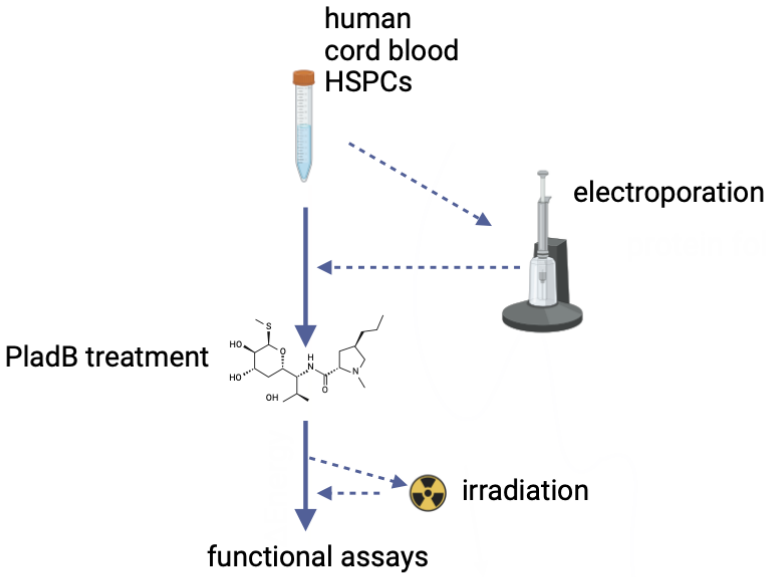
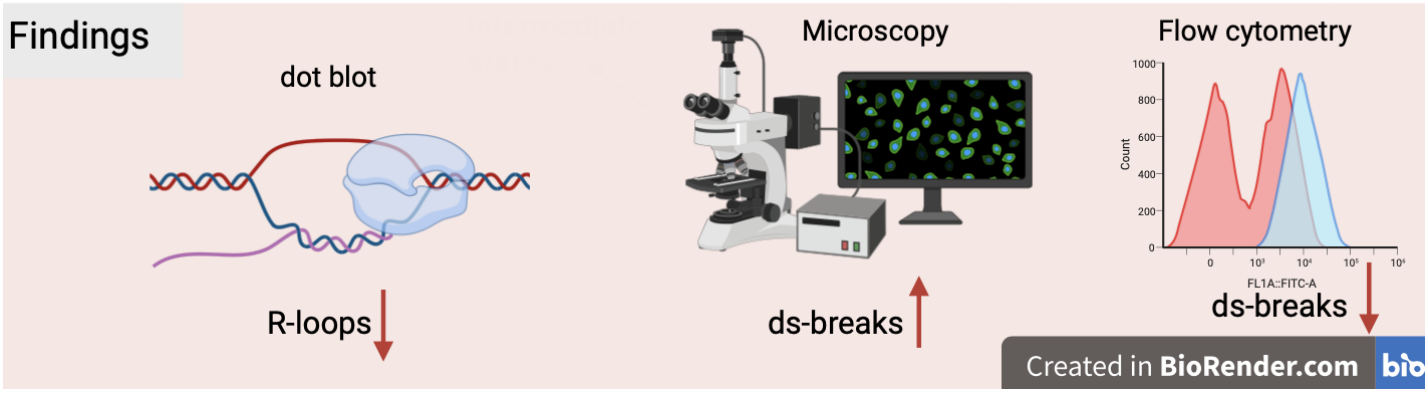
Deregulated splicing induces DNA damage and interferes with DNA repair in the hematopoietic system





Author	Suzanne van Dreumel (6383017) Master Molecular and Cellular Life Sciences, GSLS, Utrecht University
Date	November, 2022 – August, 2023
Correspondence	s.vandreumel@students.uu.nl
Supervisor	Lucca Derks, MSc
First examiner	Dr. Ruben van Boxtel
Second examiner	Dr. Hugo Snippert

Van Boxtel LAB

Princess máxima center pediatric oncology

Oncode Institute
Outsmarting cancer
Impacting lives

Abstract

Leukemia is the most prevalent form of pediatric cancer, among which acute myeloid leukemia (AML) is the second most common form of pediatric leukemia. Even though the overall survival rate increased over the past decades, still 1 out of 4 children with AML does not reach the 5-year survival after diagnosis. Even if the children survive, they often encounter severe therapy-related side effects. Recently, it was found that splicing is often disturbed in pediatric AML (pAML). This finding led to the question whether splicing deregulation is involved in the development of pAML, since splicing factor hotspot mutations are also frequently found in adult AML. In this study, we mimicked splicing deregulation by treating human hematopoietic stem- and progenitor cells (HSPCs) with the splicing inhibitor pladienolide B. In addition, we generated a DNA methyltransferase 3A (DNMT3A) knock-out, since DNMT3A is one of the most prevalent mutated genes in adult AML and recently it was found that DNMT3A is also involved in splicing regulation. We observed an increase in double-stranded breaks and a major decrease in R-loops in splicing deregulated HSPCs. Interestingly, this decrease in R-loops is contradictory to the increase that was often observed in comparable studies with different cell lines. Future studies about splicing deregulation in the hematopoietic system can eventually contribute to prevent the development of pAML and stimulate the development of new targeted therapies.

Table of contents

Abstract 2
Introduction 4
Processes underlying the development of AML 4
Functional consequences of deregulated splicing 4
DNMT3A involved in splicing regulation 5
Inducing splicing deregulation in human hematopoietic stem- and progenitor cells 6
Results 7
Generation and validation of aberrant splicing models in human HSPCs 7
Genetic engineering of human HSPCs to modulate splicing in a direct manner 7
CRISPR base editors not functional after mRNA transfection 7
Alternative models to modulate splicing 8
Functional read-outs of aberrant splicing models in human HSPCs 10
Increase in ds-breaks upon PladB treatment 10
Bulk analysis shows decrease in ds-breaks upon PladB treatment and DNMT3A knock-out 10
Decrease in R-loops upon splicing deregulation 12
Activation of the DNA damage response in HSPCs not detected upon splicing deregulation 12
XPA expression may be decreased via nonsense-mediated decay upon splicing deregulation 14
Discussion 16
Contradictory ds-breaks outcomes 16
Expression and activation of DNA damage response genes in HSPCs below detection limit 16
R-loop level may be dependent on transcriptional activity of cells 17
Altered transcriptional activity may contribute to the development of pAML 17
In vitro translation could resolve base editing issues 18
Acknowledgements 19
Materials and methods 20
Cell culture 20
Magnetic cell separation (MACS) of $CD34^+$-HSPCs 20
GC base editor plasmid maxi-prep 20
CRISPR base editor mRNA generation 20
Human cord blood-derived HSPC electroporation 21
Kasumi-1 cell electroporation and optimization 21
Sanger sequencing genome editing 21
Imaging 22
Quantification p-γH2AX foci widefield microscopy 22
FACS intracellular p-γH2AX staining 22
Dot-blot for R-loop quantification 22
Western Blot 23
qPCR DNA damage response genes 23
References 24
Supplementary 29
Layman summary 29
Supplementary figures 30
Figure S1. FACS gating examples 30
Figure S2. S9.6 Ab binds nonspecific in cytoplasm 31
Figure S3. Western Blot DDR activation 32
Supplementary tables 33
Table S1. Used plasmids 33
Table S2. Used primers 34
Table S3. Significance bulk p-γH2AX data 36
Table S4. Imaging settings 37

Introduction

Leukemia is with a prevalence of 30 percent the most common form of pediatric cancer and is characterized by disturbed differentiation of hematopoietic cells (Metayer et al., 2016; Namayandeh et al., 2020). Acute lymphoblastic leukemia (ALL) is the most common form of pediatric leukemia, but also acute myeloid leukemia (AML) and myelodysplastic syndromes (MDS) are often observed in children (Cacace et al., 2022; Hofmann, 2015). The overall survival rate of pediatric ALL (pALL) is around 90 percent nowadays, but the long-term survival of pediatric AML (pAML) is only 75 percent (Devilli et al., 2021; Reinhardt et al., 2022). Despite the relatively high survival rate, the treatment often leads to severe long-term side effects, ranging from cardiovascular diseases to the development of a second, unrelated cancer (Reinhardt et al., 2022; X. Zhang et al., 2023). To date, the development of targeted therapies against pediatric leukemia is limited by lack of knowledge about what drives the development of pediatric leukemia (Bolouri et al., 2018; Hofmann, 2015; Lee et al., 2023). An increased understanding of what processes are involved in the development of pediatric leukemia can result in a better prognosis for pediatric AML (pAML) and will also enable the development of targeted therapies with fewer side-effects.

Processes underlying the development of AML

AML is a diverse clonal blood cancer, characterized by incomplete development of hematopoietic progenitor cells in the myeloid lineage (Madhusoodhan et al., 2016; Pimenta et al., 2021). As in other cancer types, environmental agents and genetic driver mutations contribute to the development of AML (Lee et al., 2023; Pimenta et al., 2021). Contradictory to other cancers, the somatic mutation burden in AML is low (Gröbner et al., 2018; Ma et al., 2018). This indicates that the development of AML does not solely depend on the accumulation of mutations over time, but also on cellular processes that influence gene expression. A potential process that may be involved in the development of AML is splicing. In adults, splicing factor genes are frequently mutated in myeloid malignancies (Hershberger et al., 2021; Taylor & Lee, 2019; Yoshida et al., 2011a). The mutations in splicing factors were often enriched at specific positions in the gene, so-called hotspot mutations, indicating positive selection for these mutations in myeloid malignancies. The hotspot mutations in adult AML were not found in pAML, but recently it was found that despite the absence of these hotspot mutations, splicing is often disturbed in pAML (van der Werf et al., 2023). The high frequency of deregulated splicing in both adult- and pAML indicates a role of splicing in the development of AML.

Functional consequences of deregulated splicing

In adults, around 50 percent of the MDS cases and 20 percent of the AML cases harbor a mutation in a splicing factor gene (Cheruiyot et al., 2021). Among these splicing factor genes, SF3B1 is the most common mutated splicing factor in MDS and also occurs in AML to a lesser extent (Huber et al., 2022; Singh et al., 2020). Mutant SF3B1 results in exon skipping and alternative 3’ splice side usage. The alternative splicing patterns caused by mutant SF3B1 were enriched in transcripts from genes involved in the DNA damage response (DDR) and replication pathways (Lappin et al., 2022; Liu et al., 2020). Due to the altered splicing, these transcripts are prone to nonsense-mediated decay (NMD), leading to decreased expression (Liu et al., 2020). Next to downregulation of DDR genes, also an increase in DNA damage has been observed upon SF3B1 inhibition or mutation (Lappin et al., 2022; Liu et al., 2020; Singh

et al., 2020). In cells with defective SF3B1, an increase in R-loops and double-stranded breaks (ds-breaks) was observed. R-loops are three-stranded structures consisting of DNA:RNA hybrids with a single-stranded DNA strand. R-loops normally occur in the genome, in which they have a functional role in gene expression regulation (Santos-Pereira & Aguilera, 2015). However, an increase in R-loops can cause transcriptional stress and genomic instability (Crossley et al., 2019; Santos-Pereira & Aguilera, 2015). The increase in ds-breaks observed in SF3B1-defective cells co-occurred with an increase in R-loops and was rescued upon treatment with RNase H (Chen et al., 2018; Cheruiyot et al., 2021). This implicates that defects in SF3B1 increase the number of R-loops in the DNA, which results in an increase in ds-breaks. Together, the decreased expression of DDR genes and increase in R-loops upon deregulated splicing can contribute to genomic instability and the accumulation of mutations and therefore also potentially cause MDS and AML (figure 1).

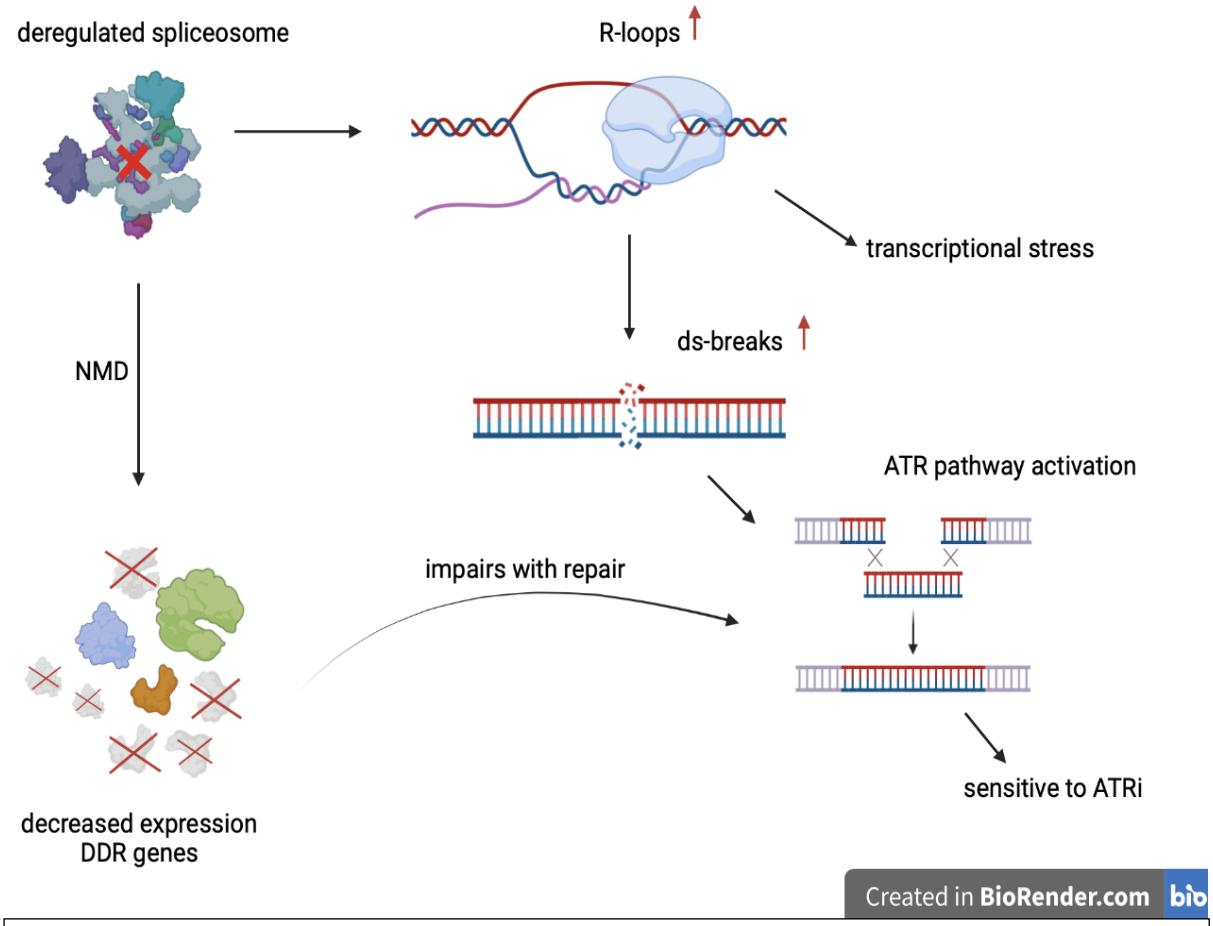


Figure 1: Overview of effects of deregulated splicing that potentially contribute to the development of AML. Deregulation of the spliceosome results in decreased expression of DDR genes via NMD. The decreased expression of DDR interferes with downstream ATR pathway, resulting in unrepaired ds-breaks that are caused by an increase in R-loops.

DNMT3A involved in splicing regulation

Next to hotspot mutations in splicing factors, splicing can also be disturbed indirectly, leading to the development of AML. For example, a knock-out of the DNA methyltransferase 3A (DNMT3A) also lead to aberrant splicing, altered DDR, and an increase in ds-breaks (Banaszak et al., 2018; Ramabadran et al., 2023). DNMT3A is indirectly involved in splicing by recruiting

splicing factors to active transcription sites and binds directly to SF3B1 (Ramabadran et al., 2023). The interaction between DNMT3A and the spliceosome governs a shift in stem cells towards a more differentiated state, which is independent of the methyltransferase function of DNMT3A. Since approximately 20 percent of the adult AML patients contain a mutation in DNMT3A and these mutations often co-occur with mutations in splicing factors, DNMT3A is likely involved in the emergence of AML via its role in splicing regulation (Banaszak et al., 2018; Dvinge et al., 2016a; Park et al., 2020). Like hotspot mutations in splicing factor genes, DNMT3A mutations are not found in pediatric AML (Bolouri et al., 2018). Nevertheless, it is hypothesized that mutations found in other genes in pAML, like WT1, have a similar effect as mutations in DNMT3A in adults, and can be used for risk stratification (Bolouri et al., 2018). Both the observation that a DNMT3A knock-out has similar effects as mutant splicing factors and the fact that DNMT3A is one of the most common mutated genes in adult AML, are indications that splicing is involved in the development of AML.

Inducing splicing deregulation in human hematopoietic stem- and progenitor cells

Already many studies have been performed to investigate the effect of splicing inhibition on the accumulation of DNA damage (Lappin et al., 2022; Singh et al., 2020; Sveen et al., 2016). These studies were often performed in leukemic cell lineages. The disadvantage of leukemic cell lineages is that the leukemic state of these cells could influence the observed effects of splicing perturbation. Besides that, it was found that splicing in tumor cells is different from splicing in healthy cells to support tumorigenesis (Y. Zhang et al., 2021). This makes it difficult to draw conclusions about a direct causal relationship between deregulated splicing and the development of AML when using leukemic cells. Next to leukemic cell lineages, also mice were used to study the effects of splicing factor hotspot mutations (Mupo et al., 2017). However, splicing in mice differs from splicing in humans, limiting the translation of those findings to human disease models. To overcome the limitations of leukemic cell lineages and mice, we made use of human cord blood-derived hematopoietic stem- and progenitor cells (HSPCs) as a model system to study the direct role of splicing deregulation in the emergence of AML. We chose HSPCs as model system, since these cells derive from a healthy donor and it is assumed that AML originates from these cells (Roboz & Guzman, 2009; Rosendahl Huber et al., 2022). We applied various methods to deregulate splicing in cord blood, in which we focused on both SF3B1 and DNMT3A, to see with different functional read-outs whether impairment with their function resulted in similar effects as shown above. Unfortunately, we did not manage to induce splicing factor hotspot mutations in HSPCs via CRISPR base editing, so we mimicked the splicing deregulation with the SF3B1 inhibitor pladienolide B (PladB). Remarkably, we saw a decrease in the number of R-loops in both splicing deregulated- and DNMT3A knock-out HSPCs. This indicates that *in vitro* in a healthy genetic background, R-loops do not lead to an increase in ds-breaks, and that there is presumably a difference in transcriptional activity between healthy- and splicing deregulated cells. Together, our results indicate that splicing deregulation can be involved in the development of AML via downregulation of R-loops, but more studies are necessary to confirm this hypothesis. This confirmation will enable the development of new targeted therapies against pAML and improve the prognosis in the future.

Results

Generation and validation of aberrant splicing models in human HSPCs

Genetic engineering of human HSPCs to modulate splicing in a direct manner

In this study, we tried to modulate splicing in different ways to answer the question whether deregulated splicing is involved in the emergence of AML. First, we aimed to obtain an *in vitro* system that closely represents splicing deregulation *in vivo*, by generating adult AML splicing factor hotspot mutations in human HSPCs (figure 2A). Here, we focused on the frequently occurring hotspot mutations U2AF1(Q157R), SF3B1(K666N) and SF3B1(K700E) (Seiler et al., 2018; Yoshida et al., 2011b). Since conventional CRISPR/Cas9 has a low homology-dependent repair editing efficiency, we used CRISPR base editing to generate the hotspot mutations, which was shown to be very efficient in HSPCs (Siegner et al., 2022). As CRISPR base editing proteins are not commercially available yet and electroporation of plasmids is toxic for HSPCs, we transfected cord blood derived HSPCs with mRNA encoding the base editor protein and the corresponding guide RNA (Lattanzi et al., 2019). We used Sanger sequencing to detect the presence of the hotspot mutation (figure 2A). Unfortunately, we did not observe any hotspot mutation in the electroporated HSPCs. This indicates that the base editing efficiency of our transfection is at least below the Sanger sequencing detection limit of five percent. Next, we troubleshooted all different steps of our transfection method, in order to solve the base editing issues and enable the generation of the splicing factor hotspot mutations.

CRISPR base editors not functional after mRNA transfection

We hypothesized that the absence of the hotspot mutations in the Sanger traces could be explained by different aspects of the protocol, like the base editor sequence, the base editing efficiency itself, or the intracellular translation of the base editor mRNA to a functional protein. First, we checked whether the base editor sequences contained a mutation that interfered with its function. Sanger sequencing of both plasmid and cDNA revealed no mutation in the sequence (data not shown). This indicated that the absence of the hotspot mutations in the Sanger traces was caused or by inefficient base editing or by defects in the intracellular translation of the base editor mRNA.

Next, we assayed the base editing efficiency in several ways. First, we aimed to test whether the percentage of base edited cells perhaps could enrich over time via a proliferative advantage of the splicing factor mutant cells. This would indicate that the editing efficiency is too low to detect with Sanger sequencing, but that the base editing itself is functional. However, even after the culturing limit of four weeks for HSPCs, we did not detect a hotspot mutation (data not shown). Second, we transfected the myeloid cell line Kasumi-1 with different settings to test whether the base editing efficiency is specifically low in HSPCs. Nevertheless, also in Kasumi-1 cells, we did not detect a hotspot mutation with Sanger sequencing (data not shown). This indicates that the absence of hotspot mutations did not depend on the used cell type. Third, we assayed the transfection efficiency with GFP mRNA, to check whether mRNA is transfected inside HSPCs with the used Neon electroporation system and settings. After 24h after electroporation, we saw that almost all HSPCs were GFP positive after electroporation, based on flow cytometry (figure 2B). On top of that, the addition of base editor mRNA did not lead to increased cell death compared to electroporated HSPCs without mRNA (figure 2C). However, the viability reached only ten percent, indicating

that the base editor mRNA is not toxic for the cells, but that the electroporation itself is toxic. Nevertheless, the fact that almost all cells were GFP positive after electroporation and the observation that the base editor did not lead to increased cell death, show that the used transfection method could not explain the absence of the hotspot mutations in the Sanger traces.

Since the base editor sequences did not contain mutations and the absence of the hotspot mutations could not be caused by the transfection efficiency of HSPCs, we hypothesized that the base editor mRNAs were not converted to functional base editor proteins inside the cells. To test this hypothesis, we checked the functionality of the base editor inside HSPCs by making use of a highly efficient control gRNA. As positive control gRNA, we made use of a TP53(Y220C) gRNA that showed high efficiency with the adenine base editor that we used (Geurts et al., 2021). Next to that, we increased the amount of base editor mRNA, as it was shown that there is a positive correlation between the amount of adenine base editor mRNA and the base editing efficiency (Jiang et al., 2020). Both the positive control gRNA and increase in amount of base editor mRNA did not result in functional base editing in HSPCs (data not shown). To rule out that the base editing efficiency of TP53 was below the Sanger sequencing detection limit, we also performed Nutlin-3a selection on the TP53-electroporated cells to enrich the bulk cellular population for TP53 mutants (Kucab et al., 2017). Even after Nutlin-3a selection we did not detect the hotspot mutation with Sanger sequencing. This confirmed our hypothesis that at least the adenine base editor does not become functional inside HSPCs after electroporation. Therefore, we assumed that the base editor mRNA was not translated after electroporation. To test this assumption, we performed a western blot targeting the Cas9 region of the base editor protein after electroporation. Antibody specificity was validated by doxycycline induced Cas9 expressing HEK293T cells as positive and negative control (figure 2D). As expected, we saw a band at 150 kDa in the positive control sample. For the GC-base editor, we also observed a band at 150 kDa, whereas we expected the mass to be around 200 kDa due to the additional domains coupled to the Cas9 protein. This indicates that the GC-base editing proteins do not contain their base editing domains and are therefore not functional. The samples electroporated with ABEmax contained a band around 200 kDa, indicating complete translation. The Cas9 antibody signal was weak, but since the control β-actin signal was also weak, this indicates that the adenine base editor is translated efficiently in HSPCs (figure 2D). Probably, something in the post-translational production of the adenine base editor is disturbed, causing the base editor to be nonfunctional despite complete translation.

Alternative models to modulate splicing

Since we were not able to generate splicing factor hotspot mutations in HSPCs, we developed two different models to deregulate splicing in order to study a causal relationship between deregulated splicing and the development of pAML. In the first model, we aimed to mimic the effect of SF3B1 hotspot mutations via the SF3B1 inhibitor pladienolide B (PladB). PladB binds to SF3B1 and interferes with its binding to the spliceosome (Kotake et al., 2007). This effect is similar to the effect of SF3B1 hotspot mutations, since the hotspot mutations alter the conformation of the SF3B1 protein and therefore alter the interactions with the spliceosome and the pre-mRNA (Canbezdi et al., 2021). Based on literature, we treated the HSPCs with 100nM PladB for 24h, after which we performed the functional read-outs (Kumar et al., 2022). In the second model, we deregulated splicing in an indirect manner via generation of a

DNMT3A knock-out (KO) in HSPCs with conventional CRISPR/Cas9. We generated this second aberrant splicing model, to test the hypothesis that a DNMT3A knock-out has a similar effect as splicing deregulation on HSPCs and to have a different model next to pharmacological inhibition of splicing (Ramabadran et al., 2023). To be able to answer the research question whether there is a causal relationship between deregulated splicing and the development of AML, we made use of different functional read-outs for both aberrant splicing models. As functional read-outs we compared the number of ds-breaks and of R-loops, and the activation and expression of several DDR genes between our aberrant splicing models and healthy HSPCs. These functional read-outs were namely altered in immortalized cell lines upon splicing deregulation and their effects could contribute to the development of AML (sources). Therefore, if we would observe a similar effect of splicing deregulation in healthy HSPCs compared to the immortalized cell lines that would further confirm our hypothesis that there is a causal relationship between splicing deregulation and the development of (p)AML.

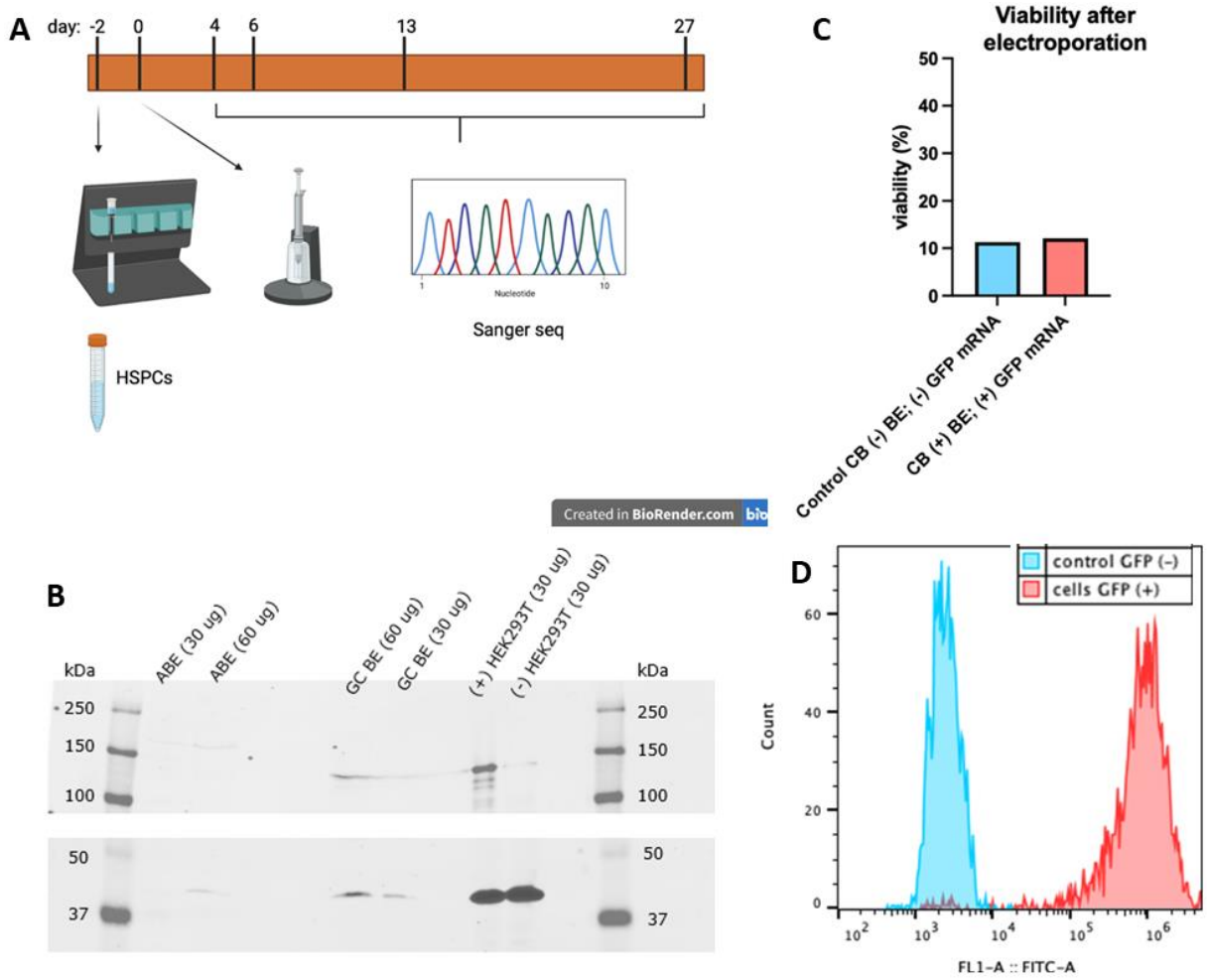


Figure 2: Troubleshooting the ineffective generation of hotspot mutations in HSPCs. **A)** Schematic representation of experimental method. HSPC isolation from CB, followed by electroporation of the CRISPR complex after two days, and Sanger sequencing for detection of the hotspot mutation. **B)** Western Blot staining of Cas9 in 30 μg or 60 μg HSPC lysates, electroporated with the adenine base editor (ABE) or GC base editor (GC BE) mRNA. Dox-induced Cas9-expressing HEK293T lysate used as both positive and negative control. Control protein staining for β-actin. **C)** CB-derived HSPC viability 24h after electroporation measured on cytoflex cell sorter. Blue control CB sample electroporation did not contain mRNA (left bar), red CB sample electroporation contained both BE mRNA and 0,4 μg GFP mRNA (right bar). **D)** Histogram of measured GFP intensity with cytoflex cell sorter. Blue histogram displays control HSPCs, red histogram displays HSPCs electroporated with GFP mRNA.

Functional read-outs of aberrant splicing models in human HSPCs

Increase in ds-breaks upon PladB treatment

As it was shown that PladB treatment resulted in an increase in ds-breaks in other immortalized cell lines, we wondered whether PladB treatment has a similar effect in healthy HSPCs (Nguyen et al., 2017). To study the number of ds-breaks in single cells, we performed immunofluorescence microscopy, in which we counted the number of individual p-γH2AX foci per cell after different treatments, which is specific for a histone modification that occurs as response to ds-breaks (figure 3A). As positive control for the ds-breaks, we irradiated a fraction of the cells. Overall, the HSPCs that were treated with PladB contained significantly more foci per nucleus compared to the control HSPCs, with a mean difference of six foci per nucleus ($p=0.012$, figure 3B). This finding indicates that pharmacologically deregulated splicing in HSPCs leads to an increase in ds-breaks, as expected. We also tested the effect of PladB treatment on irradiated HSPCs, since we hypothesized that more ds-breaks would remain after PladB treatment due to disrupted DNA repair (Pederiva et al., 2016). The mean number of ds-breaks was slightly higher in PladB-treated, irradiated HSPCs compared to the untreated, irradiated HSPCs, but this difference was not significant (figure 3B). This finding indicates that there is not a large difference in DNA repair efficiency directly after irradiation between PladB-treated HSPCs and control HSPCs. Nevertheless, it could still be that there is a difference in DNA repair efficiency between PladB-treated HSPCs and control HSPCs, which only becomes significant after a longer period after irradiation.

Bulk analysis shows decrease in ds-breaks upon PladB treatment and DNMT3A knock-out

Next, we aimed to confirm our immunofluorescence microscopy results in bulk. For bulk analysis, we made use of FACS to quantify the total intracellular p-γH2AX intensity per cell, which is a more approachable method for high-throughput analysis compared to immunofluorescence microscopy. To test the hypothesis that a DNMT3A KO has a similar effect as splicing deregulation on HSPCs, we also analyzed DNMT3A KO HSPCs in bulk for intracellular p-γH2AX intensity (Ramabadran et al., 2023). Contradictory to the microscopy results, the total intracellular p-γH2AX intensity decreased significantly upon PladB treatment in the bulk population ($p<0.00001$, figure 3C, supplementary table S3). The same effect was observed in DNMT3A KO HSPCs compared to the untreated HSPCs ($p<0.00001$, figure 3C). In addition to the single aberrant splicing models, we also aimed to test the combinatorial effect of both aberrant splicing models by treating DNMT3A KO HSPCs with PladB. The total intracellular p-γH2AX intensity of the PladB-treated DNMT3A KO HSPCs showed a significant, but minor increase compared to untreated DNMT3A KO HSPCs ($p=0.028$, figure 3D, supplementary table S3). These findings indicate that splicing deregulation via PladB treatment or via a DNMT3A KO decreases the intracellular p-γH2AX intensity compared to healthy HSPCs. However, when both aberrant splicing models are combined, the splicing deregulation seems to increase the intracellular p-γH2AX intensity compared to healthy HSPCs. These outcomes did not change by altering the gating of the viable cell fraction in the FACS analysis, indicating a true decrease in p-γH2AX intensity in the HSPCs of our aberrant splicing models compared to healthy HSPCs. Next to the effect of splicing deregulation on intracellular p-γH2AX intensity, we also aimed to further test the hypothesis that the repair of ds-breaks is disturbed upon splicing deregulation (Pederiva et al., 2016). Since we did not detect a significant effect of splicing deregulation on DNA repair directly after irradiation with microscopy, we increased the culturing period after irradiation to test whether the difference in DNA repair becomes visible over time. We pretreated HSPCs one day with PladB, irradiated

both the pretreated- and untreated HSPCs, and continued the PladB treatment of pretreated HSPCs for two days until FACS analysis. We hypothesized that the ds-breaks in the untreated, irradiated HSPCs should mostly be repaired after culturing, but the ds-breaks in the PladB-treated HSPCs would not be repaired. Indeed, we observed that the total p-γH2AX intensity in untreated, irradiated HSPCs was significantly higher but comparable to the signal of untreated HSPCs, indicating effective DNA repair in the irradiated condition (p=0.000032, figure 3E, supplementary table S3). However, we were unable to test the hypothesis that DNA repair is restricted upon splicing deregulation with our approach, as almost all cells died upon a combination of PladB treatment and irradiation. In conclusion, the microscopy data indicate an increase in ds-breaks upon splicing deregulation in HSPCs, whereas the FACS data indicate a decrease in ds-breaks in bulk both upon PladB treatment and in DNMT3A KO HSPCs. We consider the immunofluorescence microscopy results as more reliable, because with microscopy separate p-γH2AX foci are visible and quantified, whereas in FACS the effect of splicing deregulation is studied by quantifying the total p-γH2AX intensity of both background signal and ds-breaks. Therefore, it could be that the background p-γH2AX intensity measured in FACS interferes with the signal that derives from the ds-breaks, resulting in the observed differences in outcomes between FACS and microscopy.

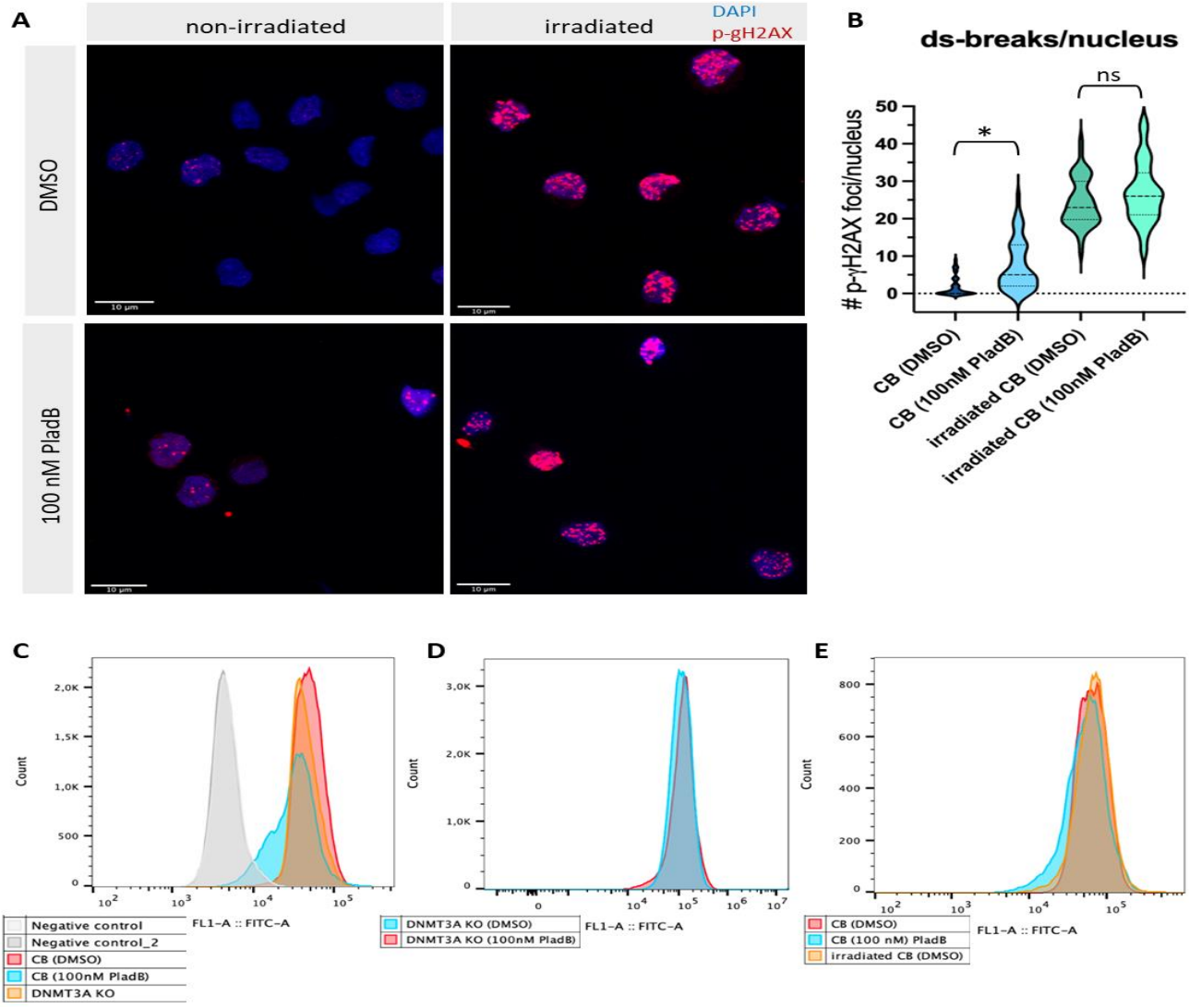


Figure 3: Number of ds-breaks in HSPCs upon PladB-treatment and DNMT3A knock-out. A) Confocal microscopy images of HSPCs, stained for ds-breaks with p-γH2AX (red) and DAPI staining (blue) indicating the nuclei of non-irradiated cells (left panels) and irradiated cells (right panels). Upper panels treated with DMSO, lower panels treated with 100nM PladB for 24h. Scale-bars 10 μm. **B)** Widefield microscopy vulcano-plot of quantified number p-γH2AX foci/nucleus for each condition. N=50 for each condition. Statistical analysis Kruskall-Wallis test. *p<0,05. ns = non-significant. **C,D,E)** Cell sorting analysis histograms of intracellular p-γH2AX staining. **C)** Comparison of unstained HSPCs in duplo (negative controls) (grey); HSPCs (DMSO) (red); HSPCs (100nM PladB) (blue); DNMT3A KO HSPCs (orange). **D)** Comparison of DNMT3A KO HSPCs (DMSO) (blue) with DNMT3A KO HSPCs (100nM PladB) (red). **E)** Comparison of HSPCs (DMSO) (red); HSPCs (100nM PladB) (blue); two days priorly irradiated HSPCs (orange).

Decrease in R-loops upon splicing deregulation

It was hypothesized that the increase in ds-breaks upon splicing deregulation was induced by an increase in R-loops (Singh et al., 2020). Therefore, we decided to test the difference in number of R-loops in HSPCs upon PladB treatment and DNMT3A KO, to explain the increase in ds-breaks that was observed with microscopy. First, we aimed to see whether there was colocalization between R-loops and ds-breaks via confocal immunofluorescence microscopy. However, the brightness the nonspecific antibody binding to cytoplasmic RNA concealed the nuclear signal, making it impossible to detect colocalization between R-loops and ds-breaks (supplementary figure S2). Others have also encountered this problem when using the S9.6 antibody and tried to solve nonspecific binding with RNases (Smolka et al., 2021). Despite the RNase treatment, the cytoplasmic S9.6 intensity was still bright in their data. Therefore, instead of immunofluorescence microscopy, we performed a dot-blot with isolated DNA and stained the blot for DNA:RNA hybrids, which was proven to be a suitable method for specific detection of R-loops in bulk (Ramirez et al., 2021). We measured fluorescence intensity from the R-loop staining for each tested DNA concentration upon both PladB treatment and in DNMT3A KO HSPCs (figure 4A,B). These measurements indicate that the overall number of R-loops in HSPC nuclei decreases upon splicing deregulation. Next, we tested the effect of PladB treatment in DNMT3A KO HSPCs. Here, we also observed a decrease in R-loop intensity upon PladB treatment (figure 4C,D). In conclusion, we saw a decrease in bulk R-loop intensity in both PladB-treated- and DNMT3A KO HSPCs compared to untreated HSPCs, indicating a decrease in R-loops upon splicing deregulation. This decrease correlated with the decrease in intracellular p-γH2AX intensity observed in bulk cell sorting, but not with the results observed with microscopy. Contradictory to our hypothesis, these findings suggest that the ds-breaks in our aberrant splicing models are not induced by R-loops, and that R-loops probably contribute to the development of AML in a different way than via the induction of genomic instability.

Activation of the DNA damage response in HSPCs not detected upon splicing deregulation

It was shown that the ds-breaks observed upon deregulated splicing specifically activates the ATR pathway, but despite the activation of the ATR pathway the ds-breaks were not repaired (Nguyen et al., 2018). Therefore, it was hypothesized that splicing deregulation causes downstream inhibition of the ATR pathway, which would explain how ds-breaks can accumulate in splicing deregulated cells. To test this hypothesis, we performed a Western Blot that was stained for both activated ATM protein and activated ATR protein, to check the activation of both homology dependent repair pathways upon splicing deregulation.

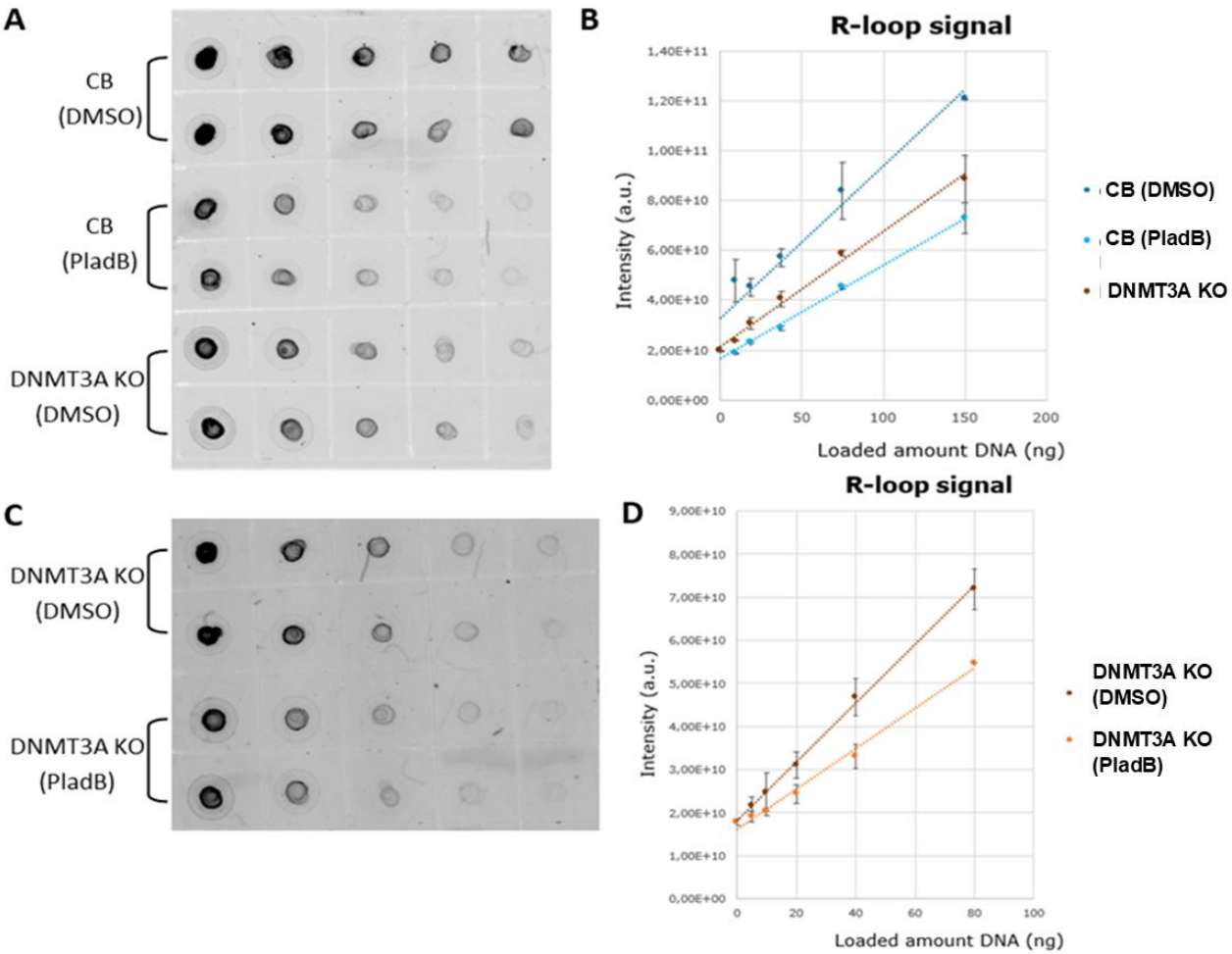


Figure 4: Decreased number of R-loops in PladB-treated- and DNMT3A KO HSPCs. A) Image of dot-blot stained for R-loops. Experimental conditions depicted on left. A DNA dilution series was pipetted *in duplo* on blot from high to low amount of DNA. **B)** Graphical representation of measured mean R-loop intensity from dots in figure A. Conditions depicted in the legend. SE error bars. **C)** Image of dot-blot stained for R-loops for conditions DNMT3A KO +/- PladB. **D)** Graphical representation of measured mean R-loop intensity from dots in figure C. Conditions depicted in the legend. SE error bars.

We expected that only the activated ATR protein would give a signal, based on the increase in ds-breaks upon PladB treatment that was observed with fluorescence microscopy, and the observation that ATM was not activated in splicing deregulated immortalized cell lines (Dvinge et al., 2016b; Singh et al., 2020). Besides that, we also stained the blot for activated P53 to check for downstream activation of the ATR pathway. Since P53 is a target of many pathways, it could be that the activated P53 signal on the Western Blot did not derive from the ATR pathway. Nevertheless, we hypothesized that if deregulated splicing interferes with the downstream part of the ATR pathway, the activated P53 signal in our aberrant splicing models would be lower compared to the untreated- or irradiated HSPC samples then. Next to the DDR proteins, we also stained the Western Blot for DNMT3A, to confirm the knock-out efficiency in addition to the Sanger sequencing results. Unexpectedly, we did not detect any signal besides the control protein β-actin, even not in the irradiated samples that were used as positive control for the activation of the ATR pathway (supplementary figure S3A). The presence of β-actin indicates that the protein transfer was successful, meaning that either the protein concentration was below the detection limit, or the ATR pathway was not activated in

our model systems. Therefore, we increased the amount of loaded protein lysate of the negative control HSPCs, PladB-treated HSPCs and DNMT3A KO HSPC lysates. Again, we did not detect any signal except from the β-actin signal (figure 5A). Next, we tested the presence of total ATR protein on the blot to check whether the ATR protein expression in HSPCs was below the detection limit. The ATR antibody gave a dim signal in each sample (supplementary figure S3B). The dim signal confirms low expression of ATR in HSPCs, meaning that the phosphorylated fraction of ATR could fall below the detection threshold of our methods. Therefore, we cannot objectively draw conclusions about whether deregulated splicing interferes with the downstream activation of the ATR pathway. Replacing the used antibodies with more sensitive antibodies should enable us to study the effect of deregulated splicing on the ATR pathway.

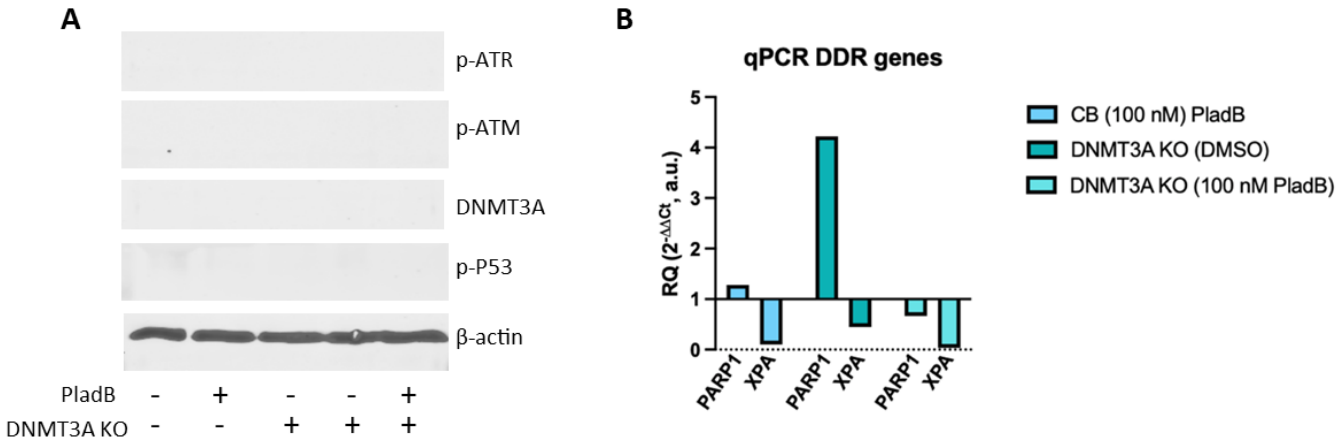


Figure 5: No detectable DDR activation upon PladB-treatment or DNMT3A knock-out. A) Image of Western Blot stained for presence of p-ATR, p-ATM, DNMT3A, p-P53 and β-actin. HSPC treatment conditions are depicted below. **B)** Bar plot presenting relative quantification (RQ) values of PARP1- and XPA expression of samples depicted in the legend compared to expression levels in untreated HSPCs. RQ = 1 means same expression level as in untreated HSPCs.

XPA expression may be decreased via nonsense-mediated decay upon splicing deregulation

A hypothesis about how deregulated splicing can interfere with the ATR pathway and why DNA damage is not repaired, is that splicing deregulation causes downregulation of DDR gene expression via nonsense-mediated decay (NMD) (Han et al., 2022). Therefore, we measured with qPCR the difference in gene expression of three DDR genes RAD51, PARP1 and XPA upon PladB treatment and in DNMT3A KO HSPCs, to study the effect of deregulated splicing on different DDR pathways. We expected a downregulation of RAD51 in both PladB-treated- and DNMT3A KO HSPCs, based on the decreased expression observed in T-ALL cells upon SF3B1 inhibition (Han et al., 2022). Unfortunately, we failed to detect the RAD51 mRNA with our method. For each treatment condition we observed a decrease in XPA expression compared to untreated HSPCs (figure 5B). Since the decrease in XPA expression was observed in each aberrant splicing model, it seems that XPA is a target of NMD induced by deregulated splicing. Contradictory to the XPA results, the expression of PARP1 was inconsistent between the treated conditions. We observed a 28 percent increase in PARP1 expression in the PladB-treated condition and a fourfold increase in PARP1 expression in the untreated DNMT3A KO HSPCs (figure 5B). However, we observed a 33 percent decrease in PARP1 expression in the PladB-treated DNMT3A KO HSPCs (figure 5B). We were unable to explain this difference between the different treatment conditions. Due to the limited cell numbers that we were

able to sample from primary HSPC, we were only able to perform one qPCR measurement for each condition. As a result, we were not able to perform any statistics on this data or confirm the decrease in XPA expression and the contradictory PARP1 results in replicative experiments. Therefore, the qPCR measurements need to be repeated with other HSPC cultures to confirm these outcomes and determine the influence of deregulated splicing on different DDR pathways. Besides that, the addition of a NMD inhibitor in these future studies should confirm that the altered expression is due to NMD and not for example due to the experimental conditions.

Discussion

In this project we aimed to investigate the effect of deregulated splicing on DNA damage and disrupted DNA repair in healthy HSPCs, to see whether deregulated splicing can contribute to the development of AML. Whereas most studies used leukemic cell lines to study the effects of splicing deregulation, we made use of HSPCs because AML probably originates from HSPCs (Roboz & Guzman, 2009). In this study, we deregulated splicing in HSPCs both pharmacologically with the SF3B1 inhibitor PladB, and indirectly via the generation of a DNMT3A KO. We observed an increase in ds-breaks upon PladB treatment, but a decrease in R-loops in both PladB-treated and DNMT3A KO HSPCs. These findings indicate that splicing deregulation could contribute to the development of AML, but in a different way than is proposed in literature (Chen et al., 2018; Cheruiyot et al., 2021; Singh et al., 2020).

Contradictory ds-breaks outcomes

In this project, we quantified the number of ds-breaks per nucleus with immunofluorescence microscopy to study the effect on single cell level and be able to see the nuclear localization. As expected based on literature, we saw a significant increase in ds-breaks upon PladB treatment with microscopy (figure 3B, Lappin et al., 2022, Nguyen et al., 2017). This observation demonstrates that also in a healthy hematopoietic background, splicing deregulation leads to an increase in ds-breaks and is not a side-effect of a leukemic genetic background. In addition, this finding indicates that splicing deregulation can contribute to the development of AML, as genomic instability is one of the hallmarks of cancer. Since our microscopy protocol was low-throughput, we aimed to confirm the increase in ds-breaks in bulk with FACS. Remarkably, FACS showed a decrease in ds-breaks upon PladB treatment, which is probably due to nonspecific binding of the p-γH2AX antibody (figure 3C,E). To determine whether the background staining indeed influenced the FACS outcomes, we could make use of a p-γH2AX antibody gradient in a follow-up experiment, and optimize our FACS staining protocol to be able to quantify the p-γH2AX foci in bulk. Nevertheless, it is also recommended to test whether the FACS detectors are sensitive enough to detect a difference of six p-γH2AX foci per nucleus on average (figure 2B). Therefore, it would be good to also optimize our immunofluorescence microscopy method in a high-throughput manner, in order to validate the FACS outcomes. After establishing a protocol to quantify the number of p-γH2AX foci per nucleus in bulk, additional testing using our aberrant splicing models and positive controls, such as DNA damaging agents, should confirm our microscopy findings that splicing deregulation leads to genomic instability.

Expression and activation of DNA damage response genes in HSPCs below detection limit

It was hypothesized that deregulated splicing can contribute to the development of pAML due to downstream interference with DNA repair, which results in the persistence of ds-breaks and therefore in an increase in mutations (Nguyen et al., 2018). Based on the increase in p-γH2AX foci in PladB-treated HSPCs and irradiated HSPCs observed with microscopy (figure 3B), we expected activation of the ATR pathway in these conditions, but decreased activation of P53 in our aberrant splicing models (Chen et al., 2018; Nguyen et al., 2018). Unfortunately, the Western Blot showed no signal besides the control protein β-actin (figure 5A, supplementary figure S3A). Because of the absence of signal in the positive control sample with irradiated HSPCs, we are not able to draw conclusions about the repair of ds-breaks in our aberrant

splicing models. Two possible explanations for the absence of p-ATR signal are that blocking with ELK interfered with detection of phosphorylated proteins, or the protein concentration was below the detection limit of the used antibodies. Since ELK blocking was proven to be successful in the detection of p-ATR and p-P53 with the same antibodies that we used in this study, we hypothesize (Christodoulou et al., 2022). Further optimization with a higher amount of loaded protein lysate and perhaps also BSA blocking should solve this issue. A way how deregulated splicing could interfere with DNA repair is via NMD of DDR genes (Han et al., 2022). Since the detection levels of the tested DDR genes were comparable to background noise in this study, further testing and optimization of the qPCR protocol is necessary to be able to confirm NMD of DDR genes induced by deregulated splicing. Eventually, comprehension of the influence of deregulated splicing on DNA repair and the DNA damage response can lead to the development of new targeted therapies against (p)AML.

R-loop level may be dependent on transcriptional activity of cells

Next to interference with the repair of ds-breaks, it was hypothesized that splicing deregulation also induces ds-breaks via an accumulation of R-loops (Singh et al., 2020). Remarkably, we observed a decrease in bulk R-loop level in our aberrant splicing models (figure 4A-D). A possible explanation for the decrease in overall R-loop level is that splicing deregulation reduces transcriptional activity in these cells (Sveen et al., 2016). If there are fewer sites of active transcription, also the number of R-loops drops, since R-loops are formed at sites of transcription. A hypothesis on how splicing deregulation can reduce transcriptional activity, is that splicing inhibition leads to RNA polymerase pausing after the promotor and interferes with transcription elongation (Castillo-Guzman et al., n.d.). However, further research is necessary to confirm this explanation, since there are also multiple studies that showed an increase in R-loops upon splicing deregulation (Cheruiyot et al., 2021; Lappin et al., 2022; Singh et al., 2020). The major difference between these studies and our study is that these studies were performed in tumor-derived cell lines or $CD34^+$ cells that derived from MDS patients, whereas we made use of cord blood-derived HSPCs. The difference in used cell line could be responsible for the contradictory results, as the overall transcriptional activity in tumor cells is higher compared to healthy cells (Bywater et al., 2013). Because of the difference in transcriptional activity, it may be that splicing deregulation has a different effect on R-loop accumulation in a cancerous background compared to a healthy background. Besides that, it could also be that the decrease in transcriptional activity upon PladB treatment is a drug-mediated effect that does not occur upon splicing factor hotspot mutations, as we used a relatively high PladB dose of 100 nM. Nevertheless, the decrease in R-loops implicates that the increase in ds-breaks observed with microscopy cannot be caused by stress induced by an accumulation of R-loops. This implicates that the increase ds-breaks observed here in a healthy background are caused by another factor. Further research on the exact mechanisms is required to understand how splicing deregulation leads to the emergence of ds-breaks.

Altered transcriptional activity may contribute to the development of pAML

Now that we discussed the main effects of PladB treatment and DNMT3A KO in healthy human HSPCs, we still do not have a clear answer to the question whether splicing deregulation can cause the development of pAML. Nevertheless, we do have a hypothesis about the working mechanism based on the results in this project. Since ds-breaks can induce somatic mutations, we hypothesize that splicing deregulation results in an increase in ds-breaks and therefore also in an accumulation of somatic mutations (Martincorena & Campbell, 2015). These

somatic mutations could interfere with specific cellular pathways and eventually transform the cells to a malignant state. If the DNA repair is indeed restricted in splicing deregulated cells, that would explain why the accumulation of somatic mutations can take place. However, the overall mutational burden in pAML is low compared to other tumor types, which implies that the induction of ds-breaks cannot be the only mechanism via which splicing deregulation causes tumor development (Gröbner et al., 2018; Ma et al., 2018). It is known that R-loops are normally present in the genome and involved in transcription regulation (Lee et al., 2023). We hypothesize that the decrease in R-loops observed in this study results in an altered gene expression pattern. This altered gene expression pattern can give the cells a proliferative advantage and decrease for example the expression of genes involved in myeloid differentiation, which stimulates the cells to become blasts. Together, the altered gene expression pattern induced by R-loops and accumulation of somatic mutations by ds-breaks could contribute to the development of leukemia and explain why splicing deregulation is often observed in pAML.

In vitro translation could resolve base editing issues

A first step towards testing our hypothesis about the role of splicing deregulation in the development of pAML, is to generate adult splicing factors hotspot mutations in HSPCs. It was shown that different splicing factor hotspot mutations have different effects on the conformation of the splicing factor and resulted in different splicing patterns, even though the same gene was mutated (Canbezdi et al., 2021). Therefore, the effect of PladB-binding to the SF3B1 could differ from hotspot mutations, because PladB-binding can alter the conformation of SF3B1 in a different way than hotspot mutations. The disadvantage of the splicing factor hotspot mutations is that they are not detected in pAML, so it could be that the consequences of altered splicing are different in children compared to adults. Nevertheless, we think that hotspot mutations more closely represent *in vivo* splicing deregulation than splicing factor inhibition with a drug. In this study we were not able to generate the adult splicing factor hotspot mutations by electroporating the base editor mRNA. We assume that the absence of hotspot mutations was due to incomplete or incorrect translation and/or protein folding of the base editor, and therefore, we propose an alternative strategy. A way to circumvent this issue is to translate the base editor mRNA *in vitro* and electroporate the base editor protein. *In vitro* production of an adenine base editor in bacteria has already been proven successful for base editing of human HSPCs (Martin-Rufino et al., 2023). Once genetic models in a wildtype background are created, similar functional assays as mentioned in this study could be executed and optimized, to confirm our findings about the direct effects of splicing deregulation in HSPCs. Besides that, it would be interesting to compare the number of R-loops and the transcriptional states of the genetic models with pAML splicing factor mutant cells, in order to test our hypothesis about the function of R-loops in the development of pAML. In addition, it would also be interesting to perform single cell whole genome sequencing with these genetic models, in order to find other processes that are potentially disrupted upon splicing deregulation and could possibly contribute to the development of pAML. These studies can eventually lead to the development of new targeted therapies with fewer side effects and also help to prevent the development of pAML in the end.

Acknowledgements

I would like to thank **Ruben van Boxtel**, for providing me an internship position in his lab. Because of this internship, I developed many new lab skills and got experience with a whole new research field. Next, I would especially like to thank **Lucca Derks** for her great supervision during my internship and her help with my experiments. Thanks to her way of supervision, I was able to develop myself from an internship student to a more independent researcher and therefore prepared me for a future PhD position. Besides that, I am grateful for her empathy for my personal circumstances during this internship. I also want to thank all **Van Boxtel group members** not only for their help or advise during my internship, but mostly for the pleasant ambience within the group, nice chats, support, and lots of laughter. Because of them I really enjoyed going to work. Besides that, I would like to thank **Jens Bunt** (Kool group, PMC) for his advice regarding the dot-blot, **Ravian van Ineveld** (imaging facility, PMC) for his help with the imaging settings, and **Maroussia Ganpat** (Drost group, PMC) for providing me Cas9-expressing control lysates for my Cas9-troubleshooting Western Blot. Lastly, I want to thank **Hugo Snippert** for being my second examiner.

Materials and methods

Cell culture

Human $CD34^+$-HSPCs derived from cord blood samples were cultured in StemSpan SFEM media (SCT, #09650), containing 100 ng/mL SCF (Miltenyi, #130-096-696), 100 ng/mL Flt3-ligand (Miltenyi, #130-096-480), 20 ng/mL IL-6 (SCT, #78050), 10 ng/mL IL-3 (SCT, #78040), 50 ng/mL TPO (Miltenyi, #130-095-754), 100 μg/mL Primocin (InvivoGen, #ant-pm-1), 500nM UM729 (SCT, #72332) and 750 nM SR1 (SCT, #72344). HSPCs were cultured in culture dishes in a density of 1E5/mL at 37 °C/ 5% CO_2. Culture media was refreshed three times a week. Kasumi-1 cells were cultured in RPMI 1640 media (ThermoFisher, #A1049101) supplemented with 10% 1x FBS at 37 °C/ 5% CO_2. The Kasumi-1 cells were cultured in T25 falcon flasks at a density of 5E5 cells/mL. The media was refreshed twice a week.

Magnetic cell separation (MACS) of $CD34^+$-HSPCs

Human $CD34^+$-HSPCs were isolated using a CD34 microbeat kit (Miltenyi, #130-100-453) according to manufacturer's protocol. In summary, human cord blood vials stored in liquid nitrogen were thawed quickly in a 37 °C water bath. The cells were washed twice with preheated IMDM media (ThermoFisher, #31980030) supplemented with 10% FBS. Afterwards, the cells were washed once with cold MACS buffer (PBS pH 7.2 supplemented with 0.5% BSA and 2mM EDTA). After washing, the cells were stained in a mixture of MACS buffer, FcR blocking reagent (Miltenyi, #130-100-453) and CD34-beads (Miltenyi, #130-100-453) for 30 minutes on ice. The $CD34^+$-cells were isolated on a LS-column (Miltenyi). The $CD34^+$-cell fraction was resuspended in StemSpan SFEM complete culture media and kept at 37 °C/ 5% CO_2.

GC base editor plasmid maxi-prep

Two different GC-base editor plasmids in bacterial stabs were ordered from Addgene based on the results from Koblan et al. (2021) (#163554, #163527, supplementary table S1). For both base editors, one plasmid was amplified overnight in LB media supplemented with ampicillin at 37 °C. The plasmids were isolated from the bacteria, using the NucleoBond Xtra Maxi kit (Macherey-nagel) according to manufacturer's protocol. The plasmids were resuspended in milliQ and stored at -24 °C until further usage. Both plasmids were sequence verified with Sanger sequencing (Macrogen, primers in supplementary table S2).

CRISPR base editor mRNA generation

For the adenine base editor ABEmax, the pCMV_ABEmax plasmid (Addgene, #124163, supplementary table S1) was linearized by SapI enzyme (Thermofisher, #ER1931) in Tango buffer (Thermofisher). The linearized plasmid was purified using QIAquickPCR purification kit (Qiagen) according to manufacturer's protocol. ABEmax mRNA was produced from the linearized plasmid using the HiScribe T7 ARCA mRNA kit (with tailing) (NEB, #E2060) according to manufacturer's protocol. The mRNA was purified using the RNeasy Mini Kit (Qiagen) according to manufacturer's protocol. The mRNA was divided into single-use aliquots and stored at -80 °C. For the GC-base editor, the pCMV_UdgX-Anc689-NG-nCas9-RBMX base editor (Addgene, #163554, supplementary table S1) was chosen based on the online tool mentioned in the article by Koblan et al. (2021). The plasmid was linearized by PciI enzyme (NEB, #R0655S) in buffer r3.1 (NEB). To produce the GC-BE mRNA the same protocol was

followed as for the ABEmax mRNA. Both base editor mRNAs were sequence verified, via cDNA production with the SuperScript II RT kit (ThermoFisher) according to manufacturer's protocol, followed by Sanger sequencing (Macrogen). In addition, gel electrophoresis was performed to confirm that no side products were present in the mRNA sample (data not shown).

Human cord blood-derived HSPC electroporation

The isolated $CD34^+$-cells were cultured for two days before electroporation. The cells were washed once with 1x PBS and resuspended in Neon electroporation buffer T (Invitrogen). For each electroporation, 1E5-2E5 cells were resuspended in 10 μL buffer T. The crRNAs (IDT, supplementary table S2) and tracrRNA (IDT, #1072533) were dissolved in nuclease-free duplex buffer (IDT). sgRNAs were produced by mixing dissolved crRNAs and tracrRNA in equal amounts, followed by heating at 95 °C for 5 minutes and cooling down to room temperature. In case of base editing, 3 μg of base editor mRNA was mixed with 3.2 μg sgRNA, in a total volume of approximately 2 μL. For the generation of a DNMT3A knock-out, in which we used the conventional CRISPR/Cas9 system, 0.9 μL of both 100 μM sgRNAs was added to 0.3 μL 10 mg/mL S.p. Cas9 nuclease V3 (IDT, #1081058). The base editing or CRISPR/Cas9 mixture was pipetted to 10 μL cell suspension per electroporation. As negative control for each electroporation, one electroporation was performed with scrambled sgRNA (IDT, #1072544). For electroporation, the Neon Transfection System (ThermoFisher) was used according to manufacturer's protocol. The HSPCs were electroporated with the HSPC-optimized settings 1600V; 10 ms; 3 pulses (Gundry et al., 2016). After electroporation, the cells were transferred to 1 mL preheated StemSpan SFEM complete media without antibiotics. After 24h, 100 μg/mL Primocin (InvivoGen, #ant-pm-1) was added to the media.

Kasumi-1 cell electroporation and optimization

Approximately 2E5 Kasumi-1 cells were resuspended in 10 μL Neon electroporation buffer R (Invitrogen). The electroporation procedure was the same as for electroporation of HSPCs, but with adjusted electroporation settings: 1350V, 20 ms, 1 pulse. The electroporated Kasumi-1 cells were pipetted in preheated RPMI with 10% FBS media. After 5h, 100 U/mL Penicillin/Streptomycin was added to the media. For the optimization of the electroporation settings, the ABEmax sgRNA mixture was replaced with nuclease-free duplex buffer (ThermoFisher), containing 0, 0.5 or 2 μg pMax-GFP plasmid (NEB). For each amount of plasmid three different electroporation settings were tested: 1350V, 35 ms, 1 pulse; 1650V, 20 ms, 1 pulse; and 1800V, 15 ms, 1 pulse. After 48h, the electroporated Kasumi-1 cells were analyzed on a Cytoflex S flow cytometer for GFP expression. Data was analyzed using FlowJo (v10.8.2). Cells were first gated for viable cells, followed by single cell gating and GFP-positive gating (representative gating in supplementary figure S1A).

Sanger sequencing genome editing

Approximately 50.000 cells were pelleted and resolved in 30 μL DirectPCR lysis buffer supplemented with 0.3 μL proteinase K (Qiagen, #19131). The cell suspension was incubated overnight at 60 °C, 250 rpm. Proteins were inactivated by heating the suspension at 96 °C for 20 minutes. The suspension was centrifuged at 12,000 rpm for 5 minutes, and the supernatant containing the DNA was transferred to a new tube. The supernatant was used as template in PCR with the 2x G2 green GoTaq Hotstart master mix (Promega) and primers covering the genome editing region (IDT, supplementary table S2). The PCR product was purified using the

QIAquickPCR purification kit (Qiagen), according to manufacturer's protocol. The purified PCR product was sent for Sanger sequencing (Macrogen).

Imaging

Round glass coverslips were covered with poly-D-lysine coating. Approximately 2E6 HSPCs were irradiated at a cell density of around 3E5 cells/mL with 2.014 Gy, to induce ds-breaks just before putting the HSPCs on the coverslips. Approximately 5E5 HSPCs were pipetted on each coverslip, and after 2h of incubation at RT the liquid was removed. The cells were fixated with 4% PFA and permeabilized with 0.1% Triton X-100 in PBS. The permeabilized cells were blocked with 5% ELK solution in PBST for 1h at RT. Primary antibodies P-Histone H2A.X (S139) (Cell Signaling Technologies, #9718T, 1:500) and S9.6 (Kerafast, #ENH001, 1:100) were diluted in 0.5% ELK in PBST, pipetted on the coverslips and incubated overnight at 4 °C. The secondary antibodies anti-rabbit IgG AF594 (Invitrogen, #BMS306F1-100, 1:500), anti-mouse IgG AF488 (Invitrogen, #2220848, 1:500) and DAPI (1:2000) were dissolved in 0.5% ELK in PBST were added to the coverslips and incubated for 1h at RT. The coverslips were attached to microscopy slides by using Fluoromount mounting media. The microscopy slides were imaged on both a Leica DMi8 widefield microscope with a 40x oil objective, and a Leica SP8 live confocal microscope with a 63x oil objective. Imaging settings can be found in supplementary table S4.

Quantification p-γH2AX foci widefield microscopy

Widefield LIF-files were loaded in FIJI (v2.0.0-rc-69/1.52i) and the z-stacks were conferred to maximum-projections. The number of p-γH2AX foci per nucleus was quantified with the FIJI plugin ComDet v.0.5.5, by selecting individual nuclei. Plugin particle size was 2 pixels and intensity threshold 19 a.u. Statistical analysis was performed by using a Kruskal-Wallis test in Prism (v9.5.1 (528)).

FACS intracellular p-γH2AX staining

Approximately 2E5 HSPCs per condition were fixated with 4% PFA for 20 minutes on ice. The cells were permeabilized in milliQ supplemented with 0.5% saponin, 0.5% BSA, 10 mM HEPES, 140 mM NaCl, 2.5 mM $CaCl_2$. Phospo-γH2AX FITC JBW301 antibody (Merck, #16-202A) was added in a 1:200 dilution and incubated for 1h on ice. The cells were washed and resuspended in 100 μL FACS buffer and fluorescent intensity was measured on the Cytoflex S flow cytometer. The same laser settings were used for each intensity measurement. The FACS data was analyzed with FlowJo (v10.8.2). First, the FACS plot was gated for viable cell fraction, followed by single-cell gating. The fluorescent intensity of the gated cell fraction was depicted in the final histogram (representative gating: supplementary figure S1B). Statistical analysis was performed on the FlowJo data, using a student's t-test and a Chi-Squared test in FlowJo (v10.8.2) (supplementary table S3).

Dot-blot for R-loop quantification

For each experimental condition, approximately 1.5E6 cells were harvested for bulk R-loops quantification. The DNA was extracted using a phenol/chloroform extraction according to the protocol in Ramirez et al. (2021). The DNA was resolved in 12 μL TE buffer. The DNA concentration was measured using the Qubit dsDNA BR Assay kit (ThermoFisher) with a Qubit 2.0 Fluorometer according to manufacturer's protocol. To compare the number of R-loops between different conditions, a two-fold DNA dilution series was made with TE buffer. The dot-blot was performed *in duplo* according to the protocol in Ramirez et al. (2021), with the

primary antibody S9.6 (Kerafast, #ENH001) and the secondary antibody anti-mouse IgG IRDye 680RB (LI-COR, #925-68070). The antibody intensity was measured on an Odyssey device. The R-loop intensities were quantified in FIJI (v2.0.0-rc-69/1.52i), by multiplying the mean intensity and area of each dot. Standard deviation was calculated based on duplicate results in Microsoft Excel.

Western Blot

Protein lysates in SDS with DTT buffer and the Precision PageRuler Prestained Protein ladder (ThermoFisher, #26616) were run on an 8% acrylamide gel. The proteins were blotted on a Trans-Blot Turbo mini nitrocellulose membrane (Biorad), using a Biorad Trans-Blot Turbo Transfer System with the standard high molecular weight blotting settings. The membrane was blocked with 5% ELK in TBST for 1h at RT. The membrane was stained overnight at 4 °C with primary antibody solution in 0.5% ELK in TBST. Used antibodies in this project: CRISPR/Cas9 [7A9] (Epigentek, #A9000-010, 1:500), DNMT3A (64B1446) (Novus Biologicals, #NB120-13888SS, 1:1000), p-ATR (S428) (Cell Signaling Technologies, #2853T, 1:1000), ATR (Abcam, #ab2905, 1:1000) p-ATM (S1981) (Cell Signaling Technologies, #5883T, 1:1000), p-P53 (S15) (Cell Signaling Technologies, #9286T, 1:1000), β-actin (Merck, #MABT523, 1:5000). The membrane was stained with the corresponding secondary antibody anti-mouse IgG IRDye 680RD (LI-COR, #925-68070, 1:10.000), or anti-rabbit IgG IRDye 800CW (LI-COR, #925-32211, 1:5000) in 0.5% ELK in TBST for 1h at RT. The membrane was imaged with an Odessey device with automated laser intensities. The blots were analyzed with FIJI (v2.0.0-rc-69/1.52i).

qPCR DNA damage response genes

RNA was extracted from the cell pellets that were stored at -80 °C, using the RNeasy Mini Kit (Qiagen) according to manufacturer's protocol. The RNA concentration was measured using both a nanodrop and the Qubit RNA broad range kit (ThermoFisher) with the Qubit 2.0 fluorometer according to manufacturer's protocol. cDNA was produced from the isolated RNA using the SuperScript II RT kit (ThermoFisher) according to manufacturer's protocol. 5 ng cDNA was used in each qPCR reaction with the 2x DyNAmo HS SYBR Green qPCR master mix (Thermo Scientific) and qPCR primers (supplementary table S2). The qPCR reaction mixtures were pipetted in a 384-wells plate, and qPCR was performed in a Biorad CFX384 Real-Time System qPCR device. The relative expression was calculated with the formula: $2^{-\Delta\Delta Ct}$.

References

Banaszak, L. G., Giudice, V., Zhao, X., Wu, Z., Gao, S., Hosokawa, K., Keyvanfar, K., Townsley, D. M., Gutierrez-Rodrigues, F., Fernandez Ibanez, M. del P., Kajigaya, S., & Young, N. S. (2018). Abnormal RNA splicing and genomic instability after induction of DNMT3A mutations by CRISPR/Cas9 gene editing. *Blood Cells, Molecules, and Diseases*, *69*. https://doi.org/10.1016/j.bcmd.2017.12.002

Bolouri, H., Farrar, J. E., Triche, T., Ries, R. E., Lim, E. L., Alonzo, T. A., Ma, Y., Moore, R., Mungall, A. J., Marra, M. A., Zhang, J., Ma, X., Liu, Y., Liu, Y., Auvil, J. M. G., Davidsen, T. M., Gesuwan, P., Hermida, L. C., Salhia, B., … Meshinchi, S. (2018). The molecular landscape of pediatric acute myeloid leukemia reveals recurrent structural alterations and age-specific mutational interactions. *Nature Medicine*, *24*(1). https://doi.org/10.1038/nm.4439

Bywater, M. J., Pearson, R. B., McArthur, G. A., & Hannan, R. D. (2013). Dysregulation of the basal RNA polymerase transcription apparatus in cancer. *Nature reviews. Cancer*, *13*(5), 299–314. https://doi.org/10.1038/nrc3496

Cacace, F., Iula, R., De Novellis, D., Caprioli, V., D'Amico, M. R., De Simone, G., Cuccurullo, R., Wierda, W. G., Mahadeo, K. M., Menna, G., & Tambaro, F. P. (2022). High-Risk Acute Myeloid Leukemia: A Pediatric Prospective. *Biomedicines*, *10*(6). https://doi.org/10.3390/biomedicines10061405

Canbezdi, C., Tarin, M., Houy, A., Bellanger, D., Popova, T., Stern, M.-H., Roman-Roman, S., & Alsafadi, S. (2021). Functional and conformational impact of cancer-associated SF3B1 mutations depends on the position and the charge of amino acid substitution. *Computational and Structural Biotechnology Journal*, *19*, 1361–1370. https://doi.org/10.1016/j.csbj.2021.02.012

Castillo-Guzman, D., Hartono, S. R., Sanz, L. A., & Chédin, F. (n.d.). *SF3B1-targeted Splicing Inhibition Triggers Global Alterations in Transcriptional Dynamics and R-Loop Metabolism*. https://doi.org/10.1101/2020.06.08.130583

Chen, L., Chen, J.-Y., Huang, Y.-J., Gu, Y., Qiu, J., Qian, H., Shao, C., Zhang, X., Hu, J., Li, H., He, S., Zhou, Y., Abdel-Wahab, O., Zhang, D.-E., & Fu, X.-D. (2018). The Augmented R-Loop Is a Unifying Mechanism for Myelodysplastic Syndromes Induced by High-Risk Splicing Factor Mutations. *Molecular Cell*, *69*(3). https://doi.org/10.1016/j.molcel.2017.12.029

Cheruiyot, A., Li, S., Nonavinkere Srivatsan, S., Ahmed, T., Chen, Y., Lemacon, D. S., Li, Y., Yang, Z., Wadugu, B. A., Warner, W. A., Pruett-Miller, S. M., Obeng, E. A., Link, D. C., He, D., Xiao, F., Wang, X., Bailis, J. M., Walter, M. J., & You, Z. (2021). Nonsense-Mediated RNA Decay Is a Unique Vulnerability of Cancer Cells Harboring *SF3B1* or *U2AF1* Mutations. *Cancer Research*, *81*(17). https://doi.org/10.1158/0008-5472.CAN-20-4016

Christodoulou, P., Boutsikos, P., Neophytou, C. M., Kyriakou, T.-C., Christodoulou, M.-I., Papageorgis, P., Stephanou, A., & Patrikios, I. (2022). Amygdalin as a chemoprotective agent in co-treatment with cisplatin. *Frontiers in Pharmacology*, *13*. https://doi.org/10.3389/fphar.2022.1013692

Crossley, M. P., Bocek, M., & Cimprich, K. A. (2019). R-Loops as Cellular Regulators and Genomic Threats. *Molecular Cell*, *73*(3). https://doi.org/10.1016/j.molcel.2019.01.024

Devilli, L., Garonzi, C., Balter, R., Bonetti, E., Chinello, M., Zaccaron, A., Vitale, V., De Bortoli, M., Caddeo, G., Baretta, V., Tridello, G., & Cesaro, S. (2021). Long-Term and Quality of Survival in Patients Treated for Acute Lymphoblastic Leukemia during the Pediatric Age. *Hematology Reports*, *13*(2), 8847. https://doi.org/10.4081/hr.2021.8847

Dvinge, H., Kim, E., Abdel-Wahab, O., & Bradley, R. K. (2016a). RNA splicing factors as oncoproteins and tumour suppressors. *Nature Reviews Cancer*, *16*(7). https://doi.org/10.1038/nrc.2016.51

Dvinge, H., Kim, E., Abdel-Wahab, O., & Bradley, R. K. (2016b). RNA splicing factors as oncoproteins and tumour suppressors. *Nature Reviews Cancer*, *16*(7), 413–430. https://doi.org/10.1038/nrc.2016.51

Geurts, M. H., de Poel, E., Pleguezuelos-Manzano, C., Oka, R., Carrillo, L., Andersson-Rolf, A., Boretto, M., Brunsveld, J. E., van Boxtel, R., Beekman, J. M., & Clevers, H. (2021). Evaluating CRISPR-based prime editing for cancer modeling and CFTR repair in organoids. *Life Science Alliance*, *4*(10), e202000940. https://doi.org/10.26508/lsa.202000940

Gröbner, S. N., Worst, B. C., Weischenfeldt, J., Buchhalter, I., Kleinheinz, K., Rudneva, V. A., Johann, P. D., Balasubramanian, G. P., Segura-Wang, M., Brabetz, S., Bender, S., Hutter, B., Sturm, D., Pfaff, E., Hübschmann, D., Zipprich, G., Heinold, M., Eils, J., Lawerenz, C., … Pfister, S. M. (2018). The landscape of genomic alterations across childhood cancers. *Nature*, *555*(7696), 321–327. https://doi.org/10.1038/nature25480

Gundry, M. C., Brunetti, L., Lin, A., Mayle, A. E., Kitano, A., Wagner, D., Hsu, J. I., Hoegenauer, K. A., Rooney, C. M., Goodell, M. A., & Nakada, D. (2016). Highly Efficient Genome Editing of Murine and Human Hematopoietic Progenitor Cells by CRISPR/Cas9. *Cell Reports*, *17*(5), 1453–1461. https://doi.org/10.1016/j.celrep.2016.09.092

Han, C., Khodadadi-Jamayran, A., Lorch, A. H., Jin, Q., Serafin, V., Zhu, P., Politanska, Y., Sun, L., Gutierrez-Diaz, B. T., Pryzhkova, M. V., Abdala-Valencia, H., Bartom, E. T., Buldini, B., Basso, G., Velu, S. E., Sarma, K., Mattamana, B. B., Cho, B.-K., Obeng, R. C., … Ntziachristos, P. (2022). SF3B1 homeostasis is critical for survival and therapeutic response in T cell leukemia. *Science Advances*, *8*(3). https://doi.org/10.1126/sciadv.abj8357

Hershberger, C. E., Moyer, D. C., Adema, V., Kerr, C. M., Walter, W., Hutter, S., Meggendorfer, M., Baer, C., Kern, W., Nadarajah, N., Twardziok, S., Sekeres, M. A., Haferlach, C., Haferlach, T., Maciejewski, J. P., & Padgett, R. A. (2021). Complex landscape of alternative splicing in myeloid neoplasms. *Leukemia*, *35*(4). https://doi.org/10.1038/s41375-020-1002-y

Hofmann, I. (2015). Pediatric myelodysplastic syndromes. *Journal of Hematopathology*, *8*(3). https://doi.org/10.1007/s12308-015-0253-4

Huber, S., Haferlach, T., Meggendorfer, M., Hutter, S., Hoermann, G., Baer, C., Kern, W., & Haferlach, C. (2022). SF3B1 mutations in AML are strongly associated with MECOM rearrangements and may be indicative of an MDS pre-phase. *Leukemia*, *36*(12), 2927–2930. https://doi.org/10.1038/s41375-022-01734-7

Jiang, T., Henderson, J. M., Coote, K., Cheng, Y., Valley, H. C., Zhang, X.-O., Wang, Q., Rhym, L. H., Cao, Y., Newby, G. A., Bihler, H., Mense, M., Weng, Z., Anderson, D. G., McCaffrey, A. P., Liu, D. R., & Xue, W. (2020). Chemical modifications of adenine base editor mRNA and guide RNA expand its application scope. *Nature Communications*, *11*(1). https://doi.org/10.1038/s41467-020-15892-8

Koblan, L. W., Arbab, M., Shen, M. W., Hussmann, J. A., Anzalone, A. V., Doman, J. L., Newby, G. A., Yang, D., Mok, B., Replogle, J. M., Xu, A., Sisley, T. A., Weissman, J. S., Adamson, B., & Liu, D. R. (2021). Efficient C•G-to-G•C base editors developed using CRISPRi screens, target-library analysis, and machine learning. *Nature Biotechnology*, *39*(11), 1414–1425. https://doi.org/10.1038/s41587-021-00938-z

Kotake, Y., Sagane, K., Owa, T., Mimori-Kiyosue, Y., Shimizu, H., Uesugi, M., Ishihama, Y., Iwata, M., & Mizui, Y. (2007). Splicing factor SF3b as a target of the antitumor natural product pladienolide. *Nature Chemical Biology*, *3*(9), 570–575. https://doi.org/10.1038/nchembio.2007.16

Kucab, J. E., Hollstein, M., Arlt, V. M., & Phillips, D. H. (2017). Nutlin-3a selects for cells harbouring *TP 53* mutations. *International Journal of Cancer*, *140*(4), 877–887. https://doi.org/10.1002/ijc.30504

Kumar, D., Kashyap, M. K., Yu, Z., Spaanderman, I., Villa, R., Kipps, T. J., La Clair, J. J., Burkart, M. D., & Castro, J. E. (2022). Modulation of RNA splicing associated with Wnt signaling pathway using FD-895 and pladienolide B. *Aging*, *14*(5), 2081–2100. https://doi.org/10.18632/aging.203924

Lappin, K. M., Barros, E. M., Jhujh, S. S., Irwin, G. W., McMillan, H., Liberante, F. G., Latimer, C., La Bonte, M. J., Mills, K. I., Harkin, D. P., Stewart, G. S., & Savage, K. I. (2022). Cancer-Associated SF3B1 Mutations Confer a BRCA-Like Cellular Phenotype and Synthetic Lethality to PARP Inhibitors. *Cancer Research*, *82*(5). https://doi.org/10.1158/0008-5472.CAN-21-1843

Lattanzi, A., Meneghini, V., Pavani, G., Amor, F., Ramadier, S., Felix, T., Antoniani, C., Masson, C., Alibeu, O., Lee, C., Porteus, M. H., Bao, G., Amendola, M., Mavilio, F., & Miccio, A. (2019). Optimization of CRISPR/Cas9 Delivery to Human Hematopoietic Stem and Progenitor Cells for Therapeutic Genomic Rearrangements. *Molecular Therapy*, *27*(1), 137–150. https://doi.org/10.1016/j.ymthe.2018.10.008

Lee, S.-Y., Miller, K. M., & Kim, J.-J. (2023). Clinical and Mechanistic Implications of R-Loops in Human Leukemias. *International Journal of Molecular Sciences*, *24*(6). https://doi.org/10.3390/ijms24065966

Liu, Z., Yoshimi, A., Wang, J., Cho, H., Chun-Wei Lee, S., Ki, M., Bitner, L., Chu, T., Shah, H., Liu, B., Mato, A. R., Ruvolo, P., Fabbri, G., Pasqualucci, L., Abdel-Wahab, O., & Rabadan, R. (2020). Mutations in the RNA Splicing Factor SF3B1 Promote Tumorigenesis through MYC Stabilization. *Cancer Discovery*, *10*(6). https://doi.org/10.1158/2159-8290.CD-19-1330

Ma, X., Liu, Y., Liu, Y., Alexandrov, L. B., Edmonson, M. N., Gawad, C., Zhou, X., Li, Y., Rusch, M. C., Easton, J., Huether, R., Gonzalez-Pena, V., Wilkinson, M. R., Hermida, L. C., Davis, S., Sioson, E., Pounds, S., Cao, X., Ries, R. E., … Zhang, J. (2018). Pan-cancer genome and transcriptome analyses of 1,699 paediatric leukaemias and solid tumours. *Nature*, *555*(7696), 371–376. https://doi.org/10.1038/nature25795

Madhusoodhan, P. P., Carroll, W. L., & Bhatla, T. (2016). Progress and Prospects in Pediatric Leukemia. *Current Problems in Pediatric and Adolescent Health Care*, *46*(7). https://doi.org/10.1016/j.cppeds.2016.04.003

Martincorena, I., & Campbell, P. J. (2015). Somatic mutation in cancer and normal cells. *Science*, *349*(6255), 1483–1489. https://doi.org/10.1126/science.aab4082

Martin-Rufino, J. D., Castano, N., Pang, M., Grody, E. I., Joubran, S., Caulier, A., Wahlster, L., Li, T., Qiu, X., Riera-Escandell, A. M., Newby, G. A., Al'Khafaji, A., Chaudhary, S., Black, S., Weng, C., Munson, G., Liu, D. R., Wlodarski, M. W., Sims, K., … Sankaran, V. G. (2023). Massively parallel base editing to map variant effects in human hematopoiesis. *Cell*, *186*(11), 2456-2474.e24. https://doi.org/10.1016/j.cell.2023.03.035

Metayer, C., Dahl, G., Wiemels, J., & Miller, M. (2016). Childhood Leukemia: A Preventable Disease. *Pediatrics*, *138*(Supplement_1). https://doi.org/10.1542/peds.2015-4268H

Mupo, A., Seiler, M., Sathiaseelan, V., Pance, A., Yang, Y., Agrawal, A. A., Iorio, F., Bautista, R., Pacharne, S., Tzelepis, K., Manes, N., Wright, P., Papaemmanuil, E., Kent, D. G., Campbell, P. C., Buonamici, S., Bolli, N., & Vassiliou, G. S. (2017). Hemopoietic-specific Sf3b1-K700E knock-in mice display the splicing defect seen in human MDS but develop anemia without ring sideroblasts. *Leukemia*, *31*(3), 720–727. https://doi.org/10.1038/leu.2016.251

Namayandeh, S. M., Khazaei, Z., Lari Najafi, M., Goodarzi, E., & Moslem, A. (2020). GLOBAL Leukemia in Children 0-14 Statistics 2018, Incidence and Mortality and Human Development Index (HDI): GLOBOCAN Sources and Methods. *Asian Pacific Journal of Cancer Prevention*, *21*(5), 1487–1494. https://doi.org/10.31557/APJCP.2020.21.5.1487

Nguyen, H. D., Leong, W. Y., Li, W., Reddy, P. N. G., Sullivan, J. D., Walter, M. J., Zou, L., & Graubert, T. A. (2018). Spliceosome Mutations Induce R Loop-Associated Sensitivity to ATR Inhibition in Myelodysplastic Syndromes. *Cancer Research*, *78*(18). https://doi.org/10.1158/0008-5472.CAN-17-3970

Nguyen, H. D., Yadav, T., Giri, S., Saez, B., Graubert, T. A., & Zou, L. (2017). Functions of Replication Protein A as a Sensor of R Loops and a Regulator of RNaseH1. *Molecular Cell*, *65*(5), 832-847.e4. https://doi.org/10.1016/j.molcel.2017.01.029

Park, D. J., Kwon, A., Cho, B.-S., Kim, H.-J., Hwang, K.-A., Kim, M., & Kim, Y. (2020). Characteristics of *DNMT3A* mutations in acute myeloid leukemia. *BLOOD RESEARCH*, *55*(1). https://doi.org/10.5045/br.2020.55.1.17

Pederiva, C., Böhm, S., Julner, A., & Farnebo, M. (2016). Splicing controls the ubiquitin response during DNA double-strand break repair. *Cell Death & Differentiation*, *23*(10), 1648–1657. https://doi.org/10.1038/cdd.2016.58

Pimenta, D. B., Varela, V. A., Datoguia, T. S., Caraciolo, V. B., Lopes, G. H., & Pereira, W. O. (2021). The Bone Marrow Microenvironment Mechanisms in Acute Myeloid Leukemia. *Frontiers in Cell and Developmental Biology*, *9*. https://doi.org/10.3389/fcell.2021.764698

Ramabadran, R., Wang, J. H., Reyes, J. M., Guzman, A. G., Gupta, S., Rosas, C., Brunetti, L., Gundry, M. C., Tovy, A., Long, H., Gu, T., Cullen, S. M., Tyagi, S., Rux, D., Kim, J. J., Kornblau, S. M., Kyba, M., Stossi, F., Rau, R. E., … Goodell, M. A. (2023). DNMT3A-coordinated splicing governs the stem state switch towards differentiation in embryonic and haematopoietic stem cells. *Nature Cell Biology*, *25*(4). https://doi.org/10.1038/s41556-023-01109-9

Ramirez, P., Crouch, R. J., Cheung, V. G., & Grunseich, C. (2021). R-Loop Analysis by Dot-Blot. *Journal of Visualized Experiments*, *167*. https://doi.org/10.3791/62069

Reinhardt, D., Antoniou, E., & Waack, K. (2022). Pediatric Acute Myeloid Leukemia—Past, Present, and Future. *Journal of Clinical Medicine*, *11*(3). https://doi.org/10.3390/jcm11030504

Roboz, G. J., & Guzman, M. (2009). Acute myeloid leukemia stem cells: seek and destroy. *Expert Review of Hematology*, *2*(6), 663–672. https://doi.org/10.1586/ehm.09.53

Rosendahl Huber, A., van Leeuwen, A. J. C. N., Peci, F., de Kanter, J. K., Bertrums, E. J. M., & van Boxtel, R. (2022). Whole-genome sequencing and mutational analysis of human cord-blood derived stem and progenitor cells. *STAR Protocols*, *3*(2). https://doi.org/10.1016/j.xpro.2022.101361

Santos-Pereira, J. M., & Aguilera, A. (2015). R loops: new modulators of genome dynamics and function. *Nature Reviews Genetics*, *16*(10). https://doi.org/10.1038/nrg3961

Seiler, M., Peng, S., Agrawal, A. A., Palacino, J., Teng, T., Zhu, P., Smith, P. G., Buonamici, S., Yu, L., Caesar-Johnson, S. J., Demchok, J. A., Felau, I., Kasapi, M., Ferguson, M. L., Hutter, C. M., Sofia, H. J., Tarnuzzer, R., Wang, Z., Yang, L., … Mariamidze, A. (2018). Somatic Mutational Landscape of Splicing Factor Genes and Their Functional Consequences across 33 Cancer Types. *Cell Reports*, *23*(1), 282-296.e4. https://doi.org/10.1016/j.celrep.2018.01.088

Siegner, S. M., Ugalde, L., Clemens, A., Garcia-Garcia, L., Bueren, J. A., Rio, P., Karasu, M. E., & Corn, J. E. (2022). Adenine base editing efficiently restores the function of Fanconi anemia hematopoietic stem and progenitor cells. *Nature Communications*, *13*(1). https://doi.org/10.1038/s41467-022-34479-z

Singh, S., Ahmed, D., Dolatshad, H., Tatwavedi, D., Schulze, U., Sanchi, A., Ryley, S., Dhir, A., Carpenter, L., Watt, S. M., Roberts, D. J., Abdel-Aal, A. M., Sayed, S. K., Mohamed, S. A., Schuh, A., Vyas, P., Killick, S., Kotini, A. G., Papapetrou, E. P., … Boultwood, J. (2020). SF3B1 mutations induce R-loop accumulation and DNA damage in MDS and leukemia cells with therapeutic implications. *Leukemia*, *34*(9). https://doi.org/10.1038/s41375-020-0753-9

Smolka, J. A., Sanz, L. A., Hartono, S. R., & Chédin, F. (2021). Recognition of RNA by the S9.6 antibody creates pervasive artifacts when imaging RNA:DNA hybrids. *Journal of Cell Biology*, *220*(6). https://doi.org/10.1083/jcb.202004079

Sveen, A., Kilpinen, S., Ruusulehto, A., Lothe, R. A., & Skotheim, R. I. (2016). Aberrant RNA splicing in cancer; expression changes and driver mutations of splicing factor genes. *Oncogene*, *35*(19). https://doi.org/10.1038/onc.2015.318

Taylor, J., & Lee, S. C. (2019). Mutations in spliceosome genes and therapeutic opportunities in myeloid malignancies. *Genes, Chromosomes and Cancer*, *58*(12). https://doi.org/10.1002/gcc.22784

van der Werf, I., Mondala, P. K., Steel, S. K., Balaian, L., Ladel, L., Mason, C. N., Diep, R. H., Pham, J., Cloos, J., Kaspers, G. J. L., Chan, W. C., Mark, A., La Clair, J. J., Wentworth, P., Fisch, K. M., Crews, L. A., Whisenant, T. C., Burkart, M. D., Donohoe, M. E., & Jamieson, C. H. M. (2023). Detection and targeting of splicing deregulation in pediatric acute myeloid leukemia stem cells. *Cell Reports Medicine*, *4*(3). https://doi.org/10.1016/j.xcrm.2023.100962

Yoshida, K., Sanada, M., Shiraishi, Y., Nowak, D., Nagata, Y., Yamamoto, R., Sato, Y., Sato-Otsubo, A., Kon, A., Nagasaki, M., Chalkidis, G., Suzuki, Y., Shiosaka, M., Kawahata, R., Yamaguchi, T., Otsu, M., Obara, N., Sakata-Yanagimoto, M., Ishiyama, K., … Ogawa, S. (2011a). Frequent pathway mutations of splicing machinery in myelodysplasia. *Nature*, *478*(7367), 64–69. https://doi.org/10.1038/nature10496

Yoshida, K., Sanada, M., Shiraishi, Y., Nowak, D., Nagata, Y., Yamamoto, R., Sato, Y., Sato-Otsubo, A., Kon, A., Nagasaki, M., Chalkidis, G., Suzuki, Y., Shiosaka, M., Kawahata, R., Yamaguchi, T., Otsu, M., Obara, N., Sakata-Yanagimoto, M., Ishiyama, K., … Ogawa, S. (2011b). Frequent pathway mutations of splicing machinery in myelodysplasia. *Nature*, *478*(7367), 64–69. https://doi.org/10.1038/nature10496

Zhang, X., Pan, X., Pan, Y., & Wang, Y. (2023). Effects of preventive care on psychological state and complications in leukemia patients receiving chemotherapy. *American Journal of Translational Research*, *15*(1).

Zhang, Y., Qian, J., Gu, C., & Yang, Y. (2021). Alternative splicing and cancer: a systematic review. *Signal Transduction and Targeted Therapy*, *6*(1), 78. https://doi.org/10.1038/s41392-021-00486-7

Supplementary

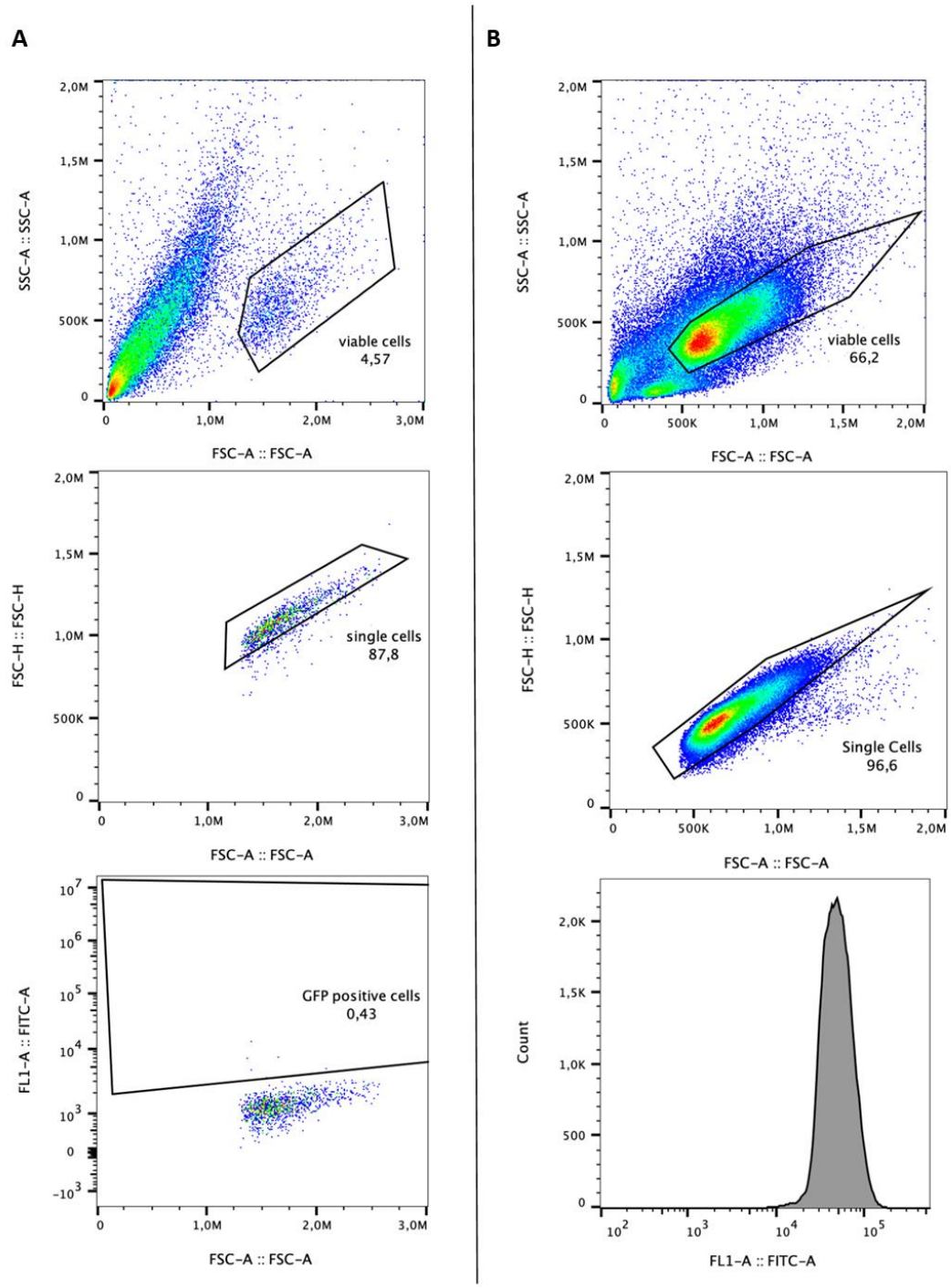
Layman summary

All different kinds of blood cells in the body are formed from blood stem cells in the bone marrow. In the case of leukemia, the transformation from a blood stem cell into a mature blood cell is disturbed. Therefore, the formed blood cells will not become functional, but do occupy space in the blood system. This leads to less space for healthy blood cells, resulting for example in anemia, fatigue, and infections. Normally in cancer, an accumulation of DNA mutations disturbs different cellular processes, resulting in a transformation of healthy cells into malignant cells. However, children have not encountered a lot of DNA damaging agents yet. Besides that, especially in a subtype of leukemia called acute myeloid leukemia (AML), the mutation load is much lower than in other cancer types. This indicates that probably another process in the cell is involved in the development of AML than solely the accumulation of mutations.

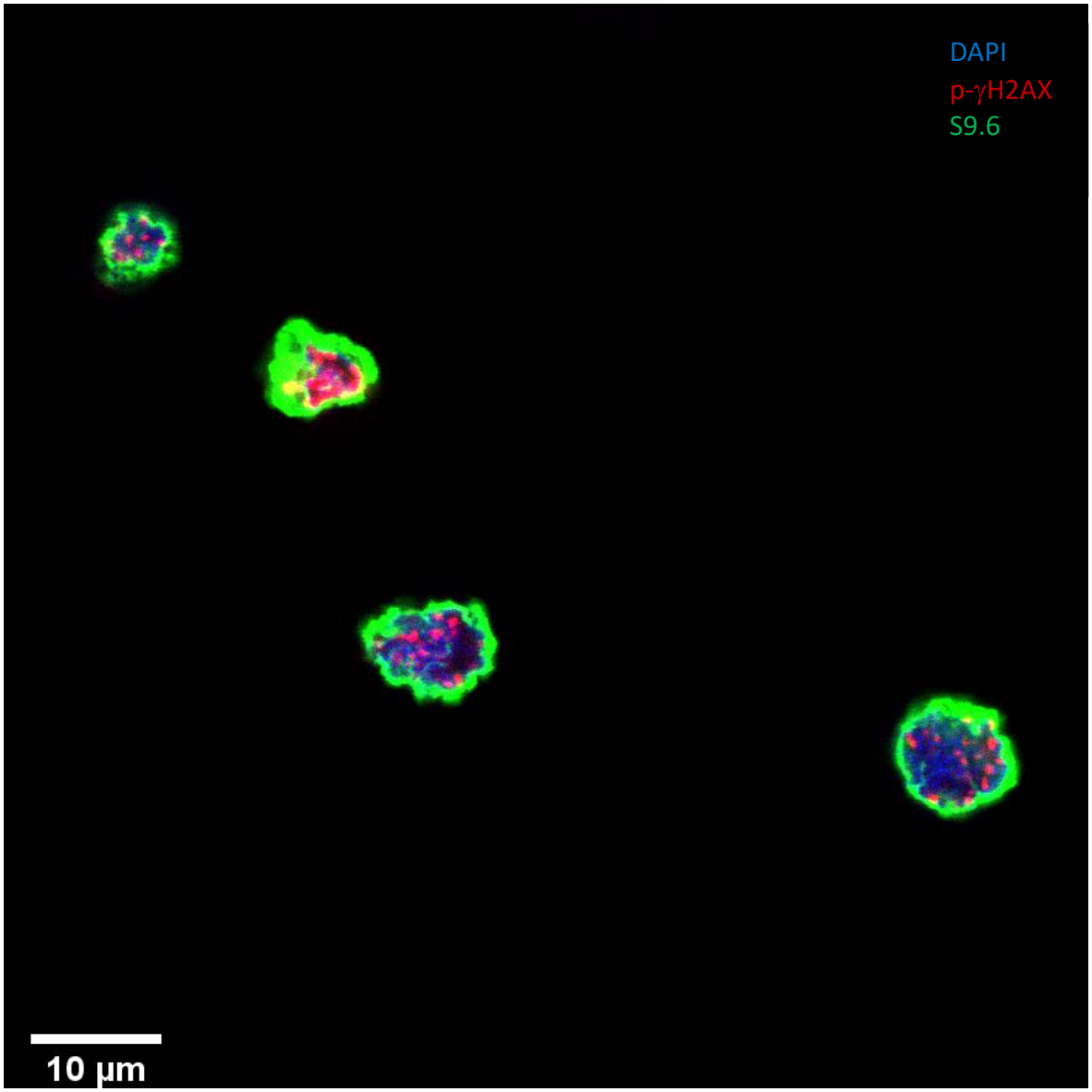
A cellular process that is suspected to contribute to the emergence of AML is splicing. Splicing enables the production of a larger number of different proteins from the same piece of DNA and takes place during transcription. In transcription, a pre-mRNA is produced from a gene. The pre-mRNA consists of alternating protein-coding parts, called exons, and sequences that are not part of the final protein, called introns. During transcription, multiple different splicing factors assemble to the pre-mRNA to form a protein complex called the spliceosome. The spliceosome removes the introns and couples the exons together to form a functional protein-coding mRNA. When a splicing factor is defective, the splicing of a pre-mRNA cannot take place properly anymore, resulting in for example retained introns or skipped exons in the mRNA. Because of the altered sequence of the mRNA, the protein sequence is also different, and therefore these proteins often lose their functionality or are degraded.

In both adult and pediatric AML, splicing is frequently disturbed. Other studies have shown that disturbed splicing leads to decreased expression of genes that are involved in the repair of DNA damage. On top of that, it was found in different cell lineages that disturbed splicing directly led to an increase in DNA damage. Therefore, it is hypothesized that mutations in splicing factors leads to an accumulation of DNA damage, and because of dysfunctional DNA repair proteins this DNA damage is not repaired. Whereas most studies used leukemic cell lineages, we used human cord blood stem- and progenitor cells (HSPCs) in this study to test the effects of splicing deregulation, because it is assumed that the development of AML starts in those cells. In addition, HSPCs are healthy cells, which enables us to study the effect of splicing deregulation more directly compared to leukemic cell lineages. Our findings show an increase in DNA damage in HSPCs upon pharmacological deregulation of splicing. Besides that, our findings indicate that splicing deregulation also could lead to altered transcriptional activity. This finding forms a relatively new explanation how splicing deregulation could contribute to the development of AML. Nevertheless, more studies are needed to directly study the effect of splicing deregulation on transcriptional activity.

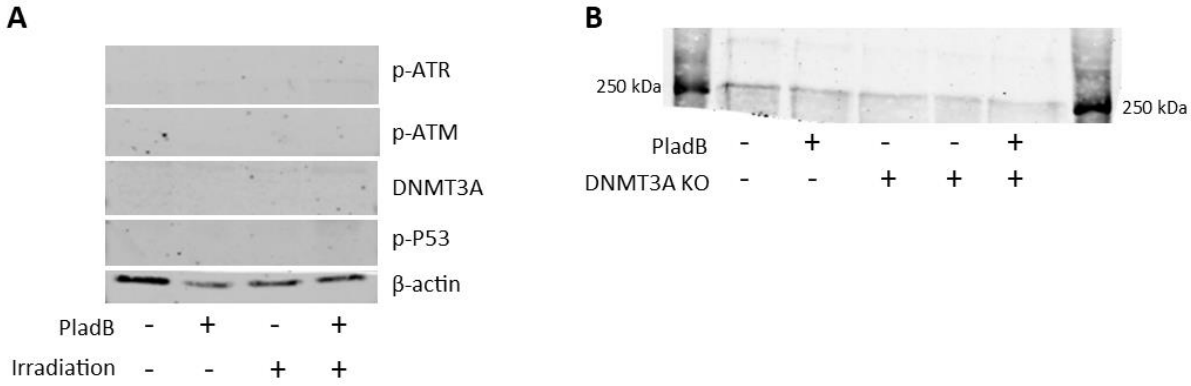
Supplementary Figures



Supplementary figure S1: FACS gating examples. A) Representative gating in FlowJo of cytoflex sorted Kasumi-1 cells 48h after electroporation with GFP plasmid. Selected panel top figure is viable cell fraction. Middle figure represents single cell selection. Selected cells in bottom figure are GFP-positive. Numbers represent selected percentage of cells of all cells in that plot. **B)** Representative gating in FlowJo of cytoflex sorted human cord blood-derived HSPCs that are stained with p-γH2AX FITC JBW301 antibody. Top panel selects viable cell fraction and middle panel selects single cell fraction of viable cells. Numbers represent selected percentage of cells of all cells in that plot. Bottom figure represents histogram of single cells selected in middle panel, with p-γH2AX FITC intensity on x-axis and cell count on y-axis.



Supplementary figure S2: S9.6 antibody binds nonspecific in cytoplasm. Representative image of human cord blood derived HSPCs analyzed with a confocal microscope. Nuclei are stained with DAPI, cells are stained for ds-breaks with p-γH2AX antibody (red) and for R-loops with S9.6 antibody (green). S9.6 antibody localizes to the cytoplasm of all cells, resulting in an overexposed green cytoplasm. Scale bar is 10 µm.



Supplementary figure S3: Western Blot DDR activation. A) Western Blot stained for p-ATR, p-ATM, DNMT3A, p-P53 and β-actin. 20 μg of total protein lysate loaded in each well. Different treatment conditions mentioned on bottom. **B)** Western Blot stained for total ATR (active and inactive). 60 μg of total protein lysate loaded in each well. Different treatment conditions mentioned on the bottom.

Supplementary tables

Supplementary table S1: Used plasmids. Table presenting the used plasmids in this study. The column 'Plasmid name' contains the name of the genetic base editing construct in the plasmid. The column 'Specifications' contains the catalog number of the plasmid, the genetic backbone in which the base editor sequence is cloned, and the antibiotics resistance gene. The column 'Aim in study' contains a description about the base editor type and which hotspot mutations it was used for in this study. *Plasmid ordered and isolated from bacteria as back-up, but not used further.

Plasmid name	Specifications	Aim in study
pCMV_ABEmax	# 124163; pCMV backbone; ampicillin resistance	Adenine base editor, used for U2AF1(Q157R), SF3B1(K700E) and TP53(Y220C)
pCMV_UdgX-Anc689-UdgX-NG-nCas9-RBMX	# 163554; pCMV backbone; ampicillin resistance	GC base editor, used for SF3B1(K666N)
pCMV_Anc689-NG-nCas9*	# 163527; pCMV backbone; ampicillin resistance	Different GC base editor, not used to generate hotspot mutations

Supplementary table S2: Used primers. Table presenting the primers used in this study. Column 'Target' contains the name of the gene or plasmid for which the primers are used. The column 'Type' contains the kind of experiment for which the primer is used, and whether the primer is in forward (Fw) direction or reversed (Rv) direction. The column 'Name' contains the name of the primer on the tube, or how the primer is listed. The column 'Sequence' contains the nucleotide sequence of the primer in 5' to 3' direction. *The sequence of this plasmid is verified, but the plasmid is not used further for cDNA production.

Target	Type	Name	Sequence
CMV promotor	Sequencing, Fw	S26_CMV_Fw	aagtacgccccctattgacg
SF3B1(K666N)	crRNA	SF3B1_K666N	gtacaatcttaataccagtg
	PCR, Fw; Sequencing	L21	accaactcatgactgtcctttct
	PCR, Rv	L22	attggtggatttacctttcctct
SF3B1(K700E)	crRNA	SF3B1_K700E	agcagaaagttcggaccatc
	PCR, Fw	L21	accaactcatgactgtcctttct
	PCR, Rv; Sequencing	L22	attggtggatttacctttcctct
U2AF1(Q157R)	crRNA	U2AF1_Q157R	gccgtcagtatgagatgggg
	PCR, Fw; Sequencing	L103, Fw2	cgccgtgaggaagatgcg
	PCR, Rv	L103, Rv2	aggagacatttactacctcgtgtg
TP53(Y220C)	crRNA	crRNA_TP53_Y220C	cctatgagccgcctgaggtc
	PCR, Fw	SvD031_TP53(Y220C)_Fw	ctgaggtgtagacgccaact
	PCR, Rv	SvD032_TP53(Y220C)_Rv	gacaaccacccttaacccctc
	Sequencing, Fw	SvD033_TP53(Y220C)_seq	cagtactcccctgccctcaac
DNMT3A knock-out (exon 10)	crRNA	crRNA_DNMT3a_KO_1	aggtggccagcagccgcgcg
	crRNA	crRNA_DNMT3a_KO_2	tgacactgccaaggccgtgg
	PCR, Fw; Sequencing	SvD034_DNMT3a_KO_Fw	gggcttgagcagaacggagg
	PCR, Rv	SvD035_DNMT3a_KO_Rv	ggttctagccaaccaacagagagc
pCMV_ABEmax	Sequencing, Fw	SvD030_pBE-general_cDNA_Fw	caccatgaaacggacagccg
		SvD006_pBE-general_seq	ggataagaagcacgagcggc
		SvD008_pBE-general_seq	agagctgcacgccattctgc
		SvD010_pBE-general_seq	agccccgccattaagaaggg
		SvD012_pBE-general_seq	caagaccgagattaccctggcc
		SvD014_pBE-general_seq	acggatcgacctgtctcagc
	Sequencing, Rv	SvD005_pBE-general_seq	gggtacttctcgtggtaggcc
pCMV_UdgX-Anc689-UdgX-NG-nCas9-RBMX (GC base editor)	Sequencing, Fw	SvD030_pBE-general_cDNA_Fw	caccatgaaacggacagccg
		SvD002_pGC-RBMX-B_seq	ctgccgatgtgagaccttct
		SvD004_pGC-RBMX-BE_seq	acacaggccgtgttcggag
		SvD006_pBE-general_seq	ggataagaagcacgagcggc
		SvD008_pBE-general_seq	agagctgcacgccattctgc
		SvD010_pBE-general_seq	agccccgccattaagaaggg
		SvD012_pBE-general_seq	caagaccgagattaccctggcc
		SvD014_pBE-general_seq	acggatcgacctgtctcagc
		SvD016_pGC-RBMX-BE_seq	caccaccaccacgagattatacttaccg
	Sequencing, Rv	SvD005_pBE-general_seq	gggtacttctcgtggtaggcc
pCMV_Anc689-NG-nCas9*	Sequencing, Fw	SvD030_pBE-general_cDNA_Fw	caccatgaaacggacagccg
		SvD026_pGC-BE_seq	tctaccagctgttccatcacatgg

		SvD006_pBE-general_seq	ggataagaagcacgagcggc
		SvD008_pBE-general_seq	agagctgcacgccattctgc
		SvD010_pBE-general_seq	agccccgccattaagaaggg
		SvD012_pBE-general_seq	caagaccgagattaccctggcc
		SvD014_pBE-general_seq	acggatcgacctgtctcagc
cDNA ABEmax	PCR, Fw; Sequencing	SvD030_pBE-general_cDNA_Fw	caccatgaaacggacagccg
		SvD006_pBE-general_seq	ggataagaagcacgagcggc
		SvD008_pBE-general_seq	agagctgcacgccattctgc
		SvD010_pBE-general_seq	agccccgccattaagaaggg
		SvD012_pBE-general_seq	caagaccgagattaccctggcc
	PCR, Rv; Sequencing	SvD005_pBE-general_seq	gggtacttctcgtggtaggcc
		SvD007_pBE-general_seq	cccggttgtccttcaggaatgg
		SvD009_pBE-general_seq	tccaccaccttcactgtctgc
		SvD011_pBE-general_seq	cttccggatctcgccgttgg
		SvD025_pGC-BE_seq	tcggcggttctttttgagcc
cDNA GC base editor	PCR, Fw; Sequencing	SvD030_pBE-general_cDNA_Fw	caccatgaaacggacagccg
		SvD002_pGC-RBMX-B_seq	ctgccgatgtgagaccttct
		SvD004_pGC-RBMX-BE_seq	acacaggccgtgttcggag
		SvD006_pBE-general_seq	ggataagaagcacgagcggc
		SvD008_pBE-general_seq	agagctgcacgccattctgc
		SvD010_pBE-general_seq	agccccgccattaagaaggg
		SvD012_pBE-general_seq	caagaccgagattaccctggcc
		SvD014_pBE-general_seq	acggatcgacctgtctcagc
		SvD016_pGC-RBMX-BE_seq	caccaccaccacgagattatacttaccg
	PCR, Rv; Sequencing	SvD001_pGC-RBMX-BE_seq	cctccagaaggtctcacatcgg
		SvD003_pGC-RBMX-BE_seq	ctcgccgatcatcatgatgcg
		SvD005_pBE-general_seq	gggtacttctcgtggtaggcc
		SvD007_pBE-general_seq	cccggttgtccttcaggaatgg
		SvD009_pBE-general_seq	tccaccaccttcactgtctgc
		SvD011_pBE-general_seq	cttccggatctcgccgttgg
		SvD013_pGC-RBMX-BE_seq	cgctagatcctccagagtcacc
		SvD015_pGC-RBMX-BE_seq	gatcacgaccatatccatctctatcgc
GAPDH	qPCR, Fw	GAPDH Fw	cacatcgctcagacaccatg
	qPCR, Rv	GAPDH Rv	tgacggtgccatggaatttg
RAD51	qPCR, Fw	R12 Fw	caacccatttcacggttagagc
	qPCR, Rv	R12 Rv	ttctttggcgcataggcaaca
PARP1	qPCR, Fw	R3 Fw	agcgtgtttctaggtcgtgg
	qPCR, Rv	R3 Rv	catcaaacatgggcgactgc
XPA	qPCR, Fw	R11 Fw	tgtgataactgcagagatgctg
	qPCR, Rv	R11 Rv	ccccattgtgaatgatgtgga

Supplementary table S3: Significance bulk p-γH2AX data. Table presenting statistics of comparison p-γH2AX signal between different conditions in bulk with FACS. Column 'Comparison' represents the two compared tested conditions. The column 'Figure' represents the figure number in this report in which the conditions are plotted in a histogram. The column 'Mean' represents the mean intensity for both conditions of comparison (first – second) analyzed with FlowJo. The column 'St.dev' represents the standard deviation for both conditions of comparison (first – second) analyzed with FlowJo. The column 'P-value' contains the p-value number for each comparison calculated with a standard student's T-test. The column 'Chi-Squared T(X)' contains the outcomes of a different statistical test for each condition performed with FlowJo. The comparison is significantly different ($p < 0.01$) if $T(X) > 4$.

Comparison	**Figure**	**Mean (a.u.)**	**St.dev (a.u.)**	**P-value**	**Chi-Squared T(X)**
CB (DMSO) – CB (100 nM PladB)	3C	50948 – 36715	23185 – 29050	<0.00001	2859.90
CB (DMSO) – DNMT3A knock-out	3C	50948 – 45699	23185 – 22102	<0.00001	385.18
DNMT3A knock-out (DMSO) – DNMT3A knock-out (100 nM PladB)	3D	137273 – 142331	$4{,}79^{E}5$ – 371716	0.027951	434.10
CB (DMSO) – CB (100nM PladB)	3E	68206 – 60935	34638 – 43673	<0.00001	532.33
CB (DMSO) – irradiated CB (DMSO)	3E	68206 – 71878	34638 – 44987	0.000032	124.90

Supplementary table S4: Imaging settings. Table presenting the used imaging settings for both the widefield- and the confocal microscope. Each column presents a different setting for the microscopy type mentioned in the first column. *Column 'Laser channel' presents light pathway chosen for imaging, which is a specific filter cube for the widefield microscope, or a laser and detector combination based on the specific used fluorophore for the confocal microscope. ** Gain in percentage as a Hyd-detector was used instead of a PMT-detector.

Microscope	**Laser channel***	**Gain (a.u.)**	**Exposure (ms)**	**Laser power (a.u.)**	**Z-stack size (μm)**
Widefield Leica DMi8 Live	Blue	1	75	-	1
	Green	1	5	-	
	Yellow	1	200	-	
Confocal Leica SP8 Live	DAPI	10%**	-	1	0,45
	488	650	-	2	
	594	650	-	1	