

Establishment of a patient-specific iPSC model of stepwise leukemogenesis in severe congenital neutropenia

Dissertation

der Mathematisch-Naturwissenschaftlichen Fakultät
der Eberhard Karls Universität Tübingen
zur Erlangung des Grades eines
Doktors der Naturwissenschaften
(Dr. rer. nat.)

vorgelegt von
Benjamin Dannenmann
aus Nagold

Tübingen

2020

Gedruckt mit Genehmigung der Mathematisch-Naturwissenschaftlichen Fakultät der Eberhard Karls Universität Tübingen.

Tag der mündlichen Qualifikation: 20.10.2020

Stellvertretender Dekan: Prof. Dr. József Fortágh

1. Berichterstatter/in: Prof. Dr. Julia Skokowa

2. Berichterstatter/in: Prof. Dr. Klaus Schulze-Osthoff

DECLARATION

Herewith I declare that this thesis is the result of my independent work. All sources and auxiliary materials used by me in this thesis are cited completely.

Tübingen, 2020

Benjamin Dannenmann

TABLE OF CONTENTS

SUMMARY	6
ZUSAMMENFASSUNG.....	8
LIST OF PUBLICATIONS.....	10
PERSONAL CONTRIBUTION	12
INTRODUCTION.....	13
Pre-leukemia bone marrow failure syndromes and leukemia progression.....	13
Severe congenital neutropenia, <i>ELANE</i> mutations and deregulated intracellular signaling	13
Leukemia progression in CN patients	14
<i>CSF3R</i> mutations and their effects on G-CSFR signaling	15
Missense and nonsense <i>RUNX1</i> mutations and their implications in leukemogenesis	15
Human induced pluripotent stem cell (hiPSC) – a reliable model to study leukemogenesis in CN.....	16
OBJECTIVES OF THE STUDY.....	18
RESULTS AND DISCUSSION	19
Establishment of an in vitro model of CN and CN/AML using EB-based myeloid differentiation of patient-derived hiPSCs (see publications A, B)	19
Increased ER stress and UPR response in CN- and CN/AML-derived HSPCs (CD34 ⁺ CD45 ⁺ cells) (see publication A)	20
Increased DNA damage response in CN-iPSC-derived HSPCs (see publication A).....	21
Elevated DNA damage and diminished granulocytic differentiation of <i>GADD45b</i> deficient HSPCs and iPSCs (see publication C and manuscript E).....	21
p21 expression is upregulated in CN- and CN/AML-derived HSPCs (see publication A)..	22
Missense but not nonsense <i>RUNX1</i> mutations are associated with trisomy 21 in CN/AML patients (see manuscript E).....	24
In vitro patient-derived hiPSC model of stepwise leukemic transformation in <i>ELANE</i> - mutant CN/AML (see manuscript E).....	24

Severely impaired myeloid differentiation and highly elevated proliferation rate of CN/AML-iPSC-derived HSPCs (see manuscript E)	25
AML-related gene expression signature is present in CN/AML blasts and in CN/AML-iPSC derived HSPCs (see manuscript E).....	26
Elevated BAALC expression is a key feature of leukemogenic transformation in CN (see manuscript E).....	28
Leukemic gene expression signature is present in CN/AML-iPSC-derived HSPCs (see manuscript E).....	28
BAALC-dependent leukemogenic signaling pathways in CN/AML (see manuscript E)	30
The selective p38 MAPK inhibitor CMPD1 suppresses proliferation of CN/AML cells (see manuscript E).....	32
OUTLOOK	34
ACKNOWLEDGEMENTS	36
REFERENCES	37

SUMMARY

Severe congenital neutropenia (CN) is a pre-leukemia bone marrow failure syndrome with profoundly diminished terminal granulocytic differentiation of hematopoietic stem cells (HSCs). Life-long treatment with granulocyte-colony stimulating factor (G-CSF) increases the number of neutrophils. About 20 % of CN patients develop AML or MDS. The mechanism of leukemia transformation in CN patients is mostly unclear. Limited numbers of available primary HSC of pediatric CN patients and a lack of animal models hinder the research attempts on the CN and CN/AML pathophysiology. In my PhD Thesis, I aimed to overcome these obstacles by implementing advanced techniques, including patient-derived induced pluripotent cells (iPSCs) in combination with CRISPR/Cas9 gene-editing, for the establishment of the *in vitro* model of congenital neutropenia and of step-wise leukemia development. Using this model, I was able to (1) reproduce “maturation arrest” of granulopoiesis seen in the bone marrow of CN patients; (2) to generate sufficient numbers of hematopoietic stem and progenitor cells (HSPCs) for the analysis of myeloid differentiation, multi-omics analysis and functional validation of identified targets. The ultimate aim was to identify novel molecules that may be further explored for their therapeutic potential for CN/AML.

Our group recently reported an extremely high frequency of cooperating acquired *CSF3R* (colony-stimulating factor 3 receptor) and *RUNX1* (runt-related transcription factor 1) mutations in CN patients who developed acute myeloid leukemia (CN/AML). I established an *in vitro* model of stepwise leukemogenesis in CN/AML through CRISPR/Cas9 mediated gene-editing of iPSCs from two CN patients with acquired *CSF3R* and *RUNX1* mutations and overt AML. Using this model, I identified *BAALC* (brain and acute leukemia, cytoplasmic) upregulation as a key leukemogenic event in CN. Importantly, CRISPR/Cas9-mediated knockout of *BAALC* in CN/AML-iPSCs derived hematopoietic cells of CN/AML patients

restored defective myeloid differentiation to the levels observed in healthy donor hematopoietic stem and progenitor cells (HSPCs). Using transcriptomics data analysis, I found that CMPD1, an inhibitor of p38-mediated MK2a phosphorylation, re-establishes gene expression signature similar to that of BAALC knockout. Intriguingly, *in vitro* treatment of primary blasts of CN/AML patients with CMPD1 resulted in a marked reduction of cell proliferation without affecting differentiation of healthy donor HSPCs. In summary, these observations suggest that targeting of BAALC in hematopoietic cells of CN patients may prevent leukemogenic transformation or eliminate AML blasts in CN/AML individuals.

ZUSAMMENFASSUNG

Schwere kongenitale Neutropenie (CN) ist ein prä-leukämisches Knochenmarksinsuffizienz-Syndrom mit stark verminderter Differenzierung von hämatopoetischen Stammzellen zu Granulozyten. Eine lebenslange Behandlung mit Granulozyten-Kolonie-stimulierendem Faktor (G-CSF) steigert die Anzahl an Neutrophilen deutlich. Ungefähr 20 % der Patienten mit kongenitaler Neutropenie entwickeln später eine akute myeloische Leukämie (AML) oder ein Myelodysplastisches Syndrom (MDS). Der Mechanismus der leukämischen Transformation in Patienten mit kongenitaler Neutropenie und das Fehlen von Tiermodellen limitieren die Erforschung der kongenitalen Neutropenie und der daraus resultierenden AML (CN/AML). Ziel meiner Doktorarbeit war es diese Probleme durch die Entwicklung eines patientenspezifischen induzierten-pluripotenten Stammzell (iPSZ) - Modells in Kombination mit CRISPR/Cas9 Genom-Editierung zu überwinden. Dieses Modell sollte kongenitale Neutropenie und den schrittweisen Prozess der Leukämie-Entstehung *in vitro* darstellen. Anhand dieses Modells konnte ich den Ausreifungsstop der Granulopoese darstellen und genügend Hämatopoetische Stammzellen (HSZ) generieren um die myeloische Differenzierung, Multi-omics Analysen und die funktionelle Validierung von identifizierten Zielproteinen durchzuführen. Das Ziel war es neue Faktoren zu identifizieren, die dann weiter auf ihre therapeutische Wirksamkeit für CN/AML untersucht werden können.

Unsere Arbeitsgruppe hat kürzlich von einer hohen Frequenz von kooperierenden *CSF3R* (Kolonie-stimulierendem Faktor 3 Rezeptor) und *RUNX1* (Runt-verwandtem Transkriptionsfaktor 1) Mutationen in CN/AML Patienten berichtet. Ich habe ein *in vitro* Modell der stufenweisen Leukämie-Entstehung in CN/AML durch CRISPR/Cas9 Genom-Editierung von iPS-Zellen von zwei CN/AML Patienten mit *CSF3R* und *RUNX1* Mutationen etabliert. Mithilfe dieses Modells habe ich die stark erhöhte Expression von BAALC (brain and acute leukemia, cytoplasmic) als Schlüssel-Ereignis der Leukämie-Entstehung bei

kongenitaler Neutropenie identifiziert. Der Ausschalten des BAALC Gens durch CRISPR/Cas9 konnte die defekte myeloische Differenzierung in CN/AML-iPS-Zellen, vergleichbar zu der Differenzierung von gesunden HSZ, wiederherstellen. Mittels Transkriptom-Analyse, identifizierte ich das Molekül, CMPD1, ein Inhibitor der p38-vermittelten MK2a Phosphorylierung, das die Genexpression zu einem ähnlichen Ausmaß wie der BAALC Knockout wiederherstellen konnte. Die *in vitro* Anwendung von CMPD1 bei primären Blasten von CN/AML Patienten führte zu einer deutlichen Reduzierung der Proliferationsrate, ohne dabei die Differenzierung von gesunden HSZ zu beeinflussen. Zusammenfassend legen diese Beobachtungen nahe, dass die Inhibition von BAALC in hämatopoetischen Stammzellen von Patienten mit kongenitaler Neutropenie möglicherweise die Leukämie-Entstehung verhindern oder AML-Blasten in CN/AML Patienten eliminieren kann.

LIST OF PUBLICATIONS

- A. **Dannenmann B**, Zahabi A, Mir P, Oswald B, Bernhard R, Klimiankou M, Morishima T, Schulze-Osthoff K, Zeidler C, Kanz L, Lachmann N, Moritz T, Welte K and Skokowa J. Human iPSC-based model of severe congenital neutropenia reveals elevated UPR and DNA damage in CD34⁺ cells preceding leukemic transformation. *Exp Hematol.* 2019 Mar;71:51-60.
- B. **Dannenmann B***, Nasri M*, Welte K, Skokowa J. CRISPR/Cas9 Genome Editing of Human-Induced Pluripotent Stem Cells Followed by Granulocytic Differentiation. *Methods Mol Biol.* 2020;2115:471-483. (*equal contribution).
- C. Nasri M*, Mir P*, **Dannenmann B**, Amend D, Skroblyn T, Xu Y, Schulze-Osthoff K, Klimiankou M, Welte K, Skokowa J. Fluorescent labeling of CRISPR/Cas9 RNP for gene knockout in HSPCs and iPSCs reveals an essential role for GADD45b in stress response. *Blood Adv.* 2019 Jan 8;3(1):63-71. (*equal contribution).
- D. Mir P, Klimiankou M, **Dannenmann B**, Aghaallaei N, Nasri M, Wingert S, Thalheimer F, Findik B, Doll L, Ritter M, Thumberger T, Kandabarau S, Zeidler C, Rieger M, Bajoghli B, Welte K, Skokowa J. Retinoic acid restores granulocytic differentiation in severe congenital neutropenia: Role of GADD45 β -dependent gene demethylation. (in revision)
- E. **Dannenmann B***, Klimiankou M*, Oswald B, Solovyeva A, Mardan J, Krahl AC, Nasri M, Ritter M, Zahabi A, Mir P, Stein F, Bernhard R, Lachmann N, Moritz T, Morishima T, Ripperger T, Steinemann D, Zeidler C, Welte K and Skokowa J. hiPSC Model of Stage-Specific Leukemogenesis in Severe Congenital Neutropenia reveals BAALC as a Key Oncogene. (in revision), (*equal contribution).

- F. Nasri M, Ritter M*, Mir P*, **Dannenmann B***, Aghaallaei N, Amend D, Makaryan V, Xu Y, Fletcher B, Bernhard R, Steiert I, Hahnel K, Berger J, Koch I, Sailer B, Hipp K, Zeidler C, Klimiankou M, Bajoghli B, Dale DC, Welte K, Skokowa J. CRISPR/Cas9-mediated *ELANE* knockout enables neutrophilic maturation of primary hematopoietic stem and progenitor cells and induced pluripotent stem cells of severe congenital neutropenia patients. *Haematologica*. 2020;105(3):598-609. (*equal contribution).
- G. Witte A, Rohlfing AK, **Dannenmann B**, Dicenta V, Nasri M, Kolb K, Sudmann J, Castor T, Rath D, Borst O, Skokowa J, Gawaz M. The chemokine CXCL14 mediates platelet function and migration via direct interaction with CXCR4. *Cardiovasc Res*. 2020 Apr 1. pii: cvaa080. doi: 10.1093/cvr/cvaa080.
- H. Xu Y, Nasri M, **Dannenmann B**, Mir P, Zahabi A, Welte K, Morishima T, Skokowa J. NAMPT/SIRT2-mediated inhibition of the p53-p21 signaling pathway is indispensable for maintenance and hematopoietic differentiation of human iPS cells. (in revision)
- I. Malak PN, **Dannenmann B**, Hirth A, Rothfuss OC, Schulze-Osthoff K. Novel AKT phosphorylation sites identified in the pluripotency factors OCT4, SOX2 and KLF4. *Cell Cycle*. 2015;14(23):3748-54.

PERSONAL CONTRIBUTION

The following section indicates the personal contribution to each of the above listed publications.

- A. **First Author.** Design of experiments, performing the main experiments, analysis of data and preparation of the manuscript.
- B. **Co-first Author.** Design of the chapter; design, performance and presentation of the experiments described in the chapter. Writing the chapter.
- C. **Co-second Author.** Generation, characterisation and hematopoietic differentiation of iPSCs before- and after CRISPR/Cas9 *GADD45B* gene-knockout.
- D. **Co-second Author.** Generation, characterisation and hematopoietic differentiation of iPSCs before- and after CRISPR/Cas9 *GADD45B* gene-knockout.
- E. **Co-first Author.** Design of experiments, performing the main experiments, analysis of data and preparation of the manuscript.
- F. **Co-second Author.** Generation, characterisation and hematopoietic differentiation of iPSCs before and after CRISPR/Cas9 *ELANE* gene-editing, qRT-PCR for iPSC-derived hematopoietic/myeloid cells.
- G. **Co-second Author.** Establishment of a protocol of iPSC-based generation of megakaryocytes and platelets. Generation of iPSC lines and production of iPSC-derived megakaryocytes and platelets.
- H. **Co-second Author.** Establishment of a protocol for hematopoietic differentiation of hiPSCs and establishment of multi-color flow cytometry panels for characterisation of immature and mature myeloid cells.
- I. **Co-second Author.** Performing the main experiments and analyse data.

INTRODUCTION

Pre-leukemia bone marrow failure syndromes and leukemia progression

The group of pre-leukemic bone marrow failure syndromes (BMFS) describes inherited hematological disorders that can lead to the development of myelodysplastic syndrome (MDS) or acute myeloid leukemia (AML). They are characterized by an abnormal proliferation and differentiation of hematopoietic stem and progenitor cells (HSPC)¹⁻⁸.

To be able to treat or prevent AML it is important to understand the mechanisms of leukemia development first. By now, it is hypothesized that the regulation, the self-renewal as well as the proliferation and differentiation of HSPCs is disturbed due to inherited mutations^{1-6,8}. Chronic stress that is induced by intrinsic defects and extrinsic abnormalities in the HSPC niche could cause an elevated DNA-damage response or diminished DNA-damage repair that increases the possibility of acquisition of mutations in leukemia-related genes or karyotypical abnormalities^{3,9}. The expansion and evolution of mutant HSPC clones can be supported by a selective pressure induced by the background of an inherited disease¹⁰.

Severe congenital neutropenia, *ELANE* mutations and deregulated intracellular signaling

Severe congenital neutropenia (CN), a pre-leukemia BMFS, is predominantly caused by inherited heterozygous missense *ELANE* mutations encoding neutrophil elastase (NE)¹¹. *ELANE* mutations are distributed throughout all five exons of the *ELANE* gene, but the majority of mutations can be found in exon 4 and exon 5¹². Other inherited gene mutations leading to CN are detected in *HAXI*¹³ (HCLS-1 associated protein X-1), *G6PC3*¹⁴ (glucose 6 phosphatase, catalytic, 3), *JAGNI* (Jagunal Homolog 1) and other genes.

HSPCs of CN patients lack the ability to differentiate into mature neutrophils, but show no differentiation defects in other blood lineages. CN patients without treatment show severe infections due to maturation arrest at promyelocytic stage^{2,3,15,16}. Mechanisms for granulocytic differentiation defects are still not fully understood. Our group and others have shown that mutant NE activates unfolded protein response (UPR) and induces endoplasmic reticulum (ER) stress by accumulation of misfolded NE protein within the ER¹⁷⁻¹⁹.

Our laboratory also described several deregulated intracellular signaling pathways in hematopoietic cells of CN patients. Thus, we found severely diminished expression levels of transcription factors LEF-1 and C/EBP α , adaptor protein HCLS1, and neutrophil elastase inhibitor SLPI^{2,3}. In parallel, chronic G-CSF therapy led to compensatory hyperactivation of phospho-STAT5a, NAMPT/SIRT1 pathway and C/EBP β transcription factor expression^{2,3,20}. All these processes may contribute to the activation/deregulation of the intracellular stress response mechanisms that may ultimately lead to the leukemogenic transformation of hematopoietic stem cells and leukemia.

Leukemia progression in CN patients

Approximately 20 % of CN patients develop MDS or AML later in life^{3,21,22}. Treatment of CN patients with high doses of granulocyte colony-stimulating factor (G-CSF) can partially rescue neutrophilic maturation defect^{3,16}. However, treatment with high doses of G-CSF also leads to a higher probability of developing MDS or AML^{21,22}.

Our research group recently reported a cooperativity of acquired *CSF3R* (encoding the G-CSF receptor) and *RUNX1* (runt-related transcription factor 1) mutations in 31 CN/AML patients²³. 65 % of them had acquired *RUNX1* mutations and 80 % co-acquired *CSF3R* mutations. Co-acquisition of *CSF3R* and *RUNX1* mutations led to enhanced proliferation and diminished

myeloid differentiation of healthy donor CD34⁺ cells in functional tests. Further, CN/AML patients with *RUNXI* mutations showed a high frequency of trisomy 21 and monosomy 7²³.

By screening 307 CN patients with *ELANE* mutations from Severe Chronic Neutropenia International Registry (SCNIR) for up to 27 years, we identified in collaboration with other groups “hot-spot” *ELANE* mutations (e.g., p.C151R, p.C151Y, p.G214R) in CN patients who developed MDS or AML¹². However, it is still unclear how *CSF3R* and *RUNXI* mutations in cooperation with these *ELANE* mutations can induce leukemia in CN patients.

***CSF3R* mutations and their effects on G-CSFR signaling**

To identify possible mechanisms of leukemogenesis it is important to know the effects of *CSF3R* and *RUNXI* mutations on their downstream signaling pathways. Acquired *CSF3R* mutations mainly occur in the intracellular domain of G-CSFR. Mutations in this domain are stop codon mutations leading to a truncated G-CSFR with interrupted signaling. This part is responsible for the termination of STAT5 (signal transducer and activator of transcription 5) - dependent proliferative signals and STAT3 - dependent activation of differentiation. Consequences are increased proliferation and diminished differentiation ability of HSPCs carrying this mutation^{2,20,24-32}. G-CSFR mutations alone are not sufficient for leukemic transformation in CN. In many cases of myeloid leukemia, only 1 or 2 additional mutations are needed to induce leukemia.³³

Missense and nonsense *RUNXI* mutations and their implications in leukemogenesis

RUNXI mutations predominantly occur in the Runt homology domain (RHD), which is responsible for DNA-binding, and in the transactivation domain (TAD), which is needed for protein-protein interaction³⁴. *RUNXI* point mutations were reported with a frequency of 6 - 33 % in *de novo* AML³⁵⁻³⁹ and are highly associated with monosomy 7 and trisomy 21^{38,40,41}.

Acquired or inherited *RUNXI* mutations are found in several pre-leukemic diseases. Fanconi anemia patients that developed AML or MDS showed acquired mutations translocations or deletions in *RUNXI*⁹ and familial platelet disorder (FPD) patients with a predisposition to AML have inherited *RUNXI* mutations^{42,43}.

In CN/AML patients, missense and nonsense *RUNXI* mutations have been described^{23,44}. Several studies proposed possible mechanisms of leukemogenesis downstream of both missense and nonsense *RUNXI* mutations^{43,45,46}. *RUNXI* mutations in the RHD might induce plenty of effects including loss of DNA and/or CBF β -binding, disturbed heterodimerization, inefficient transactivation, protein mislocalization, appearance of dominant-negative alleles or haploinsufficiency⁴⁶⁻⁵¹. Missense mutations in *RUNXI* can result in dominant-negative protein isoforms with numerous outcomes: decreased CBF β -binding, protein mislocalization and impaired DNA-binding^{46,48,50,51}.

Missense *RUNXI* mutations in the RHD can also alter the activity of other RUNX1-binding proteins and inhibit the transactivation of wild-type (WT) RUNX1 protein^{43,48}. Moreover, the *RUNXI* gene is located on chromosome 21 and most CN/AML patients with missense *RUNXI* mutations have trisomy 21, indicating that the ratio of mutant to WT *RUNXI* alleles might play a role in leukemogenic transformation in CN patients²³.

Human induced pluripotent stem cell (hiPSC) – a reliable model to study leukemogenesis in CN

Since there are no animal models to study CN and leukemogenesis in CN, establishment of an *in vitro* patient-specific iPSC-model is the only convenient alternative. Transgenic mice with knock-in of human *ELANE* mutation or *ELANE* KO mice did not show a neutropenia phenotype¹⁸. Patient-specific CN-iPSCs could already recapitulate the maturation arrest of granulopoiesis⁵²⁻⁵⁴. Moreover, iPSC-models were also used to study leukemogenesis of *de*

novo AML⁵⁵ and for drug screening. By a combination of iPSC-reprogramming and CRISPR/Cas9 gene-editing, iPSC-models can be used to study stepwise leukemogenesis⁵⁶. Such models could be also used to compare the effects of missense and nonsense mutations in endogenous genes on their functional outcomes.

OBJECTIVES OF THE STUDY

- 1) To establish an *in vitro* experimental model of CN and stepwise leukemia development in CN using patients-derived iPSC and CRISPR/Cas9 gene-editing.
- 2) To determine the functional relationship between *RUNX1* and *CSF3R* mutations and their role in leukemia development in CN patients, by analyzing myeloid differentiation potential of patient-derived iPSCs at different stages of leukemogenesis.
- 3) To identify intracellular signaling pathways responsible for leukemogenic transformation of HSPCs triggered by *CSF3R* and *RUNX1* mutations and to identify small molecules that can correct “leukemogenic” defects caused by *RUNX1* and *CSF3R* mutations.

RESULTS AND DISCUSSION

Results and discussion parts covered in three per-reviewed publications 'A-C', including one book chapter and in the manuscripts 'D, E' that are currently under revision.

Establishment of an *in vitro* model of CN and CN/AML using EB-based myeloid differentiation of patient-derived hiPSCs (see publications A, B)

We established a highly efficient and reproducible 3D spin-embryoid body (EB) based hiPSC-differentiation protocol that allows the production of high amounts of HSPCs (CD34⁺CD45⁺ cells) and mature granulocytes (more than 1x10⁶ cells/10 EBs). This consists mainly of three phases. The first non-adherent phase until day four of iPSC-differentiation leads to mesoderm induction, the second phase until day 14 produces a high amount of HSPCs, and the third phase until day 28 leads to the maturation to granulocytes. Using this protocol, we could show a strong reduction of HSPC numbers at day 14 for CN/AML-iPSCs and highly reduced numbers of mature myeloid cells at day 28 of differentiation for CN iPSC and especially for CN/AML iPSC. Flow cytometry and Wright-Giemsa stainings of cytopspin slides at day 28 also revealed significantly decreased percentages of PMNs for CN iPSCs, no PMNs for CN/AML-iPSC and highly increased percentage of monocytes for CN iPSCs that is in line with observations from CN patients which have increased peripheral blood monocyte counts^{3,57}. Using our EB-based hiPSC differentiation model, we wanted to have a closer look at the variables that could influence and enforce leukemic transformation in CN, like increased UPR stress, DNA damage response and diminished apoptosis.

Increased ER stress and UPR response in CN- and CN/AML-derived HSPCs (CD34⁺CD45⁺ cells) (see publication A)

We observed highly upregulated *ELANE* mRNA levels in CD34⁺CD45⁺ cells generated from CN/AML iPSCs, but unchanged *ELANE* mRNA levels for CN1 iPSCs and slightly increased levels for CN2 iPSCs compared to HD. NE protein levels were highly up-regulated, especially in CN/AML samples. *ELANE* expression is regulated by RUNX1⁵⁸, which could explain the highly increased NE levels in CN/AML HSPCs. It is still unclear how *RUNX1* mutations affect *ELANE* expression, but it has been shown that HSPCs of *RUNX1*^{-/-} mice have a weaker UPR response⁵⁹. Based on our results, mutated *RUNX1* may upregulate the expression of mutated NE in CN and amplify UPR response. To proof this hypothesis, we should analyze UPR response in CN/AML *RUNX1* KO iPSC - derived HSPCs.

It has been reported that mutated neutrophil elastase (NE) induces elevated ER stress and UPR in HSPCs and granulocytes of CN patients¹⁷⁻¹⁹. Consistent with these data, CN- and CN/AML-iPSC derived CD34⁺CD45⁺ cells both highly express UPR downstream regulator CHOP (DDIT3). The mRNA expression of genes involved in the other UPR pathways, like BiP, ATF4 and ATF6 was not commonly upregulated in HSPCs derived from all iPSC clones, but only from CN and CN/AML iPSCs. Thus, only CHOP, which is a downstream target of several UPR pathways, seems to be commonly upregulated in both CN and CN/AML. This observation was also confirmed on protein level: CHOP protein levels were upregulated to a similar extent in both CN- and CN/AML-derived myeloid cells at day 28 of iPSC differentiation. Increased expression of mutant NE protein, which is mainly inducing CHOP upregulation and thus increasing UPR stress, could be the reason for acquisition of further gene mutations like in *CSF3R*. Chronically elevated UPR in HSPCs of CN patients with inherited *ELANE* mutations might cause genotoxic stress which leads to an increased susceptibility of HSPCs to acquisition of secondary leukemia-initiating mutations. However,

the complete pathomechanism of leukemogenesis on the background of hyperactivated UPR is still unclear.

Increased DNA damage response in CN-iPSC-derived HSPCs (see publication A)

Next we tested the susceptibility of pre-leukemic CN-iPSC-derived HSPCs to DNA damage by five minutes treatment of cells with the cytostatic bleomycin. DNA damage repair ability was also tested after two hours of recovery time in cell culture. For both analysis of DNA damage and DNA damage repair, we performed a long-run real-time PCR-based DNA damage quantification (LORD-Q) of two nuclear DNA loci in *GAPDH* and *TP53* and additionally of mitochondrial DNA (mt DNA)⁶⁰. Highly increased DNA damage was observed for mtDNA and *TP53* loci in CN-iPSC-derived HSPCs compared to HD-iPSC-derived cells. Also, DNA damage repair was diminished after recovery time for CN-compared to HD-derived HSPCs, especially for mtDNA. Taken together, these results suggest an increased susceptibility to DNA damage in CN-iPSC-derived HSPCs and a delayed DNA damage repair.

Elevated DNA damage and diminished granulocytic differentiation of *GADD45b* deficient HSPCs and iPSCs (see publication C and manuscript E)

We next aimed to evaluate whether elevated DNA damage and deregulated expression of the proteins related to the DNA damage response, might be involved in the diminished granulocytic differentiation of HSPCs in congenital neutropenia. We found that *GADD45B* (Growth Arrest And DNA Damage Inducible Beta) expression was upregulated in HSPCs of healthy donors upon G-CSF treatment, but not in CN patients. It seems that the proper cellular response to G-CSF requires activation of signaling pathways involved in the regulation of stress/DNA damage response, e.g., *GADD45B*-triggered signaling. To test this hypothesis, we

performed CRISPR/Cas9 mediated knockout of GADD45B in iPSC derived from a healthy donor and generated pure GADD45 β knockout iPSC clones. Thus, our iPSC-based experimental model of hematopoiesis allows unique investigations of the role of candidate proteins in “clean settings”. Also, the effects of the gene dosage of the candidate proteins might be investigated using CRISPR/Cas9 gene-edited iPSC cells, by analyzing the heterozygous iPSC clones. Thus, using this model, we evaluated DNA damage response as well as the ability to differentiate into hematopoietic/myeloid cells of iPSCs in the absence of GADD45 β protein. We evaluated whether iPSCs remained pluripotent after GADD45B knockout and assessed the DNA damage response of GADD45 β deficient iPSCs and HSPCs exposed to UV irradiation. Indeed, we found that although GADD45B KO iPSCs retained pluripotency, they showed elevated DNA damage upon UV exposure. DNA damage was analysed using γ H2AX staining by FACS as well as a long-run real-time PCR-based DNA-damage quantification method for nuclear and mitochondrial genome analysis (LORD-Q⁶⁰). To study the role of GADD45 β in granulocytic differentiation, we tested pure GADD45B KO iPSC clones. Indeed, we were able to recreate CN-specific “maturation arrest” of granulopoiesis: we found that no CFU-G colonies were generated from GADD45B^{-/-} iPSC-derived CD34⁺ HSPCs, compared to healthy control iPSCs. Moreover, using the embryoid body (EB)-based hematopoietic differentiation, we observed increased amounts of immature cells and strongly diminished amounts of mature neutrophils in the culture of GADD45B-deficient iPSCs compared to healthy control iPSCs. The data indicate an essential role of GADD45 β in granulopoiesis.

p21 expression is upregulated in CN- and CN/AML-derived HSPCs (see publication A)

Since it is known that p53-p21 pathway is activated upon DNA damage, we analyzed the mRNA expression of *TP53*, the p53 inhibitor MDM2 and the p53 downstream targets p21 and

GADD45a in iPSC-derived HSPCs. We observed no difference in p53 and GADD45a mRNA expression, but a three-fold increase of p21 mRNA expression for CN and six-fold increase for CN/AML compared to HD-derived cells. MDM2 mRNA was reduced in CN/AML cells. On protein level, we could also confirm increased p21 levels in CN cells that were even further increased in CN/AML cells. Increased p21 expression can be explained by either reduced MDM2 levels, since MDM2 is a negative regulator of p53, or by p53-independent p21 activation⁶¹.

p21 has anti-apoptotic functions in AML cells and high p21 levels can enforce chemoresistance in AML cells^{62,63}. Since p21 expression is already increased at the CN stage in our model, it indicates that p21 anti-apoptotic function can play a role in both, leukemic transformation and severity of the overt AML phenotype. High p21 levels might rescue stressed and mutated CN-stage cells from apoptosis and thus support leukemic transformation.

A possible connection between diminished GADD45 β expression and activation of the p53/p21 signaling in CN is interesting but unclear. GADD45 proteins are known targets of p53⁶⁴⁻⁶⁶ and are interaction partners of p21⁶⁷. There are different scenarios for CN and CN/AML concerning p53/p21 and GADD45 β . First, GADD45 β regulation in CN HSPCs, or HSPCs in general are p53-independent, or p53, for some reason, failed to regulate GADD45 β . Next, downstream targets or cellular processes downstream of the p21:GADD45 β protein complex (e.g. cell cycle check point) are not properly regulated in CN, because GADD45 β is very low expressed, or GADD45 β is not interacting with p21 in HSPCs and myeloid cells. All these scenarios are worse to study in connection to leukemogenesis, but also tumorigenesis in general.

Missense but not nonsense *RUNXI* mutations are associated with trisomy 21 in CN/AML patients (see manuscript E)

In CN/AML patients, three different types of *RUNXI* mutations are observed: missense, nonsense and frame-shift mutations. By comparing cytogenetic abnormalities and the type of acquired *RUNXI* mutations it was revealed that ~ 50 % of CN/AML patients with missense *RUNXI* mutations also acquired trisomy 21, whereas no trisomy 21 was detected in ~ 90 % of CN/AML patients with nonsense or frame-shift *RUNXI* mutations.

Based on these observations, we hypothesized that nonsense/frameshift and missense *RUNXI* mutations may lead to different functional outcomes. We evaluated whether an additional copy of the mutated or WT *RUNXI* is co-acquired with trisomy 21 in CN/AML patients. Digital PCR (dPCR) analysis showed two copies of mutated missense *RUNXI* in three CN/AML patients with trisomy 21, revealing an essential role of the dosage of missense *RUNXI* protein for leukemia initiation and progression. It seems that missense *RUNXI* mutations, but not nonsense or frameshift mutations, need the co-occurrence of trisomy 21 to enhance the leukemogenic effects of mutated *RUNXI*.

***In vitro* patient-derived hiPSC model of stepwise leukemic transformation in *ELANE*-mutant CN/AML (see manuscript E)**

To obtain deeper insight into the process of leukemic transformation in CN patients, we used our already described hiPSC- differentiation model. We reprogrammed iPSCs from PB MNCs of AML-stage of CN/AML patients harboring *ELANE* mutations p.C151Y and p.G214R. Our aim was to obtain iPSC for three different stages of leukemia progression in CN: firstly the CN stage with only an *ELANE* mutation, secondly the MDS-stage with an additional *CSF3R* mutation and thirdly the AML-stage with an additional *RUNXI* mutation.

For CN/AML patient 1, we could receive CN and CN/AML iPSC clones after reprogramming of peripheral blood mononuclear cells (PBMNCs). The intermediate stage with only *CSF3R* mutation we had to produce by CRISPR/Cas9 gene-editing. For CN/AML patient 2, we obtained CN and CN2+*CSF3R* Q743X^{+/-} clones from reprogramming of bone marrow mononuclear cells (BMMNCs). CN/AML2 clone we had produced by adding an additional *RUNXI* frameshift mutation to CN2+*CSF3R* Q743X^{+/-} clone using the CRISPR/Cas9 gene-editing system. CN/AML1 iPSC clones with *RUNXI* missense mutations had additional trisomy 21 which was confirmed in the Array-CGH. dPCR also confirmed, according to our observation in CN/AML patients, that the mutated *RUNXI* to the WT *RUNXI* ratio was 2 to 1 in CN/AML iPSCs and iPSC-derived CD45⁺ cells.

Thus, by combining patient-specific iPSC-technology for various CN/AML patients, in the form of reprogramming of CN HSPCs and CN/AML blasts to iPSCs, with CRISPR/Cas9 gene-editing, we provided the perfect platform to study step-wise leukemogenesis in CN. These iPSCs can be used then to study hematopoietic differentiation and to identify leukemia-related gene-expression profiles and signaling pathways.

Severely impaired myeloid differentiation and highly elevated proliferation rate of CN/AML-iPSC-derived HSPCs (see manuscript E)

We differentiated all iPSC clones from both CN/AML patients with our already described EB-based iPSC differentiation protocol into HSPCs and mature myeloid cells. We found an increased proportion of CD45⁺CD34⁺ HSPCs for CN/AML clones and these cells also showed a ~3-fold increased proliferation rate compared to HD when cultured on SL/SL feeder cells with a cytokine mix supporting growth of HSPCs. Already HSCs derived from CN1+*CSF3R* Q741fs or CN2+*CSF3R* Q743X^{+/-} - iPSCs showed an increased proliferation rate, but the additional *RUNXI* mutation led to a further significant increase in the

proliferation. Thus, we could confirm that acquired *CSF3R* mutations lead to a truncated G-CSFR resulting in a proliferative advantage of HSPCs, which predisposes to leukemia, but is not sufficient to induce it^{3,28,31,68-71}. Further, we showed that acquisition of an additional *RUNX1* mutation has a multiplying effect on proliferation advantage of HSPCs, even in the absence of G-CSF, since our proliferation experiments on SL/SL feeder cells were performed without the addition of G-CSF.

Differentiation to the later myeloid stages revealed that CN/AML iPSCs were able to produce only a severely reduced amount of granulocytes (CD45⁺CD15⁺CD11b⁺) and mature neutrophils (CD45⁺CD15⁺CD16⁺). Further, we showed a step-wise reduction in the percentage of granulocytes and mature neutrophils from HD to CN and to CN with mutated G-CSFR and to CN/AML by flow cytometry and Wright-Giemsa staining. By this, we proved the strong impact of the acquisition of *CSF3R* and additional *RUNX1* mutations on leukemic progression in CN *in vitro* with increased phenotype severity in the form of highly decreased differentiation potential of mutated HSCs together with highly increased proliferation shown before.

These observations were supported by CFU Assay results that revealed the highly diminished potential of CN/AML-derived HSCs to form any kind of CFU, especially CFU-G. The total number of CFU colonies was already very low in CN with mutated G-CSFR, but the number of CFU-G colonies was still further decreased in CN/AML cells.

AML-related gene expression signature is present in CN/AML blasts and in CN/AML-iPSC derived HSPCs (see manuscript E)

Since we observed the upregulation of AML-related genes in a microarray study of AML blasts of CN/AML patient 1, we examined mRNA expression of the top four hematopoiesis-related CN/AML-related genes from this microarray, *CD34*, *BAALC*, *CD109* and *HPGDS*.

We studied primary blasts of five CN/AML patients and BMMNCs of three CN patients and found a strong upregulation of all four genes in all five CN/AML patients. Next, we wanted to evaluate whether these genes are also upregulated in CN/AML iPSC-derived HSPCs. In CN/AML patient 1, which was also used for the initial microarray experiments, mRNA of all four genes was highly upregulated in HSPCs from two CN/AML iPSC clones compared to CN-stage. In CN/AML patient 2, only *BAALC* mRNA was upregulated. BAALC protein levels were also highly upregulated in CN/AML iPSC-derived HSPCs of both CN/AML patients.

Next we wanted to test if BAALC expression is dependent on *RUNX1* mutations. Transduction of HD CD34⁺ cells with lentiviral constructs expressing either *RUNX1* wt or the *RUNX1* mut with the same *RUNX1* mutations as in our CN/AML patients revealed that both *RUNX1* mutations lead to an upregulation of *BAALC* mRNA expression, as compared to cells transduced with WT *RUNX1*.

Since we previously showed the elevation of STAT5a^{20,46,67,68} levels in CN and CN/AML patients, we analyzed mRNA and protein expression of STAT5a, RUNX1 and G-CSFR in our iPSC-derived HSPCs. Indeed with observed upregulation of STAT5a and RUNX1 on mRNA and protein levels, especially in CN/AML iPSC-derived HSPCs and to a smaller extent at the CN-stage, whereas G-CSFR expression was mainly unchanged.

Taken together, increased expression of CN/AML-related genes, increased proliferation rate and diminished differentiation ability represent major leukemic features. Thus, our iPSC-model perfectly recapitulates step-wise leukemia development in CN and becomes even more important due to the fact that mouse models with the mutated *ELANE* do not show any neutropenia phenotype.

Elevated BAALC expression is a key feature of leukemogenic transformation in CN (see manuscript E)

We generated *BAALC*, *HPGDS*, *CD109* and *RUNX1* iPSC knockout (KO) lines via CRISPR/Cas9 gene-editing using CN/AML1 iPSCs to determine which of these CN/AML-related genes is responsible for leukemogenic transformation in CN. We differentiated each KO iPSCs using our EB-based differentiation protocol and compared their differentiation ability to CN/AML iPSCs w/o KO. Strikingly, *BAALC* KO could reverse AML phenotype and induce granulocytic differentiation in CN/AML iPSCs, even resulting in mature neutrophils. The percentage of granulocytes produced by CN/AML iPSC upon *BAALC* KO was comparable to HD-iPSC derived neutrophils. Further, *BAALC* KO reduced proliferation rate of CN/AML iPSC derived HSPCs by 3-fold and re-established the potential to form CFU. KO of *HPGDS*, *CD109* or *RUNX1* was not able to restore granulocytic differentiation in CN/AML iPSCs. We could also confirm the induction of granulocytic differentiation, decreased proliferation and increased CFU-potential for the CN/AML iPSCs of the second patient. Interestingly, *BAALC* KO could even induce granulocytic differentiation of CN-iPSCs, indicating that inherited *ELANE* mutations influence BAALC expression levels. Thus, we identified BAALC as the key regulator of leukemogenic transformation in CN.

Leukemic gene expression signature is present in CN/AML-iPSC-derived HSPCs (see manuscript E)

To obtain deeper insight in the underlying gene expression profile changes and pathways of leukemic transformation in CN, we performed RNA-seq of CN and CN/AML iPSC-derived HSPCs. Differential gene expression analysis identified 132 up- and 570 down-regulated genes for CN/AML1 vs CN1. In CN/AML2 HSPCs, 579 genes were up- and 1422 down-regulated, as compared to CN2 cells.

The most significantly upregulated pathways for CN/AML1 revealed by gene set enrichment analysis (GSEA) were “E2F targets”, “oxidative phosphorylation” genes and “MYC targets”, whereas “platelet-specific genes” (classical *RUNX1* downstream targets) were downregulated. For CN/AML patient 2, major enriched gene sets were “GM2 checkpoint”, “E2F targets” and “TGF β signaling”, whereas the Gene Ontology of the “Structural Constituent of Ribosomes” (classical *RUNX1* downstream targets) was regulated in the opposite way.

Using the Transcription Factor enrichment analysis, we predicted transcription factors (TF) that control differentially expressed genes in both CN/AML patients. Among these, motifs of the RUNX1 binding partners *GATA1* and *GATA2* as well as of the AML-associated gene *SUZ12* were most significantly enriched. Motifs specifically presented in only one CN/AML patient were *TRIM28* and *TP53* for CN/AML1 and *CEBPB*, *NANOG* and *KLF4* for CN/AML2.

Using Kinase Enrichment Analysis (KEA), we predicted kinases that phosphorylate proteins of prioritized transcription factors networks. HIPK2, MAPK3/1/14, CSNK2A1, ERK1 and AKT1 were shared between both patients, whereas several unique enriched kinases like CDK1, JNK1 and ERK2 were detected in CN/AML patient 2 only.

Taken together, RNA-Seq analysis of iPSC-derived HSPCs revealed a substantial difference in the gene expression signature between CN and CN/AML stages. Many signaling pathways, especially MAPK/ERK pathways, including various MAP kinases and transcription factors that can be downstream effectors in these pathways like *GATA1* and *GATA2*, were affected by both *RUNX1* missense and frameshift mutations. However, *RUNX1* frameshift mutations leading to truncated RUNX1 protein affected many more additional pathways. These observations might help to understand the mechanisms of leukemogenesis downstream of different types of *RUNX1* mutations.

It would be interesting to investigate whether reduced GADD45 β expression, in combination with *RUNX1* and/or *CSF3R* mutations, has an impact on the leukemogenic transformation of

HSPCs. It is known that in hematopoietic cells, GADD45 β plays a role in JNK pathway inactivation by interaction with MKK4⁷². Reduced GADD45 β in CN hematopoietic cells may lead to the JNK activation. To this end, we may introduce *RUNX1*, *CSF3R* and *ELANE* mutations in the GADD45 β ^{-/-} or GADD45 β ^{+/-} iPSCs of healthy donors and study hematopoietic proliferation and differentiation of HSPCs derived from these iPSC clones. Studying intracellular pathways in HSPCs from gene-edited GADD45 β KO iPSC clones might shed light on a better understanding of the leukemia progression and the role of GADD45 β in this process.

BAALC-dependent leukemogenic signaling pathways in CN/AML (see manuscript E)

To examine BAALC-dependent AML-related gene expression, we performed RNA-seq of CN/AML-iPSC derived HSPCs before and after *BAALC* KO for both CN/AML patients. RNA-seq revealed 165 up- and 254 down-regulated genes for CN/AML1 *BAALC* KO, as well as 185 up- and 381 down-regulated genes for CN/AML2 *BAALC* KO. Venn Diagrams comparing differentially expressed genes of CN/AML vs CN/AML *BAALC* KO group with CN/AML vs CN group, showed 170 overlapping genes for CN/AML1 and 380 overlapping genes for CN/AML2, indicating that gene expression of many genes during leukemic transformation is BAALC-dependent. Additionally, GSEA showed an overlap with pathways already detected in CN vs CN/AML analysis, especially for CN/AML2. For CN/AML patient 1, oxidative phosphorylation and p53 signaling were enriched, platelet-specific genes and TCA cycle were inhibited in CN/AML cells. For CN/AML patient 2, E2F targets, G2M checkpoint-associated genes, TGF β signaling and MYC- targets were enriched in CN/AML cells. GSEA results suggested that elevated BAALC exerts its leukemogenic effects affecting different signaling pathways including Myc- or p53- downstream signaling. Moreover,

BAALC also seems to have an impact on the cell cycle regulation and transcriptional regulation. It would be essential to evaluate a connection between elevated BAALC and diminished GADD45 β expression in CN leukemogenesis, since both proteins seem to be connected to p53 and MAPK/JNK signaling.

Transcription factor enrichment analysis (TFEA) using the list of differentially expressed genes revealed that *RUNX1*, *GATA1/2* and *SUZ12* TF binding motifs were BAALC-dependent. *AR*, *TCF3*, *RAD21* and *NANOG* TF motifs were specific for CN/AML1. *EGRI*, *STAT3* and *ZC3H11a* were specific for CN/AML2. TFEA results indicated that BAALC also exerts its leukemogenic effects by deregulating the expression of major hematopoietic transcription factors like *RUNX1* and *GATA1/2* as well as tumor suppressor genes like *SUZ12*. KEA showed also BAALC-dependency of kinase, which were detected to be upregulated in CN/AML vs CN, namely *CSNK2A1*, *HIPK2* and *MAPK1* and *MAPK14*. CN/AML patient 2 showed patient-specific enrichment of *JNK1* and *ERK2* kinases again. KEA highlighted a strong impact of BAALC on various kinases of the MAPK/ERK pathway signaling, which is involved in cancerogenesis.

We also found that four key transcription factor motifs (*RUNX1*, *GATA1/2*, *SUZ12* and *SMAD4*) as well as seven kinases (e.g. *MAPK14*, *MAPK1*, *ERK1* and *AKT1*) were enriched in CN/AML-iPSC-derived HSPCs compared to either CN/AML *BAALC* KO or CN stage, thus indicating that *BAALC* KO in CN/AML HSPCs partially reproduces a CN-stage-like phenotype. *BAALC* KO induced a strong gene expression shift in CN/AML HSPCs of both patients independent of missense or frameshift *RUNX1* mutations. MAPK/ERK signaling pathways, which we already showed in the RNA-seq of CN vs CN/AML analysis to play a major role in leukemogenesis, seem to be strongly dependent on BAALC. However, there are also some pathways, like TCF3-, TP63- and STAT3-dependent signaling, which seem to be not essential for leukemic progression in CN. Taken together, it becomes clear that BAALC is exerting its leukemogenic effects in multiple ways by deregulating various signaling

pathways, expression of hematopoietic transcription factors and activation of kinases. Some of these deregulated pathways, transcription factors and kinases also seem to be dependent on the type of *RUNX1* mutation.

The selective p38 MAPK inhibitor CMPD1 suppresses proliferation of CN/AML cells (see manuscript E)

Since there are no BAALC inhibitors available and no crystal structure of the BAALC protein is known yet, we performed Connectivity Map (CMAP) analysis of RNA-Seq data comparing CN/AML vs CN/AML-*BAALC* KO samples, in order to find small molecules that could reproduce the effect of *BAALC* KO or inhibit pathways that are affected by hyperactivated BAALC. CMAP analysis revealed CMPD1 as the best small molecule capable of mimicking *BAALC* KO phenotype in CN/AML cells. CMPD1 is a selective p38 MAPK (MAPK14) inhibitor. To find an effective concentration of CMPD1, we first treated BAALC^{high} AML cell line Kasumi-1 with three concentrations of CMPD1 or the MEK1/2 inhibitor AZD-6244. The MEK1/2 inhibitor UO126, in combination with KLF4 activation, was reported to reduce the growth of BAALC-high *de novo* AML cells⁷³, but showed no effect on cell proliferation in our model. Therefore we chose the more specific MEK1/2 inhibitor AZD-6244. In Kasumi-1 cells, CMPD1 was sufficient in reducing cell proliferation rate at 1, 2 and 5 μ M concentrations. AZD-6244 had no effect at all concentrations, 1, 2 and 5 μ M.

Application of these small molecules for CN/AML iPSC-derived HSPCs also showed high effectiveness of CMPD1 for both CN/AML patients. It led to a reduction of proliferation rate down to 20 % without affecting proliferation of HD iPSC-derived HSPCs. AZD-6244 also reduced proliferation of CN/AML derived HSPCs, but to a smaller extent. Next, we tested 1 μ M CMPD1 and 1 μ M AZD-6244 for primary CN/AML blasts with high BAALC expression and could confirm an excellent efficiency of CMPD1 to reduce proliferation rate

of CN/AML blasts down to 10 %, whereas AZD-6244 could reduce proliferation rate down to 60 % only. Importantly, proliferation of HD CD34⁺ cells was not affected or only slightly inhibited after treatment with 1 μ M or 2 μ M of drugs. Finally, we applied these drugs for BAALC^{high} *de novo* AML blasts in 1 μ M and concentrations. Both drugs were effective to inhibit proliferation rate down to 60 %. The strong inhibition of proliferation of CN/AML-iPSC-derived HSPCs, as well as of primary CN/AML blasts and *de novo* AML blasts highlighted that the p38-MAPK pathway is a major downstream signaling pathway of hyperactivated BAALC in leukemogenesis.

OUTLOOK

Our iPSC-based experimental in vitro model of CN and CN/AML is a perfect system to study step-wise leukemogenesis in pre-leukemic bone marrow failure syndromes. Using this model, we were able to recreate a “maturation arrest” of granulopoiesis in vitro, which is similar to what we see in CN patients` bone marrow. Thus, this is a first experimental model that might be broadly used to study congenital neutropenia, since there are no mouse models and very low amounts of primary patients` hematopoietic cells are available for scientific projects. Using this model, we identified elevated DNA damage and connected downregulation of *GADD45b* as one of the key mechanisms of defective granulopoiesis.

We also could show that both, *RUNX1* missense mutations with additional trisomy 21 or *RUNX1* frameshift mutations, lead to a similar leukemia phenotype with increased proliferation rate and decreased differentiation potential of HSPCs in CN. This was in line with the similar leukemogenic gene expression profiles and affected signaling pathways in CN/AML-iPSC derived HSPCs of patients carrying either missense or truncated *RUNX1* mutations. Only a minor number of affected pathways was dependent on the *RUNX1* mutation type.

We identified BAALC upregulation as a key leukemogenic event sufficient to induce leukemia, downstream of acquired *RUNX1* and *CSF3R* mutations. Ultimately, BAALC inhibition in CN HSPCs or treatment with CMPD1 or other p38 MAPK inhibitors might be a possible treatment strategy for *RUNX1*-mutated CN/AML and *de novo* AML patients. Because we demonstrated that BAALC is a key deregulating protein with leukemogenic activity, it would be highly favorable to identify a specific BAALC inhibitor. This may be achieved either by small molecule drug screening or designing a specific BAALC inhibitor after solving a crystal structure of BAALC protein.

The process of stepwise leukemogenesis in CN, the role of BAALC upregulation and possible treatment strategies for CN/AML are summarized in **Figure 1**.

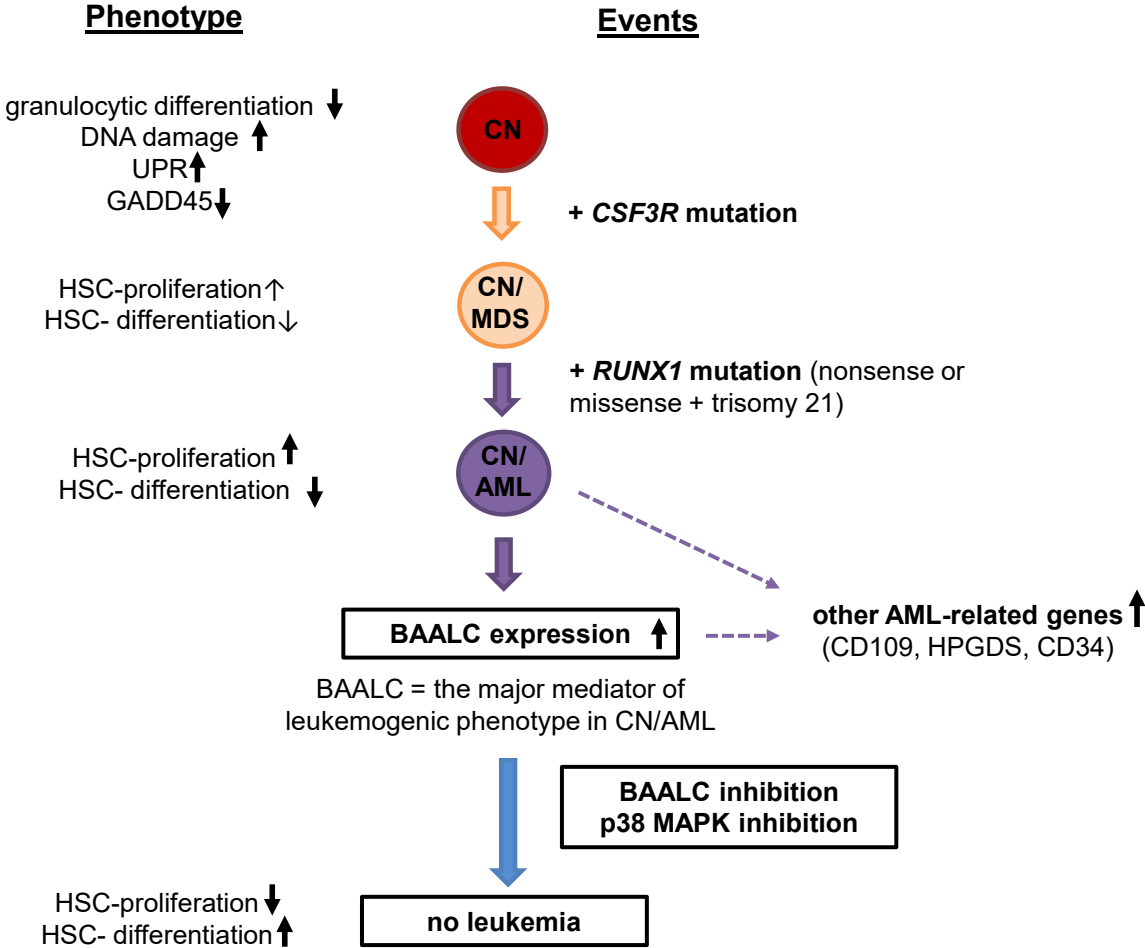


Figure 1: Scheme of the mechanism of leukemia development in CN and possible treatment strategies for CN/AML.

ACKNOWLEDGEMENTS

I would like to thank first my PhD Supervisor Prof. Dr. Julia Skokowa who supported me greatly during my PhD Thesis. I would like to thank all my colleagues who assisted in various experiments. Moreover, they were always there for discussions and for giving great suggestions.

Jose Carreras Leukemia Foundation, German-Israeli Foundation for Scientific Research and Development (GIF) and Fritz Thyssen Foundation financially supported me and my work. I also want to thank my parents for financial support outside of the lab.

REFERENCES

1. Feurstein, S., Drazer, M.W. & Godley, L.A. Genetic predisposition to leukemia and other hematologic malignancies. *Semin Oncol* **43**, 598-608 (2016).
2. Skokowa, J., Germeshausen, M., Zeidler, C. & Welte, K. Severe congenital neutropenia: inheritance and pathophysiology. *Curr Opin Hematol* **14**, 22-28 (2007).
3. Skokowa, J., Dale, D.C., Touw, I.P., Zeidler, C. & Welte, K. Severe congenital neutropenias. *Nat Rev Dis Primers* **3**, 17032 (2017).
4. Gohring, G., *et al.* Chromosomal aberrations in congenital bone marrow failure disorders--an early indicator for leukemogenesis? *Ann Hematol* **86**, 733-739 (2007).
5. Collins, J. & Dokal, I. Inherited bone marrow failure syndromes. *Hematology* **20**, 433-434 (2015).
6. Alter, B.P. Inherited bone marrow failure syndromes: considerations pre- and posttransplant. *Blood* **130**, 2257-2264 (2017).
7. Savage, S.A. & Dufour, C. Classical inherited bone marrow failure syndromes with high risk for myelodysplastic syndrome and acute myelogenous leukemia. *Semin Hematol* **54**, 105-114 (2017).
8. Wegman-Ostrosky, T. & Savage, S.A. The genomics of inherited bone marrow failure: from mechanism to the clinic. *Br J Haematol* **177**, 526-542 (2017).
9. Quentin, S., *et al.* Myelodysplasia and leukemia of Fanconi anemia are associated with a specific pattern of genomic abnormalities that includes cryptic RUNX1/AML1 lesions. *Blood* **117**, e161-170 (2011).
10. Lensch, M.W., Rathbun, R.K., Olson, S.B., Jones, G.R. & Bagby, G.C., Jr. Selective pressure as an essential force in molecular evolution of myeloid leukemic clones: a view from the window of Fanconi anemia. *Leukemia* **13**, 1784-1789 (1999).
11. Dale, D.C., *et al.* Mutations in the gene encoding neutrophil elastase in congenital and cyclic neutropenia. *Blood* **96**, 2317-2322 (2000).
12. Makaryan, V., *et al.* The diversity of mutations and clinical outcomes for ELANE-associated neutropenia. *Curr Opin Hematol* **22**, 3-11 (2015).
13. Klein, C., *et al.* HAX1 deficiency causes autosomal recessive severe congenital neutropenia (Kostmann disease). *Nat Genet* **39**, 86-92 (2007).
14. Boztug, K., *et al.* A syndrome with congenital neutropenia and mutations in G6PC3. *N Engl J Med* **360**, 32-43 (2009).
15. Welte, K. & Dale, D. Pathophysiology and treatment of severe chronic neutropenia. *Ann Hematol* **72**, 158-165 (1996).
16. Welte, K., Zeidler, C. & Dale, D.C. Severe congenital neutropenia. *Semin Hematol* **43**, 189-195 (2006).
17. Nustede, R., *et al.* ELANE mutant-specific activation of different UPR pathways in congenital neutropenia. *Br J Haematol* **172**, 219-227 (2016).
18. Nanua, S., *et al.* Activation of the unfolded protein response is associated with impaired granulopoiesis in transgenic mice expressing mutant Elane. *Blood* **117**, 3539-3547 (2011).
19. Grenda, D.S., *et al.* Mutations of the ELA2 gene found in patients with severe congenital neutropenia induce the unfolded protein response and cellular apoptosis. *Blood* **110**, 4179-4187 (2007).
20. Gupta, K., *et al.* Bortezomib inhibits STAT5-dependent degradation of LEF-1, inducing granulocytic differentiation in congenital neutropenia CD34(+) cells. *Blood* **123**, 2550-2561 (2014).
21. Rosenberg, P.S., *et al.* Stable long-term risk of leukaemia in patients with severe congenital neutropenia maintained on G-CSF therapy. *Br J Haematol* **150**, 196-199 (2010).

22. Rosenberg, P.S., *et al.* The incidence of leukemia and mortality from sepsis in patients with severe congenital neutropenia receiving long-term G-CSF therapy. *Blood* **107**, 4628-4635 (2006).
23. Skokowa, J., *et al.* Cooperativity of RUNX1 and CSF3R mutations in severe congenital neutropenia: a unique pathway in myeloid leukemogenesis. *Blood* **123**, 2229-2237 (2014).
24. Dong, F., *et al.* Mutations in the gene for the granulocyte colony-stimulating-factor receptor in patients with acute myeloid leukemia preceded by severe congenital neutropenia. *N Engl J Med* **333**, 487-493 (1995).
25. Dong, F., *et al.* Mutations in the granulocyte colony-stimulating factor receptor gene in patients with severe congenital neutropenia. *Leukemia* **11**, 120-125 (1997).
26. Zeidler, C., Germeshausen, M., Klein, C. & Welte, K. Clinical implications of ELA2-, HAX1-, and G-CSF-receptor (CSF3R) mutations in severe congenital neutropenia. *Br J Haematol* **144**, 459-467 (2009).
27. Germeshausen, M., Skokowa, J., Ballmaier, M., Zeidler, C. & Welte, K. G-CSF receptor mutations in patients with congenital neutropenia. *Curr Opin Hematol* **15**, 332-337 (2008).
28. Germeshausen, M., Welte, K. & Ballmaier, M. In vivo expansion of cells expressing acquired CSF3R mutations in patients with severe congenital neutropenia. *Blood* **113**, 668-670 (2009).
29. Germeshausen, M., Ballmaier, M. & Welte, K. Incidence of CSF3R mutations in severe congenital neutropenia and relevance for leukemogenesis: Results of a long-term survey. *Blood* **109**, 93-99 (2007).
30. Tschan, C.A., Pilz, C., Zeidler, C., Welte, K. & Germeshausen, M. Time course of increasing numbers of mutations in the granulocyte colony-stimulating factor receptor gene in a patient with congenital neutropenia who developed leukemia. *Blood* **97**, 1882-1884 (2001).
31. Liu, F., *et al.* Csf3r mutations in mice confer a strong clonal HSC advantage via activation of Stat5. *The Journal of clinical investigation* **118**, 946-955 (2008).
32. Klimiankou, M., Mellor-Heineke, S., Zeidler, C., Welte, K. & Skokowa, J. Role of CSF3R mutations in the pathomechanism of congenital neutropenia and secondary acute myeloid leukemia. *Ann N Y Acad Sci* **1370**, 119-125 (2016).
33. Welch, J.S., *et al.* The origin and evolution of mutations in acute myeloid leukemia. *Cell* **150**, 264-278 (2012).
34. Harada, Y. & Harada, H. Molecular mechanisms that produce secondary MDS/AML by RUNX1/AML1 point mutations. *J Cell Biochem* **112**, 425-432 (2011).
35. Osato, M., *et al.* Biallelic and heterozygous point mutations in the runt domain of the AML1/PEBP2alphaB gene associated with myeloblastic leukemias. *Blood* **93**, 1817-1824 (1999).
36. Christiansen, D.H., Andersen, M.K. & Pedersen-Bjerggaard, J. Mutations of AML1 are common in therapy-related myelodysplasia following therapy with alkylating agents and are significantly associated with deletion or loss of chromosome arm 7q and with subsequent leukemic transformation. *Blood* **104**, 1474-1481 (2004).
37. Harada, H., Harada, Y., Tanaka, H., Kimura, A. & Inaba, T. Implications of somatic mutations in the AML1 gene in radiation-associated and therapy-related myelodysplastic syndrome/acute myeloid leukemia. *Blood* **101**, 673-680 (2003).
38. Schnittger, S., *et al.* RUNX1 mutations are frequent in de novo AML with noncomplex karyotype and confer an unfavorable prognosis. *Blood* **117**, 2348-2357 (2011).
39. Gaidzik, V.I., *et al.* RUNX1 mutations in acute myeloid leukemia: results from a comprehensive genetic and clinical analysis from the AML study group. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology* **29**, 1364-1372 (2011).
40. Preudhomme, C., *et al.* High incidence of biallelic point mutations in the Runt domain of the AML1/PEBP2 alpha B gene in Mo acute myeloid leukemia and in myeloid malignancies with acquired trisomy 21. *Blood* **96**, 2862-2869 (2000).

41. Taketani, T., *et al.* AML1/RUNX1 mutations are infrequent, but related to AML-M0, acquired trisomy 21, and leukemic transformation in pediatric hematologic malignancies. *Genes Chromosomes Cancer* **38**, 1-7 (2003).
42. Preudhomme, C., *et al.* High frequency of RUNX1 biallelic alteration in acute myeloid leukemia secondary to familial platelet disorder. *Blood* **113**, 5583-5587 (2009).
43. Michaud, J., *et al.* In vitro analyses of known and novel RUNX1/AML1 mutations in dominant familial platelet disorder with predisposition to acute myelogenous leukemia: implications for mechanisms of pathogenesis. *Blood* **99**, 1364-1372 (2002).
44. Beekman, R., *et al.* Sequential gain of mutations in severe congenital neutropenia progressing to acute myeloid leukemia. *Blood* **119**, 5071-5077 (2012).
45. Harada, Y. & Harada, H. Molecular pathways mediating MDS/AML with focus on AML1/RUNX1 point mutations. *J Cell Physiol* **220**, 16-20 (2009).
46. Imai, Y., *et al.* Mutations of the AML1 gene in myelodysplastic syndrome and their functional implications in leukemogenesis. *Blood* **96**, 3154-3160 (2000).
47. Nagata, T. & Werner, M.H. Functional mutagenesis of AML1/RUNX1 and PEBP2 beta/CBF beta define distinct, non-overlapping sites for DNA recognition and heterodimerization by the Runt domain. *J Mol Biol* **308**, 191-203 (2001).
48. Matheny, C.J., *et al.* Disease mutations in RUNX1 and RUNX2 create nonfunctional, dominant-negative, or hypomorphic alleles. *Embo j* **26**, 1163-1175 (2007).
49. Osato, M. Point mutations in the RUNX1/AML1 gene: another actor in RUNX leukemia. *Oncogene* **23**, 4284-4296 (2004).
50. Tahirov, T.H., *et al.* Crystallization and preliminary X-ray analyses of quaternary, ternary and binary protein-DNA complexes with involvement of AML1/Runx-1/CBFalpha Runt domain, CBFbeta and the C/EBPbeta bZip region. *Acta Crystallogr D Biol Crystallogr* **57**, 850-853 (2001).
51. Tahirov, T.H., *et al.* Structural analyses of DNA recognition by the AML1/Runx-1 Runt domain and its allosteric control by CBFbeta. *Cell* **104**, 755-767 (2001).
52. Morishima, T., *et al.* Genetic correction of HAX1 in induced pluripotent stem cells from a patient with severe congenital neutropenia improves defective granulopoiesis. *Haematologica* **99**, 19-27 (2014).
53. Nayak, R.C., *et al.* Pathogenesis of ELANE-mutant severe neutropenia revealed by induced pluripotent stem cells. *J Clin Invest* **125**, 3103-3116 (2015).
54. Hiramoto, T., *et al.* Wnt3a stimulates maturation of impaired neutrophils developed from severe congenital neutropenia patient-derived pluripotent stem cells. *Proc Natl Acad Sci U S A* **110**, 3023-3028 (2013).
55. Chao, M.P., *et al.* Human AML-iPSCs Reacquire Leukemic Properties after Differentiation and Model Clonal Variation of Disease. *Cell stem cell* **20**, 329-344.e327 (2017).
56. Kotini, A.G., *et al.* Stage-Specific Human Induced Pluripotent Stem Cells Map the Progression of Myeloid Transformation to Transplantable Leukemia. *Cell stem cell* **20**, 315-328.e317 (2017).
57. Bonilla, M.A., *et al.* Long-term safety of treatment with recombinant human granulocyte colony-stimulating factor (r-metHuG-CSF) in patients with severe congenital neutropenias. *Br J Haematol* **88**, 723-730 (1994).
58. Lausen, J., Liu, S., Fliegau, M., Lubbert, M. & Werner, M.H. ELA2 is regulated by hematopoietic transcription factors, but not repressed by AML1-ETO. *Oncogene* **25**, 1349-1357 (2006).
59. Cai, X., *et al.* Runx1 Deficiency Decreases Ribosome Biogenesis and Confers Stress Resistance to Hematopoietic Stem and Progenitor Cells. *Cell stem cell* **17**, 165-177 (2015).
60. Lehle, S., *et al.* LORD-Q: a long-run real-time PCR-based DNA-damage quantification method for nuclear and mitochondrial genome analysis. *Nucleic Acids Res* **42**, e41 (2014).
61. Galanos, P., *et al.* Chronic p53-independent p21 expression causes genomic instability by deregulating replication licensing. *Nature cell biology* **18**, 777-789 (2016).

62. Schepers, H., Geugien, M., Eggen, B.J. & Vellenga, E. Constitutive cytoplasmic localization of p21(Waf1/Cip1) affects the apoptotic process in monocytic leukaemia. *Leukemia* **17**, 2113-2121 (2003).
63. Abbas, T. & Dutta, A. p21 in cancer: intricate networks and multiple activities. *Nature reviews. Cancer* **9**, 400-414 (2009).
64. Kim, Y.A., *et al.* Gadd45beta is transcriptionally activated by p53 via p38alpha-mediated phosphorylation during myocardial ischemic injury. *J Mol Med (Berl)* **91**, 1303-1313 (2013).
65. Carrier, F., *et al.* Characterization of human Gadd45, a p53-regulated protein. *The Journal of biological chemistry* **269**, 32672-32677 (1994).
66. Zhan, Q., Chen, I.T., Antinore, M.J. & Fornace, A.J., Jr. Tumor suppressor p53 can participate in transcriptional induction of the GADD45 promoter in the absence of direct DNA binding. *Molecular and cellular biology* **18**, 2768-2778 (1998).
67. Vairapandi, M., Balliet, A.G., Fornace, A.J., Jr., Hoffman, B. & Liebermann, D.A. The differentiation primary response gene MyD118, related to GADD45, encodes for a nuclear protein which interacts with PCNA and p21WAF1/CIP1. *Oncogene* **12**, 2579-2594 (1996).
68. Ward, A.C., van Aesch, Y.M., Schelen, A.M. & Touw, I.P. Defective internalization and sustained activation of truncated granulocyte colony-stimulating factor receptor found in severe congenital neutropenia/acute myeloid leukemia. *Blood* **93**, 447-458 (1999).
69. Dong, F., *et al.* Distinct cytoplasmic regions of the human granulocyte colony-stimulating factor receptor involved in induction of proliferation and maturation. *Mol Cell Biol* **13**, 7774-7781 (1993).
70. Hermans, M.H., *et al.* Signaling mechanisms coupled to tyrosines in the granulocyte colony-stimulating factor receptor orchestrate G-CSF-induced expansion of myeloid progenitor cells. *Blood* **101**, 2584-2590 (2003).
71. Hermans, M.H., *et al.* Sustained receptor activation and hyperproliferation in response to granulocyte colony-stimulating factor (G-CSF) in mice with a severe congenital neutropenia/acute myeloid leukemia-derived mutation in the G-CSF receptor gene. *J Exp Med* **189**, 683-692 (1999).
72. Gupta, M., Gupta, S.K., Hoffman, B. & Liebermann, D.A. Gadd45a and Gadd45b protect hematopoietic cells from UV-induced apoptosis via distinct signaling pathways, including p38 activation and JNK inhibition. *The Journal of biological chemistry* **281**, 17552-17558 (2006).
73. Morita, K., *et al.* BAALC potentiates oncogenic ERK pathway through interactions with MEKK1 and KLF4. *Leukemia* **29**, 2248-2256 (2015).

APPENDIX: LIST OF PUBLICATIONS AND MANUSCRIPTS

- A.** Human iPSC-based model of severe congenital neutropenia reveals elevated UPR and DNA damage in CD34⁺ cells preceding leukemic transformation

- B.** CRISPR/Cas9 Genome Editing of Human-Induced Pluripotent Stem Cells Followed by Granulocytic Differentiation

- C.** Fluorescent labeling of CRISPR/Cas9 RNP for gene knockout in HSPCs and iPSCs reveals an essential role for GADD45b in stress response

- D.** Retinoic acid restores granulocytic differentiation in severe congenital neutropenia: Role of GADD45 β -dependent gene demethylation

- E.** hiPSC Model of Stage-Specific Leukemogenesis in Severe Congenital Neutropenia reveals BAALC as a Key Oncogene

PUBLICATION A:

**Human iPSC-based model of severe congenital neutropenia reveals elevated UPR and
DNA damage in CD34⁺ cells preceding leukemic transformation**



ELSEVIER



Experimental Hematology 2019;71:51–60

**Experimental
Hematology**

Human iPSC-based model of severe congenital neutropenia reveals elevated UPR and DNA damage in CD34⁺ cells preceding leukemic transformation

Benjamin Dannenmann^a, Azadeh Zahabi^a, Perihan Mir^{a,b}, Benedikt Oswald^a, Regine Bernhard^a, Maksim Klimiankou^a, Tatsuya Morishima^a, Klaus Schulze-Osthoff^{b,c}, Cornelia Zeidler^d, Lothar Kanz^a, Nico Lachmann^e, Thomas Moritz^e, Karl Welte^f, and Julia Skokowa^a

^aDepartment of Oncology, Hematology, Immunology, Rheumatology, and Pulmonology, University Hospital Tuebingen, Tuebingen, Germany;

^bGerman Cancer Research Center (DKFZ), Heidelberg, Germany; ^cInterfaculty Institute of Biochemistry, Tübingen University, Germany;

^dDepartment of Hematology, Oncology, and Bone Marrow Transplantation, Hannover Medical School, Hannover, Germany; ^eInstitute of Experimental Hematology, Hannover Medical School, Hannover, Germany; ^fUniversity Children's Hospital Tuebingen, Tuebingen, Germany

(Received 10 October 2018; revised 23 December 2018; accepted 30 December 2018)

We describe the establishment of an embryoid-body-based protocol for hematopoietic/myeloid differentiation of human induced pluripotent stem cells that allows the generation of CD34⁺ cells or mature myeloid cells in vitro. Using this model, we were able to recapitulate the defective granulocytic differentiation in patients with severe congenital neutropenia (CN), an inherited preleukemia bone marrow failure syndrome. Importantly, in vitro maturation arrest of granulopoiesis was associated with an elevated unfolded protein response (UPR) and enhanced expression of the cell cycle inhibitor p21. Consistent with this, we found that CD34⁺ cells of CN patients were highly susceptible to DNA damage and showed diminished DNA repair. These observations suggest that targeting the UPR pathway or inhibiting DNA damage might protect hematopoietic cells of CN patients from leukemogenic transformation, at least to some extent. © 2019 ISEH – Society for Hematology and Stem Cells. Published by Elsevier Inc. All rights reserved.

Severe congenital neutropenia (CN) is a monolineage preleukemia bone marrow failure syndrome characterized by early onset of neutropenia and severe infections due to promyelocytic maturational arrest in the bone marrow [1,2]. CN is a heterogeneous disease caused by mutations in a number of genes, including *ELANE* [3] (the most common [1]), *HAXI* [4], *CSF3R* [5,6], *JAGN1* [7], *G6PC3* [8], *TCIRG1* [9], and others. In most cases, *ELANE* mutations are missense mutations that are distributed throughout all five exons of the *ELANE* gene, although a majority of mutations are found in exons 4 and 5 [10]. Hematopoietic stem and progenitor cells (HSPCs) of CN patients (CN-HSPCs) fail to differentiate into neutrophilic granulocytes, but show no severe maturation defects in other blood

lineages [1,2,11,12]. Exposure of CN-HSPCs to high concentrations of granulocyte colony-stimulating factor (G-CSF) partially reverses granulocytic maturation defects [1,12], but approximately 10% of CN patients do not respond to G-CSF doses up to 50 $\mu\text{g}/\text{kg}/\text{d}$. The mechanism underlying the granulocytic differentiation defects in bone marrow HSPCs of CN patients has only been partially elucidated. Among the relevant factors, we have identified deregulated levels of LEF-1 [13,14], C/EBP α [13], and PU.1 [15,16] transcription factors; hyperactivated JAK2 [17] and phospho-STAT5a [18]; elevated NAMPT/SIRT1 signaling [19]; abrogated expression of the anti-apoptotic genes Bcl2 and Bcl-xl [20], and markedly diminished expression of the natural inhibitor of neutrophil proteases SLPI (secretory leukocyte protease inhibitor) [21]. In addition, we and others have shown that mutant neutrophil elastase (NE) triggers activation of the unfolded protein response (UPR) and induction of endoplasmic reticulum (ER) stress caused by accumulation of altered (incompletely

Offprint requests to: Prof. Julia Skokowa, Department of Oncology, Hematology, Immunology, Rheumatology and Pulmonology, University Hospital Tuebingen, Otfried-Müller-Str. 10, Tübingen D-72076, Germany; E-mail: julia.skokowa@med.uni-tuebingen.de

folded or misfolded) NE protein within the ER or by disturbed intracellular trafficking of NE [22–24].

Despite these insights, we are still far from a clear understanding of the ultimate origin of the defective granulopoiesis and leukemic transformation in CN. There are no animal models of CN, except for rare neutropenia cases caused by inherited mutations in *GFII* (growth factor independent 1). *Elane*^{-/-} and *Hax1*^{-/-} mice, as well as transgenic mice with a knock-in of a human *ELANE* mutation [23], exhibit normal hematopoietic phenotypes and are not neutropenic. Reprogramming somatic cells of CN patients into induced pluripotent stem cells (iPSCs), followed by hematopoietic differentiation of iPSCs, provides a means for establishing an in vitro model of neutropenia and leukemic transformation in CN. iPSC hematopoietic differentiation models have limitations and cannot fully replace in vivo mouse disease models, but they represent an excellent source of immature hematopoietic cells and mature myeloid cells for further experimentation. Generation of CN-patient-specific iPSCs that recapitulate the maturation arrest of granulopoiesis has been described previously [25–28].

In the present study, we describe the establishment of an experimental in vitro model for studying CN using patient-derived iPSCs. Using this model, we were able to identify upregulation of key components of the UPR pathway and enhanced DNA damage and increased p21 protein levels in CD34⁺ cells and CD45⁺ cells of CN patients.

Methods

iPSC cell culture

iPSCs were maintained on mitomycin-C-treated SNL-feeder cells (Public Health England, GB) in iPSC medium consisting of DMEM F12 (Sigma-Aldrich, Germany) supplemented with 20% Knockout Serum Replacement (Invitrogen, USA), 30 ng/mL basic fibroblast growth factor (bFGF; Peprotech, USA), 1% nonessential amino acid solution (Invitrogen, USA), 100 μmol/L 2-mercapto-ethanol, and 2 mmol/L L-glutamine. iPSC medium was replaced every day. hiPSCs were subcultured by manual colony picking on new SNL feeder cells every 10 days.

Reprogramming of peripheral blood mononuclear cells

A total of 1.5×10^6 peripheral blood mononuclear cells were cultured after thawing for 6 days in CD34⁺ cell expansion medium (Stemline II Medium, Sigma-Aldrich, Germany) supplemented with 10% fetal calf serum, 1% penicillin/streptomycin, 1% glutamine, and the following cytokines: interleukin-3 (IL-3; 20 ng/μL), IL-6 (20 ng/μL), thrombopoietin (20 ng/μL), stem cell factor (SCF; 50 ng/μL), and FLT3L (50 ng/μL). All cytokines were purchased from R&D Systems (USA). After 1 week, cells were added to Retronectin (Clontech, USA)-coated 12-well plates together with OSKM lentiviral supernatant (pRRL.PPT.SF.hOct34.hKlf4.

hSox2.i2dTomato.pre.FRT, provided by A. Schambach) with multiplicity of infection of 2. Four days later, cells were transferred to SNL feeders and cultured in an 1:1 mixture of iPSC medium and CD34⁺ cell expansion medium supplemented with 2 mmol/L valproic acid and 50 μg/mL vitamin C. Medium was gradually changed to iPSC medium only. The first iPSC colonies appeared approximately 3 weeks after initiation of reprogramming.

Quantitative reverse transcription polymerase chain reaction

For quantitative reverse transcription polymerase chain reaction (qRT-PCR), RNA was isolated using RNeasy Micro Kit (Qiagen, Germany). cDNA was prepared from 1 μg of total RNA using the Omniscript RT Kit (Qiagen, Germany). qRT-PCR was performed using SYBR Green qPCR Master Mix (Roche, Switzerland) on Light Cycler 480 (Roche). Data were analyzed using the ddCT method. Target genes were normalized to *GAPDH* and *ACTB* as housekeeper genes. qRT-PCR primer sequences are listed in Supplementary Table E1 (online only, available at www.exphem.org).

Western blotting

A total of 1×10^6 cells were lysed in 200 μL of 3 × Laemmli buffer. Protein was denatured for 10 min at 95°C. Then, 5 μL of cell lysate in Laemmli buffer was loaded per lane. Proteins were separated on a 12% polyacrylamide gel and transferred on a nitrocellulose membrane (GE Healthcare, USA) for 1 hour at 100 V and 4°C. Membrane was blocked for 1 hour in 5% bovine serum albumin (BSA)/Tris-buffered saline + Tween 20 and incubated in primary antibody overnight at 4°C. The following primary antibodies were used: anti-p21 (Cell Signaling Technology, #2947S), anti-CHOP (Cell Signaling Technology, #2895S), and anti-NE (Santa Cruz Biotechnology, sc-55549) and anti-β-actin (Cell Signaling Technology, #13E5). Next, membranes were washed and incubated with secondary horseradish peroxidase-coupled (Santa Cruz Biotechnology) antibody for 1 hour at room temperature. Enhanced chemoluminescence solution (Thermo Fisher Scientific, USA) and Amersham Hyperfilm were used to detect chemiluminescence signal of proteins.

Embryoid-body-based hematopoietic differentiation of iPSCs

iPSCs were dissociated from SNL feeders or Matrigel (Corning, USA)-coated plates using phosphate-buffered saline (PBS)/ethylenediaminetetraacetic acid (0.02%) for 5 min. Embryoid body (EB) generation was done via centrifugation of 20,000 cells per EB in 96-well plates using APEL serum-free differentiation medium (StemCell Technologies, Canada) supplemented with bFGF (20 ng/μL) and Rho kinase (ROCK) inhibitor (10 μmol/L) (R&D Systems). The next day, bone morphogenetic protein 4 (BMP4; 20 ng/μL) was added to the culture to induce mesodermal differentiation. On day four, EBs were plated on Matrigel-coated six-well plates (10 EBs/well) in APEL medium supplemented with vascular endothelial growth factor (VEGF; 40 ng/μL), SCF (50 ng/μL), and IL-3 (50 ng/μL). For neutrophilic differentiation, medium was changed 3 days later to fresh APEL medium supplemented with IL-3 (50 ng/μL) and G-CSF (50 ng/μL). The first hematopoietic suspension cells appeared on day 12 to day 14.

Suspension cells were harvested every 3–4 days and analyzed starting from day 14 to day 32. All cytokines were purchased from R&D Systems if not otherwise indicated.

Flow cytometry

A total of 30,000 suspension cells collected from an EB-based hematopoietic differentiation system were used for flow cytometry. For cell surface staining, cells were prepared in PBS/1% BSA containing 0.05% sodium azide and stained with the mouse monoclonal antihuman antibody. For detection of mature myeloid cells, a multicolor fluorescence-activated cell-sorting (FACS) antibody panel for “late-stage” hematopoietic differentiation using the following antibodies was applied: CD15-PE, CD16-FITC, CD14-APC-H7, CD45-BV510, CD33-BV421, CD11b-PE-Cy7, and 7-AAD. Samples were analyzed using a FACSCanto II (BD Biosciences, Germany) and FlowJo version 10 software. Antibodies for flow cytometry were purchased from BD Biosciences if not otherwise indicated.

Morphological analysis

Wright–Giemsa-stained cytospin slides were prepared using Hema-Tek slide stainer (LabX, Canada). Hematopoietic cells were classified into four groups according to the differentiation state: myeloblast and promyelocyte (MB/ProM), myelocyte and metamyelocyte (Myelo/Meta), band and segmented neutrophils (Band/Seg), and monocytes/macrophages (Mo/MΦ).

DNA damage quantification

A total of 2.5×10^4 iPSC-derived CD34⁺/CD45⁺ cells were resuspended in PBS and treated with the indicated concentrations of bleomycin for 5 min at room temperature. Genomic DNA was isolated using the QIAamp DNA Mini Kit (Qiagen, Germany). Long-run real-time PCR-based DNA-damage quantification (LORD-Q) was performed and analyzed according to Lehle et al. [29].

Statistical analysis

Differences in mean values between groups were analyzed using two-sided, unpaired Student *t* tests using the SPSS (IBM, USA) version 9.0 statistical package.

Results

In vitro EB-based iPSC differentiation model reveals severely impaired myelopoiesis of CN and CN/acute myeloid leukemia patient-derived iPSCs

We generated iPSCs from two *ELANE*-CN patients, one patient with p.C151Y (CN1), and another patient with p.G214R (CN2) *ELANE* mutations. From one CN patient, we were able to generate iPSC clones harboring either an *ELANE* p.C151Y mutation only (CN-iPSC clone) or additional *CSF3R* and *RUNX1* mutations and trisomy 21 (CN/acute myeloid leukemia [AML]-iPSC clone). All iPSC lines expressed elevated protein and mRNA levels of pluripotent stem-cell-specific factors and showed inactivation of the lentiviral plasmid used for reprogramming.

To evaluate myeloid differentiation of *ELANE*-CN patients iPSCs, we slightly modified an in vitro EB-based iPSC differentiation method developed by Lachmann et al. [30] that allows generation of hematopoietic cells and mature myeloid cells for more than 30 days. EB formation was performed by first centrifuging dissociated iPSCs (20,000 cells/EB) in 96-well conical-bottomed plates in APEL serum-free differentiation medium containing bFGF and ROCK inhibitors, followed by induction of mesodermal differentiation by the addition of BMP4 on day 1 of culture. To induce hematopoietic differentiation, we plated EBs on Matrigel-coated six-well plates (10 EBs/well) in APEL medium supplemented with VEGF, SCF, and IL-3 on day 4 of culture. Neutrophilic differentiation was initiated 3 days later by replacement of cytokines with IL-3 and G-CSF (Figure 1A). In EBs cultured from healthy donor (HD)-derived iPSCs, hematopoietic cells appeared in the supernatants on day 14 and the number of cells was markedly increased at day 28 of culture (Figure 1B). In contrast, the number of hematopoietic cells in culture supernatants of EBs from CN patient-derived iPSCs was markedly diminished and almost no myeloid cells were produced from cultured CN/AML iPSCs (Figure 1B). Consistent with these observations, granulocytic differentiation into mature polymorphonuclear neutrophils (CD15⁺CD16⁺CD45⁺ cell population) was markedly reduced in iPSC lines of both *ELANE*-CN patients compared with HD-derived iPSCs and was abolished in CN/AML-iPSCs (Figure 1C). However, the number of monocytes (CD14⁺ CD11b⁺CD45⁺ cells) produced by CN patient-derived iPSCs was strongly increased compared with that produced by HD-derived iPSCs (Figure 1C). This pattern recapitulates the situation in CN patients, in whom monocyte numbers in the peripheral blood are elevated [31]. Morphological examinations of Giemsa-stained cytospin slides confirmed these results (Figure 1D). Therefore these observations demonstrate the successful establishment of an in vitro model to evaluate mechanisms of diminished granulocytic differentiation of hematopoietic cells in CN patients.

Elevated ER stress and UPR in CD34⁺ and CD45⁺ cells from CN patient- and CN1/AML patient-derived iPSCs

We next investigated whether intracellular signaling pathways operating during induction of the UPR, which is known to be hyperactivated in primary HSPCs and granulocytes of CN patients [21–24], are also affected in iPSC-derived hematopoietic cells. We found that *ELANE* mRNA and NE protein levels were highly upregulated in CD34⁺ cells generated from CN1/AML-iPSCs compared with those generated from CN-iPSCs and HD-iPSCs (Figures 2A and 2C). We detected deregulation of mRNA expression of the UPR downstream targets in CN1 and CN2 iPSC-derived CD34⁺ cells. Therefore DDIT3 (DNA-

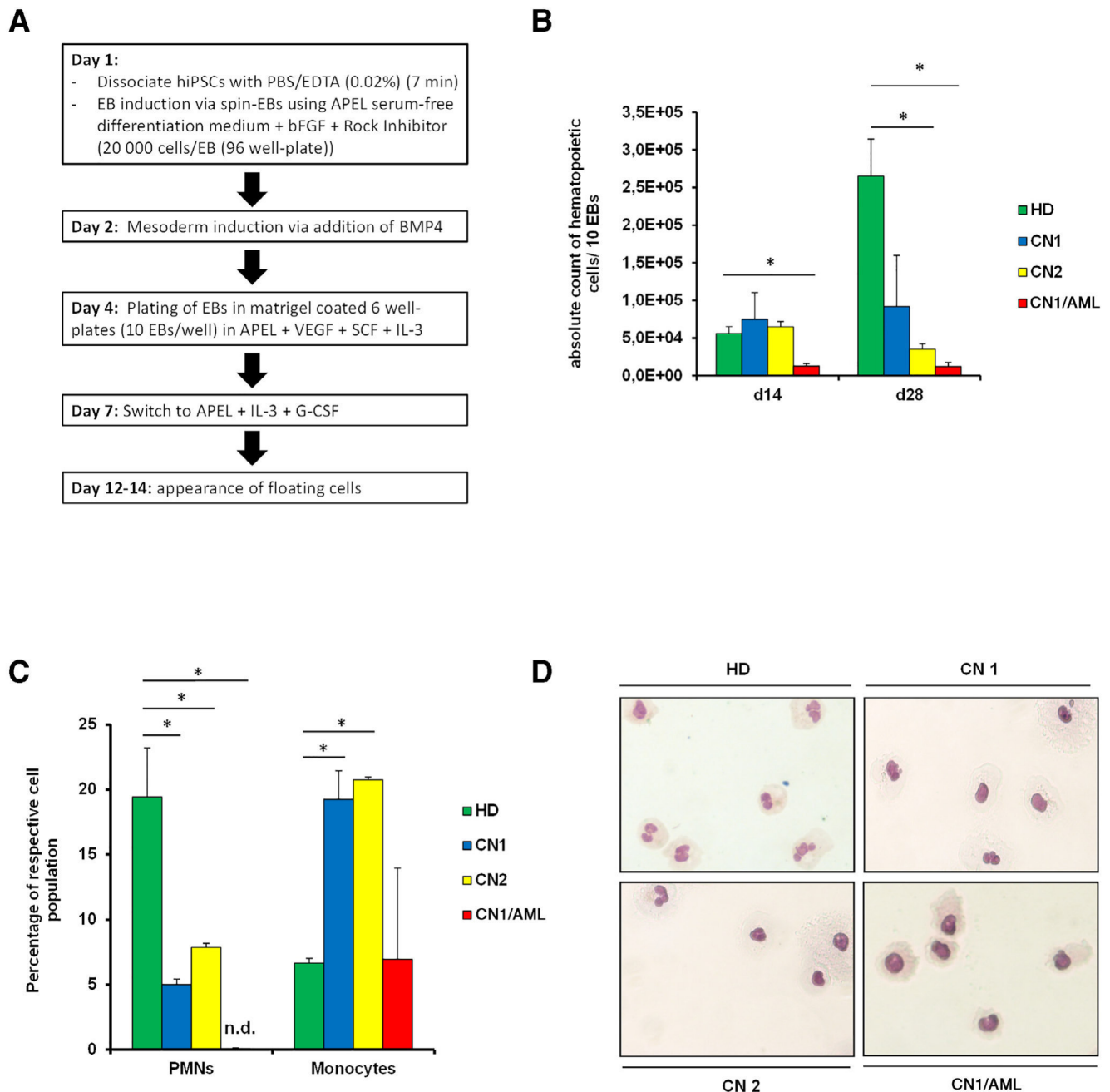


Figure 1. In vitro model of CN and CN/AML using EB-based myeloid differentiation method of patients iPSCs. **(A)** Scheme of the protocol for EB-based hematopoietic and neutrophilic differentiation of iPSCs. **(B)** Production of hematopoietic cells from iPSCs over time in the EB-based differentiation system. Hematopoietic cells were harvested from EB culture supernatants starting from day 14 to day 28 and counted using trypan blue dye exclusion. Data represent means \pm SD from two independent experiments. Two-sided, unpaired Student *t* test *p* values to HD are shown. $*p < 0.05$. **(C)** Flow cytometry analysis of suspension cells harvested from EBs culture on day 28 of differentiation. Data represent means \pm SD from two independent experiments. $*p < 0.05$. **(D)** Morphological analysis of suspension cells harvested from iPSCs at day 28 of differentiation (Wright-Giemsa Stain). Representative cytopsin slide pictures are shown. HD, healthy donor.

damage inducible transcript 3, also called CHOP) was increased in CD34⁺ cells generated from CN1-iPSCs, CN2-iPSCs, and CN1/AML-iPSCs compared with those from HD-iPSCs (Figure 2B). In contrast, expression of mRNA for BiP (binding immunoglobulin protein) was upregulated in CN2, but not CN1-iPSC- or CN1/AML-iPSC-derived

CD34⁺ cells. ATF6 (activating transcription factor 6) mRNA was highly expressed in CD34⁺ cells derived from CN1-iPSCs, but not in CN2-iPSC- or CN1/AML-iPSC-derived cells. In contrast, maximum ATF4 (activating transcription factor 4) mRNA expression was detected in CD34⁺ cells from CN2-iPSCs and CN1/AML-iPSCs, but

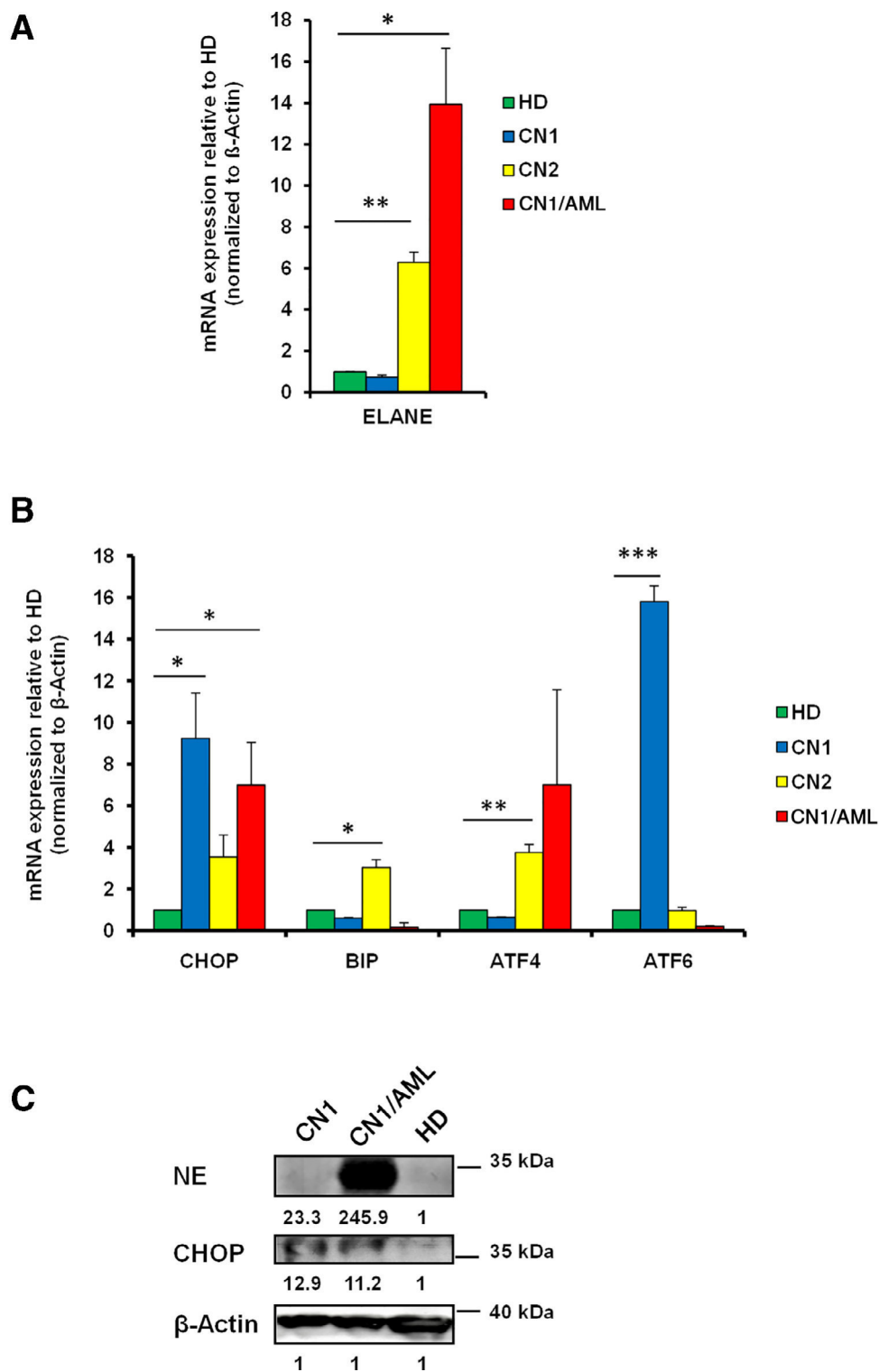


Figure 2. Analysis of ER stress and unfolded protein response (UPR) in CD34⁺ and CD45⁺ cells derived from CN- and CN/AML-iPSCs. **(A)** qRT-PCR analysis of ELANE mRNA expression in CD45⁺CD34⁺ cells at day 14 of iPSC differentiation. Data represent means ± SD from two independent experiments. **p* < 0.05, ***p* < 0.001. **(B)** qRT-PCR analysis of mRNA expression of UPR-related genes in CD45⁺CD34⁺ cells at day 14 of iPSC differentiation, as indicated. Data represent means ± SD from two independent experiments. **p* < 0.05, ***p* < 0.01, ****p* < 0.001. **(C)** Representative Western blot images of NE and CHOP protein expression in CD45⁺ cells at day 28 of iPSC differentiation, as indicated. Numbers below Western blot images indicate protein expression levels normalized to β-Actin.

not in those from CN1-iPSCs or HD-iPSCs (Figure 2B). CHOP protein levels were also elevated in CD45⁺ cells derived on day 28 of EB differentiation of CN and CN/AML iPSCs (Figure 2C).

DNA damage and DNA-repair responses in CD34⁺ cells from CN-iPSCs

We further assessed whether elevated ER stress resulting from UPR sensitizes CN-iPSC-derived CD34⁺ cells to DNA damage. DNA damage was induced by treating cells with bleomycin for 5 min, after which bleomycin was removed and DNA damage and DNA repair were assessed immediately and after 2 hours of incubation (Figure 3A). Long-run, real-time PCR-based DNA-damage quantification (LORD-Q [29]) of genomic DNA loci (*GAPDH* and *TP53*) and mitochondrial DNA (mtDNA) revealed a robust increase in mtDNA and nuclear DNA lesions in CN-iPSC-derived CD34⁺ cells compared with HD-iPSC-derived cells (Figure 3B). Damaged mtDNA and *GAPDH* DNA loci, but not *TP53* DNA loci, were also more frequent in CD34⁺ cells from CN-iPSCs 2 hours after bleomycin treatment compared with that in HD-iPSC-derived CD34⁺ cells. These data suggest delayed DNA repair and increased susceptibility to DNA damage in CD34⁺ cells from CN-iPSC lines (Figure 3C).

p21 upregulation in CD34⁺ and CD45⁺ cells from CN-iPSCs and CN/AML-iPSCs

We next sought to determine whether DNA damage pathways are differentially regulated in CD34⁺ and CD45⁺ cells generated from iPSCs of CN or CN/AML patients compared with HD-generated cells. Our assessment of the role of the p53-p21 pathway, which is typically activated upon DNA damage, showed no differences in *TP53* mRNA expression, but revealed an approximately threefold increase in p21 mRNA expression in CN-iPSC-derived blood cells compared with those derived from HD cells and a fivefold to sixfold increase in CN/AML-iPSC-derived cells (Figure 3D). p21 protein levels were also elevated in CN/AML cells (Figure 3E). The expression of *MDM2* mRNA was markedly diminished in CN/AML cells, and *GADD45a* mRNA expression was slightly induced in both CN-iPSC- and CN/AML-iPSC-derived cells compared with those derived from HD-iPSCs (Figure 3D). Interestingly, p21 mRNA expression was also upregulated in primary bone marrow CD33⁺ cells of CN patients compared with those of G-CSF-treated healthy individuals, in which p21 levels were even suppressed by G-CSF (data not shown).

Discussion

In the present study, we established an in vitro model of CN using patient-derived iPSCs. We also

successfully reprogrammed AML blasts from one CN/AML patient and were able to compare hematopoietic and myeloid differentiation of iPSCs derived from CN/AML cells, CN patients, and a healthy donor. Our data provide strong evidence that, despite some limitations, iPSCs represent a valuable resource for disease modeling, especially for investigations on inherited bone marrow failure syndromes. Primary bone marrow material from pediatric patients with bone marrow failure is extremely limited and mouse models are not available for many of these syndromes. In this latter context, *elane*^{-/-} mice, *hax1*^{-/-} mice, and transgenic mice carrying mutated *elane* do not exhibit neutropenia. In addition, transgenic mice carrying a truncated G-CSF receptor (*csf3r*) mutant never develop leukemia. These acquired *CSF3R* mutations are observed in leukemic blasts of more than 80% of CN patients with overt AML or MDS [32].

Using our iPSC model, we were able to recapitulate the hematopoietic and myeloid differentiation defects of HSPCs observed in CN patients in vivo. Specifically, we detected diminished granulocytic differentiation of CN-iPSCs, an observation consistent with the induction of UPR. iPSC lines from two CN patients showed different behavior during first 2 weeks of hematopoietic differentiation, when differentiation of EBs into CD34⁺ CD45⁺ cells occurs, but both CN patients' iPSC lines demonstrated similar markedly diminished granulocytic differentiation at later stages. These early differentiation stage differences may be explained by the varying effects of the mutated NE protein on blood cell formation, which is dependent on the mutated amino acid residues. Interestingly, different *ELANE* mutations resulted in abolished granulocytic differentiation of CN patient-derived iPSCs, an observation consistent with insights gained from CN patients [1]. Monocytic differentiation was elevated in CN-iPSC lines, an observation that parallels the common finding of peripheral blood monocytosis in CN patients [1,31]. One possible explanation for this monocytosis is a compensatory reaction of the bone marrow to diminished neutrophil counts and function that serves to induce an immune response to bacterial pathogens. An alternative explanation is deregulated expression of lineage-specific (granulocyte-specific vs. monocyte-specific) transcription factors in myeloid progenitor cells of CN patients. Deregulated expression of relevant transcription factors (e.g., diminished expression of LEF-1 and C/EBP α , but elevated PU.1 expression) has been described by us previously [15]. Elevated monocytic maturation of CN patients' iPSC lines in vitro further supports the theory that deregulation of a transcriptional program in HSPCs is a cause of neutropenia and monocytosis.

UPR hyperactivation in primary HSPCs of CN patients has been described previously [22–24]. It is presumed that

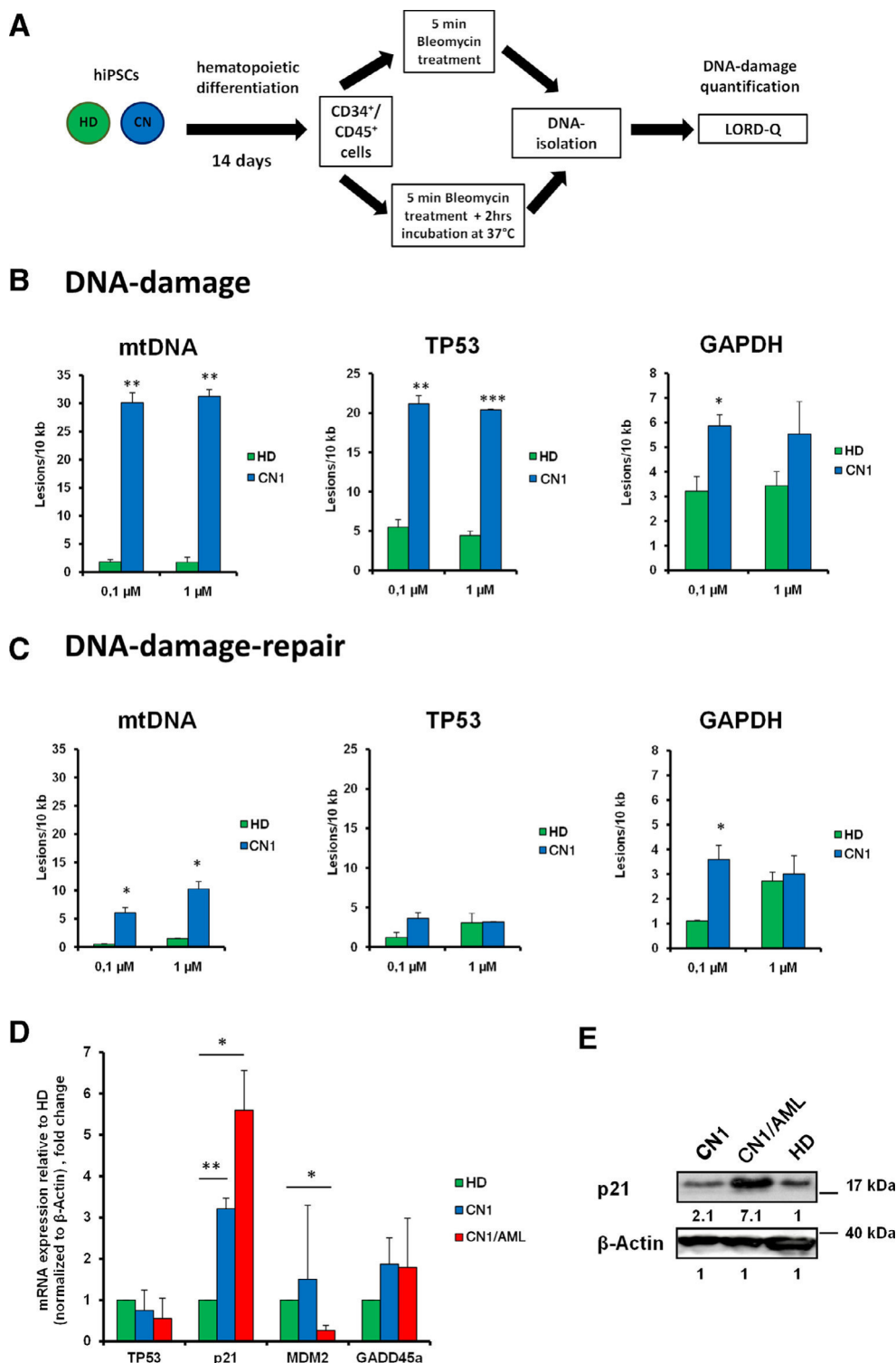


Figure 3. Quantification of DNA damage and p21 expression in CD34⁺ and CD45⁺ cells derived from patient-specific iPSCs. **(A)** Scheme for DNA damage measurements in HD- and CN-iPSCs derived CD34⁺CD45⁺ cells. **(B,C)** Measurement of DNA damage loci for mitochondrial DNA (mtDNA) and DNA of *TP53* or *GAPDH* gene loci of CN-iPSC-derived CD45⁺CD34⁺ cells upon 5-min treatment with 0.1 and 1 μmol/L bleomycin. DNA lesions were assessed directly 5 min (DNA damage **(B)**) and 2 hours (DNA repair **(C)**) after bleomycin treatment. Data are shown as means ± SD from two independent experiments. **p* < 0.05, ***p* < 0.01, ****p* < 0.001. **(D)** qRT-PCR analysis of selected genes in CD45⁺CD34⁺ cells generated from different iPSCs clones on day 14 of iPSC differentiation. mRNA expression of target genes was normalized to β-actin and shown relative to HD. Data represent means ± SD from two independent experiments. **p* < 0.05, ***p* < 0.01, ****p* < 0.001. **(E)** Representative Western blot images of p21 protein and β-actin expression in CD45⁺ cells at day 28 of iPSC differentiation. Numbers below the Western blot images indicate protein expression levels normalized to β-actin.

proper folding and intracellular localization of mutated NE is severely affected in myeloid cells of *ELANE*-CN patients. These defects ultimately lead to activation of the UPR and ER stress. We detected no activation of ATF6 and BiP in CN/AML cells compared with CN cells. These differences may be attributable to a dosage effect of mutated NE or additional coregulation of ATF4, but not ATF6, by mutated *RUNX1* and trisomy 21. Interestingly, elevated expression of ATF4 is associated with resistance to current chemotherapeutic drugs [33] and we detected hyperactivation of ATF4 in CN/AML cells. We have recently reported activation of different UPR pathways depending on the type of *ELANE* mutation [22]. Consistent with this, we demonstrate here that *ELANE* mutant p.C151Y induces expression of ATF4, ATF6, and CHOP, but not BiP. At the same time, p.G214R *ELANE* mutation caused upregulation of BiP, ATF4, and CHOP, but not ATF6. The impact of inherited CN-associated mutations (e.g., in *ELANE*) and UPR activation on leukemic progression is still unclear. It has been shown that induction of ER stress protects gastric cancer cells against apoptosis [34]. It is also known that activation of the UPR remodels the sensitivity of tumor cells to chemotherapeutic agents, making them more sensitive in some cases and more resistant in others [35]. Interestingly, upon ER stress, ATF4 is activated in mouse HSPCs, but not in more committed progenitors, leading to apoptosis of HSPCs [36]. Elevated levels of mutated NE, which we detected in CD45⁺ cells derived from CN-iPSCs and CN/AML-iPSCs, may further amplify the UPR, and additional signaling pathways (e.g., hyperactivated STAT5a [18] or mutated *RUNX1* [37]) may protect these cells from apoptosis, resulting in leukemogenic transformation of these cells. Because *ELANE* expression is regulated by *RUNX1* [38], it will be interesting to determine how missense *RUNX1* mutations affect *ELANE* expression. Cai et al. demonstrated an attenuated UPR in HSPCs from *Runx1*^{-/-} mice [39] and another study showed induction of UPR by trisomy 21 in immortalized lymphocytes and fibroblasts of Down syndrome patients [40]. The induction of UPR and ER stress in HSPCs of these patients has not yet been studied, but it is known that individuals with Down syndrome often develop AML.

We identified an increased susceptibility of CN CD34⁺ CD45⁺ cells to DNA damage, a finding consistent with the observed prolongation in DNA repair. Interestingly, ER stress suppresses DNA double-strand break repair in tumor cells [41]. Moreover, Nagelkerle et al. demonstrated that the UPR increases the resistance of tumor cells to therapeutic agents by regulating the DNA damage response [42]. In CN patients with inherited *ELANE* mutations, a permanent stress response caused by a chronically elevated UPR may cause genotoxic stress, increasing the susceptibility of HSPCs to secondary leukemia-causing events, although the

complete pathomechanism remains to be investigated. An evaluation of the expression levels of the main players in p53 signaling, a classical DNA damage response pathway, revealed strong upregulation of p21 mRNA levels, but not p53 mRNA expression, in CN and CN/AML CD34⁺ cells. p21 protein expression was also elevated in CD45⁺ cells of CN/AML patients compared with CN and HD-derived cells. The mechanism of p21 upregulation in CN and CN/AML cells remains to be investigated, but may be explained by increased p53 protein stability or diminished levels of the p53 ubiquitin E3 ligase MDM2. Another possibility is p53-independent activation of p21 expression. In this context, Galanos et al. reported p53-independent upregulation of p21 selectively in more aggressive tumor cells, which featured increased genomic instability, aggressiveness, and chemoresistance [43]. Unaffected expression of GADD45a, another p53 target, also argues for possibly inactive p53 and thus a p53-independent mechanism of p21 activation in CD34⁺ cells of CN patients. Less is known about UPR-mediated regulation of p21 expression. One study described inhibition of p21 expression by CHOP [44], but we found that both CHOP and p21 levels were elevated in CN and CN/AML CD34⁺ and CD45⁺ cells. It is known that accumulated cytoplasmic p21 exerts anti-apoptotic functions in AML cells [45,46] and high p21 levels are associated with chemoresistance in AML [47]. Elevated expression of p21 in preleukemic cells of CN patients may make these cells resistant to apoptosis, leading to an increase in their survival and an increased probability of their leukemogenic transformation. AML blasts of CN patients are resistant to conventional chemotherapy, so bone marrow transplantation is the only treatment option in these patients [1].

Taken together, our findings demonstrate that the iPSC-based model established here is a reliable in vitro model for studying defective signaling systems underlying impaired hematopoietic differentiation in patients with bone marrow failure syndromes (in our case, CN). This model may also be used for drug development or generation of isogenic iPSC lines using CRISPR/Cas9-mediated correction of inherited disease-causing mutations.

Acknowledgments

This work was supported by the Excellence Initiative of the Faculty of Medicine, University of Tuebingen (JS), the Jose Carreras Leukemia Foundation (JS, BD), Madeleine Schickedanz Kinderkrebsstiftung (JS, MK), the Deutsche Forschungsgemeinschaft (JS, MK), the Medical Faculty of the Tübingen University (intramural Fortüne funding to M.K.), the German–Israeli Foundation for Scientific Research and Development (KW, AZ, BD), the German Cancer Consortium (KS-O, JS, PM), and the Fritz Thyssen Foundation (B.D.).

Author contributions

JS and BD made initial observations, designed the experiments, and analyzed the data; BD performed the main experiments, generated, characterized, cultured, and differentiated human iPSCs, and performed FACS, qRT-PCRs, and WBs; BO differentiated iPSCs and performed FACS; PM conducted LORD-Q experiments; AZ with the help of NL and TM established iPSC generation and EB-based hematopoietic differentiation in the laboratory; RB assisted in iPSC culture and qRT-PCR; KS-O assisted with LORD-Q experiments and provided insightful comments; CZ and KW provided patient material; LK provided insightful comments; and KW and JS supervised and supported the study and wrote the manuscript (with the help of BD).

References

- Skokowa J, Dale DC, Touw IP, Zeidler C, Welte K. Severe congenital neutropenias. *Nat Rev Dis Primers*. 2017;3:17032.
- Skokowa J, Germeshausen M, Zeidler C, Welte K. Severe congenital neutropenia: inheritance and pathophysiology. *Curr Opin Hematol*. 2007;14:22–28.
- Dale DC, Person RE, Bolyard AA, et al. Mutations in the gene encoding neutrophil elastase in congenital and cyclic neutropenia. *Blood*. 2000;96:2317–2322.
- Klein C, Grudzien M, Appaswamy G, et al. HAX1 deficiency causes autosomal recessive severe congenital neutropenia (Kostmann disease). *Nat Genet*. 2007;39:86–92.
- Triot A, Järvinen PM, Arostegui JI, et al. Inherited biallelic CSF3R mutations in severe congenital neutropenia. *Blood*. 2014;123:3811–3817.
- Klimiankou M, Klimenkova O, Uenal M, et al. GM-CSF stimulates granulopoiesis in a congenital neutropenia patient with loss-of-function biallelic heterozygous CSF3R mutations. *Blood*. 2015;126:1865–1867.
- Boztug K, Järvinen PM, Salzer E, et al. JAGN1 deficiency causes aberrant myeloid cell homeostasis and congenital neutropenia. *Nat Genet*. 2014;46:1021–1027.
- Boztug K, Appaswamy G, Ashikov A, et al. A syndrome with congenital neutropenia and mutations in G6PC3. *N Engl J Med*. 2009;360:32–43.
- Makaryan V, Rosenthal EA, Bolyard AA, et al. UW Center for Mendelian Genomics. TCIRG1-associated congenital neutropenia. *Hum Mutat*. 2014;35:824–827.
- Makaryan V, Zeidler C, Bolyard AA, et al. The diversity of mutations and clinical outcomes for ELANE-associated neutropenia. *Curr Opin Hematol*. 2015;22:3–11.
- Welte K, Dale D. Pathophysiology and treatment of severe chronic neutropenia. *Ann Hematol*. 1996;72:158–165.
- Welte K, Zeidler C, Dale DC. Severe congenital neutropenia. *Semin Hematol*. 2006;43:189–195.
- Skokowa J, Cario G, Uenal M, et al. LEF-1 is crucial for neutrophil granulopoiesis and its expression is severely reduced in congenital neutropenia. *Nat Med*. 2006;12:1191–1197.
- Skokowa J, Welte K. LEF-1 is a decisive transcription factor in neutrophil granulopoiesis. *Ann N Y Acad Sci*. 2007;1106:143–151.
- Skokowa J, Welte K. Dysregulation of myeloid-specific transcription factors in congenital neutropenia. *Ann N Y Acad Sci*. 2009;1176:94–100.
- Skokowa J, Welte K. Defective G-CSFR signaling pathways in congenital neutropenia. *Hematol Oncol Clin North Am*. 2013;27:75–88. viii.
- Rauprich P, Kasper B, Tidow N, Welte K. The protein tyrosine kinase JAK2 is activated in neutrophils from patients with severe congenital neutropenia. *Blood*. 1995;86:4500–4505.
- Gupta K, Kuznetsova I, Klimenkova O, et al. Bortezomib inhibits STAT5-dependent degradation of LEF-1, inducing granulocytic differentiation in congenital neutropenia CD34(+) cells. *Blood*. 2014;123:2550–2561.
- Skokowa J, Lan D, Thakur BK, et al. NAMPT is essential for the G-CSF-induced myeloid differentiation via a NAD(+)-sirtuin-1-dependent pathway. *Nat Med*. 2009;15:151–158.
- Cario G, Skokowa J, Wang Z, et al. Heterogeneous expression pattern of pro- and anti-apoptotic factors in myeloid progenitor cells of patients with severe congenital neutropenia treated with granulocyte colony-stimulating factor. *Br J Haematol*. 2005;129:275–278.
- Klimenkova O, Ellerbeck W, Klimiankou M, et al. A lack of secretory leukocyte protease inhibitor (SLPI) causes defects in granulocytic differentiation. *Blood*. 2014;123(8):1239–1249.
- Nustede R, Klimiankou M, Klimenkova O, et al. ELANE mutant-specific activation of different UPR pathways in congenital neutropenia. *Br J Haematol*. 2016;172:219–227.
- Nanua S, Murakami M, Xia J, et al. Activation of the unfolded protein response is associated with impaired granulopoiesis in transgenic mice expressing mutant Elane. *Blood*. 2011;117:3539–3547.
- Grenda DS, Murakami M, Ghatak J, et al. Mutations of the ELA2 gene found in patients with severe congenital neutropenia induce the unfolded protein response and cellular apoptosis. *Blood*. 2007;110:4179–4187.
- Morishima T, Watanabe K, Niwa A, et al. Genetic correction of HAX1 in induced pluripotent stem cells from a patient with severe congenital neutropenia improves defective granulopoiesis. *Haematologica*. 2014;99:19–27.
- Nayak RC, Trump LR, Aronow BJ, et al. Pathogenesis of ELANE-mutant severe neutropenia revealed by induced pluripotent stem cells. *J Clin Invest*. 2015;125:3103–3116.
- Hiramoto T, Ebihara Y, Mizoguchi Y, et al. Wnt3a stimulates maturation of impaired neutrophils developed from severe congenital neutropenia patient-derived pluripotent stem cells. *Proc Natl Acad Sci U S A*. 2013;110:3023–3028.
- Pittermann E, Lachmann N, MacLean G, et al. Gene correction of HAX1 reversed Kostmann disease phenotype in patient-specific induced pluripotent stem cells. *Blood Adv*. 2017;1:903–914.
- Lehle S, Hildebrand DG, Merz B, et al. LORD-Q: a long-run real-time PCR-based DNA-damage quantification method for nuclear and mitochondrial genome analysis. *Nucleic Acids Res*. 2014;42:e41.
- Lachmann N, Ackermann M, Frenzel E, et al. Large-scale hematopoietic differentiation of human induced pluripotent stem cells provides granulocytes or macrophages for cell replacement therapies. *Stem Cell Reports*. 2015;4:282–296.
- Bonilla MA, Dale D, Zeidler C, et al. Long-term safety of treatment with recombinant human granulocyte colony-stimulating factor (r-metHuG-CSF) in patients with severe congenital neutropenias. *Br J Haematol*. 1994;88:723–730.
- Germeshausen M, Skokowa J, Ballmaier M, Zeidler C, Welte K. G-CSF receptor mutations in patients with congenital neutropenia. *Curr Opin Hematol*. 2008;15:332–337.
- Ye J, Koumenis C. ATF4, an ER stress and hypoxia-inducible transcription factor and its potential role in hypoxia tolerance and tumorigenesis. *Curr Mol Med*. 2009;9:411–416.
- Feng R, Zhai WL, Yang HY, Jin H, Zhang QX. Induction of ER stress protects gastric cancer cells against apoptosis induced by cisplatin and doxorubicin through activation of p38 MAPK. *Biochem Biophys Res Commun*. 2011;406:299–304.
- Mann MJ, Hendershot LM. UPR activation alters chemosensitivity of tumor cells. *Cancer Biol Ther*. 2006;5:736–740.

36. van Galen P, Kreso A, Mbong N, et al. The unfolded protein response governs integrity of the haematopoietic stem-cell pool during stress. *Nature*. 2014;510:268–272.
37. Skokowa J, Steinemann D, Katsman-Kuipers JE, et al. Cooperativity of RUNX1 and CSF3R mutations in severe congenital neutropenia: a unique pathway in myeloid leukemogenesis. *Blood*. 2014;123:2229–2237.
38. Lausen J, Liu S, Fliegauf M, Lübbert M, Werner MH. ELA2 is regulated by hematopoietic transcription factors, but not repressed by AML1-ETO. *Oncogene*. 2006;25:1349–1357.
39. Cai X, Gao L, Teng L, et al. Runx1 deficiency decreases ribosome biogenesis and confers stress resistance to hematopoietic stem and progenitor cells. *Cell Stem Cell*. 2015;17:165–177.
40. Aivazidis S, Coughlan CM, Rauniyar AK, et al. The burden of trisomy 21 disrupts the proteostasis network in Down syndrome. *PLoS One*. 2017;12:e0176307.
41. Yamamori T, Meike S, Nagane M, Yasui H, Inanami O. ER stress suppresses DNA double-strand break repair and sensitizes tumor cells to ionizing radiation by stimulating proteasomal degradation of Rad51. *FEBS Lett*. 2013;587:3348–3353.
42. Nagelkerke A, Bussink J, van der Kogel AJ, Sweep FC, Span PN. The PERK/ATF4/LAMP3-arm of the unfolded protein response affects radioresistance by interfering with the DNA damage response. *Radiother Oncol*. 2013;108:415–421.
43. Galanos P, Vougas K, Walter D, et al. Chronic p53-independent p21 expression causes genomic instability by deregulating replication licensing. *Nat Cell Biol*. 2016;18:777–789.
44. Mihailidou C, Papazian I, Papavassiliou AG, Kiaris H. CHOP-dependent regulation of p21/waf1 during ER stress. *Cell Physiol Biochem*. 2010;25:761–766.
45. Schepers H, Geugien M, Eggen BJ, Vellenga E. Constitutive cytoplasmic localization of p21(Waf1/Cip1) affects the apoptotic process in monocytic leukaemia. *Leukemia*. 2003;17:2113–2121.
46. Abbas T, Dutta A. p21 in cancer: intricate networks and multiple activities. *Nat Rev Cancer*. 2009;9:400–414.
47. Zhang W, Kornblau SM, Kobayashi T, Gambel A, Claxton D, Deisseroth AB. High levels of constitutive WAF1/Cip1 protein are associated with chemoresistance in acute myelogenous leukemia. *Clin Cancer Res*. 1995;1:1051–1057.

PUBLICATION B:

CRISPR/Cas9 Genome Editing of Human-Induced Pluripotent Stem Cells

Followed by Granulocytic Differentiation



CRISPR/Cas9 Genome Editing of Human-Induced Pluripotent Stem Cells Followed by Granulocytic Differentiation

Benjamin Dannenmann, Masoud Nasri, Karl Welte, and Julia Skokowa

Abstract

Research on patient-derived induced pluripotent stem cells (iPSCs) could immensely benefit from the implementation of CRISPR/Cas9 genome editing of iPSCs, creating unique opportunities such as the establishment of isogenic iPSC lines for disease modeling or personalized patient-specific drug screenings. Here we describe a stepwise protocol of safe, efficient, and selection-free CRISPR/Cas9-mediated gene correction or knockout in human iPSCs followed by 3D spin-embryoid body (EB)-based hematopoietic/neutrophilic iPSC-differentiation.

Key words CRISPR/Cas9, Human-induced pluripotent stem cells, Genome editing, Ribonucleoprotein, iPSC differentiation, Hematopoietic, and neutrophilic differentiation of iPSCs

1 Introduction

iPSC technology is a powerful tool for disease modeling, patient-specific drug screening, and paving the way for personalized medicine [1, 2]. Patient-derived iPSC cells can be used as an excellent tool to model the disease and reveal the key driving pathomechanisms behind it. An important challenge of iPSC technology is the high degree of variability between individual iPSC lines that are mainly due to differences in genetic background and reprogramming history [3]. Therefore, to have an optimal comparison in studies using iPSC, especially disease modeling, the role of the isogenic iPSC lines is indispensable. Combination of CRISPR/Cas9-based gene correction or knockout in human iPSCs (hiPSCs) derived from patients suffering from inherited or acquired hematopoietic disease with subsequent EB-based in vitro hematopoietic

Benjamin Dannenmann and Masoud Nasri contributed equally to this work.

differentiation is a valuable tool to delineate the role of the gene of interest or gene mutations on hematopoietic differentiation in a fast, efficient, and reproducible manner [4–6].

In this chapter, we first explain in detail our CRISPR/Cas9 gene-editing strategy of iPSCs which is safe, as no viral vectors or plasmids are used, efficient, and selection free. We next describe our 3D spin-embryoid body (EB)-based iPSCs-differentiation protocol, which has many advantages over standard 2D iPSCs-differentiation protocols. The main advantage is the high amounts of HSPCs and mature granulocytes derived from iPSCs (more than 1×10^6 cells/10 EBs) and improved functionality of mature cells [4]. Another advantage is that the present protocol is based on the use of STEMdiff APEL medium, a fully defined, serum-free, and animal component free-differentiation medium which is highly versatile and allows the differentiation of cells to multiple hematopoietic lineages [7, 8]. By adapting variable cytokine combinations, our strategy allows efficient generation of not only a high number of CD34⁺/CD45⁺ cells and mature granulocytes but also monocytes or megakaryocytes. Furthermore, the current protocol displays high reproducibility between different experiments.

2 Materials

The reagents describing below have been successfully used by our group, but there are other commercially available alternatives that can be used and optimized.

2.1 Transfection and Electroporation of iPSCs with Cas9 Ribonucleoprotein (RNP)

1. Guide RNAs for target sequence ordered as lyophilized single-guide RNA with or without 2'-O-methyl 3'phosphorothioate (Integrated DNA Technologies).
2. Alt-R S.p Cas9 Nuclease V3 or Alt-R S.p HiFi Cas9 Nuclease V3 (Integrated DNA Technologies).
3. Ultramer DNA oligos, 4 nm synthesis scale, and standard desalting as purification method (Integrated DNA Technologies).
4. *TransIT*-X2 transfection reagent (Mirus, Madison, USA).
5. Amaxa 4D-Nucleofector system and P3 primary Cell 4D-Nucleofector Kit (Lonza).

2.2 iPSCs Maintenance and EB-Based Myeloid Differentiation

1. StemFlex medium, RevitaCell supplement, and Geltrex LDEV-Free reduced growth factor basement membrane matrix (ThermoFisher Scientific).
2. STEMdiff APEL 2 Medium (Stem Cell Technologies).
3. ROCK inhibitor, FGF2, BMP4, VEGF, SCF, IL3, and G-CSF.

3 Methods

3.1 Design, Testing, and Validation of Guide RNA

1. Design guide RNAs using guide RNA designing web tools (e.g., CCTOP, Centre for Organismal Studies Heidelberg):
 - (a) For gene-correction purpose, select 4–6 guide RNAs in a window of 40 nt distance from the desired mutation site with good off-target activity (*see Note 1*).

The optimal window is 10 nt distance. Rank the guide RNAs based on the off-target activity profile.
 - (b) For gene knockout purpose, select 4–6 guide RNAs in the coding sequence of the first or second exon of the gene of choice. Consider all possible transcripts and isoforms for a successful knockout.
2. Cross check the guide RNAs list in the second guide RNA designing website (e.g., DESKGEN).
3. Order the list of guide RNAs as crRNA or sgRNA via provider of choice (e.g., Integrated DNA Technologies, Coralville, USA).
4. Upon receiving the guide RNAs, spin them down briefly and resuspend to the final concentration of 100 μM .
5. For validation of guide RNAs, seed 10^5 HEK 293T cells 1 day before transfection in 12-well plate.
6. On the day of transfection, assemble 0.24 μL from each single-guide RNA (100 μM Stock, 24 nM final concentration) with 0.2 μL Cas9 protein (62 μM Stock, 12 nM final concentration) in a sterile tube.
7. Add 100 μL of OptiMEM I Reduced Serum Medium to each tube and mix gently.
8. Add 2 μL *TransIT-X2* transfection reagent (*see Note 2*) to each tube, mix gently, and incubate at room temperature for 20 min.
9. Add the *TransIT-X2*: RNP complexes dropwise to the different areas of the well. Gently rock the plate forth and back and incubate at 37 °C for 48 h.
10. Harvest cells for isolation of genomic DNA and amplify the target genomic area by PCR. Upon successful PCR, send PCR products for Sanger sequencing.
11. Quantify the on-target efficiency of guide RNAs by Tideor ICE web tools [9, 10] and re-rank the guide RNA list based on highest on-target and lowest off-target activity.
12. For gene correction purpose, design the repair template as single-stranded oligonucleotide (ssODN) with homology arms of 50–90 nt surrounding the mutation site. Introduce a silent mutation at the PAM site to protect edited sequence from re-cutting by Cas9 after gene correction. Synthesize or order ssODN as 4 nmol Ultramer DNA oligonucleotide.

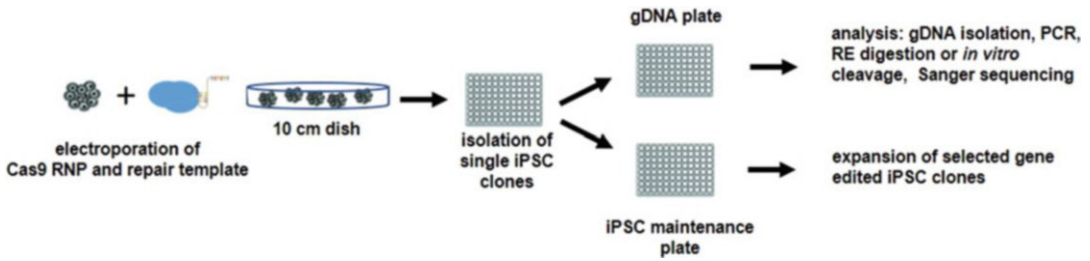


Fig. 1 Scheme of CRISPR/Cas9-based genome editing of human iPSCs

3.2 Electroporation of iPSCs with Cas9 Ribonucleoprotein (RNP)

Scheme of CRISPR/Cas9-based genome editing of human iPSCs is shown in Fig. 1.

1. Use the nucleofection protocol for the Amaxa 4D Nucleofection system with P3 Primary Cell 4D-Nucleofector X Kit.
2. For the electroporation of 10^6 cells, assemble 3.3 μ L sgRNA and 15 μ g Cas9 protein at room temperature for 30 min.
3. For detaching iPSCs, remove the medium and add Accutase, and then incubate until the cells start to detach (for approximately 3–5 min). Keep the number of pipetting to a minimum.
4. Wash cells two times with PBS prior to nucleofection to remove all traces of nucleases present in the culture medium. Centrifuge the cells at $300 \times g$ for 5 min.
5. Remove the supernatant after centrifugation. For successful electroporation, it is important to completely remove the supernatant.
6. Prepare the nucleofection buffer according to the instructions of the P3 Primary Cell 4D-Nucleofector X Kit. Add ssODN and RNP mix to the nucleofection buffer.
7. Resuspend the cell pellet in the nucleofection buffer-RNP mix and transfer iPSCs suspension to the electroporation cuvette. Try to avoid any air bubble formation.
8. Place the cuvette in the Amaxa 4D Nucleofector and choose the CA-137 program.
9. After electroporation, incubate the cells in the cuvette for 10 min at room temperature.
10. Take up the cells gently in pre-warmed StemFlex medium supplemented with RevitaCell and transfer them to the Geltrex-coated 6-well plate.

3.3 Clonal Isolation and 96-Well Plate Culture of iPSCs

1. About 48 h post-electroporation, remove the medium, wash two times with PBS, and add 2 mL Accutase; incubate until the cells start to detach. It is important to have single cells at this step. Transfer the cells in 15 mL tube, centrifuge down at

$300 \times g$ for 5 min, discard the supernatant, and resuspend cells in Stemflex medium supplemented with RevitaCell.

2. Count cells.
3. Transfer 5000 iPSCs (*see Note 3*) to 10 mL Stemflex medium supplemented with RevitaCell, mix with 10 mL pipette few times, and transfer to the Geltrex-coated 10 cm dish. Rock forth and back. Put the dish back into the incubator.
4. Incubate the 10 cm dish for 7 days. Change the medium every 24 h with Stemflex medium without RevitaCell.
5. Coat the flat bottom 96-well plate with 100 μ L Geltrex for at least 2 h at room temperature. Remove the Geltrex from the wells using a multichannel pump and add 100 μ L Stemflex medium supplemented with RevitaCell to each well using a multichannel pipette.
6. After 7 days (*see Note 4*), wash the 10 cm dish gently with PBS. Add 10 mL PBS to the dish and incubate at room temperature for 5 min.
7. Put the 10 cm dish under the microscope (*see Note 5*) and using a 100 μ L pipette that adjusted to 20 μ L, gently scratch, pick, and transfer single iPS clone to a well of 96-well plate. Pipette few times to dissociate the clone. Pick 96 clones (*see Note 6*).
8. Put the 96-well plate containing 96 single iPSC clones to the incubator for three to four for days.
9. Coat at least 2 96-well plates (use flat bottom plates) with Geltrex (*see Note 7*) before use, remove the coating, and add 100 μ L StemFlex Media supplemented with RevitaCell.
10. After 3–4 days, remove the media from 96-well plate using a multichannel pump (*see Note 8*), add 40 μ L Accutase diluted 1:3 with PBS to each well, and incubate at 37 °C for 30 min (*see Note 9*) until the iPSCs start to dissociate.
11. Prepare a sterile reservoir filled with StemFlex Media supplemented with RevitaCell. Using a multichannel pipette, add 60 μ L media to each well, pipette 4–5 times, and transfer 60 μ L iPSCs containing media to the exact same position in a new Geltrex-coated 96-well plate. This plate will be used later for genomic DNA isolation (gDNA plate). Using the same tips, transfer the remaining 35 μ L iPSCs containing media to the exact same position in the second Geltrex-coated 96-well plate. This plate will be stored in the incubator (maintenance plate).

3.4 Screening for Gene-Corrected iPSC Clones and Validation of Gene Correction

1. Using a multichannel pump, remove the media from all wells of the gDNA plate and wash each well two times with PBS.
2. Add 30 μ L of QuickExtract DNA Extraction solution.
3. Heat the 96-well plate at 65 °C for 6 min and 98 °C for 2 min. The 96-well plate can be directly used as a template DNA in PCR reaction or stored at –20 °C.
4. To amplify the target region, prepare a PCR master mix, pipette it to PCR-specific 96-well plate, and add 2 μ L of gDNA prepared in **step 11** of Subheading 3.3.
5. Perform the PCR reaction.
6. For gene correction purpose, use restriction endonucleases enzyme that produces a different pattern of digestion after gene correction (e.g., <http://insilico.ehu.es>) and perform restriction enzyme digestion of each clone.
7. For gene knockout purpose, instead of endonuclease restriction enzyme digestion, perform in vitro Cas9 cleavage using the same sgRNA that has been used for the knockout as following:
 - (a) Prepare sgRNA at a concentration of 10 μ M in a sterile microtube by diluting the sgRNA in nuclease-free buffer.
 - (b) Prepare 1 μ M RNP complex by adding 1.6 μ L Cas9 enzyme (62 μ M stock) and 88.4 μ L PBS to the microtube containing 10 μ M sgRNA.
 - (c) Incubate at room temperature for 15 min.
 - (d) Perform the in vitro digestion reaction by adding 1 μ L of 1 μ M Cas9 RNP, 1 μ L IDT nuclease-free buffer, and 3 μ L of unpurified PCR reaction from **step 5** of Subheading 3.4.
 - (e) Incubate the reaction at 37 °C for 1 h.
 PCR product of knockout candidate iPSC clones should not be digested by Cas9 RNP as the target sequence is already modified.
8. Run the reaction from **steps 6** or **7** on 1% agarose gel and perform Sanger sequencing of the PCR products with a preferred pattern in order to validate gene correction or knock out in the selected iPSC clones.

3.5 Hematopoietic and Neutrophilic Differentiation of iPSCs Via Spin-EBs

The scheme of EB-based neutrophilic differentiation is shown in Fig. 2. A list of cytokines used for EB-differentiation is shown in Table 1.

1. Two days before EB formation, seed 4×10^5 mitotically inactivated SNL feeder cells on a 3.5 cm² cell culture dish.

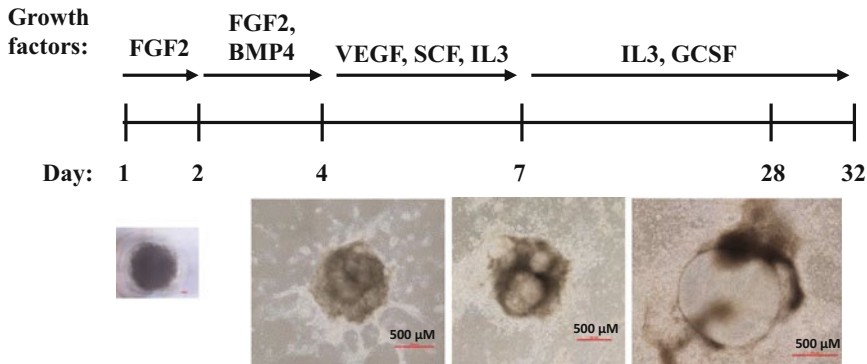


Fig. 2 Scheme of EB-based neutrophilic differentiation of human iPSCs

Table 1

List of cytokines used for EB-based hematopoietic iPSCs differentiation

Cytokine	Stock concentration	Final concentration	Function	Company
BMP4	20 $\mu\text{g}/\text{mL}$	20 ng/mL	Mesoderm induction	R&D
bFGF (FGF2)	100 $\mu\text{g}/\text{mL}$	20 ng/mL	Mesoderm differentiation	Peprotech
VEGF	40 $\mu\text{g}/\text{mL}$	40 ng/mL	Angiogenesis	R&D
SCF	100 $\mu\text{g}/\text{mL}$	50 ng/mL	Stem cell factor	Peprotech
IL-3	50 $\mu\text{g}/\text{mL}$	50 ng/mL	Myeloid differentiation	Peprotech
ROCK inhibitor (Y-27632dihydrochloride)	10 mM	10 nM	Inhibition of rho-kinase	Tocris
G-CSF	100 $\mu\text{g}/\text{mL}$	50 ng/mL	Granulocytic differentiation	Neupogen (Filgrastim)

2. One day before EB formation, seed approximately 50–70% confluent feeder-free iPSCs from one well of a 6-well plate to a 3.5 cm^2 dish with SNL feeder cells in iPSCs maintenance medium supplemented with 20 ng/mL bFGF and 10 nM ROCK inhibitor (*see Note 10*).
3. On the day of EB formation (day 1), precool centrifuge to 4 $^{\circ}\text{C}$, as it will increase efficiency of compact EB formation.
4. Prepare APEL medium for EB formation. From one 3.5 cm^2 dish, you can produce up to 30 EBs. For 30 EBs, 3 mL of APEL medium supplemented with ROCK (10 nM) and bFGF (20 ng/mL) are needed.

5. Remove iPSCs maintenance medium from iPSCs and wash once with 1 mL PBS.
6. Add 1 mL PBS/0.02% EDTA to the 3.5 cm² dish with iPSCs for 7 min, remove PBS/EDTA solution, and detach iPSCs in prepared APEL medium.
7. Collect iPSCs in a falcon tube and distribute 100 μ L per well of iPSCs suspension to a 96-well plate with conical bottom.
8. Centrifuge plate for 5 min at $435 \times g$ and 4 °C.
9. Incubate the plate for 24 h at 37 °C.
10. On day 2, add BMP4 in 10 μ L APEL medium per well for a final concentration of 20 ng/mL.
11. On day 3, prepare Geltrex-coated 6-well plates.
12. On day 4, plate EBs on Geltrex-coated wells (10 EBs/well) in 2 mL per well of APEL medium supplemented with VEGF (40 ng/mL), SCF (50 ng/mL), and IL-3 (50 ng/mL).
13. On day 7, change medium to APEL medium supplemented with IL-3 (50 ng/mL) and G-CSF (50 ng/mL). Hematopoietic cells start to appear on day 8.
14. On day 10, add 2 mL of APEL medium supplemented with IL-3 (50 ng/mL) and G-CSF (50 ng/mL).
15. On day 14, take floating cells for further analysis and exchange medium to the fresh APEL medium containing G-CSF (50 ng/mL) and IL-3 (50 ng/mL), 2 mL per well. At this stage, floating cells are mainly CD34⁺ hematopoietic stem and progenitor cells that can be used for various downstream analyses such as colony-forming unit (CFU) assay, morphological analysis, or flow cytometry analysis of early hematopoietic markers of hematopoietic stem and progenitor cells (HSPCs) (*see* Subheading 3.6).
16. From day 14, flow cytometry analysis for hematopoietic stem cells (*see* Subheading 3.6) and mature myeloid cells (*see* Subheading 3.7) as well as cell counts of floating cells should be performed every 3–4 days to follow the myeloid maturation process. Floating cells mature stepwise from CD34⁺ cells on day 14 to mature CD15⁺CD16⁺ neutrophils on day 28–32. The culture medium from day 14 to day 32 consists of APEL medium supplemented with IL-3 (50 ng/mL) and G-CSF (50 ng/mL), 2 mL medium per well.
17. On day 28 and day 32, mature neutrophils are analyzed using multicolor flow cytometry panel for mature myeloid cells (*see* Subheading 3.7 and **Note 11**).

Table 2
A list of antibodies used for the early panel FACS analysis

Antibody	Conjugation	Company	Volume in μL
CD33	BV421	BioLegend	2
CD45	BV510	BioLegend	2
KDR	AF647	BioLegend	2
CD41a	FITC	BD	0.5
CD235	FITC	BD	0.5
CD34	PE-Cy7	BD	5
CD43	PE	BD	5
7AAD		BD	2

3.6 Multicolor Flow Cytometry Analysis of the Early Stages of Hematopoietic Differentiation of iPSCs

Multicolor flow cytometry analysis for the evaluation of HSPCs in suspension cell fraction of EBs-based hematopoietic differentiation culture of iPSCs is performed on day 14 of culture.

1. Harvest 6×10^4 cells in 1 mL of ice-cold PBS and centrifuge for 5 min at $1200 \times g$. Usually, 3×10^4 cells are used for staining with specific antibody and 3×10^4 cells for unstained control.
2. Resuspend cells in 150 μL ice-cold FACS buffer.
3. Add 75 μL of cell suspension into each FACS tube on ice.
4. Add antibody as described in Table 2.
5. Incubate for 15 min on ice and protected from light.
6. Add 500 μL ice-cold FACS buffer to wash stained cells.
7. Centrifuge 5 min at $435 \times g$ and $+4^\circ\text{C}$, and then discard supernatant.
8. Resuspend cell pellet in 100 μL ice-cold FACS buffer.
9. Evaluate stained cells on BD FACS Canto II. Representative images of flow cytometry are shown in Fig. 3.

3.7 Multicolor Flow Cytometry Analysis of the Late Stages of Myeloid Differentiation of iPSCs

Multicolor flow cytometry analysis of floating cells derived from EBs-based iPSCs culture in order to evaluate surface expression of myeloid markers is performed on day 28–32.

1. Harvest 6×10^4 cells in 1 mL of ice-cold PBS and centrifuge for 5 min at $12,000 \times g$. In addition, 3×10^4 cells are used for staining with specific antibody and 3×10^4 cells for unstained control.
2. Resuspend cells in 150 μL ice-cold FACS buffer.

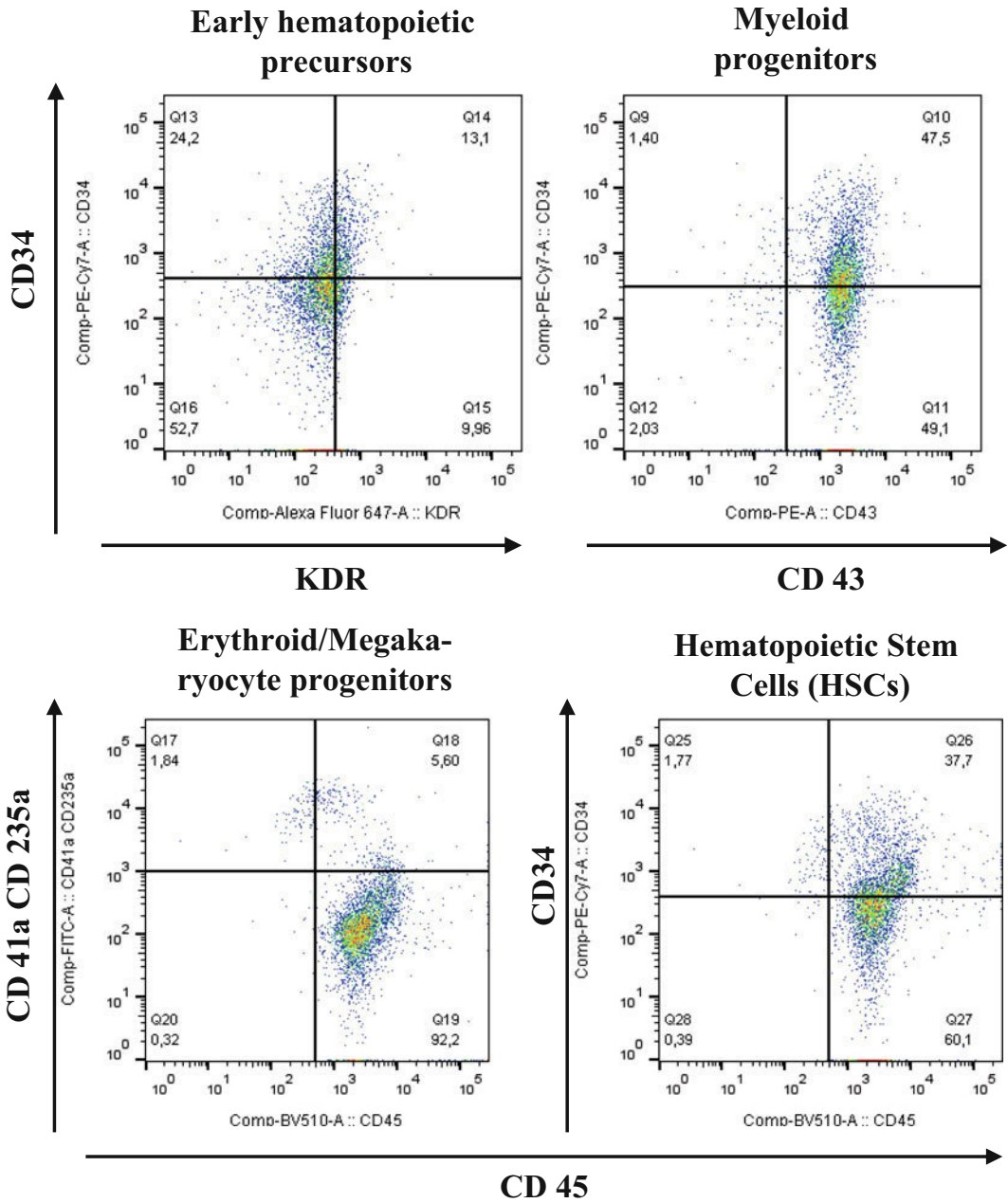


Fig. 3 Representative images of flow cytometry analysis of early stages of EB-based hematopoietic differentiation of healthy donor-derived iPSCs

3. Add 75 μ L of cell suspension into each FACS tube to a final volume of 100 μ L.
4. Add FACS antibodies as described in Table 3.
5. Incubate for 15 min on ice and protect from light.

Table 3
A list of antibodies used for the late panel FACS analysis

Antibody	Conjugation	Company	Volume in μL
CD33	BV421	BioLegend	2
CD45	BV510	BioLegend	2
CD15	PE	BD	5
CD16	FITC	BD	5
CD11b	PE-Cy7	BioLegend	2
CD14	APC-H7	BD	2
CD177	APC	BioLegend	0.5
7AAD		BD	2

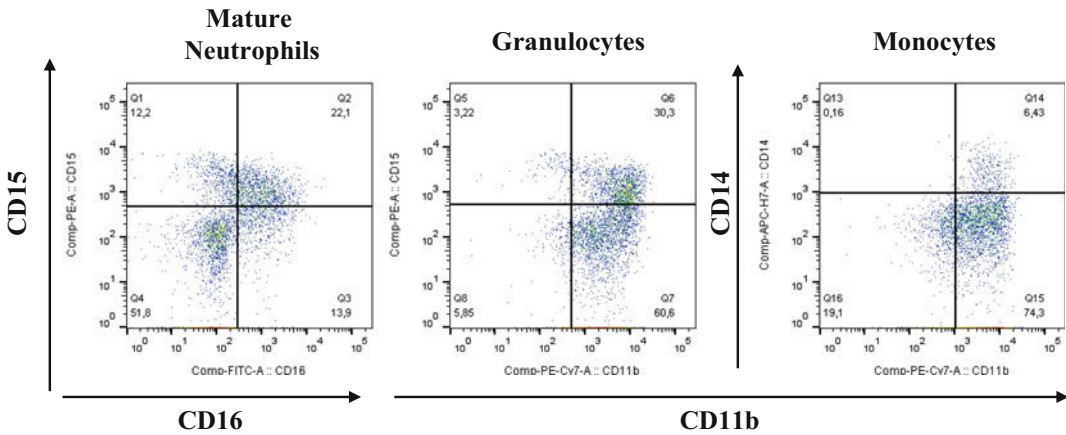


Fig. 4 Representative images of flow cytometry analysis of late stages of myeloid EB-based differentiation of healthy donor-derived iPSCs

6. Add 500 μL of ice-cold FACS buffer to wash stained cells.
7. Centrifuge 5 min at $435 \times g$ and $+4^\circ\text{C}$, and then discard supernatant.
8. Resuspend cell pellet in 100 μL of ice-cold FACS buffer.
9. Measure stained cells on BD FACS Canto II. Representative images of flow cytometry are shown in Fig. 4.

4 Notes

1. Most of the existing off-target prediction methods calculate scores based on the positions of the mismatches to the guide sequence [11]; a good off-target profile is a probability of

having no unintended cut with up to two mismatches (0:0, 1:0, 2:0). Sanger Institute webtool (<https://www.sanger.ac.uk/htgt/wge/>) can be used to assess the off-target profiles of selected sgRNAs.

2. Any RNP-compatible transfection reagent can be used at this step.
3. Cell number should be optimized between 5000 and 50,000 cells for each iPSC line.
4. The incubation time could be up to 14 days due to variations in different iPSC lines.
5. Keep the condition sterile and use a face mask. The probability of cell contamination at this step is high.
6. The number of colonies for picking can be adjusted based on the estimation of the gene modification efficiency in the total cell population.
7. The Geltrex-coated 96-well plates can be stored at 4 °C for up to 7 days.
8. Wash the tips in PBS after each use, to minimize the possibility of clonal cross-contamination.
9. This step could take up to 1 h in case of the high density of iPSC clones. Because diluted Accutase is used at this step, there will be no harm if longer incubation time is needed, but pipetting iPSC clones to accelerate the detachment could kill the cells.
10. To start EB formation, iPSCs should be more than 80% confluent. In this case, no feeder depletion is needed.
11. After day 32 of culture, the number of floating cells drops heavily.

Acknowledgments

This work was supported by the DFG (J.S.), the Fritz Thyssen Foundation (B.D.), and the Jose Carreras Leukemia Foundation (B.D.).

References

1. Kim C (2015) iPSC technology-powerful hand for disease modeling and therapeutic screen. *BMB Rep* 48(5):256
2. Hockemeyer D, Jaenisch R (2016) Induced pluripotent stem cells meet genome editing. *Cell Stem Cell* 18(5):573–586
3. de Bruin O, Lisa M, Volpi S, Musunuru K (2015) Novel genome-editing tools to model and correct primary immunodeficiencies. *Front Immunol* 6:250
4. Dannenmann B et al (2019) Human iPSC-based model of severe congenital neutropenia reveals elevated UPR and DNA damage in CD34+ cells preceding leukemic transformation. *Exp Hematol* 71:51–60

5. Lachmann N, Ackermann M, Frenzel E et al (2015) Large-scale hematopoietic differentiation of human induced pluripotent stem cells provides granulocytes or macrophages for cell replacement therapies. *Stem Cell Reports* 4 (2):282–296
6. Nasri M et al (2019) Fluorescent labeling of CRISPR/Cas9 RNP for gene knockout in HSPCs and iPSCs reveals an essential role for GADD45b in stress response. *Blood Adv* 3 (1):63–71
7. Ng ES, Davis RP, Azzola L, Stanley EG, Elefanty AG et al (2005) Forced aggregation of defined numbers of human embryonic stem cells into embryoid bodies fosters robust, reproducible hematopoietic differentiation. *Blood* 106(5):1601–1603
8. Ng ES, Davis R, Stanley EG, Elefanty AG (2008) A protocol describing the use of a recombinant protein-based, animal product-free medium (APEL) for human embryonic stem cell differentiation as spin embryoid bodies. *Nat Protoc* 3(5):768–776
9. Brinkman EK et al (2014) Easy quantitative assessment of genome editing by sequence trace decomposition. *Nucleic Acids Res* 42 (22):e168
10. Hsiao T et al. (2018) Inference of CRISPR Edits from Sanger Trace Data. *bioRxiv*. 251082
11. Lin J, Wong K-C (2018) Off-target predictions in CRISPR-Cas9 gene editing using deep learning. *Bioinformatics* 34(17):i656–i663

PUBLICATION C:

**Fluorescent labeling of CRISPR/Cas9 RNP for gene knockout in HSPCs and
iPSCs reveals an essential role for GADD45b in stress response**

Fluorescent labeling of CRISPR/Cas9 RNP for gene knockout in HSPCs and iPSCs reveals an essential role for GADD45b in stress response

Masoud Nasri,^{1,*} Perihan Mir,^{1,2,*} Benjamin Dannenmann,¹ Diana Amend,¹ Tessa Skroblyn,^{1,3} Yun Xu,¹ Klaus Schulze-Osthoff,^{2,4} Maksim Klimiankou,¹ Karl Welte,^{1,5} and Julia Skokowa^{1,2}

¹Department of Oncology, Hematology, Immunology, Rheumatology, and Pulmonology, University Hospital Tübingen, Tübingen, Germany; ²German Cancer Research Center, Heidelberg, Germany; ³Max Delbrück Center for Molecular Medicine in the Helmholtz Association, Berlin, Germany; ⁴Interfaculty Institute of Biochemistry, Tübingen University, Tübingen, Germany; and ⁵University Children's Hospital Tübingen, Tübingen, Germany

Key Points

- Fluorescent labeling of CRISPR/Cas9–gRNA RNP enables sorting of edited HSPCs and iPSCs for further applications.
- GADD45B plays a crucial role in UV stress-induced response of HSPCs and iPSCs.

CRISPR/Cas9-mediated gene editing of stem cells and primary cell types has several limitations for clinical applications. The direct delivery of ribonucleoprotein (RNP) complexes consisting of Cas9 nuclease and guide RNA (gRNA) has improved DNA- and virus-free gene modifications, but it does not enable the essential enrichment of the gene-edited cells. Here, we established a protocol for the fluorescent labeling and delivery of CRISPR/Cas9–gRNA RNP in primary human hematopoietic stem and progenitor cells (HSPCs) and induced pluripotent stem cells (iPSCs). As a proof of principle for genes with low-abundance transcripts and context-dependent inducible expression, we successfully deleted growth arrest and DNA-damage-inducible β (*GADD45B*). We found that GADD45B is indispensable for DNA damage protection and survival in stem cells. Thus, we describe an easy and efficient protocol of DNA-free gene editing of hard-to-target transcripts and enrichment of gene-modified cells that are generally difficult to transfect.

Introduction

CRISPR/Cas9-mediated gene editing^{1,2} has a tremendous potential for clinical applications, such as gene therapy of inherited disorders or boosting of immune cells for cancer immunotherapies.^{3–9} Several monogenic disorders, including life-threatening bone marrow failure syndromes, might be treated by CRISPR/Cas9-mediated gene correction in autologous hematopoietic stem and progenitor cells (HSPCs) *ex vivo*. This procedure could then be followed by transplantation of the corrected HSPCs without exposing the patient to harsh immunosuppression regimens (ie, chemotherapy, irradiation).

Two groups recently published successful gene therapy approaches to cure sickle cell disease, a common inherited blood disorder. They generated deletions in the β -globin gene locus using CRISPR/Cas9 technology to mimic the hereditary persistence of fetal hemoglobin mutations.^{10,11} First attempts at gene editing using CRISPR/Cas9 for cancer therapy have also been launched recently. CRISPR/Cas9-generated chimeric antigen receptor–modified T cells targeting the checkpoint receptor programmed cell death 1 have been injected in a patient with metastatic non-small cell lung cancer.^{12,13} In addition to HSPCs, CRISPR/Cas9-mediated gene editing was successfully applied in neurons, hepatocytes, and cardiomyocytes.^{14–20}

To further advance clinical applications of the CRISPR/Cas9 technology, unspecific integrations of viral or plasmid CRISPR/Cas9 DNA in the host genome and undesirable immune responses must be prevented. This may be achieved by transient virus- and DNA-free delivery approaches using CRISPR/Cas9–guide RNA (gRNA) ribonucleoprotein (RNP) complexes. Compared with DNA-based approaches, the direct delivery of CRISPR/Cas9–gRNA RNP complexes might have several advantages. Due to the

Submitted 24 December 2017; accepted 28 November 2018. DOI 10.1182/bloodadvances.2017015511.

*M.N. and P.M. contributed equally to this work.

The full-text version of this article contains a data supplement.

© 2019 by The American Society of Hematology

cellular degradation of the RNP complex, exposure of cells to Cas9 is generally only transient and restricted, thereby limiting potential off-target effects of endonuclease overexpression.

Gundry et al recently described the efficient delivery of CRISPR/Cas9-gRNA RNP in HSPCs²¹; however, the method used in that study did not allow the enrichment of gene-edited cells. Purification of gene-edited HSPCs early in the manufacturing process is desirable, especially for clinical applications, because HSPCs differentiate and progressively lose their long-term repopulating capacity during culture. This is especially true for genes with low-abundance messenger RNA (mRNA) transcripts or inducible mRNA expression. These genes may be difficult to target, also due to epigenetic modifications, leading to tightly packed chromatin at the time of gRNA delivery. In these cases, the selection of gene-edited cells is indispensable. Nonmodified cells may retain a proliferative advantage over gene-edited cells, especially in mixed populations.

Introduction of a fluorescent tag to the CRISPR/Cas9-gRNA RNP enables enrichment of gene-edited cells for further experimental and clinical applications. Tagging of Cas9 protein could be achieved by fusion of Cas9 with fluorescent proteins or by chemical labeling of the CRISPR/Cas9-gRNA RNP with a fluorescent dye. However, Cas9-GFP fusion proteins might affect the intracellular localization, activity, or on- and off-target specificity of the endonuclease. Therefore, in the present study, we established a safe, simple, and efficient method for CRISPR/Cas9 gene knockout using transfection of stem cells with fluorescently labeled CRISPR/Cas9-gRNA RNP complexes.

Materials and methods

Cell culture

Human embryonic kidney 293FT (HEK293FT) and Jurkat cells were cultured under standard conditions (37°C, 5% CO₂) using Dulbecco's modified Eagle medium high glucose (HEK293FT cells) or RPMI 1640 GlutaMAX (Jurkat cells) medium supplemented with 10% fetal calf serum (Sigma-Aldrich) and 1% penicillin/streptomycin (Biochrome). HEK293FT cells were detached using 0.05% Trypsin-EDTA (Gibco) and seeded at a density of 1×10^5 cells per milliliter of medium. Jurkat cells were seeded at a density of 1 to 2×10^5 cells per milliliter of medium.

Human CD34⁺ HSPCs were isolated from the bone marrow or leukapheresis mononuclear cell fraction by magnetic bead separation (Human CD34 Progenitor Cell Isolation kit; Miltenyi Biotech). CD34⁺ cells were cultured in a density of 2×10^5 cells/mL of Stemline II medium (Sigma Aldrich) supplemented with 10% fetal calf serum, 1% penicillin/streptomycin, 1% L-glutamine, and a cytokine cocktail consisting of 20 ng/mL IL-3, 20 ng/mL IL-6, 20 ng/mL thrombopoietin, 50 ng/mL SCF, and 50 ng/mL Flt-3L. Human induced pluripotent stem cells (iPSCs) were cultured on plates coated with Geltrex lactate dehydrogenase elevating virus-free reduced growth factor basement membrane matrix (cat. no. A1413201; Thermo Fisher Scientific) at a density of 2×10^5 cells/mL in StemFlex medium (cat. no. A3349401; Thermo Fisher Scientific) supplemented with 1% penicillin/streptomycin.

Generation and testing of the *GADD45B* gRNA

Specific CRISPR RNA (crRNA) for the first exon of the *GADD45B* gene (GCTCGTGGCGTGCGACAACGCGG, cut site: chr19

[+2,476,389: -2,476,389], NM_015675.3 Exon 1, 31bp; NP_056490.2 position N11) was designed using an online tool from the University of Heidelberg (<http://crispr.cos.uni-heidelberg.de>). The crRNA for *GADD45B* was first tested in transfected HEK293FT cells showing a gene modification efficiency of 67% in the total population of transfected cells.

Labeling of gRNA and plasmid DNA

Trans-activating CRISPR RNA (tracrRNA) and crRNA, obtained from IDT, were annealed at a ratio of 1:1 by incubating for 15 minutes at room temperature to generate gRNA. gRNA was fluorescently labeled using LabellIT CX-Rhodamine (cat. no. MIR7022; Mirus) or LabellIT Fluorescein (cat. no. MIR7025; Mirus) kits according to the manufacturer's instructions. Labeling reagent and nucleic acid ratio were used at a ratio of 1:1 leading to 1 label per 20 to 60 bases, which is suitable for most applications.

Generation of crRNA-tracrRNA duplexes (gRNA) was conducted by adding 800 pmol of *GADD45B*-targeting crRNA and 800 pmol tracrRNA into 40 μ L nuclease-free duplex buffer (IDT) at room temperature for 15 minutes. Labeling of the gRNA was performed by mixing gRNA, DNase- and RNase-free water, 10 \times labeling buffer A and 1:10 of LabellIT Reagent (CX-rhodamine or fluorescein). The reaction was incubated at 37°C for 1 hour while centrifuging briefly after 30 minutes to minimize evaporation and maintain the appropriate concentration of the reaction components.

Purification of labeled gRNA was conducted using the ethanol precipitation method. To this end, 5 M sodium chloride (0.1 volume) and ice-cold ethanol (2.5 volume) were added to the reaction, mixed well, and placed at -20°C for 30 minutes. Afterwards, the sample was centrifuged at 14 000g at 4°C for 30 minutes to pellet the labeled gRNA. Once pelleted, the supernatant was discarded gently without disturbing the pellet. The pellet was washed using 70% ethanol at room temperature and centrifuged at 14 000g for 30 minutes. After centrifugation, the pellet was air dried for 5 minutes and resolved in IDT nuclease-free duplex buffer. The labeled gRNA stock was stored at -20°C for up to 2 months.

Labeling of the pMAX GFP plasmid (Lonza) was carried out using LabellIT Tracker Intracellular Nucleic Acid Localization Kit (cat. no. MIR7022; Mirus) following the manufacturer's protocol.

Assessment of the RNA integrity using Agilent Bioanalyzer

Labeled and unlabeled gRNA were analyzed using the Agilent RNA 6000 Pico Kit according to the manufacturer's instructions on the Agilent 2100 Bioanalyzer using the total RNA program.

Transfection of cells with CRISPR/Cas9-gRNA RNP complexes

Transfection was carried out either using TransIT-X2 (cat. no. MIR6003; Mirus) dynamic delivery system or the Amaxa nucleofection system (P3 primary kit, cat. no. V4XP-3024) according to the manufacturers' instructions. For 0.5×10^5 HEK293FT cells, 100 pmol of labeled duplexed gRNA was mixed with 100 pmol of Cas9 protein (Alt-R S.p. Cas9 Nuclease 3NLS, cat. no. 1074182; IDT) in IDT nuclease-free duplex buffer and assembled for 30 minutes at room temperature. Afterwards, the CRISPR/Cas9-gRNA RNP was mixed with either Opti-MEM I reduced-serum medium and TransIT-X2 transfection reagent (HEK293FT) or with electroporation mix for the

Amaya nucleofection system according to the manufacturer's protocol (Jurkat, and human iPSCs and CD34⁺ HSPCs, respectively). Jurkat cells (1.0×10^6) were electroporated with 300 pmol labeled duplexed gRNA mixed with 300 pmol Cas9 protein. Human iPSCs and CD34⁺ HSPCs (1.0×10^6) were electroporated with 400 pmol labeled duplexed gRNA and 400 pmol Cas9 protein. Transfection of HEK293FT cells with CX-rhodamine-labeled pMAX GFP plasmid was performed using TransIT-LT1 transfection reagent (cat. no. MIR2304; Mirus).

Genomic DNA isolation, PCR, Sanger sequencing and TIDE assay

Genomic DNA (gDNA) was isolated using the QIAamp DNA Mini Kit (cat. no. 51306; Qiagen) according to the manufacturer's instructions. Polymerase chain reaction (PCR) with isolated gDNA and *GADD45B*-specific primers (forward: 5'-GACTACCGTTGGTTCCGCAAC-3', reverse: 5'-ATACATCAGGATACGGCAGCCC-3') was carried out using the GoTaq Hot Start Polymerase Kit (cat. no. M5006; Promega) using 50 ng of gDNA template for each PCR reaction. PCR product purification was conducted with ExoSAP (ratio 3:1), a master mix of 1 part Exonuclease I 20 U/ μ L (cat. no. EN0581; Thermo Fisher Scientific) and 2 parts of FastAP thermosensitive alkaline phosphatase 1U/ μ L (cat. no. EF0651; Thermo Fisher Scientific). Sanger sequencing of purified PCR product was performed by Eurofins Genomics and analyzed using the TIDE (Tracking of Indels by Decomposition) webtool developed by Brinkman et al.²²

Establishment of gene-edited cell lines and human iPSCs from single-cell clones using limiting dilution

Cells were serially diluted to 0.5 cells per 100 μ L by adding of 60 cells in 12 mL Dulbecco's modified Eagle medium or RPMI medium and pipetting of 100 μ L of cell suspension per well of the 96-well plate. The 96-well plate was incubated for 2 to 3 weeks until appearance of growing cells.

Human iPSCs (15 000) were plated on a Geltrex-coated 10-cm dish in StemFlex medium (cat. no. A3349401; Thermo Fisher Scientific) and RevitaCell supplement (cat. no. A2644501; Thermo Fisher Scientific). Medium was changed every 24 hours without RevitaCell supplement. After 9 to 12 days, each colony was picked and transferred on the Geltrex-coated 96-well plate.

Cloning of the PCR products for the evaluation of the gene modification mode in *GADD45B*-edited clones

gDNA was isolated from gene-edited *GADD45B*^{+/-} and *GADD45B*^{-/-} iPSCs. The Cas9 RNP-targeted region of the *GADD45B* gene was amplified from gDNA using PCR with followed primers: forward 5'-GACTACCGTTGGTTCCGCAAC-3', reverse 5'-ATACATCAGGATACGGCAGCCC-3'. PCR product was purified from the agarose gel using QIAquick Gel Extraction kit (cat no./ID: 28706; Qiagen) and cloned into the linearized pMiniT 2.0 vector using the NEB PCR Cloning Kit (cat. no. E1202S; New England Biolabs) followed by transformation of competent *Escherichia coli* and subsequent colony PCR of *E coli* colonies, according to the manufacturer's instructions (cat. no. M5006; Promega). PCR products were analyzed using Sanger sequencing.

UV exposure and cell viability assay

Cells were irradiated with UV light (7 mJ/cm²) for 5 minutes and subsequently incubated for 2 hours under standard culture conditions before measuring the percentage of live *GADD45B*-targeting

CRISPR/Cas9-gRNA RNP-transfected cells by quantitation of CX-rhodamine- or fluorescein-positive cells using a BD FACS-Canto II flow cytometer.

Intracellular staining and fluorescence-activated cell sorter analysis of γ H2AX (pSer139) protein

Intracellular γ H2AX (pSer139) protein levels were measured in UV-irradiated cells. Briefly, cells were washed with phosphate-buffered saline and stained using the IntraSure kit (cat. no. 641778; BD) according to the manufacturer's instructions and incubated with Alexa-Fluor 488 mouse anti-H2AX pSer139 antibody (1:100; cat. no. 560445; BD) for 15 minutes at room temperature, washed twice, fixed with 0.5% paraformaldehyde and analyzed using a FACSCanto II flow cytometer.

LORD-Q-DNA damage quantification

gDNA was isolated using a QIAamp DNA Mini Kit according to the manufacturer's instructions. Long-run real-time PCR-based DNA-damage quantification (LORD-Q) was performed and analyzed according to the protocol of Lehle et al.²³

Results

Design of the CRISPR/Cas9-gRNA RNP fluorescent labeling

We generated gRNA by annealing crRNA with tracrRNA. gRNA was covalently labeled with CX-rhodamine or fluorescein and incubated with recombinant Cas9 protein to generate CRISPR/Cas9-gRNA RNP complexes (Figure 1A). To assess the efficiency of fluorescent labeling, we transfected HEK293FT cells with a CX-rhodamine-labeled plasmid encoding GFP protein. GFP signals were colocalized with CX-rhodamine signals, thus proving efficient labeling of the GFP plasmid with CX-rhodamine (supplemental Figure 1A). An Agilent Bioanalyzer was used to further confirm that fluorescent labeling does not affect gRNA integrity (Figure 1B).

Specific knockout of *GADD45B* using labeled CRISPR/Cas9-gRNA RNP

To functionally validate the knockout of weakly expressed genes with inducible mRNA expression using labeled CRISPR/Cas9-gRNA RNP, we chose to disrupt the human growth arrest and DNA-damage-inducible 45 β (*GADD45B*) gene.²⁴ We designed crRNA for exon 1 of *GADD45B* (Figure 1C), generated labeled *GADD45B* CRISPR/Cas9-gRNA RNP, and transfected HEK293FT cells, the Jurkat T-ALL cell line, bone marrow CD34⁺ HSPCs, and iPSCs. We detected CX-rhodamine or fluorescein signals 6 hours (HEK293FT cells) or 12 hours (Jurkat cells, CD34⁺ HSPCs, and iPSCs) after transfection. Transfection efficiency varied between 40% and 80%, depending on the cell type (Figure 2A-B). The intracellular fluorescent signal disappeared ~48 hours after transfection. Labeling did not affect the gene-editing efficiency of CRISPR/Cas9-gRNA RNP, as assessed by Sanger sequencing and tracking of indels by decomposition (TIDE) assay analysis of HEK293FT cells, Jurkat cells, CD34⁺ HSPCs, and human iPSCs transfected with labeled or unlabeled *GADD45B*-targeting CRISPR/Cas9-gRNA RNP (Figure 2C). Using fluorescein or rhodamine signals of labeled CRISPR/Cas9-gRNA RNP, we sorted and enriched gene-edited fluorescent cells by flow cytometry. Gene-modification efficiency in sorted cells was approximately 40% in iPSCs, 60% in HSPCs, and 70% in Jurkat cells (Figure 2D).

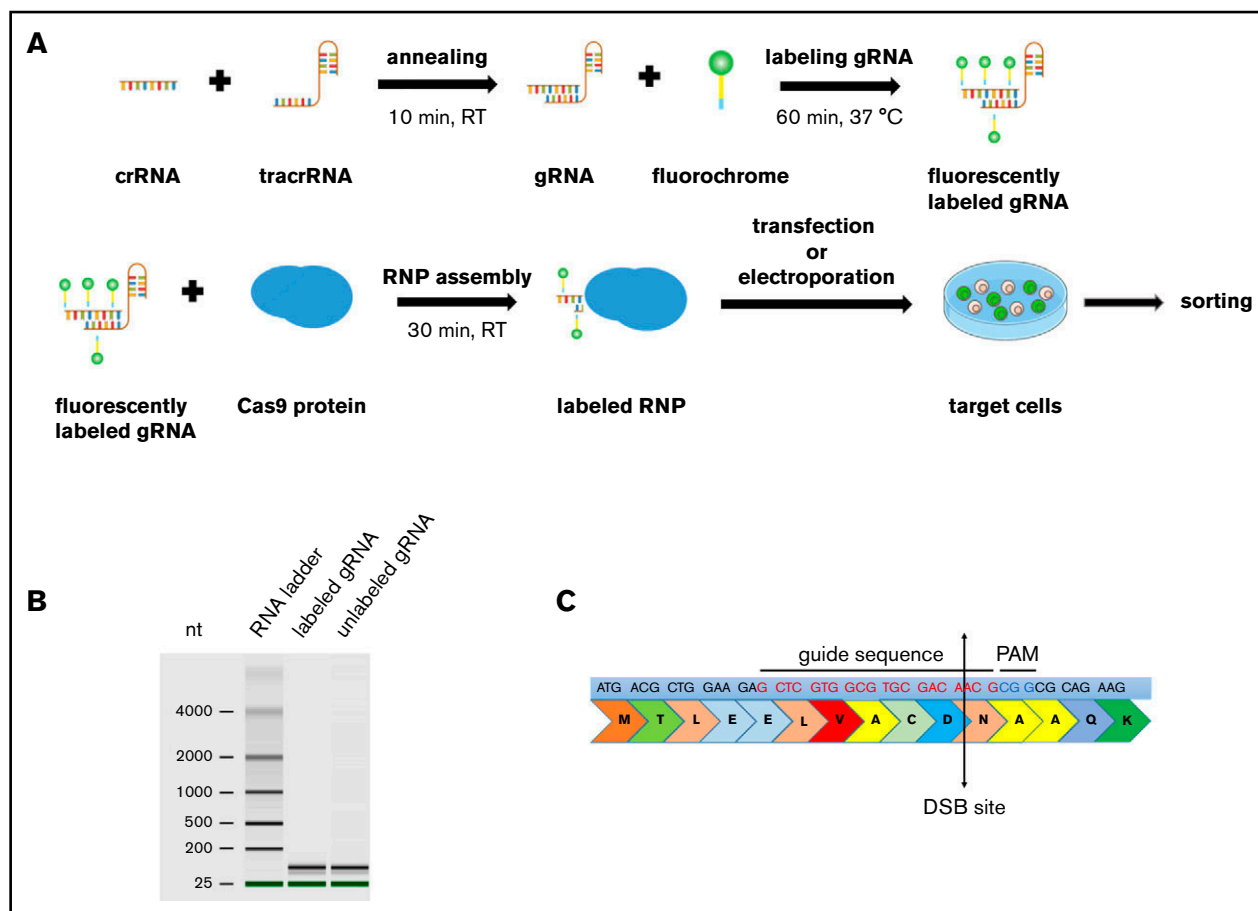


Figure 1. Scheme of CRISPR/Cas9-gRNA RNP labeling and cell transfection. (A) crRNA and tracrRNA were annealed at room temperature for 10 minutes. The resulting gRNA was labeled with fluorescein- or CX-rhodamine-coupled Label IT Tracker labeling reagent. The fluorescent *GADD45B*-targeting gRNA was assembled with recombinant Cas9 protein prior to transfection to assemble an active CRISPR/Cas9-gRNA RNP complex targeting human *GADD45B*. Cells were transfected with TransIT-X2 Transfection Reagent or by using the Amaxa Nucleofector System and were incubated for 24 hours before sorting the CX-rhodamine⁺ or fluorescein⁺ cells using a BD FACSAria II. After sorting, some of the cells were used for a single-cell culture, and the rest were used for DNA isolation or cell-based assays. (B) Virtual gel of an Agilent Bioanalyzer analysis revealing no difference in the size or quality of labeled gRNA compared with unlabeled gRNA. (C) *GADD45B* was targeted using gRNA (highlighted in red), which inserts a double-strand break at NM_015675.3 exon 1, 31 bp after ATG; NP_056490.2, p.N11.

Transfection of cells with a nontargeting RNP, consisting of tracrRNA and Casp9 alone, did not affect genome integrity (supplemental Figure 1B). We also compared fluorescent labeling of crRNA with the expression of Cas9-EGFP fusion protein. We detected much lower editing efficiency of the fused Cas9-EGFP protein assembled with *GADD45B*-targeting gRNA compared with CRISPR/Cas9-gRNA RNP (supplemental Figure 2A).

Cloning of the PCR products from genomic DNA of single-cell clones of the gene-edited Jurkat cells and iPSCs revealed compound heterozygous *GADD45B* frameshift mutations in Jurkat cells (supplemental Figure 3A), and heterozygous, as well as homozygous, *GADD45B* deletions in iPSC clones (supplemental Figure 3B). We found no off-target activities of the *GADD45B*-specific crRNA, with the possibility of 3 bp mismatches. We also detected only a small number of potential off-target sites in other genes that could be targeted (<3 mismatches) with low probability (0.2%-0.9%) by the *GADD45B*-specific crRNA (supplemental Table 1). However, we did not detect any mutations in the selected gene regions in the edited cell types used in our study (supplemental Figure 4A).

***GADD45B* is essential for the induction of DNA damage response in human hematopoietic cells and iPSCs**

We further performed functional studies of the effect of *GADD45B* knockout on cell growth and sensitivity to UV-induced DNA damage. Remarkably, we detected a strongly diminished viability of *GADD45B*-deficient HEK293FT cells, Jurkat cells, iPSCs, and CD34⁺ HSPCs compared with control transfected cells (Figure 3A). We also found markedly elevated susceptibility of *GADD45B*-deficient Jurkat cells and CD34⁺ HSPCs to UV-induced DNA damage, as documented by increased expression of the DNA damage marker γ H2AX (phospho-Ser139). Basal levels of γ H2AX (phospho-Ser139) were also elevated in *GADD45B*-modified HSPCs and Jurkat cells (Figure 3B-C). In addition, we detected an accumulation of DNA lesions in *GADD45B*-deficient cells compared with wild-type Jurkat cells (Figure 3D). As revealed by LORD-Q DNA damage-quantification analysis,²³ *GADD45B*-deficient cells exhibited increased DNA damage rates in the mitochondrial DNA (mtDNA), as well as in 2 analyzed genomic loci

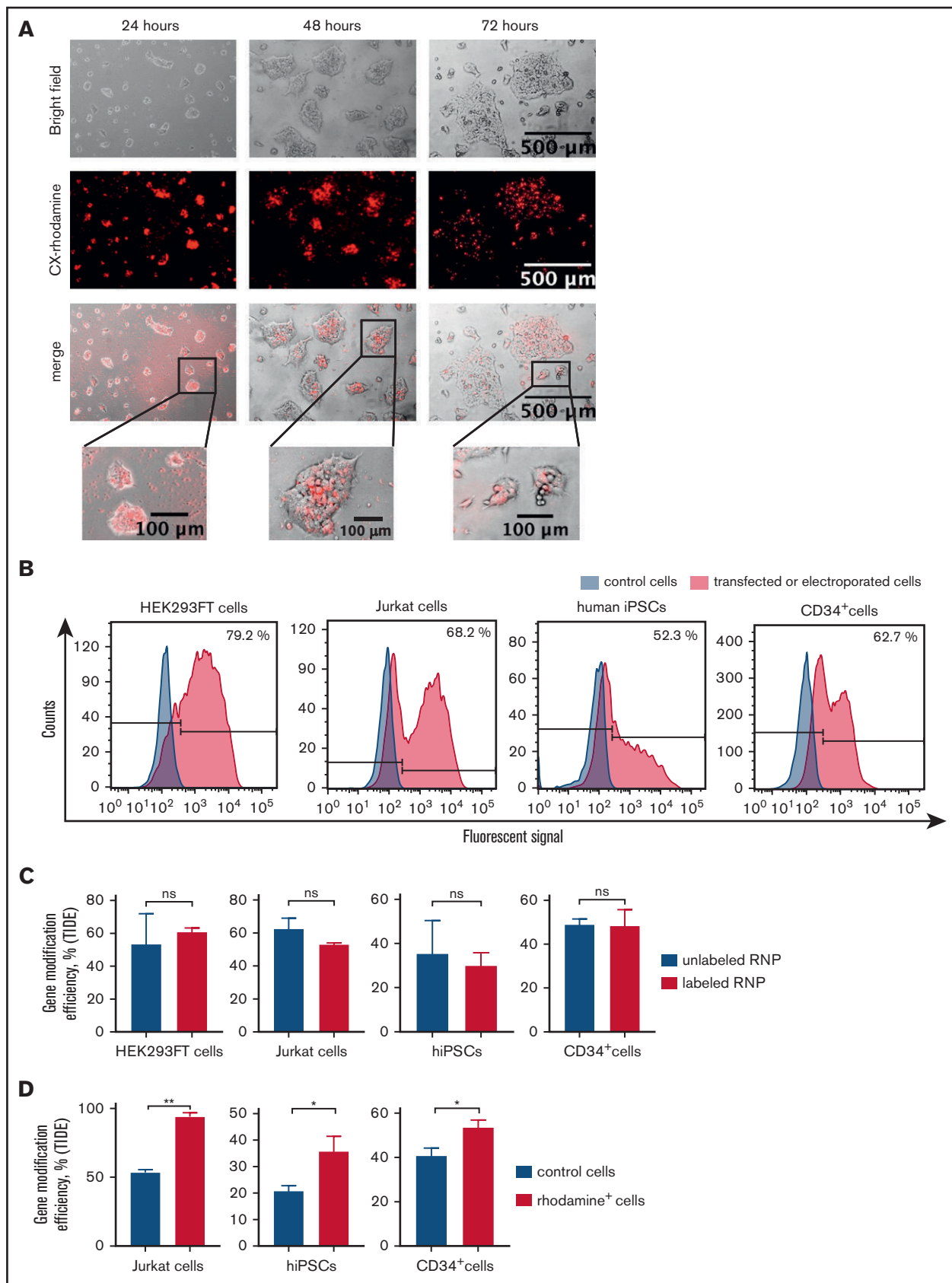


Figure 2.

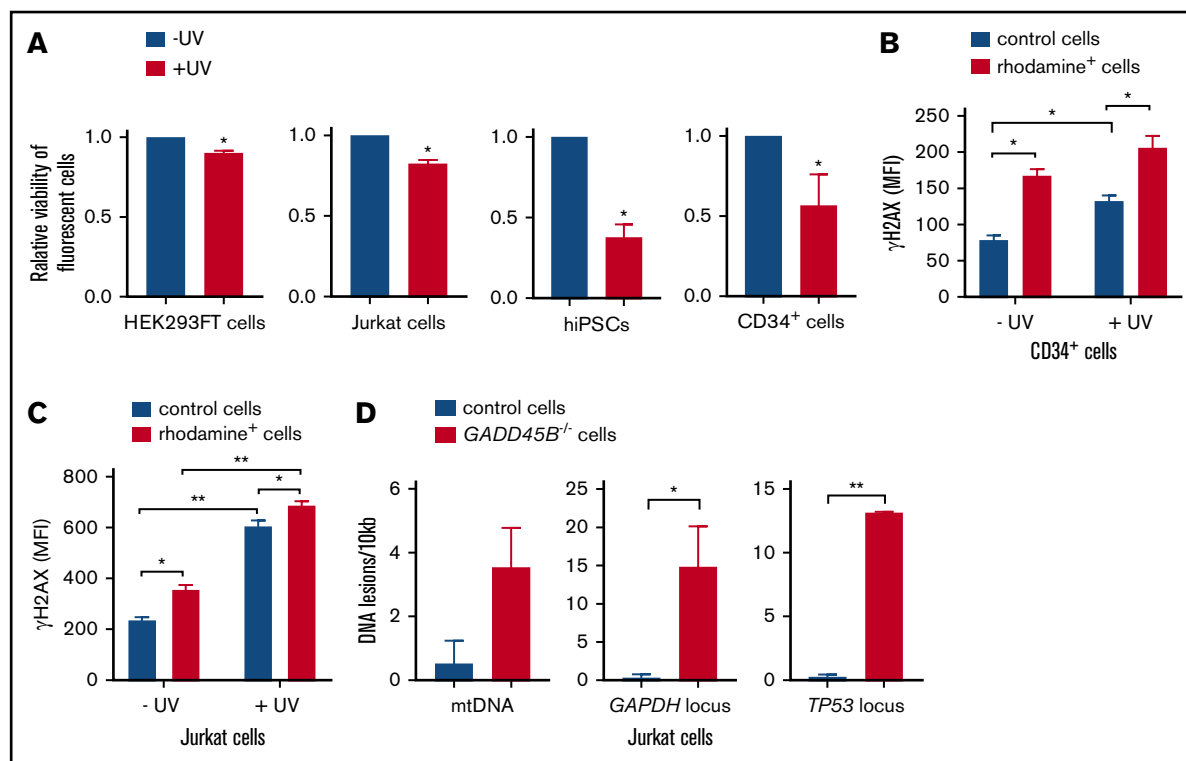


Figure 3. *GADD45B* knockout leads to reduced cell viability and increased UV-induced cellular stress. (A) Cell viability of HEK293FT cells, Jurkat cells, iPSCs, and CD34⁺ HSPCs, transfected with labeled tracr-Cas9 RNP (nontarget RNP) or with labeled *GADD45B*-targeting CRISPR/Cas9-gRNA RNP, was measured after exposing the cells to UV for 5 minutes, followed by 2 hours of additional incubation. Relative viability of nonirradiated control cells was set as 1.0. (B) CD34⁺ HSPCs were transfected with fluorescein-labeled *GADD45B*-targeting CRISPR/Cas9-gRNA RNP. After 48 hours, the cells were exposed to UV irradiation for 5 minutes. Following 2 hours of further incubation, intracellular γ H2AX (phospho-Ser139) levels were measured by flow cytometry. (C) Jurkat cells were transfected with CX-rhodamine-labeled *GADD45B*-targeting CRISPR/Cas9-gRNA RNP. After 48 hours, the total population was exposed to UV irradiation for 5 minutes, followed by 2 hours of incubation before performing intracellular staining and FACS analysis for the DNA damage marker γ H2AX (phospho-Ser139). (D) mtDNA damage (left panel) and nuclear DNA damage in the *GAPDH* locus (middle panel) and *TP53* locus (right panel) were quantified in Jurkat control cells and a *GADD45B*^{-/-} Jurkat clone using the LORD-Q method. Data are mean \pm standard deviation from 3 (A-B) or 2 (C-D) independent experiments, each performed in duplicates. * $P \leq .05$, ** $P \leq .01$, Student *t* test.

of nuclear DNA. Based on these observations, we conclude that *GADD45B* knockout in cells transfected with labeled *GADD45B*-targeting CRISPR/Cas9-gRNA RNP led to increased susceptibility to DNA damage.

***GADD45B* protects iPSCs from UV stress in a dose-dependent manner**

Interestingly, *GADD45B*^{+/-} and *GADD45B*^{-/-} iPSCs retained pluripotency (Figure 4A), but we detected markedly elevated

phospho-Ser139 γ H2AX levels in *GADD45B*-haploinsufficient and *GADD45B*-homozygous-knockout iPSCs. Elevated DNA damage was observed under steady-state conditions and, more profoundly, upon genotoxic UV exposure compared with wild-type iPSCs (Figure 4B). In line with upregulated γ H2AX (phospho-Ser139) levels, we measured elevated DNA damage in *GADD45B*^{+/-} and *GADD45B*^{-/-} iPSCs, as determined by the LORD-Q DNA damage-quantification assay (Figure 4C-D). These data are in accordance with our observations in *GADD45B*-deficient Jurkat cells and HSPCs.

Figure 2. Transfection- and genome-editing efficiency in different cell types using CX-rhodamine-labeled CRISPR/Cas9-gRNA RNP targeting *GADD45B*.

(A) HEK293FT cells were transfected with CX-rhodamine-labeled *GADD45B*-targeting CRISPR/Cas9-gRNA RNP. Fluorescence signal could be detected for up to 72 hours posttransfection. Representative images of 3 experiments are shown. (B) HEK293FT cells, Jurkat cells, human iPSCs, and CD34⁺ cells were transfected with labeled *GADD45B*-targeting CRISPR/Cas9-gRNA RNP. At 24 hours posttransfection, cells were harvested and measured for transfection efficiency using a BD FACSCanto II flow cytometer. Representative line graphs of 3 independent experiments are shown. (C) HEK293FT cells, Jurkat cells, human iPSCs, and CD34⁺ HSPCs were transfected with unlabeled or labeled *GADD45B*-targeting CRISPR/Cas9-gRNA RNP and analyzed for gene-modification efficiency using a TIDE assay. (D) Jurkat cells, human iPSCs, and CD34⁺ HSPCs were transfected with CX-rhodamine-labeled *GADD45B*-targeting CRISPR/Cas9-gRNA RNP and sorted 24 hours posttransfection using a flow cytometer. Genomic DNA was isolated 48 hours posttransfection from the total population of transfected cells and from sorted CX-rhodamine⁺ or fluorescein⁺ cells. TIDE assay analysis showed significantly higher gene modification efficiency in CX-rhodamine⁺ cells. Data in panels C and D are mean \pm standard deviations derived from 3 (HEK293FT cells, Jurkat cells, CD34⁺ HSPCs) or 4 (iPSCs) independent experiments. * $P \leq .05$, ** $P \leq .01$, Student *t* test. ns, not significant.

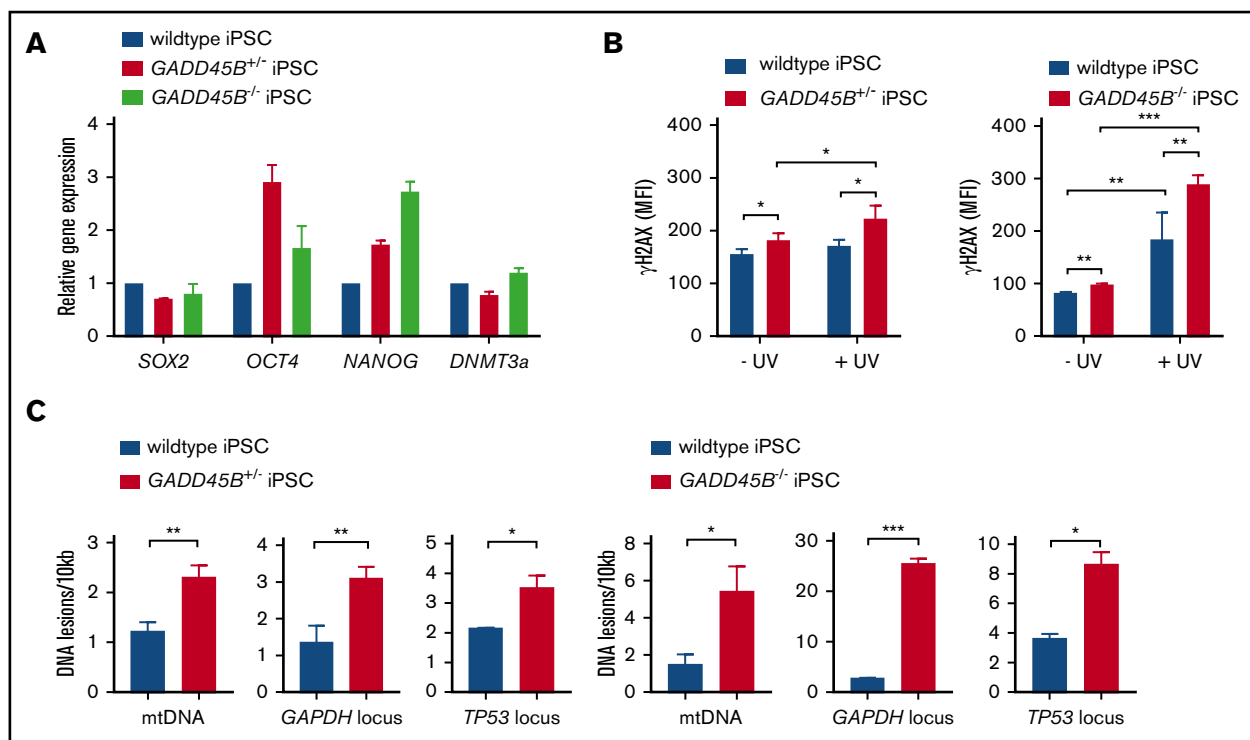


Figure 4. Heterozygous and homozygous *GADD45B* knockout in human iPSCs results in high levels of DNA damage. (A) Pluripotency state of *GADD45B*^{+/+} and *GADD45B*^{-/-} iPSCs was assessed by real-time quantitative PCR and compared with validated healthy donor-derived human iPSCs expressing wild-type *GADD45B*. (B) *GADD45B* wild-type, *GADD45B*^{+/+}, and *GADD45B*^{-/-} iPSCs were irradiated with UV light for 5 minutes, incubated under cell culture conditions for 2 hours, and stained for intracellular γ H2AX (phospho-Ser139). DNA damage in *GADD45B* wild-type and *GADD45B* heterozygous-knockout (C) or homozygous-knockout (D) iPSCs was quantified by the LORD-Q method. Cells were analyzed for mtDNA damage and nuclear DNA damage in the *GAPDH* and *TP53* gene loci. All data are mean \pm standard deviation derived from 3 independent experiments. * $P \leq .05$, ** $P \leq .01$, *** $P \leq .001$, Student *t* test.

Discussion

In the present study, we have developed a new method of CRISPR/Cas9-mediated gene editing in primary HSPCs and iPSCs using fluorescent labeling of CRISPR/Cas9-gRNA RNP complexes. Using fluorescein or CX-rhodamine signals from labeled CRISPR/Cas9-gRNA RNP, we were able to enrich gene-edited cells by fluorescent cell sorting. For clinical settings, it is essential to select and enrich gene-edited HSPCs because of the limited capacity of these cells to divide, to retain their engraftment capability, and to generate unlimited numbers of progeny cells. Sorting of the labeled cells also allows removal of untargeted HSPCs that may compete with gene-edited cells. In addition, application of CRISPR/Cas9 RNP decreases the probability and frequency of off-target effects, because CRISPR/Cas9 RNP activity is preserved in cells for only ~48 hours.

Dever et al recently reported the CRISPR/Cas9 RNP-based modification and enrichment of human HSPCs by introduction of a repair template targeting the human β -globin gene through a GFP-expressing adeno-associated virus-based vector.³ Although considered safer than retroviral constructs, adeno-associated virus-based expression constructs may induce antiviral host immune responses and may integrate into the host genome nonspecifically. Because RNP-mediated gene knockout allows the efficient virus- and DNA-free transfection and selection of edited cells, future studies should further investigate our method of fluorescent labeling of

CRISPR/Cas9 RNP in a gene-correction approach mediated by homology-directed repair, which additionally requires a donor template DNA. Again, the short exposure of cells to CRISPR/Cas9 RNP activity makes it superior to virus-based delivery techniques.

We also tested EGFP-conjugated CRISPR Cas9 RNP but found that EGFP tagging resulted in reduced editing efficiency compared with unlabeled CRISPR Cas9 RNP. At the same time, labeling of CRISPR/Cas9 RNP using our method described here does not affect editing.

Although we studied gene editing in primary hematopoietic stem cells and iPSCs, the method described here may be extended to other primary cell types. Gene-modification efficiency is dependent on the cell type, cell cycle stage, activation of DNA-repair pathways, chromatin dynamics at the gRNA-targeted gene locus, and the delivery method.²⁵⁻²⁷ Delivery and editing protocols may be further improved (eg, by synchronizing of the targeted cells), which could increase nuclear uptake of the RNP components and chromatin relaxation.

For a proof of principle, we chose to target the *GADD45B* gene and found that homo- and even heterozygous deficiency in *GADD45B* led to increased susceptibility to DNA damage. *GADD45B* belongs to a family of evolutionarily conserved *GADD45* proteins²⁸ that functions as stress sensors regulating cell cycle, survival, and apoptosis in response to various stress stimuli.²⁹ With some degree of redundancy, *GADD45* proteins exhibit specific functions, depending on the stimulus and cell type. *Gadd45b*-knockout mice

are viable, but *Gadd45b*^{-/-} primary mouse embryonic fibroblasts proliferate slowly and accumulate increased levels of DNA damage and features of premature senescence.³⁰⁻³³ It has also been shown that myeloid differentiation is compromised in *Gadd45b*^{-/-} mice.³³ Nothing was known about GADD45b functions in the DNA damage response of human HSPCs and iPSCs. Our study demonstrates an essential role for GADD45B in the survival and protection from DNA damage in human leukemia cells and CD34⁺ HSPCs that are not compensated for by other GADD45 proteins.

Our study is the first report describing transfection of iPSCs with labeled CRISPR/Cas9-gRNA RNP with high transfection and knockout efficiency. In future studies, gene correction using labeled RNP complexes for ex vivo gene therapy and transplantation could be tested in CD34⁺ cells. There are also reports describing in vivo gene correction in neurons, hepatocytes, and cardiomyocytes.^{3,14-20} Currently, there are many challenges associated with using CRISPR/Cas9 approaches. Efficient gene editing is reliant on the successful delivery of the Cas9 nuclease and the gRNA, which is especially cumbersome in primary cell types resistant to DNA transfection. Moreover, plasmid and viral delivery lead to persistent over-expression of Cas9, which can potentially result in off-targets. In contrast, direct delivery of RNP complexes, which are gradually cleared by intracellular degradation over time, does not result in Cas9 persistence and, therefore, reduces potential off-target effects. Furthermore, we did not detect toxic or inhibitory effects of the labeling on the gene-editing efficiency or on cell growth. As a result of their short exposure, it is also unlikely that the fluorochrome dyes used in our study are immunogenic in vivo. However, further optimizations are required for the in vivo application of our method. For instance, different dyes should be tested for their immunogenicity.

Taken together, chemical labeling of the gRNA and the direct transfection of RNP complexes provide a simple, safe, and efficient method that could considerably expand future therapeutic avenues for CRISPR/Cas9-mediated gene editing.

References

1. Jinek M, Chylinski K, Fonfara I, Hauer M, Doudna JA, Charpentier E. A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science*. 2012;337(6096):816-821.
2. Ran FA, Hsu PD, Wright J, Agarwala V, Scott DA, Zhang F. Genome engineering using the CRISPR-Cas9 system. *Nat Protoc*. 2013;8(11):2281-2308.
3. Dever DP, Bak RO, Reinisch A, et al. CRISPR/Cas9 β -globin gene targeting in human haematopoietic stem cells. *Nature*. 2016;539(7629):384-389.
4. Huai C, Jia C, Sun R, et al. CRISPR/Cas9-mediated somatic and germline gene correction to restore hemostasis in hemophilia B mice. *Hum Genet*. 2017;136(7):875-883.
5. Ren J, Liu X, Fang C, Jiang S, June CH, Zhao Y. Multiplex genome editing to generate universal CAR T cells resistant to PD1 inhibition. *Clin Cancer Res*. 2017;23(9):2255-2266.
6. Liu X, Zhang Y, Cheng C, et al. CRISPR-Cas9-mediated multiplex gene editing in CAR-T cells. *Cell Res*. 2017;27(1):154-157.
7. Ren J, Zhang X, Liu X, et al. A versatile system for rapid multiplex genome-edited CAR T cell generation. *Oncotarget*. 2017;8(10):17002-17011.
8. Eyquem J, Mansilla-Soto J, Giavridis T, et al. Targeting a CAR to the TRAC locus with CRISPR/Cas9 enhances tumour rejection. *Nature*. 2017;543(7643):113-117.
9. Rupp LJ, Schumann K, Roybal KT, et al. CRISPR/Cas9-mediated PD-1 disruption enhances anti-tumor efficacy of human chimeric antigen receptor T cells. *Sci Rep*. 2017;7(1):737.
10. Ye L, Wang J, Tan Y, et al. Genome editing using CRISPR-Cas9 to create the HPFH genotype in HSPCs: an approach for treating sickle cell disease and β -thalassemia. *Proc Natl Acad Sci USA*. 2016;113(38):10661-10665.

Acknowledgments

The authors thank Regine Bernhardt, Ingeborg Steiert, Karin Haehnel, and Ursula Hermanutz-Klein for the excellent technical assistance. The authors thank the FACS Core Facility of the University Hospital Tübingen, especially Kristin Bieber, Cornelia Gimmel, and Stella Autenrieth, for assistance with cell sorting. They also thank Kai Witte and D. Campana for providing 8E permeabilization reagent.

This work was supported by the Excellence Initiative of the Faculty of Medicine, University of Tübingen (J.S.), Madeleine Schickedanz Kinderkrebsstiftung (M.N.), German Cancer Consortium (P.M. and J.S.), DFG (M.K.), the Fresenius Foundation (M.K.), the Fritz Thyssen Foundation (B.D.), and the Jose Carreras Leukemia Foundation (J.S. and B.D.).

Authorship

Contribution: M.N., P.M., and J.S. made initial observations, designed the experiments, analyzed the data, and wrote the manuscript; M.N. and P.M. performed labeling of CRISPR/Cas9-gRNA RNP complexes, electroporation or transfection of cells, cell culture, treatment with UV light, FACS analysis, cell counts, cell sorting, Sanger sequencing, and LORD-Q experiments; B.D. and M.N. conducted experiments with iPSCs; M.N. and T.S. established transfection of iPSCs; Y.X. performed real-time quantitative PCR for pluripotency markers; D.A. performed image analysis of HEK293FT cells; K.S.-O. assisted with LORD-Q experiments and provided insightful comments; and M.K. and K.W. assisted with the interpretation of data and provided insightful comments.

Conflict-of-interest disclosure: The authors declare no competing financial interests.

Correspondence: Julia Skokowa, Division of Translational Oncology, Department of Hematology, Oncology, Clinical Immunology, Rheumatology and Pulmonology, University Hospital Tübingen, Otfried-Müller-Str 10, 72076 Tübingen, Germany; e-mail: julia.skokowa@med.uni-tuebingen.de.

11. Traxler EA, Yao Y, Wang YD, et al. A genome-editing strategy to treat β -hemoglobinopathies that recapitulates a mutation associated with a benign genetic condition. *Nat Med*. 2016;22(9):987-990.
12. Cyranoski D. CRISPR gene-editing tested in a person for the first time. *Nature*. 2016;539(7630):479.
13. Zhan T, Rindtorff N, Betge J, Ebert MP, Boutros M. CRISPR/Cas9 for cancer research and therapy. *Semin Cancer Biol*. 2018;S1044-579X(17)30274-2.
14. Staahl BT, Benekareddy M, Coulon-Bainier C, et al. Efficient genome editing in the mouse brain by local delivery of engineered Cas9 ribonucleoprotein complexes. *Nat Biotechnol*. 2017;35(5):431-434.
15. Xue W, Chen S, Yin H, et al. CRISPR-mediated direct mutation of cancer genes in the mouse liver. *Nature*. 2014;514(7522):380-384.
16. Yin H, Song CO, Dorkin JR, et al. Therapeutic genome editing by combined viral and non-viral delivery of CRISPR system components in vivo. *Nat Biotechnol*. 2016;34(3):328-333.
17. Wang X, Raghavan A, Chen T, et al. CRISPR-Cas9 targeting of PCSK9 in human hepatocytes in vivo—brief report. *Arterioscler Thromb Vasc Biol*. 2016;36(5):783-786.
18. Pankowicz FP, Barzi M, Legras X, et al. Reprogramming metabolic pathways in vivo with CRISPR/Cas9 genome editing to treat hereditary tyrosinaemia. *Nat Commun*. 2016;7(1):12642.
19. Liang WC, Liang PP, Wong CW, et al. CRISPR/Cas9 technology targeting Fas gene protects mice from concanavalin-A induced fulminant hepatic failure. *J Cell Biochem*. 2017;118(3):530-536.
20. Xie C, Zhang YP, Song L, et al. Genome editing with CRISPR/Cas9 in postnatal mice corrects PRKAG2 cardiac syndrome. *Cell Res*. 2016;26(10):1099-1111.
21. Gundry MC, Brunetti L, Lin A, et al. Highly efficient genome editing of murine and human hematopoietic progenitor cells by CRISPR/Cas9. *Cell Reports*. 2016;17(5):1453-1461.
22. Brinkman EK, Chen T, Amendola M, van Steensel B. Easy quantitative assessment of genome editing by sequence trace decomposition. *Nucleic Acids Res*. 2014;42(22):e168.
23. Lehle S, Hildebrand DG, Merz B, et al. LORD-Q: a long-run real-time PCR-based DNA-damage quantification method for nuclear and mitochondrial genome analysis. *Nucleic Acids Res*. 2014;42(6):e41.
24. Selvakumaran M, Lin HK, Sjin RT, Reed JC, Liebermann DA, Hoffman B. The novel primary response gene MyD118 and the proto-oncogenes myb, myc, and bcl-2 modulate transforming growth factor beta 1-induced apoptosis of myeloid leukemia cells. *Mol Cell Biol*. 1994;14(4):2352-2360.
25. Lin S, Staahl BT, Alla RK, Doudna JA. Enhanced homology-directed human genome engineering by controlled timing of CRISPR/Cas9 delivery. *eLife*. 2014;3:e04766.
26. Yang D, Scavuzzo MA, Chmielowiec J, Sharp R, Bajic A, Borowiak M. Enrichment of G2/M cell cycle phase in human pluripotent stem cells enhances HDR-mediated gene repair with customizable endonucleases. *Sci Rep*. 2016;6(1):21264.
27. Chen X, Rinsma M, Janssen JM, Liu J, Maggio I, Gonçalves MA. Probing the impact of chromatin conformation on genome editing tools. *Nucleic Acids Res*. 2016;44(13):6482-6492.
28. Zhan Q, Lord KA, Alamo I Jr, et al. The gadd and MyD genes define a novel set of mammalian genes encoding acidic proteins that synergistically suppress cell growth. *Mol Cell Biol*. 1994;14(4):2361-2371.
29. Liebermann DA, Hoffman B. Gadd45 in stress signaling. *J Mol Signal*. 2008;3:15.
30. Magimaidas A, Madireddi P, Maifrede S, Mukherjee K, Hoffman B, Liebermann DA. Gadd45b deficiency promotes premature senescence and skin aging. *Oncotarget*. 2016;7(19):26935-26948.
31. Gupta M, Gupta SK, Balliet AG, et al. Hematopoietic cells from Gadd45a- and Gadd45b-deficient mice are sensitized to genotoxic-stress-induced apoptosis. *Oncogene*. 2005;24(48):7170-7179.
32. Gupta M, Gupta SK, Hoffman B, Liebermann DA. Gadd45a and Gadd45b protect hematopoietic cells from UV-induced apoptosis via distinct signaling pathways, including p38 activation and JNK inhibition. *J Biol Chem*. 2006;281(26):17552-17558.
33. Gupta SK, Gupta M, Hoffman B, Liebermann DA. Hematopoietic cells from gadd45a-deficient and gadd45b-deficient mice exhibit impaired stress responses to acute stimulation with cytokines, myeloablation and inflammation. *Oncogene*. 2006;25(40):5537-5546.

MANUSCRIPT D (IN REVISION):

**Retinoic acid restores granulocytic differentiation in severe congenital neutropenia: Role
of GADD45 β -dependent gene demethylation**

Retinoic acid restores granulocytic differentiation in severe congenital neutropenia:

Role of GADD45 β -dependent gene demethylation

Running title: Reduced GADD45 β levels in congenital neutropenia

Perihan Mir^{1,2}, Maksim Klimiankou¹, Benjamin Dannenmann¹, Narges Aghaallaei¹, Masoud Nasri¹, Susanne Wingert³, Frederic B. Thalheimer³, Betuel Findik¹, Larissa Doll¹, Malte U. Ritter¹, Thomas Thumberger⁴, Sergei Kandabarau¹, Cornelia Zeidler⁶, Michael A. Rieger^{2,3,5}, Baubak Bajoghli¹, Karl Welte⁷, Julia Skokowa^{1,2*}

¹Department of Oncology, Hematology, Immunology, Rheumatology, University Hospital Tübingen, Tübingen, Germany; ²German Cancer Research Center (DKFZ), Heidelberg, Germany; ³Department of Medicine, Hematology/Oncology, Goethe University Frankfurt, Frankfurt am Main, Germany; ⁴Centre for Organismal Studies (COS), Heidelberg University, Heidelberg, Germany; ⁵Frankfurt Cancer Institute, Frankfurt am Main, Germany; ⁶Hematology / Oncology, SCNIR, Hannover Medical School, Hannover, Germany; ⁷The University Children's Hospital Tübingen, Tübingen, Germany

*Correspondence: Julia Skokowa, Division of Translational Oncology, Dept. of Hematology, Oncology, Clinical Immunology, Rheumatology and Pulmonology, University Hospital Tübingen, Otfried-Müller-Straße 10, Tübingen, 72076, Germany, e-mail: Julia.skokowa@med.uni-tuebingen.de

Abstract

In severe congenital neutropenia (CN) patients, elevated DNA damage in hematopoietic stem and progenitor cells (HSPCs) induces maturation arrest of granulopoiesis in the bone marrow. Here, we explored the mechanisms underlying the hypersensitivity of CN HSPCs to stress conditions. We detected markedly reduced expression of growth arrest and DNA-damage-inducible 45, beta (GADD45 β) in myeloid progenitors of G-CSF-treated CN patients compared to G-CSF-treated healthy individuals. Rescue of GADD45 β expression in CN HSPCs induced granulocytic differentiation. At the same time, CRISPR/Cas9-mediated GADD45 β knockout in HSPCs and iPSCs from healthy individuals, as well as in zebrafish embryos led to diminished granulocytic differentiation. GADD45 β is known to regulate active DNA demethylation, and we found that GADD45 β -dependent hypomethylation is essential for granulopoiesis and regulation of neutrophil functions via activation of retinoic acid signaling. Importantly, the treatment of CN HSPCs with retinoic acid receptor agonist, all-trans-retinoic acid (ATRA) restored diminished G-CSF-triggered granulocytic differentiation. Therefore, ATRA may be used for the treatment of patients with CN.

Introduction

Inherited mutations in *ELANE* (which encodes neutrophil elastase) or *HAXI* (which encodes HCLS1-associated protein X-1) cause intracellular stress responses in hematopoietic stem and progenitor cells (HSPCs) of patients with severe congenital neutropenia (CN)¹. Inappropriate activation and regulation of intrinsic stress due to inherited mutations may be a cause of the defective granulocytic differentiation of HSPCs in CN patients. Granulocyte colony-stimulating factor (G-CSF) normally induces granulocytic differentiation of HSPCs by triggering G-CSFR to activate intracellular signal transduction pathways. In CN patients, G-CSF fails to activate granulopoiesis at physiological concentrations; such patients require life-long daily treatment with very high pharmacological doses of G-CSF to maintain neutrophil numbers at a level sufficient to prevent severe infections¹. We have reported various deregulated G-CSFR signaling pathways in the HSPCs of CN patients, and propose that they may cause maturation arrest of granulopoiesis in the bone marrow²⁻¹². These alterations include the almost complete loss of LEF-1 and C/EBP α expression, diminished levels of HCLS1 and SLPI, and constitutive activation of phospho-STAT5A^{2,5,11,13}. Conversely, the G-CSF-induced activation of NAMPT/SIRT1-mediated granulopoiesis may be involved in the granulocytic maturation of CN HSPCs¹⁰. These observations suggest that the HSPCs of CN patients are continuously exposed to endo- and exogenous stress stimuli. The endogenous stress is caused by the inherited mutations and underlying signal transduction defects mentioned above, while the exogenous stress is due to life-long treatment with non-physiological doses of G-CSF. HSPCs of CN patients exhibit elevated endoplasmic reticulum (ER) stress and unfolded protein response (UPR), dysregulation of the inner mitochondrial membrane potential, and/or increased sensitivity to apoptosis due to inherited mutations in *ELANE* or *HAXI*^{6,11,14-19}. UPR and ER stress may induce DNA damage and genetic instability in HSPCs. It is known that HSPCs react to the abnormal stress activation by shifting of the differentiation program. For example, exposure of HSPCs to γ -irradiation is associated with

induction of lymphoid differentiation and inhibition of myeloid differentiation²⁰. Moreover, UPR governs the integrity of the HSC pool by depleting HSCs but not progenitors upon stress^{21,22}. An abnormal response of CN HSPCs to stress conditions, such as the persistent activation of DNA damage response pathways¹², may block granulocytic differentiation. One potential explanation for the abnormal response of CN HSPCs to stress and G-CSF is defective regulation of stress-induced factors.

Growth arrest and DNA damage-inducible 45, beta (GADD45 β), also called myeloid differentiation primary response 118 (MyD118), functions as a stress sensor in regulating the cell cycle, cell survival, and apoptosis in response to ER stress or DNA damage²³⁻²⁷. GADD45 β belongs to the evolutionarily conserved and highly homologous GADD45 protein family, along with GADD45 α and GADD45 γ ²⁸⁻³¹. The GADD45 proteins show some degree of redundancy, but also exhibit specific functions depending on the stimulus and cell type^{32,33}. Gadd45 β -deficient primary mouse embryonic fibroblasts (MEFs) proliferate slowly, accumulate increased levels of DNA damage, and exhibit signs of premature senescence^{27,34}. Gadd45 β is essential for stress-induced murine hematopoiesis, and myeloid differentiation is severely compromised in *Gadd45b* knockout mice^{27,34}. Upon TNF α activation, GADD45 β expression is induced via NF κ B³⁵. GADD45 β protects hematopoietic cells from UV-induced apoptosis by suppressing JNK signaling²⁷. We recently demonstrated that CRISPR/Cas9-mediated *GADD45B* knockout (KO) elevated the DNA damage induced by UV exposure of human HSPCs and iPSCs³⁶. GADD45 proteins have been shown to coordinate active DNA demethylation by recruiting deaminases and glycosylases to the promoter regions of target genes³⁷⁻⁴⁰. GADD45 proteins bind to a methylated promoter region proximal to an acetylated histone to recruit deaminases, which convert 5-methylcytosine to thymine to create a T:G mismatch. After that, a GADD45-recruited DNA glycosylase removes thymine from the T:G mismatch, and the thymine is replaced by an unmethylated cytosine⁴¹. Very little is known about the transcriptional programs that are regulated by the G-CSF-induced DNA

demethylation of HSPCs and during the granulocytic differentiation of HSPCs. Moreover, the regulation of hematopoiesis and granulopoiesis by GADD45 β -mediated active DNA demethylation has not yet been explored in detail.

In the present study, we examined the HSPC pool of CN patients and how stress-response regulation by GADD45 β -mediated active DNA demethylation contributes to the G-CSF-triggered granulocytic differentiation. We found that retinoic acid signaling is regulated by GADD45 β and stimulation of HSPCs with all-trans-retinoic acid (ATRA) restores defective granulocytic differentiation in CN patients.

Material and methods

Patients

Twenty-nine severe congenital neutropenia patients (19 *ELANE*-CN and 13 *HAXI*-CN patients) were used in the study. All studied CN patients were on G-CSF therapy and received G-CSF (daily dose ranged between 1.4 and 9.4 $\mu\text{g}/\text{kg}/\text{day}$) within the last 3-5 days before sampling. Bone marrow samples from patients were collected in association with an annual follow-up recommended by the Severe Chronic Neutropenia International Registry. Healthy control cells were obtained from healthy bone marrow donors. Study approval was obtained from the Ethical Review Board of the Medical Faculty, University of Tübingen. Informed written consent was obtained from all participants of this study.

CRISPR/Cas9-gRNA ribonucleoprotein (RNP) mediated *GADD45B* knockout

150 pmol of *GADD45B* CRISPR/Cas9 RNP (crRNA: GCTCGTGGCGTGCGACAACGCGG, cut site: chr19 [+2,476,389: -2,476,389], NM_015675.3:r.266, NP_056490.2:p.N11) was used for the nucleofection of 1×10^6 CD34⁺ cells or iPSCs. Nucleofection was performed with the Lonza 4D Nucleofector X using the P3 Primary Cell 4D-Nucleofector X kit (Lonza, #V4XP-3012) and program CA-137, according to the manufacturer's instructions. After nucleofection, cells were resuspended in pre-warmed culture medium. Tracking of Indels by Decomposition (TIDE) webtool⁴² and ICE webtool (Synthego) were used to estimate CRISPR/Cas9 editing efficiency.

Liquid culture differentiation of CD34⁺ cells

CD34⁺ cells (2×10^5 cells/ml) were incubated for 7 days in RPMI 1640 GlutaMAX supplemented with 10 % FBS, 1 % penicillin/streptomycin, 5 ng/ml SCF, 5 ng/ml IL-3, 5 ng/ml GM-CSF and 10 or 1 ng/ml G-CSF as previously described in Skokowa et al.¹⁰. Medium was exchanged every second day. On day 7, medium was changed to RPMI 1640 GlutaMAX supplemented with 10 % FBS, 1 % penicillin/streptomycin and 10 ng/ml or 1 ng/ml G-CSF. Medium was exchanged every second day until day 14. On day 14, cells were

analyzed by flow cytometry using following antibodies: mouse anti-human CD45 (Biolegend, #304036), mouse anti-human CD11b (BD, #557754), mouse anti-human CD15 (BD, #555402), mouse anti-human CD16 (BD, #561248) on FACSCanto II. ATRA was added to the differentiation medium in a concentration of 1 μ M and was refreshed during the medium exchange every second day.

Statistics

Statistical analysis was performed using a two-sided unpaired Student's t-test for the analysis of differences in mean values between groups. Statistical analysis of patient groups was performed using an unpaired nonparametric t-test (Mann-Whitney *U*).

Results

Elevated DNA damage levels and lymphopoiesis-biased hematopoietic differentiation of CN HSPCs

We first evaluated intrinsic stress levels in primary bone marrow hematopoietic cells of CN patients. We found that CD49f⁺ HSCs²¹ obtained from eight *ELANE*-CN and seven *HAXI*-CN patients exhibited elevated DNA damage, as assessed by γH2AX protein staining (**Figure 1A**). γH2AX may also mark proliferating cells, but since CN patients hematopoietic cells have cell cycle arrest at the G0-G1 stage, we assume, that using γH2AX staining, we measured DNA damage response.

Intrinsic stress and DNA damage may affect the hematopoietic differentiation of HSPCs, and failure of the control mechanisms may change the fate of HSPCs (e.g., via inappropriately biased differentiation)^{20,21}. We further compared the composition of bone marrow HSPCs obtained from G-CSF-treated CN patients (*ELANE*-CN n = 12, *HAXI*-CN n = 10) and healthy individuals treated (n = 5) or not (n = 7) with G-CSF. As expected, treatment of healthy individuals with G-CSF resulted in a shift towards myeloid differentiation; we observed a 5-fold increase of common myeloid progenitors (CMPs), a 3-fold increase of granulocyte-monocyte progenitors (GMPs), and a 10-fold decrease of multi-lymphoid progenitors (MLPs). In contrast, the hematopoietic differentiation of CN patients was substantially shifted from myeloid (5- and 10-fold decreases of CMPs for *ELANE*-CN and *HAXI*-CN, respectively) to lymphoid (more than 100-fold increases of MLPs for *ELANE*-CN and *HAXI*-CN) differentiation, and we detected reductions in the numbers of GMPs (3-fold for both *ELANE*-CN and *HAXI*-CN) (**Figure 1B**). These data suggest that the integrity control of HSPCs from CN patients has defects that may alter the response to G-CSF and ultimately lead to severe neutropenia.

Markedly diminished expression of the stress response protein, GADD45 β , in myeloid progenitor cells of CN patients

The DNA damage and stress response protein, Gadd45 β , is known to regulate stress-induced granulopoiesis in mice⁴³. GADD45 β is also expressed in human HSPCs with increasing levels during myeloid commitment and maximum expression in mature neutrophilic granulocytes, as judged by *Bloodspot* database⁴⁴ (**Supplemental Figure 1A**). We compared GADD45 β mRNA expression levels in bone marrow promyelocytes obtained from G-CSF-treated CN patients (n = 7), healthy controls treated (n = 2) or not (n = 3) with G-CSF for 3 days, and patients with metabolic neutropenia (n = 3) chronically treated with G-CSF. Interestingly, we found that G-CSF induced GADD45 β mRNA expression in CD33⁺ cells of healthy controls and patients with metabolic neutropenia, but not in those of CN patients (**Figure 1C**). Moreover, a DUOLINK assay revealed that nuclear GADD45 β protein was highly expressed in CD33⁺ cells of G-CSF-treated healthy controls, but it was almost entirely absent from the nuclei of promyelocytes from CN patients (**Figure 1D**). We did not detect any correlation between GADD45 β expression levels and response to G-CSF therapy in CN patients (data not shown).

Ectopic expression of GADD45 β in CD34⁺ cells of *ELANE*-CN patients restores diminished granulocytic differentiation

We tested whether restoration of the diminished GADD45 β expression in primary CD34⁺ HSPCs of CN patients would affect their granulocytic differentiation *in vitro*. We used a Venus⁺-lentiviral vector to overexpress the GADD45 β cDNA in CD34⁺ HSPCs obtained from two *ELANE*-CN patients, and differentiated the transduced cells towards neutrophils in liquid culture (**Figure 2A**). Transduction efficiency of control virus was around 60 % and of GADD45 β cDNA-containing virus around 10 %. Indeed, we detected dramatically elevated

amounts of CD11b⁺CD15⁺, CD11b⁺CD16⁺, and CD15⁺CD16⁺ cells within the Venus⁺CD45⁺ cell population of GADD45 β -transduced *ELANE*-CN CD34⁺ HSPCs, compared to control samples (**Figure 2B**). These data suggest that GADD45 β induces and supports granulocytic differentiation in CN CD34⁺ cells.

Gadd45 β promotes differentiation of mouse LT-HSCs

To evaluate the effects of GADD45 β overexpression on the *in vitro* myeloid differentiation of primary HSPCs, we transduced mouse long-term repopulating HSCs (LT-HSCs) with ectopic *Gadd45b* and differentiated the cells into myeloid lineage in liquid culture *in vitro* and analyzed differentiated cells at different time points by FACS. We found that HSCs transduced with Gadd45 β displayed accelerated and enhanced myeloid differentiation, as assessed by FACS: percentage of immature HSPCs rapidly decreased, while the number of granulocyte-monocyte progenitors (GMPs) and mature myeloid cells increased over time, as compared to control transduced cells (**Figure 2C,D**). Colony forming unit (CFU) assay and CFU re-plating assay revealed that HSCs expressing ectopic *Gadd45 β* have myeloid-biased colony forming potential and generate more mature cells with lower re-plating capacity in comparison to control virus expressing cells (**Figure 2E**).

GADD45 β KO inhibits the granulocytic differentiation of human hematopoietic cells

To study the impact of GADD45 β deficiency on the granulocytic differentiation of human HSPCs, we knocked out GADD45 β in human iPSCs and CD34⁺ HSPCs of healthy individuals using GADD45 β -specific CRISPR/Cas9-gRNA RNPs, and differentiated the gene-edited cells towards neutrophils *in vitro*. The editing efficiency of gene-modified CD34⁺ HSPCs was around 44-85% (**Supplemental Figure 2A-C and data not shown**). We found that GADD45 β knockout abrogated granulocytic differentiation in liquid culture, as assessed

by diminished percentages of CD11b⁺CD15⁺, CD11b⁺CD16⁺, and CD15⁺CD16⁺ cells in the CD45⁺ cell populations of GADD45 β -KO cells compared to wild type control cells (**Figure 3A**).

To study the role of complete GADD45 β knockout on granulocytic differentiation of human HSPCs *in vitro*, we tested granulocytic differentiation of pure GADD45 β knockout iPSC clones³⁶. In line with the observations from primary HSPCs, no CFU-G colonies were generated from GADD45 β ^{-/-} iPSC-derived CD34⁺ HSPCs, compared to healthy control iPSCs (**Figure 3B**). Moreover, during embryoid body (EB)-based hematopoietic differentiation, we observed increased amounts of immature cells on day 14 in culture and strongly diminished amounts of mature neutrophils on day 28 in culture for GADD45 β ^{-/-} iPSCs compared to healthy control iPSCs (**Figure 3C, Supplemental Figure 3A**). This was confirmed by morphological analysis of Wright-Giemsa-stained cytospin preparations of differentiated cells collected on days 21 and 28 of culture. We detected dramatically increased amounts of immature hematopoietic cells and far fewer mature neutrophils in the GADD45 β -deficient samples compared to healthy control iPSCs (**Figure 3D, E and Supplemental Figure 3B**). The presented data suggest an essential role for GADD45 β in granulopoiesis.

Zebrafish *gadd45bb* is required for neutrophil development

We next examined whether GADD45 β is required for *in vivo* granulopoiesis. We used zebrafish as an *in vivo* model system, because the mechanisms underlying hematopoiesis are evolutionarily conserved in vertebrates^{45,46}. Two orthologs of human GADD45 β , *gadd45ba* and *gadd45bb*, are present in the genome of zebrafish, but only the *gadd45bb* paralog is expressed in the embryonic hematopoietic site at 1 day post-fertilization (dpf, data not shown)⁴⁷. To test whether *gadd45bb* is required for granulopoiesis, we performed transient CRISPR/Cas9 targeting of the *gadd45bb* gene in the transgenic line, Tg(*mpo:gfp*), in which

expression of green fluorescent protein (GFP) is driven by the neutrophil-specific myeloperoxidase (*mpo*) promoter⁴⁸. Compared with non-injected siblings, *gadd45bb* crispants showed significantly fewer GFP⁺ cells in the caudal hematopoietic tissue (akin to the mammalian fetal liver) at 3 dpf (**Figure 4A**). In another set of experiment, injection of Cas9 alone did not change the neutrophil count in embryos (**Supplemental Figure 4A**). We next asked whether the reduced number of neutrophils caused by mutation of *gadd45bb* could be rescued by *csf3a* mis-expression. To test this, we used a heat-inducible system⁴⁹ to ectopically express *csf3a* in zebrafish embryos at 1 dpf. Consistent with a previous report⁵⁰, we found that ectopic expression of *csf3a* increased the number of *mpo*-expressing cells at the hematopoietic site of wild-type embryos (**Figure 4B**). When *csf3a* was induced in *gadd45bb* crispants, the number of *mpo*-expressing cells increased to a level comparable with that seen in the uninjected group (**Figure 4B**). Together, these results suggest that *gadd45bb* plays a crucial role in zebrafish granulopoiesis and *csf3a* compensates for the lack of *gadd45bb* in this process.

GADD45 β expression is directly regulated by C/EBP α

To investigate the mechanism underlying the G-CSF-mediated activation of GADD45 β expression, we performed *in silico* analysis of the GADD45 β gene promoter using Genomatix software. This analysis identified four putative binding sites for C/EBP transcription factors (**Figure 4C**). We previously reported that C/EBP α levels were markedly reduced in myeloid progenitor cells of CN patients^{2,7-9}. Therefore, we tested for possible C/EBP α -mediated regulation of GADD45 β expression that might be abolished in C/EBP α -deficient CN progenitors. We performed ChIP assays with an anti-C/EBP α antibody in lysates of THP1 cells expressing both C/EBP α and GADD45 β , and identified three C/EBP α binding sites on the GADD45 β gene promoter, located -655 bp, -214 bp, and -77 bp from the

ATG (**Figure 4C and Supplemental Figure 5A**). We further generated a reporter gene construct by cloning a GADD45 β promoter fragment (1.2 kb from the ATG) into the pGL4.10 [*luc2*] vector, and performed dual luciferase reporter gene assays in HEK293T cells co-transfected with the *C/EBPA* cDNA and a GADD45 β reporter construct. Indeed, we found that C/EBP α specifically and dose-dependently activated the GADD45 β promoter (**Figure 4D**).

In HSPCs, GADD45 β regulates mRNA expression programs that are essential for granulopoiesis and responsible for neutrophil activation

To identify intracellular signaling pathways regulated by GADD45 β in G-CSF treated human primary bone marrow HSPCs, we cultured WT or GADD45 β -KO CD34⁺ HSPCs obtained from three healthy donors for 72 hours with or without 50 ng/ml G-CSF, and performed RNA sequencing (**Figure 5A**). For GADD45 β KO, we electroporated cells with GADD45 β -specific CRISPR/Cas9-gRNA RNP complexes. The efficiency of GADD45 β KO was up to 84%, as assessed by Sanger sequencing and ICE assay (**Supplemental Figure 2A-C**). We found that 94 genes were differentially expressed between G-CSF-treated control and G-CSF-treated GADD45 β -KO cells (adjusted *p*-value < 0.05) (**Figure 5B, Supplemental Table 1**). Strikingly, most of the differentially expressed genes regulate myeloid differentiation (e.g., *RXRA*, *FGR*, *TFPI*, *HK3*, and *NLRP12*), are induced by DNA damage in G-CSF-treated cells (e.g., *BATF*), or are essential players in neutrophil adhesion, migration (*ITGAM*, *SIGLEC5*, *CX3CR1*, *CXCR1*, *FPR1*, *FPR2*, and *CORO1A*), and activation (*NCF2*, *MYL6*, and *NLRP12*). Previously published microarray data² indicate that GADD45 β -regulated genes were also up-regulated in bone marrow CD33⁺ myeloid progenitor cells of G-CSF-treated healthy individuals, but not in the same cell population of CN patients chronically treated with G-CSF (**Figure 5C**).

Using Genomatix pathway analysis of significantly upregulated genes in GADD45 β -WT and GADD45 β -KO groups treated with G-CSF, we identified response to cytokines, response to stress, regulation of intracellular signal transduction, regulation of ERK1/2 signaling, neutrophil activation, degranulation, and migration as being among the top significant biological processes differentially regulated by G-CSF via GADD45 β (**Figure 5D**). Analysis of associated signaling pathways revealed that integrin-, inflammatory-, CXCR1-, FPR1-, and SRC kinase signaling were upregulated by G-CSF via GADD45 β (**Figure 5E**). We also identified associations with various disease phenotypes, including leukocyte disorders, leukopenia, agranulocytosis, neutropenia, phagocyte bactericidal dysfunctions, and (importantly) AML/MDS (**Figure 5F**).

iRegulon analysis of the significantly enriched transcription factor binding motifs within significantly upregulated genes^{51,52} revealed that genes with hematopoietic- or myeloid-specific transcription factor binding motifs were strongly enriched in G-CSF-treated control CD34⁺ HSPCs compared to GADD45 β -deficient G-CSF-exposed samples. In particular, we identified enrichment of the binding motifs for RXRA/RARA, SPI1, and C/EBP β in G-CSF-treated control cells but not in the corresponding GADD45 β -KO samples (**Figure 5G, Supplemental Figure 6A**).

Together, these data clearly demonstrate that GADD45 β plays important roles in G-CSF-triggered granulocytic differentiation and the regulation of neutrophil functions.

GADD45 β -dependent active gene demethylation during G-CSF-triggered granulocytic differentiation of HSPCs

GADD45 proteins, including GADD45 β , regulate active DNA demethylation by promoting the coupling of deamination, glycosylation, and the recruitment of the DNA demethylation machinery to specific genomic loci^{37-40,53}. We therefore investigated whether

the G-CSF-triggered granulocytic differentiation of HSPCs requires active gene demethylation mediated by GADD45 β . We assessed methylation changes in control and GADD45 β -KO CD34⁺ HSPCs treated or not with G-CSF, using an Infinium MethylationEPIC array (Illumina, Inc.) and subsequent data analysis with the R packages, *minfi*, *limma*, and *DMRcate*⁵⁴⁻⁵⁷.

G-CSF treatment of control CD34⁺ HSPCs triggered robust changes of DNA methylation: we detected 13.516 hypomethylated and 6.236 hypermethylated CpGs (adjusted p -value < 0.05) (**Figure 6A, left part, Supplemental Table 2**). In contrast, GADD45 β KO markedly attenuated the G-CSF-induced demethylation: we detected only 8.440 hypomethylated and 4.674 hypermethylated CpGs (χ^2 test: $p = 0.0001$, **Figure 6A, right part**). We found that 5.704 CpGs and 628 were hypomethylated in G-CSF-treated WT and GADD45 β -KO cells, respectively (**Figure 6B**). In contrast, hypermethylation was seen for 2.878 sites in WT cells and 1.316 sites in GADD45 β -KO cells upon G-CSF treatment (**Figure 6B**). The highest number of differentially methylated CpGs among the studied groups can be assigned to open sea regions (χ^2 test: $p < 0.05$, **Figure 6C, left**). Gene-specific annotations of CpGs show that 5'UTR and 3'UTR are the preferential places for differentially methylated CpGs (χ^2 test: $p < 0.05$, **Figure 6C, right**).

Consistent with the gene expression changes discussed above, Genomatix analysis of methylation data revealed that GADD45 β and G-CSF significantly regulated various biological processes, including myeloid leukocyte differentiation, neutrophil activation, cell migration, and chemotaxis (**Supplemental Figure 7A**). G-CSF-dependent signaling pathways regulated by GADD45 β -mediated hypomethylation were identified, including hematopoietic cell kinase, SRC, CXCR1, G-CSFR, and RARA signaling (**Supplemental Figure 7B**). Again, diseases associated with defective G-CSF-triggered hypomethylation upon GADD45 β KO

were immune system diseases, AML, leukocyte disorders, leukopenia, pre-leukemia, and neutropenia (**Figure 6D**).

The top genes with differentially methylated regions (DMRs) were *RXRA*, *MEFV*, *FPR2*, *CSF3R*, and several genes encoding neutrophil granule proteins (e.g., *ELANE*, *MPO*, *AZU1*, and *CTPSG*) (**Figure 6E and Supplemental Figure 8A**).

Genomic locus overlap enrichment analysis⁵⁸ showed that the DMRs were associated with top-ranked datasets representing regions occupied by H3K9K14ac in ATRA-treated NB4 cells, the RARA binding sites in the PML-RAR α zinc-inducible cell line, UPR9, and SPI-associated regions (**Supplemental Table 3**).

Activation of the retinoic acid signaling by ATRA rescued defective granulocytic differentiation of CN HSPCs

We further evaluated whether the GADD45 β -mediated regulation of active DNA demethylation correlated with the mRNA expression differences that we observed between control and GADD45 β -KO cells treated with G-CSF. We selected genes that exhibited significantly reduced mRNA expression in G-CSF-treated GADD45 β -KO cells but not in G-CSF-exposed control samples, and assessed their association with differentially methylated CpGs. We found that G-CSF regulates the methylation and expression of genes essential for myeloid differentiation or neutrophil activation (e.g., *RXRA*, *MEFV*, *CXCR1*, *FPR2*, and *SERPINA1*) via GADD45 β (**Figure 7A**). Hypomethylated CpGs in these genes were located in regions with regulatory activity, such as the transcriptional start sites (TSS), first exon, and untranslated regions (UTRs). Genomatix analysis revealed that GADD45 β -deficient G-CSF-treated HSPCs exhibited dysregulation of biological processes and signal transduction cascades that regulate granulopoiesis and neutrophil functions, as well as those associated with disease entities, such as phagocyte bactericidal dysfunctions, neutropenia, leukocyte disorders, and myeloid leukemia (**Figure 7B,C, Supplemental Figure 9A**).

We identified the retinoic acid pathway as one of the top hits regulated by GADD45 β during G-CSF-triggered *in vitro* granulopoiesis (**Figures 4 and 5**). We, therefore, assumed that in the absence of GADD45 β activation in CN patients HSPCs, the retinoic acid pathway could not be induced leading to diminished granulocytic differentiation. We evaluated whether activation of the retinoic acid signaling by ATRA will bypass missing activation of GADD45 β downstream of G-CSF. We treated CN HSPCs with ATRA or DMSO during the *in vitro* liquid culture granulocytic differentiation. Strikingly, in all treated CN patient's HSPCs, (five CN patients: *ELANE*-CN n = 3 and *HAXI*-CN n = 2), we observed markedly improved granulocytic differentiation, as assessed by a drastic reduction of myeloblasts/promyelocytes and increase of neutrophils on cytopsin slides of cells assessed on day 14 of culture (**Figure 7D, E**).

These data strongly argue that both granulocytic differentiation and neutrophil functions are regulated by GADD45 β -mediated DNA demethylation upon G-CSF treatment of CD34⁺ cells via activation of retinoic acid signaling. Moreover, the treatment of CN HSPCs with ATRA rescued the GADD45 β deficiency and induced granulopoiesis.

Discussion

Granulopoiesis is a highly complex and coordinated process of differentiation of HSPCs into mature neutrophils. It is controlled by a network of regulatory mechanisms and factors, including transcription factors, epigenetic modifiers, receptors, cytokines, and protein kinases. In the present study, we demonstrated that GADD45 β plays an essential role in the G-CSF-triggered granulopoiesis. We were able to link diminished GADD45 β expression in HSPCs to CN, which is an inherited bone marrow failure syndrome with severe neutrophil differentiation defects. Rescue experiments in CN HSPCs confirmed that GADD45 β plays an important role in G-CSF-mediated granulopoiesis. Moreover, ectopic expression of Gadd45 β in murine HSPCs resulted in the elevated myelopoiesis. These data are in line with previous observations that neutrophils exhibit defects in their numbers and functions in *Gadd45 β ^{-/-}* mice under stress conditions^{27,34,43,59}. We herein observed elevated DNA damage in bone marrow HSPCs of G-CSF treated CN patients. We and others previously demonstrated that inherited CN-associated mutations deregulate intracellular signaling cascades^{1-11,60}. This may lead to replicative stress, DNA damage, and ultimately deregulated commitment of HSPCs resulting in defective granulopoiesis. GADD45 β expression is activated upon stress (e.g. DNA damage or cytokine exposure^{27,34,43}), regulating gene expression programs specific for granulopoiesis and neutrophil functions. In CN patients, in contrast, GADD45 β expression is not activated, and granulopoiesis is therefore not initiated.

Severe defects in the expression levels of the transcription factors, LEF-1 and C/EBP α , were previously described by our group in myeloid cells of CN patients harboring either *ELANE* or *HAX1* mutations and were suggested to explain the defective granulopoiesis seen in this disorder². LEF-1 binds to and activates the *C/EBPA* gene promoter under normal circumstances². We herein show that C/EBP α binds to and activates *GADD45B* promoter. Therefore, we hypothesized that GADD45 β is not activated in CN HSPCs due to the

downregulation of LEF-1 and C/EBP α . Since both investigated CN patient groups (with *ELANE* or *HAXI* mutations) showed defective GADD45 β expression upon G-CSF treatment, we assume that this results from common dysfunctional pathways downstream of the inherited mutations. The mechanism of GADD45 β deregulation in CN HSPCs downstream of e.g. UPR triggered by mutated *ELANE* or mitochondrial dysfunctions triggered by *HAXI* mutations remains to be investigated. We previously described compensatory G-CSF/NAMPT/SIRT1-mediated activation of C/EBP β -triggered granulopoiesis in CN patients in the absence of LEF-1 and C/EBP α ¹⁰. Based on these observations, we assume that G-CSF - C/EBP β - GADD45 β axis is not active in CN, and G-CSF induces granulopoiesis via NAMPT/SIRT1/C/EBP β pathway, that is GADD45 β independent. Future work will be needed to examine why C/EBP β cannot induce GADD45 β expression in CN HSPCs.

The GADD45 protein family consists of three members that have similar structures and partially redundant functions⁶¹. The expression levels of the other GADD45 proteins are not affected in CN (data not shown), but the presence of these proteins does not appear to compensate for the abrogation of GADD45 β . This might reflect that GADD45 β has tissue-specific functions and/or a specific affinity for particular stimuli, such as G-CSFR signaling. *Gadd45b*^{-/-} mice also have a neutropenic phenotype with markedly diminished stress-induced neutrophil functions^{27,34,43,59}. The role of GADD45 β in G-CSF triggered granulopoiesis appears to be evolutionarily important, given our observation that *gadd45bb*-deficient zebrafish also exhibit markedly reduced neutrophil numbers. Since *Elane*- or *Hax1*-KO mice do not develop neutropenia, we suggest that *gadd45bb*-KO zebrafish could be used as an experimental model to investigate neutropenia associated with these mutations.

The signal transduction pathways that are specifically regulated by GADD45 β in human HSCs have not yet been fully elucidated. NF κ B, p38, and MKK7 signaling have been described to be regulated by GADD45 β in mice^{27,34,43,59,62}. Our RNA-seq analysis

demonstrated that G-CSF induces e.g. *RXRA*, *ITGAM* (CD11b), *FPR1/2*, *MEFV*, *BATF*, and *FGR* in a GADD45 β dependent manner. Intriguingly, these factors were also upregulated in bone marrow myeloid progenitor cells from G-CSF-treated healthy donors, but not in those from CN patients. All these genes are key hematopoietic- and myeloid-lineage factors. For example, endogenous retinoic acid (RA) receptors are upregulated in mouse HSPCs upon G-CSF treatment⁶³ and RARA/RXRA signaling plays an important role during granulopoiesis^{64,65}. Vitamin A and D both require RARA/RXRA for proper intracellular signal transduction and, interestingly, our RNA seq and methylation analyses identified disease entities associated with neutropenia and other immunological bone marrow failure syndromes, as well as vitamin A or D deficiency among the diseases significantly correlated with GADD45 β deficiency.

GADD45 β regulates active DNA demethylation that involves the base excision repair^{38,66,67}. The GADD45 proteins remove 5-hydroxymethylcytosine (5HMC) and 5-methylcytosine (5MC) and act as a scaffold for recruited cytidine deaminases or DNA glycosylases^{39,40,68-70}. GADD45 proteins are also essential for targeting promoters for 5HMC/5MC removal through direct interactions with nuclear hormone receptors⁷¹ or hyperacetylated nucleosomes²⁵. We found here, that in GADD45 β -KO cells treated with G-CSF, the number of hypomethylated CpGs was markedly reduced, whereas much less alteration was seen for hypermethylated CpGs. Consistent with the profiles of differentially expressed genes, our analysis of DNA hypomethylated sites revealed an association with defects in myeloid cell differentiation, hematopoiesis, neutrophil activation, regulation of cell adhesion, chemotaxis, and phagocytosis in these cells. The G-CSFR, RARA/RXRA, vitamin D, and FPR1 signaling pathways were among top regulated pathways, and neutropenia was among the significant disease entities correlated with GADD45 β -dependent methylation changes in G-CSF-stimulated HSPCs. By comparing our RNA seq data with the observed DNA methylation changes, we were able to identify candidate factors whose expression

levels are regulated by GADD45 β -mediated active demethylation. These factors included *RXRA*, *MEFV*, *FPR2*, *CSF3R* and genes encoding the neutrophil granule proteins, NE, AZU1, MPO, and CTSG.

In addition, genomic locus overlap enrichment analysis showed that DMRs were associated with the top-ranked regions occupied by H3K9K14ac in ATRA-treated NB4 cells, RARA binding sites in the PML-RAR α zinc-inducible cell line, UPR9, and *SPII*/PU.1 binding sites. H3K9K14ac marks active inducible gene regulatory elements establishing a chromatin conformation that is compatible with transcription⁷². Martens et al.⁷³ demonstrated that ATRA treatment of acute promyelocytic leukemia (APL) cells strongly increases H3K9K14ac at PML-RARA/RXR target sites. The increased methylation observed in most DMRs upon GADD45 β KO might affect the induction of local DNA histone modifications, perhaps accounting for the deregulation of the myeloid differentiation program in GADD45 β -KO HSPCs.

Strikingly, to functionally validate our observations from the RNA-seq and methylation array analyses, we activated retinoic acid signaling in GADD45 β -deficient HSPCs of CN patients. Indeed, we could circumvent the induction of GADD45 β by treating CN HSPCs with ATRA to trigger granulopoiesis by activating retinoic acid signaling. ATRA is an FDA-approved drug for the treatment of human promyelocytic leukemia⁷⁴ and myelodysplastic syndrome⁷⁵ and could offer a potential therapy option for CN patients.

Gene expression and methylation changes detected in our patients were associated with pre-leukemia, MDS, and AML. Since CN is a pre-leukemic bone marrow syndrome, it would be interesting to investigate the role of diminished GADD45 β expression in leukemia development in CN. There is some controversy regarding the role of GADD45 β in tumorigenesis and leukemogenesis, with GADD45 proteins reportedly acting as tumor suppressors in some tumors (e.g., breast and prostate cancer), but as tumor-promoting factors in others (e.g., colorectal cancer)⁷⁶⁻⁸⁷. In leukemia, GADD45 β was reported to induce

apoptosis in AML³⁰, and GADD45 β deficiency was shown to accelerate BCR-ABL-driven chronic myeloid leukemia⁸⁸.

In summary, we herein show that GADD45 β plays essential roles in G-CSF-triggered granulopoiesis by regulating the active demethylation and expression of genes essential for these processes, including components of the retinoic acid signaling (**Figure 7E**). G-CSF activates GADD45 β expression via C/EBP α and since the myeloid cells of CN patients are deficient for C/EBP α , GADD45 β and downstream retinoic acid signaling are not activated leading to severe diminished granulocytic differentiation (**Figure 7E**). Strikingly, ATRA treatment rescues defective granulocytic differentiation of CN HSPCs. These data clearly demonstrate how investigations combining basic and translational science may lead to the identification of novel therapeutic opportunities. In our case, patients with severe congenital neutropenia, especially patients who do not respond to the standard G-CSF therapy may benefit from this scientific approach.

Figure legends

Figure 1. Diminished GADD45 β expression in line with elevated gH2AX levels in HSCs of CN patients may lead to “decision shift” from myeloid to lymphoid progenitors

A, Bone marrow mononuclear cells were stained for surface markers to detect human stem and progenitor cell subsets, gH2AX was included into the panel (healthy controls treated with G-CSF $n = 4$, *ELANE*-CN $n = 8$, *HAXI*-CN $n = 7$). The cell populations are as followed: multipotent progenitors (MPP), common myeloid progenitors (CMP), granulocyte-macrophage progenitors (GMP), multi-lymphoid progenitors (MLP) and B- and NK-cell progenitors (B/NK). Data represent means \pm SEM; analyzed using unpaired nonparametric t-test (Mann-Whitney U): *, $p < 0.05$. **B**, The composition of HSPCs was assessed by multicolor FACS panels (healthy controls ($n = 7$), healthy control + G-CSF ($n = 5$), *ELANE*-CN ($n = 12$), *HAXI*-CN ($n = 10$)). Data represent means \pm SEM; analyzed using unpaired nonparametric t-test (Mann-Whitney U): *, $p < 0.05$; **, $p < 0.01$; ***, $p < 0.001$; ns, not significant. **C**, GADD45 β expression was measured by qRT-PCR in CD33 $^{+}$ progenitors of healthy individuals ($n = 3$), healthy individuals treated with G-CSF ($n = 2$), patient with metabolic neutropenia (MN, $n = 3$) and congenital neutropenia patients (CN, $n = 7$). Data represent means \pm SEM from duplicates. **D**, GADD45 β protein expression was investigated by DUOLINK proximity ligation assay in healthy donor CD33 $^{+}$ cells treated with G-CSF and *ELANE*-CN patients CD33 $^{+}$ cells. Representative images are depicted.

Figure 2. Ectopic expression of GADD45 β rescues neutrophilic differentiation of CD34 $^{+}$ HSPCs of CN patients and accelerates granulopoiesis in mice

A, Schematic of the experimental procedure: CD34 $^{+}$ HSPCs of two *ELANE*-CN patients were expanded *in vitro*, transduced with lentivirus containing GADD45 β construct or a control virus with the fluorescent marker Venus and differentiated in liquid culture for 14 days. **B**, Neutrophilic surface marker expression was measured by FACS. Only Venus $^{+}$ CD45 $^{+}$ cell

fraction was considered for analysis. In both patients the ectopic GADD45 β expression led to the increased percentage of CD11b⁺CD15⁺, CD11b⁺CD16⁺ and CD15⁺CD16⁺ cells significantly. Data represent means \pm S.D.; *, $p < 0.05$; **, $p < 0.01$; ***, $p < 0.001$. **C,D**, Murine LT-HSCs were transduced with *Gadd45b*, differentiated for 8 days in liquid culture. **C**, Surface marker expression was evaluated at different time points of culture by FACS, as described in material and methods. Transduced cells were gated for immature HSPCs, GMP-like cells, mature neutrophils and macrophages, means \pm S.D. of 3 independent experiments; *, $p < 0.05$; **, $p < 0.01$. Representative FACS images of cells on day 8 of differentiation stained for granulocyte/monocyte surface markers CD11b and CD16/32 are depicted in **D**. **E**, CFU assay of mouse LT-HSCs transduced with either control or *Gadd45b* virus under permissive cytokine conditions, as described in material and methods. Colonies generated from transduced cells were counted, data represent means \pm S.D. of 3 independent experiments (left image). 2×10^4 cells isolated from CFUs were re-plated in the secondary CFUs, as described in material and methods ($n = 1$, right image).

Figure 3. Diminished *in vitro* granulocytic differentiation of HSPCs and iPSCs upon GADD45 β knockout

A, Bone marrow CD34⁺ HSPCs were nucleofected with *GADD45B*-specific CRISPR/Cas9-gRNA RNP. Granulocytic differentiation was evaluated in liquid culture differentiation for 14 days. Neutrophilic surface marker expression was assessed by FACS. **B-E**, Evaluation of granulocytic differentiation of GADD45 β -deficient iPSCs; **B**, CFU assay; **C**, EB-based myeloid differentiation; **D,E**, morphological assessment of differentiated iPSCs on day 28 of culture was conducted on Wright-Giemsa stained cytospin preparations (MB/ProM: Myeloblasts/Promyelocytes; Myelo/Meta: Myelocytes/Metamyelocytes; Band/Seg: Band/Segmented cells; MF: Macrophages); representative cytospin images (60 X magnification) of

three experiments are depicted in **E**. **A-E** show data from 3 experiments. Data represent means \pm S.D.; *, $p < 0.05$; **, $p < 0.01$; ***, $p < 0.001$.

Figure 4. Knockout of *gadd45bb* in zebrafish resulted in drastically low neutrophil numbers

A, Quantification of GFP⁺ neutrophils in uninjected controls and *gadd45bb* sgRNAs injected Tg(mpo:gfp) zebrafish embryos at 3 dpf. Each dot represents an individual embryo at 3 dpf (N). Data are mean \pm S.D., ****, $p < 0.0001$. Representative images are shown in the right panel. **B**, Number of mpo-expressing cells in the hematopoietic site of embryos in different experimental setups at 1 dpf is indicated. Each dot represents an individual embryo at 1 dpf (N). Data are mean \pm S.D., **** $p < 0.0001$, ** $p < 0.001$, n.s., not significant. Representative images are shown in the right panel. **C**, In silico promoter analysis of *GADD45B* promoter region by Genomatix with four predicted binding sites for C/EBP transcription factors. C/EBP α binding sites on *GADD45 β* promoter were analyzed by chromatin immunoprecipitation. C/EBP α binding was confirmed for binding site 1, 3 and 4. **D**, Binding of C/EBP α to promoter region (1.2 kb) of *GADD45 β* was assessed by dual luciferase reporter gene assay with different concentrations of C/EBP α -expressing construct, representative data from 3 experiments. Data represent means of 3 experiments \pm S.D.; *, $p < 0.05$; **, $p < 0.01$; ***, $p < 0.001$; ****, $p < 0.0001$.

Figure 5. CRISPR/Cas9-mediated knockout of *GADD45 β* in CD34⁺ HSPCs repressed G-CSF-triggered granulocytic differentiation program

A, Schematic of the experimental approach. **B**, Comparison between wild-type and *GADD45 β* -KO CD34⁺ HSPCs both treated with G-CSF revealed 94 significantly differentially expressed genes (adjusted p -value < 0.05), most of them having roles in myeloid differentiation (indicated with *). **C**, candidate genes of which the expression was regulated

similarly in current RNA seq data and in microarray experiments conducted and published previously on healthy individual and CN patient promyelocytes treated with G-CSF. **D-F**, Genomatix-based gene ontology analysis using a list of significantly upregulated genes in G-CSF treated GADD45 β WT HSPCs, in comparison to GADD45 β -KO HSPCs revealed top significantly enriched biological processes (**D**), signal transduction pathways (**E**) and disease entities (**F**) that are GADD45 β dependent. **G**, iRegulon transcription factor motif enrichment analysis using a list of significantly upregulated genes in G-CSF treated GADD45 β WT HSPCs, in comparison to GADD45 β -KO HSPCs, revealed enrichment of genes (Normalized Enrichment Score (NES) > 3.0 and Area Under the cumulative Recovery Curve (AUC) > 0.03) with myeloid-specific transcription motifs in control but not in knockout cells.

Figure 6. GADD45 β is essential for active DNA demethylation in response to G-CSF

A-F, DNA methylation of G-CSF-treated control or GADD45 β -KO cells was evaluated using methylation EPIC array and was analyzed using R packages *minfi* and *limma*. **A**, G-CSF treatment induced the hypomethylation of 13.516 CpGs and hypermethylation of 6.236 CpGs in control cells (left). In *GADD45B*-KO cells (right), G-CSF induced hypomethylation of 8.440 and hypermethylation of 4.674 CpGs. **B**, Venn diagram of hypo- or hypermethylated CpGs in indicated groups. **C**, Evaluation of the numbers of differentially methylated CpGs per gene region (left) revealed that top significant differentially methylated CpGs (adjusted p -value < 0.05) were located mainly in open sea and island regions of genes. N: upstream and S: downstream of CpG island, shore: 0-2 kb from island, shelf: 2-4 kb from island. Gene-oriented annotation of CpGs (CpG feature, right) revealed that the most significant differences preferentially located in intragenic regions, 5'UTR and 3'UTR. Significance was calculated with the χ^2 test: $p < 0.05$. **D**, Genomatix-based gene ontology analysis using a list of significantly differentially hypomethylated CpGs in G-CSF treated GADD45 β -WT HSPCs, in comparison to GADD45 β -KO HSPCs revealed top significantly enriched disease

entities that are GADD45 β dependent. **E**, Differentially methylated regions (DMRs) in top genes that are hypomethylated upon G-CSF but not in GADD45 β -KO cells are shown (*minfdr* <10⁻⁶).

Figure 7. Regulation of G-CSF-triggered granulopoiesis by GADD45 β -mediated gene demethylation

A, List of significant genes differentially expressed and methylated (adjusted *p*-value < 0.05) with differentially hypomethylated CpGs (adjusted *p*-value < 0.05) upon G-CSF treatment of healthy control HSPCs, as compared to GADD45 β -KO cells. For all genes, CpGs are located in essential gene regulatory regions. **B,C**, Genomatix-based gene ontology analysis using the list of genes from **A** revealed top significantly enriched biological processes (**B**) and diseases entities (**C**) that are GADD45 β dependent. **D, E**, Five CN patient HSPCs (*HAXI*-CN n =2 and *ELANE*-CN n = 3) were treated with either DMSO or ATRA during *in vitro* liquid culture granulocytic differentiation. Bar chart of the morphological analysis (**D**) and representative cytopsin images (**E**) of the cytopsin slides of cells derived on day 14 of culture are presented, MB/ProM: Myeloblasts/Promyelocytes; Band/Seg: Band/ Segmented cells, 60 X magnification. Data represent percentages of respective populations from five CN patients \pm S.D.; **, *p* < 0.01. **F**, Mechanism of “maturation arrest” of G-CSF-triggered granulopoiesis in CN patients via diminished GADD45 β -mediated regulation of granulopoiesis-activating genes, including components of the RA signaling. Treatment of cells with ATRA restored granulopoiesis in CN.

Acknowledgments

This work was supported by the J. Carreras Leukemia Foundation (J.S., P.M.), Madeleine Schickedanz Kinderkrebsstiftung (J.S., K.W., M.K.), DFG (J.S., M.K.), intramural Fortüne program of the Medical Faculty of the UKT (M.K., J.S.), GIF (K.W., B.D.), German Cancer Consortium (P.M., J.S., M.A.R.), Else Kröner-Fresenius Stiftung (M.K.), Fritz Thyssen Foundation (J.S.). We thank Hans-Georg Rammensee for the financial support of the RNA seq and methylation arrays; E. Trompouki (MPI of Immunobiology and Epigenetics, Freiburg) and P. Müller (Friedrich-Miescher Laboratory, Tübingen) for providing zebrafish wild-type TE strain and the Tg(*mpo:gfp*) fish; J. Wittbrodt (Centre for Organismal Studies, Heidelberg) for the support with CRISPR-Cas9 experiments in zebrafish. We thank K. Hähnel and R. Bernhard for the excellent technical support. We also thank A. Dick for support with zebrafish experiments; the microarray and high-throughput sequencing units of the DKFZ Genomics and Proteomics Core Facility for providing the Illumina Human Methylation EPIC array and mRNA sequencing services.

Author contributions

J.S. and K.W. made initial observations and supervised experimentation; J.S. and P.M. designed the experiments, analyzed the data and wrote the manuscript (with the assistance of M.K., B.B. and K.W.); P.M. and M.K. designed GADD45 β gRNA, P.M., M.K., S.K. and M.N. analyzed RNA seq and methylation array data; P.M. and M.N. performed CRISPR/Cas9-mediated GADD45 β knockout in iPSCs and primary HSPCs; B.D. performed differentiation of iPSCs; N.A., B.B., L.D. and P.M. performed zebrafish experiments; B.B. performed analysis and data summary of zebrafish experiments; T.T. made the *gadd45bb* CRISPR constructs; P.M. and B.F. performed CHIP and reporter gene assay; F.B.T., S.W. and M.A.R. generated Gadd45 β construct and performed *in vitro* differentiation of mouse HSCs; C.Z. and K.W. provided patients material; M.K. and M.U.R. assisted with the data analysis and provided insightful comments.

References

1. Skokowa, J., Dale, D.C., Touw, I.P., Zeidler, C. & Welte, K. Severe congenital neutropenias. *Nature reviews. Disease primers* **3**, 17032 (2017).
2. Skokowa, J., et al. LEF-1 is crucial for neutrophil granulocytopoiesis and its expression is severely reduced in congenital neutropenia. *Nat Med* **12**, 1191-1197 (2006).
3. Cario, G., et al. Heterogeneous expression pattern of pro- and anti-apoptotic factors in myeloid progenitor cells of patients with severe congenital neutropenia treated with granulocyte colony-stimulating factor. *Br J Haematol* **129**, 275-278 (2005).
4. Gupta, K., et al. Bortezomib inhibits STAT5-dependent degradation of LEF-1, inducing granulocytic differentiation in congenital neutropenia CD34(+) cells. *Blood* **123**, 2550-2561 (2014).
5. Klimentkova, O., et al. A lack of secretory leukocyte protease inhibitor (SLPI) causes defects in granulocytic differentiation. *Blood* **123**, 1239-1249 (2014).
6. Nustede, R., et al. ELANE mutant-specific activation of different UPR pathways in congenital neutropenia. *Br J Haematol* **172**, 219-227 (2016).
7. Skokowa, J. & Welte, K. Dysregulation of myeloid-specific transcription factors in congenital neutropenia. *Ann N Y Acad Sci* **1176**, 94-100 (2009).
8. Skokowa, J. & Welte, K. Defective G-CSFR signaling pathways in congenital neutropenia. *Hematol Oncol Clin North Am* **27**, 75-88, viii (2013).
9. Skokowa, J. & Welte, K. LEF-1 is a decisive transcription factor in neutrophil granulopoiesis. *Ann N Y Acad Sci* **1106**, 143-151 (2007).
10. Skokowa, J., et al. NAMPT is essential for the G-CSF-induced myeloid differentiation via a NAD(+)-sirtuin-1-dependent pathway. *Nat Med* **15**, 151-158 (2009).
11. Skokowa, J., et al. Interactions among HCLS1, HAX1 and LEF-1 proteins are essential for G-CSF-triggered granulopoiesis. *Nat Med* **18**, 1550-1559 (2012).
12. Dannenmann, B., et al. Human iPSC-based model of severe congenital neutropenia reveals elevated UPR and DNA damage in CD34(+) cells preceding leukemic transformation. *Exp Hematol* **71**, 51-60 (2019).
13. Gupta, K., et al. Bortezomib inhibits STAT5-dependent degradation of LEF-1, inducing granulocytic differentiation in congenital neutropenia CD34+ cells. *Blood* **123**, 2550-2561 (2014).
14. Kollner, I., et al. Mutations in neutrophil elastase causing congenital neutropenia lead to cytoplasmic protein accumulation and induction of the unfolded protein response. *Blood* **108**, 493-500 (2006).
15. Grenda, D.S., et al. Mutations of the ELA2 gene found in patients with severe congenital neutropenia induce the unfolded protein response and cellular apoptosis. *Blood* **110**, 4179-4187 (2007).
16. Nanua, S., et al. Activation of the unfolded protein response is associated with impaired granulopoiesis in transgenic mice expressing mutant Elane. *Blood* **117**, 3539-3547 (2011).
17. Nayak, R.C., et al. Pathogenesis of ELANE-mutant severe neutropenia revealed by induced pluripotent stem cells. *J Clin Invest* **125**, 3103-3116 (2015).
18. Xia, J. & Link, D.C. Severe congenital neutropenia and the unfolded protein response. *Curr Opin Hematol* **15**, 1-7 (2008).
19. Klein, C., et al. HAX1 deficiency causes autosomal recessive severe congenital neutropenia (Kostmann disease). *Nat Genet* **39**, 86-92 (2007).
20. Wang, J., et al. A differentiation checkpoint limits hematopoietic stem cell self-renewal in response to DNA damage. *Cell* **148**, 1001-1014 (2012).
21. van Galen, P., et al. The unfolded protein response governs integrity of the haematopoietic stem-cell pool during stress. *Nature* **510**, 268-272 (2014).
22. Matatall, K.A., et al. Chronic Infection Depletes Hematopoietic Stem Cells through Stress-Induced Terminal Differentiation. *Cell Rep* **17**, 2584-2595 (2016).
23. Wang, X.W., et al. GADD45 induction of a G2/M cell cycle checkpoint. *Proc Natl Acad Sci U S A* **96**, 3706-3711 (1999).

24. Vairapandi, M., Balliet, A.G., Hoffman, B. & Liebermann, D.A. GADD45b and GADD45g are cdc2/cyclinB1 kinase inhibitors with a role in S and G2/M cell cycle checkpoints induced by genotoxic stress. *J Cell Physiol* **192**, 327-338 (2002).
25. Carrier, F., *et al.* Gadd45, a p53-responsive stress protein, modifies DNA accessibility on damaged chromatin. *Mol Cell Biol* **19**, 1673-1685 (1999).
26. Hollander, M.C., *et al.* Genomic instability in Gadd45a-deficient mice. *Nat Genet* **23**, 176-184 (1999).
27. Gupta, M., *et al.* Hematopoietic cells from Gadd45a- and Gadd45b-deficient mice are sensitized to genotoxic-stress-induced apoptosis. *Oncogene* **24**, 7170-7179 (2005).
28. Liebermann, D.A. & Hoffman, B. Gadd45 in stress signaling. *J Mol Signal* **3**, 15 (2008).
29. Magimaidas, A., *et al.* Gadd45b deficiency promotes premature senescence and skin aging. *Oncotarget* **7**, 26935-26948 (2016).
30. Selvakumaran, M., *et al.* The novel primary response gene MyD118 and the proto-oncogenes myb, myc, and bcl-2 modulate transforming growth factor beta 1-induced apoptosis of myeloid leukemia cells. *Mol Cell Biol* **14**, 2352-2360 (1994).
31. Zhan, Q., *et al.* The gadd and MyD genes define a novel set of mammalian genes encoding acidic proteins that synergistically suppress cell growth. *Mol Cell Biol* **14**, 2361-2371 (1994).
32. Wingert, S., *et al.* DNA-damage response gene GADD45A induces differentiation in hematopoietic stem cells without inhibiting cell cycle or survival. *Stem Cells* **34**, 699-710 (2016).
33. Thalheimer, F.B., *et al.* Cytokine-regulated GADD45G induces differentiation and lineage selection in hematopoietic stem cells. *Stem Cell Reports* **3**, 34-43 (2014).
34. Gupta, M., Gupta, S.K., Hoffman, B. & Liebermann, D.A. Gadd45a and Gadd45b protect hematopoietic cells from UV-induced apoptosis via distinct signaling pathways, including p38 activation and JNK inhibition. *J Biol Chem* **281**, 17552-17558 (2006).
35. De Smaele, E., *et al.* Induction of gadd45beta by NF-kappaB downregulates pro-apoptotic JNK signalling. *Nature* **414**, 308-313 (2001).
36. Nasri, M., *et al.* Fluorescent labeling of CRISPR/Cas9 RNP for gene knockout in HSPCs and iPSCs reveals an essential role for GADD45b in stress response. *Blood advances* **3**, 63-71 (2019).
37. Ma, D.K., *et al.* Neuronal activity-induced Gadd45b promotes epigenetic DNA demethylation and adult neurogenesis. *Science* **323**, 1074-1077 (2009).
38. Gavin, D.P., *et al.* Growth arrest and DNA-damage-inducible, beta (GADD45b)-mediated DNA demethylation in major psychosis. *Neuropsychopharmacology* **37**, 531-542 (2012).
39. Rai, K., *et al.* DNA demethylation in zebrafish involves the coupling of a deaminase, a glycosylase, and gadd45. *Cell* **135**, 1201-1212 (2008).
40. Cortellino, S., *et al.* Thymine DNA glycosylase is essential for active DNA demethylation by linked deamination-base excision repair. *Cell* **146**, 67-79 (2011).
41. Guidotti, A., *et al.* Epigenetic GABAergic targets in schizophrenia and bipolar disorder. *Neuropharmacology* **60**, 1007-1016 (2011).
42. Brinkman, E.K., Chen, T., Amendola, M. & van Steensel, B. Easy quantitative assessment of genome editing by sequence trace decomposition. *Nucleic Acids Res* **42**, e168 (2014).
43. Gupta, S.K., Gupta, M., Hoffman, B. & Liebermann, D.A. Hematopoietic cells from gadd45a-deficient and gadd45b-deficient mice exhibit impaired stress responses to acute stimulation with cytokines, myeloablation and inflammation. *Oncogene* **25**, 5537-5546 (2006).
44. Bagger, F.O., Kinalis, S. & Rapin, N. BloodSpot: a database of healthy and malignant haematopoiesis updated with purified and single cell mRNA sequencing profiles. *Nucleic Acids Res* **47**, D881-D885 (2019).

45. Robertson, A.L., Avagyan, S., Gansner, J.M. & Zon, L.I. Understanding the regulation of vertebrate hematopoiesis and blood disorders - big lessons from a small fish. *FEBS Lett* **590**, 4016-4033 (2016).
46. de Pater, E. & Trompouki, E. Bloody Zebrafish: Novel Methods in Normal and Malignant Hematopoiesis. *Frontiers in cell and developmental biology* **6**, 124 (2018).
47. Kawahara, A., Che, Y.S., Hanaoka, R., Takeda, H. & Dawid, I.B. Zebrafish GADD45beta genes are involved in somite segmentation. *Proc Natl Acad Sci U S A* **102**, 361-366 (2005).
48. Renshaw, S.A., *et al.* A transgenic zebrafish model of neutrophilic inflammation. *Blood* **108**, 3976-3978 (2006).
49. Bajoghli, B., Aghaallaei, N., Heimbucher, T. & Czerny, T. An artificial promoter construct for heat-inducible misexpression during fish embryogenesis. *Dev Biol* **271**, 416-430 (2004).
50. Stachura, D.L., *et al.* The zebrafish granulocyte colony-stimulating factors (Gcsfs): 2 paralogous cytokines and their roles in hematopoietic development and maintenance. *Blood* **122**, 3918-3928 (2013).
51. Janky, R., *et al.* iRegulon: from a gene list to a gene regulatory network using large motif and track collections. *PLoS Comput Biol* **10**, e1003731 (2014).
52. Verfaillie, A., Imrichova, H., Janky, R. & Aerts, S. iRegulon and i-cisTarget: Reconstructing Regulatory Networks Using Motif and Track Enrichment. *Curr Protoc Bioinformatics* **52**, 2 16 11-39 (2015).
53. Labonte, B., *et al.* Gadd45b mediates depressive-like role through DNA demethylation. *Sci Rep* **9**, 4615 (2019).
54. Peters, T.J., *et al.* De novo identification of differentially methylated regions in the human genome. *Epigenetics Chromatin* **8**, 6 (2015).
55. Aryee, M.J., *et al.* Minfi: a flexible and comprehensive Bioconductor package for the analysis of Infinium DNA methylation microarrays. *Bioinformatics* **30**, 1363-1369 (2014).
56. Fortin, J.P., Triche, T.J., Jr. & Hansen, K.D. Preprocessing, normalization and integration of the Illumina HumanMethylationEPIC array with minfi. *Bioinformatics* **33**, 558-560 (2017).
57. Ritchie, M.E., *et al.* limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res* **43**, e47 (2015).
58. Sheffield, N.C. & Bock, C. LOLA: enrichment analysis for genomic region sets and regulatory elements in R and Bioconductor. *Bioinformatics* **32**, 587-589 (2016).
59. Salerno, D.M., Tront, J.S., Hoffman, B. & Liebermann, D.A. Gadd45a and Gadd45b modulate innate immune functions of granulocytes and macrophages by differential regulation of p38 and JNK signaling. *Journal of cellular physiology* **227**, 3613-3620 (2012).
60. Koch, C., *et al.* GM-CSF treatment is not effective in congenital neutropenia patients due to its inability to activate NAMPT signaling. *Ann Hematol* **96**, 345-353 (2017).
61. Moskalev, A.A., *et al.* Gadd45 proteins: relevance to aging, longevity and age-related pathologies. *Ageing Res Rev* **11**, 51-66 (2012).
62. Papa, S., *et al.* Gadd45 beta mediates the NF-kappa B suppression of JNK signalling by targeting MKK7/JNK2. *Nature cell biology* **6**, 146-153 (2004).
63. Niu, H., *et al.* Endogenous retinoid X receptor ligands in mouse hematopoietic cells. *Science signaling* **10**(2017).
64. Johnson, B.S., *et al.* Retinoid X receptor (RXR) agonist-induced activation of dominant-negative RXR-retinoic acid receptor alpha403 heterodimers is developmentally regulated during myeloid differentiation. *Mol Cell Biol* **19**, 3372-3382 (1999).
65. Lawson, N.D. & Berliner, N. Neutrophil maturation and the role of retinoic acid. *Exp Hematol* **27**, 1355-1367 (1999).
66. Gavin, D.P., *et al.* Gadd45b and N-methyl-D-aspartate induced DNA demethylation in postmitotic neurons. *Epigenomics* **7**, 567-579 (2015).

67. Grassi, D., *et al.* Neuronal Activity, TGFbeta-Signaling and Unpredictable Chronic Stress Modulate Transcription of Gadd45 Family Members and DNA Methylation in the Hippocampus. *Cereb Cortex* **27**, 4166-4181 (2017).
68. Guo, J.U., Su, Y., Zhong, C., Ming, G.L. & Song, H. Hydroxylation of 5-methylcytosine by TET1 promotes active DNA demethylation in the adult brain. *Cell* **145**, 423-434 (2011).
69. Morgan, H.D., Dean, W., Coker, H.A., Reik, W. & Petersen-Mahrt, S.K. Activation-induced cytidine deaminase deaminates 5-methylcytosine in DNA and is expressed in pluripotent tissues: implications for epigenetic reprogramming. *J Biol Chem* **279**, 52353-52360 (2004).
70. Spruijt, C.G., *et al.* Dynamic readers for 5-(hydroxy)methylcytosine and its oxidized derivatives. *Cell* **152**, 1146-1159 (2013).
71. Yi, Y.W., *et al.* Gadd45 family proteins are coactivators of nuclear hormone receptors. *Biochemical and biophysical research communications* **272**, 193-198 (2000).
72. Karmodiya, K., Krebs, A.R., Oulad-Abdelghani, M., Kimura, H. & Tora, L. H3K9 and H3K14 acetylation co-occur at many gene regulatory elements, while H3K14ac marks a subset of inactive inducible promoters in mouse embryonic stem cells. *BMC Genomics* **13**, 424 (2012).
73. Martens, J.H., *et al.* PML-RARalpha/RXR Alters the Epigenetic Landscape in Acute Promyelocytic Leukemia. *Cancer Cell* **17**, 173-185 (2010).
74. Huang, M.E., *et al.* Use of all-trans retinoic acid in the treatment of acute promyelocytic leukemia. *Blood* **72**, 567-572 (1988).
75. Baldus, M., Walter, H., Moller, M., Schurfeld, C. & Brass, H. All-Trans-Retinoic Acid (Atra) in the Treatment of Myelodysplastic-Syndromes - Results in 5 Cases. *Onkologie* **17**, 515-520 (1994).
76. Tront, J.S., Hoffman, B. & Liebermann, D.A. Gadd45a suppresses Ras-driven mammary tumorigenesis by activation of c-Jun NH2-terminal kinase and p38 stress signaling resulting in apoptosis and senescence. *Cancer Res* **66**, 8448-8454 (2006).
77. Tront, J.S., Huang, Y., Fornace, A.J., Jr., Hoffman, B. & Liebermann, D.A. Gadd45a functions as a promoter or suppressor of breast cancer dependent on the oncogenic stress. *Cancer Res* **70**, 9671-9681 (2010).
78. Tront, J.S., Willis, A., Huang, Y., Hoffman, B. & Liebermann, D.A. Gadd45a levels in human breast cancer are hormone receptor dependent. *J Transl Med* **11**, 131 (2013).
79. Wang, W., *et al.* Analysis of methylation-sensitive transcriptome identifies GADD45a as a frequently methylated gene in breast cancer. *Oncogene* **24**, 2705-2714 (2005).
80. Yamasawa, K., Nio, Y., Dong, M., Yamaguchi, K. & Itakura, M. Clinicopathological significance of abnormalities in Gadd45 expression and its relationship to p53 in human pancreatic cancer. *Clin Cancer Res* **8**, 2563-2569 (2002).
81. Ying, J., *et al.* The stress-responsive gene GADD45G is a functional tumor suppressor, with its response to environmental stresses frequently disrupted epigenetically in multiple tumors. *Clin Cancer Res* **11**, 6442-6449 (2005).
82. Zerbini, L.F., *et al.* JunD-mediated repression of GADD45alpha and gamma regulates escape from cell death in prostate cancer. *Cell Cycle* **10**, 2583-2591 (2011).
83. Ramachandran, K., *et al.* Methylation-mediated repression of GADD45alpha in prostate cancer and its role as a potential therapeutic target. *Cancer Res* **69**, 1527-1535 (2009).
84. Michaelis, K.A., *et al.* Identification of growth arrest and DNA-damage-inducible gene beta (GADD45beta) as a novel tumor suppressor in pituitary gonadotrope tumors. *Endocrinology* **152**, 3603-3613 (2011).
85. Na, Y.K., *et al.* Hypermethylation of growth arrest DNA-damage-inducible gene 45 in non-small cell lung cancer and its relationship with clinicopathologic features. *Mol Cells* **30**, 89-92 (2010).
86. Li, Y., *et al.* Adenoviral-mediated gene transfer of Gadd45a results in suppression by inducing apoptosis and cell cycle arrest in pancreatic cancer cell. *J Gene Med* **11**, 3-13 (2009).

87. Zhao, Z., *et al.* GADD45B as a Prognostic and Predictive Biomarker in Stage II Colorectal Cancer. *Genes* **9**(2018).
88. Sha, X., Hoffman, B. & Liebermann, D.A. Loss of Gadd45b accelerates BCR-ABL-driven CML. *Oncotarget* **9**, 33360-33367 (2018).

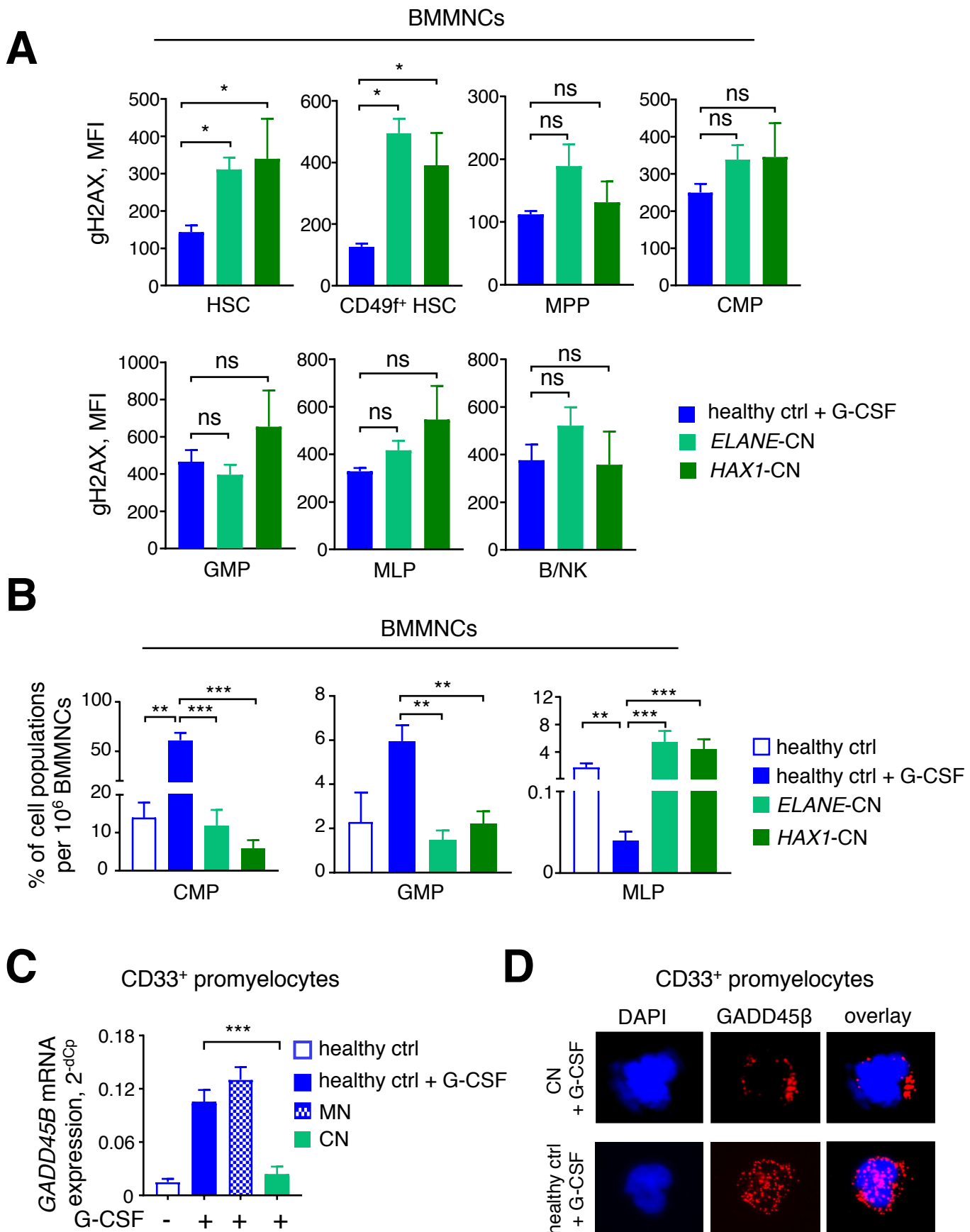
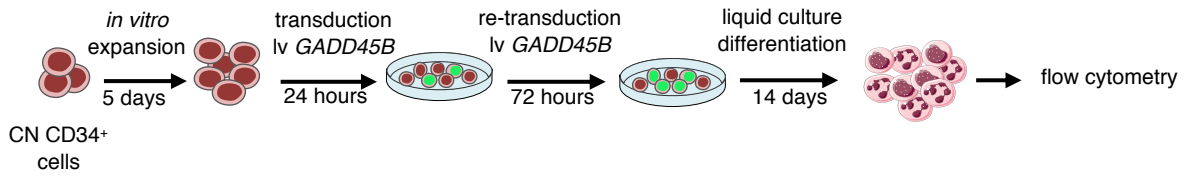
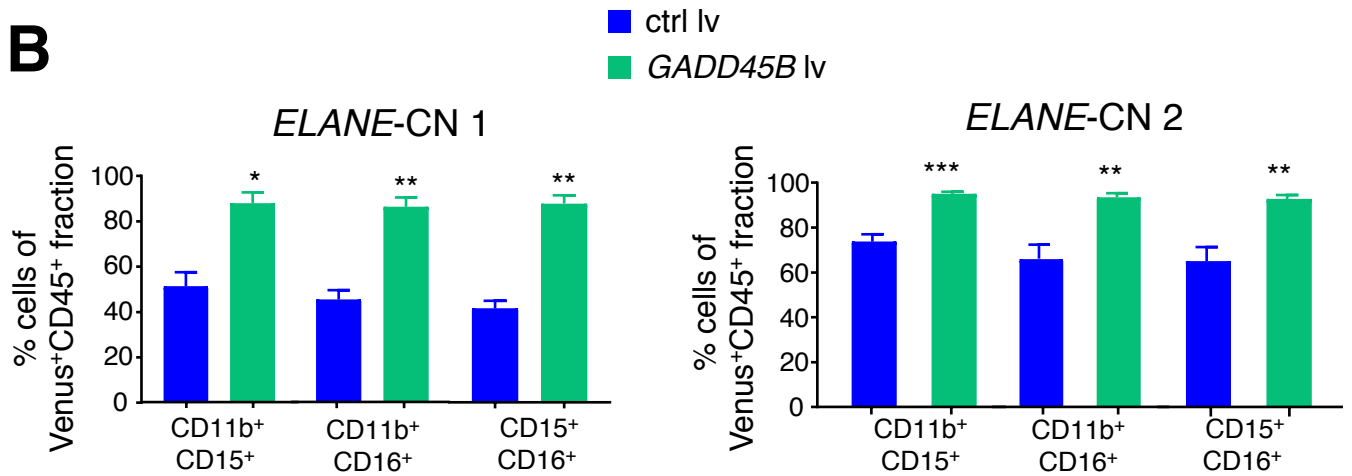
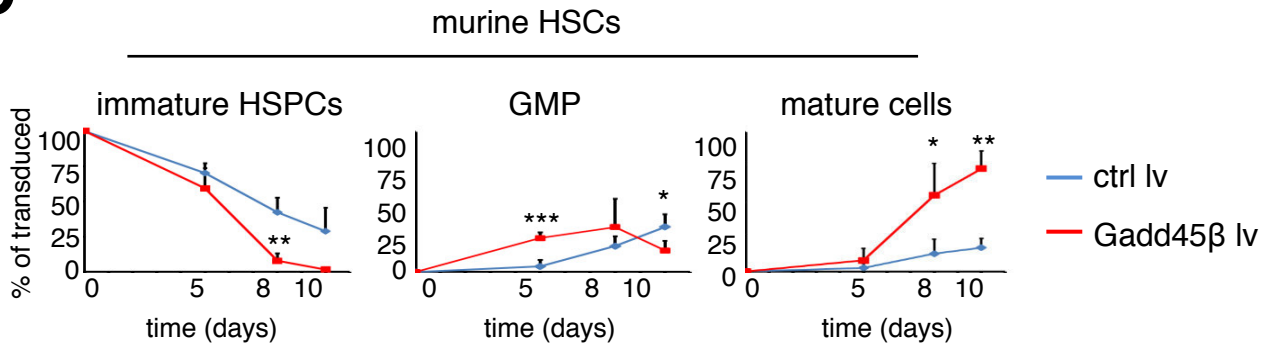
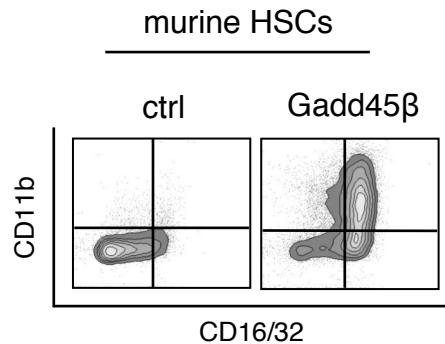
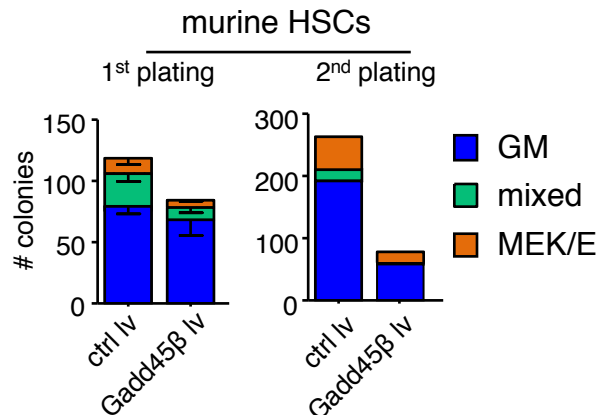
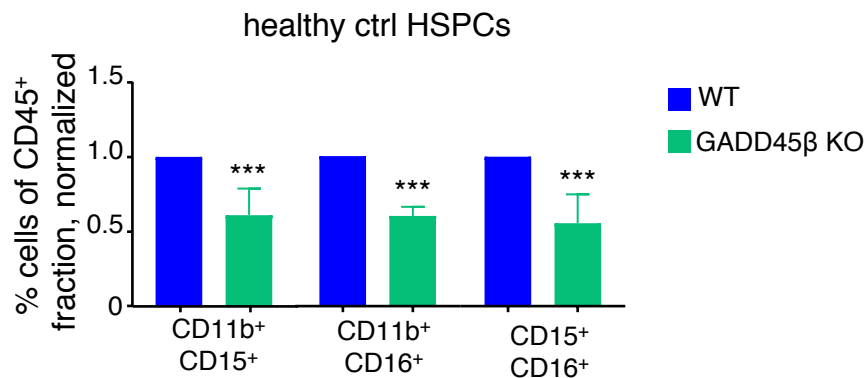
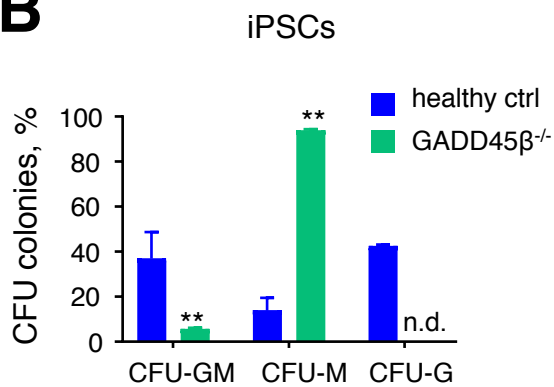
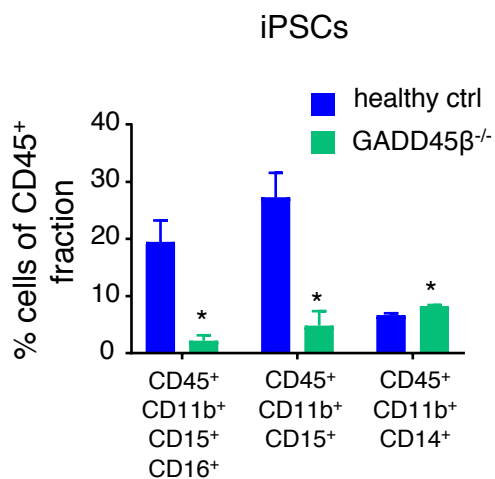
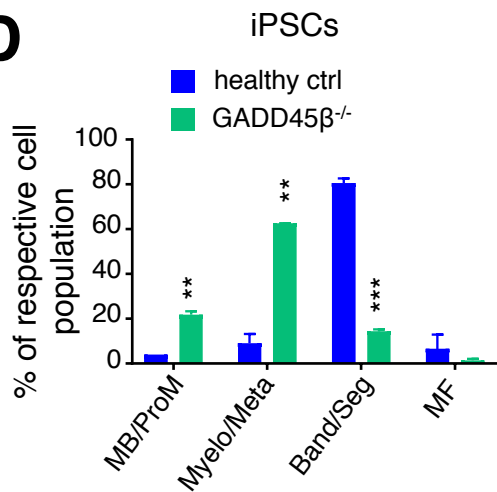
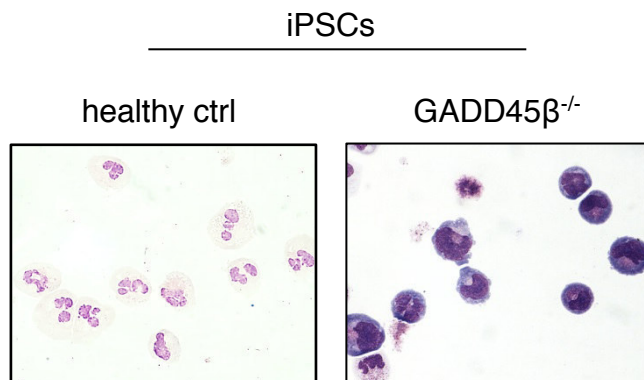
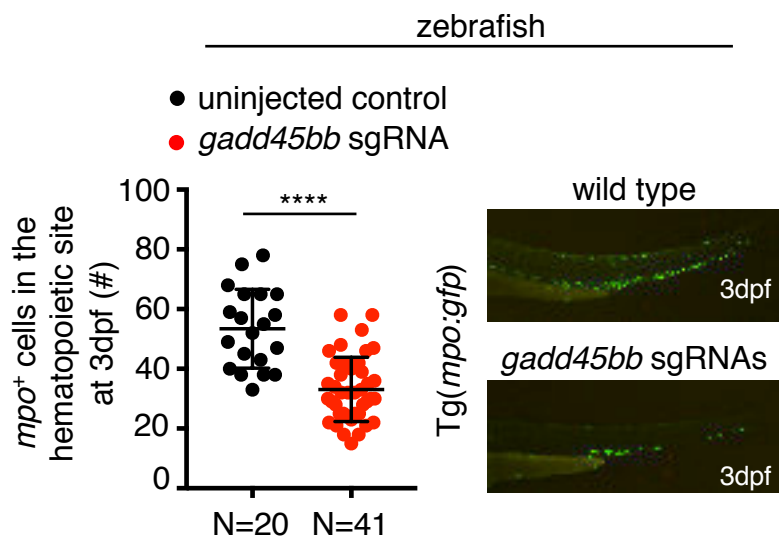
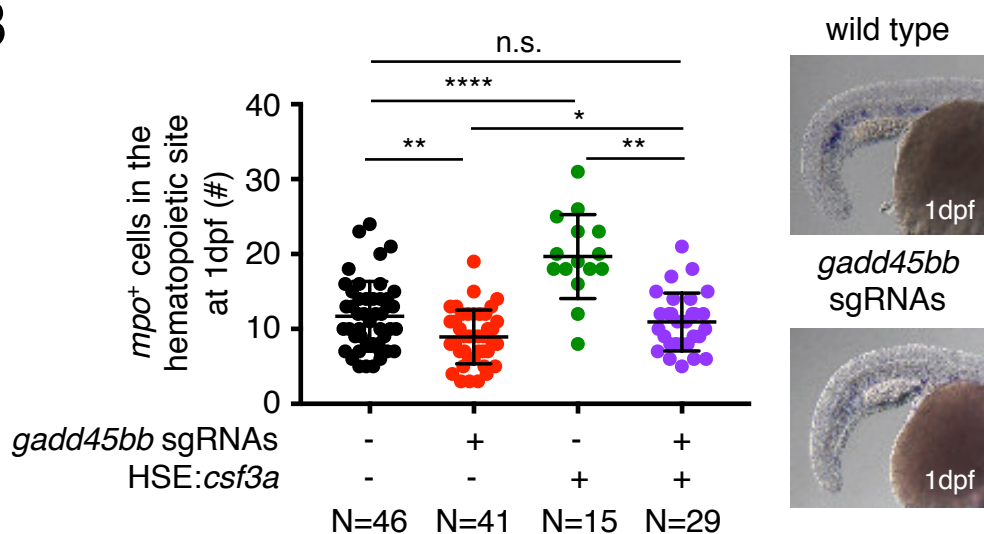
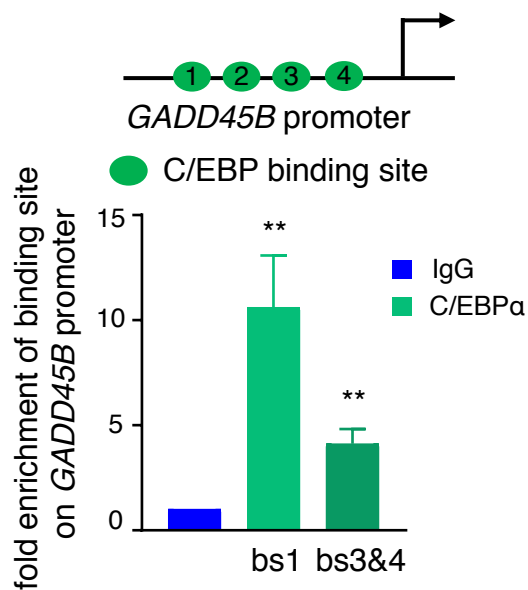
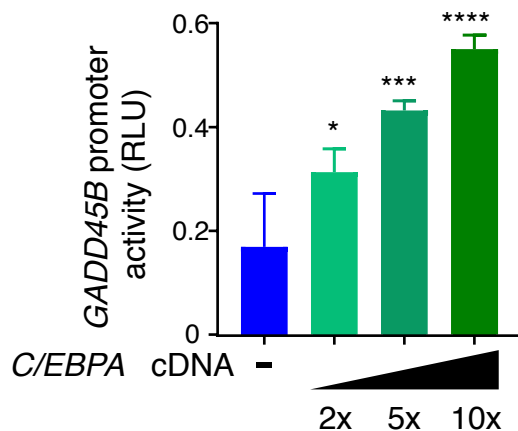
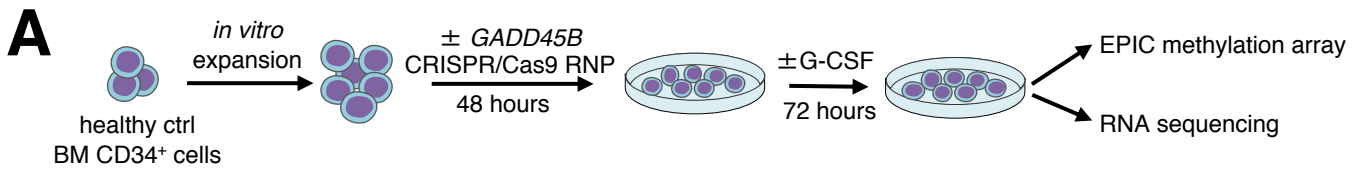


Figure 1

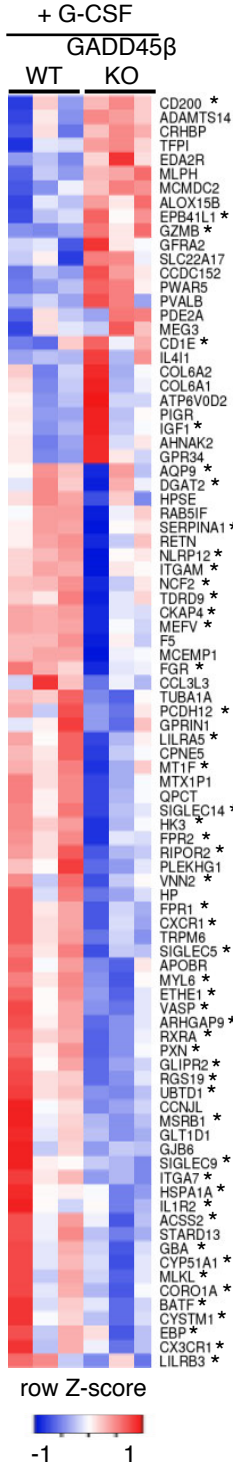
A**B****C****D****E****Figure 2**

A**B****C****D****E****Figure 3**

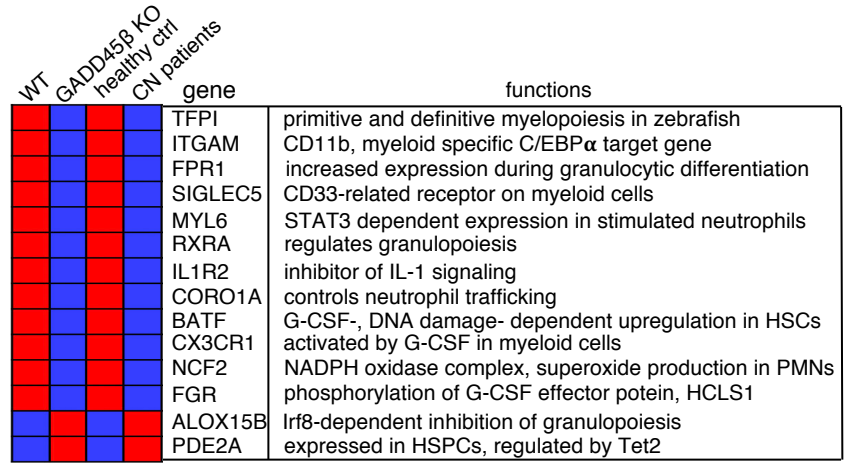
A**B****C****D****Figure 4**



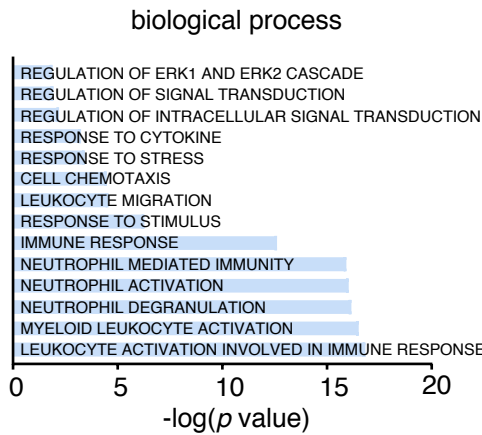
B



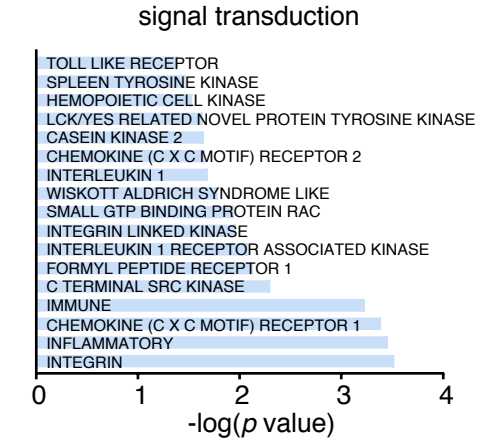
C



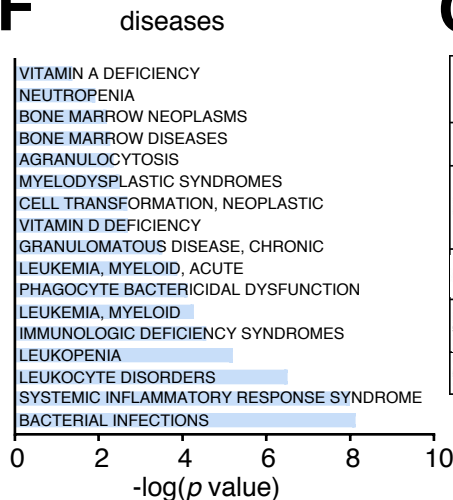
D



E



F



G

Enriched Motif	NES	AUC	Transcription factors	Targets
	5.194	0.089	RARA, RARB, ETS1	9
	3.682	0.068	RXRA, RARG, RXRB, RARA, RARB, FOS, PPARG *	7
	3.570	0.066	SPI1	21
	3.289	0.062	CEBPG	13
	3.100	0.059	CEBPB	22

Figure 5

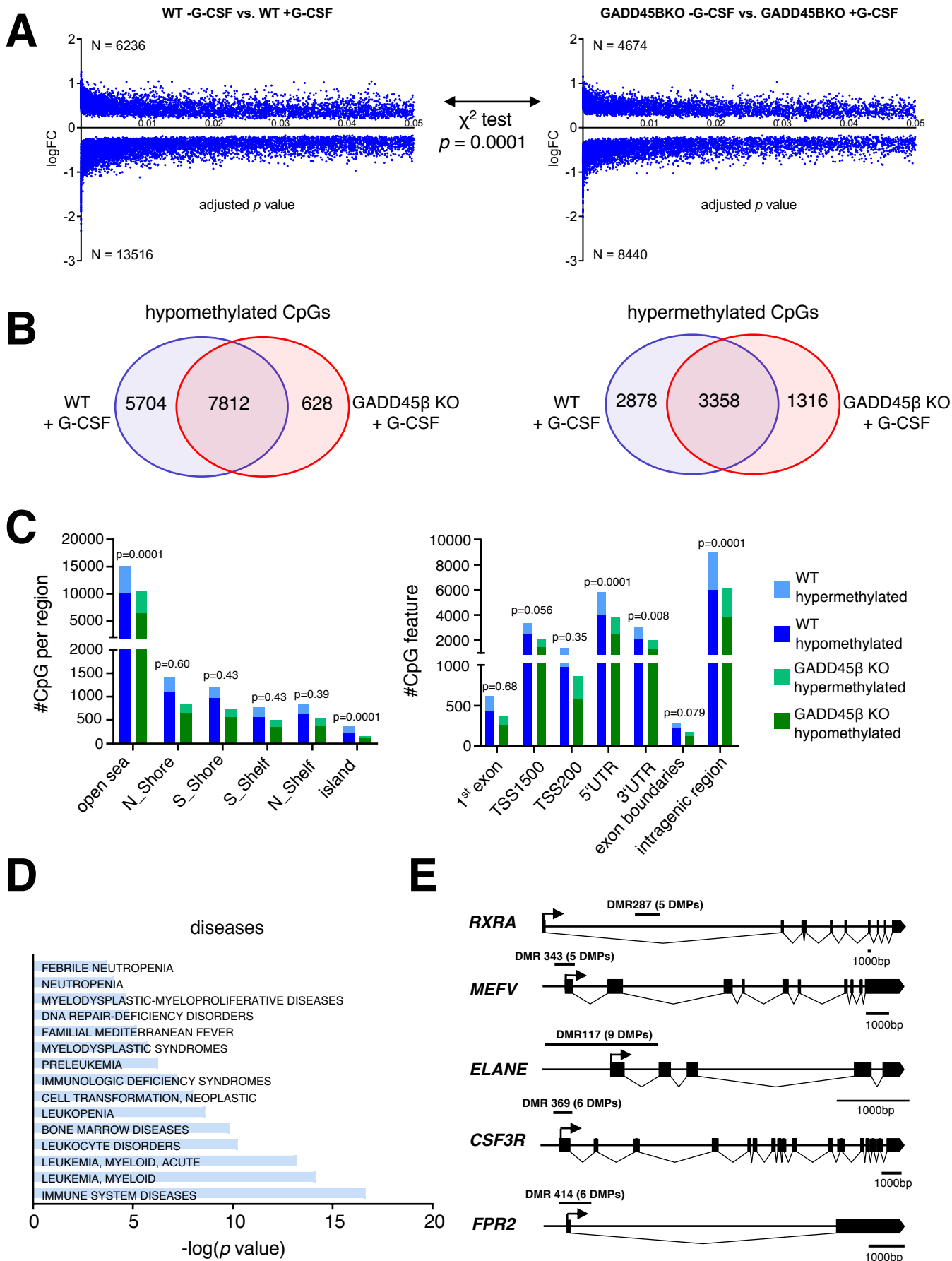


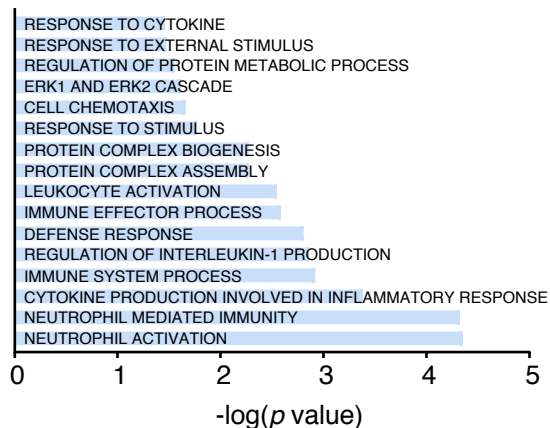
Figure 6

A

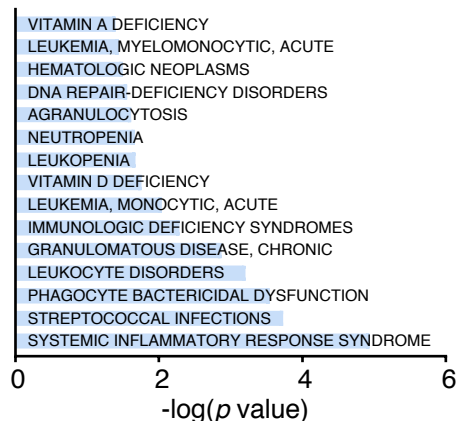
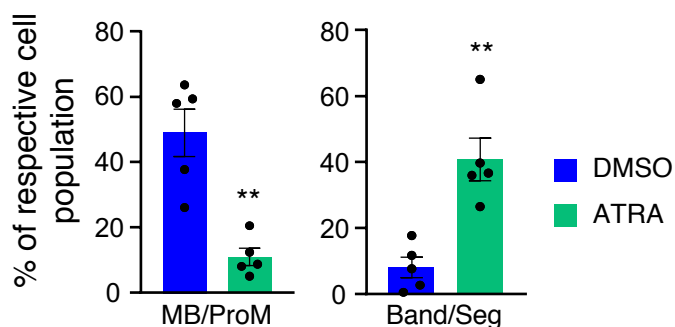
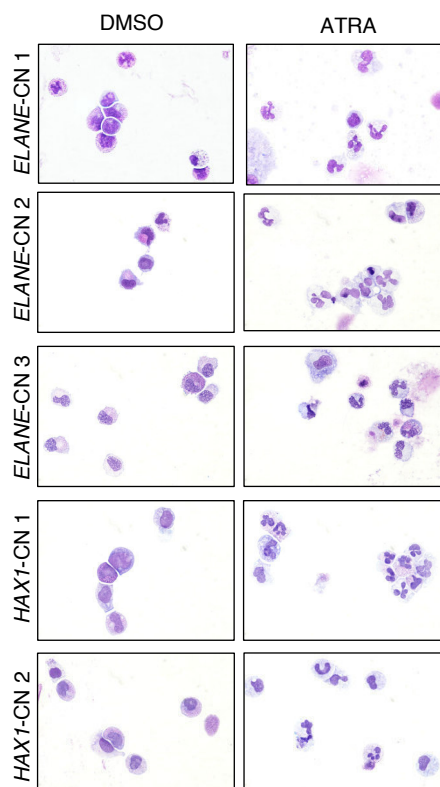
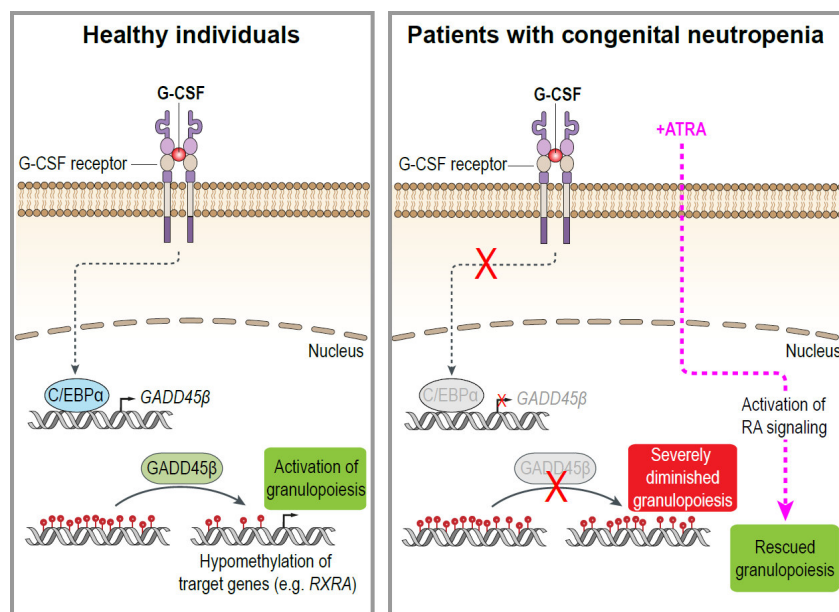
gene	CpG localization
GJB6	3' UTR
CCNJL	3' UTR
TRPM6	TSS1500
CXCR1	5' UTR
PCDH12	1st Exon
IL1R2	1st Exon
MEFV	TSS1500/1st Exon/3' UTR
FPR2	5' UTR/1st Exon
SIGLEC9	TSS1500
MSRB1	3' UTR
SERPINA1	TSS200/TSS1500
GLIPR2	TSS200
RXRA	TSS200/5' UTR
ARHGAP9	TSS1500
EPB41L1	5' UTR
IL4I1	5' UTR

B

biological process

**C**

diseases

**D****E****F****Figure 7**

Supplemental information

Supplemental material and methods

Cell culture

The human embryonic kidney 293T (HEK293T) cell line was cultured under standard conditions (37 °C, 5 % CO₂, 20 % O₂) using DMEM high glucose for HEK293T cells (Gibco, #41966-052).

Human CD34⁺ hematopoietic stem and progenitor cells (HSPCs) and CD33⁺ myeloid progenitor cells were isolated from bone marrow- or peripheral blood mononuclear cell fraction by magnetic bead separation using Human CD34 Progenitor Cell Isolation kit, (#130-046-703) and Human CD33 Myeloid Progenitor Cell Isolation kit (#130-045-501) (Miltenyi Biotech, Germany). CD34⁺ cells were cultured in a density of 2 x 10⁵ cells/ml in Stemline II Hematopoietic Stem Cell Expansion medium (Sigma Aldrich, #50192) supplemented with 10 % FBS, 1 % penicillin/streptomycin, 1 % L-glutamine and a cytokine cocktail consisting of 20 ng/ml IL-3, 20 ng/ml IL-6, 20 ng/ml TPO, 50 ng/ml SCF and 50 ng/ml FLT-3L.

Human induced pluripotent stem cells (iPSCs) were cultured on Geltrex LDEV-free reduced growth factor basement membrane matrix (Thermo Fisher Scientific, #A1413201) coated plates in a density of 2 x 10⁵ cells/ml in StemFlex medium (Thermo Fisher Scientific, #A3349401) supplemented with 1 % penicillin/streptomycin.

FACS analysis of HSPCs composition

10⁵ CD34⁺ HSPCs or 10⁶ bone marrow mononuclear cells were washed with PBS and stained with corresponding surface marker antibodies in PBS containing 2% FBS and 0.02% sodium azide. For HSPC surface marker analysis (panel adapted from van Galen *et al.*¹), following antibodies were used: mouse anti-human CD38 (BD, #563964), mouse anti-human CD34 (BD, #348811), rat anti-human CD49f (BD, #563271), mouse anti-human CD90 (BD, #562685), mouse anti-human CD45RA (BD, #560673), mouse anti-human CD10 (BD, #563734), mouse anti-human CD135 (BD, #564708). For the analysis of nuclear H2AX

protein, cells were subsequently permeabilized and fixed using IntraSure kit (BD #641776) followed by incubation with mouse anti-human gamma H2AX (BD, #560447 or #560445) antibody.

Pure knockout clone isolation and embryoid body (EB)-based iPSC differentiation

Human healthy donor iPSCs were nucleofected with CRISPR/Cas9-gRNA RNP complexes against *GADD45B* (150 pmol per 10^6 cells, program CA-137) and afterwards plated on a Geltrex-coated 10-cm dish (15×10^3 cells/dish) in StemFlex medium (ThermoFisher, #A3349401) with RevitaCell supplement (Thermo Fisher Scientific, #A2644501). Media was changed every 24 hours without RevitaCell supplement. After 9 to 12 days, iPSC colonies were picked and transferred on a Geltrex-coated 96 well plate. Knockout was assessed for each clone by Cas9 *in vitro* digestion and confirmed by Sanger sequencing. The knockout clone was expanded and cultured on feeder cells. The expanded iPSC clones were dissociated from SNL-feeders or Geltrex LDEV-free reduced growth factor basement membrane matrix (Thermo Fisher Scientific, #A1413201) coated plates using PBS/EDTA (0.02%) for 5 min. As described in Dannenmann et al.² EB induction was done via spin EBs (2×10^4 cells/EB) in 96-well plates using APEL serum-free differentiation medium (Stemcell Technologies, #5270) supplemented with bFGF (20 ng/ml) and ROCK inhibitor (Y-27632 dihydrochloride, R&D Systems, #1254). On day 1, BMP4 (40 ng/ml) was added to induce mesodermal differentiation. On day 4, EBs were plated on Geltrex coated 6-well-plates (10 EBs/well) in APEL medium supplemented with VEGF (40 ng/ml), SCF (50 ng/ml) and IL-3 (50 ng/ml). For neutrophilic differentiation, the medium was changed 3 days later to APEL supplemented with IL-3 (50 ng/ml) and G-CSF (50 ng/ml). First hematopoietic floating cells appeared on day 12 – 14. Floating cells were harvested starting from day 14 to day 32 every 3 - 4 days and used for FACS analysis and Wright-Giemsa staining for morphologic discrimination.

Colony forming unit (CFU) assay

As described previously³, CD34⁺ cells were resuspended in Iscove's MDM with 2 % FBS (Stemcell Technologies, #07700) and enriched Methocult (Stemcell Technologies, #H4435). Cell suspension was plated on 3.5 cm dishes (1 x 10³ cells/dish) followed by colony counting at day 14.

Murine *Gadd45b* expression vector

The open reading frame (ORF) of a fluorescent reporter protein (either VENUS-hImportin subunit α 1 (AA2-67) or tdTOMATO-hImportin subunit α 1 (AA2-67)) was cloned into the third generation self-inactivating lentiviral vector pRRL.PPT.SFFV.IRES.eGFP.wPRE by replacing the ORF of green fluorescent protein⁴. A multiple cloning site (MCS) was inserted after the SFFV promoter. The ORF of murine *Gadd45b* was amplified by RT-PCR (forward 5'-TTGGCCGGCCTGCATCATGACCCTGGAAGAG-3', reverse 5'-TTACTAGTCACGGGTAGGGTAGCCTTTGA-3') and then cloned into the MCS.

Lentivirus production and transduction

Human GADD45 β cDNA expressing and empty control lentiviruses were produced in HEK293T cells⁵. For this, HEK293T were transfected with lentiviral envelope plasmid pMD2.G, the lentiviral packaging plasmid psPAX2 and donor plasmid using *TransIT-LT1* transfection reagent (Mirus, #MIR2300) according to the manufacturer's protocol. Two days after transfection, the supernatant containing the virus was harvested and concentrated using Lenti-X Concentrator (Takara Clontech, #631232). Viral titers were determined by transduction of HEK293T cells with different concentrations of virus supernatant and FACS.

Vesicular Stomatitis Virus-G (VSVG)-pseudotyped lentiviral particles containing murine *Gadd45 β* cDNA were produced in a split genome approach by calcium-phosphate-mediated transient transfection of human embryonic kidney 293T producer cells⁵. After 48 hours, supernatant was collected, filtered (45 μ m), and enriched by ultracentrifugation

(50.000 g, 2 hours). Viral titers were determined by transduction of NIH3T3 cells with different concentrations of virus supernatant and FACS.

CD34⁺ cells were transduced with a multiplicity of infection (MOI) of 6 to 10 using 5 µg/ml polybrene by spinoculation at 500 x g at 4 °C for 10 minutes. Cells were re-transduced after 24 hours. One day after re-transduction medium was exchanged and cells were incubated for two more days before differentiation experiments were started.

GADD45β shRNA and control shRNA were produced in HEK293T cells and concentrated with Lenti-X Concentrator as described above. NB4 cells were transduced with MOI 5 using 5µg/ml polybrene. Medium was exchanged after 24 hours. NB4 cells were differentiated using 1 µM ATRA for 14 days.

***In vitro* differentiation of murine LT-HSCs**

Mouse BM cells from femurs, tibiae, coxae and sternum of 12-16 week old C57BL/6 mice were isolated by crushing the bones and LT-HSCs were sorted with a FACS Aria (BD) after staining with antibodies against CD117, Sca1, CD150, CD48, CD34, CD16/32, and Streptavidin. Details are explained elsewhere²⁸. 100 murine LT-HSCs were lentivirally transduced (MOI 100) and cultured in SFEM (Stemcell Technologies) supplemented with 100 ng/ml SCF and TPO (Peprotech). Cells were analyzed by FACS (antibodies against CD48, CD117, CD16/32, CD11b) at indicated time points²⁸. To read out colony-formation, LT-HSCs were lentivirally transduced (MOI 100), seeded 24 hours later in M3434 medium (Stemcell Technologies) and scored microscopically after 9 to 12 days for transduction and colony formation. After counting the cells, defined cell numbers after primary colony formation were replated to fresh M3434 medium and scored for secondary colony formation after 9 to 12 days.

Genomic DNA isolation and *GADD45B* Sanger sequencing

Genomic DNA was isolated using the QIAamp DNA Mini Kit (Qiagen, #51306) according to the manufacturer's instructions. PCR to amplify CRISPR/Cas9 sgRNA RNP-

edited region of *GADD45B* was conducted using GoTaq Hot Start Polymerase Kit (Promega, #M5006). PCR products were purified using the QIAquick PCR Purification Kit (Qiagen, #28106), and Sanger sequencing was performed at Eurofins Genomics.

qRT-PCR

RNA was isolated using the RNeasy Mini Kit (Qiagen, #74106) according to the manufacturer's instructions. Reverse transcription was conducted with the Omniscript RT kit (Qiagen, #205113) according to the protocol using a mix of random hexamers (Thermo Fisher Scientific, #S0142) and oligo-dT (Thermo Fisher Scientific, #S0132) primers. The mRNA expression of mRNA was measured with the SYBR Green quantitative PCR kit (Roche, #04887352001) and primers against *GADD45B* mRNA (forward 5'-TGCTGTGACAACGACATCAAC-3' and reverse 5'-GTGAGGGTTCGTGACCAGG-3'). *GADD45B* mRNA expression was normalized to *GAPDH* and is presented as $2^{-\Delta C_p}$ values.

Duolink *in situ* proximity ligation assay

Proximity ligation assay was described in detail elsewhere^{3,6}. Cytospin slides (1×10^4 cells per slide) were fixed in ice-cold methanol for 5 min, washed with PBS twice, permeabilized using 0.5% Triton X-100 for 10 min, washed with PBS twice and stained using primary anti-GADD45 β antibody followed by Duolink[®] *in situ* PLA probes and Duolink[®] *in situ* Detection Reagents Orange (Olink Bioscience) following the manufacturer's protocol (Olink Bioscience, Uppsala, Sweden). After staining, samples were air-dried, mounted with Duolink[®] *In Situ* Mounting Medium with DAPI (Olink Bioscience) and examined using an Axio Observer.Z1 fluorescence microscope (ZEISS, Oberkochen, Germany).

Dual luciferase reporter gene assay

GADD45B promoter (1.2 kb upstream of ATG) predicted to contain four C/EBP binding sites was cloned into the pGL4.10 [*luc2*] firefly luciferase reporter vector (Promega, #E6651). HEK293T cells were transfected with the *GADD45B* firefly luciferase reporter vector, *Renilla* vector and *C/EBPA* plasmid in different concentrations using *TransIT-LT1*

transfection reagent (Mirus, #MIR2300). Lysates were harvested 60 hours post transfection and activity of both firefly and *Renilla* luciferase was measured using a GloMax Multi Detection System (Promega).

Chromatin immunoprecipitation assay

THP1 cells (15×10^6) were fixed using 1 % formaldehyde for 5 min at room temperature, fixation reaction was stopped by adding 0.125 M glycine. Cells were rinsed twice with ice-cold PBS and resuspended in 1 mL Farnham lab (FL) buffer (5 mM PIPES pH 8, 85 mM KCl, 0.5% Igepal CA-630) containing protease inhibitors. Covaris M220 focused ultrasonicator was used to isolate nuclei (peak power 75 W, duty factor 15 % and 200 cycles/burst at 4 °C). Nuclei were centrifuged at 1000 xg for 5 min at 4 °C and washed once in FL buffer. To shear the chromatin nuclei were taken up in resuspension buffer D3 (1 mM EDTA, 10 mM Tris-HCl pH 7.6, 0.1 % SDS) containing protease inhibitors and sonicated for 25 min, at peak power 75 W, duty factor 15 % and 200 cycles/burst. Sheared chromatin was pre-cleaned with mixed Dynabeads (protein A and G, Thermo Fisher Scientific, #10001D and #10003D) and precipitated using either C/EBP α monoclonal antibody (sc-166258) or input control antibody (sc-2025). After several washing steps immunoprecipitated protein-DNA crosslinked complexes were eluted, the cross-links were reversed over night by incubation with 0.2 M NaCl and 20 μ g RNase A (Macherey Nagel, #740505) at 65 °C and de-proteinated with 40 μ g proteinase K (Qiagen, #19131). We detected C/EBP α bound fraction of DNA was measured by qRT-PCR. The protocol was adapted from Arrigoni *et al.*⁷.

Zebrafish experiments

All zebrafish experiments described in the present study were conducted on embryos younger than 5 days post-fertilization under the guidelines of the European Commission, Directive 2010/63/EU. Zebrafish husbandry was performed in accordance with European Union animal welfare guidelines⁸ and under the supervision of the local representative of the animal welfare agency (permit 35/9185.46/Uni Tü). The zebrafish strain used in this study

was *Danio rerio* wild type TE. The stable transgenic line Tg(*mpo:gfp*) was described previously⁹.

sgRNA target site selection

sgRNAs for zebrafish *gadd45bb* were designed with CCTop as described in Stemmer et al.¹⁰. Four sgRNAs were used (PAM in brackets): CGAGACAGTGTCTGCTGCAAG (AGG), CTGTGCCAGACGCCTCATGC(CGG), AGACCAAAGGAGCATCTGGG(TGG) and CTACGCTTTGCACTCCACGT(GGG). Cloning of sgRNA templates and in vitro transcription was performed as detailed in Stemmer et al.¹⁰.

Whole-mount *in situ* hybridization

RNA *in situ* hybridization of zebrafish embryos was performed as described previously¹¹ using digoxigenin-labeled RNA antisense probe for *gadd45bb* (accession number NM001012386, nucleotides 484-968) and *myeloperoxidase*, *mpo* (accession number BC056287, nucleotides 225-938).

In vitro transcription of mRNA

The pCS2+Cas9 plasmid was linearized using NotI and the mRNA was transcribed in vitro using the mMessage_mMachine SP6 Kit (ThermoFisher Scientific, AM1340).

Microinjection and heat treatment of zebrafish embryos

Zebrafish zygotes were co-injected with 15 ng/μl of the sgRNAs and 150 ng/μl of Cas9 mRNA. As negative control, 150 ng/μl of Cas9 mRNA without sgRNAs was injected into wild type embryos. Embryos were kept at 28°C in E3 medium with 200 μM 1-phenyl 2-thiourea (PTU) to prevent pigmentation. To ectopically express *csf3a*, full-length cDNA of zebrafish *csf3a* was cloned in a plasmid containing a bi-directional heat-inducible promoter¹² and Tol2 binding sites. The resulting plasmid, HSE:*csf3a*, was then co-injected at 20 ng/μl with 10 ng/μl of mRNA of Tol2 transposase. Embryos were then heat treated for 1 hour at 39 °C after one day. GFP-positive embryos were selected for whole-mount *in situ* hybridization analysis and genotyping.

Genotyping of gadd45bb crispants

To genotype gadd45bb crispants, genomic DNA was extracted from individual embryos after whole-mount in situ hybridization analysis using the QuickExtract™ DNA extraction solution (Epicentre, USA) according to the manufacturer's protocol. PCR with specific forward (5'-TTGCAAGATTTCACTGCGGC-3') and reverse (5'-ACGAAAGGTAAACATGTGCAA A-3') primers was used for subsequent sequence analysis. The investigators were blinded to allocation during experiments and outcome assessment.

RNA sequencing

RNA was isolated using the RNeasy Mini kit (Qiagen, #74106). The RNA concentration was determined by Qubit RNA High Sensitivity kit (Thermo Fisher Scientific, #Q32855). Library preparation and RNA sequencing on a HiSeq 4000 (single-read 50bp) were performed at the DKFZ Genomics and Proteomics Core Facility. RNA seq data analysis was performed on an input matrix of raw read counts loaded in R package DESeq2¹³. The counts from technical replicates were combined using the *collapseReplicates* function of DESeq2. Differences in gene expression between groups were quantified while controlling for individuals effects. The identified list of differently expressed genes was further characterized by performing analysis in Genomatix (Intrexon Bioinformatics, Germany), iRegulon^{14,15} and Heatmapper¹⁶.

Methylation analysis

DNA was isolated using QIAamp DNA Mini kit (Qiagen, #51306). The DNA concentration was determined by Qubit dsDNA High Sensitivity kit (Thermo Fisher Scientific, #Q32854). Methylation analysis using the Infinium Human MethylationEPIC beadchip, which covers 866.895 genome-wide CpGs, was performed at the DKFZ Genomics and Proteomics Core Facility. The analysis of methylation data was conducted using R packages *minfi*^{17,18} and *limma*¹⁹. Next, R packages *DMRcate* and *limma* were applied for the

identification of differentially methylated regions between groups²⁰. Genomatix and iRegulon analysis were performed to gain an insight into biological processes and gene pathways that differentially methylated CpGs might be involved in.

Supplemental figure legends

Supplemental Figure 1. GADD45 β expression is increased during granulopoiesis with highest levels in granulocytes. **A**, Analysis of GADD45 β expression using the public database *Bloodspot*²¹ revealed that early progenitors do not express much GADD45 β and during granulopoiesis it is increased with an expression peak in granulocytes (PMN).

Supplemental Figure 2. Gene modification efficiency of CRISPR/Cas9-mediated GADD45 β knockout in CD34⁺ HSPCs from three healthy donors that were used for RNA seq and EPIC methylation array

A-C, Gene modification efficiency of *GADD45B*-targeting CRISPR/Cas9 RNP was determined by Sanger sequencing (left) and indel percentage was determined by ICE webtool, Synthego (right). Representative chromatograms are depicted.

Supplemental Figure 3. EB-based myeloid differentiation of GADD45 β ^{-/-} iPSCs revealed an increased amount of immature and decreased amount of mature cells compared to control iPSCs

A, EB-based myeloid differentiation of GADD45 β -deficient iPSCs was performed and floating cells generated on day 14 of culture were analysed by FACS. **B**, Morphological evaluation using Wright-Giemsa stained cytopsin preparations of differentiated iPSCs on day 21 of culture. **A, B** Data from 3 independent experiments are depicted, means \pm SD; *, $p < 0.05$; **, $p < 0.01$; ***, $p < 0.001$.

Supplemental Figure 4. Injection of Cas9 mRNA alone does not affect neutrophil numbers. **A**, Cas9 mRNA was injected into wild-type embryos. After one day, the numbers of neutrophils were determined by whole-mount in situ hybridization using a mpo antisense

probe. The number of mpo-expressing cells in the trunk region was then compared to uninjected embryos. N indicates number of embryos. Data are means \pm SD.

Supplemental Figure 5. *GADD45 β* promoter has four putative C/EBP binding sites

A, 654 bp sequence of the *GADD45 β* gene till ATG start codon with four predicted C/EBP binding sites (b.s.1-4, indicated in red) is depicted.

Supplemental Figure 6. iRegulon motif enrichment analysis of genes upregulated after G-CSF treatment of healthy control HSPCs, but not *GADD45 β* -KO cells

A, List of significantly enriched motifs and transcription factors binding to these motifs are presented.

Supplemental Figure 7. Gene ontology and motif enrichment analysis of differentially hypomethylated CpGs upon *GADD45 β* KO

A,B, Genomatix gene ontology analysis of biological processes (**A**) and signal transduction pathways (**B**) related to *GADD45 β* -dependent DNA demethylation. **C**, Significantly enriched motifs of hypomethylated genes were identified using iRegulon motif enrichment analysis.

Supplemental Figure 8. Differentially methylated regions (DMRs) related to myeloid-specific genes are hypomethylated upon G-CSF treatment of control but not *GADD45 β* -KO cells. Presented are DMRs of *HDAC5*, *AZU1*, *MPO*, *CTSG*, *SLPI*, *CXCR2*, *GF11*, *PRTN3* and *IL1R2*. The number of DMPs in each DMR is indicated in parenthesis.

Supplemental Figure 9. Gene ontology analysis of genes that are differentially expressed and methylated in G-CSF-treated HSPCs upon *GADD45 β* knockout

A, Gene ontology analysis was performed using Genomatix software. Selected significant signal transduction pathways are presented.

Supplemental tables

Supplemental table 1. Differentially expressed genes between G-CSF treated control and GADD45 β -KO group

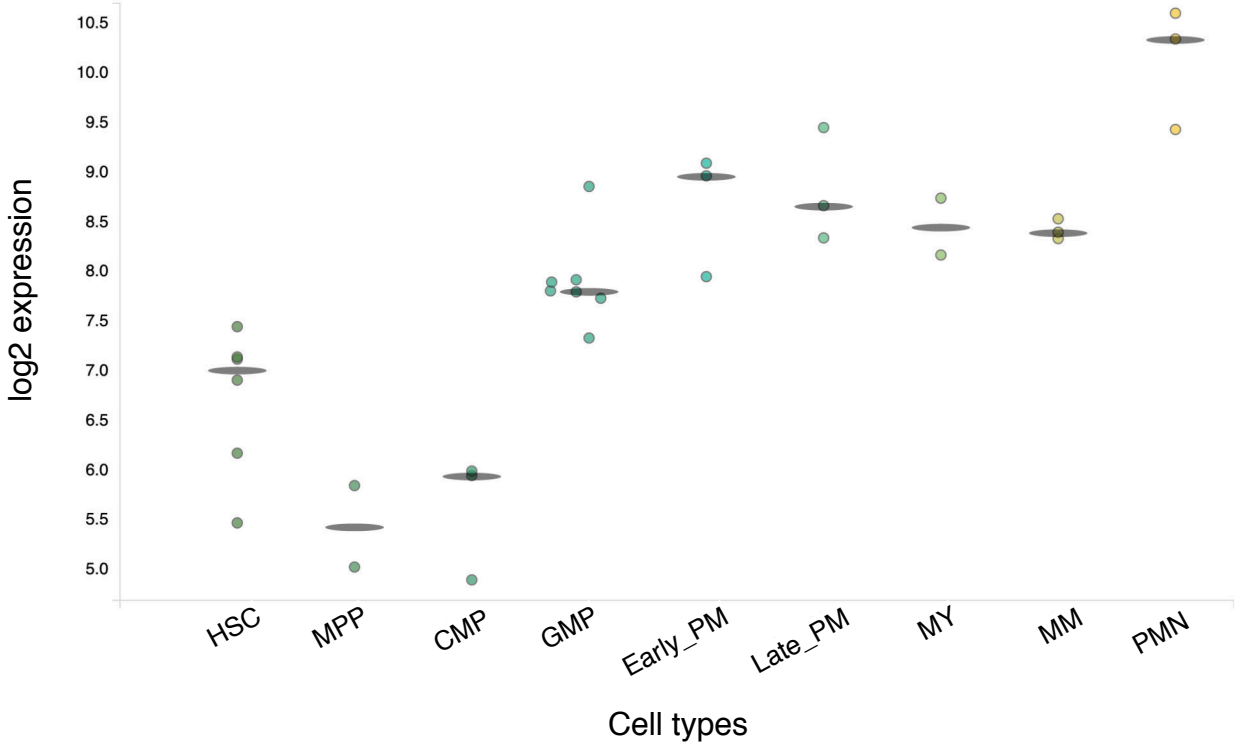
Supplemental table 2. Differentially methylated sites (DMPs) between G-CSF treated and untreated control cells

Supplemental table 3. Genomic locus enrichment analysis of differentially methylated regions between G-CSF treated control and GADD45 β -KO groups

Supplemental references

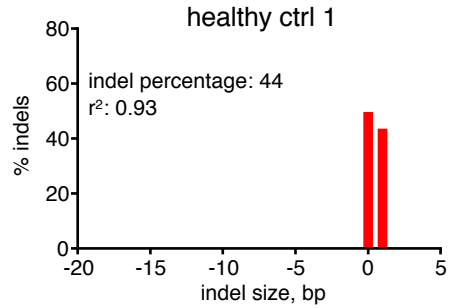
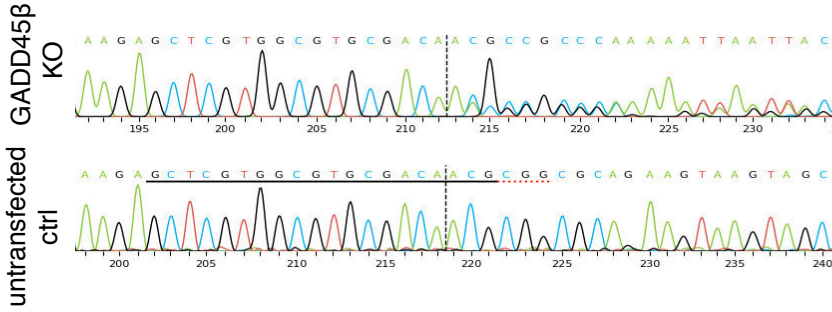
1. van Galen P, Kreso A, Mbong N, et al. The unfolded protein response governs integrity of the haematopoietic stem-cell pool during stress. *Nature*. 2014;510(7504):268-272.
2. Dannenmann B, Zahabi A, Mir P, et al. Human iPSC-based model of severe congenital neutropenia reveals elevated UPR and DNA damage in CD34(+) cells preceding leukemic transformation. *Exp Hematol*. 2019;71:51-60.
3. Skokowa J, Klimiankou M, Klimenkova O, et al. Interactions among HCLS1, HAX1 and LEF-1 proteins are essential for G-CSF-triggered granulopoiesis. *Nat Med*. 2012;18(10):1550-1559.
4. Schambach A, Mueller D, Galla M, et al. Overcoming promoter competition in packaging cells improves production of self-inactivating retroviral vectors. *Gene Ther*. 2006;13(21):1524-1533.
5. Naldini L. Lentiviruses as gene transfer agents for delivery to non-dividing cells. *Curr Opin Biotechnol*. 1998;9(5):457-463.
6. Soderberg O, Gullberg M, Jarvius M, et al. Direct observation of individual endogenous protein complexes in situ by proximity ligation. *Nat Methods*. 2006;3(12):995-1000.
7. Arrigoni L, Richter AS, Betancourt E, et al. Standardizing chromatin research: a simple and universal method for ChIP-seq. *Nucleic Acids Res*. 2016;44(7):e67.
8. Bert B, Chmielewska J, Bergmann S, et al. Considerations for a European animal welfare standard to evaluate adverse phenotypes in teleost fish. *EMBO J*. 2016;35(11):1151-1154.
9. Renshaw SA, Loynes CA, Trushell DM, Elworthy S, Ingham PW, Whyte MK. A transgenic zebrafish model of neutrophilic inflammation. *Blood*. 2006;108(13):3976-3978.
10. Stemmer M, Thumberger T, Del Sol Keyer M, Wittbrodt J, Mateo JL. CCTop: An Intuitive, Flexible and Reliable CRISPR/Cas9 Target Prediction Tool. *PLoS One*. 2015;10(4):e0124633.
11. Aghaallaei N, Bajoghli B, Czerny T. Distinct roles of Fgf8, Foxi1, Dlx3b and Pax8/2 during otic vesicle induction and maintenance in medaka. *Dev Biol*. 2007;307(2):408-420.
12. Bajoghli B, Aghaallaei N, Heimbucher T, Czerny T. An artificial promoter construct for heat-inducible misexpression during fish embryogenesis. *Dev Biol*. 2004;271(2):416-430.
13. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol*. 2014;15(12):550.
14. Janky R, Verfaillie A, Imrichova H, et al. iRegulon: from a gene list to a gene regulatory network using large motif and track collections. *PLoS Comput Biol*. 2014;10(7):e1003731.
15. Verfaillie A, Imrichova H, Janky R, Aerts S. iRegulon and i-cisTarget: Reconstructing Regulatory Networks Using Motif and Track Enrichment. *Curr Protoc Bioinformatics*. 2015;52:2 16 11-39.
16. Babicki S, Arndt D, Marcu A, et al. Heatmapper: web-enabled heat mapping for all. *Nucleic Acids Res*. 2016;44(W1):W147-153.
17. Aryee MJ, Jaffe AE, Corrada-Bravo H, et al. Minfi: a flexible and comprehensive Bioconductor package for the analysis of Infinium DNA methylation microarrays. *Bioinformatics*. 2014;30(10):1363-1369.
18. Fortin JP, Triche TJ, Jr., Hansen KD. Preprocessing, normalization and integration of the Illumina HumanMethylationEPIC array with minfi. *Bioinformatics*. 2017;33(4):558-560.
19. Ritchie ME, Phipson B, Wu D, et al. limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res*. 2015;43(7):e47.
20. Peters TJ, Buckley MJ, Statham AL, et al. De novo identification of differentially methylated regions in the human genome. *Epigenetics Chromatin*. 2015;8:6.
21. Bagger FO, Kinalis S, Rapin N. BloodSpot: a database of healthy and malignant haematopoiesis updated with purified and single cell mRNA sequencing profiles. *Nucleic Acids Res*. 2019;47(D1):D881-D885.

A

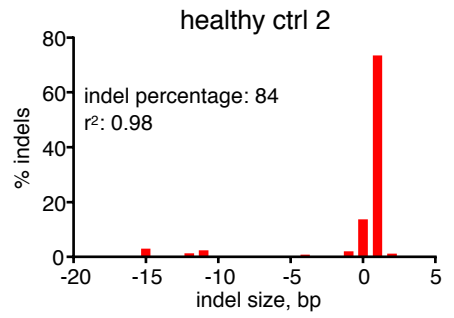
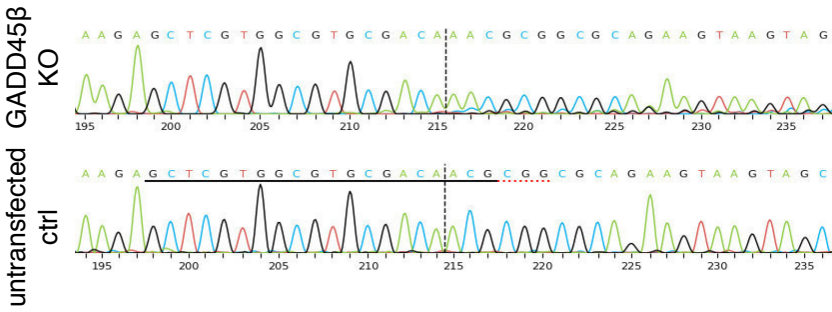


HSPCs

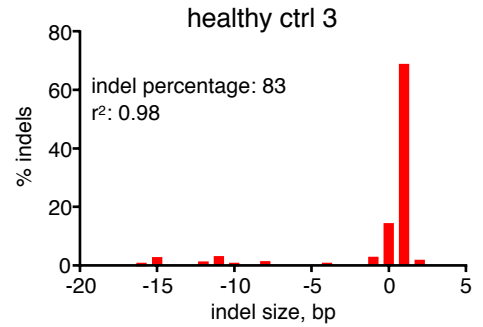
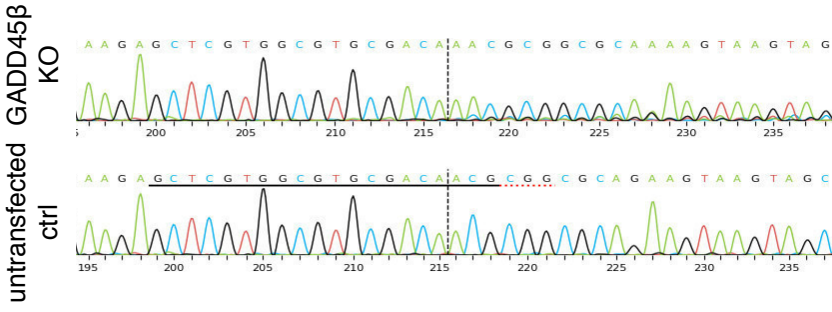
A

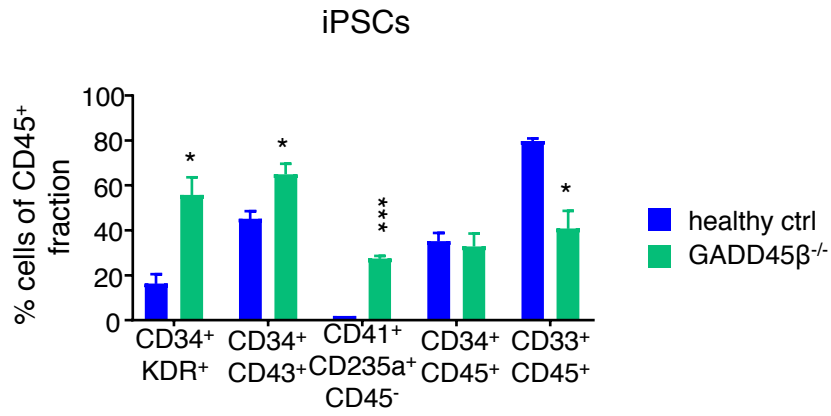
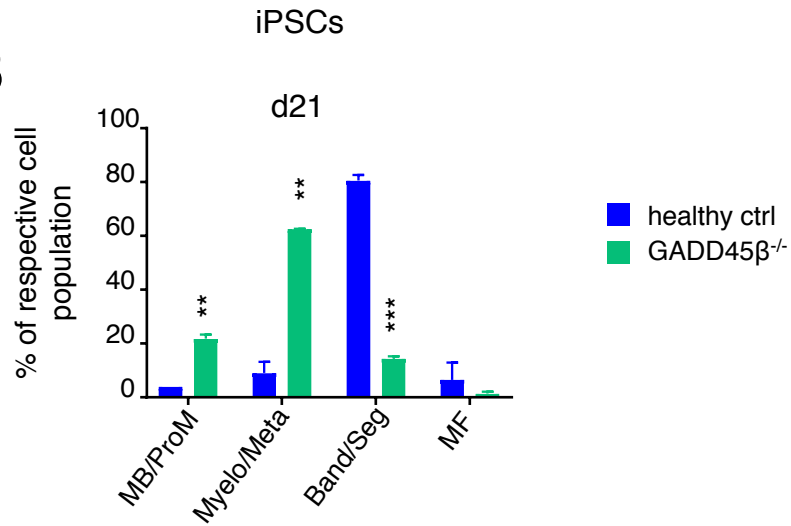


B

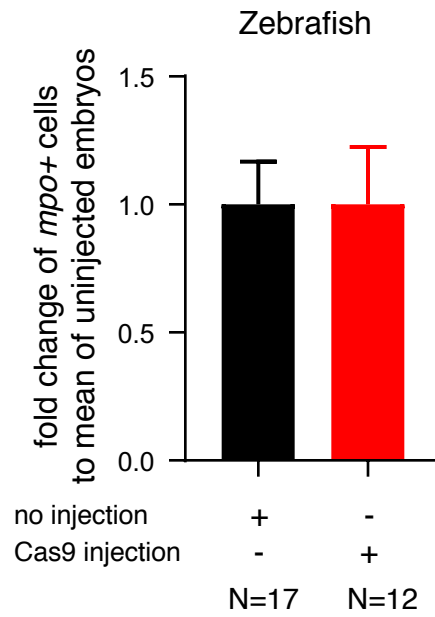


C



A**B**

A


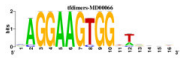


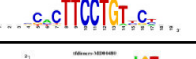
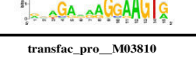
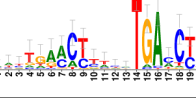



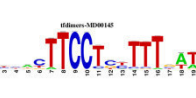
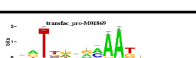

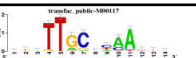


A

b.s. 1
CCTTCCCGGTGCAGCCCCTCCACCCCCAGCAGAACTTGGGAAAGGCGCG
GTCCGGGACTCTCCGCGGATCGGGAGGGGATTCCAGGCCCCCCCGAAA
GTCCGGGCCGCCTCGCGCGCTGGAAATCCCGCGCGCGCCCCGAACCGC
GGCTCGGCTGCCGGGAAATCAGGAGAAAAAACTTCTGCTTTTTTTTTCTTT
TCTGGCATTGCGGGTACCTACCCGGCCCCCGCGCGCCCTCCTCCCGGT
TCTCGCCCCACGTGGGGCGCCCCGCACGCCGCTCCTCCCCCTCCCC
b.s. 2
TCGTCGGCCAACCGCAGAGCTAGCTGCACTCGCCCTTGTCTTTCCACC
AATAGGAGGGGCGAATGACTCCACTGAGGCCACGCCCAATGTTCAAGTC
TATAAAAGTCGGTGCCGGAGGCTCCCAGCTCAGATCGCCGAAGCGTCGG
b.s. 3
ACTACCGTTGGTTTCCGCAACTTCCTGGATTATCCTCGCCAAGGACTTTG
CAATATATTTTTCCGCCTTTTCTGGAAGGATTTGCTGCTTCCCGAAGGTC
TTGGACGAGCGCTCTAGCTCTGTGGGAAGGTTTTGGGCTCTCTGGCTCG
b.s. 4
GATTTTGCATTTCTCCCTGGGGACTGCCGTGGAGCCGCATCCACTGTGG
ATTATAATTGCAACATG

A

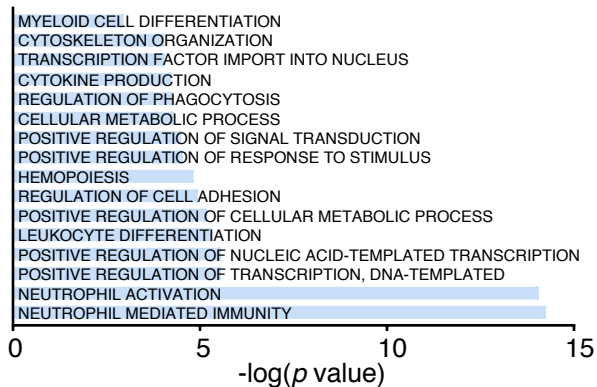
HSPCs

Enriched Motif	NES	AUC	Transcription factors	Targets
	5.194	0.089	RARA, RARB, ETS1	9
	5.117	0.088	ELF1, CBF, RUNX2	8
	4.496	0.079	ETS1	4
	4.289	0.076	LTF	7
	4.136	0.074	NR3C1, STAT1	8
	3.947	0.071	TAL1	4
	3.682	0.068	RXRA, RARG, RXRB, RARA, RARB, FOS, PPARG *	7
	3.570	0.066	SPI1	21
	3.471	0.065	NFKB1, NFKB2	5
	3.393	0.064	STAT1, LEF1	5
	3.358	0.063	ELK4, ELF2, GABPB1, TCF4, FLI1, ELK1, ELF4, ETV6, ETV7	10
	3.289	0.062	CEBPG	13
	3.219	0.061	IRF4, SPI1, ETS1, GABPB1	26
	3.100	0.059	CEBPB	22
* Binding with FDR < 2.5x10 ⁻⁵				

A

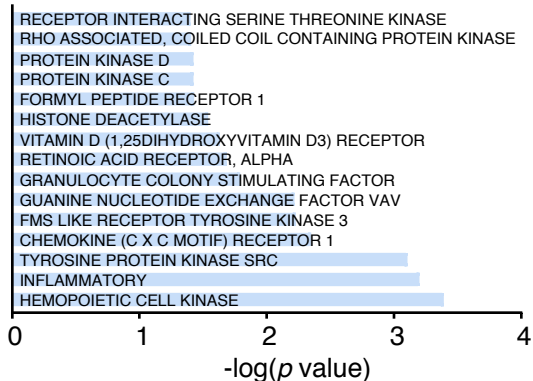
HSPCs

biological process

**B**

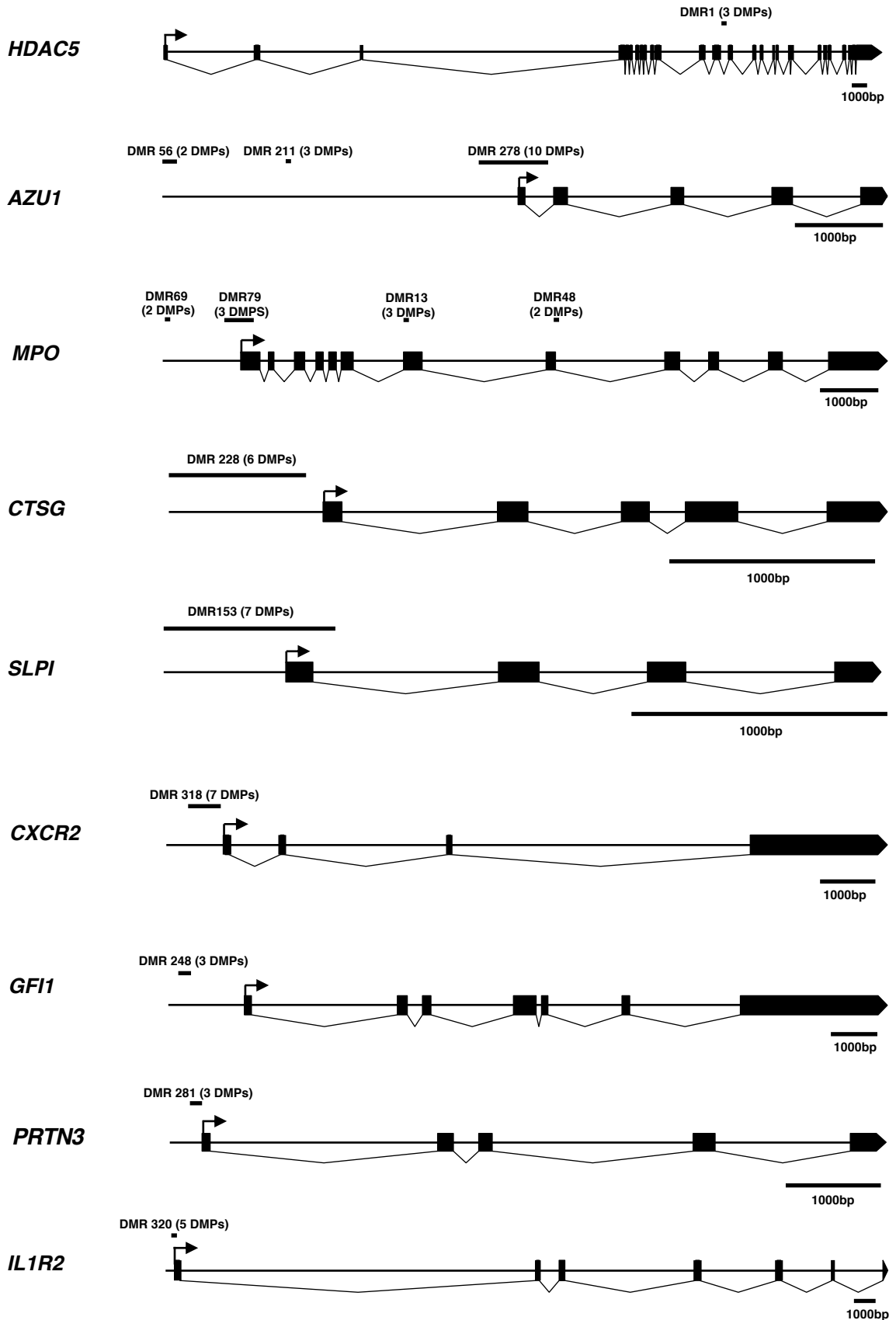
HSPCs

signal transduction

**C**

HSPCs

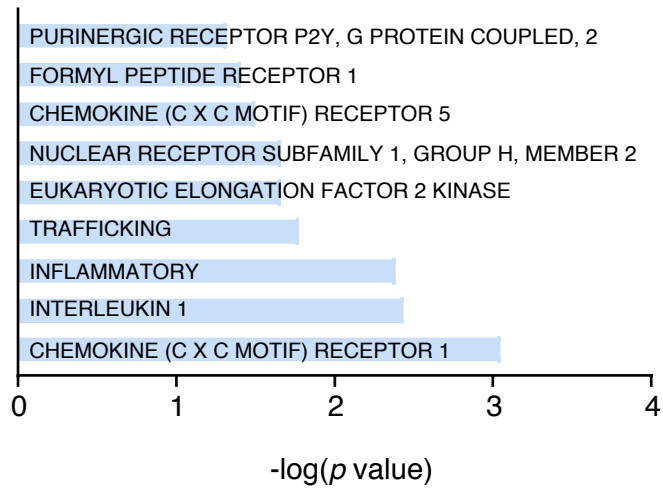
Enriched Motif	NES	AUC	Transcription factors	Targets
	3.297	0.04	FOXP3	74
	3.226	0.04	ETV7	100
	3.214	0.04	CEBPE, CEBPG, CEBPD, SOX9, CEBPB, CEBPA	111
	3.152	0.039	ZBTB3	164
	3.061	0.039	EP300	85
	3.047	0.039	IRF4	127
	3.029	0.039	FOXO1, GABPA	64

A

A

HSPCs

signal transduction



MANUSCRIPT E (IN REVISION):

**hiPSC Model of Stage-Specific Leukemogenesis in Severe Congenital Neutropenia
reveals BAALC as a Key Oncogene**

hiPSC Model of Stage-Specific Leukemogenesis in Severe Congenital Neutropenia reveals

BAALC as a Key Oncogene

Benjamin Dannenmann^{1*}, Maksim Klimiankou^{1*}, Benedikt Oswald¹, Anna Solovyeva¹, Jehan Mardan¹, Ann-Christin Krahl¹, Masoud Nasri¹, Malte Ritter¹, Azadeh Zahabi¹, Perihan Mir¹, Frederic Stein¹, Regine Bernhard¹, Nico Lachmann², Thomas Moritz², Tatsuya Morishima¹, Tim Ripperger³, Doris Steinemann³, Cornelia Zeidler⁴, Karl Welte⁵ and Julia Skokowa¹

¹Department of Oncology, Hematology, Immunology, Rheumatology and Pulmonology, University Hospital Tuebingen, Tuebingen, Germany; ²Institute of Experimental Hematology, Hannover Medical School, Hannover, Germany; ³Institute of Human Genetics, Hannover Medical School, Hannover, Germany; ⁴Department of Oncology, Hematology, Immunology and Bone Marrow Transplantation, Hannover Medical School, Hannover, Germany; ⁵University Children's Hospital Tuebingen, Tuebingen, Germany

* these authors contributed equally to the work

#**Correspondence:** Julia Skokowa, Division of Translational Oncology, Department of Hematology, Oncology and Clinical Immunology, University Hospital Tübingen, Germany, e-mail: Julia.Skokowa@med.uni-tuebingen.de

Summary

We reported an extremely high frequency of cooperating acquired *CSF3R* (colony-stimulating factor 3 receptor) and *RUNX1* (runt-related transcription factor 1) mutations in patients with pre-leukemic bone marrow failure syndrome, severe congenital neutropenia (CN), who developed acute myeloid leukemia (CN/AML). We established an *in vitro* model of stepwise leukemogenesis in CN/AML through CRISPR/Cas9-mediated gene-editing of induced pluripotent stem cells (iPSCs) from CN/AML patients. We identified *BAALC* (brain and acute leukemia, cytoplasmic) upregulation as a key leukemogenic event: *BAALC* knockout restored myeloid differentiation and inhibited proliferation of CN/AML-iPSCs derived hematopoietic stem cells. We detected a selective inhibitor of p38 α -mediated MK2a phosphorylation, CMPD1, which selectively inhibits the growth of CN/AML cells. Strikingly, the treatment of primary blasts of CN/AML patients with CMPD1 resulted in a marked reduction of cell proliferation without affecting the growth of healthy donor hematopoietic stem cells. Thus, targeting of *BAALC* may prevent leukemogenic transformation or may eliminate AML blasts in CN/AML.

Keywords

Severe congenital neutropenia, pre-leukemic bone marrow failure syndrome, AML, acquired *CSF3R* mutations, *RUNX1* mutations, iPSCs, CRISPR/Cas9 gene-editing, *BAALC*.

Highlights

- iPSC model recapitulates step-wise leukemia development in CN
- *BAALC* knockout restores differentiation and inhibits proliferation of CN/AML cells
- CMPD1 selectively inhibits the growth of CN/AML cells

eTOC Blurb

Dannenmann et al. generated iPSCs (induced pluripotent stem cells) from patients with pre-leukemia bone marrow failure syndrome, severe congenital neutropenia (CN) who overt acute myeloid leukemia (AML). Using CRISPR/Cas9 gene editing and aided by RNA-Seq data from CN/AML patients blasts, they created an *in vitro* experimental model that recapitulated leukemia development in CN. Using this model, they identified BAALC as an essential leukemogenic protein: *BAALC* knockout completely reverses the defective hematopoiesis of CN/AML-iPSCs. Furthermore, they identified a selective inhibitor of p38 α -mediated MK2a phosphorylation, CMPD1, as a potential drug capable of killing CN/AML cells.

Introduction

Pre-leukemia bone marrow failure syndromes comprise a group of inborn hematological disorders, characterized by abnormal differentiation and functions of hematopoietic stem and progenitor cells (HSPCs). They frequently culminate in the development of myelodysplastic syndrome (MDS) or leukemia, usually acute myeloid leukemia (AML)¹⁻⁸. Understanding the molecular mechanisms underlying leukemia development in inborn blood disorders would aid in the development of therapeutic tools for preventing or treating leukemia and provide insight into the pathomechanisms of *de novo* MDS and AML. The prevailing hypothesis of leukemia development in inborn pre-leukemic syndromes postulates that homeostasis, self-renewal, proliferation or differentiation of HSPCs is disturbed owing to inherited mutations^{1-4,6-8}. These “unfit” HSPCs are exposed to constant stress because of intrinsic defects and extrinsic abnormalities of the HSPC niche, which together may cause an elevated DNA-damage response or diminished DNA repair, leading to the acquisition of mutations and/or chromosomal abnormalities in leukemia-associated genes^{8,9}. The background of an inherited disease can create selective pressure that supports the outgrowth and evolution of mutant HSPCs clones¹⁰. Severe congenital neutropenia (CN) is a pre-leukemia syndrome that, in the majority of patients, is caused by heterogeneous *ELANE* mutations encoding neutrophil elastase (NE)¹¹. The HSPCs of CN patients fail to differentiate into neutrophilic granulocytes, without any severe maturation defects in other blood lineages^{7,8,12,13}. The cumulative incidence of MDS or AML in CN patients is approximately 20% after 20 years^{8,14,15}. Exposure of CN-HSPCs to high concentrations of granulocyte colony-stimulating factor (G-CSF) partially reverses granulocytic maturation defects^{8,13}. However, there is a clear correlation^{8,13} between susceptibility to G-CSF therapy and frequency of leukemia, such that patients who require more than 8 µg/kg/d of G-CSF have a higher probability of developing MDS or AML^{14,15}. We recently reported a high frequency of acquired cooperative mutations in *CSF3R* (encoding the G-CSF receptor) and *RUNX1* (runt-

related transcription factor 1) in 55 % of CN/AML patients (31 patients were studied)¹⁶. In a majority of cases, the appearance of HSPCs clones carrying a *CSF3R* mutation preceded the co-acquisition of *RUNX1* mutations in the same cells¹⁶. In another multicenter study of genotype-phenotype correlations in *ELANE*-associated CN patients, we identified “hot-spot” *ELANE* mutations (e.g., p.C151R, p.C151Y, p.G214R) in CN patients who developed MDS or AML¹⁷. The mechanism of leukemogenic transformation of HSPCs downstream of acquired *CSF3R* and *RUNX1* mutations on an *ELANE*-mutated background is unclear. *CSF3R* mutations are stop-codon mutations that are localized in the intracellular domain of the G-CSF receptor, which is responsible for the termination of SOCS3 (suppressor of cytokine signaling 3)-triggered, STAT5 (signal transducer and activator of transcription 5)-mediated proliferative signals and STAT3-dependent activation of differentiation^{7,18-27}. *RUNX1* mutations are mostly located in the DNA-binding Runt homology domain (RHD) and a transactivation domain (TAD)²⁸. Point mutations in *RUNX1* have been described in *de novo* AML or AML secondary to MDS, radiation exposure, or chemotherapy at frequencies of 6% to 33%²⁹⁻³³. A high incidence of *RUNX1* mutations has been associated with monosomy 7, trisomy 21, and trisomy 13^{32,34,35}. Acquired translocations, deletions, or mutations in *RUNX1* have been found in Fanconi anemia patients who developed AML or MDS⁹. *RUNX1* germline mutations have also been described in a familial platelet disorder (FPD) with a predisposition to AML^{36,37}. Both missense and nonsense *RUNX1* mutations have been detected in CN/AML patients^{16,38}. Numerous studies have reported a diversity of mechanisms underlying the pathogenesis of missense and nonsense *RUNX1* mutations^{37,39,40}. *RUNX1* mutations in the RHD may result in loss of DNA and/or CBF β binding, protein mislocalization, disturbed heterodimerization, inefficient transactivation, haploinsufficiency, or appearance of dominant-negative alleles⁴⁰⁻⁴⁵. Missense *RUNX1* mutations may create dominant-negative protein isoforms that exhibit severely impaired DNA binding, protein mislocalization, and slightly or moderately decreased CBF β binding^{40,42,44,45}. Moreover, *RUNX1* protein with a missense

Runt mutation may sequester or alter the activity of other RUNX1-binding proteins that are present in the cell in limiting amounts, thereby exerting inhibitory effects on transactivation mediated by wild-type (WT) RUNX1 protein^{37,42}. Interestingly, *RUNX1* maps to chromosome 21, and a majority of leukemic blasts from CN/AML patients harbor trisomy 21¹⁶. This suggests that both, trisomy 21 and the ratio of mutant to WT *RUNX1* alleles may contribute to leukemogenesis.

There are no animal models of CN; neither transgenic mice with a knock-in of the human *ELANE* mutation⁴⁶ nor *Elane*- and *Hax1*-knockout (KO) mice exhibit a neutropenia phenotype in a steady-state condition. Thus, the establishment of the viable experimental model would be essential for studying multistep leukemogenesis in CN. Despite some limitations, patient-specific induced pluripotent stem cells (iPSCs) are an excellent alternative to animal models. Generation of CN-patient-specific iPSCs that recapitulate the maturation arrest of granulopoiesis has been previously described⁴⁷⁻⁴⁹. iPSCs have also been used to study leukemogenesis of *de novo* AML samples⁵⁰ and to screen compounds that may target AML cells. Using CRISPR/Cas9 gene-editing, it is possible to introduce distinct mutations in iPSCs that allow the study of stepwise, stage-specific leukemia progression⁵¹. With this approach, it is possible to develop experimental leukemia models with distinct endogenous genetic abnormalities that otherwise cannot be generated, as for example, to compare functional outcomes of different types of mutations (e.g., missense versus nonsense) in endogenously expressed proteins (e.g., RUNX1).

In the present study, we established an *in vitro* model of the leukemia evolution in CN using patients-derived iPSCs and CRISPR/Cas9-mediated introduction of gene mutations. Using this model, we identified the upregulation of *BAALC* (brain and acute leukemia, cytoplasmic) as a key leukemogenic event. We identified a small molecule inhibitor that kills CN/AML iPSC-derived and primary CN/AML blasts without affecting healthy donor HSPCs. This strategy could be applied to treat or prevent leukemia in patients with *RUNX1* mutations.

Results

Missense *RUNXI* mutations are associated with trisomy 21 in CN/AML patients

A comparison of cytogenetic abnormalities and type of acquired *RUNXI* mutations in 23 CN/AML patients revealed that 14 patients had missense, 5 patients truncated, and 4 patients compound heterozygous *RUNXI* mutations (missense, nonsense, and intronic) (**Table S1**). 7 of 14 CN/AML patients with missense *RUNXI* mutations co-acquired trisomy 21. In contrast, trisomy 21 was detected in 4 of 5 CN/AML patients with nonsense *RUNXI* mutations. A digital PCR (dPCR) analysis revealed the presence of two copies of mutated missense *RUNXI* in three CN/AML patients with trisomy 21 (**Figure 1A, Figure S1A**), suggesting a crucial role of the dosage of missense *RUNXI* for leukemia initiation.

In vitro iPSC model of stepwise leukemogenic transformation in CN/AML

We established iPSC based model of stepwise leukemia development in CN. We generated iPSCs from two CN/AML patients harboring *ELANE* mutations, p.C151Y and p.G214R. Using PBMNCs from CN patient 1 at overt AML, we generated iPSC clones with *ELANE* p.C151Y mutation only (CN1, derived from non-leukemic PBMNCs) or acquired nonsense *CSF3R* p.Q741X in combination with missense *RUNXI* p.R139G mutation and trisomy 21 (CN/AML1.1 and 1.2). iPSC clones of second CN patient without- and with nonsense *CSF3R* p.Q743X mutation (CN2 and CN2+*CSF3R* Q743X+/-, respectively) were generated from BMMNCs isolated at the AML stage (**Figure 1B, Figure S2A**).

Array-CGH (comparative genomic hybridization) revealed no chromosomal abnormalities in CN1-iPSC clone, trisomy 21 in CN/AML-iPSC clone 1.2, and the additional gain of part of chromosome 12 in the second CN/AML-iPSCs clone 1.1 (**Figure S2B**). The gain of chromosome 12 is a frequent finding during iPSC maintenance and reflects the high proliferation rate of these cells⁵². No chromosomal abnormalities were detected in iPSC clones derived from the CN patient 2 (**Figure S2B**).

All iPSC lines expressed pluripotent stem cell markers (**Figure S3A, B, C**) displayed alkaline phosphatase activity (**Figure S3D, E**), differentiated into three germ layers (data not shown) and had inactivated lentiviral plasmid used for reprogramming (**Figure S3A-B**).

CN/AML1 iPSCs had trisomy 21 and dPCR showed one copy of WT and two copies of mutated *RUNX1* in iPSCs as well as in iPSC-derived CD45⁺ cells (**Figure 1C**).

We next introduced frame-shift *CSF3R* mutations in CN1 iPSC clone using CRISPR/Cas9 gene-editing (**Table S2**). We generated two iPSC lines, one with the heterozygous *CSF3R* mutation p.(Arg734SerfsX35) (CN1+*CSF3R* R734fs) another with the compound heterozygous *CSF3R* mutations, p.(Gln741AsnfsX59) and p.(Ser742PhefsX29) (CN1+*CSF3R* Q741fs) (**Figure 1B, Figure S4A, B**). Heterozygous frame-shift *RUNX1* mutation, p.(Glu175SerfsX7) was introduced into CN2+*CSF3R* Q743X+/- iPSC clone using CRISPR/Cas9 (CN/AML2 iPSC clone) (**Figure 1B and Figure S4C**). No off-target effects were observed in the CRISPR/Cas9 gene-edited iPSCs (**Table S3**).

Impaired myeloid differentiation and elevated proliferation of CN/AML-iPSC derived HSPCs

We performed an *in vitro* embryoid body (EB)-based iPSC differentiation into hematopoietic stem and mature myeloid cells⁵³ (**Figure S5A**). The proportion of CD45⁺CD34⁺ cells was markedly higher in iPSC lines generated from CN/AML1 patient, as compared to the healthy donor (HD) derived iPSCs. For CN/AML2 patient, the proportion of CD45⁺CD34⁺ cells was unchanged (**Figure 2B**), but CD34⁺CD43⁺ and CD45⁻CD41a⁺CD235⁺ cell populations were increased (**Figure S5B, C**). We observed dramatically increased proliferation of CN/AML iPSC-derived HSPCs from both patients, that were collected at day 14 of EB differentiation and further cultured on the FLT3-L-secreting feeder cells, as compared to HD- or CN - iPSC derived HSPCs (**Figure 2C, D**). The proliferation of HSPCs expressing mutated *CSF3R* was higher than that of CN cells, but much lower, as of CN/AML

cells (**Figure 2C, D**). Differentiation into $CD45^+CD15^+CD16^+$ and $CD45^+CD15^+CD11b^+$ granulocytes was markedly reduced in CN-iPSCs and was almost completely abolished in CN/AML-iPSCs compared with HD-iPSCs (**Figure 2E, F**). *CSF3R* mutations in CN-iPSCs also led to reduced granulocytic maturation (**Figure 2E, F**). Morphological analysis of cytopsin preparations revealed almost no mature myeloid cells in CN/AML samples (**Figure 2G, H**). Additionally, $CD34^+$ cells derived from CN-iPSC lines produced reduced numbers of CFU-Gs and CFU-GMs compared with HD cells. These numbers were further decreased in *CSF3R*-mutated and CN/AML cells (**Figure 2I, J**).

AML-like gene expression signature in CN/AML-derived HSPCs

We next examined whether genes that are upregulated in primary CN/AML blasts also elevated in the CN/AML iPSCs-derived HSPCs. We found that genes specific for the early HSPCs, such as *BAALC* (brain and acute leukemia, cytoplasmic), *HPGDS* (hematopoietic prostaglandin D synthase), *CD34* and *CD109* were increased in primary CN/AML blasts of five CN/AML patients, compared with BM $CD34^+CD33^+$ cells isolated from three CN patients prior to leukemia development (**Figure 3A, Table S1**). An increase in *BAALC*, *CD34*, *HPGDS* and *CD109* expression was also observed in $CD34^+$ cells derived from CN/AML1 iPSC lines compared with $CD34^+$ cells from CN1 iPSCs (**Figure 3B**). In HSPCs of CN/AML patient 2, we measured elevated levels of *BAALC* mRNA, as compared to CN-iPSC-derived HSPCs (**Figure 3C**). *BAALC* protein levels were increased in CN/AML-derived HSPCs of both patients (**Figure 3D**).

We further tested whether elevated *BAALC* levels are dependent on the presence of *RUNX1* mutations. Indeed, transduction of $CD34^+$ cells of healthy individuals with lentivirus-based constructs with mutated *RUNX1* (p.R139G and p.R174L) led to markedly increased *BAALC* mRNA, as compared to WT *RUNX1* transduced cells (**Figure 3E**).

We previously detected elevated levels of STAT5a in HSPCs of CN patients and primary blasts of CN/AML patients⁵⁴. Upregulated mRNA and protein levels of STAT5a

were detected in CN/AML-iPSC-derived HSPCs for both CN/AML patients. Interestingly, *RUNX1* mRNA and protein expression was markedly increased in CN/AML patient 1 cells with missense *RUNX1* mutation and trisomy 21, but in CN/AML patient 2 who harbors truncated *RUNX1* mutation, *RUNX1* mRNA only was elevated. G-CSFR expression was also enhanced in CN/AML patient 1, but not in CN/AML patient 2 (**Figure 3F-I**).

Elevated *BAALC* expression is essential for leukemogenic transformation in CN

To determine which of the candidate leukemia-associated genes, *BAALC*, *HPGDS*, *CD109* or *RUNX1*, is responsible for the leukemogenic transformation in CN, we generated knockout iPSC lines of CN/AML patient 1 for each of the selected genes (**Figure 4A**). For this, we created sgRNAs specific for the nucleotide sequence of the first exon of each gene, cloned them into a pSpCas9(BB)-2A-GFP (PX458) construct⁵⁵ and transfected CN/AML-iPSC lines (**Tables S2**). Transfected CN/AML-iPSC cells were sorted based on GFP signal, and single-cell clones with homozygous (*RUNX1*, *BAALC*), or compound heterozygous (*CD109*, *HPDGS*), mutations were selected for further analysis (**Figure S4D-G**). Pluripotency of selected iPSC clones was unaffected, and there was no evidence of off-target activity of CRISPR-Cas9 used for gene knockout (**Tables S3 and data not shown**).

Interestingly, *BAALC* knockout resulted in a dramatic induction of granulocytic differentiation and markedly reduced proliferation of CN/AML-iPSCs derived CD34⁺ cells (**Figure 4B-E**). Almost no differences were observed in the stem cell composition between gene-edited groups (**Figure S6A**). A knockout of *CD109*, *RUNX1* or *HPGDS* did not restore granulocytic differentiation of CN/AML-iPSCs (**Figure 4B-E, Figure S6A**).

Since *BAALC* mRNA and protein levels were also upregulated in CN/AML iPSC-derived hematopoietic cells of patient 2 (**Figure 3C, D**), we tested the effects of *BAALC* knockout on the proliferation and granulocytic differentiation in CN/AML patient 2 as well. Indeed, similar to CN/AML patient 1, we found an increase in granulocytic differentiation and markedly reduced proliferation of HSPCs derived from CN/AML2 iPSCs (**Figure 4F-I**) but

an increased proportion of CD45⁺CD34⁺ cells upon *BAALC* knockout (**Figure S6B**). These data strongly support the essential role of *BAALC* in the CN/AML leukemogenesis.

Of note, *BAALC* KO in healthy donor iPSCs (**Figure S6C**) did not strongly affect granulocytic differentiation (**Figure 4J**). At the same time, *BAALC* KO induced granulocytic differentiation of CN1 patient-derived iPSCs (**Figure 4J**).

Leukemia-specific gene expression signature in CN/AML-iPSCs derived HSPCs

Differential gene expression analyses of RNA-Seq of iPSC-derived HSPCs using DESeq2 R package⁵⁶ identified 132 up- and 570 down-regulated, as well as 579 up- and 1422 down-regulated genes between CN/AML and CN stages for CN patient 1 and CN patient 2, respectively ($\log_2FC > 1$ or < -1 , adj. P-value < 0.05 ; **Figure 5A-D, Table S4, S5**). To the top significant pathways in the gene set enrichment analysis (GSEA) belonged the E2F targets, the oxidative phosphorylation, an enrichment of the MYC targets and the reduction of the platelet - specific genes in CN/AML1, as compared to CN1 samples (**Figure 5C**). In CN/AML patient 2, we detected enrichment of the gene sets for the G2M checkpoint, the E2F targets as well as TGF β signaling, whereas gene set “GO:Structural Constituent of Ribosomes” was enriched in CN2 (**Figure 5D**).

To predict transcription factors that control differentially expressed genes in both CN/AML patients, in comparison to the CN stage, we performed eXpression2Kinases analysis⁵⁷. Among the transcription factor binding motifs significantly enriched in CN/AML cells (P-value < 0.05), we detected motifs for RUNX1; for “classical” RUNX1 binding partners as *GATA1*, *GATA2*; as well as for AML-associated genes such as *SUZ12* and *EZH2* (**Figure 5E**). Patient-specific motives were *TRIM28* and *TP53* for CN/AML1 as well as *CEBPB*, *NANOG*, *KLF4* for CN/AML2 (**Figure 5E**).

Using kinase enrichment analysis (KEA)⁵⁷, we predicted kinases that phosphorylate proteins regulated by the subnetwork of prioritized transcription factors. The vast majority of enriched kinases including, the HIPK2, MAPK1/3/14, CSNK2A1, ERK1 and AKT1, were

shared between both patients at the selected threshold (P-value $<10^{-8}$), (**Figure 5F**). Unique enriched kinases (e.g., CDK1, JNK1, ERK2) were detected in patient 2 only (**Figure 5F**).

BAALC-dependent leukemogenic signaling pathways in CN/AML

We further examined the BAALC-dependent leukemia-associated gene expression. We compared CN/AML cells before- and after *BAALC* KO. We identified 165 up- and 254 down-regulated, as well as 185 up- and 381 down-regulated genes ($\log_2FC > 1$ or < -1 , adj. P-value < 0.05) in CN/AML versus CN/AML *BAALC* KO comparisons for patient 1 and patient 2, respectively (**Figure 6A, B, Table S6, S7**). *BAALC* knockout in CN/AML-derived HSPCs led to a dramatic shift in the gene expression signature in both patients (**Figure 6C, D**). GSEA revealed enrichment of the oxidative phosphorylation, tricarboxylic acid cycle (TCA), p53 signaling and inhibition of platelet-specific genes in CN/AML1 samples, as compared to the CN/AML1 *BAALC* KO sample (**Figure 6E**). In CN/AML patient 2, E2F targets, G2M checkpoint-associated genes, TGF β signaling, MYC targets were enriched in CN/AML cells, as compared to *BAALC* KO cells (**Figure 6F**).

The TFEA using the list of differentially expressed genes ($\log_2FC > 1$ or < -1 , adj. P-value < 0.05) from CN/AML versus CN/AML *BAALC* KO comparisons for patient 1 and patient 2 revealed that *RUNX1*, *GATA1/2*, *SUZ12* transcription factor binding motifs are BAALC dependent. *AR*, *TCF3*, *RAD21*, *NANOG* motifs were specific for CN/AML patient 1 and *EGR1*, *STAT3* and *ZC3H11a* were CN/AML patient 2 specific (**Figure 6G**).

CSNK2A1, CDK1, GSK3B, HIPK2, MAPK14 were among the top significant kinases in CN/AML vs CN/AML *BAALC* KO comparisons in both patients (**Figure 6H**). Again, CN/AML patient 2, but not CN/AML patient 1 showed patient-specific enrichment of kinases, such as MAPK3/8, JNK1, ERK2 (**Figure 6H**).

Notably, the key transcription factors motifs such as *RUNX1*, *GATA1/2*, *SUZ12*, *SMAD4* as well as 7 out of 12 kinases (e.g. MAPK14, MAPK1, ERK1, AKT1) were enriched in CN/AML cells, compared to either CN or CN/AML *BAALC* KO stages in both patients

(**Figure 6I, J**). These data clearly indicate that *BAALC* KO at least partially reproduces CN phenotype in CN/AML cells.

A selective inhibitor of p38 α -mediated MK2a phosphorylation, CMPD-1, suppresses the growth of CN/AML cells

Using Connectivity Map (CMAP) analysis⁵⁸ of RNA-Seq data generated from the CN/AML versus CN/AML- *BAALC* KO comparisons, we evaluated the drug candidates that will induce gene expression signature similar to that shown in CN/AML cells after *BAALC* KO. A selective inhibitor of p38 α -mediated MK2a phosphorylation, CMPD-1, was the first hit with the effects similar to the *BAALC* KO seen in CN/AML cells (**Figure 7A**). We treated *BAALC*^{high} AML cell line Kasumi-1 with CMPD-1 and found markedly reduced cell proliferation, as compared to DMSO treated cells (**Figure 7B**). Moreover, CMPD1 treatment of CN/AML iPSCs derived HSPCs from both patients markedly reduced cell growth without any effect on the proliferation of HD- and CN- iPSC derived cells (**Figure 7C**), as compared to DMSO groups. Also, exposure of primary AML blasts of three *BAALC*^{high} CN/AML patients to CMPD1 led to the dramatic suppression of cell growth (**Figure 7D, E, Table S1**). The less prominent effect of CMPD1 was seen in *de novo* *BAALC*^{high} AML samples (**Figure 7F, G**). The proliferation of healthy donors' CD34⁺ cells was not affected by CMPD1 (**Figure 7E, G**).

It has been reported that MEK1/2 inhibitor U0126, in combination with KLF4 activation, reduced the growth of *BAALC*- expressed *de novo* AML cells⁵⁹. We compared the effects of CMPD1 with that of two selective MEK inhibitors, U0126 and AZD-6244, on the proliferation of CN/AML and *de novo* AML cells. U0126 had no effects on cell proliferation (**Figure S7A**), whereas AZD-6244 had a moderate effect (**Figure 7B, C, E, and G**). KLF-4 activator, in combination with AZD-6244 was either not effective in a dose of KLF4 10 nM, or was toxic at the 100 nM and 1000 nM KLF4 concentrations (data not shown).

Discussion

We present here an *in vitro* modeling of stepwise leukemia development in pre-leukemia bone marrow failure syndromes, exemplified by CN. By generating iPSCs from non-leukemic and leukemia cells of two CN patients, we were able to compare hematopoietic differentiation and gene expression differences between these two conditions. This led to the description of high *BAALC* expression, as an ultimate leukemia-causing event in CN/AML and the identification of CMPD1 as a potential drug for CN/AML and *BAALC*-expressing *de novo* AML patients.

Using CRISPR/Cas9 gene-editing, we introduced various gene alterations associated with pre-leukemia and leukemia conditions in iPSC lines carrying CN-causing inherited *ELANE* mutations. Thus, we created homo- and heterozygous CN-iPSC clones containing truncated G-CSF receptor that lacked three tyrosines, a finding frequently reported in CN patients^{22,26,60-62}. The acquisition of truncated *CSF3R* mutations results in a proliferative advantage of HSPCs and clonal selection of *CSF3R*-mutated cell clones upon G-CSF treatment, which predisposes to leukemia but does not induce it^{8,23,26,61-64}. In line with this, we observed elevated proliferation and reduced granulocytic differentiation of HSPCs generated from *CSF3R*-mutated CN-iPSC clones compared with CN-iPSCs expressing WT G-CSFR. Disturbed signal transduction downstream of a mutated *CSF3R*, in combination with inherited CN-causing mutations and life-long treatment with supra-physiological doses of G-CSF, induces genotoxic stress⁶⁵ and increases the susceptibility of HSPCs to the acquisition of additional leukemia-associated mutations, e.g. in *RUNX1*. Acquisition of *RUNX1* mutations is known to enhance the sensitivity of HSPCs to G-CSF⁶⁶.

Different types of *RUNX1* mutations were detected in CN/AML patients (missense, nonsense, frame-shift)¹⁶. In the present study, we provided evidence for dose-dependent outcomes of missense mutations versus *RUNX1* haploinsufficiency. We observed a correlation between the type of *RUNX1* mutations and trisomy 21: missense *RUNX1*

mutations, but not non-sense or frame-shift mutation in RHD, require trisomy 21 to amplify the leukemogenic effect of the mutated *RUNXI*. CRISPR/Cas9 gene-editing in iPSCs gave as a unique opportunity to modify endogenous *RUNXI* without overexpression of mutated *RUNXI* cDNA. We generated iPSC lines carrying *ELANE*, *CSF3R* and, additionally, mutations and chromosomal abnormalities mimicking CN/AML phenotype: (1) missense *RUNXI* mutation and trisomy 21 with two copies of mutated *RUNXI* (CN/AML patient 1), or (2) *RUNXI* haploinsufficiency without trisomy 21 (CN/AML patient 2). We observed almost completely abrogated myeloid differentiation and elevated proliferation of CN/AML-iPSCs compared with control HD-, or CN- iPSCs. The elevated levels of CN/AML-associated genes (e.g. *BAALC*, *STAT5*^{20,46,67,68}) in HSPCs derived from CN/AML-iPSCs as well as their increased proliferation and diminished differentiation argues for their leukemic features. Therefore, our *in vitro* model recapitulates step-wise leukemia development in CN. Given the absence of animal models of CN - transgenic mice carrying mutated *Elane* and *Hax1*-KO mice show no signs of neutropenia - *in vitro* modeling of CN and leukemia using CRISPR/Cas9 gene-editing of patient-derived iPSCs will expand our understanding of pre-leukemia bone marrow failure syndromes.

The impact of inherited CN-associated mutations (e.g., in *ELANE*) on leukemia progression is unclear. We have reported activation of UPR in *ELANE* mutated HSPCs⁶⁷. *ELANE* expression is regulated by *RUNXI*⁶⁸. It would be interesting to determine how missense *RUNXI* mutations and trisomy 21 affect UPR triggered by mutated *ELANE*. Induction of UPR by trisomy 21 in immortalized lymphocytes and fibroblasts of Down syndrome patients has been described⁶⁹. These patients often develop acute megakaryoblastic leukemia (AMKL), and the role of UPR in this process has not been studied yet.

We applied CRISPR/Cas9-mediated gene knockout, to evaluate the leukemogenic role of candidate genes found to be elevated in CN/AML cells. The most noticeable effect had *BAALC* KO, which demonstrated an inhibitory effect on cell proliferation and simultaneously

induced myeloid differentiation of CN/AML cells. *BAALC* is upregulated in *RUNXI*-mutated *de novo* AML^{70,71}, and high *BAALC* expression in AML patients is associated with poor prognosis⁷². The ultimate mechanism of *BAALC* increase by mutated *RUNXI* is unclear. We found that the transduction of healthy donor HSPCs with mutated *RUNXI* induced *BAALC* expression.

BAALC KO also improved granulocytic differentiation of CN iPSCs. Most probably, *BAALC* expression levels are already affected by the inherited *ELANE* mutations, and co-acquisition of *CSF3R* and *RUNXI* mutations further amplify these effects culminating in the extremely high *BAALC* levels in CN/AML blasts.

We found a dramatic difference in the gene expression signature between CN- and CN/AML-iPSCs derived HSPCs. We also identified signaling pathways deregulated by both missense and frame-shift *RUNXI* mutations, or affected explicitly by each mutation. The vast majority of crucial signaling leukemia-causing pathways were overlapped in HSPCs carrying missense or truncated *RUNXI* mutations. Thus, we found the activation of the *E2F* pathway, an enrichment of *RUNXI*, *GATA1/2*, *SUZ12*, *SOX2*, *SMAD4* targets and of the MAPK1/3/14, HIPK2, AKT1, ERK1 kinases in both CN/AML patients. Some transcription factors and kinases were regulated on the *RUNXI* mutation type-specific manner: missense p.R139G *RUNXI* regulated MYC targets and oxidative phosphorylation, while truncated *RUNXI* changed the constituents of ribosomes, G2M checkpoint genes, and TGFβ signaling genes. Some transcription factors and kinases were deregulated by truncated, but not missense *RUNXI*. These data might help to better understand the mechanistic outcomes of different types of the endogenously expressed mutated RUNX1 proteins in leukemogenesis.

BAALC KO led to a marked gene expression shift in CN/AML cells of both patients independent of the *RUNXI* mutation type. We detected oxidative phosphorylation, MYC-, TGFβ-, E2F- and G2M checkpoint- controlling pathways to be *BAALC*-dependent. Moreover, *RUNXI*, *SUZ12*, *SOX2*, *GATA1/2*, *SMAD4*, and *SALL4* transcription factors as well as

MAPK14, GSK3B, CSNK2A1, CDK1/2 kinases were found to be downstream of *BAALC* in both CN/AML patients. At the same time, *BAALC* KO does not affect TCF3-, TP63-, and STAT3- dependent signaling in CN/AML. These pathways are most probably not essential in the leukemic evolution in CN, since *BAALC* KO reverses the phenotype of CN/AML cells. These findings clearly demonstrate the fundamental role of *BAALC* in the leukemia development in CN and will contribute to a better understanding of the *BAALC* role in AML.

We found that the selective p38 MAPK (MAPK14) inhibitor, CMPD-1, effectively reduces the proliferation of iPSC-derived and primary CN/AML blasts without affecting the proliferation of healthy donor HSPCs. CMPD-1, or other p38 MAPK (MAPK14) inhibitors might be implemented in the treatment of CN/AML and *BAALC*^{high} AML. Morita et al. demonstrated that MEK inhibition in combination with KLF4 (Kruppel-like factor 4) induction reduces the proliferation of the *BAALC*-high AML cell line, Kasumi-1⁵⁹. In our experiments, KLF4-activator and MEK inhibitors were less effective, as CMPD-1.

In summary, our findings obtained using CN as an example, demonstrate that this *in vitro* model is a reliable system for studying step-wise leukemogenesis in pre-leukemic bone marrow failure syndromes. Using this model, we identified *BAALC* upregulation as a key leukemogenic event downstream of *RUNX1* and *CSF3R* mutations. Inhibition of *BAALC* in hematopoietic cells of CN patients or treatment with MAPK14 inhibitors might eliminate *RUNX1*-mutated AML blasts (**Figure 7H**).

Acknowledgments

We would like to thank the FACS core facility of the UKT, especially Kristin Bieber, Cornelia Grimmel and Stella Autenrieth for the assistance in the cell sorting and cATG – core facility for performing RNA-seq. This work was supported by the Excellence Initiative of the Faculty of Medicine, University of Tuebingen (J.S.), Jose Carreras Leukemia Foundation (J.S., B.D), Madeleine Schickedanz Kinderkrebsstiftung (J.S., M.K.), DFG (J.S., M.K.), internal funding program Fortüne of the Medical Faculty of the Tübingen University (M.K.), German-Israeli Foundation for Scientific Research and Development (GIF) (K.W., A.Z., B.D.), and Fritz Thyssen Foundation (B.D.).

Author contributions

B.D., M.K. and J.S. made initial observations, designed the experiments and analyzed the data; B.D. and M.K. performed the experiments; B.D. generated, characterized, cultured and differentiated human iPSCs, performed FACS, qRT-PCRs and WBs; M.K., M.N. and A.S. established transfection of iPSCs; M.K., A.S., and B.O. performed CRISPR/Cas9-mediated knockout of AML-related genes in CN/AML iPSCs; A.S. conducted CFU assay; M.K. and F. S. performed dPCR; J.M performed drug treatment of Kasumi-1 and AML samples. P.M. conducted LORD-Q for CD34⁺ from iPSCs. M.K. performed RNA-seq data analysis. A.Z. established iPSCs generation and EB-based hematopoietic differentiation in the lab. R.B. and A.Z. assisted with iPSC culture. C.L., T.M., and R.B. assisted with qRT-PCR and WB; T.R. and D.S. performed array-CGH of iPSC clones; C.Z. and K.W. provided patients material; L.K. assisted with the interpretation of data and provided insightful comments; K.W. and J.S. supervised and supported the study and wrote the manuscript (with assistance of B.D. and M.K.).

Material and Methods

iPSCs culture

iPSCs were maintained on mitomycin-C treated SNL-feeder cells (Public Health England) in iPSC-medium consisting of DMEM F12 (Sigma) supplemented with 20 Knockout Serum Replacement (Invitrogen), 30 ng/ml bFGF (PeproTech), 1 % non-essential amino acids solution (Invitrogen), 100 μ M 2-Mercapto-Ethanol and 2 mM L-Glutamine. iPSC-medium was replaced every day. For CRISPR/Cas9 gene-editing experiments and expansion of single cell derived clones, iPSC lines were cultured on Geltrex (Life Technologies).

CRISPR/Cas9 gene-editing of HD, CN, and CN/AML iPSC lines

Corresponding sgRNAs (**Table S2**) were cloned into all-in one pSpCas9(BB)-2A-GFP(PX458) plasmid which was a gift from Feng Zhang (Addgene plasmid # 48138). CN- or CN/AML iPSC lines were transfected with 1-5 μ g PX458-sgRNA construct using P3 Primary Cell 4D-Nucleofector™ X Kit L and 4D Nucleofector (Lonza) or reverse transfection with LT-Transfection Reagent (Mirus). GFP⁺ cells were sorted 48 hours post-transfection, cultured on the Geltrex and single-cell clones were analyzed by Sanger sequencing of the gene-edited target regions. *BAALC* KO in HD iPSCs was introduced by transfection with cytosine base editor plasmid pCMV-BE3 which was a gift from the David Liu lab (Addgene plasmid # 73021). Off-target sites were predicted using <http://crispor.org>⁷³. The top 3 sites with the highest off-target scores and/or exon localization were selected for Sanger sequencing.

EB-based hematopoietic differentiation of iPSCs

iPSCs were dissociated from SNL-feeders or matrigel- (Corning) coated plates using PBS/EDTA (0.02%) for 5 min. EB generation was done via centrifugation of 20,000 cells per EB in 96-well plates using APEL serum-free differentiation medium (Stemcell Technologies) supplemented with bFGF (20 ng/ μ l) and ROCK Inhibitor (R&D). Next day, BMP4 (40 ng/ml) was added to the culture induce mesodermal differentiation. On day four, EBs were plated on Matrigel-coated 6-well plates (10 EBs/well) in APEL medium supplemented with VEGF (40

ng/ml), SCF (50 ng/ml) and IL-3 (50 ng/ml). For neutrophilic differentiation, medium was changed 3 days later to fresh APEL medium supplemented with IL3 (50 ng/ml) and GCSF (50 ng/ml). First hematopoietic suspension cells appeared on day 12 – 14. Suspension cells were harvested every 3-4 days and analyzed starting from day 14 to day 32. All cytokines were purchased from R&D Systems, if nothing else is indicated.

Culture of iPSC-derived CD45⁺CD34⁺ cells, *de novo* CN/AML cells and AML cells on SL/SL feeder cells and drug treatment

1-3 x 10⁵ iPSC-derived CD45⁺CD34⁺ cells, primary CN/AML blasts, or *de novo* AML blasts were cultured on the SL/SL feeder cells producing FLT3L (kindly provided by C. Eaves, Vancouver, Canada) in HLTM/Myelocult H5100 medium (Stemcell Technologies) supplemented with 10⁻⁶M Hydrocortisone, IL-3 (20 ng/ml), Il-6 (20 ng/ml), TPO (20 ng/ml), SCF (50 ng/ml) and FLT3L (50 ng/ml) for 7 days with medium change every 3 - 4 days. For drug treatment, 1 μM AZD-6244, 1 μM CMPD1 or DMSO was added to culture medium and incubated for 7 days with medium change and addition of fresh drugs every 3 days. After 7 days, cell viability was assessed by Trypan blue staining and counting using Neubauer cell counting chambers.

References

1. Alter, B.P. Inherited bone marrow failure syndromes: considerations pre- and posttransplant. *Blood* **130**, 2257-2264 (2017).
2. Collins, J. & Dokal, I. Inherited bone marrow failure syndromes. *Hematology* **20**, 433-434 (2015).
3. Feurstein, S., Drazer, M.W. & Godley, L.A. Genetic predisposition to leukemia and other hematologic malignancies. *Semin Oncol* **43**, 598-608 (2016).
4. Gohring, G., *et al.* Chromosomal aberrations in congenital bone marrow failure disorders--an early indicator for leukemogenesis? *Ann Hematol* **86**, 733-739 (2007).
5. Savage, S.A. & Dufour, C. Classical inherited bone marrow failure syndromes with high risk for myelodysplastic syndrome and acute myelogenous leukemia. *Semin Hematol* **54**, 105-114 (2017).
6. Wegman-Ostrosky, T. & Savage, S.A. The genomics of inherited bone marrow failure: from mechanism to the clinic. *Br J Haematol* **177**, 526-542 (2017).
7. Skokowa, J., Germeshausen, M., Zeidler, C. & Welte, K. Severe congenital neutropenia: inheritance and pathophysiology. *Curr Opin Hematol* **14**, 22-28 (2007).
8. Skokowa, J., Dale, D.C., Touw, I.P., Zeidler, C. & Welte, K. Severe congenital neutropenias. *Nat Rev Dis Primers* **3**, 17032 (2017).
9. Quentin, S., *et al.* Myelodysplasia and leukemia of Fanconi anemia are associated with a specific pattern of genomic abnormalities that includes cryptic RUNX1/AML1 lesions. *Blood* **117**, e161-170 (2011).
10. Lensch, M.W., Rathbun, R.K., Olson, S.B., Jones, G.R. & Bagby, G.C., Jr. Selective pressure as an essential force in molecular evolution of myeloid leukemic clones: a view from the window of Fanconi anemia. *Leukemia* **13**, 1784-1789 (1999).
11. Dale, D.C., *et al.* Mutations in the gene encoding neutrophil elastase in congenital and cyclic neutropenia. *Blood* **96**, 2317-2322 (2000).
12. Welte, K. & Dale, D. Pathophysiology and treatment of severe chronic neutropenia. *Ann Hematol* **72**, 158-165 (1996).
13. Welte, K., Zeidler, C. & Dale, D.C. Severe congenital neutropenia. *Semin Hematol* **43**, 189-195 (2006).
14. Rosenberg, P.S., *et al.* Stable long-term risk of leukaemia in patients with severe congenital neutropenia maintained on G-CSF therapy. *Br J Haematol* **150**, 196-199 (2010).
15. Rosenberg, P.S., *et al.* The incidence of leukemia and mortality from sepsis in patients with severe congenital neutropenia receiving long-term G-CSF therapy. *Blood* **107**, 4628-4635 (2006).
16. Skokowa, J., *et al.* Cooperativity of RUNX1 and CSF3R mutations in severe congenital neutropenia: a unique pathway in myeloid leukemogenesis. *Blood* **123**, 2229-2237 (2014).
17. Makaryan, V., *et al.* The diversity of mutations and clinical outcomes for ELANE-associated neutropenia. *Curr Opin Hematol* **22**, 3-11 (2015).
18. Dong, F., *et al.* Mutations in the gene for the granulocyte colony-stimulating-factor receptor in patients with acute myeloid leukemia preceded by severe congenital neutropenia. *N Engl J Med* **333**, 487-493 (1995).
19. Dong, F., *et al.* Mutations in the granulocyte colony-stimulating factor receptor gene in patients with severe congenital neutropenia. *Leukemia* **11**, 120-125 (1997).

20. Gupta, K., *et al.* Bortezomib inhibits STAT5-dependent degradation of LEF-1, inducing granulocytic differentiation in congenital neutropenia CD34(+) cells. *Blood* **123**, 2550-2561 (2014).
21. Zeidler, C., Germeshausen, M., Klein, C. & Welte, K. Clinical implications of ELA2-, HAX1-, and G-CSF-receptor (CSF3R) mutations in severe congenital neutropenia. *Br J Haematol* **144**, 459-467 (2009).
22. Germeshausen, M., Skokowa, J., Ballmaier, M., Zeidler, C. & Welte, K. G-CSF receptor mutations in patients with congenital neutropenia. *Curr Opin Hematol* **15**, 332-337 (2008).
23. Germeshausen, M., Welte, K. & Ballmaier, M. In vivo expansion of cells expressing acquired CSF3R mutations in patients with severe congenital neutropenia. *Blood* **113**, 668-670 (2009).
24. Germeshausen, M., Ballmaier, M. & Welte, K. Incidence of CSF3R mutations in severe congenital neutropenia and relevance for leukemogenesis: Results of a long-term survey. *Blood* **109**, 93-99 (2007).
25. Tschan, C.A., Pilz, C., Zeidler, C., Welte, K. & Germeshausen, M. Time course of increasing numbers of mutations in the granulocyte colony-stimulating factor receptor gene in a patient with congenital neutropenia who developed leukemia. *Blood* **97**, 1882-1884 (2001).
26. Liu, F., *et al.* Csf3r mutations in mice confer a strong clonal HSC advantage via activation of Stat5. *J Clin Invest* **118**, 946-955 (2008).
27. Klimiankou, M., Mellor-Heineke, S., Zeidler, C., Welte, K. & Skokowa, J. Role of CSF3R mutations in the pathomechanism of congenital neutropenia and secondary acute myeloid leukemia. *Ann N Y Acad Sci* **1370**, 119-125 (2016).
28. Harada, Y. & Harada, H. Molecular mechanisms that produce secondary MDS/AML by RUNX1/AML1 point mutations. *J Cell Biochem* **112**, 425-432 (2011).
29. Osato, M., *et al.* Biallelic and heterozygous point mutations in the runt domain of the AML1/PEBP2alphaB gene associated with myeloblastic leukemias. *Blood* **93**, 1817-1824 (1999).
30. Christiansen, D.H., Andersen, M.K. & Pedersen-Bjergaard, J. Mutations of AML1 are common in therapy-related myelodysplasia following therapy with alkylating agents and are significantly associated with deletion or loss of chromosome arm 7q and with subsequent leukemic transformation. *Blood* **104**, 1474-1481 (2004).
31. Harada, H., Harada, Y., Tanaka, H., Kimura, A. & Inaba, T. Implications of somatic mutations in the AML1 gene in radiation-associated and therapy-related myelodysplastic syndrome/acute myeloid leukemia. *Blood* **101**, 673-680 (2003).
32. Schnittger, S., *et al.* RUNX1 mutations are frequent in de novo AML with noncomplex karyotype and confer an unfavorable prognosis. *Blood* **117**, 2348-2357 (2011).
33. Gaidzik, V.I., *et al.* RUNX1 mutations in acute myeloid leukemia: results from a comprehensive genetic and clinical analysis from the AML study group. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology* **29**, 1364-1372 (2011).
34. Preudhomme, C., *et al.* High incidence of biallelic point mutations in the Runt domain of the AML1/PEBP2 alpha B gene in Mo acute myeloid leukemia and in myeloid malignancies with acquired trisomy 21. *Blood* **96**, 2862-2869 (2000).
35. Taketani, T., *et al.* AML1/RUNX1 mutations are infrequent, but related to AML-M0, acquired trisomy 21, and leukemic transformation in pediatric hematologic malignancies. *Genes Chromosomes Cancer* **38**, 1-7 (2003).

36. Preudhomme, C., *et al.* High frequency of RUNX1 biallelic alteration in acute myeloid leukemia secondary to familial platelet disorder. *Blood* **113**, 5583-5587 (2009).
37. Michaud, J., *et al.* In vitro analyses of known and novel RUNX1/AML1 mutations in dominant familial platelet disorder with predisposition to acute myelogenous leukemia: implications for mechanisms of pathogenesis. *Blood* **99**, 1364-1372 (2002).
38. Beekman, R., *et al.* Sequential gain of mutations in severe congenital neutropenia progressing to acute myeloid leukemia. *Blood* **119**, 5071-5077 (2012).
39. Harada, Y. & Harada, H. Molecular pathways mediating MDS/AML with focus on AML1/RUNX1 point mutations. *J Cell Physiol* **220**, 16-20 (2009).
40. Imai, Y., *et al.* Mutations of the AML1 gene in myelodysplastic syndrome and their functional implications in leukemogenesis. *Blood* **96**, 3154-3160 (2000).
41. Nagata, T. & Werner, M.H. Functional mutagenesis of AML1/RUNX1 and PEBP2 beta/CBF beta define distinct, non-overlapping sites for DNA recognition and heterodimerization by the Runt domain. *J Mol Biol* **308**, 191-203 (2001).
42. Matheny, C.J., *et al.* Disease mutations in RUNX1 and RUNX2 create nonfunctional, dominant-negative, or hypomorphic alleles. *Embo j* **26**, 1163-1175 (2007).
43. Osato, M. Point mutations in the RUNX1/AML1 gene: another actor in RUNX leukemia. *Oncogene* **23**, 4284-4296 (2004).
44. Tahirov, T.H., *et al.* Crystallization and preliminary X-ray analyses of quaternary, ternary and binary protein-DNA complexes with involvement of AML1/Runx-1/CBFalpha Runt domain, CBFbeta and the C/EBPbeta bZip region. *Acta Crystallogr D Biol Crystallogr* **57**, 850-853 (2001).
45. Tahirov, T.H., *et al.* Structural analyses of DNA recognition by the AML1/Runx-1 Runt domain and its allosteric control by CBFbeta. *Cell* **104**, 755-767 (2001).
46. Nanua, S., *et al.* Activation of the unfolded protein response is associated with impaired granulopoiesis in transgenic mice expressing mutant Elane. *Blood* **117**, 3539-3547 (2011).
47. Morishima, T., *et al.* Genetic correction of HAX1 in induced pluripotent stem cells from a patient with severe congenital neutropenia improves defective granulopoiesis. *Haematologica* **99**, 19-27 (2014).
48. Nayak, R.C., *et al.* Pathogenesis of ELANE-mutant severe neutropenia revealed by induced pluripotent stem cells. *J Clin Invest* **125**, 3103-3116 (2015).
49. Hiramoto, T., *et al.* Wnt3a stimulates maturation of impaired neutrophils developed from severe congenital neutropenia patient-derived pluripotent stem cells. *Proc Natl Acad Sci U S A* **110**, 3023-3028 (2013).
50. Chao, M.P., *et al.* Human AML-iPSCs Reacquire Leukemic Properties after Differentiation and Model Clonal Variation of Disease. *Cell stem cell* **20**, 329-344.e327 (2017).
51. Kotini, A.G., *et al.* Stage-Specific Human Induced Pluripotent Stem Cells Map the Progression of Myeloid Transformation to Transplantable Leukemia. *Cell stem cell* **20**, 315-328.e317 (2017).
52. Liang, G. & Zhang, Y. Genetic and epigenetic variations in iPSCs: potential causes and implications for application. *Cell stem cell* **13**, 149-159 (2013).
53. Lachmann, N., *et al.* Large-scale hematopoietic differentiation of human induced pluripotent stem cells provides granulocytes or macrophages for cell replacement therapies. *Stem Cell Reports* **4**, 282-296 (2015).
54. Gupta, K., *et al.* Bortezomib inhibits STAT5-dependent degradation of LEF-1, inducing granulocytic differentiation in congenital neutropenia CD34+ cells. *Blood* **123**, 2550-2561 (2014).

55. Ran, F.A., *et al.* Genome engineering using the CRISPR-Cas9 system. *Nat Protoc* **8**, 2281-2308 (2013).
56. Love, M.I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* **15**, 550 (2014).
57. Clarke, D.J.B., *et al.* eXpression2Kinases (X2K) Web: linking expression signatures to upstream cell signaling networks. *Nucleic Acids Res* **46**, W171-W179 (2018).
58. Lamb, J., *et al.* The Connectivity Map: using gene-expression signatures to connect small molecules, genes, and disease. *Science (New York, N.Y.)* **313**, 1929-1935 (2006).
59. Morita, K., *et al.* BAALC potentiates oncogenic ERK pathway through interactions with MEKK1 and KLF4. *Leukemia* **29**, 2248-2256 (2015).
60. Tidow, N., *et al.* Clinical relevance of point mutations in the cytoplasmic domain of the granulocyte colony-stimulating factor receptor gene in patients with severe congenital neutropenia. *Blood* **89**, 2369-2375 (1997).
61. Ward, A.C., van Aesch, Y.M., Schelen, A.M. & Touw, I.P. Defective internalization and sustained activation of truncated granulocyte colony-stimulating factor receptor found in severe congenital neutropenia/acute myeloid leukemia. *Blood* **93**, 447-458 (1999).
62. Dong, F., *et al.* Distinct cytoplasmic regions of the human granulocyte colony-stimulating factor receptor involved in induction of proliferation and maturation. *Mol Cell Biol* **13**, 7774-7781 (1993).
63. Hermans, M.H., *et al.* Signaling mechanisms coupled to tyrosines in the granulocyte colony-stimulating factor receptor orchestrate G-CSF-induced expansion of myeloid progenitor cells. *Blood* **101**, 2584-2590 (2003).
64. Hermans, M.H., *et al.* Sustained receptor activation and hyperproliferation in response to granulocyte colony-stimulating factor (G-CSF) in mice with a severe congenital neutropenia/acute myeloid leukemia-derived mutation in the G-CSF receptor gene. *J Exp Med* **189**, 683-692 (1999).
65. Touw, I.P. Game of clones: the genomic evolution of severe congenital neutropenia. *Hematology Am Soc Hematol Educ Program* **2015**, 1-7 (2015).
66. Chin, D.W., *et al.* RUNX1 haploinsufficiency results in granulocyte colony-stimulating factor hypersensitivity. *Blood Cancer J* **6**, e379 (2016).
67. Nustede, R., *et al.* ELANE mutant-specific activation of different UPR pathways in congenital neutropenia. *Br J Haematol* **172**, 219-227 (2016).
68. Lausen, J., Liu, S., Fliegau, M., Lubbert, M. & Werner, M.H. ELA2 is regulated by hematopoietic transcription factors, but not repressed by AML1-ETO. *Oncogene* **25**, 1349-1357 (2006).
69. Aivazidis, S., *et al.* The burden of trisomy 21 disrupts the proteostasis network in Down syndrome. *PLoS One* **12**, e0176307 (2017).
70. Metzeler, K.H., *et al.* A stem cell-like gene expression signature associates with inferior outcomes and a distinct microRNA expression profile in adults with primary cytogenetically normal acute myeloid leukemia. *Leukemia* **27**, 2023-2031 (2013).
71. Mendler, J.H., *et al.* RUNX1 mutations are associated with poor outcome in younger and older patients with cytogenetically normal acute myeloid leukemia and with distinct gene and MicroRNA expression signatures. *J Clin Oncol* **30**, 3109-3118 (2012).
72. Weber, S., *et al.* BAALC expression: a suitable marker for prognostic risk stratification and detection of residual disease in cytogenetically normal acute myeloid leukemia. *Blood Cancer J* **4**, e173 (2014).
73. Haeussler, M., *et al.* Evaluation of off-target and on-target scoring algorithms and integration into the guide RNA selection tool CRISPOR. *Genome Biol* **17**, 148 (2016).

Figure Legends

Figure 1. Elevated allele frequency of missense *RUNXI* mutations in CN/AML with trisomy 21. Generation of stage-specific CN-, pre-leukemia, and leukemia iPSC lines.

(A) Allele frequency (AF) of wild type and corresponding mutant *RUNXI* allele in CN samples before (CN) and at the overt AML (CN/AML) measured by digital PCR (dPCR) is shown. Each bar represents a percentage of mutant *RUNXI* and wild type *RUNXI* AF quantified by dPCR. Data were analyzed using QuantStudio 3D Analysis Suite (Thermo Scientific, USA). In the CN phase the genotyping of BM or PB cells didn't reveal presence of mutant *RUNXI* allele. However, at the overt AML leukemia cells acquired *RUNXI* mutations and trisomy 21. Accordingly, percentage of mutant *RUNXI* allele changed to 59-68% indicating that the AML cells gained second mutant *RUNXI* allele.

(B) Establishment of the *in vitro* model of leukemia development in congenital neutropenia using generation of patient-specific iPSCs and CRISPR/Cas9 gene editing. Time scale of leukemic transformation of CN/AML patient and iPSCs generation for various time points during disease progression is presented. For CN/AML patient 1, CN1-iPS clone harbors *ELANE* p.C151Y mutation only (CN1), another two clones harbor additional *CSF3R* heterozygous p.R734fs or compound heterozygous p.Q741fs mutation leading to truncated G-CSF-receptor (CN+*CSF3R* p.R734fs and CN+*CSF3R* p.Q741Xfs), CN/AML iPSCs clones C1.1 and C1.2 both harbor *ELANE* p.C151Y, *CSF3R* p.Q741X and *RUNXI* p.R139G mutations and trisomy 21 (CN/AML1.1 and CN/AML1.2). For CN/AML patient 2, CN2 clones harbors *ELANE* p.G214R mutation only, CN2 *CSF3R* mut clone harbors additional *CSF3R* p.Q743X^{+/-} mutation and CN/AML2 clone harbors additional *CSF3R* p.Q743X^{+/-} mutation and *RUNXI* p.Glu175SerfsX7 mutation.

(C) *RUNXI* p.R139G and *RUNXI* wild type allele frequencies measured by dPCR in CN1-iPSCs, CN/AML1.2-iPSCs and CN/AML1.2-iPSC derived CD45⁺ cells. dPCR detected 100%

of wild type *RUNXI* allele in CN1-iPSCs, or 33 % and 66 % of wild type and mutant *RUNXI* allele in CN/AML1.2-iPSCs and CN/AML1.2-iPSC derived CD45⁺ cells, respectively.

Figure 2. Hematopoietic differentiation of the stage-specific iPSC lines derived from two CN/AML patients

(A, B) Flow cytometry analysis of suspension cells harvested from EBs culture on day 14 of differentiation for CN/AML patient 1 (A) and 2 (B). Data represent means \pm SD from two independent experiments. *P < 0.05, **P < 0.01.

(C, D) Proliferation rate of iPSC-derived CD34⁺CD45⁺ cells of CN/AML patient 1 (C) and CN/AML patient 2 (D) expanded on SL/SL feeder cells for 7 days. Proliferation of CN/AML cells is normalized to HD. Data represent means \pm SD from two independent experiments. *P < 0.05, **P < 0.01.

(E, F) Flow cytometry analysis of suspension cells harvested from EBs culture on day 32 of differentiation of CN/AML patient 1 (G) and 2 (H) iPSCs. Data represent means \pm SD from two independent experiments. *P < 0.05, **P < 0.01.

(G, H) Morphological analysis of suspension iPSC-derived hematopoietic cells harvested at day 32 of differentiation (Wright-Giemsa Stain) for CN/AML patient 1 (I) and 2 (J). Representative cytopsin slide pictures are shown.

(I, J) CFU-assay of CD45⁺CD34⁺ cells isolated from EBs culture on day 14. Total CFU counts are shown for CN/AML patient 1 (E) and 2 (F). Data represent means \pm SD from two independent experiments.

Figure 3. Up-regulation of the AML-associated genes in primary CN/AML blasts and CN/AML iPSC-derived CD34⁺ cells

(A) Expression of genes upregulated in AML with mutant *RUNXI* in BM MNCs of five CN/AML patients measured by qPT-PCR. Target gene mRNA expression is normalized to β -actin and shown relative to CN BM MNC or CD33⁺. Data represent means \pm SD from two independent experiments. *P < 0.05, **P < 0.01, ***P < 0.001.

(B, C) qRT-PCR analysis of genes upregulated in AML with mutant *RUNX1* in CD34⁺ cells generated from CN/AML-iPSCs of CN/AML patient 1 (B) and 2 (C). Target gene mRNA expression is normalized to β -actin and shown relative to CN CD34⁺. Data represent means \pm SD from two independent experiments. *P < 0.05, **P < 0.01, ***P < 0.001.

(D) Representative Western Blot images of BAALC protein expression in CD45⁺CD34⁺ cells generated from different iPSC clones of CN/AML patient 1 and 2. Numbers below Western Blot images indicate protein expression levels relative to HD and normalized to β -actin.

(E) *BAALC* mRNA expression in CD34⁺ cells of healthy individuals (n = 2) transduced with control GFP, wild type *RUNX1* or mutant *RUNX1* (p.R139G and p.R174L) lentiviral constructs was measured by qRT-PCR and normalized to β -actin. Data represent means \pm SD from two independent experiments. *P < 0.05.

(F, G) qRT-PCR analysis of CN-related hematopoietic factors in CD45⁺CD34⁺ cells generated from iPSCs clones of CN/AML patient 1 (F) and 2 (G) on day 14 of culture. Target gene mRNA expression is normalized to β -actin and shown relative to HD group. Data represent means \pm SD from two independent experiments. *P < 0.05, **P < 0.01, ***P < 0.001.

(H, I) Representative Western Blot images of CN-related hematopoietic proteins and β -actin in CD45⁺CD34⁺ cells generated from different iPSCs clones of CN/AML patient 1 (H) and 2 (I) on day 14 of culture. Numbers below Western Blot images indicate protein expression levels relative to HD and normalized to β -actin.

Figure 4. *BAALC* upregulation is essential for the leukemogenic transformation seen in CN/AML patients

(A) Workflow-scheme of the CRISPR/Cas9-mediated gene knockout performed in CN/AML1.2 iPSC line for the genes upregulated in AML with mutant *RUNX1*. EB-based hematopoietic differentiation was compared between gene-edited CN/AML- and HD-derived iPSCs.

(B) Proliferation rates of CD34⁺CD45⁺ cells derived from CN/AML1 and CN/AML1 *BAALC* KO iPSC lines cultured on SL/SL feeder cells for 7 days, normalized to HD. Data represent means ± SD from two independent experiments. **P < 0.01.

(C) Flow cytometry analysis of suspension cells harvested from EBs culture on day 32 of CN/AML1.2 KO iPSC differentiation. Data represent means ± SD from two independent experiments. *P < 0.05.

(D) Morphological analysis of hematopoietic cells differentiated from iPSCs at day 32 of CN/AML1.2 and CN/AML1.2 *BAALC* KO iPSC differentiation (Wright-Giemsa Stain). Representative images of cytopsin slides are shown.

(E) CFU-Assay of CD45⁺CD34⁺ cells generated from CN/AML1.2 KO iPSC lines. Total CFU counts are shown. Data represent means ± SD from two independent experiments. *P < 0.05, **P < 0.01.

(F) Cell proliferation rates of iPSC-derived CD34⁺CD45⁺ cells generated from CN/AML and CN/AML *BAALC* KO iPSC lines of patient 2 and cultured on SL/SL feeder cells for 7 days, normalized to HD. Data represent means ± SD from two independent experiments. **P < 0.01.

(G) Flow cytometry analysis of suspension cells harvested from EBs culture on day 32 of HD, CN/AML2, and CN/AML2 *BAALC* KO iPSC differentiation. Data represent means ± SD from two independent experiments. *P < 0.05.

(H) Morphological analysis of hematopoietic cells differentiated from iPSCs at day 32 of CN/AML and CN/AML2 *BAALC* KO iPSC differentiation (Wright-Giemsa Stain). Representative images of cytopsin slides are shown.

(I) CFU-Assay of CD45⁺CD34⁺ cells generated from HD, CN/AML2, and CN/AML2 *BAALC* KO iPSC lines. Total CFU counts are shown. Data represent means ± SD from two independent experiments. *P < 0.05, **P < 0.01, ***P < 0.001.

(J) Flow cytometry analysis of suspension cells harvested from EBs culture on day 32 of differentiation of HD, CN1, CN2.1 and corresponding *BAALC* KO clones. Data represent means \pm SD from two independent experiments. *P < 0.05, **P < 0.01.

Figure 5. Gene expression analysis of CD45⁺CD34⁺ HSPCs derived from CN/AML and CN iPSCs

(A) Supervised average linkage clustering of 702 genes differentially expressed ($\log_2\text{FC} > 1$ or < -1 , adj. P-value < 0.05) between HSPCs derived from CN/AML1 or CN1 iPSCs. RNA-seq samples were obtained from 3 independent EB-based hematopoietic differentiation experiments performed with CN1, CN/AML1.1, or CN/AML1.2 iPSC clones. Expression heatmap was created using variance stabilizing transformed read counts as an input.

(B) Supervised average linkage clustering of 1992 genes differentially expressed ($\log_2\text{FC} > 1$ or < -1 , adj. P-value < 0.05) HSPCs derived from CN/AML2 or CN2 iPSCs. RNA-seq samples were obtained from 3 independent EB-based hematopoietic differentiation experiments performed with CN2, or CN/AML2 iPSC clones. Gene expression heatmap was created using as an input variance stabilizing transformed read count data.

(C) Gene set enrichment analysis (GSEA) of datasets of HSPCs derived from CN/AML1 iPSCs compared to CN1 iPSC. Normalized read counts were analyzed using GSEA 4.0.3.

(D) Gene set enrichment analysis (GSEA) of datasets of HSPCs derived from CN/AML2 iPSCs compared to CN2 iPSC. Normalized read counts were analyzed using GSEA 4.0.3.

(E) Transcription factors predicted by the TFEA of the differentially expressed genes ($\log_2\text{FC} > 1$ or $\log_2\text{FC} < -1$, adjusted P-value < 0.05) from CN/AML1 versus CN1 and CN/AML2 versus CN2 group comparisons. Transcription factors (P-value < 0.05) which are common for both comparisons are plotted on the same graph. Data on the graphs are displayed as the negative \log_{10} transformed P-values.

(F) Kinase enrichment analysis (KEA) with differentially expressed genes ($\log_2\text{FC} > 1$ or < -1 , adjusted P-value < 0.05) from CN/AML1 versus CN1 and CN/AML2 versus CN2 group

comparisons. Kinases ($P\text{-value} < 10^{-8}$) which are common for both comparisons are plotted on the same graph. Data on the graphs are displayed as negative \log_{10} transformed P-values.

Figure 6. Gene expression analysis of CD45⁺CD34⁺ HSPCs derived from CN/AML iPSCs and CN/AML *BAALC* KO iPSCs

(A) Supervised average linkage clustering of 419 genes differentially expressed ($\log_2\text{FC} > 1$ or < -1 , adjusted P-value < 0.05) between CN/AML1 iPSC versus CN1 *BAALC* KO iPSC datasets. RNA-seq samples were obtained from 3 independent EB-based hematopoietic differentiation experiments performed with CN/AML1.2 and CN/AML1.2 *BAALC* KO iPSC clones. Gene expression heatmap was created using variance stabilizing transformed read counts as an input.

(B) Supervised average linkage clustering of 566 genes differentially expressed ($\log_2\text{FC} > 1$ or $\log_2\text{FC} < -1$, adjusted P-value < 0.05) between CN/AML2 iPSC versus CN2 *BAALC* KO iPSC datasets. RNA-seq samples were obtained from 3 independent EB-based hematopoietic differentiation experiments performed with CN/AML2 and CN/AML2 *BAALC* KO iPSC clones. Gene expression heatmap was created using variance stabilizing transformed read counts as an input.

(C) Venn diagram showing overlaps of genes differentially expressed ($\log_2\text{FC} > 1$ or < -1 , adjusted P-value < 0.05) between CN/AML1 versus CN/AML1 *BAALC* KO and CN/AML1 versus CN1 comparisons.

(D) Venn diagram showing overlaps of genes differentially expressed ($\log_2\text{FC} > 1$ or $\log_2\text{FC} < -1$, adjusted P-value < 0.05) between CN/AML2 versus CN/AML2 *BAALC* KO and CN/AML2 versus CN2 comparisons.

(E) GSEA of CN/AML1 iPSC compared to CN/AML1 *BAALC* KO iPSC datasets.

Normalized read counts were analyzed using GSEA 4.0.3.

(F) GSEA of CN/AML iPSC patient 2 versus CN/AML *BAALC* KO iPSC patient 2 datasets.

Normalized read counts were analyzed using GSEA 4.0.3.

(G) Transcription factors predicted by the TFEA of the differentially expressed genes ($\log_2FC > 1$ or $\log_2FC < -1$, adjusted P-value < 0.05) from CN/AML1 versus CN/AML1 *BAALC* KO and CN/AML2 versus CN/AML2 *BAALC* KO group comparisons. Transcription factors (P-value < 0.05) which are common for both comparisons are plotted on the same graph. Data on the graphs are displayed as negative \log_{10} transformed P-values.

(H) Results of KEA of differentially expressed genes ($\log_2FC > 1$ or < -1 , adjusted P-value < 0.05) from CN/AML1 versus CN/AML1 *BAALC* KO and CN/AML2 versus CN/AML2 *BAALC* KO group comparisons. Kinases (P-value $< 10^{-8}$) which are common for both comparisons are plotted on the same graph. Data on the graphs are displayed as negative \log_{10} transformed P-values.

(I) Venn diagram showing common transcription factors that overlap between CN/AML versus CN/AML *BAALC* KO and CN/AML versus CN groups.

(J) Venn diagram showing common kinases that overlap between CN/AML versus CN/AML *BAALC* KO and CN/AML versus CN groups.

Figure 7. Treatment with p38-MAPK inhibitor CMPD1 reduces cell proliferation in Kasumi-1, CN/AML-iPSC-derived CD34⁺CD45⁺ cells and *de novo* CN/AML or AML blasts

(A) Connectivity MAP analysis of RNA-seq data of CN/AML and CN/AML *BAALC* KO iPSC-derived CD45⁺CD34⁺ cells. **pc_selection** is the percent of total perturbagens, querying the column sample against selected rows, that exceed the given thresholds.

(B) Treatment of Kasumi-1 WT cells with 1, 2 and 5 μ M AZD-6244 or 1, 2 and 5 μ M CMPD1 for 7 days. Data represent means \pm SD from two independent experiments. *P < 0.05 , ****P < 0.0001 .

(C) Treatment of CN/AML-iPSC derived CD34⁺CD45⁺ cells with 1 μ M AZD-6244 or 1 μ M CMPD1 for 7 days. Data represent means \pm SD from two independent experiments.

(D) qRT-PCR analysis of *BAALC* mRNA expression in patient CN/AML blasts. mRNA expression is shown relative to HD CD34⁺ cells and normalized to β -actin. Data represent means \pm SD from two independent experiments **P < 0.01.

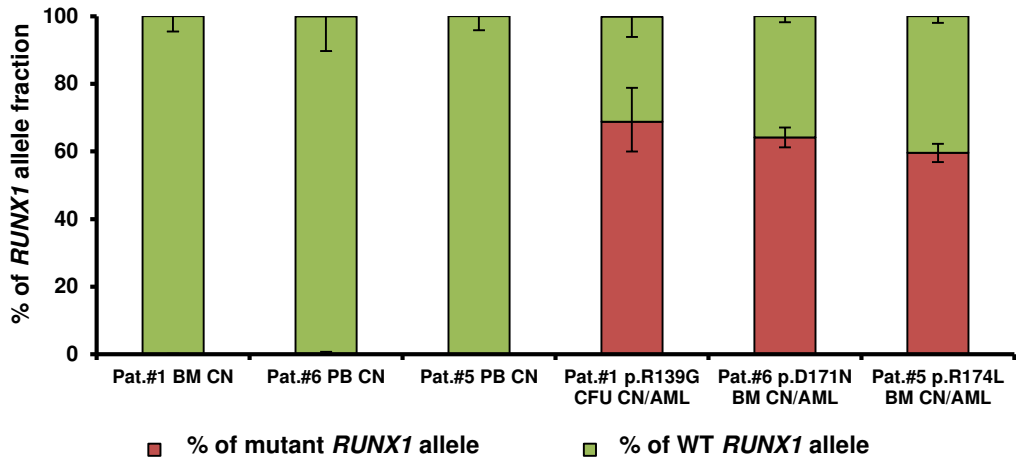
(E) Treatment of primary CN/AML-cells with 1 μ M AZD-6244 or 1 μ M CMPD1 for 7 days. Data represent means \pm SD from two independent experiments. **P < 0.01, ***P < 0.001.

(F) qRT-PCR analysis of *BAALC* mRNA expression in de novo AML blasts. mRNA expression is shown relative to HD CD34⁺ cells and normalized to β -actin. Data represent means \pm SD from two independent experiments. ***P < 0.001.

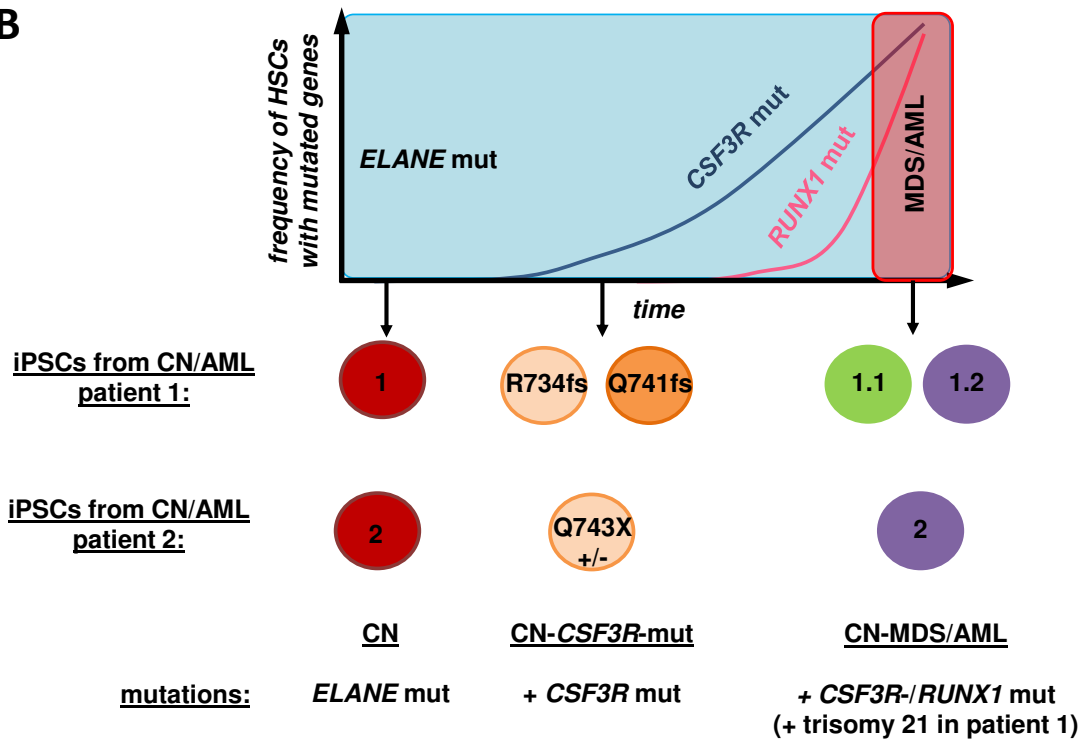
(G) Treatment of *de novo* AML blasts with 1 μ M and 2 μ M AZD-6244 or 1 μ M and 2 μ M CMPD1 for 7 days. Data represent means \pm SD from two independent experiments. **P < 0.01, ***P < 0.001.

(H) Schematic of the leukemia development in CN patients and the role of *BAALC* upregulation in this process. Prolonged exposure of the genetically “unfit” HSPCs of CN patients with inherited CN-associated mutations (e.g., *ELANE*) to the high dose of G-CSF led to the clonal selection of HSPCs (pre-leukemia HSPCs) with acquired *CSF3R* mutations that generate truncated G-CSF receptors. These cells exhibit elevated proliferation and reduced differentiation upon G-CSF treatment. Co-acquisition of missense *RUNX1* mutations in the Runt domain and trisomy 21 or *RUNX1* haploinsufficiency lead to elevated *BAALC* expression and leukemogenic transformation. *BAALC* KO in CN/AML cells restores normal hematopoietic differentiation. A selective p38-MAPK inhibitor, CMPD1, efficiently blocks the proliferation of CN/AML cells.

A



B



C

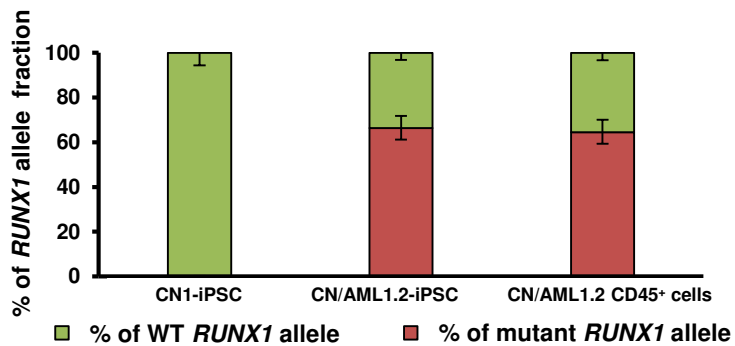


Figure 1

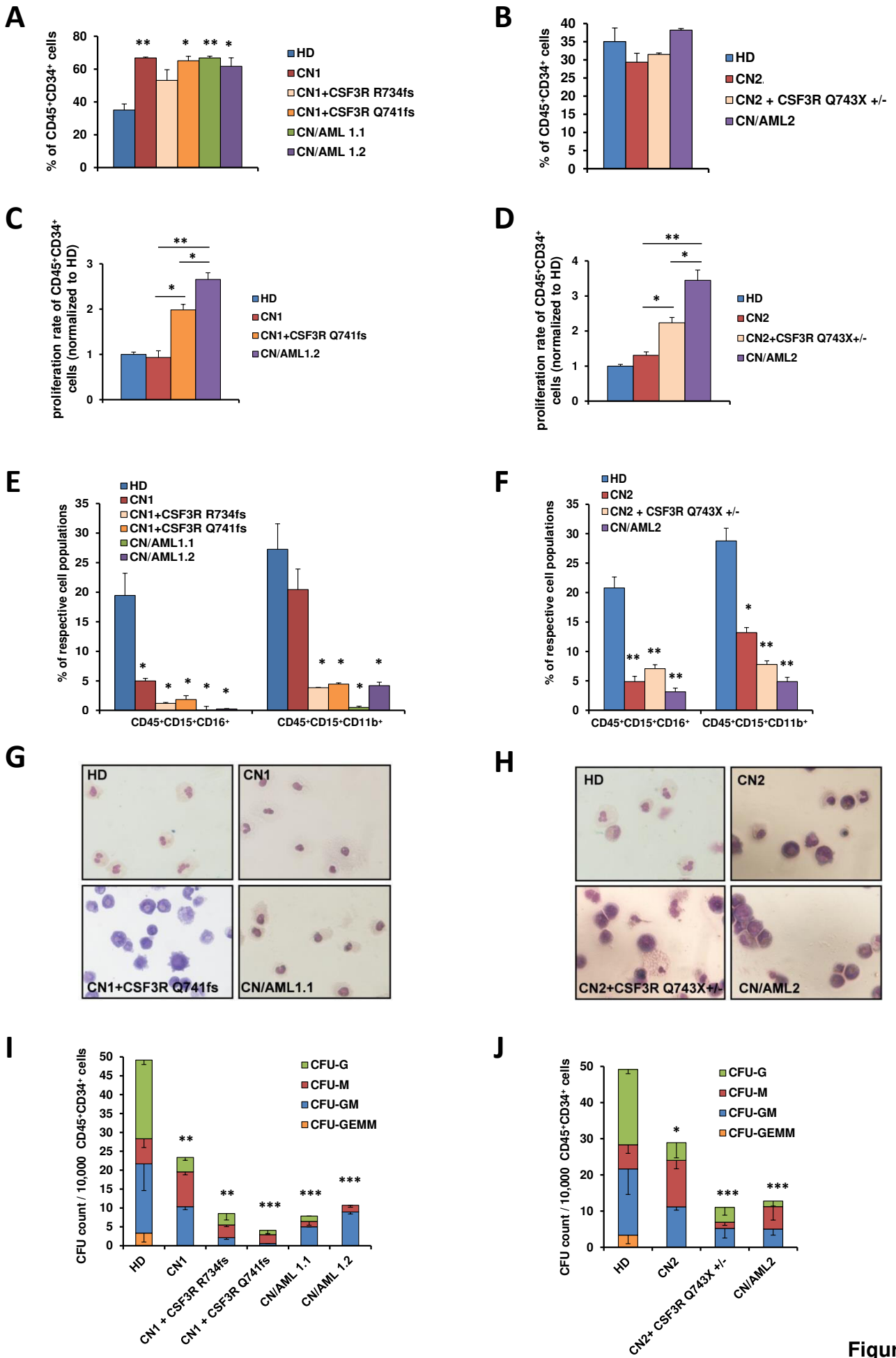
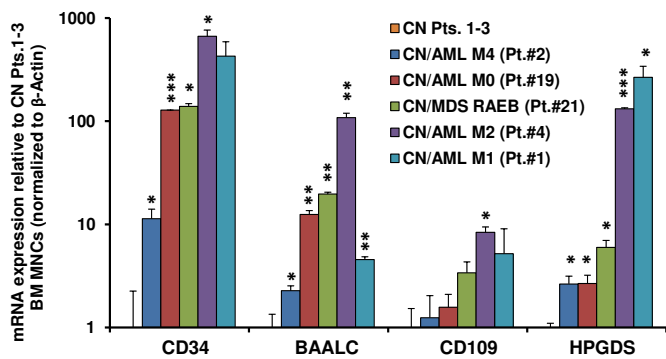
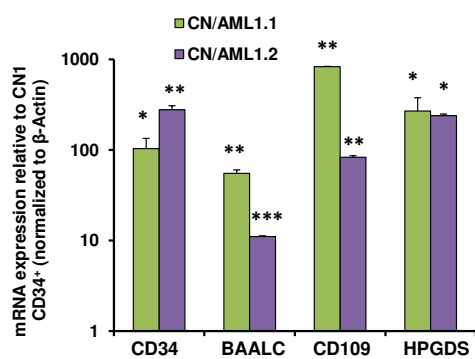
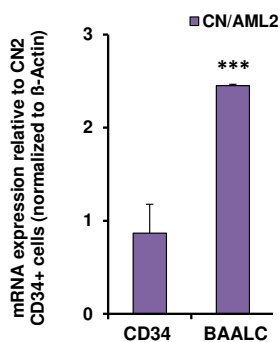
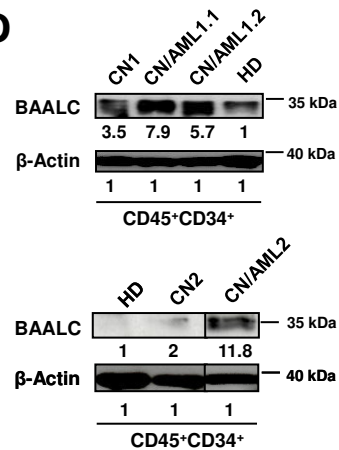
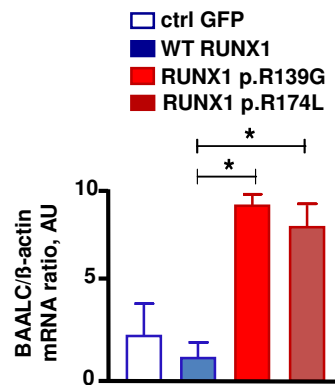
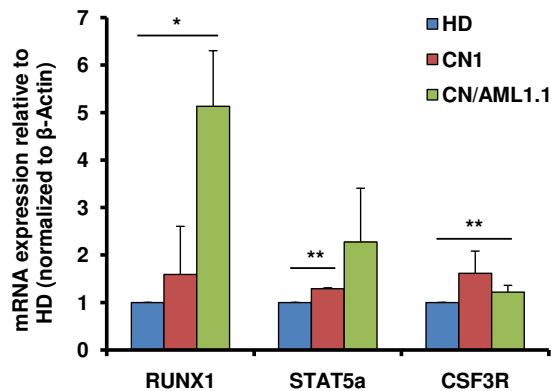
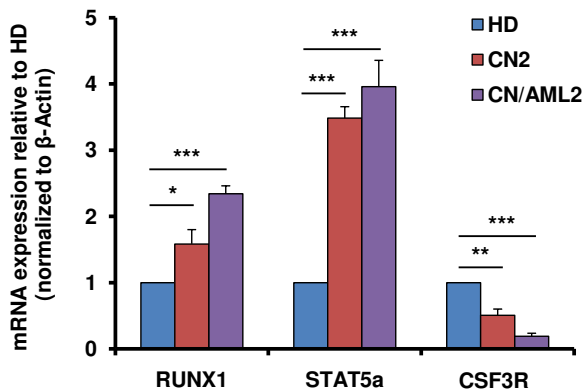
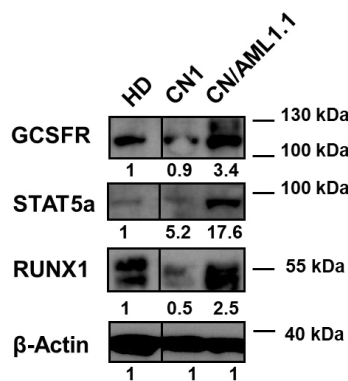
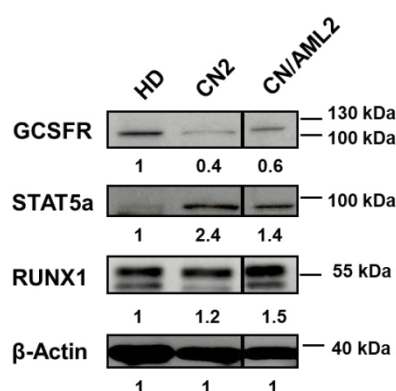


Figure 2

A**B****C****D****E****F****G****H****I****Figure 3**

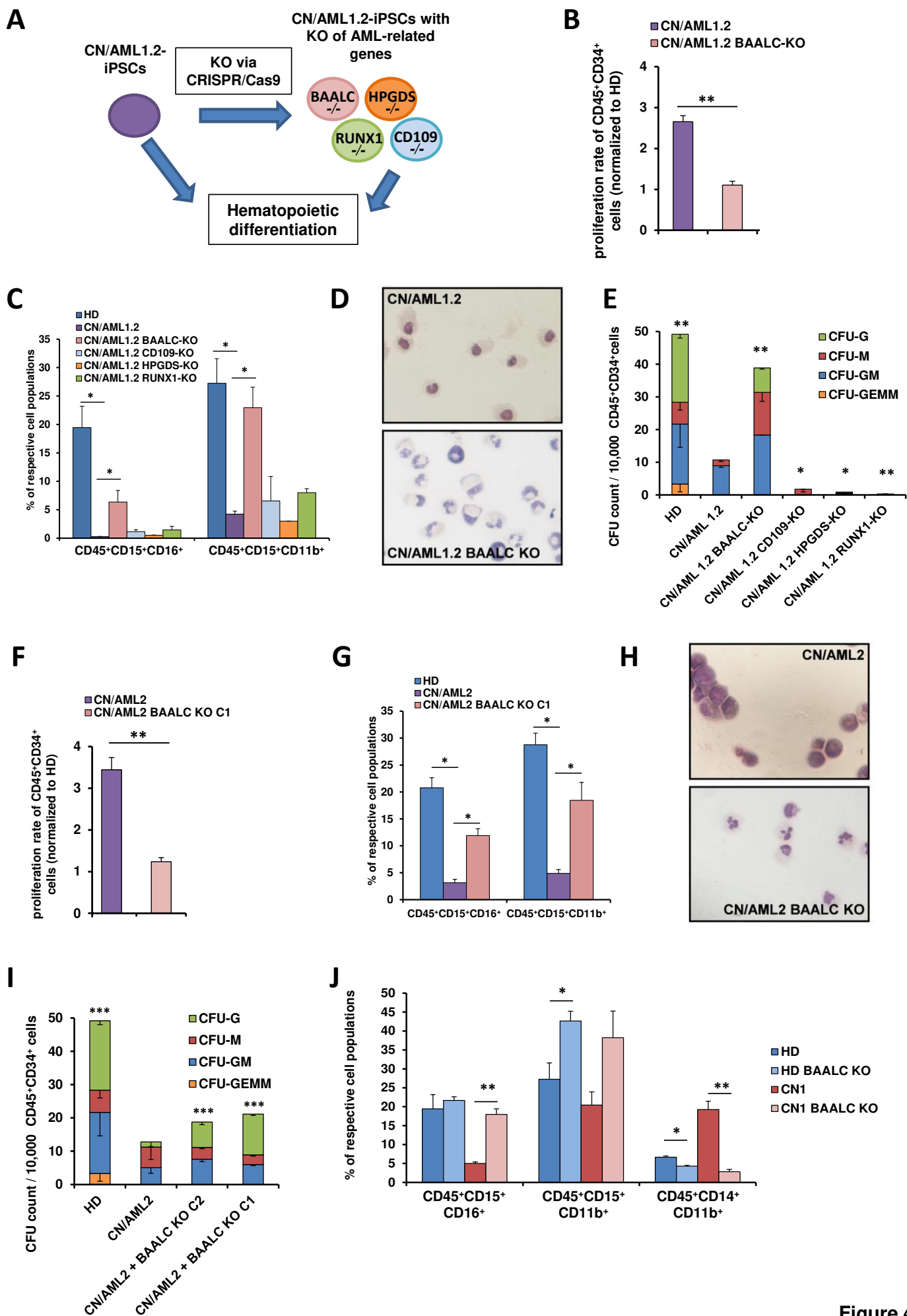


Figure 4

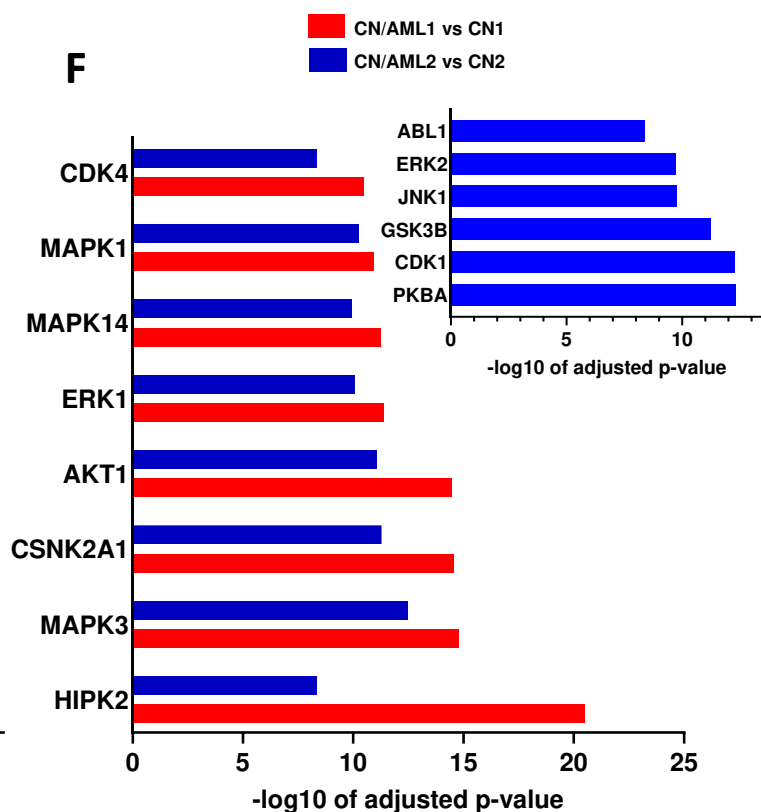
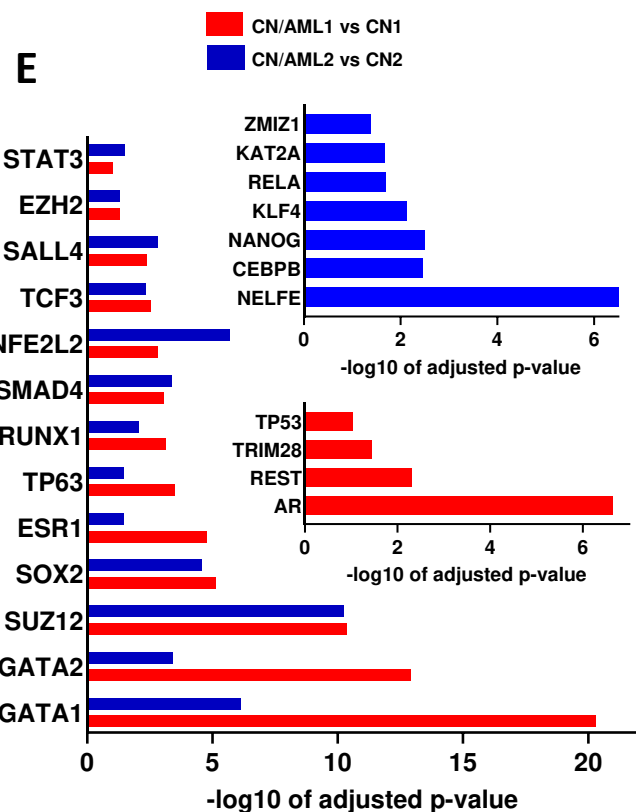
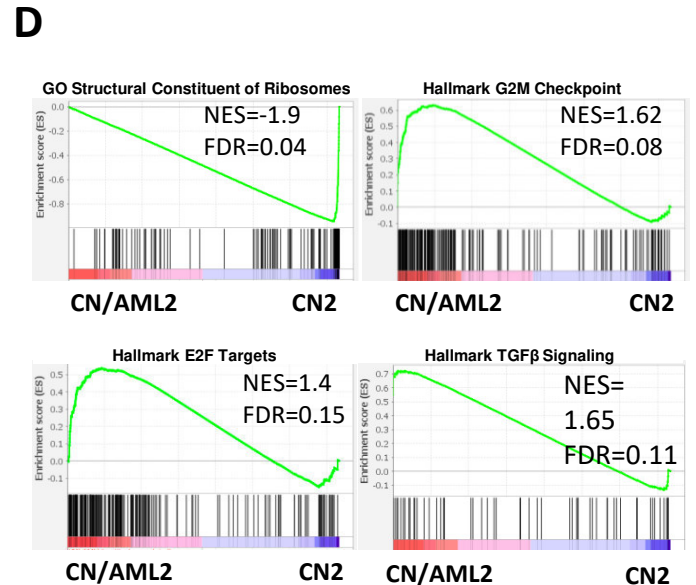
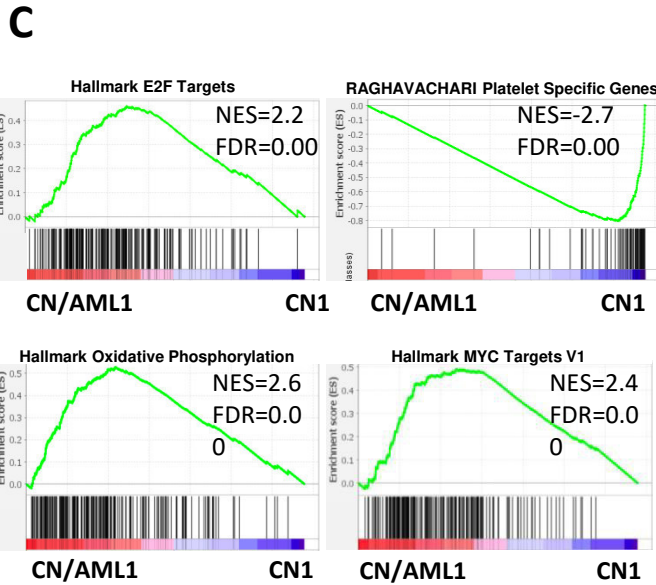
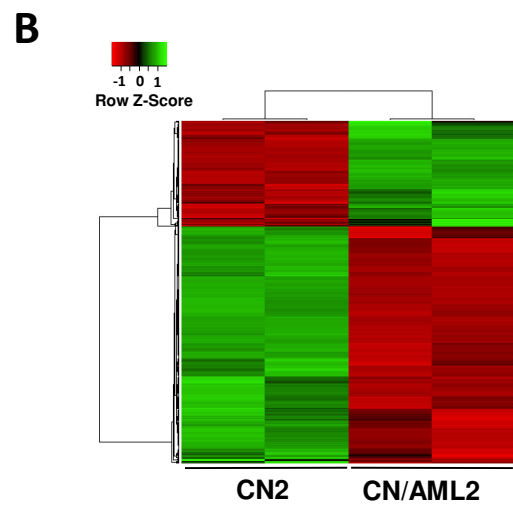
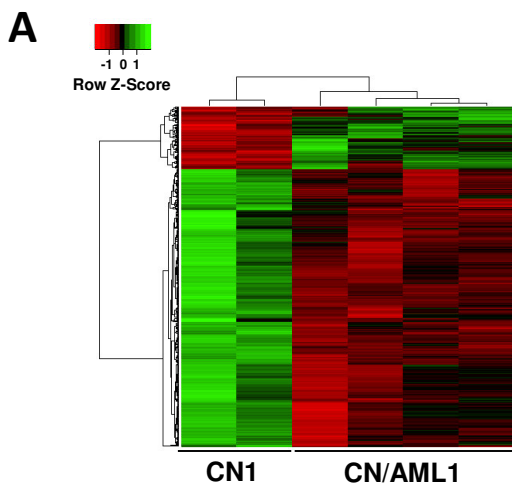


Figure 5

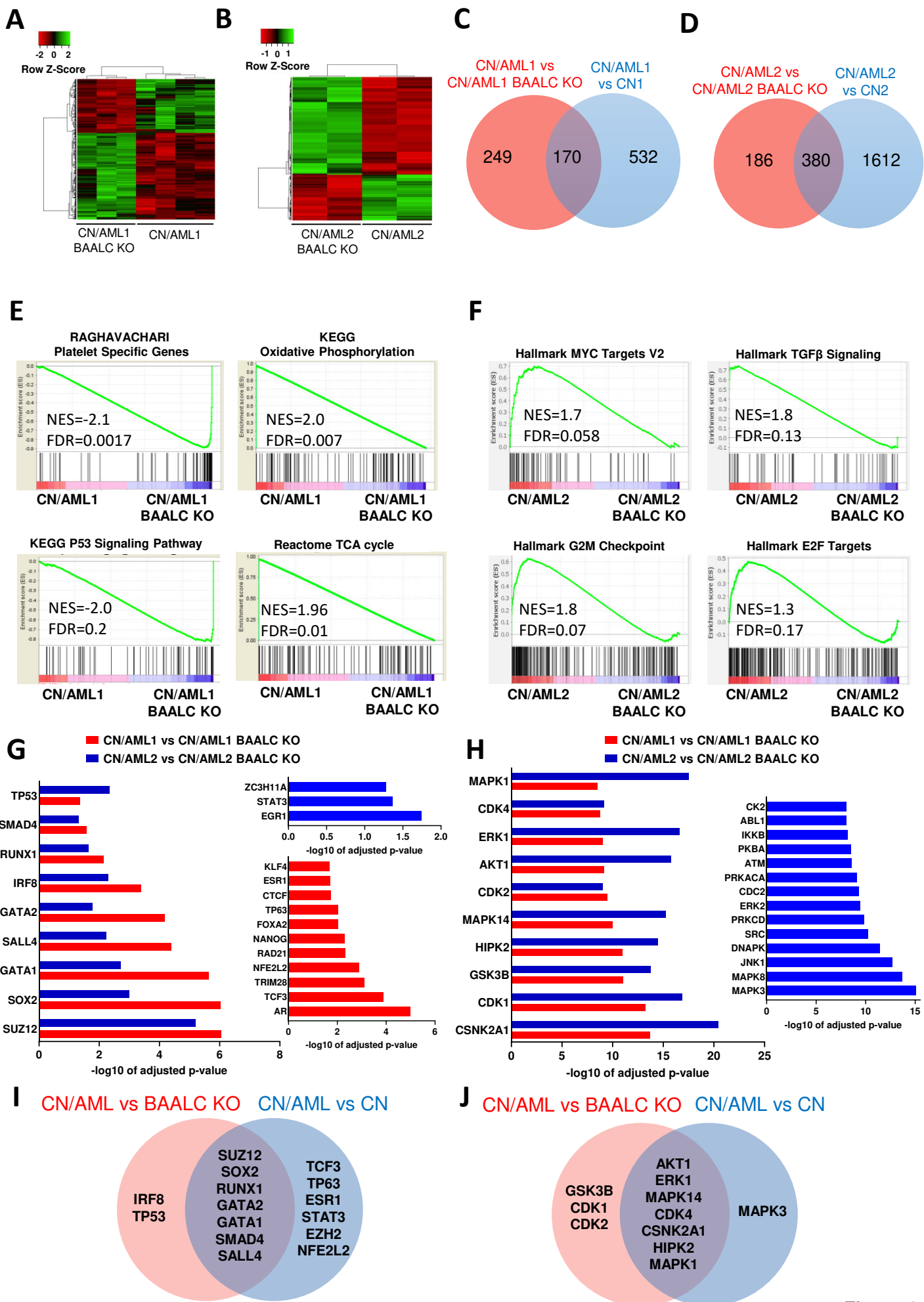
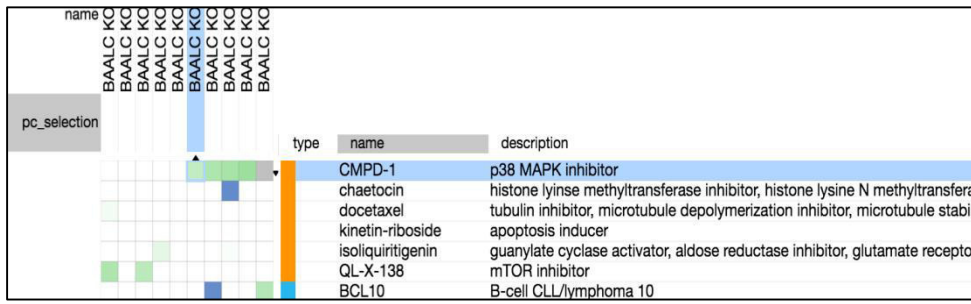
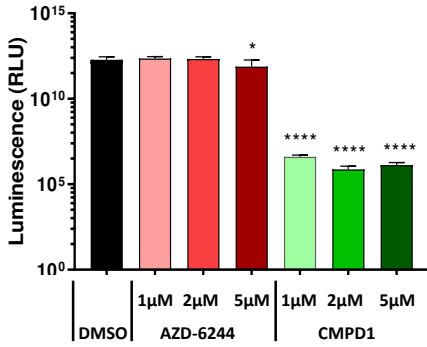
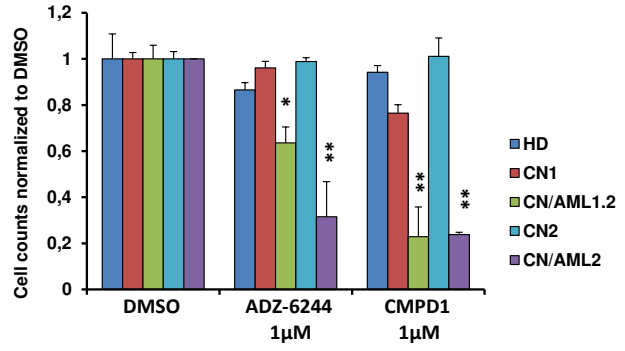
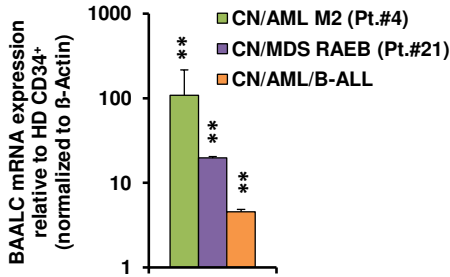
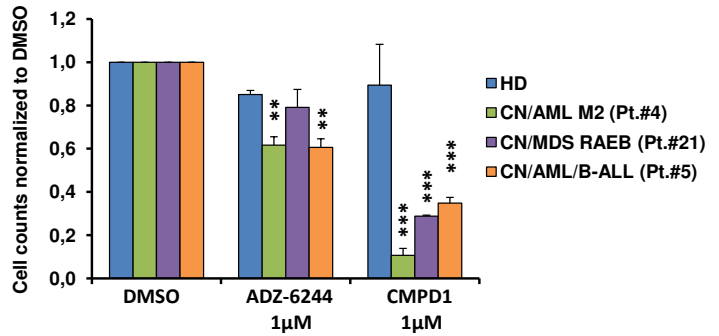
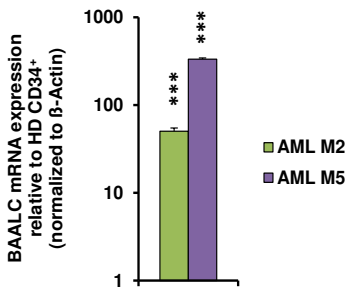
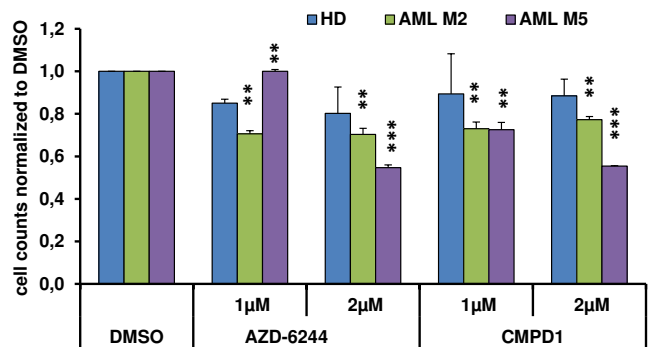
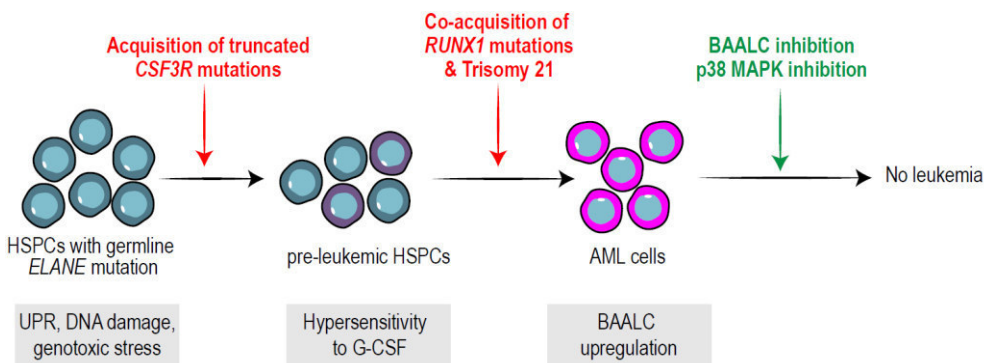


Figure 6

A**B****C****D****E****F****G****H****Figure 7**

Supplemental Information

Supplementary Methods

Reprogramming of PBMNCs

1.5x10⁶ PBMNCs were cultured after thawing for 6 days in CD34⁺ cells expansion medium (StemlineII medium, Sigma-Aldrich) supplemented with 10 % FCS, 1 % Pen/Strep, 1 % Glutamine and cytokines: IL-3 (20 ng/ml), Il-6 (20 ng/ml), TPO (20 ng/ml), SCF (50 ng/ml) and FLT3L (50 ng/ml). All cytokines were purchased from R&D Systems. After 1 week, cells were added to Retronectin (Clontech)-coated 12-well plates together with OSKM lentiviral supernatant (pRRL.PPT.SF.hOct34.hKlf4.hSox2.i2dTomato.pre.FRT, provided by A. Schambach) with multiplicity of infection (MOI) of 2. Four days later, cells were transferred to SNL-feeders and cultured in the mixture of ½ iPSCs-medium and ½ CD34⁺ cells expansion medium supplemented with 2 mM valproic acid and 50 µg/ml Vitamin C. Medium was gradually changed to iPSCs medium only. First iPSCs colonies appeared approximately three weeks after initiation of reprogramming.

Sequencing of iPSCs

Genomic DNA of iPSCs was isolated using Nucleo-Spin Tissue Kit (Machery-Nagel) and DNA regions for sequencing were amplified using the following primers: *RUNX1-F* 5'-ACATCCCTGATGTCTGCATTTGTCC-3', *RUNX1-R* 5'-TGTGGGTTTGTGGCCATGAAACGTG-3', *ELANE-F* 5'-CGCCCTGAGCCTTGGTGACG-3', *ELANE-R* 5'-AGCCACGGTGCCTGTTGCTG-3', *CSF3R-F* 5'-ATGGCATGTGTCAGGCATGT-3', *CSF3R-R* 5'-AGTCACAGCGGAGATAGTGC-3'. Sanger Sequencing was performed by GATC Biotech.

Transduction of CD34⁺ cells with RUNX1 constructs

We transduced CD34⁺ cells from healthy donors (2 × 10⁵/well) with lentiviral supernatants containing control plasmid, or cDNA with WT RUNX1 or two RUNX1 mutants (p.R139G and p.R174L) expressed in pRRL.PPT.SFFV.i2RFP vector with a multiplicity of infection

(MOI) of 5. Re-transduction was performed next day. After 72 hours post-transduction, GFP-positive cells were sorted and analyzed for RUNX1 and BAALC mRNA expression by qRT-PCR. Vector and primer information is available upon request.

Quantitative RT-PCR

RNA was isolated using RNeasy Micro Kit (Qiagen). cDNA was prepared from 0.2-1 µg of total RNA using Omniscript RT Kit (Qiagen). qPCR was performed using SYBR Green qPCR master mix (Roche) on Light Cycler 480 (Roche). Data were analyzed using ddCT-method. Target genes were normalized to *GAPDH* and/or β -Actin as housekeeper genes. qPCR Primer sequences are available upon request.

Western blotting

1×10^6 cells were lysed in 200 µl of 3x Lämmli buffer. Protein was denaturated for 10 min at 95 °C. 5 µl of cell lysate in Lämmli buffer were loaded per lane. Proteins were separated on a 12% polyacrylamide gel and transferred on a nitrocellulose membrane (GE Healthcare) (1 hour, 100V, 4°C). Membrane was blocked for 1 hour in 5 % BSA/TBST and incubated in primary antibody overnight (at 4°C). The following primary antibody were used: anti-RUNX1/AML1 (Cell Signaling, #4334, 1:500), anti-STAT5a (Cell Signaling, #4807, 1:500) anti-G-CSFR (Santa Cruz, #74026, 1:500), and β -Actin (Cell Signaling, #13E5, 1:1000). After that, membranes were washed and incubated with secondary HRP-coupled antibody (Cell Signaling, #7076 or #7074, 1:2000) for 1 hour at room temperature. Pierce ECL solution (Thermo Fisher) and Amersham Hyperfilm were used to detect chemiluminescence signal of proteins.

Alkaline Phosphatase Assay

iPSCs colonies on SNL-feeders at day 10 of culture were washed with PBS, fixed in 4 % PFA /10 % Sucrose in water and stained with NBT/BCIP staining dye (Sigma) for 20 min at RT.

Array-CGH

Array-CGH was performed using the Agilent Human Genome Microarray Kits 2x400K (Agilent Technologies, Santa Clara, CA, USA). Labelling and hybridization of genomic DNA was performed according to the protocol provided by Agilent. Microarray slides were scanned using an Agilent microarray scanner G2505B at a resolution of 2 μ m. For image analysis, default CGH settings of Feature Extraction Software (Agilent Technologies) were applied. Output files from Feature Extraction were subsequently imported into Agilent's CGH data analysis software, Genomic-Workbench. The Aberration Algorithm ADM2 was applied and Aberration Filters were set to: threshold 7.0, at least 4 probes with mean log₂ ratio of +/-0.3 leading to a resolution of approx. 20 kb.

Flow Cytometry

30.000 of suspension cells collected from EB-based hematopoietic differentiation system were used for flow cytometry. For cell surface staining, cells were prepared in PBS/1% BSA containing 0.05 % sodium azide and stained with the mouse monoclonal anti-human antibody. For detection of hematopoietic progenitor cells, a multicolor FACS antibody panel for 'early-stage' hematopoietic differentiation using the following antibody was applied: CD33-BV421 (BioLegend, BL), CD34-PeCy7, KDR-AF647 (BL), CD43-PE, CD41a-FITC, CD235a-FITC, CD45-BV510 (BL), 7-AAD. For detection of mature myeloid cells, a multicolor FACS antibody panel for 'late-stage' hematopoietic differentiation using the following antibody was applied: CD15-PE, CD16-FITC, CD14-APC-H7, CD45-BV510 (BL), CD33, BV-421 (BL), 7-AAD. For iPSCs characterization, the stem cell surface markers TRA1-60-PE (eBioscience) and SSEA4-FITC were analyzed. TRA1-85-APC (R&D) was used as a human iPSCs marker. Anti-mouse IgGk beads (BD Biosciences) were used for compensation. Samples were analyzed using FACSCanto II (BD) and FlowJo V10 (FlowJo LLC). Antibodies for flow cytometry were purchased from BD Biosciences, if nothing else is indicated.

Morphological analysis

Wright-Giemsa-stained cytopsin slides were prepared using Hema-Tek slide stainer (Ames). Hematopoietic cells were ordered into 4 groups according to the differentiation state: myeloblast and promyelocyte (MB/ProM), myelocyte and metamyelocyte (Myelo/Meta), band and segmented neutrophils (Band/Seg) and monocytes/macrophages (Mo/MΦ).

Colony Forming Unit (CFU) Assay

10.000 suspension cells from EB-based iPSC hematopoietic differentiation at day 14 were used for CFU-Assay using Methocult H4435 enriched medium (Stemcell Technologies). Colonies were counted after 10-14 days.

Digital PCR (dPCR)

Digital PCR was used for absolute endpoint quantification of gene copy numbers using TaqMan SNP genotyping Assay and Quant Studio 3D Digital PCR System (Thermo Fisher). dPCR was performed according to QuantStudio 3D Digital PCR protocol using genomic DNA. Data were processed using QuantStudio 3D Digital PCR Analysis Suite Software (Thermo Fisher).

RNA sequencing (RNA-seq)

mRNA was isolated from CD34⁺CD45⁺ cells differentiated from iPSCs using RNeasy Mini- or Micro Kit (Qiagen). RNA quality was assessed using an Agilent 2100 Bioanalyzer. Samples with high RNA integrity number (RIN > 8) were selected for library construction. Using the TruSeq RNA Sample Prep Kit (Illumina) and 100 ng of total RNA for each sequencing library, poly (A) selected single-read sequencing libraries (75 bp read length) were generated according to the manufacturer's instructions. All libraries were sequenced on an Illumina NextSeq500 platform at a depth of 20–30 million reads each. Read quality of RNA-seq data in fastq files was assessed using FastQC (v0.11.4). Reads were aligned using STAR (v2.4.2a) allowing gapped alignments to account for splicing against a custom-built genome composed of the Ensembl Homo Sapiens genome v90 and Alignment quality was analyzed using samtools (v1.1) and visually inspected in the Integrative Genome Viewer

(v2.3.67). Transcripts covered with less than 50 reads were excluded from the analysis leaving >12.000 genes for determining differential expression between the experimental groups. Normalized read counts for the genes were obtained in DESeq2 package¹ (v1.8.2) and used for GSEA to find significant differences between groups². Differential expression analysis was performed in DESeq2 package. Differentially expressed genes with $\log_2FC > 1$ or $\log_2FC < -1$, adjusted P-value < 0.05 were selected for downstream analysis. Heat maps were produced in Heatmapper³. Venn diagrams were made in Venny 2.1 (<https://bioinfogp.cnb.csic.es/tools/venny/index.html>). For identification upstream transcription factors that regulate differently expressed genes and kinase enrichment assay, we utilized eXpression2Kinases algorithm⁴.

Connectivity map (CMAP) analysis

CLUE platform (<https://clue.io/cmap>) was used for CMAP analysis to evaluate the gene expression profiles of CN/AML and CN/AML BAALC KO RNA-seq data sets for connectivity to known perturbagens⁵.

Supplementary Figure Legends

Figure S1. Representative images of digital PCR data plots for quantification of *RUNXI* allele ratio in CN and CN/AML samples.

(A) The signals from the FAM reporter dye are plotted on the Y-axis against the signal from the VIC reporter dye on the X-axis. Blue dots represent FAM-labeled mutant *RUNXI* PCR product, red dots represent VIC-labeled wild type *RUNXI* PCR product, green dots represent droplets containing both mutant and wild type DNA, and yellow dots are droplets with no DNA incorporated. Ratio of mutant and wild type *RUNXI* allele was calculated based on number of FAM and VIC positive signals on the chips. For example, DNA isolated from BM sample of Pat.#1 in CN-phase and DNA extracted from single cell derived CFU clone

(CN/AML phase) was amplified using TaqMan assay probe set against wild type *RUNXI* and *RUNXI* p.R139G allele. The presence of *RUNXI* p.R139G mutation was detected in 360 droplets representing an allelic fraction of 66.8 %. Correspondingly, the ratio of mutant to wild type allele is 2.04.

Figure S2. Genetic analysis of iPSCs clones derived from CN/AML patients.

(A) Representative Sanger Sequencing diagrams of DNA isolated from the indicated iPSCs clones demonstrating *ELANE* p.C151R, *CSF3R* p.Q741X and *RUNXI* p.R139G heterozygous mutations for CN/AML patient 1. *ELANE* p.G214R and *CSF3R* p.Q743X heterozygous mutations for CN/AML patient 2 are also shown.

(B) Representative images of array CGH of the indicated iPSCs clones. Whole genome view is shown.

Figure S3. Characterization of HD-, CN- and CN/AML-iPSCs clones

(A, B) qRT-PCR analysis of mRNA expression of pluripotency-specific genes using mRNA isolated from indicated iPSCs clones of CN/AML patient 1 (A) and 2 (B). Target gene mRNA expression is normalized to *GAPDH* and shown relative to CD34⁺ cells. PRE indicates exogenous mRNA – expression from lentiviral vector. Data represent means \pm SD from two independent experiments.

(C) Flow cytometry analysis of pluripotency-specific surface markers Tra1-60 and SSEA4 in indicated iPSCs clones of CN/AML patient 1 and 2.

(D, E) Representative images of the Alkaline Phosphatase activity analysis in iPSC colonies of CN/AML patient 1 (D) and 2 (E), indicated by purple staining.

Figure S4. Assessment of CRISPR/Cas9 gene-edited iPSC lines

(A-J) DNA isolated from single-cell derived iPSC clones was amplified by PCR and analyzed using TIDE (Tracking of Indels by DEcomposition). Graphs show the estimated composition of the inserted/deleted (in/del) nucleotide bases in iPSC lines calculated by TIDE. Additionally, genetic changes introduced by CRISPR/Cas9 have been annotated on protein

level in order to predict the consequence of in/del introduced by genome editing on the amino acid sequences of modified genes.

Figure S5. Analysis of hematopoietic differentiation of CN and CN/AML iPSC-derived HSCs by flow cytometry

(A) Schematic of the protocol for EB-based hematopoietic/neutrophilic differentiation of iPSCs.

(B, C) Flow cytometry analysis of suspension cells harvested from EBs culture on day 14 of differentiation for CN/AML patient 1 (B) and 2 (C). Data represent means \pm SD from two independent experiments. *P < 0.05, **P < 0.01, ***P < 0.001.

Figure S6. Analysis of hematopoietic differentiation in CN/AML iPSC knockout lines

(A) Flow cytometry analysis of suspension cells harvested from EBs culture on day 14 of differentiation of CN/AML iPSC knockout lines of patient 1. Data represent means \pm SD from two independent experiments. *P < 0.05.

(B) Flow cytometry analysis of suspension cells harvested from EBs culture on day 14 of differentiation of HD, CN/AML2 and CN/AML2 *BAALC* KO iPSC lines. Data represent means \pm SD from two independent experiments. **P < 0.01, ***P < 0.001.

(C) Alignment of Sanger sequencing traces from parent control iPSC line (upper panel) and base edited HD *BAALC* KO (lower panel) confirms the introduction of *BAALC* knockout in a single cell derived iPSC clone. Positions of based edited nucleotides are depicted with black arrows, PAM sequence of guide RNA is marked above the alignment by red line.

Figure S7. Treatment of Kasumi-1 cells with MEK inhibitor UO126

(A) Kasumi-1 cells were treated with 1, 5 and 10 μ M UO126 for 7 days. Data represent means \pm SD from two independent experiments.

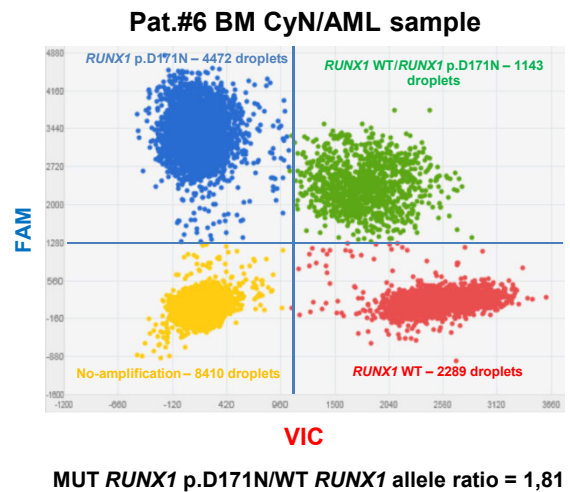
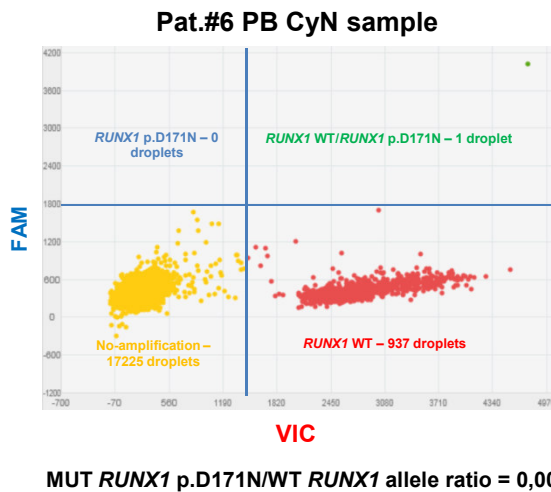
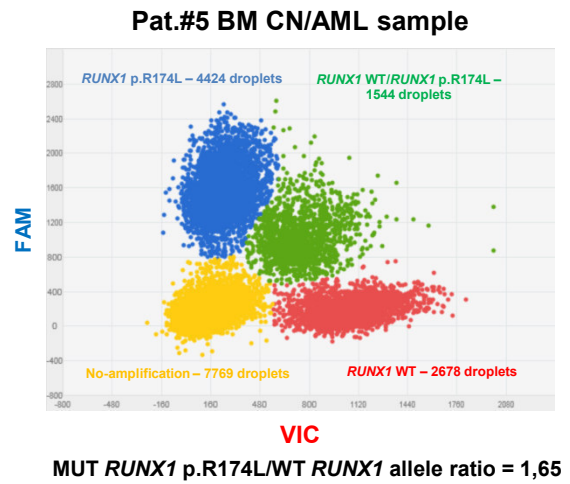
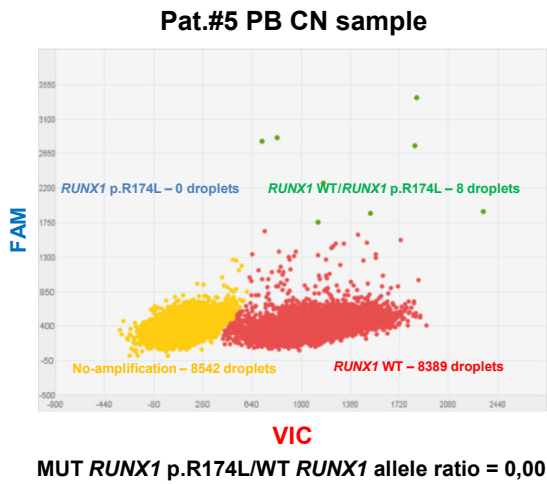
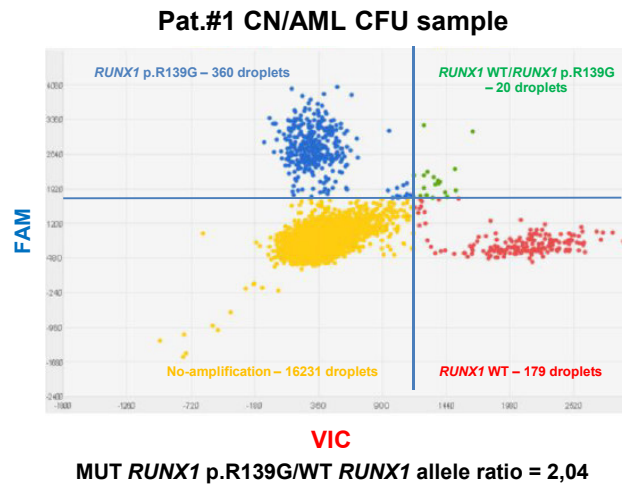
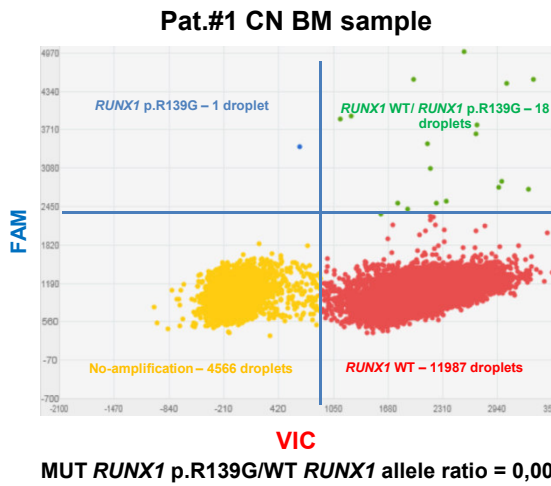
A

Figure S1

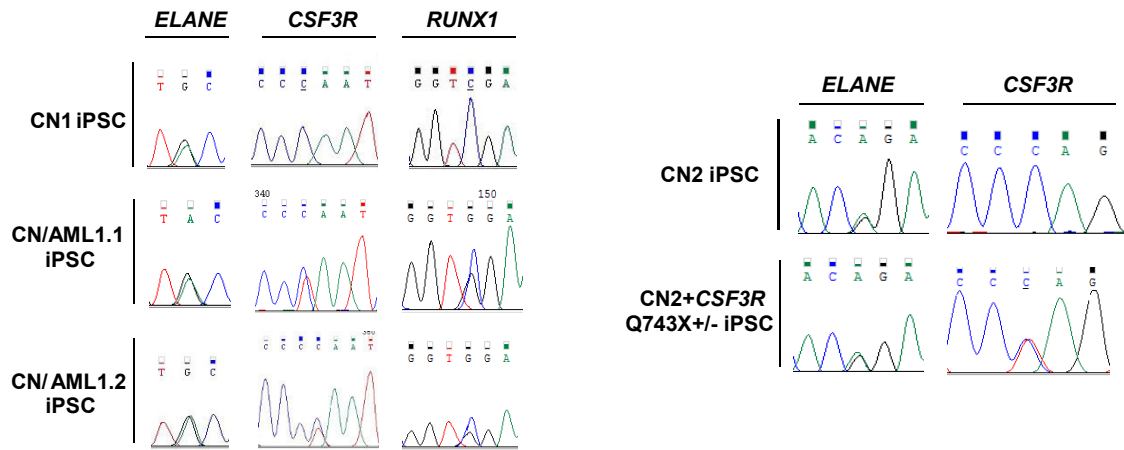
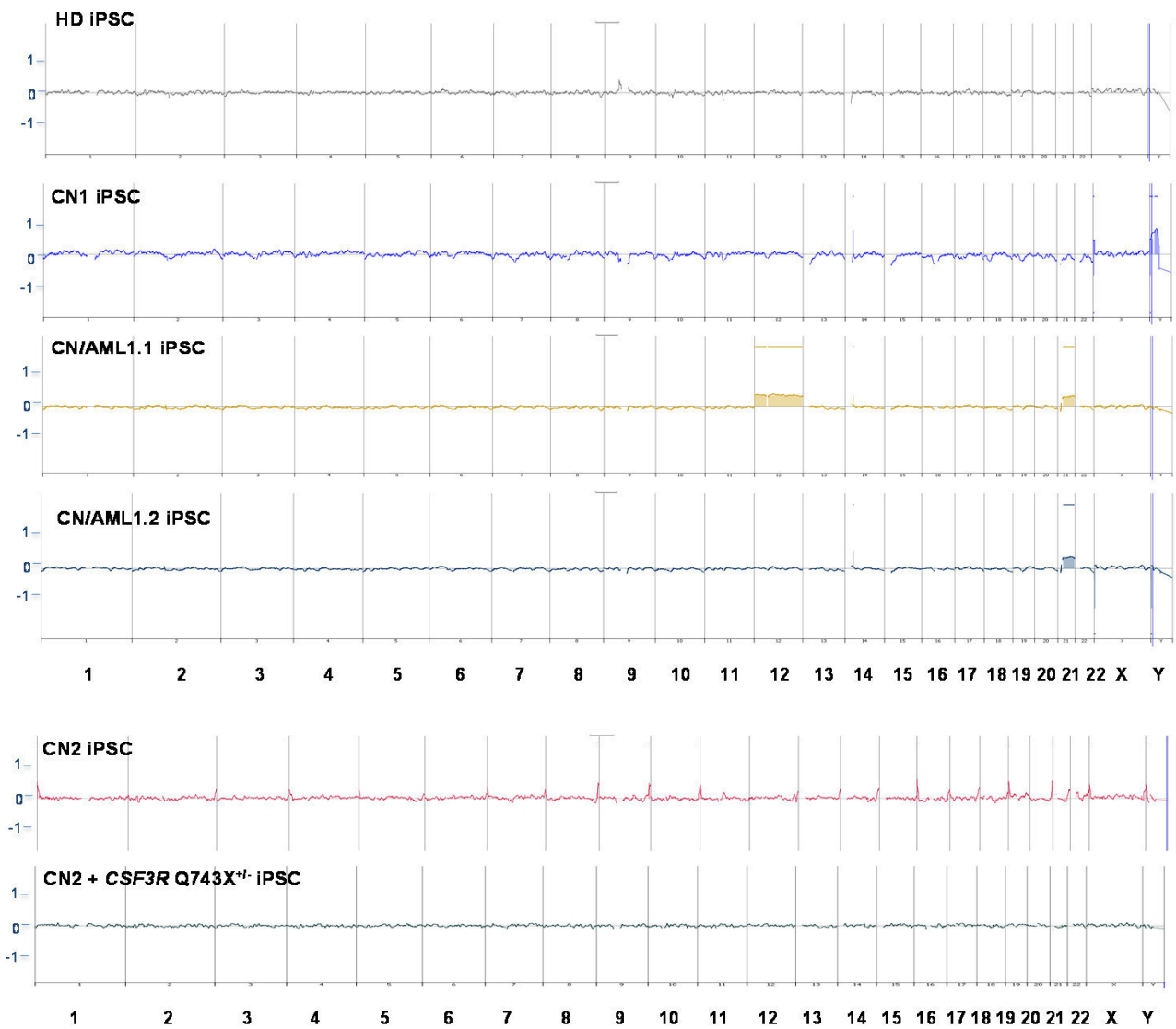
A**B**

Figure S2

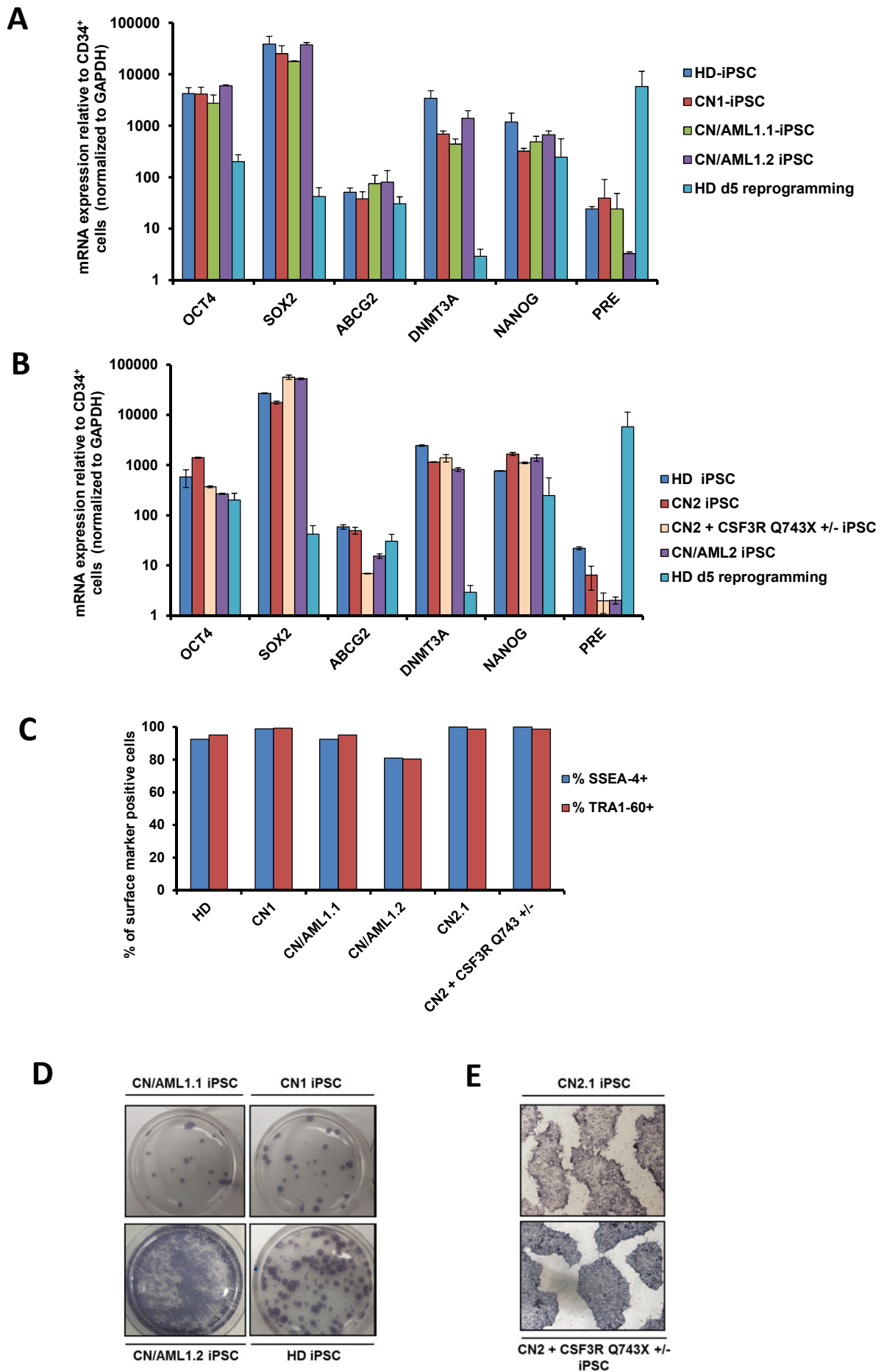
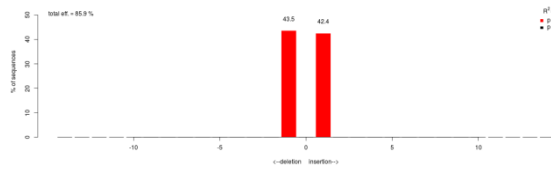


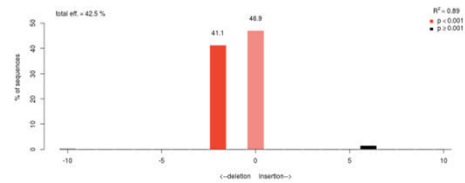
Figure S3

A

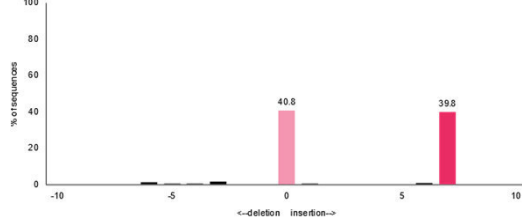
CN1 + CSF3R Q741fs
CSF3R het. NP_000751.3:p.(Ser742PhefsX29),
het. NP_000751.3:p.(Gln741AsnfsX59)

**B**

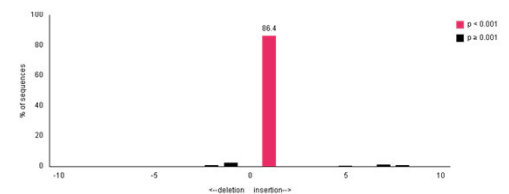
CN1 + CSF3R R734fs
CSF3R het. NP_000751.3:p.(Arg734SerfsX35)

**C**

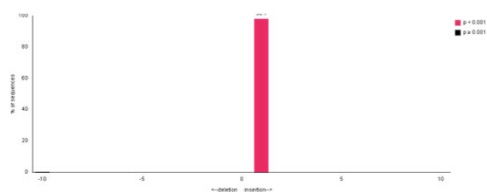
CN/AML2
RUNX1 het. NP_001001890.1:p.(Glu175SerfsX7)

**D**

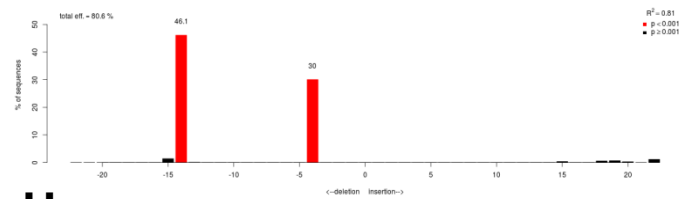
CN/AML1.2 RUNX1-KO
RUNX1 hom. NP_001001890.1:p.(Ser21GlufsX88)

**E**

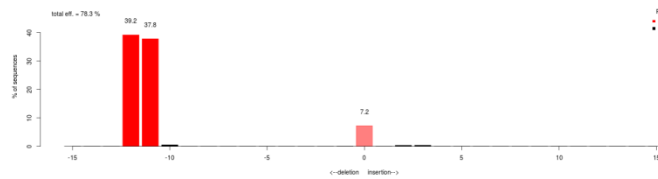
CN/AML1.2 BAALC-KO
BAALC hom. NP_079088.1:p.(Tyr15LeufsX68)

**F**

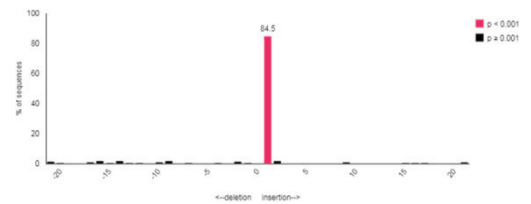
CN/AML1.2 CD109-KO
CD109 het. NP_001153060.1:p.(Leu13ArgfsX62)
het. NP_001153060.1:p.(Val15Alafs13X)

**G**

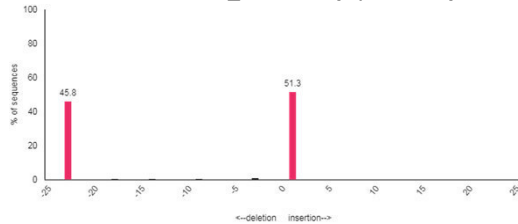
CN/AML1.2 HPGDS-KO
HPGDS het. NP_055300.1:p.(Asn3HisfsX2),
het. NP_055300.1:p.(Pro2_Lys5del)

**H**

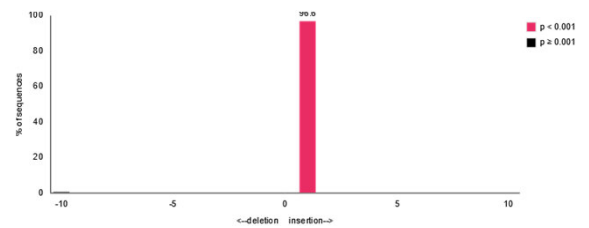
CN/AML2 BAALC-KO C1
BAALC hom. NP_079088.1:p.(Tyr15LeufsX68)

**I**

CN/AML2 BAALC-KO C2
BAALC het. NP_079088.1:p.(Tyr15LeufsX68)
BAALC het. NP_079088.1:p.(Glu12AspfsX76)

**J**

CN1 BAALC-KO
BAALC hom. NP_079088.1:p.(Tyr15LeufsX68)



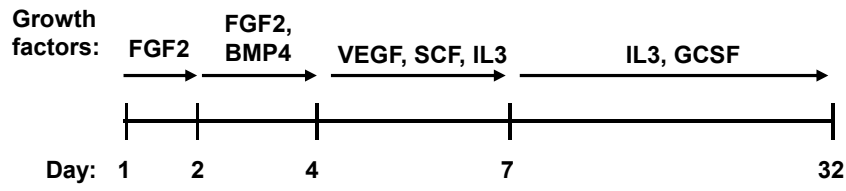
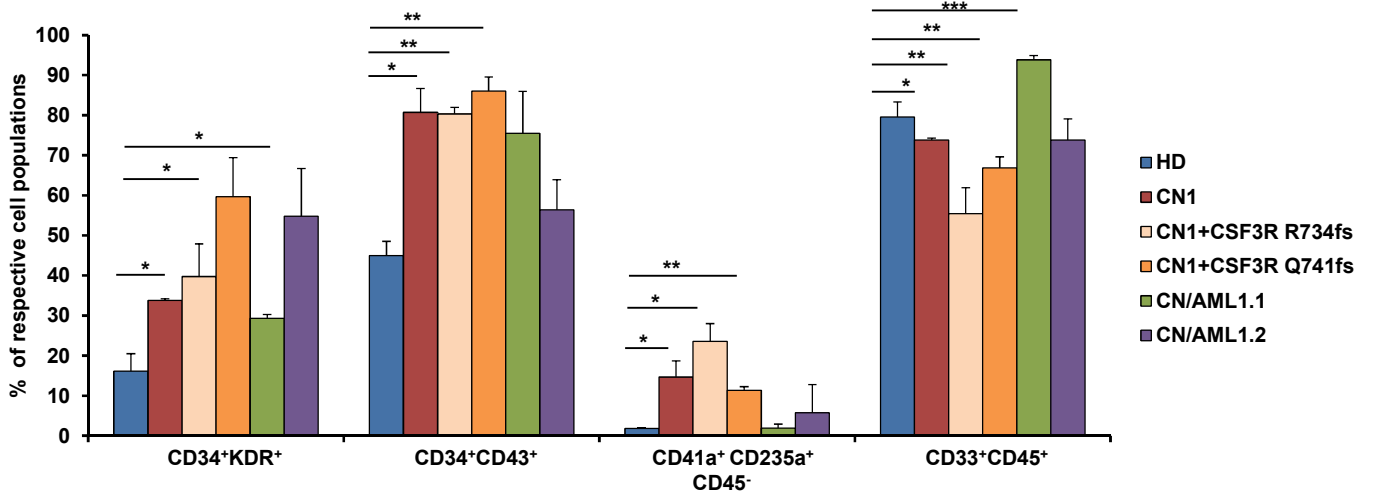
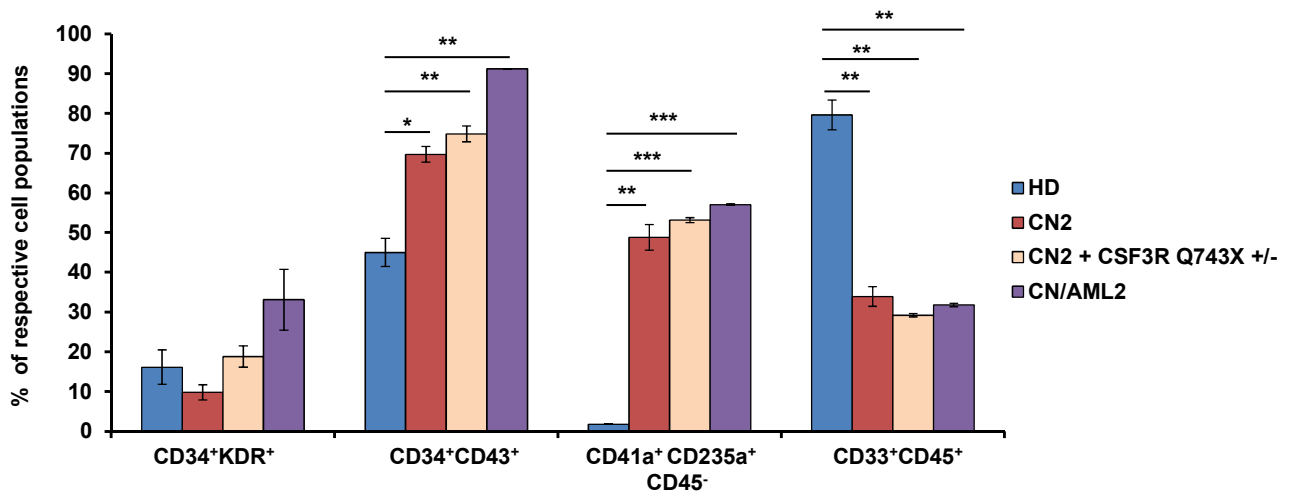
A**B****C**

Figure S5

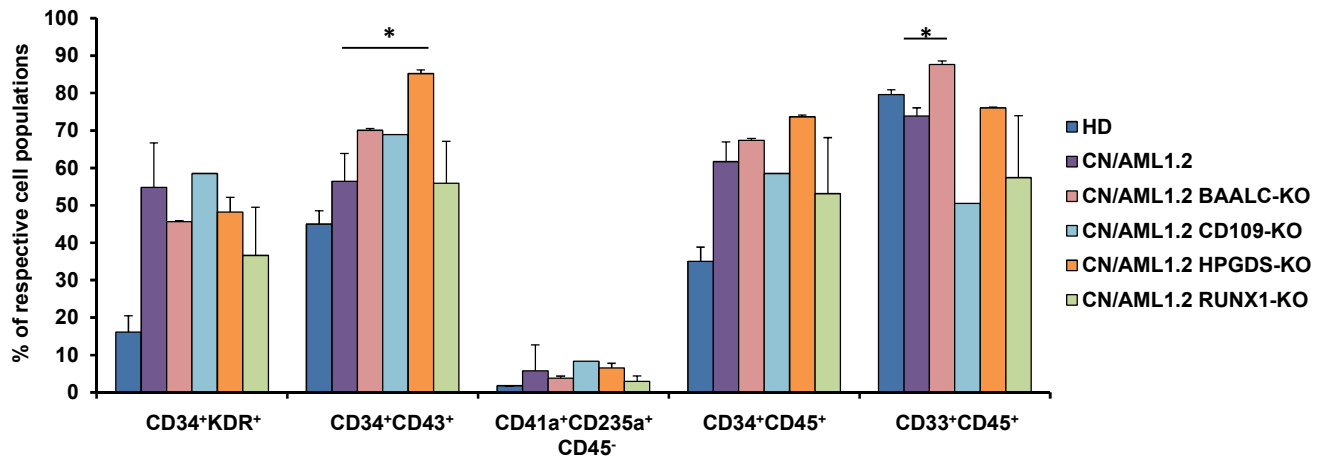
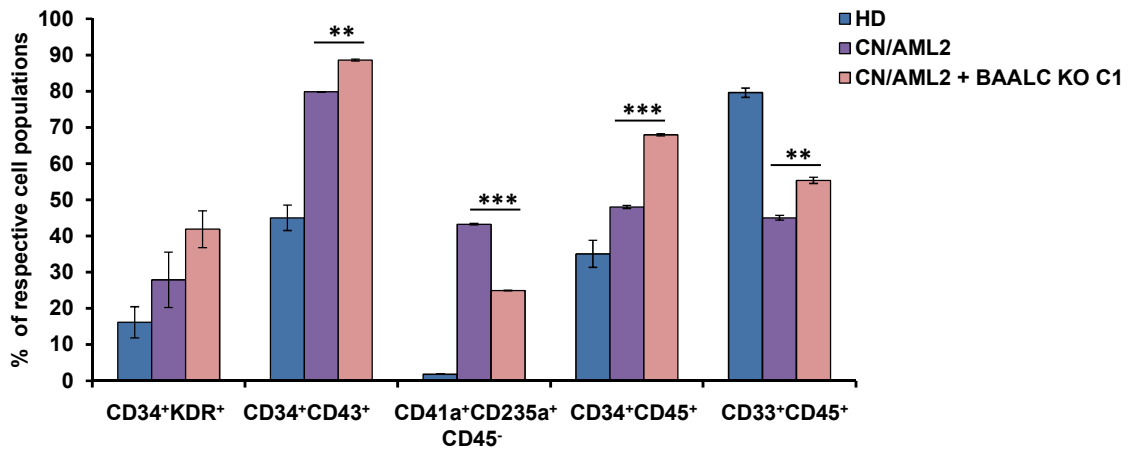
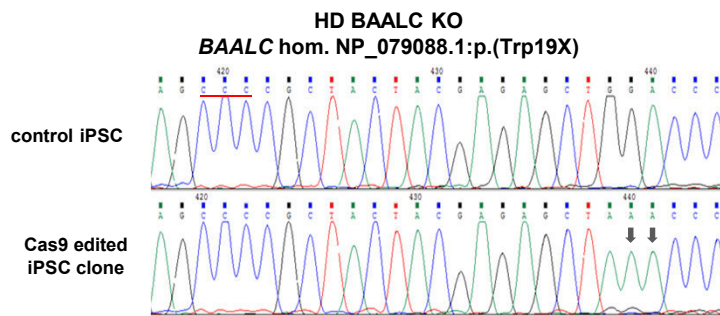
A**B****C**

Figure S6

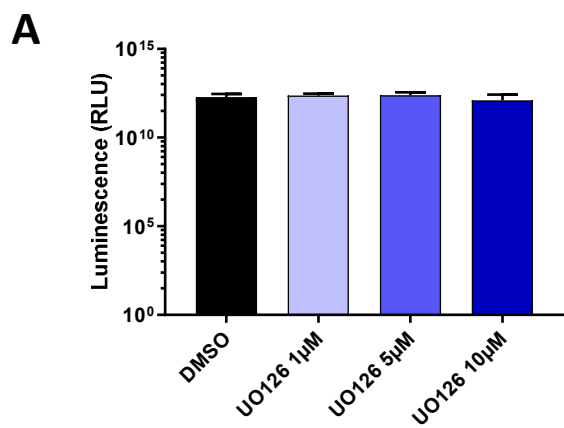


Figure S7

Supplementary Tables

Table S1. Genetic characterization of CN/AML patient group.

Table S2. Guide RNA used for CRISPR/Cas9 gene editing.

Table S3. List of off-targets sites inspected in CRISPR/Ca9-edited iPSC lines by means of Sanger sequencing.

Table S4. List of DE genes between HSPCs derived from CN AML and CN iPSCs of patient 1.

Table S5. List of DE genes between HSPCs derived from CN AML and CN iPSCs of patient 2.

Table S6. List of DE genes between HSPCs derived from CN AML and CN AML *BAALC* KO iPSCs of patient 1.

Table S7. List of DE genes between HSPCs derived from CN AML and CN AML *BAALC* KO iPSCs of patient 2.

Table S1. Genetic characterization of CN/AML patient group.

Patient number	AML subtype	Karyotype	Inherited mutation	Acquired <i>RUNX1</i> mutation	Acquired <i>CSF3R</i> mutation
1	AML M1	45,XY,-7[9];46,XY[11] (2010)47,XY +21[13];46,XY[2] (2011)	<i>ELANE</i> (C151Y)	R139G	Q718X
2	AML M4	46,XY	<i>ELANE</i> (G214R)	R139Q	Q720X
3	RAEBT/AML FAB NA	47,XY, +21 [14] /46, XY [4]	<i>ELANE</i> (G174R)	D171N	Q718X, Q726X
4	AML M2	47,XX,+21	<i>GPT1</i>	K83Q	Q720X
5	AML/B-ALL	46,XY,add(21q)	<i>ELANE</i> (A57V)	R174L	neg
6	AML M2	47,XY,+21	<i>ELANE</i> (A233P, V235TfsX)	D171N	Q741X
7	AML M1	46,XY-7, +21	<i>ELANE</i> (C151S)	R135K	Q731X
8	AML M5	45,XX,-7	<i>ELANE</i> (L152P)	R80S	Q726X
9	AML FAB NA	46, XY, t(9;11)	<i>ELANE</i> (N113K)	R64P	Q718X
10	AML FAB NA	45,XY,-7	<i>WAS</i> (L270P)	R80S	Y729X
11	MDS	46,XY	<i>HAX1</i> (V44X)	I22K	neg
12	MDS	46,XX	<i>HAX1</i> (V44X)	L29S, R64P	Y729X
13	AML M1	46,XY	<i>ELANE</i> (IVS4+1G>T)	K83Q	Q718X
14	MDS RAEB-2	46,XX,dup(21)(q22.1q22.3)[19]	<i>Neg</i>	S114P, Y380G394delinsC	Q726X
15	pre-B ALL	48,XX,del(5)(q21q34),p21,p 22(16)/46,XX[8]	<i>ELANE</i> (G185R)	A160T, S114X	Q702X
16	AML M1	t(p1;q3)	<i>ELANE</i> (C151Y)	R139X	Q731X
17	MDS RAEB	46,XX,add(2)(q37),add(7)(q22)	<i>WAS</i>	Q370X	Y729X
18	AML M2	47,XX +mar[8], 47, idem, del(10)(q32)	<i>ELANE</i> (G214R) <i>GFI1</i>	R174X, L294QfsX6	Q739X
19	AML M0	45,XX,-7[12];46,XX[11]	<i>WAS</i> (S478I)	Intron 4, c.415_427dup6 Intron 4, c.421_427dup7	Q707L
20	MDS/ AML M1	46,XY-7, +21	<i>ELANE</i> (S126L)	R135K	Q726P
21	MDS RAEB	45,XY -7 [10], 46XY [5]	<i>HAX1</i> (V44X)	F13TrpfsX14, R139ProfX47	Q726P
22	AML FAB NA	46,XX	<i>ELANE</i> (A79VfsX9)	R139X, V137D	neg
23	B-ALL/AML	49,XXdel(5)(q13q33),+21,+21,+22[7] /47,XX,-7,+21,mar1/46,XX[3]	<i>ELANE</i> (V101M)	R139fs	Q754X

Table S2. Guide RNA used for CRISPR/Cas9 gene editing.

sgRNA name	sgRNA sequence	On-target score ¹	Specificity score ²	Out of frame score ³	Number of off-targets for 0-1-2-3-4 mismatches, total number of mismatches
<i>BAALC</i> p.14	GTCCAGCTCTCGTAGTAGCG	70	99	67	0 - 0 - 0 - 0 - 13, 13 off-targets
<i>RUNXI</i> p.20	CATCTTGCCTGGGCTCAGCG	69	68	62	0 - 0 - 3 - 36 - 234, 273 off-targets
<i>RUNXI</i> p.R174	CACTTACTTCGAGGTTCTCG	61	47	68	0 - 0 - 0 - 2 - 33, 35 off-targets
<i>CSF3R</i> p.Q741	CTGGTGCCAGACTGGGATTG	50.7	66	64	0 - 0 - 2 - 26 - 229, 257 off-targets
<i>CSF3R</i> p.R734	GCTGGGTGGAAACTGCTCTT	33.2	77	71	0 - 0 - 1 - 14 - 142, 157 off-targets
<i>CD109</i> p.14	GCGCGGCGGTGCACACGCAG	62	86	65	0 - 0 - 0 - 9 - 49, 58 off-targets
<i>HPGDS</i> p.2	AGTGAGTTTGTAGTTTGCA	60.3	67	65	0 - 0 - 0 - 23 - 200, 223 off-targets

¹Doench, Fusi, et al., 2016. Score is from 0-100. Higher is better.

²Hsu et al., 2013. Score is from 0-100. Higher is better.

³Bae et al., 2014. Score is from 0-100. Higher is better.

Table S3. List of off-targets sites inspected in CRISPR/Ca9-edited iPSC lines by means of Sanger sequencing.

sgRNA name	iPSC lines tested for off-targets	Off-target sites selected for inspection by Sanger sequencing	Off-target site sequence	Positions of the mismatches on the off-target site	Off-target site sequencing results
BAALC p.14	HD <i>BAALC</i> KO, CN1 <i>BAALC</i> KO, CN/AML1.2 <i>BAALC</i> KO, CN/AML2 <i>BAALC</i> KO C1, CN/AML2 <i>BAALC</i> KO C2	4:intergenic:RP11-672A2.1-RP11-21L23.4	CTCCAGCCCTCACAGTAGCGAGG	*.....*..**.....	negative
		4:intergenic:RNA5SP358-RNA5SP359	GGCCAGCTCTAGTGGCAGCGCGG	.*.....*..*.*....	negative
		4:intron:ERC2	TTCTTGCTCTCTTAGTAGCTAGG	*..*.....*.....*	negative
RUNX1 p.20	CN/AML1.2 <i>RUNX1</i> KO	2:intergenic:AFF2-IT1-AFF2	CATCCTGCCTGGGCTCAGCTGGG*.....*	negative
		3:intergenic:RP11-282I1.2-GPR26	CATTTTGCCAAGGCTCAGCGTGG	..*.....**.....	negative
		3:intron:SORCS2	GATCTTGCTCTGGGCTCAGCTGGG	*.....*.....*	negative
RUNX1 p.R174	CN/AML2	4:exon:RP11-244N9.4	TATTTACTCTGAGGTTCTCGCGG	*.*.....**.....	negative
		4:exon:MAP4K4	CACTTACTATGTAGTTCTCGTGG**.*.....	negative
		4:exon:RUNX2	TACTTACTTCTGGGTTCCCGAGG	*.....**.*.....*	negative
CSF3R p.Q741	CN1+ <i>CSF3R</i> Q741fs	2:intron:C8B	CAGTTGCCAGACTGGGATTGAGG	.*.*****.....	negative
		3:intron:LOXHD1	CTGGCGCTGACTAGGATTGAGG*..*....*	negative
		3:intron:GRHL2	CAGGTGAAAGACTGGGATTGAAG	*.....**.....	negative
CSF3R p.R734	CN1+ <i>CSF3R</i> R7341fs	3:intron:IL15	GCTGGGTGTAAGCTGCCCTTAGG*..*....*	negative
		2:intron:PDE4D	GCTGGGAGGAAACTGCTCATGGA*.....*	negative
		3:exon:MRGPRF	GATGGCTGAAACTGCTCCTGGG	*..*.....*.....*	negative
CD109 p.14	CN/AML1.2 <i>CD109</i> KO	3:exon:CRIM1	CTGCGGGGTTGCACACGCAGCGG	**.....*.....	negative
		4:intergenic:NKX2-5-Y_RNA	AGGAGGCAGTGCACACGCAGAGG	**.*.....*	negative
		3:intergenic:EFNA2-RPS15P9	GCGGGGCGGGGCACACAGGGG	..*.....*.....*	negative
HPGDS p.2	CN/AML1.2 <i>HPGDS</i> KO	4:intron:BLMH	AGTAAGCCTGTAGTTAGGCATGG	...*..**.....*....	negative
		3:intron:SERPINI1	AGTGAGTATTTAGTTTGGCTGGG*.*.....*	negative
		3:intron:NUMBL	TGTGTGTTTGTGGTTTGGCAGGG	*..*.....*.....	negative

Supplementary references

1. Love, M.I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* **15**, 550 (2014).
2. Mootha, V.K., *et al.* PGC-1alpha-responsive genes involved in oxidative phosphorylation are coordinately downregulated in human diabetes. *Nat Genet* **34**, 267-273 (2003).
3. Babicki, S., *et al.* Heatmapper: web-enabled heat mapping for all. *Nucleic Acids Res* **44**, W147-153 (2016).
4. Clarke, D.J.B., *et al.* eXpression2Kinases (X2K) Web: linking expression signatures to upstream cell signaling networks. *Nucleic Acids Res* **46**, W171-W179 (2018).
5. Lamb, J., *et al.* The Connectivity Map: using gene-expression signatures to connect small molecules, genes, and disease. *Science (New York, N.Y.)* **313**, 1929-1935 (2006).