Direct reprogramming of murine fibroblasts to blood

A thesis submitted to The University of Manchester

for the degree of Doctor of Philosophy

in the Faculty of Medical and Human Sciences

2015

Magdalena Bronislawa Florkowska

School of Medicine

Institute of Cancer Sciences

List of contents

Word count 56745

List of figures

List of tables

List of abbreviations

AcLDL- Acetylated-low density lipoprotein
AGM - Aorta-gonad-mesonephros
BCR – B-cell receptor
BM- Bone marrow
Bmp4 – Bone morphogenic protein 4
BRY - Brachyury
CAR - Chimeric antigen receptor
CDH5 – VE-cadherin
cDNA – complementary DNA
CFU-C - Colony-forming unit count
ChIP-seq – Chromatin immunoprecipitation coupled with high-throughput sequencing
CLP - Common lymphoid progenitor
CMP - Common myeloid progenitor
DA – Dorsal aorta
DABG – Detection above background
Dam - DNA adenine methyltransferase
DamID - DNA adenine methyltransferase identification
DEAE- Diethylaminoethyl
DMEM - Dulbecco's Modified Eagle Medium
DMSO - dimethyl sulfoxide
DN – CD4/CD8 double negative fraction
DP - CD4/CD8 double positive fraction
Dox – Doxycycline
E – Embryonic day
EB – Embryoid body
EF1α - Elongation factor 1α promoter
EHT – Endothelial to haematopoietic transition
ELISA - Enzyme-linked immunosorbent assay
EML – Erythroid myeloid lymphoid cell line
EMPs – Erythro-myeloid progenitors
EPO - Erythropoietin
ESCs - Embryonic stem cells
EtOH - Ethanol
FACS – Fluorescence-activated cell sorting
FCS - Fetal calf serum
FDA – Food and Drug Administration
FL – Fetal liver
FLK1 - Fetal liver kinase 1 (VEGF receptor 2)
FLT3 - Fms-related tyrosine kinase 3

FLT3L- FLT3 ligand
GAF – Reporter mouse model with *Gfp* expression driven by *Ai467606* promoter
GFB - Reporter mouse model with *Gfp* expression driven by *Gfi1b* promoter
GFP – Green fluorescent protein
GM-CSF – Granulocyte-macrophage colony-stimulating factor
GMP - Granulocyte/macrophage progenitor
GMT – combination of cardiomyocyte reprogramming factors (GATA4, MEF2c, TBX5)
GRNs – Gene regulatory networks
GSEA – Gene set enrichment analysis
GSK3 - Glycogen synthase kinase 3
H3K4me - Monomethylation of lysine four of histone 3
H3K9me3 - Trimethylation of lysine nine of histone 3
H3K27me3 - Trimethylation of lysine 27 of histone 3
HE – Haemogenic endothelium
HEK293T - Human embryonic kidney cells
HET – Haematopoietic to endothelial transition
HLA – Human leukocyte antigen
HPC-7 - Haematopoietic precursor cell-7 line
HSCs – Haematopoietic stem cells
HSPCs – Haematopoietic stem and progenitor cells
IFNγ – Interferon γ
iHSCs – inducible HSCs
IL - Interleukin
IMDM - Iscove's Modified Dulbecco's Medium
iMPPs – inducible MPPs
IPA – Ingenuity pathway analysis
iPSCs – Induced pluripotent stem cells
IRES - Internal ribosome entry site
iScl –inducible *Scl* mouse model
IT-HSCs – Intermediate-term HSCs
kb - kilobase
KL – Kit ligand
LBM – Liquid big mix (full reprogramming medium)
LDA – Limiting dilution assay
L-Gln – L-glutamine
LIN – Lineage
LMPP - Lymphoid primed MPP
LPS - Lipopolysaccharide
LSCs - Limbal stem/progenitor cells
LSK - $LIN^+SCA1^+cKIT^+$ cells

LT-HSCs – Long-term HSCs
LTR – Long terminal repeats
MAFs – Mouse adult fibroblasts
M-CSF - Macrophage colony-stimulating factor
MEFs - Mouse embryonic fibroblasts
MEM - Minimum Essential Medium Eagle
MEP - Megakaryocyte/erythrocyte progenitor
MPP - Multipotent progenitor
mRNA – messenger RNA
MSB – *Gfi1:dtTomato* reporter cell line
MSBB – *Gfi1/Gfi1b* double reporter cell line
MTG - Alpha-monothioglycerol
NaCl – Sodium chloride
NSG - NOD-SCID IL2R$\gamma^{-/-}$ mouse
OSKM – combination of Yamanaka reprogramming factors (OCT4, SOX2, KL4, cMYC)
PAP – Pulmonary alveolar proteinosis
PARP1 - Poly ADP ribose polymerase 1
PB – Peripheral blood
PBS- Phosphate buffered saline
PCA – Principal component analysis
PCR - Polymerase chain reaction
PDS - Plasma-derived serum
PK – Proteinase K
PL – Placenta
PROM1- Prominin 1
P/S - Penicillin/streptomycin
P-Sp - Para-aortic splanchnopleura
qRT-PCR - quantitative real time PCR
RE - Restriction enzyme
RT - Room temperature
SD – Standard deviation
SDS - Sodium dodecyl sulphate
SEM – Standard error of the mean
SP – CD4 or CD8 single positive fraction
ST-HSCs – Short-term HSCs
TCR – T-cell receptor
T/E – Trypsin/EDTA solution
TF – Transcription factor
TGFβ - Transforming growth factor β
T-iPSCs – T-cell derived iPSCs
TNFα - Tumour necrosis factor α

TPO – Thrombopoietin
TRE – tetracyklin responsive element
rtTA – reverse tetracyklin transactivator
VEGF – Vascular endothelial growth factor
WT – Wild type
YS – Yolk sac

Abstract

Magdalena Bronislawa Florkowska
University of Manchester
Degree of Doctor of Philosophy

September 2015

Direct reprogramming of murine fibroblasts to blood.

Transdifferentiation is the direct conversion of somatic cells to another functionally distinct cell type with bypassing the pluripotency state. This rapid cellular fate change can be achieved by overexpression of lineage specific transcription factors, as recently shown for reprogramming to neuronal or muscle cells. The aim of this project was to evaluate whether this approach could be applied to the generation of transplantable haematopoietic cells for regenerative medicine. As a result of my investigation, I have defined a limited set of haematopoietic regulators that convert murine fibroblasts to blood. Expression of only five genes (*Scl, Lmo2, Gata2, Runx1c, Erg*) in fibroblasts led to the generation of round grape-like blood colonies within 6 days of culture. These emerging cells, when cultured with cytokines, gave rise to myeloid, erythroid, megakaryocytic and T-lymphoid lineages. A similar procedure performed on fibroblasts lacking *p53* led to even more robust reprogramming with additional generation of mature B-lymphocytes. The timely investigation of early stage reprogramming cultures revealed that transdifferentiation is mediated through two intermediates: transient endothelial VE-cadherin^{+} haemogenic endothelium and multipotent cKIT^{+} haematopoietic progenitors. Importantly, the latter exhibited robust *in vitro* clonogenic potential and short-term *in vivo* repopulation capabilities. The major focus is now to confer long-term multilineage repopulation potential on reprogrammed cells mostly by amending culture conditions and improving existing protocol. Achieving this goal will open new avenues for the generation of patient-specific haematopoietic cells for therapeutic applications.

Declaration

No portion of the work referred to in the thesis has been submitted in support of an application for another degree or qualification of this or any other university or other institute of learning.

Copyright statement

The following four notes on copyright and the ownership of intellectual property rights must be included as written below:

i. The author of this thesis (including any appendices and/or schedules to this thesis) owns certain copyright or related rights in it (the “Copyright”) and s/he has given The University of Manchester certain rights to use such Copyright, including for administrative purposes.

ii. Copies of this thesis, either in full or in extracts and whether in hard or electronic copy, may be made only in accordance with the Copyright, Designs and Patents Act 1988 (as amended) and regulations issued under it or, where appropriate, in accordance with licensing agreements which the University has from time to time. This page must form part of any such copies made.

iii. The ownership of certain Copyright, patents, designs, trade marks and other intellectual property (the “Intellectual Property”) and any reproductions of copyright works in the thesis, for example graphs and tables (“Reproductions”), which may be described in this thesis, may not be owned by the author and may be owned by third parties. Such Intellectual Property and Reproductions cannot and must not be made available for use without the prior written permission of the owner(s) of the relevant Intellectual Property and/or Reproductions.

iv. Further information on the conditions under which disclosure, publication and commercialisation of this thesis, the Copyright and any Intellectual Property and/or Reproductions described in it may take place is available in the University IP Policy (see http://documents.manchester.ac.uk/DocuInfo.aspx?DocID=487), in any relevant Thesis restriction declarations deposited in the University Library, The University Library’s regulations (see http://www.manchester.ac.uk/library/aboutus/regulations) and in The University’s policy on Presentation of Theses

Statement I: Scientific content

Part of the work presented in this thesis has been conducted in collaboration with others. This is clearly stated at each point it applies. Part of the results described in chapter 4 and 5 has been already published (Appendix 7).

Statement II: Publication

I am a first co-author* in the following publication:

Kiran Batta*, **Magdalena Florkowska***, Valerie Kouskoff and Georges Lacaud 'Direct Reprogramming of Murine Fibroblasts to Hematopoietic Progenitor Cells'

Cell Reports 2014, Dec 11, 9(5):1871-84

This publication is included in Appendix 7

Acknowledgements

First and foremost I would like to thank Dr Georges Lacaud, my supervisor, for all the patience and scientific as well as moral support throughout my PhD program. I am equally grateful to my co-supervisor Dr Valerie Kouskoff for invaluable insights into my project and all the scientific advice given to me during the last four years.

Thank you to all, current and former, members of Stem Cell Biology and Stem Cell Research Groups. The friendly and supportive atmosphere you created in the laboratory during my time in the Paterson Institute, as well as all your scientific advice and expertise, undoubtedly helped me to reach the point of submission.

Huge appreciation goes to the employees of CRUK Manchester Institute Core Facilities, especially Flow Cytometry Division.

Thank you to Monika for patience, patience and patience. I appreciate it took more effort from you than from me to get us to this point.

Thank you to Rabab, both on personal and scientific level. Particularly that you were always there for me during this demanding time.

Finally, thank you to my family for endless and unconditional support

Chapter 1

1. Introduction

1.1. Haematopoiesis

1.1.1. Embryonic sites of blood system development

Embryonic development of the blood system can be separated into three major waves. Primitive progenitors are first generated to provide rapidly blood cells for the growing conceptus. Shortly afterwards this transient population is replaced by definitive erythro-myeloid progenitors (EMPs) and haematopoietic stem cells (HSCs). The initial site of haematopoietic differentiation is the yolk sac (YS) which develops from posterolateral epiblast (Lawson et al., 1991). Between mouse embryonic day 6.5 (E6.5) and E7.25 a pool of epiblast cells migrates from the early primitive streak and contributes to the mesodermal compartment of the YS, which is subsequently overlaid with visceral endoderm (Kinder et al., 1999; Lawson and Pedersen, 1987). Later in development, at the mid-primitive streak stage, mesodermal masses are formed as a result of extensive proliferation of the mesodermal layer. With time, these local sites become a region of rapid generation of primitive erythroid cells, which leads to the emergence of a single belt-like structure known as a blood band (E7.5) (Ferkowicz and Yoder, 2005; Silver and Palis, 1997). This region of the YS is the sole site of primitive haematopoietic activity in the embryo, producing not only primitive erythrocytes but also primitive macrophages and megakaryocytes (Palis et al., 1999; Tober et al., 2007) (Fig 1.1).

Transitory primitive haematopoiesis is then gradually replaced by the long-term definitive haematopoietic system. The very first studies performed as early as the mid 1970's by Dieterlen-Lievre and colleagues suggested that this secondary wave has an intra-embryonic origin. The exact site of emergence of

definitive haematopoietic activity was originally identified as the aorta-gonad-mesonephros (AGM) region that develops from the para-aortic splanchnopleura (P-Sp), present between E8 and E9. A particularly important site of the AGM region is the dorsal aorta (DA) in which clusters of haematopoietic cells appear as early as E9 and peak at E10.5 (Yokomizo and Dzierzak, 2010). The emergence of these structures is timely associated with the establishment of HSC activity in the AGM. Notably, E8 P-Sp already contains functional definitive lymphoid and myeloid cells, whereas at the same time point cells from the YS can commit solely to the myeloid lineage (Cumano et al., 2001; Tavian et al., 2001) (Fig 1.1). More recently, it was established that autonomous B- and T-lymphoid potential is present in E9.5 YS of *Ncx*$^{-/-}$ embryos that lack circulation (Yoshimoto et al., 2011; Yoshimoto et al., 2012). Interestingly, already at E9 the developing AGM and mature YS both contain neonatal HSCs which provide multilineage engraftment when injected directly into the liver but do not support long-term repopulation of adult recipients (Yoder et al., 1997) (Fig 1.1). So far the AGM region remains the only recognized site capable of the autonomous initiation and expansion of definitive HSCs between E10 and 11, therefore it is still widely accepted that the YS and FL are colonized by cells in circulation (Medvinsky and Dzierzak, 1996). However, recent evidence demonstrated that YS cells can contribute to adult haematopoiesis. By tracking Runt-related transcription factor 1 (*Runx1*) expression at the onset of haematopoiesis Samokhalow and colleagues revealed the presence of YS-derived cells in the adult lymphoid population and HSC pool (Samokhvalov et al., 2007). Another study utilising resque of Runx1-null embryos indicated extraembryonic mesoderm as a source of HSC precursors (Tanaka et al., 2012a). This and other findings contributed to the hypothesis that definitive precursors generated in the YS may require an additional maturation step in adult haematopoietic sites (Matsuoka et al., 2001). Consistent with this notion, definitive EMP population was recently characterised in E8.5 YS. These cells were distinct from primitive haematopoiesis or immature HSCs and seeded fetal liver (FL) by E11.5 (McGrath et al., 2015). Furthermore, haematopoietic clusters with the potential to generate long-term HSCs (LT-HSCs) were recently described within the umbilical and vitelline arteries (U/VA) and even head vasculature of E11.5

embryos (Gordon-Keylock et al., 2013; Li et al., 2012) (Fig 1.1). HSCs were also found in the mouse placenta (PL) (Gekas et al., 2005). Follow up experiments on the *Ncx*$^{-/-}$ embryos confirmed that in the absence of circulation PL was the source of cells with multilineage potential *in vitro* (Rhodes et al., 2008).

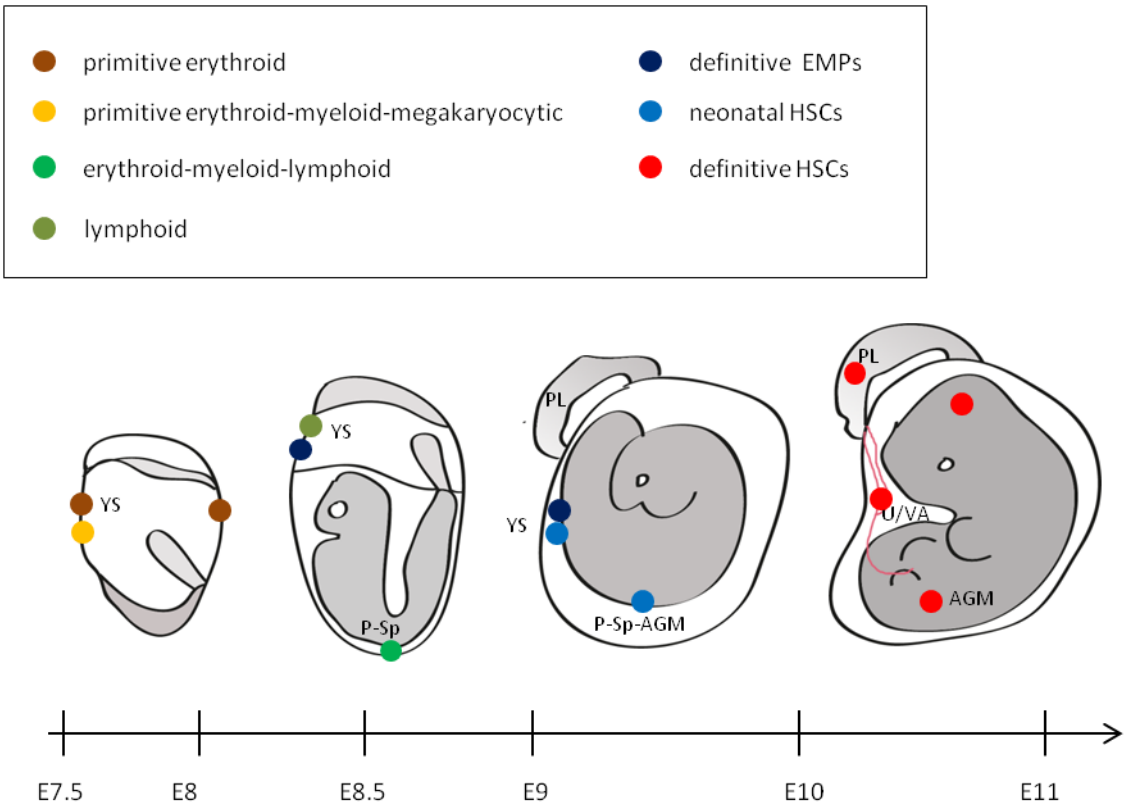


Figure 1.1. Primitive and definitive haematopoiesis in the mouse conceptus.

During the onset of haematopoiesis cell populations with a different potential occur in the temporal and spatial separation. The diagram shows localizations of haematopoietic cells with varying degrees of specialization in the fetus between E7.5, ie at the beginning of primitive haematopoiesis, and E10.5 when adult repopulating definitive HSCs appear (see text for details). AGM - aorta-gonad-mesonephros region, EMPs – erythro-myeloid progenitors; HSCs – haematopoietic stem cells; PL – placenta, P-Sp - para-aortic splanchnopleura, U/VA – umbilical/ vitteline artery.

Similar studies of the human embryonic haematopoiesis have proved to be more challenging due to the limited availability of the early tissue. Even though the particular mechanisms of human HSC development are still poorly understood, it has been established that the major biological processes of embryonic blood specification are conserved between mammalian species. In this respect the first haematopoietic progenitors with limited myeloid potential

originate from the third week human YS. Subsequently multipotent lymphomyeloid stem cells were detected in the P-Sp and AGM between week 27 and 40 (Tavian et al., 2001). Later Ivanovs and colleagues confirmed that highly potent HSCs emerge in the AGM from the ventral wall of the DA to subsequently colonize YS, FL and PL (Ivanovs et al., 2014; Ivanovs et al., 2011).

1.1.2. The cellular origins of haematopoietic cells

When examined on sagittal sections, the site of primitive haematopoiesis in the murine E8 YS is apparent as isolated islands of primitive blood cells surrounded by an endothelial covering. This observation, together with notable similarities in the molecular regulation of haematopoietic and endothelial cells, supported the hypothesis that each 'blood island' is clonally derived from a single bi-potent progenitor termed later as haemangioblast (Sabin et al., 1920; Murray et al., 1932). This notion has been disproved by Ueno and colleagues, who showed that the spatial association of these two lineages within the YS is not due to their generation from single or even multiple clonal precursors (Ueno and Weissman, 2006). However, the existence of the haemangioblast was supported by the identification of a precursor able to give rise to blast-colony containing endothelial and haematopoietic potential following the *in vitro* differentiation of embryonic stem cells (ESCs) (Choi et al., 1998). Only later was it discovered that *in vivo* the haemangioblast is not localized to the YS, as originally thought, but in the gastrulating posterior mesoderm (Huber et al., 2004). The current model suggests that haemangioblasts migrate from the posterior primitive streak and undergo specification to separate haematopoietic and endothelial lineages before entering the YS (Huber et al., 2004; Kinder et al., 1999; Padron-Barthe et al., 2014). However, the recent study performed in early frog embryo proposes that haemangioblast can reside in the YS as a bi-potential competence state. Meyers and Krieg showed that inhibition of bone morphogenic protein 4 (Bmp4) signalling, or disruption of primitive erythroid programme in ventral blood island, lead to the generation of endothelial cells within this site (Myers and Krieg, 2013). This finding suggests that haemangioblast can commit to either haematopoietic or endothelial lineage

depending on extrinsic cues. Interestingly, the equivalent of a bi-potential precursor has been found in the E10.5-11.5 AGM (Hamaguchi et al., 1999; Hara et al., 1999). Notably, cells isolated on the basis of the endothelial Podocalyxin-like protein 1 (PCLP1) marker exhibited long-term repopulating potential when injected into irradiated mice (Hara et al., 1999).

A close spatial proximity of haematopoietic cells and endothelium was also observed in the AGM region. The first evidence of haematopoietic activity in the DA is the generation of tight cellular clusters (Dieterlen-Lievre et al., 2006). Cells within this intermediate population express haematopoietic as well as endothelial markers and are enriched in pre-HSCs and definitive HSCs (Garcia-Porrero et al., 1998; North et al., 2002). Pre-HSCs are defined as direct precursors of mature HSCs. Type I pre-HSCs express vascular endothelial cadherin (VE-cadherin, CDH5) and CD41 on their surface but lack CD45, whereas type II pre-HSCs are defined as $CDH5^{+}CD45^{+}$ cells (Rybtsov et al., 2011). Importantly, only directly transplanted E11.5 AGM type II pre-HSCs can reconstitute recipient mice, whereas type I pre-HSCs require additional maturation step byre-aggregation with OP9 stroma. These data reveal a linear directionality of HSC development where type I subtype mature into type II pre-HSCs. In the further course of development, after budding off the ventral wall of the aorta and subsequent FL colonization, maturing cells lose the expression of CDH5 and their endothelial potential (Kim et al., 2005; Taoudi et al., 2005). *In vivo* lineage-tracing assays and experiments performed in *in vitro* culture systems supported the existence of haemogenic endothelium (HE) – a highly specialized subpopulation of endothelial cells that differentiate into functional haematopoietic cells in the process of endothelial to haematopoietic transition (EHT) (Boisset et al., 2010; Eilken et al., 2009; Jaffredo et al., 1998; Lancrin et al., 2009). Emerging evidence suggests that similar population of RUNX1-dependent HE giving rise to definitive EMPs is also present in the YS (Padron-Barthe et al., 2014; Yokomizo et al., 2001). Moreover, it was recently discovered that the concept of the HE is not contradictory to the notion of the haemangioblast. In the ESC differentiation system the two precursors seem to be linked through a linear pathway in which the haemangioblast gives rise to HE (Lancrin et al., 2009).

1.1.3. Early adult haematopoiesis

Following maturation and expansion in the FL, blood cells will then seed the bone marrow (BM) where they will assure lifelong adult haematopoiesis. A population of rare self-renewing HSCs will subsequently serve as a source of all mature blood cells through a process of gradual restriction of uncommitted progenitors. The hierarchy of haematopoietic differentiation has been the subject of multiple studies and has become a paradigm for the investigation of directed lineage commitment undertaken by stem cells and their downstream precursors (Fig 1.2). In adult mammals LT-HSCs reside at the apex of the blood hierarchy. They are characterised by robust self-renewal alongside long-lived engraftment capabilities (over 16 weeks) and the capacity to repopulate secondary recipients (Benveniste et al., 2010). The turnover of LT-HSCs in the BM is maintained at a very low level as on average one in 110 cells leave the niche per day (Busch et al., 2015). Subsequent asymmetrical division leads to the appearance of the intermediate-term (IT) and short-term (ST) HSCs with a finite self-renewal capability (Benveniste et al., 2010; Osawa et al., 1996). The downstream events include acquisition of myeloid, megakaryocytic, erythroid or lymphoid fate through a stepwise commitment of presumed multipotent progenitors (MPP). The first model of HSC differentiation proposed by Weissman and colleagues over a decade ago suggested bidirectional segregation of MPPs into common lymphoid and myeloid precursor (CLP and CMP respectively) (Akashi et al., 2000; Kondo et al., 1997). The latter subsequently give rise to megakaryocyte/erythrocyte (MEP) or granulocyte/macrophage (GMP) progenitors. This classical model remains applicable to early human haematopoiesis (reviewed by Chotinantakul and Leeanansaksiri, 2012), however it has been recently challenged in the murine system. More specifically Adolfsson and colleagues identified a lymphoid primed MPP (LMPP) subpopulation that gives rise to GMPs and CLPs but lack megakaryocytic/erythroid potential (Adolfsson et al., 2005). In support of this finding immunophenotypic LMPPs, presenting high levels of Fms-Related Tyrosine Kinase 3 (FLT3), were shown not to express the pro-megakaryocytic von Willebrand factor (vWF) (Mansson et al., 2007). However, a series of clonal and lineage tracing studies, the latter with the use of a dual-reporter

Flt3^{Cre}*Rosa*^{mTmG} switch mouse line, suggested that all lineages can derive from common FLT3^{+} MPPs (Boyer et al., 2011; Forsberg et al., 2006). Overall, the existence of LMPPs stands in opposition to the classical view of the immediate and complete separation of myelopoiesis and lymphopoiesis. The contradictory reports on this issue prompt further in depth investigation of the mutual relations between early multipotent blood progenitors.

Notably, the long-established model of stepwise blood differentiation has been recently challenged by the discovery of myeloid-restricted repopulation progenitors (MyRPs). In this study long-term repopulating precursors with differentiation potential toward myeloid, megakaryocytic or megakaryocytic-erythroid lineages directly originated from the asymmetrical division of the most naïve HSCs (Yamamoto et al., 2013). Therefore, the myeloid bypass model proposes that blood lineage commitment may occur in a non-stepwise manner without passing through an MPP stage (Fig 1.2).

Interestingly HSCs, generally defined as self-renewing multipotent cells, proved to be much more heterogeneous than previously thought. As already mentioned, separation into LT-, IT- and ST-HSCs is based on the degree of self-renewal and lifespan. Later it was revealed that HSCs also vary in their differentiation properties. The concept of lymphoid-biased, myeloid-biased and balanced HSCs was introduced on the basis of the skewed ratio of lymphoid to myeloid cells produced upon *in vivo* engraftment (Muller-Sieburg et al., 2004; Muller-Sieburg et al., 2002) (Fig 1.2). Subsequently, another classification was proposed that also took into consideration the contribution of competitor cells. Single cell transplantation assays allowed the distinction between α (myeloid), β (balanced) and γ/δ (lymphoid) skewed HSCs (Dykstra et al., 2007) (Fig 1.2). The following cell separation studies showed that this functional classification can be to the extent reflected at the immunophenotypical level. Lineage negative (LIN^{-})SCA1^{+}cKIT^{+}CD34^{-} cells sorted on the basis of differential CD150 surface expression exhibited variable engraftment capabilities (Morita et al., 2010). α–HSCs were represented by CD150^{hi} fraction, whereas CD150^{med} and CD150^{-} cells were enriched in balanced and lymphoid-biased subtypes respectively. Interestingly, in addition to the three-way classification of HSCs established over a decade ago a recent study revealed the presence of a new platelet-primed HSC subtype (Sanjuan-Pla et al., 2013).

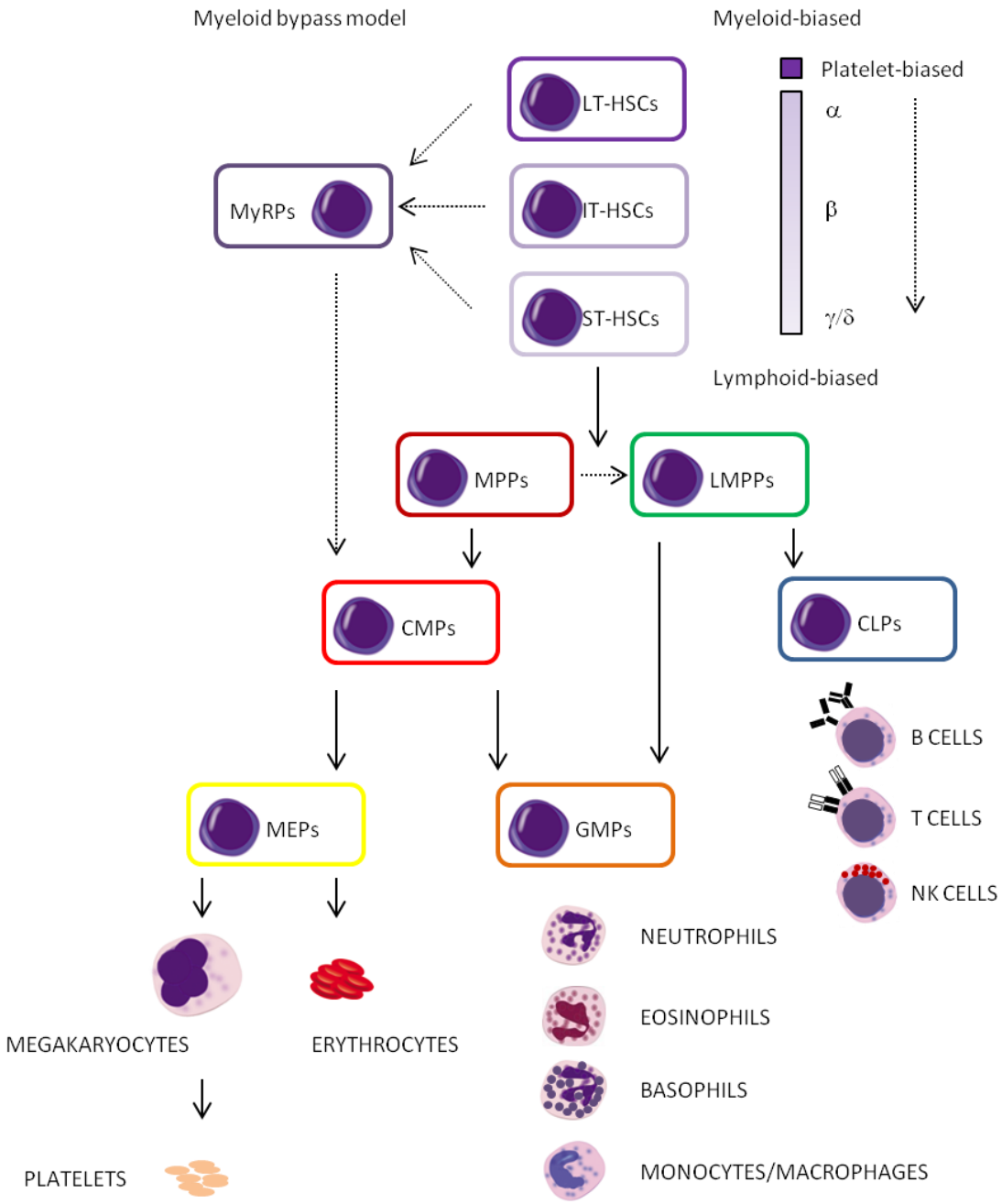


Figure 1.2. Haematopoietic hierarchy.

Blood effector cells are generated in the course of a stepwise specification of HSPCs. This scheme gathers current knowledge on differentiation potential of blood precursors as well as mutual relations between different compartments of haematopoietic hierarchy (see text for details). CMP – common myeloid progenitor, CLP – common lymphoid progenitor, GMP – granulocyte/macrophage progenitor, LMPP – lymphomyeloid progenitor, LT-, IT-, ST-HSC – long-term, intermediate-term, short-term haematopoietic stem cell, MEP – megakaryocyte/erythrocyte progenitor, MPP – multipotent progenitor, MyRP – myeloid-restricted progenitor.

These cells marked by a $LIN^-SCA1^+cKIT^+CD150^+CD48^-$ $CD34^-vWF^+$ immunophenotype generated a high platelet output alongside robust myeloid repopulation when injected into irradiated mice. Moreover, upon division platelet-biased HSCs could generate lymphoid skewed HSCs which position them at the apex of HSC hierarchy (Fig 1.2).

1.2. Transcriptional landscape of blood specification and maturation

1.2.1. Regulation of embryonic haematopoietic development

The key methods to unravel mechanisms governing blood specification are based on gene targeting in the ESC system and knock-out studies. The latter enabled the identification of FLK1 (Vascular endothelial growth factor (VEGF) receptor 2) as the marker of early blood and endothelial progenitors. *Flk*$^{-/-}$ embryos die at E8.5 due to the severe disruption of endothelial and haematopoietic development manifested by the lack of YS blood islands (Shalaby et al., 1995). Further studies of *Flk1* mutations in zebrafish allowed selecting several candidate transcriptional regulators of early blood development (Sumanas et al., 2005). Among screened transcription factors (TFs), of particular interest was the amniotes' homologue of mammalian Ets variant 2 (ETV2). Similar to the *Flk*$^{-/-}$ model the *Etv2*$^{-/-}$ mouse embryos fail to develop YS blood islands (Koyano-Nakagawa et al., 2012). More in depth studies revealed that the expression and functions of ETV2 are limited to the early embryo (before E9.5) (Wareing et al., 2012) and are particularly associated with the specification of FLK1 positive platelet-derived growth factor receptor α (PDGFRα) negative haemogenic mesoderm (Kataoka et al., 2011; Liu et al., 2015). Chromatin immunoprecipitation coupled with high-throughput sequencing (ChIP-seq) using the *in vitro* ESC differentiation system revealed that a haemogenic programme is induced by ETV2 through the activation of its downstream targets including other Ets factors, Stem Cell Leukaemia (SCL), LIM domain only 2 (LMO2) and GATA binding protein 2 (GATA2) (Liu et al., 2015). Indeed SCL and GATA2 were previously reported to positively modulate haemangioblast development, individually or in concert with ETV2 (Ismailoglu et al., 2008; Liu et al., 2013; Lugus et al., 2007; Shi et al., 2014). Importantly, *Scl*

but not *Gata2* or Friend leukaemia integration 1 (*Fli1*) re-expression in *Etv2*$^{-/-}$ cells fully rescued haematopoiesis (Wareing et al., 2012). Beyond E9.5, upon downregulation of ETV2, the expression of *Scl* and *Gata2* is still maintained by FLI1 due to the process of ETS switching on target loci (Liu et al., 2015). Established in this way SCL-GATA2-FLI1 regulatory circuit has been shown to orchestrate the further transition toward HE and HSCs and is active throughout remaining early blood development (Gottgens et al., 2002; Pimanda et al., 2007).

In particular SCL was found to be essential during the subsequent specification of HE from haemangioblasts. In the ESC system *Scl*$^{-/-}$ FLK1^{+} cells fail to generate adherent HE clusters and consequently CD41^{+} haematopoietic cells (Lancrin et al., 2009). Notably this requirement for SCL falls into a narrow time window of developmental events. *Scl*$^{-/-}$ mutants die early in embryogenesis due to the lack of haematopoietic lineages consistent with the role of SCL in the specification of haemogenic mesoderm (Ismailoglu et al., 2008; Shivdasani et al., 1995). Furthermore, rescue of *in vitro* primitive and definitive haematopoiesis by SCL was only effective before the onset of *Cdh5* and *Tie2* expression in HE (Endoh et al., 2002; Schlaeger et al., 2005). These studies suggest that even though *Scl* expression is detectable throughout blood ontogeny in all major haematopoietic sites, it is only required in a non-redundant fashion within a short time span.

In the course of EHT HE gives rise to blood cells. *In vivo* definitive HSCs are enriched in haematopoietic clusters generated in the ventral wall of the DA (North et al., 2002). This particular stage of blood specification was found to be disrupted in the *Runx1*-deleted embryos (North et al., 1999). Moreover, *Runx1*$^{-/-}$ ESCs were only capable of generating tight adherent clusters of CD31^{+} cells without subsequent production of CD41^{+} haematopoietic progenitors (Lancrin et al., 2009). These studies allowed the identification of RUNX1 as a key regulator of EHT. Using a conditional knock-out mouse model Chen and colleagues confirmed the precise window of RUNX1 requirement. Haematopoietic development was only affected when *Runx1* was inactivated in the CDH5^{+} endothelial cells of the YS and DA (Chen et al., 2009). Conversely, ablation of *Runx1* in already formed blood cells showed no adverse effect on fetal haematopoiesis. More precisely *Runx1* deficiency in the AGM does not affect

the formation of type I pre-HSCs ($CDH5^+CD41^+CD45^-$) but abrogates transition towards type II pre-HSCs ($CDH5^+CD41^+CD45^+$) (Liakhovitskaia et al., 2014). Interestingly, RUNX1 seems to have dose dependent role in haematopoiesis as *Runx1*$^{+/-}$ embryos are viable but present alterations in the kinetic of emergence of blood cells in the YS and AGM (Cai et al., 2000). Similarly *Gata2* haploinsufficient embryos exhibit decreased numbers and activity of HSCs in the AGM, whereas the *Gata2* knock-out is lethal at E10 (Ling et al., 2004; Tsai et al., 1994). Furthermore, *Runx1*$^{+/-}$*Gata2*$^{+/-}$ mutants display severe haematopoietic defects suggesting functional interplay between these two powerful blood regulators (Wilson et al., 2010a). Indeed *Gata2* is expressed in the endothelial lining of the DA at E10.5 and has been implicated in the production and expansion of HSCs in the AGM (Robert-Moreno et al., 2005).

Mechanistically EHT has been described as a stepwise process of losing endothelial identity and the concomitant acquisition of the morphology and transcriptional profile of blood cells (Eilken et al., 2009; Lancrin et al., 2009; Swiers et al., 2013). Accordingly, RUNX1-induced transcriptional changes during EHT clearly include the downregulation of endothelial genes and upregulation of a haematopoietic programme (Lichtinger et al., 2012; Tanaka et al., 2012b). Recently, two transcriptional repressors, Growth factor independent 1 and 1b (GFI1 and GFI1b), have been recognized as direct targets of RUNX1 and mediators of the loss of endothelial identity during EHT (Lancrin et al., 2012). Enforced *Gfis* expression rescues the formation of round cells with blood morphologies by *Runx1*$^{-/-}$ HE but not the subsequent acquisition of haematopoietic fate (Lancrin et al., 2012). Furthermore, the double knock-out study confirmed complete lack of AGM clusters in the embryos deprived of GFI1 and GFI1b activities (Thambyrajah et al., *submitted*).

1.2.2. Self-renewal versus commitment

In the adult, HSCs are a unique population capable of balancing self-renewal and differentiation in order to provide a virtually unexhausted source of blood cells throughout life. Many TFs, including Ets variant 6 (ETV6), CCAAT enhancer binding protein α (CEBPα) or Homeobox B4 (HOXB4) have been implicated in the regulation of survival, quiescence and cycling status of

HSCs (Hock et al., 2004b; Sauvageau et al., 1995; Ye et al., 2013). A key role in the maintenance of HSC homeostasis was also attributed to GATA2 (Fig 1.3). Indeed, haploinsufficiency experiments revealed that in the context of a reduced GATA2 dose, stem cells have a proliferative and survival disadvantage (Ling et al., 2004; Rodrigues et al., 2005). Furthermore, *in silico* studies of transcriptional statuses of self-renewing cells revealed that fluctuations in the messenger RNA (mRNA) levels of *Gata2* have the strongest impact on commitment frequency (Teles et al., 2013). Altogether these data suggest that GATA2 dynamics constitutes a major switch between the quiescence/ proliferation and self-renewal/differentiation statuses of HSCs. This role is mostly facilitated by the regulatory circuit of GATA2-SCL-FLI1. Using mathematical modelling Narula and colleagues established that the triad exists in two stable states that can be switched by physiologically relevant signals (Narula et al., 2010). At the onset of haematopoiesis Notch and Bmp4 signalling change the status of the circuit from low to high expression thereby facilitating blood specification. Later in response to GATA binding protein 1 (GATA1), which is a known regulator of erythroid and megakaryocytic commitment (Pimkin et al., 2014), the triad is switched to the OFF state which in turn promotes differentiation. Recently, the single-cell gene expression analysis of haematopoietic stem and progenitor cells (HSPCs) identified an additional regulatory circuit connecting GATA2, GFI1 and GFI1b (Moignard et al., 2013). Functions of these two paralogues are widely spread within haematopoietic system. The loss of *Gfi1* is manifested by neutropenia and lymphopenia accompanied by decreased repopulation activity of HSPC compartment arising from enhanced apoptosis (Hock et al., 2004a; Karsunky et al., 2002; Khandanpour et al., 2011). Deletion of Gfi1b cause embryonic lethality due to severe disruptions in erythropoiesis (Saleque et al., 2002). The conditional knock-out of *GFi1b* revealed its role in the maintenance of HSC dormancy (Khandanpour et al., 2010) (Fig 1.3).

Particularly interesting, although still not fully understood, is the role of HOXB4 in the expansion of the HSC compartment. Loss of *Hoxb4* had no prominent effect on haematopoiesis although its overexpression enhanced the self-renewal of fully functional HSCs (Brun et al., 2004; Sauvageau et al., 1995). The same treatment conferred a long-term repopulating capacity to the YS and

ESC-derived haematopoietic precursors (Kyba et al., 2002). It has been proposed that the observed effect is attributed to the increased responsiveness of HSCs to multiple extrinsic stimuli including Wnt and Notch signalling (Schiedlmeier et al., 2007). As HOXB4-induced expansion does not lead to malignancies and cannot be instilled on human B-leukemic cells, it was proposed that the observed process may have clinical applications in the future (Fournier et al., 2014).

Current theories on how an individual HSC decides between self-renewal and commitment span from instructive to noise-driven mechanisms. Both hypotheses are inextricably linked with the phenomenon of lineage priming established over a decade ago. This concept was based on a finding that low expression levels of globin and myeloperoxidase could be detected in uncommitted multipotential progenitors (Hu et al., 1997). This suggested that HSCs express the spectrum of lineage-specific TFs reflecting their multipotent capabilities and the 'gene expression noise' can further translate toward fluctuations of protein levels and subsequent lineage choice. The question remained whether the decision of commitment constitutes a gradual process or a discreet transition. To address this issue two studies have been conducted using Erythroid myeloid lymphoid (EML) haematopoietic cells line representing a mixture of cells with varying levels of SCA1 surface expression. The $SCA1^{lo}$ population is erythroid-biased, whereas $SCA1^{hi}$ marks myeloid-skewed cells. Both sorted populations can recreate the original spectrum of SCA1 presentation upon culture, which is associated with their self-renewal capabilities. The population-based study of EML cellular fractions suggested that lineage commitment is driven not by individual changes in specific lineage-affiliated genes but rather slow fluctuations in transcriptome-wide noise resulting in stochastic priming of multilineage cells (Chang et al., 2008). Conversely, single-cell studies in the same model led to an alternative conclusion that discrete and uncoordinated events prompt cells to exit from the self-renewal state. According to Pina and colleagues self-renewing and committed states do not coexist in a single cell (Pina et al., 2012). Instead, precursors experience independent activation of individual key modulators translating into cell-to-cell variation in the expression profile of a set of lineage-commitment genes. The coordinated lineage programme is instilled only downstream of these initial

changes. This theory is in line with the recent advent of single cell expression data unravelling significant levels of heterogeneity in the transcriptional profiles of HSCs that cannot be accounted for by stochastic noise (Glotzbach et al., 2011). Moreover, the isolation of HSC subsets with discrete predetermined lineage options proved that HSCs are indeed not unbiased in their differentiation potential toward blood lineages. For instance the platelet-biased HSCs are distinctly marked by *vWf* expression and poised to generate predominantly megakaryocytic outputs (Sanjuan-Pla et al., 2013). The fact that sporadic low-level expression of lineage-affiliated genes can exist in self-renewing cells capable of multilineage engraftment suggest, that priming can be hardwired into HSPC regulatory network (Pina et al., 2012). Instead of presenting a fluent stem cell continuum, HSCs can be therefore constrained to enter multiple independent differentiation trajectories originating from the stochastic modulation of key regulatory factors.

1.2.3. Lineage specification

Upon exiting their multilineage state, progenitors begin a stepwise specialization toward committed blood cells. According to the classical model of blood differentiation, genes associated with a single path of development are upregulated in the course of lineage commitment, whereas pathways responsible for the conduct of alternative fates are silenced leading to the appearance of more restricted haematopoietic precursors. The coherent lineage transcriptional programmes are implemented in cells upon passing through specific bifurcation checkpoints. The knock-out and knock-down studies proved invaluable in identifying TFs critical for those lineage choices and/or specification (Fig 1.3). For instance role of GFI1 in myeloid differentiation was revealed after *GFI1*$^{-/-}$ mice displayed severe neutropenia and accumulation of immature monocytes in BM and blood (Karsunky et al., 2002). Furthermore, *Gfi1b* deficiency led to embryonic lethality due to severe disruption of erythroid development, therefore defining this TF as a key regulator of erythropoiesis (Saleque et al., 2002). Such clear results, as well as lineage conversion experiments, have reinforced the long-standing model that haematopoietic differentiation is governed by a limited number of TFs that can single-handedly

drive blood specification. However, genome-wide binding techniques start to unravel the highly complicated nature of these processes. It is now clear that fate specification is more reliant on the function of multiprotein complexes generated by collaborating TFs.

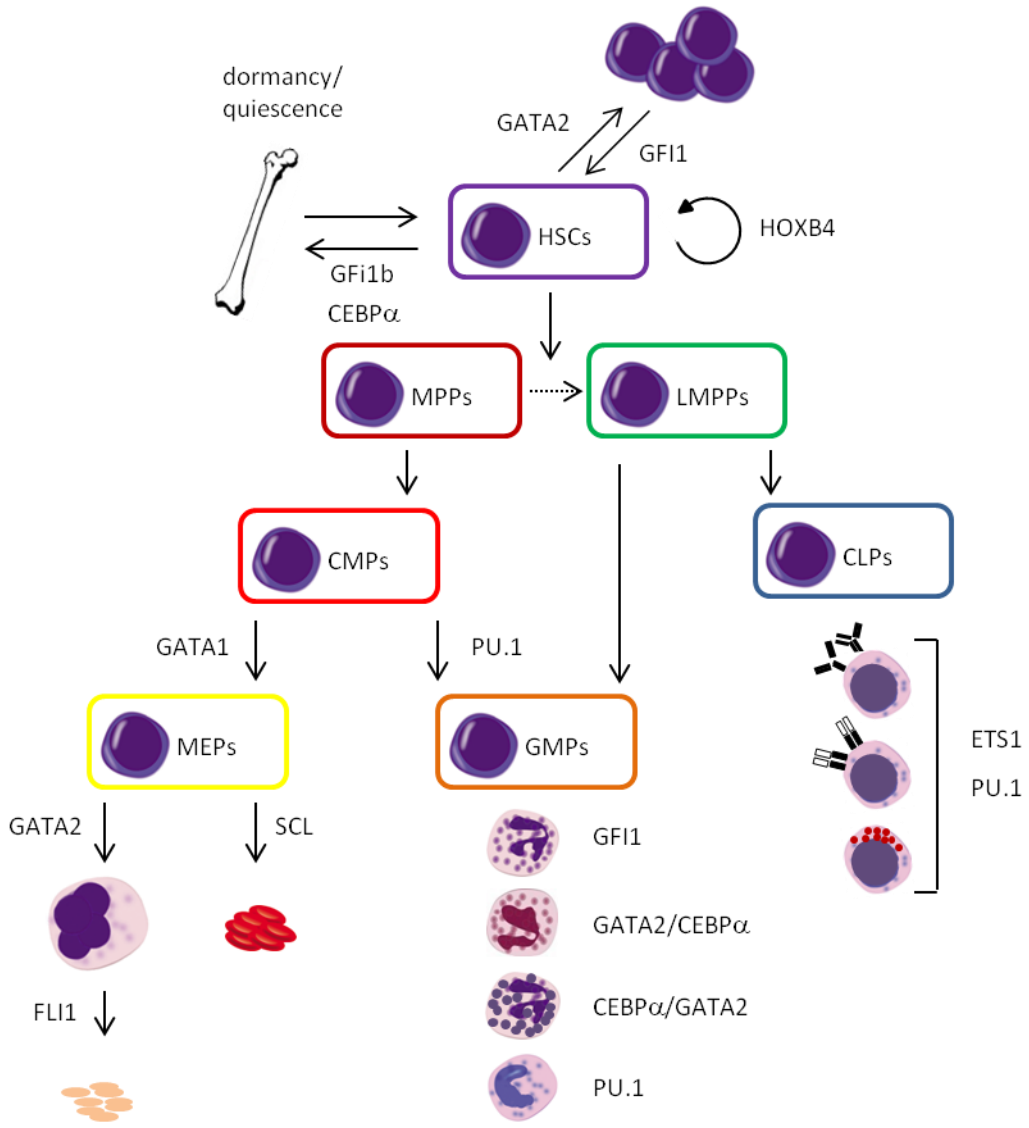


Figure 1.3. TFs governing hierarchical differentiation of haematopoietic cells.

The process of commitment of blood cells is mediated by lineage-specific TFs. The timing and degree of their expression is crucial for the specification to the given fate. The diagram comprises the main factors responsible for maintenance of HSC compartment and differentiation of progenitors to a specific lineage. However it should be noted that in fact very complex network of multiple molecules is crucial for the emergence of terminally differentiated cells. Colours and abbreviations as in Figure 1.2.

Some of these modules are restricted to specific lineages like the LMO2/SCL/LIM domain-binding protein 1 (LDB1)/E47/GATA1/ complex that occupy loci of erythroid genes (Wadman et al., 1997). Others can be reused at different levels of the haematopoietic hierarchy. The master regulatory complex of haematopoietic TFs was first recognized in the Haematopoietic precursor cell-7 (HPC-7) line. Genome-wide binding data identified combinatorial interactions between a heptad of factors (lymphoblastic leukaemia associated haematopoiesis regulator 1 (LYL1), SCL, LMO2, GATA2, RUNX1, ERG and FLI1) as key regulatory cues in maintaining haematopoietic fate (Wilson et al., 2010a). This result was later extended to primary human HSPCs (Beck D 2013). Interestingly, the same transcriptional combinations were also identified as drivers of megakaryocytic and mast cell differentiation (Calero-Nieto et al., 2014; Pimkin et al., 2014).

A significant proportion of key regulators at various stages of haematopoiesis belong to the Ets transcription factor family. ETS1 is mainly required for the survival and maturation of T- and B-cells as well as the development of functional natural killer cells (Barton et al., 1998; Bories et al., 1995) Additionally FLI1 has been recognized as indispensable for the maturation of the megakaryocytic lineage (Hart et al., 2000). Another extensively investigated member of the Ets family, which appears to be necessary for both myeloid and lymphoid development, is Spleen focus forming virus proviral integration oncogene spi1 (PU.1). In the absence of the functional factor mouse embryos lack all granulocyte/macrophage lineages as well as B-cells (Scott et al., 1994). Furthermore, development of T-cells is also severely impaired, whereas the numbers of megakaryocytes and erythroid progenitors remain unchanged. These data, supported by other *in vivo* experiments performed on adult BM, established that PU.1 is a key regulator of the development and differentiation of lymphomyeloid progenitors (Dakic et al., 2005; Iwasaki et al., 2005). More recent mechanistic studies revealed that in multilineage precursors PU.1 genomic binding sites localize in the vicinity of motifs of other lineage-specific TFs, priming them for occupancy by myeloid (CCAAT enhancer binding protein β (CEBPβ)/Activator Protein 1 (AP1)) or lymphoid (E2A/ Octamer-binding transcription factor 2 (OCT2)) fate determinants (Heinz et al., 2010). Subsequent collaboration with the bound

lineage-affiliated regulators leads to deposition of a monomethylation mark on lysine four of histone 3 (H3K4me) and stepwise acquisition of either macrophagic or B-lymphoid fate.

The interplay between PU.1 and GATA1 at the CMP stage is widely studied as a paradigm of how TFs control lineage choice. Overexpression and knock-down studies have shown that the two factors act antagonistically by repression of each other's expression or direct interaction at the protein level and mutual inhibition of transcriptional activity (Nerlov et al., 2000; Stopka et al., 2005; Zhang et al., 2000). Functional studies in zebrafish model confirmed that upregulation of *pu.1* in CMPs leads to the formation of the granulocyte/macrophage lineages, whereas activation of *gata1* results in the commitment to megakaryocytic and erythroid fates (Rhodes et al., 2005). Subsequently the phenomenon of GATA switching was also investigated in this context. It was discovered that GATA2 occupancy facilitates intermediate expression of *Pu.1* and its downstream myeloid target genes which maintains undifferentiated state (Chou et al., 2009). With replacement by GATA1 the *Pu.1* expression is extinguished and erythro/megakaryocytic fate can be determined. Conversely, the removal of GATA2 directs cells to become macrophages. Another example of cross-antagonism includes the relationship of PU.1 and GFI1 at the GMP stage, in respect to monocyte and neutrophil choice (Dahl et al., 2007). Interestingly, the order of expression and its duration also affects the fate of uncommitted progenitors. Upregulation of *Gata2* in GMPs leads to the generation of mast cell progenitors, whereas activation of CEBPα specifies neutrophil lineage commitment (Iwasaki et al., 2006). However, if *Gata2* expression is preceded by CEBPα activity solely eosinophils are formed, whereas with the reversion of this order specification is restricted to basophils.

The hypothesis that fluctuating TF networks lie at the core of fate decisions is a long-standing paradigm that has been recently challenged by continuous observation of single-cell dynamics. Schroeder and colleagues employed live imaging of double reporter *Gata1* and *Pu.1* strain to follow expression of this pair of TFs throughout specification of a single HSC (personal communication). In line with previous studies on the role of PU.1 in the commitment of HSCs, the factor was expressed in the majority of the starting population (Fukuchi et al., 2008; Iwasaki et al., 2005). However, preliminary

data suggest that throughout the timeline of subsequent fate decisions there is no concurrency between GATA1 and PU.1. More particularly, if the cell was to become MEP, *Pu.1* expression was downregulated before the occurrence of *Gata1*. Conversely, if the cell was predestined to undertake GMP fate *Gata1* was not expressed throughout the commitment process. Altogether, this group could not obtain biological data supporting the role of cross-antagonism between PU.1 and GATA1 in controlling binary lineage choices. This finding also poses the question whether TFs are truly decision-makers during lineage choice or only executors. In this context, it was recently shown that macrophage colony-stimulating factor (M-CSF) can directly induce myeloid fate in HSCs through the activation of PU.1, suggesting that extrinsic cues can have a determining influence on commitment (Mossadegh-Keller et al., 2013).

1.3. Reprogramming

1.3.1. Induced pluripotency

In 2006 Takahashi and Yamanaka opened new avenues that revolutionised the field of reprogramming. At this point, it was already established that the differentiation of cells is not irreversible. The seminal work by John Gurdon demonstrated that somatic cell from *Xenopus laevis* can be reprogrammed to embryonic state by factors present in the oocyte, which led to the recreation of adult organism (Gurdon, 1962). This technique termed somatic cell nuclear transfer (SCNT) was later employed to other vertebrates including sheep (Wilmut et al., 1997), mice (Eggan et al., 2004; Hochedlinger and Jaenisch, 2002) and dogs (Lee et al., 2005). In the late 70s another approach of cell fusion between embryonal carcinoma cells and thymocytes also led to the erasure of somatic identity in the generated hybrids (Miller and Ruddle, 1977). This methodology was later employed to investigate the molecular mechanism of acquisition of pluripotency by differentiated cells, unravelling the crucial role of Octamer-binding transcription factor 4 (OCT4) in the process (Pereira et al., 2008). Altogether, the early studies proved that cells retain a certain level of plasticity and restrictions that occur during specification are most likely transient and reversible. Subsequently, Takahashi and Yamanaka demonstrated that

cellular plasticity can be also regained by the simple forced expression of TFs, which led to the generation of induced pluripotent stem cells (iPSCs) (Takahashi and Yamanaka, 2006) (Fig 1.4). The four ESC-related TFs used by authors included OCT4, Krueppel-like factor 4 (KLF4), SRY (Sex determining Region Y)-box 2 (SOX2) and cMYC. Generated cells had the ability to form teratomas when transplanted into immunodeficient recipients and contributed to three germ layers of the developing embryo upon injection into blastocysts. However, it is noteworthy that the original iPSCs were missing a complete pluripotency genetic landscape and methylation pattern as well as failing to generate chimeras and contribute to the germline. Subsequent implementation of improved selection procedures, based on the activation of the *Nanog* promoter, led to the generation of cells with greater similarity to ESCs (Maherali et al., 2007; Okita et al., 2007).

Successful reprogramming has been achieved for cells originating from many species including humans (Takahashi et al., 2007), rats (Li et al., 2009), pigs (Esteban et al., 2009) and rhesus monkeys (Liu et al., 2008) (Esteban et al., 2009; Li et al., 2009; Liu et al., 2008; Takahashi et al., 2007). Moreover, despite differential efficiency and kinetics, iPSCs have also been derived from different somatic cell types including melanocytes (Utikal et al., 2009a), keratinocytes (Aasen et al., 2008), hepatocytes (Liu et al., 2010), B- and T-cells (Eminli et al., 2009) as well as neural stem cells (Kim et al., 2008). Specific time for reprogramming, method of gene delivery and combinations of TFs employed are optimal for each cell type. Although the Yamanaka factors (OSKM standing for OCT4, SOX2, KLF4 and cMYC) are the most frequently used, OCT4, SOX2, NANOG and LIN28 were also shown to reprogram human somatic cells to iPSCs (Yu et al., 2007) In subsequent studies cMYC, LIN28 and NANOG proved to be dispensable, although their removal significantly decreased the yield of transition (Wernig et al., 2008; Yu et al., 2007). Particular efforts were aimed to replace oncogene cMYC in a reprogramming setting. One interesting substitute is poly ADP ribose polymerase 1 (PARP1). This DNA repair-associated enzyme is a direct target of cMYC therefore it efficiently compensated for the loss of TF in the reprogramming protocol (Chiou et al., 2013). Notably, the efficiency of iPSC generation by OCT4 and SOX2 in

conjunction with PARP1 was even higher than that of OCT4, SOX2 and KLF4 combination.

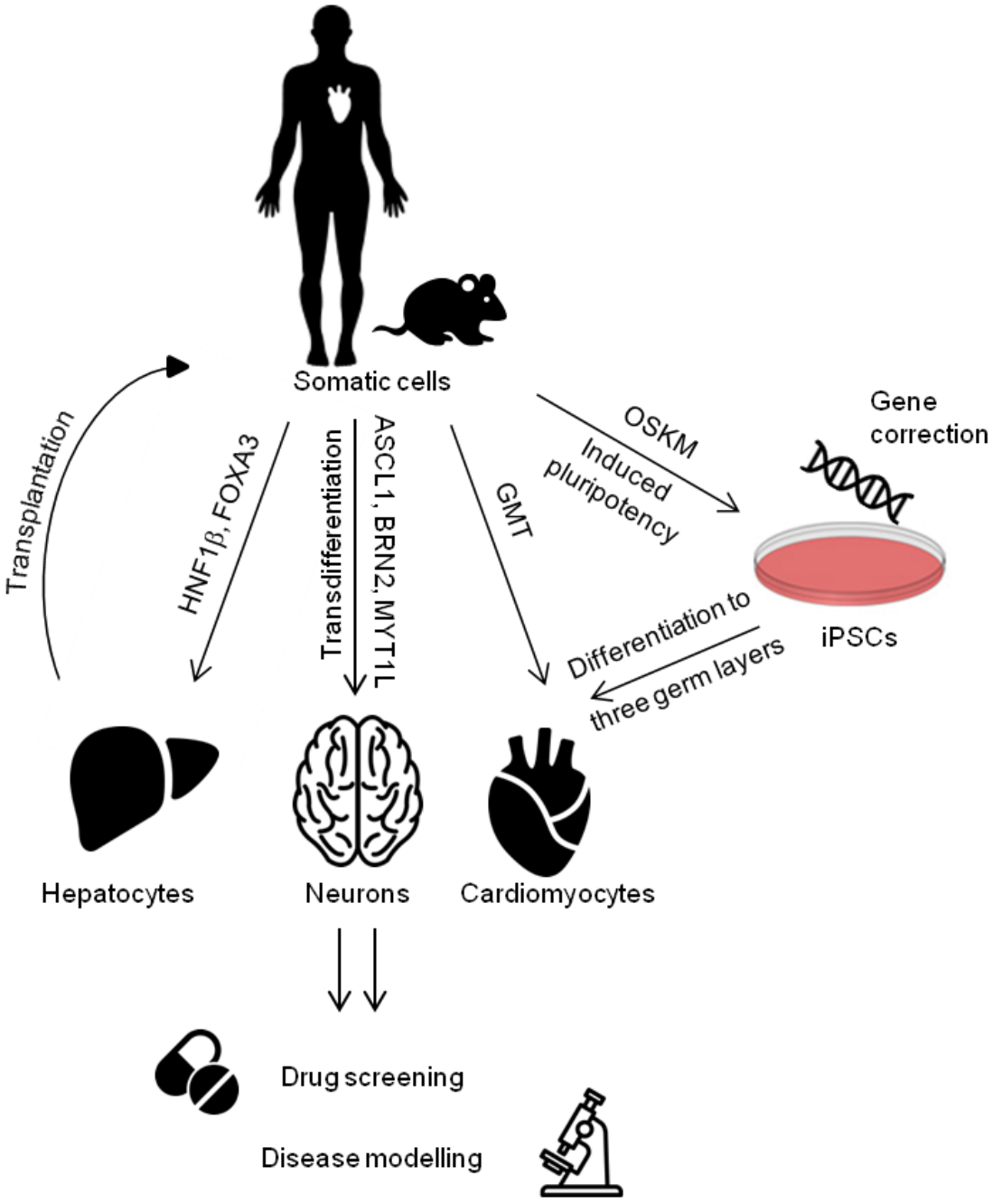


Figure 1.4. Induced cell fate change.

Patient-derived cells for regenerative medicine, drug discovery purposes and disease modeling can be generated either by differentiation of iPSCs or by direct reprogramming of primary cells. Induced pluripotency and induced transdifferentiation are achieved by ectopic expression of pluripotency genes or lineage-instructive TFs respectively. OSKM – OCT4, SOX2, KLF4, cMYC, GMT – GATA4, MEF2c, TBX5

In normal development SOX2 and KLF4 are both engaged in a molecular circuit of pluripotency as well as inhibition of differentiation. As they are also expressed endogenously after the embryonic stage, they can be omitted in reprogramming protocols under carefully selected cellular conditions as was shown for neural stem cells and melanocytes (Chiou et al., 2013; Kim et al., 2008; Utikal et al., 2009a). Furthermore, KLF4 can be substituted for by the orphan nuclear receptor Esrrb (Feng et al., 2009) or butyrate (Mali et al., 2010), whereas functions of SOX2 can be substituted by the transforming growth factor β (TGFβ) signalling inhibitor RepSox (Ichida et al., 2009). Both SOX2 and KLF4 can be also interchanged by Polycomb complex protein BMI1 (Moon et al., 2011). The latter strategy is particularly promising as BMI1 was recently shown to generate iPSCs in combination with NANOG only (Moon et al., 2013).

Initially, it was presumed that OCT4 is absolutely essential for inducing pluripotency. This notion was consistent with its critical requirement in the establishment and maintenance of pluripotency in embryonic development (Nichols et al., 1998). Notably, introduction of OCT4 alone was sufficient to generate iPSCs from neuronal stem cells, although with low efficiency (Kim et al., 2009b). Nevertheless, even though OCT4 is critical for reprogramming, we now know that it is dispensable. The first identified OCT4 substitute was unrelated orphan nuclear receptor NR5A2 which, together with KLF4, SOX2 and cMYC induced pluripotency in mouse embryonic fibroblasts (MEFs) (Heng et al., 2010). Subsequently, a few other regulators with the same capabilities were defined including deoxygenase TET1 (Gao et al., 2013) and trans-membrane protein E-cadherin (Redmer et al., 2011).

In the last few years, several studies have tried to address the issues of low efficiency, partial reprogramming and insufficient clinical grade of generated iPSCs. The latter was addressed by the development of various transgene delivery methods as an alternative to viral transduction. As a result, non-integrating techniques emerged including episomes, transposons and cationic nanoparticle carriers (Davis et al., 2013; Hu et al., 2011; Khan et al., 2013). Furthermore, the direct transfection of OSKM mRNAs or proteins was employed for the generation of true human iPSCs (Kim et al., 2009a; Warren et al., 2010). Another very interesting approach included reprogramming with the whole cell protein extract obtained from murine ESCs (Cho et al., 2010). All

these methods, however, rely on the delivery of pluripotency factors, whereas the future of reprogramming lies in the discovery of small chemical compounds capable of replacing the OSKM cocktail. Relatively quickly it became possible to generate iPSCs with a combined chemical treatment (valproic acid, glycogen synthase kinase 3 (GSK3) inhibitor, TGFβ signalling inhibitor) and OCT4 only (Li et al., 2011). Importantly, the efficiency of this transition was higher than with OSKM or OCT4, SOX2 and cMYC combinations. More recently, the same group developed chemically induced pluripotent stem cells from MEFs. Using seven small molecule compounds Hou and colleagues reprogrammed MEFs to fully functional iPSCs with an efficiency of 0.2% (Hou et al., 2013). This exciting finding opens new avenues for the generation of clinically applicable cells for regenerative medicine.

1.3.1.1. Applications of iPSC-derived blood cells

Human iPSCs are a promising tool for the treatment and study of diseases as they provide an unlimited source of patient-specific cells that reflect all the genetic features of the individual. One field in which this technology can be utilized are hematologic disorders which include a large number of hereditary and malignant conditions. Importantly, many protocols already optimised for the differentiation of ESCs toward blood could be directly implemented for iPSCs. Along this line, mesoderm specification from iPSCs can be obtained through the generation of three-dimensional structures of embryoid bodies (EBs) or culture on BM-derived OP9 stroma, both combined with stimulation by haematopoietic cytokines. OP9 co-culture was found to recapitulate early embryonic haematopoiesis by enhancing mesoderm specification from pluripotent cells followed by endothelial differentiation and EHT (Endoh et al., 2002; Hashimoto et al., 2007). These methods, despite leading to the generation of mature blood cells in relatively big quantities, do not yield long-term engrafting HSCs. Only recently two groups implemented a teratoma-based technique that allowed for the isolation of human cells with lymphomyeloid potential. In principal, iPSCs injected subcutaneously into NOD-SCID IL2R$\gamma^{-/-}$ (NSG) mice were allowed to form teratomas. Next cells differentiating within these structures were isolated and characterised for their functionality. By using this *in vivo* approach Amabile

and colleagues obtained $CD34^+CD45^+$ cells that supported multilineage reconstitution of primary (8 weeks) and secondary (at low levels) mouse recipients (Amabile et al., 2013). Simultaneously, Suzuki and colleagues proposed the variation of this methodology by co-injecting iPSCs with OP9 cells and providing *in vivo* supplementation of cytokines. Interestingly, in this study LT-HSCs were present not only in the site of teratoma formation but also migrated to the host BM (Suzuki et al., 2013b). Most recently, Elchevia and colleagues attempted to generate inducible HSCs (iHSCs) by the direct reprogramming of iPSCs. Even though the group identified two sets of haematopoietic TFs that endowed pluripotent cells with pan-myeloid (ETV2, GATA2) or erythro-megakaryocytic (GATA2, SCL) features, no accompanying engraftment potential could be detected (Elcheva et al., 2014) (Fig 1.5).

A primary objective of developing the iPSC technology was to generate vast amounts of patient-specific cells for transplantation, disease modelling and drug screening. Albeit on a small scale, these goals have been achieved. So far *in vivo* models mimicking blood disease genotypes and/or phenotypes have been developed for Fanconi anaemia, polycythemia vera or Person syndrome (Cherry et al., 2013; Muller et al., 2012; Ye et al., 2009). The latter is a mitochondrial disease leading to marrow defects. Cherry and colleagues showed that a fraction of iPSCs derived from patients lose the defective mitochondria upon selective culture and differentiate into healthy erythrocytes that could be used for treatment (Cherry et al., 2013). Many other examples of potential therapeutic uses of iPSCs are based on gene therapy. In this context, the first breakthrough has been reached for sickle cell anaemia. Murine blood progenitors differentiated from iPSCs were corrected for human β^S mutation and cured the disease after transplantation into the mouse model of the disorder (Hanna et al., 2007). In the follow up study Zou and colleagues performed the correction of human iPSCs derived from a patient suffering from sickle cell disease due to mutation of both globin alleles (β^S/β^S). The resulting cells remained heterogeneous for β^S allele and therefore could be differentiated to erythrocytes expressing 25 to 40% of wild type (WT) β-globin (Zou et al., 2011a). Gene transfer was also used for corrections of either X-linked chronic granulomatous disease or α-thalassemia. Both studies employed zinc-finger

nuclease mediated site-specific insertion of WT therapeutic transgenes into 'safe harbor' AAVS1 locus. In the first case, a minigene encoding $gp91^{phox}$ subunit of phagocyte nicotinamide adenine dinucleotide phosphate (NADPH) oxidase was introduced into one AAVS1 allele which resulted in iPSC clone differentiating into functional neutrophils (Zou et al., 2011b). In the α-thalassemia study, Chang and colleagues obtained the complete correction of globin chain imbalance in erythrocytes derived from iPSCs with homozygous insertion of therapeutic globin gene (Chang and Bouhassira, 2012).

The iPSC technology could be also used to generate lymphocytes for adoptive cell transfer therapies. This very promising strategy is based on utilising an individual's own immune system to fight cancer or viral infections. However, ex-vivo expansion of patient-derived T-cells is often limited due to exhaustion. Therefore the capabilities of iPSCs for self-renewal while maintaining pluripotency state can be used to facilitate large-scale production of therapeutic T-cells. The T-cell receptor (TCR) rearrangement occurs during T-cell maturation in the thymus. Therefore iPSCs derived from patients' T-lymphocytes (T-iPSCs) will harbour the rearranged TCR genes that should remain antigen specific upon re-differentiation to CD8 cytotoxic T-cells. This methodology has been used simultaneously by two groups to obtain T-cells directed against the melanoma MART-1 epitope (Vizcardo et al., 2013) and nef protein of HIV-1 virus (Nishimura et al., 2013). In both studies highly proliferative iPSC-derived $CD8^+$ T-cells exhibited normal cytolytic activity when stimulated with the relevant target peptide.

Recently Themeli and colleagues combined the T-iPSC methodology described above with a chimeric antigen receptor (CAR) strategy. This approach utilizes the cytotoxic activity of T-cells and high specificity of antibody fragments to recognize cell surface antigens. Importantly, the recognition of epitope by these cells is human leukocyte antigen (HLA)-independent, which alleviates the need for donor selection and laborious engineering of autologous T-cells from each individual patient. Themeli and colleagues first generated T-iPSCs from healthy T-lymphocytes to subsequently transduce them with a CAR specific for CD19, an antigen expressed by malignant B-cells (Themeli et al., 2013). Resulting re-differentiated CAR-expressing T-cells potently inhibited tumour growth in

xenograft model of human Burkitt lymphoma and conferred substantial survival advantage.

In summary, many recent studies contributed to a significant progress in the use of iPSCs for modelling and treatment of blood-related diseases. However, the lack of follow up clinical trials impedes the development of truly useful therapies for patients. The many issues still to be overcome include the residual 'epigenetic memory' of the donor tissue in iPSCs (Vaskova et al., 2013), the low efficiency of the process, the incomplete reprogramming leading to teratoma formation and the absence of robust and selective differentiation protocols. Even though the first clinical trial utilizing iPSCs for the treatment of macular degeneration has recently been initiated in Japan, these rare activities are definitely below the hopes initially associated with this technology. Consequently, alternative approaches to fate conversion have been recently revised.

1.3.2. Induced transdifferentiation

Transdifferentiation, also termed direct reprogramming, is defined as the conversion of one committed cell type into another whilst bypassing a pluripotent state (Fig 1.4). The first report showing that a single factor can activate a lineage-specific programme in a lineage-independent manner was published as early as 1987 (Davis et al., 1987). It described MYOD, a basic helix-loop-helix TF, which when expressed in fibroblasts converted them into myoblasts. Later the same outcome was also observed in neuroblastoma, melanoma, chondroblast, smooth muscle, liver, retinal pigmented epithelia and adipose cell lines (Choi et al., 1990; Weintraub et al., 1989). The discovery of direct reprogramming conflicted with the well-established concept of irreversibility and directivity of differentiation and therefore raised as much interest as disbelief. Scepticism within scientific community was majorly fuelled by early misleading research on contribution of BM-derived cells to the central nervous system (Mezey et al., 2000; Priller et al., 2001). Later it was shown that cell fusion and not direct conversion is responsible for observed formation of alternative cell types from BM cells (Alvarez-Dolado et al., 2003). As a result research on transdifferentiation has been mostly disregarded until recently,

despite the scientific evidence of physiological and pathological examples of this process. One of the most extensively investigated is the endothelial-blood transition during embryonic development (Boisset et al., 2010; Eilken et al., 2009; Lancrin et al., 2009) or the epithelial-mesenchymal transition accompanying tumour formation (Yang and Weinberg, 2008). Since 2006, as an aftermath of iPSC discovery, the number of studies reporting successful direct fate switching between two distinct cell types has risen dramatically (Xu et al., 2015). The initial reports continued to focus on reprogramming mostly within developmentally related lineages. In 2010 Vierbuchen and colleagues reported the first transdifferentiation experiment across germ layers (Vierbuchen et al., 2010). Starting from the pool of 19 TFs with established roles in neuronal development, the group identified Achaete-Scute Family BHLH transcription factor 1 (ASCL1) as single-handedly converting fibroblasts to neuron-like cells. This and other early transdifferentiation studies served as the basis for the establishment of 'pioneer factor' theory. The hypothesis claims the existence of a master regulator, specific for each lineage, which is capable of binding to lineage-affiliated regulatory elements and inducing local and global chromatin changes further facilitating comprehensive instillation of an alternative transcriptional programme. In this respect ASCL1 has been identified as a master neuronal converter (Wapinski et al., 2013), GATA binding protein 4 (GATA4) tends to be associated with reprogramming to hepatic and cardiac fate (Huang et al., 2011; Ieda et al., 2010) whereas MYOD remains the most potent inducer of myogenic genes to date (Choi et al., 1990; Weintraub et al., 1989). The insightful mechanistic studies of MYOD binding revealed that during the reprogramming of non-muscle cells TF binds to its permissive enhancer marked by H3K4me1 (Taberlay et al., 2011). Upon provision of permissive culture conditions, and as a result of subsequent promoter rearrangements, the endogenous expression of MYOD is rejuvenated. Therefore it was suggested that existence of multivalent permissive enhancers uniformly accessible for master lineage-affiliated regulators constitute one of the mechanism by which cells retain transcriptional plasticity.

It is notable that even though pioneer factors are very potent in the reprogramming of developmentally related lineages, the stable transition between germ layers requires the orchestrated cooperation of multiple factors.

As was shown for ASCL1, the initial binding of this master regulator to its authentic neuronal target genes in fibroblasts is a means to facilitate the recruitment of other exogenous factors (Wapinski et al., 2013). Importantly, the type of assisting TF directed the sub-lineage outcome. Co-transduction with LIM homeobox transcription factor 1 α (LMX1α) and orphan nuclear receptor NURR1 led to the acquisition of dopaminergic fate (Caiazzo et al., 2011). Addition of Octamer-binding transcription factor 7 (OCT7) and Myelin Transcription Factor 1-Like (MTF1L) directed cells towards glutamatergic neurons (Vierbuchen et al., 2010), whereas further expansion of reprogramming mix resulted in the generation of motor neurons (Son et al., 2011). These studies showed that the generation of mature and functional cells in the course of transdifferentiation may require not only generic fate determinants but also maturation factors. This notion was confirmed by Du and colleagues who reported the generation of first human hepatocytes with the reconstituted network of drug metabolism upon the transduction of fibroblasts with hepatic master regulators (Hepatic Nuclear Factor (HFN) 1α, 4α and 6) along with the maturation factors (Activating Transcription Factor 5 (ATF5), Prospero Homeobox 1 (PROX1) and CEBPα) (Du et al., 2014).

Despite significant progress made in the recent years the mechanisms governing transdifferentiation remain very much understudied. Interestingly, irrespective of the implemented protocol there are a few characteristics that apply to the direct conversion as a whole. Typically the yield of transition spans from 10 to even 30% and newly obtained cellular identity is stable after removal of transgenes (Huang et al., 2011; Ieda et al., 2010; Sekiya and Suzuki, 2011; Zhou et al., 2008). Moreover, the transcriptional programme of the destination lineage is upregulated very rapidly within few hours to several days posttransduction (Ieda et al., 2010; Vierbuchen et al., 2010). At the same time several studies point out the issue of residual epigenetic memory manifested by only partial downregulation of host cell gene expression and/or incomplete restoration of the target transcriptional profile (Ieda et al., 2010; Marro et al., 2011; Sekiya and Suzuki, 2011). This phenomenon may be reflective of a possible molecular mechanism governing transdifferentiation. It was proposed that some cells obtain alternative cellular fate directly by passing through an

unnatural intermediate stage in which two genetic programmes are active at the same time. This phenomenon could be observed in transitioning B-cells that presented low levels of both CD19 and CD11b when directed toward a macrophagic fate (Xie et al., 2004). Another example includes the transdifferentiation of α-cells to β-cells which was manifested by the appearance of an intermediate population producing both glucagon and insulin (Thorel et al., 2010). In general, it is tempting to attribute this mechanism to conversions of closely related cells residing within the same germ layer. For more developmentally distant cell types an alternative model was proposed, involving initial dedifferentiation to a precursor stage closely followed by the differentiation toward the desired lineage. The recent advent of studies identifying progenitor populations during the switching to neurons (Han et al., 2012), hepatocytes (Yu et al., 2013) or blood cells (Sandler et al., 2014) strongly supports this hypothesis. Importantly, the generation of stem/progenitor cells is more desirable for regenerative medicine as during transplantation naïve cells present better capabilities to incorporate into an *in vivo* microenvironment and fulfil a wide spectrum of biological functions. Moreover, unlike terminally differentiated cells, precursors are highly proliferative providing a platform for the generation of large number of cells for therapeutic use.

Conceptually transdifferentiation is a very promising approach providing possibility of one-step rapid transition between distant lineages. However, our poor understanding of mechanisms governing direct reprogramming makes it challenging to fully unleash potential of this method. The most pressing issue is inability to recapitulate full identity of the cell of interest. Conditions used in the majority of protocols don't allow complete erasure of original cellular landscape. As a result generated cells obtain only partial morphological, transcriptional and functional resemblance to alternative lineage. Normally fate commitment is governed by plethora of TFs cooperating in systematic and hierarchical manner to obtain stepwise differentiation of progenitors. Instead direct reprogramming forces unphysiological transition between terminally differentiated cells. It still remains to be established whether small pool of lineage-instructive TFs can gain access to comprehensive set of loci and overcome epigenetic barriers to activate coherent transcriptional program of interest. Additionally more systemic approach utilising delivery of extrinsic stimuli and provision of supportive niche

mimicking *in vivo* microenvironment may help to obtain full reprogramming. Another burning issue is insufficient efficiency of transdifferentiation. Even though in some cases the conversion rates can reach 30%, direct reprogramming remains transition of one cell to another without advantage of proliferative intermediate. The recent progress in the generation of expandable progenitor populations may help to address issues of clinical-scale production as well as heterogeneity of generated cells.

1.3.2.1. Cell fate conversions with clinical applications

Even though transdifferentiation is highly promising approach for the generation of patient-specific cells it is not devoid of drawbacks. Very often obtained cells are not only immature but also only partially reprogrammed, as manifested by residual gene expression of starting cell type. Additionally, similarly to iPSC approach, direct reprogramming requires overexpression of TFs normally not expressed in a starting cell type, which may lead to the generation of non-physiological cellular intermediates. Such cells of unknown or unstable identity create high risk of unpredictable behaviour when introduced *in vivo.* The main challenge is therefore to generate well characterised high quality cells for prospective use in therapy

Despite issues listed above many of the direct conversions achieved to date showed clear potential therapeutic applications. The first proof of principle that transdifferentiated cells could be used in functional regenerative therapy was obtained by the Melton group. Following *in vivo* expression of the three TFs (Neurogenin 3 (NGN3), Pancreatic and Duodenal Homeobox 1 (PDX1) and V-Maf Avian Musculoaponeurotic Fibrosarcoma Oncogene homolog A (MAFA)) differentiated pancreatic cells were reprogrammed into active β-cells (Zhou et al., 2008). Even though the induced β-cells did not create islets, their presence mitigated the hyperglycemia in a mouse model of type I diabetes. More recently, overexpression of the same combination of factors in the intestine led to the generation of glucose-responsive insulin-producing neoislets below the intestinal crypt base (Chen et al., 2014). Importantly, the presence of reprogrammed cells ameliorated hyperglycemia in diabetic mice improving glucose homeostasis also after downregulation of exogenous transgenes.

Another example of the clinical potential of transdifferentiation was delivered by Srivastava and colleagues. Ectopic expression of GATA4, Myocyte Enhancer Factor 2C (MEF2c) and T-Box 5 (TBX5) (GMT combination) in fibroblasts led to the emergence of cardiomyocytes with a sarcomeric structure within 2 weeks upon transplantation into immunodeficient mouse recipients (Ieda et al., 2010). More recently, two groups performed *in vivo* reprogramming of cardiac fibroblasts in the injured heart. The retroviral introduction of GMT combination (Qian et al., 2012) or GMT and Heart And Neural crest Derivatives expressed 2 (HAND2) (Song et al., 2012) resulted in conversion rates of 35% and 6% respectively. In both reports functional studies performed 2 to 3 weeks after gene transfer confirmed improved cardiac function and decreased scarring/fibrosis of the heart.

A satisfactory therapeutic effect has also been achieved from induced hepatocytes. When injected into the liver of the mouse model of hereditary tyrosineaemia type I, transduced fibroblasts reconstituted hepatic tissue with a clinical pattern and kinetics similar to WT hepatocyte transplantation (Huang et al., 2011; Sekiya and Suzuki, 2011). Nevertheless the reconstitution was not complete, which resulted in reduced survival rates. This clinical outcome has been recently improved by Yu and colleagues who generated hepatic stem cells with the use of Hepatocyte Nuclear Factor 1β (HNF1β) and Forkhead Box A3 (FOXA3) only (Yu et al., 2013). *In vivo* transdifferentiated cells repopulated 11.6% of total hepatocytes in the liver-failure mouse model. The engrafted cells exhibited some functionality manifested by increased levels of albumin and decreased bilirubin in blood of treated mice compared to MEF control.

Neurodegenerative diseases are characterised by the loss of functional neurons in brain and therefore constitute perfect targets for cell-based therapies. The early work by Kim and colleagues focused on the generation of dopaminergic neurons for the treatment of Parkinson's disease. In this study fibroblasts transduced with 6 neuronal TFs gave rise to cells which, when injected into the lesioned striatum, induced substantial re-innervation of the structure, clearly alleviating symptoms of the disease (Kim et al., 2011). A significant advance has been also achieved regarding cell-replacement therapies for Alzheimer’s disease. This condition is manifested by the

appearance of β-amyloid plaques and chronic reactive gliosis. Interestingly, overly amplified glial cells, including astrocytes and NG2 cells, proved to be highly susceptible to *in vivo* reprogramming with Neuronal Differentiation 1 (NEUROD1) (Guo et al., 2014). Forced expression of this TF in a mouse model for Alzheimer's disease led to the generation of a large number of functional GABAergic and glutamatergic neurons. Importantly, procedure was more efficient in older mice than in young and WT specimens.

Finally, a highly promising outcome for patients suffering from corneal surface diseases is the recent acquisition of human limbal stem/progenitor cells (LSCs) from skin epithelial stem cells via ectopic expression of Paired Box 6 (PAX6) (Ouyang et al., 2014). Reprogrammed LSCs could further differentiate into cornea epithelial cells when transplanted onto eyes in a rabbit cornea injury model. *In vivo* experiments resulted in the complete repair of the injured surface with accompanying restoration of normal clarity and transparency of the cornea.

1.3.2.2. TF-mediated reprogramming toward blood cells

1.3.2.2.1. Direct conversion within the blood lineage

The haematopoietic system is the most thoroughly investigated model of differentiation and lineage specification. Therefore not surprisingly it provided some of the first indications of direct lineage conversion. Multiple overexpression and loss-of-function studies at various stages of haematopoietic hierarchy have been shown to result in direct reprogramming. The first results were obtained with GATA1 and PU.1, two master regulators directing fate decisions in multipotent blood progenitors. Overexpression of *Pu.1* in MEPs led to the generation of myeloid cells (Nerlov and Graf, 1998). Conversely, a switch in the opposite direction was obtained through forced expression of *Gata1* in monocytic cell line (Visvader et al., 1992) and primary myeloid as well as lymphoid progenitors (Heyworth et al., 2002; Iwasaki et al., 2003). Another remarkable example of lineage conversion was delivered by the studies on CEBPα. Its expression in CLPs, MEPs and even mature B-cells, as well as committed pre-T-cells, redirected them into functional macrophages (Fukuchi et

al., 2006; Laiosa et al., 2006; Xie et al., 2004). Later members of the Graf group further investigated the transcriptional mechanism of B-cell to macrophage transition. For this, Bussmann and colleagues employed an inducible CEBPα system that allowed the monitoring of morphological and transcriptional alterations in living cells. Impressively, the first bidirectional expression changes could be detected as early as 3 hours after CEBPα induction (Bussmann et al., 2009). Moreover, cells became transgene independent within 24 to 48 hours. This study illustrated very rapid kinetics (full transition occurring within 2 to 3 days) and exceptional efficiency (reaching 100%) of fate switching governed by CEBPα. Furthermore, DiTullio and colleagues confirmed that upregulation of a macrophagic programme and the downregulation of lymphoid identity are concomitant and not associated with intermediate re-differentiation or reactivation of progenitor traits and markers (Di Tullio et al., 2011).

This series of experiments clearly showed that lineage-affiliated TFs can not only specify the developmental path of uncommitted progenitors but also override the ongoing haematopoietic programme. However, the outcome of reprogramming with single blood specifiers was clearly limited to one terminally differentiated sublineage. The Holy Grail of regenerative medicine is to obtain inducible MPPs (iMPPs) or iHSCs capable of reconstituting the complete blood system of the recipient (Fig 1.5). The first report addressing this issue was delivered by Cobaleda et al. who generated cells with multilineage myeloid and T-lymphoid potential by ablation of *Pax5* in mature B-cells (Cobaleda et al., 2007). Unfortunately, the procedure led to the appearance of highly aggressive lymphomas *in vivo* thus precluding this methodology from therapeutic use. It took a few more years to employ TF-mediated reprogramming for the generation of engrafting HSPCs. For this, human pluripotent stem cell-derived myeloid progenitors were retrovirally infected with 9 candidate HSC-specific TFs and screened for acquired self-renewal potential (Doulatov et al., 2013). The colonies remaining in the long-term *in vitro* culture consistently presented integration of Retinoid-related Orphan Receptor Alpha (RORA), ERG and HOXA9. Additional *in vivo* screening revealed the necessity for SRY (Sex determining Region Y)-box 4 (SOX4) and MYB to endow transduced cells with engrafting capabilities. Nevertheless, re-specified $CD34^+CD45^+$ progenitors

mediated only short-term myelo-erythroid engraftment dependent on continuous expression of the transgenes. This initial success in the highly challenging human system was shortly followed by the demonstration of dedifferentiation of committed myeloid and lymphoid murine cells to iHSCs. Riddell and colleagues started with the selection of 36 candidate reprogramming factors that are differentially expressed between HSCPs and mature blood cells (Riddell et al., 2014). The identification of indispensable reprogramming factors was performed *in vivo* by the immediate injection of 36-factor transduced B-cell precursors into recipient mice. After 16 weeks only very few recipients showed multilineage donor reconstitution. Subsequent polymerase chain reaction (PCR) analysis revealed a set of 6 TFs (Hepatic Leukemia Factor (HLF), Pre-B-Cell Leukemia Homeobox 1 (PBX1), PR Domain Containing 5 (PRDM5), RUNX1; Translocated to 1 (RUNX1t1), LMO2, Zinc Finger Protein 37 (ZFP37)) consistently integrated in all reconstituted lineages, therefore responsible for conferring progenitor potential onto the original Pro/Pre-B-cells. Further optimization experiments with the use of polycistronic vectors and introduction of Meis Homeobox 1(MEIS1) and MYCN led to even greater engraftment levels. Importantly, iHSCs derived from both lymphoid precursors and myeloid effector cells were capable of multilineage long-term repopulation of secondary recipients, therefore exhibiting the biological functions of bona fide HSCs.

1.3.2.2.2. Reprogramming developmentally distant cells to blood

Intuitively crossing epigenetic barriers between developmentally unrelated cells is more demanding than switching between cellular fates within the same lineage. The first study reporting conversion of primary fibroblasts to blood dates back to 2008 when Feng and colleagues demonstrated that the combination of CEBPα and PU.1 can transdifferentiate adult and embryonic skin cells to macrophages (Feng et al., 2008). Notably, despite the high efficiency rate, the conversion was not stable and required the continuous provision of exogenous *Pu.1*. Therefore this study, even though the first of its kind, questioned whether lineage-affiliated TFs are sufficient to direct an inter-lineage transition toward blood. Emerging evidence suggests that transdifferentiation protocols requiring only single maturation TF are rather

unique. Instead, reciprocal regulation of multiple generic fate determinants in the self-organizing network is usually considered a key step of reprogramming. In recent years two groups, including ours, employed this strategy to obtain haematopoietic cells from MEFs. Pereira and colleagues screened a pool of 18 TFs highly expressed in HSCs to eventually identify four regulators (GATA2, GFI1b, cFOS, ETV6) capable of activating human CD34 promoter in double-transgenic MEFs (Pereira et al., 2013). The immunostaining analysis of re-specified cells revealed that the acquisition of the definitive blood marker CD45 by haematopoietic cells was preceded by the appearance of flat adherent SCA1^{+}Prominin1(PROM1)$^{+}$ haemogenic precursors.

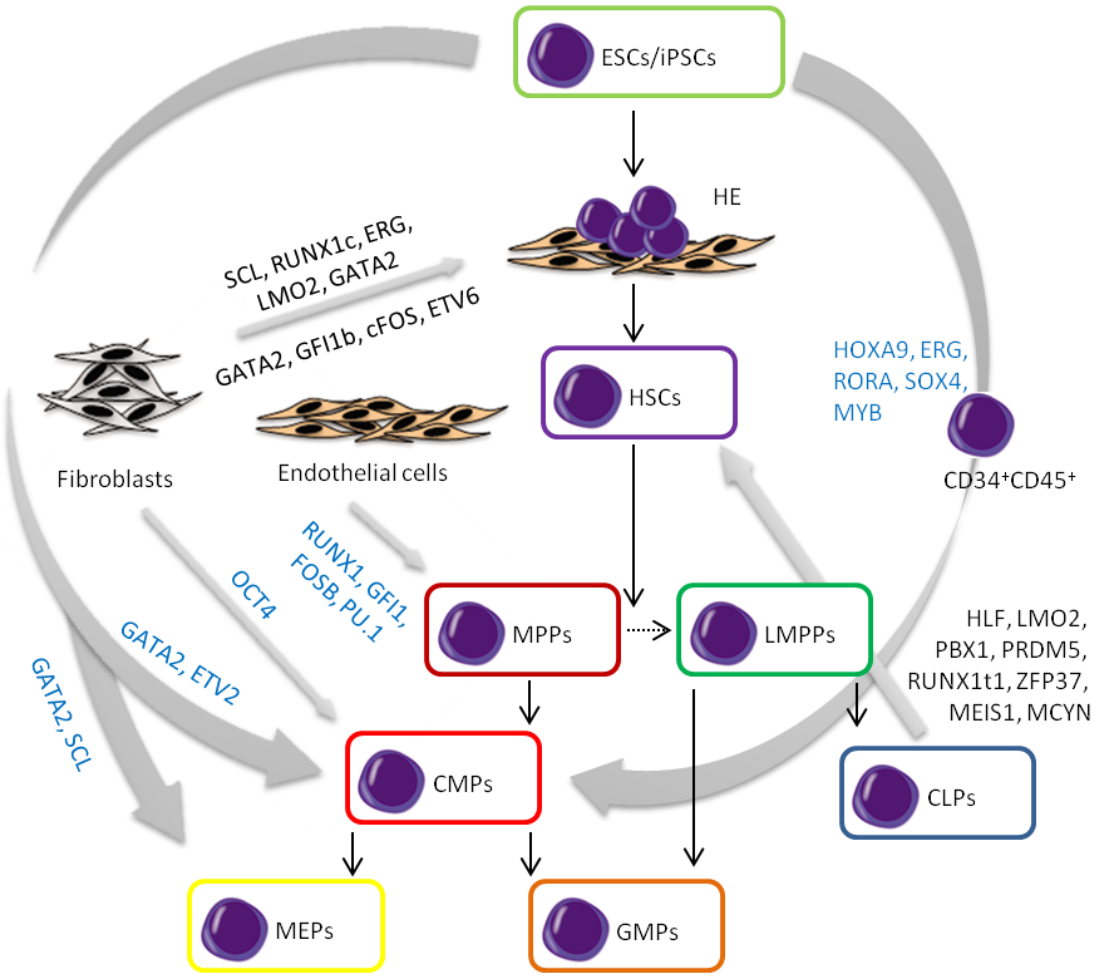


Figure 1.5. Different routes to obtain HSCPs via ectopic expression of instructive TFs.

Current protocols for the generation of blood precursors include direct transdifferentiation with the use of lineage-specific TFs and primed conversion by ectopic expression of pluripotency factors. Haematopoietic cells exhibiting various levels of biological activity has been obtained from both developmentally distant and related blood cell types. The identified combinations of TFs used in mouse (black) or human (blue) cells are listed. Colours and cell populations as in Figure 1.2

Notably, even though $CD45^+$ cells exhibited the transcriptional profile and immunostaining characteristics of HSCs, they only displayed limited myeloid differentiation. The functional assay was preceded by re-aggregation with PL cells which was supposed to provide maturation signals to presumptive pre-HSCs similar to what was observed for $CDH5^+CD45^+$ cells co-cultured with AGM (Taoudi et al., 2008). The *in vivo* potential of reprogrammed cells was not assessed in this study. More functional haematopoietic progenitors were subsequently obtained by our group through transduction of MEFs with five blood fate determinants (Batta et al., 2014). The detailed description of this study will be presented throughout this thesis.

Finally, a significant progress in the field of transdifferentiation to blood has recently been made by Sandler and colleagues. This group delivered the only report on re-specifying human cells to engraftable multipotent progenitors thus far. The crucial aspect of this protocol was the provision of a vascular niche that clearly facilitated the acquisition of an haematopoietic programme by endothelial cells derived from umbilical vein and adult dermal microvasculature (Sandler et al., 2014). The transition was driven by four TFs (GFI1, RUNX1, PU.1 and FOSB) that evoked first morphological changes and the subsequent generation of blood colonies as early as day 12 of the experiment. Upon expansion for another 3 weeks $CD45^+$ cells presented the immunophenotypic and transcriptional characteristics of HSCs and MPPs. Moreover, when transplanted into irradiated mice the iMPPs displayed robust multilineage reconstitution of primary and secondary recipients. The fact that engrafted mice showed no evidence of malignant transformation or genetic abnormalities makes this approach a highly promising cell-based treatment of blood-related disorders.

1.3.2.2.3. Transdifferentiation to blood with the use of pluripotency factors

The alternative to classical direct reprogramming is primed conversion in which somatic cells are first subjected to the action of pluripotency factors. In theory this initial step is supposed to evoke plasticity in treated cells and make them susceptible to the permissive culture conditions. This approach was firstly successfully used by Szabo and colleagues for the generation of

multilineage blood progenitors from human dermal as well as neonatal fibroblasts (Szabo et al., 2010). Cells transduced with *Oct4* and cultured in medium supplemented with FLT3 ligand (FLT3L) and kit ligand (KL) gave rise to $CD45^+$ haematopoietic colonies without the concomitant upregulation of pluripotency genes or cell surface markers. These blood cells were further shown to differentiate *in vitro* into monocytes, granulocytes, megakaryocytes and erythrocytes. Transplantation experiments revealed long-term myeloid engraftment and minimal repopulation of secondary recipients. Originally the authors suggested that instructive role of OCT4 in this setting originates from the unphysiological occupancy of binding loci normally ascribed to OCT1 and OCT2, which are expressed in haematopoietic system (Corcoran et al., 1993; Wang et al., 2004). However, considering that two-fold increase in *Oct4* expression promotes differentiation of ESCs into primitive mesoderm, it is more plausible that the formation of a mesodermal intermediate provides fibroblasts with an incentive to undergo haematopoietic commitment (Niwa et al., 2000). More recent transcriptional profiling of *Oct4*-transduced fibroblasts revealed that originally cells express genes associated with multiple lineages from across the germ layers (Mitchell et al., 2014). This plastic state, which is not, however, an iPSC intermediate, then responds to permissive cues provided by the specific culture environment.

The methodology of primed conversion was also employed by Pulecio and colleagues to obtain engrafting monocyte-like progenitors from human fibroblasts. In this study the initial conversion to CD34-presenting cells was governed by SOX2, whereas further maturation to $CD34^+CD45^+$ myeloid precursors was facilitated by micro-RNA mir-125b (Pulecio et al., 2014). Importantly, reprogrammed cells exhibited intermediate-term (8 weeks) myeloid reconstitution of mouse recipients including BM and spleen compartments.

Altogether, even though primed conversion is associated with the risk of teratoma formation due to the residual expression of pluripotency factors, it also provides an interesting alternative for the production of large numbers of patient-specific cells. The plastic intermediate formed during the process can be prospectively used to amplify re-specified cells to the therapeutically relevant scale. Moreover, it was shown that initial short-term treatment with OSKM combined with the culture in cardiogenic media increases the efficiency of

conversion to cardiomyocytes from 10% reported for the classical GMT transduction to 40% (Efe et al., 2011; Ieda et al., 2010).

1.4. Aim of the project

The direct conversion to blood is an approach, as promising as challenging, to generate transplantable cells for the cell-replacement therapies. Current treatments of blood-related diseases are based on BM transplants that are still associated with important morbidity and mortality due to infections and incidents of graft-versus-host disease. Moreover, the lack of HLA-matching donors precludes use of allogeneic transplantations for the treatment of all patients in need. My thesis project aimed to investigate the alternative route for the generation of autologous cells by direct conversion of easily accessible fibroblasts to blood. Firstly, I sought to identify the minimal set of haematopoietic regulators capable of TF-mediated transdifferentiation in the murine system. The optimized protocol was further employed for the investigation of molecular mechanisms and cellular processes governing the transition of developmentally distant fibroblasts to the haematopoietic lineage. As a result, I identified intermediate haemogenic endothelium and multilineage progenitor populations, both giving rise to a wide spectrum of terminally differentiated blood lineages. Eventually, the particularly appealing goal became obtaining iHSCs through the amendment of the initial culture conditions and transduction methodology. Even though the latter has not been achieved yet, we generated promising results that will be further explored in the course of future work. Overall, this thesis resulted in the development of one of the most robust reprogramming protocol to date which deliver multilineage blood progenitors from murine fibroblasts.

Chapter 2

2. Materials and methods

2.1. *In vitro* culture

2.1.1. Isolation of primary embryonic fibroblasts

MEFs were isolated from E14.5 embryos. Head and internal organs were removed and discarded. The remaining tissue was washed with 70% ethanol (EtOH, Fisher Scientific) and then immediately with sterile phosphate-buffered saline (PBS). Material from a single embryo was placed in a well of a 12 well plate and macerated into fine slurry with a scalpel. Tissue was then digested with 1 ml of trypsin-EDTA solution (T/E, Gibco) for 5-10 minutes at 37^{o}C accompanied by homogenization by repeated pipetting. The generated cell suspension was removed from above the tissue slurry and transferred to ϕ10 Petri dishes containing 10 ml of Iscove's Modified Dulbecco's Medium (IMDM, PAA) supplemented with 10% fetal calf serum (FCS, Sigma), 1% penicillin/streptomycin (P/S, Invitrogen), 2mM L-glutamine (L-Gln, Invitrogen) and 20 μM alpha-monothioglycerol (MTG, Sigma) (later referred to as full IMDM). After an overnight culture at 37^{o}C/5% O_2 atmosphere, medium was changed and unattached material removed. When confluent, cells were frozen in liquid nitrogen at passage 0 using 50% medium, 40% FCS and 10% dimethyl sulfoxide (DMSO, Sigma) solution. For experiments MEFs at passage 1 were used.

2.1.2. Isolation of adult skin fibroblasts

Mouse adult fibroblasts (MAFs) were prepared from ears of mice between 2 and 6 months of age. Minced tissue was suspended in 10 ml of 10% FCS/PBS containing 62.5 U/ml collagenase II (Gibco), 62.5 U/ml collagenase IV

(Gibco) and 2 U/ml DNase (Qiagen) and incubated at 37^{o} for 2 hours to overnight with occasional stirring. Dissociated tissue was passed through a 70 μm cell sieve and pelleted by centrifugation at 290 g for 5 minutes. Cell suspension was re-suspended in Dulbecco's Modified Eagle Medium (DMEM, PAA) supplemented with 10% FCS, 1% P/S, 2 mM L-Gln and 0.025 M 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (Sigma) (later referred to as full DMEM) and seeded onto T75 flasks. Cells were frozen at passage 0 or 1depending on the proliferation capacity of a batch.

2.1.3. Culture and irradiation of stromal cells.

The OP9 and OP9-DL1 cells were cultured in Minimum Essential Medium Eagle (MEM, PAA) supplemented with 20% FCS, 2 mM L-Gln and 1% P/S. Confluent cells were harvested with T/E and suspended in medium. The cells were irradiated with 30 Gy (21.9 minutes on our instrument) in 50 ml Falcon tubes. Irradiated cells were pelleted and re-suspended in 90% FCS and 10% DMSO and stored at -80^{o}C. On appropriate day cells were thawed and seeded at 18500 cells/cm^{2}.

2.1.4. Culture of control BM-derived blood cells

To obtain megakaryocytes from the BM progenitors the femurs were flushed into DMEM supplemented with 10% FCS, filtered and counted. Cell were plated at 5x10^{5}/ml in ϕ10 cm dishes for 5 to 6 days in IMDM medium supplemented with 10% FCS (Stem Cell Technologies), 2 mM L-Gln, 1% P/S, 0.5 mM MTG, 1% thrombopoietin (TPO) conditioned medium, 20 ng/ml interleukin (IL) 6, 1% IL3 conditioned medium and 50 ng/ml IL11. After the initial culture suspension of megakaryocytes was aspirated, centrifuged at 200g for 5 minutes and subjected to acetylcholinesterase staining protocol (see paragraph 2.5.6)

To obtain macrophages the femurs were flushed with DMEM containing 10% FCS, filtered, centrifuged at 290 g for 5 minutes and re-suspended in 5 ml of medium. Next mononuclear cells were separated by centrifugation on Ficoll-Paque (GE Healthcare) for 30 minutes at room temperature (RT), washed

with PBS and re-suspended in DMEM containing 10% FCS, 1% P/S, 2 mM L-Gln, 1% granulocyte-macrophage colony-stimulating factor (GM-CSF) conditioned medium and 5 ng/ml M-CSF. The equivalent of two femurs was seeded on ϕ10 cm dish and cultured overnight at 37°C in humidified atmosphere containing 5% CO_2. The next day not attached cells were recovered and transferred to a new dish. After 2 days culture medium was changed and after 5 days from seeding adherent macrophages were harvested with T/E and used in enzyme-linked immunosorbent assay (ELISA) (see paragraph 2.5.3).
If not stated otherwise cytokines and growth factors were obtained from R&D Systems.

2.1.5. Reprogramming by viral transduction

Cultures of MEFs or MAFs were started 5 days before the experiment. Cells were split once and grown to confluency. On day 0, cells expressing the haematopoietic CD45, CD41, cKIT and endothelial CD31 markers were excluded by sort. Subsequently, the cells were seeded on 0.1% gelatine-coated (Sigma) culture dishes in full DMEM (MAFs) or IMDM (MEFs) medium at a density of 30000/9cm^2. The next day, the fibroblasts were infected with lentiviruses encoding haematopoietic TFs of choice in the presence of 10 μg/ml diethylaminoethanol (DEAE)-dextran (Amersham Biosciences). For the initial experiments, querying minimal set of haematopoietic TFs capable of inducing direct conversion to blood (chapter 3), I used 19 TFs chosen on the basis of data mining. For subsequent experiments, aiming to characterise transition outcomes and mechanisms governing the process (chapters 4, 5 and 6), I used the narrower set of five TFs (SCL, RUNX1c, GATA2, ERG, LMO2) sufficient to induce transition (on the basis of results from paragraph 3.2.4). Finally to test the minimal required set of reprogramming TFs (paragraph 4.2.8) MEFs were transduced with two molecules (SCL, LMO2) selected on the basis of TF-removal experiment (paragraph 4.2.7). Overall viral aliquots of reprogramming TFs were diluted 100x in a total volume of 1 ml and 100 μl of this mix was applied on 15000 seeded cells. After 4 hours viral medium was replaced with fresh haematopoietic medium. In some experiments, the progression of

reprogramming was monitored by time-lapse microscopy using IncuCyte FLR or IncuCyte ZOOM (Essen BioScience). In other experiments, the cells were harvested at different time points with T/E and subjected to downstream analyses.

2.1.6. Reprogramming inducible *Scl* (*iScl*) MEFs

The doxycycline inducible *Scl-2A-GFP* transgenic mice (*iScl*) are carrying two transgenes. *Scl* expression cassette is inserted into the *Hprt* locus and is linked to green fluorescent protein (GFP) coding sequence via a 2A viral peptide sequence under the control of a Tet responsible element (TRE). The second transgene inserted into the *Rosa* locus encodes a reverse tetracycline transactivator (rtTA). The stimulation with doxycycline initiates the transcription of both *Scl* and *Gfp*, allowing monitoring the expression of gene of interest.

E14.5 *iScl* MEFs negative for CD45/CD41/cKIT/CD31/GFP were isolated and seeded as described in paragraph 2.1.5. On day 1, the fibroblasts were infected with a LMO2 encoding lentivirus and after 4 hours the viral supernatant was replaced with reprogramming medium containing 1 μg/ml doxycylin (dox, Sigma). Fresh dose of antibiotic was added on days 5 and 9 of experiment.
For the screening of a compound library cells were prepared as described in paragraph 2.1.5 and seeded at 1825 cells/well on a 96 well plates containing the library molecules (10 μM) in reprogramming medium with 1μg/ml dox. Antibiotic was replenished on days 5 and 9 of experiment.

2.1.7. Haematopoietic culture conditions

The basal reprogramming medium contained IMDM, 10% fetal bovine plasma-derived serum (PDS, Animal Technologies), 10% protein-free hybridoma medium (Invitrogen), 2 mM L-Gln, 20μg/ml ascorbic acid (Sigma) and 500 μM MTG. The standard full haematopoietic liquid big mix (LBM), used for all standard reprogramming experiments, was supplemented with 180 μg/ml transferrin, 1% KL conditioned medium, 1% IL3 conditioned medium, 1% TPO conditioned medium, 1% GM-CSF conditioned medium, 5 ng/ml IL11, 4 U/ml

erythropoietin (EPO, Ortho Biotech), 10 ng/ml IL6 and 10 ng/ml M-CSF. If not stated otherwise cytokines and growth factors were obtained from R&D Systems.

Erythroid cultures were carried out on OP9 stroma in the basal reprogramming medium containing 1% KL conditioned medium, 4 U/ml EPO and 180 μg/ml transferrin. Megakaryocyte conditions included supplementation with 1% IL3 conditioned medium, 10 ng/ml IL6 and 1% TPO conditioned medium. These conditions were used for lineage specification from sorted cKIT$^+$ cells (paragraph 5.2.3). In this line cells were originally reprogrammed in LBM and after day 12 sort cultured in lineage-specific medium.

Lymphoid potential of day 12 sorted cKIT$^+$ reprogrammed cells was tested by 4 weeks cultures on OP9 (B-cells) or OP-DL1 (T-cells) stroma (paragraph 5.2.1 and 5.2.3). Cells were passaged onto new support layer two times every week. The first two passages were performed in medium containing MEM, 10% FCS, 1% P/S, 2 mM L-Gln, 5 ng/ml FLT3L and 1 ng/ml IL7 (both from Peprotech). Upon the third passage, T-lymphoid cultures were subjected to low (0.25 ng/ml) IL7 dose. In some experiments, 3 μM GSK3 inhibitor CHIR99021 (R&D Systems) was added to T-cell cultures.

Serum free cultures were carried out in StemSpan medium (Stem Cell Technologies) containing 1μM hydrocortisone (Sigma), 1% KL conditioned medium, 1% IL3 conditioned medium, 10 ng/ml IL6 and 10 ng/ml FLT3L. Cells were cultured for 19 days with weekly half medium change. These conditions were used for the generation and maintenance of HSPCs (paragraph 5.2.2).

2.2. DNA and RNA manipulation

2.2.1. Cloning of TFs to lentiviral backbone

The majority of TFs used in this project were amplified using a Taq proofreading mix which contained 50 μl of 2x concentrated Green Taq (Promega), 1 μl of 50 mM magnesium chloride (Finzymes) and 1 μl of 2 U/μl Phusion DNA Polymerase (Finnzymes). Coding sequences of CEBPα and β were obtained with Phusion HotStart Polymerase in combination with

Q solution (Qiagen) and buffer for templates with high GC content (Finnzymes). Sequences of the primers used and introduced restriction enzyme (RE) sites are presented in Table 2.1. Each forward primer included a FLAG-tag and a KOZAK sequences that preceded coding region of TF in the final vector. A template for amplification of TFs was either murine complementary DNA (cDNA) or plasmid containing the sequence of interest (see table 2.1). The PCR products were sub-cloned into pCR®-Blunt II-TOPO® vector using a Zero Blunt® TOPO® PCR Cloning Kit (Invitrogen). Subsequent to amplification in competent bacteria, coding sequences of TFs were cloned into the SCBR757 vector which is a lentiviral vector compatible with the third generation HIV derived lentiviral packaging system (Fig 2.1). This vector, produced by Zufferey and colleagues, has been further modified by Seppen and colleagues to include the human hepatitis B post transcriptional regulatory element which enhances mRNA export from the nucleus and the HIV central poly purine tract which facilitates nuclear import of viral DNA (Seppen et al., 2002; Zufferey et al., 1998).

Ligations of all DNA fragments were carried out using the T4 Ligase (Roche). Transformations were performed according to standard procedures in Oneshot TOP10 chemically competent strain (Invitrogen). DNA plasmid isolations, gel extractions and purification were performed using the relevant Qiagen kits. All enzymatic digest reactions were performed using New England BioLabs (NEB) enzymes and buffers. Dephosphorylation was performed using alkaline phosphatase (Roche).

2.2.2. Virus production and purification

A third generation HIV derived lentiviral vector packaging system consisting of three plasmids (pMDG.2, rsv-REV and pMDL), encoding essential viral proteins, was used. These constructs and the lentiviral vector containing the TF were co-transfected in human embryonic kidney cells (HEK293T) to produce infectious viral particles. On day 1 HEK293T cells were seeded on ϕ 15 cm dishes (1/4 of confluent T162 flask per dish) and a mix of plasmids was prepared in the following proportions: 20 μg of the lentiviral plasmid of interest, 13 μg of pMDL, 5 μg of rsv-REV and 7 μg of pMDG.2, 1/18 volume of 5 M sodium chloride (NaCl, Sigma) and 2 volumes of 100% EtOH. After overnight

storage at -80°C the DNA pellet was washed twice in 70% EtOH and dissolved under sterile conditions in water containing 0.25 M calcium solution. The transfection procedure was performed according to the CalPhos™ Mammalian Transfection Kit protocol (Clontech). On the next day full DMEM medium was changed and during the following two days viruses were collected. Finally, the viral supernatants were filtered with 0.2 μm syringe filters and ultracentrifuged at 130000 g for 2 hours (Sorval Ultra Pro80). The resulting pellets were dissolved in 200 μl of PBS per dish, incubated at RT for 10-30 minutes and stored in aliquots at -80°C. The expression of the protein of interest was tested by Western Blot.

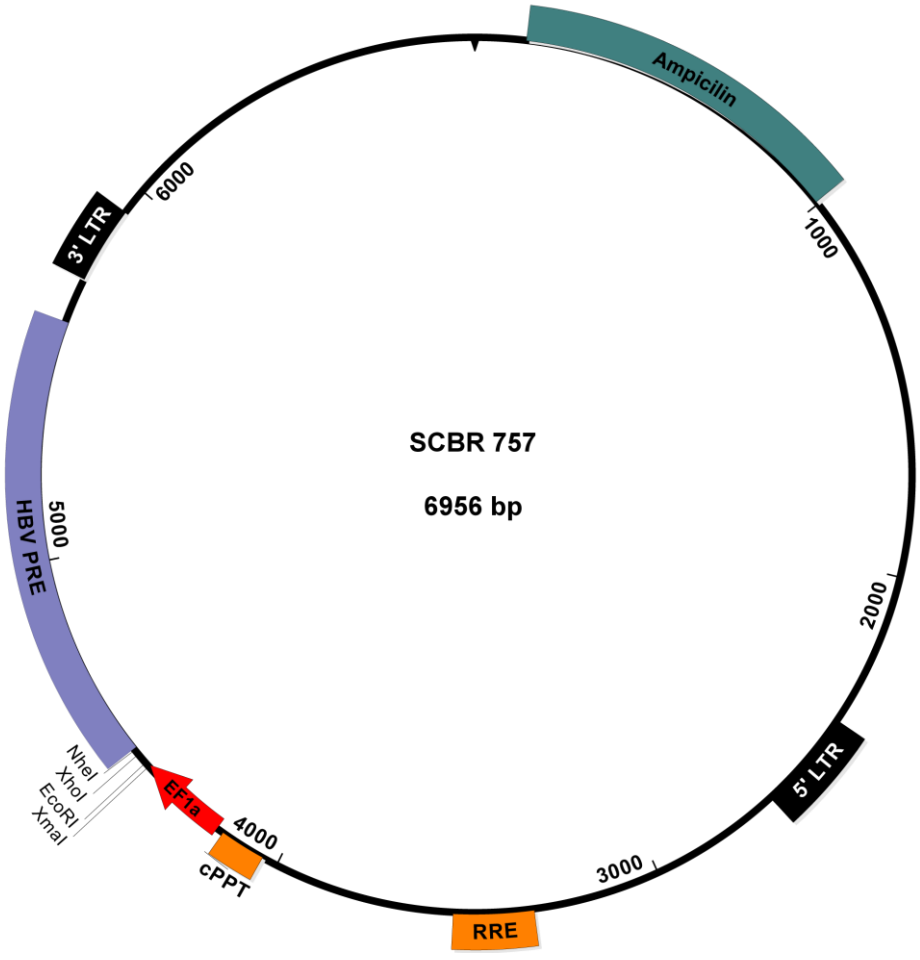


Figure 2.1. Map for vector SCBR757.

Vector SCBR757 was used as a backbone to clone coding sequences of haematopoietic TFs of interest. EF1α - Elongation factor 1α promoter, LTRs – long terminal repeats, cPPT - central polypurine tract, HBV PRE - Hepatitis virus B posttranscriptional regulatory element, RRE - Rev response element.

Table 2.1. Sequences of primers used for the cloning of TFs.

KOZAK refers to the GCCACC sequence, FLAG_ stands for ATGGATTACAAGGA TGACGACGATAAG, source of SCBR plasmids is Stem Cell Biology group plasmid database. *cloned by Kiran Batta

TF	Template	Forward primer Reverse primer	RE sites
CEBPα	murine liver	GAGA**gaatcc**_KOZAK_FLAG_gagtcggccgacttctac	EcoRI
		GAGA**gctagc**tcacgcgcagttgcccatg	NheI
CEBPβ	murine liver	GAGA**gaatcc**_KOZAK_FLAG_caccgcctgctggcctg	EcoRI
		GAGA**gctagc**ctagcagtggcccgc	NheI
cFOS*	murine BM	GAGA**gaattc**_KOZAK_FLAG_atgttctcgggtttcaac	EcoRI
		GAGA**gctagc**tcacagggccagcagc	NheI
ERG1*	SCBR-967	GAGA**cccggg**_KOZAK_FLAG_gccagcactattaag	XmaI
		GAGA**gaattc**ctagtagtaggtgcccaggtg	EcoRI
ETS1*	SCBR-968	GAGA**cccggg** _KOZAK_FLAG_ aaggcggccgtcgat	XmaI
		GAGA**tgatca**ctagtcagcatccggctttac	EcoRI
ETV2*	SCBR-1124	GAGA**cccggg**_KOZAK_FLAG_gacctgtggaactgg	XmaI
		GAGA**gaattc**ttattggccttctgcacctgg	EcoRI
ETV6*	Murine BM	GAGA**gaattc**_KOZAK_FLAG_tctgagactcctgctc	EcoRI
		GAGA**gctagc**ctattcccgggtctcttcct	NheI
FLI1*	SCBR-969	GAGA**cccggg**_KOZAK_FLAG_gacgggactattaag	XmaI
		GAGA**gaattc**ctagtagtagctgcctaagtg	EcoRI
GATA1*	SCBR-1150	GAGA**cccggg**_KOZAK_FLAG_gattttcctggtctagg	XmaI
		GAGA**ctcgag**tcaagaactgagtggggcg	XhoI
GATA2*	SCBR-1053	GAGA**gaattc**_KOZAK_FLAG_gaggtggcgcctgag	EcoRI
		GAGA**gctagc**ctagcccatggcagtcaccat	NheI
GFI1*	SCBR-388	GAGA**cccggg**_KOZAK_FLAG_ccgcgctcattcctg	XmaI
		GAGA**gaattc**CTAtttgagtccatgctgag	EcoRI
GFI1b*	SCBR-397	GAGA**gaattc**_KOZAK_FLAG_ccacggtcctttctag	EcoRI
		GAGA**gctagc**tcacttgagattgtgttgactc	NheI
HOXB4*	SCBR-521	GAGAgaattc_KOZAK_FLAG_gctatgagttctttttg	EcoRI
		GAGAgctagctcatcgagcgcgcgggg	NheI
LMO2	E9.5 YS	GAGA**gaattc**_KOZAK_FLAG_gaagggagcgcggtg	EcoRI
		GAGA**gctagc**ctagatgatcccattgatcttgg	NheI
PU.1*	SCBR-303	GAGA**cccggg**_KOZAK_FLAG_ttacaggcgtgcaaa	XmaI
		GAGA**gaattc**tcagtggggcgggaggcgc	EcoRI
RUNX1b	SCBR-042	GAGA**gaattc**_KOZAK_FLAG_cgtatccccgtagatg	EcoRI
		GAGA**gctagc**tcagtagggccgccac	NheI
RUNX1c	SCBR-043	GAGA**gaattc**_KOZAK_FLAG_gcttcagacagcatttttg	EcoRI
		GAGA**gctagc**tcagtagggccgccagac	NheI
SCL*	SCBR-225	GAGA**cccggg**_KOZAK_FLAG_acggagcggccgccg	XmaI
		GAGA**gaattc**tcaccgggggccagcccc	EcoRI
SOX7*	SCBR-967	GAGA**cccggg**_KOZAK_FLAG_gcctcgctgctgggc	XmaI
		GAGA**gaattc**ctatgacacactgtagctgttg	EcoRI

2.2.3. Quantitative real time PCR (qRT-PCR)

RNA from pelleted or sorted cells was isolated with RNeasy Mini/Micro Kit (Qiagen). AGM control cells referred to in paragraph 4.2.8 were obtained from E10.5 and E11.5 embryos upon dissection. Tissue was macerated with a scalpel and dissociated with 1 mg/ml Collagenase/Dispase (Roche) and 10 μg/ml DNAse I (Sigma) for 1 h at 37^0C with gentle shaking. Obtained single cell suspension was pelleted was further processed with RNeasy Mini/Micro Kit.

Up to 1 μg of RNA samples were used in a reverse transcription reaction carried out according to the Omniscript RT kit (Qiagen) manual. Subsequently, 1 μl of synthetized cDNA was used in the qRT-PCR reaction prepared with SYBR Green master mix (Life Technologies) or TaqMan Universal PCR Master Mix (Applied Biosystems) and carried out in an ABI 7900 system (Applied Biosystems). The reaction was performed in a total volume of 10 μl with 0.1 μl of the appropriate Exigon Universal Roche Probe (Taq Man variant), 0.2 μl of each primer (0.2 μM, Sigma) and 5 μl of the relevant master mix. Primers for the TaqMan reaction were designed using the Universal Probe Library Assay Design Center provided by Roche Applied Science (Table 2.2). Primers for the SYBR Green reaction were obtained using NCBI primer design tool (Table 2.3). Final data were processed in the SDS2.1 program (Applied Biosystems).

2.2.4. Genomic DNA isolation for PCR

Genomic DNA was isolated using proteinase K (PK) digestion. The pieces of tissue or cell pellet were placed in 24 μl of PBS in a PCR tube. Sample was boiled at 95ºC for 8 minutes. Subsequently 6 μl of 3.33 mg/ml PK (Quiagen) in 0.1x PBS was added to the sample and digestion was carried on for 30 minutes at 55ºC. PK was then inactivated at 95ºC for 8 minutes.

2.2.5. Integration levels of exogenous TFs

Genomic DNA was isolated from 20000 $TER119^+$ or $10x10^4$ $CD11b^+$ sorted reprogrammed MEFs. Subsequently, 1 μl of extracted material was used

for the PCR reaction with SYBR Green mix and primers specific to FLAG-tag and individual TF (Table 2.3).

Table 2.2. List of primers used in the TaqMan-based qRT-PCR.

Gene of interest	Forward primer	Reverse primer	Universal probe
Pu.1	ggagaagctgatggcttgg	caggcgaatctttttcttgc	94
Mpo	cccttcctaaactgaacctgac	atggcctccgtccttctc	12
Ptprc	tcagaaaatgcaacagtgacaa	ccaactgacatctttcaggtatga	100
Itga2b	tgctgctgaccctgctagt	gtcgattccgcttgaagaag	97
β-major	tgcatgtggatcctgagaac	gtgaaatccttgcccaggt	92
β-H1	tggatcctgagaacttcaagc	attggccactccaatcacc	29
Gfi1b	gttgctgaaccagagccttc	ttggggtgtcacgagagg	45
Gata1	ccctgaactcgtcataccact	gaacactggggttgaacctg	83
Cd42b	gctctgttcctccaaaggact	ctgggaatgcaagggactt	4
Pf4	catctcctctgggatccatct	ccattcttcagggtggctat	9
Fbn1	ccttcctgtggctccagat	gctgcccccattcataca	89
Acta2	atggctctgggctctgtaag	cccattccaaccattactcc	9
Fsp1	tcagcacttcctctctcttgg	tttgtggaaggtggacacaa	25
Nanog	caggtttcagaagcagaagtacc	ggttttgaaaccaggtcttaacc	67
Sox2	ggcagagaagagagtgtttgc	tcttctttctcccagcccta	34
Oct4	ctcccatgcattcaaactga	ttctcttgtctacctcccttgc	100
Scl	acagaagaggcaaacagagtga	tcatcattggggaaatctttg	7
Lmo2	cactgaacacctccccaca	cactgaacacctccccaca	95
Runx1c	caactcgcccaccaacat	atggcgctcagctcagtag	25
Erg	ccagccatatgccctctc	ttgggaaaggcctccatc	104
Gata2	acatggacatcaagggtggt	caagagtatgttccacccatcc	104
εγ globin	tggtgaagccttgggaag	ctttgacccttgggttgc	29
β1 globin	gtgacaagctgcatgtggat	gtgaaatccttgcccaggt	92
β2 globin	ctgcatgtggatcctgagaa	agcaggggtgaaatccttg	49
α1/2 globin	tgacagactcaggaagaaacca	gggaagctagcaaacatcctt	60
ζ–globin	catcatgtccatgtgggaga	ggggtagctgcagaagagc	46
β-actin	tgacaggatgcagaagaaga	cgctcaggaggagcaatg	106
Rag1	aggcctgtggagcaaggta	gctcagggtagacggcaag	46
Rag2	tgccaaaataagaaagagtatttcac	gggacatttttgattgtgaatagg	4
Pax5	acgctgacagggatggtg	ggggaacctccaagaatcat	83
λ5	aagccccagttttggtatgtc	ccaaggggtcagacttgg	77
Mb1	gatgccaggggggtctagaag	gttcaccgtcagggatgg	34
VpreB	gctggcctatctcacaggtt	tggctccaagggaagaaga	83
Jag1	tggccgaggtcctacactt	gccttttcaattatgctatcagg	22
Jag2	cgtcattccctttcagttcg	cctcatctggagtggtgtca	95
Hes1	acaccggacaaaccaaagac	cgcctcttctccatgatagg	99
Hey1	catgaagagagctcacccaga	cgccgaactcaagtttcc	17

Table 2.3. List of primers used in the SYBR Green-based PCR.

Gene of interest	Forward primer	Reverse primer
exogenous *Scl*	ggattacaaggatgacgacga	ccgttgagcaggactaggtg
exogenous *Lmo2*		tcctgcgcctcctcttg
exogenous *Runx1c*		tcgtgctggcatctctcat
exogenous *Erg*		catctctgtcttagccagg
exogenous *Gata2*		tagttatgcgccaggccc
Parp	gcctctcgctttttcatattac	tggtctcatagcctttggtctc

2.2.6. Genotyping

Genomic DNA was isolated from tail tip of E14.5 mouse embryo as described in paragraph 2.2.4. Next 1 μl of extracted DNA was used in PCR reaction with 18 μl Green Taq and 0.5 μl of each specific primer (20 μM, Sigma, Table 2.4). The PCR reactions were carried out according to the following thermal cycling protocol: 95°C for 3 minutes, then 40 cycles of 95°C for 30 seconds, 55°C for 30 seconds, 72°C for 1 minute per every 1 kilobase (kb) of product followed by the elongation step at 72° for 7 minutes. PCR reactions were performed using a DNA Engine Dyad thermal cycler (MJ Research) and PCR products were visualised on 1% agarose gels (Appleton) containing 0.1 μg/ml ethidium bromide (Sigma).

Table 2.4. Sets of primers used for genotyping of E14.5 embryos.

Modyfication	Primers
p53$^{-/-}$	gctctgatgccgccgtgttcc g gtgtttcattagttccccaccttgac atgggaggctgccagtcctaaccc
Gfi1:dtTomato	ccctgacgtgccatcttgaagc ctactgctctcaggagagtgatgatctag
Gfi1b$^{WT/GFP}$	gcgcttgtcactcttagtcac cctcaactccaaaggctagag gctgttgtagttgtactccag
Scl$^{WT/GFP}$	ctagatctcgaaggatctggag atactttctcggcaggagca
Rosa$^{WT/rtTA}$	acaaggtttttcactagagaacgcg agatcgaaatcgtctagcgcgtcg

2.2.7. B-cell receptor (BCR) and TCR rearrangements

For the establishment of VDJ and TCR gene rearrangement events *p53*$^{-/-}$ reprogrammed MEFs were sorted for B220 and CD19 (B-lymphocytes) or CD25 (T-lymphocytes). Genomic DNA was isolated from 10x10^4 pelleted cells according to the protocol described in paragraph 2.2.4. D_H to J_H1-2 and D_H to J_H3 gene rearrangements were detected by nested PCR as reported earlier (Yamane et al., 2013). TCR-β chain gene rearrangements were detected by a single round of PCR with the following primers: 5'gcacctgtggggaagaaact3'; 5'tgagagctgtctcctactatcgatt3' (Obokata et al., 2014) according to the thermal

cycler program from paragraph 2.2.5. Products of both experiments were subjected to electrophoresis in 1% agarose gel and visualised bands were excised from the gel for subsequent Sanger sequencing.

2.3. Protein based assays

2.3.1. Western Blot

HEK293T cells ($25x10^4$) were seeded in 2 ml of medium on a 6 well plate. The following day 10 μl of virus was added to cells cultured in full DMEM supplemented with 10 μg/ml of DEAE-dextran. After 4 hours medium was removed and replaced with fresh DMEM. After three more days of culture at 37ºC in 5% CO_2 atmosphere, cells were washed twice with cold PBS and lysed using RIPA buffer containing 25 mM Tris(hydroxymethyl)aminomethane-hydrochloride (Sigma), 150 mM NaCl, 1% of Nonidet P-40 (Sigma), 0.1% of sodium dodecyl sulphate (SDS, Fluka) and 1% sodium deoxycholate (Sigma). After 30 minutes of incubation at 4ºC with continuous mixing, the lysate was centrifuged at 16100 g for 20 minutes and the generated viscous pellet was removed. The whole protein extracts were separated on a 10% gel in MOPS buffer using the NuPAGE® SDS-PAGE Gel System (Invitrogen). Protein transfer to a membrane was performed using iBlot® Gel Transfer Stacks Nitrocellulose (Invitrogen). The membranes were then blocked with 5% Milk (Marvel) for 3 hours at RT and subsequently incubated with 1:1000 dilution of polyclonal anti-FLAG antibody produced in rabbit (PU.1, SOX7, RUNX1b) or 1:5000 dilution of monoclonal M2 anti-FLAG antibody produced in mouse (both from Sigma) at 4ºC overnight. The next day, the membranes were washed with 0.1% Tween20 (Sigma) in PBS 3 times for 15 minutes and incubated with 1:10000 dilution of donkey/horse anti-rabbit/mouse IgG-horseradish peroxidase antibody (Southern Biotech/Cell Signaling Technology) for 1 hour at RT. After 3 additional steps of washing with PBS/0.1% Tween20 detection was performed using the Amersham™ ECL™ Prime Western Blotting Detection Reagent (GE Healthcare). The membranes were transferred to an X-ray film cassette and films were exposed and developed in a MAS automated developing machine.

2.3.2. Immunofluorescence

For the visualization of TFs' nuclear localization (Fig 3.2) HEK293T cells were infected with the virus of interest according to the same procedure as for Western Blot. After 4 days, the cells were harvested and 15000 of them were seeded on gelatin pre-coated μ-Slide VI (Ibidi) in 50 μl of medium. The next day the attached cells were fixed with 0.4% formaldehyde (Sigma) in PBS for 20 minutes at RT. After several washes with PBS, the cells were permeabilized with 0.2% TritonX (Sigma) solution with PBS for 5 minutes at RT, washed in 20% FCS/PBS and the slides were blocked with 10% horse serum (Gibco) for 1 hour at RT. The cells were then incubated with a rabbit anti-FLAG primary antibody at 1:500 for 1 hour and washed with 20% FCS/PBS (every 10 minutes for an hour). The cells were incubated for another hour with goat anti-rabbit Alexa647 antibody (Invitrogen) in 1:2000 dilution. After another series of washes actin filaments were stained with a 2 U/ml solution of phalloidin-Alexa488 (Invitrogen) for 20 minutes followed by repeated washing with PBS. Finally, cells were covered with ProLong® Gold antifade reagent with DAPI (Invitrogen) and left in the fridge overnight.

For CDH5 staining (Fig 5.8) cells were reprogrammed for 6 days on gelatine-coated coverslips, then washed with PBS and fixed with 0.4% formaldehyde in PBS for 20 minutes at RT. After washes with PBS, the cells were permeabilized with 0.2% TritonX solution in PBS for 5 minutes at RT, washed in 20% FCS/PBS and the slides were blocked with 10% goat serum (Gibco) for 10 minutes at RT. The cells were then incubated with a CDH5-biotin primary antibody (eBioBV14, eBioscience) at 1:100 dilution overnight at 4ºC. The next day cells were incubated with the secondary anti-rat Alexa647 antibody (Invitrogen) in 1:400 dilution for 1 hour. As a final step, after series of washes, cells were covered with ProLong® Gold antifade reagent with DAPI and left at 4ºC overnight.

For the visualization of acetylated-low density lipoprotein (AcLDL) uptake the life-staining on coverslips was performed. Cells at day 5 of culture were incubated overnight with 1:100 dilution of DIL-AcLDL (Life Technologies). Next the procedure for standard CDH5 immmunostaining was performed with exception of fixation and permeabilization steps.

Day 12 cKIT staining of reprogrammed MAFs and MEFs (Fig 5.1) was performed according to CDH5 staining steps with a use of 1:200 dilution of biotinylated cKIT antibody (2B8, eBioscience).
Imaging of all slides was performed using a low-light system consisting of a Zeiss Axiovert 200M microscope and an Andor iXon DU888+ camera. The system utilises Metamorph (Molecular Devices) software. Subsequent analysis was performed in Adobe Photsohop and ImageJ programmes.

2.4. Flow cytometry and cell sorting

2.4.1. Fluorescence-activated cell sorting (FACS)

Before setting up reprogramming experiment, MEFs and MAFs were stained with CD45, CD31, cKIT (1:200; 30-F11, Mec 13-3, 2B8) and CD41 (1:300, eBio-MWReg30) biotin-conjugated antibodies and subsequently with secondary streptavidin PE-Cy7 (1:300). The population negative for these markers was isolated by sort in order to exclude haematopoietic and endothelial cells. CDH5 sort at day 6 was performed with APC-conjugated antibody (eBioBV13) in 1:200 dilution. Day 12 cKIT cells were enriched using APCeFluor780 conjugate (1:200, 2B8). To isolate cells for phagocytosis and ELISA day 21 cultures were stained with CD45-PerCyP5*5 (1:200, 30-F11, Biolegend) and CD11b-APC (1:300, M1/70) or F4/80-APC (1:300, BM8) respectively. Lymphoid cells were sorted after 4 weeks of OP9 or OP9-DL1 culture using PE-conjugated B220 (1:300, RA3-6B2) and CD19-PECy7 (1:300; eBio1D3) for B-cells and CD11b-FITC (1:500, M1/70, BD Pharminogen) together with APC-conjugated CD25 (1:300, PC61.5) for T-lymphocytes. Erythroid and myeloid day 21 reprogrammed cells for the analysis of viral integration were sorted on the basis of TER119 (PE-CY7 conjugate, 1:200, TER119) and CD11b expression respectively. The CD41 positive and negative fractions for qRT-PCR analyses were obtained by sorting on the basis of APC signal (1:300, MWReg30). Cell populations for all experiments were isolated on Influx, AriaII or AriaIII (BD Bioscience) cell sorters.

2.4.2. Flow cytometry

Flow cytometry of reprogrammed cultures was performed using a 5-laser LSRII with Diva software (BD Biosciences) and analysed using FlowJo software (v 9.6.2). A minimum of $10x10^4$ cells were harvested, stained with an appropriate antibody mix for 30 minutes at 4^oC, washed (10% FCS/PBS) and if applicable stained with secondary antibody for another 30 minutes. Dead cells were excluded from the analysis on the basis of Hoechst 33258 uptake (1:100 of 1 μg/ml stock, Invitrogen), which was added to the cells with the primary antibody. Staining was performed with different combinations of the following sets of antibodies: GR1-PE-Cy7 (1:500, RB6-8C5), CD41-PE-Cy7 (1:300, MWReg30), CD41-PE (1:200, MWReg30), cKIT-APC (1:200, 2B8), CD71-biotin (1:2000, C2, BD Pharminogen), CD11b-APC eFluor780 (1:300, M1/70), CD3-APC (1:200, 145-2C11), CD4-PE (1:200, GK1.5), CD8-biotin (1:500, 53-6.7), CD16/32- PerCyP5*5 (1:400, 93), CD24-PE (1:200, 30-F1), BP1-biotin (1:100, 6C3), CD43-APC (1:200, S7, BD Pharminogen), B220-eFluor450 (1:500, RA3-6B2), IgM-FITC (1:100, II/41, BD Pharminogen), IgM-APC (1:200, II/41), IgD-biotin (1:100, 11-26c), streptavidin-BV421 (1:100), SCA1-PE-Cy7 (1:300, D7, Biolegend), CD150- PerCyP5*5 (1:100, TC15-12F12.2, Cambridge Bioscience), CD48-APC (1:500, HM48-1), CD135-PE (1:50, A2F10), CD34-eFluor660 (1:50, RAM34), CD45-PE-Cy7 (1:500, 30-F11), CD48-biotin (1:400, HM48-1), CD201-APC (1:50, 1560), isotype control IgG1κ (1:200, eBRG1), isotype control IgG2bκ (1:200, eB149/10H5), mouse haematopoietic lineage biotin panel (1:300) as well as conjugates used for FACS and immunostaining.
Unless stated otherwise, antibodies for FACS and flow cytometry were purchased from eBioscience.

2.5. *In vitro* functional assays

2.5.1. Colony-forming-unit count (CFU-C) assay

In order to evaluate the progenitor potential of the cells subjected to reprogramming, 10000 (five and two TF reprogramming) or 50000 (19-factor reprogramming) cultured cells were re-plated in triplicate in semi-solid media containing 55% methylcellulose (VWR) and supplemented with the spectrum of cytokines used in LBM (see paragraph 2.1.7). The cultures were maintained at 37ºC/5% CO_2 atmosphere and scored for the presence of haematopoietic colonies after 10-12 days. For the qRT-PCR analysis, colonies were picked from the plate, washed twice with PBS and subjected to standard RNA manipulation procedures.

2.5.2. Re-aggregation assay

Day 6 sorted cells were mixed with $10x10^4$ irradiated OP9 cells and re-suspended in 30 μl of IMDM containing 10% PDS, 1% P/S, 2 mM L-Gln, 1% KL conditioned medium, 1% IL3 condition medium and 100 ng/ml FLT3L. Aggregates of cells were generated after overnight hanging drop culture. The next day, the intact aggregates were transferred on the top of a 0.65 μm Durapore filter (Millipore) and cultured for another 4 days. After this time, a single cell suspension was obtained by dissociation with 1 mg/ml collagenase/dispase solution (Roche) and vigorious pipetting. Cells were re-plated into semi-solid culture according to standard haematopoietic CFU-C assay protocol.

2.5.3. ELISA

ELISA assay was performed on day 21 sorted F4/80$^+$ reprogrammed cells and BM-derived control macrophages (see paragraph 2.1.4). Purified cells (30x10^4) were seeded in duplicate into well of 96 well plate and stimulated with 20 ng/ml interferon γ (IFNγ, Peprotech) and 10 ng/ml lipopolysaccharide (LPS, Invivo Gen) for 12 hours. The plate was next span down at 108 g for 10 minutes and supernatant tested for the presence of IL6 and tumour necrosis factor α (TNFα) according to the protocol provided with Murine IL6/TNFα ELISA kit (Diaclone).In this line 100 μl of supernatant or standard (2 fold dilutions) were added to appropriate wells of 96 well plate. All samples were tested in duplicate. Next 50 μl of biotinylated anti-mouse IL6 or TNFα antibody was added to each well. Plate was incubated at RT for 3 h and washed three times with provided wash solution. Next each well was filled with 100 μl of streptavidin-HRP solution and incubated for 30 minutes. After three rounds of washing 100 μl of TMB substrate solution was added and after 20 minutes reaction was stopped with 100 μl of Stop reagent. Read of the absorbance was performed with GloMax®-Multi+ Microplate Multimode Reader (Promega) using 450nm wavelength.

2.5.4. Phagocytosis assay

Active phagocytosis of macrophages was tested with a use of fluorescent red conjugated carboxylate-modified polystyrene beads (Sigma). The volume of 10 μl of 0.5 μm diameter particles was added to 100 μl aliquots of day 21 CD45$^+$CD11b$^+$ sorted reprogrammed cells (3x10^6/ml) and incubated for 2 hours at 37^{o}C/5% CO_2 atmosphere with occasional shaking. Cells were washed twice with 10% FCS/PBS, to remove any beads that were not phagocytised, and subjected to standard cytospin procedure. Images were obtained with Leica DMI3000B microscope.

2.5.5. Cytospin analysis

For cytospin analysis 30000 cells suspended in 150 μl of serum containing medium were span at 72 g for 5 minutes with low acceleration onto a glass slide using Shandon Cytospin3 Centrifuge (Thermo Scientific). The cells were next fixed in methanol for 10 minutes, washed and stained with May-Grünwald dye (VWR) for 3 minutes. Upon few rounds of washing with water slides were subjected to Giemsa stain (1:20 dilution, VWR) for 20 minutes, rinsed and air-dried. Images were obtained with a Leica Scanner and subsequently analysed with SlidePath Gateway software.

2.6. *In vivo* functional assays

E14.5 MEFs for transplantation experiments were obtained from *p53*-null mice, GAF reporter mouse line exhibiting *Gfp* expression under the control of *AI467606* promoter (Ferreras et al., 2011) or MSBB reporter cell line (*Gfi1:dTomato*/*Gfi1b*$^{WT/GFP}$). Day 12 cKIT sorted GAF or *p53*$^{-/-}$ MEF-derived haematopoietic progenitors were injected intra-femorally into sub-lethally irradiated NSG recipients. Alternatively, day 12 reprogrammed GAF or MSBB MEFs were harvested and re-plated onto OP9-DL1 stroma with subsequent culture for 2 more weeks. Upon maturation step cKIT sorted cells were intra-venously co-injected with 50000 BM helper cells into lethally irradiated NSG recipients. The presence of engraftment in peripheral blood (PB) of mice was determined at weeks 2, 4, 6 and 8 posttransplantation using flow cytometry (GAF and MSBB MEFs) or PCR (*p53*$^{-/-}$ MEFs).

2.7. Bioinformatics and statistical analysis

2.7.1. Affymetrix analysis

Global gene expression analysis was performed with a use of Mouse Exon 1.0 ST array platform. Raw data were normalised using the Robust Multi-array Average (RMA) algorithm. Only probesets that were predicted to target a single exonic locus on the genome and have all four probes mapping to the

target exon were kept. These probesets were subsequently filtered by Affymetrix Detection Above Background (DABG) scores to remove probesets with signal intensities substantially below the background. A probeset was kept if its DABG p-value was less than 0.01 in at least one sample. Then gene-level expression summary was calculated for each gene represented on the Mouse Exon 1.0 ST array by averaging the normalised expression values of its exonic probesets. For processed data the R/Bioconductor package LIMMA was used to identify genes that were differentially expressed between the 3 conditions (untransduced, CDH5 positive and cKIT positives). 868 genes were found to be significant at 5% false discovery rate (FDR). The microarray was performed by the Molecular Biology core facility and the data normalization and analysis was performed by Hui Sun Leong from the Bioinformatics group at the CRUK Manchester Institute.

2.7.2. Pathway and Gene set enrichment analysis (GSEA)

GSEA was performed using JavaGSEA desktop application (v 2.2.0) provided by Broad Institute. The normalized expression data obtained for untransduced MEFs, day 6 $CDH5^+$ and day 12 $cKIT^+$ reprogrammed cells were analysed against previously published gene data sets generated for cells at different stages of haematopoietic development (Ivanova et al., 2002). GSEA analysis was performed with default settings (the enrichment score calculated with weighted statistics, statistical significance of the enrichment score assessed by 1000 gene set permutations, FDR cutoff of 25%).

The Ingenuity Pathway Analysis (IPA) tool from Ingenuity Systems was utilized to interrogate gene sets for enrichment of molecular and cellular functions. The analysis was performed on three groups of previously identified differentially expressed genes (paragraph 2.7.1). First analyzed data set contained genes downregulated between $cKIT^+$ and control populations. Second group was represented by genes downregulated during the transition of cells from CDH5 to cKIT positive state. Finally third group contained the pool of genes upregulated between untransduced MEFs and $cKIT^+$ haematopoietic progenitors. These differentially expressed genes were chosen as most informative for biological changes accompanying transition between MEFs

through $CDH5^+$ cells to haematopoietic progenitors. I used Downstream Effects Analysis (DAE) tool which predicts downstream biological processes that are increased or decreased as a result of changes in transcriptome. Analysis was performed with default settings (Core Analysis). The Z score cutoff (correlating with the level of up (negative value) or downregulation (positive value) of biological activity) was >I1.5I.

2.7.3. Statistical significance

Statistical analysis was majorly performed using Student t-test or Paired Student t-test. One way ANOVA test was applied for experiments where more than two groups were compared (Fig 4.7A and S3C). Where significant ($p<0.05$) the follow up Dunnett's multiple comparison test was performed to identify significantly different conditions. Values considered statistically significant are marked with * for $p< 0.05$, ** for $p< 0.005$, *** for $p< 0.0005$ and **** for $p< 0.0001$). Data are displayed as mean ± standard deviation (SD) or mean ± standard error of the mean (SEM).

Chapter 3

3. Reprogramming of fibroblasts to haematopoietic lineage

3.1. Introduction

The ability to reprogram somatic cells to pluripotency or alternative cell type has challenged the long-standing scientific dogma that differentiation is unidirectional and irreversible process. The ground-breaking discovery of iPSCs became a turning point for regenerative medicine and opened new avenues for the generation of patient specific cells for transplantations (Takahashi and Yamanaka, 2006). Nevertheless, subsequent studies revealed events of partial reprogramming and substantial genome instabilities in iPSCs, which impede the use of induced pluripotency for patient treatment (Gore et al., 2011; Ohnishi et al., 2014). Such concerns prompted investigations into alternative ways for obtaining functional effector cells for medical use. One alternative strategy, that is the focus of this thesis, is based on the direct conversion of one cell type to another by overexpression of lineage-instructive TFs. The seminal study performed on transdifferentiation to myoblasts identified a single regulator MYOD that instilled alternative fate on fibroblasts (Davis et al., 1987). This result was the basis for the driver TF theory that propose the existence of a master regulator, specific for each lineage, which would be able to single-handedly drive switching of cellular fate decisions. Multiple TFs fitting this description have been reported since, including CEBPα for the conversion within haematopoietic system (Bussmann et al., 2009), ASCL1 (Chanda et al., 2014) and NEUROD1 (Guo et al., 2014) for switching to neuronal fate or NGN3 (Yechoor et al., 2009) for the generation of pancreatic cells. The outcome of these transitions is, however, limited to terminally differentiated progeny and in many cases allows switching only between related lineages. On the contrary, multiple reports indicate that the generation of more primitive stem/progenitor populations from developmentally distinct cells types requires a broader

repertoire of TFs (Batta et al., 2014; Yu et al., 2013). This doesn't preclude the need for a pioneer factor that would bind to the regulatory regions of lineage instructive genes and trigger initial chromatin remodelling facilitating access of other modulators. This role is played by OCT4 in iPSC generation (Buganim et al., 2013) or ASCL1 during conversion of fibroblasts to neurons (Wapinski et al., 2013). However, the full transition seems to be orchestrated by multiple cooperating regulators sequentially instilling new fate characteristics on the cells of origin.

Interestingly, established transdifferentiation protocols use different fate converters and methodology to obtain a similar lineage outcome (Guo et al., 2014; Heinrich et al., 2010; Vierbuchen et al., 2010). This clearly shows that the plasticity of terminally differentiated cells is far broader than previously anticipated. There seems to be also a notable difference in the ease to generate different lineages of interest. For example, whereas cardiomyocytic and neuronal fates seem to be instructed on developmentally distinct cell types quite easily, there are only few reports describing conversion to lymphocytes or hepatocytes. Both of these cell types are known to fulfil variable functions in the organism, which may be associated with more complex transcriptional network governing their ontogeny. Conceptually, the simpler the regulatory circuits involved in specific lineage specification, the easier it should be to evoke fate transition by means of induction of only limited number of master regulators.

For many years, the haematopoietic system was studied as a model of cellular hierarchy and specialization of cell potential. Starting from HSCs, blood cells are subjected to multiple extrinsic and internal signals that ultimately define their blood effector functions. Even though highly structured, the haematopoietic differentiation is a truly complex process in which balanced expression of plethora of TFs is required on multiple levels. The alternative cell fates are selected for on the basis of the ratio of lineage-affiliated regulators, their sequential appearance as well as temporal and spatial expression patterns (Dahl et al., 2007; Iwasaki et al., 2006; Zhang et al., 2000). One very well described example of TFs' instructive function is the pair of PU.1 and GATA1 that alternatively promote myeloid and erythroid fates at the level of multipotent blood progenitors (Arinobu et al., 2007). Existence of multiple tightly interconnected regulatory circuits governing blood development make it highly

challenging to select for the limited number of haematopoietic TFs that would be able to induce haematopoietic fate change in alternative cell types. Accordingly, the first successes in generating blood progenitors by transdifferentiation of cells from the same germ layer (myeloid and lymphoid cells or endothelium) were reported only recently (Riddell et al., 2014; Sandler et al., 2014). It still remains very challenging to convert more developmentally distant but easily accessible cells, like fibroblasts, to blood stem and progenitor cells.

The aim of the work presented in this chapter was to select candidate haematopoietic TFs that would drive blood developmental programme in adult murine fibroblasts. Based on data curated from the literature, I selected a pool of 19 TFs that were reported to play crucial roles during the establishment of primitive and definitive waves of haematopoiesis as well as fine-tuning of adult blood system. These candidate TFs were then overexpressed in adult fibroblasts and, following a subtraction screen, selected for the minimum pool sufficient to drive reprogramming to blood.

3.2. Results

3.2.1. Introduction of experimental strategy

Recent interest in the field of transdifferentiation resulted in the development of a large range of protocols and strategies. Despite the plethora of approaches some common features have also started to appear. In general, the generation of the desired lineage takes between 21 and 40 days and is carried out in selected supportive medium supplemented with lineage-specific growth factors. In some cases a complementary niche is also provided in order to mimic the endogenous growth environment (Riddell et al., 2014). Finally, the maximum efficiency of delivery of TFs is in general achieved by viral transduction. On the basis of these common characteristics, I designed the experimental strategy of reprogramming to haematopoietic lineage depicted in Fig 3.1A. MAFs were obtained from skin of 2 to 6 months old mice, whereas MEFs were derived from E14.5 mouse embryos. Only low passage cells were used to avoid senescence. One main concern in reprogramming experiments is

the potential contamination of the starting material with cells of the desired lineage. To eliminate this possibility, I stained fibroblasts with biotin coupled CD41, CD45, cKIT and CD31 antibodies and excluded cells positive for these markers by cell sorting (Fig 3.1Bi). The subsequent sort purity check confirmed the removal of haematopoietic and endothelial cells (Fig 3.1Bii). Purified cells were next seeded on gelatine-coated dishes and transfected with the viruses encoding haematopoietic TFs. After four hours, the viral supernatant was removed and replaced by a haematopoiesis supporting medium. Cultures were monitored for change in morphology during the following three weeks. At day 21, the cultures of transduced cells were assayed for the presence of haematopoietic cell surface markers, blood cell morphologies and expression of haematopoietic genes.

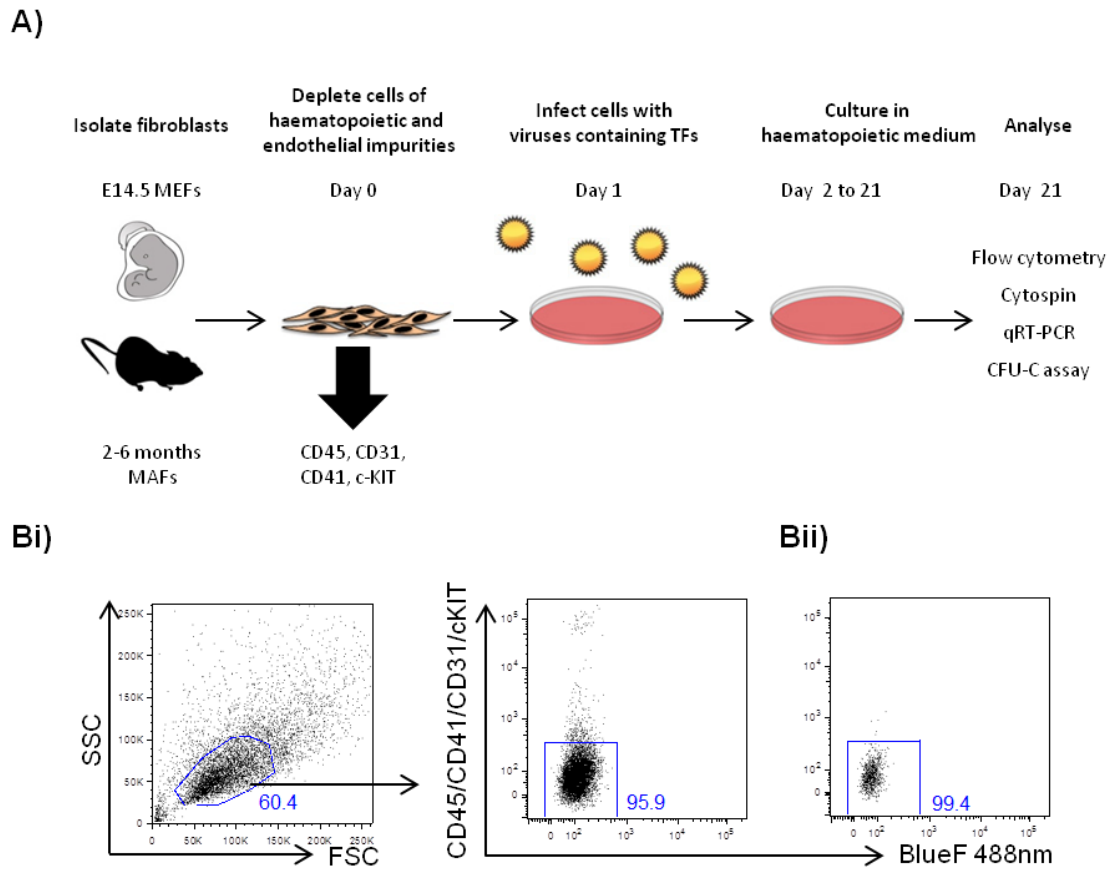


Figure 3.1. Schematic representation of experimental procedure.

A) MEFs were isolated from E14.5 embryos and MAFs were generated from ear skin of 2 to 6 months old mice. Cells expressing CD41, CD31, CD45 and cKIT were removed from the culture. Sorted fibroblasts were then transduced with cocktail of TFs and cultured in haematopoietic medium for up to 21 days. Reprogrammed cells were analysed for features of blood cells (morphology, cell surface markers, gene expression). **Bi)** Representative flow cytometry plot of day 0 sort for CD41/CD31/CD45/cKIT negative MAFs. **Bii)** Purity check of sorted negative fraction. FSC – forward scatter, SSC –side scatter.

3.2.2. Validation of infectious system

The 19 transcriptional regulators chosen for this study (Fig 3.2A) play various roles in the haematopoietic system (reviewed by Orkin and Zon, 2008). Some of them (CEBPα, PU.1, GATA1) were already identified as cell fate converters (Feng et al., 2008; Iwasaki et al., 2003). Others like ETS1, SCL, GATA2 and FLI1 have recognized roles in the establishment of myeloid, megakaryocytic erythroid and lymphoid lineages (Barton et al., 1998; Bories et al., 1995; Chou et al., 2009; Hart et al., 2000; Wadman et al., 1997)). Moreover, the haematopoietic triad (SCL, GATA2, FLI1) and RUNX1 are also well characterised key players at the onset of haematopoiesis (Chen et al., 2009; Pimanda et al., 2007). GFI1 and GFI1b are active already during the cluster formation in the AGM (Thambyrajah et al., *submitted*) but also differentially characterise effector myeloid/lymphoid and erythroid/ megakaryocytic cells respectively (reviewed by van der Meer et al., 2010)). Overall, the selected 19 factors play crucial roles at various levels of blood system development and maintenance, as demonstrated by numerous *in vivo* knock-out and knock-down studies (Gandillet et al., 2009; Karsunky et al., 2002; Koyano-Nakagawa et al., 2012; Lee et al., 2008; Loughran et al., 2008; Okuda et al., 1996; Saleque et al., 2002; Shivdasani et al., 1995; Tsai et al., 1994).

The cDNAs of selected TFs were amplified and cloned into a lentiviral vector, which did not contain any reporter to maximize the efficiency of packaging and expression. All sequences included a FLAG-tag to allow the detection of the expression of the proteins and a KOZAK sequence to optimise translation efficiency (Fig 3.2B). I first validated the functionality of the produced viruses by infecting MAFs with each single virus and detecting FLAG-tag by Western Blot. In contrast to the negative control GFP virus, transduction with infectious particles for each TF led to a significant expression of the protein of interest (Fig 3.2C). Next, to examine the correct cellular localization of each TF, HEK293T cells transduced with individual viruses were fixed and the TF protein visualised with an anti-FLAG antibody. Simultaneously nuclei were stained with DAPI and actin filaments with phalloidin-Alexa488. As shown for some selected examples, such as SCL, RUNX1c and HOXB4, the introduced exogenous

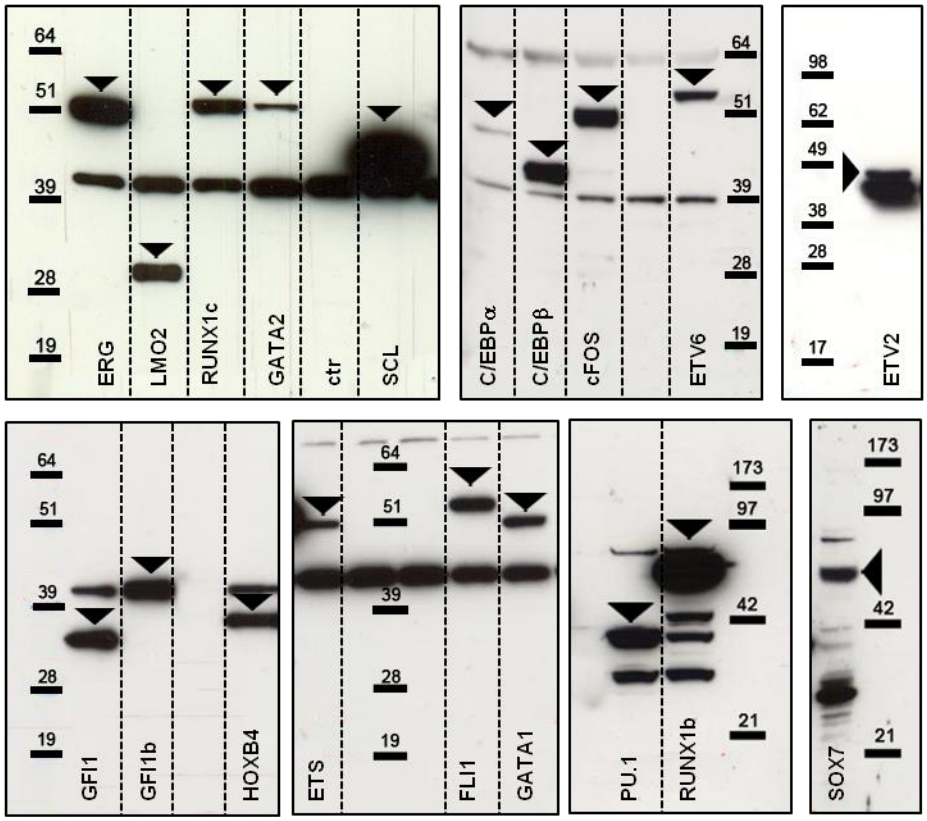
factors were correctly localized in the nuclei (Fig 3.2D). Taken together these experiments validated the ability of lentiviral vectors to efficiently deliver reprogramming TFs to cells of interest.

A)

Gene Name	Gene ID	Gene Name	Gene ID
CEbpα	NM_007678.3	*Gfi1*	NM_010278.2
CEbpβ	NM_009883.3	*Gfi1b*	NM_008114.2
cFos	NM_010234.2	*Hoxb4*	NM_010459.6
Erg	NM_133659.2	*Lmo2*	NM_008505.3
Ets1	NM_001038642.1	*Pu.1*	NM_011355.1
Etv2	NM_007959.2	*Runx1b*	NM_001111022.2
Etv6	NM_007961.3	*Runx1c*	NM_001111023.2
Fli1	NM_008026.4	*Scl*	NM_011527.2
Gata1	NM_008089.1	*Sox7*	NM_011446.1
Gata2	NM_008090.4		

B)

C)



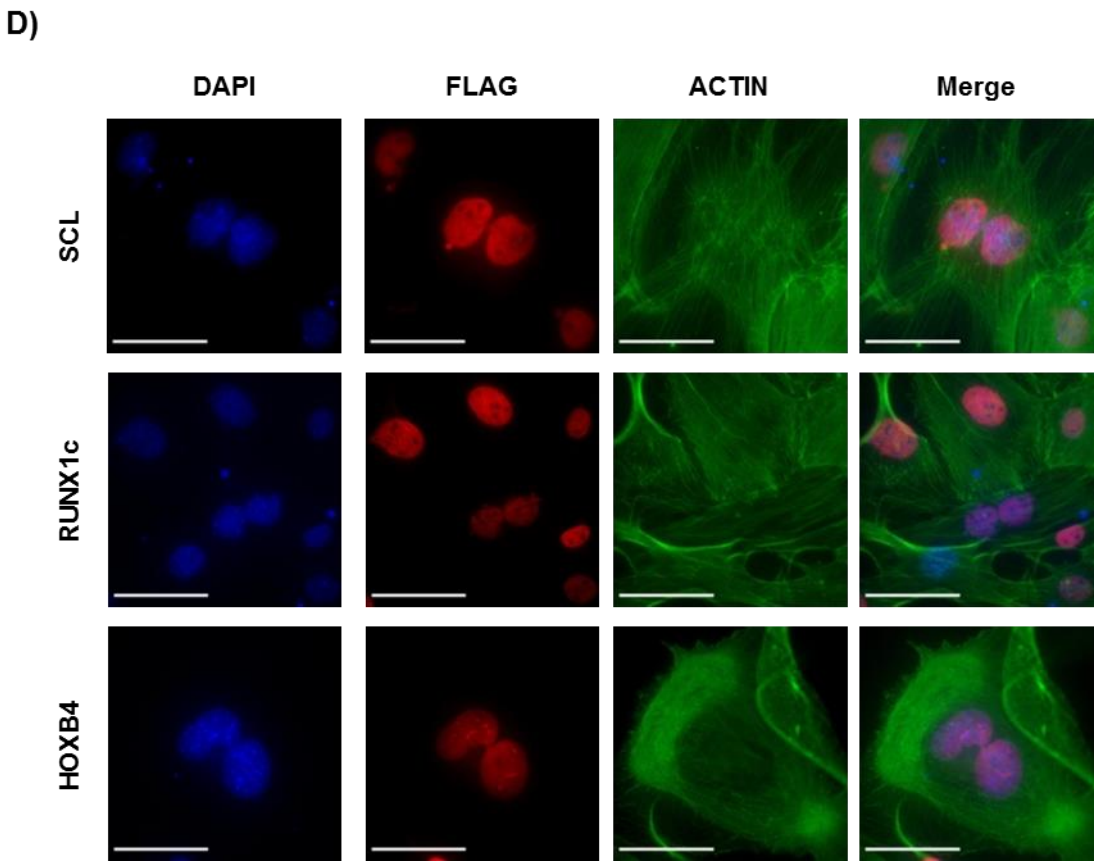


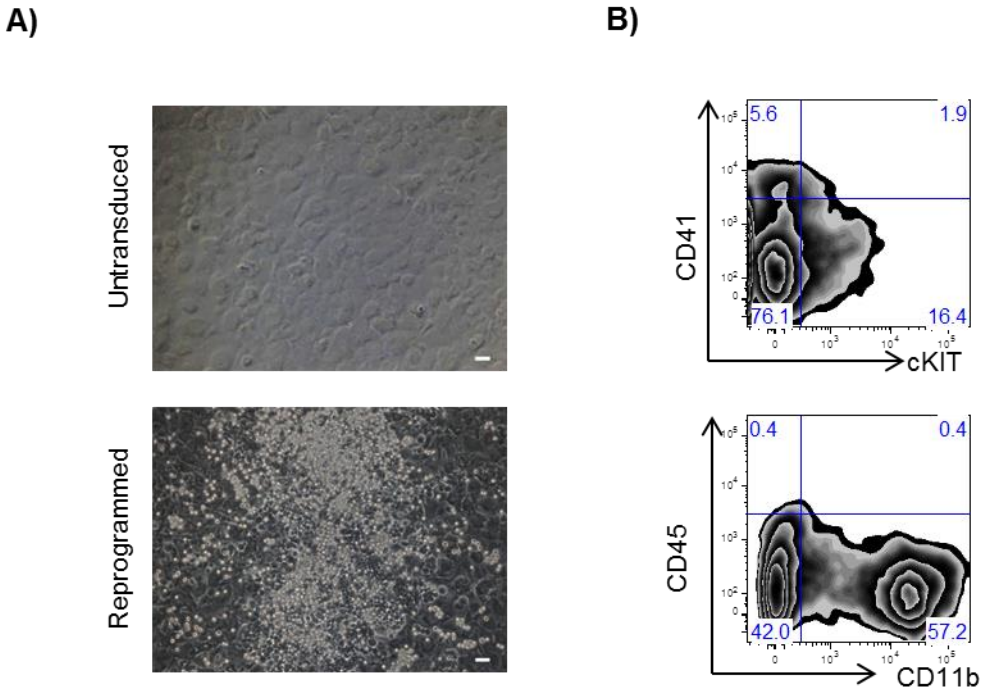
Figure 3.2. Validation of lentiviral infectious system.

A) List of haematopoietic factors used for cell fate conversion. **B)** Schematic representation of lentivirus transgene expression cassette. **C)** TFs' protein expression in cells infected with single lentiviruses. Visualization by the detection of FLAG–tagged protein in the cellular whole-protein extracts. Protein mass ladder in kiloDaltons. Arrows indicate the protein band of interest. **D)** Nuclear localization of three representative TFs overexpressed in HEK293T cells. Nuclear staining performed with DAPI, actin filaments bound by phalloidin-Alexa488 and a FLAG-tagged protein visualized by Alexa647 fluorochrome. Scale bars represent 50 μm. EF1α - Elongation factor 1α promoter, LTRs – long terminal repeats, ctr - control.

3.2.3. Reprogramming of MAFs by 19-factor cocktail of haematopoietic TFs

Upon transduction of the pool of 19 lentiviruses, fibroblasts were monitored for changes in morphology manifesting successful reprogramming events. By day 21 of experiment, I observed the generation of grape-like packed colonies of small round cells in transduced but not control MAFs (Fig 3.3A). Flow cytometry analysis of transduced experimental wells confirmed the appearance of new cell population with a specific size and granularity that

could not be detected in control cultures (Fig S1A). To better characterize originating cells, transduced cells were assayed for the presence of first embryonic haematopoietic cell surface markers CD41 and cKIT as well as the myeloid differentiation marker CD11b. Flow cytometry analysis confirmed the expression of all these surface signalling molecules on reprogrammed cells (Fig 3.2B). The myeloid population was most substantial and accounted for 57% of the cells. Concomitantly, I could not detect CD45 expression that is also associated with monocytes/macrophages. To investigate the presence of haematopoietic progenitors, I next performed CFU-C assay on the whole pool of day 21 harvested cells. Unlike control fibroblasts, the 19-factors transduced cells gave rise to myeloid colonies in semi-solid culture conditions indicating the presence of haematopoietic progenitors (Fig 3.3C). Although I did not detect any erythroid or megakaryocytic colonies in CFU-C assays, generation of these lineages was confirmed by the cytospin analysis of day 21 cultured cells (Fig 3.3D).



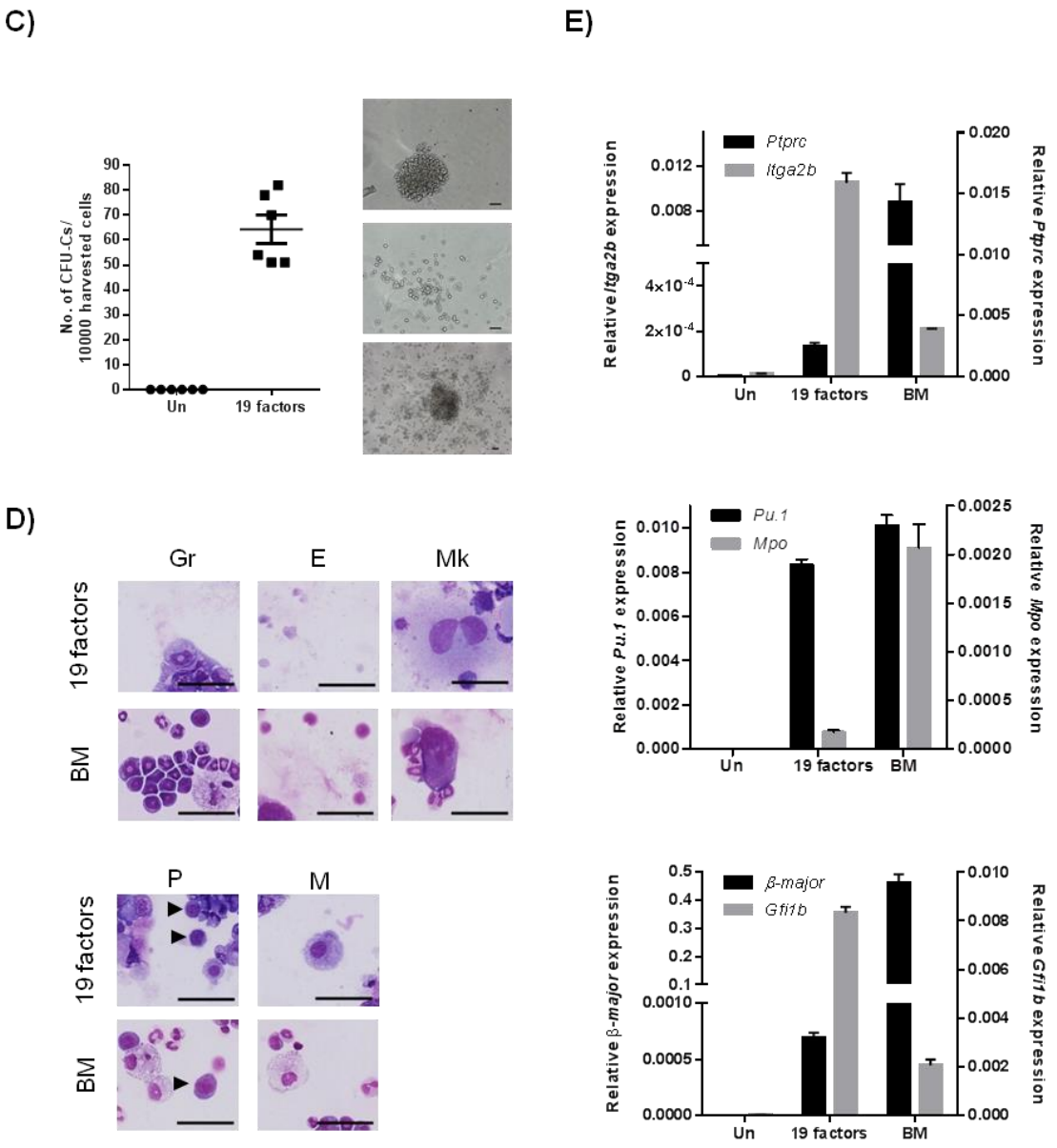


Figure 3.3. Reprogramming of MAFs with 19-factor cocktail.

A) Representative bright field images of untransduced and reprogrammed day 21 MAFs. **B)** Representative flow cytometry plots of day 21 reprogrammed MAF cultures showing surface expression of haematopoietic markers. **C)** The number of CFUs generated in semi-solid medium from 10000 day 21 19-factor transduced MAFs (left) and bright field images of representative colonies (right) (mean ± SEM , n=2). **D)** May-Grünwald staining of day 21 MAFs transduced with 19 TFs in comparison to BM-derived control cells. Morphologies were assigned to specific lineages on the basis of Atlas of haematological cytology (Hayhoe et al., 1982). Arrows indicate location of progenitors. **E)** Relative haematopoietic gene expression in control fibroblasts (Un) ,reprogrammed MAFs (19 factors) and BM with respect to *β-actin*. For *Pu.1* and *Gfi1b* total transcripts (exogenous and endogenous) were measured. Data shown are mean ± SD from a single experiment performed in triplicate (n=2). Scale bars represent 50 μm. E – erythrocyte, M – macrophage, Gr- granulocyte, P – progenitor, Mk – megakaryocyte.

Importantly, reprogrammed cells represented a wide spectrum of blood lineages and exhibited morphology similar to the equivalent cells derived from murine BM. Finally, the presence of blood cells was also confirmed by gene expression analysis. Transcripts for myeloid (*Pu.1, Mpo*), erythroid (*β-major, Gfi1b*) and signalling molecules (*Ptprc* encoding CD45*, Itga2b* encoding CD41) were readily detected in the transduced but not control cultures (Fig 3.3E). Altogether, these experiments indicate that the pool of 19 TFs can induce haematopoietic programme in fibroblasts.

3.2.4. The minimal set of five TFs is sufficient for reprogramming of MAFs

I next sought to determine which factors from the initial pool of 19 regulators are not essential or inhibitory for the efficient reprogramming of MAFs. For this, I performed an experiment in which individual TFs or groups of factors, generated on the basis of structural similarities and potential compensatory functions, were removed from the starting mix. In this respect, I grouped ETS1, ERG and FLI1, as closely related ETS factors with pointed domain, and paired GFI1 with GFI1b, RUNX1b with RUNX1c, CEBPα with CEBPβ and GATA1 with GATA2. MAFs transduced with these combinatorial mixes were assayed for the number of colonies emerging on a plate at day 12 of reprogramming (Fig 3.4A). Consequently, I identified five groups of TFs the lack of which resulted in the most significant decrease in the efficiency of transition. The critical factors included SCL, LMO2 and members of the RUNX, GATA and ETS groups. Noteworthy, even though these factors were chosen as indispensible for the transition, the character of generated colonies was not determined.

The overarching goal of this project is to generate blood progenitors for transplantation. Conceptually, in order to reprogram cells to multipotent state the factors employed should be involved in the early blood fate determination. I tried to use this criteria to identify the most pertinent reprogramming factor within preselected groups of TFs. ERG, FLI1 and ETS1 are known to bind to the same consensus DNA sequence and to the extent compensate for each other's functions (reviewed by Ciau-Uitz et al., 2013). Notably, the whole genome

binding profile of HSPCs revealed that the core regulatory complex bound to the majority of haematopoietic genes includes ERG and FLI1 (Wilson et al., 2010a). Among the two FLI1 is crucial for the terminal differentiation of megakaryocytes and subsequent platelet formation (Hart et al., 2000). Based on the fact that FLI1 is involved in terminal specification of haematopoietic lineages, I chose ERG for further optimization. Experiments performed later did not reveal any significant differences in the potency of ERG and FLI1 to reprogram fibroblasts to blood. More particularly the day 21 reprogrammed cells presented the same range of blood lineages regardless of whether the ETS factor involved was ERG or FLI1 (Fig S1B).

The distinct expression patterns and knock-out phenotypes of GATA1 and GATA2 indicate their independent contributions to haematopoietic development. In particular, GATA2 is characteristic of the HSPC stage whereas GATA1 drives erythroid lineage commitment. It is well established that GATA2, together with SCL and FLI1, form an early haematopoietic regulatory circuit, crucial during the specification of HSCs (Pimanda et al., 2007). Moreover, it was shown that GATA2 still occupies erythroid genes in the early erythroid precursors and is subsequently replaced by GATA1 in a process termed ‘GATA factors switching’, which facilitates further red blood cell development (Suzuki et al., 2013a). This wide span of regulatory functions of GATA2 at the early stages of blood formation and specification prompted me to choose this TF for further validation.

The final group of TFs consisted of RUNX1c and b, the expression of which is driven by two alternative promoters – P1 (distal) and P2 (proximal) respectively. It was established that these two isoforms are differentially expressed with P2 mRNA being dominant until E10.5 and therefore involved in EHT (Sroczynska et al., 2009a). Subsequently, expression pattern is skewed towards P1 driven mRNA that remains prevalent throughout adult life. Given that distal isoform is the most abundantly expressed in the adult haematopoiesis and can majorly complement for the other variant’s functions, I chose RUNX1c for further experiments.

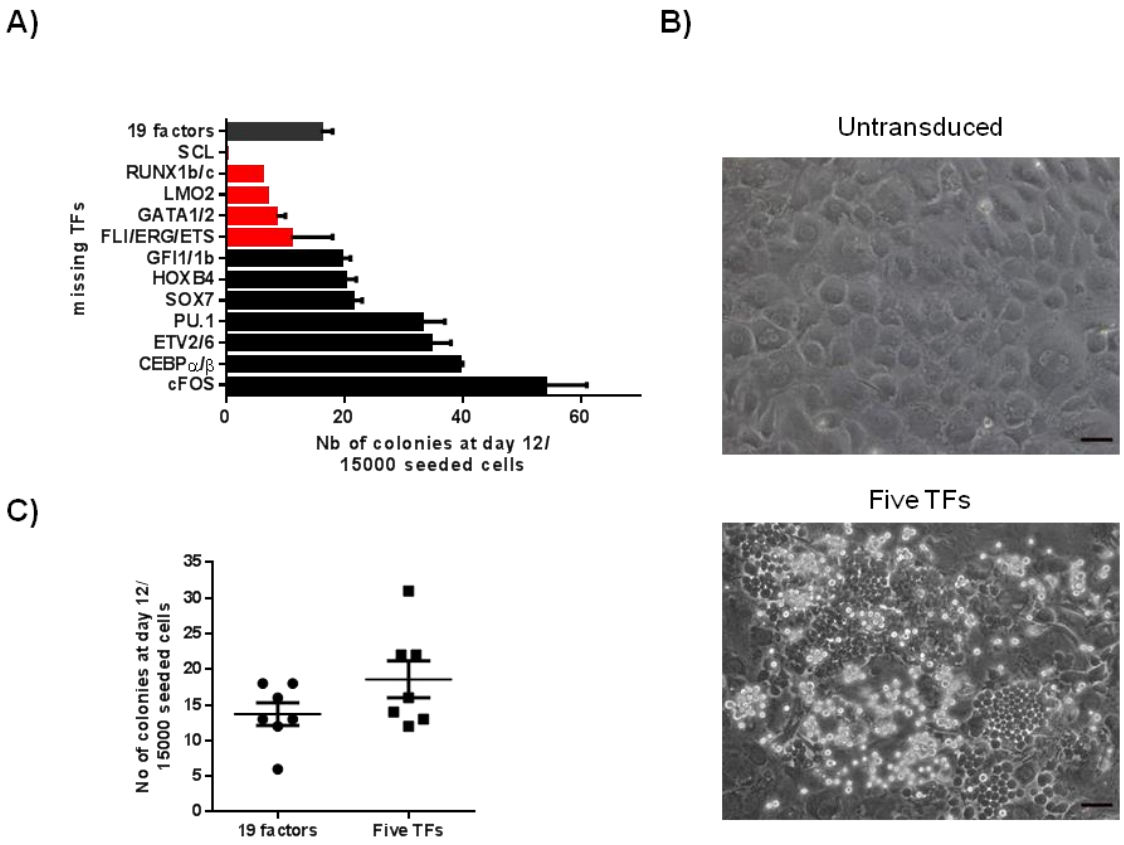


Figure 3.4. Identification of factors sufficient for fate conversion of MAFs.

A) Change in the number of reprogrammed colonies present on day 12 of transdifferentiation in MAF cultures transduced with indicated combinations of TFs. Data are presented as mean ± SEM of duplicate from a single experiment. **B)** Bright field images of untransduced and five-factor (SCL, LMO2, ERG, RUNX1c, GATA2) infected MAFs at day 21. **C)** The number of blood colonies generated at day 12 of reprogramming upon transduction of MAFs with five factors or 19-factor cocktail (mean ± SEM , n=2). Scale bars represent 50 μm.

Altogether, performed selection process led to the identification of a limited set of five factors (SCL, LMO2, RUNX1c, GATA2 and ERG) that could potentially reprogram MAFs to blood. To establish the functionality of this limited mix, I tested it against the original pool of 19 TFs. By day 21 the colonies of round cells exhibiting previously seen blood morphology were visible under both tested conditions. Importantly, the five-factor transduced MAFs generated also adherent clusters of cells exhibiting cobblestone morphology usually associated with progenitors in *in vitro* cultures (Fig 3.4B). The number of colonies generated with the set of five TFs was slightly higher but not significantly different from the 19-factor variant of experiment. Altogether, these data reveal that mere five factors can successfully reprogram MAFs to blood.

3.3. Discussion

The experiments presented in this chapter have explored the possibility of reprogramming adult fibroblasts toward haematopoietic fate. Since these studies were initiated significant progresses in transdifferentiation to blood has been reported by other groups (Fig 1.5 and 3.5). Several laboratories recently achieved the generation of inducible HSPCs from various human or mouse cell types. I will therefore discuss my results in the context of these recent reports while trying to propose possible reasons for discrepancies in observed results.

The cells generated in all mentioned studies exhibit clear blood cells characteristics but differ in biological potential. This might be reflective of the very diverse sets of TFs used in each experiment. Doulatov, Riddell and Sandler all chose their candidate reprogramming TFs on the basis of differential expression between HSCs and terminally differentiated blood cells. Still, after the screening process, there is no large overlap between the functional mix identified by each group. This is most likely attributed to the fact that distinct starting cell type namely endothelial cells (Sandler et al., 2014), committed blood cells (Riddell et al., 2014) or haematopoietic progenitors (Doulatov et al., 2013), will have variable transcriptional requirements. Interestingly, our core of five TFs (SCL, LMO2, ERG, GATA2, RUNX1c), is at the intersection of all the studies regardless of species, developmental stage and cellular identity of the cells of origin.

The initial set of 19 haematopoietic factors for transdifferentiation of MAFs was chosen on the basis of their established role in haematopoiesis. Because many of the selected factors are instructive in blood development, this approach was supposed to ensure immediate and firm induction of haematopoietic fate. Indeed already by day 12, well defined grape-like colonies could be detected. This is a significantly quicker outcome than one reported by Pereira and colleagues, where the first morphological signs of fate conversion were detected not earlier than 20 days after transduction of MEFs with four factors (GATA2, GFI1b, cFOS, ETV6) (Pereira et al., 2013). Moreover, the generated cells exhibited solely myeloid morphology and only marginal *in vitro* progenitor potential upon re-aggregation with PL cells. In contrast, MAFs

transduced with 19-factor mix, which notably also contained the four TFs selected by Pereira, were converted to cells with multilineage blood potential.

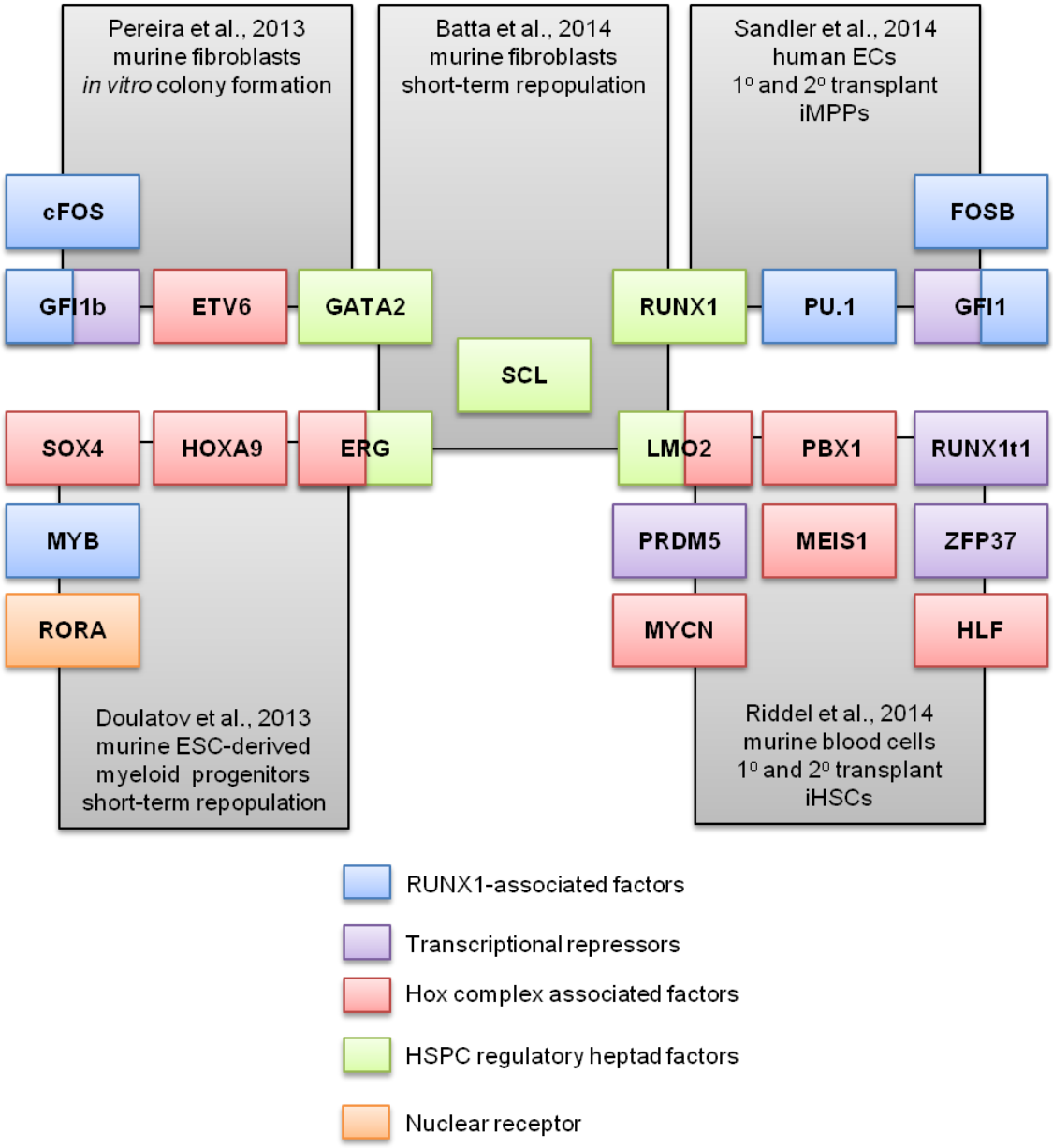


Figure 3.5. Schematic representation of recent advances in transdifferentiation to blood.

Five reports published within the last two years identified a whole plethora of TFs suitable for reprogramming of different cell types to haematopoietic lineage. This scheme gathers these recent advances including characteristics of selected TFs and biological outcomes of the respective studies.

The minimal mix selected in both studies share only GATA2. Moreover, whereas GFI1b and ETV6 were considered neutral for reprogramming of MAFs, cFOS was strongly inhibitory for the transition (Fig 3.4). These contrasting results could originate from a different transcriptional landscape provided in both settings, as only seven factors were overlapping between the two screening protocols. It is known that many haematopoietic TFs antagonize each other's functions and act on mutually exclusive basis, which in this complicated overexpression environment could provide a barrier for obtaining plasticity. To test this hypothesis, I transduced MAFs with GATA2, GFI1b, cFOS and ETV6 only, but I could not detect any reprogramming events up to 40 days post transduction. This lack of expected outcome may be attributed to the different developmental stages of the initial source of cells. Adult fibroblasts might have different reprogramming factors requirements than their embryonic counterparts (also discussed in chapter 4). Moreover, whereas Pereira and colleagues used retroviral vectors I introduced TFs of interest through lentiviral infection. The right stoichiometry of TFs was found to be important in both iPSC generation and transdifferentiation towards cardiomyocytes (Tiemann et al., 2011; Wang et al., 2015). Therefore, the use of different viral delivery system and driver promoter as well as variable multiplicity of infection (MOI) of viruses could have contributed to the variability in the overall pattern of transgene expression and my difficulty to reproduce these results.

The seminal work on reprogramming of human endothelial cells to iMPPs has been performed with a use of FOSB, GFI1, PU.1 and RUNX1 (Sandler et al., 2014). Yet again, three of these factors were present in the original 19-factor mix used in my study, however only RUNX1 was selected for the reprogramming of MAFs. In this case the main reasons for observed divergence in reprogramming mix are probably species and cell type differences. Even though mammalian TFs show high level of structural and functional homology, some discrepancies in binding patterns have been reported. For instance, the four highly conserved hepatocyte specific TFs (Forkhead Box A2 (FOXA2) Hepatocyte Nuclear Factor 1A, 4A and 6 (HNF1A, 4A and 6)) display a high divergence of binding sites with only one third of promoters being similarly occupied in human and mouse (Odom et al., 2007). The comparative transcriptome studies were also performed for different stages of erythropoiesis

(Pishesha et al., 2014) and cells of immune system (Shay et al., 2013). Both reports revealed that the expression pattern of lineage-instructive regulators is largely conserved between species. The notable divergence in global transcriptional profile has been in turn attributed to the different degree of interaction of these master blood regulators with their target loci. The modification of response rate of downstream targets can be associated with the level of promoter conservation and species-specific properties of distal enhancers. The most extensively studied example of different functional outcomes of TF's activity among species is HOXB4. The overexpression of this factor in murine HSCs confers self-renewal and proliferation advantage on treated cells with only minimal cancer transformation rate (Antonchuk et al., 2002). However, this expansion boost is majorly lost when human $CD34^+$ cells are subjected to the effect of HOXB4 secreting stroma (Amsellem et al., 2003). Similarly, HOXB4 was found to confer haematopoietic repopulation potential on mouse ESC-derived blood cells, whereas no such effect was observed in human counterparts (Kyba et al., 2002; Wang et al., 2005). These findings raise the question whether straightforward implementation of haematopoietic reprogramming TFs discovered in murine system to human environment will give the same fate switching outcome. It has been shown for induced neurons (Son et al., 2011; Caiazzo et al., 2011) or iPSCs (Takahashi et al., 2007; Takahashi and Yamanaka, 2006) that the same cocktail of TFs can drive transition in both species. Studies performed so far on transdifferentiation to blood identified large pool of haematopoietic TFs enabling transition (Fig 3.5), however no concomitant reprogramming of human and mouse cells with the same set of regulators has been reported. Still careful re-evaluation of selected factors on the basis of available data and implementation of 'mix and match' strategy may result in significant progress in the generation of human HSCs from fibroblasts.

The studies described above provide evidence for successful reprogramming to blood of developmentally and functionally distinct cell types such as endothelium or fibroblasts (upper part of Fig 3.5). Even though the selected reprogramming TFs overlap by only one factor, all of them are firmly related to blood development. This suggests that inducing an alternative cellular fate requires set of TFs that are similar to those involved in the embryonic

specification of haematopoiesis. In contrast, when the reprogramming aims to achieve fate switching within blood lineage, the selected factors are more closely associated with development, stemness and self-renewal (Doulatov et al., 2013; Riddell et al., 2014) (lower part of Fig 3.5). In this case the 'blood cell memory' and major components of haematopoietic gene regulatory networks (GRNs) may still be present and accessible in blood effector cells that only require to be reversed to stem/progenitor state by the arrest of differentiation signature. That would explain the wide range of Hox complex members and transcriptional repressors identified in reprogramming of myeloid, lymphoid and $CD34^+$ blood cells (Doulatov et al., 2013; Riddell et al., 2014). Namely, HOXA9, PBX1 and MES1, known to interact physically, have all been implicated in limb and eye regeneration as well as cancer development, implying the basal role of this complex in conferring stem/progenitor characteristics (Chen et al., 2012; Mercader et al., 2005; Roensch et al., 2013; Shen et al., 1997). Moreover, members of Hox complex have been found to regulate expression of multiple haematopoietic targets like ERG, LMO2 or MYB through interaction with their enhancers (Huang et al., 2012). Affecting expression levels of haematopoietic TFs may directly lead to the restoration of specific blood GRNs and re-specification of cellular fate. Furthermore, Riddell and colleagues identified also several transcriptional repressors, namely PRDM5, RUNX1t1 and ZFP36, with so far unrecognized roles in haematopoiesis. This suggests that fine balance between activation but also suppression of crucial targets may be necessary to obtain truly functional HSCs.

The TFs identified in studies discussed above fall into different functional categories. However, RUNX1 and its direct transcriptional targets or binding partners have been selected in every single report. RUNX1 is known to be indispensible for the specification of blood cells through facilitation of EHT in the AGM (Chen et al., 2011; Chen et al., 2009; Eilken et al., 2009; Lancrin et al., 2009). The initial loss of endothelial identity of HE has been attributed, at least partially, to the action of GFI1 and GFI1b, the expression of which is regulated by RUNX1 (Lancrin et al., 2012). The subsequent establishment of haematopoietic program involves activity of PU.1, the well-established direct transcriptional target of RUNX1, whose balanced expression is required for correct myeloid and lymphoid development (Huang et al., 2008; Rhodes et al.,

2005). At the protein level, RUNX1 co-operates with multiple binding partners including Core Binding Factor β(CBFβ) and AP1 complex (Ogawa et al., 1993; Pencovich et al., 2011; Wang et al., 1996). Notably, the FOS subunit of the latter has been selected in studies of Sandler and Pereira for reprogramming of non-haematopoietic cells to blood. Altogether, the RUNX1 associated transcriptional landscape seems to play crucial role in inducing HSPC fate both in human and mouse cells of various identities.

In the context of all the other findings the five TFs identified as drivers of the conversion of adult fibroblasts to blood create a coherent group of regulators. SCL, RUNX1, LMO2, ERG and GATA2 are members of the heptad complex together with FLI1 and LYL1 (Wilson et al., 2010a). They were found to build the core transcriptional machinery occupying the majority of haematopoietic genes in HSPCs. One of the identified target genes for this collaborative activity is *Gfi1*, the expression of which is mediated by -35 enhancer through identified SCL, RUNX1, GATA2 and ERG binding sites (Wilson et al., 2010b). Importantly, the co-occupancy sites of SLC/RUNX1/FLI1/ GATA1/2 were also detected in the primary human megakaryocytes, suggesting that this combinational activity is also present in normal cells in addition to cell lines (Tijssen et al., 2011). The whole-genome binding data revealed that even though the majority of sequences bound in HPC-7 cell line by SCL/RUNX1/LMO2/ERG/GATA2/FLI1/LYL1 heptad of factors consist of consensus motifs for all of the molecules, the most frequent sites are recognized by GATA and ETS factors, whereas Runt sequence is significantly underrepresented. This consequently led to the identification of SCL, ERG and GATA2 as the new binding partners of RUNX1 (Wilson et al., 2010a). Moreover, the members of the heptad were also found to mutually regulate each other's expression. The -75 enhancer of *Lmo2* was identified as a driver of haematopoiesis specific expression of the factor. The transactivation experiments confirmed that this process requires a positive feedback loop coming from GATA/LMO2/SCL complexes and pre-existing ETS factors-dependent promoter activity (Landry et al., 2009). Similarly, the +19 *Scl* enhancer, functional in HSPCs, is highly dependent on Ets/Ets/Gata motif (Gottgens et al., 2002; Silberstein et al., 2005). The *Runx1* +23kb enhancer, active during HSC emergence, was found to be bound and regulated by SCL,

LMO2, GATA2 and RUNX1 itself (Nottingham et al., 2007). This tight interconnection suggests that further decrease in the number of reprogramming TFs may be possible if the right positive feedback loops can be evoked by the remaining regulators (also discussed in chapter 4). Among five reprogramming factors SCL seems to be standalone master TF absolutely indispensible for the conversion of MAFs. Studies in zebrafish confirmed the crucial role of the SCL-LMO2 complex formation for haematopoiesis, as SCL mutant lacking the ability to interact with LMO2 could not rescue SCL morphant (Patterson et al., 2007). All reports to date position SCL at the core of haematopoietic GRNs and support finding that in reprogramming experiments its absence cannot be compensated by other factors. Knowing this it is even more striking that SCL was not identified as reprogramming TF in any other transdifferentiation study.

It is important to highlight that the selected set of TFs is probably not the only combination able to reprogram MAFs. Whereas SCL plays the most important role, the remaining four TFs seem to have less crucial effect. Notably, whereas the removal of cFOS only increased the efficiency of colony formation by 3 fold, this effect was not sustained during reprogramming with the minimal set of five TFs (Fig 3.4). Therefore a more systematic N-1 screen could reveal some other functional combinations of factors. Moreover selection of TFs shouldn't be based solely on their ability to generate colonies, as shown in Fig 3.4. Instead it would be more valuable to associate requirements for specific TFs with lineage composition of colonies and biological properties of generated cells. Moreover, the initial set of TFs was obtained using data curated from literature and therefore was limited by the state of knowledge regarding blood development four years ago. An alternative screening, based on differential expression patterns within the haematopoietic system, could lead to the identification of non-haematopoietic TFs with broader developmental functions like in the study of Riddell and colleagues (Riddell et al., 2014). Finally, the identification of successful reprogramming events based on the efficiency of *in vitro* colony formation allowed me to identify factors able to induce morphological and immunophenotypical changes on MAFs, but this does not directly translate to functionality in repopulation assays. Much more comprehensive efforts would be required to perform *in vivo* screen for factors that confer a repopulation potential to reprogrammed cells.

Overall, even though the set of reprogramming factors could be certainly refined and improved, the nature of identified TFs, as well as kinetics and efficiency of reprogramming, make SCL, RUNX1c, ERG, GATA2 and LMO2 very promising candidates for further investigation. By building multiprotein complexes these TFs can integrate inputs from different signals and therefore facilitate acquisition of comprehensive haematopoietic program in terminally differentiated fibroblasts. Importantly, the identified TFs are blood fate determinants implicated in early embryonic specification of haematopoiesis, suggesting that they can possess capability to induce stem/progenitor fate in somatic cells and therefore allow to generate the functional progenitor cells required for regenerative medicine.

Chapter 4

4. Reprogramming to blood by five TFs

4.1. Introduction

The five reprogramming TFs (SCL, LMO2, RUNX1c, GATA2, ERG) that I identified are very promising candidates for efficient fate switching to blood due to their well-established roles at different stages of haematopoietic development starting from embryonic specification. SCL and its essential protein partner LMO2 are both detected at the primitive stage of haematopoiesis (Patterson et al., 2007). SCL marks the establishment of haematopoietic or endothelial lineages from mesoderm at the level of haemangioblast (Gering et al., 2003). This stage of development is also tightly regulated by Bmp4 signaling which induces GATA2 and RUNX1 autoregulatory loops (Zafonte et al., 2007). The subsequent definitive blood generation is fully dependent on the upregulation of *Runx1*, as indicated by the absence of haematopoietic clusters in the DA of the *Runx1*$^{-/-}$ mice (Yokomizo et al., 2001). In fact, the close association of the selected reprogramming factors during embryonic development of HSPCs would explain why out of 19 regulators these alone were able to re-establish complex haematopoietic programme in terminally differentiated fibroblasts.

It is evident that during normal development these TFs remain active also downstream of stem cell generation, influencing fate decisions at the stage of multi- and bi-potential progenitors as well as functions of terminally differentiated effector cells. For example *Scl*, when overexpressed in HSCs, increase the output of myeloid cells. Conversely, dominant-negative mutant of SCL biased cells toward lymphoid differentiation (Kunisato et al., 2004). In the adult, conditional knock-out of *Scl* was found to perturb megakaryocytic/ erythroid haematopoietic development, as manifested by decreased levels of early progenitors for both lineages as well as blocked erythroid maturation

(Hall et al., 2003). In addition, the global approach of investigating whole genome binding patterns revealed the role of combinatorial TFs' interactions in governing these cell fate decisions. The intimate regulatory linkage between seven major haematopoietic TFs (SCL, LYL1, LMO2, GATA2, RUNX1, ERG, and FLI1) was first recognized in HSPCs (Wilson et al., 2010a). The findings from the HPC-7 cell line studies were soon after confirmed by ChIP-seq analysis of $CD34^{+}$ human progenitors (Beck et al., 2013). It was suggested that these combinatorial interactions between the heptad factors govern maintenance of undifferentiated blood fate and commitment to lineage specification. Indeed, binding of the HSPC regulatory complex to megakaryocytic genes, followed by GATA switching and FLI1 recruitment, was found to prime multipotent progenitors towards megakaryocytic fate (Pimkin et al., 2014). Conversely, disassembly of heptad factors and re-establishment of SCL/LMO2/GATA1/ LDB1 complex was suggested as a mechanism governing erythroid specification (Pimkin et al., 2014; Wadman et al., 1997). Interestingly, recent work has revealed that similar levels of expression of FL1, LMO2, SCL, ERG and RUNX1, identified as members of stem/progenitor cells regulatory complex, were detected in both progenitor cell line HPC-7 and terminally differentiated mast cells (Calero-Nieto et al., 2014). That suggests that not only differential expression, but also distinct behaviour of these master TFs, contributes to specific cellular statuses. Indeed, comparative analysis of ChIP-seq data revealed majorly non-overlapping binding patterns in both cellular environments. The motif enrichment analysis showed that Hox factors consensus binding sites were overrepresented in HPC-7 data. Conversely, AP1 and E-box motifs, normally bound by mast cells specific cFOS and Microphthalmia-Associated Transcription Factor (MITF), were largely occupied in differentiated but not precursor cells. Strikingly, even the *Kit* locus exhibited some substantial divergence in binding location of master blood TFs even though cKIT is present to the similar levels in both HPC-7 and mast cells. Changes in expression profiles of mast cell-specific genes, as well as growth retardation observed upon knock-down of shared stem/progenitor TFs (GATA2, ERG, LMO2), confirmed that bindings of these factors is not only 'opportunistic' but also 'functional' and can affect directly transcriptional programme and biology of terminally differentiated cells.

Altogether, the spectrum of roles performed by the identified reprogramming TFs at diverse levels of blood fate determination, multipotent progenitor specification and effector cell differentiation are likely to have some incidence on the spectrum of blood lineages that are generated by reprogramming from ontologically unrelated cells. In this chapter, I will investigate in more detail the haematopoietic outcome of reprogramming with five TFs and provide insights into the initial steps, kinetics and transcriptional landscape of this process in both adult and embryonic fibroblasts.

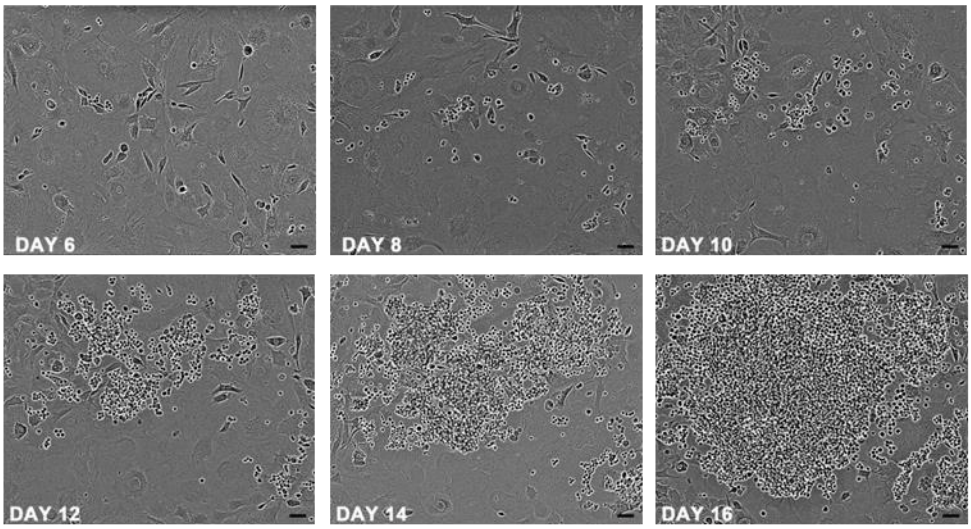
4.2. Results

4.2.1. Characterisation of blood cells generated by five TFs

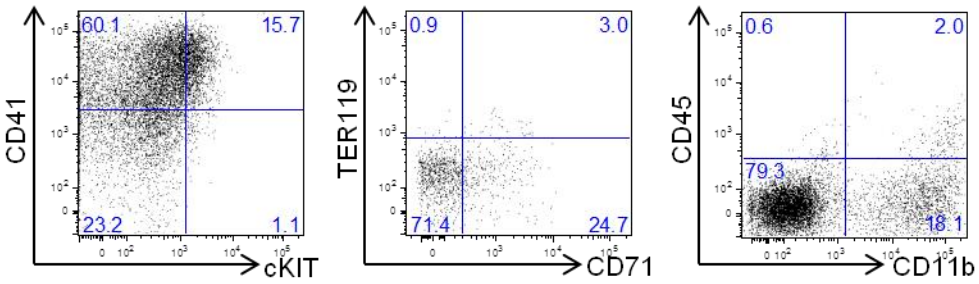
The five TFs identified in my reprogramming screen of MAFs are representative of all the major families of transcriptional regulators involved in haematopoietic development. I first re-examined their potential to transdifferentiate fibroblasts to blood using time-lapse microscopy. These experiments indicated that, as early as day 6 of reprogramming, the first phenotypical changes can be detected as rounding of fibroblasts that is shortly followed by the generation of grape-like blood colonies and proliferation (Fig 4.1A). Flow cytometry analysis of day 21 reprogrammed MAFs confirmed the acquisition of combination of haematopoietic cell surface markers associated with either progenitor (CD41/cKIT), myeloid (CD45/CD11b) or erythroid (CD71/TER119) populations (Fig 4.1B). These markers were not detected on the surface of the control untransduced MAFs as well as no signals could be detected with isotype controls (Fig S2Ai and ii). Cytospin of the cells present in day 21 cultures followed by May/Grünwald/Giemsa staining confirmed the presence of cells with a wide range of blood lineage morphologies, including three subtypes of granulocytes: basophils, eosinophils and neutrophils (Fig 4.1C and Fig S2B). Finally, qRT-PCR analysis for the expression of key haematopoietic genes associated with blood development and lineage determination revealed the upregulation of megakaryocytic (*Pf4, Cd42b*), myeloid (*Pu.1, Mpo*), erythroid (*β-major, Gata1, Gfi1b*) and pan-haematopoietic (*Ptprc, Itga2b*) genes in reprogrammed but not control

MAFs (Fig 4.1D). Importantly, the low level expression of primitive *βH1* globin in reprogrammed cells confirms obtainment of definitive erythrocytes. Collectively, these data suggest that a robust reprogramming of fibroblasts toward various blood lineages may be possible with a limited set of five TFs.

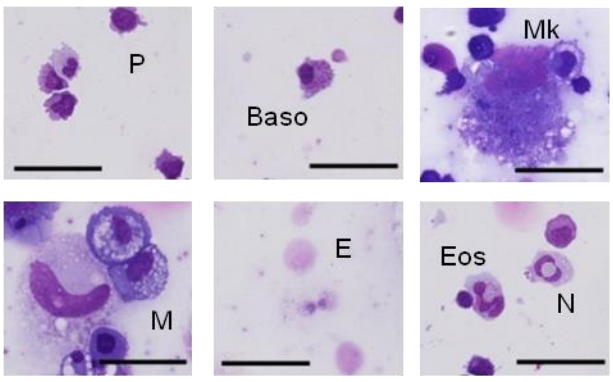
A)



B)



C)



D)

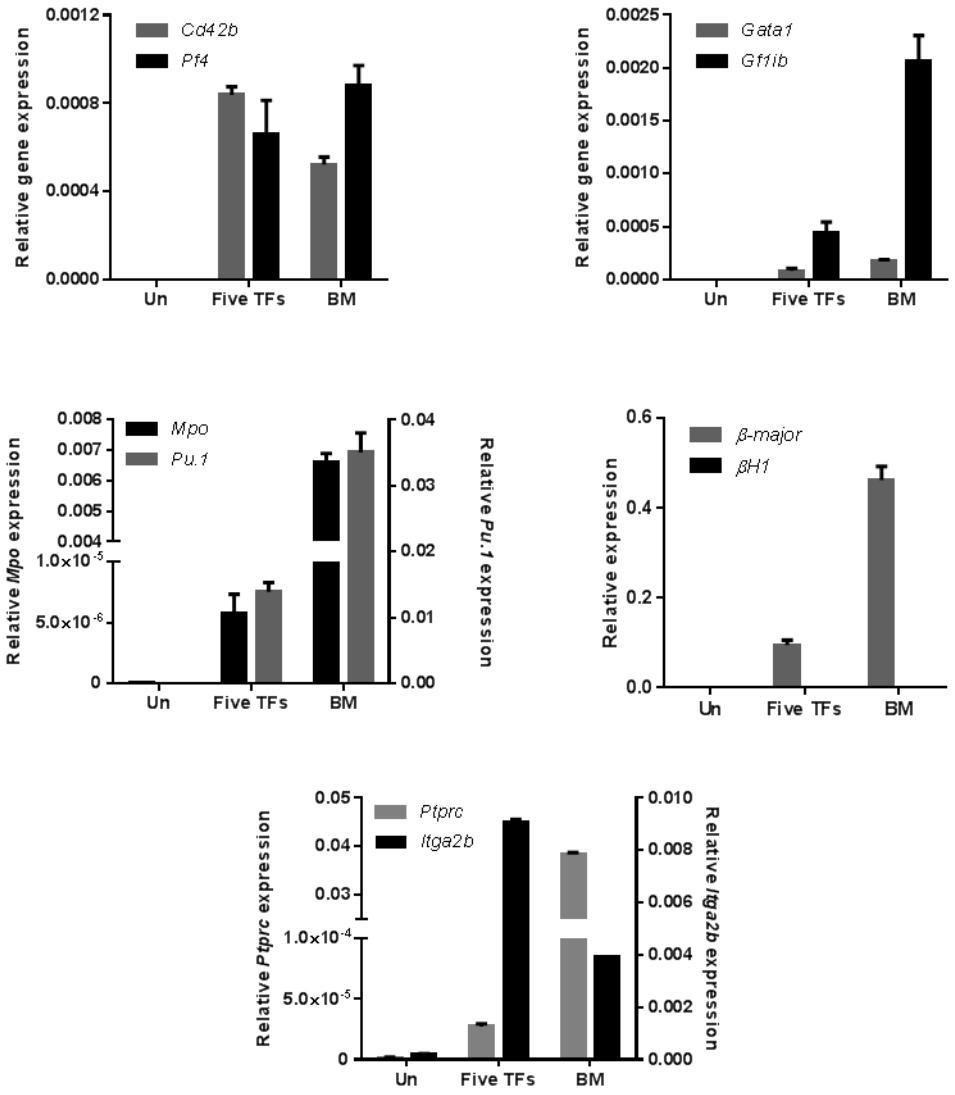


Figure 4.1. Fate conversion of MAFs by ectopic expression of five TFs.

A) Bright field images of colonies emerging from five-factor transduced MAFs between day 6 and 16 of culture. **B)** Flow cytometry analyses of day 21 reprogrammed MAFs for the expression of haematopoietic cell surface markers. CD41 and cKIT were used as early markers of haematopoietic progenitors. CD11b and CD45 expression define monocytes/macrophages and granulocytes. Collective staining with CD71 and TER119 reveal early ($CD71^{+}TER119^{-}$) and late ($CD71^{-}TER119^{+}$) erythroid commitment. **C)** Morphology of blood cells generated from day 21 MAFs transduced with five TFs. **D)** Relative to *β-actin* haematopoietic gene expression in control (Un) and day 21 reprogrammed MAFs (Five TFs) or control BM. Data shown are mean ± SD from a single experiment performed in triplicate (n=3). Scale bars represent 50 μm. E – erythrocyte, M – macrophage; P – progenitor, Mk – megakaryocyte, Eos – eosynophil, Baso – basophil; N- neutrophil

4.2.2. MEFs can be reprogrammed to blood with the same combination of factors as MAFs

The work detailed in figure 4.1 was performed on MAFs. To gain a better understanding of the versatility of the panel of TFs, I performed similar studies on MEFs. Embryonic fibroblasts should retain high level of plasticity due to the less stringent epigenetic landscape and this should translate into a higher efficiency and broader spectrum of reprogramming outcomes. When transduced with SCL, LMO2, ERG, GATA2 and RUNX1c, followed by culture in conditions favoring haematopoietic differentiation, MEFs underwent rapid change of morphology giving rise to round blood cells (Movie S1). Furthermore, MEFs were consistently reprogrammed with a higher efficiency (0.33%), up to twice in average the one observed for their adult counterparts (Fig 4.2A). Both types of fibroblasts exhibited similar immunophenotypes associated with reprogrammed cells at day 21, with comparable percentage of cells expressing CD41, cKIT and CD11b (Fig 4.2B). A significant difference was however observed in the frequency of CD45 and TER119 positive cells, reaching 10% in the reprogrammed population of MEFs versus only 3% in MAFs. This more robust development of the erythroid lineage in MEFs was also reflected in CFU-C assays in semi-solid culture conditions. At day 21, MEF cultures contained fewer functional progenitors than MAF cultures as indicated by their potential to generate haematopoietic colonies (Fig 4.2C left). In addition, the colonies derived from reprogrammed MEFs displayed a large range of morphologies associated with blood lineages, including erythroid, mixed erythroid-macrophagic and megakaryocytic colonies (Fig 4.2C right). In contrast, the CFU potential of reprogrammed MAFs was mostly limited to the generation of colonies with a myeloid morphology (macrophagic and granulocytic-macrophagic) with additional small number of compact colonies (indicated by white arrow) (Fig 4.2C). It was evident that the latter exhibit morphological resemblance to burst-forming unit erythroid (BFU-E), and were therefore better classified after performing qRT-PCR analysis of globin gene expression. Notably, the substantial levels of *β–major* gene but negligible presence of primitive *βH1* globin could be detected, which confirmed the definitive erythroid character of ‘immature’ CFUs (Fig S2C). Altogether, these data confirm that

MEFs can be reprogrammed to blood with higher efficiency than their adult counterparts and exhibit broader differentiation potential with significantly increased contribution towards erythroid lineage.

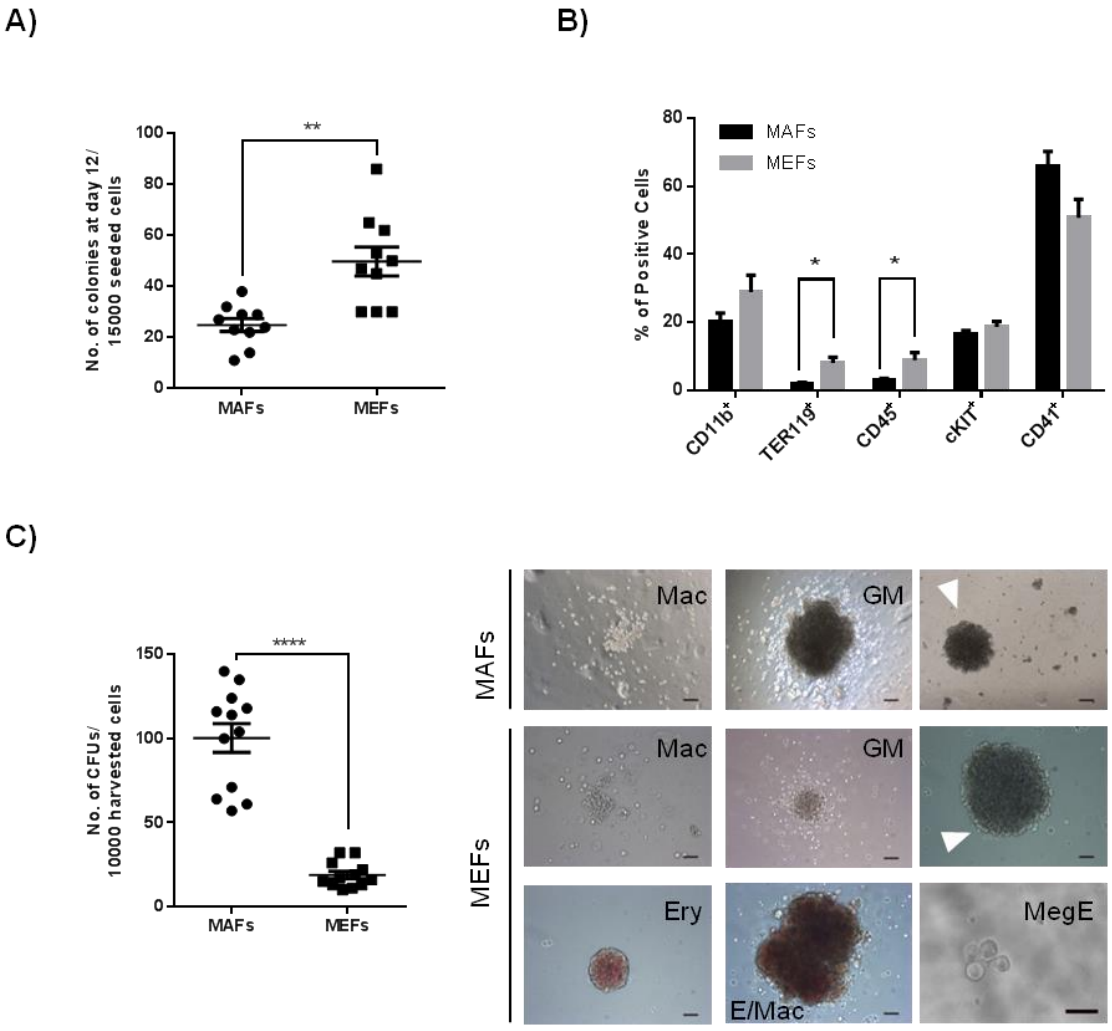


Figure 4.2. MEFs can be reprogrammed with the same factor cocktail as MAFs.

A) The number of colonies generated by five-factor transduced MAFs/MEFs on day 12 of reprogramming (mean ± SEM, n=5). **B)** Percentage of generated CD11b, TER119, CD41 or cKIT positive cells in day 21 cultures of reprogrammed MAFs/MEFs (mean ± SEM, n=6). **C)** The number of CFUs generated by 10000 day 21 five-factor transduced MAFs/ MEFs (left) and bright field images of representative colonies (right) (mean ± SEM, n=4). Scale bars represent 50 μm. Asterisk(s) represent significant differences (Student t-test, * p< 0.05, ** p< 0.005, **** p< 0.0001). Ery – erythroid CFU, E/Mac – mixed erythroid-macrophagic CFU, Mac – macrophagic CFU, GM – granulocytic-macrophagic CFU, MegE – megakaryocytic CFU.

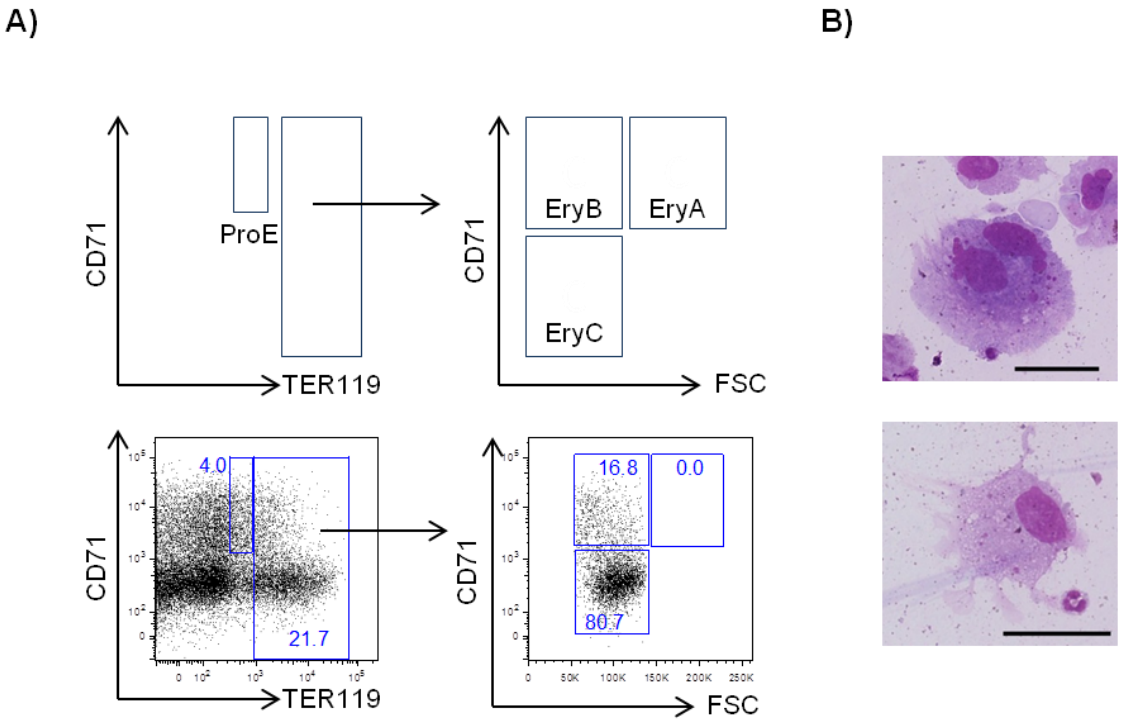
4.2.3. MEF-derived effector cells are mature and functional

Erythroid cell maturation occurs primarily within the FL between E11 and E15. Subsequent differentiation of erythroid cells can be followed through the combinatorial expression of TER119 and CD71 cell surface markers in this primary tissue (Pop et al., 2010). In this model the $Ter119^{hi}$ population contains late erythroblasts that can be further divided into three subclasses depending on their size (Fig 4.3A top). The EryA compartment contains early basophilic erythroblasts which further develop towards late basophilic erythroblasts (EryB) to finally differentiate into CD71 negative poly/ortochromatic erythroblasts (EryC). I therefore employed this flow cytometry approach to determine the maturation stage of MEF-derived erythrocytes. Combined TER119 and CD71 staining indicated that the overwhelming majority of erythroid cells are at their terminal differentiation stage and represent poly/orthochromatic erythroblasts (Fig 4.3A bottom). The presence of enucleated reticulocytes in the culture was confirmed by cytospin analysis (Fig 4.1C).

The process of platelet formation is complex and requires number of subsequent remodelling steps. To obtain the critical mass for production of numerous platelets, megakaryocytes first need to enlarge considerably by undergoing multiple rounds of endomitosis, before maturation and fractionation of internal membrane structures (reviewed by Ravid et al., 2002). The cytospin analysis of reprogrammed MEFs and MAFs revealed presence of polyploid megakaryocytes over 50 μm in diameter (Fig 4.1C and 4.3B top). Notably the subsequent steps of proplatelet formation, manifested by the presence of cells with protruding lammelipodia, could be observed in reprogrammed MEF cultures (Fig 4.3B bottom). Eventually these branched structures, when subjected to the force of blood flow *in vivo*, give rise to mature platelets.

Myeloid progeny is the most prevalent blood lineage generated from both MAFs and MEFs in the reprogramming process (Fig 4.2B, Fig S2B). Both $GR1^{+}CD11b^{+}$ granulocytes (Fig 4.4F) and $F4/80^{+}$ mature macrophages (Fig 4.3D) were detected in the cultures. The profound role of myeloid cells in innate immune response is associated with phagocytosis capabilities and secretion of inflammatory cytokines, both activated upon bacteria invasion. To assess the properties and functionality of fibroblast-derived macrophages, I first

performed a phagocytosis assay with the use of fluorescent latex beads. For this, CD11b/CD45 double positive cells were sorted from day 21 reprogrammed MEFs to isolate both macrophages and granulocytes (Fig 4.3C left). Upon incubation with beads only macrophages (large globular cells) but not neutrophils (characterised by donut-shaped nucleus) exhibited active intake of fluorescent particles (Fig 4.3C; neutrophils indicated by white arrow). I next sought to determine whether upon activation reprogrammed macrophages would secrete inflammatory cytokines IL6 and TNFα. An ELISA assay was performed on supernatants obtained from control BM-derived macrophages and test F4/80^{+} reprogrammed cells (Fig 4.3D left), both stimulated with LPS and IFNγ. Inflammatory cytokines were released by activated MEF-derived macrophages at similar (TNFα) or even higher (IL6) levels than measured for their healthy BM-derived controls (Fig 4.3D right). These assays demonstrate the physiological functionality of MEF-derived myeloid cells.



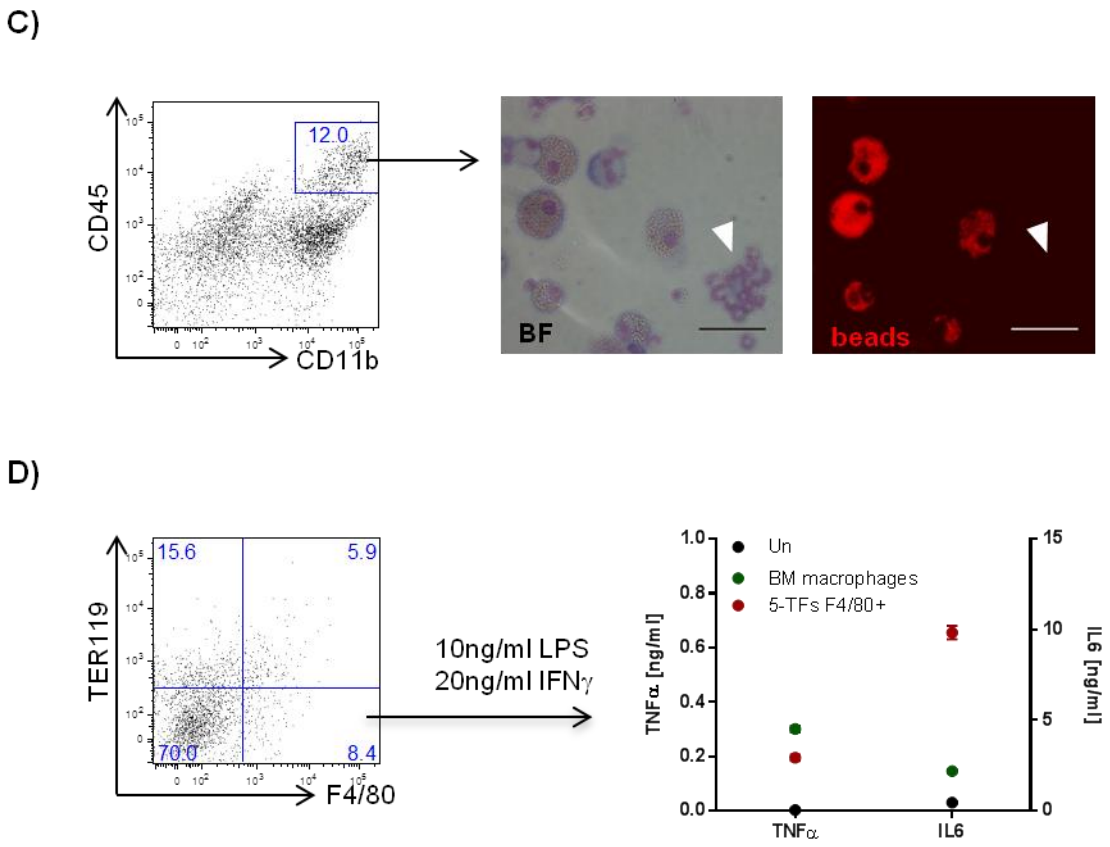


Figure 4.3. Reprogrammed MEFs give rise to mature and functional blood cells.

A) Schematic representation of cell surface markers distribution in subsets of erythroid cells (top) and flow cytometry plots showing expression of these markers on blood cells derived from MEFs at day 21 (bottom) **B)** Representative images of cytospin analysis of reprogrammed MEFs depicting polyploid megakaryocyte (top) and proplatelet (bottom). **C)** Sorting strategy (left) and bright field (BF) and fluorescence images (right) of day 21 sorted CD45/CD11b double positive reprogrammed MEFs upon phagocytosis of red-fluorescent latex-beads. **D)** Sorting strategy for obtaining F4/80$^+$ reprogrammed cells (left) and levels of TNFα and IL6 released by untreated MEFs (Un), BM-derived or reprogrammed F4/80$^+$ macrophages stimulated with 10 ng/ml LPS and 10 ng/ml IFNγ (mean ± SD) (right). Scale bars represent 50 μm. ProE - proerythroblast, EryA – early basophilic erythroblasts, EryB - late basophilic erythroblasts, EryC – poly/orthochromatic erythroblasts, FSC – forward scatter.

4.2.4. CD41 is the marker of transdifferentiation to blood

All five reprogramming TFs are involved in blood development during embryogenesis. This raises the possibility that the *in vitro* reprogramming could recapitulate to some extent the process of blood formation in the embryo. CD41 is an integrin widely expressed on HSPCs emerging during primitive and definitive waves of haematopoiesis (Boisset et al., 2013; Ferkowicz et al., 2003; Robin et al., 2011), which makes it a good candidate as a marker of

transdifferentiation. Flow cytometry analysis revealed the emergence of CD41 marker as early as day 4 of reprogramming, which then continued to be predominantly present throughout the culture period (Fig 4.4A). This data was supported by immunostaining for CD41 in day 12 experimental MAF cultures. CD41 was detected on fibroblasts undergoing fate switching as well as emerging blood colonies (Fig 4.4Bi and ii respectively). Control ectopic expression of single TF in fibroblasts did not drive CD41 upregulation by day 4 after transduction (Fig 4.4C). Similar results were also obtained for MEFs (work by Kiran Batta). In contrast GR1, the granulocyte associated cell surface marker, could be detected on fibroblasts infected with viruses expressing single TF even after initial removal of $GR1^+$ cells from the culture at day 0 (Fig 4.4D). Particularly ERG was able to most efficiently induce the expression of *Ly6g* (encoding GR1). The time course analysis of cell surface markers acquisition showed that whereas GR1 was present on ~35% of cells as early as day 6, the frequency of cells also positive for the other myeloid marker CD11b gradually increased from day 8 of culture (Fig 4.4E). Interestingly, this steady CD11b upregulation could be observed predominantly in the form of appearance of $CD11b^+/GR1^+$ cells. As a result, by day 21 of reprogramming, the majority of $GR1^+$ cells were also positive for CD11b (Fig 4.4F). One possible interpretation of these data would be that acquisition of GR1, probably induced by ERG, precedes *Itgam* (encoding CD11b) upregulation and subsequent determination of myeloid fate. A reverse situation could be observed for SCA1 marker that was ubiquitously expressed on starting MEFs (Fig S2D). Upon acquisition of haematopoietic fate, SCA1 was downregulated in the majority of blood cells both LIN negative and positive. Collectively, these data indicate that the acquisition of haematopoietic markers upon reprogramming is quite complex. This highlights the importance of choosing the right combination of cell surface markers to follow and quantify cell fate changes. Notably, I found that CD41 is the haematopoietic cell surface marker expressed specifically on cells undergoing transition towards blood and therefore represents a good indicator of successful reprogramming.

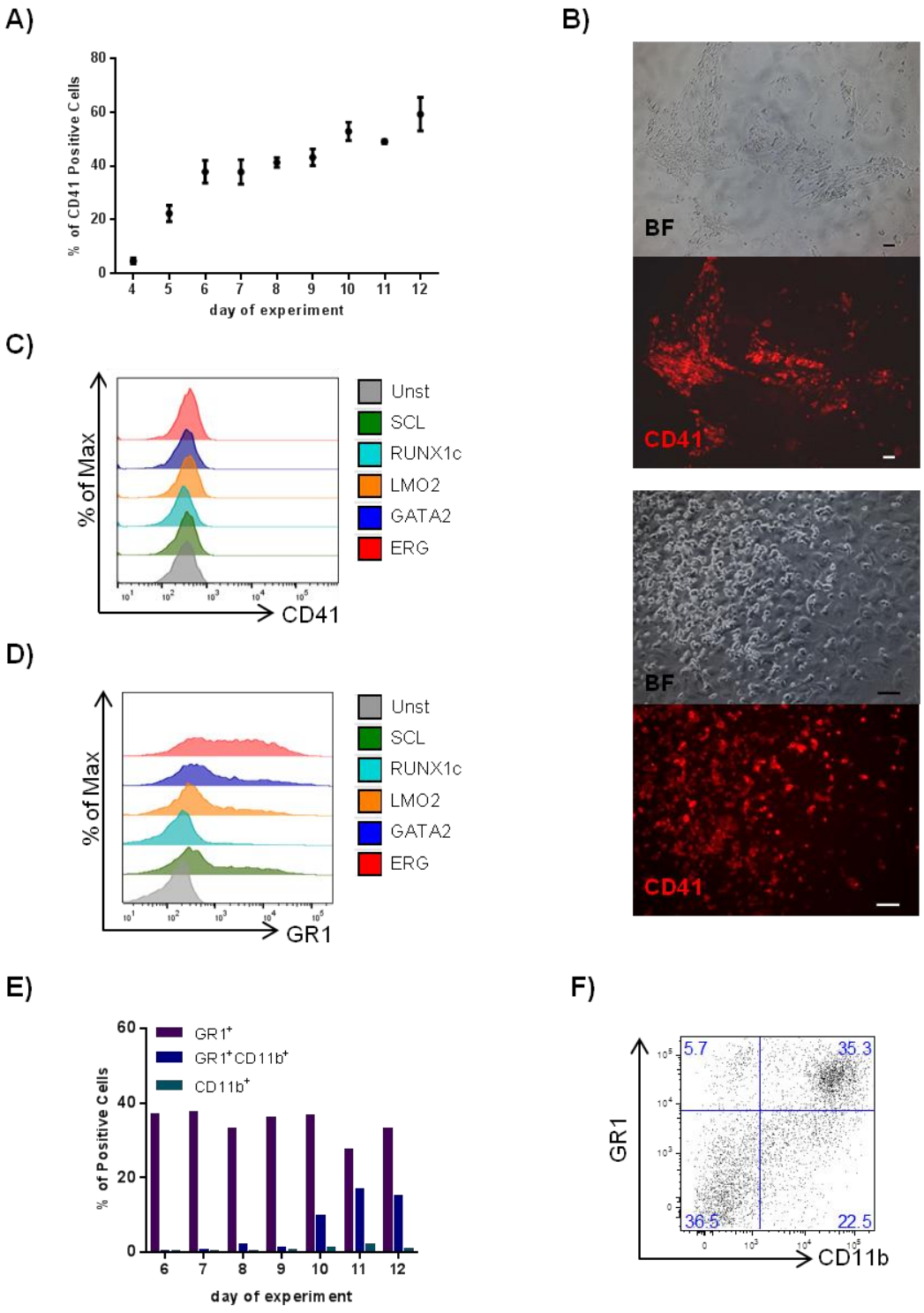


Figure 4.4. Pan-haematopoietic marker CD41 marks first reprogramming events.

A) Flow cytometry profile of CD41 expression between day 4 and 12 of reprogramming (mean ± SEM, n=2). **B)** Bright field (BF) and fluorescent images of CD41-PE stained cultures of day 12 five-factor transduced MAFs **C)** Representative histograms showing lack of CD41 expression in day 4 single-factor transduced fibroblasts. **D)** Histograms reflecting surface expression of GR1 expression on transitioning fibroblasts 4 days after transduction with single reprogramming TF or unstained control (Unst). **E)** Time course profile of GR1 and CD11b acquisition on reprogrammed fibroblasts between day 6 and 12 of culture. **F)** Flow cytometry plot presenting distribution of myeloid markers GR1/CD11b on the surface of day 21 reprogrammed MEFs Scale bars represent 50 μm.

4.2.5. Transdifferentiation to blood is a rapid transition involving concomitant downregulation of fibroblast-specific and upregulation of haematopoietic genes

Following the finding that CD41 is specifically expressed on transdifferentiating cells, I decided to use this marker to investigate whether transition to blood is an immediate process. To this end, I sorted day 6 and 12 CD41 positive and negative cells and assayed fibroblastic and haematopoietic gene expression levels. A 10 fold downregulation of *Fbn1* was observed as early as day 6 in $CD41^+$ cells (Fig 4.5 Ai). The expression of the two other fibroblastic genes *Acta2* and *Fspn1* was halved by day 12. Concomitantly, *Gfi1b* and *β-major* were detected by day 6 and 12 respectively in CD41 positive (Fig 4.5Ai) but not in the negative (Fig 4.5Aii) fraction of transduced cells. These data further indicate that CD41 marks cells undergoing haematopoietic transition. The observed trend of timely dependent upregulation of haematopoietic genes was also reflected on a more global level, without purification of $CD41^+$ cells. Wells with transduced MAFs were harvested at the indicated time points between day 4 and 21 of transition (Fig 4.5B and C). Once again a stepwise upregulation of expression of *Gfi1b* and *β-major* but also myeloid genes *Pu.1* and *Mpo* was detected, suggesting a steady transition towards the haematopoietic fate and blood differentiation (Fig 4.5Bi and ii). No significant upregulation of pluripotency genes, *Oct4, Sox2* or *Nanog,* was detected throughout reprogramming (Fig 4.5C). The same pattern of global haematopoietic and pluripotency gene expression was also obtained for MEF cultures (work by Kiran Batta). Altogether, these data illustrate that the major changes in transcriptional landscape of transitioning cells occur rapidly and as early as day 6 of the process. Moreover this immediate fate switch is direct and does not require reversal to an intermediate pluripotency state.

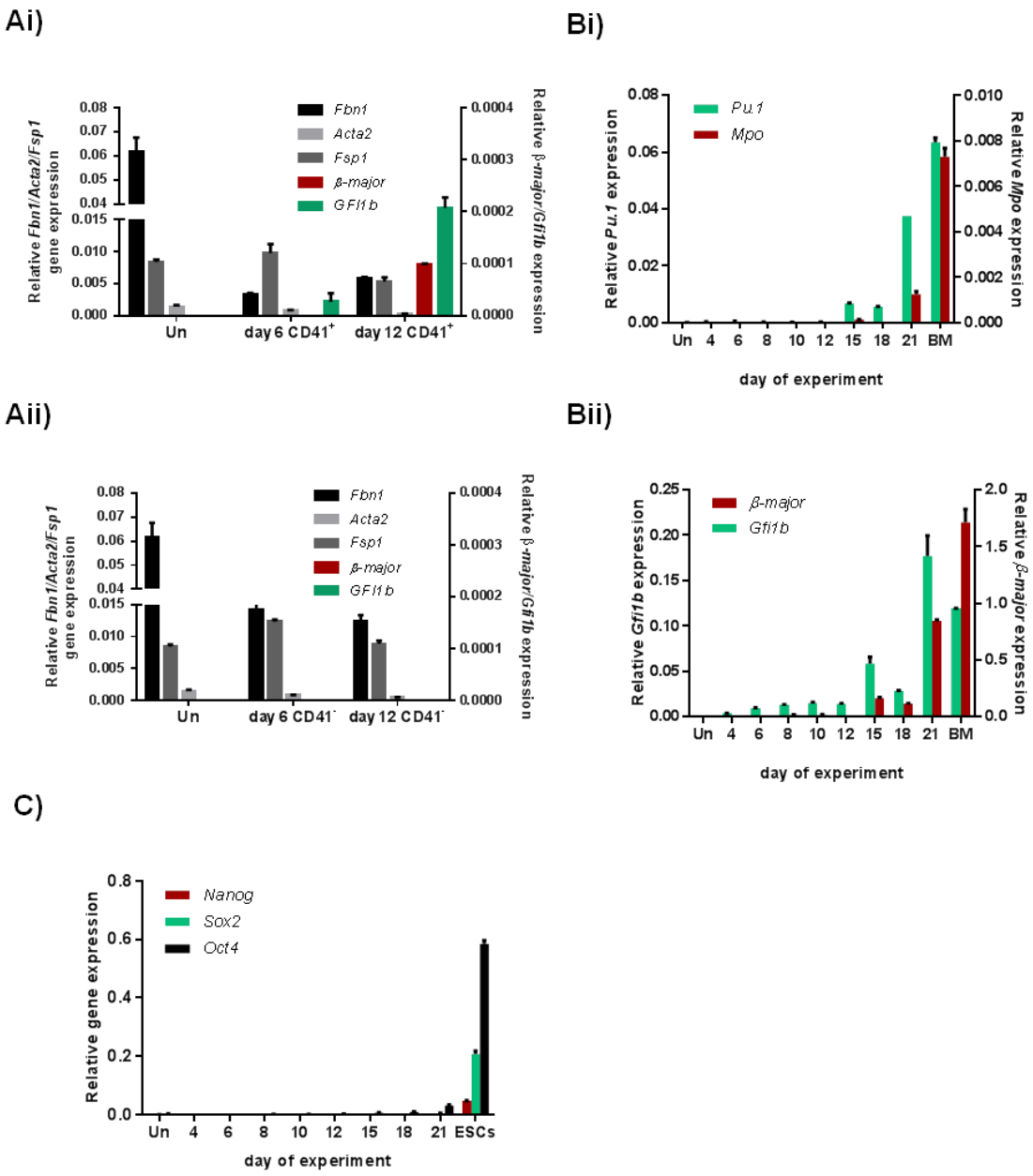


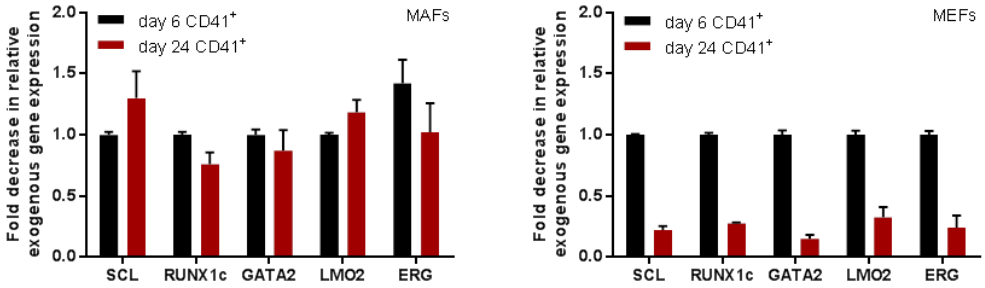
Figure 4.5. Reprogramming by five TFs is a direct conversion to blood.

A) Gene expression analysis of untransduced,(Un) day 6 and day 12 CD41positive **(i)** and negative **(ii)** sorted reprogrammed MAFs. **B, C)** Relative expression of haematopoietic myeloid **(Bi)** and erythroid **(Bii)** as well as pluripotency **(C)** genes in MAFs undergoing reprogramming. Data presented on this figure are mean ± SD of a representative experiment performed in triplicate and were analysed relative to *β-actin* (n=2).

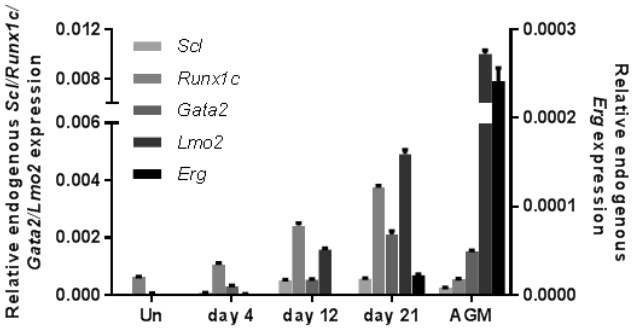
4.2.6. Reprogrammed MEFs, but not MAFs, acquire balance between exogenous and endogenous gene expression

An important indication of a successful reprogramming, leading to a complete fate conversion, is the independence of transitioning cells from exogenously introduced TFs. This indicates that the generated cells maintain their newly acquired identity by solely relying on the activated intrinsic regulatory networks. Although the lentiviral vectors used for my work did not include an intrinsic system of control of the transgene expression, passive silencing in transduced cells has been reported in our laboratory. To establish whether reprogrammed MEFs and MAFs have become independent of exogenous TFs, I compared expression levels of virally introduced *Scl, Lmo2, Runx1c, Gata2* and *Erg* in the $CD41^{+}$ cells quickly following transduction (day 6) and at the final steps of the reprogramming protocol (day 24). Interestingly, transitioning MAFs maintained similar expression levels of the exogenous factors at both time points, whereas for MEFs with the progression of reprogramming the level of all five TFs decreased by approximately 10 fold (Fig 4.6A). Furthermore, when MEF-derived progenitors were subjected to serial re-plating in semi-solid conditions colony formation was not sustained beyond the second round of culture (Fig S2E). These data confirm that residual expression of the transgenes does not confer pro-survival and proliferative advantage on reprogrammed cells. When I assessed the expression of endogenous counterparts of the reprogramming TFs, I found that the expression of the intrinsic regulators was gradually activated in reprogrammed MEFs throughout the cell culture (Fig 4.5Bi). A similar trend could also be observed in MAFs, despite the sustained expression of transgenes (Fig 4.5Bii). In both MEFs and MAFs the acquisition of endogenous genes expression was stepwise, suggesting a gradual activation of internal GRNs that maintain the newly acquired haematopoietic identity. Altogether, these data suggest that even though reprogramming of MAFs and MEFs generally result in the same blood lineage outcome, it is accompanied by different transgene transcriptional landscape.

A)



Bi)



Bii)

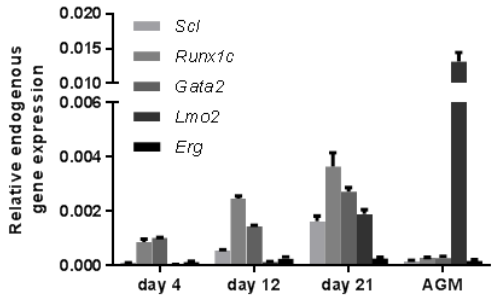


Figure 4.6. Transcriptional landscape of five-factor reprogrammed fibroblasts.

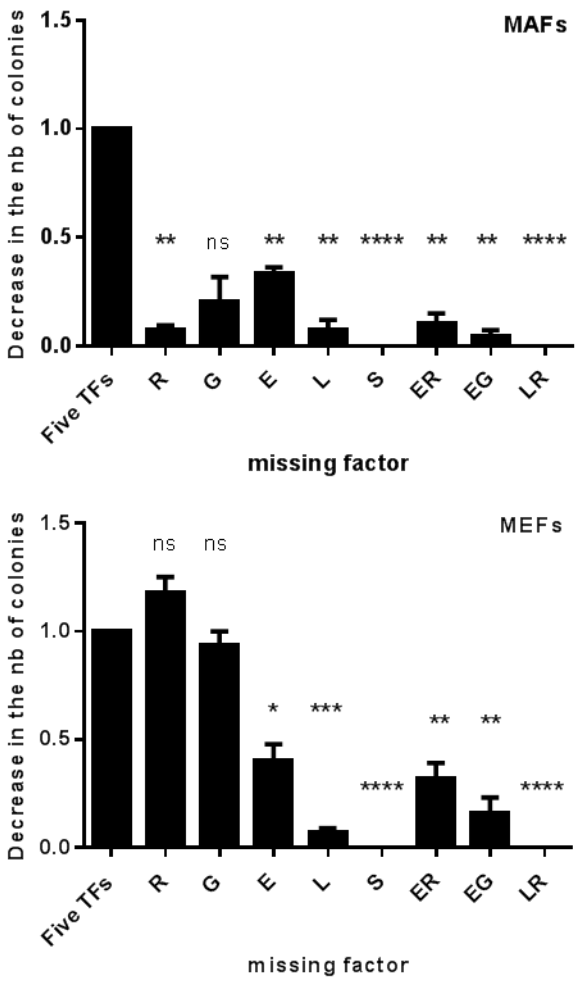
A) Difference in expression levels of virally delivered TFs in day 6 and day 24 CD41 sorted reprogrammed MAFs (left) and MEFs (right) (mean ± SEM, n=2). Data were analysed relative to *β-actin.* **B)** Relative expression levels of endogenous genes with respect to *β-actin* in day 4, 12 and 21 transduced MEFs**(i)**/MAFs**(ii)** and control E10.5 AGM cells. Mean ± SD of a representative experiment is shown (n=2)

4.2.7. MAFs and MEFs have different minimal reprogramming TFs requirements

The five selected haematopoietic fate converters are known to mutually regulate each other's expression, as discussed in chapter 3. In an aim to identify a master regulator, I sought to determine whether the number of reprogramming TFs can be decreased to a single modulator capable of inducing expression of remaining key factors. To address this question both MAFs and MEFs were first transduced with combinations of four or three TFs, and the number of emerging colonies was assayed between day 12 and 15 of reprogramming, depending on the kinetics of the specific experiment. As previously observed with TFs removal from the pool of 19 regulators, the lack of SCL and LMO2 had also the most detrimental effect on reprogramming efficiency (Fig 3.4A and 4.7A). As for the remaining factors, whereas RUNX1c and GATA2 proved to be crucial for transdifferentiation of MAFs, they were mainly dispensable for transition of MEFs. The lack of ERG decreased the efficiency of reprogramming by half in both tested sources of fibroblasts. For fate conversion of MAFs ERG was, however, the least crucial TF. Such a discrepancy in reprogramming requirements between two almost identical cell types could have originated from their inherent transcriptional landscape. For instance, generation of iPSCs from neural stem cells requires only OCT4, and not the established OSKM combination. This is due to the inherent expression of *Sox2, Klf4* and *cMyc* in neuronal tissue (Kim et al., 2009b). The qRT-PCR analysis of endogenous *Scl, Runx1c, Lmo2, Erg* and *Gata2* expression levels in the starting populations of MAFs and MEFs confirmed significant differences in their starting profile (Fig 4.7B). However, obtained results were not in line with TF removal experiment. *Runx1c* and *Gata2* were highly expressed in untransduced MAFs but not MEFs. However, the removal of the two factors from the viral mix affected only transition of MAFs. This clearly shows that intrinsic expression is not compensating for the lack of reprogramming factors. Another reason for the observed discrepancy in TFs' requirements may be higher susceptibility of MEFs to fate change. This source of cells is widely used in reprogramming experiments due to their high plasticity resulting from embryonic origins. Indeed, when MEFs were transduced with different

combinations of factors, even with the restricted haematopoietic triad SCL, FLI1 and GATA2, rare reprogramming events could still be detected (Fig 4.7C). Conversely, I have not found an alternative reprogramming mix capable of driving transdifferentiation of MAFs. This susceptibility of MEFs to fate switching may be explained by a more open chromatin status which would allow binding of minimal set of master haematopoietic regulators to a broader range of regulatory elements and, therefore, facilitate more robust instillation of alternative cellular fate. To test this hypothesis, I transduced both MAFs and MEFs with SCL and LMO2 - two critical TFs from the original mix. Not surprisingly treatment of embryonic but not adult fibroblasts resulted in the appearance of blood-like colonies under these reduced conditions with lower albeit still substantial efficiency (Fig 4.7D).

A)



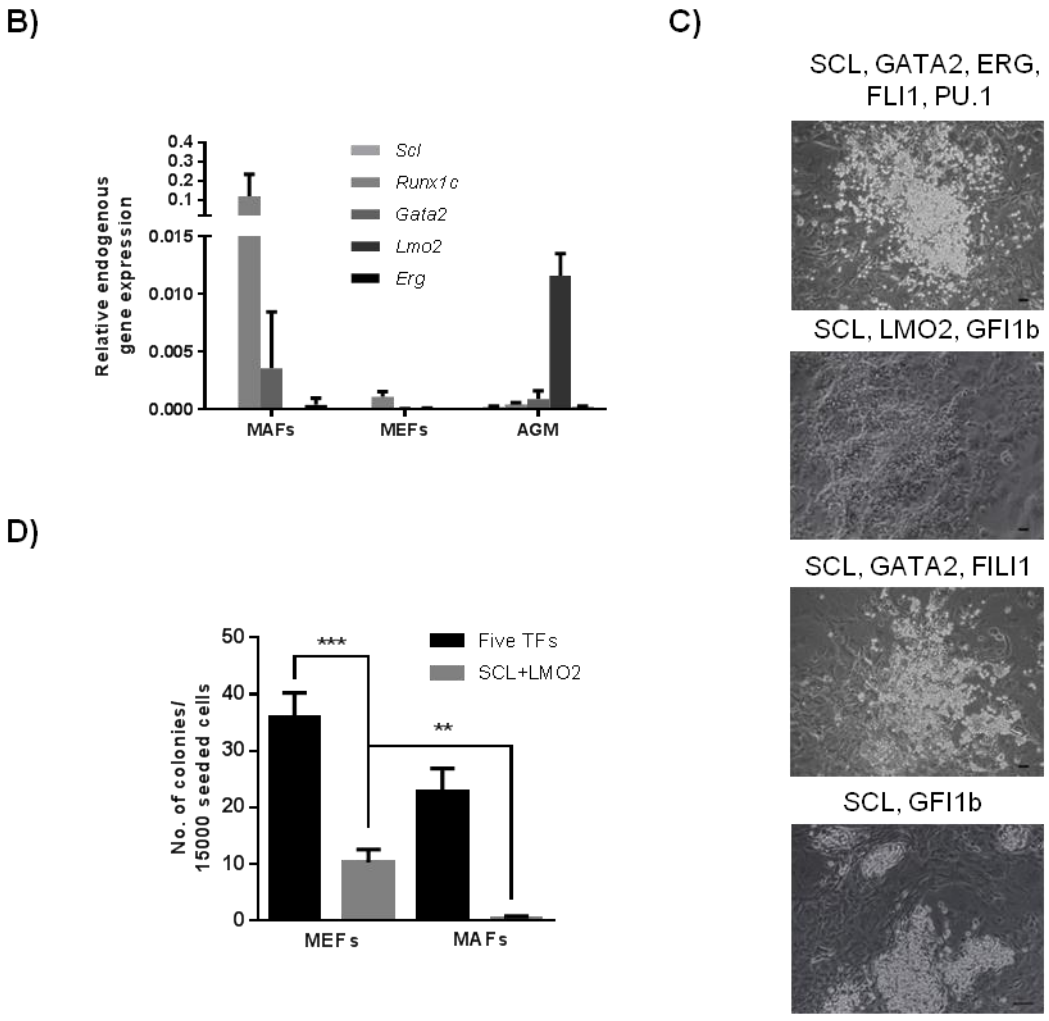


Figure 4.7. MAFs and MEFs have different requirements for reprogramming TFs.

A) The number of colonies generated from transduced MAFs (top) and MEFs (bottom) between day 12 and 15 of experiment upon removal of the indicated TFs from the five-factor viral mix (mean ± SEM, n=3 (MAFs), n=2 (MEFs)).Statistical significance was assessed with a use of one-way ANOVA followed by Dunnett's post hoc test. **B)** Relative gene expression of endogenous *Scl, Lmo2, Erg, Gata2, Runx1c* with respect to *β-actin* in untransduced MAFs/MEFs and control AGM (mean± SEM, n=3 (MAFs), n=2 (MEFs, AGM). **C)** Bright field images of colonies generated by different possible combinations of factors reprogramming MEFs to blood-like cells. **D)** The number of colonies generated from adult and embryonic fibroblasts between day 12 and 15 after transduction with SCL and LMO2 only or five TFs (mean ± SEM, n=3). Statistical significance was assessed with a use of paired Student t-test. Scale bars represent 50 μm. Asterisk(s) represent significant differences ($p< 0.05$, $^{**} p< 0.005$, $^{***} p< 0.0005$, $^{****} p< 0.0001$).G – GATA2, E- ERG, L- LMO2, R- RUNX1c, S- SCL.

4.2.8. SCL and LMO2 reprogram MEFs to blood with the same lineage outcome as five TFs

The fact that visible change of MEF morphology can be achieved with only two factors opens new possibilities for establishing more robust and simplified procedure. A more reproducible protocol represents an excellent platform for testing different alternatives to improve or understand reprogramming. To develop such alternative, I employed MEFs isolated from mice carrying a inducible *Scl* cassette (*iScl*). In these cells *Scl* expression is followed by expression of a *Gfp* reporter (internal ribosome entry site (IRES)-GFP) and is activated upon addition of dox (Fig 4.8A). The LMO2 expression cassette was cloned into a lentiviral vector together with dtTomato fluorescent protein (IRES-dtTomato) (obtained from Axel Schambach, Hannover Medical School). The presence of the dtTomato reporter facilitated the titration of the virus and consequently more consistent levels of transduction. Combining induction of *Scl* expression and transduction with the LMO2 virus resulted in the generation of blood-like colonies with varying efficiency depending on the batch of isolated *iScl* MEFs (Fig 4.8B). Notably, the system was functional provided that induction of *Scl* and *Lmo2* were concomitant (Fig 4.8C). Interestingly, untransduced MEFs exhibited some degree of uncontrolled low level of *Scl* expression (Fig 4.8D). These detected at day 0 GFP^+ cells generated low numbers of colonies even upon transduction with LMO2 only without accompanying dox induction (Fig 4.8E). Conversely, the subpopulation of MEFs that was GFP^-, and therefore displayed tight control of transgene expression, displayed change of morphology only in the presence of both LMO2 and dox stimulation (Fig 4.8E). To avoid any contamination, the ‘leaky’ cells were therefore removed from day 0 MEF cultures by sorting.

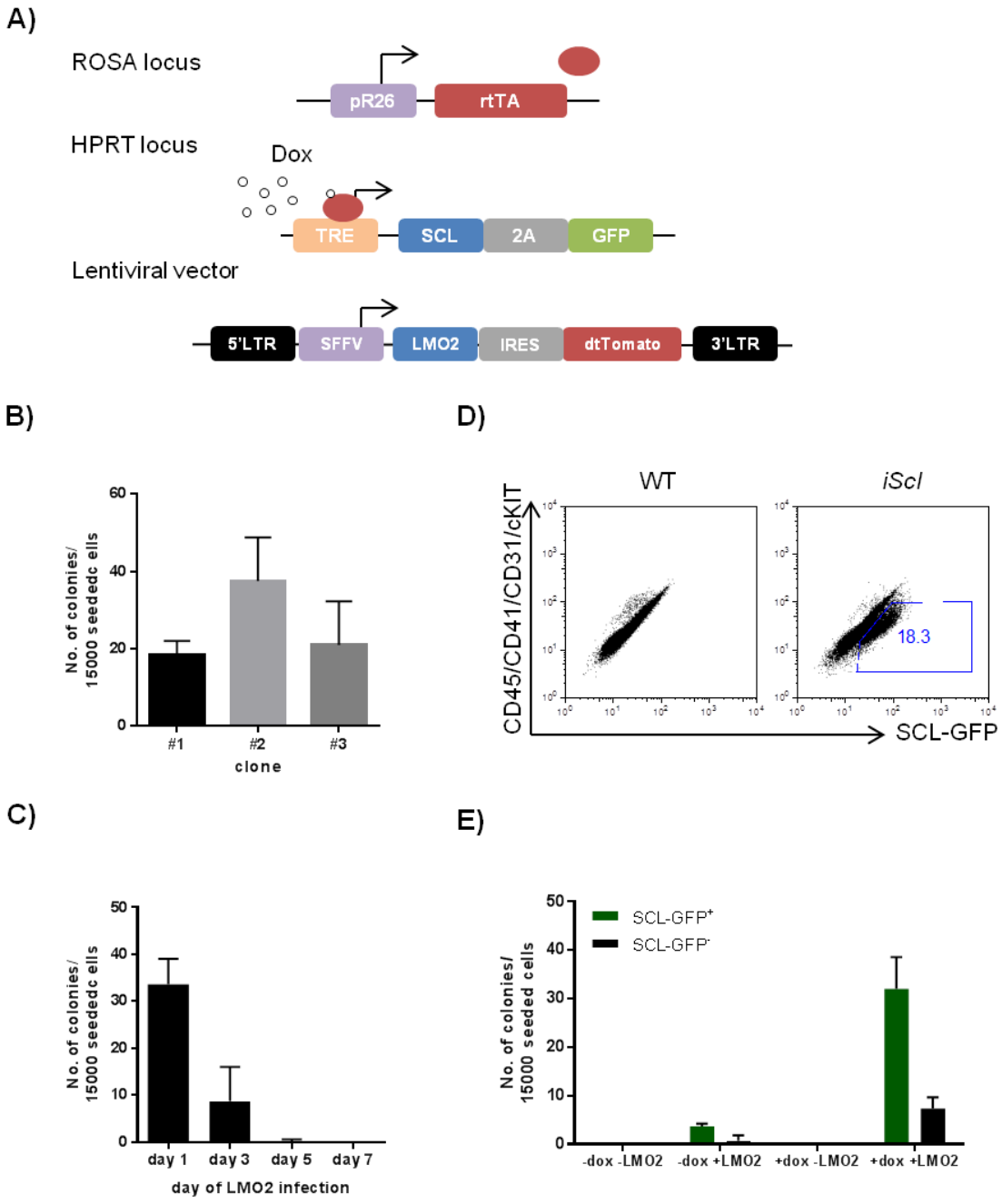
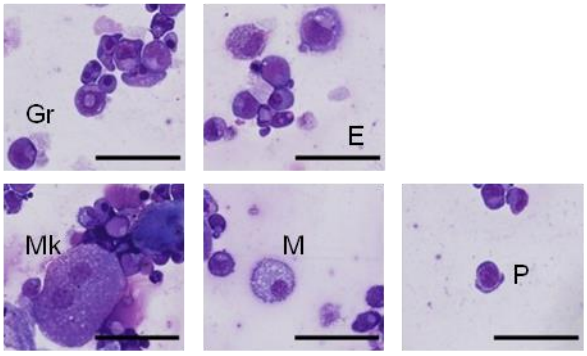


Figure 4.8. Assessment of i*Scl* system.

A) Schematic representation of *iScl* mouse model and reporter-coupled LMO2 viral backbone. **B)** Different batches of MEFs derived from *iScl* mouse model were infected with LMO2 virus and treated with dox at day 1 of experiment. Generated blood-like colonies were counted between day 12 and 19 of experiment depending on the observed efficiency of reprogramming (mean ± SD) **C)** Fraction of cells was treated as described in B (day 1). Remaining cells were infected with LMO2 lentivirus at day 3, 5 or 7 with concomitant normal pattern of *Scl* induction. Generated colonies were counted at day 19 of experiment (mean ± SD). **D)** Flow cytometry profile of day 0 WT and *iScl* MEFs presenting leaky expression of *Gfp*. **E)** At day 0 *iScl* MEF-derived CD45/CD41/GR1/cKIT/CD31 negative cells were sorted into GFP+ and GFP- fractions and subjected to specified dox/LMO2 treatments. The number of generated colonies was assessed at day 12 of reprogramming (mean ± SD). pR26 – Rosa26, rtTA – reverse tetracycline transactivator, HPRT –, Hypoxanthine-guanine phosphoribosyltransferase, TRE – tetracycline responsive element, LTRs – Long terminal repeats, SFFV – Spleen focus-forming virus promoter

The aim of the following experiments was to determine whether blood cells are indeed generated by *Scl*/*Lmo2* overexpression in MEFs. Multiple blood lineages, including myeloid, erythroid and megakaryocytic progeny, could be morphologically identified in day 21 cultures (Fig 4.9A). Flow cytometry analysis confirmed the generation of mature macrophages displaying F4/80 cell surface marker, 60% of which were also transgene independent (Fig 4.9B). To further confirm the definitive character of reprogrammed erythroblasts, I investigated the expression of primitive (*εy, βH1, ξ*) and definitive (*β1, β2, α1/2*) globin genes in TER119$^+$ sorted cells. As anticipated, all types of globins were expressed in the AGM-derived cells and only *β–* and *α-major* chains could be detected in the BM sample (Fig 4.9C). Interestingly, day 21 TER119$^+$ cells reprogrammed via five factors were expressing both primitive and definitive globins, whereas in the SCL/LMO2 reprogrammed counterparts only definitive globins were present. All the same two-factor reprogrammed cells show lower levels of β1 and α1/2 transcripts in comparison to five-TF evoked transition, which may be attributed to smaller number of reprogrammed erythroid cells or different distribution of erythroblasts and their respective ratios (compare Fig 4.3A). Finally, I investigated whether the two-factor reprogrammed blood cells contain a wide spectrum of progenitors. At day 21 of experiment cKIT$^+$ cells were readily detected in the culture (Fig 4.9D left). Moreover, when a pool of reprogrammed cell was re-plated into semi-solid conditions they gave rise to a wide range of types of haematopoietic colonies (Fig 4.9D right). This series of experiments indicates that temporal activation of SCL and LMO2 enables the induction of multiple blood fates on MEFs.

A)



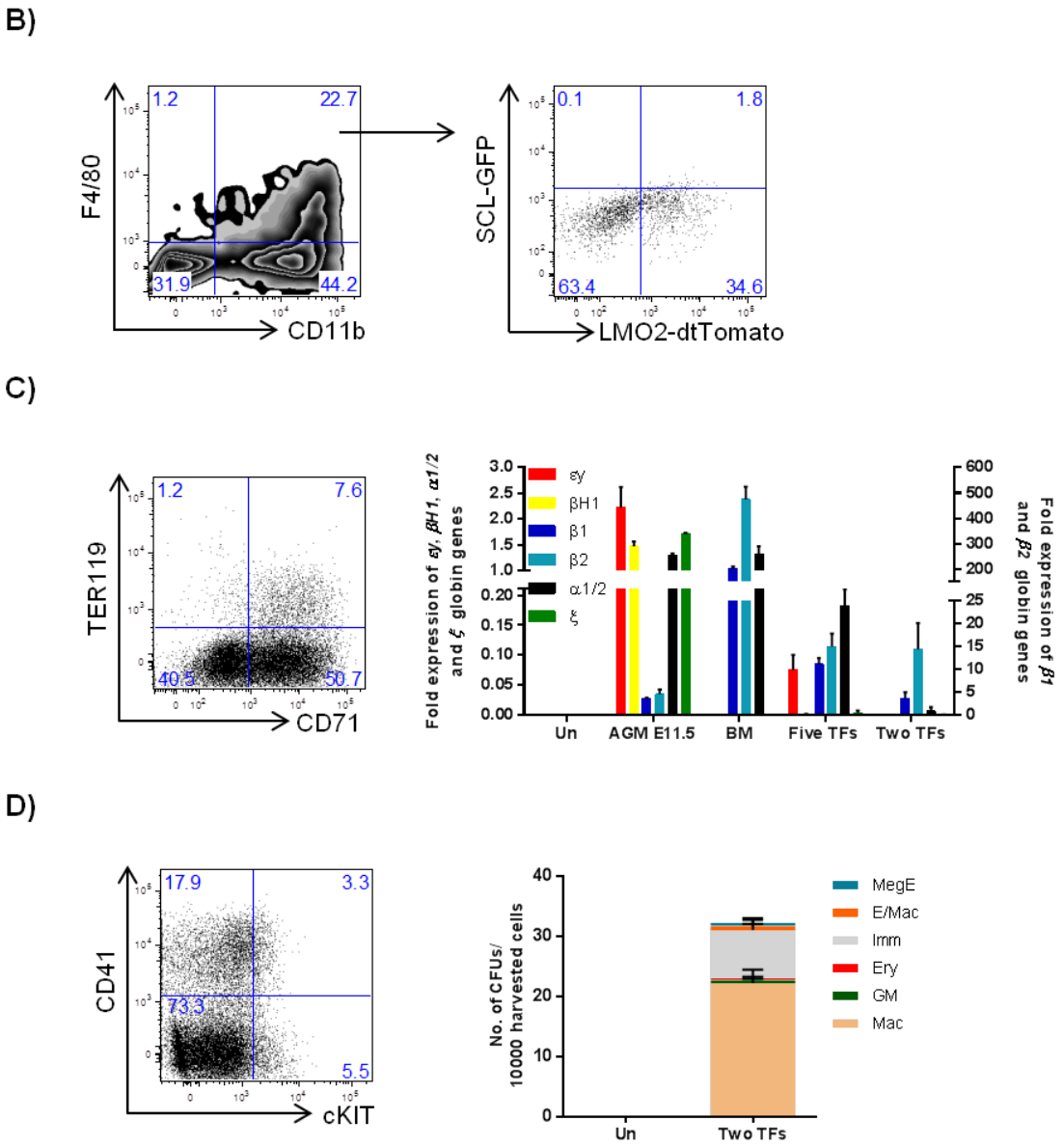


Figure 4.9. SCL and LMO2 are sufficient to reprogram MEFs.

A) Morphology of day 21 SCL/LMO2 reprogrammed cells. **B)** Surface expression of mature macrophage-specific cell surface markers on day 21 two-factor reprogrammed MEFs (left) and SCL/LMO2 transgene expression in F4/80/CD11b double positive fraction (right) **C)** Distribution of erythroid markers CD71 and TER119 on the surface of day 21 two-factor reprogrammed MEFs (left). Relative to β-actin expression of embryonic and adult globins in TER119^{+} sorted *iScl*/LMO2 or five TFs reprogrammed MEFs, control fibroblasts (Un), AGM E11.5 and BM cells presented as fold of levels detected in E10.5 AGM (mean ± SD) (right). **D)** Flow cytometry profile of early progenitor cell surface markers (left) and number and types of CFUs generated by day 21 two-factor reprogrammed MEFs (right, mean ± SEM, n=2). Scale bars represent 50 μm., E – erythrocyte, M – macrophage, Gr- granulocyte, P – progenitor, Mk – megakaryocyte, Ery – erythroid CFU, E/Mac – mixed erythroid-macrophagic CFU, Mac – macrophagic CFU, GM – granulocytic-macrophagic CFU, MegE – megakaryocytic CFU.

4.3. Discussion

The five selected reprogramming TFs function both as blood fate determinants and regulators of subsequent haematopoietic specification. It has been previously shown that combinatorial interactions of the same master TFs can control multiple levels of haematopoietic development, spanning from establishment of HSCPs, through priming towards megakaryocytic fate or governing mast cell differentiation (Calero-Nieto et al., 2014; Pimkin et al., 2014; Wilson et al., 2010a). Achieving these various functions is cell-context dependent and relies on a fine balance between members of the regulatory heptad and other determinants of lineage specification present in the cell. For example, cFOS and MITF, known mast cell regulators, facilitate differentiation of mast cells by allowing binding of the heptad factors to mast cell-specific sites not occupied in HSCPs (Calero-Nieto et al., 2014). Furthermore, the priming of multilineage progenitors towards megakaryocytic fate requires the binding of heptad factors to specific loci followed by genome wide GATA switching. Conversely, establishment of erythroid fate is mostly associated with dissociation of ETS/RUNX1/GATA2 factors and reassembly of a new regulatory complex with GATA1 component (Pimkin et al., 2014; Wadman et al., 1997). The challenge in cell fate switching towards blood lies in our ability to accurately recapitulate these intricate spatial and temporal combinatorial mechanisms. In this chapter, I established that a set of five master haematopoietic regulators (SCL, LMO2, RUNX1c, GATA2 and ERG) can induce different blood fates in fibroblasts and drive the generation of both progenitors and terminally differentiated haematopoietic cells.

The data presented in this chapter show that both adult and embryonic fibroblasts can be reprogrammed to blood with the same set of haematopoietic TFs. Upon transduction MAFs and MEFs gave rise to vast spectrum of mature blood lineages as well as progenitors with robust *in vitro* clonogenic potential (Fig 4.1 and 4.2). Importantly, MEF-derived erythrocytes and megakaryocytes clearly underwent advanced maturation process (Fig 4.3A and B). Moreover macrophages displayed functionality in relevant biological assays (Fig 4.3C and D). Nevertheless, I also identified several differences between the cells generated upon MAF and MEF transdifferentiation. This suggests that not only

the type, but also the maturation state of cells, has a profound effect on reprogramming. MAFs were consistently more difficult to reprogram, a phenomenon manifested by lower transition efficiency and more stringent requirements for fate-determining TFs (Fig 4.4 and 4.7). When converted to blood, MAF-derived cells were characterised by a less differentiated state as indicated by increased number of haematopoietic progenitors and lower percentage of LIN^+ cells present in day 21 liquid cultures (Fig 4.2). Furthermore, the generation of $TER119^+$ erythroblasts was significantly reduced in experimental MAF cultures, which was also reflected by the lack of globinized erythroid colonies in the CFU-C assay (Fig 4.2). It is possible that the sustained transgene expression detected in late reprogrammed MAFs, but not MEFs, might have contributed to these differences (Fig 4.6). Particularly, the persistent expression of exogenous *Gata2* detected in day 24 reprogrammed MAFs could result in the observed inhibition of erythroid commitment. Indeed, GATA2 is required, in addition to its initial role in HSPCs, during terminal megakaryopoiesis and mast cells generation, but not erythroid maturation (Huang et al., 2009; Tsai and Orkin, 1997; Walsh et al., 2002). This type of bias in lineage output due to the continuous overexpression of cell fate determinants has been reported previously. More precisely, transdifferentiated endothelial cells subjected to sustained expression of *Pu.1* showed decreased lymphoid differentiation potential (Sandler et al., 2014).

One important goal in reprogramming is to generate transgene-independent cells harbouring a wide spectrum of functional capabilities, using the smallest possible set of lineage-instructive TFs. In this study the endogenous blood regulators were readily upregulated during the culture suggesting that internal haematopoietic GRNs were activated in reprogrammed fibroblasts (Fig 4.6). However even in MEFs, exhibiting significant downregulation of transgene expression, the silencing was not complete. This can preclude cells from obtaining their full functionality and impede their possible use in regenerative medicine. To circumvent this issue, I aimed to decrease the minimal set of factors necessary for reprogramming to a single master regulator. The removal of TFs did not prove possible for MAFs, as they could still be reprogrammed to half the efficiency only upon removal of ERG but not any other factor. However, MEFs showed a higher flexibility, by exhibiting a

significant yield of transition upon transduction with three (SCL, LMO2, ERG) or two (SCL, LMO2) TFs (Fig 4.7). SCL proved to be the only regulator indispensible for transdifferentiation, as its removal led to complete abolition of fate conversion. Following this finding, I used a dox inducible i*Scl* system to gain more control over expression of this crucial reprogramming factor. Moreover, in this set of experiments virally delivered LMO2 was coupled with dtTomato fluorescent protein that allowed selection of cells devoid of residual transgene expression. MEFs overexpressing two factors generated a similar spectrum of differentiated lineages as five-factor reprogrammed fibroblasts (Fig 4.9A). However, only 60% of terminally differentiated F4/80^{+} macrophages passively silenced viral transgene expression (Fig 4.9B). LMO2 is a known tumorigenic factor inducing self-renewal in T-lymphoid malignancies (McCormack et al., 2010). Possibly reprogrammed cells, even though not dependant on exogenous LMO2, will still select for pro-survival advantage associated with transgene overexpression. This emphasises the need to develop, in the next stage, inducible or transposon-based/floxed transgene cassette that would allow abrupt silencing of exogenous expression.

The analysis of globin switching in TER119^{+} erythroblasts suggested that SCL/LMO2-driven transdifferentiation led to the generation of cells without residual expression of fetal globins (Fig 4.9C). However, the overall efficiency of mature erythroid lineage outcome (CD71^{-}Ter119^{+} cells) was not increased. This issue could be potentially addressed by amending the culture conditions. For instance, the removal of EPO can facilitate post-EPO-dependent erythroblasts maturation. On the other hand, introducing erythroid differentiation denucleation factors could prompt reticulocyte formation (Xue et al., 2006). Furthermore, expression of maturation factors in addition to the cell fate determinants have been shown to increase the number and functionality of mature hepatocytes generated from fibroblasts (Du et al., 2014). Based on these findings, transient transduction of *Gata1* or *Pu.1*, alongside SCL and LMO2 activation, could tip the balance in multilineage progenitors towards erythroid/megakaryocytic or lymphoid differentiation respectively. Further enhancement of lymphoid commitment can be also obtained by transduction with additional B-cell (E2A or PAX5) and T-cell (T-Cell-Specific Transcription Factor 1 (TCF7), GATA binding protein 3 (GATA3)) specific TFs. In conclusion, the two-factor approach

provides a very useful tool that in the future may facilitate the investigation of molecular mechanisms governing transdifferentiation.

The results presented in this chapter prove that multilineage blood cells can be derived from fibroblasts of both adult and embryonic origins. This transition is rapid and does not include an intermediate pluripotency stage (Fig 4.5). The identification of the intermediate steps of transdifferentiation to blood and, if there are, alternative routes leading to reprogramming success are questions that remain to be addressed. The fact that the first cell surface marker detected in *in vitro* reprogramming cultures is CD41 (Fig 4.4) suggests that the transition mimics, at least in some aspects, blood development in the embryo. This hypothesis is more thoroughly assessed by functional assays and gene expression profiling presented in the next chapter of this thesis.

Chapter 5

5. Reprogramming by enforced expression of five TFs mimics embryonic development

5.1. Introduction

During embryogenesis, haematopoietic development occurs in successive waves. This process begins in the conceptus, where primitive erythrocytes are generated in the YS (Palis et al., 1999). The production of this transient population is shortly followed by the generation of definitive YS-derived EMPs (McGrath et al., 2015) and then HSCs originating from the AGM (Medvinsky and Dzierzak, 1996). Notably, an intimate spatial relation between endothelial and blood cells has been observed in both of these embryonic sites of haematopoiesis (Sabin et al., 1920; (Tavian et al., 1999). These findings led to the hypothesis of a common precursor capable of giving rise to both endothelial and haematopoietic lineages. Later, an *in vivo* bi-potential haemangioblast marked by the expression of Brachyury (BRY) and FLK1 was indeed identified within the posterior primitive streak of E7.5 embryos (Fehling et al., 2003; Huber et al., 2004; Lacaud et al., 2004). Cells with the same bi-lineage potential were also isolated from E7 YS on the basis of dual *Gata1* and *Cdh5* expression (Yokomizo et al., 2007).

Haematopoietic emergence in the DA is first evident through the generation of tight clusters budding from the endothelial wall after E9.5 (Dieterlen-Lievre et al., 2006; Yokomizo and Dzierzak, 2010). The appearance of these intermediate structures correlates with the establishment of HSC activity within the AGM (Garcia-Porrero et al., 1995; Garcia-Porrero et al., 1998; Medvinsky and Dzierzak, 1996). Lineage tracing of this region suggests that the endothelium underlying clusters is the source of the emerging HSCs (Jaffredo et al., 1998; Zovein et al., 2008). These specialised endothelial cells, known as HE, were later visualised using live *in vivo* imaging. An elegant body of work by

Boisset and colleagues showed that the rare SCA1^{+}cKIT^{+}CD41^{+} haematopoietic cells could be clearly seen emerging by budding from the CD31^{+} ventral lining of the mouse DA (Boisset et al., 2010).

The inaccessibility of embryos at the early stages of development as well as rarity of emerging cellular populations makes it very challenging to investigate the events regulating the earliest stages of lineage induction and specification. To address this issue, the *in vitro* ESC culture system was developed. ESCs can be expanded as pure population of undifferentiated cells for extended periods of time, therefore enabling large-scale studies. Moreover, when subjected to well-established culture conditions, they acquire desired cellular fate in a controlled and stepwise manner. All these features make this system an invaluable tool for studying early events of embryonic development and characterising rare populations emerging during lineage specification (reviewed by Keller, 2005). In the presence of Bmp4, Wnt and Activin, mimicking embryonic development, ESC differentiate toward BRY^{+} mesoderm, which upon stimulation with VEGF and IL6 further differentiate towards blood (Nostro et al., 2008; Pearson et al., 2008). This process mainly recapitulates the YS stage of embryonic haematopoietic development. The *in vitro* counterpart of the haemangioblast can be detected between day 2.5-3.5 of culture and is marked by the FLK1 marker (Fehling et al., 2003) (Fig 5.1). The later stages of blood lineage specification involve EHT manifested by gradual acquisition of haematopoietic cell surface markers and downregulation of endothelial programme in HE cells (Costa et al., 2012; Eilken et al., 2009; Lancrin et al., 2009). Multiple immunophenotypes have been employed to better characterise this transition. More particularly, CDH5^{+}CD45^{-} cells have been shown to give rise to cKIT^{+}CD45^{+} progenitor and cKIT^{-}CD45^{+} or TER119^{+} mature blood cells (Nishikawa et al., 1998). Eilken and colleagues have also reported that CD45 acquisition is preceded by CD41 expression which marks the earliest events of haematopoietic commitment *in vivo* (Eilken et al., 2009; Ferkowicz et al., 2003; Mikkola et al., 2003; Sroczynska et al., 2009b). Furthermore, it has been demonstrated that a Tie2^{+}cKIT^{+}CD41^{-} population isolated from the ESC/EB system also possess haemogenic potential (Lancrin et al., 2009). Altogether, these reports identified several endothelial and haematopoietic cell surface

markers that are specifically modulated during embryonic HSPC generation (Fig 5.1).

In the previous chapter I was able to demonstrate that CD41 is haematopoietic marker detected specifically on blood cells emerging from fibroblasts during reprogramming. To further identify the intermediate stages of this process, I employed the well-defined immunophenotypical signatures of HE and HSPCs from the ESC system to isolate and characterise cellular populations emerging in reprogramming culture before terminal differentiation of blood cells.

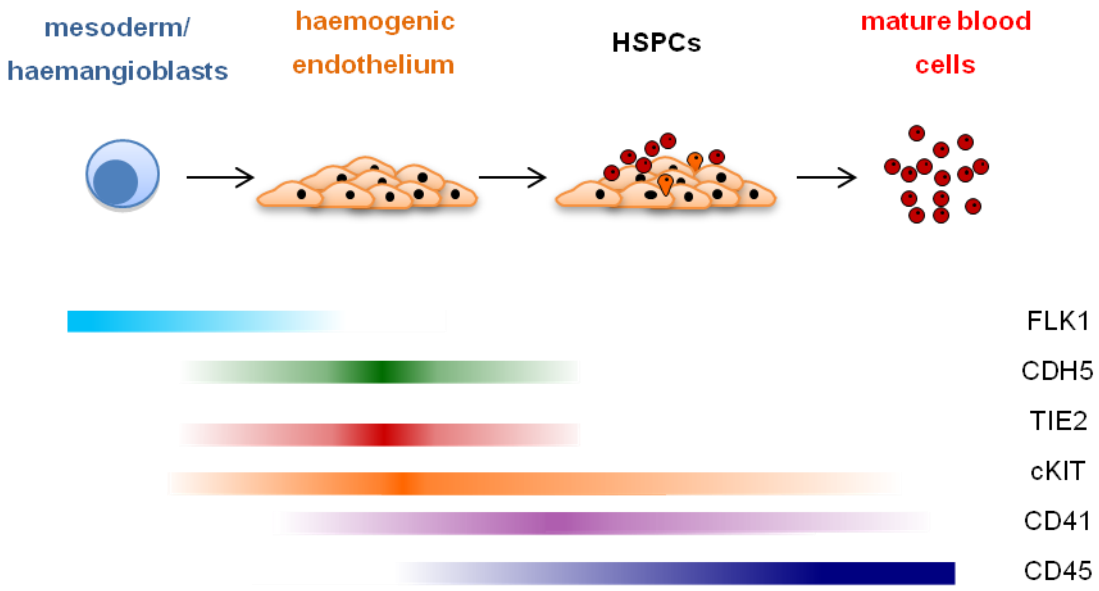


Figure 5.1. Schematic representation of cellular stages involved in the generation of haematopoietic cells during *in vitro* ESC differentiation culture.

The embryonic development of the haematopoietic system has been well recapitulated *in vitro* with the use of the ESC differentiation model. Various intermediate stages during blood specification have been separated and defined on an immunophenotypical level. Firstly, mesoderm differentiates to the FLK1^{+} equivalent of haemangioblasts. These cells, under specific culture conditions, give rise to HE marked by the expression of both endothelial (CDH5, TIE2) and haematopoietic (cKIT) markers. The earliest signs of haematopoietic commitment can be measured by increased surface expression of CD41 which is then accompanied by a marker of definitive haematopoiesis, CD45.

5.2. Results

5.2.1. The multilineage progenitor potential is restricted to early appearing cKIT$^+$ cells

I have previously shown that at day 21 reprogrammed cells contain a cKIT$^+$ subpopulation and exhibit progenitor potential (Fig 4.1 and 4.2). Subsequent flow cytometry analysis of cKIT acquisition at the initial stages of culture revealed that immunophenotypical progenitors are present as early as day 6 (Fig 5.2A). Immunostaining of the emerging colonies confirmed the association of cKIT surface expression with blood-like cell morphology (Fig 5.2B). The nature of these cells was confirmed after re-plating day 12 cKIT positive and negative fractions which showed that the CFU potential was mainly restricted to cKIT$^+$ cells (Fig 5.2C). To determine the frequency of functional MEF-derived haematopoietic progenitors within cKIT$^+$ population, I also performed limiting dilution assay (LDA). I observed that 1 out of 27 cells gave rise to haematopoietic colonies (Fig 5.3A). Still it was unclear whether haematopoietic colonies were generated from unipotent or multipotent progenitors. To address the question of a potency of reprogrammed precursors, I sorted single cKIT$^+$ cells and assessed their capability to give rise to different blood lineages. Upon culture on OP9 stroma in haematopoietic medium single precursors gave rise to myeloid, erythroid and megakaryocytic lineages, therefore demonstrating multilineage properties (Fig 5.3B). To further confirm this finding, I isolated day 21 CD11b$^+$ and TER119$^+$ reprogrammed cells and investigated the integration of exogenous factors. Interestingly, the levels of genomic transgene integration were similar between the two lineages suggesting that either they have the same TF requirements or they originate from the common precursor (Fig 5.3C).

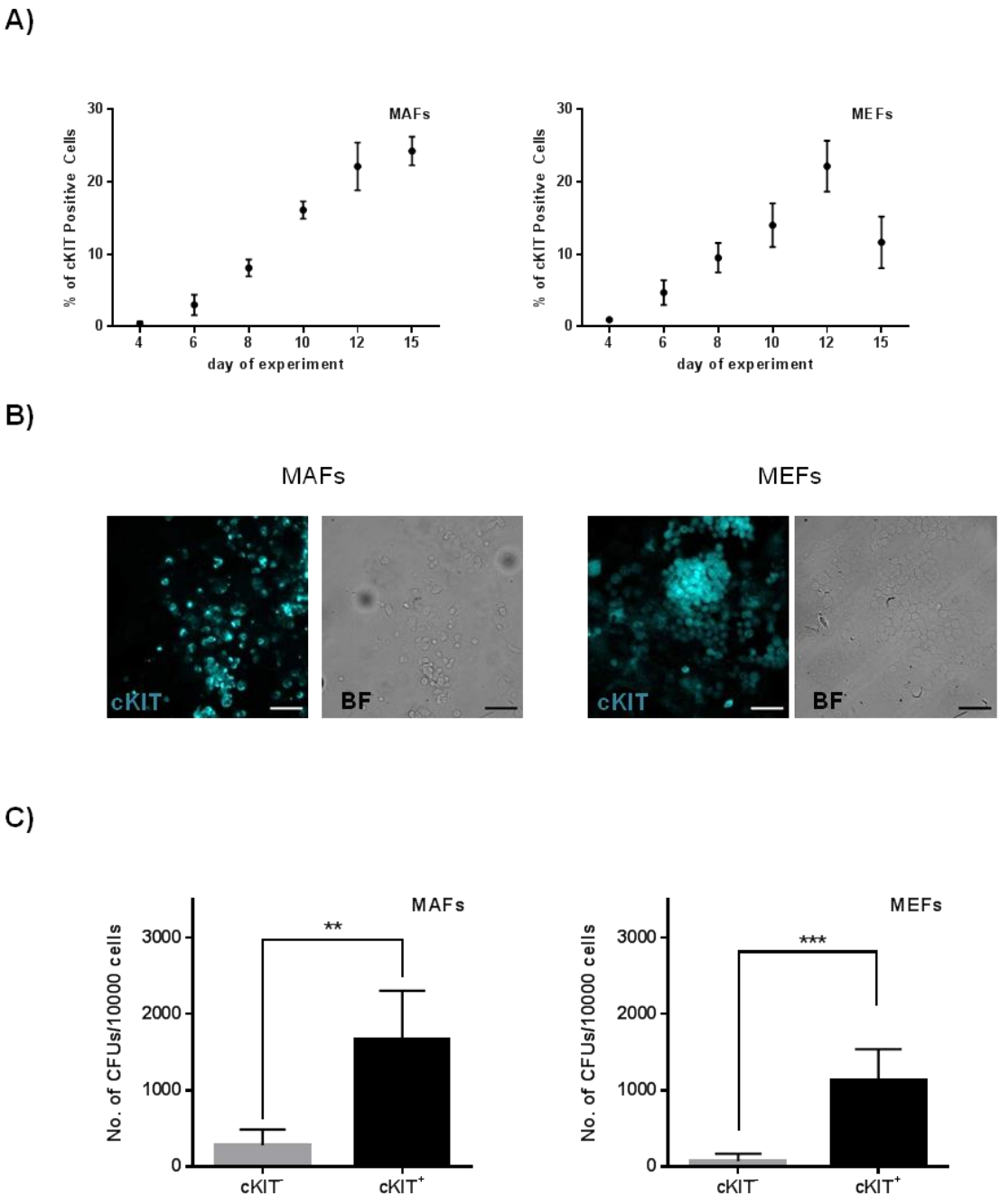


Figure 5.2. Haematopoietic progenitors are generated at the early stages of reprogramming.

A) Flow cytometry profile of cKIT surface expression on transitioning MEFs/MAFs at the indicated time points (mean ± SEM, n=3 (MAFs), n=4 (MEFs). **B)** Immunostaining of cKIT$^+$ day 12 reprogrammed MEFs/MAFs. **C)** The number of colonies generated in CFU-C assay by day 12 sorted cKIT positive and negative fractions of reprogrammed MEFs/MAFs (mean ± SEM, n=3 (MEFs), n=2 (MAFs)). Scale bars represent 50 μm. Asterisks represent significant differences (paired Student t-test, ** $p< 0.005$, *** $p< 0.0005$). BF – bright field.

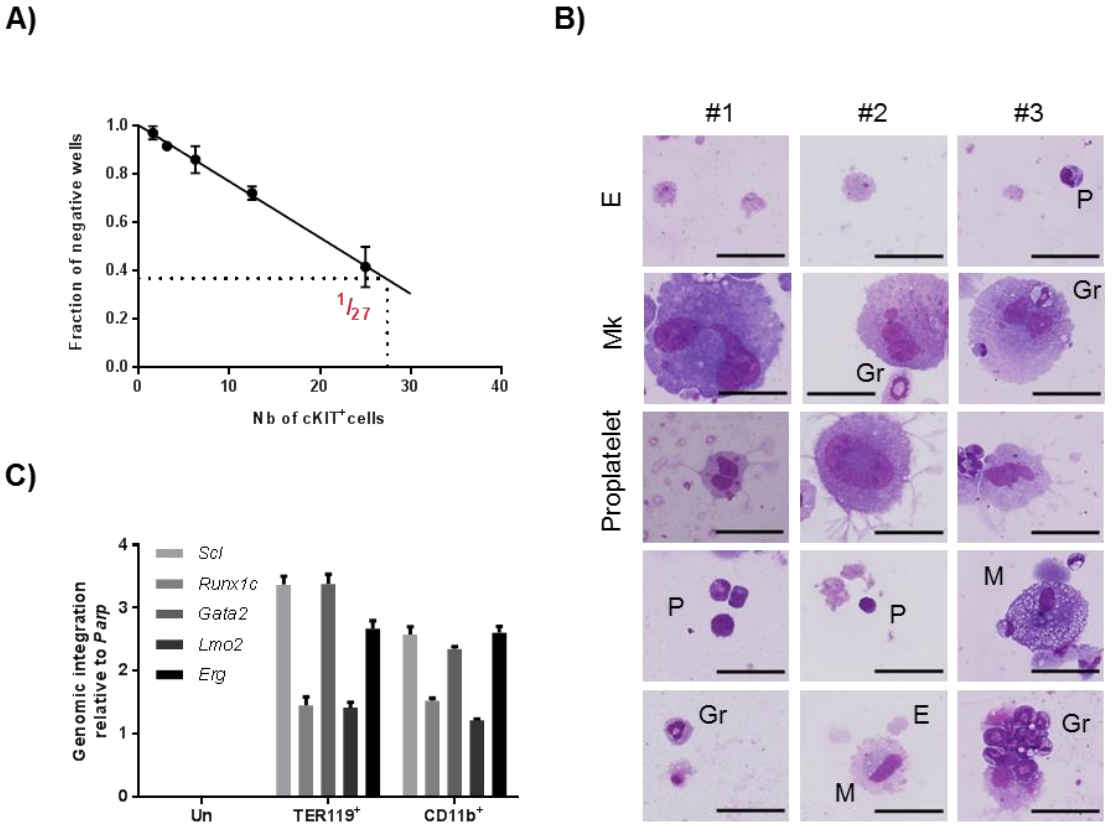


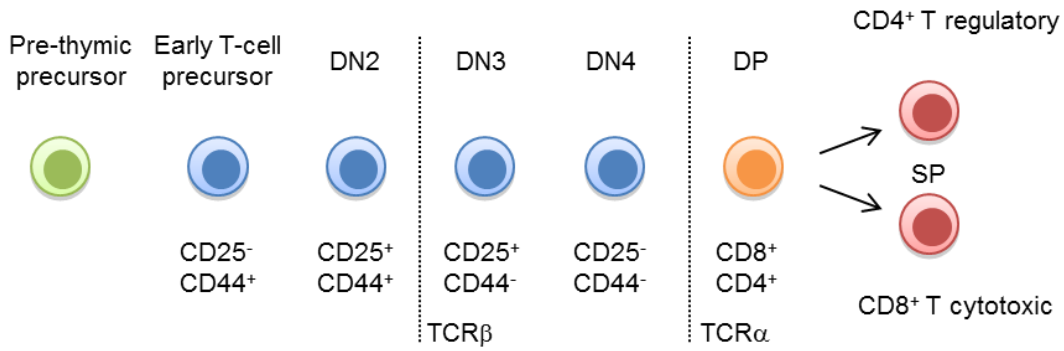
Figure 5.3. Blood progenitors generated from MEFs have multilineage potential.

A) LDA performed on cKIT$^+$ cells sorted from day 12 reprogrammed MEF cultures (mean ± SEM, n=3). **B)** Morphology of cells generated from day 12 sorted MEF-derived cKIT$^+$ cell matured on OP9 stroma in haematopoietic medium. Representative images from 3 independent experiments (#1,2,3) are shown (n=6). **C)** Relative genomic integration of vectors for all five TFs in untransduced (Un) and reprogrammed MEFs in respect to the endogenous *Parp* gene. MEFs transduced with five TFs were cultured on OP9. On day 21 cells were sorted into erythroid (TER119$^+$) or myeloid (CD11b$^+$) fractions and subjected to genomic DNA extractions and subsequent qRT-PCR. Data presented are mean ± SD of a representative experiment performed in triplicates (n=2). Scale bars represent 50 μm. E – erythrocyte, M – macrophage, Gr- granulocyte, P – progenitor, Mk – megakaryocyte,

Day 12 cKIT$^+$ progenitors were also analyzed for their lymphoid potential. T-lymphopoiesis can be supported *in vitro* by using OP9-DL1 stromal cells (Schmitt and Zuniga-Pflucker, 2002) and media containing FLT3L and low levels of IL7. Under these conditions, I observed that MEF-derived progenitors acquired some features of immature T-cells. The ontogeny of T-lymphocytes consists of multiple steps (Fig 5.4A). The initial transition, manifested by temporal CD25 expression (DN), is followed by the acquisition of both CD4 and

CD8 markers (DP). Mature T-lymphocytes then become either $CD4^+$ T regulatory or $CD8^+$ T cytotoxic cells (SP). During the initial stages of culture on OP9-DL1 stroma, $cKIT^+$ progenitors acquired the CD25 marker but did not proceed toward further stages of maturation (Fig 5.4Bi). It has been reported that GSK3 inhibitor CHIR99021 allows for specification of T-cells from human ESCs (Sturgeon et al., 2014). Addition of this compound to reprogramming cultures pushed T-cell differentiation toward the next maturation stage of $CD25^-$ progeny, whereas no-compound control continued to accumulate $CD25^+$ cells (Fig 5.4Bi). Importantly, the CHIR99021-treated $CD25^-$ subpopulation acquired both CD4 and CD8 markers (Fig 5.4Bii). Furthermore, both cultures contained cells that underwent TCR-β chain gene rearrangements attributed to DN stage of T-cell ontogenesis (Fig 5.4A and C). Taken together these results indicate that blood progenitors are present as early as day 12 of reprogramming and exhibit multilineage haematopoietic differentiation potential. The observed limited lymphoid maturation could be attributed to suboptimal *in vitro* culture conditions.

A)



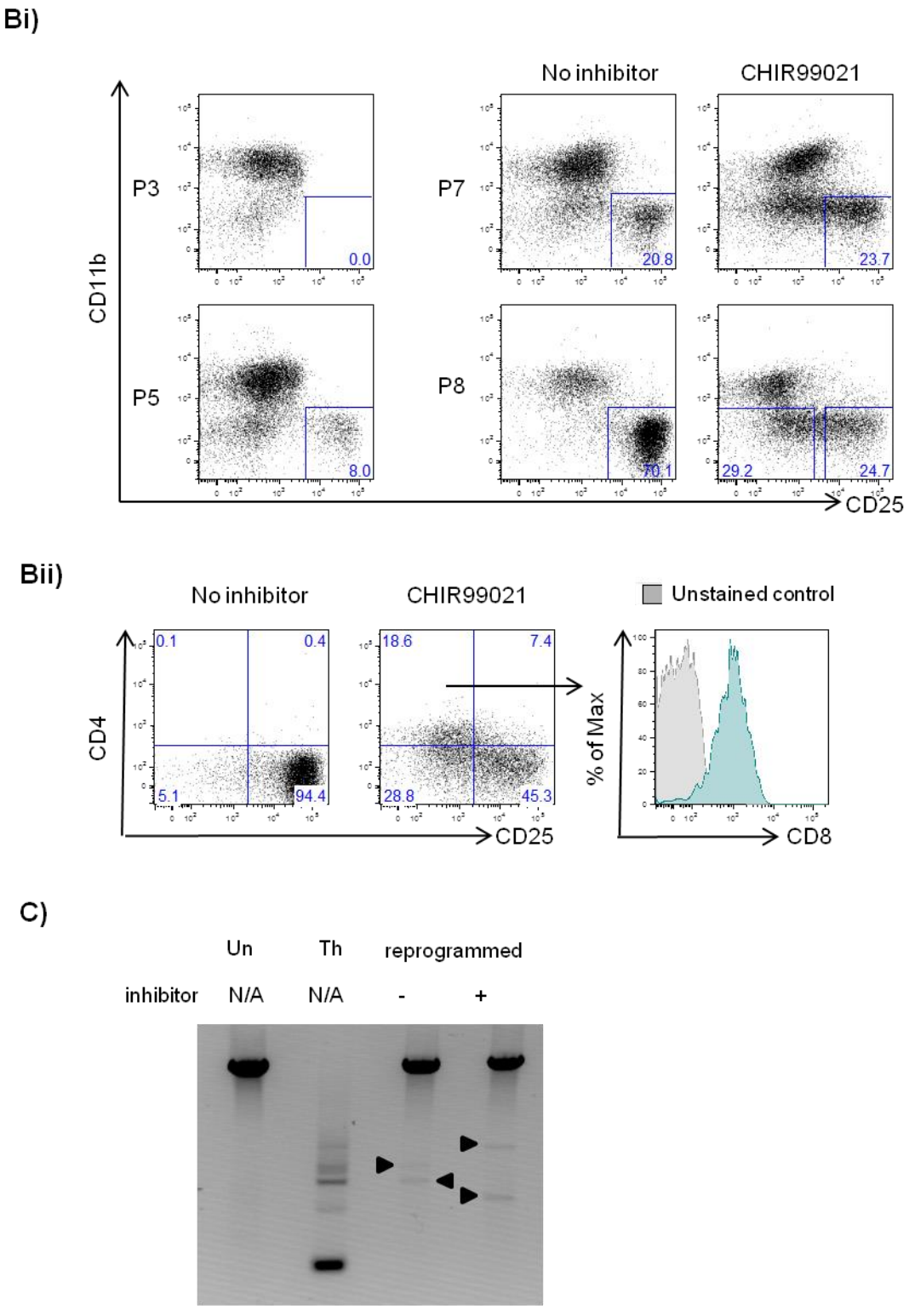


Figure 5.4. CKIT progenitors display limited T-lymphoid potential.

A) Schematic representation of T-cell differentiation stages. **B)** Representative flow cytometry plots depicting acquisition of CD25 **(i)** and CD4/CD8 **(ii)** cell surface markers by day 12 sorted cKIT$^+$ cells maintained on OP9-DL1 in lymphoid medium with or without CHIR99021 compound. **C)** TCRβ rearrangement detection in CD25 sorted reprogrammed MEFs, control untransduced MEFs (Un) and thymus cells (Th). Arrows indicate rearrangement bands. DN – CD4/CD8 double negative fraction, DP - CD4/CD8 double positive fraction, SP - CD4 or CD8 single positive fraction.

5.2.2. Reprogrammed cKIT$^+$ MEF-derived progenitors exhibit short-term engraftment potential

An essential aspect of HSC characteristics is their ability to provide long-term haematopoietic reconstitution of irradiated recipients. To facilitate the prospective use of reprogrammed cells in regenerative medicine we need to recapitulate this unique HSC ability. In this context, I performed repopulation assay by injecting cKIT sorted cells from reprogramming cultures in lethally irradiated NSG mice. Transplantations were performed with MEFs isolated from a reporter mice expressing *Gfp* under the control of the pan-haematopoietic *Ai* gene promoter (Ferreras et al., 2011). The level of engraftment was evaluated by quantifying the percentage of GFP$^+$ donor cells in PB of recipient mice. In the first *in vivo* experiment, we assayed the cKIT$^+$ cells isolated from day 12 cultures (Fig 5.5A). Furthermore, in the aim of increasing the production of engrafting cells, we also introduced a variation to the protocol in the form of OP9-DL1 stroma co-culture maturation/amplification step (Fig 5.5A). The rational to provide stroma support is that Notch signalling is well known to play a crucial role in maintaining the balance between HSPC self-renewal and differentiation. For example, soluble Delta like 1 ligand increases the proliferation of CD34$^+$CD38$^-$ human HSCs *in vitro* and enhances their long-term repopulation capabilities (Karanu et al., 2001). Moreover, in the BM, high secretion of Jagged1 ligand by osteoblasts was found to increase the pool of HSCs (Calvi et al., 2003). Altogether, Notch signalling plays a crucial role in the generation of HSCs during ontogeny and allows the maintenance of undifferentiated HSCs *in vitro*. Interestingly, gene expression analysis of ligands and downstream targets of the Notch pathway in day 4, 12 and 21 reprogrammed fibroblasts revealed a dramatic decrease in their mRNA levels, which was evident as early as day 12 (Fig 5.5B). This finding reassured me that sustained provision of Notch stimulation might result in the increased output of cells capable of murine engraftment (Benveniste et al., 2014; Kumano et al., 2003). Indeed, co-culture of the day 12 reprogrammed cells with OP9-DL1 stroma enhanced propagation/generation of immunophenotypical LIN$^-$ SCA1$^+$cKIT$^+$ (LSK) HSPCs (Fig 5.5C).

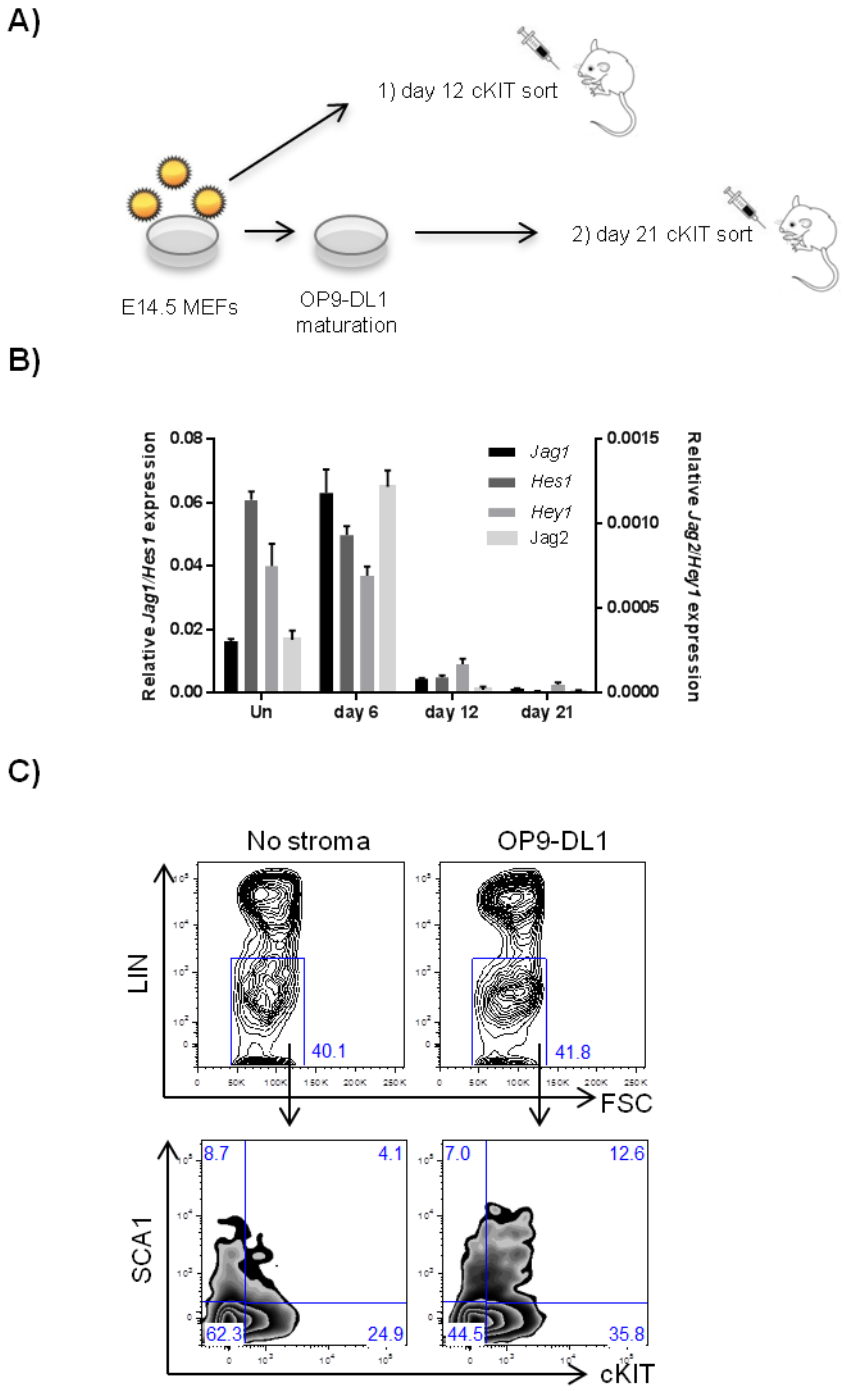


Figure 5.5. The effect of OPD-DL1 maturation on reprogrammed LSK population.

A) Schematic representation of transplantation experiment strategy. **B)** Relative gene expression levels of the Notch pathway associated molecules in respect to *β-actin* on day 6, 12 and 21 of reprogramming. Data presented in this figure are mean ± SD of a representative experiment performed in triplicate. **C)** LSK population in day 12 reprogrammed MEFs maintained on gelatin or OP9-DL1. FSC- forward scatter

A more detailed analysis of LSK and Lin$^-$SCA1$^-$cKIT$^+$ fractions revealed a skewed distribution in reprogrammed progenitors in comparison to BM cells (Fig 5.6). Almost no LSK CD50$^+$CD48$^-$ HSCs were detected within MEF-derived blood cells (Fig 5.6Ai). In these cultures CD150$^-$CD48$^-$ ST-HSCs were highly overrepresented (92%) with a concomitant significant decrease in the number of CD150$^-$CD48$^+$ MPPs. The Lin$^-$SCA1$^-$cKIT$^+$ fraction of reprogrammed cells consisted mostly of MEPs and trace amounts of CMPs and GMPs (Fig 5.6Aii).

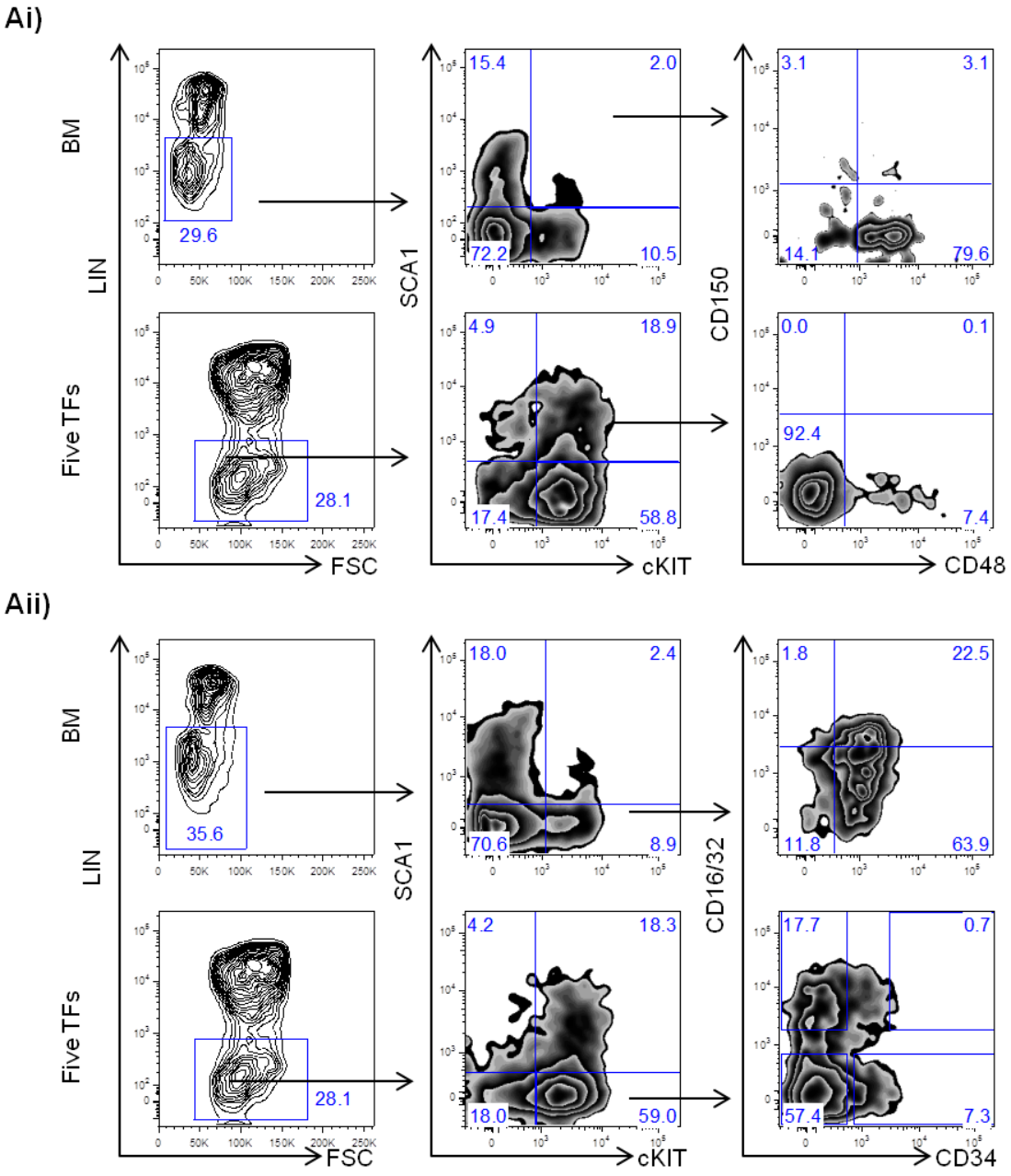


Figure 5.6. Immunophenotypical characterisation of reprogrammed HSPCs

Flow cytometry of HSPC populations within LSK **(Aii)** and LIN$^-$SCA1$^-$cKIT$^+$ **(Aii)** fractions of reprogrammed MEFs upon OPD-DL1 maturation and control BM.

Importantly, the effect of OP9-DL1 culture step was also reflected in repopulation assay. Whereas day 12 precursors did not exhibit engraftment capabilities, cells expanded on OP9-DL1 stroma contributed to the murine haematopoietic system (Fig 5.7A). However, the observed effect was very transient and limited to only 2 weeks post injection. Moreover, the reprogrammed cells contributed solely to the erythroid compartment (Fig 5.7B). These results indicate that the five reprogramming factors can only confer limited engraftment potential to MEFs. Further optimisation steps are therefore likely to be required to obtain multilineage long-term repopulation of reprogrammed cells.

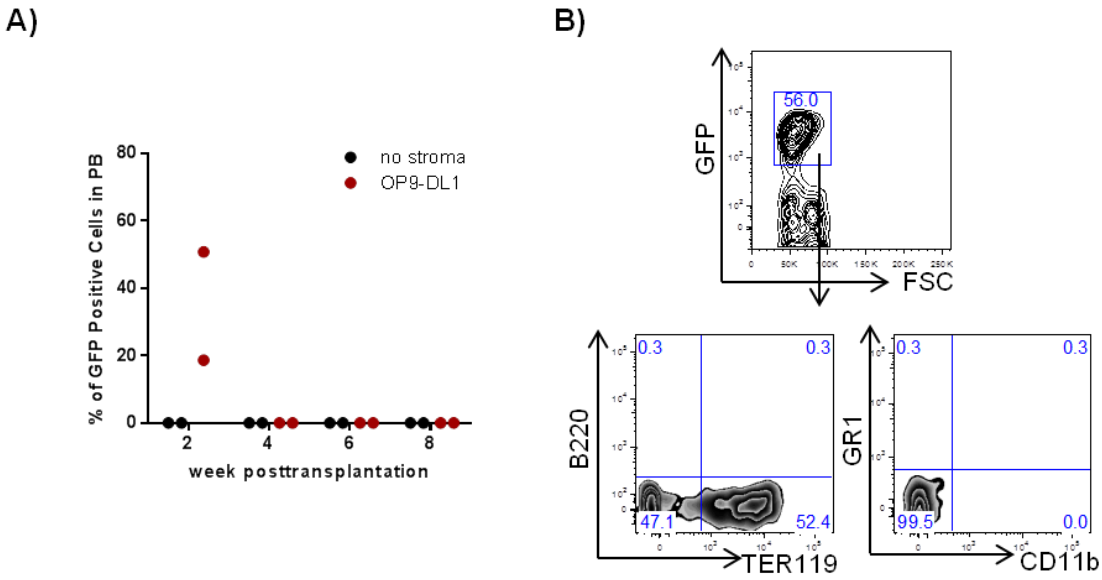


Figure 5.7. Short-term engraftment of cKIT⁺ reprogrammed MEFs.

A) Level of engraftment in recipients' PB at weeks 2 to 8 posttransplantation measured by expression of *Gfp* driven from *Ai* pan-haematopoietic gene promoter. Two separate experiments were combined on this graph – with (red) or without (black) OP9-DL1 maturation **B)** Lineage commitment of GFP⁺ reprogrammed MEFs in PB of mouse injected with cKIT⁺ cells after OP9-DL1 maturation.

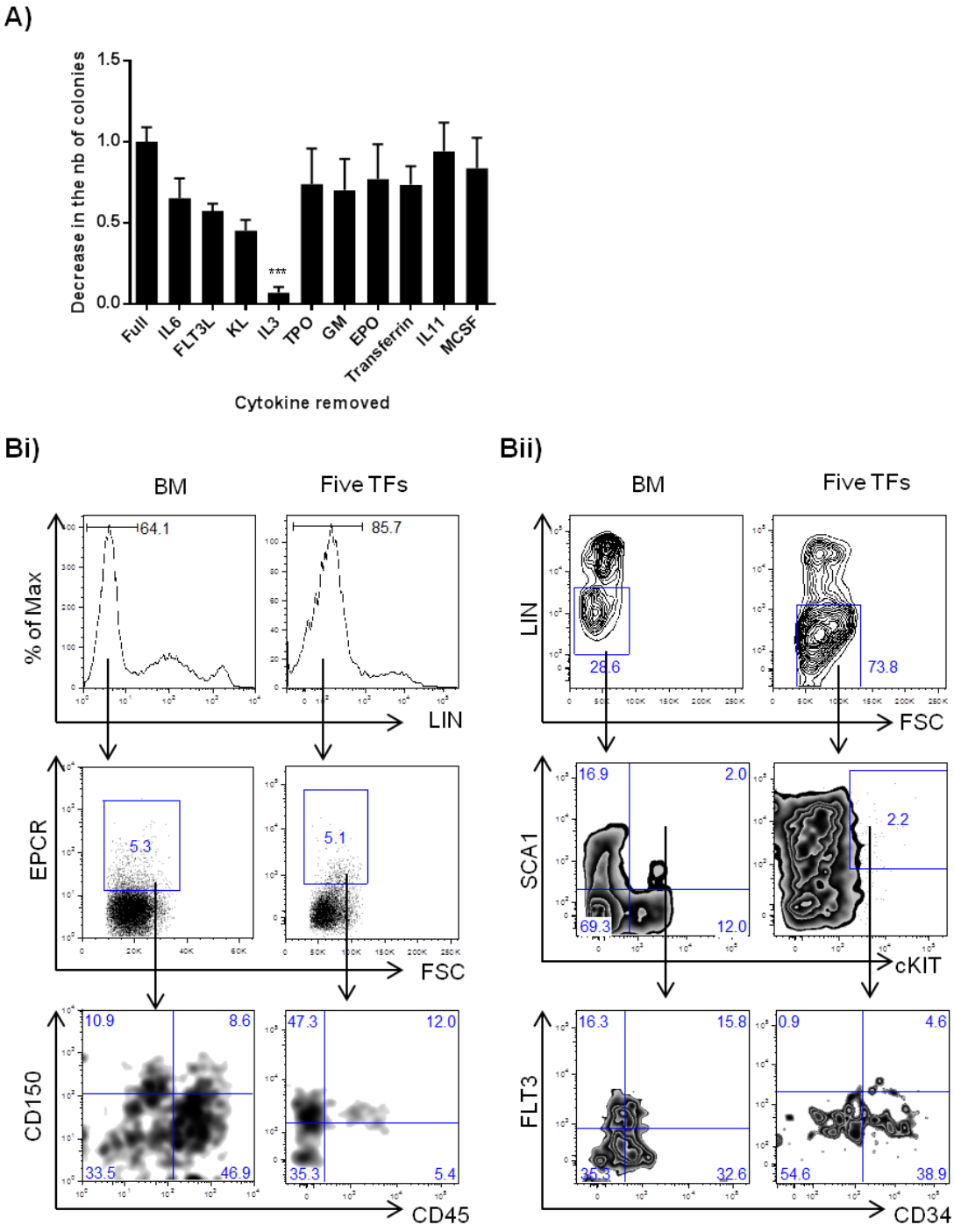


Figure 5.8. HSPC profile of reprogrammed cells.

A) Decrease in the number of blood colonies generated from transduced MEFs upon removal of the indicated cytokines from reprogramming medium (mean ± SEM, n=2). Statistical significance was established using one-way ANOVA followed by post hoc Dunnett's multiple comparison test (***p<0.0005). **B)** Immunophenotypical characterisation of HSCs generated from reprogrammed MEFs upon serum free culture with IL3, KL, IL6 and FLT3L supplementation defined as LIN$^-$EPCR$^+$CD150$^+$CD45$^+$ **(i)** or LIN-SCA1+cKIT+FLT3$^+$CD34$^+$ **(ii)** fraction.

It is known that maintaining HSPCs *in vitro* is a very challenging task due to their spontaneous differentiation, which is often driven by the presence of cytokines in the culture medium. For instance, M-CSF was shown to direct myeloid fate at the HSC level (Mossadegh-Keller et al., 2013). In order to minimize this effect, I optimized the culture conditions by stepwise removal of cytokines from the original haematopoietic culture medium. The decrease in the number of blood colonies clearly showed that IL3, KL, IL6 and FLT3L, all widely used for the maintenance of haematopoietic progenitors, were also those crucial for reprogramming (Fig 5.8A). Implementation of the limited cytokine mix, together with serum free culture conditions, led to the generation of a substantial pool of immunophenotypical HSCs by day 19 of culture (Fig 5.8B). The variability in the viral batches used for reprogramming has prevented, however, the ideal reproduction of these results.

5.2.3. *P53* knock-out facilitates reprogramming

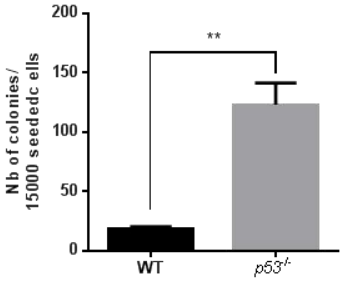
Loss of P53, which acts as a 'genome guardian', leads to the increased number of genomic mutations but also higher cell proliferation and lower susceptibility to apoptosis. It is therefore not surprising that deletion of *p53* was reported as a factor increasing the efficiency of iPSC generation (Utikal et al., 2009b). I tested if a similar effect can be observed during transdifferentiation of fibroblast to blood. *P53*$^{-/-}$ MEFs were reprogrammed with significantly higher efficiency than their WT counterparts (Fig 5.9A). Additionally, the number of cKIT progenitors peaked as early as day 8 demonstrating the accelerated kinetics of the transition (Fig 5.9B). By day 11, I was able to observe the first signs of differentiation towards mature blood cells marked by the appearance of CD45. Increased susceptibility of *p53*$^{-/-}$ MEFs for reprogramming was also translated into a higher clonogenic potential (Fig 5.9C). Similar to WT MEFs, the multilineage capability of blood precursors was confirmed by the presence of myeloid, erythroid, megakaryocytic and mixed colonies. The change in CFUs numbers on consecutive days was closely associated with cKIT expression pattern in liquid cultures.

To confirm that higher plasticity of *p53*$^{-/-}$ MEFs leads to a more profound fate switching, I sorted cKIT^{+} cells and cultured them in various conditions supporting differentiation of haematopoietic progenitors towards different lineages. In this respect, supplementation with IL3, TPO and IL6 led to significant megakaryocytic commitment which was further confirmed by acetylcholinesterase staining (Fig 5.9D). To enhance formation of erythrocytes, I cultured the cells in a limited medium (supplemented with KL, IL3, EPO, transferrin) (ERY) or re-seeded sorted cells on OP9 stroma (LBM/stroma). In comparison to default conditions (LBM), OP9 co-culture resulted in very significant erythroid differentiation reaching 40% of the total population (Fig 5.9E).

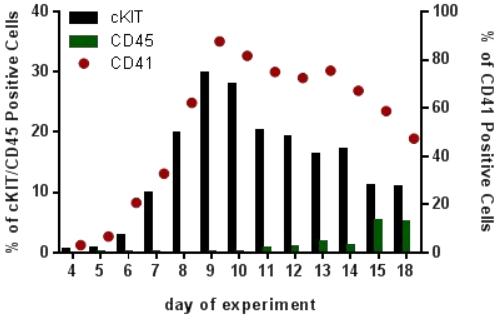
The T-lymphoid potential of cKIT^{+} cells originating from *p53*$^{-/-}$ MEFs was comparable to WT MEF-derived counterparts (Fig 5.4). Notably, at late passages I was able to detect the emergence of a small CD3^{+} population, which was, however, not expanding beyond the frequency of 1% (Fig 5.9F).

Due to more robust *in vitro* potential of *p53*$^{-/-}$ reprogrammed MEFs, I sought to establish whether they will also exhibit more substantial *in vivo* functionality than WT counterpart. For this, day 12 cKIT^{+} precursors of both genotypes were sorted and injected intra-femorally into partially irradiated congenic mice. Consequently, all recipients of *p53*$^{-/-}$ reprogrammed cells died by day 50 of experiment (Fig 5.9G). Whereas no engraftment could be detected based on CD45.2 expression, presence of transdifferentiated cells in PB was confirmed by genotyping (work by Kiran Batta). Collectively, even though detailed histological studies were not performed, it can be presumed that *p53*$^{-/-}$ MEF-derived cells proliferated in recipient mice and led to multiorgan failure and anemia.

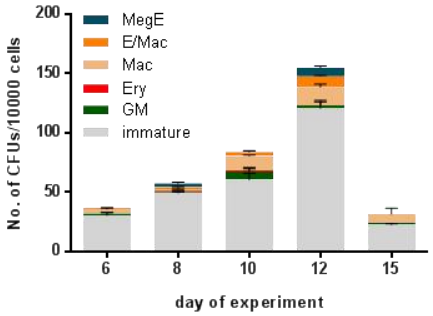
A)


Nb of colonies/ 15000 seededc ells
200
150
100
50
0
**
WT
p53-/-

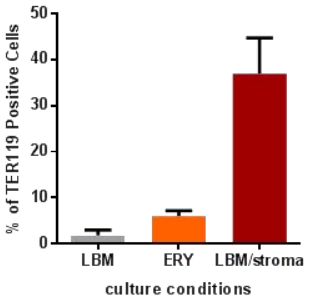
B)


% of cKIT/CD45 Positive Cells
% of CD41 Positive Cells
cKIT
CD45
CD41
40
30
20
10
0
100
80
60
40
20
0
4 5 6 7 8 9 10 11 12 13 14 15 18
day of experiment

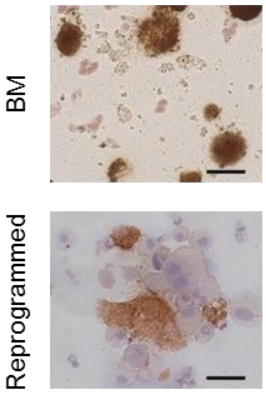
C)


No. of CFUs/10000 cells
MegE
E/Mac
Mac
Ery
GM
immature
200
150
100
50
0
6 8 10 12 15
day of experiment

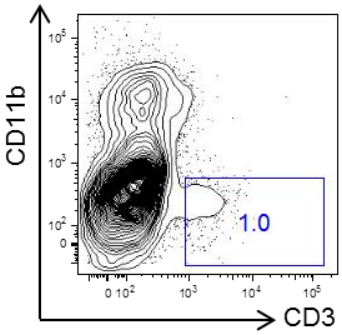
E)


% of TER119 Positive Cells
50
40
30
20
10
0
LBM
ERY
LBM/stroma
culture conditions

D)


BM
Reprogrammed

F)


CD11b
CD3
1.0

G)

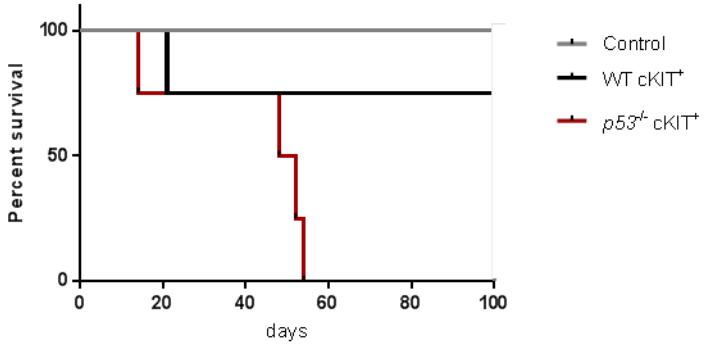
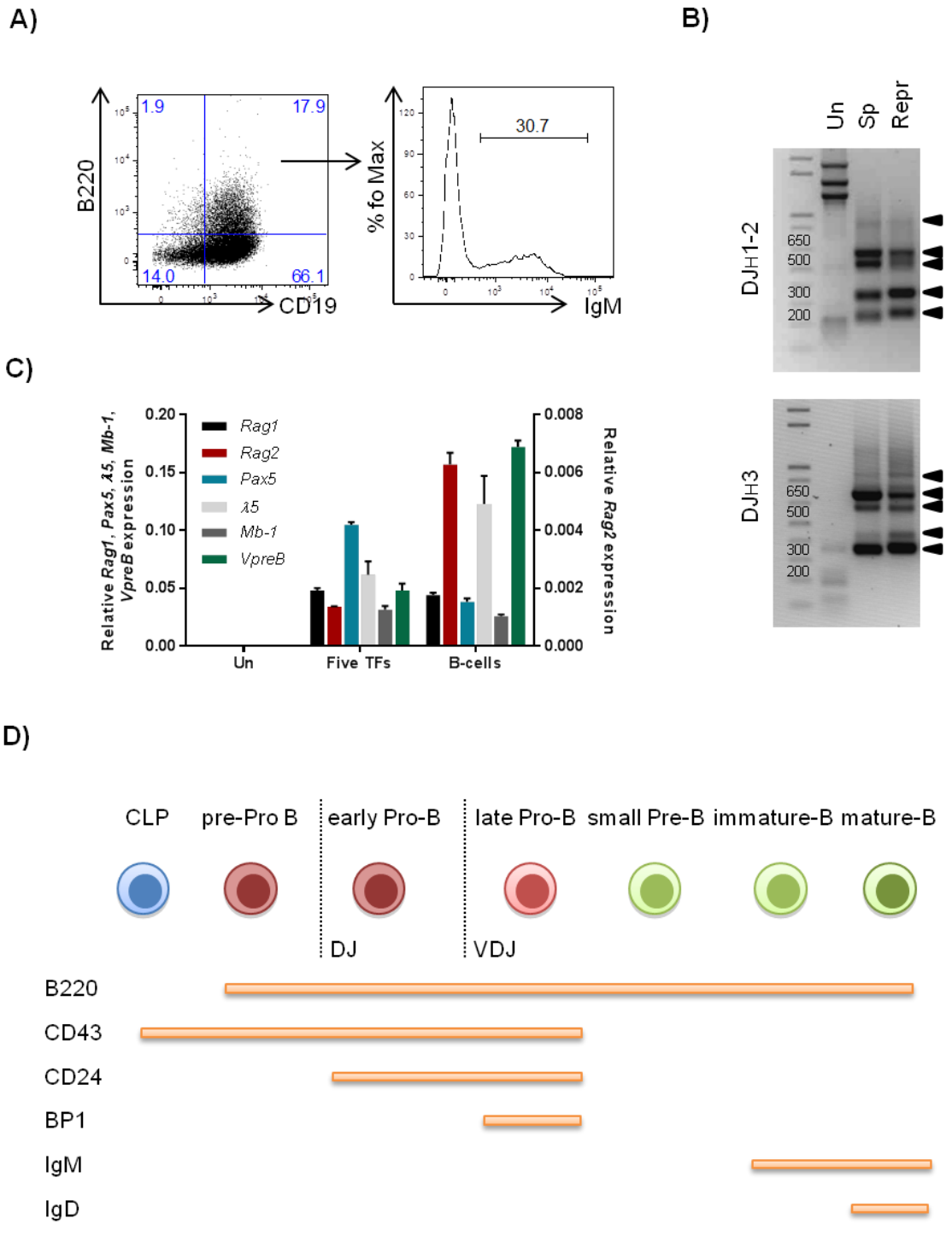


Figure 5.9. *P53* knock-out facilitates reprogramming.

A) The number of round colonies generated from WT or *p53*$^{-/-}$ MEFs between day 9 and 11 of reprogramming (mean ± SEM, n=2). **B)** Flow cytometry profile of cKIT, CD41 and CD45 acquisition during the course of *p53*$^{-/-}$ MEFs reprogramming. Data presented originate from a representative experiment (n=2) **C)** Changes in the CFU-C potential of *p53*-null MEF-derived blood cells between day 6 and 15 of reprogramming. Data presented are mean ± SEM of a representative experiment performed in triplicate (n=3) **D)** Acetylcholinesterase staining of megakaryocytes derived from BM or generated by day 12 sorted cKIT^{+} reprogrammed *p53*$^{-/-}$ MEFs. **E)** The level of erythroid commitment of *p53*$^{-/-}$ reprogrammed MEFs in different culture conditions (mean ± SEM, n=2). **F)** CD3 staining of cKIT^{+} cells cultured in T-lymphoid conditions. **G)** Survival curve of mice injected with WT MEF-derived (n=4) or *p53*$^{-/-}$ MEF-derived (n=4) cKIT^{+} progenitors. Control mice were only subjected to sublethal irradiation Scale bars represent 50 μm. Asterisks represent significant differences (paired Student t-test, ** p< 0.005). LBM – liquid big mix, ERY – erythroid medium (EPO, IL3, KL, transferrin), Ery – erythroid CFU, E/Mac – mixed erythroid-macrophagic CFU, Mac – macrophagic CFU, GM – granulocytic-macrophagic CFU, MegE – megakaryocytic CFU.

Lastly, B-cell differentiation was achieved by multiple passages of cKIT sorted cells on OP9 stroma in the presence of FLT3L and IL7. After few weeks the cultures began to generate CD19/B220 double positive cells, a fraction of which was also expressing IgM (Fig 5.10A). Cells expressing B220 were then sorted and tested for the BCR rearrangements. Recombination of DJ junctions could be detected in reprogrammed and spleen cells but not control MEFs (Fig 5.10B). The identity of bands was confirmed by IgBlast tool. To further assess the lymphoid potential of the reprogrammed cells, I tested lineage specific gene expression (Fig 5.10C). Importantly, all major B-cell fate-associated regulators, including *Pax5*, *Rag1* and *Rag2*, were present in purified B220^{+} reprogrammed cells. B-lymphoid development, similar to T-cell differentiation, is marked by multiple maturation stages (Fig 5.10D) that can be distinguished on the basis of specific cell surface markers (Fig 5.10D and Ei). The two main subpopulations are mature cells expressing both B220 and CD43 and the less mature fraction displaying only B220 (Hardy and Hayakawa, 1991). Using this strategy, I aimed

to determine the spectrum of maturation stages within the reprogrammed B-cells in comparison to WT BM. In both tested samples, mature cells marked by $B220^{+}CD43^{-}$ immunophenotype and immunoglobulin presentation contributed similarly to the overall population (Fig 5.10Eii right bottom panel). In comparison to BM control, in reprogrammed cells I observed a higher frequency of more mature IgM-expressing subpopulation than small Pro-B stage.



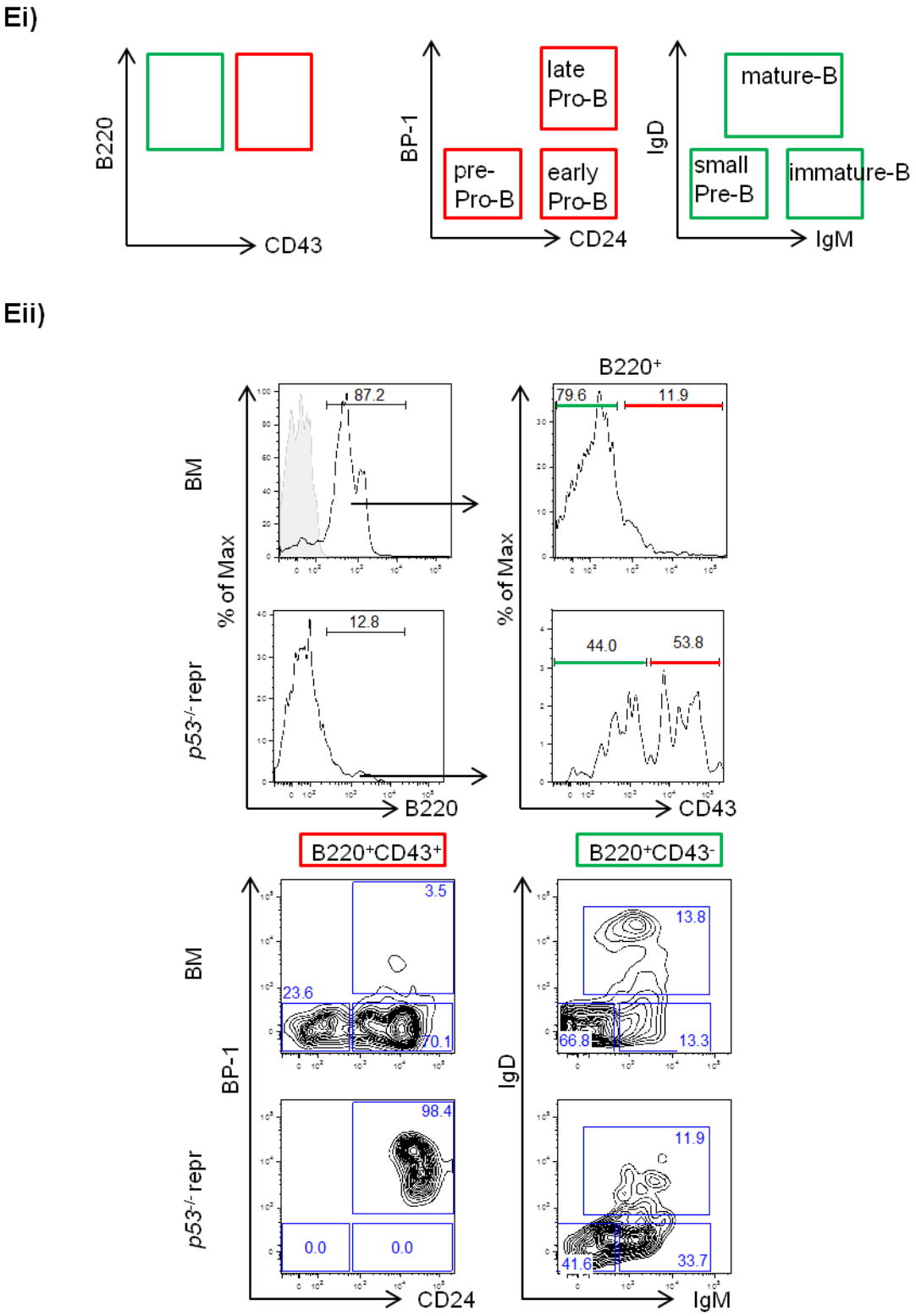


Figure 5.10. *P53*$^{-/-}$ MEF-derived cKIT^{+} progenitors generate B-lymphocytes.

A) Flow cytometry plots presenting acquisition of B-lymphoid markers by cKIT day 12 sorted *p53*$^{-/-}$ MEF-derived progenitors during long-term culture on OP9 stroma.

However, all three subpopulations were readily represented in both samples. The reprogrammed cells with mature B-cell phenotype were more frequently found in the early $B220^{+}CD43^{+}$ fraction (Fig 5.8Eii left bottom panel). Whereas BM cells were mostly consisting of pre Pro-B and Pro-B populations, the reprogrammed B-cells were solely restricted to late Pro-B stage. This shift could originate from the prolonged culture of cells before the flow cytometry analysis, from the character of introduced transgenes and/or loss of *p53* itself. Most importantly, the fact that reprogrammed *p53*$^{-/-}$ MEFs gave rise to mature B-cells serves as a proof of principal that under the right conditions SCL, LMO2, ERG, GATA2 and RUNX1c can instill the entire spectrum of haematopoietic cell fates on developmentally distinct cells.

Figure 5.10 continued

B) BCR rearrangements detection in B220/CD19 double positive reprogrammed *p53*$^{-/-}$ MEFs (Repr), control untransduced MEFs (Un) and spleen cells (Sp). The PCR reaction was designed to detect newly formed sequences after rearrangement between diversity (D) and joining (J) gene segment of heavy immunoglobulin chain. In this respect forward primer was recognizing D segment whereas reverse primes were specific to depicted J segments. Arrows indicate rearrangement bands. DNA size ladder in kb. **C)** Gene expression of B-lymphoid genes in control (Un), B220/CD19 double positive reprogrammed *p53*$^{-/-}$ MEFs and spleen-derived control B cells with respect to *β-actin*. Data presented are mean ± SD of a single experiment performed in triplicate. **D)** Schematic representation of B-cell development and cell surface markers characterising specific B-cell subsets**. E)** Schematic representation of cell surface markers distribution in subsets of B-lymphocytes **(i)** and flow cytometry plots showing expression of these markers on cells derived from BM or day 12 sorted $cKIT^{+}$ cells maintained on OP9 in lymphoid medium **(ii)**. Top panel depicts division into early (red) and late (green) B-cell populations on the basis of differential presentation of CD43. Bottom panel reflects distribution of subsets within early (left) and late (right) B-cells gated above.

5.2.4. Endothelial cells with haemogenic potential can be detected at the early stages of reprogramming

During embryogenesis definitive cKIT$^+$ HSPCs arising in the E10.5 AGM are generated as a result of EHT of specialized endothelial population (Boisset et al., 2010). Having confirmed the presence of multilineage cKIT$^+$ progenitors in the reprogramming setting, I next aimed to establish whether their formation, similar to the *in vivo* counterparts, also involves an endothelial precursor. This question was addressed by monitoring CDH5 staining that would reflect the acquisition of an endothelial immunophenotype in MAFs and MEFs undergoing reprogramming. A transient CDH5$^+$ population was indeed present at the early stages of transition in both tested cultures (Fig 5.11A). The peak of CDH5$^+$ cells could be detected by flow cytometry at different time points spanning from day 5 to day 10 of culture (Fig S3A). Furthermore, immunostaining of day 6 reprogrammed MEFs showed that surface expression of CDH5 coincided with cells undergoing a morphological change (Fig 5.11B). The endothelial nature of CDH5$^+$ cells was further confirmed by their ability to uptake AcLDL (Fig 5.11Ci and ii).

To test the haemogenic potential of the CDH5$^+$cKIT$^-$ population, cells sorted at day 6 of reprogramming were re-aggregated with OP9s (Taoudi et al., 2008) (Fig 5.12A and S3B). This strategy allows for the maturation of potential HE before being re-plated into a semi-solid culture. The obtained data showed that only cells originally positive for the endothelial marker CDH5 give rise to haematopoietic colonies in the CFU-C assay (Fig 5.12B). In addition, if the CDH5$^+$cKIT$^-$ cells were cultured for a week on OP9 stroma, they acquired CD41, cKIT and CD45 markers (Fig 5.12C). These experiments clearly demonstrate the haemogenic nature of CDH5$^+$ cells.

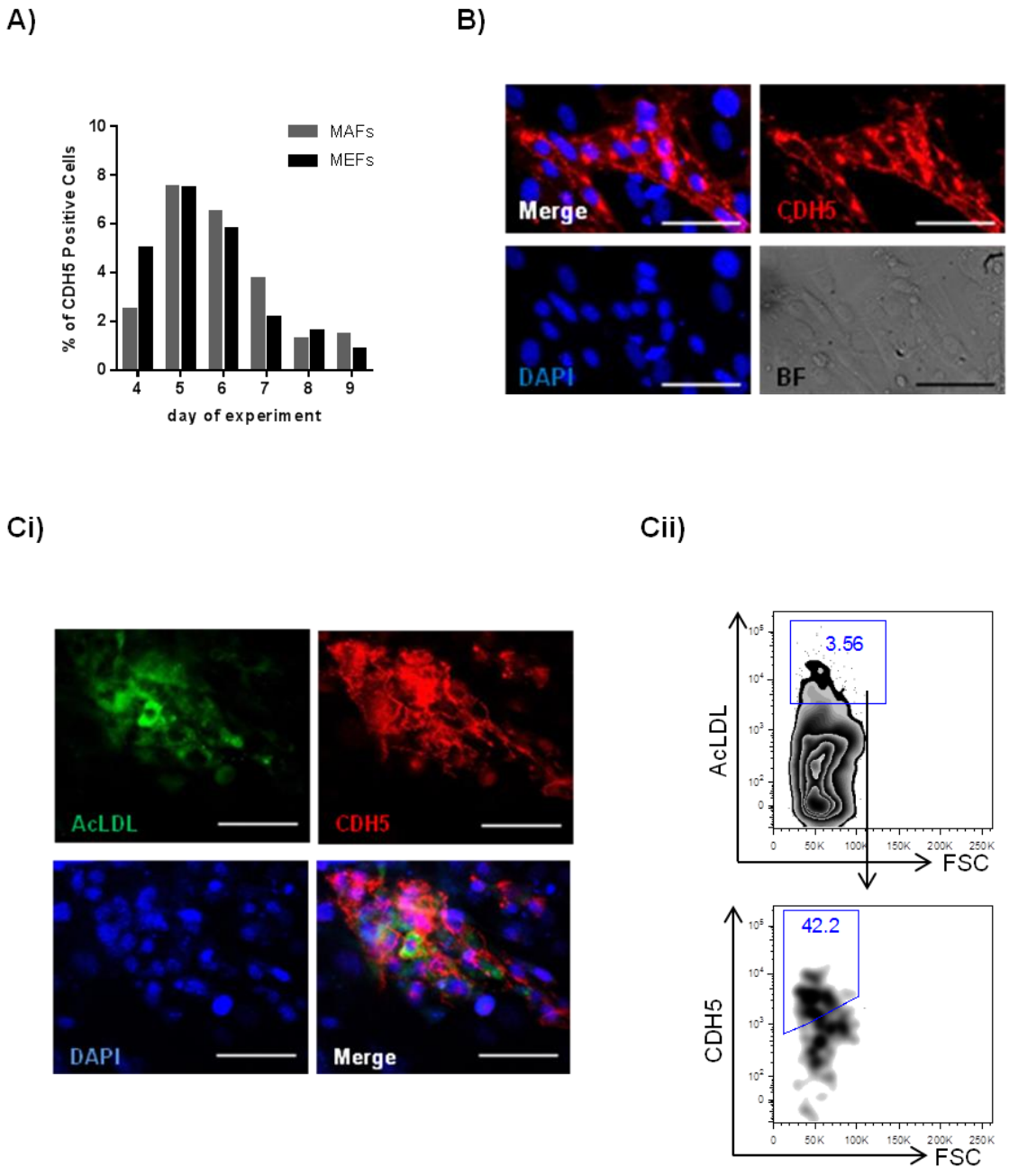


Figure 5.11. Transient endothelial population precedes peak of cKIT$^+$ progenitors.

A) Representative flow cytometry profile of CDH5 cell surface expression between day 4 and 15 of MEFs/MAFs reprogramming. **B)** CDH5 immunostaining of day 6 reprogrammed MEFs **C)** Immunostaining **(i)** and flow cytometry **(ii)** of day 6 reprogrammed MEFs subjected to AcLDL uptake assay. Scale bars represent 50 μm.BF- bright field, FSC – forward scatter.

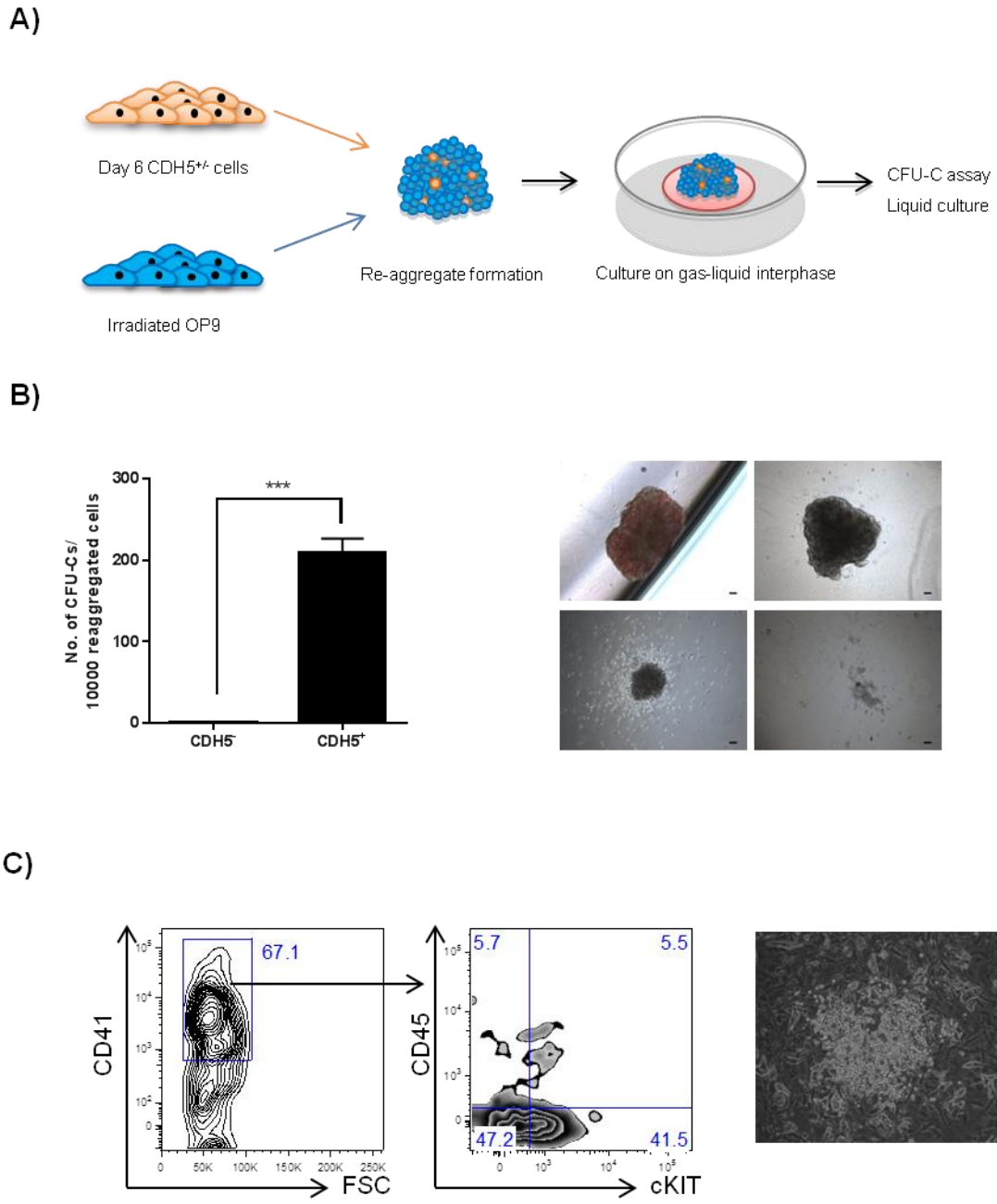


Figure 5.12. CDH5^{+} cells have haemogenic potential.

A) Schematic representation of re-aggregation assay. Cells positive or negative for CDH5 were sorted at day 6 of culture and re-aggregated with 100000 OP9 cells overnight in a hanging drop. Upon 4 days culture in gas-liquid atmosphere aggregates were dissociated to single cell and re-plated into CFU-C assay or liquid cultures. **B)** The number and morphologies of CFUs generated from CDH5^{+}cKIT^{-} or double negative fractions of reprogrammed MEFs (mean ± SEM, n=2) **C)** Representative flow cytometry plots showing acquisition of haematopoietic cell surface markers (CD45 and cKIT) (left) and bright field image of grape-like colonies (right) of blood cells generated from sorted CDH5^{+} cells (n=2). Scale bars represent 50 μm. Asterisks represent significant differences (paired Student t-test, *** $p < 0.0005$)

5.2.5. Reprogrammed cells share transcriptome signature with HSPCs

The above results suggest that transdifferentiation to blood is a stepwise process which, at least in part, proceeds through an endothelial intermediate. In order to characterize the emerging populations on a molecular level, three sets of samples were analyzed by exon array sequencing: untreated MEFs, day 8 sorted CDH5$^+$ cells and day 12 sorted cKIT$^+$ progenitors (Fig 5.13A). Data analysis identified 868 differentially expressed genes. Subsequent principal component analysis (PCA) confirmed the consistency between biological replicates as well as the differences between samples of each three distinct reprogramming stages (Fig S3C). After applying hierarchical clustering, differentially expressed genes segregated into three groups (Fig 5.13B). The first cluster contained genes only expressed in fibroblasts and absent in the remaining two populations. These genes are mostly involved in cell adhesion, extracellular matrix and cytoskeleton organization (Fig 5.13C). The second cluster contained genes of endothelial nature that were solely expressed in the CDH5$^+$ population. Specifically, enriched biological processes regulated by this gene set include angiogenesis and blood vessels formation. Finally, genes gathered in the third cluster were highly expressed in cKIT$^+$ progenitors and moderately in CDH5$^+$ cells. They constitute genes associated with multiple haematopoietic processes of immune response and wound repair. GSEA was also performed on the obtained datasets to establish the similarities between reprogrammed cells and naturally occurring haematopoietic precursors. The analysis revealed that whereas CDH5$^+$ intermediates shared a signature with LT-HSCs, cKIT$^+$ cells resembled the downstream progenitors, consistent with the sequential emergence of both populations in culture (Fig 5.14). Collectively, this transcriptome analysis confirmed the hypothesis that the transition to blood proceeds through an endothelial stage. Moreover, CDH5$^+$ intermediate displays high resemblance to HSCs at the apex of haematopoietic hierarchy.

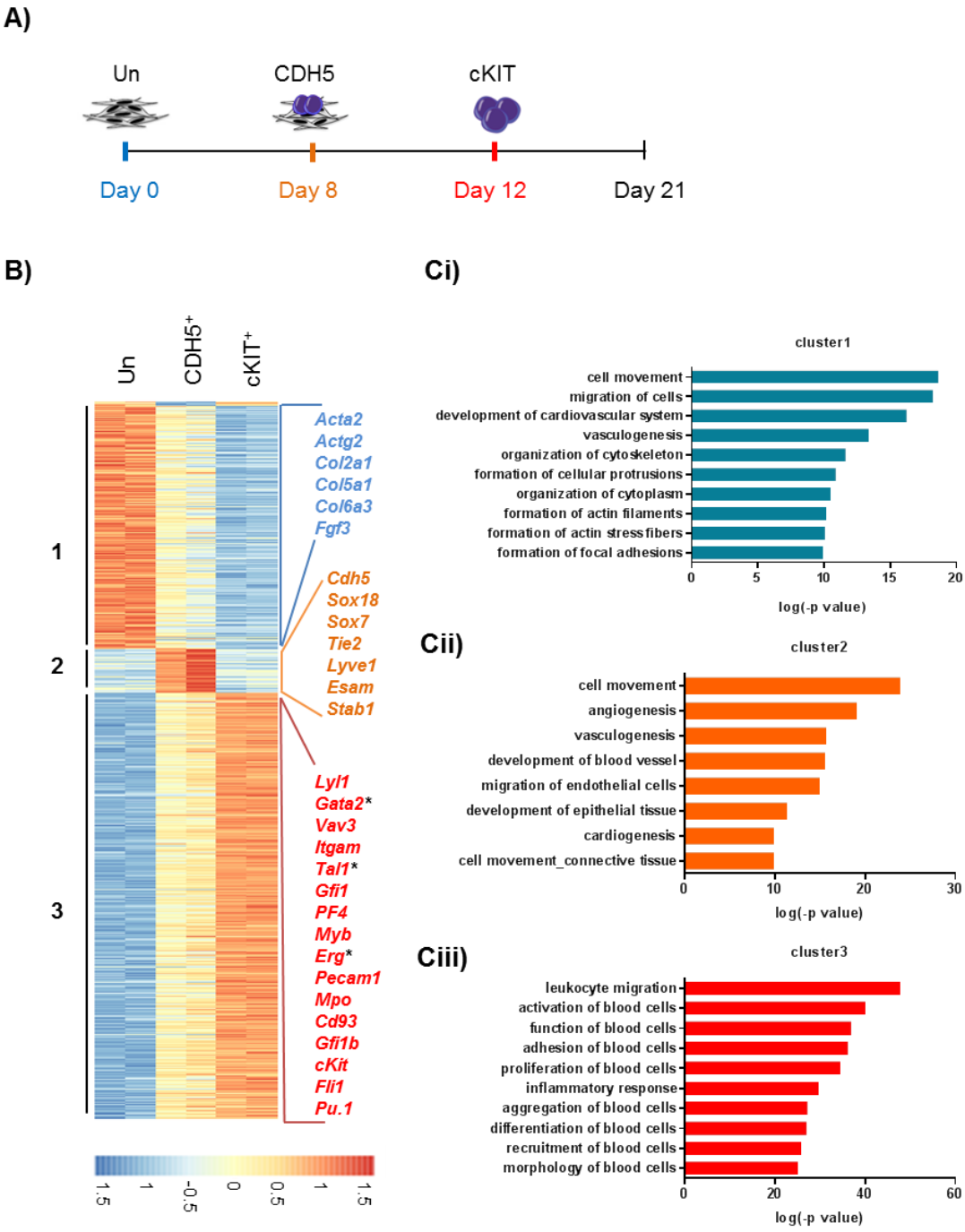


Figure 5.13. Transcriptome analysis reveals intermediate character of CDH5⁺ cells.

A) Sorting timeline for Affymetrix experiment. **B)** Hierarchical clustering of the differentially expressed genes derived from comparison of untreated (Un), day 8 CDH5⁺ and day 12 cKIT⁺ sorted populations. Cluster 1 contains fibroblastic genes downregulated during transition towards cKIT population, cluster 2 gathers endothelial genes upregulated in CDH5⁺ cells and cluster 3 enlists haematopoietic genes highly expressed in day 12 progenitors. * Both endogenous and exogenous gene expression was assessed. **C)** Biological processes affected by differences in expression of genes associated with specific gene clusters. **(i)** Cellular functions affected by downregulation of genes between cKIT and Un population. **(ii)** Processes repressed with the transition from CDH5⁺ to cKIT⁺ population. **(iii)** Biological processes activated by the set of genes upregulated between Un and cKIT populations. Core Analysis was performed with IPA software sourcing from the Ingenuity Knowledge Base.

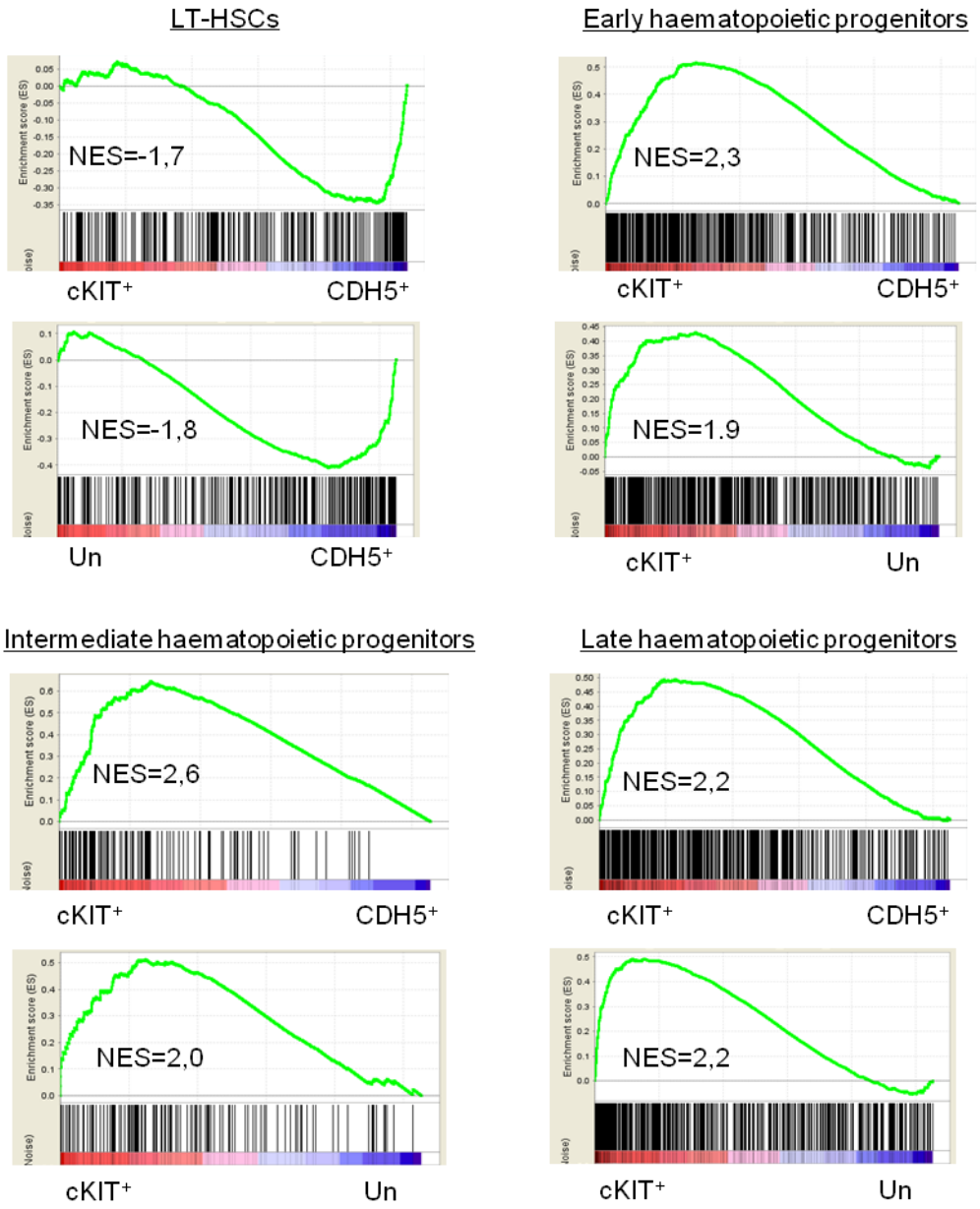


Figure 5.14. GSEA of gene data sets.

GSEA was performed on expression data obtained for untransduced MEFs, CDH5$^+$ and cKIT$^+$ purified populations. Normalised datasets were analysed against previously published gene data sets generated for cells at different stages of haematopoietic development (Ivanova et al., 2002). In this report LT-HSCs were defined as BM-derived Lin$^-$Sca1$^+$cKIT$^+$Rhodamine^{low} cells, early progenitors were sorted as Lin$^-$Sca1$^+$cKIT$^+$Rhodamine^{high} cells, whereas late and intermediate progenitors were enriched as Lin$^-$Sca1$^-$cKIT$^+$. GSEA analysis was performed with default settings and FDR cutoff of 25%.

5.3. Discussion

The interest in transdifferentiation as a strategy to generate cells for regenerative medicine has been fueled by the discovery of induced pluripotency. In contrast to iPSC generation, direct lineage switches are more efficient and rapid. The reprogrammed cells obtained are mostly transgene independent, the upregulation of the target transcriptional programme is visible within hours to days and the transition does not involve re-acquisition of pluripotency (Di Tullio et al., 2011; Vierbuchen et al., 2010; Zhou et al., 2008). These common features of the transdifferentiation process are also observed/recapitulated in my studies. However many questions about the cellular mechanisms governing reprogramming to blood remain unresolved.

Transdifferentiation protocols have included various starting and destination cell types as well as a vast spectrum of lineage-affiliated TFs. This precludes unraveling a uniform mechanism of fate switching. In an *in vivo* model of transdifferentiation hintgut U cell reprograms to Y motor neuron during the development of *Caenorhabditis elegans*. This process is cell cycle independent and is solely manifested by morphological changes and the acquisition of neural markers (Jarriault et al., 2008). Similarly, early studies on artificial transdifferentiation suggested that fate switching does not require cell division (Heinrich et al., 2010; Vierbuchen et al., 2010). Moreover, many conversion protocols reported to date do not involve progenitor stage and the outcome of reprogramming is limited to only one lineage subtype (Ieda et al., 2010; Kim et al., 2011). Only recently few groups obtained conversion to neuronal (Han et al., 2012) or hepatic (Yu et al., 2013) precursors, therefore broadening the potential therapeutic applications of transdifferentiation. Our study, in which intermediate blood progenitors give rise to the whole spectrum of blood lineages, is a further illustration of a transition toward more naïve cellular states with multilineage capabilities.

The time course flow cytometry analysis of cells arising during fate switching identified two distinct populations marked by CDH5 and cKIT respectively. The early endothelial fraction displayed endothelial features and haemogenic potential (Fig 5.11 and 5.12). Furthermore, analysis of the differentially expressed genes suggested a sequential transition from $CDH5^{+}$

precursors to cKIT$^+$ progenitors as manifested by stepwise acquisition of haematopoietic programme between the two populations (Fig 5.13). This was also confirmed by the production of cKIT$^+$ precursors from CDH5$^+$ cells in liquid culture (Fig 5.12C). Notably, Pereira and colleagues reported that the transition to blood induced by GATA2, cFOS, GFI1b and ETV6 also involves an endothelial precursor, marked by SCA1 and PROM1 (Pereira et al., 2013). This transient population is further giving rise to CD45$^+$ cells exhibiting transcriptional similarity to specifying HSCs. In our study, however, HSC profile was attributed already to endothelial CDH5$^+$ population (Fig 5.14) suggesting that the two fractions with haemogenic potential are not at the comparable developmental stages. Considering that SCA1$^+$PROM1$^+$ cells displays low expression levels of *Gata2*, *Erg* and *Lmo2,* it is plausible that this population is present in my cultures before day 6 when the transgenes expression is not yet fully prevalent.

It is noteworthy that the identified EHT process might not be the only route to successful reprogramming. Indeed, during the time-lapse imaging of reprogramming cultures captured between day 3 and 13 of the experiment, two distinct events of fate switching could be observed (Movie S1). The blue arrow marks the apparent direct conversion of adherent fibroblasts to round blood cells. In contrast, the red arrows indicate the transitions involving an intermediate adherent core structures that once formed give rise to floating blood cells. These data suggest that different routes could lead to the same haematopoietic fate, or various blood populations are generated from different preliminary sources. The observed spectrum of transdifferentiation events most likely originates from the transduction of reprogramming factors as single constructs. These conditions generate a spectrum of cells expressing different types and levels of TFs. In this respect, the use of a polycistronic vector would help in unifying mechanism of blood fate induction.

Regardless of the route followed during their generation, cKIT$^+$ cells present at day 12 of culture were clearly identified as multilineage haematopoietic progenitors giving rise to all blood lineages including B-lymphocytes, provided the presence of a permissive *p53* knock-out background (Fig 5.2 to 5.4 and 5.10). P53 is known as ‘the guardian of the genome’ regulating stress response to DNA damage, hypoxia and oncogenes (reviewed by Vousden and Lu, 2002). Not surprisingly, various mutations of *p53* were

recognized in a vast spectrum of human malignancies (Hainaut et al., 1998). Indeed, murine models with impaired P53 activity are more prone to tumour development and represent a particularly high percentage of lymphoma incidents (Jacks et al., 1994). In normal haematopoiesis *p53* is mostly expressed in HSCs to regulate their quiescence and self-renewal (Liu et al., 2009). Notably, even though *p53*$^{-/-}$ HSCs present impaired functionality, their proliferative phenotype allows them to outgrow WT cells in competitive assays (Chen et al., 2008; TeKippe et al., 2003). Nevertheless, as a consequence of this short-term engraftment advantage, the repopulated mice die quickly due to haematopoietic malignancies. The same pattern was probably observed following the injection of *p53*$^{-/-}$ MEF-derived cKIT^{+} cells to immunocompromised recipients (Fig 5.9G).

P53-nullizygous mice are often used as a model for studying lymphoma. The majority of these malignancies originates from the thymus and consist predominantly of immature (CD4^{+}CD8^{+}) T-cells (Donehower et al., 1995; Dudgeon et al., 2014). This may explain why *p53* loss did not facilitate T-cell maturation beyond DP stage obtained from WT MEF-derived cKIT^{+} cells. The generation of terminally differentiated T-cells through reprogramming could be enhanced by mimicking the thymic environment. The most comprehensive system so far developed is based on *in vitro* culture of embryonic thymus lobes. Fetal thymus organ culture (FTOC) supports the full maturation of T-cells including positive and negative selection of TCR repertoire (Ramsdell et al., 2006). Moreover, the addition of thymopentin, which is naturally secreted by epithelial cells of thymic cortex and medulla, enhances the maturation of human ESC-derived T-lymphocytes (Zhu et al., 2015). The use of these more elaborate culture conditions may therefore support further maturation of reprogrammed T-cells toward SP stage.

Interestingly, the generation of B-cells was not obtained from WT reprogrammed blood progenitors but was specifically facilitated by *p53* loss. This effect can be attributed to the deregulation of both self-cycle and apoptosis upon removal of P53. The increased proliferation and survival of cells could have facilitated expansion of rare lymphoid progenitors in culture. Furthermore, two previously reported knock-out mouse models manifesting B-lymphopenia showed association of this phenotype with significant upregulation of *p53*. More

particularly, depletion of *Mysm1* or *Rpl22* led to the activation of a P53-dependent checkpoint which in turn abrogated generation of B-lymphocytes (Belle et al., 2015; Fahl et al., 2015). Consequently, enforced loss of P53 restored B-lymphopoiesis in both models. This observation provides a possible explanation for how *p53* knock-out facilitates the generation of B-cells from reprogrammed MEFs. Moreover, obtained data suggest that this phenomenon can be possibly recapitulated also in the WT reprogrammed progenitors by the use of P53 inhibitor Pfithrin.

Phenotypically reprogrammed *p53*$^{-/-}$ cultures contained more B220^{+}CD43^{-} Pro-B cells than the BM. The *p53*(mΔpro) mouse line, in which only apoptotic activity of p53 is affected, also showed accumulation of Pro-B precursors (Slatter et al., 2010). The sole presence of a late Pro-B cell fraction might in turn be explained by imbalanced D-J and V-DJ joining events as well as the insufficient diversity of V segment repertoire (Green and Jakobovits, 1998). Altogether, the spectrum of P53 activities makes it challenging to pinpoint which aspect of the multi-stage B-lymphopoiesis may be affected by the introduced change of genotype. Nevertheless, performed experiments establish the potential of the five-factor based reprogramming system to generate mature IgD- and IgM- presenting cells from MEFs under permissive conditions.

Despite their very high potency *in vitro,* long-term haematopoietic engraftment with purified cKIT^{+} cells was not achieved *in vivo.* The co-culture of cells with OP9-DL1 stroma (provision of sustained Notch signaling) (Fig 5.7) or Veravec endothelial cells (Butler et al., 2010) (work by Kiran Batta) did not expand the detectable reconstitution beyond 2 weeks, suggesting a requirement for more substantial modifications to the protocol. A recently published study investigating early haematopoietic ontogeny characterized population of definitive EMPs arising in E8.5 YS (McGrath et al., 2015). These cells are marked by cKIT^{+}SCA1^{-}CD16/32^{+} immunophenotype and possess myeloid, megakaryocytic and erythroid but not B-lymphoid *in vitro* potential. When injected to immunocompromised recipients, EMPs supported short-term reconstitution of the erythroid compartment. Notably, the same spectrum of functionalities was also exhibited by the cKIT^{+} reprogrammed cells. Analysis of cell surface markers of OP9-DL1-matured day 21 cultures confirmed the presence of a CD16/32^{+} fraction which was not detectable in the BM control

(Fig 5.6Aii). This finding is in line with the hypothesis that transdifferentiation mimics embryonic development. As early as day 6 of culture, I was able to detect cKIT$^+$ cells not expressing CDH5, even though visible blood colonies are not yet formed at this stage (Fig S3B). Similar to the ESC/EB culture system, in which presence of EMPs was also confirmed, transdifferentiation may therefore recapitulate extra-embryonic definitive haematopoiesis characterized by the generation of cKIT$^+$CD16/32$^+$ EMPs and rare lymphoid progenitors

Notably, the injection of early day 6 CDH5$^+$ cells with a clear haemogenic potential (Fig 5.12) and transcriptional profile similar to HSCs (Fig 5.14) did not result in repopulation of irradiated mice. Further optimization steps in serum-free conditions, aimed at removing from the culture cytokines that might induce differentiation, led to a higher generation of immunophenotypical HSCs albeit with variable efficiency and reproducibility (Fig 5.8). Unfortunately these cells also exhibited lack of engraftment capabilities. A similar serum-free strategy employed by Pearson and colleagues, aiming to meet the long-standing challenge of generating engrafting blood cells from ESC cultures, proved more successful. In this study stepwise provision of carefully selected growth factors led to the generation of a transient cKIT population at the very onset of mesoderm commitment that could provide long-term engraftment (Pearson et al., 2015). As reprogramming to blood proceeds through similar steps as ESC blood specification it may be interesting to investigate and evaluate in more detail very early events of transdifferentiation occurring between day 4 and 6 of experiment. Reoccurring issues in obtaining functional HSCs through transdifferentiation may also originate from residual transgene expression blocking cells from obtaining their full functionality. In this respect implementing an inducible SCL/LMO2-based system could provide tighter control of transgene expression as well as a more substantial involvement of endogenous haematopoietic GRNs in the maintenance of the newly acquired identity. Moreover, adding other TFs (MEIS1, LYL1 or HOXB4), which play crucial roles in HSC biology, may increase the output of long-term engrafting cells.

A very important aspect of the successful generation and maintenance of HSCs are optimal *in vitro* culture conditions. Long-standing efforts to establish conditions supporting undifferentiated HSCs resulted in moderate progress however significant expansion has not been achieved yet. Numerous studies

have illustrated that stromal layers derived from haematopoietic niches like AGM (UG26), FL (AFT024) or BM (MS3) facilitate prolonged cultures of undifferentiated cells with maintained engraftment potential (Bennaceur-Griscelli et al., 2001; Moore et al., 1997; Oostendorp et al., 2002). These favorable conditions can be accompanied by a hypoxic environment, which is naturally occurring in physiological haematopoietic sites. Moreover, the recently identified small molecules StemReginin1, Pleiothropin and pyrimidoindole derivative UM171 (Boitano et al., 2010; Fares et al., 2014; Himburg et al., 2010), that were found to expand $CD34^+$ cord blood cells with repopulation capabilities, could be tested. Finally, the expanding field of bioengineering could provide 3D culture systems that might better recreate the spatial organization of cells in tissues. The synthetic organ mimic has been used to model vascular networks by injecting endothelial cells, pericytes and collagen I into a microfluidic channel (van der Meer et al., 2013). A similar approach could be used for the formation of organ-on-chip equivalent of the DA niche for the generation of iHSCs.

Despite continuous advances in the reprogramming methods, the cells generated in a dish often do not possess full biological functionality. These reoccurring issues suggest that the *in vivo* environment might be better suited to generate cells for regenerative medicine. Riddell and colleagues tested the feasibility of this approach by injecting 36-factor transduced cells shortly after their infection into recipient mice and allowing for *in vivo* selection of haematopoietic reprogramming TFs (Riddell et al., 2014). A similar strategy can be employed for early SCL/LMO2 infected MEFs to facilitate the generation of engrafting cells. Conceptually *in vivo* environment may allow for the maturation of HE similarly to OP9 re-aggregation *in vitro* assay. Another variant of the experiment could include generating an inducible SCL/LMO2 mouse model. This tool would aid robust *in vivo* reprogramming, in depth investigation of mechanisms governing transdifferentiation and easier introduction of improvements to the existing protocol.

Chapter 6

6. GFI1 and GFI1b as markers of successful reprogramming

6.1. Introduction

Mature blood cells are generated via sequential cell fate decisions which begin at the level of HSCs and extend to multi- and oligopotent progenitors. These specific subpopulations of blood precursors can be distinguished by means of multiparametric flow cytometry. HSCs have been isolated on the basis of multiple cell surface markers including LSK $CD150^{+}CD48^{-}$, $CD48^{-}EPCR^{+}CD45^{+}CD150^{+}$ or LSK $CD34^{-}Flt3^{-}$ combinations. However, cells marked by these established immunophenotypes still exhibit substantial heterogeneity in repopulation capabilities (Beerman et al., 2010; Kent et al., 2009; Morita et al., 2010). The skewed lineage outputs and diverse self-renewal capabilities have been attributed to both extrinsic and intrinsic factors. In this line, the recent work performed with the use of *vWF* reporter line allowed identifying and isolating a subfraction of platelet-primed HSCs defined as LSK $CD150^{+}CD48^{-}CD34^{-}vWF^{+}$ (Sanjuan-Pla et al., 2013). This study illustrated that integrated approach employing both immunophenotype and molecular signature of cells can result in the identification of functional subpopulations within previously purified cell fractions.

In the previous chapter, I have characterised the immunophenotypic and functional attributes of two intermediate populations occurring during the early stages of transdifferentiation. The identification of both $CDH5^{+}$ cells with haemogenic potential and $cKIT^{+}$ multilineage progenitors suggest that the transition recapitulates to some extent embryonic haematopoietic development. Despite a clear *in vitro* haematopoietic potential (Fig 5.1, 5.2, 5.9), neither $CDH5^{+}$ nor $cKIT^{+}$ cells displayed long-term repopulation potential. This may be due to the rarity of truly functional bona fide HSPCs arising during reprogramming experiment. Therefore identifying additional molecular signature

could help enriching reprogrammed endothelial and haematopoietic populations in cells with more significant *in vivo* biological activity. Recent studies performed in our laboratory have identified GFI1 and GFI1b as two haematopoietic TFs that differentially mark populations involved in the formation of blood clusters within the AGM (Thambyrajah et al., *submitted*). These highly homologous zinc finger proteins act as transcriptional repressors and have been well characterised as regulators of adult haematopoiesis. *Gfi1* is expressed in mature myeloid and lymphoid cells as well as their respective precursors (Karsunky et al., 2002; Yucel et al., 2004). Mice lacking GFI1 suffer from neutropenia and their HSCs display decreased repopulation capabilities (Hock et al., 2004a; Karsunky et al., 2002). *Gfi1b* exhibits a complementary pattern of expression and is predominantly present in erythroid and megakaryocytic lineages (Vassen et al., 2007). Accordingly, *GFI1b*$^{-/-}$ embryos die at E14.5 due to severe defects in erythropoiesis (Saleque et al., 2002). GFI1 and GFI1b have also been recently identified as direct targets of RUNX1, linking their function with initial stages of EHT (Lancrin et al., 2012). Indeed, expressing either repressor in *Runx1*$^{-/-}$ HE cells prompted downregulation of endothelial programme and acquisition of blood morphology. *In vivo* knock-out studies confirmed that the DA of embryos lacking GFI1 and GFI1b was completely devoid of haematopoietic clusters. Interestingly, blood progenitors were still present in the E10.5 YS suggesting a different mechanism governing blood formation in these two haematopoietic sites (Lancrin et al., 2012)Thambyrajah et al., *submitted*). Subsequent single-cell resolution imaging further deciphered differential roles of GFI1 and GFI1b during early haematopoiesis (Fig 6.1). GFI1 was the only paralogue present at the stage of HE formation and marked cells in the ventral lining of the DA (Thambyrajah et al., *submitted*). GFI1b could be detected only after the formation of haematopoietic clusters and the acquisition of blood cells morphology. Observation that GFIs mark different populations within AGM prompted me to use a *GFI1*/*GFI1b* double reporter system to unravel discreet subsets of cells that could be arising during transdifferentiation.

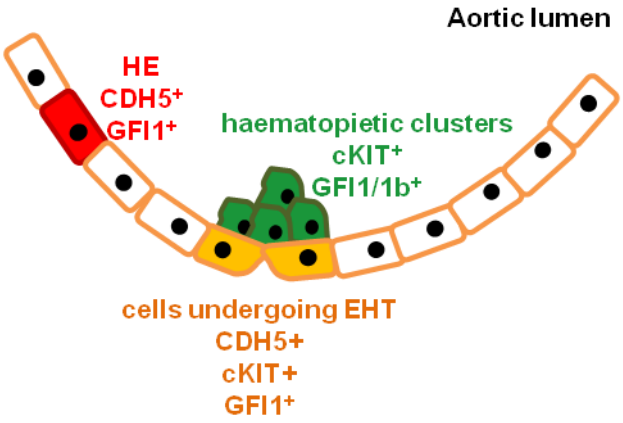


Figure 6.1. Distribution of GFI1 and GFI1b in cells of DA during embryonic haematopoiesis.

In the embryo GFI1 marks specifically endothelial cells in the lining of the DA that are predestined to give rise to definitive HSCs. During EHT GFi1^{+}CDH5^{+} cells start acquiring cKIT and subsequently become part of intra-aortic haematopoietic clusters. Generated in the process HSCs are the only reported cells expressing both *Gfis*.

6.2. Results

6.2.1. *Gfi1* and *Gfi1b* are upregulated during transdifferentiation to blood

To track GFI1 and GFI1b activities and their potential contribution to the transition of fibroblasts to blood, I used an in-house double reporter system. Firstly a MSB mouse line expressing *H2B-dtTomato* under the control of *Gfi1* promoter was generated using engineered ESCs. The cell line was established by recombining *dtTomato* into a *Gfi1* bacterial artificial chromosome followed by transfection into ESCs (work by Monika Stefanska). The mouse line that allows tracking *Gfi1b* expression with GFP reporter (GFB) was obtained from Dr Tarik Moroy (Institut de Recherches Cliniques de Montreal). In order to follow *Gfi1* expression in the context of *Gfi1b* the two mouse lines were crossed to obtain MSBB strain (Fig 6.2A). MEFs were isolated from E14.5 mice with *Gfi1:dtTomato*/*Gfi1b*$^{GFP/WT}$ genotype and used for reprogramming following the standard protocol. After transduction with five TFs, I tracked the emergence of both single and double positive populations, reflecting the acquisition of *Gfi1* and *Gfi1b* expression, in the reprogrammed cells (Fig 6.2B). Importantly, none of the single TF induced the expression of either *Gfi* gene (Fig S4A).

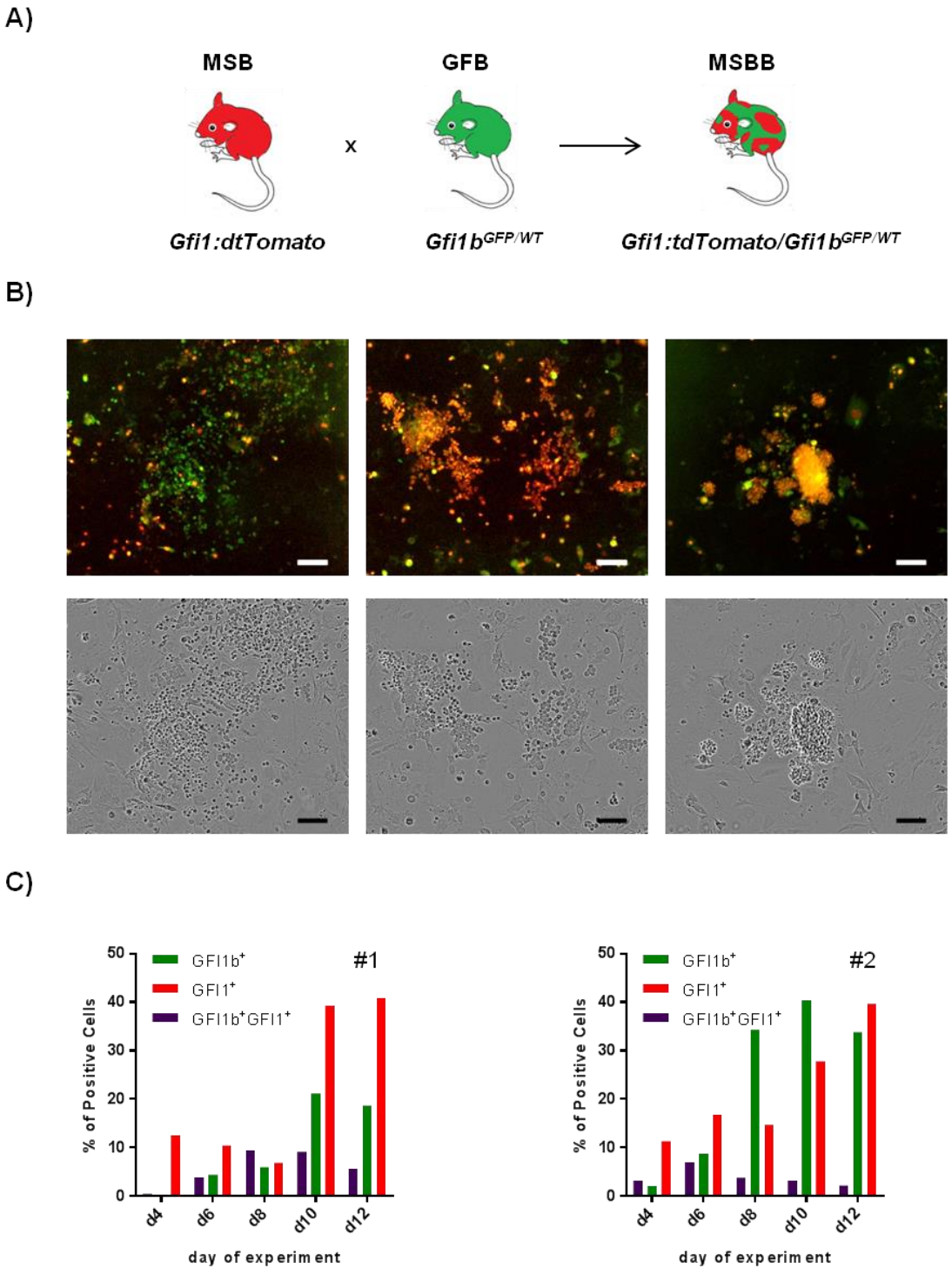


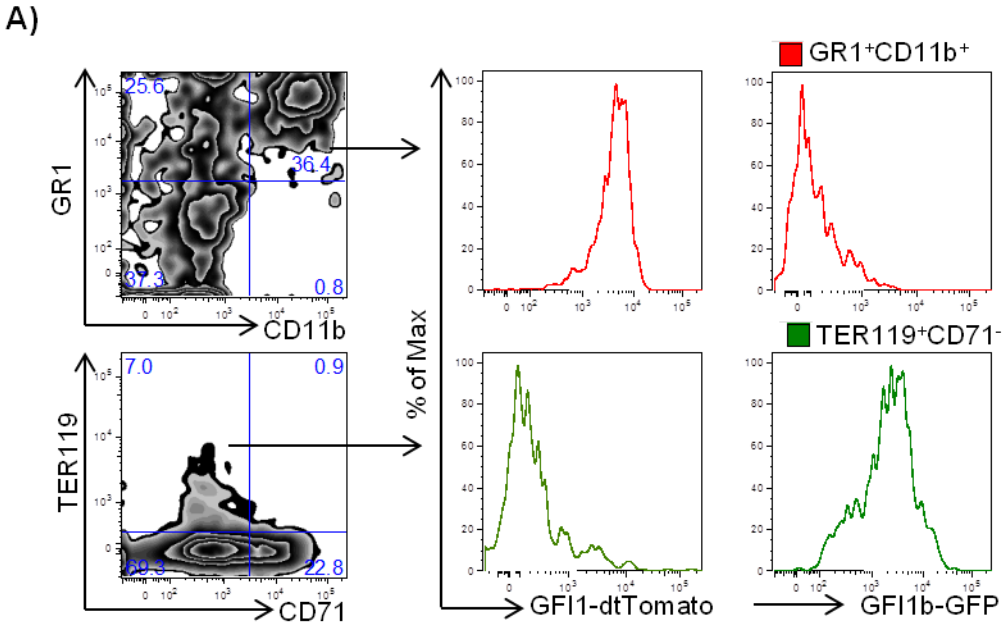
Figure 6.2. *Gfis* are upregulated during transdifferentiation to blood.

A) Schematic representation of the generation of *Gfi1*:*dtTomato*/*Gfi1b*$^{GFP/WT}$ double reporter mouse model. **B)** Bright field and fluorescent images of day 8 to day 15 cultures of reprogrammed MSBB MEFs. **C)** Flow cytometry profile of GFIs acquisition during the initial steps of transdifferentiation. Data presented originate from two biological replicates (n=4). Scale bars represent 50 μm.

Time course flow cytometry analysis of GFP and dtTomato acquisition reflected a steady upregulation of expression of both *Gfis* (Fig 6.2C). I consistently observed that *Gfi1* was the predominantly expressed paralogue at the early stages of reprogramming. This observation was confirmed by time-lapse imaging performed on MSBB MEFs, which showed that only *Gfi1* expression was detected as early as day 3 of the experiment (Movie S2).

6.2.2. Validation of a reporter tracking system

In adult mice GFI1 is present in mature lymphocytes and granulocytes (Karsunky et al., 2002; Yucel et al., 2004), whereas *Gfi1b* is mostly expressed in erythroid and megakaryocytic lineages (Vassen et al., 2007). To investigate whether this pattern is also reflected in differentiated reprogrammed blood cells, I performed flow cytometry analysis of day 21 cultures. Consistently $GR1^{+}CD11b^{+}$ myeloid cells expressed *Gfi1* but not *Gfi1b* (Fig 6.3A top and 6.3B). Conversely, the vast majority of $CD71^{-}TER119^{+}$ erythroblasts were $GFI1b^{+}$ (Fig 6.3A bottom). A small population of $dtTomato^{+}$ cells could be detected in the $TER119^{+}$ fraction, which may have originated from globin autofluorescence (Fig 6.3B). Collectively, these experiments showed that the double reporter system is functional in the reprogramming set-up. Moreover, transdifferentiated cells recapitulate the pattern of *Gfis* expression present in mature blood lineages.



B)

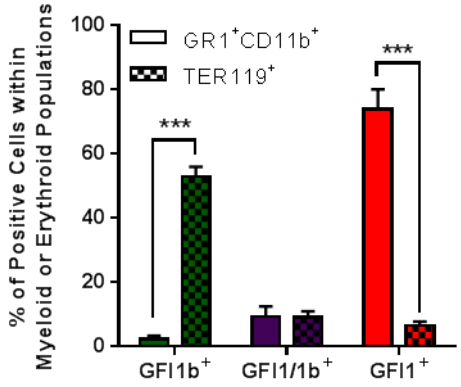


Figure 6.3. GFI1/1b distribution in differentiated cells.

A) Representative flow cytometry plots of day 21 reprogrammed MEFs and GFI1/1b distribution within mature myeloid and erythroid populations. **B)** Statistical significance of GFIs' differential distribution in day 21 reprogrammed GR1/CD11b or TER119 positive cells (mean ± SEM, n=5). Asterisks represent significant differences (paired Student t-test, *** $p < 0.0005$).

6.2.3. GFI1 marks first reprogrammed cells with haematopoietic potential

As mentioned previously, until day 6 of reprogramming *Gfi1* is the predominantly expressed GFI paralogue. It has been recently shown that in the AGM GFI1s mark successive cellular populations and are strictly associated with cells undergoing EHT in the endothelial lining of the DA ($CDH5^+$) or in its close proximity ($CDH5^+cKIT^+$) (Thambyrajah et al., *submitted*) (Fig 6.1). To investigate whether this pattern is also reflected during transdifferentiation, I performed immunostaining and flow cytometry analyses of cultures at the initial stages of reprogramming. These data showed that by day 6 cells undergoing the initial morphological changes presented CDH5, and this was associated with *Gfi1* but not *Gfi1b* expression (Fig 6.4A). Transduced MSB MEFs were also used to determine the percentage of $CDH5^+$ cells co-expressing *Gfi1*. Between day 4 and 8, GFI1 was consistently present in ~10% of $CDH5^+$ cells. Moreover, this subpopulation was not visibly increasing or decreasing throughout the culture. During embryonic development, the first *Gfi1*-expressing cells are also CDH5 positive and they mark the HE formed in the lining of the DA (Thambyrajah et al., *submitted*) (Fig 6.1). I therefore sorted

day 6 cKIT$^-$ cells on the basis of differential CDH5 and *Gfi1* expression to perform re-aggregation assays assessing their haemogenic potential (Fig S4B). Interestingly, the highest number of haematopoietic colonies originated from GFI1 only cells (Fig 6.4C). In contrast, the presence of GFI1 in the CDH5$^+$ cells did not affect the clonogenic capabilities of these populations.

After HE formation in the AGM cells begin to undergo EHT and generate haematopoietic clusters. This process is accompanied by the acquisition of cKIT on the surface of GFI1$^+$CDH5$^+$ HE (Thambyrajah et al., *submitted*) (Fig 6.1). To establish the haematopoietic potential of the reprogrammed equivalent of these cells, I again sorted day 6 cKIT$^+$CDH5$^+$ fibroblasts based on *Gfi1* expression (Fig S4C). Haematopoietic colony formation potential was mostly associated with *Gfi1*-expressing cells both in MEF and MAF-derived cells (Fig 6.4D and E). Altogether, these preliminary data illustrate that *Gfi1* expression specifically marks cells with haematopoietic potential within CDH5$^+$cKIT$^+$ but not CDH5$^+$cKIT$^-$ fraction.

6.2.4. GFIs are associated with increased haematopoietic potential of cKIT progenitors

In vivo, following the initial *Gfi1* only expression in the endothelial cells of the DA, cells start undergoing EHT and generate haematopoietic clusters with concomitant acquisition of *Gfi1b* expression (Thambyrajah et al., *submitted*) (Fig 6.1). Similarly, starting from day 6 GFI1b is steadily acquired by transitioning cells (Fig 6.2C), a pattern similar to the one established previously for cKIT (Fig 5.2A). In the aim of unravelling the role of sequential and differential *Gfi* expression in arising progenitors, I performed series of flow cytometry and clonogenic assays. I have already shown that *Gfi1* and *Gfi1b* expression is compartmentalized in the differentiated reprogrammed cells (Fig 6.3) To establish whether this pattern is present also in upstream progenitors, I sorted day 21 reprogrammed cells into GFI single and double positive/negative fractions and performed CFU-C assay. The clonogenic potential of double negative and GFI1$^+$ precursors was markedly lower than the one attributed to *Gfi1b*-expressing cells (Fig 6.5A). The latter exhibited higher

potency in the generation of erythroid, immature and granulocytic-macrophagic but not macrophagic colonies. This result correlates with the finding that during embryonic development all definitive HSCs in the AGM, as well as haematopoietic precursors originating from the YS, express *Gfi1b* (Thambyrajah et al., *submitted*). Flow cytometry analysis of cells at day 12 revealed that different combinations of GFIs are distributed equally within the cKIT$^+$ fraction (Fig 6.5B; upper panel). The same pattern was still observed at day 21 (Fig 6.5B bottom). This suggests that enrichment of cKIT$^+$ progenitors in *Gfi1b*-expressing cells is not the reason for observed higher clonogenic potential of these fractions. Another possibility is that the frequency of functional cells within cKIT populations is different depending on expressed *Gfis* combination. To test this hypothesis, I sorted day 12 cKIT$^+$ progenitors into four subfractions, according to their *Gfis* expression (Fig S4D), and performed limiting dilution analysis. The results clearly indicated that hematopoietic potential is associated with expression of either *Gfi1* or *Gfib* as double negative cells displayed poor clonogenic potential (Fig 6.5C). Conversely, functional haematopoietic precursors were three times more frequent in fractions expressing *Gfis* without notable bias towards any sorted combination. These data conclusively prove that *Gfis* expression is associated with the acquired functionality of cKIT$^+$ haematopoietic progenitors.

Finally, I aimed to determine whether selection of cells on the basis of the differential expression of *Gfis* could enrich for cells with repopulating capabilities. For this, after amplification on OP9-DL1, I sorted cells into three GFI fractions and injected them into irradiated immunocompromised mice in competitive assay (Fig 6.5D top). Similar to the previous *in vivo* experiments, I could observe a significant level of engraftment in PB at 2 weeks posttransplantation, but this contribution was gradually lost within the following weeks (Fig 6.6D bottom). It is noteworthy that the highest level of donor contribution originated from fractions expressing *Gfi1b* alone or in conjunction with *Gfi1*.

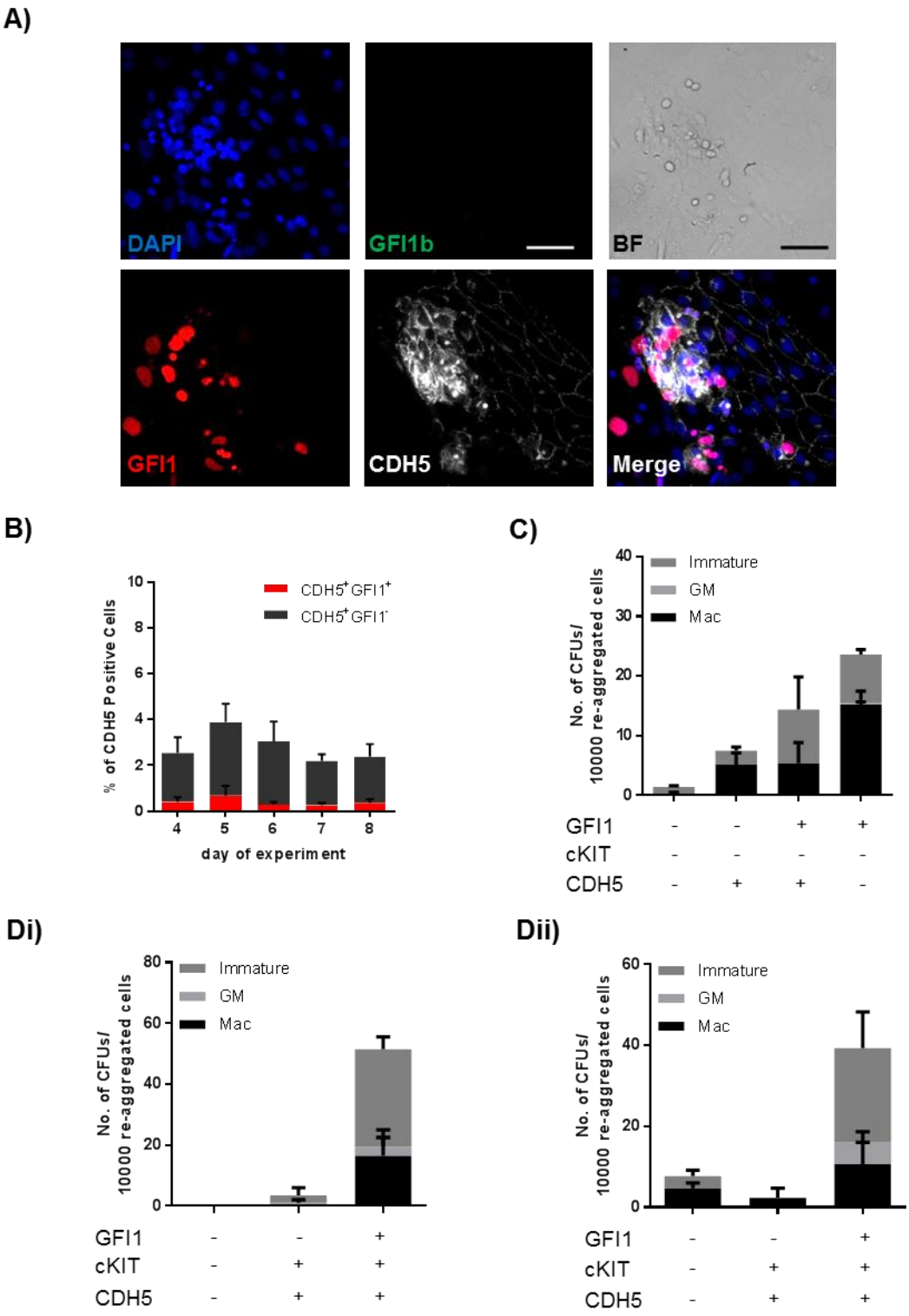
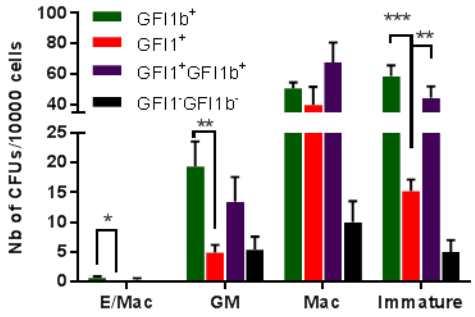


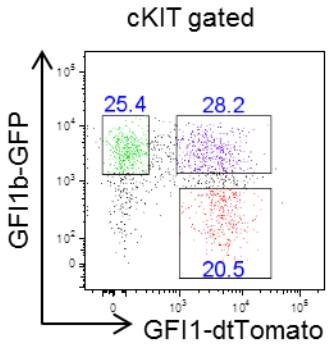
Figure 6.4. GFI1 marks haematopoietic cells at the onset of reprogramming.

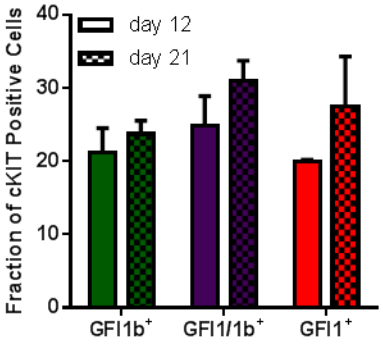
A) Immunostaining of day 6 reprogrammed MSBB MEFs expressing *Gfi1:dtTomato* and *Gfi1b*$^{GFP/WT}$ reporters **B)** Daily pattern of *Gfi1* expression within CDH5^{+} fraction (mean ± SEM, n=3) **C)** The number of CFUs generated by cKIT^{-} cells sorted on day 6 on the basis of differential CDH5 presence and *Gfi1* expression and subjected to re-aggregation assay (mean ± SEM) **D)** The number of CFUs generated by day 6 sorted re-aggregated MEFs **(i)** and MAFs **(ii)** expressing different combinations of CDH5, GFI1 and cKIT (mean ± SEM). BF – bright field, Mac – macrophagic CFU, GM – granulocytic-macrophagic CFU.

A)

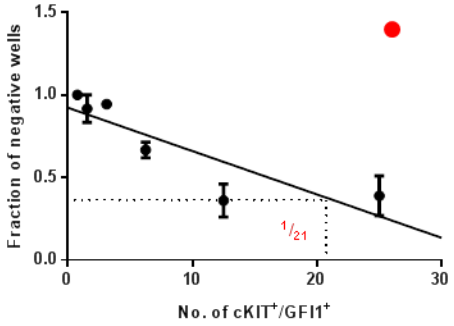

Nb of CFUs/10000 cells
GFI1b+
GFI1+
GFI1+GFI1b+
GFI1-GFI1b-
E/Mac
GM
Mac
Immature

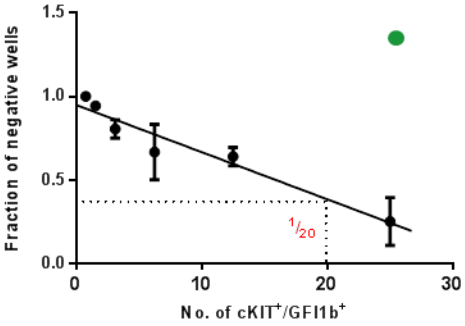
B)

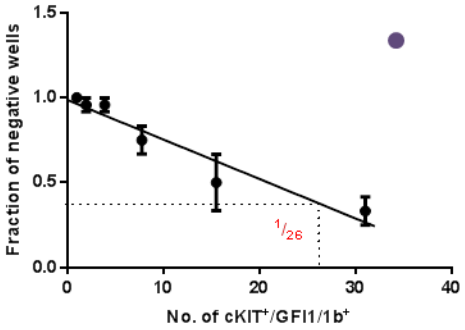

cKIT gated
GFI1b-GFP
GFI1-dtTomato
25.4
28.2
20.5

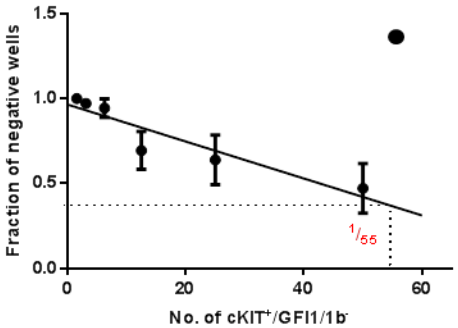

Fraction of cKIT Positive Cells
day 12
day 21
GFI1b+
GFI1/1b+
GFI1+

C)


Fraction of negative wells
1/21
No. of cKIT+/GFI1+


Fraction of negative wells
1/20
No. of cKIT+/GFI1b+


Fraction of negative wells
1/26
No. of cKIT+/GFI1/1b+


Fraction of negative wells
1/55
No. of cKIT+/GFI1/1b-

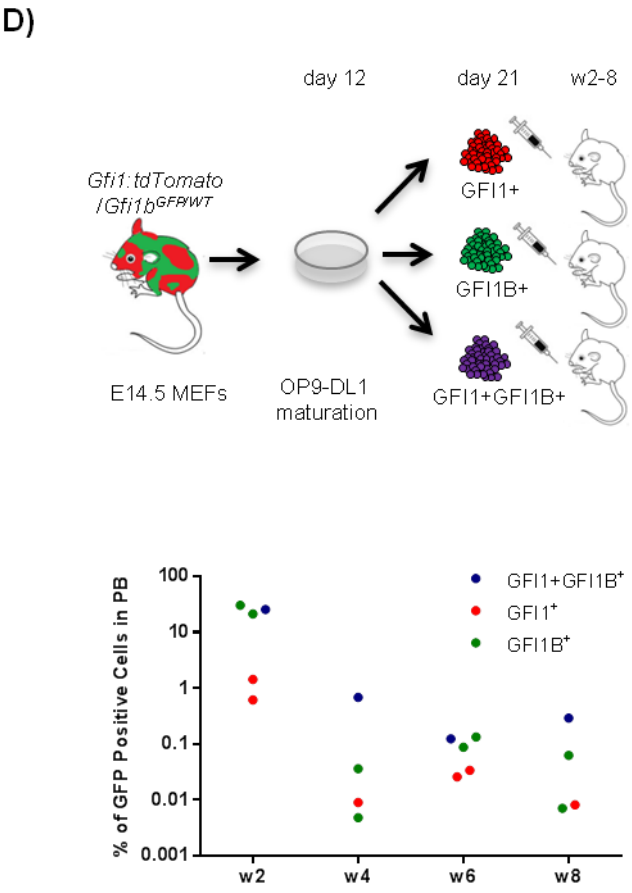


Figure 6.5. Presence of GFIs is associated with functionality of cKIT⁺ progenitors.

A) The number and type of CFUs generated from day 21 sorted GFI1/1b subpopulations (mean ± SEM, n=3) **B)** Flow cytometry plot showing distribution of *Gfi1*- and *Gfi1b*-expressing cells within cKIT⁺ population at day 12 (top) and changes in the pattern acquired by day 21 (bottom) (mean ± SEM, n=2). **C)** LDA of four subfractions of day 12 cKIT⁺ cells (mean ± SEM, n=3 (GFI1⁺, GFI1b⁺, GFI1/1b⁻); n=2 (GFI1/1b⁺). **D)** Schematic representation of transplantation experiment (top) and PB engraftment levels obtained between week 2 and 8 posttransplantation (bottom). Asterisk(s) represent significant differences (Student t-test, * p< 0.05, ** p< 0.005, *** p< 0.0005). Ery – erythroid CFU, E/Mac – mixed erythroid-macrophagic CFU, Mac – macrophagic CFU, GM – granulocytic-macrophagic CFU, MegE – megakaryocytic CFU.

6.3. Discussion

In the previous chapters I was able to show that transdifferentiation of fibroblasts to blood mimics embryonic blood development. The first haematopoietic molecular marker detected on the surface of reprogrammed cells is CD41, which is widely expressed on HSPCs emerging during the primitive and definitive waves of haematopoiesis (Ferkowicz et al., 2003; Robin et al., 2011). Moreover, the presence of a transient CDH5⁺ HE population at the

early stages of reprogramming resembles the EHT during blood development in the DA (Boisset et al., 2010). In this chapter I used a *Gfi1*/*Gfi1b* double reporter system to track the expression of *Gfi1* and *Gfi1b*, two transcriptional repressors marking different cellular populations during the formation of haematopoietic clusters in the AGM. I established that similarly to embryonic haematopoiesis GFI1 is the first paralogue present in reprogrammed cells acquiring blood fate (Movie S2). In the embryo *Gfi1* expression specifically marks the $CDH5^+$ HE in the lining of the DA (Thambyrajah et al., *submitted*). Subsequently these cells start acquiring cKIT and become part of the haematopoietic clusters (Fig 6.1). During reprogramming subpopulation of $CDH5^+$ transitioning cells also expressed *Gfi1* either in conjunction or without accompanying cKIT (Fig 6.4). Interestingly, whereas the presence of GFI1 was highly predictive of haematopoietic potential of $CDH5^+cKIT^+$ cells it was not enriching for haemogenic potential within the $CDH5^+cKIT^-$ fraction. Single-cell isolation and culture experiments would be required to establish whether the $CDH5^+cKIT^-$ cells solely give rise to the cKIT-expressing CDH5 fraction, and whether GFI1 would be the determining factor in this process. Notably, the most robust progenitor potential was associated with *Gfi1* only expressing cells. This finding further supports the hypothesis that, under specific experimental settings, transdifferentiation to blood can be achieved through many independent routes that can result in the generation of different cellular identities.

During the intra-embryonic wave of haematopoiesis *Gfi1b* is expressed within the haematopoietic clusters and together with GFI1 is associated with the HSC identity (Thambyrajah et al., *submitted*). The combined *Gfi1* and *Gfi1b*, but not single, knock-out embryos displayed both lack of blood clusters in the DA and abrogated haematopoietic potential of the AGM. These findings demonstrate the requirement for GFIs for the generation of HSCs and/or their maturation during embryogenesis. In adult mice both GFI1s play crucial roles in HSC biology. GFI1 was found to affect the self-renewal and engraftment capabilities of this cellular fraction (Zeng et al., 2004), whereas the conditional knock-out of GFI1b revealed its crucial role in HSC dormancy (Khandanpour et al., 2010). To some extent reprogramming events mimic the sequential appearance of GFIs during intra-embryonic haematopoiesis. *Gfi1b* is readily expressed around day 6 of reprogramming and continues to be

upregulated in a manner similar to cKIT acquisition (Fig 6.2 and 5.2). At day 21, the *Gfi1b*-expressing fraction exhibits higher haematopoietic potential *in vitro* and more significant engraftment capabilities *in vivo* (Fig 6.5). This finding is in line with the molecular signature of LT-HSCs exhibiting high GFI1b but low GFI1 levels (Wilson et al., 2010a). However, the cKIT$^+$ cells expressing both *Gfis* contribute to 20% of the progenitor population, which significantly exceeds the expected number of immunophenotypical HSPCs. Moreover, whereas lack of GFIs adversely affected functionality of cKIT$^+$ precursors, no bias could be detected in the potency of the purified GFI1/1b single and double positive subpopulations (Fig 6.5). This suggests that even though acquisition of GFI1b can be one of the molecular markers of functional HSPCs generated from fibroblasts, more stringent criteria need to be employed to purify functional long-term repopulating cells.

Altogether, the study of *Gfis* expression pattern provided yet another evidence that transdifferentiation mimics the physiological process of blood formation. Reprogramming cultures reflected the general pattern of sequential GFI1 and GFI1b acquisition and the compartmentalization of GFIs in mature effector cells (Fig 6.2 and 6.3). Moreover, *Gfis* expression was correlated with the functionality of early day 6 CDH5$^+$cKIT$^+$ population and more developed day 12 cKIT$^+$ progenitors. Similar to the *in vivo* pattern, *Gfi1b*-expressing reprogrammed cells exhibited substantial erythroid potential manifested by erythroid-macrophagic and immature colony formation in CFU-C assays (Fig 6.5). Nevertheless, the GFI1$^+$ population did not exhibit an expected extensive differentiation potential toward the myeloid lineage. Additionally, CDH5$^+$GFI1$^-$ reprogrammed cells exhibited haemogenic potential (Fig 6.4), a feature not observed during physiological intra-embryonic blood development (Thambyrajah et al., *submitted*). These discrepancies can be attributed to the rapid dynamics of transdifferentiation. Cells isolated as CDH5$^+$ only at day 6 may have acquired GFI1 within the few next hours of the assay. Using a polycistronic vector or reprogramming with two factors could provide stable and more synchronized environment required for detailed cellular analysis. Moreover, to gain a proper insight into roles of GFI1 and GFI1b in reprogramming an inducible system needs to be implemented. The expression of *Gfis* is regulated by SCL, RUNX1, ERG and GATA2 (Moignard et al., 2013;

Wilson et al., 2010b), therefore continuous overexpression of reprogramming factors can affect the true pattern of GFIs' acquisition. Despite these technical issues it remains clear that GFIs are crucial players during transdifferentiation towards blood. Both parologues were previously selected for reprogramming of murine MEFs or human endothelial cells (Pereira et al., 2013; Sandler et al., 2014). I also found that GFI1b in cooperation with SCL was able to induce low number of reprogramming events (Fig 4.7B). Altogether, these preliminary data highlight the potential use of GFIs as a read-out during fate switching that may provide invaluable insight into cellular mechanisms governing the reprogramming process. Double reporter model could also serve as a new tool for screening of compounds driving or enhancing transdifferentiation to blood.

Chapter 7

7. General conclusions and future work

BM transplant is the most well-established cell replacement therapy dating back to the 1950s. It has been widely used for the treatment of a variety of hematologic disorders including sickle cell anaemia, leukaemia and lymphoma. However to this day procedures are accompanied by incidents of morbidity and mortality due to infections and graft-versus-host disease. Moreover, the insufficient number of HLA-matching donors precludes the use of allogeneic transplantations in therapy of all patients in need of treatment. The big challenge has therefore been to find alternative ways in which the basic functional unit of transplant – HSC – can be generated and expanded. For years different approaches have been employed, however, with limited success. Only recently the first engrafting multilineage cells have been derived from ESCs/iPSCs (Amabile et al., 2013; Pearson et al., 2015; Suzuki et al., 2013b). The therapeutic application of these methods is however limited due to the risk of teratoma formation by incompletely reprogrammed iPSCs and ethical concerns associated with human ESC derivation. An appealing alternative to obtain iHSCs is transdifferentiation of somatic cells. In some instances, depending on the character and relation of used lineages, direct conversion was reported to proceed with high efficiency and rapidity. Moreover, potential safety advantage of this approach originates from the lack of pluripotent intermediates. Recently a remarkable progress has been made by Sandler and colleagues who generated functional iMPPs from human endothelial cell (Sandler et al., 2014). Our group shares an interest in the objective of obtaining iHSCs for therapeutic use by focusing on reprogramming of easily accessible skin fibroblasts.

In this project, my goal was to investigate the possibility of generating blood cells from murine fibroblasts of both adult and embryonic origin. In the course of selection experiments (Fig 3.4), I identified a set of five TFs, known blood fate determinants, capable of reprogramming differentiated cells to blood. Transduction of MAFs and MEFs with SCL, LMO2, GATA2, ERG and RUNX1c led to the transition towards erythroid, megakaryocytic, macrophagic and granulocytic lineages (Fig 4.1 and 4.2). Even though the lineage outcome was the same for both sources of fibroblasts, I noted that mature somatic cells are more difficult to reprogram. MAFs required all five TFs to facilitate the transition (Fig 4.7), which was consistently less efficient than for MEFs (Fig 4.2). These observations are in agreement with data suggesting that less differentiated cells are easier to reprogram. For instance, neuronal stem cells are converted to iPSCs 50-fold more efficiently than fibroblasts (Kim et al., 2008). This phenomenon has been attributed to more plastic state of undifferentiated cells originating from hypomethylation of promoters and repression of lineage-specific genes by Polycomb complex. Notably, MAF-derived blood precursors also displayed lower commitment to the erythroid lineage, which could be due to a lesser degree of transgene silencing (Fig 4.2 and 4.6). Variable biological properties of reprogrammed cells, closely associated with variability in the developmental stage of the cell of origin, were also noted in case of induced pluripotency. More particularly, tail tip fibroblast-derived iPSCs generated secondary neurospheres with higher teratoma-formation propensity than their MEF-derived counterparts (Miura et al., 2009).

Despite the above noted differences, fate switching induced by the five TFs in MAFs and MEFs is proceeding through similar $CDH5^{+}$ endothelial (Fig 5.11) and $cKIT^{+}$ blood progenitor (Fig 5.1) intermediates. As established for MEFs, the functional effector blood cells present within day 21 cultures (Fig 4.3) are majorly a result of commitment of multilineage precursors (Fig 5.2). The latter exhibited robust clonogenic potential *in vitro* (Fig 5.1) and short-term engraftment *in vivo* (Fig 5.7). The immunophenotypical analysis revealed that the haematopoietic marker present specifically on transitioning cells was CD41 (Fig 4.4) marking also initial events of embryonic haematopoiesis (Boisset et al., 2013; Ferkowicz et al., 2003; Robin et al., 2011). This suggested that direct reprogramming to blood may recapitulate the naturally occurring process of

mesoderm specification and EHT. Presence of BRY was not detected in early MEFs cultures suggesting that re-specification to mesoderm does not take place (work by Kiran Batta). However, as early as day 6, a transient population of $CDH5^+$ cells capable of AcLDL uptake was detected and could be isolated (Fig 5.11). Importantly, these $CDH5^+cKIT^-$ cells gave rise to $cKIT^+$ progenitors in liquid culture, and haematopoietic colonies in CFU-C assay (Fig 5.12). Altogether, the confirmed events of EHT as well as noise levels of *βH1* globin expression in day 21 cultures (Fig 4.1 and 4.9) would propose transdifferentiation to definitive blood cells. Furthermore, several lines of evidence imply that in particular the YS branch of haematopoiesis could be mimicked during reprogramming. My *in vivo* studies revealed solely short-term engrafting capabilities of cKIT sorted progenitors (Fig 5.7). Moreover, the repopulation was limited to the erythroid compartment, a feature recently associated with biology of definitive EMPs originating from the YS (McGrath et al., 2015). Indeed, a substantial population of immunophenotypical $SCA1^-$ $cKIT^+CD16/32^+$ EMPs, not present in the control BM, could be detected in reprogrammed cultures (Fig 5.6). In contrast, the presence of immunophenotypical HSCs (Fig 5.8), which normally emerge within the AGM region, highlights similarities between transdifferentiation and the process of intra-embryonic blood development. Additionally, the engraftment potential of reprogrammed cells was obtained only upon OP9-DL1 maturation step. Previously published data indicate that Notch1 is essential for the emergence of haematopoietic potential within the AGM (Kumano et al., 2003), however it is dispensable for the occurrence of EMPs both *in vitro* and *in vivo* (Hadland et al., 2004). It is therefore inconclusive which, if any, stage, of embryonic blood specification may be reflected during induced transdifferentiation. One way to establish similarity to the *in vivo* process would be through further analyses of $CDH5^+$ HE population. It has been shown that endothelial cells giving rise to both EMPs and HSCs share a requirement for *Runx1* expression (Chen et al., 2009; Yokomizo et al., 2001). Chen and colleagues showed that whereas *Ly6a*-mediated expression of CBFβ, a crucial RUNX1 binding partner, rescues HSC formation, it is however not sufficient to support development of EMPs (Chen et al., 2011). Therefore, the generation of these two types of blood precursors can

be uncoupled by tracking their spatial/temporal requirements of CBFβ. It is also plausible that cells presenting YS- or AGM-derived functionality could be concurrently present in the culture and only extrinsic factors dictate which population is maintained and therefore prevalent in functional assays. Recent work by Pearson and colleagues, who for the first time isolated multilineage engrafting cells from differentiating ESC culture, revealed that careful changes in culture conditions can unravel new repopulation potential of what was assumed to be a constrained culture system (Pearson et al., 2015).

Similar optimizations may be required to increase the biological relevance of transdifferentiated cells. For this a simplified but also sustainable reprogramming system needs to be implemented. A possible approach may be transduction with polycistronic vector, which have been shown to result in increased efficiency and/or functionality of reprogrammed cells during re-specification to iHSCs (Riddell et al., 2014) or direct conversion to cardiomyocytes (Inagawa et al., 2012; Wang et al., 2015). Importantly, the latter study revealed that, similar to induced pluripotency, the ratio of reprogramming TFs is of critical importance for successful transdifferentiation. More particularly, only the polycistronic construct M-G-T, delivering non equimolar levels of TFs (MEF2c>GATA4≥TBX5), yielded higher efficiency of conversion to cardiomyocytes. On the basis of N-1 experiments performed during my study (Fig 4.7) the variations of S-L-E-G combination could be tested for increased efficiency of transition to blood.

The finding that forced expression of only two factors evokes reprogramming of MEFs to blood (Fig 4.9) opens new avenues for more in-depth studies of transdifferentiation to haematopoietic lineage. This data suggests that using a simple 2-TF based approach could provide a stable and reliable tool for investigating the molecular mechanisms governing the transition, and as such allow implementation of further improvements to the existing protocol. ChIP-seq analysis can be used to determine genome-wide landscape of SCL and LMO2 binding sites in early time point transduced MEFs. Alternative method for locating accessible target promoters and enhancers of a single TF is DNA adenine methyltransferase (Dam) identification (DamID). In this approach SCL and LMO2 would be fused to the bacterial enzyme Dam which specifically methylates adenines at GATC sequences in the vicinity of the

binding sites of the fusion protein, allowing to identify binding signature of reprogramming factors in fibroblasts. Comparison of Chip/DamID data obtained from purified HSPCs and day 4 transduced MEFs, or day 12 cKIT$^+$ population may unravel haematopoietic targets not accessible for SCL/LMO2 during transdifferentiation of fibroblasts. These key regulators, differentially activated in reprogrammed cells and HSCPs, can be then introduced to reprogramming protocol in a form of lentiviral vectors. Similarly, regulators crucial for complete conversion to blood can be predicted on the basis of accessibility of chromatin. DNA fragments stripped of nucleosomes, and therefore most likely containing expressed genes, can be digested with DNase I, identified by deep sequencing and compared in between starting and target cell type as well as reprogramming intermediates.

Finally, valuable information on the completeness and coherency of transition can be obtained by following complex epigenetic changes during reprogramming. It is well known that specific histone marks are associated with active or silenced chromatin. In this line trimethylation of histone 3 on lysine 4 (H3K4me3) or acethylation of lysine 27 (H3K27ac) mark active promoters. By contrast, trimethylation of lysine 9 or 27 (H3K9 or K3K27me3) typically labels transcriptionally silent heterochromatic regions. Upon overexpression of SCL/LMO2, or full five-TF mix, major fibroblastic loci (like *Actn2* or *Fbn1*) can be surveyed for the appearance of repressive chromatin marks using ChIP. Comparing H3K9me3 or H3K27me3 enrichment in reprogrammed cells and starting fibroblasts would provide information on the completeness of cell conversion. Moreover, similar comparison of activation marks on selected haematopoietic loci (eg. *Meis1*, *Msl2, HoxA9*) between reprogrammed population and purified HSPCs could be performed to identity crucial target genes that are not activated upon transduction but are characteristic for establishment of HSC fate. These missing regulators can be further used for *in vivo* screening of potent transcriptional combinations leading to the generation of cells with engraftment potential.

Biomedical applications focus primarily on the safety aspects associated with reprogrammed cells. Since some of the factors used to reprogram fibroblasts to blood have been associated with oncogenic transformations (LMO2, SCL, RUNX1), their use in clinics might not be desirable. Therefore, screening for small molecules that could replace reprogramming TFs would be of great benefit for the prospective clinical translation of these findings. The CRUK Manchester Institute provides access to a library of 1200 molecules approved by the Food and Drug Administration (FDA) that can be used for high-throughput screening of active compounds. To evaluate them, I isolated E14.5 MEFs by crossing *iScl* mouse line carrying a dox inducible *Scl* cassette with the MSB line carrying the *Gfi1*:*dtTomato* reporter (Fig 7.1). The cells were then seeded on 96 well plates representing the complete library (compounds at 10 μM), and *Scl* expression was induced by dox. The cells were screened at day 21 for induction of *dtTomato* expression or change in morphology. A positive result would suggest that the compound is able to replace LMO2 during transdifferentiation. This large project was repeated twice, each time yielding no positive hits. This negative outcome may be due to the fact that the combined activity of few molecules is required to replace the cocktail of TFs, as was shown for the chemical induction of pluripotency (Hou et al., 2013). Moreover, the half-life of the compounds in culture (usually 24 to 48 hours) may have been insufficient to induce stable changes required for fate switching. All these issues can be addressed in the future experiments, which were not however feasible during the time frame of my PhD. A different approach, implementing an inducible SCL:LMO2 cassette, could be also used to screen for compounds that would increase the efficiency of reprogramming. Recently Li and colleagues employed the similar concept to identify chemical compounds capable of driving transdifferentiation from fibroblasts to a neuronal fate (Li et al., 2015). The low efficiency transition evoked by ASCL1 alone was used as a baseline for the screening. That allowed identifying four molecules that had a moderate effect on the yield of transdifferentiation. More importantly, it was later revealed that a combination of ISX9, I-BET191, Forskolin and CHIR99021 not only enhance efficiency of reprogramming by 10 fold but also induce the transition in the absence of ASCL1.

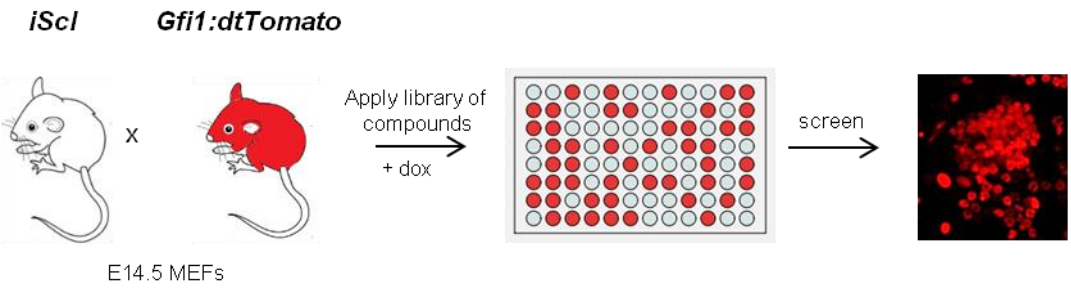


Figure 7.1. FDA library screening strategy.

The major criterion for the potential applicability of the reprogramming protocol lies in the functionality of the generated cells. Even though progenitors with sustained engraftment potential were not yet derived by means of our protocol, we have achieved a stage at which the reprogrammed blood progenitors exhibit substantial differentiation potential towards macrophages and granulocytes. Functionality of these terminally differentiated cells can be assessed in mouse models of hereditary diseases affecting maturation and/or activity of specific blood effectors. Pulmonary alveolar proteinosis (PAP) is caused by mutations in genes encoding GM-CSF receptors α/β and is affecting functions of pulmonary macrophages in a way that they are incapable of clearing surfactant from alveoli. Currently there is no pharmacological treatment for this condition, and BM transplants are ineffective due to the infection-related fatality. It has been recently shown, however, that direct pulmonary transplantation of BM-derived F4/80$^+$ macrophages alleviates symptoms of disease with a sustainable effect of at least one year (Suzuki et al., 2014). In the nearest future, in collaboration with Prof Thomas Morris from Hannover Medical School, we aim to implement the PAP mouse model to assess the biological functionality of reprogrammed macrophages. Similarly, we can test the ability of the transdifferentiated neutrophils to cure X-linked chronic granulomatous disease by transplants into the aberrant environment of the mouse model.

Nonetheless, before *in vivo* assays can be performed the exact identity of transplanted cells needs to be confirmed. The previously performed population-scale morphological and immunofluorescence comparisons with native counterparts should be further validated with more comprehensive single-cell transcriptome, proteome and epigenome studies. Additionally,

for *in vivo* functional characterisation, cells will need to be expanded in large numbers. Recently two groups working on reprogramming towards human hepatocytes developed distinct strategies to meet this requirement. Huang and colleagues implemented the co-transduction of a SV40 large T antigen, which restored proliferation capacity of target cells (Huang et al., 2014). Simultaneously the Deng group used *p53* small interfering RNA to facilitate expansion of induced hepatocytes (Du et al., 2014). This is in accordance with my study, in which reprogramming of *p53* knock-out MEFs proceeded with significantly increased efficiency and kinetics (Fig 5.9). Therefore transient loss of P53 activity could be used to obtain higher number of blood cells for transplantation experiments.

Finally, findings of this study need to be translated towards human system. Based on previous reports on reprogramming to neurons, the same mix of TFs can be applied for both mouse and human cells (Caiazzo et al., 2011; Son et al., 2011). However, our preliminary data suggest that transduction of human dermal fibroblasts with the orthologs of SCL, LMO2, GATA2, ERG and RUNX1 is not sufficient to obtain robust fate conversion (work by Kiran Batta). Therefore, it will be necessary to perform additional screens with our initial set of regulators (Fig 3.2) as well as TFs identified so far by other groups (Fig 3.5). Once the transition is obtained, the successful protocol will require multiple optimization steps in order to ensure an acceptable clinical grade of the delivered cells. That includes development of chemical reprogramming, as was done for transdifferentiation of fibroblasts to neurons (Hu et al., 2015), conducted in defined serum-free and feeder-free conditions. Ultimately, the reprogrammed cells will need be tested in the state-of-art xenograft models to collect data on their possible functionality and safety in patients. This long route to obtaining therapeutic iHSCs becomes even more challenging due to our limited knowledge of HSC biology, but also the lack of fully established tools to expand, maintain, test and control the quality of generated cells. Despite these challenges the promise of major advances in patient care is a strong drive to further explore this emerging field of transdifferentiation, and in particular toward blood.

References

Aasen, T., Raya, A., Barrero, M.J., Garreta, E., Consiglio, A., Gonzalez, F., Vassena, R., Bilic, J., Pekarik, V., Tiscornia, G.*, et al.* (2008). Efficient and rapid generation of induced pluripotent stem cells from human keratinocytes. Nat Biotech *26*, 1276-1284.

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

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

Alvarez-Dolado, M., Pardal, R., Garcia-Verdugo, J.M., Fike, J.R., Lee, H.O., Pfeffer, K., Lois, C., Morrison, S.J., and Alvarez-Buylla, A. (2003). Fusion of bone-marrow-derived cells with Purkinje neurons, cardiomyocytes and hepatocytes. Nature *425*, 968-973.

Amabile, G., Welner, R.S., Nombela-Arrieta, C., D'Alise, A.M., Di Ruscio, A., Ebralidze, A.K., Kraytsberg, Y., Ye, M., Kocher, O., Neuberg, D.S.*, et al.* (2013). In vivo generation of transplantable human hematopoietic cells from induced pluripotent stem cells. Blood *121*, 1255-1264.

Amsellem, S., Pflumio, F., Bardinet, D., Izac, B., Charneau, P., Romeo, P.H., Dubart-Kupperschmitt, A., and Fichelson, S. (2003). Ex vivo expansion of human hematopoietic stem cells by direct delivery of the HOXB4 homeoprotein. Nature medicine *9*, 1423-1427.

Antonchuk, J., Sauvageau, G., and Humphries, R.K. (2002). HOXB4-induced expansion of adult hematopoietic stem cells ex vivo. Cell *109*, 39-45.

Arinobu, Y., Mizuno, S., Chong, Y., Shigematsu, H., Iino, T., Iwasaki, H., Graf, T., Mayfield, R., Chan, S., Kastner, P.*, et al.* (2007). Reciprocal activation of GATA-1 and PU.1 marks initial specification of hematopoietic stem cells into myeloerythroid and myelolymphoid lineages. Cell stem cell *1*, 416-427.

Barton, K., Muthusamy, N., Fischer, C., Ting, C.-N., Walunas, T.L., Lanier, L.L., and Leiden, J.M. (1998). The Ets-1 Transcription Factor Is Required for the Development of Natural Killer Cells in Mice. Immunity *9*, 555-563.

Batta, K., Florkowska, M., Kouskoff, V., and Lacaud, G. (2014). Direct reprogramming of murine fibroblasts to hematopoietic progenitor cells. Cell reports *9*, 1871-1884.

Beck, D., Thoms, J.A., Perera, D., Schutte, J., Unnikrishnan, A., Knezevic, K., Kinston, S.J., Wilson, N.K., O'Brien, T.A., Gottgens, B.*, et al.* (2013). Genome-wide analysis of transcriptional regulators in human HSPCs reveals a densely interconnected network of coding and noncoding genes. Blood *122*, e12-22.

Beerman, I., Bhattacharya, D., Zandi, S., Sigvardsson, M., Weissman, I.L., Bryder, D., and Rossi, D.J. (2010). Functionally distinct hematopoietic stem cells modulate hematopoietic lineage potential during aging by a mechanism of clonal expansion. Proceedings of the National Academy of Sciences of the United States of America *107*, 5465-5470.

Belle, J.I., Langlais, D., Petrov, J.C., Pardo, M., Jones, R.G., Gros, P., and Nijnik, A. (2015). p53 mediates loss of hematopoietic stem cell function and lymphopenia in Mysm1 deficiency. Blood *125*, 2344-2348.

Bennaceur-Griscelli, A., Pondarre, C., Schiavon, V., Vainchenker, W., and Coulombel, L. (2001). Stromal cells retard the differentiation of CD34(+)CD38(low/neg) human primitive progenitors exposed to cytokines independent of their mitotic history. Blood *97*, 435-441.

Benveniste, P., Frelin, C., Janmohamed, S., Barbara, M., Herrington, R., Hyam, D., and Iscove, N.N. (2010). Intermediate-term hematopoietic stem cells with extended but time-limited reconstitution potential. Cell stem cell *6*, 48-58.

Benveniste, P., Serra, P., Dervovic, D., Herer, E., Knowles, G., Mohtashami, M., and Zuniga-Pflucker, J.C. (2014). Notch signals are required for in vitro but not in vivo maintenance of human hematopoietic stem cells and delay the appearance of multipotent progenitors. Blood *123*, 1167-1177.

Boisset, J.C., Clapes, T., Van Der Linden, R., Dzierzak, E., and Robin, C. (2013). Integrin alphaIIb (CD41) plays a role in the maintenance of hematopoietic stem cell activity in the mouse embryonic aorta. Biology open *2*, 525-532.

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

Boitano, A.E., Wang, J., Romeo, R., Bouchez, L.C., Parker, A.E., Sutton, S.E., Walker, J.R., Flaveny, C.A., Perdew, G.H., Denison, M.S.*, et al.* (2010). Aryl hydrocarbon receptor antagonists promote the expansion of human hematopoietic stem cells. Science *329*, 1345-1348.

Bories, J.-C., Willerford, D.M., Grevin, D., Davidson, L., Camus, A., Martin, P., Stehelin, D., and Alt, F.W. (1995). Increased T-cell apoptosis and terminal B-cell differentiation induced by inactivation of the Ets-1 proto-oncogene. Nature *377*, 635-638.

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

Brun, A.C., Bjornsson, J.M., Magnusson, M., Larsson, N., Leveen, P., Ehinger, M., Nilsson, E., and Karlsson, S. (2004). Hoxb4-deficient mice undergo normal hematopoietic development but exhibit a mild proliferation defect in hematopoietic stem cells. Blood *103*, 4126-4133.

Buganim, Y., Faddah, D.A., and Jaenisch, R. (2013). Mechanisms and models of somatic cell reprogramming. Nature reviews Genetics *14*, 427-439.

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

Bussmann, L.H., Schubert, A., Vu Manh, T.P., De Andres, L., Desbordes, S.C., Parra, M., Zimmermann, T., Rapino, F., Rodriguez-Ubreva, J., Ballestar, E.*, et al.* (2009). A robust and highly efficient immune cell reprogramming system. Cell stem cell *5*, 554-566.

Butler, J.M., Nolan, D.J., Vertes, E.L., Varnum-Finney, B., Kobayashi, H., Hooper, A.T., Seandel, M., Shido, K., White, I.A., Kobayashi, M.*, et al.* (2010). Endothelial Cells Are Essential for the Self-Renewal and Repopulation of Notch-Dependent Hematopoietic Stem Cells. Cell stem cell *6*, 251-264.

Cai, Z., de Bruijn, M., Ma, X., Dortland, B., Luteijn, T., Downing, R.J., and Dzierzak, E. (2000). Haploinsufficiency of AML1 affects the temporal and spatial generation of hematopoietic stem cells in the mouse embryo. Immunity *13*, 423-431.

Caiazzo, M., Dell'Anno, M.T., Dvoretskova, E., Lazarevic, D., Taverna, S., Leo, D., Sotnikova, T.D., Menegon, A., Roncaglia, P., Colciago, G.*, et al.* (2011). Direct generation of functional dopaminergic neurons from mouse and human fibroblasts. Nature *476*, 224-227.

Calero-Nieto, F.J., Ng, F.S., Wilson, N.K., Hannah, R., Moignard, V., Leal-Cervantes, A.I., Jimenez-Madrid, I., Diamanti, E., Wernisch, L., and Gottgens, B. (2014). Key regulators control distinct transcriptional programmes in blood progenitor and mast cells. EMBO J *33*, 1212-1226.

Calvi, L.M., Adams, G.B., Weibrecht, K.W., Weber, J.M., Olson, D.P., Knight, M.C., Martin, R.P., Schipani, E., Divieti, P., Bringhurst, F.R.*, et al.* (2003). Osteoblastic cells regulate the haematopoietic stem cell niche. Nature *425*, 841-846.

Chanda, S., Ang, C.E., Davila, J., Pak, C., Mall, M., Lee, Q.Y., Ahlenius, H., Jung, S.W., Sudhof, T.C., and Wernig, M. (2014). Generation of induced neuronal cells by the single reprogramming factor ASCL1. Stem cell reports *3*, 282-296.

Chang, C.J., and Bouhassira, E.E. (2012). Zinc-finger nuclease-mediated correction of alpha-thalassemia in iPS cells. Blood *120*, 3906-3914.

Chang, H.H., Hemberg, M., Barahona, M., Ingber, D.E., and Huang, S. (2008). Transcriptome-wide noise controls lineage choice in mammalian progenitor cells. Nature *453*, 544-547.

Chen, J., Ellison, F.M., Keyvanfar, K., Omokaro, S.O., Desierto, M.J., Eckhaus, M.A., and Young, N.S. (2008). Enrichment of hematopoietic stem cells with SLAM and LSK markers for the detection of hematopoietic stem cell function in normal and Trp53 null mice. Exp Hematol *36*, 1236-1243.

Chen, J.L., Li, J., Kiriluk, K.J., Rosen, A.M., Paner, G.P., Antic, T., Lussier, Y.A., and Vander Griend, D.J. (2012). Deregulation of a Hox protein regulatory network spanning prostate cancer initiation and progression. Clinical cancer research : an official journal of the American Association for Cancer Research *18*, 4291-4302.

Chen, M.J., Li, Y., De Obaldia, M.E., Yang, Q., Yzaguirre, A.D., Yamada-Inagawa, T., Vink, C.S., Bhandoola, A., Dzierzak, E., and Speck, N.A. (2011). Erythroid/myeloid progenitors and hematopoietic stem cells originate from distinct populations of endothelial cells. Cell stem cell *9*, 541-552.

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

Chen, Y.J., Finkbeiner, S.R., Weinblatt, D., Emmett, M.J., Tameire, F., Yousefi, M., Yang, C., Maehr, R., Zhou, Q., Shemer, R.*, et al.* (2014). De novo formation of insulin-producing "neo-beta cell islets" from intestinal crypts. Cell reports *6*, 1046-1058.

Cherry, A.B., Gagne, K.E., McLoughlin, E.M., Baccei, A., Gorman, B., Hartung, O., Miller, J.D., Zhang, J., Zon, R.L., Ince, T.A.*, et al.* (2013). Induced pluripotent stem cells with a mitochondrial DNA deletion. Stem Cells *31*, 1287-1297.

Chiou, S.H., Jiang, B.H., Yu, Y.L., Chou, S.J., Tsai, P.H., Chang, W.C., Chen, L.K., Chen, L.H., Chien, Y., and Chiou, G.Y. (2013). Poly(ADP-ribose) polymerase 1 regulates nuclear reprogramming and promotes iPSC generation without c-Myc. J Exp Med *210*, 85-98.

Cho, H.J., Lee, C.S., Kwon, Y.W., Paek, J.S., Lee, S.H., Hur, J., Lee, E.J., Roh, T.Y., Chu, I.S., Leem, S.H.*, et al.* (2010). Induction of pluripotent stem cells from adult somatic cells by protein-based reprogramming without genetic manipulation. Blood *116*, 386-395.

Choi, J., Costa, M.L., Mermelstein, C.S., Chagas, C., Holtzer, S., and Holtzer, H. (1990). MyoD converts primary dermal fibroblasts, chondroblasts, smooth muscle, and retinal pigmented epithelial cells into striated mononucleated myoblasts and multinucleated myotubes, pp. 7988-7992.

Choi, K., Kennedy, M., Kazarov, A., Papadimitriou, J.C., and Keller, G. (1998). A common precursor for hematopoietic and endothelial cells. Development *125*, 725-732.

Chotinantakul, K., and Leeanansaksiri, W. (2012). Hematopoietic stem cell development, niches, and signaling pathways. Bone marrow research *2012*, 270425.

Chou, S.T., Khandros, E., Bailey, L.C., Nichols, K.E., Vakoc, C.R., Yao, Y., Huang, Z., Crispino, J.D., Hardison, R.C., Blobel, G.A.*, et al.* (2009). Graded repression of PU.1/Sfpi1 gene transcription by GATA factors regulates hematopoietic cell fate. Blood *114*, 983-994.

Ciau-Uitz, A., Wang, L., Patient, R., and Liu, F. (2013). ETS transcription factors in hematopoietic stem cell development. Blood cells, molecules & diseases *51*, 248-255.

Cobaleda, C., Jochum, W., and Busslinger, M. (2007). Conversion of mature B cells into T cells by dedifferentiation to uncommitted progenitors. Nature *449*, 473-477.

Corcoran, L.M., Karvelas, M., Nossal, G.J., Ye, Z.S., Jacks, T., and Baltimore, D. (1993). Oct-2, although not required for early B-cell development, is critical for later B-cell maturation and for postnatal survival. Genes & development *7*, 570-582.

Costa, G., Mazan, A., Gandillet, A., Pearson, S., Lacaud, G., and Kouskoff, V. (2012). SOX7 regulates the expression of VE-cadherin in the haemogenic endothelium at the onset of haematopoietic development. Development *139*, 1587-1598.

Cumano, A., Ferraz, J.C., Klaine, M., Di Santo, J.P., and Godin, I. (2001). Intraembryonic, but not yolk sac hematopoietic precursors, isolated before circulation, provide long-term multilineage reconstitution. Immunity *15*, 477-485.

Dahl, R., Iyer, S.R., Owens, K.S., Cuylear, D.D., and Simon, M.C. (2007). The Transcriptional Repressor GFI-1 Antagonizes PU.1 Activity through Protein-Protein Interaction. Journal of Biological Chemistry *282*, 6473-6483.

Dakic, A., Metcalf, D., Di Rago, L., Mifsud, S., Wu, L., and Nutt, S.L. (2005). PU.1 regulates the commitment of adult hematopoietic progenitors and restricts granulopoiesis. The Journal of Experimental Medicine *201*, 1487-1502.

Davis, R.L., Weintraub, H., and Lassar, A.B. (1987). Expression of a single transfected cDNA converts fibroblasts to myoblasts. Cell *51*, 987-1000.

Davis, R.P., Nemes, C., Varga, E., Freund, C., Kosmidis, G., Gkatzis, K., de Jong, D., Szuhai, K., Dinnyes, A., and Mummery, C.L. (2013). Generation of induced pluripotent stem cells from human foetal fibroblasts using the Sleeping Beauty transposon gene delivery system. Differentiation; research in biological diversity *86*, 30-37.

Di Tullio, A., Vu Manh, T.P., Schubert, A., Castellano, G., Mansson, R., and Graf, T. (2011). CCAAT/enhancer binding protein alpha (C/EBP(alpha))-induced transdifferentiation of pre-B cells into macrophages involves no overt retrodifferentiation. Proceedings of the National Academy of Sciences of the United States of America *108*, 17016-17021.

Dieterlen-Lievre, F., Pouget, C., Bollerot, K., and Jaffredo, T. (2006). Are intra-aortic hemopoietic cells derived from endothelial cells during ontogeny? Trends Cardiovasc Med *16*, 128-139.

Donehower, L.A., Harvey, M., Vogel, H., McArthur, M.J., Montgomery, C.A., Jr., Park, S.H., Thompson, T., Ford, R.J., and Bradley, A. (1995). Effects of genetic background on tumorigenesis in p53-deficient mice. Molecular carcinogenesis *14*, 16-22.

Doulatov, S., Vo, L.T., Chou, S.S., Kim, P.G., Arora, N., Li, H., Hadland, B.K., Bernstein, I.D., Collins, J.J., Zon, L.I.*, et al.* (2013). Induction of multipotential hematopoietic progenitors from human pluripotent stem cells via respecification of lineage-restricted precursors. Cell stem cell *13*, 459-470.

Du, Y., Wang, J., Jia, J., Song, N., Xiang, C., Xu, J., Hou, Z., Su, X., Liu, B., Jiang, T., *et al.* (2014). Human hepatocytes with drug metabolic function induced from fibroblasts by lineage reprogramming. Cell stem cell *14*, 394-403.

Dudgeon, C., Chan, C., Kang, W., Sun, Y., Emerson, R., Robins, H., and Levine, A.J. (2014). The evolution of thymic lymphomas in p53 knockout mice. Genes & development *28*, 2613-2620.

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

Efe, J.A., Hilcove, S., Kim, J., Zhou, H., Ouyang, K., Wang, G., Chen, J., and Ding, S. (2011). Conversion of mouse fibroblasts into cardiomyocytes using a direct reprogramming strategy. Nat Cell Biol *13*, 215-222.

Eggan, K., Baldwin, K., Tackett, M., Osborne, J., Gogos, J., Chess, A., Axel, R., and Jaenisch, R. (2004). Mice cloned from olfactory sensory neurons. Nature *428*, 44-49.

Eilken, H.M., Nishikawa, S.-I., and Schroeder, T. (2009). Continuous single-cell imaging of blood generation from haemogenic endothelium. Nature *457*, 896-900.

Elcheva, I., Brok-Volchanskaya, V., Kumar, A., Liu, P., Lee, J.H., Tong, L., Vodyanik, M., Swanson, S., Stewart, R., Kyba, M*., et al.* (2014). Direct induction of haematoendothelial programs in human pluripotent stem cells by transcriptional regulators. Nature communications *5*, 4372.

Eminli, S., Foudi, A., Stadtfeld, M., Maherali, N., Ahfeldt, T., Mostoslavsky, G., Hock, H., and Hochedlinger, K. (2009). Differentiation stage determines potential of hematopoietic cells for reprogramming into induced pluripotent stem cells. Nature genetics *41*, 968-976.

Endoh, M., Ogawa, M., Orkin, S., and Nishikawa, S. (2002). SCL/tal-1-dependent process determines a competence to select the definitive hematopoietic lineage prior to endothelial differentiation. EMBO J *21*, 6700-6708.

Esteban, M.A., Xu, J., Yang, J., Peng, M., Qin, D., Li, W., Jiang, Z., Chen, J., Deng, K., Zhong, M*., et al.* (2009). Generation of Induced Pluripotent Stem Cell Lines from Tibetan Miniature Pig, pp. 17634-17640.

Fahl, S.P., Harris, B., Coffey, F., and Wiest, D.L. (2015). Rpl22 Loss Impairs the Development of B Lymphocytes by Activating a p53-Dependent Checkpoint. Journal of immunology *194*, 200-209.

Fares, I., Chagraoui, J., Gareau, Y., Gingras, S., Ruel, R., Mayotte, N., Csaszar, E., Knapp, D.J., Miller, P., Ngom, M*., et al.* (2014). Cord blood expansion. Pyrimidoindole derivatives are agonists of human hematopoietic stem cell self-renewal. Science *345*, 1509-1512.

Fehling, H.J.r., Lacaud, G., Kubo, A., Kennedy, M., Robertson, S., Keller, G., and Kouskoff, V. (2003). Tracking mesoderm induction and its specification to the hemangioblast during embryonic stem cell differentiation, pp. 4217-4227.

Feng, B., Jiang, J., Kraus, P., Ng, J.-H., Heng, J.-C.D., Chan, Y.-S., Yaw, L.-P., Zhang, W., Loh, Y.-H., Han, J.*, et al.* (2009). Reprogramming of fibroblasts into induced pluripotent stem cells with orphan nuclear receptor Esrrb. Nat Cell Biol *11*, 197-203.

Feng, R., Desbordes, S.C., Xie, H., Tillo, E.S., Pixley, F., Stanley, E.R., and Graf, T. (2008). PU.1 and C/EBPalpha/beta convert fibroblasts into macrophage-like cells. Proceedings of the National Academy of Sciences of the United States of America *105*, 6057-6062.

Ferkowicz, M.J., Starr, M., Xie, X., Li, W., Johnson, S.A., Shelley, W.C., Morrison, P.R., and Yoder, M.C. (2003). CD41 expression defines the onset of primitive and definitive hematopoiesis in the murine embryo. Development *130*, 4393-4403.

Ferkowicz, M.J., and Yoder, M.C. (2005). Blood island formation: longstanding observations and modern interpretations. Exp Hematol *33*, 1041-1047.

Ferreras, C., Lancrin, C., Lie, A.L.M., Kouskoff, V., and Lacaud, G. (2011). Identification and characterization of a novel transcriptional target of RUNX1/AML1 at the onset of hematopoietic development. Blood *118*, 594-597.

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

Fournier, M., Savoie-Rondeau, I., Larochelle, F., Hassawi, M., Shestakova, E.A., Roy, D.C., and Bijl, J.J. (2014). Inability of HOXB4 to enhance self-renewal of malignant B cells: favorable profile for the expansion of autologous hematopoietic stem cells. Exp Hematol *42*, 526-535 e524.

Fukuchi, Y., Ito, M., Shibata, F., Kitamura, T., and Nakajima, H. (2008). Activation of CCAAT/enhancer-binding protein alpha or PU.1 in hematopoietic stem cells leads to their reduced self-renewal and proliferation. Stem Cells *26*, 3172-3181.

Fukuchi, Y., Shibata, F., Ito, M., Goto-Koshino, Y., Sotomaru, Y., Ito, M., Kitamura, T., and Nakajima, H. (2006). Comprehensive analysis of myeloid lineage conversion using mice expressing an inducible form of C/EBP alpha. Embo J *25*, 3398-3410.

Gandillet, A., Serrano, A.G., Pearson, S., Lie-A-Ling, M., Lacaud, G., and Kouskoff, V. (2009). Sox7-sustained expression alters the balance between proliferation and differentiation of hematopoietic progenitors at the onset of blood specification. Blood *114*, 4813-4822.

Gao, Y., Chen, J., Li, K., Wu, T., Huang, B., Liu, W., Kou, X., Zhang, Y., Huang, H., Jiang, Y.*, et al.* (2013). Replacement of Oct4 by Tet1 during iPSC induction reveals an important role of DNA methylation and hydroxymethylation in reprogramming. Cell stem cell *12*, 453-469.

Garcia-Porrero, J.A., Godin, I.E., and Dieterlen-Lievre, F. (1995). Potential intraembryonic hemogenic sites at pre-liver stages in the mouse. Anatomy and embryology *192*, 425-435.

Garcia-Porrero, J.A., Manaia, A., Jimeno, J., Lasky, L.L., Dieterlen-Lievre, F., and Godin, I.E. (1998). Antigenic profiles of endothelial and hemopoietic lineages in murine intraembryonic hemogenic sites. Developmental and comparative immunology *22*, 303-319.

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

Gering, M., Yamada, Y., Rabbitts, T.H., and Patient, R.K. (2003). Lmo2 and Scl/Tal1 convert non-axial mesoderm into haemangioblasts which differentiate into endothelial cells in the absence of Gata1. Development *130*, 6187-6199.

Glotzbach, J.P., Januszyk, M., Vial, I.N., Wong, V.W., Gelbard, A., Kalisky, T., Thangarajah, H., Longaker, M.T., Quake, S.R., Chu, G.*, et al.* (2011). An information theoretic, microfluidic-based single cell analysis permits identification of subpopulations among putatively homogeneous stem cells. PloS one *6*, e21211.

Gordon-Keylock, S., Sobiesiak, M., Rybtsov, S., Moore, K., and Medvinsky, A. (2013). Mouse extraembryonic arterial vessels harbor precursors capable of maturing into definitive HSCs. Blood *122*, 2338-2345.

Gore, A., Li, Z., Fung, H.L., Young, J.E., Agarwal, S., Antosiewicz-Bourget, J., Canto, I., Giorgetti, A., Israel, M.A., Kiskinis, E.*, et al.* (2011). Somatic coding mutations in human induced pluripotent stem cells. Nature *471*, 63-67.

Gottgens, B., Nastos, A., Kinston, S., Piltz, S., Delabesse, E.C., Stanley, M., Sanchez, M.J., Ciau-Uitz, A., Patient, R., and Green, A.R. (2002). Establishing the transcriptional programme for blood: the SCL stem cell enhancer is regulated by a multiprotein complex containing Ets and GATA factors. Embo J *21*, 3039-3050.

Green, L.L., and Jakobovits, A. (1998). Regulation of B cell development by variable gene complexity in mice reconstituted with human immunoglobulin yeast artificial chromosomes. J Exp Med *188*, 483-495.

Guo, Z., Zhang, L., Wu, Z., Chen, Y., Wang, F., and Chen, G. (2014). In vivo direct reprogramming of reactive glial cells into functional neurons after brain injury and in an Alzheimer's disease model. Cell stem cell *14*, 188-202.

Gurdon, J.B. (1962). The developmental capacity of nuclei taken from intestinal epithelium cells of feeding tadpoles. Journal of embryology and experimental morphology *10*, 622-640.

Hadland, B.K., Huppert, S.S., Kanungo, J., Xue, Y., Jiang, R., Gridley, T., Conlon, R.A., Cheng, A.M., Kopan, R., and Longmore, G.D. (2004). A requirement for Notch1 distinguishes 2 phases of definitive hematopoiesis during development. Blood *104*, 3097-3105.

Hainaut, P., Hernandez, T., Robinson, A., Rodriguez-Tome, P., Flores, T., Hollstein, M., Harris, C.C., and Montesano, R. (1998). IARC Database of p53 gene mutations in human tumors and cell lines: updated compilation, revised formats and new visualisation tools. Nucleic Acids Res *26*, 205-213.

Hall, M.A., Curtis, D.J., Metcalf, D., Elefanty, A.G., Sourris, K., Robb, L., Gothert, J.R., Jane, S.M., and Begley, C.G. (2003). The critical regulator of embryonic hematopoiesis, SCL, is vital in the adult for megakaryopoiesis, erythropoiesis, and lineage choice in CFU-S12. Proceedings of the National Academy of Sciences of the United States of America *100*, 992-997.

Hamaguchi, I., Huang, X.L., Takakura, N., Tada, J., Yamaguchi, Y., Kodama, H., and Suda, T. (1999). In vitro hematopoietic and endothelial cell development from cells expressing TEK receptor in murine aorta-gonad-mesonephros region. Blood *93*, 1549-1556.

Han, D.W., Tapia, N., Hermann, A., Hemmer, K., Hoing, S., Arauzo-Bravo, M.J., Zaehres, H., Wu, G., Frank, S., Moritz, S.*, et al.* (2012). Direct reprogramming of fibroblasts into neural stem cells by defined factors. Cell stem cell *10*, 465-472.

Hanna, J., Wernig, M., Markoulaki, S., Sun, C.-W., Meissner, A., Cassady, J.P., Beard, C., Brambrink, T., Wu, L.-C., Townes, T.M.*, et al.* (2007). Treatment of Sickle Cell Anemia Mouse Model with iPS Cells Generated from Autologous Skin, pp. 1920-1923.

Hara, T., Nakano, Y., Tanaka, M., Tamura, K., Sekiguchi, T., Minehata, K., Copeland, N.G., Jenkins, N.A., Okabe, M., Kogo, H.*, et al.* (1999). Identification of podocalyxin-like protein 1 as a novel cell surface marker for hemangioblasts in the murine aorta-gonad-mesonephros region. Immunity *11*, 567-578.

Hardy, R.R., and Hayakawa, K. (1991). A developmental switch in B lymphopoiesis. Proceedings of the National Academy of Sciences of the United States of America *88*, 11550-11554.

Hart, A., Melet, F., Grossfeld, P., Chien, K., Jones, C., Tunnacliffe, A., Favier, R., and Bernstein, A. (2000). Fli-1 Is Required for Murine Vascular and Megakaryocytic Development and Is Hemizygously Deleted in Patients with Thrombocytopenia. Immunity *13*, 167-177.

Hashimoto, K., Fujimoto, T., Shimoda, Y., Huang, X., Sakamoto, H., and Ogawa, M. (2007). Distinct hemogenic potential of endothelial cells and CD41+ cells in mouse embryos. Development, growth & differentiation *49*, 287-300.

Hayhoe, F.G.J., Flemans, R.J., and Hayhoe, F.G.J. (1982). Color atlas of hematological cytology, 2nd edn (New York: Wiley).

Heinrich, C., Blum, R., Gascon, S., Masserdotti, G., Tripathi, P., Sanchez, R., Tiedt, S., Schroeder, T., Gotz, M., and Berninger, B. (2010). Directing astroglia from the cerebral cortex into subtype specific functional neurons. PLoS biology *8*, e1000373.

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

Heng, J.C.D., Feng, B., Han, J.Y., Jiang, J.M., Kraus, P., Ng, J.H., Orlov, Y.L., Huss, M., Yang, L., Lufkin, T.*, et al.* (2010). The Nuclear Receptor Nr5a2 Can Replace Oct4 in the Reprogramming of Murine Somatic Cells to Pluripotent Cells. Cell stem cell *6*, 167-174.

Heyworth, C., Pearson, S., May, G., and Enver, T. (2002). Transcription factor-mediated lineage switching reveals plasticity in primary committed progenitor cells. EMBO J *21*, 3770-3781.

Himburg, H.A., Muramoto, G.G., Daher, P., Meadows, S.K., Russell, J.L., Doan, P., Chi, J.T., Salter, A.B., Lento, W.E., Reya, T.*, et al.* (2010). Pleiotrophin regulates the expansion and regeneration of hematopoietic stem cells. Nature medicine *16*, 475-482.

Hochedlinger, K., and Jaenisch, R. (2002). Monoclonal mice generated by nuclear transfer from mature B and T donor cells. Nature *415*, 1035-1038.

Hock, H., Hamblen, M.J., Rooke, H.M., Schindler, J.W., Saleque, S., Fujiwara, Y., and Orkin, S.H. (2004a). Gfi-1 restricts proliferation and preserves functional integrity of haematopoietic stem cells. Nature *431*, 1002-1007.

Hock, H., Meade, E., Medeiros, S., Schindler, J.W., Valk, P.J., Fujiwara, Y., and Orkin, S.H. (2004b). Tel/Etv6 is an essential and selective regulator of adult hematopoietic stem cell survival. Genes & development *18*, 2336-2341.

Hou, P., Li, Y., Zhang, X., Liu, C., Guan, J., Li, H., Zhao, T., Ye, J., Yang, W., Liu, K.*, et al.* (2013). Pluripotent stem cells induced from mouse somatic cells by small-molecule compounds. Science *341*, 651-654.

Hu, K., Yu, J., Suknuntha, K., Tian, S., Montgomery, K., Choi, K.D., Stewart, R., Thomson, J.A., and Slukvin, II (2011). Efficient generation of transgene-free induced pluripotent stem cells from normal and neoplastic bone marrow and cord blood mononuclear cells. Blood *117*, e109-119.

Hu, M., Krause, D., Greaves, M., Sharkis, S., Dexter, M., Heyworth, C., and Enver, T. (1997). Multilineage gene expression precedes commitment in the hemopoietic system. Genes & development *11*, 774-785.

Hu, W., Qiu, B., Guan, W., Wang, Q., Wang, M., Li, W., Gao, L., Shen, L., Huang, Y., Xie, G.*, et al.* (2015). Direct Conversion of Normal and Alzheimer's Disease Human Fibroblasts into Neuronal Cells by Small Molecules. Cell stem cell *17*, 204-212.

Huang, G., Zhang, P., Hirai, H., Elf, S., Yan, X., Chen, Z., Koschmieder, S., Okuno, Y., Dayaram, T., Growney, J.D.*, et al.* (2008). PU.1 is a major downstream target of AML1 (RUNX1) in adult mouse hematopoiesis. Nature genetics *40*, 51-60.

Huang, P., He, Z., Ji, S., Sun, H., Xiang, D., Liu, C., Hu, Y., Wang, X., and Hui, L. (2011). Induction of functional hepatocyte-like cells from mouse fibroblasts by defined factors. Nature *475*, 386-389.

Huang, P., Zhang, L., Gao, Y., He, Z., Yao, D., Wu, Z., Cen, J., Chen, X., Liu, C., Hu, Y.*, et al.* (2014). Direct reprogramming of human fibroblasts to functional and expandable hepatocytes. Cell stem cell *14*, 370-384.

Huang, Y., Sitwala, K., Bronstein, J., Sanders, D., Dandekar, M., Collins, C., Robertson, G., MacDonald, J., Cezard, T., Bilenky, M.*, et al.* (2012). Identification and characterization of Hoxa9 binding sites in hematopoietic cells. Blood *119*, 388-398.

Huang, Z., Dore, L.C., Li, Z., Orkin, S.H., Feng, G., Lin, S., and Crispino, J.D. (2009). GATA-2 reinforces megakaryocyte development in the absence of GATA-1. Mol Cell Biol *29*, 5168-5180.

Huber, T.L., Kouskoff, V., Joerg Fehling, H., Palis, J., and Keller, G. (2004). Haemangioblast commitment is initiated in the primitive streak of the mouse embryo. Nature *432*, 625-630.

Ichida, J.K., Blanchard, J., Lam, K., Son, E.Y., Chung, J.E., Egli, D., Loh, K.M., Carter, A.C., Di Giorgio, F.P., Koszka, K.*, et al.* (2009). A small-molecule inhibitor of tgf-Beta signaling replaces sox2 in reprogramming by inducing nanog. Cell stem cell *5*, 491-503.

Ieda, M., Fu, J.D., Delgado-Olguin, P., Vedantham, V., Hayashi, Y., Bruneau, B.G., and Srivastava, D. (2010). Direct reprogramming of fibroblasts into functional cardiomyocytes by defined factors. Cell *142*, 375-386.

Inagawa, K., Miyamoto, K., Yamakawa, H., Muraoka, N., Sadahiro, T., Umei, T., Wada, R., Katsumata, Y., Kaneda, R., Nakade, K.*, et al.* (2012). Induction of cardiomyocyte-like cells in infarct hearts by gene transfer of Gata4, Mef2c, and Tbx5. Circulation research *111*, 1147-1156.

Ismailoglu, I., Yeamans, G., Daley, G.Q., Perlingeiro, R.C., and Kyba, M. (2008). Mesodermal patterning activity of SCL. Exp Hematol *36*, 1593-1603.

Ivanova, N.B., Dimos, J.T., Schaniel, C., Hackney, J.A., Moore, K.A., and Lemischka, I.R. (2002). A stem cell molecular signature. Science *298*, 601-604.

Ivanovs, A., Rybtsov, S., Anderson, R.A., Turner, M.L., and Medvinsky, A. (2014). Identification of the niche and phenotype of the first human hematopoietic stem cells. Stem cell reports *2*, 449-456.

Ivanovs, A., Rybtsov, S., Welch, L., Anderson, R.A., Turner, M.L., and Medvinsky, A. (2011). Highly potent human hematopoietic stem cells first emerge in the intraembryonic aorta-gonad-mesonephros region. J Exp Med *208*, 2417-2427.

Iwasaki, H., Mizuno, S.-i., Wells, R.A., Cantor, A.B., Watanabe, S., and Akashi, K. (2003). GATA-1 Converts Lymphoid and Myelomonocytic Progenitors into the Megakaryocyte/Erythrocyte Lineages. Immunity *19*, 451-462.

Iwasaki, H., Mizuno, S., Arinobu, Y., Ozawa, H., Mori, Y., Shigematsu, H., Takatsu, K., Tenen, D.G., and Akashi, K. (2006). The order of expression of transcription factors directs hierarchical specification of hematopoietic lineages. Genes & development *20*, 3010-3021.

Iwasaki, H., Somoza, C., Shigematsu, H., Duprez, E.A., Iwasaki-Arai, J., Mizuno, S.-i., Arinobu, Y., Geary, K., Zhang, P., Dayaram, T.*, et al.* (2005). Distinctive and indispensable roles of PU.1 in maintenance of hematopoietic stem cells and their differentiation. Blood *106*, 1590-1600.

Jacks, T., Remington, L., Williams, B.O., Schmitt, E.M., Halachmi, S., Bronson, R.T., and Weinberg, R.A. (1994). Tumor spectrum analysis in p53-mutant mice. Current biology : CB *4*, 1-7.

Jaffredo, T., Gautier, R., Eichmann, A., and Dieterlen-Lievre, F. (1998). Intraaortic hemopoietic cells are derived from endothelial cells during ontogeny. Development *125*, 4575-4583.

Jarriault, S., Schwab, Y., and Greenwald, I. (2008). A Caenorhabditis elegans model for epithelial-neuronal transdifferentiation. Proceedings of the National Academy of Sciences of the United States of America *105*, 3790-3795.

Karanu, F.N., Murdoch, B., Miyabayashi, T., Ohno, M., Koremoto, M., Gallacher, L., Wu, D., Itoh, A., Sakano, S., and Bhatia, M. (2001). Human homologues of Delta-1 and Delta-4 function as mitogenic regulators of primitive human hematopoietic cells. Blood *97*, 1960-1967.

Karsunky, H., Zeng, H., Schmidt, T., Zevnik, B., Kluge, R., Schmid, K.W., Duhrsen, U., and Moroy, T. (2002). Inflammatory reactions and severe neutropenia in mice lacking the transcriptional repressor Gfi1. Nature genetics *30*, 295-300.

Kataoka, H., Hayashi, M., Nakagawa, R., Tanaka, Y., Izumi, N., Nishikawa, S., Jakt, M.L., Tarui, H., and Nishikawa, S. (2011). Etv2/ER71 induces vascular mesoderm from Flk1+PDGFRalpha+ primitive mesoderm. Blood *118*, 6975-6986.

Keller, G. (2005). Embryonic stem cell differentiation: emergence of a new era in biology and medicine. Genes & development *19*, 1129-1155.

Kent, D.G., Copley, M.R., Benz, C., Wohrer, S., Dykstra, B.J., Ma, E., Cheyne, J., Zhao, Y., Bowie, M.B., Zhao, Y.*, et al.* (2009). Prospective isolation and molecular characterization of hematopoietic stem cells with durable self-renewal potential. Blood *113*, 6342-6350.

Khan, M., Narayanan, K., Lu, H., Choo, Y., Du, C., Wiradharma, N., Yang, Y.Y., and Wan, A.C. (2013). Delivery of reprogramming factors into fibroblasts for generation of

non-genetic induced pluripotent stem cells using a cationic bolaamphiphile as a non-viral vector. Biomaterials *34*, 5336-5343.

Khandanpour, C., Kosan, C., Gaudreau, M.C., Duhrsen, U., Hebert, J., Zeng, H., and Moroy, T. (2011). Growth factor independence 1 protects hematopoietic stem cells against apoptosis but also prevents the development of a myeloproliferative-like disease. Stem Cells *29*, 376-385.

Khandanpour, C., Sharif-Askari, E., Vassen, L., Gaudreau, M.C., Zhu, J., Paul, W.E., Okayama, T., Kosan, C., and Moroy, T. (2010). Evidence that growth factor independence 1b regulates dormancy and peripheral blood mobilization of hematopoietic stem cells. Blood *116*, 5149-5161.

Kim, D., Kim, C.H., Moon, J.I., Chung, Y.G., Chang, M.Y., Han, B.S., Ko, S., Yang, E., Cha, K.Y., Lanza, R.*, et al.* (2009a). Generation of human induced pluripotent stem cells by direct delivery of reprogramming proteins. Cell stem cell *4*, 472-476.

Kim, I., Yilmaz, O.H., and Morrison, S.J. (2005). CD144 (VE-cadherin) is transiently expressed by fetal liver hematopoietic stem cells. Blood *106*, 903-905.

Kim, J., Su, S.C., Wang, H., Cheng, A.W., Cassady, J.P., Lodato, M.A., Lengner, C.J., Chung, C.Y., Dawlaty, M.M., Tsai, L.H.*, et al.* (2011). Functional integration of dopaminergic neurons directly converted from mouse fibroblasts. Cell stem cell *9*, 413-419.

Kim, J.B., Sebastiano, V., Wu, G., Arauzo-Bravo, M.J., Sasse, P., Gentile, L., Ko, K., Ruau, D., Ehrich, M., van den Boom, D.*, et al.* (2009b). Oct4-induced pluripotency in adult neural stem cells. Cell *136*, 411-419.

Kim, J.B., Zaehres, H., Wu, G., Gentile, L., Ko, K., Sebastiano, V., Arauzo-Bravo, M.J., Ruau, D., Han, D.W., Zenke, M.*, et al.* (2008). Pluripotent stem cells induced from adult neural stem cells by reprogramming with two factors. Nature *454*, 646-650.

Kinder, S.J., Tsang, T.E., Quinlan, G.A., Hadjantonakis, A.K., Nagy, A., and Tam, P.P. (1999). The orderly allocation of mesodermal cells to the extraembryonic structures and the anteroposterior axis during gastrulation of the mouse embryo. Development *126*, 4691-4701.

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

Koyano-Nakagawa, N., Kweon, J., Iacovino, M., Shi, X., Rasmussen, T.L., Borges, L., Zirbes, K.M., Li, T., Perlingeiro, R.C., Kyba, M.*, et al.* (2012). Etv2 is expressed in the yolk sac hematopoietic and endothelial progenitors and regulates Lmo2 gene expression. Stem Cells *30*, 1611-1623.

Kumano, K., Chiba, S., Kunisato, A., Sata, M., Saito, T., Nakagami-Yamaguchi, E., Yamaguchi, T., Masuda, S., Shimizu, K., Takahashi, T.*, et al.* (2003). Notch1 but not Notch2 is essential for generating hematopoietic stem cells from endothelial cells. Immunity *18*, 699-711.

Kunisato, A., Chiba, S., Saito, T., Kumano, K., Nakagami-Yamaguchi, E., Yamaguchi, T., and Hirai, H. (2004). Stem cell leukemia protein directs hematopoietic stem cell fate. Blood *103*, 3336-3341.

Kyba, M., Perlingeiro, R.C., and Daley, G.Q. (2002). HoxB4 confers definitive lymphoid-myeloid engraftment potential on embryonic stem cell and yolk sac hematopoietic progenitors. Cell *109*, 29-37.

Lacaud, G., Keller, G., and Kouskoff, V. (2004). Tracking Mesoderm Formation and Specification to the Hemangioblast in Vitro. Trends in Cardiovascular Medicine *14*, 314-317.

Laiosa, C.V., Stadtfeld, M., Xie, H., de Andres-Aguayo, L., and Graf, T. (2006). Reprogramming of committed T cell progenitors to macrophages and dendritic cells by C/EBP alpha and PU.1 transcription factors. Immunity *25*, 731-744.

Lancrin, C., Mazan, M., Stefanska, M., Patel, R., Lichtinger, M., Costa, G., Vargel, O., Wilson, N.K., Moroy, T., Bonifer, C*., et al.* (2012). GFI1 and GFI1B control the loss of endothelial identity of hemogenic endothelium during hematopoietic commitment. Blood *120*, 314-322.

Lancrin, C., Sroczynska, P., Stephenson, C., Allen, T., Kouskoff, V., and Lacaud, G. (2009). The haemangioblast generates haematopoietic cells through a haemogenic endothelium stage. Nature *457*, 892-895.

Landry, J.R., Bonadies, N., Kinston, S., Knezevic, K., Wilson, N.K., Oram, S.H., Janes, M., Piltz, S., Hammett, M., Carter, J*., et al.* (2009). Expression of the leukemia oncogene Lmo2 is controlled by an array of tissue-specific elements dispersed over 100 kb and bound by Tal1/Lmo2, Ets, and Gata factors. Blood *113*, 5783-5792.

Lawson, K.A., Meneses, J.J., and Pedersen, R.A. (1991). Clonal analysis of epiblast fate during germ layer formation in the mouse embryo. Development *113*, 891-911.

Lawson, K.A., and Pedersen, R.A. (1987). Cell fate, morphogenetic movement and population kinetics of embryonic endoderm at the time of germ layer formation in the mouse. Development *101*, 627-652.

Lee, B.C., Kim, M.K., Jang, G., Oh, H.J., Yuda, F., Kim, H.J., Hossein, M.S., Kim, J.J., Kang, S.K., Schatten, G*., et al.* (2005). Dogs cloned from adult somatic cells. Nature *436*, 641.

Lee, D., Park, C., Lee, H., Lugus, J.J., Kim, S.H., Arentson, E., Chung, Y.S., Gomez, G., Kyba, M., Lin, S*., et al.* (2008). ER71 Acts Downstream of BMP, Notch, and Wnt Signaling in Blood and Vessel Progenitor Specification. Cell stem cell *2*, 497-507.

Li, W., Wei, W., Zhu, S., Zhu, J., Shi, Y., Lin, T., Hao, E., Hayek, A., Deng, H., and Ding, S. (2009). Generation of Rat and Human Induced Pluripotent Stem Cells by Combining Genetic Reprogramming and Chemical Inhibitors. Cell stem cell *4*, 16-19.

Li, X., Zuo, X., Jing, J., Ma, Y., Wang, J., Liu, D., Zhu, J., Du, X., Xiong, L., Du, Y.*, et al.* (2015). Small-Molecule-Driven Direct Reprogramming of Mouse Fibroblasts into Functional Neurons. Cell stem cell *17*, 195-203.

Li, Y., Zhang, Q., Yin, X., Yang, W., Du, Y., Hou, P., Ge, J., Liu, C., Zhang, W., Zhang, X.*, et al.* (2011). Generation of iPSCs from mouse fibroblasts with a single gene, Oct4, and small molecules. Cell research *21*, 196-204.

Li, Z., Lan, Y., He, W., Chen, D., Wang, J., Zhou, F., Wang, Y., Sun, H., Chen, X., Xu, C.*, et al.* (2012). Mouse embryonic head as a site for hematopoietic stem cell development. Cell stem cell *11*, 663-675.

Liakhovitskaia, A., Rybtsov, S., Smith, T., Batsivari, A., Rybtsova, N., Rode, C., de Bruijn, M., Buchholz, F., Gordon-Keylock, S., Zhao, S.*, et al.* (2014). Runx1 is required for progression of CD41+ embryonic precursors into HSCs but not prior to this. Development *141*, 3319-3323.

Lichtinger, M., Ingram, R., Hannah, R., Muller, D., Clarke, D., Assi, S.A., Lie, A.L.M., Noailles, L., Vijayabaskar, M.S., Wu, M.*, et al.* (2012). RUNX1 reshapes the epigenetic landscape at the onset of haematopoiesis. EMBO J *31*, 4318-4333.

Ling, K.W., Ottersbach, K., van Hamburg, J.P., Oziemlak, A., Tsai, F.Y., Orkin, S.H., Ploemacher, R., Hendriks, R.W., and Dzierzak, E. (2004). GATA-2 plays two functionally distinct roles during the ontogeny of hematopoietic stem cells. J Exp Med *200*, 871-882.

Liu, F., Bhang, S.H., Arentson, E., Sawada, A., Kim, C.K., Kang, I., Yu, J., Sakurai, N., Kim, S.H., Yoo, J.J.*, et al.* (2013). Enhanced hemangioblast generation and improved vascular repair and regeneration from embryonic stem cells by defined transcription factors. Stem cell reports *1*, 166-182.

Liu, F., Li, D., Yu, Y.Y., Kang, I., Cha, M.J., Kim, J.Y., Park, C., Watson, D.K., Wang, T., and Choi, K. (2015). Induction of hematopoietic and endothelial cell program orchestrated by ETS transcription factor ER71/ETV2. EMBO reports *16*, 654-669.

Liu, H., Ye, Z., Kim, Y., Sharkis, S., and Jang, Y.Y. (2010). Generation of endoderm-derived human induced pluripotent stem cells from primary hepatocytes. Hepatology *51*, 1810-1819.

Liu, H., Zhu, F., Yong, J., Zhang, P., Hou, P., Li, H., Jiang, W., Cai, J., Liu, M., Cui, K.*, et al.* (2008). Generation of Induced Pluripotent Stem Cells from Adult Rhesus Monkey Fibroblasts. Cell stem cell *3*, 587-590.

Liu, Y., Elf, S.E., Miyata, Y., Sashida, G., Liu, Y., Huang, G., Di Giandomenico, S., Lee, J.M., Deblasio, A., Menendez, S.*, et al.* (2009). p53 regulates hematopoietic stem cell quiescence. Cell stem cell *4*, 37-48.

Loughran, S.J., Kruse, E.A., Hacking, D.F., de Graaf, C.A., Hyland, C.D., Willson, T.A., Henley, K.J., Ellis, S., Voss, A.K., Metcalf, D.*, et al.* (2008). The transcription factor Erg

is essential for definitive hematopoiesis and the function of adult hematopoietic stem cells. Nat Immunol *9*, 810-819.

Lugus, J.J., Chung, Y.S., Mills, J.C., Kim, S.I., Grass, J., Kyba, M., Doherty, J.M., Bresnick, E.H., and Choi, K. (2007). GATA2 functions at multiple steps in hemangioblast development and differentiation. Development *134*, 393-405.

Maherali, N., Sridharan, R., Xie, W., Utikal, J., Eminli, S., Arnold, K., Stadtfeld, M., Yachechko, R., Tchieu, J., Jaenisch, R.*, et al.* (2007). Directly reprogrammed fibroblasts show global epigenetic remodeling and widespread tissue contribution. Cell stem cell *1*, 55-70.

Mali, P., Chou, B.K., Yen, J., Ye, Z., Zou, J., Dowey, S., Brodsky, R.A., Ohm, J.E., Yu, W., Baylin, S.B.*, et al.* (2010). Butyrate greatly enhances derivation of human induced pluripotent stem cells by promoting epigenetic remodeling and the expression of pluripotency-associated genes. Stem Cells *28*, 713-720.

Mansson, R., Hultquist, A., Luc, S., Yang, L., Anderson, K., Kharazi, S., Al-Hashmi, S., Liuba, K., Thoren, L., Adolfsson, J.*, et al.* (2007). Molecular evidence for hierarchical transcriptional lineage priming in fetal and adult stem cells and multipotent progenitors. Immunity *26*, 407-419.

Marro, S., Pang, Z.P., Yang, N., Tsai, M.C., Qu, K., Chang, H.Y., Sudhof, T.C., and Wernig, M. (2011). Direct lineage conversion of terminally differentiated hepatocytes to functional neurons. Cell stem cell *9*, 374-382.

Matsuoka, S., Tsuji, K., Hisakawa, H., Xu, M., Ebihara, Y., Ishii, T., Sugiyama, D., Manabe, A., Tanaka, R., Ikeda, Y.*, et al.* (2001). Generation of definitive hematopoietic stem cells from murine early yolk sac and paraaortic splanchnopleures by aorta-gonad-mesonephros region-derived stromal cells. Blood *98*, 6-12.

McCormack, M.P., Young, L.F., Vasudevan, S., de Graaf, C.A., Codrington, R., Rabbitts, T.H., Jane, S.M., and Curtis, D.J. (2010). The Lmo2 oncogene initiates leukemia in mice by inducing thymocyte self-renewal. Science *327*, 879-883.

McGrath, K.E., Frame, J.M., Fegan, K.H., Bowen, J.R., Conway, S.J., Catherman, S.C., Kingsley, P.D., Koniski, A.D., and Palis, J. (2015). Distinct Sources of Hematopoietic Progenitors Emerge before HSCs and Provide Functional Blood Cells in the Mammalian Embryo. Cell reports *11*, 1892-1904.

Medvinsky, A., and Dzierzak, E. (1996). Definitive Hematopoiesis Is Autonomously Initiated by the AGM Region. Cell *86*, 897-906.

Mercader, N., Tanaka, E.M., and Torres, M. (2005). Proximodistal identity during vertebrate limb regeneration is regulated by Meis homeodomain proteins. Development *132*, 4131-4142.

Mezey, E., Chandross, K.J., Harta, G., Maki, R.A., and McKercher, S.R. (2000). Turning blood into brain: cells bearing neuronal antigens generated in vivo from bone marrow. Science *290*, 1779-1782.

Mikkola, H.K., Fujiwara, Y., Schlaeger, T.M., Traver, D., and Orkin, S.H. (2003). Expression of CD41 marks the initiation of definitive hematopoiesis in the mouse embryo. Blood *101*, 508-516.

Miller, R.A., and Ruddle, F.H. (1977). Properties of teratocarcinoma-thymus somatic cell hybrids. Somatic cell genetics *3*, 247-261.

Mitchell, R.R., Szabo, E., Benoit, Y.D., Case, D.T., Mechael, R., Alamilla, J., Lee, J.H., Fiebig-Comyn, A., Gillespie, D.C., and Bhatia, M. (2014). Activation of neural cell fate programs toward direct conversion of adult human fibroblasts into tri-potent neural progenitors using OCT-4. Stem cells and development *23*, 1937-1946.

Miura, K., Okada, Y., Aoi, T., Okada, A., Takahashi, K., Okita, K., Nakagawa, M., Koyanagi, M., Tanabe, K., Ohnuki, M.*, et al.* (2009). Variation in the safety of induced pluripotent stem cell lines. Nature biotechnology *27*, 743-745.

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

Moon, J.H., Heo, J.S., Kim, J.S., Jun, E.K., Lee, J.H., Kim, A., Kim, J., Whang, K.Y., Kang, Y.K., Yeo, S.*, et al.* (2011). Reprogramming fibroblasts into induced pluripotent stem cells with Bmi1. Cell research *21*, 1305-1315.

Moon, J.H., Yun, W., Kim, J., Hyeon, S., Kang, P.J., Park, G., Kim, A., Oh, S., Whang, K.Y., Kim, D.W.*, et al.* (2013). Reprogramming of mouse fibroblasts into induced pluripotent stem cells with Nanog. Biochemical and biophysical research communications *431*, 444-449.

Moore, K.A., Ema, H., and Lemischka, I.R. (1997). In vitro maintenance of highly purified, transplantable hematopoietic stem cells. Blood *89*, 4337-4347.

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

Mossadegh-Keller, N., Sarrazin, S., Kandalla, P.K., Espinosa, L., Stanley, E.R., Nutt, S.L., Moore, J., and Sieweke, M.H. (2013). M-CSF instructs myeloid lineage fate in single haematopoietic stem cells. Nature *497*, 239-243.

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

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

Muller, L.U., Schlaeger, T.M., DeVine, A.L., and Williams, D.A. (2012). Induced pluripotent stem cells as a tool for gaining new insights into Fanconi anemia. Cell cycle (Georgetown, Tex *11*, 2985-2990.

Myers, C.T., and Krieg, P.A. (2013). BMP-mediated specification of the erythroid lineage suppresses endothelial development in blood island precursors. Blood *122*, 3929-3939.

Narula, J., Smith, A.M., Gottgens, B., and Igoshin, O.A. (2010). Modeling reveals bistability and low-pass filtering in the network module determining blood stem cell fate. PLoS computational biology *6*, e1000771.

Nerlov, C., and Graf, T. (1998). PU.1 induces myeloid lineage commitment in multipotent hematopoietic progenitors. Genes & development *12*, 2403-2412.

Nerlov, C., Querfurth, E., Kulessa, H., and Graf, T. (2000). GATA-1 interacts with the myeloid PU.1 transcription factor and represses PU.1-dependent transcription. Blood *95*, 2543-2551.

Nichols, J., Zevnik, B., Anastassiadis, K., Niwa, H., Klewe-Nebenius, D., Chambers, I., SchÄ¶ler, H., and Smith, A. (1998). Formation of Pluripotent Stem Cells in the Mammalian Embryo Depends on the POU Transcription Factor Oct4. Cell *95*, 379-391.

Nishikawa, S.I., Nishikawa, S., Hirashima, M., Matsuyoshi, N., and Kodama, H. (1998). Progressive lineage analysis by cell sorting and culture identifies FLK1+VE-cadherin+ cells at a diverging point of endothelial and hemopoietic lineages. Development *125*, 1747-1757.

Nishimura, T., Kaneko, S., Kawana-Tachikawa, A., Tajima, Y., Goto, H., Zhu, D., Nakayama-Hosoya, K., Iriguchi, S., Uemura, Y., Shimizu, T.*, et al.* (2013). Generation of rejuvenated antigen-specific T cells by reprogramming to pluripotency and redifferentiation. Cell stem cell *12*, 114-126.

Niwa, H., Miyazaki, J., and Smith, A.G. (2000). Quantitative expression of Oct-3/4 defines differentiation, dedifferentiation or self-renewal of ES cells. Nature genetics *24*, 372-376.

North, T., Gu, T.L., Stacy, T., Wang, Q., Howard, L., Binder, M., Marin-Padilla, M., and Speck, N.A. (1999). Cbfa2 is required for the formation of intra-aortic hematopoietic clusters. Development *126*, 2563-2575.

North, T.E., de Bruijn, M.F., Stacy, T., Talebian, L., Lind, E., Robin, C., Binder, M., Dzierzak, E., and Speck, N.A. (2002). Runx1 expression marks long-term repopulating hematopoietic stem cells in the midgestation mouse embryo. Immunity *16*, 661-672.

Nostro, M.C., Cheng, X., Keller, G.M., and Gadue, P. (2008). Wnt, activin, and BMP signaling regulate distinct stages in the developmental pathway from embryonic stem cells to blood. Cell stem cell *2*, 60-71.

Nottingham, W.T., Jarratt, A., Burgess, M., Speck, C.L., Cheng, J.-F., Prabhakar, S., Rubin, E.M., Li, P.-S., Sloane-Stanley, J., Kong-a-San, J.*, et al.* (2007). Runx1-mediated hematopoietic stem-cell emergence is controlled by a Gata/Ets/SCL-regulated enhancer, pp. 4188-4197.

Obokata, H., Wakayama, T., Sasai, Y., Kojima, K., Vacanti, M.P., Niwa, H., Yamato, M., and Vacanti, C.A. (2014). Stimulus-triggered fate conversion of somatic cells into pluripotency. Nature *505*, 641-647.

Odom, D.T., Dowell, R.D., Jacobsen, E.S., Gordon, W., Danford, T.W., MacIsaac, K.D., Rolfe, P.A., Conboy, C.M., Gifford, D.K., and Fraenkel, E. (2007). Tissue-specific transcriptional regulation has diverged significantly between human and mouse. Nature genetics *39*, 730-732.

Ogawa, E., Inuzuka, M., Maruyama, M., Satake, M., Naito-Fujimoto, M., Ito, Y., and Shigesada, K. (1993). Molecular cloning and characterization of PEBP2 beta, the heterodimeric partner of a novel Drosophila runt-related DNA binding protein PEBP2 alpha. Virology *194*, 314-331.

Ohnishi, K., Semi, K., Yamamoto, T., Shimizu, M., Tanaka, A., Mitsunaga, K., Okita, K., Osafune, K., Arioka, Y., Maeda, T.*, et al.* (2014). Premature termination of reprogramming in vivo leads to cancer development through altered epigenetic regulation. Cell *156*, 663-677.

Okita, K., Ichisaka, T., and Yamanaka, S. (2007). Generation of germline-competent induced pluripotent stem cells. Nature *448*, 313-317.

Okuda, T., van Deursen, J., Hiebert, S.W., Grosveld, G., and Downing, J.R. (1996). AML1, the target of multiple chromosomal translocations in human leukemia, is essential for normal fetal liver hematopoiesis. Cell *84*, 321-330.

Oostendorp, R.A., Harvey, K.N., Kusadasi, N., de Bruijn, M.F., Saris, C., Ploemacher, R.E., Medvinsky, A.L., and Dzierzak, E.A. (2002). Stromal cell lines from mouse aorta-gonads-mesonephros subregions are potent supporters of hematopoietic stem cell activity. Blood *99*, 1183-1189.

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

Osawa, M., Nakamura, K., Nishi, N., Takahasi, N., Tokuomoto, Y., Inoue, H., and Nakauchi, H. (1996). In vivo self-renewal of c-Kit+ Sca-1+ Lin(low/-) hemopoietic stem cells. Journal of immunology *156*, 3207-3214.

Ouyang, H., Xue, Y., Lin, Y., Zhang, X., Xi, L., Patel, S., Cai, H., Luo, J., Zhang, M., Zhang, M.*, et al.* (2014). WNT7A and PAX6 define corneal epithelium homeostasis and pathogenesis. Nature *511*, 358-361.

Padron-Barthe, L., Temino, S., Villa del Campo, C., Carramolino, L., Isern, J., and Torres, M. (2014). Clonal analysis identifies hemogenic endothelium as the source of the blood-endothelial common lineage in the mouse embryo. Blood *124*, 2523-2532.

Palis, J., Robertson, S., Kennedy, M., Wall, C., and Keller, G. (1999). Development of erythroid and myeloid progenitors in the yolk sac and embryo proper of the mouse. Development *126*, 5073-5084.

Patterson, L.J., Gering, M., Eckfeldt, C.E., Green, A.R., Verfaillie, C.M., Ekker, S.C., and Patient, R. (2007). The transcription factors Scl and Lmo2 act together during development of the hemangioblast in zebrafish. Blood *109*, 2389-2398.

Pearson, S., Cuvertino, S., Fleury, M., Lacaud, G., and Kouskoff, V. (2015). In vivo repopulating activity emerges at the onset of hematopoietic specification during embryonic stem cell differentiation. Stem cell reports *4*, 431-444.

Pearson, S., Sroczynska, P., Lacaud, G., and Kouskoff, V. (2008). The stepwise specification of embryonic stem cells to hematopoietic fate is driven by sequential exposure to Bmp4, activin A, bFGF and VEGF. Development *135*, 1525-1535.

Pencovich, N., Jaschek, R., Tanay, A., and Groner, Y. (2011). Dynamic combinatorial interactions of RUNX1 and cooperating partners regulates megakaryocytic differentiation in cell line models. Blood *117*, e1-14.

Pereira, C.F., Chang, B., Qiu, J., Niu, X., Papatsenko, D., Hendry, C.E., Clark, N.R., Nomura-Kitabayashi, A., Kovacic, J.C., Ma'ayan, A.*, et al.* (2013). Induction of a hemogenic program in mouse fibroblasts. Cell stem cell *13*, 205-218.

Pereira, C.F., Terranova, R., Ryan, N.K., Santos, J., Morris, K.J., Cui, W., Merkenschlager, M., and Fisher, A.G. (2008). Heterokaryon-based reprogramming of human B lymphocytes for pluripotency requires Oct4 but not Sox2. PLoS genetics *4*, e1000170.

Pimanda, J.E., Ottersbach, K., Knezevic, K., Kinston, S., Chan, W.Y., Wilson, N.K., Landry, J.R., Wood, A.D., Kolb-Kokocinski, A., Green, A.R.*, et al.* (2007). Gata2, Fli1, and Scl form a recursively wired gene-regulatory circuit during early hematopoietic development. Proceedings of the National Academy of Sciences of the United States of America *104*, 17692-17697.

Pimkin, M., Kossenkov, A.V., Mishra, T., Morrissey, C.S., Wu, W., Keller, C.A., Blobel, G.A., Lee, D., Beer, M.A., Hardison, R.C.*, et al.* (2014). Divergent functions of hematopoietic transcription factors in lineage priming and differentiation during erythro-megakaryopoiesis. Genome research *24*, 1932-1944.

Pina, C., Fugazza, C., Tipping, A.J., Brown, J., Soneji, S., Teles, J., Peterson, C., and Enver, T. (2012). Inferring rules of lineage commitment in haematopoiesis. Nat Cell Biol *14*, 287-294.

Pishesha, N., Thiru, P., Shi, J., Eng, J.C., Sankaran, V.G., and Lodish, H.F. (2014). Transcriptional divergence and conservation of human and mouse erythropoiesis. Proceedings of the National Academy of Sciences of the United States of America *111*, 4103-4108.

Pop, R., Shearstone, J.R., Shen, Q., Liu, Y., Hallstrom, K., Koulnis, M., Gribnau, J., and Socolovsky, M. (2010). A key commitment step in erythropoiesis is synchronized with the cell cycle clock through mutual inhibition between PU.1 and S-phase progression. PLoS biology *8*.

Priller, J., Persons, D.A., Klett, F.F., Kempermann, G., Kreutzberg, G.W., and Dirnagl, U. (2001). Neogenesis of cerebellar Purkinje neurons from gene-marked bone marrow cells in vivo. The Journal of cell biology *155*, 733-738.

Pulecio, J., Nivet, E., Sancho-Martinez, I., Vitaloni, M., Guenechea, G., Xia, Y., Kurian, L., Dubova, I., Bueren, J., Laricchia-Robbio, L.*, et al.* (2014). Conversion of human fibroblasts into monocyte-like progenitor cells. Stem Cells *32*, 2923-2938.

Qian, L., Huang, Y., Spencer, C.I., Foley, A., Vedantham, V., Liu, L., Conway, S.J., Fu, J.D., and Srivastava, D. (2012). In vivo reprogramming of murine cardiac fibroblasts into induced cardiomyocytes. Nature *485*, 593-598.

Ramsdell, F., Zuniga-Pflucker, J.C., and Takahama, Y. (2006). In vitro systems for the study of T cell development: fetal thymus organ culture and OP9-DL1 cell coculture. Current protocols in immunology / edited by John E Coligan [et al] *Chapter 3*, Unit 3 18.

Ravid, K., Lu, J., Zimmet, J.M., and Jones, M.R. (2002). Roads to polyploidy: the megakaryocyte example. Journal of cellular physiology *190*, 7-20.

Redmer, T., Diecke, S., Grigoryan, T., Quiroga-Negreira, A., Birchmeier, W., and Besser, D. (2011). E-cadherin is crucial for embryonic stem cell pluripotency and can replace OCT4 during somatic cell reprogramming. EMBO reports *12*, 720-726.

Rhodes, J., Hagen, A., Hsu, K., Deng, M., Liu, T.X., Look, A.T., and Kanki, J.P. (2005). Interplay of pu.1 and gata1 determines myelo-erythroid progenitor cell fate in zebrafish. Developmental cell *8*, 97-108.

Rhodes, K.E., Gekas, C., Wang, Y., Lux, C.T., Francis, C.S., Chan, D.N., Conway, S., Orkin, S.H., Yoder, M.C., and Mikkola, H.K. (2008). The emergence of hematopoietic stem cells is initiated in the placental vasculature in the absence of circulation. Cell stem cell *2*, 252-263.

Riddell, J., Gazit, R., Garrison, B.S., Guo, G., Saadatpour, A., Mandal, P.K., Ebina, W., Volchkov, P., Yuan, G.C., Orkin, S.H.*, et al.* (2014). Reprogramming committed murine blood cells to induced hematopoietic stem cells with defined factors. Cell *157*, 549-564.

Robert-Moreno, A., Espinosa, L., de la Pompa, J.L., and Bigas, A. (2005). RBPjkappa-dependent Notch function regulates Gata2 and is essential for the formation of intra-embryonic hematopoietic cells. Development *132*, 1117-1126.

Robin, C., Ottersbach, K., Boisset, J.C., Oziemlak, A., and Dzierzak, E. (2011). CD41 is developmentally regulated and differentially expressed on mouse hematopoietic stem cells. Blood *117*, 5088-5091.

Rodrigues, N.P., Janzen, V., Forkert, R., Dombkowski, D.M., Boyd, A.S., Orkin, S.H., Enver, T., Vyas, P., and Scadden, D.T. (2005). Haploinsufficiency of GATA-2 perturbs adult hematopoietic stem-cell homeostasis. Blood *106*, 477-484.

Roensch, K., Tazaki, A., Chara, O., and Tanaka, E.M. (2013). Progressive specification rather than intercalation of segments during limb regeneration. Science *342*, 1375-1379.

Rybtsov, S., Sobiesiak, M., Taoudi, S., Souilhol, C., Senserrich, J., Liakhovitskaia, A., Ivanovs, A., Frampton, J., Zhao, S., and Medvinsky, A. (2011). Hierarchical organization and early hematopoietic specification of the developing HSC lineage in the AGM region. J Exp Med *208*, 1305-1315.

Saleque, S., Cameron, S., and Orkin, S.H. (2002). The zinc-finger proto-oncogene Gfi-1b is essential for development of the erythroid and megakaryocytic lineages, pp. 301-306.

Samokhvalov, I.M., Samokhvalova, N.I., and Nishikawa, S.-i. (2007). Cell tracing shows the contribution of the yolk sac to adult haematopoiesis. Nature *446*, 1056-1061.

Sandler, V.M., Lis, R., Liu, Y., Kedem, A., James, D., Elemento, O., Butler, J.M., Scandura, J.M., and Rafii, S. (2014). Reprogramming human endothelial cells to haematopoietic cells requires vascular induction. Nature *511*, 312-318.

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.*, et al.* (2013). Platelet-biased stem cells reside at the apex of the haematopoietic stem-cell hierarchy. Nature *502*, 232-236.

Sauvageau, G., Thorsteinsdottir, U., Eaves, C.J., Lawrence, H.J., Largman, C., Lansdorp, P.M., and Humphries, R.K. (1995). Overexpression of HOXB4 in hematopoietic cells causes the selective expansion of more primitive populations in vitro and in vivo. Genes & development *9*, 1753-1765.

Schiedlmeier, B., Santos, A.C., Ribeiro, A., Moncaut, N., Lesinski, D., Auer, H., Kornacker, K., Ostertag, W., Baum, C., Mallo, M.*, et al.* (2007). HOXB4's road map to stem cell expansion. Proceedings of the National Academy of Sciences of the United States of America *104*, 16952-16957.

Schlaeger, T.M., Mikkola, H.K., Gekas, C., Helgadottir, H.B., and Orkin, S.H. (2005). Tie2Cre-mediated gene ablation defines the stem-cell leukemia gene (SCL/tal1)-dependent window during hematopoietic stem-cell development. Blood *105*, 3871-3874.

Schmitt, T.M., and Zuniga-Pflucker, J.C. (2002). Induction of T cell development from hematopoietic progenitor cells by delta-like-1 in vitro. Immunity *17*, 749-756.

Scott, E.W., Simon, M.C., Anastasi, J., and Singh, H. (1994). Requirement of transcription factor PU.1 in the development of multiple hematopoietic lineages. Science *265*, 1573-1577.

Sekiya, S., and Suzuki, A. (2011). Direct conversion of mouse fibroblasts to hepatocyte-like cells by defined factors. Nature *475*, 390-393.

Seppen, J., Rijnberg, M., Cooreman, M.P., and Oude Elferink, R.P.J. (2002). Lentiviral vectors for efficient transduction of isolated primary quiescent hepatocytes. Journal of Hepatology *36*, 459-465.

Shalaby, F., Rossant, J., Yamaguchi, T.P., Gertsenstein, M., Wu, X.F., Breitman, M.L., and Schuh, A.C. (1995). Failure of blood-island formation and vasculogenesis in Flk-1-deficient mice. Nature *376*, 62-66.

Shay, T., Jojic, V., Zuk, O., Rothamel, K., Puyraimond-Zemmour, D., Feng, T., Wakamatsu, E., Benoist, C., Koller, D., Regev, A.*, et al.* (2013). Conservation and divergence in the transcriptional programs of the human and mouse immune systems. Proceedings of the National Academy of Sciences of the United States of America *110*, 2946-2951.

Shen, W.F., Montgomery, J.C., Rozenfeld, S., Moskow, J.J., Lawrence, H.J., Buchberg, A.M., and Largman, C. (1997). AbdB-like Hox proteins stabilize DNA binding by the Meis1 homeodomain proteins. Mol Cell Biol *17*, 6448-6458.

Shi, X., Richard, J., Zirbes, K.M., Gong, W., Lin, G., Kyba, M., Thomson, J.A., Koyano-Nakagawa, N., and Garry, D.J. (2014). Cooperative interaction of Etv2 and Gata2 regulates the development of endothelial and hematopoietic lineages. Developmental biology *389*, 208-218.

Shivdasani, R.A., Mayer, E.L., and Orkin, S.H. (1995). Absence of blood formation in mice lacking the T-cell leukaemia oncoprotein tal-1/SCL. Nature *373*, 432-434.

Silberstein, L., Sanchez, M.J., Socolovsky, M., Liu, Y., Hoffman, G., Kinston, S., Piltz, S., Bowen, M., Gambardella, L., Green, A.R.*, et al.* (2005). Transgenic analysis of the stem cell leukemia +19 stem cell enhancer in adult and embryonic hematopoietic and endothelial cells. Stem Cells *23*, 1378-1388.

Silver, L., and Palis, J. (1997). Initiation of murine embryonic erythropoiesis: a spatial analysis. Blood *89*, 1154-1164.

Slatter, T.L., Ganesan, P., Holzhauer, C., Mehta, R., Rubio, C., Williams, G., Wilson, M., Royds, J.A., Baird, M.A., and Braithwaite, A.W. (2010). p53-mediated apoptosis prevents the accumulation of progenitor B cells and B-cell tumors. Cell death and differentiation *17*, 540-550.

Son, E.Y., Ichida, J.K., Wainger, B.J., Toma, J.S., Rafuse, V.F., Woolf, C.J., and Eggan, K. (2011). Conversion of mouse and human fibroblasts into functional spinal motor neurons. Cell stem cell *9*, 205-218.

Song, K., Nam, Y.J., Luo, X., Qi, X., Tan, W., Huang, G.N., Acharya, A., Smith, C.L., Tallquist, M.D., Neilson, E.G.*, et al.* (2012). Heart repair by reprogramming non-myocytes with cardiac transcription factors. Nature *485*, 599-604.

Sroczynska, P., Lancrin, C., Kouskoff, V., and Lacaud, G. (2009a). The differential activities of Runx1 promoters define milestones during embryonic hematopoiesis. Blood *114*, 5279-5289.

Sroczynska, P., Lancrin, C., Pearson, S., Kouskoff, V., and Lacaud, G. (2009b). In vitro differentiation of mouse embryonic stem cells as a model of early hematopoietic development. Methods in molecular biology *538*, 317-334.

Stopka, T., Amanatullah, D.F., Papetti, M., and Skoultchi, A.I. (2005). PU.1 inhibits the erythroid program by binding to GATA-1 on DNA and creating a repressive chromatin structure. Embo J *24*, 3712-3723.

Sturgeon, C.M., Ditadi, A., Awong, G., Kennedy, M., and Keller, G. (2014). Wnt signaling controls the specification of definitive and primitive hematopoiesis from human pluripotent stem cells. Nature biotechnology *32*, 554-561.

Sumanas, S., Jorniak, T., and Lin, S. (2005). Identification of novel vascular endothelial-specific genes by the microarray analysis of the zebrafish cloche mutants. Blood *106*, 534-541.

Suzuki, M., Kobayashi-Osaki, M., Tsutsumi, S., Pan, X., Ohmori, S., Takai, J., Moriguchi, T., Ohneda, O., Ohneda, K., Shimizu, R.*, et al.* (2013a). GATA factor switching from GATA2 to GATA1 contributes to erythroid differentiation. Genes to cells : devoted to molecular & cellular mechanisms *18*, 921-933.

Suzuki, N., Yamazaki, S., Yamaguchi, T., Okabe, M., Masaki, H., Takaki, S., Otsu, M., and Nakauchi, H. (2013b). Generation of engraftable hematopoietic stem cells from induced pluripotent stem cells by way of teratoma formation. Molecular therapy : the journal of the American Society of Gene Therapy *21*, 1424-1431.

Suzuki, T., Arumugam, P., Sakagami, T., Lachmann, N., Chalk, C., Sallese, A., Abe, S., Trapnell, C., Carey, B., Moritz, T.*, et al.* (2014). Pulmonary macrophage transplantation therapy. Nature *514*, 450-454.

Swiers, G., Baumann, C., O'Rourke, J., Giannoulatou, E., Taylor, S., Joshi, A., Moignard, V., Pina, C., Bee, T., Kokkaliaris, K.D.*, et al.* (2013). Early dynamic fate changes in haemogenic endothelium characterized at the single-cell level. Nature communications *4*, 2924.

Szabo, E., Rampalli, S., Risueno, R.M., Schnerch, A., Mitchell, R., Fiebig-Comyn, A., Levadoux-Martin, M., and Bhatia, M. (2010). Direct conversion of human fibroblasts to multilineage blood progenitors. Nature *468*, 521-526.

Taberlay, P.C., Kelly, T.K., Liu, C.C., You, J.S., De Carvalho, D.D., Miranda, T.B., Zhou, X.J., Liang, G., and Jones, P.A. (2011). Polycomb-repressed genes have permissive enhancers that initiate reprogramming. Cell *147*, 1283-1294.

Takahashi, K., Tanabe, K., Ohnuki, M., Narita, M., Ichisaka, T., Tomoda, K., and Yamanaka, S. (2007). Induction of Pluripotent Stem Cells from Adult Human Fibroblasts by Defined Factors. Cell *131*, 861-872.

Takahashi, K., and Yamanaka, S. (2006). Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. Cell *126*, 663-676.

Tanaka, Y., Hayashi, M., Kubota, Y., Nagai, H., Sheng, G., Nishikawa, S., and Samokhvalov, I.M. (2012a). Early ontogenic origin of the hematopoietic stem cell lineage. Proceedings of the National Academy of Sciences of the United States of America *109*, 4515-4520.

Tanaka, Y., Joshi, A., Wilson, N.K., Kinston, S., Nishikawa, S., and Gottgens, B. (2012b). The transcriptional programme controlled by Runx1 during early embryonic blood development. Developmental biology *366*, 404-419.

Taoudi, S., Gonneau, C., Moore, K., Sheridan, J.M., Blackburn, C.C., Taylor, E., and Medvinsky, A. (2008). Extensive hematopoietic stem cell generation in the AGM region via maturation of VE-cadherin+CD45+ pre-definitive HSCs. Cell stem cell *3*, 99-108.

Taoudi, S., Morrison, A.M., Inoue, H., Gribi, R., Ure, J., and Medvinsky, A. (2005). Progressive divergence of definitive haematopoietic stem cells from the endothelial compartment does not depend on contact with the foetal liver. Development *132*, 4179-4191.

Tavian, M., Hallais, M.F., and Peault, B. (1999). Emergence of intraembryonic hematopoietic precursors in the pre-liver human embryo. Development *126*, 793-803.

Tavian, M., Robin, C., Coulombel, L., and Peault, B. (2001). The human embryo, but not its yolk sac, generates lympho-myeloid stem cells: mapping multipotent hematopoietic cell fate in intraembryonic mesoderm. Immunity *15*, 487-495.

TeKippe, M., Harrison, D.E., and Chen, J. (2003). Expansion of hematopoietic stem cell phenotype and activity in Trp53-null mice. Exp Hematol *31*, 521-527.

Teles, J., Pina, C., Eden, P., Ohlsson, M., Enver, T., and Peterson, C. (2013). Transcriptional regulation of lineage commitment--a stochastic model of cell fate decisions. PLoS computational biology *9*, e1003197.

Themeli, M., Kloss, C.C., Ciriello, G., Fedorov, V.D., Perna, F., Gonen, M., and Sadelain, M. (2013). Generation of tumor-targeted human T lymphocytes from induced pluripotent stem cells for cancer therapy. Nature biotechnology *31*, 928-933.

Thorel, F., Nepote, V., Avril, I., Kohno, K., Desgraz, R., Chera, S., and Herrera, P.L. (2010). Conversion of adult pancreatic alpha-cells to beta-cells after extreme beta-cell loss. Nature *464*, 1149-1154.

Tiemann, U., Sgodda, M., Warlich, E., Ballmaier, M., Scholer, H.R., Schambach, A., and Cantz, T. (2011). Optimal reprogramming factor stoichiometry increases colony numbers and affects molecular characteristics of murine induced pluripotent stem cells. Cytometry Part A : the journal of the International Society for Analytical Cytology *79*, 426-435.

Tijssen, M.R., Cvejic, A., Joshi, A., Hannah, R.L., Ferreira, R., Forrai, A., Bellissimo, D.C., Oram, S.H., Smethurst, P.A., Wilson, N.K.*, et al.* (2011). Genome-wide analysis of simultaneous GATA1/2, RUNX1, FLI1, and SCL binding in megakaryocytes identifies hematopoietic regulators. Developmental cell *20*, 597-609.

Tober, J., Koniski, A., McGrath, K.E., Vemishetti, R., Emerson, R., de Mesy-Bentley, K.K., Waugh, R., and Palis, J. (2007). The megakaryocyte lineage originates from hemangioblast precursors and is an integral component both of primitive and of definitive hematopoiesis. Blood *109*, 1433-1441.

Tsai, F.-Y., Keller, G., Kuo, F.C., Weiss, M., Chen, J., Rosenblatt, M., Alt, F.W., and Orkin, S.H. (1994). An early haematopoietic defect in mice lacking the transcription factor GATA-2. Nature *371*, 221-226.

Tsai, F.-Y., and Orkin, S.H. (1997). Transcription Factor GATA-2 Is Required for Proliferation/Survival of Early Hematopoietic Cells and Mast Cell Formation, But Not for Erythroid and Myeloid Terminal Differentiation. Blood *89*, 3636-3643.

Ueno, H., and Weissman, I.L. (2006). Clonal analysis of mouse development reveals a polyclonal origin for yolk sac blood islands. Developmental cell *11*, 519-533.

Utikal, J., Maherali, N., Kulalert, W., and Hochedlinger, K. (2009a). Sox2 is dispensable for the reprogramming of melanocytes and melanoma cells into induced pluripotent stem cells, pp. 3502-3510.

Utikal, J., Polo, J.M., Stadtfeld, M., Maherali, N., Kulalert, W., Walsh, R.M., Khalil, A., Rheinwald, J.G., and Hochedlinger, K. (2009b). Immortalization eliminates a roadblock during cellular reprogramming into iPS cells. Nature *460*, 1145-1148.

van der Meer, A.D., Orlova, V.V., ten Dijke, P., van den Berg, A., and Mummery, C.L. (2013). Three-dimensional co-cultures of human endothelial cells and embryonic stem cell-derived pericytes inside a microfluidic device. Lab on a chip *13*, 3562-3568.

van der Meer, L.T., Jansen, J.H., and van der Reijden, B.A. (2010). Gfi1 and Gfi1b: key regulators of hematopoiesis. Leukemia *24*, 1834-1843.

Vaskova, E.A., Stekleneva, A.E., Medvedev, S.P., and Zakian, S.M. (2013). "Epigenetic memory" phenomenon in induced pluripotent stem cells. Acta naturae *5*, 15-21.

Vassen, L., Okayama, T., and Moroy, T. (2007). Gfi1b:green fluorescent protein knock-in mice reveal a dynamic expression pattern of Gfi1b during hematopoiesis that is largely complementary to Gfi1. Blood *109*, 2356-2364.

Vierbuchen, T., Ostermeier, A., Pang, Z.P., Kokubu, Y., Sudhof, T.C., and Wernig, M. (2010). Direct conversion of fibroblasts to functional neurons by defined factors. Nature *463*, 1035-1041.

Visvader, J.E., Elefanty, A.G., Strasser, A., and Adams, J.M. (1992). GATA-1 but not SCL induces megakaryocytic differentiation in an early myeloid line. EMBO J *11*, 4557-4564.

Vizcardo, R., Masuda, K., Yamada, D., Ikawa, T., Shimizu, K., Fujii, S., Koseki, H., and Kawamoto, H. (2013). Regeneration of human tumor antigen-specific T cells from iPSCs derived from mature CD8(+) T cells. Cell stem cell *12*, 31-36.

Vousden, K.H., and Lu, X. (2002). Live or let die: the cell's response to p53. Nature reviews Cancer *2*, 594-604.

Wadman, I.A., Osada, H., Grutz, G.G., Agulnick, A.D., Westphal, H., Forster, A., and Rabbitts, T.H. (1997). The LIM-only protein Lmo2 is a bridging molecule assembling an erythroid, DNA-binding complex which includes the TAL1, E47, GATA-1 and Ldb1/NLI proteins. EMBO J *16*, 3145-3157.

Walsh, J.C., DeKoter, R.P., Lee, H.J., Smith, E.D., Lancki, D.W., Gurish, M.F., Friend, D.S., Stevens, R.L., Anastasi, J., and Singh, H. (2002). Cooperative and antagonistic interplay between PU.1 and GATA-2 in the specification of myeloid cell fates. Immunity *17*, 665-676.

Wang, L., Liu, Z., Yin, C., Asfour, H., Chen, O., Li, Y., Bursac, N., Liu, J., and Qian, L. (2015). Stoichiometry of Gata4, Mef2c, and Tbx5 influences the efficiency and quality of induced cardiac myocyte reprogramming. Circulation research *116*, 237-244.

Wang, L., Menendez, P., Shojaei, F., Li, L., Mazurier, F., Dick, J.E., Cerdan, C., Levac, K., and Bhatia, M. (2005). Generation of hematopoietic repopulating cells from human embryonic stem cells independent of ectopic HOXB4 expression. J Exp Med *201*, 1603-1614.

Wang, Q., Stacy, T., Miller, J.D., Lewis, A.F., Gu, T.L., Huang, X., Bushweller, J.H., Bories, J.C., Alt, F.W., Ryan, G.*, et al.* (1996). The CBFbeta subunit is essential for CBFalpha2 (AML1) function in vivo. Cell *87*, 697-708.

Wang, V.E., Schmidt, T., Chen, J., Sharp, P.A., and Tantin, D. (2004). Embryonic lethality, decreased erythropoiesis, and defective octamer-dependent promoter activation in Oct-1-deficient mice. Mol Cell Biol *24*, 1022-1032.

Wapinski, O.L., Vierbuchen, T., Qu, K., Lee, Q.Y., Chanda, S., Fuentes, D.R., Giresi, P.G., Ng, Y.H., Marro, S., Neff, N.F.*, et al.* (2013). Hierarchical mechanisms for direct reprogramming of fibroblasts to neurons. Cell *155*, 621-635.

Wareing, S., Mazan, A., Pearson, S., Gottgens, B., Lacaud, G., and Kouskoff, V. (2012). The Flk1-Cre-mediated deletion of ETV2 defines its narrow temporal requirement during embryonic hematopoietic development. Stem Cells *30*, 1521-1531.

Warren, L., Manos, P.D., Ahfeldt, T., Loh, Y.H., Li, H., Lau, F., Ebina, W., Mandal, P.K., Smith, Z.D., Meissner, A.*, et al.* (2010). Highly efficient reprogramming to pluripotency and directed differentiation of human cells with synthetic modified mRNA. Cell stem cell *7*, 618-630.

Weintraub, H., Tapscott, S.J., Davis, R.L., Thayer, M.J., Adam, M.A., Lassar, A.B., and Miller, A.D. (1989). Activation of muscle-specific genes in pigment, nerve, fat, liver, and fibroblast cell lines by forced expression of MyoD, pp. 5434-5438.

Wernig, M., Meissner, A., Cassady, J.P., and Jaenisch, R. (2008). c-Myc Is Dispensable for Direct Reprogramming of Mouse Fibroblasts. Cell stem cell *2*, 10-12.

Wilmut, I., Schnieke, A.E., McWhir, J., Kind, A.J., and Campbell, K.H.S. (1997). Viable offspring derived from fetal and adult mammalian cells. Nature *385*, 810-813.

Wilson, N.K., Foster, S.D., Wang, X., Knezevic, K., Schutte, J., Kaimakis, P., Chilarska, P.M., Kinston, S., Ouwehand, W.H., Dzierzak, E.*, et al.* (2010a). Combinatorial transcriptional control in blood stem/progenitor cells: genome-wide analysis of ten major transcriptional regulators. Cell stem cell *7*, 532-544.

Wilson, N.K., Timms, R.T., Kinston, S.J., Cheng, Y.H., Oram, S.H., Landry, J.R., Mullender, J., Ottersbach, K., and Gottgens, B. (2010b). Gfi1 expression is controlled by five distinct regulatory regions spread over 100 kilobases, with Scl/Tal1, Gata2, PU.1, Erg, Meis1, and Runx1 acting as upstream regulators in early hematopoietic cells. Mol Cell Biol *30*, 3853-3863.

Xie, H., Ye, M., Feng, R., and Graf, T. (2004). Stepwise Reprogramming of B Cells into Macrophages. Cell *117*, 663-676.

Xu, J., Du, Y., and Deng, H. (2015). Direct lineage reprogramming: strategies, mechanisms, and applications. Cell stem cell *16*, 119-134.

Xue, S.P., Zhang, S.F., Ma, W., Zhang, Z., Liu, P., Zhao, Q., and Han, D. (2006). Erythroid differentiation denucleation factors (EDDFs) function as intrinsic, post-erythropoietin regulators for mammalian erythroid terminal differentiation. Cell proliferation *39*, 61-74.

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

Yamane, T., Washino, A., and Yamazaki, H. (2013). Common developmental pathway for primitive erythrocytes and multipotent hematopoietic progenitors in early mouse development. Stem cell reports *1*, 590-603.

Yang, J., and Weinberg, R.A. (2008). Epithelial-mesenchymal transition: at the crossroads of development and tumor metastasis. Developmental cell *14*, 818-829.

Ye, M., Zhang, H., Amabile, G., Yang, H., Staber, P.B., Zhang, P., Levantini, E., Alberich-Jorda, M., Zhang, J., Kawasaki, A.*, et al.* (2013). C/EBPa controls acquisition and maintenance of adult haematopoietic stem cell quiescence. Nat Cell Biol *15*, 385-394.

Ye, Z., Zhan, H., Mali, P., Dowey, S., Williams, D.M., Jang, Y.Y., Dang, C.V., Spivak, J.L., Moliterno, A.R., and Cheng, L. (2009). Human-induced pluripotent stem cells from blood cells of healthy donors and patients with acquired blood disorders. Blood *114*, 5473-5480.

Yechoor, V., Liu, V., Espiritu, C., Paul, A., Oka, K., Kojima, H., and Chan, L. (2009). Neurogenin3 is sufficient for transdetermination of hepatic progenitor cells into neo-islets in vivo but not transdifferentiation of hepatocytes. Developmental cell *16*, 358-373.

Yoder, M.C., Hiatt, K., Dutt, P., Mukherjee, P., Bodine, D.M., and Orlic, D. (1997). Characterization of Definitive Lymphohematopoietic Stem Cells in the Day 9 Murine Yolk Sac. Immunity *7*, 335-344.

Yokomizo, T., and Dzierzak, E. (2010). Three-dimensional cartography of hematopoietic clusters in the vasculature of whole mouse embryos. Development *137*, 3651-3661.

Yokomizo, T., Ogawa, M., Osato, M., Kanno, T., Yoshida, H., Fujimoto, T., Fraser, S., Nishikawa, S., Okada, H., Satake, M.*, et al.* (2001). Requirement of Runx1/AML1/PEBP2alphaB for the generation of haematopoietic cells from endothelial cells. Genes to cells : devoted to molecular & cellular mechanisms *6*, 13-23.

Yokomizo, T., Takahashi, S., Mochizuki, N., Kuroha, T., Ema, M., Wakamatsu, A., Shimizu, R., Ohneda, O., Osato, M., Okada, H.*, et al.* (2007). Characterization of GATA-1(+) hemangioblastic cells in the mouse embryo. EMBO J *26*, 184-196.

Yoshimoto, M., Montecino-Rodriguez, E., Ferkowicz, M.J., Porayette, P., Shelley, W.C., Conway, S.J., Dorshkind, K., and Yoder, M.C. (2011). Embryonic day 9 yolk sac and intra-embryonic hemogenic endothelium independently generate a B-1 and marginal zone progenitor lacking B-2 potential. Proceedings of the National Academy of Sciences of the United States of America *108*, 1468-1473.

Yoshimoto, M., Porayette, P., Glosson, N.L., Conway, S.J., Carlesso, N., Cardoso, A.A., Kaplan, M.H., and Yoder, M.C. (2012). Autonomous murine T-cell progenitor production in the extra-embryonic yolk sac before HSC emergence. Blood *119*, 5706-5714.

Yu, B., He, Z.Y., You, P., Han, Q.W., Xiang, D., Chen, F., Wang, M.J., Liu, C.C., Lin, X.W., Borjigin, U.*, et al.* (2013). Reprogramming fibroblasts into bipotential hepatic stem cells by defined factors. Cell stem cell *13*, 328-340.

Yu, J., Vodyanik, M.A., Smuga-Otto, K., Antosiewicz-Bourget, J., Frane, J.L., Tian, S., Nie, J., Jonsdottir, G.A., Ruotti, V., Stewart, R.*, et al.* (2007). Induced Pluripotent Stem Cell Lines Derived from Human Somatic Cells, pp. 1917-1920.

Yucel, R., Kosan, C., Heyd, F., and Moroy, T. (2004). Gfi1:green fluorescent protein knock-in mutant reveals differential expression and autoregulation of the growth factor independence 1 (Gfi1) gene during lymphocyte development. The Journal of biological chemistry *279*, 40906-40917.

Zafonte, B.T., Liu, S., Lynch-Kattman, M., Torregroza, I., Benvenuto, L., Kennedy, M., Keller, G., and Evans, T. (2007). Smad1 expands the hemangioblast population within a limited developmental window. Blood *109*, 516-523.

Zeng, H., Yucel, R., Kosan, C., Klein-Hitpass, L., and Moroy, T. (2004). Transcription factor Gfi1 regulates self-renewal and engraftment of hematopoietic stem cells. EMBO J *23*, 4116-4125.

Zhang, P., Zhang, X., Iwama, A., Yu, C., Smith, K.A., Mueller, B.U., Narravula, S., Torbett, B.E., Orkin, S.H., and Tenen, D.G. (2000). PU.1 inhibits GATA-1 function and erythroid differentiation by blocking GATA-1 DNA binding. Blood *96*, 2641-2648.

Zhou, Q., Brown, J., Kanarek, A., Rajagopal, J., and Melton, D.A. (2008). In vivo reprogramming of adult pancreatic exocrine cells to beta-cells. Nature *455*, 627-632.

Zhu, M.X., Wan, W.L., Li, H.S., Wang, J., Chen, G.A., and Ke, X.Y. (2015). Thymopentin enhances the generation of T-cell lineage derived from human embryonic stem cells in vitro. Exp Cell Res *331*, 387-398.

Zou, J., Mali, P., Huang, X., Dowey, S.N., and Cheng, L. (2011a). Site-specific gene correction of a point mutation in human iPS cells derived from an adult patient with sickle cell disease. Blood *118*, 4599-4608.

Zou, J., Sweeney, C.L., Chou, B.K., Choi, U., Pan, J., Wang, H., Dowey, S.N., Cheng, L., and Malech, H.L. (2011b). Oxidase-deficient neutrophils from X-linked chronic granulomatous disease iPS cells: functional correction by zinc finger nuclease-mediated safe harbor targeting. Blood *117*, 5561-5572.

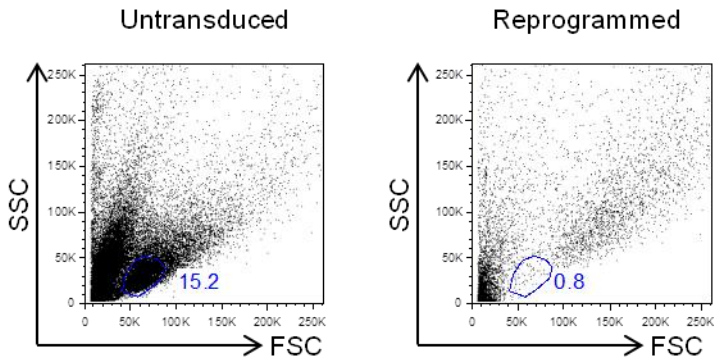
Zovein, A.C., Hofmann, J.J., Lynch, M., French, W.J., Turlo, K.A., Yang, Y., Becker, M.S., Zanetta, L., Dejana, E., Gasson, J.C.*, et al.* (2008). Fate tracing reveals the endothelial origin of hematopoietic stem cells. Cell stem cell *3*, 625-636.

Zufferey, R., Dull, T., Mandel, R.J., Bukovsky, A., Quiroz, D., Naldini, L., and Trono, D. (1998). Self-inactivating lentivirus vector for safe and efficient in vivo gene delivery. J Virol *72*, 9873-9880.

Appendices

Appendix 1

A)



B)

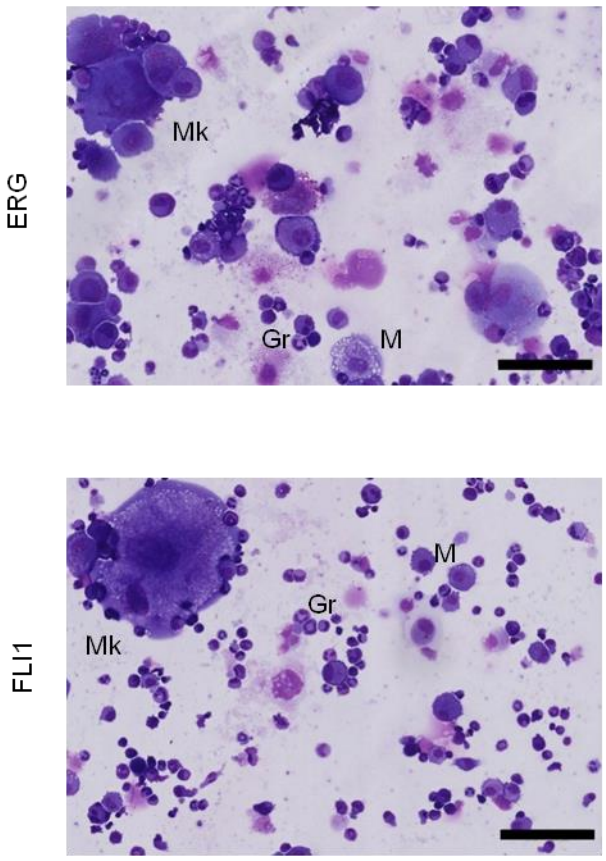


Figure S1. Related to chapter 3.

A) Flow cytometry profiles of control and 19-factor reprogrammed cells at day 21 of experiment. **B)** Representative low magnification images of blood cells generated from MAFs reprogrammed with a use of ERG or FLI1 in the five-TFs lentiviral mix. Scale bars represent 100μm. FSC – forward scatter, SSC – side scatter, M- macrophage, Gr – granulocyte, Mk – megakaryocyte.

Appendix 2

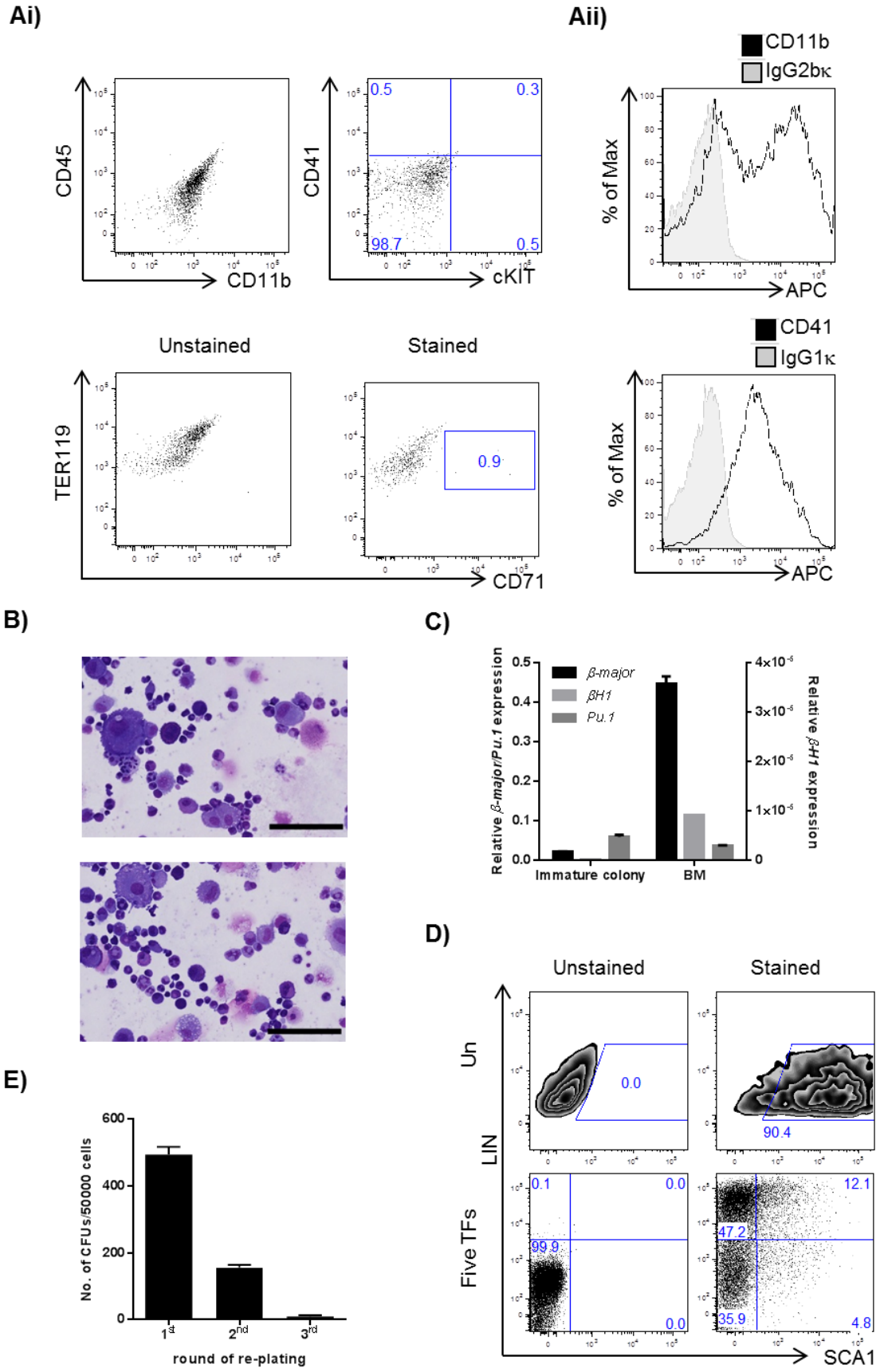

Ai)
Aii)
CD45
CD11b
CD41
cKIT
0.5
0.3
98.7
0.5
CD11b
IgG2bκ
% of Max
APC
Unstained
Stained
TER119
CD71
0.9
CD41
IgG1κ
% of Max
APC
B)
C)
Relative β-major/Pu.1 expression
Relative βH1 expression
β-major
βH1
Pu.1
Immature colony
BM
D)
Unstained
Stained
Un
LIN
Five TFs
SCA1
0.0
90.4
0.1
0.0
99.9
0.0
12.1
47.2
35.9
4.8
E)
No. of CFUs/50000 cells
round of re-plating
1st
2nd
3rd

Figure S2. Related to chapter 4.

A) Representative flow cytometry plots of day 21 control untransduced MAFs **(i)** and reprogrammed fibroblasts **(ii)** stained for haematopoietic surface markers **(i, ii)** and relevant isotype controls **(ii)**. **B)** Low magnification images of MAF-derived blood cells at day 21 of reprogramming. **C)** Relative to *β–actin* expression levels of myeloid and erythroid genes in immature CFU colonies derived from reprogrammed MAFs. Data presented are mean ± SD of a representative experiment performed in triplicate (n=2). **D)** Day 21 reprogrammed (Five TFs) and control (Un) fibroblasts stained for SCA1 and LIN markers **E)** Day 21 reprogrammed MEFs were re-plated into CFU-C assay at 50000 cells per dish. After a week blood progenitor-derived colonies were counted and cells harvested and re-plated into a new semi-solid culture. This procedure was repeated until cells were exhausted (round 3). Data presented are mean ± SEM of a single experiment performed in triplicate. Scale bars represent 100 μm.

Appendix 3

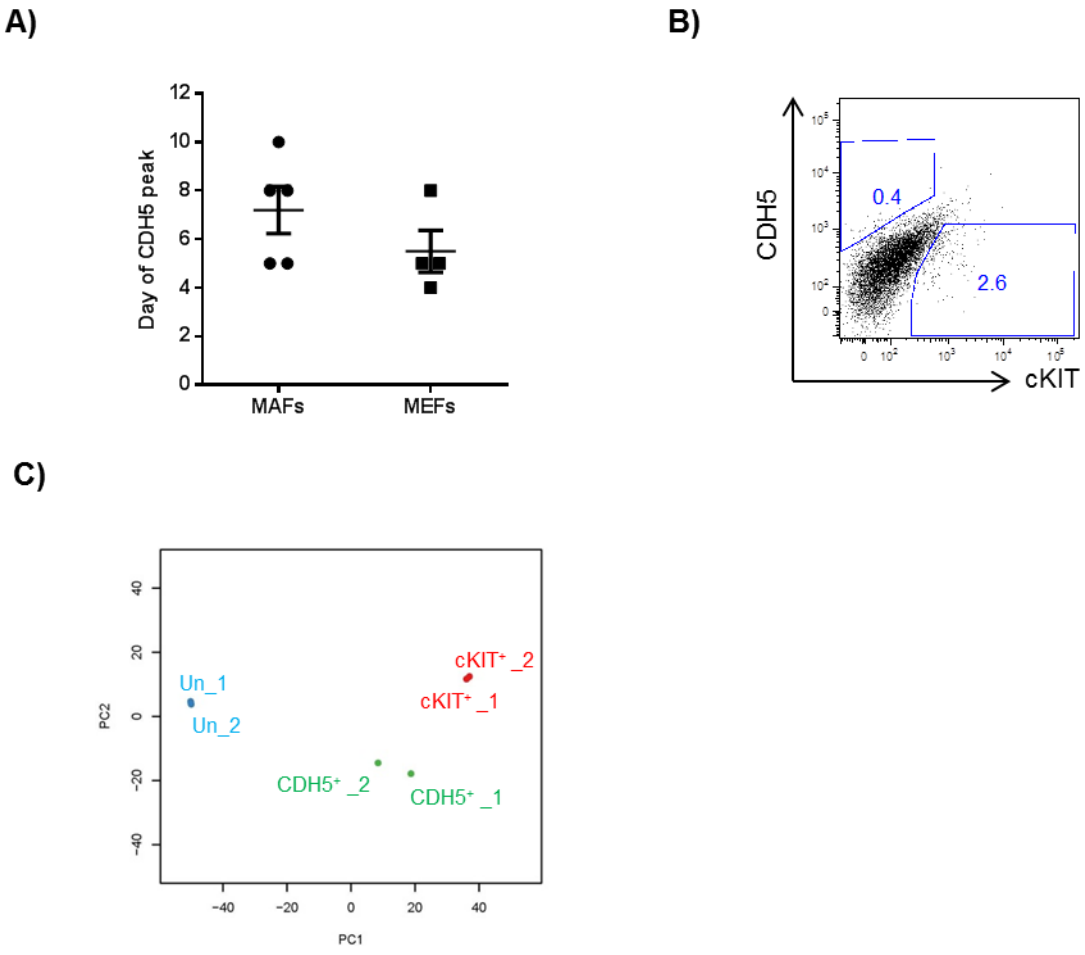


Figure S3. Related to chapter 5.

A) Distribution of peak of CDH5$^+$ cells when tracked by flow cytometry between day 4 and 12 of reprogramming culture (mean ± SEM, n=4 (MEFs), n=5 (MAFs)). **B)** Sorting strategy for re-aggregation assay of CDH5$^+$ cells **C)** PCA of differentially expressed genes among control (Un), CDH5$^+$ and cKIT$^+$ fractions of reprogrammed MEFs.

Appendix 4

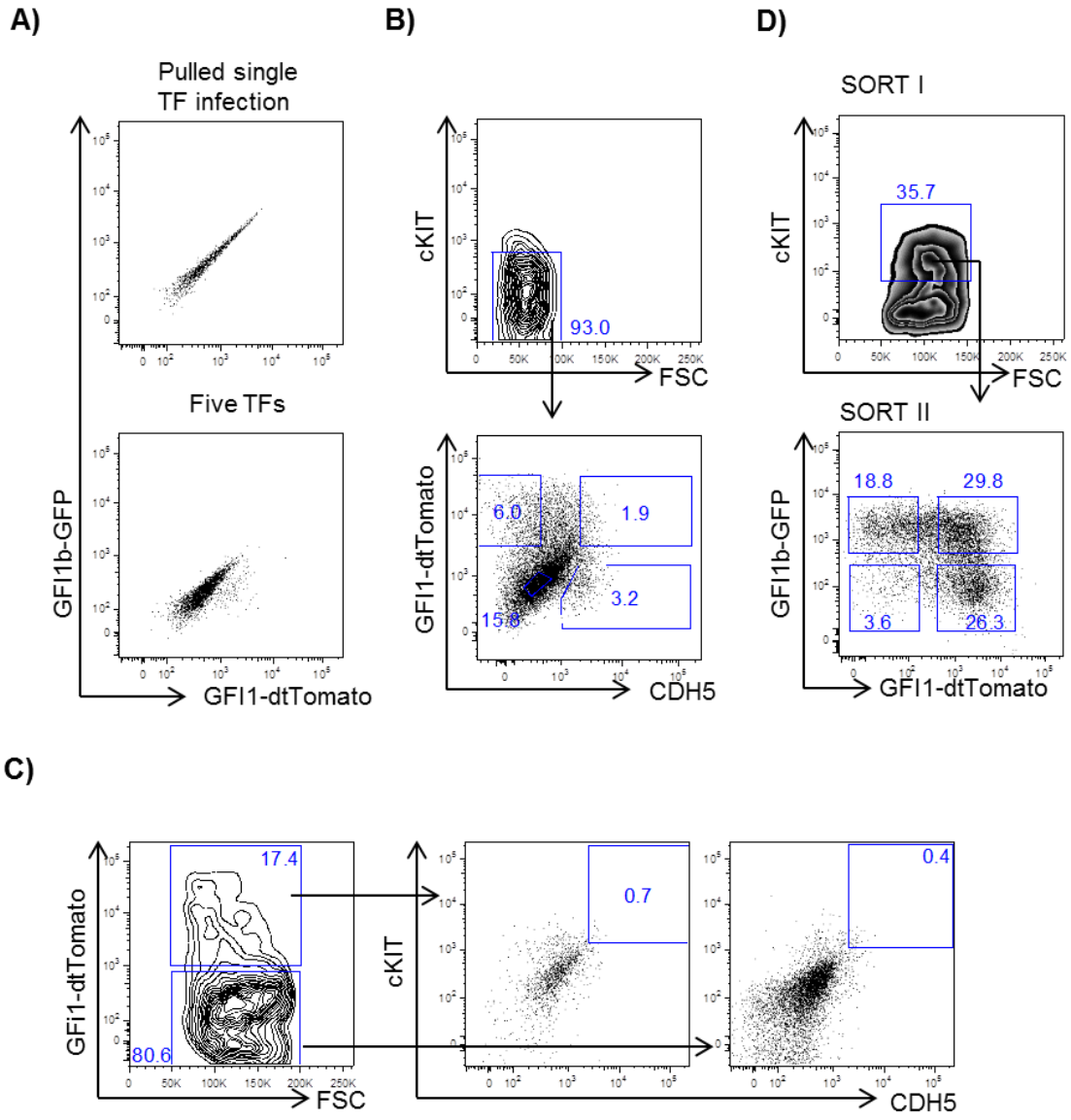


Figure S4. Related to chapter 6.

A) Representative histograms showing dtTomato and GFP reporter acquisition in day 4 pulled single-factor or five TFs trandsuced MEFs **B)** Sorting strategy employed for cKIT^{-}/CDH5$^{+/-}$/GFI1$^{+/-}$ re-aggregation assay. **C)** Sorting strategy utilized for cKIT^{+}/CDH5^{+}/GFI1$^{+/-}$ re-aggregation assay. **D)** Sorting strategy for subfractioning day 12 cKIT^{+} cells on the basis of their *Gfis* expression. Mac – macrophagic CFU, GM – granulocytic-macrophagic CFU.

Appendix 5

Movie S1: Time-lapse imaging of reprogramming MEF-culture between day 3 and 13 of experiment. Fibroblasts were transduced with five haematopoietic TFs (SCL, LMO2, RUNX1c, GATA2, ERG) and cultured in haematopoietic medium. Scanning was performed with IncuCyte FLR and images were processed using Final Cut Pro software.

Appendix 6

Movie S2: Time-lapse imaging of reprogramming MSBB MEF-(*Gfi1*:dtTomato, *Gfi1b*:GFP) culture between day 3 and 12 of experiment. Upon transduction with five haematopoietic TFs (SCL, LMO2, ERG, GATA2, RUNX1c) cells were cultured in haematopoietic medium supplemented with cytokines. GFP (green channel) allows to track expression of *Gfi1b* whereas presence of red fluorochrome is indicative of *Gfi1* expression. Image capture was performed with IncuCyte ZOOM. Data were processed using Final Cut Pro software.

The disk containing Appendices 5 and 6 is enclosed

Appendix 7

Cell Reports
Article

OPEN ACCESS
CellPress

Direct Reprogramming of Murine Fibroblasts to Hematopoietic Progenitor Cells

Kiran Batta,[1,3] Magdalena Florkowska,[1,3] Valerie Kouskoff,[2] and Georges Lacaud[1,*]
[1]CRUK Stem Cell Biology Group, Cancer Research UK Manchester Institute, The University of Manchester, Wilmslow Road, Manchester M20 4BX, UK
[2]CRUK Stem Cell Haematopoiesis Group, Cancer Research UK Manchester Institute, The University of Manchester, Wilmslow Road, Manchester M20 4BX, UK
[3]Co-first author
*Correspondence: georges.lacaud@cruk.manchester.ac.uk
http://dx.doi.org/10.1016/j.celrep.2014.11.002

SUMMARY

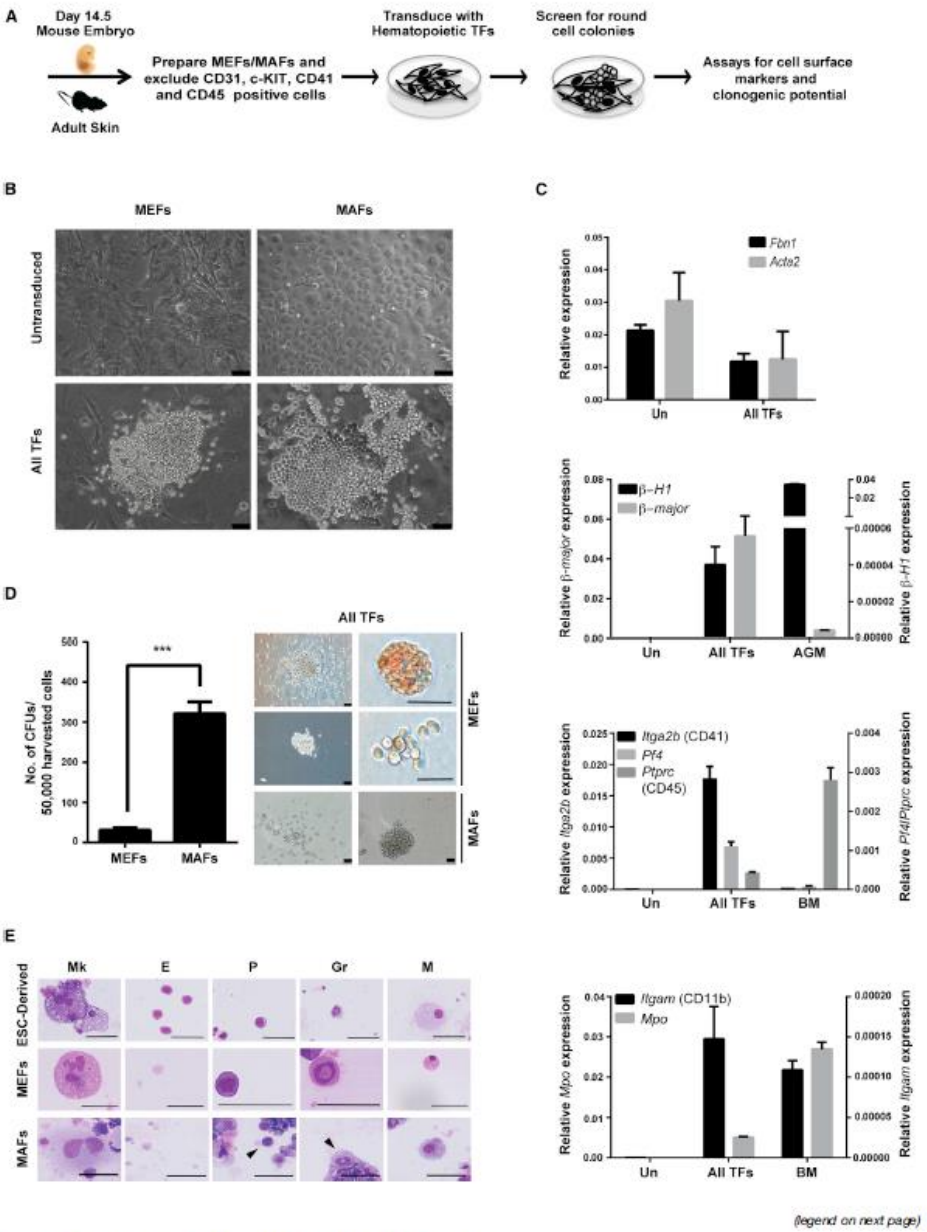
Recent reports have shown that somatic cells, under appropriate culture conditions, could be directly reprogrammed to cardiac, hepatic, or neuronal phenotype by lineage-specific transcription factors. In this study, we demonstrate that both embryonic and adult somatic fibroblasts can be efficiently reprogrammed to clonal multilineage hematopoietic progenitors by the ectopic expression of the transcription factors ERG, GATA2, LMO2, RUNX1c, and SCL. These reprogrammed cells were stably expanded on stromal cells and possessed short-term reconstitution ability in vivo. Loss of p53 function facilitated reprogramming to blood, and $p53^{-/-}$ reprogrammed cells efficiently generated erythroid, megakaryocytic, myeloid, and lymphoid lineages. Genome-wide analyses revealed that generation of hematopoietic progenitors was preceded by the appearance of hemogenic endothelial cells expressing endothelial and hematopoietic genes. Altogether, our findings suggest that direct reprogramming could represent a valid alternative approach to the differentiation of embryonic stem cells (ESCs) or induced pluripotent stem cells (iPSCs) for disease modeling and autologous blood cell therapies.

INTRODUCTION

Until recently, it was assumed that differentiation was mostly a unidirectional and irreversible route that cells undertake during lineage commitment. This dogma was rebutted by the groundbreaking discovery of Yamanaka and colleagues that the expression of four transcription factors (TFs) could reprogram mouse and human cells into a pluripotent stage (Takahashi and Yamanaka, 2006; Takahashi et al., 2007). Subsequent studies have established that cellular-fate conversion is also obtained by direct transdifferentiation between two distinct lineages. Transdifferentiation is generally achieved by overexpressing lineage-instructive TFs, as demonstrated by the effective cell-fate switching of fibroblasts into neuronal, hepatocyte, and cardiomyocyte lineages (Du et al., 2014; Huang et al., 2014; Ieda et al., 2010; Nam et al., 2013; Sekiya and Suzuki, 2011; Vierbuchen et al., 2010).

The hematopoietic system relies on the existence of a rare population of hematopoietic stem cells (HSCs) that are able to self-renew and reconstitute the entire system by generating all hematopoietic lineages. In the clinic, transfusion of HSCs and terminally differentiated blood cells (erythrocytes, platelets, and granulocytes) is used to successfully treat blood genetic disorders and malignancies. However, a major restriction to the wider use of these treatments is the limited availability of cells from donors with adequate match. An alternative strategy for the generation of patient-specific hematopoietic cells would be to differentiate induced pluripotent stem cells (iPSCs) to HSCs. Unfortunately, so far, the development of robust methods to produce blood cells and, in particular, transplantable long-term HSCs has met with limited success (Blum and Benveniste, 2008; Sturgeon et al., 2013). Therefore, direct reprogramming of patient-derived cells by transdifferentiation represents an attractive alternative strategy for the generation of transplantable blood cells.

Hematopoiesis is governed by the combined functions of numerous TFs, complicating attempts to establish simple approaches toward transdifferentiation into this lineage. This complexity is highlighted by knockout studies that identified multiple regulators of blood cell generation including SCL, RUNX1, ERG, and GATA2 (Loughran et al., 2008; Okuda et al., 1996; Robb et al., 1996; Tsai et al., 1994). Genome-wide chromatin immunoprecipitation data indicated that these four factors, in conjunction with LMO2, LYL1, and FLI1, create a regulatory complex that mediates transcription of multiple genes in hematopoietic progenitor cells (Wilson et al., 2010). Each TF of this heptad has been shown to act at multiple stages of hematopoietic specification, maturation, and differentiation (Loose et al., 2007). For example, SCL is required during the formation of hemogenic endothelium precursors from hemangioblast and mesoderm (Lancrin et al., 2010). RUNX1 is critical for the emergence of hematopoietic progenitors and HSCs from hemogenic endothelium (Chen et al., 2009; Lancrin et al., 2010). ERG is required for the maintenance of fetal HSCs and also for the



(legend on next page)

self-renewal and survival of adult HSCs (Loughran et al., 2008; Taoudi et al., 2011). Synergistic, antagonistic, and sequential relationships among these TFs create complex regulatory landscapes that shape the hematopoietic identity (Pimanda and Göttgens, 2010).

Previous studies have revealed an inherent plasticity of hematopoietic cells, as they are amenable to transdifferentiation and dedifferentiation. Recently, this approach has been remarkably employed for the generation of inducible HSCs (iHSCs) by reprogramming blood cells or endothelium (Riddell et al., 2014; Sandler et al., 2014). In both studies, cells capable of multilineage long-term engraftment were obtained by transient ectopic expression of TFs selectively expressed in HSPCs. Importantly, the successful generation of iHSCs required provision of a favorable niche for the maturation of the cells in the form of either the in vivo bone marrow environment or a vascular support mimicking the aorta-gonad-mesonephros (AGM) niche. However, reprogramming of differentiated blood cells might not be suitable for the generation of healthy transplantable cells for patients with blood malignancies or acquired genetic diseases (Pereira et al., 2014). In addition, it could prove very difficult to obtain enough endothelial cells from an adult patient to perform reprogramming. An approach more appropriate for this purpose, but also more challenging, would be to reprogram more developmentally distinct cell types, such as fibroblasts, into blood. In this context, it has been shown that the TFs PU.1 and cEBPα were capable of reprogramming fibroblasts into mature macrophage-like cells (Feng et al., 2008). However, reprogramming of fibroblasts to more immature hematopoietic progenitors has so far remained challenging. The first attempt, reported in 2010, involved the ectopic expression of the pluripotency factor OCT4 in human fibroblasts (Szabo et al., 2010). Although OCT4-induced hematopoietic progenitors exhibited myeloid and erythroid potential, lymphoid potential and long-term in vivo engraftment capacity were not achieved. Moreover, the use of the pluripotent TF OCT4 also raises concerns about tumorigenicity, as partial reprogramming induced by OCT4 could mimic a neoplastic state (Ohnishi et al., 2014). More recently, the combinatorial expression of a limited set of hematopoietic TFs, including GATA2, GFI1b, ETV6, and c-FOS in fibroblasts, was shown to induce a hemogenic endothelial cell fate (Pereira et al., 2013). However, the reprogrammed cells subsequently displayed only minimal hematopoietic potential despite coculture with placental cells.

In this report, we establish that fibroblasts can be rapidly and robustly reprogrammed to hematopoietic progenitors. The ectopic expression of five hematopoietic TFs, functionally selected from a range of 19 regulators, reproducibly induced hematopoietic fate in adult and embryonic fibroblasts within 8 days. The reprogrammed progenitors exhibited multilineage clonogenic capacity in vitro. Genome-wide transcriptional analyses of the reprogrammed cells revealed that the generation of hematopoietic progenitors was preceded by a hemogenic endothelial stage.

RESULTS

Reprogramming of Fibroblasts to Hematopoietic Fate with a Pool of Hematopoietic TFs

The aim of the present study was to investigate whether the ectopic expression of specific hematopoietic TFs could result in a direct and rapid cell fate conversion to hematopoietic precursors. First, we carefully selected a set of 19 different hematopoietic TFs based on their expression and function during hematopoiesis (Table S1). The expression of most of the selected TFs spans from the onset of blood development during ontogeny to the adult blood system (Orkin and Zon, 2008). Lentiviruses expressing each individual factor were prepared and a cocktail of all TFs was used to infect primary fibroblasts. As a source of starting material, we used either embryonic day 14.5 (E14.5) mouse embryonic fibroblasts (MEFs) or mouse adult ear skin fibroblasts (MAFs). Prior to infection, these cells were depleted of any CD41-, CD31-, c-KIT-, and CD45-positive cells to eliminate potential contamination by hematopoietic and endothelial cells (Figure 1A). Posttransduction, the infected cells were switched to a media supporting the growth of hematopoietic cells, and the appearance of colonies containing round cells was monitored daily. Starting from day 8, we observed in transduced MEFs and MAFs cultures, but not in untransduced control cells, the emergence of colonies of small round cells often associated with cobblestone-like areas (Figure 1B). These colonies continued to expand during the following weeks of culture.

The transduced cells were next tested by immunostaining for the acquisition of the hematopoietic cell-surface markers c-KIT and CD41, which are expressed on emerging blood cells during embryonic development (Mikkola et al., 2003; Mitjavila-Garcia et al., 2002), and for the expression of hematopoietic genes. Live staining at day 12 indicated that cells in emerging colonies

Figure 1. Screen for Hematopoiesis-Inducing TFs

(A) Schematic representation of experimental strategy. Murine embryonic fibroblasts (MEFs) were prepared from day 14.5 embryos, and murine adult fibroblasts (MAFs) were prepared from adult ear skin. Cells expressing the surface markers CD31, CD41, c-KIT, and CD45 were excluded from the starting populations. Sorted cells were transduced with a cocktail of all TFs. After 21 days of culture, cells were analyzed for hematopoietic cell-surface markers, clonogenic capacity, and cellular and nuclear morphology.

(B) Bright-field images of untransduced and all TFs transduced MEFs and MAFs at day 12.

(C) Relative gene expression levels of indicated genes with respect to β-actin in untransduced (Un), all-TF-transduced MEFs, control bone marrow (BM), and control E10.5 aorta-gonad-mesonephros (AGM) region cells. Data presented are representative of one out of three independent experiments performed in triplicate ($n = 3$; mean ± SD).

(D) Number of hematopoietic colonies generated by 50,000 all-TF-transduced day 21 harvested MEFs/MAFs (left). Mean ± SEM from two independent experiments performed in triplicates is shown ($n = 2$). Representative bright-field images of the different types of colonies observed (right).

(E) Cellular morphology of day 21 transduced MEFs/MAFs and ESC-derived hematopoietic cells. Arrowheads depict indicated morphologies.

Gr, granulocyte; M, macrophage; P, progenitor; E, erythrocyte; Mk, megakaryocyte. Scale bars represent 50 μm. Asterisks indicate significant differences (Student's t test; ***$p < 0.0005$).

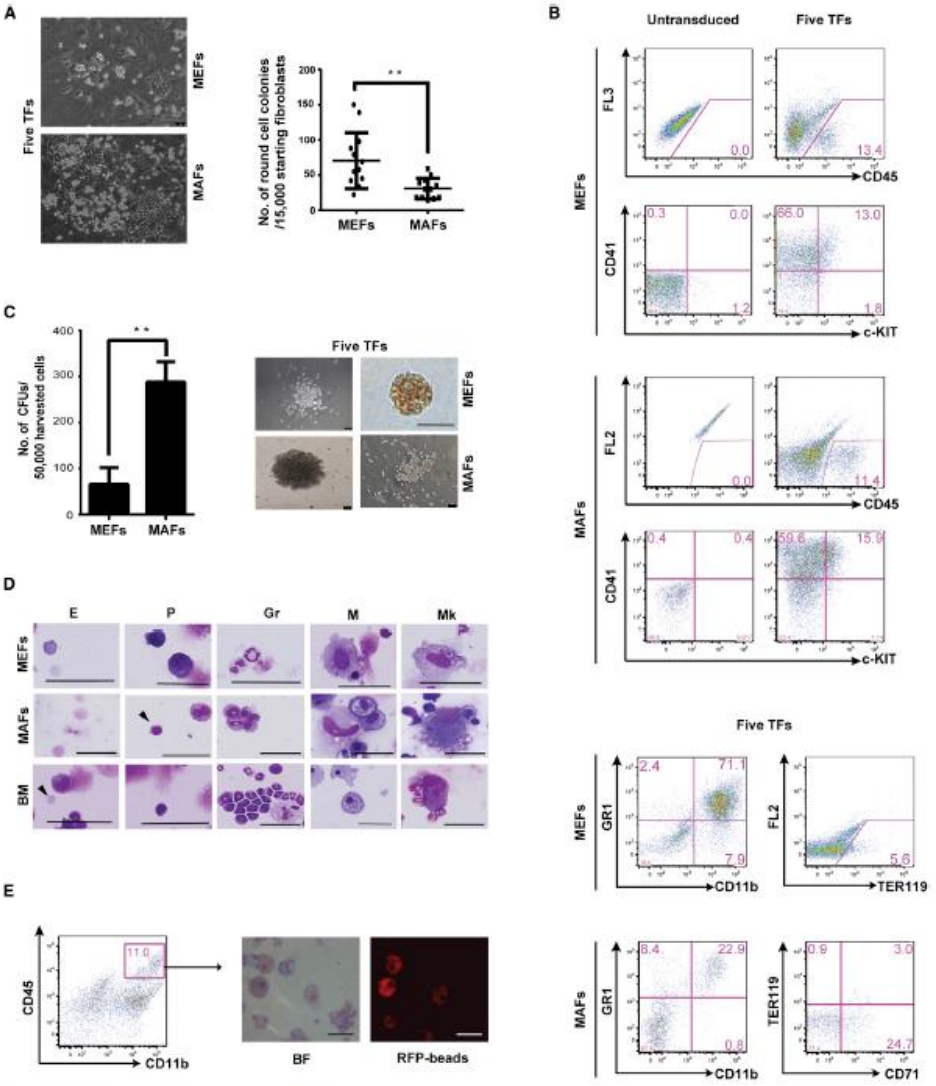


Figure 2. Five TFs Induce Reprogramming to Blood

(A) Bright-field images of five TFs transduced MEFs/MAFs at day 12 (left). Number of round cell colonies observed per 15,000 transduced MEFs/MAFs (right, n = 5, mean ± SEM).

(B) FACS analysis of untransduced and five-TF-transduced MEFs and MAFs at day 21.

(legend continued on next page)

were positive for CD41 expression (Figure S1A). Flow cytometry (fluorescence-activated cell sorting [FACS]) analyses at day 21 confirmed the acquisition of CD41 and low levels of c-KIT expression by the transduced MEFs and MAFs (Figure S1B). Gene expression analyses on 3-weeks-posttransduced cells clearly indicated the downregulation of the fibroblasts markers (*Acta2* and *Fbn1*) (Figure 1C). In contrast, the expression of genes associated with erythroid (embryonic *β-H1* and adult *β-major* hemoglobin), megakaryocytic (*Pf4*), and myeloid (*Itgam* and *Mpo*) lineages were markedly upregulated in these cultures (Figure 1C). The expression of both *β-H1* and *β-major* hemoglobin genes suggests the presence of both primitive and definitive erythroid cells in these cultures. Having demonstrated the induction of a hematopoietic gene signature, we next sought to determine whether the reprogrammed cells possessed functional hematopoietic clonogenic potential. Day 21 reprogrammed MEFs and MAFs, but not untransduced fibroblasts, generated hematopoietic colonies containing macrophage, erythroid, and granulocytic cells upon replating in clonogenic assays (Figure 1D). May-Grünwald Giemsa staining confirmed the presence of cells with erythroid, megakaryocytic, and myeloid morphologies in these colonies (Figure 1E). Collectively, these experiments established that both embryonic and adult fibroblasts could be reprogrammed to hematopoietic cells upon the ectopic expression of 19 hematopoietic TFs.

ERG, GATA2, LMO2, RUNX1c, and SCL Reprogram Fibroblasts to Blood

Our next aim was to define the minimal combination of TFs required for reprogramming to blood cells. By eliminating each TF individually from the pool, taking into account the redundancy among TFs, and reiteration, we established that a minimal set of five TFs (ERG, GATA2, LMO2, RUNX1c, and SCL) was able to robustly and reproducibly induce the generation of colonies of round cells in both adult and embryonic fibroblasts (Figure 2A, left; Movie S1). The efficiency of reprogramming was consistently higher in MEFs than in MAFs (Figure 2A, right). Several other combinations of TFs could also induce the generation of hematopoietic colonies from MEFs, albeit with lower efficiencies. The combination of ERG, FLI1, GATA2, PU.1, and SCL or the triad of FLI1, GATA2, and SCL, which controls the specification of mammalian hematopoietic progenitors (Pimanda et al., 2007), induced the generation of hematopoietic colonies in MEFs (Figure S2A). However, as the combination of the five TFs efficiently and reproducibly reprogrammed both MEFs and MAFs, we selected this set for further experiments. FACS analyses of day 21 MEF and MAF transduced cultures confirmed the acquisition of the hematopoietic markers c-KIT, CD41, CD45, CD11b, and TER119 (Figure 2B). Clonogenic assays and cytospin analysis of day 21 reprogrammed cells confirmed that five-TF-transduced MEFs and MAFs exhibited erythroid and myeloid potential (Figures 2C and 2D). Finally, sorted CD45/CD11b double-positive cells were able to uptake red fluorescent latex beads, demonstrating their phagocytic capacity and functionality (Figure 2E).

We next set out to determine the specific requirement for each TF and therefore performed "*N* minus 1" experiments. In MAFs, SCL, LMO2, and RUNX1c were more important than ERG and GATA2 for the generation of hematopoietic colonies (Figure S2B). In MEFs, SCL and LMO2 alone were sufficient, albeit at a lower efficiency, to generate hematopoietic colonies containing CD41-, CD45-, and c-KIT-positive cells (Figure S2C); in contrast, no colonies were obtained with these two TFs in MAFs (Figure S2C). To investigate if the hematopoietic phenotype of the reprogrammed cells was dependent on the sustained expression of the exogenous TFs, we performed quantitative RT-PCR (qRT-PCR) on day 4 and day 21 transduced MEFs with primers specific for exogenous transduced factors. These experiments demonstrated that the vector-driven transcription of all five TFs was silenced by day 21 (Figure S2D). In contrast, vector integration of all five viruses was enriched upon the emergence and expansion of blood cells in these cultures (Figure S2E). These results established the contribution of all five exogenous factors in the induction of the hematopoietic program. In addition, qRT-PCR specific for the endogenous genes indicated that expression of all five endogenous TFs was induced during reprogramming (Figure S2F). Finally, to determine the growth factors and cytokines essential to reprogram fibroblasts, we performed "*N* minus 1" experiments with MEFs. We observed that interleukin-3 (IL-3) was absolutely critical for reprogramming (Figure S2G), a finding consistent with its established role in the maintenance of hematopoietic progenitor cells. Collectively, our results indicate that the combination of ERG, GATA2, LMO2, RUNX1c, and SCL efficiently induces reprogramming of both embryonic and adult fibroblasts to blood.

Five-TF-Induced Reprogramming Generates Multipotent Progenitors

Both morphological analysis and cell-surface staining data suggested that a large fraction of day 21 reprogrammed hematopoietic cells were already mature and differentiated at this stage (Figure 2D). To investigate whether fibroblasts were reprogrammed through a transient hematopoietic progenitor stage, or transdifferentiated more directly to mature blood cells, we investigated the emergence of cells positive for hematopoietic precursor markers by FACS. We observed from day 10 onward a sudden rise in the frequency of c-KIT/CD41 double-positive cells followed by a similar rise in the emergence of $CD45^+$ cells from day 12 (Figure 3A). Time-course replating of transduced MEF cultures in semisolid colony assays indicated a transient peak of clonogenic potential around day 12–15 (Figure 3B), which correlated with higher frequencies of $c\text{-}KIT^+$ cells. A similar gradual increase in the number of $c\text{-}KIT^+$ cells and peak of clonogenic potential was observed in transduced MAF cultures (Figures S3A and S3B).

(C) Number of hematopoietic colonies generated by 50,000 five-TF-transduced day 21 harvested MEFs/MAFs (left, n = 4, mean ± SEM). Representative bright-field images of the different types of colonies observed (right).
(D) Cellular morphology analyses of day 21 transduced MEFs/MAFs and control bone marrow (BM) derived cells. Arrowheads depict indicated morphology. Gr, granulocyte; M, macrophage; P, progenitor; E, erythrocyte; Mk, megakaryocyte. Scale bars represent 50 μm.
(E) Phagocytic capacity of CD45/CD11b double-positive cells. Asterisks indicate significant differences (Student's *t* test; $^{**}p < 0.01$).

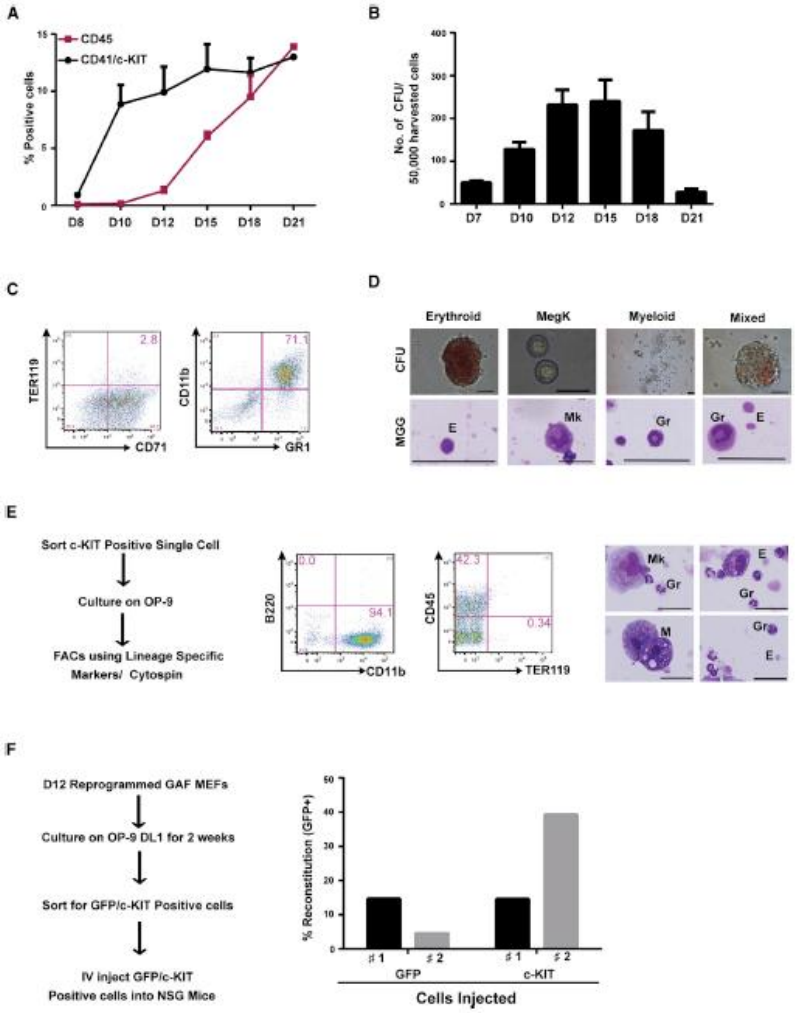


Figure 3. Multilineage Potential of Five-TF-Reprogrammed Cells

(A) Acquisition of hematopoietic cell surface markers during the course of reprogramming. Average percentages of cells expressing CD41, c-KIT, and CD45 are represented at indicated intervals after five-TF transduction of MEFs ($n = 2$, performed in duplicate; mean ± SD).

(B) Number of hematopoietic colonies generated by 50,000 reprogrammed MEFs from day 7 to day 21 ($n = 2$, performed in duplicate; mean ± SEM).

(C and D) Multilineage potential of five TFs reprogrammed day 12 sorted c-KIT$^+$ cells. (C) Sorted c-KIT$^+$ cells were cultured under erythroid and myeloid culture conditions for 1 week and FACS analyzed for cell-surface markers. (D) Morphology of colonies and cells generated by sorted c-KIT$^+$ cells.

(E) FACS and cellular morphology of cells derived from day 12 sorted single c-KIT$^+$ cell expanded on OP9 for 2 weeks ($n = 3$).

(F) Percentage reconstitution in peripheral blood determined by detection of donor-derived GFP$^+$ cells after 2 weeks of transplantation in two individual mice per group (1 and 2) either with GFP$^+$ or c-KIT$^+$ reprogrammed cells.

Confirming the relationship between c-KIT acquisition and clonogenic potential, c-KIT$^+$ cells displayed a higher clonogenic potential than c-KIT$^-$ cells (data not shown). When c-KIT$^+$ sorted cells were cultured in conditions that specifically support the growth of erythroid and myeloid cells, CD71/TER119 and CD11b/GR1 double-positive cells were respectively detected (Figure 3C). When evaluated in clonogenic replating assays, c-KIT$^+$ cells generated colonies containing erythroid, megakaryocytic, and myeloid cells (Figure 3D). Relatively similar genomic integration levels of the five vectors were observed between erythroid TER119 and myeloid CD11b-positive cells (Figure S3C). This observation suggests that both lineages have similar TF requirements for their generation and/or that they are generated through a common progenitor. Collectively, these results indicate that hematopoietic progenitors with myeloid, erythroid, and megakaryocytic potential are generated early during reprogramming.

To investigate the presence of multipotential progenitors, we evaluated the frequency of clonogenic precursors in the c-KIT$^+$ fraction, containing all hematopoietic progenitor cells, including the most immature precursors. Limiting dilution analyses indicated that 1 in 20 c-KIT$^+$ cells generated colonies when cultured on OP-9 stromal cells (data not shown). To investigate if c-KIT$^+$ cells have multilineage clonal ability, single day 12 fibroblast-derived c-KIT$^+$ cells were sorted onto OP-9 stromal cells and amplified for 2 weeks (Figure 3E). FACS and May-Grünwald Giemsa staining indicated the presence of erythroid, myeloid, and megakaryocytic cells in cultures initiated with a single c-KIT$^+$ cell (Figure 3E). These results suggest that fibroblasts are reprogrammed to the blood program, at least in part, through multipotential hematopoietic progenitor cells.

To further determine if hematopoietic stem cells might be generated during reprogramming, we performed FACS analyses for the SLAM markers on day 12 reprogrammed cells that were amplified on OP9-DL1. LIN$^-$c-KIT$^+$SCA1$^+$ (LSK) cells were detected in these cultures, and some of the LSK cells were CD48$^-$CD150$^+$, indicating the presence of phenotypic HSCs (Figure S3D). Additionally, we also observed phenotypic HSCs identified by the combination EPCR$^+$CD48$^-$CD45$^+$CD150$^+$ (Figure S4A). We next investigated if these reprogrammed cells have the capacity to engraft mice in vivo. For this, we established E14.5 MEFs from a transgenic mouse carrying the GFP reporter cDNA under the control of the pan-hematopoietic *AI467606* gene promoter (Ferreras et al., 2011). We amplified day 12 reprogrammed cells on OP9-DL1 stromal cells for 2 weeks. At this stage, GFP$^+$ cells, characterizing both progenitors and differentiated cells, or c-KIT$^+$ cells, including the most immature progenitors, were sorted and injected into two irradiated immunocompromised mice per each group. After 2 weeks, we monitored the presence of GFP$^+$ cells in the peripheral blood and detected short-term engraftment of both injected populations (Figures 3F and S4B). Engrafted cells were mostly TER119$^+$ erythroid cells (Figure S4B). However, 8 weeks postinjection, we could not detect engraftment greater than 1% in peripheral blood (data not shown). Collectively, these results suggest that the five-TF-mediated reprogrammed cells have short-term engraftment capacity and that further optimization of the culture conditions or TF composition will be necessary to obtain long-term engraftment capacity.

Loss of p53 Increases the Efficiency of Reprogramming to Blood

The loss of p53 or p16/p19 function has been shown to dramatically improve the efficiency of reprogramming to the pluripotent state (Hong et al., 2009; Li et al., 2009). We therefore investigated whether similarly deletion of these genes improved reprogramming of fibroblasts to blood. Following transduction with the five TFs, we observed significantly higher numbers of hematopoietic colonies generated by p53 and p16/p19-null MEFs than by wild-type MEFs (Figure 4A). These colonies emerged as early as 5 days after transduction, and scoring was performed on day 6, as their excessive proliferation precluded individual counting at later time points. Reprogramming of p53$^{-/-}$ MEFs resulted in the emergence of hematopoietic progenitors with multilineage potential (Figures 4B and 4C). Also, loss of p53 or p16/p19 function significantly increased the frequency of TER119$^+$ erythroid cells generated by reprogrammed cells (Figure 4D). Cells with mature megakaryocytic morphologies that stained positive for acetylcholinesterase activity were also more readily observed with p53$^{-/-}$ reprogrammed MEFs (Figure 4E). To evaluate the lymphoid potential of p53$^{-/-}$ reprogrammed cells, sorted c-KIT$^+$ cells were cultured and passaged on OP-9 or OP9-DL1 stromal cells in culture conditions supporting B or T cell growth, respectively. Cells positive for the B cell markers B220/CD19 emerged and proliferated during the culture of reprogrammed p53$^{-/-}$ MEFs (Figure 4Fi). A fraction of B220/CD19 double-positive reprogrammed p53$^{-/-}$ cells were also positive for immunoglobulin M expression (Figure 4Fi). The B cell identity of sorted B220/CD19-positive cells was further confirmed by detection of V(D)J chain rearrangements by PCR and sequencing (Figure 4Ri). Similarly, some reprogrammed p53$^{-/-}$ MEFs cells cultured on OP9-DL1 stromal cells acquired the expression of the early T cell marker CD25 and displayed T cell receptor rearrangements, confirming early T lymphoid commitment (Figures 4Gi and 4Gii). Collectively, these results demonstrate that p53$^{-/-}$ MEFs are more efficiently reprogrammed to hematopoietic progenitors with erythroid, myeloid, megakaryocyte, and B and T lymphoid lineage potential than wild-type MEFs, indicating that P53 expression is a barrier for reprogramming to blood.

Reprogramming to Blood Occurs via an Intermediate Endothelial Stage

We next investigated the molecular events and cellular processes leading to the reprogramming of fibroblasts to blood progenitors. The rapid emergence of hematopoietic colonies suggested that the reprogramming did not involve a pluripotent stem cell stage with subsequent differentiation into blood cells. Accordingly, the expression of pluripotent markers (*Oct4*, *Sox2*, and *Nanog*) was not detected during the course of reprogramming (Figure 5A). The detection of embryonic hemoglobin *β-H1* expression in reprogrammed cells (Figure 1C) suggested that the reprogramming process might recapitulate to some extent embryonic hematopoiesis. During embryonic hematopoiesis, mesodermal hemangioblast precursors generate primitive as well as definitive blood cells through an intermediate hemogenic endothelium stage. We evaluated by PCR whether the reprogramming of fibroblasts to blood process was associated

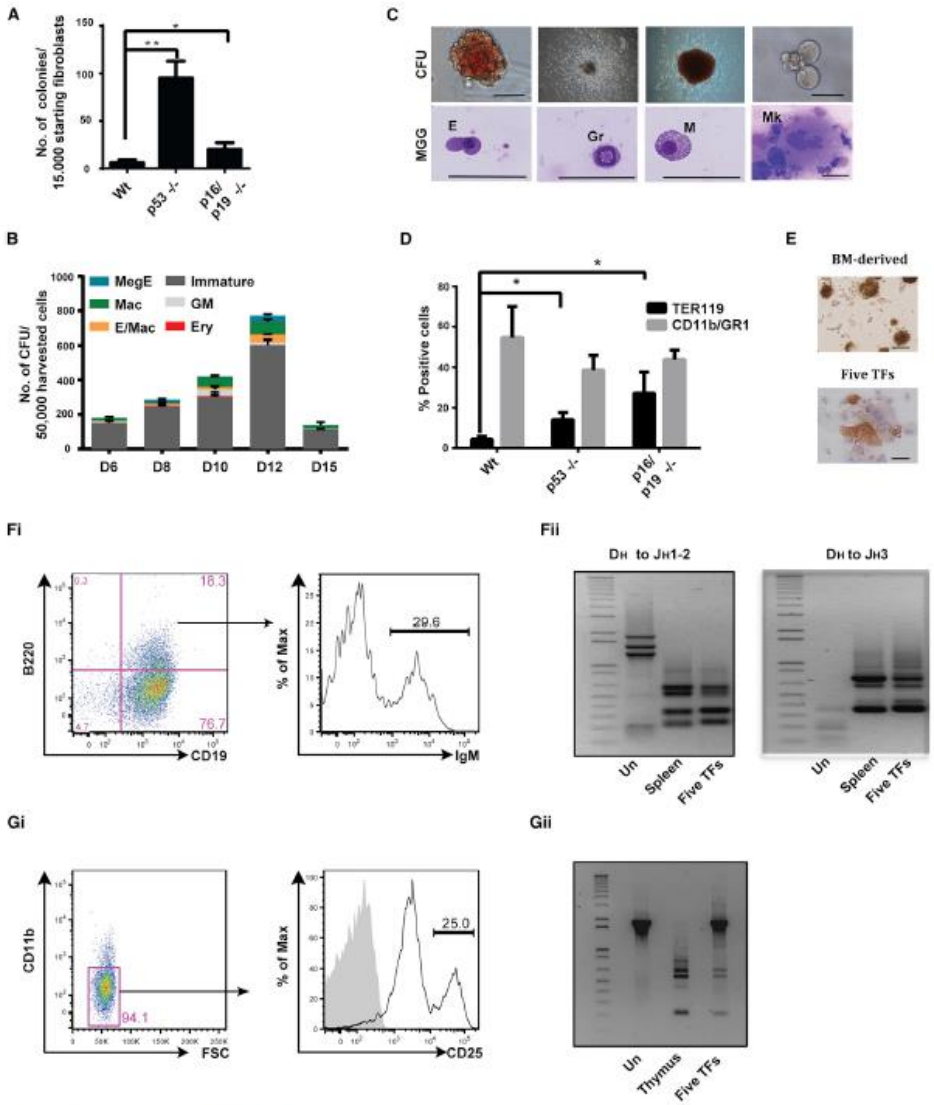


Figure 4. Reprogramming p53-Null MEFs

(A) Number of round cell colonies observed at day 6 per 15,000 transduced WT, $p53^{-/-}$, and $p16/p19^{-/-}$ MEFs (n = 4; mean ± SEM).

(B) Number and types of hematopoietic colonies generated by 50,000 five-TF-reprogrammed $p53^{-/-}$ MEFs harvested from day 6 to 15. Data presented are mean ± SEM of triplicates in a representative experiment (n = 2).

(legend continued on next page)

with any of these distinct embryonic steps. Although we did not detect any expression of the *Brachyury* mesodermal marker (data not shown), we clearly observed a peak of endothelial gene expression (*Cdh5*, *Tie2*, and *Icam1*; *Pecam1* and *Vwf*) early in reprogramming (Figures 5B and 5C). In contrast, hematopoietic genes (*Itga2b*, *Mpo*, and *Pu.1*) showed a delayed but steady increase in expression from day 12 to day 21 (Figure 5D). Live staining for vascular endothelial cadherin (CDH5) performed on day 5 cultures confirmed the early emergence of clusters of $CDH5^+$ endothelial cells (Figure 5E). To determine if these endothelial cells corresponded to hemogenic endothelium that gives rise to differentiated blood cells, we sorted $CDH5^+$ and $CDH5^-$ cells from c-KIT-depleted day 6 transduced MEFs to eliminate any potential hematopoietic precursors. The cells were then cultured in conditions that support the transition to hematopoietic progenitors before seeding them in clonogenic assays. As shown in Figure 5F, $CDH5^+$, but not $CDH5^-$, cells generated hematopoietic colonies, demonstrating their hemogenic potential. In addition, these $c\text{-}KIT^-/CDH5^+$ endothelial cells acquired the hematopoietic markers CD41, CD45, and c-KIT upon culture on OP9 (Figure S5A). We concluded from these results that reprogramming of fibroblasts to blood cells proceeds through an intermediate hemogenic endothelial cellular stage.

Transcriptome Analyses on Reprogrammed Cells

To define the global changes in gene expression driving reprogramming of fibroblasts to hematopoietic lineages, we performed exon array analyses on biological replicates of untreated MEFs (Un), day 8 sorted $CDH5^+$ cells, and day 12 sorted $c\text{-}KIT^+$ cells. We identified 868 significantly (5% false discovery rate [FDR]) differentially expressed genes (DEGs). Principal component analyses based on these DEGs confirmed the correlation between replicates and the differences between cells at the three stages of reprogramming (Figure S6A). We then performed hierarchical clustering on DEGs and observed three main clusters (Figure 6A). The first cluster mostly contained genes that were gradually downregulated from untransduced MEFs to $CDH5^+$ and then to $c\text{-}KIT^+$ populations and were associated with adhesion and muscular development (Figure 6A; Table S2). This cluster included fibroblast-specific genes (*Acta2*, *Actg2*, *Col2A*, *Col4a1*, *Col5a1*, and *Fgf3*) that were silenced during reprogramming. The second cluster encompassed genes that were upregulated in $CDH5^+$ cells but then downregulated in $c\text{-}KIT^+$ cells. These genes were mainly of endothelial nature and associated with blood vessel development and cell-cell adhesion (Figure 6A; Table S2). The third cluster included genes that were moderately expressed in $CDH5^+$ cells, highly expressed in $c\text{-}KIT^+$ populations, and associated with ontology terms such as development of the hematopoietic program, leukocyte migration, chemotaxis, and response to infection/wounding ontology terms (Figure 6A; Table S2). Pathway analyses with DAVID confirmed that hematopoietic pathways were already activated in the $CDH5^+$ fraction, a finding consistent with a hemogenic endothelium identity (Table S2). To further confirm this transient upregulation of the endothelial program, we performed gene set enrichment analyses (GSEA) with endothelial-specific genes. As expected, genes upregulated between MEFs and $CDH5^+$ cells, and downregulated between $CDH5^+$ and $c\text{-}KIT^+$ cells, were positively correlated with an endothelial gene signature (Figure 6B). These analyses also indicated an enrichment for genes specifically upregulated in $CDH5^+$ cells for early hematopoietic progenitors signature, whereas $c\text{-}KIT^+$-specific genes were more closely associated with late or differentiated hematopoietic progenitor cells (Figures 6C and S6B). Finally, we investigated to which extent our reprogrammed cells were similar to HSCs by comparing our transcriptome signature with previously published data sets from different HSC populations isolated from diverse hematopoietic tissues (McKinney-Freeman et al., 2012). We observed that $CDH5^+$ cells clustered with AGM, placental, and yolk sac HSCs that have been classified as specifying HSCs. Interestingly, $c\text{-}KIT^+$ cells clustered more closely with definitive HSCs from fetal liver, bone marrow, and embryonic stem cell (ESC)-derived HSCs (Figure 6D). Collectively, our transcriptome analyses further support the concept that reprogramming to hematopoietic progenitors proceeds through a hemogenic endothelium intermediate. This $CDH5^+$ cell population, while displaying clear evidence of its endothelial nature, also expressed hematopoietic genes and displayed similarity with emerging HSCs and early progenitors, whereas the later $c\text{-}KIT^+$ population was more closely associated with hematopoietic progenitors.

DISCUSSION

In this study, we establish that fibroblasts can be robustly reprogrammed to hematopoietic cell fate by concomitant ectopic expression of the hematopoietic TFs ERG, GATA2, LMO2, RUNX1c, and SCL. These five TFs have been shown to interact and act at diverse stages of the hematopoietic program; i.e., from the onset of its development to the more established adult hematopoietic hierarchy (Wilson et al., 2010). In particular, SCL, GATA2, and FLI1, a distinct but related-to-ERG ETS factor, form an interconnected regulatory triad that is activated during specification of HSCs (Pimanda et al., 2007). Once established, this circuit is self-maintained, providing the newly specified progenitors with a memory of their stemness in a similar fashion as OCT4, SOX2, and NANOG in pluripotent

(C) Morphology of colonies and cells obtained from five-TF-transduced day 12 sorted $c\text{-}KIT^+$ $p53^{-/-}$ MEFs.
(D) Average percentage of TER119 and CD11b/GR1 double-positive cells generated by day 12 sorted $c\text{-}KIT^+$ cells from WT, $p53^{-/-}$, and $p16/p19^{-/-}$ transduced MEFs (n = 3; mean ± SEM).
(E) Acetylcholinesterase staining of megakaryocytes derived from bone marrow or generated by day 12 sorted $c\text{-}KIT^+$ reprogrammed $p53^{-/-}$ MEFs.
(Fi) FACS analysis of $c\text{-}KIT^+$ sorted reprogrammed $p53^{-/-}$ MEFs after expansion on OP9 in lymphoid medium.
(Fii) BCR rearrangements detection in $B220^+/CD19^+$ sorted reprogrammed $p53^{-/-}$ MEFs, control untransduced MEFs (Un) and spleen cells.
(Gi) FACS analysis of $c\text{-}KIT^+$ sorted reprogrammed $p53^{-/-}$ MEFs after expansion on OP9-DL1 in lymphoid medium.
(Gii) TCR rearrangement detection in sorted $CD25^+$ reprogrammed $p53^{-/-}$ MEFs, control untransduced MEFs (Un) and thymus cells.
Asterisk(s) represents statistical significance (Student's t test; $^*p < 0.05$, $^{**}p < 0.01$).

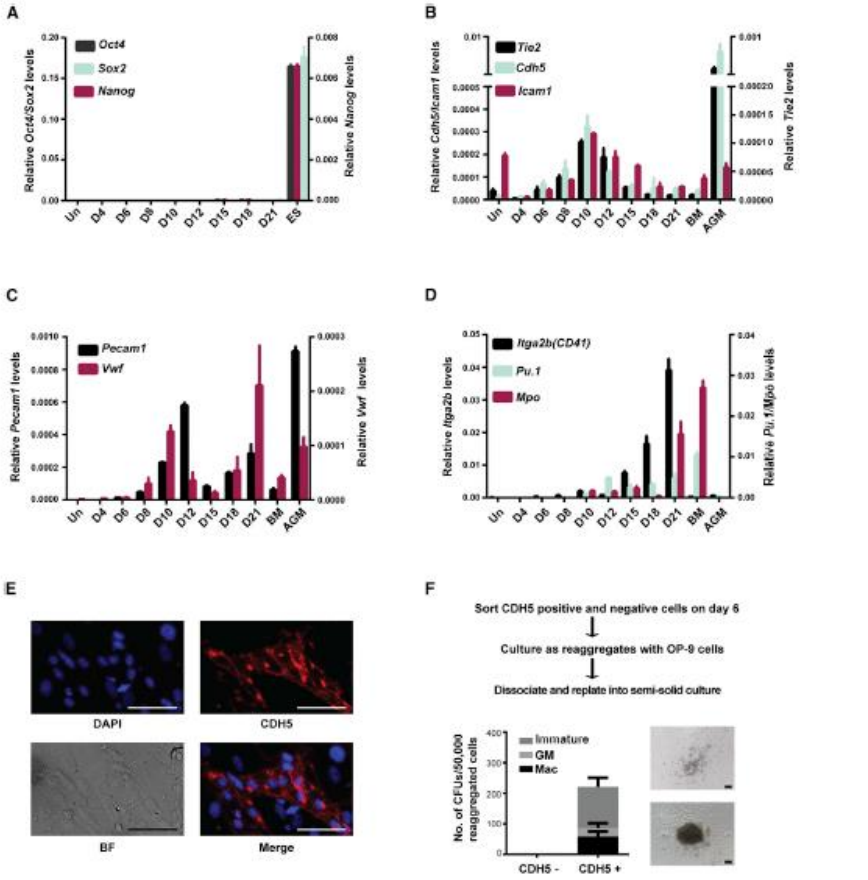


Figure 5. Fibroblasts Are Reprogrammed to Blood via an Intermediate Hemogenic Endothelial Stage

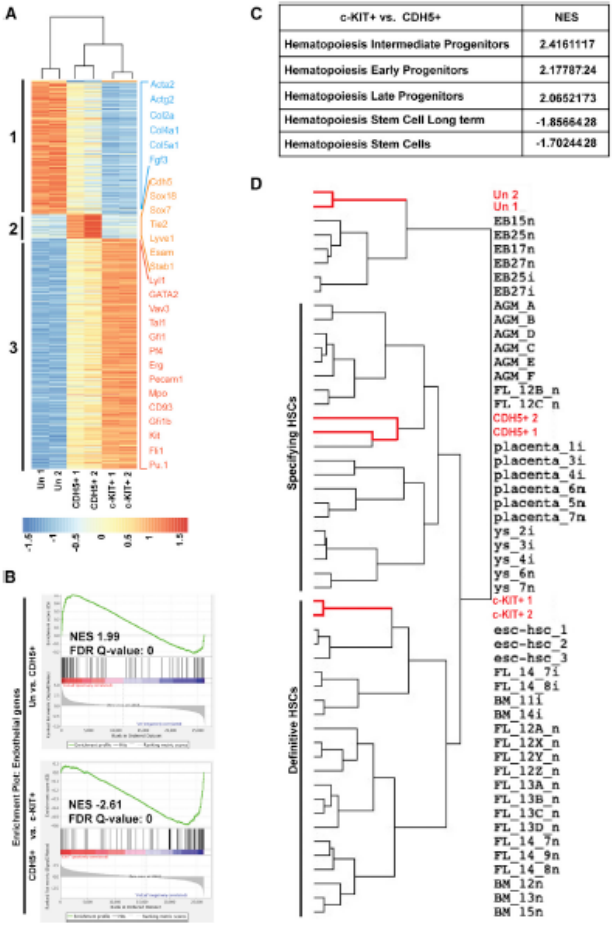
(A–D) Relative gene expression levels of pluripotent (A), endothelial (B), endothelial/ hematopoietic (C), and hematopoietic (D) markers with respect to β-actin at indicated days after five-TF transduction and in control bone marrow (BM) and aorta-gonad-mesonephros (AGM) region cells. Data presented are mean ± SD from one representative experiment ($n = 3$).

(E) Immunostaining of day 5 five-TF-transduced MEFs for CDH5 (red) and DAPI (blue).

(F) Schematic strategy to determine hemogenic potential of $CDH5^{-}cKIT^{-}$ and $CDH5^{+}cKIT^{-}$ five-TF-transduced cells (top). Number, type, and morphology of hematopoietic colonies generated by sorted cells reaggregated with irradiated OP9 stromal cells.

Data presented are mean ± SEM of triplicates of a representative experiment ($n = 2$). Scale bars represent 50 μm.

ESCs (Boyer et al., 2005). Furthermore, ERG, when expressed along with HOXA9 and RORA, has also been shown to confer multilineage potential to myeloid-restricted precursors (Doulatov et al., 2013). The transcription factor SCL and its binding partners are implicated in the induction of the hematopoietic program in mesoderm and the generation of hemogenic endothelium (Gering et al., 2003; Lancrin et al., 2009). Finally, RUNX1 physically interacts with SCL, ERG, and GATA2 and is critical for the transition from hemogenic endothelium to hematopoietic stem and progenitor cells (HSPCs) (Chen et al., 2009; Lancrin et al., 2010; Wilson et al., 2010). Furthermore, enforced expression of the isoform RUNX1a enhances



c-KIT+ vs. CDH5+	NES
Hematopoiesis Intermediate Progenitors	2.4161117
Hematopoiesis Early Progenitors	2.17787.24
Hematopoiesis Late Progenitors	2.0652173
Hematopoiesis Stem Cell Long term	-1.8566428
Hematopoiesis Stem Cells	-1.7024428

Figure 6. Global Transcriptome Analyses of Reprogrammed Cells

Day 8 CDH5$^+$ (CDH5-positive) and day 12 c-KIT$^+$ (c-KIT-positive) flow-sorted cells were compared to untreated (Un) MEFs in duplicate by mouse Affymetrix exon arrays.

(A) Hierarchical clustering of differentially expressed genes (DEGs) among untreated MEFs, day 8 sorted CDH5$^+$, and day 12 sorted c-KIT$^+$ cells. Genes specific to specified cell types are shown on the right.

(B) Gene set enrichment analysis (GSEA) for endothelial gene signature in untreated, CDH5$^+$, and c-KIT$^+$ transcriptomes.

(C) Normalized enrichment score (NES) values obtained after performing GSEA for indicated gene sets comparing CDH5$^+$ and c-KIT$^+$ transcriptome data sets.

(D) Unsupervised hierarchical clustering of DEGs in our transcriptome data sets along with expression in HSCs from different hematopoietic organs (McKinney-Freeman et al., 2012).

hematopoietic lineage commitment from human ESCs and iPSCs (Ran et al., 2013). The close functional association among these five TFs during normal HSPCs generation could explain why we functionally identified them as the best combination of TFs to robustly induce a complex hematopoietic program in differentiated fibroblasts. Interestingly we also demonstrated that SCL and LMO2 are sufficient to induce hematopoietic fate in embryonic but not adult fibroblasts. This finding raises the interesting prospect that a limited set of small-molecule modulators could reinforce reprogramming or even circumvent the need for viral transduction.

We observed that the generation of blood precursors by the ectopic expression of our set of five TFs is a very fast process. The first morphological signs of cell-fate switching were observed by day 5 followed by the emergence and proliferation of c-KIT$^+$ progenitors by 8 days. These cells exhibited a wide range of differentiation potential with robust generation of granulocytes and functional macrophages and, to a lower extent, erythrocytes and megakaryocytes. Furthermore, using fibroblasts with a p53$^{-/-}$ background broadened the range and increased the frequency of the generated lineages. In addition to larger pool of erythrocytes and megakaryocytes, we were able to generate B and T cells at different developmental stages. The results obtained with p53$^{-/-}$ MEFs provide a proof of principle that our combination of five TFs is suitable for reprogramming to all major blood lineages. The addition of small-molecule modulators that transiently increase epigenetic plasticity or briefly inhibit p53 or p16 might further extend the range and frequency of lineages generated from wild-type fibroblasts. In contrast, a permanent inactivation of p53 might not be desirable given the requirement for p53 to limit aberrant self-renewal (Zhao et al., 2010).

Establishing methods for transdifferentiation of nonhematopoietic cells to blood lineage remains challenging due to significant epigenetic barriers to overcome and the need to induce and maintain complex regulatory networks in these cells. So far, there are only two reports describing transdifferentiation of fibroblasts to blood progenitors. In the first study, hematopoietic progenitors with multilineage potential were generated

by overexpression of a single pluripotency factor, OCT4, in a hematopoietic-supportive culture environment (Szabo et al., 2010). The use of a pluripotent TF, however, raises some concerns about tumorigenicity. In the second report, Pereira et al. demonstrated that ectopic expression of c-FOS, ETV6, GATA2, and GFI1b reprogrammed fibroblasts to hemogenic endothelial cells, which upon further coculture with placental cells generated myeloid cells, albeit with a low efficiency (around 10 colonies for 100,000 starting MEFs) (Pereira et al., 2013). Although reprogramming by c-FOS, ETV6, GATA2, and GFI1b, as well as our combination of five TFs, seems to transit through a similar hemogenic endothelium intermediate stage, the efficiency and kinetics are quite different. While these authors identified a hemogenic endothelium population emerging by day 35 of reprogramming, our results indicate the emergence of an endothelial population as early as day 5. We detected c-KIT$^+$ functional hematopoietic progenitors with multilineage potential as early as 8 days after transduction, whereas Pereira et al. obtained only limited myeloid potential after more than 6 weeks of culture. Collectively, these results indicate that our combination of TFs induce the rapid generation of a multipotent blood progenitors with better efficiency than previously reported.

Several lines of evidence suggest that the reprogramming of fibroblasts to blood cells recapitulates the embryonic process of blood development. CD41, the first hematopoietic marker we detected at the surface of transitioning reprogrammed cells, is widely expressed on the earliest blood progenitors generated during ontogeny (Figure S1A) (Ferkowicz et al., 2003; Mikkola et al., 2003; Mitjavila-Garcia et al., 2002; Robin et al., 2011). The detection of primitive erythrocytes expressing the embryonic βH1-hemoglobin also suggests that these TFs are inducing early developmental steps. Moreover, the presence of a transient hemogenic endothelium population resembles the endothelial-to-hematopoietic transition that takes place during blood development in the yolk sac or dorsal aorta (Boisset et al., 2010; Jaffredo et al., 1998; Lancrin et al., 2009; Zovein et al., 2008). Finally, the detection of phenotypic HSCs, as well as the gene expression profile clustering of the CDH5$^+$ cell population with specifying HSC cell populations, further supports the hypothesis that the reprogramming mimics early hematopoietic development (Figures 6D, S3D, and S4A). The fact that we are recapitulating the normal embryonic program of blood development suggests that it may be only due to the limitations of the in vitro cultures that we cannot functionally identify an emerging HSC population. The reports by the Rossi and Rafii's laboratories emphasize the requirement for a supportive niche to obtain engraftable populations upon reprogramming of blood or endothelial cells (Riddell et al., 2014; Sandler et al., 2014). Providing a similar in vivo bone marrow or in vitro vascular support might greatly improve the repopulation potential of cells generated from fibroblasts and therefore will be a major direction for further research.

In conclusion, we provide a report of efficient and rapid reprogramming of fibroblasts to blood progenitors with a limited set of hematopoietic TFs. Fibroblasts represent an accessible and safe starting cell population for reprogramming. Despite many remaining questions, the approach reported here holds a huge promise and should be further optimized and developed for its safe use in clinical settings.

EXPERIMENTAL PROCEDURES

MEF and MAF Preparation

E14.5 MEFs were isolated as described previously (Sroczynska et al., 2009). MAFs were isolated from the ears of 2- to 6-month-old mice by collagenase digestion. Passage 0 fibroblasts were cultured until confluent and frozen. Before infection, MEFs and MAFs were depleted of cells positive for the hematopoietic/endothelial markers CD41, CD31, c-KIT, and CD45.

Lentivirus Transduction, Colony-Forming Unit Assay, and Reaggregation

MEFs/MAFs (15,000) were seeded on a gelatin-coated 12-well plate, and after 24 hr viral transductions were carried out in the presence of 10 μg/ml diethylaminoethyl dextran. After 4 hr of transduction, the cultures were placed in hematopoietic media (1× Iscove's modified Dulbecco's medium supplemented with plasma-derived serum [PDS; Antech], 10% protein-free hybridoma medium [PFM; GIBCO], 0.5 mM ascorbic acid, 4.5 × 10^{-4} M MTG, 2 mM L-glutamine, 80 mg/ml transferrin, 1% c-KIT ligand, 1% IL-3, 1% granulocyte-macrophage colony-stimulating factor, 1% thrombopoietin conditioned media, 4 U/ml erythropoietin [Ortho-Biotech], 10 ng/ml macrophage colony-stimulating factor, 10 ng/ml IL-6, 5 ng/ml IL-11 [all from R&D Systems], and 50 μg/ml penicillin-streptomycin). Colony-forming unit (cfu) assays were performed as described previously (Sroczynska et al., 2009). To evaluate the presence of hemogenic endothelium, CDH5$^+$ and CDH5$^-$ cells were sorted from day 6 cultures that were depleted for c-KIT$^+$ cells to avoid hematopoietic contamination. A total of 20,000 sorted cells were mixed with 100,000 irradiated OP9 cells, cultured overnight in hanging-drop cultures, and then on Durapore filter (Millipore) for another 4 days. Single-cell suspensions obtained following dissociation with collagenase/dispase solution were assayed in cfu assays.

Transplantation Assay

E14.5 MEFs for transplantation experiments were obtained from a reporter mouse line with GFP expression under the control of *AI467606* promoter (Ferreras et al., 2011). Day 12 transduced cells were harvested and further expanded for 2 weeks on OP-DL1 stroma cells in hematopoietic medium. GFP- or c-KIT-positive cells were sorted and intravenously injected in 6- to 8-week-old lethally irradiated NSG mice (n = 2). The level of engraftment was determined by GFP expression in peripheral blood. All animal work was performed under regulations set out by the Home Office Legislation under the 1986 United Kingdom Animal Scientific Procedures Act.

Affymetrix Analyses and Integration of Data Sets

Global gene expression analyses were done using the Mouse Exon 1.0 ST arrays. The R/Bioconductor package LIMMA was used to identify genes that were differentially expressed among the three conditions. Differentially expressed genes were obtained with significant p value and at 5% FDR. The distance weighted discrimination (Benito et al., 2004) method for cross-platform normalization was done for our transcriptome data sets along with McKinney data sets (McKinney-Freeman et al., 2012). Hierarchical clustering was performed using the complete linkage.

ACCESSION NUMBERS

Affymetrix data were deposited in the Gene Expression Omnibus under the accession number GSE59428.

SUPPLEMENTAL INFORMATION

Supplemental Information includes Supplemental Experimental Procedures, six figures, two tables, and one movie and can be found with this article online at http://dx.doi.org/10.1016/j.celrep.2014.11.002.

AUTHOR CONTRIBUTIONS

K.B. and M.F. designed and performed experiments, analyzed the data, and wrote the manuscript. V.K. and G.L. designed and supervised the research project, analyzed the data, and wrote the manuscript.

ACKNOWLEDGMENTS

We thank the staff at the Advanced Imaging, Animal Facility, Molecular Biology Core facilities, and Flow Cytometry of CRUK Manchester Institute for technical support and Hui Sun Leong for help with the bioinformatic analyses. We thank members of the Stem Cell Biology and Stem Cell Haematopoiesis group for valuable advice and critical reading of the manuscript. Work in our laboratory is supported by the Leukaemia and Lymphoma Research Foundation, the Biotechnology and Biological Sciences Research Council, and Cancer Research UK.

Received: July 1, 2014
Revised: October 10, 2014
Accepted: November 4, 2014
Published: November 26, 2014

REFERENCES

Benito, M., Parker, J., Du, Q., Wu, J., Xiang, D., Perou, C.M., and Marron, J.S. (2004). Adjustment of systematic microarray data biases. Bioinformatics *20*, 105–114.

Blum, B., and Benvenisty, N. (2008). The tumorigenicity of human embryonic stem cells. Adv. Cancer Res. *100*, 133–158.

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

Boyer, L.A., Lee, T.I., Cole, M.F., Johnstone, S.E., Levine, S.S., Zucker, J.P., Guenther, M.G., Kumar, R.M., Murray, H.L., Jenner, R.G., et al. (2005). Core transcriptional regulatory circuitry in human embryonic stem cells. Cell *122*, 947–956.

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

Doulatov, S., Vo, L.T., Chou, S.S., Kim, P.G., Arora, N., Li, H., Hadland, B.K., Bernstein, I.D., Collins, J.J., Zon, L.I., and Daley, G.Q. (2013). Induction of multipotential hematopoietic progenitors from human pluripotent stem cells via respecification of lineage-restricted precursors. Cell Stem Cell *13*, 459–470.

Du, Y., Wang, J., Jia, J., Song, N., Xiang, C., Xu, J., Hou, Z., Su, X., Liu, B., Jiang, T., et al. (2014). Human hepatocytes with drug metabolic function induced from fibroblasts by lineage reprogramming. Cell Stem Cell *14*, 394–403.

Feng, R., Desbordes, S.C., Xie, H., Tillo, E.S., Pixley, F., Stanley, E.R., and Graf, T. (2008). PU.1 and C/EBPalpha/beta convert fibroblasts into macrophage-like cells. Proc. Natl. Acad. Sci. USA *105*, 6057–6062.

Ferkowicz, M.J., Starr, M., Xie, X., Li, W., Johnson, S.A., Shelley, W.C., Morrison, P.R., and Yoder, M.C. (2003). CD41 expression defines the onset of primitive and definitive hematopoiesis in the murine embryo. Development *130*, 4393–4403.

Ferreras, C., Lancrin, C., Lie-A-Ling, M., Kouskoff, V., and Lacaud, G. (2011). Identification and characterization of a novel transcriptional target of RUNX1/AML1 at the onset of hematopoietic development. Blood *118*, 594–597.

Gering, M., Yamada, Y., Rabbitts, T.H., and Patient, R.K. (2003). Lmo2 and Scl/Tal1 convert non-axial mesoderm into haemangioblasts which differentiate into endothelial cells in the absence of Gata1. Development *130*, 6187–6199.

Hong, H., Takahashi, K., Ichisaka, T., Aoi, T., Kanagawa, O., Nakagawa, M., Okita, K., and Yamanaka, S. (2009). Suppression of induced pluripotent stem cell generation by the p53-p21 pathway. Nature *460*, 1132–1135.

Huang, P., Zhang, L., Gao, Y., He, Z., Yao, D., Wu, Z., Cen, J., Chen, X., Liu, C., Hu, Y., et al. (2014). Direct reprogramming of human fibroblasts to functional and expandable hepatocytes. Cell Stem Cell *14*, 370–384.

Ieda, M., Fu, J.D., Delgado-Olguin, P., Vedantham, V., Hayashi, Y., Bruneau, B.G., and Srivastava, D. (2010). Direct reprogramming of fibroblasts into functional cardiomyocytes by defined factors. Cell *142*, 375–386.

Jaffredo, T., Gautier, R., Eichmann, A., and Dieterlen-Lièvre, F. (1998). Intraaortic hemopoietic cells are derived from endothelial cells during ontogeny. Development *125*, 4575–4583.

Lancrin, C., Sroczynska, P., Stephenson, C., Allen, T., Kouskoff, V., and Lacaud, G. (2009). The haemangioblast generates haematopoietic cells through a haemogenic endothelium stage. Nature *457*, 892–895.

Lancrin, C., Sroczynska, P., Serrano, A.G., Gandillet, A., Ferreras, C., Kouskoff, V., and Lacaud, G. (2010). Blood cell generation from the hemangioblast. J. Mol. Med. *88*, 167–172.

Li, H., Collado, M., Villasante, A., Strati, K., Ortega, S., Cañamero, M., Blasco, M.A., and Serrano, M. (2009). The Ink4/Arf locus is a barrier for iPS cell reprogramming. Nature *460*, 1136–1139.

Loose, M., Swiers, G., and Patient, R. (2007). Transcriptional networks regulating hematopoietic cell fate decisions. Curr. Opin. Hematol. *14*, 307–314.

Loughran, S.J., Kruse, E.A., Hacking, D.F., de Graaf, C.A., Hyland, C.D., Willson, T.A., Henley, K.J., Ellis, S., Voss, A.K., Metcalf, D., et al. (2008). The transcription factor Erg is essential for definitive hematopoiesis and the function of adult hematopoietic stem cells. Nat. Immunol. *9*, 810–819.

McKinney-Freeman, S., Cahan, P., Li, H., Lacadie, S.A., Huang, H.-T., Curran, M., Loewer, S., Naveiras, O., Kathrein, K.L., Konantz, M., et al. (2012). The transcriptional landscape of hematopoietic stem cell ontogeny. Cell Stem Cell *11*, 701–714.

Mikkola, H.K., Fujiwara, Y., Schlaeger, T.M., Traver, D., and Orkin, S.H. (2003). Expression of CD41 marks the initiation of definitive hematopoiesis in the mouse embryo. Blood *101*, 508–516.

Mitjavila-Garcia, M.T., Cailleret, M., Godin, I., Nogueira, M.M., Cohen-Solal, K., Schiavon, V., Lecluse, Y., Le Pesteur, F., Lagrue, A.H., and Vainchenker, W. (2002). Expression of CD41 on hematopoietic progenitors derived from embryonic hematopoietic cells. Development *129*, 2003–2013.

Nam, Y.J., Song, K., Luo, X., Daniel, E., Lambeth, K., West, K., Hill, J.A., DiMaio, J.M., Baker, L.A., Bassel-Duby, R., and Olson, E.N. (2013). Reprogramming of human fibroblasts toward a cardiac fate. Proc. Natl. Acad. Sci. USA *110*, 5588–5593.

Ohnishi, K., Semi, K., Yamamoto, T., Shimizu, M., Tanaka, A., Mitsunaga, K., Okita, K., Osafune, K., Arioka, Y., Maeda, T., et al. (2014). Premature termination of reprogramming in vivo leads to cancer development through altered epigenetic regulation. Cell *156*, 663–677.

Okuda, T., van Deursen, J., Hiebert, S.W., Grosveld, G., and Downing, J.R. (1996). AML1, the target of multiple chromosomal translocations in human leukemia, is essential for normal fetal liver hematopoiesis. Cell *84*, 321–330.

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

Pereira, C.F., Chang, B., Qiu, J., Niu, X., Papatsenko, D., Hendry, C.E., Clark, N.R., Nomura-Kitabayashi, A., Kovacic, J.C., Ma'ayan, A., et al. (2013). Induction of a hemogenic program in mouse fibroblasts. Cell Stem Cell *13*, 205–218.

Pereira, C.F., Lemischka, I.R., and Moore, K. (2014). 'From blood to blood': de-differentiation of hematopoietic progenitors to stem cells. EMBO J. *33*, 1511–1513.

Pimanda, J.E., and Göttgens, B. (2010). Gene regulatory networks governing haematopoietic stem cell development and identity. Int. J. Dev. Biol. *54*, 1201–1211.

Pimanda, J.E., Ottersbach, K., Knezevic, K., Kinston, S., Chan, W.Y.I., Wilson, N.K., Landry, J.-R., Wood, A.D., Kolb-Kokocinski, A., Green, A.R., et al. (2007). Gata2, Fli1, and Scl form a recursively wired gene-regulatory circuit during early hematopoietic development. Proc. Natl. Acad. Sci. USA *104*, 17692–17697.

Ran, D., Shia, W.J., Lo, M.C., Fan, J.B., Knorr, D.A., Ferrell, P.I., Ye, Z., Yan, M., Cheng, L., Kaufman, D.S., and Zhang, D.E. (2013). RUNX1a enhances hematopoietic lineage commitment from human embryonic stem cells and inducible pluripotent stem cells. Blood *121*, 2882–2890.

Riddell, J., Gazit, R., Garrison, B.S., Guo, G., Saadatpour, A., Mandal, P.K., Ebina, W., Volchkov, P., Yuan, G.C., Orkin, S.H., and Rossi, D.J. (2014). Reprogramming committed murine blood cells to induced hematopoietic stem cells with defined factors. Cell *157*, 549–564.

Robb, L., Elwood, N.J., Elefanty, A.G., Köntgen, F., Li, R., Barnett, L.D., and Begley, C.G. (1996). The scl gene product is required for the generation of all hematopoietic lineages in the adult mouse. EMBO J. *15*, 4123–4129.

Robin, C., Ottersbach, K., Boisset, J.-C., Oziemlak, A., and Dzierzak, E. (2011). CD41 is developmentally regulated and differentially expressed on mouse hematopoietic stem cells. Blood *117*, 5088–5091.

Sandler, V.M., Lis, R., Liu, Y., Kedem, A., James, D., Elemento, O., Butler, J.M., Scandura, J.M., and Rafii, S. (2014). Reprogramming human endothelial cells to haematopoietic cells requires vascular induction. Nature *511*, 312–318.

Sekiya, S., and Suzuki, A. (2011). Direct conversion of mouse fibroblasts to hepatocyte-like cells by defined factors. Nature *475*, 390–393.

Sroczynska, P., Lancrin, C., Pearson, S., Kouskoff, V., and Lacaud, G. (2009). In vitro differentiation of mouse embryonic stem cells as a model of early hematopoietic development. Methods Mol. Biol. *538*, 317–334.

Sturgeon, C.M., Ditadi, A., Clarke, R.L., and Keller, G. (2013). Defining the path to hematopoietic stem cells. Nat. Biotechnol. *31*, 416–418.

Szabo, E., Rampalli, S., Risueño, R.M., Schnerch, A., Mitchell, R., Fiebig-Comyn, A., Levadoux-Martin, M., and Bhatia, M. (2010). Direct conversion of human fibroblasts to multilineage blood progenitors. Nature *468*, 521–526.

Takahashi, K., and Yamanaka, S. (2006). Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. Cell *126*, 663–676.

Takahashi, K., Tanabe, K., Ohnuki, M., Narita, M., Ichisaka, T., Tomoda, K., and Yamanaka, S. (2007). Induction of pluripotent stem cells from adult human fibroblasts by defined factors. Cell *131*, 861–872.

Taoudi, S., Bee, T., Hilton, A., Knezevic, K., Scott, J., Willson, T.A., Collin, C., Thomas, T., Voss, A.K., Kile, B.T., et al. (2011). ERG dependence distinguishes developmental control of hematopoietic stem cell maintenance from hematopoietic specification. Genes Dev. *25*, 251–262.

Tsai, F.Y., Keller, G., Kuo, F.C., Weiss, M., Chen, J., Rosenblatt, M., Alt, F.W., and Orkin, S.H. (1994). An early haematopoietic defect in mice lacking the transcription factor GATA-2. Nature *371*, 221–226.

Vierbuchen, T., Ostermeier, A., Pang, Z.P., Kokubu, Y., Südhof, T.C., and Wernig, M. (2010). Direct conversion of fibroblasts to functional neurons by defined factors. Nature *463*, 1035–1041.

Wilson, N.K., Foster, S.D., Wang, X., Knezevic, K., Schütte, J., Kaimakis, P., Chilarska, P.M., Kinston, S., Ouwehand, W.H., Dzierzak, E., et al. (2010). Combinatorial transcriptional control in blood stem/progenitor cells: genome-wide analysis of ten major transcriptional regulators. Cell Stem Cell *7*, 532–544.

Zhao, Z., Zuber, J., Diaz-Flores, E., Lintault, L., Kogan, S.C., Shannon, K., and Lowe, S.W. (2010). p53 loss promotes acute myeloid leukemia by enabling aberrant self-renewal. Genes Dev. *24*, 1389–1402.

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., et al. (2008). Fate tracing reveals the endothelial origin of hematopoietic stem cells. Cell Stem Cell *3*, 625–636.