iCre;Nsd1$^{fl/fl}$* mice (n=5, purple bar) sixteen to twenty weeks after first injection (Students t-test with (un)equal variances. * p<0.05). Data presented as mean, error bars represent ±SD.

We also crossed the *Nsd1*$^{fl/fl}$ mice to *Scl-iCre*^{ER} ablators in which *Cre* is expressed in the hematopoietic tissue and endothelial cells and is activated by administration of tamoxifen. First, we tried to induce female mice that gave birth to *Scl-iCre*^{ER}*;Nsd1*$^{fl/fl}$ mice. Unfortunately, mice did not tolerate milking by tamoxifen- fed mother animals and did not survive (data not shown). We next induced adult mice by feeding them tamoxifen-impregnated chow pellets for 8-14 weeks (**Figure 26A** & data not shown). Nine weeks after induction we bled the mice and determined *Nsd1* cleavage by PCR in peripheral blood cells. We observed very variable cleavage between mice (**Figure 26B**). Like for *Mx1-iCre*, we analyzed the mice 16- 20 weeks after first induction, a time point where *Vav1-iCre*-induced animals developed an extensive disease. We found a striking reduction of *Nsd1 exon5* mRNA expression in BM cells (**Figure 26C**). Flow cytometry analysis of CD71 and TER119 revealed accumulation of erythroid progenitor cells in peripheral blood (**Figure 26D**). Spleen and liver weights did not significantly differ between induced *Scl-iCre*^{ER}*;Nsd1*$^{fl/fl}$ and *Nsd1*$^{fl/fl}$ controls (**Figure 26E&F**). We found slight decreases in white but not red blood counts and a significant reduction of platelets in induced mice (**Figure 26G-I).** Hemoglobin levels and lymphocytes were decreased whereas reticulocytes, LUCs, monocytes, eosinophils and neutrophil remained unchanged (**Figure 26J-N**). Histopathology did not reveal any significant alterations of the hematopoietic organs (data not shown).

To conclude, similar to ablation by *Mx1-iCre*, cleavage of *Nsd1* by *Scl-iCre*^{ER} in adult mice was efficient resulting in a very mild erythroid and platelet phenotype but no overt disease.

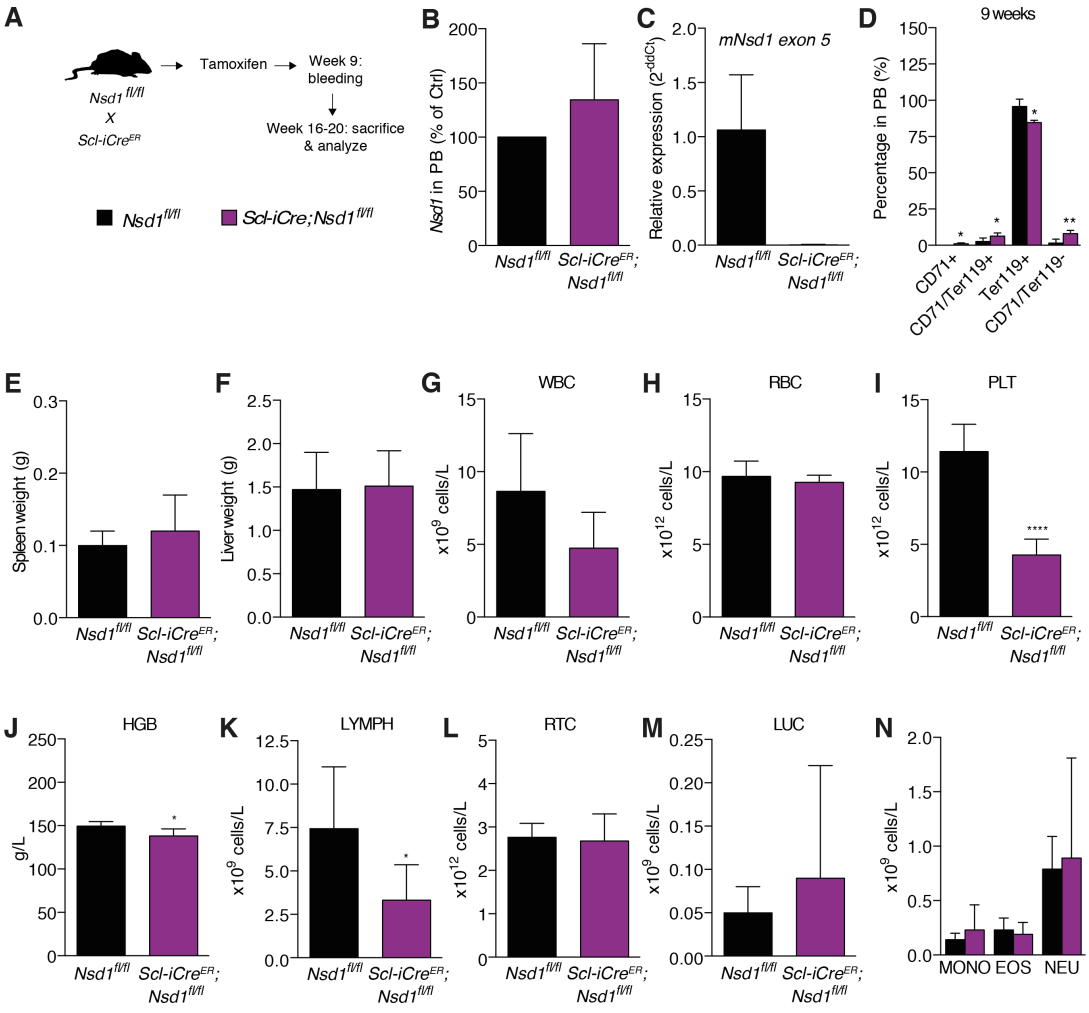

A
Nsd1 fl/fl
X
Scl-iCreER
Tamoxifen
Week 9: bleeding
Week 16-20: sacrifice & analyze
Nsd1fl/fl
Scl-iCre;Nsd1fl/fl
B
Nsd1 in PB (% of Ctrl)
C
mNsd1 exon 5
Relative expression (2-ddCt)
D
9 weeks
Percentage in PB (%)
CD71+
CD71/Ter119+
Ter119+
CD71/Ter119-
E
Spleen weight (g)
F
Liver weight (g)
G
WBC
x10^9 cells/L
H
RBC
x10^12 cells/L
I
PLT
x10^12 cells/L
J
HGB
g/L
K
LYMPH
x10^9 cells/L
L
RTC
x10^12 cells/L
M
LUC
x10^9 cells/L
N
x10^9 cells/L
MONO
EOS
NEU

FIGURE 26. Phenotype in *Scl-iCre^{ER};Nsd1$^{fl/fl}$* mice.

(A) Schematic depiction of the experimental setup using sex and age- matched *Nsd1$^{fl/fl}$* and *Scl-iCre^{ER};Nsd1$^{fl/fl}$* mice that were exposed to tamoxifen containing food for eight to forteen weeks at six to eight weeks after birth. Nine weeks after start of the experiment, mice were tail vein bleeded and thirteen to thirty weeks later sacrificed and the bone marrow was analyzed. **(B)** Calculation of percentage of *Nsd1* cleavage in red blood cell lysed peripheral blood of tamoxifen induced *Scl-iCre^{ER};Nsd1$^{fl/fl}$* compared to their matched controls based on PCR. **(C)** Quantitative RT-PCR of bone marrow fifteen to thirty weeks after first administration of tamoxifen containing food to *Nsd1$^{fl/fl}$* (n=2, black bar) and *Scl-iCre^{ER};Nsd1$^{fl/fl}$* mice (n=4, purple bar) indicating mRNA expression of *Nsd1 Exon 5*. Ct values were normalized to *Gapdh* expression and shown as relative expression using 1/dCt method (Students t-test with equal variances. ** $p<0.01$). **(D)** Percentage flow cytometry marker CD71 and Ter119 stained cell population in unlysed single cell suspensions of peripheral blood (PB) nine to eleven weeks after first administration of tamoxifen containing food to *Nsd1$^{fl/fl}$* (n=4, black bar) and *Scl-iCre^{ER};Nsd1$^{fl/fl}$* (n=5, purple bar) (Two-way ANOVA with Sidak's multiple comparisons test. * $p<0.05$, ** $p<0.01$). **(E)** Spleen and **(F)** liver weight in gram of induced *Nsd1$^{fl/fl}$* (n=5, black bar) and *Scl-iCre^{ER};Nsd1$^{fl/fl}$* mice (n=7, purple bar) thirteen to thirty weeks after first injection (Students t-test with equal variances). **(G)** Peripheral blood counts of white blood cells (WBC), **(H)** red blood cells (RBC), **(I)** platelets (PLT), **(J)** hemoglobin (HGB), **(K)** lymphocytes (LYMPH), **(L)** reticulocytes (RTC), **(M)** "large unstained cells" (LUC), **(N)** monocytes (MONO), eosinophils (EOS) and neutrophils (NEU) of induced *Nsd1$^{fl/fl}$* (n=5, black bar) and *Scl-iCre^{ER};Nsd1$^{fl/fl}$* mice (n=7, purple bar) thirteen to thirty weeks after first injection (Students t-test with (un)equal variances. * $p<0.05$, **** $p<0.0001$). Data presented as mean, error bars represent ±SD.

4.3.10 Loss of Nsd1 in the fetal liver leads to reduction of hematopoietic stem cells and accumulation of erythroid progenitors

As Vav1-promoter driven transgenes are expressed as early as E12.5 of fetal development we next characterized the fetal liver hematopoiesis of *Vav1-iCre;Nsd1*$^{fl/fl}$ mice at different days of development (**Figure 27A**) [206,207]. At E13.5, we found already significant cleavage of the floxed *Nsd1* alleles with a 95% reduced *Nsd1* mRNA expression (**Figure 27B&C**). However, the number of LSKs was only marginally reduced in fetal livers from *Vav1-iCre;Nsd1*$^{fl/fl}$ and *Nsd1*$^{fl/fl}$ littermate controls (**Figure 27D**). Nevertheless, at later time points (E16.5 and E19.5), the LSK compartment was clearly reduced in *Vav1-iCre;Nsd1*$^{fl/fl}$ animals (**Figure 27D**). Interestingly at E16.5 we could observe a significant increase of LK progenitor cells in fetal liver (**Figure 27E**). We also addressed erythroid maturation by determination of CD71 and TER119 expressing cells. We found no significant changes at E13.5 and E16.5, though at E19.5, a significant increase in early erythroid ("R0") progenitors was observed (**Figure 27F-H**). The total number of fetal liver cells remained unchanged whereas a significant increase was seen at E19.5 (**Figure 27I**). The colony forming ability of fetal liver cells generally decreased with later embryonic stages and did not display significant differences between *Vav1-iCre;Nsd1*$^{fl/fl}$ and Nsd1$^{fl/fl}$ fetal liver cells (**Figure 27J**). However, we observed some morphological differences with formation of benzidine-positive and compact colonies formed by cells from E19.5 *Vav1-iCre;Nsd1*$^{fl/fl}$ compared to littermate controls (**Figure 27K**). Histopathology revealed the appearance of clusters of large cells with a dark blue cytoplasm on sections from the fetal liver of *Vav1-iCre;Nsd1*$^{fl/fl}$ mice (**Figure 27L**).

These findings demonstrate that Vav-iCre-directed loss of *Nsd1* leads to reduction of LSKs and aberrant accumulation of erythroid progenitors during late fetal hematopoiesis.

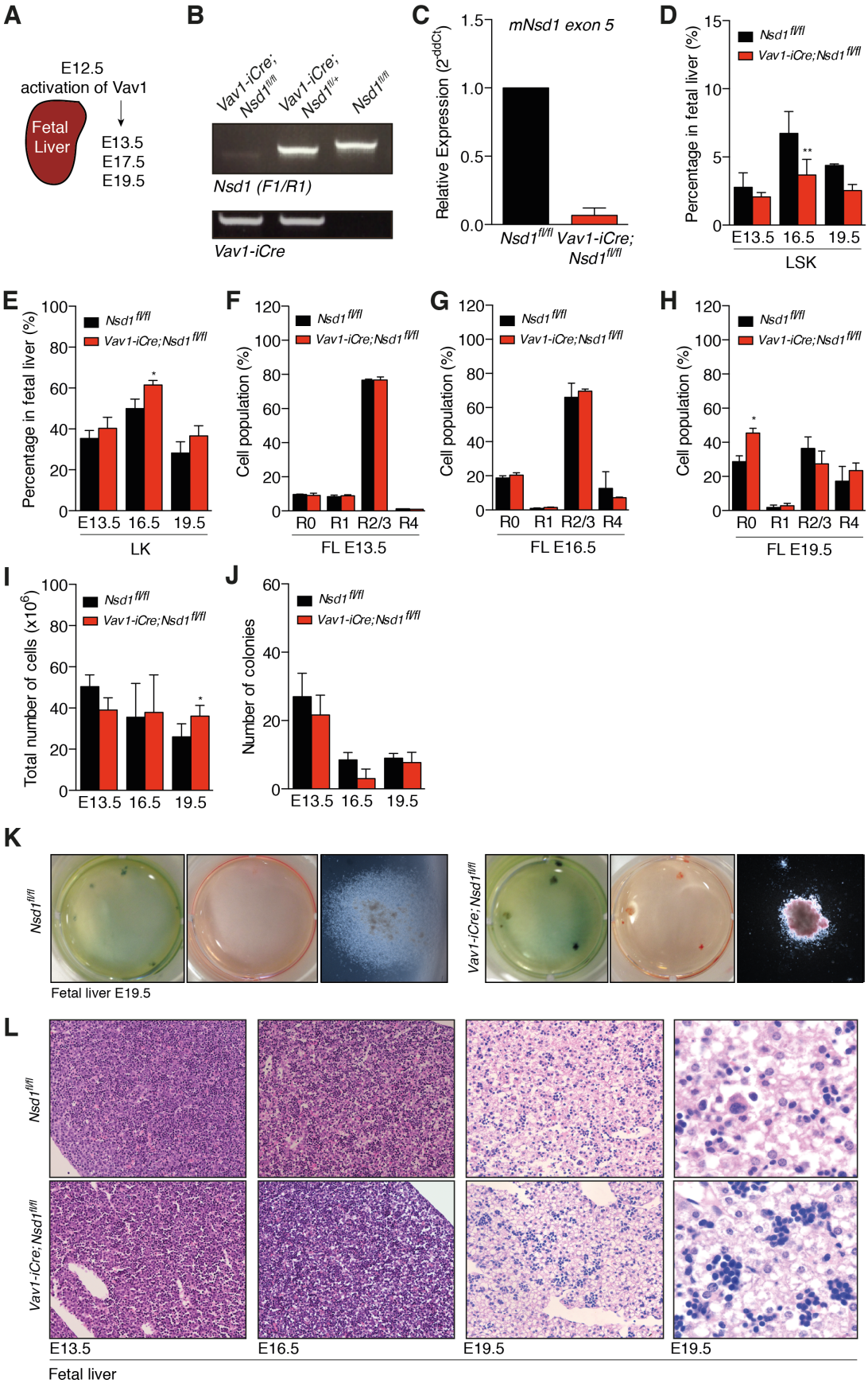

A
E12.5
activation of Vav1
Fetal Liver
E13.5
E17.5
E19.5
B
Vav1-iCre; Nsd1fl/fl
Vav1-iCre; Nsd1fl/+
Nsd1fl/fl
Nsd1 (F1/R1)
Vav1-iCre
C
mNsd1 exon 5
Relative Expression (2-ddCt)
1.5
1.0
0.5
0.0
Nsd1fl/fl
Vav1-iCre; Nsd1fl/fl
D
Nsd1fl/fl
Vav1-iCre;Nsd1fl/fl
Percentage in fetal liver (%)
15
10
5
0
**
E13.5 16.5 19.5
LSK
E
Nsd1fl/fl
Vav1-iCre;Nsd1fl/fl
Percentage in fetal liver (%)
100
80
60
40
20
0
*
E13.5 16.5 19.5
LK
F
Nsd1fl/fl
Vav1-iCre;Nsd1fl/fl
Cell population (%)
120
100
80
60
40
20
0
R0 R1 R2/3 R4
FL E13.5
G
Nsd1fl/fl
Vav1-iCre;Nsd1fl/fl
Cell population (%)
120
100
80
60
40
20
0
R0 R1 R2/3 R4
FL E16.5
H
Nsd1fl/fl
Vav1-iCre;Nsd1fl/fl
Cell population (%)
120
100
80
60
40
20
0
*
R0 R1 R2/3 R4
FL E19.5
I
Nsd1fl/fl
Vav1-iCre;Nsd1fl/fl
Total number of cells (x10^6)
100
80
60
40
20
0
*
E13.5 16.5 19.5
J
Nsd1fl/fl
Vav1-iCre;Nsd1fl/fl
Number of colonies
60
40
20
0
E13.5 16.5 19.5
K
Nsd1fl/fl
Vav1-iCre;Nsd1fl/fl
Fetal liver E19.5
L
Nsd1fl/fl
Vav1-iCre;Nsd1fl/fl
E13.5
E16.5
E19.5
E19.5
Fetal liver

FIGURE 27. CHARACTERIZATION OF PHENOTYPE IN FETAL LIVERS OF *VAV1-ICRE;NSD1$^{FL/FL}$* MICE. (A) Schematic depiction of experimental setup using fetal liver tissue. Timematings between Nsd1$^{fl/fl}$ and Vav1-iCre;Nsd1$^{fl/+}$ mice were setup and plugs controlled in the morning. Plug- positive mice were marked with E0.5 and followed for the next weeks. Pregnant mice were sacrificed at E13.5, 16.5 and 19.5 (day of birth) and fetal livers from pups dissected. **(B)** Cleavage PCR of fetal liver tissue at E13.5 in *Vav1-iCre;Nsd1$^{fl/fl}$*, *Vav1-iCre;Nsd1$^{fl/+}$* and *Nsd1$^{fl/fl}$* mice. **(C)** Verification of Nsd1 knockout in fetal liver samples from E13.5 in *Vav1-iCre;Nsd1$^{fl/fl}$* (n=2, red bar) and *Nsd1$^{fl/fl}$* (n=1, black bar) mice. Quantitative RT-PCR of *Nsd1 exon5/6* junction. Bars represent average relative expression normalized to *Gapdh* and control mice (Students t- test with equal variances. **** $p<0.0001$). **(D)** Percentage of Lin$^-$/Sca-1$^+$/c-Kit$^+$ (LSK) and **(E)** Lin$^-$/Sca-1$^-$/c-Kit$^+$ (LK) cell population in fetal liver samples of E13.5, 16.5 and 19.5 in *Nsd1$^{fl/fl}$* (n=3/E13.5, n=3/E16.5, n=2/E19.5, black bar) *Vav1-iCre;Nsd1$^{fl/fl}$* (n=5/E13.5, n=3/E16.5, n=4/E19.5, red bar) (Two-way ANOVA with Sidak's multiple comparisons test. * $p<0.05$). **(F)** Percentage flow cytometry marker stained cell population in unlysed fetal liver samples of E13.5, **(G)** E16.5 and **(H)** E19.5 in *Nsd1$^{fl/fl}$* (black bar) *Vav1-iCre;Nsd1$^{fl/fl}$* (red bar) to distinguish maturation steps in erythropoiesis: R0 fraction (CD71$^{-/dim}$, Ter119$^-$), R1 (CD71$^+$, Ter119$^-$), R2 (CD71$^+$, Ter119$^+$), R3 (CD71$^{+/dim}$, Ter119$^+$) and R4 (CD71$^-$, Ter119$^+$) (n=2/group. Two-way ANOVA with Sidak's multiple comparisons test. * $p<0.05$). **(I)** Total number of fetal liver cells in *Vav1-iCre;Nsd1$^{fl/fl}$* (red bar) and *Nsd1$^{fl/fl}$* (black bar) mice at E13.5 (n=2), E16.5 (n=5), E19.6 (n=4) (Two-way ANOVA with Sidak's multiple comparisons test. * $p<0.05$). **(J)** Total number of colonies formed in M3434 methylcellulose by 10.000 fetal liver cells derived from *Nsd1$^{fl/fl}$* (n=4/E13.5, n=2/E16.5, n=2/E19.5) and *Vav1-iCre;Nsd1$^{fl/fl}$* (n=6/E13.5, n=2/E16.5, n=4/E19.5) (Two-way ANOVA with Sidak's multiple comparisons test). **(K)** Representative images of colonies of *Nsd1$^{fl/fl}$* (left panels) and *Vav1-iCre;Nsd1$^{fl/fl}$* (right panels) fetal liver E19.5 cells in M3434 methylcellulose demonstrating morphology (right), appearance of dishes (middle) and benzidine- staining of dish (left). **(L)** Representative images of HE- stained fetal livers in *Nsd1$^{fl/fl}$* (upper row) *Vav1-iCre;Nsd1$^{fl/fl}$* (lower row) at different developmental stages ranging from E13.5, 16.5 to 19.5 at 20x magnification. Last column shows fetal livers at 60x magnification. Data presented as mean, error bars represent ±SD.

4.3.11 Loss of Nsd1 alters the gene expression signatures of hematopoietic stem and progenitor cells

As Nsd1 has been previously shown to have strong transcriptional activity, we next addressed the impact of Nsd1 loss on the gene expression program of hematopoietic stem and progenitor cells [183]. We sequenced total RNA from flow sorted LSK, MEP and GMP cells from three diseased *Vav-iCre;Nsd1$^{fl/fl}$* mice as well as LSK from young, asymptomatic mice (**Figure 28A**). Principal component analysis (PCA) revealed that LSKs from individual asymptomatic young (LSK "young") and diseased older (LSK "old") *Vav-iCre;Nsd1$^{fl/fl}$* mice shared a very similar gene expression profile. Moreover, PCA for the first two dimensions revealed that the expression differences were more based on the different cell types rather than by separation of the two genotypes (**Figure 28B**). In total we found 1918 genes differentially expressed in LSK, 2046 in MEP and 820 in GMP of *Vav-iCre;Nsd1$^{fl/fl}$* compared to littermate controls (**Figure 28C**). Differential gene expression analysis revealed no significant changes in LSK of "young" compared to "old" *Vav-iCre;Nsd1$^{fl/fl}$* mice. Therefore, LSK signatures were pooled together for further analysis (**Figure 28D**). Volcano plots reflect that 906 genes were UP- and 1012 DOWN-regulated in *Vav-iCre;Nsd1$^{fl/fl}$* LSK, 1426 genes UP- and 620 DOWN-regulated in *Vav-iCre;Nsd1$^{fl/fl}$* MEP and 421 genes UP- and 399 DOWN-regulated in *Vav-iCre;Nsd1$^{fl/fl}$* GMP (p-value<0.05) compared to littermate controls (**Figure 28E-G**). Interestingly, LSK cells lacking *Nsd1* expressed lower levels of multiple genes that are known to be associated with myelo-erythroid differentiation of hematopoietic cells including *Gata1*, a key erythroid transcription factor, as well as Erythropoietin receptor (*Epor*), and Friend of GATA protein (*Fog1*= *Zfpm1*). Furthermore, MEPs lacking *Nsd1* displayed an increase in *Spi1*, encoding for the hematopoietic transcription factor Pu.1 (**Figure 28H**).

Collectively, this data indicates that absence of Nsd1 significantly alters the gene expression profiles programs of LSK and MEP cells including activation of known oncogenes and reduced expression of known regulators of erythroid differentiation.

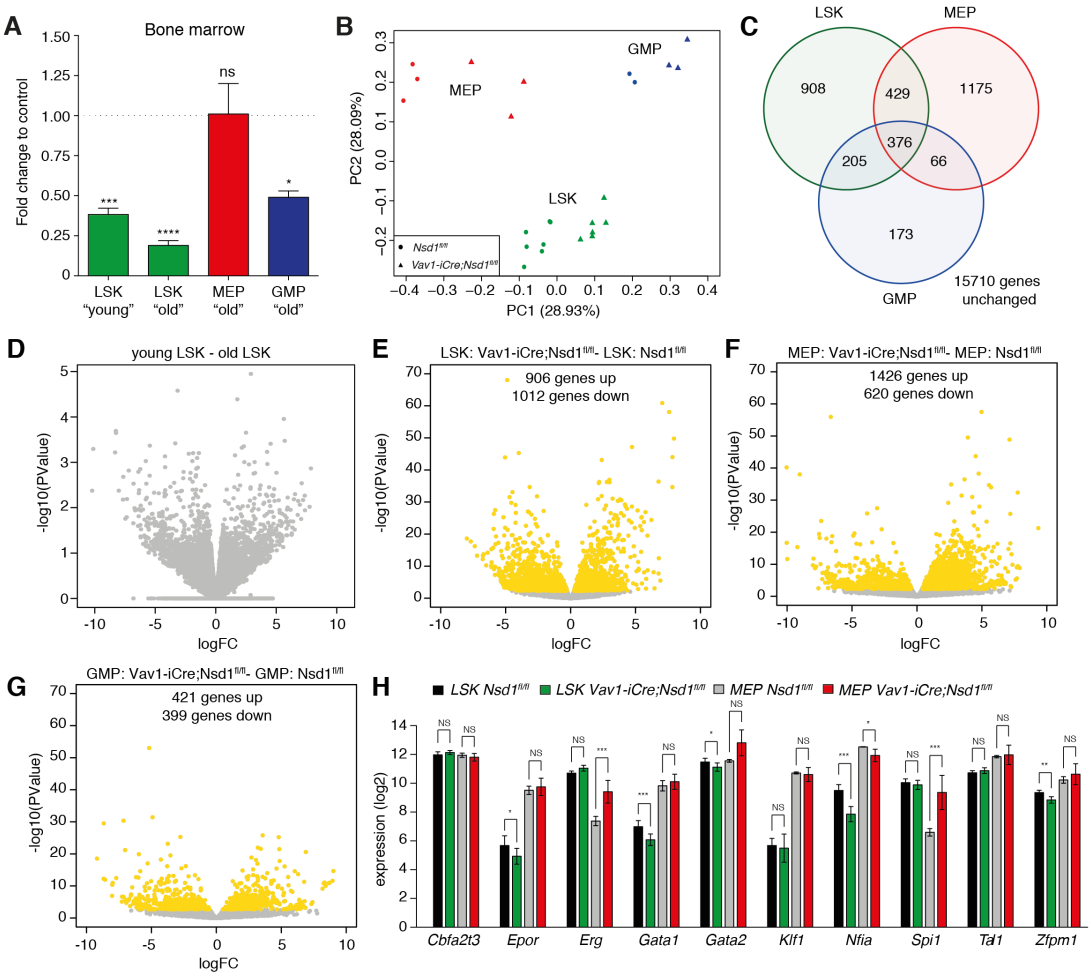

A
Bone marrow
Fold change to control
ns
***
****
*
LSK "young"
LSK "old"
MEP "old"
GMP "old"
B
GMP
MEP
LSK
PC2 (28.09%)
PC1 (28.93%)
Nsd1fl/fl
Vav1-iCre;Nsd1fl/fl
C
LSK
MEP
GMP
908
429
1175
376
205
66
173
15710 genes unchanged
D
young LSK - old LSK
-log10(PValue)
logFC
E
LSK: Vav1-iCre;Nsd1fl/fl- LSK: Nsd1fl/fl
906 genes up
1012 genes down
F
MEP: Vav1-iCre;Nsd1fl/fl- MEP: Nsd1fl/fl
1426 genes up
620 genes down
G
GMP: Vav1-iCre;Nsd1fl/fl- GMP: Nsd1fl/fl
421 genes up
399 genes down
H
LSK Nsd1fl/fl
LSK Vav1-iCre;Nsd1fl/fl
MEP Nsd1fl/fl
MEP Vav1-iCre;Nsd1fl/fl
expression (log2)
NS
Cbfa2t3
Epor
Erg
Gata1
Gata2
Klf1
Nfia
Spi1
Tal1
Zfpm1

FIGURE 28. RNA- Sequencing of sorted stem and progenitor cells of young and diseased *Vav1-iCre;Nsd1$^{fl/fl}$* mice.

(A) Calculated fold changes of LSK (n=7, green bar), MEP (n=4, red bar) and GMP (n=4, blue bar) cell populations in bone marrow of asymptomatic ("young") or diseased ("old") *Vav1-iCre;Nsd1$^{fl/fl}$* mice. Percentages were acquired using flow cytometry analysis described before and fold changes to their controls calculated manually (Students t-test with equal variances. * p<0.05, *** p<0.001, **** p<0.0001). **(B)** PCA plot of the first two components (PC1 versus PC2) indicating the variation between the three cells types, LSK (green), MEP (red) and GMP (blue), for *Nsd1$^{fl/fl}$* (round) and *Vav1-iCre;Nsd1$^{fl/fl}$* (triangle). Together, PC1 and PC2, explain 57.02% of the variance among cell types and conditions. **(C)** Venn Diagram showing the amount of differentially regulated genes (p-Value < 0.05) between for *Nsd1$^{fl/fl}$* and *Vav1-iCre;Nsd1$^{fl/fl}$* in the 3 cell types (colors as in A&B). The number of shared differentially expressed (determined from edgeR) among the cell types are shown in the intersections of the circles. **(D)** Volcano plot of the p values (-log10 scale) versus the fold change (FC, log2 scale) of differentially up or down regulated genes in „young" versus „old" LSKs and comparisons between *Vav1-iCre;Nsd1$^{fl/fl}$* - *Nsd1$^{fl/fl}$* in **(E)** „old" LSKs **(F)** „old" MEPs and **(G)** „old" GMPs. Differentially expressed genes (p < 0.05) are marked in yellow. **(H)** Bar plot of selected transcription factors for *Nsd1$^{fl/fl}$* and *Vav1-iCre;Nsd1$^{fl/fl}$* in LSK (resp. black and green) and MEP cells (resp. grey and red). Data presented as mean, error bars represent ±SD.

4.3.12 Loss of Nsd1 impairs differentiation of extensively self-renewing erythroblasts constitutively expressing high levels of GATA1 protein levels

Extensively self-renewing erythroblasts (ESREs) provide a non-transformed cellular model of erythroid maturation derived from murine fetal liver or yolk sac, and morphologically, immunophenotypically, and functionally resemble proerythroblasts that maintain both cytokine dependence and the potential to generate enucleated erythrocytes after 3-4 cell divisions [203]. To address the role of Nsd1 in erythroid maturation we extracted ESREs from *Vav1-iCre;Nsd1$^{fl/fl}$* and littermate control E17.5 fetal livers and expanded the cells in so called "maintenance medium" (MM) containing dexamethasone, Insulin-like growth factor 1 (IGF1), cholesterol, SCF and EPO for more than six days to obtain stable cell lines. We then transferred the cells into "differentiation medium" (DM) containing SCF and EPO and measured erythroid differentiation-associated surface markers and the cell size by flow cytometry, and visualized the cellular morphology on cytospots (**Figure 29A**). After 3 days in DM, cells were significantly reduced in size (FSC^{-}) and already fully mature erythrocytes were seen (**Figure 29B**). Control cells at this timepoint displayed signs of differentiation, whereas the cells lacking Nsd1 appeared mostly as proerythroblasts (**Figure 29C**). After 6 days in DM, control cells completely matured towards erythrocytes. In contrast, although exhibiting an early proliferation burst, *Vav1-iCre;Nsd1$^{fl/fl}$* ESREs were significantly impaired to differentiate and exhibited disturbed differentiation still present at day 13 in liquid differentiation culture (**Figure 29C&D**). However, important to note, very few cells lacking *Nsd1* seemed to be able to fully mature towards mature erythrocytes. Impaired maturation was confirmed by significantly decreased percentages of CD71^{-}/ FSC^{-} cells (**Figure 29E** & data not shown). *Vav1-iCre;Nsd1$^{fl/fl}$* and littermate control "ESREs" did not significantly differ in proliferation capacity, expression of erythroid surface markers and morphology in MM before induction of differentiation (data not shown). However, in DM, *Vav1-iCre;Nsd1$^{fl/fl}$* "ESREs" showed significantly increased proliferation rates (**Figure 29F**). Terminal erythroid maturation is controlled by dynamic regulation of the erythroid transcription factor Gata1 [96]. In contrast to littermate controls, we observed reduced *Gata1* mRNA expression during differentiation of *Vav1-iCre;Nsd1$^{fl/fl}$* ESREs (**Figure 29G**). However, rather surprisingly we observed significantly increased GATA1 protein levels in *Vav1-iCre;Nsd1$^{fl/fl}$* fetal liver-derived ESREs in maintenance medium (**Figure 29H**).

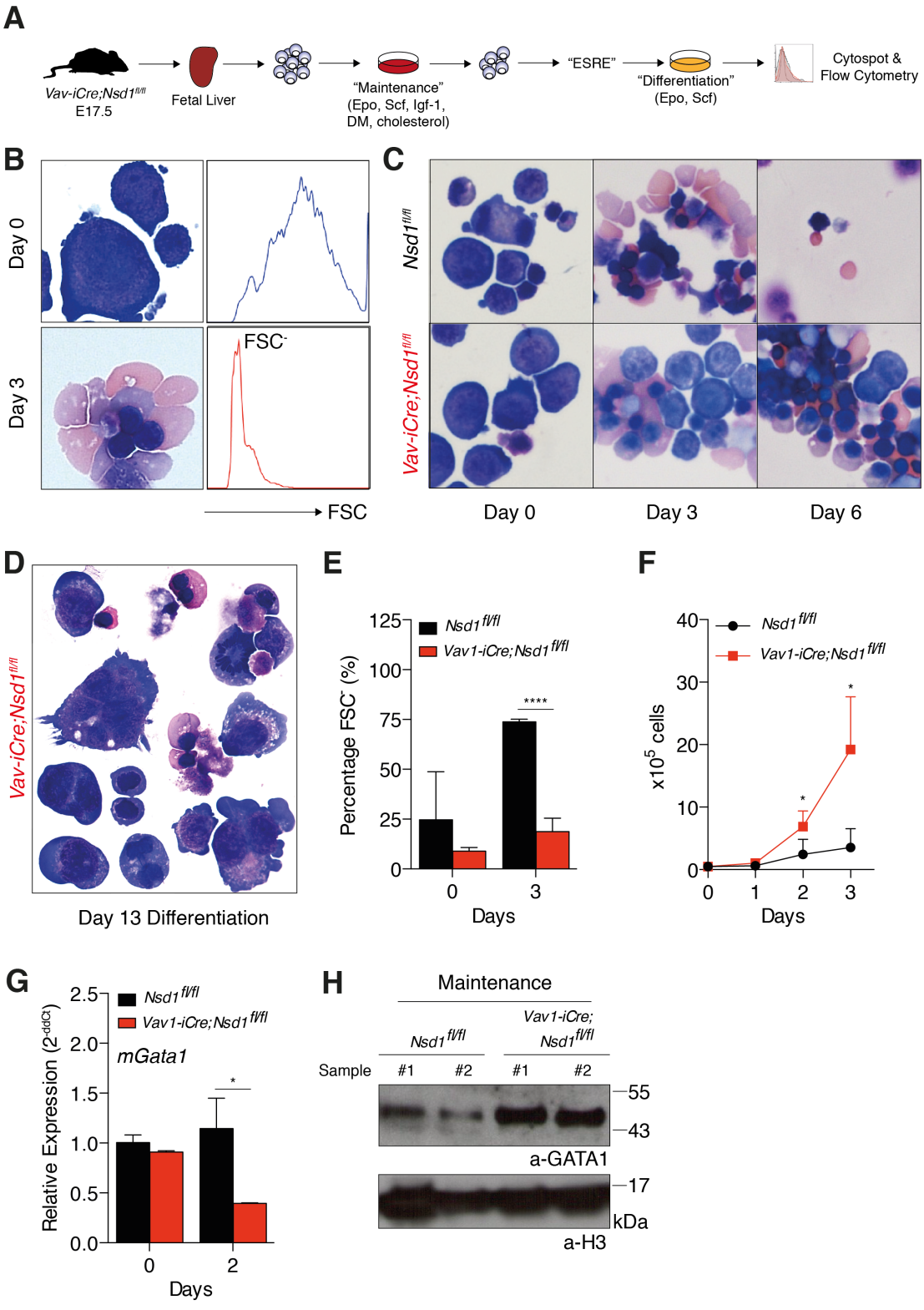

A
Vav-iCre;Nsd1fl/fl
E17.5
Fetal Liver
"Maintenance"
(Epo, Scf, Igf-1,
DM, cholesterol)
"ESRE"
"Differentiation"
(Epo, Scf)
Cytospot &
Flow Cytometry
B
Day 0
Day 3
FSC-
FSC
C
Nsd1fl/fl
Vav-iCre;Nsd1fl/fl
Day 0
Day 3
Day 6
D
Vav-iCre;Nsd1fl/fl
Day 13 Differentiation
E
Nsd1fl/fl
Vav1-iCre;Nsd1fl/fl
Percentage FSC- (%)
125
100
75
50
25
0
****
0
3
Days
F
Nsd1fl/fl
Vav1-iCre;Nsd1fl/fl
x10^5 cells
40
30
20
10
0
*
*
0
1
2
3
Days
G
Nsd1fl/fl
Vav1-iCre;Nsd1fl/fl
mGata1
Relative Expression (2-ddCt)
2.5
2.0
1.5
1.0
0.5
0.0
*
0
2
Days
H
Maintenance
Nsd1fl/fl
Vav1-iCre;
Nsd1fl/fl
Sample
#1
#2
#1
#2
55
43
a-GATA1
17
kDa
a-H3

FIGURE 29. Establishment of extensive self- renewing erythroblast cells of *Vav1-iCre;Nsd1$^{fl/fl}$* fetal livers to study erythroid differentiation *in vitro.*

(A) Schematic depiction of experimental setup using fetal liver cells of E17.5 that were grown in maintenance medium (containing dexamethyasone, SCF, EPO, IGF-1, cholesterol) to establish extensive self- renewing fetal liver derived erythryoblasts ("ESRE"). Cells were used for *in vitro* differentiation (containing serum, EPO, SCF) assays in liquid medium and subject to flow cytometry and morphologic analysis. **(B)** Representative images of wildtype HE- stained cells (left column) in maintenance (upper row) and differentiation medium (lower row) and forward scatter (FSC) profile of cells using flow cytometry. **(C)** Representative images of HE- stained cells in maintenance (left column) and at day three (middle column) and six (right column) in differentiation medium. Cells were obtained from fetal livers of *Nsd1$^{fl/fl}$* and *Vav1-iCre;Nsd1$^{fl/fl}$* mice at E17.5. **(D)** Representative selection of HE-stained *Vav1-iCre;Nsd1$^{fl/fl}$* derived "ESRE" cells at day 13 in differentiation medium. **(E)** Percentage forward scatter (FSC) negative living cells before (day 0) and after induction of differentiation (day 3) in *Nsd1$^{fl/fl}$* and *Vav1-iCre;Nsd1$^{fl/fl}$* ESRE cells (n=3/group. Two-way ANOVA with Sidak's multiple comparisons. **** $p<0.0001$). **(F)** Growth curve of fetal liver derived "ESRE" cells obtained from *Nsd1$^{fl/fl}$* (n=6, black bar) and *Vav1-iCre;Nsd1$^{fl/fl}$* mice (n=4, red bar) in differentiation medium. Living cells at indicated days were counted using trypan blue exclusion (Two-way ANOVA with Sidak's multiple comparisons. * $p<0.05$). **(G)** Quantitative RT- PCR of fetal liver cells before (day 0) and during differentiation (day 2) of *Nsd1$^{fl/fl}$* (black bar) and *Vav1-iCre;Nsd1$^{fl/fl}$* mice (red bar) indicating mRNA expression of *Gata1*. Ct values were normalized to *Gapdh* expression and control ESRE in maintenance medium. Value are shown as relative expression using 2^{-ddcT} method (n=2/group. Two way ANOVA with Sidak's multiple comparisons test. * $p<0.05$). **(H)** Western Blot analysis of protein levels of GATA1 and nuclear loading control histone 3 (H3) in "ESRE" cells in maintenance medium of *Nsd1$^{fl/fl}$* and *Vav1-iCre;Nsd1$^{fl/fl}$* mice. Data presented as mean, error bars represent ±SD.

Next we setup the “ESRE-like” proerythroblast culture system from diseased *Vav1-iCre;Nsd1$^{fl/fl}$* mice to perform functional analysis (**Figure 30A**). After six days, the majority of lineage-depleted cells was as expected TER119^{-}, CD71$^{dim/+}$, c-Kit^{+} and FSC^{+} (data not shown). Wildtype cells could be maintained in culture for more than 1 month and still formed colonies when plated into methylcellulose (data not shown). We were able to culture BM-derived proerythroblasts of diseased *Vav1-iCre;Nsd1$^{fl/fl}$* mice, which like fetal liver-derived “ESREs”, did not show any significant changes in cell growth when kept in MM for six days (**Figure 30B**). Growth of *Vav1-iCre;Nsd1$^{fl/fl}$* proerythroblasts was still dependent on SCF and EPO (**Figure 30C**). Like fetal liver-derived “ESRE-like” cells, adult derived *Vav1-iCre;Nsd1$^{fl/fl}$* proerythroblasts were also significantly impaired in terminal erythroid maturation in DM (**Figure 30D-E**). Interestingly, induced differentiation was also associated with lower Gata1 mRNA expression (**Figure 30F**). However, like the fetal liver-derived “ESRE-like” cells, *Vav1-iCre;Nsd1$^{fl/fl}$* proerythroblasts expressed constitutively high GATA1 protein levels (**Figure 30G**).

The unexpected high GATA1 protein levels led us the explore the cellular localization of GATA1 during differentiation of *Vav1-iCre;Nsd1$^{fl/fl}$* and control littermate proerythroblasts. We prepared cytospots from cells growing in MM or differentiating in DM for 48h, and stained for GATA1 by performing immunofluorescence (IF). We used murine erythroleukemia (MEL) cells known to express significant GATA1 protein levels as positive controls for IF and subsequent immunoblot experiments [56,208]. GATA1 presented as speckled pattern with fine dots throughout the nucleus of MEL cells (**Figure 31A**). Using the same conditions, we observed very low signal intensity in control erythroblasts in MM, whereas *Vav1-iCre;Nsd1$^{fl/fl}$* cells exhibited a very strong, clearly nuclear signal (**Figure 31B**). When cells were grown for 48 hours in DM, increased GATA1 expression was observed in control cells to comparable levels seen in *Vav1-iCre;Nsd1$^{fl/fl}$* cells. Interestingly, cells formed a single dot like structure that did not correlate with DAPI intensity, which has been previously observed by others (**Figure 31C** & data not shown) [56].

Taken together, we found that absence of Nsd1 impairs *in vitro* terminal erythroid maturation of self-renewing of fetal liver as well as adult BM-derived erythroblasts. Moreover we discovered that *Vav1-iCre;Nsd1$^{fl/fl}$* proerythroblasts express high levels of GATA1 protein without *Gata1* mRNA upregulation that is associated with erythroid maturation.

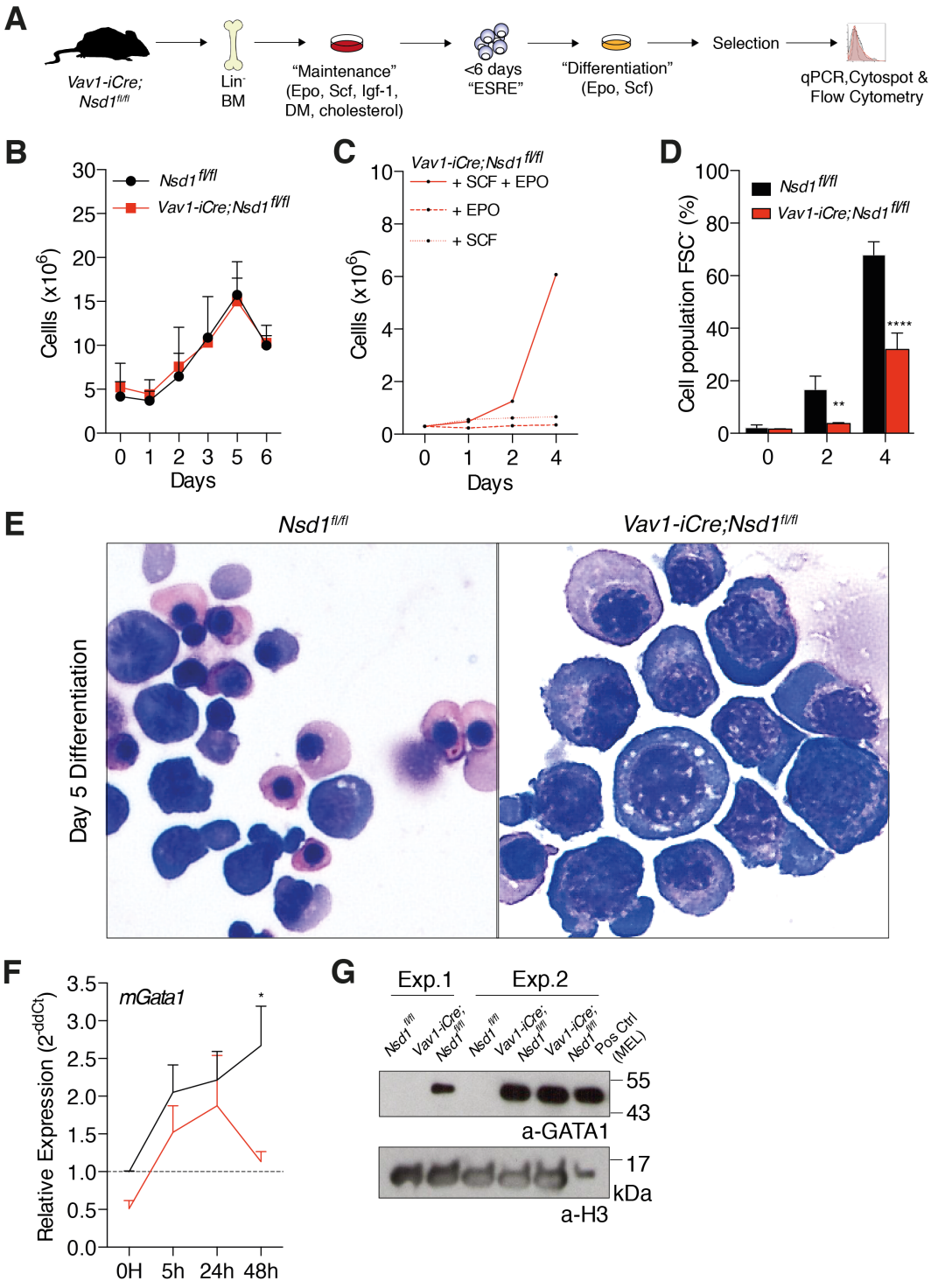

A
Vav1-iCre; Nsd1fl/fl
Lin- BM
"Maintenance" (Epo, Scf, Igf-1, DM, cholesterol)
<6 days "ESRE"
"Differentiation" (Epo, Scf)
Selection
qPCR,Cytospot & Flow Cytometry
B
Nsd1fl/fl
Vav1-iCre;Nsd1fl/fl
Cellls (x10^6)
Days
C
Vav1-iCre;Nsd1fl/fl
+ SCF + EPO
+ EPO
+ SCF
Cellls (x10^6)
Days
D
Nsd1fl/fl
Vav1-iCre;Nsd1fl/fl
Cell population FSC- (%)
**
****
E
Nsd1fl/fl
Vav1-iCre;Nsd1fl/fl
Day 5 Differentiation
F
mGata1
Relative Expression (2-ddCt)
*
0H 5h 24h 48h
G
Exp.1
Exp.2
Nsd1fl/fl
Vav1-iCre; Nsd1fl/fl
Pos Ctrl (MEL)
55
43
a-GATA1
17
kDa
a-H3

FIGURE 30. Establishment of pro- erythroblast cells of diseased *Vav1-iCre;Nsd1$^{fl/fl}$* mice to study erythroid differentiation for functional analysis.

(A) Schematic depiction of experimental setup using lineage depleted bone marrow cells of *Vav1-iCre;Nsd1$^{fl/fl}$* grown in maintenance medium for short- time (< 6 days of culture) or long- time (>6 days of liquid culture) to establish extensive self- renewing erythryoblasts ("ESRE"). Cells were used for *in vitro* differentiation assays in liquid medium and subject to flow cytometry and morphologic analysis, analyzed by qPCR or treated with inhibitors. **(B)** Short- time cell growth of bone- marrow derived "ESRE" *Nsd1$^{fl/fl}$* and diseased *Vav1-iCre;Nsd1$^{fl/fl}$* in enance medium (n=3/group. Students t-test with equal variances). **(C)** Cell growth of established (in liquid culture for more than 6 days) *Vav1-iCre;Nsd1$^{fl/fl}$* "ESRE" in different kinds of maintenance medium, either complete (plus SCF and EPO), or lacking EPO or SCF (n=1). **(D)** Percentage of forward scatter (FSC) positive living cells before (day 0) and after induction of differentiation (day 2 and 4) in *Nsd1$^{fl/fl}$* and *Vav1-iCre;Nsd1$^{fl/fl}$* ESRE cells that were kept in liquid culture more than 6 days before induction of differentiation (n=3/group. Two-way ANOVA with Sidak's multiple comparisons test. ** p<0.01, **** p<0.0001). **(E)** Representative image of ESRE cells 5 days in differentiation medium of *Nsd1$^{fl/fl}$* (upper row) *Vav1-iCre;Nsd1$^{fl/fl}$* (lower row). **(F)** Quantitative RT- PCR of bone marrow derived "ESRE" cells before (day 0) and during differentiation (5, 24 and 48 hours) of *Nsd1$^{fl/fl}$* (black line) and *Vav1-iCre;Nsd1$^{fl/fl}$* mice (red line) indicating mRNA expression of *Gata1*. Ct values were normalized to *Gapdh* expression and control ESRE in maintenance medium. Value are shown as relative expression using 2^{-ddcT} method (n=3/group. Two way ANOVA with Sidak's multiple comparisons test. * p<0.05, ** p<0.01). **(G)** Western blot analysis of protein levels of GATA1 and nuclear loading control histone 3 (H3) in bone marrow derived "ESRE" cells in maintenance medium of *Nsd1$^{fl/fl}$* and *Vav1-iCre;Nsd1$^{fl/fl}$* mice of two different experiments and murine erythroleukemia cell line (MEL) as positive control. **(H)** Percentage of forward scatter (FSC) negative living cells before (day 0) and after induction of differentiation (day 3) in *Nsd1$^{fl/fl}$* (black line) and *Vav1-iCre;Nsd1$^{fl/fl}$* (red line) ESRE cells treated with I-CBP112 at increasing concentrations (DMSO, 1, 3 and 5uM). Cells were kept in liquid culture more than 6 days before induction of differentiation. Black and red stars indicate significance to untreated control, grey stars indicate significance between two groups (n=3/group. Two way ANOVA with Sidak's multiple comparisons test. * p<0.05). Data presented as mean, error bars represent ±SD.

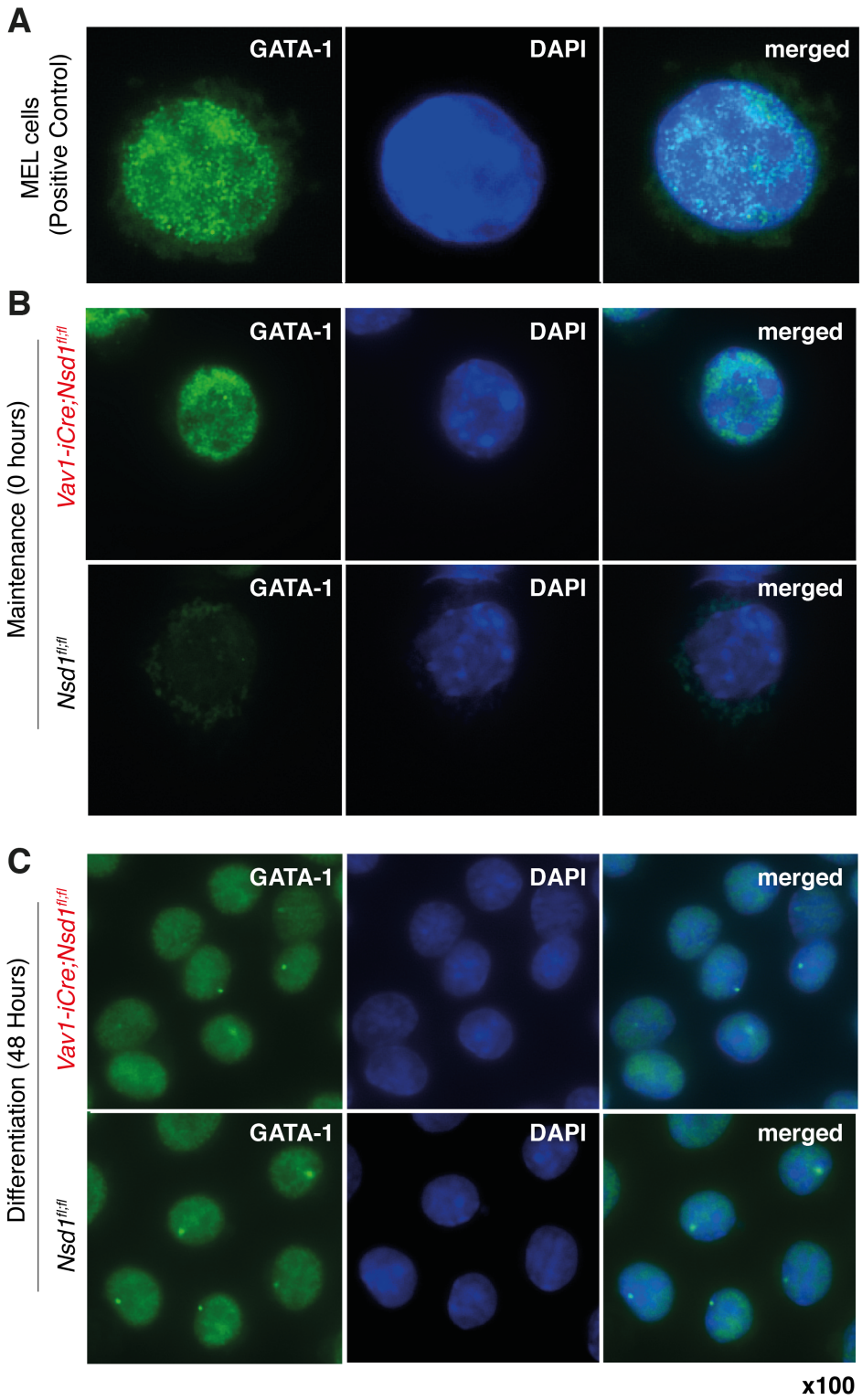


FIGURE 31. Localization of GATA1 in erythroblasts cells of diseased *Vav1-iCre;Nsd1$^{fl/fl}$* mice.

(A) Representative images of immunofluorescence staining of GATA1 and DAPI as nuclear marker in MEL cells as positive control. **(B)** *Nsd1$^{fl/fl}$* (left column) and *Vav1-iCre;Nsd1$^{fl/fl}$* (right column) "ESRE" cells in maintenance medium, **(C)** differentiation medium (48 hours). Cells are shown with 100x magnification.

4.3.13 Overexpression of GATA1 induces erythroid maturation of *Vav1-iCre;Nsd1$^{fl/fl}$* proerythroblasts

Similar to Friend-virus integration driven MEL cells, *Vav1-iCre;Nsd1$^{fl/fl}$* proerythroblasts express very high levels of GATA1 protein and are significantly impaired in erythroid maturation. As overexpression of *Gata1* is able to induce terminal maturation of MEL cells and restore erythroid differentiation in other mouse models [73,208–210]. We wondered whether exogenous alterations of Gata1 expression may also have an impact on maturation of *Vav1-iCre;Nsd1$^{fl/fl}$* proerythroblasts and/or aberrant colony formation in methylcellulose [211,212]. To address the effect on stem and progenitor dependent colony formation, we depleted BM cells of diseased *Vav1-iCre;Nsd1$^{fl/fl}$* mice from cells expressing lineage-associated markers and transduced them in a stem cell favoring medium, containing SCF and TPO, with *pMSCV-mouseGata1-pgk-puro* (mGATA1) or with the vector *pMSCV-pgk-puro* (mock) (**Figure 32A**). Mock-transduced *Vav1-iCre;Nsd1$^{fl/fl}$* derived, “stem” cells formed significantly more colonies than the *Nsd1$^{fl/fl}$* controls. Interestingly retroviral overexpression of *Gata1* significantly reduced the number of colonies in both *Nsd1$^{fl/fl}$* and *Vav1-iCre;Nsd1$^{fl/fl}$* transduced “stem” cells (**Figure 32B-C**). Cytospin preparations reflected a slight shift to erythroid differentiation in mGATA1 transduced *Vav1-iCre;Nsd1$^{fl/fl}$* cells (**Figure 32D**).

To study the impact of *Gata1* overexpression on colony formation and maturation of erythroid progenitors we transduced the cells one day after lineage depletion in maintenance medium supporting growth of proerythroblasts (**Figure 33A**). Upon plating in methylcellulose, mock transduced *Vav1-iCre;Nsd1$^{fl/fl}$* cells formed more colonies than the *Nsd1$^{fl/fl}$* control. We also observed abnormally dense appearing red colonies. Overexpression of *Gata1* resulted in a decrease of colonies formed by *Nsd1$^{fl/fl}$* cells. However, with and without *Gata1* overexpression, *Nsd1$^{fl/fl}$* colonies had a myeloid morphology resembling CFU-G/Ms. In contrast, overexpression of *Gata1* not only significantly reduced the number of colonies formed by *Vav1-iCre;Nsd1$^{fl/fl}$* cells but also induced erythroid maturation with colonies resembling most mature erythroid colonies called CFU-Es (**Figure 33B-C**).

Cytospin preparations displayed myeloid cells in *Nsd1*$^{fl/fl}$ and mostly erythroid differentiated cells in *Vav1-iCre;Nsd1*$^{fl/fl}$ derived colonies overexpressing *Gata1* (**Figure 33D**). To exclude the influence of different starting populations (myeloid progenitors in controls and proerythroblasts in *Vav1-iCre;Nsd1*$^{fl/fl}$ cells) we first synchronized the cells for six days in maintenance medium resulting in a R0-R1 erythroid maturation phase of the majority of the cells (data not shown). Lineage depleted BM cells of diseased *Vav1-iCre;Nsd1*$^{fl/fl}$ mice and the "synchronized" *Nsd1*$^{fl/fl}$ proerythroblasts were transduced, puromycin selected and plated into differentiation medium for different time points (**Figure 33A**). As previously observed in fetal livers, *Vav1-iCre;Nsd1*$^{fl/fl}$ mock transduced cells grew two times faster than *Nsd1*$^{fl/fl}$ cells during liquid differentiation assay, whereas *Gata1* transduced cells had a tendency to grow slower in differentiation medium (**Figure 33E**). When cells were spun down for further analysis, the pellets of Gata1 overexpressing *Vav1-iCre;Nsd1*$^{fl/fl}$ cells appeared red whereas mock-transduced *Vav1-iCre;Nsd1*$^{fl/fl}$ cells appeared white reflecting missing formation of erythrocytes (**Figure 33F**). We then substantiated *Gata1* induced erythroid differentiation by immunophenotyping of the cells. Transduced cells were plated into differentiation medium and the expression of TER119 and size (FSC) analyzed by flow cytometry. All conditions started the differentiation assay with a FSC^{+} population mostly negative for TER119. Upon completion of differentiation assay (96 hours), mock or *Gata1* transduced *Nsd1*$^{fl/fl}$ controls and *Gata1*-transduced *Vav1-iCre;Nsd1*$^{fl/fl}$ cells contained around 60% FSC^{-} and 15-30% TER119^{+} cells whereas mock transduced *Vav1-iCre;Nsd1*$^{fl/}$ cells only contained 35% FSC^{-} and 10% TER119^{+} cells, also depicted in flow cytometry panels (**Figure 33G-I**). Protein analysis of nuclear GATA1 levels by Western blot confirmed overexpression of HA-tagged GATA1 in differentiating *Nsd1*$^{fl/fl}$ and *Vav1-iCre;Nsd1*$^{fl/fl}$ cells. As previously observed in fetal liver and lineage depleted ESRE-like proerythroblasts (0H), *Vav1-iCre;Nsd1*$^{fl/fl}$ constitutively expressed significant GATA1 levels before induction of differentiation. Surprisingly, GATA1 levels in mock and transduced *Vav1-iCre;Nsd1*$^{fl/fl}$ cells remained almost stable, whereas they clearly increased upon differentiation of *Nsd1*$^{fl/fl}$ control cells (**Figure 33J**).

Collectively this data suggests that overexpression can overcome the erythroid maturation block of proerythroblasts lacking *Nsd1*.

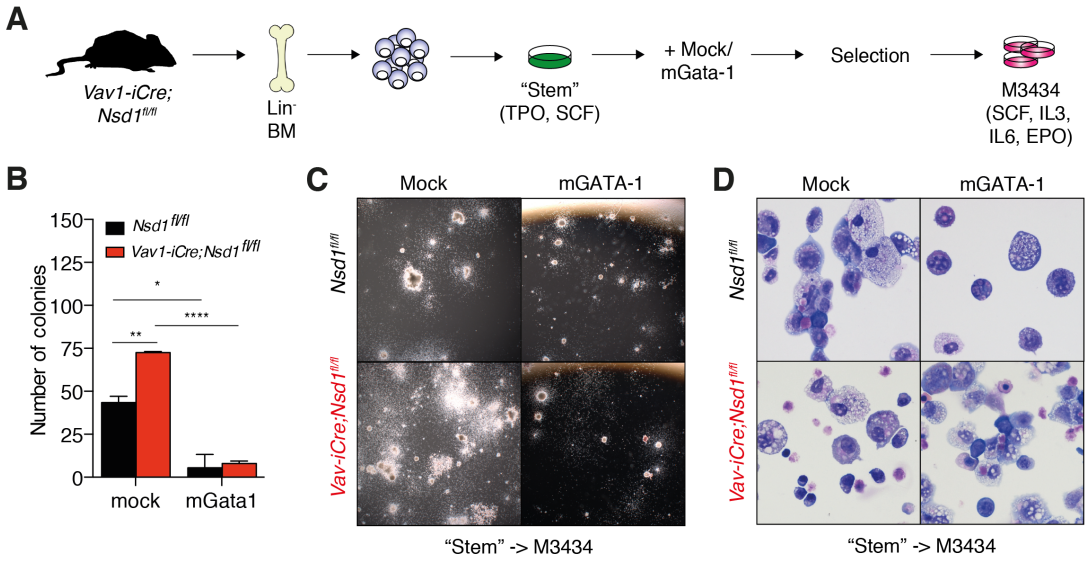


FIGURE 32. Retroviral overexpression of Gata1 in Lin- depleted BM cells of diseased *Vav1-iCre;Nsd1$^{fl/fl}$*. **(A)** Schematic depiction of experimental setup using lineage depleted bone marrow cells of *Vav1-iCre;Nsd1$^{fl/fl}$* grown in "stem" medium containing TPO and SCF, transduced with *pMSCV-pgk-puro* or *pMSCV-mGata1-pgk-puro*, puromycin selected and plated into methylcellulose M3434 containing SCF, IL3, IL6 and EPO. **(B)** Total number of colonies formed by 1.000 lineage depleted cells in M3434 methylcellulose of *Nsd1$^{fl/fl}$* (black bar) and *Vav1-iCre;Nsd1$^{fl/fl}$* (red bar) transduced in "Stem" medium with *pMSCV-pgk-puro* or *pMSCV-mGata1-pgk-puro* (n=2/group. Two way ANOVA with Sidak's multiple comparisons test. * p<0.05, ** p<0.01, *** p<0.001). **(C)** Representative images of colonies formed by 1.000 lineage depleted cells in M3434 methylcellulose of *Nsd1$^{fl/fl}$* (upper row) and *Vav1-iCre;Nsd1$^{fl/fl}$* (lower row) transduced in "Stem" medium with *pMSCV-pgk-puro* (left column) or *pMSCV-mGata1-pgk-puro* (right column). **(D)** Representative images of HE- stained cytospots obtained from resuspended cells after first round of methylcellulose plating of *Nsd1$^{fl/fl}$* (upper row) and *Vav1-iCre;Nsd1$^{fl/fl}$* (lower row) transduced in "Stem" medium with *pMSCV-pgk-puro* (left column) or *pMSCV-mGata1-pgk-puro* (right column).

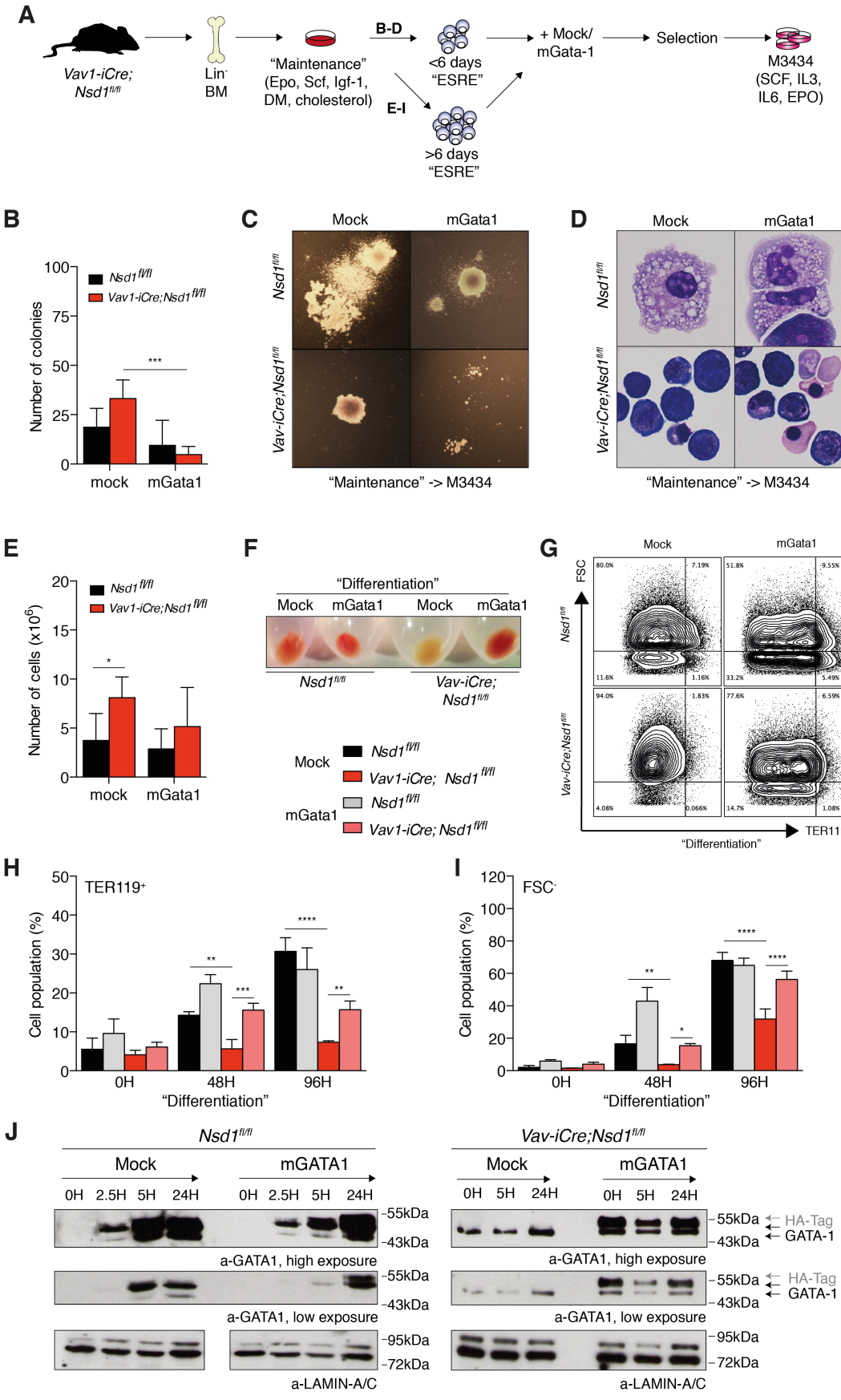

A
Vav1-iCre; Nsd1fl/fl
Lin- BM
"Maintenance" (Epo, Scf, Igf-1, DM, cholesterol)
B-D
<6 days "ESRE"
E-I
>6 days "ESRE"
+ Mock/ mGata-1
Selection
M3434 (SCF, IL3, IL6, EPO)
B
Nsd1fl/fl
Vav1-iCre;Nsd1fl/fl
Number of colonies
mock
mGata1
C
Mock
mGata1
"Maintenance" -> M3434
D
Mock
mGata1
"Maintenance" -> M3434
E
Number of cells (x10^6)
mock
mGata1
F
"Differentiation"
Mock
mGata1
Vav-iCre; Nsd1fl/fl
G
FSC
TER119
"Differentiation"
H
TER119+
Cell population (%)
0H
48H
96H
"Differentiation"
I
FSC-
Cell population (%)
0H
48H
96H
"Differentiation"
J
Nsd1fl/fl
Vav-iCre;Nsd1fl/fl
Mock
mGATA1
0H 2.5H 5H 24H
0H 5H 24H
55kDa
43kDa
HA-Tag
GATA-1
a-GATA1, high exposure
a-GATA1, low exposure
95kDa
72kDa
a-LAMIN-A/C

FIGURE 33. Retroviral Gata1 overexpression in pro-erythroblasts of diseased *Vav1-iCre;Nsd1$^{fl/fl}$* mice. (A) Schematic depiction of experimental setup using lineage depleted bone marrow cells of *Vav1-iCre;Nsd1$^{fl/fl}$* grown "maintenance" medium, transduced with *pMSCV-pgk-puro* or *pMSCV-mGata1-pgk-puro*, puromycin selected and plated into methylcellulose M3434 containing SCF, IL3, IL6 and EPO. For erythroid differentiation analysis, cells were first kept for more than six days in liquid maintenance culture and then analyzed by flow cytometry. **(B)** Total number of colonies formed by 5.000 lineage depleted cells in M3434 methylcellulose of *Nsd1$^{fl/fl}$* (black bar) and *Vav1-iCre;Nsd1$^{fl/fl}$* (red bar) transduced in "maintenance" medium with *pMSCV-pgk-puro* or *pMSCV-mGata1-pgk-puro* (n=6/group. Two way ANOVA with Sidak's multiple comparisons test. *** p<0.001). **(C)** Representative images of colonies formed by 5.000 lineage depleted cells in M3434 methylcellulose of *Nsd1$^{fl/fl}$* (upper row) and *Vav1-iCre;Nsd1$^{fl/fl}$* (lower row) transduced in "maintenance" medium with *pMSCV-pgk-puro* (left column) or *pMSCV-mGata1-pgk-puro* (right column). **(D)** Representative images of HE- stained cytospots obtained from resuspended cells after first round of methylcellulose plating of *Nsd1$^{fl/fl}$* (upper row) and *Vav1-iCre;Nsd1$^{fl/fl}$* (lower row) transduced in "maintenance" medium with *pMSCV-pgk-puro* (left column) or *pMSCV-mGata1-pgk-puro* (right column). **(E)** Number of cells after 48 hours of liquid differentiation of *Nsd1$^{fl/fl}$* (black bar) and *Vav1-iCre;Nsd1$^{fl/fl}$* (red bar) ESRE cells transduced in "maintenance" medium with *pMSCV-pgk-puro* or *pMSCV-mGata1-pgk-puro* (n=6. Two way ANOVA with Sidak's multiple comparisons test. * p<0.05). **(F)** Representative image of cell pellets after 48 hours of liquid differentiation of *Nsd1$^{fl/fl}$* and *Vav1-iCre;Nsd1$^{fl/fl}$* ESRE cells transduced in "maintenance" medium with *pMSCV-pgk-puro* or *pMSCV-mGata1-pgk-puro*. **(G)** Representative flow cytometry panel of *Nsd1$^{fl/fl}$* (upper row) and *Vav1-iCre;Nsd1$^{fl/fl}$* (lower row) ESRE cells transduced in "maintenance" medium with *pMSCV-pgk-puro* (left column) or *pMSCV-mGata1-pgk-puro* (right column) after 96 hours of liquid differentiation. **(H)** Percentage of flow cytometry Ter119 expression marker positive stained and **(I)** forward scatter negative cell population before (0 hours) and after induction of differentiation (48 and 96 hours) of *Nsd1$^{fl/fl}$* (black bar) and *Vav1-iCre;Nsd1$^{fl/fl}$* (red bar) ESRE cells transduced in "maintenance" medium with *pMSCV-pgk-puro* (resp. black or red bar) or *pMSCV-mGata1-pgk-puro* (resp. red or rose bar) (n=3. Two-way ANOVA with Turkey's multiple comparisons. * p<0.05, ** p<0.01, *** p<0.001, **** p<0.0001). **(J)** Western blot analysis of GATA1 and LAMIN-A/C protein levels as nuclear loading control in *Nsd1$^{fl/fl}$* (left panel) and *Vav1-iCre;Nsd1$^{fl/fl}$* (right panel) ESRE cells transduced in "maintenance" medium with *pMSCV-pgk-puro* (resp. black or red arrow) or *pMSCV-mGata1-pgk-puro* (resp. red or rose arrow) before and after liquid differentiation at different time points. Arrows at the right site indicate endogenous GATA1 as well as GATA1+HA- Tag sizes. Data presented as mean, error bars represent ±SD.

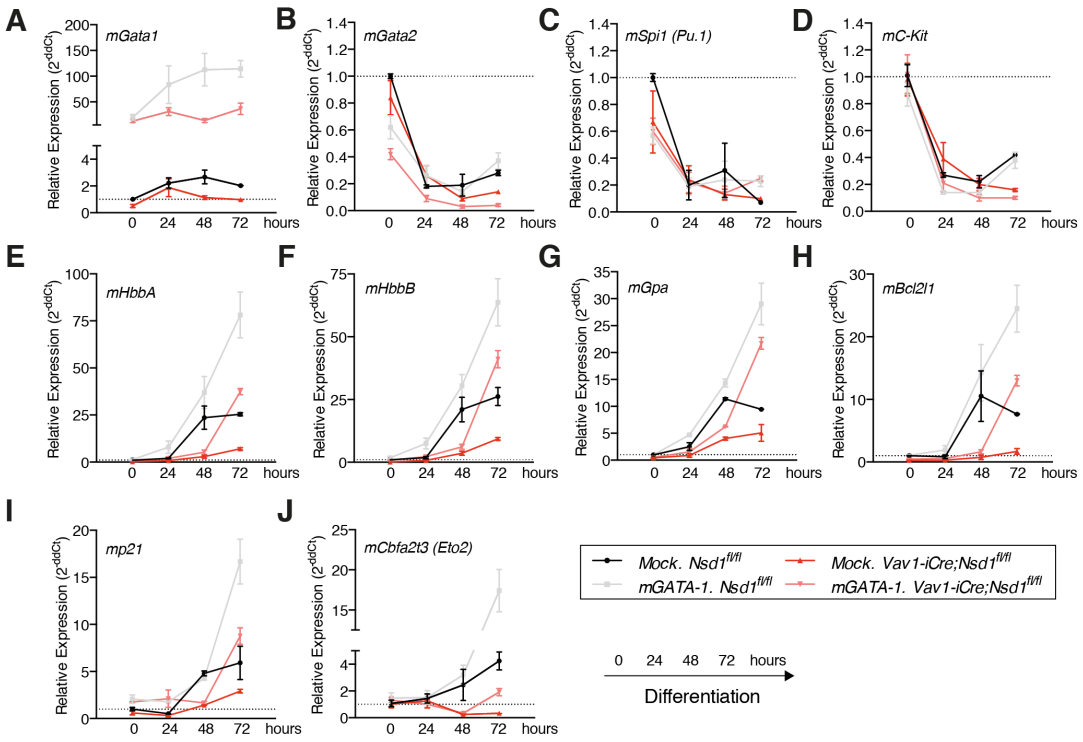


FIGURE 34. Expression of Gata1 regulated targets during erythroid differentiation of pro- erythroblasts of diseased *Vav1-iCre;Nsd1$^{fl/fl}$*.

(A) Quantification of *Gata1*, **(B)** *Gata2*, **(C)** *Spi1 (Pu.1),* **(D)** c-Kit, **(E)** *HbbA,* **(F)** *Hbb-B,* **(G)** *Gpa,* **(H)** *Bcl2l1,* **(I)** *p21,* **(J)** *Cbfa2t3* mRNA expression of bone marrow derived cells of *Nsd1$^{fl/fl}$* and *Vav1-iCre;Nsd1$^{fl/fl}$* ESRE cells transduced in "maintenance" medium with *pMSCV-pgk-puro* (resp. black or red line) or *pMSCV-mGata1-pgk-puro* (resp. red or rose line). Values are shown as relative expression using 2^{-ddcT} method by normalization to *Gapdh* and *Nsd1$^{fl/fl}$* transduced with *pMSCV-pgk-puro* mice (n=3/ 0,24,48 hours and n=1/ 72 hours group). Data presented as mean, error bars represent ±SD.

We next wondered whether overexpression of *Gata1* was associated with activation of known GATA1 target genes [32]. As GATA1 is subject to autoregulation we first compared *Gata1* mRNA levels *Nsd1*$^{fl/fl}$ and *Vav1-iCre;Nsd1*$^{fl/fl}$ cells in differentiation medium with and without Gata1 overexpression. We observed differentiation-associated upregulation of *Gata1* mRNA in *Nsd1*$^{fl/fl}$ control cells. *Vav1-iCre;Nsd1*$^{fl/fl}$ cells expressed significantly lower *Gata1* mRNA levels. Overexpression of *Gata1* resulted in high mRNA levels in both types of cells. Interestingly, lower *Gata1* mRNA levels were obtained in *Vav1-iCre;Nsd1*$^{fl/fl}$ cells (**Figure 34A**). We then followed the expression of genes that are downregulated by *Gata1*. Hereby we found that with and without overexpression of *Gata1*, mRNA levels of *Gata2*, *Spi1* (Pu.1) and *c-kit*, decreased upon differentiation of both *Vav1-iCre;Nsd1*$^{fl/fl}$ and *Nsd1*$^{fl/fl}$ controls (**Figure 34B-D**). In sharp contrast, differentiation-associated upregulation of *Gata1* induced targets such as hemoglobin alpha and beta chains (*HbbA/B*) or glycophorin A (*GPA*, also known as TER119) was significantly impaired in mock transduced *Vav1-iCre;Nsd1*$^{fl/fl}$ cells and its increase upon *Gata1* overexpression was much lower than in *Nsd1*$^{fl/fl}$ control cells (**Figure 34E-G**). Similar regulation was observed for the negative apoptosis regulator BCL2 like 1 (*Bcl2l1*) and the cyclin-dependent kinase inhibitor 1 (CDK1, also known was *p21*) (**Figure 34H-I**). Interestingly, we observed a very poor upregulation of *Cbfa2t3* (also known as Eto2) upon erythroid differentiation of *Vav1-iCre;Nsd1*$^{fl/fl}$ with or without Gata1 overexpression (**Figure 34J**). Collectively we found a minor impact of the Nsd1 status for erythroid differentiation-associated Gata1 negatively regulated targets, but impaired expression of activating target that was at least partially restored by overexpression of *Gata1*.

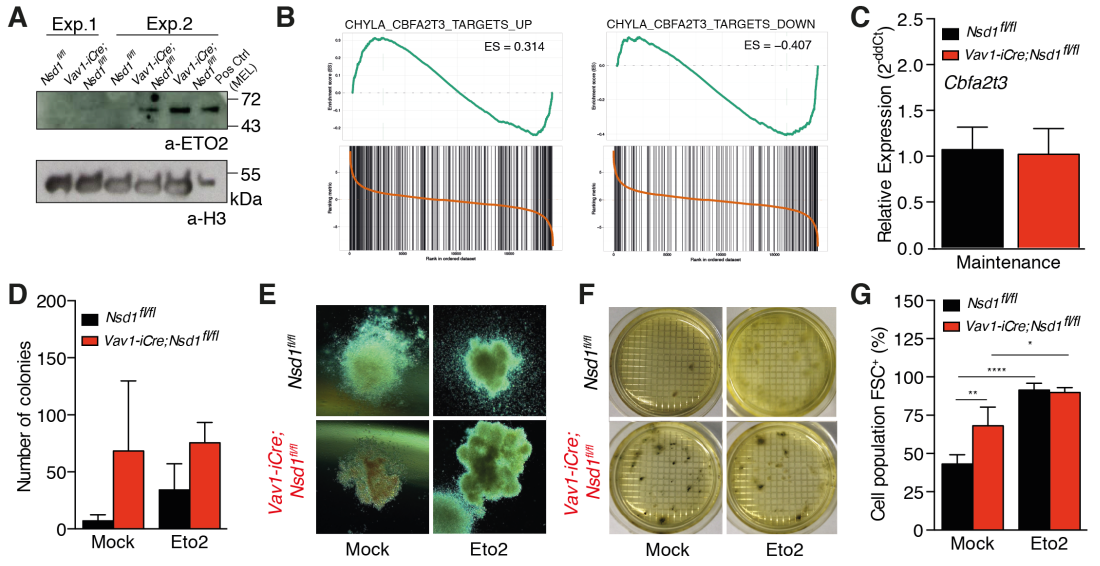


FIGURE 35. Retroviral overexpression of Eto2 in pro- erythroblasts of diseased *Vav1-iCre;Nsd1$^{fl/fl}$* mice.

(A) Western blot analysis of protein levels of ETO-2 and nuclear loading control histone 3 (H3) in bone marrow derived "ESRE" cells in maintenance medium of *Nsd1$^{fl/fl}$* and *Vav1-iCre;Nsd1$^{fl/fl}$* mice of two different experiments and murine erythroleukemia cell line (MEL) as positive control. **(B)** Gene set enrichment analysis for MEP derived signatures comparing to curated gene sets of "Chyla Genes up/down-regulated in immature bone marrow progenitor cells upon knock out of CBFA2T3". **(C)** Quantification of *Cbfa2t3* mRNA expression of bone marrow derived cells of *Nsd1$^{fl/fl}$* (black bar) and *Vav1-iCre;Nsd1$^{fl/fl}$* (red bar) ESRE cells (n=3 per group. Students t-test with equal variances). **(D)** Total number of colonies formed by 20.000 lineage depleted cells in M3434 methylcellulose of *Nsd1$^{fl/fl}$* (black bar, n=3) and *Vav1-iCre;Nsd1$^{fl/fl}$* (red bar, n=4) transduced in "Maintenance" medium with *pMSCV-pgk-puro* or *pMSCV-hETO-2-pgk-puro-IRES-GFP* (n=3/group. Two way ANOVA with Sidak's multiple comparisons). **(E)** Representative images of colonies formed by lineage depleted cells in M3434 methylcellulose of *Nsd1$^{fl/fl}$* (upper row) and *Vav1-iCre;Nsd1$^{fl/fl}$* (lower row) transduced in "Maintenance" medium with *pMSCV-pgk-puro* (left column) or *pMSCV-hETO-2-pgk-puro-IRES-GFP* (right column). **(F)** Representative images of colonies formed in methylcellulose M3434 after benzidine staining of dishes. **(G)** Percentage of forward scatter (FSC) positive living cells after induction of differentiation (day 4) in *Nsd1$^{fl/fl}$* (black bar, n=3) and *Vav1-iCre;Nsd1$^{fl/fl}$* (red bar, n=4) ESRE cells transduced in "Maintenance" medium with *pMSCV-pgk-puro* or *pMSCV-hETO-2-pgk-puro-IRES-GFP* that were kept in liquid culture more than 6 days before induction of differentiation (n=3/group. Two-way ANOVA with Sidak's multiple comparisons test. * p<0.05, ** p<0.01, **** p<0.0001). Data presented as mean, error bars represent ±SD.

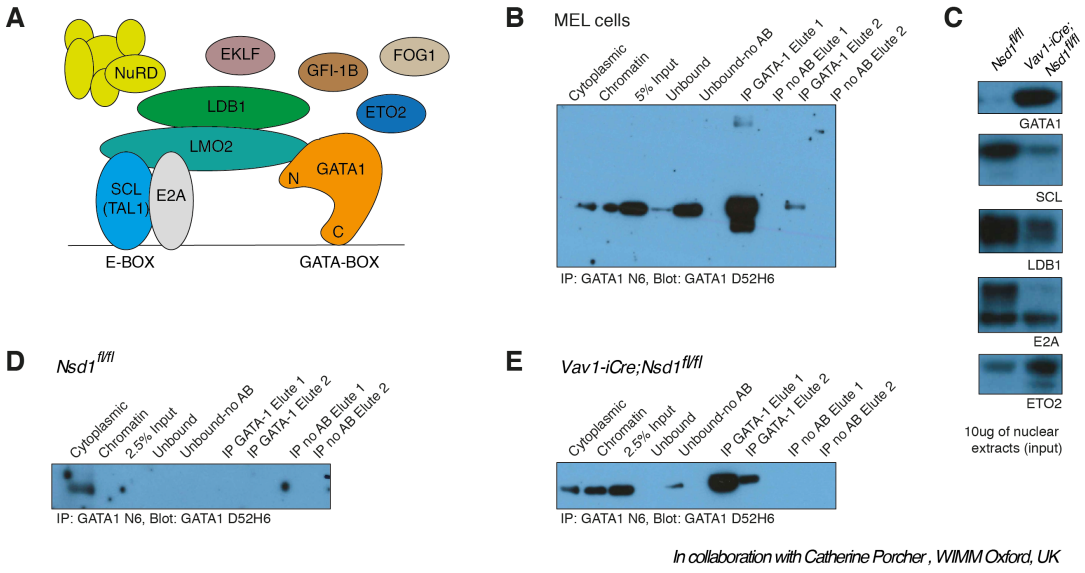
4.3.14 Absence of Nsd1 results in higher ETO2 protein levels

In addition to GATA1, we also observed increased ETO2 protein levels in erythroblasts of *Vav1-iCre;Nsd1$^{fl/fl}$* mice (**Figure 35A**). Moreover we previously found poor mRNA regulation of ETO2 in *Vav1-iCre;Nsd1$^{fl/fl}$* cells during *in vitro* differentiation (**Figure 34J**). Gene set enrichment analysis (GSEA) of the expression profiles obtained from sorted MEPs of *Vav1-iCre;Nsd1$^{fl/fl}$* mice revealed some interesting links to previous signatures derived from *Eto2 (Cbfa2t3, or Mtg16)* knockout mice. In fact MEP lacking Nsd1 seemed to correlate with previously proposed Eto2 targets (**Figure 35B**) [213]. However we found no difference in Eto2 mRNA expression in proerythroblasts of *Vav1-iCre;Nsd1$^{fl/fl}$* mice (**Figure 35C**). Since mRNA levels were not increased we investigated whether retroviral Eto-2 overexpression may affect colony formation and *in vitro* erythroid differentiation. We did not observe any significant differences in colony formation between mock and *Eto2* transduced cell. *Vav1-iCre;Nsd1$^{fl/fl}$* colonies were more abundant with and without Eto2 overexpression (**Figure 35D**). Interestingly, colonies appeared mostly white and non- hemoglobinized (as shown by benzidine staining) and the size seems larger in *Vav1-iCre;Nsd1$^{fl/fl}$* than control derived colonies (**Figure 35E&F**). Importantly, overexpression of *Eto2* clearly blocked erythroid differentiation of control and *Vav1-iCre;Nsd1$^{fl/fl}$* derived erythroblasts reflected by increasing percentages of FSC^{+} cells (**Figure 35G**). However, cells showed decreased CD71 expression after four days in differentiation medium. Our observation so far suggests that the absence of Nsd1 results in aberrant Eto2 regulation. In currently ongoing experiments we are addressing the impact of Eto2 knockdown on erythroid differentiation in presence or absence of Nsd1.

4.3.15 Absence of Nsd1 alters expression of known GATA1-complex forming proteins

As shown before, the lack of Nsd1 influenced regulation of known GATA1 target genes: repressed GATA1 target genes were downregulated in *Vav1-iCre;Nsd1$^{fl/fl}$* whereas activated targets did not increase to the same extent as seen in *Nsd1$^{fl/fl}$* control cells (**Figure 34**). Interestingly retroviral overexpression of *Gata1* induced the expression of erythroid differentiation markers resulted in erythroid maturation. Previous studies have shown that GATA1 is part of a protein-protein interaction network controlling transcription of erythroid differentiation regulating genes. Depending on the partner, it can form repressing or activating transcription factor complexes [63,65]. One of these protein interaction networks is the activating pentamer containing T-Cell Acute Lymphocytic Leukemia 1 (TAL1, also known as SCL), Transcription Factor 3 (E2A), LIM Domain Binding 1 (LDB1) and LIM Domain Only 2 (LMO2) (see **Figure 36A** for schematic overview) [214]. The LIM (Lin11, Isl-1, Mec-3) domain is protein- protein interaction motif presenting as tandem zinc finger [215,216]. Together the complex binds to erythroid- differentiation associated regulatory sequences such as the globin locus [217].

We hypothesized that the lack of *Nsd1* may result in aberrant GATA1 protein containing complexes being "repaired" upon overexpression. To obtain a first insight into formation of this GATA1 complex, we used an immunoprecipitation (IP) strategy. First we tested the efficiency of different protocols for preparation of cytoplasmic and nuclear extracts using five to ten times sonication and increasing amounts of benzonase (5-10U/µl) on murine erythroleukemia cells (MEL), which we previously used as a positive control for GATA1 IF and western blots. Increased amounts of benzonase (10U/µl) and 5 sonication steps allowed us to immunoprecipitate increased amounts of chromatin bound nuclear GATA1 (data not shown). Next, I tested the optimal antibody concentration for immunoprecipitation of GATA1 in MEL cells. Using a protocol to mildly separate cytoplasmic from nuclear fraction, we could achieve enrichment of nuclear signal (**Figure 36B**) [218]. However, standard mild nuclear lysis, including sonication and benzonase treatment, was not able to fully free GATA1 off the chromatin. We then prepared nuclear extracts from control and *Vav1-iCre;Nsd1$^{fl/fl}$* differentiating proerythroblasts and first checked the expression of some of the best studied GATA1 complex members. As expected, GATA1 protein levels were significantly increased in *Vav1-iCre;Nsd1$^{fl/fl}$* cells compared to *Nsd1$^{fl/fl}$* controls.



In collaboration with Catherine Porcher , WIMM Oxford, UK

FIGURE 36. Immunoprecipitation of GATA1.

(A) Schematic depiction of elected complex members of GATA1. GATA1 forms a pentamer with FOG1, LMO2, LDB1, SCL and E2A, generally associated with activation of transcription. Moreover it can also form repressive complex, e.g. with ETO2. **(B)** Immunoprecipitation of GATA1 using N6 antibody in MEL cells. GATA1 was detected in western blot with D52H6 antibody. **(C)** Western Blot of nuclear extractions (10ug) of *Nsd1*$^{fl/fl}$ and *Vav1-iCre;Nsd1*$^{fl/fl}$ erythroblasts for abundance of factors. **(D)** Immunoprecipitation of GATA1 using N6 antibody in *Nsd1*$^{fl/fl}$ differentiating erythroblasts. GATA1 was detected in western blot with D52H6 antibody. **(E)** Immunoprecipitation of GATA1 using N6 antibody in *Vav1-iCre;Nsd1*$^{fl/fl}$ differentiating erythroblasts. GATA1 was detected in western blot with D52H6 antibody.

Interestingly, protein levels of the activating complex such as SCL, LDB1 and E2A were decreased and repressing complex member ETO2 was increased in *Vav1-iCre;Nsd1*$^{fl/fl}$ nuclear extracts (**Figure 36C**). Using the optimized conditions, we aimed to precipitate GATA1 in differentiating control erythroblasts. However, we were not able to pull down enough protein to perform qualitative analysis of the complexes. Most of the protein was lost in cytoplasmic extracts or appeared to be tightly bound to chromatin (**Figure 36D** & data not shown). Unfortunately, control cells are not as stable in liquid culture as *Vav1-iCre;Nsd1*$^{fl/fl}$ erythroblasts therefore yielding first lower cell numbers and second smaller cells during differentiation and therefore less protein for further analysis. Next we decided to use only *Vav1-iCre;Nsd1*$^{fl/fl}$ cells, that express higher GATA1 protein levels. We could successfully pull- down GATA1 in these cells (**Figure 36E**). As we miss the appropriate controls for this experiment we are currently exploring MEL cells with Nsd1 knockdown to dissect the influence of the GATA1-complex (ongoing). Taken together, our observations suggest that the absence of Nsd1 results in formation of an aberrant GATA1 containing complex: establishing efficient immunoprecipitation protocols will help to dissect these complexes in MEL cells (ongoing).

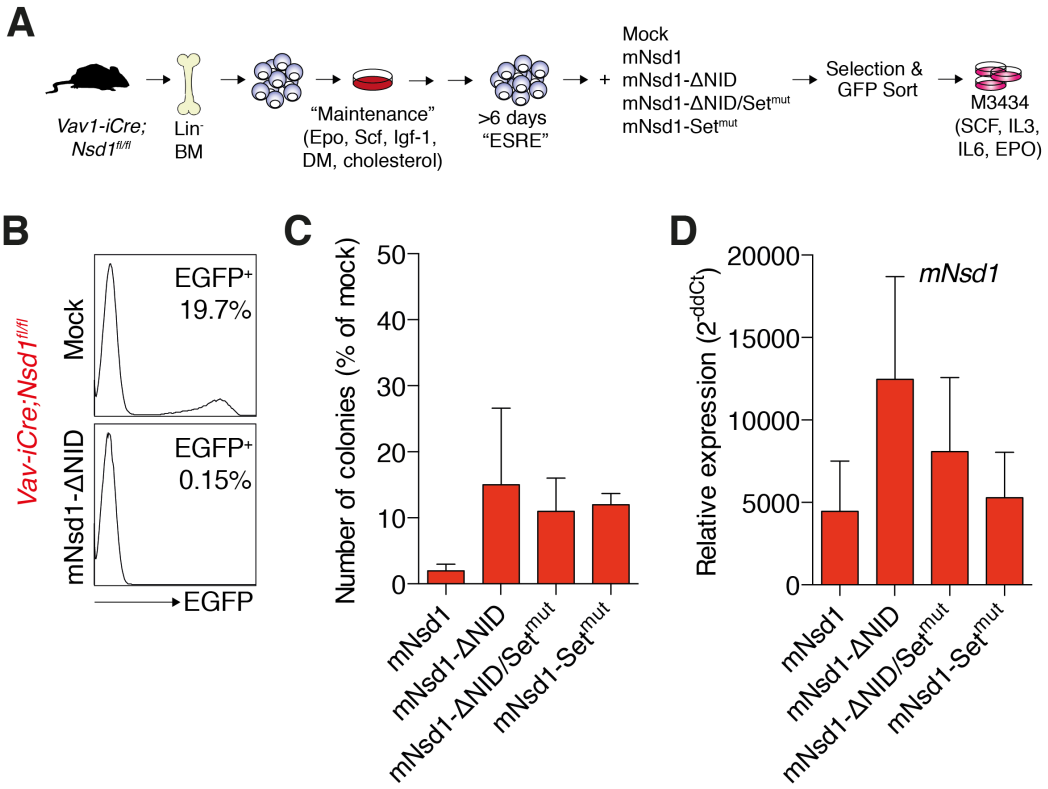


FIGURE 37. Retroviral overexpression of full-length and NSD1 mutants in pro-erythroblasts of diseased *Vav1-iCre;Nsd1$^{fl/fl}$* mice.

(A) Schematic depiction of experimental setup using lineage depleted bone marrow cells of *Vav1-iCre;Nsd1$^{fl/fl}$* that were kept in maintenance medium for more than six days to create stable ESRE cells. Cells were transduced in maintenance medium with pMSCV-pgk-puro-IRES-GFP, pMSCV-mNsd1-pgk-puro-IRES-GFP, pMSCV-mNsd1-ΔNID-pgk-puro-IRES-GFP, pMSCV-mNsd1- ΔNID /Set^{mut}-pgk-puro-IRES-GFP and pMSCV-mNsd1-Set^{mut}-pgk-puro-IRES-GFP constructs, EGFP^{+} cells sorted and kept under puromycin selection in methylcellulose M3434. **(B)** Percentage of EGFP^{+} cell population two days after pMSCV-pgk-puro-IRES-GFP or pMSCV-mNsd1-pgk-puro-IRES-GFP transduction of *Vav1-iCre;Nsd1$^{fl/fl}$* ESRE cells. **(C)** Relative number of colonies formed in methylcellulose M3434 of *Vav1-iCre;Nsd1$^{fl/fl}$* ESRE cells transduced with mNsd1 wildtype and mutants compared to mock transfected cells. **(D)** Quantitative real time PCR of *Nsd1 exon 5* after transduction of ESRE cells with mNsd1 wildtype and mutants. Ct values were normalized to *Gapdh* expression and shown as relative expression using 1/dCt method (n=2). Data presented as mean, error bars represent ±SD.

4.3.16 Attempts to rescue erythroid differentiation of *Vav1-ICre;Nsd1$^{fl/fl}$* erythroblasts by overexression of Nsd1

To demonstrate that blocked erythroid differentiation in *Vav1-iCre;Nsd1$^{fl/fl}$* cells is indeed the direct consequence of the lack of Nsd1 we tried to restore its activity by expression of wildtype or different Nsd1 mutants. We have generated retroviral constructs (*pMSCV*) expressing full-length Nsd1 and mutants lacking either the nuclear receptor interacting domain (NID), the SET domain or both. Proerythroblasts kept in maintenance culture for more than six days were transduced with either the empty vector (pMSCV-pgk-puro-IRES-EGFP = mock) or full-length Nsd1 or the different mutants: mNsd1-ΔNID, mNsd1-ΔNID/ΔSET^{mut} or mNsd1-ΔSET. We selected transduced cells with puromycin selected and plated sorted EGFP^{+} cells into methylcellulose (**Figure 37A**). Although we observed a decent transduction efficiency of about 20% with to control vector, transduction with any of the Nsd1 constructs was very poor reaching less than 1% (=1500 living EGFP^{+} cells per 10^6 starting cells) (**Figure 37B**). The size of the Nsd1 cDNA of about 8000bp seemed to significantly impair generation of viral particles. Nevertheless plating the small number of selected cells revealed significantly reduced colony formation upon transduction of wildtype or mutant Nsd1 compared to mock transduced cells (**Figure 37C**). Albeit the limitation of the small number of cells, it seemed that cells formed more colonies when transduced with the mutant than the wild type Nsd1 expressing virus. Quantitative RT-PCR revealed higher Nsd1 expression in transduced cells compared to controls (**Figure 37D**). Although limited by poor viral gene transfer these observation indicate that overexpression of Nsd1 significantly impairs colony formation of *Vav1-iCre;Nsd1$^{fl/fl}$* proerythroblasts. Further work by expressing smaller truncation mutants might be necessary to dissect the critical domains that are responsible for the effects.

4.3.17 Absence of Nsd1 alters global histone marks

Nsd1 has been previously characterized as a lysine methyltransferase of histone 3 (H3K36me$^{1/2}$) and linker histone H1, particularly on H1.5-K168. Whether Nsd1 also methylates H4K20 remains controversial (**Figure 38A**) [183,184]. Comparison of global histone alterations by western blotting of hematopoietic cells from diseased *Vav-iCre;Nsd1$^{fl/fl}$* mice to normal littermates revealed significant reduction of H3K36me$^{1-3}$. Whereas the global H3K9me^{3} mark was reduced, the H3K4me$^{2/3}$ activation marks were slightly increased (**Figure 38B**). To exclude that the observed differences are the consequence of different cellular composition due to severe tumor infiltration we also determined histone marks in CD71+ cells (enriched by magnetic beads) and found marked reduction of H3K36me^{2} and H3K9me^{2} (data not shown). To confirm altered histone methylation, we applied flow cytometry analysis and found 30% decrease of H3K36me^{1} and 60% decrease of H3K36me^{2} marks in lineage depleted bone marrow cells (**Figure 38C&D**). Moreover using H3K36me^{1} antibody on histone 1 we could also show a reduction of H1.5K168 methylation in *Vav-iCre;Nsd1$^{fl/fl}$* cells (**Figure 38E&F**).

In an attempt to dissect differential lysine methylation in *Vav1-iCre;Nsd1$^{fl/fl}$* cells we performed SDS-PAGE of cytoplasmic, nuclear and histone extracts of proerythroblasts grown in maintenance medium and probed the membranes with antibodies recognizing mono- or di/tri-methylated proteins (kind gift of A. Peters, FMI; [219]). Histone extracts did not show overall changes in mono- or di/tri- methylation (**Figure 38G**). However we observed increased signals for mono-methylated proteins 55-72 kDa and decreased bands between 43 and 55 kDa in *Vav1-iCre;Nsd1$^{fl/fl}$* proerythroblasts compared to *Nsd1$^{fl/fl}$* controls (**Figure 38H**). The blot for di/tri- methylation showed additional signals between 72- 130kDa and below 55kDa in cytoplasmic extracts and increase of signal around 55kDa in nuclear extracts (**Figure 38I**). Taken together, our data shows that absence of Nsd1 significantly alters posttranslational modification of its proposed histone substrates. In addition, the absence of Nsd1 may also influence protein methylation in cytoplasm and nucleus beyond histones. The nature of these substrates and whether Nsd1 directly regulates some of them remains to be investigated.

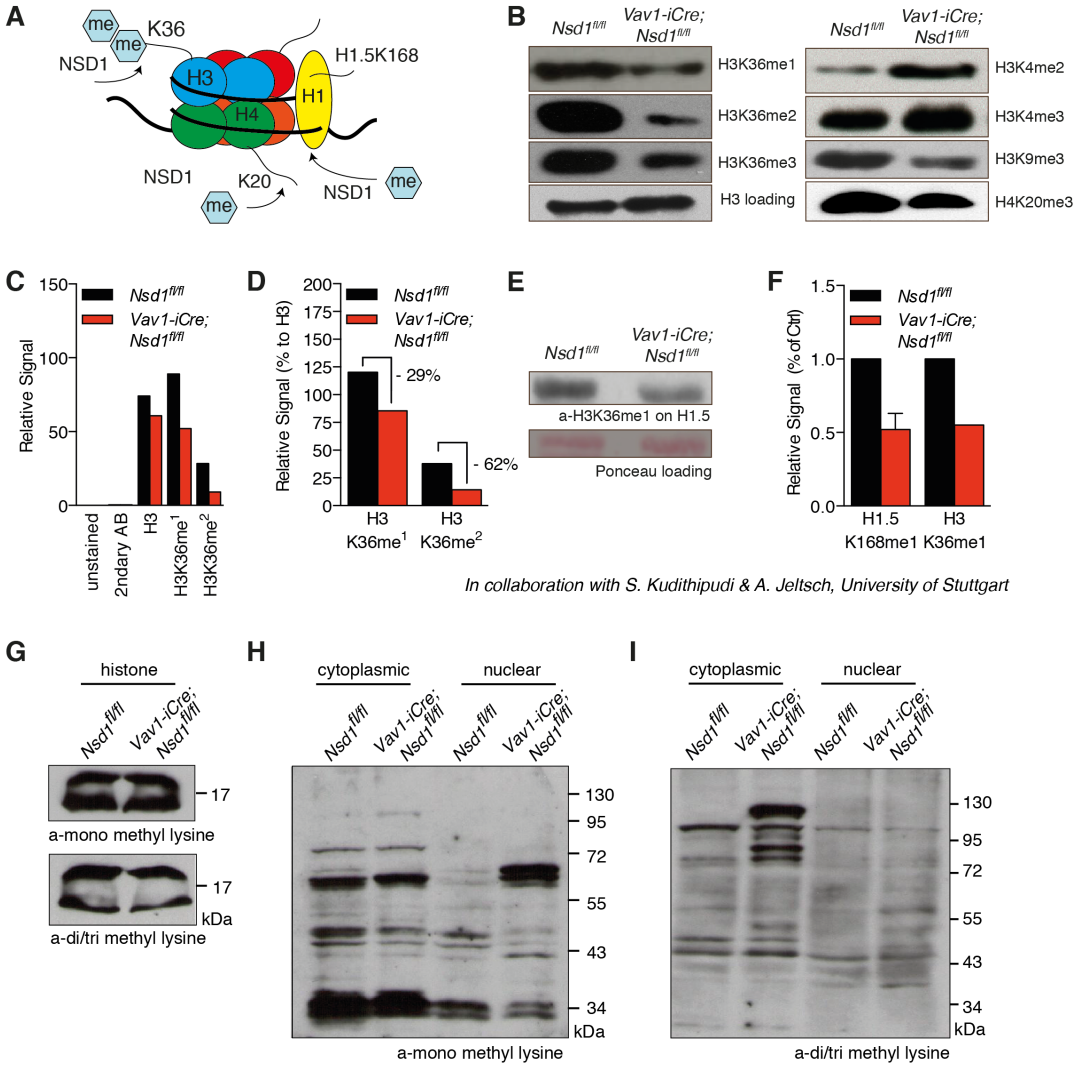

A
me
me
K36
NSD1
H3
H4
H1
H1.5K168
NSD1
K20
NSD1
me
me
B
Nsd1fl/fl
Vav1-iCre; Nsd1fl/fl
H3K36me1
H3K36me2
H3K36me3
H3 loading
H3K4me2
H3K4me3
H3K9me3
H4K20me3
C
Relative Signal
Nsd1fl/fl
Vav1-iCre; Nsd1fl/fl
unstained
2ndary AB
H3
H3K36me1
H3K36me2
D
Relative Signal (% to H3)
- 29%
- 62%
H3 K36me1
H3 K36me2
E
a-H3K36me1 on H1.5
Ponceau loading
F
Relative Signal (% of Ctrl)
H1.5 K168me1
H3 K36me1
In collaboration with S. Kudithipudi & A. Jeltsch, University of Stuttgart
G
histone
a-mono methyl lysine
a-di/tri methyl lysine
kDa
H
cytoplasmic
nuclear
a-mono methyl lysine
kDa
I
cytoplasmic
nuclear
a-di/tri methyl lysine
kDa

FIGURE 38. Global histone methylation in spleen and erythroblasts of *Vav1-iCre;Nsd1$^{fl/fl}$* mice.

(A) Schematic depiction of NSD1 histone methylation. **(B)** Western blot analysis of histone extracts obtained from spleen tissue of *Nsd1$^{fl/fl}$* and *Vav1-iCre;Nsd1$^{fl/fl}$* mice. Blots were probed with antibodies recognizing lysine mono- (me1) di- (me2) and tri- (me3) methylation on histone 3 and 4 (H3K36, H3K4, H3K9, H4K20). Histone 3 was used as loading control. **(C)** Relative signal intensity of lysine 36 mono and di- methylation on histone 3 (H3K36me^1, H3K36me^2) in lineage depleted bone marrow cells of *Nsd1$^{fl/fl}$* and *Vav1-iCre;Nsd1$^{fl/fl}$* mice. Unstained cells and secondary antibody only were used as negative controls, histone 3 as normalization control to calculate **(D)** relative signal intensity as percentage of histone 3 loading to quantify signal strength (n=1). **(E)** Relative signal of histone 1.5 lysine 168 mono- methylation (H1.5K168me^1) using antibody against histone 3 lysine 36 mono-methylation in spleen tissue of *Nsd1$^{fl/fl}$* and *Vav1-iCre;Nsd1$^{fl/fl}$* mice. Signal was normalized to Ponceau stained membrane as loading control and *Nsd1$^{fl/fl}$*. H3K36me^1 was used as technical control (n=2). **(F)** Representative image of histone 1.5 lysine 168 mono- methylation (H1.5K168me^1) using antibody against histone 3 lysine 36 mono- methylation in spleen tissue of *Nsd1$^{fl/fl}$* and *Vav1-iCre;Nsd1$^{fl/fl}$* mice. Signal was normalized to Ponceau stained membrane as loading control. **(G)** Western Blot analysis of mono- and di/tri- lysine- methylation of *Nsd1$^{fl/fl}$* and *Vav1-iCre;Nsd1$^{fl/fl}$* ESRE cells in maintenance medium in histone acid extracts, as well as in (H-I) cytoplasmic and nuclear extracts.

4.3.18 The NSD1 SET domain methylates GATA1 lysine residues 245, 246 and 308 on a peptide array

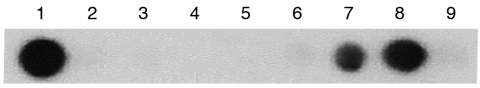
There is increasing evidence that NSD1 is able to methylate not only lysine residues of histones, but also of non- histone proteins [184,185]. We therefore asked whether the increased GATA1 protein levels observed in proerythroblasts lacking Nsd1 might be linked to aberrant methylation of its lysine residues (blue marks, **Figure 39A**). To address this question, we first performed *in vitro* methylation analysis of GATA1 peptides (spotted on a membrane) covering 9 pre- chosen lysine residues and found that the NSD1-SET domain was able to methylate a sequence located between the N- and C- finger domain (**Figure 39B**). We next aimed to dissect the whole GATA1 structure by designing 15-amino acid alternating peptides throughout the whole protein sequence and could not only confirm methylation of that N/C- interspacing region (K245, K246 and K252) but also identify another region behind the C finger (K308, K312, K314, K315 and K316) (**Figure 39C**). We next aimed to repeat rescue experiments in the *Vav1-iCre;Nsd1*$^{fl/fl}$ proerythroblasts and cloned GATA1 K246A and K252A (lysine to alanine mutations) into *pMSCV*. Both mutants were able to rescue aberrant colony formation as well as liquid differentiation of *Vav1-iCre;Nsd1*$^{fl/fl}$ proerythroblasts to the same extent than GATA1 WT transduced cells (**Figure 39D-F**). We also cloned GST-GATA1 fusions: however the NSD1-SET domain seemed unable to methylate recombinant GATA1 (whereas significant methylation of H1.5 (= control) was observed (not shown, performed by S. Kudithiphudi, Stuttgart). So far we could not confirm or exclude methylation of GATA1 or one of its complex partners regulating erythroid differentiation (ETO2, LBD1, LMO2, E2A).

A **NP_032115.1l erythroid transcription factor [Mus musculus]**

N 1-70	MDFPGLGALGTSEPLPQFVDSALVSSPSDSTGFFSSGPEGLDAASSSTSPNAATAAASALAYYREAEAYR
71- 140	HSP/VFQVYPLLNSMEGIPGGSPYASWAYGKTALYPASTVCPSHEDAPSQALEDQEGKSNNTFLDTLKTER
141 – 210	LSPDLLTLGTALPASLPVTGSAYGGADFPSPFFSPTGSPLSSAAYSSPKFHGSLPLAPCEARECVNCGAT
211 – 280	ATPLWRRDRTGHYLCNACGLYHKMNGQNRPLIRPKKRMIVSKRAGTQCTNCQTTTTTLWRRNASGDPVCN
281 – 350	ACGLYFKLHQVNRPLTMRKDGIQTRNRKASGKGKKKRGSNLAGAGAAEGPAGGFMVVAGSSSSGNCGEVA
351 – 413	SGLALGTAGTAHLYQGLGPVVLSGPVSHLMPFPGPLLGSPTTSFPTGPAPTTSSTSVIAPLSS –C terminus

N- Finger: AA 204-228
C- Finger: AA 258-282

B



Control Peptides
1. H3K36
2. H3K36A

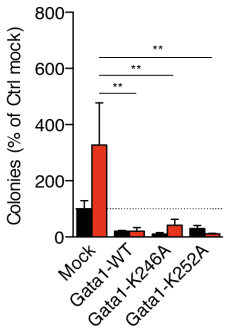
GATA-1 Sequence
3. YASWAYGKTALYPAS
4. ALEDQEGKANNTFLD
5. NTFLDTLKTERLSPD
6. SAAYSSPKFHGSLPL
7. RPLIRPKKRMIVSKR
8. KKRMIVSKRAGTQCT
9. DPVCNACGLYFKLHQ

C

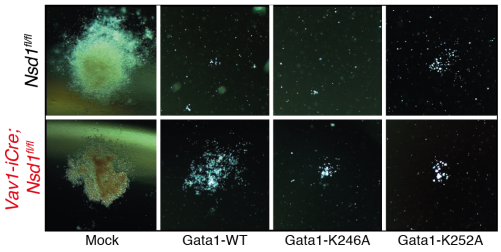
Lysine residue Nr.	Lysine methylation screen	Alanine Scan (K-> A)
137	x	
245	x	x (weak)
246	x	x (strong)
252	x	
308	x	x (strong)
312	x	
314	x	
315	x	
316	x	

In collaboration with S. Kudithipudi & A. Jeltsch, University of Stuttgart and S. Picaud & P. Filippakopoulos, SGC Oxford

D



E



F

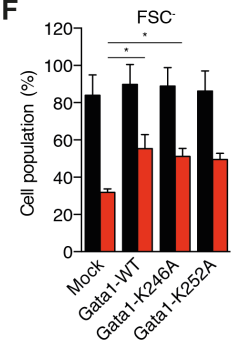


FIGURE 39. Analysis of GATA1 lysine methylation by NSD1.

(A) FASTA protein structure and accession details of murine GATA1. N- finger is highlighted in pink, C-finger in green and lysines in blue. **(B)** *In vitro* peptide methylation assay using NSD1 Set domain on membrane with peptide spots of selected lysine- containing structures of GATA1. **(C)** Summary of i*n vitro* peptide methylation and alanine scan assay using NSD1 Set domain on membrane with peptide spots of entire GATA1 protein sequence. **(D)** Relative number of colonies formed in M3434 methylcellulose by *Nsd1*$^{fl/fl}$ (black bar) and *Vav1-iCre;Nsd1*$^{fl/fl}$ (red bar) ESRE cells, kept for more than six days in maintenance medium, transduced in "maintenance medium with *pMSCV-pgk-puro, pMSCV-mGata1-pgk-puro, pMSCV-mGata1-K246A-pgk-puro,* or *pMSCV-mGata1-K252A-pgk-puro*. (n=4/group. ANOVA with Turkeys multiple comparison test. ** $p<0.01$). **(E)** Representative images of colony formation in M3434 methylcellulose by 5.000 plated *Nsd1*$^{fl/fl}$ (upper row) and *Vav1-iCre;Nsd1*$^{fl/fl}$ (lower row) ESRE cells, kept for more than six days in maintenance medium, transduced in "maintenance medium with *pMSCV-pgk-puro, pMSCV-mGata1-pgk-puro, pMSCV-mGata1-K246A-pgk-puro* or *pMSCV-mGata1-K252A-pgk-puro* (from left to right column). **(F)** Percentage forward scatter (FSC) negative cell population measured by flow cytometry in *Nsd1*$^{fl/fl}$ (black bar) and *Vav1-iCre;Nsd1*$^{fl/fl}$ (red bar) ESRE cells, kept for more than six days in maintenance medium, transduced in "maintenance medium with *pMSCV-pgk-puro, pMSCV-mGata1-pgk-puro, pMSCV-mGata1-K246A-pgk-puro* or *pMSCV-mGata1-K252A-pgk-puro* (n=3/group. ANOVA with Turkeys multiple comparison test. * $p<0.05$). Data presented as mean, error bars represent ±SD.

4.3.19 Reduced NSD1 expression impairs *in vitro* differentiation of human HSCs

Among the different hematopoietic lineages, erythropoiesis seems to be the one with the lowest degree of genetic homology between mouse and man [220]. To exclude the possibility that the Nsd1 may only regulate murine hematopoiesis, we explored the effect of altered NSD1 expression for differentiation of human HSCs. $CD34^+/CD38^-$ selected peripheral HSCs from three healthy donors were transduced with five lentiviral *NSD1*-miR-shRNA expression vectors and plated in methylcellulose favoring myeloid and erythroid differentiation. Transduction with three (353, 369, 372) out of 5 *NSD1* knockdown vectors resulted in reduced number of colonies compared to GFP or mock miR-shRNA transduced cells in the first plate. Interestingly, upon replating of the cells we observed the formation of multiple abnormally dense and rather large red colonies formed by erythroid progenitors only in the cells expressing the *NSD1* miR-shRNA 353 and 372 but not in the control or miR-shRNA 369 (**Figure 40A&B**). Despite several attempts, we were unable to grow these cells beyond the 2nd plate (data not shown). Abnormal colony growth was associated with a significant reduction of *NSD1* mRNA expression in shRNA 353 and 372 (**Figure 40C**). We decided to continue with miR-shRNA 372 giving us the most efficient knockdown, increased colony formation in second plating and therefore sufficient cells to analyze (**Figure 40D-F**). Cytospin preparations of colonies showed in contrast to the control composed of mostly monocytic cells, mainly erythroid progenitors (**Figure 40G**). Flow cytometry analysis of colonies confirmed erythroid differentiation by presence of $CD71^+$ and glycophorin A^+ (GPA) cells increased particularly in the second plating (**Figure 40H**). Quantitative real time PCR confirmed knockdown of *NSD1*mRNA below 50% in the second plating (**Figure 40I**). Very similar results were obtained upon *NSD1* knockdown in cord blood-derived $CD34^+$ HSCs (data not shown). This data suggests that reduction of NSD1 not only impairs erythroid differentiation of mouse but also of human HSCs.

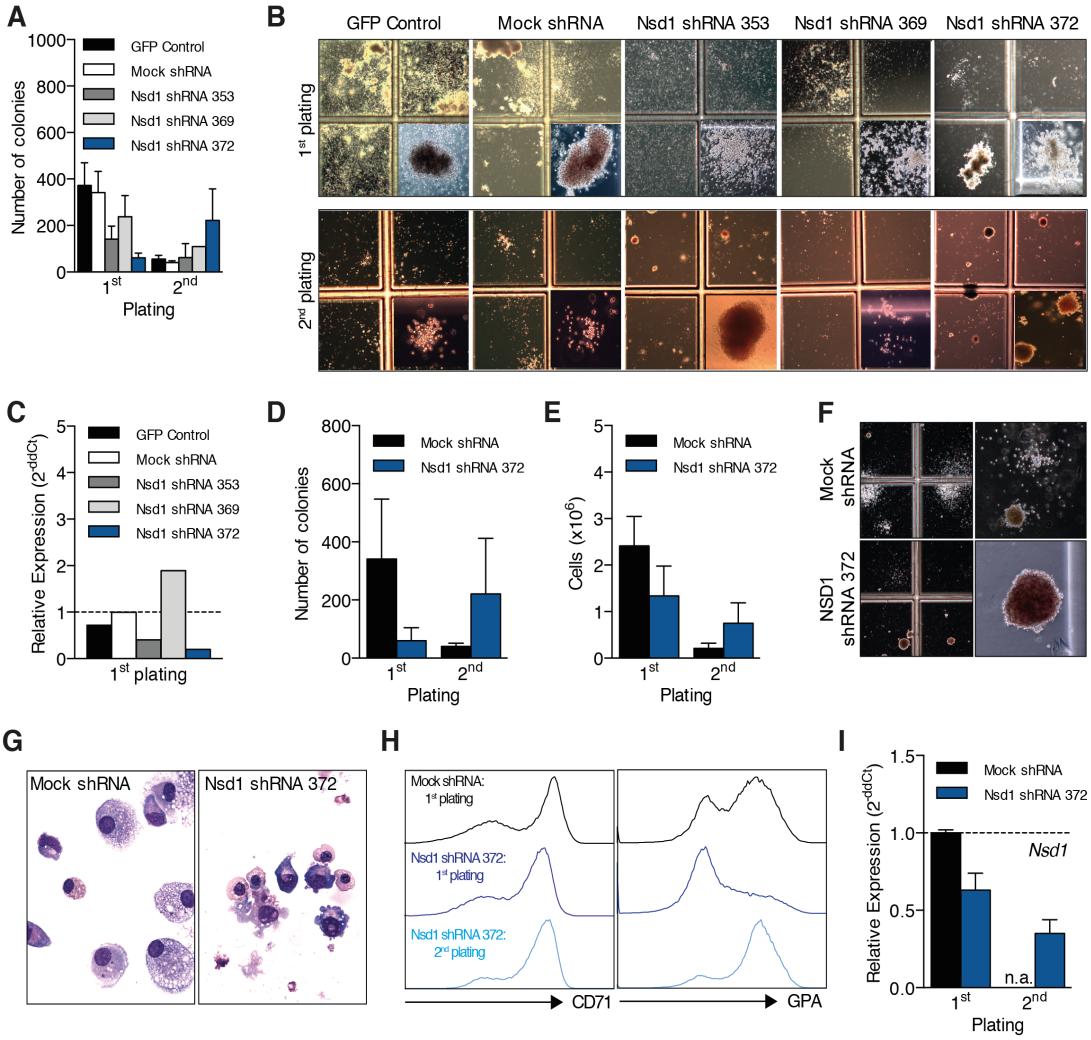

A
Number of colonies
GFP Control
Mock shRNA
Nsd1 shRNA 353
Nsd1 shRNA 369
Nsd1 shRNA 372
Plating
B
GFP Control
Mock shRNA
Nsd1 shRNA 353
Nsd1 shRNA 369
Nsd1 shRNA 372
1st plating
2nd plating
C
Relative Expression (2-ddCt)
1st plating
D
Number of colonies
Plating
E
Cells (x10^6)
Plating
F
Mock shRNA
NSD1 shRNA 372
G
Mock shRNA
Nsd1 shRNA 372
H
Mock shRNA: 1st plating
Nsd1 shRNA 372: 1st plating
Nsd1 shRNA 372: 2nd plating
CD71
GPA
I
Relative Expression (2-ddCt)
Nsd1
n.a.
Plating

FIGURE 40. Effects of NSD1 knockdown in human peripheral CD34$^+$ HSCs.

(A) Total number of colonies formed in methylcellulose H4434 by CD34$^+$ cells obtained from peripheral blood of healthy donors transduced in Stemline medium with pCMV-EGFP-pgk, pLKO.1 mock shRNA (shRNA Ctrl) or pLKO.1 shRNA targeting *NSD1* 353, 369 or 372 after first (n=5) and second plating (n=2 except for 369 n=1). **(B)** Representative images of first and second plating in methylcellulose H4434 CD34$^+$ cells obtained from peripheral blood of healthy donors transduced in Stemline medium with pCMV-EGFP-pgk, pLKO.1 mock shRNA (shRNA Ctrl) or pLKO.1 shRNA targeting *NSD1* 353, 369 or 372. **(C)** Quantitative real time PCR of *NSD1* in CD34$^+$ cells obtained from peripheral blood of healthy donors transduced in Stemline medium with pCMV-EGFP-pgk, pLKO.1 mock shRNA (shRNA Ctrl) or pLKO.1 shRNA targeting *NSD1* 353, 369 or 372 after first plating. Bars represent average relative expression normalized to *GAPDH* and shRNA Ctrl transduced cells (n=1). **(D)** Total number of colonies formed in methylcellulose H4434 by CD34$^+$ cells obtained from peripheral blood of healthy donors transduced in Stemline medium with pLKO.1 mock shRNA (shRNA Ctrl) or pLKO.1 shRNA targeting *NSD1* 372 after first and second plating (n=5). **(E)** Total number of cells formed by colonies in methylcellulose H4434 by CD34$^+$ cells obtained from peripheral blood of healthy donors transduced in Stemline medium with pLKO.1 mock shRNA (shRNA Ctrl) or pLKO.1 shRNA targeting *NSD1* 372 after first and second plating (n=5). **(F)** Representative images of first and second plating in methylcellulose H4434 CD34$^+$ cells obtained from peripheral blood of healthy donors transduced in Stemline medium with pCMV-EGFP-pgk, pLKO.1 mock shRNA (shRNA Ctrl) or pLKO.1 shRNA targeting *NSD1* 353, 369 or 372. **(G)** Representative images of colonies HE- stained resuspended cells formed in methylcellulose H4434 by CD34$^+$ cells obtained from peripheral blood of healthy donors transduced in Stemline medium with pLKO.1 mock shRNA (shRNA Ctrl) or pLKO.1 shRNA targeting *NSD1* 372 after first plating. **(H)** Flow cytometry histograms depicting CD71 and glycophorin A (GPA) marker expression in CD34$^+$ cells obtained from peripheral blood of healthy donors transduced in Stemline medium with pLKO.1 mock shRNA (shRNA Ctrl) or pLKO.1 shRNA targeting *NSD1* 372 after first and second plating. **(I)** Quantitative real time PCR of *NSD1* in CD34$^+$ cells obtained from peripheral blood of healthy donors transduced in Stemline medium with pLKO.1 mock shRNA or pLKO.1 shRNA targeting *NSD1* 372 after first and second plating. Bars represent average relative expression normalized to *GAPDH* and shRNA Ctrl transduced cells (n=5). Data presented as mean, error bars represent ±SD.

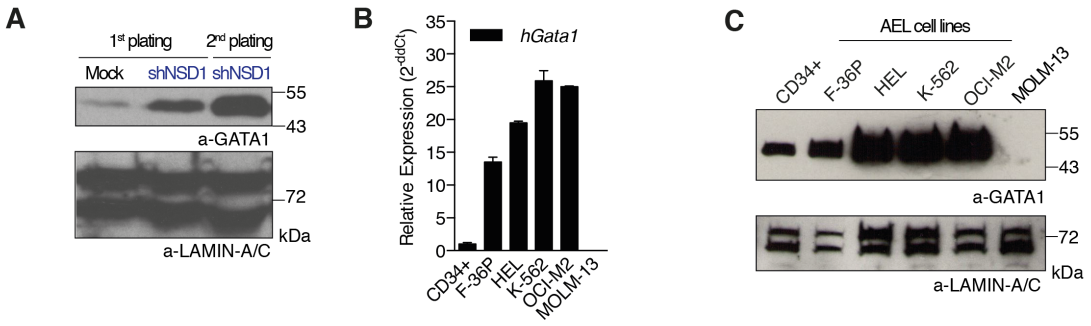
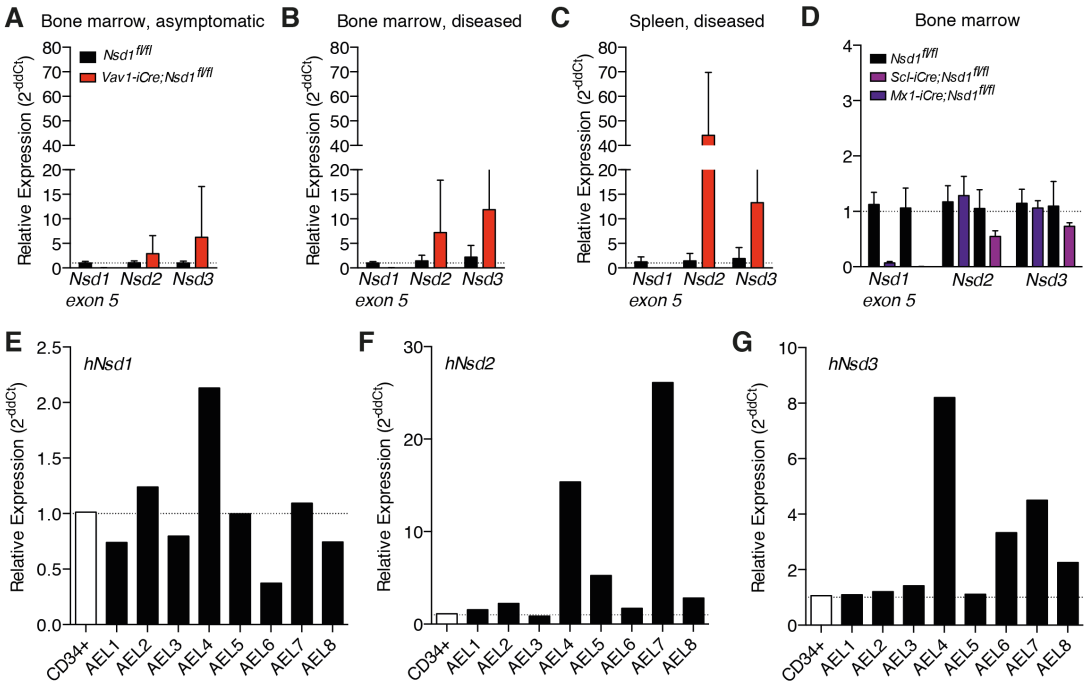


FIGURE 41. GATA1 mRNA and protein expression in human leukemic cell lines. (A) Western blot analysis of GATA1 protein nuclear expression levels in $CD34^+$ cells obtained from peripheral blood of healthy donors transduced in Stemline medium with pLKO.1 mock shRNA (shRNA Ctrl) or pLKO.1 shRNA targeting *NSD1* 372 after first and second plating. LAMIN-A/C was used as nuclear loading control. (B) Quantitative real time PCR of *GATA1* in $CD34^+$ cells obtained from peripheral blood of healthy donors and human leukemic cell lines. Ct values were normalized to *GAPDH* expression and shown as relative expression using 1/dCt method. MOLM13 cell line was used as biological negative control for GATA1 (n=1-2). (C) Western blot analysis of GATA1 protein nuclear expression levels in $CD34^+$ cells obtained from peripheral blood of healthy donors and human leukemic cell lines. MOLM13 cell line was used as biological negative control and LAMIN-A/C was used as nuclear loading control.

4.3.20 GATA1 in human leukemic cell lines

Finally, we aimed to test whether like in erythroblasts from *Vav1-iCre;Nsd1$^{fl/fl}$* mice, NSD1 knockdown in human HSCs would also result in changes of GATA1 expression. Western blot analysis of methylcellulose-derived cells revealed significantly increased of GATA1 protein expression in Nsd1 knockdown cells (**Figure 41A**). Since AEL is a rare disease with very limited access to primary cells we also determined GATA1 protein expression in 4 erythroleukemia cell lines with an erythroid phenotype [221]. Interestingly, all AEL cell lines did show increased levels of Gata1 mRNA and protein compared to another AML cell line (MOLM13) and control cells ($CD34^+$) (**Figure 41B-C**). We are currently aiming to amplify patient AEL cells in immunocompromised mice to address NSD1 and GATA1 protein expression. So far, we were able to show that similar to the phenotype in leukemic erythroblasts derived of *Vav1-iCre;Nsd1$^{fl/fl}$* mice, human AEL cell lines also express significantly increased levels of GATA1, whether GATA1 is responsible for the differentiation block remains to be studied.



H

	Diagnosis	Blasts %(from non eryth cells)
AEL 1	AML M6	>20
AEL 2	AML M6a	>50
AEL 3	AML M6	>30
AEL 4	AML M6b	>50 from erythropoesis
AEL 5	RAEB	20
AEL 6	AML M6	>50
AEL 7	AML M6	50
AEL 8	AML M6	80 from non erythropoesis

FIGURE 42. mRNA expression of NSD family members in *Nsd1$^{fl/fl}$* mouse models and human leukemia samples.

(A) Quantitative real time PCR of *Nsd1 exon 5, Nsd2 and Nsd3* in *Nsd1$^{fl/fl}$* respectively *Vav1-iCre;Nsd1$^{fl/fl}$* bone marrow samples of asymptomatic (n=6 resp. n=9), **(B)** diseased (n=13 resp. n=10) and **(C)** spleen (n=7 resp. n=4) samples of diseased mice. **(D)** Quantitative real time PCR of *Nsd1 exon 5, Nsd2 and Nsd3* in bone marrow samples of *Nsd1$^{fl/fl}$* (black bars), *Scl-iCre;Nsd1$^{fl/fl}$* (light purple bar) and *Mx1;iCre;Nsd1$^{fl/fl}$* (dark purple bar). Murine samples were normalized to *Gapdh* and control mouse tissue. **(E)** Quantitative real time PCR of *NSD1,* **(F)** *NSD2* and **(G)** *NSD3* in human AEL patient derived samples (n=1). Human samples were normalized to *GAPDH* and four healthy donor derived CD34^{+} cells. Values are shown as relative expression using 2^{-ddcT} method. **(H)** Table containing information regarding AEL patient diagnosis and blast content. Data presented as mean, error bars represent ±SD.

4.3.21 Does the lack of Nsd1 results in compensatory regulation of other Nsd family members?

As the other Nsd family members Nsd2 and Nsd3 seem to have at least partially overlapping histone substrates to Nsd1 we wondered whether the lack of Nsd1 would result in compensatory upregulation of the other family members. We first compared *Nsd2* and *Nsd3* mRNA levels in BM and spleen tissue of asymptomatic and diseased *Vav1-iCre;Nsd1*$^{fl/fl}$ with *Nsd1*$^{fl/fl}$ control mice by RT-PCR. Interestingly, we found that expression of Nsd2 and Nsd3 was on average higher in all tissues, but not statistically significant due to the high inter-individual differences between samples (**Figure 42A-C**). We tried to correlate disease severity (spleen weight) with gene expression but did also not obtain any significant correlation (data not shown). We also tested bone marrow of *Mx1-iCre* and *Scl-iCre;Nsd1*$^{fl/fl}$ mice and did not observe significant changes in gene expression of *Nsd2* and *Nsd3* (**Figure 42D**).

We wondered whether we might also be able to detect aberrant *NSD1-3* expression in cells from AEL patients. Hereby we obtained BM cell pellets and/or cDNA of 8 cases of MDS/AEL. The degree of infiltration varied significantly between different patients (**Figure 42H**). We compared *NSD1-3* mRNA levels to the average of RNA pooled from CD34^{+} cells four healthy donor CD34^{+} cells. We found decreased *NSD1* mRNA levels in 4/8 patients (**Figure 42E**). One AEL patient had *NSD1* levels below 50% (AEL patient 6). There was no strict correlation between lower levels of *NSD1* and higher levels of *NSD2* and *NSD3*. Interestingly, 2/8 patients seemed to express even higher levels of NSD1. Nevertheless some AEL patients had increased levels of *NSD2* (4/8) and *NSD3* (4/8) (**Figure 42F&G**). Although very limited by the heterogeneity of the samples, this analysis suggests aberrant expression of NSD1-3 in human leukemic disorders. However, we found no clear correlation between low NSD1 expression and AEL.

4.4 DISCUSSION

4.4.1 Ablation of Nsd1 impaired erythroid maturation during late fetal hematopoiesis

To study the role of *Nsd1* in normal hematopoiesis, we constitutively inactivated the gene in the blood forming system of the mouse using the *Vav1-iCre* transgenic ablator strain (**Figure 12**). Activation of *Vav1-iCre* starts in fetal liver starting at around E12.5-13.5 [205,222]. In addition we intercrossed the targeted *Nsd1$^{fl/fl}$* mice to transgenic strains (*Mx1-iCre* and *Scl-iCre^{ER}*) to conditionally inactivate the gene during adult hematopoiesis. In *Vav1-iCre;Nsd1$^{fl/fl}$* mice, the floxed *Nsd1 exon 5* was almost completely excised (as measured by RT-PCR) in fetal and adult tissue (**Figures 12 & 27**). In contrast, activation of *Mx1-iCre* and *Scl-iCre^{ER};Nsd1$^{fl/fl}$* by repetitive injection of poly-(pIC) and Tamoxifen respectively did not significantly reduce the floxed allele sequence in peripheral blood cells (**Figures 25 & 26**). However, the relative expression levels of *Nsd1 exon 5* mRNA were reduced to 1-10% in the BM of all models (**Figure 42**). It is tempting to speculate that the PCR of peripheral blood cells does not fully reflect the ablation status of the BM stem- and progenitor cells. Peripheral blood after red blood cell lysis contains mostly platelets and differentiated white blood cells of the myeloid (15% of WBC) and lymphoid (80% of WBC) lineage while erythrocytes and reticulocytes are destroyed during the lysis process, and platelets do not contain DNA. Therefore the signal of the cleavage PCR largely results from lymphoid cells. Red blood cell lysed BM contains in physiological situations approximately 55% Mac-1/Gr-1+ myeloid progenitor cells, 20% B220+ B cells, 10% CD71+ erythroblasts, 5% CD3+ T cells and 0.5% CD41+ Megakaryocytes. Only 1.5% are myeloid progenitors and 0.1% are stem cells (= total 92%). Moreover, memory immune cells can survive a long time in the periphery without being affected by bone marrow constitution [223,224]. Therefore, it remains to be investigated whether Nsd1 was fully ablated in all hematopoietic cells in *Mx1-* and *Scl-iCre;Nsd1$^{fl/fl}$* mice.
In contrast, in *Vav1-iCre;Nsd1$^{fl/fl}$* mice with complete excision of the floxed Nsd1 alleles we already observed a significant accumulation of erythroid progenitor cells during fetal liver hematopoiesis (**Figure 27**), progressing into a malignant AEL- like phenotype in adult mice at around 12 weeks after birth (**Figure 13**). Defective hematopoiesis was reflected by aberrant surface marker expression and blood

counts in the peripheral blood of diseased mice whereas the counts in conditional ablators remained largely stable (**Figures 13, 14, 25 & 26**). Nevertheless, some signs of defective erythropoietic maturation were observed in the periphery in all models, but not found in the BM of induced *Mx1-* and *Scl-iCre*^{ER} mice (**Figures 13, 25 & 26**). *Mx1-* and *Scl-iCreER* mice also did not develop any disease symptoms within a year of observation. For this reason, one could speculate that Nsd1 plays a critical role during fetal liver hematopoiesis and ablation in a certain window of opportunity might be necessary to induce observed phenotype. As outlined earlier, HSC maintenance and erythroblast island formation are major processes of fetal liver hematopoiesis (**Figure 1**). Interestingly, the phenotype in *Vav1-iCre;Nsd1*$^{fl/fl}$ mice starts to develop during fetal liver hematopoiesis, but needs about 12 weeks (=adult hematopoiesis in mouse) to propagate into a symptomatic and ultimately lethal disease. Future experiments will compare the effect of conditional *Nsd1* ablation on the number of HSC *in vivo* and erythroblast maturation *in vitro.*

Investigating mRNA expression, we could not find sufficient indices for a potential upregulation of other NSD family members, NSD2 and NSD3, in hematopoietic cells (**Figure 42**). Whole bone marrow extracts revealed increased mRNA expression of Nsd2 and Nsd3 in some *Vav1-iCre;Nsd1*$^{fl/fl}$ mice with a high deviation. Remarkably, higher expression was mostly seen in spleen cells of diseased animals. Importantly, gene expression signatures of LSK, MEP and GMP that were derived of bone marrow of diseased mice did not reveal expression changes. Moreover, we found global decrease in histone methylation suggesting that NSD2 and NSD3 do not have redundant roles in hematopoiesis (**Figure 38**).

4.4.2 Ablation of Nsd1 resulted in an AEL-like phenotype

All *Vav1-iCre;Nsd1*$^{fl/fl}$ mice with constitutive inactivation of *Nsd1* starting during fetal liver hematopoiesis developed a lethal disease characterized by hepato-splenomegaly, multi-organ infiltration as well as thrombocytopenia (**Figures 13 & 14**). The mice also had a significantly reduced number of HSCs and accumulation of CD71$^{dim/+}$/TER119^{-} erythroid progenitor cells (**Figures 14 & 16**). Expansion and aberrant *in vitro* clonogenic activity of *Nsd1* null erythroid progenitors was EPO- and SCF- dependent (**Figure 15**). In addition, the disease phenotype could be

transferred to irradiated recipients by transplanting BM of *Vav1-iCre;Nsd1$^{fl/fl}$* (**Figure 17**).

Vav1-iCre;Nsd1$^{fl/fl}$ proerythroblasts accumulating in diseased mice showed many similarities to cells derived from one of the earliest AEL model that was published in 1957 by Charlotte Friend. She described a murine retroviral model involving the "Friend" virus, which is a complex of the *Friend Spleen focus virus* (*SFFV*) and a replication- competent *Friend murine leukemia virus* (*F-MuLV*) [87,225]. In this model, accumulation of erythroblasts, called "erythroblastosis", developed as early as 10 days after injection of the virus. The *SFFV* produces a p55 glycoprotein (gp55) that binds to the Epo receptor (EpoR) and to the short form of the stem cell kinase receptor (sf-STK) [226]. This results in EPO-independent proliferation orchestrated through activation of JAK/STAT signaling. The second step of the disease is marked by a block in differentiation due to integration of long terminal repeat (LTR) transcriptional enhancers of the *SFFV* into the promoter region of *Spi1* (spleen focus forming virus proviral integration site 1) encoding for the PU.1 transcription factor [225,227]. In addition, the F-MuLV has been shown to integrate into the *Fli-1* gene causing AEL with a latency of six weeks [228,229].

Several stable "Friend" mouse erythroleukemia cell lines (MEL) have been derived of which the MEL cell clone 745-PC–4 is the most commonly used one (often simply referred as MEL cells) [230]. These cells expresses aberrantly high PU.1 protein levels but also very high levels GATA1 protein (**Figure 30 & 31**) [208]. Multiple studies have suggested that PU.1 most likely physically binds to and functionally interferes with GATA1 resulting in impaired activation of GATA1 targets that are necessary for terminal erythroid maturation [25,27,67–69,89,231,232]. However, the reason for the aberrant high level of GATA1 protein remains unknown.

Due to the discovery of retroviral integration of *SFFV* into the *Spi-1* locus, a *Spi-1* transgenic mouse model was developed. About 50% of mice overexpressing *SFFV-LTR* driven *Spi-1* developed erythroleukemia between 1.5 and 6 month after birth. The disease was characterized by anemia, proerythroblast accumulation and splenomegaly. Later, proerythroblasts seemed to autonomously expand due to activating mutations in stem cell factor receptor *c-kit* [88]. Important to note is the fact that neither viral integrations nor *c-kit* mutations are molecular hallmarks of human AEL [233,234].

Interestingly, the group of Dan Tenen could identify a –14-kb upstream regulatory region (URE) in the *Spi-1* locus that regulates its expression. If ablated, PU.1 protein

expression decreased to 20% of its normal levels and resulted AML, but not AEL in a transgenic mouse model. Interestingly, retroviral overexpression of *Spi-1* could rescue the myeloid differentiation block of the cells [91,235]. Another group demonstrated that a fraction of mice lacking the gene for the lymphoid specific helicase (Lsh, also known as HELLS) developed erythroleukemia. Lsh is a member of the SNF2 chromatin remodeling family and involved in *de novo* DNA methylation. Interestingly, the lack of *Lsh* was associated with hypomethylation at repetitive sequences and retroviral elements within gene locus. Moreover, the authors found in *Lsh* deficient mice decreased levels of Dnmt3b binding at the Pu.1 gene locus finally resulting in moderately elevated PU.1 protein levels [236]. In addition, Li *et al*. could demonstrate that PU.1 protein overexpression through retroviral overexpression of miR-92a is involved in development of murine erythroleukemia [237].

Next to the above-mentioned mouse models, avian leukemia viruses have been described to cause erythroid leukemia-like and sarcoma phenotypes in chicken. Hereby, the viral oncogenic counterparts for v- *Myb* and -*Ets* (in the E26 virus strain) or *ErbA* and *ErbB* proto-oncogenes (in the avian erythroblastosis virus strain) could be identified as the driver of leukemogenic transformation. When E26 was used to infect chicken, it caused mixed erythro- myeloid leukemia and resulting EPO-dependent cell lines resembled erythroid precursors [238]. Interestingly, infection of IL3-dependent murine cell lines with murine E26 (ME26) induced *Gata1* mRNA expression suggesting involvement in the blocked maturation phenotype [228].

In another transgenic mouse model, expression of the human *c-myc* proto-oncogene under the control of *Gata1* promoter elements caused rapid onset erythroleukemia, The EMY cell line obtained from diseased mice also expresses high GATA1 protein levels [239]. In addition, another group reported occurrence of GATA1-expressing erythroleukemic cells derived of *Zeta-Globin-V-Harvey-Ras* transgenic mice [240].

Male *Gata1* knockout embryos die during gestation due to defective erythropoiesis and accumulation of cells at the proerythoblast stage [39]. However, genetic modification of the murine *Gata1* locus in female mice, leading to 5% of normal expression, resulted in an AEL- like phenotype with 50% penetrance after a rather long latency of median 143 days. Similar to Nsd1 null mice, these so-called "*Gata1*$^{1.05/X}$" transgenic mice accumulated CD71^{+}/TER119^{-}/c-Kit^{+} erythroid blast in multiple organs. In addition, the erythroid differentiation arrest could also be restored upon retroviral overexpression of Gata1 [73]. In addition to the above-mentioned

models, other genetically modified mouse lines have been reported to develop an erythroleukemia-like phenotype. Interestingly, in most of the models aberrant expression of GATA1 and/or PU.1 was observed (**Table 10**).

There are several particular features of the *Vav1-iCre;Nsd1$^{fl/fl}$* AEL mouse model to be discussed. First of all, 100% of the mice lacking Nsd1 develop the AEL- like phenotype. To date, all mice developed symptoms/signs of disease after a median latency of 85 days with anemia, reticulocytosis, hepato-splenomegaly with multi-organ infiltration of CD71$^{dim/+}$/ TER119^{-}/ c-Kit^{+} expressing cells (**Figure 13 & 14**). Our experiments suggest that the disease in *Vav1-iCre;Nsd1$^{fl/fl}$* mice originates from cells of the Lin^{-}/c-Kit^{+} and most probably CD71$^{-/dim}$ compartment going along with previous reports suggesting that the target cell for leukemic transformation is a very early c-kit-positive erythroid precursor that does not express fully CD71 or TER119 at all (**Table 10**, **Figure 18**). Notably, at this developmental stage of erythroid maturation, GATA1 signaling is essential to induce erythroid maturation of progenitors [70]. Hereby, GATA1 increases the expression levels of the EPO receptor (EPO-R) starting at late CFU- E stage serving as enhancer signal once EPO is produced due to hypoxia signals in the body [241,242]. Binding of the EPO-R transmits signals via the JAK2/STAT pathway resulting in upregulation of multiple target genes including the transferrin receptor CD71, the survival protein BCL2L1, hemoglobin and other drivers of erythroid maturation. Since proliferation and *in vitro* colony formation of Nsd1 null proerythroblasts was fully depended on EPO, we concluded that these cells were arrested in a maturation stadium between BFU-E and CFU-E (**Figures 15 & 30**).

As previously described, acute leukemia is the product of a differentiation block combined with a proliferative advantage of hematopoietic stem and progenitor cells (**Figure 7**). Based on the published models we conclude that the AEL phenotype in mice is characterized by a blocked erythroid differentiation resulting in accumulation of erythroblasts, that were often EPO- dependent. The differentiation block is caused by aberrant expression or action of transcription factors involved in myelo- erythroid differentiation pathways. Moreover, some groups reported secondary events such as loss of tumor suppressor p53 or activating c- Kit mutations providing a growth advantage to the cells [243]. In conclusion, our observations collectively suggest that *Vav1-iCre;Nsd1$^{fl/fl}$* mice developed a phenotype resembling human acute erythroleukemia (AEL) in many aspects.

TABLE 10. Acute Erythroleukemia Mouse Models.

Category	Model	Phenotype	REF
Avian Leukemia Viruses	E26 (*Myb-Ets* Encoding)	AEL & Sarcoma In Chicken	238
	AEV (Mutated *Erb-A/B*)	AEL & Sarcoma In Chicken	244
Murine Retroviruses	Friend Virus (SFFV, F-Mulv)	Erythroblastosis & AEL	87
	Graffi (Gv1.4) Murine Retrovirus	AEL & AMKL	245
Transgenic Mice	Gata1-Cmyc Tg	AEL	239
	Pu.1/Spi Tg	AEL	225
	Gata1 KO	Proerythroblast arrest	246
	Zeta-Globin-V-Ha-Ras Tg	AEL	240
	Spi1 Knockdown Tg (Ablation Of Distal Enhancer)	AML (Myeloblast) & T-Cell Lymphoma	91
	GATA1-$^{1.05/X}$ Tg	AEL & ALL	73
	Lsh KO	AEL	247
	VavCre+;Asxl1$^{fl/fl}$	MPN/MDS	248
Non- transgenic mice	ERG (Mscv-LTR, BMT)	Lymphoid Leukemia/ Erythro-Megakaryocyte Leukemia	249
miRNA	Mir-19a; Mir-92a (LTR, BMT)	B- Cell Hyperplasia, AEL	237
Exogenous Factors	X- Rays (300rad) C3H Mice	Transplantable AEL	250

4.4.3 The lack of Nsd1 altered protein levels of several key erythoid transcription factors

To obtain mechanistic insights into the AEL-like phenotype in *Vav1-iCre;Nsd1*$^{fl/fl}$ mice we compared genetic signatures of flow-sorted hematopoietic stem and progenitor cells (LSK, MEP, GMP). Hereby we found reduced Gata1 mRNA expression levels in *Vav1-iCre;Nsd1*$^{fl/fl}$ LSK compared to littermate controls (**Figure 28**). We also found that *Vav1-iCre;Nsd1*$^{fl/fl}$ proerythroblasts expressed lower Gata1 mRNA levels and were significantly impaired in regulation of GATA1 target genes linked to erythroid maturation (**Figures 30 & 34**). Strikingly, GATA1 protein was constitutively expressed at high levels in these cells (**Figures 29 & 30**). Even more surprisingly, we found that retroviral overexpression of a full-length mGata1 cDNA was able to complete erythroid maturation in *Vav1-iCre;Nsd1*$^{fl/fl}$ proerthroblasts (**Figures 32-36**). Several previous studies have shown that GATA1 forms activating and repressing multi-protein complexes that act as master regulators of transcription of genes for erythroid maturation [63,65,66]. GATA1 has different functions depending on the interaction partners and the cell context. Interaction with SCL, E2A, LDB1 and LMO2 results in formation of a pentameric activating complex occupying regulatory elements of erythroid differentiation-associated genes [32,63,65,66]. Interestingly, we found decreased amounts of SCL, E2A and LDB1 protein in nuclear extracts of *Vav1-iCre;Nsd1*$^{fl/fl}$ proerythroblasts (**Figure 36**). GATA1 has also been described to form repressive complexes in which the ETO2 protein seems to be a central interaction partner [65,251]. ETO2, also called CBFA2T3 or MTG16, is a transcriptional co-repressor protein that does not bind to DNA, but recruits HDACs to chromatin [252–254]. Interestingly, we found significantly increased ETO2 protein levels in pro-erythroblasts and differentiating erythroblasts [255]. As we found several GATA1 target genes to be properly down regulated during induced erythroid maturation suggests preferred formation of such a repressive complex in absence of NSD1 (**Figure 34**). In addition, gene set enrichment analysis (GSEA) of the expression signatures from Nsd1 null stem and progenitor cells (LSKs) revealed some correlation with up- and downregulated genes in stem and progenitor cells reported in *Eto2* knockout mice again suggesting activity of such a repressive complex containing ETO2 and GATA1 (**Figure 35**). To provide formal proof of concept we are currently performing immunoprecipitation experiments to demonstrate interaction of GATA1 and ETO2 in presence and absence of Nsd1. In addition, we will determine the impact of

overexpression and down regulation of GATA1/ETO2 on complex formation and target gene expression. Although we were not able to experimentally close the ring yet, our data so far suggests that the absence of Nsd1 leads to aberrant formation of GATA1 containing complexes favoring repressive over active conformations.

The links of blocked erythroid differentiation and AEL in Nsd1-null mice to ETO2 is particularly interesting as ETO2 has been identified being involved in a chromosomal translocation t(1;16)(p31;q24) present as sole cytogenetic abnormality in rare cases of pediatric AEL [256]. This translocation leads to expression of a fusion between nuclear factor IA (NFIA) and ETO2. NFIA has been previously characterized as a transcriptional regulator of erythro-megakaryoblastic fate of BM-derived cells. It is upregulated during erythroid differentiation whereas it remains low in granulocytes [257]. In addition, ETO2 was also found in a translocation t(11;21)(q24;q22) leading to expression of a RUNX1/ETO2 fusion which seems to support leukemogenesis by blocking differentiation through transcriptional repression of RUNX1 target genes in rare cases of AML [258]. ETO2 is also rearranged by an Inv(16)(p13.3q24.3) leading to expression of an ETO2-GLIS2 fusion in pediatric patients with acute megakaryoblastic leukemia (AMKL) [259]. These data suggest that aberrant formation of ETO2 and/or GATA1 complexes are key regulators of normal and malignant erythro-megakaryocytic differentiation.

We found aberrantly high constitutive GATA1 protein levels in human $CD34^+$ HSC blocked at the proerythroblastic stage after shRNA-mediated knockdown of NSD1 but also in human AEL cell lines suggesting a common mechanism. Whether the cells also express high ETO2 levels and form aberrant GATA1 complexes is currently under investigation. Ultimately we will perform chromatin immunoprecipitation (ChIP) experiments to quantify differences in genome DNA binding properties of GATA1 and ETO2. However, this still leaves us with the question, how the lack of Nsd1 results in aberrant GATA1 and ETO2 protein accumulation and complex formation.

4.4.4 Is non-histone protein methylation by NSD1 involved in blocked erythroid differentiation and AEL?

An increasing number of methyltransferases that were initially found as histone modifiers seem to also recognize non-histone proteins as substrates [184,185]. Very recently, another H3-K36 histone methyltransferase SETD2 (also known as HYPB or KMT3A) was proposed to play a dual role in chromatin as well as cytoskeletal remodeling through methylation of alpha-tubulin at lysine 40. Interestingly, the same lysine mark can also be acetylated in the context of microtubule organization. The authors could demonstrate that methylation of alpha-tubulin mainly takes place during mitosis and cytokinesis and ablation of the enzyme caused severe cytoskeletal problems [260].

Aberrant constitutive high level GATA1 protein expression led us to think about non- histone methylation in the *Vav1-iCre;Nsd1*$^{fl/f}$ mouse model. We used peptide arrays and alanine-scans to identify three potential methylation sites (K245, K246 and K308) of GATA1 (**Figure 39**). Whether the SET domain of NSD1 methylates the full-length protein *in vivo* remains to be investigated. First attempts using GST-GATA1 fusions did not confirm methylation by the NSD1 SET domain *in vitro* (not shown). However, the GST-Tag might interfere with the conformation of the protein and/or the bacterially produced protein might lack critical glycosylation. We therefore plan to use reticulocyte lysates expressing GATA1 to be used as a substrate for recombinant NSD1. Alternatively, we could co-express both proteins in cells of the erythroid lineage, however, some preliminary experiments revealed very poor overexpression of the *Nsd1* cDNA of >8kB by retroviral vectors. Nevertheless we will try to transiently co-express *Nsd1* and *Gata1* by *CMV*-driven expression vectors in HEK-293 or COS cells to address protein interaction and potential methylation.

Methylation of GATA1 and its potential functional impact have not been reported so far. Based on other examples, one could speculate about a complex interplay between methylation and several signaling pathways. It has been previously described e.g. for the tumor suppressor p53 protein that lysine methylation might serve as a signal for other post- translational modifications, influence protein- protein interaction, change the stability of the protein, change its cellular localization or change binding to certain promoters [180]. The identified potentially NSD1-methylated lysine residues in GATA1 are highly conserved among species and were previously

shown to serve as binding sites for transcriptional co-activators like CBP and p300. Therefore one could speculate that methylation interferes with acetylation and activation of GATA1 [58,261,262]. Interestingly, in this regard was the previous observation that the transcriptional co-regulator bromodomain containing protein 3 (BRD3) recognizes acetylated GATA1 to be recruited to chromatin and to facilitate erythroid target gene expression [263]. In a first trial experiment, we observed increased BRD3 levels in nuclear extracts derived of *Vav1-iCre;Nsd1*$^{fl/fl}$ erythroblast (not shown). Whether the increased levels result from unbound BRD3 to chromatin remains to be investigated.

We also observed overall significantly altered protein methylation profile in the nucleus as well as in the cytoplasm of *Vav1-iCre;Nsd1*$^{fl/f}$ proerythroblasts compared to normal controls (**Figure 38**). For future studies, it would be interesting to identify the nature of these mono or di/tri-methylated proteins and characterize those that are directly methylated by NSD1. Our collaborators Kudithipudhi *et al.* found that the zinc finger and BTB domain-containing protein 16 (ZBTB16, also known as PLZF) is a direct target of NSD1 methylation in HEK-293 cells [184]. Interestingly, this protein was previously shown to interact with GATA1 and also linked to platelet regulation by induction of Tpo receptor (TpoR) [264]. We found PLZF mRNA significantly downregulated in *Vav1-iCre;Nsd1*$^{fl/fl}$ LSK cells (data not shown). Whether PLZF is differentially regulated at the protein level in *Vav1-iCre;Nsd1*$^{fl/f}$ differentially regulated remains to be investigated. It also remains to be studied whether PLZF is methylated in murine erythroblasts or other hematopoietic cells. A major obstacle hereby is to isolate nuclear proteins in a way to preserve extraction of histone proteins since antibodies preferentially bind methylated histones or to optimize the protocol to subtract histone methylation signals (**Figure 38**).

4.4.5 Altered histone methylation upon ablation of Nsd1

We found a global reduction of histone H3K36 methylation in *Vav1-iCre;Nsd1$^{fl/f}$* proerythroblasts. In addition we saw a reduction of H3K9me3 and an increase of H3K4me2/3 histone marks suggesting a globally increased transcriptional activity (**Figure 38**). Previous reports suggested that NSD1 mainly methylates H3K36me1/2 and H1.5K168 histones residues [183,184,186,187]. However, the observed significant reduction of the H3K36me3 marks suggests that the lack of the mono- and di-methylation marks set by NSD1 will clearly impair adding the tri-methylation marks by SETD2/HYPB. In fact, it has been experimentally demonstrated that H3K36me1 and me2 are needed as substrates in order to set the me3 mark [183]. The biological significance of H3K36 methylation is not fully understood. Decreased di- methylation was found on transcriptionally silent chromatin and increased levels at open reading frames. In general, this mark has been associated with transcriptional elongation and therefore marks actively transcribed genes [265]. Moreover, H3K36me3 levels have been shown to play an important role in cell cycle transition, DNA repair and to maintain genomic stability [266–269]. Recurrent SETD2 mutations have been reported in various hematological malignancies (including AML) but also in solid cancer resulting in a loss-of-function and therefore decreased H3K36me3 levels [270]. Interestingly, the H3K36me3 is recognized by the PWWP domain of DNA methyltransferases like DNMT3 (**Figure 43**) [271]. In fact, ES cells lacking SETD2 show decreased H3K36me3 levels and impaired recruitment of DNMT3B1 to active genes [272]. In addition, DNMT3A was also shown to have high affinity for H3K36me3 *in vitro* [151]. Both DNMT3A and -B have been shown to interact with PU.1 by forming a complex involved in *de novo* methylation of DNA and regulation of transcription [273]. Therefore, one could speculate that decreased H3K36me3 methylation initiated by ablation of *Nsd1* also influences DNMT3 binding and thereby prevents DNA methylation. This hypothesis would imply that certain loci remain unmethylated and therefore "open". Supporting this hypothesis is the observation that in blood samples of SOTOS patients (that bear heterozygous most likely dominant-negative acting mutations of *NSD1*), decreased genome-wide DNA methylation at promoter regions was observed [274]. These data implies to further study DNA methylation in *Vav1-iCre;Nsd1$^{fl/fl}$* in an unbiased way taking advantage of whole genome bisulfite sequencing approaches. Depending on the research question to be answered, it would be advisable to aim for a stable model generating large amount of material to optimize and reproduce data.

A potential model could be embryonic stem (ES) cells obtained from our conditional *Nsd1* knockout mice that can be induced into different hematopoietic and also non-hematopoietic lineages using published protocols [275–277].

Moreover, it remains to be elucidated whether NSD1 binds to particular gene loci to regulate transcription. Interestingly, NSD1 was found within the nucleus, but excluded from nucleoli and condensed heterochromatin suggesting a role in chromatin remodeling [169]. Therefore ChIP would be beneficial to investigate direct consequences. However, as there are actually no commercially available reliable ChIP-grade antibodies, we have taken advantage of a “homemade” anti-NSD1 antibody (kind gift of A. Peters, FMI Basel) that nicely recognizes the C terminus of NSD1 in Western blots. We are currently exploring the utility of this antibody for optimizing ChIP. Another direction would be to tag endogenous Nsd1 by genome editing using Crispr/Cas9. In fact, preliminary experiments by a colleague have provided proof of feasibility to add a triple FLAG tag at the C-terminus of the *Nsd1* ORF in MEL cells. Ultimately it would be important to study the consequence of *Nsd1* inactivation directly in fetal liver hematopoietic cells at different time points from E13.5 to E19.5 by comparing transcriptomes and *Nsd1* binding to chromatin and related histone marks (ChIP). Furthermore it would be of interest to investigate and compare the DNA methylome, chromatin marks and transcriptomes of patient and mouse-derived AEL cells as it was done for SOTOS patients.

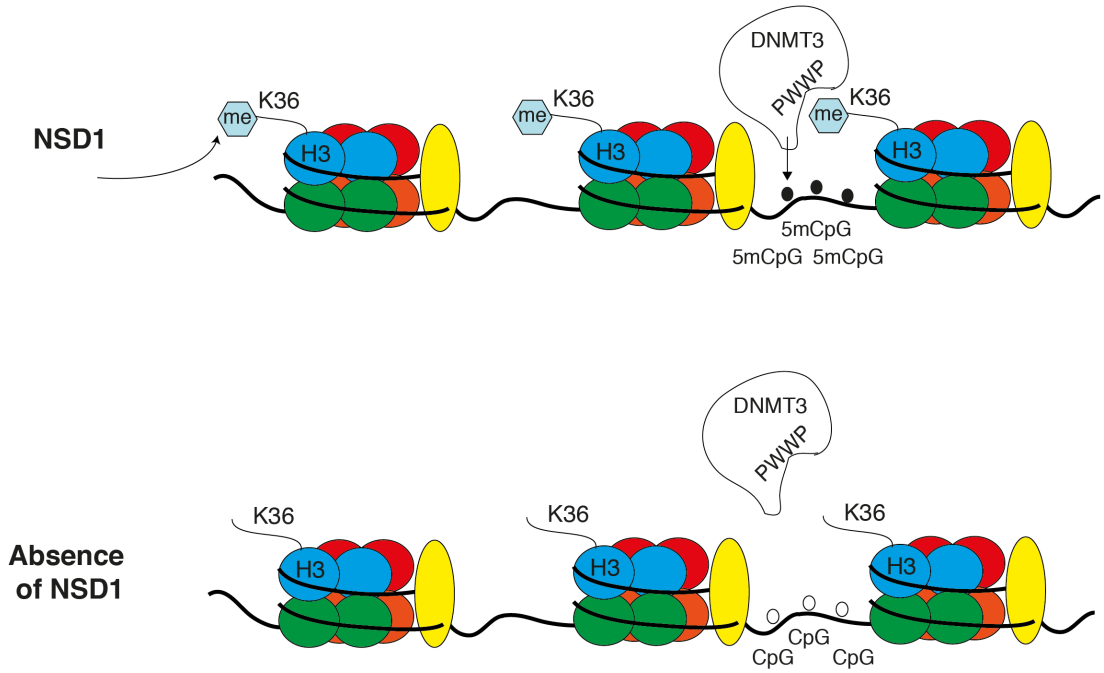


FIGURE 43. Hypothetical interplay between histone methylation and DNA methylation. Nsd1 methylates H3K36, a mark being recognized by the PWWP domain of the DNA methyltransferase 3. DNMT3 in turn methylates DNA often occuring at CpG islands (=5mCpG). In absence of NSD1, missing mono- and di- methylation of H3K36 might result in missing H3k36me3 mark. DNMT3 is not recruited to the locus and does not methylate DNA.

4.4.6 Translation into human AEL

4.4.6.1 Attempts to compare the Nsd1 KO model with human AEL

AEL is classified in humans as AML- M6 accounting for less than 5% of all acute leukemia cases. The disease was first described by Guglielmo in 1917 and often referred as “di Guglielmo disease” [278]. To date, the world health organization (WHO) classification recognizes three different forms of M6 based on the content of myelo- and proerythroblasts (**Table 11**) [279]. Myeloblasts are immature blood cells, bearing the potential to differentiate into granulocytic lineage whereas proerythroblasts are progenitor cells of the erythroid lineage.

TABLE 11. ACUTE ERYTHROLEUKEMIA SUBTYPE CLASSIFICATION

FAB M6	Acute erythroleukemia	> 50% of all nucleated cells
	Myeloblasts/NEC	Proerythroblasts/EC
MDS	< 30%	< 30%
M6A	> 30%	< 30%
M6B	< 30%	> 30%
M6C	> 30%	> 30%

AEL can occur after MDS or chronic myeloproliferations, as therapy-related neoplasia but also *de novo.* Despite this very heterogeneous ontogenicity the diseases the patients display common features such as anemia, thrombocytopenia, hepatosplenomegaly and most importantly erythroblast infiltrations [233,280,281]. In contrast to other AML subtypes, the underlying molecular cause of AEL is unknown and specific genetic aberrations phenocopying the disease have not been reported. Targeted sequencing studies revealed mutations in genes recurrently mutated in MDS or other AML subtypes, such as *p53* or *NPM1*. Interestingly, mutations giving proliferative advantages, e.g. *FLT3-ITD*, are rarely detected [233,280]. Very recently, two fusion genes have been reported to occur in childhood AEL, NFIA-CBFA2T3 t(1;16)(p31;q24) and ZMYND8-RELA t(11;20)(p11;q11) [256].

Important to note that in contrast to myelopoiesis, erythropoiesis significantly differs between mice and man [220,282]. Nevertheless, our experiments in human $CD34^+$ HSC clearly show that NSD1 is important for erythroid maturation in both species.

shRNA-mediated knockdown of NSD1 expression in $CD34^+$ cells derived from peripheral or cord blood, or whole cord blood cells altered colony formation in methylcellulose resulting in accumulation of proerythroblastic colonies upon replating (**Figure 40** & data not shown). So far, *NSD1* mutations in AEL have not been reported. Comparison of *NSD1* mRNA expression in 8 AEL patient samples revealed lower expression in some patients (**Figure 42**). This data might not be really representative, as information about the blast content of these samples was incomplete and ranged from 20 to 80%. Additionally, we sequenced the *NSD1* cDNA in 10 AEL samples and did not find mutations (data not shown). *NSD1* is located telomeric of the classical 5q- region being the most frequently found cytogenetic aberration MDS and AEL patients [283]. Hence, the F-36P cell line derived from an AEL patient developed from previous MDS bearing a 5q deletion expressed very low levels of *NSD1* (data not shown). Collectively, our data suggests that loss of NSD1 might activate key epigenetic pathways of malignant erythroid transformation in absence of genetic mutations. It is also likely that the promoter region of *NSD1* might be hypermethylated resulting in reduced expression in some AEL cases. Previous studies suggested that NSD1 might be target of aberrant hypermethylation in cancer leading to and reduced expression and increased clonogenic activity of neuroblastoma cells [179].

In order to compare the AEL phenotype of *Nsd1 null* mice with the human disease we sequenced mRNA of whole BM samples of five diseased *Vav1-iCre;Nsd1$^{fl/fl}$* and three control mice. Gene expression signatures displayed clustering according to their genotype but were rather heterogeneous between samples (data not shown). We also compared these profiles with signatures derived from AEL patients which we obtained from Catherine Carmichael (Melbourne, Australia, unpublished) and Stefan Bohlander (Auckland, New Zealand) [284]. However, we realized that the patient derived expression signatures were also very heterogeneous among each other, or did not contain appropriate controls, or they contained processed reads that could not be directly compared to our signatures. In order to overcome these limitations we started to collect samples worldwide (in collaboration with T. Mercher, Paris), to systematically determine the epigenomic landscape of a significant number of AEL patients. Taken together our data reveals a common underlying disease mechanism in mouse and human AEL that needs to be further investigated to elucidate the origin of very high GATA1 protein levels in these cells.

4.4.7 New hypotheses and experimental outlook

4.4.7.1 GATA1 levels as regulator of erythroid fate decision

Whether the increased protein levels of GATA1 in *Vav1-iCre;Nsd1*$^{fl/fl}$ mice are cause or consequence of accumulation of erythroid progenitor cells remains to be experimentally investigated. Interestingly, chromatin states seem to be set when cells undergo lineage decision for erythroid differentiation [285]. Histone modifications hereby serve as fine- tuning signals rather than an on/off switch at specific gene loci. In this context, binding of transcription factors such as GATA1 (or GATA1-containing complexes) may by altered due to differential accessibility provided by aberrant histone marks and therefore do not properly act as decision makers for differentiation [286]. Hoppe *et al*. recently suggested that "other" upstream players, activated by extracellular signals primarily decide lineage fate (**Figure 44**) [46]. We therefore plan to intercross the *Vav1-iCre;Nsd1*$^{fl/fl}$ mice with the *GATA1-mCherry* reporter mouse line to be able to directly track expression changes of the protein during fetal development and in different lineages.

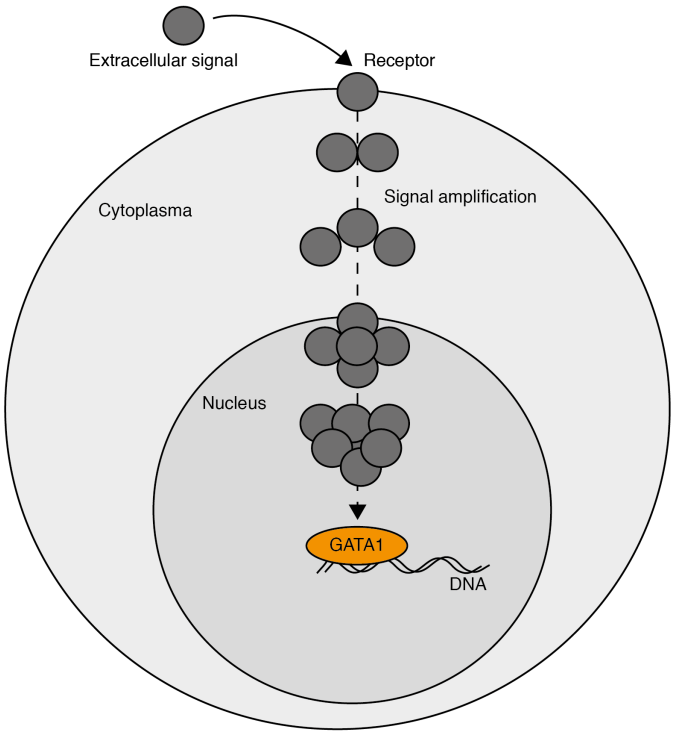


FIGURE 44. GATA1 as the downstream executer?

4.4.7.2 Indications for cell-non autonomous components of the AEL phenotype in *Vav1-iCre;Nsd1$^{fl/fl}$* mice?

In physiological situations, the spleen displays regions of white and red pulp, filled with lymphocytes and erythrocytes respectively. In case of "stress", erythropoiesis is activated in the spleen, a process called extra-medullary erythropoiesis [287]. It has been recognized that special macrophages primarily residing in the spleen form "islands" to support erythroblast maturation (**Figure 45**). These erythroid islands have been also found during fetal liver hematopoiesis [7]. Surprisingly, heavily symptomatic *Vav1-iCre;Nsd1$^{fl/fl}$* mice presented with a hypocellular BM, but their spleens were up to 16 times larger than found in littermate controls. Hereby the spleen morphology was completely disrupted by massive infiltration of CD71^{dim} erythroblasts. Although erythrocyte maturation was impaired as reflected by reduced red blood cell counts and hemoglobin values, the mice were still able to produce small and hypochromic erythrocytes (**Figure 13**).

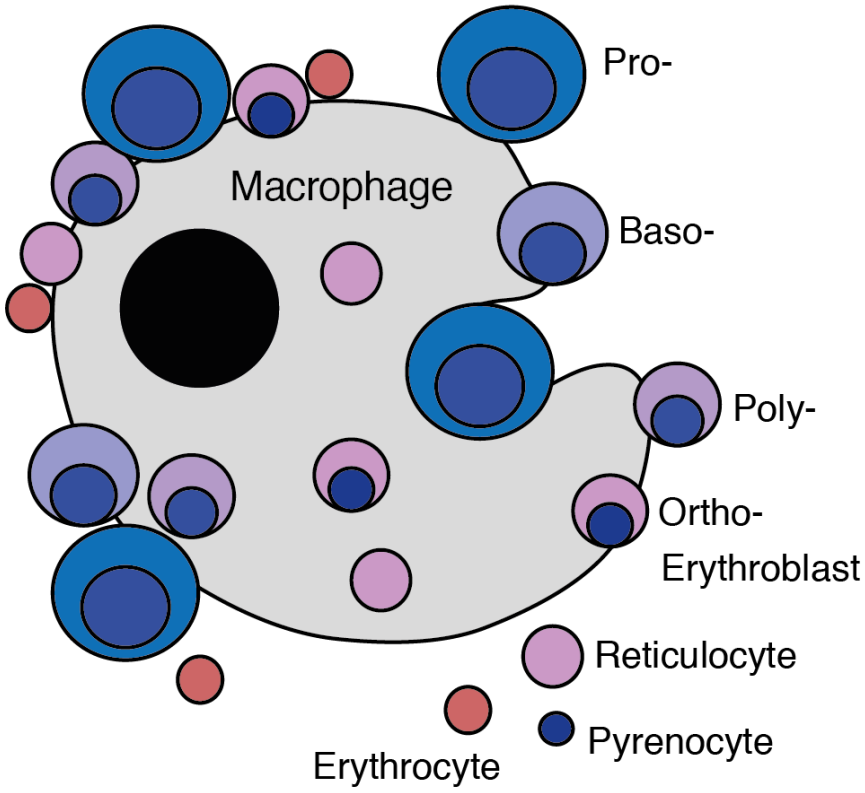


FIGURE 45. ERYTHROID ISLAND. Immature erythroblasts form an island with macrophages to enforce maturation into erythrocytes.

In addition to the blocked erythroblast differentiation, we observed an early reduction of HSC starting in fetal liver that resulted in loss of about 80% of HSC in the BM of young asymptomatic mice. Concomitantly lower numbers of GMP were present in the BM of these mice. At the same time, we found an unknown c-Kit^{+}/Sca-1^{+}/CD34^{-}/FcγRII/III^{+} expressing population suggesting a blockage at GMP/MEP bifurcation (**Figures 16, 23**). Moreover, transcription factors involved in lineage decision execution such as GATA1 and PU.1 were abnormally regulated (**Figure 28**).

Previous studies have suggested that interference with the interaction between macrophages and erythroblasts can affect erythroid differentiation. Mice deficient of retinoblastoma (Rb) tumor suppressor protein, develop anemia and die *in utero*. It has been suggested that both the intrinsic defects in erythroblasts as well as macrophage insufficiency may contribute to this phenotype [288]. In macrophages, Rb seems to counteract inhibition of PU.1 to ensure macrophage differentiation [67]. In erythroblasts, Rb binds to GATA1 directly and is responsible for erythroid maturation. Therefore a dual mechanism contributes to disease development [289]. It remains to be investigated whether and how macrophages may contribute to the phenotype we observed in *Vav1-iCre;Nsd1$^{fl/fl}$* mice. We plan to compare the macrophage content in BM and spleen, to isolate erythroid islands and to isolate island-forming macrophages and erythroblast to dissect cell (non)-autonomous contribution. A potential involvement of macrophages is of special interest in regard of potential novel therapies for AEL [290].

4.4.7.3 Mechanistic insights from the *Vav1-iCre;Nsd1$^{fl/fl}$* mouse model for future AEL-selective therapies?

Erythroblasts from *Vav1-iCre;Nsd1$^{fl/fl}$* mice expressed constitutively high GATA1 protein levels, and strikingly, overexpression of Gata1 induced terminal maturation of the cells. This picture resembles the Friend virus-induced AEL phenotypes including the MEL cell line derived from these mice. Importantly, knockdown of NSD1 in human CD34^{+} HSC also resulted in increased GATA1 protein levels. In addition, we found that all human AEL cell lines that we were able to investigate expressed higher GATA1 protein levels than other AML cell lines suggesting that aberrant GATA1 protein levels may play a critical role for development and maintenance of the disease (**Figure 46**). Current treatment of AEL is rarely curative and consists of DNA demethylating agents (such as 5-azacytitidine) with or without chemotherapy [291]. Interestingly, DNA demethylating agents were also proposed for the treatment of SOTOS syndrome since the degree of NSD1 associated silencing through DNA methylation was associated with poorer prognosis for patients [179].

Interestingly, a recent study reported aberrant CpG hypermethylation of the GATA1 gene promoter in BM samples from MDS patients. Whereas in control cells Gata1 was downregulated during erythroid differentiation correlating with gene expression on mRNA level, MDS samples expressed low GATA1 mRNA expression. The reason for the aberrant promoter methylation remained unknown [292].

These observation let me to explore the activity of hypomethylating agents such as 5- azacytidine on colony forming and replating potential as well as *in vitro* differentiation of *Vav1-iCre;Nsd1$^{fl/fl}$* erythroblasts. If cells would react to such a treatment, it would be interesting to dissect the specific DNA methylation changes that contribute to decreased leukemogenesis.

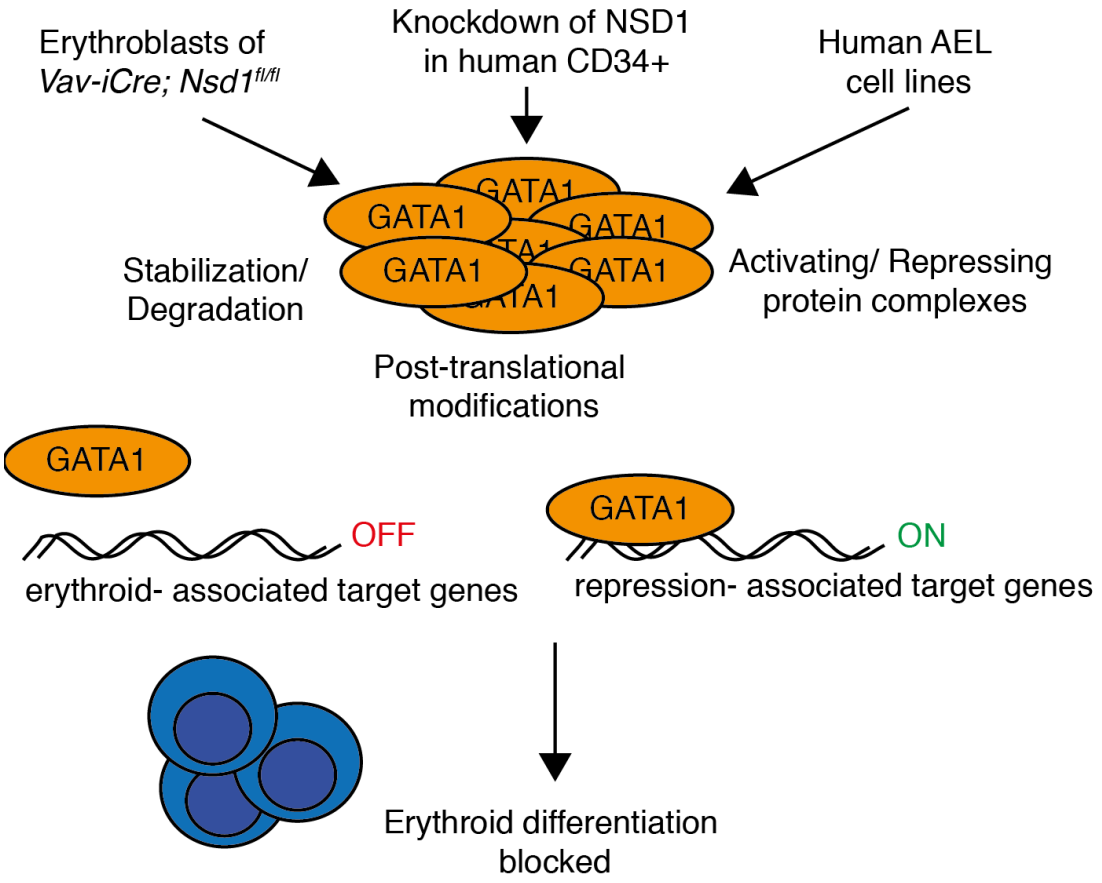


FIGURE 46. Constitutive high GATA1 protein levels: hallmark of nsd1-null erythroblasts but also human ael cell lines.

4.5 CHAPTER II. TARGETING EPIGENETIC REGULATORS OF LEUKEMIC CELLS BY SMALL MOLECULES

4.5.1 Research project and working hypothesis

Bromodomains (BRDs) are conserved protein interaction domains that recognize acetylated lysines, playing a pivotal role in chromatin remodeling and control of gene expression. Recently, our group and others have shown that bromodomains of the BET (bromo and extracellular domain) family have emerged as compelling targets for cancer therapy [293–296]. The development of selective and potent BET inhibitors and their significant activity in diverse tumor models has rapidly translated into clinical studies (e.g. NCT01713582, access through clinicaltrials.gov) and has motivated drug development efforts targeting non-BET BRDs.

We could report our work in 2015 in *Cancer Research* (see [297]) on a selective and highly potent chemical probe compound targeting the bromodomains of the histone acetyl transferases CBP/p300. Our collaborators from the SGC in Oxford have developed the acetyl-lysine mimetic oxazepine inhibitor I-CBP112 binding to CBP/p300 with nanomolar affinity and good selectivity. Exposure of human and mouse leukemic cell lines to I-CBP112 resulted in substantially impaired colony formation and induced cellular differentiation without significant cytotoxicity. The compound significantly reduced the leukemia-initiating potential of MLL-AF9 bearing AML cells in a dose-dependent manner *in vitro* and *in vivo*. Interestingly, I-CBP112 increased the cytotoxic activity of BET bromodomain inhibitor JQ1 as well as doxorubicin. Collectively, we could report the development and preclinical evaluation of a novel, potent inhibitor targeting CBP/p300 bromodomains that impairs aberrant self-renewal of leukemic cells.

4.5.2 Overall aim

The complex multidomain/subunit architecture of bromodomain protein complexes complicates predictions of consequences of their pharmacological targeting. To address this issue, our collaborator has developed a promiscuous bromodomain inhibitor (bromosporine) that broadly targets BRDs including BET with nanomolar affinity. Evaluation of bromosporine in BET- inhibitor sensitive and non-sensitive leukemic cell-lines revealed strong anti-proliferative activity in semi- solid medium in all of the tested lines. Moreover, sensitive leukemic cell lines such as MV4;11 displayed a distinct anti- proliferative response. Cells treated with BSP and JQ1 (BET inhibitor) arrested in S- cell cycle phase suggesting a BET- mediated effect. Comparison of the modulation of transcriptional profiles by bromosporine at short inhibitor exposure resulted in a BET inhibitor signature but no significant additional changes in transcription that could account for inhibition of other BRDs. Thus, non-selective targeting of BRDs identified BETs, but not other BRDs, as master regulators of a context dependent primary transcription response (see [298]).

Both projects regarding the targeting of epigenetic regulators of leukemic cells by small molecules have been published in peer- reviewed journals:

Picaud S*, Fedorov O*, Thanasopoulou A*, **Leonards K***, *et al.* (*equal contribution). Generation of a selective small molceule inhibitor of the CBP/p300 bromodomain for leukemia therapy. *Cancer Res.* 75, 5106-5120 (2015).

Picaud S, **Leonards K**, *et al.* Promiscuous targeting of bromodomains by Bromosporine identifies BET proteins as master regulators of primary transcription response in leukemia. *Sci. Adv.* **2,** e1600760 (2016).

5 REFERENCES

1. Dzierzak, E. & Speck, N. A. Of lineage and legacy : the development of mammalian hematopoietic stem cells. *Nat. Immunol.* **9,** 129–136 (2008).
2. Ferkowicz, M. J. & Yoder, M. C. Blood island formation : longstanding observations and modern interpretations. *Exp. Hematol.* **33,** 1041–1047 (2005).
3. Tober, J. *et al.* The megakaryocyte lineage originates from hemangioblast precursors and is an integral component both of primitive and of definitive hematopoiesis. *Blood* **109,** 1433–1441 (2007).
4. Potts, K. S. *et al.* A lineage of diploid platelet-forming cells precedes polyploid megakaryocyte formation in the mouse embryo. *Blood* **124,** 2725–2729 (2014).
5. Dzierzak, E. & Philipsen, S. Erythropoiesis : Development and Differentiation. *Cold Spring Harb. Perspect. Med.* **3,** 1–16 (2013).
6. Palis, J., Robertson, S., Kennedy, M., Wall, C. & Keller, G. Development of erythroid and myeloid progenitors in the yolk sac and embryo proper of the mouse. *Development* **126,** 5073–5084 (1999).
7. Chasis, J. A. & Mohandas, N. Erythroblastic islands: niches for erythropoiesis. *Blood* **112,** 470–8 (2008).
8. Nandakumar, S. K., Ulirsch, J. C. & Sankaran, V. G. Advances in understanding erythropoiesis: evolving perspectives. *Br. J. Haematol.* **173,** 206–218 (2016).
9. Orkin, S. H. & Zon, L. I. Hematopoiesis: An Evolving Paradigm for Stem Cell Biology. *Cell* **132,** 631–644 (2008).
10. Rossi, L. *et al.* Less is more: unveiling the functional core of hematopoietic stem cells through knockout mice. *Cell Stem Cell* **11,** 302–17 (2012).
11. Hattangadi, S. M., Wong, P., Zhang, L., Flygare, J. & Lodish, H. F. From stem cell to red cell : regulation of erythropoiesis at multiple levels by multiple proteins , RNAs , and chromatin modifications. *Blood* **118,** 6258–6269 (2011).
12. Wang, L. D. & Wagers, A. J. Dynamic niches in the origination and differentiation of haematopoietic stem cells. *Nat. Rev. Mol. Cell Biol.* **12,** 643–55 (2011).
13. Mendelson, A. & Frenette, P. S. Hematopoietic stem cell niche maintenance

during homeostasis and regeneration. *Nat. Med.* **20,** 833–846 (2014).

14. Baker, S. J., Rane, S. G. & Reddy, E. P. Hematopoietic cytokine receptor signaling. *Oncogene* **26,** 6724–37 (2007).
15. Hardy, R. R. & Hayakawa, K. B Cell Development Pathways. *Annu. Rev. Immunol.* **19,** 595–621 (2001).
16. Koulnis, M. *et al.* Identification and analysis of mouse erythroid progenitors using the CD71/TER119 flow-cytometric assay. *J. Vis. Exp.* 1–7 (2011). doi:10.3791/2809
17. Oguro, H., Ding, L. & Morrison, S. J. SLAM family markers resolve functionally distinct subpopulations of hematopoietic stem cells and multipotent progenitors. *Cell Stem Cell* **13,** 102–16 (2013).
18. Pronk, C. J. H. *et al.* Elucidation of the phenotypic, functional, and molecular topography of a myeloerythroid progenitor cell hierarchy. *Cell Stem Cell* **1,** 428–42 (2007).
19. Akashi, K., Traver, D., Miyamoto, T. & Weissman, I. L. A clonogenic common myeloid progenitor that gives rise to all myeloid lineages. *Nature* **404,** 193–7 (2000).
20. Yamamoto, R. *et al.* Clonal analysis unveils self-renewing lineage-restricted progenitors generated directly from hematopoietic stem cells. *Cell* **154,** 1112–26 (2013).
21. Dykstra, B., Olthof, S., Schreuder, J., Ritsema, M. & de Haan, G. Clonal analysis reveals multiple functional defects of aged murine hematopoietic stem cells. *J Exp Med* **208,** 2691–2703 (2011).
22. Sanjuan-Pla, A. *et al.* Platelet-biased stem cells reside at the apex of the haematopoietic stem-cell hierarchy. *Nature* **502,** 232–6 (2013).
23. Woolthuis, C. M. & Park, C. Y. Hematopoietic stem/progenitor cell commitment to the megakaryocyte lineage. *Blood* **127,** 1242–1248 (2016).
24. Müller-Sieburg, C. E., Cho, R. H., Thoman, M., Adkins, B. & Sieburg, H. B. Deterministic regulation of hematopoietic stem cell self-renewal and differentiation. *Blood* **100,** 1302–1309 (2002).
25. Arinobu, Y. *et al.* Reciprocal activation of GATA-1 and PU.1 marks initial specification of hematopoietic stem cells into myeloerythroid and myelolymphoid lineages. *Cell Stem Cell* **1,** 416–427 (2007).
26. Iwasaki, H. *et al.* Distinctive and indispensable roles of PU . 1 in maintenance of hematopoietic stem cells and their differentiation. *Blood* **106,** 1590–1601

(2005).

27. Graf, T. & Enver, T. Forcing cells to change lineages. *Nature* **462,** 587–594 (2009).
28. Krause, D. Regulation of hematopoietic stem cell fate. *Oncogene* **21,** 3262–3269 (2002).
29. Sainsbury, S., Bernecky, C. & Cramer, P. Structural basis of transcription initiation by RNA polymerase II. *Nat. Rev. Mol. Cell Biol.* **16,** 129–143 (2015).
30. Thomas, M. C. & Chiang, C.-M. The general transcription machinery and general cofactors. *Crit. Rev. Biochem. Mol. Biol.* **41,** 105–178 (2006).
31. Doré, L. C. & Crispino, J. D. Transcription factor networks in erythroid cell and megakaryocyte development. *Blood* **118,** 231–239 (2011).
32. Ferreira, R., Ohneda, K., Yamamoto, M. & Philipsen, S. GATA1 function, a paradigm for transcription factors in hematopoiesis. *Mol. Cell. Biol.* **25,** 1215–1227 (2005).
33. Cantor, A. B. & Orkin, S. H. Transcriptional regulation of erythropoiesis : an affair involving multiple partners. *Oncogene* **21,** 3368–3376 (2002).
34. Malinge, S. *et al.* Ikaros inhibits megakaryopoiesis through functional interaction with GATA-1 and NOTCH signaling. *Blood* **121,** 2440–51 (2013).
35. Pimkin, M. *et al.* Divergent functions of hematopoietic transcription factors in lineage priming and differentiation during erythro-megakaryopoiesis. *Genome Res.* **24,** 1932–1944 (2014).
36. Rotem, A. *et al.* Single-cell ChIP-seq reveals cell subpopulations defined by chromatin state. *Nat Biotechnol.* **33,** 1165–1172 (2015).
37. Lara-Astiaso, D. *et al.* Chromatin state dynamics during blood formation. *Science (80-.).* **345,** 943–9 (2014).
38. Zon, L. I. *et al.* The major human erythroid DNA-binding protein (GF-1): primary sequence and localization of the gene to the X chromosome. *Proc. Natl. Acad. Sci. USA* **87,** 668–672 (1990).
39. Fujiwara, Y., Browne, C. P., Cunniff, K., Goff, S. C. & Orkin, S. H. Arrested development of embryonic red cell precursors in mouse embryos lacking transcription factor GATA-1. *PNAS* **93,** 12355–12358 (1996).
40. Leonard, M., Brice, M., Engel, J. D. & Papayannopoulou, T. Dynamics of GATA transcription factor expression during erythroid differentiation. *Blood* **82,** 1071–9 (1993).
41. Martin, D. I., Zon, L. I., Mutter, G. & Orkin, S. H. Expression of an erythroid

transcription factor in megakaryocytic and mast cell lineages. *Nature* **344,** 444–447 (1990).

42. Romeo, P.-H. *et al.* Megakaryocytic and eryth- rocytic lineages share specific transcription factors. *Nature* **344,** 447–449 (1990).
43. Zon, B. L. I. *et al.* Expression of mRNA for the GATA-Binding Proteins in Human Eosinophils and Basophils: Potential Role in Gene Transcription. *Blood* **81,** 3234–3241 (1993).
44. Yomogida, K. *et al.* Developmental stage- and spermatogenic cycle-specific expression of transcription factor GATA-1 in mouse Sertoli cells. *Development* **120,** 1759–1766 (1994).
45. Imagawa, S., Yamamoto, M. & Miura, Y. Negative regulation of the erythropoietin gene expression by the GATA transcription factors. *Blood* **89,** 1430–9 (1997).
46. Hoppe, P. S. *et al.* Early myeloid lineage choice is not initiated by random PU.1 to GATA1 protein ratios. *Nature* **535,** 299–302 (2016).
47. Ito, E. *et al.* Erythroid transcription factor GATA-1 is abundantly transcribed in mouse testis. *Nature* **362,** 466–8. (1993).
48. Nishimura, S. *et al.* A GATA box in the GATA-1 gene hematopoietic enhancer is a critical element in the network of GATA factors and sites that regulate this gene. *Mol. Cell. Biol.* **20,** 713–23 (2000).
49. Martin, D. I. K. & Orkin, S. H. Transcriptional activation and DNA binding by the erythroid factor GF-1 / NF-El / Eryf 1. *Genes Dev.* **4,** 1886–1898 (1990).
50. Yang, H. Y. & Evans, T. Distinct roles for the two cGATA-1 finger domains. *Mol. Cell. Biol.* **12,** 4562–70 (1992).
51. Ko, L. J. & Engel, J. D. DNA-binding specificities of the GATA transcription factor family. *Mol. Cell. Biol.* **13,** 4011–4022 (1993).
52. Whyatt, D. J., deBoer, E. & Grosveld, F. The two zinc finger-like domains of GATA-1 have different DNA binding specificities. *Embo J* **12,** 4993–5005 (1993).
53. Yamamoto, M. *et al.* Activity and tissue-specific expression of the transcription factor NF-E1 multigene family. *Genes Dev* **4,** 1650–1662 (1990).
54. Mackay, J. P. *et al.* Involvement of the N-finger in the self-association of GATA-1. *J. Biol. Chem.* **273,** 30560–7 (1998).
55. Wall, L., DeBoer, E. & Grosveld, F. The human b-globin gene 3 ' enhancer contains multiple binding sites for an erythroid-specific protein. *Genes Dev.* **2,**

1089–1100 (1988).

56. Elefanty, A. G., Antoniou, M., Custodio, N., Carmo-fonseca, M. & Grosveld, F. G. GATA transcription factors associate with a novel class of nuclear bodies in erythroblasts and megakaryocytes. *EMBO J.* **15,** 319–333 (1996).
57. Briegel, K. *et al.* Regulation and function of transcription factor GATA-1 during red blood cell differentiation. *Development* **122,** 3839–3850 (1996).
58. Hung, H., Lau, J., Kim, A. Y. & Weiss, M. J. CREB-Binding Protein Acetylates Hematopoietic Transcription Factor GATA-1 at Functionally Important Sites. *Mol Cell Biol* **19,** 3496–3505 (1999).
59. Hernandez-Hernandez, A. *et al.* Acetylation and MAPK phosphorylation cooperate to regulate the degradation of active GATA-1. *EMBO J.* **25,** 3264–74 (2006).
60. Blobel, G. a, Nakajima, T., Eckner, R., Montminy, M. & Orkin, S. H. CREB-binding protein cooperates with transcription factor GATA-1 and is required for erythroid differentiation. *Proc. Natl. Acad. Sci. U. S. A.* **95,** 2061–2066 (1998).
61. Collavin, L. *et al.* Modification of the erythroid transcription factor GATA-1 by SUMO-1. *Proc. Natl. Acad. Sci. U. S. A.* **101,** 8870–5 (2004).
62. Crossley, M. & Orkin, S. H. Phosphorylation of the erythroid transcription factor GATA-1. *J. Biol. Chem.* **269,** 16589–16596 (1994).
63. Rodriguez, P. *et al.* GATA-1 forms distinct activating and repressive complexes in erythroid cells. *EMBO J.* **24,** 2354–2366 (2005).
64. Shimizu, R. *et al.* GATA-1 self-association controls erythroid development in vivo. *J. Biol. Chem.* **282,** 15862–15871 (2007).
65. Hamlett, I. *et al.* Characterization of megakaryocyte GATA1-interacting proteins: The corepressor ETO2 and GATA1 interact to regulate terminal megakaryocyte maturation. *Blood* **112,** 2738–2749 (2008).
66. Bresnick, E. H., Katsumura, K. R., Lee, H. Y., Johnson, K. D. & Perkins, A. S. Master regulatory GATA transcription factors: Mechanistic principles and emerging links to hematologic malignancies. *Nucleic Acids Res.* **40,** 5819–5831 (2012).
67. Rekhtman, N., Radparvar, F., Evans, T. & Skoultchi, A. I. Direct interaction of hematopoietic transcription factors PU.1 and GATA-1: functional antagonism in erythroid cells. *Genes Dev.* **13,** 1398–1411 (1999).
68. Zhang, P. *et al.* Negative cross-talk between hematopoietic regulators: GATA proteins repress PU.1. *Proc. Natl. Acad. Sci. U. S. A.* **96,** 8705–10 (1999).

69. Nerlov, C., Querfurth, E., Kulessa, H. & Graf, T. GATA-1 interacts with the myeloid PU. 1 transcription factor and represses PU. 1-dependent transcription. *Blood* **95,** 2543–2551 (2000).

70. Shimizu, R., Engel, J. D. & Yamamoto, M. GATA1-related leukaemias. *Nat. Rev. Cancer* **8,** 279–287 (2008).

71. Pevny, L. *et al.* Erythroid differentiation in chimaeric mice blocked by a targeted mutation in the gene for transcription factor GATA-1. *Nature* **349,** 257–260 (1991).

72. Weiss, M. J., Keller, G. & Orkin, S. H. Novel insights into erythroid development revealed through in vitro differentiation of GATA-1- embryonic stem cells. *Genes Dev.* **8,** 1184–1197 (1994).

73. Shimizu, R. *et al.* Leukemogenesis Caused by Incapacitated GATA-1 Function Leukemogenesis Caused by Incapacitated GATA-1 Function. *Mol Cell Biol* **24,** 10814–10825 (2004).

74. Takahashi, S. *et al.* Arrest in primitive erythroid cell development caused by promoter- specific disruption of the GATA-1 gene. *J. Biol. Chem.* **272,** 12611–12615 (1997).

75. McDevitt, M. a, Fujiwara, Y., Shivdasani, R. a & Orkin, S. H. An upstream, DNase I hypersensitive region of the hematopoietic-expressed transcription factor GATA-1 gene confers developmental specificity in transgenic mice. *Proc. Natl. Acad. Sci. U. S. A.* **94,** 7976–81 (1997).

76. Vyas, P. & Crispino, J. D. Molecular insights into Down syndrome-associated leukemia. *Curr. Opin. Pediatr.* **19,** 9–14 (2007).

77. Vyas, P. & Roberts, I. Down myeloid disorders: A paradigm for childhood preleukaemia and leukaemia and insights into normal megakaryopoiesis. *Early Hum. Dev.* **82,** 767–773 (2006).

78. Shimada, A. *et al.* Fetal origin of the GATA1 mutation in identical twins with transient myeloproliferative disorder and acute megakaryoblastic leukemia accompanying Down syndrome. *Blood* **103,** 366 (2004).

79. Hitzler, J. K., Cheung, J., Li, Y., Scherer, S. W. & Zipursky, A. GATA1 mutations in transient leukemia and acute megakaryoblastic leukemia of Down syndrome. *Blood* **101,** 4301–4304 (2003).

80. Hitzler, J. K. & Zipursky, A. Origins of leukaemia in children with Down syndrome. *Nat. Rev. Cancer* **5,** 11–20 (2005).

81. Ahmed, M. *et al.* Natural history of GATA1 mutations in Down syndrome.

Blood **103,** 2480–2489 (2004).

82. Back, J. *et al.* PU . 1 determines the self-renewal capacity of erythroid progenitor cells Plenary paper PU . 1 determines the self-renewal capacity of erythroid progenitor cells. *Blood* **103,** 3615–3623 (2004).
83. Scott, E. W., Simon, M. C., Anastasi, J. & Singh, H. Requirement of transcription factor PU.1 in the development of multiple hematopietic lineages. *Science (80-.).* **265,** 1573–1577 (1994).
84. McKercher, S. R. *et al.* Targeted disruption of the PU.1 gene results in multiple hematopoietic abnormalities. *EMBO J.* **15,** 5647–5658 (1996).
85. Spain, L., Guerriero, A., Kunjibettu, S. & Scott, E. T cell development in PU.1-deficient mice. *J. Immunol.* **163,** 2681–2687 (1999).
86. Colucci, F. *et al.* Differential requirement for the transcription factor PU.1 in the generation of natural killer cells versus B and T cells. *Blood* **97,** 2625–2632 (2001).
87. Friend, C. Cell free transmission in adult Swiss mice of a disease having the character of leukemia. *J Exp Med* **105,** 307–318 (1957).
88. Moreau- Gachelin, F. *et al.* Spi-1 / PU . 1 Transgenic Mice Develop Multistep Erythroleukemias. *Mol Cell Biol* **16,** 2453–2463 (1996).
89. Zhang, P. *et al.* PU.1 inhibits GATA-1 function and erythroid differentiation by blocking GATA-1 DNA binding. *Blood* **96,** 2641–2649 (2000).
90. Pop, R. *et al.* A Key Commitment Step in Erythropoiesis Is Synchronized with the Cell Cycle Clock through Mutual Inhibition between PU . 1 and S-Phase Progression. *Plos Biol.* **8,** e1000484 (2010).
91. Rosenbauer, F. *et al.* Acute myeloid leukemia induced by graded reduction of a lineage-specific transcription factor , PU.1. *Nat. Genet.* **36,** 624–630 (2004).
92. Mueller, B. U. *et al.* Heterozygous PU . 1 mutations are associated with acute myeloid leukemia. *Blood* **100,** 998–1008 (2002).
93. Nutt, S. L., Metcalf, D., Amico, A. D. & Polli, M. Dynamic regulation of PU . 1 expression in multipotent hematopoietic progenitors. *J Exp Med* **201,** 221–231 (2005).
94. Ling, K.-W. *et al.* GATA-2 plays two functionally distinct roles during the ontogeny of hematopoietic stem cells. *J. Exp. Med.* **200,** 871–882 (2004).
95. Tsai, F. Y. & Orkin, S. H. Transcription factor GATA-2 is required for proliferation/survival of early hematopoietic cells and mast cell formation, but not for erythroid and myeloid terminal differentiation. *Blood* **89,** 3636–43

(1997).

96. Kaneko, H., Shimizu, R. & Yamamoto, M. GATA factor switching during erythroid differentiation. *Curr. Opin. Hematol.* **17,** 163–168 (2010).
97. Theis, F. *et al.* Clinical impact of GATA2 mutations in acute myeloid leukemia patients harboring CEBPA mutations: a study of the AML study group. *Leukemia* 1–3 (2016). doi:10.1038/leu.2016.185
98. Zhang, S.-J. *et al.* Gain-of-function mutation of GATA-2 in acute myeloid transformation of chronic myeloid leukemia. *Proc. Natl. Acad. Sci. U. S. A.* **105,** 2076–81 (2008).
99. Zhang, S. J., Shi, J. Y. & Li, J. Y. GATA-2 L359 V mutation is exclusively associated with CML progression but not other hematological malignancies and GATA-2 P250A is a novel single nucleotide polymorphism. *Leuk. Res.* **33,** 1141–1143 (2009).
100. Tenen, D. G., Hromas, R., Licht, J. D. & Zhang, D. E. Transcription factors, normal myeloid development, and leukemia. *Blood* **90,** 489–519 (1997).
101. Rossetto, D., Avvakumov, N. & Cote, J. Histone phosphorylation: A chromatin modification involved in diverse nuclear events. *Epigenetics* **7,** 1098–1108 (2012).
102. Zhou, V. W., Goren, A. & Bernstein, B. E. Charting histone modifications and the functional organization of mammalian genomes. *Nat. Rev. Genet.* **12,** 7–18 (2011).
103. Ugarte, F. *et al.* Progressive chromatin condensation and H3K9 methylation regulate the differentiation of embryonic and hematopoietic stem cells. *Stem Cell Reports* **5,** 728–740 (2015).
104. Efroni, S. *et al.* Global Transcription in Pluripotent Embryonic Stem Cells. *Cell Stem Cell* **2,** 437–447 (2008).
105. Butler, J. S. & Dent, S. Y. R. The role of chromatin modifiers in normal and malignant hematopoiesis. *Blood* **121,** 3076–3084 (2013).
106. Falkenberg, K. J. & Johnstone, R. W. Histone deacetylases and their inhibitors in cancer, neurological diseases and immune disorders. *Nat. Rev. Drug Discov.* **13,** 673–91 (2014).
107. Rice, K. L., Hormaeche, I. & Licht, J. D. Epigenetic regulation of normal and malignant hematopoiesis. *Oncogene* **26,** 6697–6714 (2007).
108. Dekker, F. J. & Haisma, H. J. Histone acetyl transferases as emerging drug targets. *Drug Discov. Today* **14,** 942–948 (2009).

109. Chen, J. & Li, Q. Life and death of transcriptional co-activator p300. *Epigenetics* **6,** 957–961 (2011).
110. Delvecchio, M., Gaucher, J., Aguilar-Gurrieri, C., Ortega, E. & Panne, D. Structure of the p300 catalytic core and implications for chromatin targeting and HAT regulation. *Nat. Struct. Mol. Biol.* **20,** 1040–6 (2013).
111. Filippakopoulos, P. & Knapp, S. Targeting bromodomains: epigenetic readers of lysine acetylation. *Nat. Rev. Drug Discov.* **13,** 337–56 (2014).
112. Taverna, S. D., Li, H., Ruthenburg, A. J., Allis, C. D. & Patel, D. J. How chromatin-binding modules interpret histone modifications: lessons from professional pocket pickers. *Nat. Struct. Mol. Biol.* **14,** 1025–1040 (2007).
113. Kung, A. L. *et al.* Gene dose-dependent control of hematopoiesis and hematologic tumor suppression by CBP. *Genes Dev.* **14,** 272–277 (2000).
114. Yao, T. P. *et al.* Gene dosage-dependent embryonic development and proliferation defects in mice lacking the transcriptional integrator p300. *Cell* **93,** 361–372 (1998).
115. Chan, W.-I. *et al.* The Transcriptional Coactivator Cbp Regulates Self-Renewal and Differentiation in Adult Hematopoietic Stem Cells. *Mol. Cell. Biol.* **31,** 5046–5060 (2011).
116. Gopalakrishna Iyer, N., Özdag, H. & Caldas, C. p300/CBP and cancer. *Oncogene* **23,** 4225–4231 (2004).
117. Katsumoto, T., Yoshida, N. & Kitabayashi, I. Roles of the histone acetyltransferase monocytic leukemia zinc finger protein in normal and malignant hematopoiesis. *Cancer Sci.* **99,** 1523–1527 (2008).
118. Zhang, Y. *et al.* Characterization of genomic breakpoints in MLL and CBP in leukemia patients with t(11;16). *Genes Chromosom. Cancer* **41,** 257–265 (2004).
119. Rowley, J. D. *et al.* All patients with the T(11;16)(q23;p13.3) that involves MLL and CBP have treatment-related hematologic disorders. *Blood* **90,** 535–541 (1997).
120. Kasper, L. H. *et al.* CREB binding protein interacts with nucleoporin-specific FG repeats that activate transcription and mediate NUP98-HOXA9 oncogenicity. *Mol. Cell. Biol.* **19,** 764–76 (1999).
121. Zhao, Y. *et al.* Acetylation of p53 at Lysine 373 / 382 by the Histone Deacetylase Inhibitor Depsipeptide Induces Expression of p21 Waf1 / Cip1. *Mol. Cell. Biol.* **26,** 2782–2790 (2006).

122. Wang, L. *et al.* The Leukemogenicity of AML1-ETO Is Dependent on Site-Specific Lysine Acetylation. *Science (80-.).* **333,** 765–769 (2011).
123. Peters, A. & Schwaller, J. *Epigenetic Mechanisms in Acute Myeloid Leukemia from Epigenetics and Disease*. (2011).
124. Schuettengruber, B., Martinez, A.-M., Iovino, N. & Cavalli, G. Trithorax group proteins: switching genes on and keeping them active. *Nat. Rev. Mol. Cell Biol.* **12,** 799–814 (2011).
125. Milne, T. a, Martin, M. E., Brock, H. W., Slany, R. K. & Hess, J. L. Leukemogenic MLL fusion proteins bind across a broad region of the Hox a9 locus, promoting transcription and multiple histone modifications. *Cancer Res.* **65,** 11367–74 (2005).
126. Slany, R. K. The molecular biology of mixed lineage leukemia. *Haematologica* **94,** 984–993 (2009).
127. Dou, Y. *et al.* Regulation of MLL1 H3K4 methyltransferase activity by its core components. *Nat. Struct. Mol. Biol.* **13,** 713–719 (2006).
128. Milne, T. A. *et al.* MLL Targets SET Domain Methyltransferase Activity to Hox Gene Promoters. *Mol. Cell* **10,** 1107–1117 (2002).
129. Hublitz, P., Albert, M. & Peters, A. H. F. M. Mechanisms of transcriptional repression by histone lysine methylation. *Int. J. Dev. Biol.* **53,** 335–354 (2009).
130. Yu, B. D., Hess, J. L., Horning, S. E., Brown, G. a & Korsmeyer, S. J. Altered Hox expression and segmental identity in Mll-mutant mice. *Nature* **378,** 505–508 (1995).
131. Yokoyama, A. *et al.* The menin tumor suppressor protein is an essential oncogenic cofactor for MLL-associated leukemogenesis. *Cell* **123,** 207–218 (2005).
132. Jin, S. *et al.* c-Myb binds MLL through menin in human leukemia cells and is an important driver of MLL-associated leukemogenesis. *J. Clin. Invest.* **120,** 593–606 (2010).
133. Yokoyama, A. & Cleary, M. L. Menin Critically Links MLL Proteins with LEDGF on Cancer-Associated Target Genes. *Cancer Cell* **14,** 36–46 (2008).
134. Méreau, H. *et al.* Impairing MLL-fusion gene-mediated transformation by dissecting critical interactions with the lens epithelium-derived growth factor (LEDGF/p75). *Leukemia* **27,** 1245–53 (2013).
135. Greer, E. L. & Shi, Y. Histone methylation: a dynamic mark in health, disease and inheritance. *Nat. Rev. Genet.* **13,** 343–57 (2012).

136. Wu, H. & Zhang, Y. Reversing DNA methylation: Mechanisms, genomics, and biological functions. *Cell* **156,** 45–68 (2014).
137. Smith, Z. D. & Meissner, A. DNA methylation: roles in mammalian development. *Nat. Rev. Genet.* **14,** 204–20 (2013).
138. Okano, M., Bell, D. W., Haber, D. A. & Li, E. DNA methyltransferases Dnmt3a and Dnmt3b are essential for de novo methylation and mammalian development. *Cell* **99,** 247–257 (1999).
139. Yang, L., Rau, R. & Goodell, M. a. DNMT3A in haematological malignancies. *Nat. Rev. Cancer* **15,** 152–165 (2015).
140. Challen, G. a *et al.* Dnmt3a is essential for hematopoietic stem cell differentiation. *Nat. Genet.* **44,** 23–31 (2011).
141. Mayle, A. *et al.* Dnmt3a loss predisposes murine hematopoietic stem cells to malignant transformation. *Blood* **125,** 629–38 (2015).
142. Ley, T. et al. Genomic and epigenomic landscapes of adult de novo acute myeloid leukemia. *N. Engl. J. Med.* **368,** 2059–74 (2013).
143. Lowenberg, B., Downing, J. R. & Burnett, A. Acute Myeloid Leukemia. *N Engl J Med* **341,** 1051–1062 (1999).
144. Shih, A. H., Abdel-Wahab, O., Patel, J. P. & Levine, R. L. The role of mutations in epigenetic regulators in myeloid malignancies. *Immunol. Rev.* **263,** 22–35 (2012).
145. Thol, F. *et al.* Incidence and prognostic influence of DNMT3A mutations in acute myeloid leukemia. *J. Clin. Oncol.* **29,** 2889–2896 (2011).
146. Yan, X.-J. *et al.* Exome sequencing identifies somatic mutations of DNA methyltransferase gene DNMT3A in acute monocytic leukemia. *Nat Genet* **43,** 309–315 (2011).
147. Kim, S. J. *et al.* A DNMT3A mutation common in AML exhibits dominant-negative effects in murine ES cells. *Blood* **122,** 4086–9 (2013).
148. Walter, M. J. *et al.* Recurrent DNMT3A mutations in patients with myelodysplastic syndromes. *Leukemia* **25,** 1153–1158 (2011).
149. Genovese, G. *et al.* Clonal hematopoiesis and blood-cancer risk inferred from blood DNA sequence. *N. Engl. J. Med.* **371,** 2477–87 (2014).
150. Brown, S. J., Stoilov, P. & Xing, Y. Chromatin and epigenetic regulation of pre-mrna processing. *Hum. Mol. Genet.* **21,** 90–96 (2012).
151. Dhayalan, A. *et al.* The Dnmt3a PWWP domain reads histone 3 lysine 36 trimethylation and guides DNA methylation. *J. Biol. Chem.* **285,** 26114–26120

(2010).

152. Okada, Y. *et al.* hDOT1L links histone methylation to leukemogenesis. *Cell* **121,** 167–178 (2005).
153. Rose, N. R. & Klose, R. J. Understanding the relationship between DNA methylation and histone lysine methylation. *Biochim. Biophys. Acta - Gene Regul. Mech.* **1839,** 1362–1372 (2014).
154. Cedar, H. & Bergman, Y. Linking DNA methylation and histone modification: patterns and paradigms. *Nat. Rev. Genet.* **10,** 295–304 (2009).
155. Tan, Y.-T. *et al.* Deregulation of HOX genes by DNMT3A and MLL mutations converges on BMI1. *Leukemia* **30,** 1609–1612 (2016).
156. Papaemmanuil, E. *et al.* Genomic classification and prognosis in acute myeloid leukemia. *N Engl J Med* **374,** 2209–2221 (2016).
157. Alexandrov, L. B. *et al.* Signatures of mutational processes in human cancer. *Nature* **500,** 415–21 (2013).
158. Grove, C. S. & Vassiliou, G. S. Acute myeloid leukaemia: a paradigm for the clonal evolution of cancer? *Dis. Model. Mech.* **7,** 941–951 (2014).
159. Rosenbauer, F., Koschmieder, S., Steidl, U., Tenen, D. G. & Dc, W. Effect of transcription-factor concentrations on leukemic stem cells. *Blood* **106,** 1519–1524 (2005).
160. Schuback, H. L., Arceci, R. J. & Meshinchi, S. Somatic characterization of pediatric acute myeloid leukemia using next-generation sequencing. *Semin. Hematol.* **50,** 325–332 (2013).
161. Balgobind, B. V *et al.* Integrative analysis of type-I and type-II aberrations underscores the genetic heterogeneity of pediatric acute myeloid leukemia. *Haematologica* **96,** 1478–87 (2011).
162. Jaju, R. J. A novel gene, NSD1, is fused to NUP98 in the t(5;11)(q35;p15.5) in de novo childhood acute myeloid leukemia. *Blood* **98,** 1264–1267 (2001).
163. Hollink, I. H. I. M. *et al.* NUP98/NSD1 characterizes a novel poor prognostic group in acute myeloid leukemia with a distinct HOX gene expression pattern. *Blood* **118,** 3645–56 (2011).
164. Fasan, A. *et al.* A rare but specific subset of adult AML patients can be defined by the cytogenetically cryptic NUP98 – NSD1 fusion gene. *Leukemia* **27,** 233–258 (2013).
165. Thol, F. *et al.* Analysis of NUP98/NSD1 translocations in adult AML and MDS patients. *Leukemia* **27,** 750–4 (2013).

166. Franks, T. M. & Hetzer, M. W. The role of Nup98 in transcription regulation in healthy and diseased cells. *Trends Cell Biol.* **23,** 112–117 (2013).
167. Schmidt, H. B. roder & Görlich, D. Nup98 FG domains from diverse species spontaneously phase-separate into particles with nuclear pore-like permselectivity. *Elife* **4,** 1–30 (2015).
168. Jankovic, D. *et al.* Leukemogenic mechanisms and targets of a NUP98/HHEX fusion in acute myeloid leukemia. *Blood* **111,** 5672–82 (2008).
169. Huang, N. *et al.* Two distinct nuclear receptor interaction domains in NSD1, a novel SET protein that exhibits characteristics of both corepressors and coactivators. *EMBO J.* **17,** 3398–412 (1998).
170. Wang, X. *et al.* Identification and characterization of a novel androgen receptor coregulator ARA267-alpha in prostate cancer cells. *J. Biol. Chem.* **276,** 40417–23 (2001).
171. Wang, G. G., Cai, L., Pasillas, M. P. & Kamps, M. P. NUP98-NSD1 links H3K36 methylation to Hox-A gene activation and leukaemogenesis. *Nat. Cell Biol.* **9,** 804–12 (2007).
172. Thanasopoulou, A., Tzankov, A. & Schwaller, J. Potent co-operation between the NUP98-NSD1 fusion and the FLT3-ITD mutation in acute myeloid leukemia induction. *Haematologica* **99,** 1465–71 (2014).
173. Rayasam, G. V. *et al.* NSD1 is essential for early post-implantation development and has a catalytically active SET domain. *EMBO J.* **22,** 3153–63 (2003).
174. NP_071900 & NP_032765. Protein Blast. Available at: http://blast.ncbi.nlm.nih.gov/Blast.cgi. (Accessed: 2nd August 2016)
175. La Starza, R. *et al.* Cryptic insertion producing two NUP98/NSD1 chimeric transcripts in adult refractory anemia with an excess of blasts. *Genes Chromosom. Cancer* **41,** 395–399 (2004).
176. Walter, M. J. *et al.* Acquired copy number alterations in adult acute myeloid leukemia genomes. *Proc. Natl. Acad. Sci. U. S. A.* **106,** 12950–5 (2009).
177. Dolnik, A. *et al.* Commonly altered genomic regions in acute myeloid leukemia are enriched for somatic mutations involved in chromatin remodeling and splicing. *Blood* **120,** e83-92 (2012).
178. Garg, M. *et al.* Profiling of somatic mutations of acute myeloid leukemia, FLT3-ITD subgroup at diagnosis and relapse. *Blood* **126,** 2491–2501 (2015).
179. Berdasco, M. *et al.* Epigenetic inactivation of the Sotos overgrowth syndrome

gene histone methyltransferase NSD1 in human neuroblastoma and glioma. *Proc. Natl. Acad. Sci. U. S. A.* **106,** 21830–5 (2009).

180. Hamamoto, R., Saloura, V. & Nakamura, Y. Critical roles of non-histone protein lysine methylation in human tumorigenesis. *Nat. Publ. Gr.* **15,** 110–124 (2015).
181. Bianco-Miotto, T. *et al.* Global levels of specific histone modifications and an epigenetic gene signature predict prostate cancer progression and development. *Cancer Epidemiol. Biomarkers Prev.* **19,** 2611–2622 (2010).
182. Zhao, Q. *et al.* Transcriptome-guided characterization of genomic rearrangements in a breast cancer cell line. *Proc. Natl. Acad. Sci. U. S. A.* **106,** 1886–1891 (2009).
183. Wagner, E. J. & Carpenter, P. B. Understanding the language of Lys36 methylation at histone H3. *Nat. Rev. Mol. Cell Biol.* **13,** 115–26 (2012).
184. Kudithipudi, S., Lungu, C., Rathert, P., Happel, N. & Jeltsch, A. Substrate Specificity Analysis and Novel Substrates of the Protein Lysine Methyltransferase NSD1. *Chem. Biol.* **21,** 1–12 (2014).
185. Lu, T. *et al.* Regulation of NF-kappaB by NSD1/FBXL11-dependent reversible lysine methylation of p65. *Proc. Natl. Acad. Sci. U. S. A.* **107,** 46–51 (2010).
186. Qiao, Q. *et al.* The structure of NSD1 reveals an autoregulatory mechanism underlying histone H3K36 methylation. *J. Biol. Chem.* **286,** 8361–8 (2011).
187. Lucio-Eterovic, A. K. *et al.* Role for the nuclear receptor-binding SET domain protein 1 (NSD1) methyltransferase in coordinating lysine 36 methylation at histone 3 with RNA polymerase II function. *Proc. Natl. Acad. Sci. U. S. A.* **107,** 16952–7 (2010).
188. Kurotaki, N. *et al.* Haploinsufficiency of NSD1 causes Sotos syndrome. *Nat. Genet.* **30,** 365–6 (2002).
189. Fickie, M. R. *et al.* Adults with Sotos syndrome: review of 21 adults with molecularly confirmed NSD1 alterations, including a detailed case report of the oldest person. *Am. J. Med. Genet. A* **155A,** 2105–11 (2011).
190. Qiao, Q. *et al.* The structure of NSD1 reveals an autoregulatory mechanism underlying histone H3K36 methylation. *J. Biol. Chem.* **286,** 8361–8368 (2011).
191. Pasillas, M. P., Shah, M. & Kamps, M. P. NSD1 PHD domains bind methylated H3K4 and H3K9 using interactions disrupted by point mutations in human sotos syndrome. *Hum. Mutat.* **32,** 292–298 (2011).

192. Lucio-Eterovic, A. K. & Carpenter, P. B. An open and shut case for the role of NSD proteins as oncogenes. *Transcription* **2,** 158–161 (2011).
193. Keats, J. J. *et al.* Overexpression of transcripts originating from the MMSET locus characterizes all t(4;14)(p16;q32)-positive multiple myeloma patients. *Blood* **105,** 4060–9 (2005).
194. Kuo, A. J. *et al.* NSD2 links dimethylation of histone H3 at lysine 36 to oncogenic programming. *Mol. Cell* **44,** 609–20 (2011).
195. Morishita, M. & di Luccio, E. Structural insights into the regulation and the recognition of histone marks by the SET domain of NSD1. *Biochem. Biophys. Res. Commun.* **412,** 214–9 (2011).
196. Vougiouklakis, T. & Hamamoto, R. The NSD family of protein methyltransferases in human cancer. *Epigenomics* **7,** 1–12 (2015).
197. Morishita, M. & di Luccio, E. Cancers and the NSD family of histone lysine methyltransferases. *Biochim. Biophys. Acta* **1816,** 158–63 (2011).
198. Rosati, R. NUP98 is fused to the NSD3 gene in acute myeloid leukemia associated with t(8;11)(p11.2;p15). *Blood* **99,** 3857–3860 (2002).
199. Angrand, P. O. *et al.* NSD3, a new SET domain-containing gene, maps to 8p12 and is amplified in human breast cancer cell lines. *Genomics* **74,** 79–88 (2001).
200. Au, K. F., Jiang, H., Lin, L., Xing, Y. & Wong, W. H. Detection of splice junctions from paired-end RNA-seq data by SpliceMap. *Nucleic Acids Res.* **38,** 4570–4578 (2010).
201. Langmead, B., Trapnell, C., Pop, M. & Salzberg, S. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol.* **10,** R25 (2009).
202. Gaidatzis, D., Lerch, A., Hahne, F. & Stadler, M. B. QuasR: Quantification and annotation of short reads in R. *Bioinformatics* **31,** 1130–1132 (2015).
203. England, S. J., McGrath, K. E., Frame, J. M. & Palis, J. Immature erythroblasts with extensive ex vivo self-renewal capacity emerge from the early mammalian fetus. *Blood* **117,** 2708–2717 (2011).
204. Kudithipudi, S., Kusevic, D., Weirich, S. & Jeltsch, A. Specificity analysis of protein lysine methyltransferases using SPOT peptide arrays. *J. Vis. Exp.* **di,** e52203–e52203 (2014).
205. Georgiades, P. *et al.* VavCre transgenic mice: a tool for mutagenesis in hematopoietic and endothelial lineages. *Genesis* **34,** 251–6 (2002).

206. Ogilvy, S. *et al.* Transcriptional regulation of vav, a gene expressed throughout the hematopoietic compartment. *Blood* **91,** 419–430 (1998).
207. Gan, T., Jude, C. D., Zaffuto, K. & Ernst, P. Developmentally induced Mll1 loss reveals defects in postnatal haematopoiesis. *Leukemia* **24,** 1732–41 (2010).
208. Choe, K. S. *et al.* Reversal of tumorigenicity and the block to differentiation in erythroleukemia cells by GATA-1. *Cancer Res.* **63,** 6363–6369 (2003).
209. Whyatt, D. J. *et al.* The level of the tissue-specific factor GATA-1 affects the cell-cycle machinery. *Genes Funct.* **1,** 11–24 (1997).
210. Whyatt, D. *et al.* An intrinsic but cell-nonautonomous defect in GATA-1-overexpressing mouse erythroid cells. *Nature* **406,** 519–24 (2000).
211. Ferreira, R. *et al.* Dynamic regulation of Gata factor levels is more important than their identity. **109,** 5481–5491 (2015).
212. Gautier, E.-F. *et al.* Comprehensive Proteomic Analysis of Human Erythropoiesis. *Cell Rep.* **16,** 1470–1484 (2016).
213. Chyla, B. J. *et al.* Deletion of Mtg16, a target of t(16;21), alters hematopoietic progenitor cell proliferation and lineage allocation. *Mol. Cell. Biol.* **28,** 6234–6247 (2008).
214. Wadman, I. *et al.* Specific in vivo association between the bHLH and LIM proteins implicated in human T cell leukemia. **13,** 4831–4839 (1994).
215. Kadrmas, J. L. & Beckerle, M. C. The Lim domain: From the cytoskeleton to the nucleus. *Nat. Rev. Mol. Cell Biol.* **5,** 920–931 (2004).
216. Bach, I. The LIM domain : regulation by association. *Mech. Dev.* **91,** 5–17 (2000).
217. Anguita, E. *et al.* Globin gene activation during haemopoiesis is driven by protein complexes nucleated by GATA-1. **23,** 2841–2852 (2004).
218. Lubas, M. *et al.* Interaction Profiling Identifies the Human Nuclear Exosome Targeting Complex. *Mol. Cell* **43,** 624–637 (2011).
219. Perez-Burgos, L. *et al.* Generation and Characterization of Methyl-Lysine Histone Antibodies. *Methods Enzymol.* **376,** 234–254 (2004).
220. An, X. *et al.* Global transcriptome analyses of human and murine terminal erythroid differentiation Global Transcriptome Analyses of Human and Murine Terminal Erythroid Differentiation. *Blood* **123,** 3466–3478 (2014).
221. Drexler, H. G., Matsuo, Y. & MacLeod, R. A. F. Malignant hematopoietic cell lines: In vitro models for the study of erythroleukemia. *Leuk. Res.* **28,** 1243–1251 (2004).

222. de Boer, J. *et al.* Transgenic mice with hematopoietic and lymphoid specific expression of Cre. *Eur. J. Immunol.* **33,** 314–25 (2003).

223. Farber, D. L., Yudanin, N. a & Restifo, N. P. Human memory T cells: generation, compartmentalization and homeostasis. *Nat. Rev. Immunol.* **14,** 24–35 (2014).

224. Kurosaki, T., Kometani, K. & Ise, W. Memory B cells. *Nat. Rev. Immunol.* **15,** 149–59 (2015).

225. Moreau-Gachelin, F., Tavitian, A. and Tambourin, P. Spi-1 is a putative oncogene in virally induced murine erythroleukaemias. *Nature* **331,** 277–280 (1988).

226. Constantinescu, S. N. *et al.* Activation of the erythropoietin receptor by the gp55-P viral envelope protein is determined by a single amino acid in its transmembrane domain. *EMBO J.* **18,** 3334–3347 (1999).

227. Okuno, Y. *et al.* Potential Autoregulation of Transcription Factor PU . 1 by an Upstream Regulatory Element. **25,** 2832–2845 (2005).

228. Blair, D. G. & Athanasiou, M. Ets and retroviruses - transduction and activation of members of the Ets oncogene family in viral oncogenesis. *Oncogene* **19,** 6472–81 (2000).

229. Clausen, P. A., Mavrothalassitis, J., Watson, K. & Biair, D. G. Molecular and Structural Biology, Medical. **7,** 1525–1534 (1996).

230. Amanatiadou, E. P., Papadopoulos, G. L., Strouboulis, J. & Vizirianakis, I. GATA1 and PU . 1 Bind to Ribosomal Protein Genes in Erythroid Cells : Implications for Ribosomopathies. *PLoS One* 1–18 (2015). doi:10.1371/journal.pone.0140077

231. Burda, P. *et al.* GATA-1 Inhibits PU . 1 Gene via DNA and Histone H3K9 Methylation of Its Distal Enhancer in Erythroleukemia. *PLoS One* **11,** 1–18 (2016).

232. Chou, S. T. *et al.* Graded repression of PU . 1 / Sfpi1 gene transcription by GATA factors regulates hematopoietic cell fate. *Blood* **114,** 983–995 (2009).

233. Grossmann, V. *et al.* Acute erythroid leukemia (AEL) can be separated into distinct prognostic subsets based on cytogenetic and molecular genetic characteristics. *Leukemia* **27,** 1940–3 (2013).

234. Kasyan, A. *et al.* Acute erythroid leukemia as defined in the World Health Organization classification is a rare and pathogenetically heterogeneous disease. *Mod. Pathol.* **23,** 1113–1126 (2010).

235. Li, Y. *et al.* Regulation of the PU. 1 gene by distal elements. *Blood* **98,** 2958–2965 (2001).
236. Fan, T. *et al.* DNA hypomethylation caused by Lsh deletion promotes erythroleukemia development. *Epigenetics* **3,** 134–142 (2008).
237. Li, Y. *et al.* The miR-17-92 cluster expands multipotent hematopoietic progenitors whereas imbalanced expression of its individual oncogenic miRNAs promotes leukemia in mice. *Blood* **119,** 4486–4499 (2012).
238. Radke, K., Beug, H., Kornfeld, S. & Graf, T. Transformation of Both Erythroid and Myeloid Cells by E26 , an Avian Leukemia Virus That Contains the myb Gene. *Cell* **31,** 643–653 (1982).
239. Skoda, R. C., Tsai, S.-F., Orkin, S. H. & Leder, P. Expression of c-MYC under the Control of GATA-1 Regulatory Sequences Causes Erythroleukemia in Transgenic. *J Exp Med* **181,** 479–489 (1995).
240. Trempus, C. S. *et al.* Association of v-Ha- ras Transgene Expression with Development of Erythroleukemia in Tg.AC Transgenic Mice. *Am. J. Pathol.* **153,** 247–254 (1998).
241. Lacombe, C. & Mayeux, P. The molecular biology of erythropoietin. *Nephrol Dial Transpl.* **14,** 22–28 (1999).
242. Gregory, B. T. *et al.* GATA-1 and Erythropoietin Cooperate to Promote Erythroid Cell Survival by Regulating bcl-x expression. *Blood* **1,** 87–96 (1999).
243. Moreau-Gachelin, F. Lessons from models of murine erythroleukemia to acute myeloid leukemia (AML): proof-of-principle of co-operativity in AML. *Haematologica* **91,** 1644–1652 (2006).
244. Vennström, B. & Bishop, J. M. Isolation and Characterization of Chicken DNA Homologous to the Two Putative Oncogenes of Avian Erythroblastosis Virus. *Cell* **28,** 135–143 (1982).
245. V Voisin, C Barat, T Hoang, E. R. Novel Insights into the Pathogenesis of the Graffi Murine Leukemia Retrovirus. *J Virol.* **80,** 4026–4037 (2006).
246. McDevitt, M. A., Shivdasani, R. A., Fujiwara, Y., Yang, H. & Orkin, S. H. A 'knockdown' mutation created by cis-element gene targeting reveals the dependence of erythroid cell maturation on the level of transcription factor GATA-1. *Proc. Natl. Acad. Sci.* **94,** 6781–6785 (1997).
247. Fan, T. *et al.* DNA hypomethylation caused by Lsh deletion promotes erythroleukemia development. *Epigenetics* **3,** 134–142 (2008).
248. Shi, H. *et al.* ASXL1 plays an important role in erythropoiesis. *Sci. Rep.* **6,**

28789 (2016).

249. Carmichael, C. L. *et al.* Hematopoietic overexpression of the transcription factor Erg induces lymphoid and erythro-megakaryocytic leukemia. *Proc. Natl. Acad. Sci. U. S. A.* **109,** 15437–42 (2012).

250. Katsuhiko Itoh, Katsuhiko Ono, K. J. M. Establishment and characterization of a transplantable erythroblastic leukemia in C3H mice. *Leuk. Res.* **12,** 471–478 (1988).

251. Soler, E. *et al.* The genome-wide dynamics of the binding of Ldb1 complexes during erythroid differentiation. *Genes Dev.* **24,** 277–289 (2010).

252. Davis, J. N., Williams, B. J., Herron, J. T., Galiano, F. J. & Meyers, S. ETO-2 , a new member of the ETO-family of nuclear proteins. *Oncogene* **18,** 1375–1383 (1999).

253. Gamou, T. *et al.* The Partner Gene of AML1 in t(16;21) Myeloid Malignancies Is a Novel Member of the MTG8(ETO) Family. *Blood* **11,** 4028–4037 (1998).

254. Fujiwara, T. *et al.* Role of transcriptional corepressor ETO2 in erythroid cells. *Exp. Hematol.* **41,** 303–15.e1 (2013).

255. Micci, F. *et al.* High-throughput sequencing identifies an NFIA/CBFA2T3 fusion gene in acute erythroid leukemia with t(1;16)(p31;q24). *Leukemia* **27,** 2012–2014 (2013).

256. Micci, F. *et al.* Translocation t(1;16)(p31;q24) rearranging CBFA2T3 is specific for acute erythroid leukemia. *Leuk. Off. J. Leuk. Soc. Am. Leuk. Res. Fund, U.K* **25,** 1510–2 (2011).

257. Starnes, L. M. *et al.* NFI-A directs the fate of hematopoietic progenitors to the erythroid or granulocytic lineage and controls beta-globin and G-CSF receptor expression. *Blood* **114,** 1753–1763 (2009).

258. Hildebrand, D., Tiefenbach, J., Heinzel, T., Grez, M. & Maurer, A. B. Multiple Regions of ETO Cooperate in Transcriptional Repression. *J. Biol. Chem.* **276,** 9889–9895 (2001).

259. Gruber, T. A. *et al.* An Inv(16)(p13.3q24.3)-Encoded CBFA2T3-GLIS2 Fusion Protein Defines an Aggressive Subtype of Pediatric Acute Megakaryoblastic Leukemia. *Cancer Cell* **22,** 683–697 (2012).

260. Park, I. Y. *et al.* Dual Chromatin and Cytoskeletal Remodeling by SETD2. *Cell* **166,** 950–962 (2016).

261. Joan Boyes, Peter Byield, Yoshihiro Nakatani, V. O. Regulation of activity of the transcription factor GATA-1 by acetylation. *Nature* **428,** 153–157 (2004).

262. Blobel, G. a. CREB-binding protein and p300: molecular integrators of hematopoietic transcription. *Blood* **95,** 745–755 (2000).
263. Lamonica, J. M., Vakoc, C. R. & Blobel, G. A. Acetylation of GATA-1 is required for chromatin occupancy. *Blood* **108,** 3736–3738 (2006).
264. Labbaye, C. *et al.* PLZF induces megakaryocytic development , activates Tpo receptor expression and interacts with GATA1 protein. *Oncogene* **21,** 6669–6679 (2002).
265. Rao, B. *et al.* Dimethylation of Histone H3 at Lysine 36 Demarcates Regulatory and Nonregulatory Dimethylation of Histone H3 at Lysine 36 Demarcates Regulatory and Nonregulatory Chromatin Genome-Wide. *Mol Cell Biol* **25,** 9447–9459 (2005).
266. Aymard, F. *et al.* Transcriptionally active chromatin recruits homologous recombination at DNA double-strand breaks. *Nat. Struct. Mol. Biol.* **21,** 366–74 (2014).
267. Li, F. *et al.* The histone mark H3K36me3 regulates human DNA mismatch repair through its interaction with MutSalpha. *Cell* **153,** 590–600 (2013).
268. Carvalho, S. *et al.* SETD2 is required for DNA double-strand break repair and activation of the p53-mediated checkpoint. *Elife* **2014,** 1–19 (2014).
269. Pfister, S. X. *et al.* SETD2-Dependent Histone H3K36 Trimethylation Is Required for Homologous Recombination Repair and Genome Stability. *Cell Rep.* **7,** 2006–2018 (2014).
270. Zhu, X. *et al.* Identification of functional cooperative mutations of SETD2 in human acute leukemia. *Nat. Genet.* **46,** 287–293 (2014).
271. Li, J. *et al.* SETD2: an epigenetic modifier with tumor suppressor functionality. *Oncotarget* **7,** 50719–50734 (2015).
272. Baubec, T. *et al.* Genomic profiling of DNA methyltransferases reveals a role for DNMT3B in genic methylation. *Nature* **520,** 243–7 (2015).
273. Suzuki, M. *et al.* Site-specific DNA methylation by a complex of PU.1 and Dnmt3a/b. *Oncogene* **25,** 2477–2488 (2006).
274. Choufani, S. *et al.* NSD1 mutations generate a genome-wide DNA methylation signature. *Nat. Commun.* **6,** 10207 (2015).
275. Lim, W. F., Inoue-Yokoo, T., Tan, K. S., Lai, M. & Sugiyama, D. Hematopoietic cell differentiation from embryonic and induced pluripotent stem cells. *Stem Cell Res. Ther.* **4,** 71 (2013).
276. Im, H. Mouse Embryonic Stem Cell Differentiation to Hematopoietic

Precursors. *Bio-protocol* **2,** 3–8 (2012).

277. Carotta, S. *et al.* Directed differentiation and mass cultivation of pure erythroid progenitors from mouse embryonic stem cells Directed differentiation and mass cultivation of pure erythroid progenitors from mouse embryonic stem cells. **104,** 1873–1880 (2012).
278. Santos, F. P. S. *et al.* Adult acute erythroleukemia: an analysis of 91 patients treated at a single institution. *Leukemia* **23,** 2275–80 (2009).
279. Kowal-Vern, A. *et al.* Diagnosis and Characterization of Acute Erythroleukemia Subsets by Determining the Percentages of Myeloblasts and Proerythroblasts in 69 Cases. *Am. jorunal Hematol.* **65,** 5–13 (2000).
280. Hasserjian, R. P. *et al.* Acute erythroid leukemia : a reassessment using criteria refined in the 2008 WHO classification. *Blood* **115,** 1985–1993 (2010).
281. Bacher, U. *et al.* Comparison of genetic and clinical aspects in patients with acute myeloid leukemia and myelodysplastic syndromes all with more than 50 % of bone marrow erythropoietic cells. *Haematologica* **96,** 1284–1292 (2011).
282. Doulatov, S., Notta, F., Laurenti, E. & Dick, J. E. Hematopoiesis: a human perspective. *Cell Stem Cell* **10,** 120–36 (2012).
283. Boultwood, J., Pellagatti, A., Mckenzie, A. N. J. & Wainscoat, J. S. Review article Advances in the 5q- syndrome. *Blood* **116,** 5803–5811 (2010).
284. Greif, P. a *et al.* Identification of recurring tumor-specific somatic mutations in acute myeloid leukemia by transcriptome sequencing. *Leukemia* **25,** 821–7 (2011).
285. Popova, E. Y. *et al.* Chromatin condensation in terminally differentiating mouse erythroblasts does not involve special architectural proteins but depends on histone deacetylation. *Chromosom. Res.* **17,** 47–64 (2009).
286. Wu, W. *et al.* Dynamics of the epigenetic landscape during erythroid differentiation after GATA1 restoration. *Genome Res.* **21,** 1659–1671 (2011).
287. Hom, J., Dulmovits, B. M., Mohandas, N. & Blanc, L. The erythroblastic island as an emerging paradigm in the anemia of inflammation. *Immunol. Res.* **63,** 75–89 (2015).
288. Kadri, Z. *et al.* Direct binding of pRb/E2F-2 to GATA-1 regulates maturation and terminal cell division during erythropoiesis. *PLoS Biol.* **7,** (2009).
289. de Back, D. Z., Kostova, E. B., van Kraaij, M., van den Berg, T. K. & van Bruggen, R. Of macrophages and red blood cells; A complex love story. *Front. Physiol.* **5 JAN,** (2014).

290. Ramos, P. *et al.* Macrophages support pathological erythropoiesis in polycythemia vera and β-thalassemia. *Nat. Med.* **19,** 437–45 (2013).
291. Pierdomenico, F. & Almeida, A. Treatment of acute erythroleukemia with Azacitidine: A case series. *Leuk. Res. Reports* **2,** 41–43 (2013).
292. Hopfer, O. *et al.* Epigenetic dysregulation of GATA1 is involved in myelodysplastic syndromes dyserythropoiesis. *Eur. J. Haematol.* **88,** 144–153 (2012).
293. Filippakopoulos, P. *et al.* Histone recognition and large-scale structural analysis of the human bromodomain family. *Cell* **149,** 214–31 (2012).
294. Delmore, J. E. *et al.* BET Bromodomain Inhibition as a Therapeutic Strategy to Target c-Myc. *Cell* **146,** 904–917 (2011).
295. Picaud, S. *et al.* PFI-1 – A highly Selective Protein Interaction Inhibitor Targeting BET Bromodomains. *Cancer Res.* **73,** 3336–3346 (2013).
296. Zuber, J. *et al.* RNAi screen identifies Brd4 as a therapeutic target in acute myeloid leukaemia. *Nature* **478,** 524–528 (2011).
297. Picaud, S. *et al.* Generation of a selective small molecule inhibitor of the CBP/p300 bromodomain for leukemia therapy. *Cancer Res.* **75,** 5106–5120 (2015).
298. Picaud, S. *et al.* Promiscuous targeting of bromodomains by Bromosporine identifies BET proteins as master regulators of primary transcription response in leukemia. *Sci. Adv.* **2,** e1600760 (2016).

6 ABBREVIATIONS

5mCpG	5- methyl cytosine- phosphatidyl- guanosine island
ACK	Ammonium Chloride Potassium
AEL	acute erythroleukemia
AF	activation function
AGM	aorta–gonad–mesonephros
ALL	acute lymphoblastic leukemia
AML	acute myeloid leukemia
AR	androgen receptor
ASH2L	Ash2 (Absent, Small, Or Homeotic)-Like (Drosophila)
ATRA	all- trans retinoic acid
Bcl2l1	BCL2 like 1
BET	Bromodomain and Extra-Terminal motif
BFU-E	Burst forming unit- erythroid
BM	Bone marrow
BMP4	bone morphogenic protein 4
BRD	Bromodomain
BSP	Bromosporine
C/EBPα	CCAAT/enhancer binding protein
CBFA2T3	CBFA2/RUNX1 Translocation Partner 3
CBP	cAMP- responsive element - binding protein
CD	cluster of differentiation
CDK-1	cyclin dependent kinase inhibitor 1
CFU-E	Colony forming unit- erythroid
CFU-G/M	Colony forming unit- granulocyte/macrophage
ChIP	chromatin immuneprecipitation
ChIP-Seq	chromatin immuneprecipitation Sequencing
CLP	common lymphoid progenitors
CMP	common myeloid progenitors
CN	cytogenetically normal
DAPI	4′,6-Diamidin-2-phenylindol
DBD	DNA binding domain
DNA	Desoxyribonucleic acid
DNMT	DNA (Cytosine-5-)-Methyltransferase
DS-AMKL	Down's syndrome-related acute megakaryocytic leukemia
DTT	1,4-Dithiothreitol
E	embryonic day
E2A	Transcription Factor 3
EDTA	diaminoethanetetraaceticacid
EGFP	Enhanced green fluorescent protein
EGTA	Ethylenglycol-bis(aminoethylether)-N,N,N′,N′-tetra acetic acid
EMSA	electrophoretic mobility shift assay
EMT	epithelial- mesenchymal- transition

EPO	Erythropoietin
EPOR	Erythropoietin receptor
EryD	definitive erythroid precursor
EryP	primitive erythroid precursor
ES	embryonic stem cells
ESRE	Extensive self renewing erythroblast
ETS	E-Twenty-Six
F-MuLV	Friend murine leukemia virus
FAB	French- American- British
FcγR	Fc gamma receptor
FG	phenylalanine- glycine
FLT3-ITD	FMS-like tyrosine kinase-3 internal tandem duplication
FOG1	Friend of GATA1
G1HE	hematopoietic enhancer element in GATA1
GATA1-s	shorter GATA1 isoform
GFI1B	Growth Factor Independent 1B Transcription Repressor
GM-CSF	granulocyte- macrophage colony stimulating factor
GMP	granulocyte- macrophage progenitors
GNAT	Gcn5 N-acetyltransferases
H&E	hematoxylin and eosin
H3	Histone 3
HAT	histone acetyltransferases
HCl	Hydrochloric acid
HDAC	histone deacetylases
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
HGB	hemoglobin
HIF1α	hypoxia inducible factor 1 alpha
HIF1β	hypoxia inducible factor 1 beta
HMG	high mobility group
HMT	histone methyltransferases
HOX	homeobox
HP1	Heterochromatin Protein 1
HSC	hematopoietic stem cell
i.p.	intraperitoneal
IE	Erythroid alternative first exon of GATA1
IG	immunoglobuline
IGF-1	Insulin like growth factor 1
IT	Testis alternative first exon of GATA1
JAK2	Janus Kinase 2
KCl	Kalium chloride
KLF1	Krüppel- like- factor 1
LDB1	LIM Domain Binding 1
LEDGF	Lens-epithelial growth factor
LIM	Lin11, Isl-1, Mec-3
LMO2	LIM Domain Only 2
LS	lineage marker-negative, C-KIT^{+}/SCA-1^{-}

LSH	lymphoid specific helicase
LSK	lineage marker-negative, C-KIT$^+$/SCA-1$^+$
LT-HSC	long- term repopulating stem cells
LTR	Long tandem repeat
LUC	Large unstained cells
MDS	myelodysplastic syndromes
ME26	murine E26
MEL	murine erythroleukemia cell line
Menin	multiple endocrine neoplasia type 1
MEP	megakaryocyte- erythroid progenitors
MK	Megakaryocytes
MLL1	mixed linage leukemia 1
MPP	multi-potent progenitors
Mx1	MX Dynamin Like GTPase 1
N-/C-finger	N-C-terminal located zinc finger of GATA1
NaCl	Natrium chloride
NF-E2	nuclear factor erythroid 2
NFIA	nuclear factor 1 A
NID	nuclear receptor- interacting binding domains
NK	Natural killer cell
Npm1	Nucleophosmin 1
NSD1	Nuclear Receptor Binding SET Domain Protein 1
NUP98	nucleoporin 98
PBS	Phosphate buffered saline
PCA	Principal component analysis
PCAF	p300/CBP-associated factor
PDS	Plasma derived serum
PFHMII	Protein Free Hybridoma Medium II
pgk	phosphoglycerate kinase
PHD	plant homeodomain
PLT	platelet
PML	promyelocytic leukemia protein
pMSCV	Plasmid murine stem cell virus
poly(I:C)	polyinosinic:polycytidylic acid
pRB	Phosphorlytaed Retinoblastoma 1
PWWP	Pro-Trp-Trp-Pro
RAR	retinoid acid receptor alpha
RARα	retinoic acid receptor alpha
RBBP5	Retinoblastoma Binding Protein 5
RBC	Red blood cells
RNA	ribonucleic acid
RTC	reticulocyte
RUNX1	Runt- related transcription factor 1
RXR	retinoid X receptor gamma
SCF	stem cell factor
SDF-1	Stromal cell-derived factor 1

SET	Su(var)3-9, Enhancer-of-zeste, Trithorax
SETD2	SET Domain Containing 2
sf-STK	stem cell kinase receptor
SFFV	Spleen Focus Forming Virus
SPI-1	Spleen Focus Forming Virus Proviral Integration Oncogene 1
ST-HSC	short- tem repopulating stem cells
STAT	Signal Transducers and Activators of Transcription
SUV39H	Suppressor Of Variegation 3-9 Homolog 1
TAD	transactivating domain
TAL1	T-Cell Acute Lymphocytic Leukemia 1
TAL1=SCL	T-Cell Acute Lymphocytic Leukemia 1
TATA	TATAAA
TF	transcription factors
TMD	transient myeloproliferative disorder
TPO	Thrombopoietin
TpoR	Tpo receptor
TR	thyroid hormone receptor
TSS	transcriptional start sites
TUNEL	TdT-mediated dUTP-biotin nick end labeling
Tx	Tamoxifen
WBC	White blood cells
WDR5	WD Repeat Domain 5
WHSC1	Wolf-Hirschhorn Syndrome Candidate 1
WHSCL1	Wolf-Hirschhorn Syndrome Candidate 1-Like 1
xMELP	expedited microsphere HpaII small fragment Enrichment by Ligation-mediated PCR
ZBTB16=PLZF	Zinc finger and BTB domain-containing protein 16