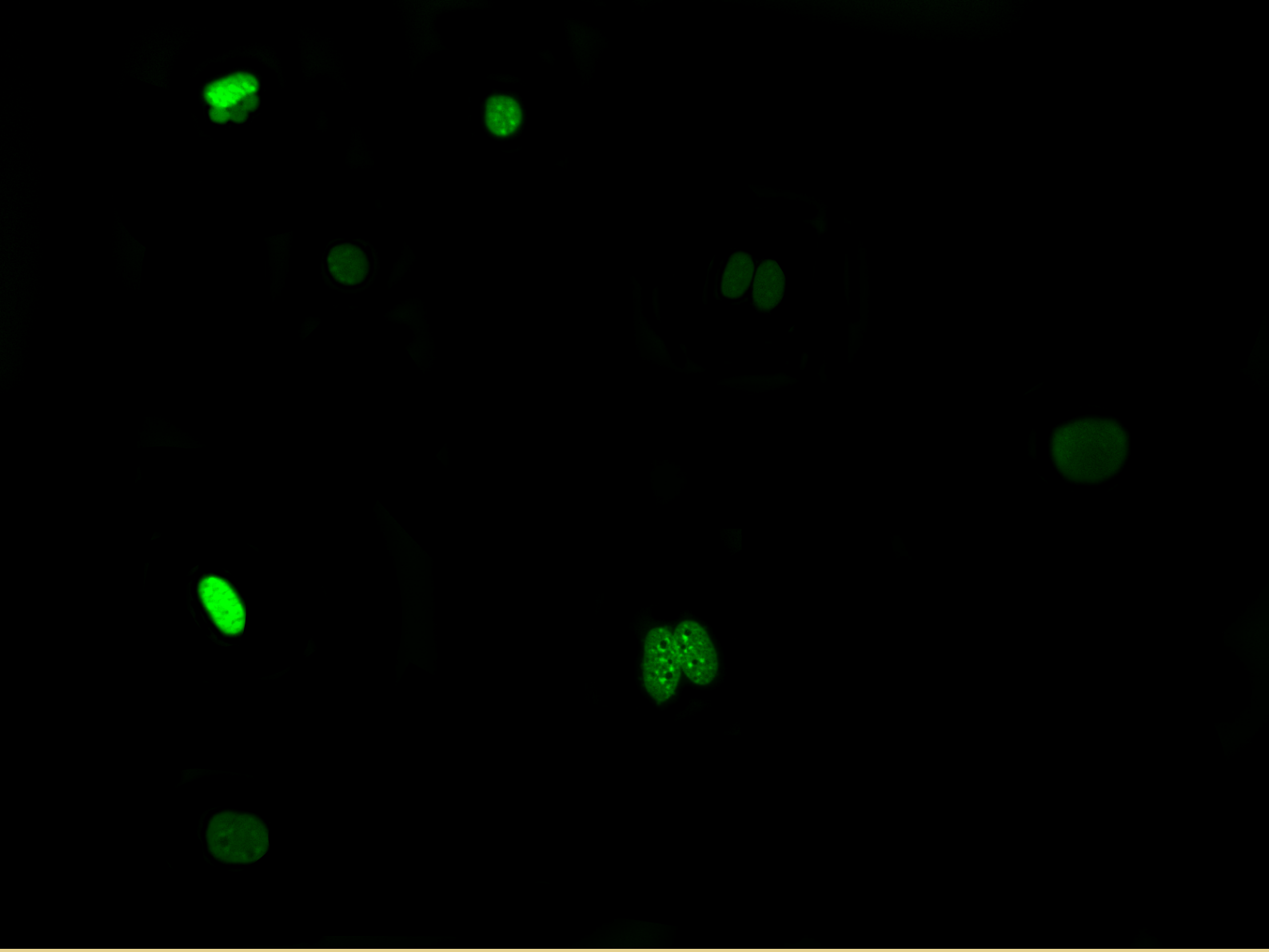


Ana Raquel Pereira Santos

THE ROLE OF LSD1/CoREST COMPLEX DURING HEMOGENIC REPROGRAMMING

Dissertação de Mestrado em Investigação Biomédica
orientada pelo Doutor Carlos Filipe Pereira e pelo Doutor Henrique Girão
e apresentada à Faculdade de Medicina da Universidade de Coimbra

Julho 2018

· U C ·

UNIVERSIDADE DE COIMBRA

The role of LSD1/CoREST complex during hemogenic reprogramming

By

Ana Raquel Pereira Santos

Thesis submitted to Faculdade de Medicina da Universidade de Coimbra for the degree of Master's in Biomedical Research

Supervisor: Dr. Henrique Girão

Co-Supervisor: Dr. Carlos Filipe Pereira

Center for Neurosciences and Cell Biology (CNC)
Universidade de Coimbra

July 2018

The most exciting phrase to hear in science, the one that heralds new discoveries, is not "Eureka!" but "That's funny"

Issac Asimov

Abstract

Hematopoietic stem cells (HSCs) are able of self-renewal and differentiation into all blood cell lineages. Due to this ability, hematopoietic stem cell transplantation (HSCT) constitutes treatment for a diversity of hematological disorders. Incompatibility between donor and host and the insufficient number of HSCs obtained for transplantation have limited the success of this cellular therapy. To overcome these limitations, expansion of HSCs *in vitro* has been explored, but this process is a changeling process as HSCs quickly lose stem cell properties upon expansion. Direct reprogramming mediated by transcription factors (TFs) of somatic cells is opening new routes for regenerative medicine and personalized HSCT. Albeit at low efficiency, combined expression of Gata2, Gfi1b and cFos induces reprogramming of fibroblasts into HSC-like cells providing a novel alternative to generate patient-specific HSCs. A better understanding of hemogenic reprogramming and the interactions between these three TFs with each other and with other players will provide valuable information to increase the efficiency of the process and to generate transplantable HSCs that can be used in the clinic.

Here, I have defined Gata2 protein domains required to interact with Gfi1b, cFos and LSD1/CoREST1. Gata2 regulates the expression of the *Kdm1a* gene that encodes LSD1 and co-immunoprecipitations experiments revealed multiple domains required for the interaction of Gata2 with cFos, Gfi1b and LSD1/CoREST1. The nuclear localization sequence of Gata2 is essential for the interaction with Gfi1b but also with LSD1/CoREST1 in the nucleus. C-terminal zinc finger and the N-terminal zinc finger of Gata2 are important for the interaction with LSD1/CoREST1 and cFos, respectively. cFos also interacts with the C-transactivator domain of Gata2, highlighting multiple regulatory pathways involved in the establishment of this “hemogenic complex”. Given this cooperation between TFs and LSD1/CoREST1 I hypothesize that this complex may be essential for the hemogenic reprogramming. LSD1 was pharmacological inhibited during hematopoietic reprogramming into HSC-like cells. Inhibition with two structurally unrelated small molecules led to a drastic decrease of reprogramming efficiency implicating the catalytic function of LSD1/CoREST complex during the hemogenic reprogramming.

Overall, this study identified functional interactions between of Gata2, Gfi1b, cFos and LSD1/CoREST1 and the vital role of this epigenetic regulator during hematopoietic reprogramming and acquisition of the HSC fate. This study paves the way for the regulation of this hemogenic complex bringing high-efficiency hemogenic reprogramming one step closer to clinical translation.

Keywords: Hematopoietic Stem Cell; Direct Reprogramming; Transcription Factors; Fibroblasts; LSD1/CoREST complex

Resumo

As células estaminais hematopoiéticas (HSCs) são capazes de se auto-renovar e diferenciar em todos os tipos de células sanguíneas. Estas características fazem com que a transplantação de HSCs seja o principal tratamento contra doenças hematológicas. A incompatibilidade entre dador-paciente e o número insuficiente de HSCs que são obtidas para transplantação limitam o sucesso deste tipo de terapia celular. De forma a ultrapassar estas limitações, a expansão destas células *in vitro* seria a solução, mas este processo continua limitado e as HSCs acabam por perder a capacidade de se auto-renovarem. A reprogramação directa de células somáticas, mediada por factores de transcrição, abriu novas portas na área da medicina regenerativa de forma a HSCT mais personalizados. Apesar da baixa eficiência, a sobre-expressão dos factores de transcrição Gata2, Gfi1b e cFos permitiu a conversão de fibroblastos em HSCs estabelecendo um novo método de geração de HSCs indicadas para o paciente. Assim, uma melhor compreensão de como o complexo de reprogramação hematopoiética funciona e as interacções entre estes três factores de transcrição com outros componentes do complexo pode fornecer informação importante que permita o aumento da eficiência deste processo de forma a gerar HSCs com maior potencial de transplantação.

Neste trabalho, defini os domínios do Gata2 necessários para interagir com o Gfi1b, cFos e com a LSD1/CoREST1 e mais importante, o papel catalítico do complexo LSD1/CoREST durante a reprogramação. Gata2 regula a expressão do gene *Kdm1a*, co-immunoprecipitações identificaram os domínios necessários para a interacção do Gata2 com os outros dois factores de transcrição, Gfi1b e cFos, mas também com a LSD1/CoREST1. O domínio que contem a sequência de localização nuclear (NLS) do Gata2 parece essencial na interacção com o Gfi1b, mas também com a LSD1/CoREST1, no núcleo. Os zinc fingers localizados no C-terminal e no N-terminal do Gata2 mostraram-se importantes na interação com a LSD1/CoREST1 e com o cFos, respectivamente. cFos também interage com o domínio transactivador do Gata2 no C-terminal, realçando as múltiplas vias necessárias à formação do complexo hemogénico. Esta cooperação entre os factores de transcrição e a LSD1/CoREST1 levanta a hipótese que este complexo pode ser essencial na reprogramação hematopoiética. LSD1 foi inibida farmacologicamente durante a reprogramação de fibroblastos em HSCs. A inibição com dois

inibidores estruturalmente diferentes levou à diminuição drástica do processo de reprogramação demostrando que a LSD1 tem um importante papel catalítico durante este processo.

Em suma, este estudo identifica as interacções funcionais entre o Gata2, Gfi1b, cFos e a LSD1/CoREST1 e o papel vital deste regulador epigenético durante a reprogramação hematopoiética e aquisição de HSCs. Mais informação sobre a regulação deste complexo hematopoiético pode aumentar a eficiência do processo aproximando esta tecnologia da translação para a clínica.

Palavras-chave: Células Estaminais Hematopoiéticas; Reprogramação celular directa; Factores de Transcrição; Fibroblastos; Complexo LSD1/CoREST

Agradecimentos

Ao concluir esta etapa tão marcante, não posso deixar de dedicar este espaço e agradecer a todas as pessoas que tornaram possível a realização desta dissertação.

Em primeiro lugar quero agradecer ao meu orientador, Dr. Carlos Filipe Pereira, por me conceder a oportunidade de integrar a sua equipa de investigação, pela partilha de conhecimentos, pela confiança e liberdade que me deu ao longo deste ano e por todos os desafios propostos.

Quero também agradecer ao Dr. Henrique Girão, mais do que coordenador do mestrado, foi um conselheiro, alguém que acreditou no meu potencial desde do início desta aventura. Obrigada por ter a porta do seu gabinete sempre aberta para os seus "MIBásticos".

Quero agradecer a todos os membros do Cell Reprogramming and Developmental Lab que me acolheram e acompanharam ao longo deste ano. Obrigada Cris por me teres ajudado a dar os primeiros passos neste desafio e por me ensinares a questionar. Um obrigada ao Fábio e à Alexandra por, apesar de longe, estarem sempre disponíveis e à Inês por aturar os meus dramas durante as viagens. À Rita deixo o meu grande obrigada! Recebeste-me no laboratório e confiaste sempre em mim, obrigada por partilhares comigo o teu entusiasmo pela ciência. À Catarina e ao Luís, obrigada por trazerem uma nova luz ao laboratório.

Um obrigada especial à Andreia e à Cláudia. Andreia, foste um dos meus maiores apoios nesta recta final, obrigada por partilhares comigo a tua paixão pela ciência e por me ensinares a pensar "outsider the box", "quando for grande quero ser como tu"! Cláudia, um obrigada não chega para te agradecer tudo! Todos os momentos sem sentido, desesperos e risos partilhados têm um espaço especial no meu coração. Sem ti não seria a mesma caturreira!

Deixo também o meu agradecimento aos membros de Engenharia de Tecidos (ET), obrigada pela simpatia e por receberem sempre como se ao grupo pertencesse, fosse para discutir ciência ou tomar um café. Aos meus ETs do coração, Ricardo e Carlos, agradeço pelas brincadeiras na sala de cultura quando o desespero já é grande, por me aturarem quando já nada funciona e por me ajudarem a ver sempre o positivo quando tudo está negro. Não podia ter escolhido melhor "família na ciência"!

Um obrigada com muito carinho à Dona Alda, Dona Adelaide e Filipa pelos momentos de descontração e pelos conselhos que levo para a vida.

Obrigada Tina e Jéssica por me receberem sempre com um sorriso na "sala dos mestrados".

Aos meus amigos de uma vida Clau, Mariana, Rita, Bernardo e Toscano, agradeço por todo o apoio e risadas ao longo deste ano mesmo estando longe, fazem-me sempre sentir em casa. Obrigada à Rita e ao Diogo pelos cafés cheios de peripécias que só eles podem ser protagonistas.

À minha Salomé, um grande obrigada! Obrigada pelos cafés, finos e afins até às 2h onde partilhámos medos, inseguranças e desesperos que acabavam em risos. És das pessoas mais puras que conheço, mas também com mais histórias cómicas porque toda a gente sabe que "é difícil ser Salomé".

Obrigada à pessoa mais descontraída do mundo, Jéssica Maravilha, por me tentar ensinar a respirar duas vezes antes de desesperar.

Aos meus amores de Coimbra, Cátia e Soraia agradeço tudo! Cátia, obrigada por estares sempre do meu lado e por me guiares quando já não vejo saída. Soras, obrigada pela amizade e por me ensinares a ser uma pessoa melhor.

Ao meu Nuno por todo o amor, apoio e compreensão. Sei que não é fácil lidar com uma Raquel mas obrigada por seres resistente o suficiente para continuares a caminhar ao meu lado.

Por fim, agradeço à minha família e em especial à minha mãe. Obrigada por nunca deixares de acreditar em mim mesmo quando a vida é injusta. Sou-te eternamente agradecida por todas as oportunidades que me deste, por me dares a liberdade de seguir o meu caminho. És uma guerreira e tenho um orgulho enorme em te chamar mãe.

Table of Contents

Table of Figures

Table of Tables

Abbreviations

AGM Aorta–gonad–mesonephros

AML Acute myeloid leukemia

Amp^r Ampicilin resistance gene

AO Amino oxidase

BJ Neonatal foreskin fibroblasts

bZIP Basic leucine zipper

CDS Coding sequences

CFU Colony forming units

CLP Common lymphoid progenitor

CMP Common myeloid progenitor

CMV Cytomegalovirus

dHSC Definitive HSC

DAPI 4′,6-Diamidine-2′-phenylindole dihydrochloride

DMEM Dulbecco's modified eagle medium

DMSO Dimethyl sulfoxide

DNA Deoxyribonucleic acid

DT Double transgenic

EF1α Human elongation factor-1 alfa

EHT Endothelial-to-hematopoietic transition

ESC Embryonic stem cells

mEpiSC mEpiblast stem cells

FAD Flavin-adenine dinucleotide

FBS Fetal bovine serum

GFP Green fluorescent protein

GGF Gata2, Gfi1b and cFos cocktail

GTLM Gata1, Tal1, Lmo2 and c-Myc cocktail

GVHD Graft-versus-host disease

huCD43 Human CD34

HDF Human dermal fibroblats

HEK Human embryonic kidney

hiPSCs Human induced PSCs

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

LSD1 Lysine-specific histone demethylase 1 protein

MBC Mature blood cells

MCS Multiple cloning site

MEF Mouse embryonic fibroblast

MPPs Multipotent progenitors

NLS Nuclear localization signal

NRD Negative regulatory domain

NuRD Mi-2/nucleosome remodeling and deacetylase

OD Optical density

PBS Phosphate-buffered saline

PCR Polymerase chain reaction

PEG Polyethylene glycol

PSC Pluripotent stem cell

Puro^r Puromycin resistance gene

RT Room temperature

SCNT Somatic cell nuclear transfer

SNAG Snail/Gfi1

TAD Transactivation domain

TAL 1 T-cell acute lymphocytic leukemia 1

Tet Tetracycline

TF Transcription factor

TRE Tet-response element

TSB Transformation and storage buffer

UCB Umbilical cord blood

UV Ultraviolet

WT Wild type

ZnF Zinc finger

Symbols

- Negative

+ Positive

% Percent (percentage)

°C Degrees Celsius

CO_2 Carbon dioxide

g Gram

h Hour

H_2O Water

E Embryonic day

L Liter

m Milli (as in milliter or milligram)

M Molar (mol/dm3)

m/v Mass/volume

MgCl2 Magnesium chloride

MgSO4 Magnesium sulphate

min Minutes

n Nano (as in nanogram)

rpm Rotations per minute

s Second

v/v Volume/volume

V Volts

α Alpha

β Beta

Δ Delta

μ Micro (as in microgram or microliter)

CHAPTER 1

General Introduction

1. Introduction

1.1 Hematopoiesis and Hematopoietic Stem Cells (HSCs)

The continuous replenishment of the blood system is possible due to a rare and multipotent population of hematopoietic stem cells (HSCs), that balance self-renewal and differentiation in order to maintain homeostasis. Hematopoiesis occurs through a series of stepwise stages in which HSCs are the starting point generating all blood cell lineages (**Figure 1**) (Metcalf, 1989; Moore and Lemischka, 2006).

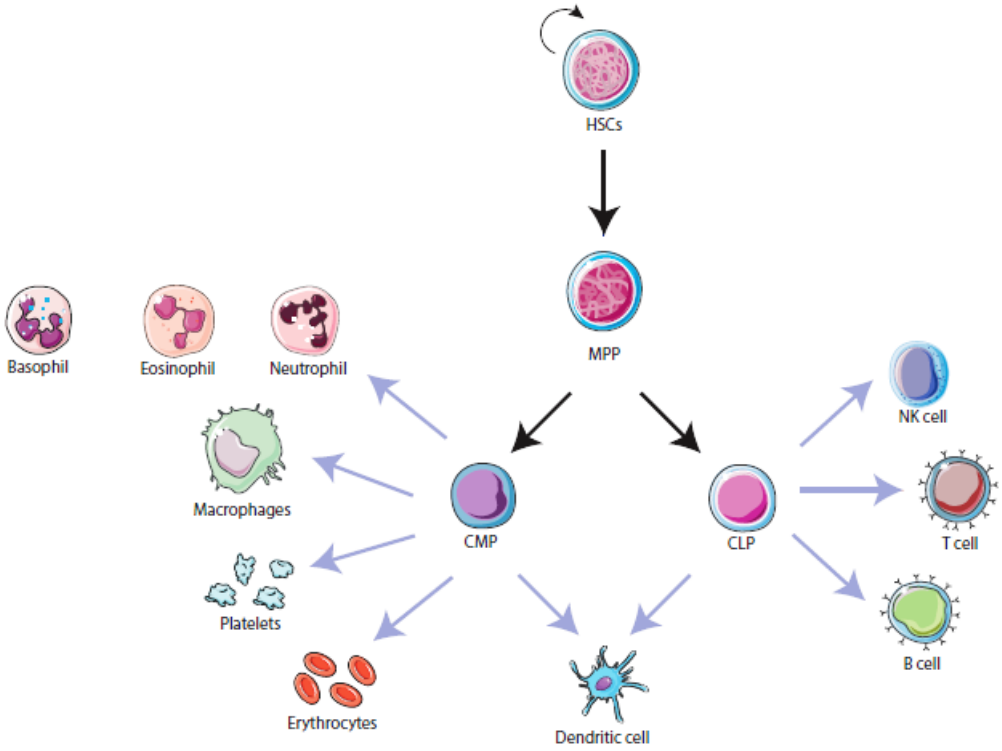


Figure 1:Hematopoiesis. The formation of all blood cell types starts with the differentiation of hematopoietic stem cells (HCS) into multipotent progenitors (MPPs). MPPs give rise to common myeloid progenitors (CMPs) and common lymphoid progenitors (CLPs) leading to the emergence of myeloid and lymphoid lineages, respectively.

During development of the mammalian embryo, blood cell specification occurs in two major waves providing the means for the embryo to rapidly obtain hematopoietic cells while generating a higher complex hematopoietic system with long-lived self-renewing HSCs (Kaimakis et al., 2013).

The generation of the first wave of blood, also called primitive hematopoiesis, begins in the yolk sac at embryonic day (E) 7.5 and primitive short-lived erythrocytes are produced to carry oxygen to the other developing tissues. The second wave, definitive hematopoiesis, starts at E 8.5 and generates *de novo* definitive hematopoietic stem cells (dHSCs) capable of self-renew and multilineage potential given raise of erythroid, lymphoid and myeloid blood cell lineages. In mice, the first dHSCs appear in the aorta-gonad-mesonephros (AGM) region and in the placenta at E 10.5. At E 11.5-12.5, the dHSCs start to migrate to the fetal liver and, at E 16.5, they further migrate, colonize and, finally, reside in the bone marrow throughout adult life (**Figure 2**) (Costa et al., 2012; Dzierzak and Speck, 2008; Medvinsky et al., 2011). In the AGM, HSC development requires a transition through the hemogenic endothelium, a rare cluster of vascular endothelial cells (Bertrand et al., 2010; Lancrin et al., 2009). The close developmental relationship between endothelial and hematopoietic lineages suggests a direct ontogeny link, wherein HSCs rise from an endothelial-to-hematopoietic transition (Kissa and Herbomel, 2010; Medvinsky et al., 2011). Imaging studies in mouse and zebrafish embryos already described the endothelial origin of HSCs (Boisset et al., 2010; Zovein et al., 2008).

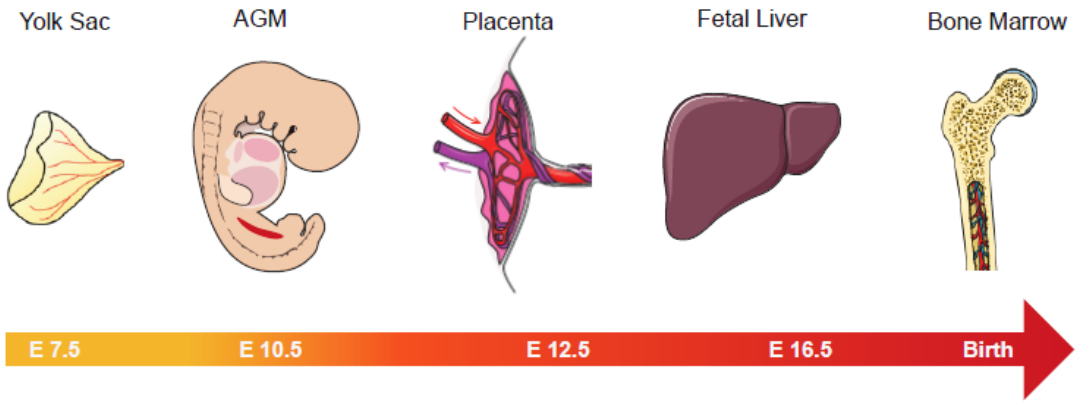


Figure 2:Figure 2: Mouse embryonic sites of hematopoietic development. In mice, at embryonic day 7.5 (E 7.5), primitive hematopoiesis starts in the yolk sac. At E 10.5, dHSCs appear in the AGM region migrating to placenta. Then, dHSCs migrate to fetal liver (E 12.5) and colonize the bone marrow at E 16.5.

Bone marrow is heterogeneous at the cellular level by including different populations of progenitors with different self-renewal and differentiation rate. Despite its high hematopoietic activity, HSCs located in the bone marrow represent only 0,01% of total nucleated cells (Rossi et al., 2011). Therefore, HSCs characterization and isolation has been possible due to either the presence,

or lack, of certain cell surface markers (Rossi et al., 2011). One of the most important surface cell markers is the glycophosphoprotein CD34 that is expressed on hematopoietic stem and progenitor cells in both human and mice (Guo et al., 2003; Krause et al., 1994; Stella et al., 1995).

1.1.1 Clinical Relevance of HSCs

Throughout an individual's lifetime, the self-renewal and differentiation properties of HSCs allow the equilibrium and maintenance of the hematopoietic system making them one of the most regenerative tissues (Doulatov et al., 2012). Because of that, HSCs are the perfect candidates for regenerative medicine purposes.

Consequently, those cells are the functional units for hematopoietic stem cell transplant (HSCT), a procedure indicated against hematological diseases such as leukemia, lymphoma, myeloma, bone marrow failures and other bone marrow/blood-related disorders (Apperley et al., 2016; Passweg et al., 2012). There are two types of HSCT: autologous and allogeneic transplantation. In autologous transplantation, HSCs are collected from the patient, stored and later returned to the same patient. This procedure is highly used during hematopoietic failure due to chemotherapy treatments, for example. In patients with acquired or congenital hematopoietic system failure, allogeneic transplantation is the most commonly used: the HSCs are obtained from related or unrelated donors to replace the blood system (Passweg et al., 2012). In the past 20 years, the annual number of HSCT performed has a constant increase for both autologous and allogeneic HSCT, according to the European Society for Blood and Marrow Transplantation 2014 (Passweg et al., 2016). In 2014, from a total of 36 469 patients, 40 829 HSCT were performed. Within the total of HSCT, 16 946 were allogeneic (42%) and 23 883 autologous (58%), representing an increase of 4.1% when compared with 2013 (Passweg et al., 2016). Data from 2012 showed an increase of 26% and 88% of HSCT performed when compared to data from the previous 5 and 15 years, respectively. (Passweg et al., 2015).

Despite all the advanced techniques to obtain HSCs from bone marrow, peripheral blood (after mobilization) and umbilical cord blood (UCB), the number of cells obtained remain limited and insufficient to treat an adult, especially in the case of UCB (Siena et al., 2000; Smith and Wagner,

2009). Regarding allogeneic transplantation, the major challenge is to find a match of the human leukocyte antigen (HLA) gene between donor and host. HLA-matched sibling has been shown to be the best clinical outcome. Unfortunately, each patient's sibling has only 25% change of being HLA-matched (Apperley et al., 2016). Therefore, since HSCs are not genetically identical to the host, there is an increase of the risk of graft-versus-host disease (GvHD) caused by the host immune system during allogeneic transplantation (Passweg et al., 2012; Petersdorf, 2013; Takizawa et al., 2011). On the other hand, in autologous transplantation, there is a risk of graft contamination with cancer cells, although there is no graft rejection nor GvHD (Batra A. Cottler-Fox M. Harville T. Rhodes-Clark B.S. Makhoul I. Nakagawa M., 2014).

To overcome the complications associated with both types of HSCT, *in vitro* expansion of HSCs has been attempted to produce a cell population suitable for bone marrow transplantation (Takizawa et al., 2011). The maintenance and expansion of HSCs *in vitro* remains challenging due to a lack of acceptable cell culture conditions (Schuster et al., 2012). Under *in vitro* conditions, HCSs tend to lose their self-renewal potential.

Therefore, efforts have been made towards to generate engraftable patient-specific HSCs from additional cell sources. The first endeavour has been the generation of HSCs by inducing pluripotent stem cells (PSCs) hematopoietic differentiation (Kim and Daley, 2009). However, PSC-derived hematopoietic cells do not fully mature, having limited multilineage long-term engraftment potential, being not able to home the bone marrow (Keller et al., 1993; Lim et al., 2013). Multiple approaches have been taken to mimic hematopoietic developmental environment to successful obtain long-term engraftment HSCs. One of the strategies was co-culturing PSCs with other cell lines to differentiate within teratomas to derive functional cells from myeloid and lymphoid lineages (Amabile et al., 2013; Suzuki et al., 2013). Both studies provide a useful tool to study the HSCs but limited for clinical application since the PSC-derived HSCs are also teratoma-derived. More recently, a report showed the generation of engraftable hematopoietic stem and progenitor cells (iHSPCs) using a single transcription factor, MLL-AF4. Despite all the promising results, the iHSPCs are susceptible to leukemic transformation during the long-term post-transplanted period, the mice

developed leukemia (Tan et al., 2018). Until now, PSC sources cannot generate functional and long-term engraftment HSCs with optimal translational value.

Thus, new alternatives are needed to obtain dHSCs *in vitro* and an exciting alternative is the direct reprogramming technologies of somatic cells.

1.2 Cellular Development to Reprogramming

During embryonic development, cells of the embryo start to commit into three germ layers- endoderm, mesoderm and ectoderm- that will differentiate into tissue-specific cell type generating a multicellular organism. As the cell differentiation levels increase, the totipotency potential is gradually lost giving rise to cells with specific functions. The specialized cell fate is controlled by a complex network of transcriptional and epigenetic mechanisms (Bernstein et al., 2007; Jenuwein and Allis, 2001). In 1942, Conrad Waddington defined epigenetics as "the branch of biology which studies the causal interactions between genes and their products which bring the phenotype into being" (Waddington, 1968). Nowadays, by the definition, Epigenetics refers to the all the pathways that modulate the expression of a genotype into a phenotype, without changing the original DNA sequence (Dupont et al., 2009).

According to Waddington's "epigenetic landscape" model, cell differentiation is a unidirectional process where the pluripotent cells stay on top of the hill and, when the differentiation starts, they "roll-down" to some different groves. Depending on the grove, they will acquire a specific cell fate. Thus, this model defends that cell differentiation is an irreversible process (Ladewig et al., 2013).

However, this model was challenged by John Gurdon in 1958 with his work in somatic cell nuclear transfer (SCNT). Using *Xenopus laevis* in his experiment, Gurdon transfer the nuclei from differentiated tadpole intestinal cells to enucleated oocytes, generating frogs with normal development (Gurdon, 1962; Gurdon et al., 1958). Thus, it has been shown that the nucleus contains all the genetic information to generate all cell types and the differentiation is a reversible process. Later, the same methodology was applied to mammalians generating the sheep "Dolly". Campbell and colleagues transplanted the nuclei of mammary glands into enucleated oocytes from a sheep. The

blastocyst was then placed in another female sheep uterus to continue the development until birth (Campbell et al., 1996; Wilmut et al., 1997).

After SCNT into enucleated oocytes, other line of research emerged, confirming that the somatic epigenome can return to pluripotency by alteration of the gene expression profile by employing cell fusion of pluripotent cells with differentiated cells.

Cell fusion of two or more cells can generate heterokaryons or hybrids. Hybrids are capable of cell division and DNA replication, leading to a fusion of the original cells nuclei. In hybrids, fusion can occur between cells of the same species (euploid hybrid) or between different species (aneuploid hybrid). In the case of heterokaryons, these cells do not proliferate and, consequently, are multinucleated (Yamanaka and Blau, 2010). In 2001, Tada and colleagues induced nuclear reprogramming of somatic cells by fusing embryonic stem cells (ESC) with adult thymocytes. They found that the Oct4-GFP transgene, normally inactivated in thymocytes, is reactivated 48h after cell fusion, demonstrating the acquisition of pluripotency by the ES-thymocyte hybrid cells (Tada et al., 2001). In another report, fusion between ES and neural stem cells, lead to the overexpression of Nanog stimulating the activation of pluripotent genes (Silva et al., 2006). In 2008, Pereira converted human-B-lymphocytes cells with ESC allowing the identification of Oct4 as critical for the establishment of the pluripotent state (Pereira et al., 2008). In summary, somatic cells can be reprogrammed through the fusion with ESC, generating heterokaryons and hybrid cells that express pluripotent genes typical of ESCs.

The overexpression of transcription factors (TF) in somatic cells can activate cohorts of certain genes that can change the cell fate. This phenomenon was first described in *D. melanogaster larvae* (Schneuwly et al., 1987) and later in mammals (Davis et al., 1987). Therefore, TF-mediated reprogramming is another type of nuclear reprogramming.

In 2006, Yamanaka surprisingly discovered that TFs that maintain stemness in ESC were sufficient to induce the recovery of the pluripotent state by somatic cells. In other words, overexpressing four TFs – Oct4, Sox2, Klf4 and c-Myc- known as OSKM or "Yamanaka factors" in mouse fibroblasts induced pluripotent stem cells (iPSCs) (Takahashi and Yamanaka, 2006;). With the same combination of TFs, human iPSCs (hiPSCs) were generated from human fibroblasts (Takahashi

et al., 2007). This exciting breakthrough has revolutionized the stem cells and reprogramming fields. Cell reprogramming can therefore be defined as the alteration of the cellular identity of one cell type to another by altering the transcription profile and epigenetic states (Yamanaka and Blau, 2010).

1.2.1 Direct Reprogramming

After Yamanaka's findings, several efforts ignited for the induction of somatic cell-types employing combinations of TFs. Direct reprogramming of one cell type to another using a combination of TFs, without passing through a pluripotent state, represent an exciting alternative for the generation of iPSCs and subsequent differentiation that remains challenging (Xu et al., 2015).

Currently, there are several direct reprogramming strategies, mostly using fibroblasts, to obtain different cell types such as cardiomyocytes (Ieda et al., 2010; Inagawa et al., 2012; Qian et al., 2012), adipocytes (Kajimura et al., 2009), hepatocytes (Huang et al., 2011), endothelial cells (Han et al., 2014), different type of neurons (Sheng et al., 2012; Son et al., 2011; Vierbuchen et al., 2010) and neural stem cells (Ring et al., 2012).

Regarding hemogenic system, direct reprogramming strategies is being used as a tool to "produce blood", somatic cells are being reprogrammed into hematopoietic stem/progenitors with multilineage potential (Capellera-Garcia and Flygare, 2017) (**Figure 3**). The overexpression of C/EBPα or C/EBPβ converted B lymphocytes into macrophages (Xie et al., 2004). The same reprogramming process happen to fibroblasts when PU.1 was added to the combination described previously (Feng et al., 2008). In both studies, the reprogrammed cells showed macrophage-like phenotype and function.

Mature B cells have already been reprogrammed into erythroid cells by enforcing the expression of Gata1, Tal1 and C/EBPα (Sadahira et al., 2012). Sill about the erythroid lineage, in 2016, Gata1, Tal1, Lmo2 and c-Myc (GTLM) converted murine and human fibroblasts into primitive-like erythroid progenitors and precursors. The presence of Klf1 or Myb to the GTLM cocktail lead to the expression of adult haemoglobin since these two TFs are related with definitive hematopoiesis (Capellera-Garcia et al., 2016). This same group also showed that when Gata2 and Runx1 were added to the GTLM cocktail, the fibroblasts were reprogrammed into megakaryocytes (Pulecio et al., 2016).

These two TFs were sufficient to change the lineage of the reprogrammed cells to the megakaryocyte lineage instead of erythrocyte lineage.

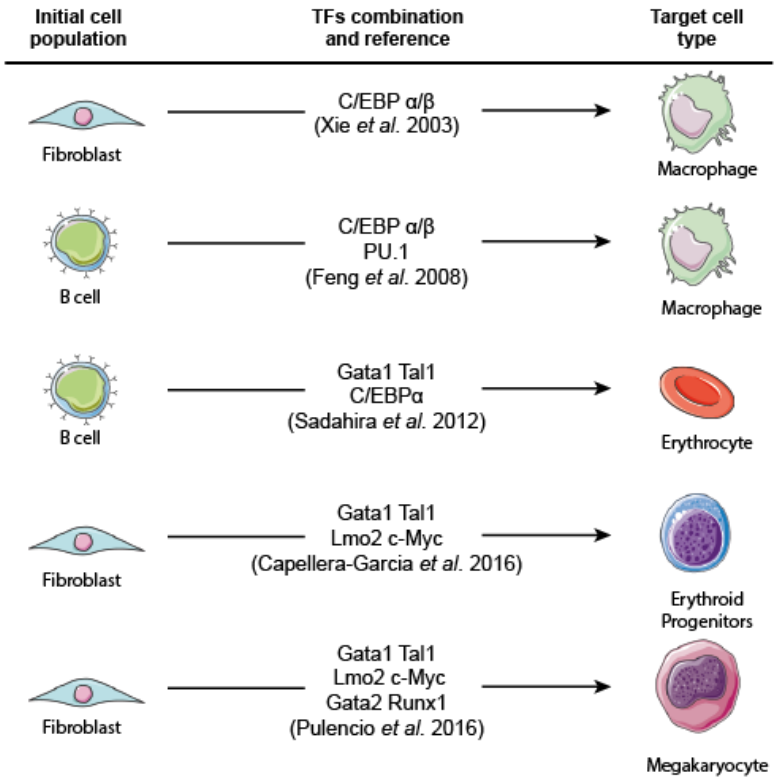


Figure 3:Several direct reprogramming strategies to obtain hematopoietic-like cells. Using fibroblast and B cells and overexpressing lineage-specific transcription factors is possible to generate macrophages, erythrocytes, progenitors and megakaryocytes.

Aiming at the generation of HSCs, Pereira and colleagues screened for specific hematopoietic transcription factors to induce hemogenic reprogramming in mouse embryonic fibroblasts (MEFs) from double transgenic mice huCD34tTA x TetO-H2BGFP (hereafter named 34/H2BGFP) (Pereira et al., 2013) (**Figure 4**).

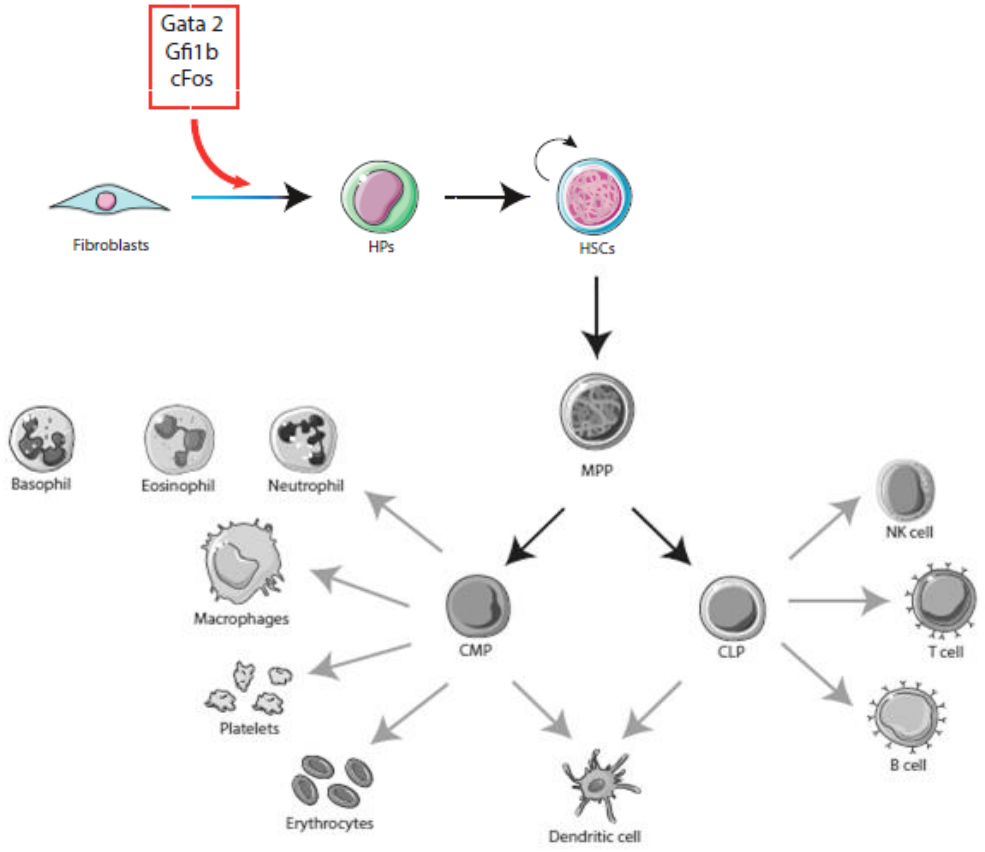


Figure 4: Direct Reprogramming of mouse embryonic fibroblasts (MEFs) into hemogenic precursors (HP). The overexpression of Gata2, Gfi1b and cFos is sufficient to induce hematopoietic stem-like cells through HP cells.

The reporter human CD34 (huCD34) is coupled with green fluorescent protein (GFP). Therefore, the acquisition of hematopoietic cell fate lead to an activation of the huCD34 and subsequently expression of GFP protein. The screening revealed that Gata2, Gfi1b, cFos and Etv6 can activate the huCD34 reporter, inducing an initial population of hemogenic precursors (Prom1+Sca1+CD34+Cd45-) with endothelial-like transcriptional program, mimetizing the endothelial-to-hematopoietic transition (EHT) that occurs during developmental hematopoiesis. A minimal combination of Gata2, Gfi1b and cFos showed to be sufficient to activate the reporter. Thus, Gata2 binds to the open chromatin and recruits Gfi1b and cFos to the target genes, inhibiting the fibroblasts genes and permitting the expression of endothelial/hematopoietic genes (Gomes *et al.* in revision). This strategy was also proved in the human system, using human dermal fibroblast with the same cocktail of TFs (Gomes *et al.* in revision). Taking together, since definitive hematopoiesis

emerges from endothelial-to-hematopoietic transition, this report demonstrates that a simple combination of four TFs can induce a complex and dynamic developmental process *in vitro* (Pereira et al., 2013).

The resemblance of EHT spotted in the hemogenic reprogrammed cells lead to another work, where mouse placentas were analysed and a population with similar phenotype and gene network was identified (Pereira et al., 2016). This study *in vivo* validated the previous results *in vitro* (Pereira et al., 2013) given rise to a new tool to study what occurs during the hemogenic development *in vivo.*

After Pereira's findings, several groups used similar approaches with different sets of TFs to convert fibroblasts into hematopoietic progenitors (Batta et al., 2014; Cheng et al., 2016), non-hemogenic endothelial cells into hematopoietic progenitors (Sandler et al., 2014) or generating HSCs from common myeloid progenitors (Riddell et al., 2014) or from non-hemogenic endothelial cells (Lis et al., 2017).

In patients with hematopoietic disorders, starting the reprogramming process from hematopoietic progenitors can be a disadvantage since the mutation will be carried through progenitor cells limiting the use of autologous HSCT (Pereira et al., 2014). Recently, Sugimura used a combination of seven TFs (ERG, HOXA5, HOXA9, HOXA10, LCOR, RUNX1 and SPI1) to obtain immature HSCs from human iPSCs. After transplantation in the mice bone marrow, this novel reprogrammed HSCs suffer a complete maturation and long-term engraftment potential (Sugimura et al., 2017).

Overall, these studies demonstrate that hemogenic reprogamming can recapitulate initial stages of hemogenic development allowing the study with more detail of a system as complex as the hematopoietic one.

1.3 The hematopoietic transcription factors: Gata2, Gfi1b and cFos

1.3.1 The role of Gata2, Gfi1b and cFos during hematopoiesis

Our group identified Gata2, Gfi1b and cFos (GGF) as the TF pool necessary to induce hemogenic precursors from mouse and human fibroblasts (Pereira et al., 2013) (Gomes et. *al.* in revision). The expression of these three TFs activated the huCD34/H2BGFP reporter and, consequently, the generation of GFP+ cells. Therefore, Gata2, Gfi1b and cFos network is fundamental for hemogenic reprogramming. These findings agree with what is known about the role of the GGF during hematopoiesis.

Gata-binding protein 2 (Gata2) is a member of the GATA family. The TFs that belong to this a family can bind to the (A/T)GATA(A/G) DNA consensus sequence of promotors and enhancers of target genes (Ko and Engel, 1993; Orkin, 1992). During hematopoiesis, Gata2 is expressed among all hematopoietic cells, especially in early progenitors, megakaryocytes and in mast cells (Tsai and Orkin, 1997; Tsai et al., 1994). More importantly, Gata2 is crucial for proliferation and self-renew of HSCs (Kosan and Godmann, 2016; Ling et al., 2004; Minegishi et al., 2003). Experimental approaches i*n vivo* already explored the role of Gata2 during different phases of hematopoiesis. $Gata2^{-/-}$ mouse embryos are anemic, the yolk sac appears pale and none embryo survived beyond embryonic day 11.5 indicating that Gata2 is essential in the early hematopoiesis (Tsai et al., 1994). Moreover, $Gata2^{-/-}$ adult mice result in defects in production and maintenance of HSCs, demonstrating the importance of this TF for the proliferation and homeostasis of HSCs in the AGM region and bone marrow, respectively (Ling et al., 2004).

Growth factor independent 1b (Gfi1b) is a transcriptional repressor with an important role in the generation of megakaryocyte and erythroid lineages (Foudi et al., 2014; Saleque et al., 2002). Gfi1b-deficient embryos maintained the expression of endothelial genes impairing the emergence of HSCs (Lancrin et al., 2012). These results indicate that Gfi1b expression is required for EHT in AGM to generate the HSCs. Others reports already showed that Gfi1b binds to endothelial genes to epigenetically silence them during the generation and specification of hematopoietic cells (Khandanpour et al., 2010; Thambyrajah et al., 2015).

FBJ osteosarcoma oncogene (cFos) is a member of Fos family (cFos, FosB, Fra1, Fra2), dimerizes with the Jun family member such as c-Jun to form the transcription factor complex Activator Protein 1 (AP-1) (Angel and Karin, 1991; Halazonetis et al., 1988). This complex is associated with differentiation, cellular proliferation and even apoptosis (Angel and Karin, 1991). Regarding hematopoietic processes, studies in the mouse placenta indicated that the absence of cFos lead to low levels of HSCs activity (Ottersbach and Dzierzak, 2005). Other report described that Fos gene was highly expressed in nascent-HSCs unlike in adult-HSCs (McKinney-Freeman et al., 2012).

Although all the reports describing the function of these TFs in hematopoiesis, little is known about the protein features and the interactions with each other during the hematopoietic reprogramming.

1.3.2 Gata2, Gfi1b and cFos protein domains and features

During hemogenic reprogramming, GGF transcriptor factors interact with each other and with other proteins through specific protein features such as the protein domains (**Figure 5**). More information about these features and how the interaction occurs will permit the optimization of the strategies already applied to HSC-induced reprogramming.

The modulation of the hematopoietic cell fate by Gata2 is possible due to the presence of two identical conserved zinc fingers (ZnF) in both C- and N-terminal that bind to the DNA (Rodrigues et al., 2012). Both ZnFs are known to interact with other proteins independently or together (Vicente et al., 2012). The N-ZnF binds independently to the DNA being mostly responsible for the stabilization of DNA-protein complexes while the C-ZnF binds to the GATA consensus sequence (Bates et al., 2008). Since, all GATA family member have zinc fingers, the non-zinc fingers domains may explain the difference between Gata genes (Vicente et al., 2012). Gata2 sequence also includes two transactivation domains (TAD), a nuclear localization signal (NLS) and a negative regulatory domain (NRD) (Minegishi et al., 2003; Tong et al., 2005; Vicente et al., 2012).

Gfi1b is a transcription factor known by its repressive function. Therefore, in the N-terminal region, the SNAG (Snail/Gfi1) domain binds to epigenetic regulators as the histone-modifying lysine-specific histone demethylase 1 protein (KDM1A or LSD1) (reviewed in 1.4), to epigenetically silence

target genes (Saleque et al., 2007; Thambyrajah et al., 2015; Velinder et al., 2016). In the C-terminal, from a total of six zinc fingers, the ZnF 1, 2 and 6 are responsible for protein-protein interaction, whereas ZnFs 3, 4 and 5 are for DNA binding at the TAAATCAC(T/A)GC(A/T) motif (Lee et al., 2010; Zweidler-Mckay et al., 1996).

cFos is a member of the protein complex AP-1 along with Jun (Angel and Karin, 1991; Halazonetis et al., 1988). All AP-1 proteins share a conserved basic leucine zipper (bZIP) domain to DNA binding to control genes containing AP-1 sites. In the C-terminal, Fos protein has a TAD domain for transcriptional regulation (Sassone-Corsi et al., 1988; Turner and Tjian, 1989).

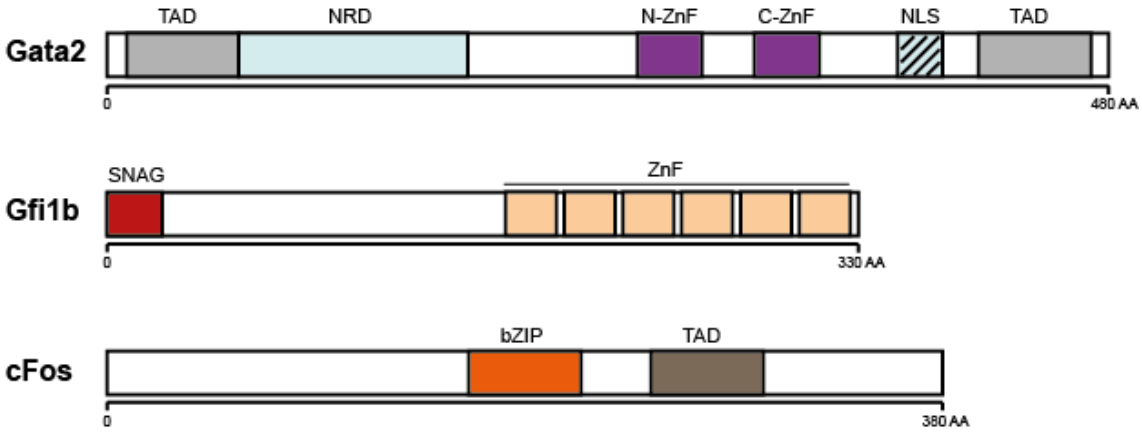


Figure 5:Schematic representation of Gata2, Gfi1b and cFos and respective domains. Gata2 has two transactivation domains (TAD) in N- and C-terminal, respectively. The TAD is followed by the negative regulatory domain (NRD), two conserved zinc fingers (N-ZnF and C-ZnF) for DNA binding and a nuclear localization sequence (NLS). Gfi1b presents a repressor domain (SNAG) and six zinc fingers (ZnF). Lastly, cFos has a basic leucine zipper (bZIP) to bind to DNA and a TAD domain as Gata2.

Since Gata2 has both dominant and independent targeting capacity during reprogramming (Gomes et *al.*, in revision), it would be interesting to unveiling the protein-protein interactions between Gata2 with the other two TFs but also with epigenetic regulators, such as LSD1.

1.4 Lysine-specific demethylase 1 (LSD1)

1.4.1 LSD1 protein complexes, structure and enzymatic activity

During the development, TFs interact with chromatin modifiers to maintain the correct cellular transcriptional profile. After binding to DNA, TFs recruit complexes to control the gene expression by modifications of histone tails allowing the activation or repression of the transcriptional profile (Maiques-Diaz and Somervaille, 2016). Methylation is the most common histone tail modification and occurs on both arginine and lysine residues. Being a dynamic process, these residues can be mono-, di-, in the case of arginine, or even trimethylated as in lysine, creating repressive or activating transcription. Lysine methylation, normally a H3K4 trimethylation, of both enhancers and promotors are often associated with active transcription (Kouzarides, 2007).

The first histone demethylase identified was the lysine-specific demethylase 1 (LSD1) (also called as KDM1A, BHC110 or AOF2), belonging to the flavin-adenine dinucleotide (FAD)-dependent amine oxidases family (Shi et al., 2004). LSD1 was identified in two different complexes: in the CoREST complex, also known as CoREST transcription repressor complex (Hakimi et al., 2003; Humphrey et al., 2001; You et al., 2001), and in the Mi-2/nucleosome remodeling and deacetylase (NuRD) complex (Wang et al., 2009). The integration of LSD1 into CoREST or NuRD complexes is essential for LSD1 nucleosome demethylation function by association with the RCOR1 or MTA2 from each complex, respectively (Lee et al., 2005; Shi et al., 2005; Wang et al., 2009).

LSD1 is characterized by the presence of a Swi3p, Rsc8p and Moira (SWIRM) domain, a coiled-coil Tower domain for protein-protein interactions and an amine oxidase (AO)-like domain (FAD-binding site)(Aravind and Iyer, 2002; Culhane and Cole, 2007). The SWIRM domain is responsible for the switch of LSD1 demethylation between H3K4 and H3K9 and strongly interact with AO-like domain resulting in an overall globular structure (Metzger et al., 2005). In the specific case of LSD1-RCOR1, this interaction between the SANT2 domain of RCOR1 and the Tower domain of LSD1 is required for the demethylase activity of this enzyme in the CoREST complex (Yang et al., 2006). All these domains give to LSD1 different structural advantages to maximize the interaction with other proteins from the complexes and, consequently, increase the catalytic activity.

1.4.2 Biologic role of LSD1

LSD1 is an essential regulator of mammalian biology, playing an important role in numerous cellular processes such as self-renewal, differentiation, cell death and metabolism, from stem cells to tumor cells (Hosseini and Minucci, 2017). This epigenetic regulator acts in a context-depending manner, depending on the cell type, association partners and function, being considered a co-activator but also a co-repressor (Hosseini and Minucci, 2017).

During development, LSD1 is mainly responsible for the stem cell maintenance. The ESC from LSD1$^{-/-}$ mice indicated a growth damage due to cell death, cell cycle and differentiation impairment. In fact, LSD1$^{-/-}$ mice perish at or before embryonic day 6 (Wang et al., 2007, 2008). Interestingly, LSD1 is expressed in undifferentiated human ESC and is downregulated during differentiation demonstrating its role in the maintenance of pluripotency and repression of lineage-specific developmental processes (Adamo et al., 2011). The transition from pluripotent to neural stem cells or from progenitors to mature neurons is also regulated by LSD1/CoREST complex (Ballas et al., 2005).

LSD1 has been reported to have a role also in DNA damage response (Mosammaparast et al., 2013), repression of mitochondrial metabolism during metabolic shift in cancer cells (Sakamoto et al., 2015) and hematopoiesis.

1.4.2.1 The role of LSD1 during hematopoiesis and its interactions with transcription regulators

The manipulation of gene expression during hematopoietic development is controlled by LSD1. In 2013, Kerenyi and colleagues reported that LSD1 is essential for epigenetic regulation during hematopoietic differentiation. Using a blood lineage-specific knockout mouse, they observed a decrease of the number of blood cells in the newborns, indicative of defects in immature bone marrow cells. Later, by evaluating the effects of LSD1 deletion in mature lineage cells, they established that LSD1 defects impairs early and terminal hematopoiesis, indicating the strong influence of LSD1 not only in initial stages but also in terminal granulocytic and erythroid

differentiation (Kerenyi et al., 2013; Sprüssel et al., 2012). Also in the primitive stage of hematopoiesis, LSD1 downregulation of a regulator of hemangioblast development- Evt2- is essential to initiate the hematopoietic commitment (Takeuchi et al., 2015).

Mechanistically, LSD1/CoREST complex interact with transcription factors to promote the differentiation of various blood cell lineages. The regulation of Gata1 and Gata2 expression, also called "GATA factor switching", is indispensable for erythroid differentiation. Since Gata2 is predominantly expressed in hematopoietic stem/progenitor cells and Gata1 regulates the expression of erythroid-specific genes, LSD1 binds to Gata2-locus, regulating the expression of both Gata proteins (Guo et al., 2015). LSD1 associates with Gata2 to repress Gata1 expression in early stages of hematopoiesis and forms a complex with TAL 1 (T-cell acute lymphocytic leukemia 1) to inhibit Gata2 expression in differentiated cells (Guo et al., 2016). The domains responsible for GATA2 interaction with LSD1 remain largely unknown.

Epigenetic regulation of hematopoiesis by Gfi1 and Gfi1b is also LSD1-mediated. LSD1/CoREST complex is recruited by Gfi1/1b via SNAG repressive domain to target gene promoters. LSD1 inhibition depresses Gfi1/1b target genes affecting the differentiation of several hematopoietic lineages (Saleque et al., 2007). The emergence of HSCs from the hemogenic endothelium is also regulated by Gfi1 proteins that further recruits LSD1 to epigenetically silence the endothelial genes, thus, triggering the EHT (Thambyrajah et al., 2015).

These studies prove that loss of LSD1 results in defects during several phases of hematopoiesis since the emergence, proliferation, and maintenance of HSC to differentiation of them into mature blood cell lineages, highlighting the importance of this regulator in normal hematopoiesis.

Moreover, LSD1 is also highly expressed in multiple malignancies due to its wide role in epigenetic regulation. LSD1 is expressed in hematological tumors such as acute myeloid leukemia (AML). In a mouse model of human MLL-AF9 AML cells, LSD1 is required to maintain the oncogenic transcriptional program (Harris et al., 2012). Other studies indicate that inhibition of LSD1 impairs acute erythroleukemia, acute megakaryoblastic leukemia and myelodysplastic syndrome by inducing cell differentiation (Ishikawa et al., 2017; Sugino et al., 2017).

Overall, the oncogenic role of LSD1 in different types of hematological diseases makes LSD1 a putative candidate for therapeutic intervention.

1.4.3 Pharmacologic inhibition of LSD1 and potential clinical use

Due to its enzymatic activity and importance during the development of hematological and neurological diseases, pharmacological inhibition of LSD1 rises as a potential therapy. Compounds with reversible or irreversible inhibition of LSD1 through the inhibition of the monoamine oxidase or impairment of the interaction LSD1-SNAG domain have been already described. Two of the most advanced LSD1 inhibitors are the RN-1 dihydrochloride and the T-3775440 (hereafter referred as RN-1 and GSK-LSD1, respectively). RN-1 is a tranylcypromine, there is it binds covalently to FAD, that showed to be highly effective in the elimination of several AML and acute lymphoblastic leukemia cell lines (McGrath et al., 2016). GSK-LSD1 induced a antiproliferative effect in small-cell lung cancer by disrupting the SNAG domain interaction between LSD1 and Gfi1b (Takagi et al., 2017).

A better understanding about the biological role of LSD1, especially in cancer, will provide the resources to ameliorate the pharmacologic inhibitors already synthetized.

1.4.4 LSD1-mediated cellular reprogramming

The importance of the epigenetic regulation during the cellular reprogramming led to an extensive study of LSD1 in this context. Since LSD1 interacts with the enhancers and promoters targets that are regulated by pluripotency-related TF such as Oct4, Sox2 and Nanog (Whyte et al., 2012), several studies emerged to evaluate the role of this epigenetic regulator in the generation of iPSCs.

In 2009, Li and colleagues reprogrammed human keratinocytes to hiPSCs using only two TFs from the "Yamanaka's cocktail", the Oct4 and Klf4. To this TFs cocktail, they added a LSD1 inhibitor mixed with a glycogen synthase kinase 3 inhibitor. This study conclude that the use of a parnate as an inhibitor of LSD1 facilitates the reprogramming process by inhibiting the H3K4 demethylation, increasing the binding of Oct4 and Klf4 to their target genes (Li et al., 2009).

In another study, using a chemical approach, the inhibition of LSD1 with small-molecules or by impairing the TGFβ pathway induced morphological changes in mouse epiblast stem cells (mEpiSC) towards to an earlier developmental pluripotency state, the mESC-like pluripotency state (Zhou et al., 2010). Also by LSD1 inhibition with a cocktail of small-molecules, Hou et *al.* converted mouse somatic cells into iPSCs inducing a chemical reprogramming process without genetic factors demonstrating that the pluripotent state can be induced in somatic cells using only small-molecules (Hou et al., 2013).

More recently, LSD1 was identified as a regulatory factor that can inhibit or delay the reprogramming process, thus, the chemical inhibition allowed the expression of a pluripotency-associated marker (TRA-1-60) in human fibroblasts (Cacchiarelli et al., 2015). During the conversion of B cells into iPSCs using C/EBPα, LSD1 showed to have a beneficial role in this type of reprogramming. C/EBPα increases the levels of several proteins including LSD1 to silence B cell genes and to induce pluripotency. The inhibition of LSD1 impaired the silencing of the B cells genes, resulting in a delay of the reprogramming process (Di Stefano et al., 2016). Regarding the studies mentioned and the potential involvement of LSD1 during the metabolic switch described in white adipose tissue (Duteil et al., 2014), Sun and colleagues explored the function of LSD1 during the iPSCs generation from reprogramming efficiency to metabolism. In this report, they demonstrated that LSD1 inhibition contributes to the reprogramming of iPSCs in two waves. First, LSD1 inhibition increases the expression efficiency of Oct4, Sox2 and Klf4 by stopping the H3K4 demethylation. Second, the genes from oxidative and glycolytic pathways are upregulated during reprogramming, thus, LSD1 inhibition induced changes in the expression of metabolic genes, facilitating a favourable metabolic switch for the reprogrammed cells (Sun et al., 2016).

These reports demonstrate the essential role of LSD1 during the reprogramming process by regulating the histone methylation and, consequently, the epigenetic state of the reprogrammed cells. With this information, the inhibition of LSD1 prevents H3K4 demethylation facilitating the exogenous expression of TFs and metabolic switch to improve the generation of iPSCs.

1.5 Challenge and aims

HSCT is the only clinical approach to re-establish a healthy hematopoietic system after a hematological disorder such as leukemia, myeloma or aplastic anemia. The low number of cells available for transplantation or the lack of compatibility between donor and host, leading to aggressive immune responses, are common clinical complications in HSCTs that hinder the wide application of this therapy. The generation of healthy, engraftable and autologous HSCs in large quantities is one of the biggest challenges in the modern regenerative medicine.

Therefore, direct reprogramming of fibroblasts into hematopoietic stem-like cells emerges as an exciting alternative to obtain patient-specific HSCs for HSCT. Nevertheless, the hematopoietic reprogramming efficiency needs to be improved for clinical translation of this technology.

The main challenge of this work is to understand the role of LSD1/CoREST complex during the hemogenic reprogramming in order to improve the reprogramming process. To do this, I will be focused in the following aims:

1. Identify Gata2 domains required for interaction with Gfi1b, cFos and LSD1/CoREST;
2. Elucidate the role of LSD1 during direct reprogramming of fibroblasts into HSCs;
3. Address Gata2 function during hemogenic reprogramming;

This work will generate valuable information on LSD1 and hemogenic reprogramming that will pave the way for novel therapeutic interventions. In other hand, a better understanding of Gata2 function, the dominant transcription factor, can also improve all the process, bringing reprogrammed HSCs on step closer to clinical application.

CHAPTER 2

Materials and Methods

2. Materials and Methods

2.1 Experimental approach:

To achieve the proposed objectives, the following strategies were used:

a) HEK293T cells transfection with FLAG-Gata2 deletion constructs, HA-Gfi1b and cFos to define which Gata2 hematopoietic domains are essential for the interaction with Gfi1b, cFos and LSD1. Includes transfection and immunoprecipitation followed by western blot.

b) Pharmacological inhibition of LSD1 during hematopoietic reprogramming with Gata2, Gfi1b and cFos. Includes lentiviral production, transduction of primary cultures of mouse embryonic fibroblasts from double transgenic mouse model (DT MEFs hCD34/H2BGFP)(Pereira et al., 2013), use the different inhibitors of LSD1 and GFP quantification by flow cytometry.

c) Hematopoietic reprogramming with Gata2 fusion domains, Gfi1b and cFos. Includes the cloning of activating or repressing domains to Gata2, lentiviral production, transduction of DT MEFs and GFP quantification by flow cytometry.

2.2 Cloning strategy

2.2.1 Lentiviral expression system

The lentiviral expression system used was the pHAGE2-MCS vector. This plasmid 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). The pHAGE2-MCS is a constitutive vector and its multiple cloning site (MCS) has restriction sites for NotI, MfeI, NheI, HpaI, XbaI, BamHI restriction enzymes. This restriction sites were used for gene cloning. pHAGE2-MCS vector also includes a human elongation factor-1 alfa (EF1α) promotor located before the MCS, an internal ribosome entry site (IRES) followed by a puromycin resistance gene ($Puro^r$) and an ampicillin resistance gene for bacterial selection (Amp^r). An additional pHAGE2 vector modified capable of expressing GFP was used as a control, the pHAGE2-MCS-IRES-eGFP.

The 2nd generation lentiviral system used is constituted by the transfer plasmids with the gene of interest, the packaging plasmid psPAX2 (Addgene, #12260) and by the envelope plasmid pMD2.G (Addgene, #12259) (**Figure 6**).

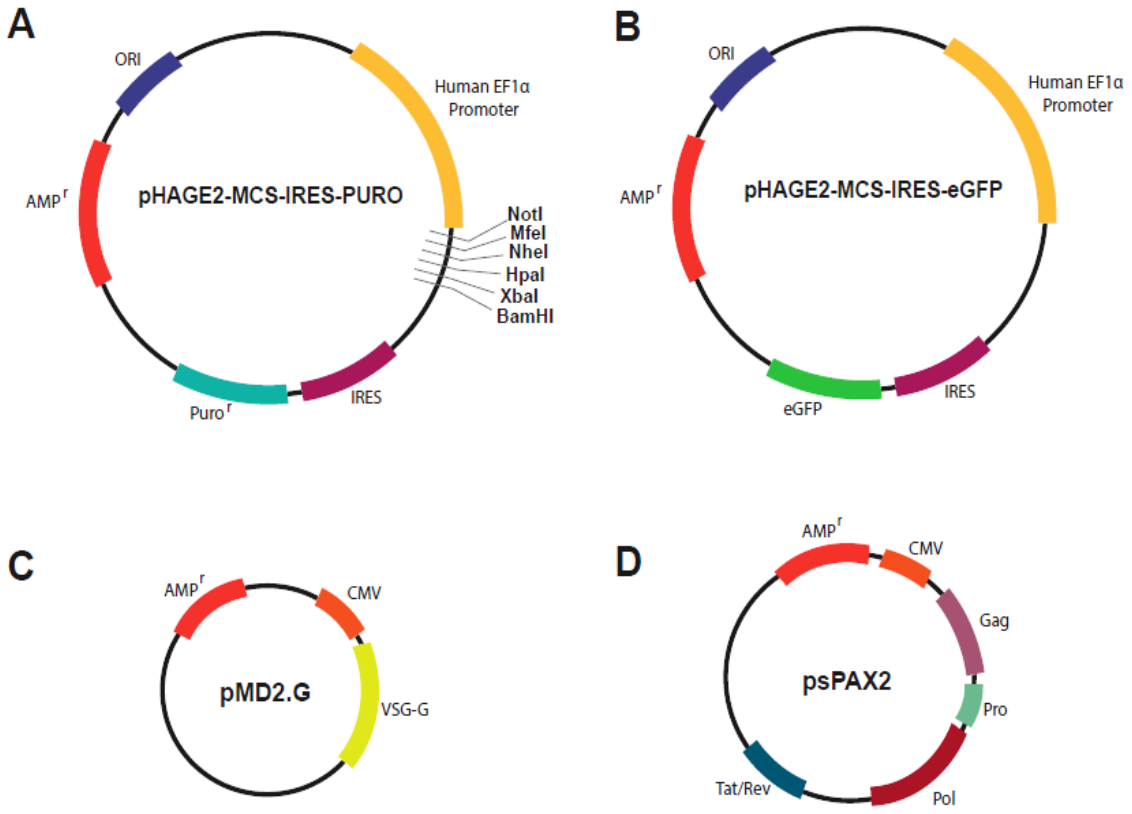


Figure 6: Lentiviral expression system. (A) pHAGE2-MCS-IRES-PURO plasmid is composed by a human EF1α promoter followed by a MCS with restriction sites for NotI, MfeI, NheI, HpaI, XbaI and BamHI. This plasmid also contains a IRES and two resistance genes against puromycin and ampicillin, respectively. **(B)** pHAGE2-MCS-IRES-eGFP plasmid was used in the control conditions. The difference between these plasmids is the presence of an eGFP gene after the IRES. For the 2nd generation lentiviral system, two more plasmids were used. **(C)** The pMD2.G plasmid has a cytomegalovirus promotor (CMV) followed by a VSV-G gene that encodes the virus envelope protein. **(D)** psPAX2 ha a CMV promoter that controls expression of the packaging proteins Gag, Pol, Rev and Tat. All plasmids used during this study had the ampicillin resistance gene allowing bacterial selection.

2.2.2 Template vectors

Genes and TFs of interest were purchased from Addgene or obtained from our lab's plasmid library. FLAG-Gata2 deletion constructs were kindly provided by Dr. Sjaak Philipsen, Erasmus MC, Rotterdam. All the genes were sent to sequencing (Sanger Sequencing Service, GATC Biotech) with

the respective sequencing primers. Then, the sequences obtained were aligned with their consensus coding sequence (CDS) from NCBI. The Vector NTI® software (Version 6) was used to do the alignment. In **Table 1**, there is the template plasmids used and the respective sequencing primer.

Table 1: Template plasmids and the specific sequencing primer used for the Sanger sequencing. In the FLAG-Gata2 deletion constructs, the symbol "Δ" stands for deletion and the numbers represent the amino acids that were deleted from the original FLAG-Gata2 sequence.

Template plasmid	Sequencing primer	Insert of interest	Source
pHAGE2-Gata2	**EF1α-F** (5´TCAAGCCTCAGACAGTGGTTC 3')	Gata2	Plasmid library
pHAGE2-Gfi1b		Gfi1b	Plasmid library
pHAGE2-cFos		cFos	Plasmid library
pcDNA3-FLAG-Gata2	**CMV-F** (5'CGCAAATGGGCGGTAGGCGTG 3')	FLAG-Gata2	Plasmid library (Dr. Sjaak Philipsen)
pcDNA3-FLAG Gata2Δ(69-153)		FLAG Gata2Δ(69-153)	
pcDNA3-FLAG Gata2Δ(154-256)		FLAG-Gata2Δ(154-256)	
pcDNA3-FLAG Gata2Δ(257-287)		FLAG-Gata2Δ(257-287)	
pcDNA3-FLAG Gata2Δ(287-342)		FLAG Gata2Δ(287-342)	
pcDNA3-FLAG Gata2Δ(343-379)		FLAG Gata2Δ(343-379)	
pcDNA3-FLAG Gata2Δ(380-440)		FLAG Gata2Δ(380-440)	
pcDNA3-FLAG Gata2Δ(1-74)		FLAG Gata2Δ(1-74)	
pcDNA3-FLAG Gata2Δ(1-235)		FLAG Gata2Δ(1-235)	
pcDNA3-FLAG Gata2Δ(440-480)		FLAG Gata2Δ(440-480)	
pTRE-Bio-TEV-AOF2	**CMV-F** (5'CGCAAATGGGCGGTAGGCGTG 3')	AOF2	
pTRE-BIO-CoREST1	**CMV-F** (5'CGCAAATGGGCGGTAGGCGTG 3')	CoREST1	
pLV-TRE-HA-Gfi1b	**CMV-F** (5'CGCAAATGGGCGGTAGGCGTG 3')	HA-Gfi1b	Plasmid library
pHR-SFFV-KRAB-dCas9-P2A-mCherry	**pHR-F** (5´GCTTCCCGAGCTCTATAAAAGAG3´)	KRAB	Addgene #60954
pAF1-VP16	**pAF1-F** (5' GGCTAACTAGAGAACCCACTG3')	VP16	Addgene #51426
pAF1-VP64		VP64	Addgene #51427

2.2.3 Polymerase Chain Reaction (PCR) Amplification

2.2.3.1 Primer design of Gata2 fusion proteins

For the primer design, each gene was analysed in the Vector NTI® software to detect the restriction sites of NotI (5' GCGGCCGC 3'), NheI (5'GCTAGC 3'), XbaI (5' TCTAGA 3') and BamHI (5' GGATCC 3') restriction enzymes. The enzymes MfeI and HpaI were excluded because the pHAGE2-MCS was cut outside the MCS by the first one and HpaI produces blunt ends.

To create fusion proteins of Gata2 with KRAB (repressive domains), VP16 or VP64 (both activating domains) (Beerli et al., 1998; Lupo et al., 2013), the sequences of these genes plus the sequence of the pHAGE2-MCS vector were analysed for the presence of the BstBI restriction site (5´TTCGAA 3´). Since this restriction site did not belong to the sequence of the vector neither any of the genes, BstBI was chosen to link the Gata2 to the other domains as a fusion protein.

The Forward primer (5') includes the restriction site of the enzyme plus the initial twenty nucleotides of the coding sequence of the gene to be cloned. The Reverse primer (3´) also had the enzyme restriction site plus the twenty nucleotides of the reverse complementary sequence. In addition, six random nucleotides were added to the primer forward (5'-GGTATC-3') and reverse (5'-CCTTAC-3'), before the enzyme restriction site, to increase the efficiency of the endonucleases during the process of amplification.

For the fusion proteins, the primer forward for KRAB, VP16 and VP64 were designed to remove the stop codon of KRAB, VP16 and VP64 genes. All these domains were added to the N-terminal of Gata2 gene

The scheme of the cloning strategy of KRAB-Gata2, VP16-Gata2 and VP64-Gata2 is shown in **Figure 7**. The primers are summarized in **Table 2**.

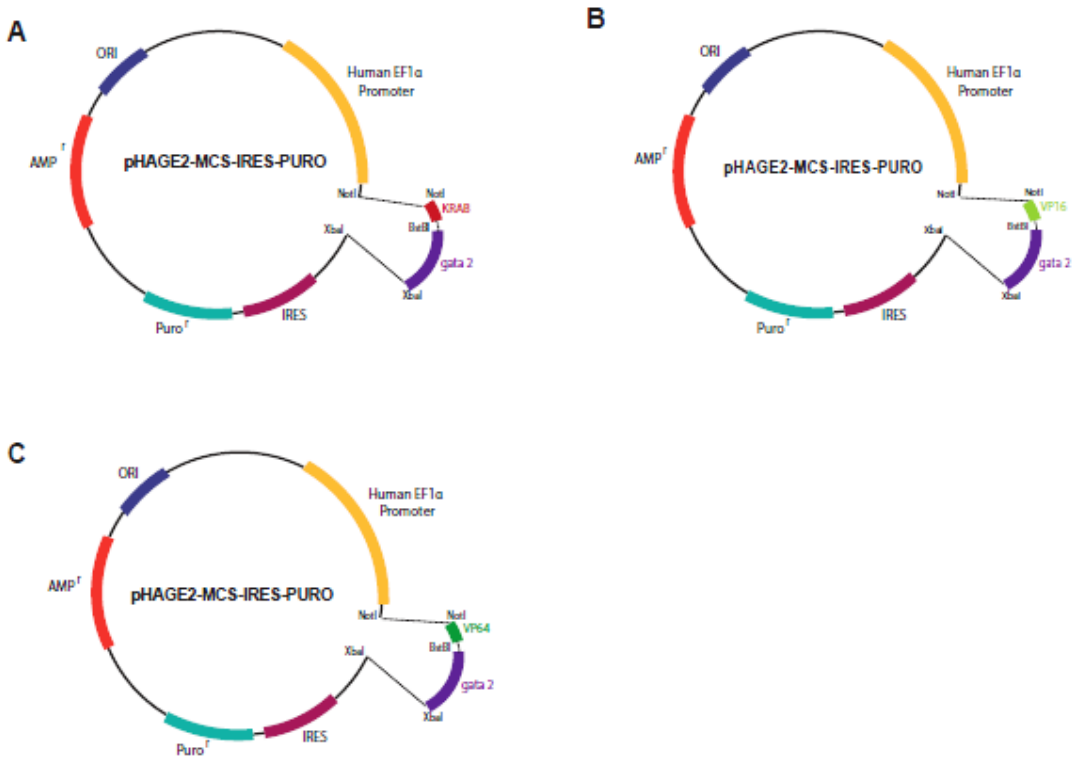


Figure 7:Scheme of the cloning strategy into pHAGE2-MCS-IRES-PURO vector: (A) (B) (C) Fusion protein of Gata2 with KRAB, VP16 and VP64 domain, respectively. Each domain was linked to the N-terminal of Gata2 through the BstBI restriction site.

Table 2: Table 2: Primers designed for cloning into pHAGE2-MCS vector.

Insert	5´ Enzyme	Primer Forward	3' Enzyme	Primer Reverse
KRAB	NotI	GGTATCGCGGCCGCatgcggacactggtgaccttcaa	BstbI	CCTTACTTCGAAggctcttctcccttctccaa
VP16	NotI	GGTATCGCGGCCGCatggccccccgaccgatgtca	BstbI	CCTTACTTCGAAcccaccgtactcgtcaattc
VP64	NotI	GGTATCGCGGCCGCatggacgcattggacgattttga	BstbI	CCTTACTTCGAAcagcatgtccaggtcgaaat
Gata2	BstbI	GGTATTCTTCGAAgaggtggcgcctaggcagcc	XBaI	CCTTACTCTAGActagccatgccatggcagtcacca

2.2.3.2 PCR Reaction and DNA purification

PCR protocol was performed to amplify genes of interest. The Phusion High-Fidelity PCR Master Mix (ThermoFisher, F-548) was one of the PCR components for a final reaction volume of 20µL. Primers and DNA template were the other components as described in **Table 3**. The PCR cycle program is defined in the **Table 4**.

Table 3:PCR components and respective concentrations for a final reaction volume of 20µL.

PCR component	Volume (µL)	Concentration
Water	8	-
Phusion High-Fidelity PCR Master Mix	10	1x
Primers (Forward+Reverse)	1	10 µM
DNA template	1	5-25 ng/µL

Table 4:PCR cycle program for Phusion High-Fidelity PCR Master Mix.

Cycle step	Temperature (°C)	Time	Cycles
Initial Denaturation	98	10 s	1
Denaturation Annealing Extension	98 67 72	1 s 5 s 30 s	30
Final Extension	72	1 min	1
	4	∞	

After PCR, the amplification products were mixed with Orange-G 1x loading buffer (VWR, E783) and loaded in 1% ultrapure grade agarose gel electrophoresis (NZYTech, MB05201) with 0,5 µg/mL of ethidium bromide. The NZYDNA Ladder III (NZYTech, #MB04402) and the samples ran for 1h under a current of 100V. Afterwards, bands with the correct size of the amplified insert were cut from the gel and DNA was purified according to NZYGelpure Kit protocol (NZYTech, MB01101). In the end, DNA was quantified using the NanoDrop (ND 1000 Spectrophotometer, Alfagene).

2.2.4 Restriction and Ligation of pHAGE2-MCS vector and inserts

The pHAGE2-MCS vector was double digested with NotI (ER0591) and XbaI (ER0682) restriction enzymes from ThermoFisher Scientific for 2h30 at 37°C. The purified inserts were restricted with different combination of enzymes for 1h30 at 37°C. KRAB, VP16 and VP64 were restricted with BstBI (ER0121) and XbaI, Gata2 was restricted with NotI and BstBI.

Next, the restricted plasmid and inserts were once again purified using a modified protocol of the NZYGelpure kit. The concentration of the purification products was measure using the NanoDrop. The components of the restriction reaction are described in the **Table 5**.

Table 5:Components necessary for the enzyme restriction protocol.

Component	Vector	Insert
Water	Add until 50 µL	13 µL
Buffer	5	5 µL
5'Enzyme	1	1 µL
3'Enzyme	1	1 µL
DNA	2,5 µg	5-25 ng/µL

Afterwards, ligation of the inserts and plasmid was performed at room temperature for 1h. To 1µL of restricted plasmid and 7µL of restricted insert was added 1µL of T4 DNA ligase (NEB, B0202) and the specific buffer. In the fusion proteins ligation protocol, the volume corresponding to the insert was divided by the two inserts. A ligation with water substituting the insert was performed as negative control of the transformation.

Recombinant DNA was immediately used for transformation with competent bacteria.

2.2.5 Production of chemically competent bacteria

The production of competent bacteria was based on two protocols from Chun et al. (Chung et al. 1989; Chung & Miller 1993).

Escherichia coli (*E. coli*) DH5α (NEB, C29871) was streaked on a plate with LB Agar (NZYTech, MB11801) without ampicillin and incubated at 37°C, overnight (≤16h). An isolated colony was picked for a liquid culture of 5mL of LB Broth (NZYTech, MB38801). The culture was incubated at 37°C, 200 rotations per minute (rpm), overnight (≤16h).

Next, the liquid culture was diluted 1:100, 1:150 and 1:200 in 50mL cultures with LB Broth and these cultures were incubated at 37°C, 200 rpm until de optical density (OD) at 600nm reach 0.5-0.6. When the optimal OD reached that range, the cultures where incubated for 10 min on ice to stop bacterial growth. Next, the cultures were centrifuged (Eppendorf Centrifuge 5810R, IL055). at 4°C,

3000rpm for 10 minutes (min). The supernatant was discarded. The pellets were resuspended in TSS buffer (**Table 6**) previously storage at 4°C. The bacterial suspensions were aliquoted and stored at -80°C.

To improve the efficiency of the competent bacteria, a transformation of 80ng of psPAX2 plasmid DNA was performed in a plate with LB Agar and ampicillin, incubated at 37°C, overnight (≤16h). The number of colony forming units (CFU) per DNA was calculated indicating the efficiency.

Table 6:TSS Buffer components.

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

2.2.6 Transformation of competent bacteria with recombinant DNA

After ligation, 10µL of ligation product were mixed with 50µL of competent bacteria, previously defrosted on ice, and incubated on ice for 30 min. The next step is the thermal shock at 42°C for 1 min followed by a cool down on ice for 5 min. LB Broth without ampicillin was added allowing bacterial growth for 25 min at 37°C. In the end, the bacterial solution was plated in LB Agar supplemented with ampicillin and incubated at 37°C, overnight (≤16h). Two additional transformations were made, one with the original non-restricted plasmid (positive control) and the restricted plasmid without insert (negative control). Additional transformations of plasmids from lab's library were also performed.

2.2.7 Detection of positive colonies by Colony PCR

After incubation, the growth colonies were picked and a PCR amplification with the primers (**Table 2**) was performed. The DNA template of the original plasmid of the inserts and the pHAGE2-

MCS were used as a positive and negative control, respectively. Besides the DNA template and primers, the NZYTaq 2x Colourless Master Mix (NZYTech, MB04003) was one of the Colony PCR components for a final reaction volume of 10µL (**Table 7**). For the fusion proteins were used two positive controls: the pHAGE2-Gata2 plasmid as a positive control for Gata2 and the original DNA templates of KRAB, VP16 and VP64 as a positive control for each domain. The Colony PCR cycle program is described in the **Table 8**.

Table 7: Colony PCR components for a final volume of 10µL.

Colony PCR component	Volume (µL)	Concentration
Water	4,5	-
NZYTaq 2x Colourless Master Mix	5	1x
Primers (Forward+Reverse)	0,5	10 µM

Table 8: Colony PCR cycle program for NZYTaq 2x Colourless Master Mix.

Cycle step	Temperature (°C)	Time	Cycles
Initial Denaturation	95	1 min	1
Denaturation	95	40 s	25
Annealing	67	40 s	
Extension	72	1 min	
Final Extension	72	7 min	1
	4	∞	

Afterwards, colony PCR products were mixed with Orange-G 1x loading buffer (VWR, E783) and loaded in 1% routine grade agarose gel electrophoresis (NZYTech, MB114) with 0,5 µg/mL of ethidium bromide. The NZYDNA Ladder III (NZYTech, #MB04402) and the samples ran for 1h under a current of 100V. Afterwards, bands were detected under ultraviolet (UV) light and the DNA bands with the matching size of the positive control were considered positive colonies. Next, the positive colonies were once again picked and grown at 37°C, overnight (≤16h), in a liquid culture of 3mL of LB Broth with ampicillin. The recombinant DNAs were then purified using the NZYTech Miniprep kit protocol (NZYTech, MB01001) and sent for sequencing (GATC Biotech Sanger Sequencing Service) to

validate final plasmid sequence. For sequencing, the sequencing primers in **Table 1** were sent together with the sample.

2.2.9 Isolation of Recombinant DNA by Midiprep

The verified positive colonies and the DNAs that were already cloned in the lab's library were diluted in 200mL of LB Broth with ampicillin at 37°C and 200 rpm, overnight (≤16h). The modified NZYMidiprep Endotoxin Free kit protocol (NZYTech, MB27901) was performed allowing the extraction and purification of the DNA. For higher DNA yield, a DNA precipitation step with 0,1 volumes of 3.0 Sodium Acetate Buffer Solution, pH 5.2 (S7899-100ML, Sigma-Aldrich) and 0,7 volumes of isopropanol at room-temperature was performed. The concentration obtained was measure using the NanoDrop.

2.3 Cell Culture of DT MEFs and HEK293T cells

Two major cellular systems were adopted to reach the aims of this project. (HEK)293T (ATCC) cells are derived from human embryonic kidney. 293T cells are recognized for their high transfection capacity and therefore they were used to produce lentiviruses.

For the direct reprogramming experiments, primary cultures of mouse embryonic fibroblasts from double transgenic mouse model (DT MEFs hCD34/H2BGFP) were used (Pereira et al., 2013). DT MEFs were isolated and sorted by flow cytometry for the absence of CD45 and GFP markers.

HEK293T and DT MEFs cells were defrosted, centrifuged at 1200 rpm and cultured in Dulbecco's modified eagle medium (DMEM) (Life Technologies, #21969035) supplemented with 10% (v/v) of fetal bovine serum (FBS) (Gibco, #10207106), 2mM L-Glutamine (Gibco, #25030081) and 10µg/mL of Penicillin-Streptomycin antibiotic (Gibco, 15070063). DT MEFs were plated in 0,1% of gelatin allowing optimal conditions for cell attachment. Cells were maintained at 37°C in 5% CO_2 until confluent.

During experiments, after transduction, DT MEFs were cultured in filtered MyeloCult M5300 medium (StemCell Technologies, #05350) supplemented with 10µg/mL of Penicillin-Streptomycin

antibiotic and 1mM of hydrocortisone (StemCell Technologies, #07904) to promote cell growth. LSD1 inhibitors such as RN-1 dihydrochloride (Tocris, #4977) and GSK-LSD1 (Sigma Aldrich, #SML1072-5MG) were added to the medium at day 2 or day 6 until day 20. Both stock solutions of 10mM were made with DMSO and kept at -20°C.

When confluent, cells were washed with phosphate-saline buffer (PBS) (Gibco, #10010-056) and detached with TrypLE Express Enzyme (1x) (ThermoFisher Scientific, #12604021). Cells were kept in 10% dimethyl sulfoxide (DMSO) and 90% of FBS solution at -80°C or liquid nitrogen until further use.

2.4 Transfection of HEK293T cells

A 2nd generation lentiviral system was used to produce lentiviral particles. This system is formed by three plasmids: the packaging plasmid (psPAX2), the envelope plasmid (pMD2.G) and the lentiviral transfer plasmid. As transfer plasmid, all the genes used were previously cloned into a pHAGE2-MCS plasmid. The transfer plasmid pHAGE2-MCS-IRES-eGFP was used as a positive control and the pHAGE2-MCS empty backbone was the negative control during this study. Both controls were modified versions from the original vector.

For viral production, 293T cells were diluted 1:6 (50-60% of confluence) in 10 cm plates. In the day after, each plate of cells was transfected with 10 µg of transfer plasmid, 10 µg of psPAX2 and 5µg of pMD2.G. First, the DNAs were mixed, and sterile ultrapure water was added until a final volume of 500 µL. Then, 62,5 µL of a 2M CaCl2 solution (Merck Millipore, #102383) were added plus 500 µL of BES-buffered saline solution (Sigma-Aldrich, #14280). This last solution was added drop-wise while a pipette controller released air bubbles with the help of a glass Pasteur pipette, allowing the formation of DNAs complex to increase de efficiency of transfection. The DNAs solutions were incubated for 30 min at RT. Meanwhile, the medium of 293T cells was replaced by 10 mL of fresh DMEM supplemented with 10% of FBS and 5% of L-Glutamine, without antibiotics. In the end, 1 mL of DNAs solution was added to each plate and the cells were incubated for 24h at 37°C, 5% of CO_2. After this incubation period, 293T GFP positive cells were observed by fluorescence microscopy in the positive controls. The medium was replaced by 4 mL of complete DMEM and cells were incubated

at 32°C, 5% of CO2. Supernatants containing the lentivirus were collect at 36, 48 and 60 hours (h) being kept at 4°C between collections. After 3rd collection, the supernatants were filtered with a 0,45µm low-protein binding filter (Corning Life Sciences, #513-3344) and kept at 4°C until use. For a maximum transduction efficiency, viruses were used fresh for the direct reprogramming experiments.

For FLAG-Gata2 deletion constructs immunoprecipitation experiments, no viral production was performed. DNA complexes were prepared with 10 µg of each combination (one FLAG-Gata2 deletion construct plus HA-Gfi1b and cFos) and sterile ultrapure water. CaCl2 solution and BES buffer were also added. The 1mL of DNAs solution was added to each plate and incubated for 48h at 37°C, 5% of CO_2.

2.5 Lentiviral transduction of DT MEFs

DT MEFs were plated in 0,1% (m/v) gelation-coated 6-well plates at a density of 3x105 cells per plate. Cells were transduced twice with a 24 h interval. To transduce, the medium was replaced by the produced lentiviruses supplemented with 8 µg/mL of polybrene (Sigma Aldrich, #TR-1003-G) that helps the integration of viruses into the cells. Th first transduction (day 0) lasted 16 h and then the viruses were removed and replaced by complete DMEM allowing cells recovery. For the second transduction (day 1), the viruses were added again for another 16 h. During the transduction protocol, the cells were kept at 37°C, 5% of CO2. After the second transduction, viruses were removed and 2mL of MyeloCult medium with Penicillin-Streptomycin antibiotic and hydrocortisone was added (day 2). To unveil the role of LSD1 during the direct reprogramming, inhibitors of this protein (RN-1 and GSK-LSD1) were added at day 2 or day 6 to the medium at different concentrations: 0, 1, 10, 100, 500 and 1000 nM. The inhibitors were dissolved in DMSO. As a control, 2 µL of DMSO was added to the medium ensuring that control cells (0nM) were exposed just to 0,1% of DMSO.

Medium with inhibitors was changed every 3-4 days until the last day of the reprogramming experiments (day 20). At day 20, cells were analysed by flow cytometry. Throughout the

experiment, the morphological changes and reporter activation of the cells were monitored by brightfield and fluorescence microscopy.

2.6 Flow Cytometry

2.6.1 hCD34/H2BGFP reporter activation analysis

To analyse reporter activation by GFP expression, transduced DT MEFs cells were washed with PBS and detached with TrypLE Express for 10 min at 37°C. The cells resuspended with PBS 2% (v/v) FBS to inactivate the trypsin and centrifuged at 1200 rpm for 5 min. The supernatant was discarded, and the pellets were resuspended in 150 µL of PBS 2% (v/v) FBS and kept on ice. Cell suspensions were filters into a cell strainer tube (BD Biosciences). To analyse cell viability, 1 µL of 4′,6-Diamidine-2′-phenylindole dihydrochloride (DAPI) (1 mg/mL) (D1306, Invitrogen) was added to the sample to exclude dead cells. GFP was measured in 530/30 filter and DAPI in 450/40 filter in BD Aria III (BD Biosciences). Flow cytometry results were analysed in the FlowJo Software.

2.7 Imaging acquisition

Bright field and fluorescent images of cells from reprogramming experiments were obtained with an inverted microscope (Zeiss AxioVert 200M). Images were processed with ZEN 2.3 software (Zeiss Microscopy) and Adobe Illustrator CS5.

2.8 Protein-protein interaction

2.8.1 Immunoprecipitation (IP) and Co-IP

Whole cell extracts were prepared from 293T cells transfected for 48h with each FLAG-Gata2 deletion construct plus HA-Gfi1b and cFos or with FLAG-Gata2, LSD1 and CoREST1. The protein G Agarose beads (Sigma Aldrich, #11719416001) were washed four times with 0,5% (m/v) BSA in

PBS (blocking solution) and incubated with 3 µg of antibody (Table 9) for at least 6h at 4°C with continuous rotation. After incubation, beads were washed again to remove the excess of antibody. Cell extracts were resuspended in lysis buffer (50 mM HEPES (pH 7.6), 250 mM NaCl, 0.1% NP-40, 0.2 mM EDTA) complemented with 1mM of PMSF (Sigma Aldrich, #10837091002) and cOmplete EDTA-free 1X (Sigma Aldrich, #11873580001) protein inhibitors and incubated on ice for 20 min. The samples were centrifuged at 13000 rpm for 30 min to remove cell debris and 10% of the supernatant was kept as input sample. The rest of the sample was added to the beads and left incubate overnight at 4°C with rotation. Next, the immune complexes were washed and SDS sample buffer was added. The sample was heated at 95°C for 5min. Short spin was performed to separate the protein from the beads. Supernatants with the protein were collected and analysed by western blot.

2.8.2 Western Blot Analysis

Samples were run into a 4% stacking gel and 10% of resolving gel at 100V for 2 h. The PVDF membranes (Amersham Hybond P 0.45, Sigma Aldrich, #10600023) were methanol activated, making them ready for the transfer.

For immunoblotting, membranes were blocked with 5% milk (NZYTech, MB26001) in TBS-T buffer (10 mM Tris-HCl (pH 7.9), 150 mM NaCl, and 0.05% Tween 20) and incubated with primary antibody (**Table 9**), overnight at 4°C. In the next day, the membranes were washed three times with TBS-T and incubated for 2h at RT with the horseradish peroxidase-conjugated (HRP) secondary antibody (**Table 10**). The membranes were washed once again and subsequently revealed by Clarity Western ECL Substrate (RioRad, #170-5060).

Table 9:Primary antibodies used during the study.

Antibody	Clone	Conjugate	Source	Application
FLAG	M2	-	Sigma Aldrich, F3165-1MG	WB, IP
HA	4C12	-	Abcam, ab9110	WB, IP
FOS	-	-	Santa Cruz Biotechnology, sc-52	
KDM1A/ LSD1	-	-	Abcam, ab37165	WB, IP
CoREST1	-	-	Santa Cruz Biotechnology, sc-135873	WB, IP

Table 10: Secondary antibodies used during the study.

Antibody	Clone	Conjugate	Source	Application
Anti-Mouse	-	HRP	Cell Signaling Technologies, # 7076S	WB
Anti-Rabbit	-	HRP	Cell Signaling Technologies, # 7074S	WB

CHAPTER 3

Results

3. Results

3.1 GATA2 binds to the KDM1A gene locus

Analysis of chromatin histone immunoprecipitation (ChIP) data from our group demonstrate that during hematopoietic reprogramming of human dermal fibroblasts into hematopoietic stem-like cells, GATA2 binds to the *KDM1A* gene (that encodes for LSD1) promoter. This binding was observed when Gata2 was expressed alone or in the presence of GFI1B and FOS, suggesting that LSD1 expression may be controlled by GATA2 during the process and therefore important for successful hemogenic reprogramming (**Figure 8**).

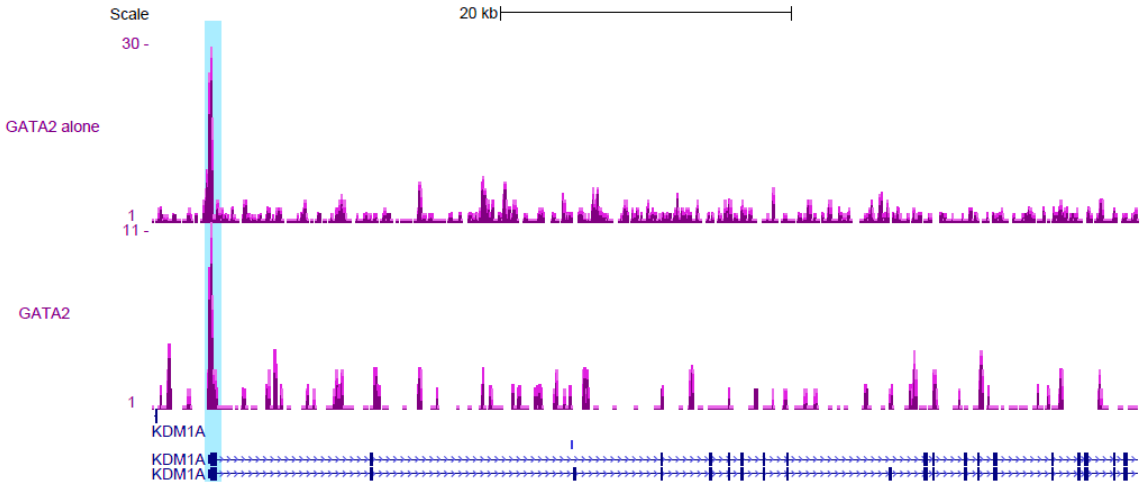


Figure 8: GATA2 binds to the *KDM1A* gene locus. Gata2 binding profile at the *KDM1A* when expressed alone (upper panel) or in the presence of GFI1B and FOS (bottom panel). The y-axis represents the number of reads detected by Chip-sequencing for the Gata2 protein. A peak located at the promoter region of Gata2 is highlighted by the blue box. KDM1A gene structure and the direction of transcription is shown.

3.2 Gata2 interacts with Gfi1b, cFos and the LSD1/CoREST1 complex with specific domains.

Gata2 has been shown to interact with LSD1 at the protein level (Guo et al., 2016). It would be important to define which Gata2 domains are required to interact with LSD1/CoREST1. For that, HEK293T cells were co-transfected with each pcDNA3.0-FLAG-Gata2 deletion constructs (Minegishi et al., 2005), pTRE-Bio-TEV-AOF2 and pTRE-Bio-CoREST1 plasmids, provided by Dr. Sjaak Philipsen, to overexpress each FLAG-Gata2 deletion constructs plus LSD1 and CoREST. After, overexpression of each FLAG-Gata2 deletion construct with HA-Gfi1b and cFos was also performed to study the cooperation between Gata2, Gfi1b and cFos. After 48h of transfection, pellets of HEK293T cells were collected and analysed by immunoprecipitation followed by western blot (**Figure 9 (A)**). Gata2 and Gfi1b genes were tagged with FLAG and HA, respectively to improve detection. In the case of FLAG-Gata2 constructs, the FLAG tag was crucial for the detection since the deletion of certain domains difficult the recognition of the protein by the antibody (**Figure 9 (B)**). As a control of transfection, HEK293T cells were transfected with pHAGE2-mCherry (**Figure 9 (C)**).

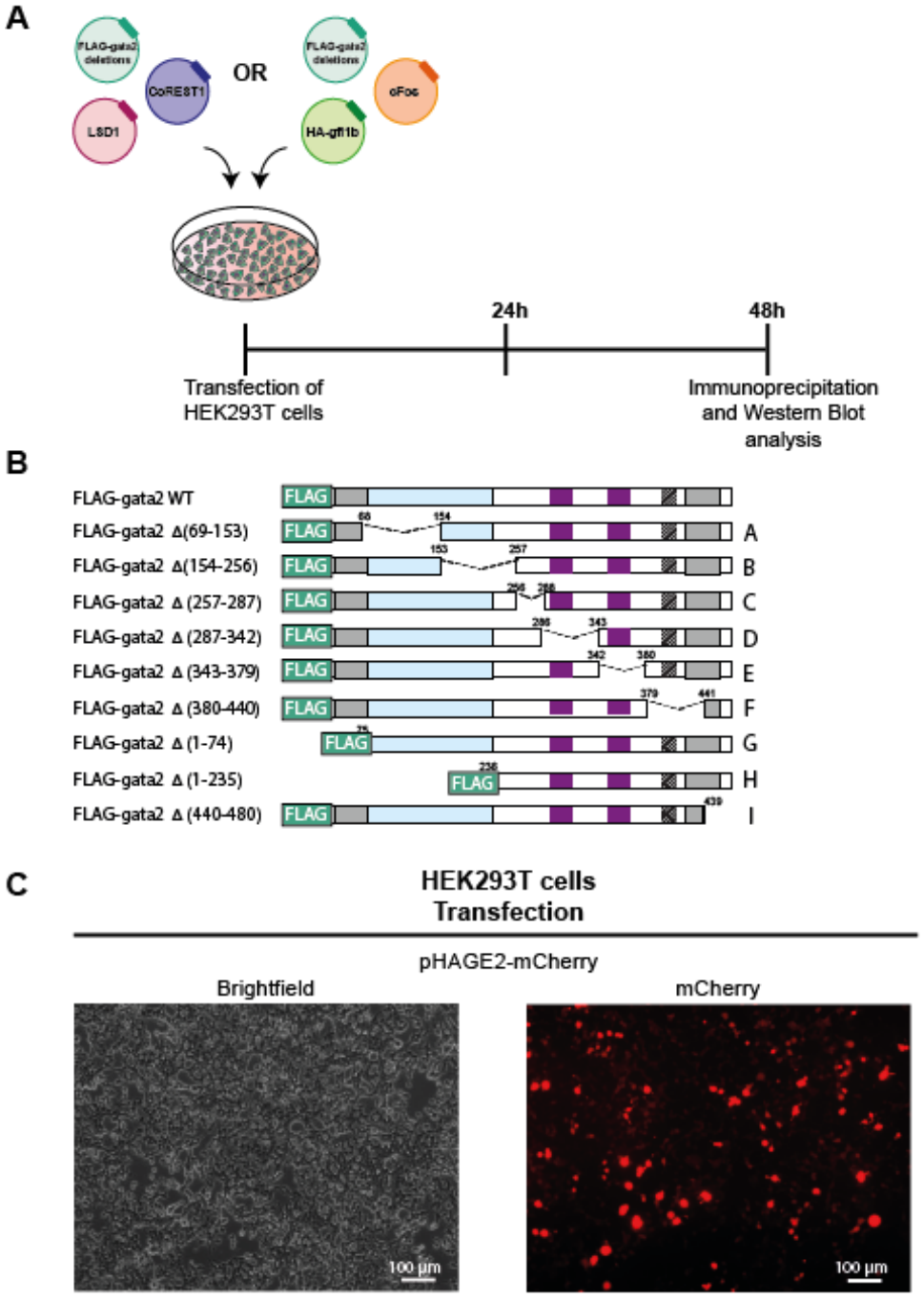


Figure 9: Investigating Gata2 domains required for interaction with Gfi1b, cFos and the LSD1/CoREST complex. (A) Scheme of the overexpression of different proteins in HEK293T cells. **(B)** Schematic representation of the protein domain organization of Gata2 wild type (WT) and Gata2 deletion mutants. Deleted amino acids are indicated between parentheses. Δ stands for "deletion". The letters A to I correspond to the construct that precedes them. **(C)** Control of the transfection of HEK293T cells with pHAGE2-mCherry. Snapshots were taken under rhodamine filter.

3.2.1 Gata2 C-terminal zinc finger is necessary for the interaction between Gata2 and LSD1/CoREST1 complex

In order to define the interactions between Gata2/LSD1/CoREST1 complex, an immunoprecipitation was performed using FLAG-Gata2 deletion constructs. As shown in **Figure 10**, there is no interaction between LSD1 and the CoREST1 with the FLAG-Gata2 deletion E and FLAG-Gata2 deletion F. In the deletion E, the C-terminal zinc finger is absent decreasing the efficiency of DNA binding. The absence of the nuclear localization sequence (NLS) domain in the FLAG-Gata2 deletion F interfere with the translocation of Gata2 to the nucleus and its interaction with nuclear LSD1. These same deletions impact the interaction with CoREST1 probably because LSD1 and CoREST1 act as a complex. In the input samples, all the proteins were detected, therefore, the lack of interactions in these two specific deletions, is due to the fundamental role that the missing domains play in this interaction.

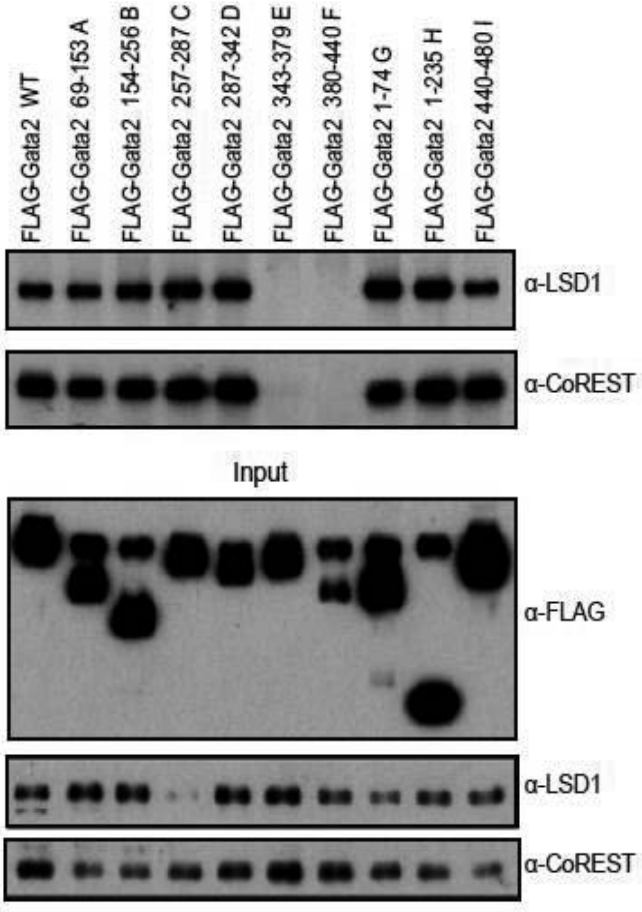


Figure 10: Interaction between Gata2 and LSD1/CoREST1 requires the C-terminal zinc-finger. Immunoblots showing the pulldown of FLAG-Gata2 and co-immunoprecipitation of LSD1 and CoREST1 (upper panel) in HEK293T cells after 48h transfection with each FLAG-Gata2 deletion constructs, LSD1 and CoREST1. Immunoblots of the 10% of total lysate (Input, bottom panel).

3.2.2 Distinct interaction mechanisms between Gata2 and additional hemogenic transcription factors

Since Gata2, Gfi1b and cFos are the three transcription factors necessary for the hemogenic reprogramming process (Pereira et al., 2013), more information about the interaction between these proteins can offer new strategies to increase the efficiency of the reprogramming.

Co-immunoprecipitation experiments were performed to identify Gata2 domains required to form a complex with Gfi1b and cFos (**Figure 11**). Regarding the interaction with Gfi1b (HA-Gfi1b), this TF can interact with all FLAG-Gata2 deletion constructs demonstrating that these proteins possibly interact indirectly. Despite this, the interaction in the deletion F, where the NLS domains is deleted, is absence because since Gata2 is not translocated to the nucleus. Since cFos was also overexpressed in these experiments, its interaction with certain Gata2 domains was also evaluated. The lack of interaction in the FLAG-Gata2 deletion D and I show that the N-terminal zinc finger and, more importantly, the C-terminal transactivation domain (TAD) of Gata2 is important for the cooperation with Gata2 (**Figure 11**). These results underscore the importance of the characterization of this reprogramming complex in more detail to design strategies to improve hemogenic reprogramming.

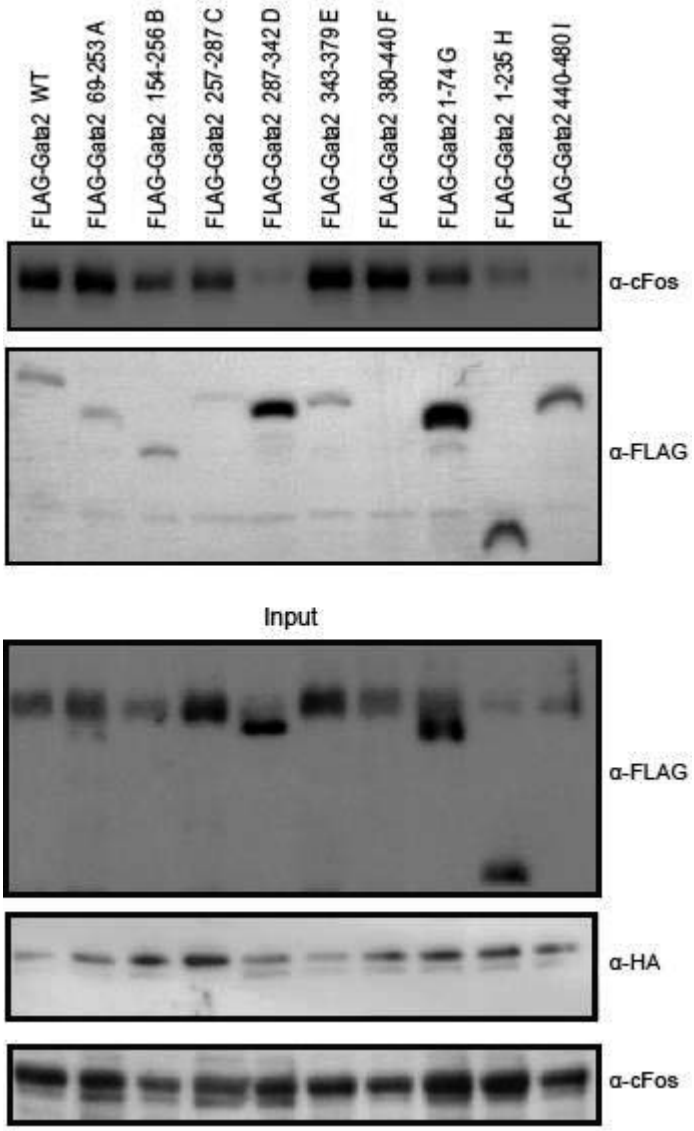


Figure 11: Interaction between Gata2 and Gfi1b requires the NLS domain. N-terminal zinc-finger and the C-TAD domain of Gata2 are necessary to interact with cFos. Immunoblots showing the pulldown of HA-Gfi1b and co-immunoprecipitation detection of cFos and FLAG-Gata2 (upper panel) in HEK293T cells after 48h transfection with each FLAG-Gata2 deletion constructs, HA-Gfi1b and cFos. Immunoblots of the 10% of total lysate (Input, bottom panel).

3.3 *Kdm1a* expression does not change along the hemogenic reprogramming process

The results above demonstrate that the three TFs interact with each other but also with epigenetic regulators such as LSD1/CoREST1 complex. As Gata2 binds to the promoter region of *KDM1A* (Figure 8) we asked whether the transcript levels of *Kdm1a* are differentially regulated during EHT and hemogenic reprogramming. Using RNA sequencing data from our lab, the expression levels of the *Kdm1a* gene were evaluated in human and mouse fibroblasts during the reprogramming process and during EHT in the mouse placenta (**Figure 12**).

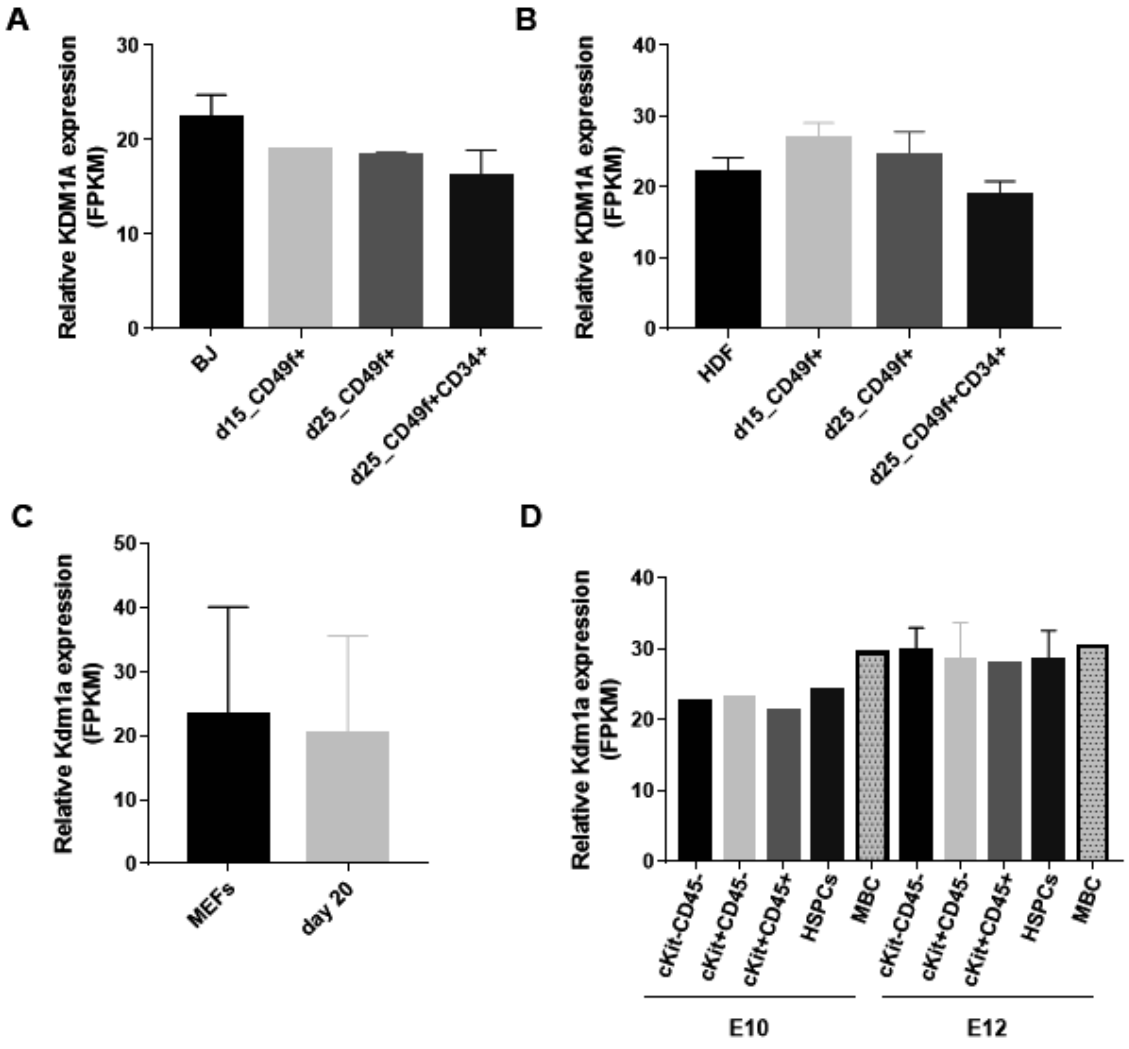


Figure 12: Expression of *Kdm1a* during hematopoietic reprogramming in human and mouse does not change. RNA sequencing data analysis of *Kdm1a* relative expression levels in **(A)** neonatal foreskin fibroblast (BJ), **(B)** human dermal fibroblasts (HDF) at day 15 and 25 of the hemogenic reprogramming and in **(C)** mouse embryonic fibroblast (MEFs) at day 20 of reprogramming. **(D)** Expression levels of *Kdm1a* gene in cKit-CD45-, cKit+CD45+ and cKit+CD45+ cells from mouse placentas with different stages of gestation (E10 and E12). Hematopoietic stem and progenitor cells (HSPCs) and mature blood cells (MBC) were used as controls.

To obtain this data, BJs and HDFs were transduced with GATA2, GFI1B and FOS and analysed after 25 days of reprogramming. Cells expressed the markers CD49f and CD34 indicating a hematopoietic stem-like cell. Analysing the relative expression of *KDM1A* gene of fibroblast through hematopoietic cell fate, the levels do not change significantly (**Figure 12 (A)(B)**). In mouse, *Kdm1a* expression is not different in MEFs and after 20 days of reprogramming (**Figure (C)**). Using mouse placentas with different stages of gestation (E10 and E12), the transcriptional profile of Prom1+Sca+CD34+CD45- (PS34CD45-) cells subsets (cKit-CD45-, cKit+CD45+ and cKit+CD45+) was evaluated and compared with hematopoietic stem and progenitor cells (HSPCs) and with mature blood cells (MBC) as controls. In this case, the expression levels of the gene *Kdm1a* does not change significantly. This data shows that in human and mouse reprogramming or in mouse *in vivo*, the expression levels of the gene that encodes LSD1 do not vary significantly, suggesting that the function of LSD1 during hematopoiesis or hematopoietic reprogramming is regulated through its catalytic function.

3.4 Abolishment of reprogramming by pharmacological inhibition of LSD1

To evaluate the role of LSD1/CoREST complex during the reprogramming process of MEFs into HSCs, pharmacological inhibition of this complex was performed. Two inhibitors were used with the same concentration range and experimental design. Double transgenic (DT) MEFs were transduced with Gata2, Gfi1b and cFos (**Figure 13 (A)**) and cultured in Myelocult medium. At day 2 or at day 6 media was supplemented with the LSD1 inhibitors RN-1 and GSK-LSD1. Media was replaced every 4 days and inhibitors were added fresh to maintain the initial concentration of 0nM, 1nM, 10nM, 100nM, 500nM and 1000nM in both inhibitors. At day 20, GFP quantification, as a measure of huCD34 reporter activation and reprogramming, was performed by flow cytometry. (**Figure 13 (B)**).

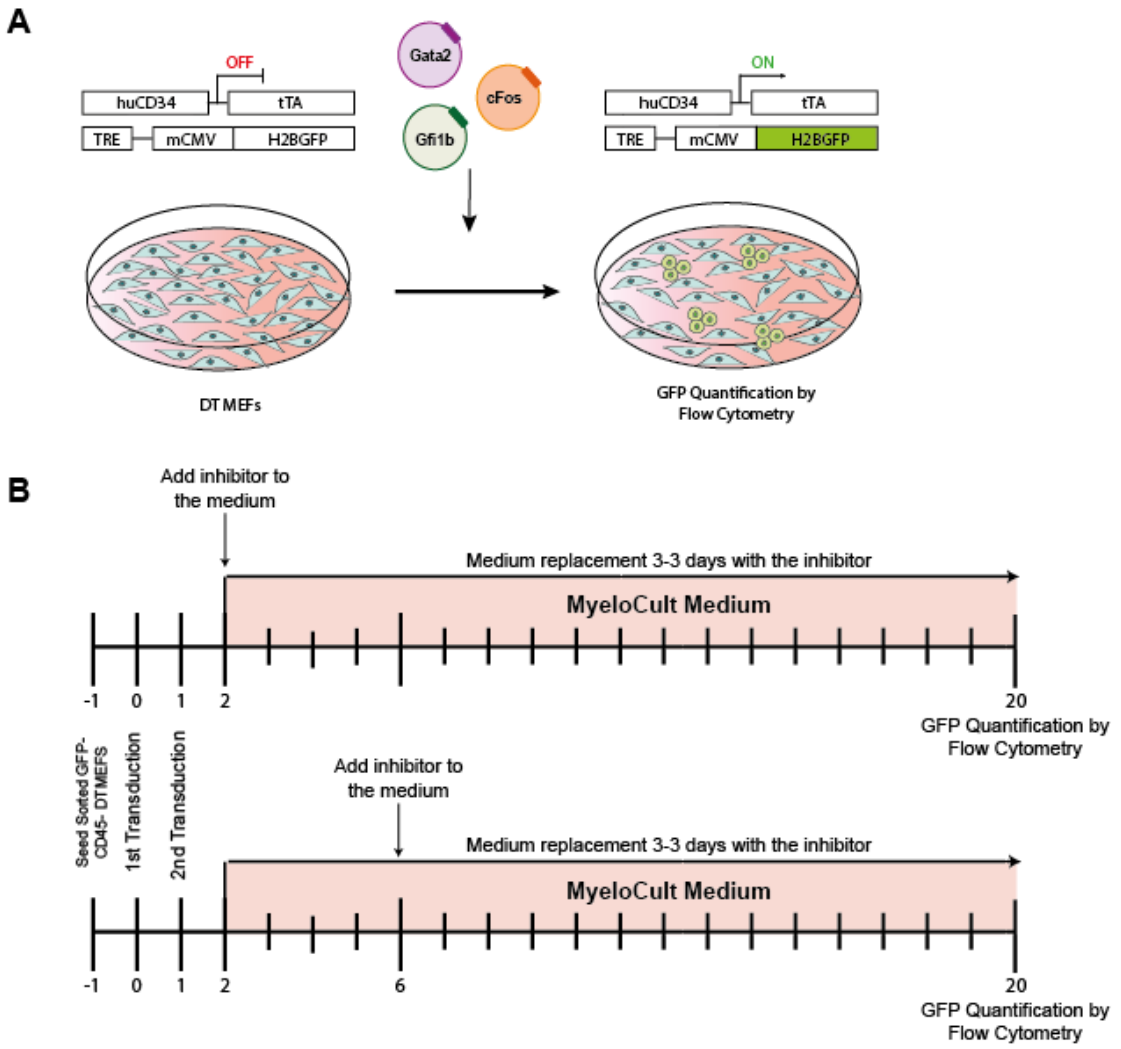


Figure 13: Experimental design of the pharmacological inhibition of LSD1 during DT MEFs hematopoietic reprogramming. **(A)** The acquisition of the hematopoietic cell fate by the overexpression of Gata2, Gfi1b and cFos leads to an activation of the huCD34 reporter and consequently GFP activation. **(B)** Scheme of the timeline used during the experiments. The addition of the inhibitor with different concentrations was made at day 2 or day 6 of reprogramming. All conditions were analysed by flow cytometry at day 20.

3.4.1 LSD1 inhibition with RN-1 hydrochloride inhibits hemogenic reprogramming

To inhibit LSD1, RN-1 hydrochloride was used and added at day 2 or 6 (**Figure 14 (A)** and **(B)**, respectively). In both conditions, when only 1nM of RN-1 was added to the medium, the percentage of GFP+ cells decrease 5-fold. The major differences are noticed in the conditions of 10nM and 100nM, the percentage of GFP+ is lower in the day 2. To study the effect of LSD1 inhibition with RN-1, a viability assay was performed to verify the rates of cell death. As presented in **Figure 14 (C)** and **(D)**, the percentage of live cells was always higher than 80%, demonstrating that RN-1 do not increase cell death even in the highest concentrations and the decrease of GFP+ cells due to the inhibition of LSD1 and consequently inhibition of the reprogramming process.

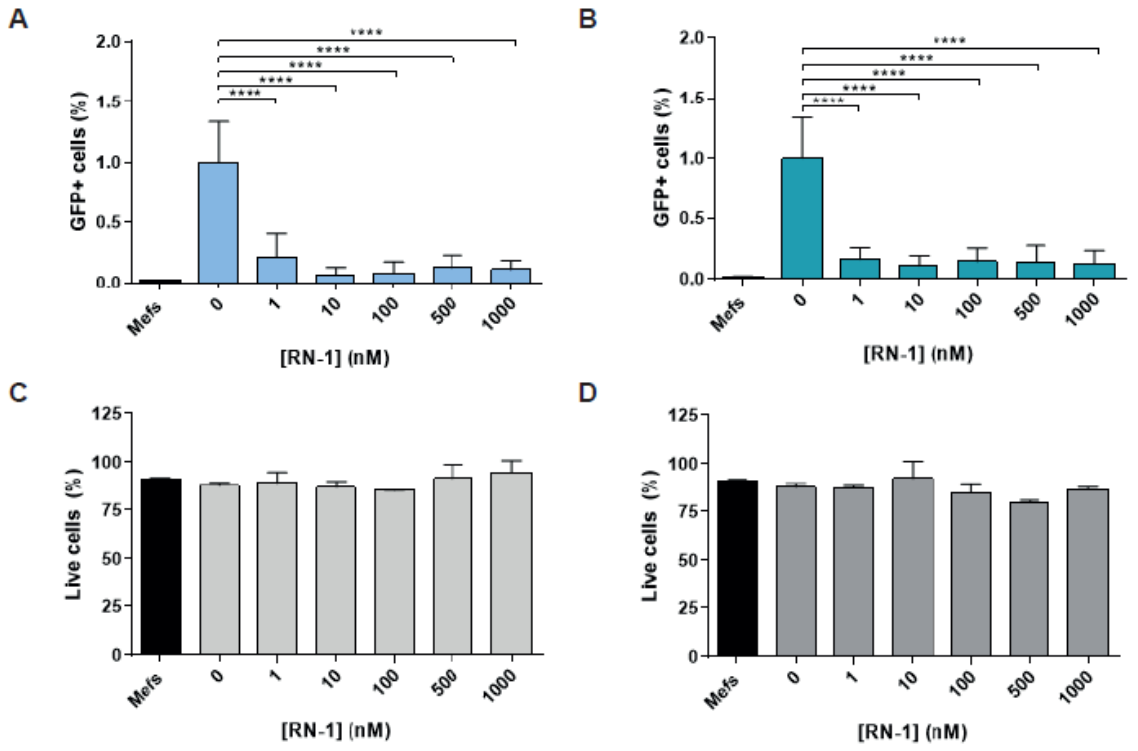


Figure 14: Pharmacological inhibition of LSD1 with RN-1 impairs the hemogenic reprogramming. **(A)** Analysis of the percentage of GFP+ cells by flow cytometry after the addition of different concentrations (1, 10, 100, 500 and 1000 nM) of the inhibitor RN-1 at day 2 of reprogramming or **(B)** at day 6. **(C)** Analysis of the percentage of live cells by flow cytometry using DAPI staining when RN-1 was added at day 2 or **(D)** at day 6. Mean ± SD; statistical analysis using one-way ANOVA with Bonferroni's multiple comparison test; ****p<0.0001.

3.4.2 LSD1 inhibition with GSK-LSD1 inhibits hemogenic reprogramming

Towards validation of the results obtained with the inhibitor RN-1 hydrochloride, another inhibitor was used - T-3775440 also known as GSK-LSD1- with the same concentrations and experimental conditions. In this experiment, when 1nM of GSK-LSD1 was added, the percentage of GFP+ cells decreased 2-fold at day 2 and also at day 6 (**Figure 15 (A)** and **(B)**, respectively).

The major differences observed between the use of RN-1 and GSK-LSD1 is that RN-1 seems to be more effective to inhibit LSD1 despite both decrease the percentage of reprogrammed cells. When using this inhibitor, the concentration of 500nM at day 2 and the concentration of 10nM at day 6 can mimic completely the effect of RN-1. In all conditions, when GSK-LSD1 is added, there is a decrease of the GFP+ cells suggesting that when LSD1 is inhibited by a different mechanism, hemogenic reprogramming is also affected. In this experiment, the percentage of live cells was also evaluated and was always higher than 80%, demonstrating that GSK-LSD1 do not increase cell death (**Figure 15 (C)** and **(D)**). This result confirms that the inhibition of LSD1 leads to a decrease of the hematopoietic reprogramming of DT MEFs in an identical way as RN-1. This point will be further discussed in **Chapter 4.**

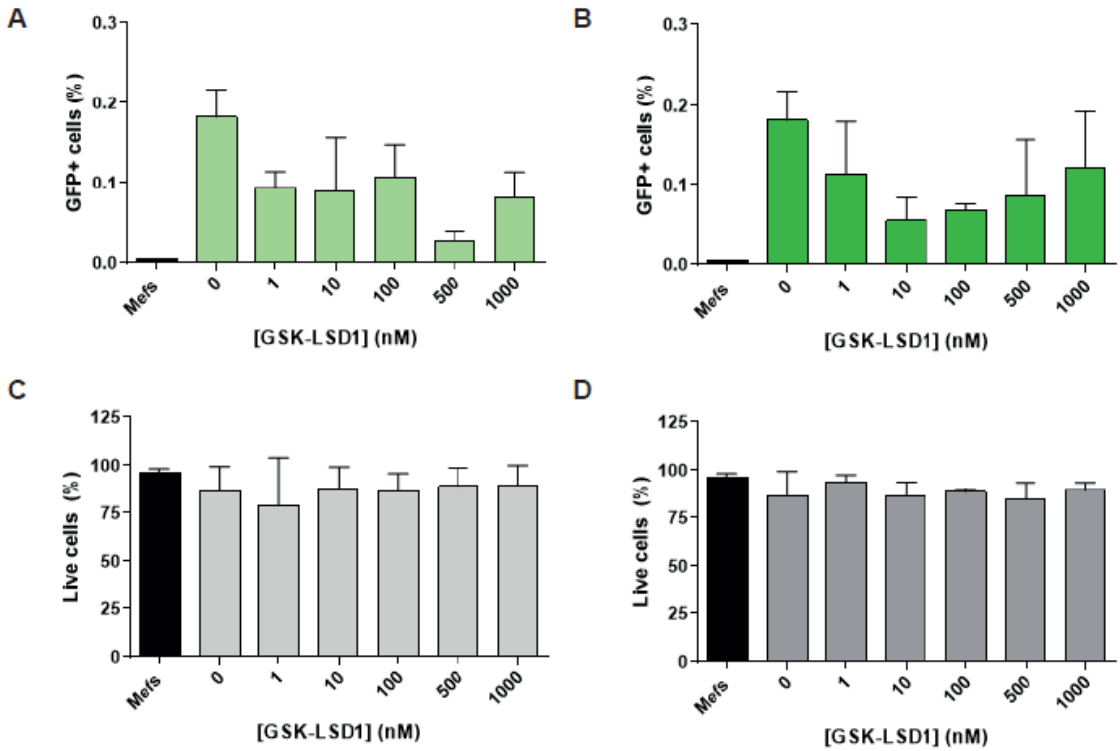


Figure 15: Pharmacological inhibition of LSD1 with GSK-LSD1 impairs the hemogenic reprogramming. **(A)** Analysis of the percentage of GFP+ cells by flow cytometry after the addition of different concentrations (1, 10, 100, 500 and 1000 nM) of the inhibitor GSK-LSD1 at day 2 of reprogramming or **(B)** at day 6. **(C)** Analysis of the percentage of live cells by flow cytometry using DAPI staining when RN-1 was added at day 2 or **(D)** at day 6.

3.5 Reprogramming DT MEFs using Gata2 with repressing or activating functions

3.5.1 PCR-based cloning strategy of fusion proteins of KRAB, VP16 or VP64 with Gata2 into a pHAGE2 plasmid

Since Gata2 showed to have an independent and strong targeting function during the reprogramming of HDFs into HSCs (Gomes et *al.*, in revision), a better understanding of Gata2 function will provide more information to improve de efficiency of all reprogramming process.

To unveil Gata2 function during reprogramming, a KRAB domain or a VP16/VP64 domain were fused in the 5'end of Gata2 generating a more repressing (KRAB-Gata2) or activating (VP16-Gata2 or VP64-gata2) transcription factor. To clone these fusion proteins, NotI restriction site was added to the forward primer of each domain and a XBaI restriction site was added to the reverse primer of Gata2 allowing the insertion into a pHAGE2 vector. BstBI restriction site allowed the link between the two sequences. KRAB, VP16, VP64 and Gata2 sequences were amplified by PCR and obtained by electrophoresis separation (**Figure 16 (A)**). After purification and restriction with the proper enzymes of pHAGE2 plasmid and inserts, the DNA sequences were ligated and transformed with competent bacteria. Colonies were submitted to a colony PCR to verify the positive ones. To make sure that both DNA segments were inserted into the vector, two colony PCRs were performed for each fusion protein using different positive controls (**Figure 16 (B) (C) (D)**).

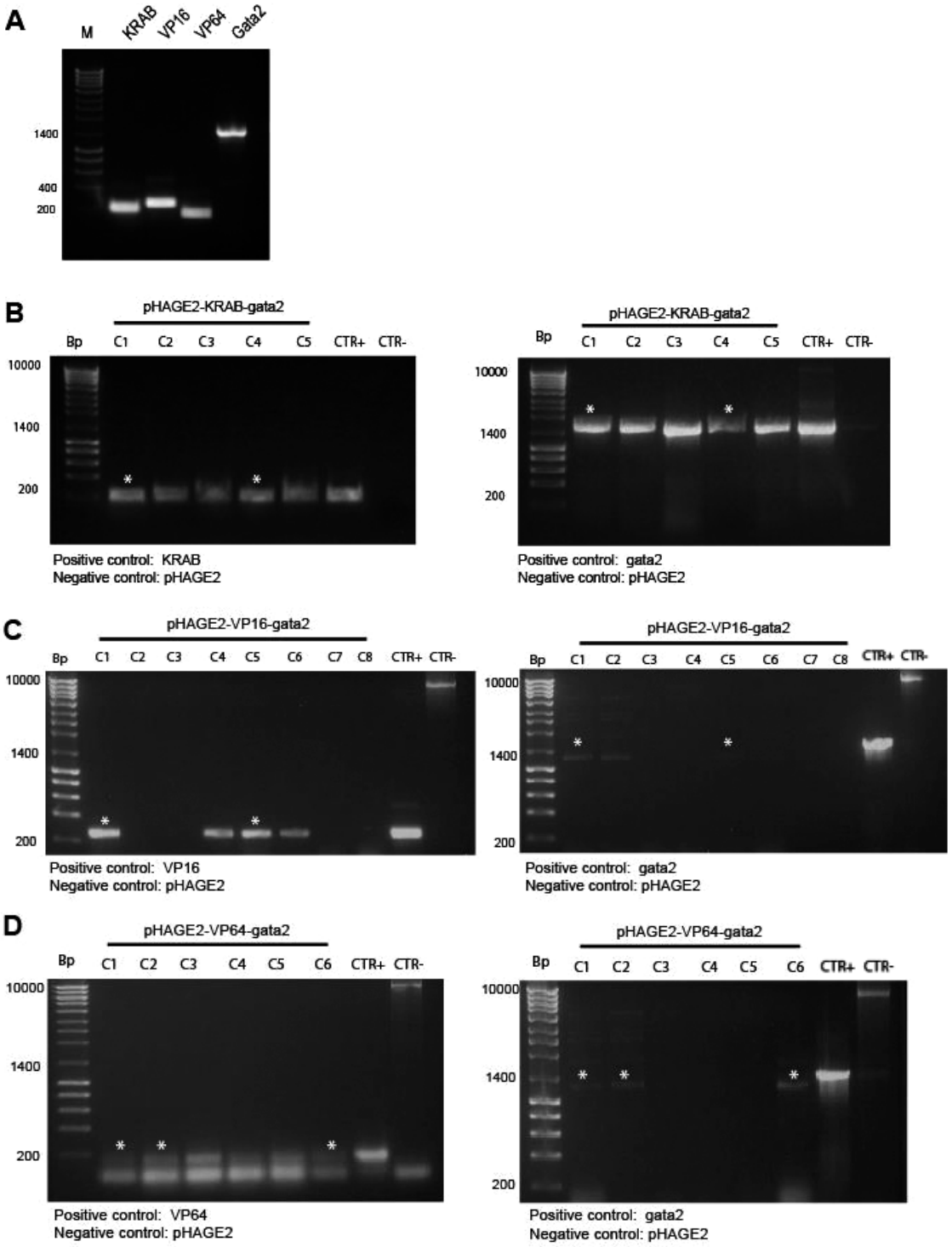
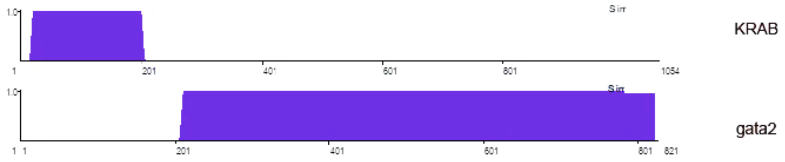
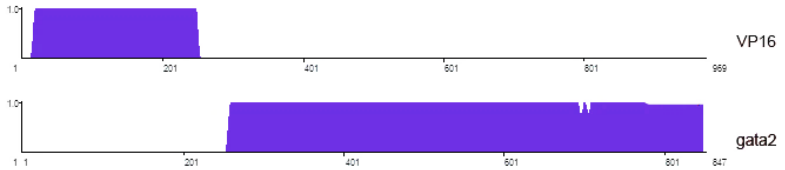
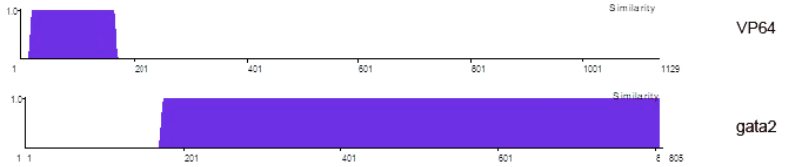


Figure 16: Construction of pHAGE2 with Gata2 fusion proteins with KRAB, VP16 and VP64 domains and screening of positive colonies by Colony PCR. **(A)** PCR amplification of KRAB, VP16, VP64 and Gata2 inserts for later restriction, ligation and competent bacteria transformation. **(B) (C) (D)** Colony PCR of pHAGE2-KRAB-Gata2, pHAGE2-VP16-Gata2 and pHAGE2-VP64-Gata2 positive colonies, respectively. Two positive controls were performed, one with the template vector of KRAB, VP16 or VP64 and the other with Gata2 sequence (CTR+). As a negative control, pHAGE2-MCS empty vector was used (CTR-). Stars (*) indicates the positive colonies sent for sequencing.

The positive colonies of Gata2 fusion proteins were sent for sequencing with the EF1α sequencing primer and the results were aligned with KRAB, VP16, VP64 and Gata2 sequence (**Table 11**). All the Gata2 fusion proteins were successfully cloned into pHAGE2 vector. More details about the cloning methodology are in the **Chapter 2**.

Table 11: Sequencing data of Gata2 fusion constructs cloned into the pHAGE2-MCS vector. Sanger sequencing results were aligned with the sequences of KRAB, VP16, VP64 and Gata2. Quality of the alignment is shown in purple. Irregularities were verified for sequencing errors.

Cloning product	Sequencing data
pHAGE2-KRAB-gata2	
pHAGE2-VP16-gata2	
pHAGE2-VP64-gata2	

3.5.2 Investigating activating and repressing transcription factor function during hemogenic reprogramming

After cloning, the strategy used for the direct reprogramming of DT MEFs with the Gata2 fusion proteins, Gfi1b and cFos (2TF) is outlined in the **Figure 17**. In this reprogramming experiment, Gata2 was replaced by KRAB-Gata2, VP16-Gata2 or VP64-Gata2 (represented in the **Figure 17 (A)** as Gata2 fusion), reconstituting different reprogramming conditions. After the transductions, Myelocult medium was added at day 2 and the percentage of GFP+ cells was measure by flow cytometry at day 20 (**Figure 17 (B)**). The condition Gata2, Gfi1b and cFos (3TFs) was used as a reprogramming control.

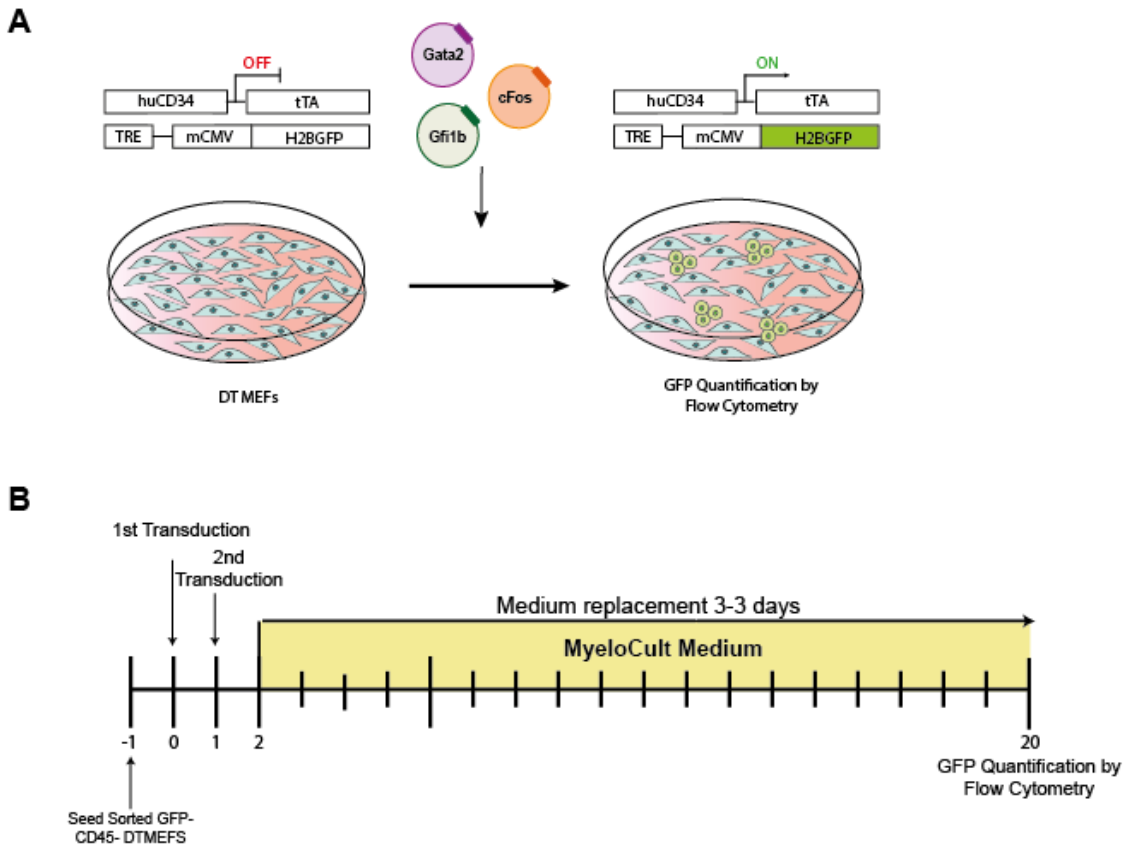


Figure 17: Experimental design to evaluate the impact in the DT MEFs reprogramming using Gata2 with an activating or repressing function. (A) The acquisition of the hematopoietic cell fate by the overexpression of Gata2 fused with KRAB, VP16 or VP64 domains, Gfi1b and cFos leads to an activation of the huCD34 reporter and consequently GFP activation. **(B)** Scheme of the timeline used during the experiments. The replacement of Gata2 by the different Gata2 fusion proteins was evaluated by flow cytometry at day 20.

After 20 days of reprogramming, the different reprogramming conditions were analysed by flow cytometry and the result is showed in **Figure 18**. Although no reprogramming condition resembles the 3TFs condition, the VP64-Gata2+2TF is the condition that can best mimetize unlike the condition that uses the KRAB-Gata2 to replace Gata2. Comparing the condition that uses VP16-Gata2 and the condition VP64-Gata2+KRAB-Gata2, the effect seems to be alike. The use of the repressive Gata2 (KRAB-Gata2) leads to a decrease of 2-fold of the percentage of GFP+ cells.

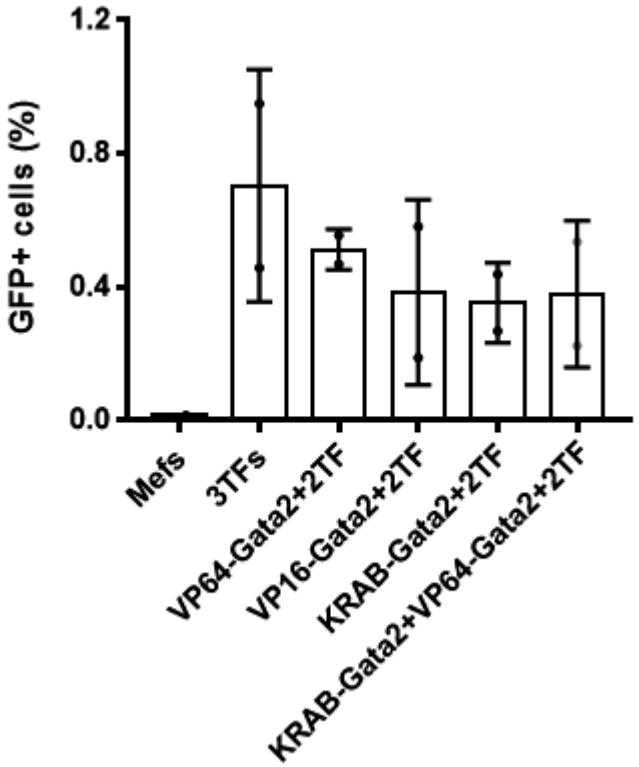


Figure 18: No Gata2 fusion protein with KRAB, VP16 or VP64 can replace Gata2 during hemogenic reprogramming. Quantification of GFP+ cells by flow cytometry at day 20 of MEFs transduced with Gfi1b, cFos (2TF) plus Gata2 (3TF), VP64-Gata2 (VP64-Gata2+2TF), VP16-Gata2 (VP16-Gata2+2TF) or KRAB-Gata2 (KRAB-Gata2+2TF).

CHAPTER 4

Discussion and Future Perspectives

4. Discussion

After Yamanaka's findings in 2012, the cell reprogramming emerged as a powerful tool in the regenerative medicine. The idea of obtaining pluripotent cells, capable of generating all cell types, from somatic cells increased the interest of many researchers to obtain different cell types in order to create a more personalized cellular therapy. The manipulation of the cell fate with direct reprogramming technologies by overexpressing TFs allowed the generation of different cell types such as cardiomyocytes, neurons, hepatocytes, adipocytes and hematopoietic cells, mainly from fibroblasts.

Regarding the hematopoietic field, the overexpression of three TFs – Gata2, Gfi1b and cFos – showed to be essential for the direct reprogramming of fibroblasts into hematopoietic stem-like cells, both in human and mouse. This discovery revealed a possible and promising alternative for the conventional therapies. However, the generation of transplantable HSCs efficiency remains too low to be considered a solid therapy. Clarifying the underlying mechanisms involved in the generation of reprogramming-derived HSCs will allow the creation of new strategies to improve the process's efficiency.

During hematopoietic development, the recruitment of LSD1/CoREST complex by the transcriptor factors such as Gata2 and Gfi1b for epigenetic regulation was already described (Guo et al., 2015, 2016; Saleque et al., 2007; Thambyrajah et al., 2015). Taking into account the similarities between the hematopoiesis and the hemogenic reprogramming in mouse (Pereira et al., 2016), this study reveals the essential role of LSD1/CoREST complex and its interactions with the Gata2, Gfi1b and cFos during the hemogenic reprogramming.

To verify the presence of LSD1/CoREST during the reprogramming, ChIP-seq data from our group was analysed and revealed that GATA2 alone but also in the presence of GFI1B and FOS binds to the *KDM1A* gene promoter and, thus, regulates de expression of the LSD1. These data suggest that GATA2, as it recruits GFI1B and FOS, promotes the expression of LSD1 creating a reprogramming complex between this epigenetic regulator and the three TFs (**Figure 8**).

Guo et. al suggested that Gata2 recruits LSD1 to regulate the expression of Gata1 during the erythroid differentiation and demonstrated the physical interaction between Gata2 and LSD1 (Guo et

al., 2016). By co-immunoprecipitation analysis, the presence of C-ZnF seems important for the interaction between Gata2 and LSD1/CoREST1 (**Figure 10**). Bates et *al.* identified this zinc finger as essential for DNA binding (Bates et al., 2008) but it can also be responsible for the binding to the promoter of *Kdm1a* gene to encode LSD1. The exclusion of the NLS domain also impaired the interaction of Gata2 with this epigenetic complex since the translocation of Gata2 to the nucleus is compromised. Therefore, NLS and the C-ZnF domains deletions defect the interaction of Gata2 with LSD1 but also with CoREST1 because these two proteins act as a complex. The CoREST1 allows the activation of the demethylase activity of LSD1 and, consequently, when Gata2 recruits LSD1, the CoREST1 is also recruited (Yang et al., 2006).

Characterisation of transcriptional networks in blood stem and progenitor cells demonstrated that Gfi1 represses Gata2 while Gata2 activates Gfi1b suggesting that Gata2 modulates the antagonism between Gfi1 and Gfi1b (Moignard et al., 2013). Gata2 domains needed for the interaction with Gfi1b were not described. The interaction between Gata2 and Gfi1b is impaired in the absence of the NLS domain (FLAG-Gata2 deletion construct F (Δ380-440)). This impairment reduces the translocation of Gata2 protein to the nucleus, thus, impairing the interaction with others nuclear proteins such as TFs. Another point to be aware is that Gfi1b seems to interact with all FLAG-Gata2 deletion constructs, independently of the domain that is missing, raising the hypothesis that exists another protein that may be bridging Gata2 and Gfi1b (**Figure 11**).

There is no interaction between Gata2 and cFos when N-ZnF is deleted (FLAG-Gata2 deletion construct D (Δ287-342)) because, as previously mentioned in Chapter 1, both ZnFs are known to interact separately or collectively with other proteins (Vicente et al., 2012). In this case, the function of N-ZnF is to bind to other proteins such as cFos. The deletion of the C-TAD domain also impaired the cooperation of Gata2 with cFos. Data from our group demonstrated that the C-TAD deletion decreased the hemogenic reprogramming in mouse indicating the importance of this domain for the recruitment of cFos to the reprogramming complex by Gata2 (Alves et al, unpublished) (**Figure 11**). The results above point to the existence of a complex formed by Gata2, Gfi1b, cFos and LSD1/CoREST and it would be interesting to perform the same co-immunoprecipitation experiments in

reprogrammed cells to confirm that these interactions are crucial during the reprogramming process.

Considering the reprogramming process, RNA sequencing data available in our group showed that the expression of the *Kdm1a* gene does not change significantly in both human (BJs and HDFs) and mouse (MEFs) reprogramming. The expression levels were also evaluated in mouse placenta to provide information relative to *in vivo* hematopoiesis and to establish a comparison with the reprogramming. In this case, no changes were also detected. Consequently, there is no alteration of the levels of the *Kdm1a* neither during the human and mouse hematopoietic reprogramming nor *in vivo* mouse hematopoiesis suggesting that the regulation of the LSD1 is catalytically controlled (**Figure 12**).

To test this hypothesis, two well-known LSD1 inhibitors were used during the direct reprogramming of fibroblasts to HSCs to evaluate the catalytic role of this enzyme. The tranylcypromine RN-1 binds covalently to FAD (McGrath et al., 2016). Since LSD1 is FAD-dependent (Culhane and Cole, 2007; Shi et al., 2005), this inhibitor disallows the catalytic activity of LSD1. The minimal addition of RN-1 to the medium (1nM) decreased 5-fold the percentage of GFP+ cells and knowing that the acquisition of the hematopoietic cell fate leads to the expression of GFP, the drastic reduction of GFP+ cells indicates that the hemogenic reprogramming is impaired when the LSD1 is not catalytic active (**Figure 14 (A)(B)**). As the concentration of the RN-1 increases, the effect remains the same meaning that the concentration of 1nM is sufficient to efficiently inhibit LSD1. To evaluate the impact of the LSD1 inhibition during different stages of reprogramming, the inhibitor was added to the day 2 or day 6 of the experiment. No differences were detected between these two timepoints demonstrating that even with four days of reprogramming, LSD1 remains an important player in the reprogramming complex.

To validate the results obtained with the RN-1, the effects of the inhibitor GSK-LSD1 were also evaluated. GSK-LSD1 is known to disrupt the interaction between LSD1 and Gfi1b, two important players of the reprogramming complex, by blocking the SNAG domain of this TF (Takagi et al., 2017). Despite the lower reprogramming efficiency compared with the RN-1 experiments, the use of this inhibitor also decreased the percentage of GFP+ cells by 2-fold at day 2 and also at day 6

(**Figure 15(A)(B)**). One of major differences between these inhibitors is the effect of some concentrations in the inhibition. In the case of RN-1, all the concentrations had a similar effect that efficiently inhibited LSD1 but only the concentrations of 500nM (day 2) and 10nM (day 6) of GSK-LSD1 can mimic RN-1 effect. Regarding the mode of action of GSK-LSD1, at day 2, the concentration of 500nM seems ideal at this point. The reprogramming complex may be in the process of being established and a higher number of Gfi1b-SNAG that is not yet interacting with LSD1 may be inhibited. At day 6, the hemogenic reprogramming is already occurring four days ago and there are fewer SNAG domains that can be inhibited and, thus, a lower concentration can induce an effect.

The percentage of live cells was also evaluated to make sure that these differences were due to the effect of the inhibition of LSD1 and not because of cell death. In both experiments, the percentage of live cells was always higher than 80% proving that the decrease of the GFP+ cells was due to the effect of the LSD1 inhibition during the hemogenic reprogramming and not because of cell death (**Figure 14 (C)(D)**) (**Figure 15 (C)(D)**).

The pharmacological inhibition of LSD1 was achieved with success by both inhibitors despite the higher effect of RN-1 compared with the GSK-LSD1. When RN-1 was used, the percentage of GFP+ cells decreased drastically when compared with the data obtained with GSK-LSD1 possibly due to the mode of action of these drugs. Despite having the same function, they interact differently with LSD1, influencing its function during the reprogramming process in different manners. These experiments showed that when the catalytic function of LSD1 or its interactions with other components of the reprogramming complex are compromised, the hemogenic reprogramming is also negatively affected proving the important catalytic role of LSD1 in the regulation of the hematopoietic reprogramming complex. Since these experiments are a loss of function studies, it would be interesting to perform a gain of function study by overexpressing LSD1 along with Gata2, Gfi1b and cFos. To increase the efficiency of the transduction, a polycistronic plasmid could be used with the three TFs plus LSD1 to allow the transduction of a single plasmid instead of four plasmids.

Our group reported GATA2 as the dominant TF in the hematopoietic reprogramming of HDF into HSCs. In this same study was demonstrated that GATA2 and GFI1B share a cohort of targets in HDFs and in the reprogrammed cells, in other words, they share targets that may be inactivated at

the beginning and other that maybe are activated at the end of the process (Gomes et. *al,* in revision). Since Gfi1b is a repressor TF, the function of Gata2 as an activator or repressor TF is not fully described. To unveil Gata2 function, repressive or activating domains were added to Gata2 as a fusion protein and reprogramming experiments were performed.

After GFP quantification, it would be expected that the condition in which Gata2 is replaced by VP64-Gata2, the percentage of GFP+ cells would increase since VP64 is commonly described as an activator domain. Nevertheless, no conditions resemble the control condition where Gata2, Gfi1b and cFos (3TF) were used. Despite that, the substitution of Gata2 by the VP64-Gata2 appears to be the condition that better mimics the effect of Gata2 and, thus, Gata2 can be acting as an activator. In the condition KRAB-Gata2, there is a decrease of 2-fold and 1.5-fold of the percentage of GFP+ cells compared with the 3TF and VP64-Gata2 conditions, respectively. These results rise the hypothesis that a more repressive Gata2 is not more efficient in the reprogramming process. Nevertheless, the strategy of using domains, especially KRAB, was already applied with other TFs. Juárez-Moreno et. *al* used the KRAB domain fused with Oct4. Surprisingly, the results of this study concluded that the KRAB, traditionally described as a repressive domain, fused with Oct4 appears to be a potent activator (Juárez-Moreno et al., 2013). Therefore, since this result is inconclusive about the repressive or activating function of Gata2 during the hemogenic reprogramming, we can hypothesize that if KRAB is acting as an activator domain, Gata2 is a repressive TFs that is supressing important targets for the generation of HSC-like cells. Concerning the condition VP16-Gata2 and the KRAB-Gata2 plus VP64-Gata2, the percentage of GFP+ cells is similar. VP16 may not be strong enough compared with the VP64, this last one is being used as the main transcriptional activator in techniques such as the CRISPR-dCas9 (La Russa and Qi, 2015). In other hand, the use of a repressive and activating Gata2 at same time probably led to activation and inhibition of the same targets, nullifying the effect of each other (**Figure 18**). Additionally, all the domains were fused in the N-terminal of Gata2 and this strategy maybe impaired the targeting of Gata2 since another DNA binding site was created in addition to the zinc fingers. The distance between these domains could lead to an alteration of the DNA binding conformation of Gata2 weakening the cooperation with the promoters and others TF/ epigenetic regulators such as Gfi1b, cFos and LSD1/CoREST. To improve this strategy,

the domains could be added to the N- and C-terminal as well as near to the DNA binding site of Gata2, in this case, close to the zinc fingers. It will be interestingly to perform the reprogramming experiment with this modified Gata2 to elucidate the role of Gata2.

Collectively, I have established that LSD1/CoREST complex has an important role during the hemogenic reprogramming of fibroblasts into HSCs acting as a complex with the Gata2, Gfi1b and cFos. This study allowed the identification of the Gata2 domains that are important for the interactions with Gfi1b, cFos and LSD1/CoREST. Future experiments to prove that the same interactions are essential for the formation of the reprogramming complex need to be performed. Hence, I have established the concept of "hemogenic complex" which not only contributes to the understanding of the process of HSC cell formation and reprogramming but also provides powerful means for clinical translation of hematopoietic reprogramming.

CHAPTER 5

References

5. References

Adamo, A., Sesé, B., Boue, S., Castaño, J., Paramonov, I., Barrero, M.J., and Belmonte, J.C.I. (2011). LSD1 regulates the balance between self-renewal and differentiation in human embryonic stem cells. Nat. Cell Biol. *13*, 652–661.

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

Angel, P., and Karin, M. (1991). The role of Jun, Fos and the AP-1 complex in cell-proliferation and transformation. Biochim. Biophys. Acta *1072*, 129–157.

Apperley, J., Niederwieser, D., Huang, X., Nagler, A., Fuchs, E., Szer, J., and Kodera, Y. (2016). Haploidentical Hematopoietic Stem Cell Transplantation : A Global Overview Comparing Asia , the European Union , and the United States. Biol. Blood Marrow Transplant. *22*, 23–26.

Aravind, L., and Iyer, L.M. (2002). The SWIRM domain: a conserved module found in chromosomal proteins points to novel chromatin-modifying activities. Genome Biol. *3*, 1–7.

Ballas, N., Grunseich, C., Lu, D.D., Speh, J.C., and Mandel, G. (2005). REST and its corepressors mediate plasticity of neuronal gene chromatin throughout neurogenesis. Cell *121*, 645–657.

Bates, D.L., Chen, Y., Kim, G., Guo, L., and Chen, L. (2008). Crystal structures of multiple GATA zinc fingers bound to DNA reveal new insights into DNA recognition and self-association by GATA. J. Mol. Biol. *381*, 1292–1306.

Batra A. Cottler-Fox M. Harville T. Rhodes-Clark B.S. Makhoul I. Nakagawa M. (2014). Autologous graft versus host disease: An emerging complication in patients with multiple myeloma. Biol. Blood Marrow Transplant. *2014*, S261–S262.

Batta, K., Florkowska, M., Kouskoff, V., and Lacaud, G. (2014). Direct Reprogramming of Murine Fibroblasts to Hematopoietic Progenitor Cells. Cell Rep. *9*, 1871–1884.

Beerli, R.R., Segal, D.J., Dreier, B., and Barbas, C.F. (1998). Toward controlling gene expression at will: Specific regulation of the erbB-2/HER-2 promoter by using polydactyl zinc finger proteins constructed from modular building blocks. Proc. Natl. Acad. Sci. *95*, 14628–14633.

Bernstein, B.E., Meissner, A., and Lander, E.S. (2007). The Mammalian Epigenome. Cell *128*, 669–681.

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

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

Cacchiarelli, D., Trapnell, C., Ziller, M.J., Soumillon, M., Cesana, M., Karnik, R., Donaghey, J., Smith, Z.D., Ratanasirintrawoot, S., Zhang, X., et al. (2015). Integrative Analyses of Human Reprogramming Reveal Dynamic Nature of Induced Pluripotency. Cell *162*, 412–424.

Campbell, K.H.S., McWhir, J., Ritchie, W.A., and Wilmut, I. (1996). Sheep cloned by nuclear transfer from a cultured cell line. Nature *380*, 64–66.

Capellera-Garcia, S., and Flygare, J. (2017). Direct lineage reprogramming: a useful addition to the blood cell research toolbox. Expert Rev. Hematol. *10*, 107–109.

Capellera-Garcia, S., Pulecio, J., Dhulipala, K., Siva, K., Rayon-Estrada, V., Singbrant, S., Sommarin, M.N.E., Walkley, C.R., Soneji, S., Karlsson, G., et al. (2016). Defining the Minimal Factors Required for Erythropoiesis through Direct Lineage Conversion. Cell Rep. *15*, 2550–2562.

Cheng, H., Ang, H.Y.K., El Farran, C.A., Li, P., Fang, H.T., Liu, T.M., Kong, S.L., Chin, M.L., Ling, W.Y., Lim, E.K.H., et al. (2016). Reprogramming mouse fibroblasts into engraftable myeloerythroid and lymphoid

progenitors. Nat. Commun. *7*.

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

Culhane, J.C., and Cole, P.A. (2007). LSD1 and The Chemistry of Histone Demethylation Jeffrey. Curr. Opin. Chem. Biol. *11*, 561–568.

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

Doulatov, S., Notta, F., Laurenti, E., and Dick, J.E. (2012). Hematopoiesis: A human perspective. Cell Stem Cell *10*, 120–136.

Dupont, C., Armant, D.R., and Brenner, C.A. (2009). Epigenetics: Definition, Mechanisms and Clinical Perspective. Semin. Reprod. Med. *27*, 351–357.

Duteil, D., Metzger, E., Willmann, D., Karagianni, P., Friedrichs, N., Greschik, H., Günther, T., Buettner, R., Talianidis, I., Metzger, D., et al. (2014). LSD1 promotes oxidative metabolism of white adipose tissue. Nat. Commun. *5*.

Dzierzak, E., and Speck, N.A. (2008). Of lineage and legacy: The development of mammalian hematopoietic stem cells. Nat. Immunol. *9*, 129–136.

Feng, R., Desbordes, S.C., Xie, H., Tillo, E.S., Pixley, F., Stanley, E.R., and Graf, T. (2008). PU.1 and C/EBPalpha/beta convert fibroblasts into macrophage-like cells. Proc. Natl. Acad. Sci. U. S. A. *105*, 6057–6062.

Foudi, A., Kramer, D.J., Qin, J., Ye, D., Behlich, A.-S., Mordecai, S., Preffer, F.I., Amzallag, A., Ramaswamy, S., Hochedlinger, K., et al. (2014). Distinct, strict requirements for Gfi-1b in adult bone marrow red cell and platelet generation. J. Exp. Med. *211*, 909–927.

Guo, Y., Lubbert, M., and Engelhardt, M. (2003). CD34– Hematopoietic Stem Cells: Current Concepts and Controversies. Stem Cells *21*, 15–20.

Guo, Y., Fu, X., Jin, Y., Sun, J., Liu, Y., Huo, B., Li, X., and Hu, X. (2015). Histone demethylase LSD1-mediated repression of GATA-2 is critical for erythroid differentiation. Drug Des. Devel. Ther. *9*, 3153–3162.

Guo, Y., Fu, X., Huo, B., Wang, Y., Sun, J., Meng, L., Hao, T., Joe Zhao, Z., and Hu, X. (2016). GATA2 regulates GATA1 expression through LSD1-mediated histone modification. Am. J. Transl. Res. *8*, 2265–2274.

Gurdon, J.B. (1962). The developmental capacity of nuclei taken from intestinal epithelium cells of feeding tadpoles. J. Embryol. Exp. Morphol. *10*, 622–640.

Gurdon, J.B., Elsdale, T.R., and Fischberg, M. (1958). Sexually mature individuals of Xenopus laevis from the transplantation of single somatic nuclei. Nature *182*, 64–65.

Hakimi, M.A., Dong, Y., Lane, W.S., Speicher, D.W., and Shiekhattar, R. (2003). A candidate X-linked mental retardation gene is a component of a new family of histone deacetylase-containing complexes. J. Biol. Chem. *278*, 7234–7239.

Halazonetis, T.D., Georgopoulos, K., Greenberg, M.E., and Leder, P. (1988). c-Jun dimerizes with itself and with c-Fos, forming complexes of different DNA binding affinities. Cell *55*, 917–924.

Han, J.K., Chang, S.H., Cho, H.J., Choi, S.B., Ahn, H.S., Lee, J., Jeong, H., Youn, S.W., Lee, H.J., Kwon, Y.W., et al. (2014). Direct conversion of adult skin fibroblasts to endothelial cells by defined factors. Circulation *130*, 1168–1178.

Harris, W.J., Huang, X., Lynch, J.T., Spencer, G.J., Hitchin, J.R., Li, Y., Ciceri, F., Blaser, J.G., Greystoke, B.F., Jordan, A.M., et al. (2012). The Histone Demethylase KDM1A Sustains the Oncogenic Potential of MLL-AF9 Leukemia Stem Cells. Cancer Cell *21*, 473–487.

Hosseini, A., and Minucci, S. (2017). A comprehensive review of lysine-specific demethylase 1 and its roles in cancer. Epigenomics *9*.

Hou, P., Li, Y., Zhang, X., Liu, C., Guan, J., Li, H., Zhao, T., Ye, J., Yang, W., Liu, K., et al. (2013). Pluripotent

stem cells induced from mouse somatic cells by small-molecule compounds. Science (80-.). *341*, 651–654.

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

Humphrey, G.W., Wang, Y., Russanova, V.R., Hirai, T., Qin, J., Nakatani, Y., and Howard, B.H. (2001). Stable Histone Deacetylase Complexes Distinguished by the Presence of SANT Domain Proteins CoREST/kiaa0071 and Mta-L1. J. Biol. Chem. *276*, 6817–6824.

Ieda, M., Fu, J.-D., Delgado-Olguin, P., Vedantham, V., Hayashi, Y., Bruneau, B.G., and Srivastava, D. (2010). Direct Reprogramming of Fibroblasts into Functional Cardiomyocytes by Defined Factors Masaki. Cell *306*, 1895.

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

Ishikawa, Y., Gamo, K., Yabuki, M., Takagi, S., Toyoshima, K., Nakayama, K., Nakayama, A., Morimoto, M., Miyashita, H., Dairiki, R., et al. (2017). A novel LSD1 inhibitor T-3775440 disrupts GFI1B-containing complex leading to transdifferentiation and impaired growth of AML cells. Molecuclar Cancer Ther. *16*, 273–284.

Jenuwein, T., and Allis, C.D. (2001). Translating the histone code. Science (80-.). *293*, 1074–1080.

Juárez-Moreno, K., Erices, R., Beltran, A.S., Stolzenburg, S., Cuello-Fredes, M., Owen, G.I., Qian, H., and Blancafort, P. (2013). Breaking through an epigenetic wall: Re-activation of Oct4 by KRAB-containing designer zinc finger transcription factors. Epigenetics *8*, 164–176.

Kaimakis, P., Crisan, M., and Dzierzak, E. (2013). The Biochemistry of Hematopoietic Stem Cell Development. Biochim. Biophys. Acta *1830*, 2395–2403.

Kajimura, S., Seale, P., Kubota, K., Lunsford, E., Frangioni, J. V., Gygi, S.P., and Spiegelman, B.M. (2009). Initiation of myoblast to brown fat switch by a PRDM16-C/EBP-β transcriptional complex. Nature *460*, 1154–1158.

Keller, G., Kennedy, M., Papayannopoulou, T., and Wiles, M. V (1993). Hematopoietic commitment during embryonic stem cell differentiation in culture. Mol. Cell. Biol. *13*, 473–486.

Kerenyi, M.A., Shao, Z., Hsu, Y.-J., Guo, G., Luc, S., O'Brien, K., Fujiwara, Y., Peng, C., Nguyen, M., and Orkin, S.H. (2013). Histone demethylase Lsd1 represses hematopoietic stem and progenitor cell signatures during blood cell maturation. Elife *2*, 1–23.

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

Kim, P.G., and Daley, G.Q. (2009). Application of induced pluripotent stem cells to hematologic disease. Cytotherapy *11*, 980–989.

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

Ko, L.J., and Engel, J.D. (1993). DNA-binding specificities of the GATA transcription factor family. Mol. Cell. Biol. *13*, 4011–4022.

Kosan, C., and Godmann, M. (2016). Genetic and Epigenetic Mechanisms That Maintain Hematopoietic Stem Cell Function. Stem Cells Int. *2016*, 1–14.

Kouzarides, T. (2007). Chromatin Modifications and Their Function. Cell *128*, 693–705.

Krause, D.S., Ito, T., Fackler, M.J., Smith, O.M., Collector, M.I., Sharkis, S.J., and May, W.S. (1994). Characterization of murine CD34, a marker for hematopoietic progenitor and stem cells. Blood *84*, 691–701.

Ladewig, J., Koch, P., and Brüstle, O. (2013). Leveling Waddington: The emergence of direct programming and the loss of cell fate hierarchies. Nat. Rev. Mol. Cell Biol. *14*, 225–236.

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

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

Lee, M.G., Wynder, C., Cooch, N., and Shiekhattar, R. (2005). An essential role for CoREST in nucleosomal histone 3 lysine 4 demethylation. Nature *437*, 432–435.

Lee, S., Doddapaneni, K., Hogue, A., McGhee, L., Meyers, S., and Wu, Z. (2010). Solution structure of Gfi-1 zinc domain bound to consensus DNA. J. Mol. Biol. *397*, 1055–1066.

Li, W., Zhou, H., Abujarour, R., Zhu, S., Joo, J.Y., Lin, T., Hao, E., Schöler, H.R., Hayek, A., and Ding, S. (2009). Generation of human-induced pluripotent stem cells in the absence of exogenous Sox2. Stem Cells *27*, 2992–3000.

Lim, W.F., Inoue-Yokoo, T., Tan, K.S., Lai, M.I., and Sugiyama, D. (2013). Hematopoietic cell differentiation from embryonic and induced pluripotent stem cells. Stem Cell Res. Ther. *4*, 71.

Ling, K.-W., Ottersbach, K., van Hamburg, J.P., Oziemlak, A., Tsai, F.-Y., Orkin, S.H., Ploemacher, R., Hendriks, R.W., and Dzierzak, E. (2004). GATA-2 Plays Two Functionally Distinct Roles during the Ontogeny of Hematopoietic Stem Cells. J. Exp. Med. *200*, 871–882.

Lis, R., Karrasch, C.C., Poulos, M.G., Kunar, B., Redmond, D., Duran, J.G.B., Badwe, C.R., Schachterle, W., Ginsberg, M., Xiang, J., et al. (2017). Conversion of adult endothelium to immunocompetent haematopoietic stem cells. Nature *545*, 439–445.

Lupo, A., Cesaro, E., Montano, G., Zurlo, D., Izzo, P., and Costanzo, P. (2013). KRAB-Zinc Finger Proteins: A Repressor Family Displaying Multiple Biological Functions. Curr. Genomics *14*, 268–278.

Maiques-Diaz, A., and Somervaille, T.C. (2016). LSD1: biologic roles and therapeutic targeting. Epigenomics *8*, 1103–1116.

McGrath, J.P., Williamson, K.E., Balasubramanian, S., Odate, S., Arora, S., Hatton, C., Edwards, T.M., O'Brien, T., Magnuson, S., Stokoe, D., et al. (2016). Pharmacological inhibition of the histone lysine demethylase KDM1A suppresses the growth of multiple acute myeloid leukemia subtypes. Cancer Res. *76*, 1975–1988.

McKinney-Freeman, S., Cahan, P., Li, H., Lacadie, S.A., Huang, H.-T., Curran, M., Loewer, S., Naveiras, O., Kathrein, K.L., Konantz, M., et al. (2012). The Transcriptional Landscape of Hematopoietic Stem Cell Ontogeny. Cell Stem Cell *11*, 701–714.

Medvinsky, A., Rybtsov, S., Taoudi, S., and Glossary, B. (2011). Embryonic origin of the adult hematopoietic system : advances and questions. *1031*, 1017–1031.

Metcalf, D. (1989). The molecular control of cell division, diferentiation commitment and maturation in haemopoietic cells. Nature *339*, 27–30.

Metzger, E., Wissmann, M., Yin, N., Müller, J.M., Schneider, R., Peters, A.H.F.M., Günther, T., Buettner, R., and Schüle, R. (2005). LSD1 demethylates repressive histone marks to promote androgen-receptor-dependent transcription. Nature *437*, 436–439.

Minegishi, N., Suzuki, N., Yokomizo, T., Pan, X., Fujimoto, T., Takahashi, S., Hara, T., Miyajima, A., Nishikawa, S. ichi, and Yamamoto, M. (2003). Expression and domain-specific function of GATA-2 during differentiation of the hematopoietic precursor cells in midgestation mouse embryos. Blood *102*, 896–905.

Minegishi, N., Suzuki, N., Kawatani, Y., Shimizu, R., and Yamamoto, M. (2005). Rapid turnover of GATA-2 via ubiquitin-proteasome protein degradation pathway. Genes to Cells *10*, 693–704.

Moignard, V., Macaulay, I.C., Swiers, G., Buettner, F., Schütte, J., Calero-Nieto, F.J., Kinston, S., Joshi, A., Hannah, R., Theis, F.J., et al. (2013). Characterization of transcriptional networks in blood stem and

progenitor cells using high-throughput single-cell gene expression analysis. Nat. Cell Biol. *15*, 363–372.

Moore, K.A., and Lemischka, I.R. (2006). Stem Cells and Their Niches. Science (80-.). *311*, 1880–1885.

Mosammaparast, N., Kim, H., Laurent, B., Zhao, Y., Lim, H.J., Majid, M.C., Dango, S., Luo, Y., Hempel, K., Sowa, M.E., et al. (2013). The histone demethylase LSD1/KDM1A promotes the DNA damage response. J. Cell Biol. *203*, 457–470.

Orkin, S.H. (1992). GATA-binding transcription factors in hematopoietic cells. Blood *80*, 575–581.

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

Passweg, J.R., Halter, J., Bucher, C., Gerull, S., Heim, D., Rovó, A., Buser, A., Stern, M., and Tichelli, A. (2012). Hematopoietic stem cell transplantation : a review and recommendations for follow-up care for the general practitioner. *2010*, 1–15.

Passweg, J.R., Baldomero, H., Bader, P., Bonini, C., Cesaro, S., Dreger, P., Duarte, R.F., Dufour, C., and Falkenburg, J.H.F. (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 Transplant. *50*, 476–482.

Passweg, J.R., Baldomero, H., Bader, P., Bonini, C., Cesaro, S., Dreger, P., Duarte, R.F., Dufour, C., Kuball, J., Farge-Bancel, D., et al. (2016). Hematopoietic stem cell transplantation in Europe 2014: More than 40 000 transplants annually. Bone Marrow Transplant. *51*, 786–792.

Pereira, C.-F., Lemischka, I.R., and Moore, K. (2014). "From blood to blood": de-differentiation of hematopoietic progenitors to stem cells. EMBO J. *33*, 1511–1513.

Pereira, C.-F., Chang, B., Gomes, A., Bernitz, J., Papatsenko, D., Niu, X., Swiers, G., Azzoni, E., de Bruijn, M.F.T.R., Schaniel, C., et al. (2016). Hematopoietic Reprogramming In Vitro Informs In Vivo Identification of Hemogenic Precursors to Definitive Hematopoietic Stem Cells. Dev. Cell *36*, 525–539.

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

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

Petersdorf, E.W. (2013). The major histocompatibility complex : a model for understanding graft-versus-host disease. Blood *122*, 1863–1872.

Pulecio, J., Alejo-Valle, O., Capellera-Garcia, S., Vitaloni, M., Rio, P., Mejía-Ramírez, E., Caserta, I., Bueren, J.A., Flygare, J., and Raya, A. (2016). Direct Conversion of Fibroblasts to Megakaryocyte Progenitors. Cell Rep. *17*, 671–683.

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

Riddell, J., Gazit, R., Garrison, B.S., Guo, G., Saadatpour, A., Mandal, P.K., Ebina, W., Volchkov, P., Yuan, G.-C., Orkin, S.H., et al. (2014). Reprogramming Committed Murine Blood Cells to Induced Hematopoietic Stem Cells with Defined Factors. Cell *157*, 549–564.

Ring, K., Tong, L., and Balestra, M. (2012). Direct reprogramming of mouse and human fibroblasts into multipotent neural stem cells with a single factor. Cell Stem Cell *11*, 100–109.

Rodrigues, N.P., Tipping, A.J., Wang, Z., and Enver, T. (2012). GATA-2 mediated regulation of normal hematopoietic stem/progenitor cell function, myelodysplasia and myeloid leukemia. Int. J. Biochem. Cell Biol. *44*, 457–460.

Rossi, L., Challen, G.A., Sirin, O., Lin, K.K., and Goodell, M.A. (2011). Hematopoietic Stem Cell Characterization and Isolation. Methods Mol. Biol. *750*, 1–10.

La Russa, M.F., and Qi, L.S. (2015). The New State of the Art: Cas9 for Gene Activation and Repression. Mol. Cell. Biol. *35*, 3800–3809.

Sadahira, K., Fukuchi, Y., Kunimono, H., Sakurai, M., Ikeda, Y., Okamoto, S., and Nakajima, H. (2012). Direct reprogramming of terminally differentiated B cells into erythroid lineage. FEBS Lett. *586*, 3645–3652.

Sakamoto, A., Hino, S., Nagaoka, K., Anan, K., Takase, R., Matsumori, H., Ojima, H., Kanai, Y., Arita, K., and Nakao, M. (2015). Lysine demethylase LSD1 coordinates glycolytic and mitochondrial metabolism in hepatocellular carcinoma cells. Cancer Res. *75*, 1445–1456.

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

Saleque, S., Kim, J., Rooke, H.M., and Orkin, S.H. (2007). Epigenetic Regulation of Hematopoietic Differentiation by Gfi-1 and Gfi-1b Is Mediated by the Cofactors CoREST and LSD1. Mol. Cell *27*, 562–572.

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

Sassone-Corsi, P., Ransone, L.J., Lamph, W.W., and Verma, I.M. (1988). Direct interaction between fos and jun nuclear oncoproteins: role of the “leucine zipper” domain. Nature *336*, 692–695.

Schneuwly, S., Klemenz, R., and Gehring, W.J. (1987). Redesigning the body plan of Drosophila by ectopic expression of the homoeotic gene Antennapedia. Nature *325*, 816–818.

Schuster, J.A., Stupnikov, M.R., Ma, G., Liao, W., Lai, R., Ma, Y., and Aguila, J.R. (2012). Expansion of hematopoietic stem cells for transplantation: current perspectives. Exp. Hematol. Oncol. *1*, 1–6.

Sheng, C., Zheng, Q., Wu, J., Xu, Z., Sang, L., Wang, L., Guo, C., Zhu, W., Tong, M., Liu, L., et al. (2012). Generation of dopaminergic neurons directly from mouse fibroblasts and fibroblast-derived neural progenitors. Cell Res. *22*, 769–772.

Shi, Y., Lan, F., Matson, C., Mulligan, P., Whetstine, J.R., Cole, P.A., Casero, R.A., and Shi, Y. (2004). Histone demethylation mediated by the nuclear amine oxidase homolog LSD1. Cell *119*, 941–953.

Shi, Y.J., Matson, C., Lan, F., Iwase, S., Baba, T., and Shi, Y. (2005). Regulation of LSD1 histone demethylase activity by its associated factors. Mol. Cell *19*, 857–864.

Siena, S., Schiavo, R., Pedrazzoli, P., and Carlo-Stella, C. (2000). Therapeutic Relevance of CD34 Cell Dose in Blood Cell Transplantation for Cancer Therapy. J. Clin. Oncol. *18*, 1360–1377.

Silva, J., Chambers, I., Pollard, S., and Smith, A. (2006). Nanog promotes transfer of pluripotency after cell fusion. Nature *441*, 997–1001.

Smith, A.R., and Wagner, J.E. (2009). Alternative haematopoietic stem cell sources for transplantation: place of umbilical cord blood. Br. J. Haematol. *147*, 246–261.

Sommer, C.A., Stadtfeld, M., Murphy, G.J., Hochedlinger, K., Kotton, D.N., and Mostoslavsky, G. (2009). Induced Pluripotent Stem Cell Generation Using a Single Lentiviral Stem Cell Cassette. Stem Cells *27*, 543–549.

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

Sprüssel, A., Schulte, J.H., Weber, S., Necke, M., Händschke, K., Thor, T., Pajtler, K.W., Schramm, A., König, K., Diehl, L., et al. (2012). Lysine-specific demethylase 1 restricts hematopoietic progenitor proliferation and is essential for terminal differentiation. Leukemia *26*, 2039–2051.

Di Stefano, B., Collombet, S., Jakobsen, J.S., Wierer, M., Sardina, J.L., Lackner, A., Stadhouders, R., Segura-Morales, C., Francesconi, M., Limone, F., et al. (2016). C/EBPα creates elite cells for iPSC reprogramming by upregulating Klf4 and increasing the levels of Lsd1 and Brd4. Nat. Cell Biol. *18*, 371–381.

Stella, C.C., Cazzola, M., De Fabritiis, P., De Vincentiis, A., Massimo Gianni, A., Lanza, F., Lauria, F., Lemoli, R.M., Tarella, C., Zanon, P., et al. (1995). CD34-positive cells: Biology and clinical relevance. Haematologica

80, 367–387.

Sugimura, R., Jha, D.K., Han, A., Soria-Valles, C., da Rocha, E.L., Lu, Y.-F., Goettel, J.A., Serrao, E., Rowe, R.G., Malleshaiah, M., et al. (2017). Haematopoietic stem and progenitor cells from human pluripotent stem cells. Nature *545*, 432–438.

Sugino, N., Kawahara, M., Tatsumi, G., Kanai, A., Matsui, H., Yamamoto, R., Nagai, Y., Fujii, S., Shimazu, Y., Hishizawa, M., et al. (2017). A novel LSD1 inhibitor NCD38 ameliorates MDS-related leukemia with complex karyotype by attenuating leukemia programs via activating super-enhancers. Leukemia *31*, 2303–2314.

Sun, H., Liang, L., Li, Y., Feng, C., Li, L., Zhang, Y., He, S., Pei, D., Guo, Y., and Zheng, H. (2016). Lysine-specific histone demethylase 1 inhibition promotes reprogramming by facilitating the expression of exogenous transcriptional factors and metabolic switch. Sci. Rep. *6*, 30903.

Suzuki, N., Yamazaki, S., Yamaguchi, T., Okabe, M., Masaki, H., Takaki, S., Otsu, M., and Nakauchi, H. (2013). Generation of Engraftable Hematopoietic Stem Cells From Induced Pluripotent Stem Cells by Way of Teratoma Formation. Mol. Ther. *21*, 1424–1431.

Tada, M., Takahama, Y., Abe, K., Nakatsuji, N., and Tada, T. (2001). Nuclear reprogramming of somatic cells by in vitro hybridization with ES cells. Curr. Biol. *11*, 1553–1558.

Takagi, S., Ishikawa, Y., Mizutani, A., Iwasaki, S., Matsumoto, S., Kamada, Y., Nomura, T., and Nakamura, K. (2017). LSD1 inhibitor T-3775440 inhibits SCLC cell proliferation by disrupting LSD1 interactions with SNAG domain proteins INSM1 and GFI1B. Cancer Res. *77*, 4652–4662.

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

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

Takeuchi, M., Fuse, Y., Watanabe, M., Andrea, C.-S., Takeuchi, M., Nakajima, H., Ohashi, K., Kaneko, H., Kobayashi-Osaki, M., Yamamoto, M., et al. (2015). LSD1/KDM1A promotes hematopoietic commitment of hemangioblasts through downregulation of Etv2. Proc. Natl. Acad. Sci. *112*, 13922–13927.

Takizawa, H., Schanz, U., and Manz, M.G. (2011). Ex vivo expansion of hematopoietic stem cells : mission accomplished ? Swiss Med. Wkly. *141*, 1–9.

Tan, Y.-T., Ye, L., Xie, F., Beyer, A.I., Muench, M.O., Wang, J., Chen, Z., Liu, H., Chen, S.-J., and Kan, Y.W. (2018). Respecifying human iPSC-derived blood cells into highly engraftable hematopoietic stem and progenitor cells with a single factor. Proc. Natl. Acad. Sci. 201718446.

Thambyrajah, R., Mazan, M., Patel, R., Moignard, V., Stefanska, M., Marinopoulou, E., Li, Y., Lancrin, C., Clapes, T., Möröy, T., et al. (2015). GFI1 proteins orchestrate the emergence of haematopoietic stem cells through recruitment of LSD1. Nat. Cell Biol. *18*, 21–32.

Tong, Q., Tsai, J., and Tan, G. (2005). Interaction between GATA and the C / EBP Family of Transcription Factors Is Critical in GATA-Mediated Suppression of Adipocyte Differentiation. Society *25*, 706–715.

Tsai, F.Y., and Orkin, S.H. (1997). Transcription factor GATA-2 is required for proliferation/survival of early hematopoietic cells and mast cell formation, but not for erythroid and myeloid terminal differentiation. Blood *89*, 3636–3643.

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

Turner, R., and Tjian, R. (1989). Leucine repeats and an adjacent DNA binding domain mediate the formation of functional cFos-cJun heterodimers. Science *243*, 1689–1694.

Velinder, M., Singer, J., Bareyan, D., Meznarich, J., Tracy, C.M., Fulcher, J.M., McClellan, D., Lucente, H., Franklin, S., Sharma, S., et al. (2016). GFI1 functions in transcriptional control and cell fate determination require SNAG domain methylation to recruit LSD1. Biochem. J. *473*, 3355–3369.

Vicente, C., Conchillo, A., García-Sánchez, M.A., and Odero, M.D. (2012). The role of the GATA2

transcription factor in normal and malignant hematopoiesis. Crit. Rev. Oncol. Hematol. *82*, 1–17.

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

Waddington, C.H. (1968). Towards a Theoretical Biology. Nature *218*, 525–527.

Wang, J., Scully, K., Zhu, X., Cai, L., Zhang, J., Prefontaine, G.G., Krones, A., Ohgi, K.A., Zhu, P., Garcia-Bassets, I., et al. (2007). Opposing LSD1 complexes function in developmental gene activation and repression programmes. Nature *446*, 882–887.

Wang, J., Hevi, S., Kurash, J.K., Lei, H., Gay, F., Bajko, J., Su, H., Sun, W., Chang, H., Xu, G., et al. (2008). The lysine demethylase LSD1 (KDM1) is required for maintenance of global DNA methylation. Nat. Genet. *41*, 125–129.

Wang, Y., Zhang, H., Chen, Y., Sun, Y., Yang, F., Yu, W., Liang, J., Sun, L., Yang, X., Shi, L., et al. (2009). LSD1 Is a Subunit of the NuRD Complex and Targets the Metastasis Programs in Breast Cancer. Cell *138*, 660–672.

Whyte, W.A., Bilodeau, S., Orlando, D.A., Hoke, H.A., Frampton, G.M., Foster, C.T., Cowley, S.M., and Young, R.A. (2012). Enhancer decommissioning by LSD1 during embryonic stem cell differentiation. Nature *482*, 221–225.

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

Xie, H., Ye, M., Feng, R., and Graf, T. (2004). Stepwise reprogramming of B cells into macrophages. Cell *117*, 663–676.

Xu, J., Du, Y., and Deng, H. (2015). Direct lineage reprogramming: Strategies, mechanisms, and applications. Cell Stem Cell *16*, 119–134.

Yamanaka, S., and Blau, H.M. (2010). Nuclear reprogramming to a pluripotent state by three approaches. Nature *465*, 704–712.

Yang, M., Gocke, C.B., Luo, X., Borek, D., Tomchick, D.R., Machius, M., Otwinowski, Z., and Yu, H. (2006). Structural Basis for CoREST-Dependent Demethylation of Nucleosomes by the Human LSD1 Histone Demethylase. Mol. Cell *23*, 377–387.

You, A., Tong, J.K., Grozinger, C.M., and Schreiber, S.L. (2001). CoREST is an integral component of the CoREST- human histone deacetylase complex. Proc. Natl. Acad. Sci. *98*, 1454–1458.

Zhou, H., Li, W., Zhu, S., Joo, J.Y., Do, J.T., Xiong, W., Kim, J.B., Zhang, K., Schöler, H.R., and Ding, S. (2010). Conversion of mouse epiblast stem cells to an earlier pluripotency state by small molecules. J. Biol. Chem. *285*, 29676–29680.

Zovein, A.C., Hofmann, J.J., Lynch, M., French, W.J., Turlo, K.A., Yang, Y., Becker, M.S., Zanetta, L., Dejana, E., Gasson, J.C., et al. (2008). Fate Tracing Reveals the Endothelial Origin of Hematopoietic Stem Cells. Cell Stem Cell *3*, 625–636.

Zweidler-Mckay, P. a, Grimes, H.L., Flubacher, M.M., and Tsichlis, P.N. (1996). Gfi-1 encodes a nuclear zinc finger protein that binds DNA and functions as a transcriptional repressor. Mol. Cell. Biol. *16*, 4024–4034.