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. Pia Sommerkamp

born in: Ahlen, Germany

Oral examination: 12.07.2019

Dissecting RNA biology in hematopoietic stem cells:

The long non-coding RNA *Meg3* is dispensable for hematopoiesis

and

alternative polyadenylation orchestrates hematopoietic stem cell function

Referees:
Prof. Dr. Andreas Trumpp
Prof. Dr. Jan Lohmann

für Alex

“Science, for me, gives a partial explanation for life.
In so far as it goes, it is based on fact, experience and experiment.”

Rosalind Franklin

I. Abstract

Hematopoietic stem cells (HSCs) reside at the top of a tightly regulated and hierarchically organized differentiation cascade. They are multipotent and able to self-renew, thereby ensuring life-long replenishment of mature blood cells. In-depth Omics analyses within the hematopoietic hierarchy, including analysis of HSCs and multipotent progenitor (MPP) cells, revealed not only differential expression of protein-coding genes, but also HSC-specific splicing variants and expression of long non-coding RNAs (lncRNAs) upon HSC commitment (Cabezas-Wallscheid et al., 2014; Klimmeck et al., 2014; Luo et al., 2015). Functional analyses revealed that non-coding RNAs and regulation of RNA biogenesis are crucial for HSC function and potency. In this thesis, the role of the lncRNA maternally expressed gene 3 *(Meg3)* in hematopoiesis is studied. In addition, I introduce the RNA regulatory mechanism alternative polyadenylation (APA) as a mechanism controlling HSC/MPP biology in homeostasis and during replicative stress response by extensive *in silico*, *in vitro* and *in vivo* analyses. Furthermore, I generated an inducible knockout mouse model, targeting the prominent APA regulator poly(A) binding protein 1 *(Pabpn1)*. Thereby, I aimed to dissect the general role of lncRNAs and RNA regulatory mechanisms in hematopoiesis.

Project 1: The role of the lncRNA *Meg3* in HSCs

The tumor suppressor lncRNA *Meg3* is encoded in the imprinted *Dlk1-Meg3* locus and expressed from the maternally inherited allele (Zhou et al., 2012). We and others found *Meg3* to be highly and specifically expressed in the HSC compartment compared to MPP cells. In this thesis, I crossed an inducible *Meg3* flox mouse model (Klibanski et al., unpublished) to MxCre mice, generating MxCre *Meg3*$^{\text{mat flox/pat wt}}$ mice. Cre-induction deleted the maternal allele in the hematopoietic compartment and thereby completely abrogated the expression of *Meg3* and its associated miRNA cluster in HSCs. Extensive *in vivo* and *in vitro* analyses of adult mice harboring a *Meg3*-deficient blood system surprisingly did not reveal any impairment of hematopoiesis or stem cell function. In addition, I performed serial transplantation assays to investigate the functional capacity of *Meg3*-deficient HSCs. Again, knockout cells did not exhibit altered blood contribution, even upon tertiary transplantation. Imprinting of the *Dlk1-Meg3* locus has recently been reported to regulate fetal liver HSC function (Qian et al., 2016). To analyze effects of the hematopoiesis-specific *Meg3* knockout in the developing embryo, we generated VavCre *Meg3*$^{\text{mat flox/pat wt}}$ mice. Cre+ offspring were born and developed normally. In-depth analysis of adult animals revealed loss of *Meg3* expression in HSCs, but again no hematopoietic impairments were detected. Next, I performed interferon-mediated stimulation in MxCre *Meg3*$^{\text{mat flox/pat wt}}$ mice. During both activation and recovery phase, *Meg3*-deficient

adult HSCs responded highly similar compared to controls. Taken together, my work shows that the highly expressed, imprinted lncRNA *Meg3* is dispensable for the function of HSCs during adulthood and embryonic development. In the adult system, loss of *Meg3* does not impair the performance in serial reconstitution assays or response to stress mediators. (Sommerkamp et al., 2019)

Project 2: HSC function, differentiation and activation are regulated by APA

The majority of mammalian genes have multiple different polyadenylation sites. The RNA editing mechanism APA controls the selection of these sites, thereby altering 3'-UTR length and isoform expression. Thus, APA modulates RNA stability, localization, protein output and even protein localization (Di Giammartino et al., 2011; Tian and Manley, 2017). So far, the role of APA in the regulation of the adult HSC/MPP compartment has not been studied. In this thesis, I show that the APA regulator *Pabpn1* is essential for HSCs, as knockdown (KD) of *Pabpn1* led to decreased HSC function *in vitro* and *in vivo*. To analyze the prevalence of APA at the top of the hematopoietic hierarchy, I established an ultra-low input 3'-Seq approach and performed analysis of HSCs, MPP cells and HSCs activated by inflammation. Bioinformatic analysis revealed dynamic APA patterns in numerous genes between HSCs and MPPs as well as during inflammation-induced HSC stress response. We observed global 3'-UTR shortening both upon HSC differentiation towards MPPs and HSC activation. Further, 3'-Seq analysis of *Pabpn1* KD cells revealed that PABPN1 regulates APA in the HSC/MPP compartment. We observed an APA-mediated glutaminase *(Gls)* isoform switch upon exit of HSCs from quiescence. *Gls* isoform switching led to enhanced relative expression of the highly active GLS GAC isoform and overall increased GLS protein levels. In line, small molecule-mediated inhibition of glutaminolysis *in vitro* enhanced HSC maintenance by limiting proliferation. I could show that *Gls* isoform switching leading to increased glutaminolysis is mediated by the APA regulator NUDT21 and in turn is required for proper HSC function. KD of *Nudt21* led to inhibition of *Gls* isoform switching, impaired HSC function and a partial block in HSC differentiation. In summary, my results install differential employment of APA and associated glutamine metabolism adaptations as novel layers in the regulation of the HSC-controlled hematopoietic hierarchy. (Sommerkamp et al., under revision)

Project 3: Generation of *Pabpn1*^{flox} mice

To enable in-depth *in vivo* analysis of the role of APA in different tissues and disease settings in the future, we generated an inducible *Pabpn1*^{flox} mouse model. Here, we used the *Easi*-CRISPR approach (Miura et al., 2018; Quadros et al., 2017) to generate transgenic animals. I performed extensive *in vitro* testing of various guide RNAs (gRNAs) to optimize recombination efficiency *in vivo*. Two different crRNA-tracrRNA:Cas9 complexes targeting upstream and

downstream genomic regions, respectively, were injected into one of the pronuclei of zygotes together with a long single-stranded DNA (ssDNA) template. By extensive genotyping, I identified 3 of 17 offspring animals to be correctly targeted. Homozygous offspring mice were successfully bred and can be used in the future to determine functionality of the *Pabpn1*^{flox} mouse model and to functionally asses the role of APA in different biological settings.

II. Zusammenfassung

Hämatopoetische Stammzellen (HSZs) befinden sich an der Spitze einer präzise regulierten, hierarchisch organisierten Differenzierungskaskade. Sie sind multipotent und können sich selbst erneuern, wodurch die lebenslange Generierung reifer Blutzellen sichergestellt wird. Detaillierte „Omics"-Analysen innerhalb der hämatopoetischen Hierarchie, einschließlich Analysen von HSZs und multipotenten Vorläuferzellen (MPPs), zeigten nicht nur eine differenzielle Expression von proteinkodierenden Genen, sondern auch von Isoformen die durch alternatives Spleißen entstehen und von langen nicht-kodierenden RNAs (lncRNAs) (Cabezas-Wallscheid et al., 2014; Klimmeck et al., 2014; Luo et al., 2015). Daher scheint die Expression von nicht-kodierenden RNAs und die Regulation der RNA-Biogenese entscheidend für die HSZ-Funktion und Potenz zu sein. In der vorliegenden Doktorarbeit habe ich die Rolle der lncRNA „maternally expressed gene 3" *(Meg3)* in der Hämatopoese analysiert. Darüber hinaus habe ich die Rolle des RNA-Regulierungsmechanismus „alternative Polyadenylierung" (APA) im HSZ/MPP-Kompartiment studiert und ein induzierbares Knockout-Mausmodell des APA-Regulators „poly(A) binding protein 1" *(Pabpn1)* generiert. Ziel meiner Arbeit war es, ein besseres Verständnis der Rolle von lncRNAs und RNA-Regulationsmechanismen im Allgemeinen in der Hämatopoese zu entwickeln.

Projekt 1: Die Rolle der lncRNA *Meg3* in HSZs

Die Tumorsuppressor lncRNA *Meg3* ist im „imprinted" *Dlk1-Meg3*-Locus kodiert und wird vom mütterlich vererbten Allel exprimiert (Zhou et al., 2012). Wir und andere Forschungsgruppen fanden heraus, dass *Meg3*, im Vergleich zu MPP-Zellen, sehr stark und spezifisch im HSZ-Kompartiment exprimiert wird. In dieser Doktorarbeit habe ich ein induzierbares *Meg3* flox-Mausmodell (Klibanski et al., unveröffentlicht) verwendet und mit MxCre Tieren verpaart. Dadurch wurden MxCre *Meg3* $^{mat\ flox/pat\ wt}$ Mäuse generiert. Die Cre-Induktion, welche zur Deletion des mütterlichen Allels führt, war ausreichend, um die *Meg3*-Expression und die Expression des zugehörigen miRNA-Clusters in HSZs vollständig zu hemmen. Umfangreiche *in vivo* und *in vitro* Analysen von adulten Mäusen mit einem *Meg3*-defizienten Blutsystem zeigten keine Beeinträchtigungen der Hämatopoese und Stammzellfunktion. Darüber hinaus habe ich serielle Transplantationsexperimente durchgeführt, um die Funktionsfähigkeit von *Meg3*-defizienten HSZs zu untersuchen. Auch hier konnte ich keine Beeinträchtigungen von Knockout-Zellen feststellen, selbst nicht in tertiären Transplantationsexperimenten. Vor Kurzem wurde gezeigt, dass korrektes „Imprinting" des *Dlk1-Meg3*-Locus essentiell für die Funktion embryonaler Leber-HSZ ist (Qian et al., 2016). Um die Auswirkungen eines Hämatopoese-spezifischen *Meg3*-Knockout auf den sich entwickelnden Embryo zu

analysieren, habe ich VavCre *Meg3* $^{mat\ flox/pat\ wt}$ Mäuse generiert. Cre+ Nachkommen wurden geboren und entwickelten sich normal. Eine eingehende Analyse der erwachsenen Tiere ergab, dass der Verlust der *Meg3* Expression in HSZs auch in diesem Kontext nicht zu einer hämatopoetischen Beeinträchtigung führt. Darüber hinaus habe ich interferonvermittelte Stimulationsexperimente in MxCre *Meg3* $^{mat\ flox/pat\ wt}$ Mäusen durchgeführt, die zu einer HSZ-Aktivierung führen. Während der Aktivierungs- und Erholungsphase reagierten *Meg3*-defiziente adulte HSZs sehr ähnlich wie Kontroll-HSZs. Zusammenfassend konnte ich zeigen, dass die stark exprimierte „imprinted" lncRNA *Meg3* für die Funktion von HSZs im adulten Stadium und in der Embryonalentwicklung nicht essenziell ist. Meine Ergebnisse zeigen, dass im adulten System der Verlust von *Meg3* die Reaktion auf Stressmediatoren und die Funktion in seriellen Transplantationsexperimenten nicht beeinträchtigt. (Sommerkamp et al., 2019)

Projekt 2: Hämatopoetische Stammzellfunktion, -differenzierung und -aktivierung werden durch APA reguliert.

Für die meisten Säugetiergene existieren mehrere verschiedene Polyadenylierungsstellen. APA kontrolliert die Selektion der Polyadenylierungsstellen und ändert dadurch die 3'-UTR Länge und die Isoform Expressionsmuster. So moduliert APA die RNA-Stabilität, RNA-Lokalisierung, Proteinproduktion und die Proteinlokalisierung (Di Giammartino et al., 2011; Tian and Manley, 2017). Die Rolle der APA bei der Regulation des adulten hämatopoetischen Systems ist derzeit unbekannt. In dieser Doktorarbeit zeigen wir, dass der APA-Regulator *Pabpn1* für HSZs essenziell ist. Der „Knockdown" (KD) von *Pabpn1* führte zu einer verminderten HSZ-Funktion *in vitro* und *in vivo*. Um die Prävalenz von APA an der Spitze der hämatopoetischen Hierarchie zu analysieren, habe ich einen 3'-Sequenzierungsansatz für geringe RNA Mengen etabliert und HSZs, MPP-Zellen und durch Entzündungsmediatoren aktivierte HSZs analysiert. Die bioinformatische Analyse ergab dynamische APA-Muster in zahlreichen Genen von HSZs und MPPs und während der entzündungsinduzierten HSZ-Stressreaktion. Wir konnten eine globale Verkürzung der 3'-UTR Länge beobachten, ausgelöst durch HSZ-Differenzierung und HSZ-Aktivierung. Weiterhin konnten wir durch die 3'-Seq-Analyse von *Pabpn1* KD-Zellen zeigen, dass *Pabpn1* APA im HSZ/MPP-Kompartiment reguliert. Wir beobachteten eine APA-vermittelte Veränderung der Glutaminase (*Gls*) Isoform-Expression nach HSZ Aktivierung. Die Veränderung der *Gls* Isoform-Expressionsmuster führte zu einer erhöhten relativen Expression der hochaktiven *Gls* Isoform GAC und zu insgesamt erhöhtem GLS Proteinlevel. Kurzfristige Hemmung der Glutaminolyse *in vitro* verbesserte die HSZ Kultivierung durch Begrenzung der Proliferation. Ich konnte zeigen, dass die veränderte *Gls* Isoform-Expression und die erhöhte Glutaminolyse durch den APA-Regulator NUDT21 vermittelt wird und für eine einwandfreie HSZ-Funktion erforderlich ist. KD von *Nudt21* hemmte das Umschalten des *Gls* Isoform-Expressionsmuster und führte damit zur Beeinträchtigung

der HSZ-Funktion und einer teilweisen Inhibition der Differenzierung. Zusammenfassend zeigen meine Daten, dass APA und die damit verbundenen Anpassungen des Glutaminstoffwechsels unterschiedlich zwischen HSZ und MPPs reguliert sind und diese Mechanismen einen Beitrag zur Regulation der hämatopoetischen Hierarchie darstellen. (Sommerkamp et al., in Revision)

Projekt 3: Generierung von *Pabpn1*^{flox}-Mäusen

Um in der Zukunft eine eingehende Analyse der Rolle von APA *in vivo* in verschiedenen Geweben und Krankheitsbildern zu ermöglichen, haben wir ein induzierbares *Pabpn1*^{flox} Mausmodell mittels *Easi*-CRISPR Ansatz generiert (Miura et al., 2018; Quadros et al., 2017). Ich habe umfangreiche *in vitro* Analysen zur Bestimmung der Effizienz verschiedener „guideRNAs“ (gRNAs) durchgeführt, um die Rekombinationseffizienz *in vivo* zu verbessern. Zwei verschiedene crRNA-tracrRNA:Cas9-Komplexe, die auf die vor und nach der entsprechenden Gensequenz lokalisierten genomischen Regionen abzielen, wurden zusammen mit einer langen einzelsträngigen DNA (ssDNA) Matrize in einen der Pronuclei der Zygoten injiziert. Durch umfangreiche Genotypisierung konnte ich 3 von 17 Nachkommen als korrekt modifiziert identifizieren. Homozygote Nachkommen konnten erfolgreich gezüchtet werden und können in der Zukunft verwendet werden, um die Funktionalität des *Pabpn1*^{flox} Mausmodells zu bestimmen und APA in verschiedenen biologischen Fragestellungen funktionell zu untersuchen.

III. Declarations

The work presented in this dissertation was performed from January 2016 until March 2019 at the German Cancer Research Center (DKFZ, Heidelberg) and the Heidelberg Institute for Stem Cell Technology and Experimental Medicine (HI-STEM gGmbH, Heidelberg) under the supervision of Prof. Dr. Andreas Trumpp.

The experiments described in project 1 of this dissertation have been published as a paper in Scientific Reports (Sommerkamp et al., 2019).

The experiments described in project 2 of this dissertation have been submitted as a paper and are currently under revision based on reviewers' comments.

I also contributed to additional published projects and publications (Cabezas-Wallscheid et al., 2017; Sommerkamp and Trumpp, 2016; Zhang et al., 2016).

Declarations according to § 8 of the doctoral degree regulations.

1) The thesis I have submitted is my own work.
2) I have only used the sources indicated and have not made unauthorized use of services of a third party. Where the work of others has been quoted or reproduced, the source is always given.
3) I have not yet presented this thesis or parts thereof to a university as part of an examination or degree.

Contents

1 Introduction

1.1 The hematopoietic system

1.1.1 Stem cell hierarchies

Development and growth of multicellular organisms ultimately depend on the constant production of terminally differentiated cells. During adulthood, production of new cells is vital to enable tissue maintenance and regeneration, as many terminally differentiated cells have a limited lifespan, for example mature cells in the gastrointestinal tract, in the skin and in the blood. The centerpiece of embryonic development, growth and tissue regeneration are stem cells. (Alberts, 2008)

Mammalian development starts from a single totipotent zygote. This cell and its progeny cells in the early morula are able to give rise to all cell types found in the embryo and extraembryonic trophoblast cells. In the blastocyst state, the cells of the inner cell mass are pluripotent. They can give rise to all the three different germ layers, namely endoderm, mesoderm and ectoderm, but not to the outer trophoblast layer. Pluripotent embryonic stem cells (ESCs), which are widely used nowadays for *in vitro* stem cell research of potency and cell fate decisions, are derived from the inner blastocyst cell mass. During development, cells become increasingly more specified and their overall stem cell potential decreases. No pluripotent stem cells can be found in the adult organism. Thus, embryonic development in itself can serve as a paradigmatic example of hierarchical cellular organization. (Alberts, 2008)

These hierarchical organization patterns are also omnipresent in the adult organism. Multipotent tissue stem cells generate progeny to ensure tissue function and regeneration throughout life (Fuchs and Chen, 2013). They can only differentiate into cell types of the respective tissue and are able to self-renew. Nowadays, *in vitro* reprogramming of these restricted multipotent adult stem cells and even of mature cell types into naïve, embryonic states is possible and led to the generation of induced pluripotent stem cells (iPSCs) (Takahashi et al., 2007; Takahashi and Yamanaka, 2006; Yu et al., 2007). Traditionally, life-long self-renewal has been defined as a hallmark of tissue stem cells (Fuchs and Chen, 2013). It has become apparent that aging, reduced regenerative capacity and disease development are associated with tissue stem cell exhaustion (Ahmed et al., 2017; Ermolaeva et al., 2018). Thus, aging is intrinsically associated with a loss of stem cell potency and function.

In general, tissue stem cells are thought to divide rather infrequently, giving rise to progenitor cells by asymmetric cell division (Morrison and Kimble, 2006). Asymmetric cell division allows

maintenance of the multipotent stem cell pool and at the same time enables production of a daughter progenitor cell (Morrison and Kimble, 2006). These progenitor cells divide rapidly and develop into short-lived oligopotent progenitor populations, which eventually give rise to terminally differentiated functional cells of the respective tissue (Orkin and Zon, 2008). This differentiation axis, stem cells – progenitor cells – differentiated cells, is the core of the classical hierarchical differentiation tree of cell production in the adult organism (Morrison and Spradling, 2008). With differentiation, cells become increasingly more specified and at the same time progressively lose cellular potency (Morrison and Spradling, 2008). It has become apparent that this unidirectional hierarchical differentiation tree can serve as a model to study and describe stem cells, differentiation and cell fate decisions. However, shortcuts and return paths seem to exist in this hierarchy, indicating towards more flexible and less unidirectional differentiation pathways (Sanchez Alvarado and Yamanaka, 2014). In addition, stem cell differentiation and lineage commitment seem to be a continuous process rather than a stepwise generation of increasingly restricted progenitors and differentiated cells (Haas et al., 2018).

Many different mammalian tissue stem cells and associated differentiation hierarchies have been studied to date. Until now, however, hematopoietic stem cells (HSCs) remain the best characterized adult tissue stem cell population in the human and murine system. Classically, HSCs are defined as multipotent stem cells that are responsible for life-long production of all mature blood cells and are considered to represent the centerpiece of hematopoiesis (Orkin and Zon, 2008; Seita and Weissman, 2010; Weissman and Shizuru, 2008).

1.1.2 Hematopoiesis

Most mature blood cells are relatively short-lived. Mature erythrocytes, for example, have a limited lifespan of approximately 120 days (Shemin and Rittenberg, 1946). Thus, HSCs are essential to ensure lifelong replenishment of mature functional blood cells, including erythrocytes, megakaryocytes/platelets, myeloid cells like monocytes, macrophages and granulocytes, mast cells, dendritic cells (DCs) and lymphocytes like T cells, B cells and natural killer (NK) cells (Seita and Weissman, 2010). In the adult organism, HSCs mainly reside in the bone marrow (BM), where also the majority of blood and immune cell maturation takes place (Mendelson and Frenette, 2014; Morrison and Scadden, 2014). If not indicated otherwise, this introduction focusses on murine hematopoiesis.

1.1.2.1 The hematopoietic hierarchy

The hematopoietic system is hierarchically organized (Figure 1.1). HSCs are at the top of the hematopoietic hierarchy, are able to self-renew and give rise to all different lineages of the

blood (Orkin and Zon, 2008; Seita and Weissman, 2010; Weissman and Shizuru, 2008). Traditionally, HSCs are identified by their ability to serially repopulate the BM of lethally irradiated recipient mice generating life-long multilineage cellular output (Purton and Scadden, 2007; Till and McCulloch, 1961). The classical hematopoiesis model is defined by unidirectional, binary cell-fate decisions and stable progenitor cell populations (Figure 1.1) (Haas et al., 2018). Multipotent HSCs beget multipotent progenitor 1 (MPP1) cells, which have slightly reduced stem cell capacity (Wilson et al., 2008; Wilson et al., 2007). This population differentiates into MPP2, MPP3 and MPP4 cells. These MPP populations are more proliferative than HSCs (Passegue et al., 2005; Wilson et al., 2008). However, their long-term engraftment in experimental settings is reduced, and they show a lineage-bias in transplantation experiments (Cabezas-Wallscheid et al., 2014; Pietras et al., 2015). MPP cells are traditionally thought to differentiate into oligopotent progenitor cells, namely common myeloid progenitors (CMPs), which give rise to the myeloid lineage, and common lymphoid progenitors (CLPs), which give rise to the lymphoid lineage (Akashi et al., 2000; Kondo et al., 1997). CMPs differentiate into more restricted oligopotent progenitors, the megakaryocyte-erythrocyte progenitors (MEPs) and the granulocyte-macrophage progenitors (GMPs) (Akashi et al., 2000). MEPs are restricted to generation of cells belonging to the megakaryocyte and erythrocyte lineage, while GMPs give rise to the granulocyte and macrophage lineages (Akashi et al., 2000). Oligopotent progenitors are even more restricted in their potency and generate lineage-restricted progenitor cells, including megakaryocyte (Mk) and erythrocyte progenitors as well as Pro-DCs, Pro-T cells, Pro-NK cells and Pro-B cells (Seita and Weissman, 2010; Weissman and Shizuru, 2008). Pro-DCs are thought to be generated by CMPs as well as CLPs, representing a myeloid/lymphoid lineage cell population (Weissman and Shizuru, 2008). These restricted progenitors later differentiate into mature effector cells generating platelets, erythrocytes, granulocytes, macrophages, DCs, T cells, NK cells and B cells.

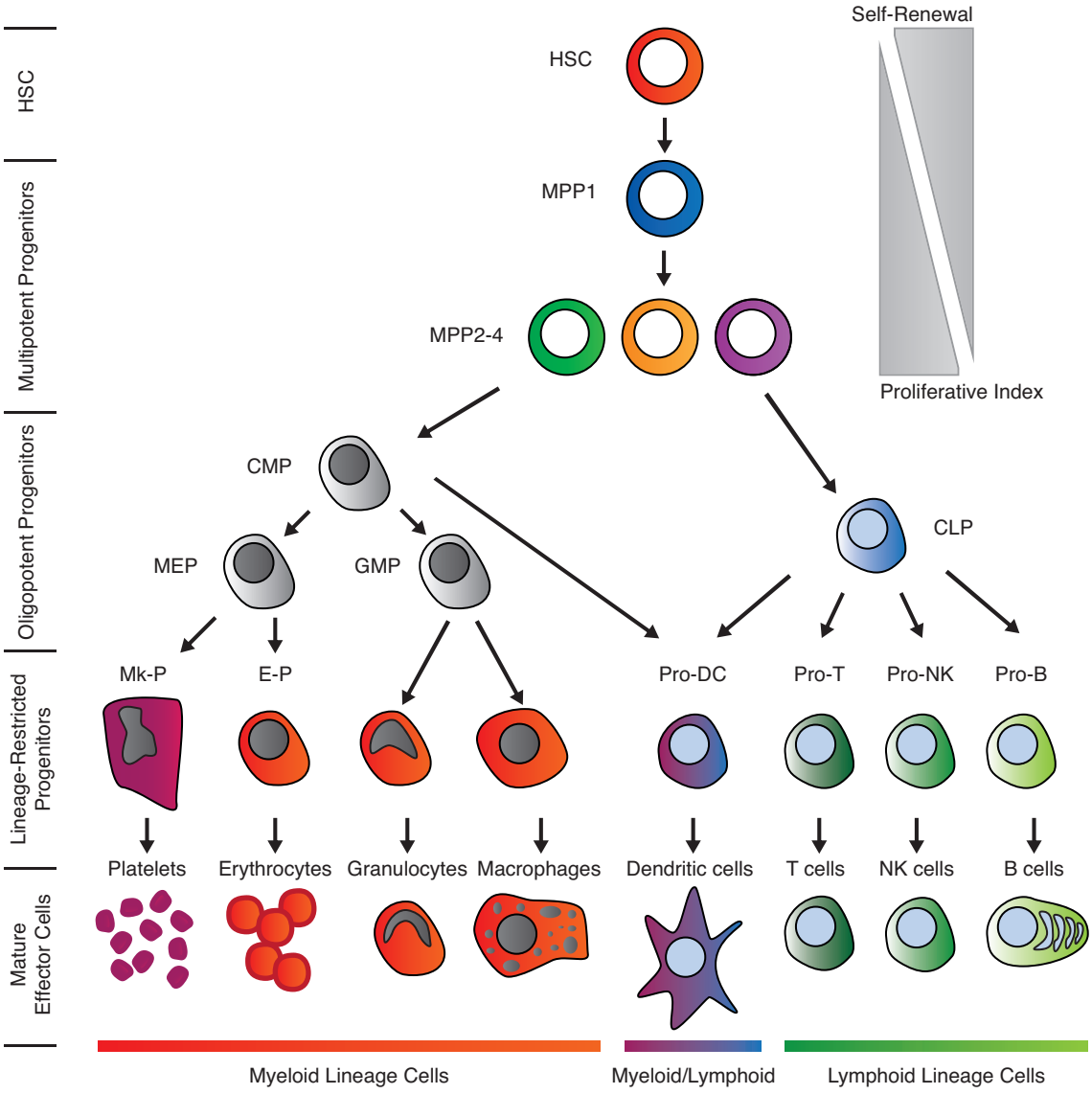


Figure 1.1: Classical hematopoietic hierarchy.
The multipotent HSC sits at the top of the hierarchy giving rise to different MPP subsets (MPP1-4). Distinct branchpoints are identifiable between the erythrocyte-myeloid lineage and the lymphoid lineage. Oligopotent progenitor populations give rise to lineage-restricted progenitors, eventually differentiating into mature effector cells. Self-renewal capacity is highest in the HSC population and gradually decreases throughout the differentiation cascade. In contrast, the proliferative index increases from HSC towards MPP1-4. Mgk-P: megakaryocyte progenitor, E-P: erythrocyte progenitor. Redrawn and modified based on (Bryder et al., 2006).

In recent years, the classical tree-like hematopoiesis model with defined stable progenitor populations harboring distinct potencies has been challenged (Haas et al., 2018). Limiting-dilution and single-cell transplantation assays revealed intrinsic HSC lineage biases in murine as well as human HSCs (Dykstra et al., 2007; Morita et al., 2010; Muller-Sieburg et al., 2002; Yamamoto et al., 2013). Only a minority of HSCs found in the immunophenotypically defined HSC compartment exhibited unbiased multilineage engraftment upon transplantation. The majority of HSCs displayed lineage-biased engraftment, which is characterized by differences in reconstitution kinetics and lineage contribution. In addition, *in silico* barcoding and fluorescent tagging approaches confirmed lineage skewing of HSCs in the unperturbed hematopoietic system (Rodriguez-Fraticelli et al., 2018; Yu et al., 2016). Single-cell

transcriptional profiling revealed transcriptional lineage priming, further supporting the presence of a highly heterogenous human and mouse HSC and progenitor compartment (Grover et al., 2016; Guo et al., 2013; Moignard et al., 2013; Velten et al., 2017). In addition, transcriptional and functional in-depth analysis of oligopotent progenitor populations, e.g. CMPs, revealed that these populations are heterogenous and seem to consist of multiple lineage-committed progenitor populations (Karamitros et al., 2018; Paul et al., 2015). Further, lineage-committed HSCs have been identified. Within the phenotypically defined HSC compartment, Mk-restricted cells with high self-renewal capacity exist (Haas et al., 2015; Rodriguez-Fraticelli et al., 2018; Yamamoto et al., 2013). Thus, these recent experimental findings have changed our view of the hematopoietic hierarchy. Lineage-commitment does not seem to be acquired in stepwise binary cell fate decisions but rather in a continuous manner (Figure 1.2) (Haas et al., 2018). Today, we can use phenotypically identified HSCs and progenitor populations to perform functional and transcriptional analysis and use the tree-like hematopoietic hierarchy as a model. However, the intrinsic heterogeneity in these defined populations always has to be considered when interpreting results.

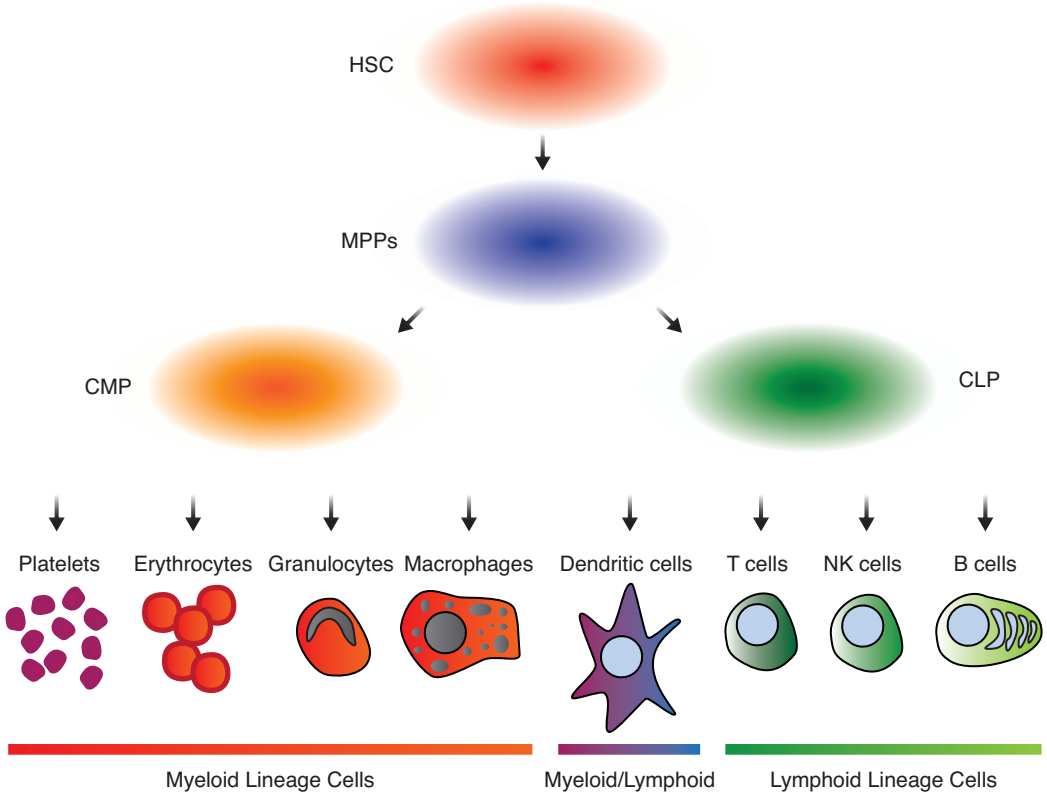


Figure 1.2: Continuous model of HSC differentiation.
In this model, no distinct subpopulations are defined. HSCs become increasingly more committed and lose self-renewal capacity as well as multipotency. MPP, CMP and CLP states are transitory states during HSC differentiation. Redrawn and modified based on (Haas et al., 2018).

1.1.2.2 Murine embryonic hematopoiesis

During murine embryonic development, distinct waves of hematopoietic activity occur in different anatomical sites (Galloway and Zon, 2003; Orkin and Zon, 2008). Primitive hematopoiesis, which primarily generates primitive erythrocytes, occurs in the yolk sac starting at embryonic day 7.5 (E7.5) (Galloway and Zon, 2003; Orkin and Zon, 2008). This initial wave of erythropoiesis is needed to meet the increasing demand of oxygen supply in the developing embryo and is characterized by the presence of embryonic globin proteins. From E9 on, definitive hematopoiesis can be detected, which gives rise not only to erythrocytes and macrophages but also to other myeloid and lymphoid cells (Galloway and Zon, 2003; Orkin and Zon, 2008). This next wave of hematopoiesis mainly takes place in the aorta-gonad-mesonephros (AGM) region. While little HSC activity is detected in this region at E10.5, transplantable HSCs are present at E11 (Muller et al., 1994). The AGM region therefore seems to be the source of the first definitive HSCs. However, HSC activity has also been detected in the placenta, adding a potential additional site for HSC development (Gekas et al., 2005; Ottersbach and Dzierzak, 2005). Subsequently, HSCs are thought to colonize the fetal liver starting E11 (Galloway and Zon, 2003; Johnson and Moore, 1975; Muller et al., 1994), where they massively expand and differentiate. However, the liver is only a transient hematopoietic site, as hematopoiesis in the liver decreases shortly after birth (Galloway and Zon, 2003). During development, the spleen is colonized at E12, and hematopoiesis in the spleen starts around E16, aiding in the transition from fetal liver to BM hematopoiesis (Galloway and Zon, 2003; Godin et al., 1999; Sasaki and Matsumura, 1988). HSCs migrate to the developing BM during embryogenesis, and B cell precursors are detected in the BM as early as E15 (Delassus and Cumano, 1996). Functional HSCs, however, are only identified in the BM starting E18 (Galloway and Zon, 2003). After birth and during the lifespan of the murine organism, the BM remains the main site of definitive hematopoiesis.

Fetal HSCs are able to engraft in long-term transplantation assays, are multipotent and able to self-renew. However, their intrinsic characteristics differ depending on the embryonic site of origin. In contrast to adult HSCs, for example, fetal liver HSCs are characterized by high proliferative activity (Bowie et al., 2006; Nygren et al., 2006; Wilson et al., 2008).

1.1.2.3 Murine adult hematopoiesis

In the adult organism, hematopoiesis mainly takes place in the BM, where HSCs reside and blood cells of all different lineages are produced, including myeloid as well as lymphoid cells. Most cells mature in the BM, including platelets, myeloid cells, B cells and reticulocytes, which advance to the erythrocyte stage in the blood (Seita and Weissman, 2010; Weissman and

Shizuru, 2008). T cell maturation, however, takes place in the thymus. After migration of lymphoid precursor cells to the thymus, T cell development and generation of different T cell lineages occurs (Alberts, 2008). In addition to the primary lymphoid organs, BM and thymus, where lymphoid cells are generated and mature, secondary lymphoid organs exist. These include lymph nodes, the spleen, tonsils, peyer's patches and mucosa-associated lymphoid tissues. In these tissues, initial antigen contact takes place, leading to activation and expansion of lymphocytes (Murphy, 2008). Recently, additional tissues have been linked to hematopoiesis. For example, HSCs have been identified in the gut, and platelet production was observed in the lung (Fu et al., 2019; Lefrancais et al., 2017). Thus, hematopoiesis is also associated to and observed in other organs.

1.1.3 Hematopoietic stem and multipotent progenitor cells

Hematopoietic stem and multipotent progenitor cells (HSPCs) are at the top of the hematopoietic hierarchy. To enable analysis and characterization of HSPCs, HSCs and different MPP populations have been phenotypically and experimentally defined in the past.

1.1.3.1 Identification of HSPCs by surface marker expression patterns

In general, surface marker combinations are used to identify different HSPC populations, for example via flow cytometry analysis. All HSCs and MPP cells are contained in the lineage-negative fraction. Lineage markers include cluster of differentiation (CD) 4, CD8, B220, Gr1, Ter119 and CD11b (Muller-Sieburg et al., 1988; Muller-Sieburg et al., 1986).

Initially, it could be shown that multipotent HSC activity is restricted to the Lin- Sca-1+ c-Kit+ (LSK) compartment (Figure 1.3 A) (Ikuta and Weissman, 1992; Ogawa et al., 1991; Spangrude et al., 1988; Wilson et al., 2007). Later on, additional markers were identified by transplantation assays, including CD34, CD135 and the signal lymphocyte activating molecule (SLAM) receptors CD150 and CD48, leading to the identification of the murine LSK CD135- CD48- CD150+ CD34- HSC population (Adolfsson et al., 2005; Forsberg et al., 2006; Kiel et al., 2005; Osawa et al., 1996; Wilson et al., 2007). In the past, alternative assays, e.g. dye efflux, have been used to isolate and study HSCs (Purton and Scadden, 2007). Furthermore, additional HSC-specific genes including α-*catulin*, *Hoxb5* and *Fgd5* were recently identified and used to enrich for stem cell function by employing reporter mouse lines (Acar et al., 2015; Chen et al., 2016; Gazit et al., 2014). The LSK CD135- CD48- CD150+ CD34- HSC surface marker code is, however, widely used and accepted.

HSCs defined using these surface markers are extremely rare and only make up around 0.0013% of all murine BM cells (Wilson et al., 2008). Nevertheless, these HSCs are extremely

potent as they can be serially transplanted, are multipotent and able to self-renew (Seita and Weissman, 2010; Weissman and Shizuru, 2008). In addition, they are highly quiescent, and less than 2% of all phenotypic HSCs are actively cycling. The majority of HSCs, up to 90%, reside in the G0 state of the cell cycle and can therefore be found in the Ki-67-negative compartment (Wilson et al., 2008). Presumably, this quiescent state protects HSCs from genetic damage (Flach et al., 2014; Walter et al., 2015). The resulting low mutational burden protects HSCs from malignant transformation and ensures lifelong production of mature blood cells (Trumpp et al., 2010).

HSCs can be further subdivided into an active HSC (aHSC) and a dormant HSC (dHSC) compartment (Figure 1.3 B) (Cabezas-Wallscheid et al., 2017; Foudi et al., 2009; Wilson et al., 2008). These subpopulations can be identified by label-retaining assays, for example using BrdU labeling approaches or the SCL-tTA:H2B-GFP transgenic mouse model (Wilson et al., 2008). dHSCs retain the cellular label for up to 300 days, indicating a highly dormant cellular state. 25-35% of all phenotypic HSCs are dormant, as identified after a chase period of 70 to approximately 150 days. All dHSCs are in the G0 state of the cell cycle. Wilson et al. could show that only dHSCs, but not aHSCs, are serially transplantable, indicating that dHSCs harbor the majority of the multilineage long-term self-renewal capacity (Wilson et al., 2008). Mathematical modeling revealed that dHSCs divide only 4-5 times per mouse lifetime (approximately every 149-193 days), while aHSCs divide every 28-36 days (van der Wath et al., 2009). It has been proposed that dHSCs function as a backup HSC population that is protected from exhaustion and acquisition of mutations by their deeply dormant state (Trumpp et al., 2010; Wilson et al., 2009). In addition, label-retaining HSCs have been proposed to persist during aging (Bernitz et al., 2016). dHSCs become activated and contribute to hematopoiesis under stress conditions or upon injury. Cabezas-Wallscheid et al. further characterized dHSCs and aHSCs on a transcriptional level (Cabezas-Wallscheid et al., 2017). In this study, it was shown that dHSCs are characterized by a state of biosynthetic shutdown. In addition, GPRC5C was identified as a surface marker expressed on dHSCs.

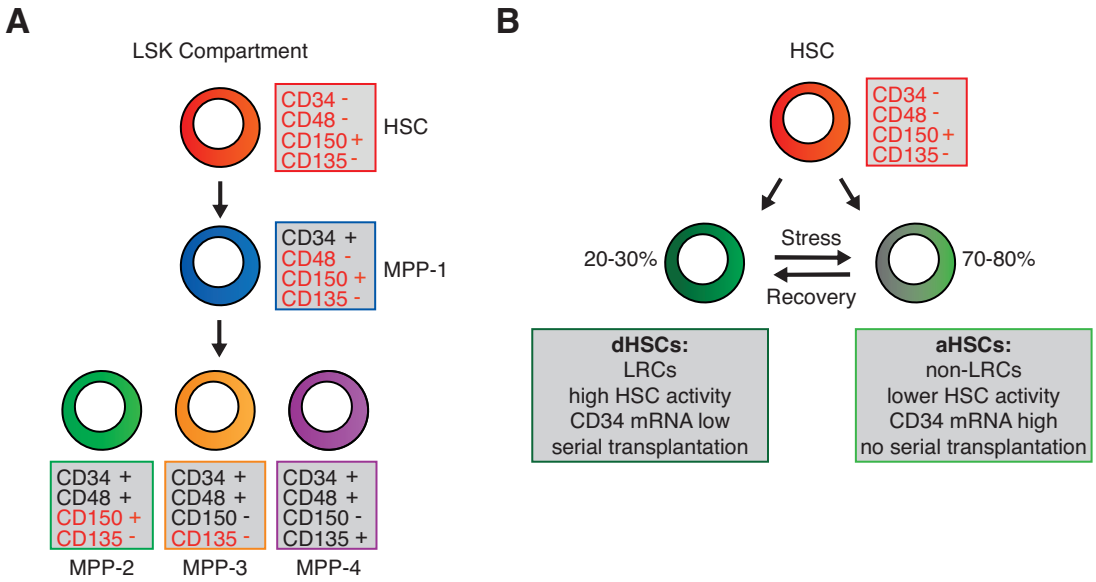


Figure 1.3: HSC-MPP populations.
(A) Surface marker codes of HSCs and MPP1-4 cells. CD34, CD48, CD150 and CD135 are used to identify the different HSPCs within the LSK population. (B) The HSC compartment can be further subdivided into a dHSC (~20-30%) and an aHSC (~70-80%) compartment. dHSCs, in contrast to aHSCs, are characterized as label retaining cells (LRCs) exhibiting low CD34 mRNA levels and high HSC activity. Only dHSCs are serially transplantable. dHSCs become activated during hematopoietic stress responses and return to their dormant state during stress recovery. Redrawn and modified based on (Cabezas-Wallscheid et al., 2017; Cabezas-Wallscheid et al., 2014).

In addition, HSC progeny cells, namely MPPs, have been described and functionally characterized (Cabezas-Wallscheid et al., 2014; Pietras et al., 2015; Wilson et al., 2007). Morrison et al. were one of the first to describe the reduced engraftment potential of MPP cells compared to HSCs (Morrison et al., 1997; Morrison and Weissman, 1994). Different surface marker combinations have been used to identify MPP cells (Wilson et al., 2008). MPPs can be found in the LSK compartment but, in contrast to HSCs, express CD34 (Osawa et al., 1996). It has been shown that increasing expression of the surface marker CD135 is accompanied by a decrease in self-renewal capacity (Adolfsson et al., 2005). In addition to CD135, the SLAM markers CD48 and CD150 are used to identify MPP subsets (Kiel et al., 2005). Wilson et al. used these markers to phenotypically identify four different MPP subsets (Figure 1.3 A) (Wilson et al., 2008). MPP1 cells are defined as LSK CD34+ CD48- CD150+ CD135-. MPP2 cells upregulate CD48 (LSK CD34+ CD48+ CD150+ CD135-) and MPP3 and MPP4 cells lose expression of CD150. They can be distinguished by CD135 expression patterns (MPP3: LSK CD34+ CD48+ CD150- CD135-; MPP4: LSK CD34+ CD48+ CD150- CD135+). To date, the MPP hierarchy is still controversially discussed and additional markers have been proposed (Oguro et al., 2013). Exact branch points could not be identified so far. Taking recent discoveries into account, it can be assumed that MPP subsets represent phenotypically identifiable populations that are part of a continuum-like developmental process from HSCs towards mature cells (Haas et al., 2018).

Functional assessment of different MPP populations was recently performed (Cabezas-Wallscheid et al., 2014; Pietras et al., 2015). Cabezas-Wallscheid et al. could show that MPP2 cells generate a large number of progeny cells of all different lineages, including B cells, T cells and myeloid cells, upon primary transplantation. MPP3 and MPP4 cells, in contrast, only generated little, but detectable progeny upon transplantation. Furthermore, MPP3 cells exhibited a myeloid and MPP4 cells a lymphoid lineage bias. In line, lymphoid-priming of MPPs has been linked to loss of expression of *Mpl*, which encodes for the thrombopoietin receptor (Luc et al., 2008). However, engraftment and lineage-contribution of MPP cells is always considered short-term compared to HSCs. In addition to engraftment potential and self-renewal capacity, the cell cycle status of MPP cells was analyzed (Wilson et al., 2008). While HSCs are highly quiescent, MPP cells become progressively more active. Less cells are in the G0 state of the cell cycle and more cells are found in the S/G2/M state, supporting the hypothesis that MPP1-4 cells are the active drivers of hematopoiesis under homeostatic conditions. This is also supported by in situ fate-tracing approaches (Busch et al., 2015; Sun et al., 2014).

1.1.3.2 The HSC bone marrow niche

HSCs are highly dependent on a functional microenvironment. To date, it is still not possible to culture HSCs *in vitro* for long time periods. Within the BM, HSCs are localized in specialized niches providing soluble signaling molecules and direct cell-cell contacts, which maintain and regulate HSCs (Crane et al., 2017; Mendelson and Frenette, 2014; Morrison and Scadden, 2014). Thereby, the niche mediates HSC self-renewal, quiescence, differentiation and stress response (Crane et al., 2017; Ehninger and Trumpp, 2011). In recent years, multiple studies have been performed to identify the cellular makeup and to precisely localize the HSC niches in the BM. However, the exact HSC location in the bone and identity of HSC niche cells are still under debate. In general, endosteal as well as perivascular HSC niches have been described (Crane et al., 2017). The majority of HSCs seems to be localized in perivascular niches, with around 80% being associated to sinusoidal blood vessels. Within the perivascular niche, only a minority of HSCs seems to be localized close to arterioles or within transition zone vessels (Acar et al., 2015; Crane et al., 2017). Multiple niche cells that are involved in HSC maintenance have been described, including megakaryocytes, monocytes and macrophages, neurons, Schwann cells, endothelial cells and leptin receptor-expressing, CXC-chemokine ligand 12 (CXCL12)-abundant reticular (CAR) perivascular stromal cells (Crane et al., 2017; Mendelson and Frenette, 2014; Morrison and Scadden, 2014).

Many different key factors in the HSC niche have been reported, including the soluble stem cell factor (SCF), which signals via the c-Kit receptor and is needed for HSC maintenance

(Barker, 1994; Crane et al., 2017; Czechowicz et al., 2007). In addition, CXCL12, signaling via CXC-chemokine receptor 4 (CXCR4), is an important soluble factor involved in retaining HSCs in the BM (Sugiyama et al., 2006; Zou et al., 1998). The CXCL12–CXCR4 signaling axis is also required for HSC homing to the BM upon transplantation by intravenous injection (Lai et al., 2014). Endothelial cells and leptin receptor-expressing, perivascular stromal CAR cells have been identified as the main sources of SCF and CXCL12 in the unperturbed young BM (Crane et al., 2017; Zhou et al., 2014). However, other sources have been reported as well. In addition, multiple other factors have been shown to be important or even essential for HSC maintenance in the niche, including thrombopoietin (TPO) (Crane et al., 2017). Overall, additional research is necessary to further analyze and better understand the organization, regulation and function of the HSC BM niche.

1.1.3.3 Functional analysis of murine HSC activity

Flow cytometry analysis of HSPCs alone cannot be used to investigate HSC capacity. To asses HSC function, self-renewal and potency, several *in vitro* and *in vivo* approaches have been established (Purton and Scadden, 2007). Today, the most common *in vitro* experiment is the colony forming unit (CFU) assay (Purton and Scadden, 2007). In this assay, total BM cells or sorted BM subsets are cultured in a semi-solid methyl-cellulose culture medium, which contains defined growth factors. Single cells are distributed in the medium, and during the incubation time of 5 to 14 days, colonies grow out depending on the initial potency of the given cell. The colony number can be quantified, and the lineage bias can be assessed. Depending on the cell type, different lineages are generated. Immature stem and progenitor cells form mixed colonies (CFU-GEMM), while restricted cell subsets give rise to myeloid (CFU-G/M/GM) or erythroid (CFU-E, BFU-E) colonies. Lymphoid potential cannot be assessed, as development depends on additional extrinsic factors and cell subsets not present in the culture conditions (Purton and Scadden, 2007). In primary platings of total BM cells, mainly the capacity of lineage-restricted progenitor cells and MPP cells is assessed. To investigate HSC capacity, serial CFU assays are needed, as HSCs primarily form colonies in the second and third round of the CFU assay.

In vivo transplantation assays have long been established to identify HSCs and assess HSC capacity. Initially, so-called spleen colony forming unit (CFU-S) assays were used (Till and McCulloch, 1961). In these assays, macroscopic colony formation in the spleen is quantified approximately 2 weeks after intravenous transplantation into irradiated recipient animals. This assay mainly allows investigation of progenitor cell potential (Purton and Scadden, 2007). Nowadays, long-term transplantation assays are the gold standard to address functional potential. A prerequisite for successful transplantation approaches is the depletion of recipient

HSCs in the BM by irradiation or high-dose cytotoxic drug treatment (Kondo et al., 2003; Weissman and Shizuru, 2008). Thereby, vacant HSC niches are created that can be taken over by the transplanted HSCs.

In general, transplantation assays rely on the identification of transplanted cells and progeny cells in the recipient mouse. Therefore, congenic mice differing in the expression of allelic variants of the Ly5 antigen (CD45) are widely used (Purton and Scadden, 2007). CD45.1, CD45.2 and CD45.1/.2 mice exist, and HSCs, MPPs and mature cells can be identified in the BM and the peripheral blood (PB) by surface antigen expression. Platelet and erythrocyte contribution in the PB cannot be assessed, as CD45 is not expressed on these cellular entities. Potency is analyzed by quantifying PB contribution over time and engraftment of HSCs and MPPs in the BM (Purton and Scadden, 2007).

Transplantation of whole BM is used to generate full chimeras and thereby asses general HSC function. In contrast, competitive transplantations can be set up by mixing BM cells with competitor BM cells. These assays are used to analyze HSC performance and potency. HSCs that exhibit minor impairments might engraft and perform normally in full chimeras, however their reduced potency becomes apparent in competitive settings. Fully functional HSCs will outperform impaired HSCs in such conditions. For HSC quantification, limiting-dilution assays can be performed. Recently, more and more single-cell transplantation assays have been used to analyze single cell HSC potency, lineage-skewing and -priming (Dykstra et al., 2007; Morita et al., 2010; Muller-Sieburg et al., 2002; Yamamoto et al., 2013). In these assays, however, co-transplantation of supportive BM cells is needed to ensure survival of the recipient mouse.

Transplantation experiments revealed that HSCs, MPP1 and MPP2 cells are capable of balanced multilineage reconstitution (Cabezas-Wallscheid et al., 2014). MPP1 cells, however, only contribute to the PB of recipient mice for a limited timeframe. Subsequently, contribution significantly declines. Only transplanted HSCs contribute to the PB for the whole lifetime of the mouse. MPP2 cells also exhibit multilineage engraftment, while PB contribution of MPP3 and MPP4 cells is reduced and exhibits a lineage-bias (Cabezas-Wallscheid et al., 2014; Pietras et al., 2015). However, only dHSCs can be serially transplanted and are therefore thought to represent the HSC subset with the highest long-term self-renewal capacity (Wilson et al., 2008). Over the years, different transplantation approaches have been used to increase the purity of the phenotypically defined HSC and MPP populations. In addition, functional properties on the single cell level have become apparent.

It has to be kept in mind that, in transplantation settings, HSC potential is always analyzed in stressed hematopoietic situations, as the recipient animals have undergone inflammation-

inducing, HSC-depleting therapies, which enable donor HSC engraftment. Alternative conditioning methods allowing transplantation without the need of myeloablative therapies have been established, including dietary valine depletion and transgenic mouse models, e.g. c-Kit mutant mice (Taya et al., 2016; Waskow et al., 2009). In addition, *in situ* fate tracing approaches that make use of HSC labeling or barcoding can be used to assess HSC PB contribution and function in homeostatic and stress conditions in a more physiological setting (Busch et al., 2015; Rodriguez-Fraticelli et al., 2018; Sun et al., 2014; Yu et al., 2016).

1.1.3.4 Cell cycle status and activation of HSCs

In situ cell fate tracing approaches and cell cycle analysis revealed that HSCs do not seem to be the main drivers of hematopoiesis under steady state conditions but rather remain in a highly quiescent, protective, cellular state with up to 90% of HSCs residing in the G0 cell cycle phase (Busch et al., 2015; Rodriguez-Fraticelli et al., 2018; Sun et al., 2014; Wilson et al., 2008). This quiescent state protects HSCs from genetic damage, potentially creating a cellular backup resource for times of hematopoietic stress (Pietras et al., 2011; Trumpp et al., 2010).

In line, deeply quiescent HSCs are activated in response to stress and injury signals, including inflammation, chemotherapy and blood loss (Trumpp et al., 2010; Wilson et al., 2009). During infection, the organism is in need of an increased production of immune cells to combat invading pathogens. Upon chemotherapy or blood loss, depleted mature cells need to be replaced. In response, HSCs exit the G0 cell cycle state, actively proliferate and increase cellular output. This activation is also apparent in MPP cells, which exhibit increased cell cycle activity potentially to sustain the HSC emergency response.

HSC activation occurs for example in the presence of lipopolysaccharide (LPS) or inflammatory mediators like IFN-gamma and IFN-alpha (Baldridge et al., 2010; Esplin et al., 2011; Essers et al., 2009; Walter et al., 2015; Zhao et al., 2013). Experimentally, polyinosinic:polycytidylic acid (pIC), a synthetic dsRNA analog, can be used to activate HSCs (Essers et al., 2009). pIC triggers an IFN-alpha response, which signals via the IFN-alpha receptor. Using genetic knockout mouse models, Essers et al. could show that downstream JAK-STAT signaling in HSCs leads to exit of the G0 cell cycle state. In addition, upregulation of the stem cell marker and interferon response gene Sca-1 on HSCs and different Lin- cell populations was reported. Activation of HSCs by pIC is only transient, the response peaks 16h after intraperitoneal pIC injections and subsequently declines. Four days after induction, a partial return to the quiescent state is observed, and after 8 days, the homeostatic status within the HSPC compartment is restored. Interestingly, Essers et al. could show that chronic activation of HSCs by interferon-signaling impairs HSC function and eventually leads to loss

of HSCs. These observations support the notion that the intrinsic quiescent status of HSCs is a prerequisite for lifelong maintenance of this highly potent stem cell population.

Intriguingly, recent studies have suggested that lineage-skewed HSCs and progenitor cells are functionally important in stress hematopoiesis. For example, emergency-megakaryopoiesis seems to be driven by Mk-restricted progenitors that can be found in the phenotypic HSC compartment (Haas et al., 2015). These novel observations in stress hematopoiesis support the theory of a rather heterogenous HSC compartment with lineage-biased HSCs (Haas et al., 2018).

1.1.3.5 The metabolic state of HSCs

Under steady-state conditions, HSCs are highly quiescent. They are characterized by a biosynthetic shutdown, a low proliferative index and low protein translation activity (Cabezas-Wallscheid et al., 2017; Signer et al., 2014; Wilson et al., 2008). Studies on human HSCs have confirmed the quiescent state of the most potent HSCs, which are characterized by low CDK6 levels (Laurenti et al., 2015). Thus, HSCs also seem to be metabolically dormant. Multiple knockout and functional studies have been performed to better understand the metabolic status of HSCs, including deletion of essential metabolic enzymes and regulators (Ito and Ito, 2018; Ito and Suda, 2014). This introduction focusses on the reported metabolic requirements for HSC self-renewal and maintenance.

Interestingly, HSCs are localized in hypoxic niches in the BM (Ito and Suda, 2014; Suda et al., 2011). This hypoxic niche promotes HSC maintenance, inhibits differentiation and protects HSCs from reactive oxygen species (ROS), which potentially induce DNA damage (Ito and Suda, 2014; Suda et al., 2011). HSCs have been reported to express high levels of hypoxia-inducible factor (HIF1) α (Figure 1.4 A) (Simsek et al., 2010; Takubo et al., 2010). Under normoxic conditions, prolyl hydroxylases (PHD) are activated by oxygen and subsequently hydroxylate HIF1α (Semenza, 2010). HIF1α is then bound by the E3 ligase von Hippel-Lindau (VHL) factor, leading to HIF1α degradation by the ubiquitin-proteasome system. Under hypoxic conditions, as observed in HSCs, HIF1α is stable and forms a complex with HIF1β. This complex then drives expression of glycolytic genes. In addition, the pyruvate dehydrogenase kinase (PDK) is activated. This kinase suppresses the pyruvate dehydrogenase (PDH) complex. Thereby, pyruvate oxidation is inhibited and mitochondrial oxidative phosphorylation (OXPHOS) is limited (Ito and Ito, 2018; Ito and Suda, 2014). Thus, HSCs highly depend on anaerobic glycolysis via production of lactate from pyruvate by the lactate dehydrogenase enzyme (Figure 1.4 A) (Ito and Suda, 2014; Takubo et al., 2013). High HIF1α levels, which are caused by HSC residency in the hypoxic niche, prevent engagement of the tricarboxylic acid

(TCA) cycle and subsequent OXPHOS. Limiting OXPHOS is essential for HSCs to remain in an undifferentiated multipotent state. HIF1α deletion, for example, leads to decreased glycolysis and increased metabolism, ultimately leading to HSC exhaustion (Takubo et al., 2010). In line, chemical mitochondrial uncouplers can be used to lower mitochondrial activity in HSCs *in vitro*, thereby increasing HSC maintenance (Vannini et al., 2016). Keeping HSCs in this metabolically dormant state is important to limit cell cycle activity and production of ROS. Excessive ROS generation induces HSC apoptosis presumably to prevent spreading of ROS-induced DNA mutations throughout the hematopoietic system (Ito and Suda, 2014). In addition to anaerobic glycolysis and inhibition of OXPHOS, HSCs seem to actively remove functional mitochondria by autophagy to reduce oxidative metabolism (Ho et al., 2017; Ito and Ito, 2018). In sum, quiescent HSCs rely on anaerobic glycolysis and actively limit the TCA cycle and OXPHOS.

While maintaining a quiescent metabolic state is beneficial for HSCs under homeostatic conditions when progeny cell production from the HSC subset is not required, HSCs need to enter the cell cycle and divide in hematopoietic stress conditions (Figure 1.4 B). This response is accompanied by metabolic adaptations that are necessary to meet the increased demand for amino acids, fatty acids and nucleotides to generate biomass in the form of proteins, lipids and DNA/RNA. Activated HSCs switch to OXPHOS to enhance energy production (Ito and Ito, 2018; Ito and Suda, 2014). Increase in OXPHOS is mediated by growth factor signaling via the PI3K-AKT pathway, leading to activation of mTOR signaling (Ito and Suda, 2014). mTOR signaling, in turn, enhances cell growth and proliferation. In addition, PI3K-AKT signaling leads to forkhead box protein O (FOXO) phosphorylation and translocation out of the nucleus, leading to reduced HSC quiescence (Ito and Suda, 2014; Yamazaki et al., 2006). FOXO proteins usually drive expression of enzymes that are required for ROS detoxification and of quiescence-promoting genes and have been shown to be important for stem cell maintenance (Miyamoto et al., 2007; Tothova and Gilliland, 2007; Tothova et al., 2007). Thereby, growth factor signaling leads to metabolic adaptations in HSCs and ultimately enhanced cell cycle progression.

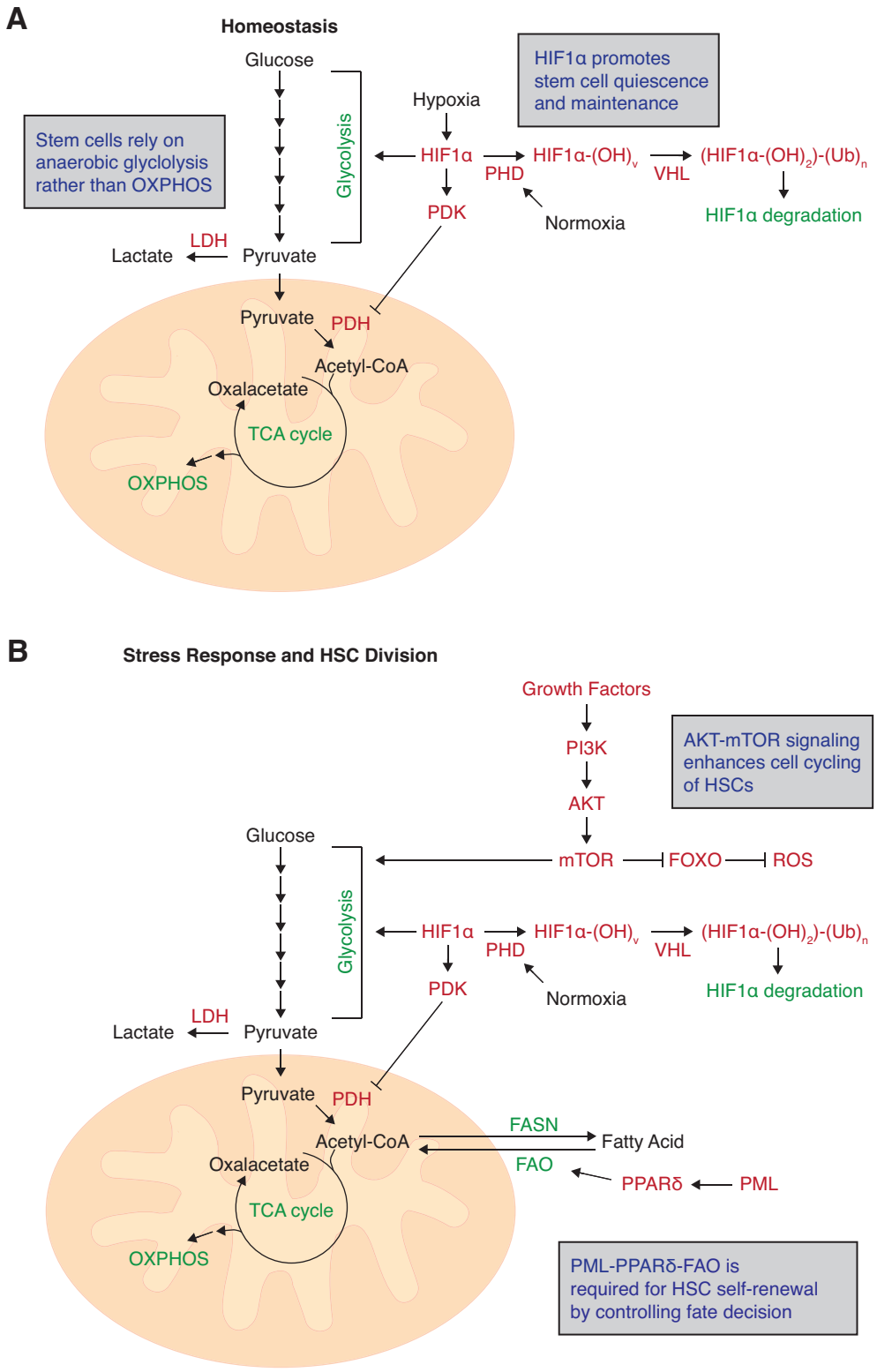


Figure 1.4: HSC metabolism during homeostasis and in response to stress.
(A) Under homeostatic conditions, HSCs rely on anaerobic glycolysis. In the hypoxic niche, HIF1α is stable and not hydroxylated by PHD, which induces HIF1α degradation by the ubiquitin-proteasome system. Active HIF1α promotes glycolysis and activates PDK. PDK inhibits PDH, thereby limiting Acetyl-CoA synthesis and in turn the TCA cycle and OXPHOS. Anaerobic glycolysis leads to the production of lactate from glucose for energy production. (B) During HSC stress response and division, metabolic adjustments take place. Growth factors activate PI3K-AKT signaling, leading to activation of mTOR. mTOR in turn enhances glycolysis and inhibits the FOXO-mediated stress response, usually inhibiting ROS. HIF1α is degraded via the PHD-VHL axis. Thereby, PDH inhibition by HIF1α-induced PDK activation is released. During stress response, the metabolic mode of HSCs switches from anaerobic glycolysis to OXPHOS. In addition, the PML-PPARδ axis promotes FAO. In cycling HSCs FAO controls cell fate decisions and limits oxidative stress. Proteins are shown in red, processes are shown in green. LDH: lactate dehydrogenase, FASN: fatty acid synthase. Redrawn and modified based on (Ito and Suda, 2014).

While entry into the cell cycle and HSC differentiation are crucial in emergency hematopoiesis, the HSC stress-response has to be balanced to maintain functional HSCs throughout life (Ito and Ito, 2018; Ito and Suda, 2014). Recently, fatty acid metabolism has been linked to stem cell self-renewal (Figure 1.4 B) (Ito et al., 2012; Ito et al., 2016). Fatty acid metabolism supports the biosynthetic and bioenergetic requirements of proliferation that are necessary during HSC activation (Ito and Ito, 2018; Ito and Suda, 2014). Activation of the promyelocytic leukemia protein (PML)-peroxisome proliferator-activated receptor (PPAR) δ axis leads to increased expression of genes involved in fatty acid oxidation (FAO) (Ito et al., 2012). The PPARδ signaling axis also enhances mitophagy via the E3 ubiquitin ligase subunit PARKIN and the PTEN-induced putative protein kinase 1 (PINK1) (Ito et al., 2016). By increasing mitophagy, oxidative stress is minimized, and HSC division under stress conditions is enabled. In addition, FAO might also act as a source for NADPH, which counteracts oxidative stress (Ito and Suda, 2014; Pike et al., 2011). Thus, FAO seems to be a key metabolic adaptation during HSC activation, enabling HSC division while minimizing ROS and subsequent DNA mutations.

Glutamine is an important energy source for proliferating cells, fueling the TCA cycle and thereby contributing to OXPHOS metabolism (Figure 1.5) (Yang et al., 2017). Glutamine (Gln) is taken up by cells and can be used to synthesize nucleotides. Deamination of Gln by the enzyme glutaminase (GLS) leads to production of glutamate (Glu). Glu can be converted into α-ketoglutarate (α-KG) by glutamate dehydrogenase or aminotransferases. α-KG is then fed into the TCA cycle, ultimately fueling OXPHOS. In addition, glutaminolysis supports the production of glutathione, which is produced from Glu and protects cells from oxidative stress. Furthermore, transaminases use Glu and α-keto acids to produce α-KG and the corresponding amino acids. Interestingly, glutaminolysis is regulated via c-Myc. It has been shown that c-Myc transcriptionally suppresses the microRNA (miRNA) miR-23, which targets the *Gls* mRNA (Gao et al., 2009). Thus, an increase in c-Myc activity is associated with increased glutaminolysis.

Little is known about the role of glutaminolysis in HSCs and during HSC activation (Ito and Suda, 2014). Interestingly, HSC activation has been linked to increased c-Myc levels (Ehninger et al., 2014). Thus, glutaminolysis could be enhanced in activated HSCs. In addition, HSCs could benefit from using glutaminolysis as an energy source, as it fuels the TCA cycle, thereby increasing energy production. At the same time, glutaminolysis enables production of antioxidants (Yang et al., 2017), e.g. glutathione, which might protect HSCs from DNA damage during the transient activation phase. However, further investigation is needed to better understand the role of glutaminolysis in HSC biology.

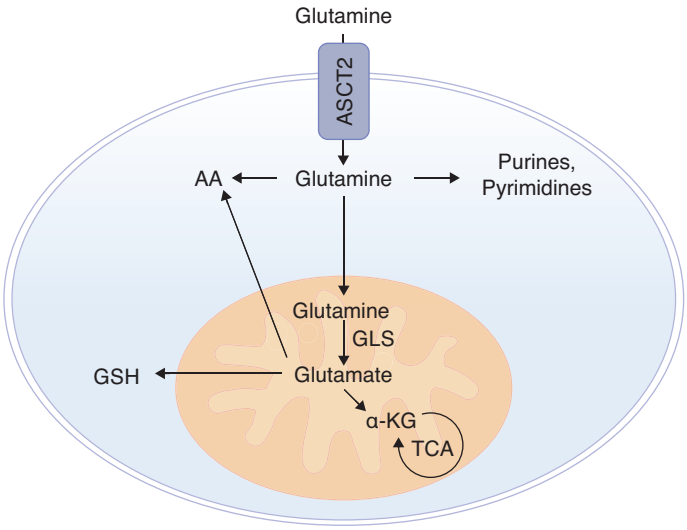


Figure 1.5: Glutaminolysis.
Glutamine is transported into cells via the ASCT2 transporter, where it can be used to synthesize nucleotides and amino acids. GLS converts glutamine to glutamate, which is then used to for glutathione (GSH) or amino acid (AA) production. Alternatively, glutamate is converted to α-KG, which is used in the TCA cycle fueling OXPHOS.

1.1.4 Knockout mouse models used in hematopoietic research

More than 100 knockout (KO) mouse models have been used to investigate the role of different proteins for HSC function, maintenance and regulation (Rossi et al., 2012). BM, PB and secondary lymphoid organ analysis can be performed in these mice. In addition, functional *in vitro* and *in vivo* assays can be executed. Originally, straight KO models were used. In these animals, expression of the respective gene is disabled in all tissues independent of the developmental stage. Thus, it is difficult to distinguish primary from secondary effects in HSCs. For example, KO of an important niche component will reduce HSC number or function. However, the respective gene is not intrinsically needed for HSC maintenance. To circumvent false interpretation of results, inducible and tissue-specific KO mouse models have been established.

1.1.4.1 The Cre/loxP system: using MxCre and VavCre mice in hematopoietic research

Different site-specific recombination systems exist, including the Cre/loxP system, which is widely used (Orban et al., 1992; Sauer, 1987; Sauer and Henderson, 1988). The Cre recombinase can be expressed under the control of a tissue-specific and/or inducible promoter. The genomic sequence of interest is flanked by two loxP sites, each consisting of a 34 bp sequence. Depending on the orientation of the loxP sites, Cre activity either leads to excision of the flanked sequence (same orientation) or inversion of the flanked sequence (opposing orientation). Thus, Cre activation can lead to KO induction in the targeted cell entities and, for example, excision of one or multiple exons.

In hematopoietic research, multiple Cre driver mouse lines have been established. One widely used inducible Cre line is MxCre (Kuhn et al., 1995). Cre expression is controlled by the interferon-response promoter *Mx1*. In healthy mice, this promoter is silent. pIC injections can be used to activated Cre expression under the *Mx1* promoter, leading to efficient deletion of the floxed allele mainly in liver, spleen and hematopoietic cells. Thus, MxCre mice can be crossed to floxed mouse lines, enabling inducible KO of the target gene in the hematopoietic system by pIC injection. As recombination also takes place in other tissues, transplantation approaches can be used to generate chimeras, which only harbor the MxCre transgene and the floxed alleles in the hematopoietic compartment.

MxCre driver mice are used for inducible deletion in adult mice. For hematopoietic deletion of floxed alleles in the embryonic stage, the VavCre mouse model can be used (Georgiades et al., 2002). This model is not inducible, and recombination is observed in hematopoietic cells, endothelial cells and germ cells during embryogenesis. Chen et al. reported VavCre-mediated deletion in E11.5 CD45+ AGM and fetal liver cells (Chen et al., 2009). In their study, deletion occurred in all cells expressing *Vav1*, leading to deletion of the floxed allele after the endothelial-cell-to-HSC transition during embryogenesis. Controversially, deletion in endothelial cells was not observed in this study (Chen et al., 2009).

1.2 Long non-coding RNAs

1.2.1 Function of RNA molecules

For decades, RNA was regarded as a pure messenger molecule, acting as a template to transfer information from stable DNA to chemically functional polypeptides (Brenner et al., 1961; Fatica and Bozzoni, 2014; Morris and Mattick, 2014; Quinn and Chang, 2016). However, rRNAs were identified as functionally important parts of ribosomes and tRNAs were discovered to act as adaptor molecules to enable polypeptide synthesis (Hoagland et al., 1958; Palade, 1955). Nonetheless, mRNA, rRNA and tRNA were considered to enable flow of information from the DNA to the protein level. For years, research was focused on protein function and transcription factor-mediated regulation of gene expression (Morris and Mattick, 2014). This DNA/protein-centric view also affected disease research, as the main focus was laid on mutations in coding DNA regions and impaired protein functions. While these studies are of high importance, investigation of additional functional and regulatory roles of RNA was missing.

Besides mRNA, rRNA and tRNA, small RNAs were discovered relatively early. However, the function of these molecules was not well understood. Later, these RNAs were termed small nuclear RNAs (snRNAs), which were shown to be involved in RNA splicing, small nucleolar RNAs (snoRNAs) and cajal-body specific RNAs (scaRNAs), involved in RNA editing, including methylation and pseudouridylation (Butcher and Brow, 2005; Dreyfuss et al., 1988; Henras et al., 2004; Jady et al., 2004; Maxwell and Fournier, 1995; Meier, 2005; Morris and Mattick, 2014).

In the 1980s, Gilbert postulated the so-called RNA-world hypothesis (Gilbert, 1986). Discovery of enzymatically active RNA, so-called ribozymes, led to the idea that RNA was the molecule used to store information and to catalyze reactions, originally (Guerrier-Takada et al., 1983; Kruger et al., 1982). According to this hypothesis, DNA and proteins evolved later to enable more stable information storage and higher catalytic activities. The underlying discoveries and this new hypothesis opened the door to further research focusing on the function and regulatory abilities of RNA.

1.2.2 Non-coding RNAs

Detailed analysis of the cellular RNA content and refined methods led to the discovery of the so-called non-coding RNA (ncRNA) world. Different ncRNA subclasses have been discovered and functionally characterized in the last 30 years (Fatica and Bozzoni, 2014; Morris and Mattick, 2014; Quinn and Chang, 2016). In total, two thirds of the genomic DNA are

transcribed, while less than 2% are translated into proteins (Djebali et al., 2012; Maeda et al., 2006). These observations indicate that RNA molecules play important functional roles in cellular systems.

MiRNAs, for example, are small, ~22 nt long ncRNAs that influence RNA translation and degradation (Fabian and Sonenberg, 2012; Lee et al., 1993; Morris and Mattick, 2014; Wightman et al., 1993). Pri-miRNAs are synthesized as dsRNA hairpin molecules in the nucleus by RNA polymerase II (Pol II). So-called pre-miRNAs are generated after cleavage by Drosha (Shukla et al., 2011). These molecules are exported into the cytoplasm by Exportin 5, where Dicer processes them to generate small (21-24 nt) dsRNA moieties. One strand of this dsRNA molecule is loaded onto the Argonaute (AGO) protein, which is a part of the multiprotein RNA-induced silencing complex (RISC). This complex binds to target mRNAs by base-pairing, using the miRNA molecule as the complementary strand (Shukla et al., 2011). miRNAs mainly exhibit incomplete homology with their target sequence, leading to translational repression of the targeted mRNA by RISC (Morris and Mattick, 2014). The majority of mRNA molecules, but also other ncRNAs, exhibit multiple miRNA target sites, and at the same time most miRNAs target multiple mRNAs. Thus, the RNA interference pathway is a complex process, fine-tuning RNA stability and translation. miRNAs are involved in many different physiological processes, including stem cell potency and cellular differentiation (Leonardo et al., 2012; Rakoczy et al., 2013). In addition to miRNAs, small interfering RNAs (siRNAs) were described, mainly in plants (Morris and Mattick, 2014). By perfect base-pairing, these molecules induce RNA degradation via AGO-mediated cleavage of target RNAs. Today, siRNAs and artificial short hairpin RNAs (shRNAs) are used to silence target gene expression in experimental settings.

In addition to miRNAs, other ncRNAs have been described. These include the telomerase RNA (TR, TER or TERC), which acts as a template for the synthesis of telomeric DNA, and PIWI-associated RNA (piRNA), which is involved in epigenetic transposon silencing, particularly in male germline cells (Feng et al., 1995; Siomi et al., 2011). Novel sequencing approaches led to the discovery of additional small ncRNA entities, namely transcription initiation RNAs (tiRNAs) and splice site RNAs (spliRNAs) (Taft et al., 2009; Taft et al., 2010). The exact function of these molecules needs to be unraveled, but they seem to be involved in nucleosome positioning and chromatin organization.

In addition to the discovery of small ncRNAs, array-based and sequencing-based analysis of the transcriptome revealed the presence of thousands of lncRNA transcripts (Morris and Mattick, 2014). The ENCODE project provided evidence that lncRNAs are highly common and broadly expressed (Derrien et al., 2012; Rosenbloom et al., 2012). In general, lncRNAs are

defined as non-protein-coding RNAs that are longer than 200 nt (Quinn and Chang, 2016). Like mRNAs, most lncRNAs are transcribed by Pol II and undergo 5'-capping, splicing and polyadenylation. LncRNAs are encoded in intronic, antisense or intergenic regions, and their expression is controlled by promoters that can be regulated by transcription factors. Like many other genes, lncRNAs often undergo alternative splicing. In general, expression levels of lncRNAs are lower compared to mRNAs, and primary sequence conservation between species is lower. However, it has been suggested that secondary structures are conserved and important for lncRNA function. LncRNAs exhibit tissue-specific and stage of development-specific expression patterns (Quinn and Chang, 2016).

1.2.3 LncRNA function

The function of distinct lncRNA molecules has been analyzed (Fatica and Bozzoni, 2014; Morris and Mattick, 2014; Quinn and Chang, 2016) but lncRNA function seems to be highly versatile. lncRNAs can form secondary structures and interact with proteins, leading to assembly of ribonucleoproteins (RNPs). By complementary base-pairing, lncRNAs can additionally interact with DNA and RNA molecules. Thus, multiple modes of action are possible (Figure 1.6).

Nuclear lncRNAs are involved in guiding epigenomic modifiers, thereby recruiting for example DNMT3, PRC2 and H3K9 methyltransferases to certain gene loci (Nagano et al., 2008; Pandey et al., 2008; Zhao et al., 2008). In general, epigenetic roles of lncRNAs have mainly been associated with transcriptional repression. However, transcriptional activation by lncRNAs has also been observed, e.g. by recruiting H3K4 methyltransferases (Bertani et al., 2011; Wang et al., 2011). Besides directing epigenetic regulators to certain gene loci, lncRNAs can also act as scaffold molecules for other chromatin remodelers, epigenetic regulators and transcription factors (Tsai et al., 2010). In epigenetic regulation, lncRNAs can act in *cis* as well as in *trans* (Fatica and Bozzoni, 2014), but their mode of action is poorly understood. In addition, lncRNAs have been reported to control alternative splicing (Morris and Mattick, 2014). Furthermore, lncRNAs exhibit enhancer properties by guiding the physical looping between enhancers and promoters (Orom et al., 2010). Additionally, lncRNAs can affect transcription by acting as decoy molecules, for example for transcription factors (Morris and Mattick, 2014). Thus, lncRNAs can execute versatile functions in the nucleus depending on the respective mode of action.

One of the most widely and longest known lncRNAs functioning in the nucleus is X-inactive specific transcript *(Xist)* (Fatica and Bozzoni, 2014). In female cells, *Xist* is transcribed from one of the two X chromosomes and acts in *cis* to initiate the formation of heterochromatin partly

by tethering PRC2 to the respective X chromosome (Zhao et al., 2008). Thereby, transcription is silenced and gene overdosage in female cells is prevented.

In the cytoplasm, lncRNAs can function as miRNA sponges, sequestering miRNAs. Especially circular RNAs (circRNAs) have a long half-life and mainly act as miRNA decoy molecules (Fatica and Bozzoni, 2014). In addition, lncRNAs have been shown to control translation and mRNA stability by recruiting regulator proteins to target mRNAs (Fatica and Bozzoni, 2014).

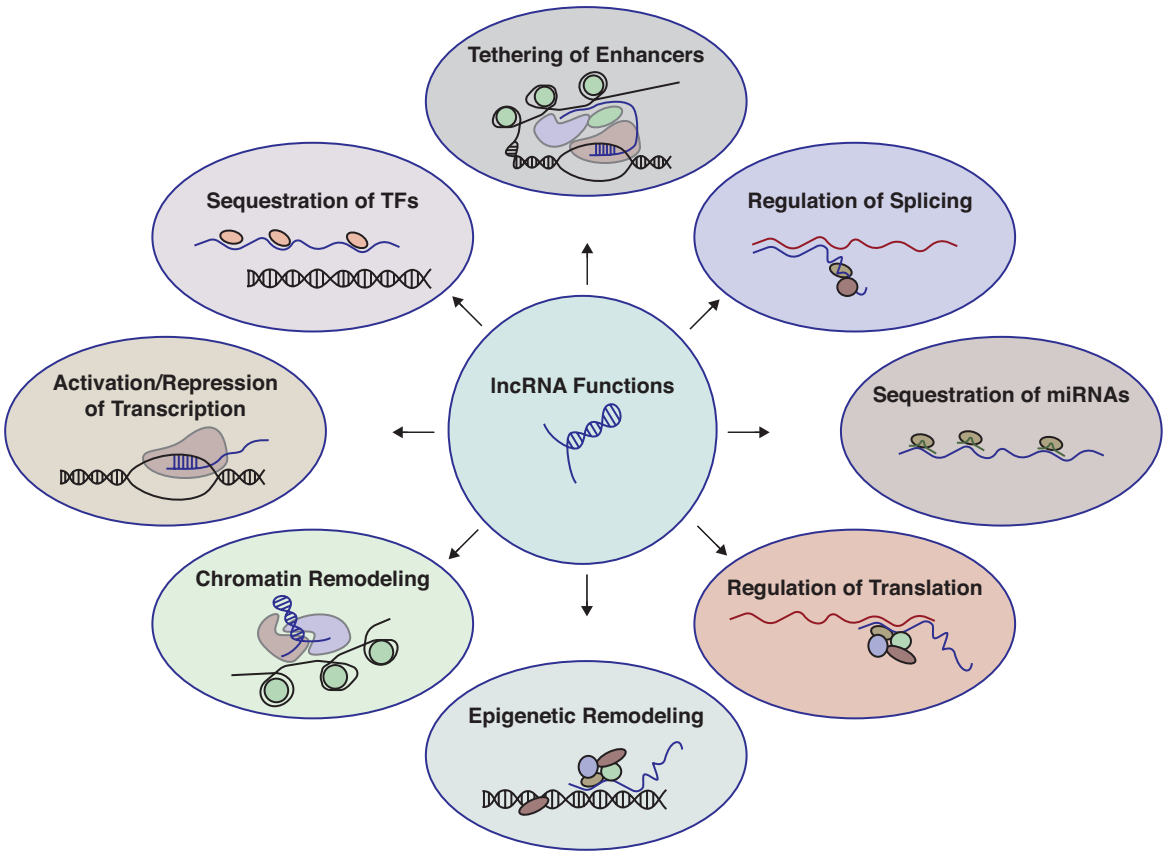


Figure 1.6: LncRNA functions.
Different modes of action have been reported for lncRNAs. In the nucleus, lncRNAs target epigenetic remodelers to certain gene loci. In addition, they act as scaffold molecules for chromatin remodelers. By interacting with proteins and DNA, lncRNAs activate or repress gene expression and guide enhancer-promoter interactions. lncRNAs have also been shown to mediate splicing and function as transcription factor sponges. In the cytoplasm, lncRNAs sequester miRNAs and act as translational regulators by recruiting specific regulatory proteins. Redrawn and modified based on (Malik and Feng, 2016).

Interestingly, many lncRNAs have been associated with imprinted regions (Fatica and Bozzoni, 2014; Morris and Mattick, 2014; Quinn and Chang, 2016). Imprinting leads to a parent-of-origin-specific gene expression pattern. During male and female gametogenesis, the respective imprinting control regions are epigenetically silenced by DNA or histone methylation in a sex-dependent manner. In the embryo and adult organism, either the paternal or the maternal regions are epigenetically silenced, leading to distinct gene expression patterns from the maternal and paternal allele, respectively. lncRNAs are often transcribed from these regions and have been reported to act in *cis*. The paternally expressed lncRNA *Airn*, for

example, is encoded antisense to the *Igf2r* gene (Latos et al., 2012; Quinn and Chang, 2016; Sleutels et al., 2002; Wutz et al., 1997). The maternal allele exhibits hypermethylated CpG islands, leading to silencing of *Airn* and expression of protein-coding genes, including *Igf2r*. On the paternal allele, however, CpG islands are not methylated, and the lncRNA *Airn* is expressed instead of the protein-coding genes. Loss of full-length *Airn* expression leads to loss of imprinting, namely hypermethylation of the paternal CpG islands and reactivation of *Igf2r* expression from the paternal allele. Thus, active transcription and/or production of the mature lncRNA are important to maintain imprinting patterns, as lncRNAs act in *cis* (Quinn and Chang, 2016). One of many lncRNAs expressed from an imprinted locus is maternally expressed gene 3 *(Meg3)*, also known as gene trap locus 2 *(Gtl2)* (Benetatos et al., 2011; Zhou et al., 2012).

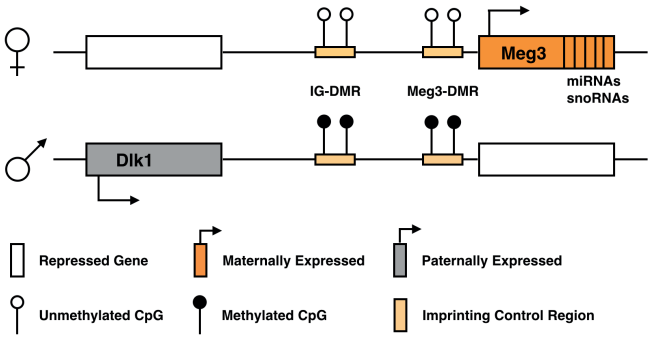
1.2.4 The imprinted *Dlk1-Meg3* locus

Meg3 is a lncRNA encoded in the imprinted *Dlk1-Meg3* gene locus (Figure 1.7 A) (Schuster-Gossler et al., 1998). Imprinted alleles are usually differentially methylated. So-called imprinting control regions (ICRs) are methylated on the paternally or maternally inherited allele, leading to allele-specific expression patterns (Benetatos et al., 2011). Uniparental disomy (upd) in humans, which leads to the presence of either two copies of the paternally or maternally inherited allele or chromosome, is associated with developmental abnormalities and an increased risk of cancer development in the case of maternal upd (Irving et al., 2010; Ogata et al., 2008; Tuna et al., 2009). Thus, correct expression patterns in the *Dlk1-Meg3* gene locus and regulation of gene-dosage effects seem to be of high importance. In the *Dlk1-Meg3* locus, three protein-coding genes, *Dlk1*, *Rtl1* and *Dio3,* are expressed from the paternally inherited allele. In contrast, multiple ncRNAs are expressed from the maternally inherited allele, including the lncRNA *Meg3* and multiple snoRNAs and miRNAs, which are organized in multiple clusters (da Rocha et al., 2008). Expression is controlled by two different intergenic differentially methylated regions (IG-DMRs), namely IG-DMR and *Meg3*-DMR (da Rocha et al., 2008). Methylation patterns of the IG-DMR are established in the germline cells (Kagami et al., 2008; Lin et al., 2007; Lin et al., 2003; Takada et al., 2002). IG-DMR is localized 13 kb upstream of the *Meg3* promoter. Tandem repeats are methylated in male germline cells, but they remain unmethylated in female germline cells. IG-DMR regulates methylation of the downstream *Meg3*-DMR (Kagami et al., 2010; Lin et al., 2003). *Meg3*-DMR methylation patterns are established respectively to the IG-DMR status in postfertilized states, indicating that methylation patterns of DMRs of the *Dlk1-Meg3* locus are hierarchically organized. The *Meg3*-DMR starts approximately 1.5 kb upstream of the *Meg3* promoter and extends into the first intron of *Meg3* (Paulsen et al., 2001; Takada et al., 2002). How these patterns are established, however, remains elusive (da Rocha et al., 2008). In summary, *Meg3* and its

associated ncRNA clusters are expressed from the unmethylated maternally inherited allele, while protein-coding genes are expressed from the methylated paternally inherited allele.

Eleven isoforms of *Meg3* in mice are currently annotated in Ensembl (Zerbino et al., 2018). Structural as well as functional studies showed that *Meg3* indeed acts as a lncRNA, rather than by generating protein output (Zhang et al., 2010b; Zhou et al., 2007). Interestingly, for human *Meg3* transcripts, it has been shown that the secondary folding structure of the different isoforms is comparable, and that the 3D-structure is more important for *Meg3* function than the nucleotide sequence (Zhang et al., 2010b).

A



B

Deletion Model	Targeting Strategy	Phenotype Maternal KO	Phenotype Paternal KO	Publication
ΔIG	deletion IG-DMR	loss of imprinting lethal after E16	no loss of imprinting viable	Lin et al., 2003
ΔMeg3-DMR/exon1-5	deletion Meg3-DMR + Exon 1-5	no loss of imprinting lethal 4 weeks after birth	no loss of imprinting 50% die perinatally 25% die few days after birth 25% show growth retardation	Takahashi et al., 2009
ΔMeg3-partDMR/exon1-5	partial deletion Meg3-DMR + Exon 1-5	loss of imprinting pre-/perinatal death	no loss of imprinting growth retardation	Zhou et al., 2010

Figure 1.7: The *Meg3-Dlk1* locus.
(A) Schematic representation of the *Dlk1-Meg3* locus. On the maternally inherited allele, IG-DMR and *Meg3*-DMR are unmethylated, leading to expression of *Meg3* and the associated miRNA/snoRNA clusters. The paternally inherited allele exhibits methylated CpG islands. From this allele, protein-coding genes like *Dlk1* are expressed. Redrawn and modified based on (Kameswaran and Kaestner, 2014). (B) Overview of different previously published KO mouse models of the *Dlk1-Meg3* locus. Targeting strategy and phenotypes are represented.

1.2.5 *Meg3* function

Many studies have focused on the broader role of the *Dlk1-Meg3* locus, including the function of the different protein-coding genes in the locus, using different mouse models (da Rocha et al., 2008). For example, the transmembrane glycoprotein *Dlk1* has been associated with cellular differentiation (Carlsson et al., 1997; Laborda, 2000; Moon et al., 2002). In addition, studies on IG-DMR and *Meg3* KO mice have also shed light on the role of the maternally

expressed lncRNA *Meg3* and the associated ncRNAs (Figure 1.7 B) (Lin et al., 2003; Takahashi et al., 2009; Zhou et al., 2010).

Lin et al. generated a mouse strain with a 4.15 kb deletion covering the IG-DMR (Lin et al., 2003). At E16, heterozygous animals were detected in the expected ratios. Later in embryonic development, the number of heterozygous live embryos was reduced in animals with maternal inheritance of the deleted IG-DMR. In these animals, bidirectional loss of imprinting on both alleles was observed in the locus, leading to loss of expression of the maternally expressed genes and gain of expression of the paternally expressed genes. These deregulations potentially lead to the observed lethality after E16. In contrast, offspring that inherited the targeted deletion from the father did not exhibit any phenotype, and no loss of imprinting was observed. In line, the model of Zhou et al. showed a highly similar phenotype (Zhou et al., 2010). Here, a 5 kb deletion covering part of the *Meg3* promoter, part of the *Meg3*-DMR and exon 1-5 was generated. Paternal deletion did not lead to loss of imprinting or deregulated expression patterns from the maternal or paternal allele. Maternal deletion, however, was associated with loss of imprinting and reactivation of paternal expression patterns, leading to loss of maternal genes and bi-allelic expression of paternal genes. Consequently, mice died perinatally and presented skeletal muscle defects. In contrast, a mouse model generated by Takahashi et al. exhibited a different phenotype (Takahashi et al., 2009). Deletion of a 10 kb region including the *Meg3*-DMR, the *Meg3* promoter and exon 1-5 was performed. Mice with maternal deletion died 4 weeks after birth, but only loss of *Meg3* expression was observed. ncRNAs from the maternal locus were still moderately expressed. In contrast, mice inheriting the deletion from the father mostly died perinatally. In these mice, expression of the paternal protein-coding genes was reduced, while ncRNAs from the maternal allele, but not *Meg3*, were upregulated. In addition, no loss of imprinting was observed, independently of deletion inheritance. Zhou et al. speculated that, in the model of Takahashi et al., the Neomycin promoter from the resistance cassette, which is oriented in the same direction as *Meg3*, leads to moderate expression of the ncRNA cluster in maternal deletion mice (Takahashi et al., 2009; Zhou et al., 2010). In line, this promoter would drive ncRNA expression from the paternal allele, leading to the observed upregulation of maternal ncRNAs in the paternal deletion mice (Takahashi et al., 2009; Zhou et al., 2010). In summary, *Meg3* itself seems to be important to control imprinting of the *Dlk1-Meg3* locus. In addition, the different phenotypes observed in these mouse models highlight the importance of study design and result interpretation.

Studies in different cancer entities revealed that *Meg3* might act as a tumor suppressor (Benetatos et al., 2011; Zhou et al., 2012). Zhang et al. could show that *Meg3* expression was reduced in nonfunctioning pituitary tumors (Zhang et al., 2003). Re-expression of the lncRNA

in cancer cells led to inhibition of cellular proliferation. Loss of *Meg3* expression has also been reported in meningioma, renal cell carcinoma, leukemia, neuroblastoma, pheochromocytoma and Wilms' tumors (Astuti et al., 2005; Benetatos et al., 2008; Benetatos et al., 2010; Kawakami et al., 2006; Khoury et al., 2010; Zhang et al., 2010a). Interestingly, loss of *Meg3* expression in cancer is not caused by mutations or genomic deletions but mainly by methylation of IG- or *Meg3*-DMR, leading to silencing of the associated maternal locus and reactivation of paternal expression patterns (Benetatos et al., 2011; Zhou et al., 2012). In summary, *Meg3* acts as a tumor suppressor gene and is epigenetically silenced in different tumor entities.

In multiple myeloma (MM), myelodysplastic syndrome (MDS) and acute myeloid leukemia (AML) patients, aberrant methylation patterns of DMRs in the *Meg3* locus were reported (Benetatos et al., 2008; Benetatos et al., 2010; Khoury et al., 2010). No association was observed between methylation patterns and AML subtypes. However, an increase in *Meg3* promoter methylation was associated with decreased overall survival of AML patients. Furthermore, an increase in *Dlk1* expression, which is linked to silencing of the maternal expression patterns, was observed in the majority of AML patients (Khoury et al., 2010). In addition, acquired upd of the *Dlk1-Meg3* locus has been shown to promote clonal hematopoiesis (Chase et al., 2015). Thus, *Meg3* and the *Meg3* locus harbor important tumor suppressor functions in the hematopoietic lineage.

Different functional studies have been performed to identify the mode of action of *Meg3*, mostly in human tumor cell lines (Benetatos et al., 2011). In general, *Meg3* seems to suppress *Mdm2* expression, leading to an increase in p53 levels (Zhou et al., 2007). Recently, Lyu et al. could show that *Meg3* reduces AML cell proliferation via a p53-dependent but also a p53-independent mechanism, presumably by downregulation of DNMT3a (Lyu et al., 2017). They also reported epigenetic silencing of the maternal locus, leading to loss of *Meg3* expression. Interestingly, they could show that TET2 activates the *Meg3* locus via WT1. Thus, TET2 mutations, which are common in AML, could lead to silencing of *Meg3*. In addition, *Meg3* seems to suppress Rb phosphorylation directly and via activation of $p16^{INK4A}$, which negatively regulates Cdk4/6-cyclin D complexes (Zhang et al., 2010c). Thus, in line with the observed tumor suppressor function, *Meg3* acts as an inhibitor of cell cycle. *Meg3* is also involved in regulating angiogenesis via vascular endothelial growth factor (VEGF) signaling (Zhou et al., 2012). An increase in brain blood vessel development in *Meg3* KO mice supports the hypothesis that *Meg3* seems to inhibit the VEGF axis, thereby limiting angiogenesis (Gordon et al., 2010). In summary, *Meg3* acts as a tumor suppressor by regulating p53 and Rb, thereby inhibiting proliferation, and by inhibition of VEGF, leading to a decrease in angiogenesis.

Meg3 expression is stimulated by cAMP response element binding protein (CREB) family activation and binding to cAMP response elements (CRE) located in the *Meg3* proximal promoter (Zhao et al., 2006). Overall, care has to be taken when analyzing *Meg3* function. Silencing of the maternally expressed genes does not only affect the lncRNA *Meg3* but also the other ncRNAs encoded in this locus. In addition, methylation of IG- and *Meg3*-DMR leads to reactivation of the paternal expression patterns and therefore a 2-fold increase in protein levels of DLK1, RTL2 and DIO3. Thus, it is very difficult to distinguish between primary and secondary effects of *Meg3* loss, especially as most studies are correlative and mainly analyze consequences of *Meg3* silencing. However, no real mechanistic studies have been performed on how *Meg3* is supposed to exert its multiple functions. As lncRNA functions are highly versatile, a better understanding of the mode of action of *Meg3* is needed.

1.2.6 The lncRNA *Meg3* in hematopoiesis

Different studies showed that specific lncRNAs play important roles in normal and malignant hematopoiesis, for example *Lnc-HSC-2, H19* and *Xist* (Alvarez-Dominguez and Lodish, 2017; Luo et al., 2015; Venkatraman et al., 2013; Yildirim et al., 2013). Qian et al. focused on the role of the *Dlk1-Meg3* locus in hematopoiesis (Qian et al., 2016). They showed that *Meg3* is highly expressed in fetal liver as well as adult murine HSCs. Using the maternal IG-DMR KO mice (Lin et al., 2003), they observed reduced fetal liver HSC numbers and reduced fetal liver HSC long-term reconstitution capacity in adult recipient mice. They also reported a comparable phenotype using the mouse model generated by Takahashi et al. (Takahashi et al., 2009). Functionally, Qian et al. showed that loss of imprinting leads to hyperactivation of the PI3K-mTOR pathway due to reduced expression of the *Meg3* miRNA cluster. This hyperactivation causes an increase in mitochondrial metabolic activity in HSCs, causing elevated ROS levels and potentially HSC apoptosis. Qian et al. could show that global deregulation of the *Meg3-Dlk1* locus in the murine embryo leads to impairment of fetal liver HSCs. However, it remains elusive if these effects are of primary or secondary nature due to the whole-body deletion approach and the broad deregulation induced in their mouse model. In addition, fetal liver HSCs greatly differ from adult HSCs, as they are actively cycling and expanding (Bowie et al., 2006; Nygren et al., 2006; Wilson et al., 2008). Thus, the role of *Meg3* and the *Meg3-Dlk1* locus in adult HSCs remains to be determined.

1.3 The mechanism of alternative polyadenylation

The majority of mRNA and lncRNA molecules undergo co-transcriptional polyadenylation in the nucleus. This process is important for nuclear export, mRNA stability and translation (Elkon et al., 2013; Tian and Manley, 2017). Polyadenylation involves endonucleolytic cleavage of the pre-mRNA strand and addition of the poly(A) tail to the 3'-UTR. In recent years, it has become apparent that the process of polyadenylation itself is an RNA regulatory mechanism influencing gene expression (Di Giammartino et al., 2011). More than 70% of mammalian mRNA-encoding genes harbor more than one functional cleavage site, and differential usage of these sites is termed alternative polyadenylation (APA) (Tian and Manley, 2017). APA regulates isoform expression and 3'-UTR length, which in turn influences translation efficiency, mRNA localization and stability and even protein localization (Di Giammartino et al., 2011; Mayr, 2017; Tian and Graber, 2012; Tian and Manley, 2017).

1.3.1 Polyadenylation

Pol II, a large protein complex consisting of multiple subunits, synthesizes all eukaryotic mRNAs and multiple ncRNAs with the exception of replication-dependent histone-encoding transcripts. This process is tightly regulated, and transcription initiation requires, for example, binding of multiple transcription factors as well as histone- and chromatin-modifying processes. In addition to its DNA-dependent RNA polymerase activity, Pol II acts as a platform and recruits, among others, epigenetic remodelers. It is also involved in RNA maturation by recruiting RNA capping, splicing and 3'-end processing factors (Gruber et al., 2014a).

All RNA maturation processes mostly occur co-transcriptionally. Polyadenylation is an essential RNA maturation step, and all mRNAs, except for histone mRNAs, and many lncRNA molecules undergo polyadenylation (Elkon et al., 2013; Gruber et al., 2014a). In 1960, Edmonds et al. first described the enzymatic addition of a sequence of adenylate units to RNA (Edmonds and Abrams, 1960). Today, it is well known that the process of polyadenylation is a prerequisite for RNA nuclear export, stability and efficient translation (Elkon et al., 2013). Polyadenylation consists of two steps: endonucleolytic cleavage of the growing RNA strand and addition of the poly(A) tail. Hence, polyadenylation is regulated by the interaction of specific sequence motifs in the RNA and numerous different multiprotein complexes (Figure 1.8) (Tian and Graber, 2012). The cleavage site is surrounded by different sequence motifs. For example, the so-called polyadenylation signal (PAS) is localized around 20 nt upstream of the respective cleavage site (Proudfoot and Brownlee, 1976; Tian and Graber, 2012). The two most common PAS variants are AAUAAA (~55%) and AUUAAA (~16%) (Beaudoing et al., 2000; Tian et al., 2005). However, multiple other conserved PAS motifs have been described

(Beaudoing et al., 2000). In addition to the upstream PAS, downstream sequence motifs have been identified, including the U/GU-rich region (Hu et al., 2005; Hutchins et al., 2008). Upstream of the PAS, so-called UGUA motifs and U-rich elements are localized (Hu et al., 2005; Hutchins et al., 2008). Important protein complexes mediating cleavage and polyadenylation are cleavage and polyadenylation specificity factor (CPSF), which binds to the polyadenylation signal AAUAAA and its variants, and the cleavage stimulatory factor (CstF), which interacts with U/GU-rich regions (Davis and Shi, 2014; Gruber et al., 2014a; Mandel et al., 2008). Cleavage factor Im (CFIm) binds to UGUA motifs upstream of the PAS and cleavage site. Interaction of these protein complexes leads to recruitment of the cleavage factor IIm (CFIIm) and poly(A) polymerase (PAP), initiating transcription termination, cleavage and polyadenylation. Moreover, other factors like Symplekin, which mediates CPSF-CstF interaction, and poly(A) binding protein 1 (PABPN1), which stabilizes and regulates the length of the poly(A) tail, are important. The C-terminal domain of Pol II acts as an assembly platform for these different multiprotein complexes (Davis and Shi, 2014; Gruber et al., 2014a; Mandel et al., 2008). Interaction of RNA motifs and different protein complexes leads to RNA cleavage at the cleavage site and addition of the poly(A) tail. Multiple different cleavage sites and the associated motifs exist in most pre-mRNAs, allowing APA.

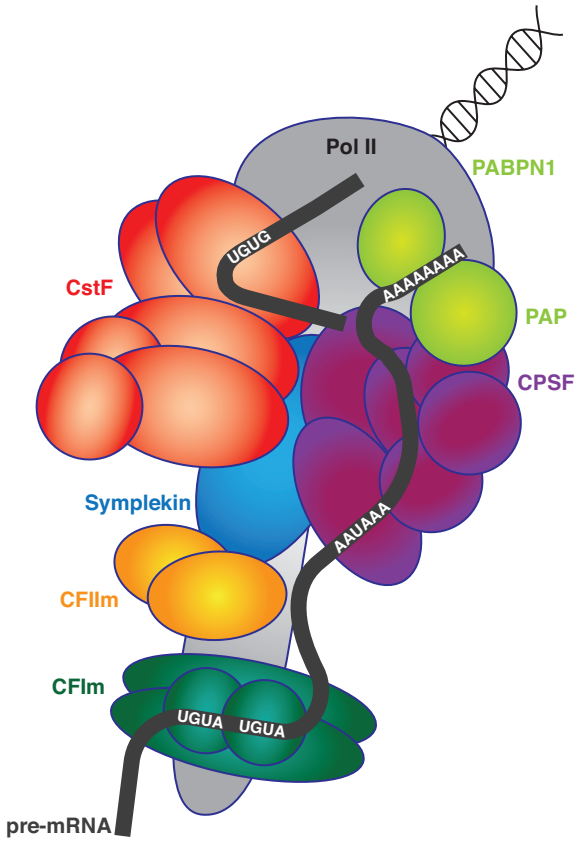


Figure 1.8: The cleavage and polyadenylation complex.
Specific motifs in the RNA strand and different protein complexes mediate RNA cleavage and polyadenylation. Redrawn and modified based on (Gruber et al., 2014a).

1.3.2 Alternative polyadenylation

Multiple PASs exists for most genes. Differential usage of these sites is termed APA and is regulated in tissue-, development-, and disease-specific patterns. APA seems to be an evolutionary conserved RNA regulatory mechanism, as the number of PASs for a given gene is conserved between species (Di Giammartino et al., 2011; Mayr, 2017). APA has been shown to influence mRNA stability, translation efficiency, nuclear export, and mRNA as well as protein localization (see 1.3.4).

1.3.2.1 Genome-wide analysis of APA

One of the first examples of functional APA was the discovery of the APA-mediated switch from the membrane-bound to the secreted form of IgM in B cells (Takagaki and Manley, 1998; Takagaki et al., 1996). Later, global analysis of APA was feasible by using expressed sequence tag (EST) databases and microarray approaches (Beaudoing et al., 2000; Gautheret et al., 1998; Sandberg et al., 2008; Tian et al., 2005). However, these approaches only allow limited analysis: EST library preparation does not enable analysis of all APA sites, and microarray analysis is based on previously identified sites for probe design (Elkon et al., 2013). Thus, early studies using these techniques underestimated the prevalence of APA and detected APA in not more than 60% of murine and human genes (Elkon et al., 2013). With the technical progress in RNA-seq, global analysis of APA sites and *de novo* discovery was simplified (Pickrell et al., 2010; Wang et al., 2008). In addition, novel bioinformatic analysis pipelines streamlined PAS prediction and APA analysis. However, it also became apparent that standard generation of RNA-seq libraries is disadvantageous for APA analysis (Chen et al., 2017; Elkon et al., 2013; Sun et al., 2012; Wilkening et al., 2013). Only few reads are localized in the 3'-UTR region, and only a limited number thereof is localized directly upstream of PAS, which are needed for bioinformatic prediction of APA patterns. Furthermore, during library amplification and sequencing, polymerase slippage leads to desynchronized, not analyzable reads within one cluster. This slippage is caused by residual parts of the poly(A) tail, which are present in libraries as A/T-stretches due to the oligo(dT)-priming step. All of these limitations hinder an in-depth analysis of APA using standard RNA-seq datasets (Chen et al., 2017; Wilkening et al., 2013). Thus, novel, specific sequencing approaches were developed to enable global APA analysis. In general, these are called 3'-enriched RNA-seq and include PAS-Seq, 3'-Seq and 3'READS (Chen et al., 2017; Elkon et al., 2013). All these methods use different strategies to specifically generate libraries that cover 3'-UTR regions and cleavage sites, for example fragmentation followed by oligo(dT)-priming or custom sequencing primers extending through the poly(A) tail part in sequencing libraries. Reliability of these methods has increased significantly during the last years, and today it is possible to

generate libraries for 3'-Seq analysis that enable global *de novo* APA analysis with high reliability (Chen et al., 2017; Elkon et al., 2013). In the dissertation at hand, a sequencing approach using a custom sequencing primer was used. Thus, read1 is derived from the terminal end of the utilized 3'-UTR, and no bioinformatic reconstitution of PAS usage is needed. The start of the read corresponds to the utilized cleavage site.

Using these novel sequencing approaches, it has become apparent that at least 70% of all mammalian mRNA-encoding genes undergo APA (Derti et al., 2012; Hoque et al., 2013; Wang et al., 2018). Thus, APA is a widely used RNA regulatory mechanism, eventually influencing gene expression patterns.

1.3.2.2 Modes of action of APA

Two general modes of action are known for APA (Figure 1.9) (Di Giammartino et al., 2011). The so-called coding region APA (CR-APA) leads to the generation of different protein isoforms. The different PASs are localized in 3'-UTR regions, introns or sometimes even in exon structures. In CR-APA, the upstream coding sequence exon of the different PASs differs. Therefore, selection of different PASs does not only influence mRNA isoform expression but also protein isoform expression (Di Giammartino et al., 2011). Thus, CR-APA influences gene expression qualitatively. This process is tightly connected to alternative splicing, and competition between APA and splicing has been described.

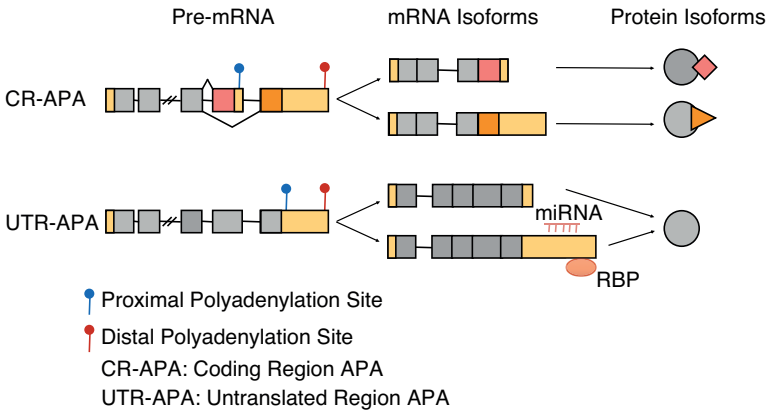


Figure 1.9: CR-APA and UTR-APA.
By using PASs associated with different coding-sequence exons, CR-APA leads to the generation of mRNA isoforms encoding different protein isoforms. UTR-APA leads to the presence of mRNA isoforms varying only in the length of the 3'-UTR, depending on the selection of proximal versus distal PASs. The encoded protein isoform is identical, but differences in 3'-UTR length influence targeting by miRNAs and RNA-binding proteins (RBPs), thereby affecting RNA localization, stability and translation efficiency. Redrawn and modified based on (Di Giammartino et al., 2011).

In addition, so-called untranslated region APA (UTR-APA) occurs. In this case, the protein isoform that is encoded is identical, only the length of the 3'-UTR differs (Di Giammartino et al., 2011). Depending on the selection of a more proximal or more distal PAS, the 3'-UTR can

be shorter or longer. For example, short 3'-UTRs have a median length of 249 nt, while long 3'-UTRs have a median length of 1773 nt (Hoque et al., 2013). In general, it is thought that differences in 3'-UTR length enable fine-tuning of gene expression and protein localization, as the 3'-UTR provides an interaction platform (Di Giammartino et al., 2011; Elkon et al., 2013; Tian and Manley, 2017). The 3'-UTR contains miRNA-target sites, protein binding sites and regulatory sequences. By 3'-UTR choice, presence of these regulatory sequences is determined. Thereby, RNA stability, protein output as well as RNA and protein localization can be controlled (Di Giammartino et al., 2011; Elkon et al., 2013; Tian and Manley, 2017). In general, longer 3'-UTRs harbor more miRNA target sequences and destabilizing sequences, leading to lower stability of these mRNA transcripts (Di Giammartino et al., 2011). However, this hypothesis is still under debate. It has to be kept in mind that multiple PASs are present in most genes and PASs are localized in introns and in the annotated 3'-UTR regions. Thus, CR- and UTR-APA occur jointly, influencing the fate of the encoded mRNAs and proteins in multiple ways.

1.3.3 PAS selection

Today, we know that the majority of genes undergo APA, leading to the expression of different protein isoforms or to the regulation of 3'-UTR length. It is still not well understood how APA is regulated and how different cleavage sites are selected. Multiple layers of regulation seem to act synergistically (Di Giammartino et al., 2011; Elkon et al., 2013; Tian and Manley, 2017). In general, different PASs as well as up- and downstream motifs have varying strength in recruiting the cleavage and polyadenylation machinery. Some PAS motifs, for example, are stronger (e.g. AAUAAA) than others (e.g. CAUAAA), leading to intrinsic differences in motif selection (Hu et al., 2005; Tian et al., 2005). Generally, PASs that are more proximal to the coding sequence exon or localized in introns are considered weaker than those located further downstream (Di Giammartino et al., 2011; Elkon et al., 2013; Tian and Manley, 2017). Usage of stronger versus weaker sites is tightly regulated.

1.3.3.1 APA regulation by 3'-end factors, RNA-binding proteins and splicing factors

Different 3'-end factors have been shown to be involved in the regulation of APA. In general, higher levels of 3'-end factors have been associated with shorter 3'-UTRs. It is thought that higher levels of these processing factors enable recruitment to weaker sites, leading to usage of, for example, cryptic or less strong intronic or more proximal cleavage sites (Di Giammartino et al., 2011; Elkon et al., 2013; Tian and Manley, 2017). One prominent example occurs during B cell maturation. CstF64, a subunit of the CstF complex, is upregulated during the maturation process, leading to usage of a weaker upstream intronic PAS in the IgM heavy transcript. As

a consequence, IgM is no longer membrane-bound but secreted (Takagaki and Manley, 1998; Takagaki et al., 1996). In line, CstF64 silencing has recently also been shown to induce 3'-UTR lengthening in HeLa cells, and its upregulation in different cancer entities is linked to global 3'-UTR shortening (Xia et al., 2014; Yao et al., 2013). Another example is the CPSF subunit Fip1, which upon knockdown leads to 3'-UTR lengthening (Lackford et al., 2014; Li et al., 2015).

Vice versa, downregulation of *Nudt21*, a component of the CFIm complex, leads to shortening of 3'-UTRs (Gruber et al., 2012; Martin et al., 2012). These observations contradict the general hypotheses that downregulation of 3'-end factors is generally associated with 3'-UTR lengthening (Tian and Manley, 2017). However, how *Nudt21* regulates APA is still not understood. In addition, *Pabpn1* has been identified as an APA regulator that enhances usage of distal PASs. *Pabpn1* knockdown leads to global 3'-UTR shortening (de Klerk et al., 2012; Jenal et al., 2012). Overall, the role of 3'-UTR factors in regulating 3'-UTR length and PAS choice seems to be highly versatile. While some seem to promote proximal PASs, others seem to enhance usage of distal PASs. In addition, changes in expression levels of some 3'-UTR factors did not affect APA patterns at all (Tian and Manley, 2017).

Besides 3'-UTR factors, RNA-binding proteins, including hnRNP H and Elav family members, as well as splicing factors have been shown to regulate APA (Di Giammartino et al., 2011; Elkon et al., 2013; Tian and Manley, 2017). Alternative splicing and APA are tightly connected. For example, alternative splicing and CR-APA seem to compete regarding the usage of intronic cleavage sites (Tian et al., 2007). Alternative splicing leads to exclusion of intronic structures harboring PASs, which in turn cannot be used for APA. In line, inhibition of splicing leads to an increase in intronic polyadenylation levels (Li et al., 2015). Thus, it remains to be determined if splicing factors directly or rather indirectly control APA. For example, the splicing factor *Nova2* seems to directly control APA. *Nova2* has been shown to inhibit PAS usage by binding in close proximity, while binding distantly from the PAS enhances respective PAS usage (Licatalosi et al., 2008).

In summary, the availability of different protein factors, including 3'-UTR processing factors, RNA-binding proteins and splicing factors, influence APA. These proteins can either enhance usage of weaker PASs or repress usage of stronger PASs. These proteins exhibit versatile modes of action, and the mechanism of how these proteins regulate APA is still poorly understood.

1.3.3.2 Chromatin status and epigenetic modifications influence APA

Interestingly, accessibility of PASs has been linked to APA. For example, nucleosome depletion is observed in PAS regions and is more pronounced at the site of and downstream of strong PAS (Spies et al., 2009). Thus, nucleosome positioning seems to influence APA by regulating accessibility of specific PASs. However, current studies are purely correlative, and experimental validation of these hypotheses is needed.

In addition to nucleosome positioning, genomic imprinting and thus specific DNA methylation patterns have been linked to APA regulation (Wood et al., 2008). Proximal as well as distal PAS clusters have been identified in the imprinted gene *H13*. These PAS clusters are separated by CpG islands. On the maternal allele, these CpG islands are methylated, and usage of the distal PASs is promoted. On the paternal allele, in contrast, the CpG island is not methylated, and usage of the proximal PASs is observed (Wood et al., 2008).

Overall, nucleosome positioning as well as DNA methylation patterns influence APA (Di Giammartino et al., 2011; Elkon et al., 2013; Tian and Manley, 2017). The mechanisms, however, are still not understood. In addition, it remains to be determined if, for example, histone modifications also play a role in APA regulation.

1.3.3.3 APA regulation by transcription

APA is also regulated by transcription. A kinetic coupling between the Pol II elongation rate and distinct APA patterns has been reported (Elkon et al., 2013; Tian and Manley, 2017). Using Drosophila models, it was shown that slower elongation rates lead to usage of more proximal PASs (Pinto et al., 2011). The idea is that a slower Pol II leads to a longer holding time of the 3'-end processing complex at the proximal PASs, leading to increased usage of these sites (Elkon et al., 2013; Tian and Manley, 2017). Again, detailed mechanistic insights into these processes are still missing.

In addition, the transcription rate and transcription-activating factors have been shown to influence APA (Elkon et al., 2013; Tian and Manley, 2017). Higher gene expression levels have been linked to shorter 3'-UTR length, and certain transcription-activating factors have been shown to promote usage of distinct PASs (Ji et al., 2011; Nagaike et al., 2011; Rosonina et al., 2003). In summary, the transcription process itself regulates APA patterns and is important for 3'-UTR choice.

1.3.3.4 Different layers of regulation control APA

Many different layers of APA regulation have been discovered in recent years, including regulation by RNA-binding proteins, 3'-end processing factors, chromatin and epigenetic regulators as well as transcription itself. However, detailed mechanistic studies on most of these processes are still lacking. In addition, novel layers of regulation are constantly described, including, for example, RNA modifications and RNA secondary structures (Di Giammartino et al., 2011; Elkon et al., 2013; Tian and Manley, 2017). Additional functional and mechanistic studies are needed to facilitate a deeper understanding of APA regulation.

1.3.4 Biological functions of APA

1.3.4.1 APA influences mRNA stability and translation

Biological functions of APA are versatile. Previous studies have shown the relevance of APA in several biological processes, such as cellular quiescence and stress responses, using cell lines (Chang et al., 2015; Elkon et al., 2012; Hollerer et al., 2016). Sandberg et al. reported in 2008 that activation of CD4+ T cells leads to an increased usage of proximal PASs, causing global 3'-UTR shortening (Sandberg et al., 2008). Forced expression of the long 3'-UTR versions led to a reduction in protein expression levels, presumably by miRNA targeting. Conclusively, Sandberg et al. showed that T cell activation is accompanied by 3'-UTR shortening, which in turn leads to an increase in protein expression. This is the first study in which an increase in proliferative status was linked to 3'-UTR shortening and in turn increased protein expression. In line with these reports, shortening of 3'-UTRs and increased protein expression has been reported in other immune cells, including for example macrophages (Jia et al., 2017; Pai et al., 2016). Further, Zheng et al. could show that cellular stress leads to a reduction in 3'-UTR length (Zheng et al., 2018). In contrast, Gruber et al. reported that 3'-UTR shortening in proliferating cells is not associated with changes in protein levels (Gruber et al., 2014b). In line, Spies et al. reported a slightly higher translational efficiency of mRNAs with longer 3'-UTRs (Spies et al., 2013). Novel studies, however, partially contradict these findings. Fu et al. could show in 2018 that mRNAs with shorter 3'-UTRs are enriched in the polysome-bound RNA fraction in most cell lines, indicating higher translation efficiency (Fu et al., 2018). However, they confirmed the findings of Spies et al. using the NIH3T3 cell line from the original study. Due to these inconsistent findings, Fu et al. performed additional analyses and could show that the cell cycle status of NIH3T3 cells is different from most other cell lines, with the majority of NIH3T3 cells in the G0/G1 state of the cell cycle due to contact inhibition. Using cells in different cell cycle phases, they could show that higher protein expression due to enhanced ribosome binding of short 3'-UTRs seems to be a property of actively proliferating cells in the G2/M phase. These findings at least resolve some of the contradictory results of

different studies. 3'-UTR length itself is not always an indicator of translation efficiency. The influence on protein expression seems to be dependent on the intrinsic cellular status of the cells. Actively proliferating cells exhibit shortened 3'-UTRs, in many cases leading to an increase in protein output from these transcripts, most likely due to evasion of miRNA targeting. Interestingly, mediators of cellular activation, like E2F transcription factors and mTOR, lead to 3'-UTR shortening and an increase in protein output (Chang et al., 2015; Elkon et al., 2012).

The 3'-UTR choice does not only influence translation and miRNA-targeting but also RNA-stability. In general, long 3'-UTRs are thought to be less stable due to destabilizing sequences present in the extended 3'-UTR and due to nonsense-mediated mRNA decay (Elkon et al., 2013). However, these hypotheses have been challenged by publications showing that longer 3'-UTRs are only slightly less stable than their shorter counterparts, presumably due to stabilizing sequences present in the longer 3'-UTR version (Elkon et al., 2013; Spies et al., 2013). As for 3'-UTR length and translation efficiency, regulation of mRNA stability by APA seems to be more complex than initially expected and appears to be influenced by the cellular status.

1.3.4.2 APA mediates mRNA and protein localization

APA has been shown to influence nuclear export efficiency and cytoplasmic mRNA localization (Elkon et al., 2013). While the role of APA for nuclear export is only poorly understood, several studies focused on the influence of 3'-UTR choice on subcellular mRNA localization, mainly in neurons. In these studies, it was shown that mRNA transcripts with the longer 3'-UTR version of a gene are transported to dendrites, while the transcripts with the shorter 3'-UTR version are retained in the cellular bodies (An et al., 2008; Andreassi and Riccio, 2009). It is thought that this mechanism is used to increase the protein concentration in certain sites by targeting the encoding mRNAs to the respective locations (Martin and Ephrussi, 2009).

Recently, it was also observed that APA can influence protein localization. In a comprehensive study, Berkovits and Mayr could show that UTR-APA regulates CD47 localization (Berkovits and Mayr, 2015). CD47 protein translated from transcripts with the shorter 3'-UTR is localized in the ER. The longer 3'-UTR version acts as a scaffold molecule, recruiting RNA-binding proteins ELAVL1 and SET to the site of translation, which in turn mediate localization of the newly synthesized CD47 protein to the surface membrane by binding of SET to the cytoplasmic CD47 domain. The group could extend their findings to additional proteins, indicating that protein localization can be determined by 3'-UTR length variation without taking the detour of altering mRNA localization.

1.3.4.3 APA in health and disease

Many studies have investigated the role of APA in different cellular systems, mostly *in vitro*, and in disease settings (Di Giammartino et al., 2011; Elkon et al., 2013; Tian and Manley, 2017). For example, the role of APA during embryonic development and its prevalence in mature cellular populations, including testis, muscle and brain tissue as well as mature immune cells, was analyzed *in vivo* (Boutet et al., 2012; Jereb et al., 2018; Ji et al., 2009; Jia et al., 2017; Lau et al., 2010; Mueller et al., 2016; Singh et al., 2018). In the hematopoietic field, mutations in PASs and deregulated intronic APA have been linked to hematological diseases, including thalassemia (Curinha et al., 2014) and leukemia (Lee et al., 2018). Furthermore, distinct individual genes undergoing APA and having functional consequences in the hematopoietic system and during hematopoietic cellular transformation have been reported (Cesana et al., 2018; Park et al., 2016). Here, I want to highlight some of the recent findings regarding the function of APA in health and disease.

Interestingly, specific APA patterns were detected in immune cells and linked to certain activation states. As previously mentioned, Sandberg et al. could show that T cell activation *in vitro* is associated with 3'-UTR shortening (Sandberg et al., 2008). In line, Gruber et al. observed similar changes in human and murine T cells (Gruber et al., 2014b). Recent studies focusing on immune cells could confirm these findings. In 2016, Pai et al. could show that bacterial infection of macrophages *in vitro* leads to 3'-UTR shortening and thereby evasion of miRNA-mediated degradation and inhibition (Pai et al., 2016). These findings were further supported by a report of Jia et al., who showed that vesicular stomatitis virus infections also lead to 3'-UTR shortening in human and murine macrophages *in vitro* (Jia et al., 2017). All these studies show that immune cell activation is linked to a global shortening of 3'-UTRs, and most studies show that these changes lead to evasion of miRNA regulation and an increase in protein output (see also 1.3.4.1). In 2018, Singh et al. published an important report, showing that APA is also differentially regulated in immune cells *in vivo* (Singh et al., 2018). They performed 3'-Seq of primary human immune cells, among others, and could show that intronic polyadenylation is common, leading to generation of truncated proteins. Thus, immune cells exhibit different APA patterns and seem to actively regulate and depend on differential APA, especially upon activation.

In addition, APA has been studied in different cancer entities (Tian and Manley, 2017). Exemplarily for many other studies, Xia et al. detected global 3'-UTR shortening in seven different tumor types (Xia et al., 2014). This leads to escape from miRNA targeting, and certain APA events can even be used as prognostic factors. By now, it has been acknowledged that cellular transformation is accompanied by changes in APA patterns, mainly 3'-UTR shortening. One could speculate that this is a consequence of the increased cellular proliferation, which

has also been linked to 3'-UTR shortening. However, we do not yet know if differential APA is a cause, a consequence or a prerequisite for cellular transformation. Notably, Lee et al. showed that intronic polyadenylation is increased in chronic lymphocytic leukemia (CLL) (Lee et al., 2018). This increase causes production of truncated proteins, leading to inactivation of tumor suppressor genes. For some of the affected genes, truncation did not only cause loss of tumor suppressor activity but also induced oncogenic potential. This was a landmark study, as it extends our current knowledge of APA in cancer and during cellular transformation, which had been characterized by mainly descriptive and correlative studies.

In recent years, APA has also been analyzed in different developmental stages, in embryonic stem cells and other stem cell populations (Mueller et al., 2013). For example, progressive lengthening of 3'-UTRs was observed during the embryonic development of mice (Ji et al., 2009). The authors hypothesized that this lengthening process enables increased fine-tuning of transcripts by a higher accessibility to post-transcriptional regulation, like miRNAs, in mature cells. Interestingly, reprogramming of differentiated cells towards the induced pluripotent stem cell (iPSC) state is accompanied by global 3'-UTR shortening (Ji and Tian, 2009). Lackford et al. could show that Fip1, a 3' processing factor, is essential for embryonic stem cell (ESC) self-renewal and somatic cell reprogramming (Lackford et al., 2014). Fip1 promotes stem cell maintenance and activates ESC-specific APA profiles. Upon differentiation, *Fip1* downregulation is accompanied by changes in APA profiles and 3'-UTR lengthening. In ESCs, Fip1 promotes usage of proximal PASs, leading to globally shortened 3'-UTRs. In addition to Fip1, other APA mediators have been identified to be important for ESC function, including NUDT21 (Brumbaugh et al., 2018).

Studies on adult stem cells have mainly focused on muscle stem cells and neural development (Boutet et al., 2012; Lau et al., 2010; Mansfield and Keene, 2012; Mueller et al., 2016). In these studies, APA events affecting specific genes have been functionally studied. However, a global in-depth analysis of APA in stem cell hierarchies is still lacking. To date, we do not know how global differential APA is affected by and affects stem cell differentiation, activation and quiescence *in vivo* (Mueller et al., 2013). The hematopoietic system with its very well defined quiescent stem cell and active progenitor populations presents a model system that can be used to study these questions and to draw conclusions that could be representative for other stem cell hierarchies *in vivo*. Studies focusing on changes in polyadenylation patterns of the gene *Hmga2* in HSCs and *Atg7* in MDS and leukemia hint that APA could indeed be a mechanism that plays an important functional role in hematopoiesis and within the hematopoietic hierarchy (Cesana et al., 2018; Park et al., 2016). In summary, the global role of APA in functional stem cell hierarchies remains elusive.

1.4 Generation of genetic mouse models using CRISPR/Cas9

1.4.1 Transgenic mouse models

Transgenic mouse models have been developed to study gene function *in vivo*. For example, global or tissue-specific knockout and overexpression models are available. In addition, these mouse models can also be designed to be inducible, for example by using tamoxifen-dependent Cre recombinases. Furthermore, specific mutations can be introduced, enabling the study of these mutations in disease settings. For studying gene expression patterns in different tissues and identification of cells expressing the respective genes, fluorescent reporter mice can be generated. These are only some of the possibilities to employ transgenic animals to better understand complex processes *in vivo* in health and disease.

Many different techniques have been developed to generate transgenic mouse models. Most commonly, homologous recombinant ESC clones are used to generate transgenic mice. ESCs can be cultured *in vitro* and can be genetically modified. A targeting vector is used to introduce the transgene by homologous recombination. Subsequently, successfully edited clones are selected, e.g. by resistance against a specific selection agent. These clones are analyzed in depth to ensure correct targeting of the sequence of interest and can then be injected into blastocysts, leading to the generation of chimeric animals, which can subsequently be used to generate homogenous transgenic animals by crossing. (Bouabe and Okkenhaug, 2013; Miao, 2013)

This approach is, however, highly time-consuming. Today, transgenic animal models can be generated in an efficient and convenient way using the CRISPR/Cas9 system.

1.4.2 The CRISPR/Cas9 system

The clustered regularly interspaced short palindromic repeat (CRISPR) system was first discovered as an antiviral mechanism in bacteria, where it cleaves foreign nucleic acids (Fineran and Charpentier, 2012; Wiedenheft et al., 2012). This system has been adapted to enable sequence-specific induction of DNA double-strand breaks (DSB) for targeted genome editing (Jinek et al., 2012).

In this system, a so-called single-molecule guide RNA (gRNA) is used to direct a nuclease, namely Cas9 (derived from *S. pyogenes*), to a specific target DNA sequence (Figure 1.10) (Sander and Joung, 2014). The gRNA is a fusion of the CRISPR RNA (crRNA), which directs the Cas9 to the target site, and the transactivating CRISPR RNA (tracrRNA), which is bound by Cas9. However, nowadays some systems start using the two-molecule crRNA:tracrRNA

hybrids as originally detected in bacteria, since this approach seems to be more efficient than single-molecule gRNAs. By complementary base-pairing, the gRNA-Cas9 complex is specifically directed to the target site. A prerequisite for this target site is the presence of a protospacer adjacent motif (PAM) immediately downstream of the target sequence. The canonical PAM signal is NGG, thus the Cas9 complex can be targeted to any N_{20}-NGG DNA sequence. Today, additional Cas9 nucleases from bacteria have been detected that utilize other PAM sequences. Therefore, the number of genomic regions that can be targeted is steadily increasing. (Sander and Joung, 2014)

After binding to the target site, the Cas9 complex cleaves the DNA, generating a DNA DSB. This introduces insertion-deletion (indel) mutations by non-homologous end-joining. These mutations can lead to amino acid deletions in the encoded proteins, insertions or frameshift mutations causing premature stop codons within the open reading frame (ORF) of the targeted gene, eventually leading to a gene knockout. By providing template DNA strands, the Cas9-induced DSB can be used for homology-directed repair (HDR). Thereby, specific gene sequences, like point mutations or even fluorescent reporter sequences, can be introduced. Furthermore, so-called dead Cas9 (dCas9), which has no nuclease activity, has been fused to activation or effector domains. With the help of the gRNA, the complex is targeted to a specific region, where it leads to gene activation or changes of chromatin or DNA modifications using the additional domains. (Sander and Joung, 2014)

In summary, the CRISPR/Cas9 approach is a highly flexible method to target and edit specific genomic sequences. Adaptations of this system extended its usage beyond the initial introduction of relatively simple indel mutations. However, care has to be taken during gRNA design to minimize off-target cleavage events.

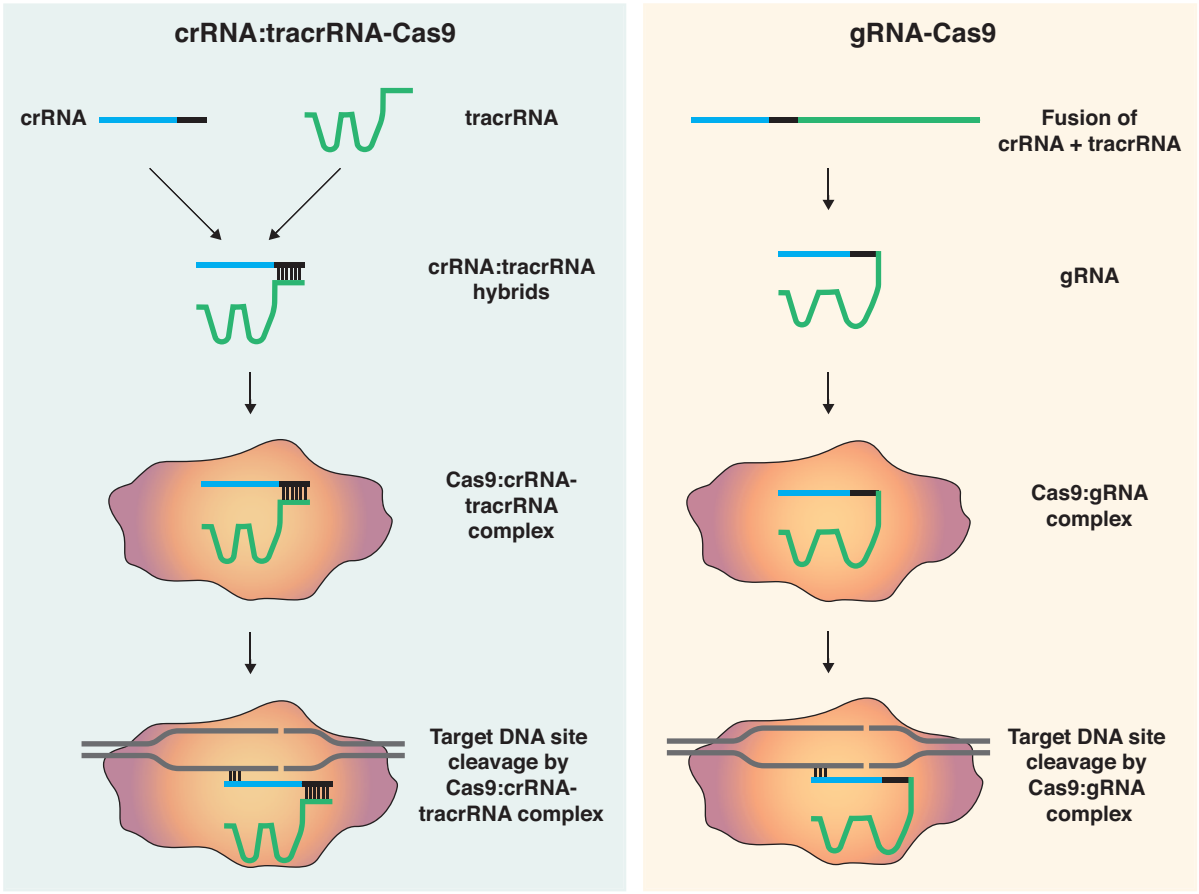


Figure 1.10: The CRISPR/Cas9 system.
The crRNA:tracrRNA hybrid assembles by base pairing (green box). The crRNA is complementary to the DNA target site, and the tracrRNA is important for complex formation with the Cas9 nuclease. Alternatively, a single-molecule gRNA can be used (orange box). In both cases, a crRNA-tracrRNA/gRNA:Cas9 complex assembles. This complex is directed to the target site by base pairing. The presence of a PAM sequence is a prerequisite for CRISPR/Cas9 targeting. The Cas9 nuclease induces a DNA DSB at the target site. Redrawn and modified based on (Sander and Joung, 2014).

1.4.3 *Easi*-CRISPR

Different approaches have been used to employ CRISPR/Cas9-mediated genome editing to generate knockout mice and targeted transgenic mice (Burgio, 2018; Liu et al., 2017). Here, I aim to introduce one of the many described systems for creating knock-in mice using CRISPR/Cas9, namely efficient additions with ssDNA inserts-CRISPR (*Easi*-CRISPR) (Miura et al., 2018; Quadros et al., 2017).

In this method, crRNA and tracrRNA molecules are precomplexed with Cas9. Long single-stranded DNA (ssDNA) donor molecules with short homology arms of only around 100 bp are used as templates (Figure 1.11). This method can for example be used to introduce loxP sites or other knock-in alleles. For introduction of loxP sites, two different crRNA-tracrRNA:Cas9 complexes are employed that cut upstream and downstream of the region of interest, respectively. The ssDNA template contains a short upper homology arm, the upper loxP site, the target exon, the lower loxP site and a short lower homology arm. The crRNA-

tracrRNA:Cas9 complexes are microinjected into one of the pronuclei of a zygote together with the long ssDNA template. The offspring can then be genotyped to identify founder animals. Using this technique, inserts of up to 1.5 kb can be introduced, and targeting efficiency is usually between 30-60%. As all custom-made components needed for the injection mix can be purchased, *Easi*-CRISPR allows fast generation of transgenic knock-in mice in only ~3 months of time.

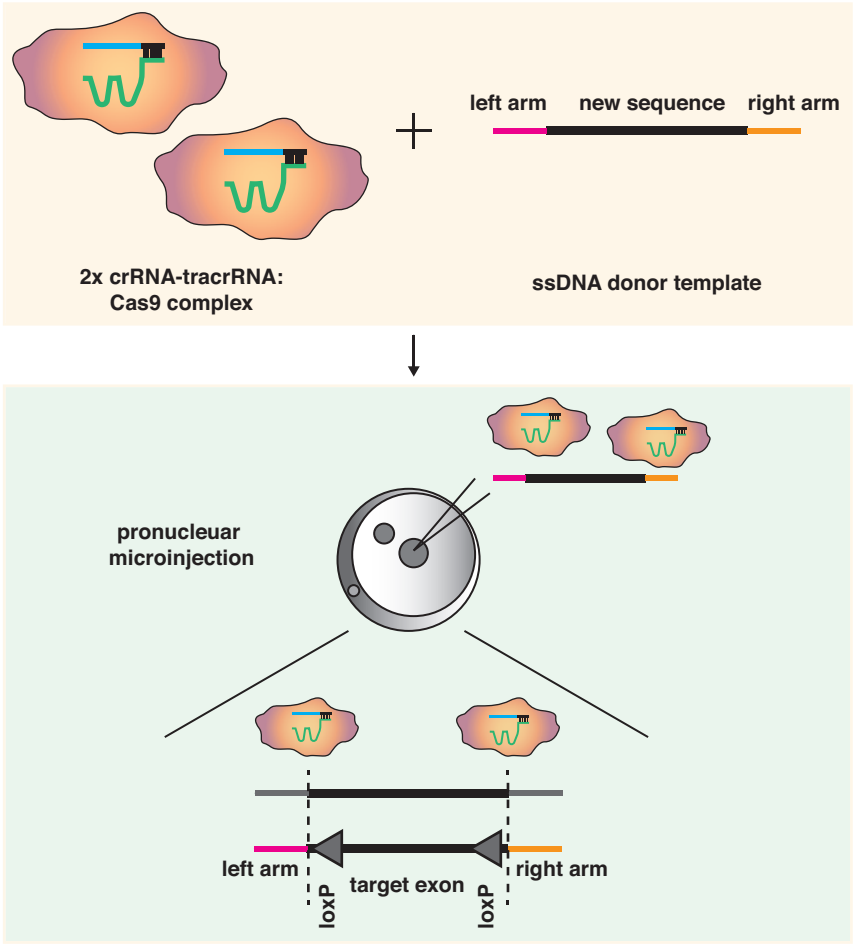


Figure 1.11: Schematic representation of *Easi*-CRISPR.
For introduction of loxP sites, 2 crRNA-tracrRNA:Cas9 complexes are precomplexed, targeting the upper and the lower integration site, respectively. In addition, a ssDNA template is provided that harbors short (~100 bp) upper and lower homology arms, the loxP sites and the target exon. These components are injected into one of the pronuclei of a zygote. In the pronucleus, the Cas9 complexes induce two DSBs and the template sequence containing the loxP sites is integrated via HDR. Redrawn and modified based on (Quadros et al., 2017).

2 Aim of the thesis

Omics analyses have recently been applied to HSCs and MPP populations to further investigate regulatory processes. The results revealed that not only proteins but also splicing variants and lncRNAs show differential expression (Cabezas-Wallscheid et al., 2014; Klimmeck et al., 2014; Luo et al., 2015). These observations suggest that expression of non-coding RNAs and regulation of RNA biogenesis could be crucial for a functional hematopoietic system. In this thesis, I aimed to uncover the role of lncRNAs and RNA regulatory mechanisms in HSCs and hematopoiesis.

The role of lncRNAs in different biological systems is attracting increasing attention (Morris and Mattick, 2014). Specific lncRNAs have been shown to be essential for HSC function and hematopoiesis (Alvarez-Dominguez and Lodish, 2017; Luo et al., 2015; Yildirim et al., 2013). Cabezas-Wallscheid et al. showed that the lncRNA *Meg3* is highly and specifically expressed in the HSC compartment (Cabezas-Wallscheid et al., 2017; Cabezas-Wallscheid et al., 2014). Therefore, we hypothesized that *Meg3* plays an important regulatory role in hematopoiesis. Recently, Qian et al. reported that the *Dlk1-Meg3* locus has important functions in fetal hematopoiesis (Qian et al., 2016). Using fetal liver HSCs, they showed that long-term repopulation capacity of these cells is impaired upon loss of imprinting of the maternally inherited *Meg3* allele. As fetal liver HSCs greatly differ from adult HSCs (Bowie et al., 2006; Nygren et al., 2006; Wilson et al., 2008), the role of *Meg3* and the *Dlk1-Meg3* locus in adult HSCs remains elusive. Thus, I aimed to uncover the function of *Meg3* in the adult hematopoietic compartment (Figure 2.1, blue part). To this end, I analyzed the hematopoietic system of induced MxCre *Meg3* $^{\text{mat flox/pat wt}}$ mice. To assess HSC functionality, I performed *in vitro* CFU assays and BM transplantations. In addition, I intended to uncover the role of *Meg3* upon acute inflammatory stress and after recovery.

In recent years, regulation of DNA methylation, gene expression, enhancer regions, protein translation as well as alternative splicing have been shown to be crucial for HSC self-renewal, function and differentiation (Bahr et al., 2018; Cabezas-Wallscheid et al., 2014; Goldstein et al., 2017; Ji et al., 2010; Lipka et al., 2014; Signer et al., 2014). We know that multiple layers of regulation act synergistically to tightly control HSC fate and function. Perturbations in those tightly regulated systems can lead to stem cell failure or even development of hematologic malignancies. An additional regulator of cells is the newly discovered RNA regulatory mechanism APA (Tian and Manley, 2017). It is still poorly understood if APA is functionally important for stem cells *in vivo*. In addition, it is not known if differential APA is essential for

stem cell differentiation and activation. In this thesis, I aimed to investigate if APA is functionally important for hematopoiesis and HSCs (Figure 2.1, yellow part). By using lentiviral knockdown of APA regulators, I analyzed the functional role of this process in hematopoiesis. In addition, I intended to perform 3'-Seq on homeostatic and activated HSPCs to investigate global APA changes upon HSC differentiation and activation. In the following, I aimed to functionally validate differential APA hits. This included APA-mediated activation of glutaminolysis in the HSC compartment upon HSC activation and differentiation. In addition, I aimed to establish a transgenic mouse model using CRISPR/Cas9 that would enable inducible knockout of the prominent APA regulator *Pabpn1 in vivo* (Figure 2.1, green part).

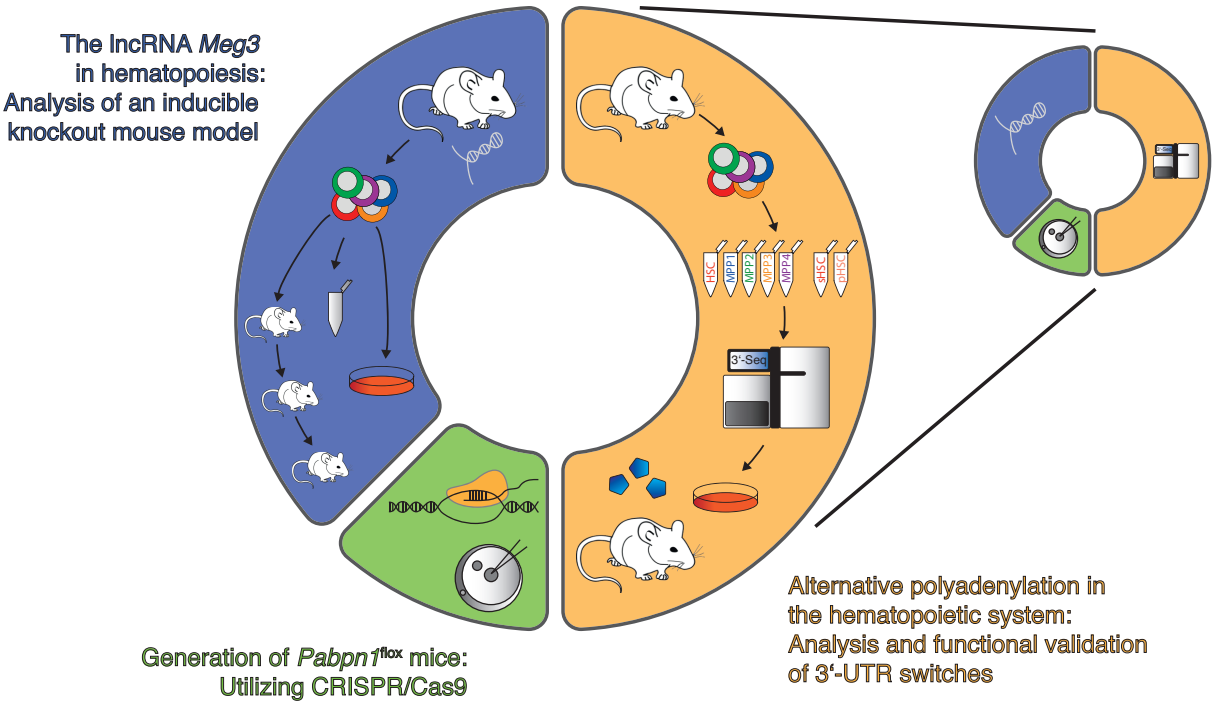


Figure 2.1: Visual summary of study aims.

3 Results

3.1 The lncRNA *Meg3* is dispensable for HSCs

3.1.1 *Meg3* expression is specifically upregulated in the HSC compartment

Previous studies found the lncRNA *Meg3* to be very highly and specifically expressed in HSCs (Cabezas-Wallscheid et al., 2017; Cabezas-Wallscheid et al., 2014; Klimmeck et al., 2014). To confirm this expression pattern, we sorted HSCs and MPP cells gating for lineage-negative (CD4, CD8, B220, Ter119, CD11b, Gr1) cells using cKit, Sca-1, CD150, CD48 and CD34 as markers and isolated total RNA (Figure 3.1 A-B). Analysis of previously generated RNA-seq data showed *Meg3* to be exclusively expressed in HSCs compared to progenitor cells (Figure 3.1 C). We could validate this expression pattern by quantitative real-time PCR (qPCR) analysis (Figure 3.1 D). In addition, analysis of *Meg3* expression in HSCs in different developmental stages was performed (Figure 3.1 E). We observed a significant decrease in *Meg3* levels in adult HSCs (8-12 weeks old mice) compared to fetal liver HSCs (E13.5; CD48- CD41+ cKit+ CD34+). However, no significant decrease in *Meg3* levels in the HSC compartment upon aging of adult mice was detected (24 months old mice). Thus, *Meg3* is highly and specifically expressed in HSCs independently of the developmental stage, but a slight reduction in overall levels is observed during development and aging.

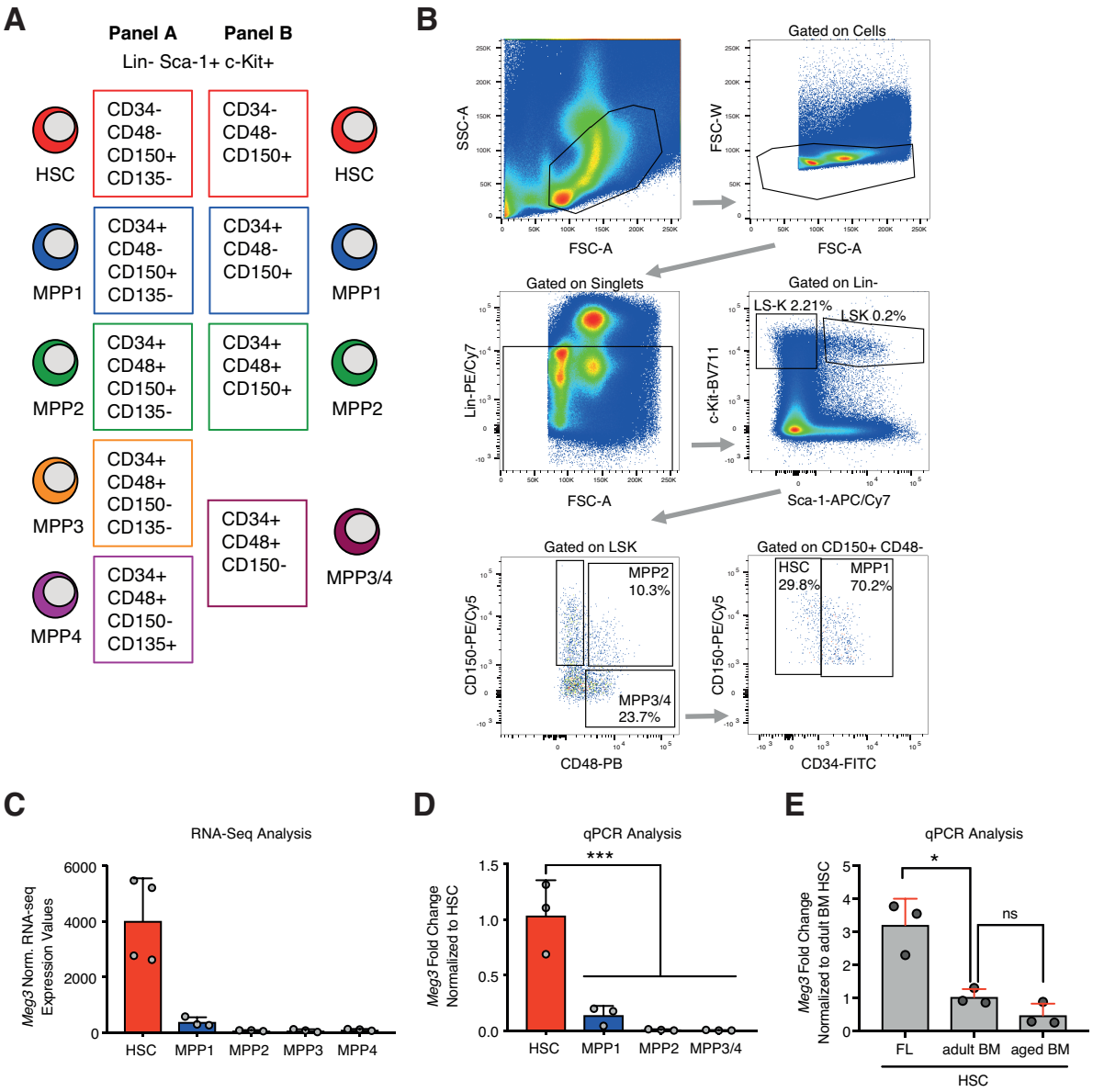
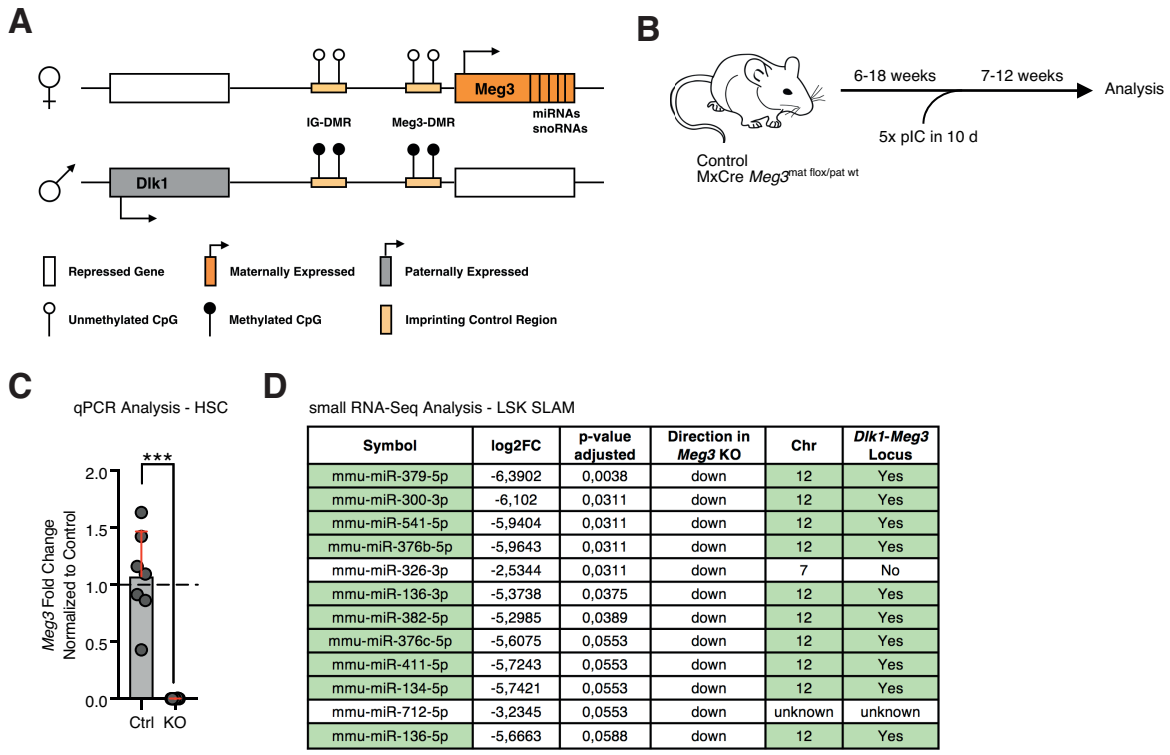


Figure 3.1: *Meg3* expression is specifically upregulated in the HSC compartment.
(A) Surface marker profiles of HSCs and MPPs used in flow cytometry analysis and sorting of BM cells. CD135 was only included for analysis in Fig. 1 C and 1 E (Panel A), for all other experiments Panel B was used. (B) Representative flow cytometry dot plots showing the gating strategy for HSCs and MPP1-4 cells using Lin, c-Kit, Sca-1, CD48, CD150 and CD34 as markers. (C) Analysis of *Meg3* expression by RNA-seq. Data derived from (Cabezas-Wallscheid et al., 2014). N = 3-4. (D) Analysis of *Meg3* expression by qPCR. N = 3, one-way ANOVA. (E) qPCR-based *Meg3* expression analysis in fetal liver (FL) at E13.5, adult BM and aged BM HSCs. N = 3, student's t-test. For all panels: *p < 0.05, **p < 0.01, ***p < 0.001, ns: not significant; N indicates number of biological replicates; 3 independent experiments were performed; mean + SD is shown. Figure is adapted from (Sommerkamp et al., 2019).

3.1.2 Loss of *Meg3* expression does not impair adult homeostatic hematopoiesis

3.1.2.1 Successful deletion of Meg3 in MxCre Meg3 mat flox/pat wt mice after Cre induction

Previously, the *Dlk1-Meg3* locus has been shown to be essential for fetal liver HSC function (Qian et al., 2016). Fetal liver HSCs and adult BM HSCs exhibit differences in cellular properties, for example fetal liver HSCs are actively cycling in contrast to adult BM HSCs (Bowie et al., 2006; Nygren et al., 2006; Wilson et al., 2008). Thus, after confirming *Meg3* expression patterns in the adult HSC/MPP compartment, we aimed to investigate the role of *Meg3* in adult hematopoiesis. To this end, we utilized *Meg3* mat flox/pat flox mice kindly provided by the laboratory of Dr. Klibanski, Boston. In these mice, exon 1 to 4 of the *Meg3* gene are floxed. As the *Meg3* locus is imprinted and *Meg3* is only expressed from the maternally inherited allele harboring unmethylated DMRs, we crossed male MxCre driver mice to female *Meg3* mat flox/pat flox mice, thereby generating MxCre *Meg3* mat flox/pat wt mice (from now on called *Meg3* KO) (Figure 3.2 A). *Meg3* KO mice and MxCre+ control mice were repeatedly injected with pIC 6 to 18 weeks after birth to induce expression of Cre (Figure 3.2 B). pIC injections lead to an inflammatory response activating the *Mx1* promoter, which controls Cre expression in hematopoietic cells (Kuhn et al., 1995). Analysis was performed after 7 to 12 weeks, allowing recovery of the hematopoietic system to a homeostatic state. First, we aimed to validate loss of *Meg3* expression in induced *Meg3* KO mice. Hence, we sorted HSCs from control and KO animals, isolated RNA and performed qPCR analysis (Figure 3.2 C). We observed complete loss of *Meg3* expression in induced MxCre *Meg3* KO animals, confirming the efficiency of our KO approach. Deletion of *Meg3* and loss of *Meg3* transcription has been shown to abolish expression of the downstream maternally expressed miRNA cluster (Lin et al., 2003; Zhou et al., 2010). To analyze expression of the miRNA cluster in our KO model, we performed small RNA-seq analysis using LSK SLAM cells (LSK CD150+ CD48-) (Figure 3.2 D). In total, around 49 miRNAs are associated with the *Meg3* cluster (Seitz et al., 2004; Winter, 2015). In our control dataset, we only observed 9 miRNAs of this cluster to be stably expressed. Presumably, detection of the other miRNAs of the cluster was not possible due to limited sequencing depth. Strikingly, we detected only 11 miRNAs to be differentially expressed in *Meg3* KO compared to control cells, with 9 of these miRNAs being the initially identified *Meg3* cluster miRNAs. All of these miRNAs were strongly downregulated. In summary, we successfully deleted *Meg3* expression and expression of the associated miRNA cluster by targeting the maternal allele in the adult hematopoietic compartment.



Symbol	log2FC	p-value adjusted	Direction in *Meg3* KO	Chr	*Dlk1-Meg3* Locus
mmu-miR-379-5p	-6,3902	0,0038	down	12	Yes
mmu-miR-300-3p	-6,102	0,0311	down	12	Yes
mmu-miR-541-5p	-5,9404	0,0311	down	12	Yes
mmu-miR-376b-5p	-5,9643	0,0311	down	12	Yes
mmu-miR-326-3p	-2,5344	0,0311	down	7	No
mmu-miR-136-3p	-5,3738	0,0375	down	12	Yes
mmu-miR-382-5p	-5,2985	0,0389	down	12	Yes
mmu-miR-376c-5p	-5,6075	0,0553	down	12	Yes
mmu-miR-411-5p	-5,7243	0,0553	down	12	Yes
mmu-miR-134-5p	-5,7421	0,0553	down	12	Yes
mmu-miR-712-5p	-3,2345	0,0553	down	unknown	unknown
mmu-miR-136-5p	-5,6663	0,0588	down	12	Yes

Figure 3.2: Successful deletion of *Meg3* in MxCre *Meg3*$^{mat\ flox/pat\ wt}$ mice after Cre induction.
(A) Schematic representation of the *Dlk1-Meg3* locus. (B) Workflow of pIC-mediated Cre induction in control and MxCre *Meg3*$^{mat\ flox/pat\ wt}$ mice and consecutive analysis. (C) Analysis of *Meg3* expression in sorted control and *Meg3* mat KO HSCs. N = 7, unpaired student's t-test. (D) Differential mature miRNA expression analysis by small RNA-seq in sorted LSK CD150+ CD48- cells (LSK SLAM). Adjusted p-value <0.1 is shown. For all panels: *p < 0.05, **p < 0.01, ***p < 0.001, ns: not significant; N indicates number of biological replicates; 3 independent experiments were performed; mean + SD is shown. Figure is adapted from (Sommerkamp et al., 2019).

3.1.2.2 Production and frequency of mature blood cells in primary and secondary hematopoietic organs is not affected by loss of Meg3 expression

First, we aimed to investigate the effects of *Meg3* KO in the hematopoietic system on the frequency of mature blood cells. Seven to twelve weeks after deletion, peripheral blood was analyzed by flow cytometry (Figure 3.3 A). We did not observe changes in the frequency of mature B cells (B220+), T cells (CD4/8+) and myeloid cells (Gr1+ CD11b+) (Figure 3.3 B). In addition, we analyzed mature cell frequencies and B cell maturation in the BM by using IgM expression as an additional marker (pre/pro B cells: B220^{low} IgM-; immature B cells: B220^{low} IgM+; mature B cells: B220^{high} IgM+) (Figure 3.3 C). Again, we did not observe major differences (Figure 3.3 D). Only the frequency of immature B cells was slightly increased in *Meg3* KO animals. Analysis of secondary hematopoietic organs also did not exhibit any changes upon loss of *Meg3* expression, as mature cell frequencies in the spleen and T cell maturation in the thymus were unaffected (Figure 3.3 E-F). T cell maturation was analyzed by investigating the CD4- CD8- compartment using CD25 and CD44 to identify DN1-DN4 populations. In conclusion, mature cell frequencies in hematopoietic organs and the peripheral blood are not perturbed by loss of *Meg3* and the associated miRNA cluster in the hematopoietic system.

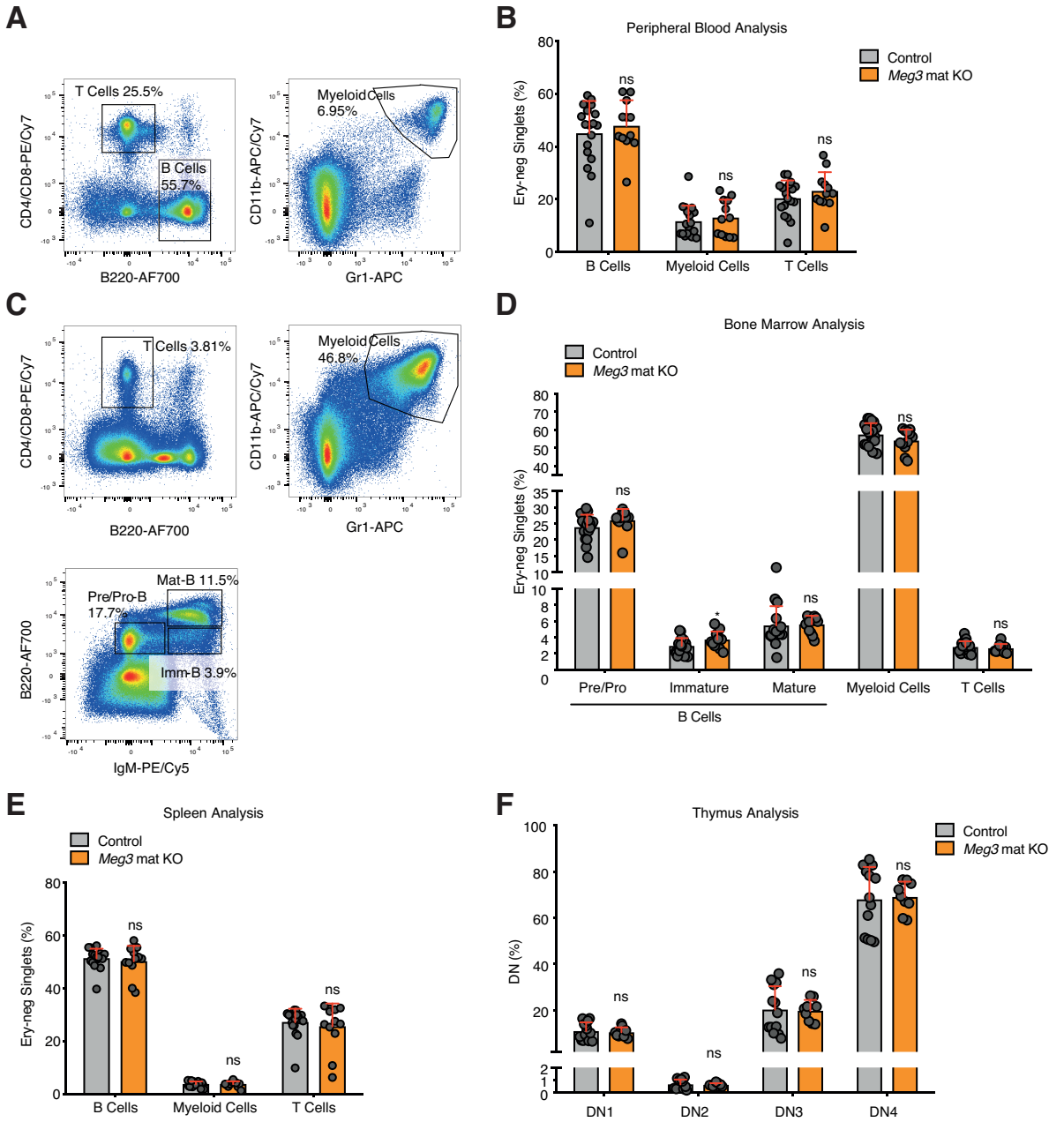


Figure 3.3: Production and frequency of mature blood cells in primary and secondary hematopoietic organs is not affected by loss of *Meg3* expression.
(A) Representative flow cytometry dot plots of PB analysis. (B) Flow cytometry-based analysis of mature cells in the peripheral blood. N = 12-18, unpaired student's t-test. (C) Representative flow cytometry dot plots showing analysis of differentiated BM cells. (D) Flow cytometry-based analysis of B cell maturation and mature cells in the BM compartment. N = 12-18, unpaired student's t-test. (E) Flow cytometry-based analysis of mature cells in the spleen. N = 12-18, unpaired student's t-test. (F) Flow cytometry-based analysis of T cell maturation in the thymus. N = 9-13, unpaired student's t-test. For all panels: *p < 0.05, **p < 0.01, ***p < 0.001, ns: not significant; N indicates number of biological replicates; 3 independent experiments were performed; mean + SD is shown. Figure is adapted from (Sommerkamp et al., 2019).

3.1.2.3 HSPC frequencies are unchanged upon hematopoietic deletion of Meg3

To investigate direct effects of *Meg3* loss on the HSC/MPP compartment, we performed flow cytometry analysis of BM cells using Lin, cKit, Sca-1, CD150, CD48 and CD34 as markers (see 3.1.1). The BM frequency of HSCs and MPP1-3/4 cells was unchanged (Figure 3.4 A). Furthermore, we aimed to analyze the cell cycle status of control and KO cells, as *Meg3* has been described to inhibit proliferation (Braconi et al., 2011; Wang et al., 2012; Zhang et al., 2010a; Zhang et al., 2003; Zhou et al., 2007). We used intracellular Ki-67 staining and the DNA dye Hoechst 33342 to analyze the cell cycle status of HSCs and MPP cells (G0: Ki-67- 1-chromatide chromosomes (1-C-C); G1: Ki-67+ 1-C-C; S/G2/M: Ki-67+ 2-C-C) (Figure 3.4 B). As previously described, HSC control cells were highly quiescent and mainly in the G0 state of the cell cycle (>80%; Figure 3.4 C) (Wilson et al., 2008). As expected, MPP cells became increasingly more activated, with MPP2 cells exhibiting the highest frequency of cells in the S/G2/M phase (>20%) (Wilson et al., 2008). *Meg3* KO HSCs did not exhibit any alterations in the cell cycle status. However, we detected slightly increased cellular quiescence in MPP1-3/4 cells. To analyze the molecular profile of HSCs in response to *Meg3* KO, we again sorted control and KO HSCs and performed qPCR analysis (Figure 3.4 D). We analyzed expression patterns of previously described HSC-specific genes and HSC regulatory genes (Cabezas-Wallscheid et al., 2017). Except for a minor reduction in *Cdk6* levels, we did not observe any deregulated expression of HSC key regulatory genes upon loss of *Meg3*. Collectively, no major changes in the HSC/MPP compartment in *Meg3* KO animals were detected. In addition, the molecular profile of HSCs remained unchanged.

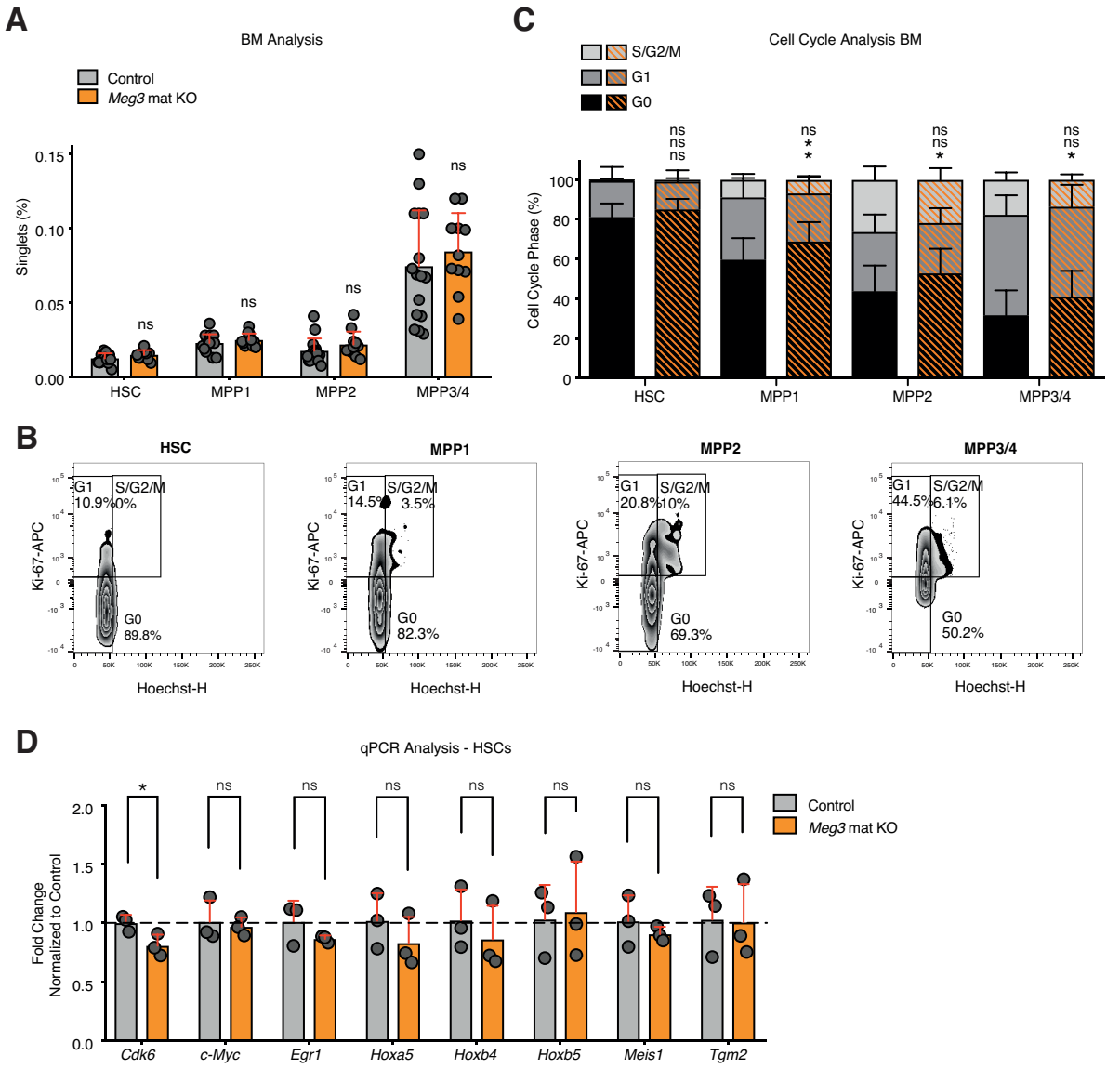


Figure 3.4: HSC and MPP frequencies are unchanged upon hematopoietic deletion of *Meg3*.
(A) Flow cytometry-based analysis of the HSC/MPP compartment in the BM. Percentage of BM HSCs/MPPs is shown. N = 11–16, unpaired student's t-test. (B) Representative flow cytometry zebra plots (including outliers) showing cell cycle analysis of BM HSC/MPP populations. (C) Analysis of cell cycle status in HSCs/MPPs. Relative percentage within each cell cycle phase is depicted. N = 9–15, two-way ANOVA. Shaded in orange: *Meg3* KO (D) qPCR-based analysis of HSC signature gene expression in sorted control and *Meg3* mat KO HSCs. N = 3, unpaired student's t-test. For all panels: *p < 0.05, **p < 0.01, ***p < 0.001, ns: not significant; N indicates number of biological replicates; 3 independent experiments were performed; mean + SD is shown. Figure is adapted from (Sommerkamp et al., 2019).

3.1.2.4 HSC function in vitro and long-term engraftment in vivo is not impaired upon loss of Meg3 expression

Hematopoietic defects in the HSC/MPP compartment might not become apparent in adult mice 7 to 12 weeks after *Meg3* deletion, as contribution to mature cellular output is low and production of blood cells can be maintained by more committed progenitors for a limited timeframe (Busch et al., 2015; Sun et al., 2014). To asses HSC function, we performed *in vitro* CFU assays and *in vivo* transplantation experiments (Figure 3.5 A). Functional HSCs are able to form colonies *in vitro* in primary and secondary platings. *In vivo*, they engraft in lethally irradiated recipient mice and contribute to peripheral blood production even in tertiary transplants (Purton and Scadden, 2007). Nine weeks after *Meg3* deletion, total bone marrow cells were isolated, and colony forming capacity was investigated. Control as well as KO cells formed colonies in primary and secondary CFU assays, and no significant differences were observed (Figure 3.5 B). To generate full chimeras, $1x10^6$ control or KO total BM cells (CD45.2+) were transplanted into lethally irradiated recipients (CD45.1+) by intravenous injection. The allelic variants of CD45 (CD45.1 and CD45.2) can be distinguished in flow cytometry analysis using specific antibodies and allow differentiation between cellular contribution of recipient and donor (Figure 3.5 C) (Purton and Scadden, 2007). Analysis of the peripheral blood exhibited equal contribution of control as well as KO BM cells to mature cellular output in primary recipients (Figure 3.5 D). Serial transplantation was performed by retransplanting $3x10^6$ BM cells into lethally irradiated recipient mice. In secondary and tertiary recipients, we again did not observe any functional defects in *Meg3* KO HSCs (Figure 3.5 D). Analysis of mature cellular lineages, including B cells, T cells and myeloid cells, did not exhibit any defects of *Meg3* KO cell contribution to mature lineage output in primary, secondary and tertiary recipient mice over time (Figure 3.5 E-G).

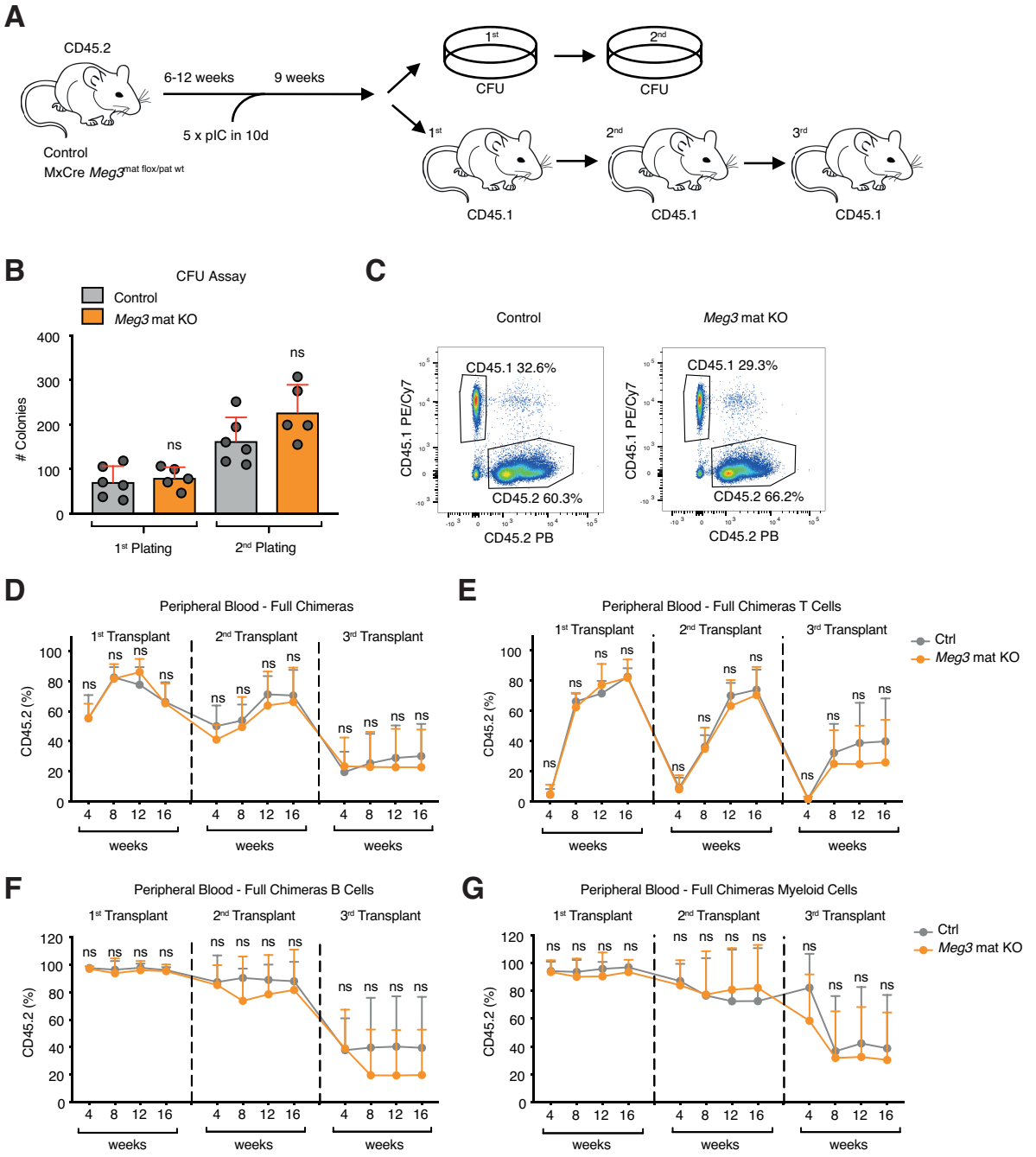


Figure 3.5: HSC function *in vitro* and long-term engraftment *in vivo* is not impaired upon loss of *Meg3*.
(A) Workflow depicting Cre induction by pIC and functional analysis by CFU assays and generation of full chimeras. (B) Serial CFU assays. N = 5–6, unpaired student's t-test. (C) Representative flow cytometry dot plots showing PB analysis of full chimeras. (D) Flow cytometry-based analysis of serial full transplantation experiments. CD45.2% outcome is shown. N = 9, two-way ANOVA. (E-G) Flow cytometry-based analysis of serial full transplantation experiments. CD45.2% outcome is shown for T cells (E), B cells (F) and myeloid cells (G). N = 7-9, two-way ANOVA. For all panels: *p < 0.05, **p < 0.01, ***p < 0.001, ns: not significant; N indicates number of biological replicates; B: 2 independent experiments were performed; D-G: 1 independent experiment was performed; mean + SD is shown. Figure is adapted from (Sommerkamp et al., 2019).

In addition to peripheral blood, we also performed analysis of BM and spleen cells of recipient animals. Here, analysis of primary recipients is shown exemplary (Figure 3.6). First, we investigated the frequency of transplanted HSCs and MPP cells by using HSC/MPP markers and analyzing CD45.2 contribution (Figure 3.6 A). We did not observe any changes in HSC/MPP engraftment of *Meg3*-deficient cells. Ki-67/Hoechst 33342 cell cycle analysis of

engrafted HSCs and MPP cells did not display significant changes in cycling activity in KO cells (Figure 3.6 B). Only MPP3 KO cells showed a slight increase in G1 and a slight decrease in S/G2/M. To confirm persistent *Meg3* deficiency, we sorted CD45.2+ HSCs of primary recipients and analyzed *Meg3* expression by qPCR analysis (Figure 3.6 C). No *Meg3* expression was detected in transplanted KO HSCs, verifying stability of our KO approach. In addition, we analyzed cellular contribution in the BM and spleen (Figure 3.6 D-E). Again, we did not observe changes in CD45.2+ contribution in hematopoietic organs. The same analysis of BM HSC/MPP and mature BM and spleen cells was also performed using material derived from secondary and tertiary recipients (data not shown). In line, no changes were observed in response to loss of *Meg3* expression. In summary, *in vitro* function as well as long-term engraftment potential of HSCs is not affected upon KO of *Meg3* and loss of expression of the associated miRNA cluster.

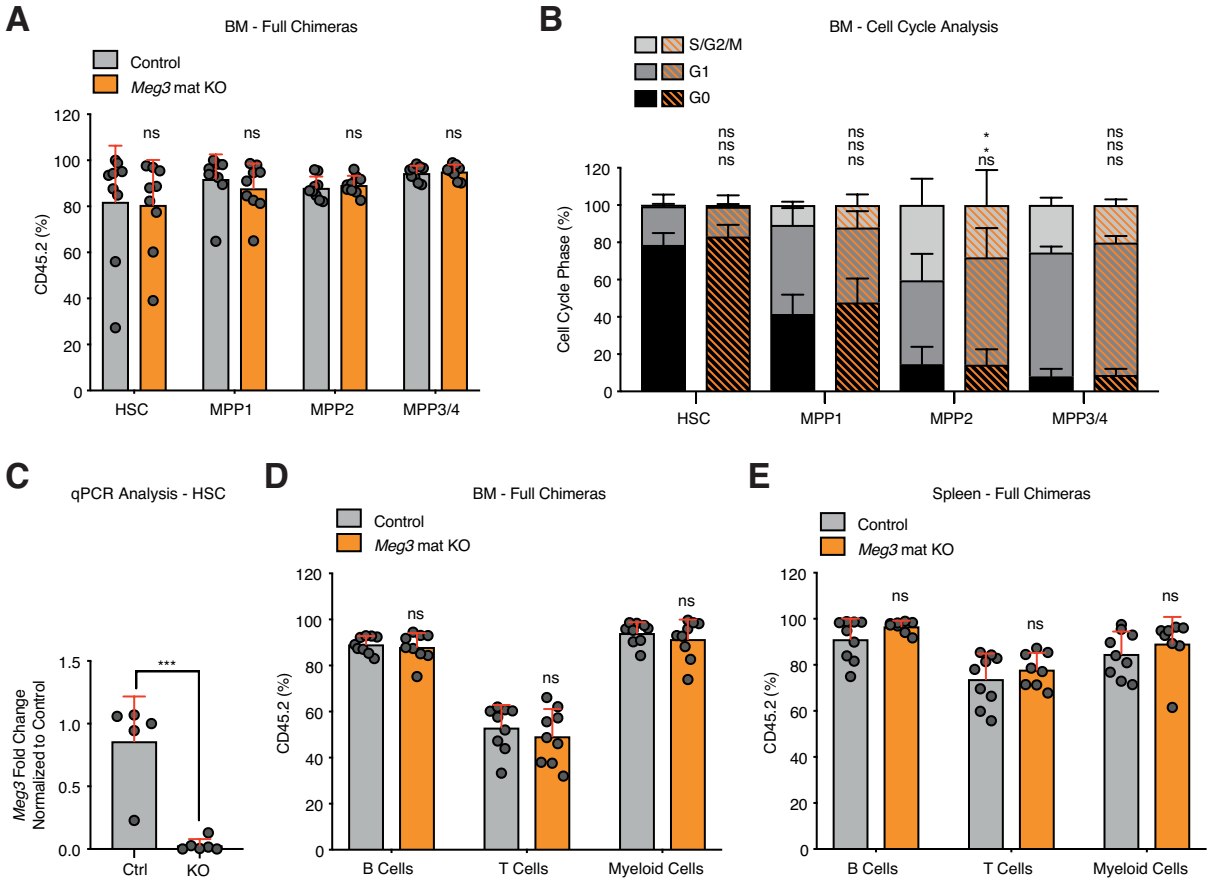


Figure 3.6: HSC long-term engraftment *in vivo* is not impaired upon loss of *Meg3*.
(A) Flow cytometry-based analysis of BM of primary recipients 18 weeks after transplantation showing the HSC/MPP compartment. CD45.2% outcome is shown. N = 9, unpaired student's t-test. (B) Cell cycle analysis of transplanted CD45.2+ HSCs/MPPs. Relative percentage within each cell-cycle phase is shown. N = 6–9, two-way ANOVA. Shaded in orange: *Meg3* KO (C) Analysis of *Meg3* expression by qPCR in sorted CD45.2+ control and *Meg3* mat KO HSCs from primary recipient animals. N = 5-6, unpaired student's t-test. (D) Flow cytometry-based analysis of mature cells in the BM of primary recipients 18 weeks after transplantation. CD45.2% outcome is shown. N = 9, unpaired student's t-test. (E) Flow cytometry-based analysis in the spleen of primary recipients 18 weeks after transplantation. Mature cells were analyzed. CD45.2% outcome is shown. N = 9, unpaired student's t-test. For all panels: *p < 0.05, **p < 0.01, ***p < 0.001, ns: not significant; N indicates number of biological replicates; 1 independent experiment was performed; mean + SD is shown. Figure is adapted from (Sommerkamp et al., 2019).

3.1.2.5 Loss of Meg3 expression does not impair the competitive potential of HSCs

If HSC function is only slightly impaired or HSC expansion kinetics are affected, defects might not become apparent in full chimera transplantation experiments (see 3.1.2.4). To ensure that we do not overlook minor changes in HSC potential upon loss of *Meg3* expression, we additionally set up competitive transplantation experiments. $2x10^5$ control or KO BM cells (CD45.2) were mixed with $2x10^5$ competitor BM cells (CD45.1/.2) and transplanted into lethally irradiated recipients (CD45.1) (Figure 3.7 A). Analysis of the peripheral blood was performed approximately every 4 weeks (Figure 3.7 B), and secondary and tertiary recipients were generated and analyzed over time ($3x10^6$ BM cells). Analysis of the peripheral blood contribution of CD45.2+ cells in long-term serial competitive chimeras did not exhibit loss of HSC engraftment upon deletion of the maternal *Meg3* allele (Figure 3.7 C). In addition, we also analyzed contribution in defined mature cell populations, including T cells, B cells and myeloid cells in serial transplantation experiments (Figure 3.7 D-F). Again, we did not observe any impairments in lineage output upon transplantation of *Meg3* KO cells. Collectively, our data shows that, even in highly competitive serial long-term transplantation experiments, *Meg3* KO HSCs do not exhibit decreased stem cell potential in comparison to control HSCs.

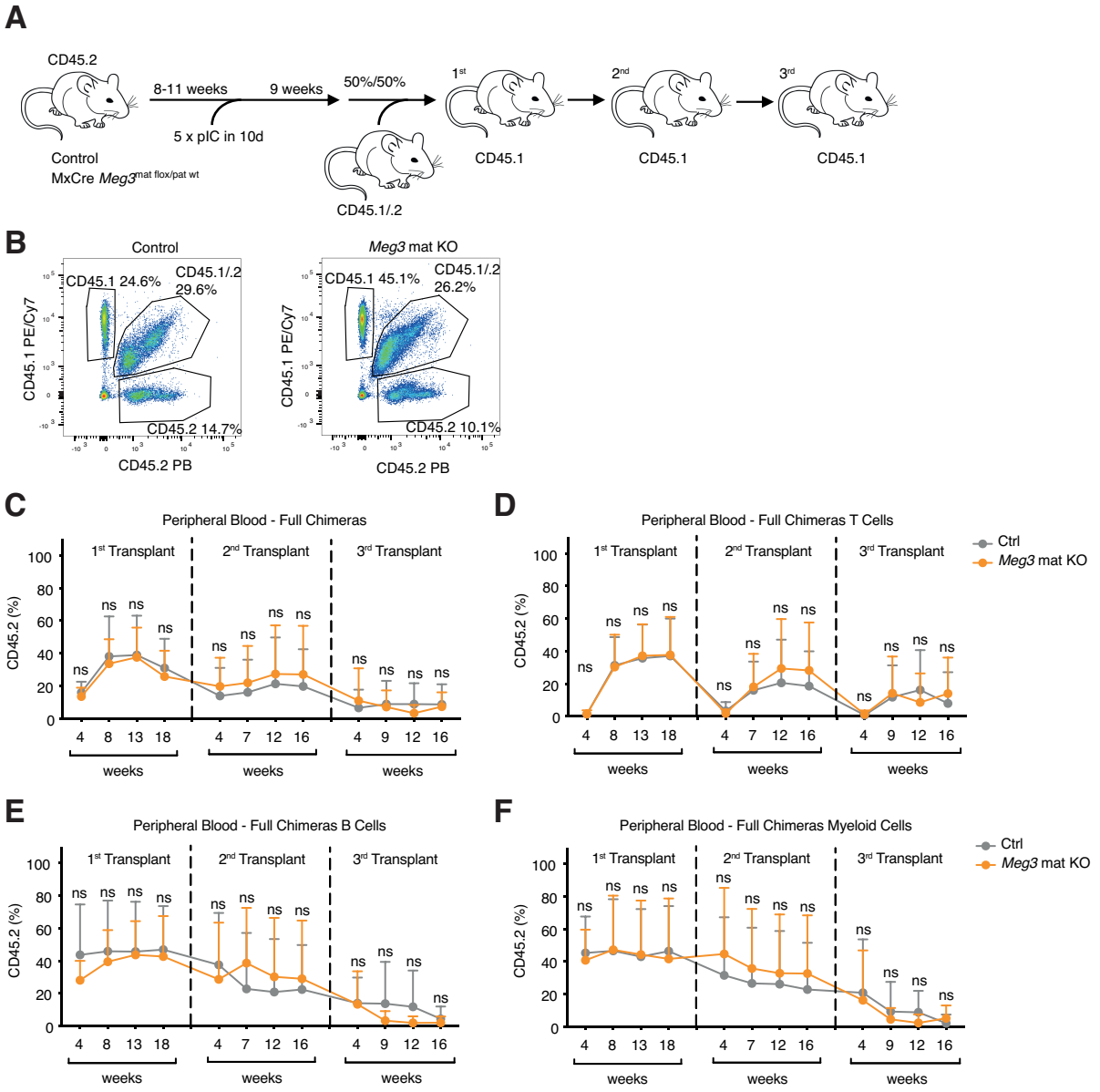


Figure 3.7: Loss of *Meg3* expression does not impair the competitive potential of HSCs.
(A) Workflow showing Cre induction and functional analysis by generation of competitive chimeras. (B) Representative flow cytometry dot plots showing PB analysis of competitive chimeras. (C) Flow cytometry-based analysis of serial competitive transplantation experiments. CD45.2% outcome is shown. N = 7-9, two-way ANOVA. (D-F) Flow cytometry-based analysis of serial competitive transplantation experiments. CD45.2% outcome is shown for T cells (D), B cells (E) and myeloid cells (F). N = 7-9, two-way ANOVA. For all panels: *p < 0.05, **p < 0.01, ***p < 0.001, ns: not significant; N indicates number of biological replicates; 1 independent experiment was performed; mean + SD is shown. Figure is adapted from (Sommerkamp et al., 2019).

3.1.3 *Meg3* is dispensable for embryonic hematopoietic cells expressing *Vav1*

A previous study focused on the role of the *Dlk1-Meg3* locus in fetal liver HSCs (Qian et al., 2016). In these studies, straight knockout models were used, leading to deletion of the locus in all embryonic cells. Thus, observed hematopoietic phenotypes might represent secondary effects generated by, for example, defective HSC niches in the developing embryos. As we did not observe any relevant biological effects in adult HSCs upon *Meg3* KO, we aimed to analyze the effects of hematopoietic specific KO of the locus in the embryo. Male VavCre driver mice were crossed to female *Meg3*$^{\text{mat flox/pat flox}}$ mice, generating VavCre *Meg3*$^{\text{mat flox/pat wt}}$ mice. Cre is expressed under the control of the *Vav1* promoter, leading to Cre expression in CD45+ AGM and CD45+ fetal liver cells starting at E11.5 (Chen et al., 2009). Thus, *Meg3* is deleted in all hematopoietic cells after the endothelial-cell-to-HSC transition, allowing embryonic deletion of *Meg3* in the hematopoietic lineage. Cre+ animals were born and appeared healthy. Analysis of hematopoiesis was performed 9 to 15 weeks after birth, Cre- littermates were used as controls (Figure 3.8 A). First, we confirmed *Meg3* KO by qPCR analysis of sorted control and *Meg3* KO HSCs (Figure 3.8 B). No expression of *Meg3* was detected in adult HSCs upon deletion of *Meg3* in the embryonic stage. Second, we analyzed peripheral blood, BM and spleen of control and KO animals. Analysis of peripheral blood did not exhibit any defects in the B cell, myeloid cell or T cell lineages (Figure 3.8 C). In line, we did not observe any changes in B cell maturation (pre/pro B cells, immature B cells, mature B cells) and the myeloid and T cell lineages in the adult BM upon embryonic loss of *Meg3* expression (Figure 3.8 D). Furthermore, no impairments in mature hematopoietic cell lineages in the spleen were detected (Figure 3.8 E). Conclusively, we performed analysis of the HSC/MPP compartment in VavCre *Meg3*$^{\text{mat flox/pat wt}}$ mice and did not observe any changes in cell frequencies compared to controls (Figure 3.8 F). Collectively, our data demonstrates that *Meg3* and the associated miRNA cluster are not only dispensable in the established adult hematopoietic hierarchy but also in the embryo after the endothelial-cell-to-HSC transition step.

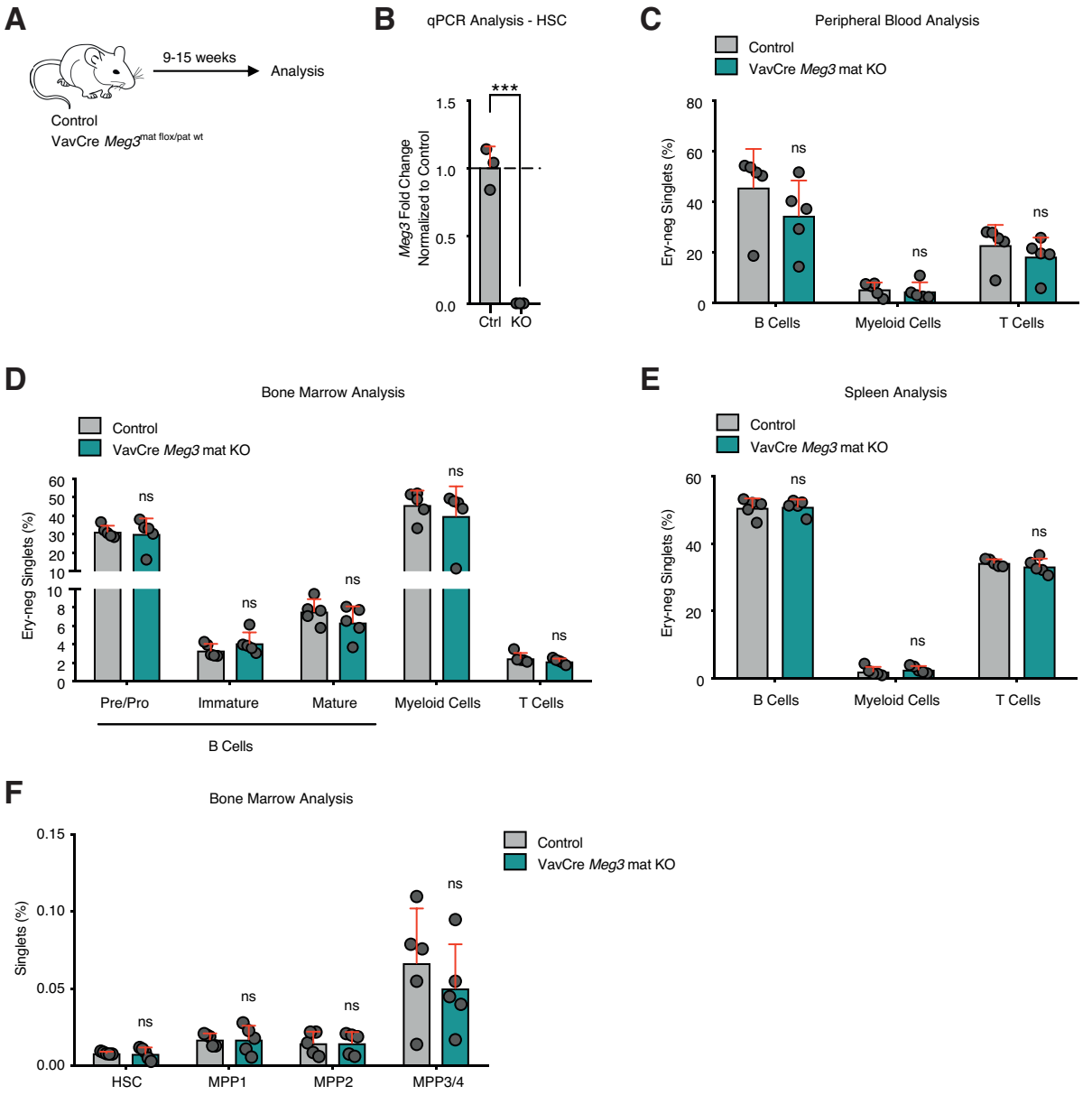


Figure 3.8: *Meg3* is dispensable for embryonic hematopoietic cells expressing *Vav1*.
(A) Workflow showing analysis of control and VavCre *Meg3*$^{\text{mat flox/pat wt}}$ mice. (B) qPCR-based analysis of *Meg3* expression in sorted HSCs. N = 3, unpaired student's t-test. (C) Flow cytometry-based analysis of mature cells in the peripheral blood compartment. N = 5, unpaired student's t-test. (D) Flow cytometry-based analysis of B cell maturation and mature cells in the BM compartment. N = 5, unpaired student's t-test. (E) Flow cytometry-based analysis of mature cells in the spleen. N = 5, unpaired student's t-test. (F) Flow cytometry-based analysis of HSCs/MPPs in the BM. Percentage of BM HSCs/MPPs is shown. N = 5, unpaired student's t-test. For all panels: *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$, ns: not significant; N indicates number of biological replicates; B: 1 independent experiment was performed; C-F: 2 independent experiments were performed; mean + SD is shown. Figure is adapted from (Sommerkamp et al., 2019).

3.1.4 Lack of *Meg3* expression marginally affects interferon-mediated HSC stress response

Meg3 has been reported to decrease cell proliferation (Braconi et al., 2011; Wang et al., 2012; Zhang et al., 2010a; Zhang et al., 2003; Zhou et al., 2007). HSCs are highly dormant, and inflammation-induced interferon signaling leads to activation of HSCs and exit from the G0 cell cycle state (Baldridge et al., 2010; Essers et al., 2009). This inflammatory response can be induced by pIC injections. We questioned if the return to quiescence in HSCs might be affected by loss of *Meg3*. We treated control and *Meg3* KO mice with pIC 6-12 weeks after deletion and analyzed the immediate cellular response in the HSC/MPP compartment using Ki-67/Hoechst 33348 staining 16 hours, 4 days and 8 days after proliferation induction (Figure 3.9 A). As expected, control HSCs and MPP cells became activated 16 hours after pIC injection and exited G0 (Figure 3.9 B). A gradual return to the intrinsic cell cycle status was observed after 4 and 8 days following cellular activation. *Meg3* KO HSCs and MPP cells did not exhibit changes in activation and return to quiescence. Only MPP1 KO cells showed slightly increased return to quiescence after 4 days and at the same time a slightly decreased return to quiescence after 8 days. In addition to cell cycle status, we analyzed BM frequencies (Figure 3.9 C). As expected, the proliferative response induced by pIC led to an increase of HSC and MPP frequencies 16 hours after injections. This increase was not affected by lack of *Meg3* expression. After 4 days, cell frequencies started to decrease again and reached homeostatic levels after 8 days. This response pattern was observed in control as well as *Meg3* KO animals. However, slightly increased HSC and MPP1 frequencies were observed in KO animals 8 days after injections.

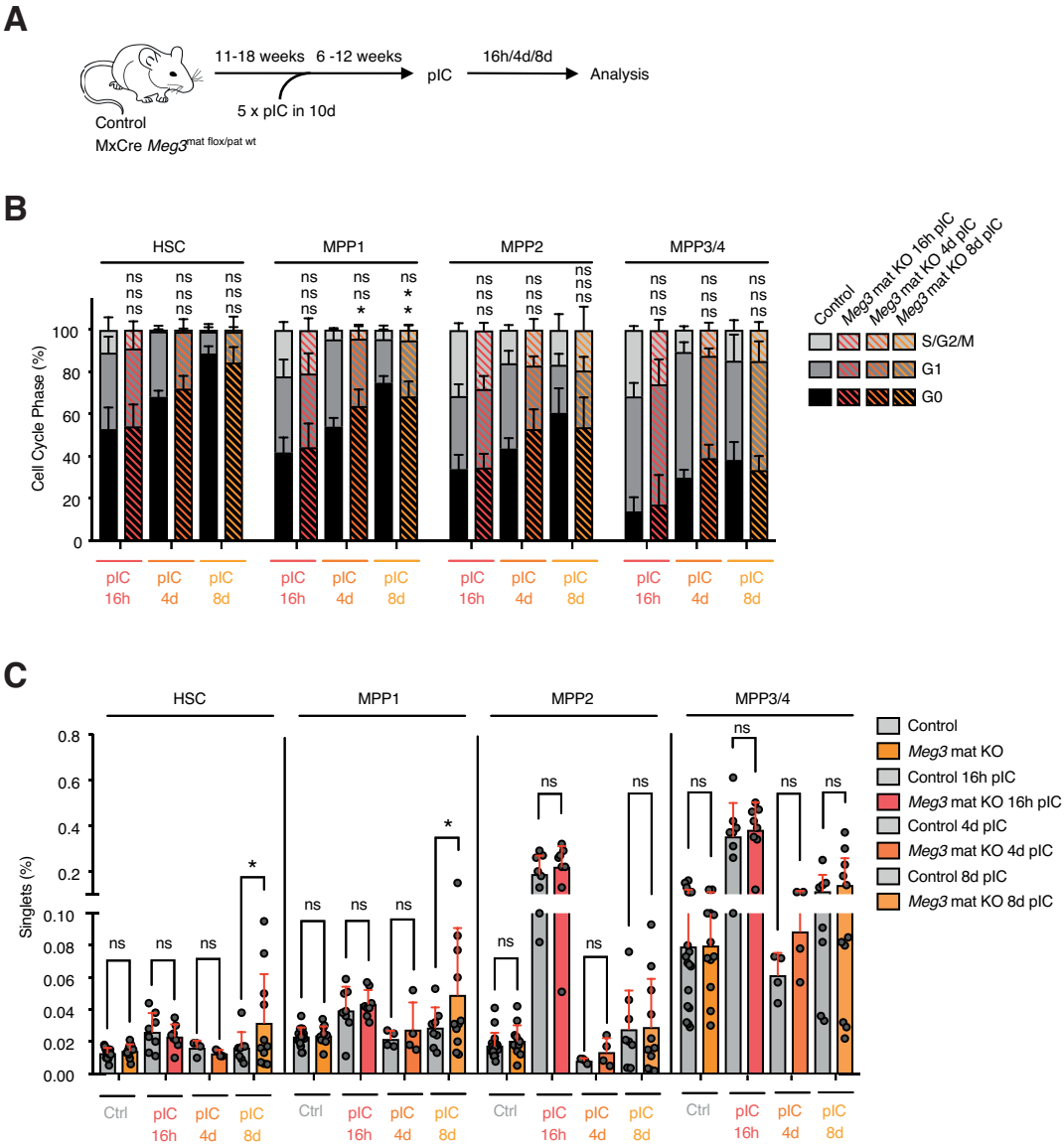


Figure 3.9: Lack of *Meg3* expression marginally affects interferon-mediated HSC stress response.
(A) Workflow showing pIC-mediated Cre induction and interferon-mediated stress induction. (B) Flow cytometry-based cell cycle analysis of HSCs/MPPs. Relative percentage within each cell cycle phase is shown. N = 4-11, two-way ANOVA. (C) HSC, MPP1, MPP2 and MPP3/4 frequencies as determined by flow cytometry analysis. Percentage of BM HSCs/MPPs is shown. N = 4-18, one-way ANOVA. For all panels: *p < 0.05, **p < 0.01, ***p < 0.001, ns: not significant; N indicates number of biological replicates; 1-4 independent experiments were performed; mean + SD is shown. Figure is adapted from (Sommerkamp et al., 2019).

To understand if this increase is of biological relevance, we decided to perform functional CFU assays using total BM cells derived from pIC-treated animals (Figure 3.10 A). We used cells derived from mice 16 hours and 8 days after pIC injection and performed serial plating experiments. CFU potential was not affected by *Meg3* KO. Even though we observed slightly increased HSC/MPP1 frequencies 8 days after pIC treatment, these changes did not lead to significantly altered colony formation capacity. This indicates that the changes observed are not of major biological relevance. To further strengthen this hypothesis, we performed long-term (LT) stress experiments. Ten weeks after *Meg3* deletion, mice were injected with pIC twice a week for a period of 4 weeks and analyzed after an 8-week recovery phase (Figure 3.10 B). As previously reported, LT stress leads to an overall decrease of HSC/MPP frequencies in control mice (Essers et al., 2009). In *Meg3* KO mice, this decrease in BM frequencies was comparable. Overall, *Meg3* does not seem to play an essential role in mediating HSC activation and return to quiescence following interferon-mediated inflammatory stress.

In summary, we performed extensive *in vitro* and *in vivo* characterization of *Meg3* KO mice harboring a specific deletion of *Meg3* in hematopoietic cells. Adult mice were analyzed, and *Meg3* is dispensable for functional hematopoiesis independently of the deletion time point (embryonic stage 11.5 or in the adult animal).

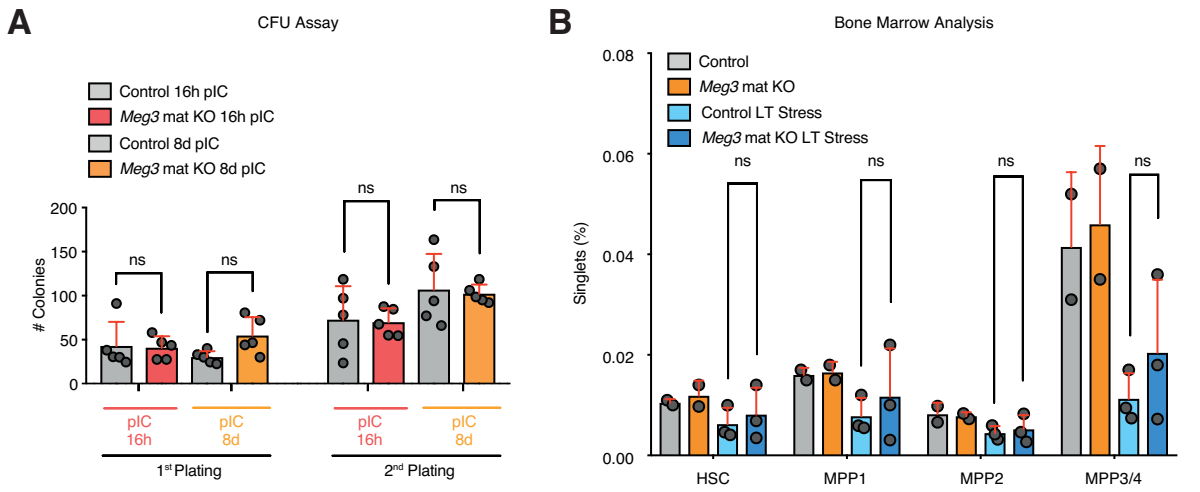


Figure 3.10: Lack of *Meg3* expression does not functionally affect interferon-mediated HSC stress response.
(A) Functional analysis of BM derived from pIC-treated mice by serial CFU assays. N = 5, one-way ANOVA. (B) HSC, MPP1, MPP2 and MPP3/4 frequencies as determined by flow cytometry analysis in LT stress experiments. Percentage of BM HSCs/MPPs is shown. N = 2-3, unpaired student's t-test. For all panels: *p < 0.05, **p < 0.01, ***p < 0.001, ns: not significant; N indicates number of biological replicates; 1 independent experiment was performed; mean + SD is shown. Figure is adapted from (Sommerkamp et al., 2019).

3.2 HSC function, differentiation and activation are regulated by APA

3.2.1 The APA regulator *Pabpn1* is an important regulator of HSC function

In addition to lncRNAs, our lab identified RNA regulatory mechanisms like alternative splicing to be differentially regulated between HSCs and MPP cells, potentially playing an important role in hematopoiesis (Cabezas-Wallscheid et al., 2014). Not only alternative splicing but also the RNA regulatory mechanism APA attracts increasing attention in different cellular entities and model systems (Di Giammartino et al., 2011; Tian and Manley, 2017). We aimed to identify if this process is functionally important for HSCs and if APA pattern changes occur upon HSC differentiation or activation. First, we investigated the functional relevance of APA in HSCs by targeting APA regulators.

3.2.1.1 Pabpn1 is differentially expressed in HSCs and MPP cells

We screened RNA-seq datasets (Cabezas-Wallscheid et al., 2014) to identify differentially expressed APA key regulators in HSCs and MPP cells. One of these key regulators was *Pabpn1*, which had previously been reported to regulate APA in cell lines and murine disease models (Jenal et al., 2012). In our RNA-seq dataset, we observed *Pabpn1* to be upregulated in HSCs compared to MPP cells (Figure 3.11 A). Expression of *Pabpn1* was especially low in MPP2 cells. This is interesting, as APA has been linked to changes in the cell cycle state and MPP2 cells are the most proliferative MPP subset (Tian and Manley, 2017; Wilson et al., 2008). We confirmed differential expression of *Pabpn1* by sorting HSC and MPP subsets and performing qPCR analysis (Figure 3.11 B).

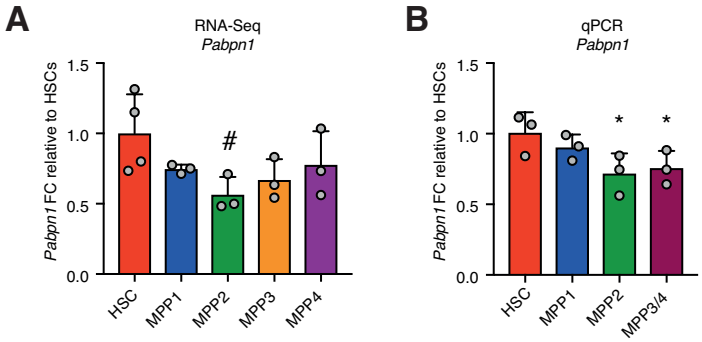


Figure 3.11: *Pabpn1* is differentially expressed in HSCs and MPP cells.
(A) RNA-seq-based expression analysis (Cabezas-Wallscheid et al., 2014) of *Pabpn1* in HSCs/MPPs (gated using LSK CD150+/- CD48+/- CD34+/- CD135+/-). Normalized relative expression values. N = 3-4, DeSeq analysis, #: significant compared to HSCs. (B) qPCR-based *Pabpn1* expression analysis in HSCs/MPPs (gated using LSK CD150+/- CD48+/-CD34+/-). Normalized relative expression values. N = 3, unpaired student's t-test between HSCs and MPP cells. For all panels: *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$, ns: not significant; N indicates number of biological replicates; 3 independent experiments were performed; mean + SD is shown.

3.2.1.2 Establishing shRNA-mediated Pabpn1 KD in LSK cells

To investigate the functional role of the APA mediator *Pabpn1* in hematopoiesis, we aimed to perform knockdown (KD) of *Pabpn1* in primary hematopoietic cells, namely LSK cells, and to perform subsequent *in vitro* and *in vivo* analyses (Figure 3.12 A). Plasmids for shRNA or Scr control RNA expression were purchased, and lentiviral particles were produced. After titration of viruses, primary LSK cells from C57Bl6/J mice were sorted by fluorescence-activated cell sorting (FACS) and transduced *in vitro* with an MOI of 35. Successfully transduced cells express the shRNA/Scr control RNA and a GFP reporter under the control of the murine EF1α promoter. 48 h and 72 h after transduction, cells were harvested and GFP+ cells were sorted by FACS (Figure 3.12 B). Up to 85% of cells were successfully transduced. Expression of the target gene *Pabpn1* was analyzed in GFP+ cells by qPCR analysis 48 h after transduction (Figure 3.12 C). Using two independent shRNAs, gene expression was reduced by approximately 50% compared to Scr control conditions (Figure 3.12 C). In addition to RNA expression, protein levels were assessed (Figure 3.12 D-E). GFP+ cells were sorted 48 h and 72 h after transduction and intracellular PABPN1 staining was performed to subsequently analyze expression levels by flow cytometry. A significant reduction in PABPN1 protein levels was observed. Reduction was higher 72 h after transduction, and protein levels were halved compared to Scr control conditions.

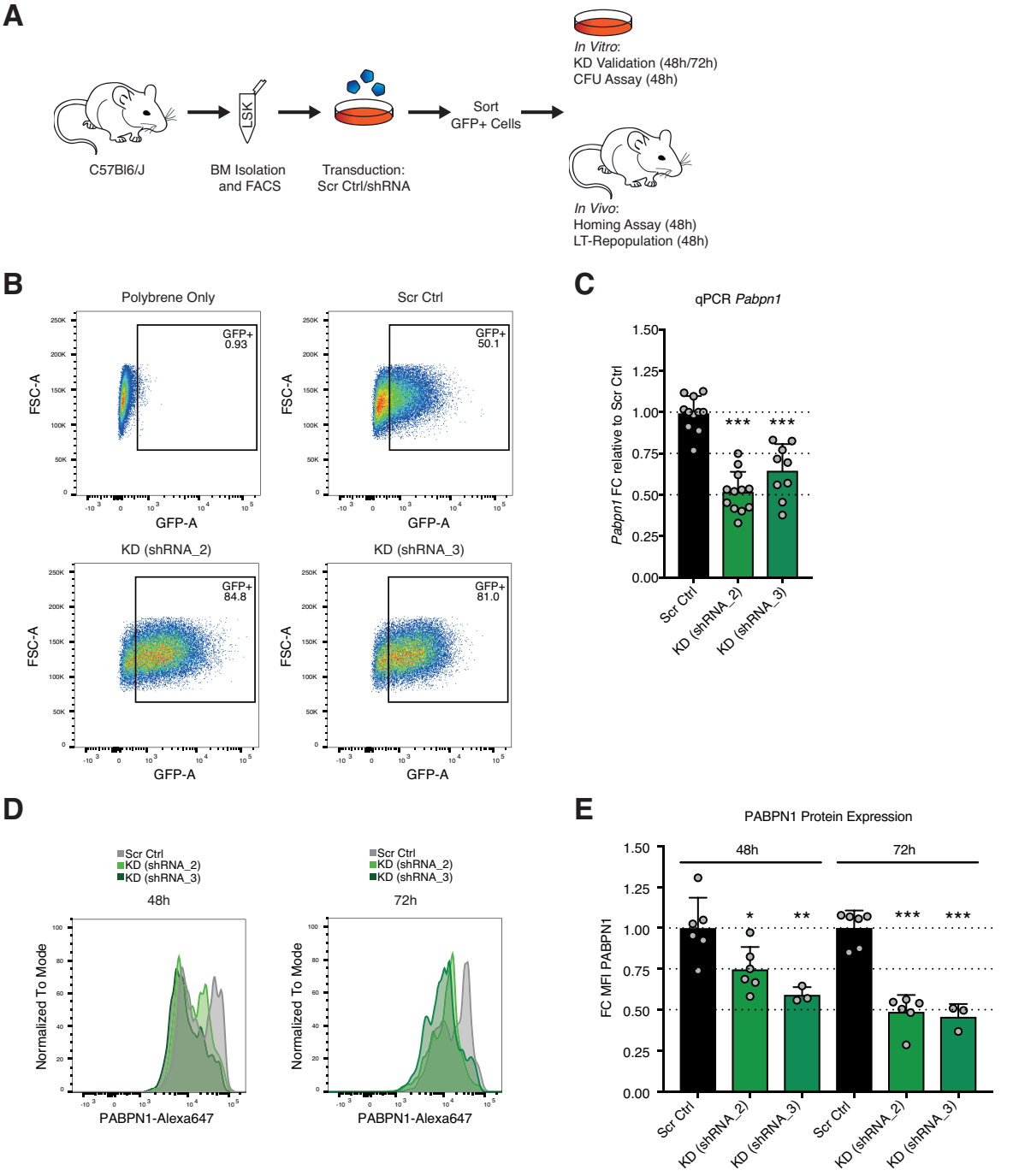


Figure 3.12: Establishing shRNA-mediated *Pabpn1* KD in LSK cells.
(A) Workflow of shRNA-mediated *Pabpn1* KD and subsequent *in vitro* and *in vivo* analyses. (B) Representative flow cytometry dot plots showing GFP expression of control and transduced LSK cells. (C) qPCR analysis of *Pabpn1* expression in Scr control and KD cells after 48 h. LSK cells were sorted and transduced, and GFP+ cells were resorted and analyzed. N = 9-12, unpaired student's t-test. (D) Representative flow cytometry histograms showing PABPN1 expression in resorted GFP+ control and KD cells after 48 h and 72 h. (E) Flow cytometry-based analysis of PABPN1 protein levels in Scr control and KD cells (48 h and 72 h). LSK cells were sorted and transduced, and GFP+ cells were resorted and analyzed. Fold change (FC) of the mean fluorescent intensity (MFI) is shown. Normalized to Scr control. N = 3-6, unpaired student's t-test, one independent experiment for KD shRNA_3. For all panels: Significance levels compared to Scr Ctrl, *p < 0.05, **p < 0.01, ***p < 0.001, ns: not significant; N indicates number of biological replicates; equal to or more than 3 independent experiments were performed unless otherwise indicated; mean + SD is shown.

3.2.1.3 shRNA-mediated Pabpn1 KD in LSK cells does not affect cellular viability or differentiation 48 h after lentiviral transduction

We aimed to perform functional assays using Scr control and *Pabpn1* KD cells 48 h after transduction. These assays are highly sensitive and if the KD approach would alter viability or cellular composition already 48 h after transduction, this might influence functional assays, as *Pabpn1* KD stem cells might already be affected and harbor cellular disadvantages. Thus, we performed viability analysis by 7AAD/Annexin staining (Figure 3.13 A). Annexin binds to phosphatidylserine, which is presented on cells undergoing apoptosis (Koopman et al., 1994). 7AAD/Annexin double negative cells are viable, 7AAD- Annexin+ cells undergo apoptosis, while 7AAD/Annexin double positive cells are no longer viable. *Pabpn1* KD did not affect cell viability 48 h after transduction. In addition, we performed flow cytometry analysis of GFP+ cells (Figure 3.13 B). 48 h after transduction, initially sorted LSK cells had partially differentiated into Lin+ Sca-1- c-Kit+ (LS-K) cells, consisting of different progenitor populations (CMP, GMP, MEP), in GFP+ Scr control samples. Analysis of *Pabpn1* KD cells did not exhibit any changes in these differentiation patterns. In summary, no changes in cell viability and differentiation were observed in *Pabpn1* KD cells 48 h after transduction, allowing unbiased functional analysis.

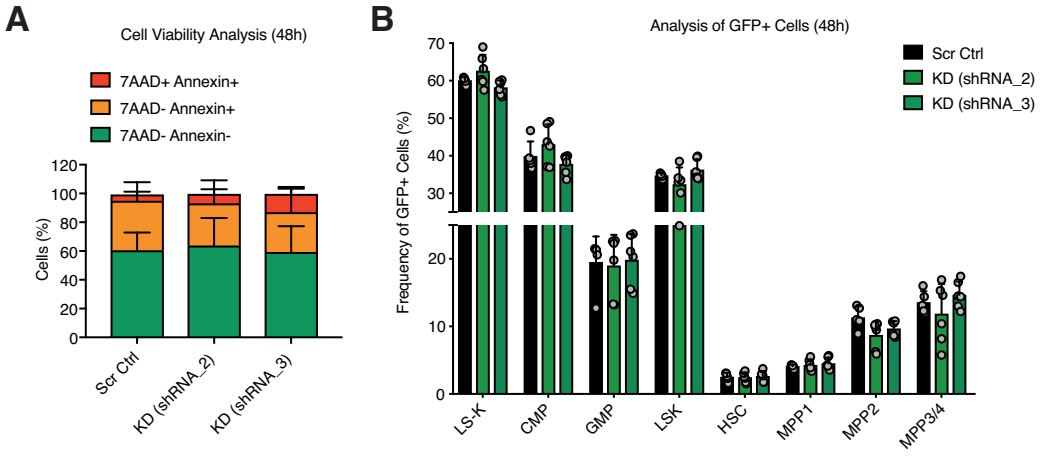


Figure 3.13: shRNA-mediated *Pabpn1* KD in LSK cells does not affect cellular viability or differentiation 48 h after lentiviral transduction.
(A) Flow cytometry-based viability analysis upon *Pabpn1* KD. N = 6. (B) Flow cytometry-based analysis of subpopulations in GFP+ cells. Percent of GFP+ cells is shown. N = 5-6. For all panels: N indicates number of biological replicates; equal to or more than 3 independent experiments were performed; mean + SD is shown.

3.2.1.4 Pabpn1 KD cells exhibit reduced CFU capacity in vitro but unchanged homing potential in vivo

To investigate the stem cell potential of *Pabpn1* KD cells, we first performed *in vitro* CFU assays (Figure 3.14 A). Interestingly, colony forming potential was reduced in primary and secondary platings using two independent shRNAs targeting *Pabpn1*. Thus, HSCs lacking the APA regulator *Pabpn1* exhibit decreased stem cell potential. In further experiments, we aimed to analyze stem cell capacity of KD cells *in vivo*. To this end, we performed *in vivo* transplantation experiments. A prerequisite to assess peripheral blood contribution is homing of the transplanted cells to the BM after i.v. injection. If the intrinsic homing capacity of the experimental cellular subject in question is already affected, this changes cellular performance in transplantation assays, leading to potential misinterpretation of results. To prevent this, we performed homing experiments using Scr control and KD cells (Figure 3.14 B). LSK cells were sorted and transduced, and 48 h later GFP+ cells were resorted. 30,000 GFP+ cells (CD45.2) were transplanted per lethally irradiated recipient mouse (CD45.1). Mice were sacrificed after 48 h, and homing was quantified by flow cytometry analysis of BM cells. CD45.2+ GFP+ cells were gated, and BM frequency of this cell population was determined (Figure 3.14 C-D). No difference in homing potential was observed in *Pabpn1* KD cells, enabling long-term transplantation experiments without an intrinsic bias in the experimental setup. Collectively, *Pabpn1* KD impairs HSC capacity *in vitro*, but does not affect *in vivo* homing capacity.

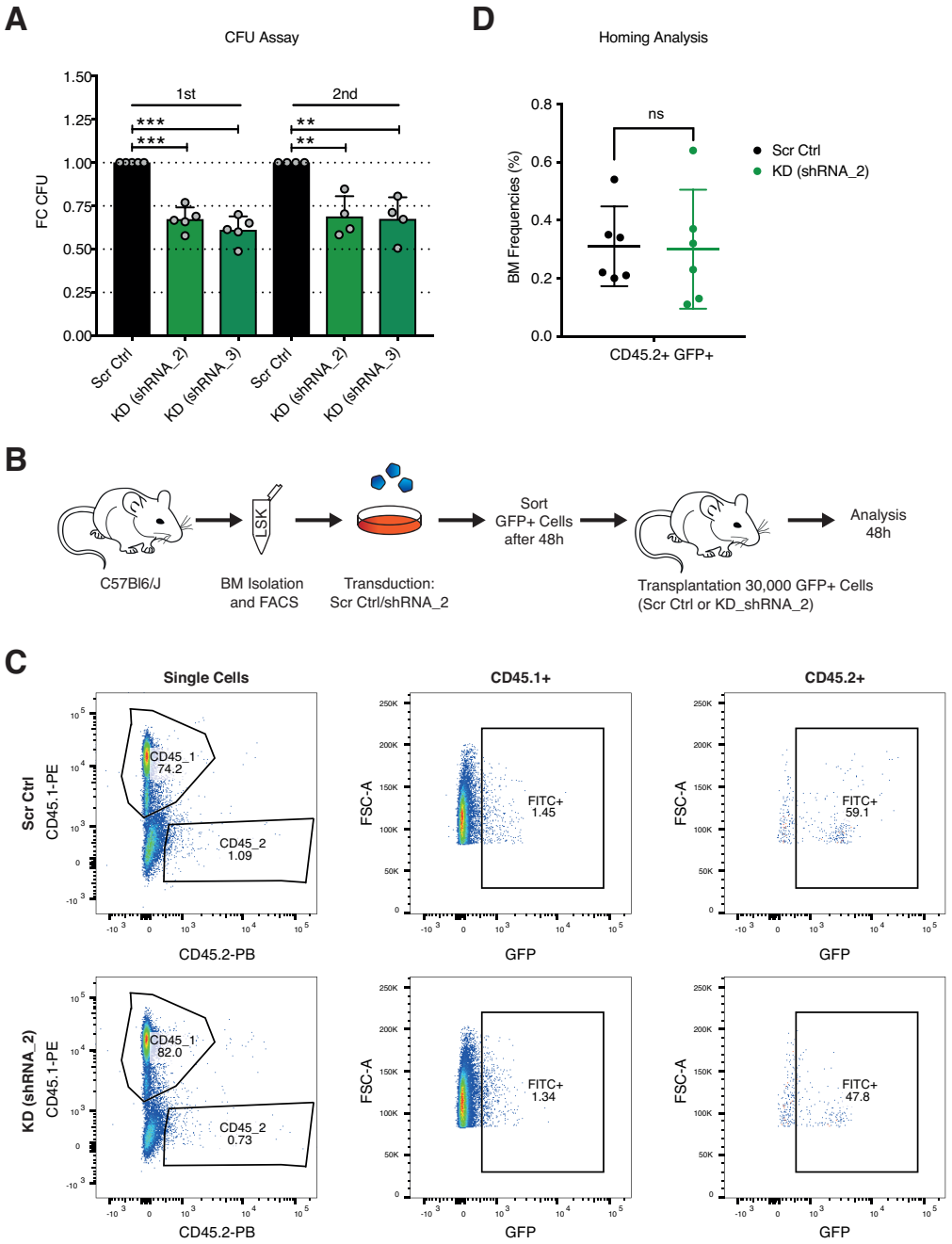
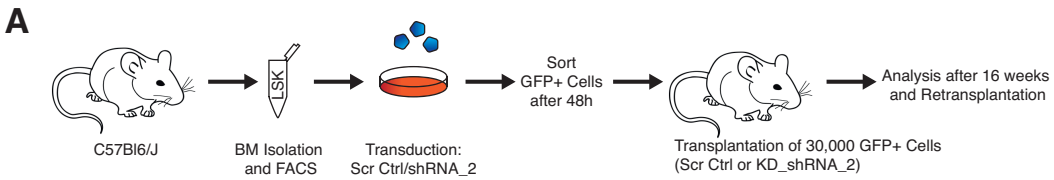
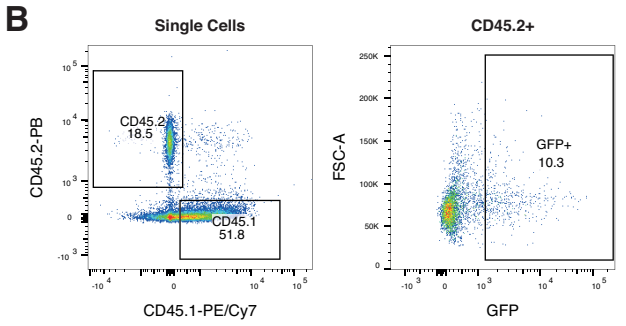


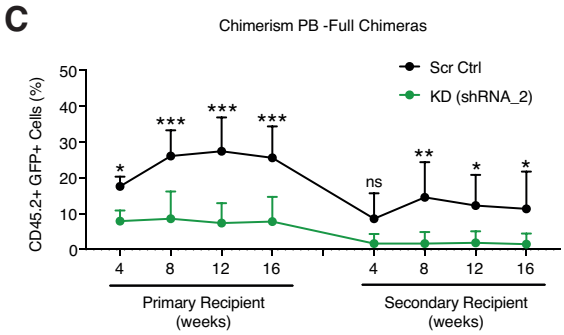
Figure 3.14: *Pabpn1* KD cells exhibit reduced colony forming unit capacity *in vitro* but unchanged homing potential *in vivo*.
(A) CFU assay of Scr control and *Pabpn1* KD cells showing fold change (FC) in colony numbers for first and second plating. N = 4-5, unpaired student's t-test. (B) Workflow showing homing assay analysis. (C) Representative flow cytometry dot plots of homing assay analysis. BM samples are shown. (D) Analysis of HSC homing. BM frequency of total CD45.2+ GFP+ cells is indicated. N = 6, unpaired student's t-test. For all panels: *p < 0.05, **p < 0.01, ***p < 0.001, ns: not significant; N indicates number of biological replicates; equal to or more than 2 independent experiments were performed; A: mean + SD is shown; D: mean ± 95% CI is shown.

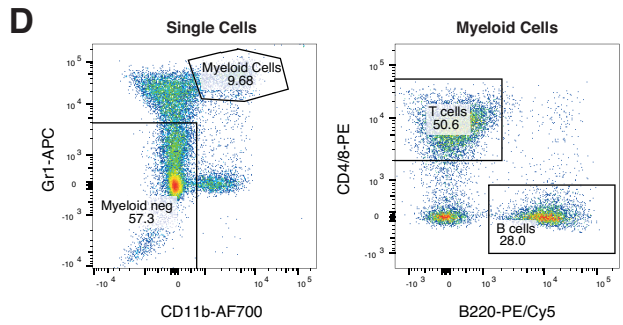
3.2.1.5 Engraftment and peripheral blood contribution of Pabpn1 KD cells in vivo is decreased

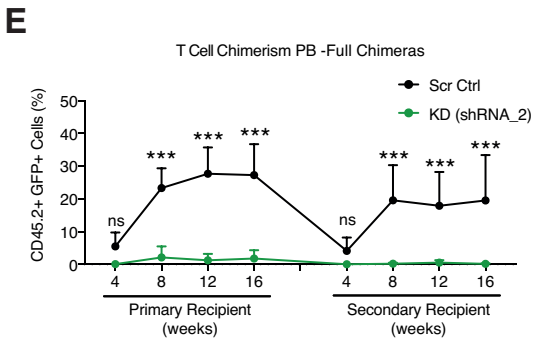
The gold standard to investigate HSC potential is transplantation, as only HSCs show serial multilineage engraftment. To asses HSC potential of *Pabpn1* KD cells *in vivo*, we generated full chimeras by transplanting either 30,000 GFP+ Scr control or 30,000 GFP+ *Pabpn1* KD (shRNA_2) cells (all CD45.2) into lethally irradiated recipient mice (CD45.1) (Figure 3.15 A). The proportion of CD45.2+ GFP+ cells in the peripheral blood was analyzed approximately every 4 weeks by flow cytometry analysis (Figure 3.15 B-C). *Pabpn1* KD cells exhibited significantly reduced contribution to the peripheral blood already in primary recipients at the first analysis time point (4 weeks). This decreased contribution was observed over time in primary as well as secondary recipients. In addition to overall peripheral blood contribution, we also analyzed contribution to the T cell, B cell and myeloid cell lineage by determining the proportion of CD45.2+ GFP+ cells in the respective population (Figure 3.15 D). Most strikingly, we did not observe CD45+ GFP+ T cells in the peripheral blood of *Pabpn1* KD recipient mice (Figure 3.15 E). We performed analysis of thymus cells to analyze T cell maturation using CD4, CD8, CD25 and CD44 as markers of different T cell development steps (Figure 3.15 F). We were not able to detect CD45.2+ GFP+ T cell populations in *Pabpn1* KD recipients. Hence, the APA regulator *Pabpn1* seems to be essential for T cell development, as successful T cell maturation is completely abrogated upon reduced *Pabpn1* expression. Contribution of *Pabpn1* KD cells to peripheral B cell levels was also significantly reduced, while myeloid cells were less affected (Figure 3.15 G-H). Further, we analyzed HSC/MPP engraftment in the BM (Figure 3.15 I). Engraftment of CD45.2+ GFP+ *Pabpn1* KD HSCs was reduced in primary and secondary recipients. Collectively, the APA regulator *Pabpn1* is essential for HSC function *in vivo* affecting all lineages, while the T cell lineage exhibits the most severe phenotype.

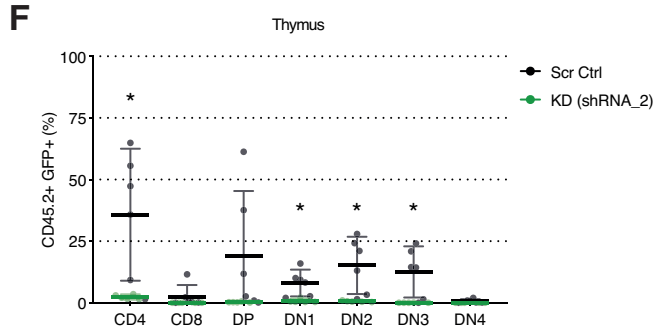

A
C57Bl6/J
BM Isolation and FACS
LSK
Transduction: Scr Ctrl/shRNA_2
Sort GFP+ Cells after 48h
Transplantation of 30,000 GFP+ Cells (Scr Ctrl or KD_shRNA_2)
Analysis after 16 weeks and Retransplantation

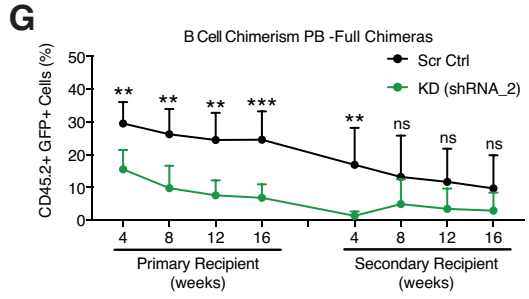

B
Single Cells
CD45.2+
CD45.2-PB
CD45.2 18.5
CD45.1 51.8
CD45.1-PE/Cy7
FSC-A
GFP+ 10.3
GFP

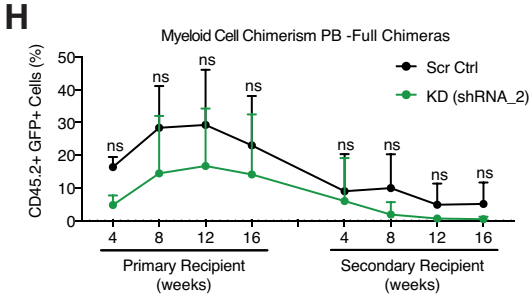

C
Chimerism PB -Full Chimeras
Scr Ctrl
KD (shRNA_2)
CD45.2+ GFP+ Cells (%)
Primary Recipient (weeks)
Secondary Recipient (weeks)


D
Single Cells
Myeloid Cells
Myeloid Cells 9.68
Myeloid neg 57.3
Gr1-APC
CD11b-AF700
Myeloid Cells
T cells 50.6
B cells 28.0
CD4/8-PE
B220-PE/Cy5


E
T Cell Chimerism PB -Full Chimeras
Scr Ctrl
KD (shRNA_2)
CD45.2+ GFP+ Cells (%)
Primary Recipient (weeks)
Secondary Recipient (weeks)


F
Thymus
Scr Ctrl
KD (shRNA_2)
CD45.2+ GFP+ (%)
CD4
CD8
DP
DN1
DN2
DN3
DN4


G
B Cell Chimerism PB -Full Chimeras
Scr Ctrl
KD (shRNA_2)
CD45.2+ GFP+ Cells (%)
Primary Recipient (weeks)
Secondary Recipient (weeks)


H
Myeloid Cell Chimerism PB -Full Chimeras
Scr Ctrl
KD (shRNA_2)
CD45.2+ GFP+ Cells (%)
Primary Recipient (weeks)
Secondary Recipient (weeks)

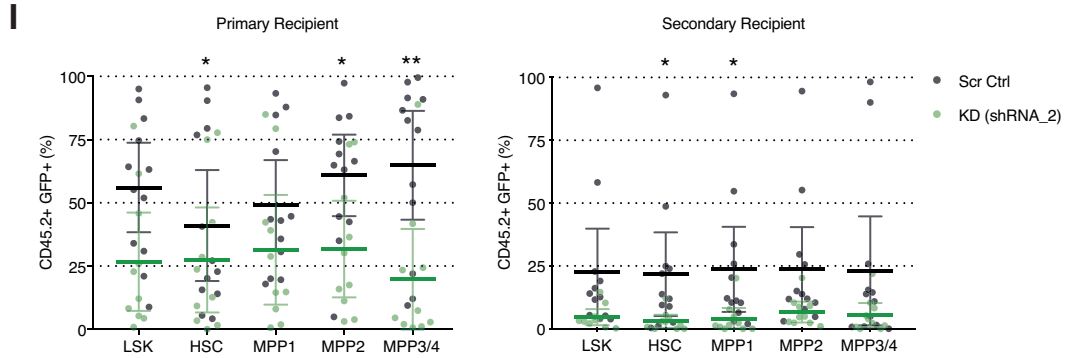
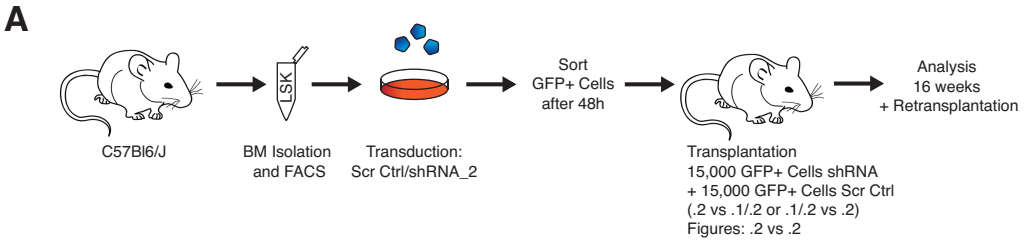
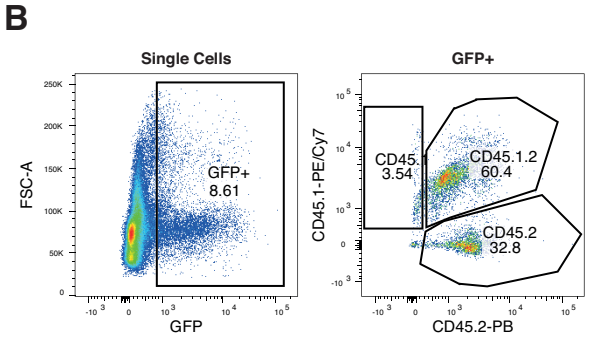

I
Primary Recipient
Secondary Recipient
Scr Ctrl
KD (shRNA_2)
CD45.2+ GFP+ (%)
LSK
HSC
MPP1
MPP2
MPP3/4

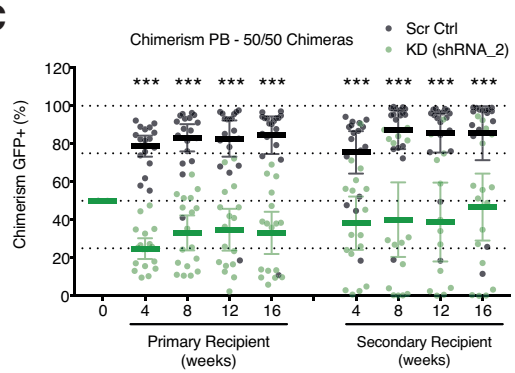
Figure 3.15: Decreased engraftment and peripheral blood contribution of *Pabpn1* KD cells *in vivo*.
(A) Workflow showing generation of full chimeras. (B) Representative flow cytometry dot plots of PB analysis are shown. (C) CD45.2+ GFP+ outcome in PB analysis of full chimeras is shown. N = 10-12, two-way ANOVA. (D) Representative flow cytometry dot plots of gating of mature cells in PB analysis are shown. (E) CD45.2+ GFP+ outcome in PB analysis for T cells of full chimeras is shown. N = 10-12, two-way ANOVA. (F) CD45.2+ GFP+ outcome in thymus of full chimeras is shown. N = 10-12, unpaired student's t-test. DP: double positive (CD4+ CD8+), DN: double negative (CD4- CD8-) (G) CD45.2+ GFP+ outcome in PB analysis for B cells of full chimeras is shown. N = 10-12, two-way ANOVA. (H) CD45.2+ GFP+ outcome in PB analysis for myeloid cells of full chimeras is shown. N = 10-12, two-way ANOVA. (I) First and second endpoint analysis of BM of full chimeras. Gated on LSK CD150+/- CD48+/- CD34+/-. CD45.2+ GFP+ outcome is shown. N=10-12, unpaired student's t-test. For all panels: Significance levels compared to Scr Ctrl, *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$, ns: not significant; N indicates number of biological replicates; 2 independent experiments were performed; mean ± 95% CI is shown.

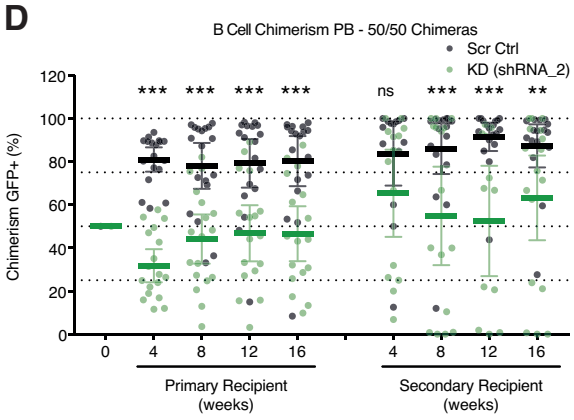
3.2.1.6 Pabpn1 KD cells exhibit competitive disadvantage in transplantation experiments in vivo

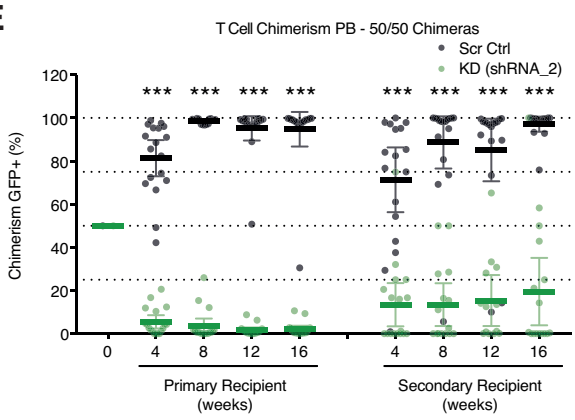
We generated 50/50 chimeras to analyze the competitive potential of *Pabpn1* KD hematopoietic cells. Therefore, we mixed 15,000 Scr control (CD45.2 or CD45.1/.2) with 15,000 *Pabpn1* KD (CD45.1/.2 or CD45.2) cells and transplanted the cells into lethally irradiated recipients (CD45.1) (Figure 3.16 A). Thus, mice were either transplanted with a mixture of Scr control CD45.2/*Pabpn1* KD CD45.1/.2 cells or Scr control CD45.1/.2/*Pabpn1* KD CD45.2 cells. This setup is necessary, as CD45.2 cells exhibit an intrinsic advantage in engraftment (Jafri et al., 2017). For analysis of peripheral blood contribution, frequencies of CD45.2+ GFP+ *Pabpn1* KD cells of one setup were compared to frequencies of CD45.2+ GFP+ Scr control cells of the other setup (Figure 3.16 B). In competitive settings, Scr control cells outperformed *Pabpn1* KD cells in serial transplantation assays (Figure 3.16 C). Analysis of mature lineages in the peripheral blood confirmed our *in vivo* results generated by full chimera analysis. Again, we observed decreased production of B cells, T cells and myeloid cells, with T cells being most affected (Figure 3.16 D-F). Results were confirmed by setting up additional competitive transplantation experiments using the independent shRNA_3 to generate *Pabpn1* KD cells (Figure 3.16 G). Competitive engraftment potential of HSCs and MPP cells in the BM was analyzed by flow cytometry (Figure 3.16 H). Again, we observed reduced engraftment of HSCs and MPP cells in primary recipients. Analysis of secondary recipients was challenging, as only few GFP+ HSCs and MPP cells could be detected 16 weeks after retransplantation. In general, we observed the tendency that *Pabpn1* KD HSC/MPP cells showed reduced engraftment in secondary recipients in competitive settings. In summary, *Pabpn1* KD cells do not only exhibit reduced engraftment potential in full chimeras but are also outperformed in competitive transplantation settings.

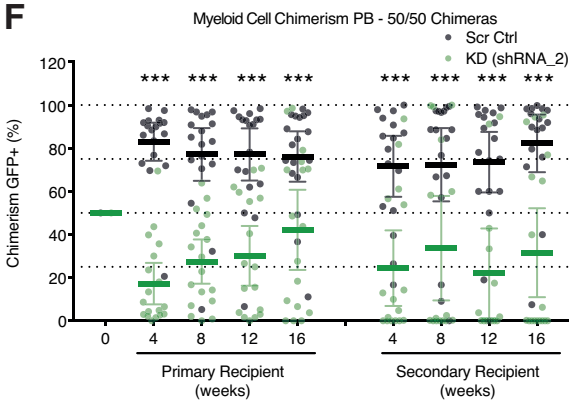

A
C57Bl6/J
BM Isolation
and FACS
LSK
Transduction:
Scr Ctrl/shRNA_2
Sort
GFP+ Cells
after 48h
Transplantation
15,000 GFP+ Cells shRNA
+ 15,000 GFP+ Cells Scr Ctrl
(.2 vs .1/.2 or .1/.2 vs .2)
Figures: .2 vs .2
Analysis
16 weeks
+ Retransplantation

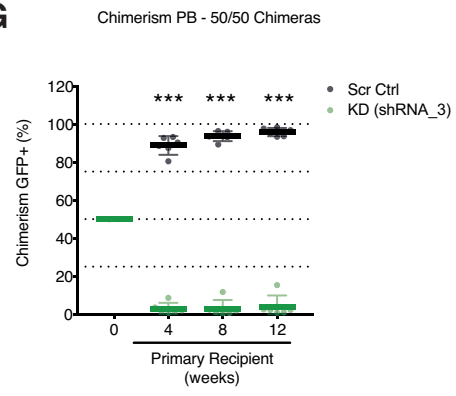

B
Single Cells
GFP+
FSC-A
GFP
GFP+
8.61
CD45.1-PE/Cy7
CD45.2-PB
CD45.1
3.54
CD45.1.2
60.4
CD45.2
32.8


C
Chimerism PB - 50/50 Chimeras
Scr Ctrl
KD (shRNA_2)
Chimerism GFP+ (%)
Primary Recipient
(weeks)
Secondary Recipient
(weeks)


D
B Cell Chimerism PB - 50/50 Chimeras
Scr Ctrl
KD (shRNA_2)
Chimerism GFP+ (%)
Primary Recipient
(weeks)
Secondary Recipient
(weeks)


E
T Cell Chimerism PB - 50/50 Chimeras
Scr Ctrl
KD (shRNA_2)
Chimerism GFP+ (%)
Primary Recipient
(weeks)
Secondary Recipient
(weeks)


F
Myeloid Cell Chimerism PB - 50/50 Chimeras
Scr Ctrl
KD (shRNA_2)
Chimerism GFP+ (%)
Primary Recipient
(weeks)
Secondary Recipient
(weeks)


G
Chimerism PB - 50/50 Chimeras
Scr Ctrl
KD (shRNA_3)
Chimerism GFP+ (%)
Primary Recipient
(weeks)

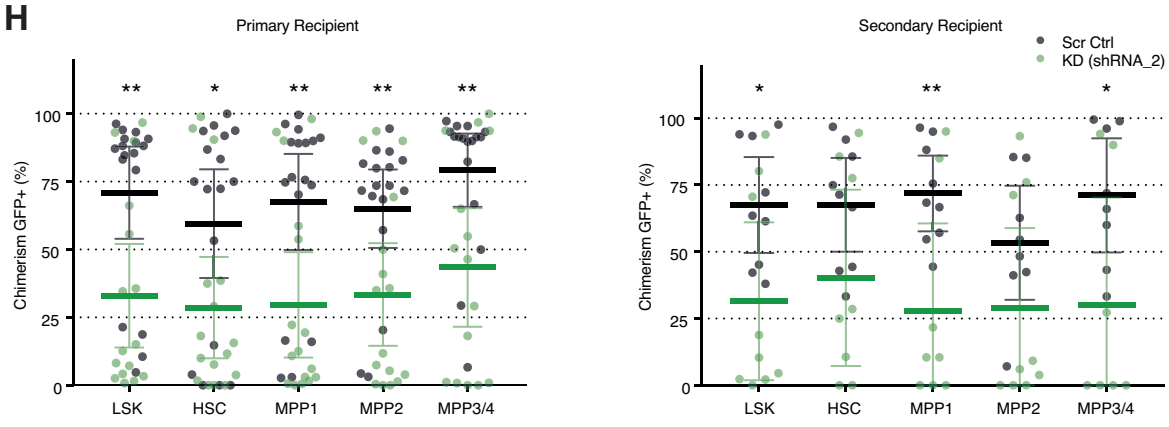

H
Primary Recipient
Secondary Recipient
Scr Ctrl
KD (shRNA_2)
Chimerism GFP+ (%)
LSK
HSC
MPP1
MPP2
MPP3/4

Figure 3.16 *Pabpn1* KD cells exhibit competitive disadvantage in transplantation experiments *in vivo*.
(A) Workflow showing generation of competitive chimeras. (B) Representative flow cytometry dot plots of PB analysis are shown. (C) PB analysis of competitive chimeras. Frequency of CD45.2+ cells in the GFP+ compartment is represented. N = 13-18, two-way ANOVA. (D-F) PB analysis showing frequency of CD45.2+ cells in the GFP+ compartment of B cells (D), T cells (E) and myeloid cells (F). N = 13-18, two-way ANOVA. (G) PB analysis of competitive chimeras using shRNA_3. Frequency of CD45.2+ cells in the GFP+ compartment is represented. N = 6, two-way ANOVA, one independent experiment. (H) First and second endpoint analysis of BM of competitive chimeras. Gated on LSK CD150+/- CD48+/- CD34+/-. Frequency of CD45.2+ cells in the GFP+ compartment is represented. N=10-12, unpaired student's t-test. For all panels: Significance levels compared to Scr Ctrl, *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$, ns: not significant; N indicates number of biological replicates; 3 independent experiments were performed unless otherwise indicated; mean ± 95% CI is shown.

3.2.2 3'-Seq analysis reveals distinct APA patterns in hematopoietic stem and progenitor populations

3.2.2.1 Establishing ultra-low input 3'-Seq

After showing that the APA regulator *Pabpn1* is essential for hematopoiesis, we aimed to analyze the prevalence of APA in HSCs and MPP cells. Standard RNA-seq data can hardly be used to analyze APA patterns, as complex bioinformatic reconstruction of cleavage site usage has to be performed (Chen et al., 2017; Wilkening et al., 2013). In addition, only few reads are localized in the 3'-UTR region to perform bioinformatic analysis. Thus, to enable in-depth analysis of APA in the hematopoietic hierarchy, we aimed to perform 3'-Seq analysis. To this end, we sorted HSC and MPP1-4 cells under homeostatic conditions and isolated total RNA (Figure 3.17 A-B). Lin, cKit, Sca-1, CD135, CD150, CD48 and CD34 were used as markers. Around 10,000 cells were sorted by FACS in quadruplicates in independent experiments.

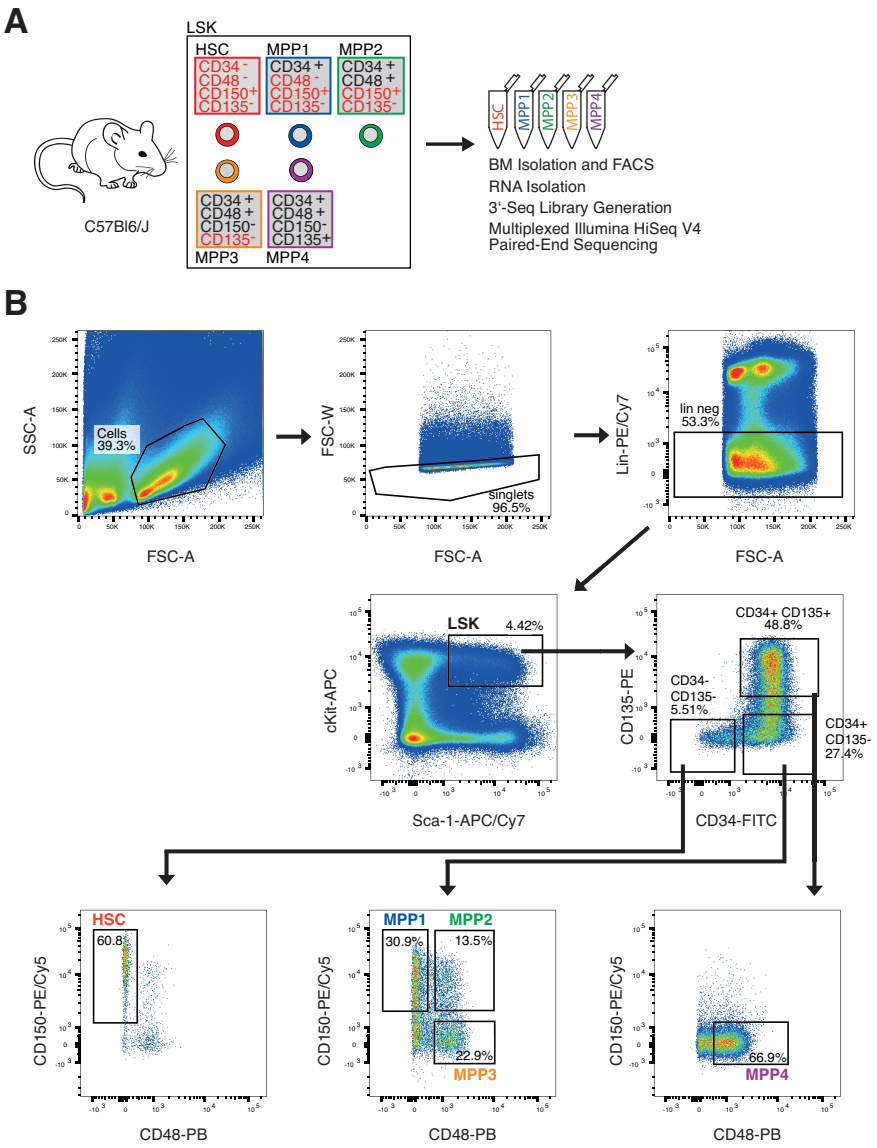


Figure 3.17: Establishing ultra-low input 3'-Seq – gating strategy.
(A) Workflow and study design of 3'-Seq analysis of homeostatic HSCs/MPPs. (B) Representative flow cytometry dot plots of HSC/MPP sorting approach.

Isolated RNA was used for library generation using the QuantSeq 3'-mRNA Library Prep Kit REV (Lexogen) (Figure 3.18 A). First, RNA quality was confirmed using an Agilent Bioanalyzer. The RNA integrity number (RIN) resembles RNA quality and integrity (Schroeder et al., 2006). Only high-quality samples with a RIN value higher than 9 were processed. During establishment of the protocol for ultra-low inputs, we started to include an RNA denaturation step to enable library preparation (see methods section). 3 ng of total RNA were reverse transcribed using oligo(dT)-priming. Residual RNA was digested, and random priming was used to generate double-stranded cDNA. PCR amplification was performed and during this step, adapter sequences for Illumina sequencing and barcodes for sample identification were introduced. The cycle number for amplification was determined to be 19 cycles with the help of PCR Add-on Kit for Illumina (Lexogen). Library quality was checked using the Agilent Bioanalyzer, and only samples with an average size below 450 bp and a molarity of at least 2

nM were used. Successfully generated libraries were pooled, and multiplexed paired-end sequencing was performed on a HiSeq 2000 device. In this sequencing approach, read1 was generated using a custom sequencing primer provided with the QuantSeq 3'-mRNA Library Prep Kit REV (Lexogen, 5' CCCTACACGACGCTCTTCCGATCTTTTTTTTTTTTTTTTTT 3'). Read1 is derived from the terminal end of the utilized 3'-UTR, providing the advantage that no bioinformatic reconstruction of PAS selection is needed. The start of the read corresponds to the utilized cleavage site. In general, two different datasets were generated: paired-end read data, which can be used for analysis of differentially expressed genes, and single-end read data, which can be used for analysis of APA.

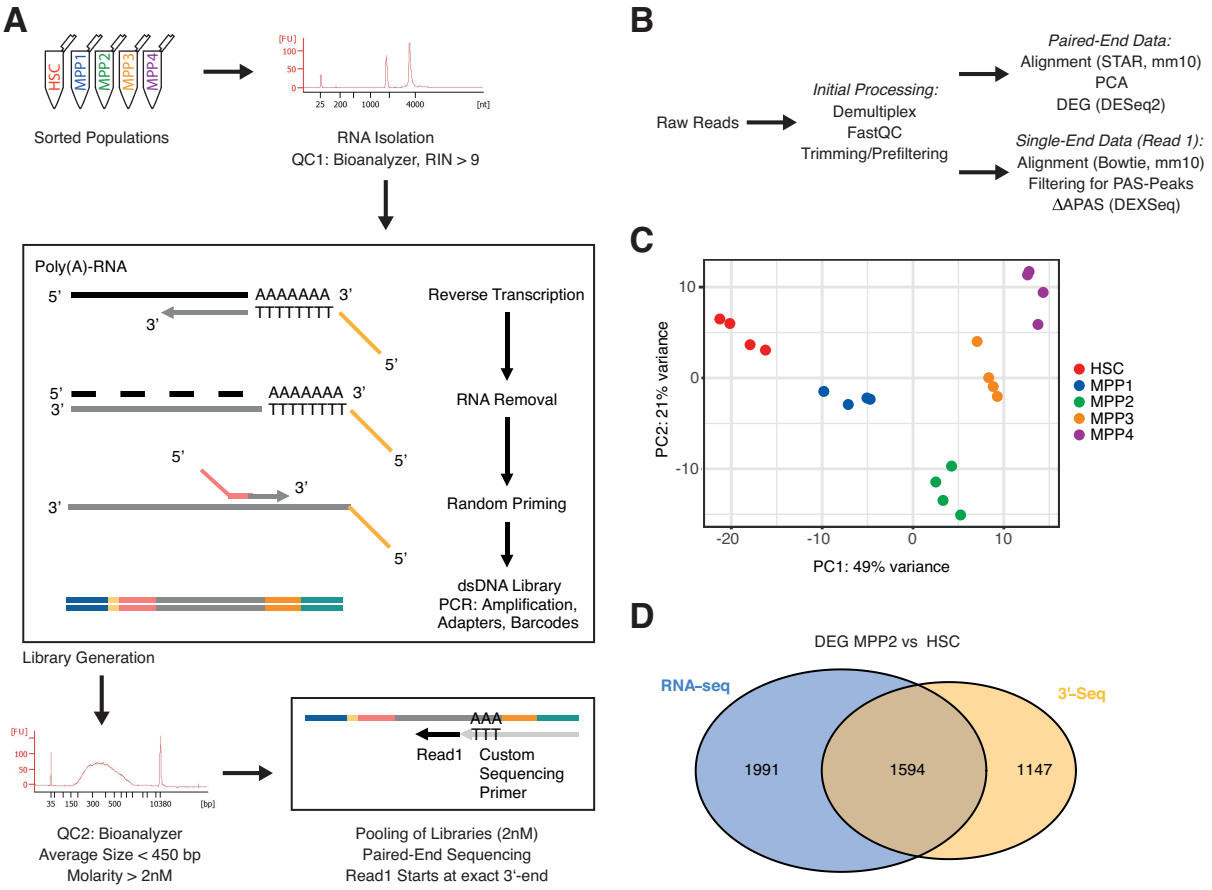


Figure 3.18: Establishing ultra-low input 3'-Seq – library generation and QC.
(A) 3'-Seq workflow of library generation and sequencing. (B) 3'-Seq bioinformatic analysis workflow. (C) PCA of the HSC/MPP 3'-Seq dataset using paired-end data. (D) Venn diagram of DEGs identified in MPP2 cells vs HSCs using standard RNA-seq (Cabezas-Wallscheid et al., 2014) or 3'-Seq analysis. 1594 shared DEGs representing 58% of DEGs detected by 3'-Seq were identified. For all panels: 3'-Seq data is derived from libraries generated in 4 independent experiments, representing 4 biological replicates per cell population, respectively.

For bioinformatic analysis (Figure 3.18 B), we collaborated with Sandro Altamura (Heidelberg University Medical Center). Initially, raw reads were demultiplexed, FastQC was performed and adapter sequences were trimmed. In addition, a prefiltering step was included. Paired-end read data was used to asses quality of our newly generated 3'-Seq dataset. As we established this sequencing approach for low inputs, we aimed to validate the reliability and robustness of

our data. Paired-end data was aligned using STAR, and principal component analysis (PCA) was performed (Figure 3.18 C). We observed robust clustering of independently generated biological HSC and MPP1-4 replicates, indicating high reliability of our 3'-Seq dataset. In addition, we performed analysis of differentially expressed genes (DEGs) using DESeq2 (Love et al., 2014). We identified 2,741 genes that were differentially expressed in HSCs and MPP2 cells. Comparison of our dataset with an RNA-seq dataset previously published by our lab revealed a strong overlap of DEGs, further supporting the quality of our generated dataset (Figure 3.18 D). Overall, we established an ultra-low RNA input 3'-Seq approach and confirmed the high quality and robustness of our dataset.

3.2.2.2 Identification of high confidence PASs using single-end read data from 3'-Seq analysis

In a next step, we aimed to use the newly generated 3'-Seq resource of primary HSCs and MPP cells to analyze APA patterns in the distinct subpopulations. To this end, we made use of the single-end read data generated using the custom sequencing primer. This allowed us to address: I) PAS usage, II) 3'-UTR length and III) associated exons for each detected peak. Reads were aligned using Bowtie, and polyadenylation peaks were identified and filtered for the presence of conserved PAS and absence of downstream internal poly-A stretches to only include peaks originating from pre-mRNA cleavage enabling high-fidelity APA analysis (Figure 3.19 A-B). As expected, analysis of peak annotation revealed that the majority of peaks is annotated in the 3'-UTR region in contrast to standard RNA-seq datasets in which almost 90% of reads are localized in coding sequence exons (Figure 3.19 C). Interestingly and in line with previous reports (Elkon et al., 2013), we also identified APA in introns, intergenic and downstream regions. Analysis of overall representation of the different conserved PASs did not reveal any differences between HSCs and MPP cells (Figure 3.19 D). The most common canonical PAS AAUAAA was used in over 50% of polyadenylation events. This observation further substantiates the quality and reliability of our APA analysis, as this PAS usage pattern was previously described (Tian and Graber, 2012). In line, we observed a distance of ~15-30 bp between the read start site, which corresponds to the cleavage site, and the conserved PAS (Figure 3.19 E) (Tian and Graber, 2012). Collectively, we were able to identify high confidence polyadenylation sites using our generated 3'-Seq analysis. Quantification of PAS usage in our dataset can subsequently be used to identify and analyze APA.

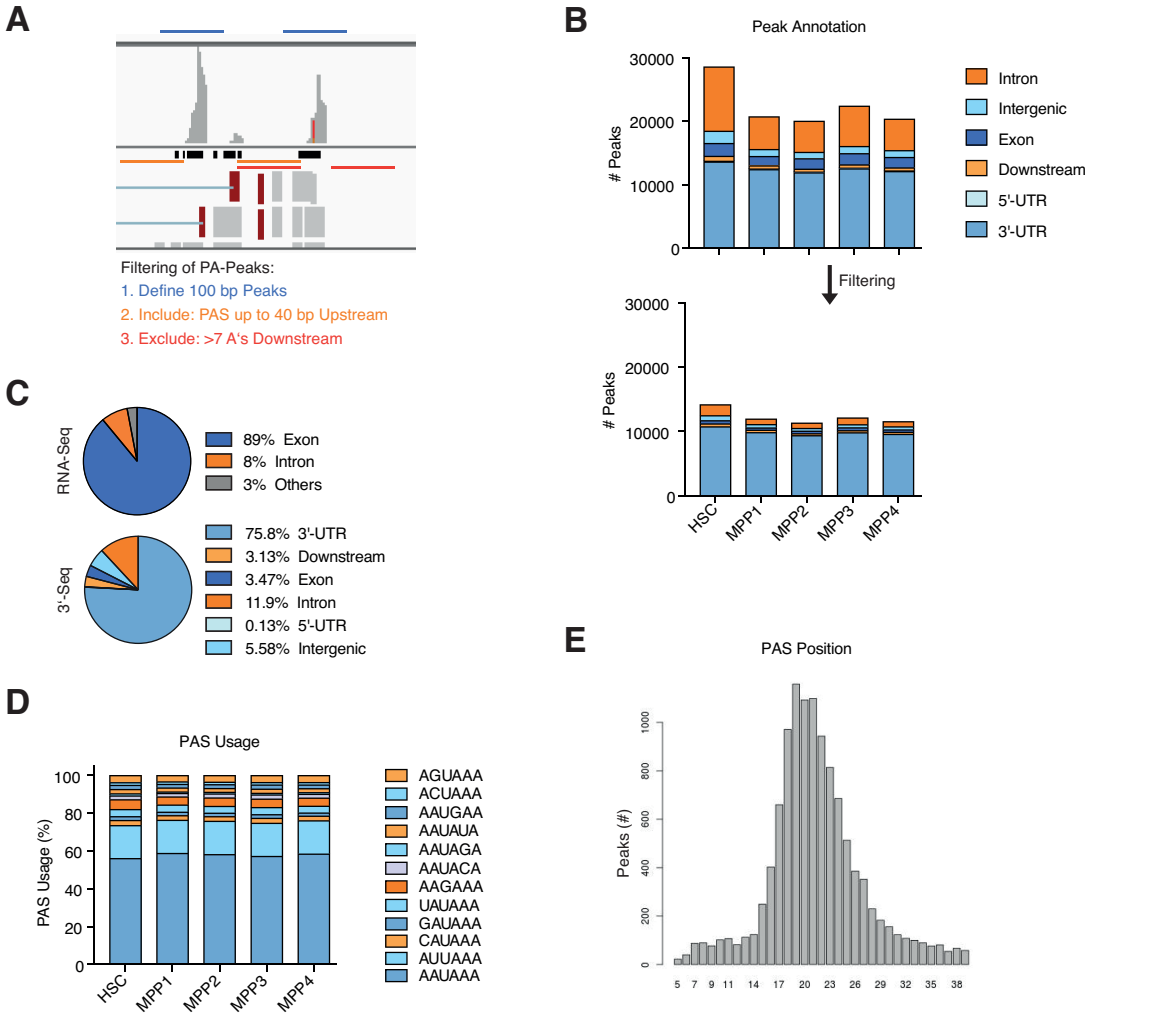


Figure 3.19: Identification of high-confidence PASs using single-end read data from 3'-Seq analysis.
(A) Workflow of peak identification in 3'-Seq data. (B) Peak annotation before and after filtering of the 3'-Seq dataset. Number of peaks per population is shown. (C) Read classification after mapping and filtering. Exemplary RNA-seq and 3'-Seq datasets are shown. (D) Usage of different conserved PASs in HSCs and MPP cells. Percentage of PAS usage is shown. (E) Representative distribution of distance from the conserved PAS to the read start site. Number of peaks is shown. For all panels: 3'-Seq data is derived from libraries generated in 4 independent experiments, representing 4 biological replicates per cell population, respectively.

3.2.2.3 Analysis of APA patterns in HSCs and MPP1-4 cells

After confirming the reliability of our PAS identification strategy, we continued with the analysis of APA patterns in primary HSC/MPP cells. First, we compared peak annotations between the different populations (Figure 3.20 A). As expected, the majority of polyadenylation events was identified in annotated 3'-UTRs. Interestingly, however, HSCs exhibited a significant, ~2-fold, increase in the intronic polyadenylation frequency compared to MPP1-4 cells. In addition, HSCs showed a significant increase in the total number of genes using more or equal to 3 polyadenylation sites compared to MPP cells (Figure 3.20 B). Polyadenylation patterns seem to be more variable in the stem cell compartment compared to progenitors with regard to localization of PASs and PAS usage per gene.

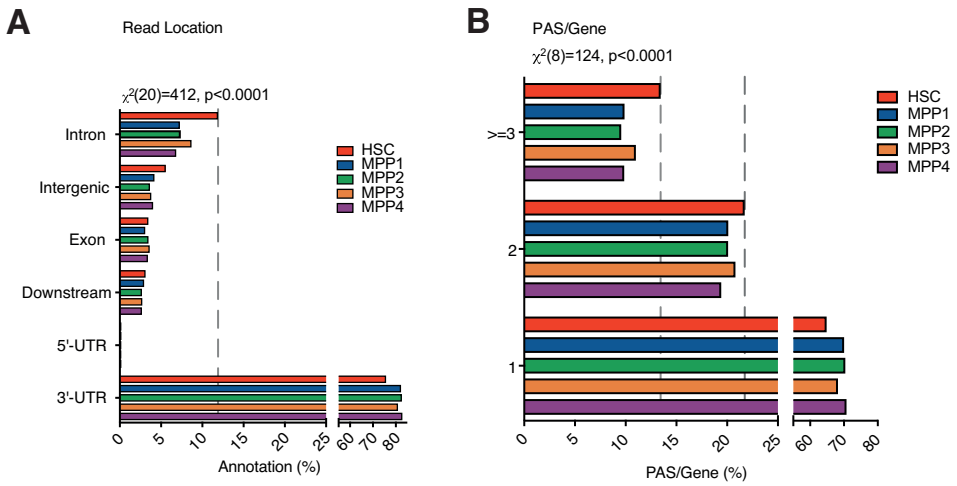


Figure 3.20: Analysis of APA patterns in HSCs and MPP1-4 cells – read localization and PAS number per gene.
(A) Read localization in percent of HSCs and MPP cells in the analyzed 3'-Seq dataset. Chi-square test. (B) PAS number per gene in percent for HSCs and MPP cells. Chi-square test. For all panels: 3'-Seq data is derived from libraries generated in 4 independent experiments, representing 4 biological replicates per cell population, respectively.

Next, we aimed to identify statistically significant differentially used APA sites (ΔAPAS) between populations. A stringent bioinformatic pipeline was established on the basis of DEXSeq (Anders et al., 2012). As ΔAPAS are defined as differentially used PASs of a given gene, analysis can only be performed for genes expressed in both cell populations. Hits can include CR-APA, affecting protein isoform expression, UTR-APA, affecting only 3'-UTR length, or a combination of both. In a first step, we called DEGs and ΔAPAS between HSCs and MPP1-4 cells (Figure 3.21 A). In total, we identified between 657 and 2,872 genes to be differentially expressed between HSCs and the different MPP populations. As expected and previously described, most DEGs were identified between the highly proliferative MPP2 subset and deeply dormant HSCs. ΔAPAS analysis revealed a similar pattern with 113 ΔAPAS hits identified in HSCs and MPP2 cells. HSCs and MPP1 cells are molecularly highly similar cell subsets and in line we identified only 11 ΔAPAS in these populations. Usage of the different PASs was evaluated by calculating the ratio of the usage of a specific PAS against the sum of the usage of all the PASs for a specific transcript, defining the percentage of polyadenylation usage (PPAU). As HSCs and MPP2 cells showed the highest difference in APA, we decided to focus on these two populations. Analysis of the top 10 ΔAPAS genes exhibited a complex mixture of CR-APA as well as UTR-APA (Figure 3.21 B-C). For example, the genes *Chd2*, *Ctsc* and *Ccnd3* displayed CR-/UTR-APA, while *Hnrnpa3*, *Minos1* and *Treml2* displayed UTR-APA only. For all of these 3'-UTRs, we observed a switch from distal to more proximal PASs in MPP2 cells. To validate our ΔAPAS identification approach, we designed 3'-UTR- and isoform-specific primers and performed analysis of polyadenylation patterns in HSCs and MPP2 cells by qPCR analysis (Figure 3.21 D). RNA expression was normalized to *Oaz1* (ΔCt),

and relative expression levels were determined as 2^-ΔCt. The proximal 2^-ΔCt value was divided by the distal 2^-ΔCt value, and the results were normalized to HSCs. Thus, values above 1 indicate a relative increase in proximal APA site usage, and values below 1 indicate a relative increase in distal APA site usage. For comparison to 3'-Seq data, proximal and distal raw read counts were added up (p+d), and the (p+d)/d ratio was calculated and normalized. This correction of proximal reads is necessary for comparison with qPCR data, as the proximal primer pair always amplifies the shortened as well as the lengthened transcript (p+d). For all ΔAPAS hits analyzed, we could confirm the APA patterns in HSCs and MPP2 cells by qPCR, independently validating our sequencing results and bioinformatic analysis pipeline. In summary, we generated a resource dataset of HSCs and MPP cells that enables analysis of ΔAPAS in primary hematopoietic cells. In addition, we showed for the first time that distinct APA patterns exist between HSCs and MPP cells.

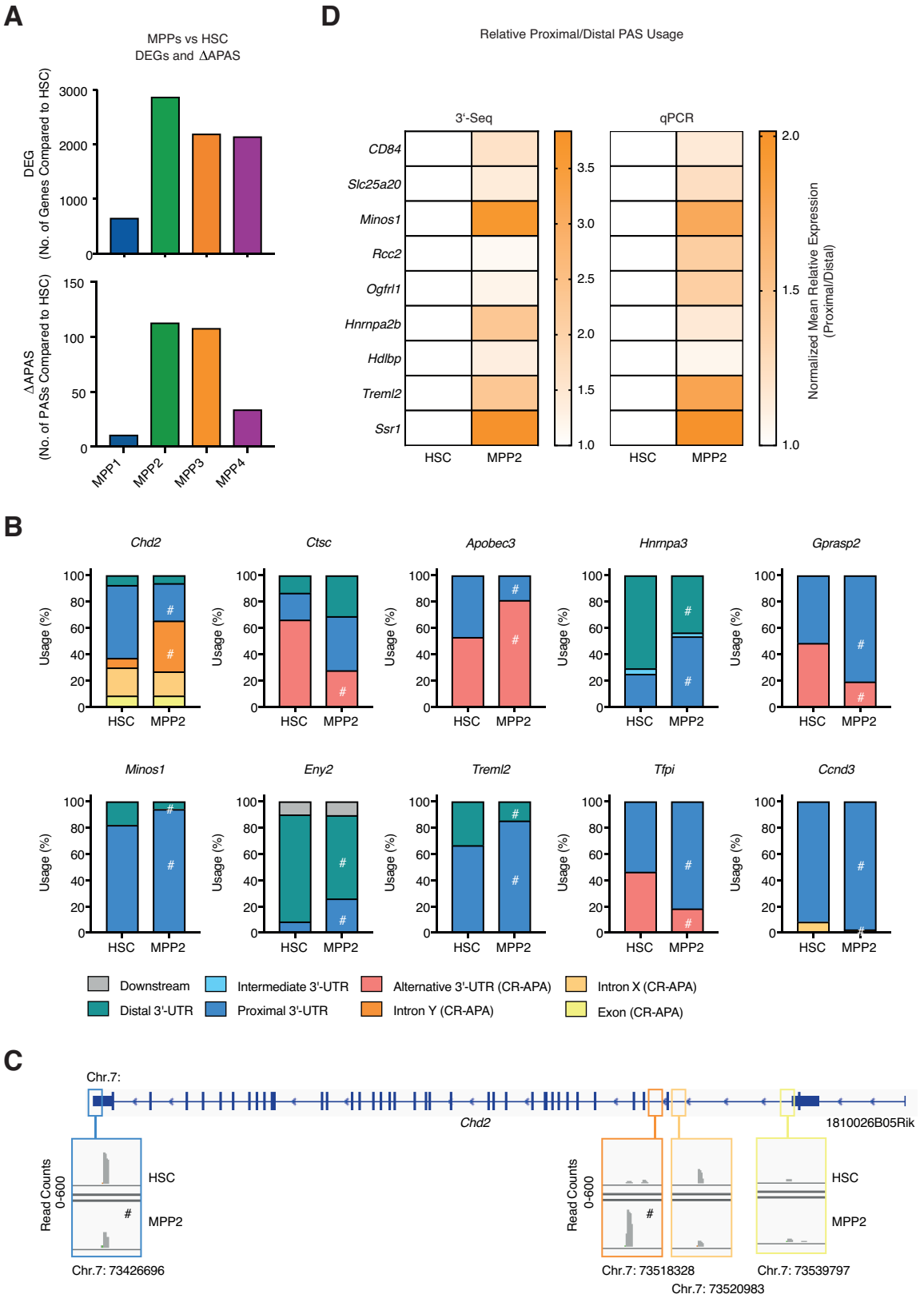


Figure 3.21: Analysis of APA patterns in HSCs and MPP1-4 cells – ΔAPAS analysis.
(A) Identification of DEGs (FDR < 0.1; -0.5 > Log2FC > 0.5) and ΔAPASs (FDR < 0.2) in MPPs compared to HSCs. (B) APA patterns of the top 10 ΔAPAS hits identified in MPP2 cells and HSCs. Usage of respective site in percent is shown (percentage of polyadenylation usage (PPAU), see methods section for details). ΔAPAS analysis: # significant. (C) Representative example of APA in MPP2 cells and HSCs. Exemplary IGV data of the gene *Chd2* is shown. (D) ΔAPAS validation by qPCR analysis of selected hits identified in MPP2 cells vs HSCs. Heatmap represents normalized mean relative expression of the proximal to the distal mRNA isoform based on 3'-Seq (left panel) and qPCR (right panel) analysis. For all panels: 3'-Seq data is derived from libraries generated in 4 independent experiments, representing 4 biological replicates per cell population, respectively.

3.2.2.4 3'-UTR shortening occurs upon HSC commitment and differentiation

Further, we aimed to analyze global 3'-UTR length in HSC/MPP subsets. To this end, we analyzed 3'-UTR hits only. For 3'-UTRs containing more than one PAS, the percentage of usage of the most distal polyadenylation site (PDUI) was considered, and we calculated its differential usage between the conditions (dPDUI). Reduced usage indicates 3'-UTR shortening, while increased usage indicates 3'-UTR lengthening. We compared MPP1-4 cells to HSCs and observed significant progressive 3'-UTR shortening in MPP1-4 cells (Figure 3.22 A-B). 3'-UTR shortening occurred upon HSC differentiation and commitment, which is tightly linked to increased cellular activity. Previously, shortening of 3'-UTRs has mainly been associated with increased protein output (Fu et al., 2018; Ji et al., 2011; Jia et al., 2017; Pai et al., 2016; Sandberg et al., 2008; Zheng et al., 2018). Therefore, we were interested in gene functions that were significantly enriched in MPPs compared to HSCs. Thus, we performed gene ontology (GO) term and REVIGO analysis on the 3'-UTR shortened gene list for MPP2 cells compared to HSCs (Figure 3.22 C). We considered only genes with a relative shortening of more than 10% in the distal PAS. Shortened genes were mainly associated with processes like *cell division*, *cell-cell signaling*, *stem cell division* and *glutamine metabolism*. Interestingly, some of these processes (e.g. *cell division*) have previously been shown to be characteristic for MPP2 cells compared to HSCs (Cabezas-Wallscheid et al., 2014). Considering that 3'-UTR shortening is linked to an increase in protein expression, it seems feasible that these MPP-related processes are enriched in the shortened gene sets. Overall, HSC commitment and differentiation are associated with global 3'-UTR shortening *in vivo*. In addition, shortened gene lists are enriched for MPP-specific processes.

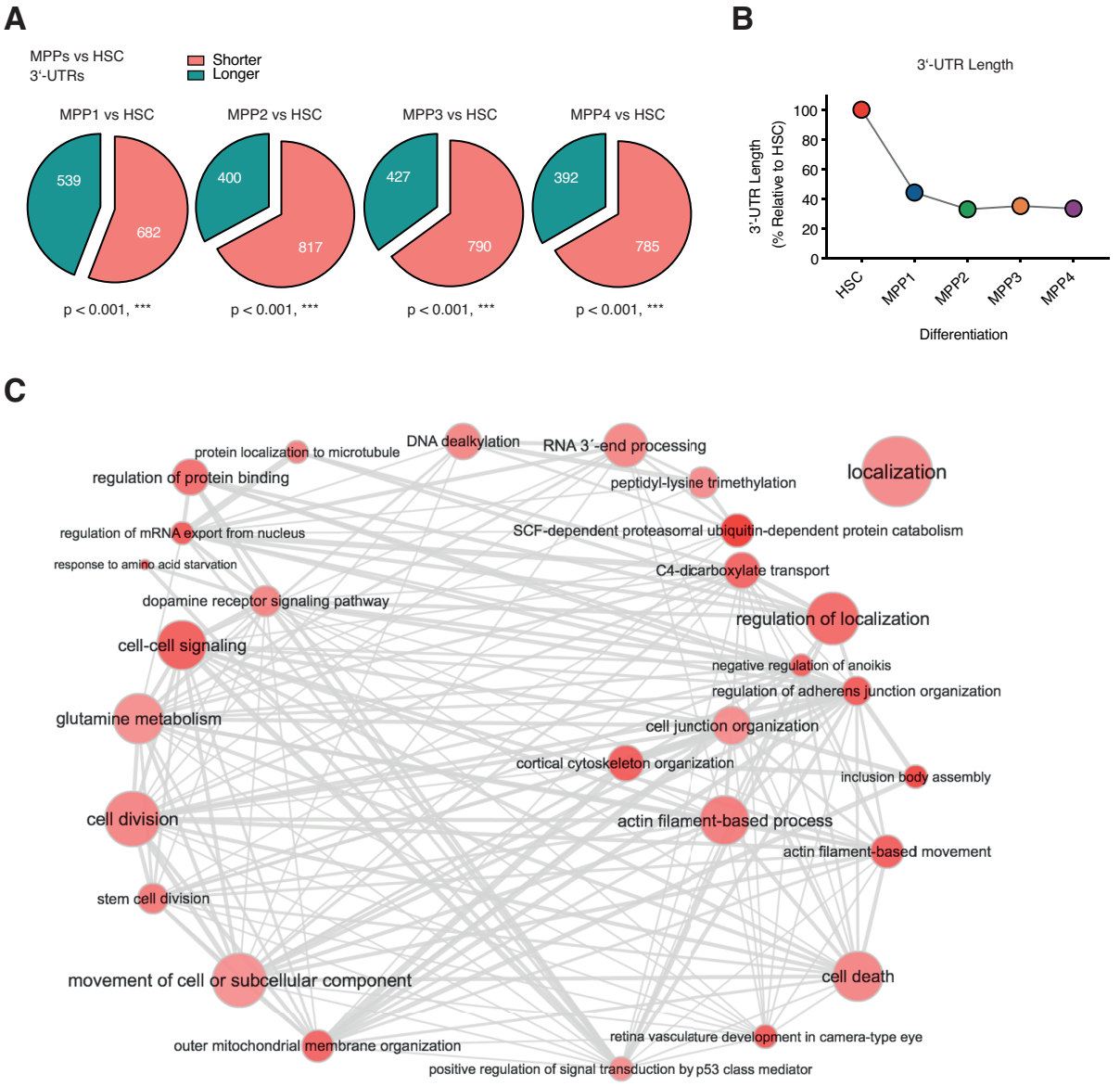


Figure 3.22: 3'-UTR shortening occurs upon HSC commitment and differentiation.
(A) Analysis of 3'-UTR shortening/lengthening. Only changes in 3'-UTRs belonging to the same exon in the respective populations were considered (UTR-APA only). Binomial test. (B) Length of 3'-UTRs in MPP populations in percent normalized to HSCs. (C) GO term and subsequent REVIGO (Supek et al., 2011) analysis of genes exhibiting 3'-UTR shortening in MPP2 cells compared to HSCs is shown. Bubble color indicates p-value; bubble size indicates frequency of the GO term. For all panels: 3'-Seq data is derived from libraries generated in 4 independent experiments, representing 4 biological replicates per cell population, respectively.

3.2.3 PABPN1 regulates APA in primary hematopoietic cells

After showing that KD of the APA regulator *Pabpn1* impairs HSC function (see 3.2.1) and performing analysis of APA patterns in HSCs and MPP cells (see 3.2.2), we aimed to show that PABPN1 regulates APA in *in vivo*-derived hematopoietic cells. Thus, we performed 3'-Seq analysis of Scr control and *Pabpn1* KD cells. GFP+ LSK cells were sorted 72 h after transduction, and libraries were generated (Figure 3.23 A). Unfortunately, analysis of transplanted HSC or LSK cells was not possible, as only few *Pabpn1* KD cells were found in recipient animals. As previously described (see 3.2.2.2), we filtered the generated datasets (Figure 3.23 B). For Scr control as well as *Pabpn1* KD conditions, the majority of reads was localized in the 3'-UTR region. Again, we could confirm reliability and robustness of our dataset by observing the expected usage patterns of conserved PASs (Figure 3.23 C). We then performed in-depth analysis and observed slightly decreased intronic polyadenylation in KD conditions (Figure 3.23 D). In addition, we observed a decrease in the number of genes showing usage of at least 3 PASs per gene (Figure 3.23 E). When considering our previous results (see 3.2.2.3), this indicates that *Pabpn1* KD LSK cells exhibit less HSC-like polyadenylation patterns. This observation is in line with the decreased HSC potential of the KD cells observed *in vivo*.

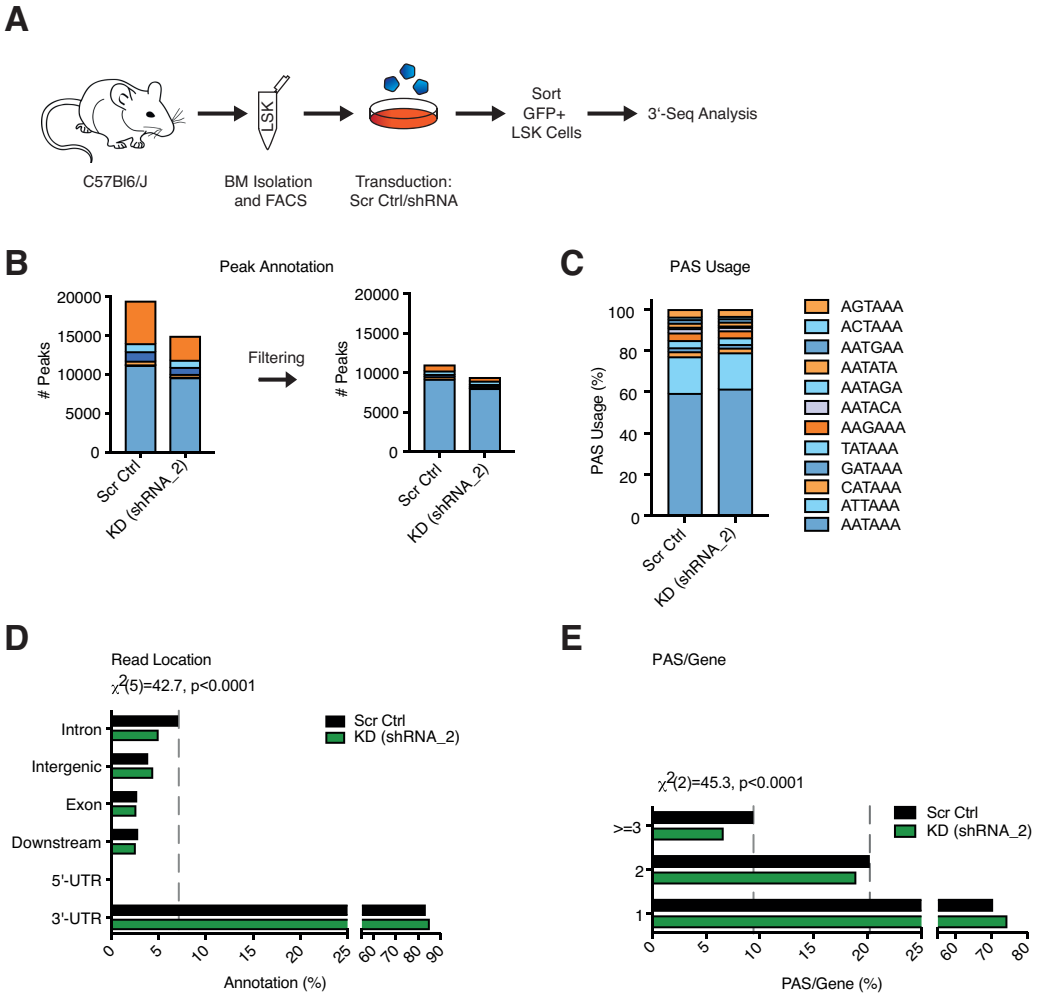


Figure 3.23: *Pabpn1* regulates APA in primary hematopoietic cells – 3'-Seq analysis.
(A) Workflow of 3'-Seq data generation of Scr control and *Pabpn1* KD cells. (B) Read annotation before and after filtering of the 3'-Seq dataset. Number of peaks per population is shown. (C) Usage of different conserved PASs in Scr control and *Pabpn1* KD cells. Percentage of PAS usage is shown. (D) Read classification in percent of Scr control and *Pabpn1* KD cells. Chi-square test. (E) PAS number per gene in percent for Scr control and *Pabpn1* KD cells. Chi-square test. For all panels: 3'-Seq data is derived from libraries generated in 3 independent experiments, representing 3 biological replicates per cell population, respectively.

Again, we called ΔAPAS and identified 40 APA switches in *Pabpn1* KD cells already 72 h after transduction (Figure 3.24 A). We found a mixture of CR-APA (e.g. *Cbx5*, *Cstf3*, *Ak2)* and UTR-APA (e.g. *Slc25a19*, *Parg*, *Cab39l)* (Figure 3.24 B). *Pabpn1* KD has previously been associated with overall 3'-UTR shortening in cell lines (Jenal et al., 2012). As described (see 3.2.2.4), we performed 3'-UTR shortening/lengthening analysis (Figure 3.25). We identified 436 genes with 3'-UTR shortening and 456 with 3'-UTR lengthening, indicating no preference in *Pabpn1* KD cells in our experimental conditions. Collectively, we were able to show that *Pabpn1* does not only regulate APA in established cell lines and disease models (Jenal et al., 2012) but also in primary *in vivo*-derived hematopoietic cells.

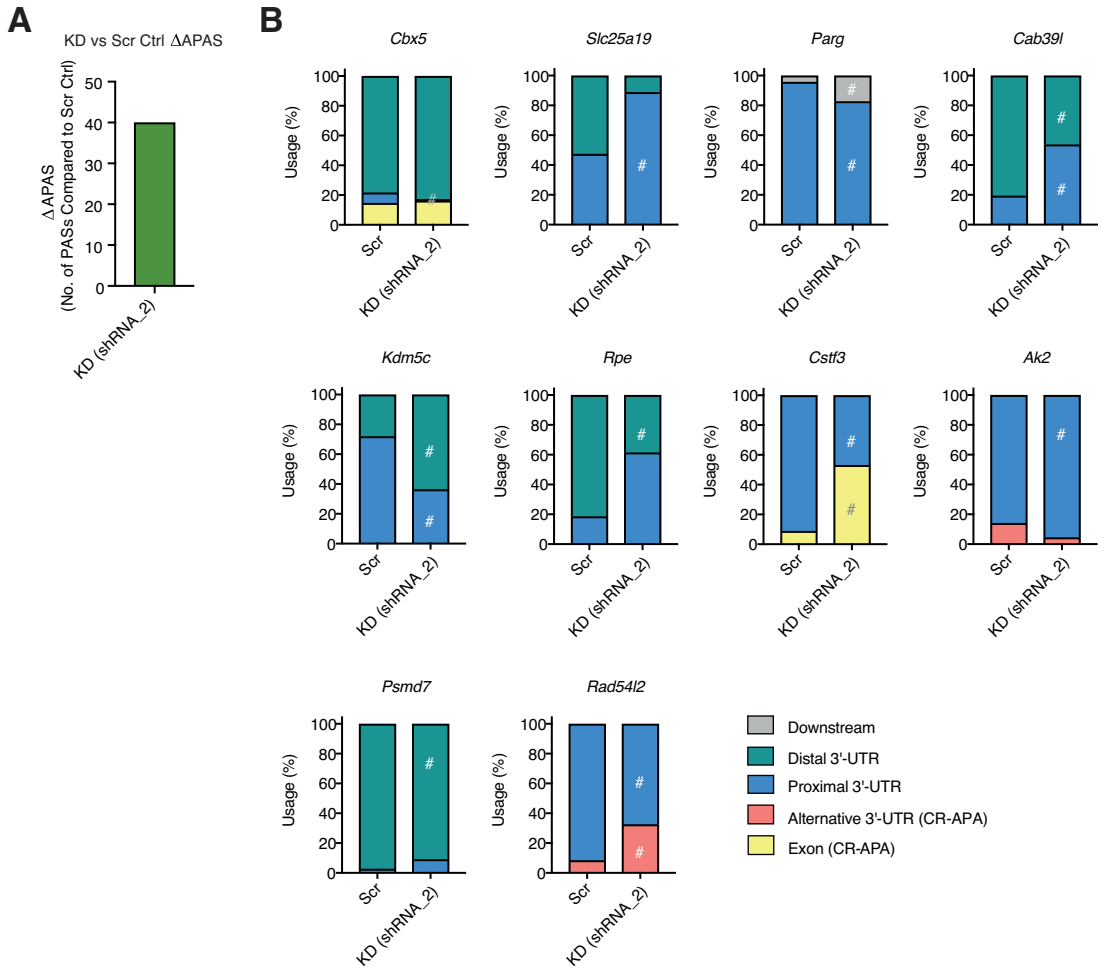


Figure 3.24: *Pabpn1* regulates APA in primary hematopoietic cells – ΔAPAS analysis.
(A) Identification of ΔAPAS in shRNA-mediated *Pabpn1* KD cells compared to Scr control cells. LSK cells were sorted and transduced, GFP+ LSK cells were resorted and used for library generation. (B) APA patterns of the top 10 ΔAPAS hits identified in shRNA-mediated *Pabpn1* KD cells and Scr control cells. Usage of respective site in percent is shown (Percentage of polyadenylation usage (PPAU), see methods section for details). ΔAPAS analysis, #: significant. For all panels: 3'-Seq data is derived from libraries generated in 3 independent experiments, representing 3 biological replicates per cell population, respectively.

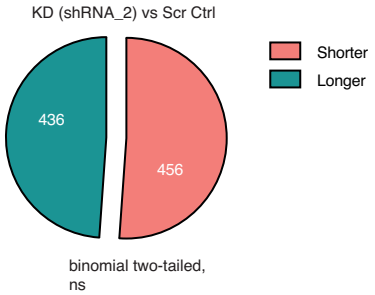


Figure 3.25 *Pabpn1* regulates APA in primary hematopoietic cells – 3'-UTR length analysis.
Analysis of 3'-UTR shortening/lengthening. Only changes in 3'-UTRs belonging to the same exon in the respective population were considered (UTR-APA only). Binomial test. For all panels: 3'-Seq data is derived from libraries generated in 3 independent experiments, representing 3 biological replicates per cell population, respectively.

3.2.4 APA and 3'-UTR shortening occur in HSCs in response to inflammation-induced activation

3.2.4.1 3'-Seq analysis of quiescent and proliferating HSCs

APA has been linked to the proliferative status of cells (Tian and Manley, 2017). In addition, we observed APA and 3'-UTR shortening upon HSC differentiation, which is tightly linked to an increase in cell cycle progression (Figure 3.26 A). Thus, we wondered if activation of HSCs induces APA changes and 3'-UTR length modifications. To this end, we injected mice with pIC, a dsRNA analog, inducing an interferon response leading to HSC activation 16h post injection (Figure 3.26 A) (Essers et al., 2009). Homeostatic steady-state HSCs (sHSC; LK CD150+ CD48- CD34-) and proliferating HSCs (pHSCs; LK CD150+ CD48- CD34-) were sorted in independent biological quadruplicates (Figure 3.26 B-C). Sca-1 was excluded as a marker as this surface protein is upregulated on different cell populations during inflammation-induced activation (Essers et al., 2009; Haas et al., 2015).

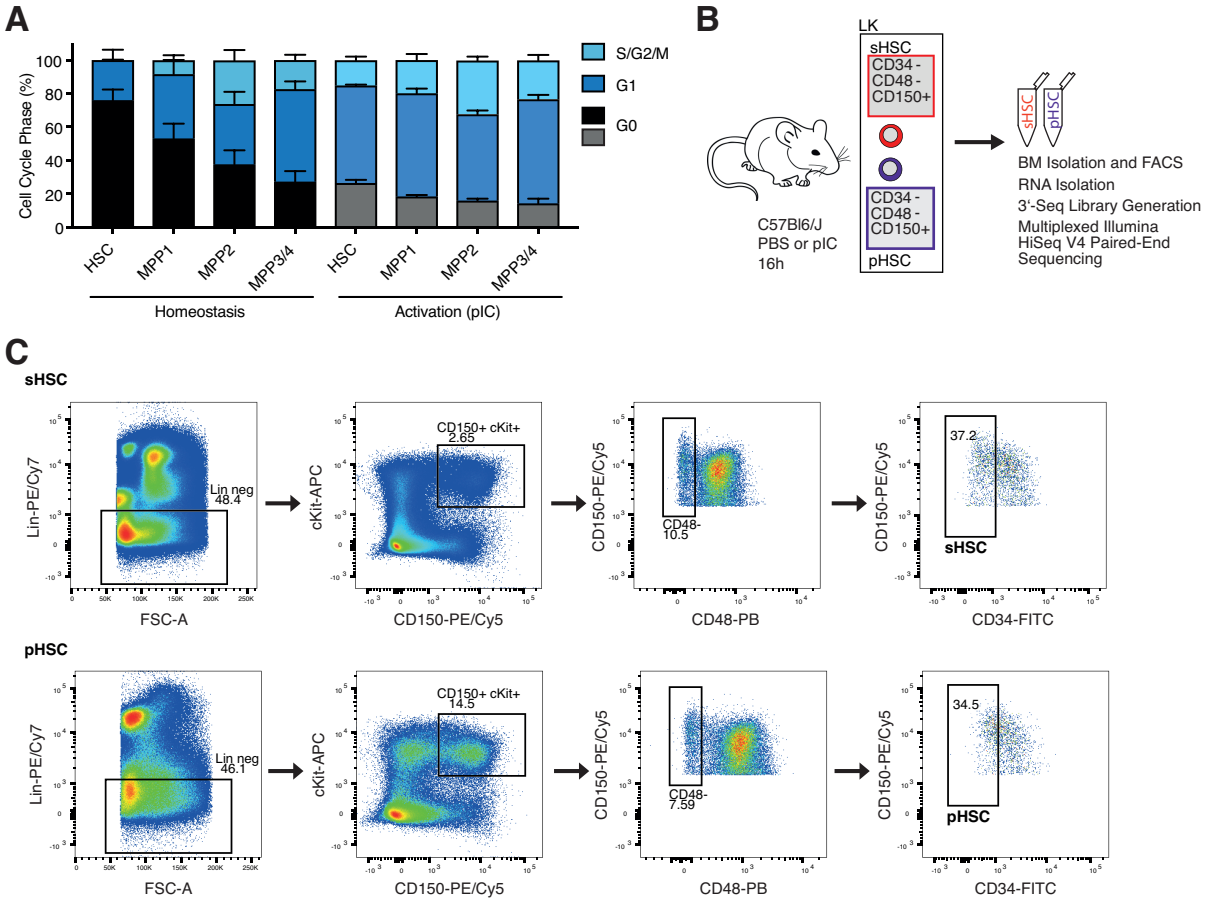


Figure 3.26: 3'-Seq analysis of quiescent and proliferating HSCs – gating strategy.
(A) Flow cytometry-based cell cycle analysis of homeostatic and activated HSCs and MPP cells. Gated on LSK CD150+/- CD48+/- CD34+/-. N = 4-6, mean + SD. (B) Workflow and study design of 3'-Seq analysis of homeostatic/proliferating HSCs (steady state HSC: sHSC; proliferating HSC: pHSC). (C) Representative flow cytometry dot plots of the homeostatic and proliferating HSC sorting approach.

As previously described (see 3.2.2.1), we used paired-end data to evaluate the quality of our generated 3'-Seq dataset. PCA showed robust clustering of biological replicates, again underlining the reliability of our sequencing approach (Figure 3.27 A). Next, we used the single-end read dataset to evaluate APA (see also 3.2.2.2 and 3.2.2.3). Filtering of the dataset led to the expected exclusion of most hits annotated in non-3'-UTR regions (Figure 3.27 B). Again, we observed no differences in PAS motif usage between conditions and observed the polyadenylation signal AAUAAA to be used in more than 50% of polyadenylation events, further underlining the quality of the generated data (Figure 3.27 C). Analysis of read localization revealed a decrease in intronic polyadenylation events in pHSCs (Figure 3.27 D). In addition, we observed a decrease in the number of genes using at least 3 PASs per gene upon activation of HSCs (Figure 3.27 E). Taking our previous results into account (see 3.2.2.3), this indicates that activated pHSCs exhibit a polyadenylation pattern that is more MPP-like compared to homeostatic sHSCs. In summary, we generated a 3'-Seq resource dataset of sHSCs and pHSCs and observed less HSC-like polyadenylation patterns upon cellular activation.

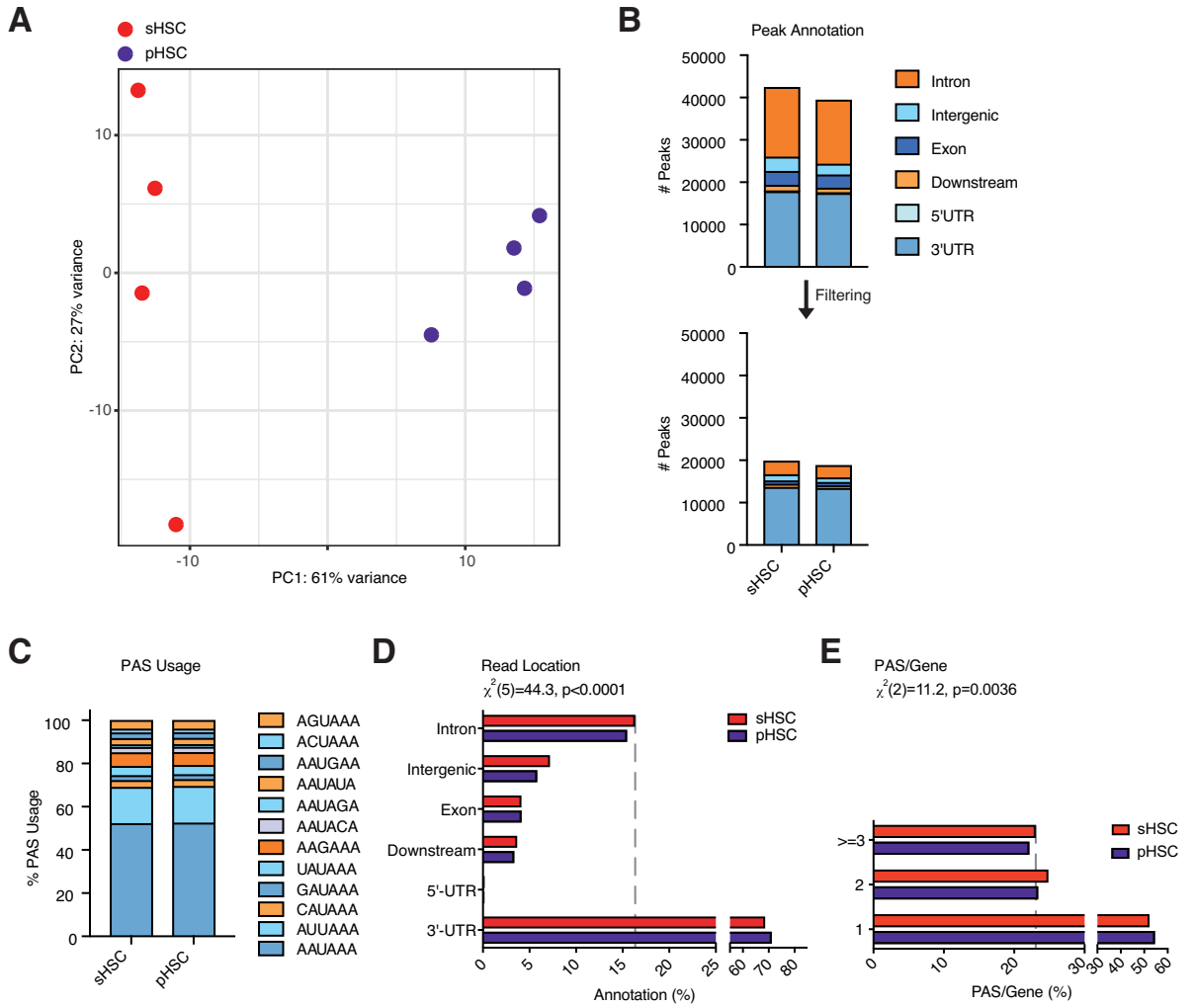


Figure 3.27: 3'-Seq analysis of quiescent and proliferating HSCs – QC.
(A) PCA of the sHSC/pHSC 3'-Seq dataset. (B) Read annotation before and after filtering of the 3'-Seq dataset. Number of peaks per population is shown. (C) Usage of different conserved PASs. Percentage of PAS usage is shown. (D) Read classification in percent. Chi-square test. (E) PAS number per gene in percent for HSCs and MPP cells. Chi-square test. For all panels: 3'-Seq data is derived from libraries generated in 4 independent experiments, representing 4 biological replicates per cell population, respectively.

3.2.4.2 Identification of distinct APA patterns and 3'-UTR shortening upon HSC activation

In the following, we performed ΔAPAS analysis (see also 3.2.2.3). 69 significant ΔAPAS events were identified upon activation of HSCs by pIC (Figure 3.28 A). Again, APA events were highly complex and consisted of CR-APA (*Helz*, *Kansl3*, *Ern1*) as well as UTR-APA (*Rps6kb1*, *Cops2*) (Figure 3.28 B). We conclude that APA is regulated in response to HSC activation and exit of G0.

In line with our expectations, analysis of 3'-UTR length (see 3.2.2.4) revealed 3'-UTR shortening upon activation of HSCs and exit of quiescence (Figure 3.29 A). Over 1300 transcripts exhibited shortening of 3'-UTR length, corresponding to approximately 60% of shared sHSC/pHSC transcripts (Figure 3.29 A-B). Collecitvely, we show that activation of HSCs is linked to APA changes and 3'-UTR shortening *in vivo*.

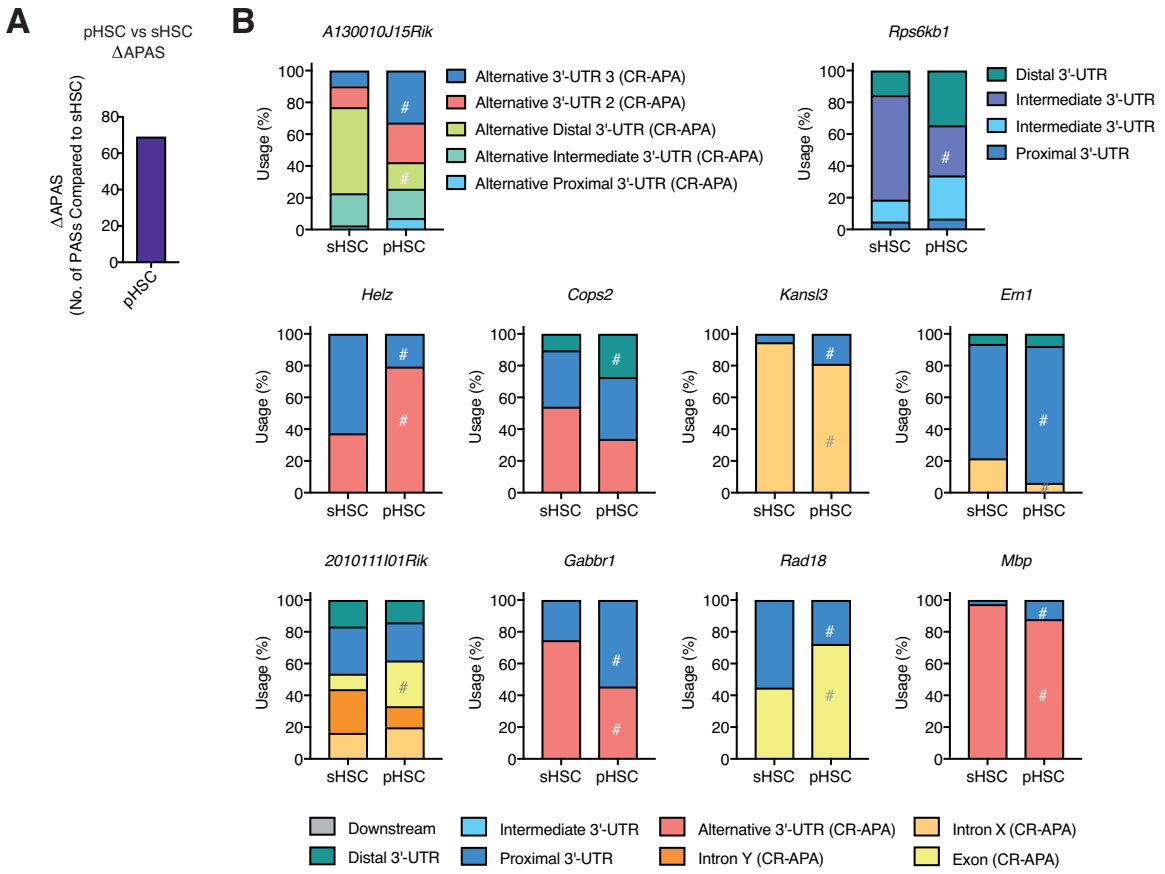


Figure 3.28: Identification of distinct APA patterns upon HSC activation.
(A) Identification of ΔAPAS in proliferating vs homeostatic HSCs by 3'-Seq analysis. (B) APA patterns of the top 10 ΔAPAS hits identified in proliferating and homeostatic HSCs. Usage of respective site in percent is shown (Percentage of polyadenylation usage (PPAU), see methods section for details). ΔAPAS analysis, #: significant. For all panels: 3'-Seq data is derived from libraries generated in 4 independent experiments, representing 4 biological replicates per cell population, respectively.

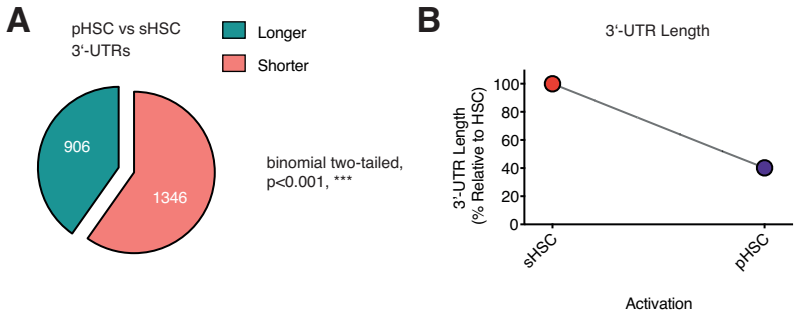


Figure 3.29: Identification of 3'-UTR shortening upon HSC activation.
(A) Analysis of 3'-UTR shortening/lengthening. Only changes in 3'-UTRs belonging to the same exon in the respective populations were considered (UTR-APA only). Binomial test. (B) Length of 3'-UTRs in pHSCs in percent normalized to sHSCs. For all panels: 3'-Seq data is derived from libraries generated in 4 independent experiments, representing 4 biological replicates per cell population, respectively.

3.2.5 The protease *Ctsc* and the immune cell surface receptor *Treml2* are regulated by APA

With robust 3'-Seq datasets at hand, we aimed to identify ΔAPAS hits with potentially functional importance for HSC differentiation and activation. To this end, we generated a Venn diagram to identify ΔAPAS overlaps between MPP1-4 cells compared to HSCs (Figure 3.30 A). The RNA-binding protein *Hnrnpa3* exhibited differential APA in MPP1-4 cells versus HSCs, while *Imp4*, involved in RNA processing, and the protease *Cathepsin C* (*Ctsc)* showed differential APA in MPP2-4 cells versus HSCs. *Ctsc* belongs to the papain enzyme family and encodes a protease (Pham and Ley, 1999). This gene caught our interest, as CTSC is important for cleavage and activation of Granzyme A and B *in vivo*, which have been linked to the LPS-induced stress response in HSCs (Carnevalli et al., 2014; Pham and Ley, 1999). We observed that expression of full-length *Ctsc* was significantly increased by CR-APA in MPP2-4 cells compared to HSCs (Figure 3.30 B-C). HSCs mainly express a shortened *Ctsc* version lacking functional domains. The same trend was observed for sHSCs compared to pHSCs. We confirmed these expression patterns by qPCR (Figure 3.30 D). It has been shown that CTSC secretion is increased in response to LPS (Hamon et al., 2016). Analysis of *in vivo* LPS-treated HSCs revealed a significant APA-mediated switch from the shortened to the functional full-length *Ctsc* isoform (Figure 3.30 D). In summary, differentiation as well as pIC- and LPS-induced activation of HSCs leads to an APA-mediated switch to the full-length *Ctsc* isoform. This switch might be important for Granzyme A and B activation during hematopoietic stress response.

In addition to CR-APA, we also identified UTR-APA only events. For example, we observed a significant switch to the more proximal PAS in MPP2 cells and pHSCs of the gene *triggering receptor expressed on myeloid cells-like 2 (Treml2)* (Figure 3.31 A-B). This switch leads to a decrease in 3'-UTR length of more than 1000 bp. Previously, *Treml2* expression has been shown to be upregulated and re-localized in response to inflammation in myeloid cells (King et al., 2006; Thomas et al., 2016). Similar processes could occur upon HSC differentiation and activation and might be mediated by UTR-APA.

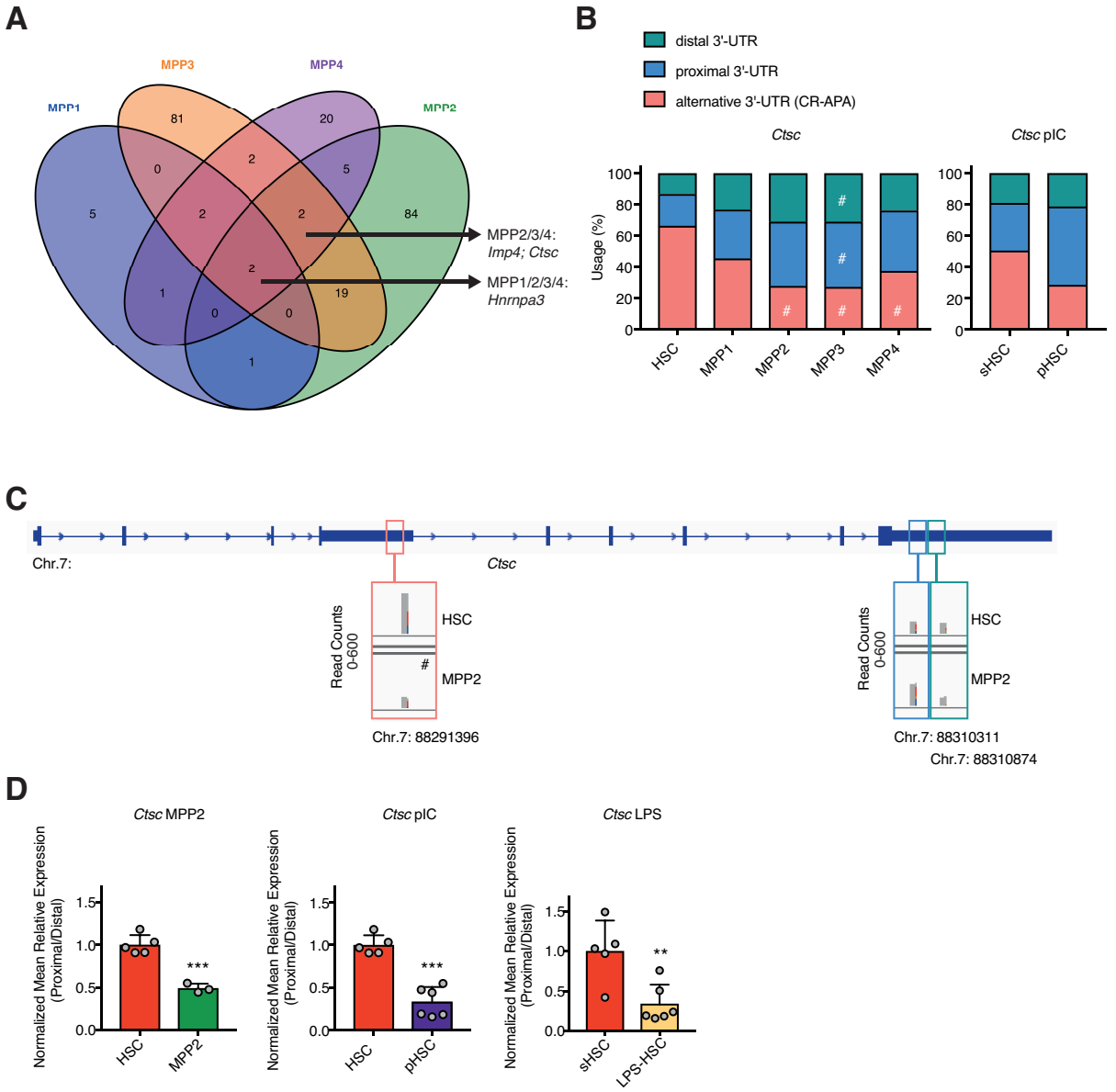


Figure 3.30: The protease *Ctsc* is regulated by APA.
(A) Venn diagram depicting overlap of ΔAPAS between HSCs and the respective MPP population. Numbers indicate ΔAPAS. (B) APA patterns of *Ctsc* in HSCs and MPP cells as well as sHSCs and pHSCs. Proliferation was induced by pIC injection. Usage of respective site in percent is shown (Percentage of polyadenylation usage (PPAU), see methods section for details). ΔAPAS analysis: # significant. (C) Representative example of APA between MPP2 cells and HSCs. Exemplary IGV data of the gene *Ctsc* is shown. (D) qPCR-based analysis of the ΔAPAS hit *Ctsc* in HSCs vs MPP2 cells; HSCs vs pHSCs; sHSCs vs *in vivo* LPS-treated HSCs. Normalized mean relative expression of the proximal to the distal mRNA isoform is shown. Values <1 indicate increase in relative usage of the full-length isoform in the respective population compared to HSCs. Gated on LSK CD150+/- CD48+/- CD34+/-. N = 3-6, unpaired student's t-test. For panels A-C: 3'-Seq data is derived from libraries generated in 4 independent experiments, representing 4 biological replicates per cell population, respectively. For panel D: *p < 0.05, **p < 0.01, ***p < 0.001, ns: not significant; N indicates number of biological replicates; 3 or more independent experiments were performed; mean + SD is shown.

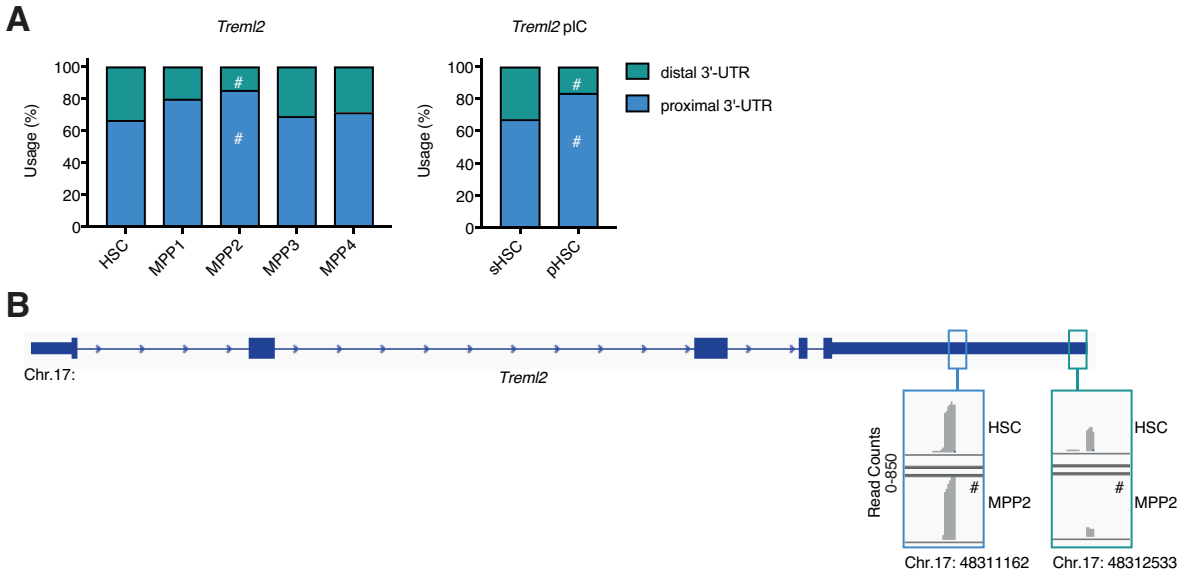


Figure 3.31: The immune cell surface receptor *Treml2* is regulated by APA.
(A) Alternative polyadenylation patterns of *Treml2* in HSCs and MPP cells as well as sHSCs and pHSCs. Usage of respective site in percent is shown (Percentage of polyadenylation usage (PPAU), see methods section for details). ΔAPAS analysis: # significant. (B) Exemplary IGV data of the gene *Treml2* is shown depicting APA between MPP2 cells and HSCs. For all panels: 3'-Seq data is derived from libraries generated in 4 independent experiments, representing 4 biological replicates per cell population, respectively.

3.2.6 APA-mediated *Gls* isoform switching occurs upon HSC activation, and GLS inhibition maintains stem cell potency

3.2.6.1 Identification of Gls isoform switching mediated by APA in response to inflammation-induced HSC activation

In addition to in-depth analysis of the HSC/MPP dataset, we also performed further analysis using the sHSC/pHSC 3'-Seq data. We observed a CR-APA switch in the gene *Gls*, leading to changes in *Gls* isoform expression (Figure 3.32 A-B). GLS is an enzyme metabolizing glutamine to glutamate, which in turn is used for energy production in the TCA cycle in the form of α-KG as well as amino acid and glutathione production (Figure 3.32 C) (Altman et al., 2016). Two different GLS protein isoforms have been described, namely GAC and KGA (Figure 3.32 B) (Altman et al., 2016), and their expression is regulated by CR-APA. We detected a relative increase in GAC levels compared to KGA levels in pHSCs vs sHSCs and observed the same trend in MPP cells (Figure 3.32 A). We could confirm this switch to the GAC isoform upon HSC differentiation and activation by qPCR analysis (Figure 3.32 D). It has previously been reported that a relative increase in GAC levels is associated with higher overall GLS protein levels (van den Heuvel et al., 2012). To analyze the effect of the GAC:KGA ratio on GLS protein expression in HSC/MPP and sHSC/pHSC populations, we performed intracellular flow cytometry staining of GLS (Figure 3.32 E). Interestingly, we observed increased GLS levels in MPP cells and activated HSCs (Figure 3.32 F). In addition to the APA-mediated *Gls* isoform switch, the process *glutamine metabolism* was enriched in the MPP2

shortened gene list (see 3.2.2.4). Thus, we decided to concentrate on the metabolic gene *Gls* and the reported isoform switch for functional validation of ΔAPAS hits.

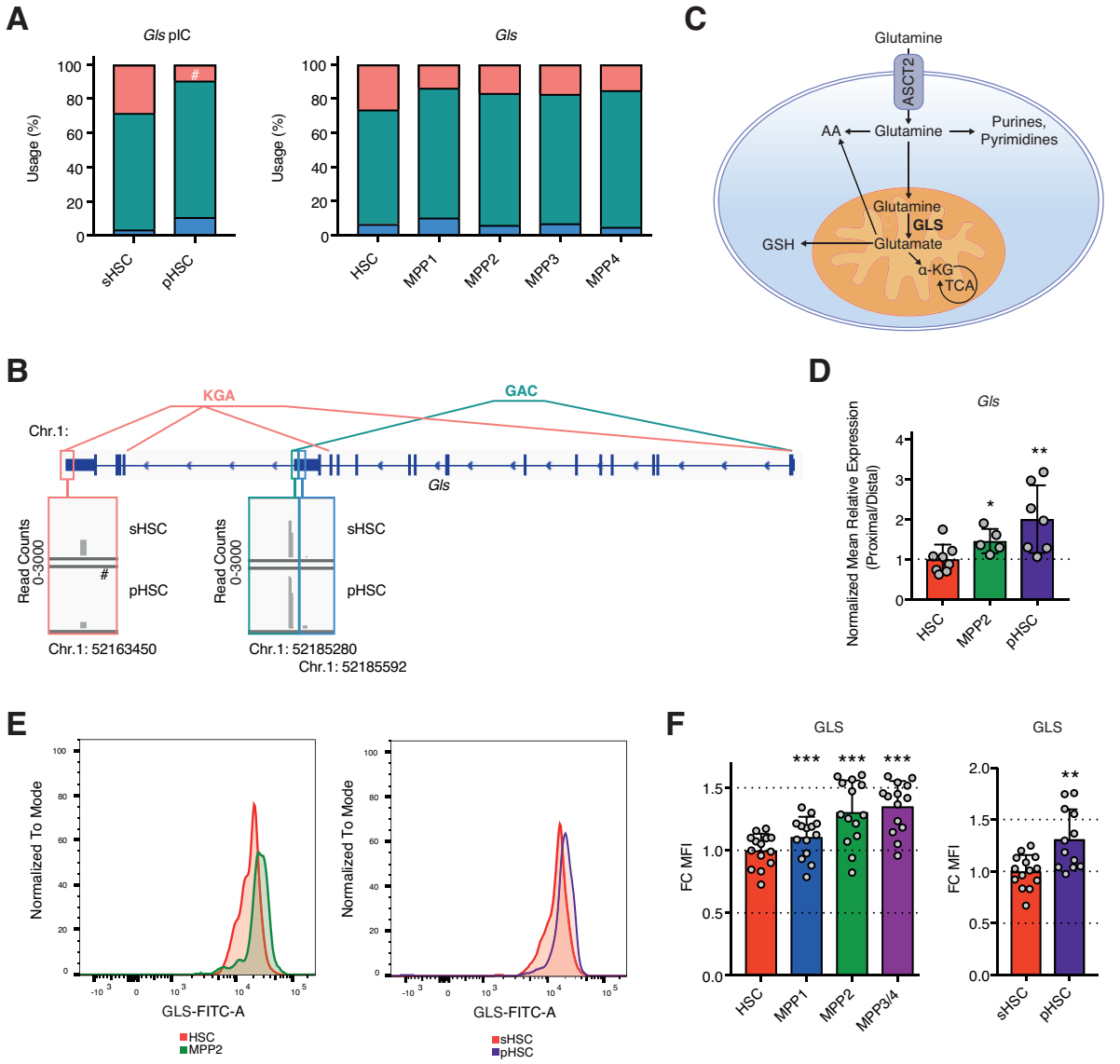


Figure 3.32: Identification of APA-mediated *Gls* isoform switching in response to inflammation-induced HSC activation.
(A) APA patterns of *Gls* in HSCs and MPP cells as well as sHSCs and pHSCs. Usage of respective site in percent is shown (Percentage of polyadenylation usage (PPAU), see methods section for details). ΔAPAS analysis, #: significant. (B) Representative example of APA between pHSCs and sHSCs. Exemplary IGV data of the gene *Gls* is shown. (C) Schematic representation of glutaminolysis. (D) qPCR-based analysis of the ΔAPAS hit *Gls* in HSCs, MPP2 cells and pHSCs. Normalized mean relative expression of the proximal to the distal mRNA isoform is shown. Values >1 indicate increased relative usage of the GAC isoform in MPP2 cells/pHSCs compared to HSCs. Gated on LSK CD150+/-CD48+/- CD34+/-. N = 5-8, unpaired student's t-test. (E) Representative histogram plots of GLS expression in HSCs and MPP2 cells as well as sHSCs and pHSCs. (F) Flow cytometry-based analysis of GLS protein expression in HSCs/MPP cells (LSK CD150 +/- CD48+/- CD34+/-) and in sHSCs/pHSCs (LK CD150+ CD48- CD34-). Fold change (FC) of the mean fluorescent intensity (MFI) is shown. Normalized to HSCs/sHSCs. N = 12-14, left panel: paired student's t-test, right panel: unpaired student's t-test. For panels A-B: 3'-Seq data is derived from libraries generated in 4 independent experiments, representing 4 biological replicates per cell population, respectively. For panel D-F: *p < 0.05, **p < 0.01, ***p < 0.001, ns: not significant; N indicates number of biological replicates; 3 or more independent experiments were performed; mean + SD is shown.

3.2.6.2 Short-term inhibition of GLS in vitro maintains HSC potency

We observed an isoform switch of the gene *Gls* to the metabolically more active GAC isoform, leading to an overall increase in GLS protein levels. Therefore, we hypothesized that these findings might be indicative of a metabolic reprogramming towards glutaminolysis upon HSC differentiation and activation. This reprogramming might be a prerequisite for HSC activation. To test our hypothesis, we first made use of the GLS inhibitor CB-839, currently in clinical trials. We aimed to investigate the effects of GLS inhibition on cell cycle status and colony forming potential of HSPCs (Figure 3.33 A). Initially, we performed depletion of Lin+ cells to enrich for HSCs and MPP cells. Depleted cells were cultured *in vitro* for 48 h and treated with different concentrations of CB-839 or DMSO. Subsequently, we performed cell cycle analysis by Ki-67/DAPI staining and viability analysis using 7AAD/Annexin (Figure 3.33 B-C). In these dose escalation studies, we observed the most prominent effects using 5 μM CB-839 (Figure 3.33 B). This concentration did not affect cell viability if compared to untreated or DMSO-treated controls (Figure 3.33 C). For cell cycle analysis, BM cells were lineage-depleted and cultured *in vitro* in the presence of 5 μM CB-839 or DMSO (Figure 3.33 A). 48 h later, we performed cell cycle analysis by flow cytometry and gated HSCs and MPP1-4 cells (Figure 3.33 D-E). In general, culture of HSCs and MPP cells leads to exit from the G0 state and active proliferation due to the differentiation-inducing *in vitro* conditions. This effect was observed in DMSO- and CB-839-treated MPP cells. Interestingly, analysis of HSCs exhibited an increase in the frequency of cells in G0 upon GLS inhibition, indicating that activation and repression of glutaminolysis are critical for controlling the balance between proliferation and quiescence of HSCs. To test if this cell cycle effect is also functionally relevant, we performed CFU assays (Figure 3.33 A). LSK cells were sorted and cultured in presence of CB-839 or DMSO. 72 h later, pre-treated cells were harvested and used for standard CFU assays in the absence of CB-839 (Figure 3.33 F). In primary platings, we observed reduced colony-forming potential in CB-839 pre-treated conditions. However, in secondary platings, CB-839 pre-treated cells formed significantly more colonies compared to DMSO pre-treated, controls indicating higher maintenance of HSC function. In CFU assays, HSC potential is assessed in secondary platings, as colony formation in primary platings is mainly dependent on MPP cells. Thus, we conclude that *in vitro* treatment of HSCs with the GLS inhibitor CB-839 prevents glutaminolysis, thereby limiting HSC proliferation. In turn, limited HSC proliferation leads to increased maintenance of HSC potential in the otherwise HSC-exhausting *in vitro* culture conditions.

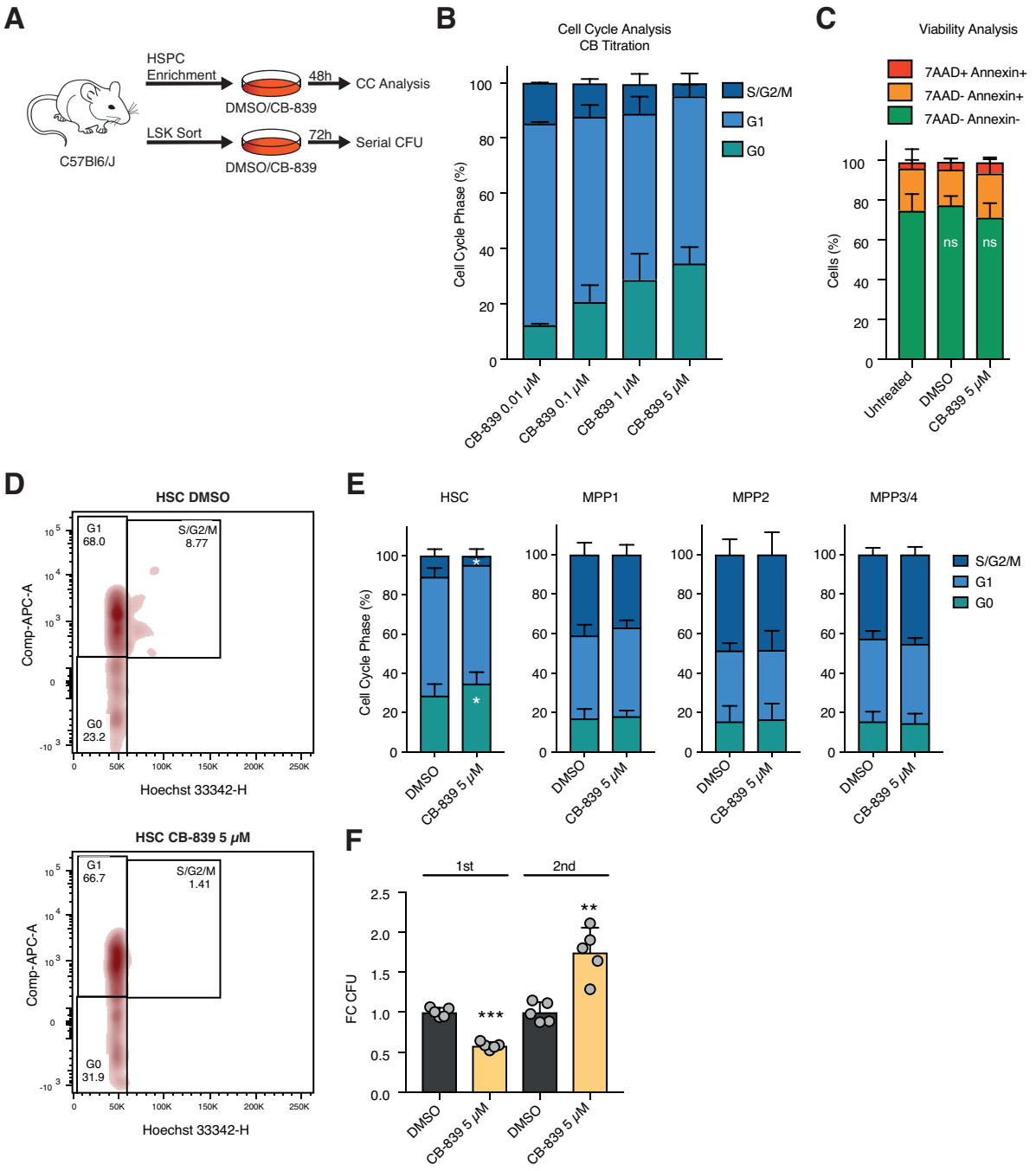


Figure 3.33: Short-term inhibition of GLS *in vitro* maintains HSC potency.
(A) Workflow of CB-839 *in vitro* treatment. (B) Flow cytometry-based cell cycle analysis for titration of the GLS inhibitor CB-839. HSCs were gated and analyzed 48 h after treatment. Greater than or equal to 1 independent experiment was performed. N = 2-7. (C) Flow cytometry-based viability analysis upon CB-839-treatment. N = 5-7, two-way ANOVA. (D) Representative flow cytometry density plots of cell cycle analysis in control and CB-839 treated HSCs. Gated on LSK CD150+ CD48- CD34-. (E) Flow cytometry-based cell cycle analysis of *in vitro* CB-839 treated HSPCs gated on HSCs and MPP cells (LSK CD150+/- CD48+/- CD34+/-). Analysis performed after 48 h. N = 7, two-way ANOVA. (F) CFU assay of control and CB-839-treated LSK cells represented as fold change (FC) in colony numbers for first and second plating. Plating was performed 72 h after start of culture and treatment. N = 5, unpaired student's t-test. For all panels: *p < 0.05, **p < 0.01, ***p < 0.001, ns: not significant; N indicates number of biological replicates; 2 or more independent experiments were performed, unless otherwise indicated; mean + SD is shown.

3.2.6.3 NUDT21 mediates Gls isoform switching in hematopoietic cells

The APA regulator NUDT21, also called CFIM25, has previously been reported to control *Gls* isoform switching in HeLa cells (Masamha et al., 2016). Masamha et al. could show that relative expression of KGA is increased upon KD of *Nudt21*, while the GAC isoform gets downregulated. We aimed to investigate if NUDT21 regulates *Gls* isoform switching in hematopoietic cells. To this end, we sorted LSK cells and performed lentiviral *Nudt21* KD (Figure 3.34 A) (see also 3.2.1.2). Successfully transduced cells express GFP and were resorted to perform KD validation and *in vitro* functional analysis. First, we performed KD validation by qPCR analysis (Figure 3.34 B). Using three different shRNAs (IDs: 10, 4, 2), we observed significant downregulation (~50%) of *Nudt21* 48 h after transduction. Next, we aimed to analyze the effect of *Nudt21* KD on *Gls* isoform expression. To this end, we analyzed GAC and KGA levels by using isoform-specific TaqMan probes for qPCR analysis (Figure 3.34 C). Normalized proximal (GAC)/distal (KGA) expression levels were analyzed. A relative increase in GAC levels is indicated by elevated numbers of proximal transcripts (GAC) leading to relative expression values above 1, while values below 1 indicate a relative increase in distal transcripts (KGA) (see also 3.2.2.3). Analysis of Scr control samples revealed a relative increase in GAC levels upon *in vitro* culture of LSK cells. These results suggest that the differentiation- and activation-inducing *in vitro* culture conditions lead to a *Gls* isoform switch to the active GAC isoform, indicating enhanced glutaminolysis. These observations are in line with our previous experiments in which we observed inhibition of glutaminolysis to be beneficial for maintaining HSCs by limiting exit of the G0 cell cycle state. In our *Nudt21* KD conditions, we observed a diminished increase in relative GAC expression, indicating that *Nudt21* is needed for *Gls* isoform switching and GAC expression in hematopoietic cells. Taken together, we showed that *Nudt21* mediates *Gls* isoform switching in the HSC/MPP compartment and reduced NUDT21 levels lead to decreased relative GAC expression.

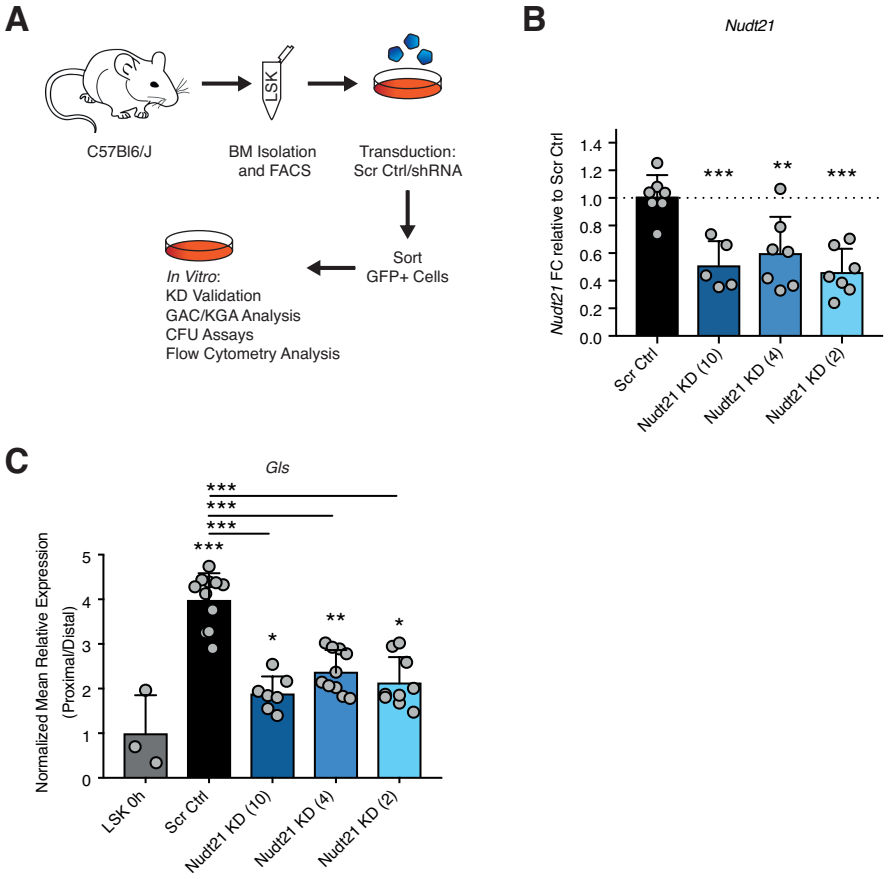


Figure 3.34: NUDT21 mediates *Gls* isoform switching in hematopoietic cells.
(A) Workflow of *Nudt21* KD and subsequent *in vitro* analysis. (B) qPCR-based analysis of *Nudt21* expression in Scr control and KD cells after 48 h. N = 5-7, unpaired student's t-test. (C) qPCR-based analysis of *Gls* expression in LSK cells (0h), Scr control cells (48 h *in vitro*, resorted GFP+ cells) and KD cells (48 h *in vitro*, resorted GFP+ cells). Normalized mean relative expression of the proximal to the distal mRNA isoform is shown. Values >1 indicate increase in relative usage of the GAC isoform in cultured cells compared to LSK cells *in vivo*. Significance levels compared to LSK 0h unless otherwise indicated. N = 3-11, unpaired student's t-test. For all panels: *p < 0.05, **p < 0.01, ***p < 0.001, ns: not significant; N indicates number of biological replicates; 3 or more independent experiments were performed; mean + SD is shown.

3.2.6.4 Nudt21 KD inhibits glutaminolysis by restricting GLS levels and impairs HSC function

In a next step, we aimed to investigate the functional importance of NUDT21-mediated *Gls* isoform switching. First, we analyzed the effect of decreased isoform switching on GLS protein levels by intracellular flow cytometry analysis 72 h after LSK transduction (Figure 3.35 A-B). *In vitro* cultured Scr control cells exhibited an overall increase in GLS protein levels compared to freshly isolated LSK cells. This is in line with the relative increase in GAC expression levels we reported (Figure 3.34 C). KD of *Nudt21* led to a diminished increase in relative GAC levels and, accordingly, we observed only a declined increase in GLS protein levels. Thus, NUDT21 seems to be essential for mediating *Gls* isoform switching to GAC, thereby leading to an increase in GLS protein levels in hematopoietic cells. These increased levels are needed, as glutaminolysis seems to be key for complete HSC activation. In turn, prevention of *Gls* switching should lead to impaired HSC function. To proof this hypothesis, we performed serial

CFU assays using Scr control and *Nudt21* KD cells 48 h after LSK transduction (Figure 3.35 C). *Nudt21* KD cells formed significantly less colonies and exhausted in serial platings. Flow cytometry analysis after the first plating revealed maintenance of LSK cell frequency in *Nudt21* KD compared to Scr control settings, indicating that lack of colony formation is caused by a block in differentiation (Figure 3.35 D). In summary, irreversible prevention of *Gls* switching by KD of *Nudt21* maintains LSK cells, which, however, are not able to form colonies. The block of differentiation might be caused by prevention of progenitor proliferation and expansion. This block is potentially caused by inhibited metabolic switching supported by our observation of reduced GLS protein levels.

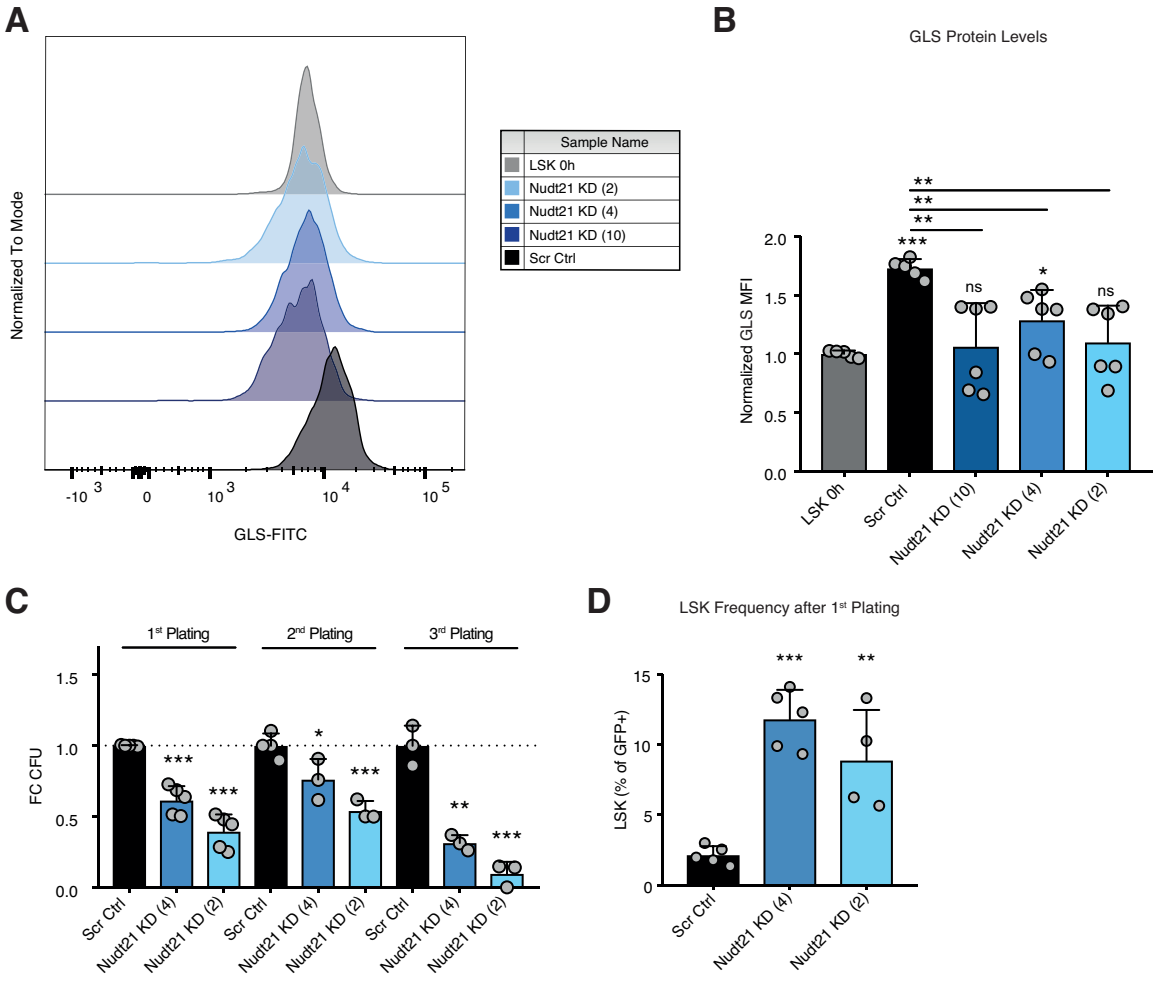


Figure 3.35: *Nudt21* KD inhibits glutaminolysis by restricting GLS levels and impairs HSC function.
(A) Representative flow cytometry plots of GLS expression in LSK cells (0h), Scr control (72 h *in vitro*) and *Nudt21* KD (72 h *in vitro*) cells. (B) Flow cytometry-based analysis of GLS protein expression in LSK cells (0h), Scr control cells (72 h *in vitro*, resorted GFP+ cells) and KD cells (72 h *in vitro*, resorted GFP+ cells). Fold change (FC) of the mean fluorescent intensity (MFI) is shown. Normalized to LSK *in vivo* cells. Significance levels compared to LSK 0h unless otherwise indicated. N = 5-6, unpaired student's t-test. (C) CFU assay of Scr control and *Nudt21* KD cells showing fold change (FC) in colony numbers for first, second and third plating. N = 3-5, unpaired student's t-test. Significance levels compared to Scr control. (D) Flow cytometry-based analysis of the LSK content in cells from CFU assays after primary plating. Percent of LSK cells of all GFP+ cells is shown. N = 4-5, unpaired student's t-test. Significance levels compared to Scr control. For all panels: *p < 0.05, **p < 0.01, ***p < 0.001, ns: not significant; N indicates number of biological replicates; 2 or more independent experiments were performed; mean + SD is shown.

Overall, we show in this study that the APA regulator PABPN1 is essential for HSC function using extensive *in vitro* and *in vivo* functional assays (Figure 3.36). The advancement in ultra-low input 3'-Seq techniques enabled us to analyze APA globally in rare BM populations, generating a 3'-UTR/APA resource dataset of HSPCs. We report APA upon HSC commitment and activation. 3'-UTR shortening occurs upon stem cell differentiation and upon exit of G0. In addition, we also demonstrate the functional importance of APA *in vivo* by reporting an APA switch in the metabolic gene *Gls* to be essential for HSC activation.

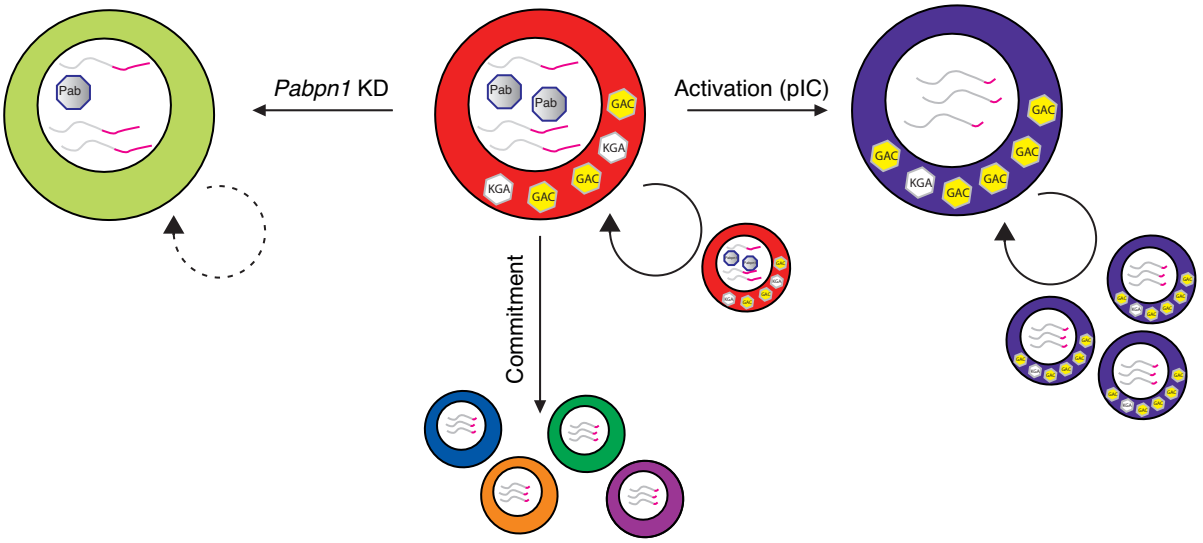


Figure 3.36: A model of the role of APA in the HSC/MPP compartment.
Schematic model of 3'-UTR shortening in the hematopoietic system upon HSC commitment/differentiation and activation during inflammatory response induced by pIC. Interfering with APA by KD of the APA regulator *Pabpn1* leads to loss of HSC function. During stress response, HSCs become activated and switch to the *Gls* isoform GAC by APA. This leads to a metabolic shift and dependency on glutaminolysis. Pab: PABPN1.

3.3 Generation of *Pabpn1*^{flox} mice

In order to facilitate functional research on APA and the APA regulator PABPN1, we aimed to generate an inducible *Pabpn1* KO mouse model. Such a model has the advantage that it can be used in combination with tissue-specific Cre drivers and in disease models, for example leukemia. Thereby, analysis of impaired APA in different cell types *in vivo* and in different functional contexts is feasible.

3.3.1 Design of a *Pabpn1*^{flox} mouse model and gRNA testing

3.3.1.1 Pabpn1 targeting approach

First, we designed our targeting approach. Deletion of the floxed *Pabpn1* region should lead to complete loss of *Pabpn1* expression from the targeted allele. To this end, we checked *Pabpn1* isoforms and functional domains using Ensembl (www.ensembl.org). Several isoforms are reported, and all of them share expression of exon 3 and 4 of the most common *Pabpn1* isoform (ENSMUST00000022808.13). In addition, the functionally important RNA binding domains are localized in these exons. Therefore, we decided to introduce loxP sites flanking exon 3 and 4. We aimed to use the CRISPR/Cas9 method as described by Miura et al. and Quadros et al. (Miura et al., 2018; Quadros et al., 2017). Briefly, long ssDNA donor templates are injected together with pre-assembled crRNA-tracrRNA:Cas9 ribonucleoprotein complexes into mouse zygotes. Excision and recombination take place in the 1-cell or early 2-cell state. As recombination occurs directly *in vivo*, we aimed to increase efficiency by pre-testing our gRNA efficiency *in vitro*. To this end, we collaborated with Dr. Kai Schönig (ZI Mannheim). 3 gRNAs per targeting locus were designed (upstream of exon 3: gRNAs A1-3; downstream of exon 4: gRNAs B1-3) using CRISPOR (http://crispor.tefor.net/) (Haeussler et al., 2016) (Figure 3.37 A). The potential targeting sequence was at least 250 bp upstream or downstream of the respective exon to exclude interference of the introduced loxP sites with splicing.

3.3.1.2 gRNA testing in vitro

To test gRNA efficiency, we used a disrupted *lacZ* reporter gene with the respective targeting sequence (gRNAs A1-3 or gRNAs B1-3) and measured reporter signal after *in vitro* targeting (Figure 3.37 B). First, we ordered gRNA oligonucleotides and reporter oligos. The reporter oligos correspond to the *Pabpn1* introns that should be targeted by the respective gRNAs. The gRNA oligonucleotides were annealed and cloned into the px330 Cas9 vector, enabling gRNA/Cas9 expression. The targeting sequences were inserted between the *lacZ* fragments in the pCRISPR report vector (targeting sequence oligo A or B). HeLa cells were transfected with the px330 Cas9 vector encoding for gRNA A1, A2, A3, B1, B2 or B3. Co-transfection with

the respective pCRISPR reporter plasmid and a luciferase-encoding control plasmid (pUHC131.1) was performed. 24 h later, cells were lysed and LacZ as well as luciferase activity were measured. The luciferase signal was used for normalization to correct for differences in transfection efficiency and cell number. The normalized LacZ signal ((LacZ signal/luciferase signal) x mean luciferase signal) of the respective tested gRNA is a direct indicator of gRNA targeting efficiency. Testing revealed the highest efficiency for gRNAs A3 and B2 (Figure 3.37 C).

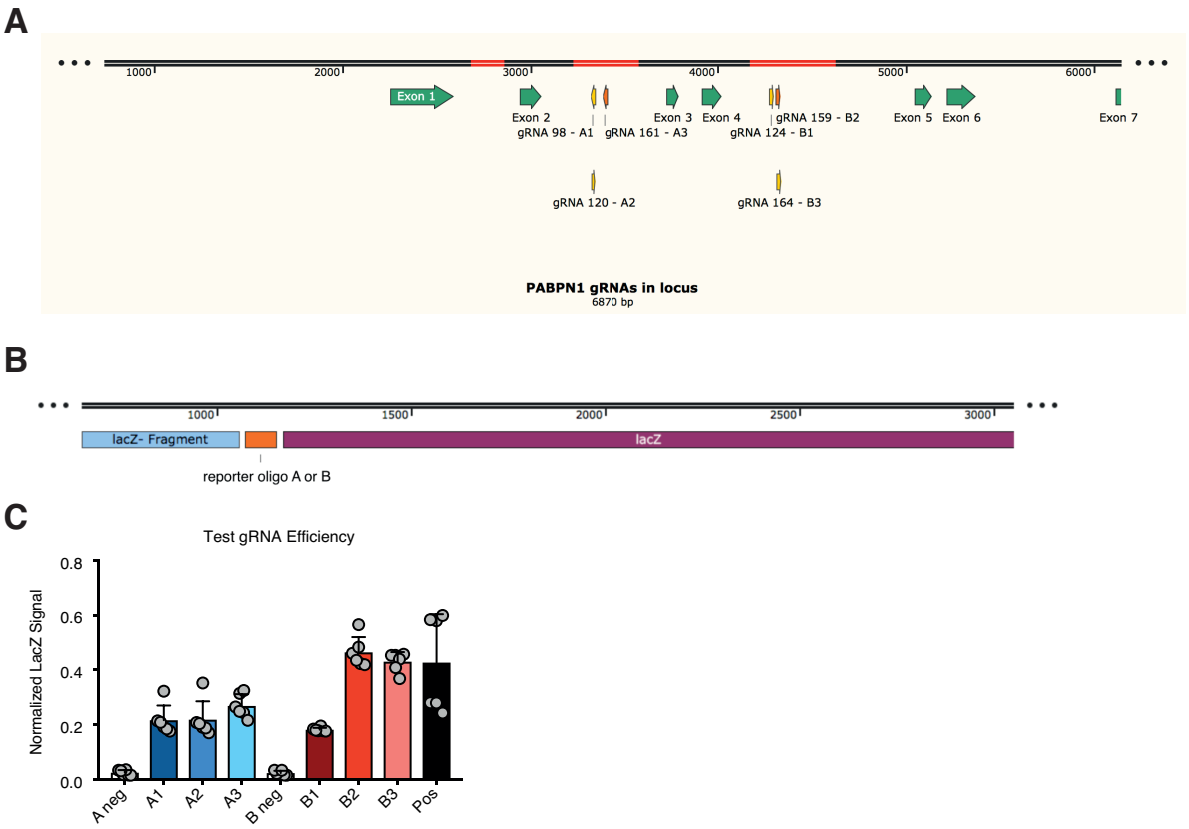


Figure 3.37: Testing of gRNA efficency.
(A) gRNA design based on *Pabpn1* isoforms and PAM localization. gRNAs were designed using CRISPOR (http://crispor.tefor.net/) (Haeussler et al., 2016). (B) Schematic representation of the *lacZ* fragments in the pCRISPR report vector containing targeting sequence oligo A or B. This vector is used to test gRNA efficiency. (C) Testing of gRNA efficiency. Normalized LacZ signal is shown. N = 6. N indicates number of biological replicates; 2 independent experiments were performed; mean + SD is shown.

3.3.2 Generation of *Pabpn1*^{flox} mouse generation

3.3.2.1 Pronucleus injections

Initially, we injected two px330 Cas9 vectors encoding gRNA A3 and B2, respectively, together with Cas9 mRNA and a circular 6 kb plasmid containing the targeting sequence with integrated loxP sites flanked by ~1300 homology arms on both sides. This approach, however, failed to generate transgenic mice. In a next step, we decided to use the CRISPR/Cas9 method as described by Miura et al. and Quadros et al. (Miura et al., 2018; Quadros et al., 2017). We ordered crRNA and tracrRNA for *in vivo* injections corresponding to gRNA A3 and B2, as these

have been shown to be more efficient than single-molecule guide RNAs (Miura et al., 2018; Quadros et al., 2017). In addition, we ordered a long ssDNA megamer®. In this 1126 nucleotide (nt) long megamer, a 60 nt homology arm was followed by the upper loxP site, exon 3 and 4 including intronic sequences, the lower loxP site and a 60 nt downstream homology arm. crRNA/tracrRNAs were annealed to obtain active gRNAs and mixed with Cas9 protein and the ssDNA template. This mix was injected into one of the pronuclei of C57Bl6/J zygotes, which were subsequently transplanted into pseudo-pregnant foster mice. Injections were performed by the transgenic service of the DKFZ.

3.3.2.2 Genotyping of transgenic mice

From two rounds of injections, 17 mice were born. In order to identify potential founder mice, we developed a stringent genotyping strategy. In a first step, integration of the upstream (us) and the downstream (ds) loxP site was evaluated by PCR1 and PCR2, respectively. Successful integration of loxP sites leads to a band shift of 36 bp. This size difference can be difficult to detect in gel analysis. Thus, in addition to loxP sites, we also introduced restriction enzyme sites during design of the ssDNA template. Only PCR products generated by amplification of the floxed sequence can be digested. Analysis revealed 7 animals with integration of both loxP sites (Table 3.1; Appendix Figures 1-2). Only 7 of these animals showed integration of upper and lower loxP and were used for further analysis (IDs: 3, 5, 7, 11, 14, 15, 16). Next, PCR3 and subsequent digestion was performed to check if both loxP sites were localized on the same allele (Table 3.1; Appendix Figure 3). Only four animals passed this QC (IDs: 5, 11, 14, 15). To check for targeted integration in contrast to random integration, we designed and performed PCR4 (upper loxP) and PCR5 (lower loxP) and subsequent digestion (Table 3.1; Appendix Figures 4-5). All four animals passed this QC step. To exclude the presence of point mutations, we performed amplicon sequencing of PCR2 and 4 (Table 3.1). In total, 3 animals (IDs: 5, 11, 15) passed this final QC. The mouse with ID5 was crossed to C57Bl6/J wildtype animals to generate heterozygous offspring. These mice were again extensively genotyped, and loxP integration was verified (data not shown). Heterozygous mice were crossed, generating homozygous Pabpn1$^{flox/flox}$ mice. These mice are currently being crossed to Cre drivers (e.g. VavCre, MxCre) to validate the function of the integrated loxP sites. In summary, we successfully generated transgenic *Pabpn1*^{flox} mice by CRISPR/Cas9-based targeting of zygotes.

Table 3.1: Genotyping of transgenic mice.
Overview of *Pabpn1*^{flox} mouse generation and genotyping results. (us: upstream; ds: downstream; ND: not determined; Het: heterozygous; Hom: homozygous; WT: wildtype)

ID	Sex	Birthdate	PCR 1 + digest (us loxP)	PCR 2 + digest (ds loxP)	PCR 3 + digest (loxP on same allele)	PCR 4 + digest (us localization correct)	PCR 5 + digest (ds localization correct)	Seq PCR 2	Seq PCR 4
1	M	23.04.18	WT	?	ND	ND	ND	ND	ND
2	F	23.04.18	WT	WT	ND	ND	ND	ND	ND
3	M	23.04.18	Hom	Hom	failed	ND	failed	ND	ND
4	M	02.05.18	WT	WT	ND	ND	ND	ND	ND
5	M	02.05.18	Het	Het	passed	passed	passed	passed	passed
6	M	02.05.18	WT	WT	ND	ND	ND	ND	ND
7	M	02.05.18	Het	Het	failed	ND	failed	ND	ND
8	F	02.05.18	WT	WT	ND	ND	ND	ND	ND
9	F	02.05.18	WT	Het	ND	ND	ND	ND	ND
10	M	02.05.18	WT	Het	ND	ND	ND	ND	ND
11	M	02.05.18	Het	Het	passed	passed	passed	passed	passed
12	M	02.05.18	WT	WT	ND	ND	ND	ND	ND
13	F	02.05.18	WT	WT	ND	ND	ND	ND	ND
14	F	02.05.18	Het	Het	passed	passed	passed	passed	failed
15	F	02.05.18	Het	Het	passed	passed	passed	passed	passed
16	M	02.05.18	Het	Hom	failed	ND	failed	ND	ND
17	F	02.05.18	WT	WT	ND	ND	ND	ND	ND

4 Discussion

4.1 The role of the lncRNA *Meg3* in HSCs

In recent years, multiple studies could show that specific lncRNAs play crucial roles in HSCs and the hematopoietic compartment (Alvarez-Dominguez and Lodish, 2017; Luo et al., 2015; Yildirim et al., 2013). As *Meg3* was very highly and specifically expressed in HSCs, we aimed to unravel the function of *Meg3* in hematopoiesis. In our experiments, we could not detect any impairments upon loss of *Meg3* expression in the hematopoietic compartment.

4.1.1 *Meg3* is dispensable for HSCs

Several KO mouse models were described in the past, showing that *Meg3* is important for embryonic development (Lin et al., 2003; Takahashi et al., 2009; Zhou et al., 2010). Depending on the targeting strategy, different developmental phenotypes were observed (see also 1.2.4). Deletion of IG-DMR (Lin et al., 2003) and deletion of the *Meg3* promoter, part of the *Meg3*-DMR and exon 1-5 (Zhou et al., 2010) exhibited highly similar phenotypes. In both models, maternal inheritance of the deletion was embryonically lethal and led to loss of imprinting. Inheritance of the deletion from the paternal side, however, was only associated with minor phenotypes. In contrast, the model of Takahashi et al., harboring a deletion of a 10 kb region including the *Meg3*-DMR, the *Meg3* promoter and exon 1-5, presented a very different developmental phenotype (Takahashi et al., 2009). In this model, no loss of imprinting was observed. Mice that inherited the deletion from the maternal side died 4 weeks after birth, while paternal inheritance led to partial embryonic lethality. This phenotype is presumably caused by residual activity of the Neomycin promoter, still present from the mouse targeting approach, which drives expression of the other ncRNA clusters in the *Dlk1-Meg3* locus in paternal and maternal deletion mice, respectively (Takahashi et al., 2009; Zhou et al., 2010).

Qian et al. recently used the Lin and Takahashi mouse models to analyze the effects of loss of *Dlk1-Meg3* imprinting on fetal liver HSCs (Qian et al., 2016). Using this approach, they reported impaired fetal liver HSC function in the IG-DMR (Lin et al., 2003) and *Meg3* (Takahashi et al., 2009) knockout mouse models. Functionally, they linked the observed phenotype to loss of expression of the miRNA cluster localized on the maternal allele. These miRNAs suppress the PI3K-mTOR pathway in healthy HSCs. Loss of this suppression led to activation of the PI3K-mTOR pathway and in turn increased OXPHOS and ROS production. ROS presumably induced HSC apoptosis, thereby causing the reported phenotype. These findings are significantly different to our findings. We neither observed impaired hematopoiesis

in adult HSCs nor during development. In the mouse models used by Qian et al., whole-body deletion of the locus takes place. Thus, all cells of the developing embryo are affected, and *Meg3* is known to be important during development (Lin et al., 2003; Takahashi et al., 2009; Zhou et al., 2010). Consequently, the observed reduced numbers of fetal liver HSCs and impaired liver HSC function could be secondary effects caused by defective HSC niche development. Importantly, Qian et al. showed impaired HSC development in the Lin and Takahashi model and linked it to loss of imprinting and loss of miRNA expression. Functional studies on the PI3K-mTOR and miRNA pathways were only performed using the Lin model. In the Takahashi model, however, imprinting is not affected, and expression patterns of the miRNA cluster in knockout animals are only reduced and not lost, especially in liver tissue (Takahashi et al., 2009). Thus, it seems reasonable that the functional impairment of fetal liver HSCs in the Takahashi model were caused by other mechanisms than those observed and described in the Lin model.

In our study, we use an inducible deletion approach, and loss of *Meg3* and the associated miRNA cluster do not impair adult hematopoiesis and HSCs, independently of the deletion time point. Using this strategy, we were able to directly analyze knockout effects on adult hematopoiesis, excluding potential developmental and microenvironmental impairments that could eventually affect HSCs. Adult HSCs are deeply quiescent, and *Meg3* is known to be involved in cell cycle regulation via the p53- and Rb-pathways (Passegue et al., 2005; Wilson et al., 2008; Zhang et al., 2010c; Zhou et al., 2007). Thus, we employed straight and competitive transplantation assays to challenge HSCs and to further investigate HSC function over time. We did not observe any biologically relevant impairment of *Meg3* KO HSCs under homeostatic or stress conditions.

As Qian et al. focused on fetal liver HSCs and we observed strikingly different results in the adult system, we aimed to additionally address the role of the *Meg3* locus during development. Of note, fetal liver HSCs greatly differ from adult HSCs, as they are actively cycling and expanding (Bowie et al., 2006; Nygren et al., 2006; Wilson et al., 2008). Thus, we deleted *Meg3* in CD45+ cells in the murine embryo by using *Vav1* as a Cre driver. Surprisingly, mice were viable, and we did not observe signs of hematopoietic impairments in adult mice. We conclude that *Meg3* is dispensable for embryonic establishment of hematopoiesis and HSC function after the endothelial-cell-to-HSC transition. We cannot exclude that *Meg3* might play an important role for this transition step during development, leading to the defects in fetal liver HSCs observed by Qian et al. (Qian et al., 2016).

4.1.2 Deletion of imprinted loci – targeting strategy matters

In our model, we target exon 1-4 of the *Meg3* locus, leading to loss of expression of both *Meg3* and the miRNA cluster. However, we do not target DMRs. Importantly, imprinting is not affected in our model (Zhou and Klibanski; unpublished observation). A broader disruption of the imprinted locus by targeting DMRs could potentially lead to globally altered gene expression patterns, which in turn could affect hematopoiesis – independently of *Meg3* and the miRNA cluster itself. It is possible that the effects observed by Qian et al. are related to such broader disruptions of gene regulation. Future studies will need to shed light on the effects of specific DMR targeting in hematopoietic cells, affecting the entire *Meg3-Dlk1* locus in adult mice. However, additional inducible mouse models are needed to answer these questions.

Further, special care has to be taken when interpreting results generated in knockout mouse models of imprinted regions. By altering imprinting or KO of imprinted regions, not only maternal genes are silenced or deleted, but paternal expression patterns are often reactivated on the maternal allele and vice versa. Thus, observed biological effects could be caused by loss of expression or increased expression of the maternal/paternal genes, respectively.

4.1.3 Loss of *Meg3* in the hematopoietic compartment does not cause development of leukemia

Meg3 has been reported to act as a tumor suppressor lncRNA in several cancer entities (Benetatos et al., 2011; Zhou et al., 2012). Interestingly, aberrant methylation patterns of DMRs in the *Meg3* locus were reported in MM, MDS and AML patients (Benetatos et al., 2008; Benetatos et al., 2010; Khoury et al., 2010). *Meg3* promoter methylation was associated with decreased overall survival (Benetatos et al., 2010) and increased *Dlk1* expression, which was further associated with silencing of the maternal expression patterns observed in the majority of AML patients (Khoury et al., 2010). Functionally*, Meg3* also has a tumor suppressor function in AML cells and was reported to act via a p53-dependent and -independent axis (Lyu et al., 2017). In this thesis, we show that loss of *Meg3* and the associated miRNA cluster alone does not cause the development of leukemia or myeloproliferative diseases even in long-term experiments such as tertiary transplantations. In addition, we did not observe any proliferative advantages of *Meg3* mat KO cells. We conclude that loss of the tumor suppressor *Meg3* alone is not sufficient for cell transformation in the hematopoietic system. Interestingly, loss of *Meg3* in cancer cells mostly occurs by epigenetic silencing (Benetatos et al., 2011; Zhou et al., 2012). Taken together, *Meg3* dysregulation might result as a secondary event during the development of leukemia but is unlikely to be an early event during leukemogenesis.

4.1.4 The role of the lncRNA *Meg3* in HSCs – summary

In this study, we show by inducible deletion of *Meg3* and the associated maternal miRNA cluster that this locus is dispensable for HSCs in the adult organism. We did not observe impaired hematopoiesis during homeostatic conditions, upon serial transplantation or in response to inflammation-induced stress. In addition, embryonic deletion of the locus after the endothelial-cell-to-HSC transition step did not affect hematopoiesis. These findings are surprising, as the lncRNA *Meg3* is highly and specifically expressed in the HSC compartment and strongly downregulated in early progenitors. Future studies are needed to understand why the locus is specifically expressed in the HSC compartment, even though it is functionally dispensable.

4.2 HSC function, differentiation and activation are regulated by APA

Different layers regulating HSC self-renewal, function and differentiation in health and disease have been discovered. These mechanisms include regulation of DNA methylation, chromatin accessibility and modifications, enhancer usage, gene expression, RNA editing, protein translation and stability as well as alternative splicing (Bahr et al., 2018; Cabezas-Wallscheid et al., 2014; Goldstein et al., 2017; Ji et al., 2010; Lipka et al., 2014; Luo et al., 2015; Signer et al., 2014; Zipeto et al., 2016). These regulatory layers act in a synergistic fashion to tightly control HSC fate and function. Perturbations in this precisely equilibrated system can lead to stem cell failure or cause development of hematological malignancies. In this thesis, we demonstrate that the RNA regulatory mechanism APA is dynamically regulated within the HSPC compartment and that correct APA is essential for proper HSC function.

4.2.1 3'-Seq of highly purified cell subsets using ultra-low input approaches

Multiple studies have used different approaches to analyze 3'-UTR length variations and occurrence of APA *in vivo* in health and disease, including EST databases, microarray approaches and 3'-enriched sequencing methods. Different developmental stages and mature tissues, including testis, muscle and brain tissue, as well as mature immune cells have been analyzed (Boutet et al., 2012; Jereb et al., 2018; Ji et al., 2009; Jia et al., 2017; Lau et al., 2010; Mueller et al., 2016; Singh et al., 2018). Most studies focused on the characterization of whole tissues, not taking cellular heterogeneity within organ structures into account. Alternatively, analysis of pre-defined APA events was performed in specific cell subsets (Boutet et al., 2012; Cesana et al., 2018; Mueller et al., 2016). Global APA analysis of cellular subsets *in vivo* is rare. Jereb et al., for example, investigated APA patterns in different neuronal cell types, including Purkinje and granule cells. They used specific mouse models that enabled pulldown of RNA from defined cell types (Jereb et al., 2018). In addition, a recent study reported 3'-Seq analysis of human *in vivo*-derived B cells (Singh et al., 2018). The so far unsolved problems of 3'-enriched sequencing analysis are caused by the prerequisites of the different techniques. Most published protocols and available kits depend on high amounts of RNA input, which can be achieved by using cell lines, whole tissues or frequent cell subsets. However, these techniques do not enable global analysis of sporadic cell subsets like stem and progenitor cells. Thus, the extent and functional relevance of APA in rare, but highly functional *in vivo* cell subsets, like stem cells, remains elusive.

Here, we were able to solve this problem. Sorting schemes and isolation of low amounts of high-quality total RNA from sorted HSC and MPP subsets were previously accomplished by our laboratory (Cabezas-Wallscheid et al., 2017; Cabezas-Wallscheid et al., 2014). We could use this approach to derive high-quality total RNA from homeostatic as well as activated HSCs

and MPP1-4 subsets. The QuantSeq 3'-mRNA Library Prep Kit REV (Lexogen) enables generation of libraries for 3'-Seq analysis using as little as 10 ng of total RNA. However, analysis of APA in HSCs and MPP subsets using 10 ng of total RNA is not feasible. Sorting of 10,000 HSCs requires pooling of 3 mice, and total RNA isolation leads to recovery of only 3 ng of total RNA. Thus, to enable time- and cost-effective APA analysis in *in vivo*-derived specific cell subsets, we further lowered the input needed for library generation by introducing an additional RNA denaturation step before random priming and reverse transcription. We think that this denaturation step leads to a higher efficiency of random priming and reverse transcription, enabling library generation from as little as 3 ng of total RNA (see also methods section). Thus, we developed a workflow that enables global analysis of 3'-UTRs in rare *in vivo* cell populations by combining widely used sorting approaches and adaptations of commercially available kits. Therefore, our approach can be broadly adapted to enable analysis of other cellular subsets in different *in vivo* systems. Here, we present the first global APA analysis of the rare *in vivo* cell subsets HSCs and MPPs.

The Rev version of the QuantSeq 3'-mRNA Library Prep Kit (Lexogen) is advantageous for APA analysis. In this approach, a custom sequencing primer is used to generate read1 during sequencing. This primer covers the residual part of the poly(A) tail, which is part of the sequencing libraries. Thereby, the read starts directly upstream of the cleavage site, facilitating detection of the utilized PAS 15 to 40 nt upstream. Using this technique, the employed PAS, the length of the 3'-UTR and the associated exon can be identified by annealing read 1 to the genome and performing bioinformatic analysis. This approach solves an issue typically occurring in 3'-UTR analysis using other sequencing methods: The base calling quality severely drops downstream of the T stretch generated during sequencing of the residual poly(A) tail (Chen et al., 2017; Wilkening et al., 2013). This is caused by polymerase slippage during library amplification, clustering or sequencing. Thus, reads within one cluster become desynchronized, and no clear fluorescent signal is detectable and the sequence cannot be identified. Other methods for 3'-enriched sequencing try to solve this issue by using so-called RNA manipulation-based approaches for library generation. In these methods, for example, splint digestion and digestion of the residual poly(A) tail (3P-Seq) or RNA fragmentation and capturing are used (3'READS) (Hoque et al., 2013; Jan et al., 2011). However, these methods are not yet feasible for analysis of ultra-low RNA inputs.

One drawback of our sequencing approach compared to 3P-Seq and 3'READS is the internal priming issue. Internal priming describes a process occurring during oligo(dT)-mediated reverse transcription. The oligo(dT) primer does not only anneal to the poly(A) tails but also to internal A stretches present in RNA molecules. During library generation, this leads to library molecules that seem to represent cleavage and polyadenylation events but, in reality, only represent genomic A stretches. In our sequencing approach, we deal with the internal priming

issue by using very strict bioinformatic filtering steps prior to APA analysis. We only include reads in our analysis that lie downstream of a conserved PAS and that do not harbor genomic poly(A) stretches downstream of the read start site. Thereby, we ensure that only real cleavage and polyadenylation events are considered in our APA analysis.

Recently, many groups have focused on the re-analysis of RNA-seq datasets for APA analysis (Ha et al., 2018; Xia et al., 2014; Ye et al., 2018). These approaches are of high interest as they allow analysis of APA without the need to generate new sequencing data. However, it has to be kept in mind that these tools depend on bioinformatic and statistical reconstruction of PAS usage. Cleavage site identification is less precise, and *de novo* PAS discovery is limited. In addition, only a fraction of reads is located in 3'-UTR regions. Thus, RNA-seq data with a higher sequencing depth is needed to enable APA analysis as compared to 3'-enriched sequencing methods. Re-analysis of RNA-seq data is a fast approach to identify APA events that occur frequently in the respective setting. However, global in-depth analysis and *de novo* identification of APA events remains dependent on 3'-enriched sequencing methodologies.

Lately, also single-cell approaches have been used for APA analysis (Velten et al., 2015; Ye et al., 2019). Library generation and sequencing of single cells frequently uses methods relying on 3'-end methods. Therefore, APA analysis and analysis of cell heterogeneity is possible using these datasets. However, sequencing depth in single cell approaches is limited (Liu and Trapnell, 2016). Hence, many low to medium expressed genes cannot be detected, and APA analysis is restricted to highly expressed genes. While highly expressed genes are useful to identify intercellular heterogeneity and cellular trajectories, APA often occurs in very specific non-housekeeping genes. Thus, for a general understanding of the prevalence and relevance of APA, bulk sequencing approaches are better suited. Single-cell approaches can be used in addition to understand to which extent APA contributes to intercellular heterogeneity.

In summary, we provide a novel 3'-Seq resource dataset enabling APA analysis in the HSPC compartment *in vivo*. The bioinformatic analysis workflow enables identification of high-confidence PAS usage patterns and allows analysis of differential APA on a single gene level and analysis of global 3'-UTR length regulation. Our library generation and bioinformatic analysis approach is applicable to other biological questions and model systems. The dataset allows us for the first time to analyze APA quantitatively and investigate the role of this RNA regulatory mechanism in stem cell hierarchies.

Previously, the role of APA in stem cells has mainly been studied in ESCs and in iPSC generation (Brumbaugh et al., 2018; Ji and Tian, 2009; Lackford et al., 2014). *In vivo* analysis was focused on germ cells and specific genes exhibiting APA in muscle stem cells and HSCs (Cesana et al., 2018; Liu et al., 2007; Mueller et al., 2016). Using our adapted sequencing

approach, we can now globally analyze APA within stem cell hierarchies. Our data provides evidence that regulation of 3'-UTR length and APA represent a novel hub to target and control stem cell function *in vivo*.

4.2.2 APA patterns identified in HSC/MPP populations are versatile

Using our 3'-Seq approach, we could identify differential APA upon HSC differentiation and activation on a single gene level. We observed differential CR-APA as well as UTR-APA in MPP1-4 cells and pHSCs compared to HSCs and sHSCs, respectively. In fact, a mixture of CR- and UTR-APA was detected for a given gene in most cases. These results indicate that APA patterns are highly versatile and that CR- and UTR-APA are tightly connected. Complex APA regulation patterns have been reported: APA is regulated by intrinsic PAS strength (Hu et al., 2005; Tian et al., 2005), the presence of 3'-processing factors (Lackford et al., 2014; Li et al., 2015; Takagaki and Manley, 1998; Takagaki et al., 1996; Xia et al., 2014; Yao et al., 2013), epigenetic marks (Spies et al., 2009; Wood et al., 2008), transcription regulators and transcription itself (Ji et al., 2011; Nagaike et al., 2011; Pinto et al., 2011; Rosonina et al., 2003) and other regulatory layers. All of these different regulatory mechanisms can be expected to work synergistically and antagonistically at any given time point. Thus, in a single cell, not only one PAS is used per actively transcribed gene. The different mechanisms controlling PAS selection slightly vary for every transcript, even if transcribed from the same gene with only minutes between start of transcription initiation. Thus, an equilibrium of PAS usage for a distinct gene is reached in a defined cellular status, e.g. in HSCs. In this status, the different PASs are used in a certain ratio, as we could observe in HSC analysis. Changes in the cellular status, in our case for example differentiation towards MPP2, do not lead to complete loss of usage of one PAS but instead to a more complex gradual shift towards usage of other PASs. PAS usage is increased at some sites, whereas a pronounced decrease in usage is observed at other sites. In addition, usage of some PASs for the same gene remain unchanged. Our data indicates that complex processes like HSC activation and differentiation are accompanied by distinct changes in APA patterns that are characterized by versatile mixtures of CR- and UTR-APA affecting PASs in different genomic locations, including introns, exons and annotated 3'-UTRs.

Interestingly, we observed enhanced intronic polyadenylation and usage of higher numbers of PASs per gene in HSCs compared to MPP cells and pHSCs. Of note, *Pabpn1* KD cells, which harbor decreased HSC function, also exhibited a decrease in intronic polyadenylation and PAS number per gene. Thus, intronic polyadenylation patterns and high PAS numbers per gene seem to be characteristic for functional quiescent HSCs. This APA status might be indicative of the highest HSC potency. However, we do not yet understand why intronic polyadenylation

and PAS selection in general are increased in the HSC compartment. Previously, intronic polyadenylation has been shown to mediate IgM switching in activated B cells from the membrane-bound to the secreted version by exclusion of the transmembrane domain (Takagaki and Manley, 1998; Takagaki et al., 1996). Thus, intronic polyadenylation in HSCs could be important to regulate localization and secretion of certain proteins. Further, specific intronic polyadenylation events have been described in resident muscle stem cells for the gene *Pdgfr*α (Mueller et al., 2016), and immune cells have been shown to exhibit globally enhanced intronic polyadenylation (Singh et al., 2018). Hence, intronic polyadenylation seems to occur in different biological settings, but additional research is needed to understand the relevance of these events. Singh et al. could already show that mRNAs generated by intronic polyadenylation are not subject to nonsense-mediated decay, as they harbor stop codons and standard 3'-UTR structures (Singh et al., 2018). In our analyses, we also observe intronic polyadenylation at sites harboring common 3'-UTR motifs, including the PAS, hinting at the generation of functional proteins.

It has to be considered that in-depth analysis of 3'-UTR regions has only recently become available. Therefore, the annotation of all expressed 3'-UTRs is not yet complete. We think that intronic polyadenylation observed in many studies should rather be considered as CR-APA. The presence of the typical 3'-UTR sequences, including the PAS and other motifs, as well as the presence of stop codons in the last exon of transcripts generated by most intronic polyadenylation events should help to differentiate between "CR-APA intronic polyadenylation" and "real intronic polyadenylation", which in turn leads to the generation of truncated non-functional proteins. Allocation to intronic structures might be caused by incomplete 3'-UTR annotation. "CR-APA intronic polyadenylation" seems to be a mechanism to modulate secretion of specific proteins and is specifically regulated during cellular differentiation processes (Singh et al., 2018). These reports are in line with our observations. Conclusively, intronic polyadenylation and enhanced PAS usage are characteristic for HSCs and become downregulated in response to HSC activation and differentiation. The functional relevance of this pattern and the effects on the proteome and individual protein function remain to be determined.

4.2.3 ΔAPAS reflect molecular distances between HSCs and MPP populations

Previous reports showed that differential gene expression analysis can be used to identify molecular differences between HSCs and MPP populations (Cabezas-Wallscheid et al., 2017; Cabezas-Wallscheid et al., 2014). In addition, cellular trajectories and molecular distances between these populations could be identified. We could validate these findings by performing differential gene expression analysis using our 3'-Seq dataset and by performing PCA. Only

few differences were identified between HSCs and MPP1 cells by differential gene expression analysis, while MPP2, MPP3 and MPP4 cells were more distinct compared to HSCs on a gene expression level. These differences also reflect functional differences between HSCs and the different MPP populations (Cabezas-Wallscheid et al., 2014; Pietras et al., 2015). Only HSCs are capable of long-term multilineage engraftment, while MPP1 cells exhibit short-term multilineage engraftment. Engraftment potential of MPP2 cells is further reduced, but remains multilineage. In contrast, MPP3 and MPP4 cells exhibit lineage-biased engraftment (Cabezas-Wallscheid et al., 2014; Pietras et al., 2015). Interestingly, ΔAPAS analysis reflected the earlier observations made by us and others in DEG analysis. Only few ΔAPAS hits were identified between HSCs and MPP1 cells, reflecting their high molecular identity. However, a greater number of ΔAPAS was identified in HSCs compared to MPP2-4 cells. These cells also exhibited greater differences in PCA and DEG analysis. Thus, the number of ΔAPAS between defined cellular populations can be used to estimate their molecular and functional distance. These differences might be caused mainly by the cell cycle status, as HSCs are highly quiescent and MPP cells are more active (see also 4.2.4).

4.2.4 APA patterns and 3'-UTR length in HSCs depend on the activation status

HSC differentiation towards progenitors as well as inflammation-mediated HSC activation rely on increased HSC proliferation (Essers et al., 2009; Wilson et al., 2008). Analysis of our 3'-Seq dataset of HSCs and MPPs revealed that HSC activation and exit of G0 is accompanied by APA switches and overall 3'-UTR shortening.

Interestingly, 3'-UTR shortening has been observed in a number of biological settings including immune cell activation (Gruber et al., 2014b; Jia et al., 2017; Pai et al., 2016; Sandberg et al., 2008), cellular transformation (Lee et al., 2018; Xia et al., 2014; Xiang et al., 2018) and during cellular stress response (Hollerer et al., 2016; Zheng et al., 2018). All of these cellular settings are accompanied by increased cell cycle progression. In line, we observed 3'-UTR shortening in activated HSCs. This shortening response was induced by cellular activity of HSCs, independently of the cause of activation, in our case differentiation or inflammation-induced stress response. 3'-UTR shortening seems to be an intrinsic cellular response observed upon increased cellular activity.

Especially during the activation of HSCs and exit of the quiescent state, the demand for functional proteins is high, as overall protein content in quiescent HSCs is extremely low. Interestingly, Signer et al. showed that an increase in protein synthesis in progenitors as well as upon HSC activation is critical for proper HSC function (Signer et al., 2014). Furthermore, we recently identified a biosynthetic shutdown as a hallmark of HSC dormancy (Cabezas-Wallscheid et al., 2017). Based on our results, HSCs might evade miRNA targeting of mRNAs

by global shortening of 3'-UTRs, potentially leading to an increase in overall protein production during differentiation and stress-induced activation processes. This increase in protein content is a prerequisite for HSC function and activation and might be influenced by APA. 3'-UTR shortening has so far mainly been analyzed in cell lines and was linked to evasion of miRNA-based regulation and increased protein output (Fu et al., 2018; Ji et al., 2011; Jia et al., 2017; Pai et al., 2016; Sandberg et al., 2008; Zheng et al., 2018); but also unchanged protein expression was observed (Gruber et al., 2014b). Studies performed by Fu et al. could partially resolve these contradictory results (Fu et al., 2018). Enhanced protein expression caused by 3'-UTR shortening seems to be characteristic for actively cycling cells, especially in the G2/M phase. These observations support our hypothesis that HSCs entering the cell cycle depend on APA-mediated 3'-UTR shortening to generate protein output, which is needed for HSC function during activation (Signer et al., 2014). In the future, the link between APA and changes in protein expression patterns in HSPCs should be further investigated. As HSPCs are extremely rare and protein content in general is low, standard analysis of protein expression of APA hits, e.g. by western blot analysis, is not feasible. Proteome analysis can be performed, however the number of detectable proteins is limited due to the low input. Thus, only intracellular flow cytometry analysis and immunofluorescence approaches are practicable to analyze protein expression, although technically challenging for a global identification approach.

How these APA adaptations are achieved in proliferating HSCs remains to be determined. In the past, mTOR signaling and E2F, two common mediators of cellular activity, have been shown to regulate APA in response to cellular activation for example by increasing levels of 3'-processing factors (Chang et al., 2015; Elkon et al., 2012). Both factors have also been shown to be important for HSC function and to mediate HSC activation, making them potential regulators of APA during HSC activation (Gan and DePinho, 2009; Julian and Blais, 2015; Kalaitzidis et al., 2012; Kim et al., 2017). Further, mTOR-mediated 3'-UTR shortening has been reported to augment protein synthesis by enhancing polysome formation (Chang et al., 2015).

We previously observed expression of many biosynthetic processes to be upregulated in HSCs upon commitment and exit of the dormant state, including transcription, mRNA processing and translation (Cabezas-Wallscheid et al., 2017). Increased levels of 3'-processing factors (Lackford et al., 2014; Li et al., 2015; Takagaki and Manley, 1998; Takagaki et al., 1996; Xia et al., 2014; Yao et al., 2013) and enhanced transcription (Ji et al., 2011; Pinto et al., 2011) have previously been linked to APA regulation. Thus, enhanced mRNA processing and transcription could be additional mechanisms regulating APA upon HSC activation. The exact mechanisms that regulate the observed changes in APA patterns upon HSC differentiation and activation remain to be elucidated.

4.2.5 PABPN1 regulates APA in the hematopoietic system and is required for HSC function

In our study, we demonstrate that the APA regulator PABPN1 is essential for HSC self-renewal and multipotency *in vitro* and *in vivo*. Initially, PABPN1 was described to be important for efficient poly(A) tail generation and regulation of poly(A) tail length (Kerwitz et al., 2003; Kuhn et al., 2009; Wahle, 1991, 1995). However, multiple recent studies could show that PABPN1 additionally acts as an important regulator of APA (Banerjee et al., 2017; de Klerk et al., 2012; Jenal et al., 2012; Riaz et al., 2016). First, Jenal et al. identified PABPN1 as an APA regulator in U2OS cells and reported 3'-UTR shortening upon *Pabpn1* KD (Jenal et al., 2012). In addition, they and others showed that mutated *Pabpn1* causes the development of the myopathic disease OPMD by deregulated APA (de Klerk et al., 2012; Jenal et al., 2012). In recent years, these initial observations were extended by studies focusing on deregulated *Pabpn1* and APA in muscle cells and during muscle wasting (Banerjee et al., 2017; Riaz et al., 2016). Our data extends these findings by showing that PABPN1 also acts as an important APA regulator in an independent, unperturbed stem cell hierarchy *in vivo*. We show that *Pabpn1* is essential for HSC function and we observed deregulated APA patterns upon *Pabpn1* KD. Of note, APA profiles, including PAS localization and PAS number per gene, were less HSC-like in *Pabpn1* KD cells, hinting that HSC-characteristic APA patterns are partially lost upon *Pabpn1* KD.

Most studies reported shortening of 3'-UTRs in response to loss of *Pabpn1* (de Klerk et al., 2012; Jenal et al., 2012). We, however, did not observe global changes in 3'-UTR length, but a decrease in the total number of detected PASs. Interestingly, Banerjee et al. reported 3'-UTR lengthening upon *Pabpn1* KD for some of their main targets (Banerjee et al., 2017). Thus, the influence of PABPN1 on global 3'-UTR length appears to depend on the cell type and the cellular status. Importantly, de Klerk et al. reported dose-dependent effects of *Pabpn1* KD on 3'-UTR length (de Klerk et al., 2012). KD of *Papbn1* that reduced protein levels by 40% did not consistently change 3'-UTR length, while a KD of 70% led to the reported 3'-UTR shortening effect. In our setting, we observed an approximately 50% KD of *Pabpn1*, raising the possibility that a stronger KD might induce global 3'-UTR shortening. In addition, we performed 3'-Seq analysis of KD cells 72 h after viral transduction. This approach was necessary to derive enough material for analysis. However, it would be interesting to analyze APA patterns in *in vivo*-derived *Pabpn1* KD HSCs. As these HSCs, however, harbor cellular defects, we were unable to derive enough cells to perform 3'-Seq analysis. While we observed deregulated APA in hematopoietic *Pabpn1* KD cells, technical challenges prevented us from performing encompassing 3'-UTR length analysis, e.g. in HSC populations.

Intriguingly, *Pabpn1* expression levels and associated APA patterns have been shown to act as prognostic factors in lung cancer (Ichinose et al., 2014). Additionally, global analysis of APA patterns in published datasets and cell lines showed that 3'-UTR shortening is a hallmark of cancer cells (Xiang et al., 2018). In this study, PABPN1 was identified as the master regulator of APA disruption in transformed cellular states. Thus, future studies should focus on the role of PABPN1 as a regulator of APA in hematological malignancies. This research is of major interest, as highly potent small molecules that modulate APA by targeting PABPN1 have recently been identified (Araki et al., 2018). These inhibitors lead to global 3'-UTR shortening, mimicking *Pabpn1* KD phenotypes. In the future, restoring tissue-specific APA patterns might be a novel therapeutic approach for cancer therapy.

4.2.6 *Ctsc* and *Treml2* are regulated by APA in HSPCs

We identified individual genes that dynamically undergo APA during HSC differentiation and pIC-mediated activation. For example, the papain family protease-encoding gene *Ctsc* exhibited significant ΔAPAS in MPPs compared to HSCs. In our analysis, we observed that a shortened, presumably enzymatically non-functional (Turk et al., 2001), CTSC version is expressed in HSCs during homeostasis. Expression of the functional full-length version is enhanced upon HSC differentiation and pIC- as well as LPS-induced activation. Interestingly, CTSC activates the proteases Granzyme A and B, and Granzyme B has been shown to play an important role in the LPS stress response in HSCs (Carnevalli et al., 2014; Pham and Ley, 1999). In addition, LPS has been shown to enhance CTSC secretion (Hamon et al., 2016). We speculate that pIC- and LPS-induced activation of HSCs leads to APA-mediated switching to the full-length *Ctsc* version. In the following, CTSC might be secreted or act intracellularly and in turn activates Granzyme A and B. Activation and secretion of granzymes in response to LPS has been reported in HSCs (Carnevalli et al., 2014). It is hypothesized that Granzymes modify the stem cell niche to allow HSC mobilization or induce cell death. In summary, we identified an APA switch in *Ctsc*, potentially coordinating Granzyme activation within the HSC/MPP compartment upon activation. Additional research is needed to functionally verify this hypothesis.

In addition, UTR-APA affecting the gene *Treml2* was identified. APA in *Treml2* led to a decrease in 3'-UTR length of more than 1000 nt in MPP2 cells and pHSCs. Interestingly, TREML2 upregulation and re-localization to the surface in response to inflammation has been reported in myeloid cells (King et al., 2006; Thomas et al., 2016). Changes in surface expression upon pIC-induced inflammatory activation and cellular commitment might occur in HSCs. This process might be regulated by APA and could thereby regulate TREML2 function in HSCs. It has been shown that modulation of 3'-UTR length alone, namely UTR-APA, alters CD47 localization (Berkovits and Mayr, 2015). The shorter 3'-UTR version mediates retention

of the CD47 protein in the ER, while the longer 3'-UTR version mediates transport to the cell membrane. This is achieved by interaction of the longer 3'-UTR with HuR and Set, which in turn mediate CD47 protein localization. A similar mechanism could regulate TREML2 localization. However, further studies are needed to functionally understand the role of ΔAPAS in *Treml2* and the general function of TREML2 in hematopoietic biology.

4.2.7 NUDT21-mediated metabolic reprogramming towards glutaminolysis is essential for HSC activation and differentiation

Another highly interesting finding of this thesis is that the metabolic enzyme GLS exhibits isoform switching by CR-APA upon HSC activation and differentiation and that the process *glutamine metabolism* is enriched in genes exhibiting 3'-UTR shortening. GLS is the first enzyme in glutaminolysis, generating glutamate from glutamine (Altman et al., 2016). Glutaminolysis provides building blocks for amino acid, nucleotide and lipid biosynthesis. In addition, it can be used for glutathione production and as a fuel for the TCA cycle by conversion to αKG (Altman et al., 2016).

Two different GLS enzymes are known: The kidney-type glutaminase (GLS) and the liver-type glutaminase (GLS2) (Aledo et al., 2000; Elgadi et al., 1999). Expression patterns of these two enzymes are highly tissue-specific (Aledo et al., 2000; Elgadi et al., 1999). In HSPCs, we observe high expression of *Gls*, while *Gls2* is not robustly expressed (Cabezas-Wallscheid et al., 2014). The *Gls* gene encodes two different protein isoforms, GAC and KGA. In the study at hand, we observed a *Gls* isoform switch favoring expression of GAC in HSCs exiting G0 or upon differentiation. This switch was accompanied by an overall increase in GLS protein levels. While both GLS isoforms are enzymatically functional, KGA contains additional ankyrin repeats at the C-terminus compared to GAC. Cassago and colleagues could show that the GLS isoform GAC, but not KGA, is found in mitochondria (Cassago et al., 2012). In addition, they could show that GAC harbors superior enzymatic kinetics for glutamine to glutamate conversion in presence of the activator phosphate compared to KGA and GLS2. Due to its localization, GAC seems to convert glutamine to glutamate preferably for mitochondrial processes like the TCA cycle. In contrast, KGA, and also GLS2, seem to be more active in the cytosol or even the nucleus (Cassago et al., 2012; Olalla et al., 2002). KGA harbors ankyrin repeats at the C-terminus enabling additional protein-protein interactions. Thus, KGA could be involved in process that are independent of glutaminolysis (Cassago et al., 2012). Transferred to our observations, this indicates that activation of HSCs is accompanied by a relative increase in GAC for metabolic adaptations, allowing increased utilization of the TCA cycle and in turn OXPHOS or glutathione production.

Different cancer entities, including leukemia, have been reported to be addicted to glutaminolysis (Gallipoli et al., 2018; Jacque et al., 2015; McBrayer et al., 2018). Interestingly, this metabolic addiction has been linked to isoform switching of *Gls* towards GAC and has been shown to be essential for cancer cell growth (Jacque et al., 2015; van den Heuvel et al., 2012; Xia et al., 2014). In our study, we found this metabolic adjustment initially identified in cancer cells to be a physiological property of untransformed HSCs in response to cellular activation. Malignant cells as well as stem cells seem to rely on glutaminolysis during enhanced cellular proliferation. Thus, cancer cells might hijack this metabolic switch providing pro-proliferative metabolic conditions usually present upon stem cell activation.

Remarkably, the oncogenic transcription factor c-Myc has been shown to be involved in regulating *Gls* expression in cancer cells (Gao et al., 2009). c-Myc transcriptionally represses miR-23a/b, which in turn leads to upregulation of *Gls*, a target of miR-23a/b. Thereby, c-Myc seems to enhance glutaminolysis in tumor cells (Gao et al., 2009). However, additional studies could show that some cancer entities rather depend on the GAC isoform, which is not targeted by miR-23a/b (Xia et al., 2014). Thus, the *Gls* isoform switch towards GAC enhances glutaminolysis and uncouples it from c-Myc. Upon HSC activation and commitment, c-Myc protein levels are known to increase, mediating cellular activation (Ehninger et al., 2014). In addition, c-Myc expression is low in dormant HSCs, and Myc-depleted ESCs are characterized by a quiescent state and low biosynthesis (Cabezas-Wallscheid et al., 2017; Scognamiglio et al., 2016). Interestingly, analysis of miRNA-seq data generated by us showed that miR-23a/b is significantly upregulated in MPP2 cells (data not shown). Thus, we consider miR-23a/b to have important roles during HSC activation and commitment. In HSCs and MPPs, enhanced c-Myc levels do not correlate with repression of miR23a/b. Therefore, HSCs seem to rely on the alternative pathway, namely switching to the more active GAC isoform, to increase glutaminolysis during stress response.

In response to stress-induced HSC activation, we observe a switch to the GAC isoform of *Gls* and enhanced total GLS protein levels. However, it remains elusive which metabolic pathways are supported by these metabolic adaptations, as glutamine and glutamate contribute to many metabolic routes (Altman et al., 2016). After conversion of glutamine to glutamate, glutamate can be used to generate glutathione, which is important to neutralize ROS. Alternatively, glutamate dehydrogenase or aminotransferases use glutamate for αKG production. αKG is used for energy generation via the TCA cycle and OXPHOS. Glutamate, glutamine and αKG are source materials to generate nucleotides, amino acids or lipids. Of note, αKG is also a very important co-substrate for many dioxygenases involved in, for example, protein hydroxylation and epigenetic modifications, e.g. Tet1/2 enzymes involved in DNA demethylation (Zdzisinska

et al., 2017). Thus, these processes could also be affected by enhanced glutaminolysis. Quiescent HSCs are biosynthetically silenced, exhibiting low cellular activity and protein translation (Cabezas-Wallscheid et al., 2017; Signer et al., 2014; Wilson et al., 2008). Thus, when HSCs become activated, these cells are in need of biomass, including amino acids and lipids. At the same time, an increased energy demand has to be met. By metabolizing glutamine to glutamate, HSCs are able to meet all of these metabolic needs, as glutamate can enter all of these biosynthetic fates. In addition, HSCs need to control ROS levels during increased cellular activity to limit DNA damage (Flach et al., 2014; Walter et al., 2015). Thus, enhanced glutathione levels are beneficial. Glutamate also provides the means to increase glutathione production. In summary, we consider increased glutaminolysis to be highly beneficial for HSCs undergoing cellular activation as it facilitates increased energy production and generation of biomass, while at the same time preventing DNA damage by producing ROS scavengers. In line, activated HSCs have been reported to switch to OXPHOS (Ito and Ito, 2018; Ito and Suda, 2014; Yu et al., 2013), which can be driven by glutaminolysis. Additional experiments are needed to understand which pathways are used to metabolize glutamate in activated HSCs.

Standard cultivation protocols for HSCs induce differentiation and loss of stem cell potential. In this thesis, I could show that short-term prevention of glutaminolysis by GLS inhibition leads to increased HSC maintenance in culture by limiting the number of HSCs exiting G0. *In vivo*, HSCs are maintained in a quiescent state, limiting for example DNA damage (Flach et al., 2014; Walter et al., 2015). HSCs seem to lose stem cell potential with every round of division, presumably to protect individuals from blood contribution of HSCs harboring DNA mutations that could eventually develop into leukemia (Ito and Suda, 2014). *In vitro*, this intrinsic HSC “clock” limits stem cell maintenance. By inhibiting GLS, we observed less HSCs exiting G0 and in turn increased stem cell maintenance as shown by CFU assay. Thus, HSC potential is indeed maintained if proliferation is prevented. In addition, these experiments show that efficient HSC proliferation is dependent on enhanced glutaminolysis as indicated by our *in vivo* discoveries. *In vivo*, glutaminolysis is enhanced by APA upon enforced HSC proliferation by pIC. Collectively, enhanced glutaminolysis is a prerequisite for HSC proliferation and activation, and *in vitro* inhibition of this pathway enhances HSC maintenance, presumably by limiting HSC cell division.

APA-mediated *Gls* isoform switching from the KGA to the GAC isoform is controlled by NUDT21 (Masamha et al., 2016). *Nudt21* KD has been shown to enhance relative KGA expression (Masamha et al., 2016). As we observed enhanced relative GAC expression upon HSC activation, we wondered if this switch is also mediated by and dependent on NUDT21. Strikingly, *Nudt21* KD limited the APA-mediated switch to the GAC isoform, leading to a

diminished increase in GLS protein levels in activating culture conditions compared to control cells. Therefore, NUDT21 is essential to mediate GAC/KGA isoform switching in HSPCs during stress response. I could additionally show that HSPC function is impaired when preventing the isoform switch. We observed a differentiation block in *Nudt21* KD cells, presumably caused by inhibition of metabolic adaptations. Without the biomass and energy provided by enhanced glutaminolysis, HSPCs seem to be incapable of undergoing differentiation. Intriguingly, and in line with our findings, Brumbaugh et al. could show that *Nudt21* KD in ESCs and myeloid cells impairs differentiation (Brumbaugh et al., 2018). In this study, the authors linked the observed phenotype to changes in APA patterns that lead to enhanced protein levels of chromatin modifiers. These chromatin modifiers in turn maintain stem cell-like chromatin states, and NUDT21 was reported to influence APA patterns of multiple additional genes (Brumbaugh et al., 2018; Masamha et al., 2014). In our study, we did not focus on the broader influence of *Nudt21* KD on APA patterns other than *Gls*. Thus, we cannot exclude that the observed inhibition of differentiation is caused by other effects of NUDT21 depletion or an interplay of different processes. Future experiments will focus on the detailed role of GAC and KGA in mediating HSC stemness and activation, for example by specific targeting of the different isoforms. Overall, we present data indicating that APA-mediated *Gls* isoform switching is important for metabolic reprogramming of HSCs and required for HSC activation.

Although the importance and prevalence of glutaminolysis in HSCs has long been discussed, functional studies are lacking (Ito and Suda, 2014). Previously, it has been shown that glutamine metabolism is essential for *de novo* nucleotide synthesis during erythroid lineage specification (Oburoglu et al., 2014). In addition, GLS has been reported to be important for T cell differentiation and proliferation, promoting Th17 while restricting Th1 cells (Johnson et al., 2018). Only recently, it was reported that skeletal stem cells depend on GLS and glutamine metabolism to regulate osteoblast and adipocyte differentiation and to enable bone formation (Yu et al., 2019). In this setting, glutaminolysis was critical for αKG production, and αKG in turn was essential for the proliferation, specification and differentiation of skeletal stem cells. Taken together, our results and the reported studies indicate that hematopoietic cells and stem cells depend on GLS and glutamine metabolism for differentiation and stem cell proliferation. We extend these findings by showing that HSCs residing at the top of the hematopoietic hierarchy depend on GLS for stem cell proliferation and differentiation as well. We further link these metabolic adaptations to changes in APA. Future studies will investigate the related metabolic pathways in depth.

4.2.8 HSC function, differentiation and activation are regulated by APA – summary

In this study, we show that the process of APA and 3'-UTR choice is tightly regulated and functionally important for hematopoiesis. We demonstrate that the APA regulator PABPN1 is essential for HSC function *in vitro* and *in vivo*. Furthermore, the advancement in ultra-low input 3'-Seq techniques enabled us to analyze APA globally in rare BM populations and to generate a 3'-UTR/APA resource dataset of all genes expressed in HSPCs. Using this novel resource, we report APA upon HSC commitment and activation. We observed 3'-UTR shortening upon stem cell differentiation and upon exit of G0. In addition, we also demonstrate the functional importance of APA by reporting an APA switch in the metabolic gene *Gls* to be essential for HSC activation.

4.3 Generation of a *Pabpn1*^{flox} mouse model

Using the *Easi*-CRISPR technique (Miura et al., 2018; Quadros et al., 2017), we generated a mouse model in which exon 3 and 4 of *Pabpn1* are floxed. Using this approach, efficiencies of 8.5-100% have been reported for generation of transgenic mice. In comparison to other CRISPR/Cas9 techniques for the generation of transgenic animals, *Easi*-CRISPR is more efficient (Miura et al., 2018). In most published techniques, dsDNA donor templates are used, leading to a reduction in targeting efficiency (Aida et al., 2016; Raveux et al., 2017). However, these methods are better suited for integrating large inserts, as *Easi*-CRISPR can only be used for inserts up to 1.5 kb (Miura et al., 2018; Quadros et al., 2017). In our initial approach, we tried to generate transgenic mice using a dsDNA plasmid as the template. However, no transgenic mice were obtained. By using *Easi*-CRISPR, 3 of 17 animals were correctly targeted (~18%). Thus, in our hands, *Easi*-CRIPSR was a quick and reliable approach to generate floxed transgenic animals.

A common issue with CRISPR-Cas9 targeting are off-target effects (Sander and Joung, 2014). To increase specificity, we chose gRNAs that are predicted to have a low number of off-targets. gRNA A3 was predicted to have 90 potential off-targets, all presenting at least 3 mismatches compared to the target site. gRNA B2 presented 100 anticipated off-targets, all with at least 2 mismatches. Due to the high number of mismatches, the likelihood of off-target mutations is low. In any case, as homology regions are not found in other parts of the genome, an off-target event would induce a mutation but not lead to loxP site integration. In addition, all of the predicted off-targets were localized on other chromosomes than *Pabpn1*, which is localized on chromosome 14. Thus, mutations caused by off-target effects would be lost during further crossing of the generated mice. We therefore did not systematically screen for off-target events. In summary, the likelihood of off-targets is low due to gRNA choice and backcrossing of mice.

Using one founder animal, we generated homozygous *Pabpn1*^{flox} mice. We are now crossing these mice to Cre drivers, including *Vav1* and *Mx1*, to first check if our targeting strategy leads to *Pabpn1* KO upon Cre induction. After the functional validation, we will apply our transgenic mouse model to answer different biological questions. For example, it could be used to analyze the effects of complete loss of *Pabpn1* on the hematopoietic system. This would also open up the possibility to perform additional 3'-Seq experiments and APA analysis of *Pabpn1* KO cells, including HSCs, *in vivo*. Our mouse model can of course also be used in additional experimental settings to analyze the role of *Pabpn1* and APA in different tissues and disease settings. For example, the mice can be used to study the role of *Pabpn1* and deregulated APA

in cancer. *Pabpn1* is downregulated in cancer cells and associated with globally deregulated APA (Ichinose et al., 2014; Xiang et al., 2018). Thus, we aim to use the generated mouse model in the future to generate leukemic mice, for example by employing the MLL-AF9 oncogene by lentiviral transduction. After initiation of leukemia, *Pabpn1* could be deleted by Cre induction, and the effects on disease progression and APA patterns could be analyzed.

4.4 Outlook

By combining transgenic knockout and conditional knockout mouse models, and novel sequencing approaches, I could show that regulatory RNAs like lncRNAs and RNA regulation itself, such as APA, are important players in HSC biology. In addition, the generation of a novel mouse model will enable further research on the role of APA in HSCs and other cellular systems.

I could show that *Meg3* and the associated miRNA cluster are dispensable for adult HSCs and during HSC development after the endothelial-cell-to-HSC transition by performing analysis of MxCre *Meg3* $^{\text{mat flox/pat wt}}$ and VavCre *Meg3* $^{\text{mat flox/pat wt}}$ mice. Additional experiments could be performed to better understand the role of the *Meg3* locus during embryonic development of the hematopoietic system. To this end, *Meg3* $^{\text{mat flox/pat flox}}$ mice could be crossed to additional Cre drivers, for example VEC-Cre. VEC-Cre enables analysis of the effects of gene deletion on HSC formation from the vasculature including the dorsal aorta, vitelline and umbilical arteries and the placenta (Chen et al., 2009). In addition, broader disruptions of the *Dlk1-Meg3* locus by targeting DMRs could be used to investigate the effects on hematopoiesis. This approach, which would not only affect *Meg3* and the miRNA cluster but also imprinting itself, could help to understand the different observations made by Qian et al. (Qian et al., 2016) and us. In addition, we could use ROSA26 Cre mice for global *Meg3* deletion in the *Meg3* $^{\text{mat flox/pat flox}}$ mice and check if we observe similar effects as described in the Qian et al. paper. All of these experiments could help to unravel why *Meg3* is so highly and specifically expressed in HSCs. Furthermore, to better understand the role of *Meg3* in leukemia, MxCre *Meg3* $^{\text{mat flox/pat wt}}$ LSK cells could be transduced with the MLL-AF9 or MLL-ENL oncogene. Subsequent transplantation of transduced cells and Cre-mediated deletion could be performed to analyze the impact of *Meg3* loss on leukemia development and maintenance. As *Meg3* functions as a tumor suppressor, we would expect faster disease progression.

I could show for the first time that APA is differentially regulated upon HSC differentiation and activation *in vivo*. In addition, I could show that APA is also important for HSC function in mice. Future studies are needed to understand exactly how APA is regulated in HSCs and MPP cells and upon inflammation-induced HSC activation. Different experimental setups are needed to understand if an increase in 3'-processing factors, enhanced transcription, epigenetic modifications or signaling pathways mediate APA changes upon HSC commitment and proliferation. In the future, we aim to link APA profiles in HSPCs to miRNA expression profiles. To this end, we have established an ultra-low input small RNA-seq approach that allows us to globally asses miRNA expression profiles in HSCs, MPPs, sHSCs and pHSCs. We already

generated the sequencing data in all of these populations, and analysis is currently ongoing. We aim to understand if 3'-UTR shortening in MPPs and upon HSC activation is induced to evade miRNA targeting. Using our novel miRNA datasets, we can compare the expression of miRNAs with their associated 3'-UTR targets and analyze if expression patterns in HSPCs are correlated. Further, the effect of APA on protein levels of different hits in HSPCs will be analyzed using microscopy-based immunofluorescence and flow cytometry analysis. In addition to *Gls*, we intend to functionally validate additional APA hits identified in our dataset. For example, we aim to understand if the longer 3'-UTR version of *Treml2* influences protein localization. Initial flow cytometry- and microscopy-based analysis of intracellular and plasma membrane expression of TREML2 in HSCs and MPP2 cells indicates that the longer 3'-UTR version could indeed mediate plasma membrane localization of TREML2. Additional experiments using fluorescent reporter constructs linked to the respective 3'-UTR will help to understand the impact of the different *Treml2* 3'-UTRs on protein localization. Besides, we also aim to understand how APA affects protein translation in hematopoietic cells. We crossed Rpl22-HA mice, which express the HA-tagged ribosomal subunit Rpl22 (Sanz et al., 2009), to CMV-Cre mice. Thereby, we generated mice, that express the tagged ribosome subunit in all body cells. Initial experiments showed that we are able to sort HSCs and MPP populations and efficiently pull down ribosome-bound RNA from low inputs. RNA-seq and 3'-Seq analysis of this RNA will allow us to determine if 3'-UTRs of a certain length exhibit enhanced ribosome association, indicating active translation. A future study will also focus on the role of APA in hematopoietic malignancies. To this end, 3'-Seq analysis of patient samples will be performed.

In addition to APA, we want to better understand the role of glutaminolysis in HSPCs. First, we aim to understand why proliferating HSCs depend on enhanced glutaminolysis and which metabolic pathways act downstream of glutamate. To this end, we aim to perform rescue experiments using *Nudt21* KD cells together with different metabolites, including αKG, nucleotides and glutathione. In addition, we plan to perform Mito Fuel Flex Seahorse experiments. This approach will help us to understand if activated HSCs indeed depend on different metabolic pathways than resting HSCs. Further, this test will be performed using *Nudt21* KD and control cells to analyze if glutaminolysis is indeed impaired in KD conditions. We also aim to establish metabolite tracing for low cellular input to investigate metabolic pathways in HSCs, activated HSCs and in response to *Nudt21* KD. Moreover, we want to focus on the specific role of the GLS isoforms GAC and KGA in HSCs. We already generated lentiviral vectors for overexpression of the GAC and the KGA isoform, respectively. We will use these for lentiviral transduction of HSCs. Afterwards, we plan to analyze functionality of GAC- and KGA-overexpressing HSCs *in vitro* and *in vivo*. Metabolic characterization of GAC-/KGA-overexpressing HSCs will also be performed. Likewise, it would be interesting to

independently target GAC and KGA in HSCs by CRISPR/Cas9 or isoform-specific KD. Functional analysis of GAC- or KGA-only HSCs will help to understand the specific roles of the isoforms in hematopoiesis. Metabolic analysis of these cells could also be performed to investigate if GAC is the isoform driving glutaminolysis in HSCs as indicated by reports in other cellular systems.

With regard to our newly generated *Pabpn1*^{flox} mouse model, we first aim to functionally test the generated mouse line by crossing it to different Cre drivers, including *Mx1* and *Vav1*. After confirming their functionality, the floxed mice can be used to interfere with APA patterns in different tissues and disease models. Thereby, analysis of the importance of APA and *Pabpn1* as an APA regulator in different settings is feasible.

5 Materials and Methods

5.1 Materials

5.1.1 Chemicals, reagents, cytokines and cell culture medium

Table 5.1: Chemicals, reagents, cytokines and cell culture medium

Item	Catalog No.	Manufacturer
2-Mercaptoethanol	M6250	Sigma-Aldrich Co. LCC
ACK Lysing Buffer	10-548E	Lonza BioWhittaker™
Agarose Standard	3810.3	Carl Roth GmbH & Co. KG
Amplicillin	10835242001	Roche Diagnostics GmbH
Boric acid	B6768	Sigma-Aldrich Co. LCC
Bromphenol blue	B0126	Sigma-Aldrich Co. LCC
CB-839	S7655	Selleckchem
Chloroform	28830L	Sigma-Aldrich Co. LCC
Chloroquine diphosphate salt	C6628	Sigma-Aldrich Co. LCC
DAPI	D1306	Thermo Fisher Scientific Inc.
Deoxyribonucleotide triphosphate (dNTP)	18427013	Thermo Fisher Scientific Inc.
Dimethyl sulfoxide (DMSO)	276855	Sigma-Aldrich Co. LCC
Direct PCR Tail Lysis Reagent	31-102-T	Peqlab
DMEM (1x)+GlutaMAX™-I	31966-021	Thermo Fisher Scientific Inc.
Dulbecco's Phosphate Buffered Saline	D8537	Sigma-Aldrich Co. LCC
Ethanol	BP2818-100	Thermo Fisher Scientific Inc.
Ethidium bromide solution	2218.4	Carl Roth GmbH & Co. KG
Ethylenediamine tetra-acetic acid-disodium salt (EDTA)	E6635	Sigma-Aldrich Co. LCC
Fetal Calf Serum (FCS)	10082147	Thermo Fisher Scientific Inc.
GeneRuler 1kb Plus DNA Ladder	SM1331	Thermo Fisher Scientific Inc.
GeneRuler 100 bp Plus DNA Ladder	SM0321	Thermo Fisher Scientific Inc.
Glycerol	A1123,1000	Applichem
GlycoBlue™ Coprecipitant	AM9515	Thermo Fisher Scientific Inc.
HEPES Buffer (1M)	H0887	Sigma-Aldrich Co. LCC
Hoechst 33342	H3570	Thermo Fisher Scientific Inc.
Isopropanol	33539	Sigma-Aldrich Co. LCC
KCl	P9541	Sigma-Aldrich Co. LCC
L-Glutamine 200 mM (100x)	25030-024	Thermo Fisher Scientific Inc.
LB Broth (MILLER) for microbiology	110285	Merck Millipore
LB-Agar (MILLER) for microbiology	110283	Merck Millipore
MethoCultTM GF M3434	03434	Stemcell Technologies
$MgSO_4$	M2643	Sigma-Aldrich Co. LCC
$Na_2HPO_4.7H_2O$	S9390	Sigma-Aldrich Co. LCC
$NaH_2PO_4.H_2O$	S5011	Sigma-Aldrich Co. LCC
NaOH	795429	Sigma-Aldrich Co. LCC
ONPG	N1127	Sigma-Aldrich Co. LCC
Opti-MEM I	31985062	Thermo Fisher Scientific Inc.
Passive Lysis 5X Buffer	E1941	Promega
Penicillin-Streptomycin	P4458	Sigma-Aldrich Co. LCC

Poly(I:C) HMW	tlrl-pic-5	InvivoGen
Polybrene	TR-1003-G	Sigma-Aldrich Co. LCC
QIAzol Lysis Reagent	79306	Qiagen Ltd.
Recombinant Human Flt3	300-19	PeproTech GmbH
Recombinant Murine Stem Cell Factor	250-03	PeproTech GmbH
Recombinant Murine Thrombopoietin	315-14	PeproTech GmbH
RetroNectin® GMP grade	T201	Takara Bio Inc.
Roti®-Fect	P001	Carl Roth GmbH & Co. KG
StemPro-34 SFM (1X)	10639011	Life Technologies™
tris(hydroxymethyl)aminomethane (Tris)	252859	Sigma-Aldrich Co. LCC
Tris HCl	10812846001	Sigma-Aldrich Co. LCC
Trypsin-EDTA Solution	T3924	Sigma-Aldrich Co. LCC
Ultrapure Water with 0.1% Gelatin	ES-006-B	Merck Millipore
UltraPure™ Distilled Water	10977-035	Life Technologies™
Xylencyanol	X4126	Sigma-Aldrich Co. LCC

5.1.2 Enzymes

Table 5.2: Enzymes

Enzyme	Manufacturer
BbsI	New England BioLabs Inc.
BstBI	New England BioLabs Inc.
DNase I	Qiagen Ltd.
EcoRI	New England BioLabs Inc.
Fisher BioReagents™ Taq DNA Polymerase	Thermo Fisher Scientific Inc.
HindIII-HF	New England BioLabs Inc.
NruI	New England BioLabs Inc.
Proteinase K	Peqlab
SnaBI	New England BioLabs Inc.
T4 DNA Ligase	New England BioLabs Inc.
XmnI	New England BioLabs Inc.

5.1.3 Buffer composition

Table 5.3: Buffer composition

Buffer	Composition
10x TBE	890 mM Tris 890 mM boric acid 20 mM EDTA pH 8.0 pH 8.0
DNA gel loading buffer	0.25% bromophenol blue 0.25% xylencyanol 15% glycerol in 10 ml dH_2O
DNase I stock solution	Lyophilized DNase I (1500 Kunitz units) dissolved in 550 µl RNase-free H_2O, aliquoted and stored at -20°C
DNase treatment	5 µl DNase solution 35 µl buffer RDD (provided in the RNase-free DNase set)

Microinjection buffer	10 mM TrisHCl 1 mM EDTA pH 7.5 in dH_2O
ONPG solution	0.55 g ONPG in 100 ml dH_2O
Z-buffer	1.60 g $Na_2HPO_4.7H_2O$ 0.55 g $NaH_2PO_4.H_2O$ 0.075 g KCl 0.012 g $MgSO_4$ pH 7.0 in 100 ml dH_2O

5.1.4 Oligonucleotides

See Table 5.18, Table 5.22, Table 5.29, Table 5.30, Table 5.33, Table 5.35 and Table 5.36 for sequences. Sequences are always indicated 5' – 3'.

Table 5.4: Oligonucleotides

Oligonucleotide	Manufacturer
crRNA	Integrated DNA Technologies (IDT)
Oligonucleotides for CRISPR/Cas9 cloning	Sigma-Aldrich Co. LCC
Primer for qPCR analysis	Sigma-Aldrich Co. LCC
Primer for qPCR genotyping	Sigma-Aldrich Co. LCC
ssDNA template	Integrated DNA Technologies (IDT)
tracrRNA (1072534)	Integrated DNA Technologies (IDT)

5.1.5 Plasmids, recombinant DNA and recombinant proteins

Table 5.5: Plasmids, recombinant DNA and recombinant proteins

Item	Catalog No.	Manufacturer
Alt-R S.p. Cas9 Nuclease 3NLS	1074181	Integrated DNA Technologies (IDT)
Packaging Plasmid pMD2.g	12259	Addgene
Packaging Plasmid pSPAX2	12260	Addgene
pCRISPR-Report	pTAL-Rep	(Wefers et al., 2013)
px330 Cas9	42230	Addgene
Scr Ctrl (pSMARTvector Non-targeting mEF1a/TurboGFP Glycerol Stock)	VSC11710	GE Helathcare Dharmacon, Inc.
shRNA Nudt21_1 (10) (pSMARTvector mEF1a/Turbo GFP Glycerol Stock; CAAATTACCTGGTGGGGAA)	V3SM11241-230917780	GE Helathcare Dharmacon, Inc.
shRNA Nudt21_2 (2) (pSMARTvector mEF1a/Turbo GFP Glycerol Stock; ACCATTAATCTGTACCCGC)	V3SM11241-234316681	GE Helathcare Dharmacon, Inc.
shRNA Nudt21_3 (4) (pSMARTvector mEF1a/Turbo	V3SM11241-237221737	GE Helathcare Dharmacon, Inc.

GFP Glycerol Stock; TCATTAGCTTTTGGTTCTA)		
shRNA Pabpn1_2 (pSMARTvector mEF1a/Turbo GFP Glycerol Stock; TAGTTATAGGATTCTCCCC)	V3SM11241-233531280	GE Helathcare Dharmacon, Inc.
shRNA Pabpn1_3 (pSMARTvector mEF1a/Turbo GFP Glycerol Stock; AATACCATGATGTCGCTCT)	V3SM11241-232142176	GE Helathcare Dharmacon, Inc.

5.1.6 Kits

Table 5.6: Kits

Item	Catalog No.	Manufacturer
Agencourt AMPure XP	A63881	Beckman Coulter
Agilent High Sensitivity DNA Kit	5067-4626	Agilent Technologies
Agilent RNA 6000 Pico Kit	5067-1513	Agilent Technologies
BD Cytofix/Cytoperm	554722	BD Biosciences
Calcium Phosphate Transfection Kit	K2780-01	Thermo Fisher Scientific Inc.
DNeasy Blood & Tissue Kits	69504	Qiagen Ltd.
DreamTaq Green PCR Master Mix (2X)	K1081	Thermo Fisher Scientific Inc.
Dynabeads™ Untouched™ Mouse T Cells Kit	11413D	Thermo Fisher Scientific Inc.
FITC Annexin V Apoptosis Detection Kit I	556547	BD Biosciences
GAC/KGA TaqMan Assay Mm01257297_m1	Mm01257297_m1	Thermo Fisher Scientific Inc.
Gapdh TaqMan Assay Mm99999915_g1	Mm99999915_g1	Thermo Fisher Scientific Inc.
KGA TaqMan Assay Mm01257298_m1	Mm01257298_m1	Thermo Fisher Scientific Inc.
NucleoBond® Xtra Midi	740410	Macherey-Nagel
NucleoSpin® Plasmid	740588	Macherey-Nagel
OneComp eBeads	01-1111-41	Thermo Fisher Scientific Inc.
PCR Add-on Kit for Illumina	SKU: 020.96.	Lexogen GmbH
PE Annexin V Apoptosis Detection Kit I	559763	BD Biosciences
PicoPure™ RNA Isolation Kit	KIT0204	Thermo Fisher Scientific Inc.
Power SYBR® Green PCR Master Mix	4367659	Thermo Fisher Scientific Inc.
QIAEX II Gel Extraction Kit	20021	Qiagen Ltd.
QIAGEN® Plasmid Maxi Kit	12163	Qiagen Ltd.
QIAquick PCR Purification Kit	28104	Qiagen Ltd.
QuantSeq 3′ mRNA-seq Library Prep Kit REV for Illumina	SKU: 016.96.	Lexogen GmbH
SMARTer® smRNA-seq Kit for Illumina®	635031	Takara Bio Inc.

SuperScript™ VILO™ Master Mix	11755050	Thermo Fisher Scientific Inc.
SUPREMERUN Tube - Supreme Sanger Sequencing	SUPREMERUN tube	Eurofins Genomics
TaqMan Fast Advanced Master Mix	4444963	Thermo Fisher Scientific Inc.

5.1.7 Cells and bacteria

Table 5.7: Cells and bacteria

Item	Catalog No.	Manufacturer
293FT cell line	R70007	Thermo Fisher Scientific Inc.
HeLa Cells	ATCC® CRM-CCL-2™	ATCC
One Shot™ Stbl3™ Chemically Competent *E. coli*	C7373-03	Life Technologies™

5.1.8 Consumables

Table 5.8: Consumables

Item	Manufacturer
0.45 µm HV Durapore® Membrane	Merck Millipore
18G Needles	Neoject Neomedic
25G Needles	Neoject Neomedic
384-well PCR plates	Thermo Fisher Scientific Inc.
96 Well Solid Polystyrene Microplate white	Sigma-Aldrich Co. LCC
Cell Strainer 40 µm, EASYStrainer	Greiner Bio One International GmbH
Conical Vials 15 ml	TPP Techno Plastic Products AG
Conical Vials 50 ml	Greiner Bio One International GmbH
Corning® tissue-culture treated culture dishes	Sigma-Aldrich Co. LCC
Disposable Scalpel	Feather Safety Razor Co Ltd.
ERSTA TBC-Syringe 1 ml	Codan
FACS tubes conical 4.5 ml	Greiner Bio One International GmbH
FACS tubes round bottom, REF 55.1579	Sarstedt AG & Co
MCT-175-C-S RNase free tubes	Axygen Inc.
Microvette® 500 K3E	Sarstedt AG & Co
Pasteur Capillary Pipettes	Wilhelm Ulbrich GdbR
PCR SingleCap 8er-SoftStrips	Biozym Scientific GmbH
Petri Dish 35 mm	Falcon Brand Products
Safe Lock Tubes	Eppendorf AG
Serological Pipettes	Corning Inc.
Stericup Quick Release-HV Sterile Vacuum Filtration System	Merck Millipore
Thinwall Polypropylene Tubes	Beckman Coulter
TipOne Filter Pipet Tips (10 µL, 20 µL, 200 µL, 1000 µL)	Starlab International GmbH
Tissue culture 6/12 well plate	Falcon Brand Products
Tissue Culture Flask 150	TPP Techno Plastic Products AG
Ultra Low Attachment 6/24/96 Well Plates	Corning Inc.
Ultra-Fine™ Needle 12.7 mm x 30 G 0.5 ml	BD Biosciences

5.1.9 Equipment

Table 5.9: Equipment

Equipment	Type	Manufacturer
2100 Bioanalyzer Instrument	G2939BA	Agilent Technologies
Cell Sorter	BD FACSAria I, BD FACSAria II, BD FACSAria III, BD FACSAria Fusion	BD Biosciences
Cell Viability Analyzer	Vi-CellXR	Beckman Coulter
Centrifuge	5415D	Eppendorf AG
Centrifuge	5810R	Eppendorf AG
Centrifuge	5427R	Eppendorf AG
Chip Priming Station	5065-4401	Agilent Technologies
CO_2 Incubator	Forma Steri-Cycle	Thermo Fisher Scientific Inc.
Digital Tube Roller	SRT9D	Stuart Bibby Scientific Limited
Electrophoresis power supply	PowerPac Basic Power Supply	Bio-Rad Laboratories
Flow Cytometer	BD LSRII, BD LSRFortessa	BD Biosciences
Gel Imaging System	UV-Transluminator and video system	Intas
Heating block	MHR23	Ditabis
Hemicytometer	Neubauer counting chamber	Braun
Incubation Shaker	Multitron	Infors HT
Luciferase Measurement Device	Perkin Elmer Wallac 1420 Victor2 Microplate Reader	GMI Inc.
Magnetic Separator	DYNAL Invitrogen Bead Separator	Thermo Fisher Scientific Inc.
Magnetic Separator	FastGene MagnaStand 0.2	Nippon Genetics Europe GmbH
Microscope	Zeiss Primo Vert inverted microscope equipped with an AxioCam Erc 5s	Zeiss
PCR Workstation	AirClean 600	Starlab International GmbH
Pipetboy	Pipetboy2	INTEGRA Biosciences AG
Pipettes	Research plus	Eppendorf AG
Real-Time PCR System	ViiA™ 7	Applied Biosystems™
Refrigerator	Refrigerator 4°C	Liebherr
Refrigerator	Refrigerator -20°C	Liebherr
Refrigerator	Refrigerator -80°C	SANYO Electric Co.
Sequencer	HiSeq 2000 Sequencing System	Illumina
Spectrophotometer	NanoDrop 2000	Thermo Fisher Scientific Inc.
Thermocycler	T3000	Biomet Inc.
Tissue Culture Hood	Safe 2020	Thermo Fisher Scientific Inc.
Ultracentrifuge	Optima L90-K	Beckman Coulter

Vortex Mixer	IKA MS 3 S36 Basic Chip Vortex	IKA-Werke
Vortex Mixer	Vortex Mixer 7-2020	neoLab Migge GmbH
Weighing balance	Precision weighing balance ALC-110.4	Acculab

5.1.10 Computer and software

Table 5.10: Computer and software

Computer/Software	Company
2100 Expert Software	Agilent Technologies
Adobe Acrobat Reader	Adobe Systems
Adobe Illustrator CS6	Adobe Systems
Endnote X8.2	Clarivate Analytics
FACSDiva Software	BD Biosciences
FlowJo 10.4.1	FlowJo, LCC
GraphPad Prism 8 for Mac	GraphPad
Integrative Genomics Viewer	Broad Institute and the Regents of the University of California
MacBookPro notebook	Apple
Microsoft Excel 2016 for Mac	Microsoft
Microsoft PowerPoint 2016 for Mac	Microsoft
Microsoft Word 2016 for Mac	Microsoft
R	R Core Team
RStudio	RStudio, Inc.
SnapGene Viewer Version 4	GSL Biotech LLC
ViiA7 Software	Thermo Fisher Scientific Inc.

5.1.11 Internet resources

Table 5.11: Internet resources

Resource	Website
Ensembl genome browser 95	https://www.ensembl.org/
National Center for Biotechnology Information	https://www.ncbi.nlm.nih.gov/
REViGO	http://revigo.irb.hr/

5.1.12 Mouse strains

Table 5.12: Mouse strains

Strain	Sex
C57Bl6/J (CD45.1, CD45.2, CD45.1/.2)	m/f
MxCre	m/f
MxCre *Meg3*$^{\text{mat flox/pat wt}}$	m/f
Meg3$^{\text{mat flox/pat wt}}$	m/f
VavCre	m/f
VavCre *Meg3*$^{\text{mat flox/pat wt}}$	m/f
Pabpn1$^{\text{flox}}$	m/f

5.1.13 Antibodies

Anti-PABPN1 and anti-GLS antibodies were purchased from abcam. All other antibodies were purchased from eBioscience and/or BD Biosciences.

Table 5.13: Antibodies

Antigen	Label	Clone
B220	AF700, PE-Cy7, PE-Cy5, biotin, PB, APC	RA3-6B2
CD117 (c-Kit)	BV-711, PE, APC	2B8
CD11b	APC-Cy7, PE-Cy7, AF700, biotin, PB, APC	M1/70
CD127	PE	A7R34
CD135	PE	RAM34
CD150	PE-Cy5	TC15-12F12.2
CD16/32	APC	93
CD25	APC	PC61
CD34	FITC, AF700	RAM34
CD4	PE-Cy7, AF700, PE, PE-Cy5, biotin, PB, APC-Cy7, APC	GK1.5
CD44	PE-Cy5	IM7
CD45.1	PE-Cy7, PE	A28
CD45.2	PB	104
CD48	PB, PE-Cy7, PE, AF700	HM48-1
CD71	PE	R17217
CD8a	PE-Cy7, PB, AF700, PE, PE-Cy5, biotin, PB, APC	53-6.7
GLS	AF488	EP7212
Gr1	APC, PE-Cy7, AF700, biotin, PB	RB6-8C5
IgM	PE-Cy5	II/41
Ki-67	AF647	B56
Ly-6A/E (Sca-1)	APC-Cy7	D7
PABPN1	AF647	EP3000Y
Streptavidin	PE-TxRed	-
Ter119	PB, PE-Cy7, AF700, FITC, biotin, APC	TER-119

5.2 Methods

5.2.1 Methods – results part 3.1

Most of this methods section (5.2.1) has been published as part of an article in *Scientific Reports* (Sommerkamp et al., 2019).

5.2.1.1 Animals

All mice were kept under specific pathogen-free (SPF) conditions in individually ventilated cages (IVC) at the animal facility of the German Cancer Research Center (DKFZ, Heidelberg). *Meg3* flox mice were generated by Klibanski et al., loxP sites flank exon 1 to 4 (Klibanski et al., unpublished). Female *Meg3* flox mice were crossed to male MxCre driver mice (Kuhn et al., 1995), generating MxCre *Meg3* $^{\text{mat flox/pat wt}}$ mice. Cre+ animals were used in experiments. MxCre mice or *Meg3* $^{\text{mat flox/pat wt}}$ mice were used as controls. All animals were intraperitoneally injected 5 times over 10 days with pIC (100 µg pIC in PBS) 6-18 weeks after birth. Analysis or further experiments were started 6-12 weeks after injections. Female *Meg3* flox mice were crossed to male VavCre driver mice (Georgiades et al., 2002) for analysis of embryonic *Meg3* deletion, generating VavCre *Meg3* $^{\text{mat flox/pat wt}}$ mice. Mice were analyzed 9-15 weeks after birth. *Meg3* $^{\text{mat flox/pat wt}}$ littermates were used as controls. C57BL/6J mice, 11-17 weeks old (CD45.2, CD45.1 or CD45.2/CD45.1) were either purchased from Harlan Laboratories (now Envigo) (the Netherlands) or Janvier Labs (France) or bred in-house. Mice were sacrificed by cervical dislocation, and all animal procedures were performed according to protocols approved by the German authorities, Regierungspräsidium Karlsruhe (Nr. A-23/17, Z110/02, DKFZ299 and G-149/16). To reduce animal numbers, remaining bone marrow/cDNA samples generated in this and previous studies were used whenever possible. Generation of miRNA-seq data was performed using residual material from *Meg3* KO and control mice generated during the reported experiments under the test number G-149/16.

5.2.1.2 Genotyping

For genotyping of mouse lines housed in the DKFZ animal core facility, tails were cut (approx. 3 mm length) and incubated with 100 µl Direct PCR tail lysis reagent containing 0.2 mg/ml proteinase K at 55°C for 3-16 h. Subsequently, proteinase was inactivated by incubation at 85°C for 45 min. Samples were stored for further PCR analysis (1 µl/reaction) at 4°C.

Genotyping of mutant mouse strains was performed by PCR. PCRs were performed according to a standard protocol (Table 5.14, Table 5.15, Table 5.16 and Table 5.17) with the primers listed in Table 5.18.

Genotyping PCR products were separated based on their length using gel electrophoresis. Gels were prepared by boiling a suspension of 1-2% agarose in 1x TBE buffer. After addition of 30 µg ethidium bromide per 100 ml, the solution was filled in a gel tray containing a comb. PCR products were loaded without addition of loading dye, as this is already included in the DreamTaq Green PCR Mastermix. PCR fragments were separated by applying 120V. Finally, DNA fragments were visualized using an UV transilluminator, and gels were documented using a camera system.

Table 5.14: Genotyping PCR mastermix

Component	Amount
DreamTaq Green PCR Mastermix	12.5 µl
Fwd primer (10 µM)	0.6 µl
Rev primer (10 µM)	0.6 µl
DNA	1 µl
H2O	11.3 µl

Table 5.15: Genotyping PCR program MxCre

Step	Temperature	Time	Number of cycles
Denaturation	95°C	2 min	
Amplification	94°C 62°C 72°C	45 s 45 s 45 s	40
Final elongation	72°C	5 min	
Storage	4°C	∞	

Table 5.16: Genotyping PCR program VavCre

Step	Temperature	Time	Number of cycles
Denaturation	95°C	2 min	
Amplification	94°C 70°C 72°C	45 s 45 s 45 s	40
Final elongation	72°C	5 min	
Storage	4°C	∞	

Table 5.17: Genotyping PCR program *Meg3* flox

Step	Temperature	Time	Number of cycles
Denaturation	95°C	2 min	
Amplification	95°C 62°C 72°C	45 s 30 s 60 s	40
Final elongation	72°C	10 min	
Storage	4°C	∞	

Table 5.18: Genotyping primers

Genotype	Primer	Sequence	Amplicon
MxCre	Mx1	GGCAGGGCTCCTCAGTGATTC	550 bp
	Cre2L21	CTGGCGATCCCTGAACATGTC	
VavCre	Vav-cre-01MR9266	AGATGCCAGGACATCAGGAACCT	236 bp
	Vav-cre-01MR9267	ATCAGCCACACCAGACACAGAGA	
Upper *Meg3* flox	Lox1	TAGTCCGTCAGAATCGGGGTAC	wt: 419 bp
	SDL2	CTATGACAAATTGCTGTAGCCGCAC	flox: 481 bp
Lower *Meg3* flox	NDEL1	CAGACTCCTGCAGCCCCTAT	wt: 557 bp
	NDEL2	CACCATCCCTAAAGAGGTGATCTCA	flox: 735 bp

5.2.1.3 pIC-induced stress experiments

Mice were intraperitoneally injected with 100 μg pIC (in PBS) or PBS either 8 days, 4 days or 16 h prior to endpoint analysis. For serial stress experiments, mice were injected 2 times a week for a period of 4 weeks. Mice were analyzed 8 weeks after the last pIC injection.

5.2.1.4 Cell suspension and flow cytometry

Peripheral blood and BM analysis

For isolation of BM, spleen and thymus cells, mice were sacrificed by cervical dislocation. Femur, tibia, hip and spine bones, spleen and thymus were isolated. Bones were cleaned from muscles and connective tissues using a scalpel and forceps and crushed in ice-cold PBS using pistil and mortar. The resulting cell suspension was filtered through a 40 μm cell strainer and pelleted by centrifugation. Total BM cells were used for colony forming unit or transplantation assays. For flow cytometry analysis and FACS, erythrocytes were lysed using ACK Lysing Buffer (5 min, RT, 3 ml/sample). Lysis was stopped by addition of 7 ml PBS. BM cells were counted using a Vi-CellXR Cell Viability Analyzer and used for flow cytometry stainings.

PB was collected into EDTA-coated tubes. Erythrocyte lysis was performed before flow cytometry analysis using ACK Lysing Buffer (5 min, RT, 1 ml/sample). Lysis was stopped by addition of 1 ml PBS. White blood cells were pelleted by centrifugation (1,500 rpm at 4°C). Depending on lysis efficiency, erythrocyte lysis was repeated as described.

For spleen and thymus, single cell suspensions were generated by passing thymus or spleen tissue through a 40 μm cell strainer. Cells were collected in PBS, and lysis of erythrocytes was

performed for spleen samples as previously described for BM samples. Afterwards, cell concentration was determined using a Vi-CellXR Cell Viability Analyzer.

FACS

For FACS, total BM cells were lineage depleted using the Dynabeads Untouched Mouse CD4 Cells Kit. Total BM was stained with a 1:5 dilution of the lineage cocktail provided in the kit in PBS for 30 min (3 ml/mouse). Cells were washed with PBS and incubated for 20 min with 1.5 ml washed Dynabeads per mouse. A magnet was used for cell depletion to enrich for the lineage-negative cell fraction. To purify HSCs, the depleted cells were stained for 30 min using the monoclonal antibodies indicated in Table 5.19. For isolation of transplanted HSCs, monoclonal antibodies were used as indicated in Table 5.19. Cell sorting was performed on a BD FACSAriaI, FACSAriaII, FACSAriaIII or FACSAria Fusion. Cell populations were identified and purified according to the cell surface phenotype listed in Table 5.20. Sorted cells were collected into RNA lysis buffer for qPCR analysis or QIAzol Lysis Reagent for small RNA-seq analysis and stored at -80°C.

Table 5.19: Antibody combinations used for FACS stainings for results part 3.1

Staining	Surface marker	Antibody clone (fluorophore)
HSCs (BM)	CD4	GK1.5 (PE-Cy7)
	CD8a	53-6.7 (PE-Cy7)
	CD11b	M1/70 (PE-Cy7)
	Gr1	RB6-8C5 (PE-Cy7)
	Ter119	Ter-119 (PE-Cy7)
	B220	RA3-6B2 (PE-Cy7)
	c-Kit	2B8 (PE)
	Sca-1	D7 (APC-Cy7)
	CD48	HM48-1 (PB)
	CD150	TC15-12F12.2 (PE-Cy5)
	CD34	RAM34 (FITC)
Transplanted HSCs (BM)	Lin (CD4, CD8a, CD11b, Gr1, Ter119, B220)	Clones see HSCs, all AF700
	CD117	2B8 (APC)
	Sca-1	D7 (APC-Cy7)
	CD48	HM48-1 (PE)
	CD150	TC15-12F12.2 (PE-Cy5)
	CD34	RAM34 (FITC)
	CD45.1	A28 (PE-Cy7)
	CD45.2	104 (PB)

Table 5.20: Hematopoietic populations and corresponding cell surface markers

Abbreviation	Population name	Cell surface phenotype
Lin	lineage	CD11b+ Gr-1+ B220+ CD4+ CD8a+ Ter119+
LSK	hematopoietic stem and progenitor cells	Lin- Sca-1+ c-Kit+
LS-K	myeloid committed progenitors	Lin- Sca-1- c-Kit+
HSC	hematopoietic stem cell	LSK CD150+ CD48- CD34- (CD135-)
MPP1	multipotent progenitor 1	LSK CD150+ CD48- CD34+ (CD135-)
MPP2	multipotent progenitor 2	LSK CD150+ CD48+ CD34+ (CD135-)
MPP3	multipotent progenitor 3	LSK CD150- CD48+ CD34+ (CD135-)
MPP4	multipotent progenitor 4	LSK CD150- CD48+ CD34+ (CD135+)
MPP3/4	multipotent progenitor 3/4	LSK CD150- CD48+ CD34+
sHSC	steady-state HSC	LK CD150- CD48+ CD34-
pHSC	proliferating HSC	LK CD150- CD48+ CD34-
CMP	common myeloid progenitor	LS-K IL7Rα- CD34+ CD16/32^{low}
GMP	granulocyte macrophage progenitor	LS-K IL7Rα- CD34+ CD16/32^{high}
MEP	megakaryocyte erythrocyte progenitor	LS-K IL7Rα- CD34- CD16/32^{low}
B cells	B lymphocytes	(CD11b-) (Gr1-) B220+
Pre/Pro B cells	pre/pro B lymphocytes	(CD11b-) (Gr1-) B220^{low} IgM-
Immature B cells	immature B lymphocytes	(CD11b-) (Gr1-) B220^{low} IgM+
Mature B cells	mature B lymphocytes	(CD11b-) (Gr1-) B220^{high} IgM+
T cells	T lymphocytes	(CD11b-) (Gr1-) CD4+ or CD8a+
Myeloid cells	myeloid cells	CD11b+ Gr1+
DP	double positive thymocytes	CD4+ CD8a+
DN1	double negative thymocytes 1	CD4- CD8a- CD44+ CD25-
DN2	double negative thymocytes 2	CD4- CD8a- CD44+ CD25+
DN3	double negative thymocytes 3	CD4- CD8a- CD44- CD25+
DN4	double negative thymocytes 4	CD4- CD8a- CD44- CD25-

Flow cytometry

Ery-lysed peripheral blood, spleen cells, thymus cells and BM cells were used for flow cytometry stainings. Table 5.21 defines the monoclonal antibodies applied for the different surface staining approaches. Cells were stained for 30 min at 4°C. For cell cycle stainings, surface stainings were performed, and cells were fixed and permeabilized by using BD

Cytofix/Cytoperm buffer for 15 min at 4°C. Subsequently, intracellular Ki-67 (B56)-AF647 staining was performed in PermWash solution for 30 min at 4°C. Then, cells were stained with Hoechst 33342 for 15 min at RT and subsequently analyzed on a BD LSRII or BD LSRFortessa flow cytometer. Cell populations were identified according to the cell surface phenotype listed in Table 5.21.

Table 5.21: Antibody combinations used for stainings for flow cytometry analysis for results part 3.1

Staining	Surface marker	Antibody clone (fluorophore)
Differentiated Cells (PB, BM, Spleen)	CD4	GK1.5 (PE-Cy7)
	CD8a	53-6.7 (PE-Cy7)
	CD11b	M1/70 (APC-Cy7)
	Gr1	RB6-8C5 (APC)
	Ter119	Ter-119 (PB)
	CD71	R17217 (PE)
	B220	RA3-6B2 (AF700)
	IgM	II/41 (PE-Cy5)
HSCs (BM)	Lin (CD4, CD8a, CD11b, Gr1, Ter119, B220)	Clones see differentiated cells, all PE-Cy7
	CD117	2B8 (BV-711)
	Sca-1	D7 (APC-Cy7)
	CD48	HM48-1 (PB)
	CD150	TC15-12F12.2 (PE-Cy5)
	CD34	RAM34 (FITC)
	CD16/32	93 (APC)
	CD127	A7R34 (PE)
Thymus	CD4	GK1.5 (PE-Cy7)
	CD8a	53-6.7 (PB)
	CD25	PC61 (APC)
	CD44	IM7 (PE-Cy5)
Cell Cycle (BM)	Lin (CD4, CD8a, CD11b, Gr1, Ter119, B220)	Clones see differentiated cells, all AF700
	CD117	2B8 (PE)
	Sca-1	D7 (APC-Cy7)
	CD48	HM48-1 (PE-Cy7)
	CD150	TC15-12F12.2 (PE-Cy5)
	CD34	RAM34 (FITC)
	Ki-67 (intracellular)	B56 (AF647)
Differentiated Cells Transplants (PB)	CD4	GK1.5 (PE)
	CD8a	53-6.7 (PE)
	CD11b	M1/70 (APC-Cy7)
	Gr1	RB6-8C5 (APC)
	B220	RA3-6B2 (PE-Cy5)
	CD45.1	A28 (PE-Cy7)
	CD45.2	104 (PB)
Differentiated Cells Transplants (BM, Spleen)	CD4	GK1.5 (PE-Cy5)
	CD8a	53-6.7 (PE-Cy5)

	CD11b	M1/70 (APC-Cy7)
	Gr1	RB6-8C5 (APC)
	Ter119	Ter-119 (FITC)
	CD71	R17217 (PE)
	B220	RA3-6B2 (AF700)
	CD45.1	A28 (PE-Cy7)
	CD45.2	104 (PB)
HSCs Transplants (BM)	Lin (CD4, CD8a, CD11b, Gr1, Ter119, B220)	Clones see differentiated cells, all AF700
	CD117	2B8 (BV-711)
	Sca-1	D7 (APC-Cy7)
	CD48	HM48-1 (PE)
	CD150	TC15-12F12.2 (PE-Cy5)
	CD34	RAM34 (FITC)
	CD45.1	A28 (PE-Cy7)
	CD45.2	104 (PB)
Cell Cycle Transplants (BM)	Lin (CD4, CD8a, CD11b, Gr1, Ter119, B220)	Clones see differentiated cells, all biotin (2^{nd} antibody Strepdavidin-PE-TR)
	CD117	2B8 (PE)
	Sca-1	D7 (APC-Cy7)
	CD48	HM48-1 (AF700)
	CD150	TC15-12F12.2 (PE-Cy5)
	CD34	RAM34 (FITC)
	CD45.1	A28 (PE-Cy7)
	Ki-67 (intracellular)	B56 (AF647)

5.2.1.5 Gene expression analysis by qPCR

HSCs (MxCre or MxCre *Meg3* $^{mat\ flox/pat\ wt}$ LSK CD150+ CD48- CD34-; fetal liver HSCs: E13.5 CD48- CD41+ cKit+ CD34+; adult HSCs (8-12 weeks): LSK CD150+ CD48- CD34- CD135-; aged HSCs (24 months): LSK CD150+ CD48- CD34- CD135-) were sorted by FACS into RNA lysis buffer for qPCR analysis and stored at -80°C. RNA was isolated using the ARCTURUS PicoPure RNA Isolation Kit following the manufacturer's instructions with a DNase digestion step. For cDNA synthesis, reverse transcription was performed using the SuperScript VILO cDNA Synthesis Kit according to manufacturer's guidelines. cDNA samples were diluted in DNase- and RNase-free water (total volume of 200 µl per 10,000 sorted cells). Primers (Table 5.22) were used with a final concentration of 0.5 µM in a total reaction volume of 10 µl. Fast SYBR Green Master Mix was used on a ViiA 7 Real-Time PCR System for qPCR analysis (see Table 5.23 for qPCR program). Fluorescent signal was detected at the end of every amplification step and continuously during the melting curve. Amplification curves were analyzed using the ViiA 7 software. RNA expression was normalized to the housekeeping gene *Oaz1* and presented as relative quantification (ratio = 2^-DDCT). Primers were designed using the Universal ProbeLibrary Assay Design Center (Roche).

Table 5.22: qPCR primer for results part 3.1 (synthesis by Sigma-Aldrich)

Primer	Sequence
Cdk6 fwd	cgagtgcagaccagtgagg
Cdk6 rev	tgtgcacacatcaaacaacct
c-Myc fwd	caccagcagcgactctga
c-Myc rev	ggggtttgcctcttctcc
Egr1 fwd	agttccaacgccctctgac
Egr1 rev	acaggaacccgaaccaca
Hoxa5 fwd	agctgcacattagtcacgaca
Hoxa5 rev	gcggttgaagtggaattctt
Hoxb4 fwd	ctggatgcgcaaagttcac
Hoxb4 rev	gtgaaactccttctccaactcc
Hoxb5 fwd	gcttcacatcagccacgata
Hoxb5 rev	caggtagcgattgaagtggaat
Meis1 fwd	gacgctttaaagagagataaagatgc
Meis1 rev	catttctcaaaaatcagtgctaaga
Meg3 fwd	cgaggacttcacgcacaa
Meg3 rev	attccagatgatggctttg
Oaz1 fwd	tttcagctagcatcctgtactcc
Oaz1 rev	gaccctggtcttgtcgttaga
Tgm2 fwd	cctgaccctggatccctact
Tgm2 rev	cacccgctgtacttctcgtag

Table 5.23: qPCR thermal cycling program

Step	Temperature	Time	Number of cycles
Initial denaturation/ activation of polymerase	95°C	10 min	
Amplification	95°C 60°C	15 s 45 s	40
Melting curve	65°C-95°C	0.5°C/s	

5.2.1.6 Small RNA sequencing

Sample preparation and library generation

Approximately 10,000-20,000 LSK SLAM cells (LSK CD150+ CD48-) were sorted by FACS into 100 µl QIAzol Lysis Reagent. Total RNA extraction was performed by adding 20 µl chloroform. Samples were incubated at RT for 3 min and centrifuged at 12,000 rpm for 5 min at RT. The aqueous phase was transferred into a new PCR tube and 0.4 µl glycoblue and 75 µl isopropanol were added per sample. Samples were stored at -20°C for at least 5 days. Samples were centrifuged at 13,000 rpm for 1 h at 4°C. Samples were washed with 70% ethanol and centrifuged at 13,000 rpm for 15 min at 4°C. Supernatant was discarded and the

pellet was dried for 1 min. The pellet was resuspended in 8 µl H_2O. RNA quality was analyzed using the Agilent RNA 6000 Pico Kit according to manufacturer's instructions. Libraries for small RNA-seq were generated using the SMARTer® smRNA-seq Kit for Illumina® according to manufacturer's instructions. 23 ng of total RNA were used, and 15 cycles were applied for PCR amplification. Quality was checked using the Agilent High Sensitivity DNA Kit according to manufacturer's instructions. Size selection was performed by using AMPure XP Beads according to manufacturer's instructions. Sequencing was performed on an Illumina HiSeq2000 device.

5.2.1.7 miRNA sequencing analysis

Genomic mapping

The adapter sequences ATGATCGGAAGAG, GATGAAGAACGCAG, AGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT, GGTCTACG, TGGAATTCTCGGGTGCCAAGG, TCGTATGCCGTCTTCTGCTTG, AGATCGGAAGAGCACACGTCTGAACTCCAGTCA were removed with Trimmomatic (Bolger et al., 2014) (version 0.32, parameters: SE ILLUMINACLIP:adapter:2:30:10 HEADCROP:4 TRAILING:3 MINLEN:17) from the fastq sequences. Additionally all reads containing N and all sequences <17 bases were removed. HomerTools (Heinz et al., 2010) (v3.18, trim -3 AAAAAAAAA) for removing poly(A) tails and fastx_artifacts_filter (FASTX Toolkit 0.0.13 http://hannonlab.cshl.edu/fastx_toolkit/) were used for further cleaning. Then, the clean fastq files were mapped against mouse genome 38 (NCBI) using Bowtie 2 version 2.2.4 (Langmead and Salzberg, 2012) (parameters -N 0 -L 8). The resulting SAM files were converted to BAM and sorted using samtools version 1.3.1.

miRNA mapping and counting

The clean fastq files were mapped with Bowtie version 0.12.9 (-f -v 1 -a --best -S) against an assembled ncRNA database consisting of a unique set of miRNAs of Ensembl version 95 (http://www.ensembl.org/) and miRBase (v21, http://www.mirbase.org/). The resulting SAM files were parsed, and the number of reads for each mature miRNA was counted using Perl.

Comparison of control Cre and KO samples

For the comparison with DESeq2 (Love et al., 2014), the input tables containing the replicates for the groups to compare were created by a custom Perl script. In the count matrix, rows with an average count number < 5 were removed. Then, DESeq2 (version 1.4.1) was run with default parameters.

Data availability

The small RNA-seq dataset generated and analyzed during the current study is available in the ArrayExpress repository, ArrayExpress accession E-MTAB-7519 (http://www.ebi.ac.uk/arrayexpress/).

5.2.1.8 Transplantation experiments and CFU assays

Full chimeras

1x10^6 total BM cells from induced MxCre control mice or induced MxCre *Meg3*$^{\text{mat flox/pat wt}}$ mice were transplanted into fully irradiated (2x5 Gy) CD45.1 mice within 24 h after irradiation by tail vein injection. Approximately 4 months after transplantation, mice were euthanized and 3x10^6 total BM cells were transplanted into fully irradiated (2x5 Gy) CD45.1 secondary recipient mice. Approximately 4 months after transplantation, mice were euthanized and 3x10^6 total BM cells were transplanted into fully irradiated (2x5 Gy) CD45.1 tertiary recipient mice. Contribution of CD45.2+ donor cells was monitored in peripheral blood approximately every 4 weeks by flow cytometry using the monoclonal antibodies indicated in Table 5.21.

50/50 chimeras

2x10^5 total BM cells from induced MxCre control mice or induced MxCre *Meg3*$^{\text{mat flox/pat wt}}$ mice were transplanted into fully irradiated (2x5 Gy) CD45.1 mice within 24 h after irradiation by tail vein injection in competition to 2x10^5 total BM cells from CD45.1/.2 mice. Approximately 4 months after transplantation, mice were euthanized and 3x10^6 total BM cells were transplanted into fully irradiated (2x5 Gy) CD45.1 secondary recipient mice. Approximately 4 months after transplantation, mice were euthanized and 3x10^6 total BM cells were transplanted into fully irradiated (2x5 Gy) CD45.1 tertiary recipient mice. Contribution of CD45.2+ donor cells was monitored in peripheral blood approximately every 4 weeks by flow cytometry using the monoclonal antibodies indicated in Table 5.21.

Colony forming unit assay

40,000 freshly prepared BM cells were mixed with 3 ml of semi-solid MethoCult M3434 medium and plated on a 35 mm culture plate in technical duplicates (1 ml/replicate). Colonies were counted after 7 days using a Zeiss Primo Vert inverted microscope. For replating, cells were harvested by washing culture plates with PBS. Cells were counted using a hemocytometer, and 30,000 cells of each technical replicate were mixed with 1.5 ml of semi-solid MethoCult M3434 medium and plated on a 35 mm culture plate (1 ml/replicate). Colony number was determined after 8 days.

5.2.1.9 Statistical analysis

Statistical analysis was performed by unpaired or paired student's t-test or two- or one-way ANOVA without correction for multiple comparison (Fisher LSD test). All data are presented as mean + SD. Please see figure legends for detailed information. Significance levels were set at *$p < 0.05$, **$p < 0.01$ and ***$p < 0.001$. For statistical analysis, GraphPad Prism was used.

5.2.2 Methods – results part 3.2

Most of this methods section (5.2.2) has been submitted as part of a paper draft.

5.2.2.1 Animals

C57BL/6J (CD45.2, CD45.1 or CD45.2/CD45.1) mice were either purchased from Envigo (the Netherlands) or Janvier Labs (France) or bred in-house. For 3'-Seq analysis, 6- to 12-weeks-old female C57BL/6J (CD45.2) mice were used. All mice were bred in-house in the animal facility of the DKFZ under SPF conditions in IVCs. According to German guidelines, mice were euthanized by cervical dislocation, and all animal procedures were performed according to protocols approved by the German authorities, Regierungspräsidium Karlsruhe (Nr. A-23/17, Z110/02, DKFZ 299, G-183/17). To reduce animal numbers, remaining bone marrow/cDNA samples generated in this and previous studies were used whenever possible.

5.2.2.2 pIC-induced inflammatory stress

Mice were intraperitoneally injected with 100 μg pIC or PBS. 16 h later, mice were sacrificed, and bone marrow cells were used for subsequent experiments.

5.2.2.3 Cell suspension and flow cytometry

Mouse BM cells were isolated, and HSCs and MPP1-4 progenitors defined by their respective immunophenotype (Lineage- Sca1+ Kit+ CD135-/+ CD150-/+ CD48-/+ CD34-/+) or LSK (Lineage- Sca1+ Kit+) or sHSCs (PBS-injected control mice)/pHSCs (pIC-injected mice; after 16h) (Lineage- Kit+ CD150+ CD48- CD34-; excluding Sca-1 as previously described (Haas et al., 2015)) were purified by FACS and subsequently subjected to 3'-Seq, qPCR analysis, *in vitro* CB-839 treatment, lentiviral transduction, CFU analysis or reconstitution assays. See also 5.2.1.4 for further details. Briefly, BM was isolated from pooled femora, tibiae, ilia and vertebrae by gentle crushing in ice-cold PBS using a mortar and pistil. If no depletion of lineage-positive cells was done, lysis of erythrocytes was performed using ACK Lysing Buffer (see 5.2.1.4). To deplete lineage-positive cells, we used the Dynabeads Untouched Mouse CD4 Cells Kit (see 5.2.1.4). To purify HSCs and MPP1-4 cells, the Lin- fraction was stained for 30 min using the monoclonal antibodies listed in Table 5.24. To purify LSK cells, the Lin- fraction was stained for 30 min using the monoclonal antibodies listed in Table 5.24. For some experiments, the fluorophore color panel was changed to enable analysis of additional markers. Cell sorting was performed on a BD FACS Aria I, II, III or FACS Aria Fusion. Cell populations were identified and purified according to the cell surface phenotype listed in Table 5.20. Cells were sorted into Complete Stem Cell Medium (see 5.2.2.9) for CFUs, lentiviral transduction, *in vitro* culture and

reconstitution experiments, or RNA lysis buffer for population 3'-Seq and qPCR analysis and stored at -80°C.

Table 5.24: Antibody combinations used for FACS stainings for results part 3.2

Staining	Surface marker	Antibody clone (fluorophore)
HSC/MPP/sHSC/pHSC	CD4	GK1.5 (PE-Cy7)
	CD8a	53-6.7 (PE-Cy7)
	CD11b	M1/70 (PE-Cy7)
	Gr1	RB6-8C5 (PE-Cy7)
	Ter119	Ter-119 (PE-Cy7)
	B220	RA3-6B2 (PE-Cy7)
	c-Kit	2B8 (APC)
	Sca-1	D7 (APC-Cy7)
	CD135	RAM34 (PE)
	CD48	HM48-1 (PB)
	CD150	TC15-12F12.2 (PE-Cy5)
	CD34	RAM34 (FITC)
LSK	Lin (CD4, CD8a, CD11b, Gr1, Ter119, B220)	Clones see HSCs, all PB
	CD117	2B8 (PE)
	Sca-1	D7 (APC-Cy7)

GLS stainings

For flow cytometry analysis of GLS protein levels, homeostatic or pIC-treated BM cells were isolated from mice and stained for HSC/MPP markers as described using antibodies indicated in Table 5.25. Cells were fixed with BD Cytofix/Cytoperm buffer. Subsequently, intracellular GLS (AF488, clone EP7212, abcam) staining was performed using BD PermWash solution (see 5.2.1.4 for details on intracellular staining procedures). Cells were subsequently analyzed on a BD LSRII flow cytometer or BD LSRFortessa cell analyzer. For analysis of GLS expression in *Nudt21* KD cells, GFP+ cells were resorted 48 h or 72 h after lentiviral transduction, and intracellular GLS staining was performed. The GFP is lost during the fixation/permeabilization step, therefore staining of GLS with anti-GLS -AF488 is possible.

Table 5.25: Antibody combinations used for stainings for flow cytometry analysis of GLS for results part 3.2

Staining	Surface marker	Antibody clone (fluorophore)
GLS in HSCs/MPP Cells	CD4	GK1.5 (PE-Cy7)
	CD8a	53-6.7 (PE-Cy7)
	CD11b	M1/70 (PE-Cy7)
	Gr1	RB6-8C5 (PE-Cy7)
	Ter119	Ter-119 (PE-Cy7)
	B220	RA3-6B2 (PE-Cy7)
	c-Kit	2B8 (PE)
	Sca-1	D7 (APC-Cy7)
	CD48	HM48-1 (PB)

	CD150	TC15-12F12.2 (PE-Cy5)
	CD34	RAM34 (AF700)
	GLS (intracellular)	EP7212 (AF488)

PABPN1 stainings

For analysis of knockdown efficiency, GFP+ cells were resorted 48 h or 72 h after lentiviral transduction. GFP+ cells were then fixed with BD Cytofix/Cytoperm buffer. Subsequently, intracellular PABPN1 (AF647, clone EP3000Y, abcam) staining was performed using BD PermWash solution (see 5.2.1.4 for details on intracellular staining procedures). Cells were subsequently analyzed on a BD LSRII flow cytometer or BD LSRFortessa cell analyzer.

Cell cycle analysis

HSC/MPP surface staining (LSK, CD150, CD48, CD34) was performed on BM cells or CB-839-treated HSPCs (Table 5.26). Cells were fixed with BD Cytofix/Cytoperm buffer. Subsequently, intracellular Ki-67 staining was performed using BD PermWash solution. Prior to flow cytometry analysis, cells were stained with Hoechst 33342 or DAPI (see 5.2.1.4 for details on intracellular staining procedures). Cells were subsequently analyzed on a BD LSRII flow cytometer or BD LSRFortessa cell analyzer.

Table 5.26: Antibody combinations used for stainings for flow cytometry analysis of cell cycle of CB-treated cells for results part 3.2

Staining	Surface marker	Antibody clone (fluorophore)
CC in CB-Treated Cells and HSCs/MPPs in BM	CD4	GK1.5 (AF700)
	CD8a	53-6.7 (AF700)
	CD11b	M1/70 (AF700)
	Gr1	RB6-8C5 (AF700)
	Ter119	Ter-119 (AF700)
	B220	RA3-6B2 (AF700)
	c-Kit	2B8 (PE)
	Sca-1	D7 (APC-Cy7)
	CD48	HM48-1 (PE-Cy7)
	CD150	TC15-12F12.2 (PE-Cy5)
	CD34	RAM34 (FITC)
	Ki-67 (intracellular)	B56 (AF647)

Viability analysis

Cells were harvested, and viability staining was performed using the FITC or PE Annexin V Apoptosis Detection Kit I (BD Biosciences) according to the manufacturer's protocol. Cells were subsequently analyzed on a BD LSRII flow cytometer or BD LSRFortessa cell analyzer.

5.2.2.4 Lentiviral KD experiments

shRNA-expressing plasmids were purchased from Dharmacon as bacterial glycerol stocks. Maxi-preps were performed to generate shRNA plasmids for transfection using Qiagen Plasmid Maxi Kit according to manufacturer's instructions. Viruses were packaged by co-transfection of shRNA plasmids (Scr GFP, *Pabpn1* shRNA_2 and _3, *Nudt21* shRNA_1(10),_2(2),_3(4); Dharmacon; see Table 5.5 for details) with pSPAX2 and pMD2.g into 293T cells using the Invitrogen Calcium Phosphate Transfection Kit according to manufacturer's guidelines and 25 μM chloroquine added additionally to the cells. After 12 h of transfection, cells were cultured in 0.1% gelatin-coated cell culture flasks in DMEM GlutaMAX (+10% FCS, 1% penicillin-streptomycin (P/S), 1% L-glutamine, 1% HEPES) at 37°C. Viral supernatants were collected 48 h post-transfection and filtered through a 45 μm filter and centrifuged at 22,000 rpm (2 h at 4°C). After centrifugation, the virus pellet was re-suspended in PBS and stored at -80°C. Viral titers were determined using 293T cells and subsequent flow cytometry analysis to quantify reporter fluorophore expression.

For lentiviral transduction of LSK cells, BM cells were sorted as described (see 5.2.2.3). Sorted LSK cells were cultured in retronectin-coated (Takara; 4 μg/cm^2) 96-/24-/6-well ultra-low attachment plates in Transduction Medium (StemPro-34 SFM, Life Technologies containing 50 ng/ml SCF, 25 ng/ml TPO, 30 ng/ml Flt3-Ligand (all Preprotech), 100 u/ml penicillin/streptomycin, 2 mM L-glutamine, 8 μg/ml Polybrene (Sigma-Aldrich) and StemPro-34 Supplement as recommended) and transduced with an MOI of 35. Medium was exchanged to Complete Stem Cell Medium (see 5.2.2.9) 24 h later. Cells were harvested 48 h or 72 h after transduction for sorting of GFP+ cells or flow cytometry analysis and viability analysis. GFP+ LSK cells were used for 3'-Seq analysis, and GFP+ cells were used for qPCR analysis, flow cytometry analysis and functional assays, including CFU and reconstitution experiments.

5.2.2.5 Colony-forming unit assays using KD cells

1,500 FACS-sorted GFP+ cells were resuspended 3 ml MethoCult M3434 (StemCell Technologies), and1 ml of Methocult was plated per technical replicate, respectively. Five days after plating of CFUs, colony formation was quantified. 20,000 cells of each replicate were used for replating and cultured in MethoCult M3434 (StemCell Technologies). Residual Cells were stained for flow cytometry analysis as described (Table 5.24, see also 5.2.2.3). 7 days after replating, colonies were quantified (for detailed procedure of CFU assays see 5.2.1.8).

5.2.2.6 Reconstitution experiments using KD cells

Full chimeras and 50/50 chimeras

For generation of full chimeras, 30,000 GFP+ cells (KD or Scr Ctrl, CD45.2) were sorted 48 h after lentiviral transduction of LSK cells per recipient mouse and injected intravenously.

For generation of 50/50 chimeras, 15,000 GFP+ cells were sorted 48 h after lentiviral transduction of LSK cells per recipient mouse. We generated two sets of chimeras: CD45.1/.2 Scr Ctrl + CD45.2 KD and CD45.1/.2 KD + CD45.2 Scr Ctrl. Always, 15,000 + 15,000 GFP+ cells were mixed. As CD45.2 cells are known to perform better in transplantation settings, we always compared contribution of CD45.2 KD and CD45.2 Scr Ctrl across our 50/50 chimeras to prevent over- or underestimation of our KD-related effects (Jafri et al., 2017).

For all experiments, cells were transplanted into fully irradiated (2x5 Gy) C57Bl6/J mice (CD45.1). Contribution of CD45.2+ donor cells was monitored in peripheral blood approximately 4, 8, 12 and 16 weeks post-transplantation in primary and secondary recipients. Outcome was analyzed by flow cytometry using the monoclonal antibodies indicated in Table 5.27. For endpoint analysis of chimeras, BM and thymus stainings were performed (Table 5.27).

For secondary transplantations, whole BM was isolated 16 weeks post-transplantation, and 1x10^6 cells were retransplanted into fully irradiated (2x5 Gy) C57Bl6/J mice (CD45.1).

Mice were only included in the 50/50 chimera analysis if >50 GFP+ cells were detected (>50 GFP+ single cells for analysis of peripheral blood and peripheral organ/BM stainings; >50 GFP+ cells of respective population or LSK cells for BM analysis). CD45-negative cells were excluded from the analysis.

Homing analysis

For analysis of homing capacity, 30,000 GFP+ cells (KD or Scr Ctrl, CD45.2) were sorted 48 h after lentiviral transduction of LSK cells per recipient mouse. Cells were transplanted into fully irradiated (2x5 Gy) C57Bl6/J mice (CD45.1) and sacrificed 48 h later. HSC/MPP and CD45.1/.2 staining was performed (Table 5.27), and engraftment was analyzed by flow cytometry analysis.

Table 5.27: Antibody combinations used for flow cytometry stainings for *Pabpn1* KD experiments in results part 3.2

Staining	Surface marker	Antibody clone (fluorophore)
Differentiated Cells Transplants (PB)	CD4	GK1.5 (PE)
	CD8a	53-6.7 (PE)
	CD11b	M1/70 (AF700)
	Gr1	RB6-8C5 (APC)
	B220	RA3-6B2 (PE-Cy5)
	CD45.1	A28 (PE-Cy7)
	CD45.2	104 (PB)
Thymus Transplants	CD4	GK1.5 (APC-Cy7)
	CD8a	53-6.7 (PE)
	CD25	PC61 (APC)
	CD44	IM7 (PE-Cy5)
	CD45.1	A28 (PE-Cy7)
	CD45.2	104 (PB)
HSCs/MPP Cells Transplants	CD4	GK1.5 (APC)
	CD8a	53-6.7 (APC)
	CD11b	M1/70 (APC)
	Gr1	RB6-8C5 (APC)
	Ter119	Ter-119 (APC)
	B220	RA3-6B2 (APC)
	c-Kit	2B8 (BV-711)
	Sca-1	D7 (APC-Cy7)
	CD48	HM48-1 (PE-Cy7)
	CD150	TC15-12F12.2 (PE-Cy5)
	CD34	RAM34 (AF700)
	CD45.1	A28 (PE)
	CD45.2	104 (PB)
Homing Transplants	Lin (CD4, CD8a, CD11b, Gr1, Ter119, B220)	Clones see HSCs/MPP cells transplants, all APC
	CD117	2B8 (BV-711)
	Sca-1	D7 (APC-Cy7)
	CD48	HM48-1 (PE-Cy7)
	CD150	TC15-12F12.2 (PE-Cy5)
	CD34	RAM34 (AF700)
	CD45.1	A28 (PE)
	CD45.2	104 (PB)

5.2.2.7 3'-Seq library generation and bioinformatic analysis

Library and data generation

Total RNA isolation was performed using the ARCTURUS PicoPure RNA isolation kit (according to the manufacturer's instructions). Total RNA was used for quality control and for normalization of the starting material. RNA quality was analyzed using the Agilent RNA 6000 Pico Assay according to manufacturer's guidelines. Sequencing libraries were generated with

3 ng of total RNA for HSCs/MPP cells, sHSCs (PBS-treated mice), pHSCs (pIC-treated mice) or Scr-/*Pabpn1*-KD LSK cells (72 h after viral transduction) using the QuantSeq 3' mRNA-seq Library Prep Kit REV according to the manufacturer's instructions except for adaptations (version 015UG009V0221: Do not skip step 2; 15 minute incubation time step 4; cycle number was adjusted to 19 using the PCR Add-on Kit for Illumina). Quality was checked using the Agilent High Sensitivity DNA Kit according to manufacturer's instructions. A multiplex of 2 nM was generated, and sequencing was performed with an Illumina HiSeq2000 device in paired-end mode reading 125 nucleotides. For sequencing, a custom sequencing primer was provided (5' CCCTACACGACGCTCTTCCGATCTTTTTTTTTTTTTTTTTTT 3')

Downstream analysis

Libraries were demultiplexed according to the barcodes, and quality control of the fastq files was performed using the fastqc package.

For regular RNA-seq gene expression analysis, paired-end libraries were mapped against the mm10 mouse genome with the STAR package (Dobin et al., 2013) to generate sorted BAM files. Reads were annotated and counted against the GRCm38.p5 reference gene set (Ensembl) using the htseq-count package (Anders et al., 2015). Differentially expressed genes were identified with the DESeq2 package (Love et al., 2014) using the independent hypothesis weighting function of the IHW package for false discovery rate (FDR) calculation (Ignatiadis et al., 2016). Genes with FDR <0.1 and with log2 fold change higher than 0.5 or lower than -0.5 were considered as differentially expressed.

For APA analysis, we utilized only the Read1 of the two paired ends, since it originates from the custom sequencing primer of the Lexogen REV kit corresponding to the terminal end of the 3'-UTR. Sequencing adapters and polyadenylation tails were removed by trimming, and sequences containing less than 80% of A and T nucleotides were mapped against the mm10 mouse genome using bowtie. Failed reads and reads with multiple hits to the genome were filtered out. Putative polyadenylation sites were clustered in a 50 bp half-window, and only polyadenylation sites present in all biological replicates for both conditions and with a minimum of 25 hits were considered for further analysis. Polyadenylation sites were annotated with the ChIPpeakAnno package (Zhu et al., 2010) using as reference the protein-coding genes reported in the GRCm38.p5 reference set. The flanking sequences were obtained with the BSgenome package (Bioconductor). High-confidence polyadenylation sites were identified as hits located within a gene or in the 3,000 bases downstream in the intergenic region that show at least one of the 12 identified consensus polyadenylation hexamers (Beaudoing et al., 2000) in the 40 bases upstream and that do not have more than seven As in the 10 bases

downstream. For each polyadenylation site, we calculated the Percentage of PolyAdenylation Usage (PPAU), defined as the ratio of its usage against the sum of the usage of all the polyadenylation sites for that specific transcript. ΔAPAS was identified using the DEXSeq package (Anders et al., 2012) using a FDR <20%.

For analysis of APA shortening/lengthening, we considered only hits present in 3'-UTRs. For each 3'-UTR containing two or more polyadenylation sites, we considered the percentage of usage of the most distal polyadenylation site (PDUI) and calculated its differential usage between the conditions (dPDUI).

GOterms (biological processes) were calculated for genes with a relative shortening or lengthening of the distal pA higher than 10% by using the topGO package (Bioconductor) with the classic Fisher statistics and using all the genes with at least one polyadenylation site in the 3'-UTR as background. To visualize and summarize GO terms, REVIGO analysis (http://revigo.irb.hr/; GO terms of shortened genes MPP2 vs HSC classicFisher < 0.05; Allowed similarity: Tiny; p-values were provided; database: whole UniProt; semantic similarity measure: SimRel) was performed (Supek et al., 2011).

Data availability

The accession number for the 3'-Seq data of HSCs and MPP1-4 cells reported in this thesis is ArrayExpress: E-MTAB-7391. The accession number for the 3'-Seq data of sHSC and pHSCs reported in this thesis is ArrayExpress: E-MTAB-7390. The accession number for the 3'-Seq data of LSK *Pabpn1* KD cells is ArrayExpress: E-MTAB-7392.

5.2.2.8 qPCR analysis

For real-time qPCR, total RNA of 5,000-10,000 cells was isolated as described (see also 5.2.1.5) and reverse-transcribed using the SuperScript VILO cDNA Synthesis Kit according to the manufacturer's guidelines (see also 5.2.1.5). For qPCR analysis, Fast SYBR Green Master Mix (see Table 5.23 for qPCR program) or TaqMan Probes and TaqMan Fast Advanced Master Mix (see Table 5.28 for TaqMan qPCR program) were used on a ViiA 7 Real-Time PCR System. RNA expression was normalized to *Oaz1* or *Gapdh* housekeeping gene expression and presented as relative quantification (ratio = 2^-DDCT). For analysis of APA switches by qPCR, the 2^-DCT value was calculated. The proximal 2^-DCT value was divided by the distal 2^-DCT value, and results were normalized. For comparison to 3'-Seq data, the proximal raw read counts were added up (proximal + distal) and divided by the distal raw read counts, followed by normalization. Adding of proximal and distal counts of the 3'-Seq data is necessary for comparison with qPCR results, as the proximal primer pair always amplifies the

distal transcript as well. Results >1 indicate relative increase of the proximal transcript, whereas results <1 indicate relative increase of the distal transcript. Primers were designed using the Universal ProbeLibrary Assay Design Center (Roche) or NCBI Primer-BLAST (NCBI). TaqMan Probes were purchased from Thermo Scientific (Table 5.6). For list of primers for SYBR Green qPCR, see Table 5.29.

Table 5.28: TaqMan qPCR thermal cycling program

Step	Temperature	Time	Number of cycles
UNG* incubation	50°C	2 min	
Activation of polymerase	95°C	20 s	
Amplification	95°C 60°C	1 s 20 s	40

*UNG: uracil-N-glycosylase; degradation of previous PCR amplifications or mis-primed, non-specific products, leaving native nucleic acid templates intended for amplification intact

Table 5.29: qPCR primer for results part 3.2 (synthesis by Sigma-Aldrich)

Primer	Forward sequence	Reverse sequence
proximal *CD84*	GCCAGCCAAGGGATTTAGGA	AGATCTGTGGGCCTGAGAGT
distal *CD84*	GTAAGGCAGTGGCCTAGCTT	AAGCCTTCACTGGTGTTCCA
proximal *Slc25a20*	TGGCTGTGATGGTGAAACTGT	CTCCAATTCTCTGGAAATCCTCT
distal *Slc25a20*	GTCCCAGCCTCCTGGGATAA	TGAGTGCTCTCATGGGGAAA
proximal *Minos1*	ACAAACAGTGGCTGGAGTTGA	AATACTGAAGTGCCCGGAGT
distal *Minos1*	TCAGATTGAACCATGACAAGCA	TAGGAAACAGACACTGGGCT
proximal *Rcc2*	GCTCCGTTTCCAAGCGATTC	AAAGGCAGGCTGGCTATGAG
distal *Rcc2*	CAAAGGGTAACAGGAAGGGAGT	AACAGCACTAAACCAGGCAC
proximal *Ogfrl1*	TTTTCTGGGCTTTGCTTCGC	GAGAGGAGAAAGCTCCTGCG
distal *Ogfrl1*	TGTCATGTTACGAGGACGCA	AGAGGAGTACACCCTGGCTT
proximal *Hnrnpa2b*	AGCTTCTTAACTCTACACACGCA	AGCCAGGATCATGGTGTAATAAGA
distal *Hnrnpa2b*	TGATGTTACCCTTTACAATCAGCA	AGTTCTTGTTGGTGCCTTCA
proximal *Hdlbp*	GTGCTGACTACATGCTGGGT	CTCCTGATGCTGTGCTCTGG
distal *Hdlbp*	TCCACCCTCAACTCAATGGC	TCATGTGGGTTAGTGCCACC
proximal *Treml2*	ACAGAGCATGGAAAGGTCCC	CAAGGCCCTAGCTGTGTCAG
distal *Treml2*	CTCAGGGGACTCTAGGCTGT	GCCATCATCACAGAACCCCA

proximal *SSR1*	ACGCAGGCAGAATATGGCTT	CCTAGCGACTTCACTTGGCA
distal *SSR1*	TTGTGACCATGACCCGCAAA	GGGATGGAGCTCAGACGTAG
proximal *Ctsc*	AGGACTTCACTGGGGCTAAGA	CCTCTGGGCTGGAATGTACC
distal *Ctsc*	TGCCTTTGAAGTCCACGATGA	TTTGTCAGCTCGAAGGGGTT
Oaz1	TTTCAGCTAGCATCCTGTACTCC	GACCCTGGTCTTGTCGTTAGA
Pabpn1	CGGAGCTAGAAGCGATCAAAG	ATTGCCAACGTAGATAGAGCG

5.2.2.9 CB-839 in vitro assays

Enrichment of HSPCs

To enrich HSPCs for *in vitro* assays, lineage marker-positive cells were depleted from BM as described in section 5.2.1.4. For bone marrow preparation, see 5.2.2.3.

HSPC plating and treatment

HSPCs were cultured in Complete Stem Cell Medium (StemPro-34 SFM, Life Technologies containing 50 ng/ml SCF, 25 ng/ml TPO, 30 ng/ml Flt3-Ligand (all Preprotech), 100 u/ml penicillin/streptomycin, 2 mM L-glutamine, StemPro-34 Supplement as recommended). Cells were cultured in 96-/24-/6-well ultra-low attachment plates and were treated with either CB-839 (5 μM) or the respective amount of DMSO. After 48 h or 72 h of culture, cells were used for cell cycle and viability analysis (see also 5.2.2.3).

CFUs of CB-839-treated cells

1,000 FACS-sorted LSK cells were cultured in 100 μl Complete Stem Cell Medium (described in “HSPCs plating and treatment”) in 96-well ultra-low attachment plates and immediately treated with CB-839 (5 μM) or the respective amount of DMSO. 72 h after treatment, all cells of one well were harvested and cultured in MethoCult M3434 (see also 5.2.2.5 for more details). Five days after plating of CFUs, colony formation was quantified. 20,000 cells of each replicate were used for replating and cultured in MethoCult M3434 (StemCell Technologies). 6-7 days after replating, colonies were quantified.

5.2.2.10 Statistical analysis

Statistical analysis was performed by unpaired or paired Student’s t-test or two- or one-way ANOVA without correction for multiple comparison (Fisher LSD test). All data are presented as mean + SD or ± 95% confidence interval. Please see figure legends for detailed information.

Significance levels were set at *$p < 0.05$, **$p < 0.01$ and ***$p < 0.001$. For statistical analysis, GraphPad Prism was used.

5.2.3 Methods – results part 3.3

5.2.3.1 Animals

C57BL/6J and generated *Pabpn1*^{flox} mice were bred in-house in the animal facility of the DKFZ under specific-pathogen-free (SPF) conditions in individually ventilated cages (IVCs). All animal procedures were performed according to protocols approved by the German authorities, Regierungspräsidium Karlsruhe (Nr. A-23/17, Z110/02, DKFZ299, G-256/16).

5.2.3.2 Testing of gRNA efficiency

Cloning of the px330 Cas9 vector

Oligos encoding the target-specific part of the crRNA were ordered for expression of full-length gRNAs (Table 5.30, Sigma-Aldrich; design gRNA A: http://crispor.tefor.net/crispor.py?batchId=aQ9aKiq6R1spZW5ypsBG; design gRNA B: http://crispor.tefor.net/crispor.py?batchId=idFOo1VhuwIGjPTeLMe3) in the px330 Cas9 vector. The respective oligos were annealed for 2 h (starting at 95°C continuously cooled to RT). The px330 Cas9 vector was digested with BbsI enzyme for 2 h at 37°C. Annealed oligos and the digested px330 Cas9 vector were purified by agarose gel electrophoresis (2% for oligos; 1% for vector). Gels were prepared by boiling a suspension of 1-2% agarose in 1x TBE buffer. After addition of 30 µg ethidium bromide per 100 ml, the solution was filled in a gel tray containing a comb. Before loading, PCR products were mixed with loading buffer containing bromophenol blue. PCR fragments were separated by applying 120 V to the gel. Finally, DNA fragments were visualized using an UV transilluminator, and gels were documented using a camera system. Subsequently, gel extraction using the QIAEX II Gel Extraction Kit according to manufacturer's guidelines (elution in 20 µl) was performed. 100 ng digested vector were ligated with 1.2 ng annealed oligos for 1.5 h at 22°C using T4 ligase. Bacterial (E. coli) transformation was performed with the ligated px330 Cas9 vectors encoding for Cas9 and the inserted oligos, which in the vector are fused to the residual gRNA sequence, leading to expression of a full-length single-molecule gRNA. Transformed bacteria were plated on agar plates (containing 100 µg/ml ampicillin) and incubated overnight at 37°C. The next day, 5 colonies were picked per construct, and colony PCR (Table 5.31) was performed using the amplification program indicated in Table 5.32. Successful integration of target sequence oligos in the px330 Cas9 vector leads to a product of approximately 500 bp. Presence of this product was verified by agarose gel electrophoresis (1%). Positive clones were used for overnight culture and subsequent DNA plasmid Mini-prep using the Macherey Nagel NucleoSpin Plasmid Kit according to manufacturer's protocols. Enzymatic digestion was used to test for oligo integration using the enzymes BbsI and EcoRI. Re-ligation of px330 Cas9 vector without

oligo integration leads to 2 products after the digestion step (3.2 kb and 5.3 kb). Successful oligo integration leads to abrogation of the BbsI restriction enzyme sites. Enzymatic digestion therefore leads to generation of only one 8.5 kb product. Digestion was performed for 2 h at 37°C. Product size was analyzed by agarose gel electrophoresis (1%).

Table 5.30: gRNA and reporter oligos

Oligo	Forward sequence	Reverse sequence
gRNA A1	CACCgagctgataatcgccccaaga	AAACtcttggggcgattatcagctC
gRNA A2	CACCgttcttggggcgattatcagc	AAACgctgataatcgccccaagaaC
gRNA A3	CACCgttcctaagtcctaactaaa	AAACtttagttaggacttaggaac
gRNA B1	CACCgttaaactgggattagacttc	AAACgaagtctaatcccagtttaaC
gRNA B2	CACCgtgactggcctaacttgagct	AAACagctcaagttaggccagtcaC
gRNA B3	CACCggcctaacttgagctgggcg	AAACcgcccagctcaagttaggcc
Reporter A	cgGTCTGTACCCTTCTTGGGGC GATTATCAGCTGGCATACgtaTG GAGACCCTTTAGTTAGGACTTA GGAACTGACTG	CAGTCAGTTCCTAAGTCCTAACTAA AGGGTCTCCAtacGTATGCCAGCTG ATAATCGCCCCAAGAAGGGTACAG AC
Reporter B	cgCTCAGAATTAAACTGGGATTA GACTTCAGGTTTATACCACCAT GACTGGCCTAACTTGAGCTGGG CGTGGTGGaattcg	cgaattCCACCACGCCCAGCTCAAGT TAGGCCAGTCATGGTGGTATAAAC CTGAAGTCTAATCCCAGTTTAATTC TGAG

Table 5.31: Colony PCR px330 Cas9 vector, Fisher BioReagents™ *Taq* DNA Polymerase, Buffer B and $MgCl_2$ provided with polymerase

Component	Amount (µl)
Buffer B	5
$MgCl_2$ (25 mM)	5
Fwd primer (respective fwd gRNA oligo)	0.3
Rev primer (px330_rev: gtactgccaagtaggaaagtcc)	0.3
dNTPs (10 mM)	0.5
TaqPol	0.5
H_2O	38.4

Table 5.32: Colony PCR amplification program

Step	Temperature	Time	Number of cycles
Denaturation	95°C	5 min	
Amplification	95°C 62°C 72°C	20 s 30 s 30 s	30
Final elongation	72°C	2 min	
Storage	4°C	∞	

Cloning of the pCRISPR-Report vector

In addition, oligos encoding for targeting sequences (Table 5.30, Sigma-Aldrich) were ordered for cloning into the pCRISPR-Report vector for gRNA testing. These oligos were annealed and purified by gel extraction as previously described (see “Cloning of the px330 Cas9 vector”). The pCRISPR-Report vector was digested using BstBI and NruI and purified by gel extraction as previously described (see “Cloning of the px330 Cas9 vector”). 100 ng digested vector were ligated with 1.2 ng annealed oligos for 1.5 h at 22°C using T4 ligase. Bacterial (E. coli) transformation was performed using the ligated pCRISPR-Report vectors encoding for lacZ fragments separated by targeting sequences that are equivalent to the mouse *Pabpn1* targeting sequences for upstream gRNAs A and downstream gRNAs B. Transformed bacteria were plated on agar plates (containing 100 µg/ml ampicillin) and incubated overnight at 37°C. The next day, 5 colonies were picked per construct, and colony PCR (Table 5.33) was performed using the amplification program indicated in Table 5.32. Successful integration of oligos in the the pCRISPR-Report vector leads to a product of approximately 900 bp length. Presence of this product was verified by agarose gel electrophoresis (1%). Positive clones were used for overnight culture and subsequent DNA plasmid Mini-prep using the Macherey Nagel NucleoSpin Plasmid Kit according to manufacturer’s protocols. Enzymatic digestion was used to test for oligo integration using the enzymes SnaBI (pCRISPR-Report for testing of A gRNAs) or EcoRI (pCRISPR-Report for testing of B gRNAs). Re-ligation of pCRISPR-Report vector without oligo integration leads to 1 product after the digestion step (7.4 kb). Successful oligo integration leads to presence of the respective restriction enzyme sites. Enzymatic digestion leads to generation of two products (A: 0.8 kb + 6.6 kb; B: 1.2 kb + 6.2 kb). Digestion was performed for 2 h at 37°C. Product size was analyzed by agarose gel electrophoresis (1%). Positive clones were selected and overnight culture and subsequent DNA plasmid Midi-prep using the Macherey Nagel NucleoBond Xtra Midi Plasmid Kit according to manufacturer’s protocols was performed.

Table 5.33: Colony PCR pCRISPR-Report vector, Fisher BioReagents™ *Taq* DNA Polymerase, Buffer B and $MgCl_2$ provided with polymerase

Component	**Amount (µl)**
Buffer B	5
$MgCl_2$ (25 mM)	5
Fwd primer (respective fwd reporter oligo)	0.3
Rev primer (lacZ_rev: ttacccgtaggtagtcacgca)	0.3
dNTPs (10 mM)	0.5
TaqPol	0.5
H_2O	38.4

Reporter assay

25,000 HeLa cells/well were seeded in 12-well plates in DMEM GlutaMAX (+10% FCS, +1% Pen/Strep). A total of 27 wells were seeded (3 wells/condition, 6 gRNAs, 2 negative controls, 1 positive control → 9*3 = 27). Cells were incubated for 2-3 h. Transfection mix was set up (Table 5.34) and mixed with 100 µl RotiFect mix (5 µl RotiFect in 100 µl Opti-MEM I) and incubated for 30 min at RT. 200 µl transfection mix were add per well, and cells were incubated overnight at 37°C. The next day, cells were washed with PBS and lysed using 1x passive lysis buffer in the wells for 3 min at RT. Plates were put on ice, and lysates were transferred to Eppendorf tubes. Lysates were centrifuged at 13,000 rpm for 5 min at 4°C. 30 µl supernatant per replicate and condition were transferred to a fresh Eppendorf tube, and 700 µl Z-buffer (0.35% 2-Mercaptoethanol) were added and mixed. Subsequently, 200 µl ONPG solution were added per tube and incubated at RT until a yellow color-shift was observed. Reaction was stopped using 500 µl 1 M NaOH per tube, and the signal was measured (420 nm). For luciferase measurements, 10 µl of lysate per replicate and condition were pipetted into a white 96-well plate. This plate was measured in a luciferase detection instrument (Perkin Elmer Wallac 1420 Victor2 Microplate Reader). To calculate normalized signals, the following equation was used: LacZ/Luciferase*mean(Luciferase).

Table 5.34: Transfection mix

Component	Amount
pUHC131.1 (Luciferase Plasmid)	200 ng
px330 Cas9	400 ng
pCRISPR-Report	400 ng
Opti-MEM I	Fill up to 100 µl

5.2.3.3 Easi-CRISPR

crRNA, tracrRNA and ssDNA template molecules were ordered from IDT (Table 5.35). crRNA and tracrRNA were reconstituted in microinjection buffer to a final concentration of 100 µM. Cas9 protein was diluted in microinjection buffer to a final concentration of 500 ng/µl. The ssDNA template was reconstituted in 30 µl H_2O. All stocks were aliquoted and stored at -80°C. crRNA and tracrRNA were diluted to 6.1 µM, and 15 µl crRNA + 15 µl tracrRNA were mixed and incubated for 2 min at 94°C and 10 min at RT. The generated 3.1 µM crRNA:tracrRNA mixes were aliquoted and stored at -80°C. For pronucleus injections, 1 µl crRNA_A3:tracrRNA (final 0.31 µM) was mixed with 1 µl crRNA_B2:tracrRNA (final 0.31 µM), 1 µl ssDNA template (final 10 ng/µl), 1 µl Cas9 protein (final 50 ng/µl) and 6 µl microinjection buffer. Injections were performed by the Transgenic Service of the DKFZ.

Table 5.35: IDT constructs

Construct	Sequence
crRNA A3	GTTCCTAAGTCCTAACTAAA
crRNA A3	TGACTGGCCTAACTTGAGCT
ssDNA megamer	GGCGATTATCAGCTGGCATTTGACCTTCAAGTCTTAACACTTTTCAGTT GGAGACCCTTTATAACTTCGTATAGCATACATTATACGAAGTTATaAGcT TAGGACTTAGGAACTGACTGGTGGCTGTAAGGGAGGGAGTTCTTTAAA AATGTTAGATTTCAGTCCATTCTGTTTTGTTTGCTTTGAGGCATTTCAGC TAAAGCTGAATAGAGACTCGACTTGAAAAACGTAAAATCTTAACACCGT TTTTTATGGTCTTTTAAGAGGTAGTTTCTCTAATGTATATTGAAACCTAC TATGTGCTAAACACTGTTATAGGTAATTGGTATAGTAGAAGTATACCCA AAATTTAGCCTGCGGGTGGAAGTGTGTGTGTATTGATCGAGCAGAAGA TGTGCTTGCCTAAGGTAATTTTTTTCCCCTCTAAAAGCTGGCCCAGTGA TCATGTCTCTTGAGGAGAAGATGGAGGCTGATGCCCGCTCTATCTACG TTGGCAATGTATGTATCAAGGAATAGTCTGCTTTGGGTTTTGGGGGGG ACTTTTGTTTCTTATTCCTTACTTTGTTATGTGTTCAAATTGGAGTTTCC CCAGTACTTGTCAAGTGTTCCTTTCCTTAGGTGGACTATGGTGCAACA GCAGAAGAGCTGGAAGCCCATTTTCATGGCTGTGGTTCAGTCAACCGT GTTACTATACTCTGTGACAAATTTAGTGGCCATCCCAAAGGGTAAGTAG GAGGATAAGTTGAGATCATTTTACTCACATTTTAAAAATACGTGAAAAAT ACATGAGCTCGGAATCGAACCTAGGACCTTGAACATGCTGATGTGTGC TGTCAATCTAGTCAAAAGAGTGTTAACTCGTTTTAAAGGTTCTTATATAT GTGGTCGGTTGGTTTTGTTTTCAAGTCAGAGTTTCTCTGTTTAACCCTG GCTGTACTGGAATTACCTCTATAGACCAGGCTGGCCTTAAACTCAGAA TTAAACTGGGATTAGACTTCAGGTTTATACCACCATGACTGGCtCTAgA CTTGAATAACTTCGTATAGCATACATTATACGAAGTTATGCTGGGCGTG GTGGTGCACGCCTTTAATCCCAGCACTTGGGAGACAGAGGCAGGGAT TTC

5.2.3.4 Genotyping

For genotyping of offspring, tails were digested, and tail DNA was purified using the Qiagen DNeasy Blood and Tissue Kit following manufacturer's instructions (elution in 100 μl). 5 different PCR and digestion approaches were used to validate loxP integration (Table 5.36). DreamTaq Green PCR Mastermix was used for PCR amplification (Table 5.37, Table 5.38, Table 5.39). PCR products were analyzed by agarose gel electrophoresis (Table 5.40) and purified using the QIAquick PCR Purification Kit according to manufacturer's guidelines. Enzyme digestion was performed at 37°C for 1.5 h and analyzed using agarose gel electrophoresis (Table 5.41). For mice exhibiting correct integration, PCR 2 and PCR 4 were repeated, and agarose gel electrophoresis was performed. The loxP band was excised and purified using the Qiagen QIAquick Gel Extraction Kit according to manufacturer's guidelines. Purified PCR products were sent for sequencing by Eurofins together with the respective rev primer (PCR 2) and fwd primer (PCR 4) to exclude presence of point mutations.

Table 5.36: Overview genotyping PCRs and digestion approaches (P: program)

PCR	Validation	Forward primer	Reverse primer	P	Enzyme
PCR 1	Integration upper loxP	CGATTATCAGCT GGCATTTGACC	GGCAAGCACATC TTCTGCTCG	P 1	HindIII-HF
PCR 2	Integration lower loxP	ACCTTGAACATG CTGATGTGTGC	GAGACACTTAACA GGCAAAACCTAC	P 1	XmnI
PCR 3	Upper and lower loxP on same allele	CGATTATCAGCT GGCATTTGACC	GAGACACTTAACA GGCAAAACCTAC	P 2	HindIII-HF + XmnI
PCR 4	Upstream localization correct	AAGCCGGGGAC CTTGAATG	GGCAAGCACATCT TCTGCTCG	P 2	HindIII-HF
PCR 5	Downstream localization correct	ACCTTGAACATG CTGATGTGTGC	ACTCTCACTGCCA CCAACTCC	P 2	XmnI

Table 5.37: Genotyping PCR mastermix

Component	Amount
DreamTaq Green PCR Mastermix	25 µl
Fwd primer (100 µM)	0.3 µl
Rev primer (100 µM)	0.3 µl
DNA	100 ng
H_2O	up to 50 µl

Table 5.38: PCR program 1 (P 1)

Step	Temperature	Time	Number of cycles
Denaturation	95°C	2 min	
Amplification	95°C 62°C 72°C	30 s 30 s 45 s	35
Final elongation	72°C	5 min	
Storage	4°C	∞	

Table 5.39: PCR program 2 (P 2)

Step	Temperature	Time	Number of cycles
Denaturation	95°C	2 min	
Amplification	95°C 62°C 72°C	30 s 30 s 1 min	35
Final elongation	72°C	5 min	
Storage	4°C	∞	

Table 5.40: Expected amplicon sizes

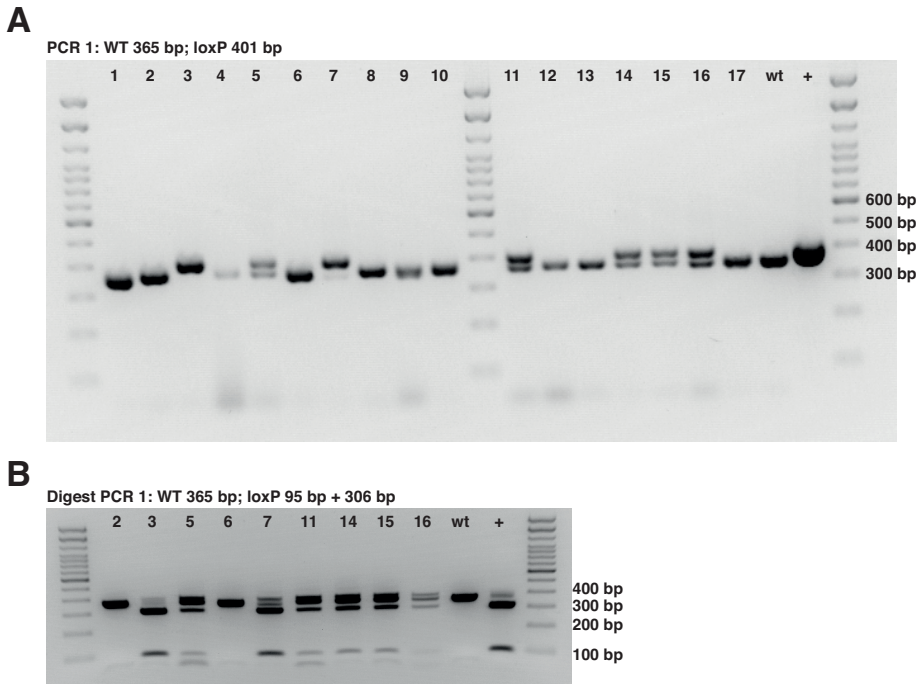
PCR	Wildtype (bp)	loxP (bp)
PCR 1	365	401
PCR 2	435	471
PCR 3	1206	1278
PCR 4	562	598
PCR 5	1477	1513

Table 5.41: Expected product sizes after digestion

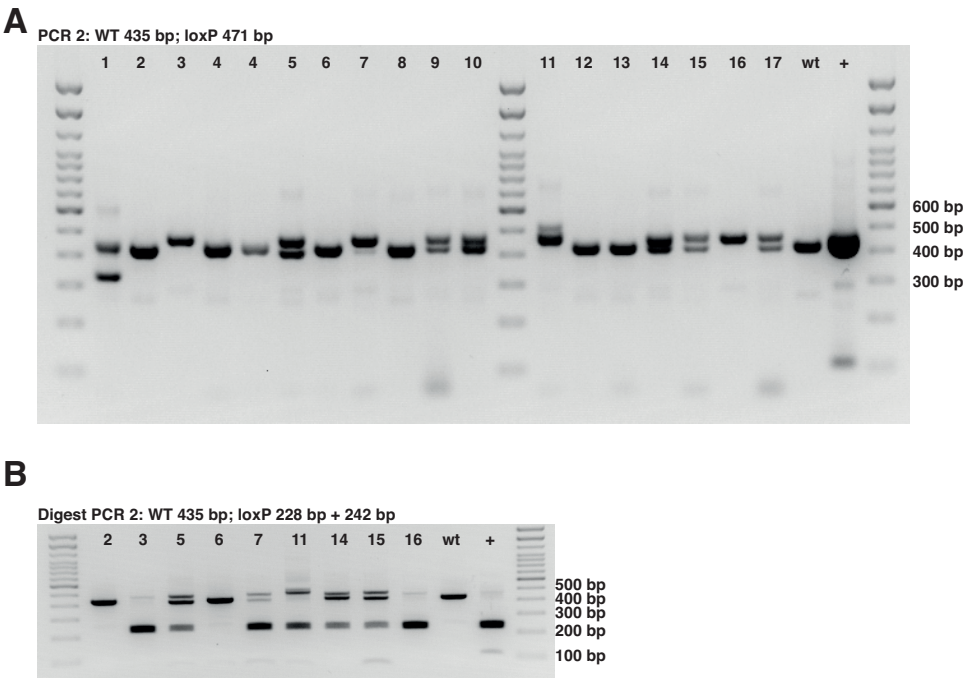
PCR	Wildtype (bp)	loxP (bp)
PCR 1	365	95 + 306
PCR 2	435	228 + 242
PCR 3	1206	Both alleles: 93 + 940 + 244 Upstream only: 93 + 1184 Downstream only: 1033 + 244
PCR 4	562	254 + 308
PCR 5	1477	231 + 1282

6 Appendix

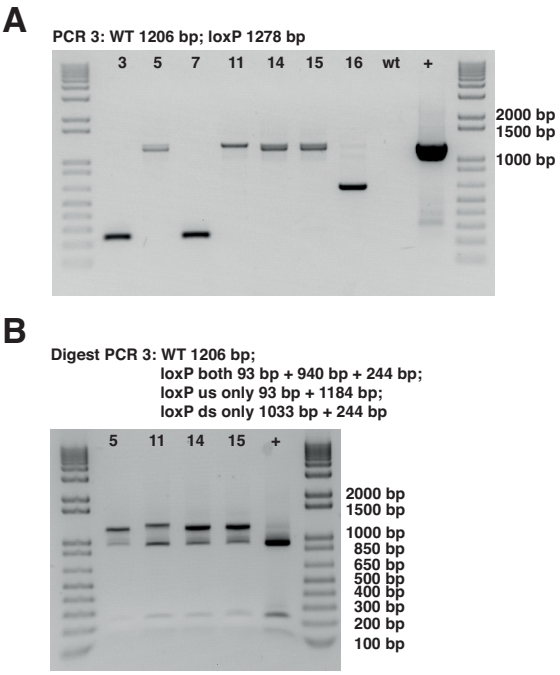
6.1 Appendix figures



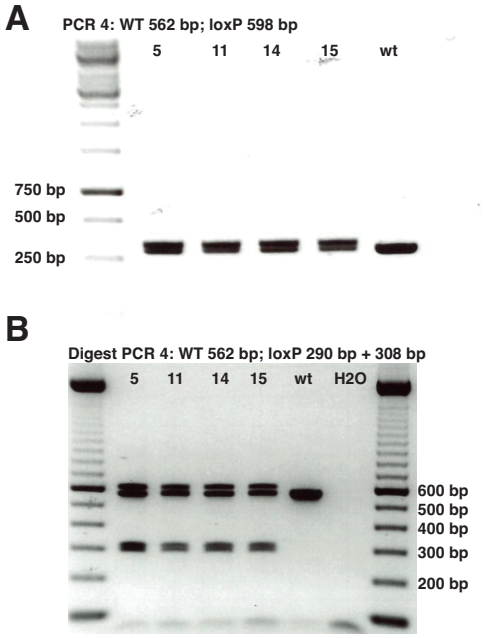
Appendix Figure 1: Gel images showing genotyping results for PCR 1 (A) and digestion of PCR 1 (B).



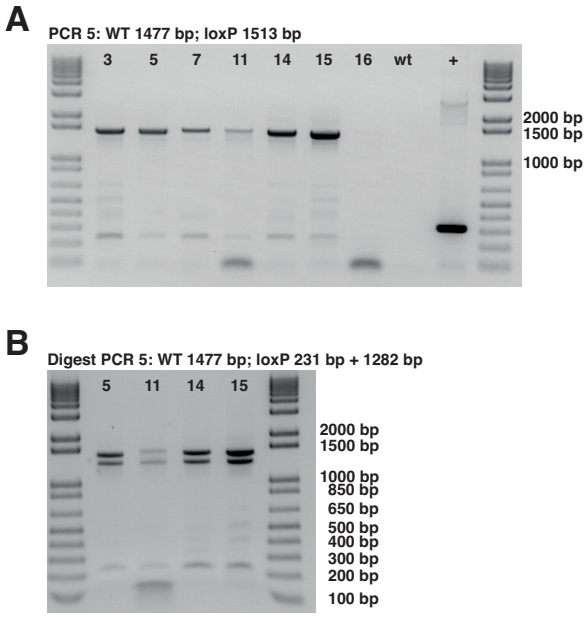
Appendix Figure 2: Gel images showing genotyping results for PCR 2 (A) and digestion of PCR 2 (B).



Appendix Figure 3: Gel images showing genotyping results for PCR 3 (A) and digestion of PCR 3 (B).



Appendix Figure 4: Gel images showing genotyping results for PCR 4 (A) and digestion of PCR 4 (B).



Appendix Figure 5: Gel images showing genotyping results for PCR 5 (A) and digestion of PCR 5 (B).

6.2 List of abbreviations

Abbreviation	Explanation
%	percent
°C	degrees Celsius
AA	amino acid
αKG	α-ketoglutarate
AGM	aorta-gonad-mesonephros
AGO	argonaute
aHSC	active HSC
AML	acute myeloid leukemia
APA	alternative polyadenylation
APC	allophycocyanin
BM	bone marrow
c-Kit	v-kit Hardy-Zuckerman 4 feline sarcoma viral oncogene homolog
c-Myc	cellular myelocytomatosis homologue
CAR cells	CXC-chemokine ligand 12-abundant reticular cells
CD	cluster of differentiation
CDK	cycline-dependent kinase
CFIIm	cleavage factor IIm
CFIm	cleavage factor Im
CFU	colony forming unit
CFU-S	spleen colony forming unit
Chr	chromosome
circRNAs	circular RNAs
CLL	chronic lymphocytic leukemia
CMP	common myeloid progenitor
CpG	cytosine phosphodiester bond guanine
CPSF	cleavage and polyadenylation specificity factor
CR-APA	coding region APA
CRE	cAMP response elements
CREB	cAMP response element binding protein
CRISPR	clustered, regularly interspaced, short palindromic repeat
crRNA	CRISPR RNA
CstF	cleavage stimulatory factor
Ct	cycle threshold
Ctsc	CathepsinC
CXCL12	CXC-chemokine ligand 12
CXCR4	CXC-chemokine receptor 4
Cy5	cyanine 5
Cy7	cyanine 7
ΔAPAS	differentially used APA sites
DC	dendritic cell

dCas9	dead-Cas9
DEG	differentially expressed gene
dHSC	dormant HSC
DNA	deoxyribonucleic acid
Dnase	deoxyribonuclease
dPDUI	differential polyA site usage index
DSB	double-strand break
Easi-CRISPR	efficient additions with ssDNA inserts-CRISPR
ESC	embryonic stem cell
EST	expressed sequence tag
FACS	fluorescence-activated cell sorting
FASN	fatty acid synthase
FAO	fatty acid oxidation
FOXO	forkhead box protein O
g	gram
GFP	green fluorescent protein
Gln	glutamine
Gls	glutaminase
Glu	glutamate
GMP	granulocyte-macrophage progenitor
GO	gene ontology
gRNA	guide RNA
Gy	Gray
h	hour(s)
HDR	homology-directed repair
HIF	hypoxia-inducible factor
HSC	hematopoietic stem cell
HSPC	hematopoietic stem and multipotent progenitor cell
ICR	imprinting control region
IFN	interferon
IG-DMR	intergenic differentially methylated region
IGV	integrative genomics viewer
indel	insertion-deletion
iPSC	induced pluripotent stem cell
IVC	individually ventilated cage
kb	kilo base pairs
KD	knockdown
KO	knockout
LacZ	β-galactosidase
LDH	lactate dehydrogenase
Lin	lineage
lncRNA	long non-coding RNA
LPS	lipopolysaccharide

LS-K	Lin+ Sca-1- c-Kit+ cells
LSK	Lin- c-Kit+ Sca-1+ cells
LSK SLAM	LSK CD150+ CD48- cells
LT	long-term
MDS	myelodysplastic syndromes
Meg3	Maternally expressed Gene 3
MEP	megakaryocyte-erythrocyte progenitor
mg	milligram
µg	microgram
miRNA	microRNA
min	minute(s)
Mk	megakaryocyte
ml	milliliter
µl	microliter
MM	multiple myeloma
mM	millimolar
MPP	multipotent progenitor
ncRNA	non-coding RNA
NK	natural killer
nm	nanometer
nt	nucleotide
ORF	open reading frame
OXPHOS	oxidative phosphorylation
Pabpn1	poly(A) binding protein 1
PAM	protospacer adjacent motif
PAP	poly(A) polymerase
PAS	polyadenylation signal
PB	peripheral blood
PBS	phosphate-buffered saline
PCA	principal component analysis
PCR	polymerase chain reaction
PDH	pyruvate dehydrogenase
PDK	pyruvate dehydrogenase kinase
PDUI	polyA site usage index
pH	power of hydrogen
pHSC	proliferating HSC
pIC	polyinosinic:polycytidylic acid
piRNA	PIWI-associated RNA
PML	promyelocytic leukemia protein
Pol II	RNA-Polymerase II
polyA	poly adenosine tail
PPAR	peroxisome proliferator-activated receptor
PPAU	percentage of polyadenylation usage

qPCR	quantitative real-time PCR
RIN	RNA integrity number
RISC	RNA-induced silencing complex
RNA	ribonucleic acid
RNA-seq	RNA sequencing
RBP	RNA-binding protein
RNP	ribonucleoproteins
ROS	reactive oxygen species
RT	room temperature
s	second(s)
Sca-1	stem cell antigen 1
scaRNA	cajal-body specific RNAs
SCF	stem cell factor
SD	standard deviation
shRNA	short hairpin RNAs
sHSC	steady-state HSC
siRNA	small interfering RNAs
SLAM	signal lymphocyte activating molecule
snoRNA	small nucleolar RNA
snRNA	small nuclear RNA
SPF	specific-pathogen-free
spliRNA	splice site RNA
ssDNA	single-stranded DNA
TCA	tricarboxylic acid
tiRNAs	transcription initiation RNA
TPO	thrombopoietin
tracrRNA	transactivating CRISPR RNA
Treml2	triggering receptor expressed on myeloid cells-like 2
TRIS	tris(hydroxymethyl)amino-methane
upd	uniparental disomy
UTR-APA	untranslated region APA
UV	ultraviolet
V	volt
VEGF	vascular endothelial growth factor
VHL	von Hippel-Lindau
WT	wildtype

6.3 List of figures

6.4 List of tables

6.5 Talks, poster presentations and publications

Parts of this dissertation have been published (Sommerkamp et al., 2019). I presented parts of the data included in this dissertation as a talk (section 6.5.1) or a poster (6.5.2) at the following conferences, seminars and summer schools. Furthermore, I contributed to the following publications (section 6.5.3).

6.5.1 Talks

Stem Cell Retreat

15.06. – 17.06.2016, in Basel, Switzerland

SFB873 “Maintenance and Differentiation of Stem Cells in Development and Disease” – Young Scientist Retreat

09.05. – 10.05.2017, in Kloster Schöntal, Germany

47th Annual Scientific Meeting of the International Society for Experimental Medicine (ISEH)

23.08. – 26.08.2018, in Los Angeles, USA

(awarded with a travel grant)

10th International Heinrich F.C. Behr-Symposium on “Stem Cells and Cancer”; Lightning Talk

16.09. – 19.09.2018, in Heidelberg, Germany

SFB873 “Maintenance and Differentiation of Stem Cells in Development and Disease” – Young Scientist Retreat

02.04. – 03.04.2019, in St. Martin, Germany

6.5.2 Poster presentations

46th Annual Scientific Meeting of the International Society for Experimental Medicine (ISEH)

26.08.2017, in Frankfurt, Germany

EMBO Workshop RNA: structure meets function

01.07. – 05.07.2018, in Stockholm Archipellago, Sweden

Helmholtz International Graduate School for Cancer Research – PhD Retreat

16.07. – 19.07.2018, in Weil der Stadt, Germany

47th Annual Scientific Meeting of the International Society for Experimental Medicine (ISEH) Pre-Meeting Workshop

23.08.2018, in Los Angeles, USA

(awarded with a price for the best poster)

10th International Heinrich F.C. Behr-Symposium on “Stem Cells and Cancer”

16.09. – 19.09.2018, in Heidelberg, Germany

Helmholtz International Graduate School for Cancer Research – Annual Poster Presentation

16.11.2018, in Heidelberg, Germany

6.5.3 Publications

Sommerkamp, P., Renders, S., Ladel, L., Hotz-Wagenblatt, A., Schonberger, K., Zeisberger, P., Przybylla, A., Sohn, M., Zhou, Y., Klibanski, A.*, et al.* (2019). The long non-coding RNA Meg3 is dispensable for hematopoietic stem cells. Sci Rep *9*, 2110.

Cabezas-Wallscheid, N., Buettner, F., **Sommerkamp, P.**, Klimmeck, D., Ladel, L., Thalheimer, F.B., Pastor-Flores, D., Roma, L.P., Renders, S., Zeisberger, P.*, et al.* (2017). Vitamin A-Retinoic Acid Signaling Regulates Hematopoietic Stem Cell Dormancy. Cell *169*, 807-823 e819.

Sommerkamp, P., and Trumpp, A. (2016). Metabolic cues for hematopoietic stem cells. Science *354*, 1103-1104.

Zhang, M., Hutter, G., Kahn, S.A., Azad, T.D., Gholamin, S., Xu, C.Y., Liu, J., Achrol, A.S., Richard, C., **Sommerkamp, P.***, et al.* (2016). Anti-CD47 Treatment Stimulates Phagocytosis of Glioblastoma by M1 and M2 Polarized Macrophages and Promotes M1 Polarized Macrophages In Vivo. PLoS One *11*, e0153550.

7 Contributions

The completion of this thesis would not have been possible without the help of many people, ranging from technical support to the contribution of ideas. In the following, the contributions of various people to the data presented in this thesis are outlined.

The Klibanski group (Massachusetts General Hospital) generated the *Meg3*^{flox} mouse model and provided it for my research.

Petra Zeisberger, Paula-Leonie Eiben, Luisa Ladel, Katharina Schönberger and Simon Renders provided excellent technical support and help (e.g. mouse work and preparation of cells).

Adriana Przybylla and Markus Sohn provided great technical support (e.g. genotyping (MxCre, *Meg3*^{flox}, MxCre *Meg3*^{flox} and VavCre *Meg3*^{flox} mice; genotyping of *Pabpn1*^{flox} mice was performed by me) and maintenance of mice).

Animal caretakers under the supervision of Anja Rathgeb maintained mice housed in the DKFZ mouse core facility. Animal caretakers and veterinarians performed irradiation of mice in the DKFZ mouse core facility.

Dr. Steffen Schmitt, Dr. Marcus Eich, Klaus Hexel, Tobias Rubner and Florian Blum supported fluorescence-activated cell sorting in the DKFZ FACS core facility.

3'-Seq and Small RNA-seq were performed by the genomics and proteomics core facility of the DKFZ. RNA was isolated and libraries were generated by me. Multiplexed libraries were provided by me.

Small RNA-seq data was analyzed by Dr. Agnes Hotz-Wagenblatt (Core Facility Omics IT and Data Management, DKFZ).

3'-Seq data was analyzed in collaboration with Dr. Sandro Altamura (Department of Pediatric Hematology, Oncology and Immunology, Heidelberg University Medical Center). In detail, bioinformatics analysis was performed by Dr. Sandro Altamura. All results were evaluated and visualized in a collaborative effort by Dr. Sandro Altamura, Dr. Nina Cabezas-Wallscheid (MPI Freiburg), Prof. Dr. Andreas Trumpp and me. For further statistical analysis, approaches were discussed with Dr. Silvia Calderazzo (Division of Biostatistics, DKFZ).

Paula Werner and Franziska Pilz of Dr. Marieke Esser's laboratory (DKFZ) provided cDNA samples of fetal liver HSCs. Dr. Simon Haas (DKFZ) provided cDNA of *in vivo* LPS-treated HSCs.

The targeting strategy for the generation of *Pabpn1*^{flox} mice was developed in a collaborative effort by Dr. Kai Schönig (ZI Mannheim), Franciscus A. van der Hoeven (Transgenic Service, DKFZ), Alexander Sommerkamp (DKFZ), Dr. Nina Cabezas-Wallscheid (MPI Freiburg), Prof. Dr. Andreas Trumpp and me. Alexander Sommerkamp provided excellent technical help (e.g. cloning and gRNA testing). Zygote injections were performed by the Transgenic Service of the DKFZ.

8 Acknowledgements

For me, developing and working on my PhD project was a remarkable experience, giving me countless opportunities to develop as a scientist, but also as a person. This journey would not have been possible without the ongoing support of so many different people.

First of all, I would like to thank you, **Andreas**, for giving me the chance to do my PhD in your lab. You have created a wonderful place at the DKFZ and within HI-STEM, combining the latest scientific technologies with a positive energetic work atmosphere. I am most thankful for the trust and confidence you put in me and for the freedom to follow up on my ideas. I am deeply grateful for your support and all the unique opportunities you gave me.

In addition, I would like to thank **Nina**. You were a wonderful supervisor during my master's thesis and since then you have never stopped supporting me in so many different ways. Thank you for all the scientific input, for all the discussions and the continuing supervision during my PhD. You always find the time to discuss experiments, results and ideas. Thanks for your enthusiasm, your positive energy, your friendship and for always believing in me.

Moreover, I would like to thank the members of my thesis examination committee, **Prof. Dr. Jan Lohmann**, **Prof. Dr. Claudia Scholl**, **PD Dr. Karin Müller-Decker and Prof. Dr. Andreas Trumpp,** for taking the time to read and evaluate my thesis.

Furthermore, I would like to thank **Prof. Dr. Jan Lohmann**, **Prof. Dr. Claudia Scholl**, **Prof. Dr. Sven Diederichs** and **Prof. Dr. Andreas Trumpp** for helpful scientific discussions during my thesis advisory committee meetings.

I want to thank **Sandro** for performing the bioinformatic analysis of the APA dataset. It was not a straightforward approach, and I am so grateful that you invested so many thoughts and so much time into the analysis. Without you, I would still have a couple of external hard drives with data and no way to unravel the secrets of APA in hematopoiesis. Thanks for always coping with my countless questions and new ideas. Without you, this work would not have been possible. I am already looking forward to bringing you the next new data… I think I already have something in mind…

In addition, I would like to thank **Simon**, **Luisa** and **Katharina**. Thanks for your support, your help during so many experiments and your friendship. Especially, I want to thank Simon for his never ending enthusiasm, crazy ideas and positive energy. We have come a long way

together, and I am looking forward to continue working with you. You are an amazing colleague and friend. And of course, thank you so much for proof-reading this thesis.

I want to thank **all technicians** of A010 and HI-STEM for keeping the lab up and running. You are doing an amazing job. Without you, none of the projects, including my PhD thesis, would be possible. In particular, I would like to thank **Petra**. You helped me with countless experiments, and my life as a PhD student would have been so much harder without you. Thanks for being so supportive, positive and flexible. In addition, I want to thank **Paula**. You were the most remarkable apprentice. Thank you for all your help. I wish you all the best for your future and your next steps, and I have no doubt that you will do an amazing job. I would also like to thank **Adriana** and **Markus** for their help with maintaining the mouse colonies and genotyping.

In addition, I would like to thank all the different people working in the **DKFZ core facilities**, including the genomics and proteomics core facility and the center for preclinical research. Especially, I want to thank **Agnes** from the Omics IT and data management core facility for the analysis of the miRNA-seq data. Further, I would like to thank **Frank** from the transgenic service for his help in developing the targeting strategy for the *Pabpn1*^{flox} mouse model and for coordinating microinjections. In addition, a huge thank you to the flow cytometry unit, including **Steffen**, **Flo**, **Klaus**, **Marcus** and **Tobi**. I spent so many hours on the 7th floor, and all of you were always available to help me.

Furthermore, I would like to acknowledge our collaborators. I want to thank **Dr. Anne Klibanski** and her group for providing us with the *Meg3*^{flox} mouse model and for helpful discussions. I also want to thank **Dr. Kai Schönig** for his help with and thoughts on the *Pabpn1*^{flox} mouse model.

A big thanks to all **present and former members of HI-STEM and A010**. I really enjoyed working with all of you. It is great to know that the right person to solve technical problems, discuss science or have a good conversation with is always just around the corner, no matter at what time of the day. A special thanks to **Marc**, **Mattia** and **Simon H**. for helpful discussions, shared protocols and technical support for experiments.

Dagmar and **Erika**, thank you so much for your help with the manifold organizational questions. Without you, I would not have been able to go on business trips, participate in advanced training or fill out all the different forms.

Moreover, I want to thank my awesome aquarium members, **Sarah** and **Aykut**. Thanks for all the great conversations, the discussions and fun moments. You are the perfect desk

neighbors, and there is no one else in the world with whom I would like to share these precious 210 cm. Honestly, I would not like to miss having you in the office and I will always stay close to you!

Further, a huge thank you to my wonderful colleagues and friends: **Andi**, **Andrea**, **Aykut**, **Felix**, **Franzy**, **Jasper**, **Kristin**, **Lisa**, **Maija**, **Manu**, **Paula**, **Sarah** and **Vera**. Thank you for all the great moments, including lunch breaks with awesome casino food, coffee breaks, the brunch club and dinner parties, skiing trips, creative activities and bouldering nights. You are amazing people, and I really enjoyed my time with you inside and outside of the lab.

Abschließend möchte ich mich besonders bei meiner ganzen Familie bedanken. **Anne** und **Axel**, vielen Dank für eure fortwährende Unterstützung in den letzten 8 Jahren. **Ines**, **Sascha**, **Simon** und **Lina**, danke, dass es bei euch immer einen Ort gibt, an dem man die Arbeit vergessen kann und an dem nur der Moment zählt. Ganz besonders möchte ich meinen Eltern, **Marlies** und **Harald**, danken. Vielen Dank für eure Unterstützung, eure Liebe und euer Vertrauen in mich. Ihr habt mir immer alle Freiheiten gegeben und mich dabei nie alleine gelassen. Vielen Dank, dass ihr mir so viele Dinge ermöglicht habt. Zum Schluss möchte ich noch **Alex**, meinem Mann, danken. Vielen Dank für all deine Hilfe und dein Verständnis. Danke für das andauernde Korrekturlesen, das Zuhören und die endlosen Gespräche. Ohne dich wäre diese Arbeit so nicht möglich gewesen.

9 Bibliography

Acar, M., Kocherlakota, K.S., Murphy, M.M., Peyer, J.G., Oguro, H., Inra, C.N., Jaiyeola, C., Zhao, Z., Luby-Phelps, K., and Morrison, S.J. (2015). Deep imaging of bone marrow shows non-dividing stem cells are mainly perisinusoidal. Nature *526*, 126-130.

Adolfsson, J., Mansson, R., Buza-Vidas, N., Hultquist, A., Liuba, K., Jensen, C.T., Bryder, D., Yang, L., Borge, O.J., Thoren, L.A.*, et al.* (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.

Ahmed, A.S., Sheng, M.H., Wasnik, S., Baylink, D.J., and Lau, K.W. (2017). Effect of aging on stem cells. World J Exp Med *7*, 1-10.

Aida, T., Nakade, S., Sakuma, T., Izu, Y., Oishi, A., Mochida, K., Ishikubo, H., Usami, T., Aizawa, H., Yamamoto, T.*, et al.* (2016). Gene cassette knock-in in mammalian cells and zygotes by enhanced MMEJ. BMC Genomics *17*, 979.

Akashi, K., Traver, D., Miyamoto, T., and Weissman, I.L. (2000). A clonogenic common myeloid progenitor that gives rise to all myeloid lineages. Nature *404*, 193-197.

Alberts, B.J., A.; Lewis, J.; Raff, M.; Roberts, K.; Walter, P. (2008). Molecular biology of the cell, 5th edition. Garland Science, New York.

Aledo, J.C., Gomez-Fabre, P.M., Olalla, L., and Marquez, J. (2000). Identification of two human glutaminase loci and tissue-specific expression of the two related genes. Mamm Genome *11*, 1107-1110.

Altman, B.J., Stine, Z.E., and Dang, C.V. (2016). From Krebs to clinic: glutamine metabolism to cancer therapy. Nat Rev Cancer *16*, 619-634.

Alvarez-Dominguez, J.R., and Lodish, H.F. (2017). Emerging mechanisms of long noncoding RNA function during normal and malignant hematopoiesis. Blood *130*, 1965-1975.

An, J.J., Gharami, K., Liao, G.Y., Woo, N.H., Lau, A.G., Vanevski, F., Torre, E.R., Jones, K.R., Feng, Y., Lu, B.*, et al.* (2008). Distinct role of long 3' UTR BDNF mRNA in spine morphology and synaptic plasticity in hippocampal neurons. Cell *134*, 175-187.

Anders, S., Pyl, P.T., and Huber, W. (2015). HTSeq--a Python framework to work with high-throughput sequencing data. Bioinformatics *31*, 166-169.

Anders, S., Reyes, A., and Huber, W. (2012). Detecting differential usage of exons from RNA-seq data. Genome Res *22*, 2008-2017.

Andreassi, C., and Riccio, A. (2009). To localize or not to localize: mRNA fate is in 3'UTR ends. Trends Cell Biol *19*, 465-474.

Araki, S., Nakayama, Y., Sano, O., Nakao, S., Shimizu-Ogasawara, M., Toyoshiba, H., Nakanishi, A., and Aparicio, S. (2018). Decoding Transcriptome Dynamics of Genome-Encoded Polyadenylation and Autoregulation with Small-Molecule Modulators of Alternative Polyadenylation. Cell Chem Biol *25*, 1470-1484 e1475.

Astuti, D., Latif, F., Wagner, K., Gentle, D., Cooper, W.N., Catchpoole, D., Grundy, R., Ferguson-Smith, A.C., and Maher, E.R. (2005). Epigenetic alteration at the DLK1-GTL2 imprinted domain in human neoplasia: analysis of neuroblastoma, phaeochromocytoma and Wilms' tumour. Br J Cancer *92*, 1574-1580.

Bahr, C., von Paleske, L., Uslu, V.V., Remeseiro, S., Takayama, N., Ng, S.W., Murison, A., Langenfeld, K., Petretich, M., Scognamiglio, R.*, et al.* (2018). A Myc enhancer cluster regulates normal and leukaemic haematopoietic stem cell hierarchies. Nature *553*, 515-520.

Baldridge, M.T., King, K.Y., Boles, N.C., Weksberg, D.C., and Goodell, M.A. (2010). Quiescent haematopoietic stem cells are activated by IFN-gamma in response to chronic infection. Nature *465*, 793-797.

Banerjee, A., Vest, K.E., Pavlath, G.K., and Corbett, A.H. (2017). Nuclear poly(A) binding protein 1 (PABPN1) and Matrin3 interact in muscle cells and regulate RNA processing. Nucleic Acids Res *45*, 10706-10725.

Barker, J.E. (1994). Sl/Sld hematopoietic progenitors are deficient in situ. Exp Hematol *22*, 174-177.

Beaudoing, E., Freier, S., Wyatt, J.R., Claverie, J.M., and Gautheret, D. (2000). Patterns of variant polyadenylation signal usage in human genes. Genome Res *10*, 1001-1010.

Benetatos, L., Dasoula, A., Hatzimichael, E., Georgiou, I., Syrrou, M., and Bourantas, K.L. (2008). Promoter hypermethylation of the MEG3 (DLK1/MEG3) imprinted gene in multiple myeloma. Clin Lymphoma Myeloma *8*, 171-175.

Benetatos, L., Hatzimichael, E., Dasoula, A., Dranitsaris, G., Tsiara, S., Syrrou, M., Georgiou, I., and Bourantas, K.L. (2010). CpG methylation analysis of the MEG3 and SNRPN imprinted genes in acute myeloid leukemia and myelodysplastic syndromes. Leuk Res *34*, 148-153.

Benetatos, L., Vartholomatos, G., and Hatzimichael, E. (2011). MEG3 imprinted gene contribution in tumorigenesis. Int J Cancer *129*, 773-779.

Berkovits, B.D., and Mayr, C. (2015). Alternative 3' UTRs act as scaffolds to regulate membrane protein localization. Nature *522*, 363-367.

Bernitz, J.M., Kim, H.S., MacArthur, B., Sieburg, H., and Moore, K. (2016). Hematopoietic Stem Cells Count and Remember Self-Renewal Divisions. Cell *167*, 1296-1309 e1210.

Bertani, S., Sauer, S., Bolotin, E., and Sauer, F. (2011). The noncoding RNA Mistral activates Hoxa6 and Hoxa7 expression and stem cell differentiation by recruiting MLL1 to chromatin. Mol Cell *43*, 1040-1046.

Bolger, A.M., Lohse, M., and Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics *30*, 2114-2120.

Bouabe, H., and Okkenhaug, K. (2013). Gene targeting in mice: a review. Methods Mol Biol *1064*, 315-336.

Boutet, S.C., Cheung, T.H., Quach, N.L., Liu, L., Prescott, S.L., Edalati, A., Iori, K., and Rando, T.A. (2012). Alternative polyadenylation mediates microRNA regulation of muscle stem cell function. Cell Stem Cell *10*, 327-336.

Bowie, M.B., McKnight, K.D., Kent, D.G., McCaffrey, L., Hoodless, P.A., and Eaves, C.J. (2006). Hematopoietic stem cells proliferate until after birth and show a reversible phase-specific engraftment defect. J Clin Invest *116*, 2808-2816.

Braconi, C., Kogure, T., Valeri, N., Huang, N., Nuovo, G., Costinean, S., Negrini, M., Miotto, E., Croce, C.M., and Patel, T. (2011). microRNA-29 can regulate expression of the long non-coding RNA gene MEG3 in hepatocellular cancer. Oncogene *30*, 4750-4756.

Brenner, S., Jacob, F., and Meselson, M. (1961). An unstable intermediate carrying information from genes to ribosomes for protein synthesis. Nature *190*, 576-581.

Brumbaugh, J., Di Stefano, B., Wang, X., Borkent, M., Forouzmand, E., Clowers, K.J., Ji, F., Schwarz, B.A., Kalocsay, M., Elledge, S.J*., et al.* (2018). Nudt21 Controls Cell Fate by Connecting Alternative Polyadenylation to Chromatin Signaling. Cell *172*, 106-120 e121.

Bryder, D., Rossi, D.J., and Weissman, I.L. (2006). Hematopoietic stem cells: the paradigmatic tissue-specific stem cell. Am J Pathol *169*, 338-346.

Burgio, G. (2018). Redefining mouse transgenesis with CRISPR/Cas9 genome editing technology. Genome Biol *19*, 27.

Busch, K., Klapproth, K., Barile, M., Flossdorf, M., Holland-Letz, T., Schlenner, S.M., Reth, M., Hofer, T., and Rodewald, H.R. (2015). Fundamental properties of unperturbed haematopoiesis from stem cells in vivo. Nature *518*, 542-546.

Butcher, S.E., and Brow, D.A. (2005). Towards understanding the catalytic core structure of the spliceosome. Biochem Soc Trans *33*, 447-449.

Cabezas-Wallscheid, N., Buettner, F., Sommerkamp, P., Klimmeck, D., Ladel, L., Thalheimer, F.B., Pastor-Flores, D., Roma, L.P., Renders, S., Zeisberger, P*., et al.* (2017). Vitamin A-Retinoic Acid Signaling Regulates Hematopoietic Stem Cell Dormancy. Cell *169*, 807-823 e819.

Cabezas-Wallscheid, N., Klimmeck, D., Hansson, J., Lipka, D.B., Reyes, A., Wang, Q., Weichenhan, D., Lier, A., von Paleske, L., Renders, S*., et al.* (2014). Identification of

regulatory networks in HSCs and their immediate progeny via integrated proteome, transcriptome, and DNA methylome analysis. Cell Stem Cell *15*, 507-522.

Carlsson, C., Tornehave, D., Lindberg, K., Galante, P., Billestrup, N., Michelsen, B., Larsson, L.I., and Nielsen, J.H. (1997). Growth hormone and prolactin stimulate the expression of rat preadipocyte factor-1/delta-like protein in pancreatic islets: molecular cloning and expression pattern during development and growth of the endocrine pancreas. Endocrinology *138*, 3940-3948.

Carnevalli, L.S., Scognamiglio, R., Cabezas-Wallscheid, N., Rahmig, S., Laurenti, E., Masuda, K., Jockel, L., Kuck, A., Sujer, S., Polykratis, A.*, et al.* (2014). Improved HSC reconstitution and protection from inflammatory stress and chemotherapy in mice lacking granzyme B. J Exp Med *211*, 769-779.

Cassago, A., Ferreira, A.P., Ferreira, I.M., Fornezari, C., Gomes, E.R., Greene, K.S., Pereira, H.M., Garratt, R.C., Dias, S.M., and Ambrosio, A.L. (2012). Mitochondrial localization and structure-based phosphate activation mechanism of Glutaminase C with implications for cancer metabolism. Proc Natl Acad Sci U S A *109*, 1092-1097.

Cesana, M., Guo, M.H., Cacchiarelli, D., Wahlster, L., Barragan, J., Doulatov, S., Vo, L.T., Salvatori, B., Trapnell, C., Clement, K.*, et al.* (2018). A CLK3-HMGA2 Alternative Splicing Axis Impacts Human Hematopoietic Stem Cell Molecular Identity throughout Development. Cell Stem Cell *22*, 575-588 e577.

Chang, J.W., Zhang, W., Yeh, H.S., de Jong, E.P., Jun, S., Kim, K.H., Bae, S.S., Beckman, K., Hwang, T.H., Kim, K.S.*, et al.* (2015). mRNA 3'-UTR shortening is a molecular signature of mTORC1 activation. Nat Commun *6*, 7218.

Chase, A., Leung, W., Tapper, W., Jones, A.V., Knoops, L., Rasi, C., Forsberg, L.A., Guglielmelli, P., Zoi, K., Hall, V.*, et al.* (2015). Profound parental bias associated with chromosome 14 acquired uniparental disomy indicates targeting of an imprinted locus. Leukemia *29*, 2069-2074.

Chen, J.Y., Miyanishi, M., Wang, S.K., Yamazaki, S., Sinha, R., Kao, K.S., Seita, J., Sahoo, D., Nakauchi, H., and Weissman, I.L. (2016). Hoxb5 marks long-term haematopoietic stem cells and reveals a homogenous perivascular niche. Nature *530*, 223-227.

Chen, M.J., Yokomizo, T., Zeigler, B.M., Dzierzak, E., and Speck, N.A. (2009). Runx1 is required for the endothelial to haematopoietic cell transition but not thereafter. Nature *457*, 887-891.

Chen, W., Jia, Q., Song, Y., Fu, H., Wei, G., and Ni, T. (2017). Alternative Polyadenylation: Methods, Findings, and Impacts. Genomics Proteomics Bioinformatics *15*, 287-300.

Crane, G.M., Jeffery, E., and Morrison, S.J. (2017). Adult haematopoietic stem cell niches. Nat Rev Immunol *17*, 573-590.

Curinha, A., Oliveira Braz, S., Pereira-Castro, I., Cruz, A., and Moreira, A. (2014). Implications of polyadenylation in health and disease. Nucleus *5*, 508-519.

Czechowicz, A., Kraft, D., Weissman, I.L., and Bhattacharya, D. (2007). Efficient transplantation via antibody-based clearance of hematopoietic stem cell niches. Science *318*, 1296-1299.

da Rocha, S.T., Edwards, C.A., Ito, M., Ogata, T., and Ferguson-Smith, A.C. (2008). Genomic imprinting at the mammalian Dlk1-Dio3 domain. Trends Genet *24*, 306-316.

Davis, R., and Shi, Y. (2014). The polyadenylation code: a unified model for the regulation of mRNA alternative polyadenylation. J Zhejiang Univ Sci B *15*, 429-437.

de Klerk, E., Venema, A., Anvar, S.Y., Goeman, J.J., Hu, O., Trollet, C., Dickson, G., den Dunnen, J.T., van der Maarel, S.M., Raz, V.*, et al.* (2012). Poly(A) binding protein nuclear 1 levels affect alternative polyadenylation. Nucleic Acids Res *40*, 9089-9101.

Delassus, S., and Cumano, A. (1996). Circulation of hematopoietic progenitors in the mouse embryo. Immunity *4*, 97-106.

Derrien, T., Johnson, R., Bussotti, G., Tanzer, A., Djebali, S., Tilgner, H., Guernec, G., Martin, D., Merkel, A., Knowles, D.G.*, et al.* (2012). The GENCODE v7 catalog of human long noncoding RNAs: analysis of their gene structure, evolution, and expression. Genome Res *22*, 1775-1789.

Derti, A., Garrett-Engele, P., Macisaac, K.D., Stevens, R.C., Sriram, S., Chen, R., Rohl, C.A., Johnson, J.M., and Babak, T. (2012). A quantitative atlas of polyadenylation in five mammals. Genome Res *22*, 1173-1183.

Di Giammartino, D.C., Nishida, K., and Manley, J.L. (2011). Mechanisms and consequences of alternative polyadenylation. Mol Cell *43*, 853-866.

Djebali, S., Davis, C.A., Merkel, A., Dobin, A., Lassmann, T., Mortazavi, A., Tanzer, A., Lagarde, J., Lin, W., Schlesinger, F.*, et al.* (2012). Landscape of transcription in human cells. Nature *489*, 101-108.

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

Dreyfuss, G., Philipson, L., and Mattaj, I.W. (1988). Ribonucleoprotein particles in cellular processes. J Cell Biol *106*, 1419-1425.

Dykstra, B., Kent, D., Bowie, M., McCaffrey, L., Hamilton, M., Lyons, K., Lee, S.J., Brinkman, R., and Eaves, C. (2007). Long-term propagation of distinct hematopoietic differentiation programs in vivo. Cell Stem Cell *1*, 218-229.

Edmonds, M., and Abrams, R. (1960). Polynucleotide biosynthesis: formation of a sequence of adenylate units from adenosine triphosphate by an enzyme from thymus nuclei. J Biol Chem *235*, 1142-1149.

Ehninger, A., Boch, T., Uckelmann, H., Essers, M.A., Mudder, K., Sleckman, B.P., and Trumpp, A. (2014). Posttranscriptional regulation of c-Myc expression in adult murine HSCs during homeostasis and interferon-alpha-induced stress response. Blood *123*, 3909-3913.

Ehninger, A., and Trumpp, A. (2011). The bone marrow stem cell niche grows up: mesenchymal stem cells and macrophages move in. J Exp Med *208*, 421-428.

Elgadi, K.M., Meguid, R.A., Qian, M., Souba, W.W., and Abcouwer, S.F. (1999). Cloning and analysis of unique human glutaminase isoforms generated by tissue-specific alternative splicing. Physiol Genomics *1*, 51-62.

Elkon, R., Drost, J., van Haaften, G., Jenal, M., Schrier, M., Oude Vrielink, J.A., and Agami, R. (2012). E2F mediates enhanced alternative polyadenylation in proliferation. Genome Biol *13*, R59.

Elkon, R., Ugalde, A.P., and Agami, R. (2013). Alternative cleavage and polyadenylation: extent, regulation and function. Nat Rev Genet *14*, 496-506.

Ermolaeva, M., Neri, F., Ori, A., and Rudolph, K.L. (2018). Cellular and epigenetic drivers of stem cell ageing. Nat Rev Mol Cell Biol *19*, 594-610.

Esplin, B.L., Shimazu, T., Welner, R.S., Garrett, K.P., Nie, L., Zhang, Q., Humphrey, M.B., Yang, Q., Borghesi, L.A., and Kincade, P.W. (2011). Chronic exposure to a TLR ligand injures hematopoietic stem cells. J Immunol *186*, 5367-5375.

Essers, M.A., Offner, S., Blanco-Bose, W.E., Waibler, Z., Kalinke, U., Duchosal, M.A., and Trumpp, A. (2009). IFNalpha activates dormant haematopoietic stem cells in vivo. Nature *458*, 904-908.

Fabian, M.R., and Sonenberg, N. (2012). The mechanics of miRNA-mediated gene silencing: a look under the hood of miRISC. Nat Struct Mol Biol *19*, 586-593.

Fatica, A., and Bozzoni, I. (2014). Long non-coding RNAs: new players in cell differentiation and development. Nat Rev Genet *15*, 7-21.

Feng, J., Funk, W.D., Wang, S.S., Weinrich, S.L., Avilion, A.A., Chiu, C.P., Adams, R.R., Chang, E., Allsopp, R.C., Yu, J.*, et al.* (1995). The RNA component of human telomerase. Science *269*, 1236-1241.

Fineran, P.C., and Charpentier, E. (2012). Memory of viral infections by CRISPR-Cas adaptive immune systems: acquisition of new information. Virology *434*, 202-209.

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.*, et al.* (2014). Replication stress is a potent driver of functional decline in ageing haematopoietic stem cells. Nature *512*, 198-202.

Forsberg, E.C., Serwold, T., Kogan, S., Weissman, I.L., and Passegue, E. (2006). New evidence supporting megakaryocyte-erythrocyte potential of flk2/flt3+ multipotent hematopoietic progenitors. Cell *126*, 415-426.

Foudi, A., Hochedlinger, K., Van Buren, D., Schindler, J.W., Jaenisch, R., Carey, V., and Hock, H. (2009). Analysis of histone 2B-GFP retention reveals slowly cycling hematopoietic stem cells. Nat Biotechnol *27*, 84-90.

Fu, J., Zuber, J., Martinez, M., Shonts, B., Obradovic, A., Wang, H., Lau, S.P., Xia, A., Waffarn, E.E., Frangaj, K.*, et al.* (2019). Human Intestinal Allografts Contain Functional Hematopoietic Stem and Progenitor Cells that Are Maintained by a Circulating Pool. Cell Stem Cell *24*, 227-239 e228.

Fu, Y., Chen, L., Chen, C., Ge, Y., Kang, M., Song, Z., Li, J., Feng, Y., Huo, Z., He, G.*, et al.* (2018). Crosstalk between alternative polyadenylation and miRNA in regulation of protein translational efficiency. Genome Res.

Fuchs, E., and Chen, T. (2013). A matter of life and death: self-renewal in stem cells. EMBO Rep *14*, 39-48.

Gallipoli, P., Giotopoulos, G., Tzelepis, K., Costa, A.S.H., Vohra, S., Medina-Perez, P., Basheer, F., Marando, L., Di Lisio, L., Dias, J.M.L.*, et al.* (2018). Glutaminolysis is a metabolic dependency in FLT3(ITD) acute myeloid leukemia unmasked by FLT3 tyrosine kinase inhibition. Blood *131*, 1639-1653.

Galloway, J.L., and Zon, L.I. (2003). Ontogeny of hematopoiesis: examining the emergence of hematopoietic cells in the vertebrate embryo. Curr Top Dev Biol *53*, 139-158.

Gan, B., and DePinho, R.A. (2009). mTORC1 signaling governs hematopoietic stem cell quiescence. Cell Cycle *8*, 1003-1006.

Gao, P., Tchernyshyov, I., Chang, T.C., Lee, Y.S., Kita, K., Ochi, T., Zeller, K.I., De Marzo, A.M., Van Eyk, J.E., Mendell, J.T.*, et al.* (2009). c-Myc suppression of miR-23a/b enhances mitochondrial glutaminase expression and glutamine metabolism. Nature *458*, 762-765.

Gautheret, D., Poirot, O., Lopez, F., Audic, S., and Claverie, J.M. (1998). Alternate polyadenylation in human mRNAs: a large-scale analysis by EST clustering. Genome Res *8*, 524-530.

Gazit, R., Mandal, P.K., Ebina, W., Ben-Zvi, A., Nombela-Arrieta, C., Silberstein, L.E., and Rossi, D.J. (2014). Fgd5 identifies hematopoietic stem cells in the murine bone marrow. J Exp Med *211*, 1315-1331.

Gekas, C., Dieterlen-Lievre, F., Orkin, S.H., and Mikkola, H.K. (2005). The placenta is a niche for hematopoietic stem cells. Dev Cell *8*, 365-375.

Georgiades, P., Ogilvy, S., Duval, H., Licence, D.R., Charnock-Jones, D.S., Smith, S.K., and Print, C.G. (2002). VavCre transgenic mice: a tool for mutagenesis in hematopoietic and endothelial lineages. Genesis *34*, 251-256.

Gilbert, W. (1986). Origin of life: The RNA world. Nature *319*, 1476-4687.

Godin, I., Garcia-Porrero, J.A., Dieterlen-Lievre, F., and Cumano, A. (1999). Stem cell emergence and hemopoietic activity are incompatible in mouse intraembryonic sites. J Exp Med *190*, 43-52.

Goldstein, O., Meyer, K., Greenshpan, Y., Bujanover, N., Feigin, M., Ner-Gaon, H., Shay, T., and Gazit, R. (2017). Mapping Whole-Transcriptome Splicing in Mouse Hematopoietic Stem Cells. Stem Cell Reports *8*, 163-176.

Gordon, F.E., Nutt, C.L., Cheunsuchon, P., Nakayama, Y., Provencher, K.A., Rice, K.A., Zhou, Y., Zhang, X., and Klibanski, A. (2010). Increased expression of angiogenic genes in the brains of mouse meg3-null embryos. Endocrinology *151*, 2443-2452.

Grover, A., Sanjuan-Pla, A., Thongjuea, S., Carrelha, J., Giustacchini, A., Gambardella, A., Macaulay, I., Mancini, E., Luis, T.C., Mead, A.*, et al.* (2016). Single-cell RNA sequencing reveals molecular and functional platelet bias of aged haematopoietic stem cells. Nat Commun *7*, 11075.

Gruber, A.R., Martin, G., Keller, W., and Zavolan, M. (2012). Cleavage factor Im is a key regulator of 3' UTR length. RNA Biol *9*, 1405-1412.

Gruber, A.R., Martin, G., Keller, W., and Zavolan, M. (2014a). Means to an end: mechanisms of alternative polyadenylation of messenger RNA precursors. Wiley Interdiscip Rev RNA *5*, 183-196.

Gruber, A.R., Martin, G., Muller, P., Schmidt, A., Gruber, A.J., Gumienny, R., Mittal, N., Jayachandran, R., Pieters, J., Keller, W*., et al.* (2014b). Global 3' UTR shortening has a limited effect on protein abundance in proliferating T cells. Nat Commun *5*, 5465.

Guerrier-Takada, C., Gardiner, K., Marsh, T., Pace, N., and Altman, S. (1983). The RNA moiety of ribonuclease P is the catalytic subunit of the enzyme. Cell *35*, 849-857.

Guo, G., Luc, S., Marco, E., Lin, T.W., Peng, C., Kerenyi, M.A., Beyaz, S., Kim, W., Xu, J., Das, P.P*., et al.* (2013). Mapping cellular hierarchy by single-cell analysis of the cell surface repertoire. Cell Stem Cell *13*, 492-505.

Ha, K.C.H., Blencowe, B.J., and Morris, Q. (2018). QAPA: a new method for the systematic analysis of alternative polyadenylation from RNA-seq data. Genome Biol *19*, 45.

Haas, S., Hansson, J., Klimmeck, D., Loeffler, D., Velten, L., Uckelmann, H., Wurzer, S., Prendergast, A.M., Schnell, A., Hexel, K*., et al.* (2015). Inflammation-Induced Emergency Megakaryopoiesis Driven by Hematopoietic Stem Cell-like Megakaryocyte Progenitors. Cell Stem Cell *17*, 422-434.

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

Haeussler, M., Schonig, K., Eckert, H., Eschstruth, A., Mianne, J., Renaud, J.B., Schneider-Maunoury, S., Shkumatava, A., Teboul, L., Kent, J*., et al.* (2016). Evaluation of off-target and on-target scoring algorithms and integration into the guide RNA selection tool CRISPOR. Genome Biol *17*, 148.

Hamon, Y., Legowska, M., Herve, V., Dallet-Choisy, S., Marchand-Adam, S., Vanderlynden, L., Demonte, M., Williams, R., Scott, C.J., Si-Tahar, M*., et al.* (2016). Neutrophilic Cathepsin C Is Maturated by a Multistep Proteolytic Process and Secreted by Activated Cells during Inflammatory Lung Diseases. J Biol Chem *291*, 8486-8499.

Heinz, S., Benner, C., Spann, N., Bertolino, E., Lin, Y.C., Laslo, P., Cheng, J.X., Murre, C., Singh, H., and Glass, C.K. (2010). Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. Mol Cell *38*, 576-589.

Henras, A.K., Dez, C., and Henry, Y. (2004). RNA structure and function in C/D and H/ACA s(no)RNPs. Curr Opin Struct Biol *14*, 335-343.

Ho, T.T., Warr, M.R., Adelman, E.R., Lansinger, O.M., Flach, J., Verovskaya, E.V., Figueroa, M.E., and Passegue, E. (2017). Autophagy maintains the metabolism and function of young and old stem cells. Nature *543*, 205-210.

Hoagland, M.B., Stephenson, M.L., Scott, J.F., Hecht, L.I., and Zamecnik, P.C. (1958). A soluble ribonucleic acid intermediate in protein synthesis. J Biol Chem *231*, 241-257.

Hollerer, I., Curk, T., Haase, B., Benes, V., Hauer, C., Neu-Yilik, G., Bhuvanagiri, M., Hentze, M.W., and Kulozik, A.E. (2016). The differential expression of alternatively polyadenylated transcripts is a common stress-induced response mechanism that modulates mammalian mRNA expression in a quantitative and qualitative fashion. RNA *22*, 1441-1453.

Hoque, M., Ji, Z., Zheng, D., Luo, W., Li, W., You, B., Park, J.Y., Yehia, G., and Tian, B. (2013). Analysis of alternative cleavage and polyadenylation by 3' region extraction and deep sequencing. Nat Methods *10*, 133-139.

Hu, J., Lutz, C.S., Wilusz, J., and Tian, B. (2005). Bioinformatic identification of candidate cis-regulatory elements involved in human mRNA polyadenylation. RNA *11*, 1485-1493.

Hutchins, L.N., Murphy, S.M., Singh, P., and Graber, J.H. (2008). Position-dependent motif characterization using non-negative matrix factorization. Bioinformatics *24*, 2684-2690.

Ichinose, J., Watanabe, K., Sano, A., Nagase, T., Nakajima, J., Fukayama, M., Yatomi, Y., Ohishi, N., and Takai, D. (2014). Alternative polyadenylation is associated with lower expression of PABPN1 and poor prognosis in non-small cell lung cancer. Cancer Sci *105*, 1135-1141.

Ignatiadis, N., Klaus, B., Zaugg, J.B., and Huber, W. (2016). Data-driven hypothesis weighting increases detection power in genome-scale multiple testing. Nat Methods *13*, 577-580.

Ikuta, K., and Weissman, I.L. (1992). Evidence that hematopoietic stem cells express mouse c-kit but do not depend on steel factor for their generation. Proc Natl Acad Sci U S A *89*, 1502-1506.

Irving, M.D., Buiting, K., Kanber, D., Donaghue, C., Schulz, R., Offiah, A., Mohammed, S.N., and Oakey, R.J. (2010). Segmental paternal uniparental disomy (patUPD) of 14q32 with abnormal methylation elicits the characteristic features of complete patUPD14. Am J Med Genet A *152A*, 1942-1950.

Ito, K., Carracedo, A., Weiss, D., Arai, F., Ala, U., Avigan, D.E., Schafer, Z.T., Evans, R.M., Suda, T., Lee, C.H.*, et al.* (2012). A PML-PPAR-delta pathway for fatty acid oxidation regulates hematopoietic stem cell maintenance. Nat Med *18*, 1350-1358.

Ito, K., and Ito, K. (2018). Hematopoietic stem cell fate through metabolic control. Exp Hematol *64*, 1-11.

Ito, K., and Suda, T. (2014). Metabolic requirements for the maintenance of self-renewing stem cells. Nat Rev Mol Cell Biol *15*, 243-256.

Ito, K., Turcotte, R., Cui, J., Zimmerman, S.E., Pinho, S., Mizoguchi, T., Arai, F., Runnels, J.M., Alt, C., Teruya-Feldstein, J.*, et al.* (2016). Self-renewal of a purified Tie2+ hematopoietic stem cell population relies on mitochondrial clearance. Science *354*, 1156-1160.

Jacque, N., Ronchetti, A.M., Larrue, C., Meunier, G., Birsen, R., Willems, L., Saland, E., Decroocq, J., Maciel, T.T., Lambert, M.*, et al.* (2015). Targeting glutaminolysis has antileukemic activity in acute myeloid leukemia and synergizes with BCL-2 inhibition. Blood *126*, 1346-1356.

Jady, B.E., Bertrand, E., and Kiss, T. (2004). Human telomerase RNA and box H/ACA scaRNAs share a common Cajal body-specific localization signal. J Cell Biol *164*, 647-652.

Jafri, S., Moore, S.D., Morrell, N.W., and Ormiston, M.L. (2017). A sex-specific reconstitution bias in the competitive CD45.1/CD45.2 congenic bone marrow transplant model. Sci Rep *7*, 3495.

Jan, C.H., Friedman, R.C., Ruby, J.G., and Bartel, D.P. (2011). Formation, regulation and evolution of Caenorhabditis elegans 3'UTRs. Nature *469*, 97-101.

Jenal, M., Elkon, R., Loayza-Puch, F., van Haaften, G., Kuhn, U., Menzies, F.M., Oude Vrielink, J.A., Bos, A.J., Drost, J., Rooijers, K.*, et al.* (2012). The poly(A)-binding protein nuclear 1 suppresses alternative cleavage and polyadenylation sites. Cell *149*, 538-553.

Jereb, S., Hwang, H.W., Van Otterloo, E., Govek, E.E., Fak, J.J., Yuan, Y., Hatten, M.E., and Darnell, R.B. (2018). Differential 3' Processing of Specific Transcripts Expands Regulatory and Protein Diversity Across Neuronal Cell Types. Elife *7*.

Ji, H., Ehrlich, L.I., Seita, J., Murakami, P., Doi, A., Lindau, P., Lee, H., Aryee, M.J., Irizarry, R.A., Kim, K.*, et al.* (2010). Comprehensive methylome map of lineage commitment from haematopoietic progenitors. Nature *467*, 338-342.

Ji, Z., Lee, J.Y., Pan, Z., Jiang, B., and Tian, B. (2009). Progressive lengthening of 3' untranslated regions of mRNAs by alternative polyadenylation during mouse embryonic development. Proc Natl Acad Sci U S A *106*, 7028-7033.

Ji, Z., Luo, W., Li, W., Hoque, M., Pan, Z., Zhao, Y., and Tian, B. (2011). Transcriptional activity regulates alternative cleavage and polyadenylation. Mol Syst Biol *7*, 534.

Ji, Z., and Tian, B. (2009). Reprogramming of 3' untranslated regions of mRNAs by alternative polyadenylation in generation of pluripotent stem cells from different cell types. PLoS One *4*, e8419.

Jia, X., Yuan, S., Wang, Y., Fu, Y., Ge, Y., Ge, Y., Lan, X., Feng, Y., Qiu, F., Li, P.*, et al.* (2017). The role of alternative polyadenylation in the antiviral innate immune response. Nat Commun *8*, 14605.

Jinek, M., Chylinski, K., Fonfara, I., Hauer, M., Doudna, J.A., and Charpentier, E. (2012). A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. Science *337*, 816-821.

Johnson, G.R., and Moore, M.A. (1975). Role of stem cell migration in initiation of mouse foetal liver haemopoiesis. Nature *258*, 726-728.

Johnson, M.O., Wolf, M.M., Madden, M.Z., Andrejeva, G., Sugiura, A., Contreras, D.C., Maseda, D., Liberti, M.V., Paz, K., Kishton, R.J.*, et al.* (2018). Distinct Regulation of Th17 and Th1 Cell Differentiation by Glutaminase-Dependent Metabolism. Cell.

Julian, L.M., and Blais, A. (2015). Transcriptional control of stem cell fate by E2Fs and pocket proteins. Front Genet *6*, 161.

Kagami, M., O'Sullivan, M.J., Green, A.J., Watabe, Y., Arisaka, O., Masawa, N., Matsuoka, K., Fukami, M., Matsubara, K., Kato, F.*, et al.* (2010). The IG-DMR and the MEG3-DMR at human chromosome 14q32.2: hierarchical interaction and distinct functional properties as imprinting control centers. PLoS Genet *6*, e1000992.

Kagami, M., Sekita, Y., Nishimura, G., Irie, M., Kato, F., Okada, M., Yamamori, S., Kishimoto, H., Nakayama, M., Tanaka, Y.*, et al.* (2008). Deletions and epimutations affecting the human 14q32.2 imprinted region in individuals with paternal and maternal upd(14)-like phenotypes. Nat Genet *40*, 237-242.

Kalaitzidis, D., Sykes, S.M., Wang, Z., Punt, N., Tang, Y., Ragu, C., Sinha, A.U., Lane, S.W., Souza, A.L., Clish, C.B.*, et al.* (2012). mTOR complex 1 plays critical roles in hematopoiesis and Pten-loss-evoked leukemogenesis. Cell Stem Cell *11*, 429-439.

Kameswaran, V., and Kaestner, K.H. (2014). The Missing lnc(RNA) between the pancreatic beta-cell and diabetes. Front Genet *5*, 200.

Karamitros, D., Stoilova, B., Aboukhalil, Z., Hamey, F., Reinisch, A., Samitsch, M., Quek, L., Otto, G., Repapi, E., Doondeea, J.*, et al.* (2018). Single-cell analysis reveals the continuum of human lympho-myeloid progenitor cells. Nat Immunol *19*, 85-97.

Kawakami, T., Chano, T., Minami, K., Okabe, H., Okada, Y., and Okamoto, K. (2006). Imprinted DLK1 is a putative tumor suppressor gene and inactivated by epimutation at the region upstream of GTL2 in human renal cell carcinoma. Hum Mol Genet *15*, 821-830.

Kerwitz, Y., Kuhn, U., Lilie, H., Knoth, A., Scheuermann, T., Friedrich, H., Schwarz, E., and Wahle, E. (2003). Stimulation of poly(A) polymerase through a direct interaction with the nuclear poly(A) binding protein allosterically regulated by RNA. EMBO J *22*, 3705-3714.

Khoury, H., Suarez-Saiz, F., Wu, S., and Minden, M.D. (2010). An upstream insulator regulates DLK1 imprinting in AML. Blood *115*, 2260-2263.

Kiel, M.J., Yilmaz, O.H., Iwashita, T., Yilmaz, O.H., Terhorst, C., and Morrison, S.J. (2005). SLAM family receptors distinguish hematopoietic stem and progenitor cells and reveal endothelial niches for stem cells. Cell *121*, 1109-1121.

Kim, E., Cheng, Y., Bolton-Gillespie, E., Cai, X., Ma, C., Tarangelo, A., Le, L., Jambhekar, M., Raman, P., Hayer, K.E.*, et al.* (2017). Rb family proteins enforce the homeostasis of quiescent hematopoietic stem cells by repressing Socs3 expression. J Exp Med *214*, 1901-1912.

King, R.G., Herrin, B.R., and Justement, L.B. (2006). Trem-like transcript 2 is expressed on cells of the myeloid/granuloid and B lymphoid lineage and is up-regulated in response to inflammation. J Immunol *176*, 6012-6021.

Klimmeck, D., Cabezas-Wallscheid, N., Reyes, A., von Paleske, L., Renders, S., Hansson, J., Krijgsveld, J., Huber, W., and Trumpp, A. (2014). Transcriptome-wide profiling and posttranscriptional analysis of hematopoietic stem/progenitor cell differentiation toward myeloid commitment. Stem Cell Reports *3*, 858-875.

Kondo, M., Wagers, A.J., Manz, M.G., Prohaska, S.S., Scherer, D.C., Beilhack, G.F., Shizuru, J.A., and Weissman, I.L. (2003). Biology of hematopoietic stem cells and progenitors: implications for clinical application. Annu Rev Immunol *21*, 759-806.

Kondo, M., Weissman, I.L., and Akashi, K. (1997). Identification of clonogenic common lymphoid progenitors in mouse bone marrow. Cell *91*, 661-672.

Koopman, G., Reutelingsperger, C.P., Kuijten, G.A., Keehnen, R.M., Pals, S.T., and van Oers, M.H. (1994). Annexin V for flow cytometry detection of phosphatidylserine expression on B cells undergoing apoptosis. Blood *84*, 1415-1420.

Kruger, K., Grabowski, P.J., Zaug, A.J., Sands, J., Gottschling, D.E., and Cech, T.R. (1982). Self-splicing RNA: autoexcision and autocyclization of the ribosomal RNA intervening sequence of Tetrahymena. Cell *31*, 147-157.

Kuhn, R., Schwenk, F., Aguet, M., and Rajewsky, K. (1995). Inducible gene targeting in mice. Science *269*, 1427-1429.

Kuhn, U., Gundel, M., Knoth, A., Kerwitz, Y., Rudel, S., and Wahle, E. (2009). Poly(A) tail length is controlled by the nuclear poly(A)-binding protein regulating the interaction between poly(A) polymerase and the cleavage and polyadenylation specificity factor. J Biol Chem *284*, 22803-22814.

Laborda, J. (2000). The role of the epidermal growth factor-like protein dlk in cell differentiation. Histol Histopathol *15*, 119-129.

Lackford, B., Yao, C., Charles, G.M., Weng, L., Zheng, X., Choi, E.A., Xie, X., Wan, J., Xing, Y., Freudenberg, J.M.*, et al.* (2014). Fip1 regulates mRNA alternative polyadenylation to promote stem cell self-renewal. EMBO J *33*, 878-889.

Lai, C.Y., Yamazaki, S., Okabe, M., Suzuki, S., Maeyama, Y., Iimura, Y., Onodera, M., Kakuta, S., Iwakura, Y., Nojima, M.*, et al.* (2014). Stage-specific roles for CXCR4 signaling in murine hematopoietic stem/progenitor cells in the process of bone marrow repopulation. Stem Cells *32*, 1929-1942.

Langmead, B., and Salzberg, S.L. (2012). Fast gapped-read alignment with Bowtie 2. Nat Methods *9*, 357-359.

Latos, P.A., Pauler, F.M., Koerner, M.V., Senergin, H.B., Hudson, Q.J., Stocsits, R.R., Allhoff, W., Stricker, S.H., Klement, R.M., Warczok, K.E.*, et al.* (2012). Airn transcriptional overlap, but not its lncRNA products, induces imprinted Igf2r silencing. Science *338*, 1469-1472.

Lau, A.G., Irier, H.A., Gu, J., Tian, D., Ku, L., Liu, G., Xia, M., Fritsch, B., Zheng, J.Q., Dingledine, R.*, et al.* (2010). Distinct 3'UTRs differentially regulate activity-dependent translation of brain-derived neurotrophic factor (BDNF). Proc Natl Acad Sci U S A *107*, 15945-15950.

Laurenti, E., Frelin, C., Xie, S., Ferrari, R., Dunant, C.F., Zandi, S., Neumann, A., Plumb, I., Doulatov, S., Chen, J.*, et al.* (2015). CDK6 levels regulate quiescence exit in human hematopoietic stem cells. Cell Stem Cell *16*, 302-313.

Lee, R.C., Feinbaum, R.L., and Ambros, V. (1993). The C. elegans heterochronic gene lin-4 encodes small RNAs with antisense complementarity to lin-14. Cell *75*, 843-854.

Lee, S.H., Singh, I., Tisdale, S., Abdel-Wahab, O., Leslie, C.S., and Mayr, C. (2018). Widespread intronic polyadenylation inactivates tumour suppressor genes in leukaemia. Nature *561*, 127-131.

Lefrancais, E., Ortiz-Munoz, G., Caudrillier, A., Mallavia, B., Liu, F., Sayah, D.M., Thornton, E.E., Headley, M.B., David, T., Coughlin, S.R.*, et al.* (2017). The lung is a site of platelet biogenesis and a reservoir for haematopoietic progenitors. Nature *544*, 105-109.

Leonardo, T.R., Schultheisz, H.L., Loring, J.F., and Laurent, L.C. (2012). The functions of microRNAs in pluripotency and reprogramming. Nat Cell Biol *14*, 1114-1121.

Li, W., You, B., Hoque, M., Zheng, D., Luo, W., Ji, Z., Park, J.Y., Gunderson, S.I., Kalsotra, A., Manley, J.L.*, et al.* (2015). Systematic profiling of poly(A)+ transcripts modulated by core 3' end processing and splicing factors reveals regulatory rules of alternative cleavage and polyadenylation. PLoS Genet *11*, e1005166.

Licatalosi, D.D., Mele, A., Fak, J.J., Ule, J., Kayikci, M., Chi, S.W., Clark, T.A., Schweitzer, A.C., Blume, J.E., Wang, X.*, et al.* (2008). HITS-CLIP yields genome-wide insights into brain alternative RNA processing. Nature *456*, 464-469.

Lin, S.P., Coan, P., da Rocha, S.T., Seitz, H., Cavaille, J., Teng, P.W., Takada, S., and Ferguson-Smith, A.C. (2007). Differential regulation of imprinting in the murine embryo and placenta by the Dlk1-Dio3 imprinting control region. Development *134*, 417-426.

Lin, S.P., Youngson, N., Takada, S., Seitz, H., Reik, W., Paulsen, M., Cavaille, J., and Ferguson-Smith, A.C. (2003). Asymmetric regulation of imprinting on the maternal and paternal chromosomes at the Dlk1-Gtl2 imprinted cluster on mouse chromosome 12. Nat Genet *35*, 97-102.

Lipka, D.B., Wang, Q., Cabezas-Wallscheid, N., Klimmeck, D., Weichenhan, D., Herrmann, C., Lier, A., Brocks, D., von Paleske, L., Renders, S.*, et al.* (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*, 3476-3487.

Liu, D., Brockman, J.M., Dass, B., Hutchins, L.N., Singh, P., McCarrey, J.R., MacDonald, C.C., and Graber, J.H. (2007). Systematic variation in mRNA 3'-processing signals during mouse spermatogenesis. Nucleic Acids Res *35*, 234-246.

Liu, E.T., Bolcun-Filas, E., Grass, D.S., Lutz, C., Murray, S., Shultz, L., and Rosenthal, N. (2017). Of mice and CRISPR: The post-CRISPR future of the mouse as a model system for the human condition. EMBO Rep *18*, 187-193.

Liu, S., and Trapnell, C. (2016). Single-cell transcriptome sequencing: recent advances and remaining challenges. F1000Res *5*.

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

Luc, S., Anderson, K., Kharazi, S., Buza-Vidas, N., Boiers, C., Jensen, C.T., Ma, Z., Wittmann, L., and Jacobsen, S.E. (2008). Down-regulation of Mpl marks the transition to lymphoid-primed multipotent progenitors with gradual loss of granulocyte-monocyte potential. Blood *111*, 3424-3434.

Luo, M., Jeong, M., Sun, D., Park, H.J., Rodriguez, B.A., Xia, Z., Yang, L., Zhang, X., Sheng, K., Darlington, G.J.*, et al.* (2015). Long non-coding RNAs control hematopoietic stem cell function. Cell Stem Cell *16*, 426-438.

Lyu, Y., Lou, J., Yang, Y., Feng, J., Hao, Y., Huang, S., Yin, L., Xu, J., Huang, D., Ma, B.*, et al.* (2017). Dysfunction of the WT1-MEG3 signaling promotes AML leukemogenesis via p53-dependent and -independent pathways. Leukemia *31*, 2543-2551.

Maeda, N., Kasukawa, T., Oyama, R., Gough, J., Frith, M., Engstrom, P.G., Lenhard, B., Aturaliya, R.N., Batalov, S., Beisel, K.W.*, et al.* (2006). Transcript annotation in FANTOM3: mouse gene catalog based on physical cDNAs. PLoS Genet *2*, e62.

Malik, B., and Feng, F.Y. (2016). Long noncoding RNAs in prostate cancer: overview and clinical implications. Asian J Androl *18*, 568-574.

Mandel, C.R., Bai, Y., and Tong, L. (2008). Protein factors in pre-mRNA 3'-end processing. Cell Mol Life Sci *65*, 1099-1122.

Mansfield, K.D., and Keene, J.D. (2012). Neuron-specific ELAV/Hu proteins suppress HuR mRNA during neuronal differentiation by alternative polyadenylation. Nucleic Acids Res *40*, 2734-2746.

Martin, G., Gruber, A.R., Keller, W., and Zavolan, M. (2012). Genome-wide analysis of pre-mRNA 3' end processing reveals a decisive role of human cleavage factor I in the regulation of 3' UTR length. Cell Rep *1*, 753-763.

Martin, K.C., and Ephrussi, A. (2009). mRNA localization: gene expression in the spatial dimension. Cell *136*, 719-730.

Masamha, C.P., Xia, Z., Peart, N., Collum, S., Li, W., Wagner, E.J., and Shyu, A.B. (2016). CFIm25 regulates glutaminase alternative terminal exon definition to modulate miR-23 function. RNA *22*, 830-838.

Masamha, C.P., Xia, Z., Yang, J., Albrecht, T.R., Li, M., Shyu, A.B., Li, W., and Wagner, E.J. (2014). CFIm25 links alternative polyadenylation to glioblastoma tumour suppression. Nature *510*, 412-416.

Maxwell, E.S., and Fournier, M.J. (1995). The small nucleolar RNAs. Annu Rev Biochem *64*, 897-934.

Mayr, C. (2017). Regulation by 3'-Untranslated Regions. Annu Rev Genet *51*, 171-194.

McBrayer, S.K., Mayers, J.R., DiNatale, G.J., Shi, D.D., Khanal, J., Chakraborty, A.A., Sarosiek, K.A., Briggs, K.J., Robbins, A.K., Sewastianik, T.*, et al.* (2018). Transaminase Inhibition by 2-Hydroxyglutarate Impairs Glutamate Biosynthesis and Redox Homeostasis in Glioma. Cell *175*, 101-116 e125.

Meier, U.T. (2005). The many facets of H/ACA ribonucleoproteins. Chromosoma *114*, 1-14.

Mendelson, A., and Frenette, P.S. (2014). Hematopoietic stem cell niche maintenance during homeostasis and regeneration. Nat Med *20*, 833-846.

Miao, X. (2013). Recent advances in the development of new transgenic animal technology. Cell Mol Life Sci *70*, 815-828.

Miura, H., Quadros, R.M., Gurumurthy, C.B., and Ohtsuka, M. (2018). Easi-CRISPR for creating knock-in and conditional knockout mouse models using long ssDNA donors. Nat Protoc *13*, 195-215.

Miyamoto, K., Araki, K.Y., Naka, K., Arai, F., Takubo, K., Yamazaki, S., Matsuoka, S., Miyamoto, T., Ito, K., Ohmura, M.*, et al.* (2007). Foxo3a is essential for maintenance of the hematopoietic stem cell pool. Cell Stem Cell *1*, 101-112.

Moignard, V., Macaulay, I.C., Swiers, G., Buettner, F., Schutte, J., Calero-Nieto, F.J., Kinston, S., Joshi, A., Hannah, R., Theis, F.J.*, et al.* (2013). Characterization of transcriptional networks in blood stem and progenitor cells using high-throughput single-cell gene expression analysis. Nat Cell Biol *15*, 363-372.

Moon, Y.S., Smas, C.M., Lee, K., Villena, J.A., Kim, K.H., Yun, E.J., and Sul, H.S. (2002). Mice lacking paternally expressed Pref-1/Dlk1 display growth retardation and accelerated adiposity. Mol Cell Biol *22*, 5585-5592.

Morita, Y., Ema, H., and Nakauchi, H. (2010). Heterogeneity and hierarchy within the most primitive hematopoietic stem cell compartment. J Exp Med *207*, 1173-1182.

Morris, K.V., and Mattick, J.S. (2014). The rise of regulatory RNA. Nat Rev Genet *15*, 423-437.

Morrison, S.J., and Kimble, J. (2006). Asymmetric and symmetric stem-cell divisions in development and cancer. Nature *441*, 1068-1074.

Morrison, S.J., and Scadden, D.T. (2014). The bone marrow niche for haematopoietic stem cells. Nature *505*, 327-334.

Morrison, S.J., Shah, N.M., and Anderson, D.J. (1997). Regulatory mechanisms in stem cell biology. Cell *88*, 287-298.

Morrison, S.J., and Spradling, A.C. (2008). Stem cells and niches: mechanisms that promote stem cell maintenance throughout life. Cell *132*, 598-611.

Morrison, S.J., and Weissman, I.L. (1994). The long-term repopulating subset of hematopoietic stem cells is deterministic and isolatable by phenotype. Immunity *1*, 661-673.

Mueller, A.A., Cheung, T.H., and Rando, T.A. (2013). All's well that ends well: alternative polyadenylation and its implications for stem cell biology. Curr Opin Cell Biol *25*, 222-232.

Mueller, A.A., van Velthoven, C.T., Fukumoto, K.D., Cheung, T.H., and Rando, T.A. (2016). Intronic polyadenylation of PDGFRalpha in resident stem cells attenuates muscle fibrosis. Nature *540*, 276-279.

Muller, A.M., Medvinsky, A., Strouboulis, J., Grosveld, F., and Dzierzak, E. (1994). Development of hematopoietic stem cell activity in the mouse embryo. Immunity *1*, 291-301.

Muller-Sieburg, C.E., Cho, R.H., Thoman, M., Adkins, B., and Sieburg, H.B. (2002). Deterministic regulation of hematopoietic stem cell self-renewal and differentiation. Blood *100*, 1302-1309.

Muller-Sieburg, C.E., Townsend, K., Weissman, I.L., and Rennick, D. (1988). Proliferation and differentiation of highly enriched mouse hematopoietic stem cells and progenitor cells in response to defined growth factors. J Exp Med *167*, 1825-1840.

Muller-Sieburg, C.E., Whitlock, C.A., and Weissman, I.L. (1986). Isolation of two early B lymphocyte progenitors from mouse marrow: a committed pre-pre-B cell and a clonogenic Thy-1-lo hematopoietic stem cell. Cell *44*, 653-662.

Murphy, K., Travers, P., Walport, M., Janeway, C (2008). Janeway's immunobiology, 6th edition. Garland Science, New York.

Nagaike, T., Logan, C., Hotta, I., Rozenblatt-Rosen, O., Meyerson, M., and Manley, J.L. (2011). Transcriptional activators enhance polyadenylation of mRNA precursors. Mol Cell *41*, 409-418.

Nagano, T., Mitchell, J.A., Sanz, L.A., Pauler, F.M., Ferguson-Smith, A.C., Feil, R., and Fraser, P. (2008). The Air noncoding RNA epigenetically silences transcription by targeting G9a to chromatin. Science *322*, 1717-1720.

Nygren, J.M., Bryder, D., and Jacobsen, S.E. (2006). Prolonged cell cycle transit is a defining and developmentally conserved hemopoietic stem cell property. J Immunol *177*, 201-208.

Oburoglu, L., Tardito, S., Fritz, V., de Barros, S.C., Merida, P., Craveiro, M., Mamede, J., Cretenet, G., Mongellaz, C., An, X.*, et al.* (2014). Glucose and glutamine metabolism regulate human hematopoietic stem cell lineage specification. Cell Stem Cell *15*, 169-184.

Ogata, T., Kagami, M., and Ferguson-Smith, A.C. (2008). Molecular mechanisms regulating phenotypic outcome in paternal and maternal uniparental disomy for chromosome 14. Epigenetics *3*, 181-187.

Ogawa, M., Matsuzaki, Y., Nishikawa, S., Hayashi, S., Kunisada, T., Sudo, T., Kina, T., Nakauchi, H., and Nishikawa, S. (1991). Expression and function of c-kit in hemopoietic progenitor cells. J Exp Med *174*, 63-71.

Oguro, H., Ding, L., and Morrison, S.J. (2013). SLAM family markers resolve functionally distinct subpopulations of hematopoietic stem cells and multipotent progenitors. Cell Stem Cell *13*, 102-116.

Olalla, L., Gutierrez, A., Campos, J.A., Khan, Z.U., Alonso, F.J., Segura, J.A., Marquez, J., and Aledo, J.C. (2002). Nuclear localization of L-type glutaminase in mammalian brain. J Biol Chem *277*, 38939-38944.

Orban, P.C., Chui, D., and Marth, J.D. (1992). Tissue- and site-specific DNA recombination in transgenic mice. Proc Natl Acad Sci U S A *89*, 6861-6865.

Orkin, S.H., and Zon, L.I. (2008). Hematopoiesis: an evolving paradigm for stem cell biology. Cell *132*, 631-644.

Orom, U.A., Derrien, T., Beringer, M., Gumireddy, K., Gardini, A., Bussotti, G., Lai, F., Zytnicki, M., Notredame, C., Huang, Q.*, et al.* (2010). Long noncoding RNAs with enhancer-like function in human cells. Cell *143*, 46-58.

Osawa, M., Hanada, K., Hamada, H., and Nakauchi, H. (1996). Long-term lymphohematopoietic reconstitution by a single CD34-low/negative hematopoietic stem cell. Science *273*, 242-245.

Ottersbach, K., and Dzierzak, E. (2005). The murine placenta contains hematopoietic stem cells within the vascular labyrinth region. Dev Cell *8*, 377-387.

Pai, A.A., Baharian, G., Page Sabourin, A., Brinkworth, J.F., Nedelec, Y., Foley, J.W., Grenier, J.C., Siddle, K.J., Dumaine, A., Yotova, V.*, et al.* (2016). Widespread Shortening of 3' Untranslated Regions and Increased Exon Inclusion Are Evolutionarily Conserved Features of Innate Immune Responses to Infection. PLoS Genet *12*, e1006338.

Palade, G.E. (1955). A small particulate component of the cytoplasm. J Biophys Biochem Cytol *1*, 59-68.

Pandey, R.R., Mondal, T., Mohammad, F., Enroth, S., Redrup, L., Komorowski, J., Nagano, T., Mancini-Dinardo, D., and Kanduri, C. (2008). Kcnq1ot1 antisense noncoding RNA mediates lineage-specific transcriptional silencing through chromatin-level regulation. Mol Cell *32*, 232-246.

Park, S.M., Ou, J., Chamberlain, L., Simone, T.M., Yang, H., Virbasius, C.M., Ali, A.M., Zhu, L.J., Mukherjee, S., Raza, A.*, et al.* (2016). U2AF35(S34F) Promotes Transformation by Directing Aberrant ATG7 Pre-mRNA 3' End Formation. Mol Cell *62*, 479-490.

Passegue, E., Wagers, A.J., Giuriato, S., Anderson, W.C., and Weissman, I.L. (2005). Global analysis of proliferation and cell cycle gene expression in the regulation of hematopoietic stem and progenitor cell fates. J Exp Med *202*, 1599-1611.

Paul, F., Arkin, Y., Giladi, A., Jaitin, D.A., Kenigsberg, E., Keren-Shaul, H., Winter, D., Lara-Astiaso, D., Gury, M., Weiner, A.*, et al.* (2015). Transcriptional Heterogeneity and Lineage Commitment in Myeloid Progenitors. Cell *163*, 1663-1677.

Paulsen, M., Takada, S., Youngson, N.A., Benchaib, M., Charlier, C., Segers, K., Georges, M., and Ferguson-Smith, A.C. (2001). Comparative sequence analysis of the imprinted

Dlk1-Gtl2 locus in three mammalian species reveals highly conserved genomic elements and refines comparison with the Igf2-H19 region. Genome Res *11*, 2085-2094.

Pham, C.T., and Ley, T.J. (1999). Dipeptidyl peptidase I is required for the processing and activation of granzymes A and B in vivo. Proc Natl Acad Sci U S A *96*, 8627-8632.

Pickrell, J.K., Marioni, J.C., Pai, A.A., Degner, J.F., Engelhardt, B.E., Nkadori, E., Veyrieras, J.B., Stephens, M., Gilad, Y., and Pritchard, J.K. (2010). Understanding mechanisms underlying human gene expression variation with RNA sequencing. Nature *464*, 768-772.

Pietras, E.M., Reynaud, D., Kang, Y.A., Carlin, D., Calero-Nieto, F.J., Leavitt, A.D., Stuart, J.M., Gottgens, B., and 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.

Pietras, E.M., Warr, M.R., and Passegue, E. (2011). Cell cycle regulation in hematopoietic stem cells. J Cell Biol *195*, 709-720.

Pike, L.S., Smift, A.L., Croteau, N.J., Ferrick, D.A., and Wu, M. (2011). Inhibition of fatty acid oxidation by etomoxir impairs NADPH production and increases reactive oxygen species resulting in ATP depletion and cell death in human glioblastoma cells. Biochim Biophys Acta *1807*, 726-734.

Pinto, P.A., Henriques, T., Freitas, M.O., Martins, T., Domingues, R.G., Wyrzykowska, P.S., Coelho, P.A., Carmo, A.M., Sunkel, C.E., Proudfoot, N.J.*, et al.* (2011). RNA polymerase II kinetics in polo polyadenylation signal selection. EMBO J *30*, 2431-2444.

Proudfoot, N.J., and Brownlee, G.G. (1976). 3' non-coding region sequences in eukaryotic messenger RNA. Nature *263*, 211-214.

Purton, L.E., and Scadden, D.T. (2007). Limiting factors in murine hematopoietic stem cell assays. Cell Stem Cell *1*, 263-270.

Qian, P., He, X.C., Paulson, A., Li, Z., Tao, F., Perry, J.M., Guo, F., Zhao, M., Zhi, L., Venkatraman, A.*, et al.* (2016). The Dlk1-Gtl2 Locus Preserves LT-HSC Function by Inhibiting the PI3K-mTOR Pathway to Restrict Mitochondrial Metabolism. Cell Stem Cell *18*, 214-228.

Quadros, R.M., Miura, H., Harms, D.W., Akatsuka, H., Sato, T., Aida, T., Redder, R., Richardson, G.P., Inagaki, Y., Sakai, D.*, et al.* (2017). Easi-CRISPR: a robust method for one-step generation of mice carrying conditional and insertion alleles using long ssDNA donors and CRISPR ribonucleoproteins. Genome Biol *18*, 92.

Quinn, J.J., and Chang, H.Y. (2016). Unique features of long non-coding RNA biogenesis and function. Nat Rev Genet *17*, 47-62.

Rakoczy, J., Fernandez-Valverde, S.L., Glazov, E.A., Wainwright, E.N., Sato, T., Takada, S., Combes, A.N., Korbie, D.J., Miller, D., Grimmond, S.M.*, et al.* (2013). MicroRNAs-140-5p/140-3p modulate Leydig cell numbers in the developing mouse testis. Biol Reprod *88*, 143.

Raveux, A., Vandormael-Pournin, S., and Cohen-Tannoudji, M. (2017). Optimization of the production of knock-in alleles by CRISPR/Cas9 microinjection into the mouse zygote. Sci Rep *7*, 42661.

Riaz, M., Raz, Y., van Putten, M., Paniagua-Soriano, G., Krom, Y.D., Florea, B.I., and Raz, V. (2016). PABPN1-Dependent mRNA Processing Induces Muscle Wasting. PLoS Genet *12*, e1006031.

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., and Camargo, F.D. (2018). Clonal analysis of lineage fate in native haematopoiesis. Nature *553*, 212-216.

Rosenbloom, K.R., Dreszer, T.R., Long, J.C., Malladi, V.S., Sloan, C.A., Raney, B.J., Cline, M.S., Karolchik, D., Barber, G.P., Clawson, H.*, et al.* (2012). ENCODE whole-genome data in the UCSC Genome Browser: update 2012. Nucleic Acids Res *40*, D912-917.

Rosonina, E., Bakowski, M.A., McCracken, S., and Blencowe, B.J. (2003). Transcriptional activators control splicing and 3'-end cleavage levels. J Biol Chem *278*, 43034-43040.

Rossi, L., Lin, K.K., Boles, N.C., Yang, L., King, K.Y., Jeong, M., Mayle, A., and Goodell, M.A. (2012). Less is more: unveiling the functional core of hematopoietic stem cells through knockout mice. Cell Stem Cell *11*, 302-317.

Sanchez Alvarado, A., and Yamanaka, S. (2014). Rethinking differentiation: stem cells, regeneration, and plasticity. Cell *157*, 110-119.

Sandberg, R., Neilson, J.R., Sarma, A., Sharp, P.A., and Burge, C.B. (2008). Proliferating cells express mRNAs with shortened 3' untranslated regions and fewer microRNA target sites. Science *320*, 1643-1647.

Sander, J.D., and Joung, J.K. (2014). CRISPR-Cas systems for editing, regulating and targeting genomes. Nat Biotechnol *32*, 347-355.

Sanz, E., Yang, L., Su, T., Morris, D.R., McKnight, G.S., and Amieux, P.S. (2009). Cell-type-specific isolation of ribosome-associated mRNA from complex tissues. Proc Natl Acad Sci U S A *106*, 13939-13944.

Sasaki, K., and Matsumura, G. (1988). Spleen lymphocytes and haemopoiesis in the mouse embryo. J Anat *160*, 27-37.

Sauer, B. (1987). Functional expression of the cre-lox site-specific recombination system in the yeast Saccharomyces cerevisiae. Mol Cell Biol *7*, 2087-2096.

Sauer, B., and Henderson, N. (1988). Site-specific DNA recombination in mammalian cells by the Cre recombinase of bacteriophage P1. Proc Natl Acad Sci U S A *85*, 5166-5170.

Schroeder, A., Mueller, O., Stocker, S., Salowsky, R., Leiber, M., Gassmann, M., Lightfoot, S., Menzel, W., Granzow, M., and Ragg, T. (2006). The RIN: an RNA integrity number for assigning integrity values to RNA measurements. BMC Mol Biol *7*, 3.

Schuster-Gossler, K., Bilinski, P., Sado, T., Ferguson-Smith, A., and Gossler, A. (1998). The mouse Gtl2 gene is differentially expressed during embryonic development, encodes multiple alternatively spliced transcripts, and may act as an RNA. Dev Dyn *212*, 214-228.

Scognamiglio, R., Cabezas-Wallscheid, N., Thier, M.C., Altamura, S., Reyes, A., Prendergast, A.M., Baumgartner, D., Carnevalli, L.S., Atzberger, A., Haas, S.*, et al.* (2016). Myc Depletion Induces a Pluripotent Dormant State Mimicking Diapause. Cell *164*, 668-680.

Seita, J., and Weissman, I.L. (2010). Hematopoietic stem cell: self-renewal versus differentiation. Wiley Interdiscip Rev Syst Biol Med *2*, 640-653.

Seitz, H., Royo, H., Bortolin, M.L., Lin, S.P., Ferguson-Smith, A.C., and Cavaille, J. (2004). A large imprinted microRNA gene cluster at the mouse Dlk1-Gtl2 domain. Genome Res *14*, 1741-1748.

Semenza, G.L. (2010). Oxygen homeostasis. Wiley Interdiscip Rev Syst Biol Med *2*, 336-361.

Shemin, D., and Rittenberg, D. (1946). The life span of the human red blood cell. J Biol Chem *166*, 627-636.

Shukla, G.C., Singh, J., and Barik, S. (2011). MicroRNAs: Processing, Maturation, Target Recognition and Regulatory Functions. Mol Cell Pharmacol *3*, 83-92.

Signer, R.A., Magee, J.A., Salic, A., and Morrison, S.J. (2014). Haematopoietic stem cells require a highly regulated protein synthesis rate. Nature *509*, 49-54.

Simsek, T., Kocabas, F., Zheng, J., Deberardinis, R.J., Mahmoud, A.I., Olson, E.N., Schneider, J.W., Zhang, C.C., and Sadek, H.A. (2010). The distinct metabolic profile of hematopoietic stem cells reflects their location in a hypoxic niche. Cell Stem Cell *7*, 380-390.

Singh, I., Lee, S.H., Sperling, A.S., Samur, M.K., Tai, Y.T., Fulciniti, M., Munshi, N.C., Mayr, C., and Leslie, C.S. (2018). Widespread intronic polyadenylation diversifies immune cell transcriptomes. Nat Commun *9*, 1716.

Siomi, M.C., Sato, K., Pezic, D., and Aravin, A.A. (2011). PIWI-interacting small RNAs: the vanguard of genome defence. Nat Rev Mol Cell Biol *12*, 246-258.

Sleutels, F., Zwart, R., and Barlow, D.P. (2002). The non-coding Air RNA is required for silencing autosomal imprinted genes. Nature *415*, 810-813.

Sommerkamp, P., Renders, S., Ladel, L., Hotz-Wagenblatt, A., Schonberger, K., Zeisberger, P., Przybylla, A., Sohn, M., Zhou, Y., Klibanski, A.*, et al.* (2019). The long non-coding RNA Meg3 is dispensable for hematopoietic stem cells. Sci Rep *9*, 2110.

Sommerkamp, P., and Trumpp, A. (2016). Metabolic cues for hematopoietic stem cells. Science *354*, 1103-1104.

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

Spies, N., Burge, C.B., and Bartel, D.P. (2013). 3' UTR-isoform choice has limited influence on the stability and translational efficiency of most mRNAs in mouse fibroblasts. Genome Res *23*, 2078-2090.

Spies, N., Nielsen, C.B., Padgett, R.A., and Burge, C.B. (2009). Biased chromatin signatures around polyadenylation sites and exons. Mol Cell *36*, 245-254.

Suda, T., Takubo, K., and Semenza, G.L. (2011). Metabolic regulation of hematopoietic stem cells in the hypoxic niche. Cell Stem Cell *9*, 298-310.

Sugiyama, T., Kohara, H., Noda, M., and Nagasawa, T. (2006). Maintenance of the hematopoietic stem cell pool by CXCL12-CXCR4 chemokine signaling in bone marrow stromal cell niches. Immunity *25*, 977-988.

Sun, J., Ramos, A., Chapman, B., Johnnidis, J.B., Le, L., Ho, Y.J., Klein, A., Hofmann, O., and Camargo, F.D. (2014). Clonal dynamics of native haematopoiesis. Nature *514*, 322-327.

Sun, Y., Fu, Y., Li, Y., and Xu, A. (2012). Genome-wide alternative polyadenylation in animals: insights from high-throughput technologies. J Mol Cell Biol *4*, 352-361.

Supek, F., Bosnjak, M., Skunca, N., and Smuc, T. (2011). REVIGO summarizes and visualizes long lists of gene ontology terms. PLoS One *6*, e21800.

Taft, R.J., Glazov, E.A., Cloonan, N., Simons, C., Stephen, S., Faulkner, G.J., Lassmann, T., Forrest, A.R., Grimmond, S.M., Schroder, K.*, et al.* (2009). Tiny RNAs associated with transcription start sites in animals. Nat Genet *41*, 572-578.

Taft, R.J., Simons, C., Nahkuri, S., Oey, H., Korbie, D.J., Mercer, T.R., Holst, J., Ritchie, W., Wong, J.J., Rasko, J.E.*, et al.* (2010). Nuclear-localized tiny RNAs are associated with transcription initiation and splice sites in metazoans. Nat Struct Mol Biol *17*, 1030-1034.

Takada, S., Paulsen, M., Tevendale, M., Tsai, C.E., Kelsey, G., Cattanach, B.M., and Ferguson-Smith, A.C. (2002). Epigenetic analysis of the Dlk1-Gtl2 imprinted domain on mouse chromosome 12: implications for imprinting control from comparison with Igf2-H19. Hum Mol Genet *11*, 77-86.

Takagaki, Y., and Manley, J.L. (1998). Levels of polyadenylation factor CstF-64 control IgM heavy chain mRNA accumulation and other events associated with B cell differentiation. Mol Cell *2*, 761-771.

Takagaki, Y., Seipelt, R.L., Peterson, M.L., and Manley, J.L. (1996). The polyadenylation factor CstF-64 regulates alternative processing of IgM heavy chain pre-mRNA during B cell differentiation. Cell *87*, 941-952.

Takahashi, K., Tanabe, K., Ohnuki, M., Narita, M., Ichisaka, T., Tomoda, K., and Yamanaka, S. (2007). Induction of pluripotent stem cells from adult human fibroblasts by defined factors. Cell *131*, 861-872.

Takahashi, K., and Yamanaka, S. (2006). Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. Cell *126*, 663-676.

Takahashi, N., Okamoto, A., Kobayashi, R., Shirai, M., Obata, Y., Ogawa, H., Sotomaru, Y., and Kono, T. (2009). Deletion of Gtl2, imprinted non-coding RNA, with its differentially methylated region induces lethal parent-origin-dependent defects in mice. Hum Mol Genet *18*, 1879-1888.

Takubo, K., Goda, N., Yamada, W., Iriuchishima, H., Ikeda, E., Kubota, Y., Shima, H., Johnson, R.S., Hirao, A., Suematsu, M.*, et al.* (2010). Regulation of the HIF-1alpha level is essential for hematopoietic stem cells. Cell Stem Cell *7*, 391-402.

Takubo, K., Nagamatsu, G., Kobayashi, C.I., Nakamura-Ishizu, A., Kobayashi, H., Ikeda, E., Goda, N., Rahimi, Y., Johnson, R.S., Soga, T.*, et al.* (2013). Regulation of glycolysis by Pdk functions as a metabolic checkpoint for cell cycle quiescence in hematopoietic stem cells. Cell Stem Cell *12*, 49-61.

Taya, Y., Ota, Y., Wilkinson, A.C., Kanazawa, A., Watarai, H., Kasai, M., Nakauchi, H., and Yamazaki, S. (2016). Depleting dietary valine permits nonmyeloablative mouse hematopoietic stem cell transplantation. Science *354*, 1152-1155.

Thomas, K.A., King, R.G., Sestero, C.M., and Justement, L.B. (2016). TREM-like transcript 2 is stored in human neutrophil primary granules and is up-regulated in response to inflammatory mediators. J Leukoc Biol *100*, 177-184.

Tian, B., and Graber, J.H. (2012). Signals for pre-mRNA cleavage and polyadenylation. Wiley Interdiscip Rev RNA *3*, 385-396.

Tian, B., Hu, J., Zhang, H., and Lutz, C.S. (2005). A large-scale analysis of mRNA polyadenylation of human and mouse genes. Nucleic Acids Res *33*, 201-212.

Tian, B., and Manley, J.L. (2017). Alternative polyadenylation of mRNA precursors. Nat Rev Mol Cell Biol *18*, 18-30.

Tian, B., Pan, Z., and Lee, J.Y. (2007). Widespread mRNA polyadenylation events in introns indicate dynamic interplay between polyadenylation and splicing. Genome Res *17*, 156-165.

Till, J.E., and McCulloch, E.A. (1961). A direct measurement of the radiation sensitivity of normal mouse bone marrow cells. Radiat Res *14*, 213-222.

Tothova, Z., and Gilliland, D.G. (2007). FoxO transcription factors and stem cell homeostasis: insights from the hematopoietic system. Cell Stem Cell *1*, 140-152.

Tothova, Z., Kollipara, R., Huntly, B.J., Lee, B.H., Castrillon, D.H., Cullen, D.E., McDowell, E.P., Lazo-Kallanian, S., Williams, I.R., Sears, C.*, et al.* (2007). FoxOs are critical mediators of hematopoietic stem cell resistance to physiologic oxidative stress. Cell *128*, 325-339.

Trumpp, A., Essers, M., and Wilson, A. (2010). Awakening dormant haematopoietic stem cells. Nat Rev Immunol *10*, 201-209.

Tsai, M.C., Manor, O., Wan, Y., Mosammaparast, N., Wang, J.K., Lan, F., Shi, Y., Segal, E., and Chang, H.Y. (2010). Long noncoding RNA as modular scaffold of histone modification complexes. Science *329*, 689-693.

Tuna, M., Knuutila, S., and Mills, G.B. (2009). Uniparental disomy in cancer. Trends Mol Med *15*, 120-128.

Turk, D., Janjic, V., Stern, I., Podobnik, M., Lamba, D., Dahl, S.W., Lauritzen, C., Pedersen, J., Turk, V., and Turk, B. (2001). Structure of human dipeptidyl peptidase I (cathepsin C): exclusion domain added to an endopeptidase framework creates the machine for activation of granular serine proteases. EMBO J *20*, 6570-6582.

van den Heuvel, A.P., Jing, J., Wooster, R.F., and Bachman, K.E. (2012). Analysis of glutamine dependency in non-small cell lung cancer: GLS1 splice variant GAC is essential for cancer cell growth. Cancer Biol Ther *13*, 1185-1194.

van der Wath, R.C., Wilson, A., Laurenti, E., Trumpp, A., and Lio, P. (2009). Estimating dormant and active hematopoietic stem cell kinetics through extensive modeling of bromodeoxyuridine label-retaining cell dynamics. PLoS One *4*, e6972.

Vannini, N., Girotra, M., Naveiras, O., Nikitin, G., Campos, V., Giger, S., Roch, A., Auwerx, J., and Lutolf, M.P. (2016). Specification of haematopoietic stem cell fate via modulation of mitochondrial activity. Nat Commun *7*, 13125.

Velten, L., Anders, S., Pekowska, A., Jarvelin, A.I., Huber, W., Pelechano, V., and Steinmetz, L.M. (2015). Single-cell polyadenylation site mapping reveals 3' isoform choice variability. Mol Syst Biol *11*, 812.

Velten, L., Haas, S.F., Raffel, S., Blaszkiewicz, S., Islam, S., Hennig, B.P., Hirche, C., Lutz, C., Buss, E.C., Nowak, D.*, et al.* (2017). Human haematopoietic stem cell lineage commitment is a continuous process. Nat Cell Biol *19*, 271-281.

Venkatraman, A., He, X.C., Thorvaldsen, J.L., Sugimura, R., Perry, J.M., Tao, F., Zhao, M., Christenson, M.K., Sanchez, R., Yu, J.Y.*, et al.* (2013). Maternal imprinting at the H19-Igf2 locus maintains adult haematopoietic stem cell quiescence. Nature *500*, 345-349.

Wahle, E. (1991). A novel poly(A)-binding protein acts as a specificity factor in the second phase of messenger RNA polyadenylation. Cell *66*, 759-768.

Wahle, E. (1995). Poly(A) tail length control is caused by termination of processive synthesis. J Biol Chem *270*, 2800-2808.

Walter, D., Lier, A., Geiselhart, A., Thalheimer, F.B., Huntscha, S., Sobotta, M.C., Moehrle, B., Brocks, D., Bayindir, I., Kaschutnig, P.*, et al.* (2015). Exit from dormancy provokes DNA-damage-induced attrition in haematopoietic stem cells. Nature *520*, 549-552.

Wang, E.T., Sandberg, R., Luo, S., Khrebtukova, I., Zhang, L., Mayr, C., Kingsmore, S.F., Schroth, G.P., and Burge, C.B. (2008). Alternative isoform regulation in human tissue transcriptomes. Nature *456*, 470-476.

Wang, K.C., Yang, Y.W., Liu, B., Sanyal, A., Corces-Zimmerman, R., Chen, Y., Lajoie, B.R., Protacio, A., Flynn, R.A., Gupta, R.A.*, et al.* (2011). A long noncoding RNA maintains active chromatin to coordinate homeotic gene expression. Nature *472*, 120-124.

Wang, P., Ren, Z., and Sun, P. (2012). Overexpression of the long non-coding RNA MEG3 impairs in vitro glioma cell proliferation. J Cell Biochem *113*, 1868-1874.

Wang, R., Zheng, D., Yehia, G., and Tian, B. (2018). A compendium of conserved cleavage and polyadenylation events in mammalian genes. Genome Res *28*, 1427-1441.

Waskow, C., Madan, V., Bartels, S., Costa, C., Blasig, R., and Rodewald, H.R. (2009). Hematopoietic stem cell transplantation without irradiation. Nat Methods *6*, 267-269.

Wefers, B., Meyer, M., Ortiz, O., Hrabe de Angelis, M., Hansen, J., Wurst, W., and Kuhn, R. (2013). Direct production of mouse disease models by embryo microinjection of TALENs and oligodeoxynucleotides. Proc Natl Acad Sci U S A *110*, 3782-3787.

Weissman, I.L., and Shizuru, J.A. (2008). The origins of the identification and isolation of hematopoietic stem cells, and their capability to induce donor-specific transplantation tolerance and treat autoimmune diseases. Blood *112*, 3543-3553.

Wiedenheft, B., Sternberg, S.H., and Doudna, J.A. (2012). RNA-guided genetic silencing systems in bacteria and archaea. Nature *482*, 331-338.

Wightman, B., Ha, I., and Ruvkun, G. (1993). Posttranscriptional regulation of the heterochronic gene lin-14 by lin-4 mediates temporal pattern formation in C. elegans. Cell *75*, 855-862.

Wilkening, S., Pelechano, V., Jarvelin, A.I., Tekkedil, M.M., Anders, S., Benes, V., and Steinmetz, L.M. (2013). An efficient method for genome-wide polyadenylation site mapping and RNA quantification. Nucleic Acids Res *41*, e65.

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.*, et al.* (2008). Hematopoietic stem cells reversibly switch from dormancy to self-renewal during homeostasis and repair. Cell *135*, 1118-1129.

Wilson, A., Laurenti, E., and Trumpp, A. (2009). Balancing dormant and self-renewing hematopoietic stem cells. Curr Opin Genet Dev *19*, 461-468.

Wilson, A., Oser, G.M., Jaworski, M., Blanco-Bose, W.E., Laurenti, E., Adolphe, C., Essers, M.A., Macdonald, H.R., and Trumpp, A. (2007). Dormant and self-renewing hematopoietic stem cells and their niches. Ann N Y Acad Sci *1106*, 64-75.

Winter, J. (2015). MicroRNAs of the miR379-410 cluster: New players in embryonic neurogenesis and regulators of neuronal function. Neurogenesis (Austin) *2*, e1004970.

Wood, A.J., Schulz, R., Woodfine, K., Koltowska, K., Beechey, C.V., Peters, J., Bourc'his, D., and Oakey, R.J. (2008). Regulation of alternative polyadenylation by genomic imprinting. Genes Dev *22*, 1141-1146.

Wutz, A., Smrzka, O.W., Schweifer, N., Schellander, K., Wagner, E.F., and Barlow, D.P. (1997). Imprinted expression of the Igf2r gene depends on an intronic CpG island. Nature *389*, 745-749.

Xia, Z., Donehower, L.A., Cooper, T.A., Neilson, J.R., Wheeler, D.A., Wagner, E.J., and Li, W. (2014). Dynamic analyses of alternative polyadenylation from RNA-seq reveal a 3'-UTR landscape across seven tumour types. Nat Commun *5*, 5274.

Xiang, Y., Ye, Y., Lou, Y., Yang, Y., Cai, C., Zhang, Z., Mills, T., Chen, N.Y., Kim, Y., Muge Ozguc, F.*, et al.* (2018). Comprehensive Characterization of Alternative Polyadenylation in Human Cancer. J Natl Cancer Inst *110*, 379-389.

Yamamoto, R., Morita, Y., Ooehara, J., Hamanaka, S., Onodera, M., Rudolph, K.L., Ema, H., and Nakauchi, H. (2013). Clonal analysis unveils self-renewing lineage-restricted progenitors generated directly from hematopoietic stem cells. Cell *154*, 1112-1126.

Yamazaki, S., Iwama, A., Takayanagi, S., Morita, Y., Eto, K., Ema, H., and Nakauchi, H. (2006). Cytokine signals modulated via lipid rafts mimic niche signals and induce hibernation in hematopoietic stem cells. EMBO J *25*, 3515-3523.

Yang, L., Venneti, S., and Nagrath, D. (2017). Glutaminolysis: A Hallmark of Cancer Metabolism. Annu Rev Biomed Eng *19*, 163-194.

Yao, C., Choi, E.A., Weng, L., Xie, X., Wan, J., Xing, Y., Moresco, J.J., Tu, P.G., Yates, J.R., 3rd, and Shi, Y. (2013). Overlapping and distinct functions of CstF64 and CstF64tau in mammalian mRNA 3' processing. RNA *19*, 1781-1790.

Ye, C., Long, Y., Ji, G., Li, Q.Q., and Wu, X. (2018). APAtrap: identification and quantification of alternative polyadenylation sites from RNA-seq data. Bioinformatics *34*, 1841-1849.

Ye, C., Zhou, Q., Hong, Y., and Li, Q.Q. (2019). Role of alternative polyadenylation dynamics in acute myeloid leukemia at single-cell resolution. RNA Biol.

Yildirim, E., Kirby, J.E., Brown, D.E., Mercier, F.E., Sadreyev, R.I., Scadden, D.T., and Lee, J.T. (2013). Xist RNA is a potent suppressor of hematologic cancer in mice. Cell *152*, 727-742.

Yu, J., Vodyanik, M.A., Smuga-Otto, K., Antosiewicz-Bourget, J., Frane, J.L., Tian, S., Nie, J., Jonsdottir, G.A., Ruotti, V., Stewart, R.*, et al.* (2007). Induced pluripotent stem cell lines derived from human somatic cells. Science *318*, 1917-1920.

Yu, V.W.C., Yusuf, R.Z., Oki, T., Wu, J., Saez, B., Wang, X., Cook, C., Baryawno, N., Ziller, M.J., Lee, E.*, et al.* (2016). Epigenetic Memory Underlies Cell-Autonomous Heterogeneous Behavior of Hematopoietic Stem Cells. Cell *167*, 1310-1322 e1317.

Yu, W.M., Liu, X., Shen, J., Jovanovic, O., Pohl, E.E., Gerson, S.L., Finkel, T., Broxmeyer, H.E., and Qu, C.K. (2013). Metabolic regulation by the mitochondrial phosphatase PTPMT1 is required for hematopoietic stem cell differentiation. Cell Stem Cell *12*, 62-74.

Yu, Y., Newman, H., Shen, L., Sharma, D., Hu, G., Mirando, A.J., Zhang, H., Knudsen, E., Zhang, G.F., Hilton, M.J.*, et al.* (2019). Glutamine Metabolism Regulates Proliferation and Lineage Allocation in Skeletal Stem Cells. Cell Metab.

Zdzisinska, B., Zurek, A., and Kandefer-Szerszen, M. (2017). Alpha-Ketoglutarate as a Molecule with Pleiotropic Activity: Well-Known and Novel Possibilities of Therapeutic Use. Arch Immunol Ther Exp (Warsz) *65*, 21-36.

Zerbino, D.R., Achuthan, P., Akanni, W., Amode, M.R., Barrell, D., Bhai, J., Billis, K., Cummins, C., Gall, A., Giron, C.G.*, et al.* (2018). Ensembl 2018. Nucleic Acids Res *46*, D754-D761.

Zhang, M., Hutter, G., Kahn, S.A., Azad, T.D., Gholamin, S., Xu, C.Y., Liu, J., Achrol, A.S., Richard, C., Sommerkamp, P.*, et al.* (2016). Anti-CD47 Treatment Stimulates Phagocytosis of Glioblastoma by M1 and M2 Polarized Macrophages and Promotes M1 Polarized Macrophages In Vivo. PLoS One *11*, e0153550.

Zhang, X., Gejman, R., Mahta, A., Zhong, Y., Rice, K.A., Zhou, Y., Cheunsuchon, P., Louis, D.N., and Klibanski, A. (2010a). Maternally expressed gene 3, an imprinted noncoding RNA gene, is associated with meningioma pathogenesis and progression. Cancer Res *70*, 2350-2358.

Zhang, X., Rice, K., Wang, Y., Chen, W., Zhong, Y., Nakayama, Y., Zhou, Y., and Klibanski, A. (2010b). Maternally expressed gene 3 (MEG3) noncoding ribonucleic acid: isoform structure, expression, and functions. Endocrinology *151*, 939-947.

Zhang, X., Zhou, Y., and Klibanski, A. (2010c). Isolation and characterization of novel pituitary tumor related genes: a cDNA representational difference approach. Mol Cell Endocrinol *326*, 40-47.

Zhang, X., Zhou, Y., Mehta, K.R., Danila, D.C., Scolavino, S., Johnson, S.R., and Klibanski, A. (2003). A pituitary-derived MEG3 isoform functions as a growth suppressor in tumor cells. J Clin Endocrinol Metab *88*, 5119-5126.

Zhao, J., Sun, B.K., Erwin, J.A., Song, J.J., and Lee, J.T. (2008). Polycomb proteins targeted by a short repeat RNA to the mouse X chromosome. Science *322*, 750-756.

Zhao, J., Zhang, X., Zhou, Y., Ansell, P.J., and Klibanski, A. (2006). Cyclic AMP stimulates MEG3 gene expression in cells through a cAMP-response element (CRE) in the MEG3 proximal promoter region. Int J Biochem Cell Biol *38*, 1808-1820.

Zhao, Y., Ling, F., Wang, H.C., and Sun, X.H. (2013). Chronic TLR signaling impairs the long-term repopulating potential of hematopoietic stem cells of wild type but not Id1 deficient mice. PLoS One *8*, e55552.

Zheng, D., Wang, R., Ding, Q., Wang, T., Xie, B., Wei, L., Zhong, Z., and Tian, B. (2018). Cellular stress alters 3'UTR landscape through alternative polyadenylation and isoform-specific degradation. Nat Commun *9*, 2268.

Zhou, B.O., Yue, R., Murphy, M.M., Peyer, J.G., and Morrison, S.J. (2014). Leptin-receptor-expressing mesenchymal stromal cells represent the main source of bone formed by adult bone marrow. Cell Stem Cell *15*, 154-168.

Zhou, Y., Cheunsuchon, P., Nakayama, Y., Lawlor, M.W., Zhong, Y., Rice, K.A., Zhang, L., Zhang, X., Gordon, F.E., Lidov, H.G.*, et al.* (2010). Activation of paternally expressed genes and perinatal death caused by deletion of the Gtl2 gene. Development *137*, 2643-2652.

Zhou, Y., Zhang, X., and Klibanski, A. (2012). MEG3 noncoding RNA: a tumor suppressor. J Mol Endocrinol *48*, R45-53.

Zhou, Y., Zhong, Y., Wang, Y., Zhang, X., Batista, D.L., Gejman, R., Ansell, P.J., Zhao, J., Weng, C., and Klibanski, A. (2007). Activation of p53 by MEG3 non-coding RNA. J Biol Chem *282*, 24731-24742.

Zhu, L.J., Gazin, C., Lawson, N.D., Pages, H., Lin, S.M., Lapointe, D.S., and Green, M.R. (2010). ChIPpeakAnno: a Bioconductor package to annotate ChIP-seq and ChIP-chip data. BMC Bioinformatics *11*, 237.

Zipeto, M.A., Court, A.C., Sadarangani, A., Delos Santos, N.P., Balaian, L., Chun, H.J., Pineda, G., Morris, S.R., Mason, C.N., Geron, I.*, et al.* (2016). ADAR1 Activation Drives Leukemia Stem Cell Self-Renewal by Impairing Let-7 Biogenesis. Cell Stem Cell *19*, 177-191.

Zou, Y.R., Kottmann, A.H., Kuroda, M., Taniuchi, I., and Littman, D.R. (1998). Function of the chemokine receptor CXCR4 in haematopoiesis and in cerebellar development. Nature *393*, 595-599.