From Department of Medicine, Huddinge - MedH
Karolinska Institutet, Stockholm, Sweden

HETEROGENEITY AND AGING OF THE HEMATOPOIETIC STEM CELLS

Tsu-Yi Su

Karolinska Institutet

Stockholm 2024

Published by Karolinska Institutet.
Printed by Universitetsservice US-AB, 2024

ISBN 978-91-8017-832-7
DOI https://doi.org/10.69622/27195504
Cover illustration: UMAP of single cell RNA-seq data

Heterogeneity and Aging of the Hematopoietic Stem Cells

Thesis for Doctoral Degree (Ph.D.)

By

Tsu-Yi Su

The thesis will be defended in public at Erna Möllersalen, NEO (Blickagången 16, Huddinge), KI Flemingsberg Campus, December 17th, 2024 at 09:00.

Principal Supervisor:
Assistant Professor Sidinh Luc, Ph.D.
Karolinska Institutet
Department of Medicine, Huddinge
Center for Hematology and Regenerative Medicine (HERM)

Co-supervisors:
Associate Professor Robert Månsson, Ph.D.
Karolinska Institutet
Department of Laboratory Medicine
Division of Clinical Immunology

Assistant Professor Petter S. Woll, Ph.D.
Karolinska Institutet
Department of Medicine, Huddinge
Center for Hematology and Regenerative Medicine (HERM)

Professor Sten Eirik W. Jacobsen, M.D., Ph.D.
Karolinska Institutet
Department of Cell and Molecular Biology
Department of Medicine, Huddinge
Center for Hematology and Regenerative Medicine (HERM)

University of Oxford
MRC Weatherall Institute of Molecular Medicine
Radcliffe Department of Medicine
Nuffield Division of Clinical Laboratory Sciences

Opponent:
Associate Professor Johan Flygare, M.D., Ph.D.
Lund University
Department of Laboratory Medicine
Division of Molecular Medicine and Gene Therapy

Examination Board:
Associate Professor Xingqi Chen, Ph.D.
Uppsala University
Department of Immunology, Genetics and Pathology

Associate Professor Andreas Lennartsson, Ph.D.
Karolinska Institutet
Department of Medicine, Huddinge
Biosciences and Nutrition Unit (BioNut)

Associate Professor Jenny Pyl Hansson, Ph.D.
Lund University
Department of Experimental Medical Science

Thank you!

Popular science summary of the thesis

The blood system is one of the largest systems in the human body and is responsible for various functions such as oxygen transport and immune defense. To ensure the maintenance of the blood system, the cells are continuously sustained by an extremely rare stem cell (SC) population. These cells have two defining properties. First, their ability to divide and produce more of themselves, and second, their ability to give rise to all other cells in the blood system.

Many different types of SCs exist in our body, and within the same kind of SCs, there are various subpopulations. The blood SC population is one of the SC types that have different subpopulations, each producing a preferred type of blood cell. However, even though different blood SCs with various functions have been observed, the reason for such functional differences remains unclear. Blood stem cell functions can only be tested using transplantation assays. To make things more challenging, there are currently no effective methods to identify different blood SCs due to the lack of standardized SC cell surface marker definition. This limitation therefore makes it very difficult to study the mechanisms causing blood SCs to perform different functions, representing a major challenge in the field.

To understand the underlying mechanisms of functionally different SCs in the blood SC pool, it is first necessary to identify cell surface markers that can reliably separate the different SCs. In study I, CD49b has been identified as a cell surface marker that can further separate the conventional blood SC pool into a $CD49b^+$ subpopulation that enriched for SCs that prefer to produce lymphoid cells, and a $CD49b^-$ subpopulation that enriched for SCs that prefer to produce myeloid cells. In addition to differences in their preference to divide to a specific cell lineage, these two subpopulations also showed differences in their DNA accessibility. This led us to hypothesize that the observed differences in DNA accessibility prepare blood SCs to produce their preferred cell types. Although CD49b is a useful marker for enriching different blood SCs, each of the CD49b population still contains other cells. Therefore, in the follow-up study II, CD229 was included as an additional cell surface marker. This allowed for the identification and exclusion of a non-functional SC population, thereby further purifying the SC pool.

Beyond understanding the underlying processes that resulted in functionally different SCs, earlier reports have shown that the blood SC pool expands with age, with an increase in the number of SCs that favor myeloid cell production. However, the cause for this aging phenomenon remains unresolved. As differences between

functionally distinct blood SCs become more prominent with age, we performed molecular analyses to investigate this potential underlying cause. Similar to the findings in study I, we observed distinct DNA accessibility differences between functionally different blood SCs, in addition to aging differences. Furthermore, it is likely that the same protein is responsible for regulating these differences. This suggested that different blood SCs can be distinguished by their DNA accessibility.

In summary, studies I-III have identified reliable cell surface markers that can separate functionally distinct blood SCs. We also demonstrated that differences in chromatin accessibility might be responsible for both blood SC aging and their preference to make different cells.

Abstract

Stem cells are necessary for generating all cells in the body and are defined by two unique properties. First, their multilineage differentiation potential and second, their self-renewal ability.

The hematopoietic stem cells (HSCs) are a rare population of cells that generate all blood cells in the human body and sit at the top of the hematopoietic hierarchy. However, the exact underlying molecular mechanisms that regulate their behavior remain partially obscure, which is a crucial part to elucidating HSC behavior. Previously, it was thought that the HSC population was homogeneous and generated all blood cells without any preference. However, this was disproved by various subsequent studies showing the presence of functionally different HSCs in the HSC pool. This included myeloid-biased (M-bi) HSCs, lymphoid-biased (L-bi) HSCs, and platelet-biased HSCs which show specific lineage differentiation preferences; as well as HSCs with a balanced output. Currently, the molecular mechanisms differentiating lineage-biased HSCs are unclear because they could not be prospectively isolated since the exact immunophenotypic definition of lineage-biased HSCs is not well characterized. As HSCs cannot be characterized based on cell surface markers, transplantation assays have to be done to confirm their long-term (LT) self-renewal and multilineage capability; highlighting the need to identify and understand the factors that discriminate them.

Study I showed that CD49b is a reliable marker to further separate the conventional LT-HSC population into subpopulations that enrich for either the $CD49b^-$ M-bi or the $CD49b^+$ L-bi HSCs. The two CD49b HSC populations differ in their chromatin accessibility even though gene expression comparisons showed no differences. This suggested that different HSCs are already primed to a preferred blood lineage before differentiation. The enrichment of lineage-biased HSCs in both CD49b populations demonstrated the usefulness of adding cell surface markers, though both CD49b subsets remain heterogeneous.

Since residual heterogeneity in the HSC pool was observed in both CD49b subsets from study I, we attempted to further purify the functional LT-HSC population by including CD229 as a cell surface marker in study II. A multipotent progenitor (MPP) population was identified when both CD49b and CD229 were used in combination to subfractionate the LT-HSC pool. This showed that the addition of CD229 can help further purify the heterogeneous HSC population by eliminating the MPPs from the conventional LT-HSC population.

Additional challenges in HSC studies, on top of a heterogeneous HSC population, is the reduction in hematopoietic system function with age, which leads to an overall decrease in HSC fitness and a higher chance of developing hematopoietic malignancies. One of the characteristics of an aging hematopoietic system is the increase in M-bi HSCs and a decrease in L-bi HSCs, but the underlying cause for such a shift in the HSC pool remains unclear.

Therefore, to identify key regulators in directing HSC lineage bias and changes in HSCs with age, we studied the aging of CD49b HSC subpopulations in study III. From the *in vivo* studies, it was found that with age, both CD49b HSC subsets become more myeloid-biased. From the molecular studies, it was found that chromatin accessibility increases with age, that the additional accessibility gained in old age was *de novo*, and this gain of chromatin accessibility progresses through the adult age. Furthermore, it was suggested that the same transcriptional factor might be regulating both HSC aging and lineage bias.

Taken together, studies I-III have shown that by adding cell surface markers in HSC selection, the HSC pool can be further purified and HSC subsets enriching for different lineage biases can be identified. Even though the subsets cannot be distinguished on the transcriptomic level, they show chromatin accessibility differences. In addition, we have found a potential transcription factor that is responsible for both HSC aging and lineage bias. These studies, taken together, not only help the field in understanding more about HSCs, but also provide the basis for future translational studies and therapeutic research.

List of scientific papers

I. Ece Somuncular, Julia Hauenstein, Prajakta Khalkar, Anne-Sofie Johansson, Özge Dumral, Nicolai S. Frengen, Charlotte Gustafsson, Giuseppe Mocci, Tsu-Yi Su, Hugo Brouwer, Christine L. Trautmann, Michael Vanlandewijck, Stuart H. Orkin, Robert Månsson, and Sidinh Luc. **CD49b identifies functionally and epigenetically distinct subsets of lineage-biased hematopoietic stem cells.** *Stem Cell Reports*. 2022 Jul 12;17(7). doi: 10.1016/j.stemcr.2022.05.014

II. Ece Somuncular*, Tsu-Yi Su*, Özge Dumral, Anne-Sofie Johansson, and Sidinh Luc. **Combination of CD49b and CD229 reveals a subset of multipotent progenitors with short-term activity within the hematopoietic stem cell compartment.** *Stem Cells Transl Med*. 2023 Nov 3;12(11). doi: 10.1093/stcltm/szad057

III. Tsu-Yi Su*, Julia Hauenstein*, Ece Somuncular*, Özge Dumral, Elory Leonard, Charlotte Gustafsson, Efthymios Tzortzis, Aurora Forlani, Anne-Sofie Johansson, Hong Qian, Robert Månsson, and Sidinh Luc. **Aging is associated with functional and molecular changes in distinct hematopoietic stem cell subsets.** *Nat Commun*. 2024 Sep 11;15(1):7966. doi: 10.1038/s41467-024-52318-1

*Equal contribution

Contents

List of abbreviations

ATAC-seq	Assay for transposase-accessible chromatin using sequencing
BM	Bone marrow
CD	Cluster of differentiation
CHIP	Clonal hematopoiesis of indeterminate potential
CLP	Common lymphoid progenitor
CMP	Common myeloid progenitor
DAR	Differentially accessible region
FACS	Fluorescence-activated cell sorting
FL	Fetal liver
FMO	Fluorescence minus one
GMP	Granulocyte-monocyte progenitor
HSC	Hematopoietic stem cell
HSPC	Hematopoietic stem- and progenitor cell
L-bi	Lymphoid-biased
L/M ratio	Lymphoid to myeloid cell ratio
LMPP	Lymphoid-primed multipotent progenitor
LT	Long-term
M-bi	Myeloid-biased
MPP	Multipotent progenitor
MSC	Mesenchymal stem cell
NK	Natural killer
PB	Peripheral blood
PCA	Principal component analysis
PCR	Polymerase chain reaction
pHSC	Primed hematopoietic stem cell
QC	Quality check

rHSC	Reserve hematopoietic stem cell
SC	Stem cell
scRNA-seq	Single cell RNA-sequencing
SP	Side population
ST	Short-term
TF	Transcription factor

Introduction

The hematopoietic system is one of the largest systems in the human body. It is responsible for essential physiological processes such as oxygen transport and immune defenses. A population of rare hematopoietic stem cells (HSCs) was discovered in the early 1960s by McCulloch and Till[1]. Since then, there has been a continuous interest in HSCs and their underlying regulatory mechanisms. Hematopoietic stem cells possess the unique multilineage differentiation potential and the ability to self-renew. However, it was unclear how they are being regulated. In addition, subsequent studies have shown that the HSC population comprises of functionally different HSCs. However, there is currently no standardized immunophenotype to prospectively isolate functional LT-HSCs. This makes mechanistic studies challenging, and further confounded by the presence of functionally different lineage-biased HSCs in the stem cell pool. As a consequence, what is known from current studies on HSC mechanisms is a summation of all the different cells in the HSC pool, rather than targeted studies on distinct HSC populations.

Furthermore, the HSC population changes with age, including an expansion of the HSC pool, along with an increase in myelopoiesis. This highlights the importance of understanding the underlying differences between lineage-biased HSCs in order to begin to unravel the cause for such differences.

The studies carried out in this thesis aimed to address current challenges in the field by elucidating the underlying mechanisms governing functional differences between lineage-biased HSCs, and how they change with age. To achieve that, reliable cell surface markers that can further purify functional LT-HSCs and identify functionally distinct HSCs have been identified in studies I-II; while study III followed up on how lineage-biased HSCs functionally and molecularly change with age. These studies would provide new information to future studies on lineage-biased HSCs and are clinically relevant for treating different leukemias and HSC malignancies.

To summarize, **studies I–III** addressed the challenges currently being faced by the field and this thesis will include discussions on how they fit into the larger context. These studies will also be evaluated from both the technical and conceptual perspectives, in addition to discussing them in the future perspectives.

1 Background

1.1 The hematopoietic system

The hematopoietic system is responsible for all the blood cells in an individual. It is well characterized that the hematopoietic stem cells (HSCs), which reside at the apex of the hematopoietic hierarchy, differentiate and give rise to all downstream progenitor and mature cells with continuous replenishment to the hematopoietic system[2].

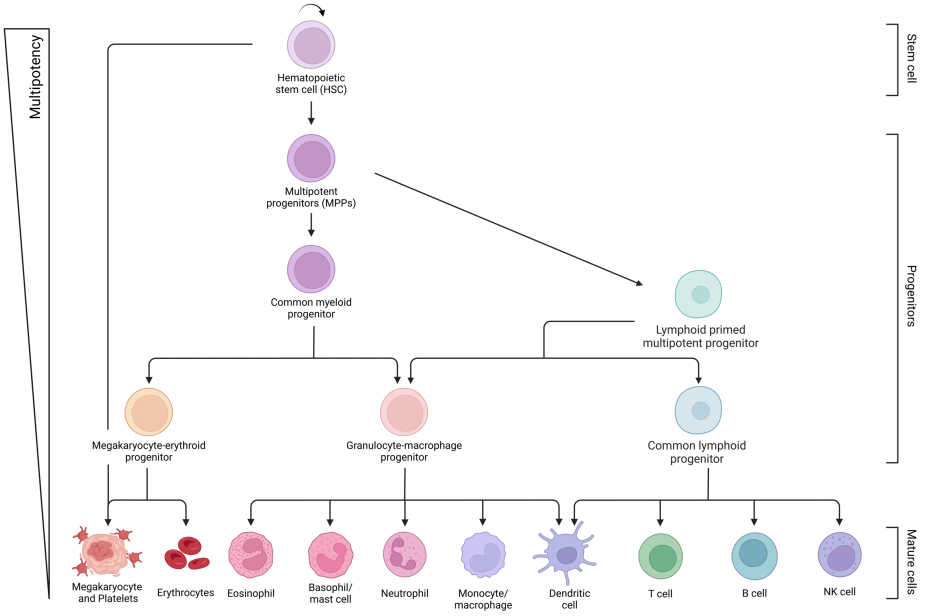

Figure 1. A simplified hematopoietic hierarchy. Image created with Biorender.

The classical model of the hematopoietic hierarchy recognizes that the HSCs differentiate into the common myeloid progenitors (CMPs) and the common lymphoid progenitors (CLPs)[3]. The CMPs give rise to megakaryocytes and their progeny, platelets, that are involved in blood clotting or coagulation[4]. CMPs also give rise to erythrocytes that deliver oxygen and scavenge for nitric oxide[5,6]. Additionally, CMPs give rise to mast cells that are involved in innate and adaptive immunity responses, allergies, and asthma[7]; as well as granulocytes, monocytes, and macrophages involved in response to tissue damages and infections[8]. The CLPs give rise to the natural killer (NK) cells involved in infections and tumor formation response[9]. Moreover, CLPs give rise to T cells for immune response and B cells for antibody production[10,11]. Dendritic cells are also part of the

hematopoietic system and are involved in the innate immune response by presenting antigens to lymphocytes[12].

As more insights are shed from ensuing studies, the hematopoietic hierarchy is constantly being revised. Differentiation from HSCs to terminally differentiated mature blood cells goes through multipotent, oligopotent, and unipotent progenitors with progressively restricted lineage fates and reduced self-renewal ability. In the most primitive pool of cells in the hierarchical tree, cells can be grouped as long-term (LT), intermediate-term, and short-term (ST) HSCs based on their ability to self-renew[13–15]. They are followed by multipotent progenitors (MPPs), granulocyte-monocyte progenitors (GMPs), CLPs, and lymphoid-primed multipotent progenitors (LMPPs)[16,17]. In the latest model, ST-HSCs are termed MPP1, which retain their multipotency but have reduced self-renewal ability. MPP2 shows a differentiation bias towards the myeloid, erythrocyte, and platelet lineages. MPP3 differentiates preferentially to the granulocytes and monocytes. MPP4, also called the LMPPs, are lymphoid primed but can give rise to myeloid and lymphoid cells equally well[16,18]. MPPs 5 and 6, which were identified recently, lie between MPP1 and MPPs 2-4. MPP5 is suggested to be an HSC reserve that is activated in emergency hematopoiesis, and showed LT lymphoid lineage reconstitution, therefore is more potent than MPPs 2-4 but less than MPP1. MPP6 are ST-HSCs that are hierarchically above MPP5 and exhibit multilineage differentiation potential[19]. However, no long-term *in vivo* analyses or serial transplantation studies were done with MPPs 5 and 6, therefore, their exact functions remain to be characterized. Regardless, with new studies revealing insights into the hematopoietic system, the model is constantly refined to help understand the intricacy of the hematopoietic tree.

1.2 Hematopoietic stem cells

The ability of the hematopoietic system to be self-replenishable was first observed when transplantation of adult bone marrow (BM) into lethally irradiated mice rescued the hematopoietic system by Jacobson and colleagues in 1951[20]. Subsequent limiting dilution assays attempting to identify the minimum number of BM cells required to repopulate the hematopoietic system have revealed a rare population of cells that is responsible for maintaining the BM in the mouse, implying the presence of HSCs[1].

Hematopoietic stem cells are a population of rare adult SCs (0,005% - 0,01% of all nucleated BM cells) that can self-renew and have multilineage differentiation

potential, thereby giving rise to all cell types in the hematopoietic system[21]. With the establishment of the hematopoietic hierarchy model, pursuits to identify and understand the rare HSC population that sits on the top of the tree have been underway. Currently, transplantation is the only way to functionally distinguish HSCs *in vivo*, where transplanted cells are determined based on their self-renewal ability and multilineage potential. Therefore, all studies in the field are retrospective as current available HSC markers can only enrich for functional HSCs, and they yet have a definitive immunophenotype.

1.2.1 Strategies to identify HSCs

To discern the different cell types in the hematopoietic system, fluorescence-activated cell sorting (FACS) is used to detect diverse cluster of differentiation (CD) molecules present on the surface of hematopoietic cells. Nakauchi's group demonstrated that negative CD34 expression can reliably further enrich for the HSC population in the mouse after selecting for the Lineage$^-$ Sca-1$^+$ c-kit$^+$ (LSK) population[22]. Similarly, Thy-1.1 was proposed to be useful in distinguishing HSCs and progenitors[23]. Flt3 was later found to be capable of further separating HSC and progenitors based on Flt3 expression level[24,25].

Non-cell surface marker identifications of HSCs have also been employed to distinguish HSCs using dyes marking intracellular activity. Hoechst staining was used to identify a rare side population (SP) that has low Hoechst fluorescence. The SP cells constitute 0.1% of the total BM population and are enriched for multilineage HSCs in transplantation experiments[26,27]. Similarly, the use of rhodamine-123, assessing mitochondria activity, was used to identify HSCs which are low on rhodamine but with multilineage potential[28].

However, it was the identification of SLAM family receptor markers (CD150, CD48) that helped the field make a significant step to further purify the HSC population, where HSCs and MPPs can be separated using SLAM marker expression levels[29,30]. Since then, the field has widely adopted the definition of LSK CD48$^-$ CD34$^-$ CD150$^+$ for HSCs[3,31]. Despite so, these markers still cannot isolate functional HSCs to absolute purity, because not all transplanted cells are able to repopulate the hematopoietic system in recipient mice[32]. The current marker combinations can only enrich the HSC population for up to 40-50% purity[33].

1.2.2 Heterogeneity within the HSC pool

In addition to the HSC pool being heterogeneous, the Müller-Sieburg group has found that when carrying out whole BM bulk cell transplantation experiments, there are HSC subpopulations with a skewed myeloid to lymphoid cell output in their reconstitution patterns. Secondary transplantation experiments further showed that the reconstitution patterns persist, suggesting that HSC lineage bias is predetermined and stable through life[34]. More recent evidence has indicated that HSC lineage predetermination happens already during developmental hematopoiesis[35–37]. With the adoption of SLAM markers in identifying the HSC population, the Nakauchi group has found that, by performing transplantation studies, CD150 expression levels can further distinguish HSCs of different self-renewal ability and lineage differentiation. HSCs with $CD150^{high}$ ($CD150^{hi}$) expression showed the most long-term self-renewal ability, as well as a preference towards myeloid cells. HSCs with $CD150^{int}$ expression showed a balanced lineage, while HSCs with $CD150^{-}$ expression showed the lowest self-renewal ability and a preference for lymphoid cells. These lineage bias patterns continue during serial transplantation, further supporting earlier reports of the persistence of HSC lineage bias[14,38].

The Goodell group has further found that within the SP cells, low SP are more myeloid biased (M-bi) and high SP cells are more lymphoid biased (L-bi)[39]. The Eaves group has combined rhodamine dye and $lineage^{-}$ cell surface markers to further enrich for HSCs in the SP. From their serial transplantation studies, they have categorized the HSC population into 4 subpopulations: α-, β-, γ-, and δ-cells. α- and β-cells showed robust self-renewal ability where α-cells showed the most M-bi lineage output and β-cells showed a balanced lineage output. Both γ- and δ-cells are mainly lymphoid-restricted cells with the least self-renewal ability[14,35,40]. Myeloid-biased HSCs showed more long-term reconstituting abilities than L-bi HSCs and exhibited LT-HSC characteristics. Lymphoid-biased HSCs are more associated with ST-HSCs with reduced self-renewal ability[14,38]. More recently, platelet-biased HSCs have also been identified, further showing the heterogeneity within the HSC pool[41]. Moreover, the Li group has identified 2 different populations of HSCs, LT reserve HSCs (rHSCs), and primed HSCs (pHSCs) that are transient cells between LT rHSCs and ST-HSCs[42]. They showed that during chemotherapy and irradiation, the pHSCs were abolished. However, the rHSCs persisted and repopulated and regenerated the HSC population in the hematopoietic system post-irradiation[42].

As new insights are being shed with more HSC studies, it is evident that HSC heterogeneity is a complex topic that needs to be resolved to study lineage-biased HSC functions. As prospective isolation of a purified lineage-biased HSC pool is currently not possible, there are continuous efforts to better separate the HSC pool and define lineage-biased HSCs based on cell surface markers such as CD61, CD229, and CD49b[30,43,44].

1.3 Hematopoiesis

Native hematopoiesis, also termed steady state hematopoiesis, is how HSCs behave *in vivo* without any external perturbation. Steady state hematopoiesis studies are challenging due to difficulties in identifying HSCs *in situ* and monitoring them through differentiation. Therefore, the gold standard of studying HSC function is through transplantation assays to assess for HSC self-renewal and regenerative abilities, as well as their homing ability. Though transplantation is the most conventional method, it might not faithfully represent what happens in steady state hematopoiesis. Firstly, because HSCs are no longer in their quiescent state and are driven to divide and differentiate. Secondly, the HSCs no longer reside in a homeostatic environment as they are being transplanted into irradiated recipients that are ablated of niche cells important for homing and cell-cell communication. Thirdly, in transplantation experiments, a limited number of HSCs are put under stress to regenerate the hematopoietic system. This is unlike an unperturbed setting, where there is a larger pool of HSCs maintaining hematopoietic homeostasis.

To study native hematopoiesis, several tools have been developed to circumvent perturbing HSCs from their quiescent state. *In situ* labeling, fate-mapping, transposon tagging, and genetic barcoding have been employed to follow HSC differentiation over time by utilizing Cre recombinase activities that are active under cell specific promoters[45–50]. Using these tools, some argued that the HSCs are largely not involved in native hematopoiesis, but rather, homeostasis of the hematopoietic system is maintained by MPPs[51,52]. However, others have shown that there is a robust contribution from the HSCs during native hematopoiesis[53]. These conflicting results might have occurred because first, HSC labeling efficiency is not absolute, which might not be reflective of the whole HSC pool. Second, labeling of the cells will be different depending on the specificity and strength of different Cre activities.

Using both *Flt3–Cre* and von Willebrand factor (*vWF)–eGFP* reporter have been successful in the identification of platelet-restricted HSCs in the Jacobsen group[54,55]. However, there remains room for further optimization to examine lineage-biased HSCs in hematopoiesis[56]. With the contribution of HSCs during native hematopoiesis from lineage tracing and fate mapping studies still under discussion, transplantation is still the most used method to prospectively study HSCs and their progenitors in hematopoiesis.

1.3.1 Fetal liver transition to adult HSC

Before HSCs obtain their characteristic quiescent nature, they undergo dynamic changes in their cellular activities from development to the adult stage. Studies done by the Eaves group have demonstrated that mice up to 3 weeks post birth showed active HSC cycling but the cells then become quiescent within a week[57,58]. This period of drastic change in HSC activity has been termed the HSC fetal liver (FL)-to-adult transition. Not only is there a change in cell cycling activities, but there are also differences between FL HSCs and adult BM HSCs in their engraftment ability[58]. This indicated that HSCs undergo major intrinsic reforms from the embryonic to the FL state, and eventually remain quiescent starting from adult to aged HSCs[59].

1.3.2 Aging hematopoiesis

Hematopoiesis alters in response to changes in microenvironment and the physiology of an individual. Consequences of an aging hematopoietic system include anemia and dysregulated immune system[60]. Hematopoietic stem cell number increases with age, but HSCs become more quiescent, divide less, and proliferate less[61–63]. Old HSCs also show functional defects including reduced engraftment and self-renewal ability[64]. It has been suggested that the increase in HSC number is to compensate for the reduced stem cell function and to maintain the overall homeostasis of the hematopoietic system[65].

The increase in HSC number observed with age is mostly due to an increased number of M-bi HSCs, leading to increased myelopoiesis[66]. The myeloid-skewed phenomenon makes an individual more susceptible to clonal hematopoiesis (CHIP) where HSCs gain mutations that give them a selective advantage. This is also observed clinically where older individuals with CHIP are more predisposed to myeloid malignancies[67].

In addition to increased M-bi HSCs in the HSC pool, there is decreased lymphopoiesis with age. Both extrinsic and intrinsic mechanisms have been suggested to influence aging. For studies on the role of extrinsic factors in HSC aging, it was proposed that L-bi HSCs do not show a decreased lymphoid differentiation potential with age, and both young and aged HSCs showed similar lymphoid differentiation capabilities. Some found that the proportion of L-bi HSCs does not decrease and factors such as niche microenvironment, cytokine signaling, cell-cell communication, and inflammation are responsible for the reduced lymphoid output in the hematopoietic system with age[68,69].

Intrinsic factors such as elevated reactive oxygen species levels, increased DNA damage, increased mitochondrial activity, and dysregulated autophagic activity, mark the aging phenotypes within an HSC[28,70–72]. Molecularly, transcription factor (TF) PU.1 activity has been proposed to be high in young HSCs and kept low in old HSCs via epigenetic regulation[73]. Likewise, the downstream TGF-β signaling pathway, involved in hematopoiesis, has been reported to be downregulated in aged HSCs, with a concomitant reduction in Egr1 TF expression[74]. It was also found that during aging, a set of lymphoid genes is downregulated while a set of myeloid genes is upregulated in aged HSCs[75]. Furthermore, data from single cell RNA-sequencing have shown that changes in cell proliferation and cell cycle were found to be intrinsically altered with age[75,76]. Currently, there are efforts to collate a library of aging signature genes by comparing transcriptomic analyses across different studies[77].

Despite the known increase in myelopoiesis with age from aging studies on the composite HSC population, heterogeneity within the HSC pool has not been extensively addressed in the context of aging since it is not possible to segregate lineage-biased HSCs.

1.4 Molecular regulation of the HSC

1.4.1 Gene expression and epigenetic mechanisms

Despite exhibiting a lineage-biased behavior *in vivo*, HSCs are still capable of giving rise to all lineages when cultured *in vitro*[41]. This suggests that the intrinsic primed lineage-biased HSCs retain their multilineage plasticity and can change their behavior when need be[78]. Molecularly, this strongly implies that epigenetic regulation is at play in HSC lineage bias due to the flexible differentiation ability of lineage-biased HSCs[79]. However, with the immunophenotype of lineage-biased HSCs yet to be fully characterized and transplantation being the only way to

assess HSC lineage bias, molecular mechanisms underlying differences among HSCs in their lineage differentiation preference are still poorly understood.

Epigenetic regulation is the regulation of the chromatin landscape by histone protein modification, DNA methylation, or chromatin remodeling that causes reversible changes in gene expression without altering the DNA sequence[80]. Transcription factor accessibility to transcription start sites and subsequent downstream gene expression are dependent on DNA methylation state and how tightly the DNA is wound to the histone proteins[81].

Epigenetic regulation is a proposed molecular mechanism in HSC lineage bias[82]. However, as the immunophenotypes of lineage-biased HSCs are still to be defined, studies nowadays are correlative because of two main reasons. First, due to the lack of prospective markers for lineage-biased HSC, it is not possible to know whether the studied cell is a lineage-biased HSC until transplantation has been carried out. Second, the same cell that is used in functional assays is not the same cell that is being used in molecular analysis. Despite so, chromatin remodeling protein SATB1 has been identified to be involved in priming the HSCs towards the lymphoid lineage as $SATB1^+$ HSCs have a strong preference to produce the B cell lineage[83]. Sirtuin 1, SIRT1, is an NAD-dependent deacetylase and another epigenetic regulator in HSC lineage bias which was found to lead to increased myelopoiesis when lost, with an acquisition of an aging phenotype[84]. Reduced PU.1 expression was suggested to lead to a decrease in B lymphocyte development and have positive feedback with cell cycle states[85]. However, the direct cause-effect relationship of how epigenetic regulators influence HSC lineage bias remains postulative.

Looking back at the very first attempt to understand the underlying processes in HSCs using single cell polymerase chain reaction (PCR), current technologies have allowed us to use more sophisticated approaches to address the question[86]. By performing the assay for transposase-accessible chromatin using sequencing (ATAC-seq) and single cell RNA-seq (scRNA-seq) methods, a few groups have shown that lineage bias could be assessed by the chromatin accessibility state exhibited by different lineage-biased HSCs[44,87]. Scadden and colleagues have aimed to directly address the question by following labeled cells *in situ* and assess their self-renewal and regenerative potentials long-term, before performing genomic profiling to ensure that the molecular characterization done on the cells

can be compared to their *in vivo* function[88]. This method provides a possibility for both functional and molecular characterization to be done on the same HSC clone.

Therefore, understanding the role that epigenetic regulation plays in influencing downstream gene expression and signaling pathways is crucial in identifying key regulators in these processes. It will also serve as a good platform for the initial steps in paralleling murine and human physiology for future translational studies.

1.4.2 HSC niche and cell signaling

Aside from HSC intrinsic processes and regulation, the HSC niche in the BM also plays a significant role in HSC cell-cell communication and signaling. The HSC niche is a specialized hypoxic microenvironment where HSCs reside[89]. It is made up of various cells including stromal cells, blood vessels, neural cells, and the extracellular matrix[42,90,91]. Furthermore, it was suggested that the LT rHSCs are localized in the endosteal region of the bone marrow and maintained by N-cadherin expressing cells, while the pHSCs reside in the perivascular region[42,92]. The crosstalk between the various cell types in the niche is essential to maintaining HSCs function and their downstream progeny[93]. Several signaling pathways have been shown to be crucial in HSC maintenance and function, including the Notch, Wnt, hedgehog, and TGF-β signaling pathways[94–96]. Disruptions in signaling between the niche and the HSCs have been shown to impact HSC function and were associated with leukemia[97,98].

1.5 Human lineage-biased HSCs and disease association

With lineage-biased HSCs also implicated in the human HSC pool[99–101], similar attempts to purify the HSC population to resolve their functional heterogeneity have been ongoing. However, it is more challenging than the mouse model due to less well-defined and undercharacterized human HSC immunophenotypes, as well as species-to-species variation in genetics[102]. This hinders progress in studying lineage-biased HSCs in human, and the underlying molecular mechanisms of human HSC lineage bias. Despite so, there are evolutionarily conserved behaviors between mouse and human in that the HSC pool showed reduced lymphopoiesis, increased bone marrow cellularity, and a higher risk of myeloid malignancies with age[100,103–105].

With declining HSC self-renewal and regenerative ability, and an increase in the frequency of M-bi HSCs, coupled with higher incidence of CHIP in aged individuals, there are still yet studies directing a relationship between HSC lineage bias and

disease development[101]. Given the presence of lineage-biased HSCs, the phenotypes of different hematopoietic malignancies may be due to a difference in cellular origin. For example, it has been suggested that platelet-biased HSCs are the target HSC in JAK^{V617F} myeloproliferative neoplasms[106]. Understanding the cell of origin in hematologic diseases may provide opportunities to explore lineage-biased HSCs for gene therapy[107].

Therefore, to begin to understand the role of lineage-biased HSCs in malignancies, it is necessary to characterize them in order to understand the cell of origin for disease development. Currently, disease treatments targeting epigenetic regulators in HSCs have yet to consider lineage-biased HSC of origin for disease development[108,109]. Therefore, more extensive epigenetic profiling of mouse lineage-biased HSCs will provide insight into their underlying mechanism, while profiling the human lineage-biased HSCs will give a comprehensive overview of the connections between HSC lineage bias and disease development.

2 Research aims

The overall research aim of the thesis is to evaluate cell surface markers in their efficiency to prospectively identify lineage-biased HSC in the heterogeneous LT-HSC pool, as well as to understand the cellular and molecular differences between the different populations. Further, the thesis aims to investigate how aging influences functional and molecular changes of lineage-biased HSCs. The specific aims for the individual studies are:

Study I – To evaluate CD49b as a marker to further subfractionate the LT-HSC population, and characterize underlying molecular mechanisms of functionally different HSC subsets.

Study II – To explore the use of CD49b and CD229 as a marker combination to further purify the HSC population.

Study III – To investigate functional and molecular changes in lineage-biased HSCs during aging.

3 Materials and methods

Detailed experimental procedures and materials specific to individual studies can be found in respective publications.

3.1 Ethical considerations

Approval by the regional ethical committee, Linköping ethical committee in Sweden, was obtained prior to any mouse work. All studies were carried out under the ethical permit Dnr 882 (2017/06/01 – 2022/06/01). Study III was additionally carried out under Dnr 02250-2022 (2022/03/03 – 2027/03/03).

All personnel had to have the required certification in laboratory animal science before being allowed to do any animal work by law. For study-specific procedures, new members were trained and assessed by experienced group members before being allowed to perform procedures independently in the animal facility.

Because *in vitro* experiments cannot fully recapitulate the *in vivo* 3D structure and cell-cell communication and signaling, the majority of the assays were carried out using mice. Furthermore, transplantation is the only method to assess for functional HSCs, leaving any substitutions unfeasible.

However, protocol optimization discussions were carried out continuously so a minimal number of mice to achieve statistically reliable data were included for each experiment to follow the 3Rs (replace, reduce, refine). The number of mice used was reduced by combining experiments whenever applicable and mice breeding was also planned strategically to prevent an unnecessary excessive number of animals.

To comply with the animal welfare guidelines, 2-5 animals were housed in the same cage and single housed animals were avoided as much as possible. Animal well-being was assessed according to the Karolinska Institutet Veterinary guidelines and plan.

3.2 Mouse models

Since transplantation is the only method to assess HSC function, it is crucial to distinguish between donor and recipient cells during the assessment of these experiments. Therefore, either the C57BL/6J or the Gata1-eGFP (CD45.2) mouse strain was used as donors, and the B6.SJL-*Ptprc*^{a}*Pepc*^{b}/BoyJ (CD45.1) mouse strain was used as recipients. The Gata1-eGFP mouse strain is advantageous as it

allows the detection of donor platelets and erythrocytes, which would otherwise be unmeasurable with the regular CD45.1/2 congenic mouse models. Mice were obtained either from the internal breeding of the animal facility or bought from the Jackson Laboratory and Charles River.

3.3 Isolation and preparation of hematopoietic cells

FACS buffer (Dulbecco's phosphate buffered saline + 2-5% fetal bovine serum + 2mM Ethylenediaminetetraacetic acid (EDTA)) was used to process the samples.

To assess the donor reconstitution and myeloid activity, tail vein bleeding was performed on recipient mice during routine peripheral blood (PB) analysis. Either heparin-coated or EDTA-coated collection tubes were used for PB collection to prevent blood clots. The platelet fraction was first separated from the leukocyte fraction using centrifugal spinning. Following platelet isolation, samples were diluted in 2% FACS buffer, mixed with 2% dextran, and incubated at 37°C to allow the separation of erythrocytes and leukocytes. The leukocyte fraction was further treated with ammonium chloride (NH_4Cl) to lyse residual erythrocytes.

To assess donor myeloid and HSC reconstitution, hind- and forelimbs were isolated and crushed into single cell suspensions using a pestle and mortar in terminal BM analysis. The single cell suspension was then passed through a 70μm filter. Unfractionated BM was counted using the XP-300-Hematology Analyzer. To enrich for the HSC population, CD117 (c-kit) MicroBeads enrichment was carried out and c-kit expressing cells were selected using immunomagnetic separation prior to antibody staining.

All samples (except platelet samples) were Fc-receptor blocked with either purified CD16/32 or fluorophore conjugated CD16/32, for at least 10 minutes. Following Fc-block, cells were stained with a mixture of fluorophore-conjugated monoclonal antibodies against cell surface marker antigens for 15 minutes at 4°C, in the dark. Bone marrow samples were filtered through a 50μm filter prior to sample run to prevent clogging during analyses.

3.4 Flow cytometry and cell sorting

Samples were analyzed on FACSymphony™ A5 and LSR Fortessa™ following cell preparation for phenotypic analyses. FACSAria™ Fusion cell sorters were used to sort HSCs for *in vitro* and functional tests. Compensation controls and fluorescence minus one (FMO) controls were included for every experiment. To

validate single cell experiments, 488-nm beads were sorted into individual wells and evaluated by the visualization of 1 bead/well using a fluorescence microscope both pre- and post-sort. Alternatively, a colorimetric assay was carried out by using the horseradish peroxidase driven oxidation of the TMB substrate, resulting in color change in the wells. For post-acquisition analyses, Flowjo software v10 was used. To gate and double check our HSC populations of interest, FMOs, and back-gating were performed.

3.5 HSC *in vitro* tests

To evaluate the physiological role that CD49b plays as an integrin in directing HSC behavior, HSC migration and cobblestone-area forming cell assays were performed[110]. $CD49b^-$ and $CD49b^+$ HSCs were bulk sorted and 100 cells were each seeded into 96-well plates pre-seeded with mesenchymal stem cells (MSCs). Non-migrated HSCs were counted 24- and 48-hours post seeding and subtracted to evaluate HSCs that have adhered and migrated through the MSC layer. The cobblestone forming ability of migrated HSCs was assessed 28 days post-seeding by counting the number of colonies under an inverted microscope. A cluster of more than 3 cells was considered a cobblestone colony.

Since HSC potential cannot be assessed based on immunophenotype alone. The CD49b HSC subsets were sorted to assess their functions *in vitro*. To test for their myeloid and B cell differentiation potential, OP9 assays were carried out by single cell sorting into 96 wells pre-coated with OP9 stromal cells, followed by myeloid and/or B cell phenotypic analysis by FACS 3 weeks post-sort.

3.6 Cell proliferation and cell cycle assays

To assess the megakaryocyte differentiation potential of different HSC subsets, HSCs were single cell sorted into individual wells of 72-well Terasaki plates and evaluated for the presence of megakaryocytes by morphology by visualization using an inverted microscope 11 days post-sort. To evaluate cell division, HSCs were single cell sorted into individual wells and the number of cells was tracked every day for 3 days post-sort. All cells were incubated under standard incubator conditions in a humidity box.

To assess cell proliferation differences, the 5-Bromo-2'deoxyuridine (BrdU) incorporation assay was used, where BrdU was incorporated into newly synthesized DNA strands during the S phase of the cell cycle, thereby allowing cell proliferation assessment. To maximize the number of labeled HSCs, mice were

intraperitoneally injected with 1 dose of BrdU by weight, followed by 3 days of BrdU given in drinking water. To evaluate HSC cell cycle states, Ki-67 staining was used since the Ki-67 protein is expressed throughout active cell cycle stages. Both experiments were performed following the manufacturer's recommended protocol.

3.7 HSC transplantation

To assess HSC self-renewal and multilineage potential, transplantation experiments were carried out as only functional HSCs could repopulate and continuously replenish the hematopoietic system.

Prior to transplantation experiments, mice were myeloblated with 12Gy irradiation split over 2 days in study I. In studies II and III, mice were myeloablated with 10Gy split over 2 days, except for adult mice in study III, where they were subjected to a total of 12Gy over 2 days.

For single cell transplantations in study I, one HSC was sorted from each of the CD49b subsets. For bulk cell transplantations, 5 cells were sorted, except for the old mice in study III, where 100 cells were sorted instead. All donor HSCs were injected together with 200 000 whole BM support cells. In secondary transplantations, 10 million whole BM cells were transplanted into 1-5 recipient mice for all studies.

3.8 RNA- and ATAC-sequencing

SmartSeq2 scRNA-seq was performed by sorting single cells individually into 384 well plates. The libraries were single-end sequenced using dual indexing. The RNA samples were then demultiplexed and aligned to the mm10 reference genome. In bulk RNA-seq, 250 or 500 cells were sorted and paired-end sequenced according to a previously published protocol[111]. The sequenced reads were mapped to the mm10 reference genome and reads were quantified using HOMER.

For bulk ATAC-seq, 500 cells were sorted and paired-end sequenced using a modified omni ATAC protocol[44]. Samples were then aligned to the mm10 reference genome using the nf-core pipeline. Peaks were called using either MACS2 or HOMER. Quality checks (QCs) were done on the samples and poor-quality samples were excluded. Following QC, the median normalized read coverage for each population was calculated and visualized. HOMER was used to annotate and quantify reads in consensus peaks before additional analyses in R.

3.9 Bioinformatics analysis

R was used to carry out all further bioinformatics analyses in both scRNA-seq and ATAC-seq. In scRNA-seq, the Seurat package was used for the analysis. QC was done on the cells before normalization and dimensionality reduction. Following that, differential expression analysis was done and gene set enrichment analysis was run. The hscScore tool was additionally used to calculate HSC-scores in study III.

In ATAC-seq analysis, read counts were $\log_{10}$ transformed and quantile normalized for visualization, followed by differential analysis. In study III, adult samples from study I were included for aging comparisons. GREAT was used for gene ontology analysis, and motif enrichment was done using HOMER.

3.10 Statistical analysis

GraphPad Prism or R was used for carrying out the statistical analyses. In studies I and II, non-parametric tests were performed and either the Kruskal-Wallis or the Mann-Whitney tests were done on all data. In study III, a preceding Shapiro-Wilk normality test was carried out to determine if the data followed a normal bell curve distribution and passed the normality assumption. Parametric tests were performed with ANOVA, t-tests, or the Wilcoxon signed-rank test if the data passed the normality assumption. Non-parametric tests were carried out using the Mann-Whitney or the Kruskal-Wallis with Dunn's post-hoc test if the data failed the normality assumption.

4 Study I

The original concept of a hematopoietic hierarchy assumed that HSCs are multipotent and generate all blood cells equally, which was later disproved by the identification of myeloid-biased (M-bi), lymphoid-biased (L-bi), platelet-biased, and lineage-balanced HSCs with varying self-renewal abilities; showing that functionally distinct HSCs are present in the HSC pool.

The HSC pool is found within the LSK ($Lin^{-}Sca\text{-}1^{+}c\text{-}kit^{+}$) fraction in the BM, though the frequency is low. Adding cell surface markers to LSK improved HSC isolation, but functionally distinct HSCs still cannot be prospectively isolated and determined until functional tests are carried out due to the lack of definitive cell surface markers for the exact immunophenotype of lineage-biased HSCs. $CD150^{hi}$ HSCs were previously shown to be M-bi with the most LT self-renewal ability, while the $CD150^{-}$ population exhibited the most short-term and transient self-renewal ability with a lymphoid bias[38]. However, variations within each population were still observed.

Therefore, one of the challenges that the field currently faces is the prospective isolation of functionally different HSCs. As a result, it impeded ensuing studies attempting to understand the underlying molecular mechanisms that functionally distinguish them.

Hence, the overall goal of this study is to resolve HSC heterogeneity by understanding the underlying molecular mechanisms of functionally different HSCs by identifying reliable markers for their prospective isolation.

4.1 Results

Shortly after the Morrison group showed that functional LT-HSCs reside in the $CD150^{+}$ population, the Rossi and Eaves groups also demonstrated strong evidence for the reliability of CD150 in enriching for the rare LT-HSC population, even with age[29,66,112]. Soon following that, the Nakauchi group managed to further subfractionate the $CD150^{+}$ population and showed that LT-HSCs reside in the $CD150^{hi}$ fraction, however, heterogeneity remained in the $CD150^{hi}$ population[38]. During our screening for potential candidates that could further subfractionate for LT functional lineage-biased HSCs in the $CD150^{hi}$ population, we observed that CD49b expression showed heterogeneity, suggesting its potential as an additional marker to resolve the heterogeneous $CD150^{hi}$ population. However, we did not know whether the observed phenotypic separation also indicated functional

differences. Therefore, we decided to further characterize the two CD49b populations to understand the functional and underlying molecular differences between them.

Both CD49b subsets were multipotent and we observed a faster cell division and cell cycle kinetics in the CD49b$^+$ cells than the CD49b$^-$ cells, though both were quiescent. This suggested that there was a distinguishable difference in cell kinetics between the two CD49b subsets despite both being highly quiescent.

However, because stem cells are assessed based on their LT self-renewal and multilineage differentiation ability, we performed transplantation experiments to assess functional HSCs. Given that it was reported that the CD150^{hi} population exhibited a M-bi pattern, the CD150^{int} population a balanced lineage output, and the CD150$^-$ cells a L-bi pattern, we carried out bulk cell transplantations from the CD49b$^-$, CD49b$^+$, CD150^{int}, and CD150$^-$ populations. We saw that the CD49b subsets were the most durable and generated all blood cells long term while CD150$^-$ cells exhibited transient myeloid repopulation and a high lymphoid repopulation, consistent with previous report[38]. Interestingly, the CD49b$^-$ subsets exhibited a myeloid skew while the CD49b$^+$ showed a lymphoid skew, even though both subsets resided within the CD150^{hi} M-bi population. Hence, the LT lymphoid bias observed in the CD49b$^+$ subset differs from the ST/transient CD150$^-$ lymphoid bias seen in previous publications[38].

Since bulk cell transplantation gave an averaged output of the transplanted cells, single cell transplantation was required to quantify and compare lineage bias between the CD49b subsets on a clonal level. Consistent with bulk cell transplantation results, both CD49b HSCs showed comparable reconstitution and repopulation levels, but CD49b$^-$ HSCs showed enrichment for myeloid bias while CD49b$^+$ HSCs showed enrichment for lymphoid bias (B, T, and NK). To evaluate their extensive LT self-renewal ability, we performed secondary transplantation of phenotypic CD49b HSC subsets generated from primary recipients. We noticed that first, lineage bias patterns persisted through secondary transplantation and second, CD49b$^+$ HSCs were less efficient than CD49b$^-$ HSCs in generating phenotypic HSCs, especially in secondary transplantation. This meant that CD49b$^-$ was the more LT and durable HSC, showing myeloid bias. The CD49b$^+$ subset, on the other hand, showed lymphoid bias but not as extensive HSC activity as the CD49b$^-$ HSCs. Overall, though both CD49b subsets sustained LT

multilineage potential and exhibited persistent lineage bias, their efficiencies differed.

Having observed functional differences between the CD49b subsets, we wanted to elucidate their underlying molecular differences using RNA-seq analysis. However, both bulk and single cell RNA-seq showed few transcriptomic differences between the CD49b populations, suggesting that CD49b functional differences could not be resolved on the gene expression level.

With little transcriptomic differences observed in RNA-seq analyses, we speculated that the differences could lie in the epigenetics of the subsets, since we observed a persistence of lineage bias from primary to secondary transplantation, consistent with published studies proposing epigenetic mechanisms at play[14,88]. Our ATAC-seq showed distinct separation within CD150^{+} populations in the principal component analysis (PCA). CD49b^{-} and CD49b^{+} also showed differentially accessible regions (DARs), where the CD49b^{+} subset showed more overlapping open chromatin regions with CD150^{int} and CD150^{-} populations, which were more downstream of the CD150^{hi} population. This suggested that CD49b^{-} was hierarchically above the CD49b^{+} subset. The distinct DARs between the CD150^{+} cluster and progenitor cells suggested that significant chromatin remodeling occurred when HSCs generate downstream progenitors.

In addition, cell quiescence and genetic imprinting pathways were more associated with the CD49b^{-} subset, while processes regulating cell number, proliferation, and differentiation were more associated with the CD49b^{+} subset. This corresponded to our cell cycle and proliferation results, where CD49b^{-} was more quiescent. Motif enrichment analysis suggested that the same group of transcription factors (TF) might regulate both CD49b subsets despite functional differences.

Taken together, our functional and molecular data suggested that the observed M-bi in CD49b^{-} HSCs and L-bi in CD49b^{+} HSCs are due to differences in their chromatin accessibility.

4.2 Conclusions and discussions

Given the functionally heterogeneous HSC pool, we have identified CD49b as a useful and reliable marker to further subfractionate the phenotypic CD150^{hi} HSC pool into a CD49b^{-} subset enriching for M-bi HSCs, and a CD49b^{+} subset enriching for L-bi HSCs with discernable differences in cell proliferation and kinetics between the CD49b subsets, despite both residing in the CD150^{hi} pool.

As opposed to the previously described CD150^{-} population, which showed a lymphoid skew with ST or transient ability[38], the CD49b^{+} L-bi HSCs in the CD150^{hi} pool in our study showed LT multilineage potential. This highlights the necessity to include CD150 in differentiating functionally distinct cells. We also showed that though the CD49b^{+} cells could sustain LT-HSC activity, they are less good at sustaining LT self-renewal ability than CD49b^{-} HSCs. Further, due to the lower efficiency in generating downstream progenitors from the CD49b^{+} HSCs, in addition to the high overlap in chromatin accessibility between CD49b^{+} and downstream CD150^{int} and CD150^{-} populations, the CD49b^{-} subset can be inferred to be hierarchically above the CD49b^{+} subset.

Even though absolute isolation of homogeneous lineage-biased HSCs was still not possible, the addition of CD49b in our study showed that it was possible to further enrich for LT M-bi and L-bi HSCs in the CD150^{hi} LT-HSC pool. Our attempt to add a combination of both CD229 (lymphoid marker) and CD41 (myeloid marker) to additionally purify functionally distinct HSCs did not show significant differences in cell proliferation and cell cycle assays. Therefore, more comprehensive functional tests are needed to resolve the remaining heterogeneity within each subset.

Since we could use CD49b to further enrich for LT M-bi and L-bi HSCs, we performed molecular analyses on them as previous genome wide analyses included the whole HSC pool and insufficient molecular analyses were done on functionally distinct HSC populations. However, our transcriptomic analyses showed high similarity between the CD49b subsets. This suggested that either very few genes are responsible for HSC lineage bias, or that RNA-seq is not sensitive enough to detect small gene expression changes. It also suggested that the differences observed in our cell kinetics analyses were not due to differences in gene expression levels, but by other factors. Hence, the differences detected in chromatin accessibility between the CD49b HSC subsets in our epigenetic analysis suggested that lineage-biased HSCs were primed and ready to differentiate in specific lineages while retaining their stem cell properties.

In addition, because CD49b is an integrin, we do not know how it influences HSC homing during transplantation, which could lead to the CD49b^{-} HSCs localizing to different locations from the CD49b^{+} HSCs. Some studies have suggested that HSCs localize to different regions in the bone, where they might receive different signaling cues from the niche microenvironment[42]. However, because we did not

investigate the role of the niche in lineage-biased HSCs, this question remains to be elucidated.

All in all, study I has shown that by incorporating CD49b as an additional marker to the conventional LT-HSC immunophenotypic definition, we were able to further subfractionate the HSC population into a subset of CD49b$^-$ HSCs that enriches for M-bi HSCs and a subset of CD49b$^+$ HSCs that enriches for L-bi HSCs which showed chromatin accessibility differences. This showed the reliability of CD49b as a marker for prospective isolation of lineage-biased HSCs.

5 Study II

True functional LT-HSCs remain not well defined, and their underlying molecular mechanisms remain obscure due to the presence of multipotent progenitors (MPPs) in the HSC pool. MPPs are cells with ST self-renewal ability that can only be distinguished from LT-HSCs through transplantation experiments assessing for self-renewal and multilineage abilities in retrospect.

It is therefore crucial to purify the HSC pool as much as possible to minimize cell contamination when carrying out HSC molecular studies. Both CD49b and CD229 have been reported to be good candidates to separate functionally different HSCs individually, though they still could not fully exclude non-HSC cells[30,44].

Therefore, this study sets out to characterize whether the CD49b and CD229 marker combination could further purify and refine the LT-HSC pool.

5.1 Results

Using previously reported CD49b and CD229 markers, we could subfractionate the $CD150^{hi}$ HSC pool into $CD49b^{-}CD229^{-}$, $CD49b^{-}CD229^{+}$, $CD49b^{+}CD229^{-}$, and $CD49b^{+}CD229^{+}$ subpopulations. Our cell cycle and proliferation studies showed that the $CD49b^{-}$ subsets were more quiescent than both $CD49b^{+}$ subsets and *in vitro* analysis showed that all 4 populations were multilineage. These observations were consistent with the results in study I and suggested that CD49b was more effective in differentiating cell kinetics than CD229.

To assess HSC function and LT self-renewal ability, we carried out limited dose bulk cell transplantation from each subset. All 4 populations showed similar reconstitution levels, implying comparable homing abilities, however, $CD49b^{+}CD229^{+}$ showed declining platelet, erythrocyte, and myeloid lineages over time. The diminishing myeloid repopulating activity indicated that $CD49b^{+}CD229^{+}$ cells had the least LT regenerative activities compared to the other 3 populations because myeloid cells are short lived and need to be constantly replenished. It was also the least efficient in regenerating HSCs and progenitor cells, again indicating ST activity. Secondary transplantation showed further reduction of reconstitution, as well as reduced myeloid and HSC repopulation of the $CD49b^{+}CD229^{+}$ population. All transplantation data indicated that $CD49b^{+}CD229^{+}$ did not exhibit LT regenerative ability, but because it exhibited multilineage potential, it was most likely a population of MPPs contaminating the LT-HSC pool.

This observation was further supported by limiting dilution calculations, which estimated the lowest functional HSC frequency in the CD49b^{+}CD229^{+} population in both primary and secondary transplantations.

Aside from the least HSC activity observed in the CD49b^{+}CD229^{+} population, the other 3 subsets showed either stable or increasing LT repopulation of all lineages. The highest myeloid repopulation was observed in the CD49b^{-}CD229^{-} population and one of the highest lymphoid repopulation in CD49b^{+}CD229^{-} in both PB and BM analyses, suggesting that M-bi HSCs might reside in the CD49b^{-}CD229^{-} population, while L-bi HSCs could reside in the CD49b^{+}CD229^{-} population.

Together, these data highlighted the effectiveness of using both CD49b and CD229 as cell surface marker combinations in further purifying the functional LT-HSC pool, but also a sustained heterogeneity in the leftover HSC populations.

5.2 Conclusions and discussions

Currently, functional LT-HSCs cannot be prospectively isolated in the field, which limits our understanding of them because of the presence of other non-HSC cell types in the HSC pool, especially their underlying molecular mechanisms. It is therefore crucial to overcome this limitation to elucidate the precise mechanisms that direct HSC behavior.

In study II, we have shown that by combining CD49b and CD229 as cell surface markers, the CD150^{hi} population can be further subfractionated into 4 subpopulations. *In vitro* and functional tests showed that all populations were multilineage with self-renewal abilities, except CD49b^{+}CD229^{+}, where it did not exhibit extensive LT reconstitution and showed the most MPP-like properties. This has thereby allowed us to further purify the LT-HSC pool by eliminating the CD49b^{+}CD229^{+} MPP cells. However, heterogeneity within each of the remaining subpopulations was observed in our functional tests, indicating that the subpopulations still contain cells that might not be true functional HSCs. Thus, more work is needed to improve the purities of the remaining subpopulations before achieving the goal of prospectively isolating true functional LT-HSCs.

Taken together, study II showed that by using CD49b and CD229 as cell surface marker combinations, we can further isolate LT-HSCs by eliminating MPPs from the HSC pool. This showed that prospective isolation of functional HSCs is possible either by identifying functional HSCs or by eliminating non-HSCs from the population.

6 Study III

With age, the HSC pool expands, along with an increase in myelopoiesis. However, it is still unclear whether the aging related myeloid expansion is due to an increase in the number of M-bi HSCs and a decrease in the number of L-bi HSCs within the HSC pool, or intrinsic alterations within individual HSCs.

Despite continuously being a topic of interest in the field, aging studies in previously published reports failed to cover a comprehensive age span in their work. Mice 3-4 weeks old (juvenile) - where a period of dynamic tissue remodeling and cell activities occur - were not included. Consequently, there is a black box of HSC activity and the underlying molecular changes that occur between juvenile and old mice during HSC aging. Furthermore, previous published studies considered the HSC population as homogeneous and the presence of lineage-biased HSCs was not taken into account. As a result, our current understanding of how functionally different HSCs behave with age is a sum of all subsets, and mechanisms underlying distinct HSC subsets have not been elucidated.

Therefore, the goal of this study is to understand the functional and molecular changes that occur from juvenile to old age in lineage-biased HSCs.

6.1 Results

Since we have shown CD49b as a reliable marker to segregate lineage-biased HSCs in the CD150^{hi} HSC population in study I, we wanted to understand it in the context of aging. The persistent subfractionation of the CD150^{hi} HSC population into CD49b^{-} and CD49b^{+} subsets indicated the stability of CD49b as a marker to phenotypically separate the heterogeneous LT-HSC pool throughout aging. We also noticed that the juvenile age group was most proliferative and most cycling, while the adult and old age groups were quiescent. Additionally, both CD49b subsets in all ages showed comparable multilineage differential potential.

Even so, transplantation assays are needed to definitively assess HSC LT self-renewal ability. Both subsets showed similar reconstitution and multilineage potential in all age groups in primary transplantation. However, significant differences in reconstitution ability and blood lineage distribution were displayed between CD49b subsets in juvenile and old ages in secondary transplantation. We also observed an increase in myeloid cells with age in both CD49b subsets. To quantify the lineage distribution, we calculated the lymphoid to myeloid cell ratio (L/M ratio) of transplanted mice using the PB of unmanipulated adult mice. We

observed that in juvenile and adult age groups, the CD49b$^-$ subset enriched for M-bi cells, while the CD49b$^+$ subset enriched for L-bi cells. However, mice transplanted with old HSCs showed an extreme myeloid skew, which made it challenging to discern any differences in L/M ratio between the CD49b subsets. Hence, we calculated the L/M ratio of old transplanted mice using the L/M ratio from old unmanipulated mice to normalize the myeloid skew and discovered that, unlike in juvenile and adult age groups, CD49b$^+$ in old age marked lineage-balanced HSCs. In brief, we observed that similar to study I, CD49b$^-$ subset enriched for a population of M-bi HSCs while CD49b$^+$ enriched for a population of L-bi HSCs in juveniles and adults. However, in old age, CD49b$^+$ marked a population of HSCs with balanced output. This suggested that with age, there was an overall shift away from lymphoid bias in the CD49b$^+$ HSC subset. We also observed *in vitro* cell adhesion and migration differences between the subsets although it was unlikely due to the BM HSC niche, since we did not see any differences in BM localization between the subsets.

Despite observed functional differences, our scRNA-seq analysis showed no subset differences, instead, a predominant age difference was observed. This suggested that on the transcriptomic level, aging played a bigger effect on HSC behavior. However, even though we could observe cluster separation after UMAP dimensionality reduction, subsequent differential analysis did not yield genes that were significantly differential between subsets.

ATAC-seq analysis also showed a major aging difference in our PCA where differential analysis showed that the most differential regions corresponded to previously published datasets. When comparing chromatin accessibilities, we observed that first, there were more open chromatin regions in the old age group than in the juvenile age group. Second, there was a progressive increase in chromatin accessibility regions with age, through the adult period. Third, the majority of the gain in chromatin accessibility in old HSCs was acquired *de novo*. Interestingly, the accessibility differences with age were only observed in HSCs and not propagated to downstream progenitors. This suggested that significant chromatin remodeling in the HSCs occurred during aging, already beginning at the juvenile to adult stage, and motif enrichment analysis showed that SpiB TF could be responsible for the chromatin accessibility differences observed in the juvenile age group.

In addition to the observed aging differences, we wanted to elucidate the molecular differences between CD49b subsets, and therefore carried out differential analysis between the subsets within each age group. However, no significant differences were detected on the epigenetic level given subset similarity, especially in the juvenile age group, thereby hindering further analyses.

Considering that we have a very M-bi subset (old CD49b$^-$) and a very L-bi subset (juvenile CD49b$^+$), we instead exploited the extreme lineage bias between them by comparing these populations in our differential analysis. However, age posed a significant variable during the analysis. Therefore, we subtracted all aging related differentially accessible regions (DARs) from lineage DARs to exclude any aging effects and observed 2 distinct clusters. Motif enrichment analysis again showed that SpiB TF was enriched in the L-bi cluster. However, further functional studies are required to validate their roles in lineage-biased HSCs during aging.

6.2 Conclusions and discussions

By studying M-bi and L-bi enriched HSC subsets using CD49b as a cell surface marker, we observed distinct functional changes in aged HSCs, as opposed to their younger counterparts. The differences observed in *in vitro* cell migration and adhesion differences between CD49b subsets were unlikely due to their localization, since there were no significant differences in their distribution patterns in the BM. Furthermore, instead of marking L-bi enriched HSCs, CD49b$^+$ marked HSCs with a balanced output in old age.

Even though transcriptomic analysis revealed more significant differences with aging than between subsets, it did not provide sufficient insight into potential gene regulation. Differential expressed gene analysis on aging differences validated previous published aging candidates while differential analysis on subset differences yielded little result due to the similarity between both subsets. Our epigenetic analysis, on the other hand, showed that chromatin remodeling started already in the juvenile age, and that the same TF, SpiB, might be regulating both HSC aging and lineage bias.

There are currently two models to explain the increased myelopoiesis in the HSC pool with age. On the one hand, it was thought that the change in the composition of the HSC pool was due to an increase in the number of M-bi HSCs and a decrease in the number of L-bi HSCs, with lineage-biased HSCs retaining their differentiation preference. On the other hand, it was theorized the increased

myelopoiesis is due to a decrease in the ability of HSCs to produce lymphoid cells. Our study looking at lineage-biased HSCs, which showed an observed shift away from lymphoid bias in both old CD49b HSCs, supported the model where HSCs undergo intrinsic alterations in lineage differentiation properties.

Taken together, study III suggests that HSCs of the same immunophenotype undergo functional and epigenetic changes with age, already beginning in the juvenile age group. These results thus support the cell intrinsic concept, where individual HSCs undergo intrinsic changes with age, and demonstrate that CD49b as a cell surface marker can not only subfractionate phenotypic HSCs, but also distinguish functional and molecular differences throughout aging. Furthermore, this is the first study to propose that the same TF might be regulating both HSC lineage bias and aging, suggesting that aging and lineage bias are perhaps not mutually exclusive events.

7 Points of perspective

The three studies included in the thesis focused on elucidating the underlying molecular mechanisms of lineage-biased HSCs in the LT-HSC pool by using cell surface marker combinations and various sequencing techniques to understand the fundamental differences that resulted in functionally different HSCs.

7.1 Current technical limitations and challenges

The idea that cellular functions should reflect transcriptomic profiles was not observed in our molecular analyses. Instead, we saw more associations between chromatin accessibility and functional differences when comparing lineage-biased HSCs. One potential reason could be the different states that the HSCs were in at the time of assay. Our functional data looked at HSCs that have undergone transplantation and therefore were under pressure to repopulate the entire hematopoietic system in myeloablated recipients, while HSCs used in sequencing analyses were from steady state mice, and hence not under stress. This highlighted one of the biggest challenges in the field, which is the inability to use the same cell in both functional and molecular tests. The HSCs might be immunophenotypically similar, but it is not possible to know if they are functionally the same. This presented a question on whether transplantation could reliably and accurately reflect what happens in steady state HSCs, since transplanted HSCs are under stress, not to mention the loss of the niche microenvironment in myeloablated recipients that will indubitably affect HSC homing and cell-cell communication. Nonetheless, though these conditions do not realistically reflect the natural environment that HSCs reside in, transplantation remains the gold standard in assessing HSC self-renew ability (the ability to undergo several cell divisions while remaining undifferentiated) and multilineage differentiation potential (the ability to give rise to all downstream progenitor and mature cells).

Using additional markers to further purify the HSC population is possible, as shown by study II, which demonstrated that with the inclusion of CD229, further phenotypic and functional subfractionation of the CD49b subsets was observed. Alternatively, using fewer cell surface markers to screen for functional HSC cell surface markers, similar to what the Göttgens group has done, will provide a more unbiased approach[33]. Nonetheless, as more cell surface markers are being identified, the HSC population becomes further enriched. However, negative selections are always more challenging than positive selections because while we know what cells express surface markers in positive selection, it is difficult to know

what cells are present in the negative population since it will contain a mixture of other cell types. Therefore, the addition of different positive selection markers in the future would be useful to help navigate the prospective isolation of functional HSCs and lineage-biased HSCs when used in combination with negative selection markers. We observed that the $CD49b^{-}CD229^{+}$ subset showed the highest reconstitution and multilineage ability among the 4 populations in study II. In addition, it showed one of the highest estimated HSC frequencies, indicating the validity of CD229 as a positive selection marker when used in combination with CD49b negative selection. However, since only bulk cell transplantation was done, follow-up single cell transplantation experiments will be required to show which lineage-biased HSCs reside in individual subsets.

The question that follows would be the purpose of prospectively isolating HSCs if we could retrospectively identify them through functional studies. Not only would the prospective isolation of lineage-biased HSCs help the field pinpoint functional HSCs prior to performing the experiments, but it would also give insight into the molecular differences between them based on immunophenotype. In order to label HSCs in their native state while retrieving mechanistic regulation information, the Weissman group combined viral transduced cell labeling, sequencing, and DNA barcoding to label HSCs and follow HSC differentiation on the single cell level[50]. The Camargo group has been working on the *in vivo* labeling of functional HSCs to trace them without perturbing HSC steady state, while the Scadden group has barcoded transplanted HSCs to understand the epigenetic regulation that underlies HSCs when under stress[88,113,114]. In a broader sense, prospective isolation of lineage-biased HSCs, when translated to the clinical setting, will help in designing personalized medicine based on patient blood malignancies. Targeted therapy of lineage-biased HSCs such as gene or base editing could help retain patient HSCs without having to irradiate the whole HSC pool by targeting only the mutated HSC subsets; thereby preserving the integrity of the remaining HSC pool and minimizing invasive treatment[115–117].

One also could not overlook the fact that the majority of the current strategies used to identify phenotypic HSCs are based on cell surface marker selections, where the cell surface proteins are involved in other cellular functions such as cell-cell communication. Additionally, one drawback of using flow cytometry is the loss of spatial information of the interaction of HSCs to their niche microenvironment[42]. Spatial transcriptomics could potentially overcome the challenge by maintaining the microenvironment that the HSCs reside in and has been done in studies

looking at the relationship between fetal liver HSCs and the niche[118,119]. When mice are lethally irradiated during transplantation, niche cells, which are essential for HSC homing, maintenance, and function, are lost[97,120]. To overcome irradiation related damages, the Thomson group generated a *kit*$^{w41/w41}$ mouse model that allowed HSC engraftment without the need for myeloablation[121]. The Forsberg group generated a model which ablates specifically the HSC population, thereby allowing the retention of other hematopoietic cells[122] while the Yamazaki group cultured HSCs in serum-free media prior to transplantation[123]. These new models have avoided the loss of important niche cells from irradiation, though further optimizations are needed to preform single cell transplantations.

Additionally, to avoid potential bias in HSC cell surface marker identification, the Göttgens group has shown that HSC and progenitor populations can be segregated without cell surface marker information as they show distinct TF expression levels[124]. However, the problem persists in that cells used to validate HSC self-renewal and multilineage differentiation potential are not the same cells being analyzed in transcriptomic analyses. Even though these methods have been proposed to be good options, they could not circumvent the challenge of the collected data being correlative.

Although the field has become more single cell driven, single cell analysis could not provide the same depth as bulk cell analysis. Since both CD49b subsets were so similar, we wanted to sequence as deep as possible to capture maximum information and differences from these cells. Therefore, although bulk cell ATAC-seq reduced the resolution that we could get from single cell analysis, it was the most suitable in our studies as it could give us an overview of the CD49b subsets that are enriched for lineage-biased HSCs as individual populations. Other sequencing techniques such as ChIP-seq, Cut&Tag, or Cut&Run could complement our ATAC-seq data and give additional information on the single cell level[125–130]. Multi-omics techniques are also now possible in the field to perform different analyses and retrieve data from the same cell instead of relying on their immunophenotypes[131,132]. From the sequencing data in studies I and III, observable differences on the bulk cell level gave us confidence for future subsequent single cell and multi-omics analyses.

Study III supported the idea of using chromatin accessibility as a method for assessing HSC age in future studies since juvenile HSCs show distinct chromatin accessibilities from old HSCs in our analysis. Horvath's epigenetic clock, which

determines the age of a cell based on DNA methylation states, has been used in various tissues and mature cell types[133] and could be a good complementation to chromatin accessibility analyses. In mice, methylation at the CpG sites in HSCs has been suggested to be a good indicator of aging, as reported by the Goodell lab[74,134]. In humans, there have been reports which looked at DNA methylation of 3 CpG sites (Itga2b, Aspa, Pde4c) in the blood and in the induced pluripotent stem cell model[135–137]. Currently, *in vivo* labeling has been used to study the hierarchical and lineage specification within the hematopoietic system[113,114]. Applying a similar concept and technique to lineage-biased HSCs could give us an alternative to resolve the underlying mechanisms of lineage-biased HSCs.

Aside from epigenetic and transcriptomic studies, proteomics analyses on HSCs have also been able to provide insight into changes in HSC metabolic pathways during HSC aging[138,139]. It was suggested that HSCs show post-transcriptional changes during aging, where the level of mRNA is unproportional to the level of proteins[140]. It was also shown that protein synthesis is lower in HSCs to minimize misfolded proteins in the cell, thereby maintaining HSC function[141]. However, limited studies on lineage-biased HSCs were performed, calling for more attention on proteomics analyses on functionally distinct HSCs.

7.2 Additional discussions on the HSC concept

Apart from the current technical limitations in the HSC field, there are also varying ideas of the HSC concept that require reconciliation. First, one would argue for the need of lineage-biased HSCs because it allows the HSCs to respond quickly to injuries and stress since they are already primed for a specific lineage. Some proposed that emergency hematopoiesis recruits MPPs instead of HSCs[142]. Others suggested that lineage-biased HSCs epigenetically alter their lineage preference during emergency granulopoiesis, though given the resources and energy that the HSCs need to expend, the rationale behind this idea remains debatable[143]. Regardless, USP22 was suggested to be involved in emergency response in HSCs[144]. The Jacobsen group has identified a population of platelet-restricted HSCs that bypasses the conventional hematopoietic hierarchy differential pathway[41,54]. These studies implied that the presence of lineage-biased HSCs allows for quick responses if needed.

With the two current concepts to explain the increased myelopoiesis with age in the HSC pool, there are several possible explanations. In the intrinsic model, compositional changes in the HSC pool could be due to HSCs intrinsically altering

their differentiation preference with age, with limited influence from the niche microenvironment[75,76,145]. On the other hand, the extrinsic model suggested that there is an increase in the number of M-bi HSCs and a decrease in the number of L-bi HSCs with age[66,68]. The possibility that M-bi HSCs outcompete lineage-balanced and L-bi HSCs with age could not be excluded, when it was proposed that MPPs compete with HSCs for the niche[146]. Additionally, some reports have suggested that different lineage-biased HSCs reside in distinct niche regions and different cell-cell communication occurs[147,148]. Several studies looking at the role of the niche in HSC aging showed that aged HSCs improved in fitness when the niche was rejuvenated into a more youthful phenotype, and showed a decrease in myeloid output when old HSCs were transplanted into young niches[149,150]. In relation, aging associated inflammation (inflammaging) has also been proposed to be another factor in decreased HSC function and increased myeloipoiesis[43,151].

7.3 Future perspectives

Though mice remain an extremely well-characterized model for HSC studies, it is not without limitations[152]. The cost, the time taken to perform the analyses, and the ST *in vivo* and *in vitro* assays, are all areas with room for optimization. The generation of humanized mouse models allowed the field to study human HSCs in the mouse, although it might not faithfully translate mouse studies into the human context[153]. Perhaps at best, we could optimize the mouse model while carrying out HSC studies to eventually reach a point where more focus is put on conceptual discussions without being limited by techniques or model systems.

Nonetheless, shared genetic similarities between mice and human permitted information from mice studies to be extrapolated into human hematopoiesis. Previous studies on human cord blood have shown that the $CD34^{+}CD38^{-}CD90^{+}CD49b^{+}$ population enriches for LT reconstituting HSCs[154]. Attempts to use whole genome sequencing to track human HSCs and their progenies have also allowed the field to understand the behaviors of different HSC clones[155]. The different LT-HSC immunophenotypes between humans and mice calls for more human HSC studies to faithfully translate what is known about mouse HSCs to human HSCs.

With our molecular characterization of the LT-HSCs in mice, it would also be very interesting to see how much of the mouse hematopoietic stem and progenitor cell (HSPC) profiles overlap with the human HSPC transcriptomic and epigenetic profiles, which might be a gateway for us to use steady state data from mice and

compare it to human HSPCs to get more comparable datasets, and therefore allow us to infer human cellular processes drawn from mouse studies, and associate mouse to human data[156].

Given that all our studies looked at normal hematopoiesis, it is also important to look at the function and role of lineage-biased HSCs in the disease context[157]. The Philadelphia chromosome positive (Ph^+) chronic myelogenous leukemia, with a hallmark signature of a BCR-ABL1 fusion oncogene, is a leukemia with stem cell-of-origin, though the underlying mechanisms remain obscure. In our preliminary data, we observed that there is a difference in leukemia development in different lineage-biased HSC subsets, as well as the type of leukemia developed. Mice with fusion oncogene induced in the $CD49b^-$ subset died rapidly while mice with fusion oncogene induced in the $CD49b^+$ subset survived up to a year. This suggested that the intrinsic lineage bias of the HSCs might be playing a role in disease development ($CD49b^-$) and disease protection ($CD49b^+$)[158]. More in-depth studies into leukemia development in different lineage-biased HSCs and understanding differences in their underlying molecular mechanisms would give us more insight into the regulation of lineage-biased HSCs and complement our studies on steady state hematopoiesis.

In addition, when put into the clinical context, clonal hematopoiesis with indeterminate potential (CHIP) is another hematological malignancy that is prevalent in aged individuals. It is the expansion of a population of HSCs with an acquired mutation that have gained selection advantage, which allowed them to expand better than wildtype HSCs in the HSC pool[159]. Several key epigenetic regulators involved in CHIP mutations have been shown to be TET2, DNMT3A, and ASXL1[160]. Individuals with CHIP are more prone to develop blood malignancies such as myelodysplastic syndromes and acute myelogenous leukemia, and often with increased risk of cardiovascular diseases[161]. It was proposed that CHIP mutant clones are more resistant to inflammation and aging when compared to wildtype clones, thereby giving them a selection advantage[162]. However, CHIP studies in the context of lineage-biased HSCs remain unclear and more need to be done.

Taken together, despite existing challenges and limitations that have to be overcome both technically and conceptually, the progress that the field has made since HSCs were first discovered cannot be discredited.

8 Acknowledgments

I am not very good at expressing feelings but will try. As with what people say that it takes a village to raise a child, it takes a department to raise a PhD student. There are so many people to thank for my PhD journey and impossible to list out everything but here goes. First, thank you, **Sidinh Luc**, for your mentorship through the years, and for giving me the opportunity to do a PhD with you, this would not have happened without you taking a chance on me. You have gone above and beyond as a mentor and I am grateful for everything that you have done to help me grow not just professionally, but also personally. **Anne-Sofie Johansson**, thank you for your sense of humor and patience in tolerating my unnecessarily dramatic reactions to minor inconveniences, and bring life everywhere you go. Thank you, **Ece Somuncular**, for your kindness through my first few years of PhD life in Sweden, and the times that we shared at the bench. **Özge Dumral**, you were the first person who brought me around and I will not forget how welcomed you made me feel on my first day. **Efthymios Tzortzis** and **Aurora Forlani**, thank you for the energy that you brought while you were here. It really helped get me going, especially on difficult days when I felt less energetic.

I will also not be who and where I am today without my co-supervisors. **Sten Eirik Jacobsen**, from whom I learnt something new during every lab meeting, the discussions were so stimulating and I always take something away. **Petter Woll**, you are so kind and I aspire to be as calm and composed as you are under any circumstances in life. **Robert Månsson**, you are a cool periodic optimist and I have learnt a lot about bioinformatics analysis from you through the lab meeting discussions, thank you.

Julia Hauenstein, thank you for being my emotional rock throughout my bioinformatics learning journey, or my life really. The patience that you have for me is unmatched and I am extremely lucky to have you as my friend and bioinformatics guide. **Nicolai Frengen**, for being the voice of reason when I say something absurd or attempt to persuade Julia into all sorts of shenanigans. **Charlotte Gustafsson**, thank you for taking me in as the bonus child of the Månsson lab and introducing me to yoga.

Thank you to the Woll and Jacobsen lab members for all the meeting and retreat discussions. **Madeleine Lehander**, your confidence is an inspiration, not only as a scientist, but also as a person. **Axel Winroth**, I always learn something new about bioinformatics, sequencing techniques, or data interpretation from you every time

we talked. Thank you, **Masafumi Seki** and **Tetsuichi Yoshizato**, you have been so kind, helpful, and patient with me whenever I come to you with questions and when you see me struggling. Thank you also to **Stefania Mazzi**, **Margs Brennan**, **Kari Högstrand**, **Karin Belander-Strålin**, **Ellen Markljung**, **Joana Carrelha**, **Affaf Aliouat, Bishan Wu**, and **Dimitra Vasileiadi** for the discussions.

Hong Qian, thank you for being so nice and encouraging, and always taking the time to answer any questions that I have. **Elory Leonard**, for the emotional support and joy that you bring. It is a feat how you managed to stay positive under stress, and your ability to look on the bright side in any situation is a gift. Thank you, **Alma Månsson**, **Karin Linderfalk** and **Runqing Zhang**, for being so kind whenever we chatted. Thank you, **Huan Cai** for the cultural exchange in language and writing. **Lakshmi Sandhow**, I will miss the conversations we had in the TCR where you shared your experiences. **Juan Pablo Medina Giménez** and **Marcus Chang**, thank you for the company.

Elisa Saccon and **Katharina Kirchhof**, you have been in room 7012 the longest with me and took in my dramatic outbursts through the years without any judgement. It is not easy having to do research and tolerate an office drama queen at the same time. Also, the happiness on your face when we talk about food, **Elisa**, I have only seen that described in books and I am convinced that you have Asian genes in you.

Next, the people that boost my day first thing in the morning. **Indira Barbosa**, for the constant optimism and joy that you bring everyday, and the quotes-of-wisdom. **Kajsa Ax**, you are adorable and always have wonderful stories to share. **Gunilla Walldin**, for being the constant that I see in the morning, it took me a while to get used to not seeing you sitting by the coffee machine with **Agneta Andersson** while I recover from climbing the stairs. **Agneta** and **Laura Sanz-Ortega**, for the morning "high-fives" to prepare us through the day. **Jelve Nejati Zendegani**, you are such a warm, caring, and encouraging person, thank you for the constant support and always being there to cheer me up.

Laura Covill, **Jonas Thier**, **Tessa Campbell**, and **Filip Segerberg**, thank you for all that you have done for the PhD students at HERM. Organizing student get-togethers and managing SAP seminars, in addition to being the go-to people for questions are no easy tasks and you carried it off so well. Thank you also, **Laura**, for being the person I go to for statistics questions, and being a reference through the defense process.

Thank you, **Eva Hellström-Lindberg**, for not only being the chair of my defense, but also a kind and inspiring leader. I am a fan and am in awe by how you lead HERM. **Yenan Bryceson**, who probably does not remember, but being the first person to welcome me at HERM. Thank you also to the other group leaders: **Petter Höglund**, **Vanessa Lundin**, **Julian Walfridsson**, **Evren Alici**, **Mattias Carlsten**, **Arnika Wagner**, **Roland Fiskesund, Nicole Marquardt**, **Jette Lengefeld**, and **William Nyberg**. Thank you to the HERM administrators, **Sri Sahlin**, **Annette Öster Fernström**, **Elin Hellsten**, and **Sara von Bahr Grebäck**. You have been so kind to me and were ever so helpful with any questions that I have. Thank you, **Annette** for also checking in on me from time to time.

Thank you, **Johanna Ungerstedt** and **Carmen Gerlach**, for being in my halftime committee. **Johanna**, I will not forget how your words shifted my perspective on self-esteem. Thank you to **Andreas Lennartsson**, for being in both my halftime and defense exam committee. **Marzia Palma**, thank you for being my mentor.

Thank you to my opponent, **Johan Flygare** and to my examination committee: **Xingqi Chen**, **Jenny Pyl Hansson**, and **Andreas** for the discussions.

Pedro Moura, thank you for being so helpful with bioinformatics discussions. Also, your advice in our chats have really helped me pull myself together. **Heinrich Schlums**, you are the funniest realistic optimist I have met, and always have hilarious comebacks. Thank you for making my interactions with you enjoyable.

Thank you, **Yang Liu**, for taking me under your wings and have been so generous in sharing your experiences. **Gloria Wu**, your Taiwanese snacks kept me going, and **Anna Pumpe**, you have been so accommodating towards me during our time in the office, thank you.

Thank you, **Sadaf Fazeli**, for the company and the emotional support alpaca; **Sigrid Wahlen**, for being so nice and for all those wonderful pastries; **Caroline Eriksson**, you are such a sweet person; **Sophia Hald**, for book recommendations; **Elisabetta Cozzi**, for the camaraderie; **Petar Mitev**, for your "what's ups". Also thank you, **Nicole Wild**, **Teresa Mortera Blanco**, **Lars Gerullat**, **Jonah Anderson**, **Francesca Grassi**, **Sophia Hofmann**, and **Luana Tomaipitinca**.

Belinda Pannagel, thank you for not only for the help with FACS, but also the regular check-ins to make sure that I am doing fine. **Narmadha Subramanian**, thank you for catching up from time to time, also for lending a listening ear and

sharing your experiences. **Francesca Grasso** and **Mahin Nikougoftar Zari**, thank you for all the FACS help, especially in my early days working with the machines.

Caroline Leijonhufvud, thank you for being so warm and caring, and **Filip**, for being such a gentle human with a contagious smile. **Giovanna Perinetti Casoni** for regularly checking in on my back pain and sharing pain-relieving exercises with me. **Lucía Peña Pérez**, you were so nice to me as a first year in one of my earliest meetings and made sure that I settled in well.

Maaike Jongen, thank you for always sharing new things with me that I would not have known otherwise. **Suria Jahan**, I will not forget the advice and experiences that you shared with me during our spontaneous chats. Thank you, **Laurent Schmied** and **Thuy Luu Thanh**, for being so welcoming during my early days.

Sofia Bengtzén, Ann Wallblom, Anna Bohlin, and **Ann-Charlotte Björklund**, and **Muhammad Kashif**, you have been so kind to me, and seeing you in the morning has become a reassuring routine.

Thank you, **Nutsa Burduli** and **Sophia Borate**, you were always so nice when we meet each other in the FACS room. **Xiufeng Xu**, thank you for being a friend after the NEO retreat. It is always nice to have a quick chat by the coffee machine. I would also like to thank all the past and present **PKL staff** for their assistance with experiments and mice housekeeping through the years. **Hooman Aghelpasand** and **Pontus Höjer**, thank you for the lab meeting discussions, though we only overlapped for a short period of time.

Chih-Chiang Chan, thank you for the guidance and the chats whenever I go back home. I will forever be grateful how you mentored me before I started my PhD.

Thank you, **Paula-Marie Ivey**, you are the longest person I have known outside of my family and went through my highs and lows. Eternally thankful and I think you will agree and say a mi fi tell yuh. **David Cabrerizo Granados**, you are a wonderful friend who not only has made my life less boring, but also side hustled as my part-time counselor.

Thank you to my sisters, **Julia** and **Mariella**, for covering things for me at home and giving me the emotional support and letting me vent. You have put up with me all my life and I would not have asked for anyone else. Thank you, **Mom** and **Dad**, for encouraging me to strive for the best I can in all I do. My deepest appreciation and

gratitude to my three **Aunts**, who each in their own way, took care of me and became my go-to person for different problems in my life.

Last but not least, **Gregory Alan Isakov**, your music is art and always brings me peace, thank you.

9 Declaration about the use of generative AI

No AI assisted tools were used in the writing and I take full responsibility for the content of the “kappa”/comprehensive summary of the thesis.

10 References

1. McCulloch, E. A. & Till, J. E. The radiation sensitivity of normal mouse bone marrow cells, determined by quantitative marrow transplantation into irradiated mice. *Radiat Res* **13**, (1960).
2. Babovic, S. & Eaves, C. J. Hierarchical organization of fetal and adult hematopoietic stem cells. *Exp Cell Res* **329**, (2014).
3. Cheng, H., Zheng, Z. & Cheng, T. New paradigms on hematopoietic stem cell differentiation. *Protein Cell* **11**, (2020).
4. Holinstat, M. Normal platelet function. *Cancer Metastasis Rev* **36**, (2017).
5. Donadee, C. *et al.* Nitric Oxide Scavenging by Red Blood Cell Microparticles and Cell-Free Hemoglobin as a Mechanism for the Red Cell Storage Lesion. *Circulation* **124**, (2011).
6. Helms, C. C., Gladwin, M. T. & Kim-Shapiro, D. B. Erythrocytes and Vascular Function: Oxygen and Nitric Oxide. *Front Physiol* **22**, (2018).
7. Krystel-Whittemore, M., Dileepan, K. N. & Wood, J. G. Mast Cell: A Multi-Functional Master Cell. *Front Immuol* **6**, (2015).
8. De Kleer, I., Willems, F., Lambrecht, B. & Goriely, S. Ontogeny of Myeloid Cells. *Front Immuol* **5**, (2014).
9. Vivier, E., Tomasello, E., Baratin, M., Walzer, T. & Ugolini, S. Functions of natural killer cells. *Nat Immunol* **9**, (2008).
10. Kumar, B. V., Connors, T. J. & Farber, D. L. Human T Cell Development, Localization, and Function throughout Life. *Immunity* **48**, (2018).
11. LeBien, T. W. & Tedder, T. F. B lymphocytes: how they develop and function. *Blood* **112**, (2008).
12. Song, L., Dong, G., Guo, L. & T Graves, D. The function of dendritic cells in modulating the host response. *Mol Oral Microbiol* **33**, (2018).
13. Benveniste, P. *et al.* Intermediate-Term Hematopoietic Stem Cells with Extended but Time-Limited Reconstitution Potential. *Cell Stem Cell* **8**, (2010).
14. Dykstra, B. *et al.* Long-Term Propagation of Distinct Hematopoietic Differentiation Programs In Vivo. *Cell Stem Cell* **16**, (2007).
15. Ema, H., Morita, Y. & Suda, T. Heterogeneity and hierarchy of hematopoietic stem cells. *Exp Hematol* **42**, (2014).
16. Adolfsson, J. *et al.* Identification of Flt3+ Lympho-Myeloid Stem Cells Lacking Erythro-Megakaryocytic Potential A Revised Road Map for Adult Blood Lineage Commitment. *Cell* **121**, (2005).

17. Iwasaki, H. & Akashi, K. Myeloid Lineage Commitment from the Hematopoietic Stem Cell. *Immunity* **26**, (2007).
18. Pietras, E. M. *et al.* Functionally Distinct Subsets of Lineage-Biased Multipotent Progenitors Control Blood Production in Normal and Regenerative Conditions. *Cell Stem Cell* **17**, (2015).
19. Sommerkamp, P. *et al.* Mouse multipotent progenitor 5 cells are located at the interphase between hematopoietic stem and progenitor cells. *Blood* **137**, (2021).
20. Jacobson, L. O., Simmons, E. L., Marks, E. K. & Eldredge, J. H. Recovery from Radiation Injury. *Science* **113**, (1951).
21. Eaves, C. J. Hematopoietic stem cells: Concepts, definitions, and the new reality. *Blood* **125**, (2015).
22. Osawa, M., Hanada, K., Hamada, H. & Nakauchi, H. Long-Term Lymphohematopoietic Reconstitution by a Single CD34-Low/Negative Hematopoietic Stem Cell. *Science* **273**, (1996).
23. Craig, W., Kay, R., L Cutler, R. & Lansdorp, P. M. Expression of Thy-1 on human hematopoietic progenitor cells. *J Exp Med* **177**, (1993).
24. Christensen, J. L. & Weissman, I. L. Flk-2 is a marker in hematopoietic stem cell differentiation: A simple method to isolate long-term stem cells. *Proc Natl Acad Sci U S A* **98**, (2001).
25. Adolfsson, J. *et al.* Upregulation of Flt3 expression within the bone marrow Lin-Sca1+c-kit+ stem cell compartment is accompanied by loss of self-renewal capacity. *Immunity* **15**, (2001).
26. Goodell, M. A., Brose, K., Paradis, G., Conner, A. S. & Mulligan, R. C. Isolation and functional properties of murine hematopoietic stem cells that are replicating in vivo. *J Exp Med* **183**, (1996).
27. Weksberg, D. C., Chambers, S. M., Boles, N. C. & Goodell, M. A. CD150- side population cells represent a functionally distinct population of long-term hematopoietic stem cells. *Blood* **111**, (2008).
28. Kim, M., Cooper, D. D., Hayes, S. F. & Spangrude, G. J. Rhodamine-123 Staining in Hematopoietic Stem Cells of Young Mice Indicates Mitochondrial Activation Rather Than Dye Efflux. *Blood* **91**, (1998).
29. Kiel, M. J. *et al.* SLAM Family Receptors Distinguish Hematopoietic Stem and Progenitor Cells and Reveal Endothelial Niches for Stem Cells. *Cell* **121**, (2005).

30. Oguro, H., Ding, L. & Morrison, S. J. SLAM Family Markers Resolve Functionally Distinct Subpopulations of Hematopoietic Stem Cells and Multipotent Progenitors. *Cell Stem Cell* **13**, (2013).
31. Purton, L. E. Adult murine hematopoietic stem cells and progenitors: an update on their identities, functions, and assays. *Exp Hematol* **116**, (2022).
32. Camargo, F. D., Chambers, S. M., Drew, E., McNagny, K. M. & Goodell, M. A. Hematopoietic stem cells do not engraft with absolute efficiencies. *Blood* **107**, (2006).
33. Wilson, N. K. *et al.* Combined Single-Cell Functional and Gene Expression Analysis Resolves Heterogeneity within Stem Cell Populations. *Cell Stem Cell* **16**, (2015).
34. Muller-Sieburg, C. E., Cho, R. H., Karlsson, L., Huang, J.-F. & Sieburg, H. B. Myeloid-biased hematopoietic stem cells have extensive self-renewal capacity but generate diminished lymphoid progeny with impaired IL-7 responsiveness. *Blood* **103**, (2004).
35. Benz, C. *et al.* Hematopoietic Stem Cell Subtypes Expand Differentially during Development and Display Distinct Lymphopoietic Programs. *Cell Stem Cell* **10**, (2012).
36. Ghersi, J. J. *et al.* Haematopoietic stem and progenitor cell heterogeneity is inherited from the embryonic endothelium. *Nat Cell Biol* **25**, (2023).
37. Strålin, K. B. *et al.* Platelet and myeloid lineage biases of transplanted single perinatal mouse hematopoietic stem cells. *Cell Res* **33**, (2023).
38. Morita, Y., Ema, H. & Nakauchi, H. Heterogeneity and hierarchy within the most primitive hematopoietic stem cell compartment. *J Exp Med* **207**, (2010).
39. Challen, G. A. & Goodell, M. A. The Side Population Contains Functionally Distinct Hematopoietic Stem Cell Subpopulations. *Blood* **112**, (2008).
40. Uchida, N., Dykstra, B., Lyons, K. J., Leung, F. Y. K. & Eaves, C. J. Different in vivo repopulating activities of purified hematopoietic stem cells before and after being stimulated to divide in vitro with the same kinetics. *Exp Hematol* **31**, (2003).
41. Carrelha, J. *et al.* Hierarchically related lineage-restricted fates of multipotent haematopoietic stem cells. *Nature* **554**, (2018).
42. Zhao, M. *et al.* N-Cadherin-Expressing Bone and Marrow Stromal Progenitor Cells Maintain Reserve Hematopoietic Stem Cells. *Cell Rep* **26**, (2019).
43. Mann, M. *et al.* Heterogeneous Responses of Hematopoietic Stem Cells to Inflammatory Stimuli Are Altered with Age. *Cell Rep* **25**, (2018).

44. Somuncular, E. *et al.* CD49b identifies functionally and epigenetically distinct subsets of lineage-biased hematopoietic stem cells. *Stem Cell Reports* **17**, (2022).
45. Boyer, S. W., Beaudin, A. E. & Forsberg, E. C. Mapping differentiation pathways from hematopoietic stem cells using Flk2/Flt3 lineage tracing. *Cell Cycle* **11**, (2012).
46. Gerrits, A. *et al.* Cellular barcoding tool for clonal analysis in the hematopoietic system. *Blood* **115**, (2010).
47. Pei, W. *et al.* Polylox barcoding reveals haematopoietic stem cell fates realized in vivo. *Nature* **548**, (2017).
48. Rodriguez-Fraticelli, A. E. *et al.* Clonal analysis of lineage fate in native hematopoiesis. *Nature* **553**, (2018).
49. Sauer, B. & Henderson, N. Site-specific DNA recombination in mammalian cells by the Cre recombinase of bacteriophage P1. *Proc Natl Acad Sci U S A* **85**, (1988).
50. Lu, R., Neff, N. F., Quake, S. R. & Weissman, I. L. Tracking single hematopoietic stem cells in vivo using high-throughput sequencing in conjunction with viral genetic barcoding. *Nat Biotechnol* **29**, (2011).
51. Busch, K. *et al.* Fundamental properties of unperturbed haematopoiesis from stem cells in vivo. *Nature* **518**, (2015).
52. Sun, J. *et al.* Clonal dynamics of native haematopoiesis. *Nature* **514**, (2014).
53. Chapple, R. H. *et al.* Lineage tracing of murine adult hematopoietic stem cells reveals active contribution to steady-state hematopoiesis. *Blood Adv* **2**, (2018).
54. Sanjuan-Pla, A. *et al.* Platelet-biased stem cells reside at the apex of the haematopoietic stem-cell hierarchy. *Nature* **502**, (2013).
55. Carrelha, J. *et al.* Alternative platelet differentiation pathways initiated by nonhierarchically related hematopoietic stem cells. *Nat Immunol* **25**, (2024).
56. Pucella, J. N., Upadhaya, S. & Reizis, B. The Source and Dynamics of Adult Hematopoiesis: Insights from Lineage Tracing. *Annu Rev Cell Dev Biol* **36**, (2020).
57. Bowie, M. B. *et al.* Identification of a new intrinsically timed developmental checkpoint that reprograms key hematopoietic stem cell properties. *Proc Natl Acad Sci U S A* **104**, (2007).

58. Bowie, M. B. *et al.* Hematopoietic stem cells proliferate until after birth and show a reversible phase-specific engraftment defect. *J Clin Invest* **116**, (2006).
59. Copley, M. R. & Eaves, C. J. Developmental changes in hematopoietic stem cell properties. *Exp Mol Med* **45**, (2013).
60. Groarke, E. M. & Young, N. S. Aging and Hematopoiesis. *Clin Geriatr Med* **35**, (2019).
61. Chambers, S. M. *et al.* Aging Hematopoietic Stem Cells Decline in Function and Exhibit Epigenetic Dysregulation. *PLoS Biol* **5**, (2007).
62. Morrison, S. J., Wandycz, A. M., Akashi, K., Globerson, A. & Weissman, I. L. The aging of hematopoietic stem cells. *Nat Med* **2**, (1996).
63. Sudo, K., Ema, H., Morita, Y. & Nakauchi, H. Age-Associated Characteristics of Murine Hematopoietic Stem Cells. *J Exp Med* **192**, (2000).
64. Dykstra, B., Olthof, S., Schreuder, J., Ritsema, M. & de Haan, G. Clonal analysis reveals multiple functional defects of aged murine hematopoietic stem cells. *J Exp Med* **208**, (2011).
65. Snoeck, H.-W. Aging of the hematopoietic system. *Curr Opin Hematol* **20**, (2013).
66. Beerman, I. *et al.* Functionally distinct hematopoietic stem cells modulate hematopoietic lineage potential during aging by a mechanism of clonal expansion. *Proc Natl Acad Sci U S A* **107**, (2010).
67. Yanada, M. & Naoe, T. Acute myeloid leukemia in older adults. *Int J Hematol* **96**, (2012).
68. Cho, R. H., Sieburg, H. B. & Muller-Sieburg, C. E. A new mechanism for the aging of hematopoietic stem cells: aging changes the clonal composition of the stem cell compartment but not individual stem cells. *Blood* **111**, (2008).
69. Montecino-Rodriguez, E. *et al.* Lymphoid-Biased Hematopoietic Stem Cells Are Maintained with Age and Efficiently Generate Lymphoid Progeny. *Stem Cell Reports* **12**, (2019).
70. Beerman, I. Accumulation of DNA damage in the aged hematopoietic stem cell compartment. *Semin Hematol* **12**, (2017).
71. de Haan, G. & Lazare, S. S. Aging of hematopoietic stem cells. *Blood* **131**, (2018).
72. Ho, T. T. *et al.* Autophagy maintains the metabolism and function of young and old stem cells. *Nature* **543**, (2017).
73. Kim, K. M. *et al.* Taz protects hematopoietic stem cells from an aging-dependent decrease in PU.1 activity. *Nat Commun* **13**, (2022).

74. Sun, D. *et al.* Epigenomic Profiling of Young and Aged HSCs Reveals Concerted Changes during Aging that Reinforce Self-Renewal. *Cell Stem Cell* **14**, (2014).
75. Rossi, D. J. *et al.* Cell intrinsic alterations underlie hematopoietic stem cell aging. *Proc Natl Acad Sci U S A* **102**, (2005).
76. Kowalczyk, M. S. *et al.* Single-cell RNA-seq reveals changes in cell cycle and differentiation programs upon aging of hematopoietic stem cells. *Genome Res* **25**, (2015).
77. Svendsen, A. F. *et al.* A comprehensive transcriptome signature of murine hematopoietic stem cell aging. *Blood* **138**, (2021).
78. Nimmo, R. A., May, G. E. & Enver, T. Primed and ready: understanding lineage commitment through single cell analysis. *Trends Cell Biol* **25**, (2015).
79. Chambers, S. M. *et al.* Hematopoietic Fingerprints: An Expression Database of Stem Cells and Their Progeny. *Cell Stem Cell* **1**, (2007).
80. Jaenisch, R. & Bird, A. Epigenetic regulation of gene expression: how the genome integrates intrinsic and environmental signals. *Nat Genet* **33**, (2003).
81. Stein, G. S. *et al.* Transcription factor-mediated epigenetic regulation of cell growth and phenotype for biological control and cancer. *Adv Enzyme Regul.* **50**, (2009).
82. Attema, J. L. *et al.* Epigenetic characterization of hematopoietic stem cell differentiation using miniChIP and bisulfite sequencing analysis. *Proc Natl Acad Sci U S A* **104**, (2007).
83. Doi, Y. *et al.* SATB1 Expression Marks Lymphoid-Lineage-Biased Hematopoietic Stem Cells in Mouse Bone Marrow. *Blood* **126**, (2015).
84. Rimmelé, P. *et al.* Aging-like Phenotype and Defective Lineage Specification in SIRT1-Deleted Hematopoietic Stem and Progenitor Cells. *Stem Cell Reports* **3**, (2014).
85. Kueh, H. Y., Champhekhar, A., Nutt, S. L., Elowitz, M. B. & Rothenberg, E. V. Positive feedback between PU.1 and the cell cycle controls myeloid differentiation. *Science* **341**, (2013).
86. Jassinskaja, M., Gonka, M. & Kent, D. G. Resolving the hematopoietic stem cell state by linking functional and molecular assays. *Blood* **142**, (2023).
87. Meng, Y. *et al.* Epigenetic programming defines haematopoietic stem cell fate restriction. *Nat Cell Biol* **25**, (2023).
88. Yu, W. C. V. *et al.* Epigenetic Memory Underlies Cell-Autonomous Heterogeneous Behavior of Hematopoietic Stem Cells. *Cell* **168**, (2017).

89. Morikawa, T. & Takubo, K. Hypoxia regulates the hematopoietic stem cell niche. *Pflugers Archiv* **468**, (2016).
90. Pinho, S. & Frenette, P. S. Haematopoietic stem cell activity and interactions with the niche. *Nat Rev Mol Cell Biol* **20**, (2019).
91. Ding, L., Saunders, T. L., Enikolopov, G. & Morrison, S. J. Endothelial and perivascular cells maintain haematopoietic stem cells. *Nature* **481**, (2012).
92. Nakamura-Ishizu, A. & Suda, T. Dynamic Changes in the Niche with N-Cadherin Revisited: The HSC "Niche Herein". *Cell Stem Cell* **24**, (2019).
93. Wei, Q. & Frenette, P. S. Niches for Hematopoietic Stem Cells and Their Progeny. *Immunity* **48**, (2018).
94. Blank, U., Karlsson, G. & Karlsson, S. Signaling pathways governing stem-cell fate. *Blood* **111**, (2008).
95. Luis, T. C., Killmann, N. M. B. & Staal, F. J. T. Signal transduction pathways regulating hematopoietic stem cell biology: Introduction to a series of Spotlight Reviews. *Leukemia* **26**, (2012).
96. Challen, G. A., Boles, N. C., Chambers, S. M. & Goodell, M. A. Distinct Hematopoietic Stem Cell Subtypes Are Differentially Regulated by TGF-β1. *Cell Stem Cell* **6**, (2010).
97. Roy, I. M. *et al.* Inhibition of SRC-mediated integrin signaling in bone marrow niche enhances hematopoietic stem cell function. *iScience* **25**, (2022).
98. Nemeth, M. J. & Bodine, D. M. Regulation of hematopoiesis and the hematopoietic stem cell niche by Wnt signaling pathways. *Cell Res* **17**, (2007).
99. Bottardi, S., Aumont, A., Grosveld, F. & Milot, E. Developmental stage–specific epigenetic control of human β-globin gene expression is potentiated in hematopoietic progenitor cells prior to their transcriptional activation. *Blood* **102**, (2003).
100. Pang, W. W. *et al.* Human bone marrow hematopoietic stem cells are increased in frequency and myeloid-biased with age. *Proc Natl Acad Sci U S A* **108**, (2011).
101. Pang, W. W., Schrier, S. L. & Weissman, I. L. Age-associated changes in human hematopoietic stem cells. *Semin Hematol* **54**, (2017).
102. Mitchell, S. J., Scheibye-Knudsen, M., Longo, D. L. & De Cabo, R. Animal models of aging research: Implications for human aging and age-related diseases. *Annu Rev Anim Biosci* **3**, (2015).
103. Lichtman, M. A. & M Rowe, J. The relationship of patient age to the pathobiology of the clonal myeloid diseases. *Semin Oncol* **31**, (2004).

104. Linton, P. J. & Dorshkind, K. Age-related changes in lymphocyte development and function. *Nat Immunol* **5**, (2004).
105. Ogawa, T., Kitagawa, M. & Hirokawa, K. Age-related changes of human bone marrow: a histometric estimation of proliferative cells, apoptotic cells, T cells, B cells and macrophages. *Mech Ageing Dev* **117**, (2000).
106. Tong, J. *et al.* Hematopoietic Stem Cell Heterogeneity Is Linked to the Initiation and Therapeutic Response of Myeloproliferative Neoplasms. *Cell Stem Cell* **28**, (2021).
107. Epah, J. & Schäfer, R. Implications of hematopoietic stem cells heterogeneity for gene therapies. *Gene Ther* **28**, (2021).
108. Chen, I. C., Sethy, B. & Liou, J. P. Recent Update of HDAC Inhibitors in Lymphoma. *Front Cell Dev Biol* **8**, (2020).
109. Lernoux, M., Schnekenburger, M., Dicato, M. & Diederich, M. Epigenetic mechanisms underlying the therapeutic effects of HDAC inhibitors in chronic myeloid leukemia. *Biochem Pharmacol* **173**, (2020).
110. Ploemacher, R. E., Van Der Sluijs, J. P., Voerman, J. S. A. & Brons, N. H. C. An in vitro limiting-dilution assay of long-term repopulating hematopoietic stem cells in the mouse. *Blood* **74**, (1989).
111. Gustafsson, C. *et al.* T-RHEX-RNAseq – a tagmentation-based, rRNA blocked, random hexamer primed RNAseq method for generating stranded RNAseq libraries directly from very low numbers of lysed cells. *BMC Genomics* **24**, (2023).
112. Kent, D. G. *et al.* Prospective isolation and molecular characterization of hematopoietic stem cells with durable self-renewal potential. *Blood* **113**, (2009).
113. Li, L. *et al.* A mouse model with high clonal barcode diversity for joint lineage, transcriptomic, and epigenomic profiling in single cells. *Cell* **186**, (2023).
114. Weinreb, C., Rodriguez-Fraticelli, A., Camargo, F. D. & Klein, A. M. Lineage tracing on transcriptional landscapes links state to fate during differentiation. *Science* **367**, (2020).
115. Ferrari, S. *et al.* Efficient gene editing of human long-term hematopoietic stem cells validated by clonal tracking. *Nat Biotechnol* **38**, (2020).
116. Liao, J. *et al.* Therapeutic adenine base editing of human hematopoietic stem cells. *Nat Commun* **14**, (2023).
117. Ferrari, S. *et al.* Gene Editing of Hematopoietic Stem Cells: Hopes and Hurdles Toward Clinical Translation. *Front in Genome Ed* **3**, (2021).

118. Gao, S. *et al.* Identification of HSC/MPP expansion units in fetal liver by single-cell spatiotemporal transcriptomics. *Cell Res* **32**, (2022).
119. Dong, R. *et al.* Using Spatial Transcriptomics to Reveal Fetal Liver Hematopoietic Stem Cell-Niche Interactions. *Blood* **138**, (2021).
120. Lu, R., Czechowicz, A., Seita, J., Jiang, D. & Weissman, I. L. Clonal-level lineage commitment pathways of hematopoietic stem cells in vivo. *Proc Natl Acad Sci U S A* **116**, (2019).
121. McIntosh, B. E. *et al.* Nonirradiated NOD,B6.SCID Il2rγ-/- kitW41/W41 (NBSGW) mice support multilineage engraftment of human hematopoietic cells. *Stem Cell Reports* **4**, (2015).
122. Rodriguez y Baena, A. *et al.* New transgenic mouse models enabling pan-hematopoietic or selective hematopoietic stem cell depletion in vivo. *Sci Rep* **12**, (2022).
123. Wilkinson, A. C. *et al.* Long-term ex vivo haematopoietic-stem-cell expansion allows nonconditioned transplantation. *Nature* **571**, (2019).
124. Moignard, V. *et al.* Characterization of transcriptional networks in blood stem and progenitor cells using high-throughput single-cell gene expression analysis. *Nat Cell Biol* **15**, (2013).
125. Rotem, A. *et al.* Single-cell ChIP-seq reveals cell subpopulations defined by chromatin state. *Nat Biotechnol* **33**, (2015).
126. Grosselin, K. *et al.* High-throughput single-cell ChIP-seq identifies heterogeneity of chromatin states in breast cancer. *Nat Genet* **51**, (2019).
127. Skene, P. J. & Henikoff, S. An efficient targeted nuclease strategy for high-resolution mapping of DNA binding sites. *eLife* **6**, (2017).
128. Zhu, Q., Liu, N., Orkin, S. H. & Yuan, G. C. CUT and RUNTools: A flexible pipeline for CUT and RUN processing and footprint analysis. *Genome Biol* **20**, (2019).
129. Kaya-Okur, H. S. *et al.* CUT&Tag for efficient epigenomic profiling of small samples and single cells. *Nat Commun* **10**, (2019).
130. Bartosovic, M., Kabbe, M. & Castelo-Branco, G. Single-cell CUT&Tag profiles histone modifications and transcription factors in complex tissues. *Nat Biotechnol* **39**, (2021).
131. Reyes, M., Billman, K., Hacohen, N. & Blainey, P. C. Simultaneous Profiling of Gene Expression and Chromatin Accessibility in Single Cells. *Adv Biosyst* **3**, (2019).

132. Chen, A. F. *et al.* NEAT-seq: simultaneous profiling of intra-nuclear proteins, chromatin accessibility and gene expression in single cells. *Nat Methods* **19**, (2022).
133. Horvath, S. DNA methylation age of human tissues and cell types. *Genome Biol* **14**, (2013).
134. Han, Y. *et al.* Epigenetic age-predictor for mice based on three CpG sites. *eLife* **7**, (2018).
135. Weidner, C. I. *et al.* Aging of blood can be tracked by DNA methylation changes at just three CpG sites. *Genome Biol* **15**, (2014).
136. Lee, Y. *et al.* Blood-based epigenetic estimators of chronological age in human adults using DNA methylation data from the Illumina MethylationEPIC array. *BMC Genomics* **21**, (2020).
137. Seale, K., Teschendorff, A., Reiner, A. P., Voisin, S. & Eynon, N. A comprehensive map of the aging blood methylome in humans. *Genome Biol* **25**, (2024).
138. Jassinskaja, M. *et al.* Comprehensive Proteomic Characterization of Ontogenic Changes in Hematopoietic Stem and Progenitor Cells. *Cell Rep* **21**, (2017).
139. Xiu, Y. *et al.* Proteomic characterization of murine hematopoietic stem progenitor cells reveals dynamic fetal-to-adult changes in metabolic-related pathways. *Biochem Biophys Res Commun* **734**, (2024).
140. Zaro, B. W. *et al.* Proteomic analysis of young and old mouse hematopoietic stem cells and their progenitors reveals post-transcriptional regulation in stem cells. *eLife* **9**, (2020).
141. Hidalgo San Jose, L. *et al.* Modest Declines in Proteome Quality Impair Hematopoietic Stem Cell Self-Renewal. *Cell Rep* **30**, (2020).
142. Munz, C. M. *et al.* Regeneration after blood loss and acute inflammation proceeds without contribution of primitive HSCs. *Blood* **141**, (2023).
143. Vanickova, K. *et al.* Hematopoietic stem cells undergo a lymphoid to myeloid switch in early stages of emergency granulopoiesis. *EMBO J* **42**, (2023).
144. Dietlein, N. *et al.* Usp22 is an intracellular regulator of systemic emergency hematopoiesis. *Sci Immunol* **7**, (2022).
145. Kuribayashi, W. *et al.* Limited rejuvenation of aged hematopoietic stem cells in young bone marrow niche. *J Exp Med* **218**, (2021).

146. Miao, R. *et al.* Competition between hematopoietic stem and progenitor cells controls hematopoietic stem cell compartment size. *Nat Commun* **13**, (2022).
147. Pinho, S. *et al.* Lineage-Biased Hematopoietic Stem Cells Are Regulated by Distinct Niches. *Dev Cell* **44**, (2018).
148. Ho, Y. H. & Méndez-Ferrer, S. Microenvironmental contributions to hematopoietic stem cell aging. *Haematologica* **105**, (2020).
149. Ramalingam, P. *et al.* Restoring bone marrow niche function rejuvenates aged hematopoietic stem cells by reactivating the DNA Damage Response. *Nat Commun* **14**, (2023).
150. Ergen, A. V., Boles, N. C. & Goodell, M. A. Rantes/Ccl5 influences hematopoietic stem cell subtypes and causes myeloid skewing. *Blood* **119**, (2012).
151. Baldridge, M. T., King, K. Y. & Goodell, M. A. Inflammatory signals regulate hematopoietic stem cells. *Trends Immunol* **32**, (2011).
152. Purton, L. E. & Scadden, D. T. Limiting Factors in Murine Hematopoietic Stem Cell Assays. *Cell Stem Cell* **1**, (2007).
153. Adigbli, G. *et al.* Development of LT-HSC-Reconstituted Non-Irradiated NBSGW Mice for the Study of Human Hematopoiesis In Vivo. *Front Immunol* **12**, (2021).
154. Wong, W. M. *et al.* Expression of integrin a2 receptor in human cord blood CD341CD382CD901 stem cells engrafting long-term in nod/scid-il2rccnull mice. *Stem Cells* **31**, (2013).
155. Lee-Six, H. & Kent, D. G. Tracking hematopoietic stem cells and their progeny using whole-genome sequencing. *Exp Hematol* **83**, (2020).
156. Zhang, X. *et al.* An immunophenotype-coupled transcriptomic atlas of human hematopoietic progenitors. *Nat Immunol* **25**, (2024).
157. Castor, A. *et al.* Distinct patterns of hematopoietic stem cell involvement in acute lymphoblastic leukemia. *Nat Med* **11**, (2005).
158. Nagel, I. *et al.* Hematopoietic stem cell involvement in BCR-ABL1–positive ALL as a potential mechanism of resistance to blinatumomab therapy. *Blood* **130**, (2017).
159. Marnell, C. S., Bick, A. & Natarajan, P. Clonal hematopoiesis of indeterminate potential (CHIP): Linking somatic mutations, hematopoiesis, chronic inflammation and cardiovascular disease. *J Mol Cell Cardiol* **161**, (2021).
160. Steensma, D. P. *et al.* Clonal hematopoiesis of indeterminate potential and its distinction from myelodysplastic syndromes. *Blood* **126**, (2015).

161. Dunn, W. G., McLoughlin, M. A. & Vassiliou, G. S. Clonal hematopoiesis and hematological malignancy. *J Clin Invest* **134**, (2024).
162. Jakobsen, N. A. *et al.* Selective advantage of mutant stem cells in human clonal hematopoiesis is associated with attenuated response to inflammation and aging. *Cell Stem Cell* **31**, (2024).