KLF1 in Erythroid Differentiation:
An essential transcription factor in health and disease

Ileana Cantù

ISBN: 978-94-6299-170-5

Cover: Ileana Cantú
Image from ©iStockphoto.com/kotoffei
Text from van Leeuwenhoek's first published observation of red blood cells, 1674
Layout: Ileana Cantú

Printing: Ridderprint BV, the Netherlands

The research presented in this thesis was performed at the Department of Cell Biology at the Erasmus University Medical Center, Rotterdam, the Netherlands, and was supported by the Landsteiner Foundation for Blood Transfusion Research (LSBR 1040), the Netherlands Genomics Initiative (NGI Zenith 93511036), the Netherlands Organization for Scientific Research (NWO/ZonMw 40-00812-98-12128 and DN 82-301), and EU fp7 Specific Cooperation Research Project THALAMOSS (306201).

Ai miei genitori

KLF1 in Erythroid Differentiation: An essential transcription factor in health and disease

KLF1 in erytroïde differentiatie:
een essentiële transcriptiefactor in ziekte en gezondheid

Thesis

to obtain the degree of doctor from the
Erasmus Universiteit Rotterdam
by command of the rector magnificus

Prof.dr. H.A.P. Pols

and in accordance with the decision of the Doctorate Board.

the public defense shall be held on
Wednesday 4 November 2015 at 11:30

by

Ileana Cantù
born in Seregno, Italy

Erasmus University Rotterdam

Erasmus

DOCTORAL COMMITTEE

Promotor: Prof.dr. J.N.J. Philipsen

Other members: Prof.dr. F.G. Grosveld
Prof.dr. M. von Lindern
Prof.dr. I.P. Touw

Copromotor: Dr. T.B. van Dijk

E quindi uscimmo a riveder le stelle

Dante Alighieri,Inferno XXXIV, 139

LIST OF ABBREVIATIONS

3C-seq Chromosome conformation capture coupled to next-generation sequencing
Bcl11a B cell CLL/lymphoma 11A
BFU-E Burst Forming Unit - Erytrhoid
bp base pairs
BSDC1 BSD domain containing protein 1
CFU Colony Forming Unit - Erytrhoid
ChIP Chromatin immunoprecipitation
CLP Common lymphoid progenitor
CMP Common myeloid progenitor
DNA Deoxyribonucleic acid
E12.5 Embryonic day 12.5
EPO Erythroipoietin
EPO-R Erythroipoietin receptor
ER Endoplasmatic reticulum
FACS Fluorescence-activated cell sorting
FB Fetal Brain
GATA 1 GATA binding protein 1
GMP Granulocyte-macrophage progenitor
HbA Adult hemoglobin
HbF Fetal hemoglobin
HSC Hematopoietic stem cell
HU Hydroxyurea
kb kilobases
KLF1 Kruppel-like factor 1 (erythroid)
LCR Locus control region
LT-HSC Long term hematopoietic stem cell
MEL Mouse erythroid leukemia cells
MEP Megakaryocyte-eyrhtoid progenitor
MFL Mouse Fetal liver
MPP Multipotent progenitor
MS Mass-spectometry
PCR Polymerase chain reaction
RBC Red blood cell
RNA Ribonucleid acid
ST-HSC Short term hematopoietic stem cell
WT Wild type
XPO7 Exportin 7

TABLE OF CONTENTS

Chapter 1

Introduction and Scope of the thesis

HEMATOPOIESIS

Blood is a specialized tissue responsible for the distribution of essential substances, such as nutrients and oxygen, to somatic cells, and for the removal of waste products, such as carbon dioxide, from those same cells. The human blood system accounts for 7% of the total body mass, with an approximate volume of 4-5 liters in the average adult.

Plasma is a protein-salt solution composed of water (92%), proteins (8%), and trace amounts of other materials, that occupies about 55% of the blood volume[1]. The major plasma proteins are albumin (the main protein constituent, responsible for maintaining colloid osmotic pressure), fibrinogen (responsible, among others, for the formation of blood clots) and globulins (including antibodies). Plasma has critical functions, such as maintaining blood pressure and volume, supplying proteins for blood clotting and immunity, and preserving a proper pH balance in the body. Plasma also circulates essential nutrients, like glucose, amino acids and fatty acids.

The remaining 45% is composed of different cell types: red blood cells (erythrocytes), white blood cells (leukocytes), and platelets (thrombocytes)[2] (Figure 1). By far, the most abundant are the red blood cells, occupying about 44% of blood total volume. In mammals, mature erythrocytes are flexible biconcave disks that have lost the cell nucleus and organelles. They are responsible for the transport of oxygen from the lungs to all tissues and the removal of carbon dioxide[1]. Their role will be discussed in more details later in this introduction. The remaining blood cell types are less abundant, but modulate important biological functions. Platelets constitute the major component of blood clots and originate from cytoplasmic fragmentation of large cells called megakaryocytes[3]. Lymphocytes, divided into B- or T-lymphocytes and Natural Killer (NK) cells, are cells of the immune system implicated in protecting the body against infections and foreign material. They perform this task by either activating other cells that eliminate infected cells (T cells and NK cells), or producing soluble immunoglobulins (antibodies) that promote the identification and elimination of foreign objects (B cells)[4]. Granulocytes, distinguished by the presence of cytoplasmic granules, are divided in neutrophils, eosinophils and basophils, and they represent the innate immune system, with a role in phagocytosis and the inflammatory response[4]. Macrophages and mast cells are not found in circulation but within tissues. Macrophages arise from circulating monocytes that enter damaged tissues through the endothelium of the blood vessel, where they terminally differentiate. They are responsible for the phagocytosis of dead cells, cellular debris and foreign substances[4]. Mast cells originate from unidentified progenitors of the hematopoietic tissue in the bone marrow. These progenitors migrate through the blood into other body sites, particularly the connective tissue, where they differentiate into mature mast cells. They are involved in wound healing and protection against pathogens, as well as in allergy and anaphylaxis[4]. Dendritic cells also arise from monocytes, and besides having roles similar to macrophages, they can also stimulate additional immune responses via

antigen presentation to lymphocytes[4].

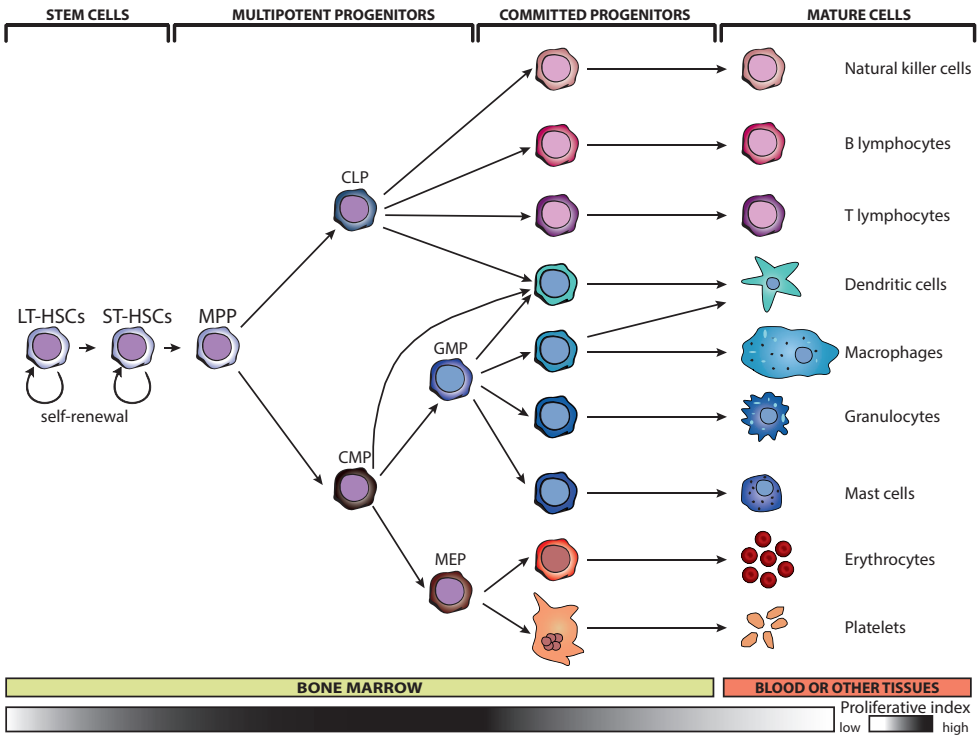


Figure 1. The hematopoietic system hierarchy. Schematic model of definitive hematopoietic development illustrating stages of differentiation. LT-HSCs (long term hematopoietic stem cells) give rise to all mature cells. The differentiation towards mature cells goes through several progenitor steps that become progressively restricted towards a specific lineage. CMPs, GMPs, CLPs and monocytes have all been reported to give rise to dendritic cells, even though the origin of dendritic cell has to be still further investigated. Cells during hematopoietic differentiation have different proliferation capability, and this is represented by the proliferation index. LT-HSCs are mainly quiescent, while multipotent progenitors are highly proliferating and this permits the generation of all the required blood cells. ST-HSCs; short term hematopoietic stem cell; MPP, multipotent progenitors with limited self-renewal; CMP, common myeloid progenitor; CLP, common lymphoid progenitor; GMP, granulocyte-macrophage progenitor; MEP, megakaryocyte-erythroid progenitor. Figure adapted and modified from ref[145].

ONTOGENY OF THE MAMMALIAN HEMATOPOIETIC SYSTEM

The hematopoietic system is one of the first major tissues to form, initially to make sure that the growing embryo has enough oxygenation and at later stages to additionally permit organogenesis[5]. The first wave of red blood cell production (also known as primitive hematopoiesis) occurs in the extraembryonic yolk sac and has the function of facilitating tissue oxygenation while the embryo rapidly grows. The large, nucleated red blood cells (RBCs) are produced within the blood islands of the yolk sac, which are pools of primitive erythroblasts derived from mesodermal cells[5]. Primitive erythroblasts enter circulation of the embryo where they replicate for several days[5]. It was believed that primitive erythroblasts remain nucleated throughout their entire life-span, but recent studies indicate that they

enucleate around E12.5-E16.5 while circulating in the mouse fetus[6,7]. The fingerprint of primitive erythroid cells is the expression of embryonic globin proteins (discussed in more detail later in this Chapter). Primitive hematopoiesis is transient and rapidly substituted by definitive hematopoiesis. Definitive hematopoiesis is responsible for the production of all hematopoietic cells present in the adult blood, including definitive erythrocytes, characterized by the lack of a nucleus and by the expression of adult globins[5]. As the anatomy of the embryo changes during development, hematopoiesis occurs in different anatomical sites. Furthermore, the hematopoietic activity often occurs simultaneously in more than one site[8,9] (Figure 2). For example, during adult life, hematopoiesis involves the colonization of thymus, spleen, lymph nodes and bone marrow.

Throughout ontogeny, the hematopoietic system requires constant production of blood cells. This entirely depends on the hematopoietic stem cells (HSCs) and their ability to self-renew (hence forming another HSC) and to differentiate into all blood lineages[10]. The first experimental evidence for the existence of HSCs in the bone marrow was published in 1961 by Till and McColluch[11]. It was then shown that in irradiated mice transplanted with bone marrow cells derived from a wild type mouse, discrete multi-lineage blood cell colonies were formed in the spleen. HSCs are extremely rare cells: for example, about 1 to 10 HSCs per 100.000 cells are estimated to exist in the mouse bone marrow[12]. Despite more than 50 years of research, it is still not possible to identify HSCs using a unique marker and to purify them at absolute purity, but technological advances now allow the purification of HSCs to a ratio of 1:2 (to total purified cells) in mice[13] or 1:4 in humans[14]. For the time being, a true HSC can only be identified by its ability to self-renew, and hence fully reconstitute the hematopoietic system in irradiated mice, both in primary and in secondary transplantations. HSCs have different characteristics, due to different developmental stages and requirements of the organism, and reflect different properties of the niches that surround them. For example, HSCs in the fetal liver are expanding, whereas HSCs in the adult bone marrow are quiescent (steady-state). Indeed, during homeostasis, most adult HSCs rarely divide, and when doing so, it is to renew the HSC pool and to ensure that there is a suitable amount of differentiated blood cells[15]. The bone marrow contains specialized

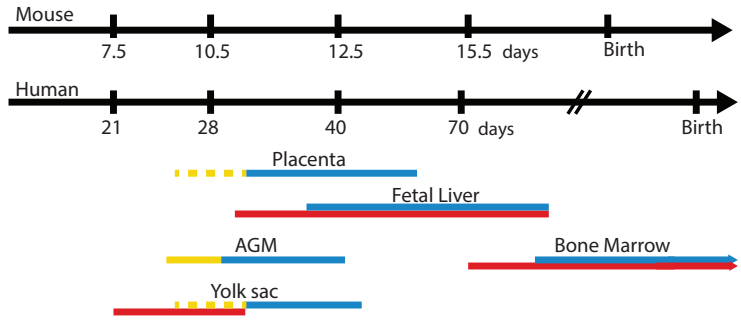


Figure 2. Human major anatomical sites of hematopoiesis during development. Different mouse/human hematopoietic sites and representation of the period during which they are active. In humans, hematopoiesis occurs mainly in the yolk sac, then in the fetal liver and finally in the bone marrow. AGM (aorta-gonad-mesonephros) and placenta also serve as transient hematopoietic organ. Red bars, active hematopoietic differentiation; yellow bars, HSC genesis; blue bars, presence of functional adult-type HSCs. Dotted yellow bars for yolk sac and placenta indicate that de-novo HSC genesis has not been experimentally proven. Figure adapted and modified from ref[146].

niches that make sure that HSCs will retain multipotency through cell divisions, whereas the progeny will be funneled to differentiation into specific lineages[16]. It is thought that HSC pool maintenance and lineage differentiation are achieved either via environmental asymmetry or by asymmetric self-renewal[16]. The former occurs when a daughter cell exits the niche, sustaining HSC self-renewal, and is then subjected to a microenvironment that induces differentiation, whereas the other daughter cells remains in the niche. The latter takes place when certain cell fate determinants are unevenly allocated to the two arising cells[16]. Nevertheless, these mechanisms have not been demonstrated in any vertebrate stem-cell type *in vivo*, and further studies are needed to prove whether these two models of asymmetric cell division exist.

Progenitors originating from HSCs lose the self-renewal ability and develop into multipotent progenitors. Progenitor differentiation takes place with sequential commitment steps, each further restraining to a certain hematopoietic lineage. The nature of such steps, and the specific lineage progenitors that actually exist, are a matter of debate and several models have been suggested[17,18]. The most widely accepted model suggests that HSCs with long-term regeneration capacity (characterized by bone marrow transplantation experiments, LT-HSCs) produce short-term HSCs (ST-HSCs). This HSC forms a multipotent progenitor (MPP)[19] that retains its multipotency, but has lost most of its self-renewal capacity. Subsequently, the first commitment step gives rise to Common Myeloid Progenitors (CMPs) [20] or Common Lymphoid Progenitors (CLPs)[21] (Figure 1). CLPs are restricted to the lymphoid lineage and produce exclusively B-cells, T-cells and natural killer cells, while CMPs give rise to megakaryocyte/erythroid progenitors (MEPs) and granulocyte/macrophage progenitors (GMPs). Subsequently GMPs produce the committed precursors of mast cells, granulocytes, and macrophages and of a specific group of myeloid derived dendritic cell, while MEPs produce the committed precursor of red blood cells and megakaryocytes (Figure 1).

It is important to keep in mind that this is a simplified and incomplete model of hematopoietic differentiation. Over the years various alternative and additional intermediate stages have been identified (for example[17,18]) and most likely more will be identified in the future.

ERYTHROPOIESIS

Erythropoiesis is the process responsible for the production of red blood cells (RBC, erythrocytes), and is part of the general process of hematopoiesis[22]. RBCs are flexible biconcave disks that have lost every cellular component (including the nucleus) not needed to fulfill their functions, therefore essentially constituting hemoglobin boxes. Erythrocytes have an average lifespan of 60 days in mice and 120 in humans. Hence, in order to guarantee tissues with a constant supply of oxygen, mature erythrocytes have to be continuously generated. Like all cells in the hematopoietic system, erythrocytes are derived from HSCs. Erythropoiesis comprises several programmed differentiation steps between HSCs and mature erythrocytes (Figure 3). The Burst Forming Unit-Erythroid (BFU-E) is the earliest erythroid-restricted progenitor to form, and is identified by a functional assay [23]: in the presence of erythropoietin (Epo), interleukin 3 (IL-3), granulocyte-macrophage colony stimulating factor (GM-CSF), thrombopoietin and stem cell factor (SCF), progenitor cell s produce large colonies containing more than a thousand hemoglobinized erythroblasts after 5-7 (mouse cells) or 14-16 (human cells) days of culture. The Colony Forming Unit-Erythroid (CFU-E), a more mature erythroid progenitor, is a cell type defined by the same colony assay, and can give rise to small colonies containing up to 64 erythroid cells after 2-4 (mouse) or 5-8 (human) days of culture[23,24]. CFU-E are only 4-5 cell divisions upstream of mature RBCs[25]. It has recently become evident that BFU-E and CFU-E can self-renew in response to acute

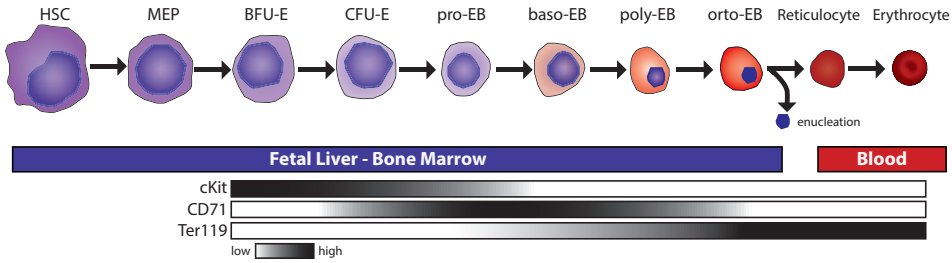


Figure 3. Stages of erythropoiesis from hematopoietic stem cell to mature erythrocytes. Schematic representation of mouse erythroid development starting from the hematopoietic stem cell (HSC) via several committed progenitor stages to an enucleated erythrocyte. Expression of surface markers KIT, CD71, and Ter119/GPA are depicted as gradients. MEP, megakaryocyte-erythroid progenitor; BFU-E, burst forming unit erythroid; CFU-E, colony forming unit erythroid; ProEB, proerythroblast; BasoEB, basophilic erythroblasts; PolyEB, polyerithroblast; OrthoEB, orthochromatic erythroblast

stress like anemia, albeit to a limited extent, and that such self-renewal is subject to EPO[26,27]. The following stages can be morphologically determined using stainings that monitor the nuclear condensation progress as well as changes in cytoplasmic components, such as hemoglobin and RNA content. The subsequent forming cell after the CFU-E, the proerythroblast, is a large cell with a massive nucleus, visible nucleoli, surrounded by a basophilic cytoplasm[25]. Basophilic erythroblasts are slightly smaller cells than the proerythroblast, with also smaller nuclei. During the poly- and ortho-chromatic erythroblast stages, robust hemoglobinization occurs. Orthochromatic erythroblasts are the smallest nucleated erythrocyte precursors. During this stage, the nucleus goes through pycnotic degeneration, that means chromatin condensation and reduction of nuclear volume. The nucleus is then extruded and removed by macrophages[28,29]. Macrophages have been demonstrated to be very important in this process, as they are also part of the erythroblastic island, composed by central macrophage cells and erythroblasts physically attached to them[29]. Such islands can induce erythroblast proliferation, especially during stress erythropoiesis[30,31]. After nuclear extrusion, the cell becomes a reticulocyte. Reticulocytes are little bit larger than a mature erythrocytes, have an irregular shape and still contain mRNA and translation proteins needed for globin protein synthesis[25,32,33]. After completing globin synthesis, reticulocytes become mature red blood cells with an average diameter of 7 -8 µm in human and 4-5 µm in mouse. This is a complex process that necessitates loss of about 20% of membrane area, reduction of cell volume, high association of the cytoskeleton to the outer cell membrane, and depletion of all residual cytoplasmic organelles by autophagy and exocytosis[34,35]. Mature erythrocytes finally maximize their surface area/volume ratio by adopting a biconcave disk shape which grants rapid oxygen diffusion in and out of the cell[33]. RBC number is maintained by constant generation of reticulocytes released in the circulation, where they mature into RBCs. This counterbalances the elimination of senescent RBCs by macrophages, occurring primarily in the liver and spleen[25]. In humans, this process results in the release of more than 2 million reticulocytes every second in the bloodstream.

HEMOGLOBIN SWITCHING

Within the erythrocytes, binding of oxygen and carbon dioxide is mediated by hemoglobin. Hemoglobin, discovered by Hünefeld in 1840, is a metalloprotein composed of two α-like and two β-like globin chains and four iron-coordinated heme moieties[1] (Figure 4). The iron ion constitutes the binding site for oxygen. The globin protein includes 8 stretches of α–helices that are folded into a globule that dimerizes, and are afterwards assembled into a tetrameric structure. These molecules are expressed at extremely high level, for example each human erythrocyte contains circa 250x10^6 hemoglobin molecules[36].

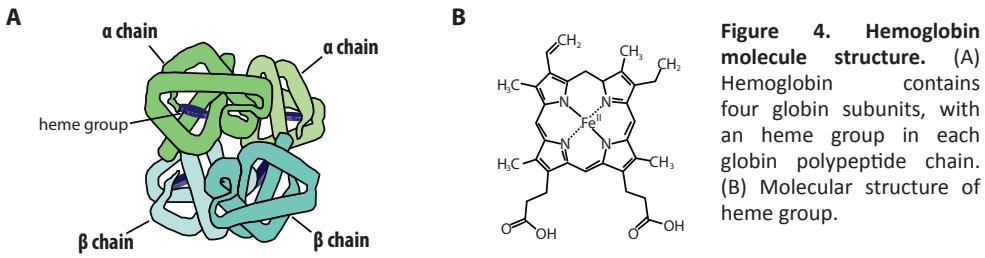


Figure 4. Hemoglobin molecule structure. (A) Hemoglobin contains four globin subunits, with an heme group in each globin polypeptide chain. (B) Molecular structure of heme group.

The human α-like and β-like globin loci, located on chromosomes 16 and 11, respectively, encode these proteins. Different α-like and β-like globin genes are active during ontogeny, and this results in developmental stage-specific hemoglobin molecules that can answer to the oxygen need of the forming organism. The human α-globin locus is composed of ζ-α1-, and α2-globin genes and an upstream major regulatory element (HS-40), while the human β-globin locus consists of five functional β-like globin genes and the locus control region (LCR), composed of five DNase I-hypersensitive sites (HSs, 1 to 5)[37,38] (Figure 5). These β-globin genes are expressed in a tissue- and developmental-specific way and are situated in the chromosome in the same order of their expression, 5'-ε-^{G}γ-^{A}γ-δ-β-3'.

During human development, two switches of β-globin gene expression occur, along with a change in the site of hematopoiesis[33] (Figure 5). The ε-globin gene is expressed during the first six weeks of gestation in primitive erythroid cells derived from the yolk sac, whereas the γ- and β-globin genes are repressed (embryonic or primitive erythropoiesis). During the first switch, hematopoiesis migrates to the fetal liver and the ^{G}γ- and ^{A}γ-globin genes are upregulated in the definitive erythrocytes (fetal definitive erythropoiesis). The ε-globin gene is silenced. During the second switch occurring around birth, the β- and the δ-globin gene are activated in the bone marrow and the spleen (adult definitive erythropoiesis). Concomitantly the γ-globin genes are silenced[33].

The expression of a specific β-like globin gene during the fetal stage evolved 35 to 55 million years ago during primate evolution[39], and thus it is considered to be a relatively recent event. Hence, except humans and old world monkeys, other species undergo only one switch, from embryonic to adult globin, that occurs during fetal development.

In the mouse for example only one switch occurs. The mouse β-like globin genes locus (ε^y, βh1, βmajor, βminor) is located on chromosome 7, while the α-like globin genes (ζ, α1, α2) on chromosome 11 (Figure 5). The ε^y and βh1 genes are expressed only during the embryonic stages. Around embryonic days E11.5 – E13.5, when mouse globin switching occurs, they are replaced by adult βmajor and βminor genes[40] (Figure 5). Interestingly, the βh1 gene initially seems to be more expressed than the ε^y gene, indicating that globin gene expression does not exactly correlate with the order of genes on the chromosome[41].

Transgenic mice that contain the human globin locus express the human genes as their

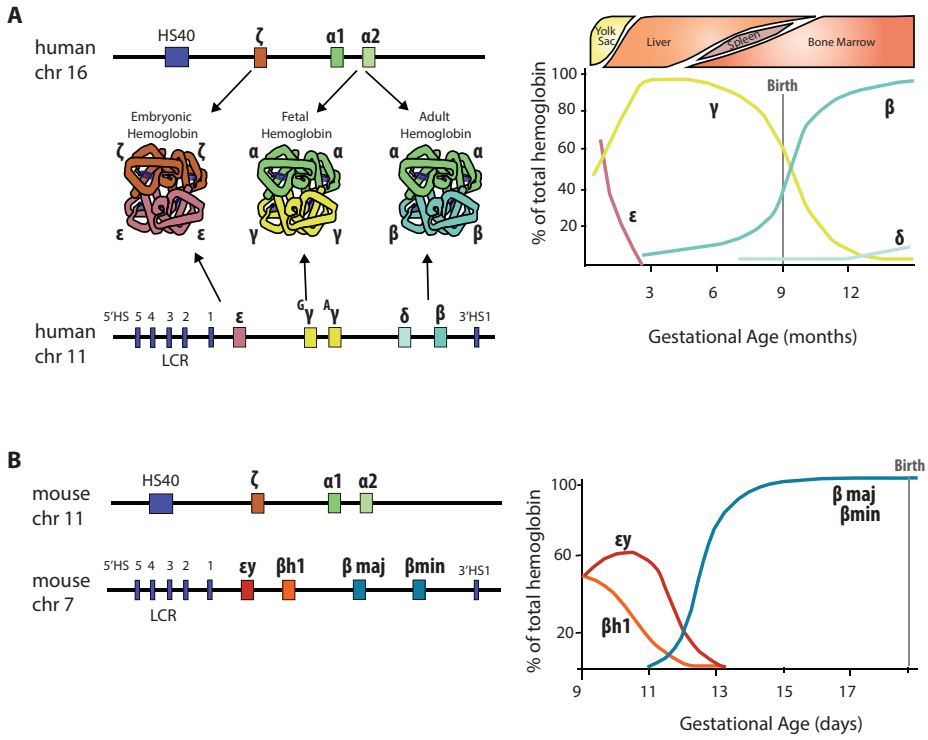


Figure 5. Hemoglobin switching. Left: schematic representation of human (A) and murine globin loci (B). The α-globin locus is composed by an upstream DNase hypersensitive site (purple boxes, HS40) and three α-like globins: ζ, a1 and a2 both in humans in mice. The β-like globin locus consists of β-like globin genes (colored boxes), upstream DNaseI hypersensitive sites 5'HS1 to 5 (purple boxes) within the locus control regions (LCR), and downstream 3'HS1. Between the human α- and β-globin loci the main hemoglobins expressed during development are illustrated: embryonic globin (ζ2ε2; HbE Gower-1); fetal hemoglobin (α2γ2; HbF) and adult hemoglobin (α2β2, HbA). Other hemoglobins not depicted in this picture are: Hemoglobin Portland (ζ2γ2), HbE Grower-2 (α2ε2), Hemoglobin A2 (α2δ2, HbA2). Right: developmental switching of the β-like globin gene expression in human (A) and mouse (B). Above the graph for the human locus the shifting sites of hematopoiesis are indicated. Figure adapted from ref[45].

murine homologs almost completely during development. Fetal γ-globin gene starts to be expressed at E10.5, while adult β-globin is activated around E12.5. γ-globin is then completely silenced around E16.5 and only β-globin is expressed throughout the entire life[42]. This result indicates that human and mouse erythroblasts go through similar developmental mechanisms. For this reason, the mouse is considered a good model to investigate human hemoglobin switching.

Properties of the LCR

The principal regulatory element of the β-globin locus is the Locus Control Region (LCR), a region essential for appropriate expression of all β-like globin genes. It was first described in the 1980's in patients suffering from anemia (γδβ-thalassemia), who displayed β-chain imbalance, although they presented a normal β-globin gene[43,44]. It was postulated that this

was due to deletion of a region important for the activation of β-globin. This region is now known as the LCR.

The LCR (Figure 5) extents 6 to 20 kb upstream of the ε-globin gene and is comprised of 5 DNaseI Hypersensitive Sites (HSs). Each HS is composed of a 250 nucleotides long core sequence, which is enriched in binding motifs for transcription factors[39]. The number of HSs in the LCR is variable among different species and in many cases there is no consensus on the exact number. However, it is generally accepted that in human and mouse, HS1–HS4 exert the strongest enhancer role, while HS5 has a structural function and has no or little effect on globin gene expression[45]. Besides HSs in the LCR, three important distal HSs exist, also with structural function: the 3' HS is positioned 20 kb downstream of the human β-globin gene, whereas two additional HSs are located upstream of the β-globin cluster at -110/-107 kb in humans and at -62/-60 kb in mice[46].

Since its discovery, many studies have been aimed at understanding how the LCR regulates globin expression. Thus far, it has become evident that, in humans, the LCR upregulates one globin gene at a time, with the different genes competing for LCR-induced activation[47]. This leads to the model that the LCR physically interacts with the induced genes. The development of two techniques, Chromosome Conformation Capture (3C) and RNA TRAP, helped demonstrating the close proximity between the LCR and the activated globin gene[48,49]. This proximity is highly dynamic and depends on the maturational status of the cell. In human progenitor cells not yet expressing β-globin, a chromatin hub (CH) is formed, a loop structure that bridges the HSs from the LCR, the upstream outer hypersensitive sites and the 3'HS[50] (Figure 6). In cells expressing β-globin, the hub co-localizes with active genes, thereby forming a transcriptionally active structure known as the active chromatin hub (ACH)[50]. Intriguingly, a similar mechanism exists in mice, but genes appear to be regulated in tandem. That means that εy and βh1 are simultaneously activated during fetal definitive hematopoiesis, and βmaj and βmin during adult definitive erythropoiesis[45] (Figure 6). Finally, it is worth mentioning that disruption of the LCR and the β-gene co-localization inhibits β-globin gene expression[45].

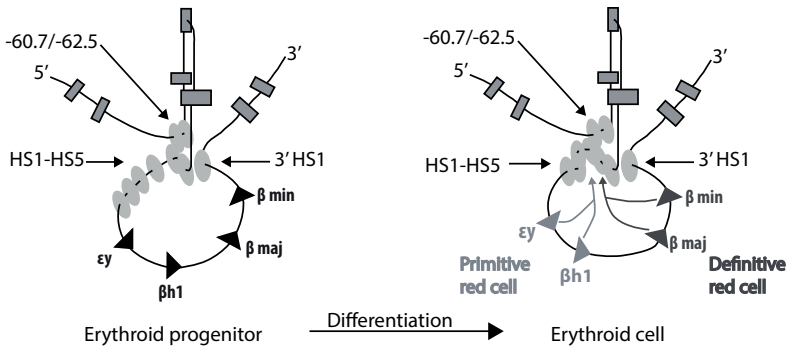


Figure 6. The chromatin hub. Structure of the mouse β-globin chromatin hub (CH, left) in erythroid progenitor cells and the active chromatin hub (ACH, right) in β-globin expressing erythroid cells. Inactive β-globin genes are depicted by black arrowheads. Embryonic genes, participating in the ACH in embryonic tissues, are depicted by light gray arrowheads. Fetal and adult genes, participating in the ACH in fetal and adult tissues, are depicted by dark gray arrowheads. Hypersensitive sites (HS1-HS5, -60.7/-62.5 kb and 3'HS) are shown as gray ovals, and olfactory receptor genes are depicted by gray boxes. Figure from ref[45].

THE ROLE OF EPO IN ERYTHROPOIESIS

Erythropoiesis is regulated by several external factors, like cytokines and hormones, that guarantee the balance between proliferation and differentiation of the cells. These factors ensure a proper number of mature cells by regulating the proper genes during different stages of erythrocyte maturation. Erythropoietin (EPO), a glycoprotein hormone, acts as the principal regulator of erythropoiesis, by increasing proliferation and survival of erythroid progenitors. It is produced by the peritubular capillary endothelial cells which are located in the kidney during adult life[51]. Hypoxia is the main regulator of EPO production[51.]

The EPO pathway plays an essential role, demonstrated by the fact that *Epo*$^{-/-}$ and *Epo receptor*$^{-/-}$ mice die *in utero* because of a deficit in definitive erythropoiesis[52]. The EPO receptor (*EpoR*) is expressed on multipotent hematopoietic progenitors and lasts until the orthochromatic erythroblast stage[53]. It is comprised of an extracellular ligand binding domain, a single transmembrane domain, and an intracellular domain. Binding of EPO induces a conformational change in the *EpoR,* resulting in phosphorylation and activation of the tyrosine kinase JAK2. In turn, the phosphorylated tyrosine residues in JAK2 and the EpoR serve as binding sites for SH2 (Src- homology 2) domain-containing signaling proteins, like the members of STAT (signal transducer and activator of transcription) family[53]. Of these, STAT5a and STAT5b are the proteins activated by the EpoR patwhay[53]. When phosphorylated, STAT5a and b dimerize, and translocate into the nucleus, where they act as activators of target genes. The anti-apoptotic BCL-XL protein is one of the several genes activated by STAT5[54], and its conditional knock-out in erythroid cells produces anemia due to breakdown of RBCs in mice[55]. Another important pathway activated by EpoR is the PI3-kinase pathway[53]. Activation of PI3-kinase stimulates the kinase PKB/AKT, which subsequently activates anti-apoptotic and proliferative signals.

Since the continuous activation of survival and proliferative signals can result in cell transformation, the modulation of the EPO pathway is extremely important. Signaling from the EPOR (and the attached JAK2) is mainly inactivated by dephosporylation by phosphatase SHP-1[56]. Secondary mechanisms involved in *EpoR* down-regulation comprise the Suppressor of Cytokine Signalling (SOCS) family proteins SOCS1, CIS and SOCS3. The genes of these proteins are activated by STAT5 following EPO stimulation, thus installing a negative feedback loop[57,58]. The SOCS proteins directly inhibit JAK2 by binding to its kinase domain, and initiate ubiquitin-mediated proteasomal degradation[59].

Other cytokines/growth factors have also been implicated in the control of erythropoiesis. For example, interleukin-3 (IL-3), Stem Cell Factor (SCF) and Insulin-Like Growth Factor-I (IGF-I) stimulate the growth and survival of erythroid progenitors[60]. Insulin receptors are expressed at later stages of erythroid development and their activation has anti-apoptotic effects on red cells that are differentiating[61]. Lastly, Transforming Growth Factor (TGF) α and β are also known to have a significant function in regulating erythrocyte numbers[62].

TRANSCRIPTIONAL REGULATION OF ERYTHROPOIESIS AND HEMOGLOBIN SWITCHING

Many proteins involved in transcriptional regulation have proven roles in erythroid development. Some of these proteins are erythroid-specific but surprisingly others are ubiquitously expressed. Three of the main transcription factors that regulate erythropoiesis and/or hemoglobin switching are described below.

GATA1

Proteins belonging to the GATA protein family (GATA1–6) are zinc finger transcription factors that bind the consensus (T/A)GATA(A/G) motif and switch on and off genes that contain this sequence in their regulatory region(s)[63,64]. The most extensively studied modulator of globin gene expression is GATA1, which has a major role both in primitive and definitive erythroid development. GATA1 binds the DNA via two zinc finger domains, both displaying a Cys-X2-Cys-X17-Cys-X2-Cys configuration[65]. The zinc finger located at the C-terminus binds with high-affinity to DNA, while the one located at the N-terminus enhances this interaction[66].

GATA1 was initially identified as a general "switch" factor for erythroid development in 1989 by Evans and Felsenfeld[65] due to its capability to associate with essential regulatory sequences residing in globin genes. Since then, GATA-binding motifs ((T/A)GATA(A/G)) have been identified in the promoters and/or enhancers of most genes with a role in erythroid and megakaryocyte development[67]. GATA1 activates erythroid development, repressing genes associated with the immature and proliferative state, including GATA2 and the KIT cytokine receptor[68-70].

GATA1 activity is crucial for erythroid cell maturation, as *Gata1* knock-out cells are unable to mature past the proerythroblast stage due to cell-autonomous induced apoptosis[71,72]. *Gata1* hemizygous male knock-out mice (the *Gata1* gene is located on the X-chromosome) die around E10.5 from severe anemia caused by defects in erythroid maturation at the proerythroblast stage[72]. Moreover, *Gata1* knock-down embryos, retaining about 5% of endogenous protein, halt primitive erythropoiesis and die around E11.5-E12.5[73]. In contrast, GATA1 overexpression induces proliferation of erythroid progenitors and inhibits differentiation into mature erythroid cells[24].

Proteins that can interact with GATA1 include important regulators of erythroid development, such as LDB1, FOG1, TAL1 and LMO2. Targeted disruption of each of these genes leads to pronounced defects in primitive and definitive erythroid cells[74-78].

BCL11A (B-cell lymphoma/leukemia 11A)

This gene, located on chromosome 2, encodes a C2H2 type zinc-finger transcription factor, that has been previously shown to be critical in lymphopoiesis for the development of B cell precursors[79]. Five different isoforms of BCL11A exist, arising from five differentially spliced transcript variants.

In 1999, through genome-wide association studies (GWAS) this locus was described to associate with HbF[80,81] (fetal hemoglobin). Subsequently, BCL11A was shown to function as a regulator of HbF expression[82]. Various single nucleotide polymorphisms (SNPs) in the BCL11A gene have been identified, which together account for approximately 15-18% of the variation in fetal hemoglobin expression levels[80,83-87]. Adult erythroid cells specifically express full-length forms of BCL11A and down-regulation of BCL11A in adult human erythroid precursors leads to up-regulation of γ-globin[82]. Knockout of BCL11A in transgenic mice carrying the human β-globin locus inhibits physiological suppression of both endogenous mouse β-like embryonic genes and exogenous human γ-globin gene expression in erythroid cells of the fetal liver[88,89]. BCL11A, together with its partners, binds the locus control region (LCR), the ε-globin, and the intergenic regions between γ-globin and δ-globin genes, where it co-localizes with GATA1 and FOG1[90,91]. Moreover, BCL11A co-operates with SOX6 (Sry-related HMG box) in silencing γ-globin transcription in adult human erythroid progenitors[90]. One mechanism by which BCL11A can negatively regulate gene expression is by attracting repressive chromatin remodelers, such as the MI2/NuRD, LSD1/CoREST and the SWI/SNF complex, to the regulatory elements where it binds[92].

These results show that BCL11A is an important developmental regulator of the fetal-to

adult hemoglobin switch in humans both by serving as a repressor of the γ-globin gene and acting directly within the β-globin locus, along with a set of partner proteins. However, the factors that allow for the developmental regulation of BCL11A expression and that may also affect other aspects of globin gene regulation are still unknown.

KLF1 / EKLF (Erythroid Krüppel-like factor)

A well-characterized erythroid-specific transcription factor is the Krüppel-like zinc finger DNA-binding protein KLF1 (previously known as EKLF). The KLF1 DNA binding domain, composed of three C2H2 zinc fingers, is located at the C-terminus of the protein and recognizes the consensus sequence 5'-NCNCNCCCN-3'[93]. This motif is often found within the regulatory regions of many erythroid genes[93]. KLF1 is mainly a transcriptional activator, but it has been reported to also behave as a transcriptional repressor[94-96]. KLF1 protein is expressed both during primitive and definitive hematopoiesis specifically in the hematopoietic organs[97,98]. KLF1 can hardly be detected in either hematopoietic stem cells or multipotent myeloid progenitors. On the contrary, KLF1 expression becomes evident in the common myeloid progenitor (CMP) and its daughter cells, and it further increases during erythropoiesis, while it is down-regulated in cells that differentiate towards megakaryocytes[99]. It hence has a dual role in hematopoiesis, promoting erythroid lineage development and simultaneously limiting megakaryocytic differentiation.

KLF1 is not required for the expression of the embryonic ε- and fetal γ-genes, but strongly affects the expression of the adult β-globin gene. KLF1 knockout mice die in utero from severe anemia at the fetal liver stage (E14-E15) because of defects in definitive erythropoiesis, partially due to defective β-globin expression[100,101]. However, the defects occurring because of KLF1 depletion appeared not to be solely generated by β-globin misregulation, as reintroduction of β-like globin alone does not rescue the phenotype. This shows that KFL1 regulates additional genes necessary for normal erythropoiesis[102]. In homozygous KLF1 knockout mice containing a transgenic human β-globin locus, β-globin expression is absent and γ-globin expression is increased. Furthermore, in these mice the interaction between the LCR and the β-major gene is abolished[103]. To demonstrate that KLF1 can directly mediate looping, 3C experiments were performed on KLF1 knockout mice rescued with an inducible version of KLF1[103], that showed reformation of the active chromatin hub and partially rescued β-major gene expression, hinting that KLF1 directly affects chromatin looping between the LCR and the β-major promoter[103]. These results pinpoint the importance of KLF1 in mediating the switching process by regulating chromatin loop formation. KLF1 allows efficient β-globin gene competition with the γ-globin genes, promoting loop creation between the β-globin gene promoter and HS2 and HS3 of the LCR. In heterozygous KLF1 mice, β-gene expression is delayed during differentiation, accompanied by increased γ-gene expression, but β-globin levels are unaffected in adult mice[104]. Halving KLF1 expression diminishes the ability of the β-globin promoter to compete with the γ-globin promoter for contact with the LCR, resulting in the upregulation of the γ-globin genes. At later stages of development, silencing of the γ-globin genes lifts the competition, granting efficient loop formation between the β-globin gene and the LCR[103]. This occurs despite the decreased KLF1 levels and culminates in normal expression of the β-globin gene.

Gene-expression profile of sorted CD71 positive populations from WT and KLF1 knock-out fetal liver cells, reveals approximately 3000 genes with altered expression, with genes involved in cell-cycle and DNA replication over-represented[105]. In addition, deregulated genes primarily belong to families that regulate virtually every aspect of red blood cell maturation.

Two studies published in 2010[106,107] further elucidated the molecular mechanisms governing

KLF1-regulation of globin switching. They demonstrated that upon KLF1 depletion the level of BCL11A protein was reduced, compared to its expression in wild-type controls. Given the fact that KLF1 also binds to the BCL11A promoter, it constitutes a direct regulator of BCL11A expression.
Mutations in KLF1 are associated with different phenotypes, such as inhibition of Lutheran blood group expression (In(Lu))[108], congenital dyserythropoietic anemia (CDA)[109], red cell zinc protoporphyria[110], and neonatal anemia (*Nan*, discussed in more detail in Chapter 3 of this thesis)[111-113]. KLF1 mutations also correlate with hereditary persistence of fetal hemoglobin(HPFH)[107], as demonstrated in 2010 by Borg and colleagues. A nonsense mutation in KLF1 was shown to be associated with hereditary persistence of fetal hemoglobin in a large Maltese family, indicating that haploinsufficiency of KLF1 can induce fetal hemoglobin expression while maintaining normal erythropoiesis[107]. However, carriers of this mutation had variable levels of fetal hemoglobin, ranging from 3 to 20%, suggesting that the role of KLF1 in these changes may also arise due to other functions.

In the last twenty years, many proteins were demonstrated to regulate erythropoiesis and globin switching, some of them having the ability to bind different cis-regulatory region within the LCR locus and to every single globin gene promoter. The most important ones are summarized in Figure 7 and Table 1.

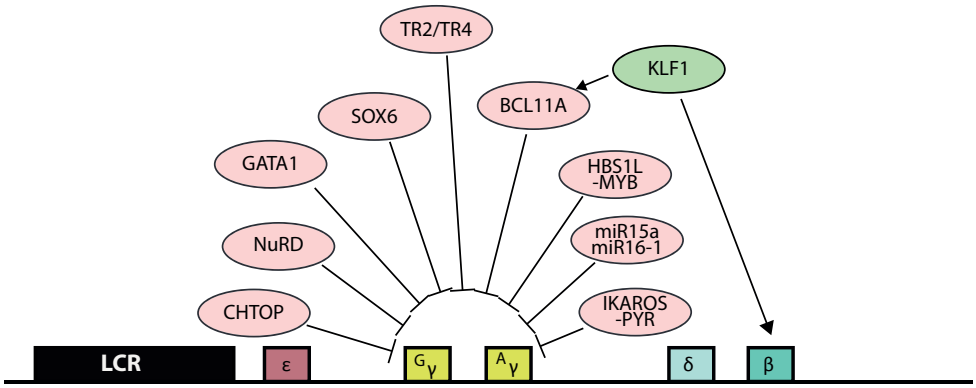


Figure 7. Regulators of globin switching. Several proteins regulate the switch between γ- and β-globin, mainly silencing the expression of the γ-chain. The known key players are depicted in this figure. KLF1 represses γ-globin via the activation of BCL11A. Notably, also mir-15a/mir16-1 have been shown to participate in this process.

Table 1. Transcription factors and other regulatory proteins in erythropoiesis. This table contains genes involved in RBCs development, with emphasis placed on genes expressed in erythroid cells. References are primarily for studies on mouse mutant phenotype. Table from ref[147].

Gene symbol(s)	Name(s)/protein	Mutant phenotype	Comments	Refs.
Gata1	GATA1	Maturation block in EryP and EryD	Erythroid progenitors form but fail to mature	[71]
Gata2	GATA2	Embryonic lethality in knockout at E10.5–E11.5 due to failure in progenitor expansion	Double KO revealed functional overlap between Gata1 and Gata2 in EryP	[145,146]
Fog1	Friend of GATA1; FOG1	Fog1 knockout mice die between E10.5 and E12.5 of severe anemia and defective megakaryopoiesis	Acts sequentially with GATA1 to specify definitive erythroid and megakaryocytic progenitors	[77,147]
Klf1/Eklf	Erythroid Kruppel-like zinc finger protein	Embryonic lethality by ~E13.5 due to defective fetal liver erythropoiesis	Defects in EryP and EryD maturation	[148-152]
Scl/Tal	Basic helix-loop-helix SCL/tal-1	Embryonic lethality at E9 –10.5 due to hematopoietic and vascular defects	Master regulator of hematopoiesis, forms complexes with Gata1 and Lmo2	[76, 153-156]
Lmo2	Rbtn2; LIM domain only 2, cysteine-rich LIM domain protein	Embryonic lethality at E10.5 due to failure in primitive erythropoiesis		[74]
Runx1	Runt-related transcription factor 1; Core Binding Factor-a2 (Cbfa2), AML1	Embryonic lethality at E11.5–12.5 due to failure in definitive erythropoiesis and hemorrhaging	Some evidence for a (nonessential) function in primitive erythropoiesis	[157-159]
Ldb1	LIM domain-binding protein 1	Embryonic lethality after E9.5 (defective yolk sac development)	Mutant displays various other patterning defects; positive regulator oflate steps in erythroid maturation. ES cell studies: decreased hemangioblast numbers and block in differentiation	[160-162]
c-Myb	c-MYB; myeloblastosis oncogene	Absence of EryD progenitors	Required for definitive but not primitive erythropoiesis; reduced megakaryocyte progenitor numbers	[163-164]
c-Myc	c-MYC; myelocytomatosis oncogene	Epiblast-restricted c-Myc-null embryos severely anemic and die before E12, largely due to defects in primitive erythropoiesis	Physiologic levels critical; downregulation required for terminal maturation offetal liver EryD	[165]
Klf2	Kruppel-like zinc finger protein 2	Critical for primitive erythropoiesis	Redundant roles with Klf1/Eklf	[166]
Klf6	Kruppel-like zinc finger protein 6	Knockouts embryonic lethal by E12	May function primarily in mesoderm specification rather than in primitive erythropoiesis	[167]
BRG1	Brahma related gene 1, SMARCA4	Embryonic lethal E10.5-11; severe EryP defect by E9.5	Tie2-Cre conditional KO	[168]
c-Maf	Proto-oncogene	Knockouts die from severe anemia E15-18	Important function in fetal liver macrophages; decreased EryD in EBIs offetal liver, effect is non cell-autonomous	[169]
ZFAT	Zinc finger and AT hook domain	Embryonic lethal E9.5, required in EryP	Regulates Tal, Lmo2, Gata1	[170]
P400/Domino	E1A binding protein p400	Embryonic lethality E11.5 due to defects in primitive erythropoiesis	SWI2/SNF2 family ATPase; N-terminal domain deletion	[171]
Zbp89	ZBP-89; Zinc finger protein 148 (ZFP-148)	Embryonic lethal between E8.5 and E10.5 due to placental insufficiency	Part of complex with GATA1 and FOG1. Required for definitive but not primitive erythropoiesis.	[172]
Gfi1b	Growth factor independent 1b; Gfi1b	Embryonic lethal at E14.5 as a result of deficiencies in erythroid and megakaryocyte development	Target of Runx1 and regulator of endothelial to hematopoietic transition	[173,174]
Sox6	SRY-box containing gene 6	Necessary for efficient physiological and stress erythropoiesis in mouse EryD; silencer of embryonic and fetal globin genes in EryD	Sox6, Gata1 and Bcl11A co-occupy human β-globin cluster; Sox6 expressed in mouse EryP progenitors but then downregulated	[90, 175, 176]
Bcl11A	B cell CLL/lymphoma 11A; Bcl11a	Knockout mice die perinatally from unknown causes. Disruption of embryonic/fetal to adult globin gene switch	Embryonic globin gene expression not silenced in knockout mice; regulated by Eklf/Klf1. Essential for normal lymphoid development	[90, 79, 88,106,177]
NuRD	Nucleosome remodeling histone deacetylase	Cofactor for Fog1; mutations in Fog1 that disrupt interactions with NuRD results in anemia and macrothrombocytopenia	Context-dependent repressor and activator functions. Complex phenotype with similarities to Gata1 or Fog1 mutations	[178]
PIT1	PIT1/SLC20A1	Embryonic lethality at E12.5, erythroid maturation defects	Defects in vitro and in vivo; target of Eklf/Klf1	[179-180]
Nr2c1/Nr2c2	NR2C1/NR2C2; TR2/TR4	Early embryonic letality in the double mutants around E7.5-E9.5. Conditional knock out of TR2/TR4 blocks the differentiation and the maturation of erythroid cell.	Increased expression of εγ and βh1 globin genes.	[181-183]

HEMOGLOBINOPATHIES

The importance of unraveling the fetal to adult globin-switch is accentuated by the clinical relevance of the β-hemoglobin disorders[39]. Hemoglobinopathies are inherited genetic disorders that culminate in the formation of distorted globin chains. The major outcome of such diseases is an imbalance between α- and β-globin polypeptides, leading to precipitation of excess unbound globins in cells and thus damage to the erythrocytes. The most common types of β-hemoglobin disorders are β-thalassemia and sickle cell disease (SCD). The homozygous condition, if untreated, leads to the death of the patients. It is probable that the high frequency of hemoglobinopathies is due to heterozygote protection against malaria, although explicit confirmation is still missing. Expectedly, β-thalassemia and SCD are common in the Mediterranean area, Africa, India, Middle east, and Southeast Asia, where malaria used to be/is endemic[114,115]. Due to population migration these two diseases are now found in most countries worldwide.

Sickle cell disease

Sickle cell disease is due to an A- to-T point mutation in the 6th codon of the β-globin gene, leading to the substitution of a glutamate (charged residue) by a valine (hydrophobic residue) in the β-globin protein, and thus producing the so-called β^s variant. The resulting hemoglobin molecule ($\alpha 2\beta^s 2$) is known as HbS or sickle hemoglobin [116,117]. In low oxygen conditions, HbS polymerizes, and this results in deformation to a sickle shape and rigidity of RBCs [118]. The misshapen RBCs can cause blockage of small blood vessels and hence local necrosis of tissues. SCD patients are indeed subject to many complications, among which moderate to severe anemia, strokes, vaso-occlusive episodes, acute chest syndrome, life-threatening infection, renal disease, retinopathy and shortened lifespan[119].

β-thalassemia

Worldwide, β-thalassemia is the second most common monogenic disorder after SCD[115]. β-thalassemia occurs because of mutations in the β-globin locus that cause decreased (β^+) or absent (β^0) production of β-globin and thus precipitation of unassembled α-chains. This causes oxidative damage to the cell membrane, subsequent recognition by macrophages and hence removal from the circulation. There are three major phenotypes of β-thalassemia, summarized in Table 2. Over 200 different mutations of the β-globin gene locus have been found in patients with β-thalassemia[114].

Table 2. Clinical classification of β-thalassemia.

Name	Alleles	Phenotype
Thalassemia minor	β^+/β β^0/β	Asymptomatic. Often no treatment needed, but risk of iron overload. Mild hypochromic anemia. Not life-threatening.
Thalassemia intermedia	β^+/β^+ β^0/β^+	Diagnosed usually in late childhood. Globin chain production moderately impaired. Patients may require blood transfusion. Transfusion dependendent patients develop iron overload and require chelation therapy. Affected individuals can manage a normal life.
Thalassemia major	β^0/β^0	Severe anemia presenting early in life. Patients require regular lifelong blood transfusions that lead to iron overload. Iron chelation therapy is essential to prevent internal organ damage. Shorter life-expectancy.

Hereditary persistence of fetal hemoglobin

Hereditary persistence of fetal hemoglobin (HPFH) is a group of genetic conditions characterized by elevated expression of the γ-globin gene in adult erythroid cells. Interestingly, patients with compound heterozygosity for sickle cell disease/β-thalassemia and HPFH mutations are largely asymptomatic[114,120]. HPFH conditions can be subdivided into deletion and non-deletion HPFH. In deletion HPFH, the rearrangement of regulatory regions, such as the deletion of suppressor sequences or the close proximity of non-canonical activation sequences to the γ-globin gene promoter, are the result of the deletion of different large regions downstream of the γ-globin genes, eventually resulting in a higher expression of γ-globin[121,122]. In non-deletion HPFH, the cause of elevated levels of fetal hemoglobin is point mutations in the γ-globin promoter[123-128]. HPFH can also be the result of mutations outside the β-globin locus, such as mutations in KLF1, BCL11A or MYB[83,84,129-132] that lead to an inefficient repression of the γ-globin. For example, as mentioned earlier in this chapter, KLF1 haploinsufficiency can lead to increased γ-globin expression[107,110].

TREATMENT FOR HEMOGLOBINOPATHIES

Different therapeutic methods are available to treat β-thalassemic and SCD patients.

- Regular blood transfusion and iron chelation therapy to remove excess of iron from the body[133]. This approach produces iron overload and is associated with many risks, due to transfusions, such as viral infection.
- Bone marrow transplantation is the only curative method, but relies on the availability of an HLA-matched donor for bone marrow cells. It has been successfully used both for SCD and thalassemia patients[134,135], but despite promising results, stem cell transplantation is only limited to a small portion of patients, due to age, care and cost considerations.
- γ-globin reactivation is achieved with chemical compounds, like Hydroxycarbamide (hydroxyurea or HU), 5-azacytidine (5-aza) and short chain fatty acid derivates. HU is a ribonucleotide reductase inhibitor that inhibits DNA replication, but the mechanisms responsible for γ-globin reactivation are still unknown. It can efficiently ameliorate disease symptoms in approximately 50% of SCD cases[136]. 5-aza is a DNA methyltransferase inhibitor known to increase levels of HbF[137,138], but also in this case the relevant mechanisms are still unknown. Sodium butyrate, a short-chain fatty acid derivative, is an HDAC inhibitor which induces HbF by increasing γ-globin mRNA transcription[139-141]. Unfortunately, none of these γ-globin reactivating drugs is efficient, specific, and safe enough to make them suitable to most hemoglobinopathy-patients worldwide.
- Gene therapy aims to express a normal β-like globin gene in erythroid cells following permanent gene transfer into hematopoietic stem cells. It has shown promising results in mice[142,143] and more recently also in humans[144], where a patient with β-thalassemia became transfusion-independent after gene therapy with autologous CD34+ cells transduced with a lentiviral vector containing an engineered β-globin gene. Since then, more patients have been recruited for clinical trials, with encouraging outcomes, but although gene therapy offers hope for future application, especially for patients with no availability of an HLA-matched donor, this approach is extremely sophisticated and limited to a number of highly specialized centers. Additionally it requires extensive financial resources that are not available to the vast majority of SCD and thalassemia patients.

SCOPE OF THIS THESIS

β-thalassemia and sickle cell disease, along with α-thalassemia, are the most common monogenic disorders, with approximately 300.000 affected children born each year. The observation that increased γ-globin expression can ameliorate the symptoms of these diseases eventually led to the discovery of the first treatments, mainly hydroxyurea-mediated induction of γ-globin expression. However, hudroxyurea is effective in only 50% of the SCD patients. Therefore, current research focuses on finding new strategies to increase fetal γ-globin induction that could be utilized to complement defective adult globin in the aforementioned diseases. Understanding the exact mechanisms driving the globin switch will help in the development of new therapeutics. However, despite increasing knowledge of how the globin genes are modulated, a comprehensive picture of these mechanism is still missing. KLF1 undoubtedly constitutes one of the major regulators of erythropoiesis, but although it has been investigated for more than 20 year, its precise role in fine-tuning erythroid maturation is still elusive. In the experimental Chapters 2 to 5, I present our work, aiming to provide in-depth insight on how KLF1 modulates diverse processes to control erythropoiesis. In particular, in **Chapter 2** I show our results on the interplay between KLF1 and BCL11A in adult erythropoiesis and globin regulation. Taking advantage of the KLF1 Nan mutant, we were able to identify new potential KLF1 target genes with a possible role in erythropoiesis. More specifically, in **Chapter 3** I describe the effects of the KFL1 Nan mutation during embryonic development, focusing in particular on the role of the exportin XPO7 in causing the overall phenotype, while in **Chapter 4** I describe our initial observation of a potential, unexpected role of KLF1 in regulating the function of the Golgi apparatus via BSDC1. Finally, identifying KLF1 binding partners would be a great step towards elucidating its multifaceted role in globin switching and erythropoiesis. Our ongoing efforts to obtain a comprehensive list of such KLF1-interacting partners by mass-spectrometry analysis of pulled-down KLF1 complexes are described in **Chapter 5**.

REFERENCES

1. Alberts, B. Molecular biology of the cell. 5th edition, Garland Science.2008.
2. Orkin, S.H. & Zon, L.I. Hematopoiesis: an evolving paradigm for stem cell biology. Cell 132, 631-44 (2008).
3. Laki, K. Our ancient heritage in blood clotting and some of its consequences. Ann N Y Acad Sci 8, 297-307 (1972).
4. Janeway, C.A., Travers, P., Walport, M., Shlomchik, M.J. Immunobiology. 5th edition, Garland Science.2001.
5. Palis, J. Ontogeny of erythropoiesis. Curr Opin Hematol 15, 155-61 (2008).
6. Kingsley, P.D., Malik, J., Fantauzzo, K.A. & Palis, J. Yolk sac-derived primitive erythroblasts enucleate during mammalian embryogenesis. Blood 104, 19-25 (2004).
7. Fraser, S.T., Isern, J. & Baron, M.H. Maturation and enucleation of primitive erythroblasts during mouse embryogenesis is accompanied by changes in cell-surface antigen expression. Blood 109, 343-52 (2007).
8. Dzierzak, E. & Medvinsky, A. Mouse embryonic hematopoiesis. Trends Genet 11, 359-66 (1995).
9. Dzierzak, E. & Speck, N.A. Of lineage and legacy: the development of mammalian hematopoietic stem cells. Nat Immunol 9, 129-36 (2008).
10. Weissman, I.L. Stem cells: units of development, units of regeneration, and units in evolution. Cell 100, 157-68 (2000).
11. Till, J.E. & Mc, C.E. A direct measurement of the radiation sensitivity of normal mouse bone marrow cells. Radiat Res 14, 213-22 (1961).
12. Abkowitz, J.L., Golinelli, D., Harrison, D.E. & Guttorp, P. In vivo kinetics of murine hemopoietic stem cells. Blood 96, 3399-405 (2000).
13. Kiel, M.J. et al. SLAM family receptors distinguish hematopoietic stem and progenitor cells and reveal endothelial niches for stem cells. Cell 121, 1109-21 (2005).
14. Notta, F. et al. Isolation of single human hematopoietic stem cells capable of long-term multilineage engraftment. Science 333, 218-21 (2011).
15. Cheshier, S.H., Morrison, S.J., Liao, X. & Weissman, I.L. In vivo proliferation and cell cycle kinetics of long-term self-renewing hematopoietic stem cells. Proc Natl Acad Sci U S A 96, 3120-5 (1999).
16. Wilson, A. & Trumpp, A. Bone-marrow haematopoietic-stem-cell niches. Nat Rev Immunol 6, 93-106 (2006).
17. Katsura, Y. Redefinition of lymphoid progenitors. Nat Rev Immunol 2, 127-32 (2002).
18. Adolfsson, J. et al. Identification of Flt3+ lympho-myeloid stem cells lacking erythro-megakaryocytic potential a revised road map for adult blood lineage commitment. Cell 121, 295-306 (2005).
19. Morrison, S.J., Wandycz, A.M., Hemmati, H.D., Wright, D.E. & Weissman, I.L. Identification of a lineage of multipotent hematopoietic progenitors. Development 124, 1929-39 (1997).
20. Akashi, K., Traver, D., Miyamoto, T. & Weissman, I.L. A clonogenic common myeloid progenitor that gives rise to all myeloid lineages. Nature 404, 193-7 (2000).
21. Kondo, M., Weissman, I.L. & Akashi, K. Identification of clonogenic common lymphoid progenitors in mouse bone marrow. Cell 91, 661-72 (1997).
22. Galloway, J.L. & Zon, L.I. Ontogeny of hematopoiesis: examining the emergence of hematopoietic cells in the vertebrate embryo. Curr Top Dev Biol 53, 139-58 (2003).
23. Wong, P.M., Chung, S.W., Chui, D.H. & Eaves, C.J. Properties of the earliest clonogenic hemopoietic precursors to appear in the developing murine yolk sac. Proc Natl Acad Sci U S A 83, 3851-4 (1986).
24. Whyatt, D. et al. An intrinsic but cell-nonautonomous defect in GATA-1-overexpressing mouse erythroid cells. Nature 406, 519-24 (2000).
25. Palis, J. Primitive and definitive erythropoiesis in mammals. Front Physiol 5, 3 (2014).
26. Peslak, S.A. et al. EPO-mediated expansion of late-stage erythroid progenitors in the bone marrow initiates recovery from sublethal radiation stress. Blood 120, 2501-11 (2012).
27. Flygare, J., Rayon Estrada, V., Shin, C., Gupta, S. & Lodish, H.F. HIF1alpha synergizes with glucocorticoids to promote BFU-E progenitor self-renewal. Blood 117, 3435-44 (2011).

28. Keerthivasan, G., Wickrema, A. & Crispino, J.D. Erythroblast enucleation. Stem Cells Int 2011, 139851 (2011).
29. Chasis, J.A. & Mohandas, N. Erythroblastic islands: niches for erythropoiesis. Blood 112, 470-8 (2008).
30. Chow, A. et al. CD169(+) macrophages provide a niche promoting erythropoiesis under homeostasis and stress. Nat Med 19, 429-36 (2013).
31. Ramos, P. et al. Macrophages support pathological erythropoiesis in polycythemia vera and beta-thalassemia. Nat Med 19, 437-45 (2013).
32. Allen, T.D. & Dexter, T.M. Ultrastructural aspects of erythropoietic differentiation in long-term bone marrow culture. Differentiation 21, 86-94 (1982).
33. Dzierzak, E. & Philipsen, S. Erythropoiesis: development and differentiation. Cold Spring Harb Perspect Med 3, a011601 (2013).
34. Waugh, R.E., Mantalaris, A., Bauserman, R.G., Hwang, W.C. & Wu, J.H. Membrane instability in late-stage erythropoiesis. Blood 97, 1869-75 (2001).
35. Griffiths, R.E. et al. The ins and outs of human reticulocyte maturation: autophagy and the endosome/exosome pathway. Autophagy 8, 1150-1 (2012).
36. Schechter, A.N. Hemoglobin research and the origins of molecular medicine. Blood 112, 3927-38 (2008).
37. Grosveld, F., van Assendelft, G.B., Greaves, D.R. & Kollias, G. Position-independent, high-level expression of the human beta-globin gene in transgenic mice. Cell 51, 975-85 (1987).
38. Forrester, W.C., Takegawa, S., Papayannopoulou, T., Stamatoyannopoulos, G. & Groudine, M. Evidence for a locus activation region: the formation of developmentally stable hypersensitive sites in globin-expressing hybrids. Nucleic Acids Res 15, 10159-77 (1987).
39. Stamatoyannopoulos, G. Control of globin gene expression during development and erythroid differentiation. Exp Hematol 33, 259-71 (2005).
40. Fantoni, A., De la Chapelle, A., Chui, D., Rifkind, R.A. & Marks, P.A. Control mechanisms of the conversion from synthesis of embryonic to adult hemoglobin. Ann N Y Acad Sci 165, 194-204 (1969).
41. Kingsley, P.D. et al. "Maturational" globin switching in primary primitive erythroid cells. Blood 107, 1665-72 (2006).
42. Strouboulis, J., Dillon, N. & Grosveld, F. Developmental regulation of a complete 70-kb human beta-globin locus in transgenic mice. Genes Dev 6, 1857-64 (1992).
43. Van der Ploeg, L.H. et al. gamma-beta-Thalassaemia studies showing that deletion of the gamma- and delta-genes influences beta-globin gene expression in man. Nature 283, 637-42 (1980).
44. Kioussis, D., Vanin, E., deLange, T., Flavell, R.A. & Grosveld, F.G. Beta-globin gene inactivation by DNA translocation in gamma beta-thalassaemia. Nature 306, 662-6 (1983).
45. Noordermeer, D. & de Laat, W. Joining the loops: beta-globin gene regulation. IUBMB Life 60, 824-33 (2008).
46. Bulger, M. et al. Comparative structural and functional analysis of the olfactory receptor genes flanking the human and mouse beta-globin gene clusters. Proc Natl Acad Sci U S A 97, 14560-5 (2000).
47. Wijgerde, M., Grosveld, F. & Fraser, P. Transcription complex stability and chromatin dynamics in vivo. Nature 377, 209-13 (1995).
48. Carter, D., Chakalova, L., Osborne, C.S., Dai, Y.F. & Fraser, P. Long-range chromatin regulatory interactions in vivo. Nat Genet 32, 623-6 (2002).
49. Tolhuis, B., Palstra, R.J., Splinter, E., Grosveld, F. & de Laat, W. Looping and interaction between hypersensitive sites in the active beta-globin locus. Mol Cell 10, 1453-65 (2002).
50. Palstra, R.J. et al. The beta-globin nuclear compartment in development and erythroid differentiation. Nat Genet 35, 190-4 (2003).
51. Koury, M.J. & Bondurant, M.C. The molecular mechanism of erythropoietin action. Eur J Biochem 210, 649-63 (1992).
52. Wu, H., Liu, X., Jaenisch, R. & Lodish, H.F. Generation of committed erythroid BFU-E and CFU-E progenitors does not require erythropoietin or the erythropoietin receptor. Cell 83, 59-67 (1995).

53. Richmond, T.D., Chohan, M. & Barber, D.L. Turning cells red: signal transduction mediated by erythropoietin. Trends Cell Biol 15, 146-55 (2005).
54. Gregoli, P.A. & Bondurant, M.C. The roles of Bcl-X(L) and apopain in the control of erythropoiesis by erythropoietin. Blood 90, 630-40 (1997).
55. Wagner, K.U. et al. Conditional deletion of the Bcl-x gene from erythroid cells results in hemolytic anemia and profound splenomegaly. Development 127, 4949-58 (2000).
56. Klingmuller, U., Lorenz, U., Cantley, L.C., Neel, B.G. & Lodish, H.F. Specific recruitment of SH-PTP1 to the erythropoietin receptor causes inactivation of JAK2 and termination of proliferative signals. Cell 80, 729-38 (1995).
57. Starr, R. et al. A family of cytokine-inducible inhibitors of signalling. Nature 387, 917-21 (1997).
58. Yoshimura, A. et al. A novel cytokine-inducible gene CIS encodes an SH2-containing protein that binds to tyrosine-phosphorylated interleukin 3 and erythropoietin receptors. EMBO J 14, 2816-26 (1995).
59. Zhang, J.G. et al. The conserved SOCS box motif in suppressors of cytokine signaling binds to elongins B and C and may couple bound proteins to proteasomal degradation. Proc Natl Acad Sci U S A 96, 2071-6 (1999).
60. Migliaccio, G. et al. In vitro mass production of human erythroid cells from the blood of normal donors and of thalassemic patients. Blood Cells Mol Dis 28, 169-80 (2002).
61. Miyagawa, S., Kobayashi, M., Konishi, N., Sato, T. & Ueda, K. Insulin and insulin-like growth factor I support the proliferation of erythroid progenitor cells in bone marrow through the sharing of receptors. Br J Haematol 109, 555-62 (2000).
62. Gandrillon, O., Schmidt, U., Beug, H. & Samarut, J. TGF-beta cooperates with TGF-alpha to induce the self-renewal of normal erythrocytic progenitors: evidence for an autocrine mechanism. EMBO J 18, 2764-81 (1999).
63. Ko, L.J. & Engel, J.D. DNA-binding specificities of the GATA transcription factor family. Mol Cell Biol 13, 4011-22 (1993).
64. Merika, M. & Orkin, S.H. DNA-binding specificity of GATA family transcription factors. Mol Cell Biol 13, 3999-4010 (1993).
65. Evans, T. & Felsenfeld, G. The erythroid-specific transcription factor Eryf1: a new finger protein. Cell 58, 877-85 (1989).
66. Blobel, G.A.W., M.J. Nuclear factors that regulate erythropoiesis. In Disorders of Hemoglobin: Genetics, Pathophysiology, and Clinical Management. 2nd edition, Cambridge University Press, 2009.
67. Weiss, M.J. & Orkin, S.H. GATA transcription factors: key regulators of hematopoiesis. Exp Hematol 23, 99-107 (1995).
68. Munugalavadla, V. et al. Repression of c-kit and its downstream substrates by GATA-1 inhibits cell proliferation during erythroid maturation. Mol Cell Biol 25, 6747-59 (2005).
69. Weiss, M.J., Keller, G. & Orkin, S.H. Novel insights into erythroid development revealed through in vitro differentiation of GATA-1 embryonic stem cells. Genes Dev 8, 1184-97 (1994).
70. Welch, J.J. et al. Global regulation of erythroid gene expression by transcription factor GATA-1. Blood 104, 3136-47 (2004).
71. Pevny, L. et al. Development of hematopoietic cells lacking transcription factor GATA-1. Development 121, 163-72 (1995).
72. Fujiwara, Y., Browne, C.P., Cunniff, K., Goff, S.C. & Orkin, S.H. Arrested development of embryonic red cell precursors in mouse embryos lacking transcription factor GATA-1. Proc Natl Acad Sci U S A 93, 12355-8 (1996).
73. Kim, S.I. & Bresnick, E.H. Transcriptional control of erythropoiesis: emerging mechanisms and principles. Oncogene 26, 6777-94 (2007).
74. Warren, A.J. et al. The oncogenic cysteine-rich LIM domain protein rbtn2 is essential for erythroid development. Cell 78, 45-57 (1994).
75. Robb, L. et al. Absence of yolk sac hematopoiesis from mice with a targeted disruption of the scl gene. Proc Natl Acad Sci U S A 92, 7075-9 (1995).
76. Shivdasani, R.A., Mayer, E.L. & Orkin, S.H. Absence of blood formation in mice lacking the T-cell leukaemia oncoprotein tal-1/SCL. Nature 373, 432-4 (1995).
77. Tsang, A.P., Fujiwara, Y., Hom, D.B. & Orkin, S.H. Failure of megakaryopoiesis and arrested

erythropoiesis in mice lacking the GATA-1 transcriptional cofactor FOG. Genes Dev 12, 1176-88 (1998).
78. Li, L. et al. A requirement for Lim domain binding protein 1 in erythropoiesis. J Exp Med 207, 2543-50 (2010).
79. Liu, P. et al. Bcl11a is essential for normal lymphoid development. Nat Immunol 4, 525-32 (2003).
80. Menzel, S. et al. A QTL influencing F cell production maps to a gene encoding a zinc-finger protein on chromosome 2p15. Nat Genet 39, 1197-9 (2007).
81. Thein, S.L., Menzel, S., Lathrop, M. & Garner, C. Control of fetal hemoglobin: new insights emerging from genomics and clinical implications. Hum Mol Genet 18, R216-23 (2009).
82. Sankaran, V.G. et al. Human fetal hemoglobin expression is regulated by the developmental stage-specific repressor BCL11A. Science 322, 1839-42 (2008).
83. Uda, M. et al. 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, 1620-5 (2008).
84. Lettre, G. et al. 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, 11869-74 (2008).
85. Galarneau, G. et al. Fine-mapping at three loci known to affect fetal hemoglobin levels explains additional genetic variation. Nat Genet 42, 1049-51 (2010).
86. Sedgewick, A.E. et al. BCL11A is a major HbF quantitative trait locus in three different populations with beta-hemoglobinopathies. Blood Cells Mol Dis 41, 255-8 (2008).
87. Danjou, F. et al. Genetic modifiers of beta-thalassemia and clinical severity as assessed by age at first transfusion. Haematologica 97, 989-93 (2012).
88. Sankaran, V.G. et al. Developmental and species-divergent globin switching are driven by BCL11A. Nature 460, 1093-7 (2009).
89. Chen, Z., Luo, H.Y., Steinberg, M.H. & Chui, D.H. BCL11A represses HBG transcription in K562 cells. Blood Cells Mol Dis 42, 144-9 (2009).
90. Xu, J. et al. Transcriptional silencing of {gamma}-globin by BCL11A involves long-range interactions and cooperation with SOX6. Genes Dev 24, 783-98 (2010).
91. Jawaid, K., Wahlberg, K., Thein, S.L. & Best, S. Binding patterns of BCL11A in the globin and GATA1 loci and characterization of the BCL11A fetal hemoglobin locus. Blood Cells Mol Dis 45, 140-6 (2010).
92. Xu, J. et al. Corepressor-dependent silencing of fetal hemoglobin expression by BCL11A. Proc Natl Acad Sci U S A 110, 6518-23 (2013).
93. Yien, Y.Y. & Bieker, J.J. EKLF/KLF1, a tissue-restricted integrator of transcriptional control, chromatin remodeling, and lineage determination. Mol Cell Biol 33, 4-13 (2013).
94. Chen, X. & Bieker, J.J. Unanticipated repression function linked to erythroid Kruppel-like factor. Mol Cell Biol 21, 3118-25 (2001).
95. Chen, X. & Bieker, J.J. Stage-specific repression by the EKLF transcriptional activator. Mol Cell Biol 24, 10416-24 (2004).
96. Siatecka, M., Xue, L. & Bieker, J.J. Sumoylation of EKLF promotes transcriptional repression and is involved in inhibition of megakaryopoiesis. Mol Cell Biol 27, 8547-60 (2007).
97. Guy, L.G., Mei, Q., Perkins, A.C., Orkin, S.H. & Wall, L. Erythroid Kruppel-like factor is essential for beta-globin gene expression even in absence of gene competition, but is not sufficient to induce the switch from gamma-globin to beta-globin gene expression. Blood 91, 2259-63 (1998).
98. Tewari, R. et al. Erythroid Kruppel-like factor (EKLF) is active in primitive and definitive erythroid cells and is required for the function of 5'HS3 of the beta-globin locus control region. EMBO J 17, 2334-41 (1998).
99. Frontelo, P. et al. Novel role for EKLF in megakaryocyte lineage commitment. Blood 110, 3871-80 (2007).
100. Nuez, B., Michalovich, D., Bygrave, A., Ploemacher, R. & Grosveld, F. Defective haematopoiesis in fetal liver resulting from inactivation of the EKLF gene. Nature 375, 316-8 (1995).
101. Perkins, A.C., Sharpe, A.H. & Orkin, S.H. Lethal beta-thalassaemia in mice lacking the erythroid CACCC-transcription factor EKLF. Nature 375, 318-22 (1995).
102. Perkins, A.C., Peterson, K.R., Stamatoyannopoulos, G., Witkowska, H.E. & Orkin, S.H. Fetal

expression of a human Agamma globin transgene rescues globin chain imbalance but not hemolysis in EKLF null mouse embryos. Blood 95, 1827-33 (2000).
103. Drissen, R. et al. The active spatial organization of the beta-globin locus requires the transcription factor EKLF. Genes Dev 18, 2485-90 (2004).
104. Wijgerde, M. et al. The role of EKLF in human beta-globin gene competition. Genes Dev 10, 2894-902 (1996).
105. Pilon, A.M. et al. Failure of terminal erythroid differentiation in EKLF-deficient mice is associated with cell cycle perturbation and reduced expression of E2F2. Mol Cell Biol 28, 7394-401 (2008).
106. Zhou, D., Liu, K., Sun, C.W., Pawlik, K.M. & Townes, T.M. KLF1 regulates BCL11A expression and gamma- to beta-globin gene switching. Nat Genet 42, 742-4 (2010).
107. Borg, J. et al. Haploinsufficiency for the erythroid transcription factor KLF1 causes hereditary persistence of fetal hemoglobin. Nat Genet 42, 801-5 (2010).
108. Singleton, B.K., Burton, N.M., Green, C., Brady, R.L. & Anstee, D.J. Mutations in EKLF/KLF1 form the molecular basis of the rare blood group In(Lu) phenotype. Blood 112, 2081-8 (2008).
109. Arnaud, L. et al. A dominant mutation in the gene encoding the erythroid transcription factor KLF1 causes a congenital dyserythropoietic anemia. Am J Hum Genet 87, 721-7 (2010).
110. Satta, S. et al. Compound heterozygosity for KLF1 mutations associated with remarkable increase of fetal hemoglobin and red cell protoporphyrin. Haematologica 96, 767-70 (2011).
111. Siatecka, M. et al. Severe anemia in the Nan mutant mouse caused by sequence-selective disruption of erythroid Kruppel-like factor. Proc Natl Acad Sci U S A 107, 15151-6 (2010).
112. Heruth, D.P. et al. Mutation in erythroid specific transcription factor KLF1 causes Hereditary Spherocytosis in the Nan hemolytic anemia mouse model. Genomics 96, 303-7 (2010).
113. White, R.A. et al. Hematologic characterization and chromosomal localization of the novel dominantly inherited mouse hemolytic anemia, neonatal anemia (Nan). Blood Cells Mol Dis 43, 141-8 (2009).
114. Weatherall, D.J. & Clegg, J.B. Inherited haemoglobin disorders: an increasing global health problem. Bull World Health Organ 79, 704-12 (2001).
115. Angastiniotis, M. & Modell, B. Global epidemiology of hemoglobin disorders. Ann N Y Acad Sci 850, 251-69 (1998).
116. Ingram, V.M. A specific chemical difference between the globins of normal human and sickle-cell anaemia haemoglobin. Nature 178, 792-4 (1956).
117. Pauling, L., Itano, H.A. & et al. Sickle cell anemia, a molecular disease. Science 109, 443 (1949).
118. Paraskevas, F. Wintrobe's Clinical Hematology. 13th edition, Lippincott Williams & Wilkins, 2014.
119. Madigan, C. & Malik, P. Pathophysiology and therapy for haemoglobinopathies. Part I: sickle cell disease. Expert Rev Mol Med 8, 1-23 (2006).
120. Weatherall, D.J. Phenotype-genotype relationships in monogenic disease: lessons from the thalassaemias. Nat Rev Genet 2, 245-55 (2001).
121. Katsantoni, E.Z. et al. Persistent gamma-globin expression in adult transgenic mice is mediated by HPFH-2, HPFH-3, and HPFH-6 breakpoint sequences. Blood 102, 3412-9 (2003).
122. Stamatoyannopoulos, G. The Molecular Basis of Blood Diseases. 3rd edition,W.B. Saunders, 2001.
123. Gumucio, D.L. et al. Nuclear proteins that bind the human gamma-globin gene promoter: alterations in binding produced by point mutations associated with hereditary persistence of fetal hemoglobin. Mol Cell Biol 8, 5310-22 (1988).
124. Ikuta, T. & Kan, Y.W. In vivo protein-DNA interactions at the beta-globin gene locus. Proc Natl Acad Sci U S A 88, 10188-92 (1991).
125. Jane, S.M., Ney, P.A., Vanin, E.F., Gumucio, D.L. & Nienhuis, A.W. Identification of a stage selector element in the human gamma-globin gene promoter that fosters preferential interaction with the 5' HS2 enhancer when in competition with the beta-promoter. EMBO J 11, 2961-9 (1992).
126. Mantovani, R. et al. A protein factor binding to an octamer motif in the gamma-globin promoter disappears upon induction of differentiation and hemoglobin synthesis in K562 cells. Nucleic Acids Res 15, 9349-64 (1987).
127. Mantovani, R. et al. The effects of HPFH mutations in the human gamma-globin promoter on binding of ubiquitous and erythroid specific nuclear factors. Nucleic Acids Res 16, 7783-97 (1988).

128. Martin, D.I., Tsai, S.F. & Orkin, S.H. Increased gamma-globin expression in a nondeletion HPFH mediated by an erythroid-specific DNA-binding factor. Nature 338, 435-8 (1989).
129. Chang, Y.C., Smith, K.D., Moore, R.D., Serjeant, G.R. & Dover, G.J. An analysis of fetal hemoglobin variation in sickle cell disease: the relative contributions of the X-linked factor, beta-globin haplotypes, alpha-globin gene number, gender, and age. Blood 85, 1111-7 (1995).
130. Chang, Y.P. et al. The relative importance of the X-linked FCP locus and beta-globin haplotypes in determining haemoglobin F levels: a study of SS patients homozygous for beta S haplotypes. Br J Haematol 96, 806-14 (1997).
131. Garner, C.P., Tatu, T., Best, S., Creary, L. & Thein, S.L. Evidence of genetic interaction between the beta-globin complex and chromosome 8q in the expression of fetal hemoglobin. Am J Hum Genet 70, 793-9 (2002).
132. Labie, D. et al. Common haplotype dependency of high G gamma-globin gene expression and high Hb F levels in beta-thalassemia and sickle cell anemia patients. Proc Natl Acad Sci U S A 82, 2111-4 (1985).
133. Steinberg, M.H.F., B.G.; Higgs, D.R.; Nagel, R.L. Disorders of Hemoglobin: Genetics, Pathophysiology, and Clinical Management. 2nd edition, Cambridge University Press, 2009.
134. Bhatia, M. & Walters, M.C. Hematopoietic cell transplantation for thalassemia and sickle cell disease: past, present and future. Bone Marrow Transplant 41, 109-17 (2008).
135. Shenoy, S. Has stem cell transplantation come of age in the treatment of sickle cell disease? Bone Marrow Transplant 40, 813-21 (2007).
136. Gambari, R. Foetal haemoglobin inducers and thalassaemia: novel achievements. Blood Transfus 8, 5-7 (2010).
137. DeSimone, J., Heller, P., Hall, L. & Zwiers, D. 5-Azacytidine stimulates fetal hemoglobin synthesis in anemic baboons. Proc Natl Acad Sci U S A 79, 4428-31 (1982).
138. Dover, G.J., Charache, S.H., Boyer, S.H., Talbot, C.C., Jr. & Smith, K.D. 5-Azacytidine increases fetal hemoglobin production in a patient with sickle cell disease. Prog Clin Biol Res 134, 475-88 (1983).
139. Glauber, J.G., Wandersee, N.J., Little, J.A. & Ginder, G.D. 5'-flanking sequences mediate butyrate stimulation of embryonic globin gene expression in adult erythroid cells. Mol Cell Biol 11, 4690-7 (1991).
140. Ikuta, T., Kan, Y.W., Swerdlow, P.S., Faller, D.V. & Perrine, S.P. Alterations in protein-DNA interactions in the gamma-globin gene promoter in response to butyrate therapy. Blood 92, 2924-33 (1998).
141. Weinberg, R.S. et al. Butyrate increases the efficiency of translation of gamma-globin mRNA. Blood 105, 1807-9 (2005).
142. Pawliuk, R. et al. Correction of sickle cell disease in transgenic mouse models by gene therapy. Science 294, 2368-71 (2001).
143. Imren, S. et al. Permanent and panerythroid correction of murine beta thalassemia by multiple lentiviral integration in hematopoietic stem cells. Proc Natl Acad Sci U S A 99, 14380-5 (2002).
144. Cavazzana-Calvo, M. et al. Transfusion independence and HMGA2 activation after gene therapy of human beta-thalassaemia. Nature 467, 318-22 (2010).
145. Weissman, I.L. & Shizuru, J.A. The origins of the identification and isolation of hematopoietic stem cells, and their capability to induce donor-specific transplantation tolerance and treat autoimmune diseases. Blood 112, 3543-53 (2008).
146. Mikkola, H.K. & Orkin, S.H. The journey of developing hematopoietic stem cells. Development 133, 3733-44 (2006).
147. Baron, M.H., Vacaru, A. & Nieves, J. Erythroid development in the mammalian embryo. Blood Cells Mol Dis 51, 213-9 (2013).
148. Fujiwara, Y., Chang, A.N., Williams, A.M. & Orkin, S.H. Functional overlap of GATA-1 and GATA-2 in primitive hematopoietic development. Blood 103, 583-5 (2004).
149. Tsai, F.Y. et al. An early haematopoietic defect in mice lacking the transcription factor GATA-2. Nature 371, 221-6 (1994).
150. Mancini, E. et al. FOG-1 and GATA-1 act sequentially to specify definitive megakaryocytic and erythroid progenitors. EMBO J 31, 351-65 (2012).
151. Drissen, R. et al. The erythroid phenotype of EKLF-null mice: defects in hemoglobin metabolism

and membrane stability. Mol Cell Biol 25, 5205-14 (2005).
152. Hodge, D. et al. A global role for EKLF in definitive and primitive erythropoiesis. Blood 107, 3359-70 (2006).
153. Isern, J., Fraser, S.T., He, Z., Zhang, H. & Baron, M.H. Dose-dependent regulation of primitive erythroid maturation and identity by the transcription factor Eklf. Blood 116, 3972-80 (2010).
154. Siatecka, M. & Bieker, J.J. The multifunctional role of EKLF/KLF1 during erythropoiesis. Blood 118, 2044-54 (2011).
155. Bieker, J.J. Putting a finger on the switch. Nat Genet 42, 733-4 (2010).
156. Porcher, C. et al. The T cell leukemia oncoprotein SCL/tal-1 is essential for development of all hematopoietic lineages. Cell 86, 47-57 (1996).
157. Robb, L. et al. The scl gene product is required for the generation of all hematopoietic lineages in the adult mouse. EMBO J 15, 4123-9 (1996).
158. Loose, M., Swiers, G. & Patient, R. Transcriptional networks regulating hematopoietic cell fate decisions. Curr Opin Hematol 14, 307-14 (2007).
159. Elefanty, A.G., Begley, C.G., Hartley, L., Papaevangeliou, B. & Robb, L. SCL expression in the mouse embryo detected with a targeted lacZ reporter gene demonstrates its localization to hematopoietic, vascular, and neural tissues. Blood 94, 3754-63 (1999).
160. Wang, Q. et al. Disruption of the Cbfa2 gene causes necrosis and hemorrhaging in the central nervous system and blocks definitive hematopoiesis. Proc Natl Acad Sci U S A 93, 3444-9 (1996).
161. Okuda, T., van Deursen, J., Hiebert, S.W., Grosveld, G. & Downing, J.R. AML1, the target of multiple chromosomal translocations in human leukemia, is essential for normal fetal liver hematopoiesis. Cell 84, 321-30 (1996).
162. Yokomizo, T. et al. Runx1 is involved in primitive erythropoiesis in the mouse. Blood 111, 4075-80 (2008).
163. Mukhopadhyay, M. et al. Functional ablation of the mouse Ldb1 gene results in severe patterning defects during gastrulation. Development 130, 495-505 (2003).
164. Soler, E. et al. The genome-wide dynamics of the binding of Ldb1 complexes during erythroid differentiation. Genes Dev 24, 277-89 (2010).
165. Mylona, A. et al. Genome-wide analysis shows that Ldb1 controls essential hematopoietic genes/pathways in mouse early development and reveals novel players in hematopoiesis. Blood 121, 2902-13 (2013).
166. Tober, J. et al. The megakaryocyte lineage originates from hemangioblast precursors and is an integral component both of primitive and of definitive hematopoiesis. Blood 109, 1433-41 (2007).
167. Tober, J., McGrath, K.E. & Palis, J. Primitive erythropoiesis and megakaryopoiesis in the yolk sac are independent of c-myb. Blood 111, 2636-9 (2008).
168. Jayapal, S.R. et al. Down-regulation of Myc is essential for terminal erythroid maturation. J Biol Chem 285, 40252-65 (2010).
169. Basu, P. et al. EKLF and KLF2 have compensatory roles in embryonic beta-globin gene expression and primitive erythropoiesis. Blood 110, 3417-25 (2007).
170. Matsumoto, N. et al. Developmental regulation of yolk sac hematopoiesis by Kruppel-like factor 6. Blood 107, 1357-65 (2006).
171. Griffin, C.T., Brennan, J. & Magnuson, T. The chromatin-remodeling enzyme BRG1 plays an essential role in primitive erythropoiesis and vascular development. Development 135, 493-500 (2008).
172. Kusakabe, M. et al. c-Maf plays a crucial role for the definitive erythropoiesis that accompanies erythroblastic island formation in the fetal liver. Blood 118, 1374-85 (2011).
173. Tsunoda, T. et al. Immune-related zinc finger gene ZFAT is an essential transcriptional regulator for hematopoietic differentiation in blood islands. Proc Natl Acad Sci U S A 107, 14199-204 (2010).
174. Ueda, T. et al. Critical role of the p400/mDomino chromatin-remodeling ATPase in embryonic hematopoiesis. Genes Cells 12, 581-92 (2007).
175. Woo, A.J. et al. Identification of ZBP-89 as a novel GATA-1-associated transcription factor involved in megakaryocytic and erythroid development. Mol Cell Biol 28, 2675-89 (2008).
176. Wilson, N.K. et al. Combinatorial transcriptional control in blood stem/progenitor cells: genome-

wide analysis of ten major transcriptional regulators. Cell Stem Cell 7, 532-44 (2010).
177. Lancrin, C. et al. GFI1 and GFI1B control the loss of endothelial identity of hemogenic endothelium during hematopoietic commitment. Blood 120, 314-22 (2012).
178. Isern, J. et al. Single-lineage transcriptome analysis reveals key regulatory pathways in primitive erythroid progenitors in the mouse embryo. Blood 117, 4924-34 (2011).
179. Dumitriu, B. et al. Sox6 is necessary for efficient erythropoiesis in adult mice under physiological and anemia-induced stress conditions. PLoS One 5, e12088 (2010).
180. Esteghamat, F. et al. Erythropoiesis and globin switching in compound Klf1::Bcl11a mutant mice. Blood 121, 2553-62 (2013).
181. Miccio, A. et al. NuRD mediates activating and repressive functions of GATA-1 and FOG-1 during blood development. EMBO J 29, 442-56 (2010).
182. Beck, L. et al. The phosphate transporter PiT1 (Slc20a1) revealed as a new essential gene for mouse liver development. PLoS One 5, e9148 (2010).
183. Forand, A. et al. EKLF-driven PIT1 expression is critical for mouse erythroid maturation in vivo and in vitro. Blood 121, 666-78 (2013).
184. Cui, S. et al. Compound loss of function of nuclear receptors Tr2 and Tr4 leads to induction of murine embryonic beta-type globin genes. Blood 125, 1477-87 (2015).
185. Shyr, C.R. et al. Roles of testicular orphan nuclear receptors 2 and 4 in early embryonic development and embryonic stem cells. Endocrinology 150, 2454-62 (2009).
186. Tanabe, O. et al. Embryonic and fetal beta-globin gene repression by the orphan nuclear receptors, TR2 and TR4. EMBO J 26, 2295-306 (2007).

Chapter 2

Erythropoiesis and globin switching in compound *Klf1::Bcl11a* mutant mice

Fatemehsadat Esteghamat[1], Nynke Gillemans[1], Ivan Bilic[2], Emile van den Akker[3], **Ileana Cantù**[1], Teus van Gent[1,4], Ursula Klingmüller[5], Kirsten van Lom[6], Marieke von Lindern[3], Frank Grosveld[1,4,7], Thamar Bryn van Dijk[1], Meinrad Busslinger[2], Sjaak Philipsen[1,4*]

1 Department of Cell Biology, Erasmus MC, Rotterdam, The Netherlands
2 Research Institute of Molecular Pathology, Vienna, Austria
3 Sanquin Research Department of Hematopoiesis, Amsterdam, The Netherlands
4 Netherlands Consortium for Systems Biology, Erasmus MC, Rotterdam, The Netherlands
5 German Cancer Research Center, Heidelberg, Germany
6 Department of Hematology, Erasmus MC, Rotterdam, The Netherlands
7 Centre for Biomedical Genetics, Erasmus MC, Rotterdam, The Netherlands
* Corresponding author

Blood. 2013 Mar 28;121(13):2553-62

ABSTRACT

B-Cell/Lymphoma 11A (BCL11A) down-regulation in human primary adult erythroid progenitors results in elevated expression of fetal γ-globin. Recent reports showed that BCL11A expression is activated by KLF1, leading to γ-globin repression. To study regulation of erythropoiesis and globin expression by KLF1 and BCL11a in an *in vivo* model, we used mice carrying a human β-globin locus transgene with combinations of *Klf1* knockout, *Bcl11a* floxed and *EpoR*^{Cre} knockin alleles. We found a higher percentage of reticulocytes in adult *Klf1*$^{wt/ko}$ mice and a mild compensated anemia in *Bcl11a*$^{cko/cko}$ mice. These phenotypes were more pronounced in compound *Klf1*$^{wt/ko}$*::Bcl11a*$^{cko/cko}$ animals. Analysis of *Klf1*$^{wt/ko}$, *Bcl11a*$^{cko/cko}$ and *Klf1*$^{wt/ko}$*::Bcl11a*$^{cko/cko}$ mutant embryos demonstrated increased expression of mouse embryonic globins during fetal development. Expression of human γ-globin remained high in *Bcl11a*$^{cko/cko}$ embryos during fetal development, and this was further augmented in *Klf1*$^{wt/ko}$*::Bcl11a*$^{cko/cko}$ embryos. After birth, expression of human γ-globin and mouse embryonic globins decreased in *Bcl11a*$^{cko/cko}$ and *Klf1*$^{wt/ko}$*::Bcl11a*$^{cko/cko}$ mice, but the levels remained much higher than those observed in control animals. Collectively, our data support an important role for the KLF1-BCL11A axis in erythroid maturation and developmental regulation of globin expression.

INTRODUCTION

Sickle cell anemia (SCA) and β-thalassemia are the most common monogenic disorders in the human population, with an estimated 300,000 seriously affected children born annually[1]. SCA is caused by expression of a pathological β-globin missense mutant (Glu6Val), while β-globin insufficiency underlies β-thalassemia. The symptoms of these diseases are ameliorated by high levels of γ-globin, a β-like globin that is expressed at the fetal stages of human development. Fetal Hemoglobin (HbF; α2γ2) is the dominant hemoglobin during fetal liver erythropoiesis. After birth, when the site of erythropoiesis shifts to the bone marrow, γ-globin is gradually replaced by β-globin and adult hemoglobin (HbA; α2β2) becomes the major hemoglobin. In adults, HbF levels normally decrease to less than 1%[2]. For this reason, newborn patients with β-thalassemia or SCA will start manifesting disease symptoms during the first year of life. Since HbF can substitute for HbA in adults, HbF induction would alleviate the symptoms of β-thalassemia and SCA. Currently, hydroxyurea is used in the clinic with considerable success, elevating HbF levels in 40-50% of patients with hemoglobin related disorders[3,4]. The variable response of individual patients and the mechanisms by which hydroxyurea induces HbF are poorly understood. Safe pharmacological reactivation of γ-globin expression therefore remains a very attractive approach[2]. Thus, understanding the molecular details of γ- to β-globin switching remains an important goal, because this will reveal novel targets for development of more specific therapeutic intervention.

To study the mechanisms underlying globin switching, human β-globin locus transgenic mice have been established as models for developmental regulation of human β-like globin gene expression[2]. The human β-globin locus contains five developmentally regulated globin genes in the order 5'-ε (embryonic) -^{G}γ-^{A}γ (fetal) -δ-β (adult) -3', while the mouse β-globin locus harbors four genes in the order 5'-εγ-βh1 (embryonic) -βmaj-βmin (fetal/adult) -3'[2]. The α-globin loci in human and mouse contain an embryonic ζ- and two fetal/adult α-globin genes[5]. In mice, εγ, βh1 and ζ are embryonic globins expressed in primitive erythrocytes. Their expression is silenced in definitive erythrocytes which express α- and β-globins at the fetal and adult stages of development[6]. In human β-globin locus transgenic mice, the γ-globin genes behave like embryonic/early fetal genes. Switching to β-globin expression takes place between embryonic day (E) 12 and E14[7], when the fetal liver is the major site of erythropoiesis. It has recently been demonstrated that this difference in developmental timing of globin switching is linked to alterations in the expression of the B-Cell/Lymphoma 11A (BCL11A) repressor protein, creating a *trans*-acting environment in the mouse fetal liver that is non-permissive for γ-globin expression[8].

In humans, genome-wide association studies (GWAS) revealed a strong correlation of HbF levels with several SNPs located in the *BCL11A* gene[9,10]. Subsequently, BCL11A was reported as a critical mediator of γ-globin silencing in the adult [8,11]. It functions as a repressor through binding to *cis*-regulatory elements in the β-globin locus, and interacts with the NuRD repressor complex and the GATA1 and FOG1 transcription factors. It binds to the third hypersensitivity site (5'HS3) in the β-globin locus control region (LCR) and a region downstream of the ^{A}γ-globin gene. BCL11A downregulation in sorted and expanded CD34^{+} human hematopoietic progenitor cells elevates γ–globin expression. BCL11A depletion does not affect the expression of well-known transcription factors regulating erythropoiesis such as KLF1 (previously known as EKLF[12]), NF-E2, GATA1 and FOG1[11]. In contrast, it was recently shown that KLF1 activates BCL11A expression[13,14]. KLF1 has a critical role in erythroid development, and in mice its expression increases three-fold upon the transition from primitive to definitive erythropoiesis[15]. Analysis of *Klf1* knockout embryos revealed lethality at E14.5 due to anemia caused by disrupted fetal liver erythropoiesis[16,17]. Remarkably, KLF1 is absolutely required for activation of β-globin expression, while the α-like and embryonic β-like genes are still highly expressed in *Klf1* knockout embryos[16,17]. Furthermore, *Klf1*

knockout embryos carrying a human β-globin locus transgene fail to activate the human β-globin gene, while the γ-globin genes are fully expressed[18,19].
Mutations in human *KLF1* are associated with a spectrum of phenotypes, such as the In(Lu) blood group[20], zinc protoporphyria[21], increased HbA2[22], congenital dyserythropoietic anemia (CDA)[23] and hereditary persistence fetal hemoglobin (HPFH)[13]. Analysis of the HPFH phenotype has led to the proposal that KLF1 has a dual role in γ-globin suppression, through its preferential activation of the β-globin gene and as a key activator of expression of the BCL11A repressor protein[13,14]. Here, we used mice carrying a human β-*globin* locus transgene (PAC8.1)[7] and crossed these with mice carrying *Klf1* knockout[16], *Bcl11a* floxed, and *EpoR-Cre* knockin[24] alleles, to interrogate the impact of these two key molecules, KLF1 and BCL11A, on erythropoiesis and globin gene regulation.

MATERIAL AND METHODS

Mice

All animal studies were approved by the Erasmus MC Animal Ethics Committee. Transgenic mouse strains used were: human β-globin locus transgenic line PAC8.1[7]; *Klf1* knockout allele[16]; knockin of Cre recombinase in the Epo receptor locus (*EpoR*^{Cre})[24] and *Bcl11a* floxed alleles (Suppl. Fig. 1A). Genotyping was performed by PCR using DNA isolated from tail snips using primers listed in Supplementary Table 1. Embryos were collected at E14.5 and E18.5; head DNA was used for genotyping by PCR. Inactivation of the *Bcl11a* gene was analyzed using PCR to detect Cre-mediated recombination at the *Bcl11a* locus. To induce stress erythropoiesis, mice were injected subcutaneously with 0.4% (w/v) phenylhydrazine (PHZ; Sigma-Aldrich, , St. Louis, MO, USA) in saline (12 μl/g body weight) for 2 consecutive days (Day 1 and 2). Mice were collected at Day 5 for analysis.

Blood analysis

Peripheral blood was collected from the mandibular vein of >10-week old mice and standard blood parameters were measured using an automated hematologic analyzer (Scil Vet ABC, Viernheim, Germany). Blood smears were stained with May-Grünwald-Giemsa and scored double-blinded by KvL. Cytospins of peripheral blood were stained with a combination of neutral benzidine and histological dyes[25].

Statistical tests

Statistical analysis of blood parameters was performed using ANOVA with Bonferroni correction; FACS data and globin expression results were analyzed using Mann-Whitney tests. All statistical tests were implemented by Stata 11.1 software (StataCorp, College Station, TX, USA). Excel 2003 (Microsoft, Redmond, WA, USA) was employed to draw the graphs. Values are displayed ± standard error of the mean (SEM).
RNA isolation, S1 nuclease protection assays and QRT-PCR analyses
RNA was isolated from mouse erythroid tissues using TRI reagent (Sigma-Aldrich). 2.5 μg of RNA was used for each S1 nuclease protection assay as described[13]. Synthesis of cDNA and quantitative real-time PCR (QRT-PCR) were performed as described[26]. Expression of different globin genes was measured using the Ct values; Ct values obtained for α-globin expression were used for normalization. The oligonucleotides used for QRT-PCR are listed in Supplementary Table 1.

Protein extraction and western blotting

Whole cell lysates were prepared using RIPA buffer[26]. To visualize γ-globin expression at the protein level, whole cell lysates of ~3x10^5 red cells were loaded on 12.5% SDS-PAGE, the gels were transferred to nitrocellulose membranes and probed with γ-globin antibody (51-7, sc-

21756, Santa Cruz Biotechnology, Santa Cruz, CA, USA). Staining for actin served as loading control (I-19, sc-1616, Santa Cruz Biotechnology).

Flowcytometry analysis

Whole blood and single cell suspensions collected from bone marrow and spleen were washed twice with phosphate buffered saline (PBS) and then resuspended in PBS containing 1% (w/v) bovine serum albumin and 1mM EDTA. ~10^6 cells were incubated for 30 min with CD71-FITC (553266, BD Biosciences, San Jose, CA, USA) and Ter119-PE (553673, BD Biosciences) antibodies (diluted 1:200) and DRAQ5 (diluted 1:500; Biostatus Ltd, Shepshed, UK) in a final volume of 100 µl. The cells were washed and live cells distinguished negatively by 7-aminoactinomycin D staining (A1310, Invitrogen, Carlsbad, CA, USA). Cells were measured on a FACScan instrument (BD Biosciences) and data were analyzed with FlowJo software (v7.6, Tree Star Inc, Stanford, USA). Bone marrow cells were sorted using Ter119 antibody and MACS separation columns according to the manufacturer's instructions (Miltenyi Biotec, Bergisch Gladbach, Germany).

Erythropoietin assays

Whole blood was centrifuged at 4 krpm for 10 min, sera were collected and stored at -20°C. Erythropoietin concentrations were measured using an ELISA-based assay (mouse/rat Epo immunoassay, Quantikine, R&D systems, Minneapolis, MN, USA).

Immunohistochemistry

Immunohistochemistry for γ-globin expression was performed as described [27].

RESULTS

Generation of compound transgenic animals

The role of KLF1 in erythropoiesis has been widely studied[12,13], while much less is known about the function of BCL11A in red cell development. The role of KLF1 and BCL11A in the switching of fetal to adult *β-like* globin gene expression is of particular interest[13,14]. Since mice do not have a fetal *β-like* globin gene, we used animals carrying a human *β-globin* locus transgene (line PAC8.1 [7]) to investigate this. In these animals, the human *γ-globin* genes are expressed at the embryonic and early fetal stages; the switch from γ-globin to β-globin expression takes place between E12.5 and E14.5 when the fetal liver is the major site of erythropoiesis[6]. We crossed the PAC8.1 mice with mice carrying a knockout allele of *Klf1*[16] and a floxed allele of *Bcl11a* (Suppl. Fig. 1A,B). To obtain erythroid-specific ablation of BCL11A expression we employed a knockin allele of Cre recombinase in the erythropoietin receptor locus (*EpoR^{Cre}*)[24]. We studied mice with four genotypes: (1) *PAC8.1::Bcl11a$^{fl/fl}$*; (2) *PAC8.1::Klf1$^{ko/wt}$::Bcl11a$^{fl/fl}$*; (3) *PAC8.1::Bcl11a$^{fl/fl}$::EpoR$^{Cre/wt}$*; and (4) *PAC8.1::Klf1$^{ko/wt}$::Bcl11a$^{fl/fl}$::EpoR$^{Cre/wt}$*. For reasons of clarity, we will refer to these mice as (1) control; (2) *Klf1$^{wt/ko}$*; (3) *Bcl11a$^{cko/cko}$* and (4) *Klf1$^{wt/ko}$::Bcl11a$^{cko/cko}$*. To ascertain erythroid-specificity of recombination, DNA was isolated form bone marrow, spleen and cells from control and *Bcl11a$^{cko/cko}$* mice. PCR amplification demonstrated Cre-mediated recombination at the *Bcl11a* locus in *Bcl11a$^{cko/cko}$* bone marrow and spleen DNA only; the PCR products were sequenced to confirm their identity (Suppl. Fig. 1C-D). Expression levels of BCL11A in Ter119-MACS selected adult bone marrow cells of *Bcl11a$^{cko/cko}$* and *Klf1$^{wt/ko}$::Bcl11a$^{cko/cko}$* mice were decreased to less than 30% of those observed in the control mice, in agreement with previous reports[24,28], while a reduction to ~60% was observed in samples from *Klf1$^{wt/ko}$* mice, consistent with the notion that expression of BCL11A is activated by KLF1[13,14] (Suppl. Fig. 1E). Mice with ubiquitous BCL11A deficiency die prenatally from unknown causes[29]. In contrast, erythroid-specific ablation of BCL11A did not affect viability since we obtained *Bcl11a$^{cko/cko}$* and

Klf1$^{wt/ko}$*::Bcl11a*$^{cko/cko}$ animals at the expected Mendelian ratios (Suppl. Table 2). These animals appeared healthy, were fertile and displayed no gross morphological abnormalities. We therefore first analyzed standard hematological parameters of adult mice.

Hematological parameters of compound mutant mice

To investigate the impact of BCL11A deficiency, either with or without *Klf1* mutation, on adult erythropoiesis we determined standard hematological parameters of the mutant mice (Fig. 1). With the exception of a small reduction in MCV, parameters of *Klf1*$^{wt/ko}$ animals were similar to those observed in the control animals. *Bcl11a*$^{cko/cko}$ animals displayed a small but significant reduction of HCT, RBC and HGB values. The reductions in these values were more pronounced in the *Klf1*$^{wt/ko}$*::Bcl11a*$^{cko/cko}$ animals. In addition, *Klf1*$^{wt/ko}$*::Bcl11a*$^{cko/cko}$ mice displayed small but significantly increased values for MCV, MCH and MCHC. We note that the observed differences were unrelated to the gender of the mice ($p > 0.05$ in all cases, data not shown).

Collectively, these results indicate that *Bcl11a*$^{cko/cko}$ and *Klf1*$^{wt/ko}$*::Bcl11a*$^{cko/cko}$ mice display a mild anemia.

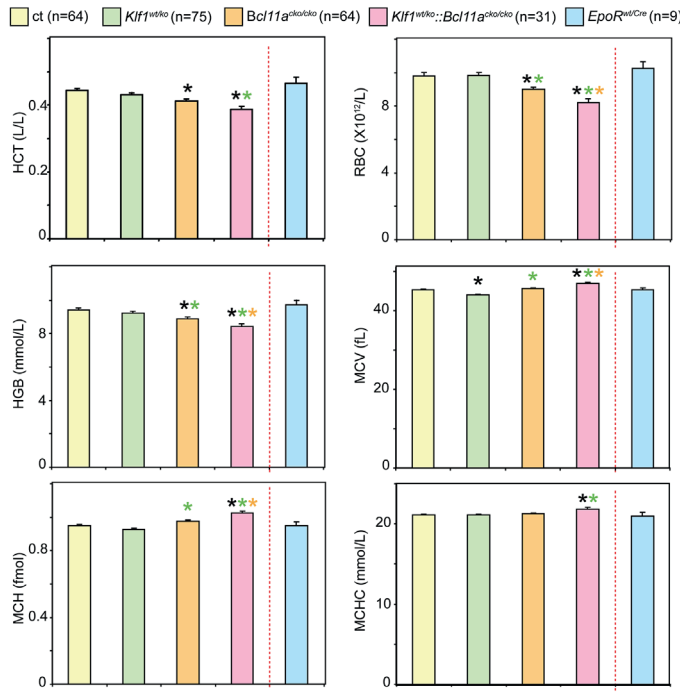


Figure 1. Analysis of hematological parameters. Hematological parameter analysis revealed a mild compensated anemia in *Bcl11a*$^{cko/cko}$ and *Klf1*$^{wt/ko}$*::Bcl11a*$^{cko/cko}$ mice.*: $p<0.05$ between mutant and control groups; when $p<0.05$ between mutant groups this is indicated by a color-matched asterisk. Abbreviations are: HCT = hematocrit; RBC = red blood cell count, HGB = hemoglobin, MCV = mean corpuscular volume, MCH = mean corpuscular hemoglobin; MCHC = mean corpuscular hemoglobin concentration. (ct n=64, *Klf1*$^{wt/ko}$ n=75, *Bcl11a*$^{cko/cko}$ n=64, K*lf1*$^{wt/ko}$*::Bcl11a*$^{cko/cko}$ n=31 and *EpoR*$^{wt/cre}$ n=9)

***Klf1*$^{wt/ko}$ and *Klf1*$^{wt/ko}$*::Bcl11a*$^{cko/cko}$ mice display reticulocytosis**

Blood smears stained with May-Grünwald-Giemsa indicated that the percentage of polychromatic erythrocytes was higher in blood from *Klf1*$^{wt/ko}$ mice and further increased in *Klf1*$^{wt/ko}$*::Bcl11a*$^{cko/cko}$ blood. Furthermore, this analysis suggested that there were more erythrocytes containing nuclear remnants (Howell-Jolly bodies) in the blood of *Klf1*$^{wt/ko}$*::Bcl11a*$^{cko/cko}$ mice, compared with that of control, *Klf1*$^{wt/ko}$ or *Bcl11a*$^{cko/cko}$ mice (data not shown). Because these results indicated that the KLF1-BCL11A axis plays a role in erythroid maturation, we investigated this in more detail using flowcytometry.

Reticulocytes are usually measured by flowcytometry analysis of the RNA content of red blood cells [30]. The transferrin receptor (CD71) is expressed by erythroid precursor cells committed

to differentiation. Its expression peaks at the basophilic erythroblast stage and gradually decreases during terminal differentiation and maturation. Although the CD71 molecule is absent on mature erythrocytes, it is still present on the surface of circulating immature erythrocytes [31]. Indeed, we found that CD71 staining correlates well with Thiazole Orange staining for RNA-containing reticulocytes (data not shown). Reticulocyte counts are higher in young mice and stabilize after 6 weeks of age [32]. Therefore, we only included >10-week old mice analysis of CD71 expression as marker for reticulocytes and Ter119 as a marker for all erythroid cells (Fig. 2A). We found increased CD71 expression in all mutant animals. This increase was moderate but significant in *Klf1*$^{wt/ko}$ mice (4.9±0.40, compared with 2.1±0.12 in control and 2.8±0.23 in *Bcl11a*$^{cko/cko}$ mice). We observed a more dramatic increase in CD71 expression in *Klf1*$^{wt/ko}$*::Bcl11a*$^{cko/cko}$ mice (8.0±1.10), this difference was statistically significant compared with all other groups (Fig. 2B). In addition, staining with Ter119 and DRAQ5, which is a vital DNA stain, revealed an increased percentage of DRAQ5-positive erythrocytes in all mutants; this was significant in the *Bcl11a*$^{cko/cko}$ and *Klf1*$^{wt/ko}$*::Bcl11a*$^{cko/cko}$ (1.2±0.16 And 1.2±0.23 respectively, compared to 0.67±0.2 in the controls; Fig. 2C).

The increase in CD71- and DRAQ5-positive cells expression suggests reticulocytosis that is usually observed in response to anemic stress. Hypoxia stimulates erythropoiesis thereby increasing the number of reticulocytes released in the circulation. We therefore determined the serum levels of erythropoietin (Epo), the major growth factor driving erythroid expansion. Despite the higher percentage of CD71-positive erythroid cells in peripheral blood, Epo levels in *Klf1*$^{wt/ko}$ mice were similar to those observed in control mice (Fig. 2D). This suggests that the increased percentage of CD71-positive cells is not due to compensated anemia, but more likely reflects prolonged erythroid maturation. In contrast, *Bcl11a*$^{cko/cko}$ and *Klf1*$^{wt/ko}$*::Bcl11a*$^{cko/cko}$ mice displayed increased Epo levels (Fig. 2D), suggesting that they were under anemic stress. In mice compensatory erythropoiesis (or stress erythropoiesis) exclusively occurs in the spleen. If the increased CD71 expression and

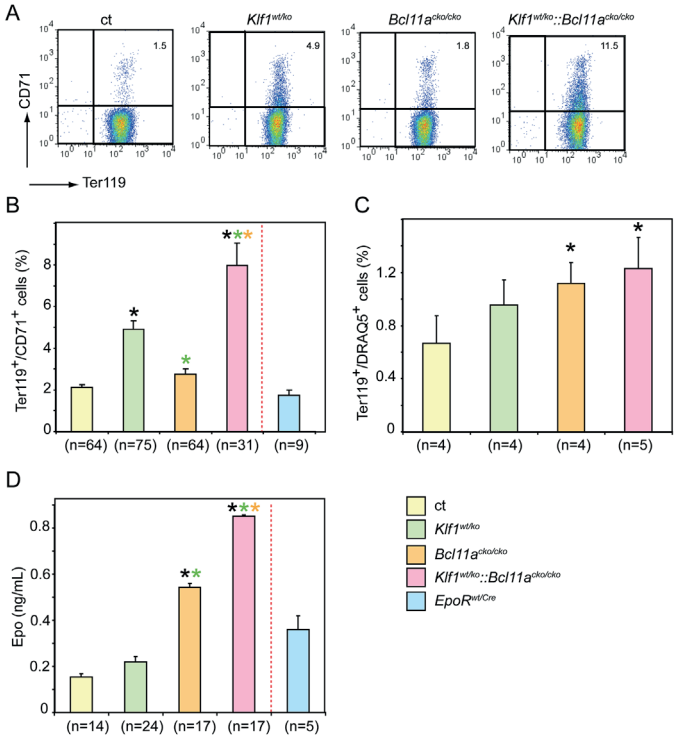


Figure 2. Flowcytometry of peripheral blood and serum erythropoietin levels. (A) Examples of flowcytometry using CD71 and Ter119 staining of peripheral blood cells. (B) Higher percentage of CD71-positive cells in mutant mice with the most pronounced effect in *Klf1*$^{wt/ko}$*::Bcl11a*$^{cko/cko}$ mice. (C) Increased percentage of DRAQ5-positive cells in peripheral blood of mutant mice. D) Compared with the controls, serum erythropoietin levels were significantly higher in *Bcl11a*$^{cko/cko}$ and *Klf1*$^{wt/ko}$*::Bcl11a*$^{cko/cko}$ mice. *: p<0.05 between mutant and control groups; when p<0.05 between mutant groups this is indicated by a color-matched asterisk. Number of mice is depicted for each experiment separately.

Epo levels indicate stress erythropoiesis, we would expect an increased number of spleen erythroblasts. In mutant animals spleen size and cellularity were not significantly different from those observed in the control animals, indicating that stress erythropoiesis had not been fully activated in order to compensate for the minor anemic condition. To assess stress erythropoiesis at the cellular level, we investigated the erythroid compartment in the spleen by flowcytometry using the criteria described by Socolovsky et al [33]. The gating strategy is depicted in Fig. 3A. We observed that the EryA and EryB populations were significantly increased in the *Bcl11a*$^{cko/cko}$ and *Klf1*$^{wt/ko}$*::Bcl11a*$^{cko/cko}$ mice (Fig. 3B,C), while in Klf1$^{wt/ko}$ animals these populations were similar to those observed in the controls. This supports the notion that a low stress erythropoiesis response has been activated in *Bcl11a*$^{cko/cko}$ and *Klf1*$^{wt/ko}$*::Bcl11a*$^{cko/cko}$ mice, but not in *Klf1*$^{wt/ko}$ animals.

To extend these observations, we analyzed embryonic blood and fetal livers at E18.5, just prior to birth. At this stage, erythropoiesis is highly active to supply the demand of the rapidly growing embryo for more oxygen-carrying capacity. Flowcytometry analysis of E18.5 blood revealed no difference in the CD71^{+}Ter119^{+} population in peripheral blood of *Bcl11a*$^{cko/cko}$ embryos (55.0±.4.2 *versus* 56.7±.4.1 in the controls). This percentage was increased in *Klf1*$^{wt/ko}$ blood samples (64.2±.5.8 *versus* 56.7±.4.1 in the controls) and was highest in blood from *Klf1*$^{wt/ko}$*::Bcl11a*$^{cko/cko}$ embryos (75.6±.4.1 *versus* 56.7±.4.1 in the controls) (Fig. 4A,B). Similar results were obtained following FACS analysis of E18.5 fetal liver cells (Fig. 4A,C). Consequently, the percentage of mature CD71^{-}/Ter119^{+} cells in fetal liver and peripheral blood of E18.5 *Klf1*$^{wt/ko}$*::Bcl11a*$^{cko/cko}$ embryos was significantly lower than that observed in *Klf1*$^{wt/ko}$, *Bcl11a*$^{cko/cko}$ and control embryos (Fig. 4B,C).

Collectively, these results suggest that haploinsufficiency for KLF1 prolongs reticulocyte maturation, and that this phenotype is further exacerbated in combination with BCL11A deficiency. Despite this, the impact on erythropoiesis is modest and none of the compound mutant mice suffer from overt anemia even at prenatal stages when the demand for

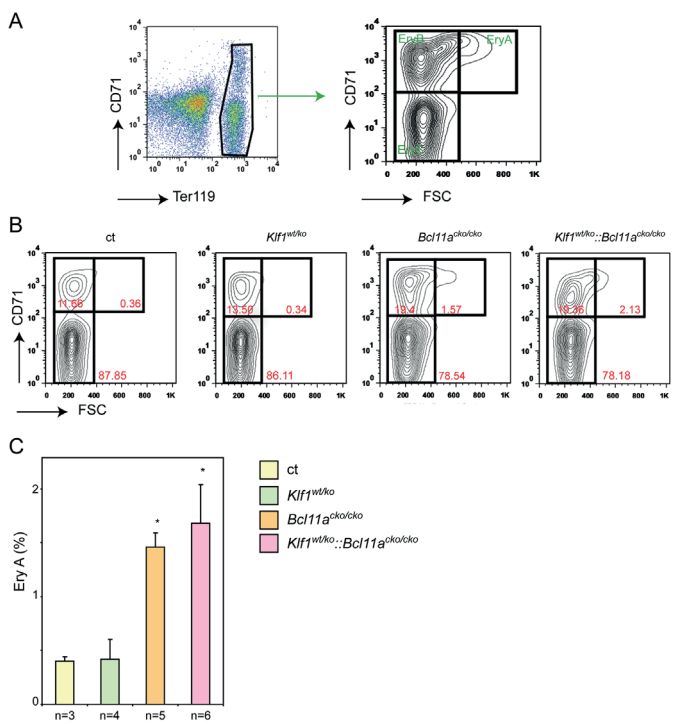


Figure 3. Flowcytometry analysis of stress erythropoiesis in the spleen. (A) Gating strategy of Ter119 and CD71-stained splenocytes to define the EryA, B and C populations[33]. The most immature cells are in the EryA population (CD71^{+}, FSC^{high}). (B) Examples of flowcytometry plots with percentages of EryA, B and C populations indicated (C) EryA populations in the spleens of mice with the four different genotypes. *: p<0.05 between mutant and control groups. (ct n=3, *Klf1*$^{wt/ko}$ n=4, *Bcl11a*$^{cko/cko}$ n=5 and *Klf1*$^{wt/ko}$*::Bcl11a*$^{cko/cko}$ n=6)

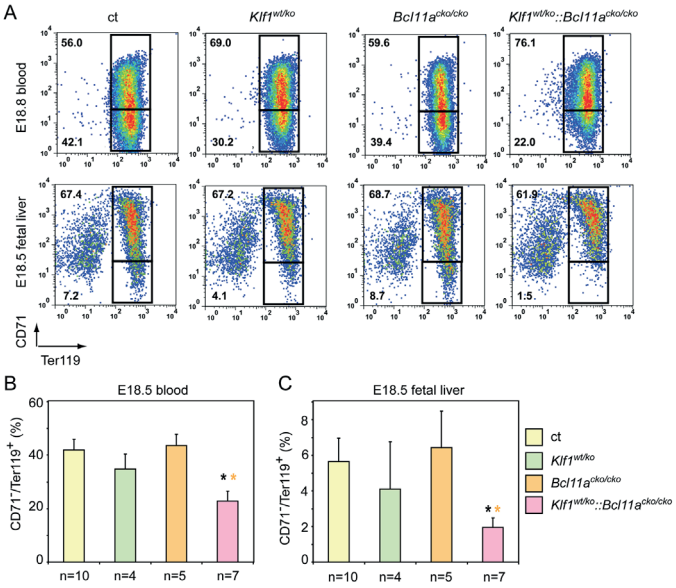


Figure 4. Flowcytometry analysis of blood and fetal liver cells at E18.5. (A) Gating strategy of Ter119 and CD71-stained blood and fetal liver cells of E18.5 embryos. Mature erythrocytes are in the lower quadrant (CD71$^-$Ter119$^+$). (B) The percentage of CD71$^-$Ter119$^+$ cells in E18.5 blood. C) The percentage of CD71$^-$Ter119$^+$ cells in E18.5 fetal livers. *: p<0.05 between mutant and control groups. (ct n=10, *Klf1*$^{wt/ko}$ n=4, *Bcl11a*$^{cko/cko}$ n=5 and *Klf1*$^{wt/ko}$*::Bcl11a*$^{cko/cko}$ n=7).

erythroid expansion is high. Furthermore, induction of anemia by phenylhydrazine treatment demonstrated that animals of all four genotypes were able to mount a normal stress response, as judged by the increase in spleen weight and the fraction of spleen cells in the EryA population (Suppl. Fig. 2A,B). Finally, we note that deformability of adult erythrocytes from mutant animals was similar to that observed in the controls, consistent with the normal morphology of the erythrocytes on cytospins and indicating the absence of structural red cell abnormalities (data not shown).

Regulation of γ-globin expression during fetal development

Previously, it has been demonstrated that KLF1 preferentially activates the adult β-globin gene[16,17]. Using mice carrying a human β-globin locus transgene, it was shown that switching from γ- to β-globin expression was delayed in a *Klf1*$^{wt/ko}$ background[19]. In addition, it was recently found that KLF1 is a direct activator of BCL11A expression[13,14]. To assess the impact of the KLF1-BCL11A axis on expression of the β-like globin genes, we determined globin expression at different developmental stages in erythroid cells derived from mice of all four genotypes. Since γ-globin expression is silenced around E13.5 in PAC8.1 mice[7], we assessed globin expression at E14.5 and at E18.5, the latest possible time before birth. We used total RNA extracted from fetal liver samples for globin expression analysis. The expression of γ- and β-globin mRNA was determined directly with S1 nuclease protection assays [13,27], and the γ/(γ+ β) ratios were calculated from data obtained with QRT-PCR assays[13]. Compared to *Bcl11a*$^{cko/cko}$ embryos, which express high levels of γ-globin[8], *Klf1*$^{wt/ko}$ embryos displayed lower levels of γ-globin expression but these levels were still significantly higher than those observed in the control samples (Fig. 5A,B), consistent with previous results[19]. Notably, the highest γ-globin mRNA expression was observed in *Klf1*$^{wt/ko}$*::Bcl11a*$^{cko/cko}$ embryos. In addition, western blot analysis revealed higher levels of γ-globin protein in E18.5 blood samples from *Klf1*$^{wt/ko}$*::Bcl11a*$^{cko/cko}$ embryos compared to those observed in *Bcl11a*$^{cko/cko}$ embryos (Fig. 5C), while for both genotypes these levels were much higher than in the control samples. Consistent with previously reported data[19], γ-globin expression was low in *Klf1*$^{wt/ko}$ E18.5 erythroid cells, although still significantly higher than in the control

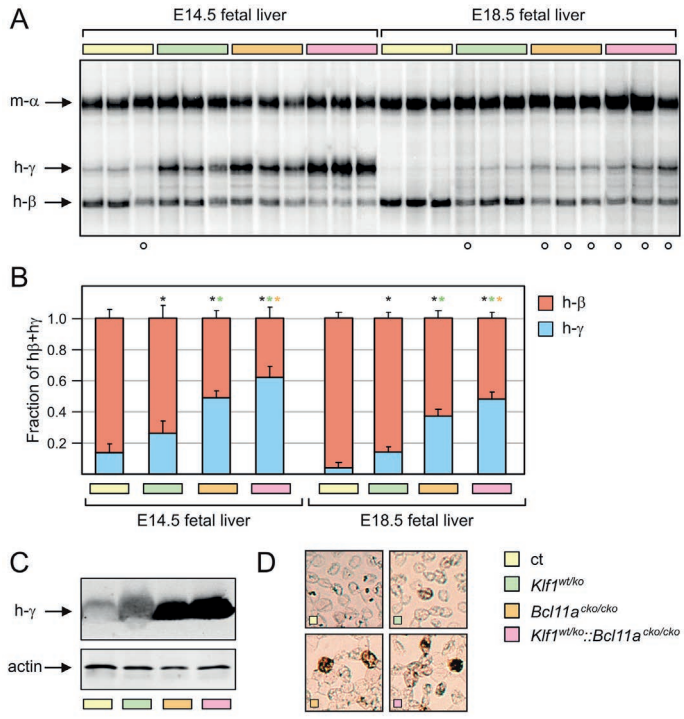


Figure 5. Analysis of γ-globin expression at E14.5 and E18.5. (A) S1 nuclease protection assay on E14.5 and E18.5 fetal liver RNA, detecting human γ- and β-globin mRNA. Mouse α-globin was used as a loading control. (B) QRT-PCR analysis of human globin expression. γ/ γ+β ratios are significantly higher in all mutant mice compared to the controls (n = 3-5 for each genotype and each time point). (C) Western blot on E18.5 blood samples detecting expression of γ-globin protein. Actin was used as a loading control. (D) Immunohistochemistry of γ-globin expression (brown) in E18.5 blood. The white circles in (A) denote embryos heterozygous for the human β-globin locus transgene. *: p<0.05 between mutant and control groups; when p<0.05 between mutant groups this is indicated by a color-matched asterisk.

samples. Finally, to determine whether the increased γ-globin expression had a pancellular or heterocellular distribution, we performed immunohistochemistry on cytospins of E18.5 blood. The number of cells staining positive for γ-globin expression correlated well with the γ-globin levels determined by the S1 nuclease protection- and QRT-PCR assays (Fig. 5D). This is consistent with a pancellular distribution of γ-globin. Collectively these data emphasize the dominant role of BCL11A in γ-globin silencing during prenatal development in mice[8]. Furthermore, the observation that γ-globin levels are highest in *Klf1*$^{wt/ko}$*::Bcl11a*$^{cko/cko}$ embryos lends support to the proposed role of the KLF1-BCL11A axis in this process[13,29].
Next, we investigated γ-globin expression post-natally in young (4-5 weeks) and adult mice (>10 weeks) by S1 nuclease protection- and QRT-PCR analysis of bone marrow RNA. Although γ-globin expression decreased after birth, it remained significantly higher in young and adult *Klf1*$^{wt/ko}$*::Bcl11a*$^{cko/cko}$ mice compared to control mice (Fig. 6A,B). In adult mutant *Klf1*$^{wt/ko}$, *Bcl11a*$^{cko/cko}$ and *Klf1*$^{wt/ko}$*::Bcl11a*$^{cko/cko}$ mice, the levels of γ-globin increased in response to phenylhydrazine-induced anemia, while such an increase was not observed in the control animals (Suppl. Fig. 2C). Collectively, the decline in γ-globin expression after birth suggests that additional silencing mechanisms exist that prevent high level γ-globin expression in adult *Klf1*$^{wt/ko}$*::Bcl11a*$^{cko/cko}$ mice.

Dynamics of mouse globin expression

The decline in γ-globin expression in adult *Klf1*$^{wt/ko}$*::Bcl11a*$^{cko/cko}$ mice raised the question whether the endogenous mouse embryonic globin genes would display a similar expression pattern. We therefore investigated the expression of mouse ζ- and εγ-globins at different developmental stages by QRT-PCR. Compared to the controls the expression of embryonic ζ- and εγ-globin was significantly higher in E14.5 and E18.5 fetal liver of *Bcl11a*$^{cko/cko}$

embryos, in agreement with previous reports [8]. In addition, we observed similarly increased expression in *Klf1*$^{wt/ko}$*::Bcl11a*$^{cko/cko}$ fetal livers (Fig. 7). Akin to the observations on γ-globin expression, we found that the absolute amount of expression of ζ- and εγ-globin declined after birth in *Bcl11a*$^{cko/cko}$ and *Klf1*$^{wt/ko}$*::Bcl11a*$^{cko/cko}$ mice (Fig. 7A,C; note logarithmic scales). In relative terms, the fold-change in ζ- and εγ-globin expression was most pronounced in *Bcl11a*$^{cko/cko}$ and *Klf1*$^{wt/ko}$*::Bcl11a*$^{cko/cko}$ mice, compared to the age-matched control mice (Fig. 7B,D; note logarithmic scales). In adult animals, this reached 150 to 200-fold increase for ζ-globin, and 30-50-fold increase for εγ-globin. For εγ-globin the fold-change tended to be highest in *Klf1*$^{wt/ko}$*::Bcl11a*$^{cko/cko}$ mice, but this was statistically significant only in 4-5 week old animals (Fig. 7C,D). In addition, we found that for *Klf1*$^{wt/ko}$ animals expression of εγ-globin was significantly increased in >10-week old animals (Fig. 7C,D), while this genotype did not have a positive effect on ζ-globin expression at any stage of development analyzed

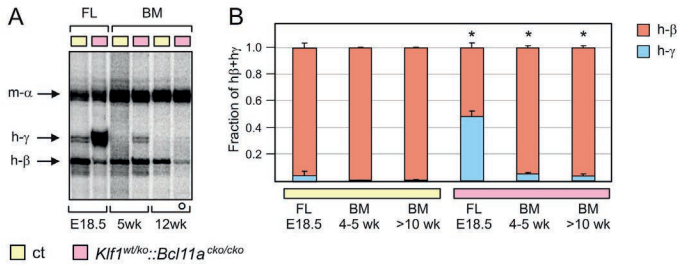


Figure 6. Expression of γ-globin in post-natal compound mutant mice. (A) S1 nuclease protection assay on bone marrow RNA from 5-week and 12-week old mice, detecting human γ- and β-globin mRNA. Mouse α-globin was used as a loading control. For comparison, fetal liver RNA from E18.5 embryos was used. Note that the probe detecting γ-globin had a 10-fold higher specific activity than the one use in Fig. 5. (B) qRT-PCR analysis of human globin expression (ct: n=3 and *Klf1*$^{wt/ko}$*::Bcl11a*$^{cko/cko}$: n=4 for each time point). The white circle in (A) denotes an animal heterozygous for the human β-globin locus transgene. *: p<0.05 between age-matched mutant and control groups.

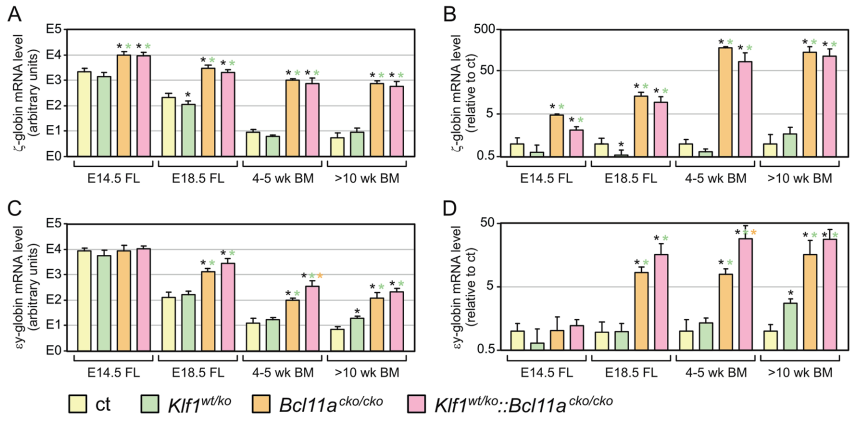


Figure 7. Expression of embryonic ζ- and εγ-globins in compound mutant mice. (A) qRT-PCR analysis of mouse ζ- and εγ-globin, absolute expression levels calculated relative to expression of mouse α-globin. (B) Expression levels relative to those observed in control mice of the same age group; expression of α-globin was used for normalization. N=3-5 for each genotype and time point. Note logarithmic scales. *: p<0.05 between mutant and control groups; when p<0.05 between mutant groups this is indicated by a color-matched asterisk.

(Fig. 7A,B). Finally, we note that the expression pattern of the human ε-globin gene is very similar to that of the mouse εy-globin gene (Suppl. Fig. 3).
Collectively, the data on expression of the endogenous mouse embryonic globin genes are congruent with the observations on expression of the γ-globin genes residing in the human β-globin locus PAC8.1 transgene. BCL11A has a dominant role in silencing of the embryonic/fetal globin genes during mouse ontogeny, which is further augmented by the preferential activation of β-globin expression by KLF1. The embryonic/fetal mouse globin genes remain expressed in adult *Bcl11a*$^{cko/cko}$ and *Klf1*$^{wt/ko}$*::Bcl11a*$^{cko/cko}$ animals, but expression declines to quantitatively low, physiologically irrelevant, levels. This shows that additional silencing mechanisms are operational in adult mouse erythropoiesis.

DISCUSSION

In this report we focused on the role of two transcription factors, KLF1 and BCL11A, in erythropoiesis and developmental regulation of globin expression. KLF1, an erythroid-specific protein, directly activates β-globin expression through binding to CACCC box sequences in the LCR and the β-globin promoter [12]. While high level expression of β-globin requires the presence of KLF1, the other globins, including γ-globin genes contained on human β-globin locus transgenes, are still highly expressed in the absence of KLF1[18,19]. In addition to a critical activator of β-globin expression, genome-wide gene expression analyses have revealed that KLF1 is a master regulator of genes activated during terminal erythroid differentiation[34-36]. These target genes include heme synthesis enzymes, cell cycle regulators and erythroid membrane and cytoskeleton proteins. It is therefore not surprising that *Klf1 null* embryos succumb to severe anemia. Embryonic lethality is not rescued by expression of a γ-globin transgene[37]; this corrects globin chain imbalance but not hemolysis caused by deregulated expression of essential erythroid membrane and cytoskeleton proteins.
In humans, *KLF1* mutations result in a spectrum of erythroid phenotypes[38]. Haploinsufficiency for KLF1 is the underlying cause of the In(Lu) blood group[20], is associated with red cell zinc protoporphyria[21], increased HbA2[22], and HPFH[13]. A dominant KLF1 missense mutation p.E325K also affects DNA binding and was reported to cause congenital dyserythropoietic anemia (CDA)[23]. The two CDA patients with the p.E325K mutation displayed HbF levels of 31.6% and 44%. Remarkably, embryonic ζ- and ε-globin were also increased to very high levels[23]. The ethylnitrosurea-induced mouse mutant *Nan* (neonatal anemia) carries a missense mutation, p.E339D, in the homologous position of mouse KLF1[39,40]. This mutation causes a dominant hemolytic anemia, with markedly increased expression of embryonic globins in adult *Nan* animals[40]. Collectively, these data support a model in which KLF1 activates β-globin expression and suppresses the embryonic/fetal β-like globin genes. Recently, it has been shown that expression of BCL11A is regulated by KLF1, suggesting an intricate mechanism for the developmental regulation of the β-like globin genes coordinately mediated by KLF1 and BCL11A[13,14]. Ablation of BCL11A in mice demonstrated that it is the major repressor of embryonic/fetal β-like globin genes during ontogeny[8]. Interestingly, the timing of expression of full-length BCL11A differs between mouse and human[8], providing an explanation for the observation that in mice carrying a human β-globin locus transgene γ-globin silencing is already completed at the fetal liver stage[41].
In this paper, we have investigated the interplay of KLF1 and BCL11A in erythropoiesis and developmental regulation of globin gene expression. We find that BCL11A is the dominant repressor of embryonic/fetal globin genes in the mouse embryo from midgestation to term. While the endogenous embryonic globin genes are normally silenced at the onset of definitive erythropoiesis in the fetal liver, their expression is markedly increased in the absence of BCL11A. Quantitatively, their expression is still considerable at E18.5, just prior to birth. These data are in agreement with those reported on mouse embryos with a systemic

BCL11A *null* mutation[8], and further establish that this phenotype is intrinsic to the erythroid lineage[28]. In combination with KLF1 haploinsufficiency, expression of the embryonic globin genes was also increased. However, expression of εγ-globin declined to quantitatively low levels, suggesting that preferential activation of β-globin expression by KLF1 still occurs in adult *Klf1*$^{wt/ko}$*::Bcl11a*$^{cko/cko}$ adults and indicating that additional adult-stage silencing mechanisms exist. To achieve efficient silencing of the embryonic globin genes in fetal liver erythropoiesis, an intact KLF1-BCL11A axis is required: KLF1 activates BCL11A expression[13,14]; BCL11A represses the embryonic globin genes [8] thereby unleashing the full potential of KLF1 to activate β-globin expression. It is interesting to note that BCL11A also represses the embryonic ζ-globin gene, while KLF1 is not essential for high levels of α-globin expression. This provides another example of the contrasting mechanisms regulating the α-like and β-like globin loci[42].

In mice carrying a human β-globin locus transgene, the switch from γ- to β-globin expression occurs at the early fetal liver stage[41]. This is dependent on the presence of BCL11A[8]; and in our experiments with erythroid-specific ablation of BCL11A we obtained similar results. In agreement with previous reports using a different human β-globin locus transgene (line 72)[19], we find that haploinsufficiency for KLF1 delays γ- to β-globin switching leading to a ~2-fold increase in the γ/(γ+β) ratio at E14.5 and E18.5. Part of this increase can be explained by diminished BCL11A expression in embryos with KLF1 insufficiency[14]. Remarkably, relative to *Bcl11a*$^{cko/cko}$ embryos, there is a further increase in the γ/(γ+β) ratio in compound *Klf1*$^{wt/ko}$*::Bcl11a*$^{cko/cko}$ embryos, e.g. from 0.37 to 0.48 at E18.5. This shows that even in the absence of BCL11A, KLF1 still preferentially activates the β-globin gene. Collectively, these data support the proposed role of the KLF1-BCL11A axis in γ-globin regulation[13,14,38].

GWAS has shown an association between *BCL11A* SNPs and erythroid parameters[43], suggesting a role for BCL11A in adult erythropoiesis beyond globin regulation. We therefore determined the effects on steady-state erythropoiesis and globin expression in adult animals. We found only minor deviations in the hematological parameters of *Klf1*$^{wt/ko}$, *Bc11a*$^{cko/cko}$ and *Klf1*$^{wt/ko}$*::Bcl11a*$^{cko/cko}$ animals. *Klf1*$^{wt/ko}$ animals display mild reticulocytosis but do not have increased Epo levels, indicating that the production of erythroid cells is adequate but that the maturation of reticulocytes takes more time, or alternatively that the reticulocytes are prematurely released in the circulation. Mild reticulocytosis is also one of the hallmarks of individuals from a Maltese pedigree with KLF1 haploinsufficiency[13]. *Bcl11a*$^{cko/cko}$ animals are displaying lower HCT, HGB and RBC counts, and slightly increased Epo levels. In the *Klf1*$^{wt/ko}$*::Bcl11a*$^{cko/cko}$ animals these traits are exacerbated and in addition increased MCV, MCH and MCHC reached statistical significance. The most straightforward explanation for these observations is that the *Bcl11a*$^{cko/cko}$ and *Klf1*$^{wt/ko}$*::Bcl11a*$^{cko/cko}$ genotypes provoke a mild anemia, which is compensated by increased Epo levels. KLF1 has a well established role in the expression of erythroid membrane and cytoskeleton proteins [12] while a putative role for BCL11A in erythroid maturation remains to be further investigated. Importantly, *Bc11a*$^{cko/cko}$ and *Klf1*$^{wt/ko}$*::Bcl11a*$^{cko/cko}$ animals were obtained at the expected Mendelian ratios and thus far none of them developed leukemia (n = 74). This is in contrast to animals reconstituted with BCL11A *null* hematopoietic cells which succumb to T cell leukemia at a high frequency[29]. We conclude that erythroid-specific ablation of BCL11A, alone or in combination with KLF1 haploinsufficiency, only mildly affects steady-state erythropoiesis in adult mice. This supports proposals for the modulation of the KLF1-BCL11A axis in β-thalassemia and SCA patients for reactivation of γ-globin expression[13,14,28].

At first glance, the observation that expression of γ-globin declines in adult *Bcl11a*$^{cko/cko}$ and *Klf1*$^{wt/ko}$*::Bcl11a*$^{cko/cko}$ animals is counterintuitive to this notion. However, there is evidence for species-specific differences in developmental regulation of globin expression. In mice, full-length BCL11A is already expressed in the fetal liver, while human fetal liver cells only express short isoforms of BCL11A[8]. The full-length isoform is first observed in human

adult erythroid progenitors. This difference in developmental timing of BCL11A expression provides a molecular explanation for the observation that γ- to β-globin switching occurs at the fetal liver stage in mouse embryos carrying a human β-globin locus transgene[8,41]. Furthermore, treatment with drugs that raise HbF cells in human subjects fail to do so in human β-globin locus transgenic mice[44,45], suggesting that silencing of the γ-globin genes is much tighter in mice than it is in humans. We therefore strongly feel that this intrinsic property of the mouse model does not provide a convincing argument against proposals to target the KLF1-BCL11A axis as a therapeutic approach to increase HbF levels in β-thalassemia and SCA patients[8,11,13,14]. This notion is confirmed by the recent elegant demonstration that erythroid-specific ablation of BCL11A corrects disease in a mouse model for SCA[28].
We suggest that the tight repression of the γ-globin genes in mice provides a window of opportunity for identification of additional factors involved in the silencing mechanism at the adult stage. Enforcement of repression of the embryonic/fetal program in adult erythropoiesis may be executed by, for instance, the transcription factors MYB[46] and SOX6[47], the chromatin-bound FOP/CHTOP protein[27] and NuRD complex[48], the orphan nuclear receptors TR2/TR4[49] and the protein arginine methyl transferase PRMT5[50], and is likely to include additional epigenetic mechanisms such as PcG complex recruitment and DNA methylation. Future work will be aimed at further elucidating the multi-layered repressive network of the embryonic/fetal program in the adult erythroid environment.

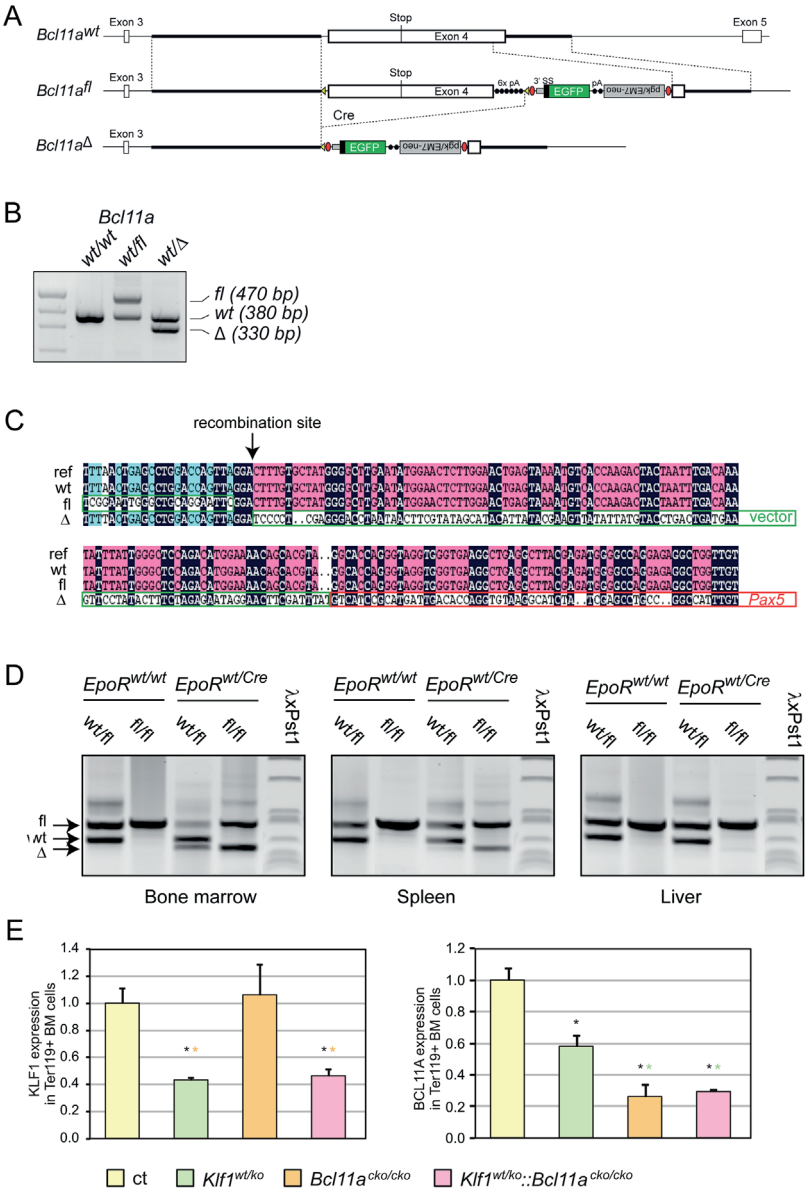
Acknowledgements
We would like to acknowledge Sylvia Dekker, Ralph Stadhouders, Rien van Haperen and Divine Kulu for technical assistance, and the animal facility of Erasmus MC for animal husbandry. This work was supported by the Netherlands Genomics Initiative (NGI). Research of the Busslinger group was supported by Boehringer Ingelheim and an Advanced Grant of the European Research Council.

Authorship and disclosures
Contribution: SP was the principle investigator and takes primary responsibility for the paper. FE, NG, IB, EvdA, IC and KvL performed experiments. TvG performed the statistical analysis. FE, IB, UK, MvL. FG, TBvD, MB, SP designed experiments. The paper was written by FE, EvdA, MB and SP.
The authors declare no potential conflicts of interest.

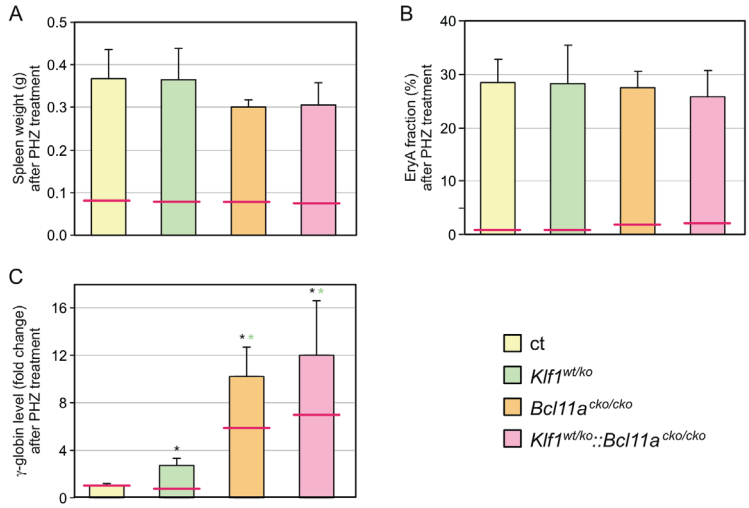
REFERENCES

1. Angastiniotis, M. & Modell, B. Global epidemiology of hemoglobin disorders. *Ann N Y Acad Sci* **850**, 251-69 (1998).
2. Stamatoyannopoulos, G. & Grosveld, F. The Molecular Basis of Blood Diseases. (eds. Stamatoyannopoulos, G., Majerus, P.W., Perlmutter, R.M. & Varmus, H.) 135-182 (WB Saunders Company, Philadelphia PA, 2001).
3. Gambari, R. Foetal haemoglobin inducers and thalassaemia: novel achievements. *Blood Transfus* **8**, 5-7 (2010).
4. Steinberg, M.H. Therapies to increase fetal hemoglobin in sickle cell disease. *Curr Hematol Rep* **2**, 95-101 (2003).
5. Higgs, D.R. & Gibbons, R.J. The molecular basis of alpha-thalassemia: a model for understanding human molecular genetics. *Hematol Oncol Clin North Am* **24**, 1033-54 (2010).
6. Palis, J. Ontogeny of erythropoiesis. *Curr Opin Hematol* **15**, 155-61 (2008).
7. de Krom, M., van de Corput, M., von Lindern, M., Grosveld, F. & Strouboulis, J. Stochastic patterns in globin gene expression are established prior to transcriptional activation and are clonally inherited. *Mol Cell* **9**, 1319-26 (2002).
8. Sankaran, V.G. et al. Developmental and species-divergent globin switching are driven by BCL11A. *Nature* **460**, 1093-7 (2009).
9. Menzel, S. et al. A QTL influencing F cell production maps to a gene encoding a zinc-finger protein on chromosome 2p15. *Nat Genet* **39**, 1197-9 (2007).
10. Uda, M. et al. 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**, 1620-5 (2008).
11. Sankaran, V.G. et al. Human fetal hemoglobin expression is regulated by the developmental stage-specific repressor BCL11A. *Science* **322**, 1839-42 (2008).
12. Siatecka, M. & Bieker, J.J. The multifunctional role of EKLF/KLF1 during erythropoiesis. *Blood* (2011).
13. Borg, J. et al. Haploinsufficiency for the erythroid transcription factor KLF1 causes hereditary persistence of fetal hemoglobin. *Nat Genet* **42**, 801-5 (2010).
14. Zhou, D., Liu, K., Sun, C.W., Pawlik, K.M. & Townes, T.M. KLF1 regulates BCL11A expression and gamma- to beta-globin gene switching. *Nat Genet* **42**, 742-4 (2010).
15. Donze, D., Townes, T.M. & Bieker, J.J. Role of erythroid Kruppel-like factor in human gamma- to beta-globin gene switching. *J Biol Chem* **270**, 1955-9 (1995).
16. Nuez, B., Michalovich, D., Bygrave, A., Ploemacher, R. & Grosveld, F. Defective haematopoiesis in fetal liver resulting from inactivation of the EKLF gene. *Nature* **375**, 316-8 (1995).
17. Perkins, A.C., Sharpe, A.H. & Orkin, S.H. Lethal beta-thalassaemia in mice lacking the erythroid CACCC-transcription factor EKLF. *Nature* **375**, 318-22 (1995).
18. Perkins, A.C., Gaensler, K.M. & Orkin, S.H. Silencing of human fetal globin expression is impaired in the absence of the adult beta-globin gene activator protein EKLF. *Proc Natl Acad Sci U S A* **93**, 12267-71 (1996).
19. Wijgerde, M. et al. The role of EKLF in human beta-globin gene competition. *Genes Dev* **10**, 2894-902 (1996).
20. Singleton, B.K., Burton, N.M., Green, C., Brady, R.L. & Anstee, D.J. Mutations in EKLF/KLF1 form the molecular basis of the rare blood group In(Lu) phenotype. *Blood* **112**, 2081-8 (2008).
21. Satta, S. et al. Compound heterozygosity for KLF1 mutations associated with remarkable increase of fetal hemoglobin and red cell protoporphyrin. *Haematologica* (2011).
22. Perseu, L. et al. KLF1 gene mutations cause borderline HbA2. *Blood* (2011).
23. Arnaud, L. et al. A dominant mutation in the gene encoding the erythroid transcription factor KLF1 causes a congenital dyserythropoietic anemia. *Am J Hum Genet* **87**, 721-7 (2010).
24. Heinrich, A.C., Pelanda, R. & Klingmuller, U. A mouse model for visualization and conditional mutations in the erythroid lineage. *Blood* **104**, 659-66 (2004).
25. Beug, H., Palmieri, S., Freudenstein, C., Zentgraf, H. & Graf, T. Hormone-dependent terminal differentiation in vitro of chicken erythroleukemia cells transformed by ts mutants of avian erythroblastosis virus. *Cell* **28**, 907-19 (1982).
26. Esteghamat, F. et al. The DNA binding factor Hmg20b is a repressor of erythroid differentiation. *Haematologica* (2011).

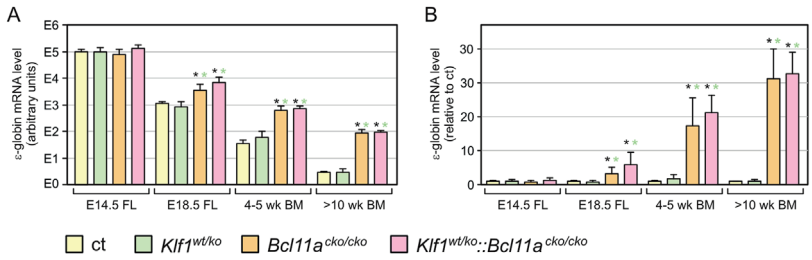
27. van Dijk, T.B. et al. Fetal globin expression is regulated by Friend of Prmt1. *Blood* **116**, 4349-52 (2010).
28. Xu, J. et al. Correction of Sickle Cell Disease in Adult Mice by Interference with Fetal Hemoglobin Silencing. *Science* (2011).
29. Liu, P. et al. Bcl11a is essential for normal lymphoid development. *Nat Immunol* **4**, 525-32 (2003).
30. Piva, E., Brugnara, C., Chiandetti, L. & Plebani, M. Automated reticulocyte counting: state of the art and clinical applications in the evaluation of erythropoiesis. *Clin Chem Lab Med* **48**, 1369-80 (2010).
31. Ponka, P. & Lok, C.N. The transferrin receptor: role in health and disease. *Int J Biochem Cell Biol* **31**, 1111-37 (1999).
32. Antonchuk, J., Hyland, C.D., Hilton, D.J. & Alexander, W.S. Synergistic effects on erythropoiesis, thrombopoiesis, and stem cell competitiveness in mice deficient in thrombopoietin and steel factor receptors. *Blood* **104**, 1306-13 (2004).
33. Socolovsky, M. Molecular insights into stress erythropoiesis. *Curr Opin Hematol* **14**, 215-24 (2007).
34. Drissen, R. et al. The erythroid phenotype of EKLF-null mice: defects in hemoglobin metabolism and membrane stability. *Mol Cell Biol* **25**, 5205-14 (2005).
35. Hodge, D. et al. A global role for EKLF in definitive and primitive erythropoiesis. *Blood* **107**, 3359-70 (2006).
36. Pilon, A.M. et al. Failure of terminal erythroid differentiation in EKLF-deficient mice is associated with cell cycle perturbation and reduced expression of E2F2. *Mol Cell Biol* **28**, 7394-401 (2008).
37. Perkins, A.C., Peterson, K.R., Stamatoyannopoulos, G., Witkowska, H.E. & Orkin, S.H. Fetal expression of a human Agamma globin transgene rescues globin chain imbalance but not hemolysis in EKLF null mouse embryos. *Blood* **95**, 1827-33 (2000).
38. Borg, J., Patrinos, G.P., Felice, A.E. & Philipsen, S. Erythroid phenotypes associated with KLF1 mutations. *Haematologica* **96**, 635-8 (2011).
39. Heruth, D.P. et al. Mutation in erythroid specific transcription factor KLF1 causes Hereditary Spherocytosis in the Nan hemolytic anemia mouse model. *Genomics* **96**, 303-7 (2010).
40. Siatecka, M. et al. Severe anemia in the Nan mutant mouse caused by sequence-selective disruption of erythroid Kruppel-like factor. *Proc Natl Acad Sci U S A* **107**, 15151-6 (2010).
41. Strouboulis, J., Dillon, N. & Grosveld, F. Developmental regulation of a complete 70-kb human beta-globin locus in transgenic mice. *Genes Dev* **6**, 1857-64 (1992).
42. Garrick, D. et al. The role of the polycomb complex in silencing alpha-globin gene expression in nonerythroid cells. *Blood* **112**, 3889-99 (2008).
43. Ganesh, S.K. et al. Multiple loci influence erythrocyte phenotypes in the CHARGE Consortium. *Nat Genet* **41**, 1191-8 (2009).
44. Buller, A.M., Elford, H.L., DuBois, C.C., Meyer, J. & Lloyd, J.A. A combination of hydroxyurea and isobutyramide to induce fetal hemoglobin in transgenic mice is more hematotoxic than the individual agents. *Blood Cells Mol Dis* **25**, 255-69 (1999).
45. Pace, B., Li, Q., Peterson, K. & Stamatoyannopoulos, G. alpha-Amino butyric acid cannot reactivate the silenced gamma gene of the beta locus YAC transgenic mouse. *Blood* **84**, 4344-53 (1994).
46. Thein, S.L. et al. 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**, 11346-51 (2007).
47. Xu, J. et al. Transcriptional silencing of gamma-globin by BCL11A involves long-range interactions and cooperation with SOX6. *Genes Dev* **24**, 783-98 (2010).
48. Gnanapragasam, M.N. et al. p66Alpha-MBD2 coiled-coil interaction and recruitment of Mi-2 are critical for globin gene silencing by the MBD2-NuRD complex. *Proc Natl Acad Sci U S A* **108**, 7487-92 (2011).
49. Cui, S. et al. Nuclear Receptors TR2 and TR4 Recruit Multiple Epigenetic Transcriptional Co-repressors that Associate Specifically with the Embryonic beta-type Globin Promoters in Differentiated Adult Erythroid Cells. *Mol Cell Biol* (2011).
50. Rank, G. et al. Identification of a PRMT5-dependent repressor complex linked to silencing of human fetal globin gene expression. *Blood* **116**, 1585-92 (2010).



Supplementary Figure 1. Deletion of conditional Bcl11a allele by EpoRCre. (A) The conditional knockout allele of the *Bcl11a* gene. (B) PCR strategy to detect the wildtype (wt), floxed (fl) and deleted (Δ) alleles. (C) Sequencing of the PCR products confirms the deletion expected after Cre-mediated recombination. (D) Recombination of the floxed *Bcl11a* allele in erythroid tissues mediated by *EpoR*^{Cre}. E) Expression of KLF1 and BCL11A in Ter119 MACS-selected bone marrow (BM) cells from adult mice (population purity ~85%). Expression of RNAse inhibitor and β-actin was used to normalize the data; levels in control mice (ct) were set to 1. For each genotype n=3. Black asterisks: p<0.05 compared to ct samples; green asterisks p<0.05 compared to *Klf1*$^{wt/ko}$ samples, orange asterisks p<0.05 compared to *Bcl1a*$^{cko/cko}$ samples.



Supplementary Figure 2. Response to phenylhydrazine-induced anemia. (A) Spleen weights at Day 5 after phenylhydrazine (PHZ) treatment. (B) EryA fractions of spleen cells at Day 5 after phenylhydrazine (PHZ) treatment. (C) Expression of human γ-globin determined by qRT-PCR in spleen cells at Day 5 after PHZ treatment. Fold-change was calculated relative to the expression levels in the control (ct) animals before PHZ treatment. Black asterisks: p<0.05 compared to ct samples; green asterisks p<0.05 compared to *Klf1*$^{wt/ko}$ samples. The red bars indicate values before PHZ treatment. For (A) and (B), ct n=3, *Klf1*$^{wt/ko}$ n=4, *Bcl11a*$^{cko/cko}$ n=6 and *Klf1*$^{wt/ko:}$*:Bcl11a*$^{cko/cko}$ n=3); for (C) n=3 for each genotype.



Supplementary Figure 3. Expression of human ε-globin. (A) qRT-PCR analysis of human ε-globin, absolute expression levels calculated relative to expression of mouse α-globin. (B) Expression levels relative to those observed in control mice of the same age group; expression of α-globin was used for normalization. N=3-5 for each genotype and time point. Note logarithmic scales in (A). *: p<0.05 between mutant and control groups. Black asterisks: p<0.05 compared to ct samples; green asterisks p<0.05 compared to *Klf1*$^{wt/ko}$ samples.

Supplementary Table 1. Oligonucleotide sequences

Name	Sequence	Purpose
Mouse α-globin S	TTGGCTAGCCACCACCCT	QRT-PCR
Mouse α-globin A	CCAAGAGGTACAGGTGCA	QRT-PCR
Mouse β_{major}-globin S	TTAAGGCTCCTGGGCAATAT	QRT-PCR
Mouse β_{major}-globin A	TGCCAACAACTGACAGATGC	QRT-PCR
Mouse ε_{y}-globin S	GTTTTGGCTAGTCACTTCGG	QRT-PCR
Mouse ε-globin A	CAAGGAACAGCTCAGTATTC	QRT-PCR
Mouse ζ-globin S	GAAGCCTGGGACAAGTTCAT	QRT-PCR
Mouse ζ-globin A	GGGTTCAATAAAGGGGAGGA	QRT-PCR
Mouse βh1-globin S	TTGCCAAGGAATTCACCCCA	QRT-PCR
Mouse βh1-globin A	CTCAATGCAGTCCCCATGGA	QRT-PCR
Human β-globin S	CTGCCTATCAGAAAGTGGTG	QRT-PCR
Human β-globin A	ATTGGACAGCAAGAAAGCGA	QRT-PCR
Human γ-globin S	AGGTGCTGACTTCCTTGGG	QRT-PCR
Human γ-globin A	GGGTGAATTCTTGCCGAA	QRT-PCR
Human ε-globin S	TGGAGATGCTATTAAAAACATGGAC	QRT-PCR
Human ε-globin A	AGAATAATCACCATCACGTTACCC	QRT-PCR
Mouse KLF1 S	CAGCTGAGACTGTCTTACCC	QRT-PCR
Mouse KLF1 A	AATCCTGCGTCTCCTCAGAC	QRT-PCR
Mouse BCL11A S	CCGGGATGAGTGCAGAATAT	QRT-PCR
Mouse BCL11A A	ATGAGTGTTCTGTGCGTGTTG	QRT-PCR
Mouse RNase-inh S	TCCAGTGTGAGCAGCTGAG	QRT-PCR
Mouse RNase-inh A	TGCAGGCACTGACTGAAGCACCA	QRT-PCR
Mouse β-actin S	GATTACTGCTCTGGCTCCT	QRT-PCR
Mouse β-actin A	TGGAAGGTGGACAGTGAG	QRT-PCR
Bcl11a common	AGTGGCACAGAGCTGAAATGAC	Genotyping
Bcl11a Flox	AAGTTGTACATGTGTAGCTG	Genotyping
Bcl11a GFP	AAGCAAGGGGAAGGGTGGTTAGAA	Genotyping

Supplementary Table 2. Normal Mendelian distributions of compound mutant mice

Genotype	Expected (%)	Observed (%)	p-value
ct	18.5 (25)	21 (28.4)	n.s.
Klf1$^{wt/ko}$	18.5 (25)	20 (27.0)	n.s.
Bcl11a$^{cko/cko}$	18.5 (25)	18 (24.3)	n.s.
Klf1$^{wt/ko}$*::Bcl11a*$^{cko/cko}$	18.5 (25)	15 (20.2)	n.s.

n.s. = not significant.

Chapter 3

The KLF1 *Nan* mutation deregulates erythroid maturation during fetal development

Ileana Cantú[1], Harmen J.G. van de Werken[1,2,3], Nynke Gillemans[1], Ralph Stadhouders[1], Steven Heshusius[4], Zelia Ozgur[5], Wilfred F. van Ijcken[5], Marieke von Lindern[4], Frank Grosveld[1,6], Sjaak Philipsen[1], Thamar B. van Dijk[1*].

1 Department of Cell Biology, Erasmus MC, Rotterdam, The Netherlands
2 Department of Urology, Erasmus MC, Rotterdam 3015 CN, The Netherlands.
3 Cancer Computational Biology Center, Erasmus MC, Rotterdam 3015 CN, The Netherlands.
4 Sanquin Research Department of Hematopoiesis, Amsterdam, The Netherlands
5 Center for Biomics, Erasmus MC, Rotterdam, The Netherlands
6 Cancer genomics Center, Erasmus MC, Rotterdam, The Netherlands
* Corresponding author

Manuscript submitted

ABSTRACT

Krüppel-like factor 1 (KLF1) is an essential transcription factor for erythroid development, as demonstrated by *Klf1* knockout mice which die around E14 due to severe anemia. Recently, several KLF1 mutations have been described, causing different erythroid phenotypes. The KLF1 *Nan* mutation, a single amino acid substitution (E339D), causes hemolytic anemia and has a dominant effect over wild type KLF1. Here we describe the effect of the *Nan* mutation during fetal development. We show that *Nan* embryos have defects in erythroid maturation. To identify genes that were differentially regulated in the presence of the *Nan* mutation, we performed RNA-sequencing analysis. Exportin 7 (*Xpo7*) was among the 782 deregulated genes (FDR < 0.01 and absolute fold change > 1.5). Expression of *Xpo7* was almost 4-fold down-regulated in E12.5 *Nan* fetal livers. *Xpo7* has been recently implicated in terminal erythroid differentiation, being involved in nuclear condensation and enucleation. Indeed, KFL1 *Nan* fetal liver cells were larger and showed less chromatin condensation. Knockdown of *Xpo7* in wild type erythroid cells caused a similar phenotype. We conclude that reduced expression of *Xpo7* is partially responsible for the erythroid defects observed in *Nan* fetal liver cells.

INTRODUCTION

Erythropoiesis is the process responsible for the production of red blood cells. Defects in this process lead to anemia and thus insufficient oxygen delivery to the tissues and subsequent organ failure. Therefore, the formation of red blood cells has to be tightly controlled during embryonic development and homeostasis in the adult organism.

KLF1 (previously known as EKLF) is a well-characterized, erythroid-specific transcription factor and one of the critical regulators of red blood cell maturation. KLF1 acts mainly as an activator and its target genes are involved in multiple processes of erythroid differentiation, including cell cycle regulation[1,2], hemoglobin metabolism[3] and expression of membrane skeleton proteins[4,5]. The importance of KLF1 is illustrated by *Klf1* knockout embryos which die around E14 owing to the lack of functional erythrocytes[6,7]. In contrast, heterozygous *Klf1*+/− mice survive into adulthood, indicating that haploinsufficiency for KLF1 is not associated with a severe phenotype[8]. KLF1 has a transactivation domain at the N-terminus and a DNA binding domain, composed of three zinc fingers, at the C-terminus. The zinc fingers mediate specific DNA binding to the 5' CACCC 3' sequence[9]. Mutations in human *KLF1* are found across the entire gene. The majority are missense mutations in the three zinc fingers, which presumably alter the DNA binding/sequence recognition properties of KLF1. Mutations in KLF1 are associated with different phenotypes in humans[10], such as In(Lu) blood group[11], hereditary persistence fetal hemoglobin (HPFH)[12], zinc protoporphyria[13], increased HbA2[14], and congenital dyserythropoietic anemia (CDA)[15].

The Neonatal anemia (*Nan*) mouse is an ethylnitrosourea (ENU)-induced semidominant hemolytic anemia model first described in 1983 by Mary Lyon[16], who, 3 years later, also positioned the mutation on chromosome 8[17]. In 2010, *Klf1* was identified as the gene responsible for this phenotype, due to a single point mutation in the second zinc finger (E339D)[18,19]. While KLF1 *Nan* homozygote mice die around E10, KLF1 *Nan* heterozygote mice survive into adulthood, although with hemolytic anemia. This indicates that the presence of the mutation also affects the function of wild type KLF1 protein, as this phenotype does not occur in *Klf1* haploinsufficient mice. Indeed, the DNA binding properties of *Nan* KLF1 are altered due to steric clash between the carboxyl group of D339 and the methyl group of thymidine, resulting in the deregulation of a subset of target genes[19].

Until now, research has focused on the effects of the *Nan* mutation in adult mice[18-20]. Given that KLF1 expression begins around E7.5[9], it is relevant to investigate the impact of aberrant KLF1 activity during development. In this study, we investigated erythropoiesis during different stages of fetal development. We observed impaired red blood cell maturation at E12.5, as assessed by flow cytometry analysis of the CD71 and Ter119 markers. Expression profiling of E12.5 fetal liver cells revealed 782 deregulated genes in *Nan* mutant samples. These included known KLF1 targets such as Dematin and E2F2. Intriguingly, the nuclear exportin XPO7, which has recently been implicated in nuclear condensation and enucleation during erythroid maturation[21], was one of the newly identified KLF1-regulated genes. The *Nan* mutation resulted in significant XPO7 downregulation. Our data indicate that the increased nuclear size observed in *Nan* erythroid progenitors is caused by reduced expression of XPO7.

MATERIAL AND METHODS

Mice

All animal studies were approved by the Erasmus MC Animal Ethics Committee. The mouse strains used were *Klf1 Nan* mutant[16] and *Klf1* knockout[6]. The *Klf1 Nan* mutant mice were

crossed with a mouse strain carrying a human β-globin locus transgene (line PAC8.1)[22]. Genotyping for KLF1 was performed by polymerase chain reaction (PCR) using DNA isolated from toe biopsies. Primers are listed below. For *Nan* genotyping, the PCR product was digested with DpnII. Embryos were collected at E12.5,E13.5, E14.5 and E18.5; tail DNA was used for genotyping.

Klf1 Nan
Fw: 5'CTGCAGGATTGCAGCTGTAGATAC3'
Rv: 5' AGTCCTTGTGCAGGATCACTCAGA3'
Approximately 340 bp PCR product for the wild type allele and 240 bp for the *Nan* allele

Klf1 knock-out
Fw1: 5'TTGCCGTTTTGCTTTGCCTG3'
Fw2: 5'CGTTGGCTACCCGTGATATTG3'
Rv: 5'GAAGTCCTCCTGGGTGTCCA3'
Approximately 250 bp PCR product for the wild type allele and 270 bp for the ko allele

Human β-globin locus
Fw: 5'GCTGCTGTTATGACCACTAGAGGG3'
Rv: 5'AGACAGGGAAGGAGGTGTGG3'
Approximately 400 bp PCR product

Blood analysis
Peripheral blood was collected from the mandibular vein of adult mice, and standard blood parameters were measured with an automated hematologic analyzer (Scil Vet ABC, Viernheim, Germany).

Cell culture and transduction
I/11 erythroid progenitors and primary mouse fetal liver cells were cultured as described[23]. To induce differentiation of I/11 cells we used StemPRO-34 SFM (10639-011, life technologies) supplemented with 500 µg/mL iron-saturated transferrin (Scipac) and 10 U/mL Epo (Janssen-Cilag).
Lentiviral shRNAs targeting Xpo7 were obtained from the Sigma MISSION shRNA library. The following clones were used: TRCN0000105410 (Xpo7 shRNA #1), TRCN0000105411 (Xpo7 shRNA #2), TRCN0000105412 (Xpo7 shRNA #3) . Lentiviral particles were produced in HEK 293T cells as described previously[24].

RNA isolation and RT-qPCR analyses
RNA was extracted using TRI reagent (Sigma-Aldrich). To synthesize cDNA, 2 µg of RNA were used together with oligo dT (Inivtrogen), RNase OUT (Invitrogen), and SuperScript reverse transciptase II (Invitrogen) in a total volume of 20 µL for 1 hour at 42 degrees. 0.2 µL of cDNA was used for amplification by RT-qPCR. Amplification was performed on a CFX96 Touch Real-Time PCR Detection System (Biorad) with the primers listed below using Platinum Taq DNA polymerase (Invitrogen) and 40 cycles consisting of 95°C for 30 sec, 60°C for 1 min. Specific polymerase chain reaction (PCR) product accumulation was monitored by SYBR Green dye fluorescence (Sigma-Aldrich). Ct values obtained for actin expression were used for normalization. Dissociation curves assessed the homogeneity of PCR products.

Xpo7 exon 1b
Fw: 5'GGGCTGTTGGTGAATTCAACC3'
Rv: 5'TTGAGACATGCGGGACTCAG3'

Actin
Fw: 5'GATTACTGCTCTGGCTCCT3'
Rv: 5'TGGAAGGTGGACAGTGAG3'

Klf1
Fw: 5'CAGCTGAGACTGTCTTACCC3'
Rv: 5'AATCCTGCGTCTCCTCAGAC3'

Protein extraction and western blotting
Total protein extracts from mouse fetal liver cells were prepared according to[25]. To visualize protein expression, cell lysates of 3×10^6 cells were loaded on 10% sodium dodecyl sulfate-polyacrylamide gels for electrophoresis, and the gels were transferred to nitrocellulose blotting membrane 0,45 µm (10600002, GE Healthcare) and probed with specific antibodies. Membranes were stained for Tubulin (T5168, Sigma-Aldrich) as loading control, and for Xpo7 (sc390025, Santa Cruz).

Flow cytometry analysis
Single-cell suspensions collected from fetal livers or I/11 cell cultures were washed twice with PBS and then resuspended in PBS containing 1% (w/v) bovine serum albumin and 1 mM EDTA. Approximately 10^6cells were incubated for 30 minutes at room temperature with Kit-PE (553355, BD Biosciences, dilution 1:200), CD44-APC (559250, BD Biosciences, dilution 1:200), CD71-BV421 (562716; BD Biosciences, dilution 1:1600) and Ter119-APC (17-5921-82; eBiosciences, dilution 1:200) antibodies and in a final volume of 100 µl. The cells were washed, and living cells were distinguished negatively by 7-aminoactinomycin D staining (A1310; Invitrogen). Cells were measured on a Fortessa instrument (BD Biosciences), and data were analyzed with FlowJo software v10 (Tree Star). E18.5 fetal liver cells were sorted by using CD71-FITC (553266; BD Biosciences) antibody and MACS separation columns according to the manufacturer's instructions (Miltenyi Biotec, Bergisch Gladbach, Germany). For enucleation analysis, cells were stained with 10 µg/ml Hoechst 33342 (H3570, Invitrogen), in addition to Ter119-APC (17-5921-82; eBiosciences, dilution 1:100) antibody, for 15 min at room temperature. The I/11 nuclear area was measure and analyzed on a ImageStreamX Mark II Imaging Flow Cytometer (Amnis). The cells were stained with CD71-FITC (333151, BD Biosciences, dilution 1:200) and 10 µg/ml Hoechst 33342 (H3570, Invitrogen).

Cell morphology
Cell morphology was analyzed using cytospins stained with May Grünwald-Giemsa (Medion Diagnostic) and O-dianisidine (Sigma-Aldrich)[26]. Pictures were taken with an Olympus BX40 microscope (40× objective, NA 0.65) equipped with an Olympus DP50 CCD camera and Viewfinder Lite 1.0 acquisition software.
Nuclear size was measured with DAPI staining (P-36931, Life technologies). Pictures were taken with Leica DMRBE microscope (40x objective, NA 0.65) equipped with Hamamatsu ORCA-ER digital camera and Hokawo Imaging Software v 2.6. Images were analyzed using Fiji[27].

RNA-sequencing and analysis

RNA-seq was performed according to manufacturer's instructions (Illumina; San Diego, CA, USA), as described[28]. The sequenced reads were mapped against the mouse genome build mm10 and the human PAC8 sequence based on build hg19 using TopHat 2.0.6[29] with the transcriptome gene annotation of Ensembl v73[30]. Raw counts were generated with HTSeq-count version 0.6.0. using the settings -m union -s no -a 20 (doi:10.1093/bioinformatics/btu638) with the Ensembl 73 gene annotation. The counted data were normalized by the size factor of the libraries using DESeq2 R package[31] and, subsequently, converted to Transcripts Per Million (TMP) as described in[32]. The differentially expressed genes were called using a generalized linear negative binomial model that controlled for the effect of each litter. The calculations were performed by the DESeq2 R package. After using the Wald significance test the False Discover Rates (FDR)/adjusted p-values were calculated with the Benjamini Hochberg method (Benjamini and Hochberg, 1995). Threshold value for differentially expressed genes was set to a FDR <0.01 with an absolute fold change of 1.5. After blind variance stabilizing transformation of the normalized counts[31], the differentially expressed genes were used to calculate the pairwise Spearman's rank correlation (rs) coefficient matrices and the subsequent dissimilarity matrices (1 – rs) for both genes and RNA-Seq samples. The Euclidean distances of the dissimilarity matrices were used to apply hierarchical clustering with complete linkage. The normalized gene expression data were scaled for each row (Z-score) and, subsequently, plotted in a heat map using the R statistical package version 3.2.0 (R-Core-Team: R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing; 2015).

Chromosome Conformation Capture Combined with high-throughput Sequencing (4C-Seq) and data analysis

4C-Seq experiments were carried out as described[33]. Briefly, 4C-seq template was prepared from E13.5 fetal liver or fetal brain cells. In total, between 1 and 8 million cells were used for analysis. The cells were cross-linked with 2% formaldehyde in PBS/10% FBS for 10 min at room temperature. After quenching the reaction with 0.125M glycine, cells were lysed in lysis buffer (50mM Tris pH 7.5, 150mM NaCl, 5mM EDTA, 0.5% NP-40, 1%Triton X-100, 1mM PMSF and protease inhibitor mixture Complete (Roche)). The cross-linked DNA was digested with DpnII as the primary restriction enzyme and BfaI as secondary restriction enzyme. Two replicate experiments were sequenced for each genotype and view point. Primers from the 4C-Seq primer database[34] with additional 50 overhangs of the adapter sequences for Illumina single-read sequencing were used as viewpoint. The sequences are listed below:

Canonical promoter
Fw: 5'TCACTGAGTCAATTTTGCTG3'
Rv: 5' TAGGCCAGACTACTCAGATC 3'

Erythroid specific Exon 1b
Fw: 5'GGTCATGTTTCCTTAGCTCTG3'
Rv: 5'GCTGAGGAGATGATGGATC3'

Twelve PCR reactions were carried out using 100 ng of the resulting 4C template per reaction. The products were pooled together and purified using QIAquick PCR purification kit (Qiagen). The samples were sent for deep sequencing using Illumina HiSeq2000 platform. 4C-Seq data analysis was performed, similarly, as described in[3.] In brief, the 4C-Seq reading primer sequences with their barcodes were used to demultiplex the reads and to trim from

the 5'-end to the first restriction enzyme recognition site. The sequences were mapped, while ignoring quality scores of the read bases and not allowing for a mismatch, to a database of digested genome fragment-ends using the mouse reference genome build mm10. All the 4C-Seq samples passed the quality control threshold values as described in[33]. We normalized the data taking the library size and 4C-Seq fragment-end types[33,34] into account. The 4C-Seq contact profiles were generated after the median value of the biological replicates (wild type fetal liver, fetal brain and *Nan* fetal liver, respectively) was calculated. To smoothen the data we applied a running trimmed(10%) mean approach using 21 fragment-ends in a single window. We further determined differentially contact profiles between 1) wild type fetal liver and brain and between 2) *Nan* fetal liver and wild type fetal liver(data not shown). Since, it is hard to estimate a proper distribution of 4C-Seq data and since 4C-Seq data are biased for each fragment-end differently, we used a non-parametric approach to test for statistical significance between the two phenotypes. We, therefore, ranked all normalized data for each fragment-end, independently, and set ties to the minimal value. Subsequently, we binned, along the locus, the ranks of 21 fragment-ends and calculated the *m* x *n* rank frequency matrix (*m* = 2; *i .e.* the number of different phenotypes; *n* = number of samples). We merged the frequency matrix columns into a 2 x 2 matrix, based on the order of each phenotype and the number of samples *e.g.* if the first ranked phenotype, based on the enumeration of the rank frequencies multiplied by the rank number, has been sampled three times, the first three columns of the *m* x *n* matrix are merged by adding up the row values, and the other columns are merged similarly resulting in a 2 x 2 matrix. Moreover, a X^{2}-test was applied on this 2 x 2 contact frequency matrix. The *p-values* were corrected for multiple hypotheses testing using the Benjamini Hochberg method (Benjamini and Hochberg, 1995) and areas with significant contact differences are indicated in gray (Figure 4B,C). The R statistical package version 3.2.0 was used for the statistical calculations and for generating the 4C-Seq contact plots. (ref: R-Core-Team: R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing; 2015). Gviz was used for plotting the annotation data(Florian Hahne, Steffen Durinck, Robert Ivanek, Arne Mueller, Steve Lianoglou, Ge Tan and Lance Parsons. Gviz: Plotting data and annotation information along genomic coordinates. R package version 1.12.1).

Statistical tests

Statistical analysis of blood parameters was performed by using analysis of variance with Bonferroni correction; flow cytometry data and gene expression results were analyzed by using Mann-Whitney tests. Excel 2003 (Microsoft, Redmond, WA) was used to draw the graphs. Values plus or minus standard deviation are displayed in the figures.

RESULTS

Characterization of *Nan* mutant fetal liver cells

The effect of KLF1 *Nan* mutation has been studied in adult mice[18-20], but data on its effect during gestation is limited. Hence, to study this mutation while hemoglobin switching occurs, we used a *Nan* mouse model carrying one mutant allele (*Nan*/+, from now on called *Nan*). In addition, these animals carried a human β-*globin* locus transgene (line PAC8.1)[22]. At the three stages of development analyzed, E12.5, E14.5, and E18.5, *Nan* embryos were paler than the wild type littermates, indicating anemia, but otherwise looked phenotypically normal. We performed flow cytometry analysis on fetal liver cells using Kit, CD71, Ter119 and CD44 markers, which help tracing red blood cell differentiation, at E12.5, E14.5, and E18.5. We observed a severe downregulation in expression of the Ter119 marker at all three stages

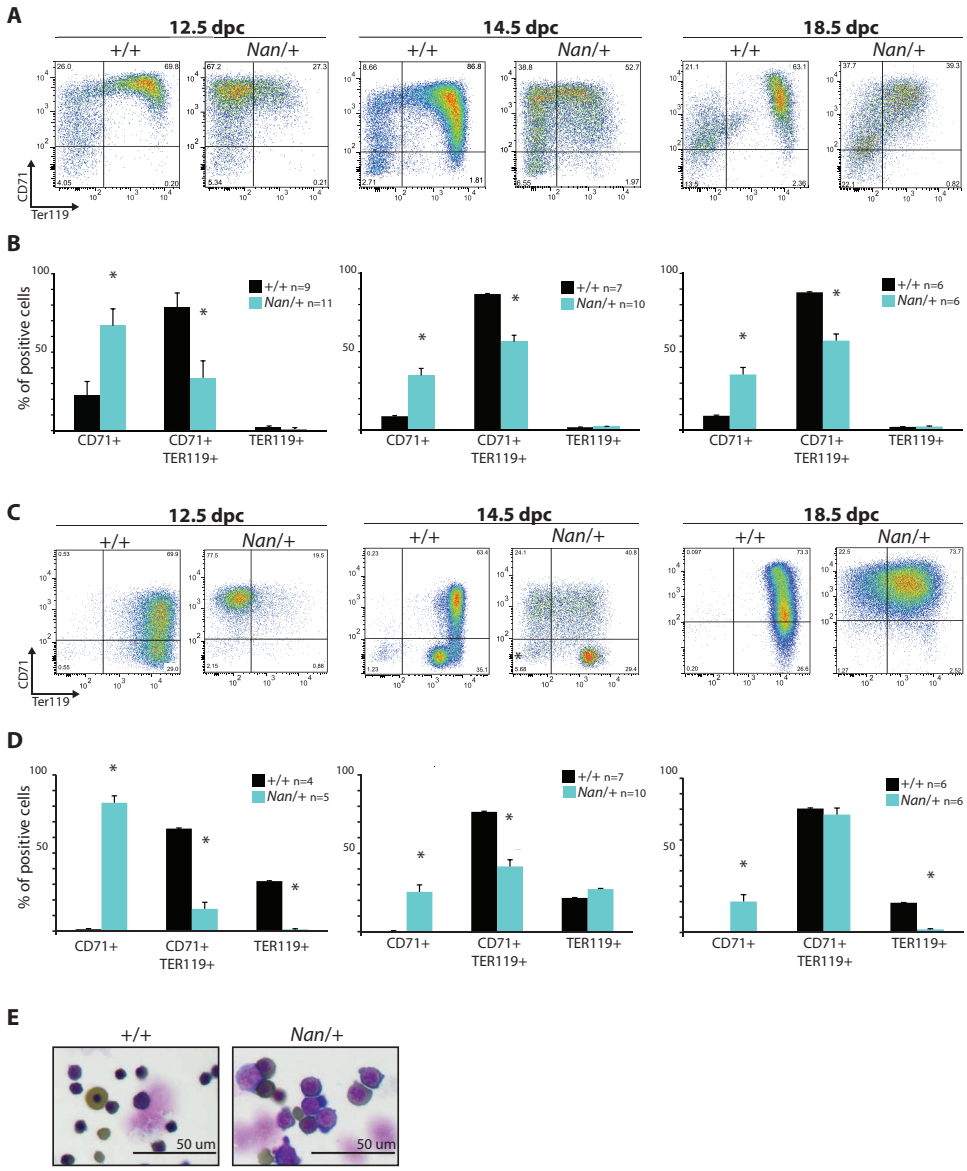


Figure 1. KLF1 Nan embryos flow cytometry. (A) Examples of flow cytometry profiles of CD71 and Ter119 staining of E12.5, E 14.5 and E18.5 wild type and *Nan* mouse fetal livers. (B) Quantification of CD71+, CD71+ Ter119+ and Ter119+ populations. n indicates the number of embryos. * indicates p value <0.01. (C) Examples of flow cytometry profiles of CD71 and Ter119 staining of E12.5, E 14.5 and E18.5 wild type and *Nan* mouse fetal blood. (D) Quantification of CD71+, CD71+ Ter119+ and Ter119+ populations. n indicates the number of embryos. * indicates p value <0.01. (E) Cytospins of E14.5 wild type and *Nan* mouse fetal liver cells stained with May Grünwald-Giemsa and O-dianisidine.

(Figure 1A and 1B). The CD71/Ter119 double positive population was significantly decreased in the *Nan* samples, while the CD71 single positive population showed an increase. No significant differences were observed for Kit and CD44 in the presence of the *Nan* mutation (Supplementary Figure 1A and 1B). In addition, we obtained similar results when assaying embryonic blood, with Ter119 being highly downregulated (Figure 1C and 1D). These results indicate that *Nan* embryos display delayed erythroid maturation compared to the wild type controls. This is in line with the observation that a higher percentage of cells is positive for CD71 in adult blood (Supplementary Figure 2A and 2B), indicative of higher percentage of circulating reticulocytes[19]. Consistent with this notion, analysis of standard blood parameters revealed a significant increase in red cell distribution width (RDW) (Supplementary Figure 2C) in the *Nan* mice. Furthermore, we observed minor, yet significant, decreases in RBC (red blood cell), HGB (total hemoglobin), HCT (hematocrit), MCH (Mean Corpuscular Hemoglobin), MCHC (Mean Corpuscular Hemoglobin Concentration) values (Supplementary Figure 2A and 2B). Interestingly, when comparing *Nan* E14.5 fetal liver cytospins to wild type controls, we observed a marked increase in the average size of the erythroid cells and their nuclei (Figure 1E). Taken together, these data strongly indicate that erythroid maturation is impaired in *Nan* mutants.

Identification of deregulated genes in E12.5 *Nan* fetal livers

In order to identify genes that are affected by the *Nan* mutation, we performed a genome-wide RNA-seq on samples derived from E12.5 *Nan* and wild type fetal livers (*N*=6 each), as at this stage the fetal liver is mainly composed by erythroid cells. 782 genes appeared to be deregulated in the *Nan* mutants (false discovery rate [FDR] <0.01, absolute fold change equal or greater than 1.5), of which 437 were upregulated and 345 downregulated (Figure 2A and 2B). Strikingly, even though KLF1 has been mainly described as a transcriptional activator, the majority of the deregulated genes displayed increased activation in the mutant mice. We postulate that this might be due to secondary effects of KLF1 on other transcriptional regulators and/or aberrant activity of KFL1 *Nan*. To validate the data, we checked the expression values of Epb4.9 and E2F2, genes known to be down-regulated in *Nan* erythroid cells[19] (Figure 2C, left panel and Supplementary Figure 3). Indeed, we detected a significant decreased expression of the transcripts of these two genes in *Nan* fetal livers. Moreover, we observed a significant 2-fold down-regulation of BCL11A, a known target of KFL1[12,35], indicating that the KLF1 *Nan* mutation affects BCL11A expression (Figure 2C, left panel and Supplementary Figure 3). Given the role of BCL11A and KLF1 in globin switching, we checked the expression levels of the β-like globin genes. We found that *Hbb-bh1* was upregulated and the known KLF1 target gene *Hbb-b* was downregulated, consistent with previous reports[19]. In addition, we observed that *Hba-x* was upregulated in E12.5 *Nan* fetal livers (Figure 2C, right panel and Supplementary Figure 3). In accordance with the notion that intact KLF1 fulfils a crucial role in globin switching, our RNA seq data showed that the KLF1 *Nan* mice retained significantly higher expression of human embryonic/fetal globins, such as HBE and HBG1, while they displayed decreased expression of adult globin HBB, albeit not at statistically significant levels (Supplementary Figure 4). In conclusion, we report that KLF1 *Nan* mice exhibit impaired globin switching, evident by the sustained expression of embryonic/fetal-globin at the expense of adult-globin.

The nuclear exportin Xpo7 is downregulated in *Nan* erythroid cells

We found that *Xpo7*, encoding a nuclear exportin, was prominently downregulated in *Nan* E12.5 fetal livers (~4-fold decrease; Figure 2A). This raised our interest since XPO7 was recently reported to be implicated in terminal erythroid differentiation, as a protein involved

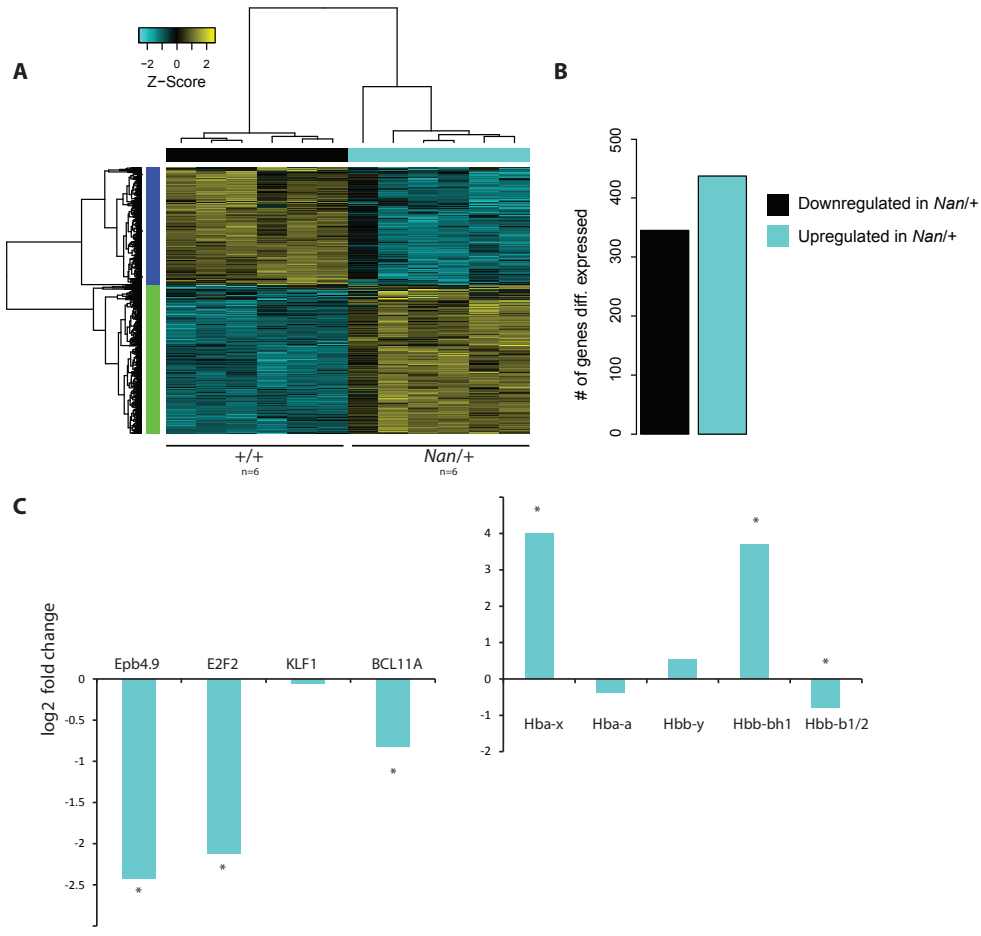


Figure 2. RNA-seq analysis. (A) Hierarchical clustered heat map with scaled Z-score color key of normalized counts of 782 differentially expressed genes in 6 WT (+/+) and 6 *Nan* (*Nan*/+) E12.5 fetal liver samples. Samples with the same genotype are indicated by black (WT) and cyan (Nan) horizontal bars; gene clusters are indicated by green (upregulated in *Nan*) and purple (downregulated in *Nan*) vertical bars. False discovery rate [FDR] <0.01, fold-change equal or greater than 1.5. (B) Schematic representation of the number of downregulated and upregulated genes in the Nan E12.5 fetal livers. (C) Log2 values of fold-change for selected genes. * indicates FDR <0.01.

in enucleation[21]. To corroborate the RNA-seq data, we investigated expression of *Xpo7* in *Nan* E14.5 fetal livers by RT-qPCR. We found that *Xpo7* transcripts including the erythroid-specific first exon were reduced by approximately 90% in the *Nan* samples (Figure 3A). Transcripts using the canonical first exon were barely detectable in all samples (not shown). Importantly, XPO7 protein levels were also reduced in *Nan* fetal liver cells (Figure 3B). To investigate whether *Xpo7* expression is dependent on KLF1, we measured *Xpo7* mRNA and protein levels in *Klf1* ko/ko erythroid cells. In E13.5 *Klf1* ko/ko fetal livers, which lack any KLF1 expression[6], expression of *Xpo7* mRNA and protein is significantly reduced (Figure 3C

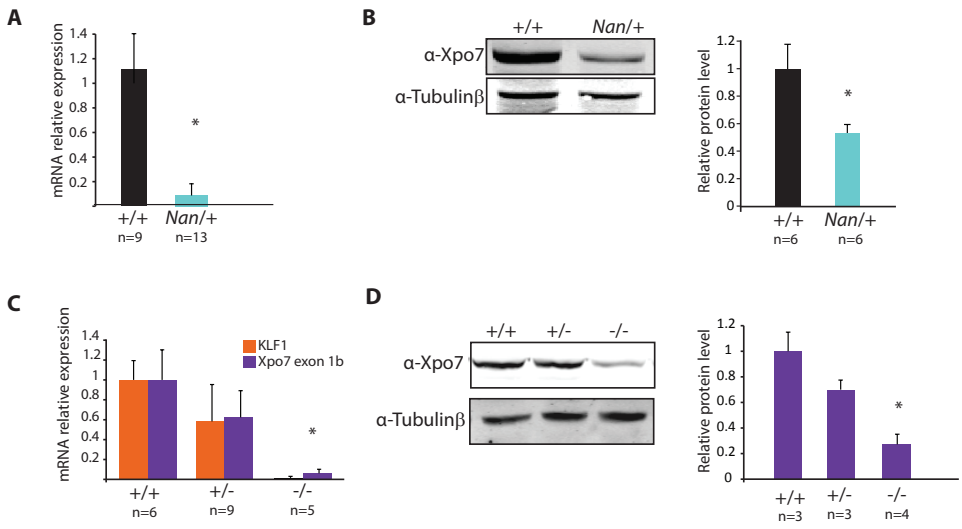


Figure 3. Xpo7 expression is activated by KLF1. (A) Xpo7 mRNA relative values in wild type and Nan E14.5 fetal livers. * indicates p value <0.01. n indicates the number of embryos. (B) Western blot analysis of Xpo7 protein in wild type and Nan E14.5 fetal livers and quantification. β-tubulin was used as loading control. * indicates p value <0.01. n indicates the number of embryos. (C) KLF1 and Xpo7 mRNA relative expression values in wild type, KLF1 heterozygotes and KLF1 knockout E13.5 fetal livers. * indicates p value <0.01. n indicates the number of embryos. (D) Western blot analysis of XPO7 protein in wild type, KLF1 heterozygotes and KLF1 knockout E13.5 fetal livers and quantification. β-tubulin was used as loading control. * indicates p value <0.01. n indicates the number of embryos.

and 3D). Remarkably, downregulation of *Xpo7* was also observed in *Klf1* wt/ko fetal livers, although to a lesser extent than observed in *Klf1* ko/ko fetal livers (Figure 3C and 3D). This indicates that, similarly to *Bcl11a* [12,35], activation of *Xpo7* by KLF1 is dose-dependent. In agreement with the notion that KLF1 is a direct activator of *Xpo7*, KLF1 binds to the canonical promoter and first intron of the *Xpo7* gene in mouse[36] and human[37] erythroid cells.

The chromatin conformation of the *Xpo7* locus is not affected in *Nan* erythroid cells

Since KLF1 is required for formation of the active chromatin hub in the β-globin locus[38], we performed 4C-seq experiments on the *Xpo7* locus in E13.5 wild type fetal livers and fetal brains and *Nan* mouse fetal livers (Figure 4A) . To investigate potential changes in chromatin conformation we used the canonical promoter of *Xpo7* as view point (Figure 4B). Interestingly, we identified a loop between the canonical promoter of *Xpo7* (situated at the beginning of exon 1a) and the exon that produces the erythroid-specific form of Xpo7 (exon 1b), indicating that these two regions are spatially in proximity to each other in erythroid cells (Figure 4B). Remarkably, this loop has less contact frequencies in fetal brain. However, we found that there were little local changes in the chromatin conformation between wild type and *Nan* samples (Figure 4B). We repeated the experiment using the erythroid-specific promoter as view point. This confirmed the results obtained with the canonical promoter as viewpoint. (Figure 4C). We suggest that this loop might recruit transcription factors binding to the area of the canonical promoter to the vicinity of the erythroid promoter, thereby facilitating the expression of the erythroid-specific *Xpo7* transcript.

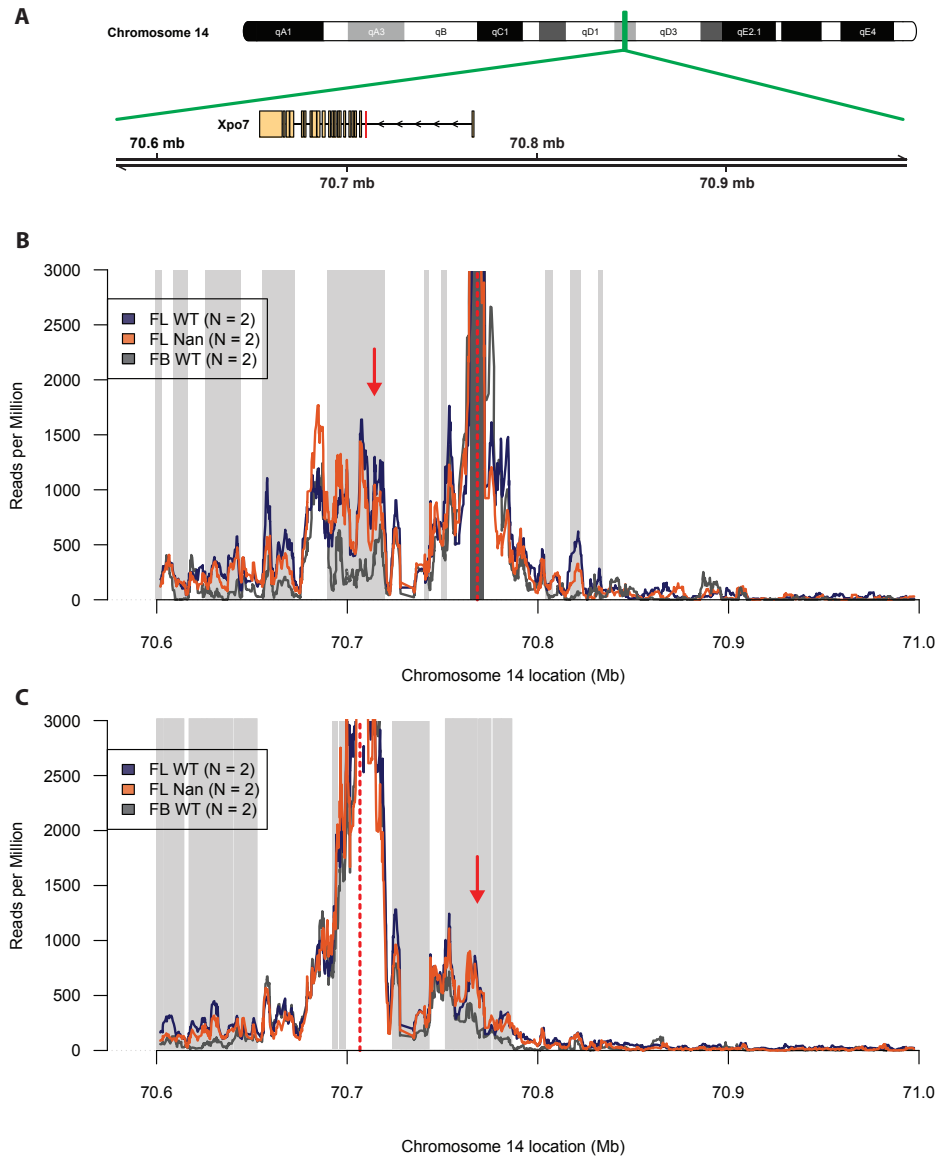


Figure 4. 4C-seq analysis of the Xpo7 locus. (A) Schematic representation of chromosome number 14. The green box indicates the zone where the Xpo7 gene resides. The RefSeq mm10 Xpo7 gene is indicated by rectangles (exons) and arrows (introns) that point to the direction of transcription. The location is indicated in Mega basepairs (Mb). The erythroid-specific first exon (exon 1b) is indicated by a red box. (B-C) 4C-Seq representation of the chromosome contact frequencies detected using the canonical promoter of Xpo7 (B) and the region of the erythroid specific Xpo7 exon (C) as viewpoints. The mean of a running window of 21 restriction fragment-ends of the median value of the biological replicates with a maximum of 3000 is indicated by colored lines. Loci with a statistically significant (FDR <0.05) higher contact frequencies and reads per million >250 in wild type fetal liver compared to the fetal brain are indicated by light grey boxes. Loci with a statistically significant (FDR <0.05) higher contact frequencies and reads per million >250 in fetal brain compared to the wild type fetal liver brain are indicated by dark grey boxes. Purple,KLF1 +/+ Fetal Liver; Orange, KLF1 Nan/+ Fetal liver; Grey, KLF1 +/+ Fetal brain.

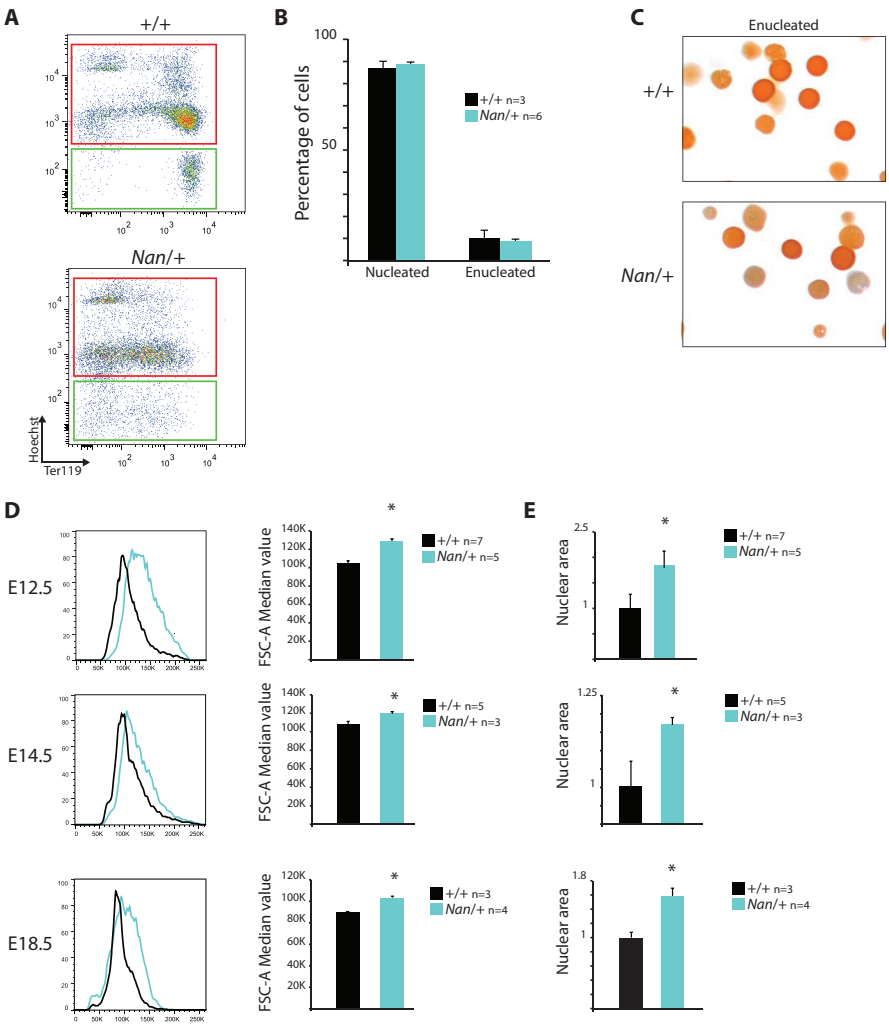


Figure 5. KLF1 Nan fetal liver cells enucleation and cell size. (A) Gating strategies of Hoechst- and Ter119-stained E14.5 fetal liver cells. Red, Hoechst+ population; Green, Hoechst- population. (B) Quantification of the number of nucleated (Hoechst+) and enucleated (Hoechst-) cells. n indicates the number of embryos. (C) Cytospins stained with May Grünwald-Giemsa and O-dianisidine of Hoechst- wild type and Nan sorted populations. (D) Representative FSC-A value flow cytometry plots of E12.5, E14.5 and E18.5 wild type and Nan fetal liver cells and quantification. * indicates p value <0.01. n indicates the number of embryos. (E) Relative nuclear area size quantification of E12.5, E14.5 and E18.5 wild type and Nan fetal liver cells. * indicates a p value <0.01. n indicates the number of embryos.

KLF1 *Nan* mouse fetal liver cells present defects in nuclear condensation

Since Xpo7 has been implicated in enucleation of erythroid cells *in vitro*[21] we analyzed enucleation in *Nan* fetal livers. Enucleation was quantified in E14.5 fetal livers by flow cytometry using the erythroid marker Ter119 and Hoechst-33342 as a nuclear stain. We observed similar percentages of enucleated cells between *Nan* and control fetal liver

samples (Figures 5A and 5B). Similar results were obtained with E12.5 and E18.5 fetal liver cells (data not shown). To check whether the flow cytometry analysis could indeed discriminate nucleated from enucleated cells, we sorted the Hoechst-positive and Hoechst-negative populations and prepared cytospins. This analysis confirmed that all the Hoechst-negative cells identified by flow cytometry were indeed enucleated (Figure 5C). In addition, assessing enucleation levels of mouse fetal liver cells cultured in proliferative medium and in differentiation medium from embryos at E12.5 and E14.5 by flow cytometry, we found similar ratios of nucleated *versus* enucleated cells when comparing the control with the *Nan* samples (Supplementary Figure 5A and 5B). Nevertheless, we observed a striking increase in the percentage of large cells in the fetal livers of the *Nan* embryos. We quantified cell sizes in control and *Nan* fetal livers using flow cytometry. This revealed a significant increase in average cell size at E12.5, E14.5 and E18.5 in the *Nan* samples (Figure 5D). In line with this finding, we found a significant increase in nuclear area of the *Nan* fetal liver cells compared to control fetal liver cells in cytospin slides stained with the nuclear dye Hoechst 33342 (Figure 5E). These data are consistent with the notion that Xpo7 is involved in nuclear condensation, a process that is thought to precede enucleation. However, despite the impaired nuclear condensation, the cells still go through enucleation.

Xpo7 knock down in I/11 cells mimics the phenotype of *Nan* cells

To further analyze the role of Xpo7 in erythroid differentiation, we knocked down Xpo7 using three different shRNAs in the factor-dependent immortalized mouse erythroid cell line I/11 [39]. We reached an efficiency of ~60% knockdown, as shown by Western blot (Figure 6A). Before differentiation, we observed a minor difference in the expression of the surface markers CD71 and Ter119 and no difference in cell size between the control and the knockdown cells (Supplementary Figure 7). In contrast, upon transfer to differentiation medium, the maturation of Xpo7 knockdown cells was impaired, as shown by CD71 and Ter119 flow cytometry analysis (Figure 6B). In addition, on average the cells were bigger (Figure 6C) and showed less condensation of the nucleus (Figure 6D), comfirming the already published observation that XPO7 is important for terminal erythroid nuclear maturation[21]. So far, our attempts to rescue the *Nan* phenotype by overexpressing exogenous XPO7 in mouse fetal liver cells has not been successful, either due to limited expression of functional XPO7 or possible contributions of other (co)factors to the observed phenotype (data not shown). Collectively, our findings indicate that *Xpo7* is partially responsible for the phenotype of the *Nan* mice, and establish *Xpo7* as a novel erythroid target gene of KLF1.

DISCUSSION

Erythropoiesis is a complex process that involves many players, whose activity ensures the production of functional red blood cells. One of these players is KLF1, a transcription factor with multiple roles during terminal erythroid differentiation. Firstly, it is essential for globin regulation, in particular for direct activation of β-globin and indirect via repression of γ-globin by the transcription factor BCL11A[12,40]. In addition, it acts as a master regulator of genes activated during differentiation of red blood cells, such as membrane proteins, heme synthesis enzymes and cell cycle regulators[2,4,41]. Hence, it comes as no surprise that *Klf1* knock out embryos die due to severe anemia, and that the phenotype is not rescued by exogenous expression of a β-like globin gene[6,7]. For the same reasons, KLF1 mutations can lead to diverse phenotypes in humans[10]. One example is a missense mutation in the second zinc finger of KLF1 (E325K) that causes CDA[15]. This mutation is believed to affect binding of KLF1 to its target genes thereby exerting a dominant-negative effect on wild type

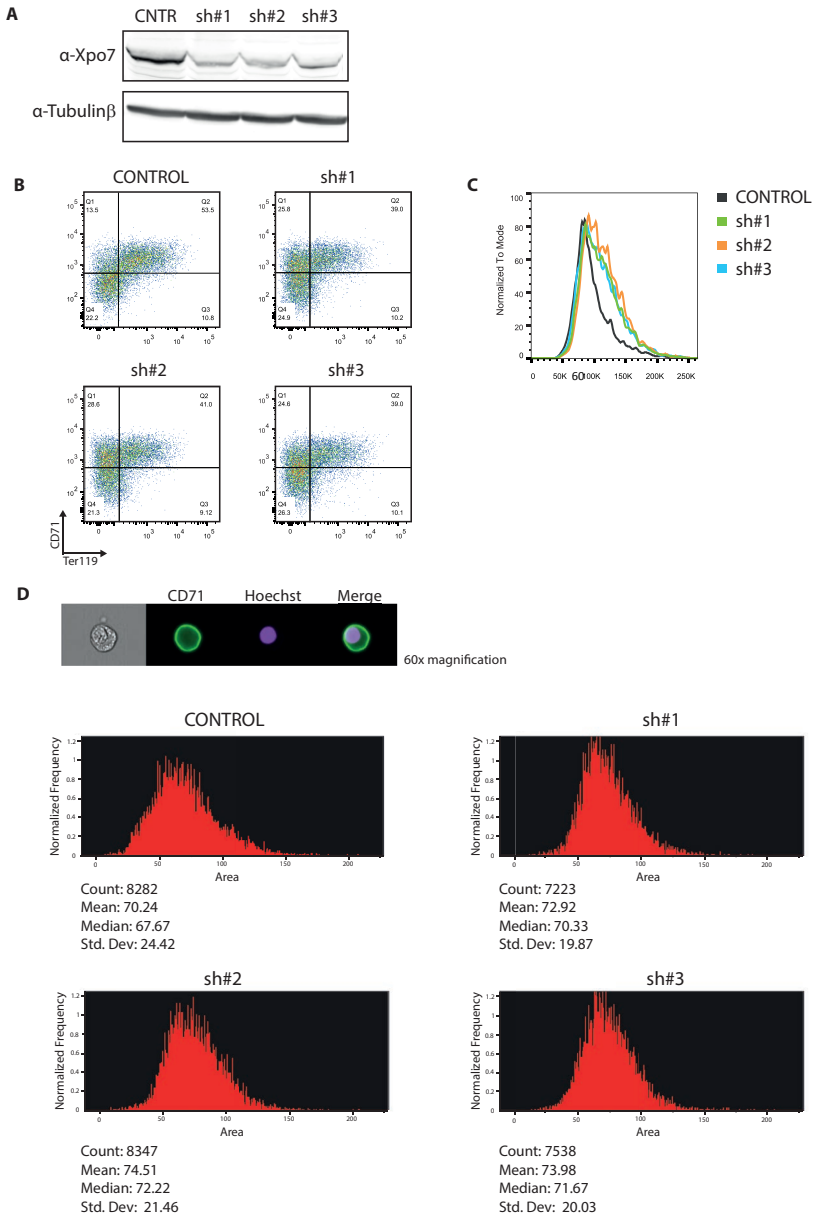


Figure 6. Xpo7 knockdown in I/11 mouse immortalized erythroid progenitor cells. (A) Western blot analysis indicating the efficiency of Xpo7 knockdown using 3 different shRNAs. β-tubulin was used as loading control. (B) Example of flow cytometry profiles of CD71 and Ter119 staining of I/11 cells transduced with either control, sh#1, sh#2 or sh#3 lentiviruses in differentiation conditions. (C) Representative FSC-A value flow cytometry plots of I/11 cells transduced with either control, sh#1, sh#2 or sh#3 lentiviruses in differentiation conditions. (D) ImageStream area quantification of I/11 cells transduced with either control, sh#1, sh#2 or sh#3 lentiviruses in differentiation conditions (arbitrary units). The total number of counted cells, the mean, the median and the standard deviation are indicated below the histograms. On the top a representative cell from the control is depicted. n=1.

KLF1 protein. Similar effects have been described for an ethylnitrosourea (ENU)-induced semidominant hemolytic anemia mouse model, termed as *Nan*. These mice have a missense mutation, p.E339D, in a position homologous to that of the human CDA mutation[18,19]. Studies on the effect of the *Nan* mutation in adult mice have revealed that these animals display life-long anemia[18-20].

In this paper we present our findings on the effects of the *Nan* mutation on definitive erythropoiesis. We showed that erythroid maturation is impaired in *Nan* fetal livers at E12.5, E14.5 and E18.5. By RNA-seq analysis we identified 782 differentially expressed genes in *Nan versus* control E12.5 fetal livers. In agreement with a previous report on erythropoiesis in adult *Nan* mice[19], we found that the expression of globin genes is impaired in *Nan* mice. In particular, the upregulation of βh1 globin can be explained by the significant lower expression of BCL11A in *Nan* embryos, as the inhibition of its transcription is dependent on the presence of BCL11A[42]. In accordance, also the human globins HBE and HBG1 were upregulated, meaning that the switch between embryonic/fetal and adult globins does not work properly in *Nan* mice. *Xpo7*, encoding a nuclear exportin, was one of the most significantly downregulated genes in *Nan* fetal livers. This gene caught our attention since a recent paper described that *Xpo7* is required for nuclear condensation and enucleation during terminal erythroid differentiation *in vitro*[21]. In addition, the observation that *Xpo7* expression was also reduced in *Klf1* knockout fetal livers indicated that *Xpo7* might be a direct target of KLF1. Supporting this notion, data mining of ChIP-seq results revealed that KLF1 binds to the *Xpo7* locus in mouse[36] and human[37] erythroid cells. Collectively, these data suggested that, similar to the β-globin locus[38], KLF1 might have a role in the spatial organization of the *Xpo7* locus. 3C-seq analysis of the *Xpo7* locus demonstrated that it adopts a different conformation in fetal liver cells compared to fetal brain cells. The presence of the *Nan* mutation doesn't appear to mediate any major changes in the chromatin conformation of the *Xpo7* locus. In the β-globin locus, KLF1 is required for the formation of the active chromatin hub[38]. In the absence of KLF1, an erythroid-specific chromatin structure is still present but the interactions between the distal regulatory elements and the β-globin promoter are specifically lacking. Notably, expression of β-globin is completely dependent on KLF1, while this is not the case for many other KLF1 target genes, including *Xpo7*. This suggests that KLF1 primarily acts as a transcriptional activator for this class of target genes. Presumably, other erythroid transcription factors such as GATA1[43] are the major players in the erythroid-specific spatial organization of these loci. An interesting observation is the presence of a loop between the promoter of the canonical Xpo7 promoter (in front of exon 1a) and the erythroid-specific promoter (in front of exon 1b), absent in the Fetal brain control. This loop is likely the consequence of recruitment of the two promoters to the same transcription factory[44]. We then further investigated the role of *Xpo7* during fetal development. *Xpo7* knock down in cultured mouse fetal liver cells impairs chromatin condensation and enucleation during terminal erythroid differentiation[21]. Although in *Nan* mice enucleation still occurs, an emerging scenario is that reduced *Xpo7* expression due to the *Nan* mutation impairs chromatin condensation during terminal erythroid differentiation. We propose that this contributes to the maturation defects of *Nan* erythrocytes in fetal and adult definitive erythropoiesis. Indeed, knockdown of *Xpo7* in immortalized mouse erythroblasts cells leads to impaired maturation of the cells, evident by the dysregulation of the flow cytometry markers CD71 and Ter119 and the presence of larger cells with bigger nuclei in the cultures. We will repeat the ImageStream analysis to check whether the observed difference in nuclear area is statistically significant. Our data are partially in agreement with the recent publication on the role of XPO7 in erythroid maturation[21]. It is important to keep in mind that we cannot compare the levels of XPO7 protein between our

system and that of Hattangadi *et al*.[21]. An emerging question is how *Nan* cells manage to enucleate in the presence of reduced levels of XPO7. One possibility is that the level of XPO7 present *in vivo* in the *Nan* mice might suffice for correct enucleation of the erythrocytes but still affects nuclear condensation. Alternatively, the downregulation of XPO7 might just slow down nuclear condensation, but cells will eventually manage to shed their nucleus when condensation is completed. Lastly, a protein with a role similar to that of XPO7 may successfully substitute for it, thus enabling cells to enucleate. We favour a scenario in which chromatin condensation is crucial for enucleation[45-47], with XPO7 as an important effector. It is not clear whether enucleation can happen before nuclear condensation is completed. Our study suggests that impaired nuclear condensation contributes to the erythroid maturation defects observed in the *Nan* mice.

Understanding the role of KLF1 during erythroid maturation and the enucleation process has clinical significance for the production of red blood cells *in vitro* for transfusion purposes. In recent years, many efforts have been made to produce erythrocytes *in vitro* starting from hematopoietic stem cells, embryonic stem cells or induced pluripotent stem cells[48-50]. Efficient enucleation is one of several challenges that have to be overcome in order to produce sufficient numbers of fully functional erythrocytes *in vitro*. More in depth knowledge of this process might guide the development of better strategies to achieve this goal.

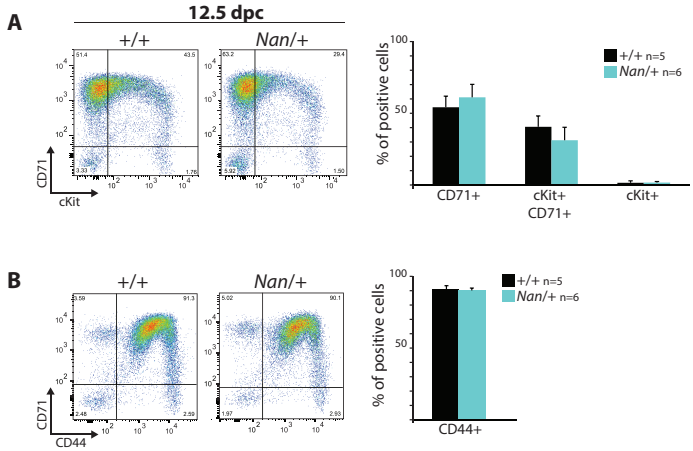
Author contributions

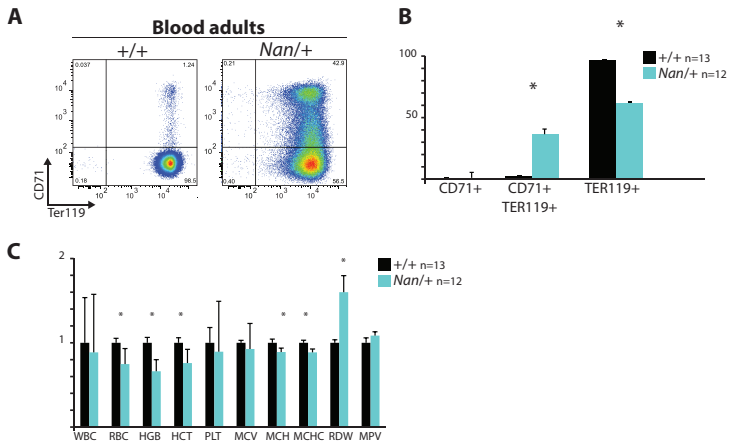
I.C., T.v.D,. H.J.G.v.d.W and S.P. designed the experiments. I.C., N.G, R.S.. and T.v.D. performed the experiments. H.J.G.v.d.W analyzed the 4C-seq and RNA-seq data and wrote the bioinformatics section. Z.O and W.F.v.I. performed high-throughput sequencing experiments. The paper was written by I.C., H.J.G.v.d.W , T.v.D. and S.P.

REFERENCES

1. Tallack, M.R., Keys, J.R., Humbert, P.O. & Perkins, A.C. EKLF/KLF1 controls cell cycle entry via direct regulation of E2f2. *J Biol Chem* **284**, 20966-74 (2009).
2. Pilon, A.M. et al. Failure of terminal erythroid differentiation in EKLF-deficient mice is associated with cell cycle perturbation and reduced expression of E2F2. *Mol Cell Biol* **28**, 7394-401 (2008).
3. Donze, D., Townes, T.M. & Bieker, J.J. Role of erythroid Kruppel-like factor in human gamma- to beta-globin gene switching. *J Biol Chem* **270**, 1955-9 (1995).
4. Drissen, R. et al. The erythroid phenotype of EKLF-null mice: defects in hemoglobin metabolism and membrane stability. *Mol Cell Biol* **25**, 5205-14 (2005).
5. Nilson, D.G., Sabatino, D.E., Bodine, D.M. & Gallagher, P.G. Major erythrocyte membrane protein genes in EKLF-deficient mice. *Exp Hematol* **34**, 705-12 (2006).
6. Nuez, B., Michalovich, D., Bygrave, A., Ploemacher, R. & Grosveld, F. Defective haematopoiesis in fetal liver resulting from inactivation of the EKLF gene. *Nature* **375**, 316-8 (1995).
7. Perkins, A.C., Sharpe, A.H. & Orkin, S.H. Lethal beta-thalassaemia in mice lacking the erythroid CACCC-transcription factor EKLF. *Nature* **375**, 318-22 (1995).
8. Wijgerde, M. et al. The role of EKLF in human beta-globin gene competition. *Genes Dev* **10**, 2894-902 (1996).
9. Siatecka, M. & Bieker, J.J. The multifunctional role of EKLF/KLF1 during erythropoiesis. *Blood* **118**, 2044-54 (2011).
10. Borg, J., Patrinos, G.P., Felice, A.E. & Philipsen, S. Erythroid phenotypes associated with KLF1 mutations. *Haematologica* **96**, 635-8 (2011).
11. Singleton, B.K., Burton, N.M., Green, C., Brady, R.L. & Anstee, D.J. Mutations in EKLF/KLF1 form the molecular basis of the rare blood group In(Lu) phenotype. *Blood* **112**, 2081-8 (2008).
12. Borg, J. et al. Haploinsufficiency for the erythroid transcription factor KLF1 causes hereditary persistence of fetal hemoglobin. *Nat Genet* **42**, 801-5 (2010).
13. Satta, S. et al. Compound heterozygosity for KLF1 mutations associated with remarkable increase of fetal hemoglobin and red cell protoporphyrin. *Haematologica* **96**, 767-70 (2011).
14. Perseu, L. et al. KLF1 gene mutations cause borderline HbA(2). *Blood* **118**, 4454-8 (2011).
15. Arnaud, L. et al. A dominant mutation in the gene encoding the erythroid transcription factor KLF1 causes a congenital dyserythropoietic anemia. *Am J Hum Genet* **87**, 721-7 (2010).
16. Lyon, M.F.G., P.H.; Loutit, J.F., Peters, J. . Dominant hemolytic anemia. *Mouse News Letter* **68:68**(1983).
17. Lyon, M.F. Position of neonatal anaemia (Nan) on chromosome 8. *Mouse News Letter* **74:95**(1986).
18. Heruth, D.P. et al. Mutation in erythroid specific transcription factor KLF1 causes Hereditary Spherocytosis in the Nan hemolytic anemia mouse model. *Genomics* **96**, 303-7 (2010).
19. Siatecka, M. et al. Severe anemia in the Nan mutant mouse caused by sequence-selective disruption of erythroid Kruppel-like factor. *Proc Natl Acad Sci U S A* **107**, 15151-6 (2010).
20. White, R.A. et al. Hematologic characterization and chromosomal localization of the novel dominantly inherited mouse hemolytic anemia, neonatal anemia (Nan). *Blood Cells Mol Dis* **43**, 141-8 (2009).
21. Hattangadi, S.M. et al. Histones to the cytosol: exportin 7 is essential for normal terminal erythroid nuclear maturation. *Blood* **124**, 1931-40 (2014).
22. de Krom, M., van de Corput, M., von Lindern, M., Grosveld, F. & Strouboulis, J. Stochastic patterns in globin gene expression are established prior to transcriptional activation and are clonally inherited. *Mol Cell* **9**, 1319-26 (2002).
23. von Lindern, M. et al. Leukemic transformation of normal murine erythroid progenitors: v- and c-ErbB act through signaling pathways activated by the EpoR and c-Kit in stress erythropoiesis. *Oncogene* **20**, 3651-64 (2001).
24. Zufferey, R., Nagy, D., Mandel, R.J., Naldini, L. & Trono, D. Multiply attenuated lentiviral vector achieves efficient gene delivery in vivo. *Nat Biotechnol* **15**, 871-5 (1997).
25. Schreiber, E., Matthias, P., Muller, M.M. & Schaffner, W. Rapid detection of octamer binding proteins with 'mini-extracts', prepared from a small number of cells. *Nucleic Acids Res* **17**, 6419

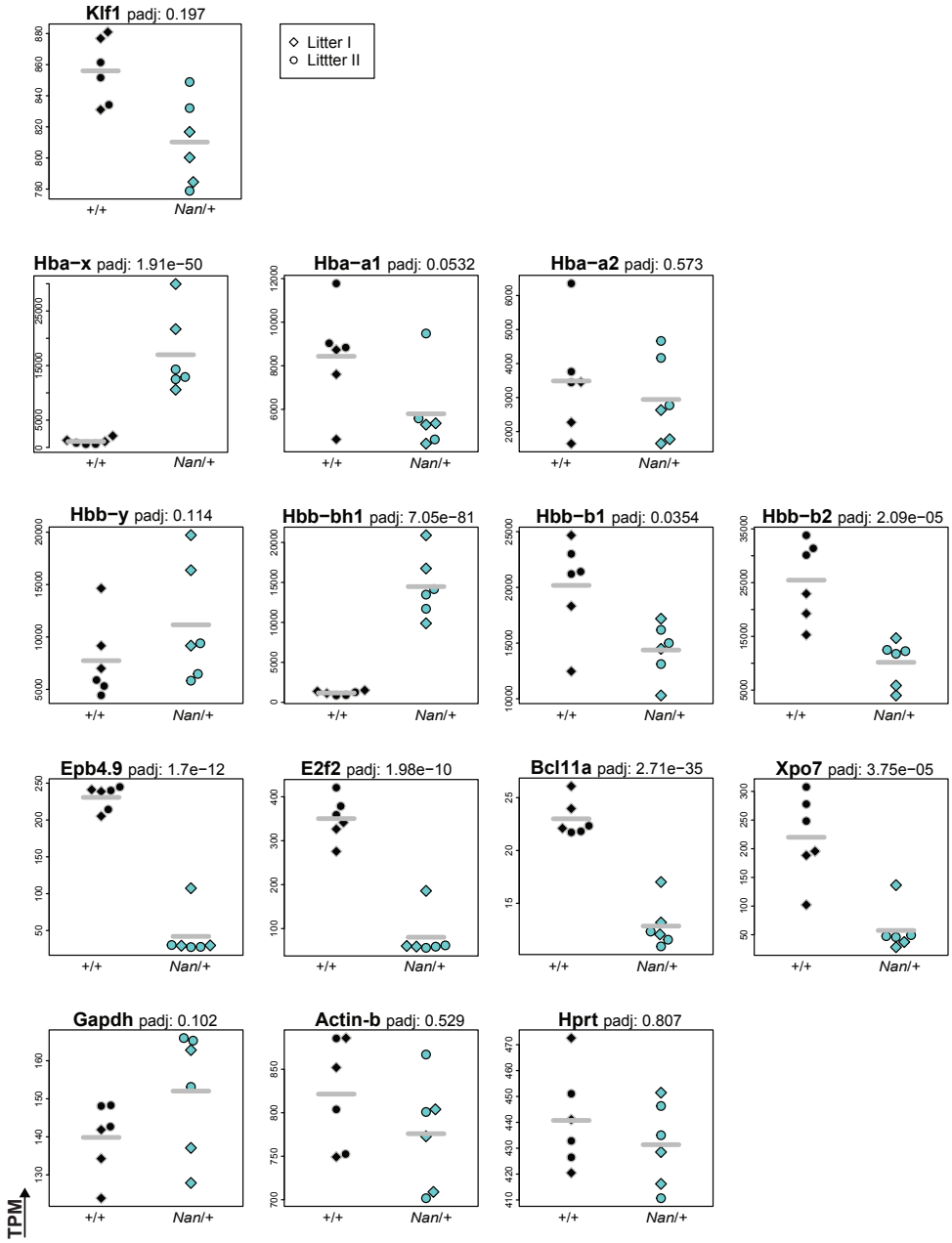
(1989).
26. Beug, H., Palmieri, S., Freudenstein, C., Zentgraf, H. & Graf, T. Hormone-dependent terminal differentiation in vitro of chicken erythroleukemia cells transformed by ts mutants of avian erythroblastosis virus. *Cell* **28**, 907-19 (1982).
27. Schindelin, J. et al. Fiji: an open-source platform for biological-image analysis. *Nat Methods* **9**, 676-82 (2012).
28. Meinders, M. et al. Sp1/Sp3 transcription factors regulate hallmarks of megakaryocyte maturation and platelet formation and function. *Blood* **125**, 1957-67 (2015).
29. Kim, D. et al. TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. *Genome Biol* **14**, R36 (2013).
30. Flicek, P. et al. Ensembl 2014. *Nucleic Acids Res* **42**, D749-55 (2014).
31. Love, M.I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* **15**, 550 (2014).
32. Wagner, G.P., Kin, K. & Lynch, V.J. Measurement of mRNA abundance using RNA-seq data: RPKM measure is inconsistent among samples. *Theory Biosci* **131**, 281-5 (2012).
33. van de Werken, H.J. et al. 4C technology: protocols and data analysis. *Methods Enzymol* **513**, 89-112 (2012).
34. van de Werken, H.J. et al. Robust 4C-seq data analysis to screen for regulatory DNA interactions. *Nat Methods* **9**, 969-72 (2012).
35. Zhou, D., Liu, K., Sun, C.W., Pawlik, K.M. & Townes, T.M. KLF1 regulates BCL11A expression and gamma- to beta-globin gene switching. *Nat Genet* **42**, 742-4 (2010).
36. Pilon, A.M. et al. Genome-wide ChIP-Seq reveals a dramatic shift in the binding of the transcription factor erythroid Kruppel-like factor during erythrocyte differentiation. *Blood* **118**, e139-48 (2011).
37. Su, M.Y. et al. Identification of biologically relevant enhancers in human erythroid cells. *J Biol Chem* **288**, 8433-44 (2013).
38. Drissen, R. et al. The active spatial organization of the beta-globin locus requires the transcription factor EKLF. *Genes Dev* **18**, 2485-90 (2004).
39. Dolznig, H. et al. Establishment of normal, terminally differentiating mouse erythroid progenitors: molecular characterization by cDNA arrays. *FASEB J* **15**, 1442-4 (2001).
40. Xu, J. et al. Transcriptional silencing of {gamma}-globin by BCL11A involves long-range interactions and cooperation with SOX6. *Genes Dev* **24**, 783-98 (2010).
41. Hodge, D. et al. A global role for EKLF in definitive and primitive erythropoiesis. *Blood* **107**, 3359-70 (2006).
42. Sankaran, V.G. et al. Developmental and species-divergent globin switching are driven by BCL11A. *Nature* **460**, 1093-7 (2009).
43. Vakoc, C.R. et al. Proximity among distant regulatory elements at the beta-globin locus requires GATA-1 and FOG-1. *Mol Cell* **17**, 453-62 (2005).
44. Ghamari, A. et al. In vivo live imaging of RNA polymerase II transcription factories in primary cells. *Genes Dev* **27**, 767-77 (2013).
45. Popova, E.Y. et al. Chromatin condensation in terminally differentiating mouse erythroblasts does not involve special architectural proteins but depends on histone deacetylation. *Chromosome Res* **17**, 47-64 (2009).
46. Ji, P., Yeh, V., Ramirez, T., Murata-Hori, M. & Lodish, H.F. Histone deacetylase 2 is required for chromatin condensation and subsequent enucleation of cultured mouse fetal erythroblasts. *Haematologica* **95**, 2013-21 (2010).
47. Zhang, L., Flygare, J., Wong, P., Lim, B. & Lodish, H.F. miR-191 regulates mouse erythroblast enucleation by down-regulating Riok3 and Mxi1. *Genes Dev* **25**, 119-24 (2011).
48. Giarratana, M.C. et al. Ex vivo generation of fully mature human red blood cells from hematopoietic stem cells. *Nat Biotechnol* **23**, 69-74 (2005).
49. Lu, S.J. et al. Biologic properties and enucleation of red blood cells from human embryonic stem cells. *Blood* **112**, 4475-84 (2008).
50. Kurita, R. et al. Establishment of immortalized human erythroid progenitor cell lines able to produce enucleated red blood cells. *PLoS One* **8**, e59890 (2013).



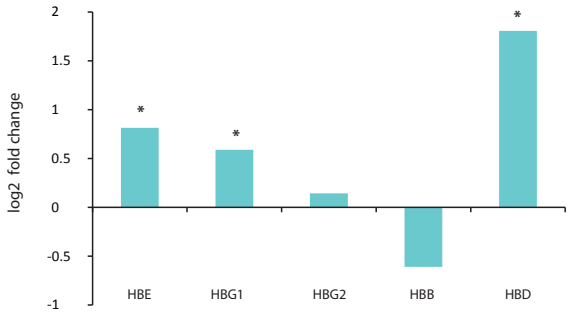


Supplementary Figure 1. (A) Example of flow cytometry profiles of CD71 and Kit staining of wild type and *Nan* E12.5 fetal livers and quantification of CD71$^+$, CD71$^+$ Kit$^+$ and Kit$^+$ populations. n indicates the number of embryos. (B) Example of flow cytometry profiles of CD71 and CD44 staining of wild type and *Nan* E12.5 fetal livers and quantification of C44$^+$ population. n indicates the number of embryos.

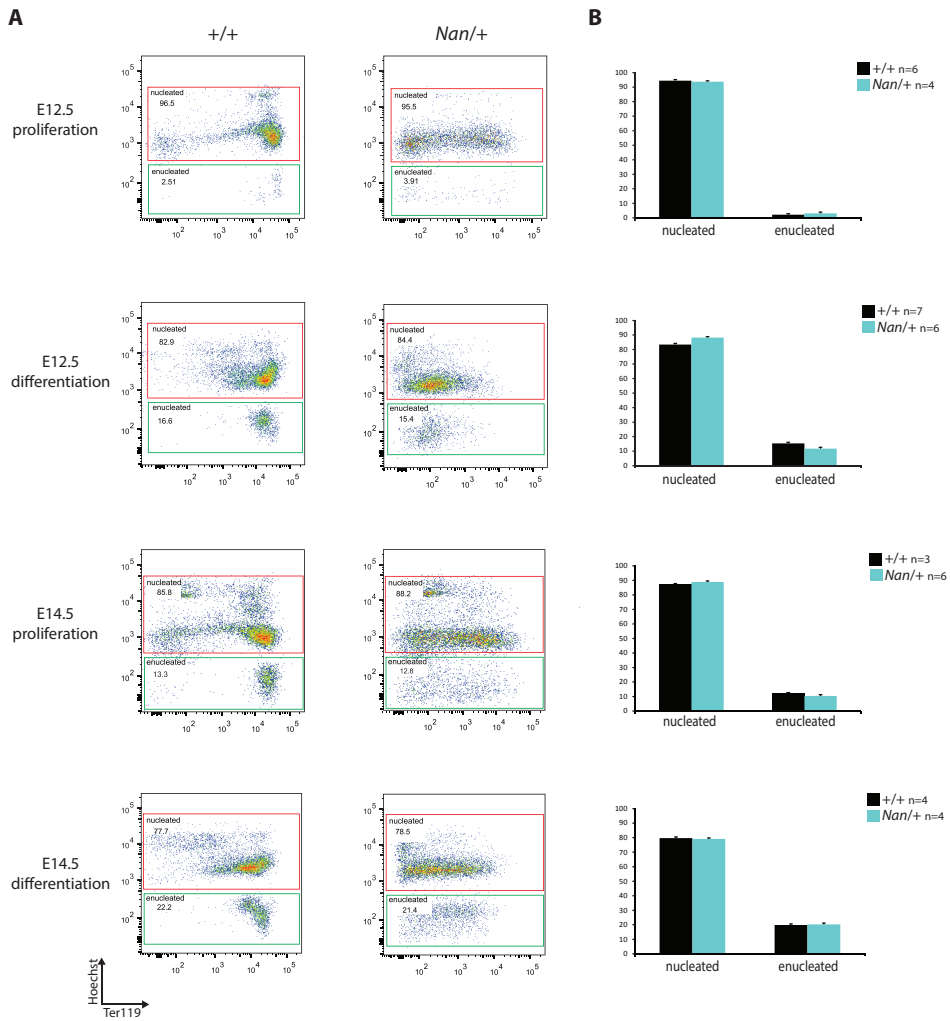
Supplementary Figure 2. (A) Example of flow cytometry profiles of CD71 and Ter119 staining of E12.5, E14.5 and E18.5 wild type and *Nan* adult blood. (B) Quantification of CD71$^+$, CD71$^+$ Ter119$^+$ and Ter119$^+$ populations. n indicates the number of embryos. * indicates p value <0.01. (C) Hematologic parameters of wild type and *Nan* adult mice. * indicates p value <0.01. WBC, with blood cells; RBC, red blood cells; HGB, hemoglobin; HCT, hematocrit; PLT, platelets; MCV, mean corpuscular volume, MCH, mean corpuscular hemoglobin, MCHC, mean corpuscular hemoglobin concentration, RDW, red cell distribution width; MPV, mean platelet volume.



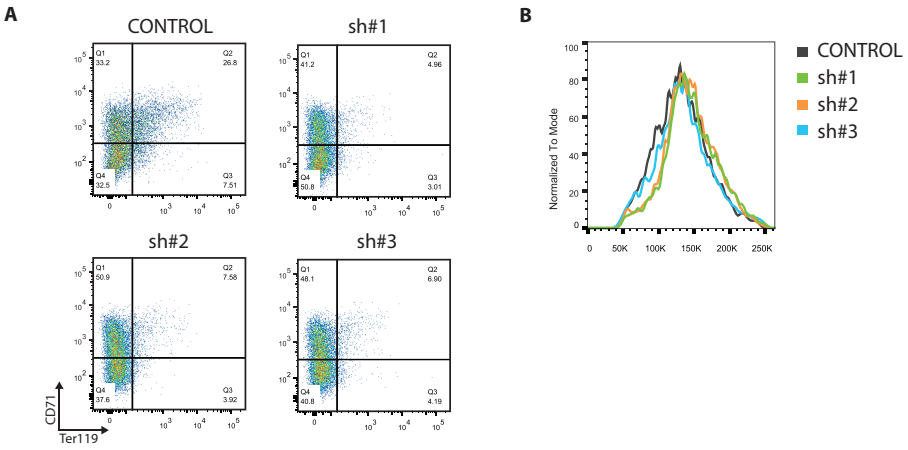
Supplementary Figure 3. Gene expression levels were determined by RNA-sequencing of two litters in Wild type (black) and *Nan*(cyan) E12.5 mouse fetal liver cells. Levels of selected erythroid cell-related genes combined with the controls Gapdh, Actin b and Hprt are shown in Transcripts Per Million(TPM) with adjusted p-values(padj).



Supplementary Figure 4. Log2 values of fold-change for human β-globin genes. * indicates p value <0.01.



Supplementary Figure 5. (A) Gating strategies of Hoechst- and Ter119-stained E12.5 and E14.5 fetal liver cells in proliferative and differentiation medium. Red, nucleated population (Hoechst+); Green, enucleated population (Hoechst-). (B) Quantification of the number of nucleated (Hoechst+) and enucleated (Hoechst-) cells. n indicates the number of embryos.



Supplementary Figure 6. (A) Example of flow cytometry profiles of CD71 and Ter119 staining of I/11 cells transduced with either control, sh#1, sh#2 or sh#3 lentiviruses in proliferating conditions. (B) Representative FSC-A value flow cytometry plots of I/11 cells transduced with either control, sh#1, sh#2 or sh#3 lentiviruses in proliferating conditions

Chapter 4

BSDC1, an unexpected link between KLF1 and the Golgi apparatus

Ileana Cantù[1], Nynke Gillemans[1], Sreya Basu[1], Jeroen Demmers[2], Niels Galjart[1], Martin Lowe[3], Sjaak Philipsen[1] and Thamar van Dijk[1].

1 Department of Cell Biology, Erasmus MC, Rotterdam, The Netherlands
2 Department of Proteomics, Erasmus MC, Rotterdam, The Netherlands.
3 Faculty of Life Sciences, University of Manchester, Manchester, United Kingdom

Manuscript in preparation

ABSTRACT

The Golgi apparatus is the organelle responsible for the modification and sorting of proteins received by the endoplasmic reticulum and destined for the lysosomes, the plasma membrane and secretion in the extracellular space. It is composed of a variable number of flattened cisternal membranes in continuous communication with each other. Owing to its complexity, many aspects of its function remain unclear. Here, we report on a novel Golgi protein called BSDC1, which we identified in a screen for potential targets of the erythroid-specific transcription factor KFL1. KLF1 regulates BSDC1 expression in erythroid cells. Furthermore, we demonstrate that BSDC1 co-localizes with GOLGB1 at the rim of the Golgi, but not with GM130. To gain more insight into this unexpected connection between KLF1 and the Golgi apparatus, we are currently focusing on the function of BSDC1.

INTRODUCTION

The Golgi apparatus is an organelle present in every eukaryotic cell. It is composed of stacks of flattened cisternal membranes and associated vesicles[1]. It is located near the rough endoplasmic reticulum (RER) and the nucleus and has the key function of modifying and sorting newly synthesized proteins received from the RER. The enzymes performing the post-translational modifications in the Golgi are all membrane-bound. Phosphorylation, glycosylation, addition of galactose and sialic acid, removal of mannose residues, are indicative examples of the reactions that occur in this organelle[2-4]. These post-translational modifications also form a signal sequence that regulates the cellular localization of the modified proteins. Depending on the imposed modification, target proteins are shuttled to the lysosomes, the plasma membrane or outside the cell. Moreover, it is also responsible for the synthesis of glycolipids and sphingomyelin, a specific type of phospholipid. The Golgi apparatus is composed of three main sections: the cis-Golgi Network (CGN), which represents the entry area of the Golgi, the Golgi stack, which is the major processing area, and the trans-Golgi Network (TGN), where the final reactions and sorting happen. Even though the Golgi apparatus was discovered more than 100 years ago, its peculiar and complicated structure has rendered its study challenging. A controversial topic is how proteins move through the Golgi. There are 3 main hypotheses: cisternal maturation, vesicular transport and rapid partitioning[3]. The cisternal maturation model proposes that vesicles sequentially deliver processing enzymes to the Golgi, and this in turn causes the maturation of the cisternae. According to the vesicular transport model, proteins are moved through the Golgi by anterograde transport from early to late cisternae. Finally, the rapid partitioning model suggests that specific phosphoglycerolipid domains present in each stack of the Golgi execute the transport. To date, no definitive proof for any of these three models exists.

The Golgi apparatus has a unique feature enabling its complex functions, which is the presence of large coiled-coil proteins, called golgins, on its cytoplasmic surface[5]. Different golgins reside at different regions of the Golgi. For instance, GM130 is mainly present in the *cis*-Golgi[6]. GM130 can interact with p115, a protein present in the cytosol, implicated in the tethering of ER-derived transport vesicles to the Golgi in mammals and yeast[7-12]. Another example is GOLGB1, which is localized on the rim of the Golgi, and adopts a coiled-coil structure for most of its length[13,14]. GOLGB1 has also been shown to interact with p115. The interactions between p115 and GM130 or GOLGB1 might create a ternary giantin-p115-GM130 complex forming tether linking vesicles and cisternae[15]. Alternatively, GOLGB1 and GM130 might compete for the interaction with p115 and hence p115- GOLGB1 and p115-GM130 interactions might mediate independent membrane tethering events[16]. Golgins have diverse functions, being involved in the Golgi matrix formation and in tethering the membranes of the Golgi apparatus to each other or to the cytoskeleton[5]. In addition, golgins recruit transport vesicles and signaling molecules to the Golgi apparatus, while they also play an important role during mitosis and apoptosis, two processes that implicate the disruption of the organelle[5]. Different golgins have different roles in the tethering process, depending on their localization in the Golgi apparatus. Nevertheless, the loss of one of these golgins often has no or minor effect on the organelle's function or causes tissue-specific defects, suggesting that golgins are redundant in their action [17-23] .

The Golgi apparatus is well developed in erythroblasts but is eventually lost during erythroid maturation[24]. Depending on their maturation level, erythroid cells have specific proteins in their membrane. We can therefore assume that the Golgi apparatus is primarily involved in the tailored modification and accurate sorting of these membrane proteins. Thus far,

only a few studies have been performed to elucidate the role of the Golgi apparatus in erythropoiesis[25-27]. An erythroid-specific transcription factor with a major role in red blood cell maturation is KLF1[28-32]. KLF1 has a diverse role and regulates, amongst others, cell-cycle progression[31,32], the expression of hemoglobins[28], and the expression of membrane skeleton proteins[29,30]. Hence, mutations in KLF1 have diverse effects and result in different phenotypes in humans[33]. The KLF1 *Nan* mouse model carries an ethylnitrosourea (ENU)-induced mutation, leading to the substitution E339D in KLF1. This in turn leads to aberrant KLF1 activity and deregulation of a subset of its downstream targets. Therefore, the KLF1 *Nan* mutation can be used as a tool to identify new potential KLF1 targets and as such help unravel its complex role during red blood cell development. By means of RNA-sequencing experiments performed using KLF1 *Nan*/+ E12.5 mouse fetal liver RNA, we identified BSDC1 (BSD domain containing 1) as a potential target gene of KLF1, downregulated in the presence of the KLF1 *Nan* mutation (Chapter 3). Although the BSD domain has been found in chromatin proteins such as the mammalian p62 subunit of the TFIIH complex and in the TFB1 subunit of the yeast RNA polymerase II[34], we observed that BSDC1 primarily localizes to the Golgi apparatus. In this paper we try to gain insight in this unexpected connection between a Golgi protein (BSDC1) and the erythroid-specific transcription factor KLF1, raising the possibility of a novel role for KLF1 in modulating the function of the Golgi apparatus.

MATERIAL AND METHODS

Mice

All animal studies were approved by the Erasmus MC Animal Ethics Committee. The mouse strains used were *Klf1 Nan* mice and *Klf1* ko mice (Chapter 3). Genotyping for *Klf1* was performed by polymerase chain reaction (PCR) using DNA isolated from toe biopsies. Primers are listed below. For *Nan* genotyping, the PCR product was digested with DpnII. Embryos were collected at E12.5, E13.5, E14.5 and E18.5; tail DNA was used for genotyping.

Klf1 Nan
Fw: 5'CTGCAGGATTGCAGCTGTAGATAC3'
Rv 5' AGTCCTTGTGCAGGATCACTCAGA3'
Approximately 340 bp PCR product for the wild type allele and 240 bp for the *Nan* allele

Kfl1 knockout
Fw1: 5'TTGCCGTTTTGCTTTGCCTG3'
Fw2: 5'CGTTGGCTACCCGTGATATTG3'
Rv: 5'GAAGTCCTCCTGGGTGTCCA3'
Approximately 250 bp PCR product for the wild type allele and 270 bp for the ko allele

Cell Culture

MEL, 293T and K562 cells were cultured in Dulbecco's Modified Eagle Medium (DMEM; Invitrogen) supplemented with 10% fetal calf serum and penicillin-streptomycin (17-602E, Lonza). Primary mouse fetal liver cells were cultured as described in[35]. The mouse fetal liver cells were differentiated using StemPRO-34 SFM (10639-011, Life Technologies) added with 500 µg/mL iron-saturated transferrin (Scipac) and Epo (Janssen-Cilag, 10 U/mL).

Plasmids

Mouse *Bsdc1* cDNA (Dharmacon) was amplified using Q5 High-Fidelity DNA Polymerase (M0491L, New England Biolabs) and cloned in pMT2 plasmid, downstream of a

biotinylation sequence and an HA-tag, using SalI/XbaI sites and the primers: Fw:5' ATATGTCGACCGCGGAAGGGGAGGACGTG3' and Rv:5' TAATGATATCTAGATCATTCCCAGTCCTCCC 3'. Bio-HA-BSDC1 was then released using NotI sites and ligated in the pEV3 vector[36]. GMAP-210-Myc[37], Myc-GM130[38] and FLAG-GolgB1[39] have been described previously.
To produce BSDC1-GFP fusion protein, mouse *Bsdc1* cDNA (Dharmacon) was amplified using Q5 High-Fidelity DNA Polymerase (M0491L,NEB) and cloned in pEGFP_N2 plasmid using XhoI and BamHI restriction sites, in frame with the GFP cDNA, using the primers: 5'CCGCTCGAGACCATGGCGGAAGGGGAGGACGTG3' and 5'CGGGATCCCTTCCCAGTCCTCCCACTCTAC3'.

Transient transfections and generation of stable clones
Transient transfection of 293T was performed as described previously[40]. The production of stable MEL clones expressing biotinylated BSDC1 was performed in MEL cells expressing the BirA biotin ligase[41]. 50x10^6 cells were electroporated with linearized pEV3 plasmid containing Bio-HA-BSDC1 with the following settings: Resistance indefinite, 960 μF, 250V. Cells were then seeded as a single cell /100 μl in a 96-well plate. Stable clones were selected with neomycin (100μg/ml, Invitrogen). After 2 weeks, growing clones were transferred to a 6-well plate and expanded. Stable U2OS clones expressing BSDC1-GFP fusion protein were produced by transfecting U2OS cells with pEGFP_N2 plasmid containing BSDC1 in addition to a pBABE-puro resistance plasmid. Transfected cells were selected adding puromycin (1μg/ml, Sigma-Aldrich) to the medium. GFP positive colonies were picked from the plate and assessed for the expression of BSDC1-GFP fusion protein.

Knockdown and transduction
Lentiviral shRNA constructs for *Bsdc1* were obtained from the Sigma MISSION shRNA library. The following clones were used: TRCN0000129543 (Bsdc1 shRna #1), TRCN0000129787 (Bsdc1 shRna #2), TRCN0000131205 (Bsdc1 shRna #3), TRCN0000147512 (Bsdc1 shRna #4). Lentivirus was produced by transient transfection of 293T cells as described before[42]. Transduction was performed overnight.

Protein extracts and Western Blot analysis
Whole cell lysate were prepared in CARIN Buffer (20 mM Tris-HCl pH 8.0, 137 mM NaCl, 10 mM EDTA, 50 mM NaF, 1% NP40, 10% glycerol) supplemented with Complete protease inhibitor mix (Roche) and Pefablock (Roche). Cells were lysed 30 minutes on ice. After centrifugation at 13000 rpm 5 minutes at 4°C, the supernatant was retrieved and the same volume of 2x Sample Buffer (125 mM Tris-HCl pH 6.8, 5 % SDS, 0.004 % Bromophenol Blue, 1.427 M (10%) β-mercaptoethanol, 20 % glycerol) was added. For cytoplasmic and nuclear extracts, cells were harvested at 1200 rpm and washed once with PBS. The cell pellet was then resuspended in Buffer A (10 mM HEPES-KOH pH 7.9, 1.5 mM $MgCl_2$, 10 mM KCl) supplemented with Complete protease inhibitor mix (Roche), Pefablock (Roche), 0.1 mM DTT, and incubated 15 minutes on ice. Cells were vortexed and spun at 13000 rpm 30 sec. The cytoplasmic fraction was retrieved and transferred to another Eppendorf. The nuclei pellet was resuspended in Buffer C (20 mM HEPES-KOH pH 7.9, 25% glycerol, 420 mM NaCl, 1.5 mM $MgCl_2$, 0.2 mM EDTA), supplemented with Complete protease inhibitor mix (Roche), Pefablock (Roche), and 0.1 mM DTT. After 30 minutes on ice, the nuclear extracts were pelleted at 13000 rpm at 4°C for 5 minutes to separate the proteins form the rest of the nucleus. The supernatant was then transferred to a new Eppendorf tube. Subcellular fractionation was performed as described[43].
Cell lysates were loaded on 10% sodium dodecyl sulfate-polyacrylamide gels for

electrophoresis, and the gels were transferred to nitrocellulose blotting membrane 0,45 μm (10600002, GE Healthcare) 1 hour at 4°C, 200 mA . Nitrocellulose membranes were blocked in 1% bovine serum albumin (BSA, Sigma) for 1 hour at room temperature, probed with the appropriate primary antibodies at 4°C overnight and analyzed using the Odyssey Infrared Imaging System (Li-Cor Biosciences, Cambridge, UK). The primary antibody signal was detected using IRDye Antibodies (Li-core). The membrane was probed with: BSDC1 (ARP-42508, Aviva Systems Biology), PRMT1 (2449S, Cell Signalling), Lamininβ1 (sc-6018, Santa Cruz Antibodies), Histone H3 (ab1791, Abcam), Tubulin (T5168, Sigma-Aldrich), Actin (sc-1616, Santa Cruz Antibodies), HA (SC 7392, Santa Cruz Antibodies), Myc (sc-40, Santa Cruz Antibodies).

Streptavidin pull down and mass spectrometry analysis

Paramagnetic streptavidin beads (Dynabeads M-280, Dynal (Great Neck, NY)) were blocked by washing three times in TBS-0.3% Nonidet P-40 with 200 ng/μl purified chicken serum albumin (Sigma-Aldrich) and 1% BSA (Sigma). We used ≈10 μl of beads per $50x10^6$ cells. For pull-down, cell lysates were combined with 3 times the cell lysate volume of TBS/0.3% Nonidet P-40, 2 times the cell lysate volume of Tris-0,45% Nonidet P-40, 250 units of Benzonase (Novagen) and the streptavidin beads and incubated 3 h at 4 °C. This was followed by three washes in TBS-0.3% Nonidet P-40 . Bound material was eluted by boiling for 5 min in Laemmli protein sample loading buffer and analyzed by immunoblotting as above or by MS.

Digestion with trypsin (Promega, sequencing grade) on paramagnetic streptavidin beads and liquid chromatography-tandem MS (LC-MS/MS) were performed as described previously[41,43]. Nanoflow LC-MS/MS was performed on an 1100 series capillary LC system (Agilent Technologies, Santa Clara, CA) coupled to an LTQ-Orbitrap mass spectrometer (Thermo), operating in positive mode and equipped with a nanospray source. Peptide mixtures were trapped on a ReproSil C18 reversed phase column (Dr Maisch GmbH; column dimensions 1.5 cm × 100 μm, packed in-house) at a flow rate of 8 μl/min. Peptide separation was performed on ReproSil C18 reversed phase column (Dr Maisch GmbH; column dimensions 15 cm × 50 μm, packed in-house) using a linear gradient from 0 to 80% B (A = 0.1% formic acid; B = 80% (v/v) acetonitrile, 0.1% formic acid) in 70 min and at a constant flow rate of 200 nl/min using a splitter. The column eluent was directly sprayed into the electrospray ionization (ESI) source of the mass spectrometer. Mass spectra were acquired in continuum mode; fragmentation of the peptides was performed in data-dependent mode. Peak lists were automatically created from raw data files using the Mascot Distiller software (version 2.3; MatrixScience). The Mascot search algorithm (version 2.2, MatrixScience) was used for searching against the IPI_mouse database (version 3.83, containing 60,010 sequences and 27,475,843 residues). The peptide tolerance was set to 10 ppm and the fragment ion tolerance was set to 0.8 Da. A maximum number of two missed cleavages by trypsin were allowed and carbamidomethylated cysteine and oxidized methionine were set as fixed and variable modifications, respectively. Search results were parsed into a home-built database system for further analysis. Entries were parsed if they had a minimum peptide Mascot score of 25, and a significance threshold of $p < 0.05$; the option “require red bold” was also selected. Using these parameters yields an estimated peptide false discovery rate (FDR) of 3–5% against a target decoy database. The data was converted using PRIDE Converter[44] (http://pride-converter.googlecode.com).

RNA isolation and RT-qPCR analyses

RNA was extracted using TRI reagent (Sigma-Aldrich), in accordance with the protocol

provided by the company. To synthesize cDNA, 2 µg of RNA was used together with oligo dT (Invitrogen), RNase OUT (Invitrogen), SuperScript reverse transciptase II (Invitrogen) in a total volume of 20 µL for 1 hour at 42 degrees. 0.2 µL of cDNA was used for further amplification by RT-qPCR. Amplification was performed on a CFX96 Touch Real-Time PCR Detection System machine (Biorad) with the primers listed below using Platinum Taq DNA polymerase (Invitrogen) and 40 cycles consisting of 95°C for 30 sec, 60°C for 1 min. Specific polymerase chain reaction (PCR) product accumulation was monitored by SYBR Green dye fluorescence (Sigma-Aldrich). Ct values obtained for actin expression were used for normalization. Dissociation curves were used to determine the homogeneity of PCR products.

Bsdc1
Fw: 5' TCATCACCCTGATGGGTACA 3'
Rv: 5' CATCAGGTTCGTTGCAGTAC 3'

Actin
Fw: 5'GATTACTGCTCTGGCTCCT 3'
Rv: 5'TGGAAGGTGGACAGTGAG 3'

Klf1
Fw: 5'CAGCTGAGACTGTCTTACCC 3'
Rv: 5'AATCCTGCGTCTCCTCAGAC 3'

Immunohistochemistry
Immunihistochemistry was performed using a ChemMate Dako EnVision Detection Kit, Peroxidase/DAB, Rabbit/Mouse (K5007, DakoCytomation B.V.) as already described[45]. Briefly, embryos were fixed with 4% PFA (paraformaldehyde) in PBS overnight at 4°C and then embedded in paraffin and sectioned in7 µm sections. Antigen retrieval was performed with microwave treatment in 10 mM citric acid buffer pH 6.0. Sections were blocked with 5% BSA in PBS for 10 min and then incubated with BSDC1 antibody (HPA031360, Sigma) or Rabbit IgG control (sc-3888, Santa Cruz Biotechnology), diluted in 5% BSA in PBS 30 min at room temperature in a humidified chamber. Peroxidase signal from the secondary antibody was detected by incubation for 30 min at room temperature with HRP-conjugated dextran polymer reagent and then by incubation with 3,3'-diaminobenzidine tetrahydrochloride (1:50 dilution of ChemMate DAB+ Chromogen in ChemMate Substrate Buffer). Slides were then counterstained with hematoxylin, dehydrated through graded alcohol and xylene, and mounted. Negative controls was performed by using Rabbit IgG antibody. Pictures were taken with an Olympus BX40 microscope (20× objective, NA 0.4) equipped with an Olympus DP50 CCD camera and Viewfinder Lite 1.0 acquisition software.

Immunofluorescence and confocal microscopy
U2OS cells were grown on 18-mm cover slips. K562 and MEL cells were spotted on Poly-Prep slides (P0425-72EA, Sigma). The cells were then fixed with 4% PFA in PBS. After permeabilization with 0.15% Triton X-100 in PBS and blocking with 1% BSA/PBS the slides were stained with the respective primary antibodies: BSDC1 (HPA031360, Sigma), Golgb1 (BioLegend, PRB-114C), GM130 (BD Transduction Laboratories), and afterwards with the secondary antibodies conjugated with Alexa Fluor 488 or 594 (Life Technologies). All cell samples were imaged using a Leica SP5 confocal laser scanning microscope using the LAS AF software (2.7.4.10100, TCS SP5) provided with the instrument. The system was equipped

with a 63x plan-apochromat oil NA1.4 DIC objective. The pinhole diameter was set to 1 airy unit. DAPI, Alexa 488 and Alexa 594 fluorochromes were excited with a 405nm diode laser, a 488nm Argon laser and a 561nm laser respectively and detected using a multi-track imaging mode of which the band pass filters were 410-485nm (DAPI), 500-575nm (Alexa 488) and 591-660nm (Alexa 594). 8-bit images with a 1024 x 1024 pixels frame size were acquired with 400Hz scan speed, 4-times line averaging and an optical sectioning of 300nm.

Flow cytometry analysis

K562 single-cell suspensions were washed twice with phosphate-buffered saline and then resuspended in phosphate-buffered saline containing 1% (w/v) bovine serum albumin and 1 mM EDTA. Approximately 10^6 cells were incubated for 30 minutes at room temperature with α-ERMAP antibody (1:100, bs12333R, Bioss Antibodies) in a final volume of 100µL. Cells were then washed once with PBS and incubated with a goat α-rabbit antibody (1:2000, ab150077, Abcam) 30 minutes at room temperature in a final volume of 100µL. The cells were washed, and measured on a Fortessa instrument (BD Biosciences), and data were analyzed with FlowJo software v10 (Tree Star).

RESULTS

BSDC1 is a potential KFL1 target

By performing RNA-sequencing on KLF1 *Nan*/+ E12.5 mouse fetal RNA, we identified approximately 780 deregulated genes, (false discovery rate [FDR] <0.01, fold change equal or greater than 1.5), of which 437 were upregulated and 345 were downregulated (Chapter 3, Figure 2). Among the downregulated genes, we focused our attention on the ones that had the highest expression values in wild type cells, as they retain the highest probability to have a physiological function in erythroid cells. We picked those genes with expression levels comparable to the already identified KLF1 target genes Ebp4.9[46], E2F2[46] and PKLR[47], and a fold-change greater than 2 (Table 1). Out of the genes meeting these criteria, we selected BSDC1 for the reasons described below. BSDC1 is a protein with unknown function and is downregulated 3-fold in the *Nan*/+ fetal livers compared to the wild type controls. It contains a BSD domain[48], composed of approximately 60 amino acids, named after the BTF2-like transcription factors, Synapse-associated proteins and DOS2-like proteins (Figure 1). Chromatin associated proteins, such as the p62 subunit of the TFIIH complex and the TFB1 subunit of yeast RNA polymerase II, contain BSD domains[34]. BSDC1 is a rather conserved protein between different species, 90% between mouse and human, 65% between mouse and Xenopus and 62% between mouse and zebrafish (Figure 1). If we take into account just the BSD domain, the conservation is even higher: 98%, 79% and 81% respectively. Furthermore, published human KLF1 ChIP-sequencing experiments show that KLF1 binds to the promoter of BSDC1[49]. This region is also occupied by GATA1 , NF-E2 and p300. Moreover, it displays an enrichment in H3K4me2 and H3K4me3, modifications associated with the promoters of active genes. Notably, BSDC1 expression levels increase during erythroid maturation, according to published data available online (mouse in[50] and human in GEO Profile ID: 70509590, Supplementary Figure 1). In summary, we selected BSDC1 because it shows high expression in mouse fetal liver cells, it is downregulated in the presence of KLF1 *Nan* mutation, it is possibly regulated by KLF1 in a direct manner, its expression increases during red blood cell development and it might be a novel chromatin associated factor.

Table 1. Downregulated genes in KLF1 Nan/wt E12.5 fetal livers. Down-regulated genes identified in the *Nan*/+ mice with a log2FC>1, [FDR] <0.01 and expression values>1000. Already known KLF1 targets are shown in yellow, in pink the two genes selected for further studies: *Xpo7* (Chapter3) and *Bsdc1*. Expression value, base mean over the row counts overlapping a given gene; log2FC, log2 of fold-change between +/+ and *Nan*/+ mice; FDR, BH adjusted p-values.

Gene Symbol	Gene ID	Expression value	log2FC	FDR	Description	Chromosome
Pklr	ENSMUSG00000041237	8831	-1.19	1.33E-49	pyruvate kinase liver and red blood cell	3
E2f2	ENSMUSG00000018983	4959	-2.01	1.98E-10	E2F transcription factor 2	4
Epb4.9	ENSMUSG00000022099	4760	-2.33	1.70E-12	erythrocyte protein band 4.9	14
Bcl11a	ENSMUSG00000000861	1500	-0.83	2.71E-35	B cell CLL/lymphoma 11A (zinc finger protein)	11
Xpo7	**ENSMUSG00000022100**	**10033**	**-1.69**	**3.75E-05**	**exportin 7**	**14**
Odc1	ENSMUSG00000011179	9629	-1.01	1.32E-39	ornithine decarboxylase, structural 1	12
Ubac1	ENSMUSG00000036352	5338	-1.64	3.30E-17	ubiquitin associated domain containing 1	2
E2f4	ENSMUSG00000014859	4989	-1.09	1.69E-29	E2F transcription factor 4	8
Slc43a1	ENSMUSG00000027075	3763	-1.15	4.16E-31	solute carrier family 43, member 1	2
Rad23a	ENSMUSG00000003813	2876	-1.13	1.44E-11	RAD23a homolog (S. cerevisiae)	8
Nxpe2	ENSMUSG00000032028	2579	-2.80	1.83E-08	neurexophilin and PC-esterase domain family, member 2	9
Mgst3	ENSMUSG00000026688	2200	-1.70	4.46E-21	microsomal glutathione S-transferase 3	1
Jhdm1d	ENSMUSG00000042599	2144	-1.02	2.91E-11	jumonji C domain-containing histone demethylase 1 homolog D (S. cerevisiae)	6
Ift140	ENSMUSG00000024169	2065	-1.55	4.01E-06	intraflagellar transport 140	17
Slc30a10	ENSMUSG00000026614	1928	-1.08	2.79E-42	solute carrier family 30, member 10	1
Mpp1	ENSMUSG00000031402	1860	-1.53	2.26E-09	membrane protein, palmitoylated	X
Cldn13	ENSMUSG00000008843	1852	-1.09	1.05E-21	claudin 13	5
Bsdc1	**ENSMUSG00000040859**	**1802**	**-1.30**	**8.18E-06**	**BSD domain containing 1**	**4**
Acp5	ENSMUSG00000001348	1774	-1.66	5.97E-17	acid phosphatase 5, tartrate resistant	9
Slc25a38	ENSMUSG00000032519	1682	-1.34	3.09E-16	solute carrier family 25, member 38	9
Fam53b	ENSMUSG00000030956	1485	-1.04	4.18E-19	family with sequence similarity 53, member B	7
Sppl2b	ENSMUSG00000035206	1386	-1.12	5.01E-31	signal peptide peptidase like 2B	10
Camsap2	ENSMUSG00000041570	1385	-1.07	8.34E-34	calmodulin regulated spectrin-associated protein family, member 2	1
Cast	ENSMUSG00000021585	1161	-1.09	1.42E-40	calpastatin	13
Ampd3	ENSMUSG00000005686	1065	-1.40	8.33E-19	adenosine monophosphate deaminase 3	7
Rgcc	ENSMUSG00000022018	1026	-2.01	1.93E-09	regulator of cell cycle	14
Fn3krp	ENSMUSG00000039253	1004	-1.51	2.88E-15	fructosamine 3 kinase related protein	11

```
human      MAEGEDVGWWRSWLQQSYQAVKEKSSEALEFMKRDLTEFTQVVQHDTACTIAATASVVKE 60
mouse      MAEGEDVGWWRSWLQQSYQAVKEKSTEALEFMKRDLTEFTQVVQHDTACTIAATASVVKE 60
xenopus    MAEGEDGSWWRSWLQQSYTSVRDKSAETLEFMKRDLTEFSRVVHHDTACTIAATASVVKE 60
zebrafish  MADGE-GGWWGGWLTQSFQSVKDKSAEAYEFIKRDLTEFSSVVQHDTACSVVATANAIKS 59
           **:**  .** .** **: :*::**:*: **:*******: **:*****::.***..:*.

human      KLATEGSSGATEKMKKGLSDFLGVISDTFAPSPDKTIDCDVITLMGTPSGTAEPYDGTKA 120
mouse      KLATEGSSGATEKVKKGLSDFLGVISDTFAPSPDKTIDCDVITLMGTPSGTAEPYDGTKA 120
xenopus    KLVVEGSSGTTEKVKKGLSDFLGVISDTFAPSPDKTIDCDVITLMATPSGTTELYDGTKA 120
zebrafish  KLAVEGSSETTEKVKKGITNILGVITDTLAPPPDKTIDCDVITLVATPTGTTEVYDSGKA 119
           **..**** :***:***::::****:**:**.***********:.**:**:* **. **

human      RLYSLQSDPATYCNEPDGPPELFDAWLSQFCLEEKKGEISELLVGSPSIRALYTKMVPAA 180
mouse      RLYSLQSDPATYCNEPDGPPELFDAWLSEFCLEEKKGEISELLVGSPSIRALYTKMVPAA 180
xenopus    RLYSLQSDPATYCNEPDGFPAEFDAWLAYWDPEQRKAEISELLVTSPSIRALFTKMVPAA 180
zebrafish  RLYSLQADPATYCNEPDGPPQQFDSWLSTFSLEERKAEISELLVSSPAIRALYTKMVPAA 179
           ******:********** *  **:**: :  *::*.******* **:****:*******

human      VSHSEFWHRYFYKVHQLEQEQARRDALKQRAEQSIS-EEPGWEEEEEE--LMGISPISPK 237
mouse      VSHSEFWHRYFYKVHQLEQEQARRDALKQRADQSIS-EEPGWEEEEEE--LEGIVP-SPK 236
xenopus    VSHSEFWQRYFYKVHQLEQEEARRDALKQRADQSVHSEEPTWEEEEED--FVGAGSAPAL 238
zebrafish  VAHAEFWQRYFYKVFQLEQEEARRVALKQRAEQTDHSESLGWEEEDEEGEFLGATSLSRL 239
           *:*:***:******.*****:*** ******:*:   *.  ****:*:  : *  . .

human      EAKVP--VAKISTFPEGEPGPQSPCEENLVTSVEPPAEVTPSESSESISLVTQIANPATA 295
mouse      EAKIPK-ETKTTTSPEDEPAPQSPCEE---TPVEPPAEATPSESSESISLVTQVANPAAA 292
xenopus    KLEEKY-VLSTPTIP--TLHVEDKSEKTAELNRDHTSFTSPSESSESISPITQIANPEYI 295
zebrafish  DFTPPVEESRVPAVTVADTPESSSPPQAVASLSNVASDVTPSVSSDSVSFPTQIENQAES 299
           .          .: .      ..   :      : .: .:** **:*:*  **: *

human      PEARVLPKDLSQKLLEASLEEQGLAVDVGETGPSPPIHSKPLTPAGHTGGPEPRPPARVE 355
mouse      PEAPELPKDLSQKLFEASLEEQSLAEDEGETGPPPPPPSKPLTPAGRASGPEPRPPARVE 352
xenopus    EQTP--SKEPSPGTLTVTKENTGAGTD--ETSAPAPLEQK----TGKSN------TQMA 340
zebrafish  VTIRVTQPSPAADELSQKLSDTGLQETFPEERPAP----------------------REE 337
                  . :   :  . .: .      *  ...                    :

human      TLREEAPTDLRVFELNSDSGKSTPSNNGKKGSSTDISEDWEKDFDLDMTEEEVQMALSKV 415
mouse      TLREEVPTDLRVFELNSDSGKSTPSNNGKKGSSTDISEDWEKDFDLDMTEEEVQMALSKV 412
xenopus    TQREDPPSDLRVFELNSDSGKSTPSNNGQKGSSTDISEDWEKEF--DMTEEEVQLALSTV 398
zebrafish  TAREDMAQDLRVFELNSDSGKSTPSNNGKKGSSTDVSEDWEKDFDLDMTEEEVQMALSKV 397
           * **: . ********************:******:******:*   ********:***.*

human      DASGELEDVEWEDWE 430
mouse      DASGELEDVEWEDWE 427
xenopus    EVSGEVEDEDWENWE 413
zebrafish  DATEELEDEDWENWE 412
           :.: *:** :**:**
```

Figure 1. BSDC1 sequence. ClustalW aligment of BSDC1 protein from four different species: Homo sapiens (human), Mus musculus (mouse), Xenopus laevis (Xenopus), Danio rerio (zebrafish). The BSD domain is highlighted in yellow. Serine 414 is indicated by a red box. The conservation of the BSD domain between species is 98% between mouse and human, 79% between mouse and Xenopus and 81% between mouse and zebrafish. Asterisk (*) indicates positions which have a single, fully conserved residue; colon (:) indicates conservation between groups of strongly similar properties - scoring > 0.5 in the Gonnet PAM 250 matrix; period (.) indicates conservation between groups of weakly similar properties - scoring =< 0.5 in the Gonnet PAM 250 matrix.

BSDC1 expression is dependent on KLF1

To confirm BSDC1 downregulation in *Nan*/+ mice, we checked *Bsdc1* expression by RT-qPCR in E12.5 mouse fetal liver cells. Indeed, KLF1 *Nan*/+ mice showed a 60% decrease in the expression of *Bsdc1* mRNA (Figure 2A). We also detected lower expression at protein level by Western blot analysis of control and KLF1 *Nan*/+ fetal liver extracts (Figure 2B). To further investigate whether BSDC1 expression is indeed dependent on the presence of functional KLF1, we checked BSDC1 levels in KLF1 heterozygotes (*Klf1* -/+) and KLF1 knock out (*Klf1* -/-) E13.5 embryos. Both RT-qPCR and Western blot analysis showed that BSDC1 is significantly downregulated in *Klf1* -/- embryos (Figure 2C and 2D). Our results, in combination with the published ChIP experiment[49], strongly support the notion that BSDC1 is a target of KLF1 in erythroid cells.

BSDC1 is upregulated during erythroid differentiation

As most KLF1 target genes are upregulated during erythroid differentiation, we measured BSDC1 expression in maturing cells. To this aim, we FACS-sorted E12.5 mouse fetal liver cells, based on the erythroid markers CD71 and Ter119. Four populations representing progressive differentiation were selected: P1 (CD71^{hi}), P2 (CD71^{hi} Ter119^{hi}), P3(CD71^{lo} Ter119^{hi}) and P4 (Ter119^{hi}) (Figure 3A). We then extracted RNA from these four different populations and used it to synthesize cDNA. Unfortunately we were not able to retrieve enough RNA from population P4, probably due to the limited number of sorted cells. By RT-qPCR analysis of *Bsdc1* transcripts we detected up-regulation of *Bsdc1* expression as maturation proceeded (Figure 3B), confirming the already published data in mouse[50] and in human (GEO Profile ID:

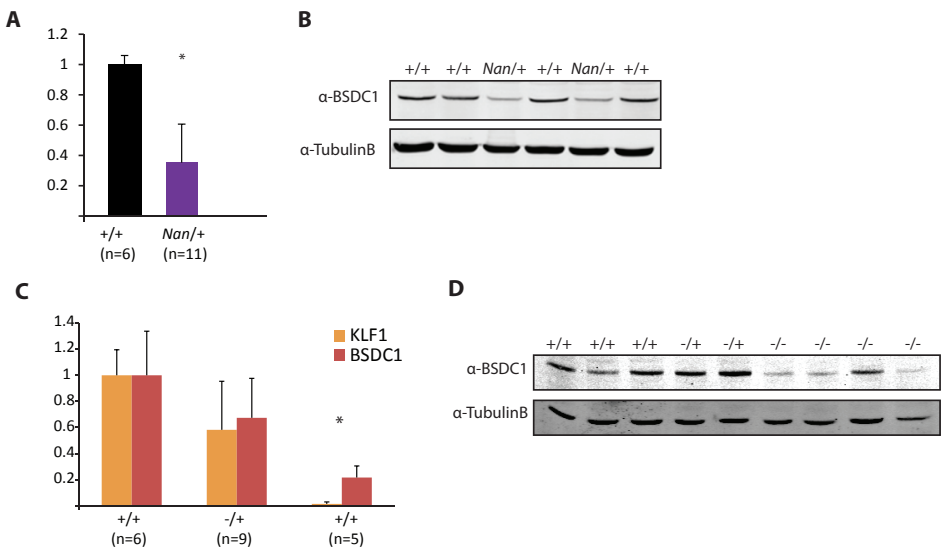


Figure 2. BSDC1 expression in KLF1 *Nan*/+ and KLF1 -/- mouse fetal livers. (A) qPCR for BSDC1 performed on E12.5 KLF1 +/+ or KLF1 *Nan*/+ mouse fetal livers. * indicates a p-value<0.01. n indicates the number of embryos used for each genotype. The values were normalized against actin. (B) BSDC1 protein levels in E12.5 KLF1 +/+ or KLF1 *Nan*/+ mouse fetal livers. Tubulin was used as loading control. (C) qPCR for BSDC1 and KLF1 performed on E13.5 KLF1 +/+ or KLF1 +/- or KLF1 -/- mouse fetal livers. * indicates a p-value<0.01. n indicates the number of embryos used for each genotype. The values were normalized against actin. (D) BSDC1 protein levels in E13.5 KLF1 +/+ or KLF1 +/- or KLF1 -/- mouse fetal livers. Tubulin was used as loading control. BSDC1 is downregulated at mRNA and protein levels both in KLF1 *Nan*/+ and in KLF1 -/- mouse fetal livers.

70509590, Supplementary Figure 1). The mRNA expression levels were normalized against *Actin* (Figure 3B), as well as the house-keeping genes *Gapdh* and *G6PD* (similar results, data not shown). To follow the protein level during maturation, we differentiated the factor dependent immortalized mouse erythroid cell line (I/11) and measured BSDC1 expression by Western blotting. We observed that BSDC1 protein levels increase during differentiation (Figure 3C).

BSDC1 is a cytoplasmic protein

Since the BSD domain is present in several chromatin-bound proteins, we tested whether BSDC1 could also be involved in DNA binding, by checking the localization of the protein. We biochemically fractionated proliferating or differentiating mouse E12.5 mouse fetal liver cells in 4 different subcellular fractions and 2 washes[43]. Surprisingly, we were not able to detect any BSDC1 in the nucleoplasm or bound to the chromatin (Figure 3D). BSDC1 appeared to be exclusively cytoplasmic, in both proliferating and differentiating cells (Figure 3D). Additional support for cytoplasmic localization of BSDC1 was obtained from immunostainings of E16.5 mouse embryos. We detected cytoplasmic BSDC1 expression in several tissues, including gut, salivary gland, pancreas, and in the developing tooth (Figure 4). As the positive cells of these tissues are specialized in secretion, we reasoned that BSDC1 might be involved in that process. In adult human tissues, BSDC1 is widely, although not ubiquitously, expressed[51,52].

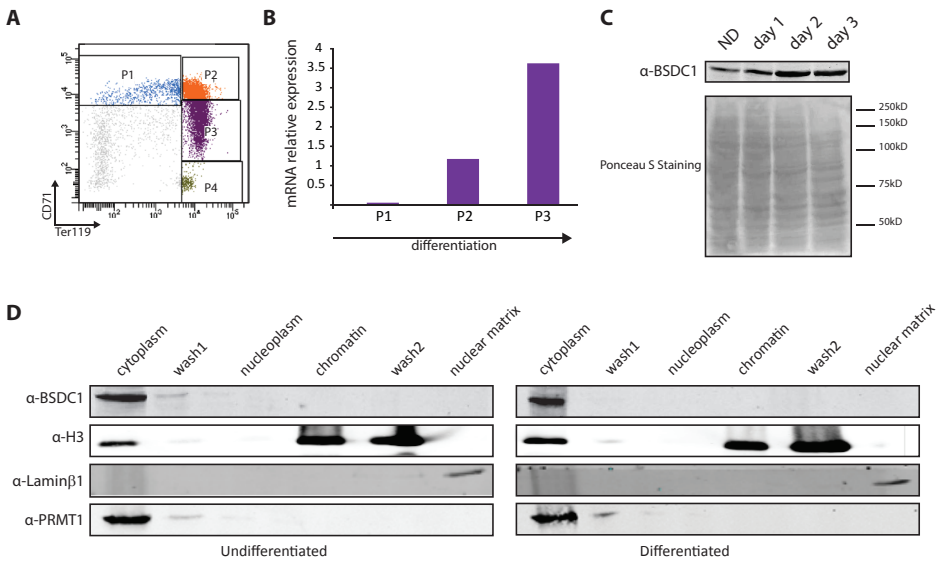


Figure 3. BSDC1 expression during differentiation and cellular localization. (A) FACS dot-plot showing the sorting strategy for E12.5 mouse fetal liver using CD71 and Ter119 antibodies. P1 represents the CD71^{hi} population, P2 the CD71^{hi} Ter119^{hi} population, P3 the CD71^{lo} Ter119^{hi} population adn P4 the Ter119^{hi} population. (B) qPCR showing BSDC1 mRNA expression levels, normalized against actin, in P1, P2 and P3populations. n=1 (C) Western blot of BSDC1 expression during 3 days of I/11 differentiation. Ponceau S staining is shown as loading control.(D) BSDC1 subcellular localization in proliferating and in differentiating cells. The cells were fractionated in 4 different fractions and 2 washes: cytoplasm,wash1, nucleoplasm, chromatin, wash2 , nuclear matrix. The blot was stained with BSDC1 antibody. Histone H3, LaminB and PRMT1 were used as control to check for proper subcellular fractionation. BSDC1 is localized exclusively in the cytoplasm and its localization does not change upon differentiation.

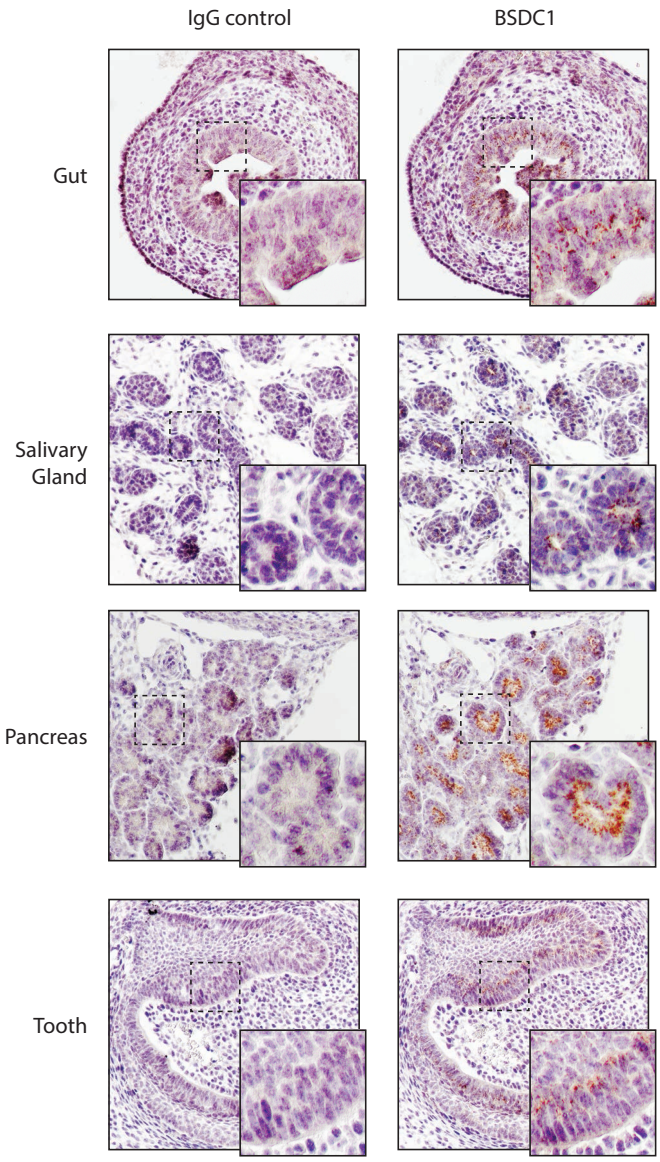


Figure 4. BSDC1 Immunostaining. Sagittal paraffin sections of E16.5 mouse embryos were incubated with rabbit IgG control (left) and BSDC1 (right) antibodies, followed by peroxidase staining. The dotted squared indicated the zoomed-in area. The BSDC1 signal in the gut, in the salivary gland, in the pancreas and in the tooth tissues is shown.

Therefore, it is very well possible that BSDC1 in mouse is expressed in other tissues below the detection level of immunostaining. Alternatively, BSDC1 expression may reach detectable levels in certain tissues at later stages of development. These observations collectively suggest that BSDC1 is a cytoplasmic protein that is relatively highly expressed in cells with a secretory function.

BSDC1 mass spectrometry analysis reveals a possible role in the Golgi apparatus.
In order to obtain more insight in the molecular role of BSDC1, we set up a strategy to identify possible interacting proteins. Tagged mouse BSDC1 (Bio-HA-BSDC1) was generated by fusing an HA epitope and a Bio-tag to its N-terminus (Figure 5A). The Bio-tag is efficiently

biotinylated by the bacterial BirA biotin ligase, which is co-expressed in stably transfected mouse erythroleukemia (MEL) cells[41]. Bio-HA-BSDC1 was expressed below endogenous levels to reduce the likelihood that non-physiological interactions would be identified (data not shown). Biotinylated BSDC1 was efficiently recovered from MEL cytoplasmic extracts and whole cell lysates with magnetic streptavidin beads (Figure 5B). BSDC1-associating proteins were identified by streptavidin pull down followed by nanoflow liquid chromatography-tandem mass spectrometry (nanoLC-MS/MS) and compared to samples from cells expressing BirA alone. Putative BSDC1-interacting proteins present in both cytoplasmic and whole cell lysates are listed in Supplementary Table 1. Along with BSDC1, we identified GOLGB1/GIANTIN and CASP, two proteins localized at the rim of the Golgi, and FUT-8, a fucosyltransferase resident in the Golgi apparatus (Table 2). In addition to these three proteins, we identified ERMAP and EMILIN2 (Table 2). ERMAP stands for Erythroblast Membrane-Associated Protein and polymorphisms in this gene are responsible for the Scianna/Radin blood group system in humans[53,54]. EMILIN2[55] (Elastin Microfibril Interfacer 2) is a an extracellular matrix glycoprotein that forms elastic fibers and is probably involved

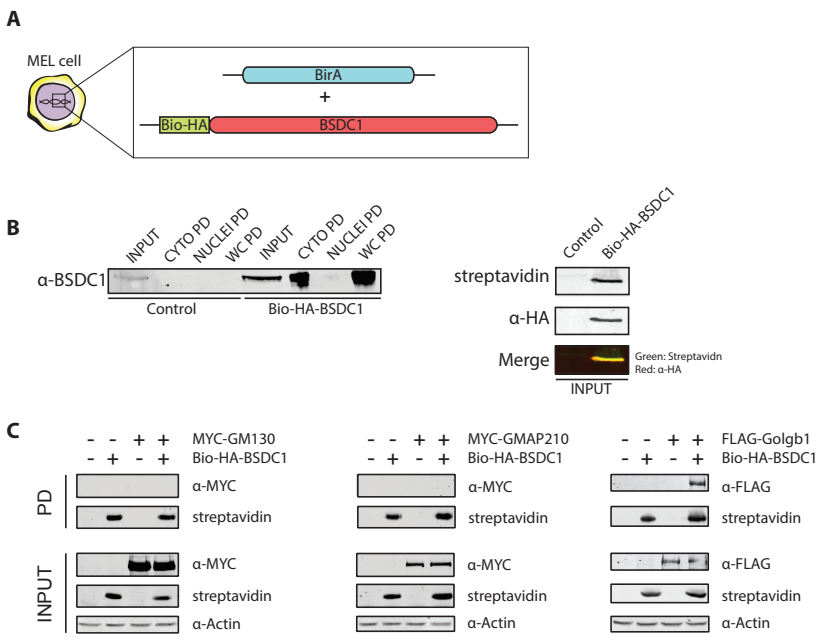


Figure 5. Bio-HA-BSDC1 streptavidin pull down. (A) schematic representation of the biotin ligase BirA and the Bio-HA-BSDC1 constructs. These constructs were electroporated MEL cells to produce a stable cell line. (B) Streptavidin pull-down of biotinylated-BSDC1 in MEL cells stably expressing Bio-HA-BSDC1. MEL cells only expressing the BirA construct were used as control. The membrane was stained with α-BSDC1 antibody, α-HA antibody (red) and streptavidin (green). Two different protein purification methods were used: cytoplasm and nuclear extracts and whole cell extracts. CYTO PD, cytoplasm pull-down; NUCLEI PD, nuclear pull-down; WC PD, whole cell pull down. (C) Streptavidin pull-down of biotinylated-BSDC1 in 293T cells co-transfected with Bio-HA-BSDC1 and Myc-tagged GM130/GMAP210 (controls) or FLAG-tagged GOLGB1. The membranes were stained with α-Myc and α-FLAG antibodies or streptavidin. α-actin staining was used as loading control. PD, pull-down. In co-transfected 293T cells, GOLGB1 but not GM130 or GMAP210 co-precipitate with Bio-HA-BSDC1.

Table 2. MS results. Table showing the number of peptides and the Mascot scores of BSDC1 and five selected proteins observed in the BSDC1 pull-down followed by MS analysis . Three proteins (GOLGB1, CUX1 and FUT-8) belong to the Golgi apparatus and two proteins (EMILIN2 and ERMAP) transit through the Golgi and therefore might represent cargo. The experiment was done in duplicate with similar results. The BirA control gave no unique peptides corresponding to these proteins. Cyto, cytoplasmic fraction; WC, whole cell lysate.

Unique peptides					mascot score	
Cyto	WC	symbol	description	length	Cyto	WC
34	**37**	**Bsdc1**	BSD domain-containing protein 1	427	2397	2746
120	**222**	Golgb1	Protein Golgb1	3238	5485	12205
15	**20**	Cux1	Homeobox protein cut-like, CASP protein	1504	556	916
9	**1**	Fut8	Alpha-(1,6)-fucosyltransferase	575	371	43
20	**19**	Emilin2	EMILIN-2	1074	910	832
17	**25**	Ermap	Erythroid membrane-associated protein	566	923	1393

in cell adhesion[56,57]. GOLGB1 was identified as the top hit of our experiments, while the other proteins were present in the top 15 potential binding partners of BSDC1. This result indicates that BSDC1 is most likely part of the Golgi apparatus, which fits with its cytoplasmic localization. We then validated the BSDC1 and GOLGB1 interaction by co-transfection experiments in 293T cells. We co-transfected Bio-HA-BSDC1 with FLAG tagged GOLGB1. Two GOLGB1 related proteins, the Golgi tethering factors GM130 and GMAP210, were also used to test the specificity. As shown by Western blot, BSDC1 specifically interacts with GOLGB1 but not with GM130 and GMAP120 (Figure 5C), validating the MS results. These preliminary results indicated that BSDC1 resides at specific domains within the Golgi apparatus.

BSDC1 co-localizes with GOLGB1

To study the potential Golgi localization of BSDC1 in more detail, we performed immunofluorescence experiments in different cell lines. First, we looked at the localization of BSDC1, GOLGB1 (localized at the Golgi stack), and GM130 (localized at the *cis*-Golgi) in the erythroleukemia cell lines MEL and K562. BSDC1 was indeed found to localize to the Golgi in structures next to, or partially overlapping with, the GM130 positive areas, a pattern very similar to the GOLGB1 staining (Supplementary Figure 2). However, the Golgi appeared to be small and compact in these spherical cells, most likely due to their small cytoplasmic volume. As a result, the resolution of the Golgi apparatus is insufficient. We therefore performed similar experiments in U2OS osteosarcoma cells, which are flat and larger. In these cells, GOLGB1 appears as 'open', circular structures, while GM130 localization is more granular (Figure 6A). The merged image shows that most GM130 signal is located within the GOLGB1 'circles'. Staining for BSDC1 revealed circular patterns around the GM130 'granules', suggesting that BSDC1 and GOLGB1 co-localize. A direct co-staining for these proteins was not possible, as the antibodies used were both raised in rabbits. We therefore stably transfected U2OS cells with a BSDC1-GFP fusion construct. GFP signal was identical to endogenous BSDC1 (Figure 6B), indicating that the GFP fusion protein had no aberrant localization. The merge of the GFP and GOLGB1 signals shows an almost perfect overlap, clearly demonstrating the co-localization of the two proteins. In conclusion, these experiments showed that BSDC1 is localized in the Golgi stack, together with GOLGB1.

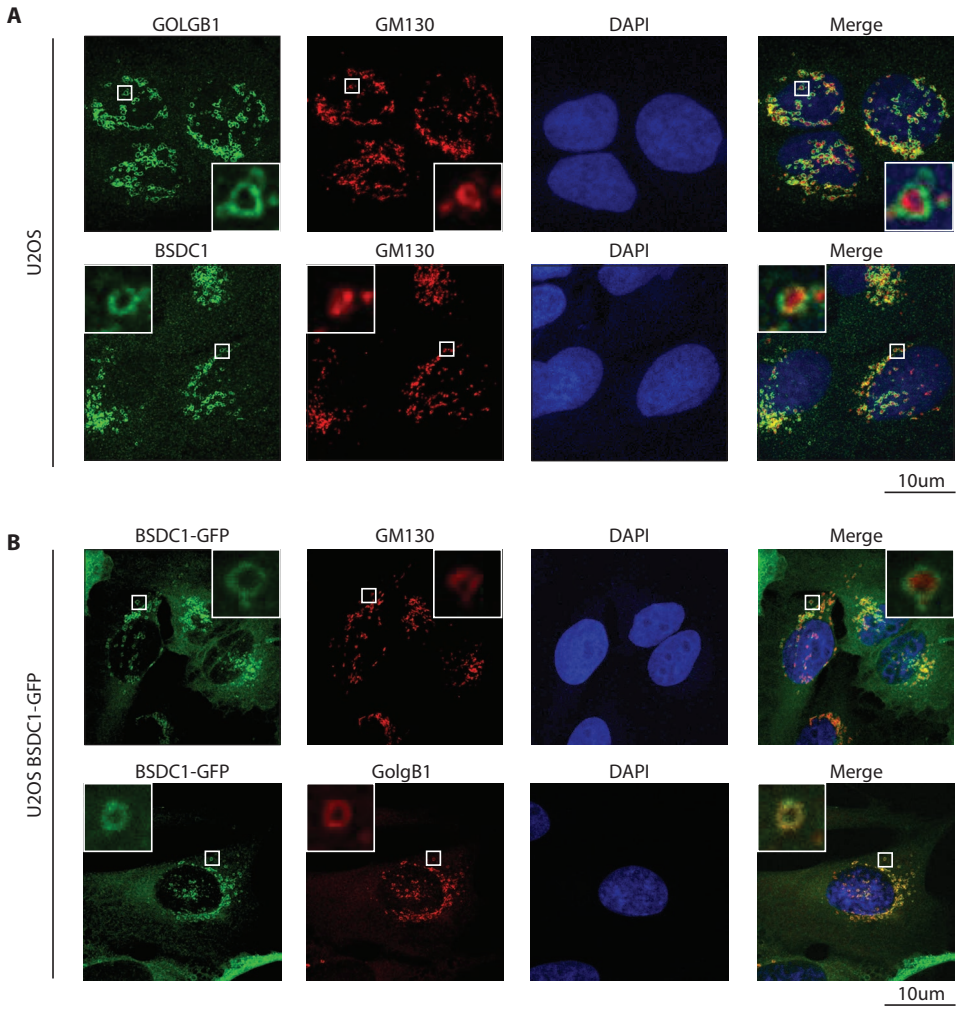


Figure 6. BSDC1 subcellular localization. (A) Confocal images of U2OS cells stained with BSDC1 or GOLGB1 together with GM130 antibodies. These images indicate that BSDC1 localizes in a perinuclear area corresponding to the Golgi apparatus and not with GM130. GOLGB1 and GM130 also do not co-localize. Scale bars, 10 µm. (B) Confocal images of U2OS cells stably expressing BSDC1-GFP fusion protein stained with GOLGB1 or GM130 antibodies. Scale bars, 10 µm. Localization of the fusion protein localization is similar to that of endogenous BSDC1 protein. BSDC1-GFP co-localizes with GOLGB1.

ERMAP localization is not affected by the knockdown of BSDC1

In our MS results we also identified ERMAP (erythroblast membrane-associated)[58]. *Ermap* gene encodes for a cell surface transmembrane protein, that may act as a mediator of adhesion in erythroid cells[53]. Polymorphisms in this gene are responsible for the Scianna/Radin blood group[54]. As a membrane protein, ERMAP needs to transit through the Golgi to its final destination. We hence reasoned that ERMAP localization might be affected by the disruption of BSDC1 protein expression. In order to test this hypothesis, we used the K562 human erythroleukemia cell line, which expresses detectable amounts of ERMAP protein on the cell surface. We tried to knockdown BSDC1 in this cell line using 4 different shRNAs

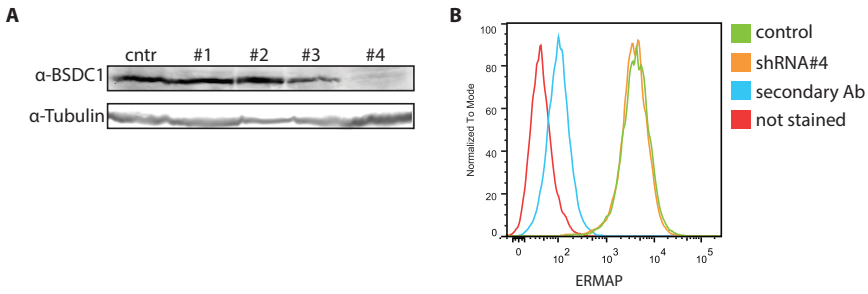


Figure 7. ERMAP membrane localization. (A) Western blot analysis of K562 cells transduced with four different lentiviral shRNAs targeting BSDC1 (#1-#4) and with a control vector. The membrane was stained with a BSDC1 antibody. A tubulinβ antibody was used as a loading control. We detected efficient BSDC1 knockdown with shRNA #4. (B) Flow cytometry analysis of ERMAP expression in K562 cells transduced with control or shRNA #4 vectors. No difference was detected between control (green line) and BSDC1 knock down (orange line) cells.

(#1-#4), obtaining a satisfactory reduction of BSDC1 protein with shRNA #4 (Figure 7A). We used an empty vector as control. We then checked by flow cytometry if ERMAP was still present on the membrane (Figure 6B). We observed that ERMAP localization was not affected by knockdown of BSDC1, as the signals were similar in the control and the shRNA#4 cells. This indicates that ERMAP transit through the Golgi and subsequent transport to the cell membrane is not affected upon BSDC1 knockdown.

DISCUSSION

The experiments described here are the first attempts to elucidate the function of BSDC1, a Golgi protein identified in the search for novel KLF1 target genes (Chapter 3). BSDC1 was first described in a screen for BSD domain-containing proteins[48], but it has never been characterized functionally. We decided to study this protein for several reasons:(1) BSDC1 was identified in the KLF1 *Nan* RNA-sequencing experiment as a downregulated gene in the mutants, hence it might represent a direct KLF1 target, (2) KLF1 has been shown to bind to the promoter region of this gene[49], (3) it showed relatively high expression values in erythroid cells (Table 1) and its expression increases during erythroid maturation (mouse[50], human GEO Profile ID: 70509590, Supplementary Figure 1), (4) since the BSD domain is present in DNA-binding proteins[34], BSDC1 could represent a chromatin associated factor, (5) BSDC1 is highly conserved among species (Figure 1). Here we show that erythroid BSDC1 expression is impaired when KLF1 has an aberrant function (*Klf1 Nan* mutant mice) or in the absence of any KLF1 activity (*Klf1* -/- mice, Figure 2), and that BSDC1 is up-regulated during erythroid maturation, indicating that it might have a role during red blood cell differentiation (Figure 3).

Cellular fractionation assays and immunohistochemistry revealed that BSDC1 is not a nuclear protein, in contrast to our assumption. Instead, we found that BSDC1 is exclusively located in the cytoplasm, a localization pattern that does not change during differentiation of erythroid cells (Figure 3). In order to obtain insight in BSDC1 function, we sought to identify its interactors. We pulled-down biotinylated-BSDC1 and identified co-purifying proteins by mass spectrometry. Among the potential BSDC1 binding partners we identified GOLGB1, a Golgi-specific protein involved in tethering processes in the Golgi apparatus, as the top

hit (Table 2). This interaction was confirmed by co-transfection experiments in 293T cells (Figure 4). We further demonstrated by immunofluorescence experiments that BSDC1 co-localizes at the Golgi apparatus with GOLGB1 (Figure 6). It would be interesting to monitor BSDC1 localization upon GOLGB1 depletion, in order to check whether its localization is dependent on the presence of GOLGB1.

Irrespectively of whether BSDC1 affects the structure/positioning of the Golgi apparatus or the efficient modification of target proteins, further experiments would provide more insight on how it operates. The N-terminal domains of GOLGB1 and GM130 interact with the C-terminal acidic domain of p115[8,59,60]. p115 is a Golgi membrane protein[61] that mainly associates with *cis*-Golgi elements and the intermediate compartment residing between the Golgi and the ER[60,62]. These interactions are phosphorylation dependent: when GM130 and/or p115 are phosphorylated, the interaction is lost[63,64]. Since BSDC1 also has an acidic domain at its C-terminus (Figure 1) we reasoned that BSDC1 might bind to GOLGB1 via this domain, possibly competing with p115. p115 is phosphorylated at serine 942 in its acidic domain[64] and when this residue is phosphorylated, p115 is found predominantly in the cytoplasm and not associated to the Golgi. BSDC1 also possesses a serine in its acidic domain (residue 415 in mouse, indicated by the red box in Figure 1) that is predicted to be phosphorylated[65]. To investigate whether the BSDC1/ GOLGB1 interaction is dependent on the acidic domain and its potential phosphorylation, we will generate 3 BSDC1 mutants: one deletion mutant, inserting a stop codon in codon 414, a point mutant (S415A) that would abolish phosphorylation and a point mutant (S415E) that would mimic phosphorylation.

Our work indicates that BSDC1 is the first BSD-domain containing protein that does not fall in the two categories of BSD-domain containing proteins described: the chromatin binding proteins and synapse associated proteins[48]. It is interesting to notice that synapses are cell structures that also have high vesicular trafficking. Hence, the BSD domain might have a role in trafficking both in the synapses and in the Golgi apparatus. However, the function of BSDC1, and that of the BSD domain in general, remains unclear. To further characterize BSDC1 behavior in the cell, we aim to check how BSDC1 responds upon the disruption of normal Golgi structure. To this end, the use of the fungal metabolite brefeldin A (BFA) and of nocodazole would be useful. The treatment of cells with BFA leads to a time-dependent disruption of the Golgi apparatus and eventually to a complete loss of Golgi structure and function. Nocodazole causes microtubule depolymerization and subsequent fragmentation of the Golgi. However, in contrast to BFA, the Golgi apparatus remains functional. It was previously shown that upon BFA treatment or in mitotic cells, the localization of erythroid 3-spectrin and FIP-2 proteins in the Golgi apparatus is disrupted, while Nocodazole treatment does not have this effect[26,27]. It would be interesting to see how BSDC1 behaves upon BFA and Nocodazole treatment. Our initial results suggest that in mitotic cells, when the Golgi structure needs to be disrupted for the cell to divide, BSDC1 relocalizes from its original position (not shown), probably being dispersed. Another possibility is that BSDC1 protein is degraded upon Golgi disruption. In order to distinguish between these two scenarios, we would have to check for BSDC1 protein levels by Western blot analysis in cells treated with BFA or Nocodazole. It is worth keeping in mind that BSDC1 is upregulated during erythroblast differentiation, while the Golgi apparatus is disassembled. A possible explanation for this upregulation might come from BSDC1 relocalization to a different cellular compartment. Such relocalization events have been previously reported for FIP-2, a chicken protein likely involved in the TNF-α signaling pathway, which relocalizes from the Golgi to the marginal band upon erythroid differentiation[26].

At present, the function of BSDC1 remains unknown. We reasoned that it could play a role in the transport of vesicles through the Golgi apparatus. However, knockdown of BSDC1 does

not affect the localization of ERMAP protein on the plasma membrane (Figure 7). ERMAP is a membrane protein that travels through the Golgi in order to be glycosylated and sorted to its final destination. The fact that BSDC1 knockdown does not affect ERMAP localization could be due to two different scenarios. One possibility is that BSDC1 knockdown alone is not sufficient to alter the structure and function of the Golgi apparatus. Golgins have been shown to play redundant roles, and the function of BSDC1 could be redundant as well. To address this question, we could combine BSDC1 and GOLGB1 knockouts, preferably using CRISPR technology. If this strategy does not significantly affect Golgi structure/function, it might be necessary to couple BSDC1 knockout to the knockout of other golgins. The second possibility is that BSDC1 knockdown does not affect ERMAP transport but affects its posttranslational modifications. In this case, it would be important to study the glycosylation state of the protein upon BSDC1 knockdown. To this end, we could assay the ERMAP glycosylation status by Western blot analysis: glycosylated proteins are expected to migrate as diffuse bands. If BSDC1 knockdown indeed affects ERMAP glycosylation, the band width and the apparent molecular weight should decrease.
In this work we showed that BSDC1 is a novel Golgi protein that constitutes a target of the erythroid-specific transcription factor KLF1. It therefore might play a role in erythropoiesis. Given the complex nature of erythropoiesis, which involves changes in cell size, high production of globins and eventually extrusion of the nuclei, it is surprising that the role of the Golgi apparatus during red blood cell maturation has been understudied. The identification of a connection between KLF1 and BSDC1 raises the possibility that the Golgi apparatus might be significant for red blood cell development and therefore should be the focus of future studies.

Author contributions

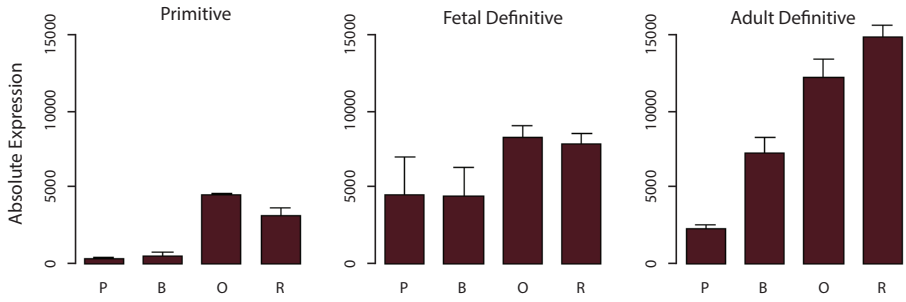
I.C., T.v.D. and S.P. designed the experiments. I.C., S.B. and T.v.D. performed the experiments. M.L provided the giantins vectors. J.D. performed the MS-analysis. The paper was written by I.C., T.v.D. and S.P.

REFERENCES

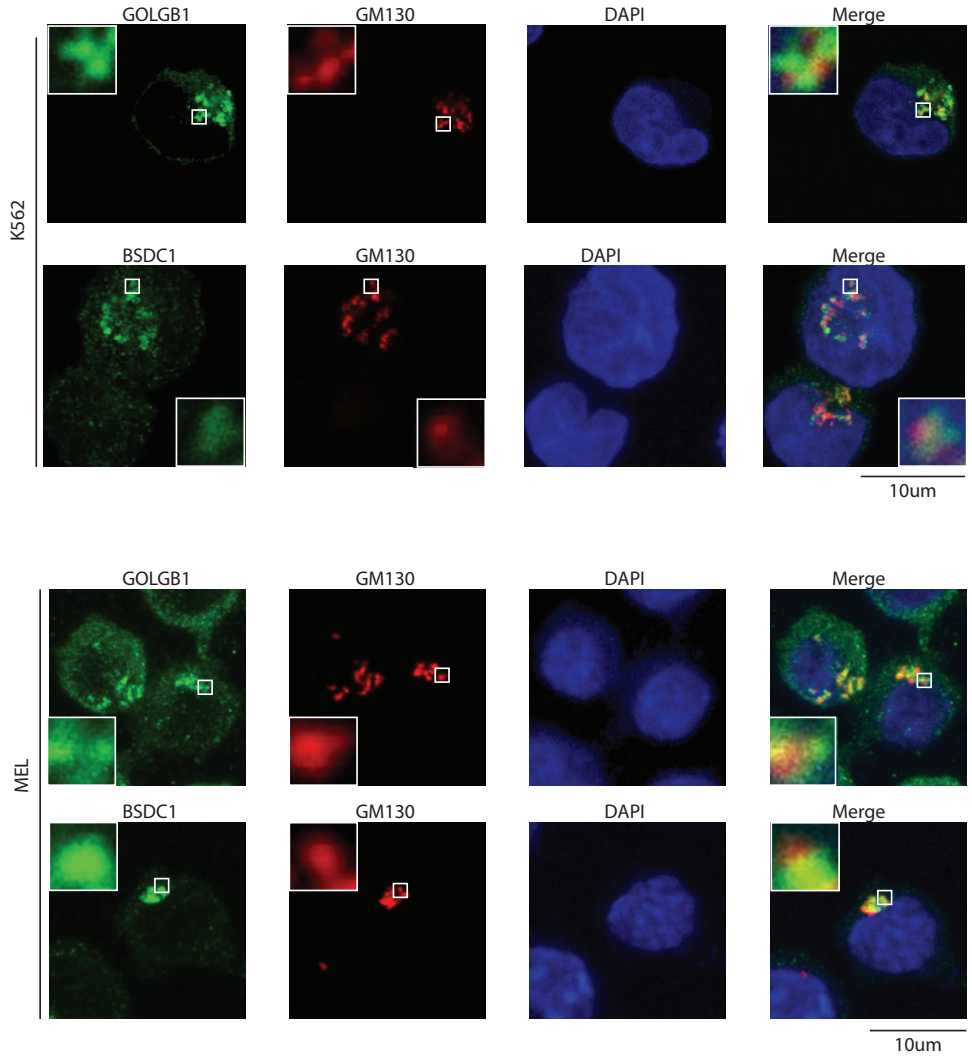
1. Lowe, M. Structural organization of the Golgi apparatus. *Curr Opin Cell Biol* **23**, 85-93 (2011).
2. Alberts, B. Molecular biology of the cell. *5th edition, Garland Science.2008*.
3. Moremen, K.W., Tiemeyer, M. & Nairn, A.V. Vertebrate protein glycosylation: diversity, synthesis and function. *Nat Rev Mol Cell Biol* **13**, 448-62 (2012).
4. Day, K.J., Staehelin, L.A. & Glick, B.S. A three-stage model of Golgi structure and function. *Histochem Cell Biol* **140**, 239-49 (2013).
5. Munro, S. The golgin coiled-coil proteins of the Golgi apparatus. *Cold Spring Harb Perspect Biol* **3**(2011).
6. Nakamura, N. et al. Characterization of a cis-Golgi matrix protein, GM130. *J Cell Biol* **131**, 1715-26 (1995).
7. Sapperstein, S.K., Walter, D.M., Grosvenor, A.R., Heuser, J.E. & Waters, M.G. p115 is a general vesicular transport factor related to the yeast endoplasmic reticulum to Golgi transport factor Uso1p. *Proc Natl Acad Sci U S A* **92**, 522-6 (1995).
8. Nakamura, N., Lowe, M., Levine, T.P., Rabouille, C. & Warren, G. The vesicle docking protein p115 binds GM130, a cis-Golgi matrix protein, in a mitotically regulated manner. *Cell* **89**, 445-55 (1997).
9. Cao, X., Ballew, N. & Barlowe, C. Initial docking of ER-derived vesicles requires Uso1p and Ypt1p but is independent of SNARE proteins. *EMBO J* **17**, 2156-65 (1998).
10. Satoh, A. & Warren, G. In situ cleavage of the acidic domain from the p115 tether inhibits exocytic transport. *Traffic* **9**, 1522-9 (2008).
11. An, Y. et al. Structural and functional analysis of the globular head domain of p115 provides insight into membrane tethering. *J Mol Biol* **391**, 26-41 (2009).
12. Striegl, H., Andrade-Navarro, M.A. & Heinemann, U. Armadillo motifs involved in vesicular transport. *PLoS One* **5**, e8991 (2010).
13. Linstedt, A.D. & Hauri, H.P. Giantin, a novel conserved Golgi membrane protein containing a cytoplasmic domain of at least 350 kDa. *Mol Biol Cell* **4**, 679-93 (1993).
14. Seelig, H.P., Schranz, P., Schroter, H., Wiemann, C. & Renz, M. Macrogolgin--a new 376 kD Golgi complex outer membrane protein as target of antibodies in patients with rheumatic diseases and HIV infections. *J Autoimmun* **7**, 67-91 (1994).
15. Lesa, G.M., Seemann, J., Shorter, J., Vandekerckhove, J. & Warren, G. The amino-terminal domain of the golgi protein giantin interacts directly with the vesicle-tethering protein p115. *J Biol Chem* **275**, 2831-6 (2000).
16. Linstedt, A.D. et al. Binding relationships of membrane tethering components. The giantin N terminus and the GM130 N terminus compete for binding to the p115 C terminus. *J Biol Chem* **275**, 10196-201 (2000).
17. Puthenveedu, M.A. & Linstedt, A.D. Evidence that Golgi structure depends on a p115 activity that is independent of the vesicle tether components giantin and GM130. *J Cell Biol* **155**, 227-38 (2001).
18. Lerer-Goldshtein, T. et al. TMF/ARA160: A key regulator of sperm development. *Dev Biol* **348**, 12-21 (2010).
19. Friggi-Grelin, F., Rabouille, C. & Therond, P. The cis-Golgi Drosophila GMAP has a role in anterograde transport and Golgi organization in vivo, similar to its mammalian ortholog in tissue culture cells. *Eur J Cell Biol* **85**, 1155-66 (2006).
20. Puthenveedu, M.A., Bachert, C., Puri, S., Lanni, F. & Linstedt, A.D. GM130 and GRASP65-dependent lateral cisternal fusion allows uniform Golgi-enzyme distribution. *Nat Cell Biol* **8**, 238-48 (2006).
21. Vasile, E., Perez, T., Nakamura, N. & Krieger, M. Structural integrity of the Golgi is temperature sensitive in conditional-lethal mutants with no detectable GM130. *Traffic* **4**, 254-72 (2003).
22. Yadav, S., Puri, S. & Linstedt, A.D. A primary role for Golgi positioning in directed secretion, cell polarity, and wound healing. *Mol Biol Cell* **20**, 1728-36 (2009).
23. Roboti, P., Sato, K. & Lowe, M. The golgin GMAP-210 is required for efficient membrane trafficking in the early secretory pathway. *J Cell Sci* **128**, 1595-606 (2015).

24. Heynen, M.J. & Verwilghen, R.L. A quantitative ultrastructural study of normal rat erythroblasts and reticulocytes. *Cell Tissue Res* **224**, 397-408 (1982).
25. Ghosh, S., Cox, K.H. & Cox, J.V. Chicken erythroid AE1 anion exchangers associate with the cytoskeleton during recycling to the Golgi. *Mol Biol Cell* **10**, 455-69 (1999).
26. Stroissnigg, H. et al. FIP-2, an IkappaB-kinase-gamma-related protein, is associated with the Golgi apparatus and translocates to the marginal band during chicken erythroblast differentiation. *Exp Cell Res* **278**, 133-45 (2002).
27. Beck, K.A., Buchanan, J.A., Malhotra, V. & Nelson, W.J. Golgi spectrin: identification of an erythroid beta-spectrin homolog associated with the Golgi complex. *J Cell Biol* **127**, 707-23 (1994).
28. Donze, D., Townes, T.M. & Bieker, J.J. Role of erythroid Kruppel-like factor in human gamma- to beta-globin gene switching. *J Biol Chem* **270**, 1955-9 (1995).
29. Drissen, R. et al. The erythroid phenotype of EKLF-null mice: defects in hemoglobin metabolism and membrane stability. *Mol Cell Biol* **25**, 5205-14 (2005).
30. Nilson, D.G., Sabatino, D.E., Bodine, D.M. & Gallagher, P.G. Major erythrocyte membrane protein genes in EKLF-deficient mice. *Exp Hematol* **34**, 705-12 (2006).
31. Pilon, A.M. et al. Failure of terminal erythroid differentiation in EKLF-deficient mice is associated with cell cycle perturbation and reduced expression of E2F2. *Mol Cell Biol* **28**, 7394-401 (2008).
32. Tallack, M.R., Keys, J.R., Humbert, P.O. & Perkins, A.C. EKLF/KLF1 controls cell cycle entry via direct regulation of E2f2. *J Biol Chem* **284**, 20966-74 (2009).
33. Borg, J., Patrinos, G.P., Felice, A.E. & Philipsen, S. Erythroid phenotypes associated with KLF1 mutations. *Haematologica* **96**, 635-8 (2011).
34. Madej, T. et al. MMDB and VAST+: tracking structural similarities between macromolecular complexes. *Nucleic Acids Res* **42**, D297-303 (2014).
35. von Lindern, M. et al. Leukemic transformation of normal murine erythroid progenitors: v- and c-ErbB act through signaling pathways activated by the EpoR and c-Kit in stress erythropoiesis. *Oncogene* **20**, 3651-64 (2001).
36. Needham, M. et al. LCR/MEL: a versatile system for high-level expression of heterologous proteins in erythroid cells. *Nucleic Acids Res* **20**, 997-1003 (1992).
37. Sato, K., Roboti, P., Mironov, A.A. & Lowe, M. Coupling of vesicle tethering and Rab binding is required for in vivo functionality of the golgin GMAP-210. *Mol Biol Cell* **26**, 537-53 (2015).
38. Diao, A., Frost, L., Morohashi, Y. & Lowe, M. Coordination of golgin tethering and SNARE assembly: GM130 binds syntaxin 5 in a p115-regulated manner. *J Biol Chem* **283**, 6957-67 (2008).
39. Lowe, M., Lane, J.D., Woodman, P.G. & Allan, V.J. Caspase-mediated cleavage of syntaxin 5 and giantin accompanies inhibition of secretory traffic during apoptosis. *J Cell Sci* **117**, 1139-50 (2004).
40. van Dijk, T.B. et al. Stem cell factor induces phosphatidylinositol 3'-kinase-dependent Lyn/Tec/Dok-1 complex formation in hematopoietic cells. *Blood* **96**, 3406-13 (2000).
41. de Boer, E. et al. Efficient biotinylation and single-step purification of tagged transcription factors in mammalian cells and transgenic mice. *Proc Natl Acad Sci U S A* **100**, 7480-5 (2003).
42. Zufferey, R., Nagy, D., Mandel, R.J., Naldini, L. & Trono, D. Multiply attenuated lentiviral vector achieves efficient gene delivery in vivo. *Nat Biotechnol* **15**, 871-5 (1997).
43. van Dijk, T.B. et al. Friend of Prmt1, a novel chromatin target of protein arginine methyltransferases. *Mol Cell Biol* **30**, 260-72 (2010).
44. Barsnes, H., Vizcaino, J.A., Eidhammer, I. & Martens, L. PRIDE Converter: making proteomics data-sharing easy. *Nat Biotechnol* **27**, 598-9 (2009).
45. Rajatapiti, P. et al. Expression of glucocorticoid, retinoid, and thyroid hormone receptors during human lung development. *J Clin Endocrinol Metab* **90**, 4309-14 (2005).
46. Siatecka, M. et al. Severe anemia in the Nan mutant mouse caused by sequence-selective disruption of erythroid Kruppel-like factor. *Proc Natl Acad Sci U S A* **107**, 15151-6 (2010).
47. Viprakasit, V. et al. Mutations in Kruppel-like factor 1 cause transfusion-dependent hemolytic anemia and persistence of embryonic globin gene expression. *Blood* **123**, 1586-95 (2014).
48. Doerks, T., Huber, S., Buchner, E. & Bork, P. BSD: a novel domain in transcription factors and synapse-associated proteins. *Trends Biochem Sci* **27**, 168-70 (2002).
49. Su, M.Y. et al. Identification of biologically relevant enhancers in human erythroid cells. *J Biol*

Chem **288**, 8433-44 (2013).
50. Kingsley, P.D. et al. Ontogeny of erythroid gene expression. *Blood* **121**, e5-e13 (2013).
51. Uhlen, M. et al. A human protein atlas for normal and cancer tissues based on antibody proteomics. *Mol Cell Proteomics* **4**, 1920-32 (2005).
52. Uhlen, M. et al. Proteomics. Tissue-based map of the human proteome. *Science* **347**, 1260419 (2015).
53. Su, Y.Y., Gordon, C.T., Ye, T.Z., Perkins, A.C. & Chui, D.H. Human ERMAP: an erythroid adhesion/receptor transmembrane protein. *Blood Cells Mol Dis* **27**, 938-49 (2001).
54. Wagner, F.F., Poole, J. & Flegel, W.A. Scianna antigens including Rd are expressed by ERMAP. *Blood* **101**, 752-7 (2003).
55. Doliana, R. et al. Isolation and characterization of EMILIN-2, a new component of the growing EMILINs family and a member of the EMI domain-containing superfamily. *J Biol Chem* **276**, 12003-11 (2001).
56. Mongiat, M. et al. Regulation of the extrinsic apoptotic pathway by the extracellular matrix glycoprotein EMILIN2. *Mol Cell Biol* **27**, 7176-87 (2007).
57. Marastoni, S. et al. EMILIN2 down-modulates the Wnt signalling pathway and suppresses breast cancer cell growth and migration. *J Pathol* **232**, 391-404 (2014).
58. Xu, H. et al. Cloning and characterization of human erythroid membrane-associated protein, human ERMAP. *Genomics* **76**, 2-4 (2001).
59. Sonnichsen, B. et al. A role for giantin in docking COPI vesicles to Golgi membranes. *J Cell Biol* **140**, 1013-21 (1998).
60. Nelson, D.S. et al. The membrane transport factor TAP/p115 cycles between the Golgi and earlier secretory compartments and contains distinct domains required for its localization and function. *J Cell Biol* **143**, 319-31 (1998).
61. Waters, M.G., Clary, D.O. & Rothman, J.E. A novel 115-kD peripheral membrane protein is required for intercisternal transport in the Golgi stack. *J Cell Biol* **118**, 1015-26 (1992).
62. Bannykh, S.I., Nishimura, N. & Balch, W.E. Getting into the Golgi. *Trends Cell Biol* **8**, 21-5 (1998).
63. Lowe, M. et al. Cdc2 kinase directly phosphorylates the cis-Golgi matrix protein GM130 and is required for Golgi fragmentation in mitosis. *Cell* **94**, 783-93 (1998).
64. Sohda, M., Misumi, Y., Yano, A., Takami, N. & Ikehara, Y. Phosphorylation of the vesicle docking protein p115 regulates its association with the Golgi membrane. *J Biol Chem* **273**, 5385-8 (1998).
65. Olsen, J.V. et al. Quantitative phosphoproteomics reveals widespread full phosphorylation site occupancy during mitosis. *Sci Signal* **3**, ra3 (2010).



Supplementary Figure 1. BSDC1 expression during erythroid development in primitive, fetal definitive, and adult definitive erythropoiesis. P, Proerythroblast ; B, Basophilic Erythroblast ; O, Poly/orthochromatic Erythroblast, R, Reticulocyte. Figure from ref[50].



Supplementary Figure 2. Confocal images of K562 and MEL cells stained with BSDC1 or GOLGB1 (green) antibodies and GM130 (red) antibody. These images indicate that BSDC1 localizes in a perinuclear area corresponding to the Golgi apparatus and, although in close proximity, it does not co-localize with GM130. Scale bars, 10 µm.

Supplementary Table 1. MS results. Table showing the number of peptides for the top 100 proteins observed in the BSDC1 pull-down followed by MS analysis . The BirA control gave very few or no unique peptides corresponding to these proteins. Cyto, cytoplasmic fraction; WC, whole cell lysate.

Unique peptides						
WC BSDC1	WC Control	Nuclei BSDC1	Nulcei Control	Cyto BSDC1	Cyto Control	symbol
222		37		120		Golgb1
50	3	13		39	2	Rpa1
37		29		34		Bsdc1
72		5		22		Blm
19				20		Emilin2
25		11		17		Ermap
32				16		Smarcal1
36		2		16		Top3a
20		7		15		Cux1
7				13		Rad52
12		5		12		Rpa2
124				11		Herc2
37				9		Fancm
1				9		Fut8
5				8		Egfl7
36				8		Peo1
19				8		Rmi1
14				7		Cltc
10		5		7		Rpa3
17				7		Ssbp1
24				6		Fanca
7		2		5		Ckap4
3				5		Pura
4				5		Purb
4				4		Bag4
3				4		Copa
				4		Ctsl
18				4		Fancb
11				4		Fancg
				4		Rdm1
				3		Adprh
2				3		Apitd1
1				3		Copb2
17				3		Faap100
12				3		Fancl
				3		Ganab
9				3		Helb
28				3		Rfc1
3				3		Stra13
				2		Alox5
				2		Atp6ap2
				2		Ccdc82
5				2		Cdca8
				2		Ces2g
				2		Cnpy2
10			2	2	2	Csnk2a1
				2		E2f1
				2		F10
4				2		Fancf
				2		Fgf15
13				2	5	Fxr1

Unique peptides						
WC BSDC1	WC Control	Nuclei BSDC1	Nulcei Control	Cyto BSDC1	Cyto Control	symbol
				2		Gm7808
2				2		Hba
6				2		Hbb-b1
				2		Hspbp1
				2		Kif2a
49				2		Lig3
52				2		Lmna
				2		Ly6k
				2		Msh2
				2		Msh6
				2		Pcmtd2
2				2		Plk1
8	1	2		2		Rfc2
10				2		Rfc5
4		3	1	2		Srsf3
3				2		Them6
2				2		Ttc13
				1		Acly
				1		Aldoart2
1				1		Anp32e
				1		Arcn1
2	1			1		Cbx1
6	1			1		Cbx3
				1		Dnajb11
1				1		Icam4
				1		Kdelc1
				1		Map2k2
8	1			1		Plec
				1		Prnp
				1		Ptpn11
13				1		Rfc3
				1		Txn
				1		Ybx1
7						Aaas
1						Aatf
2						Acaa2
2						Acad10
3						Acad12
2						Acad9
5						Acadm
5						Acads
5						Acadvl
3						Agrn
6						Ahctf1
2						Ahdc1
1						Ak4
3						Aldh18a1
3						Alkbh5
1						Ap1b1
2						Ap2a1
2						Ap2a2

Chapter 5

An MS-based approach for the identification of KLF1 binding partners

Ileana Cantù[1], Ernie de Boer[1], Dick Dekkers[2], Jeroen Demmers[2], Frank Grosveld[1,3], Sjaak Philipsen[1] and Thamar van Dijk[1].

1 Department of Cell Biology, Erasmus MC, Rotterdam, The Netherlands
2 Department of Proteomics, Erasmus MC, Rotterdam 3015 CN, The Netherlands.
3 Cancer genomics Center, Erasmus MC, Rotterdam, the Netherlands

Work in progress

ABSTRACT

Sickle Cell Disease (SCD) and β-thalassemia are monogenic disorders that respectively lead to an aberrant or insufficient production of β-globin, subsequently inducing several problems to the erythrocytes. SCD and β-thalassemia, if untreated, can lead to severe complications and eventually death of the patient. The symptoms of these two diseases are ameliorated by elevated expression of γ-globin during adult life. γ-globin is specifically expressed during the fetal stage of development and is repressed after birth, in favor of β-globin. However, in the presence of β-globin related diseases, γ-globin can substitute for it. The γ- to β- globin switch is tightly regulated by several factors, among which KLF1 plays an essential role. KLF1 is an erythroid-specific transcription factor with major functions during erythroid development. Our knowledge about its role has increased in the last two decades, but several aspects of its regulatory role remain understudied. KLF1 is predicted to work in tandem with other proteins, such as p62, CBP/p300 and BRG1. Some KLF1 binding partners have been identified over the years, but a comprehensive list of KLF1-partners has yet to emerge. In order to recognize novel interactors of KLF1 and therefore enable better understanding of its role we performed a series of pull-downs coupled to MS. The successful implementation of these isolation methods would help in the identification of KLF1-complexes *in vivo* and might reveal novel targets for drugs aimed at reactivating γ-globin expression.

INTRODUCTION

β-thalassemia and sickle cell disease (SCD) are the most common β-hemoglobinopathies and, with α-thalassemia, represent the most common human monogenic disorders[1]. They have high frequency in the Mediterranean area, North Africa, Middle East, Central Asia and Southeast Asia, even though, because of migration, they are now spread worldwide. The high frequency correlates with the presence of *Plasmodium falciparum*, a protozoan that causes malaria. Heterozygosity for SCD provides a selective advantage against malaria[2,3]. It is believed that also β-thalassemia is linked with malaria, although proper demonstration is still missing. SCD is caused by a single point mutation in the β-globin gene leading to an A to T substitution in codon 6. This results in a glutamate, a charged residue, turning into a valine, a hydrophobic residue. Subsequent aberrant polymerization of hemoglobin causes several problems to the erythrocyte and to the organism. β-thalassemia is caused by decreased or absent levels of functional β-globin, due to mutations in the β-globin gene or in its regulatory sequences. On the contrary, α-globin is physiologically expressed and in the absence of sufficient β-globin remains free and precipitates in the erythrocyte. The symptoms of these two diseases can be ameliorated by increased levels of fetal hemoglobin (HbF) during adult life. HbF is composed of 2 α- and 2 γ-globin chains, specifically expressed during the fetal stage of human development. In contrast, the major hemoglobin in adults is HbA, composed of 2 α- and 2 β-globin chains, with levels of HbF decreasing to less than 1% in healthy adults. Around birth, the switch between γ- and β-globin expression occurs, together with the switch of the site of hematopoiesis from fetal liver to bone marrow. Understanding the molecular mechanisms regulating hemoglobin switching may help to identify new potential therapeutic targets for increasing HbF levels. At present, hydroxyurea is the only approved drug used in the clinic for treatment of SCD patients, with a good response in 40-50% of patients [4,5].

KLF1 (previously known as EKLF) is a major regulator of globin switching and erythropoiesis. It is a transcription factor crucial for the proper function and development of erythrocytes, as demonstrated by *Klf1* ko/ko mouse embryos which die in utero around E14-E15 due to severe anemia[6,7]. KLF1 protein is composed of a transactivation domain at the N-terminus and a DNA binding domain at the C-terminus. Binding domains for interaction partners have been described throughout the protein, including the three zinc fingers composing the DNA binding domain[8]. Among the reported partners of KLF1 are the p62 subunit of TFIIH[9] and the TAF9 subunit of TFIID[10]. These are components of the basal transcription machinery, and they are believed to be essential for normal KLF1 function. Moreover, KLF1 associates with the N-terminal tail of histone H3, with acetylation of lysine 288 of KLF1 having a critical role in this interaction[11]. Additionally, KLF1 is thought to interact with other transcription factors. Different groups demonstrated an interaction between KLF1 and GATA1, another important erythroid transcription factor, in co-transfection experiments[12,13]. Despite that, an *in vivo* interaction between endogenous GATA1 and KLF1 proteins was not detected[14-16]. However, GATA1 and KFL1, together with SCL/TAL1, co-occupy enhancers/promoters of many erythroid genes, producing a transcriptional core network[17]. KLF1 ChIP-seq signals also often overlap with CBP/p300 signals[18], with which it also interacts[9,19]. Furthermore, KLF1 can form complexes with other chromatin modifying and remodeling factors, such as the SWI/SNF complex[20,21] through interaction with BRG1[22-25], the central ATPase subunit of the complex. These examples indicate that KLF1 activates transcription by modulating the state of the chromatin.

Adding an extra regulatory level on KLF1 function, post-translational modifications control and fine-tune its interactions with its binding partners. They can provide different protein

interaction platforms that eventually result in different downstream effects. KLF1 is post-transcriptionally modified by phosphorylation[26], ubiquitylation[27], acetylation[28] and SUMOylation[29]. Depending on the modifications and hence the interacting partners, KLF1 will have different effects in different cells. KLF1 is acetylated by CBP/p300 at position K288. Mutations in K288 lead to decreased transactivation of the β-globin promoter by KLF1[19]. K288 acetylation is necessary for optimal KLF1 DNA binding *in vivo* and for the formation of an open chromatin structure via interaction with the SWI/SNF complex[11]. Acetylation of another lysine, K302, modulates the interaction with SIN3a, which primarily takes place in primitive erythroid cells[30]. The SUMOylation at residue K74 leads to a more efficient interaction with Mi2β subunit of the NuRD complex and subsequent transcriptional repression[31]. K74 mutation, disrupting KLF1 SUMOylation, leads to loss of the inhibitory effect on megakaryocytic differentiation by KFL1[29]. Finally, KLF1 is also phosphorylated at serine and threonine residues. PPM1b, a serine/threonine phosphatase and a KLF1 interacting partner, stabilizes KLF1 but it was also shown to inhibit erythropoiesis, hence regulating erythroid homeostasis[16]. S68 phosphorylation and K74 SUMOylation also facilitate the disruption of KLF1/FOE interaction and the subsequent nuclear import of KLF1 via association-dissociation with importinβ1[32].

Though recent years have seen the identification of both novel KLF1 post-translational modifications and binding partners, they probably do not fully account for the extent of its impact on transcriptional regulation of erythroid specific genes. It is likely that several additional interactors have so far escaped detection. Moreover, the known interactions were mainly found with modelling (p62 interaction) or co-transfection experiments (TAF9 and SWI/SNF complex for instance). We therefore reasoned that establishing a purification method for KLF1 protein complexes, coupled to analysis by Mass Spectrometry, would allow identification of potential novel interaction partners. In addition to cell lines, we used an *in vivo* mouse model, which allows comparison of interacting partners between fetal (fetal liver) and adult (bone marrow) erythroid cells. By doing so, we hope to further elucidate the molecular mechanisms that dictate how KLF1 exerts transcriptional control, erythroid differentiation and globin switching.

MATERIAL AND METHODS

Mice

All animal studies were approved by the Erasmus MC Animal Ethics Committee. Bio-HA-KLF1 mice were generated using a pEV3-puro[33] vector containing mouse *Klf1* cDNA preceded by a biotinylation sequence and an HA-tag. BirA/BirA[34]::Bio-HA-KLF1::KLF1 ko/ko[6] mice and p53 ko/ko[35] mice were genotyped using the primers listed below.

BirA
Fw: 5’ TTCAGACACTGCGTGACT 3’
Rv: 5’ GGCTCCAATGACTATTTGC 3’
Approximately 500 bp PCR pruduct

Bio-KLF1
Fw: 5’ AGCCAGGGCTGGGCATA 3’
Rv: 5’ AATCCTGCGTCTCCTCAGAC 3’
Approximately 450 bp PCR pruduct

KLF1 ko
Fw1: 5' TTGCCGTTTTGCTTTGCCTG 3'
Fw2: 5' CGTTGGCTACCCGTGATATTG 3'
Rv: 5' GAAGTCCTCCTGGGTGTCCA 3'
Approximately 250 bp PCR product for the wild type allele and 270 bp for the ko allele

p53 ko
Fw: 5' TATACTCAGAGCCGGCCT 3'
Fw: 5' ACAGCGTGGTGGTACCTTAT 3'
Rv: 5' CATTCAGGACATAGCGTTGG 3'
Approximately 400 bp PCR product for the wild type allele and 700 bp for the ko allele

Cell culture
Fetal livers at E12.5 were isolated and re-suspended in 1 mL of PBS. Cells were passed through a 40µm Nylon cell strainer (352340, BD Biosystems), washed and seeded at $2x10^6$ cells/ml into StemPro-34 SFM™ (10639-011, Life Technologies) supplemented with human recombinant erythropoietin (Epo, Cilag AG, Switzerland, 2 U/ml), murine recombinant stem cell factor (SCF, R&D Systems, Minneapolis, MN, USA, 100 ng/ml), and dexamethasone (Dex, Sigma, 10^{-6} M).
MEL and 293T cells were cultured in Dulbecco's Modified Eagle Medium (DMEM; Invitrogen) supplemented with 10% fetal calf serum and penicillin-streptomycin (17-602E, Lonza).

Plasmids
Human *Klf1* and *Gfp* cDNA were amplified by PCR and cloned into pMT2_bio_HA. Bio_HA tagged constructs were subcloned into pBS containing the 2A-GFP sequence. KLF1-2A-GFP and GFP-2A-GFP were cloned alone or combined with an anti-mouse KFL1 shRNA (5'CCACTTAGCTCTGCACATGAA 3', Sigma MISSION shRNA library) in the lentiviral vector pAD5. Also V5 tagged human *Klf1* cDNA was cloned into pAD5.
The cDNA of full-length human *Paxbp1* was obtained from Open Biosystems (Clone ID: 5497083; Huntsville, AL, USA). *Paxbp1*, *p62*, *p34*, *p44* and the *Klf1* deletions were amplified by PCR and cloned to pMT2-bio-HA [36] using SalI and NotI restriction enzymes. *Klf1* deletions were constructed using different amplification primers. Bio-HA-PAXBP1 was sub-cloned into the erythroid expression vector pEV3-neo[33] and electroporated into MEL BirA cells[37].

Transient transfection and stable clone generation
Transient transfection of 293T cells was performed as described previously[38]. In order to produce stable clones expressing *Paxbp1*, MEL cells expressing the BirA biotin ligase were electroporated with the pEV3 plasmid containing Bio-Paxbp1 with the following settings: Resistance indefinite, 960 µF, 250V. Cells were then seeded as a single cell/100 µl in a 96-well plate. Stable clones were selected with neomycin (100 µg/ml, Invitrogen). After 2 weeks, clones were transferred to a 6-well plate and expanded.

Cell lysate and western blot analysis
For cytoplasmic and nuclear extracts, cells were harvested at 1200 rpm and washed once with PBS. The cell pellet was then resuspended in Buffer A (10 mM HEPES-KOH pH 7.9, 1.5 mM $MgCl_2$, 10 mM KCl) supplemented with Complete protease inhibitor mix (Roche), Pefablock (Roche), and incubated 15 minutes on ice. Cells were vortexed and spun at 13000 rpm for 30 sec. The cytoplasmic fraction was retrieved and transferred to a fresh tube. The nuclei pellet was resuspended in Buffer C (20 mM HEPES-KOH pH 7.9, 25% glycerol, 420 mM

NaCl, 1.5 mM $MgCl_2$, 0.2 mM EDTA), supplemented with Complete protease inhibitor mix (Roche), Pefablock (Roche) and 0.1 mM DTT. After 30 minutes on ice, the nuclear extracts were pelleted at 13000 rpm at 4°C for 5 minutes to separate the proteins form the rest of the nucleus. The supernatant was then transferred to a fresh tube. Total cell lysates were performed in a NP40-containing buffer (20 mM Tris-HCl pH 8.0, 137 mM NaCl, 10 mM EDTA, 50 mM NaF, 1% NP40, 10% glycerol) or RIPA buffer (20 mM Tris-HCl pH 7.4, 150 mM NaCl, 05 mM EDTA, 1% NP40, 0,1% SDS, 0,5% Sodium Deoxycholate) supplemented with Complete protease inhibitor mix (Roche) and Pefablock (Roche). Cells were lysed for 30 minutes on ice and then centrifuged 13000 rpm 5 minutes at 4°C. Cell lysates were then transferred to a fresh tube.
To visualize protein expression, cell lysates were loaded on 10% sodium dodecyl sulfate-polyacrylamide gels for electrophoresis, and the gels were transferred to nitrocellulose blotting membranes (0.45 µm, 10600002, GE Healthcare). Nitrocellulose membranes were blocked in 1% bovine serum albumin (BSA, Sigma), probed with the appropriate primary antibodies and analyzed using an Odyssey Infrared Imaging System (Li-Cor Biosciences, Cambridge, UK). The membranes were probed with KLF1 (in-house, rabbit), HA (SC 7392, Santa Cruz), NIPBL [39], PAXBP1 (21357-1-AP, ProteinTech Europe), Myc (9E10, Santa Cruz), V5 (mouse monoclonal, in-house) antibodies. The primary antibody signal was detected using IRDye secondary antibodies (Li-core).

Bio pull down and mass spectrometry analysis
Paramagnetic streptavidin beads (Dynabeads M-280, Dynal (Great Neck, NY)) were blocked by washing three times in TBS-0.3% Nonidet P-40 with 200 ng/µl purified chicken serum albumin (Sigma-Aldrich) and 1% BSA (Sigma). We used ≈10 µl of beads per $50x10^6$ cells. For pull-down, cell lysates were combined with 3 times the cell lysate volume of TBS/0.3% Nonidet P-40, 2 times the cell lysate volume of Tris-0,45% Nonidet P-40, 250 units of Benzonase (Novagen) and the streptavidin beads and incubated 3 h at 4 °C. This was followed by three washes in TBS-0.3% Nonidet P-40 . Bound material was eluted by boiling for 5 min in Laemmli protein sample loading buffer and analyzed by immunoblotting as above or by MS.
Digestion with trypsin (Promega, sequencing grade) on paramagnetic streptavidin beads and liquid chromatography-tandem MS (LC-MS/MS) were performed as described previously[36,37]. Nanoflow LC-MS/MS was performed on an 1100 series capillary LC system (Agilent Technologies, Santa Clara, CA) coupled to an LTQ-Orbitrap mass spectrometer (Thermo), operating in positive mode and equipped with a nanospray source. Peptide mixtures were trapped on a ReproSil C18 reversed phase column (Dr Maisch GmbH; column dimensions 1.5 cm × 100 µm, packed in-house) at a flow rate of 8 µl/min. Peptide separation was performed on ReproSil C18 reversed phase column (Dr Maisch GmbH; column dimensions 15 cm × 50 µm, packed in-house) using a linear gradient from 0 to 80% B (A = 0.1% formic acid; B = 80% (v/v) acetonitrile, 0.1% formic acid) in 70 min and at a constant flow rate of 200 nl/min using a splitter. The column eluent was directly sprayed into the electrospray ionization (ESI) source of the mass spectrometer. Mass spectra were acquired in continuum mode; fragmentation of the peptides was performed in data-dependent mode. Peak lists were automatically created from raw data files using the Mascot Distiller software (version 2.3; MatrixScience). The Mascot search algorithm (version 2.2, MatrixScience) was used for searching against the IPI_mouse database (version 3.83, containing 60,010 sequences and 27,475,843 residues). The peptide tolerance was set to 10 ppm and the fragment ion tolerance was set to 0.8 Da. A maximum number of two missed cleavages by trypsin were allowed and carbamidomethylated cysteine and oxidized methionine were set as fixed and

variable modifications, respectively. Search results were parsed into a home-built database system for further analysis. Entries were parsed if they had a minimum peptide Mascot score of 25, and a significance threshold of $p < 0.05$; the option "require red bold" was also selected. Using these parameters yields an estimated peptide false discovery rate (FDR) of 3–5% against a target decoy database. The data was converted using PRIDE Converter[40] (http://pride-converter.googlecode.com).

FACS sorting

Transduced MEL cells were FACS sorted on the GFP expression intensity using a FACS Aria III (BD bioscience) instrument with FACS Diva software (BD bioscience).

RESULTS

Bio-HA-KLF1 isolation from mouse fetal liver and adult bone marrow cells.

The isolation of endogenous KLF1 using antibodies is extremely inefficient. Given that the streptavidin-biotin pulldown protocol has been successfully used in the identification of interacting proteins of other transcription factors[14,36,41], we reasoned that it would be a suitable method to isolate KLF1 complexes. We therefore designed a construct containing a biotinylation sequence followed by an HA-tag at the N-terminus of mouse KLF1 (Bio-HA-KLF1, Figure 1A). This construct was cloned in a pEV3-puro vector and used to generate transgenic mice. Bio-HA-KLF1 mice were then crossed with BirA/BirA mice and *Klf1* ko/ko[37] mice to obtain Bira/Bira::*Klf1* ko/ko::Bio-HA-KLF1 (from now on called Bio-HA-KLF1 mice) strain. As KLF1 pulldowns from freshly isolated fetal livers did not yield enough material (Ernie de Boer, unpublished data), we decided to expand Mouse Fetal Liver (MFL) cells in

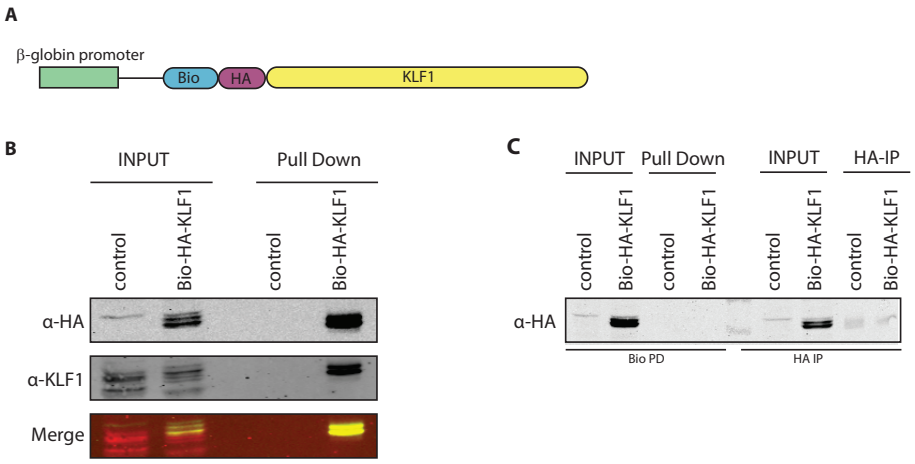


Figure 1. The Bio-KLF1 construct in primary mouse cells. (A) Bio-HA-KLF1 construct used to generate the mouse strains. (B) Streptavidin pull-down of biotinylated-KLF1 in E12.5 cultured mouse fetal liver cells. The membrane was stained with α-HA or α-KLF1 antibodies. BirA/BirA only mice were used as controls. Input lane represents 10% of total extract used in the pull-down. Ponceau S staining was used as loading control for the input material (not shown). (C) Streptavidin pull-down of biotinylated-KLF1 (Bio PD) and α-HA immunoprecipitation (HA IP) in erythroid cells from bone marrow cultures. The membrane was stained with α-HA antibody. BirA/BirA only mice were used as control. Input lane represents 10% of total extract used in the pull-down.

StemPro-34SFM medium. The cells were cultured for 4/5 days, as at later time points the majority of the cells were differentiating and eventually dying. Thus, we isolated MFL cells from Bio-HA-KLF1 mice and from BirA only mice as control. After culturing, we performed a streptavidin pull down on nuclear extracts and we obtained a noticeable enrichment of biotinylated KLF1, as shown by either the HA- or the KLF1-antibodies (Figure 1B). Subsequently, the same method was used on bone marrow cells, isolated from young adult mice, with the ultimate goal of clarifying the difference in KLF1 binding partners between fetal and adult hematopoietic cells. Unfortunately, we were unable to enrich for KLF1 in extracts from these cells, neither from freshly isolated nor from cultured cells (Figure 1C). Our data indicate that the biotinylation protocol can be used for isolation of KLF1 from cultured MFL cells, but it cannot be applied to all erythropoietic cells. This is probably related to the reduced *ex vivo* expansion capacity of bone marrow derived erythroid precursors. Following the promising results in fetal liver cells, we decided to scale up the MFL cultures, and send the samples for MS analysis. Unfortunately, and in contrast to the Western blot results, we were not able to detect KLF1 peptides in 3 individual experiments.

Bio-HA-KLF1 isolation and MS analysis in cultured MFL cells.
Given the inefficient isolation of KLF1 from primary MFL cells, we decided to establish a Bio-HA-KLF1 cell line. To achieve this we crossed Bio-HA-KLF1 mice with *p53* ko/ko mice, and we successfully established a Bira/Bira::Bio-HA-KLF1::KLF1 ko/ko::p53 ko/ko cell line from E12.5 fetal liver. Subsequently, we checked if we could efficiently pulldown biotinylated KLF1 by streptavidin beads, similar to the isolation method used for cultured primary MFL cells. Notably, we obtained a marked enrichment of biotinylated KLF1, which was not evident in BirA-only control cells (Figure 2A). Having demonstrated that we can successfully enrich for KLF1 also in this system, we increased input material by ten times (using 500 million cells per sample), performed the pull down experiment under identical conditions and send eluates for MS analysis. The results we obtained are summarized in Table 1. Unlike our initial experiments in cultured MFL cells, we were able to detect several peptides of

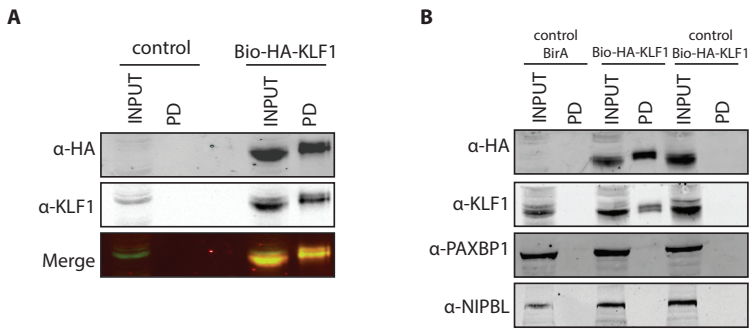


Figure 2. Bio-KLF1 pull-down in immortalized mouse fetal liver cells. (A) Streptavidin pull-down of biotinylated-KLF1 in immortalized MFL cells. An immortalized MFL cell line derived from BirA/BirA mice was used as control. The membrane was stained with α-HA and α-KLF1 antibodies. Input lane represents 10% of total extract used in the pull-down. Ponceau S staining was used as loading control for the input material (not shown). (B) Streptavidin pull-down of biotinylated-KLF1 followed by staining for two potential KLF1 binding partners. As controls, the immortalized BirA/BirA only and the immortalized Bio-KLF1 only MFL cell lines were used. The membrane was stained with α-HA, α-KLF1, α-PAXBP1 and α-NIPBL antibodies. Input lane represents 10% of total extract used in the pull-down. Ponceau S staining was used as loading control for the input material (not shown).

Table 1. Potential KFL1 binding partners identified in the Bio-HA-KLF1 MFL cell line (Unique peptides Bio-HA-KLF1 MFL). The unique peptides identified for BirA only MFL cells (unique peptides control) are shown as control. KLF1 is depicted in red and PAXBP1 in blue.

Unique peptides Bio-HA-KLF1 MFL	Unique peptides control	Protein	Description
4	1	Klf1	Krueppel-like factor 1
7		Ppp1r10	Serine/threonine-protein phosphatase 1 regulatory subunit 10
7		Tox4	TOX high mobility group box family member 4
5		Gtf2h1	General transcription factor II H, polypeptide 1
4		Paxbp1	PAX3- and PAX7-binding protein 1
4		Ruvbl2	RuvB-like 2
3	1	Snw 1	SNW domain-containing protein 1
3	1	Tada3	Transcriptional adapter 3
3		Trrap	Transformation/transcription domain-associated protein
2		Aff1	AF4/FMR2 family member 1
2		Eny2	Transcription and mRNA export factor ENY2
2		Ints1	Integrator complex subunit 1
2		Mios	WD repeat-containing protein mio
2	1	Mnat1	CDK-activating kinase assembly factor MAT1
2		Runx1	Runt related transcription factor 1, isoform CRA_c
2		Nipbl	Nipped-B-like protein
2		Gtf2h2	General transcription factor IIH subunit 2
1		Cbx5	Chromobox protein homolog 5
1		Ccdc101	SAGA-associated factor 29 homolog
1		Gfi1b	Zinc finger protein Gfi-1b
1		Gse1	Genetic suppressor element 1
1		Tceb3	Transcription elongation factor B polypeptide 3
1		Tcerg1	Transcription elongation regulator 1
1		Ccnk	Cyclin-K
1		Cdk7	Cyclin-dependent kinase 7
1		Cdk9	Cyclin-dependent kinase 9

KLF1. We also identified p62 (or GTF2H1), a known interactor of KLF1. After filtering of the data, i.e. excluding common contaminant proteins and focusing on transcription-related proteins, we narrowed the list of potential KLF1 interacting partners down to 23 factors. Among these proteins was PAXBP1 (PAX3- and PAX7-binding protein 1). PAXBP1 represented an interesting candidate as it is an adapter protein linking the transcription factors PAX3 and PAX7 to the histone methylation machinery[42]. PAXBP1 was specifically present in more than one MS results we obtained from MFL cells. We therefore thought that PAXBP1 could bridge KLF1 to other proteins and we sought to validate the KLF1-PAXBP1 interaction by streptavidin pulldown of biotinylated KLF1 followed by western blot analysis. During this experiment, we also used a Bio-HA-KLF1 :: *Klf1* ko/ko :: *p53* ko/ko cell line, lacking the BirA ligase, as a further control. Though we efficiently pulled down KLF1, we could not detect PAXBP1 as an interactor (Figure 2B). Moreover, we were also unable to co-purify KLF1 with NIPBL, another MS-identified potential KLF1 partner (Figure 2B).

MS analysis of KLF1 complexes in MEL cells stably expressing Bio-HA-KLF1

Since the experiments using cultured MFL cells or MFL-derived immortalized cells yielded inconclusive results, we turned to the MEL-BirA system which had been successfully applied to isolate erythroid transcription factor complexes previously[43]. In order to faithfully identify KLF1 complexes, we generated multiple control (GFP-GFP) or KLF1 (KLF1-GFP) constructs. More specifically, we constructed a vector with the GFP (termed G2G) or human KLF1 (termed K2G) cDNA, preceded by an HA-tag and a biotinylation sequence, in front of a 2A GFP cassette. As overexpression of KLF1 is toxic, we tried to reduce the level of endogenous KLF1. We combined the Bio-HA-KLF1-2A-GFP construct with a shRNA for mouse KLF1, in the reverse orientation compared to either the GFP-2A-GFP (termed G2G_Rv) or the KLF1-2A-GFP (termed K2G_Rv) cassette (Figure 3A). We first verified the expression of these constructs in HEK 293T cells. Western blot analysis revealed that all constructs were efficiently expressed and that ribosome skipping within the 2A-sequence, although not complete, worked to a great extent (Figure 3B).

Next, we transduced MEL-BirA cells with lentiviral vectors. Five days after transduction, cells were sorted based on GFP fluorescence, in GFP-positive (expressing the construct) or GFP-negative (not expressing the construct) populations. In order to check if the transduced cells retained stable expression, they were analyzed by flow cytometry 11 days post transduction. Only a fraction of the cells remained GFP positive (Figure 4A). In accordance with the notion that excessive KLF1 expression is detrimental, cells expressing Bio-HA-KLF1 without concomitant knockdown of endogenous KLF1 (K2G) were markedly reduced in numbers compared to those expressing Bio-HA-KLF1 with simultaneous downregulation of KLF1 (K2G-Rv) (Figure 4A). These data suggest that endogenous KLF1 depletion facilitates stable expression of exogenous KLF1, possibly because erythroid cells increasingly differentiate upon KLF1 over-expression. Subsequently, we checked for the expression levels of tagged KLF1 or GFP by sorting MEL cells 11 days post transduction in GFP positive and GFP negative

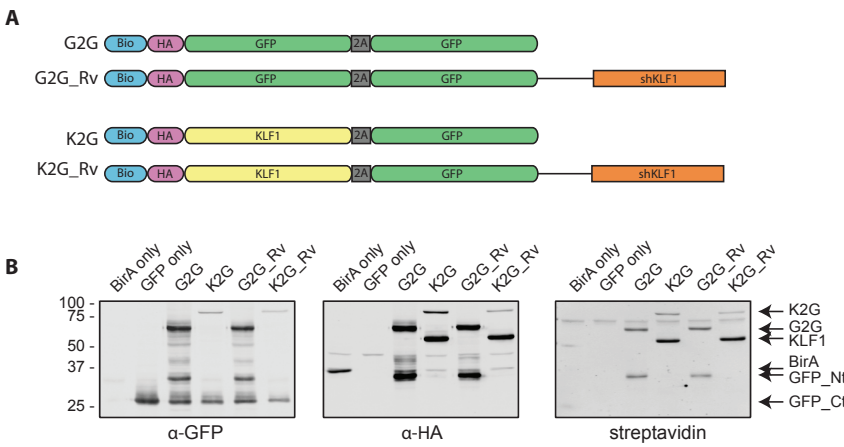


Figure 3. KLF1 vectors for expression in MEL and 293T cells. (A) Schematic overview of the KLF1 vectors used. GFP or KLF1 cDNAs were cloned in front of a 2A-GFP cassette (G2G or K2G, respectively). A mouse shRNA for KLF1 was also cloned downstream of the G2G or the K2G constructs, in the reverse orientation compared to G2G or K2G (G2G_Rv or K2G_Rv). (B) Verification of the expression of the constructs in transiently transfected 293T cells. The membrane was stained with α-HA and α-GFP antibodies or with streptavidin. HA-BirA only and GFP only were used as controls. GFP-Nt: Bio-HA-GFP; GFP-Ct: GFP downstream of the 2A sequence.

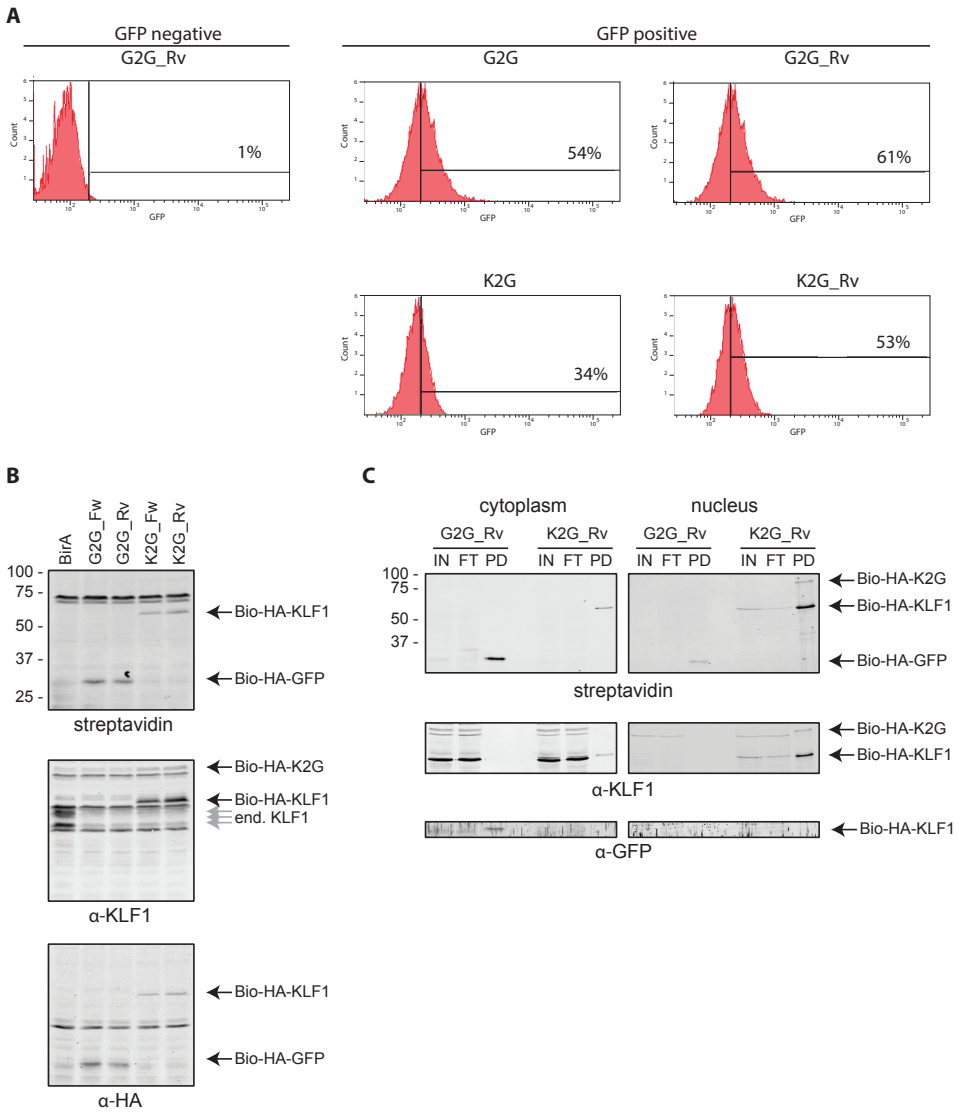


Figure 4. Bio-KLF1 purification from MEL-BirA cells. (A) FACS analysis showing the percentage of GFP positive cells (on the right) 11 days post-transduction with G2G, G2G_Rv, K2G and K2G_Rv constructs. On the left the GFP negative population for G2G_Rv is depicted. The negative populations of the other transducitons were comparable (not shown). Flow cytometry plots show that cells tend to shut down the expression of these constructs over time. Approximately 50% of the cells retain moderate to high expression. Notably, the presence of the shRNA for KFL1 results in higher levels of GFP expression. (B) Whole cell extracts of MEL cells transduced with G2G/K2G coupled with the shRNA for KLF1 in the forward or the reverse orientation (G2G/K2G_Rv or _Fw) 14 days post transduction. The membrane was probed with α-HA and α-KLF1 antibodies or with streptavidin. MEL-BirA cells were used as control. End.KLF1: endogenous KLF1. (C) Streptavidin pull-down of biotinylated-KLF1 in MEL cells 14 days after transduction using cytoplasmic and nuclear extracts. The membrane was probed with α-KLF1 and α-GFP antibodies or with streptavidin. Input lane represents 10% of total extract used in the pull-down. IN: input; FT: flow through; PD: pull-down

population. Even though only half of the cells retained moderate to high expression of GFP, we achieved good expression levels of Bio-HA-KLF1 and of the control Bio-HA-GFP (Figure 4B). In addition, we obtained adequate knockdown of endogenous KLF1 (Figure 4B). To check whether the orientation of the shRNA would affect the quality of knockdown, we transduced MEL cells with a vector containing the shRNA for KLF1 in the forward orientation (G2G_Fw or K2G_Fw). No difference was detected between the construct with the forward or reverse orientation (Figure 4B), indicating that shRNA orientation does not affect the efficiency of KLF1 knockdown. Thus, we continued with the constructs containing the shRNA for KLF1 in the reverse orientation. To determine whether we could efficiently purify Bio-HA-KLF1 or Bio-HA-GFP, we performed a pull-down experiments on transduced MEL cells using both cytoplasmic and nuclear fractions (Figure 4C). The strongest GFP signal was detected in the cytoplasmic fraction, a much fainter signal was present in the nuclear fraction (Figure 4C). In contrast, KLF1 was mainly present in the nuclear fraction (Figure 4C).
Having confirmed that we could efficiently pull-down KLF1, we sent those samples for MS analysis. Among the proteins specifically co-purified with Bio-HA-KLF1 were the p62 subunit of the TFIIH complex[9] and Ankrd17[32], two known interactors of KLF1 (Table 2). By Western blot analysis, we observed that Bio-HA-KLF1 specifically interacts with Myc-tagged p62, while similarly tagged p34 and p44 subunits did not co-precipitate with Bio-HA-KLF1 (Figure 5A). In order to identify which part of KLF1 is responsible for the interaction with p62, we generated Bio-HA-KLF1 deletion mutants, lacking specific segments of the protein (Figure 5B). We found that the C-terminal part of KLF1 mediates the binding to p62, as evident by the loss of interaction upon its removal (Figure 5C).

PAXBP1 as possible KLF1 interacting protein

Having shown that we can identify known interactors of KLF1 upon streptavidin pull-down, we shifted our attention to PAXBP1, which was specifically purified also in three MS experiments performed in the MEL cells system (Table 2). First, we tried to confirm the KLF1-PAXBP1 interaction in transduced MEL cells after streptavidin pull-down of biotinylated KLF1. Surprisingly, we were unable to detect any PAXBP1 signal using an antibody against the endogenous protein (Figure 6A). Similarly, we could not retrieve Myc-tagged PAXBP1 upon Bio-HA-KLF1 pull-down performed in transfected 293T cells (Figure 6B). We also performed the reverse experiment, pulling down Bio-HA-PAXBP1 and subsequently probing

Table 2. Number of peptides retrieved in MS analysis of Bio-KLF1-compexes for TFIIH complex, PAXBP1 and Ankrd17 (Unique peptides Bio-HA-KLF1 MFL). The unique peptides identified for BirA-only MFL cells (unique peptides control) are shown as control.

Unique peptides Bio-HA-KLF1 MFL	Unique peptides control	Protein	Description
7		KLF1	Krueppel-like factor 1
7		Gtf2h1	General transcription factor IIH subunit 1
8		Gtf2h2	General transcription factor IIH subunit 2
2		Gtf2h3	General transcription factor IIH subunit 3
2		Gtf2h4	General transcription factor IIH subunit 4
2	1	Gtf2h5	General transcription factor IIH subunit 5
4		Ercc2	TFIIH basal transcription factor complex helicase XPD subunit
15		Ercc3	TFIIH basal transcription factor complex helicase XPB subunit
11		Paxbp1	PAX3- and PAX7-binding protein 1
2		Ankrd17	Ankyrin repeat domain-containing protein 17

for V5-KLF1 protein, but again we failed to detect any interaction (Figure 6C). To rule out potential artefacts of PAXBP1 over-expression, we created MEL clones stably expressing Bio-HA-PAXBP1. After screening several single-cell clones, we selected clone#9 which appeared to have similar levels of exogenous and endogenous PAXBP1 expression (Figure 6D). In order to exclude that our inability to identify the KLF1-PAXBP1 interaction on a Western blot did not come from the stringency of the buffers used, we tried out purification under various salt and detergent conditions. We noticed that most suitable conditions for PAXBP1 and KLF1 isolation involved the use of NP40 (Figure 6E). Therefore, we repeated the Bio-HA-PAXBP1 pull-down under either the standard or optimized conditions, but again we couldn't validate the interaction with KLF1 (Figure 6F). Thus, we conclude that PAXBP1 is either a very transient partner of KLF1 or, more likely, not a true interactor.

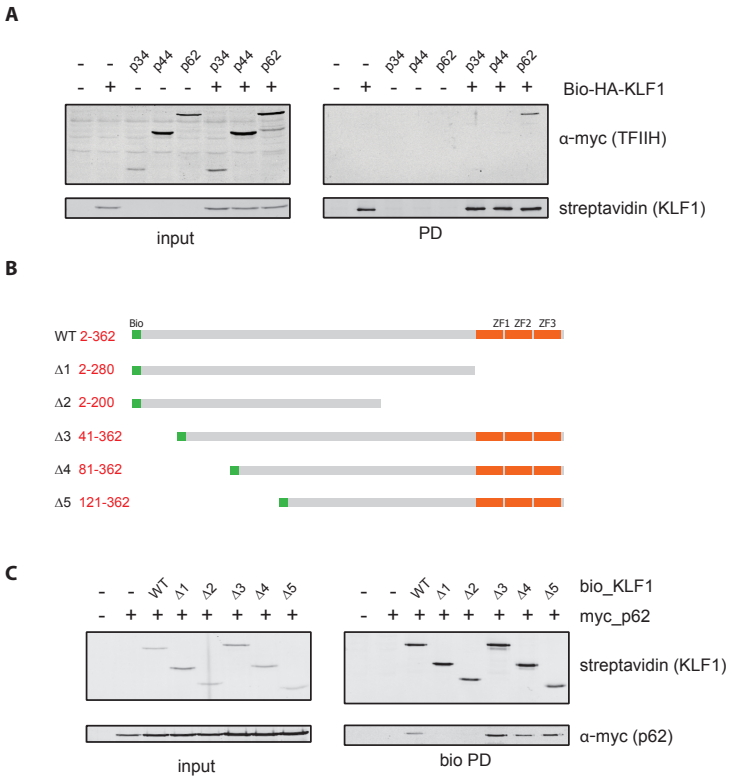


Figure 5. KLF1 interacts with p62. (A) Streptavidin pull-down of biotinylated-KLF1 in 293T cells co-transfected with different Myc-tagged components of the TFIIH complex, namely p34, p44 and p62. The membrane was probed with α-Myc antibody or streptavidin. Input lane represents 10% of total extract used in the pull-down. (B) KLF1 deletion-constructs. Note that the biotinylation sequence (Bio) was placed as a tag at the N-terminus. Amino acid sequences retained from the original protein are depicted in red. Zinc fingers (ZF) are depicted in orange. (C) Streptavidin pull-down of biotinylated-KLF1 deletion mutants in 293T cells co-transfected with Myc-tagged p62. The membrane was probed with α-Myc antibody or streptavidin. Input lane represents 10% of total extract used in the pull-down.

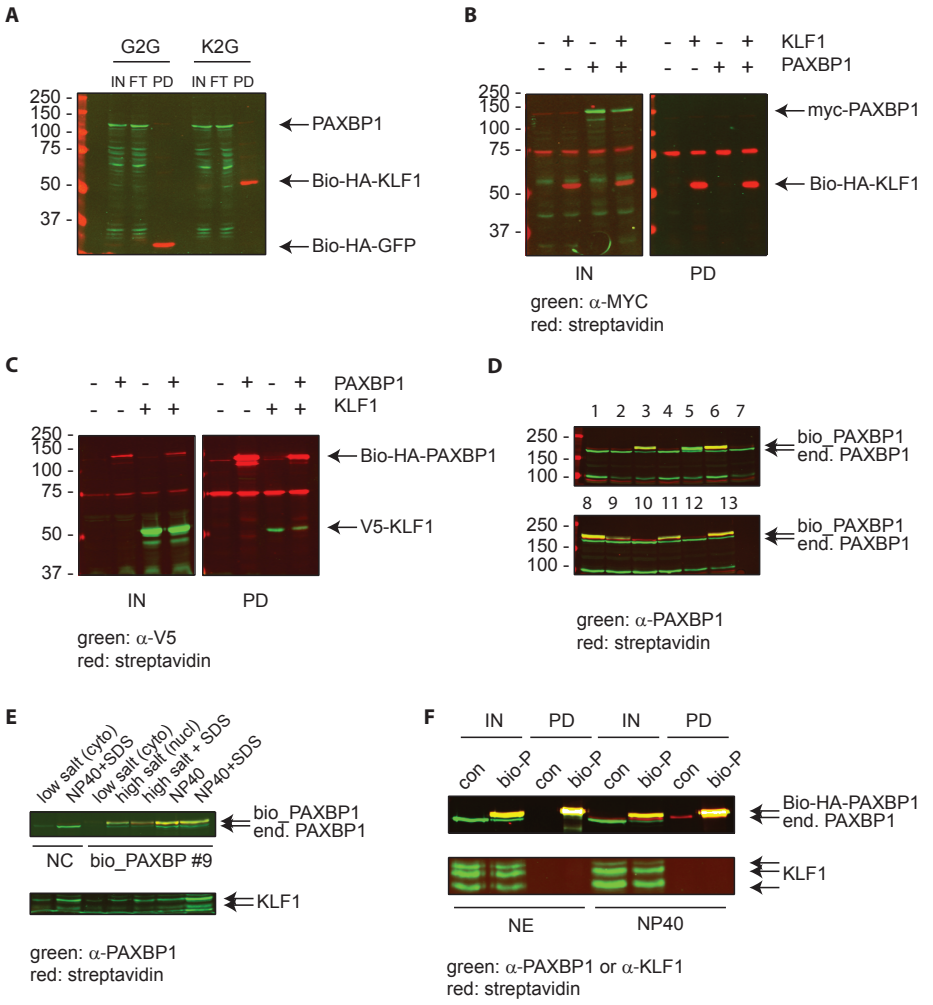


Figure 6. Investigating the potential KLF1 and PAXBP1 interaction. (A) Streptavidin pull-down of biotinylated-KLF1 in MEL cells 14 days after transduction. The membrane was stained with α-PAXBP1 antibody (green) or streptavidin (red). Input lane represents 10% of total extract used in the pull-down. IN: input; FT: flow through; PD: pull-down. (B) Streptavidin pull-down of biotinylated-KLF1 in 293T cells co-transfected with empty vector (control) or Myc-tagged PAXBP1. The membrane was stained with α-Myc antibody (green) or streptavidin (red). Input lane represents 10% of total extract used in the pull-dow.n IN: input; PD: pull-down. (C) Streptavidin pull-down of biotinylated-HA-PAXBP1 in 293T cells co-transfected with empty vector or V5-tagged KLF1. The membrane was stained with α-V5 antibody (green) or streptavidin (red). Input lane represents 10% of total extract used in the pull-down. IN: input; PD: pull-down. (D) Identification of stable MEL clones expressing Bio-HA-PAXBP1. The membrane was stained with α-PAXBP1 antibody (green) or streptavidin (red). The clone chosen for further experiments was clone #9. (E) Different protein isolation conditions on cells from clone#9. MEL-BirA cells were used as negative control (NC). The membrane was stained with α-PAXBP1 antibody (green) or streptavidin (red). (F) Streptavidin pull-down of biotinylated-PAXBP1 in stable MEL cell line #9. Two different protein isolation buffers were used: nuclear extract (NE) buffer or NP40-containing buffer (NP40). The membrane was stained with α-PAXBP1 or KLF1 antibody (green) or streptavidin (red). Input lane represents 10% of total extract used in the pull-down. Bio-P: Bio-HA-PAXBP1; IN: input; PD: pull-down; end. PAXBP1: endogenous PAXBP1.

DISCUSSION

Even though in the past years we have acquired more insight into the role of KLF1 in the regulation of erythropoiesis, still little is known about the mechanisms that govern its action. KLF1 is present in both primitive and definitive erythroid cells[44], but its function appears much more critical in definitive cells. *Klf1* ko/ko primitive cells have minor morphological abnormalities[45], while definitive cells cannot develop beyond the proerythroblast stage in the absence of KLF1. Clearly, since primitive cells are only modestly affected by the absence of KFL1, erythroid gene expression is largely independent on the presence of KLF1 in these cells. Nevertheless, KLF1 is required for the expression of genes involved in hemoglobin metabolism, cell division and membrane proteins, also important for primitive erythrocytes. During primitive erythropoiesis, KLF1 might have a redundant role with KLF2. KLF1 and KLF2 can indeed partially compensate for each other, as demonstrated by the more severe phenotype in *Klf1* -/- *Klf2* -/- double knockout embryos, compared to single knock out embryos[46]. The role of KLF1 in primitive cells is still enigmatic, it remains to be demonstrated that it selectively activates genes in these cells. In contrast, KLF1 is essential for definitive erythroid cells. These cells become dependent on KLF1 when they progress from proerythroblasts to more mature cells, and KLF1 activates the global erythroid gene expression programme during this progression[47]. This suggests that the diverse actions of KLF1 in primitive/definitive erythroid cells or during definitive erythroid maturation might involve the interaction with different interacting partner differentially expressed during red blood cells development. Alternatively, different post-translational modifications could mediate the changes in protein-protein interactions and hence in the activation of different downstream genes.

Notably, although we know that KLF1 acts as a major transcriptional activator, the currently known KLF1 interacting proteins have been identified by modelling and *in vitro* analyses. In order to find novel KLF1 partners, we reasoned that KLF1 purification followed by MS analysis might provide interesting novel candidates. Thus, we utilized *in vivo* biotinylation tagging, a system ideal for the isolation of transcription factors[14]. Unfortunately, our MS analyses did not yield conclusive results. Although we validated our pulldown by identifying known binding partners of KLF1, namely Ankrd17 and p62, we were not be able to validate the novel potential interactors. Notably, we found that p62 binds to the C-terminus of KLF1 (Figure 5C). This result is in disagreement with the previously published paper that identified p62 binding site at the N-terminus of KLF1, specifically between amino acids 51 and 90[9]. In this paper, which is mainly based on homology between the TAD domain of p53 and two similar domains found in KLF1, isothermal titration calorimetry was used to demonstrate that the TAD2 domain of KLF1 interacts with the p62 PH domain, both fused to GST. In contrast, we used full-length proteins in co-transfection experiments. These two different approaches might explain the different results; further experiments are needed to elucidate to establish the true nature of this interaction.

Even though PAXBP1 was specifically present in the Bio-HA-KLF1 MS data, subsequent experiments failed to confirm that it is a true interactor of KLF1. To further explore this, one could attempt to link PAXBP1 and KLF1 genetically by looking at the expression of known KLF1 target genes in the absence of PAXBP1, or to perform PAXBP1 ChIP experiments and look for KLF1-binding sequences.

A possible explanation for our results is that with our purification method we fail to adequately isolate the physiological DNA-bound Bio-HA-KLF1-complexes and instead primarily pulldown nucleoplasmic Bio-HA-KLF1. The use of size-exclusion chromatography (SEC), which separates molecules based on their size, would help clarifying if Bio-HA-KLF1

is present in complexes with multiple partners, in which case it would be eluted early from the column, or is mainly unbound, in which case it would require more time to pass through the column. We attempted to release larger amounts of chromatin-bound Bio-HA-KLF1, and more generally to release more chromatin proteins, by increasing the quantity of benzonase (DNA nuclease) used. Although digesting nucleic acids is essential for protein complexes purification, increasing the amount of benzonase did not improve the quality of KLF1 complexes purification. An alternative way to increase the yield of Bio-HA-KLF1 complexes would be by chemical crosslinking, thereby essentially freezing even transient partners. Using the streptavidin-biotinylation pulldown system allows for the use of very stringent washing buffers, minimizing the expected increase of background signal due to crosslinking. Finally, even though we have encountered technical problems with finding novel KLF1 partners, the quantity and purity of Bio-HA-KLF1 could be sufficient to identify post-translational modifications at different stages of erythroid differentiation.

Tagging KFL1 at the N-terminus has been proved to be successful, since KFL1 retains its endogenous activity. Bio-KLF1::KLF1 ko/ko mice, although slightly anemic (data not shown), do not die during gestation, demonstrating that KLF1 is able to adequately perform its physiological role during erythroid development. Nevertheless, Bio-KLF1 pulldown, in tandem with MS, yielded suboptimal results, possibly because KLF1 biotinylation either is partially hidden or partially inhibits normal KLF1 interactions. Moving the position of the tag might improve the efficiency of the experiment. Given that nuclear localization signal (NLS) domains are generally exposed in all cells in which a transcription factor is active, as they are needed for the proper nuclear localization of the protein, tagging in the NLS sequence might be a valid option. This approach has been proved to be successful in the case of two proteins, Fli-1 and Irf2bp2[48]. KLF1 has two nuclear localization signals near the C-terminus[49,50] which constitute potential tagging sites.

In conclusion, we have established a purification method to isolate KFL1 from erythroid cells. Further improvements are needed to identify novel KLF1 interactors faithfully. We suggest to place the purification tag at the C-terminus of the protein, and to use a chemical crosslinker to stabilize the interactions between KLF1 and its partners. Identification of *bona fide* KLF1 interaction partners is essential to understand how this transcription factor orchestrates terminal erythroid differentiation at the molecular level.

Authors contributions

I.C., T.v.D. and S.P. designed the experiments. I.C. and T.v.D. performed the experiments. E.d.B provided the mice. D.D. and J.D. performed the MS-analysis. The paper was written by I.C., T.v.D. and S.P.

REFERENCES

1. Angastiniotis, M. & Modell, B. Global epidemiology of hemoglobin disorders. *Ann N Y Acad Sci* **850**, 251-69 (1998).
2. Aidoo, M. et al. Protective effects of the sickle cell gene against malaria morbidity and mortality. *Lancet* **359**, 1311-2 (2002).
3. Nagel, R.L. & Roth, E.F., Jr. Malaria and red cell genetic defects. *Blood* **74**, 1213-21 (1989).
4. Gambari, R. Foetal haemoglobin inducers and thalassaemia: novel achievements. *Blood Transfus* **8**, 5-7 (2010).
5. Steinberg, M.H. Therapies to increase fetal hemoglobin in sickle cell disease. *Curr Hematol Rep* **2**, 95-101 (2003).
6. Nuez, B., Michalovich, D., Bygrave, A., Ploemacher, R. & Grosveld, F. Defective haematopoiesis in fetal liver resulting from inactivation of the EKLF gene. *Nature* **375**, 316-8 (1995).
7. Perkins, A.C., Sharpe, A.H. & Orkin, S.H. Lethal beta-thalassaemia in mice lacking the erythroid CACCC-transcription factor EKLF. *Nature* **375**, 318-22 (1995).
8. Yien, Y.Y. & Bieker, J.J. EKLF/KLF1, a tissue-restricted integrator of transcriptional control, chromatin remodeling, and lineage determination. *Mol Cell Biol* **33**, 4-13 (2013).
9. Mas, C. et al. Structural and functional characterization of an atypical activation domain in erythroid Kruppel-like factor (EKLF). *Proc Natl Acad Sci U S A* **108**, 10484-9 (2011).
10. Sengupta, T., Cohet, N., Morle, F. & Bieker, J.J. Distinct modes of gene regulation by a cell-specific transcriptional activator. *Proc Natl Acad Sci U S A* **106**, 4213-8 (2009).
11. Sengupta, T., Chen, K., Milot, E. & Bieker, J.J. Acetylation of EKLF is essential for epigenetic modification and transcriptional activation of the beta-globin locus. *Mol Cell Biol* **28**, 6160-70 (2008).
12. Gregory, R.C. et al. Functional interaction of GATA1 with erythroid Kruppel-like factor and Sp1 at defined erythroid promoters. *Blood* **87**, 1793-801 (1996).
13. Merika, M. & Orkin, S.H. Functional synergy and physical interactions of the erythroid transcription factor GATA-1 with the Kruppel family proteins Sp1 and EKLF. *Mol Cell Biol* **15**, 2437-47 (1995).
14. Grosveld, F. et al. Isolation and characterization of hematopoietic transcription factor complexes by in vivo biotinylation tagging and mass spectrometry. *Ann N Y Acad Sci* **1054**, 55-67 (2005).
15. Woo, A.J. et al. Identification of ZBP-89 as a novel GATA-1-associated transcription factor involved in megakaryocytic and erythroid development. *Mol Cell Biol* **28**, 2675-89 (2008).
16. Yien, Y.Y. & Bieker, J.J. Functional interactions between erythroid Kruppel-like factor (EKLF/KLF1) and protein phosphatase PPM1B/PP2Cbeta. *J Biol Chem* **287**, 15193-204 (2012).
17. Wontakal, S.N. et al. A core erythroid transcriptional network is repressed by a master regulator of myelo-lymphoid differentiation. *Proc Natl Acad Sci U S A* **109**, 3832-7 (2012).
18. Tallack, M.R. et al. Novel roles for KLF1 in erythropoiesis revealed by mRNA-seq. *Genome Res* **22**, 2385-98 (2012).
19. Zhang, W., Kadam, S., Emerson, B.M. & Bieker, J.J. Site-specific acetylation by p300 or CREB binding protein regulates erythroid Kruppel-like factor transcriptional activity via its interaction with the SWI-SNF complex. *Mol Cell Biol* **21**, 2413-22 (2001).
20. Armstrong, J.A., Bieker, J.J. & Emerson, B.M. A SWI/SNF-related chromatin remodeling complex, E-RC1, is required for tissue-specific transcriptional regulation by EKLF in vitro. *Cell* **95**, 93-104 (1998).
21. Kadam, S. et al. Functional selectivity of recombinant mammalian SWI/SNF subunits. *Genes Dev* **14**, 2441-51 (2000).
22. Brown, R.C. et al. Distinct domains of erythroid Kruppel-like factor modulate chromatin remodeling and transactivation at the endogenous beta-globin gene promoter. *Mol Cell Biol* **22**, 161-70 (2002).
23. Desgardin, A.D. et al. Regulation of delta-aminolevulinic acid dehydratase by kruppel-like factor 1. *PLoS One* **7**, e46482 (2012).
24. Kim, S.I., Bresnick, E.H. & Bultman, S.J. BRG1 directly regulates nucleosome structure and chromatin looping of the alpha globin locus to activate transcription. *Nucleic Acids Res* **37**, 6019-

27 (2009).
25. Perrine, S.P., Mankidy, R., Boosalis, M.S., Bieker, J.J. & Faller, D.V. Erythroid Kruppel-like factor (EKLF) is recruited to the gamma-globin gene promoter as a co-activator and is required for gamma-globin gene induction by short-chain fatty acid derivatives. *Eur J Haematol* **82**, 466-76 (2009).
26. Ouyang, L., Chen, X. & Bieker, J.J. Regulation of erythroid Kruppel-like factor (EKLF) transcriptional activity by phosphorylation of a protein kinase casein kinase II site within its interaction domain. *J Biol Chem* **273**, 23019-25 (1998).
27. Quadrini, K.J. & Bieker, J.J. EKLF/KLF1 is ubiquitinated in vivo and its stability is regulated by activation domain sequences through the 26S proteasome. *FEBS Lett* **580**, 2285-93 (2006).
28. Zhang, W. & Bieker, J.J. Acetylation and modulation of erythroid Kruppel-like factor (EKLF) activity by interaction with histone acetyltransferases. *Proc Natl Acad Sci U S A* **95**, 9855-60 (1998).
29. Siatecka, M., Xue, L. & Bieker, J.J. Sumoylation of EKLF promotes transcriptional repression and is involved in inhibition of megakaryopoiesis. *Mol Cell Biol* **27**, 8547-60 (2007).
30. Chen, X. & Bieker, J.J. Stage-specific repression by the EKLF transcriptional activator. *Mol Cell Biol* **24**, 10416-24 (2004).
31. Stielow, B. et al. Identification of SUMO-dependent chromatin-associated transcriptional repression components by a genome-wide RNAi screen. *Mol Cell* **29**, 742-54 (2008).
32. Shyu, Y.C. et al. Tight regulation of a timed nuclear import wave of EKLF by PKCtheta and FOE during Pro-E to Baso-E transition. *Dev Cell* **28**, 409-22 (2014).
33. Needham, M. et al. LCR/MEL: a versatile system for high-level expression of heterologous proteins in erythroid cells. *Nucleic Acids Res* **20**, 997-1003 (1992).
34. Driegen, S. et al. A generic tool for biotinylation of tagged proteins in transgenic mice. *Transgenic Res* **14**, 477-82 (2005).
35. Donehower, L.A. et al. Mice deficient for p53 are developmentally normal but susceptible to spontaneous tumours. *Nature* **356**, 215-21 (1992).
36. van Dijk, T.B. et al. Friend of Prmt1, a novel chromatin target of protein arginine methyltransferases. *Mol Cell Biol* **30**, 260-72 (2010).
37. de Boer, E. et al. Efficient biotinylation and single-step purification of tagged transcription factors in mammalian cells and transgenic mice. *Proc Natl Acad Sci U S A* **100**, 7480-5 (2003).
38. van Dijk, T.B. et al. Stem cell factor induces phosphatidylinositol 3'-kinase-dependent Lyn/Tec/Dok-1 complex formation in hematopoietic cells. *Blood* **96**, 3406-13 (2000).
39. Zuin, J. et al. A cohesin-independent role for NIPBL at promoters provides insights in CdLS. *PLoS Genet* **10**, e1004153 (2014).
40. Barsnes, H., Vizcaino, J.A., Eidhammer, I. & Martens, L. PRIDE Converter: making proteomics data-sharing easy. *Nat Biotechnol* **27**, 598-9 (2009).
41. Rodriguez, P. et al. GATA-1 forms distinct activating and repressive complexes in erythroid cells. *EMBO J* **24**, 2354-66 (2005).
42. Diao, Y. et al. Pax3/7BP is a Pax7- and Pax3-binding protein that regulates the proliferation of muscle precursor cells by an epigenetic mechanism. *Cell Stem Cell* **11**, 231-41 (2012).
43. Fanis, P. et al. Five friends of methylated chromatin target of protein-arginine-methyltransferase[prmt]-1 (chtop), a complex linking arginine methylation to desumoylation. *Mol Cell Proteomics* **11**, 1263-73 (2012).
44. Southwood, C.M., Downs, K.M. & Bieker, J.J. Erythroid Kruppel-like factor exhibits an early and sequentially localized pattern of expression during mammalian erythroid ontogeny. *Dev Dyn* **206**, 248-59 (1996).
45. Drissen, R. et al. The erythroid phenotype of EKLF-null mice: defects in hemoglobin metabolism and membrane stability. *Mol Cell Biol* **25**, 5205-14 (2005).
46. Basu, P. et al. EKLF and KLF2 have compensatory roles in embryonic beta-globin gene expression and primitive erythropoiesis. *Blood* **110**, 3417-25 (2007).
47. Siatecka, M. & Bieker, J.J. The multifunctional role of EKLF/KLF1 during erythropoiesis. *Blood* **118**, 2044-54 (2011).
48. Giraud, G. et al. NLS-tagging: an alternative strategy to tag nuclear proteins. *Nucleic Acids Res* **42**(2014).

49. Quadrini, K.J. & Bieker, J.J. Kruppel-like zinc fingers bind to nuclear import proteins and are required for efficient nuclear localization of erythroid Kruppel-like factor. *J Biol Chem* **277**, 32243-52 (2002).
50. Pandya, K. & Townes, T.M. Basic residues within the Kruppel zinc finger DNA binding domains are the critical nuclear localization determinants of EKLF/KLF-1. *J Biol Chem* **277**, 16304-12 (2002).

Chapter 9

General discussion

Sickle cell disease (SCD) and β-thalassemia constitute the most common monogenic disorders, with approximately 300,000 affected children born each year[1]. The clinical manifestation of these diseases can be alleviated by increased levels of fetal globin (HbF). In adults, HbF levels normally decrease to less than 1%[2], while in individuals with HPFH (hereditary persistence of fetal hemoglobin) it remains elevated throughout the entire life. HbF levels of around 15-20% already ameliorate the clinical severity of the disease phenotype[3,4]. Understanding the exact mechanisms driving the globin switch will help in the development/ discovery of new therapeutics aimed at γ-globin reactivation in SCD and β-thalassemic patients.

Since its discovery, KLF1 has been shown to play a key role in hemoglobin switching and red blood cell maturation in general. Nevertheless, even though KLF1 has been investigated for more than 20 years, a precise overview of its role in primitive and definitive hematopoiesis is still missing. The experiments described in this thesis represent a small step towards understanding the molecular function of KLF1 in globin switching and erythropoiesis.

KLF1 - BCL11A axis in the regulation of β-globin expression.

A unique feature of erythropoiesis is the stage-specific expression of globin chains. KLF1 and BCL11A are undoubtedly major regulators of this process. The discovery of BCL11A as the regulator of the fetal-to-adult switch represented significant progress in the field[5,6], and the finding that KLF1 in turn regulates BCL11A expression[7,8] provided better insight into the function of these two transcription factors during globin switching. *Bcl11a* conditional knock out mice embryos show an increased expression of fetal globin[6], while *Klf1* knock out embryos die around E14.4 due to severe anemia[9,10]. However, the interplay between KLF1 and BCL11A had never been studied in adult erythropoiesis.

Chapter 2 investigates the role of KLF1 and BCL11A in the regulation of globin switching and erythropoiesis in adult mice. Red blood cell development was studied in three different mouse strains: $Klf1^{wt/ko}$ heterozygous knock out, $Bcl11a^{cko/cko}$ single conditional knock out and $Klf1^{wt/ko}$ :: $Bcl11a^{cko/cko}$ compound mutant mice. In addition, these mice carried a human β-globin locus transgene[11].

The hematological parameters in $Klf1^{wt/ko}$, $Bcl11a^{cko/cko}$ and $Klf1^{wt/ko}$:: $Bcl11a^{cko/cko}$ mice were mildly but significantly affected. $Klf1^{wt/ko}$ mice displayed reticulocytosis, in accordance to the observation that human KLF1 haploinsufficiency leads to mild reticulocytosis[7] while $Bcl11a^{cko/cko}$ mice displayed slightly lower RBCs counts. In the compound $Klf1^{wt/ko}$ :: $Bcl11a^{cko/cko}$ mice, these two traits were more prominent. Taken together, these results indicate that these mice suffered from a mild, and not severe, form of anemia, as they were also able to normally respond to stress anemia induced by phenylhydrazine.

In addition, $Bcl11a^{cko/cko}$ and, to a lesser extent, $Klf1^{wt/ko}$ mice showed a higher expression of γ-globin, in accordance with published results[6]. However, the highest expression of γ-globin was observed in compound $Klf1^{wt/ko}$ :: $Bcl11a^{cko/cko}$ mice. Even though γ-globin expression remained

higher in young and adult *Klf1*$^{wt/ko}$ :: *Bcl11a*$^{cko/cko}$ mice compared to control mice, it still decreased after birth in comparison to the fetal stage, indicating that additional silencing mechanisms play a role in γ-globin silencing.
All the available data so far support the notion that an intact KLF1-BCL11A axis is required in order to achieve efficient silencing of embryonic globin genes. KLF1 activates BCL11A expression, and subsequently BCL11A represses the embryonic/fetal globin genes, enhancing KLF1 activity on β-globin. Interestingly, crossing BCL11A knock out mice with an SCD mouse model could rescue the disease pathology of SCD mice[12]. However, although the KFL1/BCL11A axis plays a crucial role in the hemoglobin switch, DNA polymorphisms at the BCL11A, MYB[13] and β-globin loci can explain only 50% of the variation observed in γ-globin levels[14]. Therefore, additional factors have to be involved.

Finding the right target for drug development
Fetal hemoglobin reactivation remains the most promising therapeutic approach for β-hemoglobinopathies, therefore the development of drugs targeting key players in globin switching is important. The properties that a "perfect target" for clinical use should have are: (1) erythroid specific expression, (2) limited targets outside globin genes, (3) high effect on fetal globin expression, (4) high accessibility to the drug. In addition, the drug should be affordable for the majority of patients in the world and, since reactivation of HbF would be required for the lifetime of the patients, chronic use should have acceptable side effects. Nevertheless, incomplete knowledge of molecular mechanisms underlying globin gene regulation has limited the progress on drug discovery. Many factors regulating globin expression that are known until now, besides being involved in hemoglobin switching, exert important effects on erythropoiesis and hematopoiesis in general, limiting their use as possible drug targets. BCL11A, for example, is a strong, specific regulator of HbF in a dose-dependent manner, has few non globin targets in erythroid cells and the knock down does not affect erythropoiesis. However, BCL11A is expressed outside the erythroid lineage, as it is essential for normal B-lymphocyte development[15,16], and also in non-hematopoietic tissues, including the brain[17]. This suggests that a potential drug targeting BCL11A could have an appreciable impact on other cell lineages. In contrast, KLF1 is highly erythroid specific with very limited, if any, role in other lineages[18], but regulates many aspects of erythroid differentiation, besides the γ- to β-globin switching. A better understanding of the mechanisms regulating KLF1 activity could help predict how KLF1 could be targeted to specifically increase HbF production. KLF1 and BCL11A are transcription factors, which have been considered "undruggable" targets, as they are not easily accessible to drugs. New developments in chemical biology may help to develop potential drugs targeting these two factors. Various examples of small molecules directed towards transcription factors are now available. These compounds modulate either protein-protein interactions or DNA-protein interactions[19-21]. It is worthwhile reminding that many of these modulators have originated from *in silico* screens of potential drugs. As examples, structurally related naphthols have been identified as antagonists of c-Myc association with MYC associated factor X (MAX)[22], while *cis* imidazoline inhibit the p53-MDM2 interaction[23]. Additional knowledge of the mechanisms underlying BCL11A and KLF1 activity during erythropoiesis would help find better modulators of hemoglobin switching, also using *in silico* strategies.

KLF1 *Nan* mutation as a tool to understand KLF1 activity in erythroid cells
Besides having a role in the γ- to β- globin switch, KLF1 has other roles in red blood cell development, regulating, for example, the cell cycle[24,25] and membrane skeleton proteins[26,27]. It is hence not surprising that different KLF1 mutations result in different phenotypes in

humans[28]. Among these, the mutation causing Congenital Dyserythropoietic Anemia (CDA) type IV (OMIM 613673)[29] is of particular interest. This disease arises from a glutamate to lysine (E325K) substitution in the second zinc finger of KLF1. It is due to a *de novo* mutation found in heterozygote state in patients. Patients carrying this mutation display high HbF levels and deficiency for CD44 and AQP1. Interestingly, a mouse ethylnitrosourea-induced mutation in KLF1, causing the Neonatal anemia (*Nan*) phenotype[30-32], is found in the homologous position of the CDA mutation, further indicating the importance of this residue. In *Nan* mice, the glutamate is substituted by an aspartate (E339D). Heterozygotes *Nan*/+ mice survive with life-long, severe hemolytic anemia, while homozygote mice die *in utero* around E10. This phenotype stems from the different DNA binding properties that KLF1 *Nan* displays when compared to wild type KLF1. This leads to an impaired expression of a subset of targets of KLF1, including erythrocyte membrane proteins and globins[31]. Moreover, KLF1 *Nan* is believed to have a dominant negative effect on wild type KLF1. To try to understand the molecular mechanisms underlying the *Nan* mutations, ultimately aiming to obtain better insight on the activity of wild type KLF1, we used the *Nan* mouse model carrying one mutant allele, as described in Chapter 3. We studied three different stages of embryonic development (E12.5, E14.5 and E18.5) and found that, already starting at day E12.5, erythropoiesis is impaired in the presence of the mutation, as exemplified by flow cytometry analysis of the erythroid-specific markers CD71 and Ter119, both in fetal livers and in blood. In order to identify deregulated genes in *Nan* erythroid cells, we performed RNA-sequencing experiments. Compared to control samples, we detected defective expression of 782 genes in *Nan* E12.5 fetal livers.

KLF1 and nuclear condensation

Among the deregulated genes upon KLF1 *Nan* mutation, we identified Xpo7, an exportin shown to be involved in enucleation of red blood cells[33]. It was of particular interest, as it could potentially explain why we observed enlarged maturing erythroid cells in *Nan* mice compared to the wild type control. Moreover, we found that *Xpo7* expression is affected by knockout of KLF1. Since KLF1 binds to the *Xpo7* locus, this data suggests that *Xpo7* expression is directly regulated by KLF1. Considering that KLF1 has an important role in the formation of the active chromatin hub in the β-globin locus[34], we investigated the chromatin conformation of the *Xpo7* locus in the presence of the *Nan* mutation. We couldn't detect any major differences between wild type and *Nan* mice. One explanation is that the KFL1 *Nan* mutation does not affect the role of KLF1 in loop formation. If this is true, it means that KLF1 *Nan* is able to recognize its binding sites in the *Xpo7* locus, but is unable to exert its transcriptional activity. Another possibility is that the interactions seen in the *Xpo7* locus are KLF1-independent, hence the presence of KLF1 *Nan* does not affect the chromatin conformation of the locus. To try to address this question, it would be essential to have ChIP (or better ChIP-sequencing) results for KLF1 in KLF1 *Nan* mouse fetal livers. Unfortunately, we experienced some difficulties in performing this technique, probably due to the inadequacy of the KLF1 antibodies for this type of assay. We are currently testing different antibodies combined with different ChIP protocols in order to set up the best conditions for KLF1 isolation. In addition, we might obtain even more insight in this matter by performing a 3C-seq experiment on KLF1 knockout fetal livers. However, in our 3C-seq experiment we identified an erythroid-specific loop between the canonical promoter of the *Xpo7* gene (situated at the beginning of exon 1a) and the exon that produces the erythroid-specific form of *Xpo7* (exon 1b). We speculate that this might help bringing the required factors to the start site of the alternative *Xpo7* promoter in erythroid cells.

In mice with the *Nan* mutation, and hence with *Xpo7* downregulation, enucleation

of erythroid cells still occurs. We did not detect nucleated red cells in the circulation. Nevertheless, the average size of the fetal liver cells and their nuclei was higher compared to the wild type, probably due to impaired nuclear condensation. We additionally showed that the knockdown of *Xpo7* in I/11 cells recapitulated the phenotype seen in mouse fetal liver cells, namely increased cell size and reduced expression of the differentiation markers CD71 and Ter119.

In conclusion, in Chapter 3 we demonstrate that KLF1 has yet another role in erythroid development, regulating nuclear condensation and consequently enucleation, presumably by affecting *Xpo7* expression. In contrast with what was reported by Hattangadi *et al*[33], we couldn't detect any noticeable defect in the enucleation capability of the *Nan* cells, but only in nuclear condensation. This might be due to the difference between the KLF1 *Nan in vivo* model and the *Xpo7* knock down *in vitro* model and/or to the difference in XPO7 protein levels between the two systems.

Enucleation is a complex process that involves asymmetric cell division ending in the extrusion of the pycnotic nucleus surrounded by the plasma membrane[35]. It is thought to represent a crucial step in mammalian evolution, as the absence of the nucleus allowed an increased hemoglobin concentration in the cells[35]. The signaling pathways and proteins behind this process are still largely unknown, although some candidate proteins[36-38] and miRNAs[39-44] have been described in the last years. Moreover, macrophages are also important for enucleation, as they are part of the erythroblastic island[45]. Within these structures, erythroid progenitors and macrophages interact through different protein-protein interactions, allowing for the correct maturation of red blood cells[46-48]. Primary erythroid progenitors can differentiate and enucleate if cultured *in vitro* in the absence of macrophages, but the rate is reduced.

Evidently, many questions regarding the enucleation process remain unanswered. It is known that membrane and cytoskeletal proteins reorganize during erythropoiesis[49,50], but we don't know whether there is a link with the extrusion of the nucleus. Additionally, the exact mechanism by which reticulocytes and nuclei separate remains unknown. The formation of a contractile actin ring suggests a process involving cytoskeletal proteins, but initial electron micrographic studies detected multiple vesicles forming in the region between the extruding nucleus and the reticulocyte[37]. Furthermore, the role of the microenvironment, and macrophages in particular, needs to be further elucidated. Last, but not least, it is not clear how and if chromatin condensation facilitates the enucleation process. According to the data presented in Chapter 3, a possible scenario is that loss of *Xpo7* causes less chromatin condensation and this in turn contributes to the overall *Nan* phenotype. Our analysis of *Nan* mice suggest that enucleation still occurs in the absence of proper nuclear condensation, but it is possible that this leads to abnormal reticulocytes/red blood cells, thus resulting in anemia.

Translational implications of efficient enucleation

Understanding the exact processes behind enucleation is crucial for the *in vitro* production of red blood cells, which would revolutionize blood transfusions. Red blood cell transfusions are, in many cases, life-saving treatments for patients with anemia, either due to underlying genetic disease, as is the case for thalassemia patients, or in the event of sudden blood loss. Globally, around 85 million units of RBCs, derived from 0.5–0.7 liters of donor blood per unit, are transfused each year[51]. Hence, the demand for donor blood is high, but availability is limited and this may cause shortages for patients, in particular for those with rare blood groups. Indeed, complications such as the potential for transmission of (unknown) infectious diseases and adverse immune reactions further emphasizes the need to develop safe and reliable alternatives to donor blood. Ideally, controlled *in vitro* production of fully functional

mature adult RBCs should lead to unlimited supplies of safe and universal transfusion units. The availability of human RBCs derived *in vitro* from hematopoietic stem/progenitor cells (HSPCs) in clinically useful quantities would have a major impact on transfusion medicine worldwide. For the moment, there are no such *in vitro*-produced alternatives to donor-derived human RBCs. Although in recent years the development of systems for production of RBCs *in vitro* has progressed enormously, the diminished capability of cultured erythroid cells to fully enucleate remains one of the major obstacles. Further deciphering and understanding of the biology of erythrocyte enucleation will in all probability result in improving the efficiency of enucleation in *in vitro* cultured erythrocytes derived from various types of stem/progenitor cells. This in turn would be a crucial step towards the generation of RBCs approximating as closely as possible to normal adult RBCs.

The Golgi apparatus in erythropoiesis

The Golgi apparatus is responsible for the modification and sorting of lysosomal, plasma membrane and extracellular proteins[52]. It is therefore of paramount importance for all cell lineages. It is composed of stacks of a variable number of flattened cisternal membranes and associated vesicles[52]. This complex structure is challenging to study. In erythroid cells, the Golgi apparatus is well developed in erythroblasts and occupies about 3% of rat erythroblast cytoplasm, but it is eventually lost during erythroid maturation[53]. Indeed, after the last mitotic division it is reduced to just 1% of the cytoplasm[53]. Nevertheless, it is reasonable to assume that the Golgi apparatus has an important role in the modification and sorting of extracellular and membrane proteins in the mature red blood cell. To date, very little is known about the Golgi function in erythropoiesis and only a few studies have described the role of Golgi proteins in erythroid cells[54-56]. In Chapter 4 of this thesis, we report the identification of a novel Golgi protein, called BSDC1. BSDC1 was originally identified through a screen for BSD domain containing proteins[57], but little is known about its localization and function. BSDC1 probably constitutes a novel, direct target of KLF1, as its expression is affected if KLF1 is mutated (KLF1 *Nan* embryos) or absent (KFL1 knock out embryos). Moreover, KLF1 was shown to bind to the promoter region of BSDC1, together with GATA1, NF-E2 and p300[58]. In order to get more insight in the possible role of BSDC1 in erythropoiesis, we performed streptavidin pull-downs followed by MS analysis. To our surprise, we identified GOLGB1, a golgin present on the rim of the Golgi, as BSDC1 interaction partner, suggesting that BSDC1 might be localized in the Golgi. Through immunofluorescence experiments we observed that BSDC1 accumulates in a perinuclear area corresponding to the Golgi apparatus where it co-localizes with GOLGB1. It is still unclear whether BSDC1 localization in the Golgi is dependent on the presence of GOLGB1. Knocking out *GolgB1* using CRISP/CAS9 technology and monitoring BSDC1 localization would provide a definitive answer to this question.

How to study BSDC1 function in the Golgi.

Even though BSDC1 cellular localization is now ascertained, its role still remains unclear. BSDC1 Golgi localization suggests that it functions either in the structure/positioning of the Golgi or in the trafficking of cargoes through the Golgi. To try to understand if and how BSDC1 contributes to Golgi function, we checked for ERMAP, a cell surface membrane protein, localization upon BSDC1 knock down. If BSDC1 would be important for trafficking through the Golgi, then ERMAP localization would be affected. Unexpectedly, we didn't detect any difference in the ERMAP localization between controls and BSDC1 knockdown cells. There are several explanations for this result. It could be that the residual BSDC1 level is sufficient to fulfill the functions of the protein. Another possibility is that BSDC1, as many other golgins, has a redundant role. In that case the single knockdown is not expected to

phenotypically affect the cell. Finally, BSDC1 might be involved in the modification and not in the trafficking of cargo proteins. Clearly, in order to elucidate the role of BSDC1 in the Golgi and specifically in erythroid development, more experiments are needed. An interesting starting point would be to study BSDC1 behavior in the presence of brefeldin A (BFA), a compound that disrupts Golgi structure and function, and Nocodazole, a compound that disrupt the function but not the structure of the Golgi apparatus. Such experiments would establish whether BSDC1 behaves like other Golgi proteins, which are dispersed in the cytoplasm during mitosis[55,56] (BFA treatment), but still associated to the membrane if the Golgi structure is kept intact (Nocodazole treatment).

In the Golgi, BSDC1 might associate with GOLGB1 via an acidic domain found at its C-terminus. An acidic domain is also found at the C-terminus of another Golgi membrane protein[59], called p115, a protein found in association with *cis*-golgi elements and in the intermediate compartment residing between the Golgi and the ER[60,61]. p115 interacts with both GOLGB1 and GM130[61-63], creating ternary giantin-p115-GM130 complexes which link vesicles and cisternae[64] or independent p115-GOLGB1 or p115-GM130 membrane tethering events[65]. We plan to explore the potential BSDC1- GOLGB1 interaction via the acidic domain by creating a mutant that lacks the C-terminal part of the protein and by inducing point mutations that either mimic or abolish the phosphorylation of Serine 415, as the interactions between Golgi proteins are destroyed by phosphorylation[66,67.] If the binding to GOLGB1 occurs in a manner similar to p115, this would suggest that BSDC1 holds a role in trafficking through the Golgi.

Since BSDC1 is located in the Golgi, and given the fact that the Golgi disappears during erythroid maturation, an emerging question is why BSDC1 expression increases during erythroid maturation. At this point, the most plausible scenario would be that BSDC1 and the Golgi apparatus hold an important role in erythroid maturation as a numerous protein modifications are needed to produce a fully mature red blood cell. Alternatively, besides in the Golgi apparatus, BSDC1 might have an additional function in the cell. Imaging experiments of undifferentiated and differentiated erythroid cells will show whether BSDC1 localization in the cytoplasm changes upon differentiation. This would be a key experiment in resolving if BSDC1 exerts its role fully as a member of the Golgi apparatus, or possibly also at other locations in the cell.

KLF1 interacting partners isolation

As previously discussed, KLF1 has transcriptional activation activity. Nevertheless, KLF1 is not working alone and is thought to have different binding partners during different stages of erythropoiesis. In the last years many KLF1 interacting partners have been identified, including the p62 subunit of TFIIH[68], the TAF9 subunit of TFIID[69], CBP/p300[68,70], histone H3[71] and the BRG1 subunit[72-75] of the SWI/SNF complex[76,77]. Despite these efforts, a comprehensive and unbiased overview of KLF1 interacting proteins is lacking. In order to address this issue, we used an in-house developed system for isolating transcription factor complexes[78,79]. As described in Chapter 5, we performed streptavidin pull-downs coupled with MS analysis from immortalized MFL cells and MEL cells containing biotinylated KLF1. Through this approach, we identified PAXBP1 as a potential KLF1 interacting protein, but we were not able to confirm it as a true interactor of KLF1. KLF1 is present in both primitive and definitive erythroid cells and one open question that still needs to be answered is how it differentially regulates genes during primitive and definitive erythropoiesis. Interactions with different partners, along with different post-translational modifications of KLF1 or its interactors, or redundant role with KLF2, presumably hold a key role in directing and fine tuning KLF1 activity. With our protocol we succeeded in purifying Bio-HA-KLF1. However, the

fraction that is pulled down might represent free KLF1, and not chromatin-bound KLF1. In order to obtain a more detailed overview of *in vivo* chromatin KLF1 complexes, a purification method for KLF1 that works efficiently and consistently is needed, and we intend to intensify our efforts to achieve this. An initial approach would entail changing the position of the tag, inserting it near the NLS[80], or at the C-terminal end downstream of the DNA-binding domain[81]. Another way to increase the quality of isolated Bio-HA-KLF1 complexes would involve chemical crosslinking, hence stabilizing all protein-protein interactions. Though this would effectively freeze even transient KLF1 interactors, the use of very stringent washing buffers would be needed to minimize the expected increase of background signal due to crosslinking. Finally, although the identification of novel KLF1 partners may be hampered by the transient nature of their interactions or limitations of the purification method, the quantity and purity of isolated Bio-HA-KLF1 could be sufficient to perform a pull-down under denaturing conditions in order to identify post-translational modifications of KLF1 at different stages of erythroid development.

Improving our purification protocol would be a crucial step toward finding novel KLF1 interacting proteins. Discovering such KLF1 interacting partners can be clinically significant as such interactors might be more suitable candidates as targets for drug therapy. Epigenetic modifiers, for example, represent an attractive candidate as 5-azacytidine has been used to reactivate HbF expression in baboons and human[82,83]. With more knowledge on the regulatory pathways involved in HbF silencing and on KLF1 interacting proteins, new epigenetic modifiers may come forward as potential targets. Drug design may also help to decrease off target effects. Since epigenetic modifiers generally have a role in multiple hematopoietic lineages and beyond, this raises concerns about the therapeutic adequacy of these targets. Therefore, the identification of post-translational modifying enzymes of KLF1 or its partners might provide alternative, less pleiotropic and thus easier controllable, targets for drug development.

Concluding remarks

Despite extensive investigation on erythropoiesis and globin switching, a comprehensive understanding of these processes is still missing. To date, hydroxyurea is the only approved drug for SCD. Nevertheless, its usage its limited by the selective effectiveness on patients, adverse effects and limited benefits for patients with β-thalassemia. A therapy combining two or more drugs, each with a distinct target, may be an efficient method for the induction of very high levels of HbF while limiting adverse effects. When the proteins that influence HbF levels will be identified comprehensively, the translation of this knowledge to the clinic may provide a breakthrough in treatment strategies for SCD and β-thalassemic patients. Such a feat would require the combined efforts of clinicians, molecular biologists, chemists, and pharmacologists.

REFERENCES

1. Angastiniotis, M. & Modell, B. Global epidemiology of hemoglobin disorders. *Ann N Y Acad Sci* **850**, 251-69 (1998).
2. Leonova, J. et al. Variability in the fetal hemoglobin level of the normal adult. *Am J Hematol* **53**, 59-65 (1996).
3. Noguchi, C.T., Rodgers, G.P., Serjeant, G. & Schechter, A.N. Levels of fetal hemoglobin necessary for treatment of sickle cell disease. *N Engl J Med* **318**, 96-9 (1988).
4. Platt, O.S. et al. Pain in sickle cell disease. Rates and risk factors. *N Engl J Med* **325**, 11-6 (1991).
5. Sankaran, V.G. et al. Human fetal hemoglobin expression is regulated by the developmental stage-specific repressor BCL11A. *Science* **322**, 1839-42 (2008).
6. Sankaran, V.G. et al. Developmental and species-divergent globin switching are driven by BCL11A. *Nature* **460**, 1093-7 (2009).
7. Borg, J. et al. Haploinsufficiency for the erythroid transcription factor KLF1 causes hereditary persistence of fetal hemoglobin. *Nat Genet* **42**, 801-5 (2010).
8. Zhou, D., Liu, K., Sun, C.W., Pawlik, K.M. & Townes, T.M. KLF1 regulates BCL11A expression and gamma- to beta-globin gene switching. *Nat Genet* **42**, 742-4 (2010).
9. Nuez, B., Michalovich, D., Bygrave, A., Ploemacher, R. & Grosveld, F. Defective haematopoiesis in fetal liver resulting from inactivation of the EKLF gene. *Nature* **375**, 316-8 (1995).
10. Perkins, A.C., Sharpe, A.H. & Orkin, S.H. Lethal beta-thalassaemia in mice lacking the erythroid CACCC-transcription factor EKLF. *Nature* **375**, 318-22 (1995).
11. de Krom, M., van de Corput, M., von Lindern, M., Grosveld, F. & Strouboulis, J. Stochastic patterns in globin gene expression are established prior to transcriptional activation and are clonally inherited. *Mol Cell* **9**, 1319-26 (2002).
12. Xu, J. et al. Correction of sickle cell disease in adult mice by interference with fetal hemoglobin silencing. *Science* **334**, 993-6 (2011).
13. Thein, S.L. et al. 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**, 11346-51 (2007).
14. Thein, S.L., Menzel, S., Lathrop, M. & Garner, C. Control of fetal hemoglobin: new insights emerging from genomics and clinical implications. *Hum Mol Genet* **18**, R216-23 (2009).
15. Liu, P. et al. Bcl11a is essential for normal lymphoid development. *Nat Immunol* **4**, 525-32 (2003).
16. Yu, Y. et al. Bcl11a is essential for lymphoid development and negatively regulates p53. *J Exp Med* **209**, 2467-83 (2012).
17. Kuo, T.Y. & Hsueh, Y.P. Expression of zinc finger transcription factor Bcl11A/Evi9/CTIP1 in rat brain. *J Neurosci Res* **85**, 1628-36 (2007).
18. Siatecka, M. & Bieker, J.J. The multifunctional role of EKLF/KLF1 during erythropoiesis. *Blood* **118**, 2044-54 (2011).
19. Arndt, H.D. Small molecule modulators of transcription. *Angew Chem Int Ed Engl* **45**, 4552-60 (2006).
20. Berg, T. Inhibition of transcription factors with small organic molecules. *Curr Opin Chem Biol* **12**, 464-71 (2008).
21. Majmudar, C.Y. & Mapp, A.K. Chemical approaches to transcriptional regulation. *Curr Opin Chem Biol* **9**, 467-74 (2005).
22. Xu, Y. et al. A credit-card library approach for disrupting protein-protein interactions. *Bioorg Med Chem* **14**, 2660-73 (2006).
23. Vassilev, L.T. et al. In vivo activation of the p53 pathway by small-molecule antagonists of MDM2. *Science* **303**, 844-8 (2004).
24. Tallack, M.R., Keys, J.R., Humbert, P.O. & Perkins, A.C. EKLF/KLF1 controls cell cycle entry via direct regulation of E2f2. *J Biol Chem* **284**, 20966-74 (2009).
25. Pilon, A.M. et al. Failure of terminal erythroid differentiation in EKLF-deficient mice is associated with cell cycle perturbation and reduced expression of E2F2. *Mol Cell Biol* **28**, 7394-401 (2008).
26. Drissen, R. et al. The erythroid phenotype of EKLF-null mice: defects in hemoglobin metabolism and membrane stability. *Mol Cell Biol* **25**, 5205-14 (2005).
27. Nilson, D.G., Sabatino, D.E., Bodine, D.M. & Gallagher, P.G. Major erythrocyte membrane protein

genes in EKLF-deficient mice. *Exp Hematol* **34**, 705-12 (2006).
28. Borg, J., Patrinos, G.P., Felice, A.E. & Philipsen, S. Erythroid phenotypes associated with KLF1 mutations. *Haematologica* **96**, 635-8 (2011).
29. Arnaud, L. et al. A dominant mutation in the gene encoding the erythroid transcription factor KLF1 causes a congenital dyserythropoietic anemia. *Am J Hum Genet* **87**, 721-7 (2010).
30. Lyon, M.F.G., P.H.; Loutit, J.F., Peters, J. . Dominant hemolytic anemia. *Mouse News Letter* **68:68**(1983).
31. Siatecka, M. et al. Severe anemia in the Nan mutant mouse caused by sequence-selective disruption of erythroid Kruppel-like factor. *Proc Natl Acad Sci U S A* **107**, 15151-6 (2010).
32. Heruth, D.P. et al. Mutation in erythroid specific transcription factor KLF1 causes Hereditary Spherocytosis in the Nan hemolytic anemia mouse model. *Genomics* **96**, 303-7 (2010).
33. Hattangadi, S.M. et al. Histones to the cytosol: exportin 7 is essential for normal terminal erythroid nuclear maturation. *Blood* **124**, 1931-40 (2014).
34. Drissen, R. et al. The active spatial organization of the beta-globin locus requires the transcription factor EKLF. *Genes Dev* **18**, 2485-90 (2004).
35. Ji, P., Murata-Hori, M. & Lodish, H.F. Formation of mammalian erythrocytes: chromatin condensation and enucleation. *Trends Cell Biol* **21**, 409-15 (2011).
36. Ji, P., Jayapal, S.R. & Lodish, H.F. Enucleation of cultured mouse fetal erythroblasts requires Rac GTPases and mDia2. *Nat Cell Biol* **10**, 314-21 (2008).
37. Koury, S.T., Koury, M.J. & Bondurant, M.C. Cytoskeletal distribution and function during the maturation and enucleation of mammalian erythroblasts. *J Cell Biol* **109**, 3005-13 (1989).
38. Lee, J.C. et al. Mechanism of protein sorting during erythroblast enucleation: role of cytoskeletal connectivity. *Blood* **103**, 1912-9 (2004).
39. Zhao, G., Yu, D. & Weiss, M.J. MicroRNAs in erythropoiesis. *Curr Opin Hematol* **17**, 155-62 (2010).
40. Dore, L.C. et al. A GATA-1-regulated microRNA locus essential for erythropoiesis. *Proc Natl Acad Sci U S A* **105**, 3333-8 (2008).
41. Fu, Y.F. et al. Mir-144 selectively regulates embryonic alpha-hemoglobin synthesis during primitive erythropoiesis. *Blood* **113**, 1340-9 (2009).
42. Pase, L. et al. miR-451 regulates zebrafish erythroid maturation in vivo via its target gata2. *Blood* **113**, 1794-804 (2009).
43. Rasmussen, K.D. et al. The miR-144/451 locus is required for erythroid homeostasis. *J Exp Med* **207**, 1351-8 (2010).
44. Zhang, L., Flygare, J., Wong, P., Lim, B. & Lodish, H.F. miR-191 regulates mouse erythroblast enucleation by down-regulating Riok3 and Mxi1. *Genes Dev* **25**, 119-24 (2011).
45. Chasis, J.A. & Mohandas, N. Erythroblastic islands: niches for erythropoiesis. *Blood* **112**, 470-8 (2008).
46. Hanspal, M., Smockova, Y. & Uong, Q. Molecular identification and functional characterization of a novel protein that mediates the attachment of erythroblasts to macrophages. *Blood* **92**, 2940-50 (1998).
47. Hanspal, M. & Hanspal, J.S. The association of erythroblasts with macrophages promotes erythroid proliferation and maturation: a 30-kD heparin-binding protein is involved in this contact. *Blood* **84**, 3494-504 (1994).
48. Soni, S. et al. Absence of erythroblast macrophage protein (Emp) leads to failure of erythroblast nuclear extrusion. *J Biol Chem* **281**, 20181-9 (2006).
49. Simpson, C.F. & Kling, J.M. The mechanism of denucleation in circulating erythroblasts. *J Cell Biol* **35**, 237-45 (1967).
50. Skutelsky, E. & Danon, D. An electron microscopic study of nuclear elimination from the late erythroblast. *J Cell Biol* **33**, 625-35 (1967).
51. Carson, J.L. et al. Red blood cell transfusion: a clinical practice guideline from the AABB*. *Ann Intern Med* **157**, 49-58 (2012).
52. Alberts, B. Molecular biology of the cell. *5th edition, Garland Science.2008*.
53. Heynen, M.J. & Verwilghen, R.L. A quantitative ultrastructural study of normal rat erythroblasts and reticulocytes. *Cell Tissue Res* **224**, 397-408 (1982).
54. Ghosh, S., Cox, K.H. & Cox, J.V. Chicken erythroid AE1 anion exchangers associate with the cytoskeleton during recycling to the Golgi. *Mol Biol Cell* **10**, 455-69 (1999).

55. Stroissnigg, H. et al. FIP-2, an IkappaB-kinase-gamma-related protein, is associated with the Golgi apparatus and translocates to the marginal band during chicken erythroblast differentiation. *Exp Cell Res* **278**, 133-45 (2002).
56. Beck, K.A., Buchanan, J.A., Malhotra, V. & Nelson, W.J. Golgi spectrin: identification of an erythroid beta-spectrin homolog associated with the Golgi complex. *J Cell Biol* **127**, 707-23 (1994).
57. Doerks, T., Huber, S., Buchner, E. & Bork, P. BSD: a novel domain in transcription factors and synapse-associated proteins. *Trends Biochem Sci* **27**, 168-70 (2002).
58. Su, M.Y. et al. Identification of biologically relevant enhancers in human erythroid cells. *J Biol Chem* **288**, 8433-44 (2013).
59. Waters, M.G., Clary, D.O. & Rothman, J.E. A novel 115-kD peripheral membrane protein is required for intercisternal transport in the Golgi stack. *J Cell Biol* **118**, 1015-26 (1992).
60. Bannykh, S.I., Nishimura, N. & Balch, W.E. Getting into the Golgi. *Trends Cell Biol* **8**, 21-5 (1998).
61. Nelson, D.S. et al. The membrane transport factor TAP/p115 cycles between the Golgi and earlier secretory compartments and contains distinct domains required for its localization and function. *J Cell Biol* **143**, 319-31 (1998).
62. Sonnichsen, B. et al. A role for giantin in docking COPI vesicles to Golgi membranes. *J Cell Biol* **140**, 1013-21 (1998).
63. Nakamura, N., Lowe, M., Levine, T.P., Rabouille, C. & Warren, G. The vesicle docking protein p115 binds GM130, a cis-Golgi matrix protein, in a mitotically regulated manner. *Cell* **89**, 445-55 (1997).
64. Lesa, G.M., Seemann, J., Shorter, J., Vandekerckhove, J. & Warren, G. The amino-terminal domain of the golgi protein giantin interacts directly with the vesicle-tethering protein p115. *J Biol Chem* **275**, 2831-6 (2000).
65. Linstedt, A.D. et al. Binding relationships of membrane tethering components. The giantin N terminus and the GM130 N terminus compete for binding to the p115 C terminus. *J Biol Chem* **275**, 10196-201 (2000).
66. Lowe, M. et al. Cdc2 kinase directly phosphorylates the cis-Golgi matrix protein GM130 and is required for Golgi fragmentation in mitosis. *Cell* **94**, 783-93 (1998).
67. Sohda, M., Misumi, Y., Yano, A., Takami, N. & Ikehara, Y. Phosphorylation of the vesicle docking protein p115 regulates its association with the Golgi membrane. *J Biol Chem* **273**, 5385-8 (1998).
68. Mas, C. et al. Structural and functional characterization of an atypical activation domain in erythroid Kruppel-like factor (EKLF). *Proc Natl Acad Sci U S A* **108**, 10484-9 (2011).
69. Sengupta, T., Cohet, N., Morle, F. & Bieker, J.J. Distinct modes of gene regulation by a cell-specific transcriptional activator. *Proc Natl Acad Sci U S A* **106**, 4213-8 (2009).
70. Zhang, W., Kadam, S., Emerson, B.M. & Bieker, J.J. Site-specific acetylation by p300 or CREB binding protein regulates erythroid Kruppel-like factor transcriptional activity via its interaction with the SWI-SNF complex. *Mol Cell Biol* **21**, 2413-22 (2001).
71. Sengupta, T., Chen, K., Milot, E. & Bieker, J.J. Acetylation of EKLF is essential for epigenetic modification and transcriptional activation of the beta-globin locus. *Mol Cell Biol* **28**, 6160-70 (2008).
72. Brown, R.C. et al. Distinct domains of erythroid Kruppel-like factor modulate chromatin remodeling and transactivation at the endogenous beta-globin gene promoter. *Mol Cell Biol* **22**, 161-70 (2002).
73. Desgardin, A.D. et al. Regulation of delta-aminolevulinic acid dehydratase by kruppel-like factor 1. *PLoS One* **7**, e46482 (2012).
74. Kim, S.I., Bresnick, E.H. & Bultman, S.J. BRG1 directly regulates nucleosome structure and chromatin looping of the alpha globin locus to activate transcription. *Nucleic Acids Res* **37**, 6019-27 (2009).
75. Perrine, S.P., Mankidy, R., Boosalis, M.S., Bieker, J.J. & Faller, D.V. Erythroid Kruppel-like factor (EKLF) is recruited to the gamma-globin gene promoter as a co-activator and is required for gamma-globin gene induction by short-chain fatty acid derivatives. *Eur J Haematol* **82**, 466-76 (2009).
76. Armstrong, J.A., Bieker, J.J. & Emerson, B.M. A SWI/SNF-related chromatin remodeling complex, E-RC1, is required for tissue-specific transcriptional regulation by EKLF in vitro. *Cell* **95**, 93-104

(1998).
77. Kadam, S. et al. Functional selectivity of recombinant mammalian SWI/SNF subunits. *Genes Dev* **14**, 2441-51 (2000).
78. Grosveld, F. et al. Isolation and characterization of hematopoietic transcription factor complexes by in vivo biotinylation tagging and mass spectrometry. *Ann N Y Acad Sci* **1054**, 55-67 (2005).
79. van Dijk, T.B. et al. Friend of Prmt1, a novel chromatin target of protein arginine methyltransferases. *Mol Cell Biol* **30**, 260-72 (2010).
80. Giraud, G. et al. NLS-tagging: an alternative strategy to tag nuclear proteins. *Nucleic Acids Res* **42**(2014).
81. Engelen, E. et al. Sox2 cooperates with Chd7 to regulate genes that are mutated in human syndromes. *Nat Genet* **43**, 607-11 (2011).
82. DeSimone, J., Heller, P., Hall, L. & Zwiers, D. 5-Azacytidine stimulates fetal hemoglobin synthesis in anemic baboons. *Proc Natl Acad Sci U S A* **79**, 4428-31 (1982).
83. Dover, G.J., Charache, S.H., Boyer, S.H., Talbot, C.C., Jr. & Smith, K.D. 5-Azacytidine increases fetal hemoglobin production in a patient with sickle cell disease. *Prog Clin Biol Res* **134**, 475-88 (1983).

Appendum

Summary

Of all the cells present in the blood, erythrocytes are the most abundant, occupying 45% of the total blood volume. The average lifespan of erythrocytes in humans is 120 days, after which they become senescent and are removed via phagocytosis by macrophages, mainly in the spleen and liver. New red blood cells need to be produced every day to ensure that all body tissues have enough oxygenation for their physiological function. Therefore, erythropoiesis has to be efficiently and tightly regulated in order to release a suitable amount of red blood cells into the blood stream. Any problem in this process can lead to severe diseases, of which thalassemia and sickle cell anemia are the most prominent, and even death of the organism. In **Chapter 1** I briefly introduce hematopoiesis, focusing in particular on erythropoiesis. Given that the main focus of my studies was the globin switching process, namely the silencing of fetal γ-globin and the concomitant activation of adult β-globin, I present a detailed overview of the topic, together with a summary of the state of the art about the three main transcription factors involved in erythropoiesis: GATA1, BCL11A, and KLF1.

KLF1 is undoubtedly a master regulator of erythropoiesis at various levels. Its importance is exemplified by KLF1 ko/ko mice, which die *in utero* at day E14.5 due to severe anemia. KLF1 is essential for the switching from γ- to β- globin, directly activating the expression of β-globin and indirectly repressing γ-globin via the induction of BCL11A. In **Chapter 2** I present in details our findings on the relationship between BCL11A and KLF1, coupling BCL11A conditional knockout with haploinsufficiency for KLF1 in adult mice. Our results demonstrate that such mice display mild compensated anemia and have increased levels of γ-globin, although these expression levels are decreased in adulthood. Our data not only confirms the important role of the KLF1-BCL11A axis during globin switching, but also indicates that additional regulators of globin switching have a role in γ-globin silencing.

Besides its role in β-globin regulation, KLF1 activates a large number of genes playing roles in virtually every aspect of red blood cell development. As expected, different KFL1 mutations can result in different phenotypes, in some cases leading to failure in completing erythroid development and hence to the development of anemia. In **Chapter 3** I introduce our results on the effect of the KLF1 *Nan* (E339D) mutation on embryonic development, utilizing the *Nan* mouse model. KLF1 *Nan* mice display impaired erythroid maturation, as demonstrated by flow cytometry analysis. By performing expression profiling experiments we identified approximately 780 differentially regulated genes between wild type and *Nan* fetal livers. We studied several of those gene in more detail, including XPO7, an exportin with a function in enucleation of the erythroid progenitors. Our data indicate that XPO7 contributes to the phenotype seen in the *Nan* mice, as its knock down in differentiating wildtype erythroid progenitor cells also leads to larger cells.

Another interesting gene downregulated in the KLF1 *Nan* mutant is BSDC1, a BSD domain containing protein with unknown function. In **Chapter 4** I put forward our results on BSDC1, which is involved in late erythropoiesis as its expression increases during erythroid maturation. Surprisingly, we demonstrate that BSDC1 mainly localizes to the Golgi apparatus. This indicates that BSDC1 might have a role in the positioning/localization of the Golgi in the cells or in the trafficking through the Golgi. We are currently performing experiments that will help us shed light on the role of this protein and of the Golgi apparatus more broadly during the final stages of erythropoiesis.

As KLF1 constitutes a crucial transcriptional factor in erythropoiesis, the mechanisms by which it exerts its regulatory effect are of particular interest. In order to better understand the mechanisms governing KLF1 function, we reasoned that identifying novel KLF1 partners would be a step in the right direction. Although the past years have seen the discovery of some potential KLF1 binding partners, a comprehensive list of KLF1 interactors is still missing. In **Chapter 5** I describe our attempts to identify such KLF1 interacting proteins, by utilizing a purification protocol based on streptavidin-dependent isolation of endogenously biotinylated KLF1 followed by mass spectrometry. Despite the fact that our protocol yielded several potential KLF1 partners, validation of the interaction of some of these proteins with KLF1 proved to be difficult. Implementing further improvements in our protocol, as well as potentially testing a few of the remaining potential interaction partners, will hopefully provide a clearer picture of the KLF1 complexes.

Finally, in **Chapter 6** I discuss the findings of this thesis in a broader context, given that erythropoiesis is a multifaceted process requiring multiple regulators. Moreover, I describe what I consider as future perspectives of this work and questions still awaiting an answer, as understanding erythropoiesis in detail can lead to the development of improved treatments for patients with hemoglobinopathies.

Samenvatting

Van alle cellen aanwezig in bloed zijn de erytrocyten de meest voorkomende, ze vormen 45% van het bloed volume. De gemiddelde levensduur van humane erytrocyten is 120 dagen, waarna ze senescent worden en via fagocytose worden verwijderd uit het bloed; dit gebeurt voornamelijk in de milt en de lever. Nieuwe rode bloedcellen moeten elke dag worden geproduceerd om te waarborgen dat alle lichaamsweefsels genoeg zuurstof krijgen voor hun fysiologische functie. Daarvoor moet de erytropoëse efficiënt en nauwkeurig gereguleerd worden, zodat er genoeg rode bloedcellen in de bloedstroom komen. Elk probleem in dit proces kan leiden tot ernstige ziekten, waarvan thalassemia en sikkelcelziekte de meest voorkomende zijn. In **Hoofdstuk 1** beschrijf ik hematopoëse globaal en erytropoëse in detail. Gegeven dat het belangrijkste onderwerp van mijn studie de globine schakeling was, namelijk de inactivatie van het foetale γ-globine en de gelijktijdige activatie van het volwassen β-globine, presenteer ik een gedetailleerd overzicht van dit onderwerp, samen met een overzicht van de huidige kennis over drie belangrijke transcriptiefactoren betrokken bij de erytropoëse: GATA1, BCL11A en KLF1.

KLF1 is zonder twijfel een hoofdregulator van erytropoëse op verschillende niveaus. Het belang van KLF1 is bewezen door het bestuderen van KLF1 'knockout' muizen, die tijdens de ontwikkeling overlijden aan de gevolgen van ernstige anemie. KLF1 is essentieel voor de schakeling van γ- naar β-globine, direct door de activatie van β-globine en indirect door de remming van γ-globine via de activatie van BCL11A. In **Hoofdstuk 2** presenteer ik onze bevindingen aangaande de relatie tussen BCL11A en KLF1. We hebben dit bestudeerd door de BCL11A conditionele knockout te koppelen met haploinsufficiëntie voor KLF1 in volwassen muizen. Onze resultaten laten zien dat deze muizen een milde gecompenseerde anemie hebben en verhoogde γ-globine niveaus, al nemen deze expressie niveaus geleidelijk af bij volwassen dieren. Onze resultaten bevestigen niet alleen het belang van KLF1-BCL11A gedurende de globine schakeling, maar tonen ook aan dat additionele regulatoren een rol spelen bij γ-globine inactivering.

Naast de rol in β-globine regulatie, activeert KLF1 vele genen die nodig zijn voor vrijwel elk aspect van de erytropoëse. Verschillende mutaties in het KLF1 eiwit resulteren daardoor in verschillende fenotypes, in sommige gevallen leidend tot falen van de erytroïde ontwikkeling en anemie. In **Hoofdstuk 3** beschrijf ik de resultaten van ons onderzoek naar de effecten van de KLF1 Nan (p.E339D) mutatie op de foetale ontwikkeling. KLF1 Nan muizen hebben vertraagde erytroïde rijping, dit hebben we aangetoond met flowcytrometrie analyse. Met behulp van expressieprofileringsexperimenten hebben we ongeveer 780 differentieel gereguleerde genen geïdentificeerd tussen wildtype en Nan foetale levers. Een aantal genen hebben we meer in detail bestudeerd, waaronder XPO7, een exportin met functie in de enucleatie van erytroïde voorlopercellen. Onze experimenten tonen aan dat verlaagde expressie van XPO7 bijdraagt aan het fenotype van de Nan muizen.

Een ander interessant omlaag gereguleerd gen in de KLF1 Nan mutant is de BSDC1, een BSD domein bevattend eiwit met nog onbekende functie. In **Hoofdstuk 4** beschrijf ik ons onderzoek naar de functie van BSDC1. De expressie van dit eiwit neemt sterk toe tijdens de erytroïde rijping. Het bleek dat BDSC1 voornamelijk lokaliseert in het Golgi apparaat van de cellen. Dit suggereert dat BSDC1 een rol speelt in de positionering/lokalisatie van het Golgi apparaat in de cellen of het transport en modificatie van eiwitten door het Golgi apparaat. Momenteel voeren we experimenten uit die een breder licht zullen werpen op de rol van

BSDC1 en de rol van het Golgi apparaat gedurende de laatste fasen van de erytropoëse.

KLF1 is een cruciale transcriptie factor voor de erytropoëse, de mechanismen waarmee het zijn regulerende effecten uitvoert zijn echter maar gedeeltelijk bekend. Om deze regulerende mechanismen beter te begrijpen, beredeneerden wij dat identificatie van nieuwe KLF1 interactiepartners een stap in de goede richting zou zijn. In de afgelopen jaren zijn een aantal potentiële KLF1 interactiepartners beschreven, maar een complete lijst met KLF1 interactiepartners ontbreekt. In **Hoofdstuk 5** beschrijf ik onze pogingen om dergelijke KLF1 interactiepartners te identificeren, gebruikmakend van een zuiveringsprotocol gebaseerd op streptavidine-afhankelijke isolatie van endogeen gebiotinyleerd KLF1 gevolgd door massaspectrometrie. Hoewel deze benadering meerdere potentiële KLF1 interactiepartners heeft opgeleverd, bleek validatie van de interacties met KLF1 een moeizaam proces en moeten we concluderen dat veel van de gevonden potentiële interactiepartners vals-positieven zijn. Implementatie van verbeteringen in het zuiveringsprotocol, evenals het testen van een aantal overgebleven potentiële interactiepartners, zal ons hopelijk in de toekomst een duidelijker beeld kunnen geven van de KLF1 eiwitcomplexen tijdens de erytroïde differentiatie.

Gegeven dat erytropoëse een veelzijdig proces is met een veelvoud aan regulerende mechanismen, plaats ik in **Hoofdstuk 6** de bevindingen beschreven in dit proefschrift in een bredere context. Verder bediscussieer ik mogelijkheden voor vervolgonderzoek om belangrijke openstaande vragen over erytropoiese te beantwoorden. Naar verwachting zal deze nieuwe kennis bijdragen aan de ontwikkeling van betere behandelingen voor patiënten met hemoglobinopathieën.

Curriculum Vitae

PERONAL INFORMATION:

Name: Ileana Cantù
Date of Birth: 13th October 1986
Place of Birth: Seregno, Milano, Italy
Nationality: Italian

EDUCATION

2011- 2015 **PhD program at Erasmus Medical center**, Department of Cell Biology, Promotor Prof dr. Sjaak Philipsen, Erasmus Medical Center, Rotterdam, The Netherlands

2008-2011 **Master in Biology**, Specialization Molecular Biology, Universitá degli Studi di Milano-Bicocca, Milano, Italy
Grade: 110/110 *cum laude*

2005-2008 **Bachelor in Biotechnology**, Specialization Health Biotechnology, Universitá degli Studi di Milano-Bicocca, Milano, Italy
Grade: 110/110 *cum laude*

RESEARCH EXPERIENCE

2011- 2015 **PhD student**, Department of Cell Biology, Promotor: Prof dr. Sjaak Philipsen, Erasmus Medical Center, Rotterdam, The Netherlands
Subject: The role of KLF1 in erythropoiesis

2009-2011 **Internship for Master Degree,** Department of Biotechnology and Biosciences, Supervisor Prof. dr. Antonella Ronchi, Università degli Studi di Milano-Bicocca, Milan, Italy
Subject: Regulation of the transcription factor Sox6 in erythroid cells

2009-2011 **Internship for Bachelor Degree,** Department of Biohemistry and Molecular Biology, Supervisor Prof. dr. Pere Suau, Universitat Autonoma de Barcelona, Barcelona, Spain.
Subject: Mutagenesis, expression and purification of histone H1

PUBLICATIONS

Cantù,I., Gillemans, N., Basu, S., Demmers,J., Galjart, N., Lowe, M., Philipsen, S., and van Dijk, T.B. *BSDC1, an unexpected link between KLF1 and the Golgi apparatus*. (Manuscript in preparation).

Cantú, I., van de Werken, H.J.G, Gillemans, N., Stadhouders, R., Heshusius, S., Ozgur, Z. , van Ijcken, W.F., Grosveld, F., von Lindern, M., Philipsen, S., van Dijk, T.B. T*he KLF1 Nan mutation deregulates erythroid maturation during fetal development*. (Manuscript submitted)

Cantú, I., Philipsen, S. *Flicking the switch: adult hemoglobin expression in erythroid cells derived from cord blood and human induced pluripotent stem cells* Haematologica. 2014 Nov; 99(11): 1647–1649.

Esteghamat, F., Gillemans, N., Bilic, I., van den Akker, E., **Cantù, I.**, van Gent, T., Klingmüller, U., van Lom, K., von Lindern ,M., Grosveld, F., Bryn van Dijk, T., Busslinger, M.,Philipsen, S. *Erythropoiesis and globin switching in compound Klf1::Bcl11a mutant mice.* Blood. 2013 Mar 28;121(13):2553-62

Cantù, C., Bosè, F., Bianchi, P., Reali, E., Colzani, M.T., **Cantù, I.**, Barbarani, G., Ottolenghi, S., Witke, W., Spinardi, L., Ronchi, A.E. *Defective erythroid maturation in gelsolin mutant mice.* Haematologica. 2012 Jul;97(7):980-8

Cantú, C., Grande, V., Alborelli, I., Cassinelli, L., **Cantú, I.**, Colzani, M.T., Ierardi, R., Ronzoni, L., Cappellini, M.D., Ferrari, G., Ottolenghi, S., Ronchi, A. *A highly conserved SOX6 double binding site mediates SOX6 gene downregulation in erythroid cells*. Nucleic Acids Res. 2011 Jan;39(2):486-501

PhD portfolio

Name PhD student: Ileana Cantu'
Erasmus MC Department: Cell Biology
Research School: Graduate School MGC
PhD period: October 2011 – November 2015
Promotor(s): Sjaak Philipsen
Co-promotor: Thamar B. van Dijk

General courses

		Workload (ECTS)
2011	Safely working in the laboratory	0.25
2012	Laboratory animal science (Artikel 9)	3
2012	Literature course	2
2012	Workshop on Photoshop and Illustrator	0.5
2013	Ensembl Workshop	1
2013	Workshop on BAGE	1
2013	Biostatistical Methods: basic Principles Part A	2

Meetings, Workshop and Symposia

2011-2015	Weekly Monday Morning Meetings, Cell Biology Department, Erasmus Medical Center (Oral presentation)	2
2012	19th MGC PhD worshop, Dusseldorf, Germany (Poster presentation)	2
2012	22th MGC-Symposium, Leiden, The Netherlands	0.5
2013	20th MGC PhD workshop, Luxembourg city, Luxembourg (Poster presentation)	2
2014	23th MGC-Symposium, Rotterdam, The Netherlands	0.5
2014	21st MGC PhD workshop, Munster, Germany (Oral presentation)	2

(Inter)National conferences

2012	18th Hemoglobin Switching meeting, Asilomar, CA, USA (Poster presentation)	2
2013	1st ThalaMoSS Scientific/Kick Off Meeting, Ferrara, Italy	2
2013	57th national meeting of the Italian society of biochemistry and molecular biology, Ferrara, Italy (Oral presentation)	2
2013	2nd symposium on Chromatin changes in differentiation and malignancies, Egmond aan Zee (Poster presentation)	2

2014	19th Hemoglobin Switching meeting, Oxford, UK (Poster presentation/ Selected for oral presentation)	2
2015	9th Dutch Hematology Meeting, Papendal, Arnhem, NL	1
2015	3rd symposium on Chromatin changes in differentiation and malignancies, Marburg (Poster presentation)	2

Additional Activities

2012	Junior science Program, ErasmusMC, 2 high school students	2
2015	Molecular Medicine master student supervision	3

Acknowledgments

First of all, my gratitude goes to **Prof**. **Dr. Sjaak Philipsen** and **Dr. Thamar van Dijk**, for giving me the unique opportunity to work in a great lab and for having guided me in this important step of my life. Dear **Sjaak,** I became a better scientist because of your supervision and the discussions we had together. Your positive attitude prodived a safe heaven even when scientific results were not that great. Dear **Thamar**, thanks for patiently answering all my questions and for being available every time I needed you. I will always be grateful for all the things that you taught me.

My appreciation also goes to **Prof. Dr. Frank Grosveld**, **Prof. Dr. Marieke von Lindern** and **Prof. Dr. Iwo Touw.** I am thankful for your critical reading and your helpful comments on my thesis. Dear **Frank**, thanks for all the useful remarks you had through these 4 years of my PhD. Dear **Marieke**, thanks for having nice words for me when I most needed them.

Prof. Dr. Antonella Ronchi, **Prof. Dr. Danny Huylebroeck** and **Dr. Marjon Cnossen**, thanks for accepting to be in my committee. Dear **Antonella**, you were the first one to guide me in my first steps of my scientific career and make me passionate about erythropoiesis. In your lab I had the possibility to learn a lot. I will always carry with me nice memories of my years with you.

To my current and former colleagues of the Philipsen's lab, I was lucky enough to have the opportunity to work with you. **Maria**, thanks for all the psychological support you gave me, and for the afternoon teas. **Silvia**, you are a funny and genuine person, thanks for all the talking and for cheering up my days at the lab. **Nynke**, no words can describe the useful help that you gave me. Simply, thanks a lot for everything! I wish you all the best. **Harmen**, thanks for answering all my (stupid) bioinformatics questions and for always being positive. My PhD thesis would not have been the same without you. **Mike**, I know you are Italian inside, don't try to hide it! Thanks for the nice chats we had. **Martijn**, good luck with your PhD. **Bella**, your positive attitude and your enthusiasm were very refreshing during my last months in the lab. **Sahar,** thanks for everything you taught me, and **Pavlos,** thanks for making me feel welcome.

A big THANK YOU goes also to the people that made my stay in Rotterdam more enjoyable. **Jessica**, you were the first person of the department that I met ("ah, ma sei italiana"). Thanks for all the support and help you gave me during my PhD but especially for all the chats and laughs we had together in the Erasmus and outside. Of course a big thank also goes to **Thomas**, my favorite French/Italian person. **Hegias**, there is no need to say how much I appreciate your friendship. Your support during our coffee beaks helped me to keep going on. Ti voglio bene. **Chiara,** dear **Chiara,** you were always there for me in my blackest periods. I am very grateful and proud to be your friend. I hope that I will see you enough in the future otherwise I will miss you a lot. **Sara**, where are you Sara? I am here drinking water all alone now, but I am glad I can see you in beautiful Milano. Thanks for taking the time to come for the defense, I appreciate it! **Luca**, the best and quietest neighbor ever (are we really sure about that?!?), I want to thank you especially for helping me through my first days here in the Erasmus. **Andrea**, thanks for all the great food you cooked for us and for all the nice time we had together. **Isa** thanks for all the nice chats (e il tessssoro dov'e?). **Simona** you always make me laugh. Many thanks also to **Thodoris**, **Martí**, **Marta**, **Fabrizia**, **Agnese**, **Friedemann, Federica**, **Aristea**, **Simone**, **Flavia**, and **Elisa** for all the fun time outside the lab.

To all my colleagues of cell biology, thanks for making the department a nice place to work. In particular, **Ralph**, **Guillame**, **Anita**, **Petros**, **Ernie**, for always answering my questions patiently and all the help you gave me with experiments. **Nesrin**, **Wouter**, **Alex**, **Johannes**, **Diana**, **Dorota**, **Derk**, **Emma**, **Christina**, **Mihaela** for having welcomed us at the 10th floor and made our stay there more enjoyable. **Nesrin**, sharing with you the last period (and all the problems related to it) of my PhD helped me go through it. **Maaike**, **Mike**, **Eric** and **Raymond** thanks for the scientific support and for trying hard to make the KLF1 pulldown work. **Sreya** and **Niels**, I am grateful for the help with the immunofluorescence. Thanks also to **Alex** (for the help with the animals), **Rick**, **Dubi** and all the other colleagues from Robbert's, Bob's, Gert's and Dies' labs. I would also like to thank the colleagues of other departments, in particular **Cheryl**, **Ruben**, **Eskeatnaf** for the fun at the (MGC) meetings and **Mateusz**, **Michael**, **Enrico** and **Elisa** for all the nice discussions.

A special thanks goes also to the secretaries **Marike**, **Bep**, **Jasperina**, **Sonja**, to **Leo**, **Melle** and **Koos**, to the IT guys and to the kitchen's ladies for all the help during these years.

Robin and **Milena,** thanks for all the lunch/dinner/coffee we had together and the fun in Poland (especially you, little **Julia**). **Bennie**, **Ana**, **Eric**, **Agata**, **Panagiotis**, **Michelle**, **Eleni**, **Thomas** and **Mona**, it was very nice to get to know you.

Most of all, thanks to **Μαρία** and **Johan** for accepting to be my paranymphs. **Μαρία** (you again!), I once told you that I was afraid my greek colleague would be a κάγκουρας. Instead, I found a great person and a true friend. I am lucky to have met you. I am sure our friendship will continue despite the distance. Και θα μιλάμε μόνο Ελληνικά σε λίγο. **Johan**, thanks for being there to share my sad days (weeks….months…). Knowing that I was not the only one helped me tons. And thanks for always trying to make me see the good aspects of life in the Netherlands that, I can admit now, are not few (I will never get over the weather and the food though, sorry). It was a pleasure to start my PhD with you two and I am honored to finish this journey with you by my side.

Thanks also to my longtime friends that I left behind (or scattered all over Europe) that were always there to support me. Whenever it's time for a (small) vacation I can't wait to see you guys. **Maria**, you are the best friend anyone could have. Even through the distance, you always made me feel your support and love. Thanks for all the years that we have passed together and, I am sure, for all the years to come. **Laura**, **Marga** and **Bea**, your support was and is irreplaceable. I know you are always there for me when I need you. When we are together it's always funny and joyful, I love spending my time with you. **Marika**, teso amo, we know each other for so many years and we have grown up together. It's amazing how things have changed, but I am so happy that you are still standing by my side. I know it was difficult for you to be at my defense, but you knew how important it was for me to have you there, and I really appreciate it. **Ilaria** and **Gloria**, what can I say girls? I treasure every moment we spent/spend/will spend together in the lab and outside. **Chiara**, **Claudia**, **Claudia** nonstante tutto quello che é successo, sono contenta di vedere che possiamo ancora contare l'una sull'altra. Finalmente siete riuscite a venire a trovarmi :) **Ale**, you have the amazing power to make me laugh when I feel very sad. **Lu**, **Ska**, **Miglio**, **Piga**, **Lale**, thanks for making me happy every time I see you. **Matt**, you stupid American, thanks for your daily presence in my life, but start saving money!

Mamma e **Papá** non c'é bisogno di dirvi che questo traguardo l'ho raggiunto solo grazie al vostro supporto e amore incondizionato. Mi mancate ogni giorno, e so che é lo stesso per voi, ma spero di avervi reso felici. **Alex**, nonostante tutte le litigate non ti cambierei con nessuno al mondo. Spero che continuerai ad essere un fratello fantastico come lo sei stato fino ad ora. **Rossana**, grazie di essere stata sempre presente ogni volta che sono ritornata in Italia. **Zio Gió**, **Zia Anna**, **Zia Tiziana** e **Ilaria** grazie di essere presenti in questo giorno cosi speciale per me. Un grazie a tutta la mia famiglia, per avermi sempre accolto a braccia aperte ad ogni mio rientro. Un pensiero in particolare va a mia **nonna Lina**, che ha visto questa avventura cominciare ma non finire. So che saresti stata orgogliosa di me.

Dimitris, thanks for being by my side, always and no matter what. Sei la mia roccia e la persona che piú di tutti mi ha aiutato ad andare avanti negli ultimi due anni del mio PhD. Με κάνεις ευτυχισμένη. Δεν τελειώνει. Καθόλου.

"Moving on is simple. It's what we leave behind that's hard."