Dissertation

Deciphering Effects of *Idh1*-R132H Mutations on the Regulation of Hematopoietic Differentiation

Jens Langstein

2020

Dissertation

Submitted to the
Combined Faculty of Natural Sciences and Mathematics
of the Ruperto Carola University Heidelberg, Germany
for the degree of
Doctor of Natural Sciences

Presented by

M.Sc. Jens Langstein
born in Frankenthal, Germany

Oral Examination: 13.03.2020

Deciphering Effects of *Idh1*-R132H Mutations on the Regulation of Hematopoietic Differentiation

Referees:

PD Dr. Odilia Popanda

Dr. Michael Milsom

"If I have seen further, it is by standing on the shoulders of giants."

Isaac Newton

Summary

The hematopoietic system is a highly versatile regenerative tissue, in which hematopoietic stem cells drive the life-long production of multiple mature blood cell types. During hematopoietic differentiation, the regulation of genome-wide epigenetic patterns of histone modification or DNA methylation marks is an essential process orchestrating cell identities, lineage decisions and developmental cell fates. In acute myeloid leukemia, mutations frequently affect direct and indirect epigenetic regulators and modifiers such as isocitrate dehydrogenase 1 (IDH1) or DNA methyltransferase 3 alpha (DNMT3A), and result in disturbed epigenetic landscapes and differentiation patterns. Here, *IDH1* mutations promote oncogenic transformation through the *de novo* production of the metabolite D2-hydroxyglutarate, which induces a genome and epigenome instability by inhibiting multiple histone and DNA demethylases. Yet, molecular details of how *IDH1* mutations alter characteristics of individual hematopoietic cell types remain poorly understood.

In the course of this thesis, combinatorial mouse models carrying specific *Idh1*-R132H and *DNMT3A*-R882H mutations, which frequently co-occur in acute myeloid leukemia patients, were extensively characterized. By integrating phenotypic readouts in combination with latest advances in high-throughput single-cell RNA-sequencing approaches, cooperativity and impact of these mutations on individual cell types of the hematopoietic system were delineated and gene regulatory networks which are altered upon the expression of an *Idh1*-R132H or a *DNMT3A*-R882H mutation were identified. At a phenotypic level, neither an *Idh1*-R132H mutation alone nor in combination with a *DNMT3A*-R882H mutation resulted in the development of myeloid malignancies, suggesting a restricted oncogenic potential of these mutations and additional intrinsic or extrinsic factors to be required for further malignant transformation. However, *Idh1*-R132H mutated hematopoietic stem cells displayed increased engraftment and reconstitution potential during serial transplantations and featured aberrant expression of genes associated with DNA damage and DNA repair. Furthermore, both *Idh1*-R132H single-mutant and *Idh1*-R132H *DNMT3A*-R882H double-mutant mice displayed aberrant differentiation patterns predominantly affecting the myelo-monocytic lineage, culminating in a favored monocytic cell fate and increased monocyte and monocyte progenitor counts in the bone marrow.

By employing a multi-layered single-cell transcriptome analysis of nearly all cell types within the hematopoietic compartment, differentiation trajectories from hematopoietic stem cells towards mature differentiated cells were reconstructed and underlying molecular defects characterized. Pseudotime-inferred myeloid lineage trajectories revealed an aberrant lineage specification in particular for *Idh1*-R132H *DNMT3A*-R882H double-mutated myeloid progenitor cells, resembling a differentiation arrest at the stage of common myeloid progenitors and an ineffective hematopoietic differentiation as seen in myelodysplastic syndromes. At the molecular level, this aberrant population was characterized by an altered metabolic signature and elevated Myc signaling, which is involved in the regulation of terminal myeloid differentiation. Importantly, we could correlate this transcriptome-defined population to a surface marker-defined population, allowing the prospective isolation of these cells for further investigation.

Independent of a *DNMT3A*-R882H mutation, the expression of an *Idh1*-R132H mutation resulted in the deregulation of several key regulatory factors which either orchestrate monocyte and macrophage development or their activation upon inflammatory stimuli. In line with this, monocyte progenitor cells displayed elevated

interferon signaling levels, suggesting that a proinflammatory environment is a common characteristic of an *Idh1*-R132H mutated hematopoietic compartment and could contribute to leukemic transformation upon additional events.

In summary, the experimental framework presented in this thesis enhanced our understanding of how *IDH1*-R132H mutations alone or in combination with a *DNMT3A*-R882H mutation in patients synergistically drive leukemia initiation and progression. The identified molecular characteristics will be of benefit in designing treatment strategies for patients carrying *IDH1*-R132H and *DNMT3A*-R882H mutations and can be used as a resource when studying these mutations in the context of altered physiological conditions and upon additional extrinsic stimuli.

Zusammenfassung

Das hämatopoetische System ist ein äußerst dynamisches regeneratives Gewebe, in dem hämatopoetische Stammzellen für die lebenslange Produktion von verschiedenen reifen Blutzelltypen verantwortlich sind. Die Regulation genomweiter epigenetischer Muster der DNA Methylierung und Histonmodifikation ist ein wichtiger Mechanismus zur Kontrolle von Zellidentitäten und der Spezifikation von Differenzierungsprozessen. Bei der akuten myeloischen Leukämie beeinflussen Mutationen häufig direkte und indirekte epigenetische Regulatoren wie z.B. Isocitrat-Dehydrogenase 1 (IDH1) oder DNA-Methyltransferase 3 alpha (DNMT3A) und führen zu veränderten epigenetischen Strukturen und Differenzierungsmustern. Hier unterstützen *IDH1* Mutationen die onkogene Transformation durch die *de novo* Produktion des Metaboliten D2-Hydroxyglutarat, welcher durch die Inhibierung verschiedener Histon- und DNA-Demethylasen eine Instabilsierung des Genoms und des Epigenoms hervorruft. Allerdings sind molekulare Grundlagen wie sich die Expression von *IDH1* Mutationen auf verschiedene hämatopoetische Zelltypen auswirkt, nur unzureichend verstanden.

Im Rahmen dieser Doktorarbeit wurde eine umfassende Charakterisierung von kombinatorischen Mausmodellen mit spezifischen, in Patienten mit akuter myeloischer Leukämie auftretenden, *Idh1*-R132H und *DNMT3A*-R882H Mutationen durchgeführt. Durch die Integration von phänotypischen Analysen in Kombination mit Hochdurchsatz-Einzelzell-RNA-Sequenzierungsmethoden konnte ein Zusammenspiel beider Mutationen und deren Einfluss auf einzelne Zelltypen des hämatopoetischen System detailliert untersucht werden. Des Weiteren wurden Genregulationsnetzwerke identifiziert, die durch die Expression einer *Idh1*-R132H oder einer *DNMT3A*-R882H Mutation beeinträchtigt werden.

Auf phänotypische Ebene führte weder eine *Idh1*-R132H Mutation alleine noch in Kombination mit einer *DNMT3A*-R882H Mutation zur Entwicklung myeloider Neoplasien, was auf ein beschränktes onkogenes Potenzial dieser Mutationen schließen lässt. Vielmehr deutet dies darauf hin, dass zusätzliche intrinsische oder extrinsische Faktoren auftreten müssen, damit es zu einer malignen Transformation im Blutsystem kommt. *Idh1*-R132H mutierte hämatopoetische Stammzellen wiesen jedoch ein erhöhtes Engraftment- und Rekonstitutionspotenzial in seriellen Transplantationsexperimenten auf, sowie eine Deregulation von Genen, die mit der Reparatur von DNA-Schäden in Verbindung gebracht werden. Darüber hinaus zeigten sowohl *Idh1*-R132H einzel-mutierte als auch *Idh1*-R132H *DNMT3A*-R882H doppel-mutierte Mäuse aberrante Differenzierungsmuster, die vorwiegend die myelo-monozytäre Linie beeinflussten und zu einer erhöhten Anzahl an monozytären Vorläuferzellen und reifen Monozyten im Knochenmark der Mäuse führten.

Durch die Anwendung von Einzelzell-Transkriptomanalysen von nahezu allen hämatopoetischen Zelltypen konnten Differenzierungsmuster von hämatopoetischen Stammzellen zu terminal differenzierten Zellen rekonstruiert und die zu Grunde liegenden molekularen Defekte analysiert werden. Insbesondere *Idh1*-R132H *DNMT3A*-R882H doppel-mutierte Zellen zeigten stark abweichende myeloide Differenzierungsmuster, die einem Differenzierungsarrest auf dem Stadium der gemeinsamen myeloiden Vorläuferzellen und einer ineffektiven hämatopoetischen Differenzierung, wie sie charakteristisch für myelodysplastische Syndrome ist, ähneln. Auf molekularer Ebene wies diese aberrante Population eine veränderte metabolische Signatur sowie eine erhöhte Myc-Aktivität, welche die terminale Differenzierung myeloider Zellen reguliert, auf. Des Weiteren konnte diese Transkriptom-definierte

Population mit einer Oberflächenmarker-definierten Population korreliert werden, sodass die prospektive Isolation dieser Zellen für weitere Analysen ermöglicht wurde.

Unabhängig von einer *DNMT3A*-R882H Mutation führte die Expression einer *Idh1*-R132H Mutation zur Deregulierung von mehreren zentralen Faktoren, die entweder die Entwicklung von Monozyten und Makrophagen, oder deren Aktivierung durch inflammatorische Stimuli, steuern. In Übereinstimmung damit zeigten monozytäre Vorläuferzellen eine erhöhte Aktitivität des Interferon Signalweges, was darauf hindeutet, dass ein entzündungsförderndes Millieu ein Merkmal eines *Idh1*-R132H mutierten hämatopoetischen Kompartiments ist und zu einer leukämischen Transformation beitragen könnte.

Zusammenfassend liefert der hier beschriebene experimentelle Datensatz neue Erkenntnisse, wie *IDH1*-R132H Mutationen alleine oder zusammen mit *DNMT3A*-R882H Mutationen in Patienten die Initiierung oder das Fortschreiten einer leukämischen Erkrankung beeinflussen. Die identifizierten molekularen Grundlagen sind entscheidend für die Entwicklung von neuen Therapieansätzen für Pateinten mit einer *IDH1*-R132H und / oder einer *DNMT3A*-R882H Mutation und dienen zudem als Ressource für Untersuchungen dieser Mutationen im Kontext veränderter physiologischer Bedingungen oder zusätzlich angewandter externer Stimuli.

Contents

4 Discussion

5 Materials & Methods

A Appendix

1

Introduction

1.1 The Hematopoietic System

The mammalian blood system represents a highly dynamic regenerative tissue aimed at the life-long production and regeneration of mature blood cell types. The renewal of hematopoietic cells is essential for the perpetuation of physiological functions and takes place in the adult bone marrow with around 10^{12} cells being produced every day [Doulatov et al., 2012]. The hematopoietic system encompasses a variety of mature cell types, including erythrocytes, platelets, B-cells, T-cells, natural killer cells, monocytes, macrophages, dendritic cells, granulocytes and mast cells [Seita and Weissman, 2010]. The functions of these cell types are very diverse; erythrocytes for instance are required for the supply of oxygen, platelets drive the coagulation of blood upon injuries, whereas other cell types are involved in adaptive and naïve immunity in response to pathogen infections. The short life-span of these cells, however, requires a continuous replenishment of mature cells, which, despite their diversity, all share one common progenitor cell.

1.1.1 Concepts of Hematopoietic Stem Cells

The existence of stem cells in the hematopoietic system was first proposed in the 1960s by Till and McCulloch in seminal studies investigating the regenerative potential of bone marrow cells [Till and McCulloch, 1961, Till et al., 1964, Becker et al., 1963]. Yet, only upon the advent of new technologies, such as fluorescence activated cell sorting (FACS) or monoclonal antibodies, these cells could prospectively be purified [Spangrude et al., 1988]. Combining cell surface markers with functional readouts and *in vivo* reconstitution experiments over the years has led to a detailed characterization of individual cell types and their functional properties within the hematopoietic system. Based on this understanding, cell types were consolidated into a cellular hierarchy that reflects developmental stages of hematopoietic differentiation with hematopoietic stem cells (HSCs) residing at the apex of the hierarchy [Laurenti and Gottgens, 2018, Haas et al., 2018].

Akin to other tissue-specific stem cells in the intestine or the skin, HSCs are defined by two essential properties. While they have the potential to produce all differentiated

cell types within the hematopoietic system, they at the same time have the ability to self-renew and perpetuate themselves by dividing and giving rise to a new stem cell [Seita and Weissman, 2010, Laurenti and Gottgens, 2018]. Maintenance of the stem cell pool through self-renewal is a crucial measure to ensure the sustained production of cells throughout the life time of an individual. In contrast, downstream cells lack any extended self-renewal potential or are confined in their capacity to regenerate only a certain subset of mature cell types [Morrison and Weissman, 1994, Christensen and Weissman, 2001, Doulatov et al., 2010].

The number of stem cells that are actively contributing to blood production at any given time is estimated to range between 50,000 and 200,000 in humans [Lee-Six et al., 2018]. Yet, under homeostatic conditions cellular output is mainly sustained by short-lived and restricted multipotent progenitors (MPPs) with only limited contribution from actual HSCs [Busch et al., 2015, Sun et al., 2014b, Sawai et al., 2016]. To prevent early stem cell exhaustion and an acquisition of genetic aberrations, HSCs reside in a quiescent state with only infrequent cell divisions, but are efficiently activated upon injury, blood loss, infection, or inflammation [Wilson et al., 2008, Essers et al., 2009, Baldridge et al., 2010, Cheshier et al., 2007, Takizawa et al., 2011]. Unfavorably, an exit from quiescence is associated with a drastic impairment of stem cell function due to the accumulation of DNA damage [Walter et al., 2015, Flach et al., 2014, Rossi et al., 2007], highlighting the importance of regulatory mechanisms that balance quiescence, activation, self-renewal and differentiation in order to assert homeostatic blood production.

1.1.2 Classical Models of Hematopoietic Differentiation

The classical hierarchical model depicts hematopoietic differentiation as a stepwise process which is defined by discrete cell stages and a gradual loss of both self-renewal potential and multipotency paralleling hematopoietic commitment to certain lineages [Seita and Weissman, 2010]. Binary branching points successively segregate lineages and cell fates, leading to a progression from multipotent via oligo- and bipotent to unipotent progenitor states [Seita and Weissman, 2010]. Within the stem cell population, several distinct subsets are distinguished based on the increasing restriction in self-renewal capacity. Whereas long-term HSCs (LT-HSCs) are characterized by the ability to reconstitute the hematopoietic system throughout the whole lifespan, descendant short-term HSCs (ST-HSCs) sustain blood production only for a limited time (usually several weeks) [Morrison and Weissman, 1994, Christensen and Weissman, 2001, Yang et al., 2005b, Kiel et al., 2005]. ST-HSCs give rise to further restricted MPPs, which still possess full lineage differentiation potential, but already lack any repopulation capacity [Morrison and Weissman, 1994, Christensen and Weissman, 2001, Morrison et al., 1997, Kiel et al., 2005].

The first branching point separates lymphoid potential from myeloid, erythroid and megakaryocytic potential and results in the formation of either common lymphoid progenitors (CLP) or common myeloid progenitors (CMPs) (Figure 1-1A) [Kondo et al., 1997, Akashi et al., 2000]. Further restricted bipotent progenitors arise from CMPs, segregated into a megakaryocytic-erythroid fate (megakaryocyte erythroid progenitor (MEP)) or a granulocytic-monocytic fate (granulocyte monocyte / macrophage progenitor (GMP)) [Pronk et al., 2007, Na Nakorn et al., 2002]. Eventually, the next branching points culminate in the emergence of unipotent progenitors which will continue to differentiate into mature megakaryocytes, erythrocytes, monocytes or granulocytes which carry out

characteristic and specialized functions within the hematopoietic compartment [Laurenti and Gottgens, 2018].

1.1.3 Revised Models of the Hematopoietic Hierarchy

Over the years, the discovery of additional surface markers combined with single-cell assays to readout characteristic lineage potentials has led to the introduction of various modifications to the initial hierarchical model. The identification of progenitor subsets that share myeloid and lymphoid potential (lymphoid-primed multipotent progenitor (LMPP)), but no megakaryocytic and erythroid potential, has indicated that myeloid and lymphoid fates are not segregated by the first branching point but only in later bifurcation steps (Figure 1-1B) [Adolfsson et al., 2005]. Additional bi- or unipotent subsets have since been identified in the myeloid-megakaryocytic-erythroid branch based on extended combinations of surface markers [Pronk et al., 2007, Guo et al., 2013]. Similarly, the origin of monocytes, macrophages and dendritic cells has been revised through the identification of monocyte / dendritic cell progenitors (MDPs) and downstream unipotent common monocyte progenitor (cMoPs) as well as common dendritic cell progenitor (CDPs) populations [Geissmann et al., 2010, Fogg et al., 2006, Hettinger et al., 2013, Onai et al., 2007].

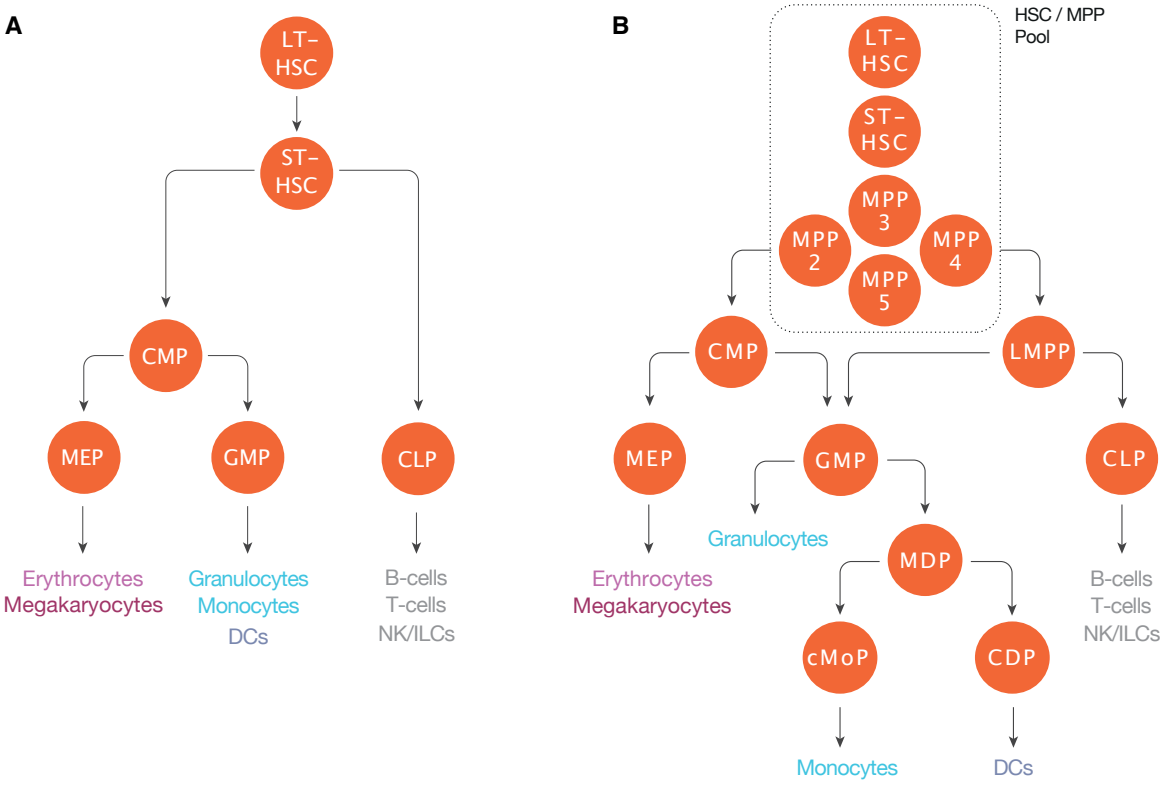


Figure 1-1: Hierarchical Models of Hematopoietic Differentiation.
(A) The earliest hierarchical model of the hematopoietic system depicted differentiation as a stepwise process in which HSCs reside at the apex of the hierarchy. The first binary branching point segregates myeloid and lymphoid potential. **(B)** Over time, additional cell populations have been identified that have led to modifications of the classical hierarchy. The hematopoietic stem and progenitor cell pool consist of heterogeneous subsets that differ in self-renewal and differentiation potential, myeloid and lymphoid potential remains associated during the first branching points and additional bi-or unipotent progenitor populations drive the production of mature cell types. The horizontal axis within the HSC / MPP pool illustrates biased differentiation potential of MPP subsets (MPP2 = megakaryocytic / erythroid, MPP3 = myeloid, MPP4 = Lymphoid) (Figure adapted from [Laurenti and Gottgens, 2018]).

Other studies revealed that the MPP compartment can be subdivided in further subsets (MPP2, MPP3, MPP4, MPP5) that are already primed to differentiate into certain lineages [Pietras et al., 2015, Wilson et al., 2008]. Here, MPP2 cells were predominantly associated with a megakaryocytic and erythroid cell fate, whereas MPP3 and MPP4 populations featured a biased differentiation into myeloid or lymphoid cell fates, respectively [Pietras et al., 2015]. Molecular priming even has been detected at the HSC level, with certain megakaryocytic restricted HSCs directly giving rise to megakaryocytes in emergency situations without progressing though any progenitor stages [Haas et al., 2015, Sanjuan-Pla et al., 2013, Yamamoto et al., 2013, Rodriguez-Fraticelli et al., 2018]. Likewise, HSCs which are dissimilar in their myeloid or lymphoid potential have been identified [Yamamoto et al., 2013, Muller-Sieburg et al., 2004, Dykstra et al., 2007, Rodriguez-Fraticelli et al., 2018], illustrating a high degree of heterogeneity already within the HSC pool.

The Continuous Model of Hematopoietic Differentiation

However, the classical model of hematopoiesis was most fundamentally challenged with the advent of single-cell transcriptomic analysis methods. Profiling gene expression at the single-cell level provides a transcriptomic snapshot as cells transition through hematopoietic differentiation landscapes and thereby allows to deconvolute cellular heterogeneity within specific populations irrespective of any surface marker combinations [Nestorowa et al., 2016, Macaulay et al., 2016, Giladi et al., 2018, Paul et al., 2015]. Integrating single-cell transcriptomic data specifically across the hematopoietic stem and progenitor cell (HSPC) compartment has led to the proposal of a continuous model of the hematopoietic hierarchy, according to which lineage specification is already determined at the level of HSPCs and acquired in a continuous process (Figure 1-2) [Velten et al., 2017, Nestorowa et al., 2016, Macaulay et al., 2016, Giladi et al., 2018]. Without transitioning any multipotent or oligopotent progenitor cell states, unilineage restricted cells originate directly from a continuum of unprimed HSCs and follow their specified differentiation trajectory (Figure 1-2A) [Velten et al., 2017, Nestorowa et al., 2016, Macaulay et al., 2016, Giladi et al., 2018]. Underpinning these findings, multipotent progenitor cells likewise were associated with a restricted unilineage output during native and unperturbed hematopoiesis [Rodriguez-Fraticelli et al., 2018]. Similarly, single-cell analysis of myeloid progenitor populations such as CMPs, GMPs, and MEPs, revealed that these seemingly homogenous populations in truth are composed of distinct and functionally heterogeneous subsets that are primed towards one particular differentiation fate and are rather defined by a unilineage potential (Figure 1-2B) [Paul et al., 2015]. Importantly, understanding the molecular characteristics that drive cellular heterogeneity and cell fate choices is in direct relevance not only to the functionality of the hematopoietic system, but also to the development of hematological malignancies [Haas et al., 2018].

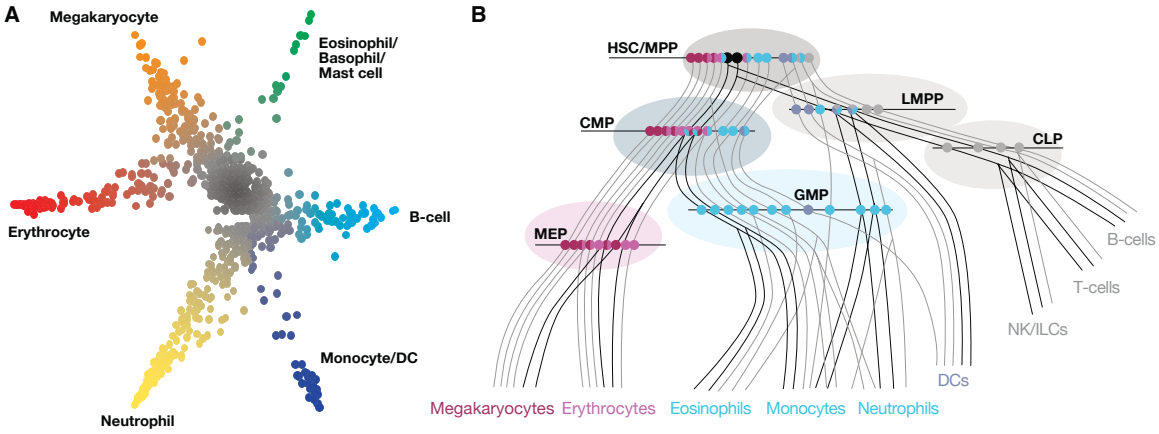


Figure 1-2: Continuous Models of Hematopoietic Differentiation.
(A) Within the HSPC compartment, unilineage restricted HSCs emerge from a continuum of unprimed cells and acquire lineage biases into designated directions in a continuous manner. Accordingly, lineage commitment decisions are already made within the HSC compartment and cells traverse differentiation trajectories without passing discrete and hierarchically organized progenitor stages (Figure adapted from [Velten et al., 2017]). **(B)** As individual and distinctly primed cells traverse denoted trajectories, they pass progenitor compartments, which are defined by certain combinations of surface markers (indicated as ellipses). Although cells share the expression of common surface markers at this stage, their predetermined transcriptional and epigenetic profile defines their developmental path (Figure adapted from [Laurenti and Gottgens, 2018].

1.1.4 Monocyte Development and Function

Monocytes belong to the 'mononuclear phagocyte system' (MPS), a part of the immune system consisting of phagocytic cells and which, beside monocytes, also encompasses conventional dendritic cells (cDCs) and macrophages [Guilliams et al., 2014]. Monocytes develop in the adult bone marrow from myeloid progenitor cells and following their generation are released into blood circulation, where they constitute around 4 % of all leukocytes [Guilliams et al., 2018]. Upon inflammation, monocytes are mobilized in large numbers and recruited to inflamed sites to function as proinflammatory mediators or aid in resolving inflammation by giving rise to several functionally distinct monocyte-derived cells [Shi and Pamer, 2011].

Monocytes are Generated through Mobilization of Distinct Progenitor Subsets

Within the bone marrow, classical monocytes were thought to arise from MDPs, which possess monocytic- and dendritic cell- (DC), but no neutrophil potential [Fogg et al., 2006, Auffray et al., 2009, Geissmann et al., 2010]. MDPs, in a binary cell fate decision, then give rise either to unipotent cMoPs, producing mature monocytes, or CDPs which are ancestral to both cDCs and plasmacytoid dendritic cells (pDCs) [Hettinger et al., 2013, Onai et al., 2007, Naik et al., 2007]. While it was initially believed that MDPs originate from the GMP population, recent evidence suggests that MDPs can directly emerge from the CMP population to give rise to monocytes without any GMP intermediate [Yanez et al., 2017]. Yet, a monocyte specific transcriptional program was identified both in CMP and GMP populations [Paul et al., 2015, Olsson et al., 2016], indicating that functionally distinct monocyte subtypes can originate from two independent pathways, either from GMPs (carrying monocyte and neutrophil potential) or from CMP-derived MDPs (carrying monocyte and DC potential) [Yanez et al., 2017]. The mobilization of different progenitor populations here was shown to depend on the microbial stimulus applied, suggesting that functionally distinct monocyte subtypes are produced only upon demand under certain

conditions and the balance between GMP and MDP differentiation defines the monocyte cell repertoire [Yanez et al., 2017]. On the transcriptional level, commitment towards a monocytic cell fate is associated with the expression of a distinct set of hematopoietic transcription factors, including PU.1, GATA2, IRF8 and KLF4, since disruption of this signaling axis severely disturbs monocyte precursor and peripheral monocyte numbers [Guilliams et al., 2018, Kurotaki et al., 2013]. On a cytokine level, the development and survival of monocytes is contingent on the growth factor colony-stimulating factor 1 / macrophage colony-stimulating factor (CSF1 / M-CSF) and downstream CSF1-receptor (CSF1R / CD115) signaling as deficiency in either CSF1 or CSF1R is associated with a severe monocytopenia [Cecchini et al., 1994, Wiktor-Jedrzejczak and Gordon, 1996, Dai et al., 2002].

Functional Heterogeneity of Monocyte Subsets

Heterogeneity within mature monocytes in humans was first described in 1989 based on their morphology and diverging expression levels of the cell surface markers CD14 and CD16 [Passlick et al., 1989]. In mice, two different monocyte subsets can be distinguished based on the expression of the surface markers Ly6C, CX_3CR1, CCR2, CD62L, and CD43. $Ly6C^{High}$ monocytes (termed classical inflammatory monocytes, corresponding to $CD14^{+}$ $CD16^{-}$ human monocytes) are defined as $Ly6C^{High}$ $CX3CR1^{Int}$ $CCR2^{+}$ $CD62L^{+}$ $CD43^{Low}$, whereas $Ly6C^{Low}$ monocytes (termed non-classical patrolling monocytes, corresponding to $CD14^{Low}$ $CD16^{+}$ human monocytes) are defined as $Ly6C^{Low}$ $CX3CR1^{High}$ $CCR2^{Low}$ $CD62L^{-}$ $CD43^{+}$ [Jung et al., 2000, Geissmann et al., 2003, Palframan et al., 2001, Jakubzick et al., 2013]. Following their generation in the bone marrow, $Ly6C^{High}$ monocytes are released into circulation, equipped with a transcriptional program enabling their migration into tissues [Jakubzick et al., 2017]. Under homeostatic conditions, $Ly6C^{High}$ monocytes remain in circulation for roughly a day and then extravasate into various tissues, such as the intestine, dermis or lung to repopulate tissue-resident macrophages, or alternatively, while remaining in circulation, convert into $Ly6C^{Low}$ monocytes [Bain et al., 2014, Tamoutounour et al., 2013, Jakubzick et al., 2013, Yona et al., 2013, Patel et al., 2017]. $Ly6C^{Low}$ monocytes are characterized by a longer life-span (between two and seven days) and by an extensive retention in circulation where they monitor the vasculature through ‘crawling’ mechanisms and surveil endothelial cell integrity [Auffray et al., 2007, Carlin et al., 2013].

Monocyte Functions During Inflammation

Upon inflammation, $Ly6C^{High}$ monocytes are avidly recruited to the injured tissues and extravasate from circulation in order to fulfill several critical effector functions such as antigen presentation, tissue remodeling, or the modulation of pro- and anti-inflammatory activities [Jakubzick et al., 2017, Guilliams et al., 2018]. Some of the effector functions can be accomplished by monocytes directly due to their potent capacity for phagocytosis and the release of inflammatory mediators upon bacterial- (e.g. reactive oxygen species, tumor necrosis factor alpha (TNFα), nitric oxide, interleukin 1 beta (IL-1β)) and viral infections (e.g. type I interferons alpha and beta (IFNα / β) [Serbina et al., 2008, Barbalat et al., 2009, Yang et al., 2014]. In many instances, however, efficiently resolving an infection or an inflammatory state further requires the differentiation of monocytes to monocyte-derived DCs (moDCs) or macrophages. Whereas moDCs play an essential role during antigen presentation and priming of T-cells [Cheong et al., 2010, Nierkens et al., 2013, Briseno et al., 2016], macrophages are critical for restoring tissue homeostasis by eliminating pathogenic insults and repairing damaged tissue [Wynn et al., 2013, Ginhoux and Guilliams, 2016]. Depending on the microenvironment, macrophages can acquire

heterogeneous activation states with customized functions, termed classically activated (M1-) macrophages, and alternatively activated (M2-) macrophages [Mosser and Edwards, 2008, Lawrence and Natoli, 2011, Murray, 2017]. Bacterial lipopolysaccharide (LPS) or interferon gamma (IFNγ) promotes the generation of M1 macrophages, which are characterized by an antimicrobial activity and the secretion of proinflammatory factors, such as IFNβ, IL-6, IL-12 or TNFα [Mosser and Edwards, 2008, Lawrence and Natoli, 2011, Murray, 2017]. On the other hand, IL-4 or IL-13 favors the generation of M2 macrophages, which are implicated in tissue remodeling, wound healing and the secretion of anti-inflammatory factors [Mosser and Edwards, 2008, Lawrence and Natoli, 2011, Murray, 2017].

Importantly, not only the conversion between monocyte subsets but also the transition from monocytes to macrophages as well as the induction of macrophage polarization programs involves substantial epigenetic remodeling, illustrating that a deregulation of the epigenetic landscape could provoke alterations in monocyte- or macrophage-specific cell fates and respective cell activation states [Mildner et al., 2017, Rico et al., 2017, Saeed et al., 2014, Chen et al., 2019].

1.2 Molecular Determinants of Cell Fate Choices

1.2.1 Lineage-Specific Transcription Factors

In contrast to external or environmental stimuli that include interactions with the stem cell niche or exposure to certain cytokines, lineage-specific transcription factors (TFs) act as the intrinsic determinants, that impact cellular identities and cell fate decisions. TFs belong to various classes of DNA binding proteins and induce the activation of transcriptional programs that lead to the successive specification of progenitor cells [Orkin, 2000]. Studies using gene knockouts in mice were of great importance to provide insights into the functionality of TFs and highlighted that disturbances in TF levels are tightly linked to the development of hematological malignancies [Orkin and Zon, 2008].

In general, factors that are essential for HSC formation or function and factors that specify differentiation into certain lineages can be distinguished. For instance, the activity of Scl, Lmo2, Runx1 or Mll is required for the production of stem cells during embryogenesis or their survival and self-renewal in the adult hematopoietic system [Orkin and Zon, 2008]. The expression of other factors such as Gata1, PU.1, Gata2, Cebpa, Gfi1, Fli1 or Klf1 on the other hand is essential for the production of myeloid, erythroid or megakaryocytic cells [Orkin and Zon, 2008]. The interactions of these TFs are highly complex and often cross-antagonistic, suggesting that the random dominance of one TF within a cell defines the eventual lineage fate [Graf and Enver, 2009, Iwasaki and Akashi, 2007, Arinobu et al., 2007]. Yet, novel reporter mouse lines coupled with single-cell based long-term quantifications of TF levels revealed that a stochastic variation in expression levels is not decisive to initiate lineage choice [Hoppe et al., 2016]. More likely, TFs irreversibly execute and reinforce lineage choices, but other upstream mechanism specify cell fates in the first place [Hoppe et al., 2016]. These could be extracellular signals and interactions with the hematopoietic niche [Pinho and Frenette, 2019], as illustrated by the direct instruction of the myeloid master regulator PU.1 by niche-derived TNF signals

[Etzrodt et al., 2019]. Alternatively, cell fates seem to be predetermined by the cell type-specific epigenetic configuration which furthermore restricts their plasticity in response to certain cues [Yu et al., 2016, Haas et al., 2018].

1.2.2 The Multi-Layered Epigenome

Conceived by Conrad Waddington in the year of 1942, the term 'epigenetics' describes how phenotypic traits and effects or nowadays heritable changes in gene expression and function are impacted by non-genetic factors that are not encoded in the genomic DNA sequence [Waddington, 1942]. Technological advances enabled a profiling of epigenetic patterns both on a genome-wide and on a single-cell level, with epigenetic configurations reflecting cell identities and activation states during differentiation, development, ageing and disease contexts [Stricker et al., 2017, Smith and Meissner, 2013, Feinberg et al., 2016].

DNA Methylation Serves as a Stable Mark for Epigenetic Memory

Within a cell, multiple layers of epigenetic regulatory mechanisms can be distinguished, which jointly determine the epigenetic configuration (Figure 1-3). On the nucleotide level, carbon-5 or nitrogen-3 positions on cytosines as well as nitrogen-6 positions on adenines can be modified, with the cytosine carbon-5 methylation being the most common and best characterized modification [Stricker et al., 2017]. Methylation usually occurs in the context of cytosine-phosphate-guanine (CpG) dinucelotides with around 60-80 % of CpGs (3-8 % of all cytosines) being methylated in mammalian cells, and is established by DNA methyltransferases (DNMTs) which catalyze the transfer of methyl groups (Figure 1-3). [Lister et al., 2009, Okano et al., 1998]. DNMT3 alpha (DNMT3A) and DNMT3 beta (DNMT3B) establish methylation marks *de novo,* whereas DNMT1 maintains methylation marks at newly synthesized DNA strands using a methylated strand as a template [Okano et al., 1998, Okano et al., 1999, Smith and Meissner, 2013]. Being a reversible modification, ten-eleven translocation (TET) methylcytosine dioxygenases catalyze the active demethylation of DNA by oxidizing methylated cytosines via several intermediates to unmethylated cytosines (Figure 1-3) [Tahiliani et al., 2009, Ito et al., 2010, Wu and Zhang, 2017, He et al., 2011].

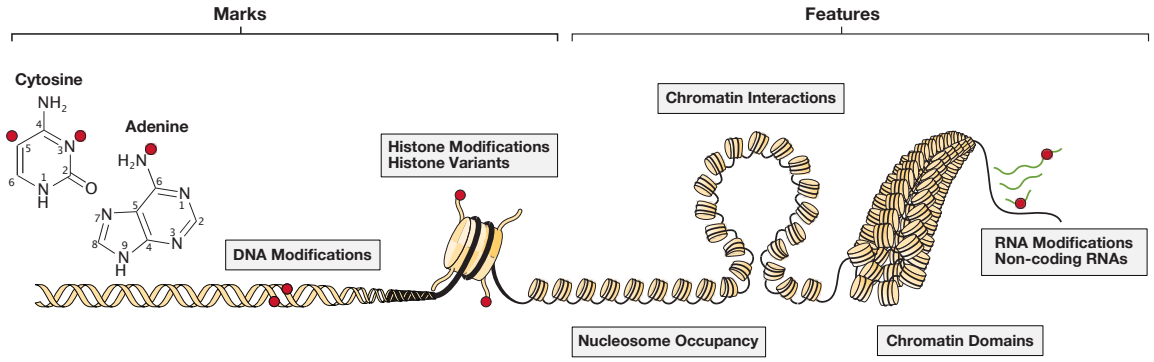
DNA methylation analysis initially focused on CpG islands, regions enriched for CpG-dinucleotides, which overlap with promoter sequences or transcriptional start sites and where methylation levels highly correlate with transcriptional activity (high methylation levels associated with gene silencing, low methylation levels associated with open chromatin, accessibility and active transcription) [Smith and Meissner, 2013, Saxonov et al., 2006, Luo et al., 2018]. Yet, high-resolution genome-wide methylation maps indicated that, while the number of dynamically regulated CpGs is limited (~20 %), they are often located in CpG-poor regions which frequently overlap with regulatory elements (e.g. enhancers or TF binding sites) [Ziller et al., 2013]. Coupling differentially methylated regions (DMRs) with binding of lineage-specific transcription factors thereby provides a framework that determines developmental cell fates and specifies lineage trajectories.

Due to the stability of DNA methylation as an epigenetic modification which is passed on to daughter cells, large-scale genomic patterns function as a molecular fingerprint which enables to define cell states, reconstruct cellular hierarchies or infer lineage specifications [Hodges et al., 2011, Bock et al., 2012, Farlik et al., 2016]. Serving as a mark which stabilizes specific cell states, many cancer entities are characterized by severely disturbed DNA methylation profiles. Yet, these profiles can be harnessed to deconvolute

intratumor heterogeneity or identify the cell of origin for a particular type of neoplasm [Brocks et al., 2014, Lutsik et al., 2017, Oakes et al., 2016, Kulis et al., 2015].

Histone Modifications Define Functional Elements

The compact organization of DNA into nucleosomes which consist of individual histone subunits, allows further epigenetic modifications without a direct modification of DNA nucleotides. Covalent post-translational histone modifications frequently encompass acetylation, methylation and phosphorylation at a large number of specific histone tail sites and by being coupled with the deposition of various histone variants amount to a complex regulatory network [Stricker et al., 2017]. Multiple evolutionary conserved enzymes classified as 'writers', 'erasers' or 'readers' establish or remove histone modifications at certain sites or recognize and bind to specific modifications, respectively (Figure 1-3) [Atlasi and Stunnenberg, 2017]. Charting histone marks across functional elements has led to the association of individual marks with certain activation states, e.g. tri-methylation of histone H3 at the lysine residue 27 (H3K27me3) and H3K9me2 with repression of gene expression, H3K27 acetylation (H3K27ac) with activation of gene expression, or H3K4me1 with priming of enhancers [Zhao et al., 2007, Leeb et al., 2010, Wen et al., 2009, Lienert et al., 2011, Creyghton et al., 2010, Kundu et al., 2000, Heinz et al., 2015].



Modification	Writer	Erasers	Location	Function
DNA methylation	DNMT1, DNMT3A, DNMT3B	TET1, TET2 and TET3	CpG dinucleotides	Silencing and others
H3K27me3	PRC2	UTX1, JMJD3	CpG-rich promoters and intergenic regions	Silencing
H3K9me2	G9A and GLP	JMJD2A, JMJD2B, JMJD2C and JMJD2D; JMJD1A, JMJD1B and JMJD1C	Gene bodies, intergenic regions and enhancers	Silencing
H3K4me3	COMPASS-like proteins (SET1, MLL1-MLL2)	JARID1A, JARID1B, JARID1C and JARID1D; KDM2B	Mainly promoters	Possibly activating
H3K27ac	HATs (including CBP/p300, GNATs and MYSTs)	HDACs and sirtuins	Promoters and enhancers	Activating
H3K4me1	COMPASS-like proteins (MLL3-MLL4)	LSD1 and LSD2	Promoters, enhancers and intergeic regions	Priming and / or activating

Figure 1-3: Epigenetic Layers and Modifications.
Several layers of epigenetic marks and features make up the epigenomic profile of a cell. Epigenome dynamics are a characteristic feature during cell state transitions, differentiation and development. Specific enzymes, termed 'writers' and 'erasers', establish or remove particular modifications and therefore are critical regulators of cellular plasticity (Figure adapted from [Stricker et al., 2017] and [Atlasi and Stunnenberg, 2017]).

The Three-Dimensional Architecture of Chromatin

Histone marks furthermore influence the nucleosome occupancy of DNA and specific chromatin interactions. Opening up of compact chromatin and removal of nucleosome occupancy is mediated by pioneering factors dependent on existing histone marks and is critical to increase DNA accessibility for the recruitment of additional factors to activate regulatory regions as enhancers or promoters [Atlasi and Stunnenberg, 2017]. Furthermore, long-range spatial chromatin interactions between promoter and enhancer elements as well as the compartmentalization of the genome into larger topologically associating domains (TADs), which facilitate interactions, provide the dynamic three-dimensional structural basis for the regulation of gene expression [Zheng and Xie, 2019, Atlasi and Stunnenberg, 2017].

Eventually, it is important to note that DNA methylation, histone marks, and higher-order structures do not govern gene expression on a solitary basis, but rather a complex interplay and crosstalk of the individual epigenetic layers is required to enforce correct transcriptomic profiles.

Epigenome Dynamics During Hematopoietic Differentiation

Early array-based technologies employed to investigate methylation changes in the hematopoietic system revealed that lineage-specific differentiation involves a modulation of DNA methylation marks [Ji et al., 2010, Hodges et al., 2011]. DMRs furthermore were enriched for lineage-specific TF binding sites, thereby correlating DNA methylation to gene expression levels and leading to the identification of regulatory programs implicated in the coordination of self-renewal and differentiation cell fates [Ji et al., 2010, Hodges et al., 2011]. First DNA methylation maps covering the whole hematopoietic hierarchy not only illustrated the inference of cellular lineage hierarchies based on methylation profiles, but also that cell type-specific methylation patterns are conserved throughout trajectories, starting from HSCs and ending in mature differentiated cells [Bock et al., 2012]. Characteristic methylation signatures were proposed to act as gatekeepers that, by obstructing aberrant TF binding, instruct and reinforce specified lineage fates throughout developmental trajectories [Bock et al., 2012]. Genome-wide single-base resolution methylation maps of the HSPC compartment further corroborated the directive role of DNA methylation at the earliest stages of hematopoietic commitment [Cabezas-Wallscheid et al., 2014, Lipka et al., 2014]. Unidirectional and progressive changes in DNA methylation were already detected during the transition from LT-HSCs to ST-HSCs and individual MPP subsets in particular at *cis*-acting regulatory regions [Cabezas-Wallscheid et al., 2014, Lipka et al., 2014]. These were enriched in TF binding sites that instruct lineage- and differentiation stage-specific transcriptional programs (e.g. Gata1, Gata2, Lmo2, Runx1 or Pbx1) [Cabezas-Wallscheid et al., 2014, Lipka et al., 2014], indicating that DNA methylation drives or at least mirrors the decisive activity of regulatory elements during hematopoietic commitment.

Assisting DNA methylation maps, genome-wide maps of histone modifications and chromatin accessibility were constructed to chart the regulatory landscape during hematopoietic differentiation [Lara-Astiaso et al., 2014, Buenrostro et al., 2018, Corces et al., 2016]. The establishment of lineage-specific enhancer repertoires precedes both lineage commitment and the activation of transcriptional programs characteristic for certain differentiated cell types [Lara-Astiaso et al., 2014]. Likewise, chromatin accessibility maps allowed the reconstruction of lineage-specific differentiation trajectories with dynamic TF activity paralleling the accessibility of chromatin around the TF binding motif [Buenrostro et al., 2018]. Eventually, analysis of the three-dimensional structure of the genome revealed long-range interactions between active promoters and

epigenetically marked active enhancers to be cell type-specific and to reflect lineage relationships between hematopoietic cell types [Javierre et al., 2016].

Unanimously, the epigenetic fingerprint of hematopoietic cells is a pivotal regulatory mechanism that shapes fundamental cellular characteristics by determining developmental fates, lineage decisions, self-renewal, quiescence or functional behavior in response to environmental stimuli. Since the integrity of the epigenome ensures the functionality of the hematopoietic system, aberrant epigenetic profiles are linked to the age-dependent decline in HSC functionality [Beerman et al., 2013, Sun et al., 2014a], as well as the development of hematological malignancies [Genovese et al., 2014, Jaiswal et al., 2014, Steensma et al., 2015, Papaemmanuil et al., 2016].

1.3 Mutational Landscapes in Preleukemia and Hematological Malignancies

1.3.1 Clonality of the Hematopoietic System

In the hematopoietic system, blood production is sustained by the activity of thousands of HSCs, which are heterogeneous and display a stereotypic clonal behavior based on their epigenetic configuration [Yu et al., 2016]. During the lifespan of an individual, the expansion of several clones will result in a clonal mosaicism where particular clones continuously contribute more to blood production than others. Clonal expansions were initially identified by investigating cytogenetic patterns in patients with chronic myeloid leukemia (CML), thereby establishing a clonal origin of hematological malignancies [Fialkow et al., 1967, Rowley, 1973]. However, only several years later a correlation between clonal hematopoiesis and the age of individuals was discovered by studying skewed X-chromosome inactivation patterns in women of different age [Fey et al., 1994, Busque et al., 1996]. Although these studies described clonal behavior within the hematopoietic system, only advances in whole-genome sequencing methods enabled to associate genomic backgrounds and specific genetic lesions with a clonal expansion.

The increasing number of somatic mutations per cell is a hallmark of aging [Lopez-Otin et al., 2013]. Environmental exposure, error-prone repair of DNA single- or double-stranded breaks, spontaneous deamination of methylcytosines, replicative stress and excessive proliferation / activation are the main factors giving rise to the accumulation of mutations over time [Jaiswal and Libby, 2019, Flach et al., 2014, Walter et al., 2015]. In humans, 50,000-200,000 HSCs actively contribute to blood production with one cell division occurring around every 25-50 weeks [Catlin et al., 2011, Lee-Six et al., 2018]. Although HSCs usually reside in a quiescent state, each cell acquires around 17 mutations in their whole genome and 0.13 ± 0.02 mutations in exons per year of life [Welch et al., 2012]. While most of these mutations will not affect the functionality or the fitness of a cell, in rare instances a mutation will convey a selective advantage, resulting in an expansion of the clone at disproportionate rates compared to other HSC clones [Jaiswal and Libby, 2019].

Whole-exome sequencing studies investigating the mutational frequency in blood samples from individuals unselected for any hematological phenotypes detected the presence of clonal hematopoiesis in combination with somatic driver mutations in a large number of otherwise healthy people [Jaiswal et al., 2014, Genovese et al., 2014, Xie et al.,

2014]. This phenomenon, termed clonal hematopoiesis of indeterminate potential (CHIP), furthermore strongly correlated with age, with detectable somatic mutations being found only rarely in individuals younger than 40 years, but frequencies steadily increasing with age (18.4 % in individuals ≥ 90 years) [Jaiswal et al., 2014]. Accordingly, the emergence of clonal hematopoiesis strongly depends on increasing aging-associated selective pressures, but also an immune-mediated clonal selection or chemotherapy-associated selective pressures were shown to promote a reduced clonality over time [Bowman et al., 2018, Coombs et al., 2017].

1.3.2 Mutational Acquisition, Cooperativity and Inflammatory Conditions Define the Evolutionary Path of Disease Development

The most recurrent mutations identified in individuals with CHIP encompass genes which are also commonly mutated in myeloid cancers, such as *DNMT3A, TET2, ASXL1, TP53, JAK2* or *SF3B1*, and thereby have linked the presence of CHIP to the development of subsequent hematological malignancies, and have led to the connotation of CHIP as a preleukemic state. [Sperling et al., 2017]. The presence of somatic mutations was associated with a ten-fold increased risk of developing hematological cancers, especially myeloid malignancies such as acute myeloid leukemia (AML) or myelodysplastic syndromes (MDS), but also with an increased overall mortality and an increased risk for cardiovascular disease [Jaiswal et al., 2014, Jaiswal et al., 2017]. Although clonal hematopoiesis has also been identified in the absence of any known driver mutations [Zink et al., 2017], CHIP carriers usually display only one driver mutation, whereas patients with overt malignancies display multiple [Jaiswal and Libby, 2019]. While a clonal expansion of mutated clones is a common characteristic of CHIP-defined preleukemic states, only after the successive acquisition of additional mutations, disease progression and malignant transformation is initiated (Figure 1-4A) [Sperling et al., 2017]. Here, different driver mutations associated with CHIP relay varying risks of eventually developing overt myeloid malignancies, thereby highlighting that specific mutations are associated with a higher or lower oncogenic potential (Figure 1-4B) [Abelson et al., 2018, Desai et al., 2018].

In CHIP, driver mutations frequently affect direct or indirect epigenetic regulators which coordinate patterns of DNA methylation (DNMT3A, TET2, IDH1/2) or histone modifications (ASXL1), suggesting that disturbed epigenetic profiles are causative for disease initiation in many cases [Bowman et al., 2018]. Likewise, more than 70 % of patients diagnosed with AML carry at least one somatic mutation in a gene involved in regulating epigenetic patterns within a cell (Figure 1-5A and Figure 1-5B) [Cancer Genome Atlas Research et al., 2013, Papaemmanuil et al., 2016, Chen et al., 2013b]. On an evolutionary timeline, the acquisition of mutations in epigenetic regulators usually precedes the acquisition of later genetic lesions affecting transcription factors or activated signaling pathways, however, malignancy progression is frequently promoted through cooperative effects between early and late arising mutations [Corces-Zimmerman et al., 2014, Guryanova et al., 2016, Loberg et al., 2019, Kameda et al., 2015]. Importantly, these early mutations usually occur in HSCs, and HSC clones with preleukemic mutations were detected in remission samples, illustrating that these clones contribute to therapy resistance and function as a reservoir from which relapse evolves [Shlush et al., 2014, Corces-Zimmerman et al., 2014].

Since the evolutionary path from preleukemia to overt hematological malignancies is shaped by the co-occurrence, cooperativity and mutual exclusivity of certain mutations (Figure 1-5C), varying combinations of genetic lesions will impact phenotype, prognosis and therapy response [Sperling et al., 2017, Papaemmanuil et al., 2016]. As preleukemic mutations alone rarely result in the progression towards an oncogenic transformation, combinatorial mouse models employed to investigate an interplay between individual mutations are critical to understand their role in disease development. Ideally, therapeutic strategies tailored to a specific mutational background then can be applied to target CHIP, diminish the selective advantage of mutated clones, and prevent a progression to hematological malignancies [Bowman et al., 2018].

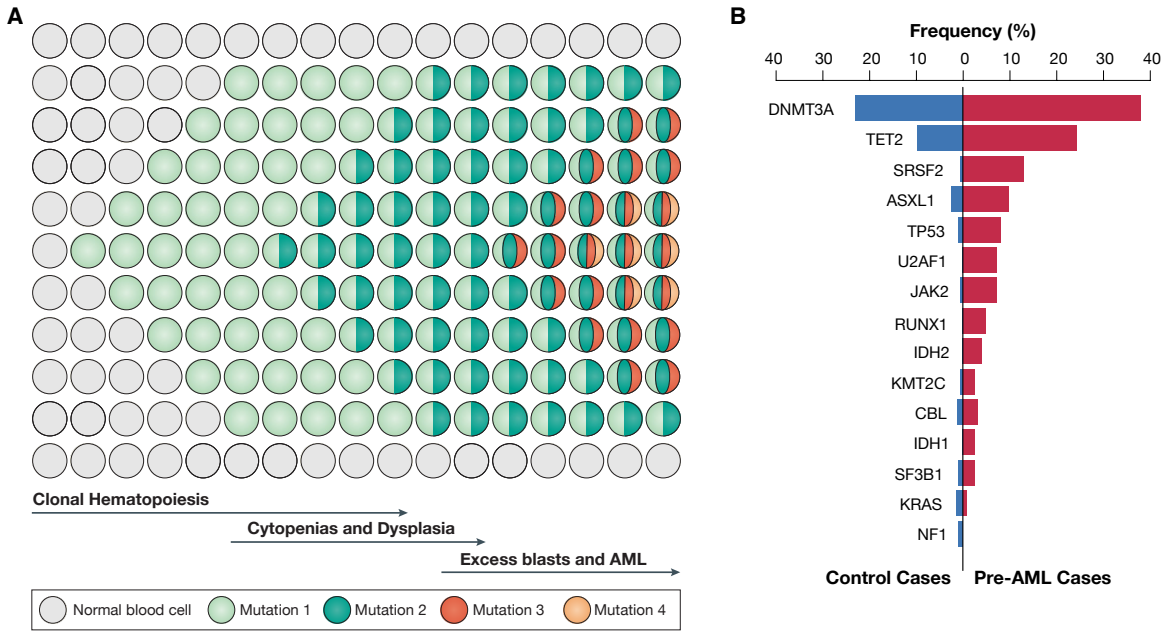


Figure 1-4: Mutation-Driven Clonal Expansion and Acute Myeloid Leukemia Risk.
(A) The development of AML is defined by the sequential acquisition of mutations. Early mutations frequently affect direct or indirect epigenetic regulators, such as DNMT3A, TET2, IDH1/2 or ASXL1 and result in a clonal expansion of the respective clone. The accumulation of additional secondary or tertiary mutations further promotes leukemic progression and the emergence of overt malignancies. Late-occurring mutations predominantly affect genes involved in cellular signaling, such as FLT3 or NRAS, and in co-operation with earlier acquired preleukemic mutations shape the evolutionary path of disease development (Figure adapted from [Sperling et al., 2017]). **(B)** Mutations in AML-associated driver genes are commonly found in healthy individuals with clonal hematopoiesis before diagnosis of any hematological malignancies. Distinct mutations confer diverging risks of future AML development, as indicated by the relative frequency of mutations in pre-AML cases (red, individuals developed AML) or in control cases (blue, individuals did not develop AML) (Figure adapted from [Abelson et al., 2018]).

Recent studies additionally linked the presence of clonal hematopoiesis and preleukemic mutations to inflammatory conditions that result in an increased risk for cardiovascular disease [Jaiswal et al., 2017, Fuster et al., 2017]. Beside catalyzing the demethylation of 5-methylcytosines, Tet2 is also required to resolve inflammation in macrophages as a loss of Tet2 (e.g. due to loss-of-function mutations) was associated with an upregulation of several proinflammatory cytokines which function as proatherogenic mediators [Zhang et al., 2015, Fuster et al., 2017, Jaiswal et al., 2017]. The proinflammatory state could be directly linked to the development of a preleukemic clonal expansion [Meisel et al., 2018], thereby establishing the importance of inflammation as a critical factor that contributes to preleukemic clonal hematopoiesis and disease progression.

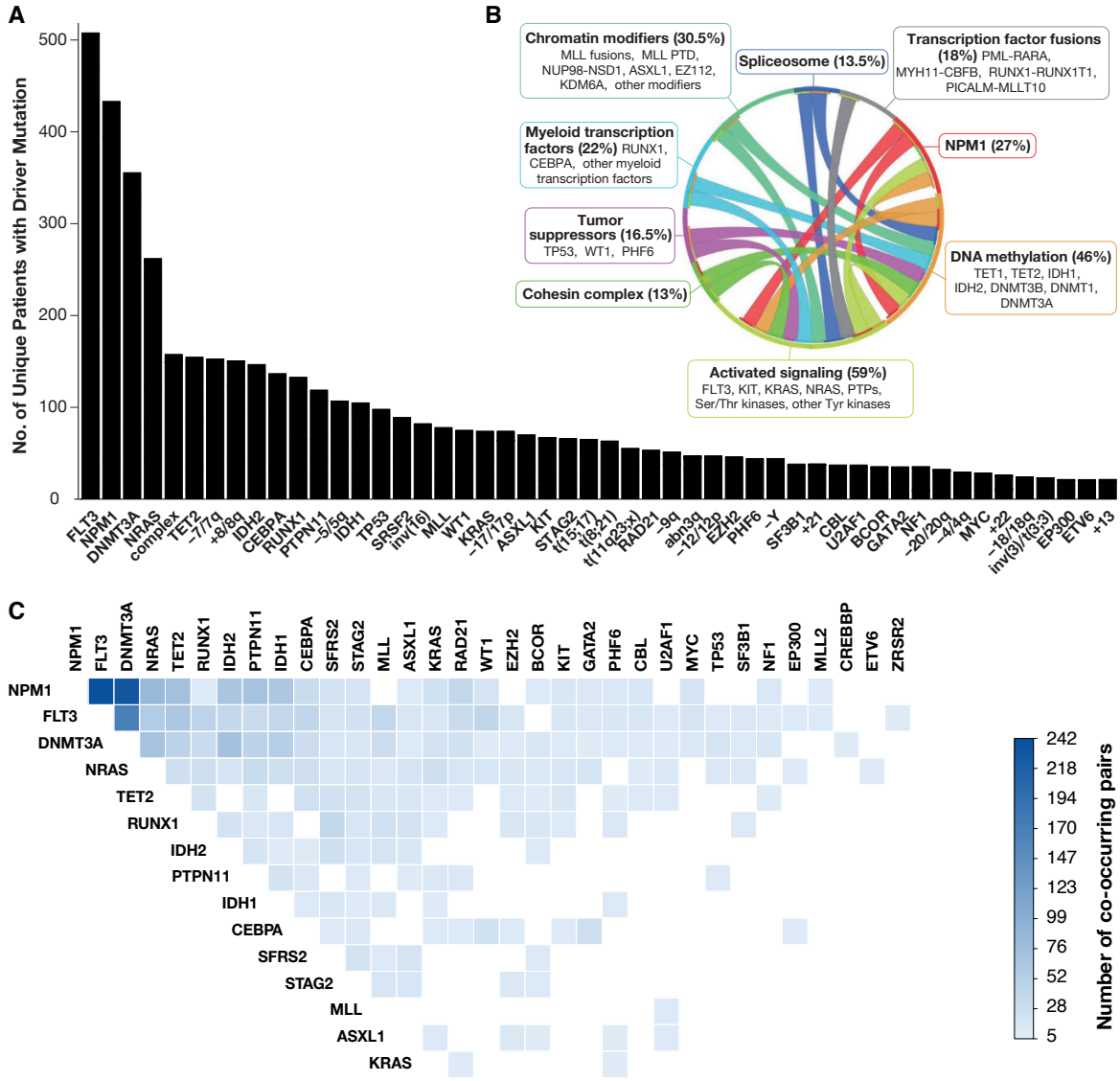


Figure 1-5: Landscape and Co-Occurrence of Driver Mutations in Acute Myeloid Leukemia Patients.
(A) Most frequently detected driver lesions in a cohort of 1540 AML patients, including gene mutations and fusions, chromosomal aneuploidies and complex karyotypes. Direct or indirect Epigenetic regulators, such as DNMT3A, TET2 or IDH1/2 are recurrently mutated (Figure adapted from [Papaemmanuil et al., 2016]). **(B)** Categories of genetic lesions detected in AML patients. More than 70 % of patents carry at least one mutation affecting an epigenetic regulator (involved in chromatin modifications or DNA methylation) (Figure adapted from [Chen et al., 2013b]). **(C)** Recurrent pairs of mutations co-occurring in a cohort of 1540 AML patients (Data from [Papaemmanuil et al., 2016]).

1.4 *IDH1*-R132H Mutations Lead to Metabolic Imbalances

1.4.1 *IDH1/2* Mutations in CHIP and AML

Mutations in isocitrate dehydrogenase 1 or 2 (*IDH1 / IDH2*) initially were only rarely found in individuals with CHIP, arguing against a preleukemic nature of these mutations [Jaiswal

et al., 2014, Genovese et al., 2014, Xie et al., 2014]. Yet, large cohort risk stratification studies revealed not only that *IDH1/2* mutations occur in healthy individuals several years before AML diagnosis, but also classified them among mutations associated with the highest risk for future AML development, with nearly all subjects carrying *IDH1/2* mutations eventually progressing to AML [Abelson et al., 2018, Desai et al., 2018]. Despite the relatively low frequency compared to other CHIP-associated mutations such as *DNMT3A* or *TET2* mutations, *IDH1/2* mutations accordingly impose a strong risk for oncogenic transformation and are recurrently detected in around 15-20 % of patients diagnosed with AML [Cancer Genome Atlas Research et al., 2013, Papaemmanuil et al., 2016]. In AML patients, *IDH1/2* mutations frequently have been detected at high variant allele frequencies in HSCs, implying that these mutations are acquired at the first stages of disease evolution and function as initiating leukemogenic event [Corces-Zimmerman et al., 2014].

1.4.2 *IDH1*-R132H Mutations Produce the 'Oncometabolite' D-2-Hydroxyglutarate

As part of the tricarboxylic acid (TCA) cycle, homodimeric IDH1/2 enzymes catalyze the conversion of isocitrate to α-ketoglutarate (αKG) while concomitantly producing reduced NADPH and CO_2 (Figure 1-6) [Cairns and Mak, 2013]. Although both enzymes catalyze the identical reaction, they differ in their cellular location, with IDH1 being localized the cytoplasm and in peroxisomes whereas IDH2 resides in the mitochondrial matrix [Cairns and Mak, 2013]. Recurrent cancer-associated *IDH1* missense mutations almost exclusively affect the arginine residue at position 132 (R132) which is required for substrate binding and is most frequently substituted by histidine (R132H mutation) in a mutated context. [Mardis et al., 2009, Parsons et al., 2008, Yang et al., 2012]. Analogously, *IDH2* mutations predominantly affect arginine residues R172 or R140, which likewise are involved in specifying substrate binding [Mardis et al., 2009, Parsons et al., 2008, Yang et al., 2012].

Due to its critical function during substrate binding, mutation of the R132 amino acid alters substrate specificity and abrogates αKG production [Dang et al., 2009, Ward et al., 2010]. Instead, *IDH1*-R132H mutations instigate a gain-of-function and result in a neomorphic enzymatic activity, catalyzing the conversion of αKG and NADPH to the D-enantiomer of 2-hydroxyglutarate (D2HG) and NADP+ [Dang et al., 2009, Ward et al., 2010]. The *IDH1*-R132H mutation-mediated formation of D2HG leads to an excessive accumulation within the cell and secretion into surrounding tissues, so that elevated D2HG levels for instance can be detected in brain tumors of glioma patients and blood sera of AML patients harboring *IDH1*-R132H mutations [Andronesi et al., 2012, Choi et al., 2012, Gross et al., 2010, DiNardo et al., 2013]. D2HG accumulation therefore serves as a predictive marker to infer an *IDH1*-R132H mutation status, as a diagnostic marker to detect *IDH1*-R132H mutated malignancies and as a prognostic marker to assess clinical outcomes [DiNardo et al., 2013]. Since oncogenic effects of *IDH1*-R132H mutations are dependent on the excessive production of D2HG, several pharmacological inhibitors, which allosterically target mutated IDH1 in order to reduce D2HG production, have been developed and are under clinical investigation [Golub et al., 2019].

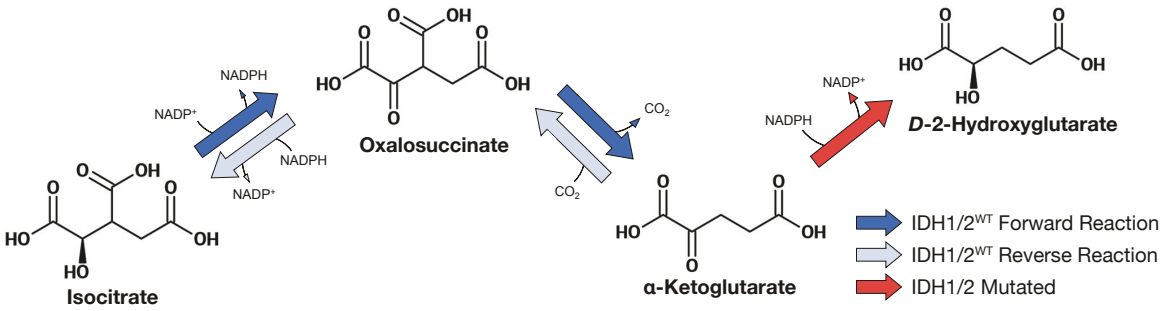


Figure 1-6: Mutated IDH1/2 Converts α-Ketoglutarate to D-2-Hydroxyglutarate.
IDH1/2 wildtype enzymes catalyze the production of α-ketoglutarate from isocitrate. However, recurrent *IDH1*-R132H, *IDH2*-R140Q and *IDH2*-R172K mutations lead to alterations in substrate specificity and a loss of regular catalytic activity, and ultimately result in mutated IDH1/2 enzymes gaining the function to convert α-ketoglutarate to D-2-hydroxyglutrate (Figure adapted from [Molenaar et al., 2018a]).

1.4.3 Excessive D2HG Levels Inhibit DNA and Histone Demethylases

Based on the structural similarity of αKG and D2HG, oncogenic functions of D2HG are mediated by acting as an inhibitor for αKG-dependent dioxygenases, including enzymes of the TET family (TET1, TET2, TET3) or histone demethylases of the Jumonji class (Figure 1-7) [Xu et al., 2011, Chowdhury et al., 2011, Lu et al., 2012]. TET enzymes are critical enzymes involved in the active demethylation of 5-methylcytosines (5-mC) by catalyzing the hydroxylation of 5-mC to 5-hydroxymethylcytosine (5-hmC), 5-formylcytosine (5-fC), and 5-carboxycytosine (5-caC) [Ito et al., 2010, Ito et al., 2011]. Owing to the inhibition of TETs, *IDH1*-R132H mutations consequently were associated with a hypermethylation or a CpG island methylator phenotype [Figueroa et al., 2010, Noushmehr et al., 2010, Turcan et al., 2012], which is reversed upon treatment with IDH1- or DNMT inhibitors [Borodovsky et al., 2013, Turcan et al., 2013, Kernytsky et al., 2015]. Interestingly, *IDH1*-R132H and *TET2* mutations usually are mutually exclusive in AML patients and aberrant DNA methylation patterns in *IDH1*-R132H mutated cells are phenocopied by *TET2* knockouts [Figueroa et al., 2010], illustrating that a D2HG-mediated inhibition of TETs is a critical mechanism contributing to oncogenic transformation.

Likewise, inhibition of Jumonji C-domain lysine demethylases (KDM2A, KDM4A, KDM4C, KDM4D, KDM5B) by D2HG results in the accumulation of methylated lysine residues, in particular of tri-methylated H3K9 and tri-methylated H3K27 [Lu et al., 2012, Chowdhury et al., 2011]. *IDH1*-R132H mutations furthermore affect chromatin structures and loops by compromising binding of the insulator protein CTCF [Flavahan et al., 2016], which is required to form boundaries between TADs [Rao et al., 2014]. Mutation-dependent DNA hypermethylation interferes with binding patterns and consequently results in deregulated enhancer-promoter interactions or deregulated enhancer activations [Flavahan et al., 2016]. Collectively, disturbed patterns of epigenetic modifications induced by *IDH1*-R132H mutations are key mechanisms that cause alterations in epigenetic-based regulatory networks and thereby drive alterations in gene expression levels.

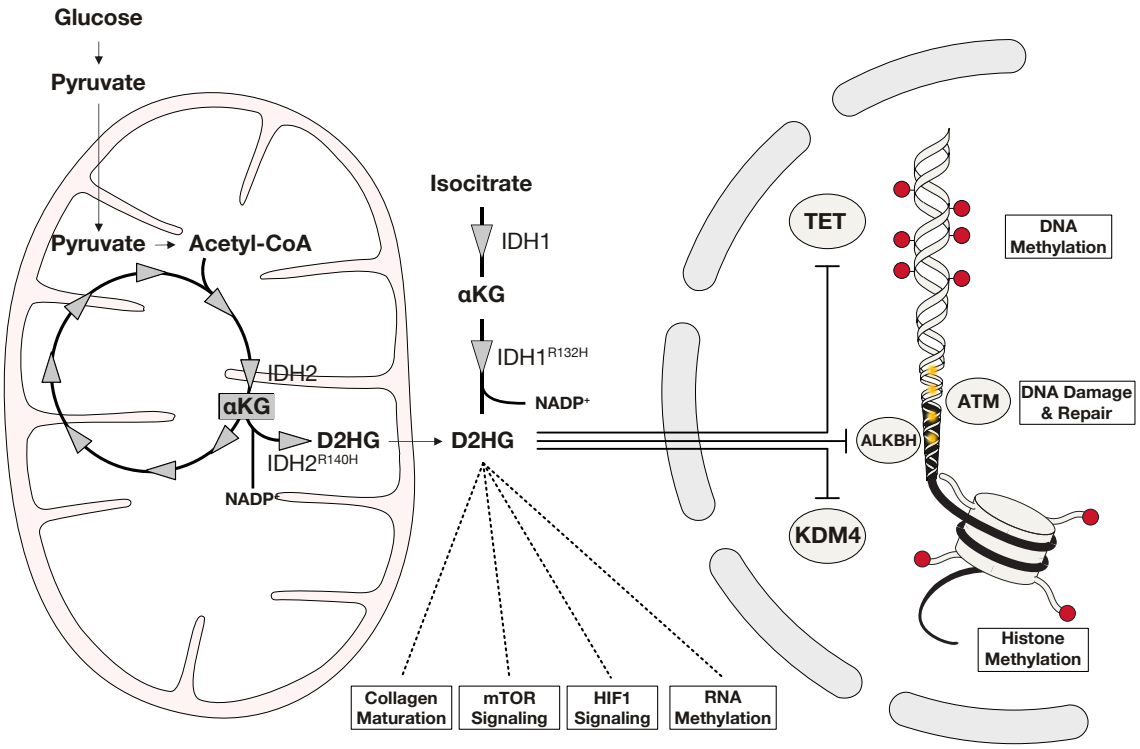
1.4.4 *IDH1*-R132H Mutations Alter DNA Damage and Repair Kinetics

In addition to a deregulated epigenetic landscape, *IDH1*-R132H mutations further affect DNA damage response and DNA repair signaling pathways. Reactive oxygen species (ROS) are a primary source of oxidative damage to DNA and therefore need to be detoxified through several NADPH-dependent mechanisms [Gagne et al., 2017]. The consumption of NADPH by mutated IDH1 consequently will result in diminished levels of damage-preventing antioxidants within the cell, and simultaneously will result in elevated levels of DNA damaging agents and a state of oxidative stress [Bleeker et al., 2010, Shi et al., 2015, Shi et al., 2014].

In response to DNA damage, the kinase ataxia-telangiectasia mutated (ATM) functions as a critical regulator of downstream signaling pathways leading to the recruitment of repair factors to the site of damage [Shiloh and Ziv, 2013]. Repression of ATM levels by D2HG, mediated through the inhibition of KDM4 lysine demethylases, was shown to result in impaired DNA damage signaling and repair kinetics [Inoue et al., 2016]. Similarly, inhibition of lysine demethylases KDM4A/B, which are involved in orchestrating DNA damage response pathways, and αKG-dependent repair enzymes of the ALKB family was shown to promote the accumulation of DNA damage in *IDH1*-R132H mutated cells (Figure 1-7) [Mallette et al., 2012, Young et al., 2013, Sulkowski et al., 2017, Wang et al., 2015].

Overall, the down-regulation of DNA damage sensing and repair pathways enables *IDH1*-R132H mutated cells to evade tumor suppression, while at the same time promoting genomic instability, random mutagenesis and the acquisition of additional driver mutations. On the other hand, these characteristics can be leveraged in treatment settings, where *IDH1*-R132H mutated cells displayed a higher sensitivity to poly (ADP-ribose) polymerase (PARP) inhibitors, genotoxic chemotherapeutics such as alkylating agents, or radiation [Sulkowski et al., 2017, Wang et al., 2015, Li et al., 2013].

Furthermore, increased D2HG levels have been highlighted to interfere with mTOR signaling, HIF1 activity, RNA methylation and collagen maturation through the inhibition of various αKG-dependent enzymes [Gagne et al., 2017], thereby emphasizing the complexity and diversity of oncogenic effects arising from an *IDH1*-R132H mutation-induced metabolic dysbalance.



Cellular Processes Influenced by *IDH1/2* Mutations through Inhibition of Enzymatic Activity

Process	Affected Enzymes	Consequences
DNA methylation	TET1, TET2	Decreased hydroxylation of 5-methylcytosine to 5-hydroxymethylcytosine, DNA hypermethylation, changes in gene expression
Histone methylation	KDM2A, KDM4A, KDM4B, KDM4C, KDM4D, KDM5B	Global increase in histone methylation (especially H3K9me2/3), changes in gene expression
DNA damage response and repair	ALKBH2, ALKBH3,	Impaired DNA repair of methylated DNA bases, accumulation of DNA damage, genomic instability
DNA damage response and repair	KDM4A, KDM4B	Down-regulation of DNA damage sensor ATM, impaired DNA repair by homologous recombination, accumulation of DNA damage and genomic instability
Collagen maturation	P4HA1/2, PLOD1/3	Defects in collagen maturation as well as basal membrane structure and function
mTOR signaling	KDM4A, ATP5B	Modulation of mTOR signaling and activity, deregulation of cell growth, proliferation and survival signals
HIF1 signaling	HIF1AN / FIH, EGLN1/2	Altered HIF-1α stability and activity, deregulation of oxygen homeostasis and response to hypoxia
RNA methylation	FTO	Decreased demethylation of RNAs, accumulation of N^6-methyladenosine modified RNAs, deregulation of RNA splicing, transport, stability and translation

Additional Effects on Cellular Homeostasis Mediated by *IDH1/2* Mutations

Effect	Consequences
Increased NADP+ Production and NADPH consumption	Dysbalanace of reducing and oxidizing agents, increased sensitivity to oxidative stress, decreased detoxification of reactive oxygen species (ROS), increased concentration of DNA damaging agents

Figure 1-7: Effects of *IDH1/2* Mutations on Cellular Processes.
Overview of cellular processes that are dysregulated upon the expression of an *IDH1* or *IDH2* mutation and the excessive production and accumulation of D2HG. The dysregulation of cellular processes is either mediated through an modification of enzymatic activities (enzymes inhibited through D2HG), or imbalances in NADP+/NADPH ratios (Figure adapted from [Molenaar et al., 2018a] and [Gagne et al., 2017]).

1.4.5 *IDH1*-R132H Mutation-Mediated Effects on Hematopoietic Stem and Progenitor Cell Function

Within the hematopoietic system, investigations have tried to segregate effects determined through the inhibition of TET enzymes and effects determined through the impairment of other pathways. In AML patients, *IDH1*-R132H mutations were associated with similar aberrant epigenetic patterns as TET2 loss-of-function mutations, featuring a global DNA hypermethylation signature [Figueroa et al., 2010]. Both mutations furthermore resulted in defects in hematopoietic differentiation, in particular of the myeloid lineage, and an increased expression of stem cell markers [Figueroa et al., 2010]. These findings were recapitulated in *Tet2* knockout mouse models, where Tet2 loss induced enhanced stem cell self-renewal, an enlargement of the HSPC compartment and skewed myeloid differentiation patterns [Moran-Crusio et al., 2011]. Equally, *Idh1*-R132H mutant mice featured enhanced numbers of early progenitors arising from changes in DNA methylation signatures which affect signaling pathways involved in hematopoietic cell proliferation, differentiation and leukemogenesis [Sasaki et al., 2012]. Yet, a direct comparison of both mouse models revealed that effects of *Idh1*-R132H mutations are cell type dependent. In HSCs, the expression of the *Idh1*-R132H mutation predominantly leads to altered histone methylation patterns, that result in the down-regulation of ATM and impaired DNA repair independent of Tet2 [Inoue et al., 2016]. On the other hand, in MPPs or myeloid progenitor populations, defects are driven by the inhibition of Tet2 and resemble defects observed in *Tet2* knockout mice [Inoue et al., 2016, Moran-Crusio et al., 2011].

Despite the overlapping DNA hypermethylation signature and mutual exclusivity, the occurrence of *IDH1/2* mutations and *TET2* mutations in hematological disorders varies substantially. Compared to *IDH1/2* mutations, *TET2* mutations are more frequently observed in CHIP, chronic myelomonocytic leukemia (CMML), MDS or myeloproliferative neoplasms (MPNs), whereas combined *IDH1/2* mutation frequencies are higher in AML and associated with a decreased survival when found in MDS [Shih et al., 2012, Thol et al., 2010]. These findings have led to the hypothesis that *IDH1* mutations are linked to more acute and aggressive diseases and confer a higher oncogenic potential than *TET2* mutations. However, this association could only partly be recapitulated in mutant mouse models. Underlining varying clinical features, *Idh1*-R132H mutated mice succumb prematurely due to anemia, splenomegaly and decreased bone marrow cellularity only after long latencies, while Tet2 loss results in a more rapid and more aggressive development of myeloid neoplasms reminiscent of CMML [Inoue et al., 2016, Moran-Crusio et al., 2011, An et al., 2015].

Due to the low transforming potential of *Idh1*-R132H mutations in mice, studies have begun to investigate the interplay of *Idh1/2* mutations with further secondary activating mutations in *Npm1, Flt3* or *Nras* [Ogawara et al., 2015, Chen et al., 2013a, Kats et al., 2014]. A co-occurrence of these mutations clearly accelerated the onset and progression of myeloproliferative diseases, thereby underpinning the preleukemic characteristic of an *Idh1*-R132H mutation alone and highlighting that the course of disease highly depends on and is determined by secondary events and their cooperation [Ogawara et al., 2015, Chen et al., 2013a, Kats et al., 2014].

1.5 *DNMT3A*-R882H Mutations Deregulate DNA Methylation Landscapes

1.5.1 DNMT3A Activity Balances HSC Activity

Early work focusing on the role of Dnmt1 within the hematopoietic system illustrated that epigenetic DNA methylation patterns maintained by Dnmts are essential for the functionality, homeostasis, self-renewal and differentiation of HSCs [Broske et al., 2009, Trowbridge et al., 2009]. In line, loss of Dnmt3a led to impaired HSC differentiation and increased HSC frequencies within the bone marrow due to the upregulation of HSC multipotency genes. [Challen et al., 2011]. Although some compensatory effects between Dnmt3a and Dnmt3b were observed, evidence suggests that Dnmt3a is the main driver of DNA methylation in HSCs in order to regulate HSC maintenance [Challen et al., 2014]. While it was initially assumed, that a loss of Dnmt3a results in a global hypomethylation phenotype, *Dnmt3a* knockout HSCs displayed DMRs defined by both hyper- and hypomethylation [Challen et al., 2011]. Conserved stretches of lowly methylated regions, termed methylation canyons, are defined by highly methylated canyon boundaries maintained by Dnmt3a [Jeong et al., 2014]. Canyon-associated genes, encompassing developmental TFs or HSC regulatory genes, become deregulated upon the erosion of boundaries through Dnmt3a loss, thereby leading to a contraction or expansion of loci-specific canyons [Jeong et al., 2014, Jeong et al., 2018]. Yet, more complex regulatory networks, including activating and repressing histone marks as well as loci-specific competition and cooperation between Dnmt3a and Tet2 balance HSC self-renewal and differentiation through the regulation of lineage-specific TFs and self-renewal genes [Jeong et al., 2014, Zhang et al., 2016]. Accordingly, double-knockout of *Dnmt3a* and *Tet2* promoted an accelerated development of malignancy and the induction of myeloid or lymphoid disease [Zhang et al., 2016]. Loss of Dnmt3a alone led to an immortalization of HSCs with indefinite self-renewal potential in serial rounds of transplantations, mediated through aberrant methylation at canyon-associated HSC regulatory elements [Jeong et al., 2018]. However, despite favoring HSC self-renewal at the expanse of differentiation, mice transplanted with *Dnmt3a* deleted HSCs eventually developed myeloid or lymphoid malignancies such as MDS, AML or T-cell acute lymphocytic leukemia (T-ALL) [Celik et al., 2015, Mayle et al., 2015], suggesting that a *Dnmt3a* knockout-mediated block in differentiation can be overcome under certain conditions or upon acquisition of cooperating mutations.

1.5.2 *DNMT3A*-R882H Mutations Affect Methyltransferase Activity and Result in a Dominant-Negative Phenotype

Research investigating the role of DNMT3A in the hematopoietic system was sparked through the detection of a wide range of nonsense, missense and frameshift mutations in AML patients [Ley et al., 2010]. Since then, *DNMT3A* mutations have been identified in a wide range of hematological malignancies, including MDS, MPN, or T-ALL [Roller et al., 2013, Celik et al., 2016], but also in individuals with CHIP, where *DNMT3A* is the most frequently mutated gene [Jaiswal et al., 2014, Genovese et al., 2014, Xie et al., 2014]. Mutations were detected with a high variant allele frequency in HSCs and result in a

preleukemic state characterized by a clonal expansion of mutated HSCs, thereby providing the basis for disease progression and remission [Shlush et al., 2014, Corces-Zimmerman et al., 2014]. The majority of *DNMT3A* mutations were confined to the methyltransferase domain and in particular enriched for missense mutations affecting the arginine residue at position 882 (R882), which in most cases is substituted by a histidine residue (R882H mutation) [Ley et al., 2010, Roller et al., 2013, Celik et al., 2016]. *DNMT3A*-R882H mutations furthermore were associated with a strong loss in methylation activity and mechanistically could be traced back to a blocked ability to form active homo-tetramers which catalyze the transfer of methyl groups to cytosines [Holz-Schietinger et al., 2012]. Importantly, R882H-mutated DNMT3A also inhibits wildtype DNMT3A activity by equally disrupting the homo-tetramer formation of wildtype enzymes in a dominant-negative fashion [Russler-Germain et al., 2014]. As a consequence, *DNMT3A*-R882H mutated AML samples are characterized by a uniform focal hypomethylation across all genomic regions (promoter, gene bodies, intergenic regions, etc.) [Russler-Germain et al., 2014]. Although CpG island hypermethylation was detected to a certain degree in mutant samples, the *DNMT3A*-R882H mutation-mediated hypomethylation is thought to function as initiating event, whereas observed hypermethylation patterns arise as a secondary effect during AML progression [Spencer et al., 2017].

In *Dnmt3a*-R878H (mouse homologue to the *DNMT3A*-R882H mutation) mutated mice, the heterozygous expression of the *Dnmt3a*-R878H mutation caused an expansion of stem and progenitor cells similar to phenotypes observed in *Dnmt3a* knockout mice, indicating that the mutation alone is sufficient to promote clonal expansion of HSCs [Guryanova et al., 2016, Loberg et al., 2019]. Yet, only the successive acquisition of *Npm1* or *Flt3* mutations resulted in the transformation of *Dnmt3a*-R878H mutated cells and the progression to AML- and MPN-like phenotypes [Guryanova et al., 2016, Loberg et al., 2019], demonstrating that a co-occurrence and cooperativity between *DNMT3A*-R882H and additional secondary mutations is required to drive hematological malignancies in a *DNMT3A* mutated context. The co-occurrence of *Idh2*-R140Q and *Dnmt3a*-R878H mutations was not sufficient to engender manifestations of hematological diseases, yet cells were distinguished by an epigenetic antagonism and the reversion of hyper- and hypomethylation patterns observed in the respective single-mutants [Glass et al., 2017].

1.6 DNA Damage and Repair Pathways Maintain Genome Integrity

1.6.1 Sources of DNA Damage

While DNA damage in mammalian cells occurs at a rate of more than 70,000 alterations per day [Tubbs and Nussenzweig, 2017], DNA repair mechanisms are critical in order to maintain genome integrity and prevent the excessive acquisition and accumulation of mutations.

Endogenous DNA damage predominantly arises from chemical reactions the DNA molecule is engaged in, for instance hydrolytic or oxidative reactions with molecules naturally present in the cell [Chatterjee and Walker, 2017]. Due to its chemical structure, DNA is predisposed to spontaneous deamination and depurination, the non-enzymatic transfer of methylgroups via S-adenosylmethionine (SAM), or oxidative damage induced

by ROS, with all reactions eventually leading to lesions within the DNA molecule [Chatterjee and Walker, 2017]. Additionally, errors occurring during DNA replication, including base substitutions, insertions or deletions, as well as during the action of topoisomerase enzymes to remove superhelical tension further contribute to endogenously derived DNA damage [Loeb and Monnat, 2008, Wang, 2002]. On the other hand, exogenous DNA damage occurs due to the interaction of the DNA molecule with environmental, physical or chemical agents, such as ionizing and ultraviolet radiation, alkylating agents, aromatic amines, etc. [Chatterjee and Walker, 2017].

Depending on its source, endogenous or exogenous damage frequently results in the formation of single-stranded breaks (SSBs) or, less frequent, in the formation of double-stranded breaks (DSBs) which are more perilous due to a higher risk of chromosomal rearrangements, amplifications or deletions [Delia and Mizutani, 2017].

1.6.2 The DNA Damage Response Pathway

Upon the formation of SSB or DSB lesions, the activation of the DNA damage response (DDR) machinery encompasses a complex network of signaling pathways to coordinate not only the detection and repair of lesions, but also to transiently arrest cell cycle activity, and, in cases of excessive damage, induce apoptotic or senescent cell programs [Ciccia and Elledge, 2010].

Forming the cornerstone of the DDR pathway, the protein kinases ATM, ATR and DNA-PK, are activated by DNA breaks and phosphorylate multiple downstream targets to regulate repair and cellular activity [Blackford and Jackson, 2017]. Here, ATM primarily is recruited to DSBs and orchestrates the activity of several cell cycle checkpoints, whereas ATR is primarily responsive to SSBs generated during replication stress and activates S-checkpoint signaling [Blackford and Jackson, 2017]. DNA-PK, similarly to ATM, is recruited to DSBs, but promotes its repair via non-homologous end joining (NHEJ) whereas ATM activation frequently results in resolving DSBs via homologous recombination (HR). The decision between NHEJ and HR depends on multiple factors, such as cell cycle phase, local chromatin environment or the structure of DNA ends [Scully et al., 2019].

1.6.3 Homologous Recombination and the Repair of Double-Strand Breaks

Following the recruitment of ATM to DSBs, ATM phosphorylates and activates downstream mediators such as MDC1, 53BP1 and BRCA1, which in part are antagonizing and counteracting each other in order to promote either NHEJ (53BP1) or HR (BRCA1) [Blackford and Jackson, 2017]. A critical step facilitating repair via HR is the resection of DSB ends to yield single stranded DNA (ssDNA) 3' ends mediated by the endonuclease activity of the MRN complex and the exonuclease activity of EXO1/BLM [Nimonkar et al., 2011, Mimitou and Symington, 2008, Symington and Gautier, 2011]. The generated ssDNA stretches are rapidly bound by RPA complexes which then are displaced by RAD51 to form a RAD51-nucleoprotein filament (facilitated by mediator RAD52) [Baumann et al., 1996, Sugiyama et al., 1997, San Filippo et al., 2008]. In contrast to NHEJ, HR relies on pervasive sequence homology between the broken DNA molecule and a donor strand template (usually the homologous chromosome), and involves templated strand synthesis during the repair process [Scully et al., 2019]. In mammalian cells, BRCA2 serves as the key recombination mediator which aids in RAD51-mediated strand invasion and

homologous pairing of the RAD51 ssDNA filaments and complementary sequences [Yang et al., 2005a, Jensen et al., 2010]. Heteroduplex formation, followed by the engagement of DNA polymerases results in the extension of the invading strand using the donor strand as a template [Chen et al., 2008, van der Heijden et al., 2008, McVey et al., 2016]. Eventually, several subpathways can be employed to resolve and disintegrate the homologous structures, depending on various fates of the RAD51-mediated recombination synapse [Scully et al., 2019].

1.6.4 PARP1 Mediates Single-Strand Break Repair and is Targeted in Patients with Defects in Homologous Recombination

Although being in part also involved in DSB repair, PARP1 is the key enzyme which detects and binds SSBs and mediates excision repair [Ray Chaudhuri and Nussenzweig, 2017]. Following SSB lesion detection, PARP1 recruits scaffold proteins (XRCC1), phosphatases (PNKP), ligases (LIG3) and DNA polymerases (Pol δ, Pol ε, and Pol β) to process and fill the strand gap [Ray Chaudhuri and Nussenzweig, 2017].

Upon bulky DNA lesions (e.g. UV damage), the global genome nucleotide excision repair (GG-NER) pathway is employed which involves lesion recognition by XPC, RAD23B as well as the DDB1-DDB2 complex and is promoted by PARP1 [Pines et al., 2012, Robu et al., 2017]. Subsequent steps, which in part are also facilitated by PARP1, involve chromatin relaxation and remodeling, lesion verification and excision, and eventually gap filling and ligation [Marteijn et al., 2014, Ray Chaudhuri and Nussenzweig, 2017].

Due to its important role in lesion recognition and repair, PARP inhibitors recently have emerged as promising targeted therapy options in the context of synthetic lethality [Lord and Ashworth, 2017]. Patients carrying germline mutations in BRCA1 or BRCA2 have elevated risks of developing breast, ovarian and other cancers, because of an only inefficient repair of DSBs by HR [Miki et al., 1994, Wooster et al., 1995, Paul and Paul, 2014]. Inhibition of PARP causes persistent SSBs, which upon replication result in the collapse of replication forks and the generation of DSBs [Lord and Ashworth, 2017]. Exploiting the deficiencies in HR repair arising through dysfunctional BRCA1/2, the overwhelming accumulation of DSBs eventually results in chromosomal instability, cell cycle arrest and cell death [Farmer et al., 2005, Bryant et al., 2005]. Although reports meanwhile suggest that PARP inhibitors function by interfering with the catalytic cycle of PARP and thereby result in cytotoxic lesions consisting of trapped PARP [Murai et al., 2012, Pommier et al., 2016], other tumor entities that likewise display defects in HR due to ATM deficiency or due to *IDH1/2* mutations were shown to be similarly hypersensitive and more susceptible to PARP inhibitors [Weston et al., 2010, Sulkowski et al., 2017].

2

Aims of the Thesis

The impact of an *IDH1*-R132H mutation on a cell is diverse, ranging from deregulation of the epigenetic landscape to defects in DNA damage and DNA repair as well as alterations in cellular signaling [Gagne et al., 2017]. Importantly, many oncogenic functions of *IDH1*-R132H mutations are derived from the analysis of brain tumor cells, whereas the effects on hematopoietic cells still remain more elusive.

In the hematopoietic system, earlier work suggests the effects of an *Idh1*-R132H mutation to depend on the cell type affected. While in LT-HSCs functionality is impaired through a prevailing inhibition of histone demethylases and a down-regulation of the damage-response gene ATM, in downstream populations (e.g. ST-HSCs, MPPs) effects of an Tet2 inhibition predominate and result in altered DNA methylation and differentiation patterns [Inoue et al., 2016]. However, although *Idh1*-R132H mutated HSCs display an accumulation of DNA damage and an impaired repair efficiency, how these characteristics affect functionality, self-renewal, differentiation and reconstitution potential in a transplantation (or serial transplantation) setting needs to be determined.
Moreover, the fraction of hematopoietic cell types analyzed in the context of *Idh1*-R132H mutations remains incomplete. For instance, as Tet2 also is involved in regulating the expression of inflammatory mediators in macrophages [Zhang et al., 2015, Fuster et al., 2017, Jaiswal et al., 2017, Meisel et al., 2018] , it seems likely that *Idh1*-R132H mutations are not confined to only affect functionality of stem and progenitor cells, but also differentiated cell populations such as monocytes or macrophages. Furthermore, how gene networks, which are dysregulated by *Idh1*-R132H mutations in different cell types, potentially contribute to cancer initiation or progression is only poorly understood. Yet, breakthroughs in single-cell RNA sequencing methods not only allow to analyze the transcriptomic landscape of a large number of single cells, they also can be applied to dissect cellular heterogeneity in the hematopoietic compartment and reconstruct differentiation trajectories [Giladi et al., 2018, Nestorowa et al., 2016, Velten et al., 2017, Laurenti and Gottgens, 2018].

Despite the fact that in humans, *IDH1*-R132H mutations represent a substantial risk factor for disease progression and AML development [Abelson et al., 2018, Desai et al., 2018], the oncogenic potential in mice is limited [Sasaki et al., 2012, Inoue et al., 2016]. Whereas cooperative effects of *Idh1/2* mutations with proliferative and activating mutations in *Npm1, Flt3* or *Nras* have been described previously [Ogawara et al., 2015, Chen et al., 2013a, Kats et al., 2014], an interplay between *Idh1*-R132H mutations and other early preleukemic mutations such as *Dnmt3a*-R878H / *DNMT3A*-R882H mutations,

has not been investigated extensively. Importantly, since *IDH1*-R132H and *DNMT3A*-R882H mutations frequently co-occur in AML patients (see Figure 1-5C) [Cancer Genome Atlas Research et al., 2013, Papaemmanuil et al., 2016], studying their interactions therefore is of particular clinical relevance.

Collectively, the main aim of this doctoral thesis was to characterize in detail the effect of *Idh1*-R132H and *DNMT3A*-R882H mutations alone or in combination on hematopoietic functionality and differentiation, and to identify gene regulatory networks which are deregulated upon the expression of these mutations in individual hematopoietic cell types.

To address this aim, the following primary objectives were devised throughout this thesis:

1. Generation of *in vitro* and *in vivo* model systems to investigate the effects of *Idh1*-R132H and *DNMT3A*-R882H mutations.

2. Establishing a conditional and inducible mouse model which expresses the *Idh1*-R132H and / or the *DNMT3A*-R882H mutation in the hematopoietic compartment.

3. Characterization of the functionality and reconstitution potential of *Idh1*-R132H and *DNMT3A*-R882H single-mutated- as well as *Idh1*-R132H *DNMT3A*-R882H double-mutated HSCs in transplantations and serial transplantations.

4. Characterization of the blood and bone marrow composition of mice transplanted with *Idh1*-R132H and *DNMT3A*-R882H single-mutated- as well as *Idh1*-R132H *DNMT3A*-R882H double-mutated cells.

5. Generation of a genotype-specific transcriptomic landscape of the hematopoietic compartment, encompassing stem cells, progenitor cells and differentiated cells.

6. Reconstruction of transcriptome-based differentiation trajectories and their plasticity in response to the expression of single or combined *Idh1*-R132H and *DNMT3A*-R882H mutations.

7. Identification of hematopoietic cell types to be most affected by *Idh1*-R132H and / or *DNMT3A*-R882H mutations.

8. Identification of hematopoietic cell type-specific gene regulatory programs differentially expressed between the individual mutated genotypes.

9. Correlation of aberrant transcriptome-defined cell populations with surface marker-defined cell populations for prospective analysis and characterization.

Overall, our phenotypic and molecular analysis of *Idh1*-R132H and / or *DNMT3A*-R882H mutated hematopoietic compartments will enhance our understanding of how these mutations alone or in combination interact to drive initiation and progression of leukemia development. The identification of aberrantly expressed gene regulatory networks will further be of benefit in designing treatment strategies tailored for patients carrying single or combined *IDH1*-R132H and *DNMT3A*-R882H mutations.

3

Results

In the course of this thesis, effects of an *Idh1*-R132H mutation were investigated in a hematopoietic cell line model as well as in a hematopoiesis-specific knock-in mouse model. Experiments performed in cell lines focused on addressing the impact of *Idh1*-R132H mutations on cell growth, proliferation and DNA damage repair, whereas experiments in mice were specifically aimed to characterize self-renewal, repopulation- and multi-lineage potential of *Idh1*-R132H mutated HSCs. In addition, we investigated the cooperativity of *Idh1*-R132H mutations and *DNMT3A*-R882H mutations in a double-mutated mouse model and assessed mutation-mediated differentiation defects phenotypically and at a molecular level by using single-cell and bulk RNA-sequencing.

3.1 Characterization of an *Idh1*-R132H Mutated Cell Line Model

In order to investigate the effects of *Idh1*-R132H mutations on myeloid cells, we expressed this mutation in the myeloid 32D bone marrow cell line. Therefore, we cloned a cDNA encoding the *Idh1*-R132H mutation into a lentiviral overexpression plasmid, where the expression of the *Idh1*-R132H mutation and an mCherry fluorescent reporter are driven from a spleen focus forming virus (SFFV) promoter (Figure 3-1A). After successful transduction and sorting of the cells based on mCherry expression, we used an enzymatic assay [Balss et al., 2012] to quantify D-2-Hydroxyglutarate (D2HG) levels in the respective 32D cell lysates. As expected, we detected a more than 20-fold increase in D2HG levels (normalized to the overall protein content) in 32D cells that were transduced with *Idh1*-R132H constructs, compared to cells that were transduced either with an empty vector construct or with an *Idh1* wildtype (*Idh1*-WT) overexpressing construct (Figure 3-1B).

Having confirmed the increased production of D2HG due to the expression of the *Idh1*-R132H mutation in our model system, we next characterized the functional consequences of this metabolite on 32D cells. Based on previous findings that linked the *Idh1*-R132H mutation and the concomitant increase in D2HG to alterations in DNA damage response and DNA repair pathways [Sulkowski et al., 2017, Inoue et al., 2016],

we specifically assessed proliferation and DNA damage rates in our 32D cell line model. By using combined 5-Ethynyl-2´-deoxyuridine (EdU) - 7-Aminoactionomycin (7AAD) labeling, we measured percentages of cells in G0-/G1- phase, S-phase or G2/M-phase of the cell cycle (Figure 3-1C). *Idh1*-R132H mutated cells displayed increased percentages of cells that are not actively cycling (G0-/G1-Phase), whereas percentages of DNA replicating cells (S-phase) were reduced (Figure 3-1D). In line with this, we could show decreased proliferation rates of *Idh1*-R132H mutated cells in a co-culture assay, where *Idh1*-R132H mutated cells and empty vector transduced cells were seeded at a 1:1 ratio. By using two different fluorescent proteins as reporter (mCherry for *Idh1*-R132H / *Idh1*-WT transduced cells, GFP for empty vector transduced cells), relative growth rates were tracked over time by quantifying mCherry:GFP ratios using flow cytometry (Figure 3-1E). Whereas co-cultured *Idh1*-WT and empty vector transduced cells displayed equal growth rates, *Idh1*-R132H transduced cells were defined by markedly reduced growth rates, that resulted in a continuous decrease of relative *Idh1*-R132H mutated cell frequencies.

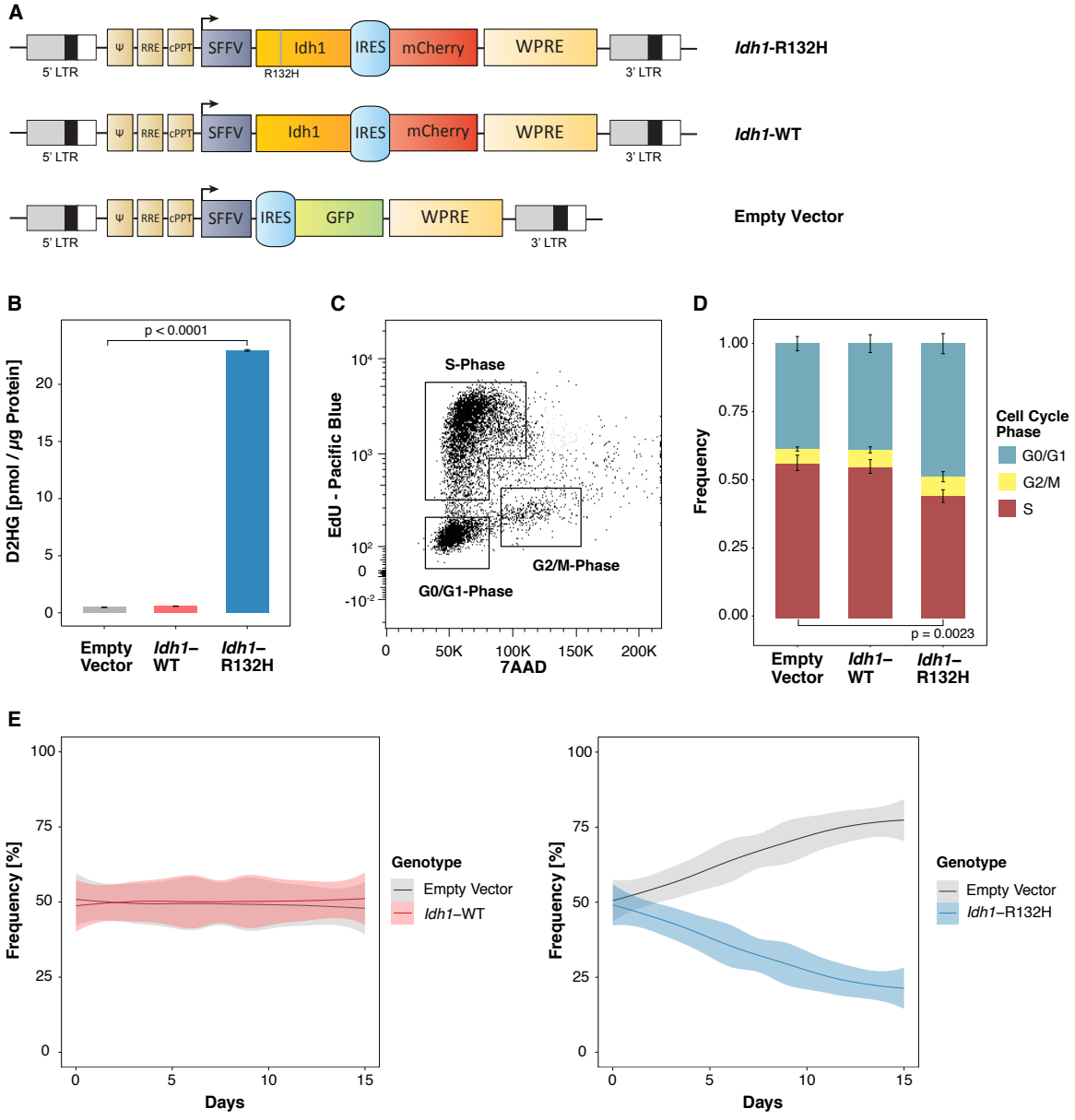


Figure 3-1: *Idh1*-R132H Mutated 32D Cells Display Aberrations in Cell Proliferation.

(A) Schematic overview of lentiviral *Idh1*-R132H, *Idh1*-WT and empty vector overexpression plasmids used for the transduction of the 32D myeloid bone marrow cell line. **(B)** Enzymatic quantification of D2HG in 32D cell lysates transduced either with an empty vector, *Idh1*-WT or *Idh1*-R132H overexpression plasmid (n = 2, mean ± SD). **(C)** Representative FACS gating scheme of an EdU-7AAD staining performed with transduced 32D cells. Pacific Blue signal correlates with incorporated EdU and allows cell cycle state assignment. **(D)** Quantification of cell cycle phases for empty vector, *Idh1*-WT and *Idh1*-R132H transduced cells. The depicted p-value was calculated by contrasting S-phase frequencies from *Idh1*-R132H cells with S-phase frequencies from empty vector cells (n = 3, mean ± SD). **(E)** Competitive proliferation assay of either *Idh1*-WT or *Idh1*-R132H transduced cells co-cultured with empty vector transduced cells seeded at a 1:1 ratio. *Idh1*-R132H or *Idh1*-WT cells were distinguished from empty vector cells based on the respective fluorescence signal (mCherry vs. GFP). Fluorescence signals were determined by flow cytometry (n = 2, depicted curves were smoothened based on a loess (locally estimated scatterplot smoothing) regression model, confidence interval (0.95) is plotted around curve).

To assess whether the reduced proliferation rates could be traced back to an increased amount of DNA damage in these cells, we performed a comet assay to measure DNA damage levels under steady state conditions, directly after irradiation, and after recovery phases following irradiation. Here, we observed that already under steady state conditions, *Idh1*-R132H mutated 32D cells feature significantly higher DNA damage levels as quantified by the olive tail moment of individual cells (Figure 3-2A). Directly after irradiation with 2- or 5 Gy, the observed effect became even more pronounced, while after a 30- or 60 min recovery phase, DNA damage levels again reached physiological levels, suggesting that repair kinetics in *Idh1*-R132H mutated cells are comparable to empty vector transduced cells.

As *Idh1*-R132H mutations were shown to result in defects in the homologous recombination (HR) pathway which is required for the efficient repair of double-stranded breaks (DSBs) [Sulkowski et al., 2017], we next investigated whether the *Idh1*-R132H mutated 32D cell line model is more sensitive to PARP inhibitors. While PARP is predominantly involved in the repair of single-stranded breaks (SSBs), inhibition of PARP causes persistent SSBs, which frequently lead to the generation of DSBs during DNA replication cycles. In case deficiencies in HR-mediated repair exist, DSBs accumulate over time and eventually lead to synthetic lethality, decreased survival rates and an increased sensitivity to PARP inhibitors. Therefore, we analyzed cell viability of *Idh1*-R132H mutated 32D cells in response to different concentrations of the PARP inhibitor Olaparib over 3 days (Figure 3-2B). We found *Idh1*-R132H mutated cells to be significantly more sensitive to PARP inhibition, especially at a concentration of 1 µM Olaparib. Increasing concentrations (4 µM Olaparib) did not aggravate discrepancies between *Idh1*-R132H mutated and empty vector transduced samples, but still resulted in significant differences in cell viability after three days of treatment.

Taken together, we established an *Idh1*-R132H mutated cell line model and could verify the expression of the mutation by measuring increased levels of the oncometabolite D2HG. On a functional level, the *Idh1*-R132H mutation resulted in reduced cell proliferation and higher DNA damage rates already at steady state. Cell viability assays after Olaparib treatment suggest that defects in HR could contribute to *Idh1*-R132H-mediated defects in DNA damage repair, yet further molecular analyses are required to determine detailed mechanisms underlying repair deficiencies.

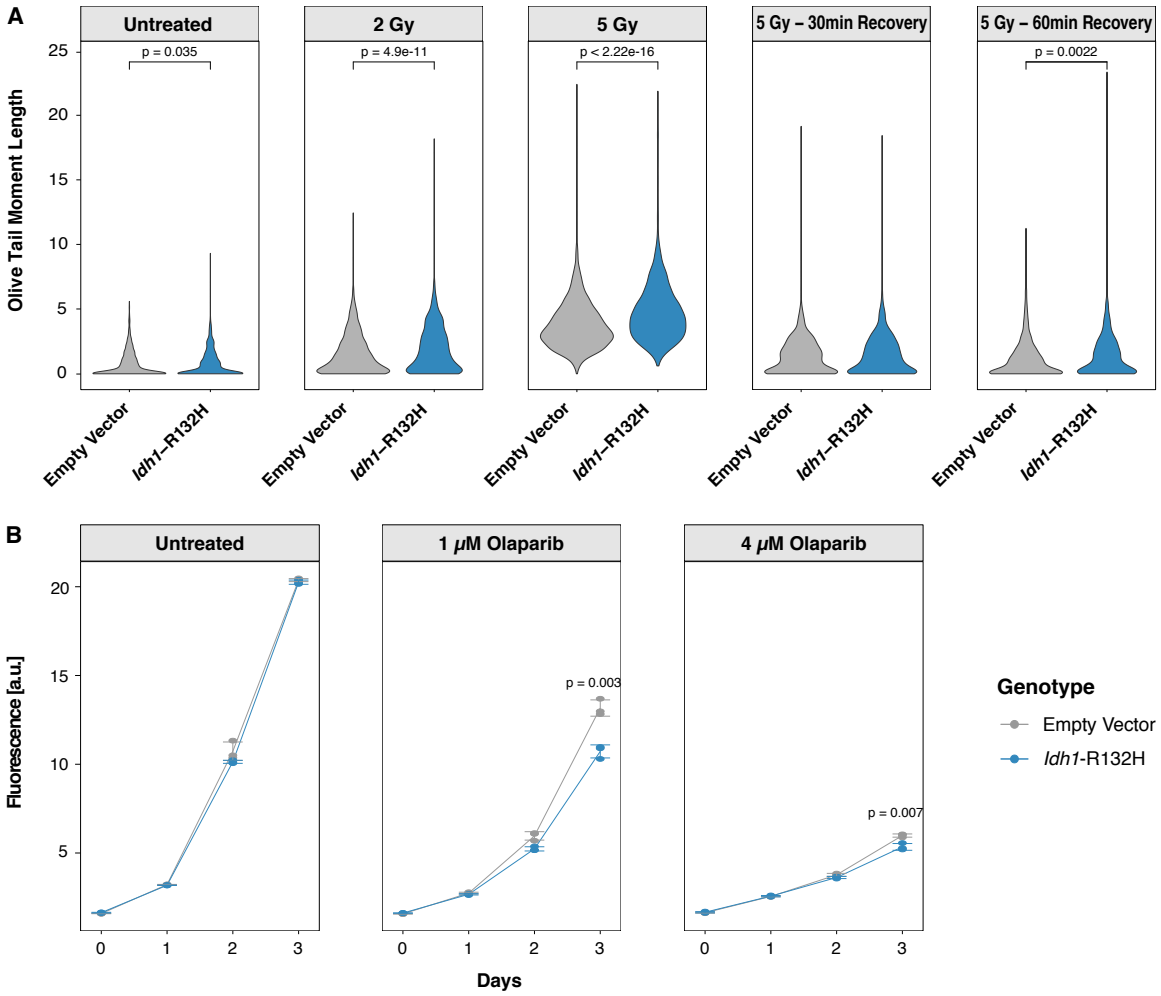


Figure 3-2: *Idh1*-R132H Mutated 32D Cells Display Increased DNA Damage Rates and Sensitivity to PARP Inhibitors. **(A)** Quantification of comet assays performed either untreated or after irradiation and recovery phases as indicated. Between 925 and 1693 cells were analyzed per sample (Student's t-test, p-values are indicated for significant results ($p < 0.05$)) (Comet assays were performed by Reinhard Gliniorz in collaboration with Dr. Peter Schmezer and PD Dr. Odilia Popanda) **(B)** Analysis of cell viability by cell titer blue assays of transduced 32D cells after treatment with the PARP-inhibitor Olaparib. Cells were treated over 3 days and cell viability was determined at each day. Fluorescence intensities correlate with cell viability (n=3, error bars indicate mean ± SD, Student's t-test, p-values are indicated for significant results ($p < 0.05$)) (Cell titer blue assays were performed by Dr. Ali Bakr).

3.2 Characterization of a Hematopoiesis-Specific *Idh1*-R132H Knock-In Mouse Model

To recapitulate previous findings from our cell line model in an *in vivo* setting, we employed an inducible *Idh1*-R132H mutated mouse model. Here, two opposing lox sites flank an R132H mutated exon 3, which is reversely inserted in the endogenous *Idh1* locus between exon 3 and exon 4 (Figure 3-3A). The regular non-mutated exon 3, which is transcribed in the absence of Cre recombinase (Cre) activity, is flanked by lox sites of equal orientation. Upon Cre activity within the cell, the non-mutated exon 3 is excised due to the orientation of the lox sites, while the inverted R132H mutated exon 3 is inserted in the correct orientation and becomes actively transcribed.

As *Idh1*-R132H mutations occur very early in AML development and often within then hematopoietic stem cell compartment [Corces-Zimmerman et al., 2014], we restricted tamoxifen-inducible CreERT2 expression to the HSPC compartment by driving its expression under the control of the stem cell-specific enhancer of the stem cell leukemia (*Scl*) locus [Gothert et al., 2005]. The *Scl* enhancer is responsible for driving Scl expression only within the most immature hematopoietic cell types [Gottgens et al., 2002, Sanchez et al., 1999, Sanchez et al., 2001], thereby confining the activity of Cre to primarily HSCs in our mouse model. To monitor Cre activity after tamoxifen induction, we made use of an enhanced yellow fluorescent protein (EYFP) reporter mouse model, where the *EYFP* gene is inserted after a lox-flanked STOP codon within the *Rosa26* locus and transcribed after the Cre-mediated excision of the STOP codon [Srinivas et al., 2001]. For all experiments described hereafter, we used transgenic mice that were heterozygous for the mutated *Idh1*-R132H allele, the *Rosa26*-EYFP allele, as well as for the *Scl*-CreERT2 transgene (*Idh1*$^{+/R132H}$, *Scl*-CreERT2$^{+/d}$, *Rosa26*-EYFP$^{+/d}$, denoted as *Idh1* mice). Control mice were heterozygous for the *Scl*-CreERT2 transgene and the *Rosa26*-EYFP allele but lacked a mutated *Idh1*-R132H allele (*Idh1*$^{+/+}$, *Scl*-CreERT2$^{+/d}$, ROSA26-EYFP$^{+/d}$, denoted as control (CTRL) mice).

To induce the expression of the *Idh1*-R132H mutation and the EYFP reporter gene, we induced Cre activity by treating our mice with tamoxifen for 5 consecutive days (2 mg / day). Two to four weeks after tamoxifen treatment, we analyzed the bone marrow of injected animals for EYFP expression (Figure 3-3B). Among total bone marrow (TBM) cells, we detected around 6-7 % EYFP$^+$ cells. However, when gating for hematopoietic stem and progenitor cells (Lineage$^-$, Sca1$^+$, cKit$^+$ (LSK) cells), HSCs, or LT-HSCs (for all gating schemes see Supplemental Figure A-1 and Supplemental Figure A-2), the percentage of EYFP$^+$ cells increased the more immature the cell type was (Figure 3-3C). LT-HSCs reside at the top of the hematopoietic hierarchy and consequently, with 35-40 % of EYFP$^+$ cells, they displayed the highest fraction of EYFP$^+$ cells, indicating that due to the *Scl* enhancer-driven expression, Cre activity indeed was highest in HSCs and decreased in more differentiated cells. It is important to note that EYFP positivity observed in more differentiated cells could in fact be derived from labelled EYFP$^+$ HSCs giving rise to EYFP$^+$ downstream cells, and not necessarily reflects the activity of Cre within these cells. To address whether EYFP$^+$ cells in the bone marrow give rise to differentiated cell types of all lineages in the periphery, we also quantified EYFP expression in the blood of injected animals. We found EYFP expression to be present in B-cells, T-cells and myeloid cells (Figure 3-3D), suggesting that EYFP positivity can be used as a marker to trace differentiation of EYFP-labelled stem cells in, for example, transplantation settings. In mice that did not carry the *Scl*-CreERT2 or the *Rosa26*-EYFP allele, we could not detect any EYFP$^+$ cells in the bone marrow at all (Figure 3-3E and Figure 3-3F).

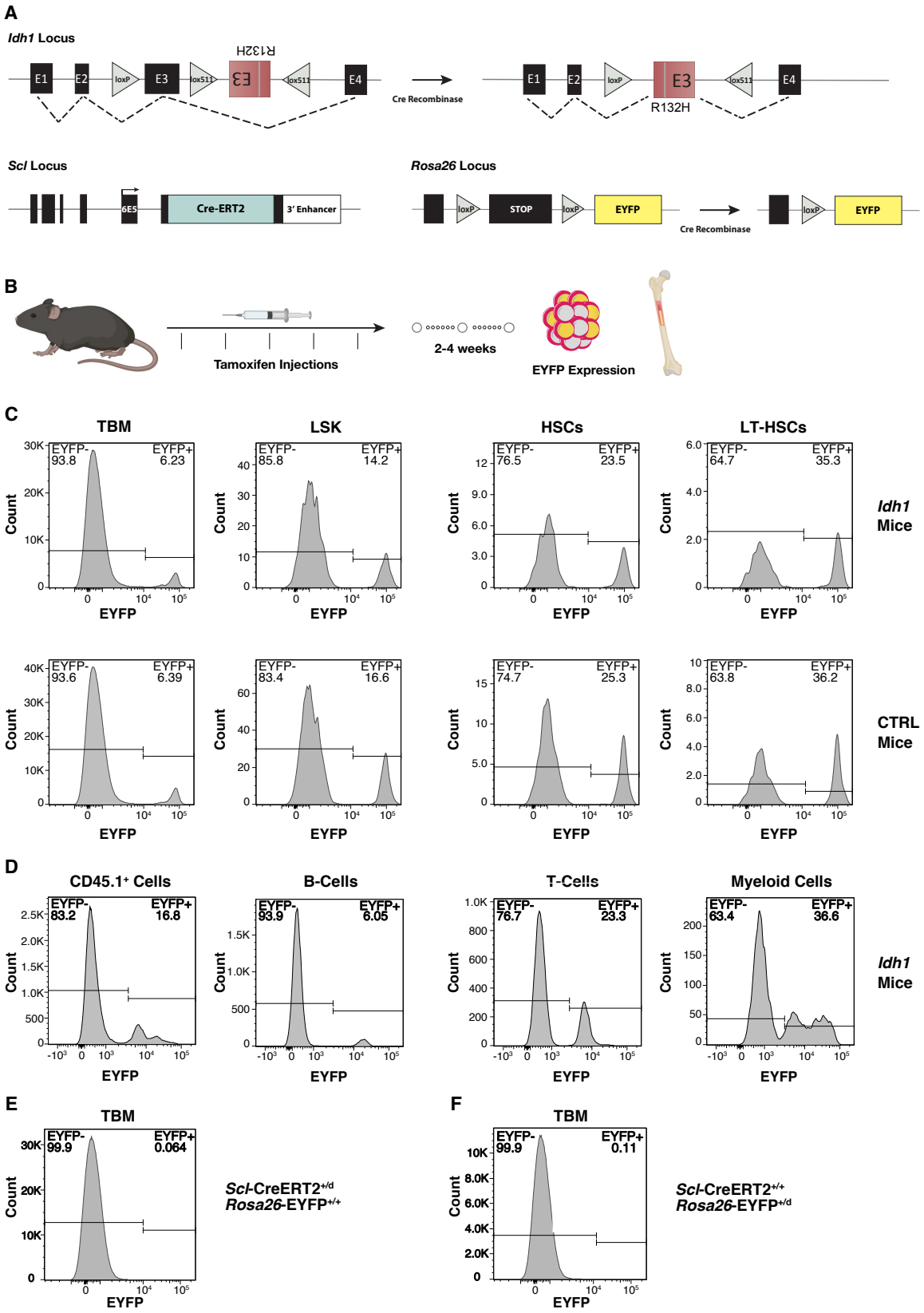


Figure 3-3: *Idh1*-R132H *Scl*-CreERT2 *Rosa26*-EYFP Mice Express EYFP in the Bone Marrow and Blood after Tamoxifen Treatment.
(A) Schematic overview of transgenic mice used throughout this study. Upon Cre activity the non-mutated *Idh1* exon 3 is excised and the R132H mutated exon 3 is inversed and transcribed. CreERT2 expression is driven from the *Scl* locus controlled by the stem cell-specific *Scl* 3' enhancer and the 6E5 exon 4 promoter. The fluorescence marker EYFP, inserted into the *Rosa26* locus, is used as a reporter for Cre activity. **(B)** Schematic overview of the experimental setup. **(C + D)**

Exemplary quantification of EYFP expression in the bone marrow (C) and blood (D) of tamoxifen-injected *Idh1* and CTRL animals by flow cytometry. **(E + F)** Exemplary quantification of EYFP expression in the bone marrow of injected animals, that lack either a *Scl*-CreERT2 transgene (E) or a *Rosa26*-EYFP allele (F) by flow cytometry.

While EYFP expression serves as a marker for Cre activity within the cell, the recombination events occurring at the *Idh1* locus and the *Rosa26* locus are not only independent of each other, they also may occur at different efficiencies or different rates. To assess how recombination at the *Rosa26* locus as measured by EYFP expression correlates with recombination at the *Idh1* locus, we sorted EYFP$^+$ and EYFP$^-$ bone marrow cells from tamoxifen-injected animals and performed a recombination-specific PCR for the *Idh1* locus. Here, *Idh1* locus recombination events occurred both in EYFP$^+$ as well as in EYFP$^-$ cells sorted from *Idh1* mice, confirming an independency of the recombination events (Figure 3-4A). As expected, cells sorted from CTRL mice did not show any *Idh1* locus recombination events.

To determine the exact recombination frequency at the *Idh1* locus in both EYFP$^+$ and EYFP$^-$ cells, we sorted single EYFP$^+$ and EYFP$^-$ HSCs from *Idh1* mice, grew colonies from these cells in liquid culture and subsequently performed the *Idh1* locus recombination-specific PCR on each of the colonies (Figure 3-4B). In EYFP$^+$ single-cell-derived colonies, we observed an *Idh1* locus recombination efficiency of 91.4 %, resulting in the expression of the *Idh1*-R132H mutation in almost all EYFP$^+$ cells (Figure 3-4C). In contrast, also 53.7 % of EYFP$^-$ single-cell derived colonies featured an *Idh1* locus recombination, implying that roughly every second EYFP$^-$ cell also expresses the *Idh1*-R132H mutation and that recombination at the *Idh1* locus is favored and occurs more likely than recombination at the *Rosa26* locus. Yet, due to the high concordance of *Idh1* locus recombination in EYFP$^+$ cells, EYFP expression can be used as a marker to strongly enrich for cells that express the *Idh1*-R132H mutation.

In order to show that the recombination of the *Idh1* locus results in transcription of the *Idh1*-R132H mutation, we designed *Idh1*-R132H mutation-specific primers that only show an amplification product when the *Idh1*-R132H mutation is expressed. This is enabled by the specific design of one of the primers, which overlaps with the mutated codon at the 3' end. Using cDNA synthesized from total RNA as a template, we could detect a mutation-specific amplification product only in cells derived from *Idh1* mice but not from CTRL mice (Figure 3-4D). An analogous *Idh1* wildtype-specific primer, however, resulted in an amplification product in cells coming from both *Idh1* and CTRL animals, since *Idh1* mice were heterozygous for the *Idh1*-R132H mutation and still carried one *Idh1* wildtype allele. Quantification of D2HG levels in the blood serum of tamoxifen-treated animals revealed that D2HG concentrations in serum of *Idh1* mice were increased around two-fold compared to concentrations in the serum of CTRL mice (Figure 3-4E). Although the exact number or fraction of *Idh1*-R132H mutated bone marrow cells cannot be determined due to the discrepancy between EYFP and *Idh1* locus recombination, the observed elevated D2HG levels suggest, that already a low percentage of *Idh1*-R132H mutated bone marrow cells (around 7% EYFP$^+$ TBM cells after tamoxifen treatment, see Figure 3-3C) is sufficient to result in a remarkable increase in D2HG levels.

In summary, we have characterized a tamoxifen-inducible hematopoiesis-specific *Idh1*-R132H mutation mouse model where EYFP expression serves as a surrogate marker for the expression of the *Idh1*-R132H mutation. Comparable to the *Idh1*-R132H 32D cell line model, the expression of the mutation resulted in increased levels of D2HG in the blood serum of tamoxifen-treated animals.

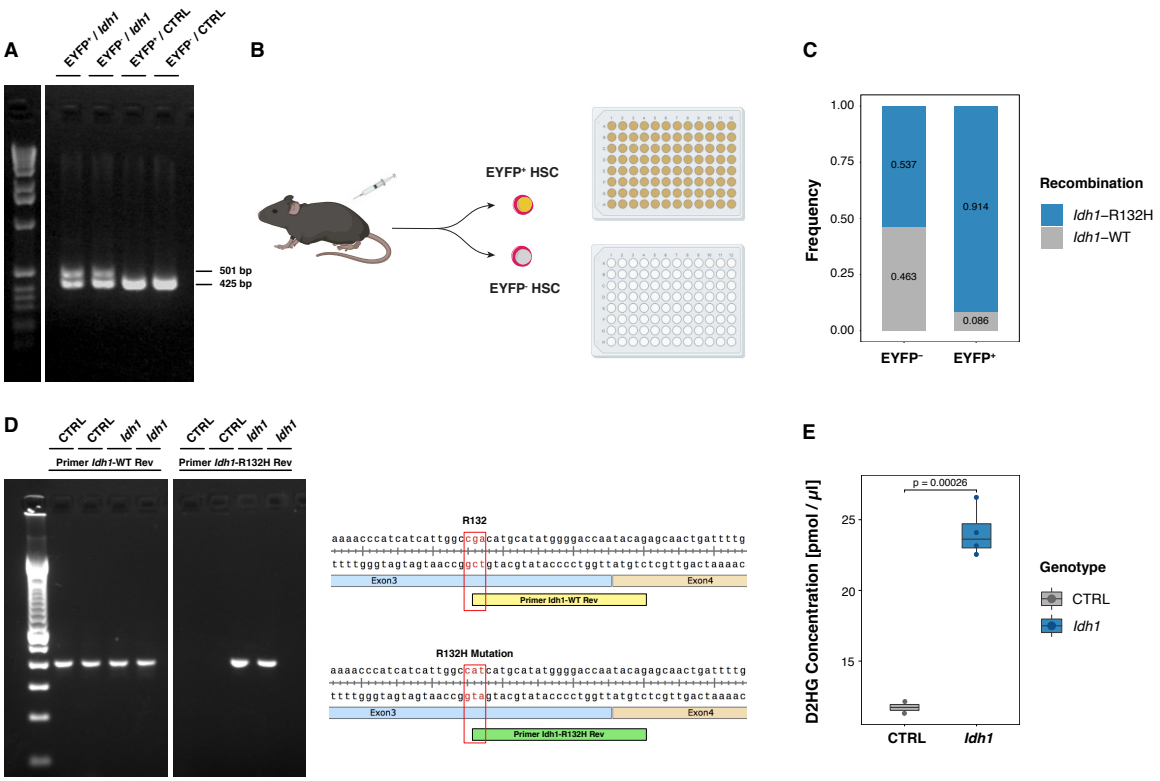


Figure 3-4: Tamoxifen-Treated *Idh1*-R132H *Scl*-CreERT2 *Rosa26*-EYFP Mice Express the *Idh1*-R132H Mutation and Accumulate Increased Levels of D2HG in the Blood Serum.
(A) EYFP⁺ and EYFP⁻ cells were sorted from *Idh1* or CTRL mice as indicated and an *Idh1* locus recombination-specific PCR was performed. The upper band at 501 bp indicates the presence of the recombined mutated allele, resulting in the expression of the *Idh1*-R132H mutation, whereas the band at 425 bp indicates the presence of an *Idh1* wildtype allele. **(B)** Experimental setup for determining the correlation between EYFP expression and recombination efficiency at the *Idh1* locus. **(C)** From single cells sorted as illustrated in (B), single-cell-derived colonies were grown and the recombination efficiency at the *Idh1* locus was determined by recombination-specific PCR (EYFP⁻ n = 41, EYFP⁺ n = 35) (PCRs performed by Umut Kilik). **(D)** *Idh1*-R132H mutation-specific primers were designed to validate the expression of the *Idh1*-R132H mutation on RNA level in blood cells from *Idh1* or CTRL mice as indicated. After isolation of total RNA and synthesis of cDNA, the PCR was performed either with wildtype- or mutation-specific reverse primer but with the same forward primer. **(E)** Enzymatic quantification of D2HG levels in the blood serum of tamoxifen-treated *Idh1* and CTRL mice (CTRL n = 2, Idh1 n = 4, Student's t-test).

3.3 Effects of *Idh1*-R132H Mutations on Bone Marrow Repopulation and Multi-Lineage Potential

Having validated the expression of the *Idh1*-R132H mutation in EYFP⁺ cells from *Idh1*-R132H *Scl*-CreERT2 *Rosa26*-EYFP mice, we next addressed functional effects of the *Idh1*-R132H mutation on proliferation and self-renewal capacity of LT-HSCs. A standard *in vitro* assay to assess these aspects is a colony-forming unit (CFU) assay, where cells are seeded in a semi-solid medium to form colonies and serially replated after a certain time. EYFP⁺ LT-HSCs were sorted from tamoxifen-treated *Idh1* and CTRL mice and the number of colonies formed was assessed in three serial platings. While in general, the number of colonies decreased with each plating, especially in the second and third plating, a trend towards a smaller number of colonies in *Idh1*-R132H mutated cells became evident (Supplemental Figure A-3).

For many years, transplantation of HSCs into lethally irradiated recipients has been considered the gold standard to assess engraftment, repopulation and self-renewal potential as well as lineage output in an *in vivo* setting. As recombination efficiency in individual animals is likely to vary due to technical confounders (injection procedure, body weight of the mouse, etc.), a transplantation setting allows us to assess functionality of a defined cell type in a reproducible way in that always the same number of cells is transplanted into irradiated recipients. Therefore, 100 EYFP$^+$ LT-HSCs from tamoxifen-treated *Idh1* or CTRL mice were isolated and transplanted into lethally irradiated recipient mice together with supportive TBM cells (Figure 3-5A). To discriminate between donor cells, supportive bone marrow and recipient cells, we used congenic strains for donor, recipient and supportive cells, that express different alleles of the pan-leukocyte marker CD45. Six weeks after transplantation of the cells into recipients, we began monitoring of recipient mice by drawing blood samples at regular intervals in order to analyze its cellular composition as a surrogate for the cellular output of transplanted LT-HSCs. The overall composition and the overall frequencies of specific cell types were determined by a HemaVet 950 FS hematology system. Here, white blood cell-, leukocyte-, neutrophil-, monocyte-, platelet- and red blood cell counts were mostly comparable for *Idh1* and CTRL transplanted mice during the monitoring phase 6-22 weeks post-transplantation (Supplemental Figure A-4).

By using flow cytometry, the engraftment and lineage output of transplanted donor cells was specifically examined by measuring the overall fraction of CD45.1$^+$ donor cells in the recipient mice. Normalized to the rate of CD45.1/2 double-positive supportive bone marrow cells, we could detect a higher ratio of CD45.1$^+$ cells in mice that were transplanted with *Idh1*-R132H mutated cells compared to mice what were transplanted with CTRL cells (Figure 3-5B). Although the increased CD45.1$^+$ cell ratio levelled out at later time points, this data suggests that *Idh1*-R132H mutated LT-HSCs possess a higher engraftment and proliferative potential in that they are capable of repopulating the recipient blood system more rapidly by generating more progeny. To determine whether the specific lineage output of *Idh1*-R132H mutated LT-HSCs is comparable to control HSCs, we also quantified the fraction of terminally differentiated CD45.1$^+$ B-cells, T-cells and myeloid cells in the blood of transplanted mice (Figure 3-5C). Especially at the end of the monitoring phase, *Idh1*-R132H mutated mice seem to produce slightly less T-cells than CTRL mice but at the same time produce marginally higher rates of myeloid cells.

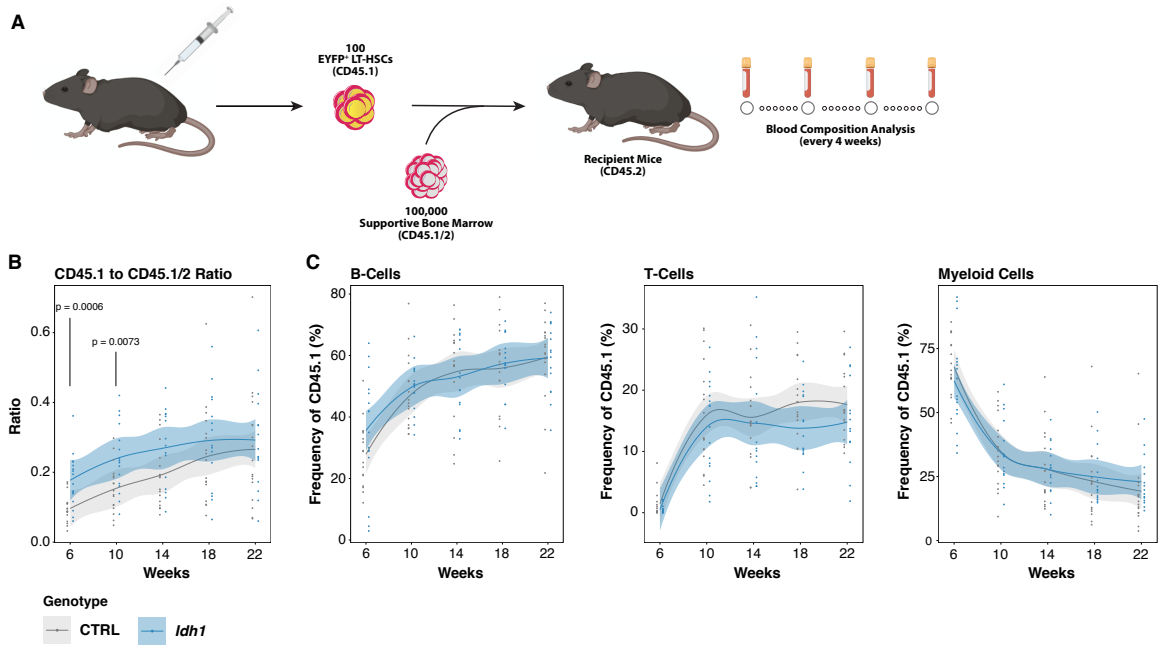


Figure 3-5: *Idh1*-R132H Mutated LT-HSCs Display an Increased Engraftment Potential after Transplantation.
(A) Schematic illustration of the experimental transplantation setup. One hundred EYFP⁺ LT-HSCs from tamoxifen-treated *Idh1* or CTRL donor mice were transplanted together with 300,000 supportive TBM cells into lethally irradiated recipients. Following transplantation, blood composition was monitored at regular intervals. The expression of different CD45 alleles was used to discriminate between donor, recipient and supportive cells. **(B + C)** Quantification of overall CD45.1⁺ cell rates and frequencies of CD45.1⁺ B-cells, T-cells and myeloid cells in the blood of transplanted mice by using flow cytometry (CTRL n = 17, *Idh1* n = 17, three independent transplantations, curves were smoothened based on a loess (locally estimated scatterplot smoothing) regression model, confidence interval (0.95) plotted around curve, Student's t-test, p-values are indicated for significant results ($p < 0.05$)).

In order to understand the molecular features underlying the observed phenotypes, we isolated EYFP⁺ LT-HSCs from transplanted recipients after the 22-week long monitoring phase and performed RNA sequencing (RNA-seq) from three LT-HSC replicates of each genotype. Global correlation analysis of the individual samples revealed a high correlation even for samples from different genotypes (Figure 3-6A). Yet, *Idh1* and CTRL samples clustered apart from each other based on calculated sample distances while the individual replicates clustered together within a given genotype (Figure 3-6B). To examine the heterogeneity of individual samples, a principle component (PC) analysis (PCA) was performed, illustrating a large heterogeneity within *Idh1*-R132H mutated samples, but not within CTRL samples, based on principle components 1 (PC1) and 2 (PC2) (Figure 3-6C). However, PC3, explaining around 20 % of the variance within the data set, clearly separated *Idh1* from CTRL samples (Figure 3-6D). We next called differentially expressed (DE) genes between *Idh1* and CTRL samples and found only a small number of genes to be deregulated in *Idh1* samples (18 genes up-regulated, 3 genes down-regulated), when applying log2 fold change (fc) ≤ -1 or ≥ 1 and adjusted p-vaule (padj) < 0.05 as filtering criteria (Figure 3-6E). To further assess the biological relevance of deregulated genes, we used a 'Gene Set Enrichment Analysis' (GSEA) to identify significantly enriched or depleted groups of genes within the *Idh1*-specific transcriptomic profile. Beside several metabolism-associated gene sets, we also found DNA repair- and inflammation-associated gene sets to be slightly enriched in *Idh1*-R132H mutated samples (Figure 3-6F). Among DNA repair-associated genes, genes which are directly involved in DNA damage sensing and DNA binding to induce downstream repair pathways (e.g. *Rad52*, *Rpa2*, *Xpc*, or *Ddb2*) as well as several DNA and RNA polymerases (e.g. *Pola2*, *Polb*, *Polr2a*) and their respective subunits were higher expressed in *Idh1*-R132H mutated samples. Interestingly, affected

genes play a crucial role in distinct repair pathways, such as HR (*Rad52, Rpa2*) or GG-NER (*Xpc*, *Ddb2*). This suggests that *Idh1*-R132H mutations do not exclusively affect a single pathway. Regarding inflammation related genes, many genes that were found to be higher expressed in *Idh1*-R132H mutated samples were genes which are induced in response to interferon (IFN) signaling or which regulate the expression of IFNs (*Ifi27*, *Ifitm1*, *Ifit1*, *Ifit3*, *Irf2*, *Irf5*, *Irf8*), implying that elevated IFN levels within the bone marrow niche could impact HSC functionality in *Idh1*-R132H mutated mice.

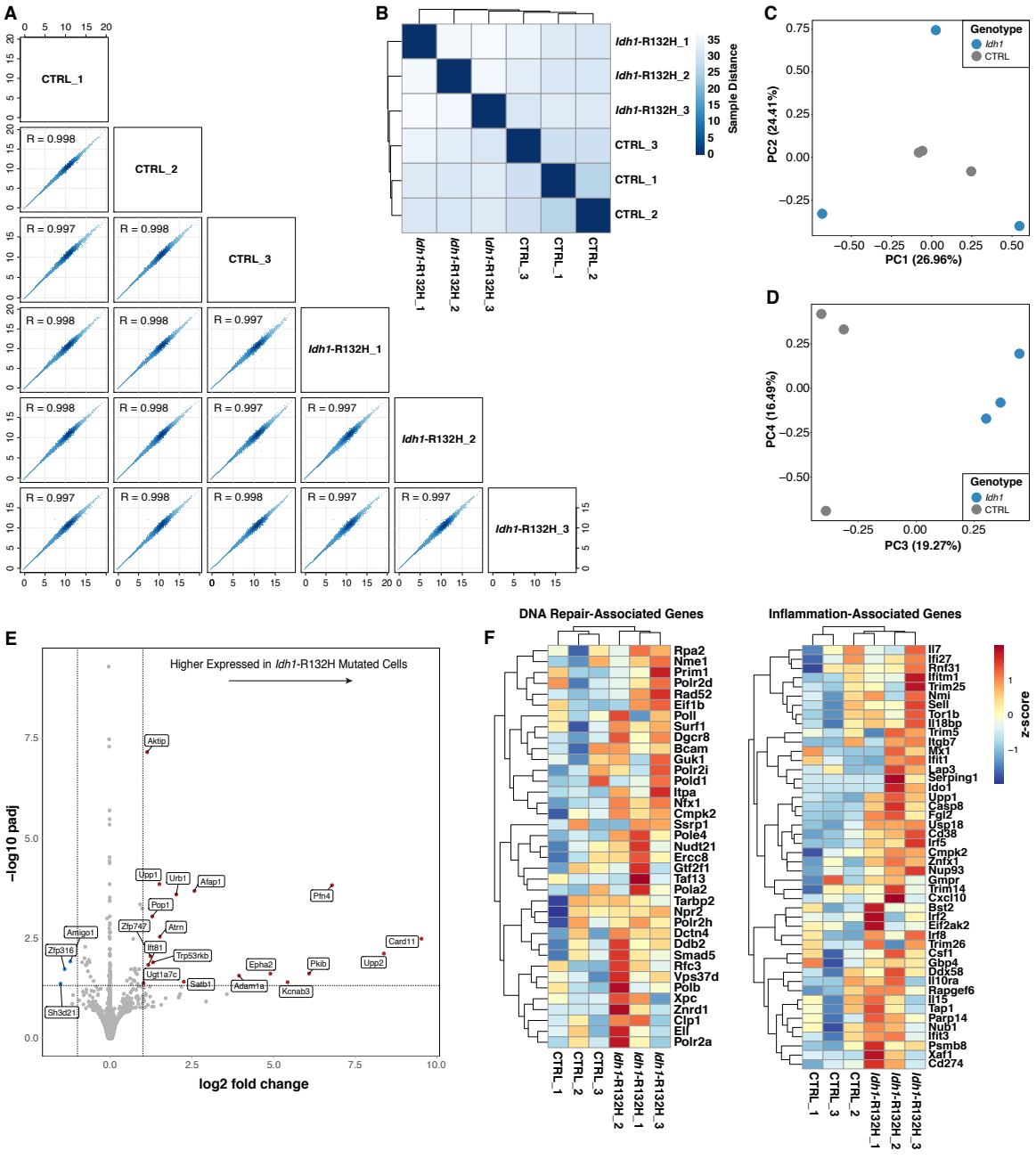


Figure 3-6: *Idh1*-R132H Mutated LT-HSCs Express Higher Levels of DNA Repair- and Inflammation-Associated Genes **(A)** Scatter plots showing normalized expression values for each gene in a pairwise comparison across individual samples. The Pearson correlation coefficient for each comparison is depicted in each plot. **(B)** Heatmap depicting sample distances across samples based on their normalized gene expression profile (using all genes). Calculation of distances and clustering of samples was computed using Euclidean distances. **(C+D)** PCA of individual samples based on their normalized gene expression profile (using all genes). PC1 and PC2 (C) or PC3 and PC4 (D) are plotted against each other. Percentages indicate the magnitude of variance within the data set which is explained by the individual PCs plotted. **(E)** Volcano plot showing DE genes between *Idh1* and CTRL samples. Genes with a log2 fc ≤ -1 or ≥ 1 and a padj < 0.05 are

highlighted. **(F)** Heatmap depicting normalized expression values of selected DNA repair and inflammation-associated genes. Displayed genes represent 'core-enriched' genes from the 'Hallmark' gene sets 'DNA_REPAIR', 'INTERFERON_ALPHA_RESPONSE' and 'INTERFERON_GAMMA_RESPONSE' and were determined by 'Gene Set Enrichment Analysis' (GSEA) between *Idh1* and CTRL samples. Clustering of rows and columns is based on Euclidean distances. The color scale represents row z-scores.

Taken together, we could show that, functionally, *Idh1*-R132H mutated LT-HSCs show a better engraftment and reconstitution potential in a transplantation setting whereas the lineage output of *Idh1*-R132H mutated LT-HSCs was mostly comparable to CTRL LT-HSCs. At the molecular level, *Idh1*-R132H mutated cells display trends towards an aberrant expression of DNA repair- and inflammation-related genes, although the heterogeneity of *Idh1*-R132H mutated replicates needs to be taken into consideration when interpreting the results.

3.4 Serial Transplantation of *Idh1*-R132H Mutated Hematopoietic Stem Cells

Similar to other regenerative tissues, the age-dependent functional degeneration of the hematopoietic system is closely linked to a decline in HSC function and self-renewal capacity over time. The 'stemness' and self-renewal potential of HSCs is commonly measured by serial transplantations, where bone marrow cells from primary recipient mice are successively transplanted into further lethally irradiated recipients. Previously, it has been shown that the self-renewal capacity and functionality of HSCs decreased with the number of serial transplantations [Harrison et al., 1978, Kamminga et al., 2005]. At the molecular level, several reports identified an accumulation of DNA damage and defects in DNA repair pathways to be critical factors for the functional decline in HSCs [Rossi et al., 2008, Flach et al., 2014, Rossi et al., 2007, Rube et al., 2011, Wang et al., 2012]. Since *Idh1*-R132H mutated HSCs exhibited an increase in DNA damage [Inoue et al., 2016] and an upregulation of DNA repair-associated genes, we hypothesized that *Idh1*-R132H mutated HSCs either exhaust prematurely or, because of the constituent accumulation of additional mutations, are prone to leukemic transformation in a serial transplantation setting.

We initially transplanted 300 $EYFP^+$ LT-HSCs isolated from *Idh1* or CTRL mice together with 300,000 supportive TBM cells into primary recipients and serially transplanted the total bone marrow of primary recipients into secondary and tertiary recipients 20-22 weeks after transplantation (Figure 3-7A). Flow cytometry analysis of donor chimerism (i.e. proportion of $CD45.1^+$ cells) in the bone marrow of primary recipients revealed that, in line with previous experiments, *Idh1*-R132H mutated LT-HSCs are characterized by a better engraftment and repopulation potential as measured by an increased ratio of $CD45.1^+$ cells to CD45.1/2 double-positive cells (Figure 3-7B). In primary recipients, the rate of $CD45.1^+$ cells further expanded in both *Idh1* and CTRL transplanted mice over time, but with CTRL mice constitutively displaying decreased rates compared to *Idh1* transplanted mice. Relative to $CD45.1^+$ cell rates from primary recipients, rates stayed level in both *Idh1* and CTRL secondary and tertiary transplanted mice. It is worth noting, that whereas only little engraftment was observed in two CTRL tertiary recipients, good engraftment was observed in all *Idh1* tertiary recipients, arguing against an *Idh1*-R132H mutation-mediated premature decline in self-renewal capacity. In order to estimate the

frequency of HSCs in each genotype, we analyzed the composition of the bone marrow by flow cytometry at the time of each bone marrow transplantation (22 and 42 weeks). In primary recipients, the frequency of LSK cells and HSCs, especially ST-HSCs, was slightly reduced in *Idh1*-R132H mutated recipients, whereas the frequency of combined MPP3 and MPP4 cells was slightly increased (Figure 3-7C, frequencies normalized to the percentage of overall EYFP$^+$ cells). In secondary recipients, a similar trend of decreased LSK and HSC frequencies in *Idh1*-R132H mutated mice was observed, with ST-HSC frequencies being significantly lower in *Idh1* transplanted mice. Comparable tendencies were noted in primary and secondary recipients when cell frequencies of HSC and MPP populations were normalized to the fraction of EYFP$^+$ LSK cells instead of overall EYFP$^+$ cells (data not shown). Overall, one can conclude that the bone marrow composition in *Idh1* and CTRL mice was rather comparable with only a minor depletion of HSCs upon serial transplantation. However, the functionality of transplanted *Idh1*-R132H mutated HSCs seems not to be compromised as engraftment and reconstitution was observed in all recipients. We furthermore quantified blood production and lineage output in secondary or tertiary recipients and found some variation in tertiary recipients, where *Idh1* transplanted mice displayed trends of decreased B-cell frequencies and increased T-cell and myeloid cell frequencies (Supplemental Figure A-5).

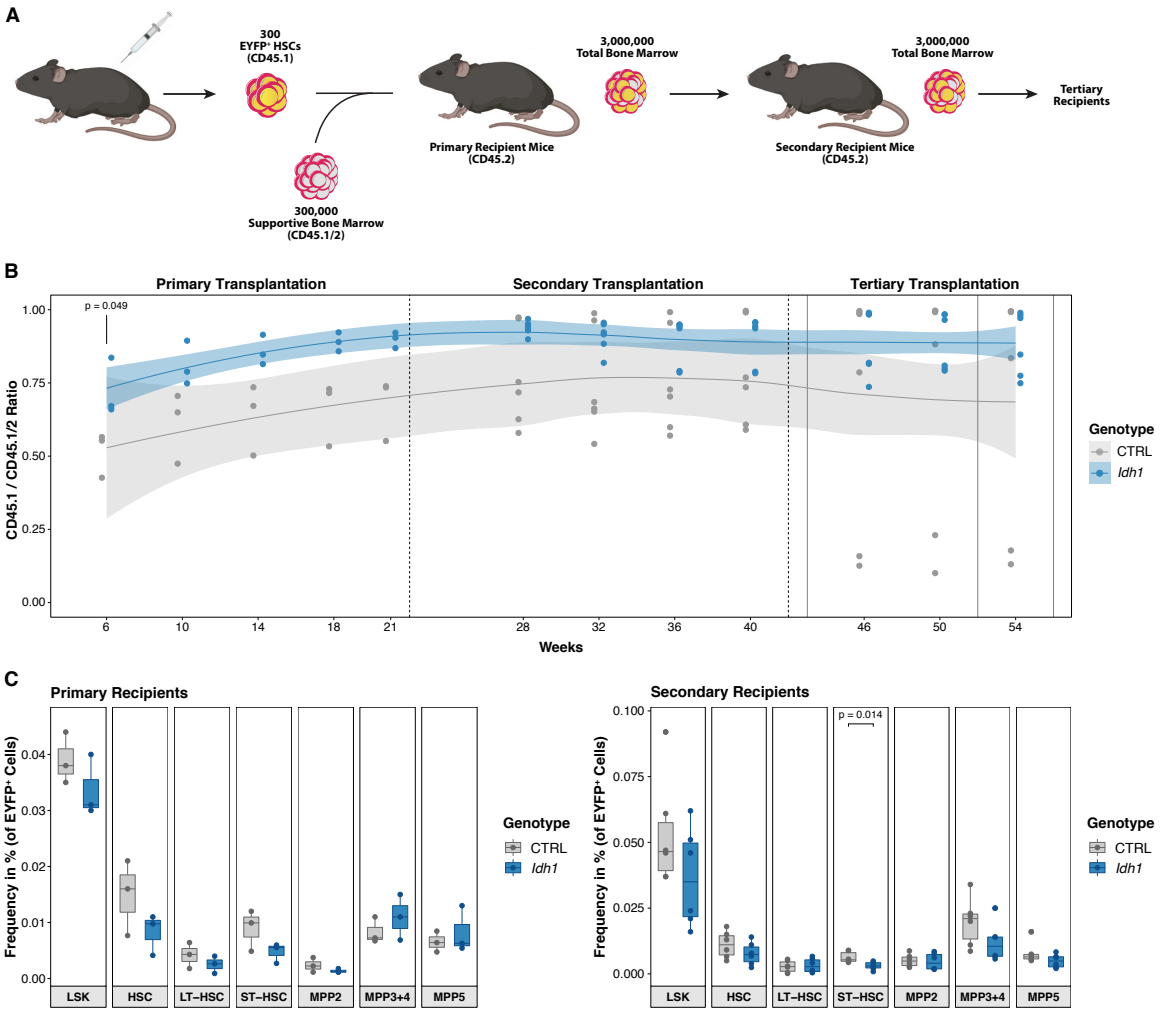


Figure 3-7: The Functionality of *Idh1*-R132H Mutated HSCs is Unaltered in Serial Transplantations.
(A) Schematic illustration of the serial transplantation experimental setup. Donor mice consisted of either *Idh1* or CTRL mice. **(B)** Quantification of CD45.1$^+$ cell rates in the blood of primary, secondary or tertiary transplanted mice by using flow cytometry (primary transplantation n=3, secondary transplantation n = 6, tertiary transplantation n = 6, curves

were smoothened based on a loess (locally estimated scatterplot smoothing) regression model, confidence interval (0.95) plotted around curve, Student's t-test, p-values are indicated for significant results ($p < 0.05$)). **(C)** Quantification of frequencies of indicated cell types in the bone marrow of primary and secondary recipient mice using flow cytometry. Frequencies are calculated in relation to the percentage of overall $EYFP^+$ cells in the bone marrow (primary recipients n = 3, secondary recipients n = 6, Student's t-test, p-values are indicated for significant results ($p < 0.05$)).

Collectively, our data suggests that the self-renewal capacity of *Idh1*-R132H mutated HSCs is unaltered relative to CTRL HSCs upon serial transplantation. Furthermore, no clear indication of oncogenic transformation or disease development could be detected in all transplanted mice so far, implying that *Idh1*-R132H mutations alone do not lead to malignant transformation and that additional genetic or epigenetic hits are necessary for the induction of hematopoietic neoplasms.

3.5 Cooperation of *Idh1*-R132H and *DNMT3A*-R882H Mutations in Hematopoietic Stem Cells

Due to the lack of evidence for malignant transformation in the context of an *Idh1*-R132H mutation as a single hit, we hypothesized that additional mutations co-occuring in *Idh1*-R132H mutated cells might lead to the induction of a leukemic phenotype. *DNMT3A* mutations in general are not only among the most frequent mutations found in individuals with clonal hematopoiesis, they are also commonly detected in AML patients where they are moreover found to co-occur with mutations in *IDH1* (see Figure 1-5C) [Cancer Genome Atlas Research et al., 2013, Papaemmanuil et al., 2016, Desai et al., 2018, Abelson et al., 2018, Shlush et al., 2014]. Although both, *IDH1* and *DNMT3A* mutations are early events in leukemia pathogenesis, they are not known to strongly activate proliferation in affected preleukemic clones [Corces-Zimmerman et al., 2014]. Nevertheless, we speculated that the induction of both an *Idh1*-R132H and a *DNMT3A*-R882H mutation in the same HSC might synergistically alter the epigenetic landscape of affected cells in such a way that it is more permissive to leukemic transformation. To address this hypothesis, we made use of a mouse model, which conditionally expresses the *DNMT3A*-R882H mutation upon activation of Cre (Figure 3-8A). We crossed this conditional *DNMT3A*-R882H knock-in mouse model with our conditional *Idh1*-R132H knock-in mouse model to obtain inducible *Idh1*-R132H *DNMT3A*-R882H double-mutant mice. For all of the experiments described hereafter, mice heterozygous for the *Idh1*-R132H and the *DNMT3A*-R882H allele as well as for the *Scl*-CreERT2 transgene and the *Rosa26*-EYFP allele were injected with tamoxifen and used as donor mice for transplantations (genotype $Idh1^{+/R132H}$, $DNMT3A^{+/R882H}$, Scl-$CreERT2^{+/d}$, $Rosa26$-$EYFP^{+/d}$, denoted as *Idh1-DNMT3A* mice). Mice heterozygous for the *DNMT3A*-R882H allele, the *Scl*-CreERT2 transgene and the *Rosa26*-EYFP allele but not for the *Idh1*-R132H allele were used as *DNMT3A*-R882H single-mutant control mice (genotype $Idh1^{+/+}$, $DNMT3A^{+/R882H}$, Scl-$CreERT2^{+/d}$, $Rosa26$-$EYFP^{+/d}$, denoted as *DNMT3A* mice)

Similar to the serial transplantation experiments described earlier, we chose an experimental setting where we transplanted 300 $EYFP^+$ LT-HSCs from tamoxifen-treated donor mice together with 300,000 supportive TBM cells into lethally irradiated recipients. To eliminate the possibility that effects caused by the mutations arise only in aged mice after long latency times, we extended the monitoring period to over a year with drawing blood samples at regular intervals and performing a detailed analysis of the bone marrow

composition at around 62 weeks post-transplantation. Consistent with previous results, we observed a better engraftment with higher CD45.1⁺ cell rates not only in *Idh1* and *Idh1-DNMT3A* transplanted mice, but also in *DNMT3A* transplanted mice compared to CTRL mice (Figure 3-8B). Regarding the lineage output of transplanted LT-HSCs, tendencies towards a reduction in B- and T-cell frequencies and a concomitant increase in myeloid cell frequencies were detected in *Idh1*-R132H single-mutant mice especially at the end of the monitoring period one year after transplantation. In *Idh1-DNMT3A* mice, frequencies of myeloid cells were comparable to frequencies of CTRL transplanted mice, suggesting that an additional *DNMT3A*-R882H mutation does not elicit synergistic effects, but rather reverts the effects observed in *Idh1*-R132H single-mutated mice. Analysis of the overall blood counts of transplanted mice, however, revealed that neutrophil and monocyte levels were markedly elevated in *Idh1-DNMT3A* mice at early analysis time points (Supplemental Figure A-6).

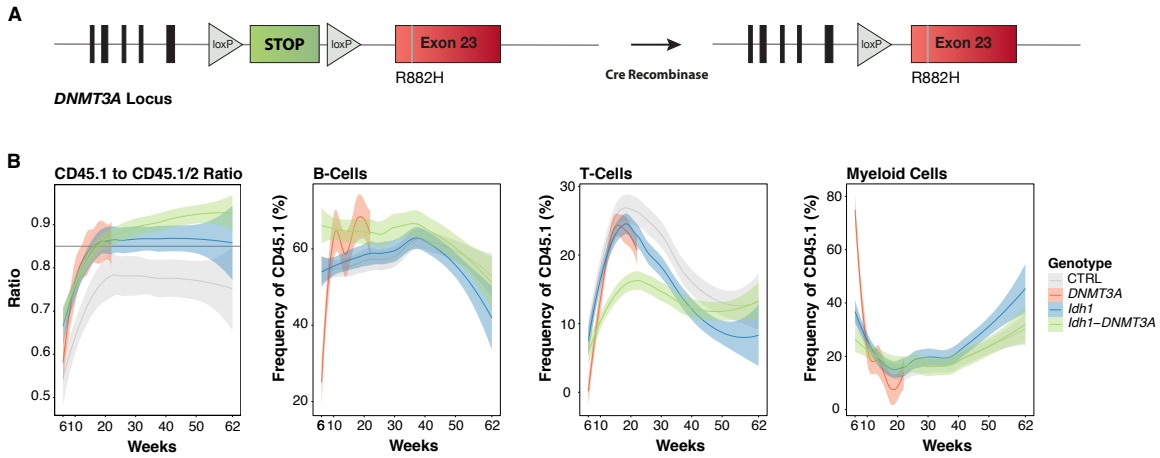


Figure 3-8: Blood Characteristics of *Idh1*-R132H *DNMT3A*-R882H Double-Mutant Mice are Similar to Control and Single-Mutant Mice.
(A) Schematic illustration of the *DNMT3A*-R882H mouse model. Upon activity of Cre, a lox-flanked stop codon in front of the R882H mutated exon 23 is excised and the *DNMT3A*-R882H mutation is expressed. **(B)** Quantification of overall CD45.1⁺ cell rates as well as terminally differentiated CD45.1⁺ B-cells, T-cells and myeloid cells in the blood of transplanted mice by using flow cytometry (CTRL n = 19, *Idh1* n =21, *Idh1-DNMT3A* n= 18, *DNMT3A* n =12, two to four independent transplantations per genotype, curves were smoothened based on a loess (locally estimated scatterplot smoothing) regression model, confidence interval (0.95) plotted around curve) (transplantation and analysis of *DNMT3A* mice was performed by Dr. Sina Stäble and Dr. Natasha Anstee).

Following the monitoring phase with periodic blood analysis, mice were sacrificed and the cellular composition of the bone marrow, including frequencies of hematopoietic stem and multipotent progenitor cells as well as that of committed progenitors (Lineage⁻, Sca1⁻, cKit⁺ (LS-K cells)), and differentiated cells was comprehensively analyzed by flow cytometry. Within the HSC and MPP compartment, a trend towards an expansion of the LSK and MPP3 + MPP4 compartment was observed in *Idh1-DNMT3A* mice, whereas all other genotypes did not display any substantial abnormalities (Figure 3-9A). The committed progenitor / LS-K compartment, encompassing CMPs, GMPs and MEPs, of mutated mice did not reveal any significant changes in cell frequencies, yet tendencies of reduced numbers of B-cells and increased numbers of myeloid cells in the bone marrow of *Idh1* and *Idh1-DNMT3A* transplanted mice were detected.

The fact that EYFP⁺ cells were detected in the bone marrow 60 weeks after transplantation suggests that no negative selection or selective depletion of *Idh1*-R132H or *DNMT3A*-R882H mutated cells is seen in any of the genotypes. Nevertheless, to rule out the possibility that the *Idh1*-R132H transgene is silenced in hematopoietic cells over

time without affecting EYFP expression, we quantified the levels of D2HG in bone marrow cell lysates of *Idh1* and CTRL transplanted mice about one year post-transplantation. Normalized to the overall protein content in the cell lysate, the concentration of D2HG was elevated on average 7.5-fold in *Idh1*-R132H mutated mice, proving the continued transcriptional activity of the transgene (Figure 3-9B). As expected, the level of increase in D2HG levels measured in bone marrow lysates highly correlated with the percentage of EYFP$^+$ cells in the bone marrow (Figure 3-9C). On the other hand, the frequency of EYFP$^+$ HSCs in the bone marrow was inversely correlated with D2HG levels. While a larger number of mice is necessary to corroborate this correlation, it implies that the amount of D2HG present could modulate the expansion or depletion of the stem cell pool. Yet, the overall number of EYFP$^+$ HSCs (normalized to the number of EYFP$^+$ cells) in the bone marrow was unaltered in *Idh1*-R132H mutant mice (Figure 3-9A). We also investigated whether secreted D2HG likewise affects frequencies of EYFP$^-$ (i.e. supportive bone marrow) HSCs which do not express an *Idh1*-R132H mutation, but found D2HG levels not correlate with EYFP$^-$ HSC counts (Pearson correlation coefficient R = 0.21, data not shown).

In summary, we found some minor deviations in cell frequencies in *Idh1*-R132H *DNMT3A*-R882H double-mutated mice, but did not observe any indications of leukemic transformation in the context of a co-occurrence of both mutations.

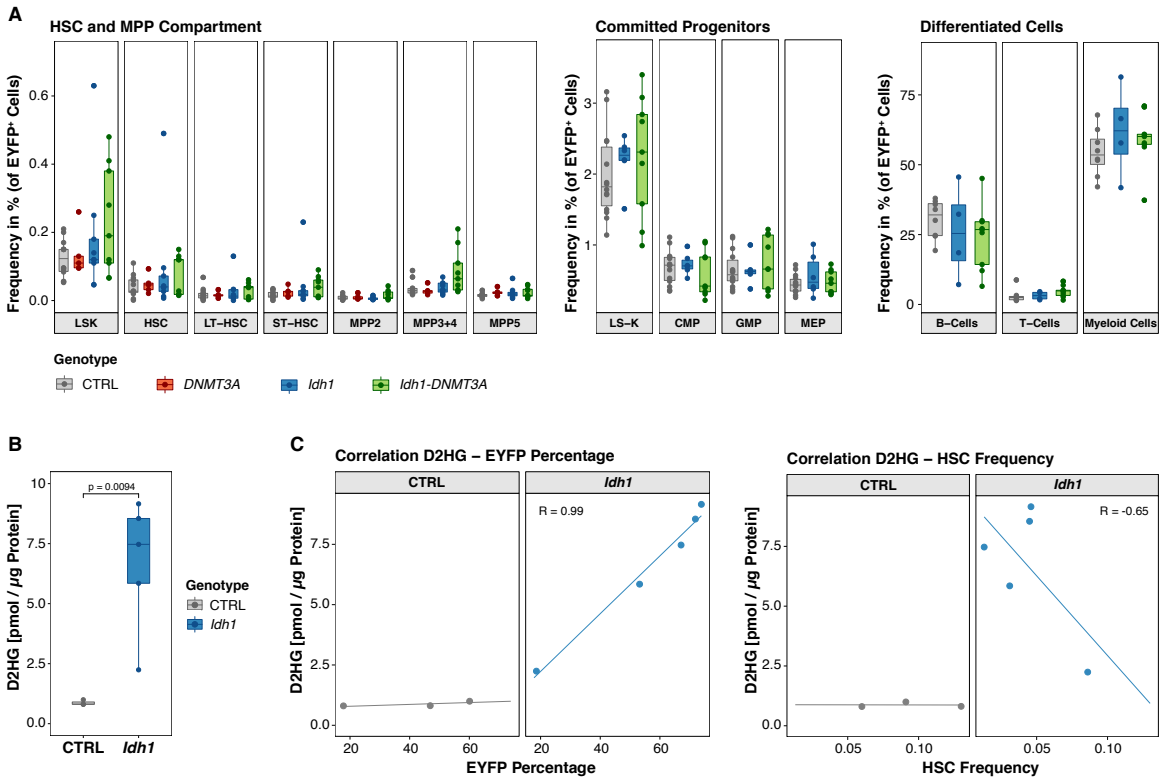


Figure 3-9: A Co-Occurrence of a *Idh1*-R132H and a *DNMT3A*-R882H Mutation Only Leads to Minor Changes in Bone Marrow Composition of LT-HSC Transplanted Mice.
(A) Quantification of frequencies of indicated cell types in the bone marrow of transplanted recipient mice using flow cytometry. Frequencies are calculated in relation to overall EYFP$^+$ cells in the bone marrow. Mice were sacrificed and analyzed 54 – 62 weeks post-transplantation (CTRL n = 8-14, *DNMT3A* n = 5, *Idh1* n = 4-9, *Idh1-DNMT3A* n = 9) (transplantation and analysis of *DNMT3A* mice was performed by Dr. Sina Stäble and Dr. Natasha Anstee). **(B)** Quantification of D2HG levels in the bone marrow cell lysate of *Idh1* and CTRL transplanted mice sacrificed 62 weeks after transplantation (CTRL n =3, Idh1 n = 5, Student's t-test). **(C)** Correlation of measured D2HG levels in bone marrow cell lysates and percentages of EYFP$^+$ cells or HSC frequencies in the bone marrow of CTRL or *Idh1* transplanted mice (CTRL n = 3, *Idh1* n = 5, lines were calculated using a generalized linear regression model, the Pearson correlation coefficient is depicted).

3.6 *Idh1*-R132H Mutations Cause Increased Monocyte Levels in Lineage$^-$ Transplanted Mice

Previous results have shown that in our hands, the *Idh1*-R132H mutation alone or in combination with a *DNMT3A*-R882H mutation does not lead to the development of hematological malignancies. Instead, *Idh1*-R132H mutations seem to evoke a differentiation bias towards the myeloid lineage at the expanse of the lymphoid lineage (see Figures 3-8B, 3-9A). In order to characterize mutation-mediated aberrant differentiation in more detail, we generated full chimeric mice, in which the entire hematopoietic system is replaced by mutated cells. This is in contrast to the transplantation experiments described previously, where the use of supportive bone marrow cells was necessary to ensure the immediate survival of the mice until transplanted LT-HSCs have engrafted and ensured the permanent production of differentiated blood cells. Even though the ratio of supportive bone marrow cells and its contribution to blood production steadily decreased over time, it is conceivable, that differentiation patterns of transplanted stem cells vary in a more drastic and stress-imposing scenario where the whole hematopoietic system immediately has to be replaced by mutated cells and patterns are not obscured by the intermediate contribution of supportive bone marrow.

To overcome this requirement for supportive bone marrow, we employed a transplantation setting where we transplanted lineage negative (Lin$^-$) cells that consist of all hematopoietic cell types except for terminally differentiated cells, so that downstream progenitors ensure the immediate blood production directly after transplantation, while long-term blood production is driven by HSCs after their engraftment (Figure 3-10A). Transplanted recipient mice were sacrificed and analyzed four months after transplantation, assuming that at this time point a stable engraftment and sustained blood production of all transplanted HSCs is reached. In addition to a surface marker-based analysis of cell types in the bone marrow by flow cytometry, we complementarily performed single-cell RNA-sequencing of selected bone marrow populations in order to comprehensively characterize hematopoietic differentiation trajectories at the molecular level (for a detailed description see chapter 3.7). To review the impact of an *Idh1*-R132H mutation on the myeloid compartment and myeloid differentiation patterns in more detail, we scrutinized this compartment by discriminating between terminally differentiated monocytes and granulocytes.

Within the HSC and MPP compartment differences in cell type frequencies were minute with *Idh1-DNMT3A* mice displaying a slight increase in LSK frequency and both *Idh1* and *Idh1-DNMT3A* mice featuring a significant reduction in ST-HSC frequency (Figure 3-10B). Furthermore, counts of LS-K and the therein enclosed CMP population were elevated in *Idh1* and *Idh1-DNMT3A* transplanted mice (Figure 3-10C). Strongest disparities, however, were observed in differentiated cells, where the bone marrow of both *Idh1* and *Idh1-DNMT3A* transplanted mice contained dramatically increased levels of terminally differentiated monocytes (Figure 3-10D).

Summarizing, in the Lin$^-$ transplantation experiments, we observed substantial differences between *Idh1* or *Idh1-DNMT3A* mice and CTRL or DNMT3A mice that were not detected in other transplantation settings, suggesting that the lack of supportive bone marrow might increase selective pressures and thus leads to pronounced phenotypes as mutated cells are forced to completely reconstitute the entire hematopoietic system after transplantation. Consistent with previous findings, *Idh1*-R132H mutations alone or in

combination with *DNMT3A*-R882H mutations lead to a myeloid differentiation bias, which manifests in the expansion of bone marrow monocytes.

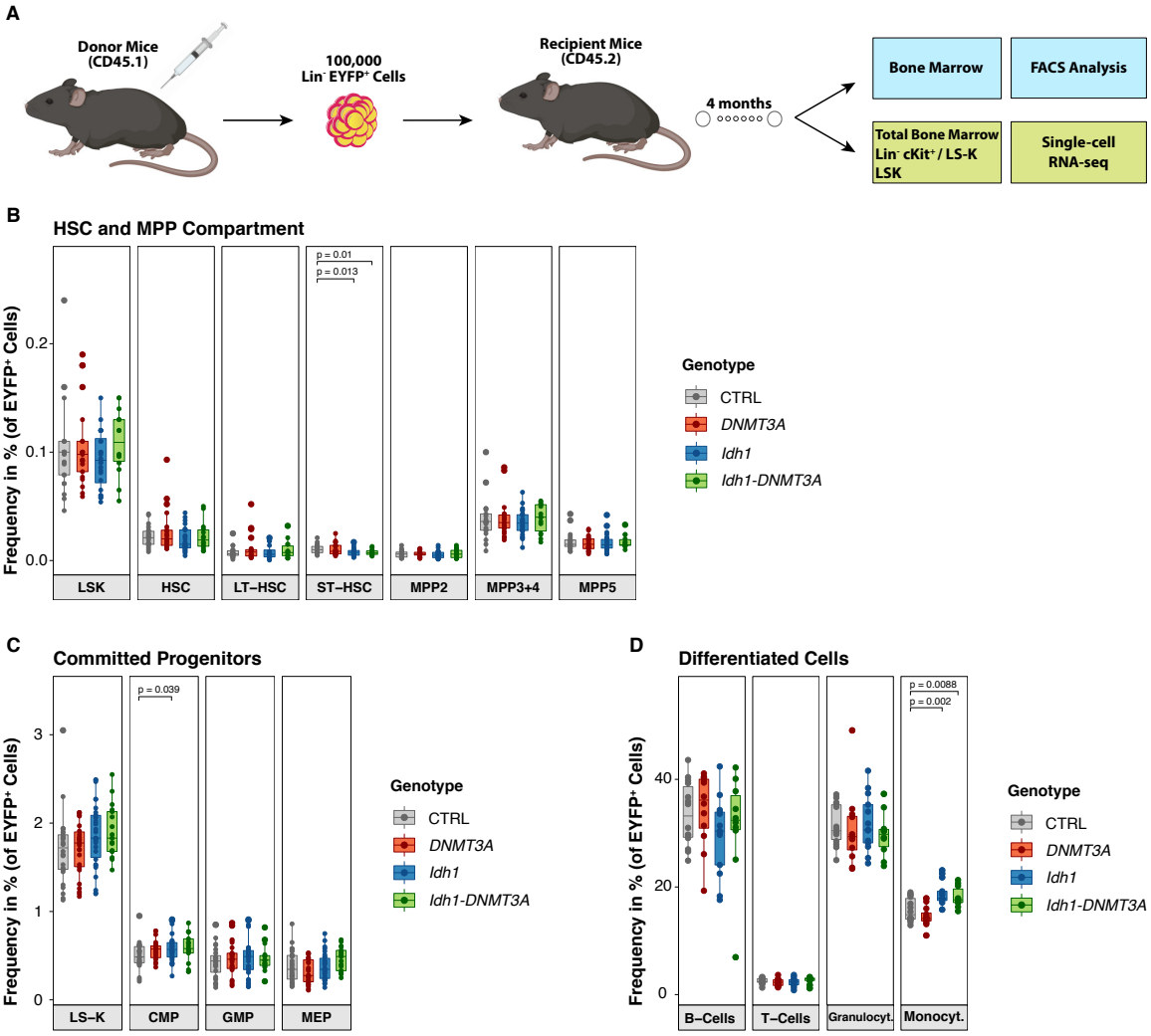


Figure 3-10: *Idh1* and *Idh1-DNMT3A* Lineage Negative Transplanted Mice Feature Increased Monocyte Counts in the Bone Marrow.
(A) Schematic illustration of the Lin⁻ transplantation setting. Donor mice consisted of tamoxifen-treated CTRL, *DNMT3A*, *Idh1* or *Idh1-DNMT3A* mice. **(B, C, D)** Flow cytometry based quantification of indicated cell types in the bone marrow of transplanted mice four months after transplantation (CTRL n = 14-28, *DNMT3A* n = 12-26, *Idh1* n = 13-30, *Idh1-DNMT3A* n = 10-18, at least three independent transplantations, Student's t-test, p-values are indicated for significant results ($p < 0.05$)).

3.7 Single-Cell RNA-Sequencing Identifies Hematopoietic Differentiation Trajectories

In recent years, high-throughput single-cell RNA-sequencing (scRNA-seq) approaches have challenged the classical hierarchical model of hematopoietic differentiation. This technology provides a high-resolution transcriptome map of individual cells that can be used for the definition of cellular states and their transitions as well as the identification

of cellular priming and lineage commitment decisions [Watcham et al., 2019, Laurenti and Gottgens, 2018]. Single-cell gene expression profiles thereby provide a transcriptomic snapshot of hundreds to thousands of cells traversing along differentiation trajectories. Ordering cells based on their gene expression profiles results in the reconstruction of differentiation trajectories from immature stem and progenitor cells towards mature differentiated blood cells [Watcham et al., 2019, Laurenti and Gottgens, 2018]. Studies of the hematopoietic system using scRNA-seq have transformed our understanding of the different cellular compartments, unraveling unanticipated heterogeneity within flow cytometry defined 'homogenous' cell populations, with different subtypes reflecting diverging lineage priming and progressive commitment along trajectories [Velten et al., 2017, Paul et al., 2015, Olsson et al., 2016, Villani et al., 2017, Rodriguez-Fraticelli et al., 2018]. While flow cytometry-based approaches are restricted to a limited set of predefined markers, scRNA-seq quantifies the expression of thousands of genes per cell, thereby providing unprecedented detail, which not only can be harnessed to define cell states, but also to understand which cell states are predominantly affected by underlying mutations and how these mutations drive aberrant differentiation on a molecular level.

In order to reconstruct the transcriptomic landscapes of the hematopoietic system in the context of *Idh1*-R132H, *DNMT3A*-R882H and co-occurring *Idh1*-R132H and *DNMT3A*-R882H mutations, we generated a multi-tiered single-cell gene expression map from Lin^- transplanted mice at the time of stable engraftment (i.e. 20 weeks post transplantation). By combining three different parts of the hematopoietic system ranging from stem and progenitor cells (i.e. LSK cells) over committed progenitors (i.e. LS-K cells) to more differentiated cells (i.e. $CD45^+$ TBM cells), we ensured to cover cells throughout the entire hematopoietic hierarchy (Figure 3-11A). This approach enables us to investigate the effects of the *Idh1*-R132H and *DNMT3A*-R882H mutations on gene expression profiles in virtually all hematopoietic cell types. Following cell sorting, single-cell gene expression profiles were generated using the droplet-based 10X Genomics platform using unique molecular identifiers (UMIs) to barcode individual RNA / cDNA molecules. For all samples, around 10,000 cells were loaded on the Chromium chip for library generation, the LSK-CTRL sample being the only exception with only 3,366 cells being loaded (Figure 3-11B). Out of the loaded cells, between 3,000 and 5,000 cells (LSK-WT: 1,108 cells) were captured per sample, yielding transcriptome information for a total of 45,971 single cells (Figure 3-11C). Initial filtering was performed based on the number of transcript counts and genes per cell as well as on the percentage of mitochondrial genes. Here, cells with low transcript and gene counts per cell or with high mitochondrial DNA contents likely represent damaged or dying cells, whereas cells with an unnatural high number of transcript and gene counts per cell are indicative of doublets. After filtering, 42,626 cells remained in the dataset. Across samples, we observed some variation in the number of transcript counts, genes and percentages of mitochondrial genes per cell with TBM samples being characterized by a lower number of detected genes and transcript counts per cell and a bimodal distribution of percentages of mitochondrial genes, probably reflecting different cell types with different mitochondria content (Figure 3-11D). For all genotypes, the highest numbers of genes per cell were detected in LS-K samples with between 3,000 and 4,000 genes being identified.

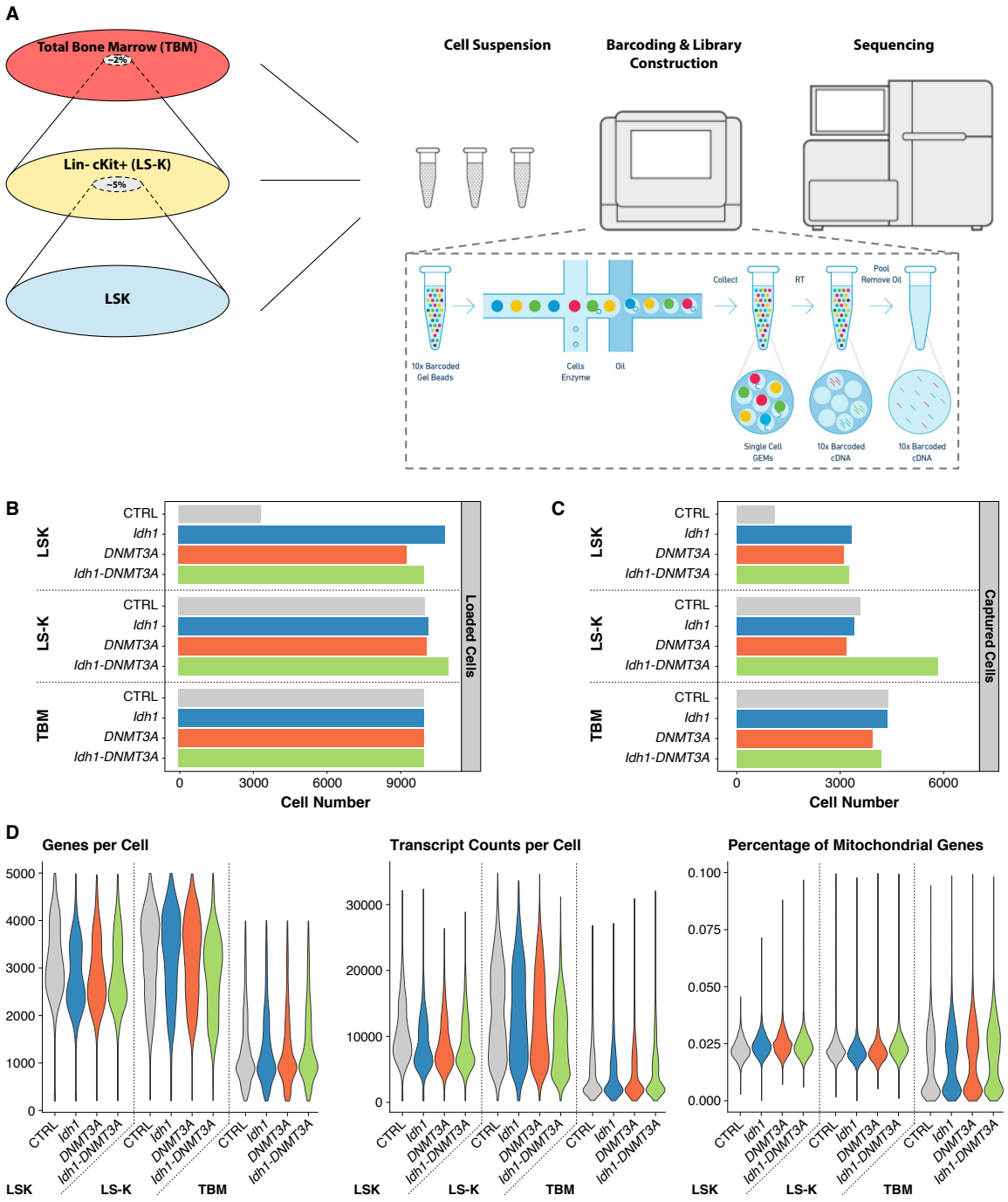


Figure 3-11: Generating a Multi-Layered Single-Cell Transcriptomic Landscape of the Hematopoietic Compartment. **(A)** Experimental workflow for the generation of single-cell transcriptomic landscapes of the hematopoietic compartment. For each genotype, three cell layers (CD45⁺ TBM, LS-K, LSK) were isolated by flow cytometry and scRNA-seq libraries generated using the 10X Genomics platform (adapted from 10X Genomics) (libraries were generated by Dr. Mark Hartmann, Maximilian Schönung and Katharina Bauer) **(B + C)** Overview of loaded and captured cell numbers per sample. Alignment, barcode and UMI counting was done using the 'Cell Ranger' software package from 10X Genomics (data processing was performed by Abdelrahman Mahmoud) **(D)** QC metrics of single-cells after data processing and filtering using the 'Cell Ranger' and 'Seurat' packages. Cells were filtered based on the number of detected genes (threshold: > 200 and < 5000 genes per cell) and the percentage of mitochondrial genes (threshold: < 10 %) (data processing was performed by Abdelrahman Mahmoud).

Subsequent to processing and quality control (QC) filtering, the R-package 'Seurat' was used for data normalization, scaling and clustering of cells. For visualization of single-cell transcriptome profiles, the 'Uniform Manifold Approximation and Projection' (UMAP)

algorithm was applied to reduce the dimensionality of the dataset and to compute cell-to-cell distances by which high-dimensional gene expression profiles of each single-cell can be distributed and ordered in a two-dimensional space. Here, annotation of the three bone marrow layers revealed a hierarchical ordering of LSK, LS-K and TBM populations with a high overlap between LSK and LS-K samples but a weak overlap of TBM and LSK or LS-K samples (Figure 3-12A). Our UMAP-based representation of the data is in line with previously published scRNA-seq data suggesting a continuous rather than a stepwise differentiation process in the hematopoietic system [Velten et al., 2017, Nestorowa et al., 2016, Macaulay et al., 2016]. Annotation of the four individual genotypes illustrated a mostly even distribution of all genotypes, yet enrichment or depletion of specific genotypes in certain regions of the map were observed (Figure 3-12B). In light of the fact, that none of the mutated mice displayed indications of leukemia development or completely aberrant differentiation patterns, a similar representation of genotypes was expected at this point (see Supplemental Table A-1 for detailed peripheral blood and bone marrow composition of mice used in the scRNA-seq experiment).

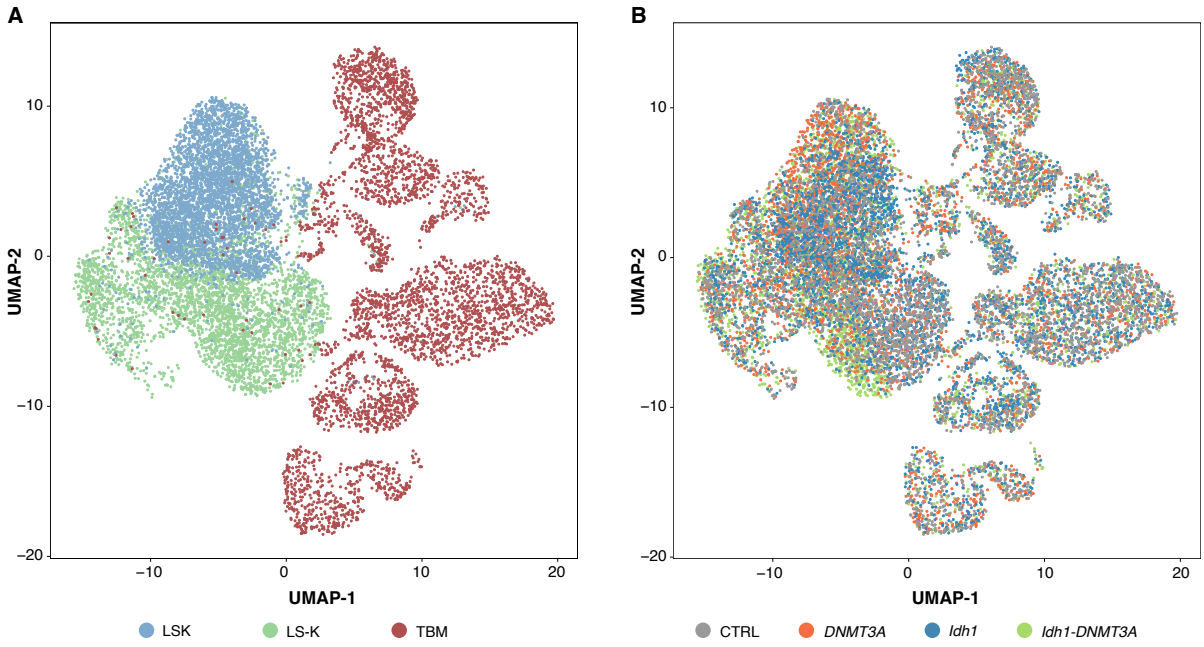


Figure 3-12: Visualization of Single-Cell Transcriptomes in a Two-Dimensional Space Reveals Hierarchical Structures and an Overlap of Sorted Cell Compartments.
(A) UMAP-based dimensionality reduction and visualization of single-cell transcriptomes with an annotation of isolated cell layers. Hierarchical structures and a high overlap between LSK and LS-K populations are seen. For representative purposes, only 4,000 cells per compartment are depicted (initial data processing was performed by Abdelrahman Mahmoud). **(B)** UMAP-based dimensional reduction and visualization of single-cell transcriptomes with an annotation of individual genotypes. For representative purposes, only 4,000 cells per genotype are depicted (initial data processing was performed by Abdelrahman Mahmoud).

Clustering of cells was performed using a shared nearest neighbor (SNN) algorithm implemented in the 'Seurat' package, yielding 49 cell clusters in total. In order to correlate these 49 clusters with specific hematopoietic cell populations, the expression of cell type-defining genes was analyzed (Figure 3-13A) and cluster-specific markers were determined by calculating DE genes between the cluster of interest and all other remaining cells (see Material and Methods, Section 5.11.2 for detailed description). Individual clusters then were annotated manually by comparing cluster-defining genes with known hematopoietic marker genes (Figure 3-13B, list of marker genes see Supplemental Table A-2). As many

distinct clusters shared the expression of the same characteristic cell type-defining genes, annotated hematopoietic cell types frequently consisted of multiple individual clusters. For instance, the ST-HSC population was aggregated from three individual clusters which all displayed high expression levels of ST-HSC-defining marker genes. Eventually, this resulted in a total of 27 different annotated cell types derived from 49 previous clusters.

Starting from LT-HSCs at top of the hematopoietic hierarchy, four diverging directions of differentiation emerged in the LSK and LS-K compartments, leading towards megakaryocytic, erythroid, myeloid and lymphoid cell fates. Although terminally differentiated cell types derived from TBM samples clustered separate in the UMAP, they were distributed in proximity to their respective lineage progenitors.

In summary, by isolating LSK, LS-K and $CD45^+$ TBM cells from transplanted mice, we have generated a multi-tier transcriptome map of the entire murine hematopoietic system. Detailed manual annotation of cell types allowed us to reconstruct trajectories into megakaryocytic, erythroid, myeloid and lymphoid lineages, which provided the basis for a detailed characterization of the molecular effects of *Idh1*-R132H and *DNMT3A*-R882H mutations throughout the entire hematopoietic system.

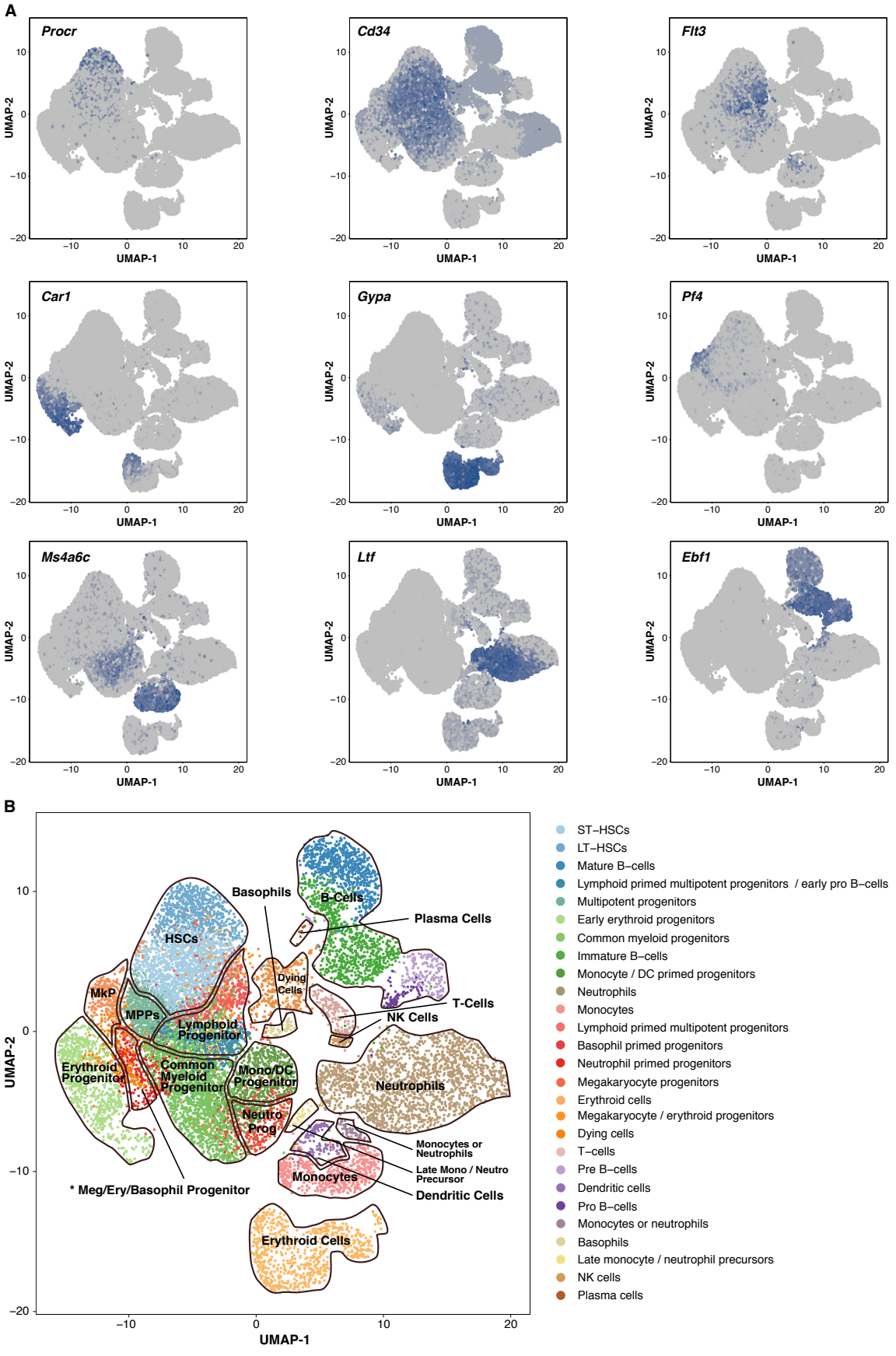


Figure 3-13: Single-Cell RNA-Sequencing Allows Reconstruction of Differentiation Trajectories.
(A) Normalized expression of exemplary hematopoietic cell type-defining marker genes highlighted within the UMAP-based representation of single-cell transcriptomic data. The color gradient from grey to dark blue indicates the level of

expression (grey = no expression, dark blue = high expression). Certain hematopoietic marker genes are only expressed in certain clusters, thereby allowing the annotation of clusters. **(B)** UMAP-based dimensional reduction and visualization of single-cell transcriptomes with an annotation of hematopoietic cell types. Clusters were annotated manually by investigating the expression of hematopoietic marker genes within cluster-defining genes. For representative purposes, only 15,000 cells are depicted (annotation of clusters was performed in collaboration with Dr. Simon Haas and Dr. Mark Hartmann).

3.8 *Idh1*-R132H *DNMT3A*-R882H Mutated Mice Feature an Aberrant Myeloid Progenitor Compartment

In order to investigate potentially diverging compositions of cell types in each genotype, we analyzed the contribution of individual genotypes to each of the annotated cell clusters. By comparing normalized cell numbers per sample and genotype within a given cluster, we found genotype-specific variation of cell type proportions to arise mainly in the LS-K and to a lesser extent also in the LSK compartment (Figure 3-14A). Differentiated cells represented in the TBM samples did not show skewed distribution across genotypes. The highest variability of genotype-specific cell type distribution was observed in clusters that were annotated as common myeloid progenitor (CMP) cells. In total, five cell clusters where annotated to the CMP compartment. Two out of the five CMP clusters were dominated by cells coming from *Idh1-DNMT3A* mice (common myeloid progenitor cluster #4 and #5), while other genotypes only marginally contributed to these clusters (Figure 3-14A, for normalized cell frequencies per cluster and genotype see Supplemental Table A-3). On the other hand, the monocyte / dendritic cell primed progenitor- (MoDCP) and the neutrophil primed progenitor (NeuP) clusters were depleted for cells from *Idh1-DNMT3A* mice. Within the LSK compartment, *Idh1*-R132H mutated mice displayed a slight tendency towards increased numbers of lymphoid primed progenitors when compared to other genotypes, but decreased numbers of LT-HSCs (Figure 3-14A, Supplemental Table A-3).

A focused UMAP representation of CMPs, MoDCPs and NeuPs clearly confirmed elevated levels of CMPs but diminished levels of MoDCPs and NeuPs in *Idh1-DNMT3A* mice (Figure 3-14B). Highlighting of the individual CMP subclusters indicates that CMP clusters #4 and #5 show the highest enrichment of *Idh1-DNMT3A* cells, whereas CMP cluster #1, #2 and #3 are not as strongly enriched for double-mutated cells (Supplemental Figure A-7).

In relation to our previous FACS-based analysis of the bone marrow composition, a minor trend of an expansion of the CMP population was also detected in *Idh1-DNMT3A* mice based on flow cytometric quantification of cell populations (see Figure 3-10C), thereby validating genotype-specific dynamics within this population. For *Idh1*-R132H single-mutant mice, our FACS analysis indicated rather a depletion of ST-HSCs (see Figures 3-7C and 3-10C) instead of LT-HSCs (as suggested by the scRNA-seq data), indicating that cell populations annotated by transcriptome profiles and cell populations defined by surface markers only overlap to a limited extent. Moreover, in our scRNA-seq dataset we failed to detect increased numbers of mature monocytes in TBM samples of *Idh1* and *Idh1-DNMT3A* mice as seen in flow cytometry-based quantifications. However, due to genotype-specific differences in particular in the composition of the myeloid progenitor cell compartment, we further focused on characterizing these cell types in more detail.

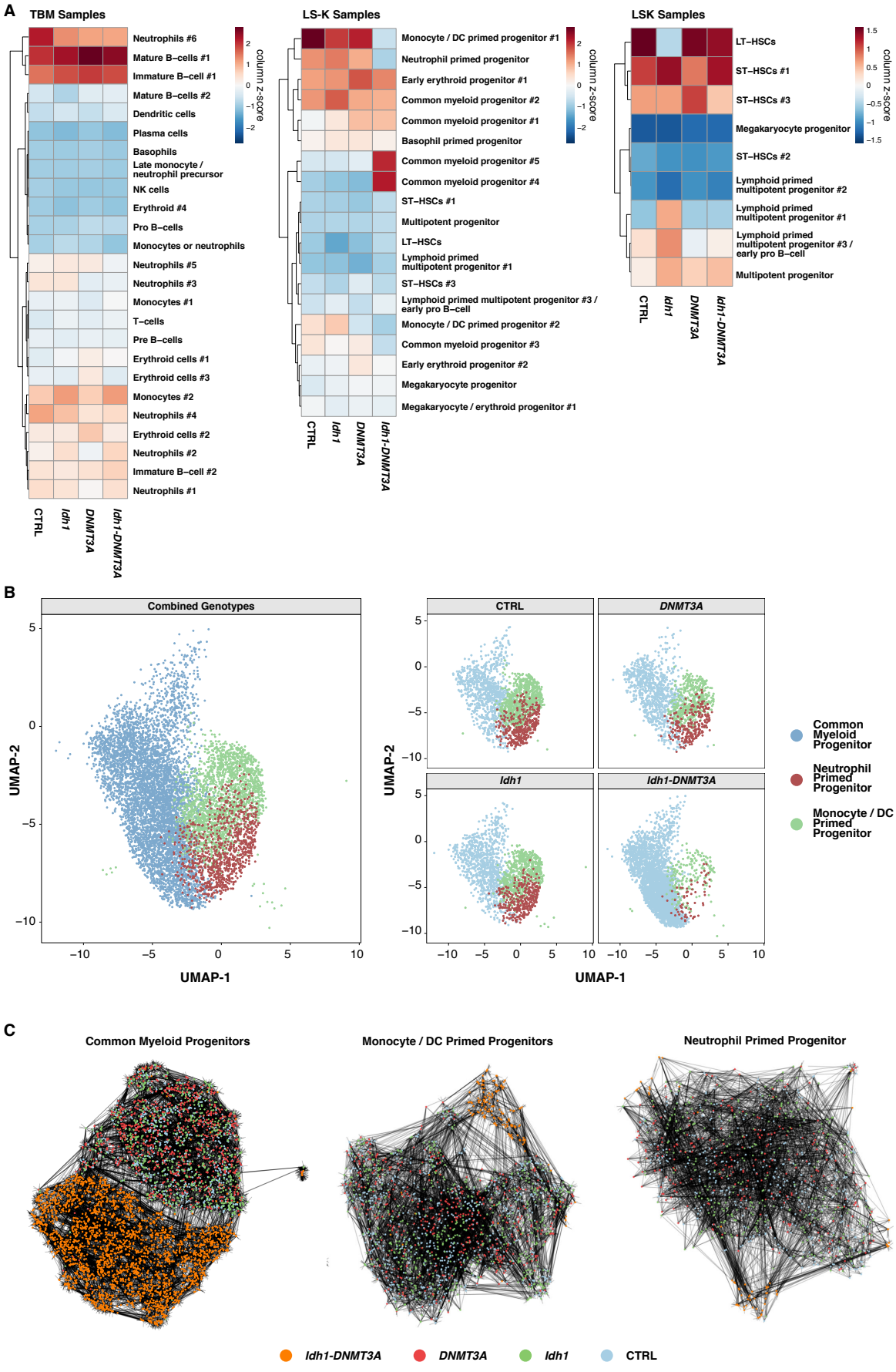


Figure 3-14: ***Idh1-DNMT3A*** **Mice Feature Increased Frequencies of Common Myeloid Progenitor Cells.**

(A) Heatmap depicting the relative distribution of cell numbers from individual genotypes within annotated clusters. Color scale depicts column z-scores, thereby disregarding differential cell numbers that were initially captured per sample. Only clusters containing more than 1 % of the total sample cell number are shown. **(B)** Focused UMAP representation of CMP, MoDCP and NeuP populations per genotype. **(C)** K-nearest neighbor analysis of CMP, MoDCP and NeuP annotated clusters (analysis was performed by Abdelrahman Mahmoud).

K-nearest neighbor (kNN) analysis of the three myeloid progenitor populations (CMPs, MoDCPs and NeuPs) revealed that for each cell type *Idh1*-R132H *DNMT3A*-R882H mutated cells cluster away from CTRL, *Idh1*-R132H or *DNMT3A*-R882H single-mutated cells, further corroborating that combined mutations strongly affect the transcriptome of myeloid progenitor cells (Figure 3-14C). When calculating sample distances of each individual myeloid progenitor cell cluster per genotype, especially CMP cluster #4 and cluster #5 cells from *Idh1-DNMT3A* mice clustered away from CMP #4 and CMP #5 cells from other genotypes. Here, *Idh1-DNMT3A* CMP#4 cells clustered together with MoDCP cells whereas *Idh1-DNMT3A* CMP#5 cells clustered together with NeuP cells (Figure 3-15), suggesting these CMP clusters to display most aberrant transcriptomic profiles in relation to corresponding CMP cluster from CTRL or single-mutant mice.

Collectively, our data implies that *Idh1*-R132H *DNMT3A*-R882H double-mutants display a dysregulated expression of myeloid differentiation programs mediating the neutrophil versus monocyte trajectory. Aberrant expression profiles lead to an increase in CMP cell numbers and predominantly affect CMP sub-cluster #4 and #5, which, based on sample distances, deviate from corresponding clusters of other genotypes.

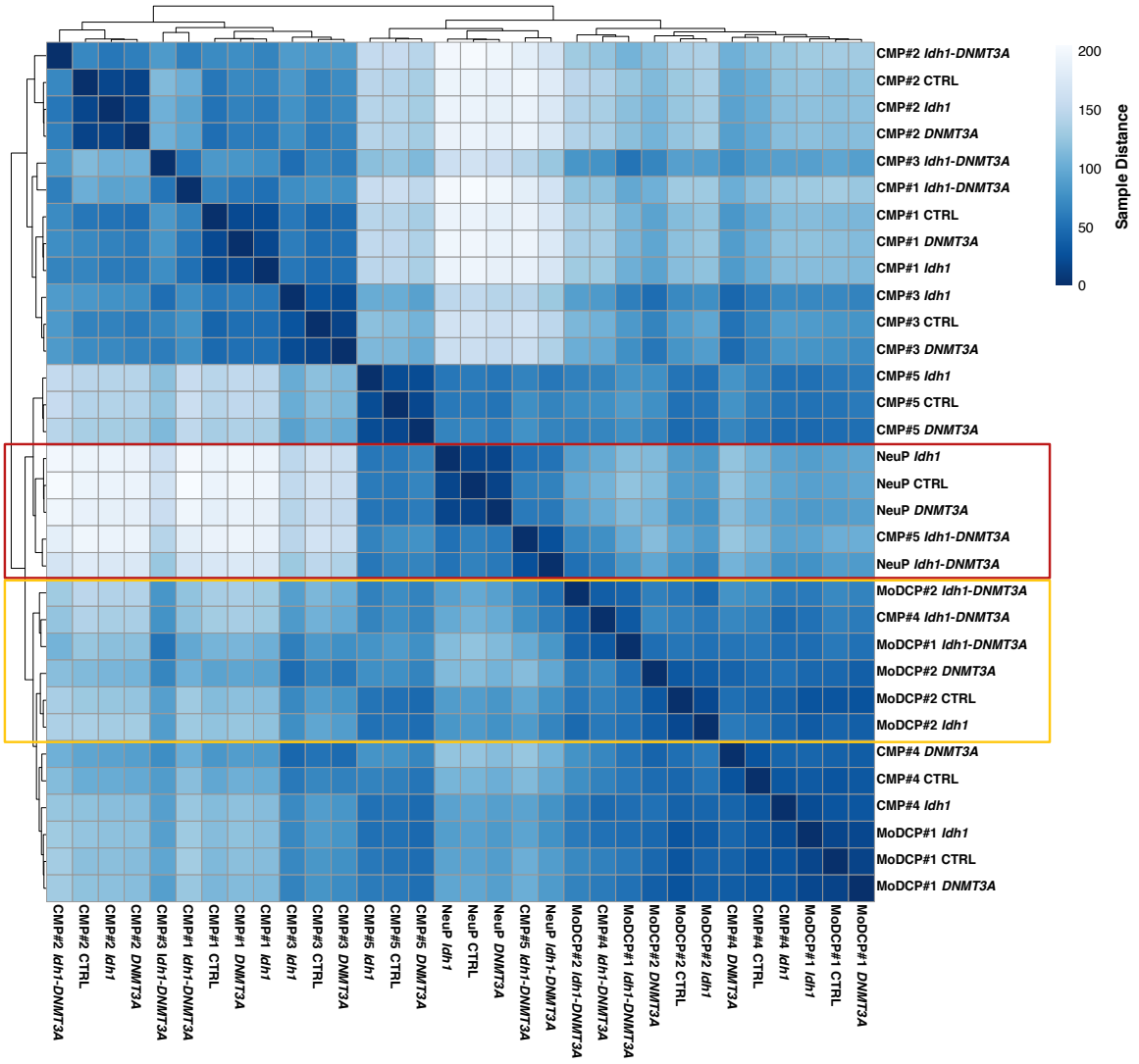


Figure 3-15: Genotype-Specific Sample Distances of Myeloid Progenitor Clusters Identify Aberrant Expression Profiles in *Idh1*-R132H *DNMT3A*-R882H Mutated Mice at the CMP Level.
Heatmap depicting sample distances for each annotated myeloid progenitor cluster. Distances were calculated based on the average expression of all cells from one specific genotype within the indicated cluster. Calculation of distances and clustering of samples was computed using Euclidean distances. UMAP-based visualization of CMP, MoDCP and NeuP subclusters is shown in Suplemental Figure A-7. *Idh1-DNMT3A* CMP #4 cells cluster together with MoDCP cells (yellow box), whereas *Idh1-DNMT3A* CMP#5 cells cluster together with NeuP cells (red box).

3.9 Diffusion Analysis of scRNA-seq Data Characterizes Aberrant Differentiation Trajectories

Similar to UMAP-based representations, diffusion maps are non-linear dimensionality reduction methods which are based on diffusion-like dynamics and can be employed to order cells along their developmental paths as they traverse from multipotent to differentiated states. In order to reconstruct in particularly the myeloid differentiation axis for individual genotypes, we applied a diffusion algorithm implemented in the R-

package ‘destiny’ [Angerer et al., 2016] on a subset of our scRNA-seq data set, containing annotated LT-HSC, ST-HSC, MPP, CMP, MoDCP and NeuP populations.

Distances of the fist diffusion component (DC1) mainly describe differentiation trajectories of cells passing from HSCs via MPPs and CMPs to MoDCPs or NeuP (Figure 3-16A). While distances of the second diffusion component (DC2) segregate MPP-associated cell fates, distances of the third diffusion component (DC3) visualize a bifurcation point where fates are split into a neutrophil primed and monocyte / DC primed direction. Plotting of DC2 versus DC3 results in a differentiation map that illustrates three different trajectories that are associated with MPP, monocyte / DC or neutrophil cell fates (Figure 3-16B, for DC1 vs. DC2 see Supplemental Figure A-8). As CMPs are located relatively in the center between ST-HSCs, MoDCPs and NeuPs, this map implies that myeloid primed cells follow a trajectory starting from LT-HSCs, and via ST-HSCs towards CMPs, but not MPPs, end in MoDCP or NeuP populations. Reconstruction of differentiation trajectories for each of the four genotypes was performed using the ‘slingshot’ algorithm [Street et al., 2018], which infers lineages based on the underlying diffusion components of each cell. For all of the four genotypes pseudotime-represented lineages could be reconstructed, that were associated with MPP-, monocyte / DC- or neutrophil-defined cell fates (Figure 3-16C, lineages indicated as lines). Notably, even for *Idh1-DNMT3A* cells, the three reconstructed lineage trajectories only marginally deviated from those observed in other genotypes. Yet, *Idh1-DNMT3A* CMP cells formed a more diffuse and cloud-like structure as compared to CMPs from other genotypes. Moreover, especially monocyte- / DC-associated trajectories appeared more constrained in *Idh1-DNMT3A* cells, suggesting that in the *Idh1*-R132H *DNMT3A*-R882H mutated myeloid compartment a partial differentiation block occurs at the CMP stage and affects monocyte / DC differentiation patterns.

In summary, our data suggests that in *Idh1-DNMT3A* mice the myeloid differentiation trajectory shows aberrant behavior. Here, cells accumulate at the CMP stage and give rise to reduced numbers of downstream NeuP and MoDCP populations. These accumulated CMP cells in part recapitulate features of NeuP and MoDCP cells (as seen in Figure 3-15) and are likely still capable of producing sufficient amounts of mature monocytes, DCs and neutrophils as evidenced by our PB and TBM analyses of theses mice. Although diffusion maps and lineage inference methods could reconstruct similar lineage trajectories for all genotypes, the observed partial differentiation block seen in *Idh1-DNMT3A* mice mirrors ineffective hematopoietic differentiation as seen for example in myelodysplastic syndromes.

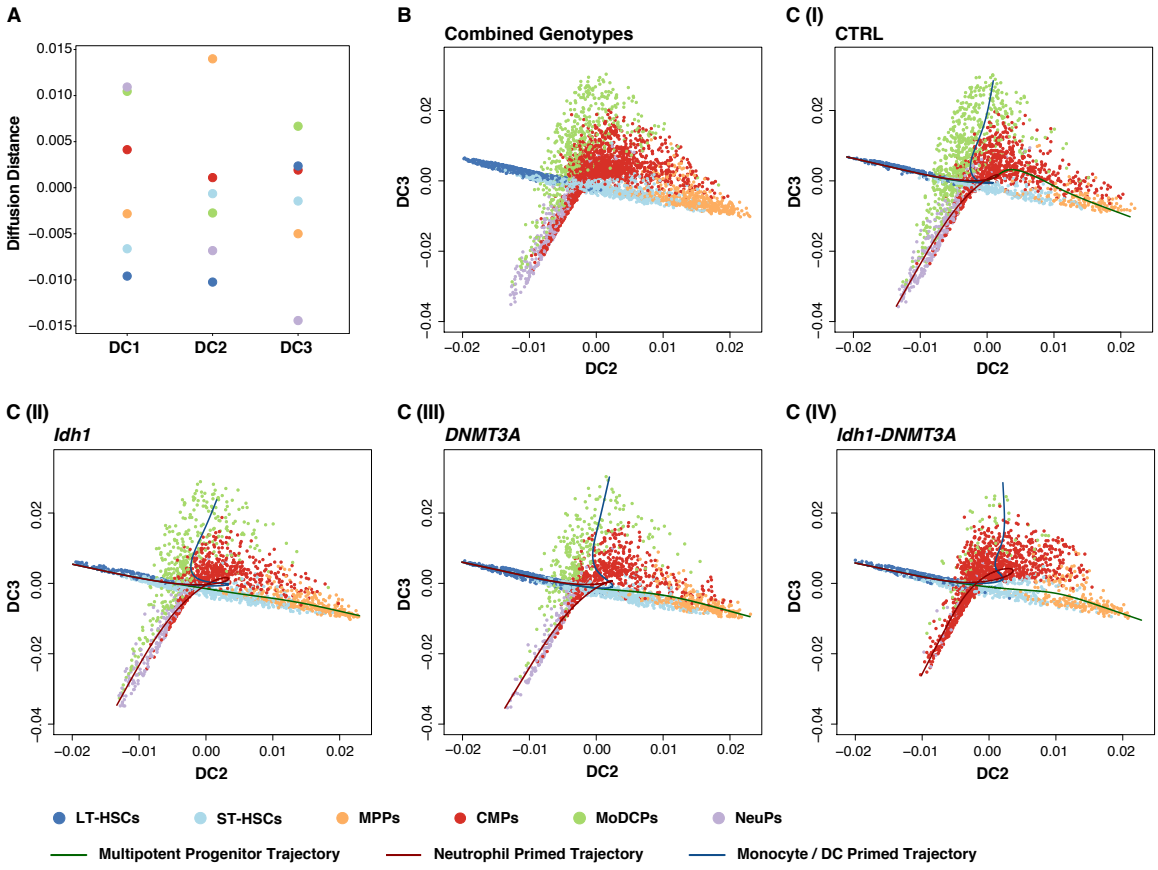


Figure 3-16: Diffusion Analysis Allows the Inference of Lineage Trajectories Within the Myeloid Compartment. Diffusion components were calculated using the R-package 'destiny' based on the normalized expression values of cells within the LT-HSC, ST-HSC, MPP, CMP, MoDCP and NeuP cluster. **(A)** Dot plot illustrating the diffusion distances for the individual diffusion components. Average distances for all cells within the indicated population were calculated irrespective of genotypes. **(B)** Diffusion map of components two and three (DC2, DC3) for combined genotypes. For representative purposes only 5,000 cells are depicted. **(C)** Diffusion map of components two and three (DC2, DC3) for each individual genotype. Lineage trajectories were inferred for each genotype separately using the R-package 'slingshot' and are depicted as lines. For representative purposes only 2,500 cells per genotype are depicted.

3.10 Molecular Signatures of Mutated Myeloid Progenitors

Having revealed differentiation defects during myeloid differentiation in *Idh1-DNMT3A* mice, we further aimed to underpin these findings at the molecular level by calling DE genes between genotypes for each annotated population. We noticed the highest differential expression between genotypes to exist in annotated CMP or MoDCP populations, where especially *Idh1*-R132H *DNMT3A*-R882H mutated cells featured a large number of DE genes (log2 fc > 0.25, padj < 0.05) (Figure 3-17A), confirming that mutation-dependent effects likely manifest in these populations or, alternatively, these cell types are most susceptible to effects mediated by combined *Idh1*-R132H and *DNMT3A*-R882H mutations. Within mature monocytes or neutrophils, the number of DE genes was rather limited.

Within the CMP population, we found several genes which are upregulated during myeloid differentiation to be significantly higher expressed in *Idh1-DNMT3A* CMPs compared to other genotypes (Figure 3-17B). These include general myeloid marker genes

such as *Fcgr3* or *Ms4a3*, granulocyte differentiation associated genes such as *Elane*, *Mpo* or *Ctsg*, but also monocyte marker genes such as *Ly6c2*. Strikingly, a large variety of transcripts encoding for ribosomal proteins (*Rps*, *Rpl*) were strongly downregulated in *Idh1-DNMT3A* CMP cells, implying translation and protein synthesis rates to be diminished in comparison to other genotypes. Inferring metabolic defects, genes which are higher expressed in *Idh1-DNMT3A* CMP cells, were often related to metabolic and biosynthesis gene ontology (GO) terms (Supplemental Figure A-9). It needs to be mentioned though, that also *DNMT3A*-R882H single-mutated cells were enriched for GO terms linked to metabolic processes (Supplemental Figure A-9).

Regarding the MoDCP population, genes which are implicated in cell cycle regulation, such as *Cdk1* or *Cdc20*, were among the highest upregulated genes in *Idh1-DNMT3A* MoDCP cells (Figure 3-17C). Consistently, GO terms associated with cell division, chromosome organization and regulation of cell cycle were enriched in *Idh1-DNMT3A* MoDCPs (Supplemental Figure A-9).

To probe whether MoDCPs indeed display higher cycling rates in *Idh1-DNMT3A* mice compared to other genotypes, we calculated S-phase and G2M-phase scores for each cell based on their gene expression profile and assigned score-based cell cycle phases in which cells are likely to reside in. As expected, HSCs were predominantly in a quiescent state assigned to a G1-phase (Figure 3-18A). Downstream progenitor cells, such as MPPs, CMPs or MoDCPs however, were predominantly assigned to the G2M- or S-phase of the cell cycle, suggesting that these populations are in a proliferative and amplifying state and responsible for driving the large-scale production of mature cell types. Comparing cell cycle assignments of myeloid progenitor populations per genotype revealed that *Idh1-DNMT3A* MoDCPs indeed featured more cells in combined G2M- or S-phases but retained less cells in the G1-phase than other genotypes (Figure 3-18B). In CMPs or NeuPs the distribution of cell cycle phases was relatively equal for all genotypes, indicating that despite alterations in metabolic processes, proliferative rates seem to be unaffected in *Idh1*-R132H *DNMT3A*-R882H double-mutated CMPs.

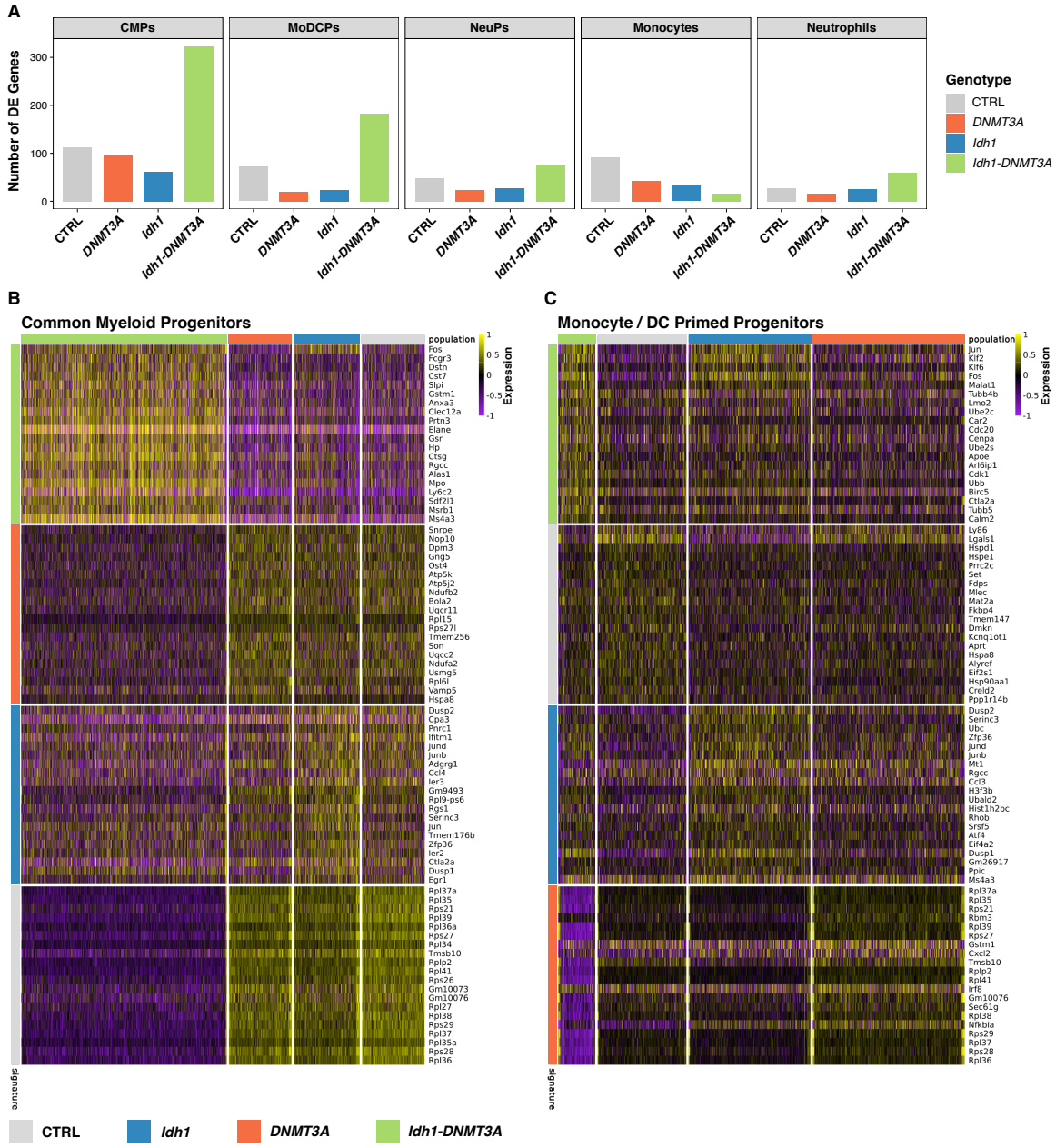


Figure 3-17: Differential Expression Analysis of Myeloid Progenitor Populations Reveals Aberrant Expression of Myeloid Differentiation-Associated Genes in *Idh1*-R132H *DNMT3A*-R882H Mutated Cells.
(A) Number of DE genes for each genotype and population. Differential expression was called between one specific genotype and all other genotypes within the same population. Genes with log2 fc > 0.25 and padj < 0.05 were considered as differentially expressed. **(B + C)** Heatmap depicting the top 20 DE genes for each genotype within CMP and MoDCP populations. Genotype-specific signatures were determined based on calling DE genes between one genotype and all other genotypes within the indicated population. Genes were ranked according to the calculated log2 fc (Differential expression analysis was performed by Abdelrahman Mahmoud).

For both *Idh1*-R132H single-mutated CMPs and MoDCPs, we noticed several inflammatory response-related genes to be among the most upregulated genes, such as *Dusp1*, *Dusp2*, *Ccl3* or *Ccl4* (Figure 3-17B and Figure 3-17C). Accordingly, gene signatures of LPS-treated monocytes or macrophages were enriched in *Idh1*-R132H mutated CMP and MoDCP cells (Supplemental Figure A-9). Collectively, this points towards an *Idh1*-R132H mutation-mediated induction of an inflammatory response in myeloid progenitors that resembles transcriptomic profiles from stimulated monocytes and macrophages.

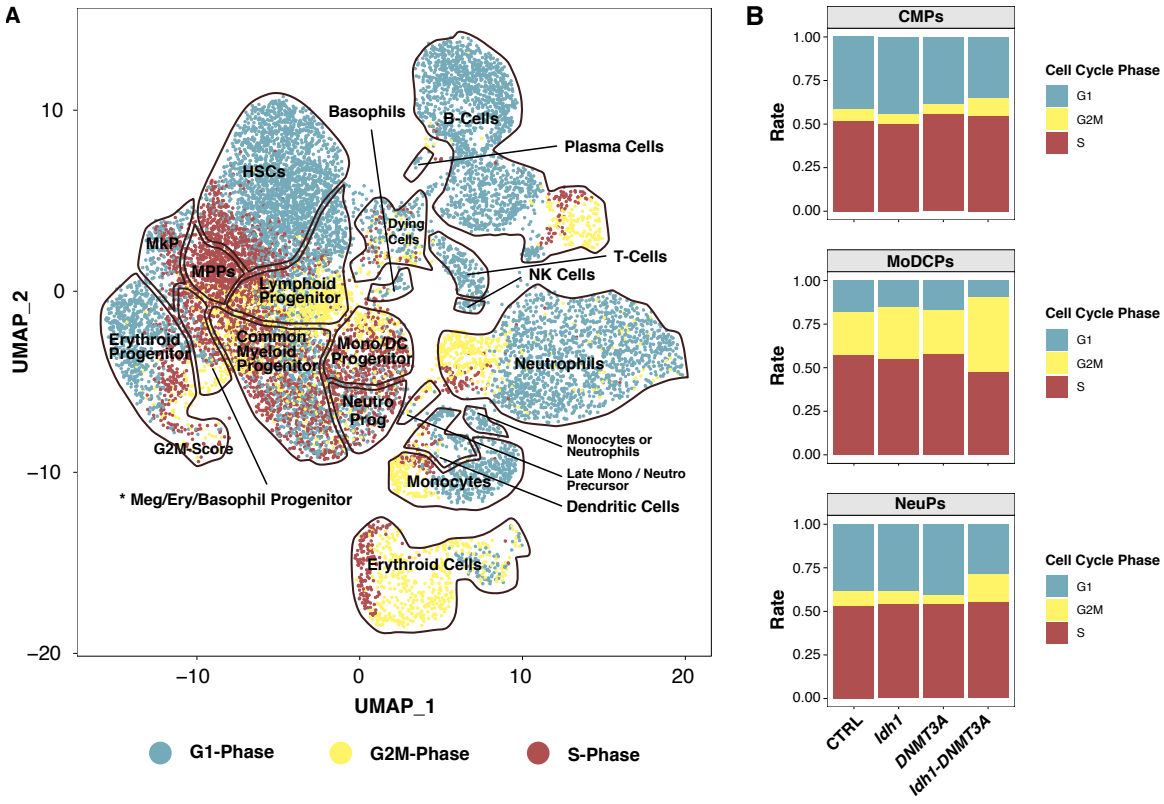


Figure 3-18: *Idh1*-R132H *DNMT3A*-R882H Mutated Monocyte / Dendritic Cell Primed Progenitors Display Deviant Cell Cycle Phase Distribution.
(A) UMAP-based representation of single-cell transcriptomes with colors illustrating the distribution of cell cycle phases across all cells. Cell cycle scores were calculated based on a predefined set of genes with characteristic expression levels during G2M- or S-phase. For each individual cell a G2M- and S-phase score was determined which was used to assign cell cycle phases. Cells with low G2M- and S-phase scores were assigned to the G1-phase. For representative purposes, only 15,000 cells are depicted. **(B)** Distribution of cell cycle phases per genotype and myeloid progenitor population.

In addition, we found the monocyte / macrophage proinflammatory stimuli-induced genes *Klf2* and *Klf6* to be consistently higher expressed in both *Idh1* and *Idh1-DNMT3A* CMPs and MoDCPs (Figure 3-19.) Similarly, expression levels of the TFs, *Fos*, *Jun*, *Junb*, *Jund* and *Myc*, which specify monocyte development, differentiation and maturation, were in part substantially increased in *Idh1*-R132H single-mutated or *Idh1*-R132H *DNMT3A*-R882H double-mutated CMPs and MoDCPs (Figure 3-19), suggesting that disturbed expression of these TFs could represent molecular determinants that drive or promote an aberrant monocytic differentiation as detected in *Idh1* and *Idh1-DNMT3A* mice.

Taken together, we could correlate the accumulation of myeloid cells at the CMP stage in *Idh1-DNMT3A* mice with a dysregulated metabolic signature and disturbed expression levels of differentiation associated genes, such as *Ly6c2* or *Elane*. Although *in silico* cycling behavior of cells at the CMP stage in *Idh1-DNMT3A* mice is comparable to other genotypes, the upregulation of myeloid marker genes suggests that cells display disturbed myeloid differentiation patterns. Yet, *Idh1-DNMT3A* MoDCP cells upregulate the expression of cell cycle-related genes to drive mature monocyte and DC production as illustrated by an increased S-/G2M-phase distribution of cells. Expression of the *Idh1*-R132H mutation alone or in combination with a *DNM3TA*-R882H mutation results in altered expression levels of several pivotal inflammatory and transcriptional regulators of the monocyte / macrophage lineage.

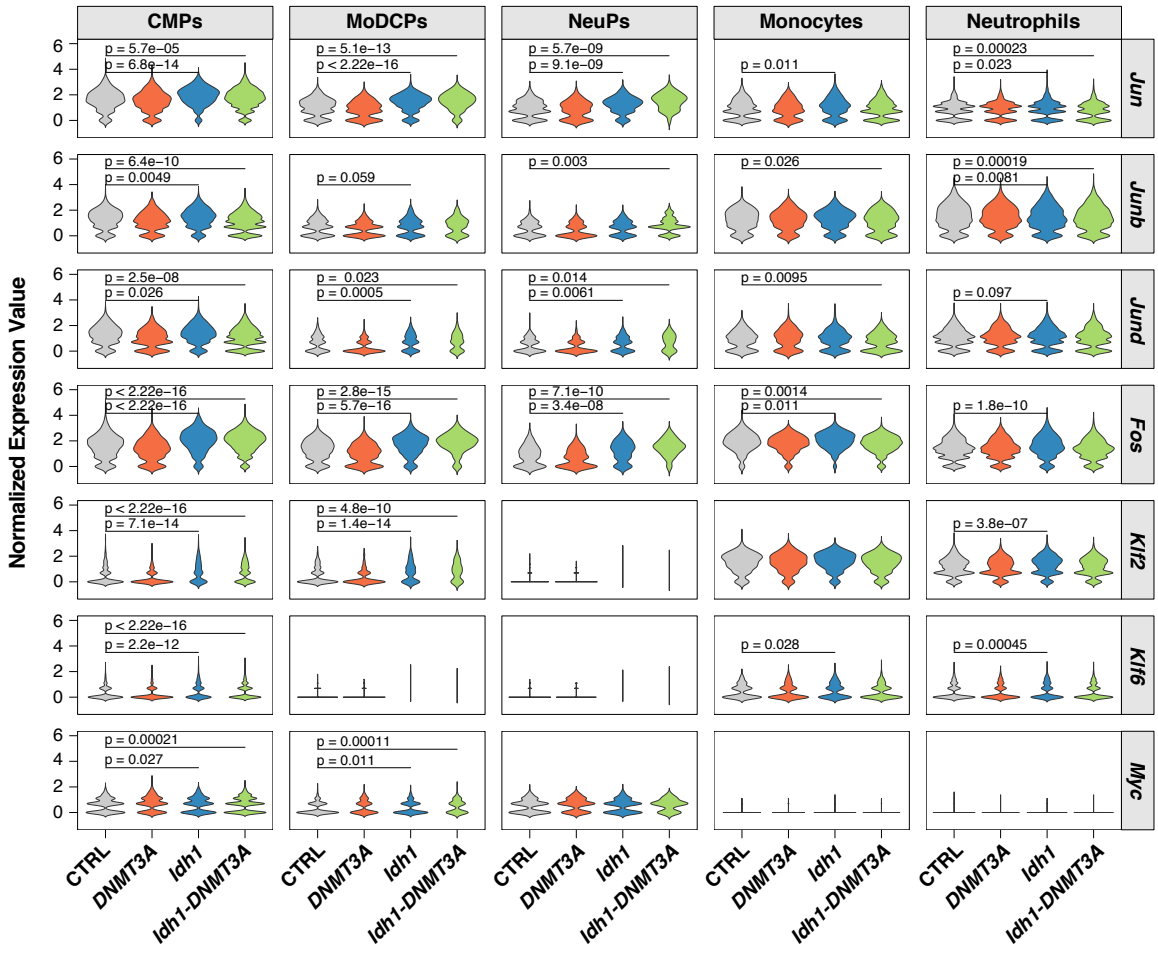


Figure 3-19: Regulators of Monocyte and Macrophage Development are Deregulated in *Idh1*-R132H Mutated Myeloid Progenitors.
Violin plot showing the distribution of normalized expression values of monocyte / macrophage regulatory TFs per genotype in myeloid progenitor populations as well as in differentiated mature monocytes and neutrophils. Violins are scaled to an equal maximum width (Student's t-test, p-values are indicated for significant results ($p < 0.05$) for the comparisons *Idh1* vs. CTRL and *Idh1-DNMT3A* vs. CTRL).

3.11 Identification of Surface Markers Specific for *Idh1*-R132H *DNMT3A-R882H* Mutated Myeloid Progenitor Cells

Hypothesizing that the observed partial differentiation block at the CMP stage in *Idh1-DNMT3A* mice mirrors a preleukemic state, we systematically assessed whether any candidate genes with aberrant overexpression in *Idh1*-R132H *DNMT3A*-R882H mutated CMPs hold potential to serve as surface markers for the early detection and diagnosis of *Idh1*-R132H *DNMT3A*-R882H mutated hematopoietic neoplasms. In order to rank genes according to the level of enrichment in *Idh1-DNMT3A* CMPs, we applied 'SurfaceGenie', a web-based application, which prioritizes genes with disparate measurements across genotypes based on the calculation of signal dispersion and signal strength coefficients [Waas et al., 2019]. By using a matrix consisting of normalized expression values of genes which are differentially expressed in *Idh1-DNMT3A* mutated CMPs, we specifically focused

on genes to be enriched in double-mutated cells. The 'OmniGenie Score' in the end served as a metric to rank genes based on the levels of disparity between *Idh1-DNMT3A* CMPs and CMPs from all other genotypes.

By far, we found the genes *Ly6c2,* encoding the Ly6C protein, and *Elane* to be expressed most divergently in *Idh1*-R132H *DNMT3A*-R882H mutated cells compared to other genotypes (Figure 3-20A) and validated the increased *Ly6c2* expression in *Idh1-DNMT3A* CMPs in our single-cell expression data set (Figure 3-20B). Ly6C is a classical surface marker which is expressed on mature monocytes, neutrophils and DC cells but also in myeloid progenitor cells as cells are specified to differentiate towards monocyte or neutrophil cell fates [Lee et al., 2013]. Within myeloid progenitor populations, Ly6C expression levels together with the expression of CD115 (Csf1r) have been employed to discriminate between cMoPs, GMPs and granulocyte progenitors (GPs) [Yanez et al., 2017, Hettinger et al., 2013, Yanez et al., 2015].

To correlate the observed increased transcriptomic *Ly6c2* mRNA expression with Ly6C surface expression levels, we extended our flow cytometric analyses of the bone marrow of Lin^- transplanted mice with Ly6C specific antibodies. Although aberrant gene expression profiles in our scRNA-seq data set were predominantly observed in transcriptome-defined CMP cells, we in particular focused on surface marker-defined GMP cells, hypothesizing that transcriptome-defined CMP cells most likely correlate with surface marker-defined GMP cells. Within the surface marker-defined GMP population (defined as Lin^-, $cKit^+$, $Sca1^-$, $CD34^+$, $CD16/32^+$) we observed elevated Ly6C mean fluorescence intensities (MFI) in both *Idh1* and *Idh1-DNMT3A* transplanted mice, indicating that increased transcriptomic levels also correlate with increased surface expression levels (Figure 3-20C). The fact that also *Idh1*-R132H single-mutants featured elevated Ly6C MFIs, however, is in contrast to our scRNA-seq data, where increased *Ly6c2* expression was exclusive to *Idh1-DNMT3A* CMPs. Quantification of cMoP frequencies (defined as Lin^-, $cKit^+$, $Sca1^-$, $CD34^+$, $CD16/32^+$, $CD115^+$, $Ly6C^+$, gating scheme in Supplemental Figure A-1), as well as of ratios of $Ly6C^+$ or $CD115^+$ GMPs to $Ly6C^-$ or $CD115^-$ GMPs revealed that both *Idh1* and *Idh1-DNMT3A* transplanted mice displayed increased cMoP counts, $Ly6C^+/Ly6C^-$ GMP ratios and $CD115^+/CD115^-$ GMP ratios (Figure 3-20D). Interestingly, increased cMoP counts and $CD115^+/CD115^-$ GMP ratios were also detected in *DNMT3A* mice, but to a much milder degree as seen in *Idh1* and *Idh1-DNMT3A* mice, while elevated $Ly6C^+/Ly6C^-$ GMP ratios were restricted to only *Idh1*-R132H single- and *Idh1*-R132H *DNMT3A*-R882H double-mutant mice.

Altogether, based on our scRNA-seq data we have identified Ly6C as an *Idh1-DNMT3A*-specific marker within transcrptome-defined CMPs. However, elevated cell surface levels of Ly6C were observed in both *Idh1* and *Idh1-DNMT3A* surface marker-defined GMPs, indicating discrepancies between our transcriptome and flow cytometric analysis. Furthermore, elevated levels of cMoPs within the myeloid progenitor cell compartment likely contribute to increased mature monocyte counts that were observed previously in *Idh1* and *Idh1-DNMT3A* Lin^- transplanted mice.

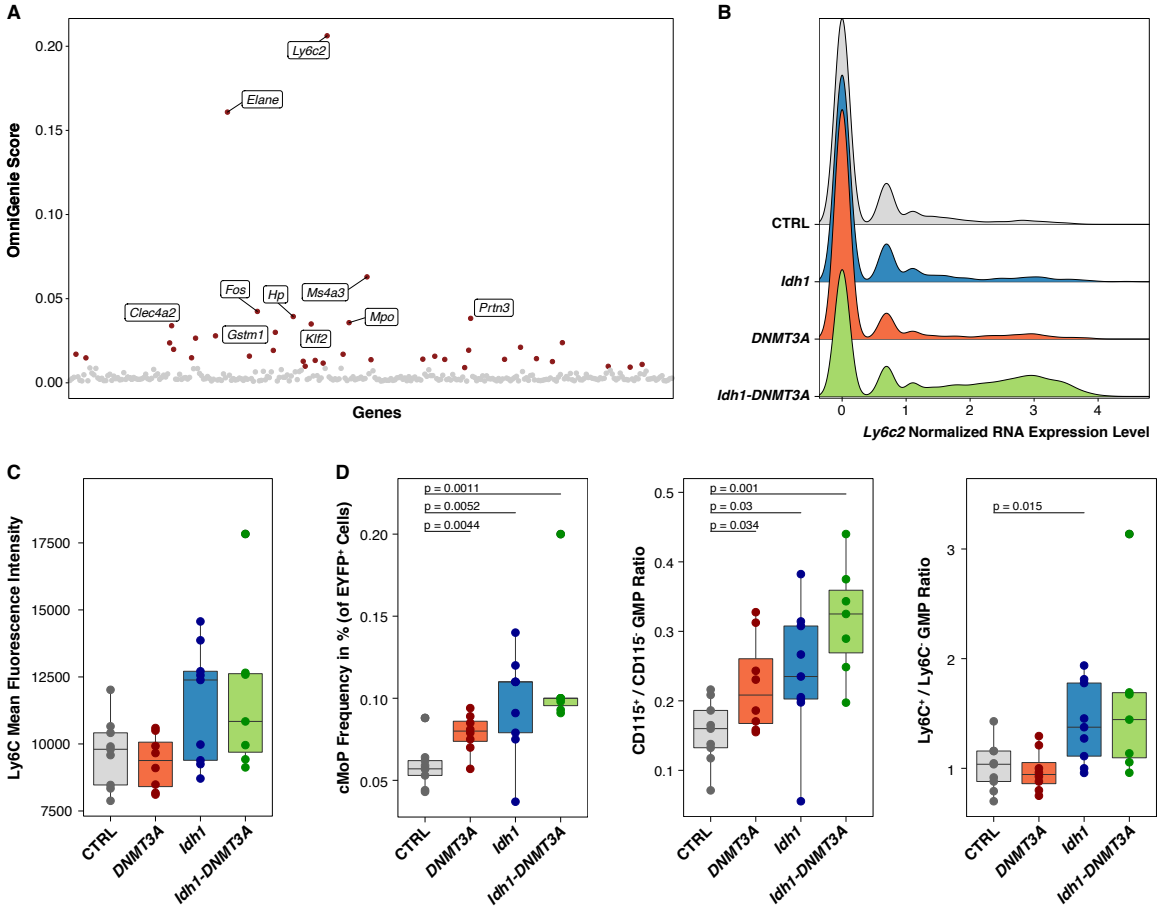


Figure 3-20: Ly6C Surface Expression is Increased in *Idh1*-R132H Mutated Myeloid Progenitor Cells.
(A) Illustration of calculated 'OmniGenie' scores as a measure for ranking disparately expressed genes between *Idh1-DNMT3A* CMPs and CMPs from all other genotypes. Scores were determined by using the 'SurfaceGenie' tool [Waas et al., 2019] with a normalized expression matrix filtered for genes differentially expressed (log2 fc > 0.25, padj < 0.05) in *Idh1-DNMT3A* mutated CMPs as input. **(B)** Quantification of normalized *Ly6c2* RNA expression levels in annotated CMP cells per genotype (data from scRNA-seq data set). **(C)** Quantification of Ly6C cell surface expression (measured by mean fluorescence intensities determined by flow cytometry) in surface marker-defined GMP populations within the bone marrow of Lin$^-$ transplanted mice (CTRL n = 9, *DNMT3A* n = 8, *Idh1* n = 9, *Idh1-DNMT3A* n = 7, two independent transplantations). **(D)** Quantification of cMoP frequencies, Ly6C$^+$/Ly6C$^-$ GMP ratios and CD115$^+$/CD115$^-$ GMP ratios in the bone marrow of Lin$^-$ transplanted mice. Frequencies of cMoPs, Ly6C$^+$ GMPs, Ly6C$^-$ GMPs, CD115$^+$ GMPs and CD115$^-$ GMPs are relative to overall EYFP$^+$ cells in the bone marrow. Gating schemes shown in Supplemental Figure A-1 (CTRL n = 9, *DNMT3A* n = 8, *Idh1* n = 9, *Idh1-DNMT3A* n = 7, two independent transplantations, Student's t-test, p-values are indicated for significant results (p < 0.05)).

To validate whether Ly6C surface expression enriches for the aberrant CMP population identified in the scRNA-seq data set in *Idh1-DNMT3A* mice, we isolated Ly6C$^+$ GMPs from Lin$^-$ transplanted mice and performed bulk RNA-seq. For each genotype, we sorted triplicates from individual transplanted mice. Clustering of samples and calculation of sample distances revealed that *Idh1* and *Idh1-DNMT3A* samples as well as CTRL and *DNMT3A* samples clustered together, with the exception being one *DNMT3A* replicate which clustered together with all three *Idh1-DMNT3A* replicates (Supplemental Figure A-10A). Accordingly, the first PC separated *Idh1* and *Idh1-DNMT3A* samples from CTRL and *DNMT3A* samples in a PCA, whereas the second PC separated *Idh1* from *Idh1-DNMT3A* samples (Supplemental Figure A-10B). It is worth noting, that with the exception of *DNMT3A* samples, replicates from all genotypes were relatively similar among each other. Calling of DE genes using log2 fc < -0.5 or > 0.5 and padj < 0.05 as cutoff resulted in 47 DE genes between *Idh1* and CTRL samples (15 upregulated, 32 downregulated), 7 DE genes between *DNMT3A* and CTRL samples (3 up, 4 down), 56 DE genes between *Idh1-*

DNMT3A and CTRL samples (23 up, 33 down), but also 66 DE genes between *Idh1-DNMT3A* and *Idh1* samples (61 up, 5 down) (Supplemental Figure A-10C).

Next, we investigated the expression of *Idh1-DNMT3A*-specific marker genes identified in scRNA-seq CMPs or MoDCPs in our bulk RNA-seq data from sorted Ly6C$^+$ GMPs. Here, the expression levels of marker genes upregulated in *Idh1-DNMT3A* scRNA-seq CMPs and MoDCPs were increased in *Idh1-DNMT3A* Ly6C$^+$ GMPs but not in *Idh1*, *DNMT3A* or CTRL Ly6C$^+$ GMPs (Figure 3-21A and Figure 3-21B). The only outlier consisted of the same *DNMT3A* replicate, which already clustered together with *Idh1-DNMT3A* samples in PCA and in sample distance analyses.

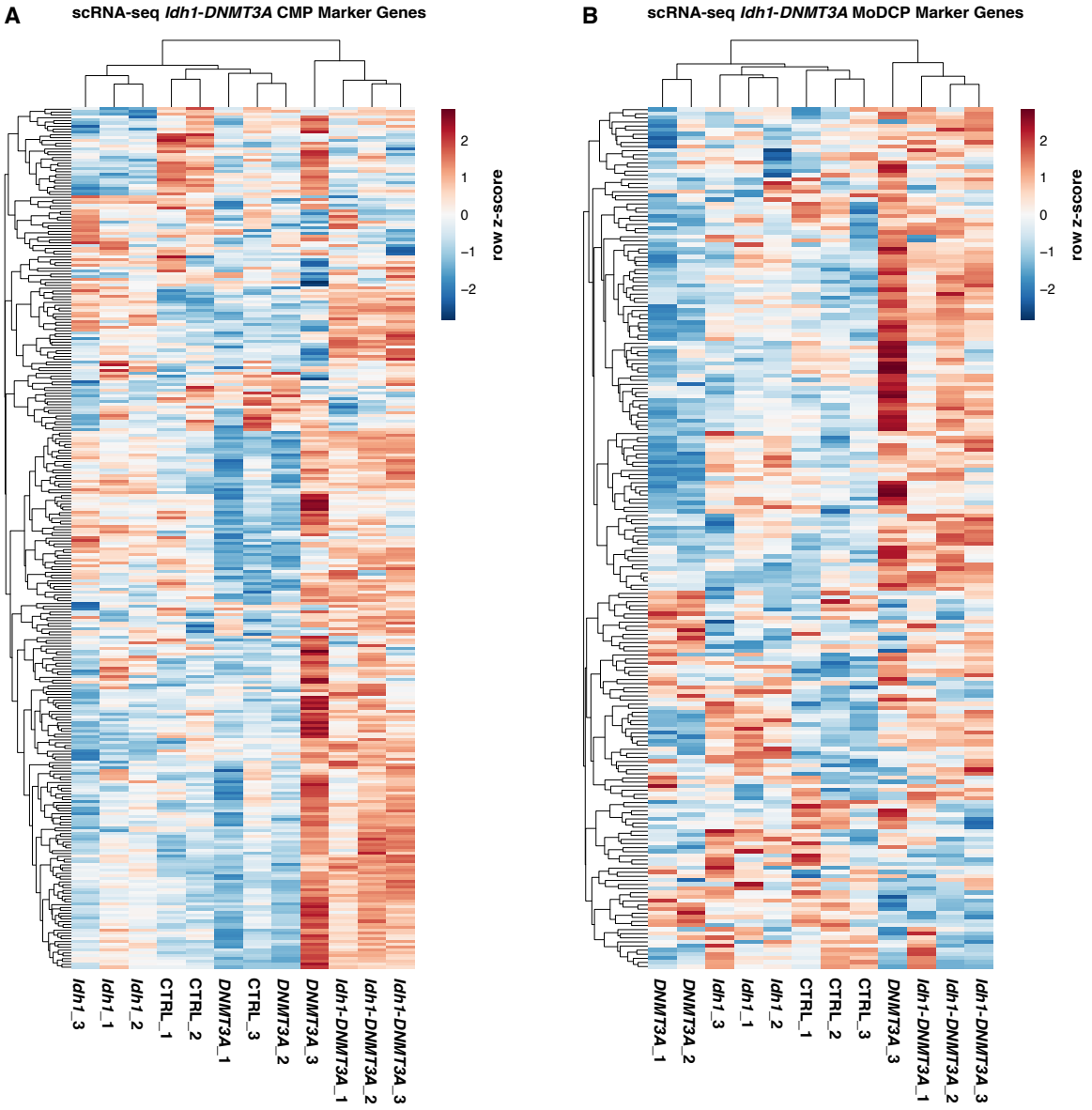


Figure 3-21: Correlation of Transcriptome- and Surface Marker-Defined Myeloid Progenitor Populations.
(A + B) Heatmap depicting normalized expression levels of *Idh1-DNMT3A*-specific marker genes, identified in scRNA-seq CMPs (A) and MoDCPs (B), in bulk RNA-seq data of isolated Ly6C$^+$ GMPs from Lin$^-$ transplanted mice. *Idh1-DNMT3A*-specific marker genes in annotated scRNA-seq CMPs were called by comparing one genotype to all other genotypes within the CMP population and DE genes with a log fc > 0.25 and padj < 0.05 were considered as differentially expressed / *Idh1*-DNMT3A specific. Each row represents a gene and clustering of rows and columns is based on Euclidean distances. Color scale represents row z-scores.

In summary, we could demonstrate that Ly6C$^+$ GMPs represent a surface marker-defined bone marrow population that carries the aberrant gene expression signature observed in *Idh1*-R132H *DNMT3A*-R882H mutated scRNA-seq CMPs. By interrogating aberrantly expressed surface markers, we correlated cell populations defined by gene expression patterns with cell surface marker-defined populations. This allowed us to prospectively isolate an aberrant cell population from *Idh1-DNMT3A* mice, which will enable us to functionally characterize these cells in future experiments.

3.12 *Idh1*-R132H Mutated Ly6C$^+$ Progenitors Show Activation of an Inflammatory Signature

In order to elucidate functional differences between mutated and CTRL Ly6C$^+$ GMPs, we performed gene set enrichment analyses, comparing expression values from *Idh1* and *Idh1-DNMT3A* replicates with CTRL replicates, but also comparing *Idh1-DNMT3A* replicates with *Idh1* replicates. Due to the low number of DE genes between *DNMT3A* and CTRL Ly6C$^+$ GMPs, we disregarded this comparison from further analysis. For both *Idh1* and *Idh1-DNMT3A* samples we found several inflammation-related gene sets, including 'Hallmark Interferon Alpha Response', 'Hallmark Interferon Gamma Response' or 'Hallmark Inflammatory Response' to be strongly enriched compared to CTRL samples (Supplemental Tables A-4 and A-5). We further aimed to identify upstream transcriptional regulators that were affected and which could explain the observed changes in gene expression. Consistently, multiple IFN signaling-related upstream regulators were activated in *Idh1* and *Idh1-DNMT3A* Ly6C$^+$ GMPs, among them IFNα, IFNβ, the IFNα/β receptor IFNAR1, the interferon regulatory factors 3 and 7 (IRF3, IRF7), but also the downstream TF STAT1 which aggregates activation signals from IFN receptors and IRFs to mediate downstream transcriptional changes (Figure 3-22A, Figure 3-22B).

Regarding differences between *Idh1*-R132H single-mutated and *Idh1*-R132H *DNMT3A*-R882H double-mutated Ly6C$^+$ progenitors, we observed a strong upregulation of Myc target genes in *Idh1-DNMT3A* samples (Figure 3-23A). Elevated Myc target gene levels were specific only for *Idh1-DNMT3A* samples with the *DNMT3A*_3 outlier-replicate again being the only exception. We next examined the expression of Myc target genes in myeloid progenitor populations of our scRNA-seq data set. Here, the highest overlap between Myc target genes and genotype-specific DE genes within the scRNA-seq CMP population was detected for *Idh1-DNMT3A*-specific DE genes, whereas the overlap between Myc target genes and DE genes from other genotypes was limited (Figure 3-23B). In line with this, Myc target genes were among gene sets enriched in *Idh1-DNMT3A* scRNA-seq CMPs (Supplemental Figure A-9).

In light of the association between Myc and cell activation, proliferation and transcriptional amplification [Wasylishen and Penn, 2010, Nie et al., 2012], elevated Myc levels could indicate an early preleukemic or even proliferative activated state. Other reports however linked Myc signaling also to myeloid and especially monocytic differentiation [Rico et al., 2017, Lee et al., 2017], implying that augmented Myc expression could propel an expansion of monocyte progenitor and mature monocyte pools.

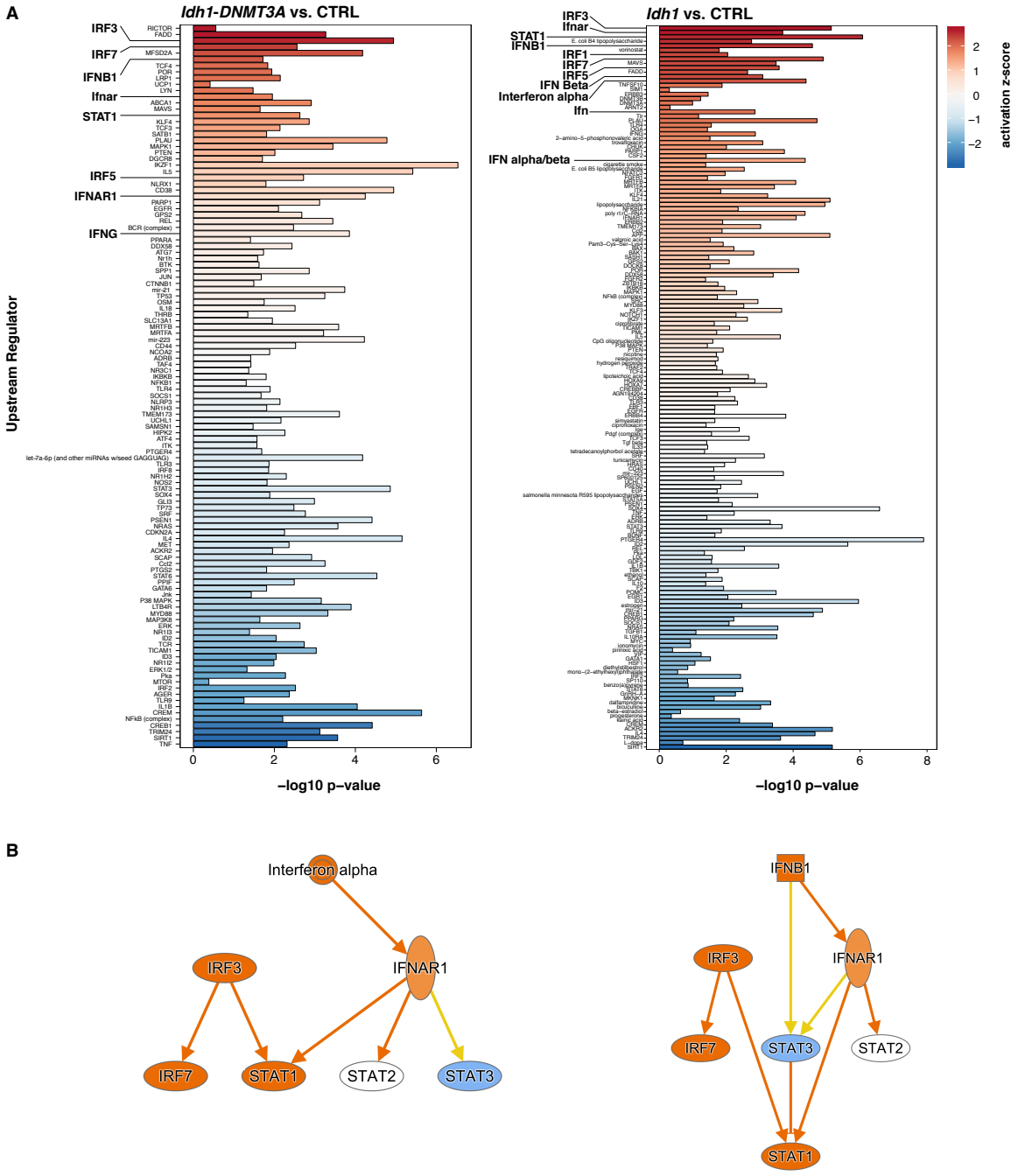


Figure 3-22: *Idh1*-R132H Mutated Ly6C⁺ Progenitors Display Elevated Expression of Interferon Signaling Components.
(A) Analysis of potential upstream regulators that could explain differential expression patterns seen between *Idh1-DNMT3A* and CTRL Ly6C⁺ GMPs as well as between *Idh1* and CTRL Ly6C⁺ GMPs. The analysis was performed using 'Ingenuity Pathway Analysis'. X-axis displays the p-value of overlap between genes from the underlying data set and genes that are regulated by the indicated regulator. Color scale represents activation z-scores which are calculated to infer activation states of indicated regulators. **(B)** Schematic illustrations of the IFNα and IFNβ signaling axis, including receptors and downstream TFs such as STAT1. Orange color indicates an activation of the regulator in the underlying dataset, blue color indicates an inhibition of the regulator in the underlying dataset. Illustrations were created with 'Ingenuity Pathway Analysis'.

To functionally test the differentiation potential of mutated myeloid progenitor cells, we sorted surface marker-defined CMPs (defined as Lin⁻ cKit⁺ Sca1⁻, CD34⁺, CD16/32⁻, Ly6C⁻, CD115⁻, CD11b⁻, Ly6G⁻), which are upstream of the aberrant surface marker-defined GMP population. Isolated CMPs were cultured in the presence of murine IL-3, IL-6 and stem cell factor (SCF). After 72 h or 96 h we investigated the expression levels of

myeloid surface markers CD11b, CD115, Ly6C and Ly6G. As expected, cells downregulated the cKit expression during culture, indicating that cells started to differentiate (Figure 3-24). We further observed slightly increased frequencies of Ly6C, CD115 and CD11b positive cells both after 72 h and 96 h of culture for *Idh1* and *Idh1-DNMT3A* cells in comparison to CTRL or *DNMT3A* cells. Rates of Ly6G$^+$ cells were also increased in *Idh1-DNMT3A* and *Idh1* cells, however the relatively low frequencies of overall Ly6G$^+$ cells indicate that Ly6G expression probably is induced only later during the myeloid differentiation process.

Taken together, both *Idh1*-R132H and *Idh1*-R132H *DNMT3A*-R882H mutated Ly6C$^+$ GMP cells are characterized by a drastically increased IFNα/β-related inflammatory signature that is driven by IRF3, IRF7 and STAT1 activity. Evidence for elevated Myc signaling was observed only in *Idh1-DNMT3A* myeloid progenitor cells and was validated both in scRNA-seq and bulk RNA-seq data from Ly6C$^+$ GMPs. Although functional effects on differentiation processes recapitulated *in vitro* appear to be only marginal, our data highlights increased expression of myeloid marker genes, which likely is mediated by the *Idh1*-R132H mutation. This observation presumably mirrors the myeloid biased differentiation patterns within the hematopoietic compartment.

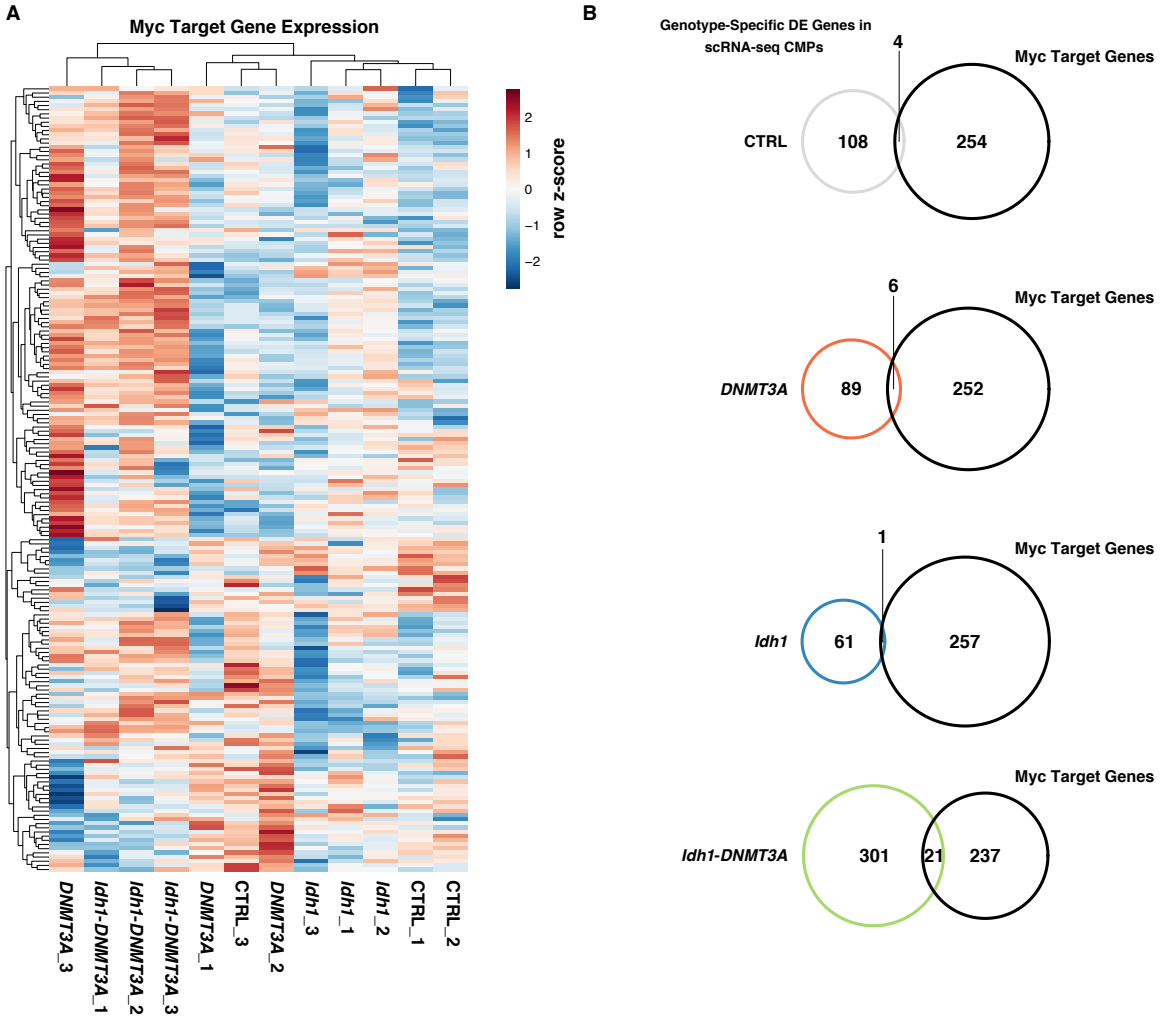


Figure 3-23: Myc Target Genes are Higher Expressed in *Idh1*-R132H *DNMT3A*-R882H Mutated Ly6C$^+$ GMPs.
(A) Heatmap depicting normalized expression levels of Myc target genes in bulk RNA-seq data of sorted Ly6C$^+$ GMPs from Lin$^-$ transplanted mice. Myc target genes were extracted from the 'Molecular Signature Database' (MSigDB). Each row represents one gene and clustering of rows and columns is based on Euclidean distances. Color scale represents

row z-scores. **(B)** Venn diagram illustrating the overlap between Myc target genes and genotype-specific DE genes in annotated scRNA-seq CMP populations. DE genes (log fc > 0.25 and padj < 0.05) were called by comparing one genotype to all other genotypes within the CMP population. Myc target genes were extracted from the 'Molecular Signature Database' (MSigDB).

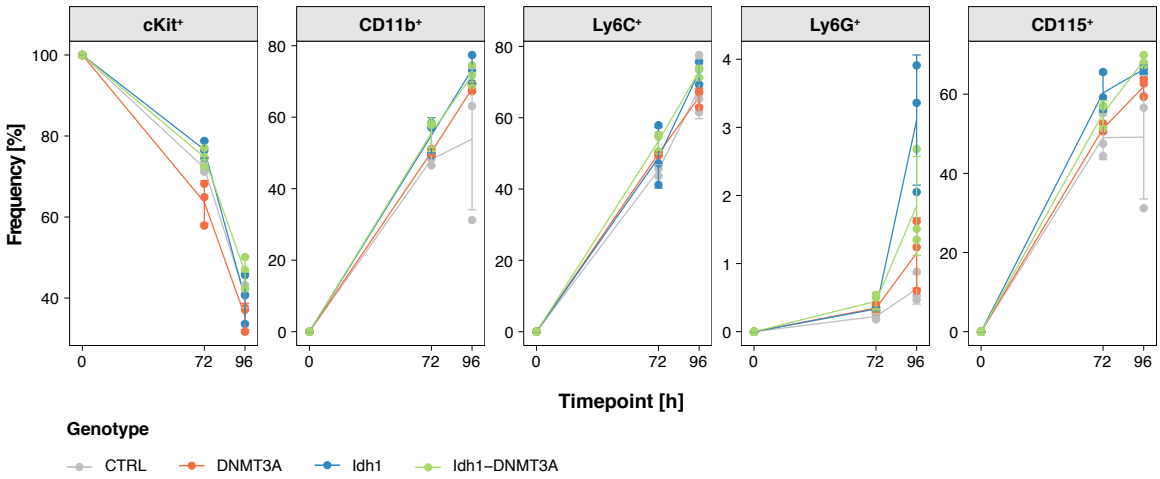


Figure 3-24: *In vitro* Differentiation Characteristics of Isolated CMPs Does Not Provide Evidence for a Differentiation Defect Towards Monocytes and Neutrophils in Mutated CMPs.
Surface marker-defined CMPs (Lin$^-$ cKit$^+$ Sca1$^-$, CD34$^+$, CD16/32$^-$, Ly6C$^-$, CD115$^-$, CD11b$^-$, Ly6G$^-$) were isolated from Lin$^-$ transplanted mice, cultured for 72 h or 96 h and then analyzed by flow cytometry for surface expression of indicated markers. Y-axis frequencies indicate the frequencies of marker positive cells at the individual time points (for all genotypes n = 3, error bars indicate mean ± SD).

4

Discussion

The process of hematopoietic differentiation and the production of mature cell types is a highly dynamic process, which is regulated on multiple levels. In this context, several studies have generated maps of transcriptome and epigenome dynamics, including DNA methylation, chromatin accessibility and histone modifications, which accompany and govern cell fate and lineage commitment decisions [Farlik et al., 2016, Buenrostro et al., 2018, Corces et al., 2016, Lara-Astiaso et al., 2014, Velten et al., 2017, Giladi et al., 2018, Lipka et al., 2014, Cabezas-Wallscheid et al., 2014]. Integration of these different molecular layers has dramatically increased our knowledge of regulatory mechanisms that are critical during hematopoiesis and serve as a reference when trying to elucidate the molecular and functional defects which are associated with hematopoietic pathologies.

Deregulation of the epigenetic machinery has been proposed for many years to be a cornerstone for the development of malignant cell growth, not only in the hematopoietic system but also in most, if not all, cancer types [Feinberg et al., 2016]. On an imaginary Waddington landscape, epigenetic regulators are thought to provide the structural basis which ensures that cells are guided to well-defined terminally differentiated cell states [Feinberg et al., 2016, Flavahan et al., 2017]. However, acquired epigenetic instability leads to the erosion of cell state-defining barriers, resulting in a more permissive state characterized by epigenetic plasticity that allows premalignant cells to adapt diverging transcriptional or developmental programs [Feinberg et al., 2016, Flavahan et al., 2017].

As a metabolic enzyme, mutations in *IDH1* indirectly disturb epigenetic patterns through the excessive production of D2HG, which inhibits several DNA and histone demethylases. However, molecular details and mechanisms of how oncogenic *Idh1*-R132H mutations drive an oncogenic transformation in the hematopoietic system are largely unknown. Within this thesis, we have comprehensively characterized the impact of oncogenic *Idh1*-R132H mutations in the hematopoietic compartment by combining functional and molecular analyses. The transcriptomic landscape of not only *Idh1*-R132H single-mutated but also *Idh1*-R132H *DNMT3A*-R882H double-mutant hematopoiesis generated here at single-cell resolution serves as a roadmap that complements previous reference transcriptomic maps of murine hematopoiesis and allowed us to identify regulatory mechanisms that are altered in the context of single *Idh1*-R132H mutations and in combination with co-occurring *DNMT3A*-R882H mutations. Critically, our analysis provided the basis to define and pinpoint aberrant cell states which are most affected by the presence of an *Idh1*-R132H mutation and will enhance our general understanding of the functional and molecular consequences induced by *Idh1*-R132H mutations.

4.1 *Idh1*-R132H Mutations in the Context of DNA Damage

Idh1-R132H mutations have been associated with an increase in DNA damage and DNA repair deficiencies, arising either through defects in the homologous recombination (HR) DNA repair pathway [Sulkowski et al., 2017], impaired sensing of DNA damage through downregulation of ATM [Molenaar et al., 2018b, Inoue et al., 2016], or through the inhibition of alkylated DNA repair enzymes of the ALKB family [Wang et al., 2015]. Here, the downregulation of ATM and the suppression of HR is mediated through the inhibition of αKG-dependent, DNA damage-induced histone demethylases KDM4A and KDM4B, supporting a model according to which epigenetic alterations represent the driving force of cellular transformation [Sulkowski et al., 2017, Inoue et al., 2016, Young et al., 2013, Mallette et al., 2012, Molenaar et al., 2018a].

We observed an increase in DNA damage rates in *Idh1*-R132H mutated as compared to wildtype 32D cells already in an untreated state, but even more pronounced after irradiation, indicating that an *Idh1*-R132H mutation *per se* is sufficient to entail an accumulation of damaged DNA. Following irradiation, we found DNA damage rates to rapidly reach steady state levels, arguing for an intact repair machinery in *Idh1*-R132H mutated 32D cells and for comparable repair kinetics to wildtype cells. It furthermore seems likely that the reduced cell cycle and proliferation rates, which we detected in *Idh1*-R132H mutated 32D cells, can be traced back to the increased amount of DNA damage which would impair survival and proliferation of cells. Yet, this direct link still needs to be established in our model system. Alternatively, the imbalance of metabolites and associated metabolic processes induced by *Idh1*-R132H mutations could simply diminish regular proliferation rates, as cells would have to cope with these unfavorable conditions. In glioma, several reports have revealed that *IDH1*-R132H mutations rather inhibit the growth and proliferation of glioma cells, an observation which could contribute to the prolonged survival of glioma patients carrying *IDH1*-R132H mutations [Bralten et al., 2011, Shi et al., 2014, Li et al., 2013, Sanson et al., 2009, Bhavya et al., 2019]. However, in these cases decreased proliferation rates were shown to be attributable to the depletion of the antioxidant glutathione (due to the decline of NADPH which is consumed during the production of D2HG), to elevated ROS levels, or to the suppressed expression of the branched-chain amino acid transaminase 1 (BCAT1), which promotes cell proliferation through initiation of the catabolism of branched-chain amino acids [Shi et al., 2014, Tonjes et al., 2013]. Yet, an increase in DNA damage rates was not shown to be causative to these phenotypes.

In transplanted LT-HSCs, several DNA damage- and DNA repair-related genes were differentially expressed in *Idh1*-R132H mutated cells, thereby validating a close interconnection of *Idh1*-R132H mutations with altered DNA damage rates and DNA repair pathways. While we did not observe any variation in *Atm* transcript levels as published previously [Inoue et al., 2016], we for instance found *Rad52*, *Rpa2, Xpc*, *Ddb2* expression levels to be increased in *Idh1*-R132H mutated LT-HSCs, with factors being pivotal for the repair of DNA by either HR (*Rad52*, *Rpa2*) or by GG-NER (*Xpc*, *Ddb2*). The fact that expression levels of these genes are elevated, could imply that increased DNA damage rates within the cell due to an expression of the *Idh1*-R132H mutation might lead to an enhanced transcription and recruitment of these repair factors to the sites of damage. However, whether *Idh1*-R132H mutated LT-HSC indeed display higher DNA damage rates in our mouse model needs to be quantified. As the *Idh1*-R132H mutation in our mice was expressed heterozygously, it seems conceivable that, because of lower expression levels,

mutation-mediated effects on DNA damage rates are not as pronounced as in cells that feature a homozygous expression or as in cells that were transduced with a mutation-expressing lentiviral overexpression plasmid.

In the hematopoietic system, the downregulation of ATM and the impairment of DNA repair were associated with reduced HSC numbers in the bone marrow and a decrease in self-renewal capacity [Inoue et al., 2016]. In our hands, we observed slight trends towards reduced HSC numbers in serially transplanted as well as in Lin^- transplanted *Idh1* mice. In line with this, in our scRNA-seq data set, the annotated LT-HSC cluster was depleted of cells from *Idh1* mice. Furthermore, we detected an inverse correlation of D2HG levels in bone marrow cell lysates and HSC frequencies in the bone marrow, although this correlation was established with a limited number of mice and needs further validation. As this indicates that the excessive production of D2HG is likely to impact the size of the stem cell pool in the bone marrow, it remains unclear whether these observations in our model system are derived from alterations in DNA damage and DNA repair rates. Alternatively, D2HG-mediated changes in the epigenetic makeup of HSCs, their cellular metabolism, their cellular signaling, or an altered cytokine environment in the bone marrow could result in reduced activities and frequencies of HSCs. Nevertheless, our results clearly support that *Idh1*-R132H mutations lead to a reduction in HSC numbers, although molecular details and causative mechanisms to this finding need to be determined.

One may speculate that the excessive accumulation of DNA damage would hinder the long-term survival and self-renewal of *Idh1*-R132H mutated HSCs, leading to a gradual and successive decline in numbers and generation of downstream progeny. In contrast, in all competitive transplantation settings, we found *Idh1*-R132H mutated transplanted HSCs to yield a better engraftment in recipients, resulting in a higher donor chimerism and arguing against a decreased self-renewal capacity. In particular in serial transplantations, which impose a great amount of selective pressure and replicative stress on transplanted HSCs, we would have expected to observe a decline in HSC functionality particularly in secondary or tertiary recipients, if defects in DNA damage repair were to play a critical role. However, even in tertiary recipients, we found *Idh1*-R132H mutated cells to showcase a better reconstitution potential, whereas in tertiary transplanted CTRL mice two out of six animals hardly displayed any donor cell engraftment. As we failed to observe any reduced engraftment or reconstitution potential in *Idh1*-R132H mutated cells, the question remains whether higher rates of DNA damage within these cells indeed impair HSC functionality. In our hands, the lack of evidence for a premature attrition of HSCs potentially could be explained by either too low expression levels of the mutation (due to the heterozygosity of the mutated allele) to induce significant alterations in DNA damage rates and DNA repair pathways, or the experimental setup applied in our serial transplantations. While we initially transplanted isolated LT-HSCs into primary recipients, we transplanted unsorted TBM into further secondary or tertiary recipients. Independent of an *Idh1*-R132H mutated context, it was shown that when transplanting isolated HSCs together with a lower number of fresh supportive bone marrow cells (i.e. 200,000 cells) in all serial transplantations, the donor cell engraftment is markedly reduced already in tertiary recipients [Challen et al., 2011, Jeong et al., 2018], suggesting that constraints in HSC self-renewal and reconstitution potential might be more evident in such a setup. Yet, even in a setup of serially transplanting TBM into further recipients, only HSCs will display long-term engraftment and contribution to blood production, so that our applied setup similarly should be appropriate to test the limits of HSC self-renewal and longevity and how these might be altered due to the presence of an *Idh1*-R132H mutation and its effects on DNA damage and repair.

On the other hand, a better engraftment of *Idh1*-R132H mutated HSCs in primary recipients could be a result from an intrinsic chronic inflammatory state. This would be in line with the inflammatory signatures we detected at the transcriptome level in these cells. In response to inflammatory signals, HSCs are known to transiently proliferate at much higher rates [Pietras, 2017, Essers et al., 2009, Baldridge et al., 2010], implying that these signatures could contribute to the rapid reconstitution of the blood system in primary recipients. In particular, it will be interesting to see how this proliferation-inducing inflammatory signature will cooperate with higher expression levels of the mutation and potentially increased rates of DNA damage in a more stringent serial transplantation setup.

Overall, while our data points at alterations in DNA damage rates within *Idh1*-R132H mutated cells, so far, we could not determine molecular details of the interplay between *Idh1*-R132H mutations and defects in DNA damage repair in our experimental model systems. Since *Idh1*-R132H mutated HSCs displayed an enhanced engraftment and reconstitution potential, a detailed functional analysis of deregulated factors involved in DNA repair as well as a correlation with methylation and histone mark signatures will be required to elucidate the exact mechanisms of how *Idh1*-R132H mutations directly or indirectly are involved in changes in DNA damage repair and how these changes impact HSC functionality.

4.2 *Idh1*-R132H Mutations Alone or in Combination with DNMT3A-R882H are not Sufficient to Induce Leukemic Transformation

In AML patients, *IDH1* mutations are considered to be early events during the pathogenesis of leukemia. They frequently occur already in HSCs and establish a pre-malignant condition that serves as a basis for disease initiation and malignant transformation [Corces-Zimmerman et al., 2014]. Recent studies identified the presence of *IDH1* mutations in healthy individuals several years before the diagnosis of AML [Abelson et al., 2018, Desai et al., 2018]. Although *IDH1* mutations here are less prevalent than for example *DNMT3A* or *TET2* mutations, they were among the mutations associated with the highest risk of disease development (measured by odds or hazard ratios) with almost all subjects carrying *IDH1* mutations eventually developing AML [Desai et al., 2018, Abelson et al., 2018]. While this showcases a strong oncogenic potential of *IDH1* mutations in humans, we failed to detect leukemic transformation in our *Idh1*-R132H mutated mouse model even after long latency times, but observed biased differentiation patterns that primarily affect the myeloid and monocytic lineage.

Critical aspects that are likely to influence disease latencies and phentoypes are expression levels of the *Idh1-R132H* mutation and of Cre recombinase within the cell and the overall mutational burden in the hematopoietic system. While a heterozygous expression of the mutated *Idh1* allele is expected to result in lower expression levels and milder phenotypes, we drove Cre expression under the control of the *Scl* enhancer. Although this resulted in the expression of the *Idh1*-R132H mutation in HSCs, the overall recombination efficiency at the *Idh1* locus was only in the range of 6-7 % (as measured by

EYFP$^+$ TBM cells). This means that the overall mutational burden in the hematopoietic system was limited in primary mice, potentially leading to prolonged disease latencies and to changes in disease phenotypes. This is supported by the fact that strongest effects on hematopoietic differentiation were observed in Lin$^-$ transplanted mice, where virtually all cells in the hematopoietic system express the *Idh1*-R132H mutation due to the lack of any supportive bone marrow. As we detected significantly increased levels of D2HG in cell lysates and blood serum of both injected and transplanted mice, thereby demonstrating that in our mouse model the mutation induction leads to an accumulation of D2HG, we nevertheless are confident that identified phenotypes reliably reflect consequences of an expression of the *Idh1*-R132H mutation within the hematopoietic system.

Previously reported mouse models for an *Idh1*-R132H mutation consisted of mice that were either crossed with LysMCre or VavCre mice. In both cases, the *Idh1*-R132H mutation is induced early on after the establishment of the adult hematopoietic system either in all hematopoietic cells (VavCre) or in developing myeloid cells (LysM) [Sasaki et al., 2012, Inoue et al., 2016]. Consistent with our findings, none of these mice developed a leukemia. However, VavCre *Idh1*-R132H mice displayed an infiltration of hematopoietic cells into the spleen, a phenotype commonly observed in mice in the context of myeloid neoplasms, and in the end succumbed prematurely with splenomegaly, anemia and thrombocytopenia [Inoue et al., 2016]. It may be that in our case the induction of the mutation in the already established adult hematopoietic system of eight- to ten-week-old mice constitutes a scenario that results in weaker phenotypes than for example a scenario where an *Idh1*-R132H mutation is expressed already in the embryo as soon as the hematopoietic system is established. Early hematopoietic cells could be more susceptible to effects mediated by the *Idh1*-R132H mutation, e.g. due to their epigenetic or transcriptomic profile, but only later in life or after certain stimuli these effects result in an abnormal hematopoietic phenotype. Alternatively, coupling of Cre to the Vav promoter could simply result in higher Cre expression levels within the cells. This might lead to higher recombination efficiencies at the *Idh1* locus, resulting in higher D2HG levels and eventually in more pronounced phenotypes. Nevertheless, we believe that our applied experimental setup, where the induction of a heterozygous *Idh1*-R132H mutation is dependent on the tamoxifen-inducible CreERT2 recombinase under the control of the *Scl* enhancer, mirrors the situation observed in humans: here, heterozygous *IDH1*-R132H mutations are typically acquired late in life within the HSC compartment and this is believed to result in the establishment and expansion of a preleukemic clone that eventually progresses to overt leukemia.

In the end, the question remains whether single *Idh1*-R132H mutations at all and upon which conditions result in the initiation or progression of a leukemic phenotype. *DNMT3A* mutations are not only the most common mutations detected in patients with CHIP or AML, but also are found to co-occur with *IDH*1-R132H mutations [Shlush et al., 2014, Abelson et al., 2018, Desai et al., 2018, Papaemmanuil et al., 2016, Cancer Genome Atlas Research et al., 2013]. Accordingly, we hypothesized that the expression of an additional *DNMT3A*-R882H mutation in *Idh1*-R132H mutated cells might result in the exacerbation of hematopoietic abnormalities. However, the observation that, similar to *Idh1*-R132H single-mutant mice, *Idh1-DNMT3A* mice displayed only minor variations in blood parameters and bone marrow composition, indicates that a co-occurrence of both mutations is not sufficient to induce malignant transformation. Yet, although our *Idh1-DNMT3A* mice in part shared similar characteristics to our *Idh1*-R132H single-mutants (i.e. increase in monocytes and monocyte progenitors, inflammatory signature in Ly6C$^+$ GMPs), our molecular data at least in part suggests a signature that is characteristic only for *Idh1-DNMT3A* mutated cells but not for single-mutants, implying that both mutations indeed cooperate to drive a unique gene expression profile. Especially in our scRNA-seq

data set we observed substantial differences between *Idh1-DNMT3A* mice and all other genotypes (discussed in detail in later sections 4.4 and 4.5).

To our knowledge, only one study investigated the interplay between *Idh1/2* and *Dnmt3a*-R878H mutations (mouse homologue to the DNMT3A-R882H mutation) in a mouse model before, however this work was focusing on the co-occurrence of an *Idh2*-R140Q mutation with a *Dnmt3a*-R878H mutation [Glass et al., 2017]. Consistent with our findings, no aberrant hematological phenotype was observed in these mice six months after the induction of the mutation [Glass et al., 2017]. At the molecular level, the authors postulated antagonizing effects of co-occurring mutations, with *Idh2*-R140Q mutations conferring a DNA hypermethylation phenotype and *Dnmt3a*-R878H mutations leading to DNA hypomethylation, but differential methylation being predominantly lost in double-mutated cells. In our model system, investigations of the epigenome will be a valuable resource to correlate transcriptomic with epigenomic signatures in order to understand how these mutations interact to drive observed aberrant gene expression profiles in specific cell types.

Published studies regarding *Dnmt3a*-R878H mutated mice indicated that the expression of a *Dnmt3a*-R878H mutation alone does not entail the development of an overt leukemia with mutated mice displaying similar survival rates as their wildtype counterparts [Guryanova et al., 2016, Loberg et al., 2019]. However, *Dnmt3a*-R878H mutated mice here featured an expansion of the HSPC compartment indicative of a situation that resembles clonal hematopoiesis, as well as a myeloid differentiation bias [Guryanova et al., 2016, Loberg et al., 2019], phenotypes that we did not observe in our *DNMT3A*-R882H mutated mice. Yet, phenotypes described in these studies were very mild and more pronounced in aged mice after long latency times since the induction of the mutation [Guryanova et al., 2016, Loberg et al., 2019], indicating that, in line with our results, the oncogenic potential of a single *DNMT3A*-R882H mutation is rather limited.

Dnmt3a knockout HSCs have been shown to exceed the lifespan and self-renewal potential of normal HSCs dramatically, being able to reconstitute the hematopoietic system in recipient mice over at least 12 rounds of serial transplantations [Jeong et al., 2018]. It will be interesting to see how our *DNMT3A*-R882H single-mutants and *Idh1*-R132H *DNMT3A*-R882H double-mutants will perform in serial transplantation settings and whether an immortalization is likewise induced by the expression of the *DNMT3A*-R882H mutation. Here, one uncertainty of our mouse model consists of the fact that we expressed the human *DNMT3A* gene carrying the R882H mutation in our mice. Although the gene is highly conserved between humans and mice, it needs to be considered that the mode of action of the mutated human version in murine cells might not be identical to the mutated murine counterpart, therefore resulting in weaker phenotypes. Similar as for *Idh1*-R132H mutations, a homozygously expressed *DNMT3A*-R882H mutation or even a homozygous knockout of *DNMT3A* has more drastic implications and consequently will result in stronger hematopoietic aberrations and changes in self-renewal capacity. However, as both *IDH1* and *DNMT3A* mutations usually are found heterozygously in patients, our approach closely mirrors situations and actual consequences of these mutations on cellular physiology as observed in patients. In contrast, a homozygous knockout or a homozygous expression of mutations might result in drastic and pronounced phenotypes, but is hardly comparable to disease scenarios found in AML patients.

In line with other reports, we conclude that, first, the oncogenic potential of *Idh1*-R132H and *DNMT3A*-R882H mutations alone or combined is not sufficient to initiate leukemic transformation and, second, whether these mutations lead to the development of leukemia seems to be context-dependent and determined by other factors that drive oncogenesis in cooperation with these mutations. Here, we hypothesize that the onset of

hematopoietic diseases is dependent either on the timing of mutational acquisition, additional activating mutations or changes in the bone marrow environment (e.g. inflammatory stress).

While we induced the expression of both *Idh1*-R132H and *DNMT3A*-R882H mutations at the same time point, one may speculate whether the temporal order of mutational acquisition might make a difference, as the initially acquired mutational hit could reshape the epigenome in a way that potentiates the effects of the second hit. However, these effects might be mitigated when mutations are acquired at the same time point as it is the case in our experimental setup.

Other reports indicated that following a preexisting *Dnmt3a*-R878H / *DNMT3A*-R882H or *Idh1*-R132H mutation, additional activating mutations in proproliferative genes such as *Flt3* and *Nras*, or a cooperation with *Npm1* mutations and Hoxa9 overexpression cause a progression to myeloproliferative disorders or AML-like phenotypes [Guryanova et al., 2016, Loberg et al., 2019, Yang et al., 2016, Meyer et al., 2016, Chen et al., 2013a, Ogawara et al., 2015, Kats et al., 2014, Chaturvedi et al., 2013]. Although *Npm1*, *Flt3*, or *Nras* mutations alone are capable of inducing myeloproliferative phenotypes in some cases, disease onset and progression was dramatically accelerated only together with *Dnmt3a*-R878H / *DNMT3A*-R882H or *Idh1*-R132H mutations, confirming that *Idh1*-R132H or *Dnmt3a*-R878H / *DNMT3A*-R882H mutations may shape an epigenetic landscape within the cell that promotes oncogenic transformation upon secondary events. We as well have begun to investigate the interplay of *Idh1*-R132H mutations and *NRAS*-G13D mutations in stem and progenitor cells using retroviral transduction. Here, preliminary data suggests that indeed both mutations cooperate to some extent and result in hematological phenotypes, although wildtype mice expressing a *NRAS*-G13D mutation alone already displayed severely reduced survival rates (data not shown). Nevertheless, further experiments will be required to corroborate these findings.

Another factor that could stimulate disease onset in the context of *Idh1*-R132H or *DNMT3A*-R882H mutations might be acute inflammatory stress. Treatments with inflammatory agents (for example LPS, polyinosinic:polycytidylic acid (pI:pC) or IFNs), provoke an activation of HSCs out of quiescence and induce cells to enter cell cycle [Walter et al., 2015, Essers et al., 2009, Baldridge et al., 2010]. Based on aberrant epigenetic patterns mediated by the presence of *Idh1*-R132H or *DNMT3A*-R882H mutations, excessive proliferation could lead to increased selective pressures followed by a clonal selection and expansion of mutated clones to initiate leukemogenesis. On the other hand, serial transplantations ideally should provoke a likewise scenario and since we have not observed strong hematological aberrations upon serial transplantations of *Idh1*-R132H mutated cells, it remains to be tested if these approaches indeed result in the development of hematopoietic neoplasms in our single-mutant or double-mutant mice.

In the end, despite the fact that the presence of an *Idh1*-R132H mutation alone or in combination with a *DNMT3A*-R882H mutation did not result in the development of an obvious hematopoietic disease phenotype, our mouse model allowed us to investigate the molecular details and signatures that are characteristic for these preleukemic cells and how these mutation-mediated alterations could incite cellular transformation upon secondary events.

4.3 Single-Cell Transcriptomic Landscapes of Mutated Hematopoietic Compartments

A centerpiece of this study consisted of a comprehensive single-cell transcriptomic analysis of the hematopoietic compartment of wildtype and mutated mice. By sampling three different layers of the hematopoietic hierarchy, we were able to obtain a transcriptomic snapshot of virtually all cell types within the hematopoietic system for each genotype. A similar approach has been applied in order to investigate native and perturbed hematopoietic transcription landscapes [Giladi et al., 2018], however, irrespective of any mutational context. In our hands, this approach proved to be a powerful tool to expose informative heterogeneity between individual cell types derived from mutated mice, rendering possible to address which cell stages are most affected by certain mutations and how transcriptional signatures are altered. Here, we relied on the droplet-based 10X Genomics platform, which allowed us to sequence a relatively high number of cells per sample (ideally ~ 10,000 cells) at the cost of covering only a limited number of genes per cell (between ~ 3,000-5,000 genes) [Wilson and Göttgens, 2018]. The low number of detected genes per cell certainly represents a confounding factor as it is not possible to distinguish whether a gene with no counts detected is truly not expressed in this cell or whether this represents a technical dropout.

By isolating only loosely defined cell populations as LSK, LS-K or $CD45^+$ TBM, we ensured to cover hematopoietic cells from all differentiation stages and lineages, including stem cells, progenitor cells and terminally differentiated cells. Contrarily, sequencing of stringently isolated individual populations such as HSCs, CMPs, GMPs, etc., inflicts potential disadvantages in that only a certain subset of cells might be defined by a particular set of surface markers used to isolate the population, whereas other subsets might be missed. For example, it was shown that classical surface marker-defined CMP and GMP populations consist of multiple transcriptionally heterogeneous subsets which show partial overlap between immunophenotypically defined CMPs and GMPs [Paul et al., 2015]. Furthermore, it turned out that the surface markers CD34 and FcgR, which are commonly used to isolate these populations, are only poor predictors of cell identity and lineage biases [Paul et al., 2015, Watcham et al., 2019]. This illustrates, that a transcriptomic definition of cell states is much more suited to dissect population heterogeneity than surface marker-based definitions. However, these discrepancies could explain divergent findings between our flow cytometry analysis and our scRNA-seq analysis.

Our scRNA-seq data set encompassed more than 40,000 cells and we could recapitulate different lineage trajectories within the hematopoietic compartment that span the erythroid, megakaryocytic, lymphoid or myeloid lineages. In line with other reports, our transcriptional analysis suggested hematopoietic differentiation to be a continuous rather than a stepwise process and the acquisition of lineage-specific fates seems to be less sequential or hierarchical [Laurenti and Gottgens, 2018, Nestorowa et al., 2016, Velten et al., 2017, Macaulay et al., 2016].

Screening a large number of cells across different compartments, populations and genotypes, allowed us to identify cell stages that are most variable between genotypes. Most importantly, further interrogation of surface markers that are expressed on these cells, enabled the prospective isolation of these cells from the bone marrow in order to characterize them in more detail. By using Ly6C expression (in addition to cKit, CD34 and CD16/32 expression) as an additional marker within the myeloid progenitor cell compartment, we could show that *Idh1*-R132H *DNMT3A*-R882H mutated cells isolated

using this marker combination exhibit an expression signature similar to aberrant *Idh1*-R132H *DNMT3A*-R882H mutated CMPs as defined in our scRNA-seq data set.

Based on our approach, it will be important to further complement and correlate transcriptional landscapes with further informative layers, such as chromatin accessibility, methylome profiles or clonal dynamics, to fully recapitulate disturbed regulatory networks of hematopoietic differentiation. For example, a method that simultaneously captures transcriptomes and somatic genotypes of cells was described recently, rendering it possible to disentangle clonal dynamics and cell ancestry information within a certain compartment and to resolve relationships between the genomic makeup and alterations in cell identities or lineage specifications [Nam et al., 2019]. In our case, a combination of functional characterizations and perturbations will be of importance to further unravel how *Idh1*-R132H mutations alone or in combination with *DNMT3A*-R882H mutations will affect functionality of certain cell types. Integration of transcriptomic states with functional readouts (e.g. index-transcriptomics together with index-cultures or transplantations) or with CRISPR-based perturbation studies have been proven to be a powerful tool to dissect regulatory circuits and dependencies at the single-cell level and will be essential to extend our understanding of the molecular mechanisms driving hematopoietic malignancies [Velten et al., 2017, Giladi et al., 2018, Jaitin et al., 2016, Dixit et al., 2016].

4.4 *Idh1*-R132H Mutations Induce Altered Myeloid Differentiation Patterns

During our analysis, we consistently observed the myeloid compartment to be most altered in either *Idh1*-R132H single-mutated or *Idh1*-R132H *DNMT3A*-R882H double-mutated mice. Prominently, our flow cytometry analysis revealed especially monocytic differentiation trajectories to be highly affected, as we observed increased cMoP frequencies, Ly6C$^+$ GMP ratios, CD115$^+$ GMP ratios, as well as increased mature monocyte frequencies in the bone marrow of *Idh1* and *Idh1-DNMT3A* mice. Our scRNA-seq analysis of the myeloid progenitor cell compartment clearly confirmed substantial discrepancies between mutant and wildtype cells. In particular, we found annotated CMP cells from double-mutated mice to be highly aberrant compared to corresponding single-mutated or CTRL CMPs. Both UMAP-based and diffusion map-based visualizations of single-cell transcriptomes suggest myeloid differentiation in *Idh1*-R132H *DNMT3A*-R882H mutated hematopoiesis to be severely compromised, resembling an accumulation of cells at the CMP stage and an ineffective hematopoietic differentiation that is characteristic for myelodysplastic syndromes. Based on these findings, it needs to be determined, why mature monocyte frequencies in the bone marrow in fact are elevated, since differentiation seems to be blocked at the CMP stage. One possible explanation could be that, although clusters were annotated as CMP cells, they still share similarities to committed monocyte progenitors, and accordingly drive the production of mature monocytes. This is supported by calculated sample distances of genotype-specific clusters, which indicates that certain *Idh1*-R132H *DNMT3A*-R882H mutated CMP cluster show similar expression profiles than MoDCP or NeuP cluster. In line with this, a differentiation trajectory associated with a monocytic cell fate could be reconstructed in double-mutated mice based on diffusion maps. Furthermore, cultured double-mutated surface marker-defined CMPs displayed comparable differentiation patterns to single-

mutant or CTRL CMPs, inferring that monocytic differentiation still is intact and not completely blocked. Nevertheless, our scRNA-seq data clearly highlights an aberrant myeloid progenitor population specific for *Idh1*-R132H *DNMT3A*-R882H mice, yet, a further and more detailed functional characterization will be required to investigate, for instance, a preleukemic nature of this population (e.g. in a transplantation setting) or whether it shows a different response to additional stimuli or conditions (e.g. in response to treatment with specific cytokines or in response to a characteristic inflammatory cytokine environment in the bone marrow).

With regard to these aspects, it furthermore needs to be determined whether this aberrant population is specific to double-mutated mice or to a certain degree even is shared with *Idh1-R132H* single-mutant mice. Although gene expression signatures specific to exclusively double-mutated myeloid progenitors were identified (e.g. increased expression of Myc target genes), increased monocyte and monocyte progenitor counts (determined by FACS), elevated expression levels of critical genes governing monocytic cell fates (e.g. *Jun*, *Fos*, *Klf2*, *Klf6*), and IFN signatures in Ly6C$^+$ GMPs were identified in both *Idh1*-R132H single-mutants and *Idh1*-R132H *DNMT3A*-R882H double-mutants. This at least in part suggests a high overlap of double-mutant phenotypes with *Idh1*-R132H single-mutant phenotypes. At this point, it needs to be noted that our scRNA-seq data was derived from a single mouse per genotype, whereas the FACS and bulk RNA-seq data is aggregated across multiple mice or even multiple transplantation cohorts, making this data more robust against experimental outliers.

Analysis of the expression of *Idh1* within the hematopoietic compartment revealed that *Idh1* is almost exclusively expressed in myeloid progenitor cell populations as well as in mature monocytes (Supplemental Figure A-11). Accordingly, the effects of an *Idh1*-R132H mutation here should be much more drastic as mutation-mediated metabolic imbalances are thought to strongly impact on cell physiology. On the other hand, cell types that express only low levels of Idh1 in the first place, will probably be only faintly affected by the expression of the mutation. Yet, it is imaginable that secreted D2HG in the bone marrow environment has an indirect impact on cell types that lack an expression of the mutated *Idh1* allele. Nevertheless, our data clearly confirms that cell stages with a high *Idh1* expression are most variable between genotypes, as we detected myeloid- and monocyte progenitor populations as well as mature monocytes to be altered in in *Idh1*-R132H mutant mice. This is in line with previously reported VavCre-*Idh1*-R132H mice, which similarly displayed an expansion of CMP and GMP populations in the bone marrow [Inoue et al., 2016].

Interestingly, *Tet2* loss in mice frequently leads to an oncogenic transformation of myeloid cells that resembles chronic myelomonocytic leukemia (CMML) and was associated with a profound increase in monocyte counts in the periphery [Moran-Crusio et al., 2011]. Since the monocytic compartment similarly seems to be highly disturbed through the expression of an *Idh1*-R132H mutation, our data suggests that *Idh1*-R132H mutations recapitulate phenotypes observed in *Tet2* knockout mice and that a D2HG-mediated inhibition of Tet2 could be a causative factor that contributes to aberrant monocytic differentiation patterns. Along these lines, further molecular (e.g. epigenomic) and functional validation experiments will be required to discriminate *Idh1*-R132H single-mutant and *Idh1*-R132H *DNMT3A*-R882H double-mutant specific characteristics or whether a shared Tet2 inhibition is the main driver of observed phenotypes.

4.5 Molecular Signatures of Mutated Myeloid Progenitors

4.5.1 Deregulation of Monocytic Cell Fate-Associated Genes

Differential expression analysis of *Idh1*-R132H *DNMT3A*-R882H double-mutant myeloid progenitor cells identified target genes of the oncogene Myc as well as a number of genes associated with various metabolic processes and mitochondrial biology to be upregulated. Besides being a critical regulator of proliferation and cell-cycle progression, Myc expression also correlates with an activation of multiple metabolic processes, such as glycolysis, nucleotide or lipid biosynthesis but also mitochondria biogenesis [Stine et al., 2015]. Elevated Myc and Myc target gene levels accordingly could explain the observed transcriptomic profile detected in *Idh1*-R132H *DNMT3A*-R882H mutated CMPs. Consistent with Myc's role in regulating cell-cycle activity, we found GO terms describing cell cycle associated processes to be enriched in *Idh1*-R132H *DNMT3A*-R882H mutated MoDCPs. Our trajectory analysis using diffusion-maps further corroborated a proliferative active, but rather diffuse and undifferentiated state of *Idh1*-R132H *DNMT3A*-R882H mutated myeloid progenitor cells that lead to an accumulation at the CMP stage.

Although it remains to be tested, we hypothesize that the accumulation of *Idh1*-R132H *DNMT3A*-R882H mutated cells at the myeloid progenitor stage, characterized by an activation of Myc signaling, could reflect initial signs of a differentiation block, which would be consistent with a preleukemic state. Yet, despite the strong oncogenic role of Myc [Dang, 2012], we think that Myc expression levels in these cells are too low to elicit tumorigenicty. Other reports have showcased a role of Myc in regulating the differentiation of myeloid cells where it is highly expressed in MDPs and cMoPs, but expression levels declining drastically in mature monocyte populations in the blood and bone marrow [Mildner et al., 2017]. Here, the down-regulation of the Myc transcriptional program by mTORC1 signaling is required for terminal myelopoieis [Selvakumaran et al., 1996, Lee et al., 2017], suggesting that a failure of cells to downregulate Myc could lead to an accumulation of cells at the progenitor stage as observed in our *Idh1-DNMT3A* mouse model.

Studies comparing the epigenomes of monocytes and granulocytes identified monocyte specific enhancers to be enriched for Myc binding sites and differentially methylated regions to overlap with Myc target gene sites [Rico et al., 2017]. Implying that the epigenomic profile in combination with Myc expression in myeloid progenitors could represent a critical step in specifying monocytic cell fate, it will be interesting to see how *Idh1*-R132H and *DNMT3A*-R882H mutations affect landscapes of DNA methylation or histone modifications to mechanistically interfere with this regulatory process, how *Idh1*-R132H mutations interact with *DNMT3A*-R882H mutations to upregulate Myc signaling, and whether the observed phenotypes are driven by the D2HG-mediated inhibition of DNA or histone demethylases (e.g. TETs, KDM4 enzymes).

Dissecting the bone marrow composition of mutated mice by flow cytometry, we found both monocyte progenitor populations as well as mature monocyte populations to be expanded in *Idh1* and *Idh1-DNMT3A* mice. On a molecular level, we could underpin this phenotype by a deregulation of several critical monocyte specific transcription factors, which display elevated expression levels in *Idh1*-R132H mutated cells, including *Jun* and *Fos*. The transcription factors c-Jun and c-Fos, together with additional subunits,

are critical parts of the TF complex AP-1, which is required for the specification of the monocytic lineage [Lord et al., 1993, Friedman, 2007b, Cai et al., 2008, Friedman, 2007a]. In combination with other myeloid determining TFs such as PU.1 and C/EBPα , AP-1 activity drives differentiation into the monocytic lineage whereas absence of AP-1 but remaining PU.1 and C/EBPα activity results in granulocyte development [Cai et al., 2008, Friedman, 2007a]. Besides the activity of TFs, specific cytokine signaling is equally important to regulate developmental programs. Here, M-CSF and granulocyte colony stimulating factor (G-CSF) are pivotal cytokines that either induce monocytic or granulocytic differentiation, respectively. Importantly, M-CSF induced signaling culminates in the induction of *Fos* expression to direct monopoiesis [Jack et al., 2009], thereby further underlining the importance of a regulation of *Jun* and *Fos* levels for a balanced monocytic and granulocytic cell output. A recent report identified the histone demethylase Kdm6b as a regulator of *Fos* and *Jun* expression, where loss of Kdm6b was followed by an increased expression of AP-1 [Mallaney et al., 2019]. Although the authors here focused on the effect of Kdm6b loss on HSCs functionality, a D2HG-mediated inhibition of Kdm6b in myeloid progenitors could provide an intriguing mechanistic link that explains how *Idh1*-R132H mutations result in elevated *Jun* and *Fos* levels to promote monocytic differentiation. To substantiate this hypothesis, further mechanistic studies will be required that specifically investigate how *Idh1*-R132H mutations influence *Jun* or *Fos* levels in myeloid progenitors.

4.5.2 Deregulation of Inflammation-Associated Genes

Within the hematopoietic system, monocytes are critical effector cells of the innate immune system that are recruited to sites of inflammation in a proinflammatory mediator-dependent manner [Jakubzick et al., 2017, Guilliams et al., 2018]. Upon inflammation, monocytes synthesize and secret proinflammatory cytokines and differentiate into proinflammatory M1-macrophages to clear pathogens and to further promote inflammatory conditions [Jakubzick et al., 2017, Guilliams et al., 2018]. Bacterial LPS and IFNγ here are critical regulators that stimulate the differentiation into this particular macrophage subtype [Orecchioni et al., 2019]. We found *Idh1*-R132H single-mutant and especially *Idh1*-R132H *DNMT3A*-R882H double-mutant mice to display increased Ly6C expression within the myeloid progenitor compartment as well as enhanced frequencies of specifically $Ly6C^{High}$ inflammatory monocytes in the bone marrow. In line with this, scRNA-seq data of *Idh1*-R132H mutated CMPs, MoDCPs and monocytes revealed *Idh1*-R132H single-mutant-specific differentially expressed genes to be enriched in gene signatures of LPS-stimulated monocytes or macrophages. Furthermore, both *Idh1*-R132H single-mutated and *Idh1*-R132H *DNMT3A*-R882H double-mutated $Ly6C^{+}$ GMPs featured elevated IFN signaling levels. Overall, this data decisively points towards a link between the expression of the *Idh1*-R132H mutation and an inflammatory signature.

Severe and acute inflammatory states have been shown to induce an emergency production of increased levels of monocytes and a recruitment to inflammatory sites [Guilliams et al., 2018]. This emergency myelopoieisis often is accompanied by a rewiring of the epigenome and acquisition of epigenetic marks which entail a state of ‘trained immunity’ to ensure a long-lasting increased production of, for example, monocytes or macrophages [Netea et al., 2016]. With respect to *Idh1*-R132H mutations, it will be important to reveal whether increased monocyte and monocyte progenitor counts are derived from intrinsic effects of the *Idh1*-R132H mutation (for example through the dysregulation of lineage-specific transcription factors such as *Jun* or *Fos*, or a rewiring of

the epigenome) or rather are only an indirect consequence of inflammatory conditions in the bone marrow. Several genes, including *Klf2*, *Klf6*, *Dusp1*, *Dusp2*, *Ccl3*, *Ccl4*, and *Junb,* which we found to be dysregulated in *Idh1*-R132H mutated myeloid progenitors, are associated with an activation of monocytes or macrophages upon inflammatory stimuli. The TF Klf2 has been shown to act as a negative regulator of a proinflammatory activation of monocytes, whereas Klf6 especially regulates the inflammatory macrophage polarization upon proinflammatory stimuli such as LPS or IFN [Das et al., 2006, Jha and Das, 2017, Nayak et al., 2013, Thomas et al., 2016, Date et al., 2014]. Similarly, the expression of *Dusp1*, *Dusp2*, *Ccl3* and *Ccl4* is induced upon the activation of macrophages to regulate pathological inflammatory responses, whereas Junb was identified as a key transcriptional modulator to govern macrophage activation [Seternes et al., 2019, Maurer and von Stebut, 2004, Fontana et al., 2015]. Collectively, although we did not sample bone marrow macrophages in our scRNA-seq experiment, our data suggests that *Idh1-R132H* mutations, either directly or indirectly through inducing an inflammatory environment, establish a gene expression profile which is associated with an excessive activation of monocytes or macrophages and which contributes to or further promotes an inflammatory response.

In macrophages, loss of Tet2 resulted in the upregulation of several inflammatory mediators such as IL-6, IL-1β, CXCL1, CXCL2 and CXCL3 as Tet2 was identified to directly be involved in the repression of these factors in order to resolve inflammation [Zhang et al., 2015, Jaiswal et al., 2017]. More specifically, the increased IL-6 production and a concomitant proinflammatory state were pivotal factors for the development of a preleukemic myeloproliferation in these mice [Meisel et al., 2018]. As *Idh1*-R132H mutated macrophages in theory are likely to phenocopy *Tet2* knockout macrophages due to the D2HG-mediated inhibition of Tet2, the expression of the *Idh1*-R132H mutation in macrophages similarly could affect activation states of macrophages, lead to an increased secretion of proinflammatory cytokines and establish an inflammatory milieu in the bone marrow. An inflammatory environment in turn is likely to affect transcriptome profiles of e.g. HSCs, myeloid progenitors or monocytes. In line with this, we detected both an inflammatory signature in *Idh1*-R132H mutated HSCs as well as a strong IFN signature in *Idh1*-R132H mutated Ly6C$^+$ GMPs. Importantly, macrophages have been shown to secrete high levels of IFNs (mainly IFNγ, but also IFNβ) upon activation [Kumaran Satyanarayanan et al., 2019] and thereby could explain observed elevated IFN signaling levels. Eventually, further functional characterizations of *Idh1*-R132H mutated macrophages will be needed to quantify their cytokine secretion and to disclose whether *Idh1*-R132H mutations mitigate the Tet2-mediated repression of proinflammatory mediators through inhibition of its enzymatic activity.

Besides changes in monocyte or macrophage activation states, inflammatory signatures as well could be derived from an activation of pathogen sensing mechanisms. Pathogen-associated molecular patterns are sensed by innate immune cells via pattern recognition receptors (PRRs) (e.g. TLRs) which, upon stimulation, induce the activation of several downstream inflammatory pathways to mediate an effective immune response [Kawai and Akira, 2010]. Whereas most of the TLRs are located on the outer cell membrane or in endosomes, a few cytosolic PRRs are activated through sensing of intracellular viral RNAs [Yoneyama and Fujita, 2009, Kawai and Akira, 2010]. IRF3 and IRF7 are key transcriptional regulators which are activated upon PRR-signaling and induce the transcription of IFNα and IFNβ in order to orchestrate an antiviral innate immunity response [Mogensen, 2018]. We identified IRF3, IRF7 as well as IFNα and IFNβ as critical upstream regulators with higher activation levels in Ly6C$^+$ GMPs from *Idh1*-R132H single-mutant and *Idh1*-R132H *DNMT3A*-R882H double-mutant mice, implying that corresponding mice either had to cope with an actual pathogen infection or experienced

an intrinsic activation of endogenous retroviruses (ERVs) which were sensed by cytosolic PRRs and resulted in the initiation of IFN signaling. Although we cannot completely rule out a bacterial or viral infection, risks were minimized by holding animals under pathogen-free conditions. Previous studies linked the upregulation of endogenous retroviruses to changes in DNA methylation, proposing a dysregulation of the epigenome to be a critical factor that could drive ERV transcription [Chiappinelli et al., 2015, Brocks et al., 2017, Daskalakis et al., 2018]. Despite the fact that in these cases ERV transcription was initiated through treatments with hypomethylating agents such as DNMT inhibitors, one could speculate that complex effects of *Idh1*-R132H mutations on DNA methylation and histone marks could cooperate to favor the expression of ERVs, eventually resulting in elevated IRF activity and IFN signaling levels. Yet, further experiments will be required to proof this hypothesis.

Chronic inflammation is a shared characteristic of many hematopoietic malignancies and plays an important role in disease progression with many patients exhibiting an excessive production of inflammatory cytokines [Craver et al., 2018]. Having detected alterations in IFN signaling at the transcriptional level, we proceeded to quantify secreted cytokine levels within the bone marrow compartment of *Idh1*-R132H mutated mice. Although we found several cytokines, among them IL-6, to be differentially expressed in a pilot experiment, these results need to be reproduced in future validation studies. In the context of *Idh1*-R132H mutations, it will be important to address which cell types are responsible for the aberrant expression of inflammatory mediators, which molecular mechanisms drive this overproduction and which role inflammatory signatures play with regard to defects in hematopoietic differentiation and malignant transformation.

4.6 Conclusions and Perspectives

In the course of this project, we have extensively studied molecular and phenotypic characteristics of an *Idh1*-R132H mutation in the hematopoietic system. Analysis of *Idh1*-R132H single-mutant mice revealed that the overall functionality of HSCs is not compromised through the heterozygous expression of an *Idh1*-R132H mutation in single or serial transplantation settings, although our molecular data at least in part indicates disturbed expression levels of DNA damage- and DNA repair-associated genes. By integrating latest advances in high-throughput single-cell technologies and combinatorial *Idh1*-R132H and *DNMT3A*-R882H mouse models, we in detail delineated the cooperativity of both mutations and the repercussion that these mutations impose on different hematopoietic cell types. Single-cell transcriptomic data of virtually all hematopoietic cell types within the bone marrow compartment here will serve as a resource when investigating *Idh1*-R132H or *DNMT3A*-R882H mutations in the context of altered physiological conditions and additional extrinsic stimuli, such as induced chronic inflammatory states, aging or the cooperation with other mutations.

Neither single-mutant, nor double-mutant mice developed spontaneous hematological malignancies, illustrating that the oncogenic potential of these mutations is limited and possibly context-dependent in response to additional intrinsic or extrinsic factors. Rather, these mutations result in subtle abnormalities of hematopoietic differentiation. *Idh1*-R132H mutations alone or in combination with a *DNMT3A*-R882H mutation preferentially lead to a myeloid differentiation bias with slight predominance of the monocytic cell fate. By and large, this phenotype could be validated both at the

phenotypic as well as at the transcriptomic level. The expression of the *Idh1*-R132H mutation resulted in a deregulation of several critical regulators of myeloid and monocytic cell fate as well as inflammatory regulators that mediate the activation of monocytes or macrophages upon inflammatory stimuli. Pseudotime-inferred differentiation trajectories revealed an aberrant lineage specification in particular for *Idh1*-R132H *DNMT3A*-R882H double-mutated myeloid progenitor cells, which exclusively were characterized by elevated Myc signaling. Furthermore, this transcriptome-defined aberrant cell population could be correlated with a surface marker-defined cell population, enabling the use of certain surface marker combinations to prospectively isolate and characterize this population. Eventually, elevated IFN signaling was found as a common characteristic in *Idh1*-R132H and *Idh1*-R132H *DNMT3A*-R882H mutated monocyte progenitor populations and could contribute to altered differentiation patterns in response to a proinflammatory state.

Based on these results, several routes need to be considered for the future progression of this project. Most importantly, underpinning the rather descriptive characterization of altered cell states with functional studies will help to enhance our understanding of molecular mechanisms underlying these phenotypes. For instance, the characterization of DNA methylome and chromatin profiles of HSCs, myeloid progenitor cell populations, monocytes or macrophages from mutated mice will assist in unraveling whether the *Idh1*-R132H mutation mediates its effects via inhibition of Tet enzymes, via inhibition of histone demethylases, or via alternative mechanisms. Furthermore, the epigenomic characterization of these cell types will disclose the potential molecular interaction between *Idh1*-R132H and *DNMT3A*-R882H mutations.

Our group recently has invested significant efforts into the generation of a genome-wide DNA methylation map of the entire murine hematopoietic system in order to investigate the DNA methylation programs which regulate hematopoietic differentiation and lineage commitment. This data will serve as a reference data set to identify how aberrant DNA methylation signatures contribute to the deregulation of hematopoietic differentiation observed in *Idh1*-R132H and / or *DNMT3A*-R882H mutated cells. The integration of epigenome and transcriptome data sets will furthermore be of use to identify disturbed regulatory networks, which we will focus on for further functional characterization.

On a functional note, a critical step will be to address if and how these findings translate to an altered cell function in affected cells. Regarding DNA damage and DNA repair, combining epigenomic and transcriptomic profiles with specific functional assays that quantify DNA damage rates, DNA repair kinetics and the mutational load of *Idh1*-R132H mutated HSCs should decipher whether mutation-mediated defects in DNA repair pathways play a critical role in these cells, which repair pathways and enzymes specifically are affected, and how this impacts HSC attrition. Assays that quantify monocyte and macrophage differentiation, migration, stimulation or cytokine production should be employed to elucidate how and to which extent *Idh1*-R132H mutations affect monocyte and macrophage recruitment, activation and the establishment of an inflammatory environment within the bone marrow niche. Having identified altered regulatory networks, their impact on monocyte or macrophage function should be tested in perturbation studies using specific inhibitors, CRISPR/Cas9 or siRNAs which allow us to specifically modulate these networks. Based on this, it is our hope to identify druggable targets that can be exploited to establish novel therapeutic strategies specifically tailored for patients with hematopoietic neoplasms driven by *IDH1*-R132H or *DNMT3A*-R882H mutations.

Ultimately, it will be of utmost importance to understand which circumstances promote the initiation of preleukemic clonal expansion and leukemic transformation in

the context of *Idh1*-R132H and *DNMT3A*-R882H mutations. Since we started to investigate cooperating effects between *Idh1*-R132H and *NRAS*-G13D mutations, we plan to further elaborate molecular characteristics that define the interaction of preleukemic *Idh1*-R132H and *DNMT3A*-R882H mutations with secondary proliferative mutations in genes such as *Flt3* or *Npm1* mutations. Based on the proinflammatory conditions specific to the bone marrow compartment of *Idh1*-R132H mutant mice, it will be interesting to see how an application of chronic inflammatory stress (for example induced through serial injections with LPS, pI:pC or IFN) will affect functionality of mutated HSCs, myeloid progenitors, monocytes or macrophages and whether these conditions promote a leukemic progression. Understanding the molecular determinants that drive this progression will be essential to develop treatment options that can be applied to healthy individuals with existing preleukemic *IDH1*-R132H or *DNMT3A*-R882H mutations in order to prevent leukemic transformation.

5

Materials & Methods

5.1 Culturing of Cell Lines and Primary Mouse Cells

32D cells were cultured in RPMI 1640 medium (Gibco) supplemented with 10 % fetal bovine serum (FBS, Biochrom), 1 % penicillin / streptomycin (Pen/Strep, Gibco) and 10 ng/µl murine IL-3 (Peprotech).

HEK293T cells and NIH/3T3 cells were cultured in DMEM medium supplemented with 10 % FBS, 1 % Pen/Strep and 1 % L-glutamine (Gibco).

For *in vitro* differentiation experiments, isolated myeloid progenitor populations (i.e. surface marker-defined CMPs) were cultured in StemSpan™ SFEM medium (Stemcell Technologies) supplemented with murine IL-3 (10 ng/µl), murine IL-6 (10 ng/µl) and murine SCF (100 ng/µl) (Peprotech).

Cultured cells were incubated at 37˚C and 5 % CO_2

5.2 Lentiviral Overexpression of *Idh1* cDNA Constructs in 32D Cells

5.2.1 Lentiviral Overexpression Plasmids

The following plasmids were used to investigate the effect of *Idh1*-R132H mutations in 32D cells:

Table 5-1: List of Lentiviral Overexpression Plasmids.

Name	Backbone	Insert	Fluorescence
LeGO-iC2 Empty Vector	LeGO-iC2	-	mCherry
LeGO-iC2 *Idh1*-WT	LeGO-iC2	Wildtype *Idh1* cDNA	mCherry
LeGO-iC2 *Idh1*-R132H	LeGO-iC2	*Idh1*-R132H mutated cDNA	mCherry
LeGO-iV2 Empty Vector	LeGO-iV2	-	GFP

5.2.2 Lentivirus Production

Lentiviral particles were produced in HEK293T cells. For each virus production, a fresh vial of cells was thawed and cells were cultured for at least one week before the start of virus production. For each 150 cm^2 (T150) flask of cells to be transfected, a transfection mix was prepared, consisting of 50 µg lentiviral plasmid of interest (either LeGO-iC2 Empty Vector, LeGO-iV2 Empty Vector, LeGO-iC2 *Idh1*-WT or LeGO-iC2 *Idh1*-R132H), 37.5 µg packaging plasmid psPAX2 (PlasmidFactory), 5 µg envelope plasmid pMD2.G (PlasmidFactory) and 90 µl 2 M $CaCl_2$ (Calcium Phosphate Transfection Kit, Invitrogen), and was filled up with sterile ddH_2O to a volume of 750 µl. The transfection mix was mixed with 750 µl 2x Hepes Buffered Saline buffer (HBS, Calcium Phosphate Transfection Kit, Invitrogen), vortexed thoroughly and incubated for 45 min at room temperature (RT). During the incubation, T150 flasks were coated with gelatin by incubating the flasks filled with 3 ml of EmbryoMax® 0,1% gelatin solution (Merck Millipore) at 37°C for 10 min. Afterwards, the gelatin solution was discarded and 2,5 x 10^7 HEK293T cells were seeded per gelatin-coated flask in 12 ml HEK293T medium supplemented additionally with 25 µM chloroquine (Calcium Phosphate Transfection Kit, Invitrogen). After the incubation of the transfection mix, 1.5 ml chloroquine containing HEK293T medium was mixed with the transfection mix and added to the cell suspension in gelatin-coated flasks. Cells were incubated overnight at 37°C and on the next morning, the medium was replaced by fresh HEK293T medium, additionally supplemented with 1 % HEPES (Merck Sigma-Aldrich). After 24 h, the medium containing the virus particles was harvested and filtered using a 0.45 µm PVDF filter unit (Merck Millipore). Fresh HEPES supplemented HEK293T medium was added to the cells and again after 24 h harvested and filtered. Afterwards, the filtered medium from the first and second harvest was pooled and centrifuged (Beckman Coulter Optima L-90K Ultracentrifuge) at 22,000 rpm for 2 h. Following centrifugation, the virus pellet was resuspended in an adequate volume of StemSpan SFEM medium (Stemcell Technologies) and stored in 10 µl aliquots at -80°C.

5.2.3 Lentivirus Titration

The titration of the produced lentiviruses was performed using the murine embryonic fibroblast cell line NIH/3T3. Cells were cultured at least one week before beginning the titration. One day before the start of the titration, 5 x 10^4 cells were seeded per well on a 6-well plate and incubated overnight. On the following day, five dilutions of the virus (ranging from 1:400 to 1:400,000) were prepared in NIH/3T3 medium containing 8 µg/ml polybrene (Merck Sigma-Aldrich) and added to the cells. The number of cells per well at the time point of infection was determined by counting cells in a comparable well, equally seeded the day before. Subsequently to adding the virus, cells were incubated for 24 h at 37°C. After this period, the medium was replaced by fresh NIH/3T3 cell medium and cells were incubated for additional 24 h. Transduced cells were harvested 48 h after infection with virus particles, and the percentage of mCherry or GFP positive cells was determined by FACS analysis. The virus titer (measured in infection units (IU) per ml) was calculated according to the following formula:

$$\mathrm{IU/ml} = \frac{(\text{Dilution factor x frequency of positive cells x number of cells})}{\text{Volume of transduction}}$$

5.2.4 Lentiviral Transduction

For the transduction of 32D cells, 1 x 10^5 cells were seeded per well on a 6-well plate in 2 ml 32D cell medium additionally supplemented with 8 µg/ml polybrene (Merck Sigma-Aldrich). The amount of virus to be added to the cells was calculated to yield a multiplicity of infection (MOI) of 2. The respective amount of virus then was added to each well and cells were incubated for 24 h at 37°C. Afterwards, the medium was replaced by fresh 32D cell medium and cells were incubated for another 24 h. Subsequently, cells were washed and mCherry or GFP positive cells were sorted on Aria I, Aria II, or Fusion cell sorter systems (BD Biosciences) using a 100 µm nozzle. Cells were sorted into 32D cell medium, taken back into culture under usual conditions and used in further experiments.

5.3 Characterization of Transduced 32D Cells

5.3.1 Cell Cycle Staining

Before cell cycle analysis, 32D cells were seeded in starvation medium (32D cell medium depleted of IL-3) for 16 h in order to synchronize cell cycle phases. Afterwards, cells were transferred to regular 32D cell growth medium for 12 h. Cell cycle staining of 32D cells was performed using the Click-iT™ Plus EdU Flow Cytometry Assay Kit (Invitrogen), as described in the manufacturer's protocol. In brief, 3 x 10^5 32D cells were seeded on a 6-well plate in medium containing 10 µM EdU and incubated for 1-2 h. Cells were harvested, fixed and permeabilized using the 1X Click-iT Saponin-based permeabilization and wash reagent. Cells subsequently were incubated with the Click-iT™ Plus reaction cocktail (containing Pacific Blue coupled picolyl acide) as described in the manufacturers protocol. Eventually, cells were treated with 2.5 µl of ribonuclease-A (10 µg/ml) and incubated for 20 min at 37°C. Subsequently, cells were pelleted, resuspended in 500 µl 7AAD solution (7AAD diluted 1:200 in PBS) and analyzed by FACS. Cell cycle phases were assigned based on 7AAD and Pacific Blue signals.

5.3.2 Competitive Proliferation Assay

Sorted 32D cells transduced either with LeGO-iC2 *Idh1*-WT or LeGO-iC2 *Idh1*-R132H constructs were seeded together with sorted LeGO-iV2 Empty Vector transduced cells at a 1:1 ratio (2.5 x 10^5 cells each in a 25 cm^2 flask). The initial rate of mCherry and GFP positive cells was determined by FACS directly after seeding. Over a period of 2-3 weeks, the relative percentage of mCherry and GFP positive cells was measured regularly every 3 days by FACS. During this period, cells were split equally for each co-seeding experiment.

5.3.3 Alkaline Comet Assay

Comet assays were conducted as described previously [Schmezer et al., 2001]. In brief, cells were counted, adjusted to an equal concentration, mixed with 0.7 % low-melting

agarose (SeaKem) and plated on comet assay slides (Trevigen). Subsequently, DNA damage was introduced by irradiation of embedded cells with 2 or 5 Gy, using a ^{134}Cs radiation source with a dose rate of 0.575 Gy / min. Afterwards, cells were incubated for indicated recovery / reparation times in RPMI 1640 (Gibco) medium. Cells were lysed overnight using a 1 % TritonX100 lysis buffer. Following lysis, electrophoresis was conducted at 25V for 20 min. Eventually, DNA was stained by incubating cells in a SybrGreen® (Biozol) solution for 30 min. DNA damage analysis and evaluation was performed by fluorescence microscopy using a Metafer-4 (MetaSystems) fully automated cell scanning and analysis platform. Comet assays described in this thesis were performed by Reinhard Gliniorz in collaboration with Dr. Peter Schmezer and PD Dr. Odilia Popanda.

5.3.4 Cell Viability Assay

Transduced and sorted 32D cells were cultured in regular 32D cell medium supplemented with 1 µM or 4 µM olaparib (Selleck Chemicals). The cell viability for each condition was determined after one, two or three days by using the CellTiter-Blue® Cell Viability Assay Kit (Promega) according to the manufacturer's protocol. In brief, cells were seeded in a 96 well plate in regular 32D cell growth medium supplemented with Olaparib at indicated concentrations and incubated for the indicated time frame. Afterwards, CellTiter-Blue® Reagent was added to each well to be analyzed and cells were incubated for 2 h at 37°C. The fluorescence eventually was analyzed using a fluorometer, recording fluorescence at $560_{Ex}/590_{Em}$ nm wave lengths. Cell viability assays described in this thesis were performed by Dr. Ali Bakr.

5.4 Validation of an Expression of the *Idh1*-R132H Mutation

5.4.1 D2HG Assay

The quantification of D2HG in cell lysates and blood serum samples was carried out using an enzymatic assay described previously [Balss et al., 2012]. Cell lysates of 32D or bone marrow cells were obtained by repeated freeze / thaw cycles of cells resuspended in NP40 cell lysis buffer. Lysates were spun down in a table-top centrifuge at 14,000 rpm for 5 min, the supernatant transferred to a new tube and stored at -80°C. Blood serum was obtained by collecting blood in a Microvette® Serum-Gel tube (Sarstedt), followed by centrifugation at 10,000 g for 5 min. The supernatant was transferred to a new tube and stored at -80°C. For the D2HG assay, stored cell lysate or blood serum supernatants were thawed and 100 µl of the respective sample used in further proceedings.

Initially, samples were deproteinized using the Deproteinization Kit (Biovision) according to the manufacturer's protocol. Briefly, 25 µl of perchloric acid solution was added to 100 µl of sample, followed by an incubation for 2 min on ice. Samples were centrifuged for 20 min at 4,000 rpm using a table-top centrifuge. From each supernatant, 95 µl were transferred to a new tube and 5 µl of neutralization solution was added. Samples were again incubated for 2 min on ice and centrifuged for 20 min at 4,000 rpm. From each supernatant, 90µl were transferred to a new tube, from which 25 µl were used

per measurement. For each reaction (reaction volume 100 µl), an assay solution was freshly prepared, containing 100 mM HEPES (Merck Sigma-Aldrich) pH 8.0, 100 µM NAD^+ (Applichem), 0.1 µg (D)-2-hydroxyglutarate dehydrogenase, 5 µM resazurin (Applichem) and 0.01 U/ml diaphorase (MP Biomedical). Twenty-five µl of sample was mixed with 75 µl of assay solution and incubated at RT in the dark for 30 min. Fluorometric detection was carried out in triplicates per sample at $560_{Ex}/610_{Em}$ nm wave lengths.

D2HG standards were prepared in NP40 lysis buffer or blood serum at concentrations of 0.5, 1, 2.5, 5, 7.5, 10, 25 and 50 µM D2HG. For each sample measurement, standards were prepared equally as the corresponding samples and measured on the same plate. The D2HG amount in each sample was determined by calculating a standard curve that was used to correlate fluorescence intensities and D2HG concentrations.

For cell lysate measurements, D2HG concentrations were normalized to the overall protein amount in each lysate, which was determined using the Pierce™ BCA Protein Assay Kit (Thermo Fisher Scientific) according to the manufacturer's protocol. Initially, 1:25 dilutions of each cell lysate were prepared by mixing 3 µl of cell lysate with 72 µl H_2O. Afterwards, a BCA working reagent was prepared by mixing BCA Reagent A with BCA Reagent B at a rate of 50:1. For measurements, 25 µl of diluted lysate was mixed with 200 µl of working reagent and incubated for 30 min at 37˚C. The absorbance was measured at 562 nm in duplicates for each sample.

Albumin standards were prepared at concentrations of 2, 1.5, 1, 0.75, 0.5, 0.25, 0.125 and 0.025 µg/µl. For each measurement, 25 µl of each standard was mixed with 200 µl of BCA working reagent and measured together with cell lysate samples. The protein amount per sample was determined by calculating a standard curve that was used to correlate absorbance and protein amounts.

5.4.2 *Idh1* Locus Recombination-Specific PCR

Genomic DNA was isolated using the QIAamp® DNA Micro Kit (Qiagen) according to the manufacturer's protocol. The following primers were used to analyze the region of the distal lox site, allowing the discrimination between conditional alleles (mutated exon 3 present but not expressed, locus not recombined), knock-in alleles (mutated exon 3 expressed, locus recombined) and wildtype alleles:

Table 5-2: List of Primers Used in *Idh1* Locus Recombination-Specific PCRs.

Primer Name	Sequence
Mf_5539	TGCAAAAATATCCCCCGGCTAGTGA
Er_5538	CACCATTACCACCAACAGCAACATCTC

The PCR was performed using a HotStarTaq DNA Polymerase (Qiagen) according to the manufacturer's protocol using the following cycling conditions: 95˚C for 15 min, 35 cycles of (i) 94˚C for 30 s, (ii) 62˚C for 30 s, (iii) 72˚C for 1 min, 72˚C for 10 min. PCR products eventually were analyzed on an 1 % agarose gel.

5.4.3 *Idh1*-R132H Mutation-Specific PCR

Total RNA was isolated using the RNeasy Micro Kit (Qiagen) according to the manufacturer's protocol including an on-column DNase digestion. Purified RNA was quantified using a NanoDrop spectrophotometer (Thermo Fisher Scientific) and 100-500

ng of total RNA was used in cDNA synthesis reactions. cDNA was synthesized with random hexamer primers using the SuperScript™ II Reverse Transcriptase Kit (Thermo Fisher Scientific) according to the manufacturer's protocol. The expression of *Idh1*-R132H mutated transcripts and wildtype *Idh1* transcripts were quantified using a SYBRGreen® Mastermix (Steinbrenner) in combination with wildtype- or mutation-specific primers on a LightCycler 480 II system (Roche). The following primers were used in combinations Idh1_Exon2_Fwd + Idh1_WT_Rev or Idh1_Exon2_Fwd + Idh1_R132H_Rev; *Gapdh* transcript expression was used for normalization:

Table 5-3: List of Primers Used in *Idh1*-R132H Mutation-Specific PCRs.

Primer Name	Sequence
Idh1_Exon2_Fwd	AGGAGGTTCTGTGGTGGAGA
Idh1_WT_Rev	CTGTATTGGTCCCCATATGCATATC
Idh1_R132H_Rev	CTGTATTGGTCCCCATATGCATAGT
Gapdh_Fwd	TCCTGCACCACCAACTGCTTA
Gapdh_Rev	GGCATGGACTGTGGTCATGAG

The following cycling conditions were used for qRT-pCRs: 95˚C for 15 min, 40 cycles of (i) 95˚C for 15 s, (ii) 56˚C for 20 s, (iii) 72˚C for 15 s, 72˚C for 10 min. Expression analysis was performed based on the $2^{-\Delta\Delta Ct}$ method and amplified PCR products eventually were visualized on an 1% agarose gel.

5.5 Animals

All animals were held and bred under specified pathogen-free conditions in individually ventilated cages at the animal facilities of the German Cancer Research Center (DKFZ). Experimental procedures were approved by the department for 'Veterinär- und Lebensmittelwesen' of the Regierungspräsidium Karlsruhe.

All mouse models (*Idh1*-R132H, *DNMT3A*-R882H, *Scl*-CreERT2, *Rosa26*-EYFP) described in this study initially were crossed with C57Bl/6-Ly5.1 mice to produce mice homozygously expressing the CD45.1 allele. *Idh1*$^{R132H/R132H}$ mice were bred with *Scl*-CreERT2$^{+/d}$ *Rosa26*-EYFP$^{d/d}$ mice to produce *Idh1*$^{+/R132H}$ *Scl*-CreERT2$^{+/d}$ *Rosa26*-EYFP$^{+/d}$ mice which were used in described experiments. *DNMT3A*$^{+/R882H}$ mice were crossed with *Scl*-CreERT2$^{+/d}$ *Rosa26*-EYFP$^{d/d}$ mice to yield *DNMT3A*$^{+/R882H}$ *Scl*-CreERT2$^{+/d}$ *Rosa26*-EYFP$^{+/d}$ mice which were used in described experiments. *Idh1*$^{R132H/R132H}$ *DNMT3A*$^{+/R882H}$ mice were bred with *Scl*-CreERT2$^{+/d}$ *Rosa26*-EYFP$^{d/d}$ mice to generate *Idh1*$^{+/R132H}$ *DNMT3A*$^{+/R882H}$ *Scl*-CreERT2$^{+/d}$ *Rosa26*-EYFP$^{+/d}$ mice which were used in described experiments. *Idh1*$^{+/+}$ *DNMT3A*$^{+/+}$ *Scl*-CreERT2$^{+/d}$ *Rosa26*-EYFP$^{+/d}$ mice were used as CTRL mice in all experiments. For all experimental mice, genotypes were verified by established genotyping PCRs. Recipient C57BL/6J mice were obtained from Harlan Laboratories.

5.5.1 Tamoxifen Injections

To induce Cre recombinase activity, six- to ten-week old mice were intraperitoneally injected with 2 mg tamoxifen per day for five consecutive days. Tamoxifen solutions were prepared by mixing 50 mg of tamoxifen (Merck Sigma Aldrich) with 500 µl ethanol and

4.5 ml sunflower oil (Merck Sigma Aldrich). Fresh tamoxifen solutions were prepared every second day.

5.6 Analysis of Murine Blood and Bone Marrow Composition

FACS antibody panels used to analyze blood and bone marrow composition are indicated in the respective tables. All antibodies were titrated and applied in appropriate dilutions. Cells were stained in PBS (Merck Sigma Aldrich) supplemented with 2 % fetal calf serum (FCS, Gibco) (PBS/FCS) for at least 20 min at 4°C, washed once, resuspended in PBS/FCS and filtered. Flow cytometric analysis of blood and bone marrow cells was carried out on LSRII or Fortessa flow cytometer (BD Biosciences). FACS data was analyzed using the FlowJo 10 software package.

5.6.1 Peripheral Blood Analysis

Peripheral blood was drawn from facial veins and collected in K3 EDTA coated tubes (Sarstedt). Cell counts were determined using a HemaVet 950 FS veterinary hematology system (Drew Scientific).

For flow cytometric analysis, 30 µl of peripheral blood was mixed with 1 ml of Ammonium-Chloride-Potassium (ACK)-lysis buffer (Lonza) to lyse red blood cells and incubated for 10 min on RT. Afterwards, cells were centrifuged at 400 g for 5 min, and stained with the respective antibody panels depicted in Table 5-4 for 30 min at 4°C. Cells were washed once, resuspended in 200 µl PBS/FCS, filtered and analyzed by FACS.
The antibody panel depicted in Table 5-4 allows the quantification of CD45.1, CD45.1/2 and CD45.2 cell rates as well as B-cell, T-cell and myeloid cell frequencies.

Table 5-4: Antibody Panel for Flow Cytometric Analysis of Peripheral Blood.

Antigen	Clone	Fluorophore	Company
CD45.1	A20	APC-Cy7	Invitrogen
CD45.2	104	Pacific Blue	BioLegend
CD4	GK1.5	PE-Cy7	Invitrogen
CD8a	53-6.7	PE-Cy7	eBioscience
B220	RA3-6B2	PE-Cy7	eBioscience
B220	RA3-6B2	APC	eBioscience
Gr1	RB6-8C5	APC	eBioscience
CD11b	M1/70	APC	Invitrogen

5.6.2 Bone Marrow Analysis

Mice were sacrificed by cervical dislocation and femora, tibiae, iliac crests and spine vertebrae were dissected and cleaned. Bones were crushed in Iscove's modified Dulbecco's medium (IMDM, Gibco) and the cell suspension filtered through a 40 µm cell strainer (Greiner Bio-One).

Cells were counted and stained with the respective antibody panels depicted in Table 5-5, Table 5-6, Table 5-7, Table 5-8 or Table 5-9 for 30 min at 4°C (maximal 2 x 10^7 cells stained in 500 µl PBS/FCS). Cells were washed once, resuspended in PBS/FCS, filtered and analyzed by FACS.

The antibody panel depicted in Table 5-5 was used for the quantification of LT-HSCs, ST-HSCs, MPP2, MPP3+4, MPP5, and LSK cell populations.

Table 5-5: Antibody Panel for Flow Cytometric Analysis of Stem Cell- and Multipotent Progenitor Populations in the Bone Marrow.

Antigen	Clone	Fluorophore	Company
CD4	GK1.5	PE-Cy7	Invitrogen
CD8a	53-6.7	PE-Cy7	eBioscience
CD11b	M1/70	PE-Cy7	eBioscience
Gr1	RB6-8C5	PE-Cy7	eBioscience
B220	RA3-6B2	PE-Cy7	eBioscience
Ter119	TER119	PE-Cy7	eBioscience
cKit	2B8	APC	Invitrogen
Sca1	D7	APC-Cy7	BD Bioscience
CD150	TC15-12F12.2	PE-Cy5	BioLegend
CD48	HM48-1	PE	eBioscience
CD34	RAM34	eFluor 450	Invitrogen

The antibody panel depicted in Table 5-6 was used for the quantification of MEP, CMP, GMP and LS-K populations.

Table 5-6: Antibody Panel for Standard Flow Cytometric Analysis of Committed Progenitor Populations in the Bone Marrow.

Antigen	Clone	Fluorophore	Company
CD4	GK1.5	PE-Cy7	Invitrogen
CD8a	53-6.7	PE-Cy7	eBioscience
CD11b	M1/70	PE-Cy7	eBioscience
Gr1	RB6-8C5	PE-Cy7	eBioscience
B220	RA3-6B2	PE-Cy7	eBioscience
Ter119	TER119	PE-Cy7	eBioscience
cKit	2B8	APC	Invitrogen
Sca1	D7	APC-Cy7	BD Bioscience
CD34	RAM34	Alexa Fluor 700	eBioscience
CD16/32	93	eFluor 450	eBioscience

The antibody panel depicted in Table 5-7 was used for the quantification of MEP, CMP, GMP, Ly6C$^+$ GMP, CD115$^+$ GMP, cMoP and LS-K populations.

Table 5-7: Antibody Panel for Extended Flow Cytometric Analysis of Committed Progenitor Populations in the Bone Marrow.

Antigen	Clone	Fluorophore	Company
CD4	GK1.5	PE-Cy7	Invitrogen
CD8a	53-6.7	PE-Cy7	eBioscience
CD11b	M1/70	PE-Cy7	eBioscience
Gr1	RB6-8C5	PE-Cy7	eBioscience
B220	RA3-6B2	PE-Cy7	eBioscience
Ter119	TER119	PE-Cy7	eBioscience
cKit	2B8	APC	Invitrogen
Sca1	D7	APC-Cy7	BD Bioscience
CD34	RAM34	Alexa Fluor 700	eBioscience
CD16/32	93	eFluor 450	eBioscience
Ly6C	HK1.4	BV605	BioLegend
CD115	AFS98	BV711	BioLegend

The antibody panel depicted in Table 5-8 was used for the quantification of B-cells, T-cells and myeloid cells.

Table 5-8: Antibody Panel for Standard Flow Cytometric Analysis of Differentiated Cells in the Bone Marrow.

Antigen	Clone	Fluorophore	Company
CD4	GK1.5	PE-Cy7	Invitrogen
CD8a	53-6.7	PE-Cy7	eBioscience
B220	RA3-6B2	PE-Cy7	eBioscience
B220	RA3-6B2	APC	eBioscience
Gr1	RB6-8C5	APC	eBioscience
CD11b	M1/70	APC	Invitrogen

The antibody panel depicted in Table 5-9 was used for the quantification of B-cells, T-cells, monocytes and granulocytes.

Table 5-9: Antibody Panel for Extended Flow Cytometric Analysis of Differentiated Cells in the Bone Marrow.

Antigen	Clone	Fluorophore	Company
CD4	GK1.5	PE-Cy7	Invitrogen
CD8a	53-6.7	PE-Cy7	eBioscience
B220	RA3-6B2	APC	eBioscience
CD11b	M1/70	BV785	BioLegend
Ly6C	HK1.4	BV605	BioLegend
Ly6G	1A8	APC-Cy7	BioLegend

5.7 Isolation and Purification of Surface Marker-Defined Bone Marrow Populations

Mice were sacrificed by cervical dislocation and femora, tibiae, iliac crests and spine vertebrae were dissected and cleaned. Bones were crushed in Iscove's modified Dulbecco's medium (IMDM, Gibco) and the cell suspension filtered through a 40 µm cell strainer (Greiner Bio-One). Obtained TBM cells were either directly stained with respective antibody panels (e.g. to sort $CD45^+$ TBM cells) or further processed to enrich Lin^- bone marrow cells.

5.7.1 Enrichment of Lineage Negative Bone Marrow Cells

A density gradient centrifugation using Histopaque 1083 (Merck Sigma-Aldrich) was carried out to enrich low-density mononuclear cells (LDMNCs). Here, 6 ml of the filtered cell suspension (maximal concentration of 3 x 10^7 cells / ml) was loaded on top of 6 ml of Histopaque 1083 and centrifuged at 400 g for 20 min at 21°C without brakes. LDMNCs were harvested by collecting upper phase, interphase and lower phase. The remaining cell pellet was resuspended in 6 ml of fresh IMDM and subjected to an additional round of density gradient centrifugation. LDMNCs from both centrifugations were pooled, centrifuged and resuspended in 3 ml PBS/FCS. Cells were incubated for 30 min at 4°C with a cocktail of biotin-conjugated rat anti-mouse lineage antibodies (Table 5-10) at a concentration of 112.8 µl lineage cocktail per 1 x 10^8 cells. Afterwards, cells were washed once in PBS/FCS and incubated with anti-rat IgG-coated magnetic Dynabeads™ from the

Dynabeads™ Untouched™ Mouse T Cells Kit (Invitrogen) for 45 min at 4˚C. Magnetic depletion of Lin^+ bone marrow cells using a DynaI MPC-6 magnet (Invitrogen) resulted in an enrichment of Lin^- bone marrow cells. Lineage-depleted cells were washed in PBS/FCS and further stained with respective antibody panels.

Table 5-10: Cocktail of Biotin-Conjugated Anti-Mouse Lineage Antibodies.

Antigen	Clone	Concentration	Company
CD5	53-7.3	4.2 µg / ml	BD Bioscience
CD45R / B220	RA3-6B2	2.8 µg / ml	BD Bioscience
CD11b / Mac1	M1/70	2.6 µg / ml	BD Bioscience
CD8a	53-6.7	4.2 µg / ml	BD Bioscience
Ly6G and Ly6C (Gr1)	RB6-8C5	2.4 µg / ml	BD Bioscience
Ter119	TER-119	2.6 µg / ml	BD Bioscience

5.7.2 Fluorescence Activated Cell Sorting of Hematopoietic Cell Populations

FACS antibody panels used to label hematopoietic cells are indicated in the respective tables. All antibodies were titrated and applied in appropriate dilutions. Cells at a maximal concentration of 1 x 10^7 cells / ml were stained in PBS/FCS for at least 20 min at 4˚C, washed, resuspended in PBS/FCS and filtered. Sorting of hematopoietic cell populations was carried out at Aria I, Aria II, or Fusion cell sorter systems (BD Biosciences) using a 100 µm nozzle. Cells were sorted into 1.5 ml tubes containing 300 µl StemSpan SFEM medium (Stemcell Technologies) in case cells were taken into culture or used in transplantations. For further RNA-seq experiments, cells were sorted directly into 200 µl Arcturus® PicoPure® Extraction Buffer (Applied Biosystems), snap-frozen in liquid nitrogen and stored at -80˚C until further processing. For further scRNA-seq experiments, cells were sorted into PBS/FCS.

Sorting of $CD45^+$ Total Bone Marrow Cells

Following crushing of bones and filtering of cell suspensions, cells were counted and centrifuged at 400 g for 5 min. Afterwards, cells were stained with the antibody panel depicted in Table 5-11 for 20 min at 4˚C. Cells were washed, resuspended in PBS/FCS + DAPI (diluted 1:1000), filtered and subjected to cell sorting.

Table 5-11: Antibody Panel for the Isolation of $CD45^+$ Total Bone Marrow Cells.

Antigen	Clone	Fluorophore	Company
CD45.1	A20	PE	eBioscience
DAPI	-	-	PeproTech

Sorting of Cell Populations from Lineage-Depleted Bone Marrow Cells

Following the enrichment of Lin^- bone marrow cells, cells were washed, centrifuged at 400 g for 5 min and stained with antibody panels depicted in Table 5-12, 5-13 or 5-14 for 20 min at 4˚C. Cells were washed, resuspended in PBS/FCS, filtered and subjected cell sorting.

The antibody panel depicted in Table 5-12 was used to isolate HSC and MPP populations from lineage-depleted bone marrow cells.

Table 5-12: Antibody Panel for the Isolation of Stem Cell- and Multipotent Progenitor Populations.

Antigen	Clone	Fluorophore	Company
CD4	GK1.5	PE-Cy7	Invitrogen
CD8a	53-6.7	PE-Cy7	eBioscience
CD11b	M1/70	PE-Cy7	eBioscience
Gr1	RB6-8C5	PE-Cy7	eBioscience
B220	RA3-6B2	PE-Cy7	eBioscience
Ter119	TER119	PE-Cy7	eBioscience
Streptavidin	-	PE-Cy7	BioLegend
cKit	2B8	APC	Invitrogen
Sca1	D7	APC-Cy7	BD Bioscience
CD150	TC15-12F12.2	PE-Cy5	BioLegend
CD48	HM48-1	PE	eBioscience
CD34	RAM34	eFluor 450	Invitrogen

The antibody panel depicted in Table 5-13 was used to isolate committed progenitor populations from lineage-depleted bone marrow cells.

Table 5-13: Antibody Panel for the Isolation of Committed Progenitor Populations.

Antigen	Clone	Fluorophore	Company
CD4	GK1.5	PE-Cy7	Invitrogen
CD8a	53-6.7	PE-Cy7	eBioscience
CD11b	M1/70	PE-Cy7	eBioscience
Gr1	RB6-8C5	PE-Cy7	eBioscience
B220	RA3-6B2	PE-Cy7	eBioscience
Ter119	TER119	PE-Cy7	eBioscience
Streptavidin	-	PE-Cy7	BioLegend
cKit	2B8	APC	Invitrogen
Sca1	D7	APC-Cy7	BD Bioscience
CD34	RAM34	Alexa Fluor 700	eBioscience
CD16/32	93	eFluor 450	eBioscience
Ly6C	HK1.4	BV605	BioLegend
CD115	AFS98	BV711	BioLegend

The antibody panel depicted in Table 5-14 was used to isolate LSK, LS-K and Lin^- populations from lineage-depleted bone marrow cells.

Table 5-14: Antibody Panel for the Isolation of LSK, LS-K and Lin^- Populations.

Antigen	Clone	Fluorophore	Company
CD4	GK1.5	PE-Cy7	Invitrogen
CD8a	53-6.7	PE-Cy7	eBioscience
CD11b	M1/70	PE-Cy7	eBioscience
Gr1	RB6-8C5	PE-Cy7	eBioscience
B220	RA3-6B2	PE-Cy7	eBioscience
Ter119	TER119	PE-Cy7	eBioscience
Streptavidin	-	PE-Cy7	BioLegend
cKit	2B8	APC	Invitrogen
Sca1	D7	APC-Cy7	BD Bioscience

5.8 Transplantations

5.8.1 Primary Transplantations

Sorted cells to be transplanted were centrifuged at 400 g for 5 min and resuspended in an appropriate volume of PBS. If necessary, cells were mixed with the indicated number of supportive bone marrow cells derived from heterozygous C57Bl/6-Ly5.1 mice. For Lin^- transplantations, no supportive bone marrow was used. Recipient mice were irradiated with 2 x 500 Rad one day prior to the transplantation. Per recipient mouse, 200 µl of cell suspension was injected into the tail vain using a 1 ml insulin syringe (BD Medical).

5.8.2 Secondary and Tertiary Transplantations

For secondary and tertiary transplantations, TBM cells from primary or secondary recipients were transplanted. Femora, tibiae, iliac crests and spine vertebrae were crushed and the cell suspension filtered. Cells were counted and the appropriate number of cells resuspended in an adequate volume of PBS. Per secondary or tertiary recipient mouse, 3 x 10^6 TBM cells were transplanted in a volume of 200 µl injected into the tail vain.

5.9 Colony-Forming Unit Assay

LT-HSCs were sorted in 300 µl StemSpan SFEM (StemCell Technologies) medium, centrifuged at 400 g for 5 min and resuspended in PBS at a concentration of 600 cells per 300 µl. Next, 300 µl of the cell suspension were mixed with 2.7 ml of MethoCult™ M3434 methylcellulose medium (StemCell Technologies) and 1 ml was seeded on a 5 cm^2 cell culture dish (Corning) in technical duplicates. Dishes were incubated for 7 days in a humidified 10 cm^2 dish (Corning) at 37°C and 5 % CO_2. Following incubation, colonies were counted manually and cells were harvested by diluting and resuspending the methylcellulose cell suspension in 2 ml PBS per plate. Harvested cells were centrifuged at 400 g for 5 min and resuspended in an appropriate volume of PBS and counted. Five thousand cells were resupended in 300 µl of PBS, mixed with fresh 2.7 ml of MethoCult™ M3434 (StemCell Technologies) methylcellulose medium and seeded on 5 cm^2 dishes. Secondary and all subsequent platings were incubated for 7-10 days at 37°C and 5 % CO_2. Replating procedures were repeated until no growth of colonies was observed. CFU assays described in this thesis were performed in collaboration with Umut Kilik.

5.10 Bulk RNA-Sequencing

5.10.1 RNA Extraction

Cells were sorted into Arcturus® PicoPure® Extraction Buffer (Applied Biosystems), snap-frozen in liquid nitrogen and stored at -80˚C until further processing. Total RNA extraction was performed using the Arcturus® PicoPure® RNA Isolation Kit (Applied Biosystems) according to the manufacturer's protocol. DNA was digested on-column using the RNase-Free DNase Set (Qiagen). RNA was eluted in 10 µl H_2O and concentrations were determined with a Qubit Fluorometer (Thermo Fisher Scientific) using a Qubit RNA High Sensitivity Kit (Invitrogen). RNA profiles were validated on an Agilent 2100 Bioanalyzer platform (Agilent Technologies) with an Agilent RNA 6000 Pico Kit (Agilent Technologies). RNA concentrations of samples were adjusted to an equal concentration for further processing.

5.10.2 Generation of RNA-seq Libraries

RNA-seq libraries were generated from extracted total RNA using the previously described Smart-seq2 protocol [Picelli et al., 2014]. In brief, cDNA was synthesized in a reverse transcription and template-switching reaction and amplified by pre-amplification PCR. cDNA was purified using Agencourt Ampure XP Beads (Beckaman Coulter) and quality-checked on an Agilent 2100 Bioanalyzer platform (Agilent Technologies) using an Agilent High Sensitivity DNA Kit (Agilent Technologies). Libraries were generated using the Nextera XT DNA Library Prep Kit (Illumina) according to the manufacturer's protocol. cDNA was tagmented and adapter-ligated fragments amplified by PCR. Amplified and indexed libraries were purified using Agencourt Ampure XP Beads (Beckaman Coulter) and quality-checked on an Agilent 2100 Bioanalyzer platform (Agilent Technologies) using an Agilent High Sensitivity DNA Kit (Agilent Technologies). Library concentrations were determined with a Qubit Fluorometer (Thermo Fisher Scientific) using a Qubit DNA High Sensitivity Kit (Invitrogen). For sequencing, libraries were pooled at equimolar ratios.

5.10.3 Sequencing of RNA-seq Libraries

Libraries were sequenced by 100 bp paired-end sequencing on a HiSeq 4000 platform (Illumina). Sequencing was performed by the Genomics and Proteomics Core Facility at the DKFZ.

5.10.4 Analysis of Bulk RNA-Sequencing Data

Alignment and Filtering

Sequencing reads initially were aligned and filtered by the Omics IT and Data Management Core Facility of the DKFZ in collaboration with Stephen Krämer. The workflow was created by Dr. Naveed Ishaque. In brief, reads were aligned the GRCm38 (mm10) genome assembly using the 'STAR' aligner [Dobin et al., 2013]. Duplicate marking of the resultant main alignment file was performed using 'sambamba' [Tarasov et al., 2015], chimeric files were sorted using 'samtools' (i.e. 'samtools sort') [Li et al., 2009] and BAM indexes were

generated using 'sambamba' [Tarasov et al., 2015]. Quality control analysis was performed using 'samtools' (i.e. 'samtools flagstat') [Li et al., 2009] and 'RNA-SeQC' [DeLuca et al., 2012]. Eventually, gene-specific read counting was performed using 'featureCounts' [Liao et al., 2014].

Data Transformation and Sample Clustering

Generated count matrices were filtered to exclude genes with low read counts (< 5) and counts were normalized using a regularized log (rlog) transformation implemented in the 'DESeq2' R-package [Love et al., 2014]. Pearson correlation coefficients, sample distances and principal components were computed using the 'cor.test', 'dist' and 'prcomp' functions implemented in the 'R stats' R-package.

Differential Expression Analysis

Differential expression between genotypes was determined using the 'DESeq2' R-package [Love et al., 2014]. The status of both *Idh1*-R132H and *DNMT3A*-R882H mutations was incorporated into the design formula (~ Mutation 1 * Mutation 2) and contrasts were specified accordingly to extract differential gene expression between genotypes in a pairwise comparison. Log2 fc shrinkage was performed using the 'ashr' shrinkage estimator [Stephens, 2017].

GSEA was performed using the 'GSEA' software from the Broad Institute [Mootha et al., 2003, Subramanian et al., 2005]. Filtered and rlog transformed count matrices were used as input and enriched gene sets were determined using 'Hallmark' gene sets as reference.

Upstream regulator analysis was performed using the Ingenuity Pathway Analysis platform (Qiagen). Differential expression results tables, containing log2 fc values, p-values, and false discovery rate (fdr) values for a given pairwise comparison were used as an input. Log2 fc < -0.2 or > 0.2 and fdr < 0.1 were used as cutoffs for the upstream regulator analysis.

5.11 Single Cell RNA-Sequencing

5.11.1 Generation and Sequencing of Single-Cell RNA-Sequencing Libraries

Hematopoietic cell layers ($CD45^+$ TBM, LS-K, LSK) were sorted into PBS/FCS as described previously. ScRNA-seq libraries were generated at the Single-Cell OpenLab of the DKFZ by Katharina Bauer, Dr. Mark Hartmann and Maximilian Schönung using a Chromium Single-Cell 3' Reagent Kit v2 (10X Genomics) according to the manufacturer's instructions. In brief, sorted cells were loaded onto a Chip A Single Cell, and mixed with a reverse transcription master mix and Single Cell 3' Gel Beads. The mixture was processed in a Chromium Controller to generate Gel Bead-In-EMulsions (GEMs) and further incubated at (i) 53°C for 45 min, (ii) 85°C for 5 min, to synthesize barcoded full-length cDNA. Silane magnetic beads were used to remove leftover reagents and primers, the cDNA was released from GEMs, purified and eventually amplified by PCR. Amplified cDNA was again purified and sequencing libraries were constructed by enzymatic fragmentation, followed by end repair, A-tailing, adaptor ligation, and a final PCR reaction. Final libraries were

purified, quality-checked and eventually sequenced by 100 bp paired-end sequencing on a HiSeq 4000 platform (Illumina) at the Genomics and Proteomics Core Facility of the DKFZ.

5.11.2 Analysis of Single-Cell RNA-Sequencing Data

Initial processing of reads and preliminary analysis of scRNA-seq data was carried out by Abdelrahman Mahmoud. Reads were demultiplexed, aligned and filtered using the 10X Genomics 'CellRanger' pipeline in order to generate feature-barcode gene expression matrices. Downstream data processing was performed using the 'Seurat' R-package v3 [Butler et al., 2018, Stuart et al., 2019].

Filtering, Normalization, Scaling and Dimensionality Reduction of Data

Low-quality cells were filtered based on the following cutoffs: < 200 genes per cell, > 5000 genes per cell, > 10 % mitochondrial genes. Data was normalized and scaled using the 'sctransform' wrapper implemented in the 'Seurat' package, regressing out the number of UMIs and the mitochondrial percentage. Principal component analysis was performed using the 'RunPCA' function and the first 50 PCAs were used in downstream analysis. Cell cluster were determined based on a shared nearest neighbor (SNN) algorithm using the 'FindNeighbors' and 'FindCluster' functions ('FindCluster' resolution = 3). Dimensionality reduction was conducted using the UMAP algorithm ('RunUMAP' function, parameters: dims = 1:25, n.neighbors = 200, min.dist = 0.99).

Annotation of Cluster

Cluster-specific signature genes were identified by running a differential expression analysis between cluster using a Wilcoxon rank-sum test ('FindAllMarker' function) with log fc and minimum fraction cutoffs of 0.25 (parameters min.pct = 0.25, logfc.threshold = 0.25). Cell cluster were annotated manually based on the list of cluster-specific signature genes. Annotation of cluster was carried out in collaboration with Dr. Simon Haas and Dr. Mark Hartmann. Cluster with equal annotation were aggregated to cell type specific clusters (e.g. ST-HSC cluster #1 and ST-HSC cluster #2). Sample distances between genotype specific cells for each cluster were determined using the 'dist' function implemented in the 'R stats' R-package.

Differential Expression Between Genotypes

Differential expression between genotypes was determined for each cell type-specific cluster using the 'FindMarker' function with a log fc cutoff of 0.25. DE genes for each genotype were called in a one vs. all comparison between all cells from the respective cluster or cell type.
To identify genotype- specific enriched gene sets, the overlap between DE genes and gene sets deposited at the Molecular Signature Database (MSigDB) [Liberzon et al., 2015, Liberzon et al., 2011] were computed. 'Hallmark' gene sets, 'Gene Ontology' gene sets and 'Immunologic Signatures' gene sets were used as reference for computing overlaps.

Diffusion Maps and Lineage Interference

Diffusion maps were created using the 'destiny' R-package and are based on a spectral non-linear dimensionality reduction method [Angerer et al., 2016]. Normalized single-cell gene expression data was used as an input. Based on the computed diffusion distances,

lineages were reconstructed based on pseudotime inference using the 'slingshot' R-package [Street et al., 2018].

Cell Cycle Assignments

Cell cycle phases for each cell were determined using the 'CellCycleScoring' function implemented in the 'Seurat' R-package. Cell cycle scores were assigned based on the expression of pre-defined G2/M- and S-phase markers.

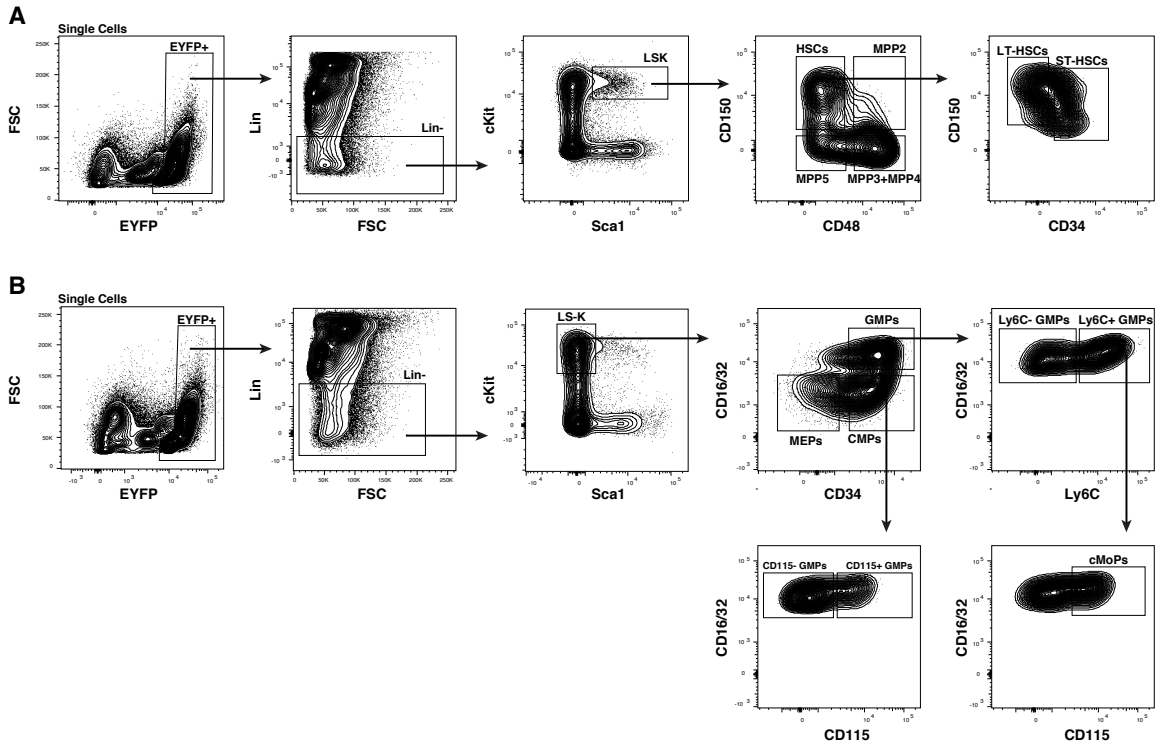
5.12 Data Processing, Visualization and Statistical Analysis

Data was processed and visualized using the R environment. Processing and transformation of data was predominantly performed using the 'dplyr' R-package. Visualization of data was predominantly performed using the 'ggplot2', 'pheatmap', or 'Seurat' R-packages. Statistical tests were computed using the 'stats' or 'ggsignif' R-packages and are stated in the respective figure legends alongside sample sizes (n) and p-values. Error bars indicate standard deviation (SD). Figures were designed using Affinity Designer 1.5.5.

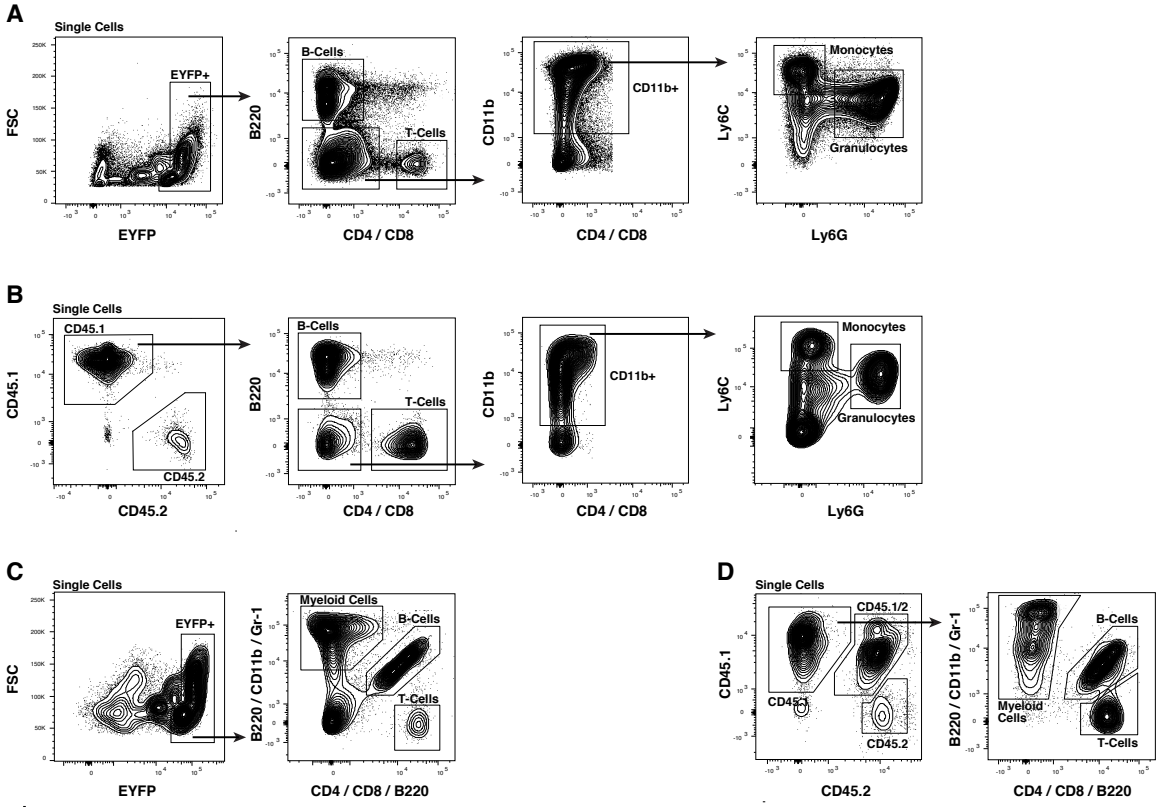
A

Appendix

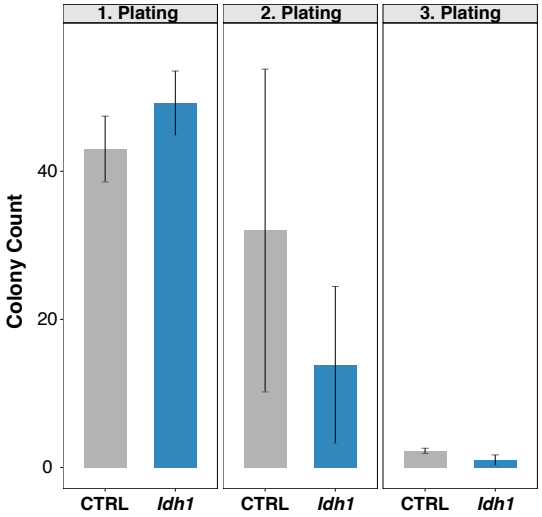
A.1 Supplemental Figures



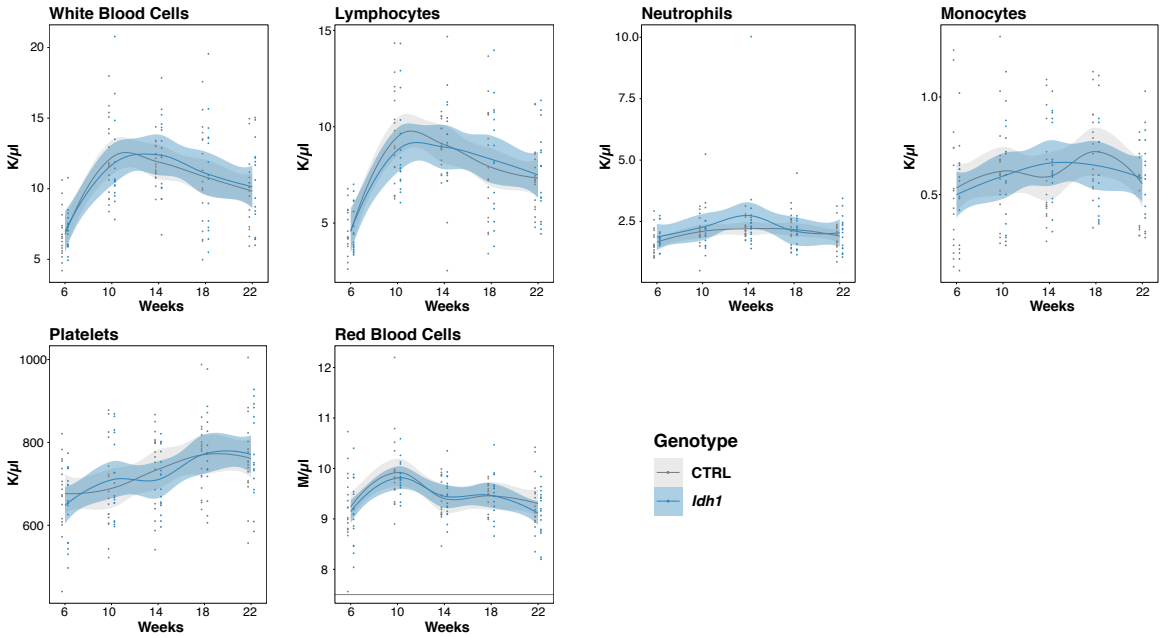
Supplemental Figure A-1: Representative FACS Analysis and Sorting Schemes to Quantify or Isolate Hematopoietic Stem- and Progenitor Cells.
(A) Representative gating scheme to isolate and quantify EYFP$^+$ stem cell- and multipotent progenitor cell populations within the bone marrow of mice. **(B)** Representative gating scheme to isolate and quantify EYFP$^+$ committed progenitor cell populations within the bone marrow of mice.



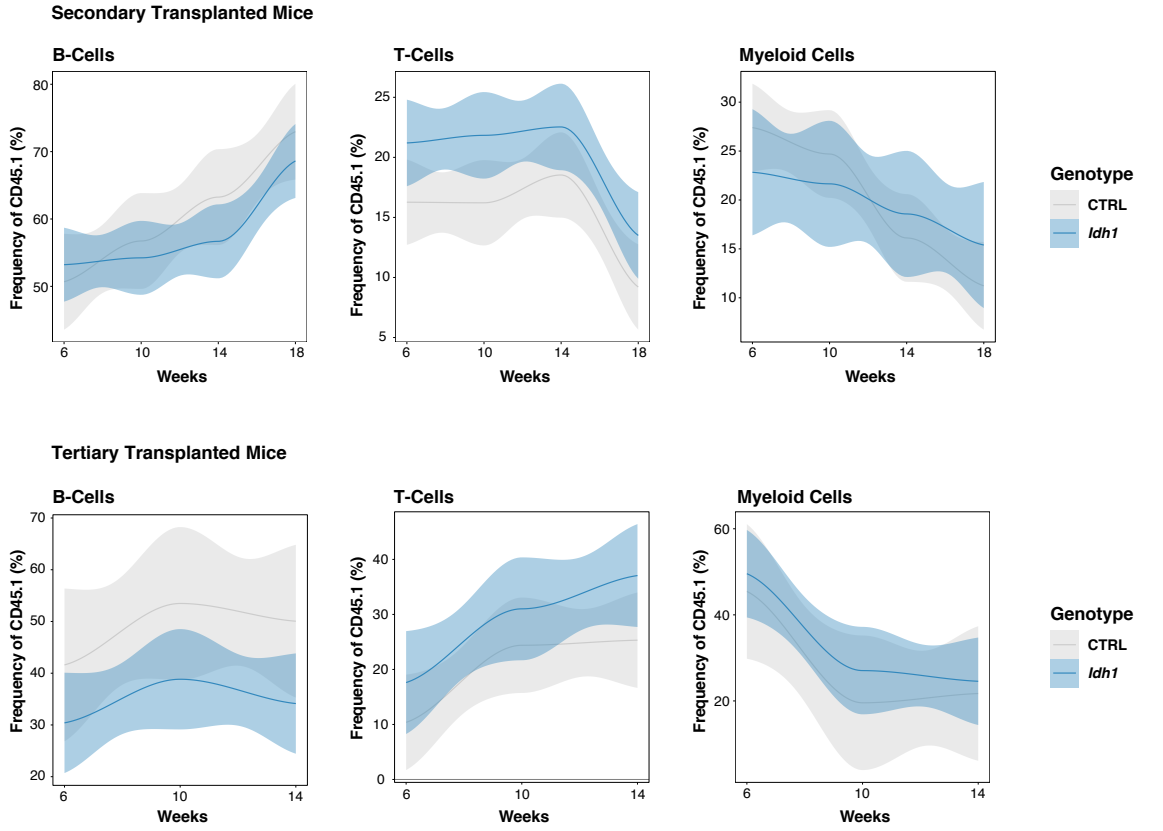
Supplemental Figure A-2: Representative FACS Analysis and Sorting Schemes to Quantify or Isolate Differentiated Hematopoietic Cells.
(A) Representative gating scheme to quantify EYFP⁺ B-cells, T-cells, granulocytes and monocytes in the bone marrow of mice. **(B)** Representative gating scheme to quantify CD45.1⁺ cell-derived B-cells, T-cells, granulocytes and monocytes in the peripheral blood of mice. **(C)** Representative gating scheme to quantify EYFP⁺ B-cells, T-cells and overall myeloid cells in the bone marrow of mice. **(D)** Representative gating scheme to quantify CD45.1⁺ cell-derived B-cells, T-cells and overall myeloid cells in the peripheral blood of mice.



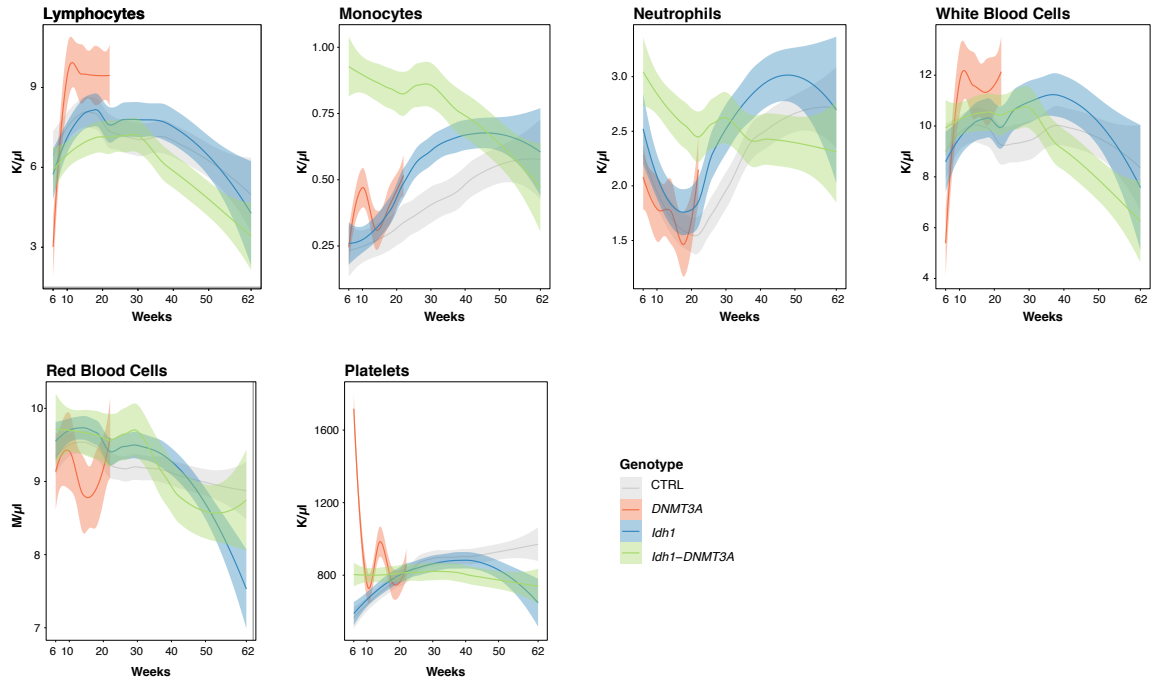
Supplemental Figure A-3: Colony-Formation Assay of *Idh1*-R132H Mutated LT-HSCs.
LT-HSCs were isolated from the bone marrow of primary tamoxifen-injected mice, seeded in methylcellulose medium and replated after 7-10 days. Before each replating the number of colonies was counted (n = 3, error bars indicate mean ± SD).



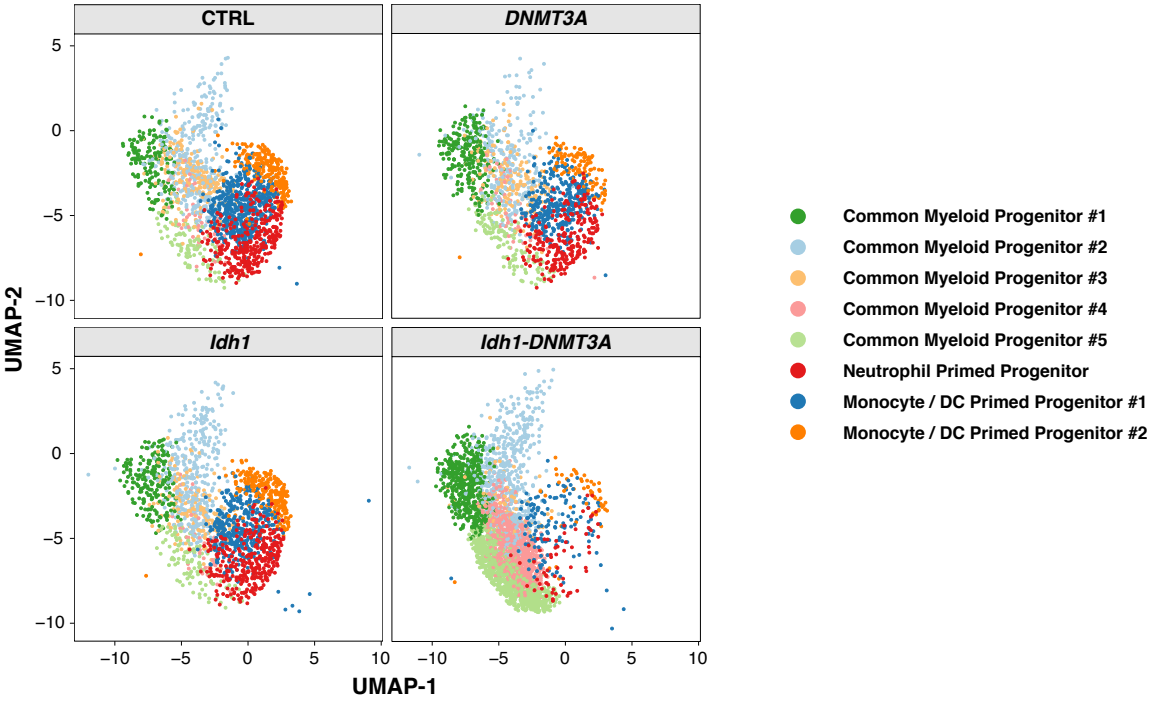
Supplemental Figure A-4: Blood Parameters of LT-HSC Transplanted *Idh1* and CTRL Mice.
Quantification of overall blood cell counts of indicated cell types using a HemaVet 950 FS hematology system. Y-axis counts are shown either as 1,000(K) / µl or 1,000,000(M) / µl (CTRL n = 17, *Idh1* n = 17, three independent transplantations, curves were smoothened based on a loess (locally estimated scatterplot smoothing) regression model, confidence interval (0.95) plotted around curve).



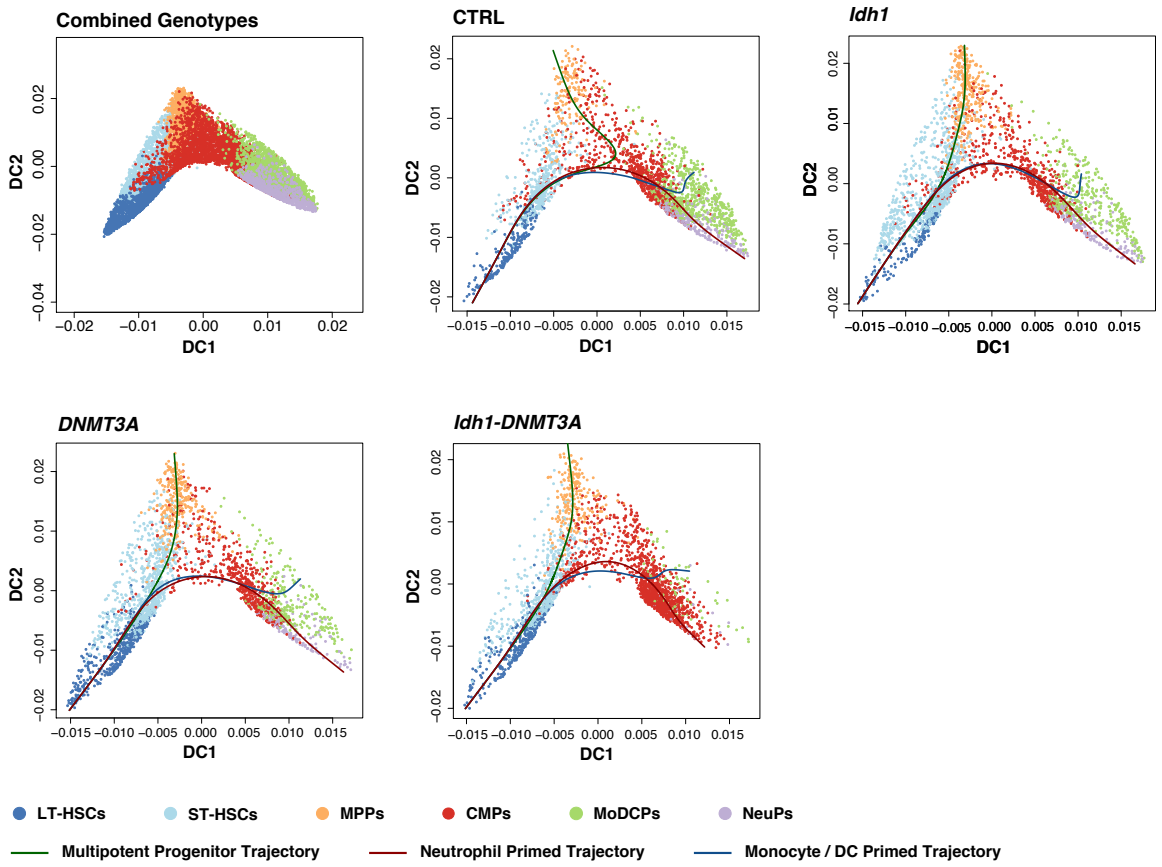
Supplemental Figure A-5: Lineage Output of Serially Transplanted *Idh1*-R132H and CTRL LT-HSCs.
Quantification of frequencies of indicated cell types in the peripheral blood of primary and secondary recipient mice using flow cytometry. Frequencies are calculated in relation to CD45.1$^+$ cells (secondary transplantation n = 6, tertiary transplantation n = 6, curves were smoothened based on a loess (locally estimated scatterplot smoothing) regression model, confidence interval (0.95) plotted around curve).



Supplemental Figure A-6: Blood Parameters of Single-Mutant and *Idh1-R132H DNMT3A-R882H* Double-Mutant LT-HSC Transplanted Mice.
Quantification of overall blood cell counts of indicated cell types using a HemaVet 950 FS hematology system. Y-axis counts are shown either as 1,000(K) / µl or 1,000,000(M) / µl (CTRL n = 19, *Idh1* n =21, *Idh1-DNMT3A* n= 18, *DNMT3A* n =12, two to four independent transplantations per genotype, curves were smoothened based on a loess (locally estimated scatterplot smoothing) regression model, confidence interval (0.95) plotted around curve) (transplantation and analysis of *DNMT3A* mice was performed by Dr. Sina Stäble and Dr. Natasha Anstee).

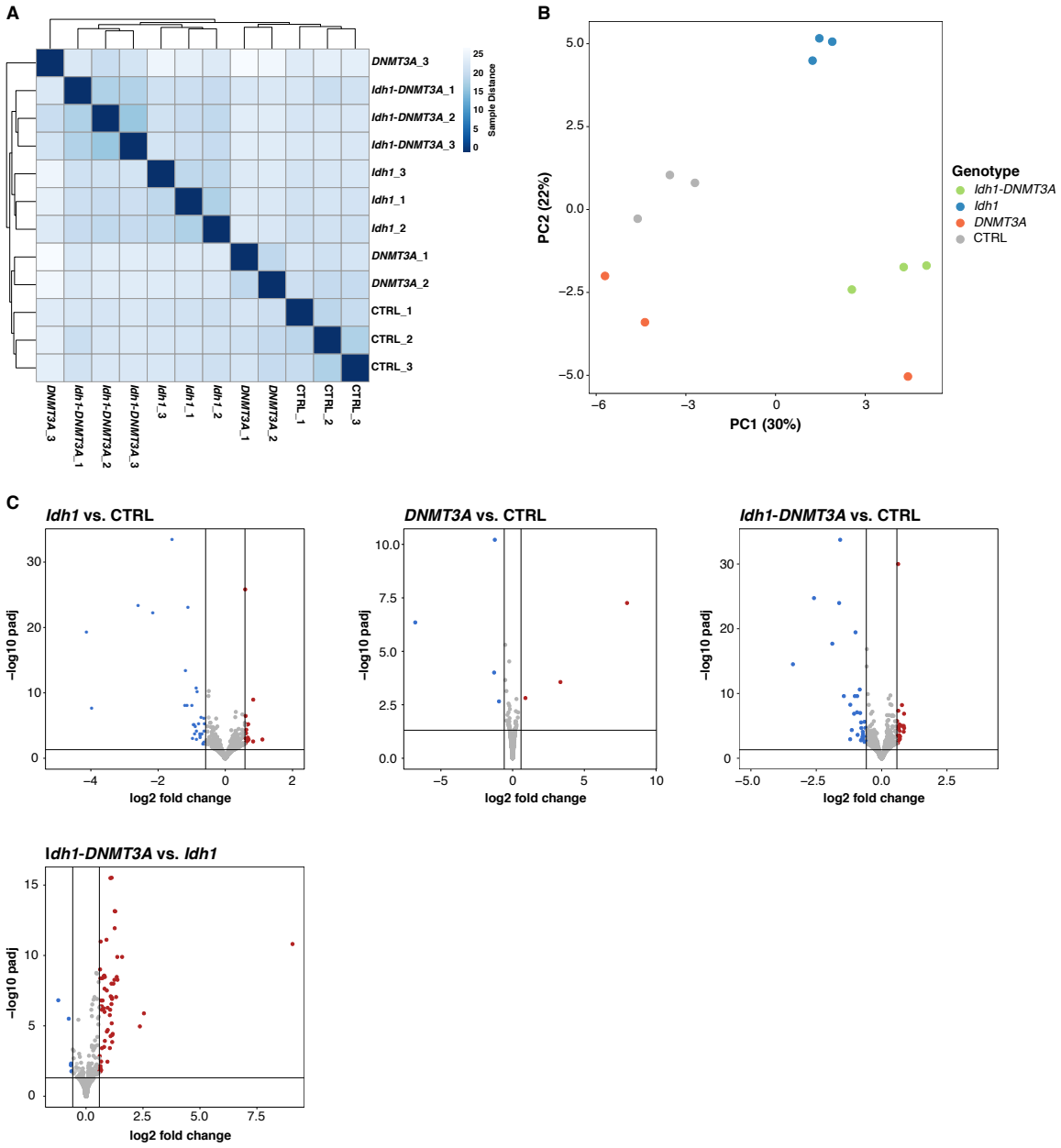


Supplemental Figure A-7: UMAP-Based Representation of Myeloid Progenitor Subclusters per Genotype.
Distribution of annotated myeloid cells per individual cluster. All cells per genotype are shown. Increased numbers of *Idh1-DNMT3A* cells are seen especially in common myeloid progenitor cluster #4 and #5.

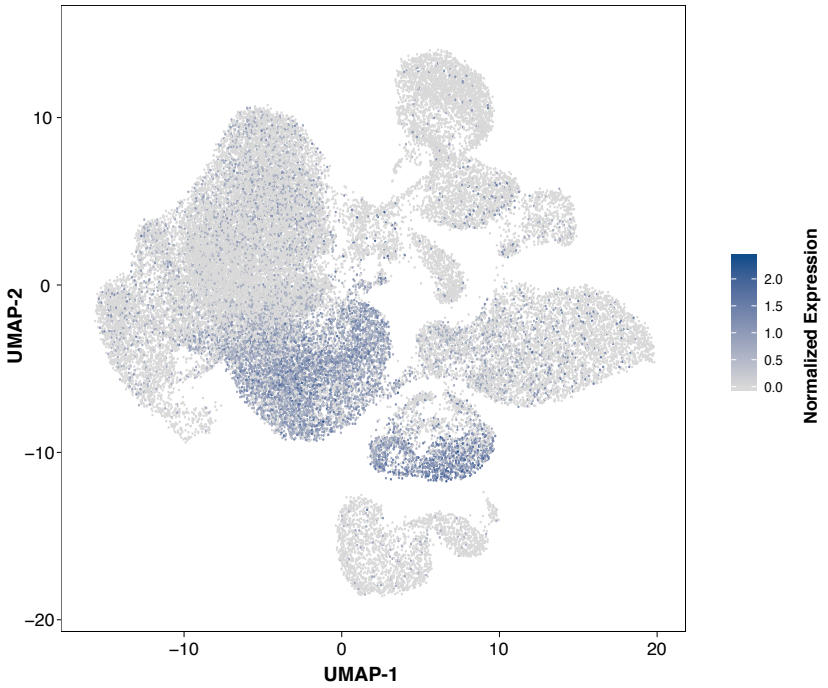


Supplemental Figure A-8: Diffusion Component 1- and 2-Based Visualization of Inferred Trajectories and Diffusion Maps.
Diffusion map of components 1 and 2 (DC1, DC2) for combined genotypes as well as separate genotypes. Lineages were inferred for each genotype separately using the R-package 'slingshot' and are depicted as lines. For representative purposes only 2,500 cells per genotype are depicted.

Supplemental Figure A-9: Enriched Gene Sets in Differentially Expressed Genes from scRNA-seq CMP and MoDCP Populations.
Differential expression was called between one specific genotype and all other genotypes within the same population. Genes with log2 fc > 0.25 and padj < 0.05 were considered as differentially expressed. For each genotype, overlaps with gene sets were calculated using the 'Molecular Signature Database' (MSigDB) with 'Hallmark', 'Gene Ontology' and 'Immunologic Signatures' gene sets as reference.



Supplemental Figure A-10: Transcriptomic Analysis of Isolated Ly6C⁺ GMPs.
(A) Heatmap depicting sample distances for each sample based on their normalized gene expression profile. Calculation of distances and clustering of samples was computed using Euclidean distances. **(B)** PCA of individual samples based on their normalized expression profile. **(C)** Volcano plot showing DE genes between indicated genotypes. Genes with a log2 fc ≤ -0.5 or ≥ 0.5 and a padj < 0.05 are highlighted and were considered as differentially expressed.



Supplemental Figure A-11: *Idh1* Expression in the Hematopoietic Compartment Based on scRNA-seq Data.
Normalized expression of *Idh1* highlighted within the UMAP-based representation of single-cell transcriptomic data. The color gradient from grey to dark blue indicates the level of expression (grey = no expression, dark blue = high expression).

A.2 Supplemental Tables

Supplemental Table A-1: Peripheral Blood and Bone Marrow Composition of Mice Used in scRNA-seq Experiments.

	CTRL	*Idh1*	*DNMT3A*	*Idh1-DNMT3A*
Peripheral Blood Counts (HemaVet 950 FS)				
White Blood Cells [K/µl]	13.12	11.82	9.36	12.08
Neutrophils [K/µl]	2.21	2.56	2.00	2.48
Lymphocytes [K/µl]	10.23	8.33	6.93	8.89
Monocytes [K/µl]	0.660	0.910	0.420	0.670
Red Blood Cells [M/µl]	9.790	9.660	9.670	10.03
Platelets [K/µl]	831	781	792	922
Peripheral Blood Cell Frequencies (FACS, Frequencies in % of CD45.1$^+$ cells)				
B-Cells	66.6	61.9	65.1	64.3
T-Cells	15.7	17.0	18.7	18.1
Myeloid Cells	15.2	18.3	14.2	15.5
Bone Marrow Cell Frequencies (FACS, Frequencies in % of EYFP$^+$ cells)				
LSK	0.11	0.086	0.18	0.15
HSC	0.028	0.00451	0.024	0.048
LT-HSC	0.00732	0.000449	0.00883	0.021
ST-HSC	0.014	0.00302	0.00839	0.013
MPP2	0.011	0.00631	0.011	0.012
MPP3+4	0.035	0.042	0.083	0.04
MPP5	0.00883	0.014	0.019	0.014
LS-K	1.89	1.82	2.11	2.15
CMP	0.6	0.61	0.65	0.77
GMP	0.49	0.38	0.51	0.45
MEP	0.39	0.41	0.52	0.52

Supplemental Table A-2: Hematopoietic Marker Genes Used for the Annotation of scRNA-seq Clusters.

Population	Marker Genes
LT-HSCs	*Mllt3, Ly6a, Procr, Mpl, Lmo2, Meis1, Hlf, Ifitm1, Ifitm3, Angpt1*
ST-HSCs	*Cd34, Mllt3, Ly6a, Procr, Mpl, Lmo2, Meis1, Hlf, Ifitm1, Ifitm3, Angpt1,*
Lymphoid primed multipotent progenitor	*Cd34, Flt3, Dntt, Il7r, Ikzf1, Ly6d*
Megakaryocyte progenitor	*Pf4, F2r, Vwf, Gp1bb, Gp5, Gp9*
Erythroid progenitor	*Car1, Car2, Gata1, Gata2, Klf1, Epor, Gfi1b*
Common myeloid progenitor	*Flt3, Spi1, Cebpa, Cebpe, Ms4a3, Ctsg, Mpo, Prtn3*
Monocyte / DC primed progenitor	*Irf8, Irf4, Csf1r, Ly6c2, Klf4, F13a1, Elane, Ctsg, Mpo, Prtn3, Ms4a3*
Neutrophil primed progenitor	*Elane, Ctsg, Mpo, Prtn3, Ms4a3, Cebpe, Gfi1*
Erythroid cells	*Hba-a1, Hbb-bh1, Hba-a2, Hbb-bt, Hbb-bs, Gypa, Car2, Epor, Tfrc*
Monocytes	*Ly6c2, S100a4, Csf1r, F13a1, Ms4a4c, Ms4a6c, Fn1, Ccr2, Clec4a3, Clec4a1*
Dendritic cells	*Cd209a, Siglech, Cd7, Cd74*
Neutrophils	*Lrg1, Ltf, Ngp, Lcn2, Mmp8, Mmp9, S100a8, S100a9, Fcnb, Cd63, Retnlg, Ly6g*
Basophils	*Prss34, Fcer1a, Ms4a2, Il6, Csf1*
T-cells	*Cd3d, Cd3g, Cd3e, Cd8a, Cd8b1*
Natural killer cells	*Gzma, Gzmb, Klra4, Klrk1, Klra7, Klre1, Klra8, Eomes, Ccl5, Tbx21*
Pro B-cells	*Dntt, Vpreb1, Igll1, Rag1, Rag2*
Pre B-cells	*Vpreb1, Igll1, Rag1, Rag2, Pcna, Tyms, Mki67, Top2a*
Immature B-cells	*Vpreb1, Igll1*
Mature B-cells	*Ms4a1, Cd74, H2-Aa, H2-Eb1, H2-Ab1, H2-DMb2, Iglc1, Iglc2, Iglc3*
Plasma cells	*Iglv1, Iglv3, Iglc1, Iglc2, Iglc3, Cd74*

Supplemental Table A-3: Varying Cell Type Frequencies per Genotype in Annotated scRNA-seq Cluster.
Frequencies are depicted as proportion of cells in indicated clusters in relation to the overall number of cells within genotype-specific LSK or LS-K samples.

	CTRL	*Idh1*	*DNMT3A*	*Idh1-DNMT3A*
LS-K Samples				
Common myeloid progenitor #1	0.046	0.060	0.074	0.088
Common myeloid progenitor #2	0.096	0.104	0.079	0.095
Common myeloid progenitor #3	0.063	0.051	0.057	0.019
Common myeloid progenitor #4	0.014	0.015	0.017	0.149
Common myeloid progenitor #5	0.032	0.034	0.041	0.145
Monocyte / DC primed progenitor #1	0.155	0.113	0.110	0.022
Monocyte / DC primed progenitor #1	0.067	0.073	0.035	0.009
Neutrophil primed progenitor	0.098	0.096	0.078	0.010
LSK samples				
LT-HSCs	0.204	0.068	0.214	0.194
Lymphoid primed multipotent progenitor #1	0.061	0.136	0.061	0.068
Lymphoid primed multipotent progenitor #2	0.040	0.020	0.030	0.032
Lymphoid primed multipotent progenitor #3	0.117	0.145	0.092	0.109

Supplemental Table A-4: Enriched Gene Sets in *Idh1*-R132H Mutated Ly6C$^+$ GMPs.

Hallmark	Normalized Enrichment Score	Nominal P-Value	False Discovery Rate
INTERFERON_ALPHA_RESPONSE	2.331	0.000	0.001
INTERFERON_GAMMA_RESPONSE	2.026	0.000	0.007
INFLAMMATORY_RESPONSE	1.378	0.054	0.180
IL6_JAK_STAT3_SIGNALING	1.173	0.225	0.348
ESTROGEN_RESPONSE_LATE	1.040	0.339	0.497
FATTY_ACID_METABOLISM	0.831	0.779	0.881
OXIDATIVE_PHOSPHORYLATION	0.657	0.989	0.970

Supplemental Table A-5: Enriched Gene Sets in *Idh1*-R132H *DNMT3A*-R882H Mutated Ly6C$^+$ GMPs.

Hallmark	Normalized Enrichment Score	Nominal P-Value	False Discovery Rate
INTERFERON_ALPHA_RESPONSE	1.995	0	0.009
INTERFERON_GAMMA_RESPONSE	1.861	0	0.021
MYC_TARGETS_V2	1.696	0.0154	0.043
MYC_TARGETS_V1	1.393	0.0235	0.211
ADIPOGENESIS	1.175	0.1561	0.543
OXIDATIVE_PHOSPHORYLATION	1.016	0.3805	0.942
UNFOLDED_PROTEIN_RESPONSE	0.991	0.4474	0.895
ANGIOGENESIS	0.982	0.4282	0.810
WNT_BETA_CATENIN_SIGNALING	0.872	0.6170	1
PEROXISOME	0.859	0.7123	1
FATTY_ACID_METABOLISM	0.835	0.7323	0.973
REACTIVE_OXIGEN_SPECIES_PATHWAY	0.637	0.9282	1
E2F_TARGETS	0.632	0.9969	1
DNA_REPAIR	0.580	1	0.992

References

Abelson, S., Collord, G., Ng, S. W. K., Weissbrod, O., Mendelson Cohen, N., Niemeyer, E., Barda, N., Zuzarte, P. C., Heisler, L., Sundaravadanam, Y., Luben, R., Hayat, S., Wang, T. T., Zhao, Z., Cirlan, I., Pugh, T. J., Soave, D., Ng, K., Latimer, C., Hardy, C., Raine, K., Jones, D., Hoult, D., Britten, A., McPherson, J. D., Johansson, M., Mbabaali, F., Eagles, J., Miller, J. K., Pasternack, D., Timms, L., Krzyzanowski, P., Awadalla, P., Costa, R., Segal, E., Bratman, S. V., Beer, P., Behjati, S., Martincorena, I., Wang, J. C. Y., Bowles, K. M., Quiros, J. R., Karakatsani, A., La Vecchia, C., Trichopoulou, A., Salamanca-Fernandez, E., Huerta, J. M., Barricarte, A., Travis, R. C., Tumino, R., Masala, G., Boeing, H., Panico, S., Kaaks, R., Kramer, A., Sieri, S., Riboli, E., Vineis, P., Foll, M., McKay, J., Polidoro, S., Sala, N., Khaw, K. T., Vermeulen, R., Campbell, P. J., Papaemmanuil, E., Minden, M. D., Tanay, A., Balicer, R. D., Wareham, N. J., Gerstung, M., Dick, J. E., Brennan, P., Vassiliou, G. S. & Shlush, L. I. 2018. Prediction of acute myeloid leukaemia risk in healthy individuals. *Nature,* 559**,** 400-404.

Adolfsson, J., Mansson, R., Buza-Vidas, N., Hultquist, A., Liuba, K., Jensen, C. T., Bryder, D., Yang, L., Borge, O. J., Thoren, L. A., Anderson, K., Sitnicka, E., Sasaki, Y., Sigvardsson, M. & Jacobsen, S. E. 2005. Identification of Flt3+ lympho-myeloid stem cells lacking erythro-megakaryocytic potential a revised road map for adult blood lineage commitment. *Cell,* 121**,** 295-306.

Akashi, K., Traver, D., Miyamoto, T. & Weissman, I. L. 2000. A clonogenic common myeloid progenitor that gives rise to all myeloid lineages. *Nature,* 404**,** 193-7.

An, J., Gonzalez-Avalos, E., Chawla, A., Jeong, M., Lopez-Moyado, I. F., Li, W., Goodell, M. A., Chavez, L., Ko, M. & Rao, A. 2015. Acute loss of TET function results in aggressive myeloid cancer in mice. *Nat Commun,* 6**,** 10071.

Andronesi, O. C., Kim, G. S., Gerstner, E., Batchelor, T., Tzika, A. A., Fantin, V. R., Vander Heiden, M. G. & Sorensen, A. G. 2012. Detection of 2-hydroxyglutarate in IDH-mutated glioma patients by in vivo spectral-editing and 2D correlation magnetic resonance spectroscopy. *Sci Transl Med,* 4**,** 116ra4.

Angerer, P., Haghverdi, L., Buttner, M., Theis, F. J., Marr, C. & Buettner, F. 2016. destiny: diffusion maps for large-scale single-cell data in R. *Bioinformatics,* 32**,** 1241-3.

Arinobu, Y., Mizuno, S., Chong, Y., Shigematsu, H., Iino, T., Iwasaki, H., Graf, T., Mayfield, R., Chan, S., Kastner, P. & Akashi, K. 2007. Reciprocal activation of GATA-1 and PU.1 marks initial specification of hematopoietic stem cells into myeloerythroid and myelolymphoid lineages. *Cell Stem Cell,* 1**,** 416-27.

Atlasi, Y. & Stunnenberg, H. G. 2017. The interplay of epigenetic marks during stem cell differentiation and development. *Nat Rev Genet,* 18**,** 643-658.

Auffray, C., Fogg, D., Garfa, M., Elain, G., Join-Lambert, O., Kayal, S., Sarnacki, S., Cumano, A., Lauvau, G. & Geissmann, F. 2007. Monitoring of blood vessels and tissues by a population of monocytes with patrolling behavior. *Science,* 317**,** 666-70.

Auffray, C., Fogg, D. K., Narni-Mancinelli, E., Senechal, B., Trouillet, C., Saederup, N., Leemput, J., Bigot, K., Campisi, L., Abitbol, M., Molina, T., Charo, I., Hume, D. A., Cumano, A., Lauvau, G. & Geissmann, F. 2009. CX3CR1+ CD115+ CD135+ common macrophage/DC precursors and the role of CX3CR1 in their response to inflammation. *J Exp Med,* 206**,** 595-606.

Bain, C. C., Bravo-Blas, A., Scott, C. L., Perdiguero, E. G., Geissmann, F., Henri, S., Malissen, B., Osborne, L. C., Artis, D. & Mowat, A. M. 2014. Constant replenishment from circulating monocytes maintains the macrophage pool in the intestine of adult mice. *Nat Immunol,* 15**,** 929-937.

Baldridge, M. T., King, K. Y., Boles, N. C., Weksberg, D. C. & Goodell, M. A. 2010. Quiescent haematopoietic stem cells are activated by IFN-gamma in response to chronic infection. *Nature,* 465**,** 793-7.

Balss, J., Pusch, S., Beck, A. C., Herold-Mende, C., Kramer, A., Thiede, C., Buckel, W., Langhans, C. D., Okun, J. G. & von Deimling, A. 2012. Enzymatic assay for quantitative analysis of (D)-2-hydroxyglutarate. *Acta Neuropathol,* 124**,** 883-91.

Barbalat, R., Lau, L., Locksley, R. M. & Barton, G. M. 2009. Toll-like receptor 2 on inflammatory monocytes induces type I interferon in response to viral but not bacterial ligands. *Nature Immunology,* 10**,** 1200-1207.

Baumann, P., Benson, F. E. & West, S. C. 1996. Human Rad51 protein promotes ATP-dependent homologous pairing and strand transfer reactions in vitro. *Cell,* 87**,** 757-66.

Becker, A. J., Mc, C. E. & Till, J. E. 1963. Cytological demonstration of the clonal nature of spleen colonies derived from transplanted mouse marrow cells. *Nature,* 197**,** 452-4.

Beerman, I., Bock, C., Garrison, B. S., Smith, Z. D., Gu, H., Meissner, A. & Rossi, D. J. 2013. Proliferation-dependent alterations of the DNA methylation landscape underlie hematopoietic stem cell aging. *Cell Stem Cell,* 12**,** 413-25.

Bhavya, B., Anand, C. R., Madhusoodanan, U. K., Rajalakshmi, P., Krishnakumar, K., Easwer, H. V., Deepti, A. N. & Gopala, S. 2019. To be Wild or Mutant: Role of Isocitrate Dehydrogenase 1 (IDH1) and 2-Hydroxy Glutarate (2-HG) in Gliomagenesis and Treatment Outcome in Glioma. *Cell Mol Neurobiol.*

Blackford, A. N. & Jackson, S. P. 2017. ATM, ATR, and DNA-PK: The Trinity at the Heart of the DNA Damage Response. *Mol Cell,* 66**,** 801-817.

Bleeker, F. E., Atai, N. A., Lamba, S., Jonker, A., Rijkeboer, D., Bosch, K. S., Tigchelaar, W., Troost, D., Vandertop, W. P., Bardelli, A. & Van Noorden, C. J. 2010. The prognostic IDH1(R132) mutation is associated with reduced NADP+-dependent IDH activity in glioblastoma. *Acta Neuropathol,* 119**,** 487-94.

Bock, C., Beerman, I., Lien, W. H., Smith, Z. D., Gu, H., Boyle, P., Gnirke, A., Fuchs, E., Rossi, D. J. & Meissner, A. 2012. DNA methylation dynamics during in vivo differentiation of blood and skin stem cells. *Mol Cell,* 47**,** 633-47.

Borodovsky, A., Salmasi, V., Turcan, S., Fabius, A. W., Baia, G. S., Eberhart, C. G., Weingart, J. D., Gallia, G. L., Baylin, S. B., Chan, T. A. & Riggins, G. J. 2013. 5-azacytidine reduces methylation, promotes differentiation and induces tumor regression in a patient-derived IDH1 mutant glioma xenograft. *Oncotarget,* 4**,** 1737-47.

Bowman, R. L., Busque, L. & Levine, R. L. 2018. Clonal Hematopoiesis and Evolution to Hematopoietic Malignancies. *Cell Stem Cell,* 22**,** 157-170.

Bralten, L. B., Kloosterhof, N. K., Balvers, R., Sacchetti, A., Lapre, L., Lamfers, M., Leenstra, S., de Jonge, H., Kros, J. M., Jansen, E. E., Struys, E. A., Jakobs, C., Salomons, G. S., Diks, S. H., Peppelenbosch, M., Kremer, A., Hoogenraad, C. C., Smitt, P. A. & French, P. J. 2011. IDH1 R132H decreases proliferation of glioma cell lines in vitro and in vivo. *Ann Neurol,* 69**,** 455-63.

Briseno, C. G., Haldar, M., Kretzer, N. M., Wu, X., Theisen, D. J., Kc, W., Durai, V., Grajales-Reyes, G. E., Iwata, A., Bagadia, P., Murphy, T. L. & Murphy, K. M. 2016. Distinct Transcriptional Programs Control Cross-Priming in Classical and Monocyte-Derived Dendritic Cells. *Cell Rep,* 15**,** 2462-74.

Brocks, D., Assenov, Y., Minner, S., Bogatyrova, O., Simon, R., Koop, C., Oakes, C., Zucknick, M., Lipka, D. B., Weischenfeldt, J., Feuerbach, L., Cowper-Sal Lari, R., Lupien, M., Brors, B., Korbel, J., Schlomm, T., Tanay, A., Sauter, G., Gerhauser, C., Plass, C. & Project, I. E. O. P. C. 2014. Intratumor DNA methylation heterogeneity reflects clonal evolution in aggressive prostate cancer. *Cell Rep,* 8**,** 798-806.

Brocks, D., Schmidt, C. R., Daskalakis, M., Jang, H. S., Shah, N. M., Li, D., Li, J., Zhang, B., Hou, Y., Laudato, S., Lipka, D. B., Schott, J., Bierhoff, H., Assenov, Y., Helf, M., Ressnerova, A., Islam, M. S., Lindroth, A. M., Haas, S., Essers, M., Imbusch, C. D., Brors, B., Oehme, I., Witt, O., Lubbert, M., Mallm, J. P., Rippe, K., Will, R., Weichenhan, D., Stoecklin, G., Gerhauser, C., Oakes, C. C., Wang, T. & Plass, C. 2017. DNMT and HDAC inhibitors induce cryptic transcription start sites encoded in long terminal repeats. *Nat Genet,* 49**,** 1052-1060.

Broske, A. M., Vockentanz, L., Kharazi, S., Huska, M. R., Mancini, E., Scheller, M., Kuhl, C., Enns, A., Prinz, M., Jaenisch, R., Nerlov, C., Leutz, A., Andrade-Navarro, M. A., Jacobsen, S. E. & Rosenbauer, F. 2009. DNA methylation protects hematopoietic stem cell multipotency from myeloerythroid restriction. *Nat Genet,* 41**,** 1207-15.

Bryant, H. E., Schultz, N., Thomas, H. D., Parker, K. M., Flower, D., Lopez, E., Kyle, S., Meuth, M., Curtin, N. J. & Helleday, T. 2005. Specific killing of BRCA2-deficient tumours with inhibitors of poly(ADP-ribose) polymerase. *Nature,* 434**,** 913-7.

Buenrostro, J. D., Corces, M. R., Lareau, C. A., Wu, B., Schep, A. N., Aryee, M. J., Majeti, R., Chang, H. Y. & Greenleaf, W. J. 2018. Integrated Single-Cell Analysis Maps the Continuous Regulatory Landscape of Human Hematopoietic Differentiation. *Cell,* 173**,** 1535-1548 e16.

Busch, K., Klapproth, K., Barile, M., Flossdorf, M., Holland-Letz, T., Schlenner, S. M., Reth, M., Hofer, T. & Rodewald, H. R. 2015. Fundamental properties of unperturbed haematopoiesis from stem cells in vivo. *Nature,* 518**,** 542-6.

Busque, L., Mio, R., Mattioli, J., Brais, E., Blais, N., Lalonde, Y., Maragh, M. & Gilliland, D. G. 1996. Nonrandom X-inactivation patterns in normal females: lyonization ratios vary with age. *Blood,* 88**,** 59-65.

Butler, A., Hoffman, P., Smibert, P., Papalexi, E. & Satija, R. 2018. Integrating single-cell transcriptomic data across different conditions, technologies, and species. *Nat Biotechnol,* 36**,** 411-420.

Cabezas-Wallscheid, N., Klimmeck, D., Hansson, J., Lipka, D. B., Reyes, A., Wang, Q., Weichenhan, D., Lier, A., von Paleske, L., Renders, S., Wunsche, P., Zeisberger, P., Brocks, D., Gu, L., Herrmann, C., Haas, S., Essers, M. A., Brors, B., Eils, R., Huber, W., Milsom, M. D., Plass, C., Krijgsveld, J. & Trumpp, A. 2014. Identification of regulatory networks in HSCs and their immediate progeny via integrated proteome, transcriptome, and DNA methylome analysis. *Cell Stem Cell,* 15**,** 507-22.

Cai, D. H., Wang, D., Keefer, J., Yeamans, C., Hensley, K. & Friedman, A. D. 2008. C/EBP alpha:AP-1 leucine zipper heterodimers bind novel DNA elements, activate the PU.1 promoter and direct monocyte lineage commitment more potently than C/EBP alpha homodimers or AP-1. *Oncogene,* 27**,** 2772-9.

Cairns, R. A. & Mak, T. W. 2013. Oncogenic isocitrate dehydrogenase mutations: mechanisms, models, and clinical opportunities. *Cancer Discov,* 3**,** 730-41.

Cancer Genome Atlas Research, N., Ley, T. J., Miller, C., Ding, L., Raphael, B. J., Mungall, A. J., Robertson, A., Hoadley, K., Triche, T. J., Jr., Laird, P. W., Baty, J. D., Fulton, L. L., Fulton, R., Heath, S. E., Kalicki-Veizer, J., Kandoth, C., Klco, J. M., Koboldt, D. C., Kanchi, K. L., Kulkarni, S., Lamprecht, T. L., Larson, D. E., Lin, L., Lu, C., McLellan, M. D., McMichael, J. F., Payton, J., Schmidt, H., Spencer, D. H., Tomasson, M. H., Wallis, J. W., Wartman, L. D., Watson, M. A., Welch, J., Wendl, M. C., Ally, A., Balasundaram, M., Birol, I., Butterfield, Y., Chiu, R., Chu, A., Chuah, E., Chun, H. J., Corbett, R., Dhalla, N., Guin, R., He, A., Hirst, C., Hirst, M., Holt, R. A., Jones, S., Karsan, A., Lee, D., Li, H. I., Marra, M. A., Mayo, M., Moore, R. A., Mungall, K., Parker, J., Pleasance, E., Plettner, P., Schein, J., Stoll, D., Swanson, L., Tam, A., Thiessen, N., Varhol, R., Wye, N., Zhao, Y., Gabriel, S., Getz, G., Sougnez, C., Zou, L., Leiserson, M. D., Vandin, F., Wu, H. T., Applebaum, F., Baylin, S. B., Akbani, R., Broom, B. M., Chen, K., Motter, T. C., Nguyen, K., Weinstein, J. N., Zhang, N., Ferguson, M. L., Adams, C., Black, A., Bowen, J., Gastier-Foster, J., Grossman, T., Lichtenberg, T., Wise, L., Davidsen, T., Demchok, J. A., Shaw, K. R., Sheth, M., Sofia, H. J., Yang, L., Downing, J. R., et al. 2013. Genomic and epigenomic landscapes of adult de novo acute myeloid leukemia. *N Engl J Med,* 368**,** 2059-74.

Carlin, L. M., Stamatiades, E. G., Auffray, C., Hanna, R. N., Glover, L., Vizcay-Barrena, G., Hedrick, C. C., Cook, H. T., Diebold, S. & Geissmann, F. 2013. Nr4a1-dependent Ly6C(low) monocytes monitor endothelial cells and orchestrate their disposal. *Cell,* 153**,** 362-75.

Catlin, S. N., Busque, L., Gale, R. E., Guttorp, P. & Abkowitz, J. L. 2011. The replication rate of human hematopoietic stem cells in vivo. *Blood,* 117**,** 4460-6.

Cecchini, M. G., Dominguez, M. G., Mocci, S., Wetterwald, A., Felix, R., Fleisch, H., Chisholm, O., Hofstetter, W., Pollard, J. W. & Stanley, E. R. 1994. Role of colony stimulating factor-1 in the establishment and regulation of tissue macrophages during postnatal development of the mouse. *Development,* 120**,** 1357-72.

Celik, H., Kramer, A. & Challen, G. A. 2016. DNA methylation in normal and malignant hematopoiesis. *Int J Hematol,* 103**,** 617-26.

Celik, H., Mallaney, C., Kothari, A., Ostrander, E. L., Eultgen, E., Martens, A., Miller, C. A., Hundal, J., Klco, J. M. & Challen, G. A. 2015. Enforced differentiation of Dnmt3a-null bone marrow leads to failure with c-Kit mutations driving leukemic transformation. *Blood,* 125**,** 619-28.

Challen, G. A., Sun, D., Jeong, M., Luo, M., Jelinek, J., Berg, J. S., Bock, C., Vasanthakumar, A., Gu, H., Xi, Y., Liang, S., Lu, Y., Darlington, G. J., Meissner, A., Issa, J. P., Godley, L. A., Li, W. & Goodell, M. A. 2011. Dnmt3a is essential for hematopoietic stem cell differentiation. *Nat Genet,* 44**,** 23-31.

Challen, G. A., Sun, D., Mayle, A., Jeong, M., Luo, M., Rodriguez, B., Mallaney, C., Celik, H., Yang, L., Xia, Z., Cullen, S., Berg, J., Zheng, Y., Darlington, G. J., Li, W. & Goodell, M. A. 2014. Dnmt3a and Dnmt3b have overlapping and distinct functions in hematopoietic stem cells. *Cell Stem Cell,* 15**,** 350-64.

Chatterjee, N. & Walker, G. C. 2017. Mechanisms of DNA damage, repair, and mutagenesis. *Environ Mol Mutagen,* 58**,** 235-263.

Chaturvedi, A., Araujo Cruz, M. M., Jyotsana, N., Sharma, A., Yun, H., Gorlich, K., Wichmann, M., Schwarzer, A., Preller, M., Thol, F., Meyer, J., Haemmerle, R., Struys, E. A., Jansen, E. E., Modlich, U., Li, Z., Sly, L. M., Geffers, R., Lindner, R., Manstein, D. J., Lehmann, U., Krauter, J., Ganser, A. & Heuser, M. 2013. Mutant IDH1 promotes leukemogenesis in vivo and can be specifically targeted in human AML. *Blood,* 122**,** 2877-87.

Chen, C., Liu, Y., Lu, C., Cross, J. R., Morris, J. P. t., Shroff, A. S., Ward, P. S., Bradner, J. E., Thompson, C. & Lowe, S. W. 2013a. Cancer-associated IDH2 mutants drive an acute myeloid leukemia that is susceptible to Brd4 inhibition. *Genes Dev,* 27**,** 1974-85.

Chen, S., Yang, J., Wei, Y. & Wei, X. 2019. Epigenetic regulation of macrophages: from homeostasis maintenance to host defense. *Cell Mol Immunol.*

Chen, S. J., Shen, Y. & Chen, Z. 2013b. A panoramic view of acute myeloid leukemia. *Nat Genet,* 45**,** 586-7.

Chen, Z., Yang, H. & Pavletich, N. P. 2008. Mechanism of homologous recombination from the RecA-ssDNA/dsDNA structures. *Nature,* 453**,** 489-4.

Cheong, C., Matos, I., Choi, J. H., Dandamudi, D. B., Shrestha, E., Longhi, M. P., Jeffrey, K. L., Anthony, R. M., Kluger, C., Nchinda, G., Koh, H., Rodriguez, A., Idoyaga, J., Pack, M., Velinzon, K., Park, C. G. & Steinman, R. M. 2010. Microbial stimulation fully differentiates monocytes to DC-SIGN/CD209(+) dendritic cells for immune T cell areas. *Cell,* 143**,** 416-29.

Cheshier, S. H., Prohaska, S. S. & Weissman, I. L. 2007. The effect of bleeding on hematopoietic stem cell cycling and self-renewal. *Stem Cells Dev,* 16**,** 707-17.

Chiappinelli, K. B., Strissel, P. L., Desrichard, A., Li, H., Henke, C., Akman, B., Hein, A., Rote, N. S., Cope, L. M., Snyder, A., Makarov, V., Budhu, S., Slamon, D. J., Wolchok, J. D., Pardoll, D. M., Beckmann, M. W., Zahnow, C. A., Merghoub, T., Chan, T. A., Baylin, S. B. & Strick, R. 2015. Inhibiting DNA Methylation Causes an Interferon Response in Cancer via dsRNA Including Endogenous Retroviruses. *Cell,* 162**,** 974-86.

Choi, C., Ganji, S. K., DeBerardinis, R. J., Hatanpaa, K. J., Rakheja, D., Kovacs, Z., Yang, X. L., Mashimo, T., Raisanen, J. M., Marin-Valencia, I., Pascual, J. M., Madden, C. J., Mickey, B. E., Malloy, C. R., Bachoo, R. M. & Maher, E. A. 2012. 2-hydroxyglutarate detection by magnetic resonance spectroscopy in IDH-mutated patients with gliomas. *Nat Med,* 18**,** 624-9.

Chowdhury, R., Yeoh, K. K., Tian, Y. M., Hillringhaus, L., Bagg, E. A., Rose, N. R., Leung, I. K., Li, X. S., Woon, E. C., Yang, M., McDonough, M. A., King, O. N., Clifton, I. J., Klose, R. J., Claridge, T. D., Ratcliffe, P. J., Schofield, C. J. & Kawamura, A. 2011. The oncometabolite 2-hydroxyglutarate inhibits histone lysine demethylases. *EMBO Rep,* 12**,** 463-9.

Christensen, J. L. & Weissman, I. L. 2001. Flk-2 is a marker in hematopoietic stem cell differentiation: a simple method to isolate long-term stem cells. *Proc Natl Acad Sci U S A,* 98**,** 14541-6.

Ciccia, A. & Elledge, S. J. 2010. The DNA damage response: making it safe to play with knives. *Mol Cell,* 40**,** 179-204.

Coombs, C. C., Zehir, A., Devlin, S. M., Kishtagari, A., Syed, A., Jonsson, P., Hyman, D. M., Solit, D. B., Robson, M. E., Baselga, J., Arcila, M. E., Ladanyi, M., Tallman, M. S., Levine, R. L. & Berger, M. F. 2017. Therapy-Related Clonal Hematopoiesis in Patients with Non-hematologic Cancers Is Common and Associated with Adverse Clinical Outcomes. *Cell Stem Cell,* 21**,** 374-382 e4.

Corces, M. R., Buenrostro, J. D., Wu, B., Greenside, P. G., Chan, S. M., Koenig, J. L., Snyder, M. P., Pritchard, J. K., Kundaje, A., Greenleaf, W. J., Majeti, R. & Chang, H. Y. 2016. Lineage-specific and single-cell chromatin accessibility charts human hematopoiesis and leukemia evolution. *Nat Genet,* 48**,** 1193-203.

Corces-Zimmerman, M. R., Hong, W. J., Weissman, I. L., Medeiros, B. C. & Majeti, R. 2014. Preleukemic mutations in human acute myeloid leukemia affect epigenetic regulators and persist in remission. *Proc Natl Acad Sci U S A,* 111**,** 2548-53.

Craver, B. M., El Alaoui, K., Scherber, R. M. & Fleischman, A. G. 2018. The Critical Role of Inflammation in the Pathogenesis and Progression of Myeloid Malignancies. *Cancers (Basel),* 10.

Creyghton, M. P., Cheng, A. W., Welstead, G. G., Kooistra, T., Carey, B. W., Steine, E. J., Hanna, J., Lodato, M. A., Frampton, G. M., Sharp, P. A., Boyer, L. A., Young, R. A. & Jaenisch, R. 2010. Histone H3K27ac separates active from poised enhancers and predicts developmental state. *Proc Natl Acad Sci U S A,* 107**,** 21931-6.

Dai, X. M., Ryan, G. R., Hapel, A. J., Dominguez, M. G., Russell, R. G., Kapp, S., Sylvestre, V. & Stanley, E. R. 2002. Targeted disruption of the mouse colony-stimulating factor 1 receptor gene results in osteopetrosis, mononuclear phagocyte deficiency, increased primitive progenitor cell frequencies, and reproductive defects. *Blood,* 99**,** 111-20.

Dang, C. V. 2012. MYC on the path to cancer. *Cell,* 149**,** 22-35.

Dang, L., White, D. W., Gross, S., Bennett, B. D., Bittinger, M. A., Driggers, E. M., Fantin, V. R., Jang, H. G., Jin, S., Keenan, M. C., Marks, K. M., Prins, R. M., Ward, P. S., Yen, K. E., Liau, L. M., Rabinowitz, J. D., Cantley, L. C., Thompson, C. B., Vander Heiden, M. G. & Su, S. M. 2009. Cancer-associated IDH1 mutations produce 2-hydroxyglutarate. *Nature,* 462**,** 739-44.

Das, H., Kumar, A., Lin, Z., Patino, W. D., Hwang, P. M., Feinberg, M. W., Majumder, P. K. & Jain, M. K. 2006. Kruppel-like factor 2 (KLF2) regulates proinflammatory activation of monocytes. *Proc Natl Acad Sci U S A,* 103**,** 6653-8.

Daskalakis, M., Brocks, D., Sheng, Y. H., Islam, M. S., Ressnerova, A., Assenov, Y., Milde, T., Oehme, I., Witt, O., Goyal, A., Kuhn, A., Hartmann, M., Weichenhan, D., Jung, M. & Plass, C. 2018. Reactivation of endogenous retroviral elements via treatment with DNMT- and HDAC-inhibitors. *Cell Cycle,* 17**,** 811-822.

Date, D., Das, R., Narla, G., Simon, D. I., Jain, M. K. & Mahabeleshwar, G. H. 2014. Kruppel-like transcription factor 6 regulates inflammatory macrophage polarization. *J Biol Chem,* 289**,** 10318-29.

Delia, D. & Mizutani, S. 2017. The DNA damage response pathway in normal hematopoiesis and malignancies. *Int J Hematol,* 106**,** 328-334.

DeLuca, D. S., Levin, J. Z., Sivachenko, A., Fennell, T., Nazaire, M. D., Williams, C., Reich, M., Winckler, W. & Getz, G. 2012. RNA-SeQC: RNA-seq metrics for quality control and process optimization. *Bioinformatics,* 28**,** 1530-2.

Desai, P., Mencia-Trinchant, N., Savenkov, O., Simon, M. S., Cheang, G., Lee, S., Samuel, M., Ritchie, E. K., Guzman, M. L., Ballman, K. V., Roboz, G. J. & Hassane, D. C. 2018. Somatic mutations precede acute myeloid leukemia years before diagnosis. *Nat Med,* 24**,** 1015-1023.

DiNardo, C. D., Propert, K. J., Loren, A. W., Paietta, E., Sun, Z., Levine, R. L., Straley, K. S., Yen, K., Patel, J. P., Agresta, S., Abdel-Wahab, O., Perl, A. E., Litzow, M. R., Rowe, J. M., Lazarus, H. M., Fernandez, H. F., Margolis, D. J., Tallman, M. S., Luger, S. M. & Carroll, M. 2013. Serum 2-hydroxyglutarate levels predict isocitrate dehydrogenase mutations and clinical outcome in acute myeloid leukemia. *Blood,* 121**,** 4917-24.

Dixit, A., Parnas, O., Li, B., Chen, J., Fulco, C. P., Jerby-Arnon, L., Marjanovic, N. D., Dionne, D., Burks, T., Raychowdhury, R., Adamson, B., Norman, T. M., Lander, E. S., Weissman, J. S., Friedman, N. & Regev, A. 2016. Perturb-Seq: Dissecting Molecular Circuits with Scalable Single-Cell RNA Profiling of Pooled Genetic Screens. *Cell,* 167**,** 1853-1866.e17.

Dobin, A., Davis, C. A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M. & Gingeras, T. R. 2013. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics,* 29**,** 15-21.

Doulatov, S., Notta, F., Eppert, K., Nguyen, L. T., Ohashi, P. S. & Dick, J. E. 2010. Revised map of the human progenitor hierarchy shows the origin of macrophages and dendritic cells in early lymphoid development. *Nat Immunol,* 11**,** 585-93.

Doulatov, S., Notta, F., Laurenti, E. & Dick, J. E. 2012. Hematopoiesis: a human perspective. *Cell Stem Cell,* 10**,** 120-36.

Dykstra, B., Kent, D., Bowie, M., McCaffrey, L., Hamilton, M., Lyons, K., Lee, S. J., Brinkman, R. & Eaves, C. 2007. Long-term propagation of distinct hematopoietic differentiation programs in vivo. *Cell Stem Cell,* 1**,** 218-29.

Essers, M. A., Offner, S., Blanco-Bose, W. E., Waibler, Z., Kalinke, U., Duchosal, M. A. & Trumpp, A. 2009. IFNalpha activates dormant haematopoietic stem cells in vivo. *Nature,* 458**,** 904-8.

Etzrodt, M., Ahmed, N., Hoppe, P. S., Loeffler, D., Skylaki, S., Hilsenbeck, O., Kokkaliaris, K. D., Kaltenbach, H. M., Stelling, J., Nerlov, C. & Schroeder, T. 2019. Inflammatory signals directly instruct PU.1 in HSCs via TNF. *Blood,* 133**,** 816-819.

Farlik, M., Halbritter, F., Muller, F., Choudry, F. A., Ebert, P., Klughammer, J., Farrow, S., Santoro, A., Ciaurro, V., Mathur, A., Uppal, R., Stunnenberg, H. G., Ouwehand, W. H., Laurenti, E., Lengauer, T., Frontini, M. & Bock, C. 2016. DNA Methylation Dynamics of Human Hematopoietic Stem Cell Differentiation. *Cell Stem Cell,* 19**,** 808-822.

Farmer, H., McCabe, N., Lord, C. J., Tutt, A. N., Johnson, D. A., Richardson, T. B., Santarosa, M., Dillon, K. J., Hickson, I., Knights, C., Martin, N. M., Jackson, S. P., Smith, G. C. & Ashworth, A. 2005. Targeting the DNA repair defect in BRCA mutant cells as a therapeutic strategy. *Nature,* 434**,** 917-21.

Feinberg, A. P., Koldobskiy, M. A. & Gondor, A. 2016. Epigenetic modulators, modifiers and mediators in cancer aetiology and progression. *Nat Rev Genet,* 17**,** 284-99.

Fey, M. F., Liechti-Gallati, S., von Rohr, A., Borisch, B., Theilkas, L., Schneider, V., Oestreicher, M., Nagel, S., Ziemiecki, A. & Tobler, A. 1994. Clonality and X-inactivation patterns in hematopoietic cell populations detected by the highly informative M27 beta DNA probe. *Blood,* 83**,** 931-8.

Fialkow, P. J., Gartler, S. M. & Yoshida, A. 1967. Clonal origin of chronic myelocytic leukemia in man. *Proc Natl Acad Sci U S A,* 58**,** 1468-71.

Figueroa, M. E., Abdel-Wahab, O., Lu, C., Ward, P. S., Patel, J., Shih, A., Li, Y., Bhagwat, N., Vasanthakumar, A., Fernandez, H. F., Tallman, M. S., Sun, Z., Wolniak, K., Peeters, J. K., Liu, W., Choe, S. E., Fantin, V. R., Paietta, E., Lowenberg, B., Licht, J. D., Godley, L. A., Delwel, R., Valk, P. J., Thompson, C. B., Levine, R. L. & Melnick, A. 2010. Leukemic IDH1 and IDH2 mutations result in a hypermethylation phenotype, disrupt TET2 function, and impair hematopoietic differentiation. *Cancer Cell,* 18**,** 553-67.

Flach, J., Bakker, S. T., Mohrin, M., Conroy, P. C., Pietras, E. M., Reynaud, D., Alvarez, S., Diolaiti, M. E., Ugarte, F., Forsberg, E. C., Le Beau, M. M., Stohr, B. A., Mendez, J., Morrison, C. G. & Passegue, E. 2014. Replication stress is a potent driver of functional decline in ageing haematopoietic stem cells. *Nature,* 512**,** 198-202.

Flavahan, W. A., Drier, Y., Liau, B. B., Gillespie, S. M., Venteicher, A. S., Stemmer-Rachamimov, A. O., Suva, M. L. & Bernstein, B. E. 2016. Insulator dysfunction and oncogene activation in IDH mutant gliomas. *Nature,* 529**,** 110-4.

Flavahan, W. A., Gaskell, E. & Bernstein, B. E. 2017. Epigenetic plasticity and the hallmarks of cancer. *Science,* 357.

Fogg, D. K., Sibon, C., Miled, C., Jung, S., Aucouturier, P., Littman, D. R., Cumano, A. & Geissmann, F. 2006. A clonogenic bone marrow progenitor specific for macrophages and dendritic cells. *Science,* 311**,** 83-7.

Fontana, M. F., Baccarella, A., Pancholi, N., Pufall, M. A., Herbert, D. R. & Kim, C. C. 2015. JUNB is a key transcriptional modulator of macrophage activation. *J Immunol,* 194**,** 177-86.

Friedman, A. D. 2007a. C/EBPalpha induces PU.1 and interacts with AP-1 and NF-kappaB to regulate myeloid development. *Blood Cells Mol Dis,* 39**,** 340-3.

Friedman, A. D. 2007b. Transcriptional control of granulocyte and monocyte development. *Oncogene,* 26**,** 6816-28.

Fuster, J. J., MacLauchlan, S., Zuriaga, M. A., Polackal, M. N., Ostriker, A. C., Chakraborty, R., Wu, C. L., Sano, S., Muralidharan, S., Rius, C., Vuong, J., Jacob, S., Muralidhar, V., Robertson, A. A., Cooper, M. A., Andres, V., Hirschi, K. K., Martin, K. A. & Walsh, K. 2017. Clonal hematopoiesis associated with TET2 deficiency accelerates atherosclerosis development in mice. *Science,* 355**,** 842-847.

Gagne, M., Boulay, K., Topisirovic, I., Huot, M. E. & Mallette, F. A. 2017. Oncogenic Activities of IDH1/2 Mutations: From Epigenetics to Cellular Signaling. *Trends Cell Biol,* 27**,** 738-752.

Geissmann, F., Jung, S. & Littman, D. R. 2003. Blood monocytes consist of two principal subsets with distinct migratory properties. *Immunity,* 19**,** 71-82.

Geissmann, F., Manz, M. G., Jung, S., Sieweke, M. H., Merad, M. & Ley, K. 2010. Development of monocytes, macrophages, and dendritic cells. *Science,* 327**,** 656-61.

Genovese, G., Kahler, A. K., Handsaker, R. E., Lindberg, J., Rose, S. A., Bakhoum, S. F., Chambert, K., Mick, E., Neale, B. M., Fromer, M., Purcell, S. M., Svantesson, O., Landen, M., Hoglund, M., Lehmann, S., Gabriel, S. B., Moran, J. L., Lander, E. S., Sullivan, P. F., Sklar, P., Gronberg, H., Hultman, C. M. & McCarroll, S. A. 2014. Clonal hematopoiesis and blood-cancer risk inferred from blood DNA sequence. *N Engl J Med,* 371**,** 2477-87.

Giladi, A., Paul, F., Herzog, Y., Lubling, Y., Weiner, A., Yofe, I., Jaitin, D., Cabezas-Wallscheid, N., Dress, R., Ginhoux, F., Trumpp, A., Tanay, A. & Amit, I. 2018. Single-cell characterization of haematopoietic progenitors and their trajectories in homeostasis and perturbed haematopoiesis. *Nat Cell Biol,* 20**,** 836-846.

Ginhoux, F. & Guilliams, M. 2016. Tissue-Resident Macrophage Ontogeny and Homeostasis. *Immunity,* 44**,** 439-449.

Glass, J., Hassane, D. C., Wouters, B., Kunimoto, H., Avellino, R., Garrett-Bakelman, F. E., Guryanova, O. A., Bowman, R., Redlich, S., Intlekofer, A., Meydan, C., Qin, T., Fall, M. P., Alonso, A., Guzman, M. L., Valk, P. J., Thompson, C. B., Levine, R. L., Elemento, O., Delwel, R., Melnick, A. & Figueroa, M. E. 2017. Epigenetic Identity in AML Depends on Disruption of Non-promoter Regulatory Elements and is Affected by Antagonistic Effects of Mutations in Epigenetic Modifiers. *Cancer Discov*.

Golub, D., Iyengar, N., Dogra, S., Wong, T., Bready, D., Tang, K., Modrek, A. S. & Placantonakis, D. G. 2019. Mutant Isocitrate Dehydrogenase Inhibitors as Targeted Cancer Therapeutics. *Front Oncol,* 9**,** 417.

Gothert, J. R., Gustin, S. E., Hall, M. A., Green, A. R., Gottgens, B., Izon, D. J. & Begley, C. G. 2005. In vivo fate-tracing studies using the Scl stem cell enhancer: embryonic hematopoietic stem cells significantly contribute to adult hematopoiesis. *Blood,* 105**,** 2724-32.

Gottgens, B., Nastos, A., Kinston, S., Piltz, S., Delabesse, E. C., Stanley, M., Sanchez, M. J., Ciau-Uitz, A., Patient, R. & Green, A. R. 2002. Establishing the transcriptional programme for blood: the SCL stem cell enhancer is regulated by a multiprotein complex containing Ets and GATA factors. *Embo j,* 21**,** 3039-50.

Graf, T. & Enver, T. 2009. Forcing cells to change lineages. *Nature,* 462**,** 587-94.

Gross, S., Cairns, R. A., Minden, M. D., Driggers, E. M., Bittinger, M. A., Jang, H. G., Sasaki, M., Jin, S., Schenkein, D. P., Su, S. M., Dang, L., Fantin, V. R. & Mak, T. W. 2010. Cancer-associated metabolite 2-hydroxyglutarate accumulates in acute myelogenous leukemia with isocitrate dehydrogenase 1 and 2 mutations. *J Exp Med,* 207**,** 339-44.

Guilliams, M., Ginhoux, F., Jakubzick, C., Naik, S. H., Onai, N., Schraml, B. U., Segura, E., Tussiwand, R. & Yona, S. 2014. Dendritic cells, monocytes and macrophages: a unified nomenclature based on ontogeny. *Nat Rev Immunol,* 14**,** 571-8.

Guilliams, M., Mildner, A. & Yona, S. 2018. Developmental and Functional Heterogeneity of Monocytes. *Immunity,* 49**,** 595-613.

Guo, G., Luc, S., Marco, E., Lin, T. W., Peng, C., Kerenyi, M. A., Beyaz, S., Kim, W., Xu, J., Das, P. P., Neff, T., Zou, K., Yuan, G. C. & Orkin, S. H. 2013. Mapping cellular hierarchy by single-cell analysis of the cell surface repertoire. *Cell Stem Cell,* 13**,** 492-505.

Guryanova, O. A., Shank, K., Spitzer, B., Luciani, L., Koche, R. P., Garrett-Bakelman, F. E., Ganzel, C., Durham, B. H., Mohanty, A., Hoermann, G., Rivera, S. A., Chramiec, A. G., Pronier, E., Bastian, L., Keller, M. D., Tovbin, D., Loizou, E., Weinstein, A. R., Gonzalez, A. R., Lieu, Y. K., Rowe, J. M., Pastore, F., McKenney, A. S., Krivtsov, A. V., Sperr, W. R., Cross, J. R., Mason, C. E., Tallman, M. S., Arcila, M. E., Abdel-Wahab, O., Armstrong, S. A., Kubicek, S., Staber, P. B., Gonen, M., Paietta, E. M., Melnick, A. M., Nimer, S. D., Mukherjee, S. & Levine, R. L. 2016. DNMT3A mutations promote anthracycline resistance in acute myeloid leukemia via impaired nucleosome remodeling. *Nat Med,* 22**,** 1488-1495.

Haas, S., Hansson, J., Klimmeck, D., Loeffler, D., Velten, L., Uckelmann, H., Wurzer, S., Prendergast, A. M., Schnell, A., Hexel, K., Santarella-Mellwig, R., Blaszkiewicz, S., Kuck, A., Geiger, H., Milsom, M. D., Steinmetz, L. M., Schroeder, T., Trumpp, A., Krijgsveld, J. & Essers, M. A. 2015. Inflammation-Induced Emergency Megakaryopoiesis Driven by Hematopoietic Stem Cell-like Megakaryocyte Progenitors. *Cell Stem Cell,* 17**,** 422-34.

Haas, S., Trumpp, A. & Milsom, M. D. 2018. Causes and Consequences of Hematopoietic Stem Cell Heterogeneity. *Cell Stem Cell,* 22**,** 627-638.

Harrison, D. E., Astle, C. M. & Delaittre, J. A. 1978. Loss of proliferative capacity in immunohemopoietic stem cells caused by serial transplantation rather than aging. *J Exp Med,* 147**,** 1526-31.

He, Y. F., Li, B. Z., Li, Z., Liu, P., Wang, Y., Tang, Q., Ding, J., Jia, Y., Chen, Z., Li, L., Sun, Y., Li, X., Dai, Q., Song, C. X., Zhang, K., He, C. & Xu, G. L. 2011. Tet-mediated formation of 5-carboxylcytosine and its excision by TDG in mammalian DNA. *Science,* 333**,** 1303-7.

Heinz, S., Romanoski, C. E., Benner, C. & Glass, C. K. 2015. The selection and function of cell type-specific enhancers. *Nat Rev Mol Cell Biol,* 16**,** 144-54.

Hettinger, J., Richards, D. M., Hansson, J., Barra, M. M., Joschko, A. C., Krijgsveld, J. & Feuerer, M. 2013. Origin of monocytes and macrophages in a committed progenitor. *Nat Immunol,* 14**,** 821-30.

Hodges, E., Molaro, A., Dos Santos, C. O., Thekkat, P., Song, Q., Uren, P. J., Park, J., Butler, J., Rafii, S., McCombie, W. R., Smith, A. D. & Hannon, G. J. 2011. Directional DNA methylation changes and complex

intermediate states accompany lineage specificity in the adult hematopoietic compartment. *Mol Cell,* 44**,** 17-28.

Holz-Schietinger, C., Matje, D. M. & Reich, N. O. 2012. Mutations in DNA methyltransferase (DNMT3A) observed in acute myeloid leukemia patients disrupt processive methylation. *J Biol Chem,* 287**,** 30941-51.

Hoppe, P. S., Schwarzfischer, M., Loeffler, D., Kokkaliaris, K. D., Hilsenbeck, O., Moritz, N., Endele, M., Filipczyk, A., Gambardella, A., Ahmed, N., Etzrodt, M., Coutu, D. L., Rieger, M. A., Marr, C., Strasser, M. K., Schauberger, B., Burtscher, I., Ermakova, O., Burger, A., Lickert, H., Nerlov, C., Theis, F. J. & Schroeder, T. 2016. Early myeloid lineage choice is not initiated by random PU.1 to GATA1 protein ratios. *Nature,* 535**,** 299-302.

Inoue, S., Li, W. Y., Tseng, A., Beerman, I., Elia, A. J., Bendall, S. C., Lemonnier, F., Kron, K. J., Cescon, D. W., Hao, Z., Lind, E. F., Takayama, N., Planello, A. C., Shen, S. Y., Shih, A. H., Larsen, D. M., Li, Q., Snow, B. E., Wakeham, A., Haight, J., Gorrini, C., Bassi, C., Thu, K. L., Murakami, K., Elford, A. R., Ueda, T., Straley, K., Yen, K. E., Melino, G., Cimmino, L., Aifantis, I., Levine, R. L., De Carvalho, D. D., Lupien, M., Rossi, D. J., Nolan, G. P., Cairns, R. A. & Mak, T. W. 2016. Mutant IDH1 Downregulates ATM and Alters DNA Repair and Sensitivity to DNA Damage Independent of TET2. *Cancer Cell,* 30**,** 337-48.

Ito, S., D'Alessio, A. C., Taranova, O. V., Hong, K., Sowers, L. C. & Zhang, Y. 2010. Role of Tet proteins in 5mC to 5hmC conversion, ES-cell self-renewal and inner cell mass specification. *Nature,* 466**,** 1129-33.

Ito, S., Shen, L., Dai, Q., Wu, S. C., Collins, L. B., Swenberg, J. A., He, C. & Zhang, Y. 2011. Tet proteins can convert 5-methylcytosine to 5-formylcytosine and 5-carboxylcytosine. *Science,* 333**,** 1300-3.

Iwasaki, H. & Akashi, K. 2007. Myeloid lineage commitment from the hematopoietic stem cell. *Immunity,* 26**,** 726-40.

Jack, G. D., Zhang, L. & Friedman, A. D. 2009. M-CSF elevates c-Fos and phospho-C/EBPalpha(S21) via ERK whereas G-CSF stimulates SHP2 phosphorylation in marrow progenitors to contribute to myeloid lineage specification. *Blood,* 114**,** 2172-80.

Jaiswal, S., Fontanillas, P., Flannick, J., Manning, A., Grauman, P. V., Mar, B. G., Lindsley, R. C., Mermel, C. H., Burtt, N., Chavez, A., Higgins, J. M., Moltchanov, V., Kuo, F. C., Kluk, M. J., Henderson, B., Kinnunen, L., Koistinen, H. A., Ladenvall, C., Getz, G., Correa, A., Banahan, B. F., Gabriel, S., Kathiresan, S., Stringham, H. M., McCarthy, M. I., Boehnke, M., Tuomilehto, J., Haiman, C., Groop, L., Atzmon, G., Wilson, J. G., Neuberg, D., Altshuler, D. & Ebert, B. L. 2014. Age-related clonal hematopoiesis associated with adverse outcomes. *N Engl J Med,* 371**,** 2488-98.

Jaiswal, S. & Libby, P. 2019. Clonal haematopoiesis: connecting ageing and inflammation in cardiovascular disease. *Nat Rev Cardiol.*

Jaiswal, S., Natarajan, P., Silver, A. J., Gibson, C. J., Bick, A. G., Shvartz, E., McConkey, M., Gupta, N., Gabriel, S., Ardissino, D., Baber, U., Mehran, R., Fuster, V., Danesh, J., Frossard, P., Saleheen, D., Melander, O., Sukhova, G. K., Neuberg, D., Libby, P., Kathiresan, S. & Ebert, B. L. 2017. Clonal Hematopoiesis and Risk of Atherosclerotic Cardiovascular Disease. *N Engl J Med.*

Jaitin, D. A., Weiner, A., Yofe, I., Lara-Astiaso, D., Keren-Shaul, H., David, E., Salame, T. M., Tanay, A., van Oudenaarden, A. & Amit, I. 2016. Dissecting Immune Circuits by Linking CRISPR-Pooled Screens with Single-Cell RNA-Seq. *Cell,* 167**,** 1883-1896.e15.

Jakubzick, C., Gautier, E. L., Gibbings, S. L., Sojka, D. K., Schlitzer, A., Johnson, T. E., Ivanov, S., Duan, Q., Bala, S., Condon, T., van Rooijen, N., Grainger, J. R., Belkaid, Y., Ma'ayan, A., Riches, D. W., Yokoyama, W. M., Ginhoux, F., Henson, P. M. & Randolph, G. J. 2013. Minimal differentiation of classical monocytes as they survey steady-state tissues and transport antigen to lymph nodes. *Immunity,* 39**,** 599-610.

Jakubzick, C. V., Randolph, G. J. & Henson, P. M. 2017. Monocyte differentiation and antigen-presenting functions. *Nat Rev Immunol,* 17**,** 349-362.

Javierre, B. M., Burren, O. S., Wilder, S. P., Kreuzhuber, R., Hill, S. M., Sewitz, S., Cairns, J., Wingett, S. W., Varnai, C., Thiecke, M. J., Burden, F., Farrow, S., Cutler, A. J., Rehnstrom, K., Downes, K., Grassi, L., Kostadima, M., Freire-Pritchett, P., Wang, F., Consortium, B., Stunnenberg, H. G., Todd, J. A., Zerbino, D. R., Stegle, O., Ouwehand, W. H., Frontini, M., Wallace, C., Spivakov, M. & Fraser, P. 2016. Lineage-Specific Genome Architecture Links Enhancers and Non-coding Disease Variants to Target Gene Promoters. *Cell,* 167**,** 1369-1384 e19.

Jensen, R. B., Carreira, A. & Kowalczykowski, S. C. 2010. Purified human BRCA2 stimulates RAD51-mediated recombination. *Nature,* 467**,** 678-83.

Jeong, M., Park, H. J., Celik, H., Ostrander, E. L., Reyes, J. M., Guzman, A., Rodriguez, B., Lei, Y., Lee, Y., Ding, L., Guryanova, O. A., Li, W., Goodell, M. A. & Challen, G. A. 2018. Loss of Dnmt3a Immortalizes Hematopoietic Stem Cells In Vivo. *Cell Rep,* 23**,** 1-10.

Jeong, M., Sun, D., Luo, M., Huang, Y., Challen, G. A., Rodriguez, B., Zhang, X., Chavez, L., Wang, H., Hannah, R., Kim, S. B., Yang, L., Ko, M., Chen, R., Gottgens, B., Lee, J. S., Gunaratne, P., Godley, L. A., Darlington, G. J., Rao, A., Li, W. & Goodell, M. A. 2014. Large conserved domains of low DNA methylation maintained by Dnmt3a. *Nat Genet,* 46**,** 17-23.

Jha, P. & Das, H. 2017. KLF2 in Regulation of NF-kappaB-Mediated Immune Cell Function and Inflammation. *Int J Mol Sci,* 18.

Ji, H., Ehrlich, L. I., Seita, J., Murakami, P., Doi, A., Lindau, P., Lee, H., Aryee, M. J., Irizarry, R. A., Kim, K., Rossi, D. J., Inlay, M. A., Serwold, T., Karsunky, H., Ho, L., Daley, G. Q., Weissman, I. L. & Feinberg, A. P. 2010. Comprehensive methylome map of lineage commitment from haematopoietic progenitors. *Nature,* 467**,** 338-42.

Jung, S., Aliberti, J., Graemmel, P., Sunshine, M. J., Kreutzberg, G. W., Sher, A. & Littman, D. R. 2000. Analysis of fractalkine receptor CX(3)CR1 function by targeted deletion and green fluorescent protein reporter gene insertion. *Mol Cell Biol,* 20**,** 4106-14.

Kameda, T., Shide, K., Yamaji, T., Kamiunten, A., Sekine, M., Taniguchi, Y., Hidaka, T., Kubuki, Y., Shimoda, H., Marutsuka, K., Sashida, G., Aoyama, K., Yoshimitsu, M., Harada, T., Abe, H., Miike, T., Iwakiri, H., Tahara, Y., Sueta, M., Yamamoto, S., Hasuike, S., Nagata, K., Iwama, A., Kitanaka, A. & Shimoda, K. 2015. Loss of TET2 has dual roles in murine myeloproliferative neoplasms: disease sustainer and disease accelerator. *Blood,* 125**,** 304-15.

Kamminga, L. M., van Os, R., Ausema, A., Noach, E. J., Weersing, E., Dontje, B., Vellenga, E. & de Haan, G. 2005. Impaired hematopoietic stem cell functioning after serial transplantation and during normal aging. *Stem Cells,* 23**,** 82-92.

Kats, L. M., Reschke, M., Taulli, R., Pozdnyakova, O., Burgess, K., Bhargava, P., Straley, K., Karnik, R., Meissner, A., Small, D., Su, S. M., Yen, K., Zhang, J. & Pandolfi, P. P. 2014. Proto-oncogenic role of mutant IDH2 in leukemia initiation and maintenance. *Cell Stem Cell,* 14**,** 329-41.

Kawai, T. & Akira, S. 2010. The role of pattern-recognition receptors in innate immunity: update on Toll-like receptors. *Nature Immunology,* 11**,** 373-384.

Kernytsky, A., Wang, F., Hansen, E., Schalm, S., Straley, K., Gliser, C., Yang, H., Travins, J., Murray, S., Dorsch, M., Agresta, S., Schenkein, D. P., Biller, S. A., Su, S. M., Liu, W. & Yen, K. E. 2015. IDH2 mutation-induced histone and DNA hypermethylation is progressively reversed by small-molecule inhibition. *Blood,* 125**,** 296-303.

Kiel, M. J., Yilmaz, O. H., Iwashita, T., Yilmaz, O. H., Terhorst, C. & Morrison, S. J. 2005. SLAM family receptors distinguish hematopoietic stem and progenitor cells and reveal endothelial niches for stem cells. *Cell,* 121**,** 1109-21.

Kondo, M., Weissman, I. L. & Akashi, K. 1997. Identification of Clonogenic Common Lymphoid Progenitors in Mouse Bone Marrow. *Cell,* 91**,** 661-672.

Kulis, M., Merkel, A., Heath, S., Queiros, A. C., Schuyler, R. P., Castellano, G., Beekman, R., Raineri, E., Esteve, A., Clot, G., Verdaguer-Dot, N., Duran-Ferrer, M., Russinol, N., Vilarrasa-Blasi, R., Ecker, S., Pancaldi, V., Rico, D., Agueda, L., Blanc, J., Richardson, D., Clarke, L., Datta, A., Pascual, M., Agirre, X., Prosper, F., Alignani, D., Paiva, B., Caron, G., Fest, T., Muench, M. O., Fomin, M. E., Lee, S. T., Wiemels, J. L., Valencia, A., Gut, M., Flicek, P., Stunnenberg, H. G., Siebert, R., Kuppers, R., Gut, I. G., Campo, E. & Martin-Subero, J. I. 2015. Whole-genome fingerprint of the DNA methylome during human B cell differentiation. *Nat Genet,* 47**,** 746-56.

Kumaran Satyanarayanan, S., El Kebir, D., Soboh, S., Butenko, S., Sekheri, M., Saadi, J., Peled, N., Assi, S., Othman, A., Schif-Zuck, S., Feuermann, Y., Barkan, D., Sher, N., Filep, J. G. & Ariel, A. 2019. IFN-beta is a macrophage-derived effector cytokine facilitating the resolution of bacterial inflammation. *Nat Commun,* 10**,** 3471.

Kundu, T. K., Palhan, V. B., Wang, Z., An, W., Cole, P. A. & Roeder, R. G. 2000. Activator-dependent transcription from chromatin in vitro involving targeted histone acetylation by p300. *Mol Cell,* 6**,** 551-61.

Kurotaki, D., Osato, N., Nishiyama, A., Yamamoto, M., Ban, T., Sato, H., Nakabayashi, J., Umehara, M., Miyake, N., Matsumoto, N., Nakazawa, M., Ozato, K. & Tamura, T. 2013. Essential role of the IRF8-KLF4 transcription factor cascade in murine monocyte differentiation. *Blood,* 121**,** 1839-49.

Lara-Astiaso, D., Weiner, A., Lorenzo-Vivas, E., Zaretsky, I., Jaitin, D. A., David, E., Keren-Shaul, H., Mildner, A., Winter, D., Jung, S., Friedman, N. & Amit, I. 2014. Immunogenetics. Chromatin state dynamics during blood formation. *Science,* 345**,** 943-9.

Laurenti, E. & Gottgens, B. 2018. From haematopoietic stem cells to complex differentiation landscapes. *Nature,* 553**,** 418-426.

Lawrence, T. & Natoli, G. 2011. Transcriptional regulation of macrophage polarization: enabling diversity with identity. *Nat Rev Immunol,* 11**,** 750-61.

Lee, P. Y., Sykes, D. B., Ameri, S., Kalaitzidis, D., Charles, J. F., Nelson-Maney, N., Wei, K., Cunin, P., Morris, A., Cardona, A. E., Root, D. E., Scadden, D. T. & Nigrovic, P. A. 2017. The metabolic regulator mTORC1 controls terminal myeloid differentiation. *Sci Immunol,* 2.

Lee, P. Y., Wang, J. X., Parisini, E., Dascher, C. C. & Nigrovic, P. A. 2013. Ly6 family proteins in neutrophil biology. *J Leukoc Biol,* 94**,** 585-94.

Lee-Six, H., Obro, N. F., Shepherd, M. S., Grossmann, S., Dawson, K., Belmonte, M., Osborne, R. J., Huntly, B. J. P., Martincorena, I., Anderson, E., O'Neill, L., Stratton, M. R., Laurenti, E., Green, A. R., Kent, D. G. & Campbell, P. J. 2018. Population dynamics of normal human blood inferred from somatic mutations. *Nature,* 561**,** 473-478.

Leeb, M., Pasini, D., Novatchkova, M., Jaritz, M., Helin, K. & Wutz, A. 2010. Polycomb complexes act redundantly to repress genomic repeats and genes. *Genes Dev,* 24**,** 265-76.

Ley, T. J., Ding, L., Walter, M. J., McLellan, M. D., Lamprecht, T., Larson, D. E., Kandoth, C., Payton, J. E., Baty, J., Welch, J., Harris, C. C., Lichti, C. F., Townsend, R. R., Fulton, R. S., Dooling, D. J., Koboldt, D. C., Schmidt, H., Zhang, Q., Osborne, J. R., Lin, L., O'Laughlin, M., McMichael, J. F., Delehaunty, K. D., McGrath, S. D., Fulton, L. A., Magrini, V. J., Vickery, T. L., Hundal, J., Cook, L. L., Conyers, J. J., Swift, G. W., Reed, J. P., Alldredge, P. A., Wylie, T., Walker, J., Kalicki, J., Watson, M. A., Heath, S., Shannon, W. D., Varghese, N., Nagarajan, R., Westervelt, P., Tomasson, M. H., Link, D. C., Graubert, T. A., DiPersio, J. F., Mardis, E. R. & Wilson, R. K. 2010. DNMT3A mutations in acute myeloid leukemia. *N Engl J Med,* 363**,** 2424-33.

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G. & Durbin, R. 2009. The Sequence Alignment/Map format and SAMtools. *Bioinformatics,* 25**,** 2078-9.

Li, S., Chou, A. P., Chen, W., Chen, R., Deng, Y., Phillips, H. S., Selfridge, J., Zurayk, M., Lou, J. J., Everson, R. G., Wu, K. C., Faull, K. F., Cloughesy, T., Liau, L. M. & Lai, A. 2013. Overexpression of isocitrate dehydrogenase mutant proteins renders glioma cells more sensitive to radiation. *Neuro Oncol,* 15**,** 57-68.

Liao, Y., Smyth, G. K. & Shi, W. 2014. featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics,* 30**,** 923-30.

Liberzon, A., Birger, C., Thorvaldsdottir, H., Ghandi, M., Mesirov, J. P. & Tamayo, P. 2015. The Molecular Signatures Database (MSigDB) hallmark gene set collection. *Cell Syst,* 1**,** 417-425.

Liberzon, A., Subramanian, A., Pinchback, R., Thorvaldsdottir, H., Tamayo, P. & Mesirov, J. P. 2011. Molecular signatures database (MSigDB) 3.0. *Bioinformatics,* 27**,** 1739-40.

Lienert, F., Mohn, F., Tiwari, V. K., Baubec, T., Roloff, T. C., Gaidatzis, D., Stadler, M. B. & Schubeler, D. 2011. Genomic prevalence of heterochromatic H3K9me2 and transcription do not discriminate pluripotent from terminally differentiated cells. *PLoS Genet,* 7**,** e1002090.

Lipka, D. B., Wang, Q., Cabezas-Wallscheid, N., Klimmeck, D., Weichenhan, D., Herrmann, C., Lier, A., Brocks, D., von Paleske, L., Renders, S., Wunsche, P., Zeisberger, P., Gu, L., Haas, S., Essers, M. A., Brors, B., Eils, R., Trumpp, A., Milsom, M. D. & Plass, C. 2014. Identification of DNA methylation changes at cis-regulatory elements during early steps of HSC differentiation using tagmentation-based whole genome bisulfite sequencing. *Cell Cycle,* 13**,** 3476-87.

Lister, R., Pelizzola, M., Dowen, R. H., Hawkins, R. D., Hon, G., Tonti-Filippini, J., Nery, J. R., Lee, L., Ye, Z., Ngo, Q. M., Edsall, L., Antosiewicz-Bourget, J., Stewart, R., Ruotti, V., Millar, A. H., Thomson, J. A., Ren, B. & Ecker, J. R. 2009. Human DNA methylomes at base resolution show widespread epigenomic differences. *Nature,* 462**,** 315-22.

Loberg, M. A., Bell, R. K., Goodwin, L. O., Eudy, E., Miles, L. A., SanMiguel, J. M., Young, K., Bergstrom, D. E., Levine, R. L., Schneider, R. K. & Trowbridge, J. J. 2019. Sequentially inducible mouse models reveal

that Npm1 mutation causes malignant transformation of Dnmt3a-mutant clonal hematopoiesis. *Leukemia*.

Loeb, L. A. & Monnat, R. J., Jr. 2008. DNA polymerases and human disease. *Nat Rev Genet,* 9**,** 594-604.

Lopez-Otin, C., Blasco, M. A., Partridge, L., Serrano, M. & Kroemer, G. 2013. The hallmarks of aging. *Cell,* 153**,** 1194-217.

Lord, C. J. & Ashworth, A. 2017. PARP inhibitors: Synthetic lethality in the clinic. *Science,* 355**,** 1152-1158.

Lord, K. A., Abdollahi, A., Hoffman-Liebermann, B. & Liebermann, D. A. 1993. Proto-oncogenes of the fos/jun family of transcription factors are positive regulators of myeloid differentiation. *Mol Cell Biol,* 13**,** 841-51.

Love, M. I., Huber, W. & Anders, S. 2014. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol,* 15**,** 550.

Lu, C., Ward, P. S., Kapoor, G. S., Rohle, D., Turcan, S., Abdel-Wahab, O., Edwards, C. R., Khanin, R., Figueroa, M. E., Melnick, A., Wellen, K. E., O'Rourke, D. M., Berger, S. L., Chan, T. A., Levine, R. L., Mellinghoff, I. K. & Thompson, C. B. 2012. IDH mutation impairs histone demethylation and results in a block to cell differentiation. *Nature,* 483**,** 474-8.

Luo, C., Hajkova, P. & Ecker, J. R. 2018. Dynamic DNA methylation: In the right place at the right time. *Science,* 361**,** 1336-1340.

Lutsik, P., Slawski, M., Gasparoni, G., Vedeneev, N., Hein, M. & Walter, J. 2017. MeDeCom: discovery and quantification of latent components of heterogeneous methylomes. *Genome Biol,* 18**,** 55.

Macaulay, I. C., Svensson, V., Labalette, C., Ferreira, L., Hamey, F., Voet, T., Teichmann, S. A. & Cvejic, A. 2016. Single-Cell RNA-Sequencing Reveals a Continuous Spectrum of Differentiation in Hematopoietic Cells. *Cell Rep,* 14**,** 966-977.

Mallaney, C., Ostrander, E. L., Celik, H., Kramer, A. C., Martens, A., Kothari, A., Koh, W. K., Haussler, E., Iwamori, N., Gontarz, P., Zhang, B. & Challen, G. A. 2019. Kdm6b regulates context-dependent hematopoietic stem cell self-renewal and leukemogenesis. *Leukemia,* 33**,** 2506-2521.

Mallette, F. A., Mattiroli, F., Cui, G., Young, L. C., Hendzel, M. J., Mer, G., Sixma, T. K. & Richard, S. 2012. RNF8- and RNF168-dependent degradation of KDM4A/JMJD2A triggers 53BP1 recruitment to DNA damage sites. *Embo j,* 31**,** 1865-78.

Mardis, E. R., Ding, L., Dooling, D. J., Larson, D. E., McLellan, M. D., Chen, K., Koboldt, D. C., Fulton, R. S., Delehaunty, K. D., McGrath, S. D., Fulton, L. A., Locke, D. P., Magrini, V. J., Abbott, R. M., Vickery, T. L., Reed, J. S., Robinson, J. S., Wylie, T., Smith, S. M., Carmichael, L., Eldred, J. M., Harris, C. C., Walker, J., Peck, J. B., Du, F., Dukes, A. F., Sanderson, G. E., Brummett, A. M., Clark, E., McMichael, J. F., Meyer, R. J., Schindler, J. K., Pohl, C. S., Wallis, J. W., Shi, X., Lin, L., Schmidt, H., Tang, Y., Haipek, C., Wiechert, M. E., Ivy, J. V., Kalicki, J., Elliott, G., Ries, R. E., Payton, J. E., Westervelt, P., Tomasson, M. H., Watson, M. A., Baty, J., Heath, S., Shannon, W. D., Nagarajan, R., Link, D. C., Walter, M. J., Graubert, T. A., DiPersio, J. F., Wilson, R. K. & Ley, T. J. 2009. Recurring mutations found by sequencing an acute myeloid leukemia genome. *N Engl J Med,* 361**,** 1058-66.

Marteijn, J. A., Lans, H., Vermeulen, W. & Hoeijmakers, J. H. 2014. Understanding nucleotide excision repair and its roles in cancer and ageing. *Nat Rev Mol Cell Biol,* 15**,** 465-81.

Maurer, M. & von Stebut, E. 2004. Macrophage inflammatory protein-1. *Int J Biochem Cell Biol,* 36**,** 1882-6.

Mayle, A., Yang, L., Rodriguez, B., Zhou, T., Chang, E., Curry, C. V., Challen, G. A., Li, W., Wheeler, D., Rebel, V. I. & Goodell, M. A. 2015. Dnmt3a loss predisposes murine hematopoietic stem cells to malignant transformation. *Blood,* 125**,** 629-38.

McVey, M., Khodaverdian, V. Y., Meyer, D., Cerqueira, P. G. & Heyer, W. D. 2016. Eukaryotic DNA Polymerases in Homologous Recombination. *Annu Rev Genet,* 50**,** 393-421.

Meisel, M., Hinterleitner, R., Pacis, A., Chen, L., Earley, Z. M., Mayassi, T., Pierre, J. F., Ernest, J. D., Galipeau, H. J., Thuille, N., Bouziat, R., Buscarlet, M., Ringus, D. L., Wang, Y., Li, Y., Dinh, V., Kim, S. M., McDonald, B. D., Zurenski, M. A., Musch, M. W., Furtado, G. C., Lira, S. A., Baier, G., Chang, E. B., Eren, A. M., Weber, C. R., Busque, L., Godley, L. A., Verdu, E. F., Barreiro, L. B. & Jabri, B. 2018. Microbial signals drive pre-leukaemic myeloproliferation in a Tet2-deficient host. *Nature,* 557**,** 580-584.

Meyer, S. E., Qin, T., Muench, D. E., Masuda, K., Venkatasubramanian, M., Orr, E., Suarez, L., Gore, S. D., Delwel, R., Paietta, E., Tallman, M. S., Fernandez, H., Melnick, A., Le Beau, M. M., Kogan, S., Salomonis, N., Figueroa, M. E. & Grimes, H. L. 2016. DNMT3A Haploinsufficiency Transforms FLT3ITD Myeloproliferative Disease into a Rapid, Spontaneous, and Fully Penetrant Acute Myeloid Leukemia. *Cancer Discov,* 6**,** 501-15.

Miki, Y., Swensen, J., Shattuck-Eidens, D., Futreal, P. A., Harshman, K., Tavtigian, S., Liu, Q., Cochran, C., Bennett, L. M., Ding, W. & et al. 1994. A strong candidate for the breast and ovarian cancer susceptibility gene BRCA1. *Science,* 266**,** 66-71.

Mildner, A., Schonheit, J., Giladi, A., David, E., Lara-Astiaso, D., Lorenzo-Vivas, E., Paul, F., Chappell-Maor, L., Priller, J., Leutz, A., Amit, I. & Jung, S. 2017. Genomic Characterization of Murine Monocytes Reveals C/EBPbeta Transcription Factor Dependence of Ly6C(-) Cells. *Immunity,* 46**,** 849-862 e7.

Mimitou, E. P. & Symington, L. S. 2008. Sae2, Exo1 and Sgs1 collaborate in DNA double-strand break processing. *Nature,* 455**,** 770-4.

Mogensen, T. H. 2018. IRF and STAT Transcription Factors - From Basic Biology to Roles in Infection, Protective Immunity, and Primary Immunodeficiencies. *Front Immunol,* 9**,** 3047.

Molenaar, R. J., Maciejewski, J. P., Wilmink, J. W. & van Noorden, C. J. F. 2018a. Wild-type and mutated IDH1/2 enzymes and therapy responses. *Oncogene,* 37**,** 1949-1960.

Molenaar, R. J., Radivoyevitch, T., Nagata, Y., Khurshed, M., Przychodzen, B., Makishima, H., Xu, M., Bleeker, F. E., Wilmink, J. W., Carraway, H. E., Mukherjee, S., Sekeres, M. A., van Noorden, C. J. F. & Maciejewski, J. P. 2018b. IDH1/2 Mutations Sensitize Acute Myeloid Leukemia to PARP Inhibition and This Is Reversed by IDH1/2-Mutant Inhibitors. *Clin Cancer Res,* 24**,** 1705-1715.

Mootha, V. K., Lindgren, C. M., Eriksson, K. F., Subramanian, A., Sihag, S., Lehar, J., Puigserver, P., Carlsson, E., Ridderstrale, M., Laurila, E., Houstis, N., Daly, M. J., Patterson, N., Mesirov, J. P., Golub, T. R., Tamayo, P., Spiegelman, B., Lander, E. S., Hirschhorn, J. N., Altshuler, D. & Groop, L. C. 2003. PGC-1alpha-responsive genes involved in oxidative phosphorylation are coordinately downregulated in human diabetes. *Nat Genet,* 34**,** 267-73.

Moran-Crusio, K., Reavie, L., Shih, A., Abdel-Wahab, O., Ndiaye-Lobry, D., Lobry, C., Figueroa, M. E., Vasanthakumar, A., Patel, J., Zhao, X., Perna, F., Pandey, S., Madzo, J., Song, C., Dai, Q., He, C., Ibrahim, S., Beran, M., Zavadil, J., Nimer, S. D., Melnick, A., Godley, L. A., Aifantis, I. & Levine, R. L. 2011. Tet2 loss leads to increased hematopoietic stem cell self-renewal and myeloid transformation. *Cancer Cell,* 20**,** 11-24.

Morrison, S. J., Wandycz, A. M., Hemmati, H. D., Wright, D. E. & Weissman, I. L. 1997. Identification of a lineage of multipotent hematopoietic progenitors. *Development,* 124**,** 1929-39.

Morrison, S. J. & Weissman, I. L. 1994. The long-term repopulating subset of hematopoietic stem cells is deterministic and isolatable by phenotype. *Immunity,* 1**,** 661-73.

Mosser, D. M. & Edwards, J. P. 2008. Exploring the full spectrum of macrophage activation. *Nat Rev Immunol,* 8**,** 958-69.

Muller-Sieburg, C. E., Cho, R. H., Karlsson, L., Huang, J. F. & Sieburg, H. B. 2004. Myeloid-biased hematopoietic stem cells have extensive self-renewal capacity but generate diminished lymphoid progeny with impaired IL-7 responsiveness. *Blood,* 103**,** 4111-8.

Murai, J., Huang, S. Y., Das, B. B., Renaud, A., Zhang, Y., Doroshow, J. H., Ji, J., Takeda, S. & Pommier, Y. 2012. Trapping of PARP1 and PARP2 by Clinical PARP Inhibitors. *Cancer Res,* 72**,** 5588-99.

Murray, P. J. 2017. Macrophage Polarization. *Annu Rev Physiol,* 79**,** 541-566.

Na Nakorn, T., Traver, D., Weissman, I. L. & Akashi, K. 2002. Myeloerythroid-restricted progenitors are sufficient to confer radioprotection and provide the majority of day 8 CFU-S. *J Clin Invest,* 109**,** 1579-85.

Naik, S. H., Sathe, P., Park, H. Y., Metcalf, D., Proietto, A. I., Dakic, A., Carotta, S., O'Keeffe, M., Bahlo, M., Papenfuss, A., Kwak, J. Y., Wu, L. & Shortman, K. 2007. Development of plasmacytoid and conventional dendritic cell subtypes from single precursor cells derived in vitro and in vivo. *Nat Immunol,* 8**,** 1217-26.

Nam, A. S., Kim, K.-T., Chaligne, R., Izzo, F., Ang, C., Taylor, J., Myers, R. M., Abu-Zeinah, G., Brand, R., Omans, N. D., Alonso, A., Sheridan, C., Mariani, M., Dai, X., Harrington, E., Pastore, A., Cubillos-Ruiz, J. R., Tam,

W., Hoffman, R., Rabadan, R., Scandura, J. M., Abdel-Wahab, O., Smibert, P. & Landau, D. A. 2019. Somatic mutations and cell identity linked by Genotyping of Transcriptomes. *Nature.*

Nayak, L., Goduni, L., Takami, Y., Sharma, N., Kapil, P., Jain, M. K. & Mahabeleshwar, G. H. 2013. Kruppel-like factor 2 is a transcriptional regulator of chronic and acute inflammation. *Am J Pathol,* 182**,** 1696-704.

Nestorowa, S., Hamey, F. K., Pijuan Sala, B., Diamanti, E., Shepherd, M., Laurenti, E., Wilson, N. K., Kent, D. G. & Gottgens, B. 2016. A single-cell resolution map of mouse hematopoietic stem and progenitor cell differentiation. *Blood,* 128**,** e20-31.

Netea, M. G., Joosten, L. A. B., Latz, E., Mills, K. H. G., Natoli, G., Stunnenberg, H. G., O'Neill, L. A. J. & Xavier, R. J. 2016. Trained immunity: A program of innate immune memory in health and disease. *Science,* 352.

Nie, Z., Hu, G., Wei, G., Cui, K., Yamane, A., Resch, W., Wang, R., Green, D. R., Tessarollo, L., Casellas, R., Zhao, K. & Levens, D. 2012. c-Myc is a universal amplifier of expressed genes in lymphocytes and embryonic stem cells. *Cell,* 151**,** 68-79.

Nierkens, S., Tel, J., Janssen, E. & Adema, G. J. 2013. Antigen cross-presentation by dendritic cell subsets: one general or all sergeants? *Trends Immunol,* 34**,** 361-70.

Nimonkar, A. V., Genschel, J., Kinoshita, E., Polaczek, P., Campbell, J. L., Wyman, C., Modrich, P. & Kowalczykowski, S. C. 2011. BLM-DNA2-RPA-MRN and EXO1-BLM-RPA-MRN constitute two DNA end resection machineries for human DNA break repair. *Genes Dev,* 25**,** 350-62.

Noushmehr, H., Weisenberger, D. J., Diefes, K., Phillips, H. S., Pujara, K., Berman, B. P., Pan, F., Pelloski, C. E., Sulman, E. P., Bhat, K. P., Verhaak, R. G., Hoadley, K. A., Hayes, D. N., Perou, C. M., Schmidt, H. K., Ding, L., Wilson, R. K., Van Den Berg, D., Shen, H., Bengtsson, H., Neuvial, P., Cope, L. M., Buckley, J., Herman, J. G., Baylin, S. B., Laird, P. W. & Aldape, K. 2010. Identification of a CpG island methylator phenotype that defines a distinct subgroup of glioma. *Cancer Cell,* 17**,** 510-22.

Oakes, C. C., Seifert, M., Assenov, Y., Gu, L., Przekopowitz, M., Ruppert, A. S., Wang, Q., Imbusch, C. D., Serva, A., Koser, S. D., Brocks, D., Lipka, D. B., Bogatyrova, O., Weichenhan, D., Brors, B., Rassenti, L., Kipps, T. J., Mertens, D., Zapatka, M., Lichter, P., Dohner, H., Kuppers, R., Zenz, T., Stilgenbauer, S., Byrd, J. C. & Plass, C. 2016. DNA methylation dynamics during B cell maturation underlie a continuum of disease phenotypes in chronic lymphocytic leukemia. *Nat Genet,* 48**,** 253-64.

Ogawara, Y., Katsumoto, T., Aikawa, Y., Shima, Y., Kagiyama, Y., Soga, T., Matsunaga, H., Seki, T., Araki, K. & Kitabayashi, I. 2015. IDH2 and NPM1 Mutations Cooperate to Activate Hoxa9/Meis1 and Hypoxia Pathways in Acute Myeloid Leukemia. *Cancer Res,* 75**,** 2005-16.

Okano, M., Bell, D. W., Haber, D. A. & Li, E. 1999. DNA Methyltransferases Dnmt3a and Dnmt3b Are Essential for De Novo Methylation and Mammalian Development. *Cell,* 99**,** 247-257.

Okano, M., Xie, S. & Li, E. 1998. Cloning and characterization of a family of novel mammalian DNA (cytosine-5) methyltransferases. *Nat Genet,* 19**,** 219-20.

Olsson, A., Venkatasubramanian, M., Chaudhri, V. K., Aronow, B. J., Salomonis, N., Singh, H. & Grimes, H. L. 2016. Single-cell analysis of mixed-lineage states leading to a binary cell fate choice. *Nature,* 537**,** 698-702.

Onai, N., Obata-Onai, A., Schmid, M. A., Ohteki, T., Jarrossay, D. & Manz, M. G. 2007. Identification of clonogenic common Flt3+M-CSFR+ plasmacytoid and conventional dendritic cell progenitors in mouse bone marrow. *Nat Immunol,* 8**,** 1207-16.

Orecchioni, M., Ghosheh, Y., Pramod, A. B. & Ley, K. 2019. Macrophage Polarization: Different Gene Signatures in M1(LPS+) vs. Classically and M2(LPS-) vs. Alternatively Activated Macrophages. *Front Immunol,* 10**,** 1084.

Orkin, S. H. 2000. Diversification of haematopoietic stem cells to specific lineages. *Nat Rev Genet,* 1**,** 57-64.

Orkin, S. H. & Zon, L. I. 2008. Hematopoiesis: an evolving paradigm for stem cell biology. *Cell,* 132**,** 631-44.

Palframan, R. T., Jung, S., Cheng, G., Weninger, W., Luo, Y., Dorf, M., Littman, D. R., Rollins, B. J., Zweerink, H., Rot, A. & von Andrian, U. H. 2001. Inflammatory chemokine transport and presentation in HEV: a remote control mechanism for monocyte recruitment to lymph nodes in inflamed tissues. *J Exp Med,* 194**,** 1361-73.

Papaemmanuil, E., Gerstung, M., Bullinger, L., Gaidzik, V. I., Paschka, P., Roberts, N. D., Potter, N. E., Heuser, M., Thol, F., Bolli, N., Gundem, G., Van Loo, P., Martincorena, I., Ganly, P., Mudie, L., McLaren, S., O'Meara,

S., Raine, K., Jones, D. R., Teague, J. W., Butler, A. P., Greaves, M. F., Ganser, A., Dohner, K., Schlenk, R. F., Dohner, H. & Campbell, P. J. 2016. Genomic Classification and Prognosis in Acute Myeloid Leukemia. *N Engl J Med,* 374**,** 2209-21.

Parsons, D. W., Jones, S., Zhang, X., Lin, J. C., Leary, R. J., Angenendt, P., Mankoo, P., Carter, H., Siu, I. M., Gallia, G. L., Olivi, A., McLendon, R., Rasheed, B. A., Keir, S., Nikolskaya, T., Nikolsky, Y., Busam, D. A., Tekleab, H., Diaz, L. A., Jr., Hartigan, J., Smith, D. R., Strausberg, R. L., Marie, S. K., Shinjo, S. M., Yan, H., Riggins, G. J., Bigner, D. D., Karchin, R., Papadopoulos, N., Parmigiani, G., Vogelstein, B., Velculescu, V. E. & Kinzler, K. W. 2008. An integrated genomic analysis of human glioblastoma multiforme. *Science,* 321**,** 1807-12.

Passlick, B., Flieger, D. & Ziegler-Heitbrock, H. W. 1989. Identification and characterization of a novel monocyte subpopulation in human peripheral blood. *Blood,* 74**,** 2527-34.

Patel, A. A., Zhang, Y., Fullerton, J. N., Boelen, L., Rongvaux, A., Maini, A. A., Bigley, V., Flavell, R. A., Gilroy, D. W., Asquith, B., Macallan, D. & Yona, S. 2017. The fate and lifespan of human monocyte subsets in steady state and systemic inflammation. *J Exp Med,* 214**,** 1913-1923.

Paul, A. & Paul, S. 2014. The breast cancer susceptibility genes (BRCA) in breast and ovarian cancers. *Front Biosci (Landmark Ed),* 19**,** 605-18.

Paul, F., Arkin, Y., Giladi, A., Jaitin, D. A., Kenigsberg, E., Keren-Shaul, H., Winter, D., Lara-Astiaso, D., Gury, M., Weiner, A., David, E., Cohen, N., Lauridsen, F. K., Haas, S., Schlitzer, A., Mildner, A., Ginhoux, F., Jung, S., Trumpp, A., Porse, B. T., Tanay, A. & Amit, I. 2015. Transcriptional Heterogeneity and Lineage Commitment in Myeloid Progenitors. *Cell,* 163**,** 1663-77.

Picelli, S., Faridani, O. R., Bjorklund, A. K., Winberg, G., Sagasser, S. & Sandberg, R. 2014. Full-length RNA-seq from single cells using Smart-seq2. *Nat Protoc,* 9**,** 171-81.

Pietras, E. M. 2017. Inflammation: a key regulator of hematopoietic stem cell fate in health and disease. *Blood,* 130**,** 1693-1698.

Pietras, E. M., Reynaud, D., Kang, Y. A., Carlin, D., Calero-Nieto, F. J., Leavitt, A. D., Stuart, J. M., Gottgens, B. & Passegue, E. 2015. Functionally Distinct Subsets of Lineage-Biased Multipotent Progenitors Control Blood Production in Normal and Regenerative Conditions. *Cell Stem Cell,* 17**,** 35-46.

Pines, A., Vrouwe, M. G., Marteijn, J. A., Typas, D., Luijsterburg, M. S., Cansoy, M., Hensbergen, P., Deelder, A., de Groot, A., Matsumoto, S., Sugasawa, K., Thoma, N., Vermeulen, W., Vrieling, H. & Mullenders, L. 2012. PARP1 promotes nucleotide excision repair through DDB2 stabilization and recruitment of ALC1. *J Cell Biol,* 199**,** 235-49.

Pinho, S. & Frenette, P. S. 2019. Haematopoietic stem cell activity and interactions with the niche. *Nat Rev Mol Cell Biol,* 20**,** 303-320.

Pommier, Y., O'Connor, M. J. & de Bono, J. 2016. Laying a trap to kill cancer cells: PARP inhibitors and their mechanisms of action. *Sci Transl Med,* 8**,** 362ps17.

Pronk, C. J., Rossi, D. J., Mansson, R., Attema, J. L., Norddahl, G. L., Chan, C. K., Sigvardsson, M., Weissman, I. L. & Bryder, D. 2007. Elucidation of the phenotypic, functional, and molecular topography of a myeloerythroid progenitor cell hierarchy. *Cell Stem Cell,* 1**,** 428-42.

Rao, S. S., Huntley, M. H., Durand, N. C., Stamenova, E. K., Bochkov, I. D., Robinson, J. T., Sanborn, A. L., Machol, I., Omer, A. D., Lander, E. S. & Aiden, E. L. 2014. A 3D map of the human genome at kilobase resolution reveals principles of chromatin looping. *Cell,* 159**,** 1665-80.

Ray Chaudhuri, A. & Nussenzweig, A. 2017. The multifaceted roles of PARP1 in DNA repair and chromatin remodelling. *Nat Rev Mol Cell Biol,* 18**,** 610-621.

Rico, D., Martens, J. H. A., Downes, K., Carrillo-de-Santa-Pau, E., Pancaldi, V., Breschi, A., Richardson, D., Heath, S., Saeed, S., Frontini, M., Chen, L., Watt, S., Müller, F., Clarke, L., Kerstens, H. H. D., Wilder, S. P., Palumbo, E., Djebali, S., Raineri, E., Merkel, A., Esteve-Codina, A., Sultan, M., van Bommel, A., Gut, M., Yaspo, M.-L., Rubio, M., Fernandez, J. M., Attwood, A., de la Torre, V., Royo, R., Fragkogianni, S., Gelpí, J. L., Torrents, D., Iotchkova, V., Logie, C., Aghajanirefah, A., Singh, A. A., Janssen-Megens, E. M., Berentsen, K., Erber, W., Rendon, A., Kostadima, M., Loos, R., van der Ent, M. A., Kaan, A., Sharifi, N., Paul, D. S., Ifrim, D. C., Quintin, J., Love, M. I., Pisano, D. G., Burden, F., Foad, N., Farrow, S., Zerbino, D. R., Dunham, I., Kuijpers, T., Lehrach, H., Lengauer, T., Bertone, P., Netea, M. G., Vingron, M., Beck, S., Flicek, P., Gut, I., Ouwehand, W. H., Bock, C., Soranzo, N., Guigo, R., Valencia, A. & Stunnenberg, H. G. 2017. Comparative analysis of neutrophil and monocyte epigenomes.

Robu, M., Shah, R. G., Purohit, N. K., Zhou, P., Naegeli, H. & Shah, G. M. 2017. Poly(ADP-ribose) polymerase 1 escorts XPC to UV-induced DNA lesions during nucleotide excision repair. *Proc Natl Acad Sci U S A,* 114**,** E6847-e6856.

Rodriguez-Fraticelli, A. E., Wolock, S. L., Weinreb, C. S., Panero, R., Patel, S. H., Jankovic, M., Sun, J., Calogero, R. A., Klein, A. M. & Camargo, F. D. 2018. Clonal analysis of lineage fate in native haematopoiesis. *Nature,* 553**,** 212-216.

Roller, A., Grossmann, V., Bacher, U., Poetzinger, F., Weissmann, S., Nadarajah, N., Boeck, L., Kern, W., Haferlach, C., Schnittger, S., Haferlach, T. & Kohlmann, A. 2013. Landmark analysis of DNMT3A mutations in hematological malignancies. *Leukemia,* 27**,** 1573-8.

Rossi, D. J., Bryder, D., Seita, J., Nussenzweig, A., Hoeijmakers, J. & Weissman, I. L. 2007. Deficiencies in DNA damage repair limit the function of haematopoietic stem cells with age. *Nature,* 447**,** 725-9.

Rossi, D. J., Jamieson, C. H. & Weissman, I. L. 2008. Stems cells and the pathways to aging and cancer. *Cell,* 132**,** 681-96.

Rowley, J. D. 1973. Letter: A new consistent chromosomal abnormality in chronic myelogenous leukaemia identified by quinacrine fluorescence and Giemsa staining. *Nature,* 243**,** 290-3.

Rube, C. E., Fricke, A., Widmann, T. A., Furst, T., Madry, H., Pfreundschuh, M. & Rube, C. 2011. Accumulation of DNA damage in hematopoietic stem and progenitor cells during human aging. *PLoS One,* 6**,** e17487.

Russler-Germain, D. A., Spencer, D. H., Young, M. A., Lamprecht, T. L., Miller, C. A., Fulton, R., Meyer, M. R., Erdmann-Gilmore, P., Townsend, R. R., Wilson, R. K. & Ley, T. J. 2014. The R882H DNMT3A mutation associated with AML dominantly inhibits wild-type DNMT3A by blocking its ability to form active tetramers. *Cancer Cell,* 25**,** 442-54.

Saeed, S., Quintin, J., Kerstens, H. H., Rao, N. A., Aghajanirefah, A., Matarese, F., Cheng, S. C., Ratter, J., Berentsen, K., van der Ent, M. A., Sharifi, N., Janssen-Megens, E. M., Ter Huurne, M., Mandoli, A., van Schaik, T., Ng, A., Burden, F., Downes, K., Frontini, M., Kumar, V., Giamarellos-Bourboulis, E. J., Ouwehand, W. H., van der Meer, J. W., Joosten, L. A., Wijmenga, C., Martens, J. H., Xavier, R. J., Logie, C., Netea, M. G. & Stunnenberg, H. G. 2014. Epigenetic programming of monocyte-to-macrophage differentiation and trained innate immunity. *Science,* 345**,** 1251086.

San Filippo, J., Sung, P. & Klein, H. 2008. Mechanism of eukaryotic homologous recombination. *Annu Rev Biochem,* 77**,** 229-57.

Sanchez, M., Gottgens, B., Sinclair, A. M., Stanley, M., Begley, C. G., Hunter, S. & Green, A. R. 1999. An SCL 3' enhancer targets developing endothelium together with embryonic and adult haematopoietic progenitors. *Development,* 126**,** 3891-904.

Sanchez, M. J., Bockamp, E. O., Miller, J., Gambardella, L. & Green, A. R. 2001. Selective rescue of early haematopoietic progenitors in Scl(-/-) mice by expressing Scl under the control of a stem cell enhancer. *Development,* 128**,** 4815-27.

Sanjuan-Pla, A., Macaulay, I. C., Jensen, C. T., Woll, P. S., Luis, T. C., Mead, A., Moore, S., Carella, C., Matsuoka, S., Bouriez Jones, T., Chowdhury, O., Stenson, L., Lutteropp, M., Green, J. C., Facchini, R., Boukarabila, H., Grover, A., Gambardella, A., Thongjuea, S., Carrelha, J., Tarrant, P., Atkinson, D., Clark, S. A., Nerlov, C. & Jacobsen, S. E. 2013. Platelet-biased stem cells reside at the apex of the haematopoietic stem-cell hierarchy. *Nature,* 502**,** 232-6.

Sanson, M., Marie, Y., Paris, S., Idbaih, A., Laffaire, J., Ducray, F., El Hallani, S., Boisselier, B., Mokhtari, K., Hoang-Xuan, K. & Delattre, J. Y. 2009. Isocitrate dehydrogenase 1 codon 132 mutation is an important prognostic biomarker in gliomas. *J Clin Oncol,* 27**,** 4150-4.

Sasaki, M., Knobbe, C. B., Munger, J. C., Lind, E. F., Brenner, D., Brustle, A., Harris, I. S., Holmes, R., Wakeham, A., Haight, J., You-Ten, A., Li, W. Y., Schalm, S., Su, S. M., Virtanen, C., Reifenberger, G., Ohashi, P. S., Barber, D. L., Figueroa, M. E., Melnick, A., Zuniga-Pflucker, J. C. & Mak, T. W. 2012. IDH1(R132H) mutation increases murine haematopoietic progenitors and alters epigenetics. *Nature,* 488**,** 656-9.

Sawai, C. M., Babovic, S., Upadhaya, S., Knapp, D. J., Lavin, Y., Lau, C. M., Goloborodko, A., Feng, J., Fujisaki, J., Ding, L., Mirny, L. A., Merad, M., Eaves, C. J. & Reizis, B. 2016. Hematopoietic Stem Cells Are the Major Source of Multilineage Hematopoiesis in Adult Animals. *Immunity,* 45**,** 597-609.

Saxonov, S., Berg, P. & Brutlag, D. L. 2006. A genome-wide analysis of CpG dinucleotides in the human genome distinguishes two distinct classes of promoters. *Proc Natl Acad Sci U S A,* 103**,** 1412-7.

Schmezer, P., Rajaee-Behbahani, N., Risch, A., Thiel, S., Rittgen, W., Drings, P., Dienemann, H., Kayser, K. W., Schulz, V. & Bartsch, H. 2001. Rapid screening assay for mutagen sensitivity and DNA repair capacity in human peripheral blood lymphocytes. *Mutagenesis,* 16**,** 25-30.

Scully, R., Panday, A., Elango, R. & Willis, N. A. 2019. DNA double-strand break repair-pathway choice in somatic mammalian cells. *Nat Rev Mol Cell Biol,* 20**,** 698-714.

Seita, J. & Weissman, I. L. 2010. Hematopoietic stem cell: self-renewal versus differentiation. *Wiley Interdiscip Rev Syst Biol Med,* 2**,** 640-53.

Selvakumaran, M., Liebermann, D. & Hoffman, B. 1996. The proto-oncogene c-myc blocks myeloid differentiation independently of its target gene ornithine decarboxylase. *Blood,* 88**,** 1248-55.

Serbina, N. V., Jia, T., Hohl, T. M. & Pamer, E. G. 2008. Monocyte-mediated defense against microbial pathogens. *Annu Rev Immunol,* 26**,** 421-52.

Seternes, O. M., Kidger, A. M. & Keyse, S. M. 2019. Dual-specificity MAP kinase phosphatases in health and disease. *Biochim Biophys Acta Mol Cell Res,* 1866**,** 124-143.

Shi, C. & Pamer, E. G. 2011. Monocyte recruitment during infection and inflammation. *Nat Rev Immunol,* 11**,** 762-74.

Shi, J., Sun, B., Shi, W., Zuo, H., Cui, D., Ni, L. & Chen, J. 2015. Decreasing GSH and increasing ROS in chemosensitivity gliomas with IDH1 mutation. *Tumour Biol,* 36**,** 655-62.

Shi, J., Zuo, H., Ni, L., Xia, L., Zhao, L., Gong, M., Nie, D., Gong, P., Cui, D., Shi, W. & Chen, J. 2014. An IDH1 mutation inhibits growth of glioma cells via GSH depletion and ROS generation. *Neurol Sci,* 35**,** 839-45.

Shih, A. H., Abdel-Wahab, O., Patel, J. P. & Levine, R. L. 2012. The role of mutations in epigenetic regulators in myeloid malignancies. *Nat Rev Cancer,* 12**,** 599-612.

Shiloh, Y. & Ziv, Y. 2013. The ATM protein kinase: regulating the cellular response to genotoxic stress, and more. *Nat Rev Mol Cell Biol,* 14**,** 197-210.

Shlush, L. I., Zandi, S., Mitchell, A., Chen, W. C., Brandwein, J. M., Gupta, V., Kennedy, J. A., Schimmer, A. D., Schuh, A. C., Yee, K. W., McLeod, J. L., Doedens, M., Medeiros, J. J., Marke, R., Kim, H. J., Lee, K., McPherson, J. D., Hudson, T. J., Consortium, H. P.-L. G. P., Brown, A. M., Yousif, F., Trinh, Q. M., Stein, L. D., Minden, M. D., Wang, J. C. & Dick, J. E. 2014. Identification of pre-leukaemic haematopoietic stem cells in acute leukaemia. *Nature,* 506**,** 328-33.

Smith, Z. D. & Meissner, A. 2013. DNA methylation: roles in mammalian development. *Nat Rev Genet,* 14**,** 204-20.

Spangrude, G. J., Heimfeld, S. & Weissman, I. L. 1988. Purification and characterization of mouse hematopoietic stem cells. *Science,* 241**,** 58-62.

Spencer, D. H., Russler-Germain, D. A., Ketkar, S., Helton, N. M., Lamprecht, T. L., Fulton, R. S., Fronick, C. C., O'Laughlin, M., Heath, S. E., Shinawi, M., Westervelt, P., Payton, J. E., Wartman, L. D., Welch, J. S., Wilson, R. K., Walter, M. J., Link, D. C., DiPersio, J. F. & Ley, T. J. 2017. CpG Island Hypermethylation Mediated by DNMT3A Is a Consequence of AML Progression. *Cell,* 168**,** 801-816 e13.

Sperling, A. S., Gibson, C. J. & Ebert, B. L. 2017. The genetics of myelodysplastic syndrome: from clonal haematopoiesis to secondary leukaemia. *Nat Rev Cancer,* 17**,** 5-19.

Srinivas, S., Watanabe, T., Lin, C. S., William, C. M., Tanabe, Y., Jessell, T. M. & Costantini, F. 2001. Cre reporter strains produced by targeted insertion of EYFP and ECFP into the ROSA26 locus. *BMC Dev Biol,* 1**,** 4.

Steensma, D. P., Bejar, R., Jaiswal, S., Lindsley, R. C., Sekeres, M. A., Hasserjian, R. P. & Ebert, B. L. 2015. Clonal hematopoiesis of indeterminate potential and its distinction from myelodysplastic syndromes. *Blood,* 126**,** 9-16.

Stephens, M. 2017. False discovery rates: a new deal. *Biostatistics,* 18**,** 275-294.

Stine, Z. E., Walton, Z. E., Altman, B. J., Hsieh, A. L. & Dang, C. V. 2015. MYC, Metabolism, and Cancer. *Cancer Discov,* 5**,** 1024-39.

Street, K., Risso, D., Fletcher, R. B., Das, D., Ngai, J., Yosef, N., Purdom, E. & Dudoit, S. 2018. Slingshot: cell lineage and pseudotime inference for single-cell transcriptomics. *BMC Genomics,* 19**,** 477.

Stricker, S. H., Koferle, A. & Beck, S. 2017. From profiles to function in epigenomics. *Nat Rev Genet,* 18**,** 51-66.

Stuart, T., Butler, A., Hoffman, P., Hafemeister, C., Papalexi, E., Mauck, W. M., 3rd, Hao, Y., Stoeckius, M., Smibert, P. & Satija, R. 2019. Comprehensive Integration of Single-Cell Data. *Cell,* 177**,** 1888-1902.e21.

Subramanian, A., Tamayo, P., Mootha, V. K., Mukherjee, S., Ebert, B. L., Gillette, M. A., Paulovich, A., Pomeroy, S. L., Golub, T. R., Lander, E. S. & Mesirov, J. P. 2005. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci U S A,* 102**,** 15545-50.

Sugiyama, T., Zaitseva, E. M. & Kowalczykowski, S. C. 1997. A single-stranded DNA-binding protein is needed for efficient presynaptic complex formation by the Saccharomyces cerevisiae Rad51 protein. *J Biol Chem,* 272**,** 7940-5.

Sulkowski, P. L., Corso, C. D., Robinson, N. D., Scanlon, S. E., Purshouse, K. R., Bai, H., Liu, Y., Sundaram, R. K., Hegan, D. C., Fons, N. R., Breuer, G. A., Song, Y., Mishra-Gorur, K., De Feyter, H. M., de Graaf, R. A., Surovtseva, Y. V., Kachman, M., Halene, S., Gunel, M., Glazer, P. M. & Bindra, R. S. 2017. 2-Hydroxyglutarate produced by neomorphic IDH mutations suppresses homologous recombination and induces PARP inhibitor sensitivity. *Sci Transl Med,* 9.

Sun, D., Luo, M., Jeong, M., Rodriguez, B., Xia, Z., Hannah, R., Wang, H., Le, T., Faull, K. F., Chen, R., Gu, H., Bock, C., Meissner, A., Gottgens, B., Darlington, G. J., Li, W. & Goodell, M. A. 2014a. Epigenomic profiling of young and aged HSCs reveals concerted changes during aging that reinforce self-renewal. *Cell Stem Cell,* 14**,** 673-88.

Sun, J., Ramos, A., Chapman, B., Johnnidis, J. B., Le, L., Ho, Y. J., Klein, A., Hofmann, O. & Camargo, F. D. 2014b. Clonal dynamics of native haematopoiesis. *Nature,* 514**,** 322-7.

Symington, L. S. & Gautier, J. 2011. Double-strand break end resection and repair pathway choice. *Annu Rev Genet,* 45**,** 247-71.

Tahiliani, M., Koh, K. P., Shen, Y., Pastor, W. A., Bandukwala, H., Brudno, Y., Agarwal, S., Iyer, L. M., Liu, D. R., Aravind, L. & Rao, A. 2009. Conversion of 5-methylcytosine to 5-hydroxymethylcytosine in mammalian DNA by MLL partner TET1. *Science,* 324**,** 930-5.

Takizawa, H., Regoes, R. R., Boddupalli, C. S., Bonhoeffer, S. & Manz, M. G. 2011. Dynamic variation in cycling of hematopoietic stem cells in steady state and inflammation. *J Exp Med,* 208**,** 273-84.

Tamoutounour, S., Guilliams, M., Montanana Sanchis, F., Liu, H., Terhorst, D., Malosse, C., Pollet, E., Ardouin, L., Luche, H., Sanchez, C., Dalod, M., Malissen, B. & Henri, S. 2013. Origins and functional specialization of macrophages and of conventional and monocyte-derived dendritic cells in mouse skin. *Immunity,* 39**,** 925-38.

Tarasov, A., Vilella, A. J., Cuppen, E., Nijman, I. J. & Prins, P. 2015. Sambamba: fast processing of NGS alignment formats. *Bioinformatics,* 31**,** 2032-4.

Thol, F., Weissinger, E. M., Krauter, J., Wagner, K., Damm, F., Wichmann, M., Gohring, G., Schumann, C., Bug, G., Ottmann, O., Hofmann, W. K., Schlegelberger, B., Ganser, A. & Heuser, M. 2010. IDH1 mutations in patients with myelodysplastic syndromes are associated with an unfavorable prognosis. *Haematologica,* 95**,** 1668-74.

Thomas, G. D., Hanna, R. N., Vasudevan, N. T., Hamers, A. A., Romanoski, C. E., McArdle, S., Ross, K. D., Blatchley, A., Yoakum, D., Hamilton, B. A., Mikulski, Z., Jain, M. K., Glass, C. K. & Hedrick, C. C. 2016. Deleting an Nr4a1 Super-Enhancer Subdomain Ablates Ly6C(low) Monocytes while Preserving Macrophage Gene Function. *Immunity,* 45**,** 975-987.

Till, J. E. & McCulloch, E. A. 1961. A direct measurement of the radiation sensitivity of normal mouse bone marrow cells. 1961. *Radiat Res,* 178**,** Av3-7.

Till, J. E., McCulloch, E. A. & Siminovitch, L. 1964. A STOCHASTIC MODEL OF STEM CELL PROLIFERATION, BASED ON THE GROWTH OF SPLEEN COLONY-FORMING CELLS. *Proc Natl Acad Sci U S A,* 51**,** 29-36.

Tonjes, M., Barbus, S., Park, Y. J., Wang, W., Schlotter, M., Lindroth, A. M., Pleier, S. V., Bai, A. H. C., Karra, D., Piro, R. M., Felsberg, J., Addington, A., Lemke, D., Weibrecht, I., Hovestadt, V., Rolli, C. G., Campos, B., Turcan, S., Sturm, D., Witt, H., Chan, T. A., Herold-Mende, C., Kemkemer, R., Konig, R., Schmidt, K., Hull, W. E., Pfister, S. M., Jugold, M., Hutson, S. M., Plass, C., Okun, J. G., Reifenberger, G., Lichter, P. &

Radlwimmer, B. 2013. BCAT1 promotes cell proliferation through amino acid catabolism in gliomas carrying wild-type IDH1. *Nat Med,* 19**,** 901-908.

Trowbridge, J. J., Snow, J. W., Kim, J. & Orkin, S. H. 2009. DNA methyltransferase 1 is essential for and uniquely regulates hematopoietic stem and progenitor cells. *Cell Stem Cell,* 5**,** 442-9.

Tubbs, A. & Nussenzweig, A. 2017. Endogenous DNA Damage as a Source of Genomic Instability in Cancer. *Cell,* 168**,** 644-656.

Turcan, S., Fabius, A. W., Borodovsky, A., Pedraza, A., Brennan, C., Huse, J., Viale, A., Riggins, G. J. & Chan, T. A. 2013. Efficient induction of differentiation and growth inhibition in IDH1 mutant glioma cells by the DNMT Inhibitor Decitabine. *Oncotarget,* 4**,** 1729-36.

Turcan, S., Rohle, D., Goenka, A., Walsh, L. A., Fang, F., Yilmaz, E., Campos, C., Fabius, A. W., Lu, C., Ward, P. S., Thompson, C. B., Kaufman, A., Guryanova, O., Levine, R., Heguy, A., Viale, A., Morris, L. G., Huse, J. T., Mellinghoff, I. K. & Chan, T. A. 2012. IDH1 mutation is sufficient to establish the glioma hypermethylator phenotype. *Nature,* 483**,** 479-83.

van der Heijden, T., Modesti, M., Hage, S., Kanaar, R., Wyman, C. & Dekker, C. 2008. Homologous recombination in real time: DNA strand exchange by RecA. *Mol Cell,* 30**,** 530-8.

Velten, L., Haas, S. F., Raffel, S., Blaszkiewicz, S., Islam, S., Hennig, B. P., Hirche, C., Lutz, C., Buss, E. C., Nowak, D., Boch, T., Hofmann, W. K., Ho, A. D., Huber, W., Trumpp, A., Essers, M. A. & Steinmetz, L. M. 2017. Human haematopoietic stem cell lineage commitment is a continuous process. *Nat Cell Biol,* 19**,** 271-281.

Villani, A. C., Satija, R., Reynolds, G., Sarkizova, S., Shekhar, K., Fletcher, J., Griesbeck, M., Butler, A., Zheng, S., Lazo, S., Jardine, L., Dixon, D., Stephenson, E., Nilsson, E., Grundberg, I., McDonald, D., Filby, A., Li, W., De Jager, P. L., Rozenblatt-Rosen, O., Lane, A. A., Haniffa, M., Regev, A. & Hacohen, N. 2017. Single-cell RNA-seq reveals new types of human blood dendritic cells, monocytes, and progenitors. *Science,* 356.

Waas, M., Snarrenberg, S. T., Littrell, J., Jones Lipinski, R. A., Hansen, P. A., Corbett, J. A. & Gundry, R. L. 2019. SurfaceGenie: A web-based application for prioritizing cell-type specific marker candidates. *bioRxiv*.

Waddington, C. H. 1942. The epigenotype. 1942. *Int J Epidemiol,* 41**,** 10-3.

Walter, D., Lier, A., Geiselhart, A., Thalheimer, F. B., Huntscha, S., Sobotta, M. C., Moehrle, B., Brocks, D., Bayindir, I., Kaschutnig, P., Muedder, K., Klein, C., Jauch, A., Schroeder, T., Geiger, H., Dick, T. P., Holland-Letz, T., Schmezer, P., Lane, S. W., Rieger, M. A., Essers, M. A., Williams, D. A., Trumpp, A. & Milsom, M. D. 2015. Exit from dormancy provokes DNA-damage-induced attrition in haematopoietic stem cells. *Nature,* 520**,** 549-52.

Wang, J., Sun, Q., Morita, Y., Jiang, H., Gross, A., Lechel, A., Hildner, K., Guachalla, L. M., Gompf, A., Hartmann, D., Schambach, A., Wuestefeld, T., Dauch, D., Schrezenmeier, H., Hofmann, W. K., Nakauchi, H., Ju, Z., Kestler, H. A., Zender, L. & Rudolph, K. L. 2012. A differentiation checkpoint limits hematopoietic stem cell self-renewal in response to DNA damage. *Cell,* 148**,** 1001-14.

Wang, J. C. 2002. Cellular roles of DNA topoisomerases: a molecular perspective. *Nat Rev Mol Cell Biol,* 3**,** 430-40.

Wang, P., Wu, J., Ma, S., Zhang, L., Yao, J., Hoadley, K. A., Wilkerson, M. D., Perou, C. M., Guan, K. L., Ye, D. & Xiong, Y. 2015. Oncometabolite D-2-Hydroxyglutarate Inhibits ALKBH DNA Repair Enzymes and Sensitizes IDH Mutant Cells to Alkylating Agents. *Cell Rep,* 13**,** 2353-61.

Ward, P. S., Patel, J., Wise, D. R., Abdel-Wahab, O., Bennett, B. D., Coller, H. A., Cross, J. R., Fantin, V. R., Hedvat, C. V., Perl, A. E., Rabinowitz, J. D., Carroll, M., Su, S. M., Sharp, K. A., Levine, R. L. & Thompson, C. B. 2010. The common feature of leukemia-associated IDH1 and IDH2 mutations is a neomorphic enzyme activity converting alpha-ketoglutarate to 2-hydroxyglutarate. *Cancer Cell,* 17**,** 225-34.

Wasylishen, A. R. & Penn, L. Z. 2010. Myc: the beauty and the beast. *Genes Cancer,* 1**,** 532-41.

Watcham, S., Kucinski, I. & Gottgens, B. 2019. New insights into hematopoietic differentiation landscapes from single-cell RNA sequencing. *Blood,* 133**,** 1415-1426.

Welch, J. S., Ley, T. J., Link, D. C., Miller, C. A., Larson, D. E., Koboldt, D. C., Wartman, L. D., Lamprecht, T. L., Liu, F., Xia, J., Kandoth, C., Fulton, R. S., McLellan, M. D., Dooling, D. J., Wallis, J. W., Chen, K., Harris, C. C., Schmidt, H. K., Kalicki-Veizer, J. M., Lu, C., Zhang, Q., Lin, L., O'Laughlin, M. D., McMichael, J. F., Delehaunty, K. D., Fulton, L. A., Magrini, V. J., McGrath, S. D., Demeter, R. T., Vickery, T. L., Hundal, J., Cook, L. L., Swift, G. W., Reed, J. P., Alldredge, P. A., Wylie, T. N., Walker, J. R., Watson, M. A., Heath, S.

E., Shannon, W. D., Varghese, N., Nagarajan, R., Payton, J. E., Baty, J. D., Kulkarni, S., Klco, J. M., Tomasson, M. H., Westervelt, P., Walter, M. J., Graubert, T. A., DiPersio, J. F., Ding, L., Mardis, E. R. & Wilson, R. K. 2012. The origin and evolution of mutations in acute myeloid leukemia. *Cell,* 150**,** 264-78.

Wen, B., Wu, H., Shinkai, Y., Irizarry, R. A. & Feinberg, A. P. 2009. Large histone H3 lysine 9 dimethylated chromatin blocks distinguish differentiated from embryonic stem cells. *Nat Genet,* 41**,** 246-50.

Weston, V. J., Oldreive, C. E., Skowronska, A., Oscier, D. G., Pratt, G., Dyer, M. J., Smith, G., Powell, J. E., Rudzki, Z., Kearns, P., Moss, P. A., Taylor, A. M. & Stankovic, T. 2010. The PARP inhibitor olaparib induces significant killing of ATM-deficient lymphoid tumor cells in vitro and in vivo. *Blood,* 116**,** 4578-87.

Wiktor-Jedrzejczak, W. & Gordon, S. 1996. Cytokine regulation of the macrophage (M phi) system studied using the colony stimulating factor-1-deficient op/op mouse. *Physiol Rev,* 76**,** 927-47.

Wilson, A., Laurenti, E., Oser, G., van der Wath, R. C., Blanco-Bose, W., Jaworski, M., Offner, S., Dunant, C. F., Eshkind, L., Bockamp, E., Lio, P., Macdonald, H. R. & Trumpp, A. 2008. Hematopoietic stem cells reversibly switch from dormancy to self-renewal during homeostasis and repair. *Cell,* 135**,** 1118-29.

Wilson, N. K. & Göttgens, B. 2018. Single-Cell Sequencing in Normal and Malignant Hematopoiesis. *HemaSphere,* 2**,** e34.

Wooster, R., Bignell, G., Lancaster, J., Swift, S., Seal, S., Mangion, J., Collins, N., Gregory, S., Gumbs, C. & Micklem, G. 1995. Identification of the breast cancer susceptibility gene BRCA2. *Nature,* 378**,** 789-92.

Wu, X. & Zhang, Y. 2017. TET-mediated active DNA demethylation: mechanism, function and beyond. *Nat Rev Genet.*

Wynn, T. A., Chawla, A. & Pollard, J. W. 2013. Macrophage biology in development, homeostasis and disease. *Nature,* 496**,** 445-55.

Xie, M., Lu, C., Wang, J., McLellan, M. D., Johnson, K. J., Wendl, M. C., McMichael, J. F., Schmidt, H. K., Yellapantula, V., Miller, C. A., Ozenberger, B. A., Welch, J. S., Link, D. C., Walter, M. J., Mardis, E. R., Dipersio, J. F., Chen, F., Wilson, R. K., Ley, T. J. & Ding, L. 2014. Age-related mutations associated with clonal hematopoietic expansion and malignancies. *Nat Med,* 20**,** 1472-8.

Xu, W., Yang, H., Liu, Y., Yang, Y., Wang, P., Kim, S. H., Ito, S., Yang, C., Wang, P., Xiao, M. T., Liu, L. X., Jiang, W. Q., Liu, J., Zhang, J. Y., Wang, B., Frye, S., Zhang, Y., Xu, Y. H., Lei, Q. Y., Guan, K. L., Zhao, S. M. & Xiong, Y. 2011. Oncometabolite 2-hydroxyglutarate is a competitive inhibitor of alpha-ketoglutarate-dependent dioxygenases. *Cancer Cell,* 19**,** 17-30.

Yamamoto, R., Morita, Y., Ooehara, J., Hamanaka, S., Onodera, M., Rudolph, K. L., Ema, H. & Nakauchi, H. 2013. Clonal analysis unveils self-renewing lineage-restricted progenitors generated directly from hematopoietic stem cells. *Cell,* 154**,** 1112-26.

Yanez, A., Coetzee, S. G., Olsson, A., Muench, D. E., Berman, B. P., Hazelett, D. J., Salomonis, N., Grimes, H. L. & Goodridge, H. S. 2017. Granulocyte-Monocyte Progenitors and Monocyte-Dendritic Cell Progenitors Independently Produce Functionally Distinct Monocytes. *Immunity,* 47**,** 890-902 e4.

Yanez, A., Ng, M. Y., Hassanzadeh-Kiabi, N. & Goodridge, H. S. 2015. IRF8 acts in lineage-committed rather than oligopotent progenitors to control neutrophil vs monocyte production. *Blood,* 125**,** 1452-9.

Yang, H., Li, Q., Fan, J., Holloman, W. K. & Pavletich, N. P. 2005a. The BRCA2 homologue Brh2 nucleates RAD51 filament formation at a dsDNA-ssDNA junction. *Nature,* 433**,** 653-7.

Yang, H., Ye, D., Guan, K. L. & Xiong, Y. 2012. IDH1 and IDH2 mutations in tumorigenesis: mechanistic insights and clinical perspectives. *Clin Cancer Res,* 18**,** 5562-71.

Yang, J., Zhang, L., Yu, C., Yang, X. F. & Wang, H. 2014. Monocyte and macrophage differentiation: circulation inflammatory monocyte as biomarker for inflammatory diseases. *Biomark Res,* 2**,** 1.

Yang, L., Bryder, D., Adolfsson, J., Nygren, J., Mansson, R., Sigvardsson, M. & Jacobsen, S. E. 2005b. Identification of Lin(-)Sca1(+)kit(+)CD34(+)Flt3- short-term hematopoietic stem cells capable of rapidly reconstituting and rescuing myeloablated transplant recipients. *Blood,* 105**,** 2717-23.

Yang, L., Rodriguez, B., Mayle, A., Park, H. J., Lin, X., Luo, M., Jeong, M., Curry, C. V., Kim, S. B., Ruau, D., Zhang, X., Zhou, T., Zhou, M., Rebel, V. I., Challen, G. A., Gottgens, B., Lee, J. S., Rau, R., Li, W. & Goodell, M. A. 2016. DNMT3A Loss Drives Enhancer Hypomethylation in FLT3-ITD-Associated Leukemias. *Cancer Cell,* 29**,** 922-34.

Yona, S., Kim, K. W., Wolf, Y., Mildner, A., Varol, D., Breker, M., Strauss-Ayali, D., Viukov, S., Guilliams, M., Misharin, A., Hume, D. A., Perlman, H., Malissen, B., Zelzer, E. & Jung, S. 2013. Fate mapping reveals origins and dynamics of monocytes and tissue macrophages under homeostasis. *Immunity,* 38**,** 79-91.

Yoneyama, M. & Fujita, T. 2009. RNA recognition and signal transduction by RIG-I-like receptors. *Immunol Rev,* 227**,** 54-65.

Young, L. C., McDonald, D. W. & Hendzel, M. J. 2013. Kdm4b histone demethylase is a DNA damage response protein and confers a survival advantage following gamma-irradiation. *J Biol Chem,* 288**,** 21376-88.

Yu, V. W., Yusuf, R. Z., Oki, T., Wu, J., Saez, B., Wang, X., Cook, C., Baryawno, N., Ziller, M. J., Lee, E., Gu, H., Meissner, A., Lin, C. P., Kharchenko, P. V. & Scadden, D. T. 2016. Epigenetic Memory Underlies Cell-Autonomous Heterogeneous Behavior of Hematopoietic Stem Cells. *Cell,* 167**,** 1310-1322 e17.

Zhang, Q., Zhao, K., Shen, Q., Han, Y., Gu, Y., Li, X., Zhao, D., Liu, Y., Wang, C., Zhang, X., Su, X., Liu, J., Ge, W., Levine, R. L., Li, N. & Cao, X. 2015. Tet2 is required to resolve inflammation by recruiting Hdac2 to specifically repress IL-6. *Nature,* 525**,** 389-393.

Zhang, X., Su, J., Jeong, M., Ko, M., Huang, Y., Park, H. J., Guzman, A., Lei, Y., Huang, Y. H., Rao, A., Li, W. & Goodell, M. A. 2016. DNMT3A and TET2 compete and cooperate to repress lineage-specific transcription factors in hematopoietic stem cells. *Nat Genet,* 48**,** 1014-23.

Zhao, X. D., Han, X., Chew, J. L., Liu, J., Chiu, K. P., Choo, A., Orlov, Y. L., Sung, W. K., Shahab, A., Kuznetsov, V. A., Bourque, G., Oh, S., Ruan, Y., Ng, H. H. & Wei, C. L. 2007. Whole-genome mapping of histone H3 Lys4 and 27 trimethylations reveals distinct genomic compartments in human embryonic stem cells. *Cell Stem Cell,* 1**,** 286-98.

Zheng, H. & Xie, W. 2019. The role of 3D genome organization in development and cell differentiation. *Nat Rev Mol Cell Biol,* 20**,** 535-550.

Ziller, M. J., Gu, H., Muller, F., Donaghey, J., Tsai, L. T., Kohlbacher, O., De Jager, P. L., Rosen, E. D., Bennett, D. A., Bernstein, B. E., Gnirke, A. & Meissner, A. 2013. Charting a dynamic DNA methylation landscape of the human genome. *Nature,* 500**,** 477-81.

Zink, F., Stacey, S. N., Norddahl, G. L., Frigge, M. L., Magnusson, O. T., Jonsdottir, I., Thorgeirsson, T. E., Sigurdsson, A., Gudjonsson, S. A., Gudmundsson, J., Jonasson, J. G., Tryggvadottir, L., Jonsson, T., Helgason, A., Gylfason, A., Sulem, P., Rafnar, T., Thorsteinsdottir, U., Gudbjartsson, D. F., Masson, G., Kong, A. & Stefansson, K. 2017. Clonal hematopoiesis, with and without candidate driver mutations, is common in the elderly. *Blood,* 130**,** 742-752.

List of Figures

List of Tables

Abbreviations

A

A	Adenosine
ACK	Ammonium-chloride-potassium
AML	Acute myeloid leukemia
APC	Allophycocyanin
APC-Cy7	Allophycocyanin-cyanine 7
ATM	Ataxia telangiectasia mutated

B

BM	Bone marrow
bp	base pair
BV605	Brilliant violett 605
BV711	Brilliant violett 711
BV785	Brilliant violett 785

C

C	Cytosine
CD	Cluster of differentiation
cDC	Conventional dendritic cell
cDNA	Complementary DNA
CDP	Common dendritic cell progenitor
CFU	Colony forming unit
CHIP	Clonal hematopoiesis of indeterminate potential
CLP	Common lymphoid progenitor
CMML	Chronic myelomonocytic leukemia
cMoP	Common monocyte progenitor
CML	Chronic myeloid leukemia
CMP	Common myeloid progenitor
CpG	Cytosine phosphate guanine
Cre	Cre recombinase
CRISPR	Clustered regularly interspaced short palindromic repeats
CSF1	Colony stimulating factor 1
CTRL	Control
Cy7	Cyanine 7

D

D2HG	D-2-Hydroxyglutarate
DC	Dendritic cell
DC1 / DC2 / DC3	Diffusion component 1 / 2 / 3
ddH_2O	Double distilled water
DDR^2	DNA damage response
DE	Differentially expressed
DKFZ	German Cancer Research Center
DMR	Differentially methylated region
DNA	Deoxyribonucleic acid
DNMT	DNA methyltransferase
DNMT3A / Dnmt3a	DNA methyltransferase 3 alpha
DSB	Double-stranded break
dsDNA	Double-stranded DNA

E

EDTA	Ethylendiaminetetraacetic acid
EdU	5-Ethynyl-2´-deoxyuridine
Em	Emission
ERV	Endogenous retrovirus
Ex	Excitation
EYFP	Enhanced yellow fluorescent protein
et.al	Et alii

F

FACS	Fluorescence-activated cell sorting
FBS	Fetal bovine serum
FCS	Fetal calf serum
fc	Fold change
fdr	False discovery rate

G

g	Gravity
G	Guanosine
GEM	Gel Bead-in Emulsion
GFP	Green fluorescent protein
GG-NER	Global genomic nucleotide excision repair
GMP	Granulocyte-monocyte progenitor
GO	Gene ontology
GP	Granulocyte progenitor
GSEA	Gene set enrichment analysis
Gy	Gray

H

h	Hour
H	Histidine
H2O	Water
H3K27ac	Histone H3 lysine 27 acetylation
H3K27me3	Histone H3 lysine 27 trimethylation
H3K4me1	Histone H3 lysine 4 monomethylation
H3K9me2	Histone H3 lysine 9 dimethylation
H3K9me3	Histone H3 lysine 9 trimethylation
HR	Homologous recombination
HSC	Hematopoietic stem cell
HSPC	Hematopoietic stem and progenitor cell

I

IDH1 / Idh1	Isocitrate dehydrogenase 1
IDH2 / Idh2	Isocitrate dehydrogenase 2
IFN	Interferon
IL	Interleukin
IRF	Interferon regulatory factor
IU	Infectious unit

K

K	1,000
KDM	Histone lysine demethylase
kg	Kilogram
kNN	K nearest neighbour

L

LDMNC	Low density mononuclear cells
Lin	Lineage
LMPP	Lymphoid-primed multipotent progenitor

loess	locally estimated scatter plot smoothing
log	Logarithm
LPS	Lipopolysaccharide
LS-K	Lineage$^-$ Sca1$^-$ cKit$^+$ cells
LSK	Lineage$^-$ Sca1$^+$ cKit$^+$ cells
LT-HSC	Long-term hematopoietic stem cell

M

M	1,000,000
M	Molar
M-CSF	Macrophage colony-stimulating factor
MDP	Monocyte-dendritic cell progenitor
MDS	Myelodysplastic syndrome
MEP	Megakaryocyte-erythroid progenitor
MFI	Mean fluorescence intensity
mg	Miligram
min	Minute
ml	Mililiter
mM	Milimolar
moDC	Monocyte-derived dendritic cell
MoDCP	Monocyte / dendritic cell primed progenitor
MOI	Multiplicity of infection
MPN	Myeloproliferative neoplasm
MPP	Multipotent progenitor
MPS	Mononuclear phagocyte system
MSigDB	Molecular Signatures Database

N

NADP+	Nicotinamide adenine dinucleotide phosphate (oxidized form)
NADPH	Nicotinamide adenine dinucleotide phosphate (reduced form)
NeuP	Neutrophil primed progenitor
ng	Nanogram
NK-cell	Natural killer cell
NHEJ	Non-homologous end joining
nm	Nanometer

P

p	p-value
padj	Adjusted p-value
PARP	Poly (ADP-Ribose) polymerase 1
PBS	Phosphate buffered saline
PBS/FCS	PBS + 2 % FCS
PC	Principle component
PCA	Principle component analysis
PCR	Poymerase chain reaction
pDC	Plasmacytoid dedritic cell
PE	Phycoerythrin
PE-Cy5	Phycoerythrin-cyanine 5
PE-Cy7	Phycoerythrin-cyanine 7
Pen/Strep	Penicillin-streptomycin
pI:pC	Polyinosinic-polycytidylic acid
PRR	Pattern recognition receptor

Q

QC	Quality control
qRT-PCR	Quantitative real-time polymerase chain reaction

R

rlog	Regularized logarithm
RNA	Ribonucleic acid
RNA-seq	RNA sequencing
ROS	Reactive oxygen species
rpm	Revolutions per minute
RT	Room temperature

S

s	Second
SAM	S-Adenosyl methionine
Sca1	Stem cell antigen 1
Scl	Stem cell leukemia
scRNA-seq	Single-cell RNA sequencing
SFFV	Spleen focus-forming virus
siRNA	Small / Short interfering RNA
SNN	Shared nearest neighbour
SSB	Single-stranded breaks
ssDNA	Single-stranded DNA
ST-HSC	Short-term hematopoietic stem cell

T

T	Thymidine
TAD	Topology associating domain
T-ALL	T-cell acute lymphoblastic leukemia
TBM	Total bone marrow
TCA	Tricarboxylic acid
TET	Ten-eleven translocation methylcytosine dioxygenase
TF	Transcription factor
TLR	Toll-like receptor
TNFα	Tumor necrosis factor alpha

U

U	Unit
UMAP	Uniform manifold approximation and projection
UMI	Unique molecular identifier

W

WT	Wildtype

%	Percent
°C	Degree Celsius
5-caC	5-Carboxylcytosine
5-fC	5-Formylcytosine
5-hmC	5-Hydroxymethylcytosine
5-mC	5-Methylcytosine
7AAD	7-Aminoactinomycin
αKG	Alpha(α)–ketoglutarate
µg	Microgram
µl	Microliter
µm	Micrometer
µM	Micromolar

Contributions

This project would not have been possible without the help and contribution of many people:

PD Dr. Daniel Lipka supervised the project.

Dr. Mark Hartmann, Dr. Sina Stäble, Mariam Hakobyan, and **Maximilian Schönung** provided help with the preparation of bone marrow cells, FACS analysis and sorting, as well as injections and transplantations of mice. **Oliver Mücke** performed genotyping of mice.

Dr. Sina Stäble and **Dr. Natasha Anstee** performed transplantation and analysis of LT-HSC transplanted *DNMT3A*-R882H single-mutant mice.

Dr. Mark Hartmann, **Maximilian Schönung** and **Katharina Bauer** generated 10X single-cell RNA-seq libraries. **Abdelrahman Mahmoud** performed processing of single-cell RNA-seq data and helped with data analysis. Single-cell RNA-seq cluster were annotated with the help of **Dr. Mark Hartmann** and **Dr. Simon Haas.**

Maximilian Schönung helped with Smart-seq2 library preparation. Initial data processing of bulk RNA-seq data was performed by the **Omics IT and Data Management Core Facility** of the DKFZ in collaboration with **Stephen Krämer**.

Dr. Stefan Pusch kindly provided *Idh1*-R132H mice and helped with D2HG measurements.

Prof. Dr. Carsten Müller-Tidow kindly provided *DNMT3A*-R882H mice.

Mice were held at and taken care of by the **Center for Preclinical Research** and the **Central Animal Laboratory** of the DKFZ**.**

Olaparib experiments and CellTiter-Blue readouts were performed by **Dr. Ali Bakr**.

Comet assays were performed by **Reinhard Gliniorz** in collaboration with **Dr. Peter Schmezer** and **PD Dr. Odilia Popanda**.

Sequencing was performed by the **Genomics and Proteomics Core Facility** of the DKFZ**.**

FACS sorting was supported by the **Flow Cytometry Core Facility** of the DKFZ**.**

Acknowledgements

I sincerely would like to thank the following people for their kind support over the years:

PD Dr. Daniel Lipka for giving me the opportunity to work in your group, your continuous support and your guidance. Thank you for always having an open door throughout the years, answering countless questions, and providing valuable input, scientific advice and fruitful discussions. As one of the earliest members of your group, you helped me to grow scientifically and over the years provided a great research and working environment that made work enjoyable. More importantly, thank you for not forgetting that there is more than a scientific site to life; thank you for occasional drinks and dinners, annual Christmas Market visits or barbecues at your house.

Dr. Michael Milsom for instructive advice, expertise and ideas that always were expedient and helped to advance the project. Moreover, thanks a lot for letting me carry out my experiments with your group.

Prof. Dr. Christoph Plass and **Prof. Dr. Andreas Trumpp** for providing a great scientific infrastructure in their divisions.

Prof. Dr. Jan Lohman, **PD Dr. Odilia Popanda**, **Dr. Michael Milsom**, and **PD Dr. Daniel Lipka** for taking the time and being part of my examination committee.

My TAC members **PD Dr. Odilia Popanda**, **Dr. Michael Milsom**, **Prof. Dr. Christoph Plass**, **PD Dr. Daniel Lipka** and **Prof. Dr. Carsten Müller-Tidow** for productive discussions during TAC meetings.

The Lipka Group members **Dr. Sina Stäble**, **Dr. Mark Hartmann**, **Maximilian Schönung**, **Mariam Hakobyan** and **Oliver Mücke** for all your support and help during experiments, countless coffee breaks and after-work beers, but most importantly, for your company and friendship.

The members of the 'Cancer Epigenomics' Division **Clarissa Feuerstein**, **Joschka Hey, Dr. Pavlo Lutsik, Alexander Kühn, Daniela Mancarella** and **Justyna Wierzbinska** for scientific help, discussions and enjoyable lunch breaks. All the remaining members of the division for a supportive atmosphere in the lab, scientific input during lab meetings and pleasant times during Happy Hours or lab excursions.
Clarissa Feuerstein and **Dr. Sina Stäble** for sharing an office for almost four years, for always having chewing gums at hand and answering all my scientific, non-scientific and especially R-related questions.

The former and present Milsom Group members **Dr. Natasha Anstee**, **Dr. Julius Gräsel**, **Megan Druce**, **Marleen Büchler-Schäff**, **Jeyan Jayarajan**, **Dr. Ruzhica Bogeska**, **Dr. Paul Kaschutnig, Dr. Ana-Matea Mikecin, Julia Knoch** (thanks for ordering all my mice) and **Melanie Ball** for all your help during early morning mouse preps and your support over the years.

All members of the divisions '**Applied Functional Genomics**', '**Translational Medical Oncology**', '**Molecular Leukemogenesis**' and '**Precision Sarcoma Research**' for making us feel welcome at TP4.

The **Cytometry Core Facility** for help, technical support and advice during sorts and the **Animal Core Facility** for taking care of mice on a daily basis.

Dr. Sina Stäble for being there since the beginning of this journey. I am deeply grateful for your support, your scientific advice and your patience to teach me all lab-related issues and techniques in the beginning. Without your help, life in the lab would have been much more difficult. Thank you for joining me in many coffee breaks or nights in front of the sorter, for having crazy nights out in town, hosting lovely dinners at your place and throwing unforgettable thesis-handing-in parties.

Dr. Mark Hartmann for all your guidance and your scientific input to the 'Inflammation' Project. Your contributions, your advice and your help always were highly appreciated and critical for the progress of the project. Thank you also for all the non-scientific advice, for making sure to drink enough beer, for having meetings with philosophers and for showing us the true 'beauty' of Frankfurt.

Maximilian Schönung for your help in the lab and for coping with all my R-related questions. Thank you for many fruitful discussions, your smart ideas, your positive attitude and your excitement for science. I always enjoyed having a coffee, a beer or a bouldering session with you.

Mariam Hakobyan for all the support during never-ending mouse preps and your decision to join the 'Inflammation' project. Your enthusiasm, curiosity and passion for the project were a great help when interpreting data or planning next experiments and I am sure that, with your help, the project will come to a successful end. Thank you for sharing a desk with me for the last couple of months and for enduring the occasional exposure to my music.

Oliver Mücke for countless genotyping of mice, for reminding us that there is more than just science and for making us smile even when times are rough.

My dearest friends **André**, **Michael** and **Felix** for all the shared times outside the lab and your support. I wish I would have made more time to spend with you, but as much more I enjoyed our joint events, holidays or evenings.

My parents **Manuela** and **Dieter**, my brother **Eric** and my **grandparents** for your adamant love, your belief in me and all your sacrifices. I will be forever grateful for your endless support.

Luisa Henkel for being my home, my shelter and my future.