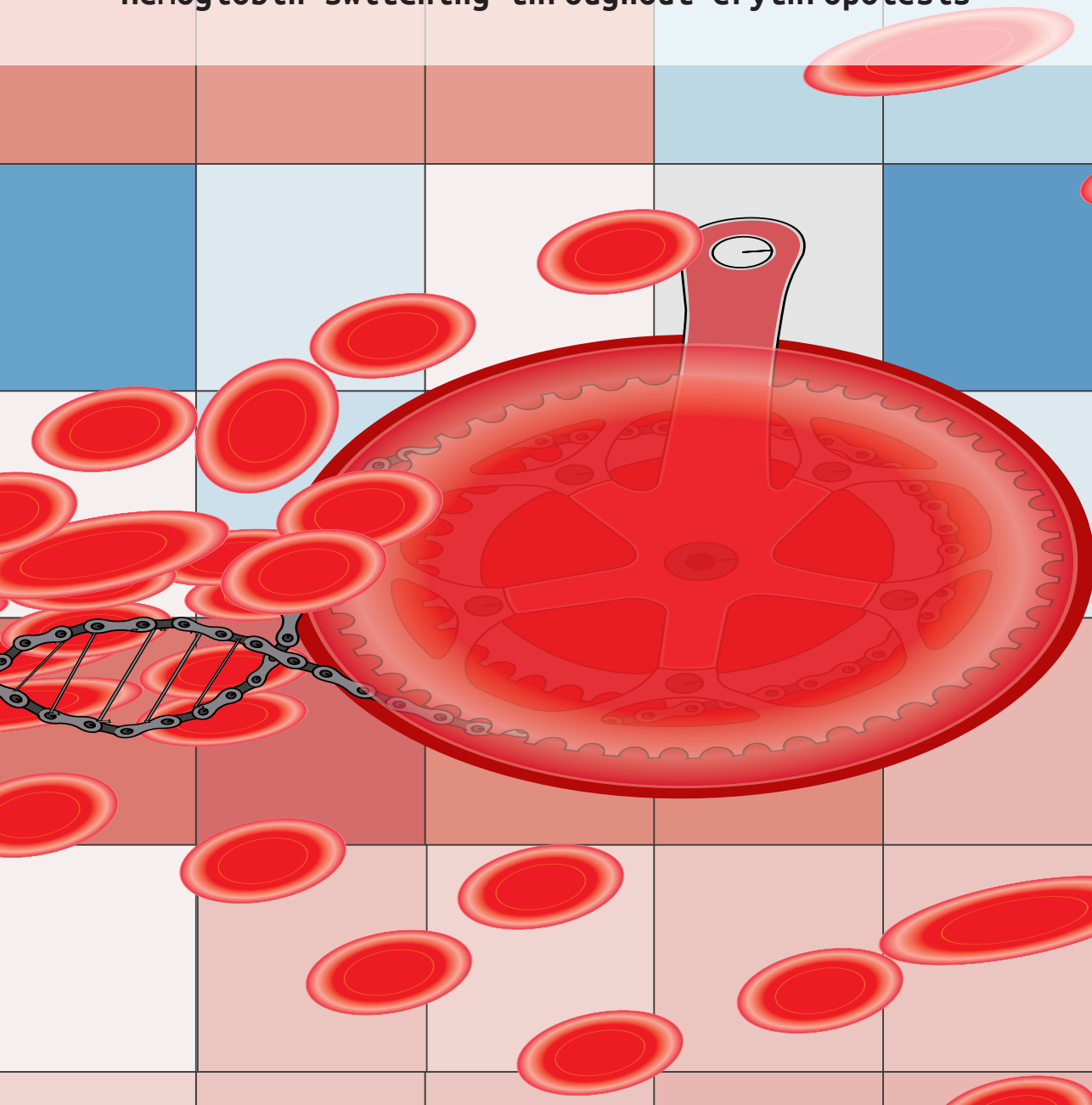
Switching Gear

Hemoglobin switching throughout erythropoiesis



Steven Heshusius

Switching Gear

Hemoglobin switching throughout erythropoiesis

Steven Janko Heshusius

Colofon

Research in this thesis was conducted at the department of hematopoiesis within Sanquin Research in Amsterdam and the department of cell biology at the Erasmus Medical Centre in Rotterdam. Printing of the thesis was financially supported by the Erasmus MC and Good! B.V.

Layout: Printservice Ede
Cover design: Steven Heshusius
Printed by: Printservice Ede

Switching Gear

Hemoglobin switching throughout erythropoiesis

Schakelmateriaal

Hemoglobin switching tijdens erythropoëse

Proefschrift

ter verkrijging van de graad van doctor aan de
Erasmus Universiteit Rotterdam
op gezag van de
rector magnificus

Prof.dr. R.C.M.E. Engels

en volgens besluit van het College voor Promoties.

De openbare verdediging zal plaatsvinden op

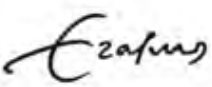
woensdag 2 december 2020 om 11.30 uur

door

Steven Janko Heshusius

Geboren te Rotterdam.

Erasmus University Rotterdam



Promotiecommissie:

Promotores:	Prof. dr. J.N.J. Philipsen Prof. dr. M.M. Von Lindern
Overige leden:	Prof. dr. N.J. Galjart Dr. M.H. Cnossen Prof. dr. J.J. Voorberg
Copromotores:	Dr. E. Van Den Akker Dr. ir. W.F.J. van IJcken

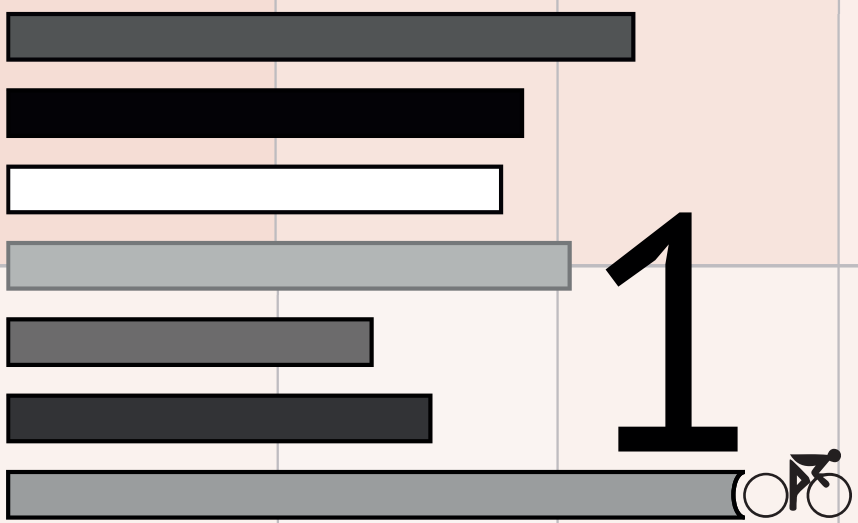
Contents

Glossary		6
Chapter 1	General Introduction	9
Chapter 2	Large-scale *in-vitro* production of red blood cells from human peripheral blood mononuclear cells	45
Chapter 3	CD14+ monocytes repress gamma globin expression at early stages of erythropoiesis	75
Chapter 4	Epigenomic analysis of KLF1 haploinsufficiency in primary human erythroblasts	95
Chapter 5	Hemoglobin switching in mice carrying the *Klf1*^{Nan} variant	123
Chapter 6	General Discussion	149
Appendix	Summary	164
	Samenvatting	166
	PhD portfolio	170
	Acknowledgements	174
	Curriculum Vitae	182

Glossary

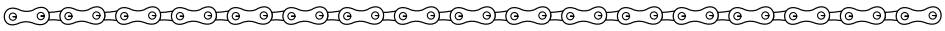
ACH	Active chromatin hub
AP-1	Activating protein 1
ATAC	Active transposon accessible chromatin
ATF	Activating transcription factor
BCL11A	B-cell lymphoma factor 11A
BCAM	Blood cell adhesion molecule
BEL-A	Bristol erythroid line A
CBP	Chromatin binding protein
C/EBP	CCAAT/enhancer binding protein
CD14	Monocyte differentiation surface antigen
CD163	Hemoglobin scavenger receptor
CD169	SIGLEC1: Sialic Acid Binding Ig Like Lectin 1
CD71	Transferrin receptor
CD206	MMR: Macrophage manose receptor
CD235a	GPA: glycophorin a receptor
CDA	Congenital dyserythropoietic anemia
ChIP	Chromatin immune preciptation
CRISPR	Clustered regularly interspaced short palindromic repeat
DNA	Deoxyribonucleic acid
DNase I	
DHS	DNase hypersensitivity site
EBL	Erythroblast
EMP	Erythroid macrophage protein
EPO	Erythropoietin
EpoR	Erythropoietin receptor
ER	Endoplasmatic reticulum
FLI1	Friend leukemia integrating transcription factor 1
FOE	Friend of EKLF protein
GR	Glucocorticoid receptor
GWAS	Genome wide association study
HbF	Fetal hemoglobin ($2\alpha2\gamma$)
HbA	Adult hemoglobin ($2\alpha2\beta$)
HBB	Gene encoding β-globin chain
HBD	Gene encoding δ-globin chain
HBE	Gene encoding ε-globin chain
HBG1/2	Genes encoding γ-globin chains
HDAC	Histone deacetylase

HPFH	Hereditary persistence of fetal hemoglobin
HRE	Hormone responsive element
HS	Hypersensitive site
HSPC	Hematopoietic stem and progenitors cell
HU	Hydroxyurea
ISR	Integrated stress response
HUDEP	Human umbilical cord derived progenitor
Kb	Kilobases
(E)KLF1	(Erythroid) Krupple like factor 1
LCR	Locus control region
LDB1	LIM-domain binding protein
LRF	Leukemia related factor
MDB	Methyl-CpG-binding domain
Nan	Neonatal anemia
NF	Nuclear factors
NLP	Natural language processing
NuRD	Nucleosome remodeling domain
PBMC	Peripheral blood mononuclear cells
PIC	Preinitiation complex
PKC	Protein kinase C
Pol III	RNA polymerase III
RBC	Red blood cell
(m)RNA	(messenger) Ribonucleic acid
ROS	Reactive oxygen species
SCF	Stem cell factor
SNP	Single nucleotide polymorphism
SCD	Sickle cell disease
SV40	Symian virus 40
TF	Transcription Factor
VCAM-1	Vascular adhesion molecule
VEGF	Vascular endothelial growth factor
ZBTB7A	Gene encoding LRF


1

Chapter 1

General introduction



Steven Heshusius

History of blood data

If the practice of bloodletting is taken into account, blood products or manipulation thereof can easily be considered the oldest healthcare application, be it not very healthy at the beginning. For example, historic records regularly mention the drawing of blood from critically ill, or transfusing animal blood into humans. At best this renders recipients of this type of treatment anemic, but more adverse effects like sepsis alloimmune reaction would mean certain death in those days.

Red blood cells, RBCs, make up the majority of our blood cells and some reports even suggest over 75% of the cells in our body cell.[1] Their main function is oxygen transport. Consequently, too few RBCs in circulation can result in insufficient oxygenation of tissues, a condition that is referred to as anemia. Anemias range from mild to lethal depending on source of RBC breakdown.

In order to be able to transport oxygen RBCs are packed with hemoglobin molecules. Hemoglobin is a tetrameric molecule that consists of two identical subdomains, the globin chains, that change throughout development in a process called hemoglobin switching. Perhaps unsurprisingly given the early fascination for blood in health and disease, the high abundance of blood cells in our bodies and hemoglobin as main molecule within these cells, the field of hemoglobin switching competes for the crown in the category oldest field in human molecular biology.

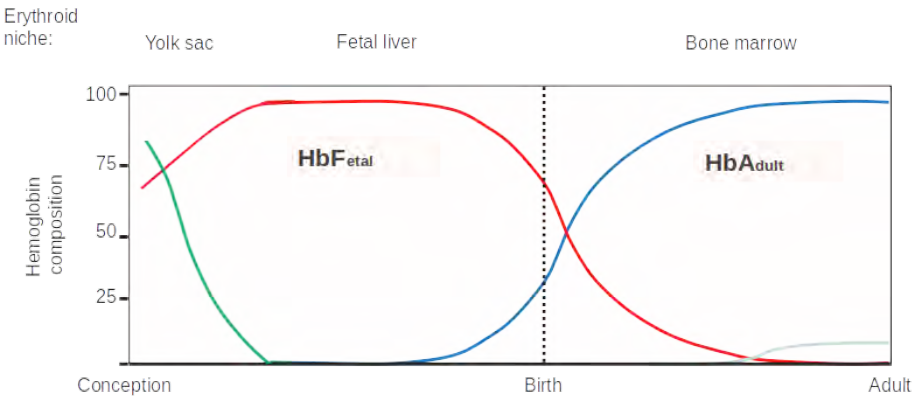


Figure 1. Schematic depiction of showing share of hemoglobin types at different stages of development.

The central premise for the hemoglobin switching field is shown in figure 1, the illustration is adapted from a figure that first appeared in Annals of New York Academy of sciences in 1974.[2] It shows how hemoglobin composition changes throughout human development in the two hemoglobin switches: an embryonic-to-fetal and a fetal-to-adult

switch. These switches occur as the main sites for erythropoiesis migrate from the yolk-sac to the fetal liver and later the bone marrow and are characterized by expression of the different globin genes. In the end this translates to most adults expressing less than 0.5% of fetal hemoglobin, HbF. Deviations from this normalcy have been of great interest since the early days of the hemoglobin switching field.

While the graph from figure 1 is commonly used to summarize the history of the entire field, it omits how research on hemoglobin switching serves as a great example of knowledge development co-evolving with technological innovation. This chapter summarizes how such combined developments have shaped the hemoglobin switching field in search of therapeutic approaches for hemoglobinopathies.

The first part of this introduction will set the stage by discussing the key discoveries that brought the hemoglobin switching field to the start of this millennium. A complete overview of the period is beyond the scope of this introduction. Instead it will focus on the core concepts that are important for the research on the fetal-to-adult switch in this thesis. It will take cues from several reviews that have provided more detailed accounts of the entire period at different time intervals.[3–6]

The two most recent decades of hemoglobin switching research will be addressed in the second part of the introduction. There I will employ natural language processing techniques to provide an overview of the literature on the hemoglobin switch. In addition to introducing the most recent progress in the field, this approach provides an example of how new means of data analysis are a central part of today's technological advances.

Blood and biology: seeing what life is made off

Starting my PhD research across the street from the Antoni van Leeuwenhoek hospital in a Landsteiner Laboratory I'll set off the introduction with the work of these two historical scientists. The Dutch lens maker Antoni van Leeuwenhoek paved the way to our knowledge on red blood cells, RBCs. Using his microscope, he was able to describe these 'red bodies' for the first time in 1695.[7] At that time, knowledge on the role of blood in health and disease was rudimentary at best, and only started to rise with the discovery of blood groups. At the start of the twentieth century Karl Landsteiner showed that blood serum from different individuals would only clump the red blood cells from certain other individuals. Initially, he reported on three different groups, a fourth group was reported shortly thereafter[8] The ABO-blood group system was born and provided the basis for the first curative blood transfusions. The method which today is still a mainstream treatment for sickle cell disease (SCD) and thalassemia patients.

The thalassemia text book, from the early '80s now in its fourth edition, is an excellent resource on the earliest history of hemoglobin research. It notes that several early reviews have placed records on a hemoglobin-associated phenomenon even before the discovery of the RBCs. For instance, in 1566 Michael Servetus described how blood changes color in the lung, which we now know results from the cells taking up oxygen.[9]

Although the earliest account of different hemoglobin types stems from the 19th century, validation of the differences between adult and newborn hemoglobin followed in the the early 20th century; the combination of alkaline sensitivity,- and electrophoretic separation assays led to the identification of two distinct types of hemoglobin.[10]

Along with the discoveries in blood this era hosted fundamental discoveries with respect to the way we understand life today; the (re)discovery of Mendel's work on hereditary traits, the discovery of the nuclear bases by Albrecht Kossel (both originating around the start of the century), the 46 chromosomes by de Winiwarter and Painter (1912 and 1923 respectively), although mistakenly quantified at 47 then 48 until the 1950 and the hereditary units on the chromosome, now genes, by Thomas Hunt Morgan in 1913.[11–14]

In the 300 years since the first record of cells within our blood, natural science had discovered the basic building blocks of the inside of cells. Still there was a long road ahead towards understanding how these parts would fit together with the different hemoglobin phenotypes.

Aberrant globin expression: nature's laboratory

Where the earliest discoveries on blood resulted from anatomical observation, much of the early knowledge on hemoglobin stems from population studies, describing the phenomena from nature's own laboratory. With the advent of x-ray crystallography, hemoglobin followed the example of whale myoglobin to become the first human protein for which the structure was identified.[15,16] This early work found that adult hemoglobin consists of two α and two β chains, fetal hemoglobin has α chains linked with two γ chains and embryonic hemoglobin pairs two ζ chains with two ε chains, or intermediate forms consisting of alpha epsilon pairing. Although the chains are similar when it comes to the subdomains, the specific amino acid sequences result in different electrophoretic properties of each chain.

Hemoglobin variants

The properties of the different globin chains were exploited with the advent of electrophoretic shift assays using paper as the carrier. Taking this assay to test the human population revealed how variants of the normal hemoglobin molecule

are at the basis of different anemic phenotypes. The molecular origin of sickle cell disease, SCD, was first described by Linus Pauling.[17] Thalassemia and other causes of anemia followed this description shortly after that and brought an increasing interest in describing the exact molecular composition of hemoglobin molecules.[18] Today's sequencing-based screening approaches have revealed the presence of different hemoglobin molecules in health as well, now over 1300 variants that result from single nucleotide polymorphisms, SNPs, to one of the globin chains are listed on the Globin server.[19,20] Combined with the thalassemia mutations this bring the total number of different hemoglobin molecules at 1815. Disease phenotypes that result from these variants are aptly called hemoglobinopathies.

β-hemoglobinopathies

Hemoglobinopathies include the most frequently occurring hereditary anemias and monogenic diseases; Sickle cell disease (SCD) and β-thalassemia.[21] These β-hemoglobinopathies result from a range of mutations in or associated with the *HBB* gene encoding β-globin.

β-thalassemia results from mutations or deletions in either the β-globin gene itself, or in the regulatory regions of the gene. Research on the latter group led to the discovery of many of the regulatory elements of the globin genes, which are a topic of a later section of the introduction. The severity of the disease ranks from minor to major depending on the type of the mutation.[22] Regardless of the severity of β-thalassemia, globin-chain imbalance, i.e. non-matching globin chain levels causing misfolded hemoglobin molecules, is one of the main causes for early decay of red blood cells and the resultant anemia in this type of disease.

SCD results from a single base-pair substitution that causes the glutamic acid residue at position 6 to be substituted by a valine residue (p.E6V).[23] The resultant hemoglobin molecules polymerize in low-oxygen environments, which leads to loss of deformability of the cells. This gives cells their characteristic sickled shape, can cause vaso-occlusive events. Repeated cycles of sickling in hypoxic and unsickling in normoxic conditions leads to a loss in membrane deformability and increased breakdown of red cells in the spleen. [24,25]

Hereditary persistence of fetal hemoglobin

Along with the characterization of alternate globin molecules it became possible to investigate why certain anemic traits did not follow Mendelian inheritance patterns. This led to the discovery of fetal hemoglobin expression in adults.[26,27] This observation was extended to healthy populations, when the screening of Swiss-army recruits revealed that around 0.01% of that population had fetal hemoglobin at levels over 0.7% of total hemoglobin.[28] The phenotype became known by the general term hereditary persistence

of fetal hemoglobin (HPFH), referring to adult individuals with HbF levels exceeding 0.5% of total hemoglobin.

Since the initial report, different forms of HPFH associated with the adult-globin locus have been described. These fall in two main categories; deletional and non-deletional HPFH. In the deletional form HPFH results from large deletions in the *HBB* locus, resulting in large stretches or even complete loss of the adult globin genes (see figure 2 for schematic representation of the locus). The non-deletional forms displays HPFH as a result of single base pair polymorphisms or small (<15bp) deletions, that occur in sites that are critical for regulation of globin expression, like for example the binding site for BCL11A at position -114 in promoter of γ-globin, that will be discussed in a later section of the introduction. A recent review summarized at least 9 forms of deletional- and 16 non-deletional HPFH.[29] In addition to mutations associated with the *HBB* locus itself, mutations in some of the hemoglobin regulators are associated with HPFH as well.

The variety of mutations that cause HPFH result in a remarkable variance of the resultant HbF levels, these range between 1% to 41% for specific cases with a single mutation in one allele. In contrast, β-thalassemia mutations that also induce HbF always results from two affected two alleles. In this the case certain deletional forms to express up to 90% HbF, due to a complete lack of a functional *HBB* gene.[30,31]

In the event of combined HPFH and β-hemoglobinopathies, disease severity was found to be inversely correlated with HbF expression and levels above 20% of total hemoglobin have been suggested to relieve patients of the need for recurrent transfusions.[32,33]

Chemical induction of fetal hemoglobin experession

The amelioration of disease symptoms with higher HbF levels, made the mechanism behind HbF variation and means of inducing HbF levels an important topic for investigation. Early success with the chemical induction of HbF came from tests with DNA-demethylating agents, 5-azacytidine and decitabine, and an inhibitor of ribonucleotide synthesis, hydroxyurea (HU). Latter was initially used as a cytotoxic drug control condition.[34,35] Later clinical trials showed that the drug induces HbF with tolerable side effects compared to a native course of the disease, making it the only drug for treating sickle cell patients that is approved by the federal drug association and still the main treatment for SCD patients today.[36–38]

Other chemical inducers of HbF showed promise on initial tests, but failed to make it beyond the stage of clinical trials. Either due to lack of efficacy, or due to cytotoxic side effects. These include dimethylbutyrate that act as histone deacetylase inhibitors, and DNA demethylating agents like 5-azacytidine.[39–42] While most alternative treatment options did not make it in to the clinic, they did suggest a role for the targeted mechanisms, like DNA methylation and Histone deacetylation, in the regulation of the hemoglobin switch.

Yet the lack of specificity for these agent makes their true relevance for the regulation of the globin genes questionable.

What started from describing the phenotype in populations, progressed with describing the proteins behind the phenotypes and turned into attempts at treating hemoglobinopathies.

With this the studies in nature's lab brought the field from the inside of patient to the interior of the cells. In the process these studies contributed to the discovery of most components of the hemoglobin switch, ranging from the genes themselves, epigenetics, regulatory elements and transcription factors that control gene expression. The following paragraphs will discuss these components as it extends the focus to the inside of the nucleus, were DNA holds the hereditary information that allows phenotypes to spread throughout populations.

Hemoglobin as a model for gene expression

While the chromosomes were already known since the start of the 20th century, electron microscopy of these structures found that they consisted of DNA fibers wrapped around nucleosomes. Winding 146 base pairs of DNA around each nucleosome allows the fibers to compact into chromatin and fit all the hereditary information inside the nucleus.[43] The discovery of the double helix structure of DNA itself provided the basis for unifying the chromosomes, nucleic acids, genes and hereditary traits in a single model. Initially, how the DNA would hold hereditary information and how this would "Translate" in to the production of proteins remained a mystery. These questions were addressed by the rise of molecular biology that followed in the decades after the discovery of the double helix marked the next important step in acquiring knowledge on hemoglobin switching.

As an interesting side note, the interdisciplinary efforts that led to these discoveries started from viewing hereditary traits in terms of code and information transfer, which crossed over from the fields of physics and mathematics in the 1940s, a period in which the term "molecular biology" was coined for the first time.[44]

Prior to the advent of molecular biology, the work on anemias had focused on the proteins at the basis of phenotypes. The discovery of DNA and its hypothesized role in storing hereditary information shifted the focus towards identifying how the information to make proteins is stored and how it is processed in cells to ultimately affect phenotypes.

From proteins to genes

Molecular cloning techniques facilitated the search for genes associated with the proteins that make up different hemoglobin molecules. Aside from efforts to find the regions on the DNA, the messenger RNA was described to carry genetic information out of the nucleus to be translated in to proteins in the cytosol.[45] Introduction of this intermediate step paved the way for questions on the regulation of gene expression that came to govern molecular biology as more genes along with their location and their regulatory elements were discovered.

The charting efforts of the DNA identified chromosome 11 as the region that harbors the genes at the basis for different globin chains involved in the fetal-to-adult switch; the *HBB* locus. Along the *HBB* locus genes line up from the embryonic epsilon- (*HBE*), to the two fetal gamma- (*HBG2* and *HBG1*), and ultimately the delta- (*HBD*) and beta-globin (*HBB*) gene (see review by Maniatis et al.[46]). Upstream of the locus there is a region that was termed locus control region. It consists of five DNase I hypersensitive sites, regions of accessible chromatin that are of pivotal importance for gene expression from the locus.[47–49]

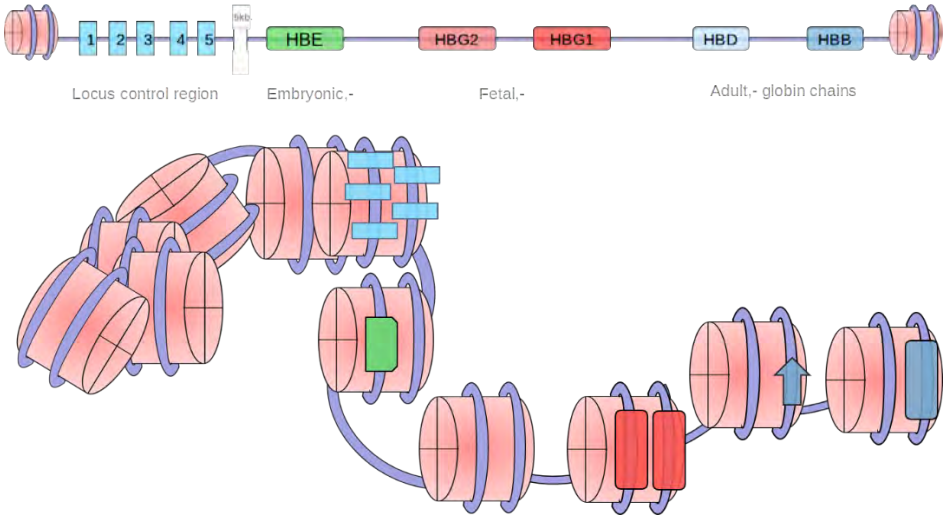


Figure 2. Schematic layout of the *HBB* locus. Top illustration shows the orientation of the genes on the adult globin locus, along with the upstream locus control region consisting of 5 DNAse I hypersensitive sites (DHS). The DHS downstream of the locus is not depicted here (for more detailed mapping of the locus see F3 of chapter 4). Bottom part shows an illustration of DNA strands wrapping around nucleosomes (147basepair per turn) to further compression of genetic material into chromosomes.

From DNA to protein

The developmental expression pattern of the globin genes along with the complete sequence map made the globin locus an invaluable model system for studies on regulation of gene expression. The smallest unit of genetic regulation of gene expression is the base pair sequence, where specific sequences make up regulatory elements that provide docking sites for the proteins that initiate transcription or have a different role in regulation.

Directly upstream of the 5' start of the transcription start site for a gene the TATA-box and initiator-sequence, together with downstream promoter elements, provide docking for the transcription preinitiation complex (PIC) and RNA polymerase II (PolII), the protein that can synthesize RNA during transcription. Further upstream in cis (on the same DNA molecule) sequences like the CCAAT-, CG-box and other enhancer elements can recruit co-activators like CCAAT-binding protein (C/EBP), activating protein 1 (AP-1). Further in cis sequences that make up binding-sites for nuclear factors (NF) and hormone responsive elements (HRE) have also been shown to contribute to the formation of the transcriptional complexes (reviewed by Lemon and Tjian[50]). Both the TATA- and CCAAT-box are shared by the proximal promoter sequences of many of the globin genes (reviewed in Philipsen and Hardison[51]). The proteins that bind these general sequences interact with those binding at sequences that are specific for lineage restricted transcription factors to result in tissue specific transcription of mRNA.

Lineage-specific regulation

Control of gene expression at sequence level extends in to lineage specific regulation. For example, the CACCC-elements, found at -80 base pairs of the transcription initiation site in globin promoters, are recognized by the erythroid specific transcription factor Krupple like factor 1 (KLF1). Along with GATA1 and NF-E2 this was described as the earliest (master)regulators of erythropoiesis.[52–55] Combinatorial interactions between the lineage specific transcription factors along with those that bind general regulatory sequences form an additional layer of transcription factor networks that control gene expression. Next to erythroid specific factors, the core transcription factors that regulate hemoglobin switching include BCL11A and LRF, which will be discussed in more detail in the second part of the introduction.[56,57]

Epigenetic regulation

In addition to regulatory sequences and transcription factor combinatorics, epigenetic modifications to the DNA form another layer that controls gene expression. To this end complexes of general and lineage specific factors can include proteins that place these modifications on the DNA or chromatin structure, for example methylation of cytidine residues, or placing various chemical side groups on the tails of histones that make

up the nucleosomes. In turn these results in altered chromatin accessibility and gene expression.

In the context of hemoglobin switching DNA methylation of CpG residues was the earliest epigenetic mechanism implicated in developmental gene regulation, while acetylation of histones tails potentially controls globin expression by affecting the accessibility of the DNA in the 30nm chromatin fiber.[58–60] In the previous section on chemical induction of HbF, we discussed how the use of DNA demethylating agents such as 5-azacytidine, and histone deacetylation inhibitors such as butyrate and trichostatin A suggests the importance of these epigenetic layers in regulation of the globin gene expression.[61–63] On the other hand, the lack of broadscale efficacy in altering globin gene expression with these agents in clinical setting raises the question to what extend epigenetics directly drives activation or inhibition of globin gene expression.

Distal regulation

The regulation mechanisms described up to this point focus on the on-site drivers and inhibitors of gene expression. DNA sequences in the promoter bind both general and specific transcription factor, that in complex with epigenetic modifiers affect local chromatin structures and gene expression.

However, beyond these local interactions, DNA sequences located further away from the transcription start sites also contribute to transcriptional regulation. Initially, such transcriptional enhancer sequences were reported within 1kb from transcription start sites, for example the SV40 enhancer that was found to increase gene expression of their associated genes. [64] Later, the β-globin locus was the first genomic region for which a long-range DNA interaction were reported to affect gene expression. The discovery of the *HBB* locus control region (LCR) showed that enhancer sequences can act from as far as 60kb in the case of the *HBB* gene.[47–49]

The potential for long-range interactions makes the 3D-organisation of the DNA-chromatin fiber a fourth layer of regulation at the level of the genome. Ultimately, chromatin loops were discovered to place the enhancer element from the LCR in trans of the beta-globin promoter to positively influence gene expression.[65] This will be discussed in a section on the more recent advances in the hemoglobin switching field.

Additional layers of regulation such as non-coding-, small RNAs at genome level and post-translational modifications and alternative isoforms of transcription factors at protein level also feed into the mix of transcriptional regulation.[66–70] Since there is limited evidence for this type of regulation affecting the globin genes, these additional modes of regulation are beyond the scope of this introduction.

Following the early mapping of the locus and the identification of different regulatory layers, hemoglobin switching emerged as a convenient model to study

regulation of gene expression throughout erythroid development. The following paragraphs discuss the process by which red blood cells are formed.

Erythropoiesis throughout development: Regulation and the environment

From a historical perspective developmental erythropoiesis is intertwined with all of the proceedings described in the previous paragraphs. From the first observations on different hemoglobin types in adult and neonatal blood, through DNA modifications at specific developmental stages, to the formation of transcription factor complexes, each of these developments aimed to understand how the hemoglobin genes are expressed as RBCs develop from hematopoietic stem-and progenitor cells.

Hematopoiesis

The description of cells in our blood other than RBCs dates back to mid-19th century, the same period as the first description of hemoglobin within red cells.[71] Contrasting to RBCs these other cells did not appear to have a color and were called leukocytes (*leuko* after white in Greek). At the beginning of the 20th century, early hypotheses by Neuman and Maximov proposed that a single cell type, residing in the bone marrow, could be the source for red and white blood cells (see review by Boisset and Robin[72]).

In comparison to red blood cells the research on hematopoietic stem- and progenitor cells was held back by problems of scale and accessibility. Not until the first bone marrow transplantation had been performed, did the experimental evidence in mouse models demonstrate the ability of donor bone marrow to reconstitute all blood lineages in a recipient.[73,74]

Hematopoiesis is classically depicted in a tree like structure; one hematopoietic stem progenitor cell (HPSC) differentiates in to more specialized branches of progenitors to give rise to five specific blood cell types as the leaves of the tree.[75,76] With recent advances in single cell sequencing approaches the classical tree-model has been challenged by a model that favors gradually shifting gene expression and cellular identity. With it a discussion has started on the origin of cell fates, the extent to which genetic information controls the future cell type(reviewed by Laurenti and Göttgens[77]).

Erythropoiesis

Erythropoiesis concerns the erythroid specification from progenitor cells. Throughout ontogeny, the process is thought to occur in different developmental waves.[78] Although the general concept of cell fate changes, a multipotent cell

that is increasingly restricted in its potential, is similar across all developmental processes, two distinct erythroid waves can be distinguished. Primitive and definitive erythropoietic waves differ in key characteristics, like their progenitor cell, cellular environment, morphology and hemoglobin expression (Figure 3).[79,80]

Hemangioblast-derived hematopoietic progenitors give rise to erythroid cells in the yolk sac during the primitive wave, while the HSC, that arises from the aorto-gonadal-mesonephros region at embryonic day E9-E10, does so in the fetal liver and the bone marrow during the definitive wave. Primitive erythroid cells are larger compared to definitive cells and mature in the circulation, where they eventually lose their nucleus. In contrast, definitive cells only enter the circulation after enucleation has occurred.[80]

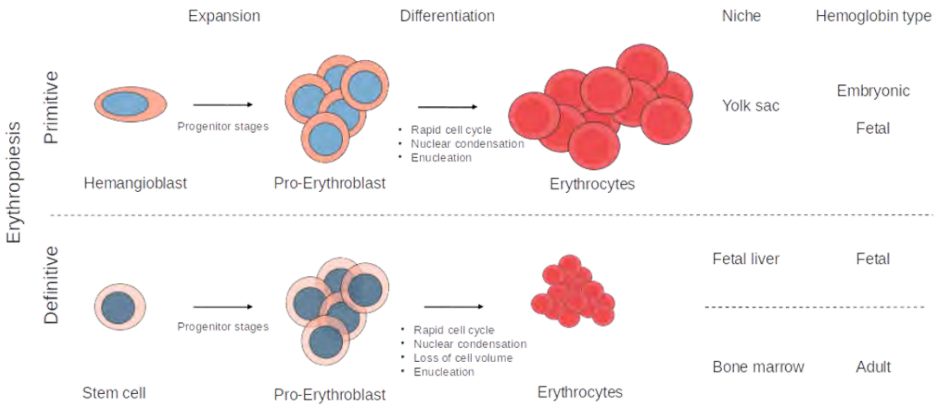


Figure 3. Key characteristics of developmental erythroid waves.

Enucleation

Enucleation is the final step in erythropoiesis and can be considered an extreme cell fate decision. The cell expels the blueprint of life, rendering it unable to divide and abrogating long-term survival.

Although the latest progenitor to maintain extensive dividing capacities is the pro-erythroblast (pro-EBL), the magnitude of this potential is unclear *in vivo*. Microscopic dissection of the bone marrow described various stages of cells that make up the transition towards the enucleated cell; terminal differentiation [81]. In this process pro-EBL undergo a limited number of cells divisions (basophilic erythroblast), gradually decrease in size, while they accumulate hemoglobin (polychromatic), condense the nucleus (orthochromatic normoblast), breakdown or expel most of their organelles and ultimately their nucleus (reticulocyte). [81–83] The reticulocytes mature in to erythrocytes in the bloodsteam.

Signaling in erythropoiesis

The extreme fate decision that occurs in erythropoiesis is thought to be controlled by balancing signals from the cellular environment. On one hand, these signals consist of (soluble) growth factors, such as stem cell factor (SCF), erythropoietin (EPO), Vascular endothelial growth factor (VEGF) and insulin like growth factor that stimulate erythropoiesis and inflammatory signals such as tumor necrosis factor, interferon gamma and tissue growth factor beta (TNF, IFNγ and TGFβ respectively) that negatively regulate erythropoiesis (see review by Tsiftsoglou et al., 2009). On the other hand, these signals originate from cell-cell interactions within a specific erythroid niche, more on that in the next section. Of the stimulatory growth factors, the role of EPO and SCF is best understood, in part because these are essential factors for *in vitro* erythropoiesis.

SCF can be produced by various cell types, in both soluble and membrane-bound forms and binds to the KIT receptor. [85–87] While the soluble form is sufficient to sustain *in vitro* erythropoiesis, mice that lack the transmembrane form are anemic.[88,89] Although SCF has a general role of accelerating cell cycle entry in hematopoiesis, in combination with other growth factors such as IL3 it greatly enhances the expansion capacity of erythroid progenitors and is suggested to positively influence long-term HSPC maintenance *in-vivo* as well. [90–93] In erythropoiesis the interaction of the KIT and EPO receptor (EpoR) signaling is critical for erythroid progenitor development and retards the differentiation of these progenitors, adding to their expansion potential and stimulating the production of fetal hemoglobin in adult erythroid cultures.[94–96]

Well before any record of a role for SCF in hematopoiesis, EPO was recognized as the body's main erythroid hormone. It is produced by the kidneys and provides the main stimulus for production of erythroid cells in the bone marrow.[97] In response to low oxygen tension in the environment or as a result of anemia EPO production increases to boost RBC production. This results in the process of stress erythropoiesis which will be discussed further on.

Aside from the synergy with SCF/KIT signaling that supports development of early erythroid progenitors, EpoR signaling promotes erythropoiesis through activation of PI3K, MAPK and JAK2 and STAT5 signaling. These pathways induce expression of Bcl-XL, an anti-apoptotic protein.[98–101]

SCF-induced PI3 kinase signaling prevents terminal differentiation of erythroid progenitors through inhibition of *Foxa3a* expression, which control genes associated with enucleation [102]. In addition, it stimulates the translation of transcripts that are crucial for self-renewal division and is implicated in a host of other functions (reviewed in Stefanetti et al.[103]).

While EPO was pursued as the first recombinant hormone to stimulate erythropoiesis, other synthetic hormone agonists were already tested as a treatment for anemia. In certain cases, glucocorticoids, a class of lipophilic hormones, were reported to relieve

anemic patients of disease symptoms.[104,105] Now these cases are all recognized to belong to Diamond Blackfan Anemia, where the curative effect results from increasing the output of stress erythropoiesis. In line with this, mice with low levels of glucocorticoid receptor (GR) expression do not show a direct erythroid phenotype, yet fail to mount a sufficient stress response.[106–108] Upon stimulation by signaling molecules nuclear hormone receptors, such as the GR, dimerize to bind specific sequences in the DNA and activate gene expression of their targets, either alone or in cooperation with other regulators like those from the KLF family or NF-E2.[109]

While the balance between pro-survival and pro-apoptotic signals in essential for the progression of erythropoiesis, the combination of SCF and EPO is sufficient to recreate the process *in vitro,* and is enhanced by glucocorticoids. A description of other soluble factors is therefore beyond the scope of this introduction.

Erythroid niche

Although soluble factors are considered the main source of signaling in cultures of erythroid cells, *in vivo* these signals make up only part of the erythroid niche. For example, membrane-bound SCF is a component that is required for erythropoiesis *in vivo*.[89]

The dissection of bone marrow that led to the description of the different stages of erythroid development also resulted in the identification of another central component of the niche for definitive erythropoiesis: the erythroblastic islands. These structures consist of central macrophages surrounded by different terminal differentiation stages of erythroid progenitors, as many as 30 around 1 macrophage in humans. [110–112] Cells in these structures can adhere to each other via a range of surface molecules, including erythroid macrophage protein (Emp), VCAM-1, ICAM-4, α5 integrin and α4-β1 (for review see Heideveld and van den Akker[113]).

Since the discovery of this part of the erythroid niche, the interaction with a central macrophage has been implicated in taking up the expelled nuclei, positively influencing erythroblast survival and contributing to iron uptake (reviewed in Manwani and Bieker[114]). Despite the range of functions that have been attributed to the interaction between erythroblasts and macrophages, the importance of the macrophage is still not fully understood. In part due to contradictory result with macrophage inhibition; depletion of macrophages by clodronate does not result in an overt erythroid phenotype. In contrast, a KO-model for *Maea*, encoding erythroid macrophage protein (EMP), is embryonic lethal at d12.5 with a severe anemic phenotype.[115,116]The embryonic lethality suggests the importance of the erythroblast island in the expansion of definitive erythropoiesis in the fetal liver. Interestingly, in fetal liver macrophages KLF1, master regulator of erythropoiesis, has been implicated in the upregulation of DNAse II, the enzyme that facilitates the breakdown of expelled nuclei after uptake by the macrophage.[117]

Like their mouse counterparts, central macrophages in human bone marrow can be characterized by the expression of CD169 and VCAM-1. Protein and mRNA analysis of patients with anemia, or after HSC transplantation have suggested a role for *EMP* in the human erythroblastic island, by showing upregulation of Emp in response to EPO treatment.[118] Recent RNA-sequencing experiments show that next to surface markers CD169, Emp and VCAM-1 mouse bone marrow and fetal liver macrophages express higher levels of the Epo receptor[119] In addition, molecules involved in iron recycling are expressed at higher levels in EpoR positive macrophages compared none EpoR macrophages, which would support the notion of their importance in erythropoietic support and possible regulation of the interaction by erythroid hormones.

In vitro approaches underline the role of macrophages in erythroid support.[120,121] Next to controlled co-cultures of isolated erythroblasts and macrophages, the interaction is also modelled in human erythroid cultures started from peripheral blood monocytes (PBMC). There the CD14+ monocyte fraction obtains a phenotype similar to that described for central macrophage in response to glucocorticoid stimulation, becoming CD163, CD169 and CD206 postive. Interestingly, these same cultures show that the CD14+ fraction positively influence the yield about 5-, to 10-fold by promoting progenitor survival.[122,123] Combined this shows the potential of cells to interact in a niche environment and influence gene expression at various stages of development.

The above paragraphs illustrate how the differentiation process is influenced by the signals a cell receives from its surroundings. Each of these signals is relayed via signaling cascades, or direct activation of transcription modifiers, that in turn form complexes with other proteins that affect different layers of regulation, which were discussed in the previous section. Our ability to combine and integrate these observations is continuously increasing, as we will see in the final part of the introduction.

Data on hemoglobin switching: recent years

In the preceding paragraphs we've seen how the knowledge on hemoglobin switching developed from the discovery of red blood cells, through discovery of normal and abnormal globin chains, to various layers that control regulation of gene expression. In the process the hemoglobin switching field developed into a well-established model for developmental gene regulation in a multi-gene locus. Even though the core regulatory mechanisms of gene expression from the locus have been identified, the exact sequence of events that initiates each developmental globin switch remains to be elucidated.

Since the hemoglobin genes and their regulators make up only a minute portion of the genes that change during erythropoiesis, a possible explanation would be that a yet to be identified factor that initiates the switching process. Indeed, new regulatory

factors have been added to the globin switching model in the recent years, as will be discussed later in this section.

Another explanation could be that we have yet to describe the interplay between regulation of gene expression at different organizational levels. At cellular level, the previous sections on erythroid signaling and the niche described how interaction between cells affects regulation of gene expression via direct cell-cell contact or via soluble signaling molecules. These signaling events have to converge at the different layers of regulation of gene expression, as the combination of transcription factors with the DNA sequence and chromatin structure controls gene expression at the level of the genome. Assuming that developmental erythropoiesis can be modeled as a deterministic process, each of the regulators involved in globin switching could potentially be affected by signals from the higher organizational level. A more extensive model of regulating gene expression during developmental erythropoiesis might help capture the interactions that drive developmental globin switching.

Historically, pioneering discoveries in each time period were guided by the ability to observe and describe the phenomena of interest with the new technologies of that time, be it the earliest microscopy in the discovery of red blood cells, electrophoretic shift assays for the different hemoglobin types or molecular cloning at the basis of all genetic components. Often, interdisciplinary efforts shaped not only new assays, but also new language to observe and describe the basis for new hypotheses.

Today, in order to deal with the higher complexity of expanded regulatory networks, biology and computer science are increasingly interconnected. They provide the basis for assays that simultaneously measure different layers of regulation at a genome wide scale; at the genome level with genome wide association studies (GWAS), at the chromatin level with ATAC sequencing, at transcript level with mRNA sequencing, or at protein level with proteomics.

By combining more elaborate data collection and improved methods to uncover patterns in data, the hemoglobin switching field is primed for further discoveries that lead to alternative treatments in hemoglobinopathies. The last chapter of this thesis will discuss this premise in more detail.

As an advance on this discussion, the second part of the introduction will borrow techniques from the field of natural language processing and network analysis to interpret the literature on hemoglobin switching from the last two decades. It will compare two ten-year periods and outline how the work in this thesis extended from the developments in these recent years and the general progress of the field described in the previous paragraphs.

Figure 4 summarizes the most frequently occurring words in the titles of publications for the search “hemoglobin switching” on Web of Science in the last two decades (see boxed text). The most frequent words in the titles capture a shift of focus in the field.

The period from 1999 – 2008 shows "locus" as top term followed by "beta" and "gamma" and terms like "alpha", "control" and "region" most frequently used after that. The period 2009 – 2018 also shows "beta" and "gamma", but now "bcl11a" and "klf1" are the most frequent terms that follow. Both periods also mention either "transgenic" or "mouse", "model" to highlight the importance of experimental systems.

The main trend captured in the word frequency analysis reflects the transition from DNA structure of the beta-locus to individual transcription factors as the central topics in publications on globin gene regulation. Along with the main components of each trend, its main implication towards the development of novel treatments will be discussed in this second part of the introduction.

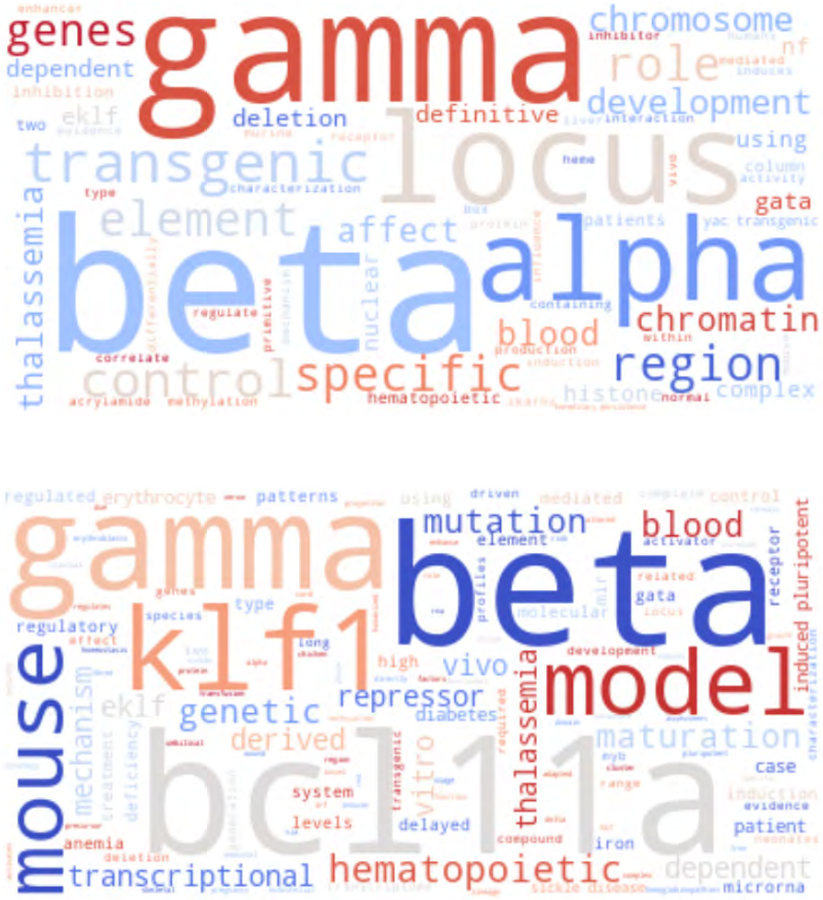


Figure 4. Word cloud summarizing most frequent terms in publication titles from ten-year time periods. Top panel period 1998 – 2008. Bottom panel 2009-2018

Locus control over the hemoglobin switch

As noted in the historical overview, the locus control region was discovered as a collection of DNAse I hypersensitive elements that conveyed erythroid-specific chromatin accessibility and copy-number dependent expression levels to linked globin genes, irrespective of the integration site of the DNA vector [49]. The different elements that make up the LCR and interacting proteins, including SPI1/KLF1, GATA1 and NF-E2, and DHS interaction were identified early on.[52,124–127] Yet, exactly how the LCR contributed to globin expression remained an outstanding question.

The working hypotheses included mechanistic models such as chromatin looping, gene competition, gene silencing and RNA polymerase II tracking. As more components of globin gene regulation were described these were found to be non-mutually exclusive (reviewed by Bank[128]). The central place of the terms "locus", "control" and "region" in the text analysis of the 1999-2008 period results from a high number of publications that described the experimental evidence of chromatin looping and its place in gene regulation. These elaborated how folding the 3D-chromatin structures, would place the LCR in close proximity to the promoters of the different globin genes in order to recruit transcription factors that co-regulate transcription of different globin genes. The hypothesis was fueled by the observations of alternating proximity of the LCR between beta- and gamma-globin promoters in RNA in-situ hybridization assays [129]. Following this observation, experiments with the mouse *Hbb* locus provided direct evidence for distal interactions *in vivo*. Using the newly developed Chromosome Conformation Capture (3C) technique; in fetal liver derived erythroid cell the LCR region was found in close proximity to the active *Hbb* genes in the locus, in a structure that was termed the Active Chromatin Hub (ACH), while *Hbb* locus was found in a linear conformation in brain cells. The interaction in the ACH results in looping out of the 30-50kb chromatin stretch between the active globin promoters and the LCR and are specific to the developmental stages in both mouse and human cells. Early developmental stages show ACH formation with the fetal globin genes, looping out the chromatin fiber with the adult stage globin genes.[65,130,131]

Forced chromatin looping

KLF1 and GATA1 were among the first transcription factors that were reported to drive ACH formation or contribute to the stabilization of the structure.[132,133] The formation of an ACH exemplifies how both the DNA sequence, transcription factors and 3d-structure interact in regulation of gene expression. The importance of transcription factors in this process has since been validated in experiments with artificial zinc fingers linked to LIM domain binding 1 protein (LDB1), that were able to reactivate fetal globin expression in adult erythroblasts, or by introduction of a KLF1 binding site in silenced *HBG* promoter that induced expression of fetal hemoglobin in HUDEP2 cells.[134–136] Both these experiments

A network approach to analysis of publication data

Along with our understanding of molecular and developmental mechanisms, the amount of scientific papers has increased exponentially in the last decade. Much like the cross-over from physic' information providing the basis for molecular biology, approaches borrowed from computer sciences can help detect structures in bodies of literature. Here a combination of natural language processing (NLP) and network analysis are combined to illustrate an alternative approach to literature review.

Data collection

Literature search was performed on Web of Science with the term "(hemoglobin OR haemoglobin) AND (switch) AND (fetal OR gamma OR mice)" and was applied to two periods of ten years, 1999-2008 and 2009-2018. The search term was created in an iterative process of trial and error with addition of, or adaptation to terms to yield a coherent set of publications on the hemoglobin switch. Only scientific articles were included to capture a process of spreading hypotheses through direct citation of experimental work.

Network analysis

Network analysis can be an intuitive to visualize connections between neurons, people, products or other concepts the analysis is performed on. In the case for publications on the hemoglobin switch, a network was constructed using first author name and publication year as an identifier for each publication. Publication that were cited at least twice were combined with the reference list from each publication to form the initial network, that was then pruned to only include nodes that were cited by at least two others in the network. Velocity (number of citations per year) was included to highlight influential papers within cliques of the network, which in turn served as the backbone for the discussed literature in the second part of the introduction.

Natural language processing

NLP revolves around allowing computers to interface, process and ultimately understand bodies of text. At its basis a bag of word approach first quantifies word frequencies. For the publications in the network the titles were split into individual words that are counted. Further processing was used to subtract stop words and most frequent co-occurring words between the two time periods analyzed. This latter step prevents the overlapping terms from obscuring differences in topics by general terms like "hemoglobin" or "erythropoiesis". For the resultant dataset most-frequent words were plotted in word clouds.

Inherent bias from selection

Each of the described steps introduces considerable biases, be it in the choice of database and search term, the inclusion criteria for two or more citations for inclusion of publications in the network, or filtering of n most frequent words. In more advanced application of this type of literature research multiple searches could be combined across databases to reduce bias there.

demonstrate the ability of the LCR to activate silenced globin genes through changing transcription factor binding to sites in the silenced promoters. Although these studies provide proof of concept for induction of HbF expression through forced chromatin looping, other genetic engineering approaches for treatments of hemoglobinopathies are at more advanced stages of the development process.

Targeting DNA sequences

The understanding of the locus layout and structure has resulted in clinical trials where hemoglobin genes are introduced in patients with sickle cell disease or β-thalassemia. The virus vectors, used to introduce the transgenes, exploit a conformation similar to the original mini locus that conveyed position independent gene expression.[137,138] Lentiviral vectors with a globin transgene linked with different assortments of the LCR hypersensitivity sites were tested for their efficacy (reviewed in Cavazzana et al[139]). Ultimately, this led to the development of the GLOBE, BB305 and βAS3-FB for clinical testing.[140–142] At their core these vectors contain the HS2, 3 and 4 elements from the locus control region linked to an *HBG* gene or a modified *HBB* gene, that prevents sickling of cells increasing treatment efficacy in sickle cell patients. These vectors have been successfully introduced in patient stem cells that upon transplantation to relieve patients from their need for blood transfusion. The initial result are promising and a number of clinical trial is still under execution (for review see Carden and Little[143], Telen et al[144]). Although this shows the power of using gene therapy to treat patients with a β-locus defect. The requirement for transplantation of modified stem cells, make broad scale application less feasible, particularly in countries with the highest incidence of beta-hemoglobinopathies.

Transcriptional regulation of hemoglobin switching

EKLF or KLF1 also shows up as a frequent term in both the '98-'08 and the '09-'18 period (figure 4). As noted earlier in the introduction, it was one of the first lineage-specific transcription factors to be directly implicated in the control of adult globin expression [145]. In addition to regulation of the globin genes, its role as a master regulator of erythropoiesis was discovered half way through the nineties. From there the mechanisms by which it can control the various processes were starting to be elucidated, as a result KLF1 shows up as one of the most frequent terms in the text analysis on the '98-'08 period (Figure 4, top panel).

KLF1 as a looping factor

Evidence for the importance of KLF1 in erythropoiesis in general and hemoglobin switching in particular came from the *Klf1* knockout mouse model which displays embryonic lethality at embryonic day 12.5 due to impaired differentiation of definitive erythroid cells. [54,55] The same mouse model was used to identify KLF1 as an important factor in establishing chromatin looping between the *Hbb* locus and its LCR. Through its three zinc finger domains KLF1 recognizes CACCC-motifs.[145] Both the LCR HS3 and the *Hbb* promoters contain such sequences, which allow KLF1 binding and formation of the ACH. *Klf1* knockout mice show no interaction between these sites and only express embryonic beta-like globin chains.[132] Combined with the observation that introduction of a KLF1 binding site in the *HBG1* promoter, mimicking the British HPFH genotype, induces γ-globin expression[136], discussed in the previous paragraph, this established the importance of KLF1 as a looping factor.

KLF1 as a co-factor

Next to the ability to establish an ACH at the *Hbb* locus, KLF1 regulates erythropoiesis as co-factor in different complexes. It can facilitate direct transcriptional activation or initiation by interaction with core transcription machinery TFIIH and TAF9. At the same time, it has been suggested to interact with epigenetic remodeling complexes, including NuRD, P300/CBP, SWI/SNF and Sin3A/HDAC, that would at allow KLF1 directed complexes to convey activating or repressive states on the chromatin at target loci.[146,147] Despite limited evidence for presence of KLF1 in these complexes, (spatie)post-translational modifications to KLF1 have been described to provide a mode through which transcriptional regulation can be altered.

Post-translational modifications to KLF1 are found throughout both the proline-rich protein interaction domain and the zinc-fingers, which suggests that different parts of the protein are required for different regulatory functions. Examples of post-translational modifications include ubiquitination to control KLF1 protein levels, phosphorylation to ensure activation, acetylation to interact with protein complexes and SUMOylation required for nuclear translocation (see review by Yien and Bieker[148]). Further importance of posttranslational modifications and nuclear localization of KLF1 has been demonstrated in the interaction between KLF1 and Friend of EKLF (FOE). FOE-knockout mice show altered nuclear localization of KLF1. While timed activation of protein kinase c (PKC) and the subsequent phosphorylation of KLF1 at the serine residue 68 (S68) was demonstrated to affect binding to FOE and importin-β1.[149]

The notion that specific regions of the protein are important for different aspects regulation of gene expression, is supported by a wide range of erythroid phenotypes that result from KLF1 variants.

KLF1 variants and HPFH

While the publications on hemoglobin switching in the '98-'08 period described the role of KLF1 in chromatin looping and transcription factor complexes, the high frequency of KLF1 in the more recent period results from publications that described various variants of KLF1 that were discovered in association with HPFH and a range of other erythroid phenotypes. A recent review describes how the advent of genetic screening in patients with different erythroid phenotypes has identified over 65 variants of KLF1 since 2010. In total the authors describe over 140 variants, that are divided over 4 classes ranked on the severity of the resultant phenotype. The lower classes included mild phenotypes, like HPFH, aberrant blood group expression and milder forms of anemia, while a single variant in class 4 is linked to dominant congenital dyserythropoietic anemia (CDA type IV; see review by KLF1 consensus group[150]).

The identification of the p.K288X variant in a Maltese family with HPFH definitively linked KLF1 to fetal hemoglobin expression in humans. Individuals in this cohort were all compound heterozygous for the mutation, that targets the KLF1 mRNA from one allele for non-sense mediated decay, rendering these individuals haploinsufficient. The lower levels of KLF1 were suggested to result in lower levels of BCL11A, the direct repressor of γ-globin expression that will be discussed further on as a KLF1 target gene.[151,152] While the HPFH individuals in this cohort shared the same KLF1 variant, they showed HbF phenotypes ranging from 18.5% to 3.5% of total hemoglobin.

Master regulator of erythropoiesis

The wide variety of erythroid phenotypes that present as a result of KLF1 variants illustrate the myriad of processes that it controls in erythropoiesis. To date roughly 700 target genes, activated by KLF1, encode proteins that are involved in globin expression, heme synthesis, membrane and cytoskeleton integrity, metabolic processes and cell cycle regulation.[132,153,154] Along with these studies, the *Klf1* knockout mice show the central importance of KLF1 during terminal erythroid differentiation.

Prior to terminal erythroid differentiation KLF1 controls the balance between the megakaryocytic and erythroid lineages. Although the exact workings of this regulatory circuit are incompletely understood, it partly results from a counter balance with FLI1, an important megakaryocytic transcription factor that is expressed at higher levels in *Klf1* knockout embryos and is downregulated upon KLF1 overexpression.[155,156] Next to the ability of KLF1 and FLI1 to interact *in vitro*, competition for limiting amounts of GATA1 as co-factor has been suggested as a possible mechanism for the cross-antagonism between the two factors.[157,158]

Targeting KLF1

The induction of HbF following the introduction of a KLF1 binding site in the *HBG1* promoter serves as an example how targeted approaches can exploit the role of KLF1 as a looping factor at the level of DNA sequences.[136] This type of approach is yet to reach the clinical testing phase, but potentially provides a gene therapy that is less prone to off-target effects of genetic manipulation by for example random integration of gene vectors that carry repaired beta-locus.

At the level of protein complexes, targeting KLF1 is an interesting candidate, because many of its variants share HPFH as a phenotypic trait.[150] Potentially, targeting specific regions of the protein might alter expression of the globin genes, without affecting other target genes. However, the central role of KLF1 in erythropoiesis and the wide variety of phenotypes observed in humans carrying KLF1 variants suggests that such an approach could come with undesirable side effects. In this light, identification of the mechanism by which specific variants of KLF1 induce HPFH, without affecting erythroid characteristics, remains of particular interest.

KLF1 targets repress HBG

While KLF1 by itself affects globin gene expression by influencing chromatin looping, it has a more pronounced effect through two of its target genes. This is reflected in the text analysis from the most recent decade, where KLF1 shows up as the main term together with BCL11A, a direct repressor of γ-globin expression (Figure 4 bottom panel).

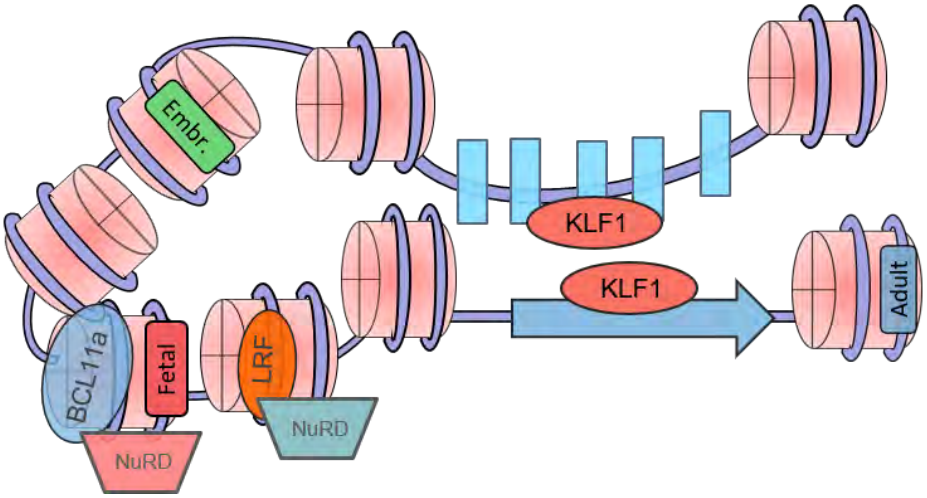


Figure 5. Schematic representation of KLF1 associated transcription factors bound across the *HBB* locus in adult conformation.

BCL11A

The importance of BCL11A was discovered in a GWAS study that showed the association of the BCL11A locus with HPFH.[159,160] Since then extensive characterization has established BCL11A as a direct repressor of γ-globin expression[57], which is reflected by the high number of article on BCL11A in the text analysis of the most recent period (figure 4B).

During the characterization process, various mechanisms for repression by BCL11A have been put forward. It was shown to interact with hematopoietic regulators, including SOX6 and GATA1, and binds to Nucleosome remodeling domain (NuRD) complexes, via interaction with CDH3/4, MDB2 or MDB3 and histone deacetylases 1 and 2.[161,162] In turn NuRD sub-complexes containing GATAD2A, HDAC2, MBD2, MTA2 and CDH4 where recently found to be essential for repression of HbF.[163] Recent insight in to the mechanism of repression by BCL11A comes from ChIP and cut-and-run sequencing experiments, which confirmed that BCL11A binding the −114 site in the HBG promoters. While the early characterization of BCL11A binding partners, like NuRD suggests that the repression by BCL11A results from the direct recruitment of epigenetic modifying complexes to these regions, the localization at the promoter also inhibits the formation of a stable chromatin loop of between the LCR and upstream regions of the globin genes.[134,164–166]

LRF

The most recent addition to the regulatory network of γ-globin does not show up in figure 4 as it was only recently described for its role in the hemoglobin switch. Leukemia related factor, LRF, which is encoded by the *ZBTB7A* gene, represses γ-globin independently of BCL11A, by affecting the histone density at the *HBG* genes.[56] Like BCL11A it is a KLF1 target, it is found in MDB2 complexes, but binds a different site located at −200 in the *HBG* promoters.[164,167] Next to the repression of γ-globin expression, LRF has pleiotropic effects in lineage specification, oncogenesis and metabolic processes (see review by Constantinou[168]). For example, in preventing cell death via inhibiting expression of pro-apoptotic BCL2 (BIM) to prevent cell death during terminal erythroid differentiation.[169]

Like for KLF1, therapeutic potential for direct targeting of LRF is limited due to its role in regulating more general processes. Approaches would have to forego reducing levels of the protein and could instead focus on disrupting binding interaction within LRF NuRD complexes.

Targeting KLF1 targets

Since its discovery BCL11A has been hailed as an ideal candidate for targeted therapy aimed at reactivating the *HBG* genes. Particularly, because there is limited evidence for lower BCL11A expression affecting erythroid characteristics apart from HbF induction and

the potential for erythroid specific targeting. Although its potential is supported BCL11A haploinsufficiency cases with curative levels of HbF, these cases are also associated with neurologic defects.[170,171] Similarly mouse models for BCL11A deficiency show defects in B-cell lymphopoiesis and early hematopoietic development.[172,173]

To limit the effects of BCL11A modulation outside of the erythroid lineage, therapy approaches use erythroid specific promoter in lentiviral constructs with short-hairpin, or micro RNA directed against *BCL11A*, or target the erythroid specific enhancer in the second intron of the gene for deletion via zinc-finger nuclease or CRISPR-Cas9.[174–176] Clinical trials using these approaches in combination with bone marrow transplantation are underway.

Scope of this thesis

Hemoglobinopathies are the most common monogenic disorder today. As noted in earlier sections, studying the hemoglobin switch has a long history in trying to find alternative treatments to recurrent blood transfusion in anemias that result from these diseases.

The work presented in this thesis aims to identify novel therapeutic alternatives, through increasing understanding of gene regulation during erythropoiesis and hemoglobin switching. To this end we developed a model system for primary erythroid cultures, characterized the model with respect to HbF expression *in vitro* and applied the model to investigate the effect of KLF1 variants in human and mouse.

In addition, future optimization and further development of the culture protocol and the corresponding culture medium can provide a basis for development of patient-tailored transfusion products (**chapter 2**).

The culture system mimics conditions that are normally found under stress erythropoiesis. Previously this mode of erythropoiesis was linked to increased HbF compared to steady-state erythropoiesis in adults. While previous work on the culture system showed that the monocyte fraction attains a central macrophage phenotype upon stimulation with dexamethasone and positively influenced the yield in cultures started from PBMCs. Combined, the HbF expression and the role of the monocyte fraction in cultures led us to the question if cell-cell interactions influence expression of HbF *in vitro* (**chapter 3**).

The range of KLF1 variants with a benign phenotype along with its central role in erythropoiesis make it an interesting target for studying both regulation of gene expression in general and the globin genes in particular. In the case of p.K288X variant in KLF1, present in a HPFH cohort from Malta, there is remaining variation in HbF levels that is not readily explained by the variant itself, or co-inheritance of other

variants of known HbF modifiers. The culture system from the first chapter was used to revisit the Maltese cohort to identify what drives the remaining variation between individuals with the p.K288X variant (**chapter 4**).

Although HPFH resulting from KLF1 variants commonly co-occurs with relatively mild erythroid phenotypes, the *KLF1* p.E325K variant is an exception. The variant affects the highly conserved second zinc finger region and causes dominant severe congenital dyserythropoietic anemia. The *Nan* mouse model can be used to study the effect of a similar KLF1 variant. It carries a *Klf1* allele with an E to D variant at the paralogous position in the protein. Previous work on this model suggests that the variant alters the transcriptional control of KLF1 over a specific group of target genes [Nebor 2018]. To what extent the variant influences the expression of the globin genes in early development has not been established. In addition to globin characterization in the earliest developmental stages, cultures of Klf^{wt}/ Klf^{nan} cells were performed to assess effect of the *Nan* variant on erythroid specification (**chapter 5**).

The general discussion of the thesis will discuss how the data generated in this thesis relates to existing insights on stress erythropoiesis, means for environmental signaling in an erythroid niche and how these observations could be used to increase our understanding of globin gene regulation during erythropoiesis (**chapter 6**).

References

1. Bianconi, E. et al. An estimation of the number of cells in the human body. Ann. Hum. Biol. 40, 463–471 (2013).
2. Kitchen, H. & Brett, I. Embryonic and fetal hemoglobin in animals. Ann. N. Y. Acad. Sci. 241, 653–71 (1974).
3. Grosveld, F., Dillon, N. & Higgs, D. The regulation of human globin gene expression. Baillieres. Clin. Haematol. 6, 31–55 (1993).
4. Weatherall, D. J. Towards molecular medicine; reminiscences of the haemoglobin field, 1960-2000. Br. J. Haematol. 115, 729–738 (2001).
5. Wilber, A., Nienhuis, A. & Persons, D. Transcriptional regulation of fetal to adult hemoglobin switching: new therapeutic opportunities. Blood 117, 3945–3953 (2011).
6. Vinjamur, D. S., Bauer, D. E. & Orkin, S. H. Recent progress in understanding and manipulating haemoglobin switching for the haemoglobinopathies. Br. J. Haematol. 180, 630–643 (2018).
7. van Leeuwenhoek, A. Arcana Naturea Detecta. Delphis Batavorum (1695).
8. Rous, F. P. Karl Landsteiner, 1868-1943. R. Soc. 5, 294–324 (1947).
9. Weatherall, D. J. & Clegg, J. B. The Thalassaemia Syndromes: Fourth Edition. The Thalassaemia Syndromes: Fourth Edition (Blackwell Science Ltd, 2008). doi:10.1002/9780470696705.
10. Brinkman, R., Wildschut, A. & Wittermans, A. On the occurrence of two kinds of hæmoglobin in normal human blood. J. Physiol. 80, 377–387 (1934).
11. Roberts, H. Plant hybridization before Mendel,. (Princeton University Press, 1929).
12. Jones, M. E. Albrecht Kossel, a biographical sketch. Yale J. Biol. Med. 26, 80–97 (1953).
13. Painter, T. S. Studies in mammalian spermatogenesis. II. The spermatogenesis of man. J. Exp. Zool. 37, 291–336 (1923).
14. Morgan, T. H. Heredity and Sex. (Columbia Univ. Press, 1913).
15. Kendrew, J. C. & Perutz, M. F. A comparative X-ray study of foetal and adult sheep haemoglobins. Proc. R. Soc. Lond. A. Math. Phys. Sci. 194, 375–398 (1948).
16. Green, D. W., Ingram, V. M. & Perutz, M. F. The structure of haemoglobin - IV. Sign determination by the isomorphous replacement method. Proc. R. Soc. London. Ser. A. Math. Phys. Sci. 225, 287–307 (1954).
17. Pauling, L., Itano, H. A., Singer, S. J. & Wells, I. C. Sickle Cell Anemia, a Molecular Disease. Science (80-.). 110, 543–548 (1949).
18. Perutz, R. R., Liquori, A. M. & Eirich, F. X-ray and solubility studies of the haemoglobin of sickle-cell anaemia patients. Nature 167, 929–931 (1951).
19. Hardison, R. C. et al. Access to a syllabus of human hemoglobin variants (1996) via the World Wide Web. Hemoglobin 22, 113–127 (1998).
20. Giardine, B. et al. Systematic documentation and analysis of human genetic variation in hemoglobinopathies using the microattribution approach. Nat. Genet. 43, 295–302 (2011).
21. Williams, T. N. & Weatherall, D. J. World distribution, population genetics, and health burden of the hemoglobinopathies. Cold Spring Harb. Perspect. Med. 2, a011692 (2012).
22. Weatherall, D. J. Phenotype—genotype relationships in monogenic disease: lessons from the thalassaemias. Nat. Rev. Genet. 2, 245–255 (2001).
23. Ingram, V. M. Gene mutations in human haemoglobin: the chemical difference between normal and sickle cell haemoglobin. Nature 180, 326–328 (1957).
24. Shriver, J. B. & Waugh, T. R. Studies on a case of sickle cell anemia. Can. Med. Assoc. J. 23, (1930).
25. Manwani, D. & Frenette, P. S. Vaso-occlusion in sickle cell disease: pathophysiology and novel targeted therapies. Blood 122, 3892–8 (2013).

26. Edington, G., Lehman, H. & Schneider, R. New results on haemoglobin G. Trans. R. Soc. Trop. Med. Hyg. 49, 309–310 (1955).
27. Jacob, G. F. & Raper, A. B. Hereditary Persistence of Foetal Haemoglobin Production, and its Interaction with the Sickle-Cell Trait. Br. J. Haematol. 4, 138–149 (1958).
28. Marti, H. R. & Buetler, R. [Hemoglobin F and hemoglobin A2 increase in the Swiss population]. Acta Haematol. 26, 65–74 (1961).
29. Wienert, B., Martyn, G. E., Funnell, A. P. W., Quinlan, K. G. R. & Crossley, M. Wake-up Sleepy Gene: Reactivating Fetal Globin for β-Hemoglobinopathies. Trends in Genetics vol. 34 927–940 (2018).
30. Forget, B. G. Molecular basis of hereditary persistence of fetal hemoglobin. in Annals of the New York Academy of Sciences vol. 850 38–44 (John Wiley & Sons, Ltd (10.1111), 1998).
31. Kulozik, A. E., Yarwood, N. & Jones, R. W. The Corfu delta beta zero thalassemia: a small deletion acts at a distance to selectively abolish beta globin gene expression. Blood 71, 457–462 (1988).
32. Perrine, R. P., Brown, M. J., Clegg, J. B., Weatherall, D. J. & May, A. Benign sickle-cell anaemia. Lancet (London, England) 2, 1163–1167 (1972).
33. Powars, D. R., Weiss, J. N., Chan, L. S. & Schroeder, W. A. Is there a threshold level of fetal hemoglobin that ameliorates morbidity in sickle cell anemia? Blood 63, 921–926 (1984).
34. Platt, O. S., Orkin, S. H. & Dover, G. Hydroxyurea enhanced fetal hemoglobin production in sickle cell anemia. J. Clin. Invest. 74, 652–656 (1984).
35. Charache, S. et al. Effect of hydroxyurea on the frequency of painful crises in sickle cell anemia. N. Engl. J. Med. 332, 1317–1322 (1995).
36. Ware, R. E. How I use hydroxyurea to treat young patients with sickle cell anemia. Blood vol. 115 5300–5311 (2010).
37. Wang, W. C. et al. Hydroxycarbamide in very young children with sickle-cell anaemia: A multicentre, randomised, controlled trial (BABY HUG). Lancet 377, 1663–1672 (2011).
38. Quarmyne, M. O. et al. Hydroxyurea effectiveness in children and adolescents with sickle cell anemia: A large retrospective, population-based cohort. Am. J. Hematol. 92, 77–81 (2017).
39. Sankaran, V. G. Targeted therapeutic strategies for fetal hemoglobin induction. Hematology Am. Soc. Hematol. Educ. Program 2011, 459–65 (2011).
40. Kutlar, A. et al. A dose-escalation phase IIa study of 2,2-dimethylbutyrate (HQK-1001), an oral fetal globin inducer, in sickle cell disease. Am. J. Hematol. 88, (2013).
41. Reid, M. E. et al. A double-blind, placebo-controlled phase II study of the efficacy and safety of 2,2-dimethylbutyrate (HQK-1001), an oral fetal globin inducer, in sickle cell disease. Am. J. Hematol. 89, 709–713 (2014).
42. Molokie, R. et al. Oral tetrahydrouridine and decitabine for non-cytotoxic epigenetic gene regulation in sickle cell disease: A randomized phase 1 study. PLoS Med. 14, (2017).
43. Abuelo, J. G. & Moore, D. E. The human chromosome. Electron microscopic observations on chromatin fiber organization. J. Cell Biol. 41, 73–90 (1969).
44. Cobb, M. 1953: When genes became 'information'. Cell vol. 153 503–506 (2013).
45. Wilson, J. T., Marotta, C. A., Forget, B. G. & Weissman, S. M. Structure of human hemoglobin messenger RNA and its relation to hemoglobinopathies. Trans. Assoc. Am. Physicians 90, 117–126 (1977).
46. Maniatis, T., Fritsch, E. F. & Lauer, J. The molecular genetics of human hemoglobins. Ann. Rev. Genet vol. 14 www.annualreviews.org (1980).

47. Groudine, M., Kohwi-Shigematsu, T., Gelinas, R., Stamatoyannopoulos, G. & Papayannopoulou, T. Human fetal to adult hemoglobin switching: changes in chromatin structure of the beta-globin gene locus. Proc. Natl. Acad. Sci. U. S. A. 80, 7551–7555 (1983).
48. Tuan, D. & London, I. M. Mapping of DNase I-hypersensitive sites in the upstream DNA of human embryonic epsilon-globin gene in K562 leukemia cells. Proc. Natl. Acad. Sci. U. S. A. 81, 2718–22 (1984).
49. 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, 975–985 (1987).
50. Lemon, B. & Tjian, R. Orchestrated response: A symphony of transcription factors for gene control. Genes and Development vol. 14 2551–2569 (2000).
51. Philipsen, S. & Hardison, R. C. Evolution of hemoglobin loci and their regulatory elements. Blood Cells. Mol. Dis. 70, 2–12 (2018).
52. Lowrey, C. H., Bodine, D. M. & Nienhuis, A. W. Mechanism of DNAse-I hypersensitive site formation within the human globin locus-control region. Proc. Natl. Acad. Sci. U. S. A. 89, 1143–1147 (1992).
53. Berry, M., Grosveld, F. & Dillon, N. A single point mutation is the cause of the greek form of hereditary persistence of fetal hemoglobin. Nature 358, 499–502 (1992).
54. Nuez, B., Michalovich, D., Bygrave, A., Ploemacher, R. & Grosveld, F. Defective hematopoiesis in fetal liver resulting from inactivation of the eklf gene. Nature 375, 316–318 (1995).
55. 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).
56. Masuda, T. et al. Gene regulation: Transcription factors LRF and BCL11A independently repress expression of fetal hemoglobin. Science (80-.). 351, 285–289 (2016).
57. Sankaran, V. G. et al. Human fetal hemoglobin expression is regulated by the developmental stage-specific repressor BCL11A.Sankaran, V. G., Menne, T. F., Xu, J., Akie, T. E., Lettre, G., Van Handel, B., … Orkin, S. H. (2008). Human fetal hemoglobin expression is regulated by th. Science 322, 1839–42 (2008).
58. van der Ploeg, L. H. T. & Flavell, R. A. DNA methylation in the human γδβ-globin locus in erythroid and nonerythroid tissues. Cell 19, 947–958 (1980).
59. Kingsley, P. D. et al. "Maturational" globin switching in primary primitive erythroid cells. Blood 107, 1665–72 (2006).
60. Zentner, G. E. & Henikoff, S. Regulation of nucleosome dynamics by histone modifications. Nature Structural and Molecular Biology vol. 20 259–266 (2013).
61. DeSimone, J., Heller, P., Hall, L. & Zwiers, D. 5-Azacytidine stimulates fetal hemoglobin synthesis in anemic baboons. Proc. Natl. Acad. Sci. 79, 4428–4431 (1982).
62. McCaffrey, P. G., Newsome, D. A., Fibach, E., Yoshida, M. & Su, M. S. Induction of gamma-globin by histone deacetylase inhibitors. Blood 90, 2075–83 (1997).
63. Hsiao, C.-H., Li, W., Lou, T.-F., Baliga, B. S. & Pace, B. S. Fetal hemoglobin induction by histone deacetylase inhibitors involves generation of reactive oxygen species. Exp. Hematol. 34, 264–273 (2006).
64. Banerji, J., Rusconi, S. & Schaffner, W. Expression of a β-globin gene is enhanced by remote SV40 DNA sequences. Cell 27, 299–308 (1981).
65. Tolhuis, B., Palstra, R. J., Splinter, E., Grosveld, F. & De Laat, W. Looping and interaction between hypersensitive sites in the active β-globin locus. Mol. Cell 10, 1453–1465 (2002).
66. Sankaran, V. G. et al. MicroRNA-15a and -16-1 act via MYB to elevate fetal hemoglobin expression in human trisomy 13. Proc. Natl. Acad. Sci. U. S. A. 108, 1519–1524 (2011).
67. Li, B. et al. MIR-144-mediated NRF2 gene silencing inhibits fetal hemoglobin expression in sickle cell disease. Exp. Hematol. 70, 85-96.e5 (2019).

68. Ghosal, S., Das, S. & Chakrabarti, J. Long Noncoding RNAs: New Players in the Molecular Mechanism for Maintenance and Differentiation of Pluripotent Stem Cells. Stem Cells Dev. 22, 2240–53 (2013).
69. Choong, M. L., Yang, H. H. & McNiece, I. MicroRNA expression profiling during human cord blood-derived CD34 cell erythropoiesis. Exp. Hematol. 35, 551–564 (2007).
70. Kim, M. J., Civin, C. I. & Kingsbury, T. J. MicroRNAs as regulators and effectors of hematopoietic transcription factors. Wiley Interdisciplinary Reviews: RNA vol. 10 (2019).
71. Hajdu, S. I. A note from history: The discovery of blood cells. Ann. Clin. Lab. Sci. 33, 237–8 (2003).
72. Boisset, J.-C. & Robin, C. On the origin of hematopoietic stem cells: Progress and controversy. Stem Cell Res. 8, 1–13 (2012).
73. Mathe, G., Amiel, J. L., Schwarzenberg, L., Cattan, A. & Schneider, m. Haematopoietic chimera in man after allogenic (homologous) bone-marrow transplantation. (Control of the secondary syndrome. Specific tolerance due to the chimerism). Br. Med. J. 2, 1633–5 (1963).
74. Till, J. E. & McCulloch, E. A. A direct measurement of the radiation sensitivity of normal mouse bone marrow cells. Radiat. Res. 14, 213–22 (1961).
75. Ogawa, M. Differentiation and proliferation of hematopoietic stem cells. Blood vol. 81 2844–2853 (1993).
76. Giebel, B. & Punzel, M. Lineage development of hematopoietic stem and progenitor cells. Biological Chemistry vol. 389 813–824 (2008).
77. Laurenti, E. & Göttgens, B. From haematopoietic stem cells to complex differentiation landscapes. Nature 553, 418–426 (2018).
78. Varga, E., Hansen, M., van den Akker, E. & von Lindern, M. Erythropoiesis and Megakaryopoiesis in a Dish. in Cell Culture (IntechOpen, 2019). doi:10.5772/intechopen.80638.
79. Haar, J. L. & Ackerman, G. A. A phase and electron microscopic study of vasculogenesis and erythropoiesis in the yolk sac of the mouse. Anat. Rec. 170, 199–223 (1971).
80. Palis, J. Primitive and definitive erythropoiesis in mammals. Front. Physiol. 5, 3 (2014).
81. Granick, S. & Levere, R. D. Heme synthesis in erythroid cells. Progress in hematology vol. 4 1–47 (1964).
82. Gifford, S. C., Derganc, J., Shevkoplyas, S. S., Yoshida, T. & Bitensky, M. W. A detailed study of time-dependent changes in human red blood cells: from reticulocyte maturation to erythrocyte senescence. Br. J. Haematol. 135, 395–404 (2006).
83. Hu, J. et al. Isolation and functional characterization of human erythroblasts at distinct stages: implications for understanding of normal and disordered erythropoiesis in vivo. Blood 121, 3246–3253 (2013).
84. Tsiftsoglou, A. S., Vizirianakis, I. S. & Strouboulis, J. Erythropoiesis: Model systems, molecular regulators, and developmental programs. IUBMB Life vol. 61 800–830 (2009).
85. Huang, E. J., Nocka, K. H., Buck, J. & Besmer, P. Differential expression and processing of two cell associated forms of the kit-ligand: KL-1 and KL-2. Mol. Biol. Cell 3, 349–362 (1992).
86. Linenberger, M. L. et al. Stem cell factor production by human marrow stromal filbroblasts. Exp. Hematol. 23, 1104–1114 (1995).
87. Broudy, V. C. Stem cell factor and hematopoiesis. Blood 90, 1345–64 (1997).
88. Chui, D. H. K. & Loyer, B. V. Erythropoiesis in Steel mutant mice - Effects of erythropoietic in vitro. Blood 45, 427–433 (1975).
89. Chui, D. H. K., Liao, S. K. & Walker, K. Fetal erythropoiesis in Steel mutant mice - Defect in differentiation from BF-E to CFU-E during early development. Blood 51, 539–547 (1978).

90. Leary, A. G., Hai Qun Zeng, Clark, S. C. & Ogawa, M. Growth factor requirements for survival in G0 and entry into the cell cycle of primitive human hemopoietic progenitors. Proc. Natl. Acad. Sci. U. S. A. 89, 4013–4017 (1992).

91. Migliaccio, G. et al. Long-term generation of colony-forming cells in liquid culture of CD34+ cord blood cells in the presence of recombinant human stem cell factor. Blood 79, 2620–2627 (1992).

92. Bernstein, I. D., Andrews, R. G. & Zsebo, K. M. Recombinant human stem cell factor enhances the formation of colonies by CD34+ and CD34+lin- cells, and the generation of colony-forming cell progeny from CD34+lin- cells cultured with interleukin-3, granulocyte colony-stimulating factor, or granulocyte-m. Blood 77, 2316–2321 (1991).

93. Miller, J. S., McCullar, V., Punzel, M., Lemischka, I. R. & Moore, K. A. Single Adult Human CD34+/Lin-/CD38- Progenitors Give Rise to Natural Killer Cells, B-Lineage Cells, Dendritic Cells, and Myeloid Cells. Blood 93, 96–106 (1999).

94. Wu, H., Klingmüller, U., Acurio, A., Hsiao, J. G. & Lodish, H. F. Functional interaction of erythropoietin and stem cell factor receptors is essential for erythroid colony formation. Proc. Natl. Acad. Sci. U. S. A. 94, 1806–1810 (1997).

95. Muta, K., Krantz, S. B., Bondurant, M. C. & Dai, C. H. Stem cell factor retards differentiation of normal human erythroid progenitor cells while stimulating proliferation. Blood 86, 572–580 (1995).

96. Peschle, C. et al. c-kit Ligand reactivates fetal hemoglobin synthesis in serum-free culture of stringently purified normal adult burst-forming unit-erythroid. Blood 81, 328–336 (1993).

97. Adamson, J. W., Eschbach, J. & Finch, C. A. The kidney and erythropoiesis. Am. J. Med. 44, 725–733 (1968).

98. Koury, M. J. & Bondurant, M. C. Erythropoietin retards DNA breakdown and prevents programmed death in erythroid progenitor cells. Science (80-.). 248, 378–381 (1990).

99. Dolznig, H. et al. Apoptosis protection by the Epo target Bcl-XL allows factor-independent differentiation of primary erythroblasts. Curr. Biol. 12, 1076–1085 (2002).

100. Burns, S. et al. Purification and characterization of the yeast-expressed erythropoietin mutant Epo (R103A), a specific inhibitor of human primary hematopoietic cell erythropoiesis. Blood 99, 4400–5 (2002).

101. Menon, M. P., Fang, J. & Wojchowski, D. M. Core erythropoietin receptor signals for late erythroblast development. Blood 107, 2662–72 (2006).

102. Liang, R. et al. A Systems Approach Identifies Essential FOXO3 Functions at Key Steps of Terminal Erythropoiesis. PLoS Genet. 11, e1005526 (2015).

103. Stefanetti, R. J., Voisin, S., Russell, A. & Lamon, S. Recent advances in understanding the role of FOXO3 [version 1; referees: 4 approved]. F1000Research vol. 7 (2018).

104. Zito, G. E. & Lynch, E. C. Prednisone-responsive congenital erythroid hypoplasia. JAMA 237, 991–2 (1977).

105. Liang, R., Chan, T. K. & Todd, D. Childhood Acute Lymphoblastic Leukaemia and Aplastic Anaemia. Leuk. Lymphoma 13, 411–415 (1994).

106. Axelrod, J. & Reisine, T. Stress hormones: their interaction and regulation. Science (80-.). 224, 452–459 (1984).

107. Tronche, F., Kellendonk, C., Reichardt, H. M. & Schütz, G. Genetic dissection of glucocorticoid receptor function in mice. Curr. Opin. Genet. Dev. 8, 532–538 (1998).

108. Bauer, A. et al. The glucocorticoid receptor is required for stress erythropoiesis. Genes Dev. 13, 2996–3002 (1999).

109. Knoedler, J. R. & Denver, R. J. Krüppel-like factors are effectors of nuclear receptor signaling. General and Comparative Endocrinology vol. 203 49–59 (2014).

110. Bessis, M. Erythroblastic island, functional unity of bone marrow. Rev. Hematol. 13, 8–11 (1958).

111. Le Charpentier, Y. & Prenant, M. Isolation of erythroblastic islands. Study by optical and scanning electron microscopy. Nouv. Rev. Fr. Hematol. 15, 119–40 (1975).

112. Lee, S. H. et al. Isolation and immunocytochemical characterization of human bone marrow stromal macrophages in hemopoietic clusters. J. Exp. Med. 168, 1193–1198 (1988).

113. Heideveld, E. & van den Akker, E. Digesting the role of bone marrow macrophages on hematopoiesis. Immunobiology 222, 814–822 (2017).

114. Manwani, D. & Bieker, J. J. Chapter 2 The Erythroblastic Island. in Current topics in developmental biology vol. 82 23–53 (2008).

115. Soni, S. et al. Absence of Erythroblast Macrophage Protein (Emp) Leads to Failure of Erythroblast Nuclear Extrusion. J. Biol. Chem. 281, 20181–20189 (2006).

116. Van Rooijen, N. & Van Kesteren-Hendrikx, E. Clodronate liposomes: Perspectives in research and therapeutics. in Journal of Liposome Research vol. 12 81–94 (2002).

117. Porcu, S. et al. Klf1 affects DNase II-alpha expression in the central macrophage of a fetal liver erythroblastic island: a non-cell-autonomous role in definitive erythropoiesis. Mol. Cell. Biol. 31, 4144–54 (2011).

118. Mao, X., Shi, X., Liu, F., Li, G. & Hu, L. Evaluation of erythroblast macrophage protein related to erythroblastic islands in patients with hematopoietic stem cell transplantation. Eur. J. Med. Res. 18, 9 (2013).

119. Li, W. et al. Identification and transcriptome analysis of erythroblastic island macrophages. Blood 134, 480–491 (2019).

120. Rhodes, M. M., Kopsombut, P., Bondurant, M. C., Price, J. O. & Koury, M. J. Adherence to macrophages in erythroblastic islands enhances erythroblast proliferation and increases erythrocyte production by a different mechanism than erythropoietin. Blood 111, 1700–1708 (2008).

121. 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, 3494–3504 (1994).

122. Heideveld, E. et al. Glucocorticoids induce differentiation of monocytes towards macrophages that share functional and phenotypical aspects with erythroblastic island macrophages. Haematologica 103, 395–405 (2018).

123. Heideveld, E. et al. CD14+ cells from peripheral blood positively regulate hematopoietic stem and progenitor cell survival resulting in increased erythroid yield. Haematologica 100, 1396–1406 (2015).

124. Philipsen, S., Talbot, D., Fraser, P. & Grosveld, F. The beta-globin dominant control region: hypersensitive site 2. EMBO J. 9, 2159–67 (1990).

125. Pruzina, S., Hanscombe, O., Whyatt, D., Grosveld, F. & Philipsen, S. Hypersensitive site 4 of the human β globin locus control region. Nucleic Acids Res. 19, 1413–1419 (1991).

126. Talbot, D. & Grosveld, F. The 5'HS2 of the globin locus control region enhances transcription through the interaction of a multimeric complex binding at two functionally distinct NF-E2 binding sites. EMBO J. 10, 1391–8 (1991).

127. Ellis, J. et al. A dominant chromatin-opening activity in 5' hypersensitive site 3 of the human beta-globin locus control region. EMBO J. 15, 562–8 (1996).

128. Bank, A. Regulation of human fetal hemoglobin: new players, new complexities. Blood 107, 435–43 (2006).

129. Wijgerde, M., Grosveld, F. & Fraser, P. Transcription complex stability and chromatin dynamics in viv. Nature 377, 209–213 (1995).

130. Palstra, R.-J. et al. The β-globin nuclear compartment in development and erythroid differentiation. Nat. Genet. 35, 190–194 (2003).

131. Dekker, J., Rippe, K., Dekker, M. & Kleckner, N. Capturing chromosome conformation. Science (80-.). 295, 1306–1311 (2002).

132. Drissen, R. et al. The active spatial organization of the beta-globin locus requires the transcription factor EKLF. Genes Dev. 18, 2485–90 (2004).

133. Im, H. et al. Chromatin domain activation via GATA-1 utilization of a small subset of dispersed GATA motifs within a broad chromosomal region. Proc. Natl. Acad. Sci. U. S. A. 102, 17065–17070 (2005).

134. Deng, W. et al. Reactivation of Developmentally Silenced Globin Genes by Forced Chromatin Looping. Cell 158, 849–860 (2014).

135. Breda, L. et al. Forced chromatin looping raises fetal hemoglobin in adult sickle cells to higher levels than pharmacologic inducers. Blood 128, 1139–1143 (2016).

136. Wienert, B. et al. KLF1 drives the expression of fetal hemoglobin in British HPFH. Blood 130, 803–807 (2017).

137. Chad, M. et al. Therapeutic haemoglobin synthesis in β-thalassaemic mice expressing lentivirus-encoded human β-globin. Nature 406, 82–86 (2000).

138. Pawliuk, R. et al. Correction of sickle cell disease in transgenic mouse models by gene therapy. Science (80-.). 294, 2368–2371 (2001).

139. Cavazzana, M., Antoniani, C. & Miccio, A. Gene Therapy for β-Hemoglobinopathies. Molecular Therapy vol. 25 1142–1154 (2017).

140. Miccio, A. et al. In vivo selection of genetically modified erythroblastic progenitors leads to long-term correction of β-thalassemia. Proc. Natl. Acad. Sci. U. S. A. 105, 10547–10552 (2008).

141. Levasseur, D. N., Ryan, T. M., Pawlik, K. M. & Townes, T. M. Correction of a mouse model of sickle cell disease: Lentiviral/antisickling β-globin gene transduction of unmobilized, purified hematopoietic stem cells. Blood 102, 4312–4319 (2003).

142. Negre, O. et al. Preclinical Evaluation of Efficacy and Safety of an Improved Lentiviral Vector for the Treatment of Beta-Thalassemia and Sickle Cell Disease. Curr. Gene Ther. 15, 64–81 (2014).

143. Carden, M. A. & Little, J. Emerging disease-modifying therapies for sickle cell disease. Haematologica vol. 104 1710–1719 (2019).

144. Telen, M. J., Malik, P. & Vercellotti, G. M. Therapeutic strategies for sickle cell disease: towards a multi-agent approach. Nature Reviews Drug Discovery vol. 18 139–158 (2019).

145. Miller, I. J. & Bieker, J. J. A novel, erythroid cell-specific murine transcription factor that binds to the CACCC element and is related to the Krüppel family of nuclear proteins. Mol. Cell. Biol. 13, 2776–2786 (1993).

146. Chen, X. & Bieker, J. J. Stage-specific repression by the EKLF transcriptional activator. Mol. Cell. Biol. 24, 10416–24 (2004).

147. 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).

148. 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).

149. Shyu, Y.-C. et al. Tight regulation of a timed nuclear import wave of EKLF by PKCθ and FOE during Pro-E to Baso-E transition. Dev. Cell 28, 409–22 (2014).

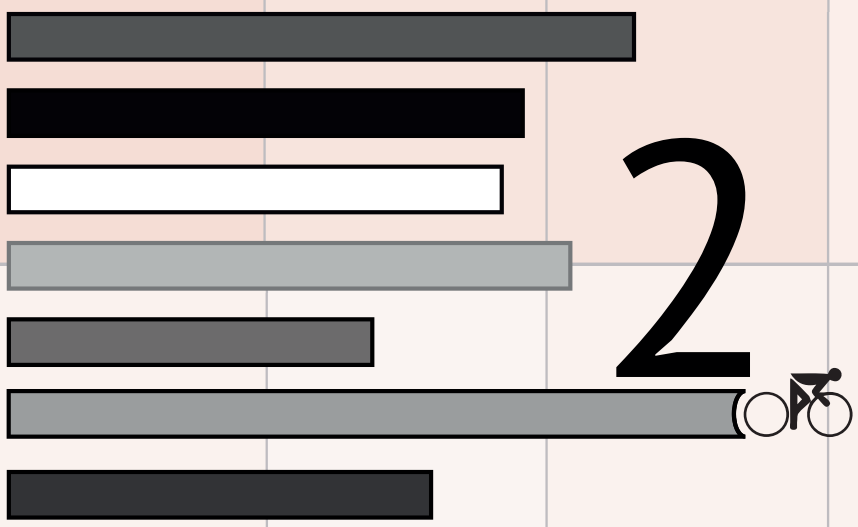
150. Perkins, A. et al. Krüppeling erythropoiesis: an unexpected broad spectrum of human red blood cell disorders due to KLF1 variants. Blood 127, 1856–62 (2016).

151. Borg, J. et al. Haploinsufficiency for the erythroid transcription factor KLF1 causes hereditary persistence of fetal hemoglobin. Nat. Genet. 42, 801–5 (2010).

152. 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).
153. Tallack, M. R. et al. A global role for KLF1 in erythropoiesis revealed by ChIP-seq in primary erythroid cells. Genome Res. 20, 1052–1063 (2010).
154. 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).
155. Frontelo, P. et al. Novel role for EKLF in megakaryocyte lineage commitment. Blood 110, 3871–80 (2007).
156. Tallack, M. R. & Perkins, A. C. Megakaryocyte-erythroid lineage promiscuity in EKLF null mouse blood. Haematologica 95, 144–7 (2010).
157. Starck, J. et al. Functional cross-antagonism between transcription factors FLI-1 and EKLF. Mol. Cell. Biol. 23, 1390–402 (2003).
158. Bouilloux, F. et al. EKLF restricts megakaryocytic differentiation at the benefit of erythrocytic differentiation. Blood 112, 576–584 (2008).
159. 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, 1197–1199 (2007).
160. Uda, M. et al. Genome-wide association study shows BCL11A associated with persistent fetal hemoglobin and amelioration of the phenotype of -thalassemia. Proc. Natl. Acad. Sci. 105, 1620–1625 (2008).
161. Mabaera, R. et al. A cell stress signaling model of fetal hemoglobin induction: what doesn't kill red blood cells may make them stronger. Exp. Hematol. 36, 1057–72 (2008).
162. Xu, J. et al. Bcl11a Deficiency Leads to Hematopoietic Stem Cell Defects with an Aging-like Phenotype. Cell Rep. 16, 3181–3194 (2016).
163. Sher, F. et al. Rational targeting of a NuRD subcomplex guided by comprehensive in situ mutagenesis. Nat. Genet. 51, 1149–1159 (2019).
164. Martyn, G. E. et al. A natural regulatory mutation in the proximal promoter elevates fetal globin expression by creating a de novo GATA1 site. Blood 133, (2019).
165. Liu, N. et al. Direct Promoter Repression by BCL11A Controls the Fetal to Adult Hemoglobin Switch. Cell 173, (2018).
166. Skene, P. J., Henikoff, J. G. & Henikoff, S. Targeted in situ genome-wide profiling with high efficiency for low cell numbers. Nat. Protoc. 13, 1006–1019 (2018).
167. Norton, L. J. et al. KLF1 directly activates expression of the novel fetal globin repressor ZBTB7A/LRF in erythroid cells. Blood Adv. 1, 685–692 (2017).
168. Constantinou, C. et al. The multi-faceted functioning portrait of LRF/ZBTB7A. Human Genomics vol. 13 (2019).
169. Maeda, T. et al. LRF Is an Essential Downstream Target of GATA1 in Erythroid Development and Regulates BIM-Dependent Apoptosis. Dev. Cell 17, 527–540 (2009).
170. Funnell, A. P. W. et al. P15-p16.1 microdeletions encompassing and proximal to BCL11A are associated with elevated HbF in addition to neurologic impairment. Blood 126, 89–93 (2015).
171. Dias, C. et al. BCL11A Haploinsufficiency Causes an Intellectual Disability Syndrome and Dysregulates Transcription. Am. J. Hum. Genet. 99, 253–274 (2016).
172. Luc, S. et al. Bcl11a Deficiency Leads to Hematopoietic Stem Cell Defects with an Aging-like Phenotype. Cell Rep. 16, 3181–3194 (2016).
173. Yu, Y. et al. Bcl11a is essential for lymphoid development and negatively regulates p53. J. Exp. Med. 209, 2467–83 (2012).
174. Bauer, D. E. et al. An erythroid enhancer of BCL11A subject to genetic variation determines fetal hemoglobin level. Science 342, 253–7 (2013).

175. Psatha, N. et al. Disruption of the BCL11A Erythroid Enhancer Reactivates Fetal Hemoglobin in Erythroid Cells of Patients with β-Thalassemia Major. Mol. Ther. - Methods Clin. Dev. 10, 313–326 (2018).

176. Esrick, E. B. et al. Validation of BCL11A As Therapeutic Target in Sickle Cell Disease: Results from the Adult Cohort of a Pilot/Feasibility Gene Therapy Trial Inducing Sustained Expression of Fetal Hemoglobin Using Post-Transcriptional Gene Silencing. Blood 134, LBA-5-LBA-5 (2019).



Chapter 2

Large-scale *in-vitro* production of red blood cells from human peripheral blood mononuclear cells

Steven Heshusius*, Esther Heideveld*, Patrick Burger*, Marijke Thiel-Valkhof, Erica Sellink, Eszter Varga, Elina Ovchynnikova, Anna Visser, Joost H.A. Martens, Marieke von Lindern, and Emile van den Akker
* contributed equally to this work

Blood advances, 2019 Nov 12;3(21):3337-3350.

Abstract

Transfusion of donor-derived red blood cells is the most common form of cellular therapy. Donor availability and the potential risk of alloimmunization and other transfusion-related complications may, however, limit the availability of transfusion units especially for chronically transfused patients. *In-vitro* cultured, customizable red blood cells would negate these concerns and further increase precision medicine. Large-scale, cost effective production depends on optimization of culture conditions. We developed a defined medium and adapted our protocols to GMP culture requirements, which reproducibly provided pure erythroid cultures from peripheral blood mononuclear cells without prior $CD34^+$ isolation, and a $3x10^7$-fold increase in erythroblasts in 25 days (or from 100 million PBMC, 2-4ml packed red cells can be produced). Expanded erythroblast cultures could be differentiated to $CD71^{dim}CD235a^+CD44^+CD117^-DRAQ5^-$ red blood cells in 12 days. More than 90% of the cells enucleated and expressed adult hemoglobin as well as the correct blood group antigens. Deformability and oxygen binding capacity of cultured red blood cells was comparable to *in-vivo* reticulocytes. Daily RNA sampling during differentiation followed by RNA-seq provided a high-resolution map/resource of changes occurring during terminal erythropoiesis. The culture process was compatible with upscaling using a G-Rex bioreactor with a capacity of 1L per reactor, allowing transition towards clinical studies and small-scale applications.

Introduction

Blood transfusion is the most applied cellular therapy, with more than 80 million transfusion units administered worldwide each year[1]. Inherent risks of donor-transfusion material are alloimmunization and presence of blood-borne diseases. Oxygen-carrier substitutes have shown to be applicable in case of immediate emergency but cannot replace long term blood transfusions[2]. The potential to culture red blood cells (RBC) for transfusion purposes has been recognized for a long time[3-10]. Transfusion medicine and the care of chronic transfusion patients with prophylactic antigen matching has already substantially decreased the rate of alloimmunization (<5%). There are many variables that result in alloimmunization, including access to centers who are molecularly typing both donors and recipients to precisely match the unit to the patient. Cultured RBC (cRBC) that are antigen-compatible will decrease the risk of alloimmunization in patients. Cost-effective, large-scale culture of blood group matched RBC will provide a degree of donor independency and minimization of donor-patient blood type variation. In addition, *c*RBC can be used as vehicles for enzyme replacement therapy[11], or as therapeutic delivery systems targeting specific body parts[12]. Several groups already cultured enucleated cRBC from cord blood $CD34^+$ cells[13-15]. However, these cells produce fetal hemoglobin with a higher tendency to denature and to cause membrane damage compared to adult hemoglobin[16]. We have previously shown that enucleated cRBC can be generated starting from adult peripheral blood mononuclear cells (PBMC), a better accessible source than cord blood $CD34^+$ cells and allows adult autologous cRBC[17]. Importantly, the erythroid yield from PBMC is 10-15 fold increased compared to $CD34^+$ cells isolated from a similar amount of PBMC, due to support from $CD14^+$ cells present in PBMC[17-19].

One transfusion unit contains about $2x10^{12}$ RBC, reflecting the high requirement for erythroblast expansion to obtain sufficient numbers of cRBC. Previous attempts to culture the required number of enucleated cRBC from $CD34^+$ cells isolated from PBMC were hampered by low expansion or poor enucleation[20,21]. Expansion of $CD71^{high}CD235a^{dim}$ erythroblasts can be prolonged by exploiting the cooperative action of erythropoietin (EPO), stem cell factor (SCF) and glucocorticoids involved in stress-erythropoiesis in a serum/plasma-free environment[7,17,18,22,23], while differentiation is induced by increasing concentrations of EPO and dispensing with SCF and glucocorticoids. Here, we describe a three stage GMP-grade culture protocol using culture dishes or G-Rex bioreactors, both with high expansion and enucleation to generate PBMC-derived cRBC. To this end we have developed a completely defined GMP-grade medium. This three-stage culture protocol can be used for small-scale GMP-grade production, yielding >90% enucleated reticulocytes with adult hemoglobinization.

Material and Methods

Cell culture

Human PBMC from whole blood were purified by density separation using Ficoll-Paque (manufacturer protocol). Informed consent was given in accordance with the Declaration of Helsinki and Dutch National and Sanquin Internal Ethic Boards. PBMC were seeded at 5-10x10^6 cells/ml (CASY® Model TCC; Schärfe System GmbH, Reutlingen, Germany) in Cellquin medium based on HEMA-Def[7,17] with significant modification (Table S1 lists all components) supplemented with erythropoietin (2U/ml; ProSpec, East Brunswick, NJ), human recombinant stem cell factor (100ng/ml; ITK Diagnostics BV, Uithoorn, The Netherlands), dexamethasone (1µM; Sigma, St. Louis, MO) and 0.1% human albumin (cHA; kindly provided by Sanquin Plasma Products, Amsterdam, The Netherlands, perturbation with albumin see Supplementary Material), referred to as Expansion Medium (EM). Interleukin 3 (IL3) was added (1ng/ml; Miltenyi Biotec, Bergisch Gladbach, Germany) to EM on the first day (stage I). Media was partially replenished every two days with EM. Around day 6, upon erythroblasts detection, the cells were maintained 1-2x10^6 cells/ml for 15-25 days (stage II). Erythroblasts differentiation (stage III) was induced in Differentiation Medium (DM) containing Cellquin supplemented with erythropoietin (10U/ml), 5% Omniplasma (Octopharma, Wien, Austria), holotransferrin (700µg/ml; Sanquin) and heparin (5U/ml; LEO Pharma BV, Breda, The Netherlands). At day two, half a medium change was performed. At day 5, storage components (Sanquin Plasma Products, Amsterdam, The Netherlands) were added and media was refreshed (half) every two days until fully differentiated at day 10-12 of differentiation. For reticulocyte filtration see Supplementary Material.

Flow cytometry

Cells were washed and resuspended in HEPES buffer supplemented with 1% HA. Cells were incubated with primary antibodies for 30 minutes at 4 °C, measured on FACS Canto II or LSRFortessa (both BD Biosciences, Oxford, UK) and analyzed using FlowJo software (FlowJo v10, Ashland, OR, USA). Reticulocyte RNA was stained with Thiazole orange (TO; Sigma) as described previously Abcam, Cambridge, UK). Antibodies are listed in Supplementary Material., Cambridge, UK). Antibodies are listed in Supplementary Material.

RBC deformability

RBC deformability was measured by the Automated Rheoscope and Cell Analyzer (ARCA) as described previously[25]. 10 Pa shear stress was used and 3000 cells were measured and grouped in 30 bins according to increasing elongation or cell projection area (as a measure of membrane surface area). Both the extent of elongation (major cell radius

divided by minor cell radius) and the area (in mm^2) was plotted against the normalized frequency of occurrence.

HPLC

Culture lysates were prepared and stored at -80 °C prior to analysis as described previously[25]. In short, hemoglobin separation was performed by high performance cation exchange liquid chromatography (HPLC) on Waters Alliance 2690 equipment (Waters, Milford, MA, USA) using30min elution over a combined 20-200mM NaCl and pH7.0-6.6 gradient in 20mM BisTris/HCl and 2mM KCN. A PolyCAT A 100/4.6mm, 3mm, 1500Å column (PolyLC, Columbia, MD, USA).

Coomassie

Ghosts (RBC membranes) were generated as described before[26], subjected to SDS-PAGE gels (Bio-Rad, Hercules, CA, USA) and total proteins were stained with Coomassie brilliant blue. In short, proteins were fixed in 30% ethanol, 2% (v/v) phosphoric acid overnight, washed two times 10minutes in 2% phosphoric acid and equilibrate for 30min in 2% phosphoric acid containing 18% ethanol and 15% (w/v) ammonium sulphate. Gels were stained by diluting Coomassie Blue G-250 dissolved in water (0.2%) slowly to a final concentration of 0.02% (0.2mg/ml).

Cytospins

Cells were cytospun using Shandon Cytospin II (Thermo Scientific), dried and fixed in methanol. Cells were stained with benzidine in combination with the Differential Quick Stain Kit (PolySciences, Warrington, PA) (manufacturers protocol). Slides were dried, subsequently embedded in Entellan (Merck-Millipore) and covered with a coverslip. Images were taken using microscope DM2500 with 40x or 10x object (Leica DM-2500; Germany).

RNA-seq analysis

Erythroid cultures on differentiation medium were sampled daily for 12 days. Sequencing libraries were prepared using Trizol RNA isolation, cDNA amplification and rRNA depletion using HyperPrep Kit with RiboErase (KAPA Biosystems, Pleasanton, CA, USA) as described by manufactures. Reads were mapped to ChGR38.v85 and differential expression analysis was performed with EdgeR[27]. A detailed description can be found in Supplementary Material.

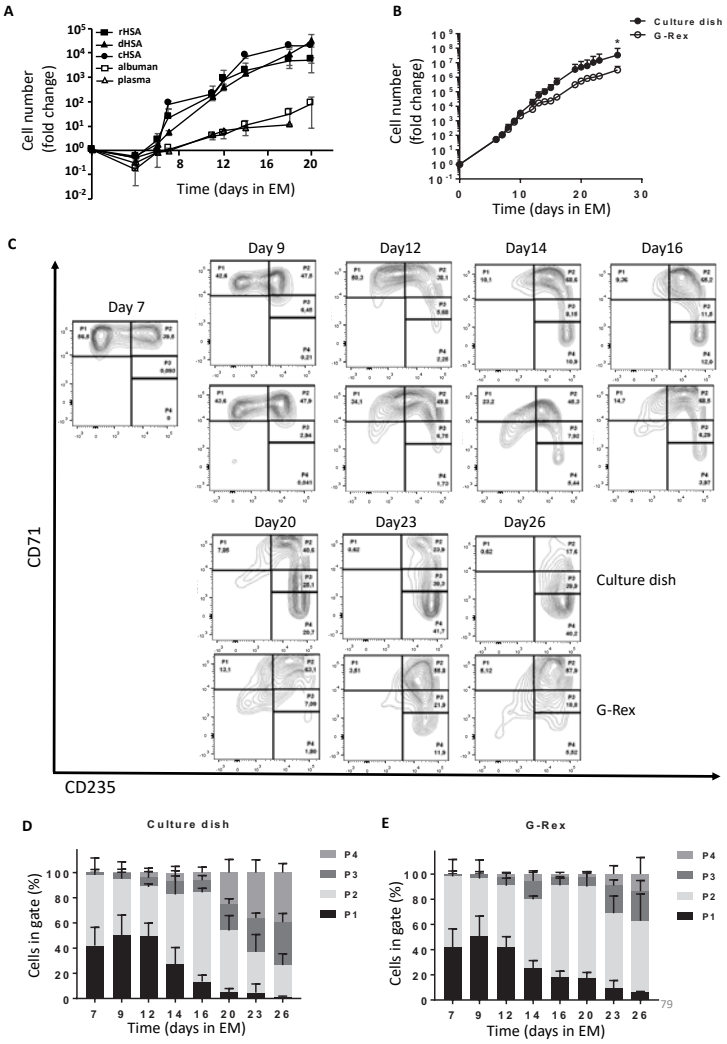


Figure 1. Efficient expansion of erythroblasts in plasma/serum-free GMP-grade medium. (A) PBMC were cultured towards erythroblasts in GMP-grade medium supplemented with EPO, SCF and Dex (EM) for 26 days. Medium was prepared using either ultra-clean HA (cHA), detoxified HA (dHA), recombinant HA (rHA), Albuman® or plasma. Cell counts at day 0 were normalized to 1 PBMC at the start of culture. Cultures were kept at 0.7 to 2x10^6 by dilution. Symbols indicate mean fold-change at any day compared to 1 PBMC seeded, error bars indicate SD (n=4). **(B)** PBMC from 4 independent donors were cultured from PBMC in Cellquin medium (cHA) supplemented with EPO (2U/ml), SCF (100ng/ml) and Dex (1µM) in culture dishes until a pure erythroblast population was obtained (at day 7). Erythroblasts were further expanded in a G-Rex bioreactor or in culture dishes. Mean fold-change (± SD) was calculated and compared (two-way ANOVA, * P<0.05; n=4). **(C)** Representative dot plots indicating cell surface expression levels of CD71 and CD235 in cultures as described in B. Quadrants are labeled (P1-P4) and relative cell numbers per quadrant indicated as percentage **(D-E)** quantification of percentages per quadrant in dot blots similar to those shown in C. Error bars indicate SD (n=4) D: cells cultured in dishes. E: cells cultured in G-Rex

Results

Large-scale erythroblast expansion from PBMC using a G-Rex bioreactor and GMP-grade medium

To establish medium conditions to obtain and prolong erythroid expansion from adult PBMC, we first tested distinct sources of albumin. Different expansion media (EM, Table S1) with 0.1% ultra-clean human albumin obtained after plasma fractionation (cHA), 2,5% plasma, 2,5% Albuman®, 0.1% detoxified Albuman (dHA), or 0.1% recombinant HA (rHA) were used (Figure 1A and Figure S1A). Plasma or Albuman® resulted in (i) low erythroblast yield, (ii) presence of non-erythroid cells (negative for CD71 and CD235) and (iii) premature differentiation of erythroblasts, indicated by a loss of CD71 expression in conjunction with increased CD235a expression (Figure S1A)[28,29]. In contrast, EM supplemented with cHA, dHA, or rHA showed a significantly increased erythroid expansion potential with limited spontaneous differentiation and a complete absence of non-erythroid cells (Figure 1A and Figure S1A). Of note, PBMC contain primarily T-cells, myeloid cells and B-cells and only on average 0.16% $CD34^+$ HSPC that are capable of differentiating into erythroid cells[17,19]. Therefore, expansion curves using PBMC as a starting material show a drop in expansion between day 0 and day 5, caused by a loss of these non-proliferating immune effector cells[17,18]. By consequence, a fold change increase of erythroid cells from PBMC of circa 10^5-fold corresponds to a circa 10^8-fold increase from the $CD34^+$ cell compartment.

Large scale cRBC production in culture dishes is practically impossible. Therefore, a G-Rex bioreactor from Wilson Wolf Manufacturing (Saint Paul, MN, USA) was used in which a gas-permeable membrane at the bottom allowed proliferation in larger volume/surface conditions[30,31]. The G-Rex bioreactor does not support adherent cells, such as the $CD14^+$ PBMC that promote erythroid yield by increasing $CD34^+$ cell survival[18], which compromises stage I yield (data not shown). Therefore, PBMC cultures were started in culture dishes/cell stacks until an erythroblast population was obtained around day 7 of culture in EM supplemented with cHA. Subsequently (stage II), cultured erythroblasts were either transferred G-Rex systems or retained in culture dishes and could be maintained for at least 26 days (Figure 1B). Transfer to a G-Rex bioreactor briefly delayed erythroid expansion but showed a similar expansion rate between day 15 and 25. Although the G-Rex bioreactor yielded slightly less cells ($3x10^6$ *vs.* $3x10^7$), less donor variation was observed. Erythroid cells can be staged from $CD71^{high}CD235^{dim}$ erythroblasts (P1) to enucleated $CD71^-CD235^+$ reticulocytes (P4), with intermediate stages in which cells are characterized by increased CD235a and reduced CD71 expression (Figure 1C)[29]. At day 7, no non-erythroid cells ($CD71^-CD235^-$) were observed, indicating a pure erythroid culture, (day 0 to day 7 culture progression was previously published by our laboratory[18]). During erythroblast expansion in stage II, erythroblasts

gained expression of CD235a from day 10 onwards resulting in a pure population of $CD71^{high}CD235^{+}$ erythroblasts, that could be expanded for at least 26 days (Figure 1C-E and Figure S1A). Previously, we showed that these $CD71^{high}CD235^{+}$ erythroblasts, cultured in presence of Epo/SCF/Dex, express low levels of hemoglobin, and have a morphology overlapping pro-erythroblasts and basophilic erythroblasts. When using the general term 'erythroblasts' in this manuscript, we indicate this stage. Expansion of erythroblasts in EM medium limits their differentiation to $CD71^{low}$ polychromatic and orthochromatic erythroblasts to a small fraction of cells that escapes the differentiation block. During prolonged culture in EM spontaneously differentiating cells accumulate (figure 1C-E). Interestingly, cells expanded in the G-Rex system maintained a less differentiated state for a prolonged culture period (>16days), as shown by delayed CD235a upregulation and maintenance of high CD71 expression (Figure 1E). In conclusion, a defined GMP-grade medium enabled pure erythroblast cultures from a mixed PBMC cell pool with significantly delayed onset of spontaneous differentiation, resulting in a large expansion potential in both culture dishes and a G-Rex bioreactor.

Erythroid cultures from G-Rex and normal culture dishes fully enucleate

Differentiation of erythroblasts to polychromatic erythroblasts, orthochromatic erythroblasts, and eventually to enucleated reticulocytes (stage III) is induced by removing SCF and dexamethasone whilst increasing the EPO and holo-transferrin concentration and supplementing with 5% Omniplasma. Heparin is added to prevent medium from clotting (differentiation medium, DM). Erythroblast differentiation is characterized by i) a transient proliferation burst with decreased cell-cycle time resulting in a reduced cell volume, ii) hemoglobinization, iii) erythroid specific protein expression of e.g., blood group antigens, and iv) enucleation[13,17,28,29,32]. During the first days of differentiation, in particular in culture dishes, proliferation was observed followed by cell growth arrest (Figure 2A). Note that the number of cells after the initial short proliferation burst did not decrease, suggesting no negative effect on culture viability. In contrast, cells differentiated in a G-Rex bioreactor did not show increased expansion.

During erythroblast differentiation, both erythroid cells cultured in dishes and the G-Rex system increased expression of CD235a and lost expression of CD71, although erythroblasts differentiated slightly faster in culture dishes (Figure 2B-D). Differentiation was accompanied with a decrease in cell size (FSC-A). Both CD71 and c-KIT (CD117) expression increased at day 1 followed by a sharp down regulation (Figure S2A-B). Furthermore, CD44 was progressively reduced in expression as reported before[17]. Although CD235 expression increases in early differentiation as cells become smaller it decreases slightly, which may be due to loss of membrane surface during enucleation.

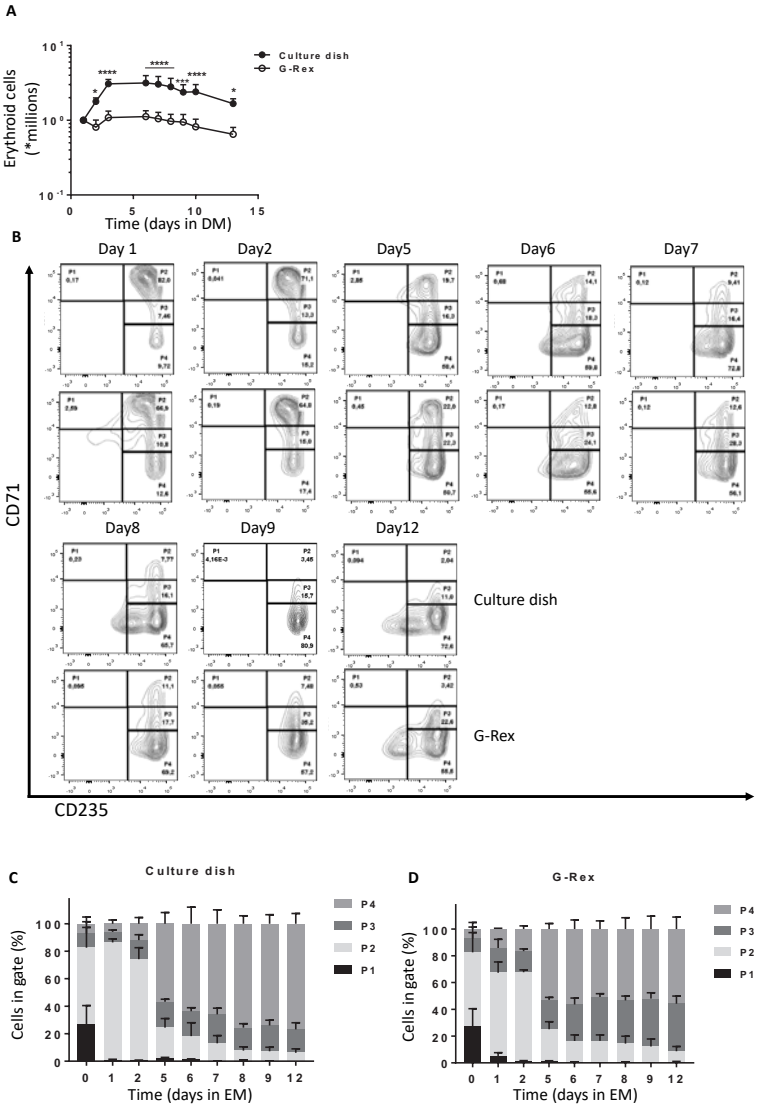


Figure 2. Differentiation of erythroblasts in culture dishes or a G-Rex bioreactor. (A) Erythroblast cultures were established in culture dishes. Erythroblast were washed and reseeded at $1x10^6$/ml in Cellquin medium supplemented with EPO (10U/ml), Transferrin (700µg/ml), 5% human plasma, and heparin (5U/ml) (Differentiation medium, DM) in culture dishes (closed symbol) or G-Rex (open symbol). Erythroblasts were differentiated for 12 days. Cell density was measured at days indicated. Mean cell numbers were calculated. Error bars indicate SD Cell expansion was compared by two-way ANOVA, **P*<0.05, ****P*<0.001, *****P*<0.0001; n=4. **(B)** Representative dot plots indicating cell surface expression levels of CD71 and CD235 in cultures as described in A. Quadrants are labeled (P1-P4) and relative cell numbers per quadrant indicated as percentage **(C-D)** quantification of percentages per quadrant in dot blots similar to those shown in C. Error bars indicate SD (n=4) C: cells cultured in dishes. D: cells cultured in G-Rex

Factors that affect enucleation at the end of erythroid differentiation

At blood banks, RBC are stored in media that contain specific components protecting viability, which were not present in DM. Therefore, a specific storage component solution was added at day 5 of differentiation, when the first reticulocytes arose in culture. This increased the number and frequency of enucleated cRBC, particularly in the G-Rex bioreactor (data not shown). Nonetheless, initial differentiation experiments in the G-Rex system yielded low numbers of enucleated cells compared to culture dishes (Figure S3A). Importantly, EM is completely replaced by DM upon initiating of differentiation in dishes, whereas only 90% EM could be replaced with DM in the G-Rex system. Indeed, replacing 90% of the culture medium in dishes, resulted in a reduction of enucleation, compared to complete medium replacement (56% vs 85%; Figure S3B). Reduction of dexamethasone (from 1μM to 10nM) two days prior to differentiation induction did not affect enucleation (Figure S3D). However, SCF removal from the culture medium two days prior to differentiation in culture dishes increased enucleation 1.5-fold (56% *vs.* 85%; Figure S3C). This indicated that residual SCF at the start of differentiation negatively affects enucleation during terminal differentiation. Enucleation was observed from day 5 onwards, resulting in >90% enucleation after 12 days of differentiation in culture dishes and almost 85% enucleation in the G-Rex system (Figure 3A and Figure S3E). A slight difference in the ratio nuclei/cRBC between culture dishes and a G-Rex bioreactor was observed (Figure 3B). The flow cytometry data and cytospin images at the end of differentiation revealed pyrenocytes (nuclei encapsulated by plasma membrane) and some nucleated cells (Figure 3C, D and Figure S3E). Filtration using neonatal leucoreduction filters resulted in 99% removal of nuclei and remaining nucleated cells and yielded a homogenous population of enucleated cRBC that is comparable to native erythrocytes on cytospins (Figure 3C, D). Of note, these cRBC resemble late reticulocytes populations as previously observed by us[19, 24].

cRBC resemble mature reticulocytes

Filtered cRBC displayed a spheroid morphology, indicative of a reticulocyte population (Figure 3C). Expression of the major membrane proteins α/β-spectrin, Band 3, protein 4.1, protein 4.2 and GAPDH were comparable between cRBC from culture dishes and peripheral blood RBC (Figure 4A). Reticulocytes released from the bone marrow have a low deformability, which increases during maturation towards RBC[33]. Both cRBC cultured in culture dishes and the G-Rex system showed a deformability index comparable to late peripheral blood reticulocytes (Figure 4B).

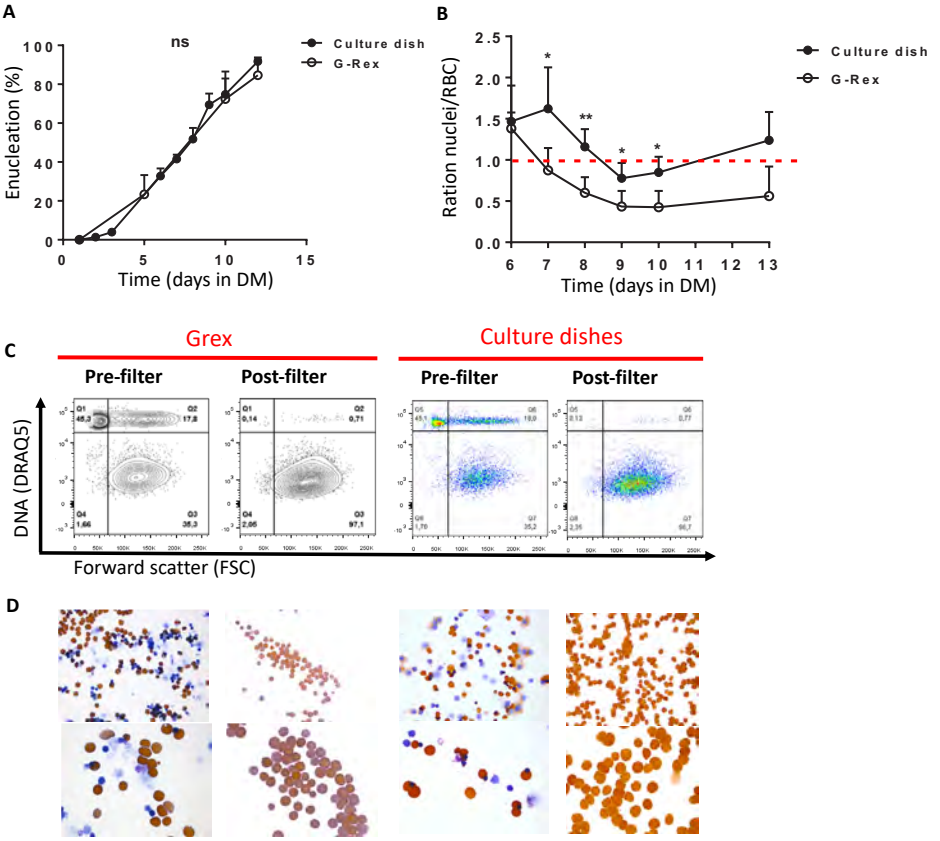


Figure 3. Efficient enucleation is observed after 10 days in differentiation medium. (A-C) enucleated cells and nuclei were measured by flow cytometry, using DRAQ5 DNA staining against size (forward scatter) as shown in C. **(A)** The percentage of enucleated erythroid cells was measured during differentiation in culture dishes (closed symbols) and a G-Rex system (open symbols). Error bars indicate SD (n=4) Comparisons were made by unpaired *t*-test. **(B)** Ratio of nuclei versus cRBC during differentiation in culture dishes or G-Rex (<1 means more cRBC than nuclei). Mean ± SD (unpaired *t*-test, **P*<0.05, ***P*<0.01; n=3-4). **(C)** Enucleation percentages of 10 days differentiated erythroid cultures before and after passage through a leukoreduction filter. Q1=nuclei; Q2=nucleated cells; Q3+Q4=enucleated cRBC. **(D)** Cytospins of the samples analysed in C stained for hemoglobin with benzidine and general cytological stains. As comparison, cytospins of native peripheral blood erythrocytes are shown.

Table 1: Blood group analysis of peripheral blood RBC (RBC) and cultured RBC (cRBC) using flow cytometry.

BLOOD GROUP	RBC 1	RBC 2	RBC 3	CRBC 1	CRBC 2	CRBC 3
A	-	-	+	-	-	+
B	-	-	-	-	-	-
RHD	+	+	-	+	+	-
RHE	-	+	-	-	+	-
RHC	+	-	-	+	-	-
RHE	+	+	+	+	+	+
RHC	+	+	+	+	+	+
K	-	-	-	-	-	-
K	+	+	+	+	+	+
KP-A	-	-	-	-	-	-
KP-B	+	+	+	+	+	+
LU-A	-	-	-	-	-	-
LU-B	+	+	+	+	+	+
P1	+	-	+	+	-	+
S	+	-	-	+	-	-
S	+	+	+	+	+	+
FYA	+	-	+	+	-	+
FYB	-	+	+	-	+	+
JKA	-	+	-	-	+	-
JKB	+	-	+	+	-	+
M	+	-	+	+	-	+
N	+	+	+	+	+	+

Furthermore, cell pellets from cRBC turned dark red and HPLC data showed that cRBC both derived from G-Rex and culture dishes mainly express adult hemoglobin (HbA1 73.4% G-Rex *vs.* 62% dish; HbA2 2.1% G-Rex *vs.* 1.4% dish) and low levels of HbF (8.8% G-Rex *vs.* 7.1% dish; Figure 4C-D). In addition, hemoglobin oxygen association and dissociation rates of cRBC and peripheral blood RBC were similar at variable oxygen pressure (Figure 4E). Blood group expression of the cRBC was assessed by flow cytometry and compared to the original donor RBC. Blood group expression was in complete agreement with the original donors (Table 1). The functional similarities between cRBC and peripheral blood RBC indicate that the three-stage culture model using culture dishes or a G-Rex bioreactor (Figure 6) yields functional enucleated erythroid cells.

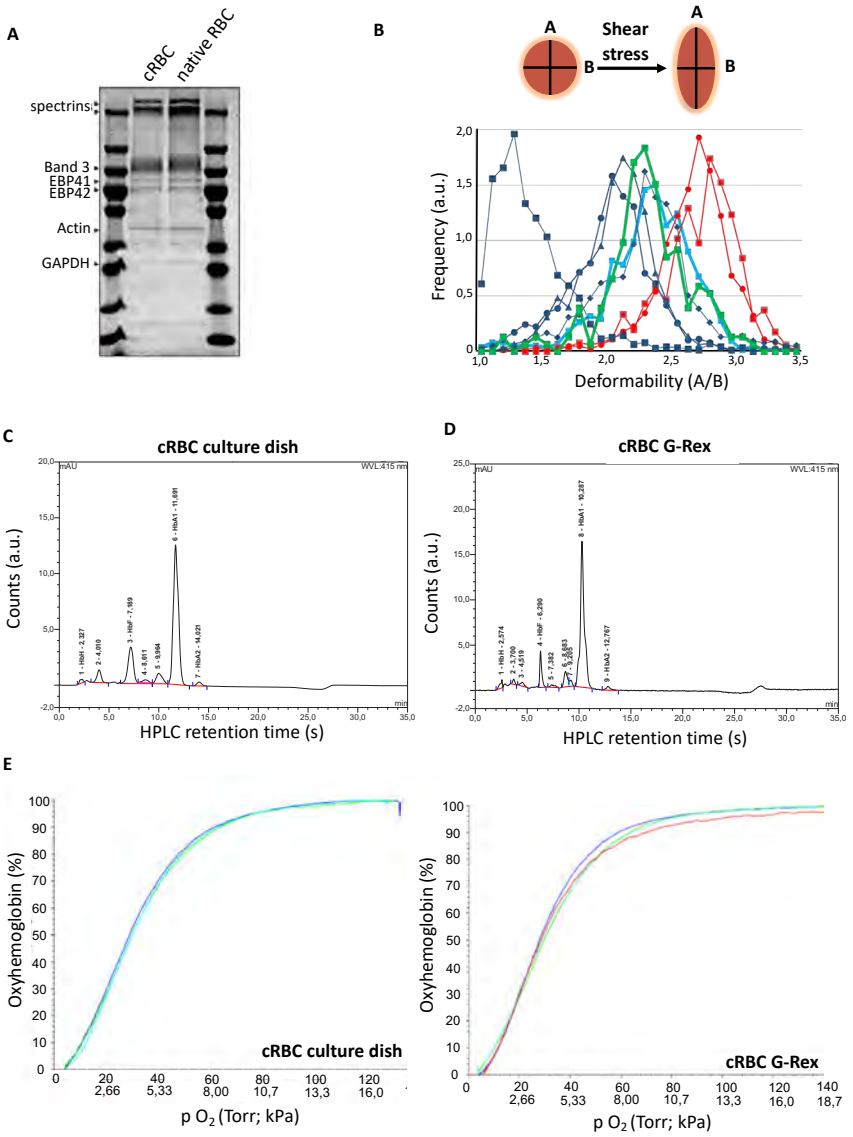


Figure 4. cRBC are highly similar to donor RBC. (A) Ghosts from cRBC and peripheral blood RBC were lysed and subjected to SDS-PAGE. Gel was stained with Coomassie and depicts the most abundant proteins in RBC membranes. **(B)** Deformability of cRBC was measured under shear stress by ARCA, which elongates cells and measures length over width as deformability parameter. Progressively maturating reticulocytes were isolated from peripheral blood (in order of maturation, R1: $RNA^{high}CD71^{high}$, R2: blue squares; $RNA^{high}CD71^{low}$, R3: blue circles; $RNA^{high}CD71^{-}$, blue triangles; R4: $RNA^{low}CD71^{-}$, blue diamonds; as described[24]) were compared with fully mature RBC (red curves) and filtered cRBC from normal culture dishes (thick green line) or the G-Rex bioreactor (thick blue line). Right bar graph represent the quantification of >1000 cells per culture condition. **(C-D)** Expression of hemoglobin variants was determined by HPLC in cRBC in culture dishes before filtering **(C)** or G-Rex after filtering **(D)**. Hemoglobin variants are indicated; exact retention time is indicated for each peak. **(E)** Oxygen association and dissociation curve for peripheral blood RBC (teal and red) and cRBC cultured in a G-Rex bioreactor (blue and green). The percentage oxygenated hemoglobin is measured at a gradient of oxygen tension given in Torr (upper line x-axes) and kPa (lower line x-axes).

Terminal differentiation of erythroblasts to enucleated reticulocytes completely changes the transcriptome

Differentiation of non-hemoglobinized erythroid progenitor cells to functional enucleated cells involves significant changes in morphology, cell volume and protein content. Identification of regulatory processes is crucial to enhance and optimize the production and the yield of *in-vitro* erythropoiesis. A comparison of the *in-vitro* transcriptome to existing databases of *in-vivo* transcriptomes allows to benchmark the *in-vitro* differentiation process. Differentiation was started from $CD117^{+}CD71^{+}CD235^{-/dim}CD44^{high}$ early erythroblast populations (phase III, day 0) from four distinct donors and daily RNA samples were collected until terminal enucleation state at day 12 (95%; Changes in surface marker expression in Figure S2). In reference to the early erythroblast population (day 0) 75% of the analyzed transcripts changed during differentiation (7792 transcripts with a FDR < 0.01 and fold-change > 4) with 5189 down regulated, 2726 upregulated and 124 that were transiently up- or downregulated (Figure 5A and Table S2). These major transcriptome alterations occurred primarily during the first 6 days (phase III). The expression of transcripts encoding proteins crucial for the function and immunological properties of RBC can be scrutinized. Figure 5B shows that RNA expression of different blood group antigen bearing proteins from these donors is differentially regulated over time. In addition, globin subunit expression dynamics indicated rapid hemoglobinization during the first 72 hours of differentiation coinciding with increased CD71 expression. Note that expression of beta globin subunits is significantly higher compared to gamma globins (FS5A). *cKIT* RNA was rapidly downregulated and *CD235* (GPA) rapidly upregulated in agreement with flow cytometry results (Figure S5A). In line with the major transcriptional changes over the course of differentiation, principal component analysis (PCA) captured ~57% of variance in the 2000 most variable genes. The first component associated with differentiation progression, sequentially grouping samples from subsequent days (Figure 5C). Cultured RBC showed good functional and morphological correspondence with *in-vivo* RBC (Figure 4). This raised the question how the cultured cells would compare to *ex-vivo* cells at the transcriptome level. Comparing published transcriptomes of bone marrow megakaryoid/erythroid progenitors ($CD38^{+}CD34^{+}CD10^{-}CD45RA^{-}CD123^{-}CD90^{-}$) and $CD71^{+}CD235^{+}$ erythroblasts[34] to *in-vitro* cultured erythroid cells revealed that the *ex-vivo* MEP grouped before the sequence of cultured cells, while the $CD71^{+}CD235^{+}$ grouped along with the cultured cells, with the same marker expression (Figure S4B). Note that the *in-vitro* cultures showed little variation between donors indicating a reproducible synchronized differentiation process.

Ex-vivo cells showed relatively more variation in the PCA, which may be a consequence of the more broadly expressed erythroid surface markers used to isolate these cells. The major differences between cultured and *ex-vivo* cells accounted for 18% of the variance (Figure S5A). Still, Pearson correlation between samples using all differentially

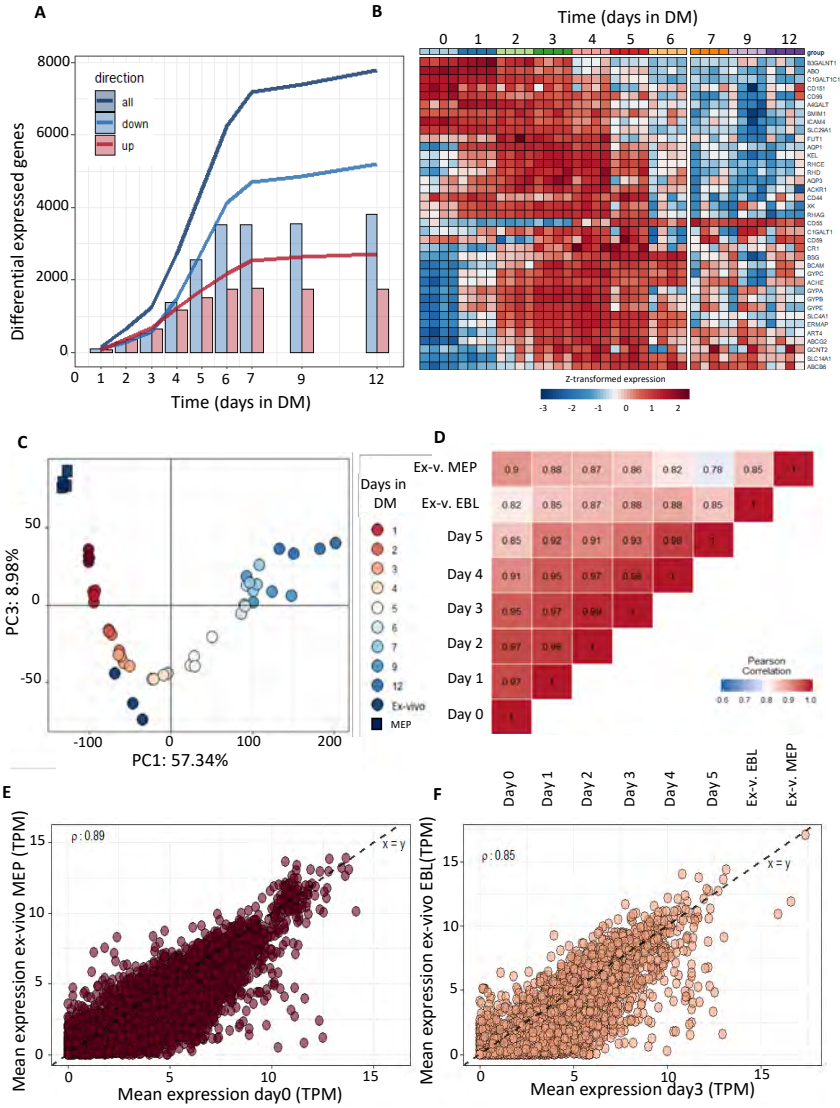


Figure 5. RNA expression profiles during *in-vitro* erythroid differentiation are similar to *ex-vivo* isolated erythroid precursors. RNA was isolated at subsequent days of differentiation to identify changes in the transcriptome (4 independent donors) **(A)** Bars represent the number of transcripts that are up- (red) or downregulated (blue) each day in reference to the start of the culture (FDR < 0.01 and log fold change > 2, or < -2). Lines reflect cumulative number of unique genes differentially expressed over the time course. **(B)** Heatmap of z-transformed expression values (log2-CPM) for all genes encoding for blood group antigens and or blood group bearing moieties. **(C)** The transcriptome of MEP (black squares) and CD71^{high}CD235^{high} erythroblasts (black circles) isolated from bone marrow was compared to the transcriptome of differentiating cRBC using principle component analysis (PC1 versus PC3). **(D)** A Pearson correlation matrix of all samples used in panel A-C. **(E-F)** Mean transcript levels (transcripts per million mapped reads) were compared between MEP and cRBC day0 **(E)**; and between freshly isolated erythroblasts and cRBC day3 **(F)**.

expressed genes indicated that *ex-vivo* MEP were most similar to erythroblasts cultured under expansion conditions (d0) whereas *ex-vivo* CD235+CD71+ erythroblasts were most similar to d3 differentiated polychromatic erythroblasts ($\rho = 0.89$ and $\rho = 0.85$ respectively, Figure 5D). Direct comparison of RNA expression between the *ex-vivo* MEP and d0 erythroid cells, and of the CD235+CD71+ *ex-vivo* erythroblast and d3 differentiated cells showed comparable transcript levels (Figure 5E-F). Overall, the main difference between cultured and freshly isolated cells originates from genes that were predominantly expressed at increased levels in cultured cells compared to *ex-vivo* cells (Table S3). Of note, transcripts with increased expression in cultured cells were further downregulated upon differentiation progression in line with decreased expression observed in the more asynchronous *ex-vivo* erythroid cells (Figure S4B). Removing the low expression filter also revealed a set of 380 genes expressed at lower levels in cultured cells, that consisted of pseudogenes and transcripts encoding mitochondrial or ribosomal proteins (Table S4), probably reflecting a difference in technical processing of samples. The similarities in the transcriptome of cultured cells and the related stage in-vivo indicates that close recapitulation of transcriptional changes is at the basis of the functional and morphological characteristics of the cultured erythroid cells.

Discussion

Widescale clinical application of cRBC is faced by several constraints, like the inability to generate large cell numbers that are required, the high costs of ill-defined media and the low yield of enucleated, biconcave cRBC[35]. The culture protocol presented here challenges these constraints by boosting advances with respect to high enucleation rates, matching erythroid characteristic at different levels, redefining medium composition and maximum expansion without the necessity to first isolate $CD34^+$ cells (Figure 6). Using a defined medium and exploiting stress erythropoiesis we achieve 10^7-fold erythroblasts expansion within 26 days. Despite starting from total PBMC, pure erythroid cultures expressing CD71 and $CD235a^{dim}$ are obtained validating the GMP-grade medium yielding similar cell numbers as previously reported using commercial Stemspan media[17,18]. Using total PBMC not only allows outgrowth of all $CD34^+$ progenitors, but also $CD34^-$ hematopoietic progenitors present in blood that have the capacity to differentiate towards erythroid cells[17,18]. Glucocorticoids are essential for stress erythropoiesis in the mouse and synergy of glucocorticoids with EPO and SCF induces erythroblasts proliferation whilst inhibiting differentiation[7,23,32,36,37]. The addition of glucocorticoids also supports erythropoiesis by differentiating peripheral blood monocytes or $CD34^+$ cells to erythroid supporting macrophages during culture from PBMC, further increasing the erythroid yield[19,38]. In contrast to our plasma/serum-free expansion, many large-scale red cell culture protocols

use glucocorticoid agonists in combination with serum and/or plasma during expansion. Here, we showed that the addition of plasma causes premature differentiation of erythroblasts also in the presence of glucocorticoids[9,10,13,39,40]. Increased spontaneous differentiation upon addition of plasma during erythroblast expansion may be due to additional growth factors or other plasma components. Optimal expansion in absence of plasma/serum is important to reach the amount of cRBC required for transfusion, but also to establish erythroid cultures from small blood aliquots of specific anemic patients. We have recently shown that 3ml of blood is sufficient to culture enough cells to reprogram erythroblasts to induced pluripotent stem cells (IPSC)[41]. In addition to direct use, the expanded erythroblast can be frozen and defrosted without loss of expansion potential (data not shown), similar to what we previously reported for starting cultures from frozen PBMCs[42]. This introduces flexibility concerning actual production of products, for example with the generation of IPSC form cryo-preserved patient material[43-45]. The use of adult PBMC also facilitates the availability of starting material and introduces the possibility to culture autologous cRBC. This is important considering alloimmunization caused by blood group mismatches and matching from cord blood derived RBC may be complicated in either conventional transfusion or with novel therapeutic blood products. The culture process was compatible with up-scaling for clinical studies and applications using a G-Rex bioreactor. The G-Rex bioreactor allowed for $3x10^6$ fold expansion per PBMC (~$1x10^9$ fold from CD34+ cells in these PBMC; taken the accepted 0,16% of CD34+ cells present in total PBMCs [heideveld, 2015]). Of note, we have submitted cells that have been expanded for 10 and the maximum of 25 days to differentiation (DM) and found that enucleation and differentiation progression is unperturbed (data not shown). Remaining nuclei and enucleated cells after differentiation could be efficiently removed using a leukoreduction filter, as generally employed by bloodbanks, to obtain a pure cRBC population. The total costs of one 5 liter GREX producing around 4.5ml GMP-grade cRBC from 100 million PBMC (roughly 80ml of blood) is approximately 28000 euros and would take 27 days (15 days of EM and 12 days of DM). Of note, it would take roughly 36 days (24 days EM and 12 days DM) to culture a transfusion unit starting from 100 million PBMCs. Two of the major challenges, besides bioreactor development, that need to be tackled are i) reducing prices of major cost drivers and ii) increasing the efficiency of filters as currently approximately 30% of cRBC can be recovered from the leucocyte filters. Costs can be significantly reduced if specific major cost drivers were to be produced in a bulk recombinant manner (e.g. growth factors). For instance, the recombinant GMP-grade SCF, one of the major cost drivers in our protocol, presently is around 3500 times more expensive compared to insulin. Exchange of human albumin by novel agents like polyvinyl alcohol that have been shown to promote HS(P)C expansion could potentially cut costs significantly[49]. Also, novel cell permeable iron chelators that could be used as replacement for transferrin may cut costs. It has been estimated that around 5-10ml of

erythrocytes will be needed for specific therapeutic delivery of cargo. Thus, currently the costs to produce this amount, taking a filter efficiency of 30%, would be between 80.000-200.000 euros. Considering expensive enzyme replacement therapies, the costs of *in vitro* cRBC may be competitively priced. However, it is clear that additional optimization and cost-reduction is needed as well as research into loading and stability of therapeutics within the *in vitro* cultured cells.

The defined IMDM-based culture medium, termed Cellquin, solely contains GMP grade components and finds its basis in HEMA-def[7]. Knowing the exact concentrations of all components within Cellquin now allows to quantitatively track erythroid requirements by combining the transcriptome/proteome with metabolomics. This may help to culture cells at higher densities and to cater specific media components exclusively erythroid need, leading to considerable cost-reduction and aiding upscaling. currently, the costs of Cellquin is lower compared to commercially available media, while performing at least similar.

One Cellquin component paramount to its effectiveness is albumin. We observed that the isolation and manufacturing process of human albumin critically influences the erythroid expansion potential. Using ultra-clean, detoxified, or recombinant additive-free HA significantly increased the erythroid expansion potential. Albumin binds substances including proteins, metabolites and fatty acids, including toxins, drugs and other therapeutics[46-48]. Replacing ultra-pure cHA in EM by Albuman® reduced erythroblast expansion potential, which could be reverted by charcoal and ion exchanger treatment of Albuman®. Interestingly, the process to manufacture Albuman® includes a saturation step to restrain the albumin binding potential, rendering it mostly inert. This suggests that the binding and/or transport function of albumin is important to ensure continued erythroblast expansion.

The technical improvements to the culture protocol result in excellent yield of *in-vitro* cultured erythroblasts combined with >90% enucleation, adult hemoglobin expression, correct blood group expression, deformability and oxygen saturation dynamics similar to donor peripheral blood enucleated cells. In addition, we present the first transcriptomic analysis from erythropoiesis originating from stress-erythropoiesis cultures. Comparison of this dataset to datasets from *ex-vivo* erythroid cells shows that, next to similarities in RBC characteristics, the cRBC are comparable to similar *ex-vivo* cells at transcript level. These observations together, make the transcriptome dataset provided here a valuable resource to address erythroid regulatory mechanisms.

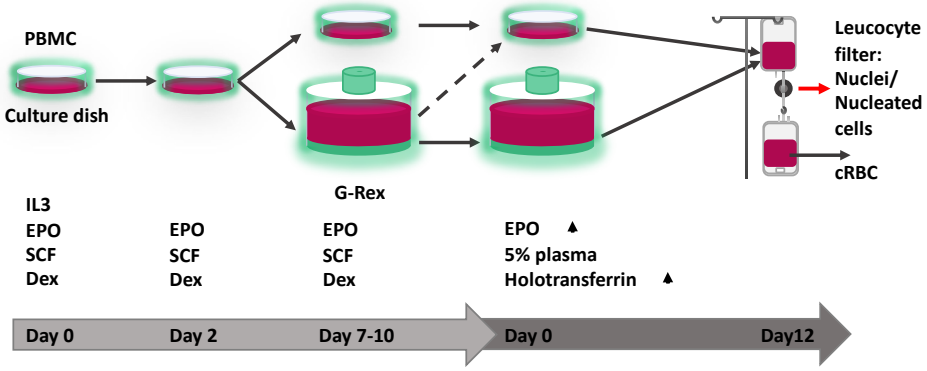


Figure 6. Progression of erythroid cells during culture. Overview of the three-phase erythroid culture system using culture dishes or a G-Rex bioreactor. A pure erythroblast culture is established in dishes from PBMC during the first 7 days. From day 7 erythroblasts are expanded in G-Rex or in dishes. When transferred from expansion to differentiation medium, erythroblasts in dishes or in G-Rex can mature to enucleated cells. The small remnant of nucleated cells and nuclei (pyrenocytes) can be removed by passage through a leucocyte filter. The growth factors and hormones used in culture are indicated at the lower half of the graph.

In 2011, Timmins et al. demonstrated an ultra-high yield of cRBC with >90% enucleation from cord blood-derived CD34+ cells in the absence of plasma[15]. This reported erythroid yield was similar to our serum/plasma-free adult PBMC-derived erythroid expansion. However, >90% enucleation during terminal differentiation using our differentiation protocol (stage III) was only recapitulated in presence of plasma. Where, adding 5% Omniplasma increases enucleation from 20-25% to more than 90% (data not shown). Whether this reflects a difference in cord blood versus adult erythroid cultures remains unknown but important to investigate along with identifying the components in plasma that are key to this increased enucleation in our system.

Combined the advances of the presented protocol facilitate both easy implementation in other laboratories and study of erythropoiesis in healthy individuals or in patients of which limited sample volumes are available. It adds to the feasibility of using adult peripheral blood as starting material for cRBC cultures, which is an important step towards precision medicine. For example, in using custom engineered cRBC as cargo vesicles for drug delivery. In addition, the synchronized cultures contain young reticulocytes that have a theoretical lifespan of about 120 days. Chronic anemia patients that receive blood transfusions every two months may benefit from transfusions with *in-vitro* cultured long-lived RBC, potentially increasing the time between transfusions and thereby reducing the costs. Currently we are working towards a clinical trial that will allow to test the *in-vivo* lifespan of transfused PBMC-derived cRBC.

Author Contributions

PB, MT-V, ES, EH, SH, EO, EV and AV performed the experiments. EA, EH, and PB designed the experiments, analyzed the data and wrote the manuscript. JM performed the RNA isolation and sequencing. ML contributed to the experiment design and writing of the manuscript. The manuscript was critically revised by all authors. The authors declare no competing financial interests.

Acknowledgments

We are grateful to Wilson Wolf Manufacturing (Saint Paul, MN, USA) for providing the G-Rex bioreactors. We would like to thank Rob van Zwieten, Martijn Veldthuis and Jeffrey Berghuis (Sanquin, dept. Blood Cell Research) for the technical assistance and data acquisition regarding the HPLC data and deformability assays, and the Central Facility of Sanquin for their assistance regarding flow cytometry analysis. This work was supported by grants from The Netherlands Organization for Health Research and Development (ZonMw-TOP, 40-00812-98-12128; SH and ZONMW-TAS, 40-41400-98-1327; PB, MT-V, ES), from the Landsteiner Foundation for Blood Transfusion Research (LSBR:1141; EA, EH),; and by Sanquin (PPOR:15-30; AV).

Supplementary material

Additional figures can be found in supplementary information. RNA-sequencing data has been submitted to NCBI's Gene Expression Omnibus and is available under accession number: GSE124363.

References

1. Carson JL, Grossman BJ, Kleinman S, et al. Red blood cell transfusion: a clinical practice guideline from the AABB*. Ann Intern Med. 2012;157(1):49-58.
2. Palmer AF, Intaglietta M. Blood substitutes. Annu Rev Biomed Eng. 2014;16:77-101.
3. Ghinassi B, Ferro L, Masiello F, et al. Recovery and Biodistribution of Ex Vivo Expanded Human Erythroblasts Injected into NOD/SCID/IL2Rgamma mice. Stem Cells Int;2011:673752.
4. Shah S, Huang X, Cheng L. Concise review: stem cell-based approaches to red blood cell production for transfusion. Stem Cells Transl Med. 2014;3(3):346-355.
5. Kim HO. In-vitro stem cell derived red blood cells for transfusion: are we there yet? Yonsei Med J. 2014;55(2):304-309.
6. Larochelle A. Generation of red blood cells in vitro: monitoring the process for improved efficiency. Cytotherapy. 2013;15(9):1043-1045.
7. Migliaccio G, Sanchez M, Masiello F, et al. Humanized culture medium for clinical expansion of human erythroblasts. Cell Transplant. 2010;19(4):453-469.
8. Singh VK, Saini A, Tsuji K, Sharma PB, Chandra R. Manufacturing blood ex vivo: a futuristic approach to deal with the supply and safety concerns. Front Cell Dev Biol. 2014;2:26.
9. Giarratana MC, Rouard H, Dumont A, et al. Proof of principle for transfusion of in vitro-generated red blood cells. Blood. 2011;118(19):5071-5079.
10. Kupzig S, Parsons SF, Curnow E, Anstee DJ, Blair A. Superior survival of ex vivo cultured human reticulocytes following transfusion into mice. Haematologica. 2017;102(3):476-483.
11. Pishesha N, Bilate AM, Wibowo MC, et al. Engineered erythrocytes covalently linked to antigenic peptides can protect against autoimmune disease. Proc Natl Acad Sci U S A. 2017;114(12):3157-3162.
12. Stowell SR, Girard-Pierce KR, Smith NH, et al. Transfusion of murine red blood cells expressing the human KEL glycoprotein induces clinically significant alloantibodies. Transfusion. 2014;54(1):179-189.
13. Giarratana MC, Kobari L, Lapillonne H, et al. Ex vivo generation of fully mature human red blood cells from hematopoietic stem cells. Nat Biotechnol. 2005;23(1):69-74.
14. Miharada K, Hiroyama T, Sudo K, Nagasawa T, Nakamura Y. Efficient enucleation of erythroblasts differentiated in vitro from hematopoietic stem and progenitor cells. Nat Biotechnol. 2006;24(10):1255-1256.
15. Timmins NE, Athanasas S, Gunther M, Buntine P, Nielsen LK. Ultra-high-yield manufacture of red blood cells from hematopoietic stem cells. Tissue Eng Part C Methods. 2011;17(11):1131-1137.
16. Kleihauer E, Braun H, Betke K. [Demonstration of fetal hemoglobin in erythrocytes of a blood smear]. Klin Wochenschr. 1957;35(12):637-638.
17. van den Akker E, Satchwell TJ, Pellegrin S, Daniels G, Toye AM. The majority of the in vitro erythroid expansion potential resides in CD34- cells, outweighing the contribution of CD34+ cells and significantly increasing the erythroblast yield from peripheral blood samples. Haematologica. 2010.
18. Heideveld E, Masiello F, Marra M, et al. CD14+ cells from peripheral blood positively regulate hematopoietic stem and progenitor cell survival resulting in increased erythroid yield. Haematologica. 2015;100(11):1396-1406.
19. Heideveld E, Hampton-O'Neil LA, Cross SJ, et al. Glucocorticoids induce differentiation of monocytes towards macrophages that share functional and phenotypical aspects with erythroblastic island macrophages. Haematologica. 2018;103(3):395-405.

20. Boehm D, Murphy WG, Al-Rubeai M. The potential of human peripheral blood derived CD34+ cells for ex vivo red blood cell production. J Biotechnol. 2009;144(2):127-134.
21. Griffiths RE, Kupzig S, Cogan N, et al. Maturing reticulocytes internalize plasma membrane in glycophorin A-containing vesicles that fuse with autophagosomes before exocytosis. Blood. 2012;119(26):6296-6306.
22. von Lindern M, Zauner W, Mellitzer G, et al. The glucocorticoid receptor cooperates with the erythropoietin receptor and c-Kit to enhance and sustain proliferation of erythroid progenitors in vitro. Blood. 1999;94(2):550-559.
23. Leberbauer C, Boulme F, Unfried G, Huber J, Beug H, Mullner EW. Different steroids co-regulate long-term expansion versus terminal differentiation in primary human erythroid progenitors. Blood. 2005;105(1):85-94.
24. Ovchynnikova E, Aglialoro F, Bentlage AEH, et al. DARC extracellular domain remodeling in maturating reticulocytes explains Plasmodium vivax tropism. Blood. 2017;130(12):1441-1444.
25. van Zwieten R, van Oirschot BA, Veldthuis M, et al. Partial pyruvate kinase deficiency aggravates the phenotypic expression of band 3 deficiency in a family with hereditary spherocytosis. Am J Hematol. 2015;90(3):E35-39.
26. van den Akker E, Satchwell TJ, Pellegrin S, et al. Investigating the key membrane protein changes during in vitro erythropoiesis of protein 4.2 (-) cells (mutations Chartres 1 and 2). Haematologica. 2010.
27. McCarthy KP, Fletcher RJ, Jr., Rota CT, Hutto RL. Predicting species distributions from samples collected along roadsides. Conserv Biol. 2012;26(1):68-77.
28. Socolovsky M, Nam H, Fleming MD, Haase VH, Brugnara C, Lodish HF. Ineffective erythropoiesis in Stat5a(-/-)5b(-/-) mice due to decreased survival of early erythroblasts. Blood. 2001;98(12):3261-3273.
29. Koulnis M, Pop R, Porpiglia E, Shearstone JR, Hidalgo D, Socolovsky M. Identification and analysis of mouse erythroid progenitors using the CD71/TER119 flow-cytometric assay. J Vis Exp. 2011(54).
30. Vera JF, Brenner LJ, Gerdemann U, et al. Accelerated production of antigen-specific T cells for preclinical and clinical applications using gas-permeable rapid expansion cultureware (G-Rex). J Immunother. 2010;33(3):305-315.
31. Bajgain P, Mucharla R, Wilson J, et al. Optimizing the production of suspension cells using the G-Rex "M" series. Mol Ther Methods Clin Dev. 2014;1:14015.
32. von Lindern M, Deiner EM, Dolznig H, 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. 2001;20(28):3651-3664.
33. Waugh RE, Mantalaris A, Bauserman RG, Hwang WC, Wu JH. Membrane instability in late-stage erythropoiesis. Blood. 2001;97(6):1869-1875.
34. Corces MR, Buenrostro JD, Wu B, et al. Lineage-specific and single-cell chromatin accessibility charts human hematopoiesis and leukemia evolution. Nat Genet. 2016;48(10):1193-1203.
35. Engert A, Balduini C, Brand A, et al. The European Hematology Association Roadmap for European Hematology Research: a consensus document. Haematologica. 2016;101(2):115-208.
36. Migliaccio G, Di Pietro R, di Giacomo V, et al. In vitro mass production of human erythroid cells from the blood of normal donors and of thalassemic patients. Blood Cells Mol Dis. 2002;28(2):169-180.
37. Bauer A, Tronche F, Wessely O, et al. The glucocorticoid receptor is required for stress erythropoiesis. Genes Dev. 1999;13(22):2996-3002.
38. Falchi M, Varricchio L, Martelli F, et al. Dexamethasone targeted directly to macrophages induces macrophage niches that promote erythroid expansion. Haematologica. 2015;100(2):178-187.

39. Neildez-Nguyen TM, Wajcman H, Marden MC, et al. Human erythroid cells produced ex vivo at large scale differentiate into red blood cells in vivo. Nat Biotechnol. 2002;20(5):467-472.
40. Zhang S, Macias-Garcia A, Velazquez J, Paltrinieri E, Kaufman RJ, Chen JJ. HRI coordinates translation by eIF2alphaP and mTORC1 to mitigate ineffective erythropoiesis in mice during iron deficiency. Blood. 2018;131(4):450-461.
41. Varga E, Hansen M, Wust T, von Lindern M, van den Akker E. Generation of human erythroblast-derived iPSC line using episomal reprogramming system. Stem Cell Res. 2017;25:30-33.
42. Masiello F, Tirelli V, Sanchez M, et al. Mononuclear cells from a rare blood donor, after freezing under good manufacturing practice conditions, generate red blood cells that recapitulate the rare blood phenotype. Transfusion. 2014;54(4):1059-1070.
43. Hansen M, Varga E, Aarts C, et al. Efficient production of erythroid, megakaryocytic and myeloid cells, using single cell-derived iPSC colony differentiation. Stem Cell Res. 2018;29:232-244.
44. Hansen M, Varga E, Wust T, et al. Generation and characterization of a human iPSC line SANi005-A containing the gray platelet associated heterozygous mutation p.Q287* in GFI1B. Stem Cell Res. 2017;25:34-37.
45. Hansen M, Varga E, Wust T, et al. Generation and characterization of human iPSC line MML-6838-Cl2 from mobilized peripheral blood derived megakaryoblasts. Stem Cell Res. 2017;18:26-28.
46. Klammt S, Brinkmann B, Mitzner S, et al. Albumin binding capacity (ABiC) is reduced in commercially available human serum albumin preparations with stabilizers. Z Gastroenterol. 2001;39 Suppl 2:24-27.
47. Klammt S, Mitzner SR, Stange J, et al. Improvement of impaired albumin binding capacity in acute-on-chronic liver failure by albumin dialysis. Liver Transpl. 2008;14(9):1333-1339.
48. Kragh-Hansen U, Chuang VT, Otagiri M. Practical aspects of the ligand-binding and enzymatic properties of human serum albumin. Biol Pharm Bull. 2002;25(6):695-704.
49. Wilkinson AC, Ishida R, Kikuchi M, Sudo K, Morita M, Crisostomo RV, Yamamoto R, Loh KM, Nakamura Y, Watanabe M, Nakauchi H, Yamazaki S. Long-term ex vivo haematopoietic-stem-cell expansion allows nonconditioned transplantation. Nature. 2019 Jul;571(7763):117-121.

Supplementary methods and material

Cell culture using different sources of human albumin

EM medium was supplemented with human bovine albumin (HSA) from different sources: Albuman® (Sanquin, Amsterdam, The Netherlands), detoxified HSA[1], recombinant HSA (Akron Biotech, Boca Raton, FL, USA), or AB+ pool plasma (Sanquin) at indicated concentrations as specified in the figure legends. In addition, the commercial medium Stemspan (Stem Cell Technologies, Germany) was used [2].

RNA-sequening

Erythroblasts were allowed to differentiate to enucleated reticulocytes as described in culture protocol and summarized in figure 6. Cells were collected daily; RNA was isolated using Trizol (Invitrogen Technologies, Carlsbad, CA) and amplified and rRNA depleted using HyperPrep Kit with RiboErase (KAPA Biosystems, Pleasanton, CA, USA) as described by manufactures.

Samples were sequenced at 30 million 75bp paired-end reads. After quality control with FastQC, read were mapped to genome ChGR38.v85 using STAR [3]. Lowly expressed genes were filtered from count data, after correction for library size (counts per million mapped reads < 3). Subsequent differential expression analysis was performed with EdgeR using a paired design for individual donors and quasi-likelihood F-test to test expression at each day compared to day 0 [4]. The differential expression threshold was set over 4-fold difference and an FDR < 0.01.

The comparison of cRBCs to *ex-vivo* cells was performed on samples aligned to ChGR38.v85 with Salmon [5]. This aligner allows for direct mapping to transcript isoforms, including these in the comparison while facilitating normalization to transcript length prior to plotting. Counts were normalized for transcript length and library size, as transcript per million mapped reads (TPM) and were plotted using ggplot2 package in R [6].

Culture filtration to obtain reticulocytes

CRBC were purified using a 60ml neonatal leucoreduction filter (Haemonetic BV, Breda, The Netherlands). The filter was primed with IMDM containing 4% Omniplasma. Cultured cells were filtered followed by a 2 times wash with 25ml priming medium. CRBC were resuspended in IMDM containing 4% Omniplasma. Cells were centrifuged at 1600rpm for 5 min in order to obtain packed red blood cells.

Flow cytometry

Antibodies used: Acris (Herford, Germany): anti-CD235a (FITC 1:400; PE 1:400); BD Biosciences: anti-CD117 (PE 1:50), anti-CD235a (PE 1:150); eBioscience (Vienna, Austria): anti-CD44 (APC 1:150); Miltenyi Biotec: anti-CD71 (APC 1:100; VioBlue 1:200), anti-

CD235a (VioBlue 1:200); and Life technologies (Carlsbad, CA, USA): anti-CD117 (FITC 1:100). The following antibodies were used for blood group analysis: IBGRL research products (Bristol, UK): anti-A (Bric145; 1:1000), anti-B (BGRL2; 1:500), anti-M (clone 6A7; 1:5000), anti-N (clone BRIC157; 1:1000); Sanquin Products (Amsterdam, The Netherlands): anti-RhC (clone MS24; RBC 1:1000, cRBC undiluted), anti-Rhc (clone MS35; 1:200), anti-RhE (clone MS260; RBC 1:80, cRBC undiluted), anti-Rhe (clone MS21/63; RBC 1:100, cRBC undiluted), P1 (clone P3NIL100; 1:500), Kp-a (undiluted), kp-b (undiluted), Jk-a (MS15; undiluted), Jk-b (MS8; undiluted), Fy-a (P3TIM; undiluted), Fy-b (undiluted), Lu-a (undiluted), Lu-b (undiluted); Merck-Millipore (Darmstadt, Germany): anti-K (MS59; 1:50), anti-k (P3A118OL67; 1:100), anti-s (P3Y326Bn5; 1:400). Ortho Clinical Diagnostics (Buckinghamshire, UK): anti-S (MNS3; undiluted). Anti-A, B, N, M were detected with an anti-mouse IgG FITC-labeled antibody (Life Technologies; 1:200) was used and anti-C, c, E, e, Jk-a, Jk-b, and P1 with an anti-human IgM PE-labeled antibody (Southern Biotech, Birmingham, AL; 1:500). The remaining blood groups were detected with an anti-human IgG PE-labeled antibody (Southern Biotech; 1:200).

Supplementary references

1. Beug H, Schroeder C, Wessely O, et al. Transformation of erythroid progenitors by viral and cellular tyrosine kinases. Cell Growth Differ. 1995;6(8):999-1008.
2. Heideveld E, Masiello F, Marra M, et al. CD14+ cells from peripheral blood positively regulate hematopoietic stem and progenitor cell survival resulting in increased erythroid yield. Haematologica. 2015;100(11):1396-1406.
3. Dobin A, Davis CA, Schlesinger F, et al. STAR: ultrafast universal RNA-seq aligner. Bioinformatics. 2013;29(1):15-21.
4. McCarthy KP, Fletcher RJ, Jr., Rota CT, Hutto RL. Predicting species distributions from samples collected along roadsides. Conserv Biol. 2012;26(1):68-77.
5. Patro R, Duggal G, Love MI, Irizarry RA, Kingsford C. Salmon provides fast and bias-aware quantification of transcript expression. Nat Methods. 2017;14(4):417-419.
6. H W. ggplot2: Elegant Graphics for Data Analysis. New York: Springer-Verlag; 2016.

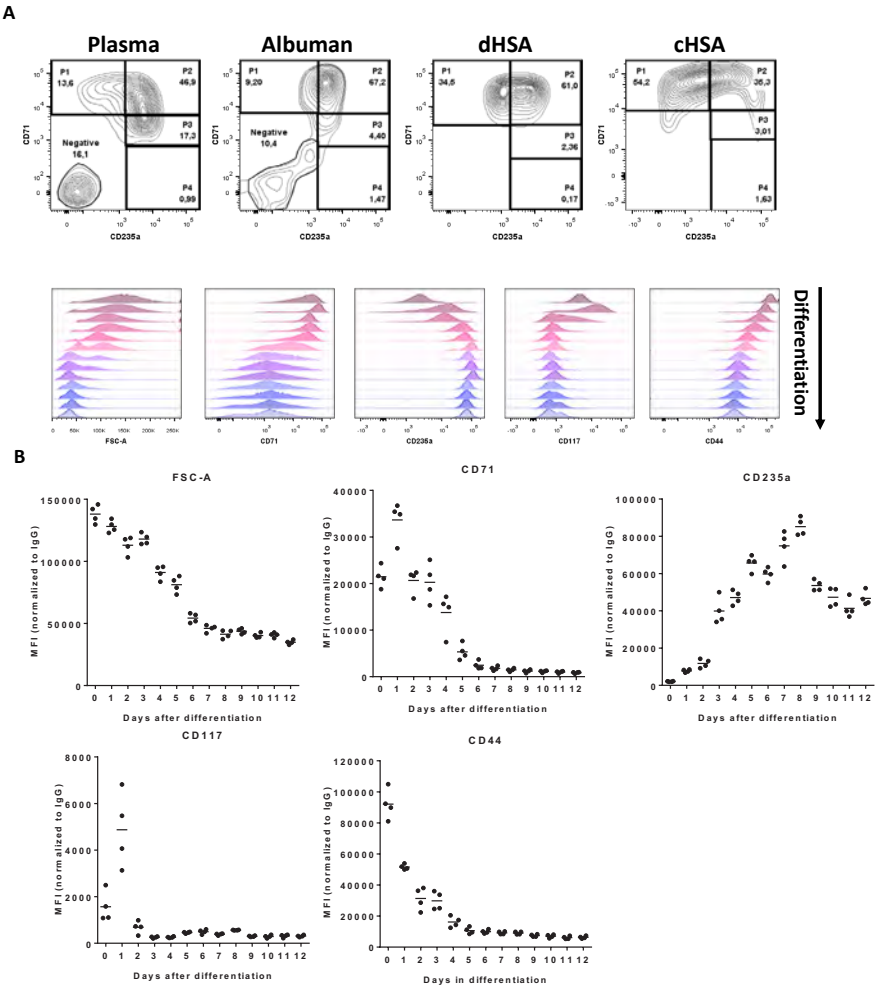


Figure S1. The presence of plasma/serum affects erythroblast expansion. (A) Representative dot plots of erythroblasts at day 12 of culture in EM in the presence of ultra-clean HSA (cHSA) or EM in which cHSA was replaced by plasma, Albuman® or detoxified HSA (dHSA) corresponding to Figure 1A (n=3-4). Note that the percentage of immature erythroblasts (P1) is 3-4 times higher in cHSA and dHSA compared to plasma and Albuman®. This suggests that cHSA and dHSA prolong the immature state of erythroid cultures. **(B)** Erythroblasts cultured in EM for 18 days in the presence of cHSA, dHSA or recombinant HSA (rHSA) showed a similar expansion potential. Mean ± SD (two-way ANOVA; n=2). **(C)** Erythroblasts at day 7 and 19 of expansion in Stemspan, as described previously[2].

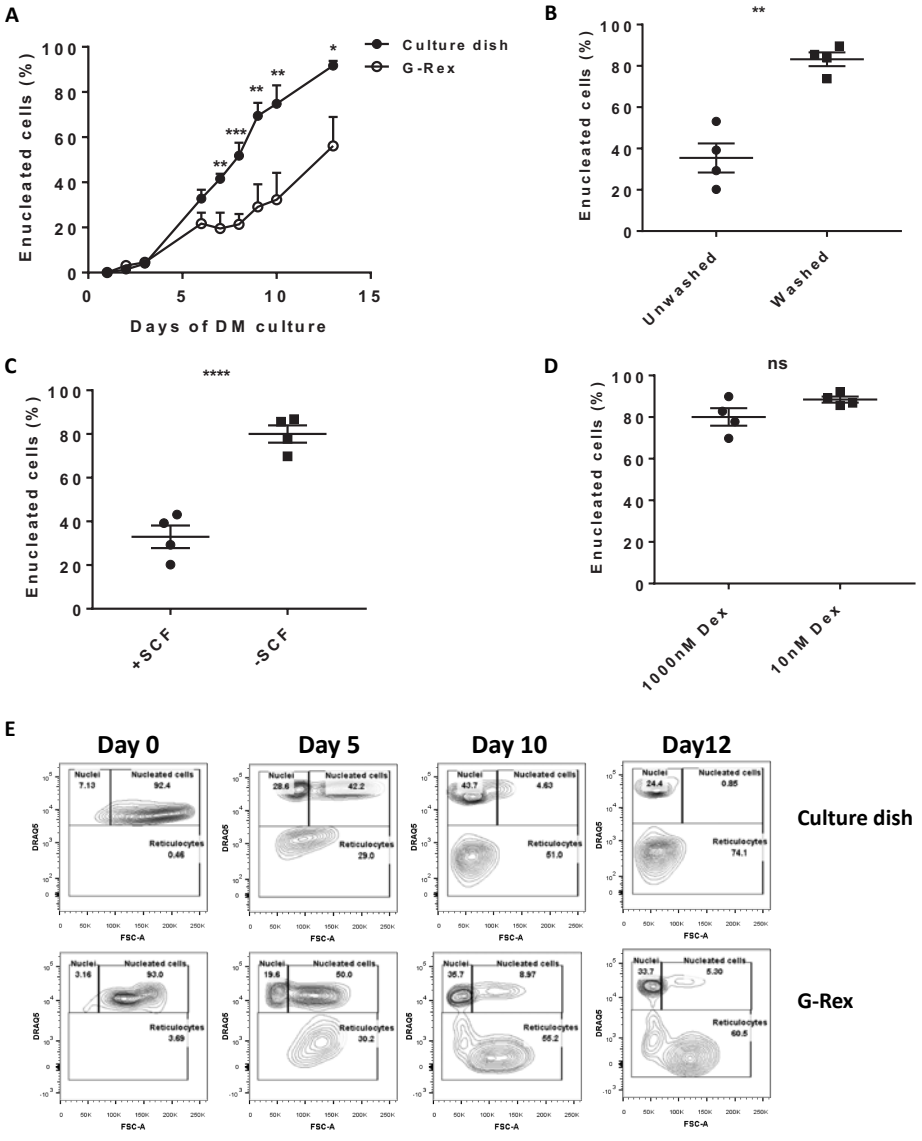


Figure S2. Marker expression during erythroblast differentiation in culture dishes. (A) Erythroblasts were cultured in EM for 8 days and subsequently differentiated for 12 days in DM. Representative overlaying histograms showing the FSC-A during differentiation and the expression levels of CD71, CD235a, CD117 and CD44 analyzed by flow cytometry. **(B)** Graphs corresponding to panel A showing the mean fluorescence intensity (MFI; n=4). **(C)** Lines depicts RNA expression over days of differentiation (log2 transformed CPM) for differentiation markers CD71 (transferrin receptor, TFRC) and CD235 (glycophorin A, GYPA) along with c-kit(SCF receptor, KIT) adult globin subunits (HBB and HBA2) and fetal globin subunit (HBG2).

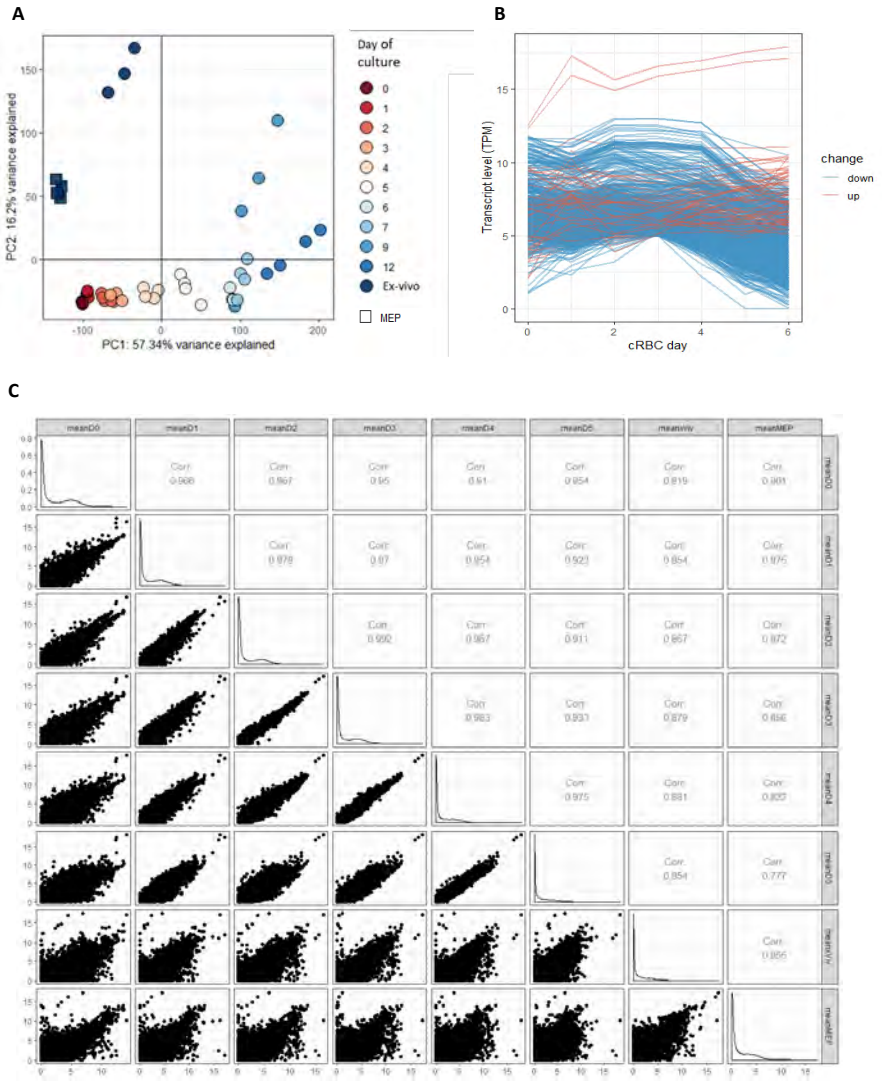


Figure S3. Enucleation of cRBC derived from culture dishes or a G-Rex bioreactor. **(A)** Graph showing the enucleation of erythroid cells during differentiation after supplementation into DM to initiate differentiation in a G-Rex bioreactor, compared to culture dishes corresponding to Figure 2E. Mean ± SD (paired *t*-test, **P*<0.05, ***P*<0.01, ****P*<0.001; n=4). **(B)** Erythroblasts expanded and differentiated in culture dishes with 90% (unwashed) or 100% (washed) replacement of expansion medium before differentiation medium. Mean ± SEM (paired *t*-test, ***P*<0.01; n=4). **(C)** Two days prior to differentiation in DM, SCF was either removed from or maintained in EM and enucleation was analyzed after 12 days of differentiation. Mean ± SEM (paired *t*-test, *****P*<0.0001; n=4). **(D)** Two days prior to differentiation 1000nM or 10nM dexamethasone was added to the expansion medium. Expansion medium was replaced with differentiation medium and cells were analyzed for enucleation after 12 days. Mean ± SEM (paired *t*-test; n=4). **(E)** Representative dot plots of erythroblasts during differentiation in DM in culture dishes or G-Rex based on the expression of DRAQ5.

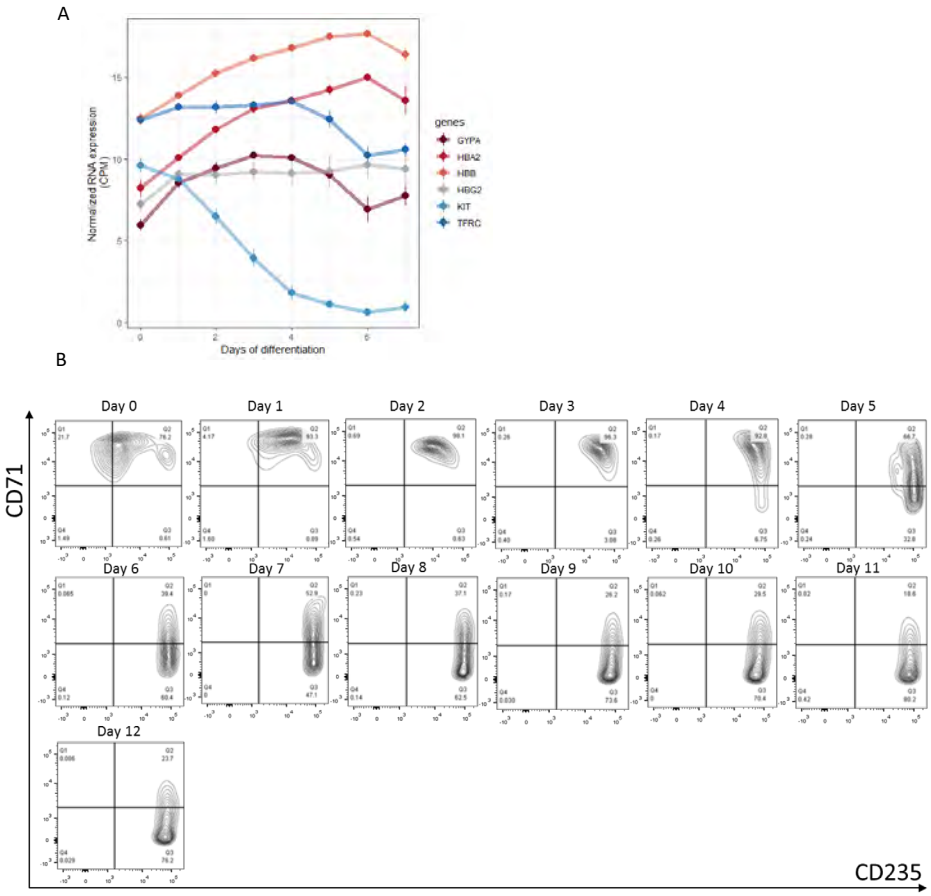
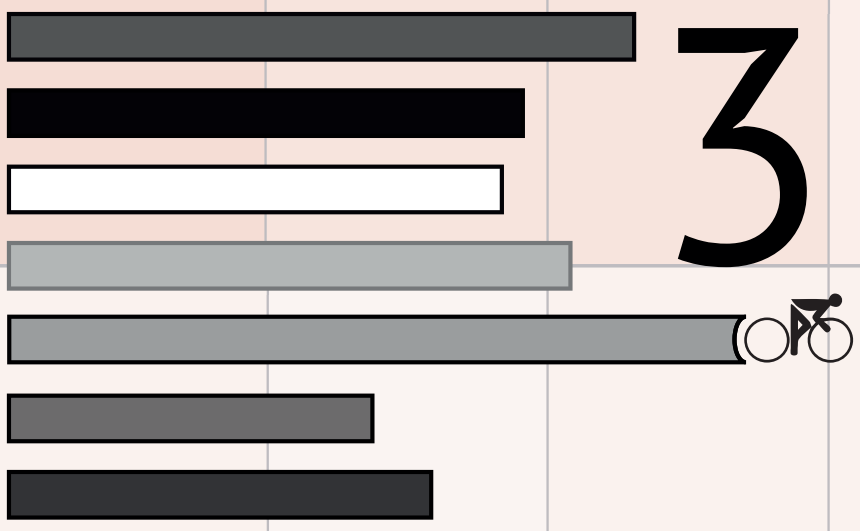


Figure S4. Variation between cRBC and *ex-vivo* cells. (A) Plot of cRBC and *ex-vivo* cells on first and second component of PCA. **(B)** Line plot shows expression levels for transcripts higher expressed in D3 cRBC compared to *ex-vivo* cells. **(C)** Correlation matrix and expression plots without filtering lowly expressed genes.


3

Chapter 3

CD14+ monocytes repress gamma globin expression at early stages of erythropoiesis

Steven Heshusius, Esther Heideveld, Marieke von Lindern and Emile van den Akker

In preparation

Abstract

In β-hemoglobinopathies, reactivation of gamma- at the expense of beta-globin is a prominent therapeutic option. Expression of the globin genes is not strictly intrinsically regulated during erythropoiesis, supported by the observation that fetal erythroid cells switch to adult hemoglobin expression when injected in mice.

We show cultured erythroblasts are a mix of HbA restrictive and HbA/HbF expressing cells and that the proportion of cells in the latter population depends on the starting material. Cultures started from CD34+ cells contain more HbA/HbF expressing cells compared to erythroblasts cultured from total peripheral blood mononuclear cells (PBMC). Depletion of CD14+ cells from PBMC resulted in higher HbF/HbA percentages. Conversely, CD34+ co-culture with CD14+ cells reduced the HbF/HbA population through cell-cell proximity, indicating that CD14+ actively repressed HbF expression in adult erythroid cultures. RNA-sequencing showed that HbA and HbA/HbF populations contain a limited number of differentially expressed genes, aside from *HBG1/2*. Co-culture of CD14+ cells with sorted uncommitted hematopoietic progenitors and CD34-CD36+ erythroblasts showed that hematopoietic progenitors prior to the hemoglobinized erythroid stages are more readily influenced by CD14+ cells to downregulate expression of *HBG1/2*, suggesting temporal regulation of these genes. This possibly provides a novel therapeutic avenue to develop β-hemoglobinopathies treatments.

Introduction

Beta-hemoglobinopathies are characterized by mutations within the beta-globin locus, leading to dysfunctional adult hemoglobin (HbA). Attempts to find treatments to these diseases has led to extensive characterization of the key molecular pathways that reactivate gamma-globin (*HBG1/2*) at the expense of beta-globin (*HBB*). Recent progress has focused on BCL11A and LRF1, which are key repressors of HBG1/2 during adult erythropoiesis [1-4]. Different approaches to genetically interfere with the expression of HbF repressors, like BCL11A, are being developed and tested in clinical trials (for review see [5]). However, the requirement for hematopoietic stem cell (HSC) transplantation to introduce the genetically manipulated HSC restricts the scale at which these treatments can be applied. Particularly, because the main endemic countries of these diseases, like African sub-Saharan regions, have limited resources available for transplantation-type treatments. Chemical treatment that targets processes regulating globin expression would allow broad scale application, either in combination with currently used treatments to increase HbF like hydroxyurea, or alone as novel treating entities. The existence of hereditary persistence of hemoglobin (HPFH) in adult erythrocytes as well as the flexibility to re-express HBG1/2 during *in vitro* erythropoiesis underscore an important degree of flexibility in HBG1/2 regulation. While cultured red blood cells are similar to *ex-vivo* red blood cells in important characteristics, like enucleation, oxygen saturation, deformability and surface protein expression, they frequently express higher levels of fetal hemoglobin (~3%) compared to adult erythrocytes [6]. Although the developmental expression pattern of the globin suggests intrinsic regulation, it can also occur extrinsically as fetal-, or induced pluripotent stem cell derived erythroid cells spontaneously switched from fetal-, to adult hemoglobin expression upon injection into irradiated NOD/SCID mice [7]. These data suggest a role for niche cells and signal transduction in extrinsic regulation of globin chains. While *in vivo* erythropoiesis and hematopoiesis is dependent on niche cells, like central macrophages [8-14], most erythroid culture models do not take into account the extrinsic signal that initiate from cell-cell interactions. We have previously reported that monocytes/macrophages within the PBMC population positively differentiate into macrophages that resemble erythroid island macrophages that positively influence the hematopoietic stem and progenitor cells leading to increased erythroid outgrowth [15,16]. In contrast to erythroid cells derived from CD34+ cells, those derived from whole blood PBMCs thus arise from HSPCs that are influenced by cells present in PBMCs and by differentiated macrophages originating from PBMC-monocytes. We used this characteristic of the culture system to study regulation of globin subunits by extrinsic factors and demonstrated that PBMC derived erythroid cells have lower HbF compared to CD34+ derived erythroid cells. This is reflected by a smaller fraction of cells expressing gamma globins. The effect depended

on cell-proximity between HSPC and the CD14+ fraction from PBMCs, but was absent upon co-culture of already committed erythroblasts with CD14+ cells. In addition, RNA sequencing showed limited differences between cells with and without HbF expression. These data suggest that CD14+ monocytes/macrophages can signal to hematopoietic progenitor cells priming globin expression to *HBB1* at the expense of *HBG1/2*.

Methods

Cell sorting

CD14 and CD34 MicroBeads (Miltenyi Biotec; Bergisch Gladbach, Germany) were used for magnetic-activated cell sorting (MACS) of PBMC fractions (manufacturer protocol). Flowcytometric cell sorting was used sort populations of progenitors based on CD34 and CD36. Sorting after intracellular staining for globin was performed on FACS Aria 4L (BD Biosciences; Oxford, UK).

Cell cultures

PBMCs were isolated buffycoat of blood donations. CD34+ cells were isolated from PBMCs by MACS. Both were cultured in a three-phase liquid erythroid culture system as previously described [6]. Low iron experiments contained 100mg/ml holotransferrin one third of the concentration normally used throughout culture.

Co-culture

CD34 and CD14 fractions were put in co-culture at the ratios indicated in the legends, (ratio in blood is ~1:100 CD34:CD14). Trans-well co-cultures were performed by seeding CD14+ cells at the bottom and CD34+ in the trans-well (0.4µm polyester membrane, Corning; NY, USA), cultured for 8 days and differentiated for two days to allow hemoglobinization as described before [6]. For co-cultures with sorted CD34CD36 fractions, freshly isolated CD34+ cells were cultured for three days, sorted and co-cultured for an additional 5 days.

Flowcytometry

Cells were washed twice with PBS and incubated in FACS-buffer (1% bovine serum albumin) with surface antibodies against CD71-VB405 (Miltenyi Biotec; 1:200), CD235-PE (Acris; 1:2500), CD14-APC (Miltenyi Biotec; 1:50), CD16-PE (BD Biosciences; 1:80); CD36-FITC (Pelicluster, Amsterdam, The Netherlands; 1:100); CD34-APC (IQ products, Groningen, the Netherlands; 1;10). Isotype controls were IgG1k-FITC (biolegends), IgG1-PE (Diaclone); IgG1k-APC (eBioscience). For hemoglobin staining on erythrocytes cells were fixed with 0.025% glutaraldehyde (sigma-Aldrich) and 0.5% paraformaldehyde

(Signa-Aldrich) in PBS and permeabilized with 0.5% NP40 (Sigma-Aldrich) prior to staining HbA-PE (Santa Cruz; 1:1000) and HbF-APC (Invitrogen; 1:1000).

HPLC

$1*10^7$ cells were collected and analysed for Hb isoform expression by high-performance cation-exchange liquid chromatography (HPLC) on Waters Alliance 2690 equipment as previously described [17].

Quantitative PCR

Samples of 48h differentiated erythroblasts were collected from three individual donors. RNA was isolated using TRIzol RNA isolation reagent (ThermoFisher Scientific).

RNA sequencing and data analysis

A technique previously described by Nicolet et al was used to reverse crosslinking after intracellular globin staining, prior to sequence library preparation [18]. Sequencing libraries were prepared using Trizol RNA isolation, cDNA amplification and rRNA depletion using HyperPrep Kit with RiboErase (KAPA Biosystems, Pleasanton, CA, USA) as described by manufactures. Samples were sequenced to a depth $30x10^6$ x 75bp paired end reads. RNA-seq reads were mapped to GRCh38v85 using STAR. Lowly expressed genes (1 count per million mapped reads in at least two samples) were filtered prior to differential expression analysis in EdgeR, testing differential expression in a paired design with a quasi-log-likelihood F-test.

Real-time quantitative PCR

cDNA was synthesized with the QuantiTect reverse transcription kit (QIAGEN, #205313) according to the manufacturers' instructions and Quantitative PCR was performed on a StepOnePlus Real-Time PCR system (ThermoFisher Scientific, #4376600) using Power SYBR green master mix (ThermoFisher Scientific, #4367659; 1µM primers; final volume 20µl; cycles: 10min 95 °C; 40 cycles: 15sec at 95 °C and 1min at 60 °C). Ct-values were normalized against housekeeping genes 18S and HPRT. Graphpad Prism V7.04 (GraphPad Software) was used for statistic testing and visualization of fold-change mRNA expression. RT-PCR-Primers used (5'-3'): *HBB*fw; ACAGCCACCACTTTCTGAT; *HBB*rv, AGCTGCACTGTGAC-AAGCTG; *HBG1/2*fw, AAACGGTCACCAGCACATTT; *HBG1/2*rv, GAAGGTGCTGACTTCCTTGG; *18S*fw, CACGGCCGGTACAGTGAAAC; *18S*rv, AGAGGAGCGAGCGACCAA; *G3PH*fw, CATCACG-CCACAGTTTCC; *G3PH*rv, TCCCATCACCATCTTCCA.

Results

CD34+ derived erythroblasts express significantly higher levels of HbF compared to PBMC-derived erythroblasts

In the majority of adults, fetal hemoglobin (HbF) makes up less than 1% of the total hemoglobin content in erythrocytes. However, erythrocytes of individuals with hereditary persistence of fetal hemoglobin (HPFH), and also *in vitro* cultured adult erythroid cells, may express HbF at levels higher than 5% as determined by HPLC [6]. HPLC measures bulk hemoglobin and cannot discriminate between increased numbers of cells expressing gamma globin or increased gamma expression in the fraction of cells that already express HbF. To this end, we analyzed hemoglobin expression in *ex-vivo* and cultured erythroid cells by flow cytometry. Upon intracellular staining for beta globin and gamma globin, components of HbA and HbF respectively, adult erythrocytes showed mostly HbA, with a fraction of cells positive for HbF (Figure 1A). Erythroblasts from human fetal liver predominantly showed HbF expression, indicating that the assay is specific for the two globin types (figure 1B). In contrast to adult and fetal *ex-vivo* erythroid cells, cultured adult erythroid cells showed two distinct populations; one expressing HbA and one expressing both HbF and HbA (Figure 1C). No cells expressing only HbF are found. Of note, erythroid cells cultured from cord blood are significantly enriched in HbF/HbA expressing cells, reflecting the ongoing globin switching occurring at this developmental stage (supplemental figure 1A). The increase in HbF expression in adult cultures was dependent on the starting source for erythroid cultures. Erythroid cultures initiated from total adult peripheral blood mononuclear cells (PBMC) showed a significantly lower HbF expression and higher HbA expression compared to cultures that were started from purified CD34+ cells from PBMC (Figure 1C, D). The increase in HbF and the linked decrease in HbA suggests an active regulation of globin chain expression. HPLC measurements confirmed that cultures from PBMC contained 1-3% HbF compared to >5% HbF in cultures started from selected CD34+ cells (supplemental figure 1B). CD34+ cells are a heterogeneous population of hematopoietic progenitors at various stages that upon culture may result in two different populations of HbA and HbA/HbF expressing cells. Thus, we tested if the two populations (HbA or HbA/HbF) can arise from one single CD34+ cell or whether the ability to express HbF is segregated at the progenitor level. Single cell sorted adult CD34+ cells were differentiated to erythroid cells. The single cell cultures still showed two populations, one expressing high HbF and one expressing low to no HbF ruling against the presence of two different parental hematopoietic progenitors. Cultures from different clones showed between 56% and 63% of cells in the HbF high population (Figure 1E), which was significantly higher compared to non-single cell derived CD34+ cultures. The presence of two populations in all clones shows that erythroid cells derived from a single CD34+ cells can express different hemoglobin expression profiles. The data

also indicated that in a selection of cells the HBG1/2 remains silent. Of note, to obtain enough cells to perform flow cytometry for HbA and HbF the CD34+ single cells were cultured for a minimum of 25 days

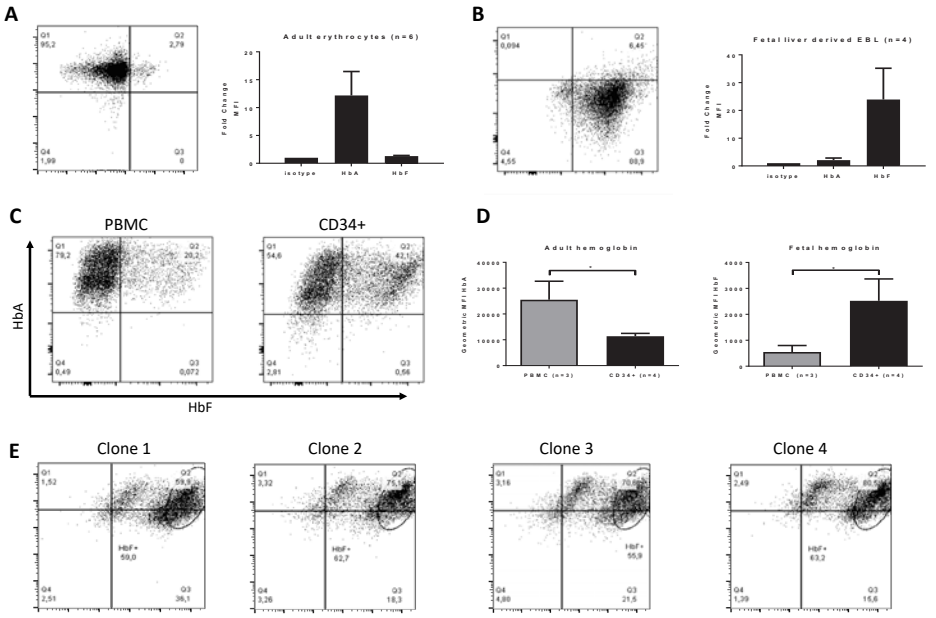


Figure 1. Expression of HbF and HbA is dependent on erythroblasts culture starting material. Erythrocytes from human adult peripheral blood (**A**; n=6) or human fetal liver (**B**; n=4) were stained with anti-HBB1 (y-axis) and anti-HBG1/2 (x-axis). Representative dot plots are show and the histograms indicate the mean fluorescent intensity of the complete globin expressing population. **(C)** erythroblasts were cultured from PBMC or from CD34+ cells isolated from PBMC. Expression of HBB1 and HBG1/2 was measured as indicated in (A). Representative dot plots are shown. **(D)** The histograms indicate the mean fluorescent intensity of the complete globin expressing population (PBMC, n=3 and CD34+ cells, n=4; *p<0.05 Student's T-test). **(E)** CD34+ cells were single cell sorted and cultured as indicated in material and methods. The dot plots represent anti-HBB1 (y-axis) and anti-HBG1/2 (x-axis) of four different clonal CD34+-single-cell-derived erythroblast populations.

CD14+ monocyte/macrophages from PBMC reduce HbF expression in erythroid cultures

Erythroid differentiation in the bone marrow occurs on erythroblast islands, structures of erythroid progenitors surrounding a central macrophage [8,9]. Defects in the interaction between macrophages and erythroid cells have been reported to result in erythropoiesis defects [14,19]. PBMC cultures contain significant amounts of cells capable of providing support and/or interactions with CD34+, erythroid progenitors and erythroblasts. In fact, 14% of PBMCs are CD14+ monocytes that can differentiate to erythropoiesis supporting

macrophages in our culture medium [15,20]. CD14+-derived macrophages from PBMC interact with, and support HSPC [15 2017,21,22]. To test if a similar interaction was responsible for the repression of HbF, PBMCs were depleted for CD14+ cells. Differentiating erythroblast from CD14+ depleted PBMCs showed a significantly increased frequency of HbA/HbF expressing cells and a higher HbF mean fluorescent intensity (MFI) compared to erythroid cultures derived from total PBMCs (figure 2A-C). The increased fetal hemoglobin expression did not result from overall lower hemoglobinization (supplemental figure 2). Of note, cultures that received suboptimal levels of holo-transferrin did show reduced hemoglobinization (supplemental figure 2).

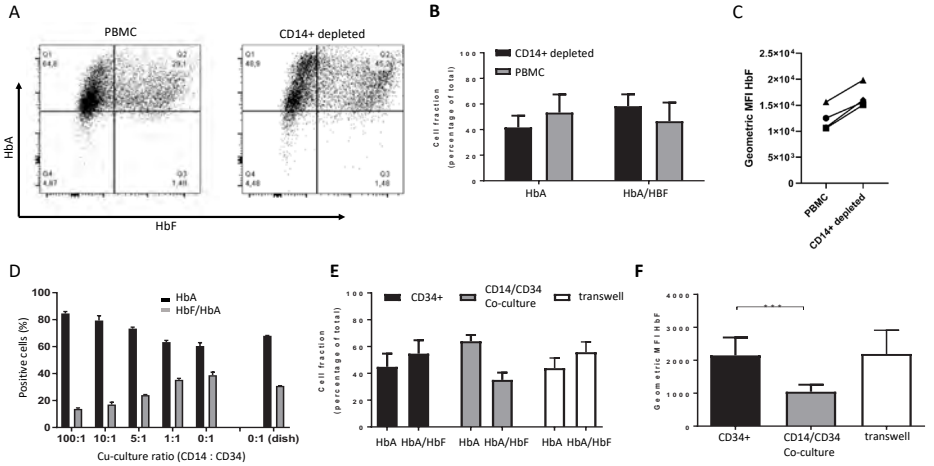


Figure 2. CD14+ monocytes promote repression of HbF expression in erythroblasts. (A) PBMC and PBMCs depleted for CD14+ cells were cultured to obtain hemoglobinized erythroblasts as indicated in material and methods. The expression of HBB1 and HBG1/2 was assessed by flow cytometry. Dot plots show representative erythroblasts cultured from PBMC (left) or PBMC depleted for CD14+ cells (right; anti-HBB1, y-axis; antiHBG1/2, x-axis). **(B)** Bar graph shows the percentage of events in Q1 (HbA cells) and Q2 (HbA/HbF cells) from **(A)** for erythroblasts derived from PBMC (grey bars) or from CD14+ depleted PBMC (black; n=3). **(C)** Mean fluorescence intensity of HBG1/2 from the data presented in (A). **(D)** CD34+ and CD14+ cells were purified from PBMC and co-cultured using specific ratios of CD14+ cells to CD34+ cells as indicated on the x-axis (CD14+:CD34+). Bar graphs represent the percentage of cells that are HBB1 positive (Black, HbA) or HBB1/HBG1/2 double positive (grey, HbA/HbF) as determined by flow cytometry (supplemental figure 3). **(E, F)** CD34+ and CD14+ cells were purified from PBMC as indicated in material and methods. Cells were co-cultured using 100:1 CD14+ to CD34+ ratio in full contact (grey bars), separated by transwells (white bars). The black bars represent CD34+ cultured without CD14+ cells as a control. Bar graphs in **(E)** represent the percentage of cells that are HBB1 positive (HbA) or HBB1/HBG1/2 double positive (HbA/HbF) as determined by flow cytometry (n=3). Bar graph in **(F)** depicts the geometric mean fluorescence of HBG1/2 expression (HBF) as measured by flow cytometry (n=3, ***p<0.01; Student's T-test).

Next, we asked if CD14+ alone would be sufficient to reduce HbF in CD34+ cultures. The CD34+ fraction and the CD14+ fraction were isolated from buffy coats and co-cultured. Addition of CD14+ cells to CD34+ at the start of culture reduced HbF expression compared to cultures started from CD34+ cells alone in a dose dependent manner (Figure 2D, supplemental figure 3). To test if reduced HbF expression required direct contact or resulted from a secreted factor, CD34+ were differentiated to erythroid cells in a transwell co-culture setup. Although co-culture in direct contact with CD14+ positive cells again reduced HbF expression in erythroid cultures, erythroblasts co-cultured with CD14+ cells in a transwell setting did not reduce HbF expression (Figure 3 E, F). This suggests that HbF repression in culture requires cell-cell proximity of CD14+ and erythroid progenitor cells.

The Transcriptomes of HbA/HbF co-expressing compared to HbA only expressing cells is highly similar.

The presence of HbF/HbA co-expressing cells and HbA single expressing cells in culture, combined with the observation that the balance between these populations can be regulated through specific interactions with monocyte/macrophage populations suggests that potential globin regulators may be differentially regulated upon comparing these two erythroid populations. To find such potential targets the transcriptome of HbF/HbA co-expressing and HbA expressing populations cultured from adult PBMCs was assessed by RNA sequencing of the erythroblast stage (CD71+/CD235+; supplemental figure 4A). RT-PCR on the RNA isolated from these population indicated beta globin expression in HbA or HbF/HbA cells, however beta-globin levels were ~4-fold lower in HbF/HbA cells (figure 3A). As expected, gamma-globin levels were nihil in HbA only cells.

Principle component analysis of the corresponding, normalized transcriptome data revealed that the first two components (45% of the variation) separated samples by donor, while the third component (15% of the variation), separated the HbF/HbA sorted population from the HbA population (figure 3B. Interestingly, out of over 16000 distinct RNA species detected only a surprisingly small number of 25 genes (8 up and 17 down in HbF/HbA vs HbA expressing cells) were differentially expressed with at least a two-fold difference between the two sorted populations (false discovery rate (FDR) < 0.01; fold change: -2<FC>2; Log count per million CPM>2; Figure 3B-D; table 1). HBG1 and HBG2, were the most differentially expressed between the two populations conforming the sorting results by HPLC and flow cytometry in figure 1 (supplemental figure 4A). Among the 8 RNAs upregulated in HbF/HbA cells, 5 were located within the beta globin locus. Besides *HBG1/2*, these were embryonic epsilon globin chain (*HBE1*), and the non- coding RNAs *BGLT3* and *HBBP1*, both non-coding RNAs were significantly higher expressed. In addition, two mitochondrial RNAs and one transmembrane protein TMEM233 were upregulated (table 1). GO analysis indicates oxygen carrying/binding

(upregulated RNA's) and transforming growth factor binding (down regulated RNA's) as enriched GO terms (data not shown). Among the genes that are expressed at lower levels in HbF/HbA cells compared to HbA only cells are *Endoglin* (*ENG*), *Thrombosondin* (*THBS1*), *Pleckstrin* (*PLEK*), *Plexin D1* (*PLXND1*), *Rap1-interacting adapter molecule* (*RIAM/ APBB1IP*) and *AHNAK,* encoding a large intracellular scaffold protein, which may all alter response to environmental factors for instance signal transduction by TGFβ (transforming growth factor beta) family members.

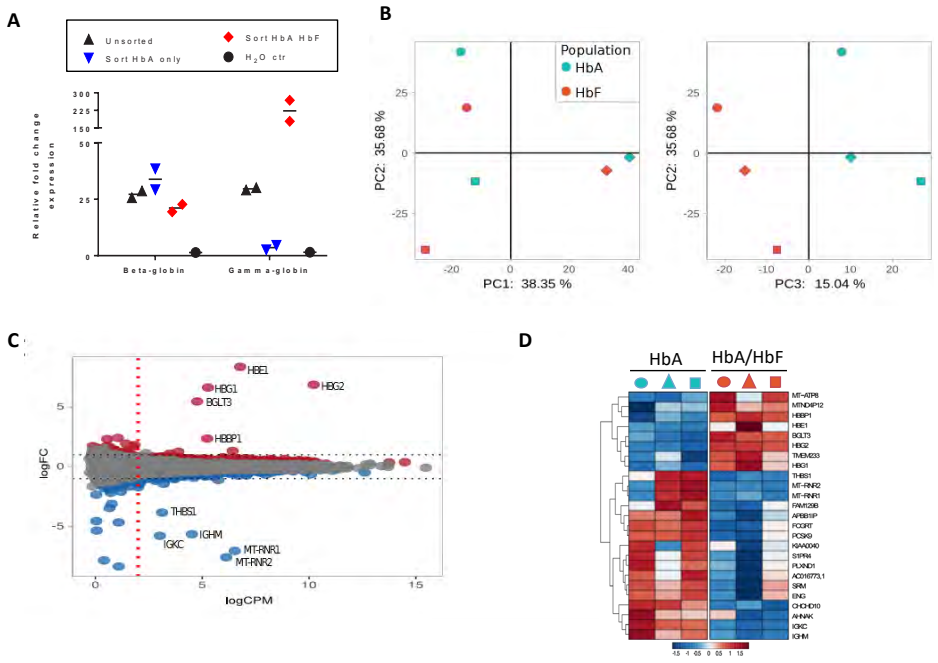


Figure 3. The RNA expression profile between sorted HbA and HbA/HbF expressing erythroblasts is highly similar. Erythroblasts cultured from three independent donors were stained for HBB1 and HBG1/2 and the HBB1 and HBB1/HBG1/2 expressing erythroblasts were FACS-sorted after which RNA was isolated and subjected to RNA-sequencing as indicated in material and methods. **(A)** RT-qPCR using primers for beta-globin (HBB1) and gamma-globin (HBG1/2) on mRNA isolated from HbA (blue) and HbA/HbF (red) sorted populations, unsorted (black) and water control (black circles) as indicated. The relative fold change expression to the water control is depicted. **(B)** principle component analysis showing the variation between the sorted populations in the three first components. **(C)** RNA sequencing expression analysis depicted as log2 fold change (y-axis) against average count per million reads per gene (CPM, x-axis). Differentially expressed genes crossing the threshold of p<0.01 are indicated in red for upregulated and blue for downregulated (dark spots represent unchanged RNAs). The blue lines indicate the log2 fold change cut-off (CPM), -1<gene>1 and the red line indicates the CPM cutoff (>2). **(D)** Heatmap depicting z-scores of the 25 differentially expressed RNAs clustered (rows and columns) using one-minus-Pearson-correlation with average linkage.

Similarly, reduced expression of the *sphingosine-1-phosphate receptor 4* (*S1PR4*) may alter response to the environment. Several transcriptional regulators are known that control or influence the expression of globin subunits. However, known globin regulators, BCL11A, KLF1, ZBTB7A and SOX6 were not found to be differentially expressed (Supplemental figure 4C). Adding to this, we do not observe a reduction in KLF1 target genes (e.g. CD44, BCAM, CARM1, BCL11A; Supplemental figure 4C). Moreover, among the deregulated RNA's no other transcriptional regulator was identified.

Repression of HbF by CD14+ cells occurs before erythroblasts stage

As the number of differentially expressed genes is surprisingly low between HbA/HbF and HbA sorted populations and cannot be attributed to RNA expression levels of known or novel transcriptional regulators, we hypothesized that posttranscriptional control mechanisms regulate the frequency of HbF/HbA cells, and/or (*ii*) globin subunit regulation occurred before the hemoglobinized CD71+/CD235+ erythroblast stage analyzed and that the molecular mechanism driving this differential globin expression may already be downregulated at the hemoglobinized erythroblast stages. The alternative explanation would fit with high HbF in the cultures from single cell CD34+ in figure 1E, if this were to result from prolonged expansion of the earliest progenitor stage. Indeed, CD34+ hematopoietic stem and progenitor cells (HSPC) are a mix of hematopoietic cells at different stages of differentiation from hematopoietic stem cells to more committed lineage progenitors [15]. Lineage specification progresses from CD34+CD36- hematopoietic stem and progenitor cells to CD34+CD36+ megakaryoid/erythroid common progenitors to CD34-CD36+ erythroblasts [15,23]. Co-cultures with these fractions and the CD14+ cells were assessed to determine the stage at which the repression occurred. Co-culture of CD34-CD36+ erythroblasts with CD14+ monocytes did not decrease HBG1/2 expression. This supports the RNA-sequencing data results that the regulation of HBG1/2 in erythroid cultures by CD14+ cells does not occur at the committed lineage restricted (pro)erythroblast stage (figure 4A). In contrast, CD14+ cells were able to repress *HBG1/2* expression in CD34+CD36-/+ progenitors with the biggest repression occurring in the early CD34+CD36- co-cultures with CD14+ cells (figure 4B). Of note, repression of HbF lead to a concomitant increase in HbA. The data indicate that regulation of globin genes by CD14+ cells can occur early during hematopoiesis before the commitment to erythroid restricted progenitors.

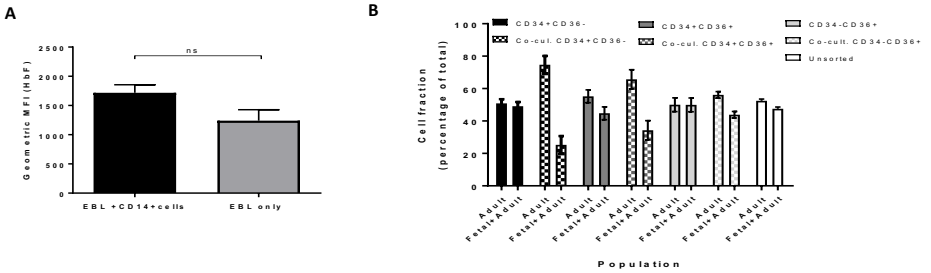


Figure 4. Inhibition of HbF expression by CD14+ cells occurs early during hematopoiesis/ erythropoiesis. (A) Erythroblasts (CD71+/CD235low) were cultured with and without PBMC-isolated CD14+ cells as indicated in material and methods. Bar graph shows the mean fluorescent intensity of gamma globin expression as measured by flow cytometry. **(B)** Hematopoietic progenitors were sorted based on their expression of CD34 and CD36 expression. In short, CD34+CD36- represent hematopoietic stem and progenitors cells, CD34+CD36+ cells are committed to a megakaryoid/erythroid fate but no myeloid fate and CD34-CD36+ cells are pro-erythroblasts [15]. Cells were cultured in presence (solid bars) or absence of PBMC-isolated CD14+ cells (striped bars). The percentage of HbA and HbA/HbF cells was measured by flow cytometry and indicated on the y-axis. Note that the reduction of HbF/HbA populations by CD14+ cells is proportional to the immaturity of the progenitors (N=3)

Discussion

Hematopoietic stem cell differentiation to erythroid cells occurs within the bone marrow. This process depends on interactions with specific cells present within the different bone marrow niches. Distinctive supporting macrophages can be identified both in the hematopoietic stem and progenitor cell niche as well as within the erythroid niche, where they can provide signals to the differentiating hematopoietic cells [8,9,24-26]. Reminiscent to these niche functions, PBMC-isolated CD14+ monocyte-derived macrophages were found to both support survival of HSPC and recapitulate central macrophage functions [15,20,21]. Here we find that the CD14+ monocyte-derived macrophages alter the balance between gamma and beta globin gene expression in erythroid cultures. Specifically, we find that gamma-globin down regulation occurs prior to the erythropoietin, stem cell factor and glucocorticoid responsive non-hemoglobinized committed erythroblast stage, defined as $CD71^{high}/CD235^{low/+}$ [6]. In fact, the highest repression was observed at the earliest CD34+CD36- hematopoietic stem and progenitor population, suggesting that erythroid specific factors like KLF1 are not involved in this regulation. Indeed, KLF1 was not deregulated between hemoglobinized erythroblasts expressing HbA-only or HbA/HbF. Moreover, the regular suspects of globin regulation, e.g SOX6, BCL11A and LRF where not differentially expressed at the erythroid committed erythroblast stage. However, as some of these regulators (e.g. BCL11A) are expressed and functional at the HSPC stage or even in HSCs [27-29], the data does not rule out a role for these factors to

regulate HbA/HbF expression in early hematopoietic progenitors. The data suggests a novel concept that globin expression induced later during erythroid differentiation can be regulated at these early HSPC progenitor stages. This would fit with the proposed flexibility in globin-chain expression observed by Kobari et al., who have reported that SCD-iPSC derived erythroid progenitor cells that *in vitro* express fetal hemoglobin (HbF) will convert to adult beta globin (HbA) expression containing the mutated adult beta sickle cell HBS peak in HPLC upon injecting *in vivo*, revert to producing beta-globin expressing erythrocytes upon injection into NOD-SCID mice [7]. In addition, Singh et al have shown that the gamma-globin promoter was highly methylated in the earliest stage of hematopoietic stem progenitor cells (CD34(+)CD36(-)) and that methylation progressively decreased as HSPC differentiation progressed in sorted adult bone marrow progenitors to committed erythroblasts [30]. Although, we did not look at the methylation status of the promoters in our erythroid cultures, combined with the flexibility in globin chain expression it substantiates a hypothesis where the niche can instruct erythropoiesis to express specific globin subunits in a process that involves epigenetic modifications at the earliest differentiation stages of erythropoiesis. In line with the role of a niche cell, the CD14+-derived macrophages from PBMCs regulate globin expression through cell-cell contact. We and others have previously shown that these PBMC-CD14-derived CD163/CD169+ macrophages resemble and function like erythroid island central macrophages [20,21]. We propose that also macrophages within the bone marrow niche may similarly exert signaling to regulate globin expression. We acknowledge that the ratio between the macrophages and the erythroid precursors is off and in favor of macrophages, whereas in the niche one macrophage is surrounded by >10 erythroid cells. However, as a clear dose response curve of macrophages on the repression of HbF is found, this clearly indicates that signaling influences specific globin subunit expression. Importantly, here we use a macrophage co-culture system with erythroid precursors recapitulating only one component of the hematopoietic/erythroid niche *in vitro* and hence a simplification of processes occurring within the bone marrow. Despite these considerations, the data indicates that hematopoietic progenitors can be influenced by signaling to control globin expression and provides a possible mechanism through which the bone marrow niche specifically produces beta globin expressing erythroid cells. The precise identity of these signals and the induced signal transduction in erythroid progenitors remains unknown but are important to elucidate. Mapping the signal transduction induced by macrophages on erythroid/hematopoietic progenitors may link to the well-known regulators of the globin genes or may uncover novel pathways. These pathways could potentially be blocked to increase gamma expression providing novel therapeutic options to increase HbF in sickle cell or beta-thalassemia either on its own or in combination with other (HbF-inducing) therapy.

Acknowledgments

Research was funded by the Landsteiner Foundation for Blood Transfusion Research (LSBR 1040 and 1627) and the Netherlands Organization for Scientific Research (ZonMw TOP 40-00812-98-12128). We would to thank Sanquin research facilities for their help with flow cytometry and sorting.

Author contributions

SH performed most of the experiments. EH helped and performed CD14 isolations and technical assistance. SH, EvdA and MvL designed, interpreted and wrote the manuscript. All authors critically read and revised the manuscript.

Table 1. Up and down regulated genes in HbF/HbA expressing cells versus HbA only expressing cells

GENE	LOGFC	LOGCPM	F	PVALUE	FDR
IGHM	-5,63562	4,517326	774,7776	1,16E-17	4,85E-14
IGKC	-5,771	3,010395	353,4626	2,52E-14	8,45E-11
THBS1	-3,83388	3,116554	259,4671	4,82E-13	1,16E-09
MT-RNR1	-7,56982	6,122425	217,4913	2,53E-12	5,01E-09
MT-RNR2	-7,0361	6,52396	216,121	2,68E-12	5,01E-09
ENG	-1,10868	5,75179	53,62633	4,02E-07	0,000181
FCGRT	-1,42587	3,190215	49,07273	7,76E-07	0,000241
PLXND1	-1,04451	3,690973	33,31091	1,13E-05	0,001075
AHNAK	-1,03942	2,854409	32,673	1,28E-05	0,001143
S1PR4	-1,02594	3,000601	31,28129	1,69E-05	0,001367
CHCHD10	-1,09501	2,231917	25,20431	6,26E-05	0,002818
SRM	-1,18123	2,815445	24,4092	7,54E-05	0,003203
FAM129B	-1,45553	2,062975	24,279	7,77E-05	0,003271
APBB1IP	-1,02367	2,427376	23,90906	8,48E-05	0,003424
KIAA0040	-1,08592	2,147614	22,67123	0,000114	0,004113
PCSK9	-1,04046	2,166769	20,12631	0,000218	0,006204
AC016773,1	-1,01497	2,064593	18,45859	0,000341	0,008104

GENE	LOGFC	LOGCPM	F	PVALUE	FDR
HBG2	6,853673	10,21744	3839,172	1,22E-24	2,04E-20
HBG1	6,617616	5,271699	1226,88	1,19E-19	9,95E-16
BGLT3	5,457426	4,761285	783,0324	1,04E-17	4,85E-14
HBBP1	2,356147	5,231101	282,5109	2,15E-13	6,02E-10
HBE1	8,343875	6,773354	194,2372	7,24E-12	1,21E-08
MT-ATP8	1,323711	6,422302	117,2473	6,81E-10	1,04E-06
TMEM233	1,253602	3,878552	88,31424	7,67E-09	8,77E-06
MTND4P12	1,085315	3,462607	21,83769	0,000141	0,004639

References

1 Borg, J. et al. Haploinsufficiency for the erythroid transcription factor KLF1 causes hereditary persistence of fetal hemoglobin. Nat Genet 42, 801-805, doi:ng.630 [pii]10.1038/ng.630 [doi].

2 Masuda, T. et al. Transcription factors LRF and BCL11A independently repress expression of fetal hemoglobin. Science 351, 285-289, doi:10.1126/science.aad3312 (2016).

3 Liu, N. et al. Direct Promoter Repression by BCL11A Controls the Fetal to Adult Hemoglobin Switch. Cell 173, 430-442.e417, doi:10.1016/j.cell.2018.03.016 (2018).

4 Garcia, M. R. et al. Monocytic suppressive cells mediate cardiovascular transplantation tolerance in mice. J Clin Invest 120, 2486-2496, doi:10.1172/jci41628 (2010).

5 Orkin, S. H. & Bauer, D. E. Emerging Genetic Therapy for Sickle Cell Disease. Annu Rev Med 70, 257-271, doi:10.1146/annurev-med-041817-125507 (2019).

6 Heshusius, S. et al. Large-scale *in vitro* production of red blood cells from human peripheral blood mononuclear cells. Blood Adv 3, 3337-3350, doi:10.1182/bloodadvances.2019000689 (2019).

7 Kobari, L. et al. Human induced pluripotent stem cells can reach complete terminal maturation: *in vivo* and *in vitro* evidence in the erythropoietic differentiation model. Haematologica 97, 1795-1803, doi:10.3324/haematol.2011.055566 (2012).

8 Lee, S. H. et al. Isolation and immunocytochemical characterization of human bone marrow stromal macrophages in hemopoietic clusters. J Exp Med 168, 1193-1198 (1988).

9 Mohandas, N. & Prenant, M. Three-dimensional model of bone marrow. Blood 51, 633-643 (1978).

10 Soni, S. et al. Absence of erythroblast macrophage protein (Emp) leads to failure of erythroblast nuclear extrusion. J Biol Chem 281, 20181-20189, doi:10.1074/jbc.M603226200 (2006).

11 Soni, S., Bala, S., Kumar, A. & Hanspal, M. Changing pattern of the subcellular distribution of erythroblast macrophage protein (Emp) during macrophage differentiation. Blood cells, molecules & diseases 38, 25-31, doi:10.1016/j.bcmd.2006.09.005 (2007).

12 Soni, S., Bala, S. & Hanspal, M. Requirement for erythroblast-macrophage protein (Emp) in definitive erythropoiesis. Blood cells, molecules & diseases 41, 141-147, doi:10.1016/j.bcmd.2008.03.008 (2008).

13 Wei, Q. et al. Maea expressed by macrophages, but not erythroblasts, maintains postnatal murine bone marrow erythroblastic islands. Blood, doi:10.1182/blood-2018-11-888180 (2019).

14 Ulyanova, T. et al. VCAM-1 expression in adult hematopoietic and nonhematopoietic cells is controlled by tissue-inductive signals and reflects their developmental origin. Blood 106, 86-94, doi:10.1182/blood-2004-09-3417 (2005).

15 Heideveld, E. et al. CD14+ cells from peripheral blood positively regulate hematopoietic stem and progenitor cell survival resulting in increased erythroid yield. Haematologica 100, 1396-1406, doi:10.3324/haematol.2015.125492 (2015).

16 Heideveld, E. et al. Glucocorticoids induce differentiation of monocytes towards macrophages that share functional and phenotypical aspects with erythroblastic island macrophages. Haematologica 103, 395-405, doi:10.3324/haematol.2017.179341 (2018).

17 van Zwieten, R. et al. Partial pyruvate kinase deficiency aggravates the phenotypic expression of band 3 deficiency in a family with hereditary spherocytosis. Am J Hematol 90, E35-39, doi:10.1002/ajh.23899 (2015).

18 Nicolet, B. P. et al. Circular RNA expression in human hematopoietic cells is widespread and cell-type specific. Nucleic Acids Res 46, 8168-8180, doi:10.1093/nar/gky721 (2018).

19 Chow, A. et al. CD169(+) macrophages provide a niche promoting erythropoiesis under homeostasis and stress. Nat Med 19, 429-436, doi:nm.3057 [pii]10.1038/nm.3057 [doi].

20 Heideveld, E. & van den Akker, E. Digesting the role of bone marrow macrophages on hematopoiesis. Immunobiology 222, 814-822, doi:10.1016/j.imbio.2016.11.007 (2017).

21 Lopez-Yrigoyen, M. et al. Genetic programming of macrophages generates an *in vitro* model for the human erythroid island niche. Nat Commun 10, 881, doi:10.1038/s41467-019-08705-0 (2019).

22 Falchi, M. et al. Dexamethasone targeted directly to macrophages induces macrophage niches that promote erythroid expansion. Haematologica 100, 178-187, doi:10.3324/haematol.2014.114405 (2015).

23 Li, J. et al. Isolation and transcriptome analyses of human erythroid progenitors: BFU-E and CFU-E. Blood 124, 3636-3645, doi:blood-2014-07-588806 [pii]10.1182/blood-2014-07-588806 [doi].

24 Winkler, I. G. et al. Bone marrow macrophages maintain hematopoietic stem cell (HSC) niches and their depletion mobilizes HSCs. Blood 116, 4815-4828, doi:10.1182/blood-2009-11-253534 (2010).

25 Luo, Y. et al. M1 and M2 macrophages differentially regulate hematopoietic stem cell self-renewal and *ex vivo* expansion. Blood Adv 2, 859-870, doi:10.1182/bloodadvances.2018015685 (2018).

26 Kaur, S. et al. Self-repopulating recipient bone marrow resident macrophages promote long-term hematopoietic stem cell engraftment. Blood 132, 735-749, doi:10.1182/blood-2018-01-829663 (2018).

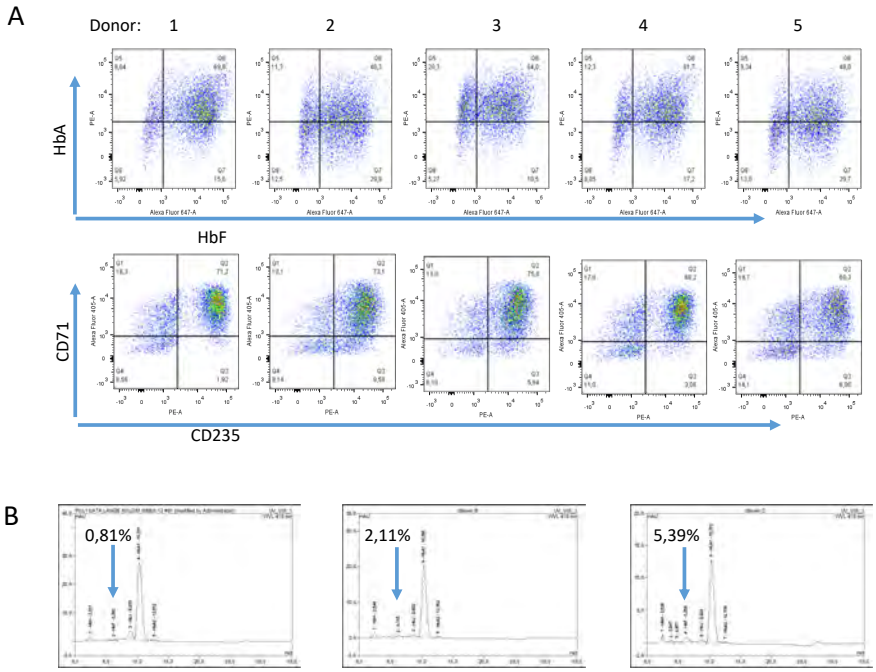
27 Lee, B. S. et al. The BCL11A transcription factor directly activates RAG gene expression and V(D)J recombination. Molecular and cellular biology 33, 1768-1781, doi:10.1128/mcb.00987-12 (2013).

28 Hoeffel, G. et al. C-Myb(+) erythro-myeloid progenitor-derived fetal monocytes give rise to adult tissue-resident macrophages. Immunity 42, 665-678, doi:10.1016/j.immuni.2015.03.011 (2015).

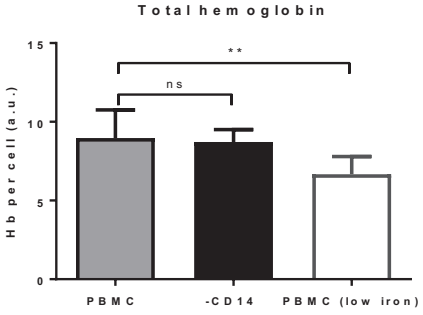
29 Turro, E. et al. A dominant gain-of-function mutation in universal tyrosine kinase SRC causes thrombocytopenia, myelofibrosis, bleeding, and bone pathologies. Sci Transl Med 8, 328ra330, doi:10.1126/scitranslmed.aad7666 (2016).

30 Singh, M. et al. The gamma-globin gene promoter progressively demethylates as the hematopoietic stem progenitor cells differentiate along the erythroid lineage in baboon fetal liver and adult bone marrow. Exp Hematol 35, 48-55, doi:10.1016/j.exphem.2006.09.001 (2007).

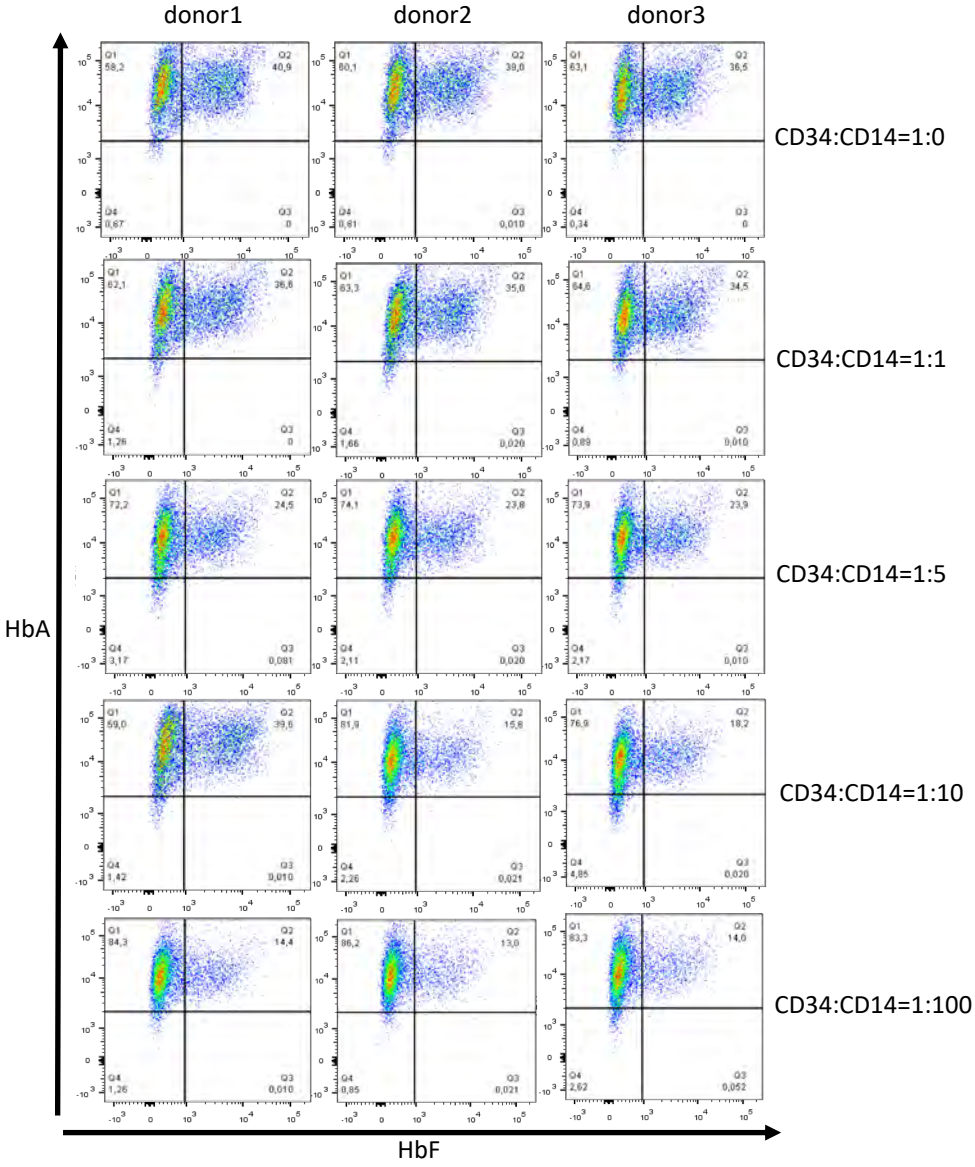
Supplemental figures



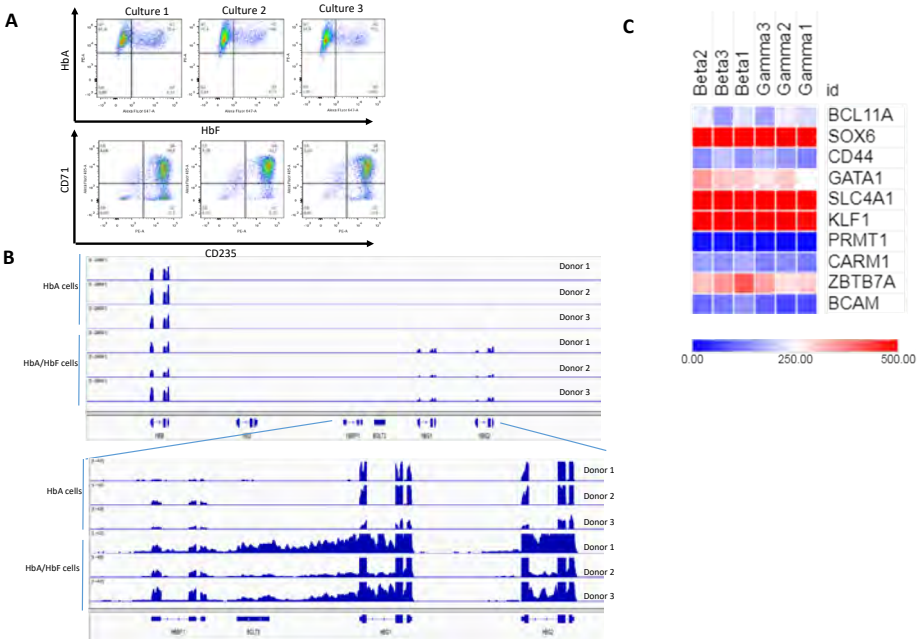
Supplemental figure 1. Globin subunit and hemoglobin measurements in cord blood and through HPLC. (A) cord blood derived CD34+ cells (n=5) were cultured to hemoglobinized erythroblasts as described in material and HBB1 and HBG1/2 were measured. Upper dot plots display HBB1 staining (y-axis, HbA) and HBG1/2 stainings (x-axis, HbF). Lower dot plots display CD71 (Y-axis) and CD235 (x-axis) stainings, indicating erythroid differentiation progression. **(B)** HPLC to fractionate the different forms of hemoglobin was performed as described in material and methods. The graphs show the absorbance at 415nm (y-axis) as a function of the retention time in minutes (x-axis). The identity of the different peaks is indicated with the blue arrow indicating the retention time of HbF with the percentage of HbF calculated as the fraction of the area under the curve of the HbF peak of the total area under the curve of all peaks.



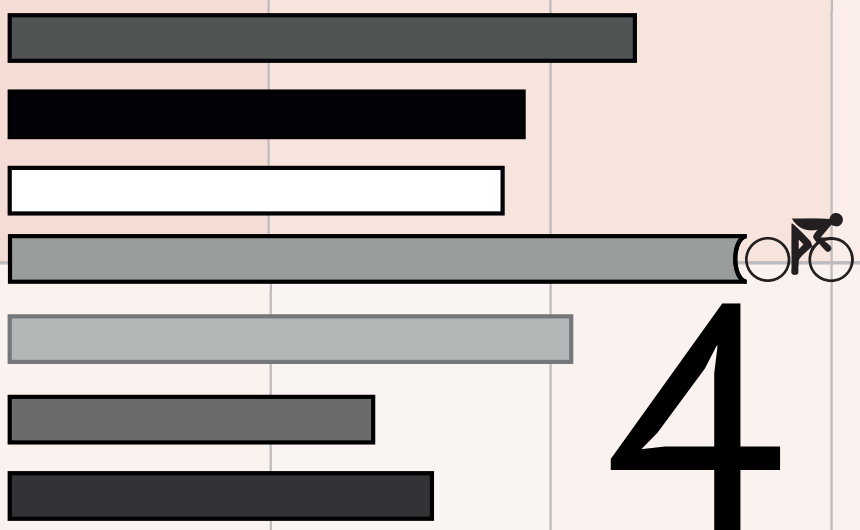
Supplemental figure 2. Erythroid cells cultured from PBMC that were CD14+ cell depleted are unchanged in total hemoglobinization. PBMCs or PBMCs depleted for CD14+ cells were cultured to hemoglobinized erythroblasts as indicated material and methods. In addition, PBMCs were cultured in low holotransferrin conditions (low iron, material and methods). Hemoglobin per cell was determined by performing o-dianosine benzidine assay to quantify the total amount of heme. No difference in the amount of heme was observed between PBMC (grey bars) and PBMC-CD14 (black bars). Lowering the total amount of available iron (through holo-transferrin) did lower total heme (PBMC-white bars). N=3, $p<0.01$ Student's T-test.



Supplemental figure 3. Increasing the ratio of CD14+ cells to CD34+ cells in co-culture results in a dose dependent effect of HbF repression. CD14+ cells and CD34+ cells from three different healthy donors were isolated and co-cultured in the ratios indicated. The dot plots represent globin expression measured using anti-HBB1 (HbA, y-axis) and anti-HBG1/2 (HbF, x-axis).



Supplemental figure 4. Comparing the RNA expression profiles between HbA and HbA/HbF expressing cells is highly similar. (A) CD34+ cells from three healthy donors were cultured to hemoglobinized erythroblasts as indicated in material and methods. The upper dot plots represent HBB1 (HbA, y-axis) as a function of HBG1/2 (HbF, y-axis) and the lower blots indicate the differentiation progression as determined by stainings for CD71 (y-axis) and CD235 (x-axis). The HbA and the HbA/HbF expressing cells were sorted and subjected to RNA-sequencing as described in material and methods. **(B)** Integrative genomics viewer (IGV) [james, 2017]) showing the mapping of RNA-sequencing reads (y-axis) onto the beta globin locus (x-axis). Note that the HbA/HbF sorted cells express high levels of HBG1/2 compared to HbA sorted cells. Lower IGV zoom in displays the HBG1 and HBG2 mRNA signal and potential read through into the neighboring BGLT3 non-coding-RNA gene in the HbA/HbF sorted cells but not in the HbA sorted cells (n=3). **(C)** Heatmap displaying a Pearson clustering of z-scores calculated from CPM for the indicated selection of globin locus regulators (y-axis). Note that the HbA and HbA/HbF populations do not segregate (x-axis) showing that variations within the expression of these genes does not explain the higher HbF expression in the HbA/HbF sorted cells.



4

Chapter 4

Epigenomic analysis of *KLF1* haploinsufficiency in primary human erythroblasts

Steven Heshusius, Laura Grech, Nynke Gillemans, Rutger W. W. Brouwer, Xander T. den Dekker, Wilfred F. J. van IJcken, Benjamin Nota, Alex E Felice, Thamar B. van Dijk, Marieke von Lindern, Joseph Borg, Emile van den Akker and Sjaak Philipsen

In preparation

Abstract

Haploinsufficiency for the erythroid-specific transcription factor KLF1 is associated with hereditary persistence of fetal hemoglobin (HPFH). Increased HbF ameliorates the symptoms of β-hemoglobinopathies and downregulation of KLF1 activity has been proposed as a potential therapeutic strategy. However, the feasibility of this approach has been challenged by the observation that KLF1 haploinsufficient individuals with the same *KLF1* variant, within the same family, display a wide range of HbF levels. This phenotypic variability is not readily explained by co-inheritance of known HbF-modulating variants in the *HBB*, *HBS1L-MYB* and/or *BCL11A* loci. We studied Maltese individuals in which *KLF1* p.K288X carriers display HbF levels ranging between 1-12% of total Hb. Using a combination of gene expression analysis, chromatin accessibility assays and promoter activity tests we find that allelic variation of wildtype *KLF1* may explain a significant part of the variability in HbF levels observed in *KLF1* haploinsufficiency. Our results have general bearing on the variable penetrance of haploinsufficiency phenotypes and on conflicting interpretations of pathogenicity of variants in other transcriptional regulators such as *EP300*, *GATA2* and *RUNX1*.

Introduction

β-hemoglobinopathies such as β-thalassemia and sickle cell disease (SCD) are caused by mutations within the β-globin subunit of adult hemoglobin (HbA, α2β2). High levels of fetal hemoglobin (HbF, α2γ2) ameliorate the symptoms of β-thalassemia and SCD. Reactivation of fetal γ-globin expression and, in the case of SCD, concomitant downregulation of sickle β-globin expression is therefore seen as an attractive approach to improve the condition of β-hemoglobinopathy patients. Although not completely resolved, fetal-to-adult hemoglobin switching depends on a core network of transcriptional regulators converging on B-cell lymphoma 11A (BCL11A) and Lymphoma Related Factor (LRF, encoded by *ZBTB7A*), that are direct repressors of the *HBG1/2* genes encoding γ-globin [1-3]. In erythroid cells, *BCL11A* and *ZBTB7A* expression is induced by the transcription factor Krüppel-like factor 1 (KLF1) [4-6]. Haploinsufficiency for *KLF1* results in hereditary persistence of fetal hemoglobin (HPFH) through a mechanism that involves reduced expression of BCL11A [4,6] and LRF [5]. Modulation of this regulatory network has been proposed as an approach to increase HbF levels, a notion that has been confirmed in animal models [3,7-10]. As BCL11A [11] and LRF [12] are tumor suppressors with important functions outside erythropoiesis, targeting the correct cells is important. In addition, KLF1 [13,14] and LRF [15] are essential for terminal erythroid differentiation. *KLF1* variants are associated with a broad spectrum of benign but also severe erythroid defects [16]. These include the In(Lu) and Indian blood types [17], pyruvate kinase deficiency [18], increased HbA2 (α2δ2) [19], increased zinc protoporphyrin [20], congenital dyserythropoiesis [21-24] and HPFH [4]. In families with HPFH caused by *KLF1* haploinsufficiency, HbF levels are variable suggesting additional regulators [4]. This variation could only be partially explained by coinheritance of known HbF-modulating genotypes at the *HBB* locus itself and *trans*-acting HPFH loci *BCL11A* and *HBS1L-MYB*. This raises the possibility that another, as yet unidentified, *trans*-acting locus is involved. To investigate the phenotypic variability of *KLF1* haploinsufficiency, we used cells carrying the *KLF1* p.K288X variant from cultures of two Maltese HPFH pedigrees [4] and control Maltese individuals. Gene expression (RNA-seq) and chromatin accessibility assays (ATAC-seq [25]) demonstrated that *KLF1* haploinsufficiency affects erythroid gene expression of KLF1 target genes, but with the exception of a small set of genes including *HBG1/2*- chromatin accessibility was not altered. The heterogeneity in HbF expression in HPFH samples correlated negatively with *KLF1*, *BCL11A* and *ZBTB7A* expression levels. We found that a polymorphism in the *KLF1* promoter modulates promoter activity in reporter assays. We suggest that allelic variation affecting wildtype *KLF1* expression explains a significant part of the HbF variability within this cohort of Maltese HPFH individuals.

Methods

Erythroid cultures

Blood samples were harvested following informed consent in accordance with the Declaration of Helsinki, under University of Malta approval (FREC 45/2014). Viable cells have been deposited in the BioBank of the University of Malta. Erythroid cultures were started as described previously [26,27]. For details see Supplementary Materials and Methods.

Flow cytometry

For globin staining cells were fixed with 0.025% glutaraldehyde, 0.5% paraformaldehyde and permeabilized with 0.05% NP40 (Sigma Aldrich), followed by incubation with primary antibodies: CD71-vio-blue421, CD235a-PE, CD44-APC, HbA-PE, HbF-APC. For details see Supplementary Materials and Methods.

Hemoglobin content and cell morphology

Aliquots ($1x10^5$) were analyzed for hemoglobin content by photometry as described [28]. Cell morphology was analyzed in cytospins stained with histologic dyes and neutral o-dianisidine [29], using an Olympus BX40 microscope (40× objective, NA 0.65), and Leica digital camera (DM-2500).

HPLC

$1*10^7$ cells were collected and analyzed for hemoglobin expression by high-performance cation-exchange liquid chromatography (HPLC) on Waters Alliance 2690 equipment. The column was purchased from PolyLC [30].

Real-time quantitative PCR

Samples were collected from pro-erythroblasts (T0) and differentiating (T48) erythroid cultures. RNA was isolated using TRIzol RNA isolation reagents (ThermoFisher Scientific) and used for RT-qPCR analysis as described in Supplementary Materials and Methods.

Next generation sequencing

For RNA sequencing libraries were prepared using the TruSeq Stranded mRNA library kit (Illumina) according to manufacturer's instructions. For ATAC sequencing samples were prepared according to [25] using the Nextera DNA transposase kit. Sequencing was performed on the Illumina HiSeq2500 sequencer. The data has been deposited in the European Nucleotide Archive (https://www.ebi.ac.uk/ena) under accession number PRJEB31712. Further details and data analysis are described in Supplementary Materials and Methods.

Promoter assays

The plasmids containing *KLF1* promoter variants (rs112943513 or rs3817621) within the pGL4.10 *Photinus* luciferase reporter construct were generated by Mutagenex, Inc. The p.GL4.74 *Renilla* luciferase construct was used as an internal control for transfection efficiency. The reporters were transfected into K562 cells and dual luciferase activity was measured according to the manufacturer's guidelines (Promega).

Results

KLF1 haploinsufficiency marginally affects *in vitro* erythropoiesis but induces HPFH

Maltese *KLF1* p.K288X carriers, further denoted as HPFH individuals, provide an opportunity to study the effect of reduced levels of this crucial transcriptional regulator on adult human erythropoiesis. To study the effect of *KLF1* haploinsufficiency on the transcriptome and chromatin state of human erythroid cells, we isolated peripheral blood mononuclear cells from four HPFH individuals and four control donors and differentiated these towards the erythroid lineage [26,27,31]. Initially, HPFH and control cultures showed similar expansion potential. However, during late expansion (>15 days) control cultures expanded two- to three-fold more compared to HPFH cultures (Fig 1A). Terminal differentiation induced a short proliferation phase (3 days) followed by growth arrest (Fig 1B) and morphological changes, which were comparable between HPFH and control cells (Fig 1C). Based on CD71 (transferrin receptor) and CD235a (glycophorin A) expression and overall hemoglobinization, no major differences in erythroid differentiation were observed (Fig 1D). The previously identified KLF1 target gene and surface marker CD44 was expressed lower on cells from HPFH cultures (Fig 1E; Suppl Fig1A,B). Flow cytometry for adult and fetal hemoglobin at basophilic/orthochromatic stage (96 hours of differentiation) showed a population expressing only adult and a population expressing both fetal and adult hemoglobin (Fig 1F). HbF levels were increased in HPFH compared to control cultures and recapitulated relative expression levels of HbA and HbF (Fig 1F; Table 1). The geometric mean fluorescence intensity (GeoMFI) correlated to the percentage of HbF measured by HPLC on both cultured erythroid cells (Fig 1G) and erythrocytes (Suppl Fig 1C), but no significant increase in HbF GeoMFI within the HbF/HbA double-positive cells was observed (Suppl Fig 1D). This shows that higher HbF results from more cells expressing HbF as opposed to increased HbF expression per cell (Fig. 1F).

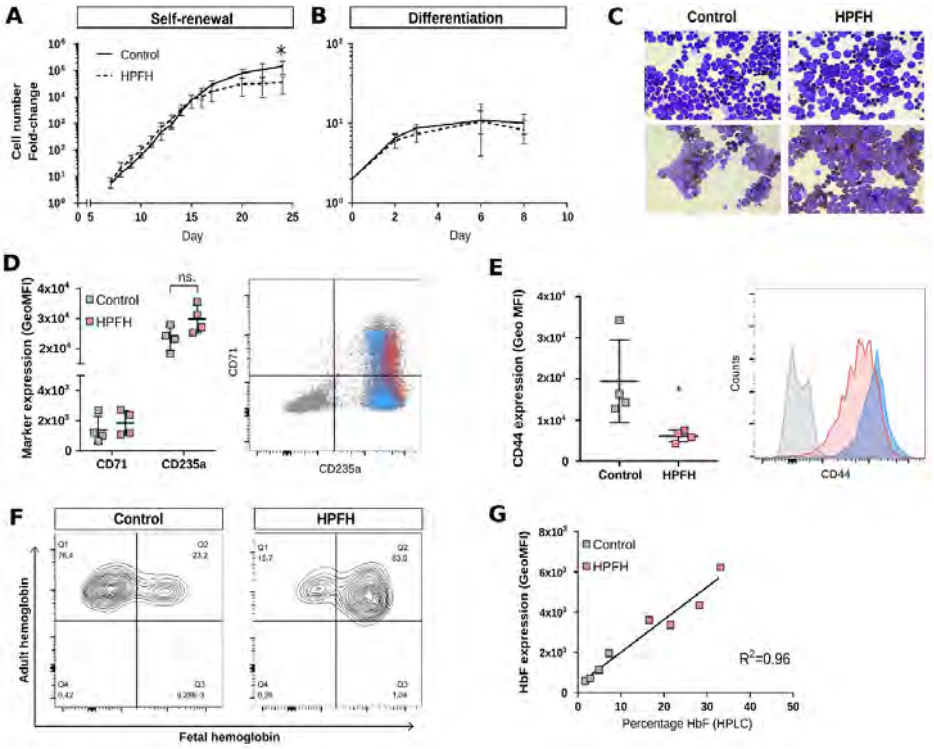


Figure 1. Cultured erythroblasts from *KLF1* p.K288X carriers reflect the HPFH phenotype. (A) Cell counts from expansion stage of erythroid cultures started from PBMCs. Final cell count after 24 days $1.4x10^{11}$± $7.3\ x10^{10}$ (controls) vs. $3.6x10^{10}$±$2.2x10^{10}$ (HPFH), mean ±sd, N= 3-4 *= p<0.05 Students T-test. **(B)** Cell counts from terminal erythroid differentiation stage of culture. **(C)** Cell morphology at two sampling time points. Top: erythroblast after 10 days of expansion from PBMCs. Bottom: poly/orthochromatic erythroblasts after 48h of terminal differentiation. **(D)** Surface expression of differentiation markers transferrin receptor (CD71) and glycophorin A (CD235a). Left: quantification of Geometric Mean Fluorescent Intensity (GeoMFI) after correction for isotype control. Right: representative dot plot for control sample in blue and HPFH in red, isotype in grey. **(E)** Surface expression for CD44 (Indian blood group). Left: quantification of GeoMFI after correction for isotype control. Right: representative histograms for control sample in blue and HPFH in red, isotype in grey. **(F)** Flow cytometry after intracellular staining for hemoglobin types. Two representative contour plots are shown. **(G)** Correlation between HbF levels determined by HPLC and flow cytometry.

As reported before, control and HPFH cultures showed higher HbF levels compared to erythrocytes obtained from the same individuals (Fig 1G; Suppl Fig1C; Table 1). For three out of four HPFH individuals, the HbF levels ranked identical for erythrocytes and cultured erythroid cells (Table 1). For HPFH individual 2, HbF ranked lowest in erythrocytes (1.3% HbF) but second highest in cultured cells (28% HbF; Table 1). Total hemoglobinization was comparable between control and HPFH samples (Suppl Fig 1B), indicating that increased HbF involved specific up-regulation of *HBG1/2* expression and was not the consequence

of overall lower HbA levels. Collectively, the reduced CD44 expression along with the observed ranking in HbF levels show that the erythroid cultures recapitulate the *KLF1* haploinsufficiency phenotypes.

KLF1 haploinsufficiency results in deregulated gene expression

Next, we asked which other genes were affected by *KLF1* haploinsufficiency. We isolated RNA from HPFH and control samples at the pro-erythroblast/basophilic stage (T0) and at the polychromatic/orthochromatic stage after 48 hours of differentiation (T48, Fig 1C). RNA samples were poly-A enriched prior to mRNA sequencing. Principal component analysis (PCA) separated the two sampling time points (first component) and HPFH from control (second component, Fig 2A; more pronounced during differentiation). This was also reflected in the correlation matrix ($\rho < 0.93$ and $\rho > 0.95$, Suppl Fig 2A). At T48, out of 9135 unique RNAs we identified 344 differentially expressed genes (3.7%; FDR<0.05) between HPFH and control samples with 173 up- and 171 downregulated genes (Suppl. Table 1A).

Table 1: HPLC analysis of HbF levels in erythrocytes and cultured cells.

SAMPLE NO.	INDIVIDUAL DENOMINATOR	KLF1 GENOTYPE	HBF IN VIVO	HBF CULTURE	HBA1 CULTURE	ATAC-SEQ	RNA-SEQ
1	FamF*, m2	wt/p.K288X	12.32	33.07	46.41	Yes	Yes
2	FamD, m8	wt/p.K288X	*1.34*	28.13	54.15	Yes	Yes
3	FamF, m5	wt/p.K288X	*7.26*	21.52	61.43	Yes	Yes
4	FamF, m4	wt/p.K288X	*3.8*	16.59	62.36	Yes	Yes
5	FamF, m1	wt/wt	*0.16*	2.53	85.55	No	Yes
6	FamF, m3	wt/wt	0.6	7.22	77.02	Yes	Yes
7	FamD, m7	wt/wt	*0.17*	5.02	82.77	Yes	Yes
8	unrelated, m9	wt/wt	*0.3*	2.18	86.44	Yes	Yes
9	unrelated	wt/wt	*< 0.5*	n.d.	n.d.	No	Yes
10	unrelated	wt/wt	*< 0.5*	n.d.	n.d.	No	Yes

*See Borg et al. [4] for description of FamF.

In line with increased HbF expression, HPFH erythroblasts expressed higher γ-globin mRNA levels (*HBG1/2)* and lower β-globin mRNA levels (*HBB)* compared to controls (Fig 2B). Expression of known KLF1 target genes *BCAM* and *CD44* was reduced in HPFH samples (Fig 2C, Suppl Table 1B).

Of a set of known globin regulators only *BCL11A* was significantly lower expressed (1.8-fold) in HPFH compared to control cultures (FDR < 0.05; Fig 2B, Suppl Fig 2B). Of note, *KLF1* and *ZBTB7A,* the gene encoding LRF, were expressed approximately 1.6-fold lower, but these genes did not pass the correction threshold for multiple testing

(p-value < 0.001, FDR > 0.05). RT-qPCR validated the different expression levels of *BCL11A, HBG1/2*, and *HBB*, but expression of *KLF1* and *ZBTB7A* genes was not significantly different comparing control to HPFH samples (Suppl Fig 2B). Heterogeneity within HPFH samples may explain the inability to detect significant expression differences for *KLF1* and *ZBTB7A*. Within individual samples, lower *KLF1* expression correlated with lower expression of *BCL11A* and *ZBTB7A* but higher expression of *HBG1/2*. (ρ=0.90 p<0.001, 0.67 p<0.05, -0.71 p<0.05 respectively, Fig 2D). In addition, *HBG1/2* expression levels correlated well with the percentage of HbF found in the cultures (see also Fig 1G). This suggests that the observed variability in HbF levels results from additional variation of *KLF1* expression in the haploinsufficiency setting of the HPFH samples. Many genes change expression at the onset of differentiation [32]. Indeed, out of the 344 differentially expressed genes between HPFH and control erythroblasts, 296 were also found to change upon differentiation induction (T0 and T48; FDR <0.05; Suppl Fig 2C). Importantly, differential expression between HPFH and control samples was not due to an altered response to induction of differentiation (Fig 1C,D).

To check for association with KLF1 DNA binding, the differentially expressed genes were referenced against a ChIP-seq dataset for KLF1 in HUDEP2 cells 5. Of the 344 deregulated genes, 186 displayed a KLF1 peak (54%; Fig 2E), most of which were located in the promoter or first intron (83%; Suppl Fig 2D). This was not enriched when compared to KLF1 peaks identified in the expressed genes of the total RNA-seq dataset (58% of genes; Fig 2E).

Next, datasets from BCL11A and LRF, ChIP-seq experiments[1] were included to check for association with KLF1 DNA binding. The differentially expressed genes were referenced against a ChIP-seq dataset for KLF1 in HUDEP2 cells [5]. Of the 344 deregulated genes, 186 displayed a KLF1 peak (54%; Fig 2E), most of which were located in the check for indirect effects of *KLF1* haploinsufficiency. Genes bound by KLF1, BCL11A and/or LRF represented 294 out of 344 (85%) differentially expressed genes (Fig 2E). Interestingly, the differentially expressed genes displayed significant enrichment for binding of either BCL11A (60% versus 50%, p<0.001) or LRF (65 versus 60%, p<0.01), and differentially expressed genes binding both BCL11A and LRF were also enriched (15% versus 9%, p<0.001). Collectively, this suggests that the 344 differentially expressed genes result from an interplay of regulation by KLF1, BCL11A and LRF. This supports the notion that *KLF1* haploinsufficiency partly acts through BCL11A and LRF [4,5]. To assess whether this is reflected in alterations in chromatin accessibility, Transposase Accessible Chromatin (ATAC) sequencing was used [25].

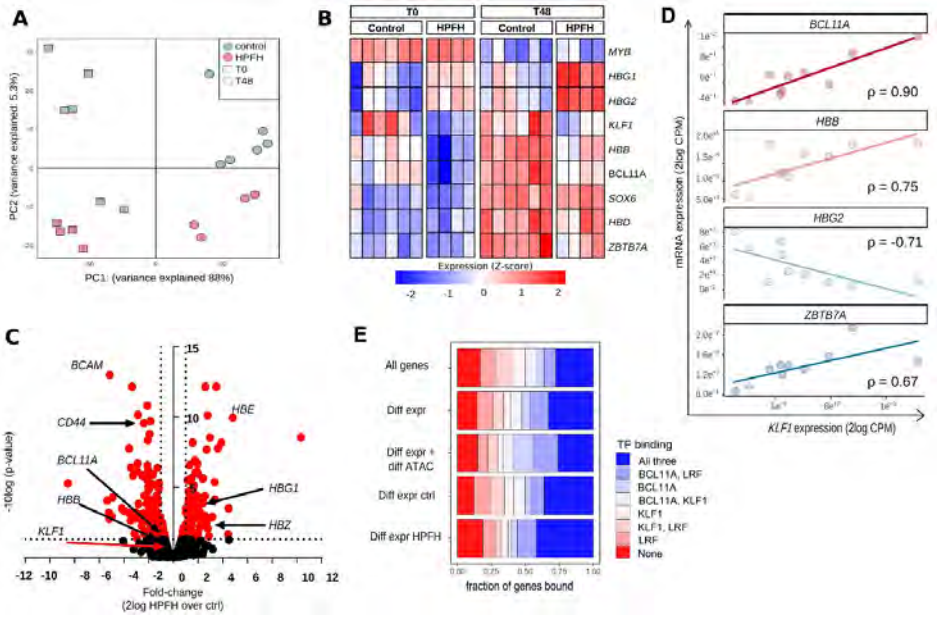


Figure 2. RNA-seq analysis of cultured erythroblasts. RNA-seq was performed on six controls and four HPFH individuals at the start of differentiation (T0) and after 48 hours of differentiation (T48). **(A)** Principal component (PC) analysis for 2000 most variable genes (blue: controls; red: HPFH individuals; squares: T0 samples, circles: T48 samples). **(B)** Heatmap for selected transcription factors and globins from *HBB* locus. Scale bar indicates Z-transformed 2log expression values (CPM). **(C)** Volcano plot with differentially expressed genes highlighted in red (FDR < 0.05). **(D)** mRNA expression for selected globins and regulators plotted against KLF1 expression. **(E)** Integration with ChIP datasets for KLF1, BCL11A and ZBTB7A. Bar plot shows fraction of genes with ChIP signal in gene sets from differential expression analysis on RNA-seq and ATAC-seq datasets. ρ: Pearson correlation coefficient. CPM: counts per million mapped reads.

Gene-proximal chromatin accessibility is largely unchanged in *KLF1* haploinsufficiency

To address if *KLF1* haploinsufficiency controls specific chromatin accessibility, for instance on the deregulated target genes, the HPFH and control cells were subjected to ATAC sequencing. Fragment length in ATAC-seq showed the expected periodicity [25], with more insert sizes at multiples of around 200bp, corresponding to the lack of one or more nucleosomes with linker DNA (Suppl Fig 3A, insert). Peaks were defined as present in HPFH- or control samples when detected in three or more replicates of that group. ATAC-seq peaks around 200bp in length were enriched just before the transcription start site (TSS; Fig 3A). To detect differences in the abundance of specific regulatory elements open regions were categorized, as belonging to proximal- or distal-regulatory elements, based on distance to the nearest TSS (ALTRE R-package) [33]. 14,654 putative regulatory elements were identified, with 76% found in both HPFH and control samples. 4% of all ATAC peaks

were unique to controls and 20% to HPFH samples (Fig 3B). Of these HPFH-specific elements the majority were identified at distal locations (>5000 bp from the TSS).

Next, the reads per called peak were quantified and annotated to the nearest TSS within a window of +5000bp upstream and -3000bp downstream, yielding 41,447 peaks. To determine the degree of variance in chromatin accessibility we performed PCA on these peaks. Samples from the two time points cluster along the first component, most likely reflecting the changes in accessible chromatin regions during differentiation, accounting for ~52% of total observed variance (Fig 3C). In the second component, the HPFH samples clustered separately from the controls (Fig 3C). However, this separation was lost upon restricting the analysis to the 2000 most variable ATAC-seq peaks (Suppl Fig3B). This suggests that *KLF1* haploinsufficiency only accounts for minor variations in chromatin accessibility in the 2000 most variable regions. The three HPFH members of the 'FamF' family (Table 1) grouped together, while the HPFH member of the 'FamD' family grouped closer to the control samples. This close grouping of HPFH 'FamF' relatives could indicate a familial genetic component as second explanation for variance between the most variable regions. At T48, 559 out of 41,447 peaks showed altered accessibility in HPFH samples compared to controls (EdgeR [34]; FDR < 0.05, Suppl Table 1C). 339 were more and 220 were less accessible (Suppl Fig 3C).

For comparison, the variation in chromatin state observed during erythroid differentiation identified 9338 differential peaks between T0 and T48. The minor differences in chromatin accessibility between HPFH and control samples against the larger scale rearrangements that occur during differentiation illustrate the limited effect of *KLF1* haploinsufficiency on chromatin accessibility. A notable exception was within the *HBB* locus with HPFH samples showing increased accessibility of the *HBG1* and *HBG2* promoter, but no differences for the other globin genes, notably *HBB*, nor the distal enhancers of the locus control region (LCR) (Fig 3D). Of note, while the HPFH samples showed increased accessibility of the *HBG1/2* promoters compared to control samples, a further subdivision of the HPFH samples into HbF levels above or below 28% (indicated as high vs. medium HbF) yielded differential accessibility of 102 peaks between the two groups, but these did not include the *HBG1/2* promoters (Fig 3D). This raises the question whether chromatin accessibility of other differentially expressed genes is affected by *KLF1* haploinsufficiency.

Chromatin accessibility at KLF1 target genes

Whereas 96% of the 9139 expressed genes and 90% of differentially expressed genes identified by RNA-seq also contained one or more ATAC peaks (Fig 4A), only 27 from the 344 differentially expressed genes displayed altered chromatin accessibility (Suppl Fig 3D). The majority of differentially expressed genes had more ATAC signal at T48 compared to T0 indicating increased chromatin accessibility upon differentiation, regardless of

sample genotype or direction of differential gene expression (Fig 4B). As discussed above, the *HBG1/2* promoters displayed differential accessibility, which was paired with differential RNA expression. In contrast, the *HBB* promoter, known to bind KLF1 [35,36], did not show differential accessibility paired with differential RNA expression (Fig 4C).

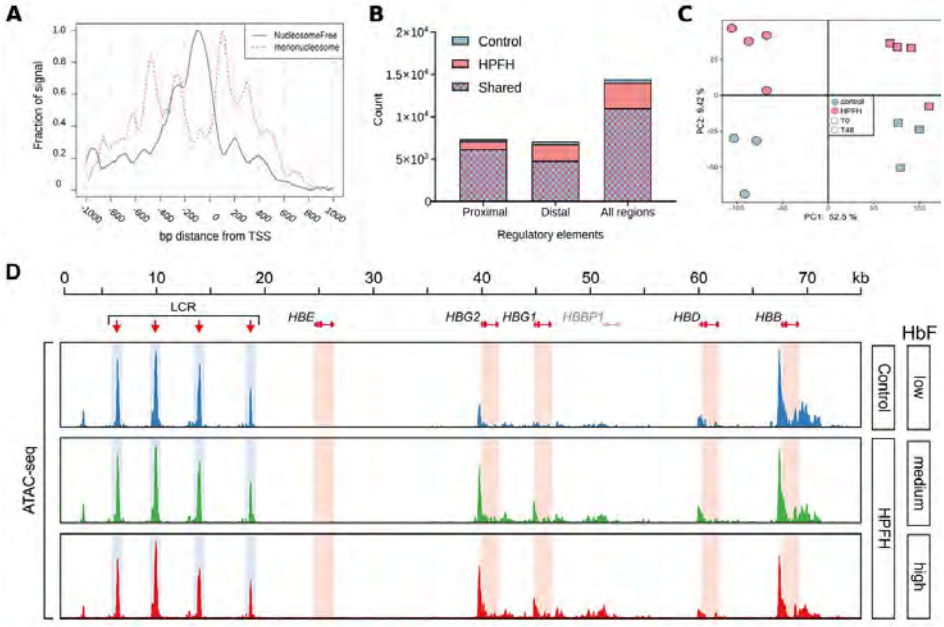


Figure 3. ATAC-seq analysis of cultured erythroblasts. ATAC-seq was performed on three controls and four HPFH individuals at the start of differentiation (T0) and after 48 hours of differentiation (T48). **(A)** Location of ATAC peaks relative to the transcription start site (TSS) **(B)** Quantification of regulatory regions determined by MACS2 and ALTRE **(C)** Principal component (PC) analysis on all ATAC peaks that change between T0 and T48 (blue: controls; red: HPFH individuals; circles: T0; squares:T48) **(D)** Genomic tracks showing ATAC signals across the *HBB* locus. Samples subdivided in control (n=3) and HPFH: medium HbF HPFH (n=2) , high HbF HPFH (n=2). LCR: locus control region.

Similar to *HBB*, KLF1 target genes that were previously identified by ChIP [5] such as *BCL11A* and *CARM1* showed differential RNA expression, but no differential chromatin accessibility. In contrast, *CD44* showed both reduced accessibility and expression (Fig 4C). However, here differential accessibility was found in the first intron, while KLF1 binding was reported at the TSS [5].

Together this suggests that *KLF1* haploinsufficiency is not resulting in an altered chromatin state of its target genes. Rather, the increased fraction of genes with BCL11A and LRF binding sites among the differentially expressed genes suggests that indirect regulation by KLF1 is an important mechanism in those cases where both gene expression and chromatin accessibility were altered.

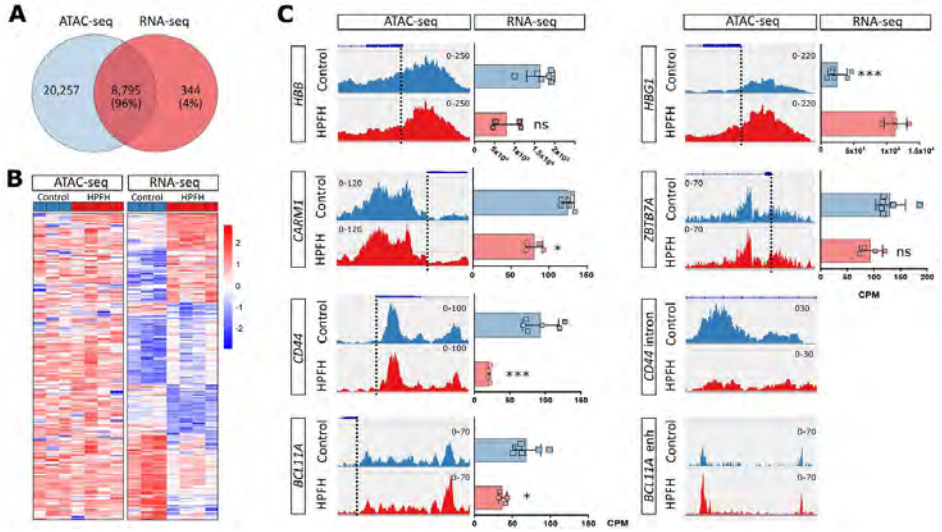


Figure 4. Integration of ATAC-seq and RNA-seq analysis. (A) Venn-diagram displaying overlap between ATAC-seq and mRNA-seq. Counts represent the number of individual genes above minimum threshold for mRNA expression (>3 CPM) or open chromatin (>1 CPM). **(B)** Heatmaps showing normalized mRNA expression and chromatin accessibility at T48 for differentially expressed genes between control and HPFH samples. **(C)** Chromatin accessibility and mRNA expression of differentially expressed genes. For each gene the genome track shows chromatin accessibility. Dotted lines indicate transcription start sites. Bar plots indicate mRNA expression values. * FDR < 0.05, *** FDR < 0.001. CPM: counts per million mapped reads; enh = erythroid enhancer in intron 2 of the *BCL11A* gene.

Specific SNPs within the promoter of *KLF1* influence promoter activity

Next, we asked if variation in residual *KLF1* expression might be linked to variability in HbF levels between HPFH cultures. Grouping RNA-seq data according to HbF levels, an inverse correlation between *HBG1/2* and *KLF1* expression was observed (Fig 5A,B; see also Fig 2D), which was confirmed by RT-qPCR (Fig 5C). Similarly, previously identified KLF1 targets *ZBTB7A*, *PLEK2,* and *CARM1* showed lower mRNA expression in high HbF HPFH samples compared to medium HbF HPFH samples and the controls. In contrast *HBB* showed the opposite pattern while *SOX6* expression was constant (Fig 5A, Suppl Fig 4). Several SNPs within the promoter of *KLF1* that may influence expression have been annotated [37].

Sanger sequencing of the *KLF1* promoter showed that high HbF expressing cells carried a discriminating SNP, rs3817621, located at -188 (Fig 5D). SNP rs112943515 further discriminated between HPFH and non-HPFH samples. To test the effect of rs3817621 and rs112943515, luciferase reporter experiments using the *KLF1* promoter were performed in K562 cells. Promoters containing minor alleles rs3817621 or

rs112943515 showed reduced luciferase expression compared to the wildtype promoter, with rs3817621 having the strongest effect (Fig 5E).

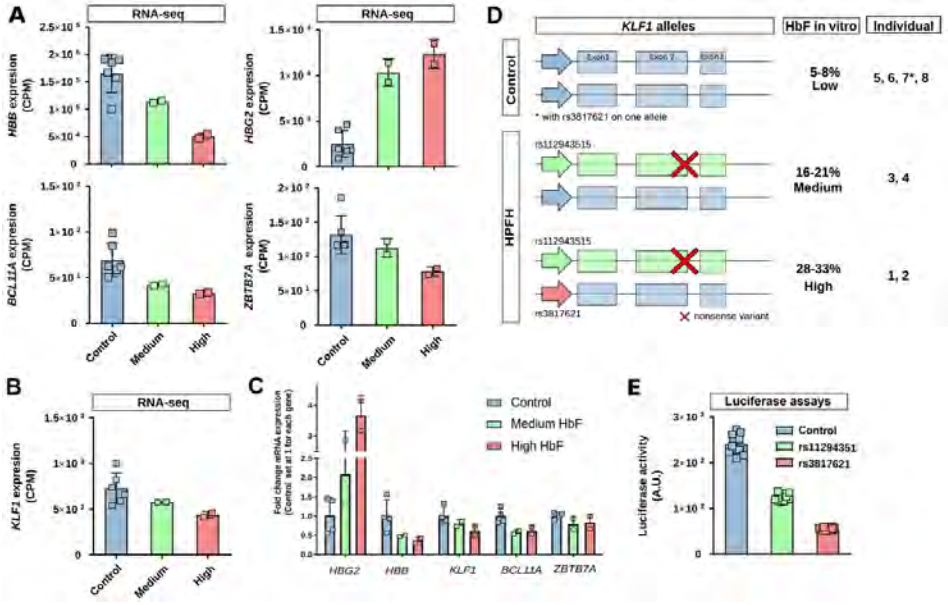


Figure 5. Variant *KLF1* alleles in Maltese HPFH individuals affect promoter activity. (A) Expression of selected genes assessed by RNA-seq plotted in relation to HbF levels. **(B)** Expression of *KLF1* gene assessed by RNA-seq plotted in relation to HbF levels. **(C)** RT-qPCR analysis stratified by HbF levels. The average expression level in the control samples was normalized to 1. **(D)** Schematic representation of genomic layout of the *KLF1* alleles present in the individuals studied here. A promoter variant, rs381762, is shared in HPFH samples with high HbF in cultured cells (individuals 1, 2) and one of the control individuals (7). The red cross indicates the *KLF1* p.K288X mutation. **(E)** Luciferase assays in K562 cells with reporter plasmids driven by *KLF1* promoter variants. A.U.: arbitrary units of luciferase activity, corrected for transfection efficiency.

This indicates that HPFH cells with higher HbF expression carry a *KLF1* promoter variant that renders the promoter less active. This particular variant is found in two HPFH individuals and one of the controls (Fig 5D). All of these individuals have higher HbF levels compared to their relatives (Table 1). Since rs3817621 is linked to the functional *KLF1* allele, our data indicate that, on top of the effect of haploinsufficiency, variation in wildtype *KLF1* expression levels affects HbF levels via modulation of *BCL11A* and *ZBTB7A* expression.

Discussion

Reduced expression of KLF1 decreases expression of the *HBG1/2* repressors BCL11A and LRF, thus enhancing HbF expression [2,4,5]. There are, however, considerable differences in HbF levels between carriers of the *KLF1* p.K288X haploinsufficiency allele, prompting us to find an explanation for this variability. We found that a SNP (rs3817621) in the promoter of the remaining wildtype *KLF1* allele, reducing *KLF1* promoter activity, correlated with high HbF levels. We surmise that subtle variations in expression of the remaining wildtype *KLF1* allele affect HbF levels through differential expression of the two major *HBG1/2* repressors BCL11A and LRF. This is consistent with previous work establishing that expression of these two repressors is under direct control of KLF1 [3-5,8]. This SNP (minor allele frequency 32.5%) is not associated with the typical reduction in Lutheran and Indian blood groups or altered globin expression caused by *KLF1* variants [16]. We propose that the HbF level is the resultant of the compound effect of the minor rs3817621 SNP on the wildtype *KLF1* allele and the p.K288X variant on the other *KLF1* allele. These individuals should therefore be labeled as compound heterozygote carriers of *KLF1* variants, as their phenotype extends beyond that expected for *KLF1* haploinsufficiency.

Increased chromatin accessibility of the HBG1/2 promoters

In erythroid progenitors from *KLF1* p.K288X-bearing individuals, the HPFH phenotype is associated with increased chromatin *HBG1/2* promoter accessibility. Notably, this increased *HBG1/2* promoter accessibility is an exception, as this is not observed for promoters of the vast majority of the >300 deregulated genes between control and HPFH erythroid cells. This overwhelming absence of locally altered chromatin states in HPFH erythroblasts shows there is no major direct impact of reduced KLF1 levels on chromatin accessibility of its target genes. This is supported by cross reference of the genes with differential expression and differential chromatin accessibility against datasets from previous ChIP-seq experiments [2,5]. These genes showed enrichment for binding sites of BCL11A and LRF but not for KLF1. In the context of HbF regulation, this is consistent with the hypothesis that reduced KLF1 expression results in downregulation of its target genes, including genes encoding the *HBG1/2* repressors LRF and BCL11A. These repressors directly bind and repress the *HBG1/2* promoters in adult erythroid cells [1-3]. Presumably, diminished expression of LRF and BCL11A results in reduced binding to the *HBG1/2* promoters. Complementary reduced interaction of the *HBB* promoter with the LCR, previously described in the context of *Klf1 null* mice [35], favors activation of the accessible *HBG1/2* promoters. We propose that this dual role of KLF1, along with micro-variation in KLF1 expression between HPFH individuals, explains a significant part of the variable penetrance of HbF expression levels associated with *KLF1* haploinsufficiency.

KLF1 haploinsufficiency does not affect erythroid chromatin accessibility broadly

In addition to the *HBG1/2* promoters there is a limited set of ATAC peaks, consisting of less than 2% of the total, that also show differential accessibility in the HPFH cells. Although most of the genes associated with these regions do not display differential expression, it remains possible that these regions act as distal regulatory elements for more remotely located genes. Since erythroid differentiation is accompanied by broad scale chromatin reorganization (reviewed by Ji et al. [38]), the limited set of regions changed during *KLF1* haploinsufficiency may potentially interfere with differentiation. However, cell morphology and expression of differentiation markers are largely unperturbed in *KLF1* haploinsufficiency indicating that the remaining KLF1 expression is sufficient to keep the erythroid chromatin state compatible with terminal erythroid differentiation. Reduced expression of the nuclear exportin XPO7, which in turn affects global chromatin condensation [39,40], provides an alternative explanation for the increased chromatin accessibility. Similar to the *HBB* locus, regulation of *Xpo7* expression by KLF1 ensues via a chromatin looping mechanism [40]. Possibly, this mode of regulation renders genes particularly susceptible to the effects of reduced KLF1 levels. This point deserves future investigation of long-range interactions using chromatin conformation capture techniques.

Phenotypic variability associated with haploinsufficiency for transcription regulators

The notion of variable penetrance resulting from mutated alleles was addressed in a study on *C. elegans*. Expression of individual transcripts was counted in a small three-layer regulatory network [41]. Mutations to the initiating node of the network caused highly variable expression of the intermediate nodes and resulted in bimodal expression of the most downstream genes. For different mutant alleles the threshold required for the downstream effects shifted, which was proposed as an explanation for incomplete penetrance. In the context of Maltese HPFH, we propose that promoter variants affecting expression of the wildtype *KLF1* allele alter indirect and direct regulation of globin gene expression from the *HBB* locus. Thus, the variable penetrance of high HbF expression in *KLF1* haploinsufficiency fits very well with the idea that micro-variation in expression of a critical factor can be amplified in a regulatory network. In summary, we propose that micro-variations in KLF1 levels are a major source of variable HbF expression observed in *KLF1* haploinsufficiency. Our findings have broader implications for understanding the phenotypic variability associated with mutations in other hematopoietic transcription factors, such as RUNX1, EP300 and GATA2. For example, variations in mono-allelic *GATA2* expression were recently shown to reduce penetrance in patients with hereditary *GATA2*-mutated MDS/AML [42]. Similarly, micro-variations in *RUNX1* and *EP300* expression might

help explain the range of phenotypes observed in response to mutations in these factors [43-45]. These cases of altered penetrance in response to an otherwise neutral SNP illustrate the importance of extended genotypic screening of transcription factor loci, in order to improve prognostic and therapeutic strategies in the event of haploinsufficiency.

Acknowledgments

Research in our laboratories is funded by the Landsteiner Foundation for Blood Transfusion Research (LSBR 1040 and 1627), the Netherlands Organization for Scientific Research (ZonMw-TOP 40-00812-98-12128; ZonMw-TAS 40-41400-98-1327), the Netherlands Genomics Initiative (NGI Zenith 93511036), and EU fp7 Specific Cooperation Research Project THALAMOSS (306201).

References

1. Liu N, Hargreaves VV, Zhu Q, et al. Direct Promoter Repression by BCL11A Controls the Fetal to Adult Hemoglobin Switch. Cell. 2018;173(2):430-442.e417.
2. Martyn GE, Wienert B, Yang L, et al. Natural regulatory mutations elevate the fetal globin gene via disruption of BCL11A or ZBTB7A binding. Nat Genet. 2018;50(4):498-503.
3. Masuda T, Wang X, Maeda M, et al. Transcription factors LRF and BCL11A independently repress expression of fetal hemoglobin. Science. 2016;351(6270):285-289.
4. Borg J, Papadopoulos P, Georgitsi M, et al. Haploinsufficiency for the erythroid transcription factor KLF1 causes hereditary persistence of fetal hemoglobin. Nat Genet. 2010;42(9):801-805.
5. Norton LJ, Funnell APW, Burdach J, et al. KLF1 directly activates expression of the novel fetal globin repressor ZBTB7A/LRF in erythroid cells. Blood Adv. 2017;1(11):685-692.
6. Zhou D, Liu K, Sun CW, Pawlik KM, Townes TM. KLF1 regulates BCL11A expression and gamma- to beta-globin gene switching. Nat Genet. 2010;42(9):742-744.
7. Borg J, Patrinos GP, Felice AE, Philipsen S. Erythroid phenotypes associated with KLF1 mutations. Haematologica. 2011;96(5):635-638.
8. Sankaran VG, Menne TF, Xu J, et al. Human fetal hemoglobin expression is regulated by the developmental stage-specific repressor BCL11A. Science. 2008;322(5909):1839-1842.
9. Esteghamat F, Gillemans N, Bilic I, et al. Erythropoiesis and globin switching in compound Klf1::Bcl11a mutant mice. Blood. 2013;121(13):2553-2562.
10. Xu J, Peng C, Sankaran VG, et al. Correction of sickle cell disease in adult mice by interference with fetal hemoglobin silencing. Science. 2011;334(6058):993-996.
11. Martin-Subero JI, Gesk S, Harder L, et al. Recurrent involvement of the REL and BCL11A loci in classical Hodgkin lymphoma. Blood. 2002;99(4):1474-1477.
12. Tanaka Y, Nagai Y, Mori M, et al. Multiple granulocytic sarcomas in essential thrombocythemia. Int J Hematol. 2006;84(5):413-416.
13. Nuez B, Michalovich D, Bygrave A, Ploemacher R, Grosveld F. Defective haematopoiesis in fetal liver resulting from inactivation of the EKLF gene. Nature. 1995;375(6529):316-318.
14. Perkins AC, Sharpe AH, Orkin SH. Lethal beta-thalassaemia in mice lacking the erythroid CACCC-transcription factor EKLF. Nature. 1995;375(6529):318-322.
15. Maeda T, Ito K, Merghoub T, et al. LRF is an essential downstream target of GATA1 in erythroid development and regulates BIM-dependent apoptosis. Dev Cell. 2009;17(4):527-540.
16. Perkins A, Xu X, Higgs DR, et al. Kruppeling erythropoiesis: an unexpected broad spectrum of human red blood cell disorders due to KLF1 variants. Blood. 2016;127(15):1856-1862.
17. Singleton BK, Burton NM, Green C, Brady RL, Anstee DJ. Mutations in EKLF/KLF1 form the molecular basis of the rare blood group In(Lu) phenotype. Blood. 2008;112(5):2081-2088.
18. Viprakasit V, Ekwattanakit S, Riolueang S, et al. Mutations in Kruppel-like factor 1 cause transfusion-dependent hemolytic anemia and persistence of embryonic globin gene expression. Blood. 2014;123(10):1586-1595.
19. Perseu L, Satta S, Moi P, et al. KLF1 gene mutations cause borderline HbA(2). Blood. 2011;118(16):4454-4458.
20. Satta S, Perseu L, Moi P, et al. Compound heterozygosity for KLF1 mutations associated with remarkable increase of fetal hemoglobin and red cell protoporphyrin. Haematologica. 2011;96(5):767-770.

21. Arnaud L, Saison C, Helias V, et al. A dominant mutation in the gene encoding the erythroid transcription factor KLF1 causes a congenital dyserythropoietic anemia. Am J Hum Genet. 2010;87(5):721-727.
22. Jaffray JA, Mitchell WB, Gnanapragasam MN, et al. Erythroid transcription factor EKLF/KLF1 mutation causing congenital dyserythropoietic anemia type IV in a patient of Taiwanese origin: review of all reported cases and development of a clinical diagnostic paradigm. Blood Cells Mol Dis. 2013;51(2):71-75.
23. Ortolano R, Forouhar M, Warwick A, Harper D. A Case of Congenital Dyserythropoeitic Anemia Type IV Caused by E325K Mutation in Erythroid Transcription Factor KLF1. J Pediatr Hematol Oncol. 2018;40(6):e389-e391.
24. Ravindranath Y, Johnson RM, Goyette G, Buck S, Gadgeel M, Gallagher PG. KLF1 E325K-associated Congenital Dyserythropoietic Anemia Type IV: Insights Into the Variable Clinical Severity. J Pediatr Hematol Oncol. 2018;40(6):e405-e409.
25. Buenrostro JD, Giresi PG, Zaba LC, Chang HY, Greenleaf WJ. Transposition of native chromatin for fast and sensitive epigenomic profiling of open chromatin, DNA-binding proteins and nucleosome position. Nat Methods. 2013;10(12):1213-1218.
26. van den Akker E, Satchwell TJ, Pellegrin S, Daniels G, Toye AM. The majority of the in vitro erythroid expansion potential resides in CD34- cells, outweighing the contribution of CD34+ cells and significantly increasing the erythroblast yield from peripheral blood samples. Haematologica. 2010.
27. Heshusius S, Heideveld E, Burger P, et al. Large-scale in-vitro production of red blood cells from human peripheral blood mononuclear cells. bioRxiv. 2019.
28. Bakker WJ, Blazquez-Domingo M, Kolbus A, et al. FoxO3a regulates erythroid differentiation and induces BTG1, an activator of protein arginine methyl transferase 1. J Cell Biol. 2004;164(2):175-184.
29. Beug H, Doederlein G, Freudenstein C, Graf T. Erythroblast cell lines transformed by a temperature-sensitive mutant of avian erythroblastosis virus: a model system to study erythroid differentiation in vitro. J Cell Physiol Suppl. 1982;1:195-207.
30. van Zwieten R, van Oirschot BA, Veldthuis M, et al. Partial pyruvate kinase deficiency aggravates the phenotypic expression of band 3 deficiency in a family with hereditary spherocytosis. Am J Hematol. 2015;90(3):E35-39.
31. Leberbauer C, Boulme F, Unfried G, Huber J, Beug H, Mullner EW. Different steroids co-regulate long-term expansion versus terminal differentiation in primary human erythroid progenitors. Blood. 2005;105(1):85-94.
32. An X, Schulz VP, Li J, et al. Global transcriptome analyses of human and murine terminal erythroid differentiation. Blood. 2014;123(22):3466-3477.
33. Baskin E, Farouni R, Mathe EA. ALTRE: workflow for defining ALTered regulatory elements using chromatin accessibility data. Bioinformatics. 2017;33(16):2609.
34. Robinson MD, McCarthy DJ, Smyth GK. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics. 2010;26(1):139-140.
35. Drissen R, Palstra RJ, Gillemans N, et al. The active spatial organization of the beta-globin locus requires the transcription factor EKLF. Genes Dev. 2004;18(20):2485-2490.
36. Miller IJ, Bieker JJ. A novel, erythroid cell-specific murine transcription factor that binds to the CACCC element and is related to the Kruppel family of nuclear proteins. Mol Cell Biol. 1993;13(5):2776-2786.
37. Gnanapragasam MN, Crispino JD, Ali AM, et al. Survey and evaluation of mutations in the human KLF1 transcription unit. Sci Rep. 2018;8(1):6587.
38. Ji P, Murata-Hori M, Lodish HF. Formation of mammalian erythrocytes: chromatin condensation and enucleation. Trends Cell Biol. 2011;21(7):409-415.

39. Hattangadi SM, Martinez-Morilla S, Patterson HC, et al. Histones to the cytosol: exportin 7 is essential for normal terminal erythroid nuclear maturation. Blood. 2014;124(12):1931-1940.
40. Cantú I, van de Werken HJ, Gillemans N, et al. The mouse Klf1 Nan variant impairs nuclear condensation and erythroid maturation. 2018:477059.
41. Raj A, Rifkin SA, Andersen E, van Oudenaarden A. Variability in gene expression underlies incomplete penetrance. Nature. 2010;463(7283):913-918.
42. Al Seraihi AF, Rio-Machin A, Tawana K, et al. GATA2 monoallelic expression underlies reduced penetrance in inherited GATA2-mutated MDS/AML. Leukemia. 2018;32(11):2502-2507.
43. Preudhomme C, Renneville A, Bourdon V, et al. High frequency of RUNX1 biallelic alteration in acute myeloid leukemia secondary to familial platelet disorder. Blood. 2009;113(22):5583-5587.
44. Churpek JE, Bresnick EH. Transcription factor mutations as a cause of familial myeloid neoplasms. J Clin Invest. 2019;129(2):476-488.
45. Zimmermann N, Acosta AM, Kohlhase J, Bartsch O. Confirmation of EP300 gene mutations as a rare cause of Rubinstein-Taybi syndrome. Eur J Hum Genet. 2007;15(8):837-842.

Supplementary Materials and Methods

Erythroid cultures

Erythroblasts were expanded from PBMCs in proprietary Cellquin culture medium (Sanquin) supplemented with stem cell factor (SCF, CHO producer cell line supernatant equivalent to 100ng/ml), EPO (2 U/ml, ProSpec) and dexamethasone (Dex, 1μM, Sigma Aldrich) for ten days. Medium was refreshed every other day, cells were counted and maintained at 0.7-2*10^6 cells/ml (CASY Model TCC, Innovatis AG). Erythroblast differentiation was induced in medium with high EPO (10 U/ml) and 5% human plasma (Sanquin). At start of differentiation cells were seeded at 2*10^6 cells/ml, and the medium was partly replenished every other day for the first four days of differentiation. Two additional control blood samples were collected according to the same guidelines and used to generate erythroid cultures according to the same protocol. However, these samples were not stored in the Malta biobank and only used in the RNA sequencing analysis to serve as extra control samples.

Flow cytometry

For hemoglobin staining cells were fixed with 0.025% glutaraldehyde, 0.5% paraformaldehyde and permeabilized with 0.05% NP40 (all from Sigma Aldrich), followed by incubation with primary antibodies. Details of the antibodies are listed below. Samples were measured on a Becton-Dickinson LSR II flowcytometer with HTS plate reader (Becton-Dickinson Bioscience) and analyzed using FlowJo software (Becton-Dickinson, v10.6)

Hemoglobin content and cell morphology

Small aliquots of the cultures were removed and analyzed for hemoglobin content by photometry as described[1] . Cell morphology was analyzed in cytospins stained with histologic dyes and o-dianisidine[2] , using an Olympus BX40 microscope (40× objective, NA 0.65), and Leica digital camera (Leica DM-2500). [3]

HPLC

1*10^7 cells were collected and analyzed for Hb isoform expression by high-performance cationexchange liquid chromatography (HPLC) on Waters Alliance 2690 equipment. The column was purchased from PolyLC[3] .

Real-time quantitative PCR

cDNA was synthesized with the QuantiTect reverse transcription kit (QIAGEN, #205313) according to the manufacturers' instructions. Quantitative PCR was performed on a StepOnePlus Real-Time PCR system (ThermoFisher Scientific, #4376600) with Power

SYBR green master mix (ThermoFisher Scientific, #4367659) and 1µM primers in a final volume of 20µl. The following program was used: 10min at 95 °C, followed by 40 cycles x 15sec at 95 °C and 1min at 60 °C. Primers used in this study are listed in Supplementary Methods. Ct-values were normalized against housekeeping genes 18S and HPRT. Graphpad Prism V7.04 (GraphPad Software) was used for statistic testing and visualization of fold-change mRNA expression. Primers used for RT-qPCR are listed below.

Supplementary Table 1. Antibodies for Flowcytometry

Antibody	Dilution	Fluorochrome	Producer	CATALOG #
CD71/TRANSFERRIN RECEPTOR	1:200	Vio-blue 421	Miltenyi	130-101-631
CD235/GLYCOPHORIN A	1:2500	PE	Acris	DM066P
CD44 /BCAM	1:200	FITC	eBioscience	11-0441-81
HBA/ADULT HEMOGLOBIN	1:1000	PE	Santa Cruz	Sc-21757
HBF /FETAL HEMOGLOBIN	1:1000	APC	Invitrogen	HBF-1
IGG1 K CONTROL	1:200	PB	Biolegend	400172
IGG1 K CONTROL	1:200	FITC	Biolegend	M1453
IGG1 CONTROL	1:1000	PE	Diaclone	857.072.010
IGG1 K CONTROL	1:1000	APC	eBioscience	P3.6.2.8.1

Supplementary table 2. Primers for RT-qPCR

PRIMER	SEQUENCE (5' – 3')
HBB FW	aca gcc acc act ttc tga t
HBB RV	agc tgc act gtg aca agc tg
HBG1/2 FW	aaa cgg tca cca gca cat tt
HBG1/2 RV	gaa ggt gct gac ttc ctt gg
KLF1 FW	tcc cac ctg aag gcg cat ct
KLF1 RV	cag ctg gtc aga gcg cga aa
BCL11A FW	gcc cca aac agg aac aca ta
BCL11A RV	ttg caa gag aaa cca tgc ac
ZBTB7A FW	gct gcg gca cct tta aga ca
ZBTB7A RV	gcg tca ctg ccc cta ca
18S FW	cac ggc cgg tac agt gaa ac
18S RV AGA	gga gcg agc gac caa
G3PH FW	cat cac gcc aca gtt tcc
G3PH RV	tcc cat cac cat ctt cca

RNA library preparation for sequencing (RNA-Seq)

RNA-Seq library was prepared for analysis according to the Illumina TruSeq stranded mRNA protocol (www.illumina.com). Briefly, 200 ng of total RNA was purified using poly-T oligoattached magnetic beads to end up with poly-A containing mRNA. The poly-A tailed mRNA was fragmented and cDNA was synthesized using SuperScript II and random primers in the presence of Actinomycin D. cDNA fragments were end repaired, purified with AMPure XP beads, A-tailed using Klenow exo-enzyme in the presence of dATP. Paired end adapters with dual index (Illumina) were ligated to the A-tailed cDNA fragments and purified using AMPure 4 XP beads. The resulting adapter-modified cDNA fragments were enriched by PCR using Phusion polymerase as followed: 30 s at 98°C, 15 cycles of (10 s at 98°C, 30 s at 60°C, 30 s at 72°C), 5 min at 72°C. PCR products were purified using AMPure XP beads and eluted in 30 µl of resuspension buffer. One microliter was loaded on an Agilent Technologies 2100 Bioanalyzer using a DNA 1000 assay to determine the library concentration and for quality check.

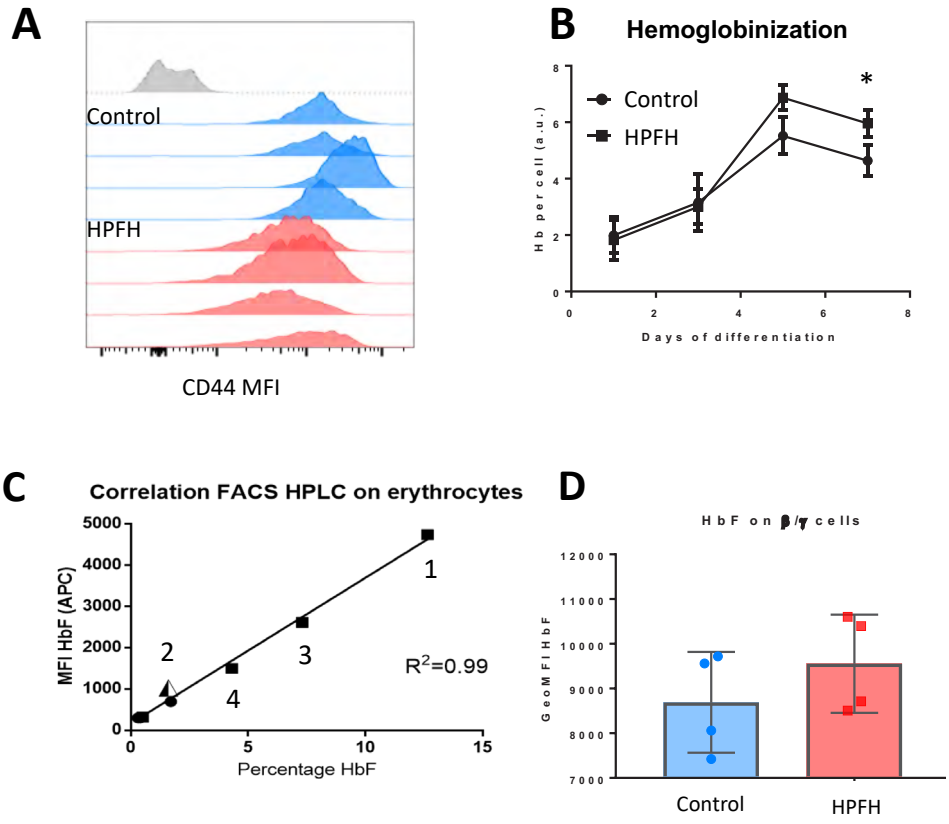
Sample preparation for ATAC-Seq

The Assay for Transposase Accessible Chromatin coupled to high-throughput sequencing (ATAC-seq) was performed as previously described 4 with slight modifications. Briefly, approximately 50,000 single cells were resuspended in 50 µl of cold lysis buffer (10 mM TrisHCl (pH 7.4) 10 mM NaCl, 3 mM MgCl2, 0.1% IGEPAL CA-630). Immediately after lysis, nuclei were centrifuged at 500xg for 10 min at 4°C. Nuclei were re-suspended in 25 µl 2x TD buffer (Illumina), 2.5 µl TDE1 (transposase, Illumina) and 20 µl nuclease-free H2O, and incubated for 30 min at 37°C. Afterward, the sample was purified using the MinElute PCR Purification Kit (QIAGEN) according to the manufacturer's protocol and eluted with 10 µl elution buffer. Samples were amplified using the NEBNext High Fidelity PCR master mix (New England Biolabs) and afterwards purified using a MinElute PCR purification kit (Qiagen) according to the manufacturer's instructions. One microliter was loaded on an Agilent Technologies 2100 Bioanalyzer using a DNA 1000 assay to determine the library concentration and for quality check.

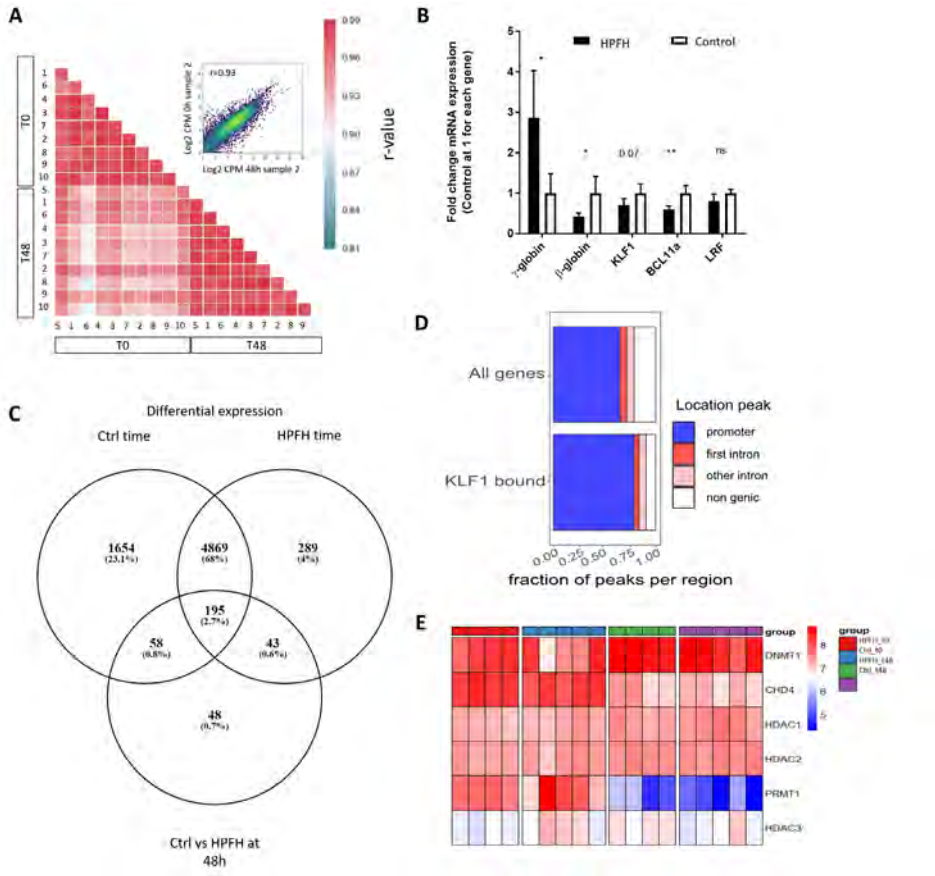
Bridge amplification and Sequencing by synthesis for RNA-Seq and ATAC-Seq

Cluster generation was performed according to the Illumina TruSeq SR Rapid Cluster kit v2 (cBot) Reagents Preparation Guide (www.illumina.com). Briefly, for sequencing libraries were pooled together to get a stock of 10 nM. One microliter of the 10 nM stock was denaturated with NaOH, diluted to 10 pM and hybridized onto the flowcell. The hybridized products were sequentially amplified, linearized and end-blocked according to the Illumina Single Read Multiplex Sequencing user guide. After hybridization of the sequencing primer, sequencing-bysynthesis was performed using the HiSeq 2500 with

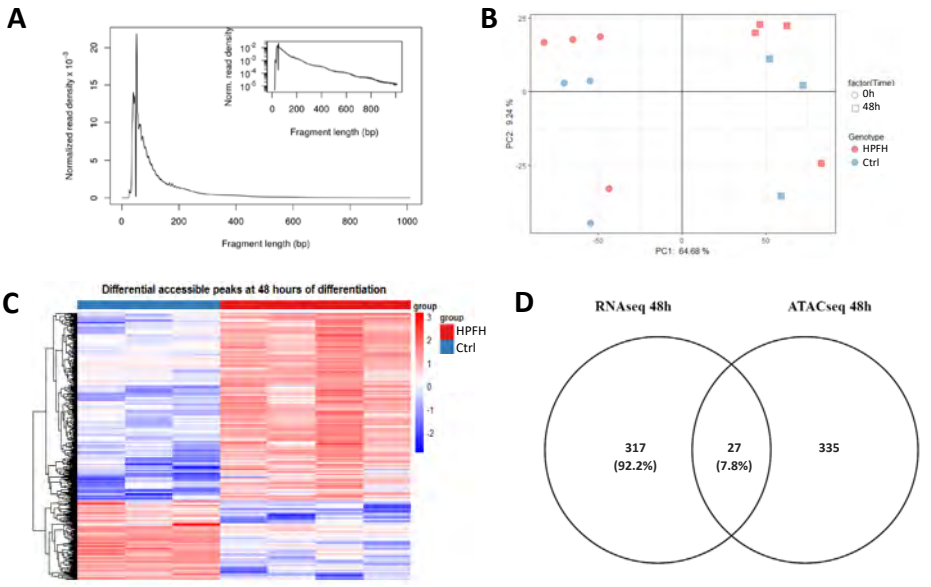
either a single read 50-cycle or paired end 50-cycle protocol followed by dual index sequencing, for RNA-Seq and ATAC-Seq respectively. Both RNA and ATAC libraries were sequenced to a depth of $15x10^6$ clusters. The raw data (fastq files) has been deposited in the European Nucleotide Archive (https://www.ebi.ac.uk/ena) under accession number PRJEB31712.



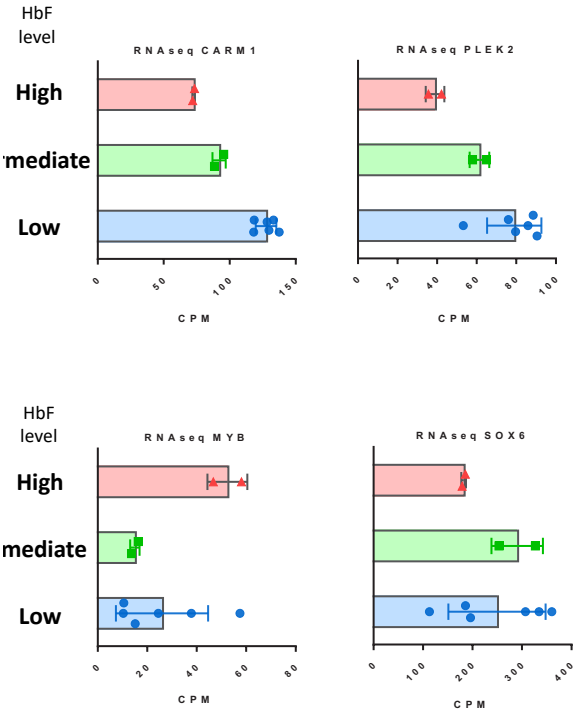
Supplementary Figure 1. Characteristics of cultured and *ex vivo* erythroid cells. (A) FACS plots for surface expression of CD44. **(B)** Lineplot showing hemoglobinization of cultured cells on subsequent days of culture **(C)** Line plot showing correlation between HPLC (x-axis) and flow cytometry mean fluorescence intensity measurements (y-axis). **(D)** Mean fluorescent intensity of cells with both HbA and HbF.



Supplementary Figure 2. Analysis of gene expression in cultured cells. (A) Correlation matrix between 2log-transformed count per million (CPM) values for all RNAseq samples. Label "Pro" indicates expansion cultures on EPO, SCF and Dex and "Dif" indicates differentiation conditions as indicated in material and methods. Inserts show 2log CPM expression plotted for a culture and differentiation sample from the same donor **(B)** RTqPCR validation of expression of selected globins and regulators of hemoglobin switching. **(C)** Venn diagram showing subdivision of RNA expression change over time for differentially expressed genes between HPFH and control samples. **(D)** Location of ChIP-seq peaks in differentially expressed genes between HPFH and control samples (184) compared to all annotated genes with a KLF1 peak (5298). **(E)** Heatmap showing 2log CPM expression values for previously identified epigenetic regulators of hemoglobin switching.

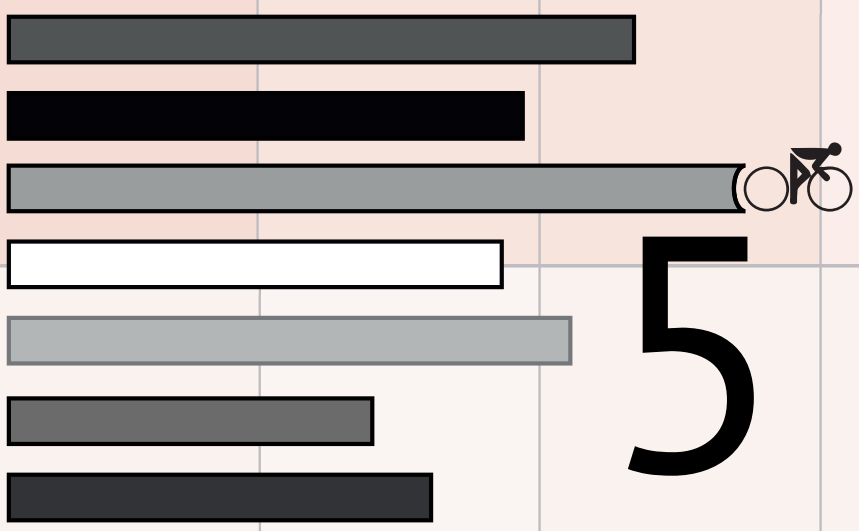


Supplementary Figure 3. Analysis of ATAC-seq data. (A) Density plots show insert size periodicity for a representative ATAC-seq sample. Insert shows log transformed insert size. Note the 200 bp periodicity. **(B)** Principal component (PC) analysis on 2000 most variable ATAC regions. **(C)** Heatmap showing z-transformed accessibility values for regions that are differentially accessible between HPFH and control cells at T48. **(D)** Venn-diagram displays the overlap between differential expression in RNAseq and differential accessibility for uniquely mapped peaks in ATACseq. Of 559 ATAC peaks differentially accessible peaks 362 mapped to unique genes.



Supplementary Figure 4. Expression analysis of selected genes. Bar plots of mRNA expression 2log CPM) from RNA-seq subdivided on γ-globin expression in cultured cells.


4



Chapter 5

Hemoglobin switching in mice carrying the *Klf1*^{Nan} variant

Anne Korporaal, Nynke Gillemans, Steven Heshusius, Ileana Cantú, Emile van den Akker, Thamar B. van Dijk, Marieke von Lindern[§] and Sjaak Philipsen[§#]
§ These authors share senior authorship

Haematologica, 2020 May 28; haematol.2019.239830.

Abstract

Haploinsufficiency for transcription factor KLF1 causes a variety of human erythroid phenotypes, such as the In(Lu) blood type, increased HbA2 levels, and hereditary persistence of fetal hemoglobin. Severe dominant congenital dyserythropoietic anemia IV (OMIM 613673) is associated with the KLF1 p.E325K variant. CDA-IV patients display ineffective erythropoiesis and hemolysis resulting in anemia, accompanied by persistent high levels of embryonic and fetal hemoglobin. The mouse *Nan* strain carries a variant in the orthologous residue, KLF1 p.E339D. *Klf1*^{Nan} causes dominant hemolytic anemia with many similarities to CDA-IV. Here we investigated the impact of *Klf1*^{Nan} on the developmental expression patterns of the endogenous α-like and β-like globins, and the human β-like globins carried on a *HBB* locus transgene. We observe that the switch from primitive, yolk sac-derived, erythropoiesis to definitive, fetal liver-derived, erythropoiesis is delayed in *Klf1*$^{wt/Nan}$ embryos. This is reflected in globin expression patterns measured between E12.5 and E14.5. Cultured *Klf1*$^{wt/Nan}$ E12.5 fetal liver cells display growth- and differentiation defects. These defects likely contribute to the delayed appearance of definitive erythrocytes in the circulation of *Klf1*$^{wt/Nan}$ embryos. After E14.5, expression of the embryonic/fetal globin genes is silenced rapidly. In adult *Klf1*$^{wt/Nan}$ animals, silencing of the embryonic/fetal globin genes is impeded, but only minute amounts are expressed. Thus, in contrast to human KLF1 p.E325K, mouse KLF1 p.E339D does not lead to persistent high levels of embryonic/fetal globins. Our results support the notion that KLF1 affects gene expression in a variant-specific manner, highlighting the necessity to characterize KLF1 variant-specific phenotypes of patients in detail.

Introduction

KLF1 is an erythroid-specific transcription factor with diverse and essential roles during terminal erythroid differentiation[1]. Cloned from mouse erythroleukemia cells it was initially called **e**rythroid **K**rüppel-**l**ike **f**actor (EKLF) in honor of its erythroid-specific expression and DNA-binding domain[2]. This domain is composed of three Cys_2-His_2 zinc fingers similar to those found in the *Drosophila* Krüppel transcription factor. KLF1 is the founding member of the KLF branch of the 27-strong SP/KLF family[3]. Despite the fact that other SP/KLF factors such as SP1, SP3, KLF2, KLF3 and KLF8 are abundantly expressed in erythroid cells[4, 5], gene inactivation in mice demonstrated that KLF1 is essential for definitive erythropoiesis[6, 7]. Specifically, activation of β-globin expression was strongly impaired leading to a severe β-thalassemia phenotype. Restoring globin chain imbalance with a human γ-globin transgene failed to rescue the embryonic lethality of KLF1 deficiency[8], indicating that other KLF1 target genes contribute to the erythroid defects. This was confirmed by genome-wide expression analyses which established that KLF1 is involved in virtually every aspect of terminal erythroid differentiation[9-12]. In humans, the first *KLF1* variants reported were identified as the molecular basis of the rare blood type In(Lu)[13]. In all cases one normal *KLF1* allele was present showing that, similar to *Klf1*$^{wt/ko}$ mice[6, 7], this is sufficient to sustain erythropoiesis. These observations were extended by the discovery that *KLF1* haploinsufficiency caused hereditary persistence of fetal hemoglobin (HPFH) in a Maltese pedigree[14]. To date, over 140 different *KLF1* variants have been reported, and these have been linked to a broad range of hitherto unrelated human red cell disorders[1]. In Sardinia, Thailand and southern China the frequency of *KLF1* variants reaches endemic proportions, e.g. 1.25% in southern China[15]. In these populations cases with compound heterozygosity for *KLF1* variants occur[16-18]. In such cases one allele invariably carries a missense variant which retains partial activity. *KLF1* compound heterozygotes display more pronounced phenotypes including HbF levels of >20%[16-18], persistence of embryonic globins[18], microcytic hypochromic anemia[16, 18], and pyruvate kinase deficiency[18]. One case of a KLF1 *null* neonate was reported[19]. This infant displayed severe non-spherocytic hemolytic anemia with elevated HbF levels (>70%). Thus, the vast majority of *KLF1* variants displays classical autosomal recessive inheritance, and *KLF1* haploinsufficiency is associated with mild erythroid phenotypes. The exception is a *KLF1* variant in which an ultra-conserved residue in the second zinc finger is affected. This KLF1 variant, p.E325K, causes congenital dyserythropoietic anemia, type IV (CDA-IV; OMIM613673)[20-22]. CDA-IV patients suffer from severe hemolytic anemia, splenomegaly, elevated HbF, iron overload, and dyserythropoiesis. Notably, the mouse **n**eonatal **an**emia (*Nan*) phenotype is caused by a variant of the orthologous residue in mouse KLF1, p.E339D[23, 24]. *Klf1*^{Nan} displays semi-dominant inheritance. *Klf1*$^{wt/Nan}$ animals suffer from hemolytic anemia while *Klf1*$^{Nan/Nan}$ embryos die around embryonic day 10.5 (E10.5) due

to failure of primitive erythropoiesis[23, 24]. This phenotype is more severe than that of *Klf1*$^{ko/ko}$ embryos which die around E14.5 due to failure of definitive erythropoiesis[6, 7]. During definitive erythropoiesis KLF1^{Nan} is thought to exert a dominant-negative effect on the function of wildtype KLF1[23, 24]. In primitive erythropoiesis, KLF1 and KLF2 have compensatory roles[25] and the early lethality of *Klf1*$^{Nan/Nan}$ embryos could therefore be due to interference of KLF1^{Nan} with normal KLF2 function. In addition, KLF1^{Nan} leads to aberrant gene expression which exerts negative effects on erythropoiesis[26, 27].

In human *HBB* locus transgenic mice, the fetal *HBG1/2* genes are expressed highly in primitive erythrocytes and early definitive erythrocytes. Switching to expression of the adult *HBB* gene occurs in the fetal liver between E12.5 and E14.5[28]. Given the profound deregulation of embryonic and fetal globin genes in CDA-IV patients, we investigated expression of the α-like and β-like globins in *Klf1*$^{wt/Nan}$ mice carrying a single-copy human *HBB* locus transgene[29] at embryonic, fetal and adult stages of development.

Methods

Animals

All animal studies were approved by the Erasmus MC Animal Ethics Committee. Mouse strains used were *Klf1*$^{wt/Nan}$ (C3H101H-*Klf1*^{Nan}/H[30]) crossed with PAC8.1 mice carrying a single-copy human *HBB* locus transgene (Tg(*HBB*)8.1Gvs[29]). For details see Supplementary Materials and Methods.

Culture of mouse erythroid progenitors

E12.5 fetal livers were disrupted and single cell suspensions were cultured as described[31]. For details see Supplementary Materials and Methods.

RNA isolation and RT-qPCR analyses

RNA was extracted using TRI reagent (Sigma-Aldrich). For details see Supplementary Materials and Methods.

Flow cytometry analysis

Single-cell suspensions were washed twice with PBS and resuspended in FACS buffer (PBS containing 1% (w/v) bovine serum albumin, 2 mM EDTA). Approximately 10^6 cells were incubated for 30 minutes at room temperature with the appropriate antibodies. Data were acquired on a Fortessa instrument (BD Biosciences), and analyzed with FlowJo software v10 (Tree Star). For details see Supplementary Materials and Methods.

Cell morphology

Cell morphology was analyzed using cytospins stained with May Grünwald-Giemsa (Medion Diagnostics) and O-dianisidine (Sigma-Aldrich)[32]. Pictures were taken with a BX40 microscope (40x objective, NA 0.65) equipped with a DP50 CCD camera and Viewfinder Lite 1.0 acquisition software (all Olympus).

Statistical tests

Statistical analysis of gene expression data was performed by using Mann-Whitney tests (GraphPad Prism). Excel 2010 was used to draw the graphs. Standard deviations and p-values <0.05 are displayed in the relevant figures (*).

Results

Expression patterns of the globin genes in Klf1$^{wt/Nan}$ mice

To assess the impact of KLF1^{Nan} on developmental regulation of globin gene expression, the *Klf1$^{wt/Nan}$* strain[30] was crossed with PAC8.1 mice carrying a single-copy human *HBB* locus transgene[29]. RT-qPCR analysis was performed to determine α-like and β-like globin expression in *Klf1$^{wt/wt::HBB}$* (control) and *Klf1$^{wt/Nan::HBB}$* (*Klf1$^{wt/Nan}$*) embryos at E11.5, E12.5, E13.5, E14.5 and E16.5. Primer pairs were designed to amplify the mouse and human embryonic, fetal and adult globin mRNAs specifically, aiming to minimize inter-globin and inter-species cross-reactivity. Of note, the human *HBG1* and *HBG2* genes, encoding Aγ- and Gγ-globin respectively, arose by a recent duplication event and expression of these genes is assessed by a single primer pair. The same is true for the mouse *Hba-a1/Hba-a2* genes, encoding α1- and α2-globin respectively, and *Hbb-b1/Hbb-b2* genes, encoding βmajor- and βminor-globin respectively (see Supplementary Information). Thus, we measured expression of mouse α-like globins (ζ and α1/2), mouse β-like globins (εy, βh1 and βmajor/minor) and human β-like globins (ε, Aγ/Gγ, and β). For reasons of clarity we will refer to mouse α-like globins as mζ and mα, mouse β-like globins as mεy, mβh1 and mβ, and human β-like globins as hε, hγ, and hβ in the remainder of this paper.

Globin gene expression patterns during development

The first erythroid cells are derived extra-embryonically from the blood islands in the yolk sac. They enter the circulation in the embryo proper after E8.5 as large nucleated cells referred to as primitive erythrocytes[33]. Expression of embryonic globins is a distinctive hallmark of primitive erythrocytes. The first intra-embryonically derived erythrocytes appear in the circulation around E12.5. These enucleated cells are generated in the fetal liver and referred to as definitive erythrocytes. They are characterized by predominant expression of adult globins; unlike the human *HBB* locus the mouse *Hbb* locus does not

harbor fetal β-like globin genes. Nevertheless, mice carrying human *HBB* locus transgenes have been extensively used to study fetal-to-adult hemoglobin switching[34].

To analyze the developmental dynamics of globin expression in *Klf1*$^{wt/Nan}$ embryos, we determined the globin expression profiles in RNA isolated from yolk sacs and fetal livers harvested between E11.5 and E16.5. Both the yolk sac and the fetal liver contain tissue cells plus circulating blood cells. First, we assessed expression of mβ-like globins. At E11.5, both the yolk sac and the fetal liver of control mice contained mainly mεy globin, reflecting the presence of primitive erythrocytes in the circulation and the fact that the fetal liver only just starts to produce definitive erythroid cells (Figure 1AC). In E12.5 fetal liver the production of large numbers of definitive erythroid cells was reflected in the dominant expression of mβ, while mεy still constituted >90% of mβ-like globins in the yolk sac (Figure 1AC). At E13.5, expression of mβ-like globins detected in yolk sac was 70% mεy and 30% mβ, whereas >95% of fetal liver mβ-like globin was mβ. Finally, mainly mβ was detected in fetal liver and yolk sac from E14.5 onward (Figure 1AC). In comparison, *Klf1*$^{wt/Nan}$ yolk sac and fetal liver expressed a larger fraction of mεy at both E12.5 and day E13.5 (Figure 1BD). The increase of mβ expression over time is delayed in *Klf1*$^{wt/Nan}$ yolk sac and fetal liver (Figure 1BD), indicating a delayed shift in expression of primitive to definitive mβ-like globins.

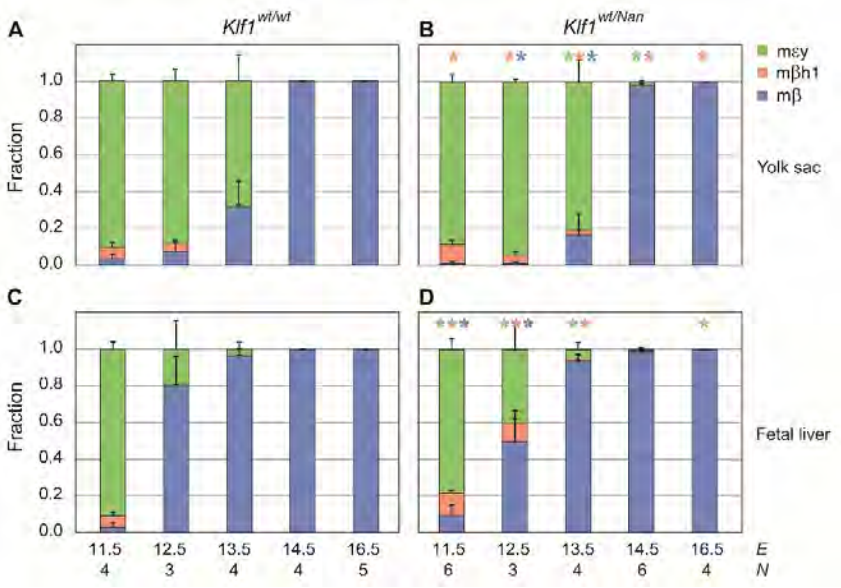


Figure 1. Developmental expression patterns of mouse β-like globins in control and *Klf1*$^{wt/Nan}$ embryos. Expression of mεy, mβh1 and mβ was determined by RT-qPCR. Data are displayed as fraction of total mβ-like globin (mεy+mβh1+mβ) expression. Embryonic day (*E*) and number of embryos (*N*) are indicated. Asterisks indicate p-values<0.05; error bars indicate standard deviations.

Next, we investigated whether this delay is specific for the *Hbb* locus, or whether it also occurs in the *Hba* locus. Yolk sacs and fetal livers from E11.5 control embryos expressed mζ as the major α-like globin, with mα contributing 25-30% to total α-like globin expression (Figure 2AC). At E12.5 mα became the dominant α-like globin in fetal liver, mζ was gradually replaced in the yolk sac by mα at E12.5 and E13.5, with the major shift to mα expression completed by E14.5 (Figure 2A), corresponding with fetal liver output in circulation. A different pattern was observed in yolk sacs and fetal livers from *Klf1*$^{wt/Nan}$ embryos. Compared to the controls, at E11.5 the contribution of mζ was increased at the expense of mα (Figure 2BD). At E12.5 and E13.5 there was no increase in mα globin expression in the yolk sac (Figure 2B), and the increase in expression of mα in fetal livers was reduced compared to control fetal livers (Figure 2BD). Thus, *Klf1*$^{wt/Nan}$ affects the developmental expression patterns of the mouse α-like and β-like globins in a very similar manner, displaying a delayed switch to expression of the adult genes.

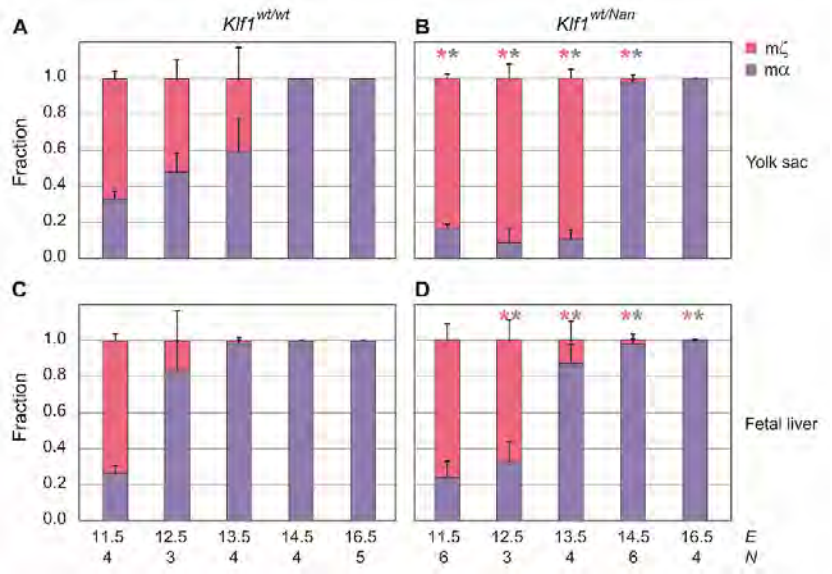


Figure 2. Developmental expression patterns of mouse α-like globins in control and *Klf1*$^{wt/Nan}$ embryos. Expression of mζ and mα was determined by RT-qPCR. Data are displayed as fraction of total mα-like globin (mζ+mα) expression. Embryonic day (*E*) and number of embryos (*N*) are indicated. Asterisks indicate p-values<0.05; error bars indicate standard deviations.

Because CDA-IV patients display very high HbF levels[20-22], we extended the observations on the mouse globins to the human β-like globin genes thus adding analysis of the developmental expression patterns of fetal-stage globin genes. The E11.5 yolk sac and fetal liver of control embryos expressed very similar ratios of hε and hγ, while expression of hβ was very low (Figure 3A). Compared to the controls, E11.5 *Klf1*$^{wt/Nan}$ yolk sac and

fetal liver displayed a small but significant shift to hε expression. Relatively increased hε expression was also observed in *Klf1*$^{wt/Nan}$ E12.5 yolk sac and fetal liver (Figure 3BD). Next to the difference in hε expression at E12.5, hβ made up ~75% of total hβ-like globins in control- compared to ~43% in *Klf1*$^{wt/Nan}$ fetal liver. This apparent lag in switching to hβ expression was also observed in *Klf1*$^{wt/Nan}$ E13.5 fetal liver, and in E13.5 and E14.5 *Klf1*$^{wt/Nan}$ yolk sacs (Figure 3BD). In control yolk sacs, expression of hβ increased rapidly to ~35% at E13.5, with hγ remaining the most abundantly expressed hβ-like globin accounting for ~50% of the total output of the *HBB* locus (Figure 3A).

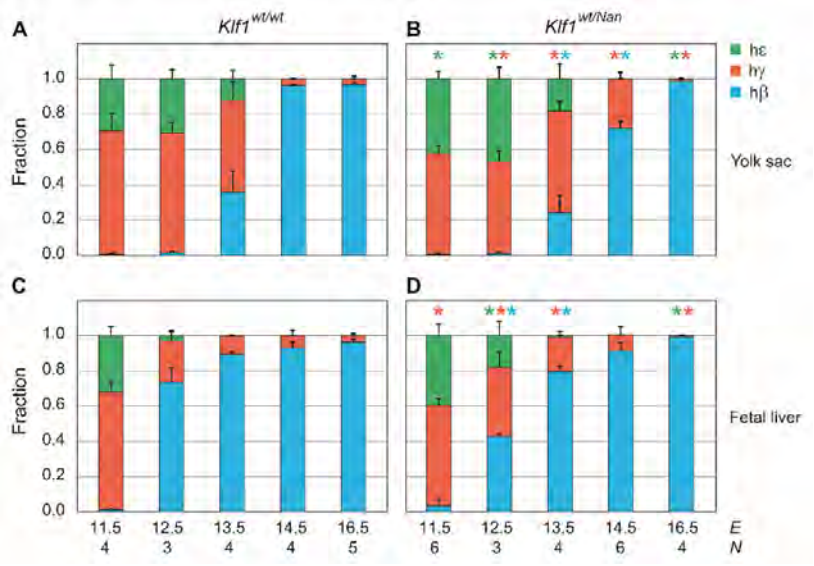


Figure 3. Developmental expression patterns of human β-like globins in control and *Klf1*$^{wt/Nan}$ embryos. Expression of hε, hγ and hβ was determined by RT-qPCR. Data are displayed as fraction of total hβ-like globin (hε+hγ+hβ) expression. Embryonic day (*E*) and number of embryos (*N*) are indicated. Asterisks indicate p-values<0.05; error bars indicate standard deviations.

Compared to control yolk sacs, expression of hγ in *Klf1*$^{wt/Nan}$ E13.5 yolk sacs was even higher at ~60%, with hβ expression also rapidly increasing but reaching a lower level of ~25% of hβ-like globins (Figure 3B). At E14.5, The hγ:hβ ratio shifted to 4:96 in control yolk sacs, while in *Klf1*$^{wt/Nan}$ yolk sacs this ratio remained higher at 28:72 (Figure 3AB). At E16.5, hβ expression accounted for >97% of total hβ-like globin in all yolk sacs and fetal livers, showing that hemoglobin switching had quantitatively proceeded to the adult profile in both genotypes (Figure 3). We conclude that expression of hε is maintained at higher levels in E11.5-E12.5 *Klf1*$^{wt/Nan}$ embryos. This is followed by a lag in switching to hβ expression, which favors expression of hγ, in E13.5-E14.5 *Klf1*$^{wt/Nan}$ embryos. Nevertheless, at E16.5 expression of hε and hγ has receded to <3% of total hβ-like globins.

Delayed appearance of definitive erythrocytes in the circulation of Klf1$^{wt/Nan}$ embryos

Having established that the shift from embryonic to fetal and adult globin expression is delayed in *Klf1*$^{wt/Nan}$ embryos, we investigated whether this could be due to a delay in embryonic development. The erythroid compartment changes dynamically during mouse development[33], and any alterations in this dynamic would be reflected in globin expression patterns. Primitive erythrocytes originate in the yolk sac and are the sole erythrocytes in the circulation until E12.5, when the first definitive erythrocytes are released from the fetal liver. As fetal liver erythropoiesis gathers momentum, the majority of cells in the circulation are definitive erythrocytes by E14.5. In contrast to primitive erythrocytes, definitive erythrocytes enucleate before they are released in the circulation. We used this characteristic to determine the contribution of primitive cells to the circulation by making cytospins of peripheral blood collected from E14.5 and E16.5 control and *Klf1*$^{wt/Nan}$ embryos. Compared to the controls, nucleated erythrocytes were more abundant in cytospins of *Klf1*$^{wt/Nan}$ blood. They were still easily detected in E16.5 cytospins of *Klf1*$^{wt/Nan}$ blood, while such cells were virtually absent in control samples (Figure 4A). To assess the switch from primitive to definitive erythropoiesis, E14.5 cytospins were split in early, mid and late times of litter harvest (Figure 4B). Consistent with the previous results, we observed that the fraction of nucleated cells declined very rapidly at this stage of development, in the controls from ~0.34 at early E14.5 to ~0.01 at late E14.5, and in the *Klf1*$^{wt/Nan}$ samples from ~0.52 at early E14.5 to ~0.13 at late E14.5. Importantly, compared to the controls the fraction of nucleated cells remained significantly higher in the *Klf1*$^{wt/Nan}$ samples in all E14.5 litters.

Using a CASY cell counter, we determined the cell size distributions in E14.5 blood samples. In E14.5 control samples, the two peaks representing primitive (large) and definitive (small) cells are clearly separated (Figure 4C). In E14.5 *Klf1*$^{wt/Nan}$ samples, these two peaks are not clearly separated. The apparent continuum of cell sizes is in agreement with the rampant anisocytosis observed in the cytospins of E14.5 *Klf1*$^{wt/Nan}$ blood (Figure 4A). Finally, we sought to use flow cytometry as an orthogonal approach to determine the contribution of primitive erythrocytes to the circulation. In an attempt to better distinguish primitive from definitive erythrocytes, we performed flow cytometry using CD71 (transferrin receptor), Ter119, and CD9 (Tetraspanin) which is a marker for primitive erythrocytes[35]. Compared to the controls, expression of CD71 was slightly increased on *Klf1*$^{wt/Nan}$ E10.5 primitive cells (Figure 4D). This might indicate a delay in maturation, similar to what has been proposed for *Klf1*$^{wt/ko}$ reticulocytes[36]. Expression of Ter119 was virtually absent in E10.5 *Klf1*$^{wt/Nan}$ erythrocytes, while CD9 expression was strongly reduced (Figure 4D). At E14.5, CD9 was unable to distinguish primitive from definitive erythrocytes in *Klf1*$^{wt/Nan}$ blood; in contrast to E14.5 *Klf1*$^{wt/wt}$ blood where a distinct fraction of CD9+ primitive cells was observed (Figure 4D, arrow). Collectively, we conclude that the contribution of primitive erythrocytes to the circulation of *Klf1*$^{wt/}$

^{Nan} embryos cannot be determined by flow cytometry using CD71, Ter119 and CD9 as markers. Despite these technical limitations, the analysis of blood samples is consistent with the notion that, compared to control embryos, the contribution of primitive erythrocytes to the pool of circulating erythrocytes in *Klf1*$^{wt/Nan}$ embryos is extended during development.

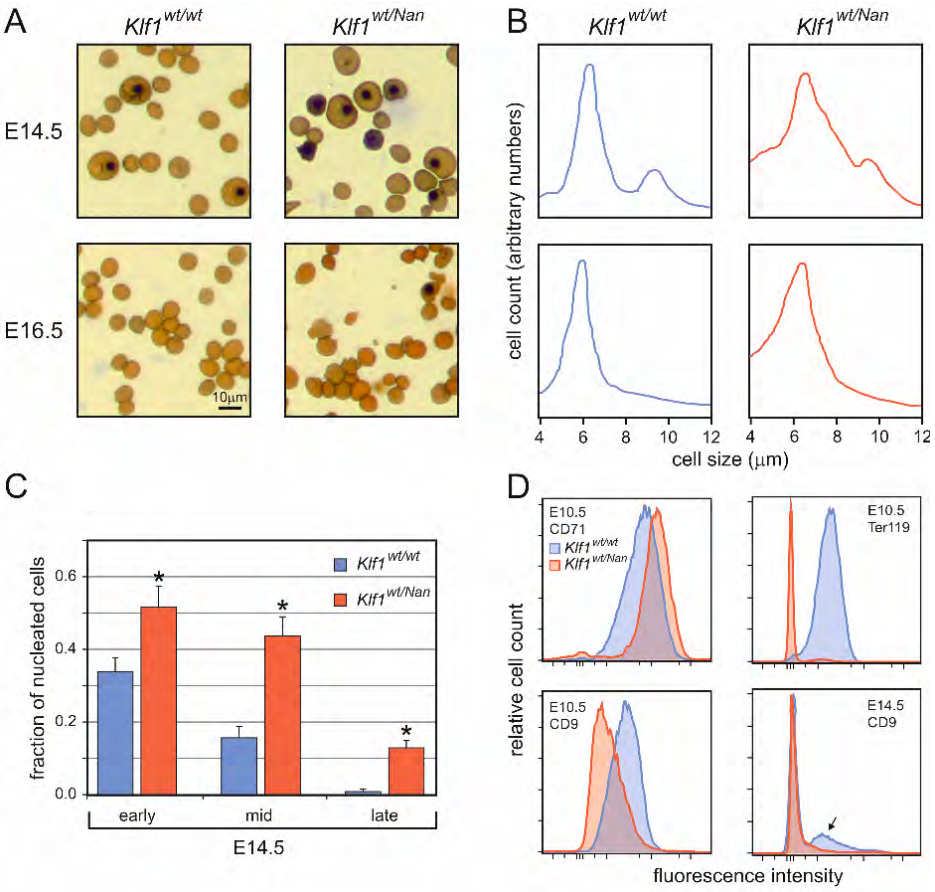


Figure 4. Analysis of blood from control and *Klf1*$^{wt/Nan}$ embryos. (A) Cytospins of E14.5 and E16.5 blood isolated from control (*Klf1*$^{wt/wt}$) and *Klf1*$^{wt/Nan}$ embryos, stained with dianisidine and histological dyes [32]. **(B)** Fraction of nucleated cells in blood isolated from control (*Klf1*$^{wt/wt}$) and *Klf1*$^{wt/Nan}$ littermates at early, mid and late E14.5. Asterisks indicate p-values<0.05; error bars indicate standard deviations. N=3 for each genotype at each developmental time point. **(C)** Cell size analysis of E14.5 and E16.5 blood isolated from control (*Klf1*$^{wt/wt}$) and *Klf1*$^{wt/Nan}$ embryos, obtained with a CASY instrument. **(D)** Histograms of flow cytometry analysis of blood isolated at E10.5 and E14.5 from control (*Klf1*$^{wt/wt}$) and *Klf1*$^{wt/Nan}$ embryos. Arrow indicates CD9+ primitive erythrocytes in *Klf1*$^{wt/wt}$ E14.5 blood.

Klf1$^{wt/Nan}$ erythroblasts display impaired proliferation and differentiation

The delays in hemoglobin switching and appearance of definitive erythrocytes in the circulation of *Klf1*$^{wt/Nan}$ embryos suggest that the production of definitive cells in the fetal liver might be affected by impaired proliferation or differentiation. Since KLF1 is critically involved in regulation of the erythroid cell cycle[11, 37, 38], central to both these processes, we cultured primary cells derived from E12.5 fetal livers to assess the proliferative capacity of *Klf1*$^{wt/Nan}$ erythroblasts. RT-qPCR analysis at Day 6 of culture showed deregulated expression of the embryonic/fetal globin genes in the *Klf1*$^{wt/Nan}$ cells compared to the controls, demonstrating that this aspect of the phenotype is maintained in the culture system (Supplementary Table 1). We have previously shown that *Klf1*$^{ko/ko}$ E12.5 fetal liver cells expand well when grown under self-renewal conditions[9]. In contrast, *Klf1*$^{wt/Nan}$ erythroblasts from E12.5 fetal liver expanded very poorly under these growth conditions (Figure 5A). Up to Day 3-4 of culture, *Klf1*$^{wt/Nan}$ fetal liver cells expanded similar to those derived from control embryos. After Day 4, expansion of the *Klf1*$^{wt/Nan}$ cultures slowed down and the percentage of smaller cells increased, suggesting spontaneous differentiation. Consistent with previously reported RT-qPCR data of *Klf1*$^{wt/Nan}$ fetal liver RNA[24], expression of cell cycle regulators E2F2, E2F4, and P18, all known KLF1 target genes[11, 37, 38], was downregulated in *Klf1*$^{wt/Nan}$ cells compared to the controls, while expression of P21 was unchanged (Figure 5B).

Next, the cells were switched to differentiation medium at Day 6. During differentiation the control cells, but not the *Klf1*$^{wt/Nan}$ cells, displayed the characteristic differentiation divisions, i.e. the cell number increased while cell size decreased (Figure 5CD). Cytospins taken at Day 2 of differentiation revealed many mature, enucleated and hemoglobinised cells in the control cultures (Figure 5Ea). Rare macrophages still present in the cultures were surrounded by healthy maturing erythroblasts (Figure 5Eb). Macrophage inclusions resembled nuclei, presumably resulting from phagocytosis of pyrenocytes (expelled erythroid nuclei[33]). In contrast, the *Klf1*$^{wt/Nan}$ cultures showed few enucleated cells and the cells displayed much larger nuclei (Figure 5Ec). Macrophages were not surrounded by enucleating erythroblasts, but appeared to engulf the entire erythroblast (Figure 5Ed). Combined these observations suggest that *Klf1*$^{wt/Nan}$ erythroblasts are impaired in both proliferation and differentiation.

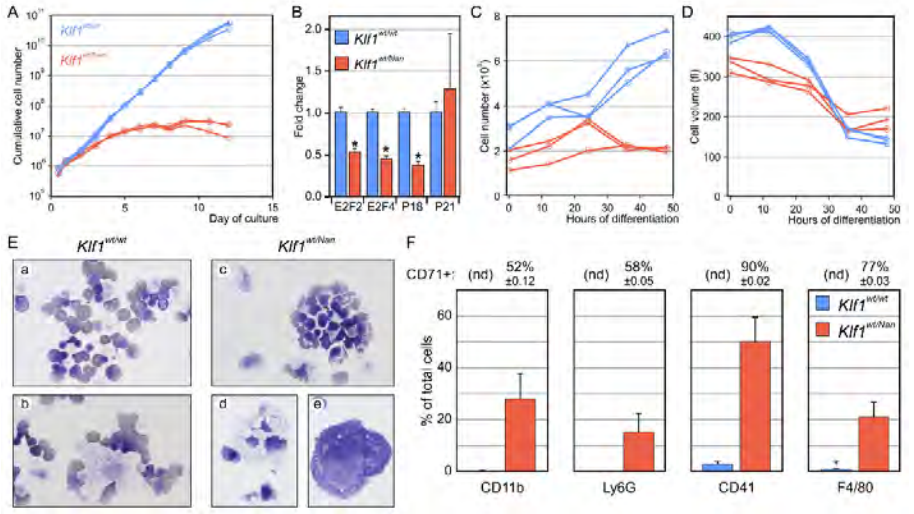


Figure 5. Erythroid progenitor cultures of control and *Klf1*$^{wt/Nan}$ E12.5 fetal liver cells. (A) Growth curves of primary erythroid progenitors cultured from control (*Klf1*$^{wt/wt}$) and *Klf1*$^{wt/Nan}$ E12.5 fetal livers in StemPro medium containing SCF, Dex and EPO. **(B)** RT-qPCR analysis of cell cycle regulators in cultured erythroid progenitors (n=3 for each group). Asterisks indicate p-values<0.05; error bars indicate standard deviations. **(C,D)** Differentiation was induced on Day 6, transferring the cells to StemPro medium supplemented with EPO and transferrin. Cell number **(C)** and cell volume **(D)** was recorded. **(E)** At Day 2 of differentiation cells were centrifuged on glass slides, fixed and stained with dianisidine and histological dyes [32]. **(F)** Flow cytometry analysis of cells cultured from control (*Klf1*$^{wt/wt}$) and *Klf1*$^{wt/Nan}$ fetal livers at Day 9 of culture. Cells were stained with antibodies indicated. The percentage of cells double-positive for the erythroid marker CD71 is shown on top, ± standard deviation. (nd): not detectable, due to the virtual absence of expression of myeloid markers on the control cells.

It is known that KLF1 blocks progression to myeloid lineages[39]. In the *Klf1*$^{wt/Nan}$ cultures we observed cells with morphological features of megakaryocytes (Figure 5Ee). Flow cytometry analysis of the cultured cells at Day 9 revealed that control cultures were essentially free of non-erythroid cells, while *Klf1*$^{wt/Nan}$ cultures displayed pan-myeloid markers on 20 to 50% of the cells (Figure 5F). Of note, the majority of these cells were positive for the erythroid marker CD71. Collectively, these data suggest that compromised lineage fidelity, reduced proliferative capacity and impaired terminal differentiation all contribute to the delay in abundance of definitive erythrocytes in the circulation of *Klf1*$^{wt/Nan}$ embryos, thus bearing weight on the observed changes in globin expression during embryonic development.

Expression of globins in adult Klf1$^{wt/Nan}$ mice

The analysis of developmental expression patterns of the globins demonstrated that by E16.5 *Klf1*$^{wt/Nan}$ embryos had quantitatively switched to expression of the adult genes (Figures 1-3). However, it remains possible that the maintenance of embryonic/fetal globin silencing is perturbed in adult *Klf1*$^{wt/Nan}$ mice. Indeed, derepression of embryonic globin genes in the spleen of *Klf1*$^{wt/Nan}$ mice has been reported[24]. To investigate this further, we isolated RNA from spleen and bone marrow derived from control and *Klf1*$^{wt/Nan}$ mice. By RT-qPCR analysis we found that mζ and mβh1, but not mε, expression was increased between 35- to 800-fold in *Klf1*$^{wt/Nan}$ samples compared to control samples (Figure 6). For the human β-like globins, we observed 4- to 9-fold increased expression of hε and hγ. In quantitative terms, even in the case of the most highly expressed embryonic globin mζ, this amounted to less than 0.3% of total α-like globin. We conclude that maintenance of embryonic/fetal globin silencing is perturbed in bone marrow and spleen of adult *Klf1*$^{wt/Nan}$ mice. Since the amount of embryonic/fetal globins produced remains below 0.3% of the total amount of globins, this is a qualitative rather than a quantitative trait.

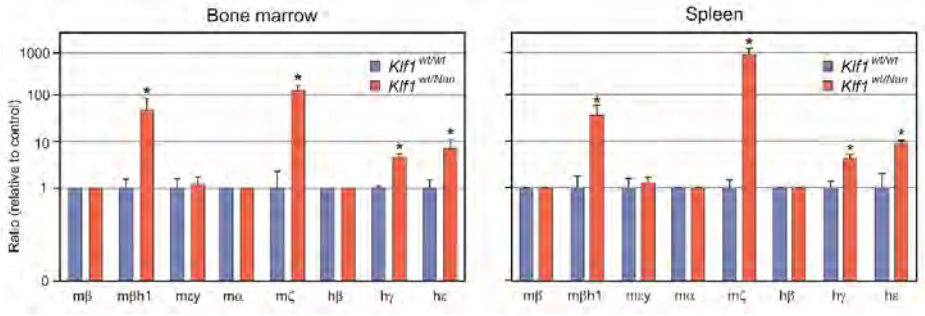


Figure 6. Expression of human and mouse globins in bone marrow and spleen of adult control and *Klf1*$^{wt/Nan}$ animals. Expression of mouse and human globins was determined by RT-qPCR. Expression ratios of individual globins in *Klf1*$^{wt/Nan}$ samples were calculated relative to those observed for the control samples (*Klf1*$^{wt/wt}$). Note logarithmic scale of the y-axis. N=3 for each group. Asterisks indicate p-values<0.05; error bars indicate standard deviations.

Discussion

In humans, reduced KLF1 activity has been associated with persistent expression of fetal hemoglobin in adults[1, 14]. A severe phenotype of hemolytic anemia characterizes patients suffering from CDA-IV, caused by the p.E325K variant in the DNA binding domain of KLF1. In the *Klf1*$^{wt/Nan}$ mouse, the orthologous glutamic acid residue (p.E339) is changed. Biochemically, these amino acid substitutions are very different. In CDA-IV, the glutamic acid (E) is replaced by a basic amino acid (lysine, K) while in KLF1^{Nan} it is replaced by aspartic acid (D), an acidic amino acid. Despite these biochemically opposing properties, the erythroid phenotypes of CDA-IV patients and *Klf1*$^{wt/Nan}$ mice share many similarities. A hallmark of CDA-IV patients is that they maintain high expression levels of embryonic and fetal globins. To investigate whether this is also the case in *Klf1*$^{wt/Nan}$ mice, we surveyed expression of the α-like and β-like globins during development. Our main observations are summarized in Supplementary Figure 1. We found that in *Klf1*$^{wt/Nan}$ embryos, switching from embryonic/fetal to adult globin genes is delayed in the endogenous *Hbb* and *Hba* loci, and in the single-copy human *HBB* locus transgene. Two mechanisms may contribute to this phenomenon. Firstly, decreased proliferation and differentiation of fetal erythroblasts delays the replacement of primitive by definitive erythroid cells. Since we isolated RNA from populations of cells, the ratio of primitive/definitive cells will have an impact on the globin levels measured. Secondly, although by E16.5 adult globins are quantitatively the dominant globins in *Klf1*$^{wt/Nan}$ embryos, the embryonic and fetal globin genes retained expression in adult spleen and bone marrow. Qualitatively, this persistent expression is another phenotypic similarity with CDA-IV patients. However, quantitatively there is a major difference. While in CDA-IV patients HbF levels of up to 37% of total hemoglobin have been reported[20], even the most highly expressed embryonic globin in adult *Klf1*$^{wt/Nan}$ mice, mζ, contributes only 0.3% to the total amount of α-like globins. KLF1 activates expression of BCL11A[14, 40] and LRF[41], two transcriptional repressors directly involved in hemoglobin switching[42, 43]. RT-qPCR analysis of BCL11A and LRF expression indicates that reduced expression of these two factors contributes to the sustained expression of the embryonic/fetal genes in *Klf1*$^{wt/nan}$ erythroid cells (Supplementary Figure 2). We note that expression of BCL1A is also significantly reduced in *Klf1*$^{wt/ko}$ cells with little effect on expression of the embryonic/fetal genes[36]. Mechanistically, this suggests that the repressor proteins are expressed well above the critical threshold level in mice, and a reduction to 40-60% of normal expression would still be sufficient for quantitative silencing of the embryonic/fetal genes.

The impaired proliferation of E12.5 fetal liver-derived *Klf1*$^{wt/Nan}$ erythroblasts was initially surprising because lack of KLF1 increased proliferation, likely due to impaired spontaneous differentiation[9]. Whereas erythropoiesis in *Klf1*$^{ko/ko}$ mice is severely affected during terminal differentiation, erythropoiesis in *Klf1*$^{wt/nan}$ mice is much less affected

with respect to terminal differentiation. Presence of KLF1^{Nan} not only results in reduced expression of KLF1 target genes, but also induces expression of genes not normally regulated by KLF1[26, 27, 44]. The combined effects of deregulation of canonical KLF1 target genes and ectopic gene expression likely underlie the observed lineage commitment and impaired proliferation and differentiation of *Klf1*$^{wt/nan}$ erythroblast cultures. Defective growth of erythroid progenitor cultures derived from a CDA-IV patient has been reported[45, 46], indicating that impaired proliferation of erythroid progenitors is another hallmark that CDA-IV patients and *Klf1*$^{wt/Nan}$ mice have in common. In contrast, adult *Klf1*$^{wt/nan}$ mice expressed mainly adult-type globin genes, as opposed to adult CDA-IV patients who maintain expression of embryonic and fetal globins at substantial levels[20-22, 45].

Importantly, the dominant effect of KLF1^{Nan} is illustrated by comparison with *Klf1*$^{wt/ko}$ mice which do not display deregulated expression of mouse embryonic globins in E14.5 yolk sac and fetal liver and adult bone morrow[36] (Supplementary Figure 3). We found that the effects of KLF1^{Nan} on developmental regulation of globin expression are very consistent, but surprisingly subtle. KLF1^{Nan} affects the dynamics of progression from primitive to definitive erythropoiesis during mouse development. Compared to control embryos, definitive erythrocytes emerge at a later stage as the dominant cell type in the circulation of *Klf1*$^{wt/Nan}$ embryos. We propose that this is at least in part due to the reduced expansion and differentiation capacity of the fetal liver progenitors, since cultured *Klf1*$^{wt/Nan}$ E12.5 fetal liver cells display growth and differentiation defects. Consistent with impaired erythroid differentiation, we observed aberrant expression of erythroid flow cytometry markers (CD71, Ter119, CD9). Furthermore, we found misexpression of myeloid markers, in particular the megakaryocyte marker CD41, indicating that *Klf1*$^{wt/Nan}$ erythroid progenitors display lineage infidelity. This is akin to the previously reported aberrantly activated megakaryocyte program in *Klf1*$^{ko/ko}$ erythroid cells[39, 47].

Collectively, our data support the notion that KLF1^{CDA} and KLF1^{Nan} present with similar but also variant-specific phenotypes. Thus, our study further highlights the need to investigate the effects of individual KLF1 variants in detail[1]. The recently developed human adult erythroid progenitor cell lines HUDEP-2[48] and BEL-A[49] could be combined with CRISPR-mediated homology-directed recombination[50] to investigate the impact of individual human KLF1 variants on the molecular control of erythropoiesis, with a view to increase understanding of the broad spectrum of human red blood cell disorders caused by KLF1 variants[1].

Acknowledgments

Research in our laboratories was funded by the Landsteiner Foundation for Blood Transfusion Research (LSBR 1040 and 1627), the Netherlands Organization for Scientific Research (ZonMw TOP 40-00812-98-12128), the Netherlands Genomics Initiative (NGI

Zenith 93511036), and EU fp7 Specific Cooperation Research Project THALAMOSS (306201).

Authorship and Conflict-of-Interest Statements

AK, NG, IC, SH and EvdA performed experiments. AK, TBvD, SH, SP, EvdA and MvL analyzed data. MvL and SP conceived the study. AK, MvL and SP made the figures and wrote the paper.

References

1. Perkins A, Xu X, Higgs DR, et al. Kruppeling erythropoiesis: an unexpected broad spectrum of human red blood cell disorders due to KLF1 variants. Blood. 2016;127(15):1856-1862.
2. Miller IJ, Bieker JJ. A novel, erythroid cell-specific murine transcription factor that binds to the CACCC element and is related to the Kruppel family of nuclear proteins. Mol Cell Biol. 1993;13(5):2776-2786.
3. Suske G, Bruford E, Philipsen S. Mammalian SP/KLF transcription factors: bring in the family. Genomics. 2005;85(5):551-556.
4. Eaton SA, Funnell AP, Sue N, Nicholas H, Pearson RC, Crossley M. A network of Kruppel-like Factors (Klfs). Klf8 is repressed by Klf3 and activated by Klf1 in vivo. J Biol Chem. 2008;283(40):26937-26947.
5. Zhang P, Basu P, Redmond LC, et al. A functional screen for Kruppel-like factors that regulate the human gamma-globin gene through the CACCC promoter element. Blood Cells Mol Dis. 2005;35(2):227-235.
6. Nuez B, Michalovich D, Bygrave A, Ploemacher R, Grosveld F. Defective haematopoiesis in fetal liver resulting from inactivation of the EKLF gene. Nature. 1995;375(6529):316-318.
7. Perkins AC, Sharpe AH, Orkin SH. Lethal beta-thalassaemia in mice lacking the erythroid CACCC-transcription factor EKLF. Nature. 1995;375(6529):318-322.
8. Perkins AC, Peterson KR, Stamatoyannopoulos G, Witkowska HE, Orkin SH. Fetal expression of a human Agamma globin transgene rescues globin chain imbalance but not hemolysis in EKLF null mouse embryos. Blood. 2000;95(5):1827-1833.
9. Drissen R, von Lindern M, Kolbus A, et al. The erythroid phenotype of EKLF-null mice: defects in hemoglobin metabolism and membrane stability. Mol Cell Biol. 2005;25(12):5205-5214.
10. Hodge D, Coghill E, Keys J, et al. A global role for EKLF in definitive and primitive erythropoiesis. Blood. 2006;107(8):3359-3370.
11. Pilon AM, Arcasoy MO, Dressman HK, 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. 2008;28(24):7394-7401.
12. Tallack MR, Magor GW, Dartigues B, et al. Novel roles for KLF1 in erythropoiesis revealed by mRNA-seq. Genome Res. 2012;22(12):2385-2398.
13. Singleton BK, Burton NM, Green C, Brady RL, Anstee DJ. Mutations in EKLF/KLF1 form the molecular basis of the rare blood group In(Lu) phenotype. Blood. 2008;112(5):2081-2088.
14. Borg J, Papadopoulos P, Georgitsi M, et al. Haploinsufficiency for the erythroid transcription factor KLF1 causes hereditary persistence of fetal hemoglobin. Nat Genet. 2010;42(9):801-805.
15. Liu D, Zhang X, Yu L, et al. KLF1 mutations are relatively more common in a thalassemia endemic region and ameliorate the severity of beta-thalassemia. Blood. 2014;124(5):803-811.
16. Huang J, Zhang X, Liu D, et al. Compound heterozygosity for KLF1 mutations is associated with microcytic hypochromic anemia and increased fetal hemoglobin. Eur J Hum Genet. 2015;23(10):1341-1348.
17. Satta S, Perseu L, Moi P, et al. Compound heterozygosity for KLF1 mutations associated with remarkable increase of fetal hemoglobin and red cell protoporphyrin. Haematologica. 2011;96(5):767-770.
18. Viprakasit V, Ekwattanakit S, Riolueang S, et al. Mutations in Kruppel-like factor 1 cause transfusion-dependent hemolytic anemia and persistence of embryonic globin gene expression. Blood. 2014;123(10):1586-1595.

19. Magor GW, Tallack MR, Gillinder KR, et al. KLF1-null neonates display hydrops fetalis and a deranged erythroid transcriptome. Blood. 2015;125(15):2405-2417.
20. Arnaud L, Saison C, Helias V, et al. A dominant mutation in the gene encoding the erythroid transcription factor KLF1 causes a congenital dyserythropoietic anemia. Am J Hum Genet. 2010;87(5):721-727.
21. de-la-Iglesia-Inigo S, Moreno-Carralero MI, Lemes-Castellano A, Molero-Labarta T, Mendez M, Moran-Jimenez MJ. A case of congenital dyserythropoietic anemia type IV. Clin Case Rep. 2017;5(3):248-252.
22. Jaffray JA, Mitchell WB, Gnanapragasam MN, et al. Erythroid transcription factor EKLF/KLF1 mutation causing congenital dyserythropoietic anemia type IV in a patient of Taiwanese origin: review of all reported cases and development of a clinical diagnostic paradigm. Blood Cells Mol Dis. 2013;51(2):71-75.
23. Heruth DP, Hawkins T, Logsdon DP, et al. Mutation in erythroid specific transcription factor KLF1 causes Hereditary Spherocytosis in the Nan hemolytic anemia mouse model. Genomics. 2010;96(5):303-307.
24. Siatecka M, Sahr KE, Andersen SG, Mezei M, Bieker JJ, Peters LL. Severe anemia in the Nan mutant mouse caused by sequence-selective disruption of erythroid Kruppel-like factor. Proc Natl Acad Sci U S A. 2010;107(34):15151-15156.
25. Basu P, Lung TK, Lemsaddek W, et al. EKLF and KLF2 have compensatory roles in embryonic beta-globin gene expression and primitive erythropoiesis. Blood. 2007;110(9):3417-3425.
26. Nebor D, Graber JH, Ciciotte SL, et al. Mutant KLF1 in Adult Anemic Nan Mice Leads to Profound Transcriptome Changes and Disordered Erythropoiesis. Sci Rep. 2018;8(1):12793.
27. Planutis A, Xue L, Trainor CD, et al. Neomorphic effects of the neonatal anemia (Nan-Eklf) mutation contribute to deficits throughout development. Development. 2017;144(3):430-440.
28. Strouboulis J, Dillon N, Grosveld F. Developmental regulation of a complete 70-kb human beta-globin locus in transgenic mice. Genes Dev. 1992;6(10):1857-1864.
29. 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. 2002;9(6):1319-1326.
30. Lyon MF, Glenister PH, Loutit JF, Peters J. Dominant haemolytic anemia. Mouse News Letters. 1983;68(68.
31. Dolznig H, Kolbus A, Leberbauer C, et al. Expansion and differentiation of immature mouse and human hematopoietic progenitors. Methods Mol Med. 2005;105(323-344.
32. Beug H, Leutz A, Kahn P, Graf T. Ts mutants of E26 leukemia virus allow transformed myeloblasts, but not erythroblasts or fibroblasts, to differentiate at the nonpermissive temperature. Cell. 1984;39(3 Pt 2):579-588.
33. Palis J. Primitive and definitive erythropoiesis in mammals. Front Physiol. 2014;5(3.
34. Peterson KR. Hemoglobin switching: new insights. Curr Opin Hematol. 2003;10(2):123-129.
35. Isern J, Fraser ST, He Z, Zhang H, Baron MH. Dose-dependent regulation of primitive erythroid maturation and identity by the transcription factor Eklf. Blood. 2010;116(19):3972-3980.
36. Esteghamat F, Gillemans N, Bilic I, et al. Erythropoiesis and globin switching in compound Klf1::Bcl11a mutant mice. Blood. 2013;121(13):2553-2562.
37. Gnanapragasam MN, McGrath KE, Catherman S, Xue L, Palis J, Bieker JJ. EKLF/KLF1-regulated cell cycle exit is essential for erythroblast enucleation. Blood. 2016;128(12):1631-1641.
38. Tallack MR, Keys JR, Humbert PO, Perkins AC. EKLF/KLF1 controls cell cycle entry via direct regulation of E2f2. J Biol Chem. 2009;284(31):20966-20974.

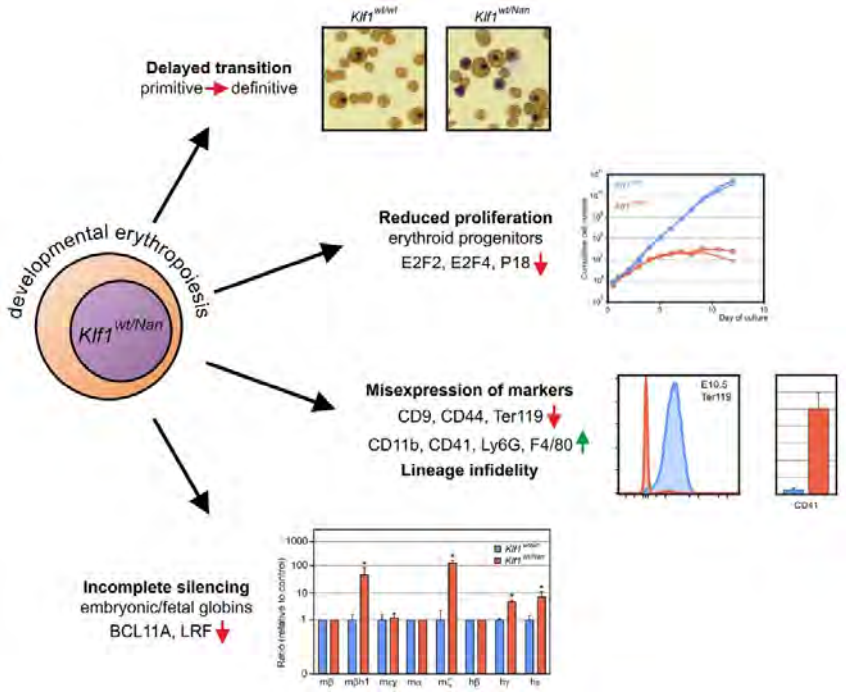
39. Tallack MR, Perkins AC. Megakaryocyte-erythroid lineage promiscuity in EKLF null mouse blood. Haematologica. 2010;95(1):144-147.
40. Zhou D, Liu K, Sun CW, Pawlik KM, Townes TM. KLF1 regulates BCL11A expression and gamma- to beta-globin gene switching. Nat Genet. 2010;42(9):742-744.
41. Norton LJ, Funnell APW, Burdach J, et al. KLF1 directly activates expression of the novel fetal globin repressor ZBTB7A/LRF in erythroid cells. Blood Adv. 2017;1(11):685-692.
42. Masuda T, Wang X, Maeda M, et al. Transcription factors LRF and BCL11A independently repress expression of fetal hemoglobin. Science. 2016;351(6270):285-289.
43. Sankaran VG, Xu J, Ragoczy T, et al. Developmental and species-divergent globin switching are driven by BCL11A. Nature. 2009;460(7259):1093-1097.
44. Gillinder KR, Ilsley MD, Nebor D, et al. Promiscuous DNA-binding of a mutant zinc finger protein corrupts the transcriptome and diminishes cell viability. Nucleic Acids Res. 2017;45(3):1130-1143.
45. Ravindranath Y, Johnson RM, Goyette G, Buck S, Gadgeel M, Gallagher PG. KLF1 E325K-associated Congenital Dyserythropoietic Anemia Type IV: Insights Into the Variable Clinical Severity. J Pediatr Hematol Oncol. 2018;40(6):e405-e409.
46. Varricchio L, Planutis A, Manwani D, et al. Genetic disarray follows mutant KLF1-E325K expression in a congenital dyserythropoietic anemia patient. Haematologica. 2019;104(12):2372-2380.
47. Frontelo P, Manwani D, Galdass M, et al. Novel role for EKLF in megakaryocyte lineage commitment. Blood. 2007;110(12):3871-3880.
48. Kurita R, Suda N, Sudo K, et al. Establishment of immortalized human erythroid progenitor cell lines able to produce enucleated red blood cells. PLoS One. 2013;8(3):e59890.
49. Trakarnsanga K, Griffiths RE, Wilson MC, et al. An immortalized adult human erythroid line facilitates sustainable and scalable generation of functional red cells. Nat Commun. 2017;8(14750.
50. Mali P, Esvelt KM, Church GM. Cas9 as a versatile tool for engineering biology. Nat Methods. 2013;10(10):957-963.

Supplementary material and methods

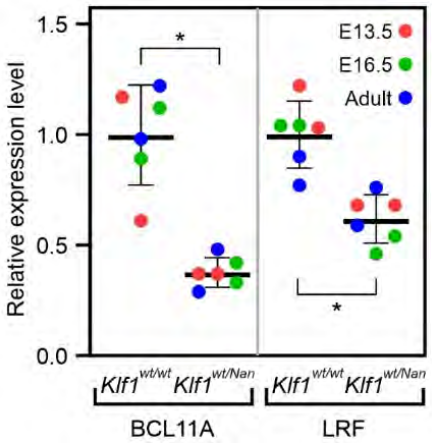
Supplementary Table 1.

	Klf1 wt/wt (n=4)	stdev	*Klf1* wt/Nan (n=4)	stdev	fold-change
mεy	0.0001	0.00003	0.0032	0.00232	54.06
mβh1	0.0000	0.00002	0.0006	0.00048	13.39
mβ	0.9999	0.00002	0.9961	0.00278	1.00
mζ	0.0000	0.00000	0.0010	0.00207	307.84
mα	1.0000	0.00000	0.9990	0.00207	1.00
hε	0.0000	0.00000	0.0007	0.00068	165.86
hγ	0.0049	0.00239	0.0494	0.01064	10.05
hβ	0.9951	0.00239	0.9498	0.01122	0.95

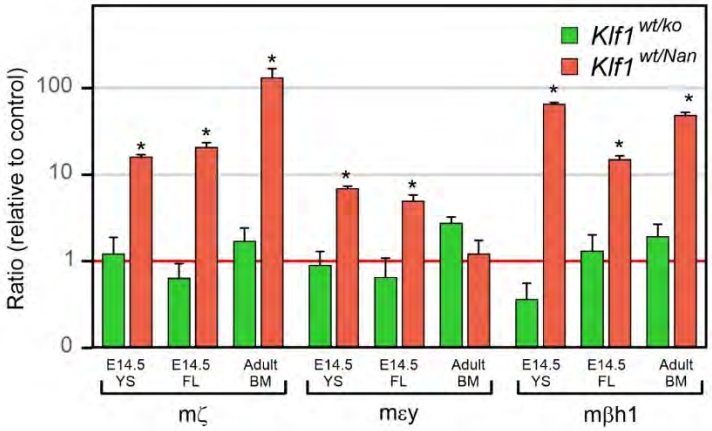
Globin expression in cultured E12.5 fetal liver cells (Day 6 of culture) was assessed by RT-qPCR. Globin expression was calculated as fraction of total mβ-like, mα-like or hβ-like globin. Fold-change is calculated as expression in *Klf1*$^{wt/Nan}$ cells divided by expression in *Klf1*$^{wt/wt}$ cells.



Supplementary Figure 1. Visual abstract summarizing the main observations reported in this paper



Supplementary Figure 2. Expression of BCL11A and LRF inerythroid tissues of control and *Klf1*$^{wt/Nan}$ embryos. Expression of BCL11A and LRF in fetal livers (E13.5 and E16.5) and adult bone marrow was determined by RT-qPCR. Expression of P112 was used as a reference to normalize the data; the average expression level in the control samples was set to 1. Asterisks indicate p-values<0.05; error bars indicate standard deviations.



Supplementary Figure 3. Expression of mouse embryonic globins in erythroid tissues of *Klf1*$^{wt/ko}$ and *Klf1*$^{wt/Nan}$ embryos. Expression of mouse embryonic globins in E14.5 yolk sac and fetal liver and adult bone marrow was determined by RT-qPCR. The average expression level in samples from control littermates (indicated by the red line) was set to 1. Asterisks indicate p-values<0.05 between *Klf1*$^{wt/ko}$ and *Klf1*$^{wt/Nan}$ samples; error bars indicate standard deviations. Note logarithmic scale. Data for *Klf1*$^{wt/ko}$ samples were obtained from [1].

Supplementary methods

Animals

Mice were maintained by breeding *Klf1*$^{wt/Nan::HBB}$ males with C57BL/6 females. Genotyping was performed by PCR using DNA isolated from toe biopsies. For *Klf1*^{Nan} genotyping, the PCR product was digested with DpnII [2].

Primers used for genotyping

All primer sequences are given in 5' to 3' direction.

Klf1^{Nan}

Fw: CTGCAGGATTGCAGCTGTAGATAC

Rv: AGTCCTTGTGCAGGATCACTCAGA

Approximately 340 bp PCR product for the wildtype *Klf1* allele, and 240 + 100 bp for the *Klf1*^{Nan} allele after DpnII digestion [2].

PAC8.1 human HBB locus single copy transgene [3]

Fw: GCTGCTGTTATGACCACTAGAGGG

Rv: AGACAGGGAAGGAGGTGTGG

PCR product 500 bp.

PCR conditions: 3 min 94°C, 28 cycles [30 sec 94°C, 30 sec 56°C, 30 sec 72°C], 2 min 72°C.

For timed pregnancies, *Klf1*$^{wt/Nan::HBB}$ males were mated with C57BL/6 and C57BL/6^{HBB} females. The day of vaginal plug discovery was considered E0.5. Embryos were collected between E11.5 and E16.5; head DNA was used for genotyping. Adult mice were analyzed at >18 weeks of age.

RT-qPCR

To synthesize cDNA, 2 µg of total RNA was used together with oligo dT, RNase OUT, and SuperScript reverse transcriptase II (all Thermo Fischer Scientific) in a total volume of 20 µl for 1 hour at 42°C. 0.2 µl of cDNA was used for amplification by RT-qPCR.

Primers were designed using Primer-BLAST, available through the NCBI web site at https://www.ncbi.nlm.nih.gov/tools/primer-blast/index.cgi?LINK_LOC=BlastHome.

PCR product size was set to 'Maximum 150 bp'. To avoid inter-globin and inter-species cross-specificity, the primer pair specificity checking parameters were set to 'Enable search for primer pairs specific to the intended PCR template' and 'Database-Genomes for selected organisms (primary reference assembly only)-Homo sapiens, Mus musculus'. Exon/intron selection was set to 'Primer pair must be separated by at least one intron on the corresponding genomic DNA' and Intron length range 'Minimum 100 bp'. All other settings were left at default values.

Mouse globins

Gene	globin encoded	Primer pair	PCR product (bp)
Hba-x	mζ	CCGGTCAACTTCAAGCTCCT	103
		TGAACTTGTCCCAGGCTTCG	
Hba-a1	mα1	GCTGAAGCCCTGGAAAGGAT	82
		CAGAGCCGTGGCTTACATCA	
Hba-a2	mα2	Identical to *Hba-a1*	
Hbb-y	mεy	TTGGCTAGTCACTTCGGCAAT	145
		GCATAGCGGACACACAGGAT	
Hbb-bh1	mβh1	TGGGCTTGGGGGTTAAGAAC	118
		AACATGTTGCCCAGGAGCTT	
Hbb-b1	mβmajor	GCTGCATGTGGATCCTGAGA	112
		CTTCTGGAAGGCAGCCTGTG	
Hbb-b2	mβminor	Identical to *Hbb-b1*	

Human globins

Gene	globin encoded	Primer pair	PCR product (bp)
HBE	hε	CCCTGGCCCATAAGTACCAC	105
		TTTCTCTCAAGGCCAAGCCC	
HBG1	hAγ	GGTGACCGTTTTGGCAATCC	106
		GTATCTGGAGGACAGGGCAC	
HBG2	hGγ	Identical to *HBG1*	
HBB	hβ	GCCCTGGCCCACAAGTATC	109
		GCCCTTCATAATATCCCCCAGTT	

Mouse transcription factors

Gene	factor encoded	Primer pair	PCR product (bp)
Bcl11a	BCL11A	CGTGTGCAGACCGAGGAGAGG	129
		GCATCCAGGTCACGCCAGAGG	
Zbtb7a	LRF	AGAAGGTGATTCAGGGTGCC	112
		AGCTTGTCCTGTCTGGTGAAT	

Mouse cell cycle regulators

Gene	factor encoded	Primer pair	PCR product (bp)
E2f2	E2F2	CTGAATTCCGGACCCCCAAG	101
		CGACGTGTCATAGCGTGTCT	
E2f4	E2F4	GCTTGGCCTACGTGACTCAT	100
		ATGGGCACCTCTAGACTGGT	
Cdkn1a	P21	GAATTGGAGTCAGGCGCAGA	89
		GAACAGGTCGGACATCACCA	

Cdkn2c	P18	AATGGATTTGGGAGAACTGCG GGAGAAGCCTCCTGGCAATC	70

Mouse reference gene

Gene	**protein encoded**	**Primer pair**	**PCR product (bp)**
Psmd1	P112	AATGTTCCAGCGATGTCTCG GACATGCAGAGTTTGAGGCTG	150

Quantitative PCR

PCR amplification was performed on a CFX96 Touch Real-Time PCR Detection System (Bio-Rad, Hercules, CA) using Platinum Taq DNA polymerase (Invitrogen). PCR conditions were 3 min 95°C, 40 cycles [30 sec 95°C, 25 sec 60°C, 15 sec 72°C], 5 sec 60°C, 5 min 95°C. Specific PCR product accumulation was monitored by SYBR Green dye fluorescence (Sigma-Aldrich, Saint Louis, MO). Using dilution series of yolk sac and fetal liver cDNAs, we found that the linear dynamic range of all globin primer pairs exceeded 256-fold dilution of the initial cDNA concentration. Dissociation curves were used to assess the homogeneity of PCR products. In addition, PCR product sizes were checked by agarose gel electrophoresis. To obtain expression values, Ct values were transformed by computing 2^{-Ct}. Total mα-like (mζ+mα), mβ-like (mεy+mβh1+mβ) and hβ-like globins (hε+hγ+hβ) expression was used to calculate the contribution of individual globins to the total for each globin in each sample, e.g. mζ/(mζ+mα) and mα/(mζ+mα) for mα-like globins. These fractional values were used to calculate the average and standard deviation for all samples belonging to the same group (e.g. *Klf1*$^{wt/Nan}$ E12.5 fetal liver). Analysis of RNA-seq data revealed that, compared to commonly used reference genes such as *GAPDH* and *ACTB*, expression of *PSMD1* (encoding P112, a proteasome 26S subunit) was relatively stable in primary human erythroblasts cultured under self-renewal or differentiation conditions (Heshusius *et al*, manuscript in preparation). We therefore used *Psmd1* as a reference gene to normalize expression levels of the *Bcl11a* and *Zbtb7a* genes.

Culture of mouse erythroid progenitors

E12.5 fetal livers were disrupted and single cell suspensions were seeded into StemPro medium (Thermo Fisher Scientific, Waltham, MA) supplemented with 1 U/ml human recombinant erythropoietin (EPO; Cilag, Schaffhausen, CH), 10 ng/ml murine recombinant stem cell factor (SCF; Supernatant of cells harbouring an expression construct) and 10^{-6} M dexamethasone (Dex; Sigma-Aldrich, Saint Louis, MO) [4]. The cultures of erythroid progenitors were subjected to daily partial medium changes and addition of fresh factors. Cell numbers and size distributions were determined daily, using an electronic cell counter (CASY; Roche, Basel, CH). Cell density was kept between 0.5-2.0x10^6 cells/ml. To differentiate erythroid progenitors, cells were washed twice in phosphate-buffered saline

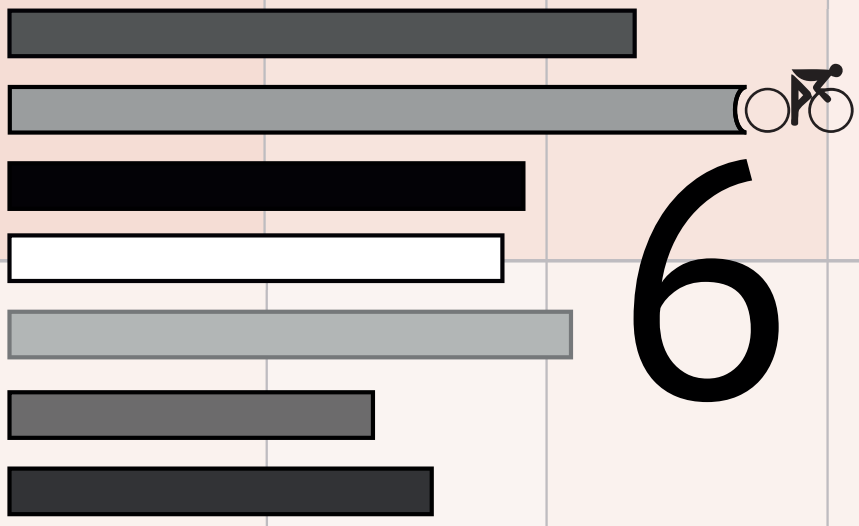
(PBS) and seeded at 1.5x10^6 cells/ml in differentiation medium: StemPro supplemented with 10 U/ml EPO and 1 mg/ml iron-saturated human transferrin (Sigma-Aldrich). Cells were maintained at densities of 2-4x10^6 cells/ml.

Flow cytometry analysis

Single-cell suspensions of peripheral blood samples were washed twice with PBS and resuspended in FACS buffer (PBS containing 1% (w/v) bovine serum albumin, 2 mM EDTA). Approximately 10^6 cells were incubated for 30 minutes at room temperature with CD71-BV421 (562716, dilution 1:400; BD Biosciences, San Jose, CA), CD9-PE (12-0019-81, dilution 1:300; Thermo Fischer Scientific) and Ter119-APC (17-5921-82, dilution 1:200; Thermo Fischer Scientific) antibodies in a final volume of 100 µl. Antibodies were diluted in FACS buffer. The cells were washed, and living cells were distinguished negatively by live/dead Aqua staining (L34965; dilution 1:400; Thermo Fischer Scientific). Cultured cells were washed with PBS and incubated in for 30 minutes at 4 ºC in FACS-buffer with CD71-APC (130-091-727, Miltenyi Bergisch Gladbach DE, dilution: 1:200), Ter119-BV450 (48-5921-82 Thermo Fischer Scientific, dilution 1:200), Ly6G-PE (RB6-8C5, Thermo Fischer Scientific, dilution 1:100), F4/80-FITC (11-4801-85, BD Biosciences, dilution 1:200), CD11b-AF488 (101217, Biolegend San Diego CA, dilution 1:100), CD41PE (12-0411-82, Thermo Fischer Scientific, dilution: 1:100).Data were acquired on a Fortessa instrument (BD Biosciences), and analyzed with FlowJo software v10 (Tree Star, Ashland, OR).

Supplementary References

1. Esteghamat F, Gillemans N, Bilic I, et al. Erythropoiesis and globin switching in compound Klf1::Bcl11a mutant mice. Blood. 2013;121(13):2553-2562.
2. Siatecka M, Sahr KE, Andersen SG, Mezei M, Bieker JJ, Peters LL. Severe anemia in the Nan mutant mouse caused by sequence-selective disruption of erythroid Kruppel-like factor. Proc Natl Acad Sci U S A. 2010;107(34):15151-15156.
3. 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. 2002;9(6):1319-1326.
4. Dolznig H, Kolbus A, Leberbauer C, et al. Expansion and differentiation of immature mouse and human hematopoietic progenitors. Methods Mol Med. 2005;105:323-344.



6

Chapter 6

General Discussion

Steven Heshusius

The aim of this thesis was to advance knowledge on the regulation of fetal hemoglobin genes and identify novel candidate modifiers of its expression. To achieve this, we improved our *in vitro* cell culture of human erythroblasts and used this cell culture model to investigate genetic determinants and environmental factors that affect HbF expression. Here I will discuss how the data generated in this thesis can be integrated with current knowledge to extend the models of developmental gene expression during erythropoiesis and what the implications are for development of novel HbF inducing agents.

Cultured cells express higher levels of fetal hemoglobin along with genes associated with the integrated stress response.

We generated transcriptome profiles that cover the entire course of terminal differentiation leading up to and beyond enucleation. Comparison of the transcriptome of the cultured cells and of *ex vivo* cells at a similar stage showed a high correspondence for most detectable transcripts (**chapter 2**). Although the transcriptome of the *ex vivo* erythroblasts were derived from a previous publication[1] the transcriptomes matched well with the exception of the *HBG* genes that show a higher expression in cultured cells. Closer examination of the transcriptome of the cultured cells, in comparison to their *ex vivo* counterparts, provides an opportunity to determine which components of the culture setup affect HbF expression *in vitro*.

Next to increased expression of the *HBG* genes, nearly 2,000 of the 13,000 detectable transcripts were expressed at least four-fold higher or lower in cultured compared to *ex vivo* cells. Referenced against the Reactome database, that catalogs information on molecular pathways[2], the top 5 enriched pathways expressed higher cultured cells included genes associated with cellular senescence and cellular response to stress (table 1), which could mean that there is indeed higher oxidative stress under the cell culture conditions used.

Table 1. Enriched term in genes expressed in cultured erythroblasts

Reactome term	Overlap	P-val	P-adj
Cellular Senescence_Homo sapiens_R-HSA-2559583	*57/161*	8.8E-19	1.35E-15
Metabolism of proteins_Homo sapiens_R-HSA-392499	*194/1074*	7.21E-18	5.51E-15
HDACs deacetylate histones_Homo sapiens_R-HSA-3214815	*32/60*	2.6E-17	1.32E-14
Cellular responses to stress_Homo sapiens_R-HSA-2262752	*89/367*	3.27E-16	1.25E-13
DNA Damage/Telomere Stress Induced Senescence_Homo sapiens_R-HSA-2559586	*30/59*	1.45E-15	4.44E-13
ERCC6 (CSB) and EHMT2 (G9a) positively regulate rRNA expression_Homo sapiens_R-HSA-427389	*25/43*	5.25E-15	1.34E-12

The induction of HbF in response to stress has been described as part of the mechanisms of HbF inducing agents like hydroxyurea.[3] This would fit with the higher levels of HbF observed in erythroid cultures resulting from increased oxidative stress under prolonged SCF and EPO stimulation or from extensive expansion of progenitors in the culture system.[4,5]

Signals from oxidative stress can be relayed to drive transcription by the cellular stress response via p38 mitogen activated kinase signaling and the integrated stress response (ISR). These pathways converge on a set of activating transcription factors, including ATF2-4, CHOP and MYC, that can in turn induce higher levels of HbF.[6,7] With the exception of ATF2 these factors are expressed between 16- and 32-fold higher in cultured cells (Fig. 1A). Also, the targets of these factors are overrepresented in the genes that are expressed higher *in vitro*; of the 351 detected CHOP and ATF4 targets[8] in our transcriptome dataset, 101 genes showed over four-fold higher expression in cultured cells (Fig. 1B p= 4.0e-8). Similarly, high expression of nearly a third of the known ATF4 targets (without CHOP) could indicate increased activity of the ISR pathway and would help explain increased HbF levels in cultured cells. Although overrepresentation of stress associated pathways and some of their core regulators implicates the ISR pathway for regulation of HbF in cultured cells, to which extent activation is triggered by higher ROS in cultured cells remains to be investigated.

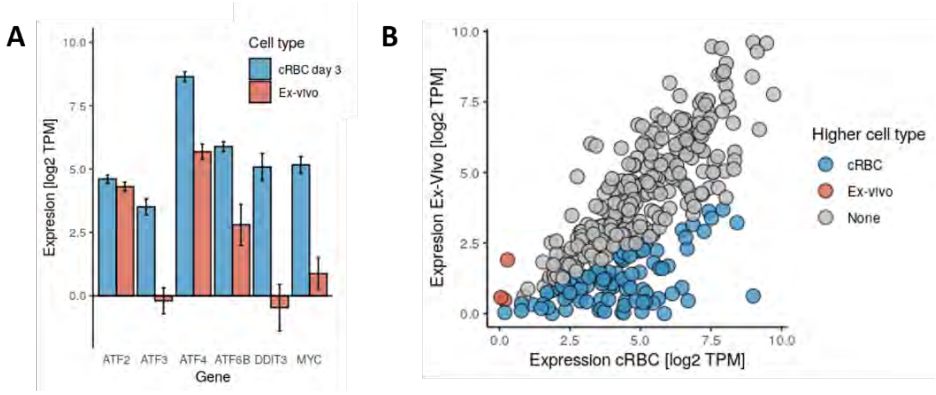


Figure 1. Expression of genes associated with ATF4 and CHOP. (A) mRNA expression for transcription factors central to integrated stress response (n=3-4 bar plot showing mean ±sd.) **(B)** Comparison of transcript levels for ATF4 and CHOP target in day cRBC and *Ex-vivo* cells.

To offset the detrimental effects of high ROS levels, cells produce antioxidizing enzymes like, heme oxygenase-1 (HO-1).[9] NRF2, encoded by *NFE2L2,* and *KEAP1* are two central transcription factors involved in the cellular response to ROS production (reviewed by Itoh et al.[10]). KEAP1 binds NRF2 and targets it for degradation thereby inhibiting activation of NRF2 target genes. Interestingly, NRF2, a master regulator of antioxidant expression, is expressed at levels comparable to *ex vivo* cells, but the inhibitory protein is expressed to much higher levels in cultured cells. This would suggest that at transcript level, the cultured cells are not showing a canonical response to offset increased ROS levels, which could mean that the ISR is activated by a different source of stress.

Three-fold higher expression of *ATF6B* in cultured compared to *ex vivo* cells could serve as an indication that the unfolded protein response (UPR) and ER stress are driving activation of stress pathways. Unlike other members of the ATF family, *ATF6B* is specifically activated by ER stress and UPR.[11] We did not explicitly measure these potential sources, but with the large amount of hemoglobin that needs to be formed during erythroid differentiation it is conceivable that misfolded proteins arise. Implicitly, the HPLC analysis picking up HbH, a tetramer of beta-globin chains, shows there is incorrect globin pairing in cultured cells (**chapter 2**). Although these molecules can precipitate to form Heinz bodies a it does not confer misfolded Hb molecules.

Another source of cellular stress that would specifically relate to our culture setup is triggering of the UPR / ER stress branch of the ISR in reaction to lipid imbalances in the medium [12]. This mode of activation has been described to occur in hepatocytes [13] and would fit with the observation from our own preliminary metabolomics experiments that show lipid precursors are rapidly depleted from Cellquin at the start of culture.

Notably, besides *HBG* and genes associated with the ISR, upregulated genes in cultured cells included core regulators of erythropoiesis and hemoglobin switching *KLF1, GATA1* and *GATA2*. While higher expression of the core regulators could help explain higher expression of many of the other genes, a notion that is supported by enrichment of KLF1 and GATA1 ChIP targets in the set of 2000 genes. This observation would be at odds with increased HbF expression: LRF and BCL11A also showed increased expression, to a maximum of 2-fold higher compared to *ex vivo* cells.

At face value the observation on the increased expression of the core regulators and the observation on increased activity of ATF4 could indicate that the HBF inducing effect of stress-associated pathways overrides that of the core regulators of hemoglobin switching in erythroid cultures.

Contradictory to this notion, recent work on the role of the HRI-EIF2a-ATF4 signaling axis cells suggests that it activates BCL11A to repress HbF expression [14]. In this study increased ATF4 is proposed to bind the +55 enhancer of BCL11A to provide direct regulation through chromatin looping to the promoter of BCL11A. The effect might be system specific as it could not be detected in mice.

Higher levels of ATF4 results from increased translation due to higher activity of heme-regulated EI2-kinase (HRI), which is the dominant EIF2 kinase in erythroid cells [15,16]. Unlike its downstream target HRI was not expressed at higher levels in our cultured cells. Combined with the discrepancy for ATF4 and the inability to detect its effect in mice this raises the question to which extent the repression of HbF by the HRI-EIF2a-ATF4 signaling axis is a biologically relevant mechanism or only present in specific culture conditions. Gene expression profiles on their own provide limited ground to discuss this question. Therefore, our observations on ATF4 and the globin regulators would have to be corroborated by protein measurements.

Even though combined RNA sequencing with proteomics approaches generally find mRNA abundance a decent predictor of protein levels in steady state conditions, the correlation between mRNA and protein become less apparent in dynamic processes, like the state transitions of differentiation (reviewed in Liu et al[17]). The temporal changes in mRNA and protein can start to diverge between up and down regulated genes due to different modes of regulation in the differentiation process. For example, higher activity of ATF4 would have to be confirmed by protein measurements, because altered translation efficiency resulting from EIF2 phosphorylation is central to the regulation of its protein levels [6]. Similarly, differences in KLF1 and GATA1 between *ex vivo* and cultured cells at transcript level might not reflect differences in protein levels, since a disconnect between the two measurements for these transcription factors was suggested in preliminary work that combines transcriptome and proteome analysis during erythroid differentiation [18]. Taking into account that both KLF1 (**chapter 4**) and GATA1 can have dose dependent effects on gene regulation the necessity to confirm the observed differences at protein level becomes even more apparent [19,20].

Collectively, the extended characterization of the Cellquin based culture setup suggests a role for stress associated pathway in inducing HbF expression. While we cannot rule out that this results from increased ROS levels in cultured cells, activation of the stress response through the unfolded protein response and ER-stress would be a better fit with the observations at transcript level.

In addition to this mechanistic insight, this culture system facilitates the analysis of human erythroid progenitor cells by reducing the cost of the culture medium, providing high yields, and robustly recapitulating erythroid characteristics. At their current scale clinical safety tests have become feasible, which brings their use for transfusion of beta-hemoglobinopathy patients or other therapeutic applications a step closer.

Environmental signaling can repress fetal hemoglobin expression

Although our culture system lacks the 3D-organisation that makes up the bone marrow niche, its components constitute an environment where a mix of signals results in cells functionally equivalent to *ex vivo* cells. Most of the success with generating cultured erythroid cells comes from exploiting the general signaling cascades that initiate from EPO and SCF signaling.[21] While *in vivo* erythropoiesis requires membrane-bound SCF for erythroid expansion, supplying the soluble form along with EPO and dexamethasone is sufficient for robust expansion of progenitor cells *in vitro*.[22,23] The cell intrinsic capacity for erythropoiesis in response to a limited set of signals makes it difficult to delineate the role of the erythroid niche, for which erythroid progenitors and central macrophages are considered the main constituents. Despite the identification of various functional interactions between these cell types, conflicting reports on the requirement of these interactions make it difficult to determine their exact contribution to steady state erythropoiesis. A major contribution of niche interactions appears to come from providing support in periods of stress erythropoiesis [reviewed by [24,25]].

Addition of dexamethasone to the Cellquin culture medium induces signaling cascades that are normally active during stress-erythropoiesis.[22,26] In addition, previous characterization of the culture system using PBMCs has shown that the monocyte fraction (CD14+) stimulates erythroid outgrowth and obtains a central macrophage like phenotype in response to dexamethasone.[27,28] In follow up to these previous observations we employed the Cellquin culture system to investigate the role of cell-cell interactions in regulation of HbF expression (**chapter 3**) and found that cultures started from peripheral blood mononuclear cells (PBMCs) express lower levels of HbF when compared to those started from hematopoietic stem- and progenitor cells (HSPCs; CD34+ fraction) alone. The HSPC derived cultures showed an increased proportion of cells that express both adult and fetal hemoglobin. Lack of this same effect in transwell cultures shows that the repression of HbF depends on cell-cell proximity and does not result from increased yield or selective outgrowth of HbA-only cells. Similar to our observation a recent publication that combined transcriptome and proteome analysis on the populations proposed that differences between the two result from transcriptional control at the level of the β-globin locus.[29] The lack of differences in the transcriptomes of the HbA and HbF populations, along with the repressive effect on HbF being most pronounced in co-cultures with the CD34+CD36- fraction suggests that potential changes to transcription influencing HbF expression occur in the earlier stages of erythroid specification.

Aside from a dependence on cell-cell proximity we did not find the mechanism through which HbF repression in PBMC cultures occurs. Following an argumentation where repression would result from an active process and combining this with a potential HbF

inducing effect of ROS or lipid imbalance by activating the integrated stress response pathway, one could argue that CD14+ cells repress HbF expression by shaping the *in vitro* environment. For example, CD14+ derived macrophages could scavenge access ROS molecules, provide lipid intermediates or balance access lipids for the early progenitors, which in turn reduces activation of the stress-response pathways. This would fit with an assumed role for the erythroid niche supporting stress erythropoiesis, by allowing cells to rapidly expand without triggering apoptosis as a result of disbalanced pro-survival and pro-apoptotic signals.[30]

Outside of a balancing effect of CD14+ derived central macrophages on the extracellular environment culture, it is interesting to speculate that cell-cell contact or proximity triggers a signaling cascade that actively represses HbF expression.

The importance of interaction between erythroid progenitors and central macrophages has been proposed from observations on embryonic lethality in mice lacking the erythroblast macrophage protein (*Emp*).[31] An inhibitory effect on HbF expression from active environmental signaling has been proposed from the observation of a fetal-to-adult globin switch of human erythroid precursors upon injection in mice [33]. A possible mechanism to relay such an active signaling event in human cultured cells would be locus priming, which has already been established during hematopoiesis with the "GATA-switch". One line of reasoning follows the premise that this switch results from gradual replacement of GATA2, required for genesis and early progenitor gene expression, by GATA1 at specific loci, which in turn drives lineage specific gene expression.[34–36]

While another line of reasoning posits spatiotemporal regulation of Gata factor levels rather than an Gata identity switch as driving force for functional erythropoiesis.[37] This spatiotemporal nature would fit with the observation that the HbF repression from the CD14+ fraction is most pronounced in co-cultures with early stage progenitors.

More specifically, at the hemoglobin locus the GATA factor and its co-factor FOG1 have been suggested to facilitate the proximity between the LCR and adult globin promoter region to regulate adult stage globin expression [38,39]. These same complexes have been shown to recruit NuRD-complexes to the distal promoter regions of the *HBG* genes providing a mode through which it could convey direct repression of the fetal stage globin genes.[40,41]

Future work, should first determine if the repressive effect indeed results from an active interaction between the CD14+ central macrophage like cells and the erythroid progenitors. If this can be established it would be interesting to see which signaling pathways initiate from the interaction and test whether it is the exchange of GATA factors or another form of priming that leads to increased HbF expression in the progeny of the CD34+ cells. This would be a first account of repressive niche effect, that shares a similarity to the fetal-to-adult globin switch that has been reported in the transfusion

of human erythroid progenitors in mice.[33] The identification of such a signaling cascade could provide novel means to interfere with previously described modes of HbF regulation.

Genomic variants affect fetal hemoglobin expression through altered KLF1 availability

Whereas it is not yet clear how signals from the environment (macrophage interactions) or generated intracellularly (activation of p38 and the ISR) control globin switching, it is clear that the transcription factor KLF1 has a central role in control of globin expression.

Our application of the culture model in combination with samples from the Maltese hereditary persistence of fetal hemoglobin (HPFH) cohort allowed the examination of the *KLF1* p.K288X variant on inducing HbF expression. While the individuals in this cohort share this haploinsufficiency inducing variant, their HbF levels show a wide variation.[20] Despite an overall increase of HbF levels in the cell culture the residual variation in HbF expression, found *in vivo*, was retained between individuals from the same family (**chapter 4**). Exploiting this characteristic, we assessed chromatin accessibility along with the transcriptome and promoter analysis in the cohort. Although the overall chromatin, including the *HBG* promoter had a more open structure in *KLF1* haploinsufficient individuals our findings suggest that the main driver for incomplete penetrance of the HbF phenotype is the rs3817621 SNP, which reduces the expression from the wildtype KLF1 promoter. We proposed that the *HBG* genes might be particularly susceptible to the resultant compound heterozygosity, because both repression of fetal stage globin and activation of adult stage globin could be affected by the microvariations in KLF1 levels.

An important consideration to these observations is that although the hypothesis holds for the cultured samples, *in vivo* the observation only generalized to the individuals from the original pedigree tested in 2010.[20] The individual with the SNP from the other pedigree was grouped as a "high" HbF donor *in vitro*, but it had the lowest levels of the HPFH cases *in vivo*. Interestingly, also the non HPFH relatives from this pedigree showed a higher induction of HbF over its *in vivo* levels; about 20-fold compared to ~5-fold in the other pedigree. From this observation one can speculate that there is yet an undetermined genetic factor that either renders individuals from this pedigree particularly susceptible to HbF inducing conditions in the culture setup, or results in lower HbF *in vivo*. Further testing of relatives, both with and without HPFH can help to determine this.

Whereas, the compound heterozygosity with the *KLF1* p.K288X variant shows the effect of KLF1 levels on the regulation of HbF expression, our analysis of *Nan*-mice, carrying the *Klf1* p.E339D variant, shows how altered DNA binding can affect HbF (**chapter 5**). Similar to the human *KLF1* p.E325K variant, that causes congenital dyserythropoietic anemia (CDA type IV), the *Nan* variant results in retained embryonic/ fetal globin expression

beyond the stage of primitive erythropoiesis and causes erythropoietic defects. The *Nan* and the CDA variant result from different mutations affecting an orthologous amino acid in the middle zinc-finger of KLF1.[42,43] It is conceivable that the lack of high levels of HbF expression in *Nan*-mice results from a variant specific effect that differs between a KLF1 protein with a lysine (CDA) and an aspartic-acid (*Nan*) in the DNA binding α-helix. Particularly, since recent work that recreated the E > K mutation in a mouse cell line suggests that the both CDA and Nan-variants share altered binding to the β-globin promoter, yet manifest different binding affinities for other target genes.[44,45]

Regardless if lower local transcription factor availability result from variants that directly alter binding affinity or those that affect availability through transcriptional regulation. the combinatorial effects of different variants on the penetrance of a trait underlines the importance of additional genetic screening when it comes to understanding disease mechanisms. Particularly, in the event of haploinsufficiency of transcription factors that serve as upstream nodes in regulatory networks.[46] While regulation of HbF by its core regulators can certainly be described in terms of a regulatory network, formally the simultaneous interplay of the ensemble of regulators and the 3-dimenstional chromatin structure at the locus has not been combined in a gene regulatory network model (GRN). Although this type of approach might not yield fundamentally new insights, it would provide a means to predict dose effects of transcription factor levels, like we observed here for compound heterozygosity for KLF1 or the proposed differences in binding affinity for two variants of *Klf1*. A better understanding of these dose effects would facilitate testing and development of combinatorial therapies that induce HbF expression.

Concluding remarks

In the general introduction I described how our knowledge on molecular biology in general and the hemoglobin switch in particular co-evolved with the tools that progressively allowed more comprehensive hypotheses on the nature of gene regulation to be formed. The strive to understand the mechanisms of hemoglobin switching was paired with a desire to control these mechanisms in order to design new treatments for β-hemoglobinopathies.

A key insight at the basis of molecular biology was to view genetic material as a means to transfer information.[47] From this premise genes could be reduced to the minimal unit required for transmission with the coding sequence, enhancers, distal control regions and higher organizational domains as main building blocks. For the hemoglobin genes this reductionist approach has resulted in a well-defined set of core regulators that are involved in their regulation. While this core set of regulators allows for a more directed approach in attempts to induce HbF expression it has also narrowed the search horizon; most efforts are directed at reversing direct repression by BCL11A. This leaves less

room for trying to understand the developmental interplay of the different regulators, which despite identification of many components is still incompletely understood. The chapters of this thesis underline how variable penetrance of HPFH in response to transcription factor microvariations, retained HbF expression as a result of reduced binding of KLF1 or lack of environmental cues serve as prime examples of research areas where treatment alternatives can still be found.

With the advance of multi-omics and analysis at single cell level, molecular biology and computer sciences are increasingly intertwined. Sharing their conceptual origin in information theory, computational analysis of multi-omics serves as an important tool to address research question in the current period. Like the use of natural language programming to help with literature review in the introduction of this thesis, it is not that previous tools have been rendered obsolete or require less moderation. Rather, these novel tools focus on identifying patterns within larger bodies of information that when combined reveal essential structures of the process. This type of approach is facing problems with respect to moderation between signal and noise in more complex regulatory networks. Still, in combination with the mechanisms that have been defined to control hemoglobin switching at the level of the locus it can result in a more holistic insight in the nature of developmental gene regulation at the level of the genome.

References

1. Corces, M. R. et al. Lineage-specific and single-cell chromatin accessibility charts human hematopoiesis and leukemia evolution. Nat. Genet. 48, 1193–1203 (2016).
2. Fabregat, A. et al. The Reactome Pathway Knowledgebase. Nucleic Acids Res. 46, D649–D655 (2018).
3. Pourfarzad, F. et al. Hydroxyurea responsiveness in β-thalassemic patients is determined by the stress response adaptation of erythroid progenitors and their differentiation propensity. Haematologica 98, 696–704 (2013).
4. Zhao, B., Mei, Y., Yang, J. & Ji, P. Erythropoietin-regulated oxidative stress negatively affects enucleation during terminal erythropoiesis. Exp. Hematol. 44, 975–81 (2016).
5. Aerbajinai, W., Zhu, J., Kumkhaek, C., Chin, K. & Rodgers, G. P. SCF induces gamma-globin gene expression by regulating downstream transcription factor COUP-TFII. Blood 114, 187–94 (2009).
6. Wek, R. C., Jiang, H.-Y. & Anthony, T. G. Coping with stress: eIF2 kinases and translational control. Biochem. Soc. Trans. 34, 7–11 (2006).
7. Harding, H. P. et al. An integrated stress response regulates amino acid metabolism and resistance to oxidative stress. Mol. Cell 11, 619–633 (2003).
8. Han, J. et al. ER-stress-induced transcriptional regulation increases protein synthesis leading to cell death. Nat. Cell Biol. 15, 481–490 (2013).
9. Li, L. et al. Nrf2/ARE pathway activation, HO-1 and NQO1 induction by polychlorinated biphenyl quinone is associated with reactive oxygen species and PI3K/AKT signaling. Chem. Biol. Interact. 209, 56–67 (2014).
10. Itoh, K., Mimura, J. & Yamamoto, M. Discovery of the Negative Regulator of Nrf2, Keap1: A Historical Overview. Antioxid. Redox Signal. 13, 1665–1678 (2010).
11. Haze, K., Yoshida, H., Yanagi, H., Yura, T. & Mori, K. Mammalian transcription factor ATF6 is synthesized as a transmembrane protein and activated by proteolysis in response to endoplasmic reticulum stress. Mol. Biol. Cell 10, 3787–3799 (1999).
12. Tam, A. B. et al. The UPR Activator ATF6 Responds to Proteotoxic and Lipotoxic Stress by Distinct Mechanisms. Dev. Cell 46, 327-343.e7 (2018).
13. Li, Z. Z., Berk, M., McIntyre, T. M. & Feldstein, A. E. Hepatic lipid partitioning and liver damage in nonalcoholic fatty liver disease: Role of stearoyl-Coa desaturase. J. Biol. Chem. 284, 5637–5644 (2009).
14. Huang, P. et al. HRI-regulated transcription factor ATF4 activates BCL11A transcription to silence fetal hemoglobin expression. Blood (2020) doi:10.1182/blood.2020005301.
15. Liu, S. et al. Haem-regulated eIF2α kinase is necessary for adaptive gene expression in erythroid precursors under the stress of iron deficiency. Br. J. Haematol. 143, 129–137 (2008).
16. Grevet, J. D. et al. Domain-focused CRISPR screen identifies HRI as a fetal hemoglobin regulator in human erythroid cells. Science (80-.). 361, 285–290 (2018).
17. Liu, Y., Beyer, A. & Aebersold, R. On the Dependency of Cellular Protein Levels on mRNA Abundance. Cell vol. 165 535–550 (2016).
18. Gillespie, M. A. et al. Absolute quantification of transcription factors reveals principles of gene regulation in erythropoiesis. bioRxiv 812123 (2019) doi:10.1101/812123.
19. Ludwig, L. S. et al. Altered translation of GATA1 in Diamond-Blackfan anemia. Nat. Med. 20, 748–753 (2014).
20. Borg, J. et al. Haploinsufficiency for the erythroid transcription factor KLF1 causes hereditary persistence of fetal hemoglobin. Nat. Genet. 42, 801–5 (2010).

21. Migliaccio, G. et al. Long-term generation of colony-forming cells in liquid culture of CD34+ cord blood cells in the presence of recombinant human stem cell factor. Blood 79, 2620–2627 (1992).
22. Bauer, A. et al. The glucocorticoid receptor is required for stress erythropoiesis. Genes Dev. 13, 2996–3002 (1999).
23. van den Akker, E., Satchwell, T. J., Pellegrin, S., Daniels, G. & Toye, A. M. The majority of the in vitro erythroid expansion potential resides in CD34(-) cells, outweighing the contribution of CD34(+) cells and significantly increasing the erythroblast yield from peripheral blood samples. Haematologica 95, 1594–8 (2010).
24. Heideveld, E. & van den Akker, E. Digesting the role of bone marrow macrophages on hematopoiesis. Immunobiology 222, 814–822 (2017).
25. Manwani, D. & Bieker, J. J. Chapter 2 The Erythroblastic Island. in Current topics in developmental biology vol. 82 23–53 (2008).
26. Migliaccio, G. et al. Humanized culture medium for clinical expansion of human erythroblasts. Cell Transplant. 19, 453–469 (2010).
27. Heideveld, E. et al. Glucocorticoids induce differentiation of monocytes towards macrophages that share functional and phenotypical aspects with erythroblastic island macrophages. Haematologica 103, 395–405 (2018).
28. Heideveld, E. et al. CD14+ cells from peripheral blood positively regulate hematopoietic stem and progenitor cell survival resulting in increased erythroid yield. Haematologica 100, 1396–1406 (2015).
29. Khandros, E. et al. Understanding Heterogeneity of Fetal Hemoglobin Induction through Comparative Analysis of F- and A-erythroblasts. Blood 135, (2020).
30. Dolznig, H. et al. Apoptosis protection by the Epo target Bcl-XL allows factor-independent differentiation of primary erythroblasts. Curr. Biol. 12, 1076–1085 (2002).
31. Soni, S. et al. Absence of Erythroblast Macrophage Protein (Emp) Leads to Failure of Erythroblast Nuclear Extrusion. J. Biol. Chem. 281, 20181–20189 (2006).
32. Javan, G. T., Salhotra, A., Finley, S. J. & Soni, S. Erythroblast macrophage protein (Emp): Past, present, and future. Eur. J. Haematol. 100, 3–9 (2018).
33. Kobari, L. et al. Human induced pluripotent stem cells can reach complete terminal maturation: in vivo and in vitro evidence in the erythropoietic differentiation model. Haematologica 97, 1795–803 (2012).
34. Iwasaki, H. et al. GATA-1 converts lymphoid and myelomonocytic progenitors into the megakaryocyte/erythrocyte lineages. Immunity 19, 451–462 (2003).
35. Galloway, J. L., Wingert, R. A., Thisse, C., Thisse, B. & Zon, L. I. Loss of Gata1 but not Gata2 converts erythropoiesis to myelopoiesis in zebrafish embryos. Dev. Cell 8, 109–116 (2005).
36. Doré, L. C., Chlon, T. M., Brown, C. D., White, K. P. & Crispino, J. D. Chromatin occupancy analysis reveals genome-wide GATA factor switching during hematopoiesis. Blood 119, 3724–3733 (2012).
37. Ferreira, R., Ohneda, K., Yamamoto, M. & Philipsen, S. GATA1 Function, a Paradigm for Transcription Factors in Hematopoiesis. Mol. Cell. Biol. 25, 1215–1227 (2005).
38. Jing, H. et al. Exchange of GATA Factors Mediates Transitions in Looped Chromatin Organization at a Developmentally Regulated Gene Locus. Mol. Cell 29, 232–242 (2008).
39. Vakoc, C. R. et al. Proximity among distant regulatory elements at the β-globin locus requires GATA-1 and FOG-1. Mol. Cell 17, 453–462 (2005).
40. Harju-Baker, S., Costa, F. C., Fedosyuk, H., Neades, R. & Peterson, K. R. Silencing of A -Globin Gene Expression during Adult Definitive Erythropoiesis Mediated by GATA-1-FOG-1-Mi2 Complex Binding at the -566 GATA Site. Mol. Cell. Biol. 28, 3101–3113 (2008).

41. Sher, F. et al. Rational targeting of a NuRD subcomplex guided by comprehensive in situ mutagenesis. Nat. Genet. 51, 1149–1159 (2019).

42. Heruth, D. P. et al. Mutation in erythroid specific transcription factor KLF1 causes Hereditary Spherocytosis in the Nan hemolytic anemia mouse model. Genomics 96, 303–307 (2010).

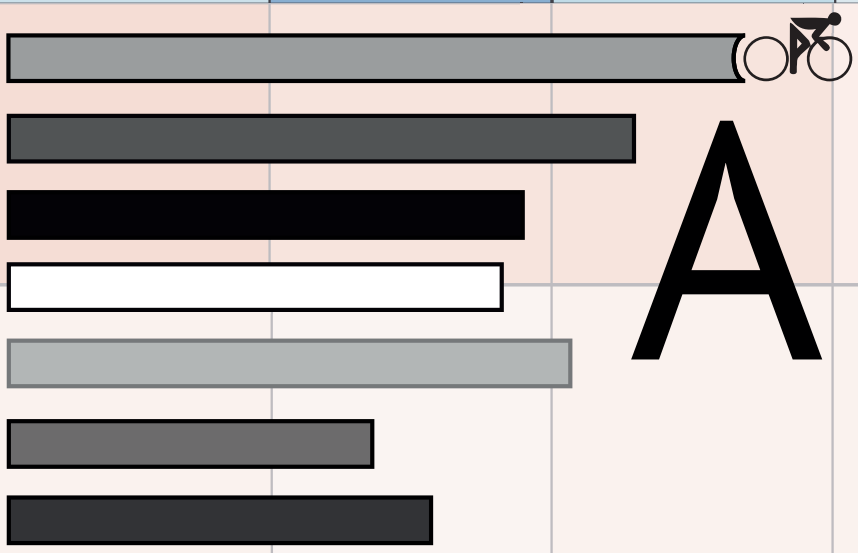
43. Siatecka, M. et al. Severe anemia in the Nan mutant mouse caused by sequence-selective disruption of erythroid Krüppel-like factor. Proc. Natl. Acad. Sci. U. S. A. 107, 15151–15156 (2010).

44. Ilsley, M. D. et al. Corrupted DNA-binding specificity and ectopic transcription underpin dominant neomorphic mutations in KLF/SP transcription factors. BMC Genomics 20, 417 (2019).

45. Singleton, B. K. et al. Mutations in the second zinc finger of human EKLF reduce promoter affinity but give rise to benign and disease phenotypes. Blood 118, 3137–3145 (2011).

46. Raj, A. & van Oudenaarden, A. Nature, Nurture, or Chance: Stochastic Gene Expression and Its Consequences. Cell vol. 135 216–226 (2008).

47. Cobb, M. 1953: When genes became 'information'. Cell vol. 153 503–506 (2013).


A

Appendix

Summary
Samenvatting
List of publications
PhD Portfolio
Curriculum Vitae

Summary

Hemoglobin molecules allow red blood cells to transport oxygen throughout our bodies. As a result of altered gene expression human hemoglobin molecules subsequently adopt embryonic, fetal and adult conformations. This phenomenon is often referred to as hemoglobin switching and is the main topic of this thesis.

As I introduced in the general introduction, the high abundance of and easy access to red blood cells in the human body/blood has placed studies on the hemoglobin molecules in health and disease, at the forefront of developments in molecular biology. Hereditary anemias, such as β-thalassemia or sickle cell disease, were linked to genetic variations that result in malformed or mal expressed hemoglobin chains. The observation that expression of fetal hemoglobin ameliorated disease symptoms in the event of these anemias (also referred to as β-hemoglobinopathies) intensified the interest in the mechanisms of gene regulation during hemoglobin switching.

The work in this thesis aims to advance our understanding of regulation of fetal hemoglobin expression by focusing on different aspects of erythropoiesis, the development from stem cell to red blood cell. Next to characterization of fetal hemoglobin expression in adult cultured cells, I describe the contribution of niche signaling and genomic background on the level of fetal hemoglobin in adults. I subsequently assessed the role of KLF1 in transcriptional regulation of hemoglobin switching in humans and in mice. Ultimately, this will contribute to development of treatments for β-hemoglobinopathies.

In the second chapter I describe a culture system that allows generation of large number of red blood cells. My colleagues and I optimized a previously described system by developing a proprietary culture medium, Cellquin, that uses GMP-grade components and is compatible with use in G-Rex bioreactors. This allows the scale up required for generating sufficient numbers of cells required for transfusion. The resultant cells are functionally and morphologically comparable to *ex-vivo* red blood cells (i.e. freshly isolated red blood cells). When these cells prove robust in clinical trials, they could provide a basis for transfusion products for hemoglobinopathy patients.

While matching most characteristics, cultured cells deviate from their *ex-vivo* counterparts by expressing higher levels of fetal hemoglobin. As a result, they provide an opportunity to study the effect of cell-cell interactions on the expression of fetal hemoglobin in this setup. Previously, transfusion experiments with human fetal globin expressing cells in mice provided proof of concept of induced switching to adult globin expression through signaling events. Similarly, we observed repression of fetal hemoglobin expression in adult cells that were grown in contact with the CD14+ monocyte fraction of PBMCs (i.e. peripheral blood mononuclear cells). In the presence of dexamethasone in the culture medium this cell fraction obtains characteristics similar to

the central macrophages, a component of the erythroid niche *in vivo*. Although we were unable to demonstrate the mechanism of the interaction beyond cell proximity, we did demonstrate temporal separation between the potential signal from the macrophage like fraction and repression of fetal hemoglobin expression.

Chapter four and five focus on the role of transcription factor KLF1. Aside from controlling expression of genes essential for erythropoiesis, KLF1 is part of a core network of transcriptional regulators that regulate expression of different globin genes. KLF1 activates adult hemoglobin expression by inducing interaction between the promoter and the locus control region. In addition, it controls the repression of fetal hemoglobin through activation of direct repressors, BCL11A and LRF. Conversely, lower KLF1 expression that results from gene expression from one instead of two alleles, ie. haploinsufficiency, causes higher levels of fetal hemoglobin. However, previous characterization of a Maltese pedigree with the same haploinsufficiency genomic variant reported different levels of fetal hemoglobin between individuals. Revisiting this donor cohort, we observed that the functional KLF1 allele contained an additional genomic variant in individuals with the highest fetal globin levels. The variant is located in the promoter of the gene and modulates KLF1 levels, which in turn affects globin expression.

Instead of focusing on the effect of genomic variants determining transcription factor levels, Chapter five explores effects of variants that result in altered binding characteristics of KLF1. We used the Nan-mouse model to characterize pre-natal hemoglobin switching in presence of the *KLF1* p.E339D variant, which has an altered DNA binding domain. The location of the amino acid substitution in the mouse model is the same as the one found in human CDA-IV, a severe anemia that features high fetal hemoglobin expression in adults. Although human and mouse hemoglobin switching differ in timing and the exact chains that are expressed, we did observe a phenotypic similarity to CDA-IV with retained expression of embryonic and fetal hemoglobins. However, quantitatively the embryonic/fetal hemoglobin levels did not increase as high as they do in CDA-IV. In addition, $KLF1^{nan}$ cells displayed lineage infidelity during *in vitro* erythroid differentiation indicated by cells expressing pan-myeloid and megakaryocytic markers. In line with the findings presented in Chapter four the simultaneous phenotypic overlap and contrasts between transcription factor variants makes the point for extended characterization of transcription factor variants in human disease.

The general discussion in Chapter six continues from the main observations in this thesis. I explore the possible influence of the integrated stress response on fetal hemoglobin expression in cultured cells, highlights the potential for environmental signals to affect globin gene expression and supports the case for modeling of the transcription factor network that affects regulatory interactions at the globin locus.

Samenvatting

Rode bloedcellen en hemoglobine

Rode bloedcellen zijn de meest voorkomende cellen in ons lichaam. Hemoglobine moleculen geven de cellen hun rode kleur en zorgen voor zuurstof transport door het lichaam. Hemoglobine moleculen bestaan uit twee paren van gelijke eiwitten, ook wel de globine ketens genoemd. ‚De combinatie van deze gepaarde ketens verandert tijdens de ontwikkeling. Zo nemen hemoglobine moleculen achtereenvolgend een embryonale (embryonale globine ketens met foetale globine ketens), foetale en volwassen (respectievelijk een combinatie foetale met volwassen globine keten en twee paar volwassen globine ketens) conformatie aan. Aan de wisseling tussen verschillende conformaties wordt vaak gerefereert met de engelse term “Hemoglobin switching” en is het hoofdonderwerp van deze thesis.

Hemoglobine genen als model voor regulatie van genexpressie

In de algemene introductie leg ik uit hoe de enorme hoeveelheid rode bloedcellen in ons lichaam ervoor heeft gezorgd, dat veel van de eerste ontdekkingen in moleculaire biologie en geneeskunde zijn beschreven aan de hand van deze cellen. Zo werd al vroeg ontdekt dat een tekort aan rode bloedcellen, anemie, bijvoorbeeld als gevolg van β-thallasemie en sikkelcelziekte, konden worden toegeschreven aan een genetisch defect. En was de locatie op het DNA als eerst bekend voor de hemoglobine genen. De ontdekking dat foetale hemoglobine genen kunnen compenseren voor het ziektebeeld bij β-thalassemie en sikkelcelziekte, die resulteren van een defect aan de volwassen globine genen, zorgde ervoor dat interesse in de regulatie van deze genen toenam. Sindsdien zijn er verschillende mechanismes van de regulatie van globine genen ontdekt met als ultiem doel een behandeling te vinden voor anemie.

Doel van het onderzoek

Het onderzoek in dit proefschrift draagt bij aan ontwikkelen van behandelingen voor erfelijke anemien door te kijken naar de regulatie van de globine genen tijdens erytropoëse, het proces waarin stamcellen differentiëren naar rode bloedcellen. Naast onderzoek aan de expressie van foetale globine genen in kweekcellen uit bloed afkomstig van volwassenen, hebben wij gekeken naar de rol die de cellulaire omgeving (niche) en genetische achtergrond spelen bij expressie van foetale globine genen. Daarnaast hebben we in meer detail gekeken naar de rol van KLF1, een belangrijk regulator eiwit voor erytropoëse, in hemoglobin switching in mensen en muizen.

Grootschalig kweekbloed maken

Op dit moment worden patiënten met zware vormen van sikkelcelziekte of β-thalassemie herhaaldelijk behandeld met bloedtransfusies. Deze behandeling heeft als nadeel dat er na herhaaldelijke transfusies allo-immunisatie op kan treden, een afweerreactie van het immuunsysteem die geactiveerd wordt in reactie op de getransfuseeerde cellen. Hierdoor wordt hetsteeds lastiger om geschikt bloed te vinden voor dit soort patiënten. Transfusies met gepersonaliseerd kweekbloed zou in dit geval uitkomst kunnen bieden omdat we in het geval van lichaamseigen cellen geen immuunreactie verwachten. Tegen deze achtergrond optimaliseerden wij in het tweede hoofdstuk een methode die kan worden gebruikt om op grote schaal kweekbloed te maken en te testen in eerste klinische studies. Dit protocol werkt zonder dierlijke componenten en in kleine G-rex bioreactoren. Daarbij beschrijf ik hoe de experimentele rode bloedcellen in de meeste functionele en morfologische aspecten lijken oprode bloedcellen direct afkomstig uit bloed. In toekomstig onderzoek zal blijken of deze cellen robuust genoeg zijn om in transfusie te worden toegepast en daadwerkelijk zouden kunnen worden in gezet als behandeling tegen anemie.

Omgevingssignalen en regulatie van hemoglobine genen

Hoewel gekweekte bloedcellen lijken op de volwassen cellen waar ze van afstammen, vertonen ze ook afwijkende eigenschappen. Gekweekte cellen hebben bijvoorbeeld een hogere expressie van foetaal hemoglobine.

In Hoofdstuk drie ga ik dieper in op het verschil in het aandeel foetaal hemoglobine cellen tussen *ex vivo* rode bloedcellen en geweekte cellen.

Uit eerder onderzoek bleek dat menselijke foetale cellen direct volwassen hemoglobin tot expressie brachten op het moment dat deze cellen in muizen werden geïnjecteerd.

Deze ogenschijnlijke flexibiliteit in hemoglobine expressie inspireerde ons om te te onderzoeken of wespecifieke omgevingssignalen konden aanwijzen die zorgen voor een verhoogd aandeel foetaal hemoglobine in volwassen cellen. Zo vonden wij dat de hoeveelheid foetaal hemoglobine in gekweekte cellen lager is wanneer stamcellen in direct contact met een ander type witte bloedcel worden opgekweekt. Hoewel het niet is gelukt om exact te beschrijven welk mechanisme er aan de basis lag van deze interactie, hebben we wel laten zien dat het specifiek lijkt te gaan om een interactie met de stamcel die word over gedragen door cel-cel contact.

Effecten van KLF1 afwijkingen op hemoglobine genexpressie

In Hoofdstuk vier en vijf kijk ik naar de rol van de transcriptie factor KLF1 in de regulatie van hemoglobine genen . Deze transcriptie factor is essentieel voor verschillende processen binnen erytropoëse en is een van centrale actoren in regulatie van expressie van verschillende globine genen. KLF1 activeert het volwassen β-globine

gen door de promotor in nabijheid van de "locus control region" te brengen. Daarnaast representeert KLF1 expressie van BCL11A and LRF1, transcriptie factoren die op hun beurt foetaal γ-globine onderdrukken. Zo leiden lagere KLF1 niveaus, als gevolg van haplo-insufficiëntie (de expressie van een in plaats van twee allelen), tot hoger foetaal hemoglobine in volwassen cellen. Desondanks zijn er binnen een Maltese familie met dezelfde haploinsufficientie mutatie aan KLF1, verschillende niveaus van foetaal hemoglobine. Familieleden met de hoogste foetale hemoglobine expressen bleken ook een mutatie in het functionele allel van KLF1 te dragen, die het KLF1 niveau verder verlaagt. Deze bevinding onderstreept het belang van bredere genotypische screening in het geval van phenotypische variatie in reactie op dezelfde mutatie.

Waar de KLF1 niveaus in Hoofdstuk vier variëren als gevolg van een afwijking in individuele individuele genetische achtergrond, kan het ook voorkomen dat een transcriptie factor als gevolg van een genetisch afwijking verminderd kan binden aan bepaalde DNA sequenties. Dit is het geval in het muis-model met de neonatale anemie mutatie, ook wel nan-model genoemd, die wij in meer detail hebben onderzocht in Hoofdstuk vijf.

In dit model resulteert een genetische mutatie in een aminozuur substitutie op een centrale plek in het DNA-bindingsdomein van KLF1. In mensen resulteert een mutatie op de overeenkomstige locatie in congenitaal dysterythropoietic anemia type IV. Dit is een dodelijke vorm van anemie die als onderdeel van het ziektebeeld naast de expressie van volwassen globine ketens ook een hoge expressie van foetaal hemoglobine kent.

Het overblijven van foetaal hemoglobine in volwassen cellen, vonden wij terug in het nan-model. Het lijkt daarmee dat hoewel hemoglobin switching tussen mensen en muizen op bepaalde punten verschilt, anemie en het overblijven van foetaal hemoglobine gedeelde fenotypische eigenschappen zijn die samenvallen met veranderingen in DNA binding. Tegelijkertijd zijn er net als in Hoofdstuk vier verschillen in de phenotypes uiting van overeenkomstige mutaties aan KLF1. Nader onderzoek zal moeten uitwijzen of dit komt door intrinsieke verschillen tussen muis en mensen of door andere DNA bindingseigenschappen van de twee KLF1 varianten.

Tot slot

In de algemene discussie (Hoofdstuk 6) bespreek ik de belangrijkste bevindingen uit de thesis in een bredere context. Ik ga in op de rol van stress erytropoëse in gekweekte cellen, de invulling van de omgevingssignalen en hun invloed op foetaal hemoglobine expressie.


A

PhD Portfolio

Name: Steven Janko Heshusius
Promotor: J.N.J Phillipsen
Co-promotores: E.A. van den Akker W.F. van IJken, M.L. von Lindern

Courses and workshops	**Year**	**ECTS***
Sanquin Science Course	2014	0.7
Embl genome browser	2014	0.2
PhD teaching program: Biophysics & Biochemistry	2014	0.5
PhD teaching program: Genetics	2015	0.5
PhD teaching program: Special Topics Course on Chromatin	2015	0.3
PhD teaching program: Special Topics Course on Signaling	2016	0.3
MolMed: Gene expression data analysis using R: How to make sense	2016	1.2
out of your RNA-Seq/microarray data Sanquin		
Good Research Practice and Scientific Integrity	2016	0.3
LUMC: Advanced bioinformatics	2018	1
Linux practical	2018	0.2
Code and data management with Git	2018	0.2
Python programming	2018	1
Datacamp: Python for Data Science track	2018-2019	2.8
Datacamp: SQL fundamentals track	2019	0.8

Seminars and masterclasses	**Year**	**ECTS**
Department meetings	2014-2018	8.0
Journal Club	2014-2018	3.0
Landsteiner Lecture and Guest Speakers	2014-2018	2.0
Lansteiner Masterclasses	2014-2018	1.5
Sanquin Science day 2014-2018	2014-2018	2.5
MGC PhD Workshop, Dortmund, 2017	2017	1
MGC PhD Workshop, Leuven, 2018	2018	1

National Conferences	**Year**	**ECTS**
Dutch Society of Stem Cell Research		
8th & 9th annual meeting, Utrecht	2015-2016	0.6
Dutch Hematology Congres (DHC)		
9th through 11th meeting, Arnhem	2015-2017	1.8
Molecular Aspects of Hematopoietic Disorders		
2014 & 17, Rotterdam	2014, 2017	1.4
Sanquin Spring Seminars,		
2017 & 18, Amsterdam	2017-2018	1.2

International conferences	**Year**	**ECTS**
International hemoglobin switching meeting		
99th & 101st, Oxford	2016, 2018	2.8
Red Cells Gordon Research		
Seminar & Conference, Newport, RI, USA	2017	2.0
European School of Hematology, Paris	2018	0.8

Selected presentations	**Year**	**ECTS**
Sanquin department meetings	2014-2018	2.0
Journal Club	2014-2018	0.3
Hemoglobin switching meeting Oxford, UK	2018	0.5
Fetal hemoglobin expression in adult erythroid cultures is repressed by CD14+ cells. Poster.		
Dutch society for hematology. Papendal	2016	0.5
Epigenetic analysis of KLF1 haploinsufficiency in primary eryhtroblasts.		
Red cells Gordon conferences. New port, RI, USA	2017	0.5
Monocyte fraction represses fetal hemoglobin expression in cultured erythroblasts.		
Molecular aspects of hematologic disorders.		
Rottterdam	2017	0.5
The role of KLF1 in erythropoiesis: when less is more; hereditary persistence of fetal hemoglobin		
Dutch society for hematology. Papendal.	2016	0.5
Epigenetic analysis of KLF1 haploinsufficiency in primary ryhtroblasts.		

Teaching	**Year**	**ECTS**
Bachelor Student, Clara Borras Eroles, 6 months	2014-2015	1
High school “snuffel” stages, multiple days	2016	0.5
HLO student, Sharon Bouw, 6 months	2016-2017	1

Awards & Prizes	**Year**
Selected poster for Presentation Gorden Red Cell seminar	2017
Poster Prize Sanquin Science day 2017	2017

** 1 ECTS equals 28 hours workload*


A

Acknowledgements

De weg kruipt onder me door, het uitzicht op de zonnige bergketens naar alle kanten is adembenemend, of was het toch de inspanning? Nog een laatste paar trappen en ik ben er. De top lag al een tijd in het zicht, maar ik moest nog wel even doorzetten om hem te bereiken.

Sinds ik begon met fietsen 8 jaar geleden heb ik verschillende bergen bedwongen en het is moeilijk om geen parallel te zien met de weg naar m'n PhD de laatste jaren. Volledig opgaand in prachtige vergezichten leidt de weg soms naar nieuwe hoogtes, gevolgd door gevaarlijke en verraderlijke afdalingen waar geen eind aan lijkt te komen. Gelukkig dreef het vooruitzicht van nieuwe panorama's mij telkens weer op de weg naar boven. Hoewel je prima in je eentje kunt fietsen, is het net als in het onderzoek veel makkelijker en leuker wanneer je het samen doet. Daarom wil ik iedereen bedanken die mij de afgelopen jaren op mijn weg hebben vergezeld, aangemoedigd, uit de wind hebben gehouden of op nieuwe routes wezen.

Emile & Marieke ik ben blij dat ik de afgelopen jaren onderdeel van het rode bloedcel team uitmaakte in de buik van het IHEP peloton. Jullie passie en enthousiasme voor onderzoek zal ik altijd bij me dragen. Ook in de soms uitgestrekte dalen wisten jullie me uiteindelijk weer op een nieuw pad naar boven te leiden. De traan werd altijd gevolgd door een lach en ik ben blij dat we altijd zo open en eerlijk hebben kunnen blijven communiceren. Andere collega's vroegen zich weleens af wat er nou toch voor bijeenkomst bezig was als er weer eens luid gelach uit onze werkbespreking naar buiten kwam. Dank voor jullie goede begeleiding en de leuke tijd bij jullie in de groep waar het plezier voor onderzoek centraal stond.

Het zou me niet verbazen als deze lol en plezier in het onderzoek al net zo aanwezig was toen jullie nog in Rotterdam werkten. De presentatie over evolutionaire oorsprong van de hemoglobine genen, die **Sjaak** gaf in de avondsessie op de Hemoglobin switching meeting, bevestigde dit vermoeden te meer. Het lijkt me sterk dat ik ooit nog een zo gevatte presentatie over een wetenschappelijk onderwerp bij zal wonen. Of het nou was in KLF1 meetings, onze achteraf bezien te schaarse een-op-een gesprekjes of tijdens het afronden van m'n proefschrift ik ben jou, Sjaak, dankbaar voor het telkens bieden van de sturing die ik nodig had.
Wilfred dank voor je gedegen commentaar op de vroege versie van dit boekje en onze prettige samenwerking aan het manuscript van Hoofdstuk vier. Ik wil ook graag mijn dank uitspreken naar de leden van de leescommissie, **Jan Voorberg**, **Marjon Cnossen** en **Niels Galjart**, voor het uittrekken van de tijd om mijn proefschrift te beoordelen en van commentaar te voorzien.

I want to thank **Joana Goncalves** who helped me analyze the erythroid differentiation RNAseq time course from Chapter 2. Even though, I was unable to carry our collaborative efforts to publication, I learned a lot in this period and I am grateful that you found the time for guiding me in the final months of my PhD.

When you're cruising in a peloton there are always those who challenge you to form a group and jump away from the bunch. Sometimes just for a quick escape and sometimes you'll end up spending a lot of time together and really get acquainted.

Benoit you're certainly one of those fellow riders. Through our animated discussion both on the bike and of, you inspired me to seek a more active contribution to changing the way our society treats the planet. I really admire your ability to think big while at same time having the guts to start small and work towards a better world for us all. Combined with your impeccable logic and wit, I'm sure, you will succeed in bringing Recircled to fruition.

Esther na de eerste keer samen cellen opwerken op de vroege zaterdagochtend, wisten we elkaar steeds vaker te vinden. Van samen fietsen over een verlaten eiland of door hartje Central Park. Ook tijdens onze trip naar de VS zaten we nooit verlegen om een goed gesprek. Dankjewel dat je daarnaast ook telkens het goede voorbeeld gaf voor de volgende stap in een PhD.

Marea dan geen mede fietser, maar zeker een van die mensen waar ik zo nu en dan mee de groep ontsnapte om goede gespreken te voeren. Je weet met je aanstekelijke lach zelfs de serieuste gespreken op te vrolijken. Daarnaast maakte je mij wegwijs tijdens m'n eerste dagen op Sanquin en bood je de basis voor het KLF1 werk, dat jij als student van **Jesse** had overgenomen van **Sahar**. Dankjewel.

Toen ik begon bij Sanquin keken sommige mede fietsers vreemd op als ik zei, dat ik op een laboratorium werkte om te experimenteren met EPO. Die experimenten horen in de fietswereld toch hopelijk tot het verleden. In de wereld van de bloedbanken is het juist de toekomst, die door verschillende brede samenwerkingen mogelijk wordt gemaakt.

Ons lichaam maakt meer dan 2 miljoen rode bloedcellen per seconden. De hoeveelheden cellen die er nodig zijn voor transfusie zijn dan ook nauwelijks voor te stellen. Niet voor niets is het kweken van bloed een team effort. **Patrick**, **Marie-Jose**, **Erica**, **Marijke**, **Anna**, **Elina**, bedankt dat ik onderdeel van jullie team kon zijn en zo een bijdrage heb kunnen leveren aan dit prachtige product voor de toekomst. And thanks **Joan** for bringing in the big guns, using real bioreactors to get closer the enormous number of cells required to make a transfusion product.

Om erytropoëse en de rol van KLF1 goed te kunnen beschrijven was ook de samenwerking met Rotterdam onmisbaar. Ik had niet gekund zonder de experimentele en conceptuele input van **Thamar, Nynke, Anna, Silvia, Martijn, Tijs, Ileana**. Dankjuliewel voor deze fijne samenwerking.

This collaboration was strengthened further by the colleagues from Malta, **Laura** and **Joseph**. Thanks for the fruitful discussions and joint efforts to learn as much as possible from the generous and special donor cohort on Malta.

Daarbij gaat mijn dank uit naar **Barbera,** voor het sequencen van KLF1 bij InLu patienten, en **Ben Nota** voor de eerste analyses op de KLF1 happloinsuffientie data.

In support of this work **Clara** and **Sharon** helped by setting up the first experiments CRISPR for KLF1 in our group. Thanks for allowing me to learn more than I could have learned on my own.

Het kweken van bloed gaat niet zonder het juiste startmateriaal. Daarom ben ik ook dankbaar voor de afdeling-brede samenwerking die, gecoördineerd door **Carlijn**, **Marion** & **Manon**, samen met iedereen op het CD34-schema ervoor zorgde dat er altijd hematopoietische stamcellen beschikbaar waren.

An interest in generating hematopoietic stem cells from induced pluripotent stem cells is part of the reason why I joined Emile & Marieke at Sanquin. Although I never quite managed to fully incorporate the IPS work in this thesis, I am thankful for the help from **Eszter**, **Tati**, **Franka** and **Marten** in trying to achieve this and allowing me to keep learning about this fascinating field.

Dat ik begon met een hoofd vol ideeën en kansen om het KLF1, IPSC en rode bloedcel onderzoek te combineren bleek ook wel uit mijn kennismaking met **Nurcan.** In een van onze eerste gesprek vroeg ze mij wat ik eigenlijk kwam doen als nieuwe postdoc in de Ery-groep. Ik bleek gewoon een nieuwe PhD en de daaropvolgende tijd bracht je mij ook gewoon de fijne kneepjes van de western blot bij. Dank voor je hulp en oprechte interesse en lieve woorden in de jaren die volgden.

Mijn eerste stappen in het biomedisch onderzoek zette ik op het AMC, waar **Rosa** mij wegwijs maakte in de fijne kneepjes van labonderzoek. Dit legde uiteindelijk het fundament voor dit proefschrift. Dankjewel. Ik bewonder je vrolijke kijk op het leven en ben blij dat we ook in mijn laatste weken op Sanquin een kantoor hebben kunnen delen en zo de cirkel voor mij rond maakte. **Christana**, als kamergenoot van Rosa en later als

directe collega op de afdeling wil ik jou ook graag bedanken voor je goede adviezen door de jaren heen. Natuurlijk moest ik op de meeste vlakken mijn eigen fouten maken, maar juist op die momenten herinnerden deze adviezen mij eraan dat ik op het goede pad zat. **Jeroen**, **Yuri**, en **Ronald** wil ik graag bedanken voor het vroege inkijkje in het postdoc leven en dat jullie mij het belang van leuke kamergenoten en collega's laten inzien.

Vaste kamergenoten was in de tijd op Sanquin een beetje lastig. Daar zie je nog zo'n parallel met het fietsen, net als in een peloton zaten we in de kantoortuin telkens naast iemand anders. Gelukkige was er aan leuke collega's geen gebrek binnen heel IHEP, elk moment van de dag.

In the early morning at my desk the sweet sounds of Italian conversation echoed through the corridors. Thanks, **Fiamma**, **Franscesca** and **Greta**, for allowing me to start those days with a holiday feeling.
And later during the day captured in fond memories of lunchtime conversations with **Corina**, **Ilse**, **Asena**, **Marieke**, **Aicha**, **Giso**, & **Roos,** casual chitchat in the hallways, or fun at different department drinks with **Rianne, Ammarina, Anna** throughout the years. Thanks for the good company.

Als ik het dan toch over gezelschap en eten heb, **Nahuel** als conference roomies stonden we regelmatig de beste route langs het buffet uit te stippelen. Thanks voor je droge humor en de leuke gespreken over sport en games (twee weken tot cyberpunk: hyped!).

Tamara in mijn beleving was jij de eerste PhD in de afdeling die het zo zat werd, dat je wel iets anders móest gaan. Maar misschien kreeg ik wel gewoon dat idee, omdat ik weer eens iets te luidruchtig rijstwafels zat te eten. Hoe dan ook bedankt voor het inzicht, dat je ook van je onderzoek kan houden en toch iets anders wilt gaan doen.

Kat you've probably paved more roads for me than I realize, for which I'm really thankful. I picked up on you breadcrumb trail of programming for biologist and am glad I did.

Arthur ik denk met plezier terug aan onze vroege ochtend gesprekken over alles op tweakers. Dank voor altijd op tijd laten weten hoe laat het was en lopende kalender te spelen voor alles dat rijmt op gemorrel. De G-unit corner groeide vlot om jou heen uit en de werkplek werd misschien niet productiever maar zeker leuker. Ik hou mooie herinneringen aan de ongein die dit bracht met **Robin, Thijs, Erik, Zoltan** & **Steven. Max** goed dat er nog iemand zo gek was op snoep als ik.

Speaking of treats and candy, what better occasional roommate to pick than **Felix.** Thanks a lot for helping me get through the 4 o'clock (or whenever really) energy dips.
Florencia een persoonlijke horoscoop verdient uiteraard een bedankje.

Han & **David** jammer dat we nooit met z'n drieën zijn gaan fietsen. Wel ben ik jullie beide dankbaar, voor het overnemen van het hemoglobine werk in de rode blode cel groep, dan wel voor de suggesties voor mijn onderzoek of tip om op een andere manier mijn wintertraining in te richten.

And off course maybe the biggest gearhead after I left, **Julien**. I'm still grateful for the t-shirt that you gave me. Today I can't help but think about the pretty classic moment your French namesake experienced this weekend (LBL lol).

Niets is zo fijn als eindeloos beschouwen op de honderden verhalen uit een lab, euh, race. Of misschien eigenlijk wel beiden. Ik heb enorm veel geleerd door de resultaten van mijn onderzoek te delen en te bespreken met collega's. Als **Monika**, **Sander** en **Aurelie** aanschoven voor de translation meeting, of tijdens de afdelingsmeeting waar de **IHEP staff**, **Ellen**, **Derk**, **Karlijn**, **Klaas**, **Martijn**, **Micha, Pleun** & **Regina**, ons werk van nieuwe inzichten voorzagen. Dank voor de waardevolle input op dit werk en de wijze lessen.

Deze wijze lessen nam ik mee naar m'n nieuwe baan, waar ik mijn plezier in data-analyse nu combineer met onderzoek naar de sociaaleconomische aspecten die spelen in de omvorming van de energievoorziening in Nederland. **Peter**, **Rolf**, **Henriette** en alle nieuwe collega's bij **DNE Research** en **Good!** dankjulliewel dat jullie mij zo snel in jullie midden hebben opgenomen en ook dit boekje helpen mogelijk te maken. Ik ben blij dat ik me, samen met jullie, kan inzetten om de wereld een stukje duurzamer te maken.

Buiten werk bracht ik veel van mijn tijd door op de fiets. Bijvoorbeeld op zondagochtend naar Sanquin voorafgegaan door een rondje met m'n fietsmaten bij wv Alcmaria Victrix. **Bas**, **René** en **Tijs** bedankt voor de gezelligheid tijdens trainingen en rondom wedstrijden. Een belangrijke rol is hierbij ook weggelegd voor **Fred**, jij bracht het groepsgevoel terug binnen de grotere ploeg en hielp mij te ontwikkelen als tijdrijder, dankjewel.

Als ik m'n PhD traject vergelijk met een fietstocht, waren er ook een hoop toeschouwers. Wanneer je maar door raast over pieken en door dalen zie je deze soms minder dan je lief is, of ben je niet met je volle aandacht aanwezig. Des te meer wil ik mijn waardering uiten voor het feit dat jullie mij zijn blijven steunen.

Mijn vroegste herinnering aan wat ik later wilde worden was uitvinder, het zal rond een jaar of zes zijn geweest. **Derk** sinds die leeftijd trekken wij al samen op en maken de ene na de andere mijlpaal mee. Ik had me dan ook niet voor kunnen stellen jou niet als paranimf aan mijn zijde te hebben. Dankjewel voor je vriendschap. Op de middelbare school groeide ons dynamische duo uit naar de BBQ-bitches. Hoewel we de laatste jaren steeds minder BBQ's hebben, blijft onze vriendschap onverminderd. **Bart, Rutger**, **Sjoerd**, **Flip**, **Sebas**,en **Davy**, ik vind het geweldig om te merken hoe we de afgelopen 18 jaar samen door allerlei fases zijn gegaan, waarin we elkaar altijd onvoorwaardelijk blijven vinden. We daagden elkaar uit in nieuwe sportieve uitdagingen of gingen als vanouds op vrijgezellenfeesten. De gezamenlijke etentjes met de vriendinnen, inmiddels grotendeels echtgenotes of verloofdes; **Shay**, **Elizabeth**, **Vera, Kim** en **Phillipa** zijn altijd meteen memorabel en in het afgelopen jaar zwijmelen we met z'n allen weg nu de eerste gezinsuitbreiding het familiegevoel compleet maken. Dankjulliewel.

Het was even wennen toen ik bij de **familie Kok** aanschoof 13 jaar geleden. De gesprekken aan tafel gingen direct in de hoogste versnelling en alles kwam voorbij. Van nationale politiek, de relaties van de klasgenoten van vijf jaar eerder, tot de beslommeringen van alle dag en welk gevoel je daarbij had. Alles kwam ongefilterd en direct uit het hart. **Gijs,** tijdens deze enerverende bijeenkomsten weet jij de boel altijd te vermaken en voor iemand die mij een maand geleden vroeg of ik nou ook een PhD deed, was je verrassend supportive. **Saar,** dat je alles in je macht doet om te helpen bleek wel toen Eef haar sleutelbeen brak vorige maand. Ik kon je nog maar net weerhouden op het vliegtuig naar Italië te stappen en toen we eenmaal thuis waren stond je telkens paraat. Dank voor je steun in die zware weken, maar ook voor alle gezellige etentje met jou en **Diek**. Lieve **Paul** & **Ans,** de gesprekken uit het hart over de beslommeringen van alle dag hebben mij erg geholpen mijn weg weten te vinden de afgelopen jaren. Dank daarvoor.

Mijn ouders en familie kan ik als fans van het eerste uur niet genoeg bedanken. Jullie maakte mij mee vanaf die kleine jongen met een uitvinders droom,tot de maatschappelijke betroken sportfanaat, die graag nadenkt over het zijn. Jullie hielpen me groeien, zagen me zoeken en vaak ook vinden wat ik zocht. **Hanneke**, al vroeg leerde jij me luisteren naar m'n hart. In de wetenschappelijke wereld waar ratio regeert is dat lang geen makkelijke opgave, maar deze vroegen lessen hielpen mij de ruimte te nemen als het kon en te volharden als het nodig was. Ze gaven mij het vertrouwen om ook na mijn periode bij Sanquin mijn hart weer te volgen en mij in te zetten voor wat mij het diepste raakt. **Bert**, van subtiele sturing d.m.v. the Infomation als kerstcadeau tot de geruststelling, dat jij net als ik het liefste dingen uit blijft stellen, maar het uiteindelijk altijd komt. Het waren vaak net de nudges die mij het gevoel gaven de laatste stappen van het afronden ook te kunnen zetten. Lieve **Oma**, van de familie zag ik jou het minste in de afgelopen jaren.

Toch wist je telkens wanneer we contact hadden precies waar ik & Eef mee bezig waren. Dank voor deze betrokkenheid, ik ben blij dat ik m'n promotie ook met jou kan delen. Lieve **Saskia**, met net zo'n honger naar sporten als ik zelf ben je onmiskenbaar mijn zusje. Ik vind het heel leuk om te zien dat je nu met **Arben** helemaal je plekje hebt gevonden en met **Mila** is het plaatje compleet. Als trotse oom wil ik jullie beiden bedanken voor jullie aanmoedigingen de afgelopen jaar in zowel m'n werk als m'n sport. **Tim,** lieve broer, dankjewel voor onze leuke gesprekken. Over werk, de maatschappij, of het laatste boek dat we hebben gelezen. Je goeie vragen weten me altijd uit te dagen om het juiste perspectief te vinden.

Allerliefste **Eva**, wauw wat een rit. Dat het een uitdaging ging worden zeiden we al hardop van tevoren, maar over de jaren van onze promoties bleek pas wat dit precies in hield. Dankjewel dat je altijd in mij bent blijven geloven en bent blijven overtuigen dat we dit samen aan konden. Zonder jouw aanmoedigingen en jouw fantastische voorbeeld promotie, afgelopen maand was ik hier niet gekomen. Ik ben trots hoe je ook met extra uitdagingen telkens je eigen torenhoge verwachting weet te overtreffen en ben blij, dat we de zware momenten de afgelopen jaren hebben kunnen afwisselen met genieten van de grote en de kleine dingen in ons leven, zoals onze fietstocht naar Italië of onze kroelende vogeltjes. Zo voelt het alsof we zelfs wanneer de weg omhoogloopt altijd de wind in de rug hebben. Dankjewel dat je ons leven zo adembenemend mooi maakt.


A

Curriculum Vitae

Steven Heshusius werd geboren in Rotterdam op 20-11-1986. Na het behalen van een gymnasiumdiploma aan het Stedelijk Gymnasium in Leiden, begon hij in 2007 met de opleiding algemene biologie aan de Universiteit van Amsterdam. Tijdens de daaropvolgende master in Biomedical sciences deed hij tijdens zijn eerste stage bij het AMC onderzoek naar de rol van co-stimulatoire molecule op Langerhans cellen. In zijn tweede stage bij Crucell droeg hij bij aan het optimaliseren van de productie van adenovirale vaccin vectoren. Hij schreef zijn afsluitende scriptie over het gebruik van geïnduceerde pluripotente stam cellen (IPSC) om verschillende hematopoietische lineages te maken. Zijn interesse voor deze cellen en processen van hematopoiese brachten hem bij Sanquin, waar hij zijn promotie onderzoek deed naar de regulatie van de hemoglobine genen. In zijn proefschrift, *Switching Gear: hemoglobin switching througout erythropoiese*, beschrijft hij de ontwikkeling van een kweekmedium voor grootschalige erythroide kweekculturen en de toepassing van dit systeem om expressie van foetaal hemoglobine in volwassen cellen beter te begrijpen. Na zijn promotieonderzoek verlegde hij zijn aandacht naar een andere vorm van onderzoek. Bij Dutch New Energy Research onderzoekt hij de verschillende markten die een rol spelen in het doorvoeren van de energietransitie.