PhD

Program in Translational and Molecular Medicine

DIMET

University of Milano-Bicocca

Department of Biotechnology and Biosciences

"Transcriptional regulation and downstream targets of the COUP-TFII transcription factor in erythroid cells"

Agata Labedz

Matr. No. 863947

PhD Coordinator: Prof. Francesco Mantegazza

Supervisor: Prof. Antonella Ronchi

XXXV cycle

Academic year: 2021/2022

This work was supported by the European Union's Horizon 2020 Research and Innovation Program under the Marie Skłodowska-Curie grant agreement No. 813091 ("ARCH, Age-related changes in hematopoiesis").

Table of contents

CHAPTER 1:

GENERAL INTRODUCTION

1.1. Erythropoiesis

Hematopoiesis is controlled both spatially and temporally during mammalian development. The production of blood cells starts in the yolk sac, transits via the fetal liver, and then is finally established in the thymus and bone marrow (Dzierzak and Bigas, 2018). Erythropoiesis is the process through which hematopoietic stem and progenitor cells (HSPCs) multiply and differentiate into enucleate reticulocytes at various functionally and morphologically defined phases. These cells continue to develop in the circulation to become red blood cells (RBCs), which are crucial carriers of oxygen and carbon dioxide for cellular respiration (Nandakumar, 2016). RBCs are the most prevalent cell type in humans, producing 2 million of them every second in the bone marrow and have an average life span of 120 days (Palis, 2014). Interestingly, in response to hypoxic stress, which happens when an adequate oxygen supply to all tissues is hampered by a lack of functional erythrocytes, the rate of erythropoiesis can dramatically increase from this baseline level. The main defense against hypoxic stress is increased red cell synthesis. Via the activity of the hypoxia-inducible transcription factor complex, tissue oxygen tension directly controls the production of erythropoietin (EPO), the main regulator of red cell production (Bunn, 2013).

1.1.1. Waves of erythropoiesis

As soon as the embryo becomes too large to receive oxygen via diffusion, the first erythroid precursors emerge from the yolk sac, a transient extraembryonic tissue (Fig. 1.1). They are derived from mesodermal cells that migrate posteriorly and, after entering the yolk sac, interact with endoderm cells. The interaction between mesodermal and endoderm cells is essential to initiate erythropoiesis. These mesodermal cells form blood islands that consist of both erythroid and endothelial cells indicating their common mesodermal precursor with bilineage potential, the hemangioblast (Dzierzak and Philipsen, 2013). Primitive erythrocytes (EryPs) are relatively large, enucleate in the bloodstream and express embryonic globin genes. Between primitive and definitive hematopoietic stem cell (HSC)-dependent hematopoiesis, a second wave of cells migrate from the yolk sac into the fetal liver and support fetal hematopoiesis until birth. Interestingly, some studies suggest that this second transient fetal hematopoietic wave of yolk sac-derived erythro-myeloid progenitors (EMPs) in mice may continue after birth (Epelman, 2014; Gomez Perdiguero, 2015). The final, third wave of erythropoiesis constitute hematopoietic stem cells (HSCs). HSCs are derived from endothelial cells as they arise from the hemogenic endothelium of the embryonic aorta-gonad-mesonephros (AGM) region, the placenta and the vitelline and umbilical arteries (Dzierzak and Philipsen, 2013; Dzierzak and Bigas, 2018). These cells move from the fetal liver to the bone

marrow, where they will sustain hematopoiesis, including erythrocyte production, throughout adulthood (Palis, 2014; Dzierzak and Bigas, 2018).

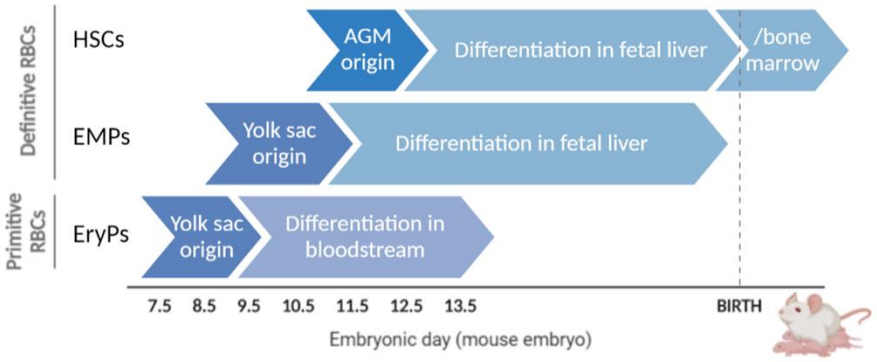


Figure 1.1. ***Waves of erythropoiesis during development.***

1.2. Differential globin genes expression during development

1.2.1. Hemoglobin switching

The molecule that carries oxygen to tissues, hemoglobin, is a heterotetramer composed of two α- and two β-like globin chains. An iron-containing heme group that can bind one molecule of oxygen is coordinated by each globin chain. The α- and β-globin loci, in human located on chromosomes 16 and 11, respectively, encode the α- and β-like globin chains (Fig. 1.2A). Different α- and β-like globin genes are expressed at different times during different stages of development. There are three active α-like globin genes: HBZ (ζ-globin), HBA2 (α_2-globin), and HBA1 (α_1-globin) controlled by MCS (multispecies conserved sequences) distal enhancer (Higgs, 1990; Stamatoyannopoulos, 2005). The β-globin locus consists of five functional genes: HBE1 (ε-globin), HBG2 (^{G}γ-globin), HBG1 (^{A}γ-globin), HBD (δ-globin), and HBB (β-globin) (Grosveld, 1987; Stamatoyannopoulos, 2005). A set of distal enhancer controlling the β-locus is called locus control region (LCR). The yolk sac-derived earliest embryonic red blood cells (RBCs) primarily express HBZ and HBE1, which form the embryonic hemoglobin (Hemoglobin Gower1, $\zeta_2\varepsilon_2$). Subsequently, the site of erythropoiesis becomes fetal liver

where fetal hemoglobin (HbF) composed of two α- and two γ-globin chains ($\alpha_2\gamma_2$), is expressed. Throughout the embryonic and fetal stages, low amounts of Hemoglobin Gower2 ($\alpha_2\varepsilon_2$) and Hemoglobin Portland1 ($\zeta_2\gamma_2$) may also be present. Around the time of birth, erythropoiesis shifts to the bone marrow. The expression of adult hemoglobin (HbA, $\alpha_2\beta_2$) predominates in adult erythroid cells, with only about 2% of HbA2 ($\alpha_2\delta_2$) and less than 1% of HbF. Changes in hemoglobin composition during development is mainly due to hemoglobin switching, a process of changes in globin gene expression at different developmental stages (Stamatoyannopoulos, 2005; Sankaran, 2010) (Fig. 1.2B).

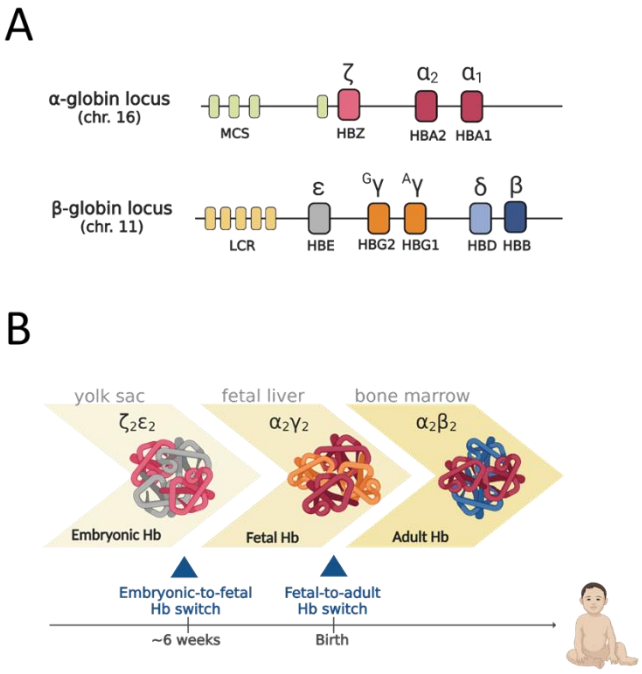


Figure 1.2. ***Organization of human globin loci (A) and developmental hemoglobin switching (B).***

1.2.2. Physiological and Aberrant γ-Globin Transcription During Development

Published by **Gloria Barbarani**, **Agata Labedz**, Sarah Stucchi, Alessia Abbiati and Antonella E. Ronchi. Frontiers in Cell and Developmental Biology, April 1, 2021.

In this review, we describe the regulation of hemoglobin switching on the β-locus with a special focus on the major transcription factors (TFs) controlling differential expression of γ-globin during development.

frontiers in Cell and Developmental Biology

MINI REVIEW
published: 01 April 2021
doi: 10.3389/fcell.2021.640060

Physiological and Aberrant γ-Globin Transcription During Development

*Gloria Barbarani†, Agata Labedz†, Sarah Stucchi, Alessia Abbiati and Antonella E. Ronchi**

Dipartimento di Biotecnologie e Bioscienze, Università di Milano-Bicocca, Milano, Italy

The expression of the fetal Gγ- and Aγ-globin genes in normal development is confined to the fetal period, where two γ-globin chains assemble with two α-globin chains to form $\alpha_2\gamma_2$ tetramers (HbF). HbF sustains oxygen delivery to tissues until birth, when β-globin replaces γ-globin, leading to the formation of $\alpha_2\beta_2$ tetramers (HbA). However, in different benign and pathological conditions, HbF is expressed in adult cells, as it happens in the hereditary persistence of fetal hemoglobin, in anemias and in some leukemias. The molecular basis of γ-globin differential expression in the fetus and of its inappropriate activation in adult cells is largely unknown, although in recent years, a few transcription factors involved in this process have been identified. The recent discovery that fetal cells can persist to adulthood and contribute to disease raises the possibility that postnatal γ-globin expression could, in some cases, represent the signature of the fetal cellular origin.

Keywords: globin genes, transcription factors, hereditary persistence of fetal hemoglobin, juvenile myelomonocytic leukemia, erythropoiesis

OPEN ACCESS

Edited by:
Claudio Cantù, Linköping University, Sweden

Reviewed by:
Annarita Miccio, INSERM U1163 Institut Imagine, France
Sjaak Philipsen, Department of Cell Biology, Erasmus MC, Netherlands

***Correspondence:**
Antonella E. Ronchi
antonella.ronchi@unimib.it

†*These authors share first authorship*

Specialty section:
This article was submitted to Stem Cell Research, a section of the journal Frontiers in Cell and Developmental Biology

Received: *10 December 2020*
Accepted: *23 February 2021*
Published: *01 April 2021*

Citation:
Barbarani G, Labedz A, Stucchi S, Abbiati A and Ronchi AE (2021) Physiological and Aberrant γ-Globin Transcription During Development. Front. Cell Dev. Biol. 9:640060. doi: 10.3389/fcell.2021.640060

ERYTHROPOIESIS DURING DEVELOPMENT

During mammalian development, hematopoiesis is regulated both spatially and temporally: it begins in the yolk sac, it goes through a transitory phase in the fetal liver and then is definitively established in the thymus and bone marrow (Dzierzak and Bigas, 2018). The first erythroid precursors emerge from the yolk sac as soon as the embryo grows too big to be supplied with oxygen by diffusion and give rise to primitive erythroid cells (EryPs). These cells are released in the bloodstream when they are still nucleated and are characterized by the expression of embryonic globins. A second wave of cells migrating into the fetal liver from the yolk sac support fetal hematopoiesis until birth, in the interval between primitive and definitive hematopoietic stem cell (HSC)-dependent hematopoiesis. Of interest, recent studies suggest that in mouse, this second fetal transient hematopoietic wave of yolk sac-derived erythro-myeloid progenitors (EMPs) may persist postnatally (Epelman et al., 2014; Gomez Perdiguero et al., 2015). HSCs arise from the hemogenic endothelium of the embryonic aorta-gonad-mesonephros (AGM), the vitelline and umbilical arteries, and from the placenta (Dzierzak and Philipsen, 2013; Lacaud and Kouskoff, 2017; Dzierzak and Bigas, 2018). These cells migrate first into the FL and then to the bone marrow (BM)—their long-term adult resident location—where they will last throughout life and will generate all types of blood cells, including erythrocytes (Orkin and Zon, 2008; Palis, 2014; Dzierzak and Bigas, 2018).

The different types of erythroid cells produced at the different hematopoietic stages have many common characteristics, including the main steps of progressive differentiation and maturation

from early progenitors to erythroblasts and finally to red blood cells (RBCs) (Kina et al., 2000; Aisen, 2004; Chen et al., 2009; Baron, 2013; Palis, 2014). However, importantly, they can be in part distinguished by differences in cell morphology and in the expression of embryo/fetal vs. adult globins.

GLOBIN GENES

In humans, the α-globin cluster contains three functional genes: the embryonic, HBZ (ζ-globin) and the two fetal/adult HBA2 and HBA1 duplicated genes (α2- and α1-globin) (Stamatoyannopoulos, 2005). The β-globin cluster contains five active genes: the embryonic HBE (ε-globin) gene, the two highly homologous fetal HBG2 and HBG1 genes (Gγ- and Aγ-globin, respectively) and the two adult HBD and HBB genes (δ- and β-globin, the latter accounting for about 98% of adult β-like globin) (**Figure 1**). Each locus is under the control of a set of distal enhancers (Grosveld et al., 1987; Higgs et al., 1990). The genes contained in the α-globin and β-globin loci are sequentially expressed in a stage-specific manner that maintains the 1:1 ratio between the α-like and β-like globin chains, in a process known as "hemoglobin switching" (Forget, 1990; Stamatoyannopoulos, 2005; Sankaran et al., 2010a).

Interestingly, the presence of fetal-specific (γ) genes and thus of a fetal (γ) to adult (β) globin switch is unique to humans and old-world monkeys: most species, including mice, have only one switch, from embryo/fetal to definitive globin genes expression, occurring early in development (Stamatoyannopoulos, 1991; Sankaran et al., 2010a; Philipsen and Hardison, 2018; **Figure 1**). In mice transgenic for the human β-globin locus, the switching of human globin genes parallels the switching of mouse genes, with γ genes being switched off between E11.5 and E13.5, together with εγ and βh1 mouse embryo/fetal genes (Strouboulis et al., 1992; Peterson et al., 1995). Each developmental switch is accompanied by a profound chromatin remodeling within the loci: the interaction ("loop") between the promoter of the gene active at a given time with the common distal enhancers [locus control region (LCR)] is progressively favored, with inactive globin genes being looped out (Tolhuis et al., 2002; Palstra et al., 2003). Whereas only EryPs of yolk sac origin expresses embryonic ζ-, human ε- and mouse εy-globin genes, the other globin genes are more promiscuously expressed by cells of different origin. McGrath et al. (2011) showed that the first cells expressing adult globins, prior to the generation of HSC-derived erythroblasts, are indeed the transient population of EMP-derived erythroid cells.

PHYSIOLOGICAL AND NON PHYSIOLOGICAL γ-GLOBIN EXPRESSION

The switching from γ- to β-globin expression is the most intensively studied because the persistence of γ-globin expression in adult stages is a hallmark of a very heterogeneous spectrum of conditions. These can be benign, as in the case of the few F cells (cells expressing HbF) found in normal adults (Boyer et al., 1975) and in HPFH (hereditary persistence of fetal hemoglobin) (Forget, 1998), or associated with disease, such as in response to transient or chronic anemias (Weatherall, 2001) or leukemias (Weatherall et al., 1968; Sheridan et al., 1976). As an additional reason of interest, the ability of reactivating γ-globin in the adult is considered as a possible strategy to cure β-hemoglobinopathies (Wienert et al., 2018). The cause of γ-globin expression in adult cells remains largely unknown and is thought to rely on different mechanisms, both maturational and/or directly related to defects intrinsic to the HBB locus (Zago et al., 1979; Stamatoyannopoulos, 2005). Although these aspects are strictly intertwined, and it is almost impossible to sharply separate them, this review will focus on the latter, in particular on the major transcription factors (TFs) that, by directly binding to the HBB locus, act as selective on/off switches of γ-globin expression in normal and aberrant conditions.

THE TRANSCRIPTIONAL SWITCHES OF γ-GLOBIN EXPRESSION

The observation that the main differentiation and maturation steps leading to RBC formation are common to YS-, EMP-, and HSC-derived erythroblasts (Kina et al., 2000; Aisen, 2004; Chen et al., 2009; Palis, 2014) suggests that these cells may also rely on a common set of transcription factors directing erythroid differentiation and globin gene regulation. Indeed, embryonic/fetal and adult cells share a common set of ubiquitous (such as NF-Y, that binds all globin promoters, although with different affinity (Liberati et al., 1998; Zhu et al., 2012; Martyn et al., 2017)) and erythroid-specific activators/coactivators [first of all GATA1 (Ferreira et al., 2005; Love et al., 2014; Katsumura et al., 2017; Barbarani et al., 2019), NFE2(Gasiorek and Blank, 2015; Kim et al., 2016), KLF1(Perkins et al., 2016), and TAL1 (Kang et al., 2015)]. Moreover, the expression of γ-globin genes in adult cells, as in HPFH, suggests that adult cells represent an environment permissive for the expression of both embryo/fetal and adult globin genes. It is thus likely that the specific timing of γ-globin expression might require specific activators/repressors acting at different times. The number of transcription factors directly involved in the activation/repression of γ-globin transcription is surprisingly small, and their main characteristics are briefly reviewed here below (**Figure 2**).

TRANSCRIPTION FACTORS AFFECTING THE EMBRYO/FETAL TO ADULT SWITCHING

γ-Globin Repressors

BCL11A, also known as CTIP1 (Coup-TFII interacting protein), is a C2H2-type zinc-finger protein (Avram et al., 2000). Alternative splicing generates four major protein isoforms, sharing a common N-terminus: eXtra-Long (XL), Long (L), Short (S), and eXtra-Short (XS) (Liu et al., 2006). BCL11A plays important roles in non-erythroid hematopoietic cells,

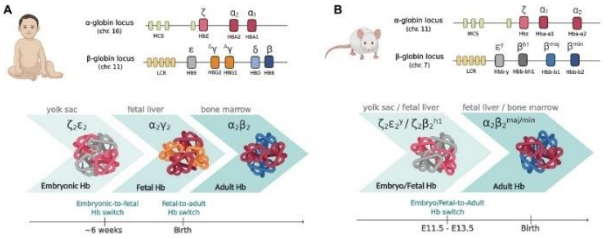


FIGURE 1 | The globin gene loci in humans **(A)** and mice **(B)**. Upper panels: the globin gene loci. MCS (multispecies conserved sequences) and LCR (locus control region) are the distal enhancers coordinating the developmental expression of the genes within the cluster. Lower panels: the different types of globin chains expressed at the different developmental stages and the corresponding sites of hematopoiesis are shown. Please note that whereas in humans two subsequent switching events take place, in mouse, as the majority of mammals, there is only one switch from embryo/fetal to adult genes, occurring early in development. Created with BioRender.com.

including B cells (Liu et al., 2003), dendritic cells (Ippolito et al., 2014), and hematopoietic stem cells (HSCs) (Tsang et al., 2015; Luc et al., 2016). Outside hematopoiesis, BCL11A is essential for central nervous system development (Basak et al., 2015; Funnell et al., 2015; Greig et al., 2016) and possibly for the differentiation of other lineages, such as breast (Khaled et al., 2015) and pancreas (Peiris et al., 2018) cells. Its specific role in γ-globin silencing was identified by genome-wide association studies (GWAS) aiming to identify eQTLs (expression quantitative trait loci) associated with high levels of postnatal HbF (Menzel et al., 2007; Lettre et al., 2008; Uda et al., 2008). Its conditional knockout within the erythroid compartment [obtained by disrupting the erythroid-specific BCL11A enhancer (Bauer et al., 2013; Canver et al., 2015)] impairs HbF silencing in adult erythroid cells, without altering erythropoiesis in the mouse (Sankaran et al., 2008, 2010b; Xu et al., 2013). This latter observation made BCL11A-targeted inactivation in erythroid cells a promising approach to reactivate γ-globin in β-hemoglobinopathies (Sankaran et al., 2010a; Wienert et al., 2018; Zeng et al., 2020). Interestingly, in human cells, S and XS isoforms are specific of YS primitive and FL erythropoiesis, whereas XL and L are specific of BM definitive erythropoiesis (Sankaran et al., 2009). Notably, the mouse *Bcl11A* pattern of expression is different, with the XL isoform being already present in FL definitive erythroid cells. This delay in BCL11A-XL in humans could explain the different timing of γ-globin switching in human vs. mouse (Sankaran et al., 2009). The specific γ-globin repression in adult cells is mediated by its binding to the consensus sequence GGTCA, present in several discrete sites within the β-locus (Liu et al., 2018). Among them, the site in the distal CCAAT box region of the γ-promoter, which contains the −117 residue, whose G > A mutation causes HPFH, indeed abolishes BCL11A-XL binding (Martyn et al., 2018).

BCL11A is activated by *KLF1/EKLF* (Borg et al., 2010; Zhou et al., 2010) (Kruppel-Like Factor-1), an erythroid-specific zinc finger TF originally identified because of its ability to bind to CACCC motifs (Miller and Bieker, 1993) and now recognized as a critical regulator of many aspects of erythropoiesis (Perkins et al., 2016). The KLF1 gene knockout in the mouse results in embryonic lethality at around stages E14–E15 due to lethal anemia because of the inability to activate β-globin (Nuez et al., 1995; Perkins et al., 1995) and mutations in the CACCC box of the β-globin promoter that abolish its binding causing thalassemia (Feng et al., 1994). Thus, KLF1 promotes the switching from γ- to β-globin gene expression both directly, by activating the β-globin promoter and indirectly, by activating BCL11A. Finally (Basak et al., 2020), recently demonstrated that the oncofetal protein LIN28B (Piskounova et al., 2011; Shyh-Chang et al., 2013), already known to increase γ-globin expression (Lee et al., 2013, 2015; de Vasconcellos et al., 2014), blocks BCL11A translation. The failure of the above regulatory circuits converging on BCL11A results in elevated HbF.

SOX6 is a HMG box transcription factor, characterized by the presence of a high-mobility group domain (HMG) (Wegner, 1999). SOX6 is expressed in several tissues, including cartilage, testis, neural cells, and erythroblasts (Hagiwara, 2011). Mice with a chromosomal inversion (p^{100H}) disrupting the *Sox6* gene, or carrying a targeted inactivation of Sox6 die perinatally, secondary to cardiac or skeletal myopathy (Hagiwara et al., 2000). *Sox6*-null mouse fetuses and pups are anemic and have defective RBCs (Dumitriu et al., 2006). In erythroid cells, SOX6 has indeed a dual role: it stimulates erythroid cell survival, proliferation, and terminal maturation during definitive murine erythropoiesis (Cantu et al., 2011), and it directly silences embryo/fetal globin genes (Yi et al., 2006; Xu et al., 2010). The first aspect is mediated by the activation of SOCS3, whose overexpression recapitulates the proliferation arrest imposed by SOX6 (Cantu et al., 2011); the second requires the direct binding and repression of the embryonic εγ-globin promoter and the cooperation with BCL11A, via direct physical interaction, to silence γ-globin in adult erythroid cells (Yi et al., 2006; Xu et al., 2010).

The DRED complex (direct repeat erythroid-definitive) is a 540-kDa complex containing the nuclear orphan receptors

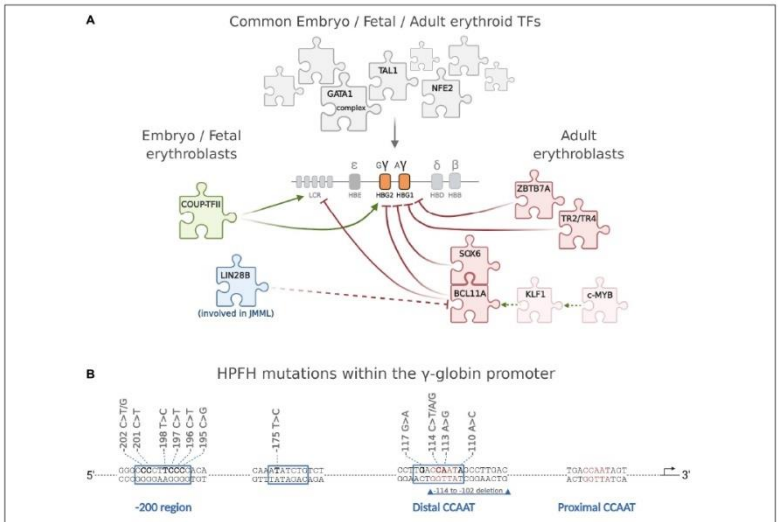


FIGURE 2 | The major transcription factors directly regulating the differential expression of γ-globin during development. **(A)** The expression of γ-globin in embryonic, fetal, and adult cells is regulated by a large common set of ubiquitous (such as, for example, NF-Y) and erythroid-specific transcription factors, the most important of which are GATA1 and its complexes (Love et al., 2014), NFE2, and TAL1. A very small number of TFs, discussed in the text, are instead directly involved in the time-specific expression of γ-globin and in its deregulation when it persists in the adult. c-MYB activates KLF1 that in turn activates BCL11A, which cooperates with SOX6 in repressing γ-globin. ZBTB7A and TR2/TR4 repress γ-globin independently from BCL11A. Of interest, whereas different γ-globin-specific repressors have been identified so far, little is known about early specific activation of embryo/fetal globin genes, and COUP-TFII is the only γ-globin-specific activator identified so far. The oncofetal LIN28B protein, expressed at high levels in JMML cells concomitantly expressing high γ-globin, blocks BCL11A translation. **(B)** Schematic representation of the γ-promoter showing the position of HPFH mutations. Created with BioRender.com.

TR2 and TR4, expressed in many tissues, including erythroid cells (Tanabe et al., 2002; Lin et al., 2017). TR2 and TR4 form homodimers or heterodimers binding to GGTCA repeat sequences with variable spacing, a consensus common to non-steroid nuclear receptors (Lee et al., 1998). The double conditional knockout of TR2 and TR4 in mouse erythroid cells results in increased embryonic εy and βh1 globins (Cui et al., 2015). In line with this result, the −117HPFH point mutation, associated with high HbF, reduces TR2/TR4 binding (Tanabe et al., 2002). Finally, TR2 and TR4 have been proposed directly repress GATA1 transcription, suggesting a wider role in erythroid maturation (Tanabe et al., 2007b).

ZBTB7A/LRF is a C2H2 zinc finger TF belonging to the POK (BTB/POZ and Krüppel) group of transcriptional regulators (Davies et al., 1999). ZBTB7A is expressed in various hematopoietic lineages (Maeda, 2016). However, its knockout shows a specific erythroid defect, with mouse embryos dying around E16.5 because of severe anemia, demonstrating that ZBTB7A is required for definitive erythropoiesis (Maeda et al., 2009). Adult-stage knockout of Zbtb7a results in erythropoietin-unresponsive macrocytic anemia, reversed by BIM knockout (Maeda et al., 2009). ZBTB7A specifically represses embryonic and fetal globin gene expression, independently from BCL11A, probably through the interaction with components of the nucleosome remodeling deacetylase (NuRD) complex (Masuda et al., 2016).

COUP-TFII, THE SELECTIVE ACTIVATOR OF γ-GLOBIN IN YOLK SAC-DERIVED CELLS

The COUP-TFII gene (Chicken Ovalbumin Upstream Promoter Transcriptional Factor II, also known as NR2F2/ARP1) encodes for an orphan nuclear receptor. Its expression is high in the mesenchymal component of developing organs

and overall decreases after the completion of organogenesis (Pereira et al., 1995; Lin et al., 2011). Its knockout results in early embryonic lethality (E9.5–E10) (Pereira et al., 1999) caused by defects in angiogenesis and heart development. In the erythroid lineage, Coup-TFII is expressed in the early embryo in the YS and in FL, and it declines around day E12.5 (Filipe et al., 1999; Cui et al., 2015; Fugazza et al., 2020). Although originally identified as ε- and γ-globin gene repressor on the basis of *in vitro* binding data, its functional role remained elusive (Ronchi et al., 1995; Liberati et al., 2001). Our group recently demonstrated that COUP-TFII is co-expressed with embryonic globin genes in cells of YS origin, where it acts as specific γ-globin activator by binding to the GGTCA motifs present within the β-locus (Fugazza et al., 2020). This last observation opens many questions since the same consensus is bound by BCL11A (Liu et al., 2018) and possibly by TR2/TR4 (Tanabe et al., 2007a) in adult cells. Importantly, COUP-TFII is able to activate γ-globin when overexpressed in adult cells, suggesting once again a very similar cellular environment of fetal vs. adult cells (Fugazza et al., 2020).

INAPPROPRIATE TIMING OF γ-GLOBIN EXPRESSION IN "ADULT-TYPE" CELLS: HEREDITARY PERSISTENCE OF FETAL HEMOGLOBIN

HPFH is a benign condition in which γ-globin remains expressed at high levels in adult life (Forget, 1998). HbF can be high in all red blood cells (pancellular HPFH) or restricted to a small subset of erythroid cells (heterocellular HPFH) (Thein et al., 2009). HPFHs, based on the different types of causative mutations, can be broadly divided in three main categories: deletional HPFH, non-deletional HPFH, and HPFH non-linked with the β-locus. Deletional HPFH is associated with the deletion of large regions of DNA between the γ- and β-globin genes within the β-globin locus, as it happens, for example, in the Sicilian ≈13 kb and in the Italian ≈40-kb deletions[1] (Kountouris et al., 2014). Many deletions include the loss of δ- and β-globin genes resulting in $(\delta\beta)^0$-thalessemia and HPFH (Ottolenghi et al., 1982). The molecular mechanism underlying the elevated HbF is complex and involves the concomitant deletion of the β promoter, which removes its competition with the γ promoters for the upstream LCR and for limiting TFs (Forget, 1998). Non-deletional HPFHs are caused by point mutations described in both γ-globin promoters. These mutations fall into three distinct clusters: the −200 region, the −175 site, and the distal CCAAT box region, around −115 (Forget, 1998; Martyn et al., 2018), where they either disrupt binding sites for γ-globin repressors or create *de novo* binding sites for γ-globin activators. For example, the −196 mutation abolishes the binding of ZBTB7A (Martyn et al., 2018), whereas mutations in the distal CCAAT box region impair BCL11A binding (Liu et al., 2018; Martyn et al., 2018). Instead, the −175T > C mutation creates a *de novo* binding site for the activator TAL1 (Wienert et al., 2015). Despite these evidences, the function of these sequences is more complex. The −198T > C mutation, for example, although being located within the cluster of HPFH mutations that impair the binding of ZBTB7A, is a gain-of-function mutation creating a binding site for the erythroid activator KLF1 (Wienert et al., 2017). The same is true for the mutation at position −113 A > G, which, although lying within the −115 region bound by the repressor BCL11A, creates a new binding site for the activator GATA1, without altering the BCL11A binding (Martyn et al., 2019).

[1]https://www.ithanet.eu/

THE ROLE OF MODIFIERS LOCI

Genome-wide association studies (GWAS) have revealed other two loci, not linked to the β-locus, consistently associated with HbF levels, and with β-globin disorder severity, across various ethnic backgrounds: a region on 2p (Menzel et al., 2007; Lettre et al., 2008; Uda et al., 2008) and the HBS1L-MYB intergenic region on 6q (Craig et al., 1996; Thein et al., 2007). The 2p region turned out to correspond to BCL11A, and the fine mapping of the single nucleotide polymorphisms associated to HbF within this region led to the identification of the intronic enhancer driving the expression of BCL11A in erythroid cells (Bauer et al., 2013; Canver et al., 2015).

The variants within the HBS1L-MYB intergenic region on 6q impair the binding of LDB1, GATA1, TAL1, and KLF1 to the enhancer controlling *c-MYB* expression (Stadhouders et al., 2014). c-Myb is the cellular homolog of *v-Myb*, the avian retroviral oncogene causing myelomas and lymphomas in birds (Wolff, 1996). Of the two major isoforms, isoform 2 (72 kDa) is the dominant one in human erythroid cells (Baker et al., 2010; Wang et al., 2018). In hematopoiesis, *c-MYB* is expressed in immature cells of all hematopoietic lineages (Wang et al., 2018); in erythropoiesis, it is required for the expansion of erythroid progenitors and must be downregulated to allow differentiation (Emambokus et al., 2003). c-Myb-null murine embryos are normal until E13.5, but by E15, they become severely anemic and die, suggesting that c-Myb is required for definitive erythropoiesis (Mucenski et al., 1991). The reduced c-MYB level has a twofold impact on globin genes. Low c-MYB levels, by accelerating the kinetics of erythroid differentiation, would favor the release of early erythroid progenitor cells still synthetizing HbF (Stamatoyannopoulos, 2005; Jiang et al., 2006). In addition, the reduced activation of KLF1 (Bianchi et al., 2010; Suzuki et al., 2013) by MYB would promote γ-globin expression by reducing BCL11A levels.

INAPPROPRIATE TIMING OF γ-GLOBIN EXPRESSION IN "FETAL-TYPE" CELLS PERSISTING AFTER BIRTH: THE CASE OF JUVENILE MYELOMONOCYTIC LEUKEMIA

High HbF levels are a hallmark of different leukemias (Sheridan et al., 1976). However, the expression of γ-globin in juvenile myelomonocytic leukemia (JMML) is peculiar. JMML is a rare and aggressive blood cancer of early childhood

(Loh, 2011; Niemeyer and Flotho, 2019). About 90% of the patients present hyperactivation of the RAS pathway, as a result of mutations in KRAS, NRAS, PTPN11, NF1, or CBL genes, and about 25% of the patients carry chromosome 7 monosomy (Flotho et al., 1999; de Vries et al., 2010). JMML is considered a stem cell disease (Inoue et al., 1987; Busque et al., 1995; Flotho et al., 1999; Cooper et al., 2000). Of interest, increased HbF levels and the presence of fetal red cell traits (Weinberg et al., 1990; de Vries et al., 2010; Helsmoortel et al., 2016a) are present in more than half of JMML patients. These evidences suggested a fetal origin for JMML, confirmed by the retrospective analysis of JMML patient samples collected at birth (Kratz et al., 2005; Matsuda et al., 2010; Stieglitz et al., 2015). However, yolk sac EMPs expressing gain of function PTPN11 mutations recapitulate part of the characteristics of JMML, but they are not able to cause disease in mice (Tarnawsky et al., 2017). Recently, gene expression profiling of JMML samples identified a subgroup characterized by high LIN28B expression and higher HbF levels (Helsmoortel et al., 2016b). LIN28B is an oncofetal protein (Shyh-Chang and Daley, 2013) that induces γ-globin expression (Lee et al., 2013, 2015; de Vasconcellos et al., 2014) and is highly expressed in fetal HSCs (Copley et al., 2013). Interestingly, LIN28B was shown to repress BCL11A-XL by blocking its translation (Basak et al., 2020). Thus, high levels of LIN28B decrease the amount of BCL11A protein, and this results in high HbF. This observation suggests a mechanistic link between LIN28B and high HbF observed in JMML. Whether the 50% of JMML with high HbF arose from a fetal EMP subpopulation expressing high LIN28B deserves further investigation.

CONCLUSION

The existence of γ-globin genes is specific to humans and old-world monkeys. Physiologically, γ-globin expression is confined to fetal life, tightly regulated by a complex network of transcription factors (activators and repressors) and co-regulators. Nevertheless, normal subjects present rare F cells, whose nature is still unclear (Rochette et al., 1994). During development, γ-globin genes are expressed in two different types of erythroblasts, with independent origin, the first originating from the transient, yolk sac-derived EMP population and the second from definitive HSCs (McGrath et al., 2011). These two cell types are very similar in their maturational pathway, and adult cells indeed represent a permissive environment for γ-globin expression, as shown by HPFH. Here, alterations in the few specific transcription factors that regulate γ-globin transcription (or in the sequences bound by them within the HBB locus) allow substantial γ-globin expression. Moreover, recent RNA-seq studies on A and F cells from the same healthy donors show that these cells do not significantly differ in the expression of known γ-globin regulators, suggesting that differences in the transcription of globin genes within the HBB locus itself account for the HbF trait (Khandros et al., 2020).

However, new evidence from the study of JMML unveil an alternative possible scenario, where γ-globin genes could be the marker of the fetal origin of these leukemic cells. The discovery that fetal progenitor-derived cells can persist to adulthood and contribute to disease raises the possibility that, especially in childhood malignancies, such as JMML, HbF can indeed be the signature of the fetal origin of cancer cells.

AUTHOR'S NOTE

While this minireview was in press, an article was published by Liu and colleagues (doi: 10.1038/s41588-021-00798-y) showing that BCL11A competes with NF-Y binding to initiate γ-globin repression at the CCAAT box region.

AUTHOR CONTRIBUTIONS

AR conceived and wrote the manuscript. GB and AL contributed with ideas and discussion. SS and AA contributed in the organization of the manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

This project has received funding from the European Union's Horizon 2020 Research and Innovation Program under the Marie Skłodowska Curie grant agreement no. 813091 (ARCH, Age-related changes in hematopoiesis).

REFERENCES

Aisen, P. (2004). Transferrin receptor 1. *Int. J. Biochem. Cell Biol.* 36, 2137–2143.

Avram, D., Fields, A., Pretty On Top, K., Nevrivy, D. J., Ishmael, J. E., and Leid, M. (2000). Isolation of a novel family of C(2)H(2) zinc finger proteins implicated in transcriptional repression mediated by chicken ovalbumin upstream promoter transcription factor (COUP-TF) orphan nuclear receptors. *J. Biol. Chem.* 275, 10315–10322. doi: 10.1074/jbc.275.14.10315

Baker, S. J., Kumar, A., and Reddy, E. P. (2010). p89c-Myb is not required for fetal or adult hematopoiesis. *Genesis* 48, 309–316.

Barbarani, G., Fugazza, C., Strouboulis, J., and Ronchi, A. E. (2019). The pleiotropic effects of GATA1 and KLF1 in physiological erythropoiesis and in dyserythropoietic disorders. *Front. Physiol.* 10:91. doi: 10.3389/fphys.2019.00091

Baron, M. H. (2013). Concise review: early embryonic erythropoiesis: not so primitive after all. *Stem Cells* 31, 849–856. doi: 10.1002/stem.1342

Basak, A., Hancarova, M., Ulirsch, J. C., Balci, T. B., Trkova, M., Pelisek, M., et al. (2015). BCL11A deletions result in fetal hemoglobin persistence and neurodevelopmental alterations. *J. Clin. Invest.* 125, 2363–2368. doi: 10.1172/jci81163

Basak, A., Munschauer, M., Lareau, C. A., Montbleau, K. E., Ulirsch, J. C., Hartigan, C. R., et al. (2020). Control of human hemoglobin switching by LIN28B-mediated regulation of BCL11A translation. *Nat. Genet.* 52, 138–145. doi: 10.1038/s41588-019-0568-7

Bauer, D. E., Kamran, S. C., Lessard, S., Xu, J., Fujiwara, Y., Lin, C., et al. (2013). An erythroid enhancer of BCL11A subject to genetic variation determines fetal hemoglobin level. *Science* 342, 253–257. doi: 10.1126/science.1242088

Bianchi, E., Zini, R., Salati, S., Tenedini, E., Norfo, R., Tagliafico, E., et al. (2010). c-myb supports erythropoiesis through the transactivation of KLF1 and LMO2 expression. *Blood* 116, e99–e110.

Borg, J., Papadopoulos, P., Georgitsi, M., Gutierrez, L., Grech, G., Fanis, P., et al. (2010). Haploinsufficiency for the erythroid transcription factor KLF1 causes hereditary persistence of fetal hemoglobin. *Nat. Genet.* 42, 801–805. doi: 10.1038/ng.630

Boyer, S. H., Belding, T. K., Margolet, L., and Noyes, A. N. (1975). Fetal hemoglobin restriction to a few erythrocytes (F cells) in normal human adults. *Science* 188, 361–363. doi: 10.1126/science.804182

Busque, L., Gilliland, D. G., Prchal, J. T., Sieff, C. A., Weinstein, H. J., Sokol, J. M., et al. (1995). Clonality in juvenile chronic myelogenous leukemia. *Blood* 85, 21–30. doi: 10.1182/blood.v85.1.21.bloodjournal85121

Cantu, C., Ierardi, R., Alborelli, I., Fugazza, C., Cassinelli, L., Piconese, S., et al. (2011). Sox6 enhances erythroid differentiation in human erythroid progenitors. *Blood* 117, 3669–3679. doi: 10.1182/blood-2010-04-282350

Canver, M. C., Smith, E. C., Sher, F., Pinello, L., Sanjana, N. E., Shalem, O., et al. (2015). BCL11A enhancer dissection by Cas9-mediated in situ saturating mutagenesis. *Nature* 527, 192–197. doi: 10.1038/nature15521

Chen, K., Liu, J., Heck, S., Chasis, J. A., An, X., and Mohandas, N. (2009). Resolving the distinct stages in erythroid differentiation based on dynamic changes in membrane protein expression during erythropoiesis. *Proc. Natl. Acad. Sci. U.S.A.* 106, 17413–17418. doi: 10.1073/pnas.0909296106

Cooper, L. J., Shannon, K. M., Loken, M. R., Weaver, M., Stephens, K., and Sievers, E. L. (2000). Evidence that juvenile myelomonocytic leukemia can arise from a pluripotential stem cell. *Blood* 96, 2310–2313. doi: 10.1182/blood.v96.6.2310.h8002310_2310_2313

Copley, M. R., Babovic, S., Benz, C., Knapp, D. J., Beer, P. A., Kent, D. G., et al. (2013). The Lin28b-let-7-Hmga2 axis determines the higher self-renewal potential of fetal haematopoietic stem cells. *Nat. Cell Biol.* 15, 916–925. doi: 10.1038/ncb2783

Craig, J. E., Rochette, J., Fisher, C. A., Weatherall, D. J., Marc, S., Lathrop, G. M., et al. (1996). Dissecting the loci controlling fetal haemoglobin production on chromosomes 11p and 6q by the regressive approach. *Nat. Genet.* 12, 58–64. doi: 10.1038/ng0196-58

Cui, S., Tanabe, O., Sierant, M., Shi, L., Campbell, A., Lim, K. C., et al. (2015). Compound loss of function of nuclear receptors Tr2 and Tr4 leads to induction of murine embryonic beta-type globin genes. *Blood* 125, 1477–1487. doi: 10.1182/blood-2014-10-605022

Davies, J. M., Hawe, N., Kabarowski, J., Huang, Q. H., Zhu, J., Brand, N. J., et al. (1999). Novel BTB/POZ domain zinc-finger protein, LRF, is a potential target of the LAZ-3/BCL-6 oncogene. *Oncogene* 18, 365–375. doi: 10.1038/sj.onc.1202332

de Vasconcellos, J. F., Fasano, R. M., Lee, Y. T., Kaushal, M., Byrnes, C., Meier, E. R., et al. (2014). LIN28A expression reduces sickling of cultured human erythrocytes. *PLoS One* 9:e106924. doi: 10.1371/journal.pone.0106924

de Vries, A. C., Zwaan, C. M., and van den Heuvel-Eibrink, M. M. (2010). Molecular basis of juvenile myelomonocytic leukemia. *Haematologica* 95, 179–182. doi: 10.3324/haematol.2009.016865

Dumitriu, B., Patrick, M. R., Petschek, J. P., Cherukuri, S., Klingmuller, U., Fox, P. L., et al. (2006). Sox6 cell-autonomously stimulates erythroid cell survival, proliferation, and terminal maturation and is thereby an important enhancer of definitive erythropoiesis during mouse development. *Blood* 108, 1198–1207. doi: 10.1182/blood-2006-02-004184

Dzierzak, E., and Bigas, A. (2018). Blood development: hematopoietic stem cell dependence and independence. *Cell Stem Cell* 22, 639–651. doi: 10.1016/j.stem.2018.04.015

Dzierzak, E., and Philipsen, S. (2013). Erythropoiesis: development and differentiation. *Cold Spring Harb. Perspect. Med.* 3:a011601. doi: 10.1101/cshperspect.a011601

Emambokus, N., Vegiopoulos, A., Harman, B., Jenkinson, E., Anderson, G., and Frampton, J. (2003). Progression through key stages of haemopoiesis is dependent on distinct threshold levels of c-Myb. *EMBO J.* 22, 4478–4488. doi: 10.1093/emboj/cdg434

Epelman, S., Lavine, K. J., Beaudin, A. E., Sojka, D. K., Carrero, J. A., Calderon, B., et al. (2014). Embryonic and adult-derived resident cardiac macrophages are maintained through distinct mechanisms at steady state and during inflammation. *Immunity* 40, 91–104. doi: 10.1016/j.immuni.2013.11.019

Feng, W. C., Southwood, C. M., and Bieker, J. J. (1994). Analyses of beta-thalassemia mutant DNA interactions with erythroid Kruppel-like factor (EKLF), an erythroid cell-specific transcription factor. *J. Biol. Chem.* 269, 1493–1500. doi: 10.1016/s0021-9258(17)42283-6

Ferreira, R., Ohneda, K., Yamamoto, M., and Philipsen, S. (2005). GATA1 function, a paradigm for transcription factors in hematopoiesis. *Mol. Cell Biol.* 25, 1215–1227. doi: 10.1128/mcb.25.4.1215-1227.2005

Filipe, A., Li, Q., Deveaux, S., Godin, I., Romeo, P. H., Stamatoyannopoulos, G., et al. (1999). Regulation of embryonic/fetal globin genes by nuclear hormone receptors: a novel perspective on hemoglobin switching. *EMBO J.* 18, 687–697. doi: 10.1093/emboj/18.3.687

Flotho, C., Valcamonica, S., Mach-Pascual, S., Schmahl, G., Corral, L., Ritterbach, J., et al. (1999). RAS mutations and clonality analysis in children with juvenile myelomonocytic leukemia (JMML). *Leukemia* 13, 32–37. doi: 10.1038/sj.leu.2401240

Forget, B. G. (1990). Developmental control of human globin gene expression. *Prog. Clin. Biol. Res.* 352, 313–322.

Forget, B. G. (1998). Molecular basis of hereditary persistence of fetal hemoglobin. *Ann. N. Y. Acad. Sci.* 850, 38–44. doi: 10.1111/j.1749-6632.1998.tb10460.x

Fugazza, C., Barbarani, G., Elangovan, S., Marini, M. G., Giolitto, S., Font-Monclus, I., et al. (2020). The Coup-TFII orphan nuclear receptor is an activator of the gamma-globin gene. *Haematologica* 106, 474–482. doi: 10.3324/haematol.2019.241224

Funnell, A. P., Prontera, P., Ottaviani, V., Piccione, M., Giambona, A., Maggio, A., et al. (2015). 2p15-p16.1 microdeletions encompassing and proximal to BCL11A are associated with elevated HbF in addition to neurologic impairment. *Blood* 126, 89–93. doi: 10.1182/blood-2015-04-638528

Gasiorek, J. J., and Blank, V. (2015). Regulation and function of the NFE2 transcription factor in hematopoietic and non-hematopoietic cells. *Cell. Mol. Life Sci.* 72, 2323–2335. doi: 10.1007/s00018-015-1866-6

Gomez Perdiguero, E., Klapproth, K., Schulz, C., Busch, K., Azzoni, E., Crozet, L., et al. (2015). Tissue-resident macrophages originate from yolk-sac-derived erythro-myeloid progenitors. *Nature* 518, 547–551. doi: 10.1038/nature13989

Greig, L. C., Woodworth, M. B., Greppi, C., and Macklis, J. D. (2016). Ctip1 controls acquisition of sensory area identity and establishment of sensory input fields in the developing neocortex. *Neuron* 90, 261–277. doi: 10.1016/j.neuron.2016.03.008

Grosveld, F., van Assendelft, G. B., Greaves, D. R., and Kollias, G. (1987). Position-independent, high-level expression of the human beta-globin gene in transgenic mice. *Cell* 51, 975–985. doi: 10.1016/0092-8674(87)90584-8

Hagiwara, N. (2011). Sox6, jack of all trades: a versatile regulatory protein in vertebrate development. *Dev. Dyn.* 240, 1311–1321. doi: 10.1002/dvdy.22639

Hagiwara, N., Klewer, S. E., Samson, R. A., Erickson, D. T., Lyon, M. F., and Brilliant, M. H. (2000). Sox6 is a candidate gene for p100H myopathy, heart block, and sudden neonatal death. *Proc. Natl. Acad. Sci. U.S.A.* 97, 4180–4185. doi: 10.1073/pnas.97.8.4180

Helsmoortel, H. H., Bresolin, S., Lammens, T., Cave, H., Noellke, P., Caye, A., et al. (2016a). LIN28B overexpression defines a novel fetal-like subgroup of juvenile myelomonocytic leukemia. *Blood* 127, 1163–1172. doi: 10.1182/blood-2015-09-667808

Helsmoortel, H. H., De Moerloose, B., Pieters, T., Ghazavi, F., Bresolin, S., Cave, H., et al. (2016b). LIN28B is over-expressed in specific subtypes of pediatric leukemia and regulates lncRNA H19. *Haematologica* 101, e240–e244.

Higgs, D. R., Wood, W. G., Jarman, A. P., Sharpe, J., Lida, J., Pretorius, I. M., et al. (1990). A major positive regulatory region located far upstream of the human alpha-globin gene locus. *Genes Dev.* 4, 1588–1601. doi: 10.1101/gad.4.9.1588

Inoue, S., Shibata, T., Ravindranath, Y., and Gohle, N. (1987). Clonal origin of erythroid cells in juvenile chronic myelogenous leukemia. *Blood* 69, 975–976. doi: 10.1182/blood.v69.3.975.bloodjournal693975

Ippolito, G. C., Dekker, J. D., Wang, Y. H., Lee, B. K., Shaffer, A. L. III, Lin, J., et al. (2014). Dendritic cell fate is determined by BCL11A. *Proc. Natl. Acad. Sci. U.S.A.* 111, E998–E1006.

Jiang, J., Best, S., Menzel, S., Silver, N., Lai, M. I., Surdulescu, G. L., et al. (2006). cMYB is involved in the regulation of fetal hemoglobin production in adults. *Blood* 108, 1077–1083. doi: 10.1182/blood-2006-01-008912

Kang, Y., Kim, Y. W., Yun, J., Shin, J., and Kim, A. (2015). KLF1 stabilizes GATA-1 and TAL1 occupancy in the human beta-globin locus. *Biochim. Biophys. Acta* 1849, 282–289. doi: 10.1016/j.bbagrm.2014.12.010

Katsumura, K. R., Bresnick, E. H., and Group, G. F. M. (2017). The GATA factor revolution in hematology. *Blood* 129, 2092–2102. doi: 10.1182/blood-2016-09-687871

Khaled, W. T., Choon Lee, S., Stingl, J., Chen, X., Raza Ali, H., Rueda, O. M., et al. (2015). BCL11A is a triple-negative breast cancer gene with critical functions in stem and progenitor cells. *Nat. Commun.* 6:5987.

Khandros, E., Huang, P., Peslak, S. A., Sharma, M., Abdulmalik, O., Giardine, B. M., et al. (2020). Understanding heterogeneity of fetal hemoglobin induction through comparative analysis of F and A erythroblasts. *Blood* 135, 1957–1968. doi: 10.1182/blood.2020005058

Kim, Y. W., Yun, W. J., and Kim, A. (2016). Erythroid activator NF-E2, TAL1 and KLF1 play roles in forming the LCR HSs in the human adult beta-globin locus. *Int. J. Biochem. Cell Biol.* 75, 45–52. doi: 10.1016/j.biocel.2016.03.013

Kina, T., Ikuta, K., Takayama, E., Wada, K., Majumdar, A. S., Weissman, I. L., et al. (2000). The monoclonal antibody TER-119 recognizes a molecule associated with glycophorin A and specifically marks the late stages of murine erythroid lineage. *Br. J. Haematol.* 109, 280–287. doi: 10.1046/j.1365-2141.2000.02037.x

Kountouris, P., Lederer, C. W., Fanis, P., Feleki, X., Old, J., and Kleanthous, M. (2014). IthaGenes: an interactive database for haemoglobin variations and epidemiology. *PLoS One* 9:e103020. doi: 10.1371/journal.pone.0103020

Kratz, C. P., Niemeyer, C. M., Castleberry, R. P., Cetin, M., Bergstrasser, E., Emanuel, P. D., et al. (2005). The mutational spectrum of PTPN11 in juvenile myelomonocytic leukemia and Noonan syndrome/myeloproliferative disease. *Blood* 106, 2183–2185. doi: 10.1182/blood-2005-02-0531

Lacaud, G., and Kouskoff, V. (2017). Hemangioblast, hemogenic endothelium, and primitive versus definitive hematopoiesis. *Exp. Hematol.* 49, 19–24. doi: 10.1016/j.exphem.2016.12.009

Lee, C. H., Chinpaisal, C., and Wei, L. N. (1998). A novel nuclear receptor heterodimerization pathway mediated by orphan receptors TR2 and TR4. *J. Biol. Chem.* 273, 25209–25215. doi: 10.1074/jbc.273.39.25209

Lee, Y. T., de Vasconcellos, J. F., Byrnes, C., Kaushal, M., Rabel, A., Tumburu, L., et al. (2015). Erythroid-specific expression of LIN28A is sufficient for robust gamma-globin gene and protein expression in adult erythroblasts. *PLoS One* 10:e0144977. doi: 10.1371/journal.pone.0144977

Lee, Y. T., de Vasconcellos, J. F., Yuan, J., Byrnes, C., Noh, S. J., Meier, E. R., et al. (2013). LIN28B-mediated expression of fetal hemoglobin and production of fetal-like erythrocytes from adult human erythroblasts ex vivo. *Blood* 122, 1034–1041. doi: 10.1182/blood-2012-12-472308

Lettre, G., Sankaran, V. G., Bezerra, M. A., Araujo, A. S., Uda, M., Sanna, S., et al. (2008). DNA polymorphisms at the BCL11A, HBS1L-MYB, and beta-globin loci associate with fetal hemoglobin levels and pain crises in sickle cell disease. *Proc. Natl. Acad. Sci. U.S.A.* 105, 11869–11874. doi: 10.1073/pnas.0804799105

Liberati, C., Cera, M. R., Secco, P., Santoro, C., Mantovani, R., Ottolenghi, S., et al. (2001). Cooperation and competition between the binding of COUP-TFII and NF-Y on human epsilon- and gamma-globin gene promoters. *J. Biol. Chem.* 276, 41700–41709. doi: 10.1074/jbc.m102987200

Liberati, C., Ronchi, A., Lievens, P., Ottolenghi, S., and Mantovani, R. (1998). NF-Y organizes the gamma-globin CCAAT boxes region. *J. Biol. Chem.* 273, 16880–16889. doi: 10.1074/jbc.273.27.16880

Lin, F. J., Qin, J., Tang, K., Tsai, S. Y., and Tsai, M. J. (2011). Coup d'Etat: an orphan takes control. *Endocr. Rev.* 32, 404–421. doi: 10.1210/er.2010-0021

Lin, S. J., Yang, D. R., Yang, G., Lin, C. Y., Chang, H. C., Li, G., et al. (2017). TR2 and TR4 orphan nuclear receptors: an overview. *Curr. Top. Dev. Biol.* 125, 357–373. doi: 10.1016/bs.ctdb.2017.02.002

Liu, H., Ippolito, G. C., Wall, J. K., Niu, T., Probst, L., Lee, B. S., et al. (2006). Functional studies of BCL11A: characterization of the conserved BCL11A-XL splice variant and its interaction with BCL6 in nuclear paraspeckles of germinal center B cells. *Mol. Cancer* 5:18.

Liu, N., Hargreaves, V. V., Zhu, Q., Kurland, J. V., Hong, J., Kim, W., et al. (2018). Direct promoter repression by BCL11A controls the fetal to adult hemoglobin switch. *Cell* 173, 430–442.e17.

Liu, P., Keller, J. R., Ortiz, M., Tessarollo, L., Rachel, R. A., Nakamura, T., et al. (2003). Bcl11a is essential for normal lymphoid development. *Nat. Immunol.* 4, 525–532. doi: 10.1038/ni925

Loh, M. L. (2011). Recent advances in the pathogenesis and treatment of juvenile myelomonocytic leukaemia. *Br. J. Haematol.* 152, 677–687. doi: 10.1111/j.1365-2141.2010.08525.x

Love, P. E., Warzecha, C., and Li, L. (2014). Ldb1 complexes: the new master regulators of erythroid gene transcription. *Trends Genet.* 30, 1–9. doi: 10.1016/j.tig.2013.10.001

Luc, S., Huang, J., McEldoon, J. L., Somuncular, E., Li, D., Rhodes, C., et al. (2016). Bcl11a deficiency leads to hematopoietic stem cell defects with an aging-like phenotype. *Cell Rep.* 16, 3181–3194. doi: 10.1016/j.celrep.2016.08.064

Maeda, T. (2016). Regulation of hematopoietic development by ZBTB transcription factors. *Int. J. Hematol.* 104, 310–323. doi: 10.1007/s12185-016-2035-x

Maeda, T., Ito, K., Merghoub, T., Poliseno, L., Hobbs, R. M., Wang, G., et al. (2009). LRF is an essential downstream target of GATA1 in erythroid development and regulates BIM-dependent apoptosis. *Dev. Cell* 17, 527–540. doi: 10.1016/j.devcel.2009.09.005

Martyn, G. E., Quinlan, K. G. R., and Crossley, M. (2017). The regulation of human globin promoters by CCAAT box elements and the recruitment of NF-Y. *Biochim. Biophys. Acta Gene Regul. Mech.* 1860, 525–536. doi: 10.1016/j.bbagrm.2016.10.002

Martyn, G. E., Wienert, B., Kurita, R., Nakamura, Y., Quinlan, K. G. R., and Crossley, M. (2019). A natural regulatory mutation in the proximal promoter elevates fetal globin expression by creating a de novo GATA1 site. *Blood* 133, 852–856. doi: 10.1182/blood-2018-07-863951

Martyn, G. E., Wienert, B., Yang, L., Shah, M., Norton, L. J., Burdach, J., et al. (2018). Natural regulatory mutations elevate the fetal globin gene via disruption of BCL11A or ZBTB7A binding. *Nat. Genet.* 50, 498–503. doi: 10.1038/s41588-018-0085-0

Masuda, T., Wang, X., Maeda, M., Canver, M. C., Sher, F., Funnell, A. P., et al. (2016). Transcription factors LRF and BCL11A independently repress expression of fetal hemoglobin. *Science* 351, 285–289. doi: 10.1126/science.aad3312

Matsuda, K., Sakashita, K., Taira, C., Tanaka-Yanagisawa, M., Yanagisawa, R., Shiohara, M., et al. (2010). Quantitative assessment of PTPN11 or RAS mutations at the neonatal period and during the clinical course in patients with juvenile myelomonocytic leukaemia. *Br. J. Haematol.* 148, 593–599. doi: 10.1111/j.1365-2141.2009.07968.x

McGrath, K. E., Frame, J. M., Fromm, G. J., Koniski, A. D., Kingsley, P. D., Little, J., et al. (2011). A transient definitive erythroid lineage with unique regulation of the beta-globin locus in the mammalian embryo. *Blood* 117, 4600–4608. doi: 10.1182/blood-2010-12-325357

Menzel, S., Garner, C., Gut, I., Matsuda, F., Yamaguchi, M., Heath, S., et al. (2007). A QTL influencing F cell production maps to a gene encoding a zinc-finger protein on chromosome 2p15. *Nat. Genet.* 39, 1197–1199. doi: 10.1038/ng2108

Miller, I. J., and Bieker, J. J. (1993). A novel, erythroid cell-specific murine transcription factor that binds to the CACCC element and is related to the Kruppel family of nuclear proteins. *Mol. Cell Biol.* 13, 2776–2786. doi: 10.1128/mcb.13.5.2776

Mucenski, M. L., Mclain, K., Kier, A. B., Swerdlow, S. H., Schreiner, C. M., Miller, T. A., et al. (1991). A Functional C-Myb gene is required for normal murine fetal hepatic hematopoiesis. *Cell* 65, 677–689. doi: 10.1016/0092-8674(91)90099-k

Niemeyer, C. M., and Flotho, C. (2019). Juvenile myelomonocytic leukemia: who's the driver at the wheel? *Blood* 133, 1060–1070. doi: 10.1182/blood-2018-11-844688

Nuez, B., Michalovich, D., Bygrave, A., Ploemacher, R., and Grosveld, F. (1995). Defective haematopoiesis in fetal liver resulting from inactivation of the EKLF gene. *Nature* 375, 316–318. doi: 10.1038/375316a0

Orkin, S. H., and Zon, L. I. (2008). Hematopoiesis: an evolving paradigm for stem cell biology. *Cell* 132, 631–644. doi: 10.1016/j.cell.2008.01.025

Ottolenghi, S., Giglioni, B., Taramelli, R., Comi, P., and Gianni, A. M. (1982). delta beta-Thalassemia and HPFH. *Birth Defects Orig. Artic. Ser.* 18, 65–67.
Palis, J. (2014). Primitive and definitive erythropoiesis in mammals. *Front. Physiol.* 5:3. doi: 10.3389/fphys.2014.00003
Palstra, R. J., Tolhuis, B., Splinter, E., Nijmeijer, R., Grosveld, F., and de Laat, W. (2003). The beta-globin nuclear compartment in development and erythroid differentiation. *Nat. Genet.* 35, 190–194. doi: 10.1038/ng1244
Peiris, H., Park, S., Louis, S., Gu, X., Lam, J. Y., Asplund, O., et al. (2018). Discovering human diabetes-risk gene function with genetics and physiological assays. *Nat. Commun.* 9:3855.
Pereira, F. A., Qiu, Y., Tsai, M. J., and Tsai, S. Y. (1995). Chicken ovalbumin upstream promoter transcription factor (COUP-TF): expression during mouse embryogenesis. *J. Steroid Biochem. Mol. Biol.* 53, 503–508. doi: 10.1016/0960-0760(95)00097-j
Pereira, F. A., Qiu, Y., Zhou, G., Tsai, M. J., and Tsai, S. Y. (1999). The orphan nuclear receptor COUP-TFII is required for angiogenesis and heart development. *Genes Dev.* 13, 1037–1049. doi: 10.1101/gad.13.8.1037
Perkins, A., Xu, X., Higgs, D. R., Patrinos, G. P., Arnaud, L., Bieker, J. J., et al. (2016). Kruppeling erythropoiesis: an unexpected broad spectrum of human red blood cell disorders due to KLF1 variants. *Blood* 127, 1856–1862. doi: 10.1182/blood-2016-01-694331
Perkins, A. C., Sharpe, A. H., and Orkin, S. H. (1995). Lethal beta-thalassaemia in mice lacking the erythroid CACCC-transcription factor EKLF. *Nature* 375, 318–322. doi: 10.1038/375318a0
Peterson, K. R., Li, Q. L., Clegg, C. H., Furukawa, T., Navas, P. A., Norton, E. J., et al. (1995). Use of yeast artificial chromosomes (YACs) in studies of mammalian development: production of beta-globin locus YAC mice carrying human globin developmental mutants. *Proc. Natl. Acad. Sci. U.S.A.* 92, 5655–5659. doi: 10.1073/pnas.92.12.5655
Philipsen, S., and Hardison, R. C. (2018). Evolution of hemoglobin loci and their regulatory elements. *Blood Cells Mol. Dis.* 70, 2–12. doi: 10.1016/j.bcmd.2017.08.001
Piskounova, E., Polytarchou, C., Thornton, J. E., LaPierre, R. J., Pothoulakis, C., Hagan, J. P., et al. (2011). Lin28A and Lin28B inhibit let-7 microRNA biogenesis by distinct mechanisms. *Cell* 147, 1066–1079. doi: 10.1016/j.cell.2011.10.039
Rochette, J., Craig, J. E., and Thein, S. L. (1994). Fetal hemoglobin levels in adults. *Blood Rev.* 8, 213–224. doi: 10.1016/0268-960x(94)90109-0
Ronchi, A. E., Bottardi, S., Mazzucchelli, C., Ottolenghi, S., and Santoro, C. (1995). Differential binding of the NFE3 and CP1/NFY transcription factors to the human gamma- and epsilon-globin CCAAT boxes. *J. Biol. Chem.* 270, 21934–21941. doi: 10.1074/jbc.270.37.21934
Sankaran, V. G., Menne, T. F., Xu, J., Akie, T. E., Lettre, G., Van Handel, B., et al. (2008). Human fetal hemoglobin expression is regulated by the developmental stage-specific repressor BCL11A. *Science* 322, 1839–1842. doi: 10.1126/science.1165409
Sankaran, V. G., Xu, J., and Orkin, S. H. (2010a). Advances in the understanding of haemoglobin switching. *Br. J. Haematol.* 149, 181–194. doi: 10.1111/j.1365-2141.2010.08105.x
Sankaran, V. G., Xu, J., and Orkin, S. H. (2010b). Transcriptional silencing of fetal hemoglobin by BCL11A. *Ann. N. Y. Acad. Sci.* 1202, 64–68. doi: 10.1111/j.1749-6632.2010.05574.x
Sankaran, V. G., Xu, J., Ragoczy, T., Ippolito, G. C., Walkley, C. R., Maika, S. D., et al. (2009). Developmental and species-divergent globin switching are driven by BCL11A. *Nature* 460, 1093–1097. doi: 10.1038/nature08243
Sheridan, B. L., Weatherall, D. J., Clegg, J. B., Pritchard, J., Wood, W. G., Callender, S. T., et al. (1976). The patterns of fetal haemoglobin production in leukaemia. *Br. J. Haematol.* 32, 487–506. doi: 10.1111/j.1365-2141.1976.tb00952.x
Shyh-Chang, N., and Daley, G. Q. (2013). Lin28: primal regulator of growth and metabolism in stem cells. *Cell Stem Cell* 12, 395–406. doi: 10.1016/j.stem.2013.03.005
Shyh-Chang, N., Zhu, H., Yvanka de Soysa, T., Shinoda, G., Seligson, M. T., Tsanov, K. M., et al. (2013). Lin28 enhances tissue repair by reprogramming cellular metabolism. *Cell* 155, 778–792. doi: 10.1016/j.cell.2013.09.059
Stadhouders, R., Aktuna, S., Thongjuea, S., Aghajanirefah, A., Pourfarzad, F., van Ijcken, W., et al. (2014). HBS1L-MYB intergenic variants modulate fetal hemoglobin via long-range MYB enhancers. *J. Clin. Invest.* 124, 1699–1710. doi: 10.1172/jci71520
Stamatoyannopoulos, G. (1991). Human hemoglobin switching. *Science* 252:383. doi: 10.1126/science.2017679
Stamatoyannopoulos, G. (2005). Control of globin gene expression during development and erythroid differentiation. *Exp. Hematol.* 33, 259–271. doi: 10.1016/j.exphem.2004.11.007
Stieglitz, E., Taylor-Weiner, A. N., Chang, T. Y., Gelston, L. C., Wang, Y. D., Mazor, T., et al. (2015). The genomic landscape of juvenile myelomonocytic leukemia. *Nat. Genet.* 47, 1326–1333.
Strouboulis, J., Dillon, N., and Grosveld, F. (1992). Developmental regulation of a complete 70-kb human beta-globin locus in transgenic mice. *Genes Dev.* 6, 1857–1864. doi: 10.1101/gad.6.10.1857
Suzuki, M., Yamazaki, H., Mukai, H. Y., Motohashi, H., Shi, L., Tanabe, O., et al. (2013). Disruption of the Hbs1l-Myb locus causes hereditary persistence of fetal hemoglobin in a mouse model. *Mol. Cell Biol.* 33, 1687–1695. doi: 10.1128/mcb.01617-12
Tanabe, O., Katsuoka, F., Campbell, A. D., Song, W., Yamamoto, M., Tanimoto, K., et al. (2002). An embryonic/fetal beta-type globin gene repressor contains a nuclear receptor TR2/TR4 heterodimer. *EMBO J.* 21, 3434–3442. doi: 10.1093/emboj/cdf340
Tanabe, O., McPhee, D., Kobayashi, S., Shen, Y., Brandt, W., Jiang, X., et al. (2007a). Embryonic and fetal beta-globin gene repression by the orphan nuclear receptors, TR2 and TR4. *EMBO J.* 26, 2295–2306. doi: 10.1038/sj.emboj.7601676
Tanabe, O., Shen, Y., Liu, Q., Campbell, A. D., Kuroha, T., Yamamoto, M., et al. (2007b). The TR2 and TR4 orphan nuclear receptors repress Gata1 transcription. *Genes Dev.* 21, 2832–2844. doi: 10.1101/gad.1593307
Tarnawsky, S. P., Yoshimoto, M., Deng, L., Chan, R. J., and Yoder, M. C. (2017). Yolk sac erythromyeloid progenitors expressing gain of function PTPN11 have functional features of JMML but are not sufficient to cause disease in mice. *Dev. Dyn.* 246, 1001–1014. doi: 10.1002/dvdy.24598
Thein, S. L., Menzel, S., Lathrop, M., and Garner, C. (2009). Control of fetal hemoglobin: new insights emerging from genomics and clinical implications. *Hum. Mol. Genet.* 18, R216–R223.
Thein, S. L., Menzel, S., Peng, X., Best, S., Jiang, J., Close, J., et al. (2007). Intergenic variants of HBS1L-MYB are responsible for a major quantitative trait locus on chromosome 6q23 influencing fetal hemoglobin levels in adults. *Proc. Natl. Acad. Sci. U.S.A.* 104, 11346–11351. doi: 10.1073/pnas.0611393104
Tolhuis, B., Palstra, R. J., Splinter, E., Grosveld, F., and de Laat, W. (2002). Looping and interaction between hypersensitive sites in the active beta-globin locus. *Mol. Cell* 10, 1453–1465. doi: 10.1016/s1097-2765(02)00781-5
Tsang, J. C., Yu, Y., Burke, S., Buettner, F., Wang, C., Kolodziejczyk, A. A., et al. (2015). Single-cell transcriptomic reconstruction reveals cell cycle and multi-lineage differentiation defects in Bcl11a-deficient hematopoietic stem cells. *Genome Biol.* 16:178.
Uda, M., Galanello, R., Sanna, S., Lettre, G., Sankaran, V. G., Chen, W., et al. (2008). Genome-wide association study shows BCL11A associated with persistent fetal hemoglobin and amelioration of the phenotype of beta-thalassemia. *Proc. Natl. Acad. Sci. U.S.A.* 105, 1620–1625. doi: 10.1073/pnas.0711566105
Wang, X., Angelis, N., and Thein, S. L. (2018). MYB – a regulatory factor in hematopoiesis. *Gene* 665, 6–17. doi: 10.1016/j.gene.2018.04.065
Weatherall, D. J. (2001). Phenotype-genotype relationships in monogenic disease: lessons from the thalassaemias. *Nat. Rev. Genet.* 2, 245–255. doi: 10.1038/35066048
Weatherall, D. J., Edwards, J. A., and Donohoe, W. T. (1968). Haemoglobin and red cell enzyme changes in juvenile myeloid leukaemia. *Br. Med. J.* 1, 679–681. doi: 10.1136/bmj.1.5593.679
Wegner, M. (1999). From head to toes: the multiple facets of Sox proteins. *Nucleic Acids Res.* 27, 1409–1420. doi: 10.1093/nar/27.6.1409
Weinberg, R. S., Leibowitz, D., Weinblatt, M. E., Kochen, J., and Alter, B. P. (1990). Juvenile chronic myelogenous leukaemia: the only example of truly fetal (not fetal-like) erythropoiesis. *Br. J. Haematol.* 76, 307–310. doi: 10.1111/j.1365-2141.1990.tb07891.x
Wienert, B., Funnell, A. P., Norton, L. J., Pearson, R. C., Wilkinson-White, L. E., Lester, K., et al. (2015). Editing the genome to introduce a beneficial naturally occurring mutation associated with increased fetal globin. *Nat. Commun.* 6:7085.

Wienert, B., Martyn, G. E., Funnell, A. P. W., Quinlan, K. G. R., and Crossley, M. (2018). Wake-up sleepy gene: reactivating fetal globin for beta-hemoglobinopathies. *Trends Genet.* 34, 927–940. doi: 10.1016/j.tig.2018.09.004

Wienert, B., Martyn, G. E., Kurita, R., Nakamura, Y., Quinlan, K. G. R., and Crossley, M. (2017). KLF1 drives the expression of fetal hemoglobin in British HPFH. *Blood* 130, 803–807. doi: 10.1182/blood-2017-02-767400

Wolff, L. (1996). Myb-induced transformation. *Crit. Rev. Oncog.* 7, 245–260. doi: 10.1615/critrevoncog.v7.i3-4.60

Xu, J., Bauer, D. E., Kerenyi, M. A., Vo, T. D., Hou, S., Hsu, Y. J., et al. (2013). Corepressor-dependent silencing of fetal hemoglobin expression by BCL11A. *Proc. Natl. Acad. Sci. U.S.A.* 110, 6518–6523. doi: 10.1073/pnas.1303976110

Xu, J., Sankaran, V. G., Ni, M., Menne, T. F., Puram, R. V., Kim, W., et al. (2010). Transcriptional silencing of {gamma}-globin by BCL11A involves long-range interactions and cooperation with SOX6. *Genes Dev.* 24, 783–798. doi: 10.1101/gad.1897310

Yi, Z., Cohen-Barak, O., Hagiwara, N., Kingsley, P. D., Fuchs, D. A., Erickson, D. T., et al. (2006). Sox6 directly silences epsilon globin expression in definitive erythropoiesis. *PLoS Genet.* 2:e14. doi: 10.1371/journal.pgen.0020014

Zago, M. A., Wood, W. G., Clegg, J. B., Weatherall, D. J., O'Sullivan, M., and Gunson, H. (1979). Genetic control of F cells in human adults. *Blood* 53, 977–986. doi: 10.1182/blood.v53.5.977.977

Zeng, J., Wu, Y., Ren, C., Bonanno, J., Shen, A. H., Shea, D., et al. (2020). Therapeutic base editing of human hematopoietic stem cells. *Nat. Med.* 26, 535–541. doi: 10.1038/s41591-020-0790-y

Zhou, D., Liu, K., Sun, C. W., Pawlik, K. M., and Townes, T. M. (2010). KLF1 regulates BCL11A expression and gamma- to beta-globin gene switching. *Nat. Genet.* 42, 742–744. doi: 10.1038/ng.637

Zhu, X., Wang, Y., Pi, W., Liu, H., Wickrema, A., and Tuan, D. (2012). NF-Y recruits both transcription activator and repressor to modulate tissue- and developmental stage-specific expression of human gamma-globin gene. *PLoS One* 7:e47175. doi: 10.1371/journal.pone.0047175

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

The handling editor declared a past co-authorship with one of the authors AR.

1.2.3. COUP-TFII, an orphan nuclear receptor controlling γ-globin expression

As described above, numerous γ-globin transcriptional repressors have been found in adult erythroid progenitor cells, with BCL11A serving as the greatest example (Sankaran and Orkin, 2013). On the other hand, knowledge of transcriptional activators of embryonic/fetal genes throughout the early stages of development as well as globins' posttranscriptional regulation are less explored.

A recent study of Elegooz and colleagues (2022) revealed Pumilo 1 (PUM1), an RNA-binding protein transcriptionally regulated by KLF1 (see section 1.2.2), as a direct posttranscriptional regulator of γ-globin. PUM1 reduces HBG messenger RNA (mRNA) stability and translational efficiency that results in reduced γ-globin protein levels.

In addition, HIF1α, a protein involved in adaptation to hypoxia and an ubiquitination target of von Hippel-Lindau (VHL) E3 ubiquitin ligase, has been demonstrated to induce γ-globin genes expression (Feng, 2022). Disruption of the VHL E3 ubiquitin ligase was shown to stabilize HIF1α and promote transcription of HBG genes by binding of HIF1α–HIF1β heterodimers to BGLT3 long non-coding RNA gene located downstream of the HBG1 and HBG2 genes. Further chromatin rearrangement and increased interaction between the γ-globin genes and the β-locus LCR result in an increased HbF production.

COUP-TFII was identified by *in vitro* DNA binding studies as a component of the NF-E3 complex proposed to function as a repressor

of the ε- and γ-globin genes (Ronchi, 1995; Liberati, 1998; Filipe, 1999). Transfection experiments further revealed that the binding of COUP-TFII to the promoters of both ε- and γ-globin genes interferes with NF-Y and perhaps other transcription factors and/or co-factors, leading to either cooperation or competition (Liberati, 2001; Zhu, 2012). Finally, COUP-TFII was identified in our laboratory as a γ-globin activator (Fugazza, 2021). Fugazza and colleagues (2021) demonstrated that the orphan nuclear receptor COUP-TFII, expressed in erythroid progenitor cells of yolk sac origin during early development, together with embryonic/fetal globin genes, is able to induce γ-globin expression in different erythroid cellular models. The overexpression of COUP-TFII in human peripheral blood cells derived from healthy donor and thalassemic patient, overcame the repressive adult erythroid environment and activated the embryonic/fetal globin genes. COUP-TFII alters the γ/β expression ratio and contributes to the interactions within the β-locus, further demonstrating that it favors the expression of embryonic/fetal globins. Interestingly, the COUP-TFII consensus found by MEME (GGTCA) in the ChIP-seq experiment nearly exactly matches the motifs described for BCL11A and TR2/TR4 (Fugazza, 2021), two known repressors of γ-globin in adult cells (described in the section 1.2.2), suggesting a possible competition between these globin regulators and COUP-TFII.

The newly discovered function of COUP-TFII as a fetal globin gene activator brings new insight on the hemoglobin switching mechanism. These findings suggest that COUP-TFII might serve as a target for γ-globin re-activation. However, despite substantial genetic research on

COUP-TFII, little is known about the upstream regulatory signals and potential ligands.

1.3. COUP-TFII

1.3.1. Role of Nuclear Receptors in Controlling Erythropoiesis

Published by **Valentina Pastori**, Serena Pozzi, **Agata Labedz** , Sajeela Ahmed and Antonella Ellena Ronchi. International Journal of Molecular Sciences, March 3, 2022.

This review highlights the involvement of nuclear receptor (NR) family members in different crucial steps of erythropoiesis, including the regulation of the differential expression of globin genes during development.

International Journal of
Molecular Sciences

MDPI

Review

Role of Nuclear Receptors in Controlling Erythropoiesis

Valentina Pastori, Serena Pozzi, Agata Labedz, Sajeela Ahmed and Antonella Ellena Ronchi *

Department of Biotechnology and Biosciences, University of Milano-Bicocca, 20126 Milan, Italy; valentina.pastori@unimib.it (V.P.); s.pozzi49@campus.unimib.it (S.P.); agata.labedz@unimib.it (A.L.); sajeela.ahmed@unimib.it (S.A.)
* Correspondence: antonella.ronchi@unimib.it

Abstract: Nuclear receptors (NRs), are a wide family of ligand-regulated transcription factors sharing a common modular structure composed by an N-terminal domain and a ligand-binding domain connected by a short hinge linker to a DNA-binding domain. NRs are involved in many physiological processes, including metabolism, reproduction and development. Most of them respond to small lipophilic ligands, such as steroids, retinoids, and phospholipids, which act as conformational switches. Some NRs are still "orphan" and the search for their ligands is still ongoing. Upon DNA binding, NRs can act both as transcriptional activators or repressors of their target genes. Theoretically, the possibility to modulate NRs activity with small molecules makes them ideal therapeutic targets, although the complexity of their signaling makes drug design challenging. In this review, we discuss the role of NRs in erythropoiesis, in both homeostatic and stress conditions. This knowledge is important in view of modulating red blood cells production in disease conditions, such as anemias, and for the expansion of erythroid cells in culture for research purposes and for reaching the long-term goal of cultured blood for transfusion.

Keywords: nuclear receptors; erythropoiesis; stress erythropoiesis; globins genes expression during development

Citation: Pastori, V.; Pozzi, S.; Labedz, A.; Ahmed, S.; Ronchi, A.E. Role of Nuclear Receptors in Controlling Erythropoiesis. *Int. J. Mol. Sci.* **2022**, *23*, 2800. https://doi.org/10.3390/ijms23052800

Academic Editor: Pengfei Xu

Received: 4 February 2022
Accepted: 27 February 2022
Published: 3 March 2022

1. Introduction

1.1. Erythropoiesis

Erythropoiesis leads to the formation of the most abundant cells of our body, red blood cells (RBCs), specialized in the oxygen transport to tissues and in the removal of carbon dioxide. In steady state conditions, RBCs are produced at a rate of about 2×10^6 per second. This number can greatly increase upon stress conditions. During development, their production becomes indispensable as soon as the growing size of the embryo makes passive gases exchange insufficient.

Different waves of erythropoiesis take place during development: the first two (hematopoietic stem cell (HSC)-independent) originate from the yolk sac (YS), whereas definitive, HSC-dependent erythropoiesis, is generated from HSCs emerging from the aorta-gonad-mesonephros (AGM) region, the placenta and possibly from other anatomical sites [1]. The first erythroid cells are called primitive erythroid cells (EryP) and are released into the bloodstream while they are still nucleated. The second wave derives from erythroid-myeloid progenitors (EMPs) of YS origin that colonize the fetal liver and generate the first enucleated RBCs. Finally, definitive adult erythrocytes derive from HSCs that first colonize the fetal liver and then the bone marrow, the anatomical site of adult hematopoiesis.

The differentiation of RBCs from HSCs entails several steps leading to the appearance of erythroid-committed progenitors (erythroid burst- and colony-forming units, BFU-E and CFU-E) which give rise to proerythroblasts that progressively differentiate and mature into RBCs, through morphologically well-defined stages. Erythropoiesis is regulated by external signals provided by the microenvironment and by intrinsic factors that coordi-

nate the expression of the genetic program typical of the erythroid cell type, including developmental stage-specific globins.

1.2. Nuclear Receptors

The nuclear receptors, NRs, represent a family of evolutionary conserved proteins acting as ligand-activated transcription factors. NRs are classified according to their evolutionary distance in six subfamilies numbered from one to six and each divided in several groups indicated with letters [2]. They are composed by two main domains, a zinc-finger DNA-binding domain (DBD), necessary for the binding to their target regions on DNA (the response elements, RE), and a ligand-binding domain (LBD), Figure 1. Despite the presence of a structural LBD, for some NRs the ligand is still unknown, and they are then classified as "orphan receptors". NRs present an extremely various pattern of expression and are translated into multiple isoforms with different transcriptional regulatory potential. Moreover, NRs can complex as homodimers or heterodimers and thus recruit different cofactors to promote the assembly of multiprotein complexes with different specificity and transcriptional output. An additional layer of plasticity is represented by their binding sites on DNA, which can be formed by monomeric sites or by two repeats separated by spacers of different length and with different possible orientations (direct repeats, DR, or inverted repeats, IR). Finally, NRs can bind to a wide spectrum of variant DNA consensus sequences. This flexibility generates a very complex combinatorial regulatory code [3–5] (Figure 1).

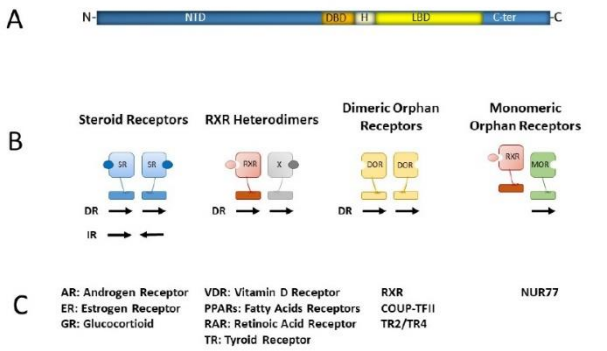


Figure 1. Nuclear receptors: (**A**) schematic structure of nuclear receptors. NTD: N-terminal domain; DBD: DNA-binding domain; H: hinge; LBD: ligand-binding domain. (**B**) The four classes of nuclear receptors. (**C**) The nuclear receptors acting at the different stages of erythropoiesis and discussed in the text are listed under the class they belong to.

The aim of this review is to highlight the involvement of different nuclear receptor family members in different crucial steps for erythropoiesis, ideally following the trajectory leading from the emergency of hematopoietic stem cells to differentiated red blood cells (Figure 2). In particular, we will first focus on NRs having a role in the emergence of erythropoietin-dependent erythropoiesis in early development and on NRs involved in the control of the proliferation-differentiation balance of erythroid progenitors in steady state and stress conditions. Finally, we will introduce NRs important for the regulation of the differential expression of globins genes during development.

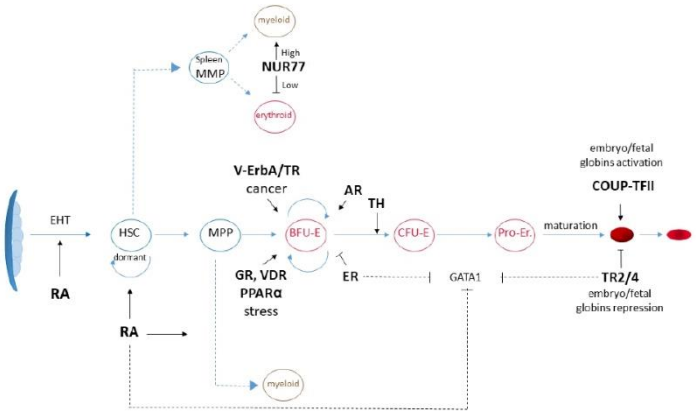


Figure 2. Schematic representation of the major NR signaling pathways affecting erythropoiesis and of their role in erythroid cell lineage specification and maturation. EHT: endothelial hematopoietic transition; HSC: hematopoietic stem cell; MMP: multiple myeloid progenitor; BFU-E: blast-forming unit, erythroid; CFU-E: colony-forming unit, erythroid; Pro Er: proerythroblast; MPPS: splenic multipotent progenitors; RA: retinoic acid; ER: estrogen receptor; AR: androgen receptor; GR: glucocorticoid receptor; GR: granulocytes; VDR: vitamin D receptor; PPARα: peroxisome proliferator-activated receptor; v-ErbA: mutated version of thyroid hormone receptor-α, responsible for avian erythroblastosis; T3: triiodothyronine; TR2/4: testicular receptors 2/4; COUP-TFII: chicken ovalbumin upstream promoter transcription factor II.

2. NRs Controlling Normal Erythropoiesis

2.1. The Developmental Control by Retinoic Acid: From the Emergence of HSCs to the Establishment of Erythropoietin-Dependent Erythropoiesis

Vitamin A acting through its main active metabolite, retinoic acid (RA), is essential for patterning during embryonic development and for adult tissue homeostasis [5]. RA signaling acts through two families of nuclear receptors: retinoic acid receptors (RARs) and retinoic X receptors (RXRs), including three members α, β and γ for each family. RXR and RAR assemble in RAR/RXR heterodimers, although RXR can also form homodimers or heterodimers with other nuclear receptors, such as the vitamin D receptor (VDR) or Peroxisome proliferator-activated receptors (PPARs). RAR/RXR recognize on DNA a specific consensus sequences named "RARE", retinoic acid response element, composed by DRs with variable spacing (DR1, DR2, DR5) [6]. When they are bound to DNA, RARs act as a transcriptional repressor or activators by recruiting cofactors, thus allowing repression or activation of target genes [7,8]. In hematopoiesis, RA signaling is essential for the specification of embryonic hematopoietic stem cells (HSCs) from the haemogenic endothelium within the AGM region [9–13]. RA acquires a transient specific role in erythropoiesis in early fetal liver (E9.5–E10.5 in mouse), where it activates the expression of the Epo gene, coding for erythropoietin, the major hormone stimulating erythroid cells production, thus marking the switch from yolk-sac-derived, Epo-independent, to the HSC-derived, Epo-dependent erythropoiesis, see below [14]. In adults, RA is important for maintaining HSCs in a dormant state (dHSCs), preventing their activation and entry into the cell cycle in response to activating signals [15].

The correlation between Vitamin A deficiency and anemia is well known in humans and in different animal models. However, the molecular mechanisms of RA action is not clearly understood and it likely involves both direct and indirect effects on erythroid progenitors, mainly related to the modulation of iron metabolism and inflammation [16].

Early studies on ex vivo human cultures showed that RA stimulates granulocytic differentiation at the expenses of the erythroid differentiation program by inhibiting the expression of GATA1, the master erythroid transcription factor, in early erythroid progenitors [17]. In line with this, defective RA/RARα axis downstream to the PML-RARα translocation causes acute promyelocytic leukemia due to a failure of promyelocytes differentiation. The pharmacological administration of RA together with arsenic trioxide, restores terminal differentiation of these leukemic progenitors and it is considered the most successful example of "differentiation therapy" [18].

Mouse Models

The role of RAR/RXR has been addressed in detail by several mouse models including loss of function mutants, mutants in specific protein domains and compound double or triple mutants. RA deficiency caused by the knock-out of the RALDH2 gene (retinaldehyde dehydrogenase 2), which catalyzes the synthesis of retinoic acid (RA) from retinaldehyde, leads to death at E10.5 because of multiple defects, including yolk sac malformations. Its conditional deletion in the endothelium (VE-CadherinCre/RALDH2flox model) hampers the formation of HSCs from the hemogenic endothelium [10]. Surprisingly, the RARα knock-out itself leads to a normal development, possibly because of compensatory effects by other RARs [19]. Instead, RARγ knock-out mice are anemic and prone to develop myeloplastic syndromes, although the RARγ conditional erythroid-specific deletion (EpoRCre/RARγ−/−flox model) has no effect on erythropoiesis, suggesting a non-cell-autonomous defect of erythroid cells [20,21]. RXRα deficient mice exhibit a transient anemia (E9.5–E11.5 in mouse) due to the inability to induce erythropoietin (Epo) gene expression, but they look normal at E12.5. The transient nature of anemia likely relies on the differential contribution during development of two regulatory regions controlling the expression of the Epo gene: the enhancer containing an RXR/RAR binding site (the DR2 site) adjacent to a hypoxia response element bound by the hypoxia-inducible factor (HIF1) and the promoter, containing a GATA binding site. In the hypoxic early fetal liver environment, the enhancer plays a dominant role in Epo gene transcription. However, this same element becomes dispensable at later stages, when Epo gene expression mainly depends on the GATA site within the promoter [14,22].

The RXRα/RXRβ double mutant embryos die between E9.5 and E10.5 due to placental abnormalities and multiple malformations [23]. Finally, the double RXRα/RXRγ and RXRβ/RXRγ compound knock-out have a phenotype identical to that of the single RXRα and RXRγ knock-out, respectively. Surprisingly, the triple RXRα+/−/RXRβ−/−/RXRγ−/− mutant is vital, suggesting a large functional redundancy of RXRs [24].

2.2. NRs Affecting Erythroid Differentiation

2.2.1. The Thyroid Hormone Receptor: NR1A1/2

Triiodothyronine (T3), the active form of the thyroid hormone, TH, exerts its action by binding to the nuclear thyroid hormone receptors (TRs), encoded by the TRα and TRβ loci, which generate different isoforms with different pattern of expression and functions [25–27].

The role of TH and TRs in erythropoiesis has been suggested during years by the observation that T3 stimulates erythropoiesis in humans and in animal models [28,29]. Moreover, the avian erythroblastosis v-ErbA oncoprotein, a dominant negative form of c-ErbA/TRα, promotes erythroid progenitors' expansion while blocking terminal differentiation in chicken cells. Under physiological conditions, c-ErbA/TRα acts as a switch: in the absence of a ligand, it supports the proliferation of erythroid progenitors by cooperating with the stem cell factor receptor c-kit, whereas in response to the T3 ligand it unlocks synchronous erythroid terminal differentiation [30–32]. Because of TRs' role in erythro-

poiesis, in patients, hypothyroidism is frequently associated with mild anemia [31–33] and erythroid cultures from hypothyroid patient cells show delayed differentiation of progenitors, high levels of c-kit (expressed on the surface of immature cells) and reduced GpA (glycophorin A, expressed by more mature erythroid cells) [34]. On the contrary, hyperthyroidism is associated with erythrocytosis [35].

Mouse Models

TRα1 knock-out mice show defective fetal and adult erythropoiesis, with a decreased number of early erythroid progenitors in the fetal liver and an impaired stress erythropoiesis response in the adults [36]. This phenotype is similar to that of the mouse model (PV/+ mice [37]), carrying a copy of the dominant negative frame-shift mutation, originally identified in patients, which abolishes the triiodothyronine (T3)-binding domain. However, the underlying mechanism is probably different in the two cases: the lack of TRα1 would lead to the loss of the regulation of erythroid TRα1 target genes, whereas the dominant TRα1PV protein main action consists in the repression of Gata1 expression [37]. Interestingly, the cross of PV/+ mice with mice expressing a mutant version of the corepressor NCOR1, unable to bind to the dominant negative PV protein, partially rescues terminal erythroid differentiation, further indicating that the action of the TRα1PV protein involves transcriptional repression [38].

As for TRβ, GWAS studies identified variants [39] and mutations [40] in the TRβ locus associated to red blood cells parameters, although anemia was not originally observed in patients with TRβ mutations [41]. The role of TRβ and of its coactivator NCOA4 in stimulating terminal differentiation has been shown in primary cells and in anemia mouse models, following the administration of TRβ agonists [42].

Interestingly, the NCOA4 knock-out results in transient anemia during embryonic development but this phenotype must take into account the role of NCOA4 in regulating iron metabolism that could, in turn, contribute to regulate red cells production [42,43]. The discrepancy of the results obtained in constitutive versus tamoxifen-induced NCOA4 knock-out mice add a further layer of complexity in the interpretation of the role of NCOA4 in erythropoiesis [44,45].

The mouse TRβ knock-out model has a normal erythropoiesis [46–49], whereas TRα1 knock-out in mice results in a reduced number of erythroid progenitors in a cell-autonomous way [37]. Interestingly, TH and TRα, but not TRβ, are specifically required for normal spleen erythropoiesis during early postnatal development [50], suggesting a predominant TRα role in early development. Finally, although a role of TH in stimulating hemoglobin production it has long been known [29], a recent observation describes the specific effect of TRIAC, a bioactive thyroid hormone metabolite, in the direct regulation of the embryonic ζ-globin gene expression [51].

2.3. The Complex Role of Sexual Hormones in Erythropoiesis

Several reports have demonstrated over the years that the administration of sex hormones (estrogens, progesterone and testosterone) have an impact on erythropoiesis (see below). The molecular basis of their action, however, is not completely clear and very likely relies, at the physiological level, on both cell-autonomous and non-cell-autonomous mechanisms. The very general message from the available data is that whereas testosterone stimulates RBCs production, estrogen has the opposite effect.

2.3.1. Estrogen Receptors (ERα: NR3A1 and ERβ: NR3A2)

Estradiol promotes human stem/progenitor cells expansion [52] and estrogen receptor alpha (ERα) is clearly required for hematopoietic differentiation of human stem cells [53]. The first experiments based on the treatment of human and animal erythroid cultures with estrogen instead suggested that this female hormone inhibits erythropoiesis [54,55]. These experiments were confirmed in vivo, where high dosage of estrogens causes anemia in humans and animal models [54,56]. Interestingly, it has been observed that in people living

at high altitudes, females are more protected than males against chronic mountain sickness erythrocytosis (CMS or Monge's disease); of note, the incidence of CMS clearly increases after menopause, confirming the link with estrogen levels [57]. The same study reported that ER acts by repressing GATA1, with subsequent erythroid progenitor apoptosis, as originally observed by G. Blobel and colleagues [58,59]. These authors demonstrated that GATA1 repression by ER occurs in a ligand-dependent manner by direct protein–protein interaction.

Mouse Models

Mouse knock-out models carrying loss of function mutations of ERα, ERβ or both show, as expected, severe reproductive phenotypes, with other systems being also affected, such as bone, brain, immune, adipose and cardiovascular, but no overt erythroid defects have been described so far [60–62].

Regarding the hematopoietic system, ERβ knock-out mice develop a chronic myeloproliferative disease during aging, similar to human chronic myeloid leukemia, uncovering the ERβ involvement in the differentiation of pluripotent hematopoietic progenitor cells [63].

2.3.2. Androgen Receptor (AR: NR3C)

The first evidence of a role of AR in erythropoiesis derives from the observation of the sexual difference in hematocrit [64]. The addition of testosterone to ex vivo erythroid progenitors confirmed that this hormone supports their expansion [65] and the in vivo administration of testosterone to different animal models supported this first finding leading to the development of androgen-based therapies to treat anemia [66]. The rationale of these approaches is that AR induces serum Epo [67,68]. Recent bone marrow reconstitution experiments with cells from mice lacking the DNA-binding domain of AR (ARΔZF2) suggested that the DNA-binding domain is indeed required to stimulate kidney erythropoietin [69].

2.3.3. Progesterone Receptor (PR: NR3C3)

Progesterone was originally reported to increase fetal hemoglobin (HbF) in fetal calf liver erythroid cells and adult human erythroid progenitors, when administrated together with Epo [70,71]. In line with these results, uteroferrin (UF), a progesterone-induced acid phosphatase was shown to enhance fetal erythropoiesis [72].

3. NRs Controlling Stress Erythropoiesis Response

In steady state condition, about 10^{11} RBCs per day are produced to maintain homeostasis but this number can increase dramatically in response to hypoxia. This phenomenon is known as "stress erythropoiesis" [73,74].

Two NRs play a fundamental role in stress erythropoiesis: the glucocorticoid receptor (GR, NR3C1) and the Vitamin D receptor (VDR: NR1I1).

3.1. The Glucocorticoid Receptor (GR: NR3C1)

The glucocorticoid receptor (GR) is the constitutively expressed receptor of cortisol and other glucocorticoids (GC). Upon ligand binding, GR homodimerizes and is transferred to the nucleus, where it activates its target genes. In the absence of ligand, GR can bind to other transcription factors, such as AP1 and NFkB, or signal transducers, such as TGFβ and STATs, activating or inhibiting them [75–77]. Because of the combinatorial assembly of different GR complexes, depending on cellular and physiological contexts, GR controls several gene networks [78–81]. Cortisol and synthetic glucocorticoids, such as dexamethasone (DEX), are known regulators of erythroid progenitors, in vitro and in vivo [82–84]: DEX is currently used together with Epo and SCF to sustain the expansion of erythroid progenitors and to delay their terminal differentiation in liquid cultures, a condition mimicking stress erythropoiesis [84–86]. Moreover, pathological conditions altering glucocoticoids metabolism are associated with dysregulated erythropoiesis: insufficient adrenal corticosteroid production is often associated with anemia (Morbus Addison),

whereas patients with elevated GC levels (Cushing's syndrome) have increased red blood cells count, hemoglobin, and hematocrit [87]. In the clinical practice, DEX has been used to treat Diamond–Blackfan anemia (DBA). Recent work based on single cell analysis of bone marrow progenitors from DBA patients suggests that DBA cells fail to activate the endogenous glucocorticoid-pathway of stress erythropoiesis, which sustains the expansion of erythroid progenitors [88]. These authors suggest that the administration of DEX to DBA cells help restoring a proper stress response rather than re-establishing a normal erythropoiesis. This would explain the therapeutic effect of DEX on DBA patients, although with heterogeneous results [89,90]. In this regard, the search for GR single nucleotide polymorphisms (SNPs) associated with DBA response to glucocorticoids failed so far to find significant associations although it is possible that some SNPs could influence the onset of the disease [91].

Mouse Models

GR knock-out mice die perinatally, preventing the analysis of adult postnatal erythropoiesis [92]. However, adult erythroid cells from knock-out E14.5 fetal livers fail to expand in ex vivo cultures and undergo accelerated differentiation [93]. To assess GR-deficient erythropoiesis in vivo, mice carrying a GR mutation in the homodimerization domain, unable to bind DNA, were generated [94]. These animals grow to adulthood and are not anemic but fail to recover from induced anemias, a response sustained by stress erythropoiesis [93]. GR synergizes with another nuclear hormone receptor, PPARα, to enhance BFU-E production: both in vitro and in vivo, PPARα agonists and glucocorticoids enhance RBCs production [95]. Interestingly, PPARα knock-out mice have no defects in steady state or stress erythropoiesis. However, PPARα wild type but not PPARα−/− respond to PPARα agonists under induced acute haemolytic anemia. The observation that PPARα is recruited to GR-adjacent sites on chromatin suggests that PPARα could facilitate GR-dependent BFU-E proliferation [95], thus explaining the above result.

3.2. Vitamin D Receptor (VDR: NR1I1)

The vitamin receptor, VDR, is activated upon binding of Vitamin D3, which induces its translocation into the nucleus and its heterodimerization with RXR. The VDR/RXR complex recruits the transcriptional coregulatory factors that activate the vitamin D hormone-regulated genes [96].

Several observations associate vitamin D deficiency with anemia and vitamin D3 is known to improve anemia condition [97,98]. Vitamin D supports erythropoiesis by increasing burst-forming unit-erythroid (BFU-E) proliferation and has a synergistic effect with Epo and GR to further enhance erythroid progenitor cell proliferation [99–101]. Importantly, VDR is expressed during fetal and adult, but not embryonic erythropoiesis in early $cKit^{+}CD71^{low/neg}$ erythroid definitive progenitors and it is progressively down-regulated during erythroid maturation [100]. According to this expression pattern, the vitamin D3 agonist calcitriol increases the proliferation of these progenitors and delays their terminal maturation. Although the targets of VDR in erythroid cells are still unknown, by analogy with other cell types where Vitamin D controls cell differentiation, it is plausible that genes essential to cell cycle progression are involved in erythroid progenitors' expansion [102,103].

Mouse Models

In mice, the lack of VDR results in a perturbed HSCs homeostasis and in an increased propensity to develop leukemia caused by a block in myeloid differentiation [104], with no overt effects on steady-state erythropoiesis [105–107]. The synergistic effect of GR and VRD in supporting erythroid progenitors [100] raises the hypothesis that VDR could be involved in the stress erythropoiesis response. Indeed, a direct regulation of glucocorticoids on the transcriptional activation of VDR has been described in systems other than erythroid cells [108–110].

4. Orphan Nuclear Receptors

Orphan nuclear receptors owe their name to the absence of a known ligand. The resolution of their crystallographic structure revealed that in the majority of cases they possess an auto-inhibitory structure with side chains occupying the ligand pocket, thus preventing their activation by ligands. In some cases, as discussed below, they can nevertheless be activated by supramolecular doses of retinoic acid or by small molecules.

4.1. The Orphan Nuclear Receptors Controlling the Hemoglobin Switching

During development, the different waves of erythropoiesis generate RBCs characterized by the expression of different globin chains producing different hemoglobin tetramers. This phenomenon, known as hemoglobin switching, has an important clinical relevance since it is known that the reactivation of the fetal globin gene γ in adult cells could substitute for the defective adult globin gene β, whose mutations are the cause of β-hemoglobinopathies, the most common monogenic disease [111].

4.1.1. Testicular Receptors 2 and 4 (TR2: NR2C1; TR4: NR2C2)

TR2 and TR4 are expressed in many tissues, including erythroid cells. They can form homodimers or heterodimers with one another and recognize on DNA directly repeated (DR) AGGTCA sequences, separated by 0–6 nucleotides. The involvement of the testicular receptor 2/4 (TR2/TR4) heterodimer in the regulation of the transition from embryo/fetal to adult globins expression was first described when these two proteins were identified as part of the DRED (direct repeat erythroid-definitive) complex, binding to and repressing the embryonic ε and fetal γ-globin promoters in adult cells [112,113]. Within the large macromolecular repressor complex DRED, TR2 and TR4 act as the core DNA-binding subunits. In addition to binding to the β-locus, TR2 and TR4 also bind to an evolutionally conserved DR element within the GATA1 hematopoietic enhancer (G1HE) and directly repress GATA1 transcription, suggesting that GATA1 may be directly silenced by TR2/TR4 during terminal erythroid maturation [114].

Mouse Models

The direct regulation of GATA1 expression by TR2 and TR4 described above, would explain why transgenic mice overexpressing TR4 alone (TgTR4) or both TR4 and TR2 (TgTR2/TR4) present a transient midembryonic anemia and defects in primitive erythroid precursor formation, which could not be explained simply by the effects of the TR2/TR4 protein complex on globins gene transcription [114].

Mice lacking TR2 are viable and show no evident mutant phenotypes, while TR4 germline mutants display reproductive and neurological deficiencies [115]. However, the careful analysis of erythropoiesis in heterozygous TR4+/− mutant erythroid cells evidenced reduced proliferation and incomplete differentiation [116]. Due to the early embryonic lethality of the double knock-out mice [117], erythropoiesis was studied in a conditional knock-out model in bone marrow cells induced to erythroid differentiation [115]. In these cells, the double TR2/TR4 ablation induced an embryonic globins increase accompanied by an impaired erythroid differentiation, confirming that beside their role in the hemoglobin switching, TR2 and TR4 have a broader role in erythroid differentiation.

4.1.2. Chicken Ovalbumin Upstream Promoter Transcription Factor II (COUP-TFII, NR2F2)

NR2F2 owes its more common alias, COUP-TFII (chicken ovalbumin upstream promoter transcriptional factor II) to the name of the promoter where it was first identified to bind as transcription factor [118]. Structural data suggest that the COUP-TFII protein is capable of an autorepressive conformation, due to side chain amino acids occupying its own ligand pocket [119]. Nevertheless, it can be activated by supraphysiological concentration of retinoids and synthetic agonists and antagonist of COUP-TFII have been recently published [120–122]. COUP-TFII is essential for many biological processes, as

demonstrated by the early embryonic lethality (around E9.5–E10) of COUP-TFII knock-out mice [118]. Conditional knock-out models further demonstrated the requirement of COUP-TFII in different developmental programs, including angiogenesis and cardiac cells specification [123]. Within the erythroid lineage, COUP-TFII is expressed in the early embryo in the yolk sac and in fetal liver and it declines around day E12.5 [124,125]. Similarly to TR2 and TR4, its involvement in the control of the hemoglobin switching, was first discovered because of its ability to bind to the DR1 DNA sequence within the γ-globin promoter [125–127]. What is more, COUP-TFII overexpression in adult cells specifically activates γ-globin, overcoming the repressive adult cellular environment [124]. Chromatin immunoprecipitation experiments confirmed the ability of COUP-TFII to bind to various positions within the locus control region (LCR) of the β-locus, with a pattern overlapping to that of Bcl11a-XL, a known repressor of γ-globin expression [128,129]. This last evidence suggests that the same DNA binding sites could be bound by a repressor or an activator at different times during development.

4.2. Nurr77 (NR4A1) and the Negative Control of the Erythroid Potential of Murine Splenic Progenitors

The NR subfamily 4 (NR4A) group contains three proteins with no known physiologic ligands: NR4A1 (Nur77/Tr3), NR4A2 (Nurr1), and NR4A3 (Nor1) [130–133]. Their crystallographic structure suggests that the ligand binding pocket is occupied by hydrophobic amino acid side chains that preclude the binding with endogenous ligands [134]. NR4s can bind DNA as monomers, homodimers, heterodimers with RXR (NR4A1 and NR4A2) or in complex with non-NR proteins. This flexibility allows NR4A proteins to regulate the expression of diverse targets in different cell types and accounts for their complex role in tumorigenesis, where they can act as tumor suppressors or oncogenes depending on the context [131,132]. In hematopoiesis, NR4As regulate the differentiation and function of various lymphoid [135,136] and myeloid cells [137,138].

Mouse Models

The involvement of NR4As in the repression of mouse spleen erythropoiesis has been highlighted by the observation that the NR4A1 knock-out mice have higher numbers of erythroid progenitors in the spleen. Indeed, NR4A1 regulates the potential of splenic multipotent progenitors (MPPS: $cKit^{+}CD71^{low}CD24^{high}$ cells): NR4A1 expressing cells will differentiate into myeloid lineages, whereas NR4A1 nonexpressing cells will become erythroid cells [139]. Interestingly, in NR4A1 knock-out mice the number of erythroid progenitors in the bone marrow is normal. This "niche-specific" effect likely relies on the different expression level of NR4A1 in the two compartments: the NR4A1 level is higher in spleen than in bone marrow and this different requirement would render these cells more sensitive to its depletion [139]. The abnormal expansion of the myeloid compartment in the double NR4A1/NR4A3 knock-out mice results in the development of acute myeloid leukemia within the first months of life [137].

5. Conclusions

Various members of the NRs family play a role in erythropoiesis during development (as in the case of retinoic acid), in adult steady state (thyroid hormones receptors, sexual hormones) or under stress conditions (glucocorticoid receptors and vitamin D receptor) (Figure 2). In many cases, the treatment of human and animal erythroid cell cultures with NRs agonists (retinoic acid, dexamethasone, thyroid hormone, sexual hormones) has been the first evidence of the role of NRs signaling in different aspects of erythropoiesis. These in vitro data have been further supported by the observation that hormone/vitamin unbalance can cause anemia or erythrocytosis in patients and by the analysis of mutant mouse phenotypes. The overall emerging picture shows an extreme complexity of NRs signaling. This is due to NRs' ability to both activate or repress transcription depending on the presence/absence of their ligand, on the cellular context, on the cellular niche and

on the plethora of interacting proteins/signaling pathways. Despite this complexity, some agonists acting on NRs are already in use: the synthetic glucocorticoid dexamethasone and thyroid hormone are currently used to expand erythroid cultures, a precondition to reach the final goal to obtain large numbers of cultured red blood cells for transfusion [140]. Dexamethasone is also used in clinics to expand erythroid progenitors in DBA patients, although with heterogeneous results.

The advent of biochemical, biophysical, structural, and genomic approaches will help to better understand the role of NRs in physiological erythropoiesis and in disease conditions. This knowledge is essential to design therapeutic strategies based on NRs manipulation in order to improve the expansion/differentiation of erythroid cells.

Funding: This project has received funding from the European Union's Horizon 2020 Research and Innovation Program under the Marie Skłodowska Curie grant agreement No. 813091 (ARCH, Age-related changes in hematopoiesis).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: No new data were created or analyzed in this study. Data sharing is not applicable to this article.

Acknowledgments: We thank Silvia Barabino for discussion and suggestions. We apologize to the many of our colleagues whose work we were not able to cite due to space limitations.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Dzierzak, E.; Philipsen, S. Erythropoiesis: Development and Differentiation. *Cold Spring Harb. Perspect. Med.* **2013**, *3*, a011601. [CrossRef] [PubMed]
2. Mangelsdorf, D.J.; Thummel, C.; Beato, M.; Herrlich, P.; Schütz, G.; Umesono, K.; Blumberg, B.; Kastner, P.; Mark, M.; Chambon, P.; et al. The nuclear receptor superfamily: The second decade. *Cell* **1995**, *83*, 835–839. [CrossRef]
3. Bookout, A.L.; Jeong, Y.; Downes, M.; Yu, R.T.; Evans, R.M.; Mangelsdorf, D.J. Anatomical Profiling of Nuclear Receptor Expression Reveals a Hierarchical Transcriptional Network. *Cell* **2006**, *126*, 789–799. [CrossRef] [PubMed]
4. Chambon, P. A decade of molecular biology of retinoic acid receptors. *FASEB J.* **1996**, *10*, 940–954. [CrossRef] [PubMed]
5. Ghyselinck, N.B.; Duester, G. Retinoic acid signaling pathways. *Development* **2019**, *146*, dev167502. [CrossRef]
6. Al Tanoury, Z.; Piskunov, A.; Rochette-Egly, C. Vitamin A and retinoid signaling: Genomic and nongenomic effects. *J. Lipid Res.* **2013**, *54*, 1761–1775. [CrossRef] [PubMed]
7. McKenna, N.J. EMBO Retinoids 2011: Mechanisms, Biology and Pathology of Signaling by Retinoic Acid and Retinoic Acid Receptors. *Nucl. Recept. Signal.* **2012**, *10*, e003. [CrossRef]
8. Cañete, A.; Cano, E.; Muñoz-Chápuli, R.; Carmona, R. Role of Vitamin A/Retinoic Acid in Regulation of Embryonic and Adult Hematopoiesis. *Nutrients* **2017**, *9*, 159. [CrossRef]
9. Goldie, L.C.; Lucitti, J.L.; Dickinson, M.E.; Hirschi, K.K. Cell signaling directing the formation and function of hemogenic endothelium during murine embryogenesis. *Blood* **2008**, *112*, 3194–3204. [CrossRef]
10. Chanda, B.; Ditadi, A.; Iscove, N.N.; Keller, G. Retinoic Acid Signaling Is Essential for Embryonic Hematopoietic Stem Cell Development. *Cell* **2013**, *155*, 215–227. [CrossRef]
11. Niederreither, K.; Subbarayan, V.; Dollé, P.; Chambon, P. Embryonic retinoic acid synthesis is essential for early mouse post-implantation development. *Nat. Genet.* **1999**, *21*, 444–448. [CrossRef] [PubMed]
12. Lai, L.; Bohnsack, B.; Niederreither, K.; Hirschi, K.K. Retinoic acid regulates endothelial cell proliferation during vasculogenesis. *Development* **2003**, *130*, 6465–6474. [CrossRef] [PubMed]
13. Bohnsack, B.L.; Lai, L.; Dolle, P.; Hirschi, K.K. Signaling hierarchy downstream of retinoic acid that independently regulates vascular remodeling and endothelial cell proliferation. *Genes Dev.* **2004**, *18*, 1345–1358. [CrossRef] [PubMed]
14. Makita, T.; Hernandez-Hoyos, G.; Chen, T.H.-P.; Wu, H.; Rothenberg, E.V.; Sucov, H.M. A developmental transition in definitive erythropoiesis: Erythropoietin expression is sequentially regulated by retinoic acid receptors and HNF4. *Genes Dev.* **2001**, *15*, 889–901. [CrossRef] [PubMed]
15. Cabezas-Wallscheid, N.; Buettner, F.; Sommerkamp, P.; Klimmeck, D.; Ladel, L.; Thalheimer, F.B.; Pastor-Flores, D.; Roma, L.P.; Renders, S.; Zeisberger, P.; et al. Vitamin A-Retinoic Acid Signaling Regulates Hematopoietic Stem Cell Dormancy. *Cell* **2017**, *169*, 807–823.e19. [CrossRef] [PubMed]
16. Semba, R.D.; Bloem, M.W. The anemia of vitamin A deficiency: Epidemiology and pathogenesis. *Eur. J. Clin. Nutr.* **2002**, *56*, 271–281. [CrossRef] [PubMed]

17. Labbaye, C.; Valtieri, M.; Testa, U.; Giampaolo, A.; Meccia, E.; Sterpetti, P.; Parolini, I.; Pelosi, E.; Bulgarini, D.; Cayre, Y.E. Retinoic acid downmodulates erythroid differentiation and GATA1 expression in purified adult-progenitor culture. *Blood* **1994**, *83*, 651–656. [CrossRef]
18. Ablain, J.; de The, H. Retinoic acid signaling in cancer: The parable of acute promyelocytic leukemia. *Int. J. Cancer* **2014**, *135*, 2262–2272. [CrossRef]
19. Li, E.; Sucov, H.; Lee, K.F.; Evans, R.; Jaenisch, R. Normal development and growth of mice carrying a targeted disruption of the alpha 1 retinoic acid receptor gene. *Proc. Natl. Acad. Sci. USA* **1993**, *90*, 1590–1594. [CrossRef]
20. Dewamitta, S.R.; Joseph, C.; Purton, L.E.; Walkley, C.R. Erythroid-extrinsic regulation of normal erythropoiesis by retinoic acid receptors. *Br. J. Haematol.* **2014**, *164*, 280–285. [CrossRef]
21. Walkley, C.R.; Olsen, G.H.; Dworkin, S.; Fabb, S.A.; Swann, J.; McArthur, G.A.; Westmorel, S.V.; Chambon, P.; Scadden, D.T.; Purton, L.E. A microenvironment-induced myeloproliferative syndrome caused by retinoic acid receptor gamma deficiency. *Cell* **2007**, *129*, 1097–1110. [CrossRef] [PubMed]
22. Makita, T.; Duncan, S.A.; Sucov, H.M. Retinoic acid, hypoxia, and GATA factors cooperatively control the onset of fetal liver erythropoietin expression and erythropoietic differentiation. *Dev. Biol.* **2005**, *280*, 59–72. [CrossRef] [PubMed]
23. Wendling, O.; Chambon, P.; Mark, M. Retinoid X receptors are essential for early mouse development and placentogenesis. *Proc. Natl. Acad. Sci. USA* **1999**, *96*, 547–551. [CrossRef] [PubMed]
24. Krezel, W.; Dupe, V.; Mark, M.; Dierich, A.; Kastner, P.; Chambon, P. RXR gamma null mice are apparently normal and compound RXR alpha +/−/RXR beta −/−/RXR gamma −/− mutant mice are viable. *Proc. Natl. Acad. Sci. USA* **1996**, *93*, 9010–9014. [CrossRef] [PubMed]
25. Aranda, A.; Pascual, A. Nuclear hormone receptors and gene expression. *Physiol. Rev.* **2001**, *81*, 1269–1304. [CrossRef]
26. Aranda, A.; Alonso-Merino, E.; Zambrano, A. Receptors of thyroid hormones. *Pediatr. Endocrinol. Rev.* **2013**, *11*, 2–13.
27. Pascual, A.; Aranda, A. Thyroid hormone receptors, cell growth and differentiation. *Biochim. Biophys. Acta* **2013**, *1830*, 3908–3916. [CrossRef]
28. Dainiak, N.; Hoffman, R.; Maffei, L.A.; Forget, B.G. Potentiation of human erythropoiesis in vitro by thyroid hormone. *Nature* **1978**, *272*, 260–262. [CrossRef]
29. Krause, R.L.; Sokoloff, L. Effects of thyroxine on initiation and completion of protein chains of hemoglobin in vitro. *J. Biol. Chem.* **1967**, *242*, 1431–1438. [CrossRef]
30. Gandrillon, O.; Jurdic, P.; Pain, B.; Desbois, C.; Madjar, J.J.; Moscovici, M.G.; Moscovici, C.; Samarut, J. Expression of the v-erbA product, an altered nuclear hormone receptor, is sufficient to transform erythrocytic cells in vitro. *Cell* **1989**, *58*, 115–121. [CrossRef]
31. Bauer, A.; Mikulits, W.; Lagger, G.; Stengl, G.; Brosch, G.; Beug, H. The thyroid hormone receptor functions as a ligand-operated developmental switch between proliferation and differentiation of erythroid progenitors. *EMBO J.* **1998**, *17*, 4291–4303. [CrossRef] [PubMed]
32. Bartunek, P.; Zenke, M. Retinoid X receptor and c-cerbA/thyroid hormone receptor regulate erythroid cell growth and differentiation. *Mol. Endocrinol.* **1998**, *12*, 1269–1279. [CrossRef] [PubMed]
33. Green, S.T.; Ng, J.P. Hypothyroidism and anaemia. *Biomed. Pharmacother.* **1986**, *40*, 326–331. [PubMed]
34. Van Gucht, A.L.M.; Meima, M.E.; Moran, C.; Agostini, M.; Tylki-Szymanska, A.; Krajewska-Walasek, M.; Chrzanowska, K.; Efthymiadou, A.; Chrysis, D.; Demir, K.; et al. Anemia in Patients With Resistance to Thyroid Hormone α: A Role for Thyroid Hormone Receptor α in Human Erythropoiesis. *J. Clin. Endocrinol. Metab.* **2017**, *102*, 3517–3525. [CrossRef] [PubMed]
35. Liu, X.; Liu, J.; Fan, L.; Shi, B. Erythrocytosis associated with hyperthyroidism: A rare case report and clinical study of possible mechanism. *Endocr. Res.* **2015**, *40*, 177–180. [CrossRef]
36. Kendrick, T.S.; Payne, C.J.; Epis, M.R.; Schneider, J.R.; Leedman, P.J.; Klinken, S.P.; Ingley, E. Erythroid defects in TRalpha−/− mice. *Blood* **2008**, *111*, 3245–3248. [CrossRef]
37. Park, S.; Han, C.R.; Park, J.W.; Zhao, L.; Zhu, X.; Willingham, M.; Bodine, D.M.; Cheng, S.-Y. Defective erythropoiesis caused by mutations of the thyroid hormone receptor α gene. *PLoS Genet.* **2017**, *13*, e1006991. [CrossRef]
38. Han, C.R.; Park, S.; Cheng, S.-Y. NCOR1 modulates erythroid disorders caused by mutations of thyroid hormone receptor α1. *Sci. Rep.* **2017**, *7*, 18080. [CrossRef] [PubMed]
39. Van der Harst, P.; Zhang, W.; Leach, I.M.; Rendon, A.; Verweij, N.; Sehmi, J.; Paul, D.S.; Elling, U.; Allayee, H.; Li, X.; et al. Seventy-five genetic loci influencing the human red blood cell. *Nature* **2012**, *492*, 369–375. [CrossRef]
40. Wu, S.Y.; Sadow, P.M.; Refetoff, S.; Weiss, R.E. Tissue responses to thyroid hormone in a kindred with resistance to thyroid hormone harboring a commonly occurring mutation in the thyroid hormone receptor beta gene (P453T). *J. Lab. Clin. Med.* **2005**, *146*, 85–94. [CrossRef]
41. Dumitrescu, A.M.; Refetoff, S. The syndromes of reduced sensitivity to thyroid hormone. *Biochim. Biophys. Acta* **2013**, *1830*, 3987–4003. [CrossRef] [PubMed]
42. Gao, X.; Lee, H.-Y.; Li, W.; Platt, R.; Barrasa, M.I.; Ma, Q.; Elmes, R.R.; Rosenfeld, M.G.; Lodish, H.F. Thyroid hormone receptor beta and NCOA4 regulate terminal erythrocyte differentiation. *Proc. Natl. Acad. Sci. USA* **2017**, *114*, 10107–10112. [CrossRef] [PubMed]

43. Santana-Codina, N.; Gableske, S.; Del Rey, M.Q.; Małachowska, B.; Jedrychowski, M.P.; Biancur, D.E.; Schmidt, P.J.; Fleming, M.D.; Fendler, W.; Harper, J.W.; et al. NCOA4 maintains murine erythropoiesis via cell autonomous and non-autonomous mechanisms. *Haematologica* **2019**, *104*, 1342–1354. [CrossRef]
44. Nai, A.; Pettinato, M.; Federico, G.; Olivari, V.; Carlomagno, F.; Silvestri, L. Tamoxifen erythroid toxicity revealed by studying the role of nuclear receptor co-activator 4 in erythropoiesis. *Haematologica* **2019**, *104*, e383–e384. [CrossRef]
45. Santana-Codina, N.; Gableske, S.; Fleming, M.D.; Harper, J.W.; Kimmelman, A.C.; Mancias, J.D. The role of nuclear receptor co-activator 4 in erythropoiesis (Reply to Nai et al.). *Haematologica* **2019**, *104*, e585–e586. [CrossRef] [PubMed]
46. Forrest, D.; Hanebuth, E.; Smeyne, R.J.; Everds, N.; Stewart, C.L.; Wehner, J.M.; Curran, T. Recessive resistance to thyroid hormone in mice lacking thyroid hormone receptor beta: Evidence for tissue-specific modulation of receptor function. *EMBO J.* **1996**, *15*, 3006–3015. [CrossRef] [PubMed]
47. Wikström, L.; Johansson, C.; Saltó, C.; Barlow, C.; Barros, A.C.; Baas, F.; Forrest, D.; Thorén, P.; Vennström, B. Abnormal heart rate and body temperature in mice lacking thyroid hormone receptor alpha 1. *EMBO J.* **1998**, *17*, 455–461. [CrossRef] [PubMed]
48. Gauthier, K.; Chassande, O.; Plateroti, M.; Roux, J.P.; Legrand, C.; Pain, B.; Rousset, B.; Weiss, R.; Trouillas, J.; Samarut, J. Different functions for the thyroid hormone receptors TRalpha and TRbeta in the control of thyroid hormone production and post-natal development. *EMBO J.* **1999**, *18*, 623–631. [CrossRef]
49. Göthe, S.; Wang, Z.; Ng, L.; Kindblom, J.M.; Barros, A.C.; Ohlsson, C.; Vennström, B.; Forrest, D. Mice devoid of all known thyroid hormone receptors are viable but exhibit disorders of the pituitary-thyroid axis, growth, and bone maturation. *Genes Dev.* **1999**, *13*, 1329–1341. [CrossRef]
50. Angelin-Duclos, C.; Domenget, C.; Kolbus, A.; Beug, H.; Jurdic, P.; Samarut, J. Thyroid hormone T3 acting through the thyroid hormone α receptor is necessary for implementation of erythropoiesis in the neonatal spleen environment in the mouse. *Development* **2005**, *132*, 925–934. [CrossRef]
51. Chen, H.; Wang, Z.; Yu, S.; Han, X.; Deng, Y.; Wang, F.; Chen, Y.; Liu, X.; Zhou, J.; Zhu, J. 3,3′,5-Triiodothyroacetic acid (TRIAC) induces embryonic zeta-globin expression via thyroid hormone receptor alpha. *J. Hematol. Oncol.* **2021**, *14*, 99. [CrossRef] [PubMed]
52. Fañanas-Baquero, S.; Orman, I.; Aparicio, F.B.; de Miguel, S.B.; Merino, J.G.; Yañez, R.; Sainz, Y.F.; Sánchez, R.; Dessy-Rodríguez, M.; Alberquilla, O.; et al. Natural estrogens enhance the engraftment of human hematopoietic stem and progenitor cells in immunodeficient mice. *Haematologica* **2020**, *106*, 1659–1670. [CrossRef] [PubMed]
53. Kim, H.-R.; Lee, J.-H.; Heo, H.-R.; Yang, S.-R.; Haengseok, S.; Park, W.S.; Han, E.-T.; Song, H.; Hong, S.-H. Improved hematopoietic differentiation of human pluripotent stem cells via estrogen receptor signaling pathway. *Cell Biosci.* **2016**, *6*, 50. [CrossRef] [PubMed]
54. Mirand, E.A.; Gordon, A.S. Mechanism of estrogen action in erythropoiesis. *Endocrinology* **1966**, *78*, 325–332. [CrossRef] [PubMed]
55. Adler, S.S.; Trobough, F.E., Jr. Effects of estrogen on erythropoiesis and granuloid progenitor cell (CFU-C) proliferation in mice. *J. Lab. Clin. Med.* **1978**, *91*, 960–968.
56. Dukes, P.P.; Goldwasser, E. Inhibition of erythropoiesis by estrogens. *Endocrinology* **1961**, *69*, 21–29. [CrossRef]
57. Azad, P.; Villafuerte, F.C.; Bermudez, D.; Patel, G.; Haddad, G.G. Protective role of estrogen against excessive erythrocytosis in Monge's disease. *Exp. Mol. Med.* **2021**, *53*, 125–135. [CrossRef]
58. Blobel, G.A.; Sieff, C.A.; Orkin, S.H. Ligand-dependent repression of the erythroid transcription factor GATA-1 by the estrogen receptor. *Mol. Cell. Biol.* **1995**, *15*, 3147–3153. [CrossRef]
59. Blobel, G.A.; Orkin, S.H. Estrogen-induced apoptosis by inhibition of the erythroid transcription factor GATA-1. *Mol. Cell. Biol.* **1996**, *16*, 1687–1694. [CrossRef]
60. Lubahn, D.B.; Moyer, J.S.; Golding, T.S.; Couse, J.F.; Korach, K.S.; Smithies, O. Alteration of reproductive function but not prenatal sexual development after insertional disruption of the mouse estrogen receptor gene. *Proc. Natl. Acad. Sci. USA* **1993**, *90*, 11162–11166. [CrossRef]
61. Dupont, S.; Krust, A.; Gansmuller, A.; Dierich, A.; Chambon, P.; Mark, M. Effect of single and compound knockouts of estrogen receptors alpha (ERalpha) and beta (ERbeta) on mouse reproductive phenotypes. *Development* **2000**, *127*, 4277–4291. [CrossRef] [PubMed]
62. Hewitt, S.C.; Korach, K.S. Estrogen Receptors: New Directions in the New Millennium. *Endocr. Rev.* **2018**, *39*, 664–675. [CrossRef] [PubMed]
63. Shim, G.J.; Wang, L.; Andersson, S.; Nagy, N.; Kis, L.L.; Zhang, Q.; Mäkelä, S.; Warner, M.; Gustafsson, J.-A. Disruption of the estrogen receptor beta gene in mice causes myeloproliferative disease resembling chronic myeloid leukemia with lymphoid blast crisis. *Proc. Natl. Acad. Sci. USA* **2003**, *100*, 6694–6699. [CrossRef] [PubMed]
64. Vahlquist, B. The cause of the sexual differences in erythrocyte, hemoglobin and serum iron levels in human adults. *Blood* **1950**, *5*, 874–875. [CrossRef] [PubMed]
65. Leberbauer, C.; Boulmé, F.; Unfried, G.; Huber, J.; Beug, H.; Müllner, E.W. Different steroids co-regulate long-term expansion versus terminal differentiation in primary human erythroid progenitors. *Blood* **2005**, *105*, 85–94. [CrossRef] [PubMed]
66. Coviello, A.D.; Kaplan, B.; Lakshman, K.M.; Chen, T.; Singh, A.B.; Bhasin, S. Effects of Graded Doses of Testosterone on Erythropoiesis in Healthy Young and Older Men. *J. Clin. Endocrinol. Metab.* **2008**, *93*, 914–919. [CrossRef] [PubMed]
67. Shahani, S.; Braga-Basaria, M.; Maggio, M.; Basaria, S. Androgens and erythropoiesis: Past and present. *J. Endocrinol. Investig.* **2009**, *32*, 704–716. [CrossRef]

68. Al-Sharefi, A.; Mohammed, A.; Abdalaziz, A.; Jayasena, C.N. Androgens and Anemia: Current Trends and Future Prospects. *Front. Endocrinol.* **2019**, *10*, 754. [CrossRef]
69. McManus, J.F.; Nguyen, N.N.; Davey, R.A.; MacLean, H.E.; Pomilio, G.; McCormack, M.P.; Chiu, W.S.; Wei, A.H.; Zajac, J.D.; Curtis, D.J. Androgens stimulate erythropoiesis through the DNA-binding activity of the androgen receptor in non-hematopoietic cells. *Eur. J. Haematol.* **2020**, *105*, 247–254. [CrossRef]
70. Lavrijsen, K.L.; Verwilghen, R.L. The effect of progesterone on hemoglobin synthesis in suspension cultures of fetal erythroid cells from calf liver. *Biochim. Biophys. Acta* **1984**, *803*, 290–301. [CrossRef]
71. Duarte, A.D.S.S.; Sales, T.S.I.; Mengel, J.O.; Costa, F.F.; Saad, S.T.O. Progesterone Upregulates GATA-1 on Erythroid Progenitors Cells in Liquid Culture. *Blood Cells Mol. Dis.* **2002**, *29*, 213–224. [CrossRef] [PubMed]
72. Ying, W.; Wang, H.; Bazer, F.W.; Zhou, B. Pregnancy-Secreted Acid Phosphatase, Uteroferrin, Enhances Fetal Erythropoiesis. *Endocrinology* **2014**, *155*, 4521–4530. [CrossRef] [PubMed]
73. Socolovsky, M. Molecular insights into stress erythropoiesis. *Curr. Opin. Hematol.* **2007**, *14*, 215–224. [CrossRef] [PubMed]
74. Paulson, R.F.; Shi, L.; Wu, D.-C. Stress erythropoiesis: New signals and new stress progenitor cells. *Curr. Opin. Hematol.* **2011**, *18*, 139–145. [CrossRef] [PubMed]
75. Biddie, S.C.; John, S.; Sabo, P.J.; Thurman, R.E.; Johnson, T.A.; Schiltz, R.L.; Miranda, T.B.; Sung, M.-H.; Trump, S.; Lightman, S.; et al. Transcription Factor AP1 Potentiates Chromatin Accessibility and Glucocorticoid Receptor Binding. *Mol. Cell* **2011**, *43*, 145–155. [CrossRef]
76. Kumar, R.; Thompson, E.B. Gene regulation by the glucocorticoid receptor: Structure:function relationship. *J. Steroid Biochem. Mol. Biol.* **2005**, *94*, 383–394. [CrossRef]
77. Wikström, A.-C. Glucocorticoid action and novel mechanisms of steroid resistance: Role of glucocorticoid receptor-interacting proteins for glucocorticoid responsiveness. *J. Endocrinol.* **2003**, *178*, 331–337. [CrossRef]
78. Weikum, E.R.; Knuesel, M.T.; Ortlund, E.A.; Yamamoto, M.T.K.K.R. Glucocorticoid receptor control of transcription: Precision and plasticity via allostery. *Nat. Rev. Mol. Cell Biol.* **2017**, *18*, 159–174. [CrossRef]
79. Li, G.; Wang, S.; Gelehrter, T.D. Identification of glucocorticoid receptor domains involved in transrepression of transforming growth factor-beta action. *J. Biol. Chem.* **2003**, *278*, 41779–41788. [CrossRef]
80. Stöcklin, E.; Wissler, M.; Gouilleux, F.; Groner, B. Functional interactions between Stat5 and the glucocorticoid receptor. *Nature* **1996**, *383*, 726–728. [CrossRef]
81. Cidlowski, J.A.; Malchoff, C.D.; Malchoff, D.M. *Glucocorticoid Receptors, Their Mechanisms of Action, and Glucocorticoid Resistance Physiology*; Jameson, J.L., De Groot, L.J., de Kretser, D.M., Eds.; Saunders: Philadelphia, PA, USA, 2015.
82. Flygare, J.; Estrada, V.R.; Shin, C.; Gupta, S.; Lodish, H.F. HIF1α synergizes with glucocorticoids to promote BFU-E progenitor self-renewal. *Blood* **2011**, *117*, 3435–3444. [CrossRef] [PubMed]
83. Zhang, L.; Prak, L.; Rayon-Estrada, V.; Thiru, P.; Flygare, J.; Lim, B.; Lodish, H.F. ZFP36L2 is required for self-renewal of early burst-forming unit erythroid progenitors. *Nature* **2013**, *499*, 92–96. [CrossRef] [PubMed]
84. Von Lindern, M.; Zauner, W.; Mellitzer, G.; Steinlein, P.; Fritsch, G.; Huber, K.; Löwenberg, B.; Beug, H. The glucocorticoid receptor cooperates with the erythropoietin receptor and c-Kit to enhance and sustain proliferation of erythroid progenitors in vitro. *Blood* **1999**, *94*, 550–559. [CrossRef] [PubMed]
85. Wessely, O.; Deiner, E.M.; Beug, H.; Von Lindern, M. The glucocorticoid receptor is a key regulator of the decision between self-renewal and differentiation in erythroid progenitors. *EMBO J.* **1997**, *16*, 267–280. [CrossRef] [PubMed]
86. Golde, D.W.; Bersch, N.; Cline, M.J. Potentiation of erythropoiesis in vitro by dexamethasone. *J. Clin. Investig.* **1976**, *57*, 57–62. [CrossRef]
87. Miller, W.L.; Tyrell, J.B. *The Adrenal Cortex*; Miller, W.L., Tyrell, J.B., Felig, P., Baxter, J., Frohman, L., Eds.; McGraw-Hill: New York, NY, USA, 1995; pp. 555–711.
88. Iskander, D.; Wang, G.; Heuston, E.F.; Christodoulidou, C.; Psaila, B.; Ponnusamy, K.; Ren, H.; Mokhtari, Z.; Robinson, M.; Chaidos, A.; et al. Single-cell profiling of human bone marrow progenitors reveals mechanisms of failing erythropoiesis in Diamond-Blackfan anemia. *Sci. Transl. Med.* **2021**, *13*, 0113. [CrossRef]
89. Vlachos, A.; Muir, E. How I treat Diamond-Blackfan anemia. *Blood* **2010**, *116*, 3715–3723. [CrossRef]
90. Bartels, M.; Bierings, M. How I manage children with Diamond-Blackfan anaemia. *Br. J. Haematol.* **2019**, *184*, 123–133. [CrossRef]
91. Lonetti, A.; Indio, V.; Dianzani, I.; Ramenghi, U.; Da Costa, L.; Pospíšilová, D.; Migliaccio, A.R. The Glucocorticoid Receptor Polymorphism Landscape in Patients with Diamond Blackfan Anemia Reveals an Association Between Two Clinically Relevant Single Nucleotide Polymorphisms and Time to Diagnosis. *Front. Physiol.* **2021**, *12*, 745032. [CrossRef]
92. Cole, T.J.; Blendy, J.A.; Monaghan, A.P.; Krieglstein, K.; Schmid, W.; Aguzzi, A.; Fantuzzi, G.; Hummler, E.; Unsicker, K.; Schütz, G. Targeted disruption of the glucocorticoid receptor gene blocks adrenergic chromaffin cell development and severely retards lung maturation. *Genes Dev.* **1995**, *9*, 1608–1621. [CrossRef]
93. Bauer, A.; Tronche, F.; Wessely, O.; Kellendonk, C.; Reichardt, H.M.; Steinlein, P.; Schütz, G.; Beug, H. The glucocorticoid receptor is required for stress erythropoiesis. *Genes Dev.* **1999**, *13*, 2996–3002. [CrossRef] [PubMed]
94. Reichardt, H.M.; Kaestner, K.H.; Tuckermann, J.; Kretz, O.; Wessely, O.; Bock, R.; Gass, P.; Schmid, W.; Herrlich, P.; Angel, P.; et al. DNA Binding of the Glucocorticoid Receptor Is Not Essential for Survival. *Cell* **1998**, *93*, 531–541. [CrossRef]
95. Lee, H.-Y.; Gao, X.; Barrasa, M.I.; Li, H.; Elmes, R.R.; Peters, L.L.; Lodish, H.F. PPAR-α and glucocorticoid receptor synergize to promote erythroid progenitor self-renewal. *Nature* **2015**, *522*, 474–477. [CrossRef] [PubMed]

96. Pike, J.W.; Meyer, M.B. Fundamentals of vitamin D hormone-regulated gene expression. *J. Steroid Biochem. Mol. Biol.* **2014**, *144 Pt A*, 5–11. [CrossRef]
97. Smith, E.M.; Tangpricha, V. Vitamin D and anemia: Insights into an emerging association. *Curr. Opin. Endocrinol. Diabetes Obes.* **2015**, *22*, 432–438. [CrossRef] [PubMed]
98. Napolitano, L.M. Vitamin D supplementation and hemoglobin: Dosing matters in prevention/treatment of anemia. *Nutr. J.* **2021**, *20*, 23. [CrossRef] [PubMed]
99. Aucella, F.; Scalzulli, R.P.; Gatta, G.; Vigilante, M.; Carella, A.M.; Stallone, C. Calcitriol increases burst-forming unit-erythroid proliferation in chronic renal failure. A synergistic effect with r-HuEpo. *Nephron Clin. Pract.* **2003**, *95*, c121–c127. [CrossRef]
100. Barminko, J.; Reinholt, B.M.; Emmanuelli, A.; Lejeune, A.N.; Baron, M.H. Activation of the vitamin D receptor transcription factor stimulates the growth of definitive erythroid progenitors. *Blood Adv.* **2018**, *2*, 1207–1219. [CrossRef]
101. Ben Alon, D.B.; Chaimovitz, C.; Dvilansky, A.; Lugassy, G.; Douvdevani, A.; Shany, S.; Nathan, I. Novel role of 1,25(OH)(2)D(3)in induction of erythroid progenitor cell proliferation. *Exp. Hematol.* **2002**, *30*, 403–409. [CrossRef]
102. Campbell, M.J.; Trump, D.L. Vitamin D Receptor Signaling and Cancer. *Endocrinol. Metab. Clin. N. Am.* **2017**, *46*, 1009–1038. [CrossRef]
103. Samuel, S.; Sitrin, M.D. Vitamin D's role in cell proliferation and differentiation. *Nutr. Rev.* **2008**, *66* (Suppl. 2), S116–S124. [CrossRef] [PubMed]
104. Paubelle, E.; Zylbersztejn, F.; Maciel, T.T.; Carvalho, C.; Mupo, A.; Cheok, M.; Lieben, L.; Sujobert, P.; Decroocq, J.; Yokoyama, A.; et al. Vitamin D Receptor Controls Cell Stemness in Acute Myeloid Leukemia and in Normal Bone Marrow. *Cell Rep.* **2020**, *30*, 739–754.e4. [CrossRef] [PubMed]
105. Li, Y.C.; Pirro, A.E.; Amling, M.; Delling, G.; Baron, R.; Bronson, R.; Demay, M.B. Targeted ablation of the vitamin D receptor: An animal model of vitamin D-dependent rickets type II with alopecia. *Proc. Natl. Acad. Sci. USA* **1997**, *94*, 9831–9835. [CrossRef] [PubMed]
106. Sakai, Y.; Kishimoto, J.; Demay, M.B. Metabolic and cellular analysis of alopecia in vitamin D receptor knockout mice. *J. Clin. Investig.* **2001**, *107*, 961–966. [CrossRef] [PubMed]
107. O'Kelly, J.; Hisatake, J.; Hisatake, Y.; Bishop, J.; Norman, A.; Koeffler, H.P. Normal myelopoiesis but abnormal T lymphocyte responses in vitamin D receptor knockout mice. *J. Clin. Investig.* **2002**, *109*, 1091–1099. [CrossRef]
108. Hidalgo, A.A.; Trump, D.L.; Johnson, C.S. Glucocorticoid regulation of the vitamin D receptor. *J. Steroid Biochem. Mol. Biol.* **2010**, *121*, 372–375. [CrossRef]
109. Hidalgo, A.A.; Deeb, K.K.; Pike, J.W.; Johnson, C.S.; Trump, D.L. Dexamethasone Enhances 1α,25-Dihydroxyvitamin D3 Effects by Increasing Vitamin D Receptor Transcription. *J. Biol. Chem.* **2011**, *286*, 36228–36237. [CrossRef]
110. Zella, L.A.; Meyer, M.B.; Nerenz, R.D.; Lee, S.M.; Martowicz, M.L.; Pike, J.W. Multifunctional Enhancers Regulate Mouse and Human Vitamin D Receptor Gene Transcription. *Mol. Endocrinol.* **2010**, *24*, 128–147. [CrossRef]
111. Wang, X.; Thein, S.L. Switching from fetal to adult hemoglobin. *Nat. Genet.* **2018**, *50*, 478–480. [CrossRef]
112. Tanabe, O.; Katsuoka, F.; Campbell, A.D.; Song, W.; Yamamoto, M.; Tanimoto, K.; Engel, J.D. An embryonic/fetal beta-type globin gene repressor contains a nuclear receptor TR2/TR4 heterodimer. *EMBO J.* **2002**, *21*, 3434–3442. [CrossRef]
113. Tanabe, O.; McPhee, D.; Kobayashi, S.; Shen, Y.; Brandt, W.; Jiang, X.; Campbell, A.D.; Chen, Y.-T.; Chang, C.S.; Yamamoto, M.; et al. Embryonic and fetal beta-globin gene repression by the orphan nuclear receptors, TR2 and TR4. *EMBO J.* **2007**, *26*, 2295–2306. [CrossRef] [PubMed]
114. Tanabe, O.; Shen, Y.; Liu, Q.; Campbell, A.D.; Kuroha, T.; Yamamoto, M.; Engel, J.D. The TR2 and TR4 orphan nuclear receptors repress Gata1 transcription. *Genes Dev.* **2007**, *21*, 2832–2844. [CrossRef] [PubMed]
115. Lee, M.P.; Tanabe, O.; Shi, L.; Jearawiriyapaisarn, N.; Lucas, D.; Engel, J.D. The orphan nuclear receptor TR4 regulates erythroid cell proliferation and maturation. *Blood* **2017**, *130*, 2537–2547. [CrossRef] [PubMed]
116. Shyr, C.-R.; Kang, H.-Y.; Tsai, M.-Y.; Liu, N.-C.; Ku, P.-Y.; Huang, K.-E.; Chang, C. Roles of Testicular Orphan Nuclear Receptors 2 and 4 in Early Embryonic Development and Embryonic Stem Cells. *Endocrinology* **2009**, *150*, 2454–2462. [CrossRef] [PubMed]
117. Cui, S.; Tanabe, O.; Sierant, M.; Shi, L.; Campbell, A.; Lim, K.C.; Engel, J.D. Compound loss of function of nuclear receptors Tr2 and Tr4 leads to induction of murine embryonic beta-type globin genes. *Blood* **2015**, *125*, 1477–1487. [CrossRef]
118. Pereira, F.A.; Qiu, Y.; Tsai, M.-J.; Tsai, S.Y. Chicken ovalbumin upstream promoter transcription factor (COUP-TF): Expression during mouse embryogenesis. *J. Steroid Biochem. Mol. Biol.* **1995**, *53*, 503–508. [CrossRef]
119. Kruse, S.W.; Suino-Powell, K.; Zhou, X.E.; Kretschman, J.E.; Reynolds, R.; Vonrhein, C.; Xu, Y.; Wang, L.; Tsai, S.Y.; Tsai, M.-J.; et al. Identification of COUP-TFII Orphan Nuclear Receptor as a Retinoic Acid–Activated Receptor. *PLOS Biol.* **2008**, *6*, e227. [CrossRef]
120. Yoon, K.; Chen, C.-C.; Orr, A.A.; Barreto, P.N.; Tamamis, P.; Safe, S. Activation of COUP-TFI by a Novel Diindolylmethane Derivative. *Cells* **2019**, *8*, 220. [CrossRef]
121. Le Guével, R.; Oger, F.; Martinez-Jimenez, C.P.; Bizot, M.; Gheeraert, C.; Firmin, F.; Ploton, M.; Kretova, M.; Palierne, G.; Staels, B.; et al. Inactivation of the Nuclear Orphan Receptor COUP-TFII by Small Chemicals. *ACS Chem. Biol.* **2017**, *12*, 654–663. [CrossRef]
122. Wang, L.; Cheng, C.-M.; Qin, J.; Xu, M.; Kao, C.-Y.; Shi, J.; You, E.; Gong, W.; Rosa, L.P.; Chase, P.; et al. Small-molecule inhibitor targeting orphan nuclear receptor COUP-TFII for prostate cancer treatment. *Sci. Adv.* **2020**, *6*, eaaz8031. [CrossRef]
123. Lin, F.-J.; Qin, J.; Tang, K.; Tsai, S.Y.; Tsai, M.-J. Coup d'Etat: An Orphan Takes Control. *Endocr. Rev.* **2011**, *32*, 404–421. [CrossRef] [PubMed]

124. Fugazza, C.; Barbarani, G.; Elangovan, S.; Marini, M.G.; Giolitto, S.; Font-Monclus, I.; Marongiu, M.F.; Manunza, L.; Strouboulis, J.; Cantu, C.; et al. The Coup-TFII orphan nuclear receptor is an activator of the gamma-globin gene. *Haematologica* **2021**, *106*, 474–482. [CrossRef] [PubMed]
125. Filipe, A.; Li, Q.; Deveaux, S.; Godin, I.; Romeo, P.-H.; Stamatoyannopoulos, G.; Mignotte, V. Regulation of embryonic/fetal globin genes by nuclear hormone receptors: A novel perspective on hemoglobin switching. *EMBO J.* **1999**, *18*, 687–697. [CrossRef]
126. Liberati, C.; Cera, M.R.; Secco, P.; Santoro, C.; Mantovani, R.; Ottolenghi, S.; Ronchi, A. Cooperation and competition between the binding of COUP-TFII and NF-Y on human epsilon- and gamma-globin gene promoters. *J. Biol. Chem.* **2001**, *276*, 41700–41709. [CrossRef] [PubMed]
127. Ronchi, A.E.; Bottardi, S.; Mazzucchelli, C.; Ottolenghi, S.; Santoro, C. Differential binding of the NFE3 and CP1/NFY transcription factors to the human gamma- and epsilon-globin CCAAT boxes. *J. Biol. Chem.* **1995**, *270*, 21934–21941. [CrossRef]
128. Sankaran, V.G.; Menne, T.F.; Xu, J.; Akie, T.E.; Lettre, G.; Van Handel, B.; Mikkola, H.K.A.; Hirschhorn, J.N.; Cantor, A.B.; Orkin, S.H. Human Fetal Hemoglobin Expression Is Regulated by the Developmental Stage-Specific Repressor BCL11A. *Science* **2008**, *322*, 1839–1842. [CrossRef] [PubMed]
129. Liu, N.; Hargreaves, V.V.; Zhu, Q.; Kurland, J.V.; Hong, J.; Kim, W.; Sher, F.; Macias-Trevino, C.; Rogers, J.M.; Kurita, R.; et al. Direct Promoter Repression by BCL11A Controls the Fetal to Adult Hemoglobin Switch. *Cell* **2018**, *173*, 430–442.e17. [CrossRef]
130. Pearen, M.A.; Muscat, G.E. Minireview: Nuclear hormone receptor 4A signaling: Implications for metabolic disease. *Mol. Endocrinol.* **2010**, *24*, 1891–1903. [CrossRef]
131. Safe, S.; Karki, K. The Paradoxical Roles of Orphan Nuclear Receptor 4A (NR4A) in Cancer. *Mol. Cancer Res.* **2021**, *19*, 180–191. [CrossRef]
132. Lee, S.-O.; Li, X.; Khan, S.; Safe, S. Targeting NR4A1 (TR3) in cancer cells and tumors. *Expert Opin. Ther. Targets* **2011**, *15*, 195–206. [CrossRef]
133. Hsu, H.C.; Zhou, T.; Mountz, J.D. Nur77 family of nuclear hormone receptors. *Curr. Drug Targets Inflamm. Allergy* **2004**, *3*, 413–423. [CrossRef] [PubMed]
134. Wang, Z.; Benoit, G.; Liu, J.; Prasad, S.; Aarnisalo, P.; Liu, X.; Xu, H.; Walker, N.P.C.; Perlmann, T. Structure and function of Nurr1 identifies a class of ligand-independent nuclear receptors. *Nature* **2003**, *423*, 555–560. [CrossRef] [PubMed]
135. Woronicz, J.D.; Calnan, B.; Ngo, V.; Winoto, A. Requirement for the orphan steroid receptor Nur77 in apoptosis of T-cell hybridomas. *Nature* **1994**, *367*, 277–281. [CrossRef] [PubMed]
136. Sekiya, T.; Kashiwagi, I.; Inoue, N.; Morita, R.; Hori, S.; Waldmann, H.; Rudensky, A.Y.; Ichinose, H.; Metzger, D.; Chambon, P.; et al. The nuclear orphan receptor Nr4a2 induces Foxp3 and regulates differentiation of CD4+ T cells. *Nat. Commun.* **2011**, *2*, 269. [CrossRef]
137. Mullican, S.E.; Zhang, S.; Konopleva, M.; Ruvolo, V.; Andreeff, M.; Milbrandt, J.; Conneely, O.M. Abrogation of nuclear receptors Nr4a3 andNr4a1 leads to development of acute myeloid leukemia. *Nat. Med.* **2007**, *13*, 730–735. [CrossRef]
138. Freire, P.R.; Conneely, O.M. NR4A1 and NR4A3 restrict HSC proliferation via reciprocal regulation of C/EBPalpha and inflammatory signaling. *Blood* **2018**, *131*, 1081–1093. [CrossRef]
139. Mumau, M.D.; Vanderbeck, A.N.; Lynch, E.D.; Golec, S.B.; Emerson, S.G.; Punt, J.A. Identification of a Multipotent Progenitor Population in the Spleen That Is Regulated by NR4A1. *J. Immunol.* **2018**, *200*, 1078–1087. [CrossRef]
140. Pellegrin, S.; Severn, C.E.; Toye, A.M. Towards manufactured red blood cells for the treatment of inherited anemia. *Haematologica* **2021**, *106*, 2304–2311. [CrossRef]

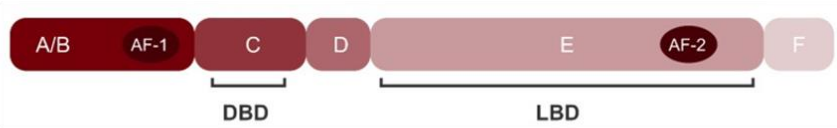
1.3.2. The COUP-TFII gene and protein structure

In humans, the COUP-TFII (Chicken Ovalbumin Upstream Promoter Transcription Factor II) gene is located at chromosome 15 and is composed of three exons and a newly discovered exon at the putative promoter's 5' end (Qiu, 1995).

COUP-TFII gene encodes a transcription factor belonging to Nuclear Receptor subfamily 2 group F (NR2F) (Qiu, 1995). Nuclear receptors (NRs) are a family of evolutionary conserved proteins acting as a ligand-activated transcription factors (described in the review above). In vertebrates, there are three homologous NR2F subtypes identified so far: COUP-TFI (EAR-3, NR2F1), COUP-TFII (NR2F2, ARP1) and COUP-TFIII (EAR-2, NR2F6). Human NR2F family members have highly conserved molecular structures and their main isoforms are similar in length. Two members of the NR2F family with the highest levels of homology are COUP-TFII (NR2F2) and COUP-TFI (NR2F1) (Qiu, 1995; Bertacchi, 2019).

COUP-TFs proteins are composed of six regions from N-terminal to C-terminal: the A/B region contains the activating function 1 (AF1) necessary for the recruitment of co-activators, the DNA-binding domain (DBD, C region), the D (or hinge) region that can work as the nuclear localization sequence (NLS), the E region that contains the activating function 2 (AF2) and the ligand binding domain (LBD), and the C-terminal F region (Fig. 1.3A).

A



B

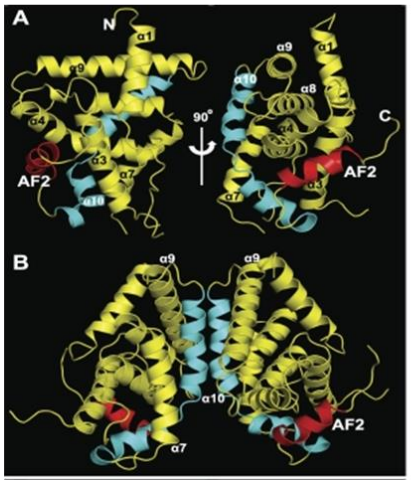


*Figure 1.3**. COUP-TFII protein structure. A** The domain organization of COUP-TF proteins. From N- to C-terminal, A/B: AF-1, activating function 1; C: DBD DNA binding domain; D: hinge region that can act as the nuclear localization sequence (NLS); E: AF-2 activating function 2 and LBD ligand binding domain; F: C-terminal region (Polvani, 2019). **B** Crystal structure of the ligand-free COUP-TFII LBD. Top: front and side views of the COUP-TFII LBD monomer with its AF2 helix colored in red. Bottom: organization of the COUP-TFII LBD dimer, showing that its dimer interface is formed predominantly by helix α10 (blue) (Kruse, 2008).*

The COUP-TFII-DBD is composed of two zinc fingers that resemble the structure of other proteins in the family and recognizes direct repeats and palindromes of the GGTCA motif with different spacings (1-5bp) (Tsai, 1997). Two GGTCA half-sites allow the active dimer form of COUP-TFII to bind to functional DNA while undergoing structural adaptation to identify the interspaced GGTCA motifs (Cooney, 1992). The COUP-TFII-LBD domain contains ten α-helices which folds into a nuclear receptor-specific conformation of a three-layered helical sandwich (Fig. 1.3B). The N-terminal portion of helix α10 is essential for dimer formation by hydrophobic interaction, which is crucial for the whole activity of the protein (Kruse, 2008). Comparing binding domains of COUP-TFII and COUP-TFI, DBD and LBD show 97% and 99% homology, respectively (Qiu, 1995; Alfano, 2014). Although they can attach to homologous DNA regions, their homology in the AF1 part is lower (45%), suggesting that they may have different functions (Pereira, 2000).

Three exons are present in each of the four COUP-TFII transcripts, however three recently discovered isoforms are shorter than the canonical COUP-TFII (COUP-TFII variant 1). The shorter isoforms have different N-terminal region and lack DNA binding domain (DBD), whereas the ligand binding domain (LBD) is identical between all four isoforms. Interestingly, the new isoforms have been demonstrated to either upregulate or downregulate activity of the canonical COUP-TFII in a cell-dependent manner (Rosa, 2011; Yamazaki, 2013).

COUP-TFII is highly conserved across species: its orthologues are found in mice (Qiu, 1995; Jonk, 1994), chicken (Lutz, 1994), Xenopus frog

(Matharu, 1992), and zebrafish (Love, 2012). They constitute a cluster with a high similarity (>95%) in the phylogenic tree of the NR2F-family (Tsai, 1997).

1.3.3. Transcriptional regulation of COUP-TFII

Despite numerous studies of the function of COUP-TFII, the signaling pathways activating the COUP-TFII transcription in various cell types and its regulatory sequences remain largely unknown.

In vitro studies in hepatic HepG2 cells identified a region upstream of the COUP-TFII gene that functions as a specific hepatic COUP-TFII enhancer (Baroukh, 2005). The region located 61 kb upstream to the COUP-TFII transcription start site (TSS) increased COUP-TFII expression *in vivo* and after being injected into the tail vein of an adult mouse (Baroukh, 2005). In P19 mouse tetracarcinoma cells treated with the main active metabolite of vitamin A, retinoic acid (RA), COUP-TFII levels increased, demonstrating that the RA signaling pathway can regulate COUP-TFII gene expression (Qiu, 1996).

In addition, it has been found that COUP-TFII is a target of sonic hedgehog, an essential component of neuronal development, and that Shh response element (ShhRE) is present within the COUP-TFII promoter (Krishnan, 1997).

Finally, E26 transformation specific 1 (ETS-1), a transcription factor belonging to the ETS family of transcription factors and involved in the

regulation of angiogenesis, was demonstrated to directly control COUP-TFII expression (Petit, 2004). ETS-1 binds to the COUP-TFII promoter together with the ETS factor ETV1 and induces COUP-TFII expression, suggesting that COUP-TFII is a potential mediator of ETS-controlled angiogenesis.

At the embryonic stage of vertebrate development, COUP-TFII has a dynamic expression pattern and is expressed in the mesenchymal cells of the developing organs. Mouse early embryonic lethality caused by the targeted disruption of COUP-TFII indicates that COUP-TFII is essential for embryo viability and may play a role in organogenesis (Pereira, 1999; Polvani, 2019). In adult healthy tissues, COUP-TFII expression declines until reaching basal levels. An increase in COUP-TFII expression occurs under some pathological conditions, such as cardiovascular diseases or cancer where it regulates tumor growth and metastasis by modulating tumor angiogenesis (Wu, 2015; Qin, 2010). The expression pattern and profile of COUP-TFII suggests that different regulatory regions are likely involved in its cell-specific and time-specific expression.

1.3.4. Transcriptional activation mediated by the COUP-TFII protein

As a transcription factor, COUP-TFII regulates gene expression. COUP-TFII was originally identified as transcriptional activator of the chicken ovalbumin gene. COUP-TFII can act as a transcriptional activator or

repressor in a cell type-dependent and gene-specific way by directly binding to its AGGTCA DNA motif (Cooney, 1992; Polvani, 2019). The transcriptional activity of COUP-TFII can occur via different mechanisms. COUP-TFII can form a homodimer or heterodimer with other transcription factors to activate or repress its target (Cooney, 1992; Park, 2003; Polvani 2019). COUP-TFII binds directly, indirectly as an auxiliary factor, or through protein-protein interactions to hormone response elements in activated target genes. For example, it physically interacts with HNF4 to activate the HNF1 gene promoter in the liver (Ktistaki and Talianidis, 1997; Ko, 2019). It can act as a repressor in different ways: directly by interacting with the target's promoter region, indirectly by occupying the binding sites of other transcription factors, or transrepressorily by competing with other nuclear receptors to form heterodimers with retinoid X receptors (RXR) (Park, 2003; Polvani 2019). It can also obstruct the transactivation of other nuclear hormone receptors like VDR and RAR by directly occupying their DNA-binding sites or by sequestering RXR, the preferred heterodimeric partner of many NRs. COUP-TFII can act as a transcriptional inhibitor via activating co-repressors or by transrepression the target gene by directly binding to the LBD of nuclear hormone receptors (Leng, 1996; Okamura, 2009). It has been suggested that transrepression can be the predominant mechanism used by COUP-TFII to repress gene expression (Park, 2003; Achatz, 1997).

COUP-TFII displays an auto-repressed conformation in the absence of ligands as a result of the contact between the AF-2 region and the

cofactor binding site, which physically prevents the recruitment of co-activator or co-repressor proteins. The C-terminal part of helix α10 of COUP-TFII protein collapses into the ligand-binding site in the absence of a ligand, revealing its autorepressed conformation (Kruse, 2008). The COUP-TFII activity was significantly reduced when the C-terminal AF2 region was altered, demonstrating that this region's intact structure is likewise necessary for COUP-TFII activation. All COUP-TFII isoforms contain the LBD, which therefore is likely important for the control of COUP-TFII function.

Several pieces of evidence hint to a process in which ligands activate COUP-TFII by releasing the receptor from its auto-inhibited state. Crystallographic analysis revealed that COUP-TFII structure lacks a conventional ligand-binding pocket indicating that it is most likely a true orphan receptor (Kruse, 2008). Even though COUP-TFII was classified as an orphan nuclear receptor that lacks of a known natural ligand, Kruse and colleagues (2008) identified COUP-TFII as a low affinity retinoic acid receptor. These authors demonstrated that the ligand-binding domain (LBD) of COUP-TFII protein has a putative active ligand-binding pocket, and eventually that COUP-TFII can be released from its autorepressed conformation by high concentrations of retinoic acids. COUP-TFII activity can be also regulated by small chemicals. Guevel and colleagues (2017) identified 4-methoxynaphthol (4-MNol), a synthetic inhibitor belonging to naphthol family, as an inhibitor of COUP-TFII. Treatment of hepatic HepG2 and fibroblastic 3T3-L1 cells with 4-MNol, reversed known add details effects of COUP-TFII in these cellular models (Guevel, 2017).

However, 4-MNol needs a high concentration (up to 100µM) to reduce COUP-TFII activity, more recently another inhibitor that functions at the nanomolar concentration and directly binds to the LBD of COUP-TFII protein was identified (Wang, 2020). This small-molecule inhibitor targets specifically COUP-TFII and prevents its interaction with other transcription factors. Finally, Wang and colleagues (2021) performed two sequential screens of metabolic enzymes, and subsequently of substrates and products of their pathways, and identified 1-deoxysphingosines as COUP-TFs regulator. 1-deoxysphingosines, non-canonical alanine-based sphingolipids, bind to COUP-TFII LBD and positively modulate the transcriptional activity of COUP-TFII, when used at physiological concentration (Wang, 2021). Thus, all these findings revealed that COUP-TFII is a ligand-controlled nuclear receptor and could be treated pharmacologically.

1.3.5. Physiological roles of COUP-TFII

In vertebrates, COUP-TFII is expressed mainly at the embryonic stage, in mesenchymal cells of developing organs, where it may be critical for organogenesis as demonstrated by an early embryonic death of COUP-TFII homozygous mouse models (Pereira, 1999; Polvani, 2019). The majority of information about COUP- TFII's physiological roles is derived from conditional knockouts in different tissues.

Cell differentiation and fate decision

Throughout development, cell fate decisions are crucial and are controlled by transcription factor networks. COUP-TFII is widely expressed in differentiating and adult mesenchymal stem cells where it impacts on cell terminal differentiation and fate choice, a role related to the regulation of stemness (Xie, 2011). In response to cytokine stimulation, COUP-TFII regulates the transition from a dormant to an active state of hematopoietic stem cells (HSCs). COUP-TFII plays a role in cell fate decision of bone marrow stromal cells (BMSCs) and in the early phases of adipose differentiation of C3H10T1/2 and BMSC cells. The crucial function of COUP-TFII in differentiation is further highlighted by the relationship and reciprocal control of COUP-TFII with stemness factors, particularly OCT4. OCT4 (octamer-binding transcription factor 4) keeps stem cells pluripotent by actively suppressing COUP-TFII transcription in undifferentiated cells (Rosa, 2011).

Cardiovascular development

A mouse study revealed that the COUP-TFII homozygous knockout (COUP-TFII$^{-/-}$) causes mortality around embryonic day 10 (E10), whereas heterozygous knockout (COUP-TFII$^{+/-}$) mice are smaller than wild-type littermates (Pereira, 1999). Immunostaining for platelet endothelial cell adhesion molecules 1 (PECAM-1), an endothelial cell marker revealed defects in vascular and atrial growth: the primitive capillary plexus develops but does not undergo remodeling that would results in defective angiogenesis (Pereira, 1999). Another study found

that COUP-TFII mutants had thin-walled myocardium, atrioventricular septal defect, and poor angiogenesis as cardiovascular abnormalities (Lin, 2012). Endothelial-specific COUP-TFII deficient embryos presented endocardial cushion hypoplasia due to a failure in epithelial-mesenchymal transition (EMT) (Lin, 2012). During typical development, COUP-TFII inhibits the expression of NP-1 and other Notch signaling molecules to confer vein identity. In contrast, the arterial endothelium lacks COUP-TFII, allowing the expression of Notch signaling molecules to activate the expression of artery-specific genes to confer artery identity.

COUP-TFII also regulates lymphangiogenesis during embryonic development and neolymphangiogenesis in adults (Lin, 2010; Ceni, 2017). COUP-TFII is crucial for the development of lymphatic vessels, which in mammals have a venous origin, and it is expressed in both lymphatic endothelial cells and venous endothelial cells (LECs). Conditional knockout of COUP-TFII in mice results in malformed primitive lymphatic arteries and a lack of LEC progenitor cells. In the cardiovascular system, COUP-TFII is known to participate in angiogenesis (Pereira, 1999) and it is predicted to interact with PROX1 that is required to induce the endothelium to adopt a lymphatic vascular phenotype (Srinivasan, 2010; Risebro, 2009). Additionally, it was demonstrated that by binding to the HEY1/HEY2 promoter areas, COUP-TFII/PROX1 heterodimers induce lymphatic endothelial cell lineage while COUP-TFII homodimers induce venous endothelial cell lineage and block an arterial phenotype (Aranguren, 2013). COUP-TFII

is also involved in the heart's atrial/ventricular cardiomyocyte identification determinations (Wu, 2013).

Reproduction

COUP-TFII is necessary for typical female reproductive processes: in response to exogenous gonadotropins, female mice carrying the COUP-TFII heterozygote exhibit a reduced ability to produce progesterone. The absence of uterine decidualization and unsuccessful embryo attachment in KO females suggest that the endometrium is not functioning properly before implantation (Takamoto, 2005). COUP-TFII is essential for normal female reproductive processes: it is one of the key regulators of the uterine epithelial-stromal cross talk and of the placentation. It also plays a role in male reproductive system: the differentiation of Leydig cells, which produce the testosterone required for normal gonad development and function, is regulated by COUP-TFII. The ablation of COUP-TFII during pre-pubertal phases of male development results in infertility, hypogonadism, and spermatogenetic arrest. Contrary to adult male homozygous mutants, which have difficulties in testosterone synthesis, when COUP-TFII is deleted in the adult stage, there are no abnormalities in differentiation and function of Leydig cells. Hence, COUP-TFII is essential for differentiation but not for the survival of Leydig cells (Qin, 2008). Humans with COUP-TFII mutation also have other characteristics such as ovotesticular or testicular variations in sex development, virilization of the external genitalia, dysmorphic

features, developmental delay, and undeveloped limbs (Bashamboo, 2018; Arsov, 2017; Upadia, 2018; Carvalheira, 2019).

Metabolism and energy homeostasis

Increased metabolism in heterozygous mice suggests that COUP-TFII contributes to the energy homeostasis. The importance of COUP-TFII in glucose homeostasis is demonstrated by the presence of COUP-TFII responsive elements in the promoters of numerous genes producing metabolic enzymes. Moreover, upregulation of COUP-TFII expression by Wnt/β-catenin signaling in response to the administration of Wnt3a in 3T3-L1 preadipocytes, resulting in an inhibition of adipogenesis, demonstrated a role of COUP-TFII as a regulator of adipogenesis (Nostro, 2008; Okamura, 2009).

Other functions

Furthermore, COUP-TFII plays a role during eye development where it regulates the specification and differentiation of mouse neural progenitor cells (Lin, 2011). Moreover, it is required for the formation of GABAergic interneurons in the mouse brain, for the radial and anteroposterior patterning of the stomach during organogenesis, the secretion of insulin and pancreatic function (Lin, 2011; Kanatani, 2015).

The maturation of renal cells also requires COUP-TFII expression. In mice, COUP-TFII deficiency hampered the development of the metanephric mesenchyme, and deletion at later stages decreased the

expression of numerous metanephric mesenchymal genes, including Eya1, Pax2, Six2, and Wt1 (Ishii, 2022).

A dose-balanced expression of COUP-TFII is also necessary for the development of skeletal muscles (Lee, 2017). COUP-TFII is required for myoblast proliferation in early mouse development and its expression decreases as the myoblasts differentiate, it is less important since COUP-TFII can prevent the fusion of myoblasts into myotubes, and therefore the formation of skeletal muscle (Lee, 2017).

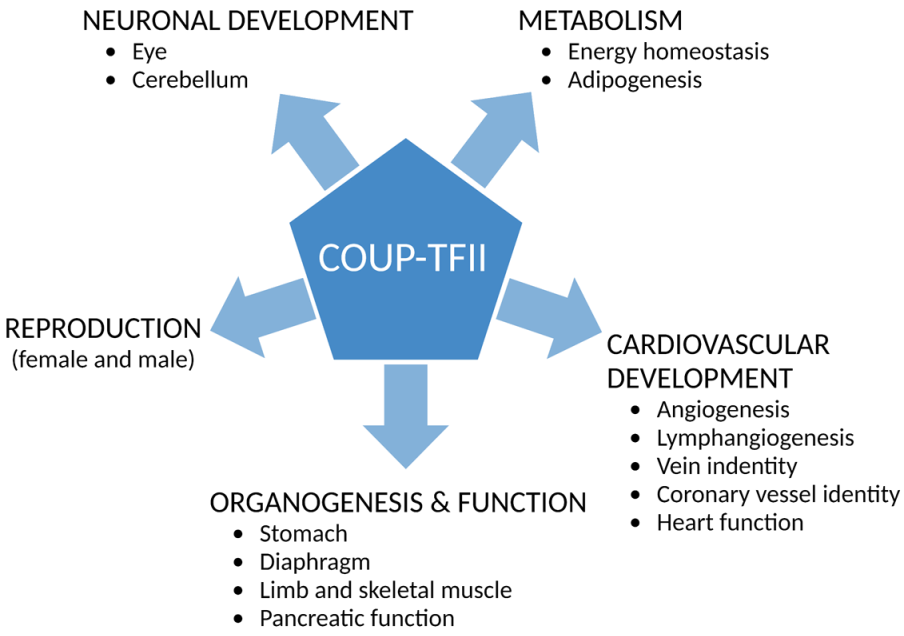


Figure 1.4. ***Physiological roles of COUP-TFII*** *(adapted from Lin, 2011).*

1.4. β-hemoglobinopathies and therapeutical approaches

1.4.1. Sickle Cell Disease (SCD) and β-thalassemia

Hemoglobinopathies are diseases caused by quantitative or qualitative defects in the globin genes expression. The most common β-hemoglobinopathies worldwide are sickle cell disease (SCD) and β-thalassemia (Modell and Darlison, 2008; Piel, 2013; Origa, 2017). SCD, a monogenic disorder, is characterized by a homozygous single base change (A->T) in the seventh codon of the β-globin chain that results in the presence of a valine residue at this location in place of glutamic acid (β6Glu->Val). Characteristic sickle shape of red blood cells (RBCs) is caused by sickle hemoglobin (HbS, $\alpha_2\beta S_2$) that polymerizes in the deoxy-hemoglobin state (Eaton and Bunn, 2017). Cellular dehydration, membrane damage, and finally premature erythrocyte clearance are all effects of RBC sickling. In the end, this causes an aberrant leucocyte, platelet, and endothelial cell interaction, inflammation, susceptibility to infection, vaso-occlusion, recurring pain crises, organ damage, and early mortality. Homozygous βS and a number of compound heterozygous βS diseases with additional β-globin mutations, such as βC (β6Glu->Lys) or β-thalassaemia alleles, that produce comparable clinical symptoms are included in SCD.

β-thalassemia results in significant anemia and inefficient erythropoiesis because there is insufficient or drastically reduced synthesis of the β-globin chain. Despite the fact that these disorders have been the subject of research for more than a century and have well-established molecular genetic underpinnings, the majority of treatment approaches still focus on symptoms rather than the underlying pathophysiology. The main treatments are regular blood transfusions and iron chelation that are essential for survival of homozygous β-thalassemia patients. The best medication currently available, hydroxycarbamide, only partially improves the hematological and clinical symptoms. Allogeneic HSC transplantation remains the only curative option. Nevertheless, the majority of patients do not have a sibling donor who matches their human leucocyte antigen (HLA) (Angelucci, 2014).

1.4.2. Hereditary Persistence of Fetal Hemoglobin (HPFH)

Hereditary Persistence of Fetal Hemoglobin (HPFH) is a benign condition characterized by continuous expression of fetal γ-globin throughout adulthood (Forget, 1998; Steinberg, 2020). HbF can be elevated in all erythroid cells (pancellular HPFH) with HbF levels ranging from 10% to 30%, or in a limited subset of red blood cells (heterocellular HPFH) with HbF levels between 1% and 10% (Thein,

2009). HPFHs can be divided into three categories: deletional HPFH, non-deletional HPFH, and HPFH not linked to the β-locus. In deletional HPFHs a large DNA region between γ and β globin genes is deleted, whereas point mutations described in both γ-globin promoters result in non-deletional HPFHs (described in section 1.2.2). Patients with sickle cell disease or β-thalassemia who also have pancellular HPFH have benign clinical symptoms (Serjeant, 1977).

1.4.3. β-Hemoglobinopathies: The Test Bench for Genome Editing-Based Therapeutic Strategies

Published by **Gloria Barbarani**, Agata Labedz and Antonella Ellena Ronchi. Frontiers in Genome Editing, December 3, 2020.

Here, we describe the different genome editing approaches for treatment of β-hemoglobinopathies.

frontiers in Genome Editing

MINI REVIEW
published: 03 December 2020
doi: 10.3389/fgeed.2020.571239

β-Hemoglobinopathies: The Test Bench for Genome Editing-Based Therapeutic Strategies

*Gloria Barbarani, Agata Łabedz and Antonella Ellena Ronchi**

Dipartimento di Biotecnologie e Bioscienze, Università di Milano-Bicocca, Milan, Italy

Hemoglobin is a tetrameric protein composed of two α and two β chains, each containing a heme group that reversibly binds oxygen. The composition of hemoglobin changes during development in order to fulfill the need of the growing organism, stably maintaining a balanced production of α-like and β-like chains in a 1:1 ratio. Adult hemoglobin (HbA) is composed of two α and two β subunits (α2β2 tetramer), whereas fetal hemoglobin (HbF) is composed of two γ and two α subunits (α2γ2 tetramer). Qualitative or quantitative defects in β-globin production cause two of the most common monogenic-inherited disorders: β-thalassemia and sickle cell disease. The high frequency of these diseases and the relative accessibility of hematopoietic stem cells make them an ideal candidate for therapeutic interventions based on genome editing. These strategies move in two directions: the correction of the disease-causing mutation and the reactivation of the expression of HbF in adult cells, in the attempt to recreate the effect of hereditary persistence of fetal hemoglobin (HPFH) natural mutations, which mitigate the severity of β-hemoglobinopathies. Both lines of research rely on the knowledge gained so far on the regulatory mechanisms controlling the differential expression of globin genes during development.

Keywords: β-hemoglobinopathies, genome editing, globin genes, hereditary persistence of fetal hemoglobin, programmable endonucleases

OPEN ACCESS

Edited by:
Carsten Werner Lederer,
The Cyprus Institute of Neurology and Genetics, Cyprus

Reviewed by:
Joseph Borg,
University of Malta, Malta
Anne Galy,
Institut National de la Santé et de la Recherche Médicale (INSERM), France

****Correspondence:***
Antonella Ellena Ronchi
antonella.ronchi@unimib.it

Specialty section:
This article was submitted to Genome Editing in Blood Disorders, a section of the journal Frontiers in Genome Editing

Received: *10 June 2020*
Accepted: *29 October 2020*
Published: *03 December 2020*

Citation:
Barbarani G, Łabedz A and Ronchi AE (2020) β-Hemoglobinopathies: The Test Bench for Genome Editing-Based Therapeutic Strategies. Front. Genome Ed. 2:571239. doi: 10.3389/fgeed.2020.571239

INTRODUCTION

Historically, because of the abundance and accessibility of red blood cells, globins served as a model for major discoveries later extended to other genes. In 1967, hemoglobin was the first human complex protein crystallized (Muirhead et al., 1967); in 1980, the β-locus was the first cloned gene cluster (Fritsch et al., 1980) and soon became the prototypical model of tissue-specific and developmentally regulated genes. In 1987, the β-locus control region (LCR) was the first long-distance position-independent enhancer characterized (Grosveld et al., 1987), and the current looping model for the interaction of far apart regulatory regions owes much to the study of globin gene sequential activation during development (Stamatoyannopoulos, 1991; Fraser and Grosveld, 1998). The wealth of data accumulated on globin genes put them now at the frontline of development of genome-editing approaches with therapeutic purposes.

THE GLOBIN GENES

In man, globin genes are organized in two clusters lying on chromosomes 16 (α cluster) and 11 (β cluster). A fine-tuned regulation maintains a 1:1 ratio of α-like and β-like chains during

development to produce first HBZ (ζ2ε2, ζ2γ2), then HbE (α2ε2), HbF (α2γ2), and finally HbA (α2β2) together with a small amount of HbA2 (α2δ2), in a process called hemoglobin switching.

At the molecular level, the hemoglobin switching involves the establishment of sequential long-range chromatin physical interactions between a common LCR and the different globin promoters active at a given developmental time (with inactive genes being looped out) in a structure called active chromatin hub (ACH) (Carter et al., 2002; Tolhuis et al., 2002; Palstra et al., 2003). The formation of ACH requires the presence of transcription factors/cofactors that, by binding with the correct affinity to their consensus on DNA, creates the favorable condition for the expression of the gene of interest (Wilber et al., 2011).

β-HEMOGLOBINOPATHIES

Qualitative or quantitative defects in the β-globin production cause the most common monogenic diseases: sickle cell disease (SCD) and β-thalassemia (Weatherall, 2008; Thein, 2013); both diseases, in particular β-thalassemias, are very severe in homozygous subjects, whereas symptoms are mild in carriers. In SCD, the amino acid β6Glu>Val substitution leads to the formation of long hydrophobic polymers of HbS that precipitate within the cell under hypoxic conditions, conferring the typical sickle shape. Sickle cells tend to stick, causing vessel obstruction and, because of their fragility, they frequently undergo hemolysis, finally leading to anemia.

In β-thalassemias, a wide spectrum of mutations causes the reduction of β-globin, which can range in severity from total absence (β^0) to partial reduction (β^+). Causative mutations vary from large deletions to small insertions or deletions (indels) and point mutations within the β gene. β-thalassemia mutations impact on all the different steps of the β gene expression regulation (Thein, 2013): transcription (mutations within regulatory regions), RNA processing (splicing mutations), and translation (ATG mutations, non-sense and missense mutations). In rare cases, β-thalassemia is caused by mutations outside the β-locus, in genes involved in the basal transcription machinery XPD (Viprakasit et al., 2001) or in the erythroid-specific transcription factor GATA1 (Yu et al., 2002). The common output of β-thalassemia mutations is a reduced production of functional β chains with the consequent precipitation of the excess α chains causing hemolysis and anemia. The presence of dysfunctional erythroid progenitors causes ineffective erythropoiesis (Rivella, 2012) and impacts on hematopoietic stem cells (HSCs) self-renewal (Aprile et al., 2020).

The definitive cure for β-hemoglobinopathies is HSC transplantation, a treatment available only for the few patients who have an HLA-matched donor. Despite intense efforts, the only drug of some efficacy remains hydroxyurea (Yu et al., 2020), used to treat SCD (Platt, 2008) and, less successfully, β-thalassemia (Koren et al., 2008; Pourfarzad et al., 2013). Although the condition of β-diseased patients have greatly improved in the last years (Taher et al., 2009), there is a clear need of new approaches, the most innovative of them being based on genome modifications.

THE LESSON FROM NATURE: HEREDITARY PERSISTENCE OF FETAL HEMOGLOBIN

The term hereditary persistence of fetal hemoglobin (HPFH) indicates a heterogeneous spectrum of spontaneous mutations, collectively named by their effect, i.e., the maintenance of the expression of fetal γ-globin in adult stages (Forget, 1998). HPFH alleles, when coinherited with β-hemoglobinopathies, greatly improve the condition of patients, 30% of HbF expression being considered a significant curative threshold able to prevent α free chains polymerization in β-thalassemias and HbS precipitation in SCD (Steinberg et al., 2014). Moreover, δβ-thalassemias, in which δ and β genes are deleted and γ-globin is reactivated, in general to a lesser extent than in HPFH, show that even a relatively low level of γ-globin has beneficial effects on β-thalassemias (Ottolenghi et al., 1982).

HPFH mutations can be broadly divided in three categories: large deletions affecting the structure of the β-locus; point mutations within the γ promoter that identify "hot-spot" HPFH sequences (−200, −175, −158, -distal CCAAT box); and mutations non-linked with the β-locus (Forget, 1998). Two of these non-linked loci, identified by genome-wide association analysis (GWAS), correspond to BCL11A gene, the most important repressor of γ-globin (Menzel et al., 2007; Sankaran et al., 2008; Uda et al., 2008) and to its key activator KLF1 (Borg et al., 2010; Zhou et al., 2010). More recently, knock-out studies in HUDEP cells led to the identification of LRF gene (also known as ZBTB7A), which represses γ-globin independently from BCL11A (Masuda et al., 2016). HPFH mutations within the β-locus greatly increase the expression of one or both γ genes in cis, whereas non-linked HPFH are associated with lower γ-globin levels (Forget, 1998).

The integration of the genetic data on HPFH with molecular studies led to the identification of the target sequences amenable for therapeutic genome editing (see below). In 1992, a pioneer study in mice transgenic for the human β-locus first demonstrated that it is indeed possible to reproduce HPFH (Berry et al., 1992).

THERAPEUTIC GENOME MODIFICATIONS: THE CHOICE OF THE MODIFICATION

In principle, different therapeutic genomic modifications can be envisaged to cure β-hemoglobinopathies (**Figure 1A**):

(1) The addition to the defective cell of an intact β gene, that, once delivered and integrated in the genome of the target hematopoietic stem cell, will produce the missing β chain under the control of an exogenous regulatory cassette, designed to ensure stable, erythroid-specific, and high-level expression. This approach (not discussed in this review), thanks to intensive efforts in developing safe and efficient vectors and in the improvement of their delivery, reached very significant results

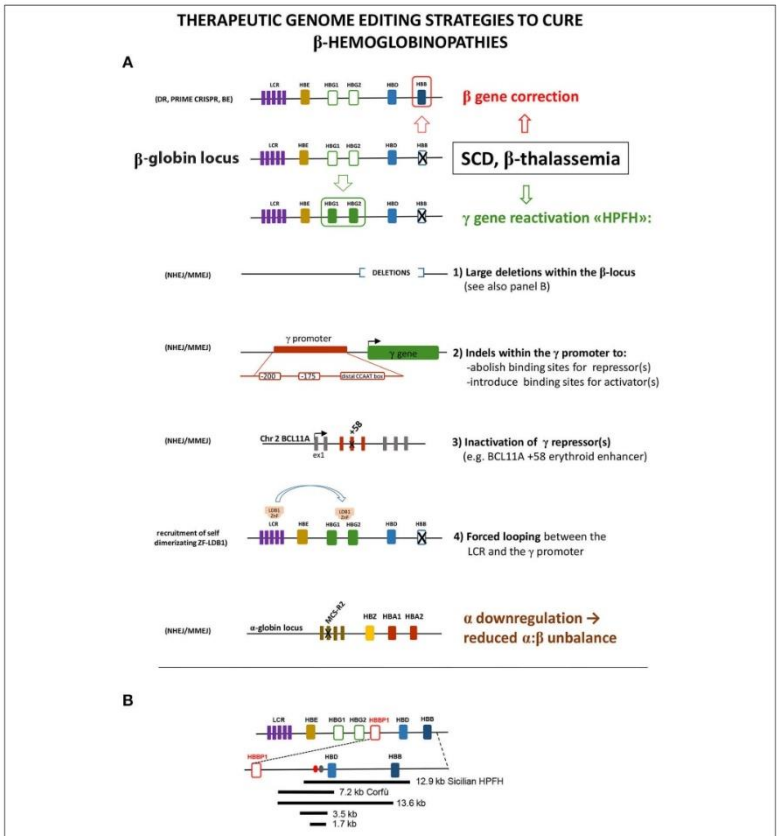


FIGURE 1 | (**A**) Schematic representation of the different editing strategies developed to cure β-hemoglobinopathies. These approaches are oriented to correct the mutated β-globin gene, to reactivate the expression of the fetal γ gene in adult cells or to lower α-globin expression. The reactivation of γ can be obtained via the introduction of different modifications mimicking HPFH caused by large deletions within the β locus or by mutations within the γ promoter that disrupt the binding site of a repressor or create the binding site of an activator. As an alternative, the disruption of the erythroid-specific enhancer of the γ repressor BCL11A abolishes its expression in erythroid cells, thus resulting in γ overexpression. A different approach relies on the forced interaction of the LCR with the γ promoter obtained by exploiting the self-dimerization property of a fusion ZnF-LDB1 protein recruited on these regions. The detrimental effect of the α:β chain imbalance can be mitigated by mutating the MCS-R2 α-globin enhancer, in order to reduce α expression. These different strategies exploit different enzymes/cellular pathway described in the text: HDR, homology-directed repair; BE, base editing; NHEJ, non-homologous end joining; MMEJ, microhomology end joining. **(B)** Schematic representation of the deletions within the β-locus discussed in the text. Ovals represent the BCL11A (red) and the polypyrimidine (gray) sites thought to be responsible for γ repression.

(Ikawa et al., 2019; Magrin et al., 2019). Importantly, the optimization of protocols developed by gene addition approaches represents a knowledge asset fundamental for bringing genome editing approaches to clinical application.

(2) The exact correction of the mutation causing the disease (gene editing), with the advantage of having the β gene expressed under the endogenous regulatory sequences, thus ensuring a perfectly regulated and stable expression with no risk of insertional mutagenesis. This procedure is particularly attractive for point mutations such as the β6Glu>Val SCD-causing mutation or for some β-thalassemia mutations, with the caveat that their extreme heterogeneity would require the design of patient-specific editing strategies. Under this aspect, this approach is worth only for mutations with high frequencies in given populations, as for the *HBB* −28A>G β-thal mutation in Southeast Asia or for the $\beta^0$39C>T and β^+thal IVS-I-110G>A mutations in the Mediterranean area (https://www.ithanet.eu/) (Kountouris et al., 2014).

In an extended perspective, the substitution of the mutated β gene with a wild-type β gene by homologous recombination could combine gene addition and gene editing to create a "universal" substitution cassette, which would minimize the risk of insertional mutagenesis (Cai et al., 2018).

(3) The introduction within the genome of modifications mimicking HPFH, in order to reactivate the expression of the fetal γ-globin gene and to compensate for the missing/defective β-globin expression.

(4) The reduction of the expression of α-globin, an important β-thalassemia modifier, as demonstrated by the milder clinical outcome of patients coinheriting α- and β-thalassemia (Thein, 2008; Mettananda et al., 2015). Reduced α levels indeed reduce the α:β chain imbalance, which represent a major problem in β-hemoglobinopathies. This effect has been achieved experimentally by deleting the MCS-R2 α enhancer (Mettananda et al., 2017).

THE ADVENT OF PROGRAMMABLE ENDONUCLEASES IN THE EDITING OF GLOBIN GENES: THE SEARCH FOR THE BEST COMPROMISE BETWEEN PRECISION AND EFFICIENCY

In 1985, Oliver Smithies first exploited homologous recombination (HR) to introduce an exogenous DNA sequence within the β-locus (Smithies et al., 1985), demonstrating the feasibility of this approach. Since then, HR was used to generate gene knock-out models (including the KO of GATA1 (Pevny et al., 1991) and KLF1 (Nuez et al., 1995; Perkins et al., 1995)), by inserting exogenous DNA in the desired target. However, the very low efficiency of gene targeting, the consequent need of selecting the modified cells, and the technical difficulties of the method discouraged clinical applications (Vega, 1991).

The scenario radically changed with the advent of programmable endonucleases: zinc finger (ZnF) and TALENs first and now, CRISPR/Cas9 and its derivatives (Cornu et al., 2017; Komor et al., 2017). These nucleases introduce double-strand breaks (DSBs) with extreme specificity at the target genomic position. CRISPR/Cas9 is the most flexible system: its cutting specificity relies on a short guide RNA (sgRNA) and only requires the additional presence of an adjacent genomic proto-spacer adjacent motif (PAM) for its cut [this limit is actually being solved by the "near-PAM-less"-engineered CRISPR-Cas9 variants (Walton et al., 2020)]. In order to minimize possible off targets, different solutions are under study: better algorithms for the prediction of optimal DNA targets, optimized sgRNAs, engineered proto-spacers and Cas enzymes improved on the basis of thermodynamical models (Chen, 2019).

Once generated, DSBs are resolved by different DNA repair cellular pathways:

(i) The homology-directed repair (HDR) high-fidelity system that uses a donor template (the sister identical chromatid in physiological conditions) to repair DSBs when cells are in S and G2 phases. The implication is that HDR is poorly efficient in non-dividing HSC (Dever and Porteus, 2017), the target cell for therapeutic correction of β-hemoglobinopathies.

(ii) The non-homologous end joining (NHEJ) error-prone system acting in all cell cycle phases that inserts small indels at the site of the lesion, resulting in the disruption of the target sequence.

(iii) The microhomology end-joining (MMEJ) (Wang and Xu, 2017) error-prone system, which exploits small homology domains to align the broken filaments and close the gap. This molecular mechanism introduces deletions encompassing the microhomology regions flanking the break sites.

On these premises, the design of HDR recombination-based therapeutic strategies is difficult because of the requirement for a codelivered donor DNA template, of the low efficiency of HDR in HSCs and of the competition of the unwanted NHEJ and MMEJ error-prone repair systems. Despite these problems, the correction of the SCD mutation in HSCs was obtained by using both ZnF (Hoban et al., 2015) and CRISPR nucleases (Dever et al., 2016). However, the efficiency of the correction, assessed in HSCs *in vitro*, dramatically decreased after transplantation *in vivo*, confirming that HSCs are more resistant to HDR-based editing than more mature progenitors (Hoban et al., 2015) and that a selection step could be required to enrich for HSC-edited cells, capable of long-term correction *in vivo* (Dever et al., 2016). Instead, NHEJ is more flexible and allow to reach an efficiency up to ≈90% of edited HSCs that is maintained *in vivo* (Genovese et al., 2014; Chang et al., 2017; Charlesworth et al., 2018; Psatha et al., 2018; Wu et al., 2019).

The "perfect" editing should leave no trace, to avoid unintended off-target mutations and should at the same time guarantee high editing efficiency with reduced toxicity for HSCs. To reach this goal, an intense optimization work has been focused on the different steps of the genome editing procedure: the development of new editing reagents [single-strand DNA donor templates (Park et al., 2019), modified sgRNA (De Ravin et al., 2017; Park et al., 2019), pre-complexed ribonucleoproteins (RNPs) (Gundry et al., 2016)] and their integration in improved platforms for their delivery (Lino et al., 2018; Lattanzi et al., 2019; Schiroli et al., 2019). This massive effort finally led to

the generation of selection-free HSCs of therapeutic potential (Genovese et al., 2014; DeWitt et al., 2016; Porteus, 2016; Yu et al., 2016; Wu et al., 2019).

THE EDITING OF GLOBIN GENES INSPIRED BY HPFH

NHEJ has been used to generate two classes of HPFH-inspired mutation: large deletions within the β-locus, to remove putative γ-globin repressive regions, and small indels within the γ-globin promoter or within the regulatory regions driving the erythroid expression of the γ-globin repressor BCL11A (Bauer et al., 2013; Canver et al., 2015). Large deletions focus around the critical "HPFH γδ-region" 5′ to the δ gene, deleted with different breakpoints in several HPFH [https://www.omim.org/entry/141749; https://www.ithanet.eu/; (Kountouris et al., 2014)]. This region is generally lost in deletions involving δ and β genes and causing HPFH, whereas it is retained in δβ-thalassemia deletions, which similarly remove δ and β genes but with little increase of γ-globin expression. This observation led to hypothesize that this region contains an element capable to repress the γ genes in *cis*. The CRISPR-mediated deletion corresponding to the 12.9-kb Sicilian HPFH, spanning from 3.2 kb upstream of the δ gene to the 3′ flanking region of the β gene, gave indeed a HPFH phenotype (increase in γ-globin with concomitant drop of β-globin expression) in HUDEP cells and in human *ex vivo* HSC-derived erythroblasts (Ye et al., 2016). The same result (γ-globin increased and β-globin decreased) was obtained by the CRISPR-mediated deletion (or inversion) of a large 13.6-kb region starting downstream to the pseudo-β1 (HBBP1) gene and extending into the β gene (Antoniani et al., 2018). The 5′ border of this deletion corresponds to the 5′ breakpoint of the Corfù δβ-thal 7.2-kb deletion (Wainscoat et al., 1985) that ends in the δ gene and is not associated with HPFH *in vivo* in humans (except in some rare cases, in homozygotes, in which an additional independent mutation in the downstream β gene is present (Kulozik et al., 1988). The CRISPR-mediated deletion of the 7.2-kb Corfù region and of two smaller internal regions of 3.5 and 1.7 kb, centered around a BCL11A binding site and a polypyrimidine stretch (**Figure 1B**), thought to mediate γ-globin repression (Sankaran et al., 2011), resulted in a very little γ-globin increase (Antoniani et al., 2018; Chung et al., 2019). These results indicate that the 1.7-kb element and its surrounding sequences *per se* are not an autonomous γ-globin silencer, as also suggested by previous studies (Galanello et al., 1990; Calzolari et al., 1999; Gaensler et al., 2003; Chakalova et al., 2005). Instead, they suggest a more complex scenario, where the competition with β-globin expression, the perfect distance/order between intergenic enhancer/repressor, and the enhancers delimitating the locus (the LCR and the 3′DNAseI hypersensitive site), all together concur to the correct γ/β gene expression (and to γ-globin increase, when perturbed in HPFH).

The effects of distorting the architecture of the β-locus can be turned in an advantage: Dr. Blobel and colleagues obtained a great increase in γ-globin (with β-globin reduction) by tethering LDB1 to the LCR and to the γ-globin promoter, thus forcing their looping (Deng et al., 2014). This result again highlights the importance of the competition between γ and β genes for the LCR.

HPFH mutation mapping within the γ-globin promoter alters the binding of transcription factors/cofactors. Theoretically, the γ-globin upregulation can be obtained either by increasing the binding of an activator or by decreasing the binding of a repressor. Both cases are observed in HPFH. Mutations at positions −198, −175, and −113 create new binding sites for erythroid transcriptional activators [KLF1 (Wienert et al., 2017), TAL1 (Wienert et al., 2015), GATA1 (Martyn et al., 2019), respectively]. Other mutations clustered around position −200 and around the distal CCAAT box (−115) reduce the binding of the γ-globin repressors LRF and BCL11A, respectively (Liu et al., 2018; Martyn et al., 2018).

Consistently, the CRISPR-mediated disruption of these two binding sites resulted in a relevant increase in γ-globin expression (Traxler et al., 2016; Weber et al., 2020). Of note, the editing of the −158 ("XmnI-Gγ-site"), known to be influenced by a QTL on chromosome 8 (Garner et al., 2002), only marginally increased γ-globin expression (Weber et al., 2020), suggesting that possible background effects might be taken into account when considering editing for therapeutic purposes.

The existence of two highly homologous γ-globin genes poses specific editing issues: the double-stranded DNA cut at the gRNA recognition sites in the HBG2 and HBG1 promoters could result in NHEJ-mediated joining of the two ends with loss of the intergenic (≈5 kb) genomic sequence in variable proportion (Traxler et al., 2016; Antoniani et al., 2018). Thus, the editing of these γ-globin regions can result either in the mutation of a single or both HBG genes or in the deletion of the intergenic region, with different resulting percentages of γ-globin induction. Moreover, given the presence of short repeats within the promoter, MMEJ can also occur (Traxler et al., 2016; Weber et al., 2020).

NHEJ can also be used to destroy the specific erythroid expression of repressors, such as BCL11A or, in principle, of LRF (both proteins have important roles in other hematopoietic cell types that must be preserved). On this front, four clinical trials based on targeting a GATA1-binding site within the intronic +58 (Canver et al., 2015) erythroid-specific BCL11A enhancer are ongoing (Hirakawa et al., 2020).

Theoretically, all the different genes involved in the γ-globin repression identified so far, including BCL11A, LRF, SOX6, and DRED are possible targets for genome editing, with the general caveat that their ablation should not perturb stem cell viability, their engraftment and differentiation potential. For example, the ubiquitous knockdown of BCL11A impairs normal HSC function and lymphopoiesis (Luc et al., 2016); LRF (Maeda et al., 2009), SOX6 (Cantu et al., 2011), and KLF1 (Nuez et al., 1995; Perkins et al., 1995) are instead required for proper erythroid differentiation. In this latter case, the need of fine-tuning the downregulation of the γ-globin repressor in order to lead to an appreciable γ-globin increase while maintaining a correct erythroid differentiation could represent an insurmountable obstacle.

Overall, the success obtained in reactivating γ-globin expression at therapeutic levels demonstrates that this strategy can work. Theoretically, the possibility to generate multiple HPFH mutations could further increase γ-globin expression.

Importantly, beside the final goal of its clinical application, the relative ease-of-use of the CRISPR-based editing techniques represents a formidable tool to answer the unsolved questions on the molecular mechanisms regulating the hemoglobin switching.

BEYOND CRISPR/CAS9: PRIME-CAS AND BE-CAS

The use of HDR to correct β-disease mutations is limited by its low efficiency and by the downstream activation of p53, which can induce toxicity and, even worse, the possible selection of potentially harmful $p53^{low}$ cells (Haapaniemi et al., 2018; Ihry et al., 2018). To overcome this problem, a new generation of engineered Cas9 that do not introduce DBSs (also avoiding unwanted NHEJ/MMEJ events triggered by the DBSs) and do not require donor DNA are under development. They rely on catalytically inactive Cas9 fused to a modified reverse transcriptase (prime editing) or to base-specific DNA deaminase enzymes [base editors (BEs)]. As for many innovations, these newcomers in the CRISPR toolbox have been tested on β-disease mutations. Prime editing has been used to correct the β6Glu>Val SCD mutation artificially introduced in HEK-293T cells (Anzalone et al., 2019). Dr. Bauer and colleagues recently demonstrated the versatility of BE by disrupting the GATA1-binding site within the +58 BCL11A erythroid-specific enhancer. The obtained HSC-edited cells express HbF at levels similar to those obtained by the NHEJ-mediated disruption of the same site and are capable of multi-lineage repopulation in serial transplantation experiments (Zeng et al., 2020). In addition, the simultaneous multiplex edit of the β-thal −28 A>G mutation in the TATA box of the β promoter increased β-globin production in the same cells.

Instead, Beam Therapeutics recently presented data relative to two therapeutic approaches based on BE, the first recreating an HPFH mutation and the second converting HbS into HbG-Makassar, a naturally occurring human variant that does not cause sickling[1].

Although at present the issues of unwanted bystander/off target mutations remain to be explored, it is clear that Prime editing and BE represent important new instruments for genome

[1]https://investors.beamtx.com/news-releases/news-release-details/beam-therapeutics-reports-additional-data-asgct-annual-meeting

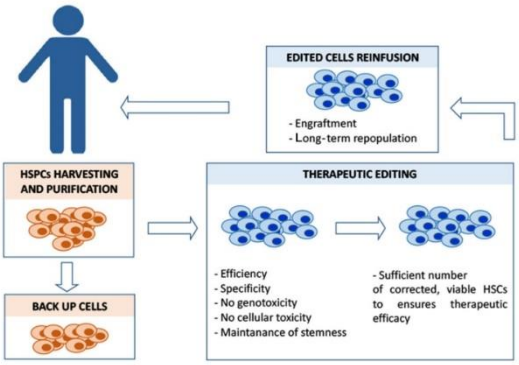


FIGURE 2 | Overview of the different steps of HSC autologous transplantation and of its major critical issues. In the case of β-hemoglobinopathies, ineffective erythropoiesis and a compromised bone marrow microenvironment sensibly reduce the yield of $CD34^+$ HSCs amenable for the editing process, posing a serious problem in a clinical-scale setting. Different mobilization protocols are currently used to maximize the yield of harvested hematopoietic stem progenitor cells (HSPCs) that must also include backup cells to be reinfused into the patient in case of engraftment failure. Editing should ensure efficiency (in terms of complete allelic correction and percentage of edited cells) and, at the same time, minimize the exposure to editing reagents, to reduce the risk of unwanted mutations. Editing manipulations must preserve the population of $CD34^+$ long-term repopulating cells. Before the reinfusion of the edited cells, the patient is treated with myeloablative agents to maximize the engraftment of the edited cells within the bone marrow niche. The conditioning regimen should be designed to guarantee the optimal risk-benefit balance between toxicity and efficacy of the engraftment, in order to achieve a stable, long-term therapeutic bone marrow repopulation.

editing, with the perspective to become even more attractive with the ongoing development of BE enabling more transition substitutions (Komor et al., 2018).

CONCLUDING REMARKS

The number of genome-editing tools is rapidly increasing (Papasavva et al., 2019; Doudna, 2020), holding the promise to reach in the near future a safe, precise, and efficient editing of β-disease mutations, via different strategies. The availability of different molecular options (HDR, NHEJ, and BE based, **Figure 1**) poses the problem of the evaluation of the pros and cons of each strategy (Ikawa et al., 2019; Papasavva et al., 2019): HDR-based approaches could ensure a higher precision at the expenses of HSC correction efficiency, whereas NHEJ is more efficient but less precise. Base editing, which does not require double-strand breaks, could be a safer option when a nucleotide substitution is required. Beside the choice of the optimal genetic modification, other issues remain open, first of all those related to unforeseeable genotoxicity (with the serious concern of inducing hyperproliferative/leukemic mutations in HSCs), the efficiency of the correction and the optimization of the delivery of genome editing reagents to target cells in conditions that preserve their stemness. Moreover, the clinical translation of these approaches requires the definition of scalable protocols to obtain under non-invasive conditions, a sufficient number of autologous HSCs amenable for the editing procedures and capable of optimal engraftment (in addition to backup cells to be reinfused in the patient in the case of engraftment failure) (**Figure 2**). This last point involves the identification of the best preparative conditioning regimen of the patient to allow efficient engraft of the corrected HSCs within the recipient niche (Psatha et al., 2016). Despite these difficulties, the recent announcement of the curative response of the first three patients (carrying a transfusion-dependent β-thalassemia and the SCD mutation) with CRISPR-Cas9-edited cells targeting BCL11A (CRISPR Therapeutics and Vertex CTX001 clinical trial[2&3]), clearly highlights the clinical potential of gene therapy. The advent of this new era urges the need to make these approaches affordable and available in low-resource settings/countries, where a large number of patients is waiting for a cure.

AUTHOR CONTRIBUTIONS

AR conceived and wrote the manuscript. GB contributed with ideas and discussion. GB and AL contributed in the organization of the manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

This project has received funding from the European Union's Horizon 2020 research and innovation program under the Marie Skłodowska-Curie grant agreement No. 813091 [age-related changes in hematopoiesis (ARCH)].

ACKNOWLEDGMENTS

We are indebted with Dr. Sergio Ottolenghi for the generous sharing of his deep knowledge of the β-locus and of its mutations and with Christina Pitsillidou and Sarah Stucchi for their suggestions in preparing figures.

[2]https://www.globenewswire.com/news-release/2020/06/12/2047260/0/en/CRISPR-Therapeutics-and-Vertex-Announce-New-Clinical-Data-for-Investigational-Gene-Editing-Therapy-CTX001-in-Severe-Hemoglobinopathies-at-the-25th-Annual-European-Hematology-Associ.html (accessed October 5, 2020).

[3]http://ir.crisprtx.com/news-releases/news-release-details/crispr-therapeutics-and-vertex-pharmaceuticals-announce-priority (accessed October 5, 2020).

REFERENCES

Antoniani, C., Meneghini, V., Lattanzi, A., Felix, T., Romano, O., Magrin, E., et al. (2018). Induction of fetal hemoglobin synthesis by CRISPR/Cas9-mediated editing of the human beta-globin locus. *Blood* 131, 1960–1973. doi: 10.1182/blood-2017-10-811505

Anzalone, A. V., Randolph, P. B., Davis, J. R., Sousa, A. A., Koblan, L. W., Levy, J. M., et al. (2019). Search-and-replace genome editing without double-strand breaks or donor DN. A. *Nature* 576, 149–157. doi: 10.1038/s41586-019-1711-4

Aprile, A., Gulino, A., Storto, M., Villa, I., Beretta, S., Merelli, I., et al. (2020). Hematopoietic stem cell function in beta-thalassemia is impaired and is rescued by targeting the bone marrow niche. *Blood* 136–610–22. doi: 10.1182/blood.2019002721

Bauer, D. E., Kamran, S. C., Lessard, S., Xu, J., Fujiwara, Y., Lin, C., et al. (2013).An erythroid enhancer of BCL11A subject to genetic variation determines fetal hemoglobin level. *Science* 342, 253–257. doi: 10.1126/science.1242088

Berry, M., Grosveld, F., and Dillon, N. (1992). A single point mutation is the cause of the Greek form of hereditary persistence of fetal haemoglobin. *Nature* 358, 499–502. doi: 10.1038/358499a0

Borg, J., Papadopoulos, P., Georgitsi, M., Gutierrez, L., Grech, G., Fanis, P., et al. (2010). Haploinsufficiency for the erythroid transcription factor KLF1 causes hereditary persistence of fetal hemoglobin. *Nat. Genet.* 42, 801–805. doi: 10.1038/ng.630

Cai, L., Bai, H., Mahairaki, V., Gao, Y., He, C., Wen, Y., et al. (2018). A universal approach to correct various HBB gene mutations in human stem cells for gene therapy of beta-thalassemia and sickle cell disease. *Stem Cells Transl. Med.* 7, 87–97. doi: 10.1002/sctm.17-0066

Calzolari, R., McMorrow, T., Yannoutsos, N., Langeveld, A., and Grosveld, F. (1999). Deletion of a region that is a candidate for the difference between the deletion forms of hereditary persistence of fetal hemoglobin and deltabeta-thalassemia affects beta- but not gamma-globin gene expression. *EMBO J.* 18, 949–958. doi: 10.1093/emboj/18.4.949

Cantu, C., Ierardi, R., Alborelli, I., Fugazza, C., Cassinelli, L., Piconese, S., et al. (2011). Sox6 enhances erythroid differentiation in human erythroid progenitors. *Blood* 117, 3669–3679. doi: 10.1182/blood-2010-04-282350

Canver, M. C., Smith, E. C., Sher, F., Pinello, L., Sanjana, N. E., Shalem, O., et al. (2015). BCL11A enhancer dissection by Cas9-mediated *in situ* saturating mutagenesis. *Nature* 527, 192–197. doi: 10.1038/nature15521

Carter, D., Chakalova, L., Osborne, C. S., Dai, Y. F., and Fraser, P. (2002). Long-range chromatin regulatory interactions *in vivo*. *Nat. Genet.* 32, 623–626. doi: 10.1038/ng1051

Chakalova, L., Osborne, C. S., Dai, Y. F., Goyenechea, B., Metaxotou-Mavromati, A., Kattamis, A., et al. (2005). The Corfu deltabeta thalassemia deletion disrupts gamma-globin gene silencing and reveals post-transcriptional regulation of HbF expression. *Blood* 105, 2154–2160. doi: 10.1182/blood-2003-11-4069

Chang, K. H., Smith, S. E., Sullivan, T., Chen, K., Zhou, Q., West, J. A., et al. (2017). Long-term engraftment and fetal globin induction upon BCL11A gene editing in bone-marrow-derived CD34(+) hematopoietic stem and progenitor cells. *Mol Ther Methods Clin Dev.* 4, 137–148. doi: 10.1016/j.omtm.2016.12.009

Charlesworth, C. T., Camarena, J., Cromer, M. K., Vaidyanathan, S., Bak, R. O., Carte, J. M., et al. (2018). Priming human repopulating hematopoietic stem and progenitor cells for Cas9/sgRNA gene targeting. *Mol. Ther. Nucleic Acids* 12, 89–104. doi: 10.1016/j.omtn.2018.04.017

Chen, S. J. (2019). Minimizing off-target effects in CRISPR-Cas9 genome editing. *Cell Biol. Toxicol.* 35, 399–401. doi: 10.1007/s10565-019-09486-4

Chung, J. E., Magis, W., Vu, J., Heo, S. J., Wartiovaara, K., Walters, M. C., et al. (2019). CRISPR-Cas9 interrogation of a putative fetal globin repressor in human erythroid cells. *PLoS ONE* 14:e0208237. doi: 10.1371/journal.pone.0208237

Cornu, T. I., Mussolino, C., and Cathomen, T. (2017). Refining strategies to translate genome editing to the clinic. *Nat. Med.* 23, 415–423. doi: 10.1038/nm.4313

De Ravin, S. S., Li, L., Wu, X., Choi, U., Allen, C., Koontz, S., et al. (2017). CRISPR-Cas9 gene repair of hematopoietic stem cells from patients with X-linked chronic granulomatous disease. *Sci Transl Med.* 9, 1–10. doi: 10.1126/scitranslmed.aah3480

Deng, W., Rupon, J. W., Krivega, I., Breda, L., Motta, I., Jahn, K. S., et al. (2014). Reactivation of developmentally silenced globin genes by forced chromatin looping. *Cell* 158, 849–860. doi: 10.1016/j.cell.2014.05.050

Dever, D. P., Bak, R. O., Reinisch, A., Camarena, J., Washington, G., Nicolas, C. E., et al. (2016). CRISPR/Cas9 beta-globin gene targeting in human haematopoietic stem cells. *Nature* 539, 384–389. doi: 10.1038/nature20134

Dever, D. P., and Porteus, M. H. (2017). The changing landscape of gene editing in hematopoietic stem cells: a step towards Cas9 clinical translation. *Curr. Opin. Hematol.* 24, 481–488. doi: 10.1097/MOH.0000000000000385

DeWitt, M. A., Magis, W., Bray, N. L., Wang, T., Berman, J. R., Urbinati, F., et al. (2016). Selection-free genome editing of the sickle mutation in human adult hematopoietic stem/progenitor cells. *Sci. Transl. Med.* 8:360ra134. doi: 10.1126/scitranslmed.aaf9336

Doudna, J. A. (2020). The promise and challenge of therapeutic genome editing. *Nature* 578, 229–236. doi: 10.1038/s41586-020-1978-5

Forget, B. G. (1998). Molecular basis of hereditary persistence of fetal hemoglobin. *Ann. N. Y. Acad. Sci.* 850, 38–44. doi: 10.1111/j.1749-6632.1998.tb10460.x

Fraser, P., and Grosveld, F. (1998). Locus control regions, chromatin activation and transcription. *Curr. Opin. Cell Biol.* 10, 361–365. doi: 10.1016/S0955-0674(98)80012-4

Fritsch, E. F., Lawn, R. M., and Maniatis, T. (1980). Molecular cloning and characterization of the human beta-like globin gene cluster. *Cell* 19, 959–972. doi: 10.1016/0092-8674(80)90087-2

Gaensler, K. M., Zhang, Z., Lin, C., Yang, S., Hardt, K., and Flebbe-Rehwaldt, L. (2003). Sequences in the (A)gamma-delta intergenic region are not required for stage-specific regulation of the human beta-globin gene locus. *Proc. Natl. Acad. Sci. U.S.A.* 100, 3374–3379. doi: 10.1073/pnas.0634132100

Galanello, R., Melis, M. A., Podda, A., Monne, M., Perseu, L., Loudianos, G., et al. (1990). Deletion delta-thalassemia: the 7.2 kb deletion of Corfu delta beta-thalassemia in a non-beta-thalassemia chromosome. *Blood* 75, 1747–1749. doi: 10.1182/blood.V75.8.1747.1747

Garner, C. P., Tatu, T., Best, S., Creary, L., and Thein, S. L. (2002). Evidence of genetic interaction between the beta-globin complex and chromosome 8q in the expression of fetal hemoglobin. *Am. J. Hum. Genet.* 70, 793–799. doi: 10.1086/339248

Genovese, P., Schiroli, G., Escobar, G., Tomaso, T. D., Firrito, C., Calabria, A., et al. (2014). Targeted genome editing in human repopulating haematopoietic stem cells. *Nature* 510, 235–240. doi: 10.1038/nature13420

Grosveld, F., van Assendelft, G. B., Greaves, D. R., and Kollias, G. (1987). Position-independent, high-level expression of the human beta-globin gene in transgenic mice. *Cell* 51, 975–985. doi: 10.1016/0092-8674(87)90584-8

Gundry, M. C., Brunetti, L., Lin, A., Mayle, A. E., Kitano, A., Wagner, D., et al. (2016). Highly efficient genome editing of murine and human hematopoietic progenitor cells by CRISPR/Cas9. *Cell Rep.* 17, 1453–1461. doi: 10.1016/j.celrep.2016.09.092

Haapaniemi, E., Botla, S., Persson, J., Schmierer, B., and Taipale, J. (2018). CRISPR-Cas9 genome editing induces a p53-mediated DNA damage response. *Nat. Med.* 24, 927–930. doi: 10.1038/s41591-018-0049-z

Hirakawa, M. P., Krishnakumar, R., Timlin, J. A., Carney, J. P., and Butler, K. S. (2020). Gene editing and CRISPR in the clinic: current and future perspectives. *Biosci. Rep.* 40, 1–37. doi: 10.1042/BSR20200127

Hoban, M. D., Cost, G. J., Mendel, M. C., Romero, Z., Kaufman, M. L., Joglekar, A. V., et al. Correction of the sickle cell disease mutation in human hematopoietic stem/progenitor cells. *Blood* (2015). 125, 2597–604. doi: 10.1182/blood-2014-12-615948

Ihry, R. J., Worringer, K. A., Salick, M. R., Frias, E., Ho, D., Theriault, K., et al. (2018). p53 inhibits CRISPR-Cas9 engineering in human pluripotent stem cells. *Nat. Med.* 24, 939–946. doi: 10.1038/s41591-018-0050-6

Ikawa, Y., Miccio, A., Magrin, E., Kwiatkowski, J. L., Rivella, S., and Cavazzana, M. (2019). Gene therapy of hemoglobinopathies: progress and future challenges. *Hum. Mol. Genet.* 28:R24–30. doi: 10.1093/hmg/ddz172

Komor, A. C., Badran, A. H., and Liu, D. R. (2017). CRISPR-based technologies for the manipulation of eukaryotic genomes. *Cell* 168, 20–36. doi: 10.1016/j.cell.2016.10.044

Komor, A. C., Badran, A. H., and Liu, D. R. (2018). Editing the genome without double-stranded DNA breaks. *ACS Chem. Biol.* 13, 383–388. doi: 10.1021/acschembio.7b00710

Koren, A., Levin, C., Dgany, O., Kransnov, T., Elhasid, R., Zalman, L., et al. (2008). Response to hydroxyurea therapy in beta-thalassemia. *Am. J. Hematol.* 83, 366–370. doi: 10.1002/ajh.21120

Kountouris, P., Lederer, C. W., Fanis, P., Feleki, X., Old, J., and Kleanthous, M. (2014). IthaGenes: an interactive database for haemoglobin variations and epidemiology. *PLoS ONE* 9:e103020. doi: 10.1371/journal.pone.0103020

Kulozik, A. E., Yarwood, N., and Jones, R. W. (1988). The Corfu delta beta zero thalassemia: a small deletion acts at a distance to selectively abolish beta globin gene expression. *Blood* 71, 457–462. doi: 10.1182/blood.V71.2.457.457

Lattanzi, A., Meneghini, V., Pavani, G., Amor, F., Ramadier, S., Felix, T., et al. (2019). Optimization of CRISPR/Cas9 delivery to human hematopoietic stem and progenitor cells for therapeutic genomic rearrangements. *Mol. Ther.* 27, 137–150. doi: 10.1016/j.ymthe.2018.10.008

Lino, C. A., Harper, J. C., Carney, J. P., and Timlin, J. A. (2018). Delivering CRISPR: a review of the challenges and approaches. *Drug Deliv.* 25, 1234–1257. doi: 10.1080/10717544.2018.1474964

Liu, N., Hargreaves, V. V., Zhu, Q., Kurland, J. V., Hong, J., Kim, W., et al. (2018). Direct promoter repression by BCL11A controls the fetal to adult hemoglobin switch. *Cell* 173, 430–42 e17. doi: 10.1016/j.cell.2018.03.016

Luc, S., Huang, J., McEldoon, J. L., Somuncular, E., Li, D., Rhodes, C., et al. (2016). Bcl11a Deficiency leads to hematopoietic stem cell defects with an aging-like phenotype. *Cell Rep.* 16, 3181–3194. doi: 10.1016/j.celrep.2016.08.064

Maeda, T., Ito, K., Merghoub, T., Poliseno, L., Hobbs, R. M., Wang, G., et al. (2009). LRF is an essential downstream target of GATA1 in erythroid development and regulates BIM-dependent apoptosis. *Dev. Cell.* 17, 527–540. doi: 10.1016/j.devcel.2009.09.005

Magrin, E., Miccio, A., and Cavazzana, M. (2019). Lentiviral and genome-editing strategies for the treatment of beta-hemoglobinopathies. *Blood* 134, 1203–1213. doi: 10.1182/blood.2019000949

Martyn, G. E., Wienert, B., Kurita, R., Nakamura, Y., Quinlan, K. G. R., and Crossley, M. (2019). A natural regulatory mutation in the proximal promoter elevates fetal globin expression by creating a de novo GATA1 site. *Blood* 133, 852–856. doi: 10.1182/blood-2018-07-863951

Martyn, G. E., Wienert, B., Yang, L., Shah, M., Norton, L. J., Burdach, J., et al. (2018). Natural regulatory mutations elevate the fetal globin gene via disruption of BCL11A or ZBTB7A binding. *Nat. Genet.* 50, 498–503. doi: 10.1038/s41588-018-0085-0

Masuda, T., Wang, X., Maeda, M., Canver, M. C., Sher, F., Funnell, A. P., et al. (2016). Transcription factors LRF and BCL11A independently

repress expression of fetal hemoglobin. *Science* 351, 285–289. doi: 10.1126/science.aad3312

Menzel, S., Garner, C., Gut, I., Matsuda, F., Yamaguchi, M., Heath, S., et al. (2007). A QTL influencing F cell production maps to a gene encoding a zinc-finger protein on chromosome 2p15. *Nat. Genet.* 39, 1197–1199. doi: 10.1038/ng2108

Mettananda, S., Fisher, C. A., Hay, D., Badat, M., Quek, L., Clark, K., et al. (2017). Editing an alpha-globin enhancer in primary human hematopoietic stem cells as a treatment for beta-thalassemia. *Nat. Commun.* 8:424. doi: 10.1038/s41467-017-00479-7

Mettananda, S., Gibbons, R. J., and Higgs, D. R. (2015). alpha-Globin as a molecular target in the treatment of beta-thalassemia. *Blood* 125, 3694–3701. doi: 10.1182/blood-2015-03-633594

Muirhead, H., Cox, J. M., Mazzarella, L., and Perutz, M. F. (1967). Structure and function of haemoglobin. 3. A three-dimensional fourier synthesis of human deoxyhaemoglobin at 5.5 Angstrom resolution. *J. Mol. Biol.* 28, 117–56. doi: 10.1016/S0022-2836(67)80082-2

Nuez, B., Michalovich, D., Bygrave, A., Ploemacher, R., and Grosveld, F. (1995). Defective haematopoiesis in fetal liver resulting from inactivation of the EKLF gene. *Nature* 375, 316–318. doi: 10.1038/375316a0

Ottolenghi, S., Giglioni, B., Taramelli, R., Comi, P., and Gianni, A. M. (1982). Delta beta-Thalassemia and HPF. H. *Birth Defects Orig. Artic. Ser.* 18, 65–67.

Palstra, R. J., Tolhuis, B., Splinter, E., Nijmeijer, R., Grosveld, F., and de Laat, W. (2003). The beta-globin nuclear compartment in development and erythroid differentiation. *Nat. Genet.* 35, 190–194. doi: 10.1038/ng1244

Papasavva, P., Kleanthous, M., and Lederer, C. W. (2019). Rare opportunities: CRISPR/Cas-based therapy development for rare genetic diseases. *Mol. Diagn. Ther.* 23, 201–222. doi: 10.1007/s40291-019-00392-3

Park, S. H., Lee, C. M., Dever, D. P., Davis, T. H., Camarena, J., Srifa, W., et al. (2019). Highly efficient editing of the beta-globin gene in patient-derived hematopoietic stem and progenitor cells to treat sickle cell disease. *Nucleic Acids Res.* 47, 7955–7972. doi: 10.1093/nar/gkz475

Perkins, A. C., Sharpe, A. H., and Orkin, S. H. (1995). Lethal beta-thalassaemia in mice lacking the erythroid CACCC-transcription factor EKL. F. *Nature* 375, 318–322. doi: 10.1038/375318a0

Pevny, L., Simon, M. C., Robertson, E., Klein, W. H., Tsai, S. F., D'Agati, V., et al. (1991). Erythroid differentiation in chimaeric mice blocked by a targeted mutation in the gene for transcription factor GATA-1. *Nature* 349, 257–260. doi: 10.1038/349257a0

Platt, O. S. (2008). Hydroxyurea for the treatment of sickle cell anemia. *N. Engl. J. Med.* 358, 1362–1369. doi: 10.1056/NEJMct0708272

Porteus, M. (2016). Genome editing: a new approach to human therapeutics. *Annu. Rev. Pharmacol. Toxicol.* 56, 163–190. doi: 10.1146/annurev-pharmtox-010814-124454

Pourfarzad, F., von Lindern, M., Azarkeivan, A., Hou, J., Kia, S. K., Esteghamat, F., et al. (2013). Hydroxyurea responsiveness in beta-thalassemic patients is determined by the stress response adaptation of erythroid progenitors and their differentiation propensity. *Haematologica* 98, 696–704. doi: 10.3324/haematol.2012.074492

Psatha, N., Karponi, G., and Yannaki, E. (2016). Optimizing autologous cell grafts to improve stem cell gene therapy. *Exp. Hematol.* 44, 528–539. doi: 10.1016/j.exphem.2016.04.007

Psatha, N., Reik, A., Phelps, S., Zhou, Y., Dalas, D., Yannaki, E., et al. (2018). Disruption of the BCL11A erythroid enhancer reactivates fetal hemoglobin in erythroid cells of patients with beta-thalassemia major. *Mol Ther Methods Clin Dev.* 10, 313–326. doi: 10.1016/j.omtm.2018.08.003

Rivella, S. (2012). The role of ineffective erythropoiesis in non-transfusion-dependent thalassemia. *Blood Rev.* 26(Suppl. 1):S12–S15. doi: 10.1016/S0268-960X(12)70005-X

Sankaran, V. G., Menne, T. F., Xu, J., Akie, T. E., Lettre, G., Van Handel, B., et al. (2008). Human fetal hemoglobin expression is regulated by the developmental stage-specific repressor BCL11A. *Science* 322, 1839–1842. doi: 10.1126/science.1165409

Sankaran, V. G., Xu, J., Byron, R., Greisman, H. A., Fisher, C., Weatherall, D. J., et al. (2011). A functional element necessary for fetal hemoglobin silencing. *N. Engl. J. Med.* 365, 807–814. doi: 10.1056/NEJMoa1103070

Schiroli, G., Conti, A., Ferrari, S., Della Volpe, L., Jacob, A., Albano, L., et al. (2019). Precise gene editing preserves hematopoietic stem cell function following transient p53-mediated DNA damage response. *Cell Stem Cell* 24, 551–65 e8. doi: 10.1016/j.stem.2019.02.019

Smithies, O., Gregg, R. G., Boggs, S. S., Koralewski, M. A., and Kucherlapati, R. S. (1985). Insertion of DNA sequences into the human chromosomal beta-globin locus by homologous recombination. *Nature* 317, 230–234. doi: 10.1038/317230a0

Stamatoyannopoulos, G. (1991). Human hemoglobin switching. *Science* 252:383. doi: 10.1126/science.2017679

Steinberg, M. H., Chui, D. H., Dover, G. J., Sebastiani, P., and Alsultan, A. (2014). Fetal hemoglobin in sickle cell anemia: a glass half full? *Blood* 123, 481–485. doi: 10.1182/blood-2013-09-528067

Taher, A. T., Musallam, K. M., and Cappellini, M. D. (2009). Thalassaemia intermedia: an update. *Mediterr. J. Hematol. Infect. Dis.* 1:e2009004. doi: 10.4084/MJHID.2009.004

Thein, S. L. (2008). Genetic modifiers of the beta-haemoglobinopathies. *Br. J. Haematol.* 141, 357–366. doi: 10.1111/j.1365-2141.2008.07084.x

Thein, S. L. (2013). The molecular basis of beta-thalassemia. *Cold Spring Harb. Perspect. Med.* 3:a011700. doi: 10.1101/cshperspect.a011700

Tolhuis, B., Palstra, R. J., Splinter, E., Grosveld, F., and de Laat, W. (2002). Looping and interaction between hypersensitive sites in the active beta-globin locus. *Mol. Cell* 10, 1453–1465. doi: 10.1016/S1097-2765(02)00781-5

Traxler, E. A., Yao, Y., Wang, Y. D., Woodard, K. J., Kurita, R., Nakamura, Y., et al. (2016). A genome-editing strategy to treat beta-hemoglobinopathies that recapitulates a mutation associated with a benign genetic condition. *Nat. Med.* 22, 987–990. doi: 10.1038/nm.4170

Uda, M., Galanello, R., Sanna, S., Lettre, G., Sankaran, V. G., Chen, W., et al. (2008). Genome-wide association study shows BCL11A associated with persistent fetal hemoglobin and amelioration of the phenotype of beta-thalassemia. *Proc. Natl. Acad. Sci. U.S.A.* 105, 1620–1625. doi: 10.1073/pnas.0711566105

Vega, M. A. (1991). Prospects for homologous recombination in human gene therapy. *Hum. Genet.* 87, 245–253. doi: 10.1007/BF00200899

Viprakasit, V., Gibbons, R. J., Broughton, B. C., Tolmie, J. L., Brown, D., Lunt, P., et al. (2001). Mutations in the general transcription factor TFIIH result in beta-thalassaemia in individuals with trichothiodystrophy. *Hum. Mol. Genet.* 10, 2797–2802. doi: 10.1093/hmg/10.24.2797

Wainscoat, J. S., Thein, S. L., Wood, W. G., Weatherall, D. J., Metaxotou-Mavromati, A., Tzotos, S., et al. (1985). A novel deletion in the beta-globin gene complex. *Ann. N. Y. Acad. Sci.* 445, 20–27. doi: 10.1111/j.1749-6632.1985.tb17171.x

Walton, R. T., Christie, K. A., Whittaker, M. N., and Kleinstiver, B. P. (2020). Unconstrained genome targeting with near-PAMless engineered CRISPR-Cas9 variants. *Science* 368, 290–296. doi: 10.1126/science.aba8853

Wang, H., and Xu, X. (2017). Microhomology-mediated end joining: new players join the team. *Cell Biosci.* 7:6. doi: 10.1186/s13578-017-0136-8

Weatherall, D. J. (2008). Hemoglobinopathies worldwide: present and future. *Curr. Mol. Med.* 8, 592–599. doi: 10.2174/156652408786241375

Weber, L., Frati, G., Felix, T., Hardouin, G., Casini, A., Wollenschlaeger, C., et al. (2020). Editing a gamma-globin repressor binding site restores fetal hemoglobin synthesis and corrects the sickle cell disease phenotype. *Sci Adv.* 6:eaay9392. doi: 10.1126/sciadv.aay9392

Wienert, B., Funnell, A. P., Norton, L. J., Pearson, R. C., Wilkinson-White, L. E., Lester, K., et al. (2015). Editing the genome to introduce a beneficial naturally occurring mutation associated with increased fetal globin. *Nat. Commun.* 6:7085. doi: 10.1038/ncomms8085

Wienert, B., Martyn, G. E., Kurita, R., Nakamura, Y., Quinlan, K. G. R., and Crossley, M. (2017). KLF1 drives the expression of fetal hemoglobin in British HPF. H. *Blood* 130, 803–807. doi: 10.1182/blood-2017-02-767400

Wilber, A., Nienhuis, A. W., and Persons, D. A. (2011). Transcriptional regulation of fetal to adult hemoglobin switching: new therapeutic opportunities. *Blood* 117, 3945–3953. doi: 10.1182/blood-2010-11-316893

Wu, Y., Zeng, J., Roscoe, B. P., Liu, P., Yao, Q., Lazzarotto, C. R., et al. (2019). Highly efficient therapeutic gene editing of human hematopoietic stem cells. *Nat. Med.* 25, 776–783. doi: 10.1038/s41591-019-0401-y

Ye, L., Wang, J., Tan, Y., Beyer, A. I., Xie, F., Muench, M. O., et al. (2016). Genome editing using CRISPR-Cas9 to create the HPFH genotype in HSPCs:

an approach for treating sickle cell disease and beta-thalassemia. *Proc. Natl. Acad. Sci. U.S.A.* 113, 10661–10665. doi: 10.1073/pnas.1612075113

Yu, C., Niakan, K. K., Matsushita, M., Stamatoyannopoulos, G., Orkin, S. H., and Raskind, W. H. (2002). X-linked thrombocytopenia with thalassemia from a mutation in the amino finger of GATA-1 affecting DNA binding rather than FOG-1 interaction. *Blood* 100, 2040–2045. doi: 10.1182/blood-2002-02-0387

Yu, K. R., Natanson, H., and Dunbar, C. E. (2016). Gene editing of human hematopoietic stem and progenitor cells: promise and potential hurdles. *Hum. Gene Ther.* 27, 729–740. doi: 10.1089/hum.2016.107

Yu, L., Myers, G., and Engel, J. D. (2020). Small molecule therapeutics to treat the beta-globinopathies. *Curr. Opin. Hematol.* 27, 129–140. doi: 10.1097/MOH.0000000000000579

Zeng, J., Wu, Y., Ren, C., Bonanno, J., Shen, A. H., Shea, D., et al. (2020). Therapeutic base editing of human hematopoietic stem cells. *Nat. Med.* 26, 535–541. doi: 10.1038/s41591-020-0790-y

Zhou, D., Liu, K., Sun, C. W., Pawlik, K. M., and Townes, T. M. (2010). KLF1 regulates BCL11A expression and gamma- to beta-globin gene switching. *Nat. Genet.* 42, 742–744. doi: 10.1038/ng.637

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

1.5. Aims of the thesis

The understanding of the molecular mechanism of the differential globins genes expression during development is crucial to design strategies to treat patients with β-hemoglobinopathies. One of the therapeutical approaches is the reactivation of the fetal γ-globin in adult cells to substitute the defective adult β-globin. While the repressors of γ-globin were widely studied, less is known about its activators. The COUP-TFII transcription factor has been identified in our lab as a specific γ-globin activator, and therefore it represents a very attractive therapeutical target. However, the mechanisms that regulates the expression of COUP-TFII, as well as its role in erythroid cells remain largely unknown. Hence, the main aim of this study is to characterize COUP-TFII erythroid networks. This would be crucial to design a future strategy to reactivate COUP-TFII (and therefore induce γ-globin) expression in patients' adult erythroid cells.

In Chapter 2, in order to identify COUP-TFII regulatory regions in erythroid cells, I selected genomic sequences that interact with the COUP-TFII promoter as candidate distance regulatory elements, to further analyze their potential to modulate COUP-TFII expression in K562 erythroid cells. The region with the highest enhancer activity was next deleted by using a CRISPR/Cas9 approach to confirm its impact on the COUP-TFII expression level. Moreover, I describe a positive autoregulation of COUP-TFII in erythroid cells.

In Chapter 3, I investigated the effects of the COUP-TFII overexpression or knockout on proliferation, differentiation and cell cycle of K562 erythroid cells.

Chapter 4 introduces the downstream network of COUP-TFII in erythroid cells, with a special focus on relevant erythroid targets: the β-globin locus and the MYB transcription factor. Moreover, I identified other candidate direct target genes of COUP-TFII by merging RNA-seq and ChIP-seq datasets from K562 erythroid cells. Selected target genes were thereupon knocked-out in a CRISPR/Cas9 mini-screening, in order to assess their effect on the erythroid program.

In Chapter 5, I discuss all the results presented in the thesis and I address the future perspective of their significance in molecular and translational medicine.

1.6. References

Dzierzak, E., & Bigas, A. (2018). Blood development: hematopoietic stem cell dependence and independence. Cell stem cell, 22(5), 639-651.

Nandakumar, S. K., Ulirsch, J. C., & Sankaran, V. G. (2016). Advances in understanding erythropoiesis: evolving perspectives. British journal of haematology, 173(2), 206-218.

Palis, J. (2014). Primitive and definitive erythropoiesis in mammals. Frontiers in physiology, 5, 3.

Bunn, H. F. (2013). Erythropoietin. Cold Spring Harbor perspectives in medicine, 3(3), a011619.

Dzierzak, E., & Philipsen, S. (2013). Erythropoiesis: development and differentiation. Cold Spring Harbor perspectives in medicine, 3(4), a011601.

Epelman, S., Lavine, K. J., Beaudin, A. E., Sojka, D. K., Carrero, J. A., Calderon, B., ... & Mann, D. L. (2014). Embryonic and adult-derived resident cardiac macrophages are maintained through distinct mechanisms at steady state and during inflammation. Immunity, 40(1), 91-104.

Gomez Perdiguero, E., Klapproth, K., Schulz, C., Busch, K., Azzoni, E., Crozet, L., ... & Rodewald, H. R. (2015). Tissue-resident macrophages originate from yolk-sac-derived erythro-myeloid progenitors. Nature, 518(7540), 547-551.

Higgs, D. R., Wood, W. G., Jarman, A. P., Sharpe, J., Lida, J., Pretorius, I. M., & Ayyub, H. (1990). A major positive regulatory region located far upstream of the human alpha-globin gene locus. Genes & development, 4(9), 1588-1601.

Stamatoyannopoulos, G. (2005). Control of globin gene expression during development and erythroid differentiation. Experimental hematology, 33(3), 259-271.

Grosveld, F., van Assendelft, G. B., Greaves, D. R., & Kollias, G. (1987). Position-independent, high-level expression of the human β-globin gene in transgenic mice. Cell, 51(6), 975-985.

Sankaran, V. G., Xu, J., & Orkin, S. H. (2010). Advances in the understanding of haemoglobin switching. British journal of haematology, 149(2), 181-194.

Sankaran, V. G., & Orkin, S. H. (2013). The switch from fetal to adult hemoglobin. Cold Spring Harbor perspectives in medicine, 3(1), a011643.

Elagooz, R., Dhara, A. R., Gott, R. M., Adams, S. E., White, R. A., Ghosh, A., ... & Gnanapragasam, M. N. (2022). PUM1 mediates the posttranscriptional regulation of human fetal hemoglobin. Blood Advances, 6(23), 6016-6022.

Feng, R., Mayuranathan, T., Huang, P., Doerfler, P. A., Li, Y., Yao, Y., ... & Weiss, M. J. (2022). Activation of γ-globin expression by hypoxia-inducible factor 1α. Nature, 610(7933), 783-790.

Ronchi, A. E., Bottardi, S., Mazzucchelli, C., Ottolenghi, S., & Santoro, C. (1995). Differential Binding of the NFE3 and CP1/NFY Transcription Factors to the Human γ-and∈-Globin CCAAT Boxes (∗). Journal of Biological Chemistry, 270(37), 21934-21941.

Liberati, C., Ronchi, A., Lievens, P., Ottolenghi, S., & Mantovani, R. (1998). NF-Y organizes the γ-globin CCAAT boxes region. Journal of Biological Chemistry, 273(27), 16880-16889.

Filipe, A., Li, Q., Deveaux, S., Godin, I., Romeo, P. H., Stamatoyannopoulos, G., & Mignotte, V. (1999). Regulation of embryonic/fetal globin genes by nuclear hormone receptors: a novel perspective on hemoglobin switching. The EMBO journal, 18(3), 687-697.

Liberati, C., Cera, M. R., Secco, P., Santoro, C., Mantovani, R., Ottolenghi, S., & Ronchi, A. (2001). Cooperation and competition between the binding of COUP-TFII and NF-Y on human ε-and γ-globin gene promoters. Journal of Biological Chemistry, 276(45), 41700-41709.

Zhu, X., Wang, Y., Pi, W., Liu, H., Wickrema, A., & Tuan, D. (2012). NF-Y recruits both transcription activator and repressor to modulate tissue-and developmental stage-specific expression of human γ-globin gene.

Fugazza, C., Barbarani, G., Elangovan, S., Marini, M. G., Giolitto, S., Font-Monclus, I., ... & Ronchi, A. E. (2021). The Coup-TFII orphan nuclear receptor is an activator of the γ-globin gene. haematologica, 106(2), 474.

Qiu, Y., Krishnan, V., Zeng, Z., Gilbert, D. J., Copeland, N. G., Gibson, L., ... & Tsai, S. V. (1995). Isolation, characterization, and chromosomal localization of mouse and human COUP-TF I and II genes. Genomics, 29(1), 240-246.

Bertacchi, M., Parisot, J., & Studer, M. (2019). The pleiotropic transcriptional regulator COUP-TFI plays multiple roles in neural development and disease. Brain research, 1705, 75-94.

Tsai, S. Y., & Tsai, M. J. (1997). Chick ovalbumin upstream promoter-transcription factors (COUP-TFs): coming of age. Endocrine reviews, 18(2), 229-240.

Cooney, A. J., Tsai, S. Y., O'Malley, B. W., & Tsai, M. J. (1992). Chicken ovalbumin upstream promoter transcription factor (COUP-TF) dimers bind to different GGTCA response elements, allowing COUP-TF to repress hormonal induction of the vitamin D3, thyroid hormone, and retinoic acid receptors. Molecular and cellular biology, 12(9), 4153-4163.

Kruse, S. W., Suino-Powell, K., Zhou, X. E., Kretschman, J. E., Reynolds, R., Vonrhein, C., ... & Xu, H. E. (2008). Identification of COUP-TFII orphan nuclear receptor as a retinoic acid–activated receptor. PLoS biology, 6(9), e227.

Alfano, C., Magrinelli, E., Harb, K., & Studer, M. (2014). The nuclear receptors COUP-TF: a long-lasting experience in forebrain assembly. Cellular and molecular life sciences, 71(1), 43-62.

Pereira, F. A., Tsai, M. J., & Tsai*, S. Y. (2000). COUP-TF orphan nuclear receptors in development and differentiation. Cellular and Molecular Life Sciences CMLS, 57, 1388-1398.

Yamazaki, T., Suehiro, J. I., Miyazaki, H., Minami, T., Kodama, T., Miyazono, K., & Watabe, T. (2013). The COUP-TFII variant lacking a DNA-binding domain inhibits the activation of the Cyp7a1 promoter through physical interaction with COUP-TFII. Biochemical Journal, 452(2), 345-357.

Rosa, A., & Brivanlou, A. H. (2011). A regulatory circuitry comprised of miR-302 and the transcription factors OCT4 and NR2F2 regulates human embryonic stem cell differentiation. The EMBO journal, 30(2), 237-248.

Jonk, L. J., de Jonge, M. E., Pals, C. E., Wissink, S., Vervaart, J. M., Schoorlemmer, J., & Kruijer, W. (1994). Cloning and expression during development of three murine members of the COUP family of nuclear orphan receptors. Mechanisms of development, 47(1), 81-97.

Lutz, B., Kuratani, S., Cooney, A. J., Wawersik, S., Tsai, S. Y., Eichele, G., & Tsai, M. J. (1994). Developmental regulation of the orphan receptor COUP-TF II gene in spinal motor neurons. Development, 120(1), 25-36.

Matharu, P. J., & Sweeney, G. E. (1992). Cloning and sequencing of a COUP transcription factor gene expressed in Xenopus embryos. Biochimica et Biophysica Acta (BBA)-Gene Structure and Expression, 1129(3), 331-334

Love, C. E., & Prince, V. E. (2012). Expression and retinoic acid regulation of the zebrafish nr2f orphan nuclear receptor genes. Developmental Dynamics, 241(10), 1603-1615..

Baroukh, N., Ahituv, N., Chang, J., Shoukry, M., Afzal, V., Rubin, E. M., & Pennacchio, L. A. (2005). Comparative genomic analysis reveals a distant liver enhancer upstream of the COUP-TFII gene. Mammalian genome, 16, 91-95.

Qiu, Y., Krishnan, V., Pereira, F. A., Tsai, S. Y., & Tsai, M. J. (1996). Chicken ovalbumin upstream promoter-transcription factors and their regulation. The Journal of steroid biochemistry and molecular biology, 56(1-6), 81-85.

Krishnan, V., Elberg, G., Tsai, M. J., & Tsai, S. Y. (1997). Identification of a novel sonic hedgehog response element in the chicken ovalbumin

upstream promoter-transcription factor II promoter. Molecular Endocrinology, 11(10), 1458-1466.

Petit, F. G., Salas, R., Tsai, M. J., & Tsai, S. Y. (2004). The regulation of COUP-TFII gene expression by Ets-1 is enhanced by the steroid receptor co-activators. Mechanisms of ageing and development, 125(10-11), 719-732.

Pereira, F. A., Qiu, Y., Zhou, G., Tsai, M. J., & Tsai, S. Y. (1999). The orphan nuclear receptor COUP-TFII is required for angiogenesis and heart development. Genes & development, 13(8), 1037-1049.

Polvani, S., Pepe, S., Milani, S., & Galli, A. (2019). COUP-TFII in health and disease. Cells, 9(1), 101.

Wu, S. P., Kao, C. Y., Wang, L., Creighton, C. J., Yang, J., Donti, T. R., ... & Tsai, S. Y. (2015). Increased COUP-TFII expression in adult hearts induces mitochondrial dysfunction resulting in heart failure. Nature communications, 6(1), 8245.

Qin, J., Chen, X., Yu-Lee, L. Y., Tsai, M. J., & Tsai, S. Y. (2010). Nuclear receptor COUP-TFII controls pancreatic islet tumor angiogenesis by regulating vascular endothelial growth factor/vascular endothelial growth factor receptor-2 signaling. Cancer research, 70(21), 8812-8821.

Park, J. I., Tsai, S. Y., & Tsai, M. J. (2003). Molecular mechanism of chicken ovalbumin upstream promoter-transcription factor (COUP-TF) actions. The Keio journal of medicine, 52(3), 174-181.

Ktistaki, E., & Talianidis, I. (1997). Chicken ovalbumin upstream promoter transcription factors act as auxiliary cofactors for hepatocyte nuclear factor 4 and enhance hepatic gene expression. Molecular and cellular biology, 17(5), 2790-2797.

Ko, H. L., Zhuo, Z., & Ren, E. C. (2019). HNF4α combinatorial isoform heterodimers activate distinct gene targets that differ from their corresponding homodimers. Cell Reports, 26(10), 2549-2557.

Leng, X., Cooney, A. J., Tsai, S. Y., & Tsai, M. J. (1996). Molecular mechanisms of COUP-TF-mediated transcriptional repression:

evidence for transrepression and active repression. Molecular and Cellular Biology, 16(5), 2332-2340.

Okamura, M., Kudo, H., Wakabayashi, K. I., Tanaka, T., Nonaka, A., Uchida, A., ... & Sakai, J. (2009). COUP-TFII acts downstream of Wnt/β-catenin signal to silence PPARγ gene expression and repress adipogenesis. Proceedings of the National Academy of Sciences, 106(14), 5819-5824.

Achatz, G., Hölzl, B., Speckmayer, R., Hauser, C., Sandhofer, F., & Paulweber, B. (1997). Functional domains of the human orphan receptor ARP-1/COUP-TFII involved in active repression and transrepression. Molecular and cellular biology, 17(9), 4914-4932.

Le Guével, R., Oger, F., Martinez-Jimenez, C. P., Bizot, M., Gheeraert, C., Firmin, F., ... & Salbert, G. (2017). Inactivation of the nuclear orphan receptor COUP-TFII by small chemicals. ACS chemical biology, 12(3), 654-663.

Wang, L., Cheng, C. M., Qin, J., Xu, M., Kao, C. Y., Shi, J., ... & Tsai, M. J. (2020). Small-molecule inhibitor targeting orphan nuclear receptor COUP-TFII for prostate cancer treatment. Science Advances, 6(18), eaaz8031.

Wang, T., Wang, Z., de Fabritus, L., Tao, J., Saied, E. M., Lee, H. J., ... & Santori, F. R. (2021). 1-deoxysphingolipids bind to COUP-TF to modulate lymphatic and cardiac cell development. Developmental Cell, 56(22), 3128-3145.

Xie, X., Qin, J., Lin, S. H., Tsai, S. Y., & Tsai, M. J. (2011). Nuclear receptor chicken ovalbumin upstream promoter-transcription factor II (COUP-TFII) modulates mesenchymal cell commitment and differentiation. Proceedings of the National Academy of Sciences, 108(36), 14843-14848.

Lin, F. J., You, L. R., Yu, C. T., Hsu, W. H., Tsai, M. J., & Tsai, S. Y. (2012). Endocardial cushion morphogenesis and coronary vessel development require chicken ovalbumin upstream promoter-transcription factor II. Arteriosclerosis, thrombosis, and vascular biology, 32(11), e135-e146.

Lin, F. J., Chen, X., Qin, J., Hong, Y. K., Tsai, M. J., & Tsai, S. Y. (2010). Direct transcriptional regulation of neuropilin-2 by COUP-TFII

modulates multiple steps in murine lymphatic vessel development. The Journal of clinical investigation, 120(5), 1694-1707.

Ceni, E., Mello, T., Polvani, S., Vasseur-Cognet, M., Tarocchi, M., Tempesti, S., ... & Galli, A. (2017). The orphan nuclear receptor COUP-TFII coordinates hypoxia-independent proangiogenic responses in hepatic stellate cells. Journal of hepatology, 66(4), 754-764.

Srinivasan, R. S., Geng, X., Yang, Y., Wang, Y., Mukatira, S., Studer, M., ... & Oliver, G. (2010). The nuclear hormone receptor Coup-TFII is required for the initiation and early maintenance of Prox1 expression in lymphatic endothelial cells. Genes & development, 24(7), 696-707.

Risebro, C. A., Searles, R. G., Melville, A. A., Ehler, E., Jina, N., Shah, S., ... & Riley, P. R. (2009). Prox1 maintains muscle structure and growth in the developing heart.

Aranguren, X. L., Beerens, M., Coppiello, G., Wiese, C., Vandersmissen, I., Lo Nigro, A., ... & Luttun, A. (2013). COUP-TFII orchestrates venous and lymphatic endothelial identity by homo-or hetero-dimerisation with PROX1. Journal of cell science, 126(5), 1164-1175.

Wu, S. P., Cheng, C. M., Lanz, R. B., Wang, T., Respress, J. L., Ather, S., ... & Tsai, S. Y. (2013). Atrial identity is determined by a COUP-TFII regulatory network. Developmental cell, 25(4), 417-426.

Takamoto, N., Kurihara, I., Lee, K., DeMayo, F. J., Tsai, M. J., & Tsai, S. Y. (2005). Haploinsufficiency of chicken ovalbumin upstream promoter transcription factor II in female reproduction. Molecular endocrinology, 19(9), 2299-2308.

Qin, J., Tsai, M. J., & Tsai, S. Y. (2008). Essential roles of COUP-TFII in Leydig cell differentiation and male fertility. PLoS One, 3(9), e3285.

Bashamboo, A., Eozenou, C., Jorgensen, A., Bignon-Topalovic, J., Siffroi, J. P., Hyon, C., ... & McElreavey, K. (2018). Loss of function of the nuclear receptor NR2F2, encoding COUP-TF2, causes testis development and cardiac defects in 46, XX children. The American Journal of Human Genetics, 102(3), 487-493.

Upadia, J., Gonzales, P. R., & Robin, N. H. (2018). Novel de novo pathogenic variant in the NR2F2 gene in a boy with congenital heart

defect and dysmorphic features. American Journal of Medical Genetics Part A, 176(6), 1423-1426.

Carvalheira, G., Malinverni, A. M., Moysés-Oliveira, M., Ueta, R., Cardili, L., Monteagudo, P., ... & Dias-da-Silva, M. R. (2019). The natural history of a man with ovotesticular 46, XX DSD caused by a novel 3-Mb 15q26. 2 deletion containing NR2F2 gene. Journal of the Endocrine Society, 3(11), 2107-2113.

Nostro, M. C., Cheng, X., Keller, G. M., & Gadue, P. (2008). Wnt, activin, and BMP signaling regulate distinct stages in the developmental pathway from embryonic stem cells to blood. Cell stem cell, 2(1), 60-71.

Lin, F. J., Qin, J., Tang, K., Tsai, S. Y., & Tsai, M. J. (2011). Coup d'Etat: an orphan takes control. Endocrine reviews, 32(3), 404-421.

Ishii, S., & Koibuchi, N. (2022). COUP-TFII in Kidneys, from Embryos to Sick Adults. Diagnostics, 12(5), 1181.

Lee, H. J., Kao, C. Y., Lin, S. C., Xu, M., Xie, X., Tsai, S. Y., & Tsai, M. J. (2017). Dysregulation of nuclear receptor COUP-TFII impairs skeletal muscle development. Scientific reports, 7(1), 3136.

Modell, B., & Darlison, M. (2008). Global epidemiology of haemoglobin disorders and derived service indicators. Bulletin of the World Health Organization, 86(6), 480-487.

Piel, F. B., Patil, A. P., Howes, R. E., Nyangiri, O. A., Gething, P. W., Dewi, M., ... & Hay, S. I. (2013). Global epidemiology of sickle haemoglobin in neonates: a contemporary geostatistical model-based map and population estimates. The Lancet, 381(9861), 142-151.

Origa, R. (2017). β-Thalassemia. Genetics in Medicine, 19(6), 609-619.

Eaton, W. A., & Bunn, H. F. (2017). Treating sickle cell disease by targeting HbS polymerization. Blood, The Journal of the American Society of Hematology, 129(20), 2719-2726.

Angelucci, E., Matthes-Martin, S., Baronciani, D., Bernaudin, F., Bonanomi, S., Cappellini, M. D., ... & Peters, C. (2014). Hematopoietic stem cell transplantation in thalassemia major and sickle cell disease:

indications and management recommendations from an international expert panel. haematologica, 99(5), 811.

Forget, B. G. (1998). Molecular basis of hereditary persistence of fetal hemoglobin. Annals of the New York Academy of Sciences, 850(1), 38-44.

Steinberg, M. H. (2020). Fetal hemoglobin in sickle hemoglobinopathies: high HbF genotypes and phenotypes. Journal of Clinical Medicine, 9(11), 3782.

Thein, S. L., Menzel, S., Lathrop, M., & Garner, C. (2009). Control of fetal hemoglobin: new insights emerging from genomics and clinical implications. Human molecular genetics, 18(R2), R216-R223.

Serjeant, G., Serjeant, B., & Mason, K. (1977). Heterocellular hereditary persistence of fetal haemoglobin and homozygous sickle-cell disease. The Lancet, 309(8015), 795-796.

CHAPTER 2:

REGULATION OF THE COUP-TFII GENE EXPRESSION IN HUMAN ERYTHROID CELLS

2.1. Introduction

The transcription factor COUP-TFII was identified in our laboratory as a novel specific activator of fetal γ-globin (Fugazza, 2021). Hence, it is an attractive therapeutic target for the treatment of β-hemoglobinopathies, diseases characterized by an impaired expression of adult β-globin. The continuous production of fetal γ-globin in adult erythroid cells, as it happens in a benign condition called hereditary persistence of fetal hemoglobin (HPFH), has been well described to ameliorate the clinical severity of β-hemoglobinopathies (Forget, 1998; Thein and Craig, 1998; Wilber, 2011; Steinberg, 2020). However, despite extensive studies on COUP-TFII biological functions in different tissues, little is known about mechanisms that drive its expression in erythroid cells and how it is functionally activated.

COUP-TFII is an essential regulator of embryogenesis and organogenesis as demonstrated by the early embryonic lethality (around embryonic day 9.5) of mouse mutants carrying the COUP-TFII homozygous deletion (Qiu, 1997; Pereira, 2000; Lin, 2011). Its expression peaks during embryonic development, whereas it is expressed at a low level in adult tissues. Nonetheless, in the adult, COUP-TFII expression is upregulated in multiple diseases, such as prostate cancer, muscular dystrophy, and heart failure, implicating its role in adult tissues homeostasis (Polvani, 2019). Within the erythroid

lineage, COUP-TFII is expressed in erythroid progenitors of yolk sac origin, most probably in a very narrow timeframe. Its expression drops around embryonic day 13 (E13) of mouse development, together with the expression fetal globin genes (Fugazza, 2021). Such expression pattern suggests a developmental regulation of the COUP-TFII gene, possibly including fetal-specific enhancer elements and/or adult-specific silencers. Indeed, transcriptional enhancers enable precise spatial and temporal control of gene expression and direct the programming of cell identity during development (de Laat, 2013; Rickels and Shilatifard, 2018).

Even though the upstream regulatory network of COUP-TFII remains poorly understood, it has been demonstrated in different cellular models that COUP-TFII can be transcriptionally regulated by retinoic acid (RA) (Jonk, 1994; Qiu, 1996; Soosar, 1996), Sonic hedgehog (Shh) (Krishnan, 1997) and Ets-1 (Petit, 2004). Retinoic acids (all-trans-retinoic acid, ATRA and 9-cis-retinoic acid, 9cRA) upregulated the activity of COUP-TFII promoter in P19 and PCC7 embryocarcinoma cells, when used at concentrations higher than physiological (Jonk, 1994; Qiu, 1996; Soosar, 1996). Moreover, Soosar and colleagues (1996) suggested that the effect of RA on the COUP-TFII promoter is not mediated by the direct binding of retinoic acid receptor (RAR) and/or retinoid X receptor (RXR), but rather indirectly. Shh and Ets-1 were also shown to act on the COUP-TFII promoter activity, in P19 and HeLa cells, respectively. Both Shh response element (ShhRE) and ETS-

binding site were mapped within the COUP-TFII promoter (Krishnan, 1997; Petit, 2004).

In this chapter, I describe the study of upstream regulatory sequences of the COUP-TFII gene.

In particular, I focused my research on erythroid-specific elements controlling COUP-TFII expression: the promoter of COUP-TFII and distal upstream regulatory regions. The understanding of the mechanism by which COUP-TFII is transcriptionally regulated is crucial to design strategies of COUP-TFII expression manipulation, and therefore of reactivation of γ-globin in adult erythroid cells.

First, I analyzed the sequence of the minimal promoter of COUP-TFII and identified two conserved elements that could be important for the regulation of COUP-TFII transcription mediated by retinoic acid.

Second, I selected the genomic regions within the COUP-TFII locus as candidate COUP-TFII regulatory sequences by merging data available in the public databases: the GeneHancer database (Fishilevich, 2017), the Evolutionary Conserved Regions (ECR) Browser (http://ecrbrowser.dcode.org) and the ENCODE ChIP-seq data. The criteria for the selection were the following features: the engagement in looping long-range interactions, as mapped by the GeneHancer track in UCSC Genome Browser, their evolutionary conservation, the presence of the active enhancer H3K27ac epigenetic mark and of transcription factor (TF) motifs bound *in vivo* by TFs known to control erythropoiesis. As a next step, I evaluated the regulatory potential of

the selected regions *in vitro* by cloning them in a reporter construct, upstream to the COUP-TFII minimal promoter. The most promising candidate enhancer elements were further functionally characterized by deletional mapping of TF motifs and CRISPR/Cas9-mediated targeting.

2.2. Materials and methods

Reporter vector construction

All cloned DNA sequences were obtained from UCSC Genome Browser (https://genome.ucsc.edu/, human GRCh37/hg19 assembly) and amplified from genomic DNA of K562 cells by using Q5® High-Fidelity DNA Polymerase (New England BioLabs, M0491S) or Phusion® High-Fidelity DNA Polymerase (New England BioLabs , M0530S) according to standard laboratory protocols. The cloned promoter region corresponds to the region between ApaI and SacII sites on mouse COUP-TFII promoter that was previously described as a COUP-TFII minimal promoter (Soosar, 1996). COUP-TFII minimal promoter (CP) and its shorter version (Cut CP=CCP) were subcloned into pGL3-Basic reporter plasmid (Promega, E1751) on HindIII sites. Candidate COUP-TFII regulatory sequences (-61 kb, -57 kb, -52 kb and -40 kb) were cloned immediately upstream to CP, between NheI and XhoI sites. Deletions of -52 kb-pGL3 construct (ΔSTAT, ΔGATA1, ΔNF-E2, ΔCOUP-TFII) were performed by using a Gibson Assembly® Master Mix (New England BioLabs, E2611SS) according to the manufacturer's protocol. All the primers used for the construction of reporter vectors are listed below.

Table 2.1. ***Primers used for cloning.***

Construct	F/R	Sequence (5'–3')
CP	F	ACGTAAGCTTCCGCTCGCGCTAGGACC
	R	ACGTAAGCTTGAAAAGAACAGAGAATCAAAGTGATC
CCP	F	ACGTAAGCTTTATATAGAGAGCTCAGT
	R	ACGTAAGCTTGAAAAGAACAGAGAATCAAAGTGATC
-61 kb	F	ACGTGCTAGCCTAGAGGCCCCTTCAAAG
	R	ACGTGTCGACCCCAGCAAATTGCACAC
-57 kb	F	ACGTGCTAGCGGCACAAATCGCTCATTAGC
	R	ACGTGTCGACTGCTTAACTCAAGCCTCAGAAAGC
-52 kb	F	ACGTGCTAGCGTCACTGCAGCCTTTACCCC
	R	ACGTGTCGACGGAGGCATGGATTATACTCG
-40 kb	F	ACGTGCTAGCAACACACAGCATCGTGCATG
	R	ACGTGTCGACTATGCATACACTGTTTCCTCATG
ΔSTAT	F	GAGCTCTTACGCGTGCTAGCCATTGACTGCTCTCGTG
	R	CTTACTTAGATCGCAGATCTCGAGGGAGGCATGGATTATAC
ΔGATA1	F	GAGCTCTTACGCGTGCTAGCCACAACTAGAAGATGC
	R	CTTACTTAGATCGCAGATCTCGAGGGAGGCATGGATTATAC
ΔNFE2	F	GAGCTCTTACGCGTGCTAGCGCTGAGCGTCTTGAG
	R	CTTACTTAGATCGCAGATCTCGAGGGAGGCATGGATTATAC
ΔCOUP-TFII	F	GAGCTCTTACGCGTGCTAGCTCATGACACACTCCTG
	R	CTTACTTAGATCGCAGATCTCGAGGGAGGCATGGATTATAC

Cell cultures

K562: K562 is a human immortalized erythroid cell line derived from a female patient with chronic myelogenous leukemia, highly undifferentiated of the granulocytic series (Lozzio and Lozzio, 1975). Cells were cultured in RPMI 1640 (Lonza, BE12-167F) supplemented

with 10% FBS (Sigma-Aldrich, F7524-500ML), 4mM L-Glutamine (Euroclone, ECB3000D) and 100µg/mL Penicillin-Streptomycin (Euroclone, ECB3001D), at an optimal concentration of 0.2 x10^6 cells/mL, at 37°C in 5% CO_2.

K562-Cas9: K562 cell line stably expressing Cas9 was kindly provided by Jürg Schwaller and Samantha Tauchmann (University of Basel, Switzerland). Cells were grown in RPMI (Gibco, 11875093) with 10% FBS (Sigma-Aldrich, F9665-500ML), 2 mM L-Glutamine (Gibco, 25030024) and 100µg/mL Penicillin-Streptomycin (Gibco, 15140122), at 37°C in 5% CO_2 **.**

HepG2: HepG2 is a human liver cancer cell line derived from the liver tissue of a 15-year-old Caucasian male from Argentina with a well-differentiated hepatocellular carcinoma (Aden, 1979). Cells were grown in DMEM (Sigma-Aldrich, D0819-500ML) with 10% FBS (Sigma-Aldrich, F7524-500ML), 4mM L-Glutamine (Euroclone, ECB3000D) and 100µg/mL Penicillin-Streptomycin (Euroclone, ECB3001D), at 37°C in 5% CO_2.

HUDEP-2: human umbilical cord blood-derived erythroid progenitor (HUDEP-2) cell line was maintained in Iscove's modified Dulbecco's medium (IMDM, PAN, P04-20250) supplemented with 0.2% Human Serum Albumin (HAS, PAN, P06-27100), 0.1% Lipid Mixture (Sigma-Aldrich, L0288), 300 µg/mL transferrin (Sigma-Aldrich, T1408), 10 µg/mL insulin (Sigma-Aldrich, I2643), 100 µM Sodium pyruvate (Sigma-Aldrich, P5280), 2 U/mL Erythropoietin (EPO, ImmunoTools, 11344795), 100 ng/mL Stem Cell Factor (SCF, ImmunoTools,

11343325), 1 µM dexamethasone (Sigma-Aldrich, D4902) and 2 µg/mL doxycycline (Sigma-Aldrich, D9891).

LentiX: Lenti-X 293T (Takara Bio, 632180) cells were used for the lentiviral production in CRISPR/Cas9 experiment. Cells were cultured in DMEM (Gibco, 11960044) with 10% FBS (Sigma-Aldrich, F9665-500ML), 4 mM L-Glutamine (Gibco, 25030024) and 100µg/mL Penicillin-Streptomycin (Gibco, 15140122), at 37°C in 5% CO_2.

HEK-293T: HEK-293T cells were used for the production of lentiviral vector overexpressing COUP-TFII. Cells were cultured in DMEM (Sigma-Aldrich, D6429-500ML) with 10% FBS (Sigma-Aldrich, F9665-500ML) and 100µg/mL Penicillin-Streptomycin (Gibco, 15140122), at 37°C in 5% CO_2.

Transfection

K562: 3 x10^5 cells in 0.6mL of Opti-MEM Reduced-Serum Medium (Gibco, 31985-047) were seeded on a 24-well plate and transfected with 0.8µg of reporter plasmid DNA and 2µL of Lipofectamine 2000 Transfection Reagent (Invitrogen, 11668019) per well. 4 h after transfection, 0.6mL of RPMI 1640 with 20% FBS, L-Glutamine and Penicillin-Streptomycin per well were added.

HepG2: 5 x10^5 cells were seeded in 1mL of DMEM with 10% FBS, L-Glutamine and Penicillin-Streptomycin per well, on a 12-well plate. After 24 h, 500µL of Opti-MEM Reduced-Serum Medium (Gibco,

31985-047) with 1µg of reporter plasmid DNA and 2µL of Lipofectamine 2000 Transfection Reagent (Invitrogen, 11668019) were added per well.

All the transfection experiments were done in technical and biological triplicates, using at least two different DNA preparations.

Treatment with retinoic acid

4 h after transfection of reporter plasmid DNA, K562 cells were treated with 5µM solution of retinoic acid (RA) in DMSO, or only DMSO (as a negative control).

Luciferase assay

K562 and HepG2 cells were collected 48 h after transfection and centrifuged (1200 rpm, 5 min). Pellet was resuspended in 80µL of Reporter Lysis Buffer 1X (Promega, E4030), followed by two freeze-thaw cycles. Cell lysates were centrifuged (12000 rcf, 5 min, at 4°C) and the supernatant (containing total protein extract) was collected. Next, 25µL of LAR reagent (Promega, 3040) were added to 10µL of protein extract and the luciferase activity was measured by GloMax® 20/20 Luminometer (Promega), integration time of 10 sec. To calculate the relative luciferase activity, amount of relative luciferase units (RLU) of non-transfected control was subtracted from each sample RLU, and normalized to either CP or -52 kb reference samples.

RNA isolation and cDNA synthesis

TriFast reagent (Euroclone) and the Direct-zol™ RNA Miniprep kit (Zymo Research, R2050), or Monarch® Total RNA Miniprep Kit (New England BioLabs, T2010S) were used to isolate total RNA. Next, RNA was reverse-transcribed into cDNA with the High-Capacity cDNA Reverse Transcription Kit (Applied Biosystems, 4368814).

Real-time quantitative PCR analysis

On the basis of sequences from the Ensembl database (http://www.ensembl.org), primers were designed to amplify 100 to 200 bp amplicons, spanning an exon-exon junction when possible. All the RT-qPCR primers used in this study are listed in the table below. Quantitative real-time PCR reaction was performed using StepOnePlus Real-Time PCR System (Applied Biosystems). Specific PCR product accumulation was monitored by SsoAdvanced Universal SYBR Green Supermix (Bio-Rad, 1725274) in a 12µL reaction volume. Gene expression levels were calculated by normalizing the values to the expression of housekeeping GAPDH gene. Triplicate samples from each experiment were analyzed.

Table 2.2. ***Primers used for RT-qPCR.***

Gene	F/R	Sequence (5'–3')
hGAPDH	F	ACGGATTTGGTCGTATTGGG
	R	TGATTTTGGAGGGATCTCGC
hCOUP-TFII	F	TTGACTCAGCCGAGTACAGC
	R	AAAGCTTTCCGAATCTCGTC
mCOUP-TFII	F	TCCAAGAGCAAGTGGAGAAG
	R	CTTCCAAAGCACACTGGGAC
hCYP26A1	F	ACTTACCTGGGGCTCTACC
	R	GGGATTCAGTCGAAGGGTCT

CRISPR/Cas9-mediated targeting

Cloning of the guide RNAs, lentiviral production and transduction were performed as described (Proietti, 2022). Briefly, single guide RNAs (sgRNAs, Table 2.3) were designed using CRISPR Targets track at the UCSC Genome Browser (https://genome.ucsc.edu/) and cloned into a lentiviral expression vector containing GFP (LentiGuide-Puro-P2A-EGFP (Addgene, 137729) or into a lentiviral expression vector containing iRFP670 (pLenti hU6-sgRNA-IT-PGK-iRFP670). For lentiviral production, Lenti-X cells were transfected with indicated constructs together with psPAX2 and pMD2.G packaging plasmids (Addgene, 12260 and 12259) using polyethyleneimine (PEI, Polysciences Europe, 24765-1). After 24 h, Lenti-X cell medium was substituted with the target cell medium. The supernatant containing lentiviral pseudo-particles was collected 48 h after transfection. Next, approximately 3

x10^6 of K562 cells stably expressing Cas9 (K562-Cas9) were seeded per well of a 6-well plate and transduced by spinfection (1000 rcf, 90 min) with lentiviral supernatants (diluted 1:2) supplemented with Polybrene Transfection Reagent (Merck, TR-1003-G). Cells were collected 6 days after transduction and sorted for GFP and/or iRFP670 expressing cells, followed by RNA and genomic DNA extraction.

Table 2.3. ***sgRNAs used for CRISPR/Cas9-mediated targeting.***

sgRNA	Sequence
1	AGCTCTTTGAAGGCCGGACG
2	TGGAGCTCACTGCTGAGTGC
3	TGCATGCTCTTCCTGTTCTG
COUP-TFII site	TGACTCAGGAGTGTGTCATG

Genomic DNA PCR analysis

Genomic DNA was isolated using a Quick-gDNA Miniprep Kit (Zymo Research, D3025). PCR amplification of the genomic region with expected deletion in the -52 kb element was performed using Q5® High-Fidelity DNA Polymerase (New England BioLabs, M0491S) with the following primers: F (5'-3') CTGCCAGCATTGAATGGAGA and R (5'-3') CTCGCGAAGAACTCACAGGA. The PCR product was analyzed by agarose gel electrophoresis.

COUP-TFII lentiviral overexpression

COUP-TFII murine cDNA (gift from Dr. Michèle Studer) was cloned into a bicistronic IRES-eGFP as described previously (Fugazza, 2021). COUP-TFII overexpressing vector and control Empty vector (EV) were produced in HEK-293T cells, lentiviral vectors were co-transfected together with the packaging plasmids psPAX2 and pMD-VSVG (www.lentiweb.com). Viral vector supernatants were collected 72 h after transfection and concentrated, followed by transduction of K562 cells.

ChIP-seq assay and ChIP-seq data analysis

ChIP-seq experiment was performed as previously (Fugazza, 2021). Briefly, K562 cells were fixed with 1% formaldehyde for 10 minutes, at room temperature. Chromatin was sonicated to a size of about 500 bp. DNA immunoprecipitation was obtained by first incubating with the proper antibodies and then isolating with protein A-agarose beads. DNA was sequenced on an Illumina platform. The sequencing data were uploaded to the Galaxy web platform. BWA was used to map reads to the human genome (GRCh37/hg19), and the HOMER toolkit was used to identify peaks. Peaks in different experiments were called as the same if the minimal overlap of intervals was 1bp.

Statistical analyses

For statistical analyses was used GraphPad Prism 8.4.3 software (GraphPad Software). Data are expressed as mean ±SEM and all experiments were performed in biological triplicates, unless stated otherwise. To calculate the p-value, the paired Student's t-test was used. $P<0.05$ was considered to be statistically significant (*$p<0.05$, **$<p<0.01$, ***$p<0.001$, ****$p<0.0001$, if not stated data not significant).

2.3. Results and discussion

2.3.1. Characterization of the COUP-TFII promoter: the -72 bp direct repeat element mediates its response to retinoic acid

Retinoic acid (RA) was identified as a positive regulator of COUP-TFII transcriptional activity in different cellular models (Jonk, 1994; Qiu, 1996; Soosar, 1996), yet never in erythroid cells. I sought to investigate the effect of RA treatment on the activity of COUP-TFII promoter in K562 cells. K562, a myelogenous leukemia cell line derived from the highly undifferentiated progenitor of the erythrocytic and megakaryocytic lineages, is the only cellular model identified so far, that has erythroid features and expresses COUP-TFII (Lozzio and Lozzio, 1975).

The COUP-TFII 320 bp promoter region, corresponding to the region identified by Soosar and colleagues (1996) as a COUP-TFII minimal promoter (CP), was cloned into a pGL3-Basic reporter vector. Interestingly, the COUP-TFII minimal promoter contains two conserved nuclear receptor (NR) binding motifs (RGGTCA, where R is a purine, Fig. 2.1A): the more distal direct repeat 1 (DR1), located 72 bp upstream to the COUP-TFII transcription start site (TSS)

(GTGTCAAAGTTCA), and the proximal inverted repeat (IR) located 49 bp upstream to the COUP-TFII TSS (AGCTCAGTGAGCTG) (Fig. 2.1B). Although these elements do not contain perfect consensus sequences, it is known that NRs are able to recognize, with variable affinity, a large variety of degenerate motifs, including the ones lying within the COUP-TFII minimal promoter.

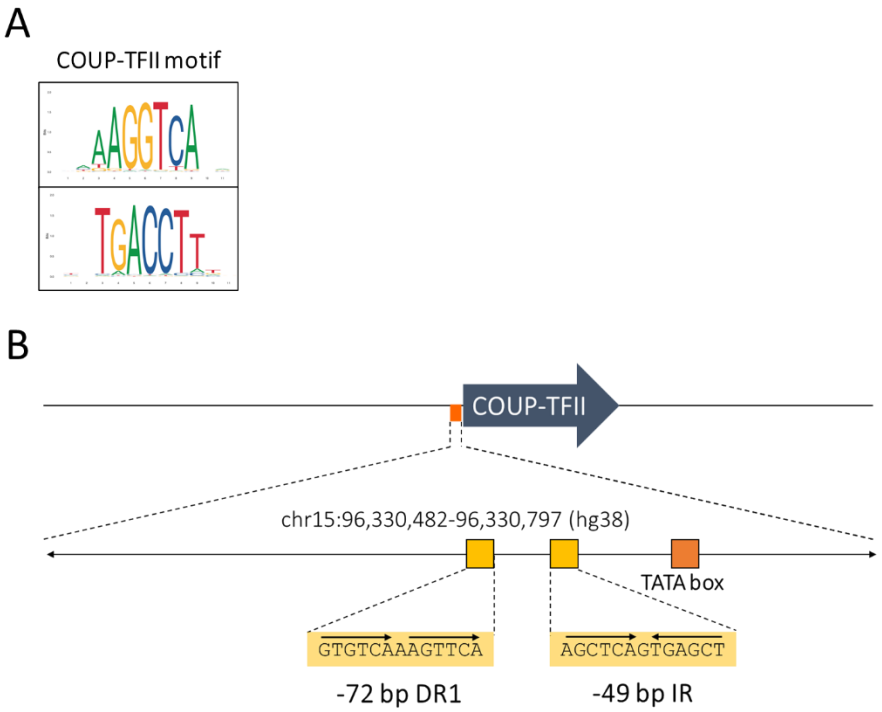


Figure 2.1. ***A direct repeat 1 (DR1) and an inverted repeat (IR) are present in the sequence of the COUP-TFII minimal promoter. A*** *DNA motif recognized by COUP-TFII (top) and its reverse complementary sequence (bottom), the sequence logos obtained from JASPAR.* ***B*** *A schematic overview of the COUP-TFII minimal promoter region. Direct Repeat 1 (DR1) and Inverted Repeat (IR) are indicated in yellow.*

Therefore, to compare the role of these two motifs in RA-mediated response, a shorter version of the promoter (127 bp) was cloned. The shorter COUP-TFII promoter (CCP), containing only the proximal IR site, was also inserted in the pGL3-Basic reporter vector (Fig. 2.2).

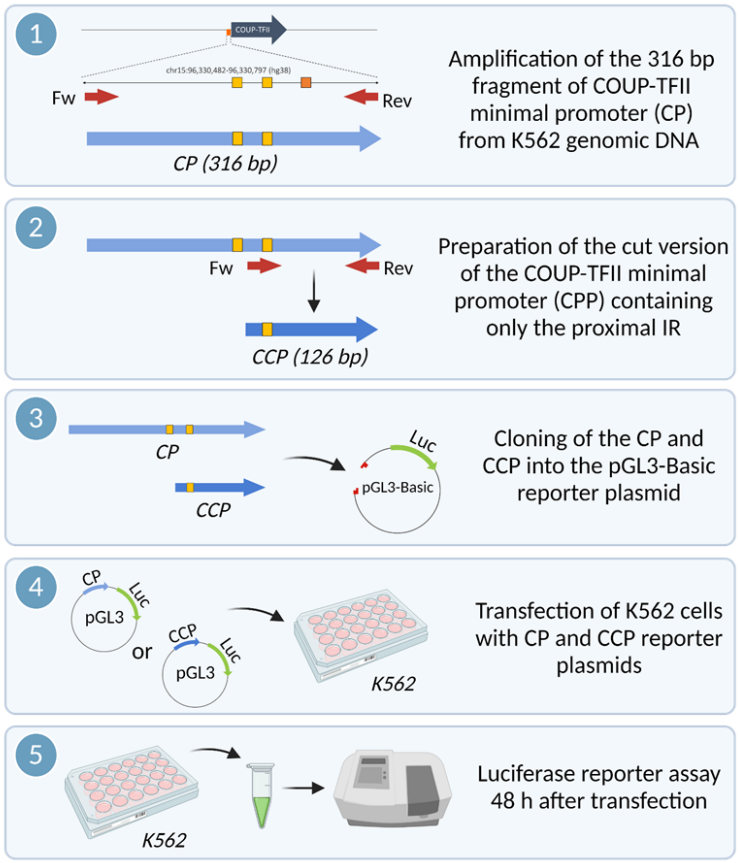


Figure 2.2. ***Experimental workflow of the cloning and analysis of the COUP-TFII minimal promoter (CP) and its cut version (CCP).*** *The functionality of both promoter sequences was confirmed in K562 erythroid cells by using a luciferase reporter assay.*

The transcriptional activity of both promoters was tested by luciferase reporter assay in K562 erythroid cells. As shown in figure 2.3, the COUP-TFII minimal promoter (CP) induced the luciferase expression at the similar level as its shorten version (cut COUP-TFII promoter, CCP). The pGL3 reporter vector containing the functional thymidine kinase promoter (pTK) was used as a control.

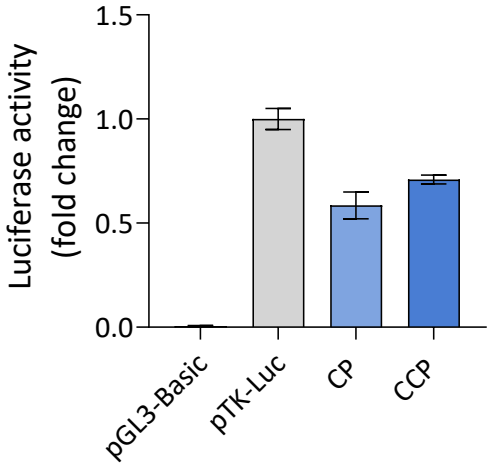


Figure 2.3. ***Functional analysis of the COUP-TFII minimal promoter's reporter constructs.*** *Activity of the COUP-TFII minimal promoter (CP) and its cut version (CCP) in K562 erythroid cells, measured by a luciferase assay, 48 h after transfection. Luciferase activity in Relative Light Units (RLU), RLU values were normalized to RLU of the pTK-Luc control. Data as mean ± SEM, n≥3.*

Next, to evaluate the effect of retinoic acid on the activity of the two versions of the COUP-TFII promoter (CP and CCP), I performed a luciferase assay experiment as shown in figure 2.4. K562 cells were transfected with CP, CCP, or pGL3-Basic plasmids. After 4 h, cells were treated with 5µM RA, or DMSO, and 44 h after transfection cells were analyzed. The pGL3-Basic plasmid was used as a negative control, and DMSO (the vehicle used to dissolve RA) was the control of the drug treatment.

Of note, 5µM is a supraphysiological concentration of retinoic acid, nevertheless, as previous studies demonstrated, the positive regulation of COUP-TFII expression was observed only with high RA concentration.

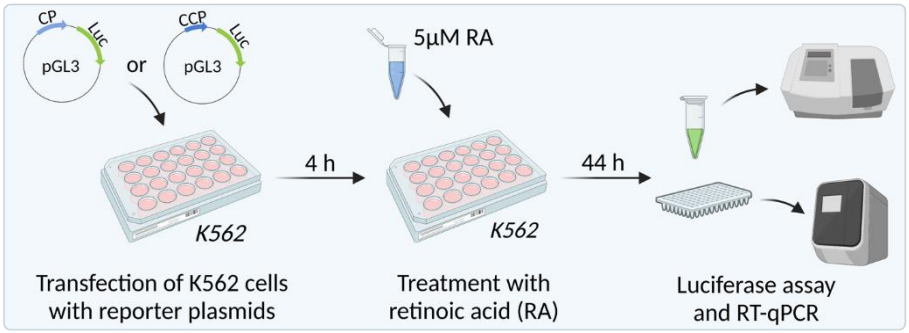


Figure 2.4. ***Experimental workflow to define the effect of DR1 deletion on RA-mediated upregulation of COUP-TFII expression.***

Luciferase assay revealed that retinoic acid significantly upregulated the activity of the COUP-TFII minimal promoter (CP) (2-fold increase) (Fig 2.5A). However, this effect was not observed for the shorter version of the promoter (CCP). The difference in the responsiveness of the COUP-TFII promoter to RA treatment, is the presence or absence of the distal DR1 site, suggesting that the upregulation of COUP-TFII by retinoic acid is mediated by this element.

To assess if the endogenous COUP-TFII expression level increased upon the RA-treatment, the level of COUP-TFII mRNA in K562 cells treated with retinoic acid for 44 h was measured by RT-qPCR. As a positive control for the RA-treatment, I checked the level of CYP26A1 mRNA, a cytochrome P450 gene known to be upregulated by retinoic acid (Sun and Wang, 2021). Although the difference is not significant, the endogenous COUP-TFII level seems to be higher in RA-treated cells when compared to DMSO-treated control cells (Fig. 2.5B).

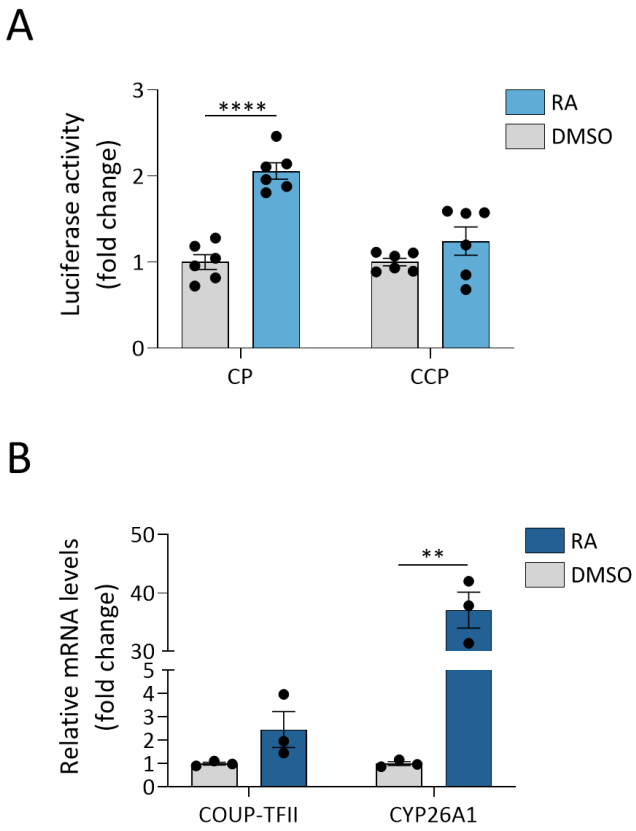


Figure 2.5. ***Retinoic acid (RA) fails to upregulate the activity of COUP-TFII minimal promoter in the absence of the distal direct repeat 1 (DR1) element. A*** *Activity of the COUP-TFII minimal promoter (CP) and its shorten version (CCP) upon the treatment with retinoic acid (RA, 5µM) in K562 erythroid cells, measured by a luciferase assay. Relative Light Unit (RLU) value was normalized to RLU of DMSO control. Data as mean ± SEM, n≥3, ****=p<0.0001.* ***B*** *The level of COUP-TFII mRNA in cells treated for 44 h with retinoic acid analyzed by RT-qPCR, fold change to DMSO control. CYP26A1, known to be activated by RA, was used as a positive control, and the expression level was normalized to hGAPDH. Data as mean ± SEM, n≥1, **=p<0.01.*

To conclude, high concentration of retinoic acid results in about 2-fold increase of the expression of COUP-TFII in K562 erythroid cells. Luciferase assay results suggest that the transcriptional regulation of COUP-TFII by retinoic acid is, at least partially, mediated by the DR1 element located within the COUP-TFII promoter, 72 bp upstream to the transcription start site.

2.3.2. Identification of a long-range enhancer element controlling COUP-TFII expression in erythroid cells

After I identified the DR1 site of the COUP-TFII promoter as an important element for the retinoic acid-mediated activation of the COUP-TFII gene expression, I sought to find COUP-TFII distal regulatory regions, possibly including erythroid enhancers. As a first step in the search for candidate distal enhancers controlling COUP-TFII expression in erythroid cells, I looked at the predictions of high regulatory potential calculated by the GeneHancer tool (Fishilevich, 2017). GeneHancer is a database of human enhancers and their inferred target genes that integrates four different genome-wide databases: the Encyclopedia of DNA Elements (ENCODE), the Ensembl regulatory build, the functional annotation of the mammalian genome (FANTOM) and the VISTA Enhancer Browser. GeneHancer tracks on the UCSC Genome Browser are visualized as loops indicating long-range interactions between predicted enhancers and genes. As shown in figure 2.6, two potential enhancer elements located approximately 52 and 40 kb upstream of the COUP-TFII transcription start site (TSS) are predicted to interact with the COUP-TFII promoter, both with high confidence score (confidence score indicated by an intensity of grey color of the blocks).

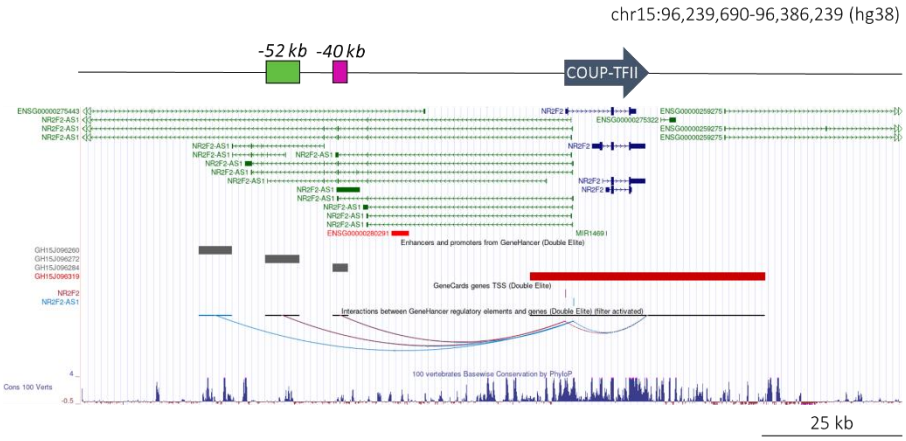


Figure 2.6. ***Identification of the intergenic regions potentially interacting with the COUP-TFII promoter in erythroid cells.*** *Long-range interactions within the COUP-TFII locus were mapped by the GeneHancer database to the reference human genome GRCh38/hg38 in UCSC Genome Browser. Regions located 52 and 40 kb upstream to the COUP-TFII promoter (green and purple boxes) were selected as candidate regulatory elements and further analyzed.*

Regulatory regions are often located within mutational "coldspots", genomic regions with relatively low mutation rate since they are often conserved among species (Harmston, 2013). For this reason, I investigated the evolutionary conservation of the genomic region in the surrounding of COUP-TFII gene. I aligned the genomic region upstream to the COUP-TFII promoter of human genome with genomes of Danio rerio, Takifugu rubripes, Xenopus Laevis, Gallus Gallus and Mus Musculus on the Evolutionary Conserved Regions (ECR) browser (https://ecrbrowser.dcode.org/) (Fig. 2.7). The -40 kb region is

conserved between human and mouse (Mus Musculus), whereas the -52 kb shows only low human-mouse conservation. Two regions located 61 kb and 57 kb upstream to the COUP-TFII transcription start site (TSS) are highly conserved between all selected species. Of note, the -61 kb and -57 kb regions were previously identified and studied as COUP-TFII candidate liver enhancers by Baroukh and colleagues (2005). I therefore decided to include these two regions in further *in vitro* analysis as potential erythroid enhancer elements of COUP-TFII and to use hepatic HepG2 cells as controls of the experiment (same cellular model that was used in the Baroukh's study).

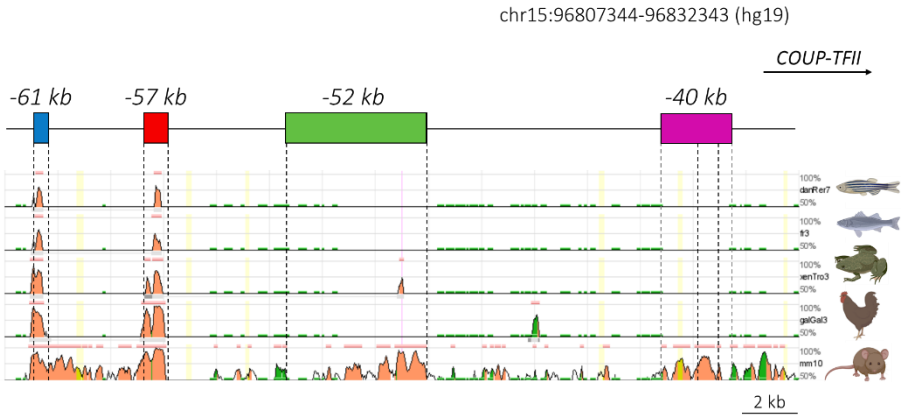


Figure 2.7. ***Schematic conservation of the candidate COUP-TFII regulatory sequences.*** *The conservation of candidate enhancers was analyzed in ECR Evolutionary Conserved Regions browser (https://ecrbrowser.dcode.org). The genome of Homo sapiens was aligned with genomes of (from the top): Danio rerio, Takifugu rubripes, Xenopus laevis, Gallus Gallus and Mus musculus. The analysis was performed using the default parameters set by ECR, minimal length of the peak: 100 bp, minimal identity: 70%, human GRCh37/hg19 assembly.*

Next, I analyzed the H3K27ac histone modification and the opening chromatin state of the selected potential enhancer elements in K562 and HepG2 cells, by using publicly available ENCODE ChIP-seq data. The H3K27ac epigenetic mark (acetylation of the lysine residue at N-terminal position 27 of the histone H3 protein) is a characteristic feature of active enhancers. The assay for transposase-accessible chromatin sequencing (ATAC-seq) provides information about the chromatin accessibility, since active gene regulatory elements are associated with open chromatin state (Klemm, 2019). H3K27ac ChIP-seq revealed an increase of H3K27ac level of the -52 kb element in erythroid K562 cells. Whereas the -61 kb region is H3K27ac-marked in hepatic HepG2 cells. Instead, the -57 and -40 kb regions are not acetylated in none of the analyzed cell lines (Fig. 2.8). Similarly, ATAC-seq showed that the -52 kb and -61 kb are chromatin accessible regions, in K562 and HepG2 cells, respectively (Fig. 2.8).

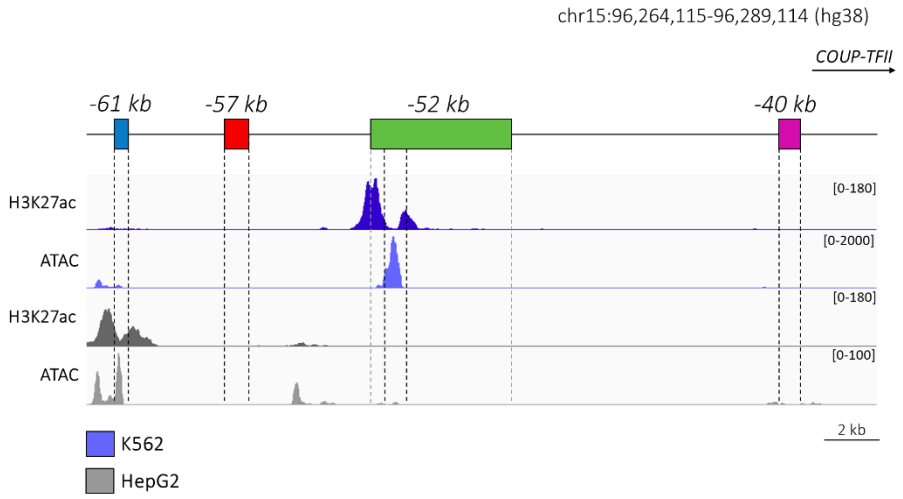


Figure 2.8. ***-52 kb region shows chromatin accessibility and active regulatory element features in K562 cells.*** *H3K27ac ChIP-seq and ATAC-seq profiles in K562 (blue, H3K27ac GEO: ENCFF465GBD, ATAC GEO: ENCFF357GNC) and HepG2 (grey, H3K27ac GEO: ENCFF493VUL, ATAC GEO: ENCFF285FQS) cells at candidate COUP-TFII regulatory elements.*

To conclude, these data suggest that the most promising prospective transcriptional enhancer of COUP-TFII is the region located 52 kb upstream of the COUP-TFII TSS. Even though its nucleotides sequence is not evolutionary conserved, it is a predicted GeneHancer enhancer element, with high level of H3K27ac and open chromatin features in erythroid K562 cells. However, I decided to investigate the regulatory potential of all four candidate sequences, where -61kb and -57kb served as controls of the experiment when analyzed in hepatic HepG2 cells. The summary of regions' characteristics is presented in table 2.4.

Table 2.4. ***A summary of the regions' characteristics selected as candidate COUP-TFII regulatory elements in erythroid cells.***

Distance from COUP-TFII TSS [kb]	Length [bp]	Conservation	Active enhancer H3K27ac mark	Open chromatin
-61	476	Yes	Yes (very low)	Yes (very low)
-57	790	Yes	No	No
-52	742	No	Yes	Yes
-40	670	Yes (Human-Mouse only)	No	No

To assess the ability of the candidate enhancer elements to activate the transcription of COUP-TFII, I cloned these regions in a reporter pGL3-based construct containing the luciferase gene driven by a COUP-TFII minimal promoter (CP, see Fig. 2.1B). For the -61 kb, -57 kb and -40 kb regions, I amplified the genomic regions corresponding to the conserved sequences from K562 genomic DNA, as shown in figure 2.7 by black dashed lines. -52 kb construct consisted of the 742 bp fragment corresponding to the non-repetitive genomic DNA with increased H3K27 acetylation in K562 cells (298 bp) and some flanking repetitive DNA (see Fig. 2.8).

The regulatory potential of four candidate enhancer elements in K562 erythroid cells and HepG2 hepatic control cells was measured by luciferase reporter assay, 48 h after transfection (Fig. 2.9).

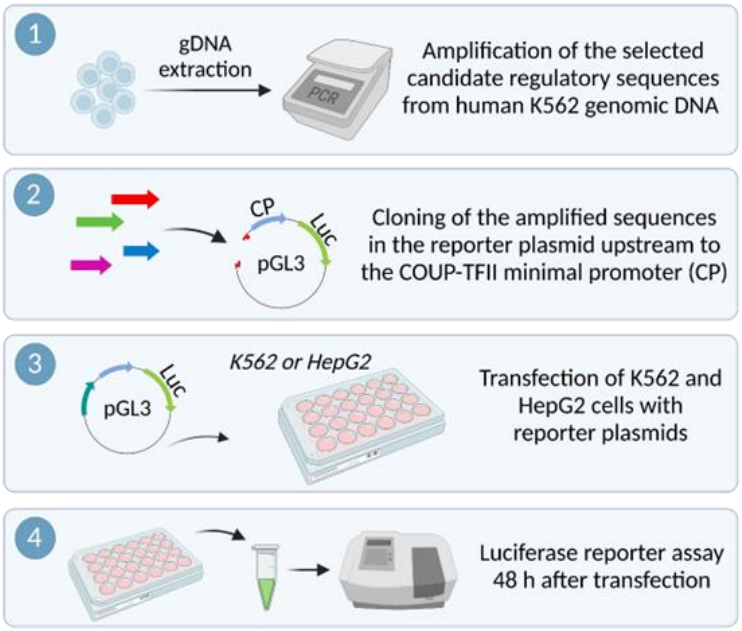


Figure 2.9. ***Experimental workflow for an analysis of the regulatory potential of selected COUP-TFII candidate regulatory elements by using a luciferase reporter assay.***

The -52 kb candidate enhancer element showed a 5-fold increase of the reporter gene expression when compared to CP control (vector containing only COUP-TFII minimal promoter), in both K562 and HepG2 cells. The effect of other three candidate enhancer elements was not as strong as in the case of -52 kb region (Fig. 2.10). In erythroid K562 cells, the -61 kb and -40 kb regions slightly decreased the COUP-TFII promoter activity, whereas -57 kb region did not change it. In hepatic HepG2 cells, -57 kb region showed ~3-fold increase of the reporter gene expression, when compared to CP. Interestingly, in previous studies (Baroukh, 2005) the -57 kb region did not increase the

promoter activity in HepG2 cells, whereas the -61 kb region showed the strongest activation effect (6-fold). However, even if we used the same cellular model, in the work by Baroukh and colleagues, the regions were cloned upstream the SV40 promoter and this could account for the observed difference.

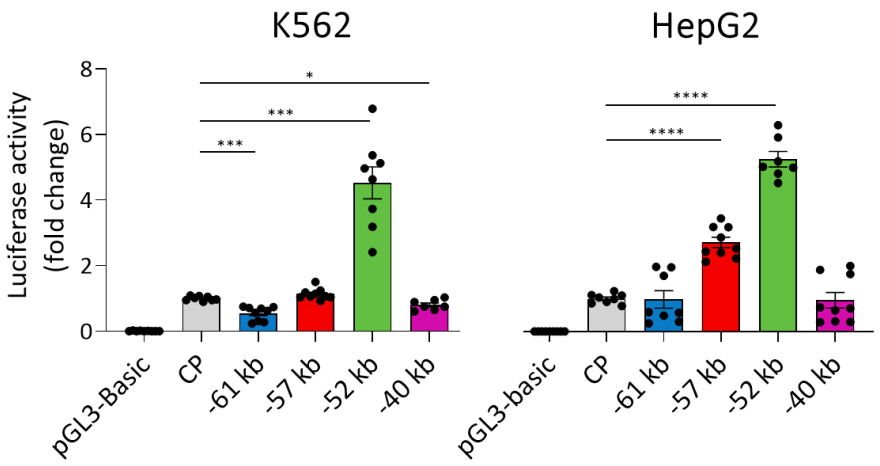


Figure 2.10. ***The -52 kb candidate enhancer element induces the activity of the COUP-TFII promoter.*** *Luciferase assay experiment showed changes in the activity of the COUP-TFII minimal promoter (CP) upon the addition of candidate regulatory sequences (upstream to CP) in K562 erythroid (left) and HepG2 hepatic (right) cells. Data as mean ± SEM, n≥3, *=p<0.05, ***=p<0.001, ****=p<0.0001.*

Taken together, as the preliminary *in silico* analysis suggested and the functional *in vitro* experiment confirmed, the region located 52 kb from the COUP-TFII TSS might be a novel erythroid/hepatic COUP-TFII enhancer element.

To further explore the mechanism by which the -52 kb enhancer element can regulate the expression of COUP-TFII, the analysis of transcription factors (TFs) that were shown to bind *in vivo* to this region was performed by using ENCODE ChIP-seq data. Candidate TFs regulating COUP-TFII expression by binding to the -52 kb enhancer element are listed in the table 2.5, together with the positions of their binding sites (EncodeStart-EncodeEnd). Among TFs that were found to bind to the -52 kb enhancer element are some key erythroid regulators: TAL1, GATA2, NFE2, STAT5A, and COUP-TFII (also called NR2F2) itself.

Table 2.5. ***Transcription factors binding to the -52 kb candidate enhancer element, identified by the ENCODE (aggregate data from different cellular models).***

Chr	EncodeStart	EncodeEnd	EncodeTF
chr15	96273298	96274340	DPF2
chr15	96273378	96274276	CBFA2T3
chr15	96273381	96274262	ARID1B
chr15	96273409	96274173	NCOR1
chr15	96273417	96274189	SMARCE1
chr15	96273443	96274186	HDAC2
chr15	96273445	96274232	SMARCA4
chr15	96273450	96274227	POLR2A
chr15	96273452	96274160	TCF12
chr15	96273458	96274183	KDM1A
chr15	96273465	96274170	TAL1
chr15	96273469	96274144	CBFA2T2
chr15	96273471	96274109	TRIM24
chr15	96273473	96274124	EGR1
chr15	96273477	96274072	LEF1
chr15	96273485	96274181	HDAC1
chr15	96273488	96274126	NR2F1
chr15	96273489	96274067	TBL1XR1
chr15	96273491	96274101	NCOA1
chr15	96273496	96274191	ATF3
chr15	96273513	96274251	DACH1
chr15	96273514	96274145	RNF2
chr15	96273528	96274053	MITF
chr15	96273532	96274198	MTA2
chr15	96273536	96274110	MCM7
chr15	96273537	96274149	MEIS2
chr15	96273540	96274096	SOX6
chr15	96273542	96274110	TEAD4
chr15	96273542	96274123	CTBP1

Chr	EncodeStart	EncodeEnd	EncodeTF
chr15	96273545	96274067	STAT5A
chr15	96273545	96274023	HES1
chr15	96273545	96274054	TCF7
chr15	96273547	96273980	CC2D1A
chr15	96273547	96274035	BCOR
chr15	96273548	96274070	SMARCC2
chr15	96273553	96274134	ARNT
chr15	96273573	96274107	ZEB2
chr15	96273573	96274029	NR2F6
chr15	96273576	96274041	BHLHE40
chr15	96273576	96274087	EP300
chr15	96273584	96273820	ZNF318
chr15	96273586	96274079	PML
chr15	96273588	96274067	NR2F2
chr15	96273590	96274000	MBD2
chr15	96273599	96274010	KLF16
chr15	96273631	96273995	PHB2
chr15	96273631	96274069	TRIM28
chr15	96273639	96274181	JUN
chr15	96273641	96274181	JUNB
chr15	96273653	96274173	GATA2
chr15	96273679	96274019	RCOR1
chr15	96273684	96274119	FOS
chr15	96273689	96274038	HMBOX1
chr15	96273690	96274110	MYC
chr15	96273691	96274217	JUND
chr15	96273723	96274136	NFE2
chr15	96273730	96274081	MAX
chr15	96273736	96274166	CBX3
chr15	96273739	96274063	RUNX1
chr15	96273747	96274164	ZNF184
chr15	96273758	96274288	NBN
chr15	96273789	96273881	FOSL1
chr15	96273807	96274157	GATAD2A

Nuclear factor, erythroid 2 (NFE2) is essential for erythroid maturation and differentiation, as well for hemoglobin synthesis (Andrews, 1994; Mutschler, 2009). Signal transducer and activator of transcription 5 (STAT5) is a key protein for terminal erythropoiesis and erythropoietin signalization (Grebien, 2008; Tothova, 2021). TAL1 (T-cell acute lymphocytic leukemia protein 1, formerly called SCL) is a basic helix-loop-helix hemopoietic transcriptional regulator that binds to a consensus sequence called an E-box (CANNTG) and is required during definitive erythropoiesis (Ravet, 2004; Kassouf, 2010). GATA-binding factor 1 (GATA1) is a master erythroid transcription factor, controlling all aspects of erythropoiesis, at both early and late stages (Shimizu, 2008; Romano, 2020). Even though GATA1 did not emerge among transcription factors directly binding to the -52 kb enhancer element (table 2.5), it is known to form a complex with TAL1. GATA1/TAL1 complexes were found to bind motifs consisting of a TG 7 or 8 bp upstream of a WGATAA motif (Han, 2016). Indeed, the -52 kb enhancer element contains the TGnnnnnnnWGATAA motif. What is more, another GATA factor, GATA2, is present in the list. GATA2 is essential for the early stages of erythropoiesis and is known to bind the same elements as GATA1 at different stages of erythroid differentiation (Tsai and Orkin, 1997; Romano, 2020) Further analysis of the ENCODE ChIP-seq profiles in K562 cells identified peaks of STAT5A, GATA1, TAL1 and NFE2, and the ChIP-seq data obtained in our laboratory in K562 cells showed also a COUP-TFII peak at the -52 kb enhancer (Fig. 2.11).

Although it activates transcription in luciferase assay in HepG2 hepatic cells (Fig. 2.10), the -52 kb element is located within in inaccessible

chromatin region (Fig. 2.8), in contrast to K562 erythroid cells. It is therefore possible that the -52 kb element in transfection experiments is bound in HepG2 and K562 cells by different transcription factors, for example different GATA proteins. In hepatocytes, GATA4 and GATA6 have been implicated in gene regulation (Zhao, 2005; Watt, 2007; Divine, 2004). GATA4 has previously been shown to act as an activator of liver-related genes such as coagulation factor X and hepcidin in hepatoma cell lines (Hung, 2011; Island, 2011) as well as a repressor in mouse hepatocytes (Zheng, 2013). Whether one of these two TFs induced the activity of the COUP-TFII promoter by binding the -52 kb enhancer element remains to be elucidated.

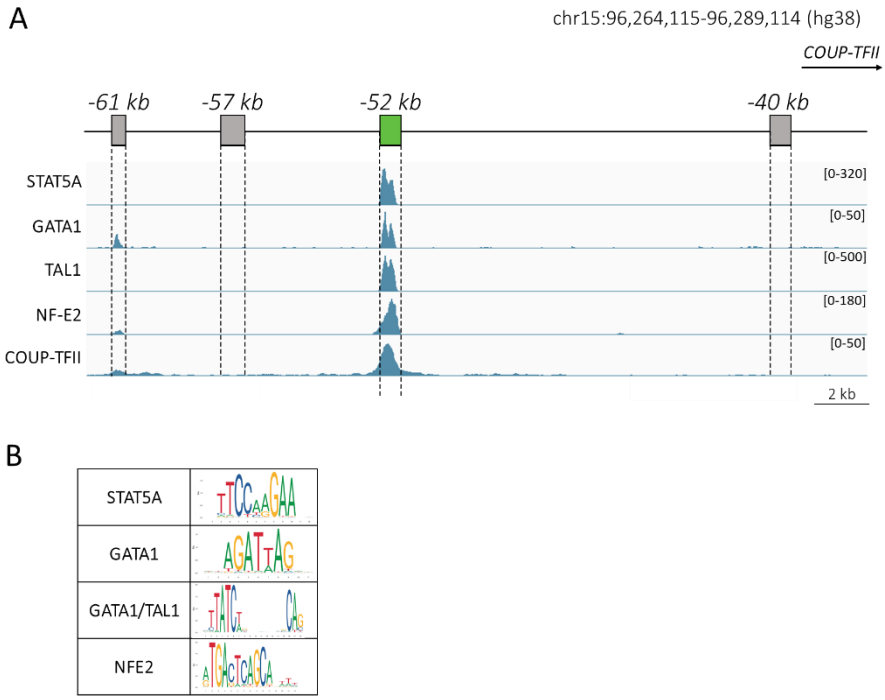


Figure 2.11. ***Key erythroid transcription factors (TFs) bind at the -52 kb regulatory element.*** ***A*** *ENCODE ChIP–seq data of (from the top): STAT5A (GEO: ENCFF253XZP), GATA1 (GEO: ENCFF331URE), TAL1 (GEO: ENCFF700NBW) and NFE2 (GEO: ENCFF398LBP) at the −52 kb COUP-TFII enhancer in K562 cells. COUP-TFII ChIP-seq experiment was performed in our laboratory.* ***B*** *JASPAR DNA motifs of the selected erythroid TFs binding the -52 kb element.*

Next, I sought to investigate the effect of the removal of STAT (ΔSTAT), GATA1 (ΔGATA1), NFE2 (ΔNFE2) and COUP-TFII (ΔCOUP-TFII) motifs on the regulatory potential of the -52 kb enhancer element. 5' repetitive sequence of the region was deleted together with a subsequent deletion of the indicated TF motifs (Fig. 2.12A). To compare the effect

of shorter versions of the -52 kb enhancer element with its full version, a luciferase assay was used. The experiment was performed in K562 erythroid cells, as previously described (Fig. 2.9). As shown in the figure 2.12B, the removal of the STAT motif led to a significant (by 60%) reduction in enhancer activity of the -52 kb enhancer element.

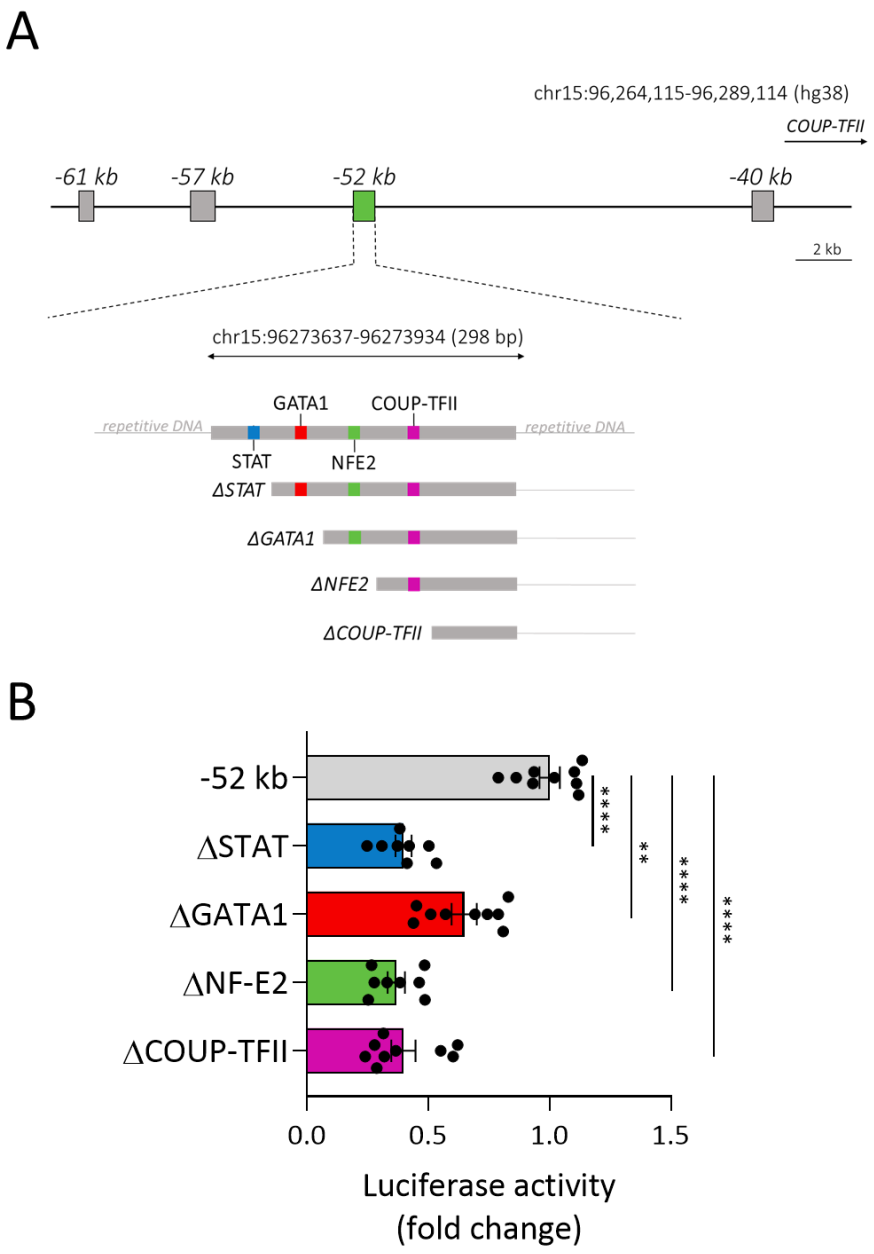

A
chr15:96,264,115-96,289,114 (hg38)
COUP-TFII
-61 kb
-57 kb
-52 kb
-40 kb
2 kb
chr15:96273637-96273934 (298 bp)
GATA1
COUP-TFII
repetitive DNA
repetitive DNA
STAT
NFE2
ΔSTAT
ΔGATA1
ΔNFE2
ΔCOUP-TFII
B
-52 kb
ΔSTAT
ΔGATA1
ΔNF-E2
ΔCOUP-TFII
****
**
****
****
0.0
0.5
1.0
1.5
Luciferase activity
(fold change)

Figure 2.12. ***Progressive deletions of the erythroid transcription factor (TF) motifs within the -52 kb enhancer result in reduced activity of this element. A*** *Schematic overview of TF binding sites within the non-repetitive DNA sequence of the -52 kb element. Shorter versions of the -52 kb reporter construct: ΔSTAT, ΔGATA1, ΔNFE2 and ΔCOUP-TFII were created as indicated.* ***B*** *Effects of TF motif deletions on the regulatory potential of the -52 kb enhancer element were analyzed in K562 erythroid cells by luciferase reporter assay (as shown in Fig. 2.9). Data as mean ± SEM, n≥3, **=p<0.1, ***=p<0.001, ****=p<0.0001.*

The progressive deletion of the -52 kb enhancer element showed that this STAT site might be significant for the enhancer activity of this element. STAT5 is a key protein of erythropoietin (EPO) signalization and a key regulator of erythroid terminal maturation and regulators of the canonical EPOR/JAK2/STAT5 signaling pathway have been widely studied (Grebien, 2008; Tothova, 2021; Fasouli and Katsantoni, 2021). Thus, it would be interesting to further investigate how modulation of STAT5 affects the COUP-TFII expression levels in erythroid cells. Although, in the study of Gillinder and colleagues (2017), COUP-TFII did not emerge as one of the direct target genes of STAT5 signaling in erythroid cells, either EPO-dependent or EPO-independent, they used a different cellular model (J2E cells). Interestingly, in the same study STAT5 has been shown to co-occupy genomic regions linked to erythroid-related genes together with GATA1, KLF1, and/or TAL1 transcription factors that is in line in co-possible co-binding of GATA1/TAL1 and STAT5 to the -52 kb COUP-TFII

enhancer element (Gillinder, 2017). Taken together, STAT5 is an attractive target in attempt to COUP-TFII gene expression modulation, however the direct regulation of COUP-TFII gene by STAT5 needs to be further confirmed.

To determine whether the deletion of the -52 kb COUP-TFII enhancer element impacts on the endogenous COUP-TFII transcription, in collaboration with L. Proietti and F. Grebien (Vetmeduni, Vienna, Austria), I performed a CRISPR/Cas9-mediated targeting in K562 erythroid cells. To delete the non-repetitive "core" sequence of the -52 kb enhancer, I transduced K562 cells stably expressing Cas9 with a pair of single guide RNAs (sgRNAs) targeting upstream and downstream of the non-repetitive DNA sequence (Fig. 2.13).

The guide RNAs were cloned into a lentiviral expression vector containing either GFP or iRFP670 (as indicated by green or blue colors in the figure 2.13). The sgRNA combination 1+2 deleted a 315 bp-fragment, and the combination 1+3 deleted a 335 bp-fragment. In addition, I designed one sgRNA to disrupt the COUP-TFII binding site within the -52 kb enhancer element (Fig. 2.13).

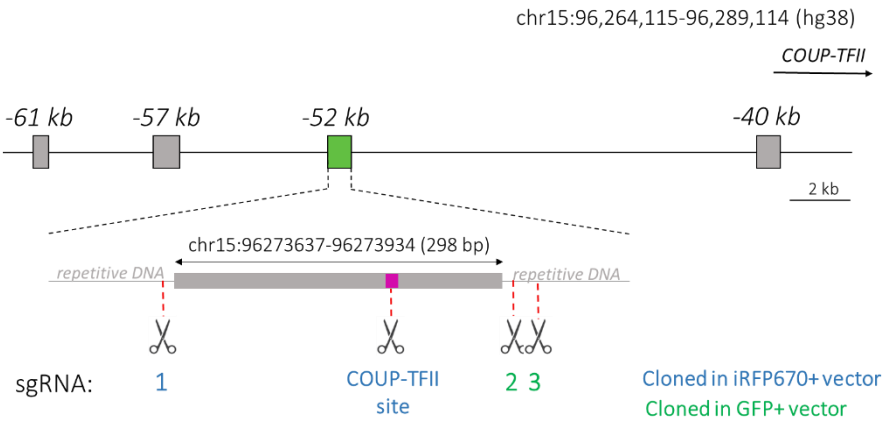


Figure 2.13. ***CRISPR/Cas9-mediated approach to target the -52 kb COUP-TFII enhancer element in K562-Cas9 cells.*** *Recognition sites of sgRNAs within the COUP-TFII enhancer region are indicated by dashed lines. Single guide RNAs were cloned in the lentiviral vector expressing iRFP670 or GFP, in blue or in green, respectively.*

K562-Cas9 cells were single- or double-transduced and 6 days post infection were sorted for iRFP670^{+} cells, or for double-positive (iRFP670^{+} GFP^{+}) cells, prior to genomic and RT-qPCR analyses (Fig. 2.14). The non-targeting sgAAVS1 was used as a control for transduction. Deletion was confirmed by PCR as shown in figure 2.15A and the shorter bands (with the deletion) were further cloned in the pGEM-T vector and sequenced.

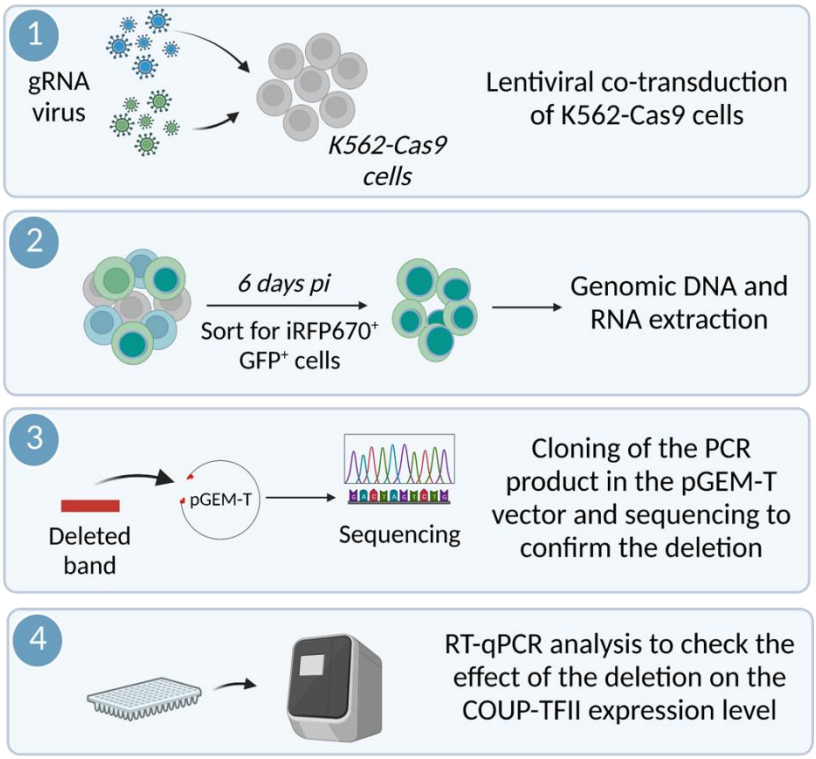


Figure 2.14. ***Schematic representation of the CRISPR/Cas9-mediated deletion of the -52 kb COUP-TFII enhancer element in K562-Cas9 cells.*** *The enhancer element was targeted using 2 sgRNAs: one in a lentiviral vector expressing iRFP670, another in a vector expressing GFP. The COUP-TFII binding site was targeted with one sgRNA (in a vector containing iRFP670) in Cas9-expressing K562 clone. Six days post infection (pi), iRFP67^{+}GFP^{+} or iRFP670^{+} cells were sorted, followed by RNA and genomic DNA analyses.*

The deletion of both 315 bp- and 335 bp-sequences of the -52 kb enhancer element led to 80% reduction of the COUP-TFII mRNA level (Fig. 2.15B). COUP-TFII mRNA level was also significantly reduced upon the disruption of COUP-TFII binding site (40% reduction) when compared to control cells (Fig.2.15B). This result clearly indicates a crucial role of the -52 kb enhancer in the transcriptional regulation of the COUP-TFII, as well as it suggests an autoregulatory mechanism as an important part of the COUP-TFII transcriptional regulation. Further analysis is ongoing in order to further characterize the -52 kb enhancer element.

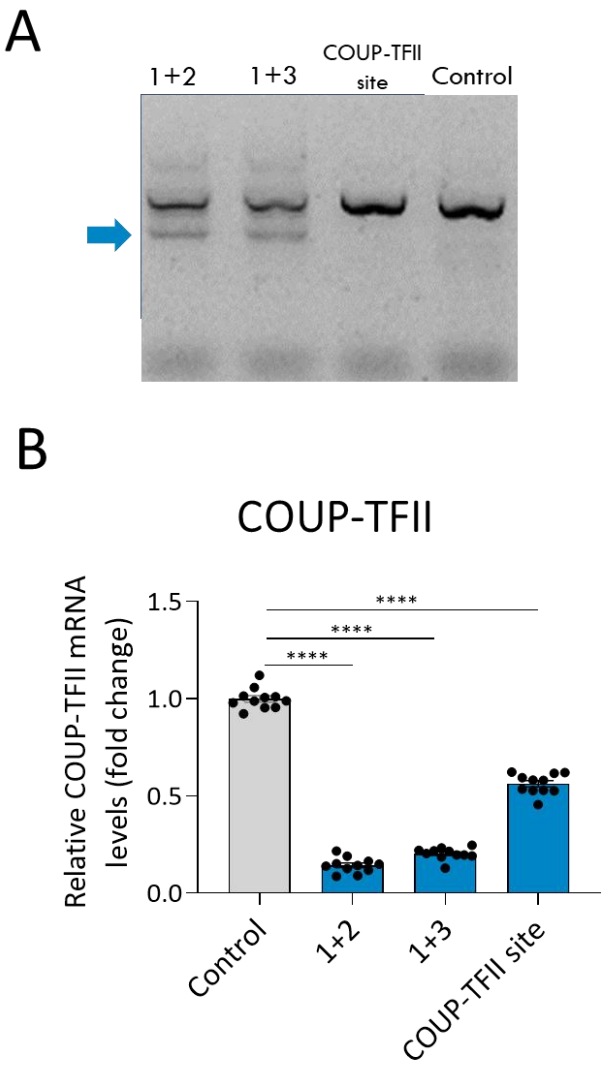


Figure 2.15. ***CRISPR/Cas9-mediated targeting of the -52 kb COUP-TFII enhancer element in K562-Cas9 cells leads to reduced COUP-TFII mRNA level. A*** *PCR analysis of the genomic DNA confirms the deletion of 315 and 335 bp fragments of the -52 kb COUP-TFII enhancer element.* ***B*** *RT-qPCR analysis of the level of COUP-TFII mRNA 6 days post infection, normalized to non-targeting sgRNA control (sgAAVS1). Data as mean ± SEM, n≥2, ****=p<0.0001.*

2.3.3. An autoregulatory mechanism controls the erythroid COUP-TFII gene expression

To assess whether an autoregulatory mechanism is involved in the transcriptional regulation of the COUP-TFII erythroid expression, K562 erythroid cells were transduced with the mouse COUP-TFII overexpressing lentiviral vector (COUP-TFII OE), or with the corresponding empty vector (EV) as a control. 72 h after transduction the level of endogenous COUP-TFII was analyzed by RT-qPCR. As shown in figures 2.16A and 2.16B, the endogenous COUP-TFII mRNA level was significantly induced (1.5-fold) upon the exogenous expression of COUP-TFII. K562 are fetal-like erythroid cells spontaneously expressing COUP-TFII, therefore I wondered if the overexpression of COUP-TFII could re-activate the expression of endogenous COUP-TFII also in adult-like erythroid cells, such as HUDEP-2 cells which express adult globin genes and not expressing COUP-TFII (Kurita, 2013). The relative expression levels of endogenous COUP-TFII in HUDEP-2 cells infected with the same COUP-TFII overexpressing vector as used in K562 cells (COUP-TFII OE) and with control EV cells remained close to 0 (Fig. 2.16A). However, even if COUP-TFII levels in both COUP-TFII OE and EV control HUDEP-2 cells were marginal, the overexpression of COUP-TFII led to a 2-fold increase (not significant) when compared to the control (Fig. 2.16B).

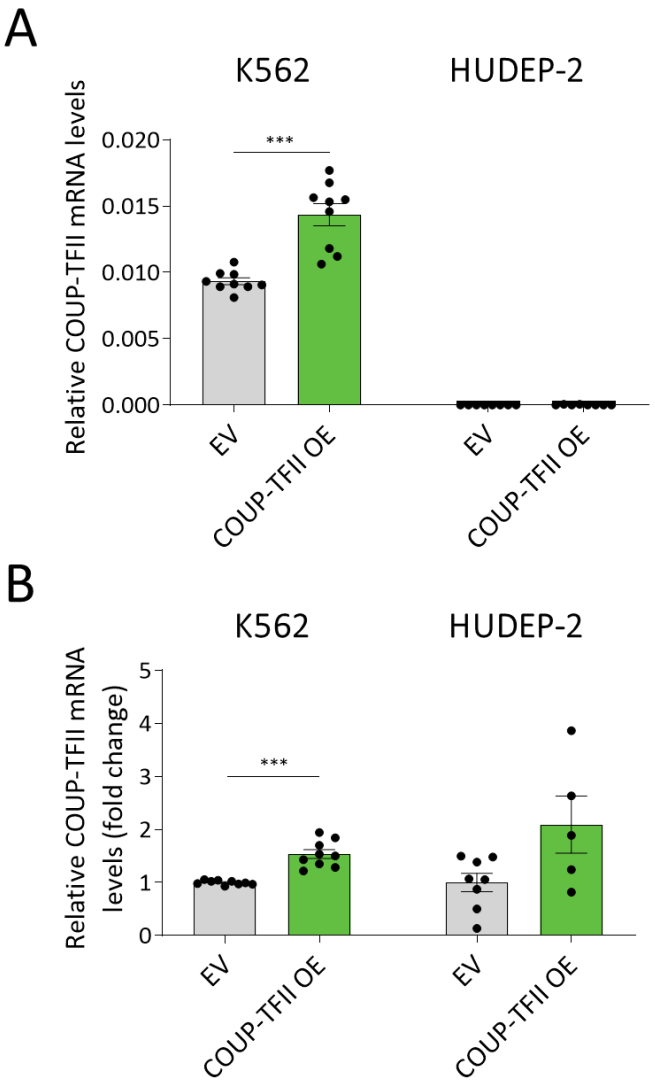


Figure 2.16. ***Exogenous expression of COUP-TFII led to increased levels of endogenous COUP-TFII mRNA in fetal-like K562 cells, but did not induced its expression in adult-like erythroid HUDEP-2 cells.*** *Expression levels of endogenous COUP-TFII after COUP-TFII overexpression (COUP-TFII OE, 72 h after transduction) in K562 and HUDEP-2 cells, relative expression to GAPDH* ***(A)*** *and its fold change* ***(B)****. RT-qPCR analysis, values normalized to empty vector (EV) control. Data as mean ± SEM, n≥3, ***=p<0.001.*

Thus, COUP-TFII upregulates its own expression in fetal-like erythroid cells already expressing COUP-TFII, although its exogenous overexpression is not sufficient to reactivate the expression of endogenous COUP-TFII in adult environment. In agreement with this interpretation, COUP-TFII CUT&RUN experiment in HUDEP-2 cells overexpressing COUP-TFII showed no peaks within the -52 kb enhancer. ATAC-seq and H3K27ac profiles in HUDEP-2 cells confirmed that this region is inactive in adult erythroid cells. COUP-TFII was not reported to act as a pioneer factor, therefore is not able to initiate transcription events in closed chromatin regions. Overall, this implicates that the identified COUP-TFII enhancer is specific to fetal erythroid cells, where COUP-TFII is physiologically expressed.

Interestingly, COUP-TFII profile in HepG2 cells (Fig. 2.17) shows its binding to the -61 kb element that was previously shown to function as COUP-TFII hepatic enhancer *in vivo* after a tail vein injection in adult mice (Baroukh, 2005). It is therefore possible that COUP-TFII positively regulates its own expression also in hepatic cells.

Taken together, I described here a novel enhancer element that is important for the regulation of COUP-TFII expression in fetal-like K562 erythroid cells. This highly acetylated, human specific enhancer is bound by key erythroid transcriptional regulators including GATA1, TAL1, NFE2, STAT5A and COUP-TFII itself. Hence, the transcriptional regulation of COUP-TFII in erythroid cells involves an autoregulatory mechanism. Whether the -52 kb COUP-TFII enhancer element is specific to erythroid lineage, and which transcription factors are

essential to drive the COUP-TFII expression through this element, remains unclear.

In hepatocytes, another cell type expressing COUP-TFII, the -52 kb region is not active as demonstrated by the ATAC-seq and H3K27ac ChIP-seq ENCODE data (Fig. 2.8). Instead the -61 kb region that has previously been identified by the liver COUP-TFII enhancer (Baroukh, 2005) shows high acetylation and open chromatin features. Interestingly, GATA4, a known activator of liver-specific genes (Hung, 2011; Island, 2011), binds *in vivo* within this region as demonstrated by the ENCODE ChIP-seq data (Fig. 2.17).

COUP-TFII is also expressed in endothelial cells where it plays a key role in specification of endothelial cell (EC) fate during vascular development (Aranguren, 2013). What is more, endothelial and erythroid progenitor cells are believed to have the same origin deriving from hemogenic endothelial cells (HECs) in mid-gestational embryos (Dzierzak and Bigas, 2018). Therefore, it would be interesting to see if in endothelial cells the -52 kb element can act as an active enhancer region. H3K27ac histone marks show the same pattern (high acetylation) in primary Human Umbilical Vein Endothelial Cells (HUVEC) as in K562 erythroid cells. Among GATA family members, GATA2 is an important regulator of endothelial expression program, controlling the expression of for example platelet/endothelial cell adhesion molecule-1 (PECAM-1) (Gumina, 1997). Indeed, GATA2 emerged in the list of predicted transcription factors binding to the -52 kb enhancer element (Table 2.5) and the ENCODE ChIP-seq data

from HUVEC cells revealed a GATA2 peak of binding mapped within this region (Fig. 2.17).

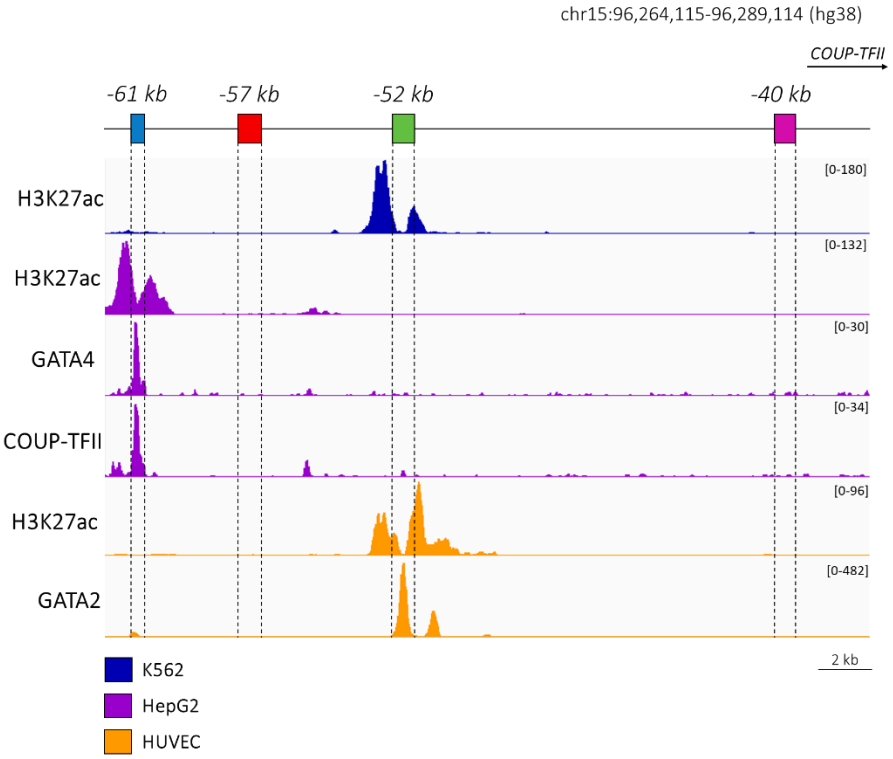


Figure 2.17. ***The -52 kb enhancer element is active in erythroid and endothelial cells, whereas the -61 kb element is active in hepatic tissue.*** *ENCODE ChIP–seq data of (from the top): H3K27ac in K562 (GEO: ENCFF465GBD), H3K27ac in HepG2 (GEO: ENCFF493VUL) , GATA4 in HepG2 (GEO: ENCFF206MPL), COUP-TFII in HepG2 (GEO: ENCFF037OTB), H3K27ac in HUVEC (GEO: ENCSR000ALB) and GATA2 in HUVEC (GEO: ENCFF375GUY) cells.*

2.4. References

Fugazza, C., Barbarani, G., Elangovan, S., Marini, M. G., Giolitto, S., Font-Monclus, I., ... & Ronchi, A. E. (2021). The Coup-TFII orphan nuclear receptor is an activator of the γ-globin gene. haematologica, 106(2), 474.

Forget, B. G. (1998). Molecular basis of hereditary persistence of fetal hemoglobin. Annals of the New York Academy of Sciences, 850(1), 38-44.

Thein, S. L., & Craig, I. E. (1998). Genetics of Hb F/F cell variance in adults and heterocellular hereditary persistence of fetal hemoglobin. Hemoglobin, 22(5-6), 401-414.

Wilber, A., Nienhuis, A. W., & Persons, D. A. (2011). Transcriptional regulation of fetal to adult hemoglobin switching: new therapeutic opportunities. Blood, The Journal of the American Society of Hematology, 117(15), 3945-3953.

Steinberg, M. H. (2020). Fetal hemoglobin in sickle hemoglobinopathies: high HbF genotypes and phenotypes. Journal of Clinical Medicine, 9(11), 3782.

Qiu, Y., Pereira, F. A., DeMayo, F. J., Lydon, J. P., Tsai, S. Y., & Tsai, M. J. (1997). Null mutation of mCOUP-TFI results in defects in morphogenesis of the glossopharyngeal ganglion, axonal projection, and arborization. Genes & development, 11(15), 1925-1937.

Pereira, F. A., Tsai, M. J., & Tsai*, S. Y. (2000). COUP-TF orphan nuclear receptors in development and differentiation. Cellular and Molecular Life Sciences CMLS, 57, 1388-1398.

Lin, F. J., Qin, J., Tang, K., Tsai, S. Y., & Tsai, M. J. (2011). Coup d'Etat: an orphan takes control. Endocrine reviews, 32(3), 404-421.

Polvani, S., Pepe, S., Milani, S., & Galli, A. (2019). COUP-TFII in health and disease. Cells, 9(1), 101.

De Laat, W., & Duboule, D. (2013). Topology of mammalian developmental enhancers and their regulatory landscapes. Nature, 502(7472), 499-506.

Rickels, R., & Shilatifard, A. (2018). Enhancer logic and mechanics in development and disease. Trends in cell biology, 28(8), 608-630.

Jonk, L. J., de Jonge, M. E., Vervaart, J. M., Wissink, S., & Kruijer, W. (1994). Isolation and developmental expression of retinoic-acid-induced genes. Developmental biology, 161(2), 604-614.

Qiu, Y., Krishnan, V., Pereira, F. A., Tsai, S. Y., & Tsai, M. J. (1996). Chicken ovalbumin upstream promoter-transcription factors and their regulation. The Journal of steroid biochemistry and molecular biology, 56(1-6), 81-85.

Soosaar, A., Neuman, K., Nornes, H. O., & Neuman, T. (1996). Cell type specific regulation of COUP-TF II promoter activity. FEBS letters, 391(1-2), 95-100.

Krishnan, V., Elberg, G., Tsai, M. J., & Tsai, S. Y. (1997). Identification of a novel sonic hedgehog response element in the chicken ovalbumin upstream promoter-transcription factor II promoter. Molecular Endocrinology, 11(10), 1458-1466.

Petit, F. G., Salas, R., Tsai, M. J., & Tsai, S. Y. (2004). The regulation of COUP-TFII gene expression by Ets-1 is enhanced by the steroid receptor co-activators. Mechanisms of ageing and development, 125(10-11), 719-732.

Fishilevich, S., Nudel, R., Rappaport, N., Hadar, R., Plaschkes, I., Iny Stein, T., ... & Cohen, D. (2017). GeneHancer: genome-wide integration of enhancers and target genes in GeneCards. Database, 2017.

Lozzio, C. B., & Lozzio, B. B. (1975). Human chronic myelogenous leukemia cell-line with positive Philadelphia chromosome.

Aden, D. P., Fogel, A., Plotkin, S., Damjanov, I., & Knowles, B. B. (1979). Controlled synthesis of HBsAg in a differentiated human liver carcinoma-derived cell line. Nature, 282(5739), 615-616.

Proietti, L., Manhart, G., Heyes, E., Troester, S., & Grebien, F. (2022). Arrayed CRISPR/Cas9 Screening for the Functional Validation of Cancer Genetic Dependencies. Bio-protocol, 12(24).

Sun, D., & Wang, S. (2021). Sphingosine kinases are involved in the regulation of all-trans retinoic acid sensitivity of K562 chronic myeloid leukemia cells. Oncology Letters, 22(2), 1-9.

Harmston, N., Barešić, A., & Lenhard, B. (2013). The mystery of extreme non-coding conservation. Philosophical Transactions of the Royal Society B: Biological Sciences, 368(1632), 20130021.

Klemm, S. L., Shipony, Z., & Greenleaf, W. J. (2019). Chromatin accessibility and the regulatory epigenome. Nature Reviews Genetics, 20(4), 207-220.

Baroukh, N., Ahituv, N., Chang, J., Shoukry, M., Afzal, V., Rubin, E. M., & Pennacchio, L. A. (2005). Comparative genomic analysis reveals a distant liver enhancer upstream of the COUP-TFII gene. Mammalian genome, 16, 91-95.

Andrews, N. C. (1994). Erythroid transcription factor NF-E2 coordinates hemoglobin synthesis. Pediatric research, 36(4), 419-423.

Mutschler, M., Magin, A. S., Buerge, M., Roelz, R., Schanne, D. H., Will, B., ... & Pahl, H. L. (2009). NF-E2 overexpression delays erythroid maturation and increases erythrocyte production. British journal of haematology, 146(2), 203-217.

Tóthová, Z., Tomc, J., Debeljak, N., & Solár, P. (2021). STAT5 as a key protein of erythropoietin signalization. International journal of molecular sciences, 22(13), 7109.

Grebien, F., Kerenyi, M. A., Kovacic, B., Kolbe, T., Becker, V., Dolznig, H., ... & Moriggl, R. (2008). Stat5 activation enables erythropoiesis in the absence of EpoR and Jak2. Blood, The Journal of the American Society of Hematology, 111(9), 4511-4522.

Ravet, E., Reynaud, D., Titeux, M., Izac, B., Fichelson, S., Roméo, P. H., ... & Pflumio, F. (2004). Characterization of DNA-binding-dependent and-independent functions of SCL/TAL1 during human erythropoiesis. Blood, 103(9), 3326-3335.

Kassouf, M. T., Hughes, J. R., Taylor, S., McGowan, S. J., Soneji, S., Green, A. L., ... & Porcher, C. (2010). Genome-wide identification of TAL1's functional targets: insights into its mechanisms of action in primary erythroid cells. Genome research, 20(8), 1064-1083.

Romano, O., Petiti, L., Felix, T., Meneghini, V., Portafax, M., Antoniani, C., ... & Miccio, A. (2020). GATA factor-mediated gene regulation in human erythropoiesis. Iscience, 23(4), 101018.

Shimizu, R., Engel, J. D., & Yamamoto, M. (2008). GATA1-related leukaemias. Nature Reviews Cancer, 8(4), 279-287.

Han, G. C., Vinayachandran, V., Bataille, A. R., Park, B., Chan-Salis, K. Y., Keller, C. A., ... & Pugh, B. F. (2016). Genome-wide organization of GATA1 and TAL1 determined at high resolution. Molecular and cellular biology, 36(1), 157-172.

Tsai, F. Y., & Orkin, S. H. (1997). Transcription factor GATA-2 is required for proliferation/survival of early hematopoietic cells and mast cell formation, but not for erythroid and myeloid terminal differentiation. Blood, The Journal of the American Society of Hematology, 89(10), 3636-3643.

Zhao, R., Watt, A. J., Li, J., Luebke-Wheeler, J., Morrisey, E. E., & Duncan, S. A. (2005). GATA6 is essential for embryonic development of the liver but dispensable for early heart formation. Molecular and cellular biology, 25(7), 2622-2631.

Watt, A. J., Zhao, R., Li, J., & Duncan, S. A. (2007). Development of the mammalian liver and ventral pancreas is dependent on GATA4. BMC developmental biology, 7(1), 1-11.

Divine, J. K., Staloch, L. J., Haveri, H., Jacobsen, C. M., Wilson, D. B., Heikinheimo, M., & Simon, T. C. (2004). GATA-4, GATA-5, and GATA-6 activate the rat liver fatty acid binding protein gene in concert with HNF-1α. American Journal of Physiology-Gastrointestinal and Liver Physiology, 287(5), G1086-G1099.

Hung, H. L., Pollak, E. S., Kudaravalli, R. D., Arruda, V., Chu, K., & High, K. A. (2001). Regulation of human coagulation factor X gene expression by GATA-4 and the Sp family of transcription factors. Blood, The Journal of the American Society of Hematology, 97(4), 946-951.

Island, M. L., Fatih, N., Leroyer, P., Brissot, P., & Loréal, O. (2011). GATA-4 transcription factor regulates hepatic hepcidin expression. Biochemical Journal, 437(3), 477-482.

Zheng, R., Rebolledo-Jaramillo, B., Zong, Y., Wang, L., Russo, P., Hancock, W., ... & Blobel, G. A. (2013). Function of GATA factors in the adult mouse liver. PLoS One, 8(12), e83723.

Journal of the American Society of Hematology, 89(4), 1260-1269.

Fasouli, E. S., & Katsantoni, E. (2021). JAK-STAT in early hematopoiesis and leukemia. Frontiers in Cell and Developmental Biology, 9, 669363.

Gillinder, K. R., Tuckey, H., Bell, C. C., Magor, G. W., Huang, S., Ilsley, M. D., & Perkins, A. C. (2017). Direct targets of pSTAT5 signalling in erythropoiesis. PLoS One, 12(7), e0180922.

Kurita, R., Suda, N., Sudo, K., Miharada, K., Hiroyama, T., Miyoshi, H., ... & Nakamura, Y. (2013). Establishment of immortalized human erythroid progenitor cell lines able to produce enucleated red blood cells. PloS one, 8(3), e59890.

Aranguren, X. L., Beerens, M., Coppiello, G., Wiese, C., Vandersmissen, I., Lo Nigro, A., ... & Luttun, A. (2013). COUP-TFII orchestrates venous and lymphatic endothelial identity by homo-or hetero-dimerisation with PROX1. Journal of cell science, 126(5), 1164-1175.

Dzierzak, E., & Bigas, A. (2018). Blood development: hematopoietic stem cell dependence and independence. Cell stem cell, 22(5), 639-651.

Gumina, R. J., Kirschbaum, N. E., Piotrowski, K., & Newman, P. J. (1997). Characterization of the human platelet/endothelial cell adhesion molecule-1 promoter: identification of a GATA-2 binding element required for optimal transcriptional activity. Blood, The

CHAPTER 3:

CHARACTERIZATION OF THE IMPACT OF COUP-TFII OVEREXPRESSION AND KNOCKOUT ON K562 ERYTHROID CELLS

3.1. Introduction

COUP-TFII is an important regulator of cell terminal differentiation and fate choice in different tissues (Xie, 2011; Yu, 2012; Xie, 2017; Ceni, 2017). In endothelial cells, reduced expression of COUP-TFII affects G1/S progression and decreases proliferation. Conversely, COUP-TFII was shown to induce proliferation of endothelial cells (You, 2005; Chen, 2012). On the other hand, several pieces of evidence indicate that COUP-TFII may function in part as a tumor suppressor. Prahalad and colleagues (2010), suggested COUP-TFII as a mediator of retinoic acid-induced reduction of cell proliferation in breast cancer. Another study demonstrated that COUP-TFII silencing leads to decreased proliferation and increased apoptosis of renal cell carcinoma cells (Zheng, 2016; Fang, 2019). Therefore, the impact of COUP-TFII on cell proliferation is context-dependent. Within erythroid lineage, COUP-TFII seems to not have any strong impact neither on proliferation nor differentiation in adult cells, as demonstrated in human peripheral blood cells overexpressing COUP-TFII (Fugazza, 2021).

Being an activator of fetal γ-globin, COUP-TFII is considered as an attractive potential therapeutic target in β-hemoglobinopathies. Therefore, modulation of COUP-TFII levels in adult erythroid progenitors should not significantly affect their proliferation/differentiation balance.

To this end, I evaluated the short-term effect of the overexpression or depletion of COUP-TFII on K562 erythroid cells' proliferation, differentiation and cell cycle by using flow cytometry up to day 7 post infection (pi). First, to assess the cell proliferation rate, I measured dilutions of Cell Trace Violet (CTV) dye. CTV dye covalently binds to amine groups in proteins and it gets diluted through subsequent cell divisions allowing the analysis of the fluorescence intensities of labeled cells, and therefore the determination of the number of generations. Second, I analyzed erythroid differentiation by measuring surface expression levels of erythroid differentiation markers: CD71 (transferrin receptor) and glycophorin A (GpA): K562 cells that undergo erythroid differentiation progressively loose CD71 and increase GpA protein levels. Third, I evaluated the impact of COUP-TFII overexpression or knockout on K562 cell cycle by using propidium iodide (PI), a fluorescent dye that intercalates with DNA. Therefore, the proportion of cells in each stage of the cell cycle is calculated by measuring variations in fluorescence intensity.

3.2. Materials and methods

Cell cultures

K562: K562 is a human immortalized erythroid cell line derived from a female patient with chronic myelogenous leukemia, highly undifferentiated of the granulocytic series (Lozzio and Lozzio, 1975). Cells were cultured in RPMI 1640 (Lonza, BE12-167F) supplemented with 10% FBS (Sigma-Aldrich, F7524-500ML), 4mM L-Glutamine (Euroclone, ECB3000D) and 100µg/mL Penicillin-Streptomycin (Euroclone, ECB3001D), at an optimal concentration of 0.2 $x10^6$ cells/mL, at 37°C in 5% CO_2.

HEK-293T: HEK-293T cells were used for the production of lentiviral vector overexpressing COUP-TFII. Cells were cultured in DMEM (Sigma-Aldrich, D6429-500ML) with 10% FBS (Sigma-Aldrich, F9665-500ML) and 100µg/mL Penicillin-Streptomycin (Gibco, 15140122), at 37°C in 5% CO_2.

CRISPR/Cas9-mediated COUP-TFII knockout

The sequences of the sgRNA targeting COUP-TFII (5'-CCCAACCAGCCGACGAGATT-3') was obtained from the human GeCKOv2 library (https://www.addgene.org/crispr/libraries/geckov2) and the sgRNA oligos were cloned into the LentiCrisprV2 plasmid

(Addgene, 52961). HEK-293T cells were used for the lentiviral production. 24 h after transduction, K562 cells were selected with puromycin and analyzed 96 h and 7 days after transduction. The empty LentiCrisprV2 plasmid (empty vector, EV) was used as a negative control.

COUP-TFII lentiviral overexpression

COUP-TFII murine cDNA (gift from Dr. Michèle Studer) was cloned into a bicistronic IRES-eGFP as described previously (Fugazza, 2021). COUP-TFII overexpressing vector and control empty vector (EV) were produced in HEK-293T cells, lentiviral vectors were co-transfected together with the packaging plasmids psPAX2 and pMD-VSVG (www.lentiweb.com). Viral vector supernatants were collected 72 h after transfection and concentrated, followed by infection of K562 cells. Cells were analyzed 72 h, 5 days and 7 days post infection.

RNA isolation and cDNA synthesis

TriFast reagent (Euroclone) and the Direct-zol™ RNA Miniprep kit (Zymo Research, R2050) were used to isolate total RNA. Next, RNA was reverse-transcribed into cDNA with the High-Capacity cDNA Reverse Transcription Kit (Applied Biosystems, 4368814).

Quantitative real-time PCR analysis

On the basis of sequences from the Ensembl database (http://www.ensembl.org), primers were designed to amplify 100 to 200 bp amplicons, spanning an exon-exon junction when possible. All the RT-qPCR primers used in this study are listed in the table below. Quantitative real-time PCR reaction was performed using StepOnePlus Real-Time PCR System (Applied Biosystems). Specific PCR product accumulation was monitored by SsoAdvanced Universal SYBR Green Supermix (Bio-Rad, 1725274) in a 12µL reaction volume. Gene expression levels were calculated by normalizing the values to the expression of housekeeping GAPDH gene. Triplicate samples from each experiment were analyzed.

Table 3.1. ***Primers used for RT-qPCR.***

Gene	F/R	Sequence (5'–3')
hGAPDH	F	ACGGATTTGGTCGTATTGGG
	R	TGATTTTGGAGGGATCTCGC
hCOUP-TFII	F	TTGACTCAGCCGAGTACAGC
	R	AAAGCTTTCCGAATCTCGTC
hCD71	F	CTTCAAGCTCCTGGGAAATGT
	R	GCAGAATAAAGCCTATCCTTGAAAG
hGpA	F	TCAGGAAACTTCTTCTGCTCACA
	R	AGGTTCCCAGGTACTGCT

Western blot

Nuclear protein extracts were resolved by SDS/PAGE in a 10% gel and blotted on Hybond ECL Nitrocellulose membrane (GE healthcare life science-Amersham) for 90 min at 4°C. Next, membranes were blocked for 1 h at room temperature by using 5% milk in TBST 1X (Tris Buffered Saline, pH 7.6 and 0,1% Tween 20 by Sigma) and incubated with the primary antibodies overnight at 4°C. On the next day, membranes were washed and incubated with secondary antibodies for 1 h at room temperature, followed by revealing.

Table 3.2. ***Antibodies used for Western blot.***

Antibody	Catalog number	Manufacturer
Anti-COUP-TFII (mouse/human)	A10251	Abclonal
Anti-U2AF (mouse/human)	U4758	Sigma-Aldrich

Flow cytometry

Cells were washed twice in PBS and fixed with 1% PFA at room temperature for 20 min, followed by incubation with CD71 (PE conjugated, Biolegend, 334105) and GpA/CD235ab (APC conjugated, Biolegend, 306607) antibodies for 20 min, at room temperature, protected from the light. Next, cells were washed in PBS and acquired. Data were recorded on a FACS LSR Fortessa X-20 (BD Life Sciences) and analyzed with FlowJo™ v10.8.2 Software (BD Life Sciences) by C. Pitsillidou and A. Roberto (FlowMetric Europe, Milan, Bresso, Italy).

Cell proliferation assay

To assess the cell proliferation CellTrace™ Violet (CTV) Cell Proliferation Kit (Invitrogen, C34557) was applied. Briefly cells were washed with PBS, resuspended in 2μM CTV solution and incubated for 15 min at 37^{o}C, protected from the light. In order to remove any residual dye remaining in the solution, cells were next incubated for 5 min with complete cell medium, followed by lentiviral transduction. Cells were acquired at indicated timepoints after transduction on a FACS LSR Fortessa X-20 (BD Life Sciences) and analyzed with FlowJo™ v10.8.2 Software (BD Life Sciences) by C. Pitsillidou and A. Roberto (FlowMetric Europe, Milan, Bresso, Italy).

Cell cycle analysis

Cells were washed twice with PBS, followed by the incubation with 1% PFA (paraformaldehyde) for 45 min. The same volume of PBS with 0.2% TritonX (Sigma, xxx) was added and cells were incubated for 15 min more. Next, cells were centrifuged (5 min, 300 xg), resuspended in PBS with propidium iodide (PI, 1mg/mL) and RNase (100 μg/mL), vortexed for 5 sec and incubated at room temperature for 1 h, protected from the light. Data were recorded on a FACS LSR Fortessa X-20 (BD Life Sciences) and analyzed with FlowJo™ v10.8.2 Software (BD Life Sciences) by C. Pitsillidou and A. Roberto (FlowMetric Europe, Milan, Bresso, Italy).

Statistical analyses

For statistical analyses was used GraphPad Prism 8.4.3 software (GraphPad Software). Data are expressed as mean ±SEM and all experiments were performed in biological triplicates, unless stated otherwise. To calculate the p-value, the paired Student's t-test was used. $P<0.05$ was considered to be statistically significant (*$p<0.05$, **$<p<0.01$, ***$p<0.001$, ****$p<0.0001$, if not stated data not significant).

3.3. Results and discussion

To characterize the overall impact of COUP-TFII on phenotype of K562 erythroid cells, I analyzed by flow cytometry changes in cell proliferation, erythroid differentiation markers' expression and cell cycle upon COUP-TFII depletion or overexpression. These experiments were performed in collaboration with C. Pitsillidou and A. Roberto (FlowMetric Europe, Milan, Bresso, Italy).

K562 cells were labelled with CellTrace Violet (CTV) dye, followed by lentiviral transduction with either a lentiviral vector overexpressing COUP-TFII (OE) or targeting COUP-TFII gene (KO). Cells transduced with the empty vector (EV) were used as a control. Cell proliferation, expression levels of erythroid markers CD71 and GpA and cell cycle phases of COUP-TFII OE and KO K562 cell bulk population were measured for 7 days post infection (pi) (Fig. 3.1).

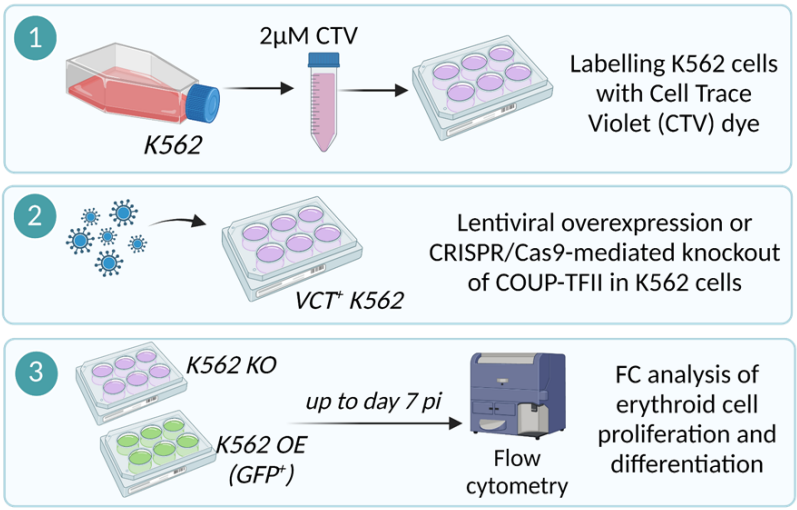


Figure 3.1. ***Scheme of the experiment to assess the effect of COUP-TFII on cell proliferation and differentiation of K562 erythroid cells****.*

The COUP-TFII OE and control (empty vector, EV) lentiviral vectors express GFP: the efficiency of the transduction was monitored over time and reached 80% in OE cells, and over 90% in EV cells (Fig. 3.2A). Cells transduced with single guide RNA expressing vector (KO) and EV control cells were selected with puromycin 24 h post infection. COUP-TFII overexpression and knockout were confirmed at protein and mRNA levels (Fig. 3.2B and 3.2C, respectively).

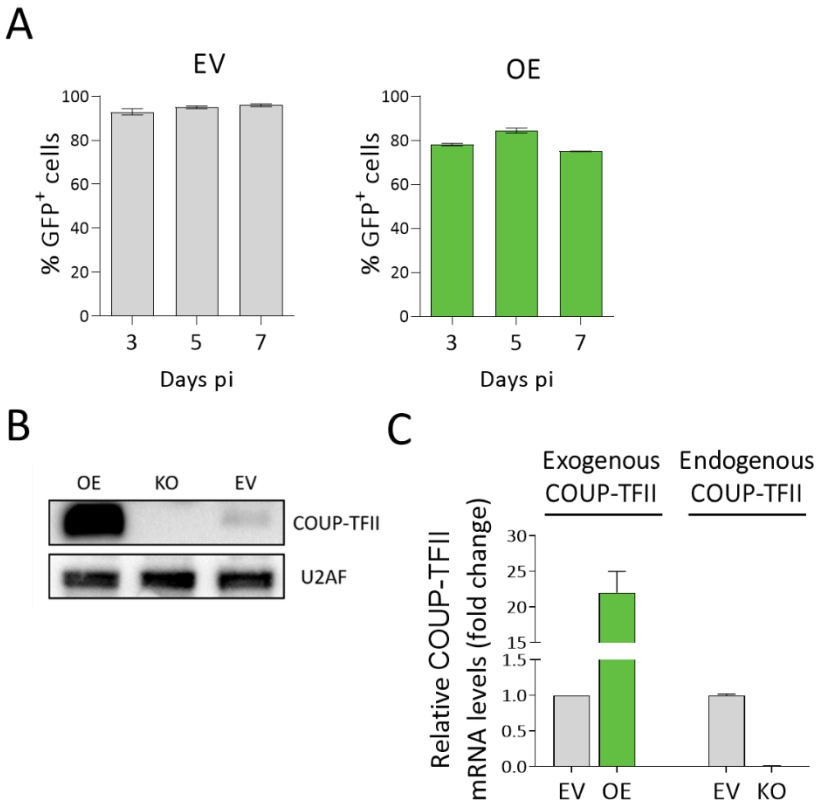


Figure 3.2. ***COUP-TFII overexpression and knockout in K562 erythroid cells. A*** *The efficiency of the lentiviral overexpression of COUP-TFII. Percentage of GFP⁺ cells was measured by flow cytometry at indicated timepoints post infection (pi) with the COUP-TFII overexpressing lentiviral vector (OE) and the control empty vector (EV), both expressing GFP, MOI=30. Data as mean ± SEM, n≥2.* ***B*** *Western blot analysis by using an anti-COUP-TFII antibody that detects both endogenous and exogenous COUP-TFII of K562 cells with COUP-TFII OE (72 h pi) and COUP-TFII KO (96 h pi), EV – control cells. Anti-U2AF antibody was used as loading control.* ***C*** *Left: expression level of the exogenous (mouse) COUP-TFII (72 h pi), right: expression level of the endogenous COUP-TFII (96 h pi), measured by RT-qPCR. Data as mean ± SEM, n≥2.*

Cell proliferation rate was assessed by CellTrace Violet (CTV) dye dilution: the lower mean fluorescence intensity (MFI) value of CTV, the more cells proliferate. As shown in the figure 3.3A, K562 erythroid cell proliferation was not affected by the overexpression of COUP-TFII, whereas COUP-TFII knockout slightly reduced proliferation of K562 cells (*=$p<0.05$ on day 7 pi, Fig. 3.3B).

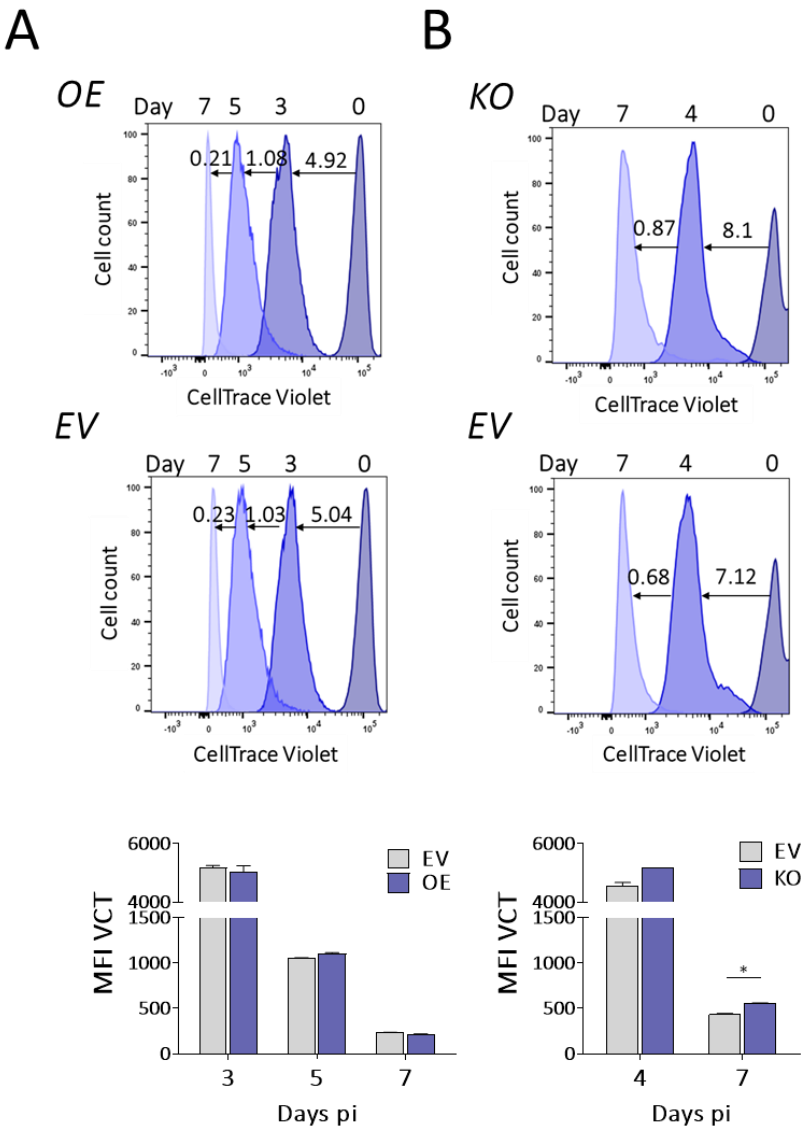

A
B
OE
Day 7 5 3 0
0.21 1.08 4.92
Cell count
CellTrace Violet
KO
Day 7 4 0
0.87 8.1
Cell count
CellTrace Violet
EV
Day 7 5 3 0
0.23 1.03 5.04
Cell count
CellTrace Violet
EV
Day 7 4 0
0.68 7.12
Cell count
CellTrace Violet
MFI VCT
6000
4000
1500
1000
500
0
EV
OE
3 5 7
Days pi
MFI VCT
6000
4000
1500
1000
500
0
EV
KO
*
4 7
Days pi

Figure 3.3. ***Cell proliferation of K562 erythroid cells expressing different COUP-TFII levels.*** *Histograms show dilution of CellTrace Violet (CTV) over 7 days indicating the proliferation rate of K562 cells upon COUP-TFII OE* ***(A)*** *or KO* ***(B).*** *Cells were labelled with 2µM CTV prior to infection and acquired by flow cytometry 3, 5 and 7 (OE) or 4 and 7 (KO) day. Dilutions of CTV dye normalized to day 0 are indicated above arrows. Bar charts show median fluorescence intensity (MFI) values of CTV at each timepoint. Data as mean ± SEM, n≥2, *=p<0.05.*

To assess the impact of COUP-TFII on erythroid differentiation, cell surface expression of early (CD71) and late (GpA) erythroid markers was measured by flow cytometry.

Upon COUP-TFII overexpression (OE), the percentage of $CD71^{+}GpA^{-}$ K562 cell population decreased from 4.3% on day 3 post infection (pi) to 1% on day 7 pi, while the percentage of $CD71^{-}GpA^{+}$ cells increased from 0.6% (day 3 pi) to 3.9% (day 5 pi) and 3.3% (day 7 pi) (Fig. 3.4A). However, the percentage of infected GFP^{+} K562 cells remained stable during the experiment (Fig. 3.2A) suggesting that COUP-TFII OE even if prompt differentiation of K562 erythroid cells, it did not result in their terminal differentiation. The depletion of COUP-TFII (KO) did not change the percentage of CD71/GpA expressing cells (Fig. 3.4B). These results demonstrated that COUP-TFII level variations in K562 cells do not cause any significant percentage changes in neither erythroid cell proliferation nor differentiation.

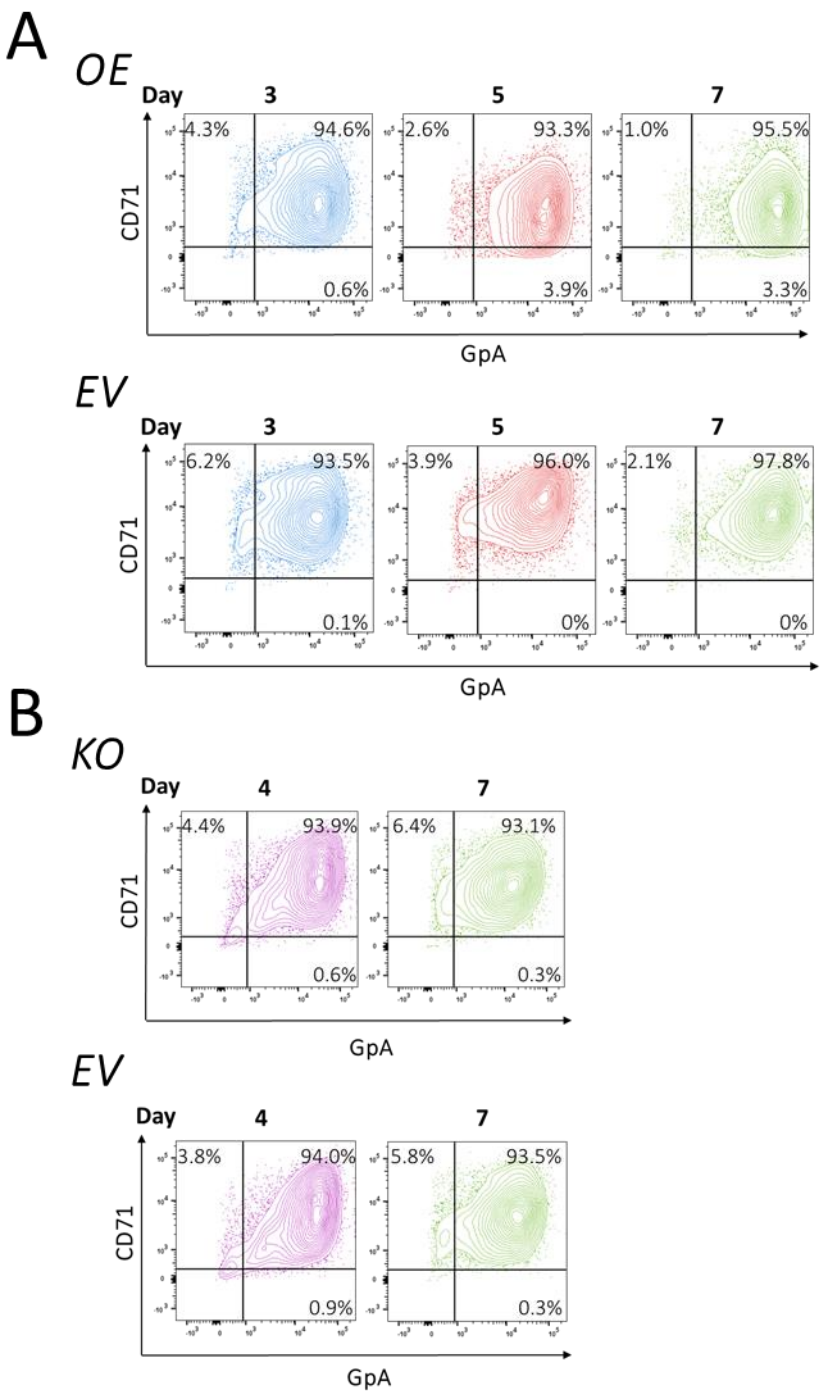

A
OE
Day
3
5
7
4.3%
94.6%
0.6%
2.6%
93.3%
3.9%
1.0%
95.5%
3.3%
CD71
GpA
EV
Day
3
5
7
6.2%
93.5%
0.1%
3.9%
96.0%
0%
2.1%
97.8%
0%
CD71
GpA
B
KO
Day
4
7
4.4%
93.9%
0.6%
6.4%
93.1%
0.3%
CD71
GpA
EV
Day
4
7
3.8%
94.0%
0.9%
5.8%
93.5%
0.3%
CD71
GpA

Figure 3.4. ***Cell differentiation of K562 erythroid cells upon overexpression or knockout of COUP-TFII.*** *Flow cytometric analysis of the levels of erythroid differentiation surface markers: CD71 and GpA in OE* ***(A)****, KO* ***(B)*** *cells, EV - control. Percentage of cells in gates, data as mean ± SEM, n≥2.*

To have a better understanding of COUP-TFII-induced changes in erythroid differentiation, I compared the mean fluorescence intensity (MFI) values (that indicates the protein expression level) of GpA within the most abundant double-positive ($CD71^{+}GpA^{+}$) K562 cell population. GpA MFI was 1.5-fold increased in cells in which COUP-TFII was overexpressed (OE) when compared to control EV cells (*p<0.05 on day 3, **p<0.01 on day 5 pi) (Fig. 3.5A) and decreased in COUP-TFII knockout (KO) K562 cells (**p<0.01 on day 4 pi) (Fig. 3.5B).

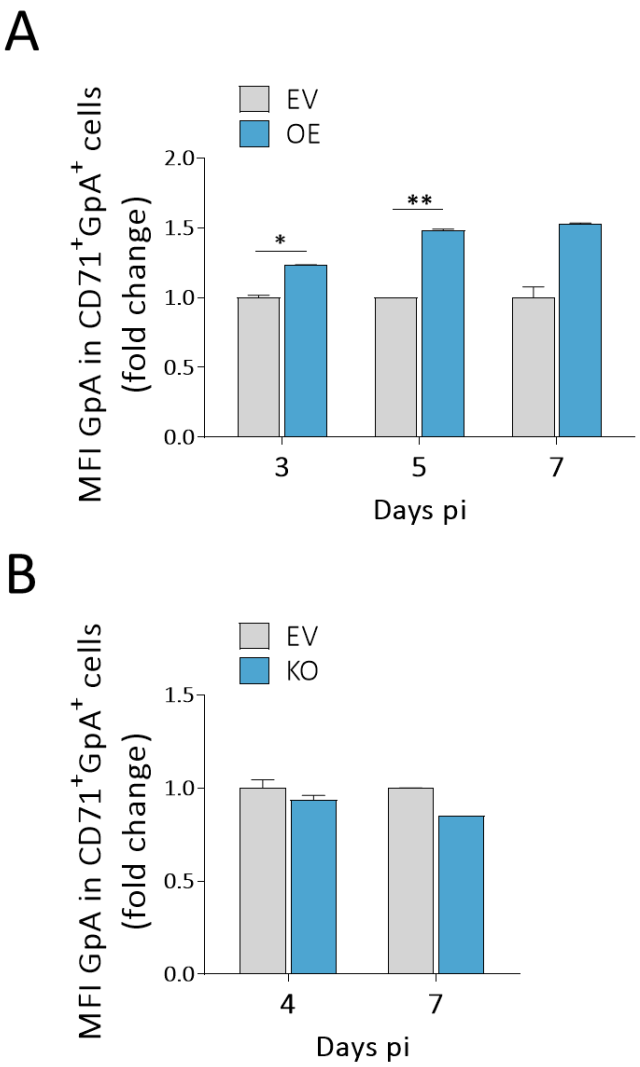


Figure 3.5. ***COUP-TFII overexpression results in increased surface expression of glycophorin A (GpA) in K562 cells.*** *Flow cytometry analysis of GpA levels in COUP-TFII OE/KO K562 cells gated as $CD71^{+}GpA^{+}$, EV – control cells. Data as mean ± SEM, n≥2, *=p<0.05, **=p<0.01.*

K562 erythroid cells die as a result of their terminal differentiation. Although the number of COUP-TFII overexpressing cells (% of GFP^+ cells) cells did not change during the experiment (Fig. 3.2A) and therefore it is unlikely that COUP-TFII induced K562 terminal erythroid differentiation. At the mRNA level, CD71 showed 20% reduction in COUP-TFII KO cells, and the same trend (not significant) was observed in COUP-TFII OE cells (Fig. 3.6A). GpA mRNA level was significantly upregulated in OE cells (4-fold when compared to EV control), and downregulated (20% reduction) when COUP-TFII was not present (Fig. 3.6B).

Hence, COUP-TFII overexpression leads to increase expression of glycophorin A (GpA) at both protein and mRNA level. However, the positive effect of COUP-TFII on erythroid differentiation is mild and does not result in a significant change in differentiation in term of number of cells undergoing terminal erythroid differentiation.

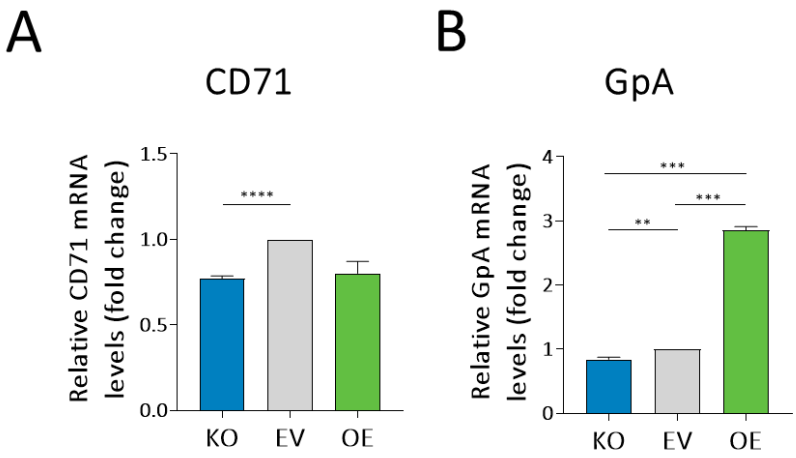


Figure 3.6. ***COUP-TFII regulates the expression of erythroid differentiation markers CD71 and GpA.*** *Effects of COUP-FII overexpression (OE, 72 h pi) and COUP-TFII knockout (KO, 96 h pi) on CD71* ***(A)*** *and GpA* ***(B)*** *mRNA levels. Data as mean ± SEM, n≥3, **=p<0.01, ***=p<0.001, ****=p<0.0001.*

Finally, the propidium iodide (PI) staining did not show any significant changes in K562 cell cycle upon COUP-TFII overexpression or knockout (Fig. 3.7A,B).

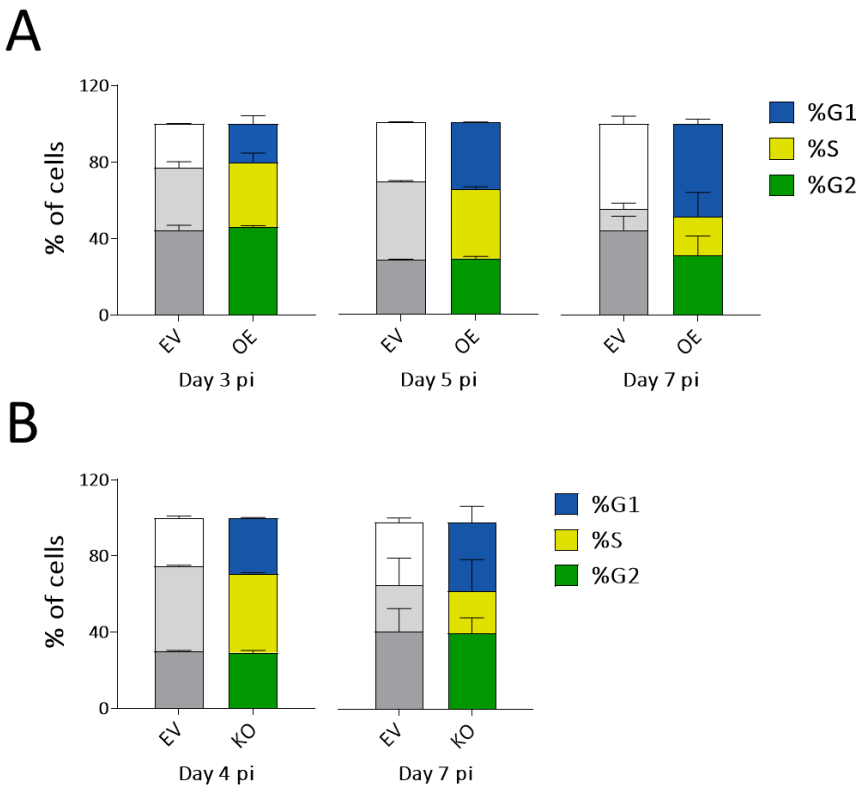


Figure 3.7. ***K562 cell cycle is not affected by COUP-TFII.*** *Cell cycle analysis (by using propidium iodide) of K562 cells upon COUP-TFII overexpression (OE) and knockout (KO) at indicated timepoints. Data as mean ± SEM, n≥2.*

Overall, COUP-TFII overexpression in the short-term modulates mRNA levels of erythroid differentiation markers by reducing CD71 and inducing GpA. However, the effect of COUP-TFII on both K562 erythroid cell proliferation and CD71/GpA profile is moderate. This suggests that COUP-TFII could be an attractive therapeutic target in β-hemoglobinopathies. Of note, the detailed analysis reported here is in line with previous results from our lab showing that the exogenous expression of COUP-TFII in human adult peripheral blood cells from healthy donors and from thalassemic patients did not cause any significant maturational changes (Fugazza, 2021).

3.4. References

Xie, X., Qin, J., Lin, S. H., Tsai, S. Y., & Tsai, M. J. (2011). Nuclear receptor chicken ovalbumin upstream promoter-transcription factor II (COUP-TFII) modulates mesenchymal cell commitment and differentiation. Proceedings of the National Academy of Sciences, 108(36), 14843-14848.

Yu, C. T., Tang, K., Suh, J. M., Jiang, R., Tsai, S. Y., & Tsai, M. J. (2012). COUP-TFII is essential for metanephric mesenchyme formation and kidney precursor cell survival. Development, 139(13), 2330-2339.

Xie, X., Wu, S. P., Tsai, M. J., & Tsai, S. (2017). The role of COUP-TFII in striated muscle development and disease. Current Topics in Developmental Biology, 125, 375-403.

Ceni, E., Mello, T., Polvani, S., Vasseur-Cognet, M., Tarocchi, M., Tempesti, S., ... & Galli, A. (2017). The orphan nuclear receptor COUP-TFII coordinates hypoxia-independent proangiogenic responses in hepatic stellate cells. Journal of hepatology, 66(4), 754-764.

You, L. R., Lin, F. J., Lee, C. T., DeMayo, F. J., Tsai, M. J., & Tsai, S. Y. (2005). Suppression of Notch signalling by the COUP-TFII transcription factor regulates vein identity. Nature, 435(7038), 98-104.

Chen, X., Qin, J., Cheng, C. M., Tsai, M. J., & Tsai, S. Y. (2012). COUP-TFII is a major regulator of cell cycle and Notch signaling pathways. Molecular endocrinology, 26(8), 1268-1277.

Prahalad, P., Dakshanamurthy, S., Ressom, H., & Byers, S. W. (2010). Retinoic acid mediates regulation of network formation by COUP-TFII and VE-cadherin expression by TGFβ receptor kinase in breast cancer cells. PloS one, 5(4), e10023.

Zheng, J., Qin, W., Jiao, D., Ren, J., Wei, M., Shi, S., ... & Wen, W. (2016). Knockdown of COUP-TFII inhibits cell proliferation and induces apoptosis through upregulating BRCA1 in renal cell carcinoma cells. International Journal of Cancer, 139(7), 1574-1585.

Fang, X., Liu, C. X., Zeng, X. R., Huang, X. M., Chen, W. L., Wang, Y., & Ai, F. (2020). Orphan nuclear receptor COUP-TFII is an oncogenic gene in renal cell carcinoma. Clinical and Translational Oncology, 22, 772-781.

Fugazza, C., Barbarani, G., Elangovan, S., Marini, M. G., Giolitto, S., Font-Monclus, I., ... & Ronchi, A. E. (2021). The Coup-TFII orphan nuclear receptor is an activator of the γ-globin gene. haematologica, 106(2), 474.

Lozzio, C. B., & Lozzio, B. B. (1975). Human chronic myelogenous leukemia cell-line with positive Philadelphia chromosome.

CHAPTER 4: DIRECT TARGET GENES OF COUP-TFII IN K562 ERYTHROID CELLS

4.1. Introduction

COUP-TFII functions as a transcriptional activator or repressor, in a context-dependent manner, through its binding to AGGTCA direct repeat (DR) elements, both as a homo- or heterodimer (Kliewer, 1992; Cooney, 1992; Schaeffer, 1993). COUP-TFII is a known positive regulator of several genes, for example the rat cholesterol 7α-hydroxylase (CYP7A1) on which it acts synergistically with hepatocyte nuclear factor 4 (HNF4) (Stroup and Chiang, 2000). Similarly, COUP-TFII can induce the HNF-1α gene promoter by physical interaction with HNF4 in the liver (Ktistaki and Talianidis, 1997). What is more, COUP-TFII can inhibit transcription by the recruitment of co-repressors, or by transrepression when directly bound to the ligand binding domain (LBD) of nuclear receptors (NRs) (Okamura, 2009; Leng, 1996).

Within the β-locus, COUP-TFII directly binds to the hypersensitive sites HS2, HS3 and HS4 of the locus control region (LCR) distal enhancer and induce the expression of fetal γ-globin, most likely by favoring the interaction of the LCR with the γ-globin genes' promoters (Fugazza, 2021). The induction of γ-globin by COUP-TFII was observed in adult erythroid cultures from peripheral blood cells, both healthy and thalassemic. However, the mechanism by which COUP-TFII can overcome the repressive adult environment to induce the expression of HBG genes, is not well characterized. It is hypothesized that COUP-TFII competes with BCL11A, a γ-globin repressor, for the binding to the

same consensus sequence. The GGTCA consensus sequence of COUP-TFII almost perfectly matches the BCL11A motif (Sankaran, 2008; Sankaran, 2011; Liu 2018), therefore the same regions of the β-locus are likely bound by COUP-TFII or BCL11A at different developmental stages, resulting in either activation or repression.

Therapeutic strategies to treat β-hemoglobinopathies aim to reactivate fetal hemoglobin (HbF) in adult erythroid, and, in consequence, to ameliorate the clinical symptoms of the disease. Therefore, one of the possible approaches could be the reactivation of COUP-TFII that would in turn induce the γ-globin expression in adult erythroid cells. Thus, it is crucial to better characterize the mechanism by which COUP-TFII controls the β-locus, as well as the downstream network of COUP-TFII in erythroid cells that, up-to-date, remains greatly unknown.

To better characterize downstream target genes of COUP-TFII in erythroid cells, I first focused on the targets with a known and important role in erythropoiesis.

Next, I extended the COUP-TFII downstream network characterization to other genes, that were not previously linked with erythroid program.

In the section 4.3.1, I described the CUT&RUN experiment performed in collaboration with G. Zambanini and C. Cantu (Linköping University, Linköping, Sweden) in adult-like HUDEP-1 and HUDEP-2 erythroid cells expressing exogenous COUP-TFII in order to map COUP-TFII binding

within the β-locus. HUDEP-1 and HUDEP-2 cells do not express COUP-TFII and only the second ones express the adult γ-globin repressor, BCL11A. Thus, by using these cellular models, it was possible to assess the mechanism by which COUP-TFII re-activates the expression of fetal γ-globin, in the presence or absence of γ-globin repressor.

In the section 4.3.2, I introduced MYB, an important erythroid transcription factor, that emerged as a candidate direct downstream target of COUP-TFII. By cloning the genomic regions located upstream to MYB gene that were bound by COUP-TFII in the ChIP-seq experiment performed in K562 erythroid cells, I evaluated their relevance in the possible control mechanism of MYB by COUP-TFII.

In the section 4.3.3, potential novel direct target genes were selected by integrating ChIP-seq (described previously in Fugazza, 2021) and RNA-seq data from K562 erythroid cells expressing different levels of COUP-TFII. Further analysis narrowed down the list to twelve genes, that were either upregulated or downregulated by COUP-TFII and have unknown (or poorly understood) role in erythropoiesis. By CRISPR/Cas9-mediated knockout of these genes and further flow cytometry analysis, I evaluated their impact on the erythroid cell proliferation and differentiation. Furthermore, I described more in details a possible direct mechanism by which COUP-TFII could control APOE gene encoding apolipoprotein E, that emerged as the most interesting novel direct target of COUP-TFII in erythroid cells.

4.2. Materials and Methods

Reporter vector construction

All cloned DNA sequences were obtained from UCSC Genome Browser (https://genome.ucsc.edu/, human GRCh37/hg19 assembly) and amplified from genomic DNA of K562 cells by using Phusion® High-Fidelity DNA Polymerase (New England BioLabs, M0530S) according to standard laboratory protocols. P1 and P4 regions were cloned into pGL3-Basic reporter plasmid (Promega, E1751) containing Thymidine Kinase promoter (pTK), upstream to the promoter, between SmaI and KpnI sites. Primers used for the amplification of the P1 and P4 regions are listed below.

Table 4.1. ***Primers used for cloning.***

Construct	F/R	Sequence (5'–3')
P1	F	ACGTGGTACCATTGAATTCTGATGATAC
	R	GATTAAAAGATATCACACCCAC
P4	F	ACGTGGTACCAAAGGAATAAAGTAAAGGGG
	R	AGTCGATATCTCCTAGGTCCTACAGATAG

Cell cultures

K562: K562 is a human immortalized erythroid cell line derived from a female patient with chronic myelogenous leukemia, highly

undifferentiated of the granulocytic series (Lozzio and Lozzio, 1975). Cells were cultured in RPMI 1640 (Lonza, BE12-167F) supplemented with 10% FBS (Sigma-Aldrich, F7524-500ML), 4mM L-Glutamine (Euroclone, ECB3000D) and 100µg/mL Penicillin-Streptomycin (Euroclone, ECB3001D), at an optimal concentration of 0.2 $x10^6$ cells/mL, at 37°C in 5% CO_2.

K562-Cas9: K562 cell line stably expressing Cas9 was kindly provided by Jürg Schwaller and Samantha Tauchmann (University of Basel, Switzerland). Cells were grown in RPMI (Gibco, 11875093) with 10% FBS (Sigma-Aldrich, F9665-500ML), 2 mM L-Glutamine (Gibco, 25030024) and 100µg/mL Penicillin-Streptomycin (Gibco, 15140122), at 37°C in 5% CO_2 **.**

HUDEP-1 and HUDEP-2: human umbilical cord blood-derived erythroid progenitor (HUDEP-1 and HUDEP-2) cell lines were maintained in Iscove's modified Dulbecco's medium (IMDM, PAN, P04-20250) supplemented with 0.2% Human Serum Albumin (HAS, PAN, P06-27100), 0.1% Lipid Mixture (Sigma-Aldrich, L0288), 300 µg/mL transferrin (Sigma-Aldrich, T1408), 10 µg/mL insulin (Sigma-Aldrich, I2643), 100 µM Sodium pyruvate (Sigma-Aldrich, P5280), 2 U/mL Erythropoietin (EPO, ImmunoTools, 11344795), 100 ng/mL Stem Cell Factor (SCF, ImmunoTools, 11343325), 1 µM dexamethasone (Sigma-Aldrich, D4902) and 2 µg/mL doxycycline (Sigma-Aldrich, D9891).

LentiX: Lenti-X 293T (Takara Bio, 632180) cells were used for the lentiviral production in CRISPR/Cas9 experiment. Cells were cultured in DMEM (Gibco, 11960044) with 10% FBS (Sigma-Aldrich, F9665-

500ML), 4 mM L-Glutamine (Gibco, 25030024) and 100µg/mL Penicillin-Streptomycin (Gibco, 15140122), at 37°C in 5% CO_2.

HEK-293T: HEK-293T cells were used for the production of lentiviral vector overexpressing COUP-TFII. Cells were cultured in DMEM (Sigma-Aldrich, D6429-500ML) with 10% FBS (Sigma-Aldrich, F9665-500ML) and 100µg/mL Penicillin-Streptomycin (Gibco, 15140122), at 37°C in 5% CO_2.

Transfection of K562 cells

3 x10^5 cells in 0.6mL of Opti-MEM Reduced-Serum Medium (Gibco, 31985-047) were seeded on a 24-well plate and transfected with 0.8µg of reporter plasmid DNA and 2µL of Lipofectamine 2000 Transfection Reagent (Invitrogen, 11668019) per well. For the co-transfection, 0.8µg of reporter plasmid DNA and indicated amounts (0, 0.4 or 1 µg) of the COUP-TFII expressing plasmid (pCAG-COUP-TFII) per well were used. The pUC19 empty vector was added at the concentration required to balance the total amount of DNA transfected in each reaction. 4 h after transfection, 0.6mL of RPMI 1640 with 20% FBS, L-Glutamine and Penicillin-Streptomycin per well were added.

Luciferase assay

K562 cells were collected 48 h after transfection and centrifuged (1200 rpm, 5 min). Pellet was resuspended in 80µL of Reporter Lysis Buffer

1X (Promega, E4030), followed by two freeze-thaw cycles. Cell lysates were centrifuged (12,000 rcf, 5 min, at 4°C) and the supernatant (containing total protein extract) was collected. Next, 25µL of LAR reagent (Promega, 3040) were added to 10µL of protein extract and the luciferase activity was measured by GloMax® 20/20 Luminometer (Promega), integration time of 10 sec. To calculate the relative luciferase activity, amount of relative luciferase units (RLU) of non-transfected control was subtracted from each sample RLU, and normalized to pTK reference samples.

RNA isolation and cDNA synthesis

TriFast reagent (Euroclone) and the Direct-zol™ RNA Miniprep kit (Zymo Research, R2050) were used to isolate total RNA. Next, RNA was reverse-transcribed into cDNA with the High-Capacity cDNA Reverse Transcription Kit (Applied Biosystems, 4368814).

Quantitative real-time PCR analysis

On the basis of sequences from the Ensembl database (http://www.ensembl.org), primers were designed to amplify 100 to 200 bp amplicons, spanning an exon-exon junction when possible. All the RT-qPCR primers used in this study are listed in the table below. Quantitative real-time PCR reaction was performed using StepOnePlus Real-Time PCR System (Applied Biosystems). Specific PCR product

accumulation was monitored by SsoAdvanced Universal SYBR Green Supermix (Bio-Rad, 1725274) in a 12µL reaction volume. Gene expression levels were calculated by normalizing the values to the expression of housekeeping GAPDH gene. Triplicate samples from each experiment were analyzed.

Table 4.2. ***Primers used for RT-qPCR.***

Gene	F/R	Sequence (5'–3')
hGAPDH	F	ACGGATTTGGTCGTATTGGG
	R	TGATTTTGGAGGGATCTCGC
hCOUP-TFII	F	TTGACTCAGCCGAGTACAGC
	R	AAAGCTTTCCGAATCTCGTC
mCOUP-TFII	F	TCCAAGAGCAAGTGGAGAAG
	R	CTTCCAAAGCACACTGGGAC
HBG	F	CTTCAAGCTCCTGGGAAATGT
	R	GCAGAATAAAGCCTATCCTTGAAAG
hMYB	F	TCAGGAAACTTCTTCTGCTCACA
	R	AGGTTCCCAGGTACTGCT
hAPOE	F	GGTCGCTTTTGGGATTACCT
	R	ACACGTCCTCCATGTCCG
hBAMBI	F	ATCGCCACTCCAGCTACATC
	R	CTAGAGAAGCAGGCGCTGAG
hENC1	F	ACTGCATTTGTCAGCACCTG
	R	GGTGCAGCAGATTTTATTCG
hGALM	F	ACACCTGCAGGACTTCCATC
	R	CCCTTTAATGTGCCATCCAG
hPLA2G6	F	CTCGGTGCAACATCATGG
	R	GTTCTTGGCCCAGTGGAG

CRISPR/Cas9-mediated COUP-TFII knockout

The sequences of the sgRNA targeting COUP-TFII (5'-CCCAACCAGCCGACGAGATT-3') was obtained from the human GeCKOv2 library (https://www.addgene.org/crispr/libraries/geckov2) and the sgRNA oligos were cloned into the LentiCrisprV2 plasmid (Addgene, 52961). Cells were selected with puromycin and analyzed 96 h after transduction, or after 7 days single clones were isolated, by limiting dilution (to a theoretical concentration of 0.5 cells/well) in a 96-well plate.

COUP-TFII lentiviral overexpression

COUP-TFII murine cDNA (gift from Dr. Michèle Studer) was cloned into a bicistronic IRES-eGFP as described previously (Fugazza, 2021). COUP-TFII overexpressing vector and control Empty vector (EV) were produced in HEK-293T cells, lentiviral vectors were co-transfected together with the packaging plasmids psPAX2 and pMD-VSVG (www.lentiweb.com). Viral vector supernatants were collected 72 h after transfection and concentrated, followed by transduction of K562 or HUDEP-2 cells.

ChIP-seq assay and ChIP-seq data analysis

ChIP-seq experiment was performed as previously (Fugazza, 2021). Briefly, K562 cells were fixed with 1% formaldehyde for 10 minutes, at

room temperature. Chromatin was sonicated to a size of about 500 bp. DNA immunoprecipitation was obtained by first incubating with the proper antibodies and then isolating with protein A-agarose beads. DNA was sequenced on an Illumina platform. The sequencing data were uploaded to the Galaxy web platform. BWA was used to map reads to the human genome (GRCh37/hg19), and the HOMER toolkit was used to identify peaks. Peaks in different experiments were called as the same if the minimal overlap of intervals was 1bp. ChIP-seq experiment was performed in K562 cells with COUP-TFII overexpression (OE, 25 days after transduction), COUP-TFII knockout (KO, 30 days after transduction), or non-treated.

CUT&RUN assay

The CUT&RUN experiment was performed to detect COUP-TFII occupancy in HUDEP-2 cells (non-treated, NT or overexpressing COUP-TFII, OE, 72 h after transduction), as described by Zambanini and colleagues (2022). 5 $\times 10^5$ cells were analyzed for each sample. To perform the library construction for DNA, the KAPA Hyper Prep Kit for Illumina platforms (KAPA Biosystems, KK8504) was used, according to the manufacturer's guidelines with some modifications. Libraries were run on an E-Gel EX 2% agarose gel (Invitrogen, G402022) for 10 min using the E-Gel Power Snap Electrophoresis System (Invitrogen). Bands of interest between 150 and 500 bp were cut out and purified using the QIAquick Gel Extraction Kit (QIAGEN, 28706) according to manufacturer's instructions. Next, libraries were quantified with the

Qubit (Thermo Scientific) using their high sensitivity DNA kit (Thermo Scientific, Q32854), pooled and analyzed by 36 bp paired-end sequencing on the NextSeq 550 (Illumina) using the Illumina NextSeq 500/550 High Output Kit v2.5 (75 cycles) (Illumina, 20024906). Trimming was performed using bbmap bbduk (Bushnell, 2017, version 39.0) removing adapters, artifacts, poly AT, G and C repeats. Reads were aligned to the hg38 genome with bowtie (Langmead, 2009, version 1.3.1) using options -v 0 -m 1 -X 500. Reads that could not be uniquely mapped to the human genome were removed by SAMtools (Li, 2009, version 1.6). Peaks were called using SEACR (MEERS et al, version 1.3) against the corresponding negative control. Motif analysis and peak set gene annotation were performed. The antibodies used are listed in the table below.

Table 4.3. ***Antibodies used for the ChIP-seq and CUT&RUN assays.***

Antibody	Catalog number	Manufacturer
Anti-COUP-TFII (mouse/human)	A10251	Abclonal
Anti-COUP-TFII	ABIN6928040	

CRISPR/Cas9 screening and competition assay

Cloning of the guide RNAs (gRNAs), lentiviral production and transduction were performed as described (Proietti, 2022). Briefly, the gRNAs targeting the genes of interest were selected by using VBC

Score (https://www.vbc-score.org/). High targeting efficiency and low off-target effects are indicated by a high VBC score. All the gRNAs are listed below (table 4.4). Guide RNAs were cloned into a lentiviral expression vector containing iRFP670 (pLenti hU6-sgRNA-IT-PGK-iRFP670), on BsmBI sites. For lentiviral production, Lenti-X cells were transfected with indicated constructs together with psPAX2 and pMD2.G packaging plasmids (Addgene, 12260 and 12259) using polyethyleneimine (PEI, Polysciences Europe, 24765-1). After 24 h, Lenti-X cell medium was substituted with the target cell medium. The supernatant containing lentiviral pseudo-particles was collected 48 h after transfection. Next, approximately 1 $x10^6$ of K562 cells stably expressing Cas9 (K562-Cas9) were seeded per well of a 24-well plate and transduced by spinfection (1000 x g, 90 min) with lentiviral supernatants (diluted 1:3) supplemented with Polybrene Transfection Reagent (Merck, TR-1003-G). Transduction efficiencies ranged from 10 to 60%. To assess the cell proliferation rate, the percentages of $iRFP670^+$ cells were monitored throughout time (competition assay). The changes in percentages of iRFP670 positive cells were calculated by normalizing the values to the to the first measurement (Day3 post infection, pi) and to a negative control (sgAAVS1).

Table 4.4. ***sgRNAs used for CRISPR/Cas9 screening.***

sgRNA	Target	Sequence
ADAMTSL4	1	AAGGAGAGCCAACTTCCCAATGG
	2	GGATTCCAGCGGGAGCCTTGCGG
ANKRD55	1	GACATGTGTACATATCGCAGCGG
	2	TTGGCCTATGCCCTGTACTGCGG
APOE	1	GCAGGTCATCGGCATCGCGGAGG
	2	ACAGTGTCTGCACCCAGCGCAGG
BAMBI	1	TATAACCAGTGGCTACACAGTGG
	2	AGATTTACACATATAACCAGTGG
CCL26	1	AAGGTGGAGACTCAGGAGGGAGG
	2	GAATTCATAGCTTCGCACCCAGG
CCRL2	1	GAAGATGGCCAATTACACGCTGG
	2	GGTACCTTTGCACAGTCAGAAGG
DYSF	1	TGGGGAGGTAAACCACAGGAAGG
	2	TGTATGGGGAGGTAAACCACAGG
ENC1	1	GCAGCTCCAAGACTTCTGGGTGG
	2	GCAAAGTGTACATTACTGGGGGG
GALM	1	GAAGCCAAGCACCACGTCCGAGG
	2	GAATGCAGTGCCTTGCACTGGGG
PLA2G6	1	TGAACCAGGTGAATAACCAAGGG
	2	ATGAAGGATGAGGTGTTCCGGGG
RBPMS2	1	TGTTGTAGATGACCTGATGGGGG
	2	GATGGCTGGGGTCAGCTCTGTGG
RNASE1	1	GCCGGAATATGACACAGGGGCGG
	2	TACCAGGGGCTCGTGCACAAAGG
COUP-TFII	1	CCCAACCAGCCGACGAGATT
	2	AGGTACGAGTGGCAGTTGAGGGG
RUNX1	1	GCTCCGTGCTGCCTACGCACTGG
	2	GCCATCTGGAACATCCCCTAGGG
AAVS1	1	GCTCCGGAAAGAGCATCCT

RPL17	1	GTACCATTCCGACGTTACAA

Flow cytometry

Cells were washed twice in PBS with 0.1% FBS in a 96-well plate and incubated with CD71 (PE conjugated, Biolegend, 334105) and GpA/CD235ab (FITC conjugated, Biolegend, 306609) antibodies for 30 min, at 4^{o}C, protected from the light. Next, cells were washed in PBS with 0.1% FBS and immediately acquired. Data were recorded on a Cytoflex S (Beckman Coulter) and analyzed with CytExpert 2.5 software (Beckman-Coulter). For the CRISPR/Cas9 screening experiment, the percentages of CD71/GpA expressing cells were calculated by normalizing the values to the first measurement (Day3 post infection, pi) and to a negative non-targeting control sgRNA (sgAAVS1).

Statistical analyses

For statistical analyses was used GraphPad Prism 8.4.3 software (GraphPad Software). Data are expressed as mean ±SEM and all experiments were performed in biological triplicates, unless stated otherwise. To calculate the p-value, the paired Student's t-test was used. P<0.05 was considered to be statistically significant (*p<0.05, **<p<0.01, ***p<0.001, ****p<0.0001, if not stated data not significant).

4.3. Results and Discussion

4.3.1. The overexpression of COUP-TFII in adult-like HUDEP-2 erythroid cells increases the fetal γ-globin levels

The expression of COUP-TFII in mouse fetal liver erythroid cultures (from E13.5 cells) induces the expression of fetal globin genes. Moreover, COUP-TFII positively regulates the expression of fetal γ-globin in different cellular models, not only fetal-like erythroid cells but also adult. In particular, the overexpression of COUP-TFII in erythroid cultures from human peripheral blood progenitors led to elevated γ-globin levels, in both healthy and β-thalassemic erythroid progenitors (Fugazza, 2021). ChIP-seq experiments in fetal-like K562 erythroid cells revealed that COUP-TFII binds to the hypersensitive sites 2, 3 and 4 (HS2, HS3 and HS4) of the β-locus control region (LCR).

In adult erythroid cells, the same regions of the β-locus LCR are bound by BCL11A, an adult erythroid transcription factor and a known γ-globin repressor that shares the same consensus sequence with COUP-TFII (Sankaran, 2008).

Therefore, I asked whether I could map in more detail the COUP-TFII binding to the β-locus in fetal and adult erythroid environment. To this

aim, CUT&RUN experiment was performed in HUDEP-1 and HUDEP-2 erythroid cells to fine map the *in vivo* COUP-TFII binding. HUDEP-2 (human umbilical cord blood-derived erythroid progenitors) is an adult-like erythroid cell line expressing BCL11A together with adult globin genes. HUDEP-1 fetal-like cell line expressing fetal globin genes. Both HUDEP cell lines do not express COUP-TFII. Phenotype of the cell lines used here are summarized in the table 4.5. HUDEP-1 and HUDEP-2 cells were grown in the proliferation cell medium that does not induced their erythroid differentiation (Kurita, 2013).

Table 4.5. ***The summary of the erythroid cell lines used in this study.***

Cells	COUP-TFII expression	γ-globin expression	β-globin expression	Phenotype
K562	YES	YES	NO	*FETAL-LIKE*
HUDEP-1	NO	YES	NO	*FETAL-LIKE*
HUDEP-2	NO	NO	YES	*ADULT-LIKE*

HUDEP-1 and HUDEP-2 cells were transduced with a COUP-TFII overexpressing lentiviral vector (COUP-TFII-IRES-eGFP) or with the corresponding empty vector (EV), as a control. The efficiency of transduction was monitored by flow cytometry (FC) analysis (Fig. 4.1A) and the expression of exogenous COUP-TFII was confirmed by RT-qPCR analysis (Fig. 4.1B).

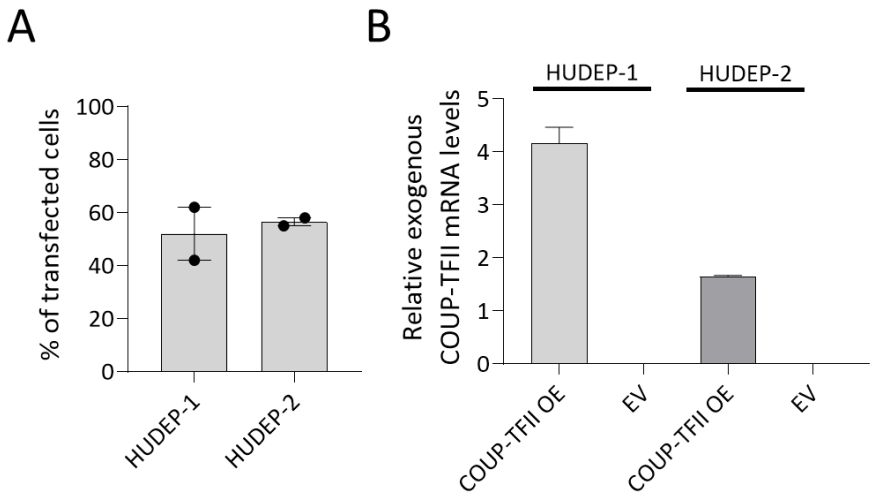


Figure 4.1. ***Efficiency of the lentiviral overexpression of COUP-TFII in HUDEP-1 and HUDEP-2 cells. A*** *Percentage of GFP⁺ cells was measured by flow cytometry 72 h post infection (pi) with COUP-TFII overexpressing lentiviral vector (COUP-TFII OE) and control empty vector (EV), both expressing GFP. Data as mean ± SEM, n≥2.* ***B*** *Expression level of the exogenous (mouse) COUP-TFII (72 h pi) measured by RT-qPCR. Data as mean ± SEM, n≥2.*

First, we analyzed the effect of the COUP-TFII overexpression (OE) on γ-globin: RT-qPCR analysis 72 h after transduction, showed a 2-fold (however statistically not significant) increase in the γ-globin level upon COUP-TFII OE when compared to control (Fig. 4.2). In HUDEP-2 cells, COUP-TFII OE significantly increased γ-globin levels (15-fold increase), when compared to the control (EV) (Fig. 4.2).

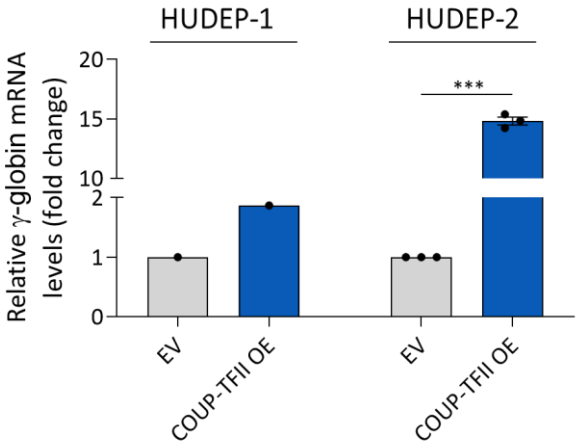


Figure 4.2. ***COUP-TFII overexpression in fetal-like HUDEP-1 and adult-like HUDEP-2 erythroid cells results in the re-activation of γ-globin expression.*** *Level of γ-globin mRNA upon exogenous expression of COUP-TFII (COUP-TFII OE, 72 h after transduction) in HUDEP-1 and HUDEP-2 cells. Cells transduced with the empty vector (EV) were used as a control. Data as mean ± SEM, n≥3, ***=p<0.001.*

After I confirmed the ability of COUP-TFII to overcome the adult environment of HUDEP-2 cells and to reactivate the γ-globin expression in both HUDEP-2 and HUDEP-1 cells, a CUT&RUN experiment was performed to map the COUP-TFII binding within the β-locus in collaboration. COUP-TFII binding profiles in HUDEP-1 and HUDEP-2 cells expressing exogenous COUP-TFII were compared with control non-overexpressing HUDEP-2 cells (when COUP-TFII was absent). As shown in figure 4.3, in COUP-TFII OE HUDEP-2 cells, COUP-TFII is bound to the HS1 HS2, HS3 and HS4 of the LCR, showing the

same pattern as in K562 OE cells. In HUDEP-1 COUP-TFII OE cells, COUP-TFII is bound to the HS2 and HS3, but to less extent to HS1 and HS4 (Fig. 4.3). Taken together, the CUT&RUN experiment demonstrate that the fetal γ-globin induction in adult erythroid cells by COUP-TFII is due to its direct binding to the β-locus LCR, therefore the mechanism is likely the same as in fetal-like K562 cells spontaneously expressing COUP-TFII. Of interest, differences in binding of COUP-TFII to HS sites of the β-locus LCR between HUDEP-2 and HUDEP-1 cells, suggest that COUP-TFII binding is, at least in part, influenced by the cellular context.

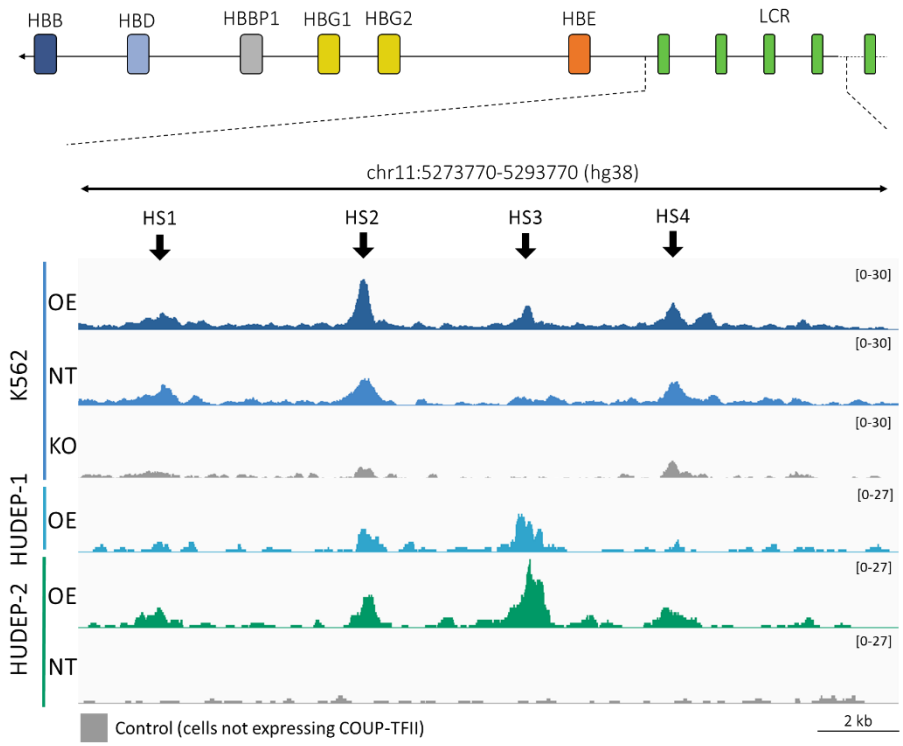


Figure 4.3. ***COUP-TFII activates the expression of γ-globin by direct binding to locus control region (LCR) of the β-locus.*** *COUP-TFII ChIP-seq profiles in K562 cells: overexpressing COUP-TFII (OE, dark blue) and non-treated (NT, blue). In light blue and green, COUP-TFII CUT&RUN profiles in HUDEP-1 and HUDEP-2 cells overexpressing COUP-TFII (OE), respectively. K562 cells with COUP-TFII knockout (KO) and HUDEP-2 non-treated (NT), both in grey, were used as controls. Black arrows indicate hypersensitive sites of the β-locus LCR (HS1-HS4).*

Moreover, as COUP-TFII and BCL11A and bind to the same DNA consensus sequences, it is possible that these two regulators of globin gene expression compete for the binding within the β-locus. Indeed,

the comparison of the COUP-TFII CUT&RUN profile with the BCL11A binding profile (Liu, 2018) demonstrated that COUP-TFII and BCL11A bind to the same HS sites (Fig. 4.4). This observation supports the idea of the competition between COUP-TFII and BCL11A in binding the β-locus LCR. Interestingly, exogenous COUP-TFII in HUDEP-2 cells shows the strongest peak of binding in the HS3, the site that has the most fetal characteristic as it was demonstrated to be independent of the adult erythroid transcription factor KLF (Jackson, 2003).

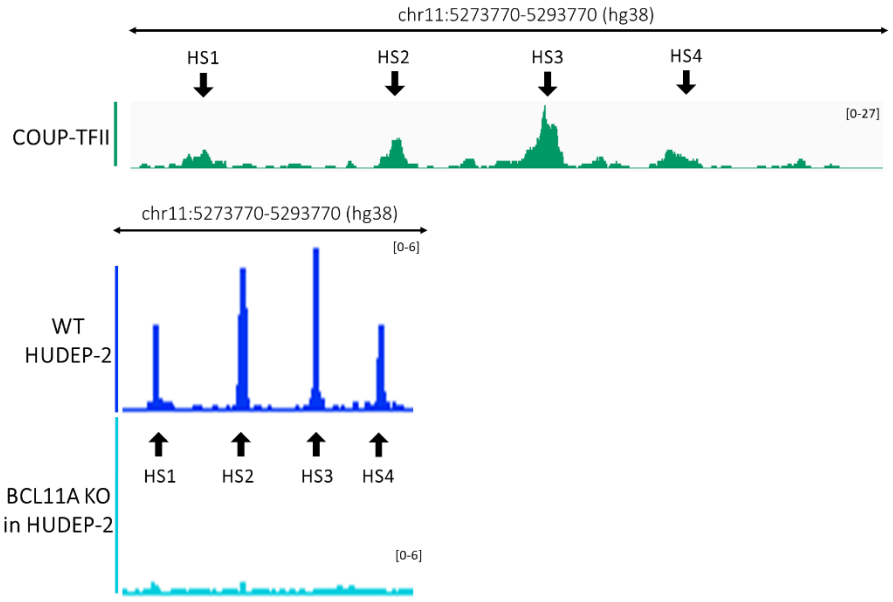


Figure 4.4. ***COUP-TFII and BCL11A bind the same regions of the β-locus control region (LCR).*** *COUP-TFII CUT&RUN profile in HUDEP-2 cells (green), overexpressing COUP-TFII. In blue, the BCL11A binding profile in HUDEP-2 cells, WT and BCL11A knockout as a control (Liu, 2018). BCL11A, an adult erythroid transcription factor, is a known repressor of γ-globin gene.*

What is more, COUP-TFII binding profiles of the entire β-locus revealed that COUP-TFII bound to β-globin and δ-globin genes only when overexpressed in adult-like HUDEP-2 erythroid cells (Fig. 4.5). This was not observed in fetal-like K562 and HUDEP-1 erythroid cells (Fig. 4.5).

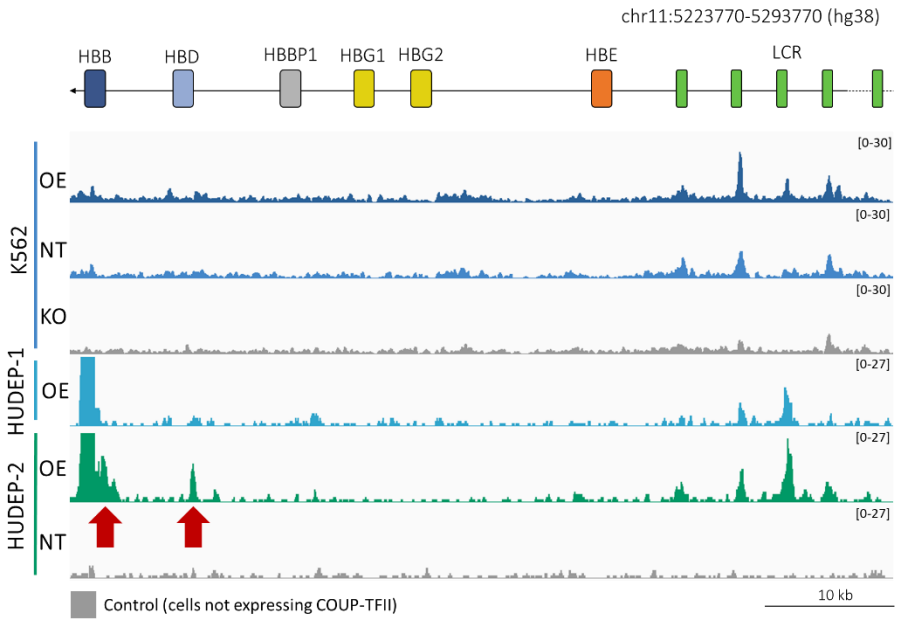


Figure 4.5. ***COUP-TFII binds to the promoter regions of HBB and HBD genes, encoding adult globins in HUDEP-2 cells.*** *COUP-TFII ChIP-seq profiles in K562 cells overexpressing COUP-TFII (OE, dark blue), non-treated (NT, blue) or knocked out for COUP-TFII (KO, grey). CUT&RUN profiles in HUDEP-1 (light blue) and HUDEP-2 cells (green), both overexpressing COUP-TFII (OE) and in non-treated HUDEP-2 cells (NT). Red arrows indicate the COUP-TFII peaks within HBB and HBD genes.*

To conclude, the CUT&RUN experiment in HUDEP-2 cells overexpressing COUP-TFII confirms the hypothesis that COUP-TFII is able to bind to the β-locus, despite the repressive adult environment. The overlap between the BCL11A and COUP-TFII binding sites suggests the competition between these two erythroid transcription factors for the binding the same regions of the β-locus control region.

4.3.2. MYB, a transcription factor important for γ-globin regulation is a potential direct target of COUP-TFII

MYB is a transcription factor essential for the control of proliferation/differentiation balance in erythroid cells and a known γ-globin regulator (Jiang, 2006; Sankaran, 2011; Gu, 2022). Common genetic variants within the intergenic region HBS1L-MYB region are associated with elevated level of fetal hemoglobin (HbF) (Thein, 2007). HBS1L-MYB intergenic polymorphisms affect regulatory elements that interact with MYB (Stadhouders, 2014). These variants were found to reduce the binding of the key erythroid transcription factor in the regulatory elements of MYB, and therefore to affect the expression level of MYB. Interestingly, ChIP-seq experiment performed in K562 cells revealed that COUP-TFII binds within the HBS1L-MYB intergenic region (Fig. 4.6).

I therefore analyzed the H3K27ac and ATAC ChIP-seq profiles of the regions bound by COUP-TFII (Fig. 4.6). Regions called P1 and P4 are acetylated in K562 cells, show an open chromatin feature, and the COUP-TFII binding motifs within these regions are conserved. What is more, P1 corresponds to the region identified as the -71 kb MYB enhancer element by Stadhouders and colleagues (2014). Regions P2 and P3 instead do not show any of the features that characterize active regulatory sequences.

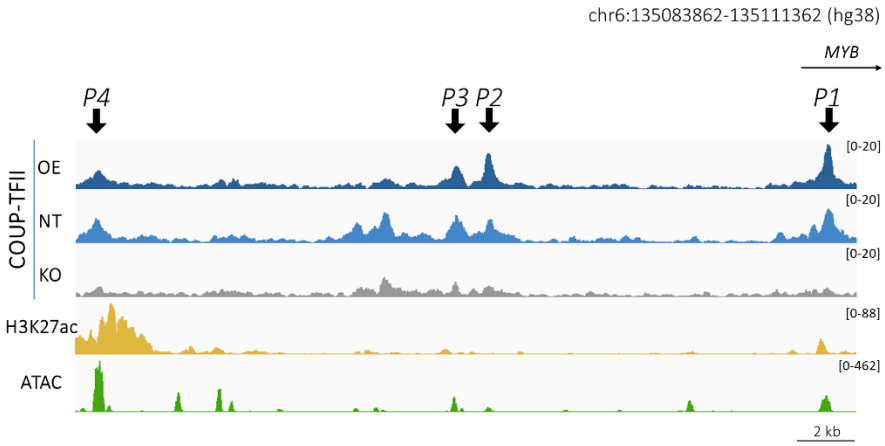


Figure 4.6. ***COUP-TFII directly binds to the HBS1L-MYB intergenic region.*** *COUP-TFII ChIP-seq profiles in K562 cells overexpressing COUP-TFII (OE, dark blue), non-treated (NT, blue) and with COUP-TFII knockout (KO, grey). Black arrows indicate positions of the COUP-TFII peaks. In yellow: H3K27ac ChIP-seq (GEO: ENCFF465GBD) and in green: ATAC-seq (GEO: ENCFF357GNC) profiles in K562 cells.*

To assess to what extent COUP-TFII contributes to the MYB transcription, I checked the levels of MYB mRNA in K562 erythroid cells overexpressing COUP-TFII (OE), knocked out for COUP-TFII (KO), or control cells transfected with the empty vector (EV). The RT-qPCR analysis revealed that in cells overexpressing COUP-TFII for 72 h, MYB levels were significantly decreased, when compared to EV control, suggesting a potential repressive role (Fig. 4.7A). On the other hand, the long-term (25 days after transduction) effect of COUP-TFII overexpression was the opposite to the short-term one (Fig. 4.7B),

possibly because of downstream indirect effects. Instead, the depletion of COUP-TFII does not affect the MYB expression level neither in short-term nor long-term conditions (Fig. 4.7A, 4.7B).

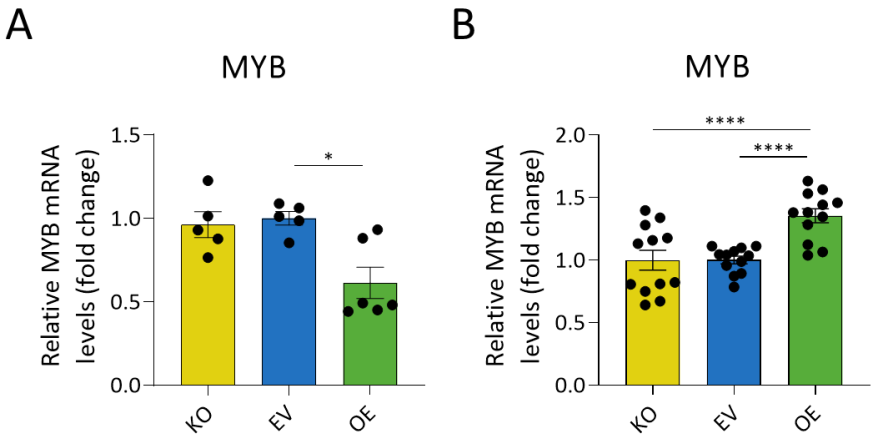


Figure 4.7. ***The opposite effect of short-term and long-term COUP-TFII overexpression on the expression of MYB in erythroid K562 cells. A*** *Level of MYB mRNA after 96 h of COUP-TFII knockout (KO) and 72 h of COUP-TFII overexpression (OE).* ***B*** *The expression level of MYB after 30 days of COUP-TFII KO and 25 days of COUP-TFII OE. Analyzed by RT-qPCR and normalized to Empty vector (EV) control. Data as mean ± SEM, n≥6, *=p<0.05, ****=p<0.0001.*

Therefore, I selected P1 and P4 regions, located 71 kb and 97 kb, respectively, upstream to the MYB gene, to further analyzed the potential of COUP-TFII to regulate MYB expression by binding to these regions.

Sequences of P1 and P4 were cloned in a reporter vector upstream to the minimal TK promoter and their regulatory potential was tested in K562 cells by luciferase assay. P1 and P4 reporter constructs were transfected in K562 cells with different level of COUP-TFII (OE, EV and KO) and their effect on the reporter gene expression was measured 48 h after transfection (Fig. 4.8). As a positive control, I used the reporter vector containing the GATA1 enhancer upstream to the minimal TK promoter, and the vector carrying only the minimal TK promoter as a negative control.

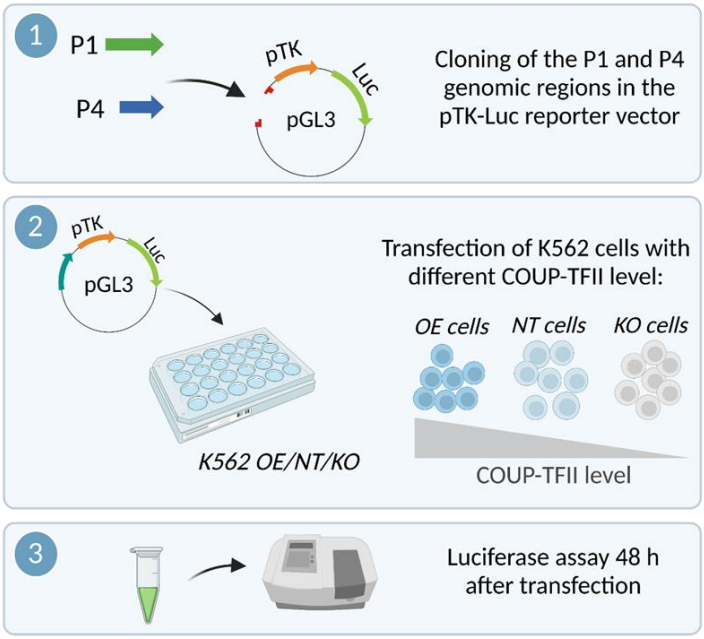


Figure 4.8. ***Transfection experiment with reporter constructs containing P1 and P4 regions to define their involvement in the regulation of MYB expression by COUP-TFII.***

As shown in figure 4.9, both P1 and P4 regions can act as activators in K562 erythroid cells. However, they increased the luciferase gene expression about 3-fold irrespectively of COUP-TFII OE or KO, thus the regulatory potential of P1 and P4 seem to not depend on the COUP-TFII level.

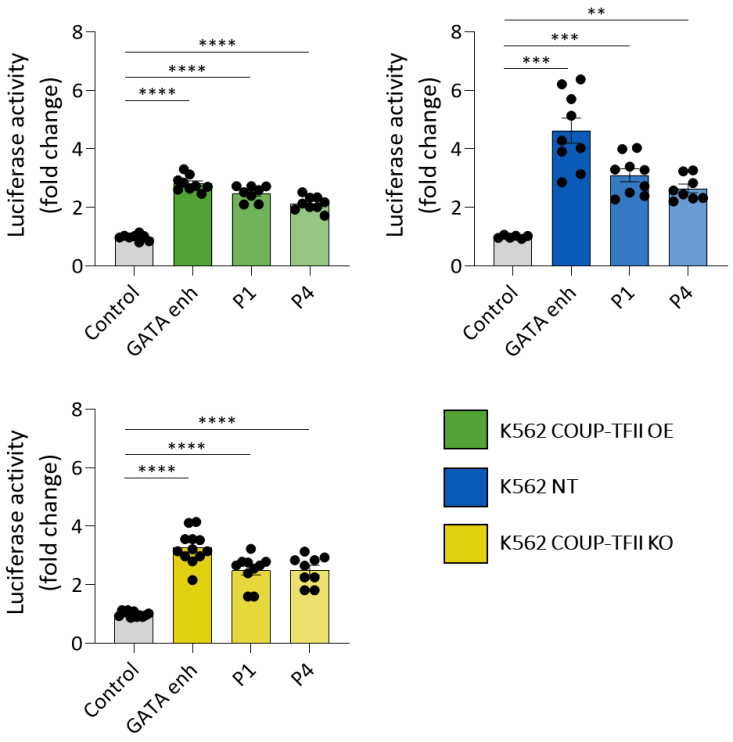


Figure 4.9. ***Both P1 and P4 regions upregulates the reporter gene expression, independently from the level of COUP-TFII.*** *Luciferase assay experiment performed in K562 cells overexpressing COUP-TFII (OE) or with COUP-TFII knockout (KO). NT – non-treated K562 cells, Control – pTK vector was used as a negative control of transfection, GATA enh – pTK vector with GATA enhancer was used as a positive control. Data as mean ± SEM, n≥3, **=p<0.01, ***=p<0.001, ****=p<0.0001.*

As demonstrated by RT-qPCR, elevated COUP-TFII levels show an opposite effect on the MYB mRNA level in short-term and long-term conditions. Therefore, to confirm the previous observation that P1 and P4 act as enhancer independently from the level of COUP-TFII, I co-transfected the P1 or P4 reporter construct together with a vector expressing COUP-TFII (pCAG-COUP-TFII expression vector). This allowed to check more direct effect of COUP-TFII on P1 and P4 regions' activity, in a dose-dependent manner. K562 cells were transfected with P1 or P4 reporter vectors together with increasing amount of the COUP-TFII expressing vector, and the reporter gene expression was analyzed after 48 h as in the previous experiment (Fig. 4.8). Again, I observed no significant differences in the enhancer potential of P1 and P4 regions between cells with different COUP-TFII level (Fig. 4.10).

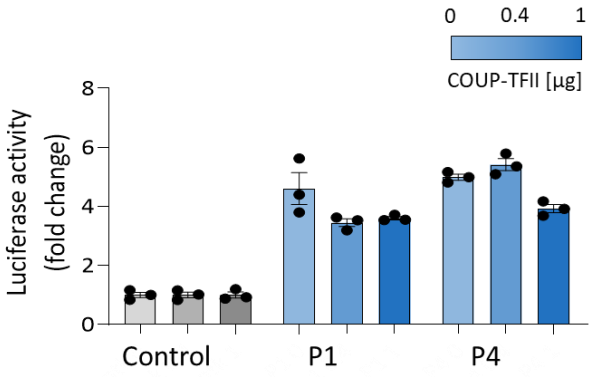


Figure 4.10. ***Co-transfection of the COUP-TFII overexpressing vector and P1 or P4 confirms that their regulatory potential do not depend on COUP-TFII.*** *Luciferase assay experiment performed in K562 cells co-transfected with a vector overexpressing COUP-TFII (pCAG-COUP-TFII) and with P1 or P4 reporter constructs. Control – pTK vector. Data as mean ± SEM, n≥1.*

To conclude, COUP-TFII binds *in vivo* within the HBS1L-MYB intergenic region, but our analysis on the effect of COUP-TFII levels on P1 and P4 regulatory regions does not show a direct relationship. However, COUP-TFII short-term overexpression affects the expression level of MYB suggesting that sequences other than P1 and P4 could mediate a repressive effect of COUP-TFII on MYB. Of note, COUP-TFII is an activator of γ-globin expressed at early developmental stages, while in contrast MYB represses γ-globin in adult erythroid cells. Both COUP-TFII and MYB are expressed in the first, pro-definitive wave of hematopoiesis that consists erythro-myeloid progenitors (EMPs) in mouse fetal liver (Fugazza, 2021). EMPs represent a rare and transient cell population, and therefore it is not completely clear if there is a spatial and temporal overlap in the expression of COUP-TFII and MYB in these cell. Hence, if there is a direct regulatory mechanism of COUP-TFII on MYB, most probably is transient and it would be difficult to identify *in vivo*.

4.3.3. Identification of novel candidate direct target genes of COUP-TFII in K562 erythroid cells

As described above, COUP-TFII regulates the expression of globin genes by directly binding to the β-locus, in both fetal-like and adult-like erythroid cells. However, it remains unknown what other genes are controlled by COUP-TFII in the erythroid cells. To better characterize the downstream network of COUP-TFII within the erythroid lineage, I sought to identify genes that are regulated by COUP-TFII in K562 erythroid cells. In collaboration with R. Piazza and L. Mologni (UNIMIB, Milan, Italy), we performed the integrative analysis of RNA-sequencing (RNA-seq) and ChIP-sequencing (ChIP-seq) datasets from K562 cells expressing COUP-TFII at different levels. To this end, we exploited data from K562 cells expressing endogenous COUP-TFII, overexpressing COUP-TFII (by lentiviral overexpression, OE) or non-expressing COUP-TFII (CRISPR/Cas9-mediated knockout, KO) (Fig. 4.11).

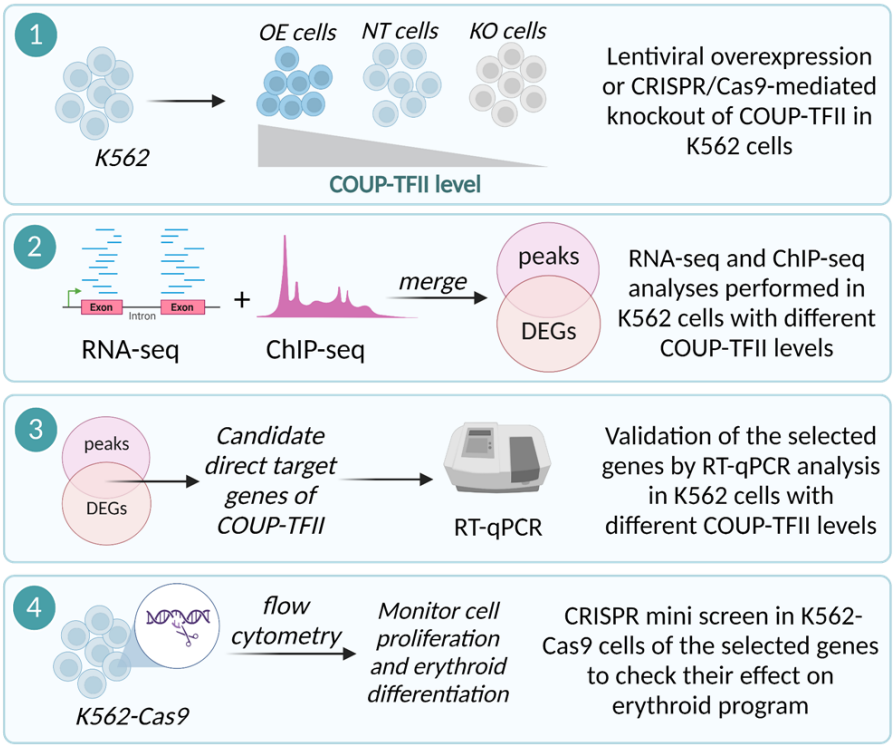


Figure 4.11. ***Experimental workflow for the identification of direct target genes of COUP-TFII in K562 erythroid cells.***

Genes that were differentially regulated in OE vs. EV and KO vs. EV cells were divided into two lists: genes upregulated and downregulated by COUP-TFII. RNA-seq analysis showed the upregulation of 178 genes ($p_{adj} < 0.01$, fold change > 2) in cells overexpressing COUP-TFII when compared to control EV cells, while only 23 genes were strongly downregulated (Fig. 4.12A, left). COUP-TFII knockout caused the upregulation of 199 genes, and downregulation of 217 genes ($p_{adj} < 0.01$, fold change > 2) (Fig. 4.12A, right).

Analysis of ChIP-seq data identified 1783 COUP-TFII peaks common to the OE and NT K562 cells. COUP-TFII *in vivo* binding sites were associated with genes when mapped 500 kb upstream or downstream to their transcription start sites. Background peaks present in COUP-TFII KO cells were subtracted. Among genes associated with the genomic regions bound by COUP-TFII, 66 genes were also differentially regulated in the RNA-seq, and therefore considered as candidate COUP-TFII direct targets, of which 33 were upregulated and 33 were downregulated by COUP-TFII (Fig. 4.12B and table 4.6).

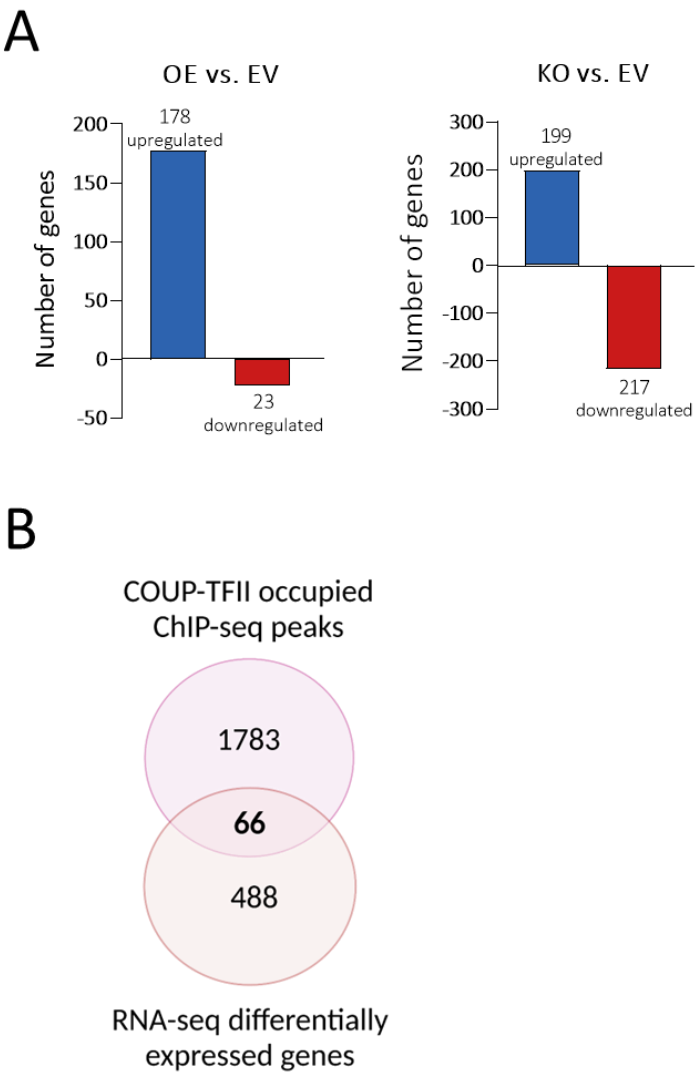


Figure 4.12. **The impact of COUP-TFII on erythroid transcription.** ***A*** *Number of differentially regulated genes between K562 cells overexpressing COUP-TFII (OE) and control (EV) cells (left), right: K562 cells with COUP-TFII knockout (KO) and control (EV) cells.* ***B*** *Integrative analysis of RNA-seq and ChIP-seq identified 66 genes as potential direct targets of COUP-TFII.*

Table 4.6. ***Genes selected as COUP-TFII candidate direct target genes.*** *Upregulated and downregulated by COUP-TFII genes were identified by RNA-seq analysis. Data as Log2Fold change OE vs. EV and KO vs. EV, p_{adj}<0.1. The most differentially expressed genes selected for further analysis are highlighted in blue.*

GENES UPREGULATED BY COUP-TFII

refseqID	Symbol	Description	Fold change KO vs. EV	Fold change OE vs. EV
NM_024669	ANKRD55	Ankyrin Repeat Domain 55	-6,32	1,64
NM_004273	CHST3	Carbohydrate Sulfotransferase 3	-4,57	2,78
NM_005045	RELN	Reelin	-4,09	0,55
NM_022127	SLC28A3	Solute Carrier Family 28 Member 3	-3,68	1,80
NM_001256574	ENC1	Ectodermal-Neural Cortex 1	-3,58	0,38
NM_005723	TSPAN5	Tetraspanin 5	-3,28	0,96
NM_002763	PROX1	Prospero Homeobox 1	-3,26	1,34
NM_033058	TRIM55	Tripartite Motif Containing 55	-2,96	4,82
NM_012342	BAMBI	BMP And Activin Membrane Bound Inhibitor	-2,92	1,69
NM_007249	KLF12	Krupper Like Transcription Factor 12	-2,51	0,91
NM_014293	NPTXR	Neuronal Pentraxin Receptor	-2,48	3,37
NM_031310	PLVAP	Plasmalemma Vesicle Associated Protein	-2,26	4,20
NM_003054	SLC18A2	Solute Carrier Family 18 Member A2	-2,20	1,72
NM_001201575	CTSV	Cathepsin V	-2,11	0,80
NM_003127	SPTAN1	Spectrin Alpha, Non-Erythrocytic 1	-1,93	1,92
NM_003856	IL1RL1	Interleukin 1 Receptor Like 1	-1,83	-0,76
NM_016540	GPR83	G Protein-Coupled Receptor 83	-1,82	2,66
NM_001190986	CRISP3	Cysteine Rich Secretory Protein 3	-1,56	1,61
NR_038357	ZSWIM8-AS1	ZSWIM8 Antisense RNA 1	-1,54	-0,35
NM_017793	RPP25	Ribonuclease P And MRP Subunit P25	-1,44	0,89
NM_198682	GYPE	Glycophorin E (MNS Blood Group)	-1,15	0,31
NM_003940	USP13	Ubiquitin Specific Peptidase 13	-1,13	0,95
NR_034005	LINC00891	Long Intergenic Non-Protein Coding RNA 891	-1,03	0,64
NM_001256026	ARHGAP22	Rho GTPase Activating Protein 22	-0,83	3,31
NM_001752	CAT	Catalase	-0,75	1,23
NM_015184	PLCL2	Phospholipase C Like 2	-0,72	0,37
NM_014616	ATP11B	ATPase Phospholipid Transporting 11B (Putative)	-0,64	0,09
NM_001288706	IL1R1	Interleukin 1 Receptor Type 1	-0,62	0,30
NM_006763	BTG2	BTG Anti-Proliferation Factor 2	-0,57	1,01
NM_014568	GALNT5	Polypeptide N-Acetylgalactosaminyltransferase 5	-0,49	1,07
NM_001282667	MICAL2	Protein-Methionine Sulfoxide Oxidase MICAL2	-0,35	0,65
NM_001243782	PRXL2A	Peroxiredoxin Like 2A	0,05	5,13
NM_024901	DENND2D	DENN Domain Containing 2D	0,40	5,80

GENES DOWNREGULATED BY COUP-TFII

refseqID	Symbol	Description	Fold change KO vs. EV	Fold change OE vs. EV
NM_016633	AHSP	Alpha Hemoglobin Stabilizing Protein	4,62	-0,70
NM_194272	RBPMS2	RNA Binding Protein, MRNA Processing Factor 2	3,33	-1,88
NM_024577	SH3TC2	SH3 Domain And Tetratricopeptide Repeats 2	2,28	-1,73
NM_001302690	APOE	Apolipoprotein E	2,20	-0,39
NM_005332	HBZ	Hemoglobin Subunit Zeta	2,19	-0,46
NM_006072	CCL26	C-C Motif Chemokine Ligand 26	2,08	-1,44
NM_001130976	DYSF	Dysferlin	2,00	-1,96
NM_138801	GALM	Galactose Mutarotase	1,91	-1,39
NM_001003841	SLC6A19	Solute Carrier Family 6 Member 19	1,84	-0,49
NM_198232	RNASE1	Ribonuclease A Family Member 1, Pancreatic	1,83	-1,22
NM_000420	KEL	Kell Metallo-Endopeptidase (Kell Blood Group)	1,64	-1,07
NM_003965	CCRL2	C-C Motif Chemokine Receptor Like 2	1,55	-2,10
NM_001645	APOC1	Apolipoprotein C1	1,54	-0,57
NM_003560	PLA2G6	Phospholipase A2 Group VI	1,45	-0,75
NM_001288608	ADAMTSL4	ADAMTS Like 4	1,27	-0,55
NR_015379	UCA1	Urothelial Cancer Associated 1	1,26	-1,22
NM_001161354	PLEKHA4	Pleckstrin Homology Domain Containing A4	1,23	1,23
NM_001927	DES	Desmin	1,13	-1,54
NM_004568	SERPINB6	Serpin (serine proteinase inhibitor) Family B Member 6	1,11	-1,04
NM_001303050	MAPRE3	Microtubule Associated Protein RP/EB Family Member 3	1,05	-0,81
NM_173175	PTK2B	Protein Tyrosine Kinase 2 Beta	0,98	-0,70
NM_030651	PRRT1	Proline Rich Transmembrane Protein 1	0,94	-1,36
NM_004458	ACSL4	Acyl-CoA Synthetase Long Chain Family Member 4	0,87	-0,69
NM_000676	ADORA2B	Adenosine A2b Receptor	0,80	-0,81
NM_006163	NFE2	Nuclear Factor, Erythroid 2	0,72	-0,09
NM_080752	ZSWIM3	Zinc Finger SWIM-Type Containing 3	0,56	-0,77
NM_138717	PPT2	Palmitoyl-Protein Thioesterase 2	0,50	-0,94
NM_004613	TGM2	Transglutaminase 2	0,42	-1,33
NM_003670	BHLHE40	Basic Helix-Loop-Helix Family Member E40	0,42	-0,97
NM_005469	ACOT8	Acyl-CoA Thioesterase 8	0,40	-0,39
NM_018660	ZNF395	Zinc Finger Protein 395	0,34	-0,27
NM_001256584	GYPC	Glycophorin C (Gerbich Blood Group)	0,30	-0,28
NM_022110	FKBPL	FKBP Prolyl Isomerase Like	0,25	-0,51

Most of the listed candidate direct target genes of COUP-TFII (Table 4.6) have unknown role in erythropoiesis, however there are some exceptions:

- Alpha Hemoglobin Stabilizing Protein (AHSP), which encodes molecular chaperone that binds to monomeric α-hemoglobin before it is transferred to β-hemoglobin to form a heterodimer.

AHSP is essential for normal erythroid cell development, by acting to prevent the harmful aggregation of free α-hemoglobin (Kihm, 2002). The depletion of AHSP results in an increased apoptosis of erythroid precursors, elevated ROS, and finally in the worsened β-thalassemia phenotype, a disease characterized by excess of free α-hemoglobin (Kong, 2004).

- NFE2 (Nuclear Factor, Erythroid derived 2) is critical for erythroid maturation and globin gene expression (Andrews, 1994; Mutschler, 2009), and interestingly it might be involved in the transcriptional regulation of COUP-TFII by binding to its distal enhancer element (chapter 2).
- GYPC encoding glycophorin C, an integral membrane glycoprotein that plays an important role in regulating the mechanical stability of red blood cells (Lopez, 2021).
- GYPE which encodes glycophorin E, a sialoglycoprotein belonging to a gene family with GPA and GPB genes that might carry the M blood group antigen (Hollox and Louzada, 2022).
- HBZ encoding embryonic α-like ζ-globin.

To narrow down the list of the candidate COUP-TFII direct targets, I selected the most differentially expressed genes in COUP-TFII KO cells when compared to EV cells and I validated the changes in their expression upon COUP-TFII OE or KO in K562 cells by RT-qPCR. Only genes that showed a perfect profile of up- or downregulated target (significantly increased or reduced mRNA levels) were selected for further functional analysis, excluding those with marginal expression

level in K562 cells (for example PROX1) or with already known effect on erythropoiesis (like AHSP) (Table 4.6).

Next, I aimed to investigate whether the selected 12 prospective direct target genes identified by integrative RNA-seq and ChIP-seq analysis (table 4.6) are of functional importance on erythroid program. In collaboration with L. Proietti and F. Grebien (Vetmeduni, Vienna, Austria), I performed CRISPR/Cas-mediated knockout of these genes in K562 cells stably expressing Cas9.

I used the VBC Score tool (https://www.vbc-score.org/) to select the gRNAs with high targeting efficiency and low off-target effects. All 12 candidate COUP-TFII target genes were targeted with two different sgRNAs. Competing growth kinetics of sgRNA-expressing iRFP670-positive ($iRFP670^{+}$) cells vs. uninfected iRFP670-negative ($iRFP670^{-}$) K562-Cas9 cells was monitored by a competition assay. The efficiency of transduction for all sgRNAs was around 50%. Changes in the percentages of iRFP670 expressing cells were measured for 20 days post infection (pi) (Fig. 4.13). Data was normalized to non-targeting sgRNA control (sgAAVS1) and the first timepoint (day 3 post infection, pi). sgRPL17, targeting gene essential to cell proliferation, results in a decrease of the percentage of iRFP670 positive cells over time and was used as a positive control. Effects of the gene knockouts were considered significant only when the same effect was observed with two different sgRNAs.

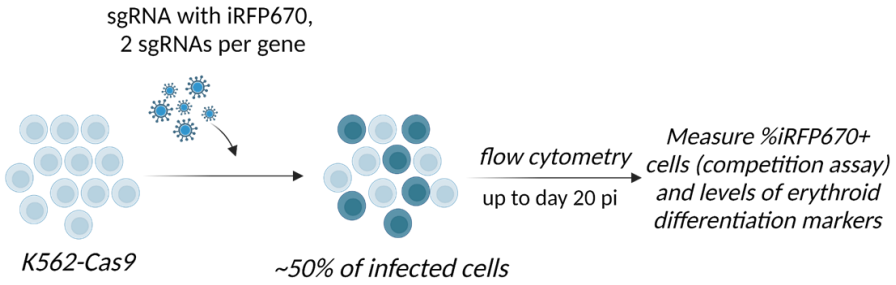


Figure 4.13. ***Schematic representation of workflow of CRISPR mini screen.*** *Each gene was targeted with 2 sgRNAs in Cas9-expressing K562 clone. Percentage of iRFP670⁺ cells, as well the levels of CD71 and GpA erythroid markers in iRFP670⁺ cells were monitored over time.*

As shown in figure 4.14, in the competition assay both ENC1-targeting sgRNAs reduced K562-Cas9 cell proliferation (40% reduction). Similar anti-proliferative effects were observed for two APOE-targeting sgRNAs (30% reduction) and BAMBI-targeting sgRNAs (20% reduction). Both ENC1 and BAMBI were upregulated by COUP-TFII and their elevated levels were reported in several cancers and positively regulate cell proliferation (Cui, 2021; Onichtchouk, 1999; Sekiya, 2004; Togo, 2008; Fritzmann, 2009). Therefore, reduced cell proliferation of K562-Cas9 cells upon ENC1 and BAMBI depletion is in line with the studies in other cellular models. On the contrary, apolipoprotein E, encoded by APOE, identified by Murphy and colleagues (2011) as a negative regulator of proliferation in mouse hematopoietic stem and multipotential progenitor cells (HSPCs), in K562 cells shows the opposite effect. Interestingly, the knockout of COUP-TFII itself

significantly induced proliferation of K562-Cas9 cells (40% increase) (Fig. 4.14). The induction of K562 cell proliferation was not observed in the previous experiment (see chapter 3). It is plausible that the effect of COUP-TFII on proliferation is diverse in different K562 erythroleukemic cell clones, and therefore it would be of interest to confirm this effect in more physiological erythroid cellular model. Additionally, the CRISPR mini screen was performed in K562 clone stably expressing Cas9 that could have some indirect impact on cell proliferation.

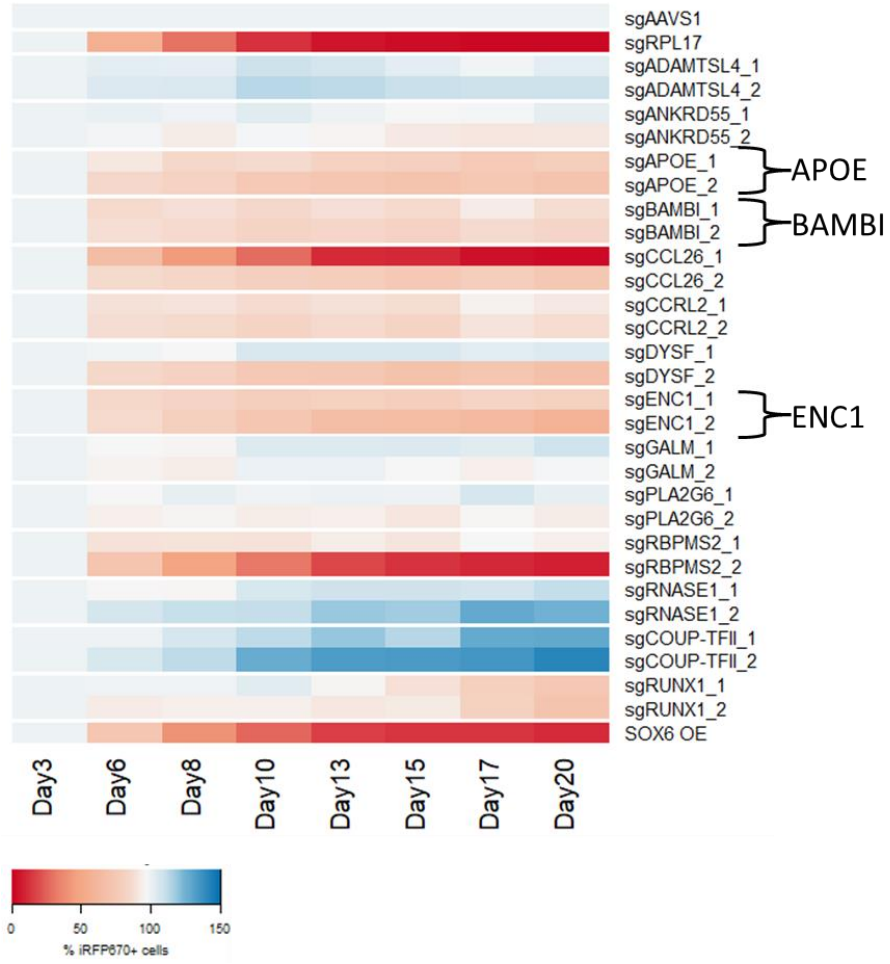


Figure 4.14. ***CRISPR/Cas9 competition assay of selected candidate direct targets of COUP-TFII.*** *The effect on cell proliferation of each sgRNA was assessed, sgAAVS1 - negative control, sgRPL17 - positive control of cell proliferation, sgRUNX1 and SOX6 OE – positive controls of erythroid differentiation. Data as %iRFP670^{+} cells$_{normAAVS1+d3}$, normalized to sgAAVS1, and day 3, n=3.*

To explore the effects of candidate COUP-TFII target genes on erythroid differentiation, I analyzed surface expression of erythroid differentiation markers glycophorin A (GpA) and CD71 (transferrin receptor). Changes in CD71/GpA profile were monitored over time in parallel with the competition assay by flow cytometry. CD71 is an early erythroid marker, whereas glycophorin A (GpA) is a late erythroid marker. As a control, I included two RUNX1-targeting sgRNAs and lentiviral overexpression of SOX6 (SOX6 OE). RUNX1 transcription factor was previously described as a negative regulator of both CD71 and GpA surface expression in K562 cells (Kuvardina, 2015). The overexpression of SOX6 in K562 results in an enhanced erythroid differentiation and hemoglobinization of K562 cells (Cantù, 2011). CD71 and GpA expression levels were measured every two days up to day 10 after transduction, as the control SOX6 OE K562 cells died within 9 days after transduction as a consequence of their terminal differentiation (Cantù, 2011) (Fig. 4.15).

Consistent effects on the GpA level with the two sgRNAs was observed for BAMBI, ENC1, GALM and PLA2G6 genes. The depletion of BAMBI, ENC1 and PLA2G6 resulted in an increased level of GpA, whereas the knockout of GALM reduced GpA surface expression (Fig. 4.15). COUP-TFII knockout in K562-Cas9 cell showed a negative effect on GpA level (Fig. 4.15) which is in line with an increased proliferation in the same cells (Fig. 4.14). The positive effect of BAMBI and ENC1 depletions on erythroid differentiation was also consistent with their negative effect on cell proliferation (Fig. 4.14). Loss of GALM and PLA2G6, did not

cause any changes in K562-Cas9 cell proliferation but led to the induction of erythroid differentiation, as shown by increased GpA expression in these cells (Fig. 4.15). Unlike in the competition assay (Fig. 4.14), APOE-targeting sgRNAs did not show a strong effect on GpA levels consistent for both sgRNAs, therefore the effect of APOE on K562 erythroid cell differentiation could not be assessed (Fig. 4.15). APOE was previously suggested to be essential for terminal erythroid maturation ($CD71^{-}GpA^{+}$ cells) (Crespo, 1993; Holm, 2002) therefore not at the stage as the majority of K562-Cas9 cells in this experiment ($CD71^{+}GpA^{+}$). All the other candidate COUP-TFII target genes that did not show a consistent effect on proliferation and/or erythroid differentiation for both gene-targeting sgRNAs and were not further considered.

Taken together, CRISPR/Cas9-mediated targeting of candidate COUP-TFII direct target genes showed only mild changes in K562-Cas9 erythroid cell proliferation and differentiation. This observation is in line with the previous study in our laboratory that reported no major changes in cell proliferation and differentiation upon COUP-TFII overexpression in adult human peripheral blood cells (Fugazza, 2021).

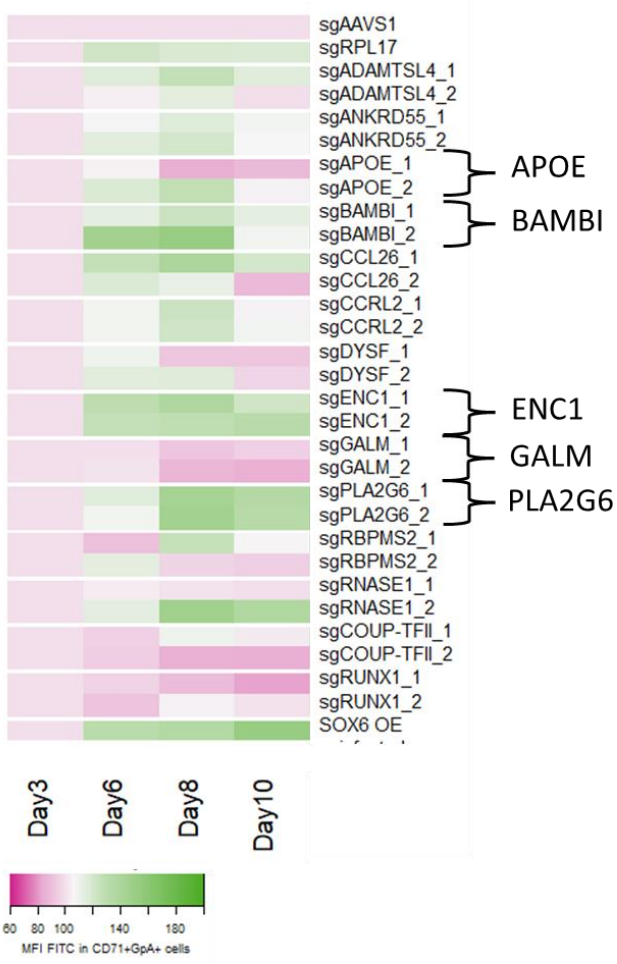
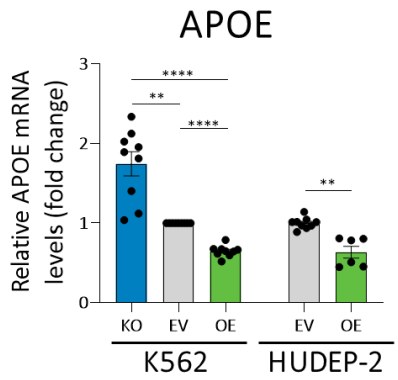
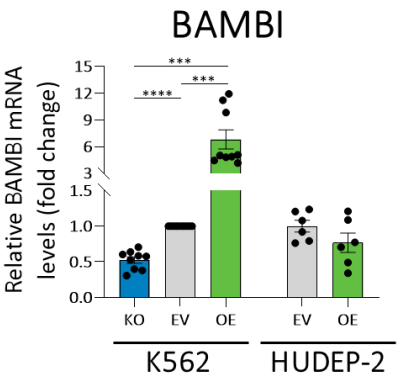


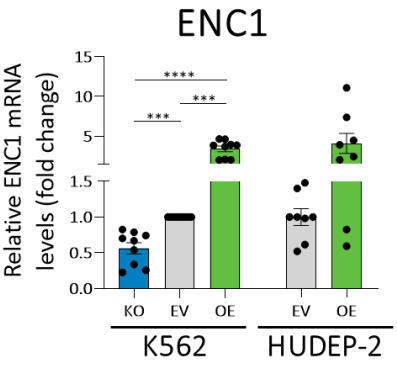
Figure 4.15. ***Changes in GpA levels of CD71⁺GpA⁺ K562-Cas9 cells upon CRISPR/Cas9 targeting of candidate COUP-TFII direct target genes****, sgAAVS1 - negative control, sgRPL17 - positive control of cell proliferation, sgRUNX1 and SOX6 OE – positive controls of erythroid differentiation. Data as MFI FITC-GpA in CD71⁺GpA⁺ cells$_{normAAVS1+d3}$, normalized to sgAAVS1, and day 3, n=3.*

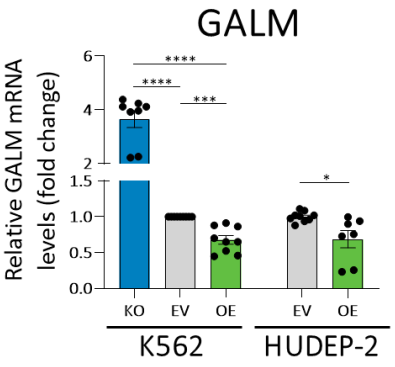
Next, to understand if COUP-TFII regulates identified candidate target genes independently on embryo/fetal erythroid cellular environment, I analyzed the impact of COUP-TFII overexpression (OE) on APOE, BAMBI, ENC1, GALM and PLA2G6 expression in HUDEP-2 adult-like erythroid cells (introduced in chapter 3.3.1).

I analyzed mRNA levels APOE, BAMBI, ENC1, GALM and PLA2G6 in K562 cells upon COUP-TFII overexpression (OE) or COUP-TFII knockout (KO) and in control cells transduced with an empty vector (EV). HUDEP-2 cells were transduced with COUP-TFII overexpressing vector and an empty vector (EV) as a control. APOE and GALM were downregulated by COUP-TFII in both K562 and HUDEP-2 cells (APOE: 40% reduction, GALM: 30% reduction), while PLA2G6, identified as a downregulated target of COUP-TFII in K562 cells (30% reduction) was significantly upregulated in HUDEP-2 COUP-TFII OE cells (2-fold increase) (Fig. 4.16). Therefore, as demonstrated by PLA2G6 gene regulation, the effect of COUP-TFII on its targets genes depends on the cellular erythroid context. Of note, among five candidate COUP-TFII target genes, two of them (BAMBI and ENC1) are expressed in HUDEP-2 cells at the marginal level, thus it is not surprising that we did not see any significant difference upon COUP-TFII OE in these cells (Fig. 4.16).


APOE
Relative APOE mRNA levels (fold change)
KO
EV
OE
EV
OE
K562
HUDEP-2


BAMBI
Relative BAMBI mRNA levels (fold change)
KO
EV
OE
EV
OE
K562
HUDEP-2


ENC1
Relative ENC1 mRNA levels (fold change)
KO
EV
OE
EV
OE
K562
HUDEP-2


GALM
Relative GALM mRNA levels (fold change)
KO
EV
OE
EV
OE
K562
HUDEP-2

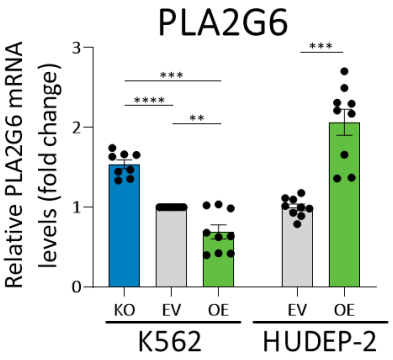

PLA2G6
Relative PLA2G6 mRNA levels (fold change)
KO
EV
OE
EV
OE
K562
HUDEP-2

Figure 4.16. ***COUP-TFII can regulate the expression of its direct targets differently in fetal-like K562 and adult-like HUDEP-2 erythroid cells.*** *Effects of COUP-TFII overexpression (OE) or knockout (KO) (72 h or 96 h post infection, pi, respectively) in K562 cells and of COUP-TFII OE (72 h pi) in HUDEP-2 cells on its candidate direct target genes, EV-control cells. Data as mean ± SEM, n≥3, *=p<0.05, **=p<0.01, ***=p<0.001, ****=p<0.0001.*

Taken together, the CRISPR/Cas9-mediated screening identified new genes through which COUP-TFII can regulate erythroid program. Importantly, depletion of none of them led to drastic changes in cell proliferation or differentiation in K562 erythroid cells. In particular APOE expression is repressed by COUP-TFII in both fetal-like (K562) and adult-like (HUDEP-2) cells, whereas PLA2G6 gene is repressed in K562 but activated in HUDEP-2 cells, suggesting a cell context-dependent regulation by COUP-TFII.

To test whether this difference is due to a differential binding of COUP-TFII in fetal vs. adult erythroid cells, I investigated the chromatin status and the presence of *in vivo* binding of COUP-TFII at these two loci.

COUP-TFII, when overexpressed (OE) in K562 cells, indeed binds to three genomic regions within the ApoE locus (Fig. 4.17). In control, non-treated (NT) K562 cells, COUP-TFII peak is present only within TOMM40 promoter region. The analysis of H3K27ac ChIP-seq and ATAC-seq profiles in K562 erythroid cells revealed that all COUP-TFII bound genomic regions show active regulatory region features (H3K27ac acetylation and open chromatin state). However, the

intergenic genomic region under the COUP-TFII peak has a very low acetylation mark level and chromatin-accessibility signature. In fact, this region is known hepatic control region (HCR.2) of APOE gene where the peak in intergenic region is located (Fig. 4.17). What is more, COUP-TFII binds to the same genomic regions also in adult-like erythroid cellular model as demonstrated by CUT&RUN experiment in HUDEP-2 cells overexpressing COUP-TFII (OE) (Fig. 4.17). The same binding pattern of COUP-TFII in fetal and adult cellular environment is consistent with the impact of COUP-TFII on ApoE mRNA levels in fetal-like K562 and adult-like HUDEP-2 erythroid cells (Fig. 4.16).

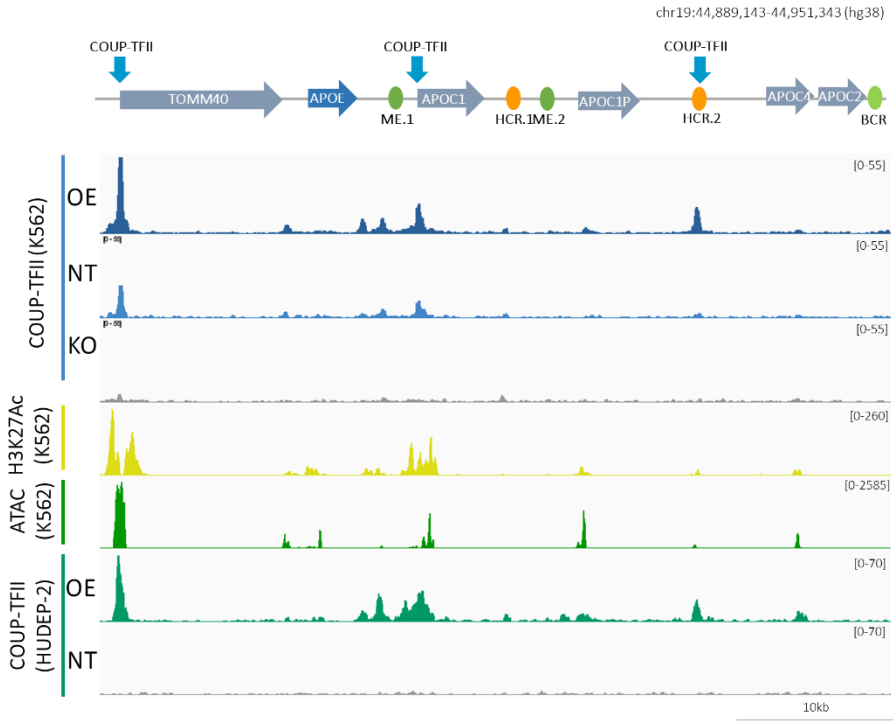


Figure 4.17. ***COUP-TFII regulates ApoE mRNA level by directly binding to ApoE locus.*** *TOMM40-APOE-APOC2 gene cluster on the chromosome 19, together with the distal regulatory elements: ME, multienhancer; HCR, hepatic control region; BCR, brain control region. Blue arrows indicate COUP-TFII binding sites. ChIP-seq profiles for K562 cells at ApoE locus: COUP-TFII (from the top: K562 overexpressing COUP-TFII (OE), non-treated (NT) and with COUP-TFII knockout (KO), H3K27ac (GEO: ENCFF465GBD) and ATAC-seq (GEO: ENCFF357GNC). COUP-TFII CUT&RUN profile for HUDEP-2 cells (overexpressing COUP-TFII (OE) and non-treated (NT)).*

Scanning of the core sequence of the three elements by using the JASPAR software identified several potential COUP-TFII consensus within these regions (Fig. 4.18).

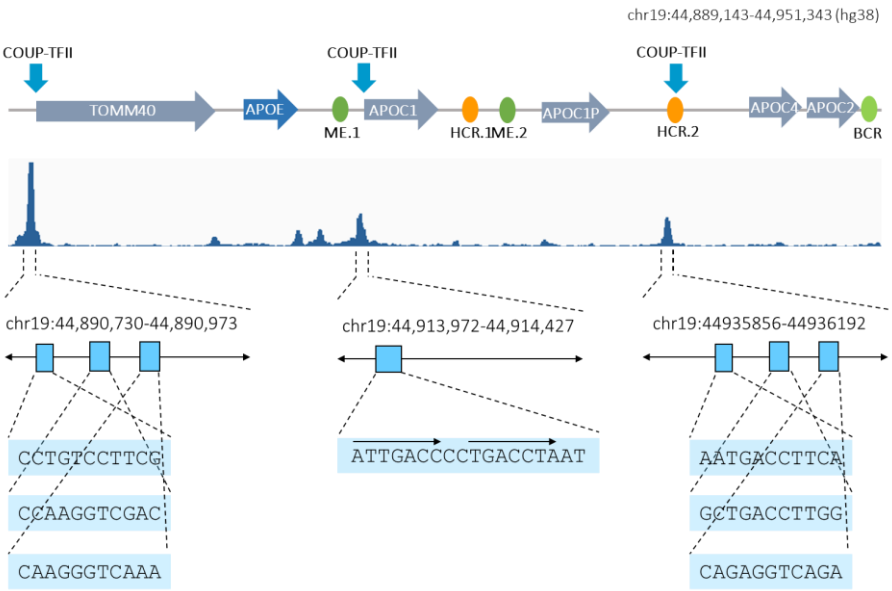


Figure 4.18. ***COUP-TFII motifs within the DNA sequences of ApoE locus that are bound in vivo by COUP-TFII.*** *Sequences under the COUP-TFII ChIP-seq peaks in the TOMM40 and APOC promoter regions, and the distal regulatory element HCR.2 were identifed by JASPAR.*

Within PLA2G6 locus, COUP-TFII binds to two intronic regions: intron 2 and intron 3 in both K562 and HUDEP-2 erythroid cells (Fig. 4.19). Additionally, COUP-TFII binds to the promoter region of PLA2G6 gene in HUDEP-2 cells.

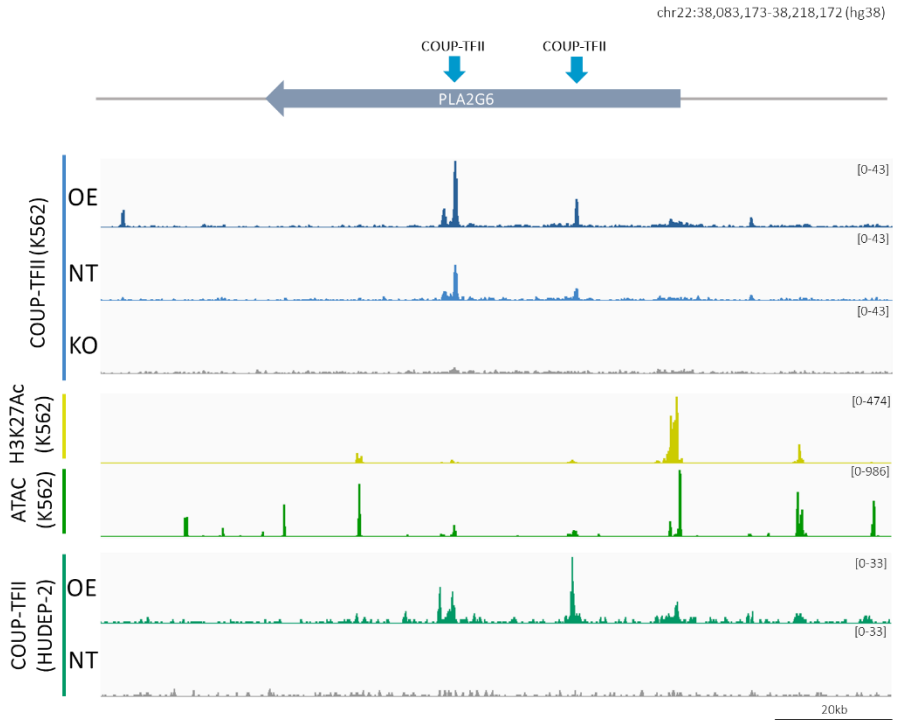


Figure 4.19. ***COUP-TFII binds intronic regions of PLA2G6 gene.*** *ChIP-seq profiles for K562 cells at PLA2G6 gene: COUP-TFII (from the top: K562 overexpressing COUP-TFII (OE), non-treated (NT) and with COUP-TFII knockout (KO), H3K27ac (GEO: ENCFF465GBD) and ATAC-seq (GEO: ENCFF357GNC). COUP-TFII CUT&RUN profile for HUDEP-2 cells (overexpressing COUP-TFII (OE) and non-treated (NT)).*

To conclude, I identified the possible mechanism by which COUP-TFII downregulates the erythroid expression of APOE gene. Whereas the negative effect of COUP-TFII on ApoE levels is due to COUP-TFII binding to all three regions remains to be further investigated.

4.3. References

Kliewer, S. A., Umesono, K., Noonan, D. J., Heyman, R. A., & Evans, R. M. (1992). Convergence of 9-cis retinoic acid and peroxisome proliferator signalling pathways through heterodimer formation of their receptors. Nature, 358(6389), 771-774.

Cooney, A. J., Tsai, S. Y., O'Malley, B. W., & Tsai, M. J. (1992). Chicken ovalbumin upstream promoter transcription factor (COUP-TF) dimers bind to different GGTCA response elements, allowing COUP-TF to repress hormonal induction of the vitamin D3, thyroid hormone, and retinoic acid receptors. Molecular and cellular biology, 12(9), 4153-4163.

Schaeffer, E., Guillou, F., Part, D., & Zakin, M. M. (1993). A different combination of transcription factors modulates the expression of the human transferrin promoter in liver and Sertoli cells. Journal of Biological Chemistry, 268(31), 23399-23408.

Stroup, D., & Chiang, J. Y. (2000). HNF4 and COUP-TFII interact to modulate transcription of the cholesterol 7α-hydroxylase gene (CYP7A1). Journal of lipid research, 41(1), 1-11.

Ktistaki, E., & Talianidis, I. (1997). Chicken ovalbumin upstream promoter transcription factors act as auxiliary cofactors for hepatocyte nuclear factor 4 and enhance hepatic gene expression. Molecular and cellular biology, 17(5), 2790-2797.

Okamura, M., Kudo, H., Wakabayashi, K. I., Tanaka, T., Nonaka, A., Uchida, A., ... & Sakai, J. (2009). COUP-TFII acts downstream of Wnt/β-catenin signal to silence PPARγ gene expression and repress adipogenesis. Proceedings of the National Academy of Sciences, 106(14), 5819-5824.

Leng, X., Cooney, A. J., Tsai, S. Y., & Tsai, M. J. (1996). Molecular mechanisms of COUP-TF-mediated transcriptional repression: evidence for transrepression and active repression. Molecular and Cellular Biology, 16(5), 2332-2340.

Fugazza, C., Barbarani, G., Elangovan, S., Marini, M. G., Giolitto, S., Font-Monclus, I., ... & Ronchi, A. E. (2021). The Coup-TFII orphan nuclear receptor is an activator of the γ-globin gene. haematologica, 106(2), 474.

Sankaran, V. G., Menne, T. F., Xu, J., Akie, T. E., Lettre, G., Van Handel, B., ... & Orkin, S. H. (2008). Human fetal hemoglobin expression is regulated by the developmental stage-specific repressor BCL11A. Science, 322(5909), 1839-1842.

Sankaran, V. G., Xu, J., Byron, R., Greisman, H. A., Fisher, C., Weatherall, D. J., ... & Bender, M. A. (2011). A functional element necessary for fetal hemoglobin silencing. New England Journal of Medicine, 365(9), 807-814.

Liu, N., Hargreaves, V. V., Zhu, Q., Kurland, J. V., Hong, J., Kim, W., ... & Orkin, S. H. (2018). Direct promoter repression by BCL11A controls the fetal to adult hemoglobin switch. Cell, 173(2), 430-442.

Lozzio, C. B., & Lozzio, B. B. (1975). Human chronic myelogenous leukemia cell-line with positive Philadelphia chromosome.

Zambanini, G., Nordin, A., Jonasson, M., Pagella, P., & Cantù, C. (2022). A new CUT&RUN low volume-urea (LoV-U) protocol optimized for transcriptional co-factors uncovers Wnt/β-catenin tissue-specific genomic targets. Development, 149(23), dev201124.

Proietti, L., Manhart, G., Heyes, E., Troester, S., & Grebien, F. (2022). Arrayed CRISPR/Cas9 Screening for the Functional Validation of Cancer Genetic Dependencies. Bio-protocol, 12(24).

Liu, N., Hargreaves, V. V., Zhu, Q., Kurland, J. V., Hong, J., Kim, W., ... & Orkin, S. H. (2018). Direct promoter repression by BCL11A controls the fetal to adult hemoglobin switch. Cell, 173(2), 430-442.

Jackson, D. A., McDowell, J. C., & Dean, A. (2003). β-Globin locus control region HS2 and HS3 interact structurally and functionally. Nucleic acids research, 31(4), 1180-1190.

Jiang, J., Best, S., Menzel, S., Silver, N., Lai, M. I., Surdulescu, G. L., ... & Thein, S. L. (2006). cMYB is involved in the regulation of fetal hemoglobin production in adults. Blood, 108(3), 1077-1083.

Sankaran, V. G., Menne, T. F., Šćepanović, D., Vergilio, J. A., Ji, P., Kim, J., ... & Lodish, H. F. (2011). MicroRNA-15a and-16-1 act via MYB to elevate fetal hemoglobin expression in human trisomy 13. Proceedings of the National Academy of Sciences, 108(4), 1519-1524.

Gu, Q., Palani, C. D., Smith, A., Li, B., Amos-Abanyie, E. K., Ogu, U., ... & Starlard-Davenport, A. (2022). MicroRNA29B induces fetal hemoglobin via inhibition of the HBG repressor protein MYB in vitro and in humanized sickle cell mice. Frontiers in Medicine.

Thein, S. L., Menzel, S., Peng, X., Best, S., Jiang, J., Close, J., ... & Lathrop, M. (2007). Intergenic variants of HBS1L-MYB are responsible for a major quantitative trait locus on chromosome 6q23 influencing fetal hemoglobin levels in adults. Proceedings of the National Academy of Sciences, 104(27), 11346-11351.

Stadhouders, R., Aktuna, S., Thongjuea, S., Aghajanirefah, A., Pourfarzad, F., van IJcken, W., ... & Soler, E. (2014). HBS1L-MYB intergenic variants modulate fetal hemoglobin via long-range MYB enhancers. The Journal of clinical investigation, 124(4), 1699-1710.

Kihm, A. J., Kong, Y. I., Hong, W., Russell, J. E., Rouda, S., Adachi, K., ... & Weiss, M. J. (2002). An abundant erythroid protein that stabilizes free α-haemoglobin. Nature, 417(6890), 758-763.

Kong, Y. I., Zhou, S., Kihm, A. J., Katein, A. M., Yu, X., Gell, D. A., ... & Weiss, M. J. (2004). Loss of α-hemoglobin–stabilizing protein impairs erythropoiesis and exacerbates β-thalassemia. The Journal of clinical investigation, 114(10), 1457-1466.

Andrews, N. C. (1994). Erythroid transcription factor NF-E2 coordinates hemoglobin synthesis. Pediatric research, 36(4), 419-423.

Mutschler, M., Magin, A. S., Buerge, M., Roelz, R., Schanne, D. H., Will, B., ... & Pahl, H. L. (2009). NF-E2 overexpression delays erythroid maturation and increases erythrocyte production. British journal of haematology, 146(2), 203-217.

Lopez, G. H., Hyland, C. A., & Flower, R. L. (2021). Glycophorins and the MNS blood group system: a narrative review. Ann Blood, 6, 39.

Hollox, E. J., & Louzada, S. (2022). Genetic variation of glycophorins and infectious disease. Immunogenetics, 1-6.

Cui, Y., Yang, J., Bai, Y., Li, Q., Yao, Y., Liu, C., ... & Zhang, Y. (2021). ENC1 facilitates colorectal carcinoma tumorigenesis and metastasis via JAK2/STAT5/AKT axis-mediated epithelial mesenchymal transition and stemness. Frontiers in Cell and Developmental Biology, 9, 616887.

Onichtchouk, D., Chen, Y. G., Dosch, R., Gawantka, V., Delius, H., & Niehrs, C. (1999). Silencing of TGF-β signalling by the pseudoreceptor BAMBI. Nature, 401(6752), 480-485.

Sekiya, T., Adachi, S., Kohu, K., Yamada, T., Higuchi, O., Furukawa, Y., ... & Akiyama, T. (2004). Identification of BMP and activin membrane-bound inhibitor (BAMBI), an inhibitor of transforming growth factor-β signaling, as a target of the β-catenin pathway in colorectal tumor cells. Journal of Biological Chemistry, 279(8), 6840-6846.

Togo, N., Ohwada, S., Sakurai, S., Toya, H., Sakamoto, I., Yamada, T., ... & Akiyama, T. (2008). Prognostic significance of BMP and activin membrane-bound inhibitor in colorectal cancer. World journal of gastroenterology: WJG, 14(31), 4880.

Fritzmann, J., Morkel, M., Besser, D., Budczies, J., Kosel, F., Brembeck, F. H., ... & Birchmeier, W. (2009). A colorectal cancer expression profile that includes transforming growth factor β inhibitor BAMBI predicts metastatic potential. Gastroenterology, 137(1), 165-175.

Murphy, A. J., Akhtari, M., Tolani, S., Pagler, T., Bijl, N., Kuo, C. L., ... & Tall, A. R. (2011). ApoE regulates hematopoietic stem cell proliferation, monocytosis, and monocyte accumulation in atherosclerotic lesions in mice. The Journal of clinical investigation, 121(10), 4138-4149.

Kuvardina, O. N., Herglotz, J., Kolodziej, S., Kohrs, N., Herkt, S., Wojcik, B., ... & Lausen, J. (2015). RUNX1 represses the erythroid gene expression program during megakaryocytic differentiation. Blood, The Journal of the American Society of Hematology, 125(23), 3570-3579.

Cantù, C., Ierardi, R., Alborelli, I., Fugazza, C., Cassinelli, L., Piconese, S., ... & Ronchi, A. (2011). Sox6 enhances erythroid differentiation in human erythroid progenitors. Blood, The Journal of the American Society of Hematology, 117(13), 3669-3679.

Holm, T. M., Braun, A., Trigatti, B. L., Brugnara, C., Sakamoto, M., Krieger, M., & Andrews, N. C. (2002). Failure of red blood cell maturation in mice with defects in the high-density lipoprotein receptor SR-BI. Blood, The Journal of the American Society of Hematology, 99(5), 1817-1824

Crespo, P., Delgado, M. D., Gómez-Casares, M. T., Cuadrado, M. A., Richard, C., & León, J. (1993). Induction of apolipoprotein E expression during erythroid differentiation of human K562 leukemia cells. Leukemia research, 17(9), 771-776.

CHAPTER 5: FINAL DISCUSSION. MOLECULAR AND CLINICAL CONSIDERATIONS

5.1. COUP-TFII transcription factor: an attractive therapeutic target with widely unknown erythroid signalling networks

In diseases such as sickle cell anemia and β-thalassemia, the re-activation of the fetal γ-globin expression can balance the defective β-globin production, as it is observed in hereditary persistence of fetal hemoglobin (HPFH), a benign condition in which the production of fetal hemoglobin continues in adulthood (Steinberg, 2020).

Therapeutic approaches that reactivate fetal hemoglobin (HbF) production remain the most effective strategy to ameliorate the clinical symptoms of β-hemoglobinopathies. Most of the new and experimentally proposed strategies are based on genome editing of γ-globin repressor or on pharmacological methods (Barbarani, 2020; Bou-Fakhredin, 2022).

The COUP-TFII transcription factor is an essential regulator of many developmental processes, as demonstrated by the early embryonic lethality of COUP-TFII$^{-/-}$ mice (Qiu, 1997; Pereira, 2000; Lin, 2011). Even though the biological function of COUP-TFII has been widely studied in different tissues by the use of conditional knockouts, the role of COUP-TFII and the upstream signals controlling its expression in erythroid cells remain poorly understood.

By *in vitro* binding studies, COUP-TFII was initially proposed as a repressor of embryo/fetal globins, as a part of the NF-E3 complex (Ronchi, 1995; Liberati, 1998). More recently, COUP-TFII was instead demonstrated by *in vivo* studies to directly bind within the locus control region (LCR) of the β-locus and to act as an activator of γ-globin in different types of erythroid cells, including adult human cells (Fugazza, 2021) from healthy and thalassemic donors.

Thus, COUP-TFII as one of the regulators of the hemoglobin switching process is an attractive prospective target in β-hemoglobinopathies' therapies. In fact, the understanding of COUP-TFII erythroid fetal transcription factor networks in erythroid cells will open up the possibilities to re-activate its expression in adult erythroid progenitors of β-thalassemia or sickle cell disease (SCD) patients with and therefore to reactivate the expression of fetal γ-globin, thus and mimicking the HPFH phenotype.

To this end, it is crucial to identify regulatory elements controlling the COUP-TFII developmentally regulated expression in erythroid cells, either through fetal-specific enhancer(s) or adult-specific silencer(s), or both. In addition, it is important to characterize the COUP-TFII target genes other than γ-globin in order to understand whether/how the modulation of COUP-TFII level could affect the erythroid program.

5.2. The transcriptional regulation of COUP-TFII

5.2.1. A direct repeat (DR1) element within the COUP-TFII promoter mediates its positive regulation in K562 erythroid cells

Little is known about the transcriptional regulation of COUP-TFII, even though its context-dependent expression pattern suggests a complex regulatory mechanism.

COUP-TFII gene expression was shown to be positively regulated by supraphysiological concentration of retinoic acid (RA) in P19 embryonal carcinoma (EC) cells and PCC7 teratocarcinoma cells (Jonk, 1994; Qiu, 1996; Soosar, 1996), Sonic hedgehog (Shh) in P19 cells (Krishnan, 1997) and Ets-1 in HeLa cells (Petit, 2004).

Yet, COUP-TFII mRNA level was shown to be upregulated in different cellular models by retinoic acid, this effect was never confirmed in erythroid cells. Therefore, I treated K562 erythroid cells (expressing COUP-TFII together with fetal hemoglobin) with RA and 48 h after the mRNA level of endogenous COUP-TFII was indeed increased in the treated cells, when compared to the control (Fig. 2.5B). To identify the sequences mediating RA activation of the COUP-TFII promoter, I

cloned the minimal COUP-TFII promoter described by Soosar and colleagues (1996) in the pGL3 reporter vector. Interestingly, the COUP-TFII minimal promoter of contains two conserved hormone response elements (HREs), known to be bound by nuclear receptors. The distal direct repeat 1 (DR1) is located 72 bp upstream to the COUP-TFII transcription start site (TSS), and the proximal inverted repeat (IR) is located 49 bp upstream to the COUP-TFII TSS (Fig. 2.1B). By removing the distal -72 bp DR1 site, The reporter gene expression was significantly reduced in luciferase assays (Fig. 2.5A) demonstrating that this element (and not the -49 bp IR element) is important for the COUP-TFII activation by retinoic acid. Which nuclear receptors can bind the DR1 of the COUP-TFII promoter remains to be elucidated. It would be interesting to perform a co-transfection experiment by using the reporter vector containing the COUP-TFII minimal promoter and vectors expressing different nuclear receptors activated by RA, such as RAR, RXR or COUP-TFII itself. However ChIP-seq profiles of COUP-TFII in erythroid cells, neither in K562 nor in HUDEP-1/2 cells show a direct COUP-TFII binding to its own promoter.

Related to this, it has been shown that in pancreatic β-cells this same DR1 element in promoter, is indeed directly bound by COUP-TFII, but this results in a negative autoregulation of COUP-TFII gene expression (Perilhou, 2008). This suggests that COUP-TFII binding to its on promoter is cell context-dependent and can result in opposite functional outputs.

It is interesting to note that retinoic acid could also be involved in COUP-TFII protein activation, although it was originally classified as an orphan NR, due to the lack of a known natural ligand. However, Kruse and colleagues (2008) demonstrated that COUP-TFII is a low-affinity retinoic acid receptor, activated only by supra physiological doses of RA. Moreover, other studies identified small-molecule synthetic inhibitors of COUP-TFII that bind directly to the COUP-TFII ligand binding domain (LBD) (Guevel, 2017; Wang, 2020). Finally, more recently, 1-deoxysphinosines were identified as candidate natural ligands of COUP-TFs (Wang, 2021). These non-canonical alanine-based sphingolipids positively regulate the transcriptional activity of COUP-TFII protein by binding to its LBD. All together, these findings demonstrate that the "orphan" nuclear receptor COUP-TFII could be instead a ligand-regulated and therefore a pharmacologically modulable protein.

5.2.2. The -52 kb COUP-TFII enhancer element mediates a COUP-TFII positive autoregulation in K562 erythroid cells

According to the ChIP-seq and CUT&RUN data, COUP-TFII does not bind to the DR1 in the minimal promoter discussed above. Instead, it binds to the -52 kb COUP-TFII enhancer element (Fig. 2.11) and

CRISPR/Cas9-mediated targeting of the COUP-TFII motif within this element reduces the COUP-TFII expression level by 40% (Fig. 2.15B). These results point to a possible positive autoregulation of COUP-TFII expression in erythroid cells through this element. Moreover, the exogenous COUP-TFII overexpression in K562 cells leads to increased mRNA levels of endogenous COUP-TFII, supporting the concept of a positive feedback loop (Fig. 2.16A). Of note, the overexpression of COUP-TFII in HUDEP-2 cells, which do not express endogenous COUP-TFII, did not result in the gene reactivation (Fig. 2.16B). However, the H3K27ac and ATAC-seq profiles in HUDEP-2 cells from the ENCODE data confirmed that COUP-TFII locus is not transcriptionally active in these cells. Therefore, COUP-TFII can increase its own expression only in fetal-like erythroid cells in which the COUP-TFII site is already open, but not in the adult erythroid environment when the locus is not accessible.

Overall, these data suggest an autoregulatory mechanism of COUP-TFII in which COUP-TFII can either activate or repress its own gene expression in a cell-dependent manner.

5.2.3. Identification of the human-specific COUP-TFII enhancer element in erythroid cells

To identify potential enhancer elements of COUP-TFII in erythroid cells, I analyzed the evolutionary conservation, epigenetic marks and transcription factor binding sites (TFBSs) within the predicted enhancers that interact with the COUP-TFII promoter according to the GeneHancer database (Fishilevich, 2017) (Fig. 2.6, 2.7, 2.8). Four regions located between 61 and 40 kb upstream of the COUP-TFII TSS (Table 2.4) were further analyzed by luciferase reporter assays in K562 erythroid and HepG2 hepatic cell lines, both expressing COUP-TFII. Since two of the selected regions (-61 and –57 kb) had been previously studied by Baroukh and colleagues (2005) in HepG2 cells, I used HepG2 cells as a control. Surprisingly, the -61 kb region identified as hepatic COUP-TFII enhancer (Baroukh, 2005), in our experiments did not show any activating activity in HepG2 cells, and in K562 cells it decreased the COUP-TFII promoter activity. The -57 kb, region previously reported as an inactive element, instead increases the reporter gene expression by 3-fold in HepG2 cells (Fig. 2.10). It is possible that the discrepancies in the evaluation of the regulatory potential of the -61 and -57 kb regions are due to the fact that in the study of Baroukh's group, these regions were cloned upstream to the SV40 promoter, and not to the COUP-TFII promoter, as we did. The -40 kb element in HepG2 cells did not affect the COUP-TFII promoter activity significantly, while in K562 it showed mild decrease. Finally, the -52 kb element increased the

reporter gene expression significantly (5-fold), in both cellular models (Fig. 2.10).

The -52 kb region has both H3K27ac mark and open chromatin feature in K562, but not in HepG2 cells (Fig. 2.8). This suggests that the -52 kb region can be indeed an erythroid-specific enhancer of COUP-TFII. Interestingly, the nucleotide sequence of this region is not evolutionary conserved, and thus the enhancer activity of the -52 kb region is also human-specific.

5.2.4. The -52 kb COUP-TFII enhancer element is bound by key erythroid transcription factors

We demonstrated (see above) that COUP-TFII itself binds within the -52 kb region (Fig. 2.11). To obtain more information about the transcription factors regulating the identified COUP-TFII enhancer activity, I analyzed the transcription factors reported by ENCODE ChIP-seq data to bind within the -52 kb region in K562 erythroid cells (Table 2.5). Among them, there are some key erythroid regulators: STAT5 (Grebien, 2008; Tothova, 2021), TAL1 (Kassouf, 2010) and NFE2 (Andrews, 1994; Mutschler, 2009) (Fig. 2.11). In addition, ChIP-seq profiles of binding of GATA1, the "master" erythroid transcription factor (Shimizu, 2008; Romano, 2020) published in other datasets, confirm a peak of binding within this region, even though GATA1 was

not present in the obtained list (Table 2.5). What is more, it is known that GATA1/TAL1 complex is strongly associated with the activation of erythroid gene expression (Wu, 2014; Han, 2016). Indeed, the -52 kb enhancer element contains the motif consisting of a TG 7 or 8 bp upstream of a WGATAA, that was described by Han and colleagues (2016) to be bound by GATA1/TAL1 complex. In addition, GATA2, an essential regulator of early steps in erythroid differentiation (Tsai and Orkin, 1997), emerged in the list of TF binding the -52 kb element. The two factors belong to the same family, recognize the same core motif and it is known that GATA2 binding is often replaced by GATA1 on the same genomic sites during the erythroid differentiation process (Romano, 2020). The progressive deletion analysis of the -52 kb COUP-TFII enhancer element in luciferase assays allowed to evaluate the contribution of the identified consensus motifs to the enhancer activity of this region. In particular, the removal of the STAT motifs significantly reduced the activity of the luciferase reporter (Fig. 2.12B).

Finally, the CRISPR/Cas9-mediated deletion of the -52kb enhancer element in K562-Cas9 cells, confirmed the importance of the identified region in the regulation of the COUP-TFII expression level. The level of COUP-TFII mRNA was reduced by 80% in cells carrying the deletion of the non-repetitive core sequence of the -52kb element (Fig. 2.15B). When only the COUP-TFII binding site within this region was disrupted, the COUP-TFII mRNA level is reduced by 40% (Fig. 2.15B).

5.3. Role of COUP-TFII in the regulation of the β-locus

In K562 cells, a fetal-like erythroid cellular model, COUP-TFII induces γ-globin expression by direct binding to the locus control region (LCR) of the β-locus, possibly by favoring the interaction of the LCR with HBG genes, encoding γ-globin (Fugazza, 2021). To better characterize the COUP-TFII binding to the β-locus, I used two human umbilical cord-derived erythroid progenitor (HUDEP) cell lines: the fetal-like HUDEP-1 and the adult-like HUDEP-2 cells. Although both cell lines do not express COUP-TFII, HUDEP-1 cells express fetal globin genes, whereas HUDEP-2 cells express adult globin genes. The exogenous expression of COUP-TFII in both HUDEP-1 and HUDEP-2 cells, as expected, reactivated HBG gene transcription (Fig. 4.2). Therefore, CUT&RUN experiment was performed in these two cellular models overexpressing COUP-TFII to map regions bound by COUP-TFII within the β-locus. In these cells, COUP-TFII binds to the same hypersensitive sites (HS2, HS3 and HS4) as in K562 cells (Fig. 4.3), with the strongest peak observed at HS3. Indeed, HS3 element was previously reported to be involved in the β-locus regulation of embryo/fetal globins (Navas, 1998). What is more, COUP-TFII sites overlap within those bound by BCL11A (Liu, 2018) (Fig. 4.4). COUP-TFII is an embryo/fetal erythroid transcription factor, whereas BCL11A is an adult regulator of β-locus. Thus, by binding to the same consensus sequence, they could

functionally contribute to the differential regulation of embryo/fetal vs. adult goblins.

5.4. MYB transcription factor, a γ-globin repressor, is a candidate direct target of COUP-TFII

MYB is an adult erythroid transcription factor and a known repressor of γ-globin. Increased levels of MYB in K562 erythroleukemic cells reduce the expression of γ-globin (Jiang, 2006), whereas the inhibition of MYB in different cellular and mouse models results in increased fetal hemoglobin (HbF) (Sankaran, 2011; Gu, 2022).

The ChIP-seq experiment performed in this study in K562 cells demonstrated the binding of COUP-TFII to the HBS1L-MYB intergenic region, a quantitative trait locus (QTL) controlling HbF level (Thein, 2007) (Fig. 4.6). What is more, one of the regions bound by COUP-TFII was identified by Stadhouders (2014) as an enhancer element of MYB gene. The RT-qPCR analysis revealed that in short-term K562 cells overexpressing COUP-TFII, the level of MYB mRNA is reduced, whereas its long-term overexpression (25 days post transduction), leads to a MYB level increase (Fig. 4.7). On the other hand, COUP-TFII knockout in K562 cells did not cause any variations in MYB protein expression

(Fig. 4.7). The discrepancy between short-term and long-term effect of elevated levels of COUP-TFII on MYB could be explained by some indirect downstream effects of COUP-TFII overexpression that are established over time.

To check if COUP-TFII can control the expression of MYB by the HBS1L-MYB intergenic region, I evaluated the regulatory potential of the P1 and P4 elements within this region depends on the COUP-TFII level by luciferase assay in K562 cells expressing different COUP-TFII levels. Both regions increased the expression of the reporter gene, confirming their relevance for MYB expression. However, their activity does not depend on the presence or absence of COUP-TFII (Fig. 4.9). These data are confirmed by co-transfection experiments of P1 and P4 reporter vectors together with a vector expressing COUP-TFII (Fig. 4.10).

5.5. ApoE is a candidate direct COUP-TFII target relevant for erythroid cells metabolism

COUP-TFII overexpression in adult cells does not significantly change the proliferation/differentiation profile (Fugazza, 2021), suggesting that its major effect in these cells is the activation of fetal γ-globin gene.

However, nothing is known about additional direct COUP-TFII targets that could be relevant in embryo/fetal cells during early development. Although COUP-TFII overexpression or knockout in K562 erythroid cells did not result in any dramatic alterations of cell proliferation/differentiation balance, it is possible that depletions of candidate targets have a stronger impact on erythroid program than only upon COUP-TFII level modulation. To get insight on this point, I performed the integrative analysis of RNA-seq and ChIP-seq in K562 erythroid cells with different levels of COUP-TFII. 66 genes were identified as candidate direct target genes of COUP-TFII of which 33 were upregulated and 33 downregulated by COUP-TFII (Fig. 4.12B, table 4.6). After further validations, 12 genes were selected for CRISPR/Cas9 screening experiment in K562-Cas9 cells. None of these genes has a well characterized role in erythroid cells. Hence, I sought to explore effects of targeting them with sgRNAs on K562 erythroid cell proliferation and differentiation.

Among 12 genes that were sgRNA targeted in K562-Cas9, APOE, BAMBI, ENC1, GALM and PLA2G6 showed consistent effect on cell proliferation and/or erythroid differentiation marker levels (Fig. 4.14, 4.15). APOE and PLA2G6, identified as downregulated direct targets of COUP-TFII, are involved in lipid metabolism (Marais, 2019; Turk, 2019), GALM functions in galactose metabolism (Timson and Reece, 2003), and BAMBI and ENC1, both upregulated by COUP-TFII are involved in cellular developmental processes (Higashihori, 2008; Hernandez, 1997). APOE, GALM and PLA2G6 are expressed both in K562 and in

adult HUDEP-2 erythroid cells. RT-qPCR analysis of expression levels of these 3 genes in HUDEP-2 cells overexpressing COUP-TFII showed that APOE, GALM and PLA2G6 can be also regulated by COUP-TFII in adult erythroid progenitor cells (Fig. 4.16). The effect of COUP-TFII on APOE and GALM gene expression in adult-like HUDEP-2 cells was the same as in fetal-like K562 cells, whereas PLA2G6, downregulated in K562 cells is upregulated by COUP-TFII in HUDEP-2 cells.

What is more, PLA2G6 and APOE affect cell phenotype in the opposite manner as COUP-TFII – as expected since they are both downregulated targets of COUP-TFII (Fig. 4.14, Fig. 4.15).

BAMBI (BMP And Activin Membrane-Bound Inhibitor Homolog) is a pseudoreceptor that negatively regulates TGFβ signaling by preventing the activation of the TGFβ receptor (Onichtchouk, 1999). TGFβ inhibits proliferation of erythroid progenitors and accelerates their differentiation (Zermati, 2000). Indeed, a role of BAMBI in cell proliferation was suggested by numerous observations: in tumors, elevated levels of BAMBI correlate with tumor growth and metastasis (Sekiya, 2004; Togo, 2008; Fritzmann, 2009). Interestingly, BAMBI mRNA expression is reduced upon COUP-TFII shRNA targeting in breast cancer cells (Erdős and Bálint, 2020).

PLA2G6 gene encodes a calcium-independent group VI phospholipase A2 (iPLA2β), an ubiquitously expressed enzyme that catalyzed the release of fatty acids from phospholipids. iPLA2β regulates cell membrane permeability and mitochondrial integrity (Kinghorn, 2015). Mutations of PLA2G6 have been associated with neurodegenerative

disease with iron dyshomeostasis (Morgan, 2006; Schaeffer, 2008; Paisan-Ruiz, 2012). Knockout of PLA2G6 was shown to cause accelerated lipid peroxidation, mitochondrial membrane abnormalities and an increase in reactive oxygen species (ROS) level (Kinghorn, 2015). PLA2G6 was never described to have a particular role within erythroid lineage. In PLA2G6 KO K562 cells, we did not observe changes in the proliferation rate (Fig. 4.14). In contrast, the expression of GpA, a late erythroid marker, was clearly upregulated in these cells (Fig. 4.15). The percentage of PLA2G6 KO cells remains stable in the competition assay, with no loss in CD71 marker and increase in GpA, suggesting that the PLA2G6 KO promotes partial erythroid differentiation. Interestingly, the upregulation of iPLA2β was linked to higher levels of TNFα, a cytokine that was shown to suppress GpA^+ population of cells (Xiao, 2002). Therefore, a possible explanation of the GpA upregulation in PLA2G6 KO cells, is that PLA2G6 depletion led to a decrease in TNFα level, and, in consequence, an increase in GpA level.

Apolipoprotein E (ApoE) is a secreted glycoprotein with a wide range of roles, in particular in lipid metabolism and in the pathogenesis of Alzheimer's disease. The APOE gene has three polymorphic alleles (ε2, ε3, and ε4) and encoded ApoE isoforms show different affinity for lipoproteins particles and low-density lipoprotein receptor (Giau, 2015; Husain, 2021).

The regulation of APOE gene expression is highly complex and tissue-specific (Kockx, 2018). It involves proximal and distal regulatory

elements, various transcription factors and epigenetics. Within the extended ApoE locus (TOMM40-APOE-APOC2), the genes are interspersed with the regulatory regions: homologous pair of hepatic control region (HCR.1 and HCR.2) (Simonet, 1993; Shachter, 1993; Allan, 1995; Allan, 1997), the copies of multienhancer (ME.1 and ME.2) (Shih, 2000; Tse, 2001) and the brain control region (BCR) (Zheng, 2004). ME.1 and ME.2 drive Apoe expression in astrocytes, macrophages and adipocytes, while they remain inactive in hepatocytes. Evolutionary duplicates HCR.1 and HCR.2 are active in both hepatic and non-hepatic tissues.

Among transcriptions factors that were shown to regulate APOE gene expression by binding to ME.1/ME.2 are LXR/RXR and STAT1 (Liang, 2004, Trusca, 2011; Trusca, 2012). Recently, C/EBPβ was identified as a key regulator of ApoE expression in brain, enhancing APOE gene expression by binding to the ApoE promoter (Xia, 2021). In hepatic cells, ApoE level is regulated by TR4 and PPARγ (Kim, 2003; Subramanian, 2017).

Our experiments showed that COUP-TFII negatively regulates ApoE level in fetal-like K562 and adult-like HUDEP-2 erythroid cells (Fig. 4.16). COUP-TFII binds to HCR.2 element and to promoter regions of TOMM40 and APOC1 genes when overexpressed in K562 cells (Fig. 4.17). All of three regions bound by COUP-TFII are acetylated and show chromatin accessibility signature in K562 cells. HCR.2 and TOMM40 promoter region were shown to regulate APOE gene. The methylation of TOMM40 promoter has been reported to influence ApoE mRNA

level (Shao, 2018). COUP-TFII binds also to APOC1 promoter. Interestingly, APOC1 was indicated in the list of candidate direct target genes from RNA-seq and ChIP-seq integrative analyses as downregulated by COUP-TFII (Table 4.6). However, it was not selected to be further analyzed as RT-qPCR analysis did not show significant decrease in ApoC1 mRNA level upon COUP-TFII overexpression in K562 cells.

The ApoE$^{-/-}$ mouse model was established over 30 years ago, followed by a mouse knockin model with human ApoE and up to now both are a widely used (Zhang, 1992; Jawien, 2012; Pendse, 2009). While the aspects of ApoE function and its transcriptional regulation were mainly studied in hepatocytes, macrophages and neural cells, there is also quite some evidence that ApoE is important in erythroid cells.

ApoE has a role in modulating reactive oxygen species (ROS) production, cellular damage and alteration of proteins' production and folding (Ramassamy, 1999). High concentration of oxygen and hemoglobin make erythrocytes particularly sensitive to oxidative damage. Indeed, it was observed that Aβ levels and oxidative stress was increased in erythrocytes from ApoE E4 carriers (Piccarducci, 2019). Particularly, enhanced oxidative damage has been linked to elevated lipid (membrane phospholipid) peroxidation, resulting in altered composition of the cellular membrane, and thus its properties (Miyazawa, 1996; Ramassamy, 1999). In line with this, decreased levels of erythrocyte membrane fluidity have been observed in subjects carrying the ApoE ε4 (Piccarducci, 2019). This is consistent

with a study in which erythrocyte membrane fluidity decreased in response to a drop in phospholipid content (Saha, 2012; Ortiz, 2013).

Holm and colleagues (2002) showed that autophagocytosis and phagolysosome expulsion are essential steps in erythroid maturation, and that these steps are inhibited by altered plasma lipoprotein metabolism and elevated cellular cholesterol levels, leading to maturation disruption of red blood cells (RBCs). Interestingly, ApoE level was upregulated during the differentiation of K562 cells by ara-C or hydroxyurea, indicating the important role of ApoE in terminal maturation process of RBCs (Crespo, 1993).

In our study, ApoE knockout led to a decrease in cell proliferation of K562 erythroid cells. It is possible that oxidative stress has different impact on progenitor and K562 cells, since erythroid cells are more sensitive to oxidative damage due to high concentration of oxygen and hemoglobin, as mentioned before. Negative regulation of K562 cell proliferation by oxidative stress was already reported by Kulkarni and colleagues (2009). Figure 5.1 shows a hypothetical model of downstream effects of APOE gene negative regulation by COUP-TFII in erythroid cells.

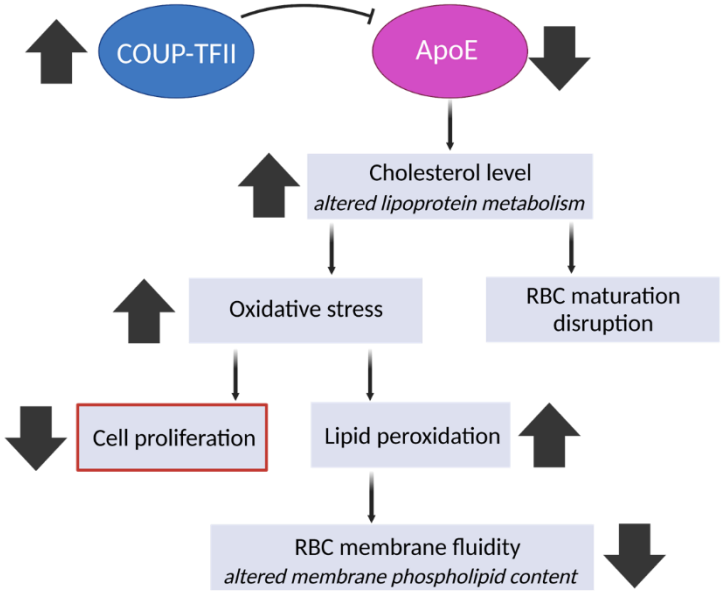


Figure 5.1. ***Hypothetical model of APOE gene expression regulation by COUP-TFII and its downstream effects.***

The knowledge on the ApoE role in controlling erythroid metabolism, suggests its relevant role in ROS response and in membrane fluidity. To what extent ApoE is a relevant target in erythroid cells of yolk sac origin in which COUP-TFII is physiologically expressed during development, remains to be elucidated.

5.6. Final remarks and future perspective

In conclusion, the transcription factor COUP-TFII acts as a γ-globin activator in both fetal and adult cellular environments. Since the re-activation of γ-globin expression is a therapeutic target in the treatment of β-hemoglobinopathies, the induction of COUP-TFII in adult erythroid progenitor cells is a potential approach to activate fetal γ-globin. Consequently, the re-activation of γ-globin expression by COUP-TFII would compensate for defective β-globin production in patients with β-hemoglobinopathies. In this study, I identified the enhancer element that most likely regulates the expression of COUP-TFII in erythroid cells. This finding opens up new possibilities for the modulation of COUP-TFII gene expression. Further characterization of the mechanism by which this element controls COUP-TFII levels would be crucial in order to design strategies for the COUP-TFII re-activation in adult erythroid cells.

Moreover, I described possible COUP-TFII direct target genes of functional relevance in erythroid cells. This knowledge and further characterization of COUP-TFII erythroid networks are important to elucidate the molecular mechanism of globin genes' control and in light of the potential of COUP-TFII as a target for treating β-hemoglobinopathies.

5.7. References

Steinberg, M. H. (2020). Fetal hemoglobin in sickle hemoglobinopathies: high HbF genotypes and phenotypes. Journal of Clinical Medicine, 9(11), 3782.

Barbarani, G., Labedz, A., & Ronchi, A. E. (2020). β-Hemoglobinopathies: the test bench for genome editing-based therapeutic strategies. Frontiers in Genome Editing, 2, 571239.

Bou-Fakhredin, R., De Franceschi, L., Motta, I., Cappellini, M. D., & Taher, A. T. (2022). Pharmacological Induction of Fetal Hemoglobin in β-Thalassemia and Sickle Cell Disease: An Updated Perspective. Pharmaceuticals, 15(6), 753.

Qiu, Y., Pereira, F. A., DeMayo, F. J., Lydon, J. P., Tsai, S. Y., & Tsai, M. J. (1997). Null mutation of mCOUP-TFI results in defects in morphogenesis of the glossopharyngeal ganglion, axonal projection, and arborization. Genes & development, 11(15), 1925-1937.

Pereira, F. A., Tsai, M. J., & Tsai*, S. Y. (2000). COUP-TF orphan nuclear receptors in development and differentiation. Cellular and Molecular Life Sciences CMLS, 57, 1388-1398.

Lin, F. J., Qin, J., Tang, K., Tsai, S. Y., & Tsai, M. J. (2011). Coup d'Etat: an orphan takes control. Endocrine reviews, 32(3), 404-421.

Ronchi, A. E., Bottardi, S., Mazzucchelli, C., Ottolenghi, S., & Santoro, C. (1995). Differential Binding of the NFE3 and CP1/NFY Transcription Factors to the Human γ-andϵ-Globin CCAAT Boxes (∗). Journal of Biological Chemistry, 270(37), 21934-21941.

Liberati, C., Ronchi, A., Lievens, P., Ottolenghi, S., & Mantovani, R. (1998). NF-Y organizes the γ-globin CCAAT boxes region. Journal of Biological Chemistry, 273(27), 16880-16889.

Fugazza, C., Barbarani, G., Elangovan, S., Marini, M. G., Giolitto, S., Font-Monclus, I., ... & Ronchi, A. E. (2021). The Coup-TFII orphan

nuclear receptor is an activator of the γ-globin gene. haematologica, 106(2), 474.

Jonk, L. J., de Jonge, M. E., Vervaart, J. M., Wissink, S., & Kruijer, W. (1994). Isolation and developmental expression of retinoic-acid-induced genes. Developmental biology, 161(2), 604-614.

Qiu, Y., Krishnan, V., Pereira, F. A., Tsai, S. Y., & Tsai, M. J. (1996). Chicken ovalbumin upstream promoter-transcription factors and their regulation. The Journal of steroid biochemistry and molecular biology, 56(1-6), 81-85.

Soosaar, A., Neuman, K., Nornes, H. O., & Neuman, T. (1996). Cell type specific regulation of COUP-TF II promoter activity. FEBS letters, 391(1-2), 95-100.

Krishnan, V., Elberg, G., Tsai, M. J., & Tsai, S. Y. (1997). Identification of a novel sonic hedgehog response element in the chicken ovalbumin upstream promoter-transcription factor II promoter. Molecular Endocrinology, 11(10), 1458-1466.

Petit, F. G., Salas, R., Tsai, M. J., & Tsai, S. Y. (2004). The regulation of COUP-TFII gene expression by Ets-1 is enhanced by the steroid receptor co-activators. Mechanisms of ageing and development, 125(10-11), 719-732.

Perilhou, A., Tourrel-Cuzin, C., Zhang, P., Kharroubi, I., Wang, H., Fauveau, V., ... & Vasseur-Cognet, M. (2008). The MODY1 Gene for Hepatocyte Nuclear Factor 4α and a Feedback Loop Control COUP-TFII Expression in Pancreatic Beta Cells. Molecular and cellular biology, 28(14), 4588-4597.

Kruse, S. W., Suino-Powell, K., Zhou, X. E., Kretschman, J. E., Reynolds, R., Vonrhein, C., ... & Xu, H. E. (2008). Identification of COUP-TFII orphan nuclear receptor as a retinoic acid–activated receptor. PLoS biology, 6(9), e227.

Le Guével, R., Oger, F., Martinez-Jimenez, C. P., Bizot, M., Gheeraert, C., Firmin, F., ... & Salbert, G. (2017). Inactivation of the nuclear orphan receptor COUP-TFII by small chemicals. ACS chemical biology, 12(3), 654-663.

Wang, L., Cheng, C. M., Qin, J., Xu, M., Kao, C. Y., Shi, J., ... & Tsai, M. J. (2020). Small-molecule inhibitor targeting orphan nuclear receptor COUP-TFII for prostate cancer treatment. Science Advances, 6(18), eaaz8031.

Wang, T., Wang, Z., de Fabritus, L., Tao, J., Saied, E. M., Lee, H. J., ... & Santori, F. R. (2021). 1-deoxysphingolipids bind to COUP-TF to modulate lymphatic and cardiac cell development. Developmental Cell, 56(22), 3128-3145.rnal of Clinical Medicine, 9(11), 3782.

Fishilevich, S., Nudel, R., Rappaport, N., Hadar, R., Plaschkes, I., Iny Stein, T., ... & Cohen, D. (2017). GeneHancer: genome-wide integration of enhancers and target genes in GeneCards. Database, 2017.

Baroukh, N., Ahituv, N., Chang, J., Shoukry, M., Afzal, V., Rubin, E. M., & Pennacchio, L. A. (2005). Comparative genomic analysis reveals a distant liver enhancer upstream of the COUP-TFII gene. Mammalian genome, 16, 91-95.

Grebien, F., Kerenyi, M. A., Kovacic, B., Kolbe, T., Becker, V., Dolznig, H., ... & Moriggl, R. (2008). Stat5 activation enables erythropoiesis in the absence of EpoR and Jak2. Blood, The Journal of the American Society of Hematology, 111(9), 4511-4522.

Tóthová, Z., Tomc, J., Debeljak, N., & Solár, P. (2021). STAT5 as a key protein of erythropoietin signalization. International journal of molecular sciences, 22(13), 7109.

Kassouf, M. T., Hughes, J. R., Taylor, S., McGowan, S. J., Soneji, S., Green, A. L., ... & Porcher, C. (2010). Genome-wide identification of TAL1's functional targets: insights into its mechanisms of action in primary erythroid cells. Genome research, 20(8), 1064-1083.

Andrews, N. C. (1994). Erythroid transcription factor NF-E2 coordinates hemoglobin synthesis. Pediatric research, 36(4), 419-423.

Mutschler, M., Magin, A. S., Buerge, M., Roelz, R., Schanne, D. H., Will, B., ... & Pahl, H. L. (2009). NF-E2 overexpression delays erythroid maturation and increases erythrocyte production. British journal of haematology, 146(2), 203-217.

Shimizu, R., Engel, J. D., & Yamamoto, M. (2008). GATA1-related leukaemias. Nature Reviews Cancer, 8(4), 279-287.

Romano, O., Petiti, L., Felix, T., Meneghini, V., Portafax, M., Antoniani, C., ... & Miccio, A. (2020). GATA factor-mediated gene regulation in human erythropoiesis. Iscience, 23(4), 101018.

Wu, W., Morrissey, C. S., Keller, C. A., Mishra, T., Pimkin, M., Blobel, G. A., ... & Hardison, R. C. (2014). Dynamic shifts in occupancy by TAL1 are guided by GATA factors and drive large-scale reprogramming of gene expression during hematopoiesis. Genome research, 24(12), 1945-1962.

Han, G. C., Vinayachandran, V., Bataille, A. R., Park, B., Chan-Salis, K. Y., Keller, C. A., ... & Pugh, B. F. (2016). Genome-wide organization of GATA1 and TAL1 determined at high resolution. Molecular and cellular biology, 36(1), 157-172.

Tsai, F. Y., & Orkin, S. H. (1997). Transcription factor GATA-2 is required for proliferation/survival of early hematopoietic cells and mast cell formation, but not for erythroid and myeloid terminal differentiation. Blood, The Journal of the American Society of Hematology, 89(10), 3636-3643.

Navas, P. A., Peterson, K. R., Li, Q., Skarpidi, E., Rohde, A., Shaw, S. E., ... & Stamatoyannopoulos, G. (1998). Developmental specificity of the interaction between the locus control region and embryonic or fetal globin genes in transgenic mice with an HS3 core deletion. Molecular and cellular biology, 18(7), 4188-4196.

Liu, N., Hargreaves, V. V., Zhu, Q., Kurland, J. V., Hong, J., Kim, W., ... & Orkin, S. H. (2018). Direct promoter repression by BCL11A controls the fetal to adult hemoglobin switch. Cell, 173(2), 430-442.

Jiang, J., Best, S., Menzel, S., Silver, N., Lai, M. I., Surdulescu, G. L., ... & Thein, S. L. (2006). cMYB is involved in the regulation of fetal hemoglobin production in adults. Blood, 108(3), 1077-1083.

Sankaran, V. G., Menne, T. F., Šćepanović, D., Vergilio, J. A., Ji, P., Kim, J., ... & Lodish, H. F. (2011). MicroRNA-15a and-16-1 act via MYB to

elevate fetal hemoglobin expression in human trisomy 13. Proceedings of the National Academy of Sciences, 108(4), 1519-1524.

Gu, Q., Palani, C. D., Smith, A., Li, B., Amos-Abanyie, E. K., Ogu, U., ... & Starlard-Davenport, A. (2022). MicroRNA29B induces fetal hemoglobin via inhibition of the HBG repressor protein MYB in vitro and in humanized sickle cell mice. Frontiers in Medicine.

Thein, S. L., Menzel, S., Peng, X., Best, S., Jiang, J., Close, J., ... & Lathrop, M. (2007). Intergenic variants of HBS1L-MYB are responsible for a major quantitative trait locus on chromosome 6q23 influencing fetal hemoglobin levels in adults. Proceedings of the National Academy of Sciences, 104(27), 11346-11351.

Stadhouders, R., Aktuna, S., Thongjuea, S., Aghajanirefah, A., Pourfarzad, F., van IJcken, W., ... & Soler, E. (2014). HBS1L-MYB intergenic variants modulate fetal hemoglobin via long-range MYB enhancers. The Journal of clinical investigation, 124(4), 1699-1710.

Marais, A. D. (2019). Apolipoprotein E in lipoprotein metabolism, health and cardiovascular disease. Pathology, 51(2), 165-176.

Turk, J., White, T. D., Nelson, A. J., Lei, X., & Ramanadham, S. (2019). iPLA2β and its role in male fertility, neurological disorders, metabolic disorders, and inflammation. Biochimica et Biophysica Acta (BBA)-Molecular and Cell Biology of Lipids, 1864(6), 846-860.

Timson, D. J., & Reece, R. J. (2003). Identification and characterisation of human aldose 1-epimerase. FEBS letters, 543(1-3), 21-24.

Higashihori, N., Song, Y., & Richman, J. M. (2008). Expression and regulation of the decoy bone morphogenetic protein receptor BAMBI in the developing avian face. Developmental Dynamics: An Official Publication of the American Association of Anatomists, 237(5), 1500-1508.

Hernandez, M. C., Andres-Barquin, P. J., Martinez, S., Bulfone, A., Rubenstein, J. L., & Israel, M. A. (1997). ENC-1: A Novel MammalianKelch-Related Gene Specifically Expressed in the Nervous System Encodes an Actin-Binding Protein. Journal of Neuroscience, 17(9), 3038-3051.

Onichtchouk, D., Chen, Y. G., Dosch, R., Gawantka, V., Delius, H., & Niehrs, C. (1999). Silencing of TGF-β signalling by the pseudoreceptor BAMBI. Nature, 401(6752), 480-485.

Zermati, Y., Fichelson, S., Valensi, F., Freyssinier, J. M., Rouyer-Fessard, P., Cramer, E., ... & Hermine, O. (2000). Transforming growth factor inhibits erythropoiesis by blocking proliferation and accelerating differentiation of erythroid progenitors. Experimental hematology, 28(8), 885-894.

Sekiya, T., Adachi, S., Kohu, K., Yamada, T., Higuchi, O., Furukawa, Y., ... & Akiyama, T. (2004). Identification of BMP and activin membrane-bound inhibitor (BAMBI), an inhibitor of transforming growth factor-β signaling, as a target of the β-catenin pathway in colorectal tumor cells. Journal of Biological Chemistry, 279(8), 6840-6846.

Togo, N., Ohwada, S., Sakurai, S., Toya, H., Sakamoto, I., Yamada, T., ... & Akiyama, T. (2008). Prognostic significance of BMP and activin membrane-bound inhibitor in colorectal cancer. World journal of gastroenterology: WJG, 14(31), 4880.

Fritzmann, J., Morkel, M., Besser, D., Budczies, J., Kosel, F., Brembeck, F. H., ... & Birchmeier, W. (2009). A colorectal cancer expression profile that includes transforming growth factor β inhibitor BAMBI predicts metastatic potential. Gastroenterology, 137(1), 165-175.

Erdős, E., & Bálint, B. L. (2020). NR2F2 orphan nuclear receptor is involved in estrogen receptor alpha-mediated transcriptional regulation in luminal a breast cancer cells. International Journal of Molecular Sciences, 21(6), 1910.

Kinghorn, K. J., Castillo-Quan, J. I., Bartolome, F., Angelova, P. R., Li, L., Pope, S., ... & Partridge, L. (2015). Loss of PLA2G6 leads to elevated mitochondrial lipid peroxidation and mitochondrial dysfunction. Brain, 138(7), 1801-1816.

Morgan, N. V., Westaway, S. K., Morton, J. E., Gregory, A., Gissen, P., Sonek, S., ... & Hayflick, S. J. (2006). PLA2G6, encoding a phospholipase A2, is mutated in neurodegenerative disorders with high brain iron. Nature genetics, 38(7), 752-754.

Schaeffer, E. L., & Gattaz, W. F. (2008). Cholinergic and glutamatergic alterations beginning at the early stages of Alzheimer disease: participation of the phospholipase A 2 enzyme. Psychopharmacology, 198, 1-27.

Paisán-Ruiz, C., Li, A., Schneider, S. A., Holton, J. L., Johnson, R., Kidd, D., ... & Houlden, H. (2012). Widespread Lewy body and tau accumulation in childhood and adult onset dystonia-parkinsonism cases with PLA2G6 mutations. Neurobiology of aging, 33(4), 814-823.

Xiao, W., Koizumi, K., Nishio, M., Endo, T., Osawa, M., Fujimoto, K., ... & Sawada, K. I. (2002). Tumor necrosis factor-α inhibits generation of glycophorin A+ cells by CD34+ cells. Experimental hematology, 30(11), 1238-1247.

Giau, V. V., Bagyinszky, E., An, S. S. A., & Kim, S. Y. (2015). Role of apolipoprotein E in neurodegenerative diseases. Neuropsychiatric disease and treatment, 1723-1737.

Husain, M. A., Laurent, B., & Plourde, M. (2021). APOE and Alzheimer's disease: from lipid transport to physiopathology and therapeutics. Frontiers in Neuroscience, 15, 630502.

Kockx, M., Traini, M., & Kritharides, L. (2018). Cell-specific production, secretion, and function of apolipoprotein E. Journal of molecular medicine, 96, 361-371.

Simonet, W. S., Bucay, N., Lauer, S. J., & Taylor, J. M. (1993). A far-downstream hepatocyte-specific control region directs expression of the linked human apolipoprotein E and CI genes in transgenic mice. Journal of Biological Chemistry, 268(11), 8221-8229.

Shachter, N. S., Zhu, Y. U. Y. I. N. G., Walsh, A., Breslow, J. L., & Smith, J. D. (1993). Localization of a liver-specific enhancer in the apolipoprotein E/CI/C-II gene locus. Journal of lipid research, 34(10), 1699-1707.

Allan, C. M., Walker, D., & Taylor, J. M. (1995). Evolutionary Duplication of a Hepatic Control Region in the Human Apolipoprotein E Gene Locus: IDENTIFICATION OF A SECOND REGION THAT CONFERS HIGH LEVEL AND LIVER-SPECIFIC EXPRESSION OF THE HUMAN

APOLIPOPROTEIN E GENE IN TRANSGENIC MICE (∗). Journal of Biological Chemistry, 270(44), 26278-26281.

Allan, C. M., Taylor, S., & Taylor, J. M. (1997). Two hepatic enhancers, HCR. 1 and HCR. 2, coordinate the liver expression of the entire human apolipoprotein E/CI/C-IV/C-II gene cluster. Journal of Biological Chemistry, 272(46), 29113-29119.

Shih, S. J., Allan, C., Grehan, S., Tse, E., Moran, C., & Taylor, J. M. (2000). Duplicated downstream enhancers control expression of the human apolipoprotein E gene in macrophages and adipose tissue. Journal of Biological Chemistry, 275(41), 31567-31572.

S., Tse, E., & Taylor, J. M. (2001). Two distal downstream enhancers direct expression of the human apolipoprotein E gene to astrocytes in the brain. Journal of Neuroscience, 21(3), 812-822.

Zheng, P., Pennacchio, L. A., Le Goff, W., Rubin, E. M., & Smith, J. D. (2004). Identification of a novel enhancer of brain expression near the apoE gene cluster by comparative genomics. Biochimica et Biophysica Acta (BBA)-Gene Structure and Expression, 1676(1), 41-50.

Liang, Y., Lin, S., Beyer, T. P., Zhang, Y., Wu, X., Bales, K. R., ... & Paul, S. M. (2004). A liver X receptor and retinoid X receptor heterodimer mediates apolipoprotein E expression, secretion and cholesterol homeostasis in astrocytes. Journal of neurochemistry, 88(3), 623-634.

Trusca, V. G., Fuior, E. V., Florea, I. C., Kardassis, D., Simionescu, M., & Gafencu, A. V. (2011). Macrophage-specific up-regulation of apolipoprotein E gene expression by STAT1 is achieved via long range genomic interactions. Journal of Biological Chemistry, 286(16), 13891-13904.

Trusca, V. G., Florea, I. C., Kardassis, D., & Gafencu, A. V. (2012). STAT1 interacts with RXRα to upregulate ApoCII gene expression in macrophages. PLoS One, 7(7), e40463

Xia, Y., Wang, Z. H., Zhang, J., Liu, X., Yu, S. P., Ye, K. X., ... & Wang, X. C. (2021). C/EBPβ is a key transcription factor for APOE and preferentially mediates ApoE4 expression in Alzheimer's disease. Molecular Psychiatry, 26(10), 6002-6022.

Kim, E., Xie, S., Yeh, S. D., Lee, Y. F., Collins, L. L., Hu, Y. C., ... & Chang, C. (2003). Disruption of TR4 orphan nuclear receptor reduces the expression of liver apolipoprotein E/CI/C-II gene cluster. Journal of Biological Chemistry, 278(47), 46919-46926.

Subramanian, S., Gottschalk, W. K., Kim, S. Y., Roses, A. D., & Chiba-Falek, O. (2017). The effects of PPARγ on the regulation of the TOMM40-APOE-C1 genes cluster. Biochimica et Biophysica Acta (BBA)-Molecular Basis of Disease, 1863(3), 810-816.

Shao, Y., Shaw, M., Todd, K., Khrestian, M., D'Aleo, G., Barnard, P. J., ... & Bekris, L. M. (2018). DNA methylation of TOMM40-APOE-APOC2 in Alzheimer's disease. Journal of human genetics, 63(4), 459-471.

Zhang, S. H., Reddick, R. L., Piedrahita, J. A., & Maeda, N. (1992). Spontaneous hypercholesterolemia and arterial lesions in mice lacking apolipoprotein E. Science, 258(5081), 468-471.

Jawien, J. (2012). The role of an experimental model of atherosclerosis: apoE-knockout mice in developing new drugs against atherogenesis. Current pharmaceutical biotechnology, 13(13), 2435-2439.

Pendse, A. A., Arbones-Mainar, J. M., Johnson, L. A., Altenburg, M. K., & Maeda, N. (2009). Apolipoprotein E knock-out and knock-in mice: atherosclerosis, metabolic syndrome, and beyond. Journal of lipid research, 50, S178-S182.

Ramassamy, C., Averill, D., Beffert, U., Bastianetto, S., Theroux, L., Lussier-Cacan, S., ... & Poirier, J. (1999). Oxidative damage and protection by antioxidants in the frontal cortex of Alzheimer's disease is related to the apolipoprotein E genotype. Free Radical Biology and Medicine, 27(5-6), 544-553.

Piccarducci, R., Daniele, S., Fusi, J., Chico, L., Baldacci, F., Siciliano, G., ... & Martini, C. (2019). Impact of ApoE polymorphism and physical activity on plasma antioxidant capability and erythrocyte membranes. Antioxidants, 8(11), 538.

Miyazawa, T., Suzuki, T., Fujimoto, K., & Kinoshita, M. (1996). Age-related change of phosphatidylcholine hydroperoxide and phosphatidylethanolamine hydroperoxide levels in normal human red blood cells. Mechanisms of Ageing and Development, 86(3), 145-150.

Saha, S. S., Chakraborty, A., Ghosh, S., & Ghosh, M. (2012). Comparative study of hypocholesterolemic and hypolipidemic effects of conjugated linolenic acid isomers against induced biochemical perturbations and aberration in erythrocyte membrane fluidity. European journal of nutrition, 51, 483-495.

Ortiz, G. G., Pacheco-Moisés, F. P., Flores-Alvarado, L. J., Macías-Islas, M. A., Velázquez-Brizuela, I. E., Ramírez-Anguiano, A. C., ... & Cortés-Enríquez, F. (2013). Alzheimer disease and metabolism: role of cholesterol and membrane fluidity. inMedicine, Mental and Behavioural Disorders and Diseases of the Nervous System, 145-173.

Holm, T. M., Braun, A., Trigatti, B. L., Brugnara, C., Sakamoto, M., Krieger, M., & Andrews, N. C. (2002). Failure of red blood cell maturation in mice with defects in the high-density lipoprotein receptor SR-BI. Blood, The Journal of the American Society of Hematology, 99(5), 1817-1824

Crespo, P., Delgado, M. D., Gómez-Casares, M. T., Cuadrado, M. A., Richard, C., & León, J. (1993). Induction of apolipoprotein E expression during erythroid differentiation of human K562 leukemia cells. Leukemia research, 17(9), 771-776.

Kulkarni, A. P., Mittal, S. P., Devasagayam, T. P., & Pal, J. K. (2009). Oxidative stress perturbs cell proliferation in human K562 cells by modulating protein synthesis and cell cycle. Free radical research, 43(11), 1090-1100.

Acknowledgments

In the following lines, I would like to express my gratitude to all the wonderful individuals who supported me during the past three years, and before.

Firstly, I would like to thank my supervisor Professor Antonella Ronchi for giving me the opportunity to do the PhD in her lab and to be a part of the ARCH MSCA Project. I am thankful for all the knowledge and experience that she shared with me, her advice, and her guidance. I would also like to thank all members of my lab and colleagues from terzo piano, especially Dr. Gloria Barbarani who was a great guide from my very first day in the 3021 lab, and Dr. Valentina Pastori who was always there to help, bringing peace and harmony in our lab. I am also glad to have had two brilliant students, Serena Pozzi and Alessia Abbiati, that lit up my every day there.

Next, I would like to thank our collaborators: Dr. Emanuele Azzoni and Dr. Cristiana Barone from UNIMIB for helping me to perform immunofluorescence experiments, Dr. Rocco Piazza and Dr. Luca Mologni from UNIMIB for supporting me with the bioinformatics analysis of RNA-seq and ChIP-seq, and Dr. Claudio Cantù and Gianluca Zambanini from Linköping University for performing the CUT&RUN experiment.

I would also like to express my gratitude to all the ARCH consortium, it had been great to be part of it! To the whole Flowmetric Europe team, especially to Dr. Alessandra Roberto and Christina Pitsillidou for sharing their immense knowledge and experience on flow cytometry and for the time that they spent designing with me the experiments. To Professor Florian Grebien and Ludovica Proietti from Vetmeduni Vienna for their great help with all CRISPR/Cas9 experiments presented in this thesis. I am truly thankful to Florian for his support, advice and kindness, and to Ludo for her help with all my experiments and patience when I was overloading her with questions and doubts.

Last but not least, to the whole Grebien lab where I spent three amazing months.

I would also like to thank Dr. Elvan Böke and Dr. Kinga Majchrzak-Kuligowska with whom I had a chance to work during MSc study and who inspired me to try myself in research.

I am thankful to Agnieszka Kaczmar, my therapist, that helped me to find a balance and motivation during those challenging three years.

To my family: my mum, my dad and my brother that were supporting me, sharing my frustrations and accepting my decisions. I am really happy that I could always come back home and you have always been there, in my beloved Halinów.

To my Italian, Cypriot, Spanish and Polish friends. Especially, to Francesca and Nick, my lovely Milanese neighbours, to Chris and Chris(tina) that always made me laugh, it has been great to have you all around! Cri, you were an incredible support during those three years understanding and sharing my every worry and struggle. I am grateful for every single beer at La Buttiga that we had together (and there were a lot)! I am grateful to Sofia, thanks to her I discovered the Marie Curie program. To Blanca, my Spanish soulmate, for her warm support and that she always believed that I can make it. To my forever friends, Joasia, Marlenka and Hania, to my university friends Kasiasiasia, Monia and Gosia, and to K(aro)lina, my very first teacher, who was teaching me maths when I was 4, and Italian when I was 15 (and she did not know yet how much it will be helpful when doing a PhD at an Italian university!).

Lastly, I dedicate all this work to myself. I am so proud that I made it despite all difficulties and I hope it will remind me that I am enough, and I always was.

HemaSphere

EUROPEAN HEMATOLOGY ASSOCIATION

Editorial
Open Access

A Journey Into the Unknown: PhD Students in a European Training Network on Age-related Changes in Hematopoiesis Conduct Their Project During a Global Pandemic

Christina Pitsillidou[1,2], Sandra Alonso-Rubido[1,3], Andrea Ávila-Ávila[4,5], Mari Carmen Romero-Mulero[6], Agata Labedz[1], Athanasios Oikonomou[7,8], Ludovica Proietti[9], Maria Eleni Psychoyiou[10], Clara Tellez-Quijorna[11], Maniriho Hillary[12], Eirini Sofia Fasouli[1,13], Guillermo Fernández-Rodriguez[14], Natalia Giner-Laguarda[15], Natalia Skinder[16], Chiara Taroni[17], John Strouboulis[10], Eleni Katsantoni[13], Antonella Ellena Ronchi[1]

Correspondence: Christina Pitsillidou (c.pitsillidou@campus.unimib.it).

The age-related changes in hematopoiesis (ARCH) project is part of the Innovative Training Network (ITN) of the Marie-Sklodowska Curie Actions (MSCA) program, which provides doctorate training of excellence based on the exchange of ideas and competencies from the academic and private sectors.[1] The ARCH project intends to outline hematopoietic stem cell (HSC) alterations with aging, how hematopoietic cell individuality is controlled at the transcriptional and epigenetic levels in normal hematopoiesis and in leukemias, and understand the crosstalk between intrinsic and extrinsic indications that support the proliferation of preleukemic and leukemic cells within the hematopoietic niche.

We are 15 PhD students funded by this network, based around Europe, and our common aim is to understand functional changes in the hematopoietic system with age, how these changes link to the development of age-associated diseases and in parallel work towards the development of new treatments.[2] Our projects kicked off just when the severe acute respiratory coronavirus-2 (SARS-CoV-2) emerged. Two and a half years later, SARS-CoV-2 continues to infect millions of people and has taken the lives of at least 6 million people worldwide.[3] The COVID-19 outbreak brought along social isolation and feelings of uncertainty to everyone around the world, including doctorate students.[4,5] Ironically, our projects have been more relevant than ever, as the pandemic has highlighted the important relationship between age-related changes in hematopoiesis and disease severity. Below, we aim to discuss the timeline of the ARCH project throughout the pandemic and how we managed to courageously pull through the hardships of doing research during a global pandemic within different settings (academia/institutes and industry). We provide recommendations to future PhD students on how to manage their PhD projects during global emergencies.

PANDEMIC TIMELINE OF ARCH

The rapid spread of SARS-CoV-2 at the beginning of 2020 in Europe led the World Health Organization to declare a global pandemic on March 11, 2020. The ARCH opening conference,

[1]Dipartimento di Biotecnologie e Bioscienze, Università degli Studi Milano-Bicocca, Milano, Italy
[2]FlowMetric Europe, Milan, Bresso, Italy
[3]Diagenode SA, Ougrée, Belgium
[4]Institute Curie, Orsay, France
[5]Centre National de la Recherche Scientifique UMR3348, Centre Universitaire, Orsay, France
[6]Max Planck Institute of Immunobiology and Epigenetics, Freiburg, Germany
[7]Fondazione Tettamanti, Centro di Ricerca Tettamanti, Centro Maria Letizia Verga, Monza (MB), Italy
[8]Dipartimento di Medicina e Chirurgia, Università degli Studi Milano-Bicocca, Monza (MB), Italy
[9]Institute of Medical Biochemistry, University of Veterinary Medicine, Vienna, Austria
[10]Molecular Haematology, Comprehensive Cancer Centre, School of Cancer and Pharmaceutical Sciences, King's College London, United Kingdom
[11]Team of Epigenetics in Normal and Abnormal Hematopoiesis, CRCM (Centre de Recherche en Cancèrologie de Marseille), France
[12]Department of Human Molecular Genetics and Biochemistry, Sackler Faculty of Medicine, Tel Aviv University, Tel Aviv-Yafo and Schneider Children's Medical Centre of Israel, Petach, Tikvah
[13]Basic Research Center, Biomedical Research Foundation, Academy of Athens, Greece
[14]Department of Biology and Biotechnology "Charles Darwin", Sapienza University of Rome, Italy
[15]Department of Cellular and Molecular Biology, Centro de Investigaciones Biológicas Margarita Salas, Madrid, Spain
[16]European Research Institute for the Biology of Ageing, University Medical Center Groningen, University of Groningen, Netherlands
[17]Institut de Génétique et de Biologie Moléculaire et Cellulaire (IGBMC), Institut National de la Santé et de la Recherche Médicale (INSERM) U1258, Centre National de la Recherche Scientifique (CNRS) UMR 7104, Université de Strasbourg, Illkirch, France
CP, SA-R, AÁ-Á, and MCR-M have contributed equally to this work.

HemaSphere (2022) 6:8(e763).
http://dx.doi.org/10.1097/HS9.0000000000000763.

planned for the same month was canceled. At the same time, the first lockdowns were put in place around Europe, while most of us were just starting our new project or were moving to a different country to begin. Some of us started on this project during or after the first lockdown, causing further delays, which led to a multiphasic pattern of recruitment (Figure 1). During the first lockdown, we individually set up our strategies to progress our research projects at home by either reading scientific literature, writing review papers,[6-9] or planning experiments for when we would eventually return to work. By summer 2020, many of us began to slowly get back to the laboratory in shifts or for a limited time to avoid overcrowding of spaces. Times were still uncertain and the pandemic was of great concern to the global scientific community. With the whole world waiting for vaccine developments to stop the pandemic, our slow training process commenced. However, by fall, new lockdown measures were set against a new surge of infections and deaths (Figure 1). This second lockdown stretched until Christmas, marking the completion of our first year with minimal hands-on training and very limited experimental progress. This brings us to today, where most of us have entered the final year of our projects and it seems that COVID-19 is here to stay.

COVID-19: PHDS INTO THE UNKNOWN

Engaging in a PhD project on its own is stressful and overwhelming. The main objectives when beginning a 3-year PhD project are to develop new laboratory, academic, and research-based skills, yet we were unable to advance on any of these. Close and continuous supervision is required to meet these goals. Communication with supervisors at the beginning of our PhDs, which coincided with the pandemic, made our start more challenging than expected. Coming out of isolation in the summer of 2020, we did not expect those upcoming months would continue to affect our projects and to make us experience how vulnerable we are.

Working hours were adjusted to allow for social distancing, while meetings were still being held online. Even when we were physically at the laboratory, there was an inability for close supervision due to safety restrictions. In addition to communication challenges, technical difficulties were also at play during the first lockdown. For example, students that had already begun to breed mice a few weeks before the lockdown were forced to interrupt their experiments and redirect their projects as it would take too much time to re-establish mouse breeding. The high demand for equipment, reagents, and consumables for the advancement of COVID-19 research and diagnosis added to

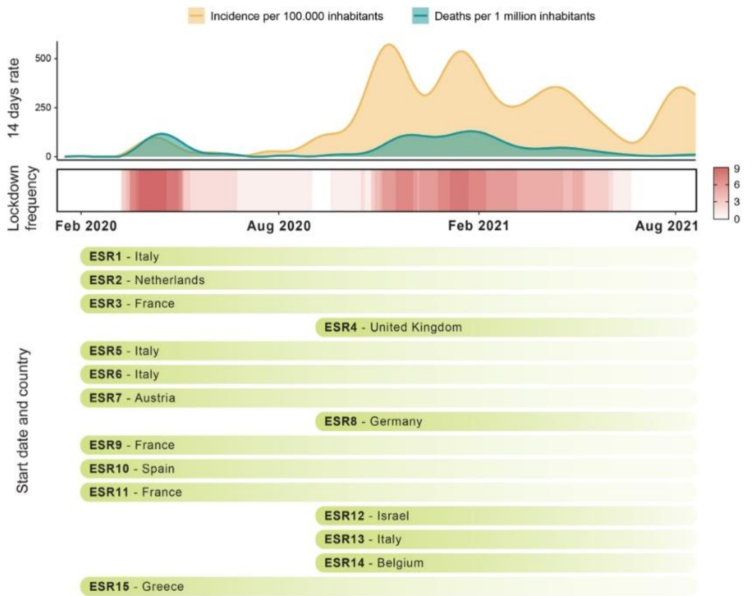


Figure 1. Timeline of the ARCH project during the COVID-19 pandemic. Average 14-day notification rates of newly reported COVID-19 cases per 100,000 inhabitants and COVID-19 related deaths per 1 million inhabitants in the participating countries of the ARCH project are represented during the first year and a half of the pandemic in Europe (top graph). By fall 2020, there is an all-time high of the infection and death rates. The lockdown frequency considered as the daily number of ARCH participating countries under measurements of lockdown is shown in a heat-map along the same period (middle graph). Starting time and country of the beneficiary institution where each early stage researcher (ESR) performs the PhD project is shown in bars (bottom graph). These data have been collected from official governmental reports from each country. The lockdown definition and concrete measures differ between countries and regions within the countries.

these difficulties and made it challenging for us to obtain the necessary consumables for our research. Reagents and consumables were not the only shortage, but also sample availability. As our projects are based on hematology, many of our experiments require fresh samples from healthy or diseased participants to progress. Such fresh materials were merely impossible to find as clinics halted unnecessary visits and hospitals gave sole priority to COVID-19 patients. Again, this inevitably led to the redirection of some projects and added to the uncertainties and delays of the projects.

Some institutions acknowledged the loss of time for many PhD students and offered time extensions for thesis submissions. In some institutions, time extensions were only applicable to students that were enrolled in their second and third year, a situation that we consider is not justified as the pandemic had a strong impact on all PhD projects regardless of the year of enrollment. However, those that were offered an extension were offered no additional financial support. This made us feel that our PhD training status was overlooked and shows once more that PhD positions are a precarious form of employment. As these are typically 3 to 4 year training periods, where PhD students need to provide evidence of strong scientific skills, not being able to work is a major limitation not only to obtain the PhD degree, but also for the success in making next step in the career.

COVID-19 repercussions created disruptions also to life science companies. Twenty percent of ARCH students work in the private sector. Companies had to adapt to the COVID-19 pandemic to survive, thus made COVID-19 research and treatment innovation their uppermost priority. Industry supervisors were understandably busy involved in clinical research projects to meet client requests. Lockdown measures stalled many investigations, having a domino effect on doctorate projects based there. At the same time, the detachment of industry PhDs from research institutions and universities negatively affected their learning progress.

The main scientific activities were mostly virtual and we missed the stimulus of having an audience present, while many such events were inadequately organized and prone to technical glitches. Virtual learning left gaps in milestone competencies like collaboration and the development of a critical approach to research questions and results. However, we are thankful for online platforms. Regardless of the complexity of our position, many of us took advantage of the lockdown by completing online courses and taking part in virtual conferences. Virtual platforms at least allowed us to have some communication and to develop more valuable computational skills than ever before, which were also applied to our projects.

WHAT PHD STUDENTS CAN DO DURING GLOBAL EMERGENCIES?

Doctorate students aim to find the answers to complex scientific problems and need to excel in critical thinking and correct decision-making, which already imposes pressure on them. Experiencing external pressures through a global pandemic makes it harder to face the challenges associated with research. Here we offer our advice to students that could be experiencing similar global emergencies:

Mental wellbeing is one of the biggest challenges one will face when conducting a PhD project in such peculiar times. Even under the best circumstances, one cannot be productive at all times. Therefore, we recommend PhD students to avoid harsh self-criticism if they are lack motivation. We strongly encourage PhD students to share feelings with their peers and colleagues in the same consortium or institution and thus to be open and talk about all problems. We created a group chat where we shared our thoughts and worries. This exercise made us realize that we were not alone and others could relate to our emotions, which was comforting. It is important to rely on other students and researchers to find personal support. Do not hesitate to reach out to your supervisor for help regarding your mental wellbeing as this can have an impact on your work. Good supervision and communication are crucial for the proper development and success of your project.

To make the most out of a PhD project we recommend that PhD students take advantage of all opportunities and plan. Design experiments to be performed during times when you can enter the laboratory and plan to generate enough data that you can analyze if a new lockdowns arises. We recommend students to start writing anything related to their research field. For example, write the introduction of the PhD thesis, methodologies that will be used, write a review article, and so on. PhD students will have to write at some point, so they should not hesitate to start even with no results or experience in the field they have just embarked on. Additionally, online seminars can be followed as well as online courses, for example to develop new computational or experimental skills.

CONCLUSION

Despite all the difficulties, we feel fortunate to be part of the ARCH MSCA ITN program. In December 2021, after 2 years of travel restrictions, we finally met for the first time in person in Rome. Here students and supervisors used this opportunity to begin collaborations, present projects, discuss, and overcome limitations.

The COVID-19 pandemic continues to have a significant impact on scientists. PhD students completing their projects throughout the global pandemic have suffered immense isolation and anxiety. Despite recurring delays throughout our project, we feel we have emerged as efficient workers, excellent planners and have developed strong social skills virtually and in person. Those experiencing similar global emergencies should share their experience with peers and colleagues, and find ways to efficiently transit towards effective solutions.

Finally, we hope that PhD students experiencing such situations are acknowledged for the time lost and hope that funding agencies recognize that time extensions should be accompanied by financial compensation.

DISCLOSURES

The authors have no conflicts of interest to disclose.

SOURCE OF FUNDING

This project has received funding from the European Union's Horizon 2020 Research and Innovation Program under the Marie Skłodowska-Curie grant agreement no. 813091 (ARCH, age-related changes in hematopoiesis).

REFERENCES

1. Marie Sklodowska-Curie Actions. *MSCA: Developing talents, advancing research*. [Online]. Available at: https://marie-sklodowska-curie-actions.ec.europa.eu Accessed April 12, 2022.
2. Age-related changes in hematopoiesis: ARCH project. [Online]. Available at: https://arch-project.eu/ Accessed March 20, 2022.
3. Wang H, Paulson K, Pease S, Watson S, Comfort H, Zheng P et al. Estimating excess mortality due to the COVID-19 pandemic: a systematic analysis of COVID-19-related mortality, 2020–21. *Lancet*. 2022;399:1513–1536.
4. Keszei Z, Vendrell X, Soultsioti M, Molina P, Wirth L, Weiss H, et al. 100 days of solitude: The spring of COVID-19 through the eyes of 15 young virologists of the INITIATE program. *Virus Res*. 2020;287:198093.
5. Gavriilaki E, Maiques-Diaz A. "Long COVID-19" of researchers: what to do next? *HemaSphere*. 2022;6:e673.
6. Barbarani G, Łabedz A, Ronchi A. β-Hemoglobinopathies: the test bench for genome editing-based therapeutic strategies. *Front Genome Ed*. 2020;2:571239.
7. Barbarani G, Labedz A, Stucchi S, Abbiati A, Ronchi A. Physiological and aberrant γ-globin transcription during development. *Front Cell Dev Biol*. 2021;9:640060.
8. Giner-Laguarda N, Vidal M. Functions of polycomb proteins on active targets. *Epigenomes*. 2020;4:17.
9. Fasouli E, Katsantoni E. JAK-STAT in early hematopoiesis and leukemia. *Front Cell Dev Biol*. 2021;9:669363.