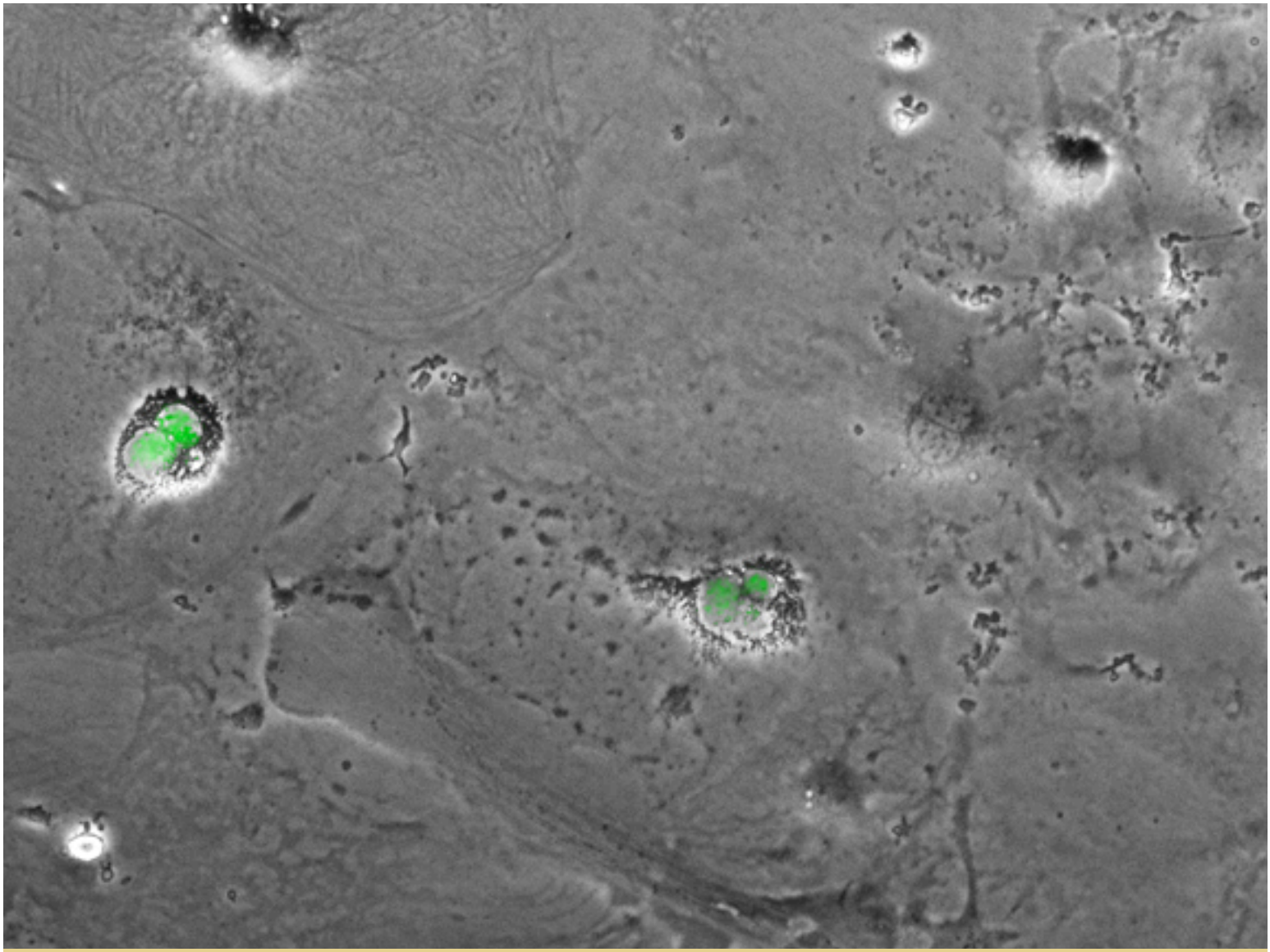

Rita Alexandra Silvério Alves

Designing Transcription Factors for Efficient Hematopoietic Reprogramming

Dissertação de Mestrado em Bioquímica, orientada pelo Dr. Carlos Filipe Pereira, apresentada ao Departamento de Ciências da Vida da Faculdade de Ciências e Tecnologia da Universidade de Coimbra

Agosto de 2017

· U C ·

Universidade de Coimbra

Designing Transcription Factors for Efficient Hematopoietic Reprogramming

By
Rita Alexandra Silvério Alves

Thesis submitted to Faculdade de Ciências e Tecnologias da Universidade de Coimbra for the degree of Master in Biochemistry

Supervisor: Dr. Carlos Filipe Pereira
Co-supervisor: Dr. Paula Veríssimo

Centro de Neurociências e Biologia Celular (CNC)
Universidade de Coimbra

August of 2017

ABSTRACT

Hematopoietic stem cell transplantation (HSCT) has been used as a treatment for a variety of haematological disorders, due to the ability of hematopoietic stem cells (HSCs) to self-renew and differentiate into all blood cell lineages. Insufficient number of cells and matching incompatibilities between donors and recipients hinder the broad application of this therapy. Expansion of HSCs has met limited success and additional strategies for the *in vitro* generation of HSCs are required to overcome transplant-associated limitations. Somatic cell reprogramming mediated by transcription factors (TFs) is opening new avenues for regenerative medicine and allowed the design of new approaches to convert one differentiated cell type directly into another. In the hematopoietic system, direct reprogramming of fibroblasts to HSC-like cells has been shown through ectopic expression of Gata2, Gfi1b and cFos, providing an alternative method to generate patient tailored HSCs *in vitro*. A better understanding of the mode of action of these three critical TFs during reprogramming is needed in order to increase the efficiency of the process.

Here, I have defined potential reprogramming domains of hematopoietic TFs by homologous gene (paralog) and deletion construct substitution during hematopoietic reprogramming. First, paralogs of Gata2, Gfi1b and cFos and Gata2 deletion constructs were cloned into a lentiviral gene delivery system to induce fibroblast cell identity towards the hematopoietic lineage. Secondly, hematopoietic reprogramming efficiency was assessed by hCD34/H2BGFP reporter activation. Interestingly, Gata1 did not substitute Gata2 for hematopoietic reprogramming, despite evidences of overlapping function during hematopoiesis. Notwithstanding, Gfi1b and cFos were partially replaced by their respective paralogs, indicating a determinant role for non-homologous domains of Gata2 during reprogramming. Consistently, hematopoietic reprogramming with Gata2 deletion constructs revealed the requirement of the transactivation domains (TADs), the negative regulatory domain (NRD) and the C-terminal zinc finger (C-ZF) for successful reprogramming. Remarkably, I have also unveiled that Gata2 display mitotic bookmarking activity. This epigenetic feature may be important for the acquisition and maintenance of the HSC fate as well as the inheritance of the reprogrammed cell state to daughter cells.

Overall, this study identified functional reprogramming features of Gata2, Gfi1b and cFos and sheds new light on how the HSC fate is acquired and preserved. Hereafter, these reprogramming modules will be critical for the design of enhanced synthetic TFs to increase hematopoietic reprogramming efficiency bringing this technology one step closer to clinical translation.

ACKNOLEGMENTS

I would like to start by thanking my supervisor, Dr Carlos Filipe Pereira, for giving me the opportunity to perform this challenging thesis project in his laboratory, alongside an amazing team, and for providing me with the right environment to take the first steps into science. Without his guidance, motivation and expertise, none of this would be possible and for that I am deeply grateful. You are truly someone that I admire and who has become a role model for me.

I would also like to acknowledge all the members of the Cell Reprogramming and Developmental Hematopoiesis Lab who supported me with their precious help whenever needed, specially Cristiana (Cris), Alexandra, Andreia, Fábio, Kritika, Rigveda and Raquel. Cris, you were like a mother to me. You taught me how to "crawl, walk and run" in the lab and nurtured my scientific spirit. Thank you for encouraging me with your contagious good mood. Alexandra, who more than a lab partner, became a close friend. Although we have similar backgrounds, you taught me a great deal and helped me out so many times that I cannot thank you enough. I hope we can pursue new challenges together, in the future. Andreia, who would gladly share her priceless knowledge and whose sense of responsibility made me respect her even more. Fábio, whose energy and resolve are always mesmerizing. When I first met you, I said I would like to know as much as you did as a master's student. I hope I have gotten closer to that goal. Kritika and Rigveda, thank you for keeping me company during our trips. I am glad you came to Portugal and took part in this adventure. Last but not least, Raquel, who brought a new light to the lab. You probably taught me more than I taught you, during those couple of weeks of your lab rotation.

I will take this opportunity to thank my friends Bárbara, Beatriz and Renata for the coffee time moments where we shared our worries and dreams, and to my dear friend Margarida for listening to my every doubt patiently. No matter how far apart we may grow, your friendship will never be forgotten.

A special thanks to my boyfriend João, who accompanied me through the whole year and inspired me to always give my best. Thank you for your understanding, for listening to me, for making me smile and for supporting me in the worst moments. No one knows better than you how much this year meant to me. Thank you for being part of my life.

Por fim, gostaria de agradecer à minha família por me ter apoiado incondicionalmente e por me ter providenciado todos os recursos necessários à realização deste projeto. Em especial à minha mãe e irmã, que lidaram recorrentemente com as minhas inconstâncias, mas que nunca deixaram de acreditar em mim. Se sou como sou, a vocês o devo.

TABLE OF CONTENTS

TABLE OF FIGURES

TABLE OF TABLES

ABREVIATIONS

AGM	Aorta–gonad–mesonephros
bZIP	Basic leucine zipper
cDNA	Complementary DNA
CDS	Coding sequences
CFU	Colony forming units
CLP	Common lymphoid progenitor
CMP	Common myeloid progenitor
CMV	Cytomegalovirus
dHSC	Definitive HSC
DMEM	Dulbecco's modified eagle medium
DMSO	Dimethyl sulfoxide
DNA	Deoxyribonucleic acid
DT	Double transgenic
EHT	Endothelial-to-hematopoietic transition
ESC	Embryonic stem cells
FACS	Fluorescence-activated cell sorting
FBS	Fetal bovine serum
FSC	Forward scatter
gDNA	Genomic DNA
GFP	Green fluorescent protein
GVHD	Graft-versus-host disease
hCD43	Human CD34
HEK	Human embryonic kidney
HLA	Human leukocyte antigen
HP	Hemogenic precursor
HSC	Hematopoietic stem cell
HSCT	Hematopoietic stem cell transplantation
HSPCs	Hematopoietic stem and progenitor cells
iPSCs	Induced PSCs
IRES	Internal ribosome entry site
LT	Long-term
mCMV	Minimal CMV (promoter)
MCS	Multiple cloning site
MD	Mitotic degron
MEF	Mouse embryonic fibroblast
MPPs	Multipotent progenitors
NLS	Nuclear localization signal
NRD	Negative regulatory domain
OD	Optical density

PBS	Phosphate-buffered saline
PCR	Polymerase chain reaction
PEG	Polyethylene glycol
PI	Propidium iodide
PSC	Pluripotent stem cell
RT	Room temperature
SCNT	Somatic cell nuclear transfer
SNAG	Snail/Gfi1
SSC	Side scatter
TAD	Transactivation domain
Tet	Tetracycline
TF	Transcription factor
TRE	Tet-response element
TSB	Transformation and storage buffer
tTat	Tet-transactivator
UV	Ultraviolet
WT	Wild type
ZF	Zinc finger

SYMBOLS

-	Negative
%	Percent (percentage)
°C	Degrees Celsius
+	Positive
A	Ampere
CO_2	Carbon dioxide
E	Embryonic day
g	Gram
h	Hour
H_2O	Water
L	Liter
m	Milli (as in milliter or milligram)
M	Molar (mol/dm3)
m/v	Mass/volume
$MgCl_2$	Magnesium chloride
$MgSO_4$	Magnesium sulphate
min	Minutes
n	Nano (as in nanogram)
rpm	Rotations per minute
s	Second
v/v	Volume/volume
α	Alpha
β	Beta
Δ	Delta
μ	Micro (as in microgram or microliter)

CHAPTER 1

1. Introduction

1.1 Hematopoietic stem cells

Hematopoietic stem cells (HSCs) are a rare population of multipotent, self-renewing cells that give rise to all differentiated blood cell types, through hematopoiesis (**Figure 1**) (Metcalf 1989).

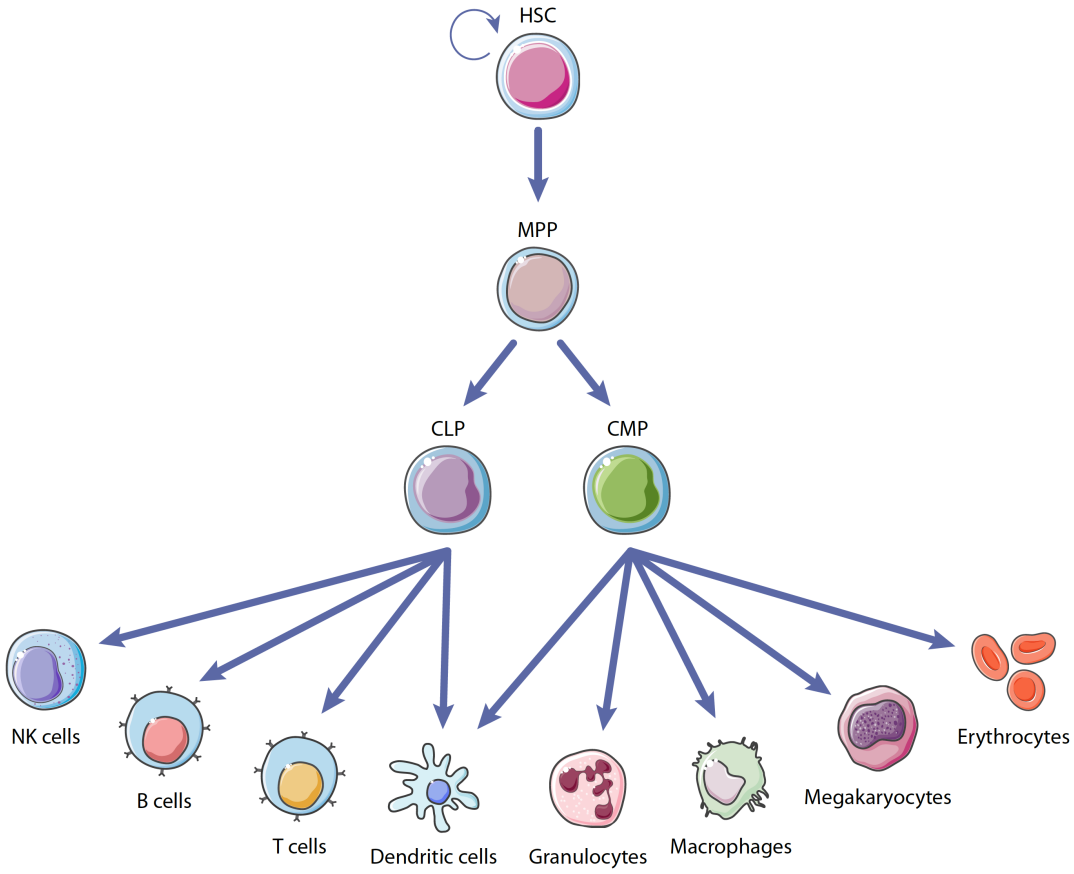

Figure 1. Hematopoiesis. Hematopoietic stem cells (HSCs) differentiate into all differentiated blood cells. Multipotent progenitors (MPPs) give rise to the lymphoid and myeloid lineages through common lymphoid progenitors (CLPs) and common myeloid progenitors (CMPs), respectively.

Hematopoiesis starts in embryonic development and is characterized by two major waves (Kaimakis et al. 2013). The first wave, primitive hematopoiesis, generates primitive short-lived erythroid and myeloid progenitors. The second wave, definitive hematopoiesis, appears later in development and generates *de novo* definitive HSCs (dHSCs) capable of self-

renewing and with the ability to differentiate into all adult blood cell lineages (erythroid, myeloid and lymphoid). In mice embryos, dHSCs appear first in the aorta–gonad–mesonephros (AGM) region and placenta at embryonic day (E) 10.5. Around E11.5, dHSCs migrate to the fetal liver to expand and finally colonize the bone marrow, at E16.5, where they mature and remain throughout adult life (**Figure 2**) (Medvinsky et al. 2011; Costa et al. 2012). In the AGM region, dHSCs arise from rare clusters of vascular endothelial cells named "hemogenic endothelium" (Bertrand et al. 2010). Imaging studies with mouse and zebrafish embryos suggested that HSCs were generated as a result of an endothelial-to-hematopoietic transition (EHT), though detailed characterization of HSCs direct precursor cells remains unclear (Boisset et al. 2010; Zovein et al. 2008).

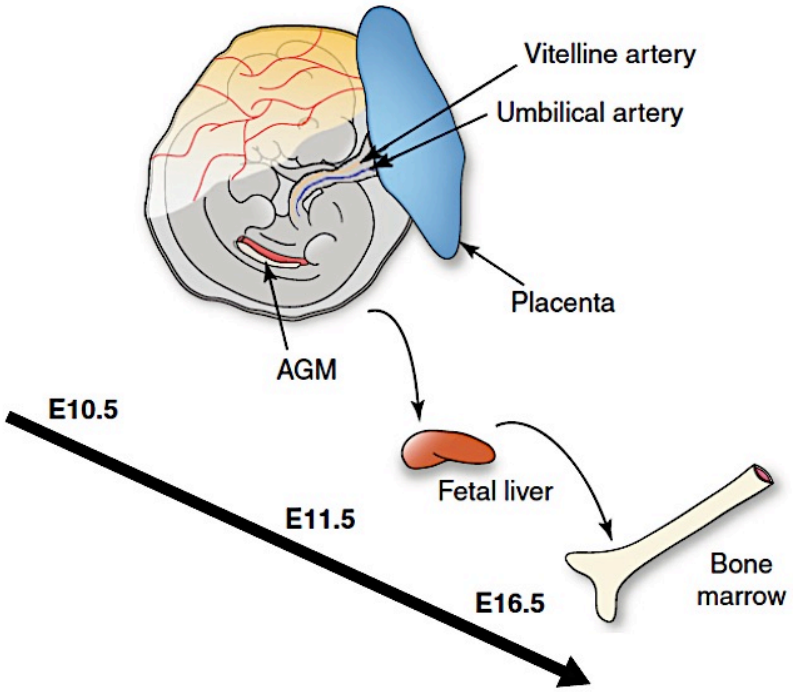

Figure 2. Mouse embryonic sites of definitive hematopoietic stem cells. At embryonic day (E) 10.5, definitive hematopoietic stem cells (dHSCs) are generated from the aorta–gonad–mesonephros (AGM) region and placenta, in mice. Then, dHSCs expand in the fetal liver around E11.5 and colonize the bone marrow at E16.5. *Adapted from Costa et al., (2012).*

HSCs identification has been possible due to the expression, or lack of expression, of a combination of cell surface markers (Berardi et al. 1995; Rossi et al. 2011). CD34 is a glycophosphoprotein expressed on hematopoietic stem/progenitor cells as well as on endothelial cells (Siena et al. 2000; Stella et al. 1995). The pool of human HSCs is CD34+, since CD34+ cells obtained from

bone marrow are able to reconstitute the hematopoietic system of humans and mice. In adult mammalians, HSCs reside predominantly in the bone marrow representing 0,01% of the nucleated cells (Suzuki et al. 2006; Rossi et al. 2011) and much lower amounts can also be found in peripheral blood (Jansen et al. 2005).

1.1.1 HEMATOPOIETIC STEM CELLS: CLINICAL APPLICATIONS

Constant replacement of blood cells during an individual's lifetime is possible due to an equilibrium between HSCs self-renewal and differentiation abilities, which makes this type of cells perfect candidates for regenerative medicine purposes. For this reason, hematopoietic stem cell transplantation (HSCT) has been used as treatment for a variety of hematological diseases since the late 50s. There are two main types of HSCT: autologous and allogeneic (Passweg et al. 2012). In autologous transplantation, HSCs are isolated from patients and stored until used. In allogeneic transplantation, HSCs are retrieved from related or unrelated donors. Data from the European Group for Blood and Marrow Transplantation 2013 annual survey revealed that in that year 39 209 transplants were performed (Passweg et al. 2015). Of those, 14 950 were allogeneic and 19 859 autologous. The major indications for HSCT include leukemia, lymphomas, myelomas and bone marrow failures (Apperley et al. 2016; Passweg et al. 2012).

Currently, sources of HSCs include bone marrow and peripheral blood (after mobilization), as well as umbilical cord blood, although the number of cells retrieved from the latter source is still limited and insufficient to treat an adult (Smith & Wagner 2009). With regard to allogeneic transplantation, the procedure requires human leukocyte antigen (HLA) gene matching between the host and the donor, and despite best clinical outcomes are observed when the donor is an HLA-matching sibling, each patient's sibling has only 25% chance of having identical HLA alleles (Apperley et al. 2016). Consequently, allogeneic HSC transplantation is usually associated with graft rejection by the host immune system, graft-versus-host disease (GVHD) caused by an immune response of contaminating T cells from transplanted tissues against the host, and high risk of transplant-related mortality (Szydlo et al. 1997). In terms of autologous transplantation, even though there is no graft rejection or GVHD, there is a risk of graft contamination with cancer cells that can provoke relapse of the disease (Forman & Nakamura 2011). Furthermore, it has very little to no clinical applicability in the treatment of inherited hematopoietic disorders.

In order to surpass HSCT limitations and to obtain a sufficient cell number, efforts have being made towards the *in vitro* expansion of HSCs for transplantation (Takizawa et al. 2011). Attempts to expand and maintain HSCs

in vitro have met limited success, mostly because of the lack of suitable culture conditions (Schuster et al. 2012). Experimental data implies that HSCs lose their self-renewal potential with continual divisions and long culture periods. Indeed, the number of multipotent cells decreases as the number of daughter-cells increases (Ema et al. 2000). Therefore, several groups have been working on different approaches to generate engraftable patient-specific HSCs from alternative cell sources. It has been attempted to generate HSCs upon inducing the differentiation of pluripotent stem cells (PSCs), but with limited success despite of 20 years of research (Daniel et al. 2016). PSCs are easily expanded and maintained *in vitro*, however differentiation protocols resulted in hematopoietic cells that engrafted poorly (Sturgeon et al. 2013; Choi et al. 2012; Rafii et al. 2013). These cells were more similar to short-lived progenitors generated during primitive hematopoiesis, than to *bona fide* HSCs (Vo & Daley 2015). To date, it has not yet been possible to generate fully functional dHSCs from PSCs.

Additional strategies to generate dHSCs *in vitro* are needed and direct cellular reprogramming technologies from somatic cells offer an exciting alternative.

1.2 Cellular development and reprogramming

During embryonic development, pluripotent stem cells differentiate into each and every tissue-specific cell type of multicellular organisms. The maintenance and transmittance of cell fate is controlled by complex transcriptional and epigenetic mechanisms (Allis & Jenuwein 2016; Goldberg et al. 2007). Epigenetics can be described as the study of heritable changes in gene expression profiles without altering the underlying DNA sequence. Until the first half of the 20th century, lineage commitment and cell differentiation were seen as a unidirectional and irreversible process, as defined by Conrad Waddington's "epigenetic landscape" model (Ladewig et al. 2013). In his model, cells metaphorically behaved like marbles rolling down a hill, separating into different paths until they reached their final destination, in other words, a differentiated cell state. From this angle, one may consider cell differentiation an epigenetic process itself, since starting from one genotype, multicellular organisms develop various cell types with distinct gene expression patterns and functions (Goldberg et al. 2007). However, the idea of differentiation as an one-way process was challenged by John Gurdon's pioneering work on somatic cell nuclear transfer (SCNT), in frogs (Briggs & King 1952; Gurdon 1962a; Gurdon 1962b). In one of his experiments, nuclei from differentiated tadpole intestinal cells were transferred into enucleated eggs, resulting in adult frogs that were genetically identical to the respective somatic cell nucleus

donor (Gurdon 1962b). Only later, in the 90s, was SCNT used to clone the first mammalians (Wakayama et al. 1998; Wilmut et al. 1997).

In the course of the second half of the 20th century, another line of research came to light, this time focused on the fusion of two different cell types to evaluate changes in gene expression profiles. The fusion of mouse muscle cells with human amniotic cells produced non-dividing heterokaryons that expressed human muscle proteins (Blau et al. 1983), demonstrating for the first time that silent genes could be activated in cells where they are normally not expressed. It was not until the 21st century that scientists were able to reprogram somatic cells to pluripotency through fusion with embryonic stem cells (ESCs) (Tada et al. 2001; Cowan et al. 2005). These stem cells, which are found in the inner mass of mammalian blastocysts, can divide indefinitely in culture and, since they are pluripotent, differentiate into cells of all embryonic germ layers (endoderm, mesoderm and ectoderm) (Evans & Kaufman 1981; Martin 1981). Tada (Tada et al. 2001), Cowan (Cowan et al. 2005) and Pereira (Pereira et al. 2008) reported that, not only somatic cell-ESC hybrids and heterokaryons differentiated into cells of the three germ layers, they also expressed pluripotent genes typical of ESCs. Overall, studies on SCNT and cell-fusion with ESCs have shown that the differentiated state of somatic cells was not static or irreversible. Instead, it could be reprogramed back to the embryonic state, indicating that enucleated eggs and ESCs held factors that were capable of rewriting the epigenetic networks in control of cell identity (Egli et al. 2008) Nonetheless, the underlying mechanisms responsible for the shift in cell fate remained poorly understood.

In 2006, Yamanaka and colleagues postulated that the factors that were involved in maintaining ESCs' stemness could be sufficient to induce pluripotency in somatic cells. By simply overexpressing four transcription factors (TFs) Oct4, Sox2, c-Myc, and Klf4 Yamanaka's group was able to reprogram fibroblasts into induced pluripotent stem cells (iPSCs) (Takahashi et al. 2007; Takahashi & Yamanaka 2006). Although this was a breakthrough in the cell reprogramming field, there are several limitations regarding the clinical applicability of iPSCs in human therapies due to the risk of tumorigenesis.

In summary, cell reprogramming can be referred as the act of changing the cellular identity of one cell type to another, by altering epigenetic states and transcription profiles (Hanna et al. 2008).

1.2.1 Direct cellular reprogramming

Since Yamanaka's group findings were published, following studies have established TFs combinations that could directly reprogram one somatic cell lineage into another, without passing through a pluripotency state (Xu et al.

2015). Direct reprogramming strategies using mainly fibroblasts have been applied to obtain several different cell types such as cardiomyocytes (Ieda et al. 2010; Protze et al. 2012; Inagawa et al. 2012), neurons (Vierbuchen et al. 2010; Liu et al. 2012; Son et al. 2011), hepatocytes (Sekiya & Suzuki 2011; Huang et al. 2011) and even multipotent stem cells (i.e., hepatic stem cells (Yu et al. 2013) and neural stem cells (Ring et al. 2012)).

In the hematopoietic system, reprogramming strategies were also implemented to generate both specialized and stem/progenitor cells (Capellera-Garcia & Flygare 2017) (**Figures 3 and 4**). Macrophages were obtained from both fibroblasts (Feng et al. 2008) and B cells (Bussmann et al. 2009) using PU.1 plus C/EBPα/β or C/EBPα alone, respectively. The reprogrammed cells acquired macrophage-like phenotype, morphology and function. Sadahira *et al.* reprogrammed differentiated mature B cells into erythroid cells, by enforcing the expression of Gata1, Scl and C/EBPα factors (Sadahira et al. 2012). Later on, Gata1, Tal1, Lmo2, and c-Myc were found to convert murine and human fibroblasts to primitive-like erythroid progenitors and precursors (Capellera-Garcia et al. 2016). Addiction of Klf1 or Myb to the previous TFs combination resulted in the expression of adult hemoglobin in reprogrammed cells. Moreover, a study from the same group has shown that *bona fide* megakaryocyte progenitors were obtained after overexpressing the earlier four TFs (Gata1, Tal1, Lmo2, and c-My) plus Gata2 and Runx1, which biased the reprogramming process toward the megakaryocyte lineage (Pulecio et al. 2016). Nonetheless, generating hematopoietic stem-like cells still stands as the holy grail of regenerative medicine for the treatment of blood diseases.

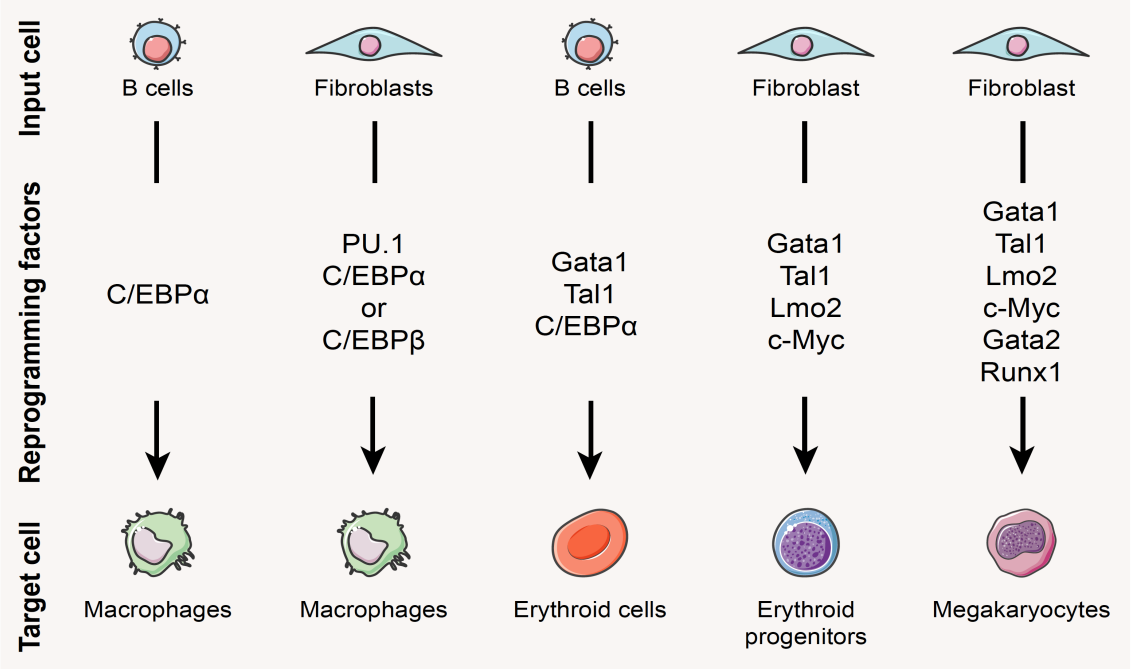

Figure 3. Examples of direct reprogramming strategies applied in the hematopoietic system. Lineage-specific transcription factors reprogramed B cells to macrophages and erythroid cells, and fibroblasts to macrophages, erythroid progenitors and megakaryocytes.

In 2013, Pereira and colleagues used the same approach to induce a hemogenic program in mouse embryonic fibroblasts (MEFs) from double transgenic mice (hCD34tTAxTetO-H2BGFP, to shorten hCD34/H2BGFP) (**Figure 4**) (Pereira et al. 2013). In this model, expression of histone H2B coupled with green fluorescent protein (GFP) is controlled by the activation of the human CD34 (hCD34) promoter. The acquisition of a hematopoietic or endothelial cell fate prompts huCD34 reporter activation and subsequently nuclear expression of GFP. Together, Gata2, Gfi1b and cFos activated the hCD34/H2BGFP reporter, and reprogramed fibroblasts into hemogenic endothelial-like precursor cells from which hematopoietic colonies emerged, albeit with low efficiency. Hemogenic precursors (HPs) exhibited a Prominin1+ Sca1+ CD34+ CD45- cell surface phenotype and endothelial-like transcriptional programs. Notably, budding hematopoietic cells expressed HSC markers and gene-expression profile (Pereira et al. 2013). Mechanistically, it is thought that Gata2 binds first to its targets genes and then recruits the other factors to both inhibit the expression of fibroblast genes and enable the expression of endothelial and hematopoietic genes (Pereira *et al*., in preparation). Hemogenic/Hematopoietic reprogramming methodology has also been proven in the human system with human fibroblasts and the human versions of the same set of TFs (Pereira et al, in preparation).

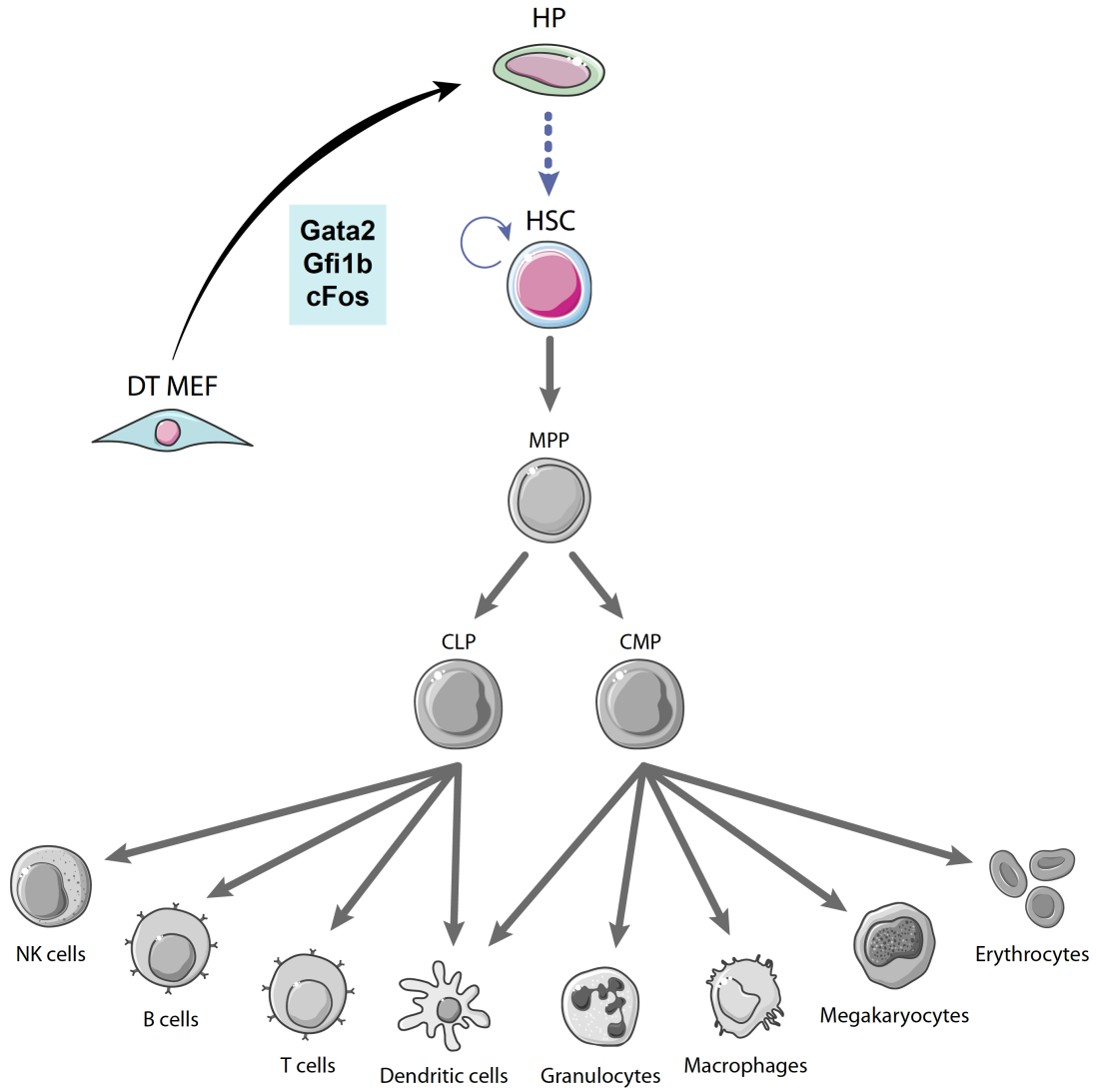

Figure 4. Direct reprogramming of mouse embryonic fibroblasts (MEFs) into hemogenic precursors (HPs). Gata2, Gfi1b and cFos transcription factors were sufficient to directly reprogram double transgenic (DT, hCD34/H2BGFP) MEFs into hematopoietic stem-like cells through hemogenic precursor cells.

This transition between cell types resembled the EHT that occurs in the embryo and placenta, during specification of definitive hematopoiesis. Following work of the same author helped to unveil these same hemogenic precursors *in vivo* (Pereira et al. 2016). Using mouse placentas, Pereira *et al.* have identified a population of cells with similar phenotypes and global gene expression. Ultimately, the hematopoietic reprogramming of fibroblasts provided an *in vitro* system for the study of a developmental process that occurs *in vivo*, and that would hardly be informed by other means.

Subsequent studies have also aimed to derive hematopoietic stem and progenitor cells (HSPCs) from fibroblasts (Batta et al. 2014; Cheng et al. 2016),

non-hemogenic endothelial cells (Sandler et al. 2014; Lis et al. 2017) and lineage committed blood cells (Riddell et al. 2014), employing different culture systems and sets of TFs. Nevertheless, starting with committed blood precursors, rather than unrelated cell type such as fibroblasts, might be unviable when patients suffer from hematological disorders caused by mutations in progenitor or stem cell pools, as mutations are passed on to the differentiated progeny (Pereira & Moore 2014). A novel hybrid strategy consisting of PSC differentiation in hemogenic endothelium and subsequent reprogramming into HSPCs was currently accomplished (Sugimura et al. 2017).

1.3 GATA2, GFI1B AND cFOS ROLE IN HEMATOPOIESIS

As reported by our group, Gata2, Gfi1b and cFos represented the minimal TF pool for (hCD34/H2BGFP) reporter activation and establishment of hematopoietic colonies from induced HPs (Pereira et al. 2013). The removal of any of the three factors impaired the generation of GFP+ colonies, proving their absolute need for hemogenic/hematopoietic reprogramming. These results are in agreement with what is known about the function of these factors in hematopoiesis.

Gata-binding protein 2 (Gata2) is a member of the GATA family of TFs named after the consensus nucleotide sequence (A/T)GATA(A/G) to which they bind in promoter and enhancer regions of target genes (Orkin 1992; Ko & Engel 1993; Wu et al. 2007). In the hematopoietic system, the Gata2 protein is highly expressed in mast cells (Tsai & Orkin 1997), megakaryocytes (Visvader & Adams 1993) and in the immature hematopoietic cell compartment, specially in HSCs, where it controls cell quiescence, self-renew and proliferation (Minegishi et al. 2003; Minegishi et al. 1999; Suzuki et al. 2006; Tsai & Orkin 1997; Ling et al. 2004; Kosan & Godmann 2016). Gata2 knockout/mutation impairs definitive hematopoiesis and embryos survival, since $Gata2^{-/-}$ mutant mouse embryos do not survive past embryonic day 11.5 (Tsai et al. 1994).

The growth factor independent 1b or Gfi1b is a transcriptional repressor fundamental for the development of megakaryocytic and erythroid blood cell lineages (Saleque et al. 2002; Foudi et al. 2014). Experiments in mice demonstrated that $Gfi1b^{-/-}$ embryos died due to failure in yielding definitive erythrocytes. More importantly, Gf1ib is a central regulator of the EHT, during embryogenesis (Lancrin et al. 2012). Together with Gfi1, Gfi1b represses endothelial genes and allows the emergence of HSCs, where its expression remains increased (Khandanpour et al. 2011; Thambyrajah et al. 2015).

cFos, also known as FBJ osteosarcoma oncogene, dimerizes with c-Jun protein to form the AP-1 (activator protein-1) TF complex (Halazonetis et al.

1988). The AP-1 complex acts as a general regulator of cell proliferation and differentiation with specific functions in bone, cartilage and in the hematopoietic compartment (Angel & Karin 1991; Wagner 2005). In the hematopoietic system, AP-1 controls the onset and development of myeloid differentiation. When it comes to cFos, cFos null mice have abnormal hematopoiesis and exhibit complications in bone remodeling (Zhao-Qi et al. 1992). Plus, in the absence of cFos, mice placentas showed low levels of HSC activity (Ottersbach & Dzierzak 2005).

Despite the data available regarding the normal function of each of the three TFs during hematopoiesis, specific protein features responsible for hemogenic/hematopoietic reprogramming are yet to be determined.

1.4 Defining GATA2, GFI1B and cFOS features for hematopoietic reprogramming

Although successful, direct conversion of fibroblasts to hemogenic/hematopoietic cells, by overexpression of Gata2, Gfi1b and cFos exhibited rather low reporter activation efficiency with only 2,82% of GFP+ cells. In order to optimize or create strategies to increase reporter activation and HSC-induced reprogramming, first it is important to assess relevant features of the three TFs for this particular reprogramming process. Such features can be directly associated with the protein domain organization, which depends on the genetic background of the protein, or with epigenetic (non-genetic) functions that these factors may display in the context of cell reprogramming.

1.4.1 Genetic features: protein domains

Methodologies to dissect the molecular function of reprogramming factors have been attempted by a substitution approach with divergent factors and/or homologous factors (Soufi 2014; Montserrat et al. 2013; Shu et al. 2013). Genes that share a common ancestral DNA (homologous genes) can be divided in two subclasses: homologous genes related by gene duplication are called *paralogs*, while genes resulting from speciation are considered *orthologs* (Fitch 1970). In this context, for example, Gata1 and Gata2 within a species are paralogs, whereas Gata2 in mouse and Gata2 in human are orthologs.

Inside the GATA family of transcription factors (Gata1-6), Gata1 and its paralog proteins Gata2 and Gata3 are considered hematopoietic factors (Vicente et al. 2012; Orkin 1992). Gata1 is a major regulator of the erythroid

lineage. It is highly expressed in cells of the myeloid lineage (primitive and definitive erythroid cells, megakaryocytes and mast cells) and is particularly important for the differentiation of hematopoietic progenitors into definitive erythrocytes and megakaryocytes (Romeo et al. 1990; Martin et al. 1990; Evans & Felsenfeld 1989; Takahashi et al. 1998). Gata1 null ESCs cannot differentiate into erythrocytes *in vitro* (Simon et al. 1992). Gata3 is abundant in T cells and thymocytes (Ko et al. 1991; Ho et al. 1991; Joulin et al. 1991; Ting et al. 1996). This protein regulates the expression of T-cell specific genes such as TCR alpha and delta genes. Gata3 null ESCs fail to differentiate into thymocytes and mature peripheral T lymphocytes (Ting et al. 1996). GATA factors ability to bind to DNA resides in two identical zinc finger (ZF) regions of the Cys-X_2-Cys-X_{17}Cys-X_2-Cys type. Although ZFs are highly conserved among GATA proteins, the non-finger domains are quite different, which may explain their unique functions (Pan et al. 2000; Rodrigues et al. 2012). In addition to the ZFs, Gata2 has two transactivation domains (TADs), located in the N- and C-terminal, a nuclear localization signal (NLS) and a negative regulatory domain (NRD) (Vicente et al. 2012; Minegishi et al. 2003), whereas Gata1 and Gata3 have just two TADs each (**Figure 5**) (Kaneko et al. 2012; Yang et al. 1994). Gata1 and Gata3 nuclear localization signals reside within each respective DNA binding domain and adjacent regions (Yang et al. 1994; Shimizu et al. 2001)

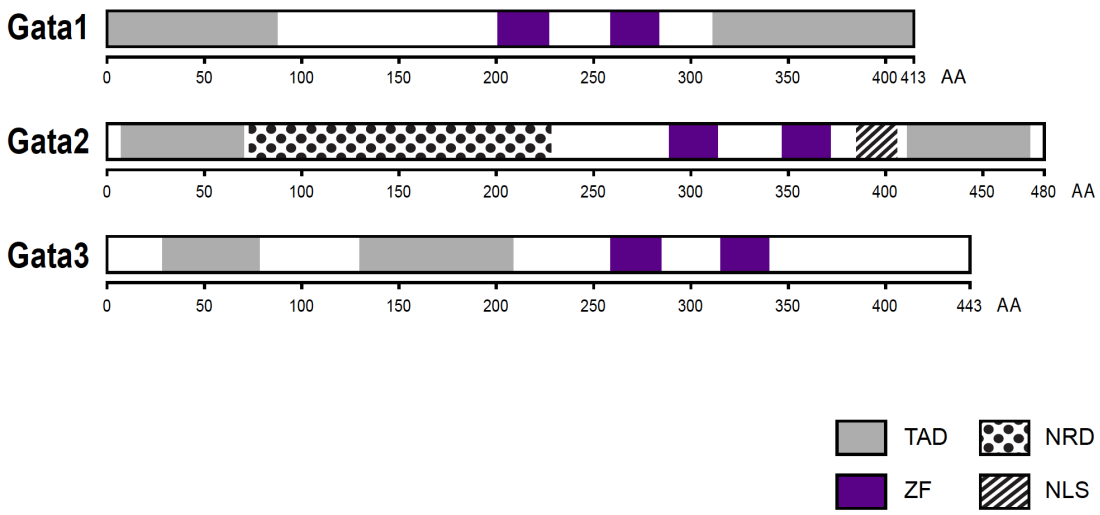

Figure 5. Schematic representation of GATA hematopoietic factor Gata1, Gata2 and Gata3 proteins. GATA factors have two conserved zinc finger (ZF) regions, which form the DNA-binding domain. Additionally, Gata2 has two transactivation domains (TADs), a nuclear localization signal (NLS) and a negative regulatory domain (NRD). Gata1 and Gata3 have two TADs each, in addition to the ZFs.

Growth factor independent 1 (Gfi1) and its paralog Gfi1b have similar structures and mechanisms of action. Even though these proteins play partially overlapping roles, they also show distinct cell expression patterns and particular functions, during hematopoiesis (Möröy et al. 2015). Both proteins are expressed in HSPCs, but Gfi1b expression is much higher in the earliest HSC compartment when compared to Gfi1 (Khandanpour et al. 2011). Upon differentiation to MPPs, Gfi1 expression is progressively up regulated, while Gfi1b expression is gradually down-regulate. Gfi1 is also expressed in common lymphoid progenitors (CLPs) (Zeng et al. 2004) neutrophils (Hock et al. 2003), thymic progenitors, early B-cells and during T-cells development (Yücel et al. 2004; Schmidt et al. 1998). Gfi1 and Gfi1b share two major domains with over 95% homolgy: a repressor SNAG (Snail/Gfi1) domain in the N-terminal and a DNA-binding domain with six highly conserved Cys_2His_2 type ZFs in the C-terminal region (**Figure 6**) (Zweidler-Mckay et al. 1996; Grimes et al. 1996; Tong et al. 1998). The N-terminal region, the SNAG domain binds to the histone-modifying proteins lysine-specific histone demethylase 1A (KDM1A also known as Lsd1) and RCOR1/2 (CoREST1/2), causing the transcriptional repression of target genes (Saleque et al. 2007). In the C-terminal region, ZFs 1, 2, and 6 are required for protein interaction, whereas zinc-fingers 3 to 5 are necessary for DNA binding (Lee et al. 2010; Zweidler-Mckay et al. 1996). The two domains are separated by a less characterized middle region, responsible for the size difference between proteins and possibly for the distinctive cell functions they execute.

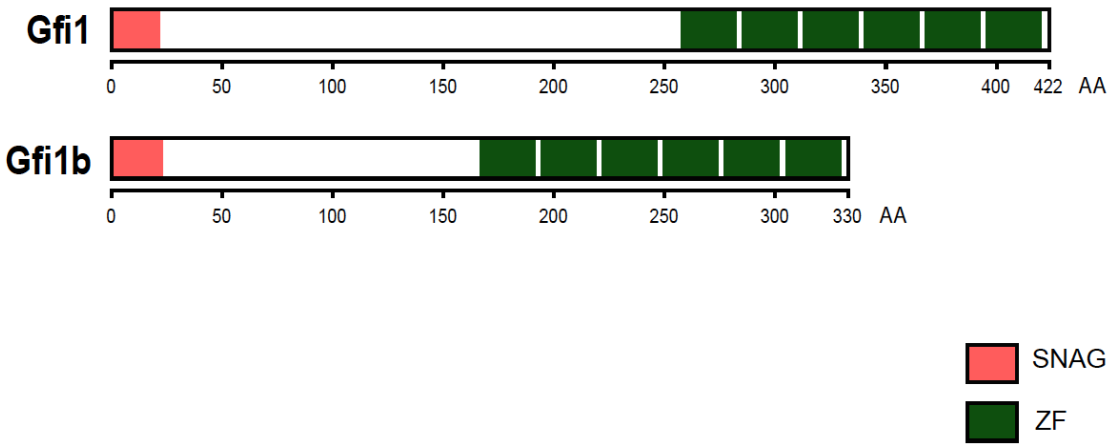

Figure 6. Schematic representation of the repressor transcription factors Gfi1 and Gfi1b. Gfi1 and Gfi1b share a repressor SNAG (Snail/Gfi1) domain in the N-terminal and six conserved zinc fingers (ZFs) in the C-terminal DNA-binding domain, separated by a non-homologous middle region.

The AP-1 transcription factor is a protein complex consisting of Jun and Fos heterodimers (Angel & Karin 1991; Halazonetis et al. 1988). This association was first discovered between cJun and cFos proteins, but it is also observed among other Jun proteins (Nakabeppu et al. 1988) and cFos paralogs FosB (Zerial et al. 1989), Fra1 (Cohen et al. 1989) and Fra2 (Nishina et al. 1990). It is thought that Fos proteins regulate gene expression by using Jun proteins as anchors to bind to regulatory regions of target genes, since Fos proteins cannot bind to DNA on their own (Angel & Karin 1991). All AP-1 proteins share a highly conserved basic leucine zipper (bZIP) domain, which comprises two adjacent genetic regions, a basic motif that interacts with DNA and a leucine-zipper region for protein dimerization (Sassone-Corsi et al. 1988; Turner & Tjian 1989). However, unlike cFos and FosB, that have an additional conversed C-terminal TAD, the smaller paralogs Fra1 and Fra2 lack this region (**Figure 7**) (Wisdon & Verma 1993; Milde-Langosch 2005). The nuclear localization signal in Fos proteins is not restricted to only one protein region as several regions in N and C-terminal are required for that purpose (Campos et al. 1999).

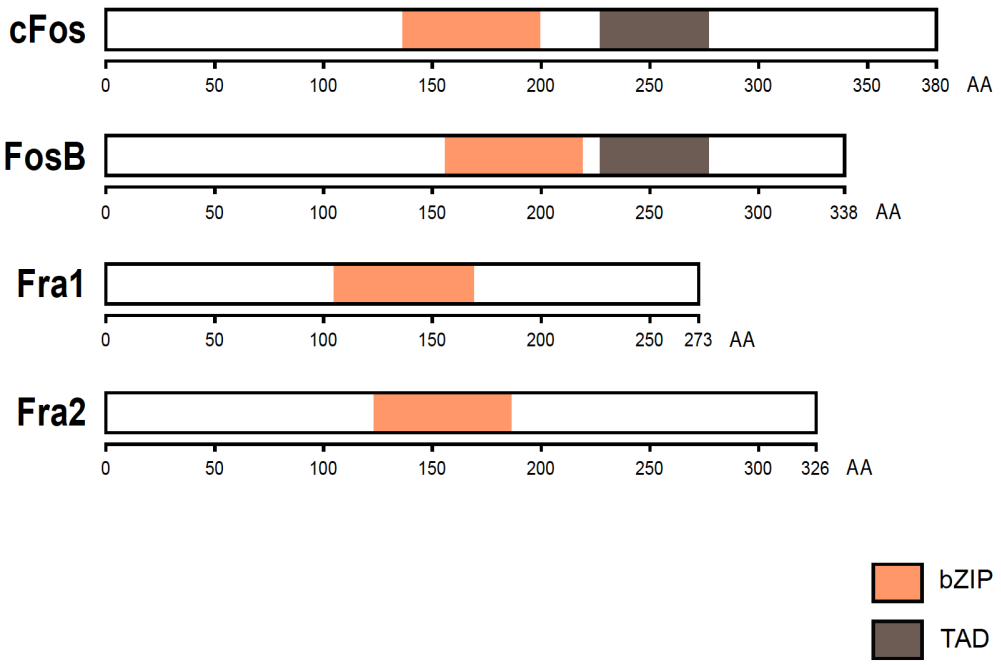

Figure 7. Schematic representation of Fos family of transcription factors cFos, FosB, Fra1 and Fra2 proteins. Fos proteins display a highly conserved basic leucine zipper (bZIP) domain. cFos and FosB, that have an additional transactivation domain (TAD) that is not present in Fra1 and Fra2.

By replacing Gata2, Gfi1b and cFos with paralog factors one at the time in hematopoietic reprogramming experiments, one can compare protein

domains and infer which regions may be fundamental for the reprogramming function exerted by the three TFs. In fact, Gfi1 and FosB were part of a TF cocktail used to directly reprogram non-hemogenic endothelial cells into engraftable hematopoietic stem and progenitor-like cells (Sandler et al. 2014; Lis et al. 2017). Moreover, Sox1 and Sox3 were able to replace Sox2 for reprogramming to pluripotency, showing that reprogramming factors may be replaced by homologous proteins in reprogramming studies (Montserrat et al. 2013; Shu et al. 2013).

A complementary strategy to do structure-function analysis is through the creation of deletion mutants, in other words, protein constructs lacking specific sequences of aminoacids. This strategy was initially applied to study the established roles of GATA (Yang & Evans 1992; Visvader et al. 1995; Shimizu et al. 2001; Minegishi et al. 2005; Minegishi et al. 2003; Kaneko et al. 2012; Yang et al. 1994), Gfi1/1b (Tong et al. 1998; Grimes et al. 1996; Zweidler-Mckay et al. 1996) and AP-1 protein domains (Gentz et al. 1989; Sassone-Corsi et al. 1988; Turner & Tjian 1989). Gata2 appears to be the most critical TF for our reprogramming system since, in addition to recruiting other factors, it is the most dominant TF i.e., binds to the majority of its targets in the genome independently of Gfi1b and cFos (Pereira *et al.,* in preparation). Thus, it would be interesting to focus on unveiling Gata2 domains for hematopoietic reprogramming through this approach.

1.4.2 Epigenetic features: mitotic bookmarking

As previously stated, throughout hematopoietic reprogramming, reprogrammed cells ultimately establish hCD34/GFP+ hematopoietic colonies, suggesting that cell division might be necessary for the success of this particular reprogramming process (**Figure 8**) (Pereira et al. 2013). So far, it remains unclear how transcriptional programs of GFP+ cells are established and inherited.

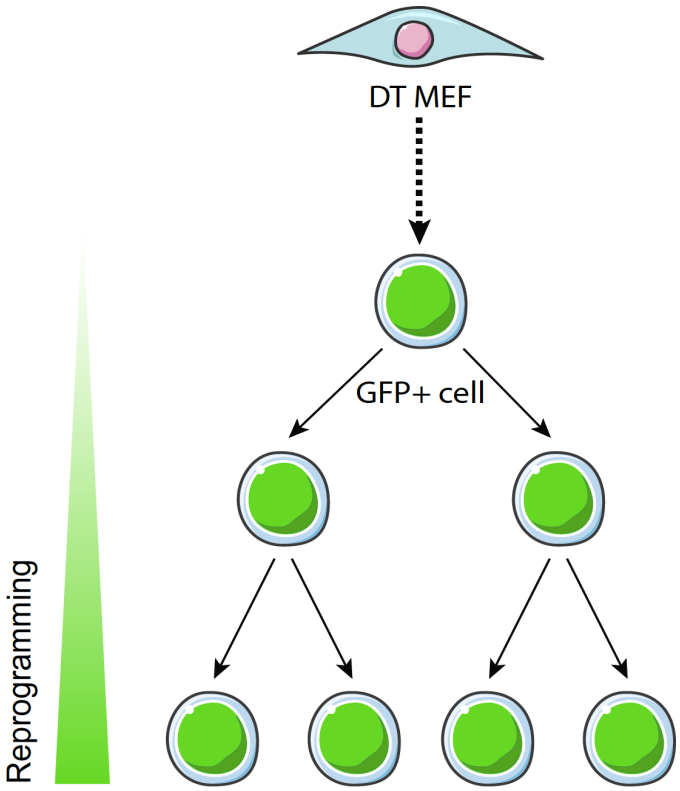

Figure 8. Cell division and morphologic changes in mouse embryonic fibroblast (MEF) during hematopoietic reprogramming. Reprogrammed cell activate hCD34/H2BGFP reporter in double transgenic (DT) MEFs and start to express nuclear green fluorescent protein (GFP). Hematopoietic-like stem cells (GFP+ cells) exhibit green nuclei. The number of reprogrammed cells and GFP levels increase with cell division.

During mitosis, cells undergo dramatic changes in nuclear organization and gene expression. Nuclear envelope disintegrates, interphase chromatin condenses, RNA polymerase detaches from chromosomes and TFs disperse across the cytoplasm, leading to a global silencing in gene transcription (Prescott & Bender 1962; Martínez-Balbás et al. 1995; Taylor 1960; Terasaki et al. 2001; Gottesfeld & Forbes 1997). Following mitosis and nuclear reassembly, previous transcriptional patterns of gene activation and repression must be re-established in daughter cells, according to lineage phenotype (Egli et al. 2008). These transitions between different states of gene expression impose a challenge for the preservation of cell identity. Therefore, several epigenetic mechanisms should be put in action to ensure the correct transmittance of lineage commitment (Margueron & Reinberg 2010; Allis & Jenuwein 2016). These include DNA methylation, histone post-translational modifications, histone variants and small interfering RNA-mediated gene silencing. DNA methylation and histone modifications are two well-characterized epigenetic

mechanisms that control gene expression without altering DNA sequences, and are inherited through cell division (Wigler 1981; Nowak & Corces 2004; Trojer & Reinberg 2006). Reports on histone acetylation and methylation in nucleosomes proposed that these mechanisms could function as keepers of cell memory for propagating gene expression patterns to daughter cells, by marking particular chromatin sites and contributing for gene reactivation, during the mitotic phase of the cell cycle (Jeppesen 1997; Valis et al. 2005; Kouskouti & Talianidis 2005).

These findings brought to light an additional epigenetic mechanism that because it resembled the way a bookmark marks the last read page of a book was entitled “mitotic bookmarking” (Michelotti et al. 1997; Sarge & Park-Sarge 2009; Lodhi et al. 2016). Of note, mitotic bookmarking of genes is not a mechanism exclusive of histone modifications (Kadauke & Blobel 2013; Liu et al. 2017; Lodhi et al. 2016). Contrary to previous data, several general and lineage-specific TFs (**Table 1**) were shown to remain bound to specific chromatin regions during mitosis. Interestingly, Oct4, Sox2, Klf4 (Takahashi et al. 2007; Takahashi & Yamanaka 2006) and Esrrb (Feng et al. 2009) are TFs that were implicated in somatic cell reprogramming to pluripotency. Also, Gata1 (Capellera-Garcia et al. 2016), Gata4 (Ieda et al. 2010; Protze et al. 2012; Inagawa et al. 2012) and Hnf1β (Yu et al. 2013) were used to directly reprogram fibroblasts into erythroid progenitor cells, cardiomyocyte-like cells and bipotent hepatic stem-like cells, respectively. Mitotic bookmarking by TFs was proposed to enable rapid target gene activation entering interphase (G1) in newborn cells, and the preservation of cell fate (Kadauke & Blobel 2013). Several studies revealed that the depletion of factors that are retained on mitotic chromatin retarded transcription reactivation of target genes upon mitotic exit and impaired normal cell fate commitment (Kadauke et al. 2012; Caravaca et al. 2013; Yang et al. 2008; Dudek et al. 2009; Young et al. 2007; Deluz et al. 2016; Liu et al. 2017).

Table 1. Transcription factors with mitotic bookmarking function.

General TFs	Lineage-specific TFs
TFIID (Christova & Oelgeschläger 2002)	**NFE2** (Xin et al. 2007)
TFIIB (Christova & Oelgeschläger 2002)	**Runx2** (Young et al. 2007)
HSF2 (Xing et al. 2005)	**Hnf1β** (Verdeguer et al. 2010)
FoxI1 (Yan et al. 2006)	**Gata1** (Kadauke et al. 2012)
TBP-PP2A (Xing et al. 2008)	**FoxA1** (Caravaca et al. 2013)
PARP1 (Lodhi et al. 2014)	**Gata4** (Caravaca et al. 2013)
RBPJ (Lake et al. 2014)	**Sox2** (Teves et al. 2016; Deluz et al. 2016)
CTCF (Shen et al. 2015)	**Oct4** (Deluz et al. 2016)
	Esrrb (Festuccia et al. 2016)
	Klf4 (Liu et al. 2017)

The hematopoietic transcription factor Gata1, a major regulator of the erythroid lineage and a paralogs of Gata2, remains bound to a subset of its target genes, during mitosis (Kadauke et al. 2012). In order to define the mitotic bookmarking purpose of Gata1, Kadauke *et al.*, created fusion proteins of Gata1 with a mitotic degron (MD), targeting the destruction of Gata1 precisely in mitosis. Mitotic-specific degradation of Gata1 in erythroid cells led to delayed restart of bookmarked genes expression, along with increased transcription of Gata1 repressed genes, such as Gata2 and Kit, which are typically present in immature cell compartments. These findings suggest that Gata1 occupancy of hematopoietic genes in mitotic chromatin favours their prompt transcription so that lineage-specific gene expression patterns remain constant. Furthermore, the destruction of the pluripotent regulators Sox2 and Oct4 at the metaphase-to-anaphase transition compromised their ability to maintain pluripotency (Deluz et al. 2016; Liu et al. 2017). Importantly, the role of Oct4 mitotic bookmarking for somatic cell reprogramming to pluripotency was also investigated (Liu et al. 2017). Ectopic expression of Oct4 marked for mitotic destruction (similar to MD-Gata1 fusion proteins) together with Sox2, Klf4 and cMyc, in MEFs, resulted in a lower number of iPSCs colonies plus defects in upregulating early pluripotent markers, when compared to controls.

Mitotic bookmarking, particularly by transcription factors, has proven to be an important epigenetic mechanism in transmitting transcriptional information to daughter cells, as well as in allowing cell reprogramming. Hence, I hypothesise that Gata2 acts as a mitotic bookmarker to convey newly acquired transcriptional programs of cells to their reprogrammed progeny.

1.5 CHALLENGE AND AIMS

Treatment of several life-threatening blood diseases such as leukemia, myelomas or aplastic anemia relies on the success of HSCT to reconstruct a healthy hematopoietic system. However clinical complications must be taken into account when performing this technique. On one hand, using stem and/or progenitor cells from patients may not be feasible if the disease is inherited or if they already lack healthy cells. On the other hand, many patients do not have proper matched donors and even when they do, aggressive immune responses can occur during the process. Ideally, healthy autologous HSCs would be obtained for every patient for transplantation. In that sense, direct reprogramming of fibroblasts to hematopoietic stem-like cells emerges as an appealing alternative to currently performed therapies. Nevertheless, this technology is still taking its first steps and a lot has to be done to optimize the overall process and its efficiency.

The main challenge of this work is to design enhanced transcription factors to improve hematopoietic reprogramming efficiency in order to enable clinical application. To do so, I first aim to uncover Gata2, Gfi1b and cFos functional features for hematopoietic reprogramming of fibroblasts. The central goal encompass the follow specific objectives:

- Address whether Gata2, Gfi1b and cFos paralog genes can substitute for hematopoietic reprogramming;
- Define Gata2 hematopoietic reprogramming domains using deletion constructs;
- Investigate whether Gata2 hematopoietic reprogramming factor acts as a mitotic bookmarker;

This work will provide valuable insights on how Gata2, Gfi1b and cFos impose and maintain HSC fate by defining critical characteristics for hemogenic/hematopoietic reprogramming. In the future, overexpression of improved TFs might be translated to the clinic to generate HSCs in sufficient numbers for autologous transplantation.

CHAPTER 2

2. MATERIALS AND METHODS

2.1 EXPERIMENTAL APPROACH

To address the proposed objectives, experimental strategies were divided in three major groups:

a) Hematopoietic reprogramming with Gata2, Gfi1b and cFos paralogs. Includes cloning of Gata2, Gfi1b and cFos paralogs, lentiviral production and transduction of DT MEFs for GFP quantification by flow cytometry.
b) Hematopoietic reprogramming with Gata2 deletion constructs. Includes cloning of Gata2 deletion constructs, lentiviral production and transduction of DT MEFs for GFP quantification by flow cytometry.
c) HEK293T cell transfection with Gata2 fusion constructs for mitotic bookmarking assessment. Includes cloning of mCherry either at 5' end or 3' end of Gata2, mitotic cell cycle arrest, DNA staining and fluorescence microscopy analysis.

2.2 CLONING PROCESS

2.2.1 PHAGE2-MCS – LENTIVIRAL EXPRESSION SYSTEM

The pHAGE2-MCS vector used in this study was modified in our lab from pHAGE2-EF1αFull-hOct4-F2A-hKlf4-IRES-hSox2-P2A-hcMyc-W-loxP vector, also known as STEMCCA (Sommer et al. 2009). pHAGE2-MCS is a constitutive vector that has a multiple cloning site (MCS), which includes the restriction sites for NotI, MfeI, NheI, HpaI, XbaI, BamHI restriction enzymes, preceded by the human EF1α promoter. Restriction sequences of these enzymes were used to design primers for gene cloning. Furthermore, it also includes an internal ribosome entry site (IRES) followed by a puromycin resistance gene (selectable mark for eukaryotic cells) and an ampicillin resistance gene for bacterial selection. A similar pHAGE2-MCS-IRES-eGFP was used in control conditions. These vectors (transfer plasmids) are included in a 2nd-generation lentiviral system composed by other two plasmids, as shown in **Figure 9**: the packaging plasmid psPAX2 (Addgene, #12260) and envelope plasmid pMD2.G (Addgene, #12259).

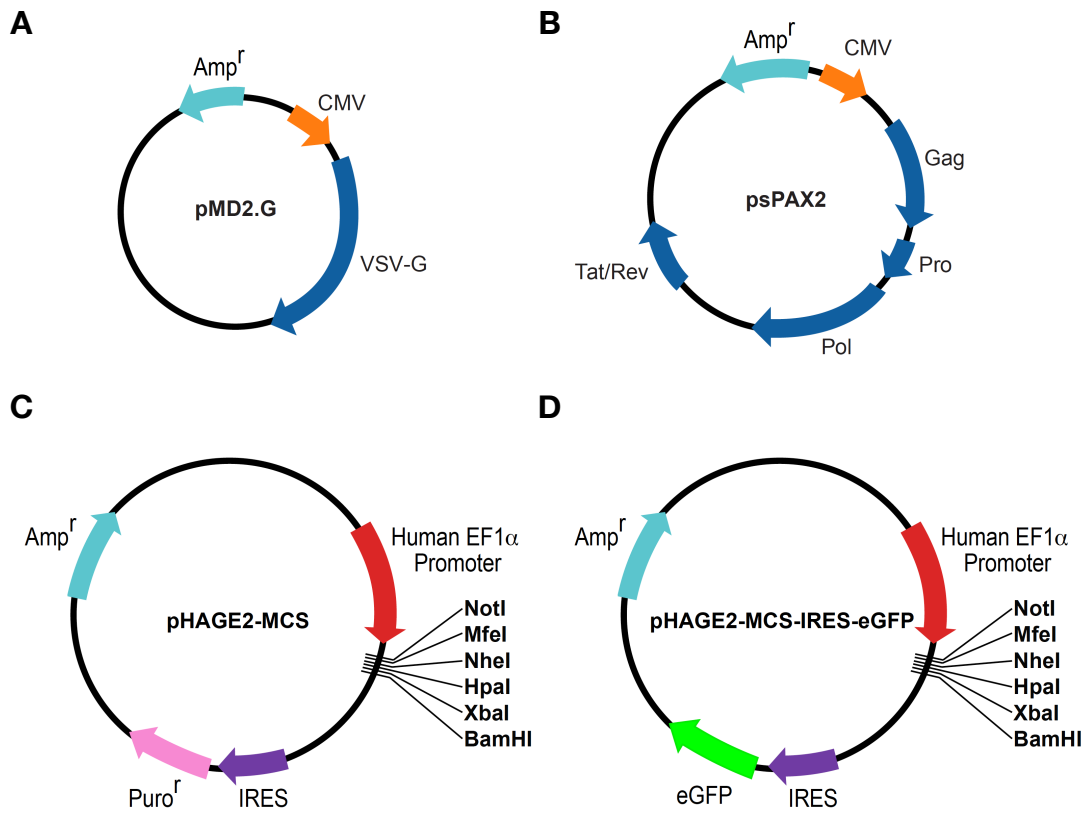

Figure 9. Lentiviral expression system of the study. (**A**) pMD2.G encodes the virus envelope protein VSV-G under the control of the cytomegalovirus (CMV) promoter. (**B**) psPAX2 encodes the virus packaging proteins Gag, Pol, Rev and Tat, under the control of the CMV promoter. (**C**) pHAGE2-MCS plasmid. This plasmid includes the human EF1α promoter followed a multiple cloning site (MCS) with restriction sites for NotI, MfeI, NheI, HpaI, XbaI, BamHI restriction enzymes. These sites were used for gene cloning. In addition, it has a puromycin resistance gene (Puro^r) after an internal ribosome entry site (IRES). (**D**) pHAGE-MCS-IRES-eGFP plasmid used in control condition. Similar to pHAGE2-MCS but with an eGFP gene after the IRES. All plasmids comprise an ampicillin resistance gene (Amp^r) for bacterial selection.

2.2.2 TEMPLATE VECTORS

The original coding sequences (CDSs) of the genes/TFs used in this study were purchased from Addgene and Open Biosystems plasmid collections, or already existed in our lab's plasmid library, with the exception of Gata2 plasmids, which were kindly provided by Dr. Sjaak Philipsen, Erasmus MC, Rotterdam. Protein CDSs were aligned with their consensus CDS (NCBI) with Vector NTI® software (Version 6.0.0.0), after being sequenced (Sanger Sequencing Service, GATC Biotech) with the sequencing primers shown in **Table 2**.

Table 2. Template plasmid vectors and sequencing primers designed for Sanger sequencing. The symbol Δ stands for "deletion" and the numbers between parentheses represent the aminoacids that were deleted from Gata2 wild type sequence.

Plasmid	Primer	Sequence
pFUW-Gfi1 pFUW-Gata1 pFUW-FosB	pFUW-F	5'-TCCACGCTGTTTTGA-3'
pcDNA-Gata3 pYX-Asc-Fra2	T7	5'-TAATACGACTCACTATAGGG-3'
pcDNA3-Flag-Gata2 pcDNA3-Flag-Gata2Δ(69-153) pcDNA3-Flag-Gata2Δ(154-256) pcDNA3-Flag-Gata2Δ(257-287) pcDNA3-Flag-Gata2Δ(287-342) pcDNA3-Flag-Gata2Δ(343-379) pcDNA3-Flag-Gata2Δ(380-440) pcDNA3-Flag-Gata2Δ(1-74) pcDNA3-Flag-Gata2Δ(1-235) pcDNA3-Flag-Gata2Δ(440-480)	CMV-F	5'-CGCAAATGGGCGGTAGGCGTG-3'
pHAEGE2-MCS-IRES-mCherry	IRES-F	5'-TGGCTCTCCTCAAGCGTATT-3'

2.2.3 POLYMERASE CHAIN REACTION (PCR) AMPLIFICATION

2.2.3.1 PRIMER DESIGN

Each gene was analysed with Vector NTI® Software for the presence of the restriction sites recognized by NotI (5'-GCGGCCGC-3'), NheI (5'-GCTAGC-3'), XbaI (5'-TCTAGA-3') and BamHI (5'-GGATCC-3') restriction enzymes. In addition, to create Gata2 fusion constructs, pHAGE2-MCS vector, Gata2 and mCherry were also analysed for the presence of BstBI restriction site (5'-TTCGAA-3'). BstBI restriction site was chosen to link both Gata2 and mCherry sequences because it neither belonged to the pHAGE2 plasmid nor existed in any of the genes. MfeI was not taken into consideration since pHAGE2-MCS vector is cut by this enzyme outside the MCS. HpaI was not desired because it produces blunt ends when cutting DNA. Restriction sites that did not exist within gene sequences were used for primer design. Cloning strategies are summarized in **Figure 10**.

Forward (5') and reverse (3') primers included the enzymes restriction site plus the first 20 nucleotides of the protein coding sequence (CDS) or its reverse complementary sequence, respectively. Six random nucleotides were also inserted upstream of restriction sites for efficient endonuclease cutting in forward (5'-GGTATC-3') and reverse (5'-CCTTAC-3') primers. Forward primers for Gata2-mCherry and mCherry-Gata2 fusion constructs were design without stop codons. Primer information is shown in **Table 3**.

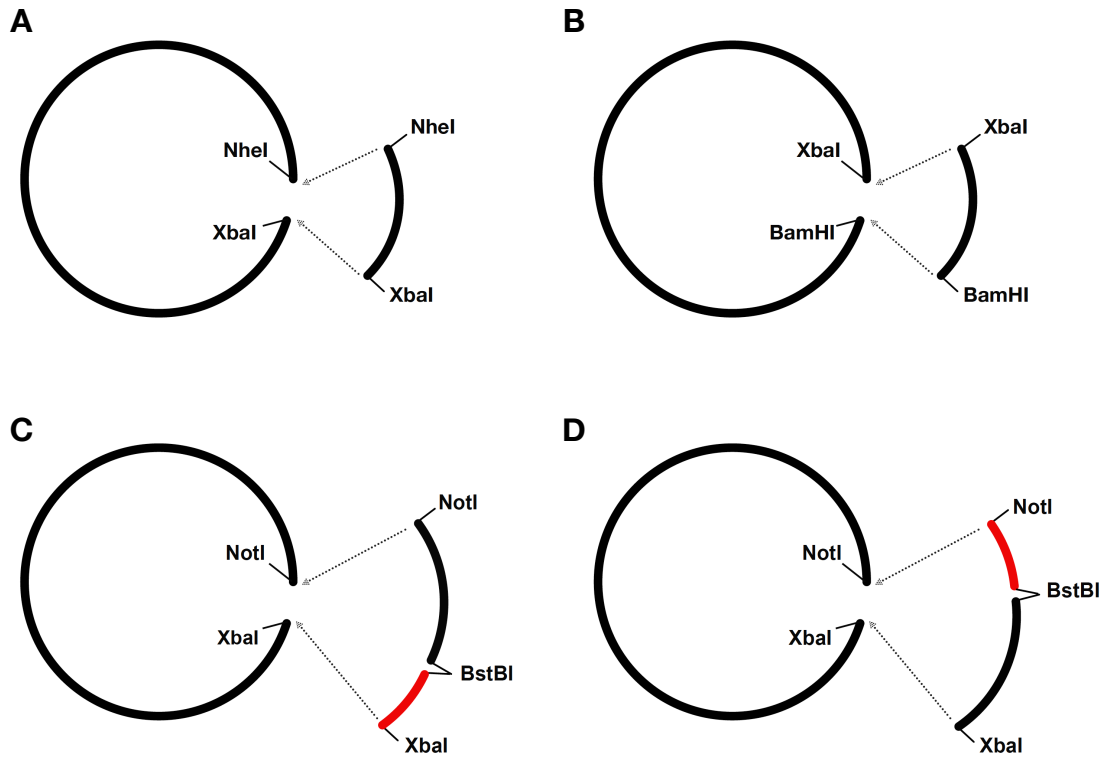

Figure 10. Cloning strategies used in the study. (**A**) Restriction sites considered for Gata3, Gfi1, Gata2 and Gata2 deletion constructs cloning. (**B**) Restriction sites used for Gata1, FosB and Fra2 cloning. (**C**) mCherry (red trace) cloning strategy at the 3' end of Gata2 (black trace). (**D**) mCherry (red trace) cloning strategy at the 5' end of Gata2 (black trace).

Table 3. Primers designed for gene cloning into the pHAGE2-MCS vector according to each experimental approach.

Insert	5'Primer	5'Enzyme	3'Primer	3'Enzyme	
Gata1	GGTATCTCTAGAatggattttcctggtctagg	XbaI	CCTTACGGATCCtcaagaactgagtggggcga	BamHI	Paralogs
Gata3	GGTATCGCTAGCatggaggtgactgcggacca	NheI	CCTTACTCTAGActaacccatggcggtgacca	XbaI	
Gfi1	GGTATCGCTAGCatgccgcgctcatttctcgt	NheI	CCTTACTCTAGAtcatttgagcccatgctgcg	XbaI	
FosB	GGTATCTCTAGAatgtttcaggctttccccgg	XbaI	CCTTACGGATCCtcacagagcgaggagggagg	BamHI	
Fra2	GGTATCTCTAGAatgtaccaggattatcccgg	XbaI	CCTTACGGATCCttacagggctagaagtgtgg	BamHI	
Gata2	GGTATCGCTAGCatggaggtggcgcctgagca	NheI	CCTTACTCTAGActagcccatggcagtcacca	XbaI	Gata2 deletion constructs
Gata2Δ(69-153)	GGTATCGCTAGCatggaggtggcgcctgagca	NheI	CCTTACTCTAGActagcccatggcagtcacca	XbaI	
Gata2Δ(154-256)	GGTATCGCTAGCatggaggtggcgcctgagca	NheI	CCTTACTCTAGActagcccatggcagtcacca	XbaI	
Gata2Δ(257-287)	GGTATCGCTAGCatggaggtggcgcctgagca	NheI	CCTTACTCTAGActagcccatggcagtcacca	XbaI	
Gata2Δ(287-342)	GGTATCGCTAGCatggaggtggcgcctgagca	NheI	CCTTACTCTAGActagcccatggcagtcacca	XbaI	
Gata2Δ(343-379)	GGTATCGCTAGCatggaggtggcgcctgagca	NheI	CCTTACTCTAGActagcccatggcagtcacca	XbaI	
Gata2Δ(380-440)	GGTATCGCTAGCatggaggtggcgcctgagca	NheI	CCTTACTCTAGActagcccatggcagtcacca	XbaI	
Gata2Δ(1-74)	GGTATCGCTAGCatgccggcgcatgcccgtct	NheI	CCTTACTCTAGActagcccatggcagtcacca	XbaI	
Gata2Δ(1-235)	GGTATCGCTAGCatgaccatgggcacccagcc	NheI	CCTTACTCTAGActagcccatggcagtcacca	XbaI	
Gata2Δ(440-480)	GGTATCGCTAGCatggaggtggcgcctgagca	NheI	CCTTACTCTAGActacacaggtgccatgtgtc	XbaI	
Gata2	GGTATCGCGGCCGCatggaggtggcgcctgagca	NotI	CCTTACTTCGAAgcccatggcagtcaccatgc	BstBI	Gata2 fusion constructs
mCherry	GGTATCTTCGAAgtgagcaagggcgaggaggt	BstBI	CCTTACTCTAGActacttgtacagctcgtcca	XbaI	
mCherry	GGTATCGCGGCCGCatggtgagcaagggcgagga	NotI	CCTTACTTCGAActtgtacagctcgtccatgc	BstBI	
Gata2	GGTATCTTCGAAgaggtggcgcctgagcagcc	BstBI	CCTTACTCTAGActagcccatggcagtcacca	XbaI	

2.2.3.2 PCR REACTION AND DNA PURIFICATION

All inserts were amplified by PCR method using a Phusion High-Fidelity PCR Master Mix (TermoFisher, F-548) in 20μL reactions, inside the GeneAMP PCR System 9700 thermal cycler (Life Technologies). PCR mix components and PCR cycle program are described in **Table 4** and **Table 5**, respectively.

Table 4. Volume and concentrations of PCR mix components in a 20μL reaction.

PCR	Volume	Final concentration
Water	8μL	
Phusion High-Fidelity PCR Master Mix	10μL	1X
Primers (forward+reverse)	1μL	0,5μM
DNA template	1μL	5 – 25ng/ μL

Table 5. PCR cycle program for Phusion High-Fidelity PCR Master Mix.

Step	Temperature	Time	Cycles
Initial denaturation	98°C	10s	1
Denaturation	98°C	1s	
Annealing	58°C	5s	30
Extension	72°C	30s	
Final extension	72°C	1min	1
	4°C	Hold	

After PCR, amplification products were mixed with Orange-G (VWR, E783) loading buffer 1x concentrated and loaded in a 1% ultrapure grade agarose (NZYTech, MB05201) gel electrophoresis with 0.5μg/mL of ethidium bromide. Samples were run alongside with 5μL of a NZYDNA Ladder III (#MB04402) for 40 minutes under a current of 180mA. Afterwards, visible bands corresponding to amplified inserts were excised from the gel and purified according to NZYGelpure kit (NZYTech, MB01101) protocol. Purified DNA was quantified using NanoDrop (ND 1000 Spectrophotometer, Alfagene).

2.2.4 Restriction and ligation of vector and inserts

Double restrictions of 2,5μg of pHAGE2-MCS plasmid were performed with three combinations of restriction enzymes from Thermo Fisher Scientific: NheI (ER0971) plus XbaI (ER0681), XbaI plus BamHI (ER0051) and NotI (ER0591) plus XbaI, for 3h at 37ºC. Total volume (30μL) of purified amplified inserts were restricted with the same combinations of enzymes or another two different combinations, NotI plus BstBI (ER0121) or BstBI plus XbaI, in agreement with the information available on **Table 3**. Double restriction of inserts was performed for 2h30min at 37ºC. Restricted vector and inserts were once again purified using an adapted protocol of NZYGelpure kit. DNA purity and concentration were measured by spectrophotometry (NanoDrop). Details of the restriction process are present in **Table 6**. Optimal enzymatic reaction buffer was determined by Thermo Fisher Scientific's Double Digest Calculator available online (Anon n.d.).

Table 6. Enzymatic restriction reaction mix (50μL) for vector and inserts.

	Vector	Inserts
Water	Add to 50μL	13μL
Buffer	5μL	5μL
5'Enzyme	1μL	1μL
3'Enzyme	1μL	1μL
DNA	2,5μg	30μL

Ligation of inserts and plasmids was done at room temperature (RT) for 1h with 1μL of T4 DNA ligase (NEB, B0202) and respective buffer, 1μL of restricted plasmid and 7μL of restricted insert. To generate Gata2 fusion constructs, 3,5μL of Gata2 and mCherry were used. Additional ligations were performed using water instead of insert DNA to serve as negative control of transformation. Recombinant DNA was immediately used to transform competent bacteria or stored at -20ºC.

2.2.5 PRODUCTION OF CHEMICALLY COMPETENT BACTERIA

Chemically competent bacteria were produced based on two protocols from Chun *et al.* (Chung et al. 1989; Chung & Miller 1993). A sample of *Escheriscia coli (E. coli)* DH5α (NEB, C2987I) was streaked on a plate with LB (Lennox) Agar (Pronadisa, #1040203900) without antibiotic and incubated at 37ºC, overnight (16h). A liquid culture was done with 5 mL of LB broth LB broth (Pronadisa, #1040203800) and an isolated colony from the bacterial plate, then let grow at 37ºC, overnight, at 200 rotations per minute (rpm). The liquid culture was diluted 1:50 and 1:100 in 50mL cultures with LB broth, and incubated again at 37ºC, 200rpm, until optical density (OD) at 600 nm was between 0,5 and 0.6. When desired OD was reached, cultures were placed on ice, for 10 minutes to stop bacterial growth. Cell pallets were obtain after performing a 10-minute centrifugation at 4ºC and 3000 rpm (Eppendorf Centrifuge 5810R, IL055). Supernatant was discard and pellets were resuspended in sterilized transformation and storage buffer (TSB) which composition is detailed in **Table 7**. The volume of TSB added was 10% of the volume of the cultures palleted before. Cell suspensions were immediately stored at -80ºC, until used.

Competent bacteria were transformed with 80ng of the ampicillin resistant psPAX2 plasmid, in a plate with LB agar with ampicillin, and grown

overnight at 37ºC. The number of colony forming units per µg of plasmid DNA was calculated to measure transformation efficiency.

Table 7. Enzymatic restriction reaction mix (50mL) for vector and inserts.

TSB	Initial concentration	Final concentration
Polyethylene glycol 8000 (PEG) (VWR, 0159)		10% (m/v)
$MgSO_4$ (VWR, E541)	1M	10mM
$MgCl_2$ (Ambion, AM9530G)	1M	10mM
DMSO (Fisher Scientific, BP231)	99%	5% (v/v)
LB Broth	1x	1x

2.2.6 TRANSFORMATION OF COMPETENT BACTERIA WITH RECOMBINANT PLASMIDS

After ligation, competent bacteria were defrosted and transformed with recombinant plasmids. 50mL of bacteria were incubated on ice for 30 minutes with the total volume of ligated DNA (10µL). Then, they were submitted to a thermal shock for 1 minute at 42°C, followed by a cool down on an ice bath for 5 minutes. LB broth without ampicillin was added to bacteria and cells were allowed to grow for 30 minutes at 37ºC. Finally, bacterial solutions were plated with beads in LB Agar plates supplemented with ampicillin and grew overnight (16h) at 37ºC, for colony formation. Grown plates were compared with two controls done at the time of transformation: a positive control with the original non-restricted pHAGE2-MCS plasmid and a negative control with the same restricted vector but with no insert.

2.2.7 SCREENING OF POSITIVE COLONIES BY COLONY PCR

After incubation, up to ten isolated colonies at a time were picked from the plates and amplified by PCR with the same set of primers used to amplify the respective DNA templates and the NZYTaq 2x Colourless Master Mix (NZYtech, MB04003), in 10µL reactions. Amplification of pHAGE2-MCS original

vector and template vector of the inserts were used as negative control and positive controls, respectively. For Gata2 fusion constructs, two positive controls were done per transformed bacterial plate: one with the template vector of mCherry and another with the template vector of Gata2. Colony PCR mix components are found in **Table 8** and PCR cycle program in **Table 9**.

Table 8. Volume and concentrations of colony PCR mix components (except DNA) in a 10μL reaction.

Colony PCR	Volume	Final concentration
Water	4,5μL	
NZYTaq 2x Colourless Master Mix	5μL	1X
Primers (forward+reverse)	0,5μL	0,5μM

Table 9. PCR cycle program for NZYTaq 2x Colourless Master Mix.

Step	Temperature	Time	Cycles
Initial denaturation	95°C	1min	1
Denaturation	95°C	40s	
Annealing	58°C	40s	25
Extension	72°C	1min	
Final extension	72°C	7min	1
	4°C	Hold	

PCR samples were mixed with Orange-G loading buffer 1x concentrated and run in an electrophoresis 1% (m/v) standard agarose (NZYTech, MB14401) gel with 0.5mg/mL ethidium bromide, together with a DNA Ladder for 40 minutes at 180 mA. DNA bands were then visualized under ultraviolet (UV) light to screen for bands matching the size of the positive control (positive colonies). Colonies that were positive for the insert were grown in 3mL of LB broth with ampicillin overnight at 37°C and 200rpm. Recombinant plasmids were extracted and purified from cell pellets using an NZYMiniprep kit (NZYtech, MB01001) according to manufacturer's protocol, and sent for sequencing (GATC Biotech Sanger Sequencing Service) to confirm colony PCR results. The sequencing primer used to confirm the success of the cloning process into

pHAGE2-MCS vector was the EF1α-F primer, whose sequence aligns with the pHAGE2-MCS promoter (human EF1α) allowing the amplification of part of the plasmid and the adjacent cloned insert.

2.2.8 Isolation of recombinant plasmids by midiprep

Liquid cultures of positive colonies verified for the presence of recombinant plasmids (pHAGE2-insert) by colony PCR and gene sequencing were diluted 1:150 and grown on 40 mL of LB broth with ampicillin overnight at 37ºC and 200rpm. NZYMidipreps Endotoxin Free kits, (NZYtech, MB27901) were used in agreement with manufacturer's protocol to extract and purify recombinant DNA. DNA concentration and purity was measured using NanoDrop and sequences were again confirmed by Sanger sequencing.

2.3 Cells and culture conditions

Two major cellular systems were adopted in the scope of this study. 293T highly transfectable cell line, derived from human embryonic kidney (HEK)293T cells (ATCC) were used for the production of lentiviruses. Primary cultures of MEFs from a double transgenic (DT, hCD34/H2BGFP) mouse model (Pereira et al. 2013) were isolated, sorted by flow cytometry and used in all direct reprogramming experiments. C57BL/6 MEFs (ATCC) were used once as a negative control for the presence of CD45 marker and GFP.

HEK293T cells, DT MEFs and C57BL/6 MEFs were defrosted, centrifuged at 1200rpm (Eppendorf 5810R, A-4-62) and cultured in Dulbecco's modified eagle medium (DMEM) (Life Technologies, #21969-035) supplemented with 10% (v/v) fetal bovine serum (FBS) (Gibco, 11573397), 2mM L-Glutamine (Gibso, 25030081) and 10μg/mL of Penicillin-Streptomycin (Gibco, 15070063) antibiotics (hereafter complete DMEM) until desired confluence was reached. Cells were maintained at 37ºC in 5% CO_2. Cells were frozen after dissociation with TrypLE Express Enzyme (1X) (Thermo Fisher Scientific, 12604021) in a 10% dimethyl sulfoxide (DMSO) and 90% FBS solution and kept at -80ºC or liquid nitrogen until further use.

HEK293T cells were also submitted to a Nocodazole (Sigma-Aldrich, M1404) treatment (100ng/mL Nocodazole in complete DMEM) for mitotic arrest in G2/M (G2 to mitosis) transition. Nocodazole stock solutions of 5mg/mL were made with DMSO.

After transduction, DT MEFs were cultivated in filtered MyeloCult M5300 (StemCell Technologies, #05350) medium supplemented with 1mM of hydrocortisone (StemCell Technologies, #07904) to promote cell grow.

2.4 TRANSFECTION OF HEK293T CELLS

Lentiviral particles were produced through a 2nd-generation lentiviral system organized in three plasmids: a lentiviral transfer plasmid, a packaging plasmid (psPAX2.G) and an envelope plasmid (pMD2). As the lentiviral transfer plasmid, I have used the pHAGE2-MCS adaptation of the original STEMCAA vector cloned with the inserts (recombinant plasmid). Additionally, the pHAGE2-MCS-IRES-eGFP, also derived from the aforementioned original vector, was used for positive controls and pHAGE2-MCS empty backbone for negative controls.

To allow viral production, HEK293T cells were transfected on the day after a 1:6 dilution (40-50% confluence) in 10cm plates with 10μg of transfer plasmid(s), 10μg of psPAX2.G and 5μg of pMD2 plasmids per HEK293T cell plate. First, DNAs were mixed in and adjusted to a final volume of 500μL, with sterile ultrapure water. Next, 62,5μL of a 2M $CaCl_2$ solution (Merck Millipore, 102383) were added to the previous solution and 500μL of BES-buffered saline solution (Sigma-Aldrich, 14280) was added drop-wise while a pipette controller released air bubbles with the aid of a glass Pasteur pipette. This step allowed the formation of DNA complexes required for efficient cell transfection. Plasmid DNA solutions were incubated at RT for 30 minutes. In the meantime, HEK293T cell plates medium was aspirated and replaced with 10mL of fresh complete DMEM without antibiotics. Finally, DNA mixtures were evenly distributed throughout each plate and cells were incubated overnight at 37ºC, 5% of CO_2. 24 hours after transfection, the presence of GPF in HEK293T cells was confirmed in positive controls by fluorescence microscopy. Culture medium was replaced by 4mL of complete DMEM and HEK293T cells were incubated at 32ºC, 5% CO_2. Supernatants containing lentiviral particles were collected 36, 48 and 60 hours after transfection and stored at 4ºC between and after collections. Collected supernatants were filtered with a 0,45μm low-protein binding filter (Corning), and kept at 4ºC. Viruses were preferred fresh for direct reprogramming experiments.

No viral production was performed for recombinant plasmids with Gata2 fusion constructs for mitotic bookmarking experiments. DNA complexes were prepared with 10μg of each individual recombinant plasmid (pHAGE2-Gata2-mCherry or pHAGE2-mCherry-Gata2) and water, plus $CaCl_2$ solution and BES.

Each HEK293T cell plate was transfected with DNA solutions containing only one type of recombinant plasmid, and incubated in identical conditions.

2.5 LENTIVIRAL TRANSDUCTION OF DT MEFS

DT MEFs were plated in 0,1% (m/v) gelatine-coated 6-well plates at a density of $0{,}5\times10^6$ cells per plate. Cells in each well were then transduced twice, in a 24h interval. To do so, culture medium was aspirated and 2mL of previously produced lentiviruses in the presence of 8μg/mL of polybrene were added. 16 hours after the first transduction (day 0), viruses were replaced by 2mL of complete DMEM to help cells recover. For the second transduction (day1), medium was discard and the same volume of viruses and polybrene was added. Polybrene helps the integration of viruses in cells by decreasing electrostatic repulsions between viral particles and cell membranes (Davis et al. 2002). During intervals between transductions, cells were kept at 37ºC in 5% CO_2 conditions. 16 hours after the second transduction, viruses were removed and cells were incubated with MyeloCult medium, suitable for the maintenance of hematopoietic cells, supplemented with hydrocortisone (day 2) to sustain hematopoietic reprogramming. Medium was replaced every 3-4 days until day 20 (final day) of the reprogramming experiments, when cells were analysed by flow cytometry. In the course of the experiments, plates were visualised several times by bright field and fluorescence microscopy to follow experiments' evolution in terms of morphological alterations and report activation.

2.6 MITOTIC ARREST OF HEK293T CELLS

After 24h of transfection, HEK293T plates for mitotic bookmarking analysis were submitted to a 16h Nocodazole treatment (100ng/mL in complete DMEM) followed by dissociation of cells from plates and centrifugation. Nocodazole arrest efficiency was tested prior to this experiment by flow cytometry analysis of cell cycle with propidium iodide DNA staining (see section 2.7.1 of this chapter). Following centrifugation, cells were resuspended in phosphate-buffered saline (PBS) solution (Life Technologies, #10010-056) and plated in poly-L-lysine (Sigma-Aldrich, P4707) coated μ-Slide 8 Well (Ibidi, 80826) plates with appropriate cell density for imaging. Then, cells were incubated with Hoechst 33342 (Life Technologies, #H3570) solution, protected from light, for 10 minutes and washed. Finally, mitotically arrested cells were assessed for the presence of red signal from fusion proteins, which

contrasts with the blue signal from Hoechst DNA staining, via fluorescence confocal microscopy.

2.7 Flow cytometry

2.7.1 Cell cycle arrest analysis with propidium iodide (PI) DNA staining

To test the efficiency of Nocodazole in arresting HEK293T cells in G2/M transition, HEK293T cells were first plated in 6 well plates until they reached 50-60% confluence, and then were treated with 2mL of both 50 and 100ng/mL of Nocodazole in complete DMEM during 4, 16 and 24 hours. A negative control corresponding to each time-point was also included in the assay. After the treatment (4, 16 or 24 hours), cells were dissociated from the respective well to a 15mL centrifuge tube (Corning), pelleted (5 minutes at 1200 rpm), washed in PBS and fixated with 2mL ice-cold 70% ethanol solution. Fixated cells were stored at 4ºC post fixation and stained afterwards. Prior to staining, fixated cells were washed in PBS and centrifuged at 1700 rpm for 7 minutes. Next, 1mL of staining buffer containing PI (Thermo Fisher Scientific, P1304MP) was added to each pellet and cells incubated first for 10 minutes at RT and then 20 minutes on ice, always protected from light. The Staining Buffer content is presented in **Table 10**. Lastly, cells suspensions were filtered into a cell strainer tube (BD Biosciences) and analyzed in a BD Accuri C6 Flow Cytometer (BD Biosciences) using the standard laser configuration of 3-blue-1-red with the 533/30 filter in FL1, 585/40 in FL2, 670 LP in FL3, and 675/25 in FL4. PI was detected in FL2 channel. Flow cytometry result analysis was performed via FlowJo Software (version 7.8, FlowJo, LLC).

Table 10. Staining Buffer composition and concentrations of the components

Staining buffer	Initial concentration	Final concentration
PBS	70%	1X
PI		1mg/mL
TERGITOL solution (Sigma-Aldrich, NP40S)		5%
NZY RNase A (NZYTech, MB18701)		10mg/mL

2.7.2 FLUORESCENCE-ACTIVATED CELL SORTING (FACS)

MEFs used in reprogramming experiments were isolated from E13.5 hCD34/H2BGFP mouse embryos following removal of hematogenic regions, expanded and sorted. Before sorting, plated MEFs were dissociated from plates as previously described and pelleted in FACS tubes. 1mL of cell suspension was saved before centrifugation to serve as the unstained control. Pellets were resuspended in 200μL of PBS 2% (v/v) FBS and incubated with 2μL of PerCP-Cy5.5 rat anti-mouse CD45 (BD Biosciences, Clone 30 F-11) for 15 minutes on ice, protected from light. MEFs were washed with the same solution and resuspended again for acquisition. Double negative MEFs for GFP and CD45 were sorted in BD FACS Aria III (BD Biosciences). 100 μL of sorted population were submitted to a purity check. Sorted cells were plated until confluent, and stored in liquid nitrogen until used.

2.7.3 HCD34/H2BGFP REPORTER ACTIVATION ANALYSIS

To analyse hCD34-driven GFP expression, transduced MEFs were washed with PBS, dissociated with TrypLE Express for 10 minutes at 37ºC, resuspended in PBS 2% (v/v) FBS for trypsin inactivation and pelleted (centrifugation at 1200rpm, 5min). Cell pellets were resuspended in 200μL of PBS 2% (v/v) FBS and kept at 4ºC prior analysis. GFP was measured in FL1 channel. Flow cytometry result analysis was performed via FlowJo Software.

2.8 DT MEF GENOTYPING BY PCR

To further confirm the presence of hCD34 and GFP genes in DT MEFs genome, genomic DNA (gDNA) was extracted from small pellets of sorted DT MEFs and C57BL/6 MEFs (for negative control) via an NZY Tissue gDNA Isolation kit (NZYTech, MB13502) according to manufacture's protocol. Purified gDNA was amplified by PCR (formerly indicated settings) with GFP and hCD34 forward (5'-AGCTGACCCTGAAGTTCATCTG-3'; 5'-AGAAGAGATGAGGTGTGAGGAT-3') and reverse (5'-GTCGGCCATGATATAGACGTTG-3'; 5'-GGATCCACAAGAATGAGCATGTA-3') primers, respectively. PCR samples were loaded and run in a 1% agarose gel as above-mentioned. DNA bands were visualized under UV light.

2.9 IMAGE ACQUISITION AND ANALYSIS

Bright field and fluorescent images of cells from reprogramming experiments were obtained with an inverted microscope (Zeiss AxioVert 200M). Fluorescent cell imaging of mitotic bookmarking experiments was performed using a confocal microscope (Zeiss Lsm710). Images were processed with Adobe Photoshop CS5 and Adobe Illustrator CS5.

CHAPTER 3

3. RESULTS

3.1 DOUBLE TRANSGENIC (hCD34/H2GFP) MEF ISOLATION AND PURIFICATION

The reporter system used in this study was based on the activation of the hCD34 promoter and consequent expression of nuclear GFP (H2BGFP) upon hematopoietic cell fate acquisition (Schaniel & Moore 2009). This reporter system was already been proven efficient for the identification of reprogrammed cells resulting from the induction of a hemogenic/hematopoietic program in MEFs (Pereira et al. 2013). Thus, before starting any reprogramming experiment, fibroblasts from double transgenic (hCD4/H2BGFP) mice embryos were separated from other tissues and purified. For that reason, the head and hematogenic regions of E13.5 embryos were removed and MEFs were expanded (**Figure 11, A**). To ensure that reporter activation was due to reprogramming experiments, contaminant cells expressing the hematopoietic marker CD45 (Hermiston et al. 2003) and GFP were excluded by cell sorting (**Figure 11, B**). The sorted cell population was then submitted to a purity check showing 99,2% of double negative (CD45- and GFP-) cells (**Figure 11, C**). Furthermore, the presence of the two transgenes (hCD34 and H2BGFP) in the sorted cell population was confirmed by genotyping (**Figure 11, D**). The resulting purified DT MEFs were used for the validation of the study's gene delivery system and in direct reprogramming experiments.

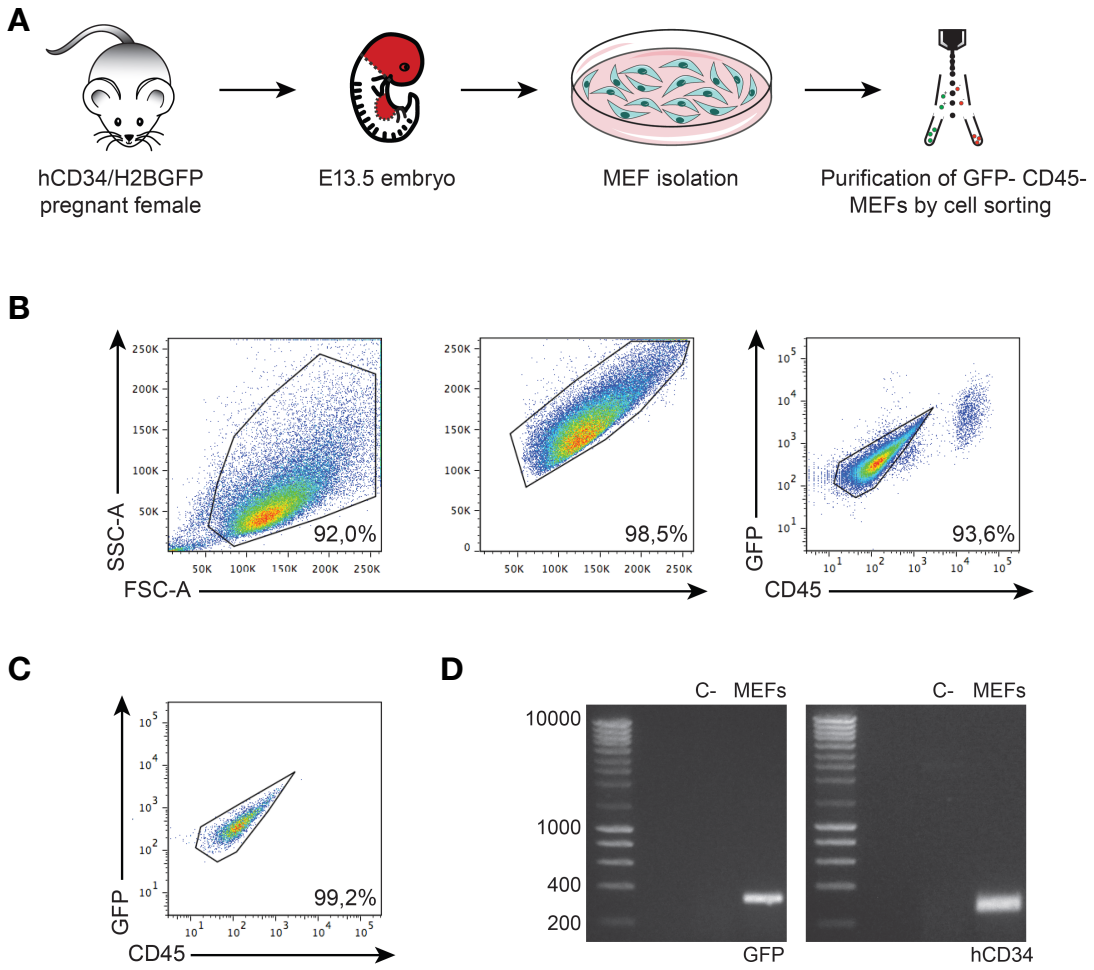

Figure 11. MEF isolation and purification from double transgenic (DT, hCD34/H2BGFP) mice. (**A**) Scheme of the isolation and purification of mouse embryonic fibroblasts (MEF) from hCD34/H2BGFP double transgenic mice embryos at E (embryonic day) 13.5. Hematogenic regions and head (in red) were removed prior isolation and MEFs were grown until confluent. MEFs were sorted to exclude GFP+ and CD45+ (hematopoietic marker) contaminant cells. (**B**) Gating strategy to limit double negative cell for sorting. (**C**) Sorted cell population purity check. (**D**) Genotyping of a sorted cell sample. MEFs were submitted to genomic PCR to confirm the presence of both transgenes. C57BL/6 MEFs were used as negative control (C-). Band sizes shown in base pairs (Bp).

3.2 Establishing a lentiviral system to induce hematopoietic reprogramming

In 2013, Pereira *et al* have used a retroviral system to deliver Gata2, Gfi1b and cFos into fibroblast in order to convert them into hemogenic/hematopoietic stem-like cells (Pereira et al. 2013). Retroviral delivery vectors have been used for decades to express genes of interest in

target cells due to their ability to stably integrate the host genome and promote long-term gene expression (Nayerossadat et al. 2012; Yi et al. 2011). However, retroviral vectors produce low viral titers in addition to the inability of transducing quiescent non-dividing cells. To overcome this limitations, lentiviral vectors, a subtype of retroviral vectors, were developed. Lentiviruses, like their retroviruses counterpart, are able to induce permanent transgene expression by integrating the cells' genome, with the advantage of also transducing non-dividing cells (Escors & Breckpot 2010). Moreover, lentiviral gene delivery systems have suffered several improvements to increase viral titers. Therefore, the lentiviral vector pHAGE2-MCS, a modified version of the STEMCCA (pHAGE2-EF1αFull-hOct4-F2A-hKlf4-IRES-hSox2-P2A-hcMyc-W-loxP), was chosen as the gene delivery system of this study. Therefore, pHAGE2-MCS must be proven as efficient as the previously used retroviral system in hematopoietic reprogramming experiments.

3.2.1 pHAGE2 IS EFFICIENT FOR GENE EXPRESSION IN FIBROBLATS

To deliver specific TFs using a lentiviral system, viral packaging cells must be transfected with the appropriate plasmids for the ensemble of lentiviral particles. Then, resulting viruses must be capable of transducing host cells and deliver the genetic material that will allow cell fate reprogramming. For this purpose, pHAGE2-MCS-IRES-eGFP (to shorten pHAGE2-GFP), a control plasmid, was first tested for its ability to transfect HEK293T cells and, as lentiviral particles, transduce DT MEFs. The general protocol for lentiviral production (HEK293T cell transfection) and DT MEF transduction is detailed in sections 2.4 and 2.5 of chapter 2, respectively.

A 2nd-generation lentiviral system composed by the transfer plasmids (pHAGE2-MCS empty or pHAGE2-GFP), the packaging plasmid psPAX2 and the envelope plasmid pMD2.G was used to co-transfect HEK293T cells for lentiviral production. After 24 hours, expression of GFP was detected by fluorescent microscopy only in plates with pHAGE2-GFP transfer plasmid (**Figure 12, A**), confirming plasmid uptake. Supernatants containing lentiviral particles were retrieved from HEK293T cell plates, filtered and used to transduce DT MEFs twice, in a period of 24 hours. GFP expression was observed under fluorescent microscopy and quantified by flow cytometry, 4 days after first transduction (**Figure 12, B**) showing a percentage of 92,7% transduced cells. In conclusion, pHAGE2 lentiviral system is efficient for MEF transduction and protein expression.

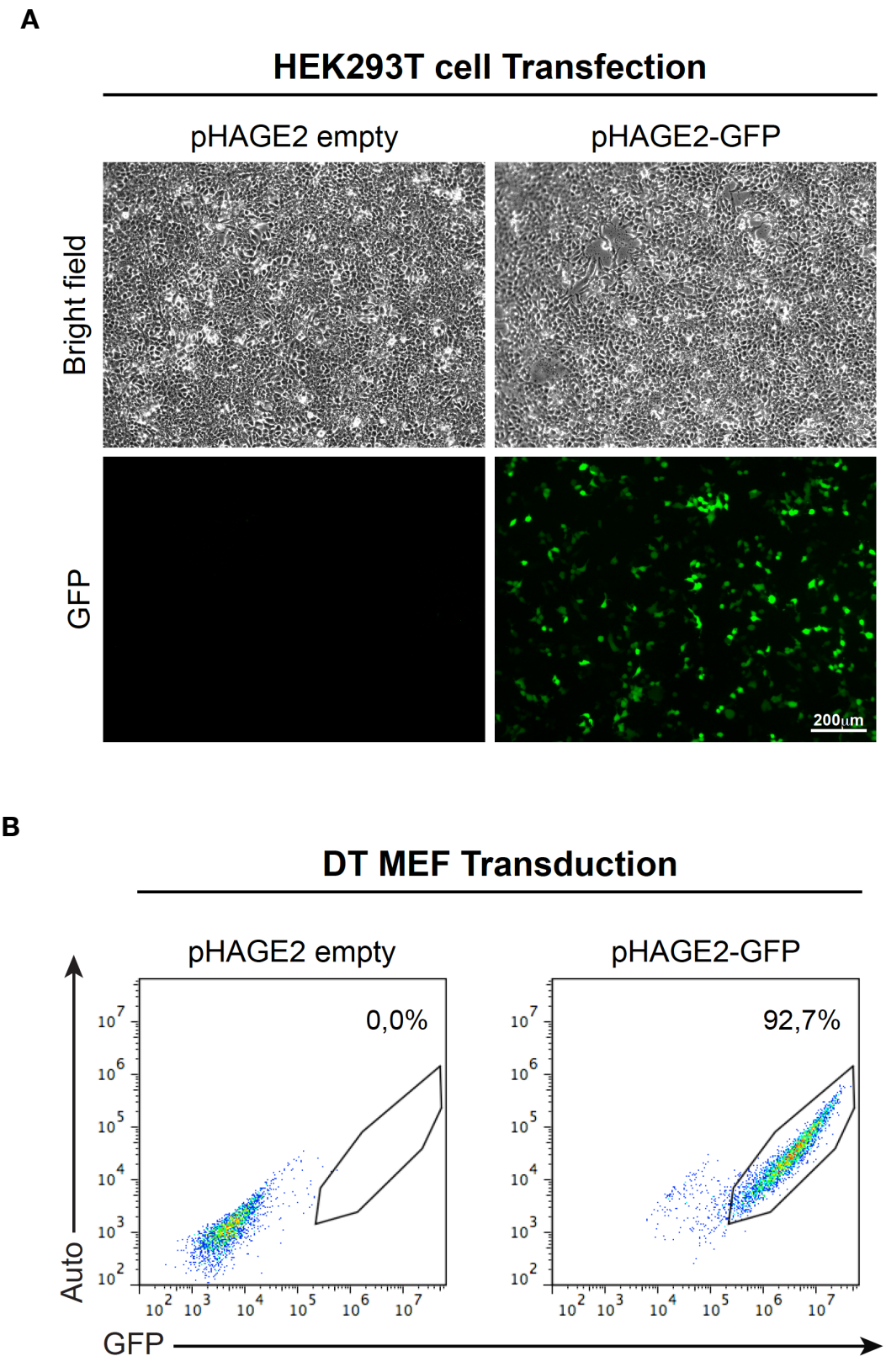

Figure 12. pHAGE2 lentiviral vectors successfully transfected the HEK293T cell line and delivered GFP into MEFs. (**A**) GFP expression 24 hours after HEK293T cell transfection with pHAGE2-GFP or pHAGE2-MCS empty vectors, plus pMD2.G and psPAX2 for lentiviral production. Fluorescent snapshots were taken under the FITC

filter. (**B**) Percentage of double transgenic MEFs expressing GFP 4 days after first transduction, measured by flow cytometry.

3.2.2 GATA2, GFI1B AND CFOS LENTIVIRAL PARTICLES INDUCE HEMATOPOIETIC REPROGRAMMING

After testing the ability of pHAGE2-MCS plasmids in transfecting HEK293T cells and transducing DT MEFs, I proceeded with direct reprogramming experiments to assess if pHAGE2 lentiviral vectors encoding Gata2, Gfi1b and cFos TFs could substitute the reprogramming proven retroviral vectors. Hence, pHAGE2 plasmids already available in the laboratory encoding Gata2 (human version), Gfi1b and cFos (mouse versions) were used to transfect HEK293T cells for lentiviral production.

Two transfection methodologies were attempted to investigate co-transduction efficiency. Factors were added individually or simultaneously alongside with the packaging (psPAX2) and envelope plasmids (pMD2.G). Supernatant of co-transfected cells were used to transduce DT MEFs (3 TFs pool) and individual supernatants were mixed in a 1:1 manner (3 TFs mix) before transduction. DT MEFs were transduced with 1 or 2mL of viral supernatant. Reporter activation was followed through fluorescent microscopy and measured by flow cytometry at the end of the experiment. The shift in cell identity towards the hematopoietic lineage activated the hCD34 promoter that prompted the transcription of a tetracycline (Tet)-transactivator (tTat), which in turn interacted with a tet-response element (TRE) that drove the transcription of H2BGFP by a minimal cytomegalovirus (mCMV) promoter (**Figure 13, A**). Flow cytometry analysis of GFP expression 20 days after first transduction showed that it was dependent on the methodology of transfection (3 TFs pool or 3TFs mix) and volume of viruses utilized for DT MEF transduction (1 or 2mL) (**Figure 13, B**). The highest percentage of reprogramed GFP+ cells (2,56%) was obtained when 2mL of lentiviral suspension were used and lentiviruses were produced simultaneously as a pool (**Figure 13, B and C**). This percentage of reprogrammed cells was comparable to the 2,82% of GFP+ cells obtained by Pereira *et al*. after hematopoietic reprogramming. Summarily, pHAGE2-TFs transfer plasmids were able to induce hCD34/H2BGFP reporter activation in fibroblasts.

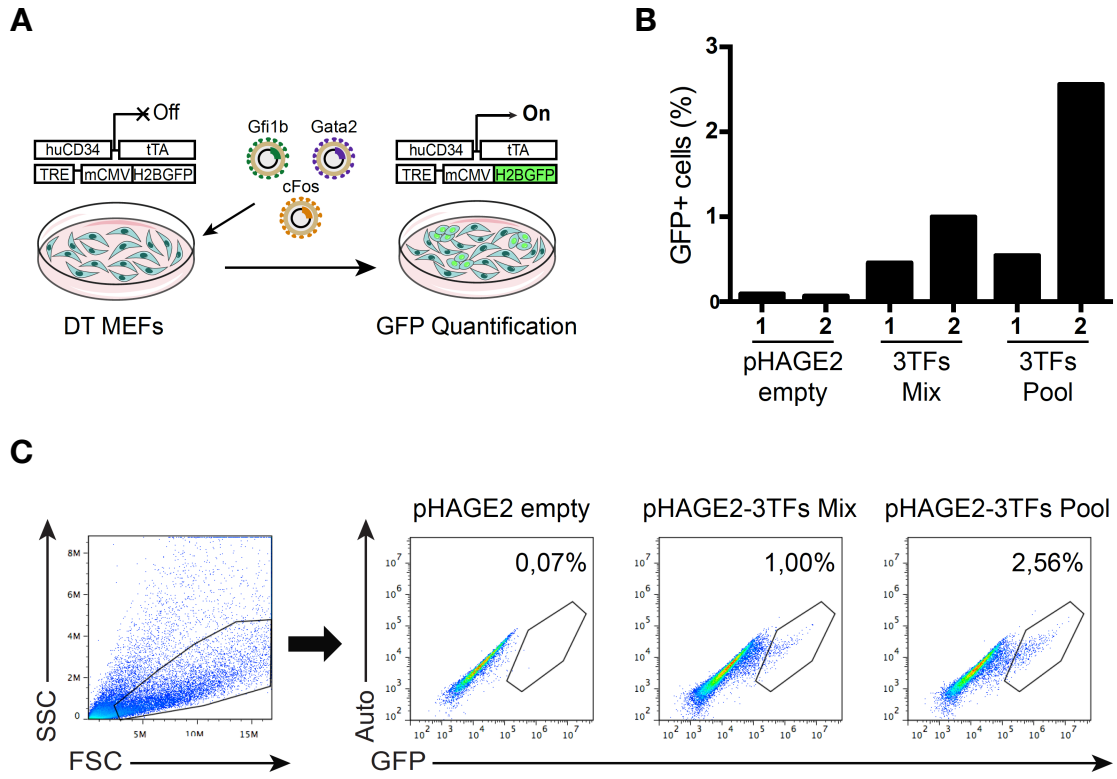

Figure 13. Hematopoietic reprogramming of fibroblasts with pHAGE2 lentiviral plasmids. (**A**) Schematic representation of hCD34/H2BGFP reporter activation after DT MEF transduction with pHAGE2 lentiviral plasmids encoding Gata2, Gfi1b and cFos. (**B**) Percentage of GFP+ cells (reprogrammed cells) depending on the volume of lentiviral supernatant (1 or 2mL per transduction) and transfection methodology (3TFs mix or pool). Flow cytometry analysis was performed 20 days after DT MEFs transduction. (**C**) Gating strategy to select GFP+ cells from the live cell population when optimal lentiviral supernatant volume (2mL) was used. SSC, side scatter. FSC, forward scatter. Auto, autofluorescence.

3.3 Impact of GATA2, GFI1B and cFOS paralogs in hematopoietic reprogramming

In order to assess what protein domains of Gata2, Gfi1b and cFos are involved in hematopoietic reprograming, a substitution approach using paralog genes of these three genes was developed. Firstly, coding sequences (CDSs) of Gata2, Gfi1b and cFos paralogs were cloned into the pHAGE2-MCS vector. Secondly, recombinant plasmids were used to transfect HEK293T cells for lentiviral production. Thirdly, DT MEFs were transduced with different combinations of lentiviral particles to investigate changes in reporter activation.

3.3.1 PCR-BASED CLONING OF GATA2, GFI1B AND cFOS PARALOG GENES INTO THE pHAGE2-MCS VECTOR

Gata2, Gfi1b and cFos paralog genes were individually sub-cloned into the pHAGE2-MCS backbone. Gata1 and Gata3 (Gata2 paralogs), Gfi1 (Gfi1b paralog) and FosB and Fra2 (cFos paralogs) were cloned. Fra1 was not available. CDSs of paralog genes came in different template vectors and were acquired from various cDNA libraries. Each CDS was verified by Sanger sequencing. **Table 11** summarizes the information for each available gene regarding the vector and species of origin, source and sequencing primer used to confirm the sequences before cloning.

Table 11. Gata2, Gfi1b and cFos paralog genes available for cloning into the pHAGE2-MCS vector.

Insert	Vector	Species	Source	Primer
Gfi1	pFUW-TetO	*Homo sapiens*	Laboratory	pFUW-F
Gata1	pFUW-TetO	*Mus musculus*	Laboratory	pFUW-F
Gata3	pcDNA3.1	*Mus musculus*	Addgene	T7
FosB	pFUW-TetO	*Mus musculus*	Laboratory	pFUW-F
Fra2	pYX-Asc	*Mus musculus*	Open Biosystems	T7

All genes were amplified and PCR products were separated by agarose gel electrophoresis (**Figure 14, A**). Subsequently, DNA bands were extracted, purified and ligated with the pHAGE2-MCS vector, after double enzymatic restriction of both inserts and vector. A ligation control was performed to confirm its success (**Figure 14, B**). A faint band corresponding to the recombinant plasmid is shown in the presence of T4 DNA Ligase, whereas in the absence of ligase, a band corresponding to the non-ligated insert is clearly seen. Following ligation, recombinant plasmids were used to transform chemically induced competent bacteria (*E. coli* DH5α) (**Figure 14, C**). Bacterial transformation efficiency, in colony forming units (CFUs)/μg of DNA, was assessed beforehand with 88ng of psPAX2 plasmid. Competent *E. coli* were successfully transformed with psPAX2 ($27x10^3$ CFUs/μg of psPAX2) and this batch was used in all transformation processes of the study. LB agar plates supplemented with ampicillin were used for bacterial selection. Bacteria which

uptook the recombinant plasmids (that comprise an ampicillin resistance gene) grew and formed colonies, however incomplete restriction of the plasmid could cause the appearance of bacterial colonies without the insert but with the selective marker, thus positive colony screening was necessary. A negative control of transformation (see section 2.2.4, chapter 2) was performed to assess the amount of colonies that derived from non-restriction of the plasmid. A positive control of transformation was also performed (see section 2.2.6). To complete the paralog genes' cloning process, colonies were screened for the presence of the corresponding genes by (colony) PCR. The original DNA template was used as positive control and a colony of bacteria transformed with the original pHAGE2-MCS backbone as negative control. Positive colonies displayed a band similar to the positive control (**Figure 14, D**). Finally, recombinant plasmids were extracted from positive colonies and sequences were confirmed via GATC Biotech Sanger sequencing service. Sequencing results, presented in **Table 12** were compared with each individual CDS.

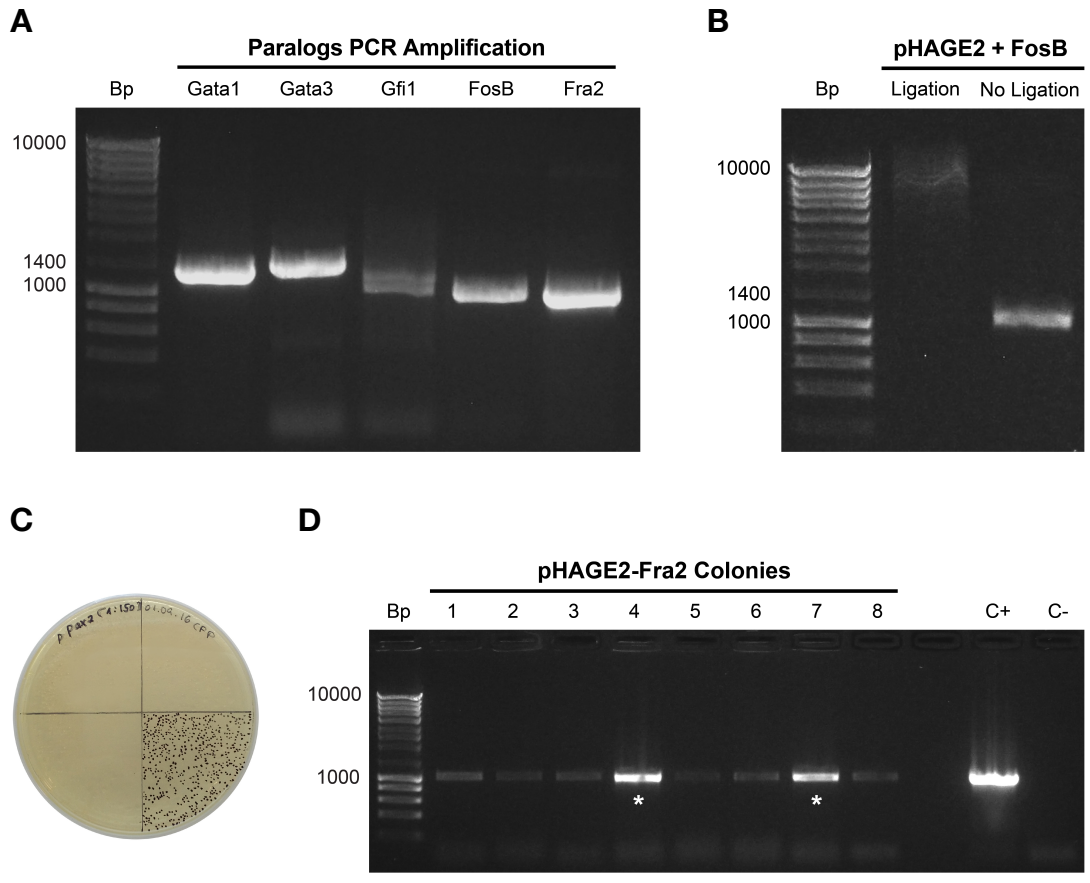

Figure 14. Gata2, Gfi1b and cFos paralog genes were successfully cloned into the pHAGE2-MCS vector. (**A**) Electrophoretic gel of paralog genes amplified by PCR. DNA bands correspond to the size of each individual coding sequence. (**B**) Ligation control example. In the presence of ligase, FosB restricted sequence binds to the

restricted pHAGE2 plasmid and no FosB band appears on the gel. In the absence of ligase the band is shown. (**C**) Chemically-induced competent bacteria (*E. coli* DH5α) yielding 27x10^3 CFUs/μg of psPAX2 DNA. These bacteria were transformed with pHAGE2 encoding each paralog individually (recombinant plasmids) for gene cloning. (**D**) Representative gel of colony PCR performed to screen for positive colonies. Eight colonies from a plate containing bacteria transformed with pHAGE2-Fra2 were picked and submitted to a PCR. The positive control (C+) corresponds to the template vector of Fra2 and the negative control (C-) to a colony picked from a plate after transformation with pHAGE2-MCS empty vector. Starts (*) indicate positive colonies. DNA band sizes are presented in base pairs (Bp).

Table 12. Sequencing data of paralog genes cloned into the pHAGE2-MCS vector. Sanger sequencing results were aligned with the corresponding cDNA sequences via Vector NTI software. Quality of the alignment is shown in purple. Irregularities were verified for sequencing errors.

Cloning product	Sequencing data
Gata1	
Gata3	
Gfi1	
FosB	
Fra2	

3.3.2 PARALOG GENES AFFECT HEMATOTPOIETIC REPROGRAMMING OF FIBROBLASTS DIFFERENTLY

Strategy of the direct reprogramming of DT MEFs with Gata2, GFi1b and cFos paralogs is outlined in **Figure 15.** Each HEK293T cell plate was transfected with one paralog gene and the remaining unrelated hematopoietic reprogramming factors, plus psPAX2 and pMD2.G viral plasmids. This way Gata2, Gfi1b and cFos factors were replaced one at the time by their respective paralog(s). DT MEFs that were sorted initially were transduced with the various pools of filtered lentiviruses, constituting different reprogramming

conditions, and GFP expressing cells were screened (**Figure 15, A**). Cells were cultivated in Myelocult and the percentage of reprogrammed cells (GFP+ cells) was quantified by flow cytometry 20 days past the first transduction (**Figure 15, B**).

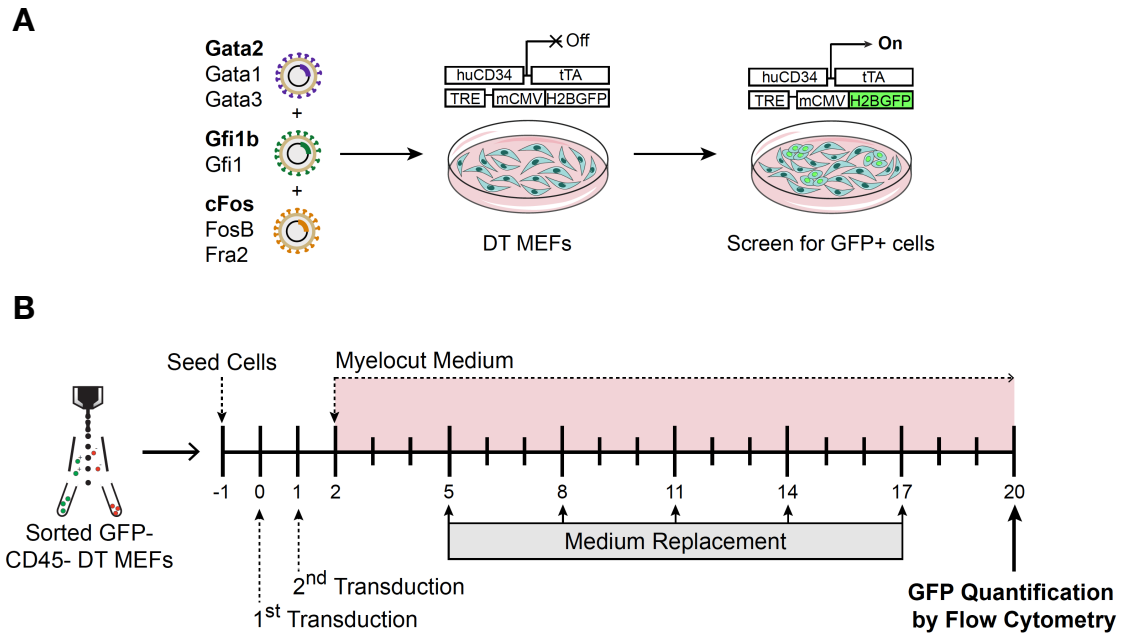

Figure 15. Paralog gene substitution strategy for direct reprogramming of DT MEF. (**A**) Scheme of the DT MEF transduction with pHAGE2 lentiviral plasmids encoding Gata2, Gfi1b, cFos and respective paralogs. Different combinations of both hematopoietic TFs and their paralogs were added to DT sorted MEFs in order to study the impact on GFP expression. (**B**) Experimental strategy outline. DT sorted MEFs were seeded the day before the first transduction and transduced twice. Following viruses' removal, cells were cultivated in myelocult medium, which was replaced every 3 days. Percentage of GFP+ cells (from reporter activation) was quantified by flow cytometry 20 days after the first transduction (day 0).

Results from GFP+ cell quantification by flow cytometry regarding paralog genes substitution for hematopoietic reprogramming are shown in **Figure 16**. DT MEF plates transduced with pHAGE2-MCS empty vector served as negative control (C-) for every condition. To investigate the extent of paralog genes substitution for hematopoietic reprogramming, the percentage of GFP+ cells obtained per experimental condition was compared to the percentage of reprogrammed cells generated by the three TFs: Gata2 (Ga), Gfi1b (G) and cFos (F), which constituted the positive control. Flow cytometry analysis has shown that Gata1 was unable to replace Gata2 for hematopoietic

reprogramming. Transduction with Gata1 resulted in a 6-fold decrease in the percentage of GFP expressing cells, when compared to Gata2, Gfi1b and cFos-induced reprogramming. On the other hand, Gata3 showed less than 1-fold decrease in GFP percentage (**Figure 16, A**), suggesting that Gata3, but not Gata1, can substitute, at least to some extent, Gata2 for hematopoietic reprogramming. Since Gata1, Gata2 and Gata3 have equivalent zinc fingers in their DNA-binding domains, non-conversed regions of Gata2 and Gata3 may be required for hematopoietic reprogramming. Interestingly, the NDR of Gata2 shares no homology with Gata1, but is 46% homologous to the correspondent region of Gata3 (Minegishi et al. 2003). Similarly to Gata3, direct reprogramming with Gfi1 decreases, by less than 1-fold, the percentage of reprogrammed cells, proposing that Gfi1 can replace Gfi1b in hematopoietic reprogramming experiments (**Figure 16, B**). In fact, Gfi1b and Gfi1 exhibit some overlapping function during hematopoiesis (Fiolka et al. 2006), besides sharing two highly conserved domains (SNAG domain and the six ZFs in the DNA-binding domain). This result suggests that the non-conserved region responsible for the differences in size of both proteins may not be required for hematopoietic reprogramming. As far as cFos paralogs are concerned, FosB successfully induced hematopoietic reprogramming with a percentage of GFP+ cells equivalent to the positive control (an average of 1,9% versus the 2,0% from the positive control), implying that FosB and cFos display similar roles during reprogramming (**Figure 16, C**). Fra2 could still replace cFos, but with less efficiency when compared to FosB, leading to a decrease by 0,7-fold, in the percentage of reprogrammed cells. While FosB shares the same conserved regions with cFos (bZIP domain and the TAD), Fra2 only exhibits the bZIP domain. These results suggest that both homologous domains are necessary for hematopoietic reprogramming. Moreover, hematopoietic reprogramming with Gfi1 and FosB is consistent with the literature, since these two factors were already implemented to convert endothelial cells into HSPCs (Sandler et al. 2014; Lis et al. 2017).

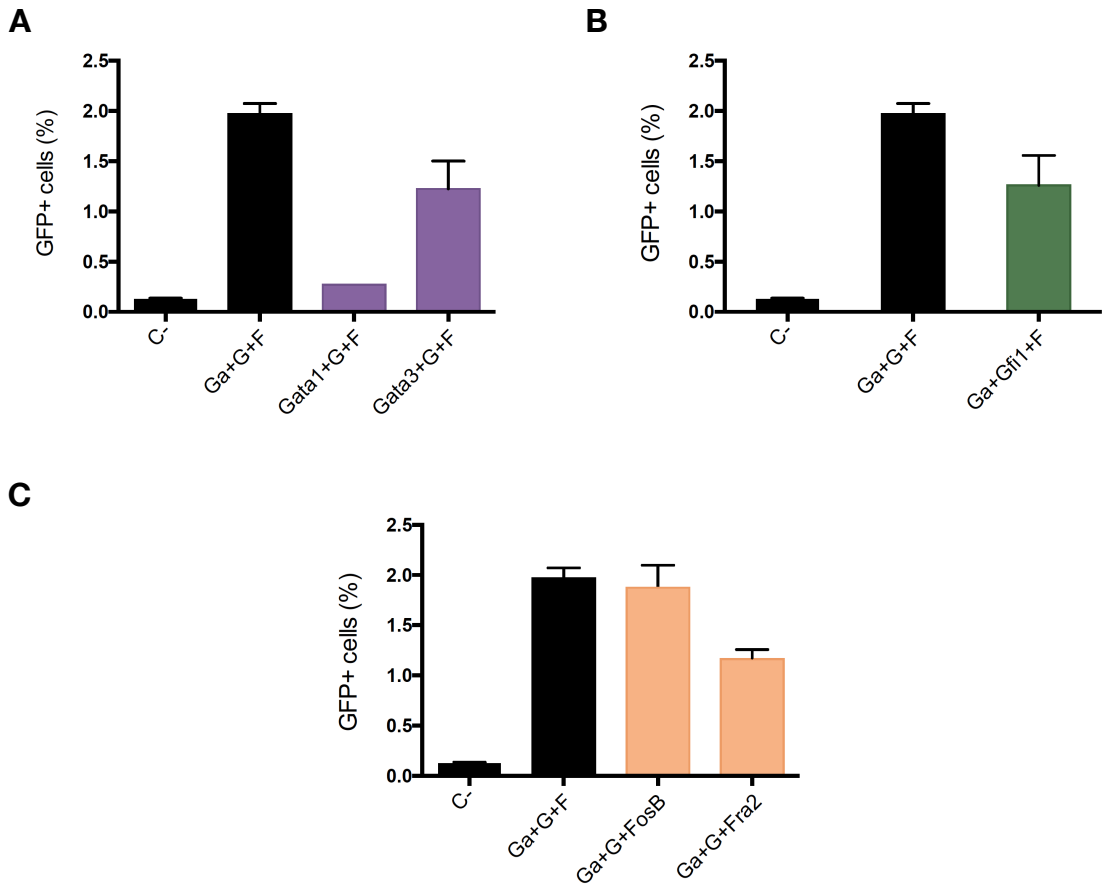

Figure 16. Gata1 does not replace Gata2 for hematopoietic reprogramming. (**A**) Percentage of GFP+ cells (resulting from hCD34/H2BGFP reporter activation) after DT MEFs transduction with Gata2 paralogs. Gata1 or Gata3 were used to substitute Gata2 for hematopoietic reprogramming. Gata1 did not replace Gata2 reprogramming function. (**B**) Percentage of GFP+ cells from Gfi1-induced hematopoieitc reprogramming. Gfi1 substituted Gfi1b reprogramming function. (**C**) Percentage of GFP+ cells when FosB or Fra2 were used to substitute cFos for hematopoietic reprogramming. FosB and Fra2 replaced cFos reprogramming function. Hematopoietic reprogramming with Gata2 (Ga), Gfi1b (G) and cFos (F) (Ga+G+F) constitutes the positive control. The negative control (C-) corresponds to the pHAGE2-MCS empty vector. Results from one experiment with 1 to 3 replicates per condition.

3.4 Functional analysis of Gata2 domains for hematopoietic reprogramming

Gata2 seems to be the dominant TF for the hematopoietic reprogramming process. Deletion constructs of this TF were used to investigate for hematopoietic reprogramming domains. CDSs of Gata2 wild

type (mouse version) and Gata2 deletion constructs were individually cloned into the pHAGE2-MCS vector. Then, recombinant plasmids were used to transfect HEK293T cells for lentiviral production and finally, DT MEFs were transduced to evaluate variations in reporter activation.

3.4.1 PCR-BASED CLONING OF GATA2 DELETION CONSTRUCTS INTO THE pHAGE2-MCS VECTOR

Gata2 wild type (WT) and Gata2 deletion constricts were kindly provided by Dr. Sjaak Philipsen at the University Medical Center Rotterdam. Each CDS was verified by Sanger sequencing. Information concerning the original vector, species and sequencing primer are presented in **Table 13**.

Table 13. Gata2 wild type and Gata2 deletion constructs template vector information. The symbol Δ stands for "deletion" and the numbers between parentheses represent the aminoacids that were deleted from Gata2 wild type sequence to create each mutant.

Insert	Vector	Species	Source	Primer
Gata2	pcDNA3-Flag	*Mus musculus*	Sjaak Philipsen	CMV-F
Gata2Δ(69-153)	pcDNA3-Flag	*Mus musculus*	Sjaak Philipsen	CMV-F
Gata2Δ(154-256)	pcDNA3-Flag	*Mus musculus*	Sjaak Philipsen	CMV-F
Gata2Δ(257-287)	pcDNA3-Flag	*Mus musculus*	Sjaak Philipsen	CMV-F
Gata2Δ(287-342)	pcDNA3-Flag	*Mus musculus*	Sjaak Philipsen	CMV-F
Gata2Δ(343-379)	pcDNA3-Flag	*Mus musculus*	Sjaak Philipsen	CMV-F
Gata2Δ(380-440)	pcDNA3-Flag	*Mus musculus*	Sjaak Philipsen	CMV-F
Gata2Δ(1-74)	pcDNA3-Flag	*Mus musculus*	Sjaak Philipsen	CMV-F
Gata2Δ(1-235)	pcDNA3-Flag	*Mus musculus*	Sjaak Philipsen	CMV-F
Gata2Δ(440-480)	pcDNA3-Flag	*Mus musculus*	Sjaak Philipsen	CMV-F

Gata2 WT and Gata2 mutants were individually sub-cloned into the pHAGE2-MCS backbone through the same PCR-based cloning process as defined for the paralog genes. In short, constructs were amplified by PCR, using the primers described in section 2.2.3.1 of chapter 2, separated by agarose gel electrophoresis (**Figure 17, A**) and purified. Then, both pHAGE2-MCS vector and inserts were restricted with the same restriction enzymes, followed by ligation and *E. coli* transformation with the recombinant plasmids.

In the presence of ligase, only a band corresponding to the recombinant plasmid is seen, while in its absence, both linear plasmid and ligand are seen (**Figure 17, B**). Colonies form transformed bacterial plates were screened for the presence of Gata2 and deletion mutants, by colony PCR (**Figure 17, C**). Positive and negative controls consisted in the original DNA template of each construct and a colony of bacteria transformed with empty pHAGE2, respectively. Lastly, recombinant plasmids were extracted from positive bacterial colonies, purified and Sanger sequencing confirmed sequences of the inserts. Sequencing results, organized in **Table 14,** were aligned with each individual CDS provided by our collaborator Dr. Sjaak Philipsen.

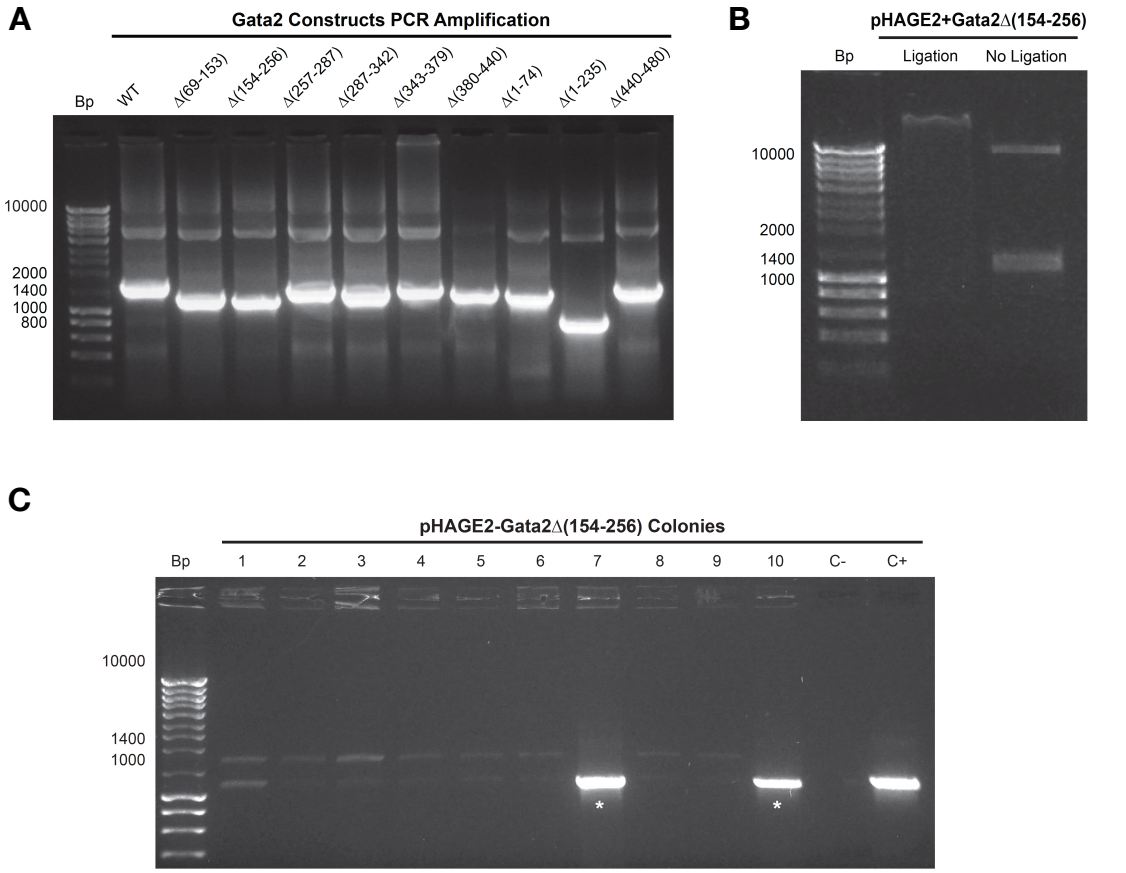

Figure 17. Gata2 constructs were successfully cloned into the pHAGE2-MCS vector. (**A**) Electrophoretic gel of Gata2 mouse gene and respective mutants after PCR amplification. DNA bands correspond to the size of each construct coding sequence. (**B**) Ligation process example. In the presence of ligase, Gata2 deletion (Δ) construct (154-256) is inserted into the pHAGE2 plasmid and no band for the mutant appears on the gel. The band of the Gata2 construct is shown when ligase is not present (**C**) Colony PCR performed to screen for positive colonies. Ten colonies from a plate containing bacteria transformed with pHAGE2-Gata2Δ(154-265) were picked and submitted to a PCR. The positive control (C+) corresponds to the template vector

of the Gata2 construct and the negative control (C-) to a colony picked from a plate after transformation with pHAGE2-MCS empty vector. Starts (*) indicate positive colonies. DNA band sizes are presented in base pairs (Bp).

Table 14. Sequencing data of Gata2 (deletion) constructs cloned into the pHAGE2-MCS vector. Sanger sequencing results were aligned with the sequences provided by Dr Skaak Philipsen via Vector NTI software. Quality of the alignment is shown in purple. Irregularities were verified for sequencing errors.

Cloning product	Sequencing data
Gata2	
Gata2Δ(69-153)	
Gata2Δ(154-256)	
Gata2Δ(257-287)	
Gata2Δ(287-342)	
Gata2Δ(343-379)	
Gata2Δ(380-440)	
Gata2Δ(1-74)	
Gata2Δ(1-235)	
Gata2Δ(440-480)	

3.4.2 SPECIFIC GATA2 DOMAINS ARE NECESSARY FOR HEMATOPOIETIC REPROGRAMMING

To produce the necessary lentiviral particles for DT MEF transduction, HEK293T cells were transfected with pHAGE2 transfer plasmids encoding Gata2 WT or deletion constructs, Gfi1b and cFos, plus the viral plasmids psPAX2 and pMD2.G. DT MEFs were transduced with the resulting lentiviral particles to induce activation of hCD34/H2BGFP reporter in a way that each construct replaced Gata2 WT for hematopoietic reprograming (**Figure 18, A**). Impact in nuclear GFP expression was assessed at day 20, for every condition (C-, WT, A-I), each letter corresponding to a determinate Gata2 mutant (**Figure 18, B**) used to replace Gata2 WT.

The percentage of GFP+ cells from each test condition was compared to the positive control Gata2 WT (plus Gfi1b and cFos) and negative control (C-) pHAGE2-MCS empty vector (**Figure 18, C**). Flow cytometry analysis shows that the deletion of the entire NRD, plus the N-terminal TAD (Gata2 Δ1-235) resulted in an evident decrease of GFP expression by approximately 6-fold. Furthermore, deletion of half NRD represented a decrease by half in the percentage of reprogrammed cells (Gata2 Δ69-153 and Gata2 Δ154-256), with Gata2 Δ(154-256) construct showing a slightly greater negative impact, when compared to Gata2 (Δ69-153) construct (**Figure 18, D**). Interestingly, Gata1 that couldn't replace Gata2 function for hematopoietic reprogramming does not contain the NRD or any related homologous region (Kaneko et al. 2012; Vicente et al. 2012), suggesting that the non-conserved NRD may play a critical role during hematopoietic reprogramming. Moreover, the percentage of GFP+ cells suffered a 4-fold decrease, when the N-terminal TAD alone was not present (Gata2 Δ1-74), as well as, when the C-terminal TAD was partially deleted (Gata2 Δ440-480), which may indicate that intact N and C-terminals are needed for Gata2 function of inducing HSC fate, or that the TADs are required for the activation of the transcription of hematopoietic genes through interaction with co-activator proteins. As expected, the deletion of the NLS domain (Gata2 Δ380-440) also impacted DT MEF reprogramming, causing a decrease by nearly 4-fold in the percentage of GFP+ cells, probably due to lack of nuclear translocation. On the contrary, the adjacent region upstream the N-ZF (Gata2 Δ257-287) does not seem to contribute for Gata2 function in our reprogramming system, since its deletion does not decrease the nuclear GFP expression. Regarding the conserved regions between GATA proteins, deletion of the C-ZF (Gata2 Δ343-379), causes a decrease in GFP expression by 7-fold, the highest decrease presented by any of the deletion constructs, while the deletion of the N-ZF (Gata2 Δ287-342) resulted in less than 1-fold decrease, suggesting that the C-terminal ZF (C-ZF), but not the N-terminal ZF

(N-ZF) is crucial for a successful hematopoietic reprogramming. Interestingly, despite homologous, the N-ZF and the C-ZF do not display the same DNA-binding properties (Whyatt et al. 1993; Yang & Evans 1992; Trainor et al. 2000).

Summarily, these results suggest that both conserved and non-conserved domains of Gata2 are required for Gata2 function during hematopoietic reprogramming, with particular interest the conserved C-ZF, and the non-conserved TADs and NRD, specifically its second half (**Figure 18, D**).

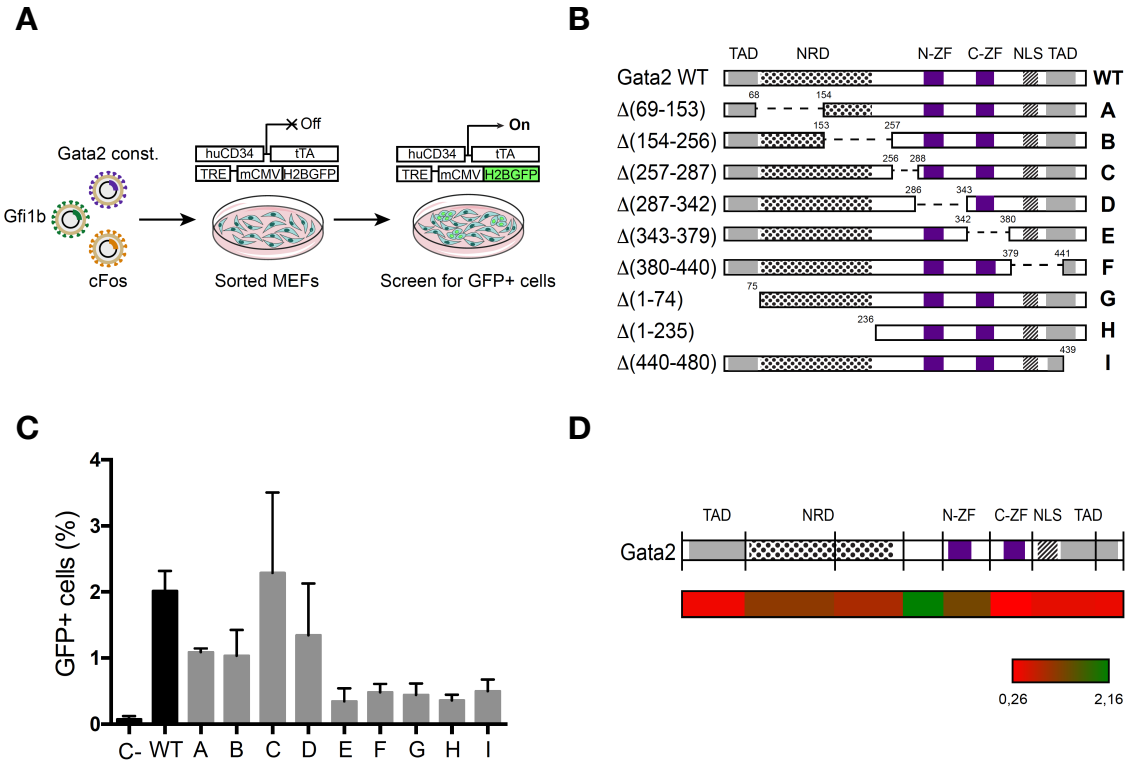

Figure 18. The transactivation domains (TADs), the negative regulatory domain (NRD), the C-terminal zinc finger (C-ZF) and the nuclear localization signal (NLS) of Gata2 are necessary for hematopoietic reprogramming. (**A**) Scheme of hematopoietic reprogramming with Gata2 deletion constructs. (**B**) Schematic representation of the protein domain organization of Gata2 wild type (WT) and Gata2 deletion mutants. Deleted aminoacids are indicated between parentheses. Δ stands for "deletion". The letters A to I correspond to the construct that precedes them. (**C**) Flow cytometry analysis of HSC-induced reprogramming by Gata2 deletion construct (A-I) and controls. The WT condition represents the positive control (MEFs transduced with Gata2 WT, Gfi1b and cFos) and pHAGE2-MCS empty vector the negative control (C-). (**D**) Heat map showing the impact of each deletion construct in hematopoietic reprogramming and the correspondent deleted region of Gata2 protein. Red indicates lack of reprogramming and green indicates reprogramming. Numbers represent the lowest and the highest percentage, on average, of reprogrammed cells. TAD,

transactivation domain. NRD, negative regulatory domain. N-ZF, N-terminal zinc finger. C-ZF, C-terminal zinc finger. NLS, nuclear localization signal. Results from one experiment with two replicates per condition.

3.5 Mitotic bookmarking by hematopoietic reprogramming factor GATA2

In order to investigate if Gata2 exerts mitotic bookmarking activity as an epigenetic feature, subcellular localization of Gata2 in mitotically arrested cells was assessed. For this, Gata2-fluorescent protein fusion constructs were generated. Constructs were achieved by fusing the CDS of the red fluorescent protein mCherry with the CDS of our target protein Gata2, via PCR-based cloning into the pHAGE2-MCS vector. Mitotic bookmarking ability of Gata2 TF was assessed after HEK293T cell transfection and arrest in mitosis. To allow the visualization of mitotic chromosomes, DNA was stained and cells were analyzed by fluorescence microscopy.

3.5.1 PCR-based cloning of GATA2 fusion constructs into pHAGE2-MCS vector

mCherry CDS was fused to either the 5' end (mCherry-Gata2) or the 3' end (Gata2-mCherry) of Gata2 CDS to take into account possible interferences of mCherry on Gata2 protein folding and function (Snapp 2005). For this purpose, the CDSs of Gata2 (mouse), which was previously cloned and the CDS of the already available mCherry (pHAGE2-MCS-IRES-mCherry) were amplified by PCR twice with two different sets of primers, as detailed in section 2.2.3.1 of chapter 2. To clone Gata2-mCherry fusion construct into the MCS of pHAGE2 vector, NotI restriction site was added to the forward primer of Gata2 and XbaI restriction site to the reverse primer of mCherry. An additional restriction site was created to link both sequences. BstBI restriction site was added to the reverse primer of Gata2 and to the forward primer of mCherry. The opposite was done for mCherry-Gata2 fusion constructs. After amplification, two different PCR products for each gene were obtained and separated by electrophoresis. Bands were purified and double restriction with the corresponding restriction enzymes was performed on both plasmid and inserts. Following restriction, DNAs with cohesive ends were all ligated in the same reaction (vector plus Gata2 and mCherry). A prove of the success of ligations lies in the inexistence of DNA bands corresponding to the inserts when ligase is in the reaction, as opposed to the control situation when ligase

is lacking (**Figure 19, A**). Competent bacteria were transformed with the recombinant plasmids pHAGE2-Gata2-mCherry or pHAGE2-mCherry-Gata2, and isolated colonies were submitted to a colony PCR. This time, two positive controls were done to certify the cloning of both ligands into the vector (**Figure 19, B**). To ensure fusion constructs were generated, an extra PCR was performed after recombinant plasmid extraction form positive colonies, using the forward primer that contained the NotI restriction site and the reverse primer that comprised the XbaI restriction site. Electrophoretic gel showed only one band with the sum of the inserts' sizes, confirming the assembling of the fusion constructs (**Figure 19, C**). Gata2 fusion constructs from positive colonies were also sent for sequencing and sequencing results were aligned with Gata2 or mCherry, depending of which sequence was downstream the human EF1-α promoter (data not shown).

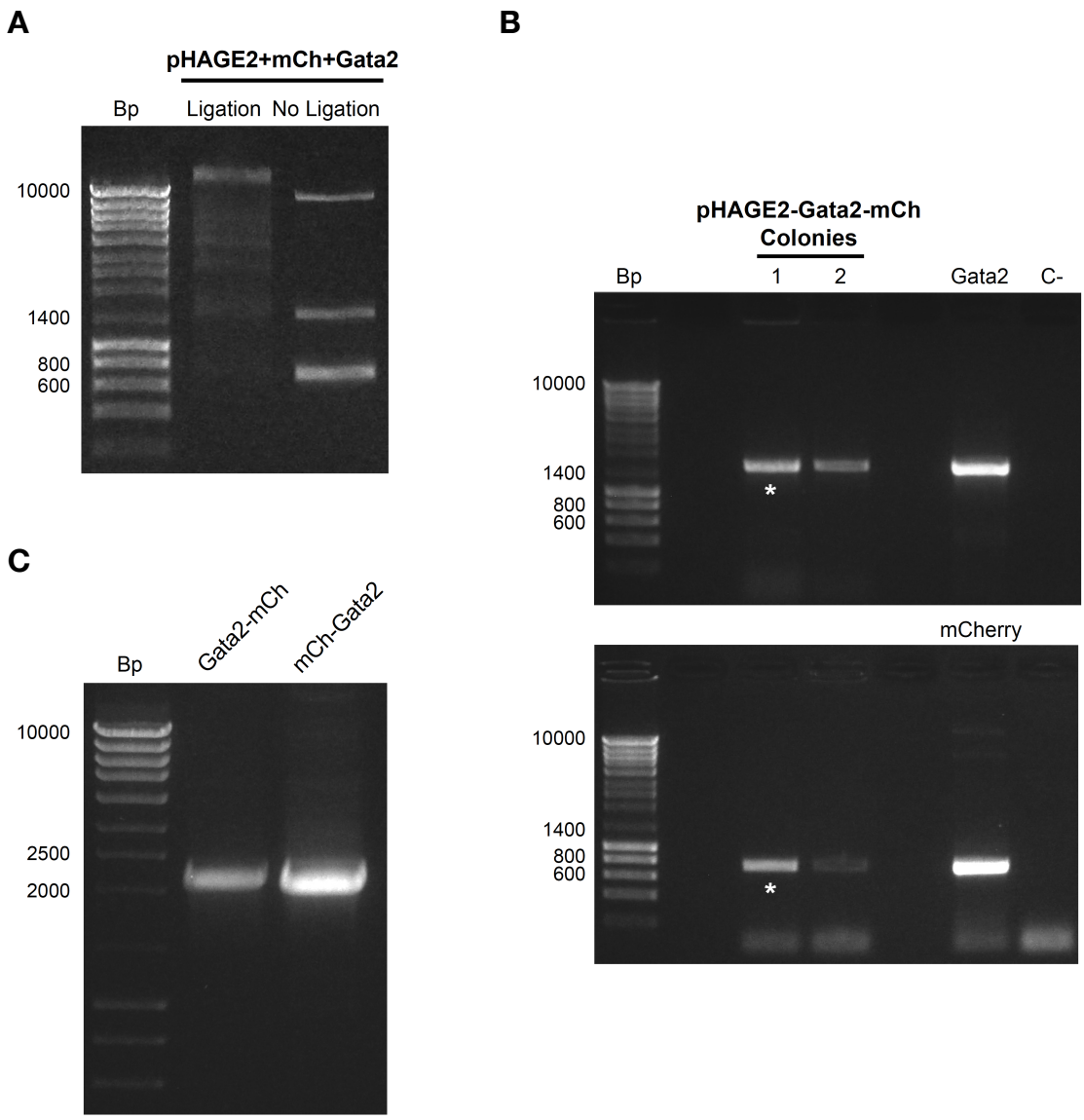

Figure 19. Screening of bacterial clones harbouring pHAGE2 with Gata2 fusion constructs and confirmation of constructs creation by PCR. (**A**) Ligation of mCherry and Gata2 to pHAGE2-MCS vector. When no ligase was added, bands for the three restricted DNAs appeared in the electrophoretic gel. In the "No Ligation" situation, the first band (top to bottom) corresponds to the vector, the second one to Gata2 and the last on to mCherry (**B**) Colony PCR performed to screen for positive colonies for both genes. Two colonies from a plate containing bacteria transformed with pHAGE2-Gata2-mCherry were picked and submitted to colony PCR. Two positive controls were performed, one with each template vector of the mCherry or Gata2 sequences. pHAGE2-MCS empty vector was the colony PCR's negative control (C-). Starts (*) indicate the positive colony. (**C**) Verification of the creation of Gata2 fusion constructs, by PCR. Fusion constructs cloned into the pHAGE2-MCS vector were amplified PCR with the forward primer of the first gene and the reverse primer of the second gene (5' to 3' direction). Bands have the combined size of the two genes. DNA band sizes are presented in base pairs (Bp). mCh, mCherry.

3.5.2 Optimization of HEK293T mitotic arrest with nocodazole

Nocodazole is commonly used in Cell Biology to synchronize cells at the G2 to mitosis (G2/M) transition of the cell cycle by reversibly inhibiting the polymerization of microtubules (Matsui et al. 2012; Zieve et al. 1980). Concentrations raging from 40 to 200ng/mL and short exposure times are advised to avoid cytotoxicity (Schorl & Sedivy 2007). G2 or mitotic DNA content can be easily assessed by flow cytometry after staining with propidium iodide (Krishan 1975; Harper 2005). To address the effect of nocodazole in HEK293T cell cycle, cells were treated with 50 or 100ng/mL of nocodazole during 4, 16 or 24h, after which cells were fixated and stained with propidium iodide for flow cytometry analysis. Percentages of HEK293T cells synchronized in G2/M transition after nocodazole treatment are displayed in **Figure 20**. Control conditions without nocodazole show more cells in interphase (G1) than in G2/M. Exposure to Nocodazole for 4h was not enough to increase significantly the G2/M cell population (maximum of 32,8% with 100ng/mL of nocodazole). Contrarily, 16 or 24h of treatment resulted in a greater enrichment of cell population in G2/M, when compared to the 4-hour treatment. The percentage of cells in G2/M was slightly higher when cells were submitted to a 16-hour treatment, for both concentrations (56,9 to 70,8% and 66,2 to 72,5%). Regarding nocodazole concentration, 100ng/mL lead to a higher percentage of arrested cells in all exposure times. 16-hour treatment with 100ng/mL of nocodazole was selected as the optimal combination generating 72,5% of cells arrested in G2/M transition.

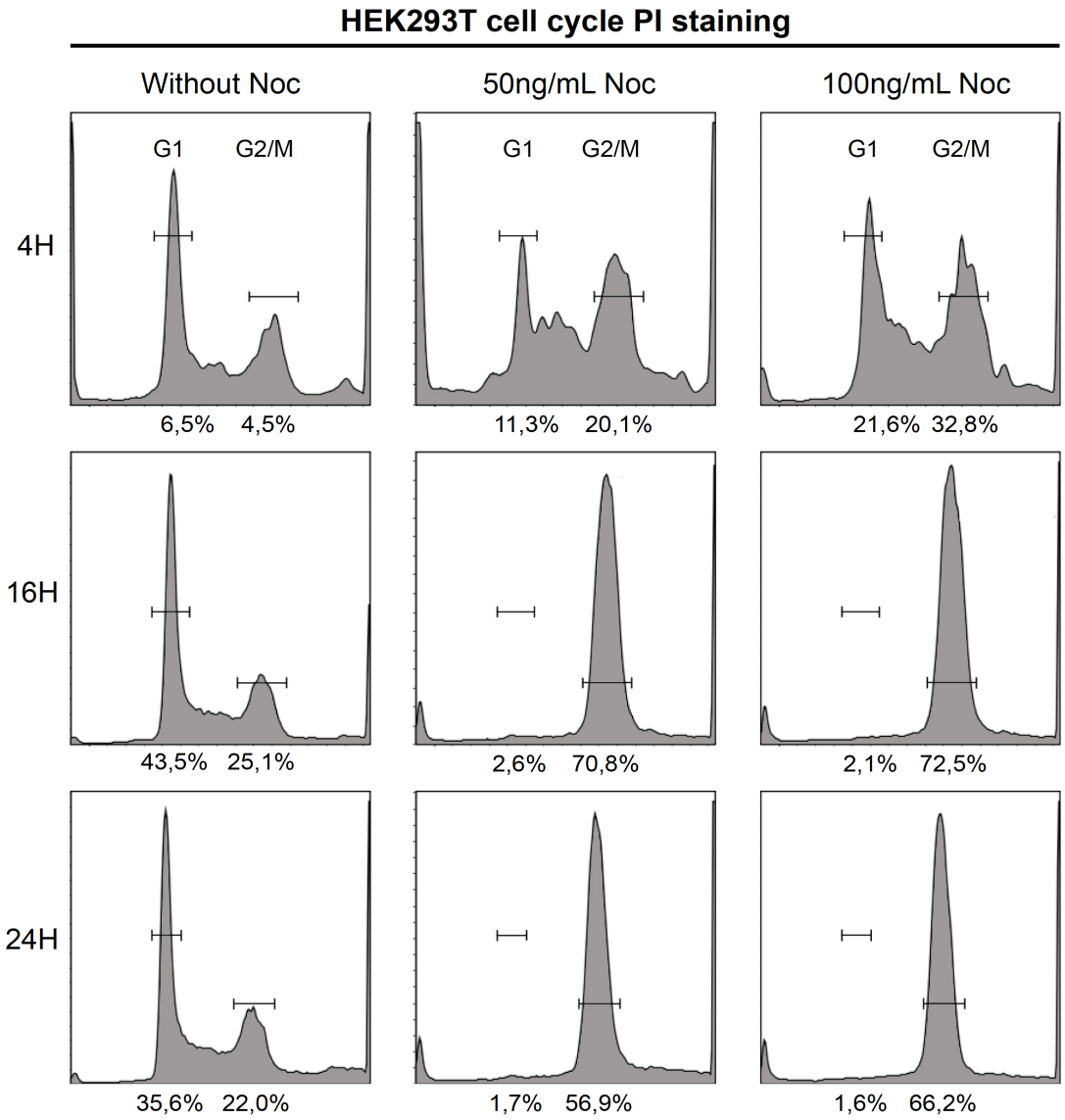

Figure 20. Analysis with propidium iodide (PI) staining of HEK293T arrested cells after nocodazole treatment. Graphic representation of HEK293T cell cycle after 4, 16 or 24h of incubation in DMEM medium containing 100ng/mL or 50ng/mL of nocodazole.

3.5.3 GATA2 DISPLAYS MITOTIC BOOKMARKING ACTIVITY

HEK293T cell plates were transfected with pHAGE2 recombinant plasmids encoding Gata2 fusion constructs or the pHAGE2-MCS-IRES-mCherry control, followed by mitotic arrest with 100ng/mL of nocodazole for 16h. After nocodazole treatment, DNA was stained with Hoechst blue fluorescent dye and live cells were visualized under confocal fluorescence microscopy (**Figure 21, A**). Confocal fluorescent snapshots of synchronized cells transfected with the Gata2 fusion constructs showed co-localization of

Gata2 with mitotic chromosomes (in blue). Inversely, mCherry protein alone was dispersed throughout the cell cytoplasm (**Figure 21, B**). 100% of transfected cells visualized demonstrated the same pattern of Gata2 localization, as well as mCherry (control) distribution. These demonstrate mitotic bookmarking activity by Gata2 and suggest that this might be an important feature for HSC fate maintenance and for the transmittance of the reprogrammed cell state during hematopoietic reprogramming.

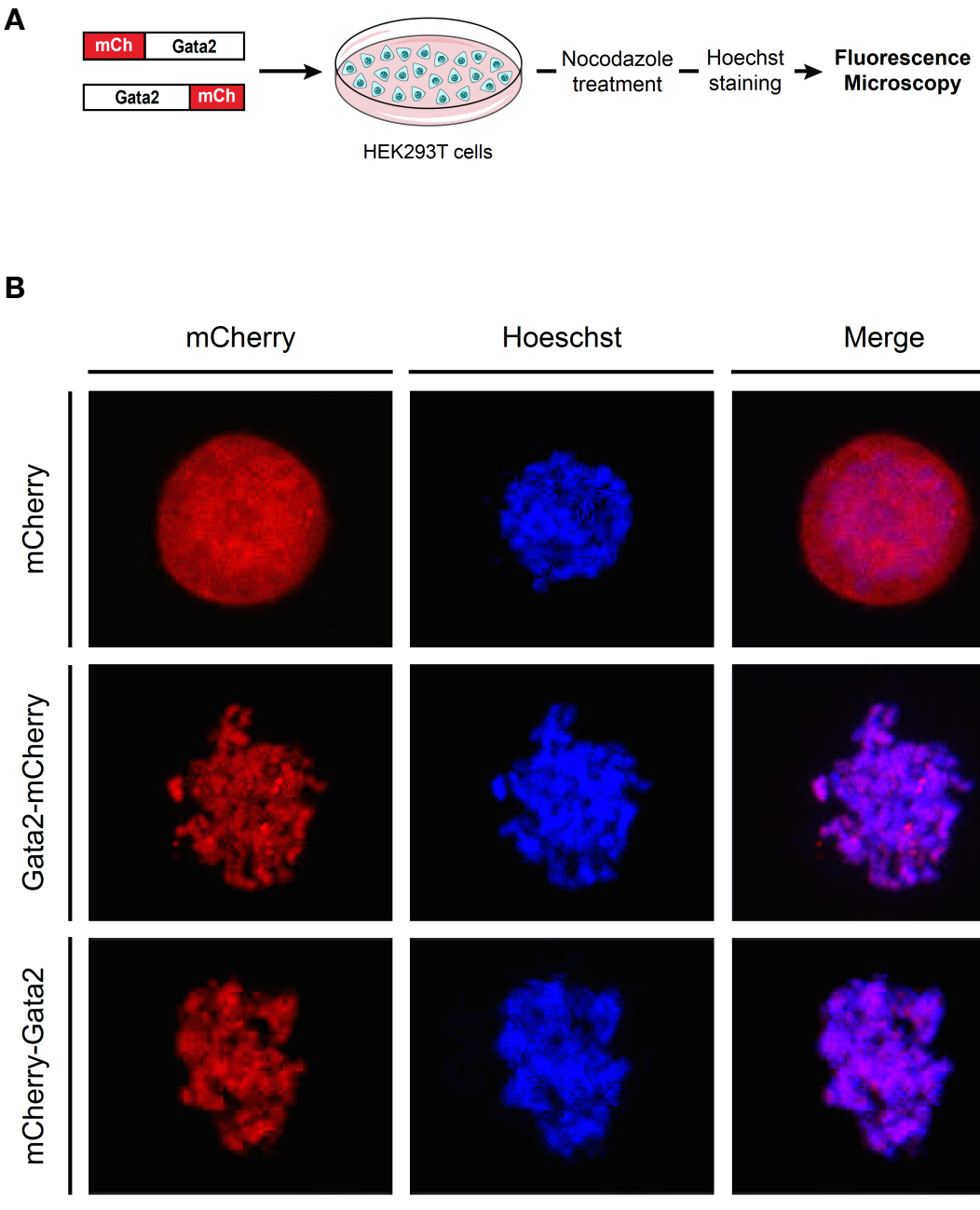

Figure 21. Gata2 co-localizates with mitotic chromosomes. (**A**) Experimental approach for HEK293T cell transfection with Gata2 fusion constructs. Transfected plates were submitted to a 100ng/mL nocodazole treatment for 16h to allow mitotic

arrest and stained with Hoechst before imaging. mCh, mCherry. (**B**) Confocal microscopy of live, synchronized (G2/M arrested) HEK293T cells expressing mCherry-Gata2 or Gata2-mCherry. Gata2 (red) co-localized with mitotic chromatin (blue). Fluorescent snapshot correspond to maximum intensity projection images. mCherry was used as control.

CHAPTER 4

4. DISCUSSION

Conversion of somatic cells to iPSCs by ectopic expression of specific TFs granted, in 2012, the Noble Prize in Medicine to Shinya Yamanaka. This landmark study revolutionized the cell reprogramming field and introduced a powerful tool for the creation of personalized cellular therapies. In theory, derivation of pluripotent cells from donors would enable the generation of any type of tissue required for regenerative medicine applications. The implications of Takahashi and Yamanaka's work were far from over, as their achievements inspired researchers all over the world to manipulated cell fate in exciting new ways. Since then, overexpression of TFs was applied in the direct reprogramming of one differentiated cell type to another. Starting mainly from fibroblasts, it has already been possible to obtain several desired cell types, including cardiomyocites, neurons, hepatocytes and blood cells. Inside the hematopoietic field, HSCs are certainly an interesting target to achieve via direct reprogramming, due to their self-renewal and differentiation capacities. Enforced expression of three TFs i.e., Gata2, Gfi1b and cFos, induced a hemogenic program in mouse fibroblasts, which led to the ultimately emergence of hematopoietic stem-like cells. Although hematopoietic reprogramming constitutes a promising alternative to current therapies for the generation of transplantable HSCs, the efficiency of the process remains low. A better understanding of the mode of action of the hematopoietic reprogramming factors will allow the creation of strategies to improve the process's efficiency. In that sense, this study aimed to unveil Gata2, Gfi1b and cFos functional (genetic and epigenetic) features for hematopoietic reprogramming of fibroblasts.

To reprogram fibroblast cellular identity towards the hematopoietic lineage, the CDSs of the three factors were initially delivered via retroviral delivery system. However, I have established a comparable lentiviral system for gene deliver in order to overcame the retrovirus-use related drawbacks. Lentiviral plasmids pHAGE2-TFs were able to reprogram fibroblasts to hematopoietic stem-like cells with a reprogramming efficiency similar to retroviral plasmids.

In order to dissect Gata2, Gfi1b and cFos function during hemogenic/hematopoietic reprogramming of fibroblasts, each factor was replaced by their respective paralog genes (homologous genes originated from gene duplication). This strategy was based on TF substitution approaches performed to study the function of the original cell reprogramming factors: Oct4, Sox2, Klf4, and c-Myc in reprogramming to pluripotency (Zeng et al. 2004; Montserrat et al. 2013; Shu et al. 2013). Paralog genes share specific regions and domains, thus they can be used as tools to infer which parts of

each TF are necessary for hematopoietic reprogramming by domain-domain comparison. Gata1, Gata3 (Gata2 paralogs), Gfi1 (Gfi1b paralog), FosB and Fra2 (cFos paralogs) were cloned into the lentiviral pHAGE2-MCS vector and assessed for their ability to activate the hCD34/H2BGFP reporter.

Flow cytometry analysis showed that Gata1 was not able to substitute Gata2, as opposed to Gata3, which was able to partially replace Gata2 function for hematopoietic reprogramming. Hence, one can assume that domains that are not conserved among proteins may be contributing for the disparities in the activation of the reporter. In fact, the NRD that is present in the Gata2 protein does not exist in Gata1 or Gata3. The NRD is responsible for an inhibitory function of Gata2 in hematopoietic differentiation, helping maintaining the population of immature hematopoietic cells, however the mechanism behind such function is yet to be uncovered (Minegishi et al. 2003). On the contrary, Gata1 is crucial for erythroid differentiation and maturation, and instead of the NRD it has a N-terminal TAD, that together with the recently discovered C-terminal TAD, regulates the transcription of erythroid genes (Kaneko et al. 2012). Moreover, despite Gata1 and Gata2 share overlapping functions during primitive hematopoiesis (Fujiwara et al. 2004), Gata2 expression is predominant in HSPCs, whereas Gata1 expression increases upon the onset of erythroid differentiation, replacing Gata2 at GATA binding sites in Gata2 locus and inhibiting its transcription (Ohneda & Yamamoto 2002; Suzuki et al. 2011). Inversely, Gata2 suppresses Gata1 expression in HSCs. That said, it was not surprising that Gata1 couldn't replace Gata2 function for the generation of hematopoietic stem-like cells. Gata3, on the other hand, besides being essential to T cell development, is also found in murine long-term (LT)-HSCs (Zhong et al. 2005) were it controls LT-HSC homeostasis and cell-cycle entry (Ku et al. 2012). Furthermore, although the NRD is absent in Gata3 protein, it has a homology of 46% for the corresponding region, which may explain why Gata3, but not Gata1 was capable to partially compensate the hematopoietic reprogramming role of Gata2.

Gfi1b substitution by Gfi1 for hematopoietic reprogramming resulted in less than 1-fold decrease in the percentage of reprogrammed cells, suggesting that Gfi1 shares enough similarities with Gfi1b to prompt changes in fibroblast's cell identity towards the hematopoietic lineage. In fact, Gfi1 and Gfi1b display overlapping functions in the hematopoietic system. Gfi1b almost completely recued Gfi1 function for blood cell formation, in Gfi1 double knockout mice (Fiolka et al. 2006). Moreover, at least one of the proteins is absolutely necessary for HSC survival (Khandanpour et al. 2011). This may be due to the existence of highly homologous regions (the SNAG domain and the ZFs in the DNA binding domain). Mechanistically, both proteins repress their target genes through the recruitment of members of the CoREST complex, namely Lsd1 and RCOR1/2 cofactors, to promoter sites and regulatory

regions, and this interaction is mediated by the SNAG domain. In fact, Thambyrajah and colleagues have recently reported that Gfi1 and Gfi1b recruit Lsd1 to inhibit the transcription of endothelial genes in hemogenic endothelium and allow EHT (Thambyrajah et al. 2015). With that in mind, Gfi1 derived hematopoietic reprogramming results seem to be in agreement with the literature, in a way that they can be explained by the structural, mechanistic and functional resemblances between proteins. However, despite the similarities, Gfi1 and Gfi1b have distinct expression patterns and unique functions that may be due to the existence of the non-conserved region. While Gfi1b is crucial for erythroid and megakaryocyte differentiation, Gfi1 regulates neutrophil differentiation and early T-cell development. Perhaps, it would be interesting to create deletion constructs of the intermediate region and investigate the impact of each mutant in our hematopoietic reprogramming system. That would not only complement our understanding about Gfi1b's role for hematopoietic reprogramming of fibroblasts, but would also provide new insights for the functional characterization of that specific region, during the specification of definitive hematopoiesis.

Regarding hematopoietic reprogramming with cFos paralogs, FosB was shown to completely substitute cFos, whereas Fra2 only replaced cFos reprogramming function partially. Gfi1 and FosB ability to substitute for hematopoietic reprograming is in agreement with two other direct reprogramming studies, in which both factors were part of the TF pool used to convert endothelial cells into engraftable mouse and human HSPCs (Sandler et al. 2014; Lis et al. 2017). The major difference between the domain organization of cFos paralog proteins is the absence of a conserved TAD in Fra2, present in cFos and FosB. As a matter of fact, cFos was shown to interact through this domain with the N-terminal TAD of GATA proteins, in order to potentiate tissue-specific gene transcription (McBride et al. 2003). This is in agreement with the mode of action of the three TFs, proposed by Pereira and colleagues. These data suggest that, in part, hematopoietic reprogramming relies on cFos interaction with GATA proteins, independently of the bZIP domain (comprises the leucine zipper and the DNA-binding motif). Nevertheless, Fra2 partially rescued cFos function for hematopoietic reprogramming, thus the conserved bZIP, necessary for the assemble of the AP-1 complex or other unexplored homologous regions in the N or C-terminal of Fos proteins may be implicated during hematopoietic reprogramming.

Since Gata2 is allegedly the dominant factor of our hematopoietic reprogramming system and Gata1 failed to substitute Gata2, I sought to further dissect Gata2 domains relevance for hematopoietic reprogramming using Gata2 deletion constructs. Hence, each Gata2 mutant without a specific coding region replaced Gata2, in direct reprogramming of fibroblasts. Deletion of the first 235 aminoacids (Gata2 D1-235), which removed the entire NRD and

N-terminal TAD, led to a decrease in the percentage of reprogramed cells by 6-fold, the same fold decreased observed when Gata2 was replaced by Gata1. It is noteworthy that partial deletions of the NRD were sufficient to negatively impact the percentage of GFP+ cells. Furthermore, deletion of the N-terminal TAD alone (Gata2 D1-74) caused a 4-fold decrease in GFP expression, which is consistent with the literature, since it has been reported that N-terminal TAD from GATA proteins interact with the N-terminal TAD of cFos to induce GATA-dependent gene transcription. Interestingly, partial removal of the C-terminal TAD (Gata2 D440-480) caused the same fold decrease as the N-terminal TAD deletion, possibly meaning that Gata2 contacts through that domain with the same or other proteins to, individually or together with the N-terminal TAD, activate transcription from target promoters. Truncation of Gata2 in the N or C-terminal could also impair the overall folding and function of the protein, and thus explained the reduction in GFP levels. Gata2 D380-440 decreased by 4-fold the percentage of reprogramming, probably due to lack of protein translocation to the nucleus caused by exclusion of the NLS domain, or the deletion of the first half of the C-terminal TAD. It is also possible that this construct failed in replacing Gata2 WT due to a combination of impaired nuclear translocation and transactivation potential.

Inside the DNA binding domain of Gata2, two deletion constructs were tested, one without the N-terminal ZF (Gata2 D287-342) and other without the C-terminal ZF (Gata2 D343-379). Cytometry results suggested distinct functions of the ZFs for Gata2 reprogramming function. While deleting the N-ZF resulted in less than 1-fold decrease in the percentage of reprogrammed cells, removal the C-ZF provoked a 7-fold decrease, proposing a prominent role of C-ZF of Gata2 for hematopoietic reprogramming. This is consistent with the features described for each ZF in GATA proteins. The N-ZF is responsible for stabilizing DNA-protein complexes and providing specificity to DNA binding, whereas C-ZF recognizes and binds to GATA consensus sequence (Whyatt et al. 1993; Yang & Evans 1992; Trainor et al. 2000). Actually, the C-ZF of Gata1 or Gata2 alone was enough to promote megakaryocytic differentiation from a primitive myeloid cell line (Visvader et al. 1995). Therefore, erasing the zinc finger that contributes the most to DNA binding may indeed reduce Gata2 role in promoting HSC fate in fibroblasts. Besides the abovementioned functions, both ZFs are known to interact separately or collectively with other proteins (Vicente et al. 2012), namely Lsd1 and CoREST1. Co-immunoprecipitation after overexpressing the same Gata2 deletion constructs in HEK293T cells revealed the formation of Gata2 complexes with endogenous Lsd1 and CoREST1 proteins, except when Gata2 mutants lacking the C-ZF or the NLS domain (Braun, unpublished). As a member of the CoREST complex, Lsd1 interacts with co-repressors to inhibit the transcription of target genes (Lan et al. 2008) and has been implicated in the transcriptional repression of

multipotency-related genes during hematopoietic differentiation (Kerenyi et al. 2013). Recent work from Guo *et al.* reported a connection between Gata2, Lsd1 and Gata1 essential for a tight control of erythropoiesis. Lsd1 knockdown in MEL (murine erythroleukemia) cells and ESCs led to augmented expression of Gata2 and diminished expression of Gata1, whereas control cells exhibited low levels of Gata2 and high levels of Gata1, once differentiation was induced(Guo et al. 2015). It was proposed that during erythroid differentiation, Lsd1 is recruited to Gata2 promoter site via T cell acute lymphocytic leukemia 1 (Tal1) protein, causing the demethylation of histone H3 in lysine 4 (H3K4) and epigenetically decreasing Gata2 expression. Inversely, Gata2 would form protein complexes with Lsd1 to repress Gata1 expression, in order to maintain the undifferentiated cell state (Guo et al. 2016). Co-immunoprecipitation assays from nuclear extracts of K562 cells proved the physical association between Gata2 and Lsd1, as well as between Gata2 and other constituents of the CoREST complex including CoREST proteins plus histone deacetylase 1 (Hdac1) and 2 (Hdac2). In undifferentiated cells, the knockdown of Lsd1, Gata2, or both genes caused a substantial increase in Gata1 expression. The authors suggested that Gata2 recruits Lsd1 to regulatory regions of the Gata1 and inhibits its expression through Lsd1-mediated H3K4 demethylation. Upon induced erythroid differentiation, the interaction between Gata2 and Lsd1 decreased. These studies not only support the results from Gata1 substitution for hematopoietic reprogramming, but also indicate that the C-ZF of Gata2 is responsible for the formation of Gata2-Lsd1 repressive complexes, which seem fundamental for the preservation of the multipotent state of HSCs. From this point of view, hematopoietic reprogramming may also require the establishment of such protein complexes for reporter activation, and so it would be interesting to study the effect of Lsd1 inhibition with our system. Furthermore, since Gfi1b also interacts with the CoREST complex, it is possible that, during hematopoietic reprogramming, Gfi1b and Gata2 are indirectly linked through the LSD1/CoREST complex. Data from our laboratory has already revealed a physical association between Gata2 and Gfi1b, however the domains responsible for that interaction remain unknown. Taken together, these results help elucidate functional aspects of Gata2, Gfi1b and cFos domains for hematopoietic reprogramming, and allowed the design of modified reprogramming factors for efficient hematopoietic reprogramming (**Figure 22**). The Gata2-based improved factor was design to comprise the protein domains which deletion led to a decrease in HSC-induced reprogramming (TADs, NRD, C-ZF and NLS) plus the N-ZF to stabilize the protein binding to DNA. (**Figure 22, A**). Duplication of the C-ZF was considered, since the deletion of this region caused the highest decrease in the percentage of reprogrammed cells. An extra C-ZF will possibly enhance TF DNA binding to target gene promoters. Since, Gfi1b has a shorter middle

region comparing to Gfi1 and is more successful in yielding GFP+ cells, it would be interesting to address whether the deletion of the entire intermediate region would result in higher percentage of reprogramed cells (**Figure 22, B**). As mentioned before, cFos forms the AP-1 complex with cJun to play general and tissue-specific roles, hence fusion proteins of cFos-cJun could be generated to potentiate the assemble and the performance of the protein complex for inducing hematopoietic reprogramming (**Figure 22, C**). FosB, which is smaller than cFos can be used to replace its function in the AP-1 complex, as the results from paralog substitution for hematopoietic reprogramming indicate. Nonetheless, the direct reprogramming experiments performed under the scope of this study should be repeated to increase statistical power.

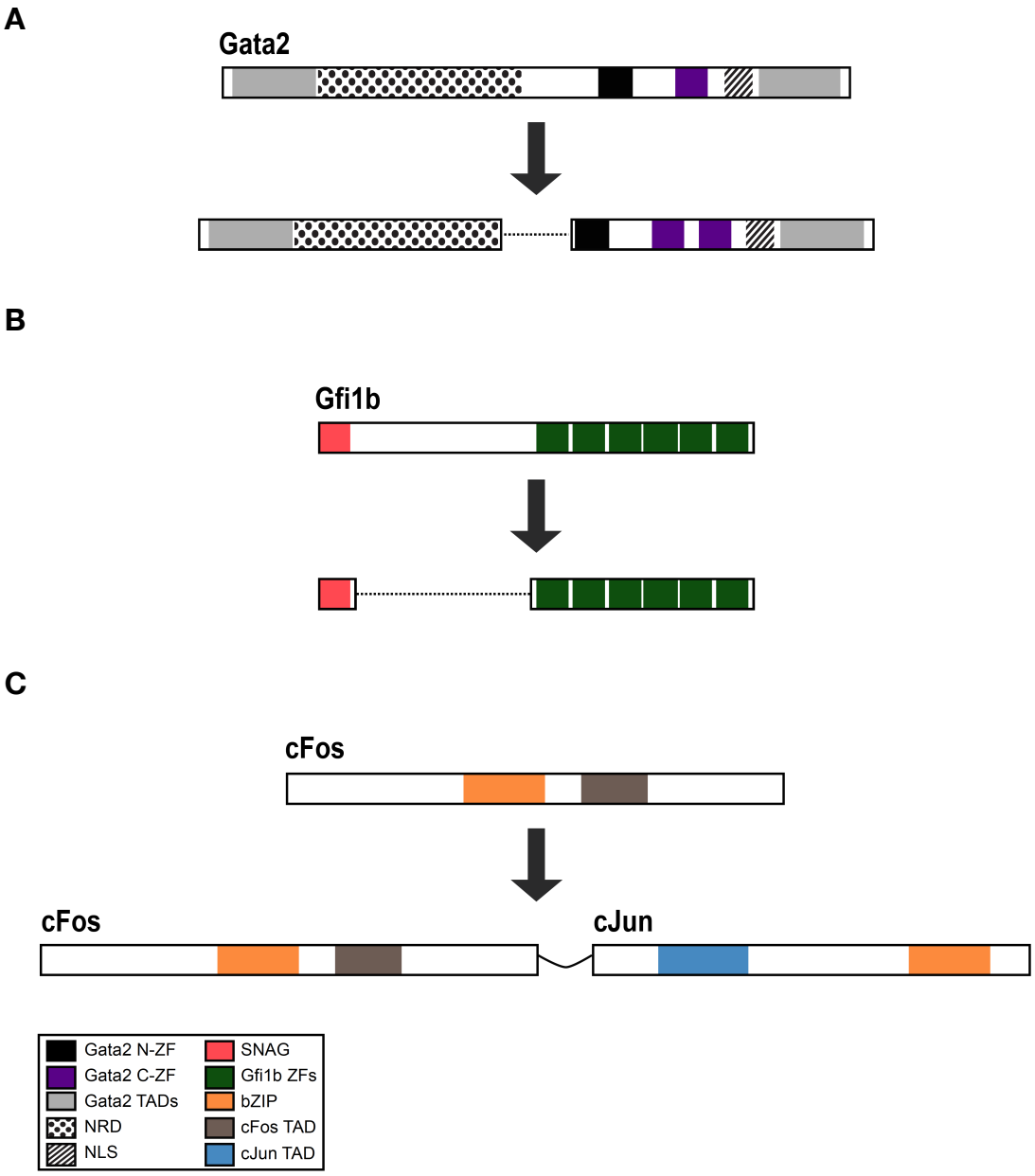

Figure 22. Schematic representation of modified reprogramming factors for efficient hematopoietic reprogramming. (**A**) Gata2-based improved reprogramming factor. This factors contains Gata2 domains that were shown relevant for hematopoietic reprogramming, and C-ZF is duplicated. (**B**) Enhanced Gfi1b-based factor. This putative improved factor includes only the two conserved regions, as the middle non-homologous region of Gfi1b and Gfi1 was removed. (**C**) cFos-cJun fusion protein. cFos can be fused with cJun to unable the expression of the AP-1 factors a whole.

Like histone modifications, transcription factors can leave their mark on mitotic chromatin in a process termed mitotic bookmarking, to preserve epigenetic cell memory. Recent studies suggest that the same happens during cell reprogramming. Therefore, I sought to investigate whether Gata2 exhibited

a mitotic bookmarking feature that could be associated with the transmittance of transcription patterns in reprogrammed cells. My efforts were directed towards Gata2 because of the importance that it has demonstrated throughout this study and also because it most likely recruits the other factors to its targets to allow fibroblast reprogramming. Additionally, two other GATA proteins: Gata1 and Gata4 were already reported to bind to mitotic chromatin, during mitosis. Of note, both proteins served as reprogramming factors in direct reprogramming into distinct cell types. To assess a possible mitotic bookmarking activity, Gata2 CDS was fused to mCherry CDS to create fluorescent fusion constructs/proteins, and mitotically arrested HEK293T cells were transfected. Live-cell imaging revealed that Gata2 co-localizes with mitotic chromosomes when cells are transfected with either Gata2-mCherry or m-Cherry-Gata2 constructs. mCherry position in the fusion construct did not affect Gata2 chromosome binding ability. Cell fixation was intentionally not performed, since general formaldehyde fixation protocols resulted in the exclusion of TFs from mitotic chromosomes, which were later demonstrated to remain bound by other techniques, such as live cell imaging used in this work (Teves et al. 2016). It was proposed that the problem resided in the rapid inward formation of cross-links between formaldehyde molecules and the TF which would deplete the cytoplasmic pool of TFs available to bind to chromatin. By using fluorescent images of living cells, I avoided misleading results that could derive from formaldehyde fixation artefacts, and demonstrated mitotic bookmarking by Gata2. The relevance of a mitotic bookmarking activity displayed by Gata2 are yet to be investigated, however there are evidence that this epigenetic feature is required for proper maintenance of pluripotency in ESCs and to preserve lineage commitment in differentiated cells, through rapid reactivation of bookmarked genes' expression, in daughter cells. For example, mitotic bookmarking by the pluripotency factor Esrrb leads to up-regulation of ESC self-renewal genes, such as Klf4, when cells restart interphase (Festuccia et al. 2016). Chromatin immunoprecipitation followed by DNA sequencing (Chip-seq) analysis of mitotically arrested ESCs indicated that Esrrb, Kfl4, Sox2 and Oct4 pluripotency TFs remained mainly bound to their corresponding DNA-binding motifs, implying a direct and specific binding to mitotic chromatin sites (Festuccia et al. 2016; Liu et al. 2017). Likewise, Gata1 requires its DNA-binding domain to interact with mitotic chromatin and to prompt the transcriptional activation of erythroid genes (Kadauke et al. 2012). In that sense, one could hypothesize a similar mechanism for mitotic bookmarking by Gata2 in HSCs, as this factor regulates self-renewal and proliferation (**Figure 23**). Gata2 DNA-binding ability provided by the ZF domains should be sufficient to keep Gata2 retained in GATA consensus motifs of HSC self-renewal gene promoters and regulatory regions, throughout cell division.

Mitotic bookmarking by Gata2 could constitute a valuable epigenetic regulatory mechanism for the production and control of a stable pool of undifferentiated hematopoietic cells during development.

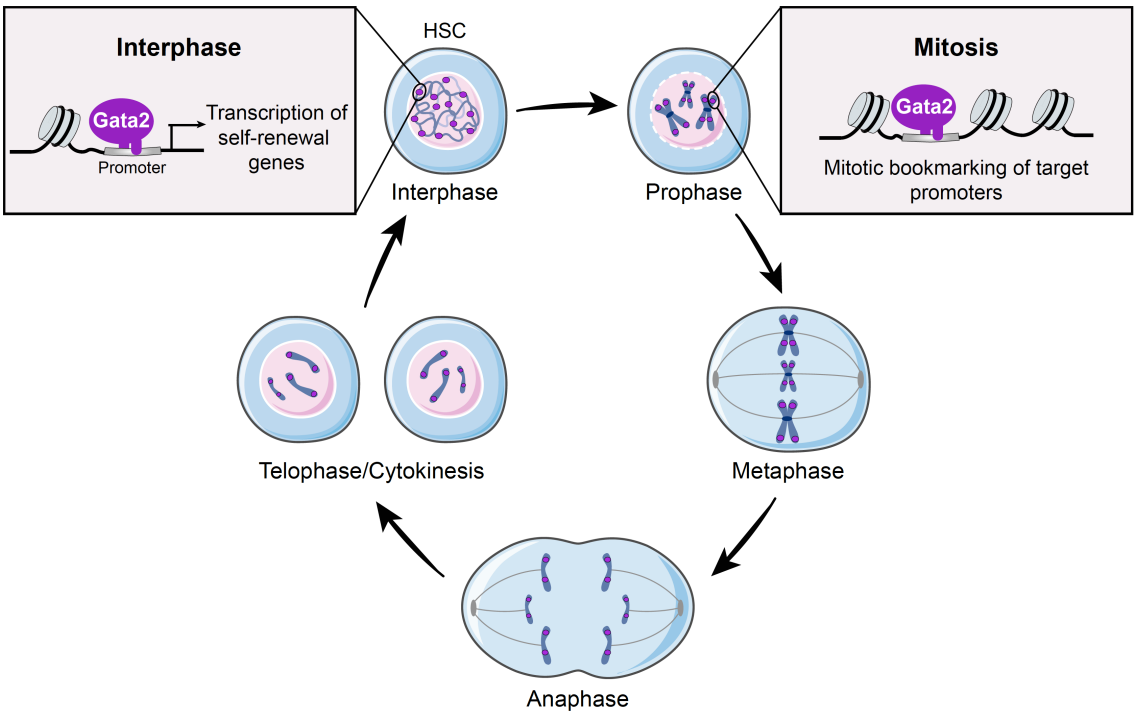

Figure 23. Schematic representation of a possible role for mitotic bookmarking by Gata2 in HSC self-renewal. During interphase, Gata2 would bind to GATA consensus sites in promoters or enhancers of target self-renewal/proliferation-associated genes, through the zinc fingers of the DNA-binding domain. In mitosis, Gata2 would remain bound to its targets in condensed chromatin the same way to allow the rapid transcriptional activation of gene as cells re-enter interphase.

The hypothesis that Gata2 mitotic bookmarking is important for HSC self-renewal has to be tested *in vivo* using transgenic mice with fluorescent fusion Gata2 protein or mice where Gata2 is replaced by Gata2 fused to a cyclin mitotic degron (MD) (Holloway et al. 1993; Glotzer et al. 1991) to prompt mitosis-specific destruction of Gata2. Chip-seq can be done for Gata2 in mitotic reprogrammed cells to define the genes that are bookmarked. To confirm specific protein domains that interact with mitotic chromatin, Gata2 deletion constructs fused to mCherry could be generated. Furthermore, since there is indication that mitotic bookmarking of genes is associated with TF's ability to induce cell reprogramming, the impact of Gata2 during mitosis should be studied. If its possible that Gata2 binds to chromatin during mitosis to transmit the transcriptional program of HSCs to the progeny, it is also possible that the same happens during hematopoietic reprogramming for the

conveyance of the reprogramed cell state, once cells achieve a hemogenic or an immature hematopoietic cell fate. An easy approach to study this would rely on the mitotic destruction of Gata2 protein. MD-Gata2 constructs could be tested in our hCD34/H2BGFP reporter system to evaluate the percentage of GFP expressing cells. Having proven the utility of mitotic bookmarking for hematopoietic reprogramming, the domains associated with this epigenetic feature would be taken into account to further improve the proposed enhanced reprogramming factors. Moreover, fluorescent fusion proteins with Gfi1b and cFos should also be generated to investigate mitotic bookmarking by these proteins. It would be interesting to use fusion constructs with fluorescent proteins that emit in distinct wavelengths or tags with different fluorophores, to evaluate the mitotic bookmarking activity of Gfi1b and cFos, individually or together with Gata2.

Collectively, I have established relevant features of Gata2, Gfi1b and cFos for hematopoietic reprogramming that broadened our understanding regarding HSC fate acquisition and preservation. Furthermore I have proposed models of improved reprogramming factors that may increase the current hematopoietic reprogramming efficiency and allow clinical translation. Future experiments will aim to validate the current results, generate and test the proposed improved factors and prove the relevance of mitotic bookmarking by Gata2 for HSC fate perpetuation and for hematopoietic reprogramming.

REFERENCES

Allis, C.D. & Jenuwein, T., 2016. The molecular hallmarks of epigenetic control. *Nature Reviews Genetics*, 17(8), pp.487–500.

Angel, P. & Karin, M., 1991. The role of Jun, Fos and the AP-1 complex in cell-proliferation and transformation. *Biochimica et biophysica acta*, 1072(2–3), pp.129–157.

Anon, Double Digest Calculator—Thermo Scientific. Available at: https://www.thermofisher.com/pt/en/home/brands/thermo-scientific/molecular-biology/thermo-scientific-restriction-modifying-enzymes/restriction-enzymes-thermo-scientific/double-digest-calculator-thermo-scientific.html.

Apperley, J. et al., 2016. Haploidentical Hematopoietic Stem Cell Transplantation: A Global Overview Comparing Asia, the European Union, and the United States. *Biology of Blood and Marrow Transplantation*, 22(1), pp.23–26.

Batta, K. et al., 2014. Direct Reprogramming of Murine Fibroblasts to Hematopoietic Progenitor Cells. *CellReports*, 9(5), pp.1871–1884.

Berardi, A.C. et al., 1995. Functional Isolation and Characterization of Human Hematopoietic Stem Cells. *Science*, 267(5194), pp.104–8.

Bertrand, J.Y. et al., 2010. Hematopoietic stem cells derive directly from aortic endothelium during development. *Nature*, 464(7285), pp.108–111.

Blau, H.M., Chiu, C.P. & Webster, C., 1983. Cytoplasmic activation of human nuclear genes in stable heterocaryons. *Cell*, 32(4), pp.1171–1180.

Boisset, J.-C. et al., 2010. In vivo imaging of haematopoietic cells emerging from the mouse aortic endothelium. *Nature*, 464(7285), pp.116–120.

Briggs, R. & King, T.J., 1952. Transplantation of Living Nuclei from Blastula Cells into Enucleated Frogs' Eggs. *PNAS*, 38(5), pp.455–463.

Bussmann, L.H. et al., 2009. A Robust and Highly Efficient Immune Cell Reprogramming System. *Cell Stem Cell*, 5(5), pp.554–566.

Campos, M.A.S. et al., 1999. Protein Domains Involved in Nuclear Transport of Fos. *Cell Biology International*, 23(1), pp.81–88.

Capellera-Garcia, S. et al., 2016. Defining the Minimal Factors Required for Erythropoiesis through Direct Lineage Conversion. *Cell Reports*, 15(11), pp.2550–2562.

Capellera-Garcia, S. & Flygare, J., 2017. Direct lineage reprogramming: a useful addition to the blood cell research toolbox. *Expert Review of Hematology*, 10(2), pp.107–109.

Caravaca, J.M. et al., 2013. Bookmarking by specific and nonspecific binding of FoxA1 pioneer factor to mitotic chromosomes. *Genes and Development*, 27(3), pp.251–260.

Cheng, H. et al., 2016. Reprogramming mouse fibroblasts into engraftable

myeloerythroid and lymphoid progenitors. *Nature Communications*, 7(13396), pp.1–15.

Choi, K. et al., 2012. Identification of the Hemogenic Endothelial Progenitor and Its Direct Precursor in Human Pluripotent Stem Cell Differentiation Cultures. *CellReports*, 2(3), pp.553–567.

Christova, R. & Oelgeschläger, T., 2002. Association of human TFIID-promoter complexes with silenced mitotic chromatin in vivo. *Nature cell biology*, 4(1), pp.79–82.

Chung, C.T. & Miller, R.H., 1993. Preparation and Storage of Competent Escherichia coli Cells. *Methods in enzymology*, 218(1979), pp.621–627.

Chung, C.T., Niemela, S.L. & Miller, R.H., 1989. One-Step Preparation of Competent Escherichia coli: Transformation and Storage of Bacterial Cells in the Same Solution. *PNAS*, 86(7), pp.2172–2175.

Cohen, D.R. et al., 1989. The product of a fos-related gene, fra-1, binds cooperatively to the AP-1 site with Jun: transcription factor AP-1 is comprised of multiple protein complexes. *Genes & Development*, 3(2), pp.173–184.

Costa, G., Kouskoff, V. & Lacaud, G., 2012. Origin of blood cells and HSC production in the embryo. *Trends in Immunology*, 33(5), pp.215–223.

Cowan, C.A. et al., 2005. Nuclear Reprogramming of Somatic Cells after Fusion with Human Embryonic Stem Cells. *Science*, 309(5739), pp.1369–1373.

Daniel, M.G. et al., 2016. Making a Hematopoietic Stem Cell. *Trends in Cell Biology*, 26(3), pp.202–14.

Davis, H.E., Morgan, J.R. & Yarmush, M.L., 2002. Polybrene increases retrovirus gene transfer efficiency by enhancing receptor-independent virus adsorption on target cell membranes. *Biophysical Chemistry*, 97(2–3), pp.159–172.

Deluz, C. et al., 2016. A role for mitotic bookmarking of SOX2 in pluripotency and differentiation. *Genes and Development*, 30(22), pp.2538–2550.

Dudek, S.M. et al., 2009. Brd4 Marks Select Genes on Mitotic Chromatin and Directs Postmitotic Transcription. *Molecular biology of the cell*, 20(23), pp.4899–4909.

Egli, D., Birkhoff, G. & Eggan, K., 2008. Mediators of reprogramming: transcription factors and transitions through mitosis. *Nature reviews. Molecular cell biology*, 9(7), pp.505–516.

Ema, B.H. et al., 2000. In Vitro Self-Renewal Division of Hematopoietic Stem Cells. , 192(9), pp.0–7.

Escors, D. & Breckpot, K., 2010. Lentiviral Vectors in Gene Therapy: Their Current Status and Future Potential. *Archivum Immunologiae et Therapiae Experimentalis*, 58(2), pp.107–119.

Evans, M.J. & Kaufman, M.H., 1981. Establishment in culture of pluripotential

cells from mouse embryos. *Nature*, 292(5819), pp.154–156.
Evans, T. & Felsenfeld, G., 1989. The Erythroid-Specific Transcription Factor Eryf1: A New Finger Protein. *Cell*, 58(5), pp.877–885.
Feng, B. et al., 2009. Reprogramming of fibroblasts into induced pluripotent stem cells with orphan nuclear receptor Esrrb. *Nature Cell Biology*, 11(2), pp.197–203.
Feng, R. et al., 2008. PU.1 and C/EBPalpha /beta convert fibroblasts into macrophage-like cells. *PNAS*, 105(16).
Festuccia, N. et al., 2016. Mitotic binding of Esrrb marks key regulatory regions of the pluripotency network. *Nature Cell Biology*, 18(11), pp.1139–1148.
Fiolka, K. et al., 2006. Gfi1 and Gfi1b act equivalently in haematopoiesis, but have distinct, non-overlapping functions in inner ear development. *EMBO reports*, 7(3), pp.326–333.
Fitch, W.M., 1970. Distinguishing Homologous from Analogous Proteins. *Systematic Zoology*, 19(2), p.99.
Forman, S.J. & Nakamura, R., 2011. Hematopoietic Cell Transplantation. *Oncology*, pp.1–15.
Foudi, A. et al., 2014. Distinct, strict requirements for Gfi-1b in adult bone marrow red cell and platelet generation. *The Journal of experimental medicine*, 211(5), pp.909–927.
Fujiwara, Y. et al., 2004. Functional overlap of GATA-1 and GATA-2 in primitive hematopoietic development. *Blood*, 103(2), pp.583–586.
Gentz, R. et al., 1989. Parallel Association of Fos and Jun Leucine Zippers Juxtaposes DNA Binding Domains. *Science*, 243(4899), pp.1695–1699.
Glotzer, M., Murray, A.W. & Kirschner, M.W., 1991. Cyclin is degraded by the ubiquitin pathway. *Nature*, 349(6305), pp.132–138.
Goldberg, A.D., Allis, C.D. & Bernstein, E., 2007. Epigenetics: A Landscape Takes Shape. *Cell*, 128(4), pp.635–638.
Gottesfeld, J.M. & Forbes, D.J., 1997. Mitotic repression of the transcriptional machinery. *Trends in Biochemical Sciences*, 22(6), pp.197–202.
Grimes, H.L. et al., 1996. The Gfi-1 proto-oncoprotein contains a novel transcriptional repressor domain, SNAG, and inhibits G1 arrest induced by interleukin-2 withdrawal. *Molecular and cellular biology*, 16(11), pp.6263–6272.
Guo, Y. et al., 2016. GATA2 regulates GATA1 expression through LSD1-mediated histone modification. *American Journal of Translational Research*, 8(5), pp.2265–2274.
Guo, Y. et al., 2015. Histone demethylase LSD1-mediated repression of GATA-2 is critical for erythroid differentiation. *Drug Design, Development and Therapy*, 9, pp.3153–3162.
Gurdon, J.B., 1962a. Adult frogs derived from the nuclei of single somatic cells. *Developmental Biology*, 4(2), pp.256–273.

Gurdon, J.B., 1962b. The developmental capacity of nuclei taken from intestinal epithelium cells of feeding tadpoles. *Journal of embryology and experimental morphology*, 10(4), pp.622–640.

Halazonetis, T. et al., 1988. c-Jun dimerizes with itself and with c-Fos, forming complexes of different DNA binding affinities. *Cell*, 55(5), pp.917–924.

Hanna, J., Carey, B.W. & Jaenisch, R., 2008. Reprogramming of somatic cell identity. *Cold Spring Harb Symp Quant Biol*, 73, pp.147–155.

Harper, J. V, 2005. Synchronization of Cell Populations in G1/S and G2/M Phases of the Cell Cycle. *Methods in Molecular Biology*, 296(1), pp.157–166.

Hermiston, M.L., Xu, Z. & Weiss, A., 2003. CD45: A Critical Regulator of Signaling Thresholds in Immune Cells. *Annu. Rev. Immunol*, 21, pp.107–137.

Ho, I.-C. et al., 1991. Human GATA-3: a lineage-restricted transcription factor that regulates the expression of the T cell receptor alpha gene. *Embo J*, 10(5), pp.1187–1192.

Hock, H. et al., 2003. Intrinsic requirement for zinc finger transcription factor Gfi-1 in neutrophil differentiation. *Immunity*, 18(1), pp.109–120.

Holloway, S.L. et al., 1993. Anaphase is Initiated by Proteolysis Rather than by the Inactivation of Maturation-Promoting Factor. *Cell*, 73(7), pp.1393–1402.

Huang, P. et al., 2011. Induction of functional hepatocyte-like cells from mouse fibroblasts by defined factors. *Nature*, 475(7356), pp.386–9.

Ieda, M. et al., 2010. Direct Reprogramming of Fibroblasts into Functional Cardiomyocytes by Defined Factors. *Cell*, 142(3), pp.375–386.

Inagawa, K. et al., 2012. Induction of cardiomyocyte-like cells in infarct hearts by gene transfer of Gata4, Mef2c, and Tbx5. *Circulation Research*, 111(9), pp.1147–1156.

Jansen, J. et al., 2005. Transplantation of hematopoietic stem cells from the peripheral blood. *Journal of cellular and molecular medicine*, 9(1), pp.37–50.

Jeppesen, P., 1997. Histone acetylation: A possible mechanism for the inheritance of cell memory at mitosis. *BioEssays*, 19(1), pp.67–74.

Joulin, V. et al., 1991. A T-cell specific TCR delta DNA binding protein is a member of the human GATA family. *EMBO Journal*, 10(7), pp.1809–1816.

Kadauke, S. et al., 2012. Tissue-specific mitotic bookmarking by hematopoietic transcription factor GATA1. *Cell*, 150(4), pp.725–737.

Kadauke, S. & Blobel, G. a, 2013. Mitotic bookmarking by transcription factors. *Epigenetics & Chromatin*, 6(1), p.6.

Kaimakis, P., Crisan, M. & Dzierzak, E., 2013. The biochemistry of hematopoietic stem cell development. *Biochimica et biophysica acta*, 1830(2), pp.2395–2403.

Kaneko, H. et al., 2012. N- and C-terminal transactivation domains of GATA1 protein coordinate hematopoietic program. *Journal of Biological Chemistry*, 287(25), pp.21439–21449.

Kerenyi, M.A. et al., 2013. Histone demethylase Lsd1 represses hematopoietic stem and progenitor cell signatures during blood cell maturation. *eLife*, 2(633).

Khandanpour, C. et al., 2011. Evidence that Growth factor independence 1b regulates dormancy and peripheral blood mobilization of hematopoietic stem cells. *Blood*, 116(24), pp.5149–5161.

Ko, L. et al., 1991. Murine and human T-lymphocyte GATA-3 factors mediate transcription through a cis-regulatory element within the human T-cell receptor delta gene enhancer. *Molecular and cellular biology*, 11(5), pp.2778–2784.

Ko, L.J. & Engel, J.D., 1993. DNA-Binding Specificities of the GATA Transcription Factor Family. *Molecular and Cellular Biology*, 13(7), pp.4011–4022.

Kosan, C. & Godmann, M., 2016. Genetic and Epigenetic Mechanisms That Maintain Hematopoietic Stem Cell Function. *Stem Cells International*, 2016.

Kouskouti, A. & Talianidis, I., 2005. Histone modifications defining active genes persist after transcriptional and mitotic inactivation. *The EMBO Journal*, 24(2), pp.347–357.

Krishan, A., 1975. Rapid Flow Cytofluorometric Analysis of Mammalian Cell Cycle by Propidium Iodide Staining. *The journal of Cell Biology*, 66(1), pp.188–193.

Ku, C.J. et al., 2012. GATA-3 regulates hematopoietic stem cell maintenance and cell-cycle entry. *Blood*, 119(10), pp.2242–2251.

Ladewig, J., Koch, P. & Brüstle, O., 2013. Leveling Waddington: the emergence of direct programming and the loss of cell fate hierarchies. *Nature Reviews Molecular Cell Biology*, 14(4), pp.225–236.

Lake, R.J. et al., 2014. RBPJ, the Major Transcriptional Effector of Notch Signaling, Remains Associated with Chromatin throughout Mitosis, Suggesting a Role in Mitotic Bookmarking. *PLoS Genetics*, 10(3).

Lan, F., Nottke, A.C.A.A.C. & Shi, Y., 2008. Mechanisms involved in the regulation of histone lysine demethylases. *Current opinion in cell biology*, 20(3), pp.316–325.

Lancrin, C. et al., 2012. GFI1 and GFI1B control the loss of endothelial identity of hemogenic endothelium during hematopoietic commitment. *Blood*, 120(2), pp.314–322.

Lee, S. et al., 2010. Solution structure of Gfi-1 zinc domain bound to consensus DNA. *Journal of Molecular Biology*, 397(4), pp.1055–1066.

Ling, K. et al., 2004. GATA-2 Plays Two Functionally Distinct Roles during the

Ontogeny of Hematopoietic Stem Cells The Journal of Experimental Medicine. *The Journal of Experimental Medicine*, 200(7).
Lis, R. et al., 2017. Conversion of adult endothelium to immunocompetent haematopoietic stem cells. *Nature*, 545(7655), pp.439–445.
Liu, X. et al., 2012. Direct reprogramming of human fibroblasts into dopaminergic neuron-like cells. *Cell Research*, 22(2), pp.321–332.
Liu, Y. et al., 2017. Widespread Mitotic Bookmarking by Histone Marks and Transcription Factors in Pluripotent Stem Cells. *Cell Reports*, 19(7), pp.1283–1293.
Lodhi, N., Ji, Y. & Tulin, A., 2016. Mitotic bookmarking: maintaining post-mitotic reprogramming of transcription reactivation. *Current molecular biology reports*, 2(1), pp.10–16.
Lodhi, N., Kossenkov, A. V & Tulin, A. V, 2014. Bookmarking promoters in mitotic chromatin: Poly(ADP-ribose)polymerase-1 as an epigenetic mark. *Nucleic Acids Research*, 42(11), pp.7028–7038.
Margueron, R. & Reinberg, D., 2010. Chromatin structure and the inheritance of epigenetic information. *Nature Reviews Genetics*, 11(4), pp.285–296.
Martin, D. et al., 1990. Expression of an erythroid transcription factor in megakaryocytic and mast cell lineages. *Nature*, 344(6265), pp.444–447.
Martin, G.R., 1981. Isolation of a pluripotent cell line from early mouse embryos cultured in medium conditioned by teratocarcinoma stem cells. *PNAS*, 78(12), pp.7634–7638.
Martínez-Balbás, M.A. et al., 1995. Displacement of sequence-specific transcription factors from mitotic chromatin. *Cell*, 83(1), pp.29–38.
Matsui, Y. et al., 2012. Enrichment of cell populations in metaphase, anaphase, and telophase by synchronization using nocodazole and blebbistatin: A novel method suitable for examining dynamic changes in proteins during mitotic progression. *European Journal of Cell Biology*, 91(5), pp.413–419.
McBride, K. et al., 2003. Interaction with GATA transcription factors provides a mechanism for cell-specific effects of c-Fos. *Oncogene*, 22(52), pp.8403–8412.
Medvinsky, A., Rybtsov, S. & Taoudi, S., 2011. Embryonic origin of the adult hematopoietic system: advances and questions. *Development*, 138(6), pp.1017–1031.
Metcalf, D., 1989. The molecular control of cell division, differentiation commitment and maturation in haemopoietic cells. *Nature*, 339(6219), pp.27–30.
Michelotti, E.F., Sanford, S. & Levens, D., 1997. Marking of active genes on mitotic chromosomes. *Nature*, 388(6645), pp.895–899.
Milde-Langosch, K., 2005. The Fos family of transcription factors and their role in tumourigenesis. *European Journal of Cancer*, 41(16), pp.2449–2461.
Minegishi, N. et al., 2003. Expression and domain-specific function of GATA-2

during differentiation of the hematopoietic precursor cells in midgestation mouse embryos. *Blood*, 102(3), pp.896–905.

Minegishi, N. et al., 2005. Rapid turnover of GATA-2 via ubiquitin-proteasome protein degradation pathway. *Genes to Cells*, 10(7), pp.693–704.

Minegishi, N. et al., 1999. The mouse GATA-2 gene is expressed in the para-aortic splanchnopleura and aorta-gonads and mesonephros region. *Blood*, 93(12), pp.4196–4207.

Montserrat, N. et al., 2013. Reprogramming of Human Fibroblasts to Pluripotency with Lineage Specifiers. *Cell Stem Cell*, 13(3), pp.341–350.

Möröy, T. et al., 2015. From cytopenia to leukemia: The role of Gfi1 and Gfi1b in blood formation. *Blood*, 126(24), pp.2561–2569.

Nakabeppu, Y., Ryder, K. & Nathans, D., 1988. DNA binding activities of three murine Jun proteins: Stimulation by Fos. *Cell*, 55(5), pp.907–915.

Nayerossadat, N., Ali, P. & Maedeh, T., 2012. Viral and nonviral delivery systems for gene delivery. *Advanced Biomedical Research*, 1(1), p.27.

Nishina, H. et al., 1990. Isolation and characterization of fra-2, an additional member of the fos gene family. *PNAS*, 87(9), pp.3619–3623.

Nowak, S.J. & Corces, V.G., 2004. Phosphorylation of histone H3: A balancing act between chromosome condensation and transcriptional activation. *Trends in Genetics*, 20(4), pp.214–220.

Ohneda, K. & Yamamoto, M., 2002. Roles of Hematopoietic Transcription Factors GATA-1 and GATA-2 in the Development of Red Blood Cell Lineage. *Acta Haematologica*, 108(4), pp.237–245.

Orkin, S.H., 1992. GATA-Binding Transcription Factors in Hematopoietic Cells. *Blood*, 80(3), pp.575–581.

Ottersbach, K. & Dzierzak, E., 2005. The murine placenta contains hematopoietic stem cells within the vascular labyrinth region. *Developmental Cell*, 8(3), pp.377–387.

Pan, X. et al., 2000. Identification of Human GATA-2 Gene Distal IS Exon and Its Expression in Hematopoietic Stem Cell Fractions. *J. Biochem.*, 127(1), pp.547–552.

Passweg, J. et al., 2012. Hematopoietic stem cell transplantation: A review and recommendations for follow-up care for the general practitioner. *Swiss Medical Weekly*, 142(13696), pp.1–15.

Passweg, R. et al., 2015. Hematopoietic SCT in Europe 2013: recent trends in the use of alternative donors showing more haploidentical donors but fewer cord blood transplants. *Bone marrow transplantation*, 50(4), pp.476–82.

Pereira, C.F. et al., 2016. Hematopoietic Reprogramming In Vitro Informs In Vivo Identification of Hemogenic Precursors to Definitive Hematopoietic Stem Cells. *Developmental Cell*, 36(5), pp.525–539.

Pereira, C.F. et al., 2008. Heterokaryon-based reprogramming of human B

lymphocytes for pluripotency requires Oct4 but not Sox2. *PLoS Genetics*, 4(9).

Pereira, C.F. et al., 2013. Induction of a Hemogenic Program in Mouse Fibroblasts. *Cell Stem Cell*, 13(2), pp.205–218.

Pereira, C.F. & Moore, K., 2014. "There will be blood" from fibroblasts. *Cell Cycle*, 13(1), pp.30–31.

Prescott, D.M. & Bender, M.A., 1962. Synthesis of RNA and protein during mitosis in mammalian tissue culture cells. *Experimental Cell Research*, 26(2), pp.260–268.

Protze, S. et al., 2012. A new approach to transcription factor screening for reprogramming of fibroblasts to cardiomyocyte-like cells. *Journal of Molecular and Cellular Cardiology*, 53(3), pp.323–332.

Pulecio, J. et al., 2016. Direct Conversion of Fibroblasts to Megakaryocyte Progenitors. *Cell Reports*, 17(3), pp.671–683.

Rafii, S. et al., 2013. Human ESC-derived hemogenic endothelial cells undergo distinct waves of endothelial to hematopoietic transition. *Blood*, 121(5), pp.770–781.

Riddell, J. et al., 2014. Reprogramming committed murine blood cells to induced hematopoietic stem cells with defined factors. *Cell*, 157(3), pp.549–564.

Ring, K.L. et al., 2012. Direct Reprogramming of Mouse and Human Fibroblasts into Multipotent Neural Stem Cells with a Single Factor. *Cell Stem Cell*, 11(1), pp.100–109.

Rodrigues, N., Tipping, A. & Enver, T., 2012. GATA-2 mediated regulation of normal hematopoietic stem/progenitor cell function, myelodysplasia and myeloid leukemia. *International Journal of Biochemistry and Cell Biology*, 44(3), pp.457–460.

Romeo, P.-H. et al., 1990. Megakaryocytic and erythrocytic lineages share specific transcription factors. *Nature*, 344(6265), pp.447–449.

Rossi, L. et al., 2011. Hematopoietic Stem Cell Characterization and Isolation. *Methods in Molecular Biology*, 750(4), pp.117–129.

Sadahira, K. et al., 2012. Direct reprogramming of terminally differentiated B cells into erythroid lineage. *FEBS Letters*, 586(20), pp.3645–3652.

Saleque, S. et al., 2007. Epigenetic Regulation of Hematopoietic Differentiation by Gfi-1 and Gfi-1b Is Mediated by the Cofactors CoREST and LSD1. *Molecular Cell*, 27(4), pp.562–572.

Saleque, S., Cameron, S. & Orkin, S.H., 2002. The zinc-finger proto-oncogene Gfi-1b is essential for development of the erythroid and megakaryocytic lineages. *Genes & Development*, 16(3), pp.301–306.

Sandler, V. et al., 2014. Reprogramming human endothelial cells to haematopoietic cells requires vascular induction. *Nature*, 511(7509), pp.312–318.

Sarge, K. & Park-Sarge, O.-K., 2009. Mitotic bookmarking of formerly active genes: Keeping epigenetic memories from fading. *Cell Cycle*, 8(6), pp.818–823.

Sassone-Corsi, P. et al., 1988. Direct interaction between fos and jun nuclear oncoproteins: role of the “leucine zipper” domain. *Nature*, 336(6200), pp.403–405.

Schaniel, C. & Moore, K.A., 2009. Genetic Models to Study Quiescent Stem Cells and Their Niches. *Hematopoietic Stem Cell VII*, 1176, pp.26–35.

Schmidt, T. et al., 1998. Zinc finger protein GFI-1 has low oncogenic potential but cooperates strongly with pim and myc genes in T-cell lymphomagenesis. *Oncogene*, 17(20), pp.2661–2667.

Schorl, C. & Sedivy, J.M., 2007. Analysis of cell cycle phases and progression in cultured mammalian cells. *Methods*, 41(2), pp.143–150.

Schuster, J.A. et al., 2012. Expansion of hematopoietic stem cells for transplantation: current perspectives. *Experimental hematology & oncology*, 1(1), p.12.

Sekiya, S. & Suzuki, A., 2011. Direct conversion ofmouse fibroblasts to hepatocyte-like cells by defined factors. *Nature*, 475(7356), pp.390–393.

Shen, W. et al., 2015. A possible role of Drosophila CTCF in mitotic bookmarking and maintaining chromatin domains during the cell cycle. *Biological Research*, 48(1), pp.1–8.

Shimizu, R. et al., 2001. In vivo requirements for GATA-1 functional domains during primitive and definitive erythropoiesis. *EMBO Journal*, 20(18), pp.5250–5260.

Shu, J. et al., 2013. Induction of Pluripotency in Mouse Somatic Cells with Lineage Specifiers. *Cell*, 153(5), pp.963–975.

Siena, S. et al., 2000. Therapeutic relevance of CD34+ cell dose in blood cell transplantation for cancer therapy. *Journal of Clinical Oncology*, 18(6), pp.1360–1377.

Simon, M.C. et al., 1992. Rescue of erythroid development in gene targeted GATA-1- mouse embryonic stem cells. *Nature genetics*, 1(2), pp.92–98.

Smith, A.R. & Wagner, J.E., 2009. Alternative Hematopoietic Stem Cell Sources for Transplantation: Place of Umbilical Cord Blood. *Br J Haematol.*, 147(2), pp.246–261.

Snapp, E., 2005. Design and Use of Fluorescent Fusion Proteins in Cell Biology. In *Current Protocols in Cell Biology*. p. 21.4.1-21.4.17.

Sommer, C.A. et al., 2009. Induced Pluripotent Stem Cell Generation Using a Single Lentiviral Stem Cell Cassette. *Stem cells*, 27(3), pp.543–9.

Son, E.Y. et al., 2011. Conversion of mouse and human fibroblasts into functional spinal motor neurons. *Cell Stem Cell*, 9(3), pp.205–218.

Soufi, A., 2014. Mechanisms for enhancing cellular reprogramming. *Current Opinion in Genetics and Development*, 25(1), pp.101–109.

Stella, C.C. et al., 1995. CD34-Positive Cells: Biology and Clinical Relevance. *Haematologica*, 80(4), pp.367–387.

Sturgeon, C.M., Ditadi, A. & Clarke, R.L., 2013. Defining the path to hematopoietic stem cells. *Nature Biotechnology*, 31(5), pp.416–418.

Sugimura, R. et al., 2017. Haematopoietic stem and progenitor cells from human pluripotent stem cells. *Nature*, 545(7655), pp.432–438.

Suzuki, M., Shimizu, R. & Yamamoto, M., 2011. Transcriptional regulation by GATA1 and GATA2 during erythropoiesis. *International Journal of Hematology*, 93(2), pp.150–155.

Suzuki, N. et al., 2006. Combinatorial Gata2 and Sca1 expression defines hematopoietic stem cells in the bone marrow niche. *Proceedings of the National Academy of Sciences of the United States of America*, 103(7), pp.2202–2207.

Szydlo, R. et al., 1997. Results of allogeneic bone marrow transplants for leukemia using donors other than HLA-identical siblings. *Journal of Clinical Oncology*, 15(5), pp.1767–1777.

Tada, M. et al., 2001. Nuclear reprogramming of somatic cells by in vitro hybridization with ES cells. *Current Biology*, 11(19), pp.1553–1558.

Takahashi, B.S. et al., 1998. Role of GATA-1 in Proliferation and Differentiation of Definitive Erythroid and Megakaryocytic Cells In Vivo. *Blood*, 92(2), pp.434–442.

Takahashi, K. et al., 2007. Induction of Pluripotent Stem Cells from Adult Human Fibroblasts by Defined Factors. *Cell*, 131(5), pp.861–872.

Takahashi, K. & Yamanaka, S., 2006. Induction of Pluripotent Stem Cells from Mouse Embryonic and Adult Fibroblast Cultures by Defined Factors. *Cell*, 126(4), pp.663–676.

Takizawa, H., Schanz, U. & Manz, M.G., 2011. Ex vivo expansion of hematopoietic stem cells: Mission accomplished? *Swiss Medical Weekly*, 141.

Taylor, J.H., 1960. Nucleic Acid Synthesis in Relation To the Cell Division Cycle. *Annals of the New York Academy of Sciences*, 90(2), pp.409–421.

Terasaki, M. et al., 2001. A new model for nuclear envelope breakdown. *Mol Biol Cell*, 12(2), pp.503–510.

Teves, S.S. et al., 2016. A dynamic mode of mitotic bookmarking by transcription factors. *eLife*, 5(22280), pp.1–24.

Thambyrajah, R. et al., 2015. GFI1 proteins orchestrate the emergence of haematopoietic stem cells through recruitment of LSD1. *Nature cell biology*, 18(1), pp.21–32.

Ting, C.-N. et al., 1996. Transcription factor GATA-3 is required for development of the T-cell lineage. *Nature*, 384(6608), pp.474–478.

Tong, B. et al., 1998. The Gfi-1B proto-oncoprotein represses p21WAF1 and inhibits myeloid cell differentiation. *Molecular and cellular biology*, 18(5),

pp.2462–2473.
Trainor, C.D., Ghirlando, R. & Simpson, M.A., 2000. GATA Zinc Finger Interactions Modulate DNA Binding and Transactivation. *Journal of Biological Chemistry*, 275(36), pp.28157–28166.
Trojer, P. & Reinberg, D., 2006. Histone Lysine Demethylases and Their Impact on Epigenetics. *Cell*, 125(2), pp.213–217.
Tsai, F.-Y. et al., 1994. An early haematopoietic defect in mice lacking the transcription factor GATA-2. *Nature*, 371(6494), pp.221–6.
Tsai, F.-Y. & 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(10), pp.3636–3643.
Turner, R. & Tjian, R., 1989. Leucine Repeats and an Adjacent DNA Binding Domain Mediate the Formation of Functional cFos-cJun Heterodimers. *Science*, 243(4899), pp.1689–1694.
Valis, E., Sánchez-Molina, S. & Martínez-Balbás, M.A., 2005. Role of histone modifications in marking and activating genes through mitosis. *Journal of Biological Chemistry*, 280(52), pp.42592–42600.
Verdeguer, F. et al., 2010. A mitotic transcriptional switch in polycystic kidney disease. *Nature medicine*, 16(1), pp.106–110.
Vicente, C. et al., 2012. The role of the GATA2 transcription factor in normal and malignant hematopoiesis. *Critical Reviews in Oncology/Hematology*, 82(1), pp.1–17.
Vierbuchen, T. et al., 2010. Direct conversion of fibroblasts to functional neurons by defined factors. *Nature*, 463(7284), pp.1035–41.
Visvader, J. & Adams, J.M., 1993. Megakaryocytic differentiation induced in 416B myeloid cells by GATA-2 and GATA-3 transgenes or 5-azacytidine is tightly coupled to GATA-1 expression. *Blood*, 82(5), pp.1493–1501.
Visvader, J.E. et al., 1995. The C-terminal zinc finger of GATA-1 or GATA-2 is sufficient to induce megakaryocytic differentiation of an early myeloid cell line. *Mol Cell Biol*, 15(2), pp.634–641.
Vo, L. & Daley, G., 2015. De novo generation of HSCs from somatic and pluripotent stem cell sources. *Blood*, 125(17), pp.2641–2648.
Wagner, E.F., 2005. Fos/AP-1 proteins in bone and the immune system. *Immunological Reviews*, 208, pp.126–140.
Wakayama, T. et al., 1998. Full-term development of mice from enucleated oocytes injected with cumulus cell nuclei. *Nature*, 394(6691), pp.369–374.
Whyatt, D.J., DeBoer, E. & Grosveld, F., 1993. The two zinc finger-like domains of GATA-1 have different DNA binding specificities. *EMBO Journal*, 12(13), pp.4993–5005.
Wigler, M., 1981. The somatic replication of DNA methylation. *Cell*, 24(1), pp.33–40.

Wilmut, I. et al., 1997. Viable offspring derived from fetal and adult mammalian cells. *Nature*, 385(6619), pp.810–813.

Wisdon, R. & Verma, I., 1993. Transformation by Fos proteins requires a C-terminal transactivation domain. *Molecular and cellular biology*, 13(12), pp.7429–7438.

Wu, X. et al., 2007. GATA-1 , -2 and -3 genes expression in bone marrow microenviroment with chronic aplastic anemia. *Hematology*, 12(4), pp.331–335.

Xin, L. et al., 2007. Exploring cellular memory molecules marking competent and active transcriptions. *BMC Molecular Biology*, 8(1), p.31.

Xing, H. et al., 2005. Mechanism of hsp70i Gene Bookmarking. *Science*, 307(5708), pp.421–423.

Xing, H., Vanderford, N.L. & Sarge, K.D., 2008. The TBP-PP2A mitotic complex bookmarks genes by preventing condensin action. *Nature cell biology*, 10(11), pp.1318–1323.

Xu, J., Du, Y. & Deng, H., 2015. Direct Lineage Reprogramming: Strategies, Mechanisms, and Applications. *Cell Stem Cell*, 16(2), pp.119–134.

Yan, J. et al., 2006. The Forkhead Transcription Factor FoxI1 Remains Bound to Condensed Mitotic Chromosomes and Stably Remodels Chromatin Structure. , 26(1), pp.155–168.

Yang, H.Y. & Evans, T., 1992. Distinct roles for the two cGATA-1 finger domains. *Molecular and cellular biology*, 12(10), pp.4562–4570.

Yang, Z. et al., 1994. Human GATA-3 trans-activation, DNA-binding, and nuclear localization activities are organized into distinct structural domains. *Molecular and cellular biology*, 14(3), pp.2201–2212.

Yang, Z., He, N. & Zhou, Q., 2008. Brd4 Recruits P-TEFb to Chromosomes at Late Mitosis To Promote G1 Gene Expression and Cell Cycle Progression. *Molecular and Cellular Biology*, 28(3), pp.967–976.

Yi, Y., Noh, M.J. & Lee, K.H., 2011. Current Advances in Retroviral Gene Therapy. *Current gene therapy*, 11(3), pp.218–28.

Young, D.W. et al., 2007. Mitotic retention of gene expression patterns by the cell fate-determining transcription factor Runx2. *PNAS*, 104(9), pp.3189–3194.

Yu, B. et al., 2013. Reprogramming Fibroblasts into Bipotential Hepatic Stem Cells by Defined Factors. *Cell Stem Cell*, 13(3), pp.1–13.

Yücel, R. et al., 2004. Gfi1:Green Fluorescent Protein Knock-in Mutant Reveals Differential Expression and Autoregulation of the Growth Factor Independence 1 (Gfi1) Gene During Lymphocyte Development. *Journal of Biological Chemistry*, 279(39), pp.40906–40917.

Zeng, H. et al., 2004. Transcription factor Gfi1 regulates self-renewal and engraftment of hematopoietic stem cells. *The EMBO Journal*, 23(20), pp.4116–4125.

Zerial, M. et al., 1989. The product of a novel growth factor activated gene, fos B, interacts with JUN proteins enhancing their DNA binding activity. *EMBO Journal*, 8(3), pp.805–813.

Zhao-Qi, W. et al., 1992. Bone and haematopoietic defects in mice lacking c-fos. *Nature*, 31(360), p.741–5.

Zhong, J.F. et al., 2005. Gene expression profile of murine long-term reconstituting vs. short-term reconstituting hematopoietic stem cells. *PNAS*, 102(7), pp.2448–53.

Zieve, G.W. et al., 1980. Production of Large Numbers of Mitotic Mammalian Cells by Use of the Reversible Microtubule Inhibitor Nocodazole. *Experimental Cell Research*, 126(2), pp.397–405.

Zovein, A.C. et al., 2008. Fate Tracing Reveals the Endothelial Origin of Hematopoietic Stem Cells. *Cell Stem Cell*, 3(6), pp.625–636.

Zweidler-Mckay, P. et al., 1996. Gfi-1 encodes a nuclear zinc finger protein that binds DNA and functions as a transcriptional repressor. *Molecular and cellular biology*, 16(8), pp.4024–4034.