From the Department of Medicine, Huddinge
Karolinska Institutet, Stockholm, Sweden

HETEROGENEITY IN PLATELET DIFFERENTIATION PATHWAYS

Axel Winroth

Karolinska Institutet

Stockholm 2025

Published by Karolinska Institutet.
Printed by Universitetsservice US-AB, 2025

ISBN 978-91-8017-630-9
DOI 10.69622/29064818
Cover illustration: Platelet differentiation pathways initiated by distinct hematopoietic stem cells

HETEROGENEITY IN PLATELET DIFFERENTIATION PATHWAYS

Thesis for Doctoral Degree (Ph.D.)

By

Axel Winroth

The public defense of this thesis will be held in Erna Möllersalen, Neo, Blickagången 16, 141 52 Huddinge

September 30th 2025 at 09:00

Principal Supervisor:
Dr Petter S. Woll
Karolinska Institutet
Department of Medicine Huddinge
Center for Hematology and Regenerative Medicine (HERM)

Co-supervisor:
Professor Sten Eirik W. Jacobsen
Karolinska Institutet
Department of Medicine Huddinge
Center for Hematology and Regenerative Medicine (HERM)
Department of Cell and Molecular Biology

Opponent:
Professor David Kent
University of York
Department of Biology

Examination Board:
Dr Martin Enge
Karolinska Institutet
Department of Medicine, Solna

Dr Sofia Ramström
Örebro Universitet
School of Medical Sciences

Professor Cristian Bellodi
Københavns Universitet
Faculty of Health and Medical Sciences
Biotech Research & Innovation Centre

Those who search for gold dig up much earth and find little.

Heraclitus B22 in Early Greek Philosophy, Volume III: Early Ionian Thinkers, Part 2. Edited and translated by André Laks, Glenn W. Most. Loeb Classical Library 526. Cambridge, MA: Harvard University Press, 2016.

Popular science summary of the thesis

Blood cells are vital for oxygen transport, immune function, and blood clotting. Billions of blood cells are generated every day by so-called hematopoietic stem cells that reside in the bone marrow. This process is called hematopoiesis and has been thoroughly investigated for over 60 years. During hematopoiesis mature blood cells develop from the hematopoietic stem cells through a series of intermediate cell types. Studies on hematopoiesis have led to new insights in fundamental stem cell biology and new treatments for severe diseases such as blood cancer. Several established treatments for blood cancer, for example chemotherapy, can lead to acute depletion of blood cells. Low platelet numbers lead to impaired blood clotting, which increases the risk of severe bleeding that can be fatal.

Not all hematopoietic stem cells are created the same. The most common type is the multilineage hematopoietic stem cell that gives rise to all blood cell types. Other stem cells are biased towards some lineages. Our group has previously identified a subset of hematopoietic stem cells that upon transplantation preferentially generates platelets whilst retaining the capacity to produce all types of blood cells. We call these cells platelet-biased or -restricted hematopoietic stem cells and have found them enriched in the fraction of the hematopoietic stem cells expressing the protein von Willebrand factor. This protein is also found in platelets and blood vessel walls and plays a key role in hemostasis. The work in this thesis further investigates these platelet-restricted and -biased hematopoietic stem cells.

In **Study I** we investigated the onset of lineage restriction and bias in hematopoietic stem cells during fetal development. During development, hematopoietic stem cells come from blood vessel walls in the aorta and expand in the fetal liver before settling in the bone marrow. Previous studies have indicated that hematopoietic stem cells in the fetal liver all have high expression of von Willebrand factor, which suggests a high abundance of platelet-biased and -restricted hematopoietic stem cells. To investigate the prevalence of lineage bias and restriction before birth we performed multiple single hematopoietic stem cell transplantations from fetal liver. Contrary to our hypothesis, we found that most hematopoietic stem cells at this early stage of development were multilineage. We also performed single-cell RNA sequencing, a molecular technique that gives the expression of all genes in a cell, on the hematopoietic stem cells from the fetal

liver just before birth. This analysis showed that several platelet genes correlated with von Willebrand factor expression, which is indicative of transcriptional priming of a subset of hematopoietic stem cells towards the platelet fate already at the time of birth.

In **Study II** we turned our attention to the question of how hematopoietic stem cells generate platelets. Studies from other laboratories have suggested that platelet-biased and -restricted hematopoietic stem cells can generate platelets through a shortcut, skipping some established progenitor cell intermediates. In theory this would allow for faster platelet generation, which would be useful in situations where platelets are depleted. To address this question, we performed single-cell transplantations of von Willebrand factor expressing and non-expressing hematopoietic stem cells and observed their output. These experiments showed that von Willebrand factor-positive platelet-restricted and von Willebrand factor-negative multilineage hematopoietic stem cells do not generate each other. We next sorted the hematopoietic stem and progenitor cells originating from platelet-restricted and multilineage hematopoietic stem cells and again performed single-cell RNA sequencing. This identified several genes that separated the differentiation path of multilineage and platelet-restricted hematopoietic stem cells. Among the genes was *Flt3* that we validated in a mouse model in which the cells switch color dependent on *Flt3* expression. Transplantation of hematopoietic stem cells from this mouse model showed that multilineage hematopoietic stem cells always generate platelets through *Flt3* expressing intermediate cells, whereas their platelet-restricted counterparts do not. Finally, we showed that the *Flt3* independent pathway is used more after treatment with chemotherapy. Taken together, the results from **Study II** confirmed the existence of a novel platelet generation pathway, initiated by a distinct type of platelet-biased hematopoietic stem cells, which increases in usage upon treatment with chemotherapy.

Platelets play an important role in normal hemostasis and in the development of clots that halt blood flow in the heart and the brain. Over the last decade it has been shown that not all platelets are the same and respond differently to vessel injury. We were interested in finding out what role the hematopoietic stem cells and differentiation pathways from **Study II** play in this process. To address this question, we developed in **Study III** an RNA-seq protocol for sorted platelets. We optimized platelet collection and sorting and could prepare sequencing data from

as little as 300,000 platelets, which enables future studies on platelet subsets within the same individual and can give insights on platelet subpopulations.

In conclusion, the studies included in this thesis have further characterized platelet-biased and -restricted hematopoietic stem cells and their differentiation path during development and in adult life. It also lays the groundwork for further studies on platelet heterogeneity which has implications for prevention and treatment of cardiovascular disease.

Populärvetenskaplig sammanfattning

Blodets celler är livsnödvändiga för syretransport, immunsystemet och blodstillning. Miljarder blodceller produceras varje dag av så kallade blodstamceller, vilka har sin hemvist i benmärgen. Blodbildningen kallas med ett annat ord för hematopoes och denna process har studerats ingående under de senaste 60 åren. I hematopoesen bildas mogna blodceller från stamceller genom en serie av alltmer specifika mellansteg. Studier av hematopoesen har givit nya insikter om fundamental stamcellsbiologi och lett till nya behandlingar för allvarliga sjukdomar som till exempel mot blodcancer. Flera etablerade behandlingar mot blodcancer som kemoterapi kan leda till akut förlust av blodceller, något som är ett viktigt kliniskt problem. Lågt antal av blodplättar, även kallade trombocyter, kan leda till bristande blodstillning och högre risk för allvarliga blödningar.

Blodstamceller är inte en homogen grupp. Den vanligaste typen är multilinjeblodstamcellen som kan bilda alla celltyper i blodet. Andra blodstamcellstyper producerar mer av vissa celltyper. Vår forskargrupp har tidigare identifierat en undergrupp av blodstamceller som efter transplantation främst bildar trombocyter, fastän de har kvar förmågan att bilda alla typer av blodceller. Detta står i kontrast till de vanliga blodstamcellerna som bildar alla celltyper efter transplantation. För fenomenet att stamceller till största delen producerar en eller bara ett fåtal av de celltyper som är möjliga finns inget bra ord på svenska men ett alternativ skulle kunna vara ”linjevinkling”, från engelskans ”lineage bias”. Proteinet von Willebrand faktor är en markör för den trombocytvinklade stamcellen och kan användas för att anrika dem vid cellsortering. Denna avhandling undersöker egenskaper hos de trombocytvinklade blodstamcellerna och hur dessa bildar trombocyter i jämförelse med vanliga blodstamceller.

I **Delarbete I** undersökte vi när och var linjevinkling uppkommer. Under fosterstadiet uppstår blodstamceller från kärlväggen i aortan för att sedan expandera i levern innan de slutligen migrerar till benmärgen. Tidigare studier har visat att nästan alla blodstamceller i levern under fosterstadiet uttrycker von Willebrand faktor, vilket skulle kunna innebära att det där finns en stor mängd linjevinklade blodstamceller. Vi hittade dock inga bevis för att dessa celler finns i levern under fosterstadiet, utan att i princip alla stamceller beter sig som multilinjeblodstamceller. Vi gjorde också RNA-sekvensering, en molekylärbiologisk teknik som kvantifierar uttrycket av alla gener i en cell, av blodstamceller från levern under fosterstadiet. Denna analys visade att många trombocytgener korrelerade med genuttrycket av von Willebrand faktor i blodstamcellerna, vilket tyder på att de är predisponerade att bilda trombocyter.

I **Delarbete II** riktade vi vår uppmärksamhet på hur blodstamceller bildar trombocyter. Andra forskargrupper har föreslagit att trombocyter kan bildas via en genväg och alltså hoppar över flera etablerade celltyper. Detta skulle kunna innebära att de kan bilda trombocyter snabbare, något som skulle vara användbart i situationer med trombocytbrist. För att undersöka detta sorterade vi stamceller baserat på deras von Willebrand faktoruttryck och övervakade deras cellproduktion. Dessa experiment visade att de två typerna av blodstamceller, multilinje och trombocytvinklade, inte kan ge upphov till varandra. Sedan sorterade vi blodstam- och progenitorceller som genererats av de två olika stamcellstyperna och genomförde återigen RNA-sekvensering för att identifiera nya markörer. Denna analys identifierade flertalet gener som separerade differentieringsvägen från multilinje och trombocytvinklade blodstamceller. Bland dessa gener fanns *Flt3* som en markör för differentieringsvägen för vanliga stamceller men inte för de trombocytvinklade. Vi bekräftade att *Flt3* var en ny markör med hjälp av en musmodell där cellerna byter färg beroende på om de någon gång uttryckt *Flt3*. Transplantation av blodstamceller från denna musmodell visade att multilinjeblodstamceller alltid bildar trombocyter genom mellansteg som uttrycker *Flt3*, medan trombocytvinklade stamceller gör tvärtom. Till sist visade vi att denna nya differentieringsväg expanderar vid behandling med kemoterapi, vilket kan ha betydelse vid cancerbehandling. Tillsammans visade resultaten från **Delarbete II** förekomsten av en ny utvecklingsväg att bilda trombocyter, vilken utgår från en distinkt typ av blodstamceller, och används i högre utsträckning efter behandling med kemoterapi.

Trombocyter spelar en stor roll i de proppar som kan stoppa blodflödet till hjärtat och hjärnan. Det senaste decenniet har det visat sig att inte alla trombocyter är likadana och att de kan reagera olika som svar på skada. Vi är intresserade av att veta vilken roll blodstamcellerna och differentieringsvägen från **Delarbete II** spelar i detta. Som ett steg mot detta mål utvecklade vi i **Delarbete III** en metod för RNA-sekvensering av sorterade trombocyter. Vi optimerade provinsamling, hantering och sortering och kunde på så vis generera sekvensdata från så lite som 300 000 trombocyter, vilket möjliggör studier av trombocytundergrupper i samma individ. Vi hoppas att metoden kommer att kunna användas för att vidare undersöka trombocytheterogenitet.

Sammanfattningsvis har denna avhandling undersökt trombocytvinklade blodstamceller och deras differentieringsväg dels under fosterstadiet, dels i vuxen ålder. Dessutom lägger den grunden för fortsatta studier av trombocytheterogenitet, vilket kan komma att ha betydelse för behandling av kardiovaskulära sjukdomar.

Abstract

Hematopoietic stem cells replenish billions of mature blood cells every day through a series of increasingly lineage-restricted progenitor cells and are therefore vital for immune response, oxygen transport and hemostasis. Upon acute depletion of the hematopoietic system, such as after cytotoxic treatment for blood cancer, one of the most troublesome complications is bleeding due to thrombocytopenia.

Through mouse models, our lab has characterized a novel and direct pathway of platelet generation, initiated by a hitherto unrecognized subset of HSCs, that are activated in response to cytotoxic stress. **Study I** investigates the emergence of lineage restriction and bias in perinatal HSCs. **Study II** describes the non-hierarchical relationship of two distinct types of HSCs, von Willebrand factor positive platelet-restricted and -biased HSCs (P-HSCs) and von Willebrand factor negative multilineage HSCs (M-HSCs). The transcriptional expression of FMS-like tyrosine kinase (*Flt-3*) was found to be a marker that distinguishes the differentiation pathways of P-HSCs and M-HSCs. This finding was validated by fate mapping using a *Flt3*-Cre/Rosa26-tdTomato mouse model. Furthermore, we performed molecular characterization of the megakaryocyte progenitors, the unilineage progenitors responsible for platelet generation, finding distinct molecular profiles depending on the type of HSC that generated them. We therefore speculate that the different pathways can in part explain platelet heterogeneity, an issue that has gained interest due to expansion of prothrombotic subsets upon aging. To this end, **Study III** describes a protocol for RNA-seq of sorted platelet populations allowing for molecular studies of platelet heterogeneity within the same individual.

Taken together, these studies show that a mature cell can be generated through more than one differentiation pathway, suggesting a revision of the established hematopoietic hierarchy that currently posits only one unique sequence of differentiation steps for each mature cell type.

List of scientific papers

I. Platelet and myeloid lineage biases of transplanted single perinatal mouse hematopoietic stem cells (Cell Research, 2023, 33, 883-886)

Karin Belander Strålin*, Joana Carrelha*, **Axel Winroth**, Christoph Ziegenhain, Michael Hagemann-Jensen, Laura M. Kettyle, Amy Hillen, Kari Högstrand, Ellen Markljung, Francesca Grasso, Masafumi Seki, Stefania Mazzi, Yiran Meng, Bishan Wu, Edwin Chari, Madeleine Lehander, Rickard Sandberg, Petter S. Woll, Sten Eirik W. Jacobsen (*equal contribution)

II. Alternative platelet differentiation pathways initiated by non-hierarchically related hematopoietic stem cells (Nature Immunology, 2024, 25, 1007-1019)

Joana Carrelha*, Stefania Mazzi*, **Axel Winroth***, Michael Hagemann-Jensen, Christoph Ziegenhain, Kari Högstrand, Masafumi Seki, Margs S. Brennan, Madeleine Lehander, Bishan Wu, Yiran Meng, Ellen Markljung, Ruggiero Norfo, Hisahi Ishida, Karin Belander Strålin, Francesca Grasso, Christina Simoglu Karali, Affaf Aliouat, Amy Hillen, Edwin Chari, Kimberly Silletti, Supat Thongjuea, Adam J. Mead, Sten Linnarsson, Claus Nerlov, Rickard Sandberg, Tetsuichi Yoshizato, Petter S. Woll, Sten Eirik W. Jacobsen, (*Equal Contribution)

III. The platelet transcriptome is stable across age and is minimally responsive to spontaneous CD62 externalization (Manuscript)

Masafumi Seki*, **Axel Winroth***, Margs Brennan, Kari Högstrand, Tetsuichi Yoshizato, Michael Hagemann-Jensen, Christoph Ziegenhain, Amie Waller, Holly Foster, John W. Semple, Cedric Ghevaert, Sten Eirik W. Jacobsen^, Petter S. Woll^ (*Equal contribution, ^Shared senior authorship)

CONTENTS

List of abbreviations

5-FU	5-Fluorouracil
AGM	Aorta-Gonad-Mesonephros
CFU	Colony forming unit
CFU-S	Spleen colony forming unit
CP	Cyclophosphamide
Cre-ERT	Tamoxifen inducible Cre
DE	Differential Expression
EGFP	Enhanced green fluorescent protein
EHT	Endothelial to hematopoietic transition
FACS	Fluorescence activated cell sorting
FL	Fetal liver
FMO	Fluorescence minus one
GO	Gene Ontology
GSEA	Gene Set Enrichment Analysis
HSC	Hematopoietic stem cell
ITP	Immune thrombocytopenic purpura
LMPP	Lymphoid primed multipotent progenitors
LT-HSC	Long term hematopoietic stem cell

MFI	Mean Fluorescence Intensity
Mk	Megakaryocyte
MkP	Megakaryocyte progenitor
MPP	Multipotent progenitor
P	Platelet
PCA	Principal Component Analysis
PE	Platelet, Erythroid
PEM	Platelet, Erythroid, Myeloid
PEMBT	Platelet, Erythroid, Myeloid, B-cell, T-Cell
PRP	Platelet-rich plasma
RT	Reverse transcription
SCF	Stem cell factor
SI	Stain Index
ST-HSC	Short-term hematopoietic stem cell
tdTomato	Tandem dimer Tomato
THPO	Thrombopoietin
TI	Trajectory Inference
TO	Thiazole orange
t-SNE	t-distributed Stochastic Neighborhood Embedding

TSO	Template Switching Oligonucleotide
UMAP	Uniform Manifold Approximation and Projection
UMI	Unique molecular identifier
VTGG	Von Willebrand factor-tdTomato/Gata1-EGFP
Vwf	Von Willebrand factor
YS	Yolk Sac

1 Introduction

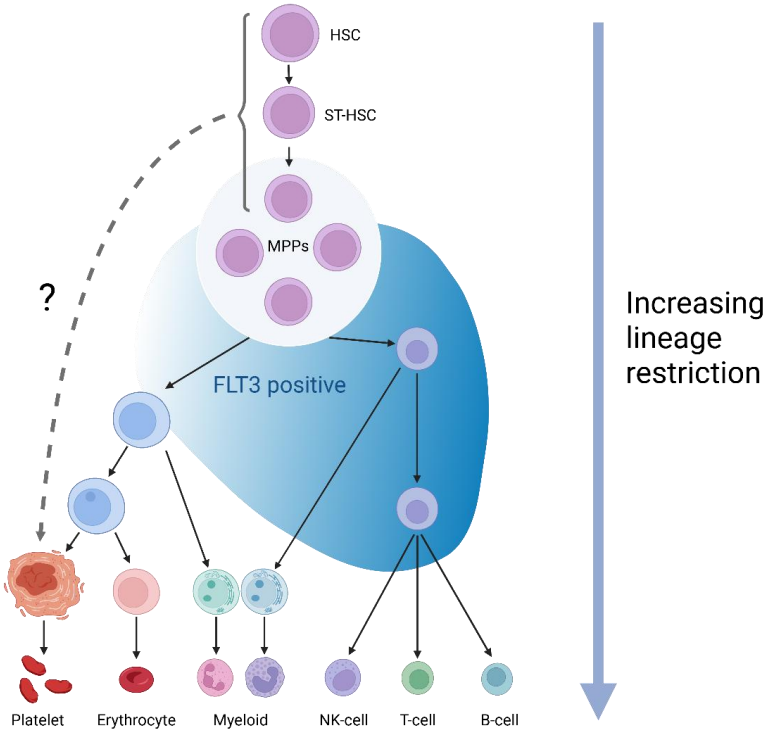

Figure 1 The hematopoietic hierarchy – each mature blood lineage is generated by a series of increasingly lineage-restricted progenitor cells. The thesis investigates the existence of the pathway with a question mark. (HSC – hematopoietic stem cell, ST-HSC – short-term hematopoietic stem cell, MPP – multipotent progenitors.) Created with BioRender.com

All eukaryotic organisms develop from a single cell, the fertilized egg, which gives rise to all tissues through a carefully orchestrated differentiation process. Once a tissue has been established, it is maintained throughout life by a small pool of tissue-specific stem cells. The hematopoietic system, responsible for, among other things, immune response, oxygen transport, and vascular integrity, is one of the most well-characterized stem cell systems (Orkin & Zon, 2008). It has several properties that make it amenable to study, including non-destructive bone marrow and peripheral blood sampling, established *in vivo* and *in vitro* single-cell or clonal assays to study and quantify hematopoietic stem and progenitor cell (HSPC) function, and model systems allowing for genetic modification to identify critical intrinsic and extrinsic hematopoietic regulators.

Blood cells turn over at a rate of 10^{11}–10^{12} cells per day in humans (Gordon et al., 2002). Particularly durable cells are required to sustain this highly dynamic system. The discovery of the hematopoietic stem cell (HSC) led to the establishment of the hematopoietic hierarchy **(Figure 1)**, with the HSC residing at the apex and maintaining the hematopoietic system through generation of increasingly lineage-restricted progenitor cells (Jacobsen & Nerlov, 2019). The self-renewal and multilineage potential of the HSCs also form the basis for bone marrow transplantation (Copelan, 2006), the only curative treatment for several hematological cancers and inherited blood disorders. Initially the HSC compartment was believed to be a homogeneous pool of quiescent cells with similar propensity for self-renewal, multi-lineage differentiation and fate choices. However, as isolation techniques and functional assays have improved, this idea has been challenged leading to several revisions of the hematopoietic hierarchy.

This thesis presents work that has investigated heterogeneity in both fetal and adult HSCs, and how these cells differentiate into mature blood cell lineages. In particular, a novel pathway for platelet differentiation is established.

The first section is a historical overview of how HSCs were first defined. The following sections describe how technological advancements have allowed for continued sub-fractionation of HSC populations, a process that is still ongoing. The section on developmental hematopoiesis is related to **Study I** and deals with the emergence of HSCs during development. In the materials and methods section, the experimental procedures used in the constituent studies are discussed in general with emphasis on technical aspects. Section 5 describes and discusses the main findings from the three constituent papers. The comprehensive summary concludes with some points of perspective where possible follow up studies are suggested together with a broader discussion on the conceptual underpinnings of stem cell research.

2 Background

2.1 Historical perspectives on hematopoietic stem cells

Studies of hematopoiesis were initially motivated by the observation of bone marrow failure after radiation exposure. Mice exposed to lethal doses of ionizing radiation could be protected by lead shielding of the spleen or rescued by transplantation of bone marrow from another mouse (Jacobson et al., 1951). It was later confirmed that transplanted cells were responsible for radioprotection rather than a humoral or chemical component in the transplant, after identifying a marker chromosome both in the donor and recipient bone marrow (Ford et al., 1956). Transplantation also gave rise to splenic nodules, the number of which correlated with the number of injected cells, indicating that nodules were of single cell origin (Till & McCulloch, 1961; McCulloch & Till, 1962). These colonies were subsequently shown to be clonal by induction of chromosomal marks (Becker et al., 1963) in the recipient, after transplantation but before the donor-derived cells had divided. The cells forming the noduli were given the name splenic colony forming unit (CFU-S). Further studies showed that the spleen colonies contained multiple lineages of mature cells at different ratios (McCulloch, 1963; Wu et al., 1968). Some of these colonies could give rise to new spleen nodules in secondary recipients, which implied the existence of self-renewing cells. Taken together, multipotency and self-renewal became the commonly accepted stem cell definition (Siminovitch et al., 1963).

Around the same time, clonal *in vitro* assays were established to test properties of bone marrow cells (Pluznik & Sachs, 1965; Bradley & Metcalf, 1966). By using feeder cells that provide the appropriate microenvironment of both secreted and contact-dependent factors, hematopoietic cells can survive for many divisions. The culture systems also enabled identification of the cytokines by analysis of the conditioned medium generated by the feeder cells. Culture assays provide a complement to the transplantation assay to test for lineage potentials, but are not considered to capture self-renewal potential (Bock, 1997).

A large variation in the frequency of splenic nodules in secondary transplantation was interpreted as heterogeneity at the stem cell level and two models were developed as explanations, the stochastic model which suggests that stem cell heterogeneity is intrinsically regulated (Till et al., 1964), and the microenvironmental model which in contrast argues that local microenvironment determines cell fate (Curry & Trentin, 1967). Both had to be revised when it was

demonstrated that the bone marrow could be functionally subdivided by gradient centrifugation or sedimentation before transplantation (Worton et al., 1969a, 1969b). It was shown that cell size was a determinant of engraftment potential and that engraftable cells are not the same as cells that expand in culture. However, the models were still valid when applied to the individual fractions and form our reasoning about stem cells today. Most studies on hematopoiesis follow this pattern of sub-fractionation and functional evaluation.

It was later shown that the CFU-S assay itself, while indicating stem cell function, did not recapitulate long-term stem cell properties. Transplanted spleen cells generated more CFU-S but gave less radioprotection than bone marrow-derived cells (Kretchmar & Conover, 1970) and treatment with 5-fluorouracil (5-FU), a cytotoxic drug that kills dividing cells, reduced direct CFU-S formation but enhanced bone marrow repopulating ability (Hodgson & Bradley, 1979). These studies suggested the existence of a pre-CFU-S cell. In summary, the CFU-S assay, still actively used in hematopoietic research today (Rodriguez et al., 2021), created the paradigm for HSC research and formed the basis for our current HSC definition: self-renewal and multipotency.

Following these pioneering studies, the continued fractionation of HSPCs was furthered by technological advancements. Monoclonal antibody technology (Kohler & Milstein, 1975) and molecular probes combined with flow cytometry (Herzenberg et al., 2002) enabled prospective isolation of cell populations based on expression of surface proteins in combination with size measurement (Visser et al., 1984). Furthermore, the drawbacks of the CFU-S assay motivated the development of the competitive repopulation assay (Harrison, 1980) which tests bone marrow repopulation after transplantation over extended periods in time rather than short-term CFU-S formation. There are many approaches to separate donor and recipient cells, but they all have in common the introduction of a heritable genetic marker that is passed on to daughter cells during division. Following the chromosomal marker approach from earlier decades, new co-isogenic (congenic) mouse strains allowed separation of transplanted donor cells from endogenous cells. Briefly, congenic mice have near-identical genomes that differ at one locus. The first study on competitive transplantation to assess HSC functional capacity used the s and d variants of the Hemoglobin beta locus, carried by the C57BL/6J and WB/Re mouse strains, to evaluate contribution from the two donors. Later studies used the same approach with the *Ptprc* locus (Spangrude et al., 1988) instead of hemoglobin. The *Ptprc* gene encodes CD45, a

protein tyrosine phosphatase receptor expressed on all nucleated hematopoietic cells and has two alleles *Ptprc*^{a} and *Ptprc*^{b} that encode the CD45.1 and CD45.2 proteins which can be distinguished by monoclonal antibodies (Shen et al., 1986). By transplanting CD45.2-expressing cells into a CD45.1 recipient, the donor-derived cells can be identified and isolated with flow cytometry. As neither mature platelets nor erythrocytes contain nuclei, these two lineages are not captured by analysis of the CD45 congenic markers. However, the development of genetically modified mouse models wherein fluorescent reporters such as enhanced green fluorescent protein (EGFP) or tandem dimer Tomato (tdTomato) are expressed from lineage/stage-specific or pan-hematopoietic promoters has allowed lineage tracing in platelets and erythrocytes (Sanjuan-Pla et al., 2013; Yamamoto et al., 2013).

Alternatively, a fluorescent reporter gene can be introduced inside a housekeeping gene such as *Rosa26* (Srinivas et al., 2001) for constitutive expression, or conditional expression by insertion of a stop codon flanked by *loxP* sites, which are recognized and removed by Cre recombinase expressed under the control of the promoter of a gene of interest (Sauer & Henderson, 1988). The latter approach allows for a binary switch mechanism allowing reverse fate mapping of a mature cell – if the stop signal has been excised and the reporter is detected, the cell has expressed the promoter gene at some point, and all progeny generated from this cell will also express the reporter. In this way the expression history of the protein in each cell can be determined. In theory, this approach can be used to track the output of HSCs given that the selected promoter is completely HSC-specific. However, no such promoter has been identified.

After realizing that primitive cells should not express surface antigens found on mature cells (Lineage negative, Lin^{-}) and, in contrast, be expressing antigens not found on mature cells (Stem cell antigen 1, SCA-1) together with a low Thy1.1 expression, a 1,000-fold enrichment of radioprotection could be achieved (Spangrude et al., 1988). The fractionation was further refined, with the introduction of positive selection on the receptor for stem cell factor, cKIT (Okada et al., 1991), and negative selection on CD34, to allow isolation of a single HSC (Osawa et al., 1996) in limiting dilution with the ability to repopulate lymphoid and myeloid blood lineages in serial transplantations. Additional markers have been identified to further enrich for long-term reconstitution such as SLAM (Kiel et al., 2005) and EPCR (Balazs et al., 2006). In addition to surface markers, stem cells

can be isolated based on low rhodamine-123 retention (Li & Johnson, 1992), or Hoechst efflux (Goodell et al., 1996).

Congenic approaches lack the resolution to trace the contribution of individual clones in a bulk transplant. Transduction of genetic barcodes before transplantation allows for separation of individual clones generated by individual cells after transplantation. This approach is limited by detection sensitivity and clone size – many HSCs contribute very little and most get outcompeted during reconstitution of the hematopoietic system. Like transgenesis, the integration of barcodes might disrupt normal function of the cell. The other option for following single cell clones is to transplant single cells, which is considered the gold standard stem cell assay. Unfortunately, it is very labor intensive and requires prospective enrichment of HSCs before transplantation.

The boundaries of hematopoietic compartments move every time one of them is sub-fractionated into sets with varying self-renewal and lineage potentials. In general, HSPCs lose lineage potential in a stepwise manner until they are fully committed uni-lineage progenitors, but self-renewal and multipotency is not lost at the same rate. Many intermediate stages between long-term HSCs (LT-HSC) and mature cells have been identified. The short-term HSC (ST-HSC) is an HSC with limited self-renewal capacity, meaning that while it reconstitutes all blood lineages, it becomes exhausted and fails to engraft in secondary transplants (Yang et al., 2005). Below the ST-HSC are a set of multipotent progenitors (MPP) with different self-renewal and lineage properties. Among these is the FLT3-expressing lymphoid-biased MPP (LMPP) (Adolfsson et al., 2005), that revised the notion of a strict separation between the myeloid (Akashi et al., 2000) and lymphoid (Kondo et al., 1997) lineages and their respective progenitors **(Figure 1)**.

Modes of cell division are closely related to the stem cell properties. At some point in the division history of a true stem cell, an asymmetric division must occur. Self-renewal requires that one of the daughter cells have the same properties as the dividing cell for the stem cells to be maintained and the other daughter cell must be able to lose its self-renewal and differentiate into all hematopoietic lineages. This has been exploited *in vitro* to show that some cells are multipotent and generate daughter cells that have different lineage potentials (Suda et al., 1984). It has also been used to show both self-renewal and asymmetric division of HSCs *in vitro* (Ema et al., 2000) and refined in transplant experiments where paired

daughters were either both LT-HSCs or one LT-HSC and one cell with less self-renewal potential and/or reduced lineage potential (Yamamoto et al., 2013).

2.2 Hematopoiesis beyond transplantation

Whilst an essential tool for studies of hematopoiesis, the transplantation assays have some important limitations. The means of conditioning in transplantation-based lineage tracing is likely to alter both the transplanted cells and the recipient bone marrow microenvironment (Dong et al., 2020) and it is plausible that cells lacking reconstitution capacity nevertheless contribute to long-term hematopoiesis in an unperturbed setting. An ideal system for lineage tracing in unperturbed hematopoiesis would therefore depend on efficient labeling of all cells in a predefined population without perturbing the cells' current and future states.

Several approaches using genetic engineering have been developed to avoid transplantation altogether and allow monitoring of HSC fate in steady-state hematopoiesis, including the Cre fate-mapping system outlined above. The *Sleeping Beauty* transposon system allows for random integration of small DNA sequences into the genome (Sun et al., 2014; Rodriguez-Fraticelli et al., 2018). This has been exploited in genetically modified mice with an inducible hyperactive transposase, which when activated results in unique integration sites of a known DNA sequence for every targeted cell. Tracking the integration site through next-generation sequencing technologies allows for monitoring the fate of single cells *in vivo*. The Poly*loxp*-express system instead relies on multiple segments flanked by *loxP* sites (Pei et al., 2020). After induction, Cre randomly recombines the segments resulting in unique sequence of inversion events in every cell. The CRISPR array repair lineage tracing (CARLIN) (Bowling et al., 2020) and the improved version DARLIN (Li et al., 2023) models use inducible CRISPR-Cas9-mediated editing to create heritable barcodes. Taken together, molecular barcoding studies have so far identified proliferative HSCs which have a high barcode overlap with mature cell populations and quiescent HSCs where the barcodes are only found within the HSCs pool. Additionally, the lineage biases discovered in classical transplantation experiments have been recapitulated, including HSCs with biased or restricted contribution to mature lineages. Spontaneously occurring variants can also be used as barcodes. Whole genome sequencing of single cell-derived HSC and MPP colonies allows for the construction of a phylogeny of HSPC clones (Kapadia et al., 2025).

Some barcoding studies showed low concordance between HSC barcodes and those detected in fully differentiated populations, challenging the conventional view that adult hematopoiesis is dependent on HSCs (Sun et al., 2014). Subsequent fate-mapping studies addressing this question using different tamoxifen inducible Cre (Cre-ERT) lines initially gave contradicting results dependent on *cis*-regulatory elements. *Tie2* (Busch et al., 2015) labelled a smaller subset of HSCs and showed little contribution to mature lineages. On the contrary, *Fgd5* (Säwen et al., 2018), *Krt18* (Chapple et al., 2018), and *Pdzk1ip1* (Sawai et al., 2016) labelled more HSCs and these studies claimed that HSCs do contribute to steady state hematopoiesis. These conflicting results can be explained by considering that the models label different HSC subsets (Takahashi et al., 2021), highlighting the impact of HSC heterogeneity and choice of Cre promoter elements.

2.3 HSC heterogeneity and lineage bias

HSC heterogeneity is often postulated to explain variability in stem cell assays. For example, the initial observation of different numbers of spleen colonies in secondary transplants of CFU-S was attributed to HSC heterogeneity. Further heterogeneity was identified by studies using random retroviral integration to track individual HSCs' contribution to different hematopoietic lineages. From these studies the HSC compartment could be divided into subsets with different contribution towards lymphoid or myeloid cells (Williams et al., 1984; Dick et al., 1985; Lemischka et al., 1986; Jordan & Lemischka, 1990). HSC transplantation in limiting dilution (Smith et al., 1991) and later single HSC transplantation have also revealed the existence of HSC subsets with different contributions towards different lineages, where distinct lineage patterns of transplanted HSCs are maintained following serial transplantations (Muller-Sieburg et al., 2002; Muller-Sieburg et al., 2004; Dykstra et al., 2007; Kent et al., 2009; Yamamoto et al., 2013). Although the authors of the studies used their own definition of lineage bias, the HSC subsets identified by these different approaches have been reconciled (Ema et al., 2014). Broadly speaking, HSCs are either balanced, myeloid-biased, or lymphoid-biased and have different long-term reconstitution potential. However, lymphoid bias can also be explained by long term B- and T-cell progenitors instead of true HSCs. Several different markers are associated with the different HSC subsets.

Gene expression profiling of HSCs has revealed not only transcriptional heterogeneity but also gene expression signatures overlapping with cells

belonging to mature lineages, suggesting an early lineage priming initiated already at the HSC stage. In particular, many genes associated with the megakaryocyte (Mk) lineage, including expression of von Willebrand factor (Vwf), were also detected in the most primitive HSC (Mansson et al., 2007). This led to the development of the *Vwf*-EGFP reporter mouse (Sanjuan-Pla et al., 2013), which in addition to labelling all platelets with EGFP and thereby allowing reliable tracking of donor-derived platelets in a transplantation setting, also allowed separation of the HSC compartment into Vwf^{+} and Vwf^{-} fractions. Vwf^{+} HSCs showed greater capacity for reconstitution and generated Vwf^{-} HSCs, suggesting elevated stemness. In addition, Vwf^{+} HSCs showed a predisposition to generate the platelet lineage, adding platelet bias as a possible lineage bias. It was also shown that platelet depletion using a CD42b antibody selectively activated quiescent Vwf^{+} HSCs. By combining fluorescent reporters for platelets and erythrocytes in the *Vwf*-tdTomato/*Gata1*-EGFP (VTGG) mouse and CD45.1/CD45.2 analysis, all mature blood lineage could be tracked in reconstitution experiments (Carrelha et al., 2018). Single-cell transplantation of VTGG immunophenotypic HSCs confirmed a hierarchical relationship between multilineage and several types of biased HSCs. The data from that study suggests that lineage bias appears in a stepwise fashion, whereby lineage potential is lost one lineage at a time until only the platelet lineage remains. These studies subdivided the myeloid-biased pattern into platelet-erythroid-myeloid (PEM), platelet-erythroid (PE), and platelet (P) biases. The most pronounced cases of P-bias only showed reconstitution after transplantation in the platelet lineage and were as such the only uni-lineage HSCs detected in these studies. Platelet- and myeloid-biased HSCs increase in frequency with age, shown both by output in transplantation and in transcriptional priming (Grover et al., 2016; Flohr Svendsen et al., 2021). In line with these studies single-cell transplantations using Kusabira-Orange (KuO) for lineage tracing in platelets and erythrocytes suggested the direct generation of platelet-repopulating cells that bypass the multipotent progenitor stage (Yamamoto et al., 2013), although in this study their version of the platelet-biased HSC failed to give long term reconstitution in secondary transplants.

Building on the idea that committed progenitors can be generated independently of multipotent progenitors, a couple of studies have investigated the molecular profile of the platelet differentiation pathway in mice. Single-cell RNA sequencing in a *CALR* mutated mouse model of myeloproliferative disease suggested a transcriptional intermediate cell in-between classical megakaryocyte progenitors

(MkPs), and HSCs, a population that was separated from the established progenitor compartments. This novel population was marked by low CD48 expression, indicating a shortcut between HSCs and MkPs (Prins et al., 2020). Another study combined competitive transplantation with fluorescent label propagation in a mathematical model to infer that there exist two parallel pathways for platelet development (Morcos et al., 2022). These studies did not, however, provide definitive evidence of the pathway based on fate mapping nor were they able to isolate the cells from which the pathways emanate.

The mechanism underlying lineage restriction and bias is an open question. The transplantation studies indicate that it is a long-lasting cell-intrinsic property, supported by single-cell transplantations and its persistence in secondary transplantations. Stated differently, it falls within the stochastic model of hematopoiesis. On the other hand, the microenvironment might be involved in the regulation since *Vwf*$^{+}$ HSCs localize together with sinusoids and Mks rather than in the arteriolar niche (Pinho et al., 2018). Furthermore, low mitochondrial activity is associated with lineage output and stem cell age (Mansell et al., 2021), which is compatible with the reactive oxygen species (ROS) gradient observed in the niche (Itkin et al., 2016). Still, the precise causal relationship between stem cell aging, metabolism and lineage bias remains to be determined.

The studies that simultaneously capture clonal information and mRNA have curated gene signatures that correlate with HSC output patterns linking function to phenotype in a novel way (Rodriguez-Fraticelli et al., 2020). However, genetic barcoding is, as discussed above, limited by low labelling frequency and low barcode diversity making the absence of a barcode in some lineage difficult to interpret — absence of evidence is not evidence of absence. As a result, lack of lineage contribution might in effect reflect a technical dropout. Single-cell HSC transplantation studies are therefore more definitive with regard to interpretation of lineage restriction and lineage bias of HSCs.

The existence of a platelet-biased or -restricted HSC and the corresponding platelet differentiation pathway in humans is also an open question. It has been suggested that HSCs can differentiate into Mks without division; however, it could in fact be a MkP that was selected for differentiation (Roch et al., 2015). Although it is possible to perform single-cell transplantation of human HSCs, these models do not support generation of the Mk lineage. However, Mk lineage expansion has been observed in patients with myeloproliferative disease (Psaila et al., 2020).

2.4 Developmental hematopoiesis in mice

Fetal hematopoiesis occurs in multiple waves in distinct locations: the yolk sac (YS), aorta-gonad-mesonephros (AGM), fetal liver (FL), and the bone marrow. Broadly speaking, the first wave of nucleated erythroblasts emanating from the YS provides the developing embryo with oxygen once circulation has been established. The second wave generates erythro-myeloid progenitors (EMP) in the AGM and other major arteries through endothelial hematopoietic transition (EHT). Finally, the third wave produces definitive HSCs in the AGM that colonize the FL before migrating to the bone marrow just before birth.

Definitive, that is transplantable, HSCs emerge primarily in the AGM. This was first suggested by the observation of intravascular clusters of cells, resembling blood cells, in the aortic arch in several vertebrates. Here, HSCs capable of multilineage differentiation and self-renewal in transplantation emerge (Medvinsky et al., 1993; Muller et al., 1994; Medvinsky & Dzierzak, 1996). On the other hand, the YS does not generate definitive HSCs (Ganuza et al., 2018), despite producing cells that are capable of multilineage proliferation. HSCs do not migrate into the bone marrow until it has been established and the HSCs have matured in the FL. Transplantation into fetal sites increases repopulation efficiency (Arora et al., 2014), highlighting both the role of the microenvironment and that direct application of methods for adult hematopoiesis might yield biased results. The rate of HSC expansion in the FL is less than previously suggested from transplantation experiments (Ganuza et al., 2022), suggesting that it mainly promotes maturation of pre-hematopoietic stem cells.

The precise mechanism of HSC generation from the AGM was a source of much debate including distinguishing whether the HSC originated from endothelial cells or the underlying connective tissue. Expression profiling, imaging, and lineage tracing from endothelial cells using VE-cadherin-CreERT2 at embryonic day 9.5 (E9.5) confirmed HSC origin in embryonal endothelial cells. The underlying connective tissue could contribute but only through an endothelial intermediate (Zovein et al., 2008). However, *in vivo* live imaging of the transition from endothelium to HSC, so-called EHT in zebrafish (Bertrand et al., 2010; Kissa & Herbomel, 2010) and mouse embryonal AGM explants (Boisset et al., 2010) argued in favor of direct differentiation from endothelium to HSCs while not ruling out a small contribution of the underlying mesenchyme.

Complementary to the imaging studies, several studies using lineage tracing have attempted to clarify the origin and frequency of HSCs, yielding different results depending on Cre-ERT *cis*-regulatory elements and induction day. Using the Confetti mouse, a model with multiple fluorescent reporters that act as a barcode, it was estimated that hematopoiesis is sustained by around 500 clones of multilineage hematopoietic progenitors (Ganuza et al., 2017). Further genetic barcoding experiments with the *Sleeping Beauty* transposon system suggest that adult hematopoiesis is maintained by embryonal MPPs that persist throughout life (Patel et al., 2022). DARLIN barcoding also has shown a high degree of overlap of barcodes in different blood compartments when induced at E10, and that around 500 clones contribute to adult hematopoiesis (Li et al., 2023). Interestingly, this study found localized clonality in different hematopoietic sites when the barcodes were induced post-birth.

In summary, while there have been many conflicting studies regarding the spatiotemporal emergence of definitive HSCs, there is now agreement that most, if not all, clones contributing to adult hematopoiesis emerge at the latest at E9.5 in the AGM and then migrate to the FL for maturation before finally colonizing the bone marrow around birth. This is followed by a final burst of proliferation before the HSCs enter a quiescent state.

2.5 Human hematopoiesis

Mouse models are used to understand hematopoiesis in the hope that the findings will generalize to humans. Although not the focus of this thesis, this section summarizes the current understanding of hematopoiesis in humans.

The methods and ideas developed for studies in mice have been applied to study human hematopoiesis. Immunodeficient mice supportive of human hematopoiesis have allowed for xenotransplantation of human bone marrow cells (Doulatov et al., 2012). Together with clonal colony assays and stromal cell co-culture assays, this has allowed the identification of the human hematopoietic hierarchy. Like studies in mice, immunophenotypes have been defined for several HSPC compartments. Xenotransplantation studies are very dependent on the choice of mouse model, since transplantation across species leads to rejection of the transplanted cells. Furthermore, some hematopoietic cytokines and chemokines are cross-reactive whereas others are not. As a solution, humanized transgenic mice expressing human cytokines have been developed, allowing for increased engraftment.

Although it has proven difficult to assess HSC heterogeneity in other species beyond the mouse, several observations support functional HSC heterogeneity in the human and primate HSC compartment. Xenotransplantation of barcoded human cord-blood derived HSPCs showed differential lineage contribution (Cheung et al., 2013), further supported by non-random X-inactivation in different lineages (Busque et al., 1996). Clonal hematopoiesis, defined by expansion of cells sharing the same somatic variants at detectable level in peripheral blood without hematological malignancy (Jaiswal et al., 2014), is also an indication of stem cell heterogeneity but could be attributed to the fitness advantage of the expanded clone. Finally, mimicking the mouse studies, barcoded autologous transplants have implicated different contribution of clones to mature lineages (Kim et al., 2014; Koelle et al., 2017; Calabria et al., 2024) in rhesus macaques and in human gene therapy trials (Biasco et al., 2016; Scala et al., 2018; Six et al., 2020). So far, no evidence of platelet-biased or -restricted HSCs has been found in species other than the mouse.

Somatic variants, which constantly accumulate in each HSC throughout the life, provides a natural barcode system that has been used to infer key parameters about the human HSC population (Lee-Six et al., 2018). This approach could also be used for lineage tracing if the sensitivity becomes high enough. Barcodes that persist over time are likely to originate in the stem cell compartment since variants that occur in progenitors will be lost as the progenitor is exhausted and replaced by the division of an HSC from higher up in the hierarchy. Variants that are detected only within the Mk lineage and in HSCs would then be an indication of platelet-biased/restricted HSCs.

2.6 Local and systemic signals regulate blood homeostasis

Hematopoietic cell function and behavior are carefully regulated by both the internal machinery of the cell and the external environment. Cells respond to local signals emanating from the bone marrow microenvironment and systemic signals such as hypoxia, inflammation, and metabolic status. Regulatory signals may be manipulated to influence cell behavior. Several cytokines have been demonstrated to be vital for HSC maintenance, of which stem cell factor (SCF) and its receptor cKIT have received the most attention. The cKIT receptor, expressed in many tissues, is a tyrosine kinase that upon activation by the dimerized SCF ligand induces an anti-apoptotic signaling cascade. SCF exists as a soluble protein, which is thought to mediate long range and systemic signaling, and a membrane-bound form which is important for local signaling (Lennartsson

& Ronnstrand, 2012). Thrombopoietin (THPO) via its receptor MPL is a main driver of thrombopoiesis but is also vital for maintenance of HSCs (Qian et al., 2007). The receptor is expressed on the megakaryocyte lineage and on HSCs. The ligand for MPL, THPO, is mainly secreted by the liver upon binding of desialylated platelets, but is also expressed in the kidney, endothelium, and osteoblastic bone marrow. THPO stimulates platelet production via a direct effect on Mks and by driving HSCs to differentiate into the Mk lineage (Nakamura-Ishizu & Suda, 2020). MPL signaling is critical for HSC function, where loss of MPL leads to reduced HSC numbers in mice (Qian et al., 2007) and MPL variants are the cause for congenital amegakaryocytic thrombocytopenia in humans, which can only be treated by HSC transplantation (Ballmaier & Germeshausen, 2011). The stem cell niche concept aims to capture both spatial organization and function of the microenvironment surrounding stem cells (Scadden, 2006). Local and systemic signals are integrated in the niche to regulate stem cell renewal and differentiation. One study (Schofield, 1978) is often credited for the niche hypothesis but the idea of an instructive microenvironment is, as mentioned, a decade older (Curry & Trentin, 1967).

Early studies found that CFU-S activity in the femoral bone marrow was not distributed uniformly but rather increased toward the inner surface of cortical bone (Lord & Hendry, 1972; Lord et al., 1975). Marrow isolated from the endosteum, the inner lining of bone, was enriched for CFU-S (Gong, 1978). In consequence, it was proposed that HSCs are maintained by osteoblasts, either by paracrine signaling or direct contact (Taichman & Emerson, 1998). However, using SLAM markers to stain HSC-enriched bone marrow *in situ* revealed that a fraction of immunophenotypic HSCs rather localize with sinusoidal endothelium, suggesting an endothelial niche (Kiel et al., 2005). Further studies using conditional knockouts of chemokines in defined niche cell populations suggested important roles of Cxcr-12 abundant reticular (CAR) cells, endothelial cells, and perivascular cells (Sugiyama et al., 2006). Deletion of THPO from these bone marrow stromal populations did not alter HSC function but deletion from albumin expressing hepatocytes almost completely ablated the HSC population, in line with previous studies, suggesting that the THPO regulating HSCs is not produced by HSC niche cells (Decker et al., 2018). The Mk niche is relevant both for HSC maintenance and platelet development and has been investigated in parallel to the HSC niche. So far, Mks are the only mature hematopoietic cells that have been shown to contribute to the niche by localizing near HSCs and secreting Cxcl-4 (Platelet

factor 4, Pf4) (Bruns et al., 2014) as well as TGF-beta (Zhao et al., 2014), both of which promote HSC quiescence. It was first proposed that Mk maturation is accompanied by migration from endosteum to sinusoids (Malara et al., 2015), since that is the site for proplatelet shedding, and that THPO promotes expansion whereas chemokines such as Cxcl-12 drive migration (Avecilla et al., 2004). However, *in situ* imaging found little migration of mature Mks within the niche indicating that Mks are instead replenished by perisinusoidal progenitors (Stegner et al., 2017), in line with the studies on the HSC quiescence gradient within the niche. Importantly, the vascular network within the niche is very dense, minimizing the need for migration. The Mk/HSC niche signaling could potentially be of relevance to platelet bias and restriction.

Much effort has been made to understand how internal and external signaling can cause transitions between the states defined by the hematopoietic hierarchy, or in other terms how cells are directed to their fate during differentiation. The molecular basis for these cell state changes is thought to be gene regulatory circuits where transcription factors instruct expression of lineage-specific genes, thereby directing differentiation. This mechanism has been studied in detail for most mature cell types (Orkin & Zon, 2008). For example, the erythroid versus myeloid fate decision is guided by a co-inhibitory loop between the transcription factors Gata1 and PU.1, which support erythroid/megakaryocytic and myeloid fates respectively. Data accumulated over many studies has implicated the same transcription factors in several lineage choices, suggesting that their action is context-dependent. Consequently, integration of global expression profiling, lineage tracing, and transcription factor binding site analysis has begun to disentangle gene regulatory circuits (Gottgens, 2015).

The stochastic and instructive microenvironment models have been used for describing responses to cytokines as well (Robb, 2007). According to the stochastic model, in this context called the permissive model, in which fate choice is directed by internal transcriptional processes alone and cytokines provide a general survival and proliferation signal. In this view, the observed lineage distribution is the result of selection of already committed cells. In contrast, the microenvironment model has been adapted to the cytokine problem as the instructive model, which posits that cytokines have a direct influence on lineage choice. This issue is reminiscent of the age-old "nature or nurture" and "randomness or determinism" debates (Enver et al., 1998; Metcalf, 1998). A practical consequence of these models is the interpretation of absence of lineage

in determining lineage potential in culture experiments with cytokine supplements – it could be the case that the progenitors did indeed have ability to generate many lineages but some of them died off, leading to the false conclusion of a restricted progenitor. This once again underscores the importance of *in vivo* assays.

2.7 Megakaryocytes and platelets

Mks are the largest and rarest cells in the bone marrow. Each Mk generate thousands of platelets over the course of its lifespan and is vital for hemostasis (Machlus & Italiano, 2013). They develop from HSCs in a similar process to other progenitors. However, the MkP further matures through endomitosis, a halted cell cycle in which the genome is replicated but division never occurs. Mks next to sinusoidal vessels extend cytoplasmic protrusions with the help of cytoskeletal remodeling, called proplatelets, into the vessel lumen. The proplatelets are loaded with granule content and bud off, possibly by shear forces in the blood stream, where they split first into proplatelets and then into functional platelets in a barbell fashion (Machlus & Italiano, 2013). Apart from platelet generation, Mks have been implicated in regulation of HSCs and the immune response. Interestingly, gene expression profiling of Mk populations with different ploidies allows separation of Mks into different distinct populations defined by HSC maintenance, immune involvement, and platelet production, among other criteria (Choudry et al., 2021; Sun et al., 2021).

Platelets are small anucleate cells critical for the coagulation system (Gremmel et al., 2016). They persist in circulation for around 10 days before being cleared in the spleen and liver (van der Meijden & Heemskerk, 2019). Vascular damage exposes tissue factor that binds to and activates circulating platelets, which then form a thrombus and recruit coagulation factors that form the definitive clot through the coagulation cascade (Versteeg et al., 2013). During systemic inflammation, the platelet count in peripheral blood is an early variable to be altered together with other acute phase reactants, including C-reactive protein (Morrell et al., 2014). Although spontaneous bleeding is only associated with platelet levels below 30×10^9/L, reduced platelet counts can be detected by enhanced susceptibility to bruising. In patients with immune thrombocytopenic purpura (ITP), platelet clearance mediated by autoantibodies results in reduced platelet counts, which in severe cases can be treated by immunosuppression or splenectomy to reduce platelet destruction, or by MPL agonists (Cines & Blanchette, 2002). Because of their high turnover rate, transient platelet deficiency and subsequent bleeding are

common complications in the acute phase after exposure to chemotherapy and/or HSC transplantation (First et al., 1985; Kuzmina et al., 2012). Identification and activation of HSC and progenitor cells to stimulate a faster recovery of platelet counts is therefore an important clinical goal in transplantation, thrombocytopenia, and severe bleeding.

Although platelets, due to their small size and lack of nuclei, were initially believed to be a homogeneous and inert cell population, platelets initiate translation of Mk-derived unspliced RNA upon activation and undergo conformational changes when in circulation. The incorporation of fluorescent dyes, such as thiazole orange (TO), has allowed identification of nascent or so-called reticulated platelets containing higher levels of RNA and granules (Kienast & Schmitz, 1990; Ault et al., 1992), proposed to play an important role in cardiovascular disease (Martin et al., 2012) and sepsis (Middleton et al., 2019). Bulk RNA sequencing of reticulated platelets has revealed an upregulation of genes associated with hemostasis compared to platelets that have spent time in circulation (Bongiovanni et al., 2019), a plausible finding considering that platelet production increases in response to tissue injury and inflammation. Furthermore, platelet activation in thrombus formation is not uniform, as layers of activation levels radiate outward from the site of injury (Brass et al., 2016).

Assays for platelet functions have mostly been developed for clinical testing in bleeding disorders and anticoagulation monitoring. Physiological activation and aggregation of platelets can be simulated *in vitro* by treating platelet-rich plasma (PRP) with activating agents such as adenosine diphosphate, thrombin, and thrombin receptor activating peptide. Activation can be assessed by upregulation of surface markers such as CD62P, CD63 (Ramstrom et al., 2016) and the activated form of the fibrinogen receptor, granule fusion with the cell membrane, or by aggregation testing. Light transmission aggregometry is considered to be the gold standard and is based on the increased transmission of light through PRP as aggregates clear the solution (Born & Cross, 1963). In impedance aggregometry, platelets clump together on two electrodes inserted into the solution, which causes an impedance change between them (Cardinal & Flower, 1980). The two methods are used clinically and show comparable results both in normal samples (Seyfert et al., 2007) and with anticoagulation (Velik-Salchner et al., 2008). Finally, aggregation can be tested using flow cytometry (De Cuyper et al., 2013), where two aliquots of the same sample are stained with different fluorochromes and the mixed back together before activation.

3 Research aims

The overall aim of this thesis is to investigate properties of lineage-biased HSCs and their progeny, both during development and in adult life, with the following specific aims:

1. Previous studies have investigated the lineage bias of perinatal HSCs but have not tracked platelets and erythrocytes. Since most FL cells at E14.5 are *Vwf* positive, we explored the hypothesis that perinatal HSCs are a source of platelet bias and whether the platelet-biased HSCs were more abundant in the perinatal period. In addition, we aimed to molecularly characterize HSCs from FL in the perinatal period, to find evidence of transcriptional priming for lineage fates.

2. Building on previous work from our group, we sought to characterize the cellular hierarchies generated by *Vwf*$^{-}$ multilineage and *Vwf*$^{+}$ platelet restricted HSCs and reconcile this with the current tree model of hematopoiesis. The first aim was to determine the hierarchical relationship between *Vwf*$^{-}$ multilineage and *Vwf*$^{+}$ platelet restricted HSCs. The second aim was to map the molecular and immunophenotypic differentiation intermediates used by *Vwf*$^{-}$ multilineage and *Vwf*$^{+}$ platelet restricted HSCs to generate platelets. The third and final aim was to investigate the role of *Vwf*$^{-}$ multilineage and *Vwf*$^{+}$ platelet-restricted HSCs and their respective differentiation pathways during hematopoietic stress conditions.

3. Recent studies have indicated the existence of platelet subsets that have different roles in clot formation and immune function. However, the cellular and molecular basis for the observed platelet heterogeneity remains to be unraveled. So far, no method allows for molecular characterization of small numbers of platelets due to their small size and low RNA and protein content. Our aim was to develop a method that allows for sensitive and reproducible capture and analysis of the transcriptome of platelet subsets within the same individual.

4 Materials and methods

In this section the main methods used in the studies included in the thesis will be discussed with emphasis on potential pitfalls and interpretation of results.

4.1 Flow cytometry and fluorescence activated cell sorting (FACS)

FACS is used extensively in all studies in this thesis and this section will describe the method in general. Specific protocols and antibody lists are available in the individual manuscripts.

A flow cytometer operates by collecting cells in suspension into a single cell stream using hydrodynamic focusing, whereby the cell flow is enveloped by a saline sheath flow with different pressure, forcing the cells in line. The cells then pass through the flow cell, illuminated in sequence by lasers of different wavelengths. Emitted light from each laser interaction with fluorophore conjugated antibodies or fluorescent proteins is captured by a photomultiplier tube array and converted to a digital signal that is recorded by a computer.

Antibodies, both monoclonal and polyclonal, are used to label a target molecule with high specificity. Monoclonal antibodies are generated through hybridomas that is a B-cell from a host immunized with a target antigen fused to a malignant plasma cell.

A fluorophore is a molecule that can absorb and then emit light. Fluorophores exist in nature, for example in bioluminescent algae or bacteria. The first generation of fluorophores used for applications in research were proteins isolated from such naturally occurring compounds. Synthetic fluorophores have expanded the number of available colors in a FACS experiment. Every fluorophore is characterized by several parameters. Each has a unique excitation and emission spectrum. The difference between the maximum emission and excitation wavelengths is called the Stokes shift. Fluorophores differ in how bright they are. Brightness is the product of the extinction coefficient and quantum yield of the fluorescent molecule. The extinction coefficient reflects how many photons can be absorbed per unit fluorescent molecule and the quantum yield specifies what fraction of absorbed photons are re-emitted. These properties are measured empirically by the manufacturer. For example, phycoerythrin, an organic fluorophore found in algae, is considered one of the brightest fluorophores because it has both a high extinction coefficient and quantum yield. Brilliant Violet 421 is a newer synthetic polymer-based fluorophore that has lower quantum yield

but a larger extinction coefficient resulting in similar brightness. Bright fluorophores should be used for dim populations or where clear separation of populations is needed.

The stain index (SI), defined as the mean fluorescence intensity (MFI) difference between a positive population and a negative population divided by twice the standard deviation of the negative population is an experiment-derived measure of the resolution of a target/antibody complex. It can be used to determine optimal antibody concentration for a particular experiment. Too low an antibody concentration obscures the positive population, but too high a concentration results in high background signal from the negative population.

Multiple laser lines with different wavelengths intersect the cell stream in sequence. The detectors either capture a small band of the emission spectrum (conventional flow cytometry) or the whole spectrum (spectral flow cytometry). The spectral overlap between dyes limits combinations in the same experiment. Since more spectral information is retained in spectral flow cytometry, highly overlapping dyes can still be deconvoluted.

The combinations of excitation and emission spectra can be increased by using tandem dyes, in which two fluorophores are covalently bound, combining the excitation spectrum of one with the emission spectrum of the other, creating a wider Stokes shift. These dyes are less stable than single molecule dyes and degradation can lead to unexpected spectral properties as the acceptor fluorophore also starts to emit light.

Beads or cells stained with a single antibody are used to assess how much signal from each fluorophore is captured in each detector. This information is then used to derive a compensation matrix, in which all pairwise spillover values are tabulated. Importantly, spillover is a property of the spectrum of a fluorochrome and is independent of experimental parameters such as antibody concentration, laser power, or photomultiplier tube (PMT) voltage.

The spillover reduces sensitivity in the channel in which it occurs by increasing data spread in the negative population. Compensation subtracts the average spillover from the signal in the secondary detector. However, each event might deviate from this average value so that more or fewer photons in the secondary detector are counted. For physical reasons, photon counting follows a Poisson distribution, for which the variance increases with the average. This is why

correctly compensated data with high spillover can produce a characteristic funnel or trumpet shape in the fluorescence intensity distribution. In practice this spillover spread (SS) can lead to erroneous conclusions about the abundance of populations in a sample (Nguyen et al., 2013). Most stem cell markers show low continuous expression and do not generate a distinct FACS population, making their expression difficult to separate from background noise. Therefore, spillover into the stem cell marker channels should be avoided.

Fluorescence minus one (FMO) controls have all fluorophores except one and can be used to quantify SS and to set negative gates (Roederer, 2001). However, if SS is high, positive events might be called as negative. FMOs do not account for unspecific antibody binding, which would also result in a higher signal in the negative population. Unspecific binding can instead be determined using an isotype control, an antibody of the same species and class conjugated with the same fluorophore but without any specific affinity to the target. Isotype controls are not identical with the experimental antibody, for example their number of fluorophore molecules can differ. An alternative measure of unspecific antibody binding is to use a negative cell population that is as similar as possible to the population of interest. This is exemplified by cytotoxic and helper T-cells, which are mutually exclusive for the CD8 and CD4 surface markers. Therefore, CD8 T-cells can be used to set the gates for CD4 T-cells and vice versa.

As evidenced by the preceding sections, there are many parameters in a flow cytometry experiment, which pose two important problems. How can optimal experimental parameters and conditions be determined, and how can reproducibility be ensured? For the first point, our lab titrates all new antibody lots on the target cells to determine optimal concentration and all new panels are validated by comparison to previous results using the relevant tissue sample and discussed before real experiments. We strive to use bright fluorophores for dim markers and where separation is crucial, for example to separate the congenic CD45.1 and CD45.2 marking donor and recipient cells in transplantation experiments. For the second point, BD cytometer setup and tracking (CST), a standardized bead solution containing beads of three brightness levels, is run on all experimental days to account for instrument variation.

The limit of detection for targeted cells in flow cytometry experiments depends the number of cells analyzed. For our lineage determination experiments in peripheral blood, we aimed to record at least 50,000 events in each blood lineage.

Therefore, the term "negative" must be interpreted as a value below the limit of detection.

Flow cytometry can be used to sort cells based on the signal from the fluorophores. A piezo-electric oscillator generates a standing wave that breaks up the stream into droplets. The cell concentration in the stream is optimized so that there is a low probability of multiple cells in one droplet. Selected droplets receive an electrical charge so that they can be deflected by a static electrical field into a microtube or plate. To ensure purity when sorting, different stringency conditions can be set using sort masks, maximizing either yield or purity. In our lab, we have a custom sort mask that provides maximum yield while sacrificing some purity. Purity is controlled by reanalyzing a sorted sample ensuring that the captured cells fall within the sort gate. Droplet deposition accuracy in the plate is important for single-cell experiments, both transplantations and sequencing. Deposition accuracy is validated by sorting fluorescent beads into 96-well plates which then can be visualized under a microscope. For 384-well plates, we sort cells suspended in buffer supplemented with horseradish peroxidase into wells containing 3,3',5,5'-tetramethylbenzidine. Successful depositions initiate a reaction that turns the well blue. We aim for more than 95% successful cell depositions. Inclusion of cell type controls and wells with no cells that serve as no template controls also enables retrospective evaluation of sorting alignment and purity. The cell type control is sorted from a predefined population such as granulocyte progenitors that have a distinct transcriptome. No template controls estimate background mRNA and should show very few, if any, reads. The well position of these controls can be used to ensure proper plate orientation during analysis.

4.2 Mouse models used in the thesis

Lab mice are vital tools for hematopoiesis studies and started being used in research more than 100 years ago to answer questions about tumor development (Paigen, 2003a). Many inbred strains have been created to remove unwanted background genetic variation.

Transplantation is considered the assay that most faithfully captures true HSC behavior and has been used since the 1950s. Transplantation is used to colloquially describe several different assays that have evolved together with the field (see section 2.1). In competitive repopulation, a sample of test cells and a fixed number of competitor bone marrow cells are transplanted together into a

myeloablated host by tail vein injection **(Figure 2)**. To engraft, the cells need to home to the bone marrow and the efficiency of this process is reasonably high. The chimerism of the test, competitor, and host cells is measured using congenic markers. Chimerism ratios between test and competitor cells give a relative measurement of stem cell potency that allows for comparisons across experiments as long as the competitor bone marrow cells are of the same type and quantity. Both the conditioning and the source of the competitor cells shape the relative advantage of the competitor and host cells. Mouse strains are differentially sensitive to radiation and doses need to be titrated. Higher doses ablate more of the host blood system, providing more "space" for the transplanted cells. Unfractionated bone marrow as competitor cells provides stronger competition but bone marrow from cKit-mutant mice or serially transplanted bone marrow, which both have impaired long term regenerative capacity, can also be used. Doses between 1×10^5 and 4×10^5 competitor cells are commonly used. Studies before prospective HSC isolation measured HSC reconstitution in repopulating units (RUs), where 1 RU was defined as the reconstitution capacity of 10^5 competitor cells. Assuming an HSC frequency 0.002%, it is estimated that around 5–30 HSCs are present in the competitive bone marrow. This amount is also enough for radioprotection.

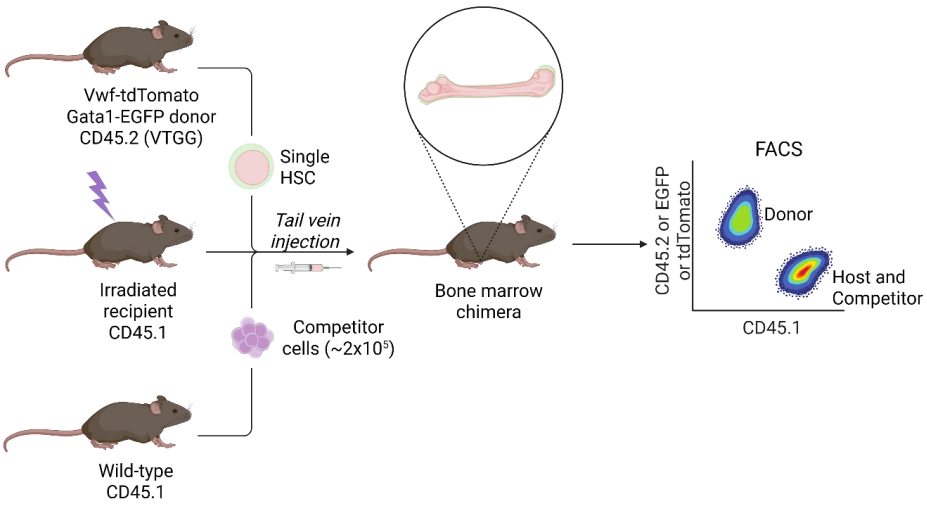

Figure 2 Single-cell transplantation. Single immunophenotypic HSC-expressing fluorescent reporters from a mouse on a CD45.2 background are isolated using FACS and transplanted together with support cells from a mouse on a CD45.1 background into a lethally irradiated CD45.1 mouse. This generates CD45.1/2 bone marrow chimeras. Peripheral blood contribution of the transplanted cells is tracked using flow cytometry. Donor, support and host cells can be separated using reporter proteins or antibodies against CD45.1/2. Created with BioRender.com

Improved cell isolation strategies resulted in higher HSC purity and could eventually be used for prospective HSC isolation. This enabled limiting dilution experiments in which Poisson statistics allow inference about the reconstitution capacity of single cells (Szilvassy et al., 1990). Single-cell deposition using FACS further refines this approach by ensuring that only one cell is transplanted. Although the experimental setup is the same as for competitive repopulation, the role of the competitor cells changes. Single HSCs have the capacity to regenerate the whole hematopoietic system at high chimerism levels, but this process takes time. The competitor bone marrow serves as a bridge until the donor cells have reached a high enough level to sustain the hematopoietic system. In a single-cell transplant, there are just a few HSCs that compete; the transplanted cell, the HSCs present in the competitors, and host HSCs that survived the conditioning.

In our experimental system, the transplanted cell is followed as a single clone, whereas the competitor and host HSCs are treated as a background population. In fact, they collectively act as a single hematopoietic system. Variation in chimerism levels of the transplanted HSCs could either be a result of intrinsic variation in the transplanted HSC or sampling variation in the composition of the competitor cells. As a consequence, it has been suggested that chimerism levels are unsuitable measures of HSC capacity and heterogeneity in the single-cell transplant assay (Ema et al., 2016). Instead, a binary measurement of reconstitution should be used, for which we set a threshold to 0.1% in our studies. Our strictest definition of platelet restriction is based on the absence (<0.01%) of all lineages except the platelet lineage. A looser definition of platelet restriction allows detection of myeloid and erythroid cells as long as the platelet chimerism is 50× higher. On the other hand, definitions of bias are based on ratios of chimerism exceeding 3× in each lineage when compared to the other lineages. Ratios <2× between all lineages were used in **Study II** to identify the multilineage pattern. One consideration is that when high levels of lineage restriction (>66% platelet chimerism) are detected in the transplanted cell, the competitor cell population exhibits lineage-biased behavior since its chimerism is one minus the chimerism of the single cell.

The goal of the CD45.1/2 congenic system is for strains to differ at only one locus. However, the congenic interval contains multiple genes, and it has been shown that CD45.2 HSCs outcompetes CD45.1 HSCs in competitive repopulation (Waterstrat et al., 2010). To address this issue the CD45.1 (STEM) strain was developed through a single targeted exon mutation making it functionally

equivalent to the CD45.2 strain (Mercier et al., 2016). For single cell transplantation this is not an issue since the readout is not dependent on relative measurement to the competitor cells. In our studies we use CD45.2 cells as test cells and CD45.1 as support cells since that would give the highest probability of single cell reconstitution.

To show self-renewal ability of HSCs, we performed secondary transplantations in Studies I and II. This was performed in two different ways, either by direct unfractionated transplantation of the primary bone marrow cells or by HSC isolation using FACS. There are competitor cells in the first approach, but they vary in number due to the chimerism in the primary transplant, making comparisons between samples difficult. Additionally, progenitors are transplanted that can mask HSCs in the secondary transplant (Rundberg Nilsson et al., 2015). The second approach requires co-transplantation with support cells and is in essence a competitive repopulation experiment.

The molecular biology revolution in the 1980s led to the introduction of genetically modified mice which have been a vital tool for stem cell research (Paigen, 2003b). Stable integration of foreign genetic material allows for *in vivo* detection of gene expression and conditional recombination. Several approaches exist to generate genetically modified mice. All mouse models in this thesis are either bacterial artificial chromosome (BAC) transgenes or targeted knock-ins. **Table 1** is a summary of mouse models used in the studies for this thesis.

BACs are large genetic constructs that theoretically exist as single copies and replicate independently of the host bacterial chromosome. Their main advantage is that large genetic fragments can be incorporated into them, which allows for cloning of whole genes and the corresponding *cis*-regulatory elements. For *Mus musculus*, the whole genome exists as a BAC library as a part of the Mouse Genome Project. The *Vwf*-tdTomato, *Gata1*-EGFP and *Flt3*-Cre strains used in this thesis are all transgenes of this type.

BAC transgenes integrate randomly into the host genome with varying copy numbers giving rise to variability in transgene expression among founder lines. Host genes can potentially be disrupted by the integration and the transgene expression can be influenced by the surrounding genetic context, for example adjacent *cis*-regulatory elements and chromatin accessibility. In addition, transgenes sometimes contain additional genes that can impact phenotype or bias molecular analysis. For example, the *Vwf* BAC contains the *Cd9* gene as well.

In our studies, comparisons are made within the same strain so this should not affect the results since all cells carry the BAC at the same copy numbers.

The Cre-*Lox* system (McLellan et al., 2017) was isolated in the P1 bacteriophage. In this system, the 343 amino acid-long protein Cre recognizes 34 base pair long locus of recombination (*loxP*) sites and either excises or inverts the sequence between them. The expression of the *Cre* transgene can be driven by cell-type-specific promoters allowing for targeted gene knockouts.

Many reporter strains for Cre activity have been created and are used to determine tissue specificity of different promoter-Cre constructs. They can also be used for lineage tracing to determine if a cell at some point has expressed the gene of interest. The Rosa26 locus is ubiquitously expressed and susceptible to gene targeting without disrupting normal gene function. Several Rosa26 reporter lines have been developed using targeted knock-ins. In our studies, we have used the Rosa26-Ai9 or Rosa26-Ai14 lines (Madisen et al., 2010), referred to as Rosa26-tdTomato.

While an essential tool for developmental biology, the Cre-*Lox* system has important limitations. Like all transgenes, the integration sites and copy numbers of the *Cre* transgenes are often unknown and can lead to both disruption of normal function and variable expression depending on the chromatin state and surrounding genetic elements. Moreover, many promoters believed to be tissue-specific are transiently expressed during development, leading to mosaic knockout/reporter patterns, and consequently which can result in incorrect interpretation of gene function or lineage relationships. If recombination occurs in the germline, the recombined allele will be inherited in all cells of the offspring (Song & Palmiter, 2018). Therefore, a careful breeding strategy and monitoring of the colony by appropriate genotyping are necessary.

Since Cre recombination requires the formation of a DNA loop, its efficacy is dependent on the distance between the *loxP* sites (Ringrose et al., 1999). This explains why the same *Cre* transgenes can show variable efficiency when crossed to different reporter lines. Moreover, some studies have observed varied levels of reporter expression in supposedly recombined cells from Rosa26-dTomato mice. The STOP cassette, an SV40 sequence repeated three times, has been proposed to serve as a non-canonical recombination site for Cre (Bapst et al., 2020).

Table 1 Mouse models used in the thesis (BAC – Bacterial artificial chromosome, CP – Cyclophosphamide, 5-FU – 5-Fluorouracil)

Mouse model	**Characteristics of the model**	**Labeled cell populations and type of experiments**	**Used in**
***Gata1-E*GFP** Full name: B6.129P2-Tg(Gata1-EGFP)Cn	BAC transgene with EGFP under control of *Gata1 cis*-regulatory elements	Labeling of *Gata1* expressing cells including platelets and erythrocytes. Used for bone marrow donor cells for single HSC transplantations in combination with labeling of *Vwf* (**Study I**, **Study II**) or both *Vwf* and *Flt3* (**Study II**) expression. Used for steady state fate mapping and anti CD42b treatment experiments in combination with labeling of *Vwf* and *Flt3* expression (**Study II**).	**Study I**, in combination with *Vwf*-tdTomato **Study II**, primarily in combination with *Vwf*-tdTomato or *Vwf*-EGFP
***Vwf-E*GFP** Full name: B6.Tg(VWF2-eGFP)#2	BAC transgene with EGFP under control of *Vwf cis*-regulatory elements	Labeling of *Vwf* expressing cells including platelets. Used for bone marrow donor cells for single HSC transplantations in combination with labeling of *Gata1* or both *Gata1* and *Flt3* expression. Used for steady state fate mapping and anti CD42b treatment experiments in combination with labeling of *Gata1* and *Flt3* expression.	**Study II**, primarily in combination with *Gata1*-EGFP
***Vwf*-tdTomato** Full name: Not registered	BAC transgene with tdTomato under control of *Vwf cis*-regulatory elements	Labeling of *Vwf* expressing cells including platelets. Used for bone marrow donor cells for single HSC transplantations in combination with labeling of *Gata1* (**Study I**, **Study II**) or both *Gata1* and *Flt3* (**Study II**) expression. Used for single-cell RNA sequencing of liver pnHSCs (**Study I**).	**Study I**, in combination with *Gata1*-EGFP **Study II**, in combination with *Gata1*-EGFP or *Gata1*-EGFP and *Flt3*-Cre
Rosa26-tdTomato (Ai9 or Ai14) Full name: Ai9:B6.Cg-*Gt(ROSA)26Sor*$^{tm9(CAG-tdTomato)Hze}$/J Ai14:B6.Cg-*Gt(ROSA)26Sor*$^{tm14(CAG-tdTomato)Hze}$/J	Mice carrying the fluorescent reporter gene tdTomato, driven by the ubiquitous Rosa26 promoter. A stop codon flanked by two *loxP* sites is before the reporter. When bred to mice expressing Cre, listed below, the fluorescent reporter is expressed in the Cre-expressing cells as well as their progeny	Depending on the Cre model.	**Study II** (Ai 9 or Ai14), in combination with *Flt3*-Cre or *Vav*-Cre

***Flt3*-Cre** Full name: B6(129X1)-Tg(Flt3cre)#CCb	Cre expression is under the control of the *Flt3* promoter.	When crossed with Rosa26-tdTomato report strain, labeling of HSPC populations expressing *Flt3* and their progeny. Used for bone marrow donor cells for single HSC transplantations in combination with labeling of *Vwf* and *Gata1* expression. Used for fate-mapping, CP and 5-FU treatment experiments, anti CD42b treatment experiments, and analysis of reticulated platelets with TO.	**Study II**, primarily in combination with Rosa26-tdTomato
***Vav*-Cre** Full name: B6.Cg-Tg(VAV1cre)1Graf/MdfJ	Cre expression is under the control of the *Vav* promoter	When crossed with Rosa26-tdTomato reporter strain, labeling of all hematopoietic cells in the bone marrow and blood. Used for fate-mapping of platelets and CP treatment experiments.	**Study II**, in combination with Rosa26-tdTomato

4.3 Bulk and Single-cell RNA sequencing

The central dogma of molecular biology states that genes encoded in DNA are transcribed to messenger RNA (mRNA), and subsequently translated into protein which exerts the biological function of the gene. The set of all transcripts, the transcriptome, can be thought of as a representation of the cell's overall state.

Over the years, many methods have been developed to assay the transcriptome of a cell — hybridization techniques, quantitative PCR, microarrays, and sequencing, to name a few. The discovery of RNA-dependent DNA polymerases, reverse transcriptases (RTs), in tumor viruses enabled capture of mRNA *in vitro*. The general approach is exemplified by the SMART (**S**witching **M**echanism **a**t 5′ End of **R**NA **T**emplate) family of protocols (Zhu et al., 2001). In this process, mRNA is converted to complementary DNA (cDNA) by the RT enzyme. The reaction can be primed either with random k-mer primers or an oligo-dT primer, which preferentially hybridizes to the poly-A tail of processed mRNA. The RT enzyme adds three cytosine bases at the 5′ end of the newly synthesized cDNA strand. A template-switching oligo (TSO) can then hybridize and act as a new template giving both strands a common sequence. Primer binding sites in the oligo-dT and the TSO allow for preamplification of cDNA libraries before sequencing library construction.

In this thesis, all sequencing was done using Illumina next generation sequencing. This platform is based on sequencing by synthesis. In brief, cDNA is fragmented either physically or by enzymes. Next, additional primer sites and binding sequences are ligated to the fragments using the sequences from the previous step. After loading the library on the flow cell, the binding sequences adhere the fragments to the flow cell and each fragment is subsequently amplified by PCR to form clusters of identical molecules. During the following sequencing cycles fluorescently labeled bases are added in a stepwise manner to the cluster. Each letter (ATCG) has its own color. The reaction cleaves off the fluorophore, and the color is recorded by a camera before the flow cell is washed, and the reaction begins anew. The generated images are then analyzed to determine the full DNA sequence of the fragment that generated the cluster.

Some important considerations for sequencing are the type of flow cell and detection optics. Patterned flow cells, like the ones used in Illumina NovaSeq 6000, increase cluster density but are more prone to index hopping, that is when combinatorial indexing is incorrectly mapped to the wrong library. Patterned flow cells also preferentially bind shorter fragments, often comprised of primer dimers. These issues can be mitigated by using unique dual indexing and careful library cleanup and size selection before sequencing. The latest generation of Illumina sequencers use 2-color detection chemistry and as a consequence, absence of signal cannot be distinguished from G nucleotides. The advantages of this system are reduced costs and increased speed of sequencing. Our approach is to first sequence libraries on the more accessible 2-color NextSeq 550 system, and if the sample passes quality control, it is sequenced on the NovaSeq 6000 at the National Genomics Infrastructure (NGI). We did not assess index hopping, but it usually appears as a plate-associated batch effect which was not observed.

By miniaturizing and optimizing library preparation, the transcriptome of a single cell can be sequenced (single-cell RNA sequencing). This enables the deconvolution of bulk expression profiles into their constituent cells. Importantly, this circumvents the problem of prior assignment of cell populations based on flow cytometry or other methods. Cell types can instead be inferred from unsupervised clustering based on gene expression patterns. The first single cell was sequenced more than 15 years ago (Tang et al., 2009) and in the following years, the field has seen a tremendous expansion in both throughput and sensitivity (Svensson et al., 2018; Svensson et al., 2020). A single mammalian cell contains approximately 30 pg of RNA, which is composed of 80% ribosomal RNA

and less than 10% mRNA. Some landmark inventions have improved sensitivity, reduced cost, and made protocols easier, such as unique molecular identifiers (UMIs) to correct PCR bias (Kivioja et al., 2011), combinatorial indexing (Islam et al., 2011), and molecular crowding to increase sensitivity (Bagnoli et al., 2018). The initial protocols were plate-based, requiring single-cell deposition by FACS. The introduction of droplet-based methods, in which the cells and reagents are encapsulated together with barcoded beads in a water in oil emulsion, greatly increased throughput but introduced other problems such as unwanted doublets and reduced sensitivity.

The hematopoietic system is particularly well-suited for single-cell RNA sequencing since its cells are inherently non-adherent. Viable single-cell suspension samples of hematopoietic cells can be collected with little effort from peripheral blood and bone marrow, allowing transcriptome analysis of all cells within the hematopoietic hierarchy from many species. Several studies have investigated the transcriptomic landscape of human and mouse hematopoiesis, both in homeostasis and under stress (Nestorowa et al., 2016; Dahlin et al., 2018; Dong et al., 2020; Rodriguez-Fraticelli et al., 2020; Weinreb et al., 2020). In steady-state, molecular phenotype, defined by clustering based on similarity in expression profiles, has a consistent overlap with already established immunophenotypes (Wilson et al., 2015). However after a challenge, such as administration of cytotoxic agents or transplantation, this correspondence is lost for some cell types, as there is a discrepancy between the surface marker expression and the molecular profile. For example, the surface marker SCA1, as well as the expression of critical cytokine receptors commonly applied to the isolation of mouse HSCs by flow cytometry, are transiently gained or lost upon treatment with certain cytotoxic drugs.

For the studies in this thesis, we have used the 2, 3 and the 3xpress variants of the Smart-seq protocols (Picelli et al., 2014; Hagemann-Jensen et al., 2020; Hagemann-Jensen et al., 2022), together with index sorting that allows integration with immunophenotypes.

After sequencing, reads are mapped to a reference mouse genome. There is seldom perfect match between a read sequence and the genome due to errors in library preparation and sequencing, and from natural genetic variation, so any assignment is done based on a scoring algorithm. Since mRNA is spliced, the aligner must be aware of splice junctions as the introns are still present in the

reference genome. The zUMIs pipeline (Parekh et al., 2018) was designed together with the Smart-seq3 protocol and performs all steps from alignment of sequences up to and including counting mRNA molecules. Under the hood, it uses the splice-aware aligner STAR and featureCounts. The key novel components are barcode demultiplexing and UMI deduplication. These important steps have tunable parameters that influence the results, mostly relating to the level of error that is tolerated within the same barcode or UMI. In the studies performed in this thesis, the default parameters were used.

In practice, there is immense analytical flexibility in analyzing RNA sequencing data, stemming from the choice of algorithms, their implementation and user defined parameters. For single-cell data, there are numerous pipelines to align, quantify, and normalize reads and the work in this thesis uses one of many possible flows of analysis.

Analysis of RNA sequencing data is often performed in a scripting language such as R or Python using software packages. The most common ones are Seurat in R and Scanpy in Python. For the work in this thesis, the Bioconductor suite of single-cell RNA sequencing tools was used because of their well-written documentation. Although the front facing interfaces are different, many core algorithms are shared between packages. This is the case for principal component analysis (PCA) and uniform manifold approximation and projection (UMAP), and the most common differential expression analysis methods. However, packages do differ in data integration algorithms, normalization and advanced differential expression analysis.

As the number of sequenced reads is not equal across cells, potentially reflecting difference in RNA content in the original cell and/or variation in loading concentration for the sequencing since molecules from all cells in the pool compete for clustering on the flow cell, counts are typically normalized to adjust for sequencing depth and facilitate comparison across cells. As a result, the values that are used for the downstream analysis therefore typically reflect fractions of the total library size rather than absolute counts. The normalization procedure can have a considerable influence on the subsequent steps in the pipeline (Vieth et al., 2019).

The procedures to identify differences in expression between groups of interest are collectively known as differential expression (DE) analysis. Many models of varying complexity have been developed for this purpose and many of them

originate from the microarray era, persisted through bulk RNA sequencing and are now being applied in the single-cell setting. These models range from the regular two group t-test or a non-parametric alternative to generalized mixed models (Soneson & Robinson, 2018). Of particular importance is the selection of appropriate distribution for the count data, as this will significantly influence which genes are identified as differentially expressed. The transcriptional process can be modeled by a stochastic process and its stationary distribution is used for differential expression analysis (Peccoud & Ycart, 1995). Many such distributions have been derived; depending on different modeling assumptions; they often incorporate a binomial or Poisson component due to the similarity between transcription and random sampling from an urn. This is also recapitulated by empirical count distributions generated from experimental data.

Absence of reads from a gene can be due to the gene being "turned off" or because of low sensitivity. This is sometimes referred to as zero-inflation or dropout. By separately modeling the expression and sampling of transcripts, many commonly used distributions can be recovered (Sarkar & Stephens, 2021). In bulk RNA sequencing experiments, little dropout is typically observed since the transcriptional process is sampled across hundreds of cells. The commonly used packages, EdgeR and DESeq2, for bulk RNA sequencing data analysis, both assume a negative binomial count distribution, also called as a gamma-Poisson model. This is an overdispersed Poisson distribution, which is justifiable on theoretical grounds and also because greater variance than expected by a Poisson distribution is observed in experiments. Inference of the overdispersion parameter is a key step in DE analysis, as it influences the level of difference considered significant between two populations. DESeq2 and EdgeR were designed to control false positive rates in studies with very few replicates. Genes with low expression can give exceedingly high estimates of log-fold-change and both packages have approaches to shrink both variance estimates and log-fold-changes. In general, their performance is similar in controlling false positive rates. Both packages were used in Study 3.

With linear models, it is possible to include additional covariates beyond the experimental condition, such as sequencing batch and day of sample collection, which can be unwanted sources of variation.

Another subtlety of single-cell data is that each cell is not an independent sample. There is a high correlation of expression profiles in cells originating from the same

biological sample which violates the independence assumption of many models. Additionally, since thousands of cells can be sequenced, the subsequent statistical tests have artificially high power resulting in unreasonably low p-values when considering how many biological replicates were included in in the test, a phenomenon known as pseudoreplication (Zimmerman et al., 2021). One potential solution to address this in the analysis is to add up counts from individual cells and perform the testing on pseudobulk samples (Squair et al., 2021) using methods for bulk RNA sequencing analysis. Another alternative is to use mixed models to incorporate the hierarchical structure of the experiment. With this approach, the biological sample is assigned a random effect, and the condition of interest, for example cluster or treatment, is the fixed effect. The mixed model is computationally more expensive, and both methods have been shown to produce comparable results.

For single-cell data analysis in this thesis, we used a combination of a Fisher's exact test and a non-parametric Wilcoxon rank-sum test for DE analysis (Rodriguez-Meira et al., 2019). Informally, the Fisher's exact test accounts for differences in zero-inflation whereas the Wilcoxon test assesses differences in median expression. This approach does not account for pseudoreplication or batch effects but has the advantage of being fast to compute and easy to understand.

Implicit in most differential expression procedures is a cell type assumption which highlights a conceptual difference between a type and a state. Clustering serves as a way of discretizing a continuous population of cells into separate groups. These clusters are then assumed to reflect actual cell types. In studies focusing on hematopoiesis, there is already a large body of evidence about which cell types exist in normal bone marrow and this categorization has been in good agreement across data modalities, but new clusters must be validated by functional assays (Villani et al., 2017). RNA sequencing requires lysis of cells, which poses a problem to establish the link between molecular phenotype and cellular function (Jassinskaja et al., 2023). Single genes can be validated using fluorescent reporter mice or index sort data can be used to find immunophenotypes that are conserved within molecular clusters (Wilson et al., 2015).

Underlying both clustering and dimensionality reduction algorithms is a nearest neighbor (NN) graph constructed based on cell similarity. Here, similarity means the distance between transcriptional profiles. Cells of the same type are

supposed to have similar transcriptomes and should therefore belong to the same cluster and be next to each other within a low-dimensional embedding. Defining similarity between cells measured across many variables is challenging because distances become very similar when there are more variables than samples, a phenomenon known as the curse of dimensionality (Altman & Krzywinski, 2018). PCA is a dimensionality reduction method that finds composite variables that maximize variation within a dataset, through diagonalization of the covariance matrix. It is used to find a lower dimensional representation of the data that can be used for nearest neighbor graph construction.

Both t-distributed stochastic neighbor embedding (t-SNE) (van der Maaten & Hinton, 2008) and UMAP (McInnes et al., 2018) are non-linear, graph-based dimensionality reduction techniques that strive to preserve the structure of high dimensional data when projected in a 2D plane. Implementations of the published algorithms vary across programming languages and there is a component of randomness in the final plot layout which can be mitigated by specifying a fixed random seed before running the algorithm.

Furthermore, both clustering and dimensionality reduction algorithms have tunable hyperparameters that influence the number of detectable clusters. One extreme end of the spectrum would be that each cell forms an individual cluster, while the case that all cells belong to a single cluster is the other (Kiselev et al., 2019).

The goal of experiments is to explain variation. This can be done by repeated measurements of an experimental system and decomposing the variability into technical and biological variability. Technical variability is the inherent variation that arises from randomness, such as pipetting error or incubation times and can be estimated using technical replicates; that is, the same samples analyzed several times with the same techniques. The remaining variation should, in theory, be explained by biological differences between experimental conditions, such as treatment or genotype. If there is systematic experimental variability such as a batch effect, this cannot be separated from a true biological effect. Ideally all samples should be collected, sorted, and sequenced in the same session to avoid batch effects. However, in practice this is difficult to achieve. We include a control from the same frozen bone marrow sample in each plate for all our single-cell RNA sequencing experiments. This control should yield similar results across experimental runs, and the cells from a control sample should always be located

within the same cluster. Batch effects can also be detected using the distribution of batches on the NN-graph.

To correct batch effects, there are two main approaches; model- and graph-based. The model-based approaches assume a linear batch effect on read count levels, which essentially translates to a differential expression problem and can be addressed by linear regression with batch as a categorical covariate. Graph-based approaches instead strive to assign the same cell types to the same clusters, assuming that the same cell types are present in each batch. In Study 2, the first approach was used to remove the cell cycle effect, and the second approach was used to remove the actual sequencing batch effect.

In contrast to clustering, which segregates high density populations in the molecular landscape, trajectory inference (TI) seeks to identify patterns of development by modeling how the cell state might evolve over time (Saelens et al., 2019). From a single timepoint sample of one data modality, it is impossible to infer a dynamical process but by using a parsimony assumption, where cells are assumed to transition to as similar states as possible, a tree of transitions can be constructed (Saelens et al., 2019). Another option is to model differentiation as a diffusion process on the NN-graph, where the transcriptional distances between cells determine the rate of diffusion between states. This process assumes no direction of the diffusion and a starting cell must be specified. This limitation can be overcome by calculating a linear approximation of the state change in mRNA by measuring spliced and unspliced RNA in combination with a model for transcription, splicing, and degradation, a technique known as RNA velocity (La Manno et al., 2018; Bergen et al., 2020; Bergen et al., 2021). The resulting vector field can be used in conjunction with the earlier similarity-based methods for unsupervised determination of the direction of state change. The cell order inferred by TI methods is referred to as pseudotime and RNA velocity provides a means to convert it to real time. These methods have been successful in systems where cell types are highly diverse but have proven difficult to apply in hematopoiesis. For example, the computed direction of erythropoiesis is inverted possibly due to changing transcriptional kinetics which violates the underlying steady state assumption of the velocity models (Bergen et al., 2021). By including pseudotime as a continuous covariate to a generalized additive negative-binomial model, it is possible to perform differentiation dependent DE analysis (Van den Berge et al., 2020). This was used in **Study II** to find genes associated with platelet differentiation pathways linked to subsets of stem cells.

An important difference in comparison with lineage tracing is that the TI methods provide information on how states will change but can only provide probabilities for the changes of a single cell whereas lineage tracing provides information on the actual relationship between cells and clones.

Interpretation of RNA sequencing data can be aided by summarizing gene expression values into molecular signatures. This is also referred to as finding expression "modules" or "programs". Conceptually, this is the inverse problem of differential expression analysis. Instead of finding genes that separate two subsets of a dataset, the objective is to find samples that match a predefined set of genes. First all expression values are ranked in each sample and then enrichment of the genes of interest at the top of the rank lists is tested. This approach is implemented in gene set enrichment analysis (GSEA) (Subramanian et al., 2005), which has been the prototypical enrichment software for 20 years. AUCell (Van de Sande et al., 2020) relies on a similar approach but instead it computes a ranking list for each cell so that a distribution of rankings in a data set is created. Outlier cells from this distribution are called as being enriched.

Curated gene lists categorized by both cell type and function have been amassed in the omics era. The gene ontology (GO) knowledgebase aims to structure gene symbols in a hierarchy of function and location. Enrichment of a particular set of gene symbols associated with a specific function within a cluster can then be interpreted as the "function" of that cluster. Single-cell atlases provide cell-type annotations that can be used in subsequent studies. While a useful tool for describing data, gene sets do not provide actual functional evidence. Furthermore, there is likely to be reporting bias in the gene set databases. Genes that are frequently studied, for example interferon-related genes, tend to accumulate more annotations and are therefore more implicated in enrichment analysis leading to ever more annotation enrichment in the database (Haynes et al., 2018).

As the previous sections have outlined, there is large analytical flexibility in analyzing RNA sequencing data, especially at the single-cell level and when there are multiple data modalities. However, from a conceptual standpoint a distinct pattern emerges. Normalization and differential expression analysis are based on statistical models of gene expression and sequencing. This is followed by a linear dimensionality reduction in the form of a matrix factorization. These two steps are variable/gene centric in the sense that the cells are not the target of the algorithm. Next, the perspective shifts to the cells as cell similarity is computed using a

measure of similarity. This is summarized in a neighbor graph, which is the main object of analysis. The flexibility of analysis arises from the choice of models and algorithms for normalization, matrix factorization, metric and nearest neighbor detection.

4.4 Induction of hematopoietic stress conditions

Several approaches were used in the studies to induce hematopoietic stress.

Transplantation induces proliferation of the transplanted cells and a damage response in the bone marrow microenvironment of the host caused by the conditioning treatment. It is known to alter the molecular phenotype of the transplanted cells (Dong et al., 2020).

Cyclophosphamide (CP) is converted to nitrogen mustard, an alkylating agent, upon administration (Emadi et al., 2009). It is used for treatment of solid and hematological tumors and as conditioning before bone marrow transplantation and after to suppress graft versus host disease. Importantly, the active form is metabolized by aldehyde dehydrogenase and expression of this enzyme determines in part the efficacy of the drug. Progenitor cells have in general a low expression of this enzyme and are therefore susceptible to the treatment since they are slow to metabolize it whereas HSCs have higher expression (Levi et al., 2009).

5-fluorouracil (5-FU), a thymidylate synthase inhibitor cause cells to undergo thymidine starvation and subsequent apoptosis (Longley et al., 2003). 5-FU is also directly incorporated into DNA and RNA which disrupts their normal function. The drug is commonly used to treat solid tumors but not hematological malignancies.

ITP is a bleeding disorder where autoantibodies against platelet surface proteins lead to platelet destruction in the reticuloendothelial system of the spleen. This has been modeled in mice and single doses of CD42 antibody injection are quickly recovered whereas repeated injections lead to long and profound thrombocytopenia (Nieswandt et al., 2000). We used single doses of anti CD42 antibodies to induce thrombocytopenia and evaluated the response by measuring thrombocyte number in an automated cell counter or with flow cytometry.

4.5 Functional assessment of hematopoietic stem and progenitor cells *in vitro*

In addition to the transplantation assays, there is a plethora of *in vitro* assays to determine the properties of HSPCs (Bock, 1997). Short-term cultures supplemented with defined hematopoietic cytokines, with or without supporting stromal cells, are used to investigate lineage-specific potential of cells (Miller et al., 2008), and were used in **Study II**. Long-term cultures can also be used to identify candidate cells with long term self-renewal potential (Ploemacher et al., 1989), including culture conditions which allows the expansion of LT-HSCs with ability to engraft non-conditioned recipients (Wilkinson et al., 2019). Although self-renewal potential is best evaluated *in vivo*, the culture assays can be combined with flow cytometry, imaging, or high throughput molecular profiling methods to link phenotype to function.

For our studies, *in vitro* cultures are used to confirm lineage potentials of sorted populations. The cultures are evaluated by microscopy of May-Giemsa-Grünwald stained cytospins.

4.6 Optimization of platelet collection and handling

Platelets are naturally very prone to activation, which profoundly alters their state. Therefore, uncontrolled activation has the potential to severely confound results of functional experiments. Additionally, platelets are sticky, especially when activated, increasing the probability of contamination by adjacent cells.

To identify ways of minimizing spontaneous activation, a literature review was conducted to determine feasible collection conditions. There are several established protocols and locations for peripheral blood collection in mice with cardiac puncture, tail vein, inferior vena cava (IVC), tibial vein, and retro orbital plexus being the most common. In our lab, we have primarily performed tail vein bleeding. Cardiac puncture, due to the heart muscle's high collagen levels, carries a considerable risk of inducing platelet activation when damaged. IVC sampling requires anesthesia and is also a terminal procedure, precluding any re-sampling from the same mouse. Sampling from the retro orbital plexus is technically challenging; although used internationally, it is associated with ethical concerns by the Swedish regulatory authorities for animal welfare. With these considerations,

we decided to optimize our already established tail vein sampling procedure to allow for collection of unactivated samples for RNA sequencing.

To avoid clotting during blood collection, an anticoagulant must be utilized. Several anticoagulants are commonly used. We have experience with ethylenediaminetetraacetic acid (EDTA), heparin and sodium citrate. EDTA binds calcium ions that are vital for platelet activation and requires recalcification before functional experiments. Heparin is a polysaccharide that activates the coagulation inhibitor antithrombin. It interferes with thrombin dependent platelet activation. Sodium citrate works by both binding ions and acidification of the blood. We have used both heparin and EDTA with success. Our protocol incorporates a wash step, so any remaining anticoagulant is removed.

4.7 General experimental design considerations and statistics

Careful experimental design improves the validity of results and reduces the required number of replicates. What constitutes an experimental replicate varies depending on the experiment and misassignment of replicates leads to pseudo replication, which is present in as many as 50% of all published research (Lazic et al., 2018). In fact, this is the same issue described above for single-cell data analysis. One clarifying distinction for animal experiments is the division into biological units (BU), experimental units (EU), and observational units (OU) (Lazic et al., 2018). The BU is the entity about which the hypothesis is tested, for example a HSPC populations. The EU is the unit assigned to experimental conditions and the sample size is equal to the number of experimental units, for example one mouse. Measurements are made on the OU. Importantly, increasing the number of OUs does not increase sample size unless it is the same as the EU. For example, measuring the MPP population in two femurs from the same mouse is still one sample.

Single-cell transplantations are susceptible to pseudoreplication since all transplanted cells (OU) in an experiment come from the same donor (EU). We ensured post analysis that the study contained replicates from several experiments involving different donors.

Several experiments in this thesis such as the lineage reconstitution or cyclophosphamide treatment follow the same individual over time, a repeated measures design. These experiments are analyzed using a hierarchical model where the individual mouse is assigned a random effect to account for mouse-to-

mouse variation, similar to the use of hierarchical models to analyze single cell transcriptomes isolated from the same mouse.

Inferences in the biological sciences are usually made based on statistical tests derived from a model of the data and the arbitrary significance threshold of 0.05. This threshold is used throughout the thesis. One clear interpretation of p-values comes from the American statistical association:

> *"Specifically, the distance between the data and the model prediction is measured using a test statistic (such as a t-statistic or a Chi squared statistic). The P value is then the probability that the chosen test statistic would have been at least as large as its observed value if every model assumption were correct, including the test hypothesis"* (Greenland et al., 2016)

In general, for experiments in the biological sciences, it is often taken for granted or at least assumed that all model assumptions are correct. Null hypothesis testing and p-values have also been criticized because they often are misrepresented, do not correspond to strength of evidence and can be manipulated (Stefan & Schonbrodt, 2023). When many statistical tests are performed on the same data some p-values will turn out significant by change. If p-values are used to quantify the evidence in a controlled experiment, adjustment for multiple testing is justified. On the other hand, if p-values are used to rank differentially expressed genes, the adjustment will not change the ordering and can be omitted.

4.8 Ethical considerations

The studies in this thesis were performed under the following ethical permits at Karolinska Institutet (KI): N29/15 with amendments N69/16 and N195/16, 17978-2018 with amendments 18539-2021, and 19970-2022, and 17034-2023. At the University of Oxford: PPL 3003103, PPL P2FF90EE8, and PPL PP5432093.

Animal models are an integral part of research. The mouse has been the main model for hematopoiesis research since the field was established in the 1950s. As discussed in the preceding sections, *ex vivo* models such as culture systems or organoid models fail to capture the *in vivo* behavior of hematopoiesis, likely explained by unaccounted microenvironmental effects and the culture induced

alterations in stem cell state. Furthermore, findings in mice model are more likely generalizable to the human hematopoietic system than those from *ex vivo* models.

Single-cell transplantations require a large number of mice and HSCs cannot be prospectively isolated, but refined FACS panels have allowed for 1,000-fold enrichment of HSCs. In our lab, we continuously try to Replace, Reduce, and Refine (3R) our animal experiments following the literature for improved HSC isolation strategies, for example by including the CD201 surface marker. Reconstitution is evaluated early and mice without the expected pattern are not kept unnecessarily. Serial sampling from the same mouse reduces numbers as well.

Optimized experimental procedures such as marker panels, sex matched donors and recipients, titration of conditioning radiation, close monitoring, and antibiotic treatment have resulted in 20-30% of transplanted mice showing reconstitution from a single cell, with almost all survive the procedure.

All experimental protocols and data are archived and made public whenever possible. Before publication, data and figures are independently checked by another member of the lab to ensure data integrity and reproducibility.

5 Results and Discussion

This section describes the main results of the papers included in this thesis. For more detail the reader is referred to the original manuscripts found after the comprehensive summary.

5.1 Platelet and myeloid lineage biases of transplanted single perinatal mouse hematopoietic stem cells

Definitive HSCs emerge in the AGM and expand in the FL before migrating to the bone marrow prior to birth. The fetal HSCs give rise to adult HSCs which show variability in the lineage contribution upon transplantation. Previous work has identified multiple patterns of HSC lineage bias as well as the rare platelet restricted pattern in bone marrow from adult mice (Sanjuan-Pla et al., 2013; Carrelha et al., 2018). The regulation and time of emergence of lineage bias and restriction are not known. On embryonal day 14.5 (E14.5), virtually all HSCs in the fetal liver express the *Vwf* reporter, which in adult bone marrow enriches for platelet biased HSCs. This could suggest that FL is a source of lineage biased HSCs. In this study, we sought to investigate lineage restriction and bias by transplantation of perinatal HSCs (pnHSC) on ED19.5/postnatal day 0 (PD0).

In brief we followed the approach of previous studies from our group: Lethally irradiated male and female B6.SJL-Ptprca Pepcb/BoyJ and B6.SJL-Ptprca Pepcb/BoyCrl (CD45.1) mice were transplanted with single perinatal $Lin^{-}Sca1^{+}cKit^{+}$ (LSK) $CD150^{+}CD48^{-}$ cells isolated from the FL or bone marrow of CD45.2 *Vwf*-tdTomato/*Gata*1-EGFP (VTGG) mice **(Figure 2)**. Indexed marker expression was collected at the time of sort to allow for retrospective grouping of *Vwf* reporter expression of the individually transplanted HSCs. Lineage contribution of the transplanted single HSC was followed for up to 26 weeks using flow cytometry. Using the definitions of mature populations and lineage bias outlined in the methods section, 22% of single-cell transplanted mice had >0.1% contribution in at least one lineage at the end of the experiment (25-26 weeks). In contrast to similar experiments in adult mice where less than 50% of reconstituting HSCs contributed to all lineages (Dykstra et al., 2007; Yamamoto et al., 2013; Carrelha et al., 2018; Carrelha et al., 2024), more than 90% of reconstituting pnHSCs contributed to all lineages **(Study 1, Figure 1a)**.

Surprisingly, only a single HSC isolated from perinatal bone marrow and none from FL exhibited platelet bias in contrast to the comparatively frequent P-bias and

restriction found in adult HSCs. Secondary transplantations of unfractionated bone marrow from single-cell transplanted mice confirmed primary reconstitution pattern and self-renewal at >17 weeks **(Study 1, Figure 1d)**. Similar to bone marrow from adult mice, the *Vwf*-tdTomato reporter was expressed by 59% of FL pnHSCs and 47% in bone marrow pnHSCs, a clear reduction from the level at E14.5. Analysis of index sort data showed no difference in overall reconstitution potential associated with *Vwf* reporter expression, but P and PEM bias was enriched in the *Vwf*$^{+}$ bone marrow pnHSC fraction **(Study 1, Figure 1k)**.

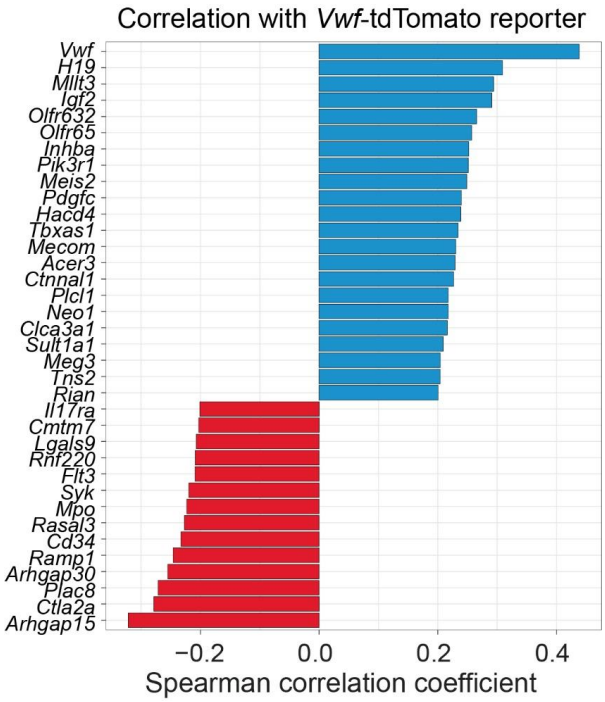

Figure 3 Spearman correlation analysis between genes and the *Vwf-tdTomato reporter in fetal liver pnHSCs. The figure shows genes that came out significant in the analysis. Blue indicates positive correlation with the Vwf-tdTomato reporter, red indicates negative correlation with the Vwf-tdTomato reporter.*

Functional output in transplantation indicated no P-restriction but we wanted to investigate if we could identify early evidence for priming of lineage bias. To this end, we performed single-cell RNA sequencing of *Vwf*$^{+}$ and *Vwf*$^{-}$ FL pnHSCs, finding no striking separation of the two subsets after dimensionality reduction and clustering **(Study 1, Figures S5, S6a, b)**. Further analysis identified only 32 differentially expressed genes between *Vwf*$^{+}$ and *Vwf*$^{-}$ pnHSCs. *Vwf* expression pattern in HSCs is continuous, making a correlation analysis more suitable than regular DE since this removes the need for a cutoff on *Vwf* expression level.

Spearman correlation further confirmed the association between *Vwf*-reporter expression and several genes previously found to be associated with HSC function and bias **(Figure 3; Study 1, Figure 1l)**. *Ctnnal1* labels almost all HSCs (Acar et al., 2015) and *Neo1* is a marker for myeloid-biased HSCs (Gulati et al., 2019). Additionally, standard DE using DESeq2 comparing the highest quantile of *Vwf*-reporter positive to *Vwf*$^{-}$ cells resulted in 162 DEGs. Additionally, among the highly expressed genes in *Vwf*-high cells were the maternally imprinted long non-coding RNAs *H19, Meg3*, and *Rian* which have been shown to be essential for establishment and maintenance of HSCs. Relevant to **Study II,** *Flt3* was one of the top genes associated with *Vwf*$^{-}$ HSCs **(Study 1, Figure 1i, m)**. This fraction is also similar to the embryonal MPP (eMPP) identified by CRISPR lineage tracing (Patel et al., 2022). The eMPPs are lymphoid biased and have been suggested to sustain hematopoiesis independent of HSCs months after birth and lymphopoiesis throughout life. In contrast, *Vwf*-high cells were enriched for Mk-biased stem cell signatures **(Study 1, Figure 1o)**.

Like most work from our lab, the experiments in this study were performed in a transplantation setting. Lineage patterns remained stable after six months, when hematopoiesis is assumed to have returned to normal, and after secondary transplantations, increasing evidence that lineage bias is intrinsically regulated. Contrary to our expectations, little lineage bias was found in pnHSCs. Nevertheless, some lineage priming at the molecular level was observed in the *Vwf*$^{+}$/high fraction of the FL pnHSCs. One plausible interpretation of this data is that lineage bias is a continuous process that is initiated early in life and expands as the individual ages with lineage being lost in sequence.

5.2 Alternative platelet differentiation pathways initiated by distinct stem cells

The classical model of hematopoiesis posits a linear branching hierarchy with the HSCs residing at the apex. Several studies have suggested revisions to this hierarchy by identifying novel progenitor subsets. Implicit in most such models is a unique differentiation pathway for each mature cell type. Previous studies on the heterogeneity of HSCs identified a subset of HSCs that are transcriptionally primed to the Mk fate (Mansson et al., 2007). In particular, they exhibit high expression of the Mk marker von Willebrand factor. Transplantation of such HSCs revealed functional lineage bias toward the platelet fate (Sanjuan-Pla et al., 2013; Carrelha et al., 2018). In parallel several studies have suggested by molecular and

immunophenotypic evidence that HSCs can differentiate directly into Mks, by passing established progenitor intermediates (Morcos et al., 2022).

As a continuation of previous studies from the lab, this study aimed to further characterize the hierarchical relationships between lineage biased HSCs. Current data would suggest that lineage bias is a gradual trait in which an HSCs loses lineage potential in a stepwise manner (Carrelha et al., 2018). *Vwf*-positive platelet-restricted HSCs (P-HSCs), and *Vwf*-negative multilineage HSCs (M-HSCs) represent the two extremes of the spectrum and we hypothesized that any functional or molecular difference would be most pronounced between them.

First, we performed transplantation of single LSK Gata1-EGFP^{-}CD34^{-}CD150^{+}CD48^{-} *Vwf*-tdTomato^{-} or *Vwf*-tdTomato^{+} HSCs into lethally irradiated recipient mice and observed their reconstitution patterns recapitulating known reconstitution patterns **(Study II, Figure 1a)**. Analysis of the bone marrow of the primary recipients showed that the *Vwf*$^{-}$ M-HSCs reconstituted all progenitor compartments whereas from the *Vwf*$^{+}$ P-HSCs only the LT-HSC and MkP compartments were reliably and robustly reconstituted whereas the ST-HSC, MPP3/4 and MPP2 compartments were missing **(Study II, Figure 1b)**.

To determine the hierarchical relationship between the two HSC types, we performed secondary transplantation of *Vwf* positive and negative HSCs isolated from the primary recipients of a single HSC **(Study II, Figure 2a)**. No conversion of *Vwf* expression status or reconstitution pattern was observed demonstrating that *Vwf*$^{-}$ M-HSCs and *Vwf*$^{+}$ P-HSCs do not generate each other, which implies that they either share a common ancestor cell or develop independently **(Study II, Figure 2c-e)**.

Since we observed the absence of most established immunophenotypic progenitor cell types in mice reconstituted by P-HSCs **(Study II, Figure 1a),** we hypothesized the existence of an alternative differentiation path from HSCs to committed Mk progenitors. To better resolve the progenitor hierarchies generated by *Vwf*$^{-}$ M-HSCs and *Vwf*$^{+}$ P-HSCs, we performed single-cell RNA sequencing of their respective HSPC progeny after transplantation. From *Vwf*$^{-}$ M-HSCs, an enrichment sort strategy was used to ensure that all cell types were captured whereas the more general Kit-positive population was sorted from *Vwf*$^{+}$ P-HSCs due to the absence of established progenitor populations.

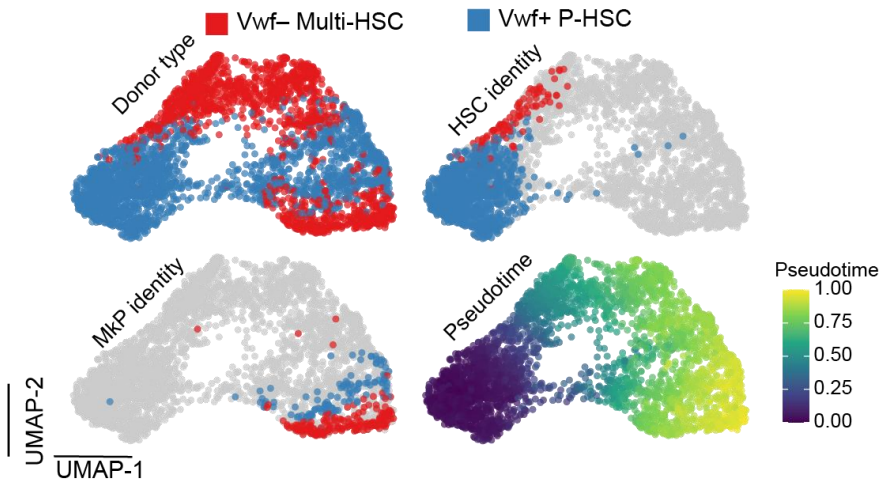

Figure 4 UMAP of HSPCs generated by *Vwf*- Multi-HSCs and *Vwf*+ P-HSCs after removal of committed progenitors. HSC and MkP identities indicate cells above the respective AUCell cutoff. The color scale indicates position on diffusion pseudotime.

After library preparation, sequencing, and standard bioinformatic processing, 2,290 HSPCs generated from P-HSCs and 2,478 HSPCs from M-HSCs were kept for further analysis. PCA followed by UMAP showed two non-overlapping populations of progenitor cells that connected the HSC cluster with the MkP cluster **(Figure 4; Study II, Figure 3b)**. Multi- and P-HSCs were mutually exclusive in their contribution to these populations. After ordering cells along their hypothetical differentiation trajectory using diffusion pseudotime **(Study II, Figure 3f)**, we performed differential expression analysis along the trajectory using TradeSeq. This analysis revealed 217 genes that drove the separation between the two pathways. Pearson correlation of transcriptional profiles along pseudotime showed a marked divergence at the intermediate progenitor stage **(Figure 5 left**

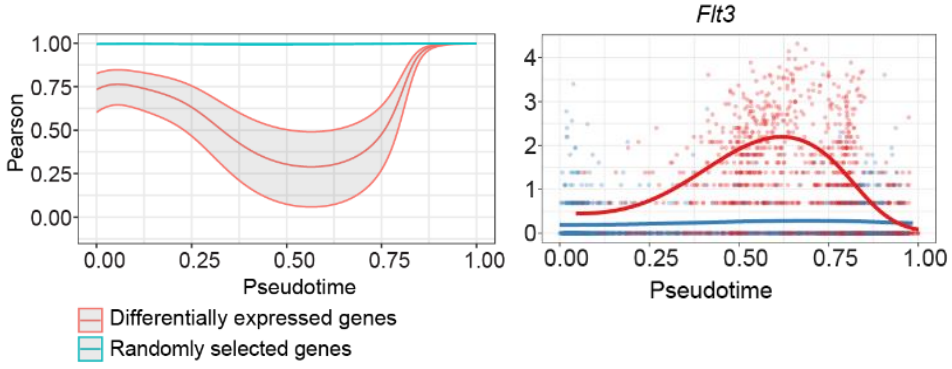

Figure 5 Left: Pearson correlation of differentially expressed genes between HSPCs generated by *Vwf*- M-HSCs and *Vwf*+ P-HSCs along pseudotime. Right: Flt3 expression along pseudotime. Red indicate cells from a *Vwf*- M-HSCs and blue cells from a *Vwf*+ P-HSCs

panel; Study II, Figure 3k). Notably, *Vwf* was highly expressed throughout the P-HSC trajectory, whereas *Vwf* only rose at the end of the M-HSC trajectory as it approached the MkP cluster. On the contrary, *Flt3* expression was completely absent from the P-HSC trajectory, while showing a distinct peak in the M-HSC trajectory **(Figure 5 right panel; Study II, Figure 3l)**. *Cd48* expression followed a similar pattern **(Study II, Figure 3l)**.

Flt3 together with *Cd48* has previously been implicated as a marker for a unique platelet differentiation pathway (Morcos et al., 2022), and we could observe the same pattern using flow cytometry **(Study II, Figure 3m, 5c)**. To functionally validate the novel pathway, we performed a new cohort of single-cell transplantations, this time using *Flt3*-Cre/Rosa26-tdTomato donor HSCs, and reconstitution patterns were determined as in the previous experiments. Platelets and all other blood cell lineages generated by M-HSCs were all recombined and expressed tdTomato, showing that their progenitors at some point had expressed *Flt3*. On the other hand, platelets generated by P-HSCs showed no tdTomato signal at all **(Figure 6; Study II, Figure 7b)**. Interestingly, ten percent of platelets in *Flt3*-Cre/Rosa26-tdTomato$^+$ mice are not recombined before treatment suggesting a low contribution of the direct pathway in steady state **(Study II, Figure 7a)**.

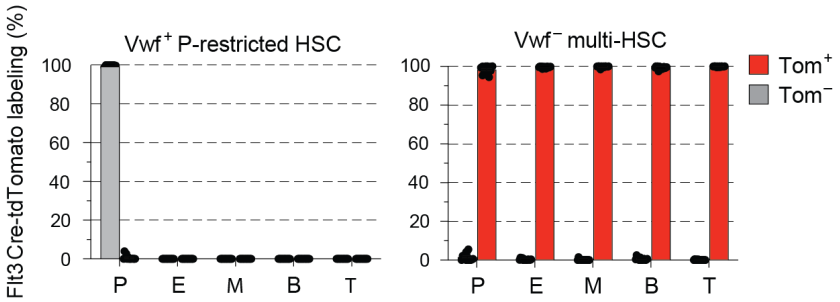

Figure 6 Recombination frequency in peripheral blood lineages generated by single transplanted *Vwf+ P-HSCs (left) and Vwf- M-HSCs (right) isolated from Flt3-Cre mice. (Tom, tdTomato expression.)*

Taken together with the molecular data this proves the existence of two differentiation pathways for platelets.

Finally, we sought to investigate the physiological role, if any, of the *Flt3* negative pathway. Injection with CD42b antibody to induce acute thrombocytopenia, led to a marked decrease in platelet numbers **(Study II, Figure 8e)**, in line with published data. However, no change in recombination rates was observed **(Study II, Figure 8f)**.

We hypothesized that antibody-depleted platelets were quickly replenished by progenitors that were already present at the time of antibody administration. We therefore injected *Flt3*-Cre/Rosa26-tdTomato mice with single doses of CP and 5-FU to induce an acute depletion of progenitor cells. A dramatic reduction in MkPs in response to CP treatment was confirmed **(Study II, Extended Figure 8b)**. This led to a decrease in recombination rate in platelets suggesting that the direct pathway is recruited in response to acute depletion of progenitor populations **(Study II, Figure 8ad)**. One possible confounder in the CP experiment is that the drug is metabolized by aldehyde dehydrogenases including *Aldh1a1*, which was one of the upregulated genes in P-HSCs. It is therefore possible that the direct pathway is more resistant to the treatment and what is observed is not an increase in the usage of the pathway but rather survival of its progenitors.

TO labels nucleotides and granules in platelets. Nascent platelets have a higher content of both and TO staining can therefore be used to investigate platelet replenishment kinetics. TO staining of *Flt3*-Cre/Rosa26-tdTomato platelets after treatment with CP showed a significant increase in the TO^+ fraction within tdTomato- platelets when compared to $tdTomato^+$ platelets, suggesting that the direct pathway replenishes platelets faster **(Study II, Figure 8c)**.

In parallel to our studies on the *Flt3* independent pathway, the Forsberg lab at the University of California, Santa Cruz performed similar experiments. Their results are in overall agreement with ours but there are some points of difference that would benefit from further discussion (Boyer et al., 2011; Boyer et al., 2012; Poscablo et al., 2024). Our study shows that the direct pathway exists in adult mice and initiates from HSCs. The Forsberg lab first argued that it only emerges in aged mice (>2 years), but then, with refined experiments, showed that it exists in adults as well. However, they did not link the pathway to a particular HSC subset.

We and the Forsberg lab use the same *Flt3*-Cre transgene (Benz et al., 2008), but our lab has used either Rosa26-EYFP (Buza-Vidas et al., 2011), or Rosa26-tdTomato, whereas the Forsberg lab uses the Rosa26-mT/mG switch mouse. In the original publication (Boyer et al., 2011), they found both high and low recombining mice and that recombination frequency increased during backcrossing. Recombination frequency in mature cell population was compared to the MPP compartment to remove the variability in baseline recombination. Since all compartments, apart from B- and T-cells, showed a similar recombination frequency as the MPPs, the authors concluded that all

hematopoietic cells are generated through a *Flt3*-positive step. However, their definition of MPP included FLT3 overlooking the fact that some MPP populations lack *Flt3* expression whilst still contributing to mature lineages. Importantly, our studies link the direct pathway to a distinct stem cell subset through the single-cell transplantation assay.

Later, the Forsberg lab crossed *Flt3*-Cre with a floxed EGFP construct driven by the pan-hematopoietic *Vav* promoter to increase sensitivity in *Flt3* detection (Perez-Cunningham et al., 2016). Conceptually, this model is the inverse of the *Flt3*-Cre/Rosa26-tdTomato model used in our study. Mature lineages were only analyzed in the bone marrow and in contrast no reporter signal was found in platelets indicating that all of them had gone through an *Flt3*-positive stage.

Since all *Flt3*-Cre mice come from the same founder, two possible explanations for differences in recombination rates are genetic drift of the *Flt3*-Cre transgene or different efficiencies in recombination of the reporter allele. In the original publication and in our lab, *Flt3*-Cre activity is only observed in males, indicating that the transgene is integrated in the Y-chromosome. In most mammals, the Y chromosome is small and exists in a heterochromatic state, but in the mouse, it is substantially larger (Hughes & Page, 2015). Like previous reports, we have also observed variable recombination in the same colony using the same reporter suggesting the contribution of genetic drift as well. However, our observation that all cells generated from an M-HSC in transplantation have recombined provides strong evidence that the mice used in this thesis faithfully capture *Flt3* expression.

Inter-*Lox* distance influences recombination efficiency (Ringrose et al., 1999). The Rosa26-tdTomato inter-*Lox* sequence is less than 1 kb, whereas the distance of Rosa26-Yfp and Rosa26-mt/mg is more than twice as long (Van Hove et al., 2020; Faust et al., 2023). This implies that, all things being equal, the Ai9/14 system would be more sensitive in detecting *Flt3*-expression, especially in early progenitor populations. In conclusion, the more sensitive reporter used in our studies, together with our transplantation data, enables us to conclude that the *Flt3*-negative pathway is present earlier in life than claimed by the Forsberg group.

The physiological role of FLT3 signaling was not investigated in our or the Forsberg study; instead, expression of it is used as a marker for the two platelet pathways. The FLT3 receptor was initially discovered by partial cloning of the granulocyte colony-stimulating factor (G-CSF) receptor or from fetal liver cells and its ligand was isolated from culture medium using a soluble form of the receptor (Lyman &

Jacobsen, 1998). Knockout of the receptor (Mackarehtschian et al., 1995) and the ligand (Hilary J. McKenna et al., 2000) lead to variable reductions in immune cell frequencies but left the myeloid compartments relatively unchanged. These studies did not track platelets and erythrocytes. Part of the variable response, at least in the dendritic cell compartment, could be explained by increased sensitivity to other cytokines (Durai et al., 2018). Although some HSCs express *Flt3* transcripts, surface detection of the protein correlates with exit from the HSC compartment and the FLT3 signaling axis is not indispensable for maintenance of HSCs (Sitnicka et al., 2002).

We and others argue that the *Flt3*-independent pathway is recruited in response to increased demand for platelets such as infection, chemotherapy, or injury (Morcos et al., 2022). The TO experiment in this study would suggest that this is the case. We also followed reconstitution kinetics of single transplanted HSCs (Carrelha et al, unpublished). After normalization to final reconstitution levels, the data was compatible with the idea that P-HSCs have faster reconstitution kinetics. A parallel study investigated chromatin accessibility features and their association with platelet bias and found that it is driven by epigenetic priming (Meng et al., 2023) and that the differentiation of a P-HSC into platelets is faster than that of an M-HSC.

Cd48 followed the same expression pattern as *Flt3* and could also be used to immunophenotypically separate MkPs generated by the two pathways. It has been suggested that a *Cd48* negative pathway generates transcriptionally distinct Mk subsets by using a combination of Cre and Dre fate-mapping (Li et al., 2024).

Since publishing our study, new data has emerged that complements it. Single-cell RNA sequencing of HSPCs from *Flt3*-Cre/Rosa26-tdTomato mice shows a distinct cluster of unrecombined cells bridging the HSC and Mk clusters thus recapitulating, on the transcriptional level, the direct differentiation pathway in a steady state setting (Singh et al., 2025). Another approach using mathematical modeling of single-cell transplantation data shows that a shortcut platelet differentiation pathway is consistent with observed reconstitution kinetics (Iwanami et al., 2025). Finally, one of the top genes associated with the *Flt3* negative pathway was *Clusterin (Clu)*. HSCs expressing *Clu* expand with age and are the main contributors to age associated platelet and myeloid bias (Koide et al., 2025).

5.3 The platelet transcriptome is stable across age and is not responsive to spontaneous CD62 externalization

Platelets are small anucleate cell fragments that are shed from Mks in the bone marrow. They are key components of hemostasis by homing to injury sites and initiating the wound healing process. Additionally, they have roles in inflammation and defense against foreign pathogens.

Cardiovascular disease is one of the leading causes of mortality in the world. The established model for ischemic damage in the heart and brain is that inflamed vessel walls calcify over time. Calcified vessels cause turbulent blood flow that can activate platelets through increased shear forces. A calcified plaque can also rupture, exposing sub endothelial matrix components that are potent platelet activators. Arterial thrombosis risk increases with age but the precise contribution of vascular dysfunction, soluble coagulation factors and intrinsic platelet function is difficult to determine (Wilkerson & Sane, 2002).

As with most cell types, recent years have seen an emergence of studies investigating platelet heterogeneity but the precise mechanism underlying this heterogeneity is not known. High-throughput molecular biology methods such as RNA sequencing and mass spectrometry-based proteomics have been successful in delineating cell and disease types. While both methods are well developed for platelets, they require large amounts of input material due to the small size of and low RNA and protein content of platelets. Furthermore, they have not been adapted for within sample measurements such as fluorescent lineage tracing where two cell populations are present withing the same individual.

In **Study II**, we confirmed the presence of a novel platelet differentiation pathway that bypasses several multipotent progenitor stages. This shortcut pathway expands in aged mice and platelets generated by it have been suggested to be responsible for the age-associated increase in thrombosis risk (Poscablo et al., 2024). Incomplete convergence in the transcriptional profile of the MkPs that sustain the two different platelet populations raises the possibility that platelet heterogeneity can be further explored by transcriptional profiling **(Study II, Figure 4a)**.

To address these questions, we developed a protocol for RNA sequencing from small amounts of FACS sorted mouse platelets. After establishing a collection

procedure using tail vein blood sampling, which allowed for repeated collection of peripheral blood without overwhelming spontaneous activation, we isolated platelets from young adult (8-10 weeks) and old (2 years) mice and subjected them to RNA sequencing using a modified single-cell Smart-seq2 protocol (**Study III, Figure 1A,B)**. While 500 platelets allowed for consistent RNA amplification and sequencing, high variability in detected genes was observed, indicated by low sample-to-sample correlation (**Study III, Figure 1C**). We attributed this to the low RNA content of platelets.

To overcome the low gene detection, we increased the number of sorted platelets to 300,000. Each sorted cell adds a small amount of sheath fluid to the final volume. By sorting into buffer supplemented with bovine serum albumin and using normal binding tubes the platelets could be pelleted and resuspended in lysis buffer and be subjected to library preparation as usual.

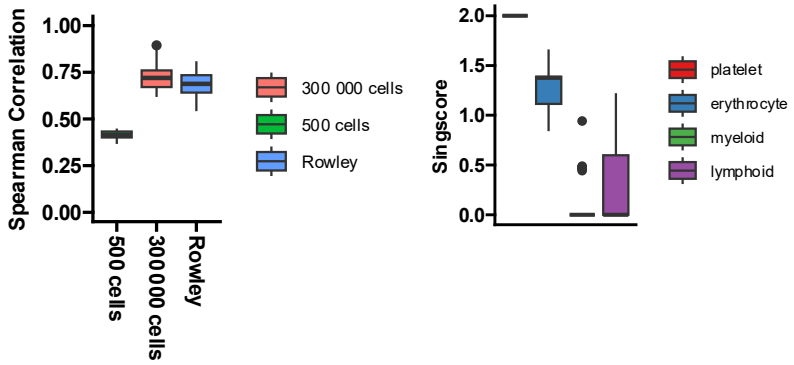

Figure 7 Left: Sample to sample read correlation between libraries prepared from 500 and 300.000 platelet and a published dataset (Rowley et al., 2011). Right: Signature scores for mature blood lineages in platelet transcriptomes prepared from 300.000 platelets.

After sequencing of libraries prepared from 300,000 platelets, higher gene detection rate and sample-to-sample correlations were observed **(Figure 7; Study III, Figure 1C,D)**. Next, we assessed contamination by other peripheral blood cell types. As expected, a considerable number of erythroid transcripts were detected, consistent with previous reports **(Figure 7; Study III, Figure 1E)**. They were primarily from the genes of hemoglobin sub-units. Fewer reads were mapped to other erythroid genes indicating that the libraries were contaminated by circulating or ambient mRNA rather than intact cells **(Study III, Figure 1F)**. This could be explained by the high abundance of erythrocytes and their lysis induced

by centrifugation. We reasoned that hemoglobin contamination was consistent across samples and would therefore not impact downstream analysis.

We next sought to investigate if the age of the mouse or spontaneous platelet activation measured by CD62P externalization had any effect on the platelet transcriptome. Old age is associated with increased thrombosis risk and spontaneous activation is a possible confounder in platelet experiments. Principal component analysis showed little separation between age or CD62P groups indicating minor impact of those variables **(Study III, Figure 2A)**. This was confirmed in differential expression analysis using both DESeq2 and EdgeR which both identified a few genes of importance **(Study III, Figure 2B,C, Supplementary Figure 3A,B)**. Notably the EdgeR analysis found far fewer DE genes for which the baseline expression was low, owing to the stronger default regularization implemented in that package.

In conclusion, this method allows isolation of relatively pure platelet populations from the sample. We intend to apply this method to characterize the platelets generated through the two differentiation pathways identified in **Study II**. Furthermore, the isolation technique can be applied to other types of experiments, such as mass spectrometry, which recently has seen a major leap in sensitivity.

6 Conclusions

Study I

Transcriptional priming of the platelet lineage occurs at the FL stage before migration to the bone marrow, but functional lineage bias emerges post-natally, despite most fetal liver HSCs being *Vwf*-positive.

Study II

Vwf-positive P-HSCs and *Vwf*-negative M-HSCs are not hierarchically related and replenish platelets through mutually exclusive differentiation pathways marked by *Flt3* expression. In steady state, this pathway is used at a low frequency, likely by many types of stem cells. Ablation of hematopoietic progenitors increases the utilization of the *Flt3*-negative pathway. There are distinct transcriptional differences between MkPs generated by *Vwf*-positive P-HSCs and *Vwf*-negative M-HSCs which could contribute to Mk and platelet heterogeneity.

Study III

Pure platelet transcriptomes can be prepared through low volume sampling and FACS sorting using a modified single-cell RNA sequencing protocol. Spontaneous upregulation of CD62P and increased age does not predict transcriptional changes. This method has potential applications for resolving platelet subsets.

7 Points of perspective

7.1 Future studies on platelet biased hematopoiesis

Cardiovascular disease is one of the leading causes of death in the world. It is believed that aging-associated inflammation increases the risk of thrombosis through inflammatory atherosclerotic plaques and hyperreactive platelets. This is the rationale behind the successful prevention of cardiovascular diseases with platelet inhibition and lipid lowering reagents.

Our results raise the exciting possibility that platelet function is determined by the type of megakaryocyte that generated it and by extension its differentiation path and parent HSC. Stated differently, platelet subsets are encoded already at the HSC stage. Poscablo et al (Poscablo et al., 2024) reported that platelets generated through the *Flt3*-negative pathway are quicker to upregulate CD62P and spread more quickly upon stimulation with thrombin. However, no difference in contribution to *in vivo* thrombus formation was found. An alternative explanation is that the excess of platelets generated through the *Flt3*-negative pathway leads to increased thrombus risk rather than some intrinsic property of the *Flt3*-negative platelets. Selective depletion of the cells contributing to the *Flt3*-negative pathway could potentially be a targeted therapy for cardiovascular thrombosis, as depletion of myeloid-biased HSCs has been suggested to rejuvenate the hematopoietic system (Ross et al., 2024).

The relative importance of intrinsic platelet hyperreactivity in comparison to external factors such as vascular calcification and inflammation is difficult to determine. Patients suffering from cardiovascular disease often have severely narrowed vessels, as shown by radiological imaging, and calcification is a main driver of cardiovascular risk (Mortensen et al., 2020). In addition, platelet aggregation *in vitro* does not increase with age in individuals older than 65 (Chan et al., 2020). In clinical practice, aspirin and statins have been highly successful in prevention of thrombosis. Aspirin is a platelet inhibitor and statins lower blood cholesterol. In contrast to mice where platelet count increases with age, it decreases with each decade in humans (Segal & Moliterno, 2006). It is unlikely that the small fraction of platelets generated from the shortcut pathway has any meaningful role, in comparison to the vast number of platelets generated through the canonical pathway, especially when considering the modest effect size in activation experiments (Poscablo et al., 2024), unless reflecting a role distinct from

coagulation. In that regard, platelets have also been suggested to play an immunomodulatory role (Morrell et al., 2014).

Considering pre-existing therapies and global vascular disease in aged individuals, the added benefit of targeted Mk progenitor therapies is probably minimal. On the other hand, there are conditions in which platelet production is deficient, for example after HSC transplantation or in autoimmune thrombocytopenia. If P-HSCs could be selectively activated or expanded *in vitro* or *in vivo* they could serve as bridge therapy until canonical thrombopoiesis has caught up. New expansion protocols for both murine (Wilkinson et al., 2019) and human (Sakurai et al., 2023) HSCs have been published and they could potentially be adapted to instruct lineage bias or by enriching for P-HSCs before transplantation.

The existence of platelet-restricted HSCs in humans is unknown, although P-bias has been suggested in HSCs isolated from patients with myeloproliferative neoplasia, by xenotransplantation as well as in culture experiments (Notta et al., 2016; Psaila et al., 2020; Aksoz et al., 2024). These studies are conducted during stress conditions and investigation of lineage bias hematopoiesis in steady state remains challenging. Tracing of spontaneously occurring somatic (Mitchell et al., 2022) or mitochondrial (Lareau et al., 2019) variants might be technically feasible but would suffer from the same problem of detecting small HSC clones as mutation and barcode-based lineage tracing does in mice. Of note, one study (Weng et al., 2024) claimed to have identified lineage-biased HSC clones using mitochondrial variants but the accuracy of mutational calling and subsequent clonal clustering has been called into question (Lareau et al., 2024). Epivariants have also been shown to allow identification of clones in human hematopoiesis (Scherer et al., 2025). This allows for detection of expanded clones in aged individuals independent of clonal hematopoiesis variants. No evidence of platelet-biased stem cells was found using this approach. The proposed mechanism of epivariant barcoding is the imperfect propagation of methylation state in early embryo development, possibly during HSC specification. Since the barcoding happens during the establishment of the hematopoietic system, the barcode diversity is unlikely to capture small HSC clones. Furthermore, as indicated in **Study I**, neither platelet-biased nor -restricted HSCs are present at this stage of development.

Finally, the role of human FLT3 signaling is less investigated than in the mouse. A case series of a family carrying a FLT3L mutation showed substantial impairment

of multiple progenitor compartments and most notably a severe immune cell deficiencies leading to recurrent infections (Momenilandi et al., 2024), but no significant alterations in the platelet lineage, as is also the case in mice (H. J. McKenna et al., 2000). This suggests that absence of FLT3 is a marker for the shortcut pathway but has no causal influence.

7.2 Evolutionary origins of the blood and vascular system

Studies on the development of HSCs have shown that definitive HSCs emerge from hemogenic endothelium in the AGM region through endothelial to hematopoietic transition, as described in the literature review. Intriguingly, adult HSCs still share markers with endothelial cells such as CD201 and Vwf, highlighting the close developmental origin of the vascular and hematopoietic systems. These markers are not exclusive to the hematopoietic and vascular systems but have high expression there throughout life.

Evolutionary developmental biology is useful as a framework to understand how and why different structures and functions in an organism have evolved. Comparative studies have tried to identify developmental stages of the hematopoietic system in other organisms. The key difference between invertebrate and vertebrate vascular systems is the presence of endothelium. A plausible explanation is that larger bodies with extremities surpass the diffusion limit of oxygen and nutrients, requiring a specialized transport system. Endothelium allows local regulation of blood flow through paracrine signaling as well as facilitated laminar flow. Furthermore, the endothelial cells are vanguards for tissue injury and recruit both immune cells and platelets upon damage.

One compelling theory is that in early invertebrates, cells constituting the coelomic (cavity) epithelium detached and became hemocytes (early blood cells), which then re-epithelialized becoming endothelial cells. (Munoz-Chapuli et al., 2005). These cells retained the ability to detach from the basal membrane during angiogenesis. Importantly, this hypothesis singles out EHT as a highly conserved bi-directional mechanism, which could be one possible explanation of the similarities between the vascular and hematopoietic systems and the Mk lineage in particular.

Another mystery is why mammals have platelets generated by polynucleated megakaryocytes whereas other vertebrates have nucleated thrombocytes. One possible explanation is that placental detachment during birth requires improved

hemostasis (Martin & D'Avino, 2022). Mk polyploidization occurs through endoreplication and has been described as a karyo- and cytokinesis deficiency (Mazzi et al., 2018). HSCs are known to expand during aging, in part due to inflammation, and redirecting hematopoiesis through the Mk-pathway could alleviate replicative pressure without increasing cell numbers.

A potential future study would be to look for thrombocyte biased HSCs in non-mammals such as the zebrafish through fate mapping, which could indicate if platelet bias is dependent on Mk polyploidization.

7.3 Stem cell concepts in hematopoiesis and other tissues

Hematopoiesis has been considered the prototypical stem cell system (Orkin & Zon, 2008), and as such it has been used as a blueprint when investigating other stem cell systems. The hematopoiesis paradigm is in many ways useful to give biologists a common framework, however the notion of multipotency and self-renewal as defined through transplantation is difficult to apply in systems where no transplantation assays exist. Furthermore, new techniques have allowed studies on hematopoiesis in an unperturbed setting. The inferred phenotypes and number of stem cells vary greatly among these studies. The number of transplantable HSCs in a mouse is estimated at 3,000–5,000 (Challen et al., 2009; Eaves, 2015), whereas the number of cells sustaining hematopoiesis, as determined by lineage tracing, is 17,000 with around 5,000 actively contributing (Busch et al., 2015). This is in contrast with results from phylodynamic inferences which estimate the number to be around 70,000 (Kapadia et al., 2025). This discrepancy calls for a reconsideration of the transplantation-derived stem cell concepts and by extension which definition of stem cells should be generalized to other tissues.

There is a diverse range of physical organizations of tissues as well as their physiology, for example colonic crypts or hair follicles. It’s not certain that all tissues are organized in the same way as hematopoiesis. The hematopoietic system is the only liquid organ in the body setting it apart from other tissues, although the bone marrow could also be thought of as a solid organ. Solid organs can broadly be classified into two groups: fast and slow dividing. Epithelium of the skin and the gut get replaced at a high rate, whereas muscle and heart tissue turn over slowly. Tissue-specific stem cell properties are likely tailored to the specific demands of their tissue.

Inspired by the hallmarks of cancer, the hallmarks of stem cells attempts to encompass all the properties that a stem cell has (Beumer & Clevers, 2024). Out of these six hallmarks, "self-renewal" is argued to be most important for the stem cell definition, followed by "multipotency". "Transplantability" is presented as a way to measure these properties but is only available for stem cells from some tissues. Stem cells from solid tissues have the ability to change their phenotype when transplanted suggesting "niche dependency" and "plasticity" as additional hallmarks but this is not the case for HSCs. The final hallmark is "mitigation of DNA damage", since stem cells must persist across a long time and several divisions. Taken together, while hallmarks like this serve as a guide to which properties are important, current definitions of stem cells from different tissues are difficult to reconcile.

An alternative is to define stemness as a state rather than a set of fixed properties of cells (Zipori, 2009). The stem state is characterized by lack of expression of markers for mature cell populations and has the unrealized potential to attain all lineages within a tissue. Transcriptional lineage priming of HSCs is therefore not consistent with this definition. The stem state can be propagated through symmetric cell division to maintain constant stem cell numbers. Ideally, this minimal definition could be applied to all types of stem cells, both embryonal and adult.

Building on this work, Lucie Laplane has suggested that the field, in fact, employs four definitions of stemness; two intrinsic and two extrinsic (Laplane, 2016; Laplane & Solary, 2019). The first is *categorical* stemness, an intrinsic property independent of external factors such as the microenvironment. The second is *dispositional* stemness, an intrinsic property that emerges within the right environment. Dispositional stemness is contrasted with *relational* stemness, an extrinsic property that is given to a non-stem cell by its microenvironment. Finally, *systemic* stemness is a property of an entire system, like a tissue. The HSCs studied in this thesis belong to the first category since lineage bias persists through transplantation. However, it cannot be ruled out that microenvironment in the donor instructed the lineage bias before transplantation. Laplane's analysis has relevance for designing therapies against cancer stem cells. If stemness turns out to be an extrinsic property, therapies that only target cancer stem cells will not be sufficient for a cure.

Another perspective on the stem cell field contends that it has several models of stem cells but lacks a unifying theory (Fagan, 2013). As a result, studies are centered around explaining and reconciling observations from different models and experimental systems rather than testing hypotheses generated by theory. Some mathematical models have been developed to model data from fate mapping and transplantation experiments. These serve to estimate parameters representing cell numbers and flux between cell compartments, but they have not yet been used to make predictions of the same parameters that can be validated in experiment.

7.4 Trees and atlases describe tissue organization

In the bigger perspective, single-cell RNA sequencing has revealed further heterogeneity within hematopoietic cell populations than was previously captured by flow cytometry and transplantation experiments. This necessitates a revision of our model for hematopoiesis from a tree with discrete cell types to a continuous branching structure where cells exist in continuous states and where traditional cell types rather are stable states of the epigenetic and transcriptional network ("What Is Your Conceptual Definition of "Cell Type" in the Context of a Mature Organism?," 2017; Laurenti & Gottgens, 2018). Importantly, attention should be given to the problem of linking cellular function such as lineage bias to the continuous states identified in this model.

Despite these recent advancements, I would argue that the field is still very much rooted in “tree thinking”. In fact, trees have served as tools for structuring data and knowledge even before modern science (Hellström, 2019). As outlined in the methods section, clustering of cells using any data creates discrete cell populations. Combined with trajectory inference, this creates a differentiation hierarchy similar to a tree. Furthermore, functional experiments will invariably create discrete cell populations as long as cells are isolated using flow cytometry. It should be noted that single-cell experiments such as single-cell transplantation could be performed in a continuous manner if all markers are index sorted. Retrospective analysis of marker expression could then be correlated with functional outcomes such as chimerism as in **Study I**.

As omics technologies become cheaper and easier to access, cell biology is finding itself drowning in data. Cell annotation has improved from unstructured searching for gene symbols to data driven label transfer. However, cell annotation requires validation in orthogonal experiments, or it is at risk of becoming circular.

Atlases have evolved together with the sciences for hundreds of years (Daston & Gailson, 2007). They are not only reference materials, they also dictate what is relevant to look for; that is, which cell types are considered canonical. The term "atlas" derives from the Titan Atlas who carries the world on his shoulders and was first used to describe a collection of maps in 1595. Atlases reflect the values of objectivity in science. The first scientific atlases, which were not map collections, showed artists' idealistic drawings of objects from nature objects, such as flowers and birds. Later, after the introduction of photography, as little interference from the experimenter and atlas-maker as possible was desired. However, it became apparent that photography was still experimenter-dependent and riddled with artifacts. A synthesis of the previous modes of objectivity — trained judgement – emerged. With trained judgement, an expert refines measurements, selects the relevant data points, and gives them an interpretation. According to the authors, this is the way scientists are objective today, exemplified by the Human Cell Atlas. In this historical perspective, cell atlases are rather a shredded atlas since, owing to the destructive nature of measurement, they are not images of something that exists like a flower or a location or a traditional histology image. Nevertheless, they still define which cell types are possible when used to annotate a new dataset. With the advent of spatial transcriptomics, the atlas nature could be recovered by integrating molecular data with images.

To address these issues, creation of a reference cell tree (Domcke & Shendure, 2023) for experimental organisms has been suggested to facilitate comparison across studies and interpretation of data. This program has been successful in *C. elegans* where the genealogy of all cells in the organism has been mapped. Novel whole organism lineage tracing techniques could, in theory, provide the data for a whole organism tree of other species than *C. elegans*. The series of papers from the Sanger institute also has established cell phylogenetic trees as powerful models for a stem cell system. New lineage tracing models that allow for sequential recording of cellular events together with lineage tracing have the potential to give measurement of state and lineage from end point sampling (Askary et al., 2025) when they become available in mouse models. As sequencing technologies become less costly and available to more researchers (Gonzalez-Pena et al., 2021), this new "tree thinking" (Stadler et al., 2021) might result in a new paradigm for developmental biology and by extension hematopoiesis research. A type of tree that allows for joint modeling of phenotypic and genealogical data might enable a future synthesis of the cell lineage, type, and state viewpoints.

8 Acknowledgements

This thesis would not have been possible without the supervision from Petter Woll and Sten Eirik Jacobsen, whose joint lab provided excellent training opportunities and room for scientific discussion. The studies constituting the thesis were all collaborative efforts within our lab and with others. I am especially grateful for the collaborations with Joana Carrelha, Stefania Mazzi, Karin Belander Strålin and Masafumi Seki. I also want to thank Tetsuichi Yoshizato and Warren Kretzschmar for our discussions on bioinformatics and statistics. Finally, I would like to thank Michael Hagemann-Jensen and Christoph Ziegenhain for the joint work on RNA sequencing, which forms a major part of the thesis.

In the Woll lab Madeleine Lehander and Stina Virding Culleton shared the PhD-journey with me and made long days in the lab enjoyable. In addition, I would like to thank past and previous members of the Jacobsen lab at KI and in Oxford, Margs Brennan, Hisashi Ishida, Affaf Aliouat, Bishan Wu. Furthermore, Kari Högstrand, Francesca Grasso, Ellen Markljung, Amy Hillen, Edwin Chari all contributed to the studies and made the lab run smoothly. People at HERM and Quarter 6A in Biomedicum all contributed to a supportive working environment. Annette Öster Fernström and Dimitra Vasileiadi for administrative support

Furthermore, I would like to thank Claus Nerlov, Supat Thongjuea, Rickard Sandberg, Sten Linnarsson, Kimberly Siletti and Adam Mead for fruitful discussions on RNA sequencing methodology and data interpretation. Tomas Lindahl and Ankit Macwan, John S. Semple, Karl Johansson and Johan Rebetz together with Petter Höglund and Maaike Jongen and finally Cedric Ghevaert, Holly Foster and Amie Waller for advice on platelet experiments. Hong Qian, Sidinh Luc and Robert Månsson with their labs for our regular meeting and pleasant retreats. Marianne Farnebo and Sofie Bergstrand for the early collaboration on WRAP53 mice that introduced me to lab work.

My formal and informal mentors, Lukas Orre, Eva Hellström Lindberg and Björn Meister have all given valuable advice on how to navigate academia and work. Kristina Rydlund at Södersjukhuset has shown great flexibility in arranging research time.

Special thanks to Laura Covill for proof-reading my thesis.

Finally, thank you to all my friends and family who have supported me during these five years.

9 Declaration about the use of generative Artificial Intelligence

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

10 References

Acar, M., Kocherlakota, K. S., Murphy, M. M., Peyer, J. G., Oguro, H., Inra, C. N., Jaiyeola, C., Zhao, Z., Luby-Phelps, K., & Morrison, S. J. (2015). Deep imaging of bone marrow shows non-dividing stem cells are mainly perisinusoidal. *Nature*, *526*(7571), 126-130.

Adolfsson, J., Mansson, R., Buza-Vidas, N., Hultquist, A., Liuba, K., Jensen, C. T., Bryder, D., Yang, L., Borge, O. J., Thoren, L. A., Anderson, K., Sitnicka, E., Sasaki, Y., Sigvardsson, M., & Jacobsen, S. E. (2005). Identification of Flt3+ lympho-myeloid stem cells lacking erythro-megakaryocytic potential a revised road map for adult blood lineage commitment. *Cell*, *121*(2), 295-306.

Akashi, K., Traver, D., Miyamoto, T., & Weissman, I. L. (2000). A clonogenic common myeloid progenitor that gives rise to all myeloid lineages. *Nature*, *404*(6774), 193-197.

Aksoz, M., Gafencu, G. A., Stoilova, B., Buono, M., Zhang, Y., Turkalj, S., Meng, Y., Jakobsen, N. A., Metzner, M., Clark, S. A., Beveridge, R., Thongjuea, S., Vyas, P., & Nerlov, C. (2024). Hematopoietic stem cell heterogeneity and age-associated platelet bias are evolutionarily conserved. *Sci Immunol*, *9*(98), eadk3469.

Altman, N., & Krzywinski, M. (2018). The curse(s) of dimensionality. *Nature Methods*, *15*(6), 399-400.

Arora, N., Wenzel, P. L., McKinney-Freeman, S. L., Ross, S. J., Kim, P. G., Chou, S. S., Yoshimoto, M., Yoder, M. C., & Daley, G. Q. (2014). Effect of developmental stage of HSC and recipient on transplant outcomes. *Dev Cell*, *29*(5), 621-628.

Askary, A., Chen, W., Choi, J., Du, L. Y., Elowitz, M. B., Gagnon, J. A., Schier, A. F., Seidel, S., Shendure, J., Stadler, T., & Tran, M. (2025). The lives of cells, recorded. *Nature Reviews Genetics*, *26*(3), 203-222.

Ault, K. A., Rinder, H. M., Mitchell, J., Carmody, M. B., Vary, C. P., & Hillman, R. S. (1992). The significance of platelets with increased RNA content (reticulated platelets). A measure of the rate of thrombopoiesis. *Am J Clin Pathol*, *98*(6), 637-646.

Avecilla, S. T., Hattori, K., Heissig, B., Tejada, R., Liao, F., Shido, K., Jin, D. K., Dias, S., Zhang, F., Hartman, T. E., Hackett, N. R., Crystal, R. G., Witte, L., Hicklin, D. J., Bohlen, P., Eaton, D., Lyden, D., de Sauvage, F., & Rafii, S. (2004). Chemokine-mediated interaction of hematopoietic progenitors with the bone marrow vascular niche is required for thrombopoiesis. *Nature Medicine*, *10*(1), 64-71.

Bagnoli, J. W., Ziegenhain, C., Janjic, A., Wange, L. E., Vieth, B., Parekh, S., Geuder, J., Hellmann, I., & Enard, W. (2018). Sensitive and powerful single-cell RNA sequencing using mcSCRB-seq. *Nature Communications*, *9*(1), 2937.

Balazs, A. B., Fabian, A. J., Esmon, C. T., & Mulligan, R. C. (2006). Endothelial protein C receptor (CD201) explicitly identifies hematopoietic stem cells in murine bone marrow. *Blood*, *107*(6), 2317-2321.

Ballmaier, M., & Germeshausen, M. (2011). Congenital amegakaryocytic thrombocytopenia: clinical presentation, diagnosis, and treatment. *Seminars in Thrombosis and Hemostasis*, *37*(6), 673-681.

Bapst, A. M., Dahl, S. L., Knopfel, T., & Wenger, R. H. (2020). Cre-mediated, loxP independent sequential recombination of a tripartite transcriptional stop cassette allows for partial read-through transcription. *Biochim Biophys Acta Gene Regul Mech*, *1863*(8), 194568.

Becker, A. J., McCulloch, E. A., & Till, J. E. (1963). Cytological demonstration of the clonal nature of spleen colonies derived from transplanted mouse marrow cells. *Nature*, *197*(4866), 452-454.

Benz, C., Martins, V. C., Radtke, F., & Bleul, C. C. (2008). The stream of precursors that colonizes the thymus proceeds selectively through the early T lineage precursor stage of T cell development. *J Exp Med*, *205*(5), 1187-1199.

Bergen, V., Lange, M., Peidli, S., Wolf, F. A., & Theis, F. J. (2020). Generalizing RNA velocity to transient cell states through dynamical modeling. *Nature Biotechnology*, *38*(12), 1408-1414.

Bergen, V., Soldatov, R. A., Kharchenko, P. V., & Theis, F. J. (2021). RNA velocity-current challenges and future perspectives. *Molecular Systems Biology*, *17*(8), e10282.

Bertrand, J. Y., Chi, N. C., Santoso, B., Teng, S., Stainier, D. Y., & Traver, D. (2010). Haematopoietic stem cells derive directly from aortic endothelium during development. *Nature*, *464*(7285), 108-111.

Beumer, J., & Clevers, H. (2024). Hallmarks of stemness in mammalian tissues. *Cell Stem Cell*, *31*(1), 7-24.

Biasco, L., Pellin, D., Scala, S., Dionisio, F., Basso-Ricci, L., Leonardelli, L., Scaramuzza, S., Baricordi, C., Ferrua, F., Cicalese, M. P., Giannelli, S., Neduva, V., Dow, D. J., Schmidt, M., Von Kalle, C., Roncarolo, M. G., Ciceri, F., Vicard, P., Wit, E.,...Aiuti, A. (2016). In Vivo Tracking of Human Hematopoiesis Reveals Patterns of Clonal Dynamics during Early and Steady-State Reconstitution Phases. *Cell Stem Cell*, *19*(1), 107-119.

Bock, T. A. (1997). Assay systems for hematopoietic stem and progenitor cells. *STEM CELLS*, *15 Suppl 1*(S2), 185-195.

Boisset, J. C., van Cappellen, W., Andrieu-Soler, C., Galjart, N., Dzierzak, E., & Robin, C. (2010). In vivo imaging of haematopoietic cells emerging from the mouse aortic endothelium. *Nature*, *464*(7285), 116-120.

Bongiovanni, D., Santamaria, G., Klug, M., Santovito, D., Felicetta, A., Hristov, M., von Scheidt, M., Aslani, M., Cibella, J., Weber, C., Moretti, A., Laugwitz, K. L., Peano, C., & Bernlochner, I. (2019). Transcriptome Analysis of Reticulated Platelets Reveals a Prothrombotic Profile. *Thrombosis and Haemostasis*, *119*(11), 1795-1806.

Born, G. V., & Cross, M. J. (1963). The Aggregation of Blood Platelets. *J Physiol*, *168*(1), 178-195.

Bowling, S., Sritharan, D., Osorio, F. G., Nguyen, M., Cheung, P., Rodriguez-Fraticelli, A., Patel, S., Yuan, W. C., Fujiwara, Y., Li, B. E., Orkin, S. H., Hormoz, S., & Camargo, F. D. (2020). An Engineered CRISPR-Cas9 Mouse Line for

Simultaneous Readout of Lineage Histories and Gene Expression Profiles in Single Cells. *Cell*, *181*(6), 1410-1422 e1427.

Boyer, S. W., Beaudin, A. E., & Forsberg, E. C. (2012). Mapping differentiation pathways from hematopoietic stem cells using Flk2/Flt3 lineage tracing. *Cell Cycle*, *11*(17), 3180-3188.

Boyer, S. W., Schroeder, A. V., Smith-Berdan, S., & Forsberg, E. C. (2011). All hematopoietic cells develop from hematopoietic stem cells through Flk2/Flt3-positive progenitor cells. *Cell Stem Cell*, *9*(1), 64-73.

Bradley, T. R., & Metcalf, D. (1966). The growth of mouse bone marrow cells in vitro. *Aust J Exp Biol Med Sci*, *44*(3), 287-299.

Brass, L. F., Diamond, S. L., & Stalker, T. J. (2016). Platelets and hemostasis: a new perspective on an old subject. *Blood Adv*, *1*(1), 5-9.

Bruns, I., Lucas, D., Pinho, S., Ahmed, J., Lambert, M. P., Kunisaki, Y., Scheiermann, C., Schiff, L., Poncz, M., Bergman, A., & Frenette, P. S. (2014). Megakaryocytes regulate hematopoietic stem cell quiescence through CXCL4 secretion. *Nature Medicine*, *20*(11), 1315-1320.

Busch, K., Klapproth, K., Barile, M., Flossdorf, M., Holland-Letz, T., Schlenner, S. M., Reth, M., Hofer, T., & Rodewald, H. R. (2015). Fundamental properties of unperturbed haematopoiesis from stem cells in vivo. *Nature*, *518*(7540), 542-546.

Busque, L., Mio, R., Mattioli, J., Brais, E., Blais, N., Lalonde, Y., Maragh, M., & Gilliland, D. G. (1996). Nonrandom X-Inactivation Patterns in Normal Females: Lyonization Ratios Vary With Age. *Blood*, *88*(1), 59-65.

Buza-Vidas, N., Woll, P., Hultquist, A., Duarte, S., Lutteropp, M., Bouriez-Jones, T., Ferry, H., Luc, S., & Jacobsen, S. E. (2011). FLT3 expression initiates in fully multipotent mouse hematopoietic progenitor cells. *Blood*, *118*(6), 1544-1548.

Calabria, A., Spinozzi, G., Cesana, D., Buscaroli, E., Benedicenti, F., Pais, G., Gazzo, F., Scala, S., Lidonnici, M. R., Scaramuzza, S., Albertini, A., Esposito, S., Tucci, F., Canarutto, D., Omrani, M., De Mattia, F., Dionisio, F., Giannelli, S., Marktel, S.,...Montini, E. (2024). Long-term lineage commitment in haematopoietic stem cell gene therapy. *Nature*, *636*(8041), 162-171.

Cardinal, D. C., & Flower, R. J. (1980). The electronic aggregometer: a novel device for assessing platelet behavior in blood. *J Pharmacol Methods*, *3*(2), 135-158.

Carrelha, J., Mazzi, S., Winroth, A., Hagemann-Jensen, M., Ziegenhain, C., Hogstrand, K., Seki, M., Brennan, M. S., Lehander, M., Wu, B., Meng, Y., Markljung, E., Norfo, R., Ishida, H., Belander Stralin, K., Grasso, F., Simoglou Karali, C., Aliouat, A., Hillen, A.,...Jacobsen, S. E. W. (2024). Alternative platelet differentiation pathways initiated by nonhierarchically related hematopoietic stem cells. *Nature Immunology*, *25*(6), 1007-1019.

Carrelha, J., Meng, Y., Kettyle, L. M., Luis, T. C., Norfo, R., Alcolea, V., Boukarabila, H., Grasso, F., Gambardella, A., Grover, A., Hogstrand, K., Lord, A. M., Sanjuan-Pla, A., Woll, P. S., Nerlov, C., & Jacobsen, S. E. W. (2018). Hierarchically related lineage-restricted fates of multipotent haematopoietic stem cells. *Nature*, *554*(7690), 106-111.

Challen, G. A., Boles, N., Lin, K. K., & Goodell, M. A. (2009). Mouse hematopoietic stem cell identification and analysis. *Cytometry A*, *75*(1), 14-24.

Chan, M. V., Chen, M. H., Barwari, T., Huffman, J. E., Armstrong, P. C., Hwang, S. J., Santer, P., Wierer, B., Mayr, M., Kiechl, S., Johnson, A. D., Willeit, J., & Warner, T. D. (2020). Platelet Reactivity in Individuals Over 65 Years Old Is Not Modulated by Age. *Circ Res*, *127*(3), 394-396.

Chapple, R. H., Tseng, Y. J., Hu, T., Kitano, A., Takeichi, M., Hoegenauer, K. A., & Nakada, D. (2018). Lineage tracing of murine adult hematopoietic stem cells reveals active contribution to steady-state hematopoiesis. *Blood Adv*, *2*(11), 1220-1228.

Cheung, A. M., Nguyen, L. V., Carles, A., Beer, P., Miller, P. H., Knapp, D. J., Dhillon, K., Hirst, M., & Eaves, C. J. (2013). Analysis of the clonal growth and differentiation dynamics of primitive barcoded human cord blood cells in NSG mice. *Blood*, *122*(18), 3129-3137.

Choudry, F. A., Bagger, F. O., Macaulay, I. C., Farrow, S., Burden, F., Kempster, C., McKinney, H., Olsen, L. R., Huang, N., Downes, K., Voet, T., Uppal, R., Martin, J. F., Mathur, A., Ouwehand, W. H., Laurenti, E., Teichmann, S. A., & Frontini, M. (2021). Transcriptional characterization of human megakaryocyte polyploidization and lineage commitment. *Journal of Thrombosis and Haemostasis*, *19*(5), 1236-1249.

Cines, D. B., & Blanchette, V. S. (2002). Immune thrombocytopenic purpura. *New England Journal of Medicine*, *346*(13), 995-1008.

Copelan, E. A. (2006). Hematopoietic stem-cell transplantation. *New England Journal of Medicine*, *354*(17), 1813-1826.

Curry, J. L., & Trentin, J. J. (1967). Hemopoietic spleen colony studies. I. Growth and differentiation. *Dev Biol*, *15*(5), 395-413.

Dahlin, J. S., Hamey, F. K., Pijuan-Sala, B., Shepherd, M., Lau, W. W. Y., Nestorowa, S., Weinreb, C., Wolock, S., Hannah, R., Diamanti, E., Kent, D. G., Gottgens, B., & Wilson, N. K. (2018). A single-cell hematopoietic landscape resolves 8 lineage trajectories and defects in Kit mutant mice. *Blood*, *131*(21), e1-e11.

Daston, L., & Gailson, P. (2007). *Objectivity*. Zone Books.

De Cuyper, I. M., Meinders, M., van de Vijver, E., de Korte, D., Porcelijn, L., de Haas, M., Eble, J. A., Seeger, K., Rutella, S., Pagliara, D., Kuijpers, T. W., Verhoeven, A. J., van den Berg, T. K., & Gutierrez, L. (2013). A novel flow cytometry-based platelet aggregation assay. *Blood*, *121*(10), e70-80.

Decker, M., Leslie, J., Liu, Q., & Ding, L. (2018). Hepatic thrombopoietin is required for bone marrow hematopoietic stem cell maintenance. *Science*, *360*(6384), 106-110.

Dick, J. E., Magli, M. C., Huszar, D., Phillips, R. A., & Bernstein, A. (1985). Introduction of a selectable gene into primitive stem cells capable of long-term reconstitution of the hemopoietic system of W/Wv mice. *Cell*, *42*(1), 71-79.

Domcke, S., & Shendure, J. (2023). A reference cell tree will serve science better than a reference cell atlas. *Cell*, *186*(6), 1103-1114.

Dong, F., Hao, S., Zhang, S., Zhu, C., Cheng, H., Yang, Z., Hamey, F. K., Wang, X., Gao, A., Wang, F., Gao, Y., Dong, J., Wang, C., Wang, J., Lan, Y., Liu, B., Ema, H., Tang, F., Gottgens, B.,...Cheng, T. (2020). Differentiation of transplanted

haematopoietic stem cells tracked by single-cell transcriptomic analysis. *Nature Cell Biology*, *22*(6), 630-639.

Doulatov, S., Notta, F., Laurenti, E., & Dick, J. E. (2012). Hematopoiesis: a human perspective. *Cell Stem Cell*, *10*(2), 120-136.

Durai, V., Bagadia, P., Briseno, C. G., Theisen, D. J., Iwata, A., Davidson, J. T. t., Gargaro, M., Fremont, D. H., Murphy, T. L., & Murphy, K. M. (2018). Altered compensatory cytokine signaling underlies the discrepancy between Flt3(-/-) and Flt3l(-/-) mice. *J Exp Med*, *215*(5), 1417-1435.

Dykstra, B., Kent, D., Bowie, M., McCaffrey, L., Hamilton, M., Lyons, K., Lee, S. J., Brinkman, R., & Eaves, C. (2007). Long-term propagation of distinct hematopoietic differentiation programs in vivo. *Cell Stem Cell*, *1*(2), 218-229.

Eaves, C. J. (2015). Hematopoietic stem cells: concepts, definitions, and the new reality. *Blood*, *125*(17), 2605-2613.

Ema, H., Morita, Y., & Suda, T. (2014). Heterogeneity and hierarchy of hematopoietic stem cells. *Exp Hematol*, *42*(2), 74-82 e72.

Ema, H., Takano, H., Sudo, K., & Nakauchi, H. (2000). In vitro self-renewal division of hematopoietic stem cells. *J Exp Med*, *192*(9), 1281-1288.

Ema, H., Uchinomiya, K., Morita, Y., Suda, T., & Iwasa, Y. (2016). Repopulation dynamics of single haematopoietic stem cells in mouse transplantation experiments: Importance of stem cell composition in competitor cells. *Journal of Theoretical Biology*, *394*, 57-67.

Emadi, A., Jones, R. J., & Brodsky, R. A. (2009). Cyclophosphamide and cancer: golden anniversary. *Nature Reviews Clinical Oncology*, *6*(11), 638-647.

Enver, T., Heyworth, C. M., & Dexter, T. M. (1998). Do stem cells play dice? *Blood*, *92*(2), 348-351; discussion 352.

Fagan, M. B. (2013). *Philosophy of Stem Cell Biology : Knowledge in Flesh and Blood* (1st 2013. ed.). Palgrave Macmillan UK.

Faust, T. E., Feinberg, P. A., O'Connor, C., Kawaguchi, R., Chan, A., Strasburger, H., Frosch, M., Boyle, M. A., Masuda, T., Amann, L., Knobeloch, K. P., Prinz, M., Schaefer, A., & Schafer, D. P. (2023). A comparative analysis of microglial inducible Cre lines. *Cell Rep*, *42*(9), 113031.

First, L. R., Smith, B. R., Lipton, J., Nathan, D. G., Parkman, R., & Rappeport, J. M. (1985). Isolated thrombocytopenia after allogeneic bone marrow transplantation: existence of transient and chronic thrombocytopenic syndromes. *Blood*, *65*(2), 368-374.

Flohr Svendsen, A., Yang, D., Kim, K., Lazare, S., Skinder, N., Zwart, E., Mura-Meszaros, A., Ausema, A., von Eyss, B., de Haan, G., & Bystrykh, L. (2021). A comprehensive transcriptome signature of murine hematopoietic stem cell aging. *Blood*, *138*(6), 439-451.

Ford, E., Hamerton, L., Barnes, D. W. H., & Loutit, F. (1956). CYTOLOGICAL IDENTIFICATION OF RADIATION-CHIM/ERAS. 3.

Ganuza, M., Chabot, A., Tang, X., Bi, W., Natarajan, S., Carter, R., Gawad, C., Kang, G., Cheng, Y., & McKinney-Freeman, S. (2018). Murine hematopoietic stem cell activity is derived from pre-circulation embryos but not yolk sacs. *Nature Communications*, *9*(1), 5405, Article 5405.

Ganuza, M., Hall, T., Finkelstein, D., Chabot, A., Kang, G., & McKinney-Freeman, S. (2017). Lifelong haematopoiesis is established by hundreds of precursors throughout mammalian ontogeny. *Nature Cell Biology*, *19*(10), 1153-1163.

Ganuza, M., Hall, T., Myers, J., Nevitt, C., Sanchez-Lanzas, R., Chabot, A., Ding, J., Kooienga, E., Caprio, C., Finkelstein, D., Kang, G., Obeng, E., & McKinney-Freeman, S. (2022). Murine foetal liver supports limited detectable expansion of life-long haematopoietic progenitors. *Nature Cell Biology*, *24*(10), 1475-1486.

Gong, J. K. (1978). Endosteal marrow: a rich source of hematopoietic stem cells. *Science*, *199*(4336), 1443-1445.

Gonzalez-Pena, V., Natarajan, S., Xia, Y., Klein, D., Carter, R., Pang, Y., Shaner, B., Annu, K., Putnam, D., Chen, W., Connelly, J., Pruett-Miller, S., Chen, X., Easton, J., & Gawad, C. (2021). Accurate genomic variant detection in single cells with primary template-directed amplification. *Proc Natl Acad Sci U S A*, *118*(24).

Goodell, M. A., Brose, K., Paradis, G., Conner, A. S., & Mulligan, R. C. (1996). Isolation and functional properties of murine hematopoietic stem cells that are replicating in vivo. *J Exp Med*, *183*(4), 1797-1806.

Gordon, M. Y., Lewis, J. L., & Marley, S. B. (2002). Of mice and men...and elephants. *Blood*, *100*(13), 4679-4680.

Gottgens, B. (2015). Regulatory network control of blood stem cells. *Blood*, *125*(17), 2614-2620.

Greenland, S., Senn, S. J., Rothman, K. J., Carlin, J. B., Poole, C., Goodman, S. N., & Altman, D. G. (2016). Statistical tests, P values, confidence intervals, and power: a guide to misinterpretations. *Eur J Epidemiol*, *31*(4), 337-350.

Gremmel, T., Frelinger, A. L., 3rd, & Michelson, A. D. (2016). Platelet Physiology. *Seminars in Thrombosis and Hemostasis*, *42*(3), 191-204.

Grover, A., Sanjuan-Pla, A., Thongjuea, S., Carrelha, J., Giustacchini, A., Gambardella, A., Macaulay, I., Mancini, E., Luis, T. C., Mead, A., Jacobsen, S. E., & Nerlov, C. (2016). Single-cell RNA sequencing reveals molecular and functional platelet bias of aged haematopoietic stem cells. *Nature Communications*, *7*(1), 11075.

Gulati, G. S., Zukowska, M., Noh, J. J., Zhang, A., Wesche, D. J., Sinha, R., George, B. M., Weissman, I. L., & Szade, K. (2019). Neogenin-1 distinguishes between myeloid-biased and balanced Hoxb5(+) mouse long-term hematopoietic stem cells. *Proc Natl Acad Sci U S A*, *116*(50), 25115-25125.

Hagemann-Jensen, M., Ziegenhain, C., Chen, P., Ramskold, D., Hendriks, G. J., Larsson, A. J. M., Faridani, O. R., & Sandberg, R. (2020). Single-cell RNA counting at allele and isoform resolution using Smart-seq3. *Nature Biotechnology*, *38*(6), 708-714.

Hagemann-Jensen, M., Ziegenhain, C., & Sandberg, R. (2022). Scalable single-cell RNA sequencing from full transcripts with Smart-seq3xpress. *Nature Biotechnology*, *40*(10), 1452-1457.

Harrison, D. E. (1980). Competitive repopulation: a new assay for long-term stem cell functional capacity. *Blood*, *55*(1), 77-81.

Haynes, W. A., Tomczak, A., & Khatri, P. (2018). Gene annotation bias impedes biomedical research. *Scientific Reports*, *8*(1), 1362.

Hellström, P. (2019). *Trees of Knowledge : Science and the Shape of Genealogy* (Publication Number 51) [Doctoral thesis, monograph, Acta Universitatis Upsaliensis]. DiVA. Uppsala

Herzenberg, L. A., Parks, D., Sahaf, B., Perez, O., Roederer, M., & Herzenberg, L. A. (2002). The history and future of the fluorescence activated cell sorter and flow cytometry: a view from Stanford. *Clin Chem*, *48*(10), 1819-1827.

Hodgson, G. S., & Bradley, T. R. (1979). Properties of haematopoietic stem cells surviving 5-fluorouracil treatment: evidence for a pre-CFU-S cell? *Nature*, *281*(5730), 381-382.

Hughes, J. F., & Page, D. C. (2015). The Biology and Evolution of Mammalian Y Chromosomes. *Annu Rev Genet*, *49*, 507-527.

Islam, S., Kjallquist, U., Moliner, A., Zajac, P., Fan, J. B., Lonnerberg, P., & Linnarsson, S. (2011). Characterization of the single-cell transcriptional landscape by highly multiplex RNA-seq. *Genome Res*, *21*(7), 1160-1167.

Itkin, T., Gur-Cohen, S., Spencer, J. A., Schajnovitz, A., Ramasamy, S. K., Kusumbe, A. P., Ledergor, G., Jung, Y., Milo, I., Poulos, M. G., Kalinkovich, A., Ludin, A., Kollet, O., Shakhar, G., Butler, J. M., Rafii, S., Adams, R. H., Scadden, D. T., Lin, C. P., & Lapidot, T. (2016). Distinct bone marrow blood vessels differentially regulate haematopoiesis. *Nature*, *532*(7599), 323-328.

Iwanami, S., Sato, T., Haeno, H., Xu, L., Imamura, K., Ooehara, J., Lan, X., Nakauchi, H., Iwami, S., & Yamamoto, R. (2025). Stable platelet production via the bypass pathway explains long-term hematopoietic stem cell reconstitution. *iScience*, *28*(6), 112547.

Jacobsen, S. E. W., & Nerlov, C. (2019). Haematopoiesis in the era of advanced single-cell technologies. *Nature Cell Biology*, *21*(1), 2-8.

Jacobson, L. O., Simmons, E. L., Marks, E. K., & Eldredge, J. H. (1951). Recovery from radiation injury. *Science*, *113*(2940), 510-511.

Jaiswal, S., Fontanillas, P., Flannick, J., Manning, A., Grauman, P. V., Mar, B. G., Lindsley, R. C., Mermel, C. H., Burtt, N., Chavez, A., Higgins, J. M., Moltchanov, V., Kuo, F. C., Kluk, M. J., Henderson, B., Kinnunen, L., Koistinen, H. A., Ladenvall, C., Getz, G.,...Ebert, B. L. (2014). Age-related clonal hematopoiesis associated with adverse outcomes. *New England Journal of Medicine*, *371*(26), 2488-2498.

Jassinskaja, M., Gonka, M., & Kent, D. G. (2023). Resolving the hematopoietic stem cell state by linking functional and molecular assays. *Blood*, *142*(6), 543-552.

Jordan, C. T., & Lemischka, I. R. (1990). Clonal and systemic analysis of long-term hematopoiesis in the mouse. *Genes Dev*, *4*(2), 220-232.

Kapadia, C. D., Williams, N., Dawson, K. J., Watson, C., Yousefzadeh, M. J., Le, D., Nyamondo, K., Kodavali, S., Cagan, A., Waldvogel, S., Zhang, X., De La Fuente, J., Leongamornlert, D., Mitchell, E., Florez, M. A., Sosnowski, K., Aguilar, R., Martell, A., Guzman, A.,...Nangalia, J. (2025). Clonal dynamics and somatic evolution of haematopoiesis in mouse. *Nature*.

Kent, D. G., Copley, M. R., Benz, C., Wohrer, S., Dykstra, B. J., Ma, E., Cheyne, J., Zhao, Y., Bowie, M. B., Zhao, Y., Gasparetto, M., Delaney, A., Smith, C., Marra, M., & Eaves, C. J. (2009). Prospective isolation and molecular characterization

of hematopoietic stem cells with durable self-renewal potential. *Blood*, *113*(25), 6342-6350.

Kiel, M. J., Yilmaz, O. H., Iwashita, T., Yilmaz, O. H., Terhorst, C., & Morrison, S. J. (2005). SLAM family receptors distinguish hematopoietic stem and progenitor cells and reveal endothelial niches for stem cells. *Cell*, *121*(7), 1109-1121.

Kienast, J., & Schmitz, G. (1990). Flow cytometric analysis of thiazole orange uptake by platelets: a diagnostic aid in the evaluation of thrombocytopenic disorders. *Blood*, *75*(1), 116-121.

Kim, S., Kim, N., Presson, A. P., Metzger, M. E., Bonifacino, A. C., Sehl, M., Chow, S. A., Crooks, G. M., Dunbar, C. E., An, D. S., Donahue, R. E., & Chen, I. S. (2014). Dynamics of HSPC repopulation in nonhuman primates revealed by a decade-long clonal-tracking study. *Cell Stem Cell*, *14*(4), 473-485.

Kiselev, V. Y., Andrews, T. S., & Hemberg, M. (2019). Challenges in unsupervised clustering of single-cell RNA-seq data. *Nature Reviews Genetics*, *20*(5), 273-282.

Kissa, K., & Herbomel, P. (2010). Blood stem cells emerge from aortic endothelium by a novel type of cell transition. *Nature*, *464*(7285), 112-115.

Kivioja, T., Vaharautio, A., Karlsson, K., Bonke, M., Enge, M., Linnarsson, S., & Taipale, J. (2011). Counting absolute numbers of molecules using unique molecular identifiers. *Nature Methods*, *9*(1), 72-74.

Koelle, S. J., Espinoza, D. A., Wu, C., Xu, J., Lu, R., Li, B., Donahue, R. E., & Dunbar, C. E. (2017). Quantitative stability of hematopoietic stem and progenitor cell clonal output in rhesus macaques receiving transplants. *Blood*, *129*(11), 1448-1457.

Kohler, G., & Milstein, C. (1975). Continuous cultures of fused cells secreting antibody of predefined specificity. *Nature*, *256*(5517), 495-497.

Koide, S., Oshima, M., Kamiya, T., Zheng, Z., Liu, Z., Rizq, O., Nishiyama, A., Murakami, K., Yamada, Y., Nakajima-Takagi, Y., Rahmutulla, B., Kaneda, A., Yokoyama, K., Yusa, N., Imoto, S., Miura, F., Ito, T., Tamura, T., Nerlov, C.,...Iwama, A. (2025). Tracking clusterin expression in hematopoietic stem cells reveals their heterogeneous composition across the life span. *Blood*, *146*(1), 62-75.

Kondo, M., Weissman, I. L., & Akashi, K. (1997). Identification of clonogenic common lymphoid progenitors in mouse bone marrow. *Cell*, *91*(5), 661-672.

Kretchmar, A. L., & Conover, W. R. (1970). A difference between spleen-derived and bone marrow-derived colony-forming units in ability to protect lethally irradiated mice. *Blood*, *36*(6), 772-776.

Kuzmina, Z., Eder, S., Bohm, A., Pernicka, E., Vormittag, L., Kalhs, P., Petkov, V., Stary, G., Nepp, J., Knobler, R., Just, U., Krenn, K., Worel, N., & Greinix, H. T. (2012). Significantly worse survival of patients with NIH-defined chronic graft-versus-host disease and thrombocytopenia or progressive onset type: results of a prospective study. *Leukemia*, *26*(4), 746-756.

La Manno, G., Soldatov, R., Zeisel, A., Braun, E., Hochgerner, H., Petukhov, V., Lidschreiber, K., Kastriti, M. E., Lonnerberg, P., Furlan, A., Fan, J., Borm, L. E., Liu, Z., van Bruggen, D., Guo, J., He, X., Barker, R., Sundstrom, E., Castelo-

Branco, G.,...Kharchenko, P. V. (2018). RNA velocity of single cells. *Nature*, *560*(7719), 494-498.
Laplane, L. (2016). *Cancer stem cells : philosophy and therapies*. Harvard University Press.
Laplane, L., & Solary, E. (2019). Towards a classification of stem cells. *eLife*, *8*.
Lareau, C. A., Chapman, M. S., Penter, L., Nawy, T., Pe'er, D., & Ludwig, L. S. (2024). Artifacts in single-cell mitochondrial DNA mutation analyses misinform phylogenetic inference. *bioRxiv*, 2024.2007.2028.605517.
Lareau, C. A., Ludwig, L. S., & Sankaran, V. G. (2019). Longitudinal assessment of clonal mosaicism in human hematopoiesis via mitochondrial mutation tracking. *Blood Adv*, *3*(24), 4161-4165.
Laurenti, E., & Gottgens, B. (2018). From haematopoietic stem cells to complex differentiation landscapes. *Nature*, *553*(7689), 418-426.
Lazic, S. E., Clarke-Williams, C. J., & Munafo, M. R. (2018). What exactly is 'N' in cell culture and animal experiments? *PLoS Biol*, *16*(4), e2005282.
Lee-Six, H., Obro, N. F., Shepherd, M. S., Grossmann, S., Dawson, K., Belmonte, M., Osborne, R. J., Huntly, B. J. P., Martincorena, I., Anderson, E., O'Neill, L., Stratton, M. R., Laurenti, E., Green, A. R., Kent, D. G., & Campbell, P. J. (2018). Population dynamics of normal human blood inferred from somatic mutations. *Nature*, *561*(7724), 473-478.
Lemischka, I. R., Raulet, D. H., & Mulligan, R. C. (1986). Developmental potential and dynamic behavior of hematopoietic stem cells. *Cell*, *45*(6), 917-927.
Lennartsson, J., & Ronnstrand, L. (2012). Stem cell factor receptor/c-Kit: from basic science to clinical implications. *Physiol Rev*, *92*(4), 1619-1649.
Levi, B. P., Yilmaz, O. H., Duester, G., & Morrison, S. J. (2009). Aldehyde dehydrogenase 1a1 is dispensable for stem cell function in the mouse hematopoietic and nervous systems. *Blood*, *113*(8), 1670-1680.
Li, C. L., & Johnson, G. R. (1992). Rhodamine123 reveals heterogeneity within murine Lin-, Sca-1+ hemopoietic stem cells. *J Exp Med*, *175*(6), 1443-1447.
Li, J. J., Liu, J., Li, Y. E., Chen, L. V., Cheng, H., Li, Y., Cheng, T., Wang, Q. F., & Zhou, B. O. (2024). Differentiation route determines the functional outputs of adult megakaryopoiesis. *Immunity*, *57*(3), 478-494 e476.
Li, L., Bowling, S., McGeary, S. E., Yu, Q., Lemke, B., Alcedo, K., Jia, Y., Liu, X., Ferreira, M., Klein, A. M., Wang, S. W., & Camargo, F. D. (2023). A mouse model with high clonal barcode diversity for joint lineage, transcriptomic, and epigenomic profiling in single cells. *Cell*, *186*(23), 5183-5199 e5122.
Longley, D. B., Harkin, D. P., & Johnston, P. G. (2003). 5-fluorouracil: mechanisms of action and clinical strategies. *Nature Reviews Cancer*, *3*(5), 330-338.
Lord, B. I., & Hendry, J. H. (1972). The distribution of haemopoietic colony-forming units in the mouse femur, and its modification by x rays. *Br J Radiol*, *45*(530), 110-115.
Lord, B. I., Testa, N. G., & Hendry, J. H. (1975). The relative spatial distributions of CFUs and CFUc in the normal mouse femur. *Blood*, *46*(1), 65-72.
Lyman, S. D., & Jacobsen, S. E. (1998). c-kit ligand and Flt3 ligand: stem/progenitor cell factors with overlapping yet distinct activities. *Blood*, *91*(4), 1101-1134.

Machlus, K. R., & Italiano, J. E., Jr. (2013). The incredible journey: From megakaryocyte development to platelet formation. *J Cell Biol*, *201*(6), 785-796.

Mackarehtschian, K., Hardin, J. D., Moore, K. A., Boast, S., Goff, S. P., & Lemischka, I. R. (1995). Targeted disruption of the flk2/flt3 gene leads to deficiencies in primitive hematopoietic progenitors. *Immunity*, *3*(1), 147-161.

Madisen, L., Zwingman, T. A., Sunkin, S. M., Oh, S. W., Zariwala, H. A., Gu, H., Ng, L. L., Palmiter, R. D., Hawrylycz, M. J., Jones, A. R., Lein, E. S., & Zeng, H. (2010). A robust and high-throughput Cre reporting and characterization system for the whole mouse brain. *Nat Neurosci*, *13*(1), 133-140.

Malara, A., Abbonante, V., Di Buduo, C. A., Tozzi, L., Currao, M., & Balduini, A. (2015). The secret life of a megakaryocyte: emerging roles in bone marrow homeostasis control. *Cell Mol Life Sci*, *72*(8), 1517-1536.

Mansell, E., Sigurdsson, V., Deltcheva, E., Brown, J., James, C., Miharada, K., Soneji, S., Larsson, J., & Enver, T. (2021). Mitochondrial Potentiation Ameliorates Age-Related Heterogeneity in Hematopoietic Stem Cell Function. *Cell Stem Cell*, *28*(2), 241-256 e246.

Mansson, R., Hultquist, A., Luc, S., Yang, L., Anderson, K., Kharazi, S., Al-Hashmi, S., Liuba, K., Thoren, L., Adolfsson, J., Buza-Vidas, N., Qian, H., Soneji, S., Enver, T., Sigvardsson, M., & Jacobsen, S. E. (2007). Molecular evidence for hierarchical transcriptional lineage priming in fetal and adult stem cells and multipotent progenitors. *Immunity*, *26*(4), 407-419.

Martin, J. F., & D'Avino, P. P. (2022). A theory of rapid evolutionary change explaining the de novo appearance of megakaryocytes and platelets in mammals. *J Cell Sci*, *135*(24).

Martin, J. F., Kristensen, S. D., Mathur, A., Grove, E. L., & Choudry, F. A. (2012). The causal role of megakaryocyte-platelet hyperactivity in acute coronary syndromes. *Nature Reviews Cardiology*, *9*(11), 658-670.

Mazzi, S., Lordier, L., Debili, N., Raslova, H., & Vainchenker, W. (2018). Megakaryocyte and polyploidization. *Exp Hematol*, *57*, 1-13.

McCulloch, E. A. (1963). Les Clones De Cellules Hematopoietiques Invivo. *Revue Francaise D Etudes Cliniques Et Biologiques*, *8*(1), 15-19.

McCulloch, E. A., & Till, J. E. (1962). The sensitivity of cells from normal mouse bone marrow to gamma radiation in vitro and in vivo. *Radiation Research*, *16*, 822-832.

McInnes, L., Healy, J., & Melville, J. (2018). UMAP: Uniform Manifold Approximation and Projection for Dimension Reduction. arXiv:1802.03426.

McKenna, H. J., Stocking, K. L., Miller, R. E., Brasel, K., De Smedt, T., Maraskovsky, E., Maliszewski, C. R., Lynch, D. H., Smith, J., Pulendran, B., Roux, E. R., Teepe, M., Lyman, S. D., & Peschon, J. J. (2000). Mice lacking flt3 ligand have deficient hematopoiesis affecting hematopoietic progenitor cells, dendritic cells, and natural killer cells. *Blood*, *95*(11), 3489-3497.

McKenna, H. J., Stocking, K. L., Miller, R. E., Brasel, K., De Smedt, T., Maraskovsky, E., Maliszewski, C. R., Lynch, D. H., Smith, J., Pulendran, B., Roux, E. R., Teepe, M., Lyman, S. D., & Peschon, J. J. (2000). Mice lacking flt3 ligand have deficient

hematopoiesis affecting hematopoietic progenitor cells, dendritic cells, and natural killer cells. *Blood*, *95*(11), 3489-3497.

McLellan, M. A., Rosenthal, N. A., & Pinto, A. R. (2017). Cre-loxP-Mediated Recombination: General Principles and Experimental Considerations. *Curr Protoc Mouse Biol*, *7*(1), 1-12.

Medvinsky, A., & Dzierzak, E. (1996). Definitive hematopoiesis is autonomously initiated by the AGM region. *Cell*, *86*(6), 897-906.

Medvinsky, A. L., Samoylina, N. L., Muller, A. M., & Dzierzak, E. A. (1993). An early pre-liver intraembryonic source of CFU-S in the developing mouse. *Nature*, *364*(6432), 64-67.

Meng, Y., Carrelha, J., Drissen, R., Ren, X., Zhang, B., Gambardella, A., Valletta, S., Thongjuea, S., Jacobsen, S. E., & Nerlov, C. (2023). Epigenetic programming defines haematopoietic stem cell fate restriction. *Nature Cell Biology*, *25*(6), 812-822.

Mercier, F. E., Sykes, D. B., & Scadden, D. T. (2016). Single Targeted Exon Mutation Creates a True Congenic Mouse for Competitive Hematopoietic Stem Cell Transplantation: The C57BL/6-CD45.1(STEM) Mouse. *Stem Cell Reports*, *6*(6), 985-992.

Metcalf, D. (1998). Lineage commitment and maturation in hematopoietic cells: the case for extrinsic regulation. *Blood*, *92*(2), 345-347; discussion 352.

Middleton, E. A., Rowley, J. W., Campbell, R. A., Grissom, C. K., Brown, S. M., Beesley, S. J., Schwertz, H., Kosaka, Y., Manne, B. K., Krauel, K., Tolley, N. D., Eustes, A. S., Guo, L., Paine, R., 3rd, Harris, E. S., Zimmerman, G. A., Weyrich, A. S., & Rondina, M. T. (2019). Sepsis alters the transcriptional and translational landscape of human and murine platelets. *Blood*, *134*(12), 911-923.

Miller, C. L., Dykstra, B., & Eaves, C. J. (2008). Characterization of mouse hematopoietic stem and progenitor cells. *Curr Protoc Immunol*, *Chapter 22*(1), 22B 22 21-22B 22 31.

Mitchell, E., Spencer Chapman, M., Williams, N., Dawson, K. J., Mende, N., Calderbank, E. F., Jung, H., Mitchell, T., Coorens, T. H. H., Spencer, D. H., Machado, H., Lee-Six, H., Davies, M., Hayler, D., Fabre, M. A., Mahbubani, K., Abascal, F., Cagan, A., Vassiliou, G. S.,...Campbell, P. J. (2022). Clonal dynamics of haematopoiesis across the human lifespan. *Nature*, *606*(7913), 343-350.

Momenilandi, M., Levy, R., Sobrino, S., Li, J., Lagresle-Peyrou, C., Esmaeilzadeh, H., Fayand, A., Le Floc'h, C., Guerin, A., Della Mina, E., Shearer, D., Delmonte, O. M., Yatim, A., Mulder, K., Mancini, M., Rinchai, D., Denis, A., Neehus, A. L., Balogh, K.,...Beziat, V. (2024). FLT3L governs the development of partially overlapping hematopoietic lineages in humans and mice. *Cell*, *187*(11), 2817-2837 e2831.

Morcos, M. N. F., Li, C., Munz, C. M., Greco, A., Dressel, N., Reinhardt, S., Sameith, K., Dahl, A., Becker, N. B., Roers, A., Hofer, T., & Gerbaulet, A. (2022). Fate mapping of hematopoietic stem cells reveals two pathways of native thrombopoiesis. *Nature Communications*, *13*(1), 4504.

Morrell, C. N., Aggrey, A. A., Chapman, L. M., & Modjeski, K. L. (2014). Emerging roles for platelets as immune and inflammatory cells. *Blood*, *123*(18), 2759-2767.

Mortensen, M. B., Dzaye, O., Steffensen, F. H., Bøtker, H. E., Jensen, J. M., Rønnow Sand, N. P., Kragholm, K. H., Sørensen, H. T., Leipsic, J., Mæng, M., Blaha, M. J., & Nørgaard, B. L. (2020). Impact of Plaque Burden Versus Stenosis on Ischemic Events in Patients With Coronary Atherosclerosis. *Journal of the American College of Cardiology*, *76*(24), 2803-2813.

Muller-Sieburg, C. E., Cho, R. H., Karlsson, L., Huang, J. F., & Sieburg, H. B. (2004). Myeloid-biased hematopoietic stem cells have extensive self-renewal capacity but generate diminished lymphoid progeny with impaired IL-7 responsiveness. *Blood*, *103*(11), 4111-4118.

Muller-Sieburg, C. E., Cho, R. H., Thoman, M., Adkins, B., & Sieburg, H. B. (2002). Deterministic regulation of hematopoietic stem cell self-renewal and differentiation. *Blood*, *100*(4), 1302-1309.

Muller, A. M., Medvinsky, A., Strouboulis, J., Grosveld, F., & Dzierzak, E. (1994). Development of hematopoietic stem cell activity in the mouse embryo. *Immunity*, *1*(4), 291-301.

Munoz-Chapuli, R., Carmona, R., Guadix, J. A., Macias, D., & Perez-Pomares, J. M. (2005). The origin of the endothelial cells: an evo-devo approach for the invertebrate/vertebrate transition of the circulatory system. *Evol Dev*, *7*(4), 351-358.

Nakamura-Ishizu, A., & Suda, T. (2020). Multifaceted roles of thrombopoietin in hematopoietic stem cell regulation. *Ann N Y Acad Sci*, *1466*(1), 51-58.

Nestorowa, S., Hamey, F. K., Pijuan Sala, B., Diamanti, E., Shepherd, M., Laurenti, E., Wilson, N. K., Kent, D. G., & Gottgens, B. (2016). A single-cell resolution map of mouse hematopoietic stem and progenitor cell differentiation. *Blood*, *128*(8), e20-31.

Nguyen, R., Perfetto, S., Mahnke, Y. D., Chattopadhyay, P., & Roederer, M. (2013). Quantifying spillover spreading for comparing instrument performance and aiding in multicolor panel design. *Cytometry A*, *83*(3), 306-315.

Nieswandt, B., Bergmeier, W., Rackebrandt, K., Gessner, J. E., & Zirngibl, H. (2000). Identification of critical antigen-specific mechanisms in the development of immune thrombocytopenic purpura in mice. *Blood*, *96*(7), 2520-2527.

Notta, F., Zandi, S., Takayama, N., Dobson, S., Gan, O. I., Wilson, G., Kaufmann, K. B., McLeod, J., Laurenti, E., Dunant, C. F., McPherson, J. D., Stein, L. D., Dror, Y., & Dick, J. E. (2016). Distinct routes of lineage development reshape the human blood hierarchy across ontogeny. *Science*, *351*(6269), aab2116.

Okada, S., Nakauchi, H., Nagayoshi, K., Nishikawa, S., Nishikawa, S., Miura, Y., & Suda, T. (1991). Enrichment and characterization of murine hematopoietic stem cells that express c-kit molecule. *Blood*, *78*(7), 1706-1712.

Orkin, S. H., & Zon, L. I. (2008). Hematopoiesis: an evolving paradigm for stem cell biology. *Cell*, *132*(4), 631-644.

Osawa, M., Hanada, K., Hamada, H., & Nakauchi, H. (1996). Long-term lymphohematopoietic reconstitution by a single CD34-low/negative hematopoietic stem cell. *Science*, *273*(5272), 242-245.

Paigen, K. (2003a). One hundred years of mouse genetics: an intellectual history. I. The classical period (1902-1980). *Genetics*, *163*(1), 1-7.

Paigen, K. (2003b). One hundred years of mouse genetics: an intellectual history. II. The molecular revolution (1981-2002). *Genetics*, *163*(4), 1227-1235.

Parekh, S., Ziegenhain, C., Vieth, B., Enard, W., & Hellmann, I. (2018). zUMIs - A fast and flexible pipeline to process RNA sequencing data with UMIs. *GigaScience*, *7*(6).

Patel, S. H., Christodoulou, C., Weinreb, C., Yu, Q., da Rocha, E. L., Pepe-Mooney, B. J., Bowling, S., Li, L., Osorio, F. G., Daley, G. Q., & Camargo, F. D. (2022). Lifelong multilineage contribution by embryonic-born blood progenitors. *Nature*, *606*(7915), 747-753.

Peccoud, J., & Ycart, B. (1995). Markovian Modeling of Gene-Product Synthesis. *Theoretical Population Biology*, *48*(2), 222-234.

Pei, W., Shang, F., Wang, X., Fanti, A. K., Greco, A., Busch, K., Klapproth, K., Zhang, Q., Quedenau, C., Sauer, S., Feyerabend, T. B., Hofer, T., & Rodewald, H. R. (2020). Resolving Fates and Single-Cell Transcriptomes of Hematopoietic Stem Cell Clones by PolyloxExpress Barcoding. *Cell Stem Cell*, *27*(3), 383-395 e388.

Perez-Cunningham, J., Boyer, S. W., Landon, M., & Forsberg, E. C. (2016). Hematopoietic stem cell-specific GFP-expressing transgenic mice generated by genetic excision of a pan-hematopoietic reporter gene. *Exp Hematol*, *44*(8), 755-764 e751.

Picelli, S., Faridani, O. R., Bjorklund, A. K., Winberg, G., Sagasser, S., & Sandberg, R. (2014). Full-length RNA-seq from single cells using Smart-seq2. *Nature Protocols*, *9*(1), 171-181.

Pinho, S., Marchand, T., Yang, E., Wei, Q., Nerlov, C., & Frenette, P. S. (2018). Lineage-Biased Hematopoietic Stem Cells Are Regulated by Distinct Niches. *Dev Cell*, *44*(5), 634-641 e634.

Ploemacher, R. E., van der Sluijs, J. P., Voerman, J. S., & Brons, N. H. (1989). An in vitro limiting-dilution assay of long-term repopulating hematopoietic stem cells in the mouse. *Blood*, *74*(8), 2755-2763.

Pluznik, D. H., & Sachs, L. (1965). The cloning of normal "mast" cells in tissue culture. *J Cell Physiol*, *66*(3), 319-324.

Poscablo, D. M., Worthington, A. K., Smith-Berdan, S., Rommel, M. G. E., Manso, B. A., Adili, R., Mok, L., Reggiardo, R. E., Cool, T., Mogharrab, R., Myers, J., Dahmen, S., Medina, P., Beaudin, A. E., Boyer, S. W., Holinstat, M., Jonsson, V. D., & Forsberg, E. C. (2024). An age-progressive platelet differentiation path from hematopoietic stem cells causes exacerbated thrombosis. *Cell*, *187*(12), 3090-3107 e3021.

Prins, D., Park, H. J., Watcham, S., Li, J., Vacca, M., Bastos, H. P., Gerbaulet, A., Vidal-Puig, A., Gottgens, B., & Green, A. R. (2020). The stem/progenitor landscape is reshaped in a mouse model of essential thrombocythemia and causes excess megakaryocyte production. *Sci Adv*, *6*(48).

Psaila, B., Wang, G., Rodriguez-Meira, A., Li, R., Heuston, E. F., Murphy, L., Yee, D., Hitchcock, I. S., Sousos, N., O'Sullivan, J., Anderson, S., Senis, Y. A., Weinberg, O. K., Calicchio, M. L., Center, N. I. H. I. S., Iskander, D., Royston, D., Milojkovic, D., Roberts, I.,...Mead, A. J. (2020). Single-Cell Analyses Reveal

Megakaryocyte-Biased Hematopoiesis in Myelofibrosis and Identify Mutant Clone-Specific Targets. *Mol Cell*, *78*(3), 477-492 e478.

Qian, H., Buza-Vidas, N., Hyland, C. D., Jensen, C. T., Antonchuk, J., Mansson, R., Thoren, L. A., Ekblom, M., Alexander, W. S., & Jacobsen, S. E. (2007). Critical role of thrombopoietin in maintaining adult quiescent hematopoietic stem cells. *Cell Stem Cell*, *1*(6), 671-684.

Ramstrom, S., Sodergren, A. L., Tynngard, N., & Lindahl, T. L. (2016). Platelet Function Determined by Flow Cytometry: New Perspectives? *Seminars in Thrombosis and Hemostasis*, *42*(3), 268-281.

Ringrose, L., Chabanis, S., Angrand, P. O., Woodroofe, C., & Stewart, A. F. (1999). Quantitative comparison of DNA looping in vitro and in vivo: chromatin increases effective DNA flexibility at short distances. *The EMBO Journal*, *18*(23), 6630-6641.

Robb, L. (2007). Cytokine receptors and hematopoietic differentiation. *Oncogene*, *26*(47), 6715-6723.

Roch, A., Trachsel, V., & Lutolf, M. P. (2015). Brief Report: Single-Cell Analysis Reveals Cell Division-Independent Emergence of Megakaryocytes From Phenotypic Hematopoietic Stem Cells. *STEM CELLS*, *33*(10), 3152-3157.

Rodriguez-Fraticelli, A. E., Weinreb, C., Wang, S. W., Migueles, R. P., Jankovic, M., Usart, M., Klein, A. M., Lowell, S., & Camargo, F. D. (2020). Single-cell lineage tracing unveils a role for TCF15 in haematopoiesis. *Nature*, *583*(7817), 585-589.

Rodriguez-Fraticelli, A. E., Wolock, S. L., Weinreb, C. S., Panero, R., Patel, S. H., Jankovic, M., Sun, J., Calogero, R. A., Klein, A. M., & Camargo, F. D. (2018). Clonal analysis of lineage fate in native haematopoiesis. *Nature*, *553*(7687), 212-216.

Rodriguez-Meira, A., Buck, G., Clark, S. A., Povinelli, B. J., Alcolea, V., Louka, E., McGowan, S., Hamblin, A., Sousos, N., Barkas, N., Giustacchini, A., Psaila, B., Jacobsen, S. E. W., Thongjuea, S., & Mead, A. J. (2019). Unravelling Intratumoral Heterogeneity through High-Sensitivity Single-Cell Mutational Analysis and Parallel RNA Sequencing. *Mol Cell*, *73*(6), 1292-1305 e1298.

Rodriguez, Y. B. A., Manso, B. A., & Forsberg, E. C. (2021). CFU-S assay: a historical single-cell assay that offers modern insight into clonal hematopoiesis. *Exp Hematol*, *104*, 1-8.

Roederer, M. (2001). Spectral compensation for flow cytometry: visualization artifacts, limitations, and caveats. *Cytometry*, *45*(3), 194-205.

Ross, J. B., Myers, L. M., Noh, J. J., Collins, M. M., Carmody, A. B., Messer, R. J., Dhuey, E., Hasenkrug, K. J., & Weissman, I. L. (2024). Depleting myeloid-biased haematopoietic stem cells rejuvenates aged immunity. *Nature*, *628*(8006), 162-170.

Rowley, J. W., Oler, A. J., Tolley, N. D., Hunter, B. N., Low, E. N., Nix, D. A., Yost, C. C., Zimmerman, G. A., & Weyrich, A. S. (2011). Genome-wide RNA-seq analysis of human and mouse platelet transcriptomes. *Blood*, *118*(14), e101-111.

Rundberg Nilsson, A., Pronk, C. J., & Bryder, D. (2015). Probing hematopoietic stem cell function using serial transplantation: Seeding characteristics and the impact of stem cell purification. *Exp Hematol*, *43*(9), 812-817 e811.

Saelens, W., Cannoodt, R., Todorov, H., & Saeys, Y. (2019). A comparison of single-cell trajectory inference methods. *Nature Biotechnology*, *37*(5), 547-554.

Sakurai, M., Ishitsuka, K., Ito, R., Wilkinson, A. C., Kimura, T., Mizutani, E., Nishikii, H., Sudo, K., Becker, H. J., Takemoto, H., Sano, T., Kataoka, K., Takahashi, S., Nakamura, Y., Kent, D. G., Iwama, A., Chiba, S., Okamoto, S., Nakauchi, H., & Yamazaki, S. (2023). Chemically defined cytokine-free expansion of human haematopoietic stem cells. *Nature*, *615*(7950), 127-133.

Sanjuan-Pla, A., Macaulay, I. C., Jensen, C. T., Woll, P. S., Luis, T. C., Mead, A., Moore, S., Carella, C., Matsuoka, S., Bouriez Jones, T., Chowdhury, O., Stenson, L., Lutteropp, M., Green, J. C., Facchini, R., Boukarabila, H., Grover, A., Gambardella, A., Thongjuea, S.,...Jacobsen, S. E. (2013). Platelet-biased stem cells reside at the apex of the haematopoietic stem-cell hierarchy. *Nature*, *502*(7470), 232-236.

Sarkar, A., & Stephens, M. (2021). Separating measurement and expression models clarifies confusion in single-cell RNA sequencing analysis. *Nature Genetics*, *53*(6), 770-777.

Sauer, B., & Henderson, N. (1988). Site-specific DNA recombination in mammalian cells by the Cre recombinase of bacteriophage P1. *Proc Natl Acad Sci U S A*, *85*(14), 5166-5170.

Sawai, C. M., Babovic, S., Upadhaya, S., Knapp, D., Lavin, Y., Lau, C. M., Goloborodko, A., Feng, J., Fujisaki, J., Ding, L., Mirny, L. A., Merad, M., Eaves, C. J., & Reizis, B. (2016). Hematopoietic Stem Cells Are the Major Source of Multilineage Hematopoiesis in Adult Animals. *Immunity*, *45*(3), 597-609.

Säwen, P., Eldeeb, M., Erlandsson, E., Kristiansen, T. A., Laterza, C., Kokaia, Z., Karlsson, G., Yuan, J., Soneji, S., Mandal, P. K., Rossi, D. J., & Bryder, D. (2018). Murine HSCs contribute actively to native hematopoiesis but with reduced differentiation capacity upon aging. *eLife*, *7*.

Scadden, D. T. (2006). The stem-cell niche as an entity of action. *Nature*, *441*(7097), 1075-1079.

Scala, S., Basso-Ricci, L., Dionisio, F., Pellin, D., Giannelli, S., Salerio, F. A., Leonardelli, L., Cicalese, M. P., Ferrua, F., Aiuti, A., & Biasco, L. (2018). Dynamics of genetically engineered hematopoietic stem and progenitor cells after autologous transplantation in humans. *Nature Medicine*, *24*(11), 1683-1690.

Scherer, M., Singh, I., Braun, M. M., Szu-Tu, C., Sanchez Sanchez, P., Lindenhofer, D., Jakobsen, N. A., Korber, V., Kardorff, M., Nitsch, L., Kautz, P., Ruhle, J., Bianchi, A., Cozzuto, L., Fromel, R., Beneyto-Calabuig, S., Lareau, C., Satpathy, A. T., Beekman, R.,...Velten, L. (2025). Clonal tracing with somatic epimutations reveals dynamics of blood ageing. *Nature*.

Schofield, R. (1978). The relationship between the spleen colony-forming cell and the haemopoietic stem cell. *Blood Cells*, *4*(1-2), 7-25.

Segal, J. B., & Moliterno, A. R. (2006). Platelet Counts Differ by Sex, Ethnicity, and Age in the United States. *Annals of Epidemiology*, *16*(2), 123-130.

Seyfert, U. T., Haubelt, H., Vogt, A., & Hellstern, P. (2007). Variables influencing Multiplate(TM) whole blood impedance platelet aggregometry and turbidimetric platelet aggregation in healthy individuals. *Platelets*, *18*(3), 199-206.

Shen, F. W., Tung, J. S., & Boyse, E. A. (1986). Further definition of the Ly-5 system. *Immunogenetics*, *24*(3), 146-149.

Siminovitch, L., McCulloch, E. A., & Till, J. E. (1963). The Distribution of Colony-Forming Cells among Spleen Colonies. *J Cell Comp Physiol*, *62*(3), 327-336.

Singh, I., Fernandez-Perez, D., Sanchez, P. S., & Rodriguez-Fraticelli, A. E. (2025). Pre-existing stem cell heterogeneity dictates clonal responses to the acquisition of leukemic driver mutations. *Cell Stem Cell*, *32*(4), 564-580 e566.

Sitnicka, E., Bryder, D., Theilgaard-Monch, K., Buza-Vidas, N., Adolfsson, J., & Jacobsen, S. E. (2002). Key role of flt3 ligand in regulation of the common lymphoid progenitor but not in maintenance of the hematopoietic stem cell pool. *Immunity*, *17*(4), 463-472.

Six, E., Guilloux, A., Denis, A., Lecoules, A., Magnani, A., Vilette, R., Male, F., Cagnard, N., Delville, M., Magrin, E., Caccavelli, L., Roudaut, C., Plantier, C., Sobrino, S., Gregg, J., Nobles, C. L., Everett, J. K., Hacein-Bey-Abina, S., Galy, A.,...Bushman, F. D. (2020). Clonal tracking in gene therapy patients reveals a diversity of human hematopoietic differentiation programs. *Blood*, *135*(15), 1219-1231.

Smith, L. G., Weissman, I. L., & Heimfeld, S. (1991). Clonal analysis of hematopoietic stem-cell differentiation in vivo. *Proc Natl Acad Sci U S A*, *88*(7), 2788-2792.

Soneson, C., & Robinson, M. D. (2018). Bias, robustness and scalability in single-cell differential expression analysis. *Nature Methods*, *15*(4), 255-261.

Song, A. J., & Palmiter, R. D. (2018). Detecting and Avoiding Problems When Using the Cre-lox System. *Trends Genet*, *34*(5), 333-340.

Spangrude, G. J., Heimfeld, S., & Weissman, I. L. (1988). Purification and characterization of mouse hematopoietic stem cells. *Science*, *241*(4861), 58-62.

Squair, J. W., Gautier, M., Kathe, C., Anderson, M. A., James, N. D., Hutson, T. H., Hudelle, R., Qaiser, T., Matson, K. J. E., Barraud, Q., Levine, A. J., La Manno, G., Skinnider, M. A., & Courtine, G. (2021). Confronting false discoveries in single-cell differential expression. *Nature Communications*, *12*(1), 5692.

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

Stadler, T., Pybus, O. G., & Stumpf, M. P. H. (2021). Phylodynamics for cell biologists. *Science*, *371*(6526), eaah6266.

Stefan, A. M., & Schonbrodt, F. D. (2023). Big little lies: a compendium and simulation of p-hacking strategies. *R Soc Open Sci*, *10*(2), 220346.

Stegner, D., vanEeuwijk, J. M. M., Angay, O., Gorelashvili, M. G., Semeniak, D., Pinnecker, J., Schmithausen, P., Meyer, I., Friedrich, M., Dutting, S., Brede, C., Beilhack, A., Schulze, H., Nieswandt, B., & Heinze, K. G. (2017). Thrombopoiesis is spatially regulated by the bone marrow vasculature. *Nature Communications*, *8*(1), 127.

Subramanian, A., Tamayo, P., Mootha, V. K., Mukherjee, S., Ebert, B. L., Gillette, M. A., Paulovich, A., Pomeroy, S. L., Golub, T. R., Lander, E. S., & Mesirov, J. P. (2005). Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci U S A*, *102*(43), 15545-15550.

Suda, J., Suda, T., & Ogawa, M. (1984). Analysis of differentiation of mouse hemopoietic stem cells in culture by sequential replating of paired progenitors. *Blood*, *64*(2), 393-399.

Sugiyama, T., Kohara, H., Noda, M., & Nagasawa, T. (2006). Maintenance of the hematopoietic stem cell pool by CXCL12-CXCR4 chemokine signaling in bone marrow stromal cell niches. *Immunity*, *25*(6), 977-988.

Sun, J., Ramos, A., Chapman, B., Johnnidis, J. B., Le, L., Ho, Y. J., Klein, A., Hofmann, O., & Camargo, F. D. (2014). Clonal dynamics of native haematopoiesis. *Nature*, *514*(7522), 322-327.

Sun, S., Jin, C., Si, J., Lei, Y., Chen, K., Cui, Y., Liu, Z., Liu, J., Zhao, M., Zhang, X., Tang, F., Rondina, M. T., Li, Y., & Wang, Q. F. (2021). Single-cell analysis of ploidy and the transcriptome reveals functional and spatial divergency in murine megakaryopoiesis. *Blood*, *138*(14), 1211-1224.

Svensson, V., da Veiga Beltrame, E., & Pachter, L. (2020). A curated database reveals trends in single-cell transcriptomics. *Database (Oxford)*, *2020*, baaa073.

Svensson, V., Vento-Tormo, R., & Teichmann, S. A. (2018). Exponential scaling of single-cell RNA-seq in the past decade. *Nature Protocols*, *13*(4), 599-604.

Szilvassy, S. J., Humphries, R. K., Lansdorp, P. M., Eaves, A. C., & Eaves, C. J. (1990). Quantitative assay for totipotent reconstituting hematopoietic stem cells by a competitive repopulation strategy. *Proc Natl Acad Sci U S A*, *87*(22), 8736-8740.

Taichman, R. S., & Emerson, S. G. (1998). The role of osteoblasts in the hematopoietic microenvironment. *STEM CELLS*, *16*(1), 7-15.

Takahashi, M., Barile, M., Chapple, R. H., Tseng, Y. J., Nakada, D., Busch, K., Fanti, A. K., Sawen, P., Bryder, D., Hofer, T., & Gottgens, B. (2021). Reconciling Flux Experiments for Quantitative Modeling of Normal and Malignant Hematopoietic Stem/Progenitor Dynamics. *Stem Cell Reports*, *16*(4), 741-753.

Tang, F., Barbacioru, C., Wang, Y., Nordman, E., Lee, C., Xu, N., Wang, X., Bodeau, J., Tuch, B. B., Siddiqui, A., Lao, K., & Surani, M. A. (2009). mRNA-Seq whole-transcriptome analysis of a single cell. *Nature Methods*, *6*(5), 377-382.

Till, J. E., & McCulloch, E. A. (1961). A direct measurement of the radiation sensitivity of normal mouse bone marrow cells. *Radiation Research*, *14*, 213-222.

Till, J. E., McCulloch, E. A., & Siminovitch, L. (1964). A Stochastic Model of Stem Cell Proliferation, Based on the Growth of Spleen Colony-Forming Cells. *Proc Natl Acad Sci U S A*, *51*(1), 29-36.

Van de Sande, B., Flerin, C., Davie, K., De Waegeneer, M., Hulselmans, G., Aibar, S., Seurinck, R., Saelens, W., Cannoodt, R., Rouchon, Q., Verbeiren, T., De Maeyer, D., Reumers, J., Saeys, Y., & Aerts, S. (2020). A scalable SCENIC workflow for single-cell gene regulatory network analysis. *Nature Protocols*, *15*(7), 2247-2276.

Van den Berge, K., Roux de Bezieux, H., Street, K., Saelens, W., Cannoodt, R., Saeys, Y., Dudoit, S., & Clement, L. (2020). Trajectory-based differential expression analysis for single-cell sequencing data. *Nature Communications*, *11*(1), 1201.

van der Maaten, L., & Hinton, G. (2008). Visualizing Data using t-SNE. *Journal of Machine Learning Research*, *9*, 2579-2605.

van der Meijden, P. E. J., & Heemskerk, J. W. M. (2019). Platelet biology and functions: new concepts and clinical perspectives. *Nature Reviews Cardiology*, *16*(3), 166-179.

Van Hove, H., Antunes, A. R. P., De Vlaminck, K., Scheyltjens, I., Van Ginderachter, J. A., & Movahedi, K. (2020). Identifying the variables that drive tamoxifen-independent CreERT2 recombination: Implications for microglial fate mapping and gene deletions. *Eur J Immunol*, *50*(3), 459-463.

Velik-Salchner, C., Maier, S., Innerhofer, P., Streif, W., Klingler, A., Kolbitsch, C., & Fries, D. (2008). Point-of-care whole blood impedance aggregometry versus classical light transmission aggregometry for detecting aspirin and clopidogrel: the results of a pilot study. *Anesth Analg*, *107*(6), 1798-1806.

Versteeg, H. H., Heemskerk, J. W., Levi, M., & Reitsma, P. H. (2013). New fundamentals in hemostasis. *Physiol Rev*, *93*(1), 327-358.

Vieth, B., Parekh, S., Ziegenhain, C., Enard, W., & Hellmann, I. (2019). A systematic evaluation of single cell RNA-seq analysis pipelines. *Nature Communications*, *10*(1), 4667.

Villani, A. C., Satija, R., Reynolds, G., Sarkizova, S., Shekhar, K., Fletcher, J., Griesbeck, M., Butler, A., Zheng, S., Lazo, S., Jardine, L., Dixon, D., Stephenson, E., Nilsson, E., Grundberg, I., McDonald, D., Filby, A., Li, W., De Jager, P. L.,...Hacohen, N. (2017). Single-cell RNA-seq reveals new types of human blood dendritic cells, monocytes, and progenitors. *Science*, *356*(6335), eaah4573.

Visser, J. W., Bauman, J. G., Mulder, A. H., Eliason, J. F., & de Leeuw, A. M. (1984). Isolation of murine pluripotent hemopoietic stem cells. *J Exp Med*, *159*(6), 1576-1590.

Waterstrat, A., Liang, Y., Swiderski, C. F., Shelton, B. J., & Van Zant, G. (2010). Congenic interval of CD45/Ly-5 congenic mice contains multiple genes that may influence hematopoietic stem cell engraftment. *Blood*, *115*(2), 408-417.

Weinreb, C., Rodriguez-Fraticelli, A., Camargo, F. D., & Klein, A. M. (2020). Lineage tracing on transcriptional landscapes links state to fate during differentiation. *Science*, *367*(6479), eaaw3381.

Weng, C., Yu, F., Yang, D., Poeschla, M., Liggett, L. A., Jones, M. G., Qiu, X., Wahlster, L., Caulier, A., Hussmann, J. A., Schnell, A., Yost, K. E., Koblan, L. W., Martin-Rufino, J. D., Min, J., Hammond, A., Ssozi, D., Bueno, R., Mallidi, H.,...Sankaran, V. G. (2024). Deciphering cell states and genealogies of human haematopoiesis. *Nature*, *627*(8003), 389-398.

What Is Your Conceptual Definition of "Cell Type" in the Context of a Mature Organism? (2017). *Cell Syst*, *4*(3), 255-259.

Wilkerson, W. R., & Sane, D. C. (2002). Aging and thrombosis. *Seminars in Thrombosis and Hemostasis*, *28*(6), 555-568.

Wilkinson, A. C., Ishida, R., Kikuchi, M., Sudo, K., Morita, M., Crisostomo, R. V., Yamamoto, R., Loh, K. M., Nakamura, Y., Watanabe, M., Nakauchi, H., & Yamazaki, S. (2019). Long-term ex vivo haematopoietic-stem-cell expansion allows nonconditioned transplantation. *Nature*, *571*(7763), 117-121.

Williams, D. A., Lemischka, I. R., Nathan, D. G., & Mulligan, R. C. (1984). Introduction of new genetic material into pluripotent haematopoietic stem cells of the mouse. *Nature*, *310*(5977), 476-480.

Wilson, N. K., Kent, D. G., Buettner, F., Shehata, M., Macaulay, I. C., Calero-Nieto, F. J., Sanchez Castillo, M., Oedekoven, C. A., Diamanti, E., Schulte, R., Ponting, C. P., Voet, T., Caldas, C., Stingl, J., Green, A. R., Theis, F. J., & Gottgens, B. (2015). Combined Single-Cell Functional and Gene Expression Analysis Resolves Heterogeneity within Stem Cell Populations. *Cell Stem Cell*, *16*(6), 712-724.

Worton, R. G., McCulloch, E. A., & Till, J. E. (1969a). Physical separation of hemopoietic stem cells differing in their capacity for self-renewal. *J Exp Med*, *130*(1), 91-103.

Worton, R. G., McCulloch, E. A., & Till, J. E. (1969b). Physical separation of hemopoietic stem cells from cells forming colonies in culture. *J Cell Physiol*, *74*(2), 171-182.

Wu, A. M., Till, J. E., Siminovitch, L., & McCulloch, E. A. (1968). Cytological evidence for a relationship between normal hemotopoietic colony-forming cells and cells of the lymphoid system. *J Exp Med*, *127*(3), 455-464.

Yamamoto, R., Morita, Y., Ooehara, J., Hamanaka, S., Onodera, M., Rudolph, K. L., Ema, H., & Nakauchi, H. (2013). Clonal analysis unveils self-renewing lineage-restricted progenitors generated directly from hematopoietic stem cells. *Cell*, *154*(5), 1112-1126.

Yang, L., Bryder, D., Adolfsson, J., Nygren, J., Mansson, R., Sigvardsson, M., & Jacobsen, S. E. (2005). Identification of Lin(-)Sca1(+)kit(+)CD34(+)Flt3-short-term hematopoietic stem cells capable of rapidly reconstituting and rescuing myeloablated transplant recipients. *Blood*, *105*(7), 2717-2723.

Zhao, M., Perry, J. M., Marshall, H., Venkatraman, A., Qian, P., He, X. C., Ahamed, J., & Li, L. (2014). Megakaryocytes maintain homeostatic quiescence and promote post-injury regeneration of hematopoietic stem cells. *Nature Medicine*, *20*(11), 1321-1326.

Zhu, Y. Y., Machleder, E. M., Chenchik, A., Li, R., & Siebert, P. D. (2001). Reverse Transcriptase Template Switching: A SMART™ Approach for Full-Length cDNA Library Construction. *BioTechniques*, *30*(4), 892-897.

Zimmerman, K. D., Espeland, M. A., & Langefeld, C. D. (2021). A practical solution to pseudoreplication bias in single-cell studies. *Nature Communications*, *12*(1), 738.

Zipori, D. (2009). *Biology of stem cells and the molecular basis of the stem state*. Humana.

Zovein, A. C., Hofmann, J. J., Lynch, M., French, W. J., Turlo, K. A., Yang, Y., Becker, M. S., Zanetta, L., Dejana, E., Gasson, J. C., Tallquist, M. D., & Iruela-Arispe, M. L. (2008). Fate tracing reveals the endothelial origin of hematopoietic stem cells. *Cell Stem Cell*, *3*(6), 625-636.