KU Leuven
Biomedical Sciences Group
Faculty of Medicine
Department of Human Genetics

DOCTORAL SCHOOL
BIOMEDICAL SCIENCES

GENETIC MODULATORS OF SICKLE CELL DISEASE IN THE DEMOCRATIC REPUBLIC OF CONGO

Mamy Ngole Zita

Jury:

Supervisor: Prof. Gert Matthijs

Co-supervisors: Prof Aimé Lumaka and Prof Prosper Lukusa

Chair examining committee: Prof Kathleen Freson

Chair public defense: Prof Patrick Callaerts

Jury members: Prof Marc Jacquemin (KULeuven)

Prof Jean Marie Kayembe (UNIKIN)

Prof Solomon Ofori-Acquah (University of Accra, Ghana)

Prof François Boemer (Université de Liège)

Dissertation presented in partial fulfilment of the requirements for the degree of Doctor in Biomedical Sciences (PhD)

Leuven, 25 March 2023

Author : Mamy Ngole Zita
Cover design : Mamy Ngole Zita
Printed by : PROCOPIA

Table of content

List of abbreviations

AI	:	Arab-Indian
ARG1	:	Arginase
ARMS	:	Amplification-Refractory Mutation System
ASA	:	Allele-Specific Amplification
ASS1	:	Argininosuccinate synthetase 1
ASO	:	Allele-Specific Oligonucleotide
BCL11A	:	B-cell lymphoma /leukemia 11A
BEN	:	Benin
CAM	:	Cameroon
CAR	:	Central African Republic
CMMASS	:	Centre de Médecine Mixte et d'Anémie SS
CRISPR-Cas9	:	Clustered Regularly Interspaced Short Palindromic Repeats
DBS	:	Dried blood spot
DNA	:	Deoxyribonucleic Acid
DRC	:	Democratic republic of Congo
EDTA	:	Ethylenediamine tetra-acetic acid
FDA	:	US Food and Drug Administration
FLT1	:	Fms-related tyrosine kinase 1
GLP2R	:	Glucagon like peptide 2 receptor
Glu	:	Glutamic acid
GWAS	:	Genome -Wide Association Studies
HAO2	:	Hydroxyacid oxidase 2
Hb	:	Hemoglobin
HbA	:	Adult hemoglobin

HbF	:	Fetal hemoglobin
HbS	:	Sickled hemoglobin
HGVS	:	Human Genome Variation Society
HLA	:	Human Leukocyte Antigens
HPLC	:	High-Performance Liquid Chromatography
HRM	:	High-Resolution Melting
HSCT	:	Hematopoietic stem cell transplantation
HU	:	Hydroxyurea
IEF	:	Iso Electric Focusing
KLF10	:	Krüppel-like factor 10
LCR	:	Locus Control Region
Lys	:	Lysine
MAF	:	Minor Allele Frequency
MAGEB18	:	Melanoma-Associated Antigen B18
MAP3K5	:	Mitogen-activated protein kinase kinase kinase 5
MCV	:	Mean Corpuscular Volume
MIP	:	Molecular Inversion Probe
MLPA	:	Multiplex ligation-dependent probe amplification
NADP	:	Nicotinamide adenine dinucleotide
NBS	:	Newborn screening
NGS	:	Next Generation Sequencing
NO	:	Nitric oxide
NOS1	:	Nitric oxide synthase 1
OR51B5	:	Olfactory Receptor Family 51 Subfamily B Member 5
PDE7B	:	Phosphodiesterase 7B
PCR	:	Polymerase Chain Reaction

PGT	:	Preimplantation Genetic Testing
PHEX	:	Phosphate regulating endopeptidase X-linked
PNLCD	:	Programme National de Lutte Contre la Drépanocytose
POC	:	Point-Of-Care
PRS	:	Polygenic Risk Score
RBC	:	Red Blood Cell
RFLP	:	Restriction Fragment Length Polymorphism
ROI	:	Region of interest
ROS	:	Reactive Oxygen Species
RTD	:	Rapid Test Diagnostic
SALL2	:	Spalt Like Transcription Factor 2
SAR1A	:	Secretion Associated Ras Related GTPase 1A
SCA	:	Sickle cell anemia
SCD	:	Sickle cell disease
SCT	:	Sickle cell trait
SEN	:	Senegal
SNP	:	Single Nucleotide Polymorphism
SSA	:	Sub-Saharan Africa
ssMIP	:	Single Stranded Molecular Inversion Probe
TOX	:	Thymus high mobility group box protein
UCB	:	Umbilical cord blood
Val	:	Valine
VEGFA	:	Vascular endothelial growth factor A
WBC	:	White blood cells
WES	:	Whole Exome Sequencing

WGS	:	Whole Genome Sequencing
WHO	:	World Health Organization

1. GENERAL INTRODUCTION

1.1 Overview of the topic

1.1.1 Definition and history

Hemoglobin (Hb) is the most abundant protein of vertebrate red blood cells (RBC) where it functions as an oxygen carrier (Kabanova et al. 2009). The hemoglobin molecule is a tetramer comprising four subunits of two different types: two α globin peptide chains and two non-α globin peptide chains. While the α globin chains remain part of human hemoglobin from embryonic to adult life, the non-α chains are variable depending on the life period. The fetal Hb (HbF) has two α and two γ chains whereas adult Hb (Hb A) has two α and two β chains (α2 β2) (Schechter 2008).

Abnormal production or aberrant structure of the Hb molecule in the cause of a group of inherited disorders named hemoglobinopathies. The hemoglobinopathies are classified into two types: qualitative hemoglobinopathies such as sickle cell disease (SCD) are characterized by changes in the hemoglobin structure; whereas quantitative hemoglobinopathies or thalassemia exhibit a decrease in globin chain production (Weatherall et al., 2008). Depending on the peptide chain that is deficient, thalassemia is mainly classified as alpha thalassemia (the alpha chain is deficient) and beta thalassemia (the beta chain is deficient) (Marengo-Rowe 2007).

The concept of SCD encompasses multiple disorders characterized by the presence of an abnormal hemoglobin, either in the homozygous state or in double (compound) heterozygous state (Bunn 1997). Sickle cell anemia (SCA) applies to the homozygous SCD (Hb SS), with both alleles carrying the mutation. The sickle cell trait or sickle cell carrier (Hb AS) refers to the heterozygous state with one allele carrying the mutation and the other allele carrying the normal allele. The double heterozygous forms of SCD combine one allele with the S variant and the other with a non-S variant on the β-globin gene. The spectrum of mutations includes HbS/C (β7Glu>Val /β6Glu>Lys), HbS/OArab (β7Glu>Val/ β121Glu>Lys), HbS/β⁰ thalassemia (β7Glu>Val / -), HbS/E (β7Glu>Val / β26Glu>Lys) and others (Piel et al. 2013b; Rees et al. 2010). However, SCA is the most frequent and the most severe form of SCD (Rees et al. 2010).

Historically, the first references to sickle cell symptomatology date back thousands of years ago within some West African tribes (Konotey-Ahulu 1968). Although controversial to date, the famous Egyptian pharaoh Tutankhamun reportedly died from complications of sickle cell disease (Pays 2010). The first formal description of SCD was provided in 1904 in Chicago by Dr James B. Herrick, a cardiologist, and professor of medicine.

He observed unusual sickle-shaped RBC on microscopic analysis of blood samples from Walter Clement Noel, a 20 years old student from Grenada admitted at Chicago Presbyterian Hospital for anemia (Herrick 2001).

1.1.2 Epidemiology

Sickle cell disease is the most common hemoglobinopathy worldwide, both as a trait (heterozygous carriers, sickle cell trait, SCT) and as sickle cell anemia (homozygous individuals, SCA).

The disease is mainly observed in tropical regions of Sub-Saharan Africa, the Mediterranean basin, the Caribbean, Saudi Arabia, and India (Serjeant 1997). Although the disease was initially locally distributed, the frequency of SCA has currently spread across the world. Slave trade and voluntary migration have played a major role in changing the map of SCA (Modell 2008; Modell et al. 2007; Roberts & de Montalembert 2007). The prevalence of SCD is estimated to be 25 million patients worldwide and 12 to 15 million (48 to 60%) of them are living in Africa (Ware et al. 2017). Each year, approximately 312,000 babies homozygous for hemoglobin S are born worldwide, around 230,000 of these births occur in Sub-Saharan African (SSA) countries (Figure 1). The incidence of SSA is around 1% of total births (McGann 2014; Piel et al. 2013c). This incidence is expected to increase to 400,000 annual births by 2050 (Piel et al. 2013a). Nigeria is reported to have the highest prevalence of SCA, followed by India. The Democratic Republic of Congo holds third place in the world and second in Africa (Odame 2014; Piel et al. 2013b). Beside HbS, the HbC and HbE are other frequent hemoglobinopathies. The HbC is the second most frequent abnormal Hb in Africa after HbS, especially in West African countries (Ghana, Togo, Benin, Burkina Faso) (Piel et al. 2013b). The HbE is frequent in some regions of India, Myanmar and throughout Southeast Asia (Fucharoen & Winichagoon 2011).

The burden of SCA is very heavy in Africa to such an extent that SCA is responsible for 6 to 15% of child mortality. About 50 to 80% of SCA children in SSA die, often undiagnosed, before their 5th birthday due essentially to infections and anemia (Mulumba & Wilson 2015). This related morbi-mortality is so important that the disease was recognized as a global public health problem by the United Nations and the World Health Organization (WHO), respectively in 2009 and 2010. The 19th of June was instituted as the international sickle cell disease awareness day (United Nations 2009; WHO Regional Office for Africa 2010).

Several theories have been proposed to explain the geographical distribution of the disease. The most plausible theory to date remains the survival advantage conferred by the sickle-cell trait (SCT) against severe forms of malaria given the region of high SCA prevalence are also characterized by malaria endemicity (Elguero et al. 2015). Malaria is a life-threatening condition caused by Plasmodium parasites that are transmitted through bites of infected female Anopheles mosquitoes. In its severe forms, malaria is accompanied by severe anemia, respiratory distress, and neurological damage, that may lead to death. Multiple factors support malaria endemicity in tropical regions, including the tropical climate. This climate promotes the breeding of Anopheles mosquitoes.

Contrary to the disadvantage of SCA mutation in homozygous individuals, there is a selective advantage of HbS carriers against malaria. This is illustrative of balanced polymorphism, in which SCA patients are likely to die due to the disease at an early age, whereas SCT individuals are prone to survive till reproduction age (Luzzatto 2012). Various mechanisms are reported to explain resistance to severe malaria in the sickle cell trait. Immunological pathways are the most involved, including higher splenic phagocytosis of Plasmodium-infected erythrocytes and a reduced expression of Plasmodium falciparum Erythrocyte Membrane protein 1 (PfEMP1), an endothelium cytoadherence protein at the surface of HbS RBC infected by *Plasmodium falciparum* (Baruch et al. 1997; Luzzatto et al. 1970; Williams et al. 2005). Additionally, more recently reported mechanisms include the translocation of HbS-specific parasite-growth inhibiting microRNAs as well as inhibition of intra-RBC parasite growth related to HbS polymerization in hypoxic microcirculation (Archer et al. 2018; LaMonte et al. 2012). The entire range of mechanisms is not fully elucidated yet (Depetris-Chauvin & Weil 2018).

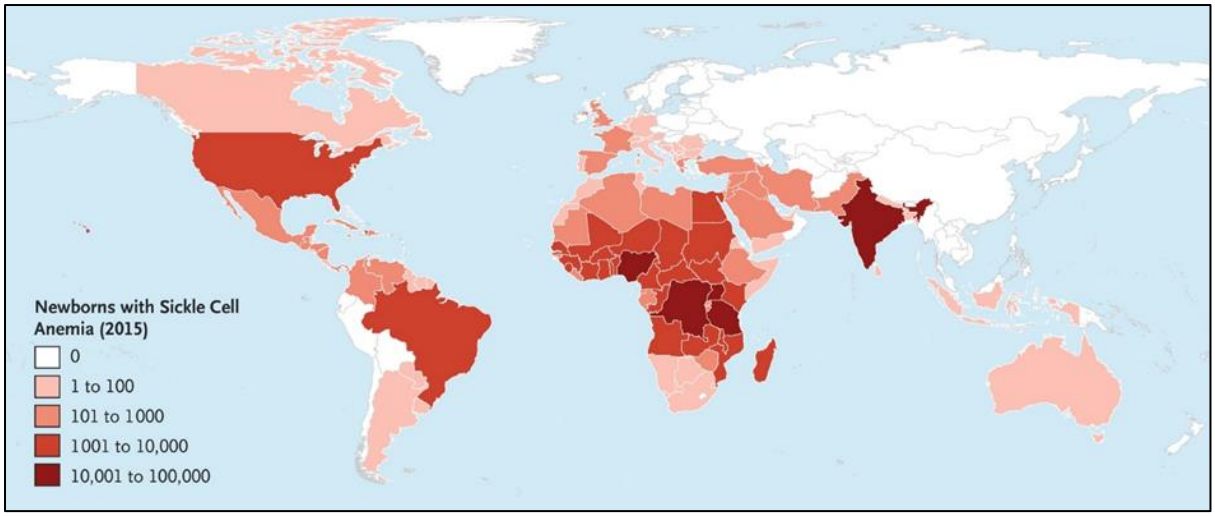

Figure 1. Sickle cell anemia newborn incidence in 2015
(Reproduced with permission from Piel FB, Steinberg MH, Rees DC. Sickle cell disease. N Engl J Med 2017 ; 376:1561-1573), Copyright Massachusetts Medical Society). Note that DR Congo features among the countries with the highest incidence in Africa and in the world.

1.1.3 Etiology and pathophysiology

Etiology

Sickle hemoglobin (HbS) is a structural variant of the hemoglobin A (HbA), in which the normal beta-globin subunit is replaced with a mutant form of the beta chain called ***Beta-S chain*** (βS). This variant resides in the seventh codon of the hemoglobin β gene (*HBB*), which encodes for a glutamic acid (E), in which the second nucleotide (G**A**G), an adenine, is substituted with a thymine (G**T**G) (rs334). At the protein level, the glutamic acid (E) is substituted by a valine (V) (Sundd et al. 2019). According to the current HGVS nomenclature, this variant is described asNM_000518.5(HBB):c.20A>T;p.Glu7Val (https://www.ncbi.nlm.nih.gov/clinvar/113462171/). However, the variant is commonly referred to as E7V.

The genes coding for the different Hb peptide chains are clustered into two loci, the alpha-like cluster on chromosome 16 and the beta-like cluster, also known as HBB locus, on chromosome 11. There are six globin genes distributed over both loci : the zeta (ζ) and alpha (α) genes on chromosome 16, and epsilon (ε), gamma (γ), delta (δ) and beta (β) genes on chromosome 11 (Figure 2) (Stamatoyannopoulos 2005). The genes at both loci share the same transcriptional orientation and are ordered in the same sequence in which they are expressed during development and life (Nussbaum 2007). Each gene in both α globin and β globin clusters is expressed during a specific period throughout life. The embryonic hemoglobin includes Hb Gower-1(ζ2ε2) , Hb Portland-1 (ζ2 γ2) , Hb Portland-3 (ζ2δ2), Hb Portland-2 (ζ2β2) and Hb Gower-2 (α2ε2). At the fetal stage, the HbF (α_2 γ_2) appears. It starts decreasing after birth to

around 1% of the total hemoglobin by the age of 6 months to 2 years (Maier-Redelsperger et al. 1994). The synthesis of adult Hb (HbA) (α2 β2) starts in the perinatal period and persists throughout life (Figure 2) (Stamatoyannopoulos 2005).

The sequential change observed in the expression of the globin genes named "globin switching" is under the control of a Locus Control Region (LCR) located 6 kb upstream of the ε gene. The expression of a selective β globin gene is regulated by interactions involving the LCR, the promoter of that active gene, and transcription factors acting as activators or repressors (figure 3) (Barbarani et al. 2021). To date, the γ/β globin switch appears more interesting because of its therapeutic implication in β hemoglobinopathies.

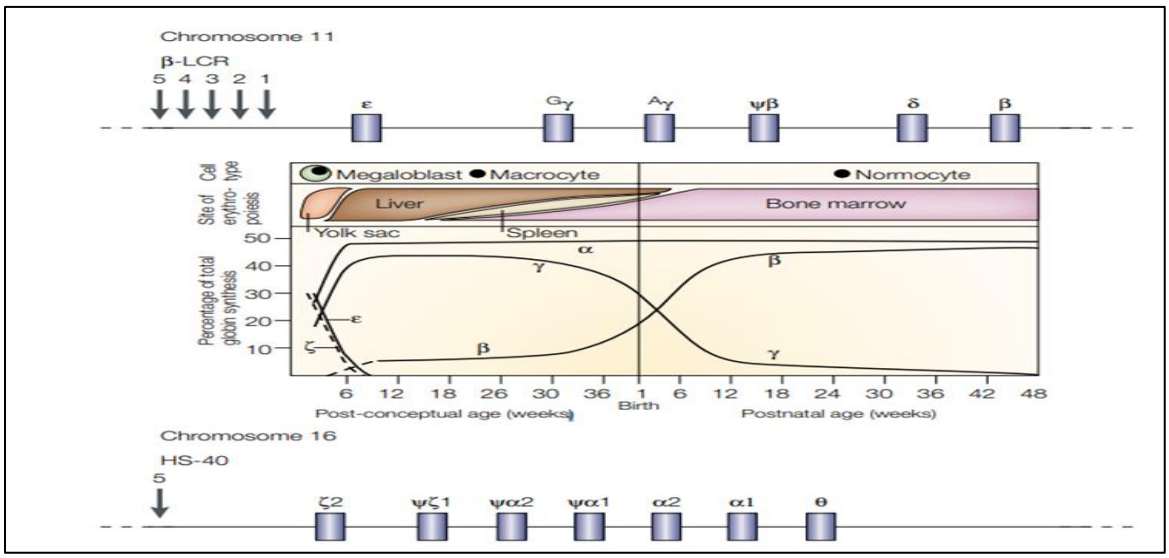

Figure 2. Schematic presentation of the β-globin and α-globin clusters respectively on chromosome 11 and chromosome 16

From Weatherall, 2001. Reproduced with permission from Springer Nature. Upper panel: schematic presentation of β-globin cluster on chromosome 11 with the locus control region (LCR) and the different globin genes. Middle panel: expression of the different Hb illustrates the first switch from embryonic (Hb Gower 1 (ζ_2 ε_2) and Hb Gower 2 (α_2 ε_2)) to fetal hemoglobin (HbF (α_2 γ_2)) followed by the second switch from fetal Hb to adult Hb (HbA (α2 β2)). Lower panel: schematic presentation of the α- globin cluster on chromosome 16.

The switching is tightly regulated by different transcription factors. SNPs in several other genomic regions in *cis* to the β-globin cluster have been associated with variation in clinical and biological expression (figure 3).

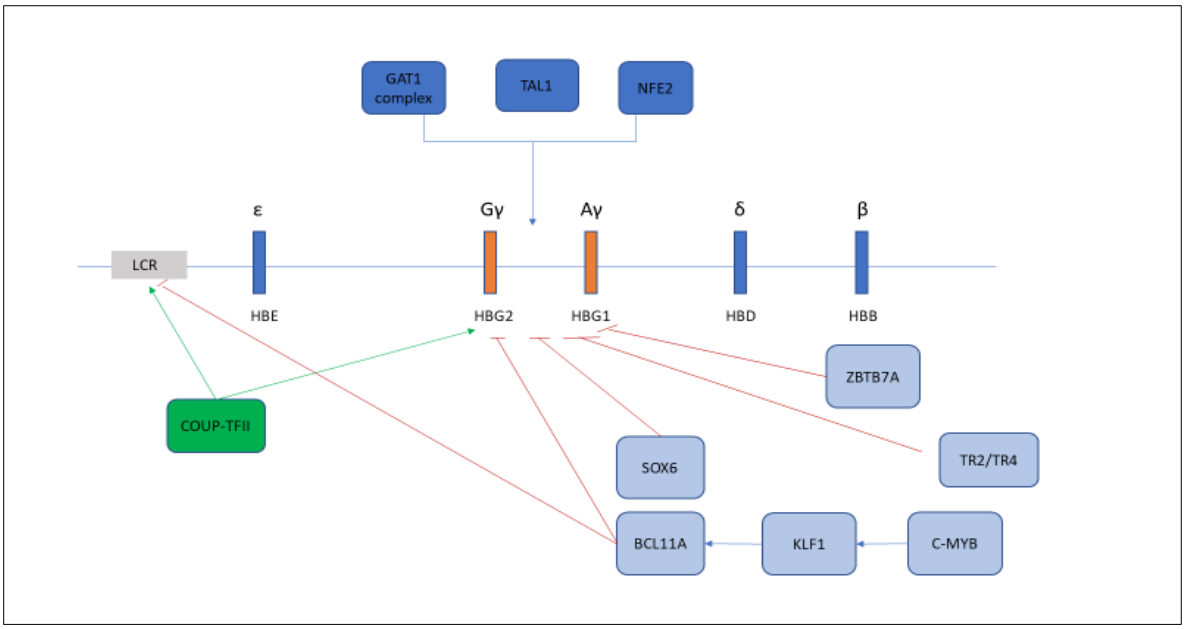

Figure 3. Transcription factors in β globin genes expression

Adapted from Barbarani et al., 2021. Upper panel: Transcription factors common for all embryonic, fetal and adult hemoglobin. Lower panel, right: Transcription factor for embryonic and fetal hemoglobin; left: γ globin-specific repressor complex with BCL11A

Pathophysiology

The SCA pathophysiology is centered around four main pathways: HbS polymerization, blood vessel obstruction, hemolysis and inflammation (Sundd et al. 2019). However, there is an intricacy that makes the separation of those pathways from one another difficult.

The starting point of SCA pathophysiology is the polymerization of HbS under hypoxia, dehydration, or acidosis conditions. The deoxy-HbS molecules stick together to form linear elongated insoluble polymers. This leads to the rigid sickle-shape deformation of RBC instead of physiological discoid deformable RBC. As a result, sickle RBCs are dehydrated cells that hemolyze prematurely (Zago & Pinto 2007).

The polymerization of HbS causes occlusion in the microcirculation by sickle RBCs. Intermediate mechanisms include increased blood viscosity; the unusual expression on sickle RBC membrane of adhesive and linking molecules such as phosphatidylserine (PS), basal cell adhesion molecule-1/Lutheran (BCAM-1/Lu), integrin-associated protein (IAP) and intercellular-adhesion-molecule-4 (ICAM-4), and the release of stress reticulocytes that express adhesive molecules as α4β1 integrin (VLA-4) and CD36 (Barabino et al. 2010; Kaul et al. 2009). Adhesive interactions between sickle RBCs, endothelial cells, leukocytes, and platelets contribute to the observed vaso-occlusion.

The excess of free Hb generated by intravascular hemolysis consumes nitric oxide (NO), thus reducing its bioavailability. The NO is not only a powerful vasodilator but also an inhibitor of adhesive molecules' expression on cellular membranes and of platelet activation (Gladwin & Kato 2008). The inflammation state consists of the activation of other blood cells (leukocytes and platelets), endothelial cells and the generation of reactive oxygen species (ROS) by several pathways involving some interleukins (Wang & Zennadi 2021).

Infections represent another important factor in the pathophysiology of SCA. Due to functional asplenia associated with a dysfunctional immune system by the age of five years old, SCA children are susceptible to infections by encapsulated germs such as Streptococcus pneumoniae, Salmonella, Haemophilus influenzae and Meningococcus (Battersby et al. 2010).

1.1.4 Clinical presentation

Due to its complex pathophysiology, SCA is considered to be a multisystem disease, beyond its original definition as a hematologic disease. According to the pathophysiological mechanisms, the clinical manifestations can be classified in four groups: hemolysis and hematological complications, vaso-occlusion, infection, and organ dysfunction (Figure 4) (Houwing et al. 2019; Makani et al. 2013). However, clinical expression of the disease is highly variable, depending on genetic, environmental, socio economic and climatic factors.

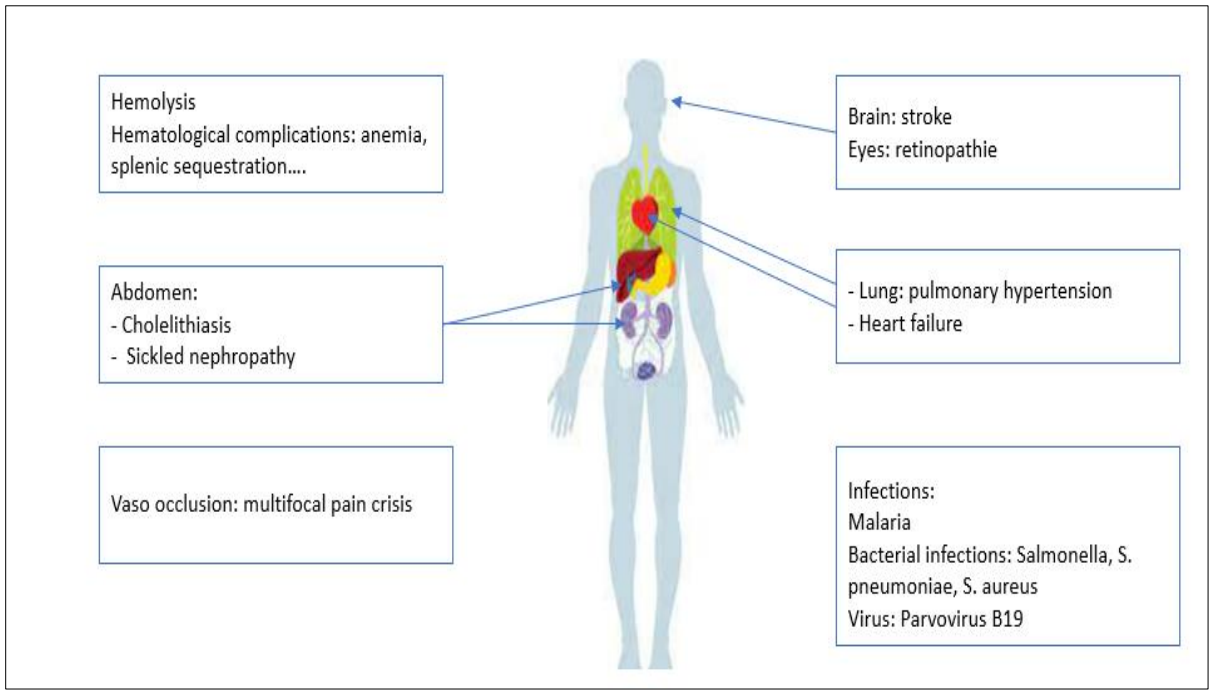

Figure 4. SCA clinical manifestations

Adapted from Makani et al., 2013.

Due to its inhibiting mechanical effect on HbS polymerization, HbF is among the main modulators of SCA severity (Akinsheye et al. 2011). The HbF level depends on the life period. The reduced HbF level after birth allows the onset of clinical manifestations of SCA. Therefore, in SCA patients high levels of HbF are reported to be associated with a reduction in the disease severity (Thein, 2008). Co-inheritance of α-thalassemia reduces SCA severity further by decreasing total Hb concentration in RBC, and consequently occurrence of HbS polymerization (Steinberg & Sebastiani 2012).

Non-genetic circumstances that promote the occurrence of hemoglobin polymerization like hypoxia, RBC dehydration, and acidosis are influenced by environmental factors (Tewari et al. 2015). Consequently, extreme temperatures and rain seasons with high humidity are among the environmental factors leading to an increased rate of acute pain episodes in SCA patients (Mekontso Dessap et al. 2014). Socio-economic factors such as nutrition, parents' or patients' incomes, and home environment are potential influencers of SCA patients' health outcomes (Rees et al. 2010).

1.1.5 Diagnostic approaches

The detection of the presence of HbS is the main approach to diagnosing SCA. This detection commonly utilizes the physicochemical properties of HbS, such as decreased solubility and sickling under deoxy conditions, mobility on an electric field, and rate of elution from solution onto adsorbents (Thom et al. 2013).

Different technical strategies for the detection of HbS may be applied depending on whether the test purpose is screening or confirmatory testing (Makani et al. 2013; Wajcman & Moradkhani 2011). The screening tests previously included the HbS solubility and the Emmel sickling test, whereas the confirmatory tests, available only in some specific laboratories, rely on electrophoresis, high-performance liquid chromatography (HPLC), and Mass Spectrometry (Joutovsky et al. 2004).

Screening tests

The HbS solubility test or Itano test

This test is based on the precipitation of deoxygenated HbS in a saturated salt solution, leading to observable turbidity of the solution (Arishi et al. 2021) . Despite its easy use and low cost,

the Itano test cannot differentiate sickle cell trait from SCA. Also, it is not indicated for newborn screening since high HbF levels lead to false negative results (Tubman & Field 2015).

The Emmel test or sickling test

This test investigates the sickling of RBC containing HbS under hypoxic conditions artificially created by using sodium metabisulfite. Unfortunately, HbAS cannot be differentiated from HbSS with this test (Okwi et al. 2009).

Confirmatory tests

Hb electrophoresis techniques

Generally, electrophoresis involves the separation of electrically charged molecules (Hb) under the influence of an electric field. Different HbS variants are easily identifiable using cellulose acetate and/or citrate agar at alkaline or acidic pH, respectively (Kumar & Derbigny 2019). The alkaline electrophoresis allows easy separation of HbA and HbS, whereas some other Hb variants such as HbS, HbF (high amount as in newborn), HbD, and HbG co-migrate. The acidic pH electrophoresis allows the separation of the co-migrated HbS (Clarke & Higgins 2000).

But those methods have the disadvantage of being time-consuming and labor-intensive, in addition to lacking accuracy for the quantification of low-concentrated hemoglobins such as the HbA2. They are now mostly used in combination with another more accurate method (Ou & Rognerud 2001).

Capillary electrophoresis separates Hb variants according to their electrophoretic mobility, the electrolyte's pH, and the electro-osmotic flow within capillaries (Chen et al. 1991). Fully automated equipment has become available for capillary electrophoresis. This equipment has the advantage of handling multiple samples simultaneously for analysis with reusable columns. The benefits of this technique are balanced by the cost and requirement of trained technicians for results interpretation (Cotton et al. 2013).

The isoelectric focusing (IEF) technique separates Hb variants across a pH gradient according to their isoelectric point (Reddy & Franciosi 1994). Due to its advantages such as an excellent resolution in separation of variants with a difference of at least 0.02 pH units in their isoelectric point, detection of abnormal hemoglobins in newborns as well as analysis of large blood samples series, especially dried blood spot, the IEF appears to be the choice method for

neonatal screening (Frömmel 2018). However, the main limitations of this method are a complex implementation, labor intensive, time consuming and interpretation by trained personnel (Clarke & Higgins 2000).

High performance liquid chromatography (HPLC)

With this chromatographic method, different hemoglobins are separated according to their specific retention time as they interact with the stationary phase (Sharma & Das 2016). The HPLC has a higher sensitivity to detect Hb variants than classic electrophoresis as well as a higher speed of analysis (about 3 minutes per sample), enabling the handling of a larger number of samples (Wajcman & Moradkhani 2011). However, the cost of the machine and its technical requirements limit access to resource-limited countries (Nair 2018).

Molecular methods to detect the E7V mutation

The molecular biology techniques consist in the identification of the βs mutation (Clark & Thein 2004). These techniques can be classified into primer-specific PCR-based methods, Restriction Fragment Length Polymorphism (RFLP), and sequencing methods (Arishi et al. 2021).

Primer-specific polymerase chain reaction (PCR)-based methods

These methods include the high-resolution melting (HRM), allele-specific amplification (ASA), amplification-refractory mutation system (ARMS), and allele-specific oligonucleotide (ASO) (Newton et al. 1989; Toye et al. 2018; Yue et al. 2014).

Restriction Fragment Length Polymorphism (RFLP)

The E7V mutation abolishes the recognition site of restriction enzymes such as DdeI, MnI I, and Mst II, resulting in observable bands of specific lengths depending on the number of affected alleles and the enzymes used (Wilson et al. 1982). Compared to others, this one is very easy to perform and less expensive, but it is time-consuming and difficult to automate (Saiki et al. 1985).

DNA sequencing methods

They provide the most comprehensive data for the HBB gene. Sanger sequencing remains the gold standard in term of sequencing. Sequencing allows the sequence of a gene to be deciphered

at single the base level. Usually, DNA sequencing methods, especially next generation sequencing (NGS), are indicated for identification of genetic modifiers related to disease severity, to drug response rather than SCD diagnostics (Fakher et al. 2007).

Point of care techniques

Recently, point-of-care (POC) devices for SCD detection have been developed. According to the WHO, an ideal POC device for SCD testing should meet the « ASSURED » criteria, with :

- A as Affordable, with an ideally cost of less than 1 $ per test
- S as Sensitive, particularly in the neonatal period with a high level of HbF
- S as Specific, differentiation of HbAS with HbSS, HbSC, and HbS-βthalassemia
- U as User-friendly, related to the ease of use of the device
- R as Rapid / Robust, in reference respectively to reduced turnaround time and device storage at room temperature
- E as Equipment free, no equipment to be added
- D as Delivered, available in high prevalence settings (McGann & Hoppe 2017).

The POC tests are based on several principles: qualitative detection of AA, AS, SS and SC profiles based on e.g. a lateral flow immuno-chromatographic assay using polyclonal (Sickle Scan™) or monoclonal antibodies (HemoTypeSC™); precipitation of insoluble deoxy-Hb on a paper-based support;

adapted cellulose acetate electrophoresis on a microchip support (HemeChip); discrimination between normal and sickled RBC based on difference of density in an aqueous multiphase system (AMPS) (Alapan et al. 2016; Kumar et al. 2014; Piety et al. 2016; Quinn et al. 2016; St John & Price 2014).

Despite advantages of those devices such as reliability in newborn period (e.g., Sickle Scan™), low cost, portable devices, timeliness (less than 30 minutes) and storage at room temperature; some limitations are pointed out such as possible misinterpretation due to visual interpretation of results, misidentification of SS individuals related to recent blood transfusion, cross-reactivity of polyclonal antibodies (Bond et al. 2017; Nnodu et al. 2019).

1.1.6 SCD newborn screening

The need for a newborn screening program (NBS) is based on the very high mortality during the first three years of SCA patients' life (Makani et al. 2013). The implementation of NBS in high-income settings has allowed a remarkable decrease in the morbi-mortality of SCA and improvement in SCA patients' survival (King et al. 2007; Quinn et al. 2010; Yanni et al. 2009). This was made possible by the integration of multiple activities into a workflow, including sample collection and transportation to the laboratory facilities where testing is performed, timely reporting of results, follow-up for diagnostic confirmation, and appropriate care for SCA-confirmed infants as well as parents' education.

Unfortunately, in low-income settings, particularly in SSA where the vast majority of SCA infants are born, only a few pilot NBS for SCD have been conducted (McGann et al. 2013; Rahimy et al. 2009; Tshilolo et al. 2009). A special tribute should be paid to the late K. Ohene-Frempong for his pilot implementation of NBS for SCA in Ghana, which demonstrated the feasibility of a NBS program in a SSA country. He led the pilot implementation of SCA NBS in Ghana in the early 1995 in the framework of a collaboration with several institutions (Ohene-Frempong et al. 2008). But to date, efforts are still needed for the sustainability of this program.

Insufficient financial and infrastructure resources, as well as the lack of equipment and trained personnel, are among the main limitations to a sustainable national-level implementation of the NBS program (Tshilolo et al. 2009).

1.1.7 Therapeutics

Prophylactic therapy

Preventive therapy includes prophylactic administration of penicillin and pneumococcal vaccination, as SCA children are susceptible to developing severe infections caused by encapsulated germs (Battersby et al. 2010). Lifestyle advice is also given to SCA patients in order to prevent heavy sports or any situation that could cause dehydration, and to prevent buccal infections and carries by having good buccal hygiene. In some settings, folic acid is systematically given to patients, especially adult female patients.

Blood transfusion

RBC transfusion is indicated to increase blood oxygen transport capacity in case of acute anemia and to reduce the proportion of HbS in case of a vaso-occlusion event (Davis et al.

2017). The RBC can be administered by simple transfusion or by exchange transfusion. Simple transfusion consists in the administration of RBC from a healthy donor (HbAA) without concurrently removing the patient's blood. Conversely, the exchange transfusion, which can be performed manually or using a machine (apheresis), combines the removal of an amount of blood from a patient and a concurrent injection of blood from a healthy unaffected donor (Fort 2019; Rees et al. 2018).

Transfusion therapy often exposes recipient patients to alloimmunization, hemolytic transfusion reactions, hyperviscosity, and iron overload. However, even if exchange transfusion prevents hyperviscosity and iron overload, it remains costly and requires a trained staff (Fasano et al. 2016; Zheng & Chou 2021).

Hydroxycarbamide therapy

Hydroxycarbamide or hydroxyurea (HU) is the first drug approved by the US Food and Drug Administration (FDA) for SCA in 1998 due to its clinical and hematological benefits observed in both adults and young SCA patients (Charache et al. 1995; Wang et al. 2011). HU is an antineoplastic S-phase-specific cytotoxic drug that inhibits DNA synthesis through its iron-chelator effect on ribonucleotide diphosphate reductase enzyme. This enzyme catalyzes the biosynthesis of deoxyribonucleotides from corresponding ribonucleotides (Timson 1975). For this purpose, HU was originally used as a myelosuppressive drug in the treatment of hematologic cancers.

The HU treatment increases the percentage of HbF (%HbF) in conjunction with significant increases in the hemoglobin concentration, mean corpuscular volume (MCV), and mean corpuscular hemoglobin as well as simultaneous significant decreases in white blood cell (WBC) count, absolute neutrophil count (ANC), absolute reticulocyte count (ARC), platelets and total bilirubin levels. Clinical benefits of HU include a decrease in vaso-occlusion complications as well as increased life expectancy, preventing end-organ damage (Rees 2011; Ware et al. 2017) .

HU increases HbF levels through intermittent inhibition of erythroid progenitors and cell stress pathway, resulting in the release of erythroid progenitors with high content of HbF (Agrawal et al. 2014). HU induces γ-globin expression through different mechanisms as reported by Shah & Dwivedi.(Shah & Dwivedi 2020). . The epigenetic mechanism consists in reduction of CpG islands methylation.

Through signal transduction pathway, HU interacts with the NO-cGMP pathway , the MAPK pathway and the SAR1 pathway. At the post-transcriptional level, HU indirectly acts on MYB, an important regulator of HbF expression.

Although the main therapeutic mechanisms of HU include enhancing HbF production, several other beneficial effects are reported. HU prevents vaso-occlusion by inhibiting adhesive interactions between sickled RBC, reticulocytes, and endothelial cells through the decrease of adhesive markers expression (Odievre et al. 2008); prevents endothelial dysfunction by either enhancing NO availability (Halsey & Roberts 2003); boosts the antioxidant enzymes to scavenge the ROS produced (Santana et al. 2020).

Other drugs

More recently, three other compounds have been approved for SCA treatment by the FDA: L-glutamine, crizanlizumab, and voxelotor. These drugs are indicated as alternatives for HU in non-responder or those who developed side effects for HU. They may also be used in combination with HU (Minniti 2018).

L-glutamine, approved by the FDA in 2017, is an amino acid that reduces the oxidative stress within sickled RBC by raising the concentration of reduced nicotinamide adenine dinucleotides (NADP) and nitric oxide (NO) (Quinn 2018). The clinical benefit outcomes include a lower frequency of acute chest syndrome and vaso- occlusive pain events ((Niihara et al. 2018).

The clinical benefit observed with the P-selectin monoclonal antibody Crizanlizumab consists essentially in reducing the frequency of vaso-occlusive pain crises (Ataga et al. 2017).

Voxelotor (previously called GBT 440), approved by the FDA in 2019, is an allosteric activator of HbS affinity for oxygen. Thus, by promoting oxy-HbS to the detriment of deoxy-Hbs, this drug prevents HbS polymerization and RBC sickling (AlDallal 2020; Howard et al. 2019).

Gene therapy

Gene therapy involves replacing the defective gene with a functional one. Viral particles or plasmids are used as vectors to release the functional gene and facilitate its integration in the host genome (Gonçalves & Paiva 2017). Depending on the technique used, gene therapy is available for SCA in the form of gene correction (correction of the A>T substitution), addition (γ- globin gene addition or functional β- globin gene addition), silencing (silencing repressor of γ- globin gene or targeted γ- globin induction through genome editing tool as the Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR)-Cas9) (Kanter & Falcon 2021).

Despite successful results being reported in SCA patients overseas, gene therapy is limited by the cost and the related chemotherapy toxicity (Kanter & Falcon 2021; Ribeil et al. 2017). To the best of our knowledge, gene therapy for SCA patients is not available yet in Africa.

Stem cells transplantation

The hematopoietic stem cell transplantation (HSCT), the only curative treatment for sickle cell anemia, consists in removing the patient's bone marrow followed by its replacement by normal globin stem cells from an unaffected donor (Shenoy 2013). Due to the associated high risk of complication, it is indicated in patients with severe complications such as acute chest syndrome, stroke, and multiple organ damage (Krishnamurti et al. 2019). Moreover, barriers to the large use of HSCT include the cost, the need for a Human leukocyte antigen (HLA)-matched donor, the side effects of immunosuppressive drugs, and the high risk of graft-versus-host disease (Robinson & Fuchs 2016). Preimplantation genetic testing (PGT) of in vitro fertilized embryos may appear as an approach to overcome the difficulty of finding a compatible donor. This method allows the selection of unaffected embryos to be implanted. Consequently, the "PGT baby born" could be a potential donor of HLA-matched hematopoietic stem cells for SCA-affected siblings. However, this practice raises important ethical questions that are not fully elucidated yet (de Rycke & Berckmoes 2020).

Natural treatment for SCA

Local research on some plants used as nutraceuticals are reporting positive results. In addition to their anti-sickling property, most of these plants are food plants. Therefore, they can easily be incorporated into the daily diet of SCA patients.

To date, researchers from the University of Kinshasa have completed *in vitro* evaluation of 15 nutraceuticals for their anti-sickling, anti-hemolytic and antioxidant activities, with promising results (Mpiana et al. 2016). Some of those explorations will soon be continued as clinical trials.

1.2 The Democratic Republic of Congo and SCA

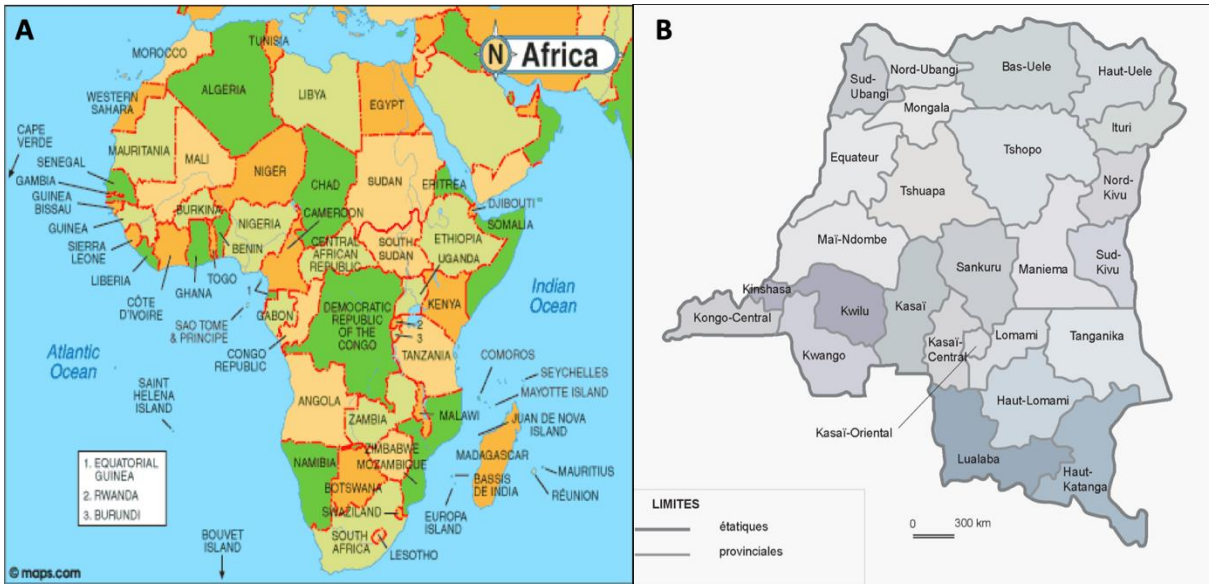

Figure 5. Maps of Africa and DR Congo
Adapted from Wikipedia. Panel A: RD Congo is at the heart of sub-Saharan Africa. Panel B: location of Kinshasa and Kisantu in Kongo Central provinces.

The Central-African Democratic Republic of the Congo (DRC) is the second largest African country, with a surface area of 2 345 000 095 km ^{2} and more than 90 million inhabitants (figure 5).

1.2.1 History of SCA in DR Congo

SCA research studies conducted in the DRC have contributed to the initial understanding of the disease. Those researches may be divided into 3 different phases along the time line. The early phase consists of researches conducted from the 1920s to 1960s. Those studies were conducted mainly by researchers from the KU Leuven (Belgium) (Ager et al. 1958; Lambotte-Legrand & Lambotte-Legrand 1958; Vandepite & Delaisse 1920; Vandepite et al. 1955; Vandepitte & Pieters 1952) . The oldest publication on SCA in DRC retrieved during our literature review dates from 1920, which reported on sickle cell disease and malaria. The main focus of the early research was the clinical description of the rather new disease at the time and the implementation of early electrophoresis testing techniques such as paper electrophoresis.

The second phase started in the early 70s with Congolese researchers Kabakele, Lurhuma, and Ngandu-Kabeya who led the implementation of the National Reference Center for Sickle Cell Anemia in 1990, known today as the Centre de Médecine Mixte et d'Anémie SS (CMMASS), The end of this era may be dated to 2010 with the creation of the Centre Médical Monkole by

Léon Tshilolo. This phase marked the constitution of research and care institutions, and the start of cohort studies.

While Kabakele and his team focused more on therapeutic approaches to treat ischemic necrosis of femur head and malleolar ulcers in SCA patients (Feola et al. 1992), Tshilolo et al. introduced the IEF method in the country and offered the first glimpse into the efficacy of hydroxyurea and the first pilot NBS in 5 provinces in the DRC. This second also provided evidence of clinical and therapeutic variability of SCA in Congolese patients (Aloni et al. 2013; Luboya et al. 2014; Mikobi et al. 2018a; Musumari et al. 2021; Tshilolo et al. 1996, 2019).

The third phase is exemplified by the work of Tshilolo et al. and Lukusa et al. The former group works on the effectiveness of HU treatment (Tshilolo et al. 2019). The latter team is leveraging genetics in the improvement of diagnostics and in the better understanding of clinical and therapeutic variability in SCA (Mikobi et al. 2018a,b). Our work is performed in this framework.

1.2.2 Epidemiology of SCA in the DR Congo

DR Congo is among the most affected countries worldwide. Nigeria, India, and DR Congo share half of the world's SCA patients (Odame 2014; Tshilolo et al. 2008). About 20 to 30% of the population is heterozygous AS whereas 0.97 to 1.4% is homozygous SS. The estimated incidence is 30 000 to 40 000 SCA newborns per year (Agasa et al. 2010; Tshilolo et al. 2009).

In 2001 the DRC Government created a program named "Programme National de Lutte Contre la Drépanocytose" (PNLCD) as the national response to SCA. This program aims to raise awareness as well as to supervise research and prevent SCD in DR Congo. Despite the existence of the PNLCD, the SCA-related morbi-mortality remains important, almost at the same rate as previously reported in other countries in the sub-Saharan region. Factors that contribute to high mortality include delayed diagnosis, scarcity of adequate infrastructures such as appropriate laboratories, poor access to adequate medical care due to limited financial resources of patients, and inadequate supply system.

1.2.3 Local socio-cultural environment

The general population is well aware of the existence of SCA. However, the basic scientific knowledge regarding this disease is limited in the population, even among students, as recently reported (Kambale-Kombi et al. 2020). Such limited awareness was also reported among

patients and their families. A study of 50 Congolese families affected with SCA showed that 70% of parents did not know about the hereditary nature of the disease. The disease itself is referred to with names that often contain a secondary meaning such as « *maladi ya kibeka*» or splenomegaly in Lingala, or « *maladi ya mikuwa* » meaning bone disease in reference to the disease-associated painful crises.

None of the local names indicates that the primary problem is hematological. The origin of the disease is rather attributed either to bad fate, to curse, or to divine punishment (Mukinayi et al. 2018).

This explains the stigmatizations suffered by SCA patients, often resulting in the rejection of affected children. The situation is exacerbated by the high recurrence risk (25%). The recurrence or the death of affected offspring often leads to the separation of families and divorces (Luboya et al. 2014).

Conversely to the negative pressure exerted by society, there is a brighter side that also deserves to be highlighted. The religious communities are currently playing a major role in disease prevention. Nowadays the major religious communities request premarital testing for SCA before wedding celebrations. Since the population values the wedding celebration by churches, such preventive measures have gained momentum and are expected to influence the incidence of the disease in the coming decades. However, because premarital testing for SCA is not a legal mandate, registration of marriage in court is not subjected to testing. SCT individuals or individuals who are not aware of their carrier status can still marry in the country, maintaining the likelihood of perpetuating the high incidence of the disease in the country.

The Congolese health system is organized as a pyramid, with health centers at the basic level and tertiary hospitals at the high level. At the level of health centers, operational, clinical expertise and reliable SCD tests are lacking (Mukinayi et al. 2018). The public health system does not provide subsidies for public hospitals. Without such support, hospitals are unable to purchase new or modern equipment or maintain existing ones. Hospitals and physicians are then often unable to respond appropriately to the demands of patients with SCA.

There is no public health insurance in DRC. The country practices the '*out-of-pocket*' health care system, meaning that the burden of covering medical expenses is put on patients and their parents. The only assistance comes from either family members or some employers

(Wembonyama et al. 2007). Limited accessibility to quality health care encourages patients and their families to use natural herbal medicine. Such an approach could be cost-effective, provided evidence of the efficacy of such medicines is available. To date, most of these medicines have not been tested in randomized clinical trials. Researchers from the University of Kinshasa are testing some of the local natural plants seeking to identify the most efficient ones (Mpiana et al. 2007, 2008, 2010, 2016).

Although WHO has recognized Hydroxyurea (HU) as an essential drug, HU is not routinely prescribed in DRC. The main reason for the limited access to HU treatment would include the cost of HU treatment.

It was reported that HU treatment for a SCA patient weighing 25 kg, would cost around 215 US dollars annually (Mbiya et al. 2021), which is above the budgetary capacity of Congolese families who live with less than 1$ per day (Kabinda et al. 2019). A second reason would be the limited drug availability of HU in DR Congo. It was demonstrated that HU was available in only 22% of pharmacies in the two main cities of DRC (Mukinayi et al. 2021). A third reason would be the insufficiency of knowledge regarding HU management among physicians. A survey of physicians reported that only 9% of physicians prescribe HU to their patients (Mukinayi Mbiya et al. 2020). The WHO is calling for a full supply of HU to the populations in need, such as Congolese SCA patients.

Malaria, malnutrition, intestinal parasites, and other infections are frequent in DR Congo. These diseases are known to be major causes of severe anemia (Aloni et al. 2013; Tshilolo et al. 2007), leading to a high rate of transfusion often early, at the mean age of 22.4 months (Tshilolo et al. 2016). These diseases often overshadow the SCA making diagnosis and care complicated.

1.2.4 Diagnostic challenge

A timely diagnosis is challenging and there is no newborn screening program yet in the country. It is worth noting that the current standards in newborn screening rely on hemoglobin-based techniques. Because of the limitations of usual methods, positive or inconclusive cases need to be recalled for a second test after the age of 3 months to confirm the diagnosis (Tshilolo et al. 2009).

The lack of accuracy of hemoglobin-based tests after the transfusion is ascribed to the presence of the donor's hemoglobin in the receiver's blood. It is worth noting that most of the patients arriving at healthcare facilities with severe anemia are transfused prior to the hemoglobin profile being determined for SCA diagnosis. After blood transfusion, screening of homozygous patients can result in a false heterozygous condition. In a recent report on the management of sickle cell anemia in DRC, the Emmel test was the main method used in 40% of cases; the other method available is Hb electrophoresis on cellulose acetate in 12% of cases, and for the remaining cases, diagnosis is based on clinical patterns (Kambale-Kombi et al. 2021). In a few specific laboratories, mainly in Kinshasa, other techniques for HbS detection are available including Isoelectrofocusing and capillary electrophoresis (Kasai et al. 2020). And recently, some Rapid Test Diagnostic (RTD) for SCA have been validated (Kasai et al. 2020; Katawandja et al. 2021). None of these technics is able to offer reliable results in patients who underwent a recent transfusion. The time until the re-conversion back to the patient's profile and for the hemoglobin-based tests to become reliable again is dependent on the clearance of donor red blood cells (RBC) from the receiver's bloodstream.

Currently, little is known about this timing, especially in African SCA patients. It is not clear how this may vary between patients and what other variables are involved. In common medical practice, clinicians in DR Congo recommend waiting 3 months after a transfusion before performing a hemoglobin-based test. This window period corresponds to the RBC half-life time.

1.2.5 The added value of genetic studies of SCA for DRC

The post-transfusion period as well as the first 3 months of life represent challenges for hemoglobin-based tests. A DNA-based diagnostic test is therefore an attractive alternative in those instances. Because of limited resources, it is essential to use cheap and easy collection, storage, and analysis solutions. Dried Blood Samples (DBS) and buccal swabs might offer that solution.

SCA patients present important phenotypic heterogeneity outcomes both in term of of types and severity of manifestations (Driss et al.,2009; Aghajani et al.,2017). This phenotypical diversity was also previously observed in Congolese SCA patients (Mikobi et al., 2017). Even if the complete mechanism of this diversity is not yet fully elucidated, beta globin haplotypes as well as some polymorphisms associated with the baseline HbF levels are involved (Steinberg

et al.,2009; Ndugwa et al., 2012) . The HbF level, as the most powerful modulator of the clinical expression of SCA (Akinsheye et al. 2011), is an heritable trait whose expression is mainly regulated by 3 quantitative trait loci (QTLs) (Garner et al. 2000; Gibney et al. 2008). So far, only the study by Mikobi et al evaluated association of a set of 8 candidate SNPs with HbF levels in a cohort of Congolese patients (Mikobi et al.2018). Consequently, It appear interesting to enlarge the number of SNPs and to fill the gap of β globin haplotypes distribution, for a better characterization of the phenotypic heterogeneity of Congolese SCA patients. Genetic factors contribute strongly to the heterogeneity in HU response (Santana et al. 2020). The most frequent reported are SNPs, alpha-thalassemia co-inherited and beta S-globin haplotypes (Ma et al. 2007; Okumura et al. 2016 ; Sales et al.2022). There is a correlation between baseline and HU-induced HbF levels. Some genetic polymorphisms are associated with HU-induced HbF levels. Interestingly, some of those are also associated with basal HbF levels (Friedrisch et al. 2016). Unfortunately, these studies originate from outside the DRCongo. To date, only the study by Tshilolo et al has evaluated clinical and biological response to HU treatment in a cohort of Congolese SCA patients (Tshilolo et al.2019). But the study did not evaluate genetic determinants of HU response. The Congolese SCA population is suited for these studies because of the high prevalence of the disease in the country, the genetic diversity of the local population, and the absence so far of systematic HU treatment. Consequently, before the national-wide implementation of HU to become reality, pharmacogenomics studies might be useful to predict HU response in terms of tolerance and HU-induced HbF levels and implement suitable guidelines for DR Congolese SCA patients.

2. AIMS AND OBJECTIVES

2.1 General aim

This research aimed to contribute to the improvement of the care for SCA patients in DRC through robust molecular testing and an improved understanding of the influence of genomic polymorphisms on clinical and biological profiles, as well as on the response to HU treatment in SCA patients.

The first challenge is a timely diagnosis. A nationwide newborn screening program would address this issue, but unfortunately, this might not be a priority for a developing country. Hemoglobin-based techniques, when available, are used to confirm the clinical diagnosis. Early testing would allow for better prophylactic and therapeutic management. The usefulness and applicability of the different genetic assays in the context of the societal, economic, and healthcare setting in DRC will be considered.

Despite being a monogenic disease and caused by a single point mutation, SCA is characterized by a great inter-individual clinical variability in terms of range and severity of clinical manifestations. The clinical variability has been correlated with both coinheritance with alpha thalassemia and variations in the levels of fetal hemoglobin (HbF). Knowledge on the contribution of genetic variants in the DRC population could better determine the prognosis of disease in Congolese patients.

At present, hydroxyurea (HU) is an essential drug used to alleviate the manifestations of SCA. Genetic factors contribute strongly to the heterogeneity in HU response. Eventually, it would be worthwhile to know whether the use of polygenic risk scores (PRS) could also improve clinical management.

The study of the genetic determinants underlying the clinical variability of SCA and insight into the pharmacogenetics of HU treatment in Central Africa are becoming topics of great interest to the scientific community.

We have assembled a large cohort of Congolese SCA patients in an attempt to generate genetic data, and fill the gaps of haplotypes, polymorphisms and their influence on reported variability in severity and HU response. This research is carried out jointly with two other Congolese

Ph.D. students, my colleagues Dr. Paul Lumbala and Dr. Gloire Mbayabo who conducted patients' recruitment and the clinical trial for HU.

2.2 Specific aims

We plan to:

1. Implement a DNA-based SCA test that will be suited for a limited resource setting and will overcome limitations of hemoglobin-based tests commonly used in the current standard of care. Because of limited resources, it is essential to use cheap and easy collection, storage and analysis solutions. Dried blood samples (DBS) and buccal swabs might offer that solution.

2. Evaluate the influence of SCA haplotypes and selected polymorphisms on basal levels of hematological parameters and clinical presentation and progression in SCA patients. Variations in the baseline HbF level are associated with genetic polymorphisms within the HBB cluster as well as polymorphisms at other quantitative trait loci like e.g. BCL11A, HBS1L-MYB and HBB. Data on the frequency and effect of SCA haplotypes on disease severity are not available for the DRC.

3. Evaluate the influence of SCA haplotypes , alpha thalassemia and selected polymorphisms on the response to HU therapy. Genetic factors contribute strongly to the heterogeneity in HU response. There is a correlation between baseline and HU-induced HbF levels. These studies originate from outside Central Africa.

3. RESULTS

3.1 PART 1. A DNA-based SCA test for a limited-resource setting

Sickle Cell Anemia is a genetic disease, meaning that the causal defect is a change in the sequence of the DNA. This implies that the identification of the mutation provides irrefutable proof of the disease. Thus far, the diagnosis is based on the detection of the phenotypic change, which is the presence of the mutant hemoglobin protein known as S hemoglobin (HbS). Although very popular, these phenotypic approaches have some limitations that are exacerbated in tropical regions such as in DR Congo. Moreover, there is a need for a test that can be financially accessible to all and feasible on all types of samples. We evaluated the feasibility and the benefit of deploying the SCA test using DNA. The rationale, the methods, and the results and presented below and have been published in the Journal of Clinical Laboratory Analysis 2022; 36(5): e24398. (doi: 10.1002/jcla.24398. PMID: 35405024).

Besides, we are currently exploring the feasibility of the molecular test in the hours following a blood transfusion since in most pediatric patients SCA is suspected after one or more transfusions. The study will continue beyond this thesis. However, the preliminary results are presented in appendix to the research article 1.

RESEARCH ARTICLE 1

3.1.1 DNA testing for sickle cell anemia in Africa: Implementation choices for the Democratic Republic of Congo

Mamy Ngole [1,2,3] | Valerie Race[1] | Gloire Mbayabo [1,2,4] | Paul Lumbala [1,2,4] | Cathy Songo [2] | Prosper Tshilobo Lukusa [1,2,4] | Koenraad Devriendt [1] | Gert Matthijs [1] | Aimé Lumaka [2,4,5]

[1] Center for Human Genetics, Faculty of Medicine, KU Leuven, Leuven, Belgium

[2] Center for Human Genetics, Faculty of Medicine, University of Kinshasa, Kinshasa, Democratic Republic of Congo

[3] Department of Medical Biology, Faculty of Medicine, University of Kinshasa, Kinshasa, Democratic Republic of Congo

[4] Department of Pediatrics, Faculty of Medicine, University of Kinshasa, Kinshasa, Democratic Republic of Congo

[5] GIGA-R, Laboratoire de Génétique Humaine, University of Liège, Liège, Belgium

Abstract

Background: Hemoglobin-based tests form the reference diagnostic test for SCA. In limited resource countries, these tests face limitations including cost, low sensitivity due to recurrent transfusions in endemic malaria region, and interference from fetal hemoglobin in neonatal diagnostic. This study aimed at adapting DNA-based SCA tests to limited resource countries and evaluating the economic benefit.

Methods: 338 participants were recruited in the Democratic Republic of Congo, sorted in 3 cohorts based on venous blood, umbilical cord blood (UCB) and buccal swab sampling. RFLP was performed to identify mutated allele. The feasibility and technical validity of this RFLP was evaluated for specimens collected on DBS cards and on EDTA tubes. RFLP on DBS stored at room temperature was regularly repeated to assess sample conservation. Finally, the cost analysis was performed.

Results: DBS cards yielded identical results to extracted DNA. Repeated testing returned the same result after four years. The DBS-based test performed on UCB or on buccal swab had a sensitivity and a precision of 100%. Cost comparison indicated that our approach costs half price of the widely used isoelectrofocussing of hemoglobin.

Conclusion: The implemented DNA-based test approach overcomes the limitations faced by hemoglobin-based tests while being more affordable. We propose to implement the RFLP test as a first-line diagnostic test after transfusion and as second tiers for newborn screening. However, users should be aware that this test is unable to differentiate HbC from HbS or identify other point mutation of gene deletion of HBB gene.

Keywords: buccal swab, DNA based-tests, sickle cell anemia, umbilical cord blood, venous blood

Introduction

Low-income countries such as the Democratic Republic of Congo (DRC) suffer a heavy burden of sickle cell anemia (SCA) with about 2% of newborns being homozygous for the classical SCA mutation, and 17%–24% being heterozygous carriers (Tshilolo et al. 2009).

The disease-related mortality remains high in many sub-Saharan African countries as 50% to 90% of affected children die before 5 years of age (Grosse et al. 2011), often before being formally diagnosed. Thus far, nationwide newborn screening is not available in DRC. The most commonly available SCA tests in many localities in DRC include the Itano solubility test and the Emmel sickling test. These tests detect only the presence of hemoglobin from the abnormal allele (HbS) and are unable to detect the hemoglobin from the reference allele (HbA). Therefore, these tests are unable to distinguish sickle cell trait (AS profile from heterozygotes individuals) from SCA (SS profile from homozygotes patients) (Piccin et al. 2008 ; Kitenge et al. 2018). Thus, they are of limited diagnostic and clinical value. Electrophoresis on acetate cellulose and isoelectrofocussing (IEF) has been introduced in DRC since 1954 and 2009, respectively (Vandepitte et al. 1963 ; Tshilolo et al. 2009).

However, these technologies are not widely available across the country. Samples need to be shipped from rural areas to the laboratories located in urban areas. Such sample transfer and storage represent big challenges and may impact on the cost and quality of the test. These technical and logistic challenges as well as the cost of the test should be considered when implementing a diagnostic test for SCA in a resource-limited setting. In addition, the hemoglobin electrophoresis-based techniques have low accuracy during the neonatal period because of the presence of fetal hemoglobin, or after a transfusion because of the presence of donor's hemoglobin (Bardakdjian-Michau et al. 2003).

Studies on Congolese SCA patients have reported that anemia is the most common sign and often the first manifestation (Tshilolo et al. 1996 ; Shongo et al. 2014). Thus, when SCA is suspected, patients have usually a recent history of transfusion, forcing practitioners to delay the diagnostic test for 3 months. Unfortunately, because of other intercurrent factors, such as malaria and other infections, that are very frequent in this setting, it is common for these children to return to the hospital and receive another transfusion within these 3 months' window, further delaying SCA confirmation by hemoglobin electrophoresis-based methods.

Therefore, the optimal diagnostic test for SCA should retain high accuracy regardless of recent transfusion or in neonatal period.

Previous studies have shown that DNA-based tests are not affected by transfusion or fetal hemoglobin (El-Haj et al. 2018). Thus, DNA-based tests would be preferable for a resource-limited setting. However, the implementation of such a technique has to address important challenges for sample collection, transfer, and cold chain requirements for storage, as well as the test accuracy in the neonatal period and following a recent transfusion.

This study aimed to evaluate the technical validity of a DNA-based SCA diagnostic test suited for DRC in particular and for low-income settings in general.

Methods

Study design

This study was conducted on three cohorts and different biological specimen types (venous blood, umbilical cord blood, and saliva) in each phase.

Cohort 1. Venous blood

Study population

A cohort of 166 participants was recruited, both at the Saint Luc Hospital and the Centre de Médecine Mixte et d'Anémie SS (CMMASS), respectively, in Kisantu (Kongo Central Province) and Kinshasa (the capital city) in the DRC. These patients were previously diagnosed as homozygous (SS; n = 145), heterozygous (AS; n = 12), and wild-type (AA; n = 9) using a hemoglobin electrophoresis-based technique. The three hemoglobin types represent the three genotypes for the SCA mutation. The SS individuals are SCA patients, AS are heterozygous carriers while the AA are homozygous wild-type individuals.

Samples collection and transportation

For each participant, 4 ml of peripheral venous blood was collected in a EDTA-coated tube (BD Diagnostics). Before removing the sampling needle, a drop of blood (approximately 50 µl of blood) was deposited on a circle of a FTA (Flinders Technology Associates) Elute card (Whatman WB120206). This type of FTA card is chemically treated filter papers that lyse cells, denature proteins and stabilize nucleic acids on the cellulose fibers (Figure 1A). The FTA cards have four circular areas where the sample should be deposited. More details about the FTA

Elute card are provided by the manufacturer: at: https://webshop.fishersci.com/webfiles/fr/web/FS2013/EU_FR_13LAB_P0167.pdf?_ga=2.10467506.685819103.1647441468-337705928.1647441468

The spotted FTA cards were kept at room temperature for 90 minutes allowing the sample to dry prior to transportation in an individual plastic bag, adequately labeled. In order to reduce the cost of the experiment, each individual FTA card was divided into two pieces, allowing one card to be used for two patients (Figure 1A).

The spotted FTA cards (dried blood spot, DBS) were stored in zip bags at room temperature on a clean shelve until being used for applications as further described below. Blood samples in EDTA tubes were transported in isotherm boxes from the sampling site to the laboratory of the Center for Human Genetics of the University of Kinshasa (UNIKIN, www.coshg.org).

DNA extraction and molecular test

The venous blood collected in EDTA tubes was used for DNA extraction by salt-induced precipitation of cellular protein method also known as the standard salting out procedure (Miller et al., 1988). The extracted DNA was then re-suspended in 300 µl of TE buffer and kept at −20°C until the analysis. The standard manual salting-out extraction procedure lasts around 24 hours. In our study, DNA extraction was performed within two days following the sample collection. A restriction fragment length polymorphism (RFLP) test for SCA was implemented in our laboratory as previously described (Mikobi et al. 2017). This test uses the DdeI restriction enzyme to differentiate between the mutant allele (S) and the wild-type allele (A) on a PCR product spanning the SCA mutation site (Figure 1B).

Cohort 2. Umbilical cord blood

Study population

Umbilical blood cord was collected from 102 newborns in two maternities in Kinshasa, the Maternity Hospital of Binza (MHB) in the north of Kinshasa, and the Maternity Hospital of Kingasani (MHK) in the east of Kinshasa.

Samples collection and transportation

Umbilical cord blood (UCB) was obtained “in utero,” meaning before the placenta was expelled, the fetal extremity of the umbilical cord was disinfected, clamped, and sectioned.

Blood sample was collected by gravity into a 4 ml EDTA tube. A drop of UCB was directly spotted onto two circles of a FTA card. Samples were transported to the laboratory of the Center

for Human Genetics of the Kinshasa University at room temperature packed in plastic zip bags for spotted FTA cards and in isotherm boxes for UCB in EDTA tubes. The UCB sample collected in EDTA tubes was used for DNA extraction following the standard salting out procedure, whereas spotted FTA cards were used as described below.

Cohort 3. Saliva

Study population

A group of 70 participants was recruited, among which 20 were from the CMMASS and 50 from the University Hospitals of the University of Kinshasa (CUK). Previous molecular diagnosis using DNA extracted from peripheral blood returned homozygous mutation genotype (SS hemoglobin type) in 43, heterozygous (AS hemoglobin type) in 13, and homozygous wild type (AA hemoglobin type) in 14.

Samples collection and transportation

For each participant, buccal cells were collected using sterilized cotton buds made with wooden sticks. Patients were recommended not to eat, drink, smoke, chew gum, or use toothpaste or mouthwash during the 30 minutes preceding sample collection. Samples were collected by scratching while rotating the swab on the mucosa of the cheeks for 30–60 seconds. After sampling, the swab was scratched against a circle of the FTA elute card for 10 seconds in order to transfer the cells from the swab onto the FTA card.

DNA testing

Each sample collected in EDTA tubes underwent DNA extraction by manual salting-out method (SO) (Miller et al. 1988). Extracted DNA was later used for a PCR reaction and the RFLP using an Applied Biosystems 2720 Thermal Cycler. We refer to this protocol as SO-RFLP across the article.

For samples collected on FTA cards, a DNA extraction step was not performed. One 1.2 mm disk was punched from the spotted FTA card, washed with purified water in a 2 ml reaction tube as recommended by the manufacturer, then transferred into the PCR reaction tube and mixed with the PCR reaction mix (Table 1). The PCR product was used for enzymatic restriction as described below. In order to prevent cross-contamination while punching, an unused card was punched three times in between two samples as recommended by the manufacturer. The remainder of the spotted FTA card was stored at room temperature in a zip bag in a dust-free closet. This second protocol is referred to as FTA-RFLP in this article.

The PCR targeted a fragment of 440 bp in the first exon of the hemoglobin beta gene (HBB; GRCh38 chr11:5225464-5229395, Figure 1B) using the following primer pair:
Forward TGTGGAGCCACACCCTAGGGTTG and Reverse CATCAGGAGTGGACAGATCC.

For each run, two controls were used, including a No DNA control (disk punched from blank unspotted paper) and a positive control (disk punched from FTA card spotted with blood sample of a Congolese SCA patient with known genetic profile). The composition of the PCR mix is presented in Table 1.

Either 1 µl of DNA extracted by SO (100–300 ng) or a 1.2 mm washed FTA punch was added into the PCR mix for the SO-RFLP or FTA-RFLP, respectively.

The PCR program consisted of an initial denaturation step at 95°C for 5 min, followed by 32 amplification cycles, each cycle included a quick denaturation at 95°C for 30 s, annealing at 58°C for 30 s, and extension at 72°C for 30 s. The reaction was concluded with a final extension step at 72°C for 5 min.

The PCR product was digested using the DdeI enzyme (Promega Corporation, Belgium, Cat: R6291) according to a standard protocol. The DdeI enzyme has two restriction sites (ctgag/) within the PCR product (Figure 1B). The digestion products were visualized on 2% agarose gel (Figure 1C). A non-mutated or wild-type allele produced three fragments with 201 bp, 167 bp, and 72 bp size, respectively. Conversely, on an allele with the S mutation (mutated allele), the first restriction site is abolished, resulting in only two fragments of 368 bp and 72 bp, respectively. For homozygotes SS individuals, only these two bands appear, three bands are observed for homozygotes AA, whereas four bands are identified in heterozygotes AS, that is, 368 bp, 167 bp, 201 bp, and 72 bp bands.

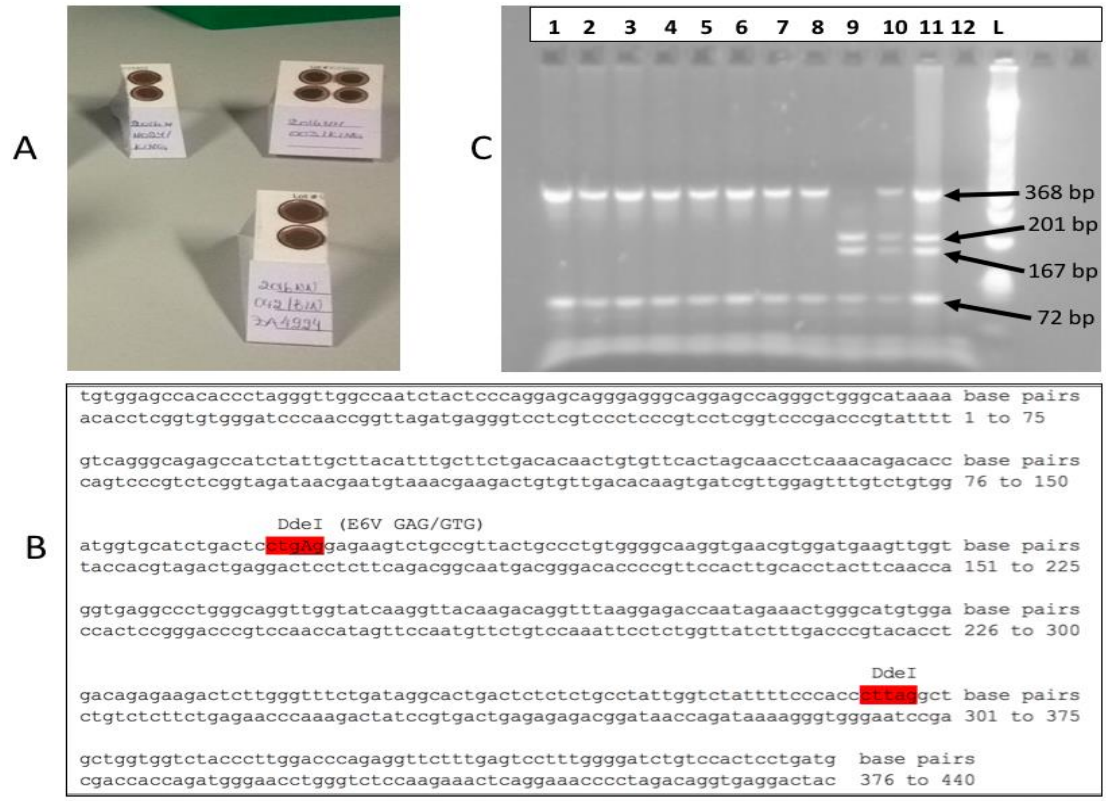

Figure 1. Sample collection, experimental design, and interpretation

A. Blood sample spotted onto FTA elute cards. Here are illustrated the full FTA card and two pieces of a split FTA card, all spotted with different samples.

B. The figure shows the target PCR sequence. The two restriction site sequences are underlined with a dashed line. The S mutation is located on the first restriction site.

C. Pattern of restriction products on 2% agarose gel. Lanes 1 to 8 shows homozygous mutation SS profiles, lane 9 shows homozygous wild type AA, lane 10 illustrates heterozygous AS, lane 11 contains a positive control, whereas lane 12 is no DNA. The size ladder is on the L lane.

Table 1. Composition of the PCR Mix

Solution	Amount per sample
Primer F + R (2.5µm)	2 µl
Amplification buffer (Roche)	2 µl
dNTP's 2mM	2 µl
Taq DNA polymerase (5U/µl) (Roche)	0.1 µl
H_2O	12.9µl
Total volume	19 µl

Sensitivity and precision

Sensitivity, considered as the ability of the test to identify variants that are present in a sample, was calculated using the formula TP/TP + FN. The precision, defined as the fraction of variant calls that match the expected, reflecting the number of FP per test, was determined using the formula TP/TP +FP. (Marshall et al., 2020). As previously demonstrated, the sensitivity threshold is influenced by the prevalence of the disease in the studied population (Leeflang et al. 2013). For the previously reported incidence of 1.4% in DRC (Tshilolo et al. 2009) an error rate of 5% will result in multiple children being denied timely care. Therefore, we set the minimum threshold at 99%.

Evaluation of repeatability and reproducibility

Repeatability, defined as the percent agreement between the results of successive tests carried out under the same conditions of measurement,(Marshall et al.2020) was assessed in two ways: the intra-run and inter-run repeatability. Five FTA cards spotted with venous blood were randomly selected. Three disks from each in five cards were included a run to evaluate intra-run repeatability. This was repeated three times to assess inter-run repeatability. We also assessed the reproducibility, which evaluates whether changes in testing platforms, reagent supplies, and operators cause significant change in test results. A second operator tested one disk from the test set of five FTA cards using a different PCR thermocycler. Reproducibility was calculated as the percent agreement between the results of tests under different conditions (Marshall et al.2020).

Repeatability and reproducibility were assessed for UCB and saliva spotted on FTA cards using six randomly selected cards for each specimen and running the same experiment as described for the venous blood. The considered cut off was 90% (Clinical Laboratory Improvement Amendments CLIA).

Stability of spotted FTA cards in local conditions Each FTA card was stored in an individual plastic zip bag at room temperature, in a clean closet, whereas the extracted DNA was stored at −20°C. We randomly selected 15 FTA cards from study group 1. The selected FTA cards were repeatedly tested every 6 months by FTA-RFLP as described above. Thus far, we have completed testing after 4 years of conservation.

Data analysis

The SO-PCR-RFLP was considered as the gold standard to which results from the FTA cards were compared. The FTA-RFLP method was compared to the gold standard regarding the cost and conditions. Sensitivity and precision were computed. Statistical analysis was performed with SPSS version 21. Kappa statistics was used to evaluate the agreement between two different testing methods. We also assessed the stability of FTA-spotted venous blood samples after 4 years of storage at room temperature.

Ethical compliance

This study was compliant to international ethics laws and regulations. Prior to the inclusion, parents and patients were fully informed and provided signed informed consent/assents. The study protocol was approved by the Ethical Committee of the Public Health School at the University of Kinshasa (ESP/CE/079/2016).

Results

FTA-RFLP on peripheral blood

This test was performed on the 166 participants with known electrophoresis-based hemoglobin profile recruited in Kinshasa and Kisantu. The FTA-RFLP identified 145 homozygotes SS, 12 heterozygotes AS and nine homozygotes wild-type AA. The SO-RFLP, performed on DNA extracted from venous blood collected on EDTA tubes, showed the same results for each participant. Therefore, there was a perfect agreement between FTA-RFLP and SO-RFLP (kappa = 1; Figure 2).

FTA-RFLP on umbilical cord blood

The 102 FTA cards containing UCB samples were tested by FTA-RFLP and returned 1 SS, 14 AS, and 87 AA. Later, the genomic DNA extracted from the 102 UCB samples collected in the EDTA tubes was also tested by SO-RFLP. This returned the exact same result for each participant, corresponding to a perfect agreement between the two techniques (kappa = 1; Figure 2).

FTA-RFLP on buccal swab samples

The FTA-RFLP test performed on swabs rubbed on FTA cards from 70 participants with known SO-RFLP genotypes, identified 43 SS, 13 AS, and 14 AA. These results were in

agreement with the profiles previously obtained by SO-RFLP performed on DNA extracted from blood sampled from the same participants (Figure 2). Thus, there was a perfect agreement between swab-based FTA test and blood-based DNA test (kappa = 1). This validated the sensitivity and precision of the DNA-based Sickle Cell Anemia diagnostic test on buccal swab samples.

Sensitivity and precision

The results from venous blood collected on FTA cards, UCB spotted on FTA cards and swabs rubbed on FTA cards, all returned identical results to the respective outcome obtained with the SO-RFLP. No false positives or false negatives were observed. Therefore, the sensitivity, as well as the precision, was 1 for the DNA-based SCA test using either of those three specimens.

Repeatability and reproducibility

Intra- and inter-run replicates returned the exact same profiles for each of the randomly selected individuals. Moreover, the change of the operator and the PCR thermocycler did not alter the outcome of the test.

Stability of spotted FTA cards in local conditions

The results of 4 years of storage indicated that the profile remained clearly recognizable, suggesting the DNA quality was not altered by the local storage conditions: shelve storage in a dust-free closet at room temperature (between 25 and 35 Celsius degrees). The results remained identical over the years up to 4 years when the last test was performed (Figure 3).

Conditions and cost comparison

The conditions and the cost of both tests are presented on Table 2. The cost of the SO method was based on estimation by Chacon-Cortes et al [14]. While the personnel cost for both SO-RFLP and FTA-RFLP was estimated according to the average monthly salary of a laboratory technician in DRC.

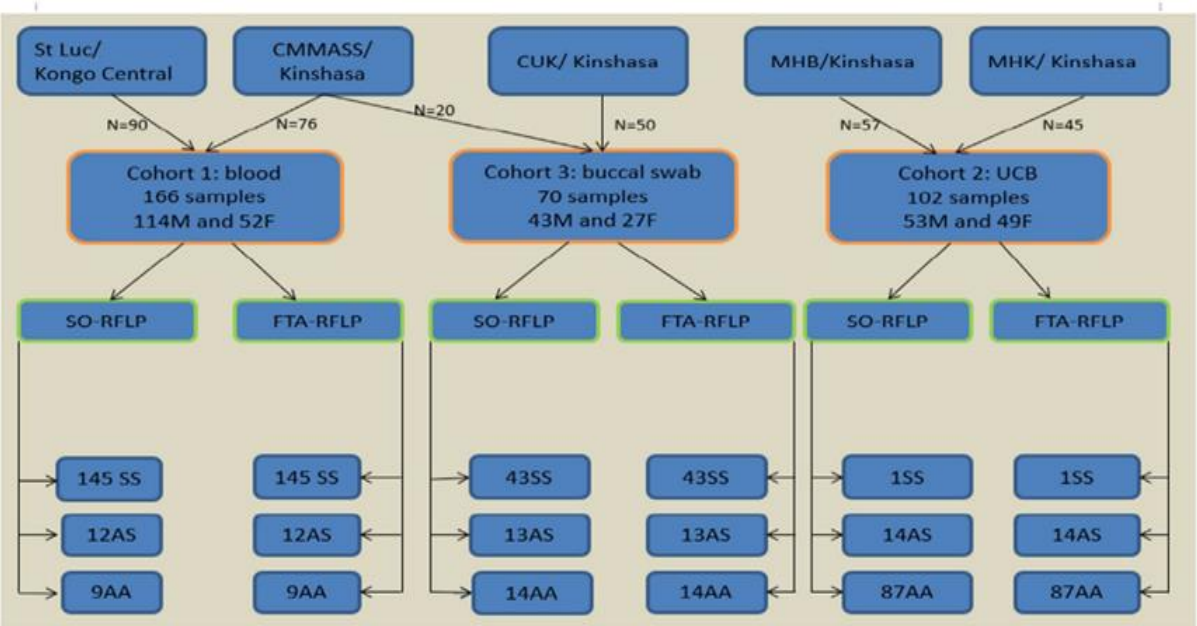

Figure 2. Study algorithm and results

The top row presents the five institutions where participants were recruited; the second row shows the composition and specimens for the three cohorts; the third row indicates the two testing approaches applied to each of the cohort. The distribution of the genotypes obtained by either of the testing approaches for each cohort is presented at the bottom of the figure.

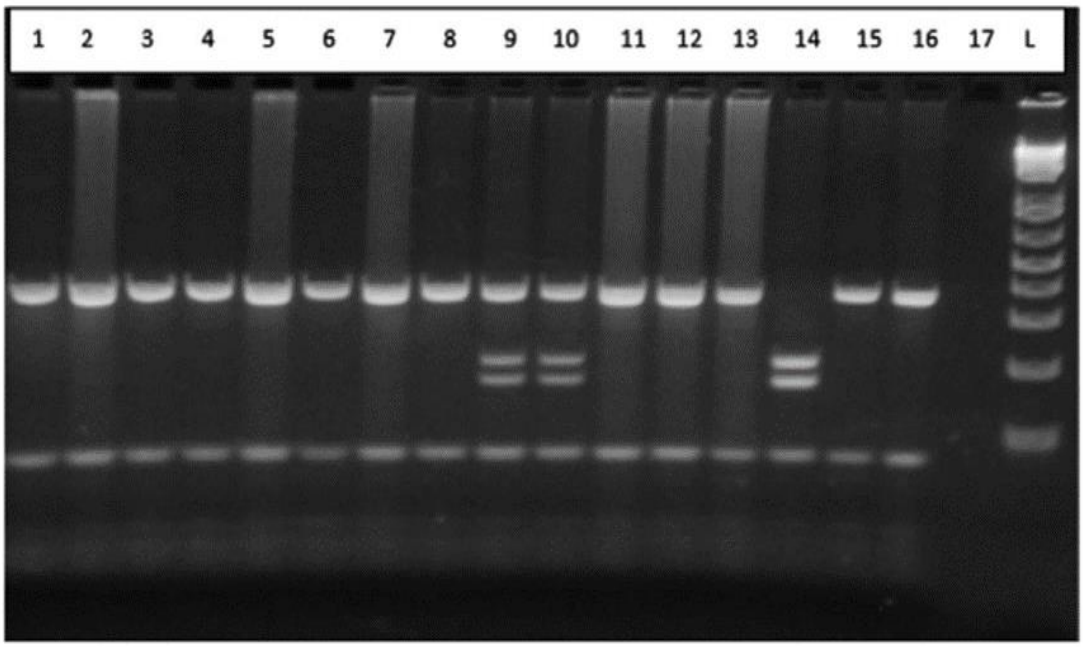

Figure 3. Restriction profile after 4 years of storage of some FTA Elute blood samples

Lane 1 to 8 shows homozygous mutation SS profiles, lanes 9 and 10 show heterozygous AS, lane 11 to 13 show homozygous SS, lane 14 represents wild-type AA, lane 15 shows homozygous SS, lane 16 contains a positive control whereas lane 17 is the no DNA. The size ladder is on the L lane.

Discussion

We conducted a prospective technical validation study aiming at implementing a DNA-based test for SCA that is feasible in a limited resource setting such as the DRC, despite infrastructure, environmental and economic challenges. Among the challenges we wanted to overcome were (1) sample transfer conditions in a country that is as big as the Western Europe and (2) instability of electricity on which depend biorepositories. In addition, we wanted the test to be efficient on non-invasive specimen such as umbilical cord blood and buccal swabs. Finally, we also wanted this test to remain within the cost range of the most used test, the hemoglobin electrophoresis. To the best of our knowledge, neither the DNA-based SCA test nor the use of non-invasive specimen is currently offered as first line test in routine for SCA diagnostic in Central Africa.

Due to the ease in sample collection, transport and handling, the use of DBS cards such as FTA cards is gaining in acceptance for genetic analysis in multiple fields including infectious diseases, agriculture and pharmacogenetic (Mas et al. 2007 ; Fata et al. 2009 ; Abdelwhab et al. 2011 ; Rabodoarivelo et al. 2015 ; Dong et al. 2017 ; Siegel et al. 2017) .

The RFLP was previously identified as a reliable and cheap approach for DNA-based diagnostic of SCA (Bardakdjian-Michau et al. 2003) . However, users should note that this test cannot differentiate HbS from HbC since both mutations affect the same restriction site and result in the same profile on agarose gel electrophoresis. Also, this test is not designed for the identification of others mutations that may cause SCA phenotype, such as compound heterozygous HbS/βthal or HbS/HbO. Fortunately, such mutations are rare in the DRC and absent from this cohort (based on sequencing data, not included in this manuscript).

The FTA cards were tested as an alternative to the EDTA tube sampling and DNA extraction by salting out. In this study, we obtained the same genomic profile using a 1.2 mm FTA discs or 1µl of DNA extracted by salting out from the same participant. We conclude that using FTA cards is technically as good as the extracted DNA for sickle cell testing by RFLP. The amount of material on a dry blood spot may vary from one individual to another. In this study all participants had clear PCR and digestion bands on agarose gel irrespectively of their disease status. This suggests that a 1.2 mm disc of FTA card retains enough DNA sample to allow the test to be carried out without DNA extraction.

The value of the FTA-based test was verified using two other specimens namely UCB and buccal swabs in comparison to genomic DNA extracted from peripheral blood. The UCB is an easily accessible material. This prevented us from sampling newborns by heel punctures. Therefore, this was the ideal and non-invasive material to test the FTA-based approach in newborns. Results from UCB samples on FTA cards proved to be very reliable. Considering that electrophoresis-based diagnostic test for SCA is unreliable before the age of 3 months , (Burton et al. 2014) the FTA-based test may be offered as a confirmatory test for suspect electrophoresis-based newborn screening test results or as a first line test in high-risk couples. This is of high value when considering both the reported delay in diagnostic and the heavy burden of sickle cell in Sub-Saharan countries. As previously reported (Abraham et al., 2012) , our study has proven that buccal swab is a good alternative of DNA for a genetic test such as SCA diagnostic. Proposing such non-invasive approach may increase acceptability of SCA test and newborn screening.

The sensitivity and precision of FTA-RFLP for either venous blood, UCB or buccal swab were 100%, confirming that irrespectively of the specimen and the amount of DNA, FTA-based test is reliable.

It is common to centralize genetic testing in a resource limited setting in order to ensure efficient resource sharing. The Center for Human Genetics of the University of Kinshasa is positioned to become the national reference center for molecular testing not only for SCA but also for various genetic tests.

Whole blood, the most used material for genetic analyses, is a type of sample that can be stored only in certain conditions with respect to the cold chain and for a relatively short period. But in DRC, the unreliability of the electricity supply and the lack of quick and safe sample transportation would not always guarantee an appropriate transport of whole blood samples to the laboratory. Moreover, transport from the collection site to the testing laboratory is laborious and requires an expensive investment for transport. Hence, it would be appropriate to use a consumable that can provide safe blood sample transportation, such as DBS cards.

The issue of power supply is critical for biorepositories or any laboratory storage in resource-limited countries. Moreover, the storage capacity is often very limited, which might be a problem for a country with about 2 840 000 births per year including an estimate of 40000 SCA babies (1.4%) per year (Tshilolo et al. 2009).

Our results show that DBS cards can facilitate sample collection, ensure safer transportation of samples, shorten the turn-around-time by skipping the DNA extraction step and return reliable results even after several years of storage at room temperature (Rajendram et al.2006 ; Peluso et al. 2015 ; Barth et al. 2016; Corradini et al. 2019) . In addition, based on the possibility of stacking the zip bags, it becomes possible for the biorepository to store enormous amount of samples. This implies that a national screening program for SCA should consider using DBS cards or similar material for testing and for safe and long-term storage of samples at minimal cost.

The FTA-RFLP was estimated to cost almost half of the SO-RFLP (Table 2). This is mainly due to the fact that the SO-RFLP is labor intensive and time consuming. This resulted in a high personnel cost that made significant difference between the two approaches. Interestingly, the estimated cost for SO-RFLP is also the half of the price commonly charged in Kinshasa for the IEF. Hence, our proposed approach of FTA-RFLP is not only more accurate than IEF, it is also more efficient on challenging specimen such as UCB, it is suited for non-invasive testing using buccal swab and, most importantly, four times cheaper than the IEF. Therefore, switching to DBS DNA-based test as confirmatory approach in the newborn screening program, and as first line test for SCA diagnostic in other aged groups will significantly reduce the financial burden while offering the highest accuracy. This approach can also be offered in mass campaigns as a non-invasive method.

Table 2. Comparison of conditions and cost

	Salting out	FTA card
CONDITIONS		
• Sample collection	EDTA tubes	FTA cards
• Sample transport • DNA extraction	RT Yes	RT No
• Storage • Turn-around-time (DNA extraction +PCR-RFLP)	Freezer -20°C 24h + 5h	RT 5h
COST ESTIMATION		
• Consumables for sample collection and DNA purification	2.3 €[a]	1.51€ (half a card/patient)
• PCR reaction per sample	2 €	2 €
• Restriction enzyme digestion	1.2 €	1.2 €
• Gel electrophoresis • Equipment depreciation • Personnel cost[b]	0.8 € 0.40 € 4.85€	0.8 € 0.10 € 0.9 €
TOTAL COST	11.55 €	6.51 €

Abbreviations: NA, not applicable; RT, room temperature.
a Based on cost estimation from Chacon-Cortes et al. (Chacon-Cortes et al. 2012)
b Based on the average monthly salary of 500$ for a laboratory technician in DRC. The daily salary average is 25$ per day. The hourly salary is 25$/8 = 3.152 $. The hourly salary is based on the legal limit of 8 working hours a day. For the SO-RFLP, the full run includes 16 samples. The hourly salary per sample is 3.125$/16 = 0.195$. For the 29 hours of the run, each sample costs 0.20$ × 29 = 5.8$ (4.85€) as personnel costs. For the FTARFLP, the full run includes 14 samples. The hourly salary per sample is 3.1$/14 = 0.2$. For the 5 hours of the run, each sample costs 0.2$ × 5 = 1.0$ (0.9€) as personnel cost.

In addition, this test can be used as a first-line test for screening sickle cell traits in blood donors and blood donations. This will increase blood transfusion safety.

CONCLUSION

The study aimed at implementing a DNA-based SCA test in resource-limited settings and evaluating the economic benefit. We have shown that this procedure is feasible in resource-limited settings, retains high accuracy for various sample types, including venous blood, umbilical cord blood, and buccal swab, while remaining affordable. We recommend this test as a first-line diagnostic test after transfusion as well as in blood donors and blood donations, and as a second line for newborn screening.

AUTHOR CONTRIBUTIONS

Prosper Lukusa, Koenraad Devriendt, Gert Matthijs and Aimé Lumaka designed the project and corrected the manuscript. Mamy Ngole, Paul Lumbala, and Gloire Mbayabo conducted patients' recruitment. Mamy Ngole and Cathy Songo performed laboratory analysis under supervision of Valerie Race. Mamy Ngole prepared the manuscript; Prosper Lukusa, Koenraad Devriendt, Gert Matthijs, Aimé Lumaka, and Valerie Race edited the manuscript. All authors read and approved the final manuscript.

DATA AVAILABILITY

We included all the data related to this research in the article. Moreover, all data are available from the corresponding author.

FINANCIAL SUPPORT

"Initiatives Sud" et projet "TEAM 2015" from VLIR-UOS

CONFLICT OF INTEREST

None

ACKNOWLEDGMENTS

The authors are grateful to patients, parents, and the families for agreeing to participate in this study. We thank Bruno Vankeirsbilck of UZ Leuven Molecular Laboratory and his team for the technical and logistic assistance.

References

Abdelwhab, E.M, Luschow, D., Harder, T.C.& Hafez, H.M. (2011). The use of FTA(R) filter papers for diagnosis of avian influenza virus. *J Virol Methods* , 174(1-2), 120-122.

Abraham, J.E., Maranian, M.J., Spiteri, I., Russell, R., Ingle, S., Luccarini, C., Earl, H.M., Pharoah, P.P., Dunning, A.M., Caldas, C. (2012). Saliva samples are a viable alternative to blood samples as a source of DNA for high throughput genotyping. *BMC Med Genomics*, 5, 19.

Bardakdjian-Michau, J., Dhondt, L., Ducrocq, R., Galacteros, F., Guyard, A., Huchet, F.X., Lahary, A., Lena-Russo, D., Maboudou, P., North, M.L., Prehu, C., Soummer, A.M., Verschelde, M. & Wajcman, H.(2003).Good practices for the study of hemoglobin. *Ann Biol Clin*, 61(4), 401-409.

Barth, H., Morel, A., Mougin, C., Averous, G., Legrain, M., Fender, M., Risch, S., Fafi-Kremer, S., Velten, M., Oudet, P., Baldauf, J.J.& Stoll-Keller, F. (2016). Long-term storage and safe retrieval of human papillomavirus DNA using FTA elute cards. *J Virol Methods*, 229, 60-65.

Burton, G.J., Sebire, N.J., Myatt, L., Tannetta, D., Wang, Y.L. , Sadovsky, Y., Staff, A.C. & Redman, C.W. (2014). Optimising sample collection for placental research. *Placenta*,35(1), 9-22.

Chacon-Cortes, D., Haupt, L.M., Lea, R.A. & Griffiths,L.R.(2012). Comparison of genomic DNA extraction techniques from whole blood samples: a time, cost and quality evaluation study. *Mol Biol Rep*,39 (5), 5961-6.

Corradini, B., Alu, M., Magnanini, E., Galinier, M.E. , Silingardi, E. (2019). The importance of forensic storage support: DNA quality from 11-year-old saliva on FTA cards. *Int J Legal Med* 133(6), 1743-1750.

Dong, L., Lin, C., Li, L., Wang, M., Cui, J., Feng, R., Liu, B., Wu, Z., Lian, J., Liao, G., Chen, W.& Qiao, Y. (2017). An evaluation of clinical performance of FTA cards for HPV 16/18 detection using cobas 4800 HPV Test compared to dry swab and liquid medium. *J Clin Virol* , 94, 67-71.

El-Haj, N. & Hoppe, C.(2018). Newborn Screening for SCD in the USA and Canada. *Int. J. Neonatal Screen*, 4(4),36.

Fata,A., Khamesipour, A., Mohajery, M., Hosseininejad, Z., Afzalaghaei, M., Berenji, F., Ganjbakhsh, M., Akhavan, A., Eskandari, E.& Amin-Mohammadi, A.(2009). Whatman Paper (FTA Cards) for Storing and Transferring Leishmania DNA for PCR Examination." *Iran J Parasitol*, 4(4), 1.

Grosse, S.D., Odame, I., Atrash, H.K., Amendah, D.D., Piel, F.B. & Williams, T.N. (2011). Sickle cell disease in Africa: a neglected cause of early childhood mortality. *Am J Prev Med*, 41(6 Suppl 4), S398-405.

Kitenge, R., Tshilolo, L., Loko, G., Wamba, G. & Gonzalez, J.P.(2018). Diagnostic tools and follow-up of sickle-cell anemia in Central Africa. *Med Sante Trop*,28(2), 124-127.

Leeflang, M.M., Rutjes, A.W., Reitsma, J.B., Hooft, L. & Bossuyt, P.M. (2013) . Variation of a test's sensitivity and specificity with disease prevalence. *CMAJ*,185(11), E537-544.

Marshall, C.R., Chowdhury,S., Taft, R.J., Lebo, M.S.,Buchan, J.G., Harrison, S.M., Rowsey, R., Klee, E.W., Liu, P., Worthey, E.A., Jobanputra, V., Dimmock, D., Kearney, H.M.,Bick,D., Kulkarni, S., Taylor, S.L., Belmont, J.W., Stavropoulos, D.J. & Lennon, N.J (2020) .Best practices for the analytical validation of clinical whole-genome sequencing intended for the diagnosis of germline disease. *NPJ Genom Med*, 5, 47.

Mas, S., Crescenti, A., Gasso, P., Vidal-Taboada, J.M.& Lafuente, A. (2007). DNA cards: determinants of DNA yield and quality in collecting genetic samples for pharmacogenetic studies. *Basic Clin Pharmacol Toxicol*, 101(2), 132-137.

Mikobi TM, Lukusa Tshilobo P, Aloni MN, Akilimali PZ, Mvumbi-Lelo G & Mbuyi-Muamba, J.M.(2017). Clinical phenotypes and the biological parameters of Congolese patients suffering from sickle cell anemia: A first report from Central Africa. *J Clin Lab Anal* , 31(6).

Miller, S.A. , Dykes, D.D. & Polesky, H.F. (1988). A simple salting out procedure for extracting DNA from human nucleated cells. *Nucleic Acids Res*,16(3), 1215.

Peluso, A.L., Cascone, A.M., Lucchese, L., Cozzolino, I., Ieni, A., Mignogna, C., Pepe, S.& Zeppa, P. (2015). Use of FTA cards for the storage of breast carcinoma nucleic acid on fine-needle aspiration samples. *Cancer Cytopathol* , 123(10), 582-592.

Piccin, A.(2010). Do we need to test blood donors for sickle cell anaemia? *Blood Transfus*, 8 (3),137-8.

Piccin, A., Fleming, P., Eakins, E., McGovern, E., Smith, O.P.& McMahon, C.(2008). Sickle cell disease and dental treatment. *J Ir Dent Assoc*,54 (2),75-79.

Rabodoarivelo, M.S., Imperiale, B., Andrianiavomikotroka, R., Brandao, A., Kumar, P., Singh, S., Ferrazoli, L., Morcillo, N., Rasolofo, V., Palomino, J.C., Vandamme, P.& Martin, A.(2015). Performance of Four Transport and Storage Systems for Molecular Detection of Multidrug-Resistant Tuberculosis. PLoS One ,10(10), e0139382.

Rajendram, D., Ayenza, R., Holder, F.M., Moran, B., Long, T. & Shah, H.N. (2006). Long-term storage and safe retrieval of DNA from microorganisms for molecular analysis using FTA matrix cards. *J Microbiol Methods* , 67(3), 582-592.

Siegel, C.S., Stevenson, F.O.& Zimmer, E.A. (2017) . Evaluation and comparison of FTA card and CTAB DNA extraction methods for non-agricultural taxa. *Appl Plant Sci*,15,5(2),apps.1600109.

Shongo, M.Y., Mukuku, O., Lubala,T.K., Mutombo, A.M., Kanteng, G.W., Umumbu, W.S., Lukamba, R.M., Wembonyama, S.O. , Luboya, O.N. (2014) . Sickle cell disease in stationary phase in 6-59 months children in Lubumbashi: epidemiology and clinical feature. *Pan Afr Med* J,19: 71.

Tshilolo, L., Mukendi, R.& Girot, R. (1996). Sickle cell anemia in the south of Zaire. Study of two series of 251 and 340 patients followed-up 1988-1992. *Arch Pediatr*, 3(2), 104-111.

Tshilolo, L. , Aissi, L.M., Lukusa, D., Kinsiama, C., Wembonyama, S., Gulbis, B. & Vertongen, F. (2009). Neonatal screening for sickle cell anaemia in the Democratic Republic of the Congo: experience from a pioneer project on 31 204 newborns. *J Clin Pathol*, 62(1), 35-38.Vandepitte, J. & Stijns, J. (1963). Hemoglobinoses in the Congo (Léopoldville) and in Ruanda-Urundi. *Ann Soc Belg Med Trop*,43, 271-281.

APPENDIX TO RESEARCH ARTICLE 1 1.

Validity and utility of DNA-based test on DBS in post-transfusion period

In DR Congo, like in other tropical countries, multiple infectious diseases such as malaria and other parasitic diseases are endemic. Those diseases have a high potential of causing severe anemia, which is treated with blood transfusion. Hemoglobin-based tests are generally not recommended after transfusion. It is a common practice to delay the test for 3 months after a blood transfusion. The reason is that the test would also pick the normal hemoglobin from the donor in the bloodstream of the receiver, and falsely report a SCA patient rather as a heterozygote. This situation creates a niche that is inaccessible by most hemoglobin-based tests and represents an opportunity for DNA-based tests.

We are currently conducting a study to determine the reliability of DNA-based tests in the immediate aftermath of a transfusion and the kinetics of the depletion of donors' blood in Congolese adult SCA patients who received a transfusion.

Adult SCA patients are recruited from "CMMASS" and St Crispin Hospitals, in Kinshasa, DR Congo. To be included, patients had to be prescribed a transfusion, and agree to be sampled multiple times according to our study test schedule. SCA patients undergoing a blood transfusion exchange program and those with previous transfusion within 3 months prior to the recruitment were excluded. The sampling schedule is as follows: (1) before transfusion, (2) 1 hour, (3) 24 hours, (4) 48 hours, (5) 7 days, (6) 30 days, (7) 45 days, (8) 60 days, (9) 75 days, (10) 90 days post-transfusion. At each sampling point, peripheral blood is collected in a 1ml EDTA tube for Hb capillary electrophoresis (Minicap, SEBIA, France) and total blood cells count; whereas 2 drops are collected on each the FTA Elute cards and the Rapid Diagnostic Test (RDT) Sickle SCAN™ (BioMEDOMICS, USA). Samples collected on FTA cards are tested with FTA-RFLP.

Six SCA patients are recruited thus far. Samples collected before transfusion (red blood cell transfusion) confirmed SS diagnostic in all patients for all the 3 testing technics: FTA-RFLP, RDT, and capillary electrophoresis. The DNA-based test indicated that the molecular profile of patients remains the SS after the transfusion. Conversely, capillary electrophoresis and RDT returned instead the AS profile in the 6 individuals (figure 4).

These preliminary results indicate that a DNA-based test is reliable as it is not influenced by the donor's DNA. This test can be offered immediately after transfusion, avoiding delaying the test until the Hb-based test becomes reliable.

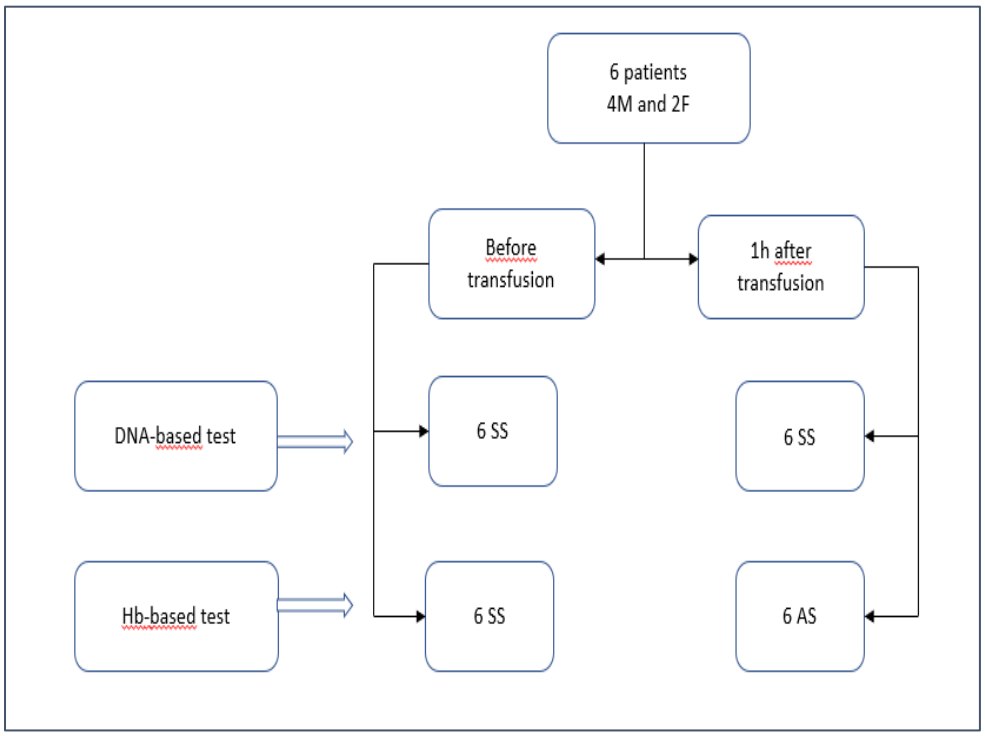

Figure 4. DNA and Hb-based tests profile after transfusion

3.2 PART 2. Genetic factors associated with HbF levels and HU response

Introduction

In the second part of the project, we focused on the evaluation of genetic factors involved in modulating the clinical and biological expression of SCA, as well as the pharmacogenetics of HU. Due to the sample size and the available funds, we opted for a candidate SNP approach exploring only SNPs with previously reported associations in the literature and public databases.

In the sections below we have first described the general methodology applied for the study of selected SNPs (research article 2). Then, we report specific sub-analyses in the format of original articles to be submitted after our thesis defense (research articles 3 and 4).

General methodology

Selection of patients, sample collection and preparation, and compilation of clinical parameters

Patients selection

A cross-sectional study was conducted between January 2017 and December 2018 at Kisantu St Luc Hospital (KSLH), located in the Kongo Central Province, and at the "Centre de Médecine Mixte et d'Anémie SS" (CMMASS), in the Kinshasa, both in the Democratic Republic of Congo (DRC). Pediatric patients were recruited at KSLH, whereas adult patients were recruited at "CMMASS" Hospital. In both sites, recruitment was performed during regular SCA follow-ups.

Prior to the inclusion in this study, SCA homozygous status had to be confirmed by molecular testing as previously described (Mikobi et al. 2017). In addition, we excluded confirmed SCA patients, those not in a steady state of the disease, those who received a transfusion within 4 months prior to blood sample collection, and those under Hydroxyurea treatment or other medications that may affect the blood counts taken during the previous 3 weeks.

Sample collection

For each participant, peripheral blood was collected in two EDTA-coated tubes of 4ml for hematologic tests and DNA extraction.

Hematological and clinical variables

A blood count was performed on an automated counter (Sysmex hematology counter, Japan). Measured hematologic parameters were: hemoglobin (Hb), white blood cells (WBC), platelets (PLT), and reticulocytes (RETIC). The HbF level was quantified by capillary electrophoresis (Minicap, SEBIA, France).

DNA extraction and quality control

Genomic DNA was extracted from blood samples using the salting out method (Miller et al. 1988) and stored at -20°C in the Center for Human Genetics of the University of Kinshasa (DR Congo). The concentration and purity of the extracted DNA were evaluated locally using a NanoDrop 2000 (Thermo Fisher Scientific, Wilmington, USA).

SNPs selection

Candidate SNPs (Table 1) associated with the baseline HbF level were retrieved from literature, searching PubMed using “fetal hemoglobin polymorphisms”. We also extended the search to relevant papers from the reference lists of articles retrieved through Pubmed search. We also searched in GWAS Catalog using “ fetal hemoglobin” AND “sickle cell anemia”. The candidate SNPs associated with Hydroxyurea-induced HbF level were also retrieved through the same process using “fetal hemoglobin” AND “Hydroxyurea”. We also extend the research to the PharmGKB database using “fetal hemoglobin hydroxyurea”. The last accession to all those databases and sites was in December 2018.

Our SNP set contained the 4 polymorphic positions defining the five βs- globin haplotypes (Shaikho et al.2017). In addition, 63 SNPs were included in the study of other genomic modifiers exerting pharmacogenomic variants. The 63 SNPs were distributed in 16 SNPs associated with HbF baseline level; 34 SNPs associated with a change in HbF levels in response to Hydroxycarbamide treatment; 13 SNPs associated with both HbF baseline level and HbF level in response to Hydroxycarbamide treatment.

Table 1. Candidate SNPs

Gene	SNP	Chromosomic location GR Ch37	Alleles	Mechanisms
HAO2	rs10494225	1:119918103	C>G	HU metabolism
BCL11A	rs10189857	2:60713235	A>G	HbF baseline
	rs1427407	2:60718043	T>G	HbF baseline HU-induced HbF
	rs7599488	2:60718347	C>T	HbF baseline
	rs766432	2:60719970	C>A	HbF baseline HU-induced HbF
	rs11886868	2:60720246	C>T	HbF baseline HU-induced HbF
	rs4671393	2:60720951	A>G	HbF baseline HU-induced HbF
	rs7557939	2:60721347	G>A	HbF baseline HU-induced HbF
	rs7606173	2:60725451	G>C	HbF baseline HU-induced HbF
VEGFA	rs2146323	6:43745095	C>A	HU-induced HbF
	rs3024997	6:43745107	G>A	HU-induced HbF
	rs10434	6:43753212	A>G	HU-induced HbF
ARG1	rs17599586	6:131904719	C>T	HbF baseline HU-induced HbF HU metabolism
HBS1L-MYB	rs28384513	6:135376209	T>G	HbF baseline HU-induced HbF HU metabolism
	rs9376090	6:135411228	T>C	HbF baseline
	rs7776054	6:135418916	A>G	HbF baseline
	rs9399137	6:135419018	T>C	HbF baseline HU-induced HbF
	rs4895440	6:135426558	A>T	HbF baseline
	rs4895441	6:135426573	A>G	HbF baseline
	rs9389269	6:135427159	T>C	HbF baseline
	rs9402686	6:135427817	G>A	HbF baseline
	rs9494142	6:135431640	T>C	HbF baseline
	rs11154849	6:136353305	C>T	HU-induced HbF
PDE7B	rs9376173	6:136359446	A>C	HU-induced HbF
	rs1480642	6:136499528	C>T	HU-induced HbF
	rs487278	6:136501828	G>A	HU-induced HbF

MAP3K5	rs9376230	6:137102365	C>A	HU-induced HbF
	rs9483947	6:137105400	C>T	HU-induced HbF
TOX	rs2693430	8:59725048	A>G	HU-induced HbF
	rs826729	8:59738913	G>A	HU-induced HbF
	rs765587	8:59790903	T>C	HU-induced HbF
	rs12155519	8:59848830	G>A	HU-induced HbF
	rs9693712	8:59947423	C>T	HU-induced HbF
	rs172652	8:59958141	A>G	HU-induced HbF
	rs380620	8:59982532	C>G	HU-induced HbF
KLF10	rs3191333	8:103662219	G>A	HU-induced HbF
ASS1	rs10901080	9:133358656	G>A	HU-induced HbF
	rs10793902	9:133377910	C>T	HU-induced HbF
SAR1A	rs4282891	10:71930069	A> C	HbF baseline HU-induced HbF
	rs2310991	10:71931646	C>A	HbF baseline HU-induced HbF
HBB	rs968857	11:5260458	T>C	βs haplotypes
	rs10128556	11:5263683	C>T	βs haplotypes
	rs28440105	11:5269799	A>C	βs haplotypes
	rs3834466	11:5291564	G>T	βs haplotypes
	rs7482144	11:5276169	G>A	HbF baseline HU-induced HbF
	rs7130110	11:5296104	G>C	HbF baseline
OR51B5	rs5006884	11:5373251	C>T	HbF baseline
	rs5024042	11:5373562	C>A	HbF baseline
NOS1	rs816361	12:117655131	C>G	HU-induced HbF
	rs7309163	12:117729274	C>G	HU-induced HbF
	rs7977109	12:117730340	G>A	HU-induced HbF
FLT1	rs3751395	13:28958955	C>A	HU-induced HbF
	rs9319428	13:28973621	G>A	HU-induced HbF
	rs2182008	13:28987061	A>G	HU-induced HbF
	rs2387634	13:28990428	T>C	HU-induced HbF
FLT1	rs8002446	13:28997400	G>A	HU-induced HbF
SALL2	rs61743453	14:21991343	G>C	HU-induced HbF
ARG2	rs2295644	14:68066559	A>T	HU-induced HbF
	rs10483801	14:68117006	C>A	HU-induced HbF
	rs10483802	14:68117421	T>C	HU-induced HbF
SIN3A	rs11072544	15:75666081	G>A	HbF baseline HU-induced HbF
	rs7166737	15:75682513	A>G	HbF baseline HU-induced HbF
GLP2R	rs12103880	17:9698342	G>A	HbF baseline HU-induced HbF
NOS2	rs1137933	17:26105932	G>A	HU-induced HbF
	rs944725	17:26109571	C>T	HU-induced HbF
PHEX	rs12559632	X:22160992	G>A	HbF baseline
MAGEB18	rs6630120	X:25736480	A>G	HbF baseline

SNPs genotyping method

Probe design

We designed 64 Molecular Inversion Probes (MIPs) to cover the 67 SNPs, including 2 MIPs covering 2 SNPs each. The genomic coordinates of the design were based on GRCh37. The MIPs were 108 to 120 bp long. Our final Region of Interest (ROI) was 1,768 bp large.

Next Generation Sequencing and data analysis

Samples were normalized to 100 ng/µl at the Genomics Core of the Center for Human Genetics of the University Hospitals Leuven (UZ Leuven) and checked on a Qubit 2.0 Fluorometer (Life Technologies, Bleiswijk, Netherlands). Patient-only targeted sequencing was performed using Single Stranded Molecular Inversion Probe (ssMIP) approach using our custom library preparation kit. Single Stranded Molecular Inversion Probe (ssMIP) is a Next Generation Sequencing (NGS) technique using single-stranded molecular probes containing at their extremities two hybridization sequences, complementary to the target DNA fragment (Cantsilieris et al. 2017), allowing the targeted sequence to be captured and the remaining linear DNA to be degraded by exonucleases.

A total of 100 ng DNA was used for library preparation. After capturing with ssMIPs, samples were pooled either by 76 for pools 1 and 2 or by 86 or 100 for pools 3 and 4, respectively. Pool 4 contained both the remaining samples and samples that failed in the three other 3 pools.

To reach our target coverage of 500x, pool 1, 2, and 3 were sequenced on a MiSeq v3 PE300 cycles flow cell (13.2–15 Gb output), whereas pool 4 was run on a MiSeq v2 PE300 cycles (4.5–5.1 Gb output). A concentration of 12.5pM was loaded on the MiSeq, as per the recommendations in MiSeq manual. A PhiX spike-in of 3% was also added.

The quality control of sequencing reads was performed using FastQC. Overlapping paired-end reads from the same fragment were merged using FLASH 1.2.11 (Magoc et al. 2011). Reads were then mapped against the reference genome build GRCh37 with BWA-mem 0.7.17 (Li et al. 2009). Duplicate reads were kept. All positions were genotyped with GATK HaplotypeCaller 4.0.11.0 with the option EMIT_ALL_SITES (McKenna et al. 2010). All samples were genotyped together using GATK GenotypeGVCFs 3.8-0.

Hardy-Weinberg equilibrium was tested with PLINK.

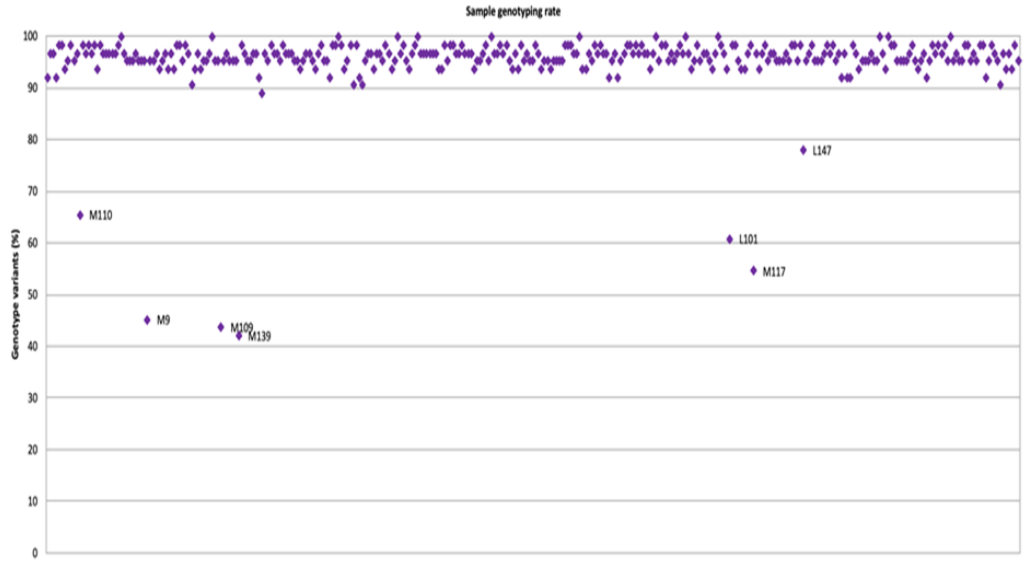

Figure 1. Quality of genotype call per SNPs

Note. The Y axis indicates the % of variants passing the genotype quality check for samples in the X axis. Samples M110, M9, M109, M139, L101, M117, L147, and L88 have less than 90% of their SNPs passing the genotype QC.

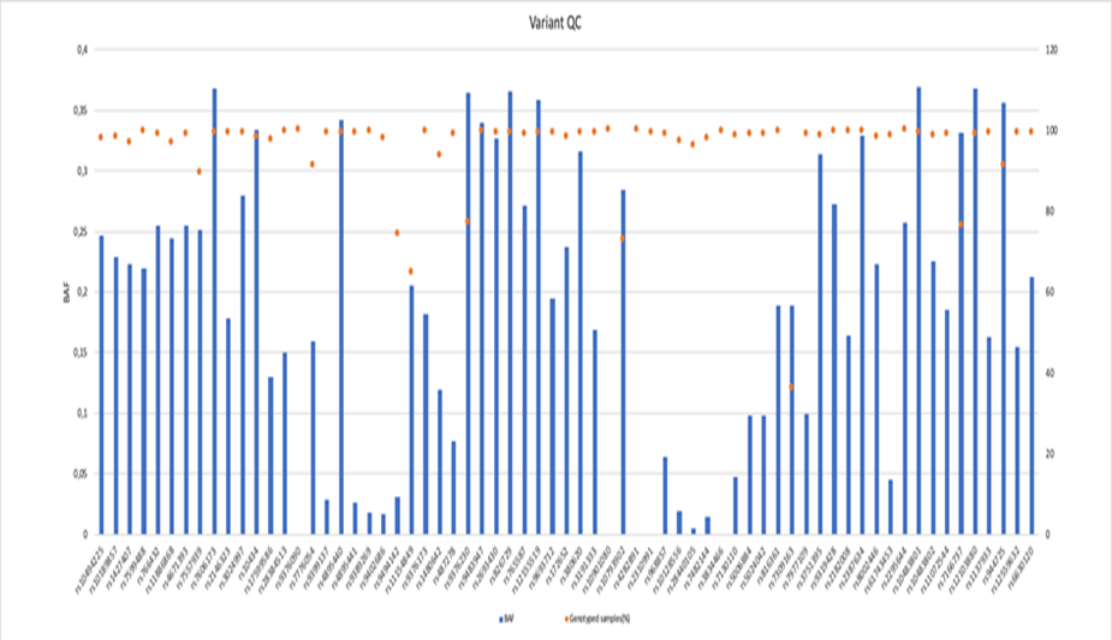

Figure 2. Variant quality control for each SNP

Note. The Y-axis indicates both the percentage (orange stars) in which an SNP (X-axis) has been well-genotyped and the B-allele frequency of the same SNP.

No deviation from Hardy-Weinberg was observed.

References

Cantsilieris, S, Stessman, H A, Shendure, J.& Eichler, E E. (2017). Targeted Capture and High-Throughput Sequencing Using Molecular Inversion Probes (MIPs). *Methods Mol Bio,*1492,95-106.

Li, H. & Durbin, R. (2009). Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics,*25(14), 1754-176).

Magoč, T. & Salzberg, S.L.(2011). FLASH: fast length adjustment of short reads to improve genome assemblies. *Bioinformatics,*27(21),2957-63.

McKenna, A., Hanna, M., Banks, E., Sivachenko, A., Cibulskis, K., Kernytsky, A., Garimella, K., Altshuler, D., Gabriel, S., Daly, M. & DePristo, M.A. (2010). The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res,*20(9),1297-303.

Mikobi, T.M., Lukusa, T.P., Aloni, M.N., Akilimali, P.Z., Mvumbi-Lelo, G.& Mbuyi-Muamba, J.M. (2017). Clinical phenotypes and the biological parameters of Congolese patients suffering from sickle cell anemia: A first report from Central Africa. J Clin Lab Anal,31(6),e22140)

Miller, S.A., Dykes, D.D. & Polesky, H.F . (1988). A simple salting out procedure for extracting DNA from human nucleated cells. *Nucleic Acids Res,*16(3) ,1215.

Shaikho, E.M., Farrell, J.J., Alsultan, A., Qutub, H., Al-Ali, A.K., Figueiredo, M.S., Chui, D.H.K., Farrer, L.A., Murphy, G.J., Mostoslavsky, G., Sebastiani, P. & Steinberg, M.H. (2017). A phased SNP-based classification of sickle cell anemia HBB haplotypes. *BMC Genomics*, 18(1),608.

RESEARCH ARTICLE 2

3.2.1 The betaS-globin haplotypes do not likely explain clinical and biological diversity in DR Congolese sickle cell anemia patients

TO BE SUBMITTED

Mamy Ngole[1,2,3], Paul Lumbala[1,2,4], Gloire Mbayabo[1,2,4], Valerie Race[1], Nono Mvuama[5], Isabelle Cleynen[1], Stephanie Deman[6], Erika Souche[6], Prosper Tshilobo Lukusa[2,4], Tite Mikobi [2],Chris Van Geet [7], Koenraad Devriendt[1], Gert Matthijs[1], Aimé Lumaka [2,4,8]

[1]Center for Human Genetics, Faculty of Medicine, KU Leuven, Leuven, Belgium

[2]Center for Human Genetics, Faculty of Medicine, University of Kinshasa, Kinshasa, Democratic Republic of Congo

[3]Department of Medical Biology, Faculty of Medicine, University of Kinshasa, Kinshasa, Democratic Republic of Congo

[4]Department of Pediatrics, Faculty of Medicine, University of Kinshasa, Kinshasa, Democratic Republic of Congo

[5]Kinshasa School of Public Health, Faculty of Medicine, University of Kinshasa, Kinshasa, Democratic Republic of Congo

[6]Genomics Core KULeuven - UZLeuven, Leuven, Belgium

[7]Department of Cardiovascular Sciences and Pediatrics (Hemato-oncology), KU Leuven and UZ Leuven, Belgium

[8]GIGA-R, Laboratoire de Génétique Humaine, University of Liège, Liège, Belgium

Abstract

Background: Sickle cell anemia (SCA) is characterized by an important phenotypic heterogeneity that is not fully understood yet.

This study aims to determine the distribution of βS-globin haplotypes among Congolese SCA patients and assess the correlation between haplotypes and baseline hematological and clinical profiles.

Method : A cross-sectional study was conducted with SCA patients recruited at St Luc Hospital in Kisantu and at the "Centre de Médecine Mixte et d'Anémie SS" in Kinshasa, both in the Democratic Republic of Congo. Blood count was performed on an automated counter. SCA homozygous status was confirmed by restriction fragment length polymorphism. βS haplotypes were determined manually on data generated by Molecular Inversion Probes (MIPS) sequencing. Patients were classified according to clinical phenotypes scores.

Results: Among 269 homozygous SCA patients, the haplotype CAR was the predominant allele (95.92%), followed by BEN (1.67%), SEN (1.31%), and AI (0.37%). Atypical haplotypes accounted for 0.74 % of alleles. Six haplotypic genotypes were identified, including homozygous CAR (92.2%), homozygous BEN (0.37%), and double heterozygous CAR/BEN (2.59%), CAR/SEN (2.59%), CAR/AI (0.74%) and CAR/ATYP (1.48%). Less than 15% of patients within the CAR/CAR haplotype presented a severe clinical phenotype, representing 11% among children and 8% among adults. No correlation was observed between the haplotypes and either hematological or clinical parameters.

Conclusion : The CAR haplotype is predominant among DR Congolese SCA patients while patients still have variable hematologic and clinical profiles. However , the almost homogeneous distribution of CAR haplotypes did not allow the evaluation of the influence of other haplotypes.

Key words : sickle cell anemia, β-globin gene haplotypes, hematological parameters, clinical severity score, DR Congo

Introduction

Despite being a monogenetic disease and almost always caused by the same single-point mutation (NM_000518.5(HBB): c.20A>T; *p*. Glu7Val), Sickle Cell Anemia (SCA), is characterized by a great inter-individual phenotypic variability, both in terms of types and severity of manifestations (Driss et al. 2009; Aghajani et al. 2017). Although this variability is not fully understood yet, βS haplotypes are included among genetic factors explaining the phenotypic heterogeneity observed in SCA (Steinberg et al. 2009; Ndugwa et al. 2012).

Traditionally, βS-globin haplotypes are defined by the profile of restriction fragments generated from polymorphic restriction sites in a genomic region of ~60kb located within the HBB cluster on chromosome 11. These haplotypes are inherited along with the SCA mutation (Okumura et al.2019). The five classical haplotypes of the βS-globin gene are named according to the geographical region where they were originally identified. Four βS haplotypes constitute the major types in Africa, named Cameroon haplotype (CAM), Benin haplotype (BEN), Senegal haplotype (SEN), and Bantu or Central African Republic (CAR) haplotype (Pagnier et al.1984; Lapouméroulie et al. 1992). In addition, the Arabian Indian haplotype (AI) was identified in India and the Arabian peninsula (Kulozik et al.1986).

Using RFLP for the identification of these haplotypes is not only time-consuming but may also result in unusual restriction patterns in the presence of unexpected SNP at the restriction site, without precisely determining the nucleotide change at the polymorphic site (Joly et al. 2011). The unusual restriction patterns are known as atypical βS haplotypes (Srinivas et al.,1988; Zago et al. 2000; Okumura et al. 2019). Nowadays, the application of Next Generation Sequencing (NGS) techniques offers the advantages of multiplexing samples and interrogating multiple loci in a single experiment, and the possibility of fully characterizing atypical βs haplotypes (Shaikho et al. 2017).

Modifying effects of SCA haplotypes are in part related to their association with the level of fetal hemoglobin (HbF) (Al-Saqladi et al. 2010). HbF acts by interrupting HbS sickling through a mechanic effect, resulting in reduced severity of the disease (Powars et al. 1989; El-Hazmi et al. 2011). The CAR haplotype is known to present a severe phenotype due to its association with a low average HbF level of about 5%. Compared to the CAR haplotype, the haplotypes BEN and CAM, with an intermediate level of HbF of 7%, have moderate symptomatology (Piel et al.2017). The SEN and AI haplotypes, which are associated with an average of 10 to

20% HbF in adult SCA patients, present milder phenotypic patterns (Rees et al. 2010; Habara et al. 2017).

The Democratic Republic of Congo (DRC) is among the most affected countries by SCA worldwide, with an incidence of around 2% in newborns (Tshilolo et al,2009). Surprisingly, data on the effect of SCA haplotypes on disease severity are unavailable for the DRC. Yet this knowledge could be helpful to explain the phenotypical diversity previously observed in Congolese SCA patients (Mikobi et al. 2017). This study aimed at determining the distribution of SCA haplotypes and investigating their correlation with baseline hematological parameters and clinical severity scores.

Methods

Patient selection, sample treatment, selection of SNPs set, and SNPs genotyping technique are presented in point 2.1 above.

In this study, patients recruited at "CMMASS" were classified according to the clinical phenotype score as described by Mikobi et al. 2015, whereas the severity score developed by Adegoke et al. 2013 was used for those from KSLH.

Data analysis

The data were analyzed with SPSS version 25 (IBM, USA), and a p-value less than 0.05 was considered significant. The Kolmogorov–Smirnov and the Shapiro–Wilk tests were used to evaluate the normal distribution of quantitative variables. The median ± interquartile space of selected hematological parameters was compared for each haplotype. When comparing hematologic parameters between haplotypes, the statistical threshold was determined using the non-parametric Mann–Whitney U test for 2 groups and Kruskal–Wallis test for more than two groups.

For this study, the SCA haplotypes were manually derived based on the genotype of the 4 SNPs of interest.

Ethics statement

Before the inclusion, parents and adult patients were fully informed of the aims, procedures, and risks of this study. Written informed consent was obtained from adult patients or parents

of minor children. Privacy was protected by the de-identification of samples. An individual alphanumeric code was used instead.

The study protocol was approved by the Ethical Committee of the Public Health School at the University of Kinshasa, DRC (Protocol number ESP/CE/079/2016).

Results

A total of 277 SCA patients were included in the study, aged between 2 and 40 years. The gender distribution was 112 (40%) males and 165 (60%) females.

After removing 8 samples with less than 90% of SNPs correctly genotyped , only samples of 269 patients were considered: 112 patients (42%) from KSLH, including 54 males and 58 females aged between 2 and 17 years; and 157 patients (58%) from « CMMASS », comprising 56 males and 101 females aged between 18 and 40 years.

Distribution of SCA haplotypes

Among the 269 patients (538 alleles), the haplotype CAR was the major allele in this study with 95.92% (516/538). The other classic haplotypes were BEN 1.67% (9/538), SEN 1.31% (7/538), and AI 0.37% (2/538). When grouped at genotype levels (i.e., the composition of the two haplotype alleles for a particular SNP), a total of 5 different genotypes were identified, including homozygous CAR 92.23% (248/269), homozygous BEN 0.37% (1/269), and double heterozygous CAR/BEN 2.59% (7/269), CAR/SEN 2.59% (7/269) and CAR/AI 0.74% (2/269) (Figure 1).

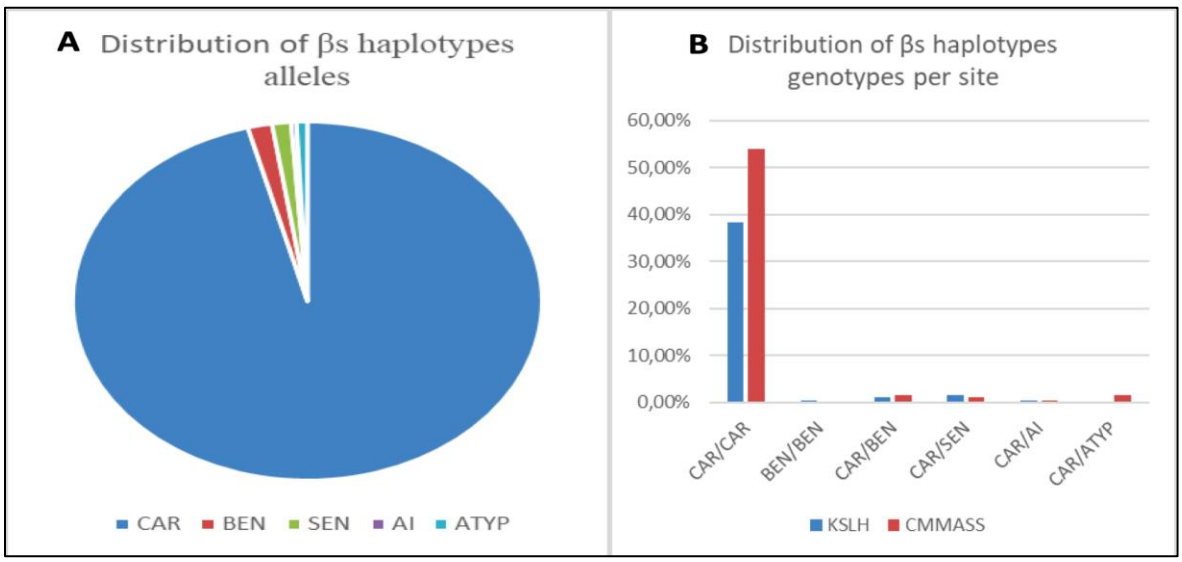

Figure 1. Distribution of βs haplotypes

Three different atypical haplotypes were identified, and one of them was observed in two individuals. Atypical haplotypes accounted for 0.74 % (4/538) of the alleles. Genotypically, all patients with atypical alleles were heterozygous CAR/ATYP (1.48%, 4/269). The atypical profiles were confirmed by Sanger sequencing to exclude the possibility of a false call. The profile and the frequency of the three atypical haplotypes are presented in Table 1.

Table 1. Description and frequency of typical and atypical haplotype alleles

HAPLOTYPES	SNPs				ALLELE FREQUENCY	
	rs3834466	rs28440105	rs10128556	rs968857		
AI	GT	C	T	T	2/538	0.37%
SEN	G	C	T	T	7/538	1.31%
BEN	G	C	C	T	9/538	1.67%
CAR	G	C	C	C	516/538	95.92%
CAM	G	A	C	T	0	0
Atypical 1	GT	C	C	C	2/538	0.37%
Atypical 2	G	A	C	C	1/538	0.18%
Atypical 3	GT	C	T	C	1/538	0.18%

We attempted to determine the nucleotide difference between typical and atypical haplotypes . We build a haplotype tree (Figure 2) based on the work by Okumura et al.2019.

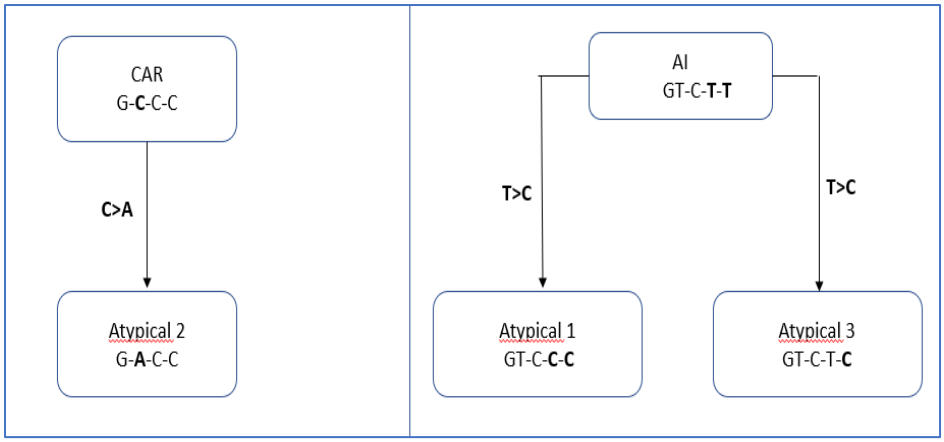

Figure 2. Atypical haplotype tree

The atypical 2 appeared to be the closest to the CAR haplotype, from which these may derive after a single event. The atypical 1 and 3 were closer to the AI and a single event may have caused the development of either atypical haplotype.

Influence of haplotypes on hematological parameters and clinical score

Due to incomplete hematological data, 8 of 248 CAR/CAR patients, and 1 of 7 CAR/BEN patients, were removed from this analysis.

The almost homogeneous distribution of βs haplotypes in our cohort do not allow evaluation of influence of other haplotypes in hematological and clinical parameters.

Table 2. Distribution of haplotypes, clinical severity and hematological parameters at KSLH

	N (%)	Severity score			HbF (%)	Hb (g/dl)	WBC (n/µl)	Platelets (n/µl)	Reticulocytes (n/µl)
		MI	Mo	S	Med ± IQR				
CAR-CAR	96	38	47	11	7.7 ± 2.5	7.3 ± 1.8	12745 ± 6875	396000 ± 238250	160418.5 ± 148733
BEN-BEN	1	0	1	0	5	5.3	21540	211000	157927
p (M-W) CAR-CAR vs BEN-BEN					0.381	0.175	0.239	0.175	1
CAR-BEN	3	0	3	0	15.1	6.6	12105	222500	183736.5
p (M-W) CAR-CAR VS CAR-BEN					0.547	0.894	0.307	0.589	0.651
CAR-SEN	4	0	3	1	10.3	5.8	10140	256000	224411.5
p (M-W) CAR-CAR vs CAR-SEN					0.629	0.159	0.232	0.054	0.333
CAR-AI	1	0	1	0	9.5	8.8	10260	260000	79208
p (M-W) CAR-CAR vs CAR-AI					0.761	0.192	0.344	0.292	0.276
p (K-W)					0.937	0.285	0.351	0.191	0.565

Med ± IQR: median ± interquartile range M-W: Mann–Whitney U K-W: Kruskal–Wallis
Mi=Mild Mo=Moderate S=severe
The Mann–Whitney U test was used for pairwise comparisons between the CAR-CAR genotype and each of the other six genotypes; the Kruskal–Wallis test for multiple analyses between the six haplotype groups.

Table 3. Distribution of haplotypes, clinical severity and hematological parameters at CMMASS

Haplotypes	N (%)	Severity score			HbF (%)	Hb (g/dl)	WBC (n/µl)	Platelets (n/µl)	Reticulocytes (n/µl)
		A	M	S	Med ± IQR				
CAR-CAR	144	44	86	14	6.46 ± 3.2	7.2 ± 1.9	10270 ± 5678	453409 ± 308020	117211 ± 99459
CAR-BEN	4	1	2	1	13.22	7.6	7910	380000	92781
p (M-W) CAR-CAR vs CAR-BEN					0.530	0.453	0.945	0.662	0.304
CAR-ATYP	4	1	2	1	3.5	8.2	10255	464000	93954
p (M-W) CAR-CAR vs CAR-ATYP					0.49	0.66	0.398	0.99	0.595
CAR-SEN	2	0	2	0	10.3	7.6	14045	458500	100929
p (M-W) CAR-CAR vs CAR-SEN					0.79	0.448	0.906	0.983	0.722
CAR-AI	1	0	1	0	14.7	9	10160	423000	84932
p (M-W) CAR-CAR vs CAR-AI					0.19	0.360	0.901	0.976	0.401
p (K-W)					0.614	0.70	0.945	0.977	0.720

Med ± IQR: median ± interquartile range M-W: Mann–Whitney U K-W: Kruskal–Wallis

A: Asymptomatic clinical phenotype; M: Moderate clinical phenotype; S: Severe clinical phenotype

Discussion

The βs haplotypes are among the genetic factors linked the phenotypic variability observed in SCA in the literature. In this study, we assessed the distribution of βs haplotypes and their correlation with baseline hematologic parameters in a cohort of Congolese SCA patients.

We observed a predominance of the CAR haplotype (at a frequency higher than 90%) similar to studies in neighboring African countries, including Angola (CAR-CAR at 92,15% in 192 SCA children; CAR-CAR at 82,2% in 12 SCA newborns) (Borges et al. 2019; Delgadinho et al. 2021), Uganda (CAR-CAR at 89% in 104 SCA patients) (Ndugwa et al. 2012), Congo Brazzaville (CAR-CAR at 85,7% in 54 SCA pygmies) (Mouélé et al. 2000). The same trend has been recently reported in a multicentric study involving DRC, Angola, Uganda, and Kenya (CAR-CAR at 92% in 635 SCA children) (Tshilolo et al., 2019) .

This observation of a nearly homogeneous haplotypic genotype contrasts with studies from other African, European, and American countries. In many American countries, the heterogeneity of βS haplotypes is explained by the flow of Africans from different regions during the slave trade period (Fong et al. 2013; Lindenau et al.2016). In Europe, the diversity of βS haplotypes is essentially due to migrations from several African regions (Bernaudin et al.2018). In Africa, the distribution of βS haplotypes is highly diversified. In some countries such as Cameroun, Sudan, Egypt, all African βS haplotypes are represented, with a higher frequency of the Benin haplotype (Elderdery et al. 2012; Bitoungui et al.2015; Abou-Elew et al.2018).

The βS haplotype homogeneity in our region may reflect limited admixture with other populations. However, the presence of the BEN, SEN and AI haplotypes even at very low frequency could result from recent admixture of West Africans, Portuguese, Lebanese and Indians with the local population through commercial transactions (Flahaux & Schoumaker, 2016). We could not find any anthropological explanation for the absence of the CAM haplotype in our study sample.

The frequency of 0.74% of atypical haplotypes reported in our study is lower than the 5 to 10% often observed (Zago et al. 2000). For instance, 5.7% of atypical haplotypes with 11 different patterns were reported in Brazil (Srinivas et al.1988) and 11.8 % of 3 different patterns of atypical haplotypes in Yemeni SCA children (Al-Saqladi et al. 2010). Population’s admixture

is among factors involved in the multiplicity of atypical βS haplotypes. Therefore, our finding can be explained by the relative homogeneity of the studied population.

The atypical haplotypes were mainly co-inherited with the most frequent haplotype encountered, the CAR haplotype. The study by Adekile et al. in Kuwait reported an atypical haplotype with a profile close to the Atypical 1 identified in our study: GTCCC genotype in our study, and TTCCC in the Kuwait SCA cohort (Adekile et al.2021). More βs haplotype distribution studies are needed in the Central Africa region to evaluate the presence of this particular atypical haplotype in the DRC neighborhood.

The likely lack of significant association of βs haplotypes with hematologic parameters was also reported by Bitoungi et al. in a study involving one of the largest national cohorts of Sub-Saharan African SCA patients where the authors found mainly the BEN haplotype (Bitoungui et al.,2015). Both our studies could not identify significant differences in hematological profile among haplotypes, probably because of the large predominance of one specific haplotype and very low frequency of others.

Due to the previously reported association of the CAR haplotype with lower HbF levels, one would expect the CAR/CAR haplotype to be characterized by a more pronounced clinical severity (Piel et al. 2017). Interestingly, our results show that only less than 15% of patients with the CAR/CAR haplotype (11% in children and 8% in adults) present a severe clinical severity, while the majority of CAR/CAR patients exhibited a moderate or mild phenotype. This could as well reflect an ascertainment bias as we have not conducted a population-wide study, and it is likely that in adults, patients with the most severe spectrum of the disease have already died.

The levels of HbF in SCA individuals homozygous for CAR/CAR varies widely in different studies from different populations, as shown on Table 4.

Similar variations in HbF levels were reported in patients homozygous for other haplotypes. For instance, Alsultan et al identified significant differences in HbF levels between African and Saudi SCA patients, all sharing homozygous Benin haplotype. This suggests that other factors could be involved such as age, gender, environmental factors, other associated hemoglobinopathies (Alsultan et al.,2011) or other genetic markers.

Table 4. Levels of HbF within SCA patients homozygous CAR/CAR haplotype

Country	N	Age	CAR/CAR N (%)	Mean HbF (%)	Reference
DRC	269	2-40 years	240 (89.2%)	7.1 ±2.5 (median values)	This study
Brazil (Sao Paolo)	117	3 to 71 months	41 (35,04%)	14.6 (1.8 ±9.1) (mean ± SD)	(Camilo-Araujo et al.2014)
Panama	95SS	6 months – 15 years	39 (41%)	15.98 ± 2.26 (mean ± SD)	(Rusanova et al.2011)
France	580	5.8 ±4.8 years	216 (37.4%)	10.1 ±6.7 (mean ± SD)	(Bernaudin et al.2018)
Angola	192	3 – 12 years	176 (92,15%)	5,43 ±3,70 (mean ± SD)	(Delgadinho et al.2021)
Angola, DRC, Uganda, Kenya	600	5.4±2.4 years	552 (92%)	10,8 ±6,7% (mean ± SD)	(Tshilolo et al. 2019)

Conclusion

The CAR haplotype is predominant among DR Congolese SCA patients. Furthermore, the evaluation of influence of βs haplotypes on hematological and clinical parameters is limited by the very low representativeness of other haplotypes .

FINANCIAL SUPPORT

"Initiatives Sud" et projet "TEAM 2015" from VLIR-UOS

CONFLICT OF INTEREST

None

AUTHOR CONTRIBUTIONS

Prosper Lukusa, Koenraad Devriendt, Gert Matthijs and Aimé Lumaka designed the project and corrected the manuscript. Mamy Ngole, Paul Lumbala, and Gloire Mbayabo conducted patients' recruitment. Mamy Ngole and Cathy Songo performed laboratory analysis under the supervision of Valerie Race. Nono Mvuama and Isabelle Cleynen performed statistical analyses, Mamy Ngole prepared the manuscript; Prosper Lukusa, Koenraad Devriendt, Gert Matthijs, Aimé Lumaka, and Valerie Race edited the manuscript. All authors read and approved the final manuscript

ACKNOWLEDGMENTS

The authors are grateful to patients, parents, and the families for agreeing to participate in this study. We thank Bruno Vankeirsbilck of UZ Leuven Molecular Laboratory and his team for the technical and logistic assistance.

References

Abou-Elew H H, Youssry I, Hefny S, Hashem, R.H., Fouad, N., Zayed, R.A. (2018). β^S globin gene haplotype and the stroke risk among Egyptian children with sickle cell disease. *Hematology*, 23(6),362-367

Adegoke, S.A. & Kuti, B.P. (2013). Evaluation of clinical severity of sickle cell anemia in Nigerian children. *J Appl Hematol*, 4,58-64.

Aghajani, F, Mahdavi, M.R., Kosaryan, M., Mahdavi, M., Hamidi, M.& Jalali, H. (2017). Identification of β-globin haplotypes linked to sickle hemoglobin (Hb S) alleles in Mazandaran province, Iran. *Genes Genet Syst*,91(6),311-313.

Al-Saqladi, A.W., Brabin, B.J., Bin-Gadeem, H.A., Kanhai, W.A., Phylipsen, M.& Harteveld, C.L. (2010). Beta-globin gene cluster haplotypes in Yemeni children with sickle cell disease. *Acta Haematol*, 123(3),182-5.

Alsultan, A., Solovieff, N., Aleem, A., AlGahtani, F.H., Al-Shehri, A., Osman, M.E., Kurban, K., Bahakim, H., Al-Momen, A.K., Baldwin, C.T., Chui, D.H.& Steinberg, M.H. (2011). Fetal hemoglobin in sickle cell anemia: Saudi patients from the Southwestern province have similar HBB haplotypes but higher HbF levels than African Americans. *Am J Hematol*, 86(7),612-4.

Bernaudin, F., Arnaud, C., Kamdem, A., Hau, I., Lelong, F., Epaud, R., Pondarré, C. & Pissard, S. (2018). Biological impact of α genes, β haplotypes, and G6PD activity in sickle cell anemia at baseline and with hydroxyurea. *Blood Adv*,2(6),626-637.

Bitoungui,V.J., Pule, G.D., Hanchard, N., Ngogang, J. & Wonkam, A. (2015). Beta-globin gene haplotypes among cameroonians and review of the global distribution: is there a case for a single sickle mutation origin in Africa? OMICS, 19(3),171-9.Borges, E., Tchonhi, C., Couto, C.S.B., Gomes, V., Amorim, A., Prata, M.J. & Brito, M. (2019).Unusual β-Globin Haplotype Distribution in Newborns from Bengo, Angola. *Hemoglobin*, 43(3),149-154.

Camilo-Araújo, R.F., Amancio, O.M., Figueiredo, M.S., Cabanãs-Pedro, A.C.& Braga, J.A.(2014) Molecular analysis and association with clinical and laboratory manifestations in children with sickle cell anemia. *Rev Bras Hematol Hemoter*,36(5),334-9.

Delgadinho, M., Ginete, C., Santos, B., Miranda, A. & Brito, M.(2021). Genotypic Diversity among Angolan Children with Sickle Cell Anemia. *Int J Environ Res Public Health*,18(10),5417.

Driss, A. ,Asare, K.O., Hibbert, J.M., Gee, B.E., Adamkiewicz, T.V.& Stiles, J.K.(2009) . Sickle Cell Disease in the Post Genomic Era: A Monogenic Disease with a Polygenic Phenotype. *Genomics Insights*, (2),23-48.

Elderdery, A.Y., Mills, J., Mohamed, B.A., Cooper, A.J., Mohammed, A.O., Eltieb, N. & Old, J. (2012). Molecular analysis of the β-globin gene cluster haplotypes in a Sudanese population with sickle cell anaemia. *Int J Lab Hematol*, 34(3),262-6.

El-Hazmi MA, Al-Hazmi AM, Warsy AS. Sickle cell disease in Middle East Arab countries. Indian J Med Res. 2011 ;134(5) :597-610

Flahaux, M.L. & Schoumaker, B. (2016). Democratic Republic of the Congo: A Migration History Marked by Crises and Restrictions. The online Journal of Migration Policy Institute;2016: April 20. Available from www.migrationpolicy.org (consulted on march 15th, 2022)

Fong, C., Lizarralde-Iragorri, M.A., Rojas-Gallardo, D. & Barreto, G. (2013). Frequency and origin of haplotypes associated with the beta-globin gene cluster in individuals with trait and sickle cell anemia in the Atlantic and Pacific coastal regions of Colombia. *Genet Mol Biol*, 36(4),494-7.

Habara, A.H., Shaikho, E.M.& Steinberg, M.H. (2017). Fetal hemoglobin in sickle cell anemia: The Arab-Indian haplotype and new therapeutic agents. *Am J Hematol*, 92(11),1233-1242

Joly, P., Lacan, P., Garcia, C., Delasaux, A. & Francina, A. (2011). Rapid and reliable β-globin gene cluster haplotyping of sickle cell disease patients by FRET Light Cycler and HRM assays. *Clin Chim Acta*, 412(13-14),1257-61.

Kulozik, A.E., Wainscoat, J.S., Serjeant, G.R., Kar, B.C., Al-Awamy, B., Essan, G.J., Falusi, A.G., Haque, S.K., Hilali, A.M. & Kate, S. (1986). Geographical survey of beta S-globin gene haplotypes: evidence for an independent Asian origin of the sickle-cell mutation. *Am J Hum Genet*,39(2),239-44.

Lapouméroulie, C., Dunda, O., Ducrocq, R., Trabuchet, G., Mony-Lobé, M., Bodo, J.M., Carnevale, P., Labie, D., Elion, J. & Krishnamoorthy, R. (1992). A novel sickle cell mutation of yet another origin in Africa: the Cameroon type. *Hum Genet, 8*9(3), 333-7.

Lindenau, J.D., Wagner, S.C., Castro, S.M.& Hutz, M.H. (2016). The effects of old and recent migration waves in the distribution of HBB*S globin gene haplotypes. *Genet Mol Biol*,39(4),515-523.

Mikobi, T.M., Lukusa, T.P., Aloni, M.N., Mvumbi, L.G., Akilimali, P.Z., Muyembe-Tamfum, J.J., Race, V., Matthijs, G. & Mbuyi, M.J.M.(2015). Correlation between the Lactate Dehydrogenase Levels with Laboratory Variables in the Clinical Severity of Sickle Cell Anemia in Congolese Patients. *PLoS One*,10(5),e0123568.

Mouélé R, Pambou O, Feingold J, Galactéros F. (2000) Alpha-thalassemia in Bantu population from Congo-Brazzaville: its interaction with sickle cell anemia. Hum Hered. 2000 Mar-Apr;50(2):118-25.

Ndugwa, C., Higgs, D., Fisher, C., Hambleton, I., Mason, K., Serjeant, B.E. & Serjeant, G.R. (2012). Homozygous sickle cell disease in Uganda and Jamaica a comparison of Bantu and Benin haplotypes. *West Indian Med J*, 61(7), 684-91.

Ndugwa, C., Higgs, D., Fisher, C., Hambleton, I., Mason, K., Serjeant, B.E.& Serjeant, G.R.(2012). Homozygous sickle cell disease in Uganda and Jamaica a comparison of Bantu and Benin haplotypes. *West Indian Med J*,61(7),684-91.

Okumura, J.V., Silva, D.G.H., Torres, L.S., Belini-Junior, E., Venancio, L.P.R., Carrocini, G.C.S., Nascimento, P.P., Lobo, C.L.C. & Bonini-Domingos, C.R. (2019). Atypical β-S haplotypes: classification and genetic modulation in patients with sickle cell anemia. *J Hum Genet*,64(3), 239-248.

Pagnier, J., Mears, J.G., Dunda-Belkhodja, O., Schaefer-Rego, K.E., Beldjord, C., Nagel, R.L. & Labie, D. (1984). Evidence for the multicentric origin of the sickle cell hemoglobin gene in Africa. *Proc Natl Acad Sci U S A*,81(6),1771-3.

Piel ,F.B., Steinberg, M.H.& Rees, D.C. (2017). Sickle Cell Disease. *N Engl J Med*,376(16),1561-1573.

Powars DR, Chan L, Schroeder WA. The influence of fetal hemoglobin on the clinical expression of sickle cell anemia. *Ann N Y Acad Sci*. 1989 ;565 :262-278

Rees, D.C., Williams, T.N.& Gladwin, M.T. (2010). Sickle-cell disease. *Lancet*,376(9757) ,2018-2031

Rusanova, I., Cossio, G., Moreno, B., Javier, Perea, F., De Borace, R.G., Perea, M., Escames, G., Acuña-Castroviejo, D. (2011). β-globin gene cluster haplotypes in sickle cell patients from Panamá. *Am J Hum Biol*, 23(3),377-80.

Shaikho, E.M., Farrell, J.J., Alsultan, A., Qutub, H., Al-Ali, A.K., Figueiredo, M.S., Chui, D.H.K., Farrer, L.A., Murphy, G.J., Mostoslavsky, G., Sebastiani, P. & Steinberg, M.H. (2017). A phased SNP-based classification of sickle cell anemia HBB haplotypes. *BMC Genomics*,18(1),608.

Srinivas R, Dunda O, Krishnamoorthy R, Fabry ME, Georges A, Labie D, Nagel RL.(1988). Atypical haplotypes linked to the beta S gene in Africa are likely to be the product of recombination. *Am J Hematol* ,29(1):60-2.

Steinberg, M.H. (2009). Genetic etiologies for phenotypic diversity in sickle cell anemia. *Scientific World Journal*,9 ,46-67.

Tshilolo, L, Aissi, L.M., Lukusa, D., Kinsiama, C., Wembonyama, S., Gulbis, B. & Vertongen, F. (2009). Neonatal screening for sickle cell anaemia in the Democratic Republic of the Congo: experience from a pioneer project on 31 204 newborns. *J Clin Pathol*, 62(1),35-8.

Tshilolo, L., Tomlinson, G., Williams, T.N., Santos, B., Olupot-Olupot, P., Lane, A., Aygun, B., Stuber, S.E., Latham, T.S., McGann, P.T., Ware, R.E.; REACH Investigators. (2019). Hydroxyurea for Children with Sickle Cell Anemia in Sub-Saharan Africa. *N Engl J Med*,380(2),121-131.

Zago, M.A., Silva, W.A. Jr, Dalle, B., Gualandro, S., Hutz, M.H., Lapoumeroulie, C., Tavella, M.H., Araujo, A.G., Krieger, J.E., Elion, J.& Krishnamoorthy, R. (2000). Atypical beta(s) haplotypes are generated by diverse genetic mechanisms. *Am J Hematol*,63(2),79-84.

RESEARCH ARTICLE 3

3.2.2 Genetic variants associated with fetal hemoglobin and hematological profile in a cohort of DR Congolese SCA patients

TO BE SUBMITTED

Mamy Ngole[1,2,3], Paul Lumbala[1,2,4], Gloire Mbayabo[1,2,4], Valerie Race[1], Nono Mvuama[5], Isabelle Cleynen[1], Stephanie Deman[6], Erika Souche[6], Prosper Tshilobo Lukusa[2,4], Tite Mikobi [2],Chris Van Geet [7], Koenraad Devriendt[1], Gert Matthijs[1], Aimé Lumaka [2,4,8]

[1]Center for Human Genetics, Faculty of Medicine, KU Leuven, Leuven, Belgium

[2]Center for Human Genetics, Faculty of Medicine, University of Kinshasa, Kinshasa, Democratic Republic of Congo

[3]Department of Medical Biology, Faculty of Medicine, University of Kinshasa, Kinshasa, Democratic Republic of Congo

[4]Department of Pediatrics, Faculty of Medicine, University of Kinshasa, Kinshasa, Democratic Republic of Congo

[5]Kinshasa School of Public Health, Faculty of Medicine, University of Kinshasa, Kinshasa, Democratic Republic of Congo

[6]Genomics Core KULeuven - UZLeuven, Leuven, Belgium

[7]Department of Cardiovascular Sciences and Pediatrics (Hemato-oncology), KU Leuven and UZ Leuven, Belgium

[8]GIGA-R, Laboratoire de Génétique Humaine, University of Liège, Liège, Belgium

Abstract

Background: Fetal hemoglobin (HbF) level, the most recognized modulator of sickle cell anemia (SCA) clinical expression, is mainly regulated by quantitative trait loci, *BCL11A*, *HBS1L-MYB,* and *HBB* loci. However, SNPs associations with HbF levels in those regulatory regions are not consistent across different populations.

This study assessed a set of SNPs in a cohort of Congolese SCA patients, for their potential association with HbF levels and the hematological parameters involved in the disease severity.

Method: SCA patients were recruited from St Luc Hospital in Kisantu and the "Centre de Médecine Mixte et d'Anémie SS" in Kinshasa in a cross-sectional study. Blood count data were obtained on an automated counter. Restriction fragment length polymorphism was used to confirm homozygous SCA status. Molecular Inversion Probes (MIPS) sequencing was used for SNPs genotyping. Data analysis was performed by regression analysis using the PLINK software with age and sex as covariates.

Results: The most significant association with HbF level was found for rs1427407 (G>T) at *BCL11A* locus , with the homozygous TT genotype having the highest level of HbF (13.8%). Other SNPs within the *BCL11A* locus (rs766432, rs11886868, rs4671393,rs7606173) and *HBS1L-MYB* intergenic region (rs28384513) were also significantly associated with HbF levels. No simultaneous association with both HbF and a particular blood cell count was observed for any of these SNPs. Only the OR51B5/6 locus with rs5006884 and rs5024042 were associated with platelet count.

Conclusion: In our cohort of Congolese SCA patients, SNPs at *BCL11A* and HBS1L-MYB loci show significant influence on the HbF level but no correlation with blood cell count.

Keywords: sickle cell anemia; DR Congo; single nucleotide polymorphisms; fetal hemoglobin level

Introduction

Sickle cell anemia (SCA) is recognized mainly as a single mutation disease with marked interindividual variability with regards to the clinical manifestations and severity, as well as the response to the HU treatment. The basis of this variability is not fully elucidated yet (Driss et al. 2009). So far, the fetal hemoglobin (HbF) level is largely considered the most powerful modulator of the clinical expression of the disease (Akinsheye et al. 2011). Multiple genome-wide studies have established a significant association between high levels of HbF on one hand, and milder manifestations of the disease and a survival advantage on the other hand (Akinsheye et al. 2011; Maitra et al. 2017; Platt et al. 1994; Steinberg & Sebastiani 2012). This beneficial effect of HbF consists in inhibiting the polymerization of deoxyHbS (Charache 1990), which is a major element in the pathophysiology of SCA (Bunn 1997; Steinberg 2020). In addition, SCA patients with high levels of HbF also present a decreased mortality related to the disease (Dadheech et al. 2014; Thein 2008).

The level of HbF is in itself a heritable trait whose expression is mainly regulated by 3 quantitative trait loci (QTLs) identified by Genome-Wide Association Studies (GWAS) (Garner et al. 2000; Gibney et al. 2008). Those QTLs are localized in the *BCL11A* (B-cell lymphoma/leukemia 11A) gene on chromosome 2p16.2, the *HBS1L-MYB* intergenic region [a126-kb intergenic region between *HBS1L* (Hsp70 subfamily B suppressor 1-like gene) and the *MYB* (MYB proto-oncogene) gene on chromosome 6q23.3], and the β-like-globin gene cluster (*HBB* cluster) on chromosome 11p15.5, especially the polymorphic loci Xmn1-*HBG2* and the OR51B5 (Lettre et al. 2008; Nuinoon et al. 2010; Solovieff et al. 2010; Thein 2008, 2013; Uda et al. 2008). Together those 3 QTLs explain 20 to 50% of the interindividual variation of HbF in SCA patients (Lionel et al. 2018; Thein 2008). Other studies have identified additional regulators of HbF expression such as the *OR51B6* (olfactory receptor) gene on chromosome 11p15.4 (Solovieff et al. 2010), *SIN3A* gene on chromosome 15q24.2 (Gravia et al. 2016), and GLP2R gene (glucagon-like peptide-2 receptor) (Bhatnagar et al. 2011). In SCA GWAS, the selection of SNPs for HbF levels association evaluation is function of the MAF in the studied population. Hence, the association could be variable depending on the studied populations (Delgadinho et al. 2021; Fong et al. 2015; Mikobi et al. 2018; Sales et al. 2020; Wonkam et al. 2014).

In DRC, one of the most SCA-affected countries worldwide, only the preliminary study by Mikobi et al. has evaluated the association of 8 selected SNPs with HbF levels in a cohort of

DR Congolese SCA patients thus far (Mikobi et al. 2018). In the present study, we assessed a larger set of selected SNPs on a larger cohort of Congolese SCA patients, for their potential association with HbF levels and the hematological parameters involved in the disease severity.

Methods

Patients' selection, sample treatment, selection of SNPs set, and SNP genotyping technique are presented in the general methodology.

Data analysis

Quantitative data of WBC, platelet, and reticulocyte counts were normalized by $\log_{10}$-transformation prior to the analysis. All genotypes underwent a Hardy-Weinberg law check using the chi-squared test. A regression analysis was performed for SNP genotype with age and sex as covariates using the software PLINK v1.9 (https://www.cog-genomics.org/plink/1.9 or https://zzz.bwh.harvard.edu/plink). For statistical analysis, a p-value ≤ 0.05 was considered significant.

Results

Description of the cohort and tested SNPs

As per the QC presented on page 54, a total of 215 patients (85 males and 130 females) aged between 5 and 40 years (median age = 20 years; range: 12.3 - 27), were retained for the study. Nineteen SNPs passed the QC for this study. None of these showed significant deviation from the HW equilibrium.

Genotypes and alleles frequencies

The candidate SNPs and the allele frequencies are presented in table 1. The lowest MAF was found for the SNPs localized in the *HBS1L-MYB* intergenic region and those in the HBB locus.

Table 1. Liste of SNPs

Gene	Candidate SNP	Chromosomal location hg19 GRCh37	Allele change	MAF
BCL11A	rs10189857	2:60713235	A>G	G=0.26
BCL11A	rs1427407	2:60718043	G>T	T=0.27
BCL11A	rs7599488	2:60718347	C>T	T=0.25
BCL11A	rs766432	2:60719970	A>C	C=0.32
BCL11A	rs11886868	2:60720246	T>C	C=0.31
BCL11A	rs4671393	2:60720951	G>A	A=0.32
BCL11A	rs7606173	2:60725451	G>C	C=0.45
HBS1L-MYB	rs17599586	6:131904719	C>T	T=0.12
HBS1L-MYB	rs28384513	6:135376209	T>G	G=0.13
HBS1L-MYB	rs7776054	6:135418916	A>G	G=0,16
HBS1L-MYB	rs9399137	6:135419018	T>C	C=0.03
HBS1L-MYB	rs4895440	6:135426558	T>A	A=0.41
HBS1L-MYB	rs4895441	6:135426573	A>G	G=0.03
HBS1L-MYB	rs9389269	6:135427159	T>C	C=0.01
HBS1L-MYB	rs9402686	6:135427817	G>A	A=0.02
HBG	rs7482144	11:5276169	G>A	A=0.00
HBE	rs7130110	11:5296104	G>C	C=0.01
OR51B5/6	rs5006884	11:5373251	C>T	T=0.06
OR51B5/6	rs5024042	11:5373562	C>A	A=0.06

Table 2. Association of selected variants with HbF and hematological parameters

Genes	SNP	HbF β(p)	Hb β (p)	WBC β(p)	Platelets β (p)	Reticulocyte β(p)
BCL11A	rs10189857	-0.17(0.75)	0.009 (0.960)	0.007(0.70)	0.003(0.873)	0.001(0.95)
BCL11A	rs1427407	2.63(**8.25e-06)**	0.38 (0.068)	-0.022(0.29)	-0.01(0.571)	0.03(0.22)
BCL11A	rs7599488	0.06(0.90)	0.13(0.502)	0.007(0.72)	-0.003(0.881)	-0.007(0.7)
BCL11A	rs766432	2.06(**0.00016**)	0.35 (0.072)	-0.027(0.17)	-0.02(0.249)	0.01(0.52)
BCL11A	rs11886868	2.34(**2.01.e-05**)	0.31 (0.10)	-0.028 (0.14)	-0.015(0.463)	0.01 (0.44)
BCL11A	rs4671393	2.11(**0.0001355**)	0.33 (0.087)	-0.029(0.14)	-0.02(0.214)	0.01(0.50)
BCL11A	rs7606173	-2.13(**2.31e-05**)	-0.381(0.034)	0.021(0.23)	0.008(0.68)	-0.008(0.7)
ARG1	rs17599586	-0.45 (0.56)	0.27 (0.331)	-0.006 (0.81)	0.03(0.283)	-0.06 (0.06)
HBS1L-MYB	rs28384513	-2.13(**2.31e-05**)	-0.09(0.69)	0.012(0.60)	0.002(0.935)	0.01(0.57)
HBS1L-MYB	rs7776054	0.41(0.57)	0.09(0.71)	-0.003(0.88)	-0.02(0.467)	-0.01(0.63)
HBS1L-MYB	rs9399137	3.76(0.011)	0.29(0.57)	0.010(0.84)	-0.11(0.051)	0.02(0.68)
HBS1L-MYB	rs4895440	0.14(0.77)	0.13 (0.46)	-0.004(0.82)	0.007(0.704)	-0.01(0.67)
HBS1L-MYB	rs4895441	1.01(0.51)	0.33(0.53)	-0.048(0.38)	8.014e-05 (0.998)	-0.06(0.38)
HBS1L-MYB	rs9389269	2.31(0.21)	0.87(0.18)	-0.040(0.55)	0.02(0.761)	-0.11(0.18)
HBS1L-MYB	rs9402686	2.87(0.14)	0.822 (0.23)	-0.036(0.60)	0.02(0.780)	0.08(0.22)
HBG	rs7482144	0.39(0.89)	-1.00(0.33)	0.02(0.37)	0.08(0.540)	0.04(0.59)
HBE	rs7130110	0.42(0.38)	0.21(0.23)	0.02(0.21)	0.01(0.462)	0.01(0.84)
OR51B5	rs5006884	-0.83(0.36)	0.26 (0.30)	-0.02(0.39)	-0.10 (**0.0030**)	0.02(0.55)
OR51B5	rs5024042	-0.83(0.36)	-0.33 (0.30)	-0.02(0.39)	-0.10 (0.**0030**)	0.02(0.55)

Regression analysis was performed with age and sex as covariates

Hb: hemoglobin; WBC: white blood cells β: correlation coefficient
underlined values: no correlation after Bonferroni correction

Significant association with HbF levels was observed for rs1427407 , rs766432, rs11886868 , rs4671393 , and rs7606173 in *BCL11A* locus; and rs28384513 in HBS1L-MYB intergenic locus (Table 2). Two SNPs (rs7606173 and rs28384513) were in LD with variant rs1427407, with r^2 >0.50 and p<0.05. The most significant association with HbF level was found for rs1427407 (G>T) at *BCL11A* locus .

The average HbF level increased from 5.9 % for the GG genotype to 13.8% for the TT genotype, while the average was 7.05% for TG genotype (Figure 1).

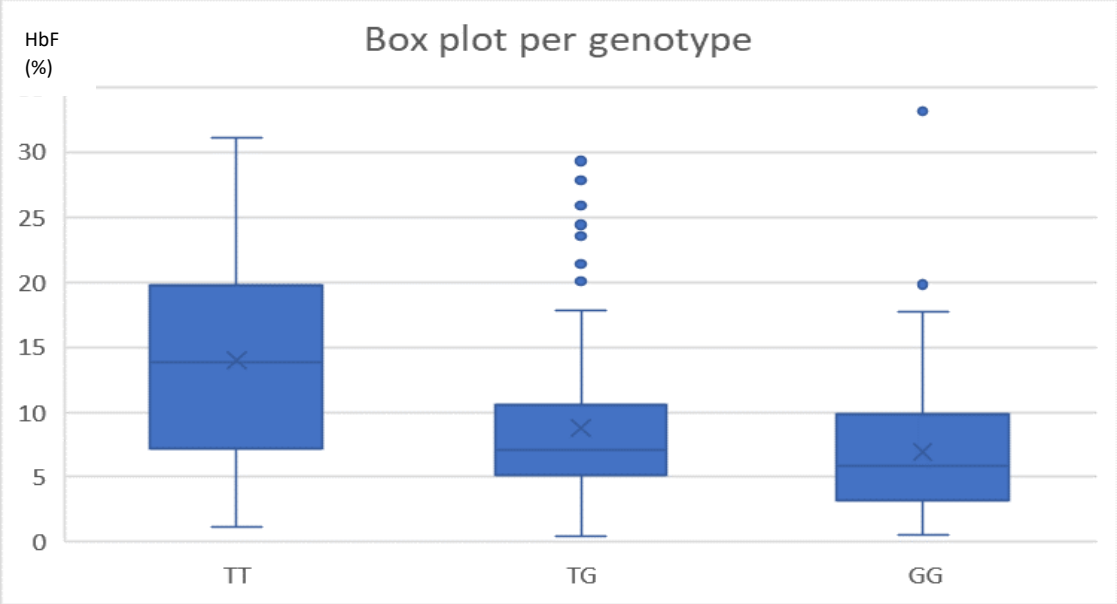

Figure 1. Distribution of HbF levels across the 3 genotypes of rs1427407 in the BCL11A gene

Lines at each box represent the lower, the median, and the upper quartile. Points outside the whiskers indicated as circles are outliers.

None of these SNPs show a simultaneous association with HbF and hematologic variables. Only the OR51B5/6 locus with rs5006884 and rs5024042 was associated with platelet count.

Discussion

The search for potential therapeutic targets for sickle cell disease has motivated studies on variants influencing the HbF level, the major determinant of the severity of the disease thus far. In this study involving 215 Congolese SCA patients, 19 SNPs passed the QC and were evaluated for previously reported association with HbF and hematological parameters. Six SNPs in the BCL11A locus and HBS1L-MYB intergenic region were significantly associated with HbF levels.

BCL11A locus

The BCL11A locus appears to be the most significant influent of HbF levels, as reported in our study and others describing African patients (Makani et al. 2011; Mikobi et al. 2018b; Wonkam et al. 2014), as well as in African descendants in America (Lettre et al. 2008). The BCL11A gene is considered as the major regulator of the Hb switch.

However, rare microdeletions affecting the BCL11A locus result in both increased HbF levels and a neurodevelopmental disorder as previously reported (Funnell et al. 2015; Lumaka et al. 2018; Peter et al. 2014). Of note, the BCL11A SNPs presenting a significant association in our

study are rather common among African individuals. The MAF of HbF increasing allele T of rs1427407 was 0.27 in our study, 0.23 in Nigeria (Adeyemo et al. 2018), 0.20 in Tanzania (Mtatiro et al. 2014), and 0.24 in Cameroon (Pule et al. 2015). The MAF of rs7606173 was 0.45 in our study versus 0.42 reported in Cameroon (Pule et al. 2015) and 0.30 in a previous study on Congolese SCA patients (Mikobi et al. 2018a). These SNPs have not been associated with developmental phenotypes, suggesting that those SNPs carry only positive effects, making them good therapeutic targets.

Like in our study, a significant association of rs11886868 and rs4671393 with HbF levels was previously reported in other studies involving African-Americans (Lettre et al. 2008), Tunisian (Chaouch et al. 2016) and DR Congolese SCA patients (Mikobi et al. 2018a).

HBS1L-MYB intergenic region

In this study, only rs28384513 was significantly associated with HbF level. This association was also reported in Brazilian, Tanzanian SCA populations, and British African descent (Creary et al. 2009; Lettre et al. 2008; Makani et al. 2011). With regards to the MAF, rs9389269 has a MAF of 0.01, which is close to the 0.03 reported in the Tanzanian cohort but higher than the 0.18 reported in the Cameroonian cohort (Makani et al. 2011; Wonkam et al. 2014). Our study reports a low MAF (0.03) for rs9399137, similar to other African cohorts: Cameroonian (0.04%), and Tanzanian (0.01%). Compared to African populations, this SNP has a strong association with HbF levels in European populations as reported by Thein et al (Thein et al. 2007). It is not clear whether these MAF differences influence the association with HbF. We believe that these differences should be explored in relation to possible environmental effects.

Xmn1-HBG2 and the OR51B5/6 locus

The rs7482144 was homozygous for the wild type allele (GG) in all patients. Thus, a correlation could not be evaluated. The T allele of XmnI polymorphism is linked to the Arab-Indian and Senegal haplotypes and explains the observed high HbF levels in those haplotypes (Gilman & Huisman 1984). For instance, the rs7481244 has a strong association with HbF in the Iraqi Kurds SCA patients as well as in Senegalese SCA patients where respectively the Arab Indian and the Senegal haplotypes are predominant (Al-Allawi et al. 2019; Bhanushali et al. 2015). Those two haplotypes are almost scarce among Congolese SCA where the Central African Republic (CAR) haplotype is predominant. This could be the explanation for the observed genotype of the rs7482144.

In general, SNPs presenting significant association with HbF levels were expected to also impact hematological parameters. However, that was not observed in our study. Surprisingly, independently of the effect on HbF levels, two variants at OR51B5/6 locus variants were associated significantly with low platelet counts in our cohort (β= -0.1036 p=0.0030). The reasons for this disconnect remain to be elucidated.

Our approach to selected SNPs and the small size of our cohort might have prevented some correlations from being identified. A GWAS study including Congolese SCA patients might allow the discovery of others loci as well as the best tagging of SNPs of interest.

Conclusion

This study replicated the previously reported associations between SNPs in the BCL11A and HBS1L-MYB loci and the HbF levels in a moderate-size cohort of Congolese SCA patients. We did not observe a combined association between our tested SNPs and both the HbF and the hematological parameters. Those results confirm the biological variability that characterizes SCA is the same genomic loci across the Sub-Saharan regions.

References

Adeyemo TA, Ojewunmi OO, Oyetunji IA, Rooks H, Rees DC, Akinsulie AO, Akanmu AS, Thein SL & Menzel S.(2018). A survey of genetic fetal-haemoglobin modifiers in Nigerian patients with sickle cell anaemia. *PLoS One*. 13(6):e0197927

Akinsheye I, Alsultan A, Solovieff N, Ngo D, Baldwin CT, Sebastiani P, Chui DH & Steinberg MH. (2011). Fetal hemoglobin in sickle cell anemia. *Blood*. 118(1):19–27

Al-Allawi N, Qadir SMA, Puehringer H, Chui DHK, Farrell JJ & Oberkanins C. 2019. The association of HBG2 , BCL11A, and HMIP polymorphisms with fetal hemoglobin and clinical phenotype in Iraqi Kurds with sickle cell disease. *Int J Lab Hematol*. 41(1):87–93

Bhanushali AA, Patra PK, Pradhan S, Khanka SS, Singh S & Das BR. (2015). Genetics of fetal hemoglobin in tribal Indian patients with sickle cell anemia. *Translational Research*. 165(6):696–703

Bhatnagar P, Purvis S, Barron-Casella E, DeBaun MR, Casella JF, Arking DE & Keefer JR. (2011). Genome-wide association study identifies genetic variants influencing F-cell levels in sickle-cell patients. *J Hum Genet*. 56(4):316–23

Bunn HF. (1997). Pathogenesis and Treatment of Sickle Cell Disease. *NEJM*. 337(11):762–69

Charache S. 1990. Fetal hemoglobin, sickling, and sickle cell disease. *Adv Pediatr*. 37:1–31

Creary LE, Ulug P, Menzel S, McKenzie CA, Hanchard NA , Taylor V, Farrall M, Forrester TE & Thein SL. (2009). Genetic Variation on Chromosome 6 Influences F Cell Levels in Healthy Individuals of African Descent and HbF Levels in Sickle Cell Patients. *PLoS One*. 4(1):e4218

Dadheech S, Jain S, Madhulatha D, Sharma V, Joseph J, Jyothy A & Munshi A. (2014). Association of Xmn1 −158 γG variant with severity and HbF levels in β-thalassemia major and sickle cell anaemia. *Mol Biol Rep*. 41(5):3331–37

Delgadinho M, Ginete C, Santos B, Miranda A & Brito M. (2021). Genotypic Diversity among Angolan Children with Sickle Cell Anemia. *Int J Environ Res Public Health*. 18(10):5417

Driss A, Asare K, Hibbert J, Gee B, Adamkiewicz T & Stiles J. (2009). Sickle Cell Disease in the Post Genomic Era: A Monogenic Disease with a Polygenic Phenotype. *Genomics Insights*. 2:23–48

Fong C, Menzel S, Lizarralde MA & Barreto G. (2015). Genetic variants associated with fetal hemoglobin levels show the diverse ethnic origin in Colombian patients with sickle cell anemia. *Biomédica*. 35(3)

Funnell APW, Prontera P, Ottaviani V, Piccione M, Giambona A, Maggio A, Ciaffoni F, Stehling-Sun S, Marra M, Masiello F, Varricchio L, Stamatoyannopoulos JA, Migliaccio AR, Papayannopoulou T. (2015). 2p15-p16.1 microdeletions encompassing and proximal to BCL11A are associated with elevated HbF in addition to neurologic impairment. *Blood*. 126(1):89–93

Garner C, Tatu T, Reittie JE, Littlewood T, Darley J, Cervino S, Farrall M, Kelly P, Spector TD, Thein SL. (2000). Genetic influences on F cells and other hematologic variables: a twin heritability study. *Blood.* 95(1):342–46

Gibney GT, Panhuysen CIM, So JCC, Ma ESK, Ha SY, Li CK, Lee AC, Li CK, Yuen HL, Lau YL, Johnson DM, Farrell JJ, Bisbee AB, Farrer LA, Steinberg MH, Chan LC & Chui DH. (2008). Variation and heritability of Hb F and F-cells among β-thalassemia heterozygotes in Hong Kong. Am J Hematol. 83(6):458–64

Gilman J & Huisman T. (1984). Two independent genetic factors in the beta-globin gene cluster are associated with high G gamma-levels in the HbF of SS patients. *Blood*. 64(2):452–57

Gravia A, Chondrou V, Kolliopoulou A, Kourakli A, John A, Symeonidis A, Ali BR, Sgourou A, Papachatzopoulou A, Katsila T & Patrinos GP. (2016). Correlation of SIN3A genomic variants with β-hemoglobinopathies disease severity and hydroxyurea treatment efficacy. *Pharmacogenomics*. 17(16):1785–93

Lettre G, Sankaran VG, Bezerra MAC, Araújo AS, Uda M, Sanna S, Cao A, Schlessinger D, Costa FF, Hirschhorn JN & Orkin SH . (2008). DNA polymorphisms at the BCL11A , HBS1L-MYB , and β- globin loci associate with fetal hemoglobin levels and pain crises in sickle cell disease. *Proceedings of the National Academy of Sciences*. 105(33):11869–74

Lionel AC, Costain G, Monfared N, Walker S, Reuter MS, Hosseini SM, Thiruvahindrapuram B, Merico D, Jobling R, Nalpathamkalam T, Pellecchia G, Sung WWL, Wang Z, Bikangaga P, Boelman C, Carter MT, Cordeiro D, Cytrynbaum C, Dell SD, Dhir P, Dowling JJ, Heon E, Hewson S, Hiraki L, Inbar-Feigenberg M, Klatt R, Kronick J, …, Shuman C, Carew C, Szego MJ, Hayeems RZ, Basran R, Stavropoulos DJ, Ray PN, Bowdin S, Meyn MS, Cohn RD, Scherer SW & Marshall CR. (2018). Improved diagnostic yield compared with targeted gene sequencing panels suggests a role for whole-genome sequencing as a first-tier genetic test. *Genetics in Medicine*. 20(4):435–43

Lumaka A, Race V, Peeters H, Corveleyn A, Coban-Akdemir Z, Jhangiani SN, Song X, Mubungu G, Posey J, Lupski JR, Vermeesch JR, Lukusa P & Devriendt K. (2018). A comprehensive clinical and genetic study in 127 patients with ID in Kinshasa, DR Congo. *Am J Med Genet* A. 176(9):1897–1909

Maitra P, Caughey M, Robinson L, Desai PC, Jones S, Nouraie M, Gladwin MT, Hinderliter A, Cai J & Ataga KI. (2017). Risk factors for mortality in adult patients with sickle cell disease: a meta-analysis of studies in North America and Europe. *Haematologica.* 102(4):626–36

Makani J, Menzel S, Nkya S, Cox SE, Drasar E, Soka D, Komba AN, Mgaya J, Rooks H, Vasavda N, Fegan G, Newton CR, Farrall M & Thein SL. (2011). Genetics of fetal hemoglobin in Tanzanian and British patients with sickle cell anemia. *Blood.* 117(4):1390–92

Mikobi TM, Lukusa PT, Aloni MN, Lumaka A, Akilimali PZ, Devriendt K, Matthijs G, Mbuyi Muamba JM & Race V. (2018a). Association between sickle cell anemia and alpha thalassemia reveals a high prevalence of the α 3.7 triplication in congolese patients than in worldwide series. *J Clin Lab Anal*. 32(1):e22186

Mikobi TM, Tshilobo Lukusa P, Aloni MN, Lumaka AZ, Kaba DK, Devriendt K, Matthijs G, Mbuyi Muamba JM, Race V. (2018b). Protective BCL11A and HBS1L-MYB polymorphisms in a cohort of 102 Congolese patients suffering from sickle cell anemia. *J Clin Lab Anal*. 32(1):e22207

Mtatiro SN, Singh T, Rooks H, Mgaya J, Mariki H, Soka D, Mmbando B, Msaki E, Kolder I, Thein SL, Menzel S, Cox SE, Makani J & Barrett JC. (2014). Genome Wide Association Study of Fetal Hemoglobin in Sickle Cell Anemia in Tanzania. *PLoS One*. 9(11):e111464

Nuinoon M, Makarasara W, Mushiroda T, Setianingsih I, Wahidiyat PA, Sripichai O, Kumasaka N, Takahashi A, Svasti S, Munkongdee T, Mahasirimongkol S, Peerapittayamongkol C, Viprakasit V, Kamatani N, Winichagoon P, Kubo M, Nakamura Y & Fucharoen S . (2010). A genome-wide association identified the common genetic variants influence disease severity in β0-thalassemia/hemoglobin E. *Hum Genet*. 127(3):303–14

Peter B, Matsushita M, Oda K & Raskind W. (2014). De novo microdeletion of BCL11A is associated with severe speech sound disorder. *Am J Med Genet* A. 164(8):2091–96

Platt OS, Brambilla DJ, Rosse WF, Milner PF, Castro O, Steinberg MH & Klug PP .(1994). Mortality In Sickle Cell Disease -- Life Expectancy and Risk Factors for Early Death. *NEJM*. 330(23):1639–44

Pule GD, Ngo Bitoungui VJ, Chetcha Chemegni B, Kengne AP, Antonarakis S, Wonkam A. (2015). Association between Variants at BCL11A Erythroid-Specific Enhancer and Fetal Hemoglobin Levels among Sickle Cell Disease Patients in Cameroon: Implications for Future Therapeutic Interventions. *OMICS.* 19(10):627–31

Sales RR, Belisário AR, Faria G, Mendes F, Luizon MR & Viana MB. (2020). Functional polymorphisms of BCL11A and HBS1L-MYB genes affect both fetal hemoglobin level and clinical outcomes in a cohort of children with sickle cell anemia. *Ann Hematol*. 99(7):1453–63

Solovieff N, Milton JN, Hartley SW, Sherva R, Sebastiani P, Dworkis DA, Klings ES, Farrer LA, Garrett ME, Ashley-Koch A, Telen MJ, Fucharoen S, Ha SY, Li CK, Chui DH, Baldwin CT & Steinberg MH. (2010). Fetal hemoglobin in sickle cell anemia: genome-wide association studies suggest a regulatory region in the 5′ olfactory receptor gene cluster. *Blood*. 115(9):1815–22

Steinberg MH. (2020). Fetal hemoglobin in sickle cell anemia. *Blood*. 136(21):2392–2400

Steinberg MH & Sebastiani P. (2012). Genetic modifiers of sickle cell disease. *Am J Hematol*. 87(8):795–803

Thein SL. (2008). Genetic modifiers of the β-haemoglobinopathies. *Br J Haematol*. 141(3):357–66

Thein SL. (2013). Genetic association studies in β-hemoglobinopathies. *Hematology*. 2013(1):354–61

Thein SL, Menzel S, Peng X, Best S, Jiang J, Close J, Silver N, Gerovasilli A, Ping C, Yamaguchi M, Wahlberg K, Ulug P, Spector TD, Garner C, Matsuda F, Farrall M, Lathrop M. (2007). Intergenic variants of HBS1L-MYB are responsible for a major quantitative trait locus on chromosome 6q23 influencing fetal hemoglobin levels in adults. *Proceedings of the National Academy of Sciences*. 104(27):11346–51

Uda M, Galanello R, Sanna S, Lettre G, Sankaran VG, Chen W, Usala G, Busonero F, Maschio A, Albai G, Piras MG, Sestu N, Lai S, Dei M, Mulas A, Crisponi L, Naitza S, Asunis I, Deiana M, Nagaraja R, Perseu L, Satta S, Cipollina MD, Sollaino C, Moi P, Hirschhorn JN, Orkin SH, Abecasis GR, Schlessinger D & Cao A. (2008). Genome-wide association study shows BCL11A associated with persistent fetal hemoglobin and amelioration of the phenotype of β-thalassemia. *Proceedings of the National Academy of Sciences*. 105(5):1620–25

Wonkam A, Ngo Bitoungui VJ, Vorster AA, Ramesar R, Cooper RS, Tayo B, Lettre G, Ngogang J. (2014). Association of Variants at BCL11A and HBS1L-MYB with Hemoglobin F and Hospitalization Rates among Sickle Cell Patients in Cameroon. *PLoS One*. 9(3):e92506

RESEARCH ARTICLE 4

3.2.3 Genetic markers associated with hydroxyurea response in Democratic Republic Congolese sickle cell anemia patients

TO BE SUBMITTED

Mamy Ngole[1,2,3], Paul Lumbala[1,2,4], Gloire Mbayabo[1,2,4], Valerie Race[1], Nono Mvuama[5], Isabelle Cleynen[1], Stephanie Deman[6], Erika Souche[6], Prosper Tshilobo Lukusa[2,4], Tite Mikobi[2],Chris Van Geet[7], Koenraad Devriendt[1], Gert Matthijs[1], Aimé Lumaka[2,4,8]

[1]Center for Human Genetics, Faculty of Medicine, KU Leuven, Leuven, Belgium

[2]Center for Human Genetics, Faculty of Medicine, University of Kinshasa, Kinshasa, Democratic Republic of Congo

[3]Department of Medical Biology, Faculty of Medicine, University of Kinshasa, Kinshasa, Democratic Republic of Congo

[4]Department of Pediatrics, Faculty of Medicine, University of Kinshasa, Kinshasa, Democratic Republic of Congo

[5]Kinshasa School of Public Health, Faculty of Medicine, University of Kinshasa, Kinshasa, Democratic Republic of Congo

[6]Genomics Core KULeuven - UZLeuven, Leuven, Belgium

[7]Department of Cardiovascular Sciences and Pediatrics (Hemato-oncology), KU Leuven and UZ Leuven, Belgium

[8]GIGA-R, Laboratoire de Génétique Humaine, University of Liège, Liège, Belgium

Abstract

Background: The beneficial effect of Hydroxyurea (HU) treatment has been demonstrated in sickle cell anemia (SCA) patients as an increase in fetal hemoglobin (HbF) level. However, the response is variable among patients. Genetic markers play a role in this variability. So, the characterization of genetic profiles of SCA patients appears useful for a successful treatment.

This study aimed at evaluating the contribution of α-thalassemia coinheritance, the βs haplotypes, and 63 candidate SNPs in the HbF quantitative trait loci (QTLs) and other loci in HU response in a cohort of 58 Congolese patients included in a HU clinical trial.

Method: The α-thalassemia status was determined using the MLPA method, whereas SNPs genotyping was performed by the MIPs method. The HU response was defined as the increment of the HbF level from the baseline level.

Results: Results demonstrated that α-thalassemia status as well as the βs haplotypes did not significantly influence the positive response to HU treatment. Conversely, the SNPs in five genes contributed significantly to the HbF increase: rs1480642 and rs 487278 in PDE7B , rs 9483947 in MAP3K5 , rs 765587 in TOX , rs 9319428 in FLT1 and rs 28384513 in HBS1L-MYB intergenic space.

Conclusion: Those markers may serve as potential predictors of HU response in our setting.

Keywords: Hydroxyurea response, alpha thalassemia, betaS haplotypes, single nucleotide polymorphisms, DRC

Introduction

Hydroxyurea (HU) is the first drug approved by the Food and Drug Administration (FDA) for Sickle Cell Anemia (SCA) treatment (Yawn et al. 2014). Treated patients benefit from both clinical and biological improvements, leading to a better quality of life (Charache et al. 1996; McGann & Ware 2015; Ofakunrin et al. 2020; Tshilolo et al. 2019; Youssry et al. 2017). The drug was originally developed to treat leukemia. The increase of F cells (RBC with high content of HbF) in patients treated with HU led to its repurposing to treat SCA despite its cytotoxicity (Dover et al. 1985). Several mechanisms are involved in the beneficial effects of HU in SCA, including HbF induction, reducing leucocytes and reticulocytes count, reducing hemolysis, improvement of vascular rheology (Ware 2010). However, the increase of HbF levels remains the primary effect of HU, through several pathways (Lebensburger et al. 2010).

The protocol for the administration of HU requires a progressive dose escalation from a minimal dose of 10-15mg/kg/d up to the maximum tolerated dose (MTD), which is around 30-35 mg/kg/d. However, despite good adherence to HU treatment, an inter-individual variability in the response to treatment has been observed, both in terms of tolerated dose and in terms of achieved % HbF at MTD, referred to as HU-induced HbF level (Baliga et al. 2000; Hankins et al. 2014). Based on the level of HU-induced HbF levels, patients are classified into good responders and poor responders (Charache et al. 1992; Steinberg et al. 1997). SCA patients non-responders to HU represent 25 to 30% of treated patients (Charache et al. 1995; McGann & Ware 2011; Steinberg et al. 1997).

Among factors likely to influence HU response among SCA patients, adherence to treatment and environment (socioeconomic status, adherence to treatment, infections) , and the individuals' genetics appear to be the most relevant (Santana et al. 2020; Yahouédéhou et al. 2018). Genetic markers reported to influence HU response include some SNPs, alpha thalassemia, and beta-globin haplotypes (Ma et al. 2007; Okumura et al. 2016). There is a correlation between the baseline of HbF and HU-induced HbF levels. Some of the genetic polymorphisms associated with HU-induced HbF levels are already known to influence basal-HbF levels of HbF (Sales et al. 2022). In addition, other polymorphisms, not concerning the baseline HbF level, have also shown a significant association with HU response [https://www.pharmgkb.org].

Due to this observed variability as well as potentially toxic effects related to HU, it appears important to target patients eligible for HU treatment based on the individual genetic profile of

the patient. This approach is particularly important in Sub-Saharan Africa in general and in the Democratic Republic of Congo, where the majority of SCA patients live.

Before the onset of this study, such an investigation did not exist in DRC. Since HU response and adherence have been evaluated in a large cohort of Congolese patients (Tshilolo et al. 2019). However, that study did not investigate genetic determinants.

This present study aimed to fill the gap of genetic determinants of HU response, including candidate SNPs, alpha thalassemia status, and beta S haplotypes.

Material and methods

Patients selection, sample treatment, selection of SNPs set and SNPs genotyping technique is presented on point 2.1 above.

The HU clinical trial was conducted by my colleagues Paul and Gloire (Paul Lumbala. Diagnosis and Treatment by Hydroxyurea of Adults with Sickle Cell Anemia in the Democratic Republic of Congo**)** (Gloire Mbayabo. Diagnosis and Treatment by Hydroxyurea of Children with Sickle Cell Anemia in the Democratic Republic of Congo)

Particularly in our study, we focused on patients' genetic profiles after being categorized as HU responders and non-responders after HU treatment.

Specifically for this study, we used the Multiplex Ligation-dependent Probe Amplification (MLPA) to define α-thalassemia status among the cohort of patients. MPLA consists of multiplex PCR using probes specific to the genomic sequence to evaluate the relative copy number of each sequence. The SALSA MLPA Probemix kit (MRC-Holland, Netherlands) was used to detect the copy number variations in the α-globin cluster on chromosome 16p13.3 defining α-thalassemia. Each DNA sample concentrated at 100 ng was used to perform the test according to manufacturer instructions. Fragments obtained were detected on ABI Hitachi DNA Analyzer 3730 XL (Hitachi America, USA) and analyzed using the Coffalyser.NET software (MRC-Holland).

Data analysis

Delta HbF level, as the difference between HbF baseline level and HbF HU induced level, was calculated. The t-student test was used to evaluate differences in HbF values between patients according to their alpha-thalassemia status and beta-globin haplotypes. Genetic analysis was

performed by regression analysis, with age and sex as covariates using the PLINk software. A p-value of 0.05 was considered significant.

Results

A cohort of 58 SCA patients included in this study comprised 28 females and 30 males aged between 3 and 33 years.

Influence of alpha-thalassemia status

The α 3.7 kb alpha thalassemia deletion, the most frequent, was co-inherited in 19 SCA patients. But no statistical difference was observed in both groups neither in terms of HbF baseline nor HbF (Table 1).

Table 1. HbF level according to α thalassemia status before and after two years of HU clinical trial

	SCA patients (N=39)	SCA – α2.7 thalassemia patients (N=19)	p values
Baseline HbF (%)	7.4 ± 3.6	7 ± 3.1	0.647
HU induced HbF level (%)	20.8 ± 12.4	20.5 ± 9.2	0.908

Results are presented as mean ±standard deviation

Influence of βS haplotypes

Almost all patients were homozygous CAR haplotypes and this did not allow comparison with other heterozygous CAR haplotypes (Table 2).

Table 2 . HbF values at baseline and after HU treatment according to βs haplotype

	CAR/CAR (N=54)	P	CAR/BEN (N=3)	CAR/SEN (N=1)
HbF baseline level	7.30 ±3.4		-	-
HbF induced HU	20.7 ± 11.4	<0.001	-	-

Values are presented as mean ± standard deviation.

SNPs association with HU response

Six SNPs in five loci (PDE1B, MAP9K5, TOX, FLT1 and HBS1L-MYB) were significantly associated with HbF level induced by HU treatment. In addition, the SNP rs 28384513 in the HBS1L-MYB intergenic region was previously reported to be associated with the baseline HbF level and HU metabolism (Table 3).

Table 3. SNPs significantly associated with HU-induced HbF level

Gene	SNP	Mechanism associated	β	P value
HBS1L-MYB	rs 28384513	HbF baseline + HU-induced HbF level + HU metabolism	-6.94	0.011
PDE7B	rs 1480642	HU-induced HbF level	-7.53	0.012
PDE7B	rs 487278	HU-induced HbF level	-7.39	0.026
MAP3K5	rs 9483947	HU-induced HbF level	-4.90	0.041
TOX	rs 765587	HU-induced HbF level	-5.34	0.026
FLT1	rs 9319428	HU-induced HbF level	-6.4	0.004

Discussion

Hydroxyurea pharmacogenomics studies are useful to anticipate the drug's response. In this study, 58 patients followed up for an HU clinical trial, neither the α-thalassemia coinheritance nor the βs haplotypes influenced HU treatment. Six SNPs were associated significantly with HU response.

α-thalassemia influence

Coinheritance with α-thalassemia deletion concerns around 30% of SCA patients and improves both biological and clinical manifestations of SCA (Darbari et al. 2012; Embury et al. 1982).

Controversial results are reported concerning the influence of α-thalassemia on HU response. The study by Vasavda et al observed a negative effect of α-thalassemia in HU response in terms of the high level of plasma cell-free DNA instead of the expected reduced level induced by HU treatment (Ulug et al. 2008; Vasavda et al. 2012). In contrast, Sheehan et al reported an increase

in HbF levels after HU treatment independently of α-thalassemia status as well as Figueiredo et al (Figueiredo et al. 2017; Sheehan et al. 2013).

Our results sustain the trends by Sheehan and Figueiredo as in our cohort of patients the increase in HbF level was observed regardless of α-thalassemia status. Factors such as age and reaching of MTD are among possible factors that may explain this variability according to Figueiredo et al (Figueiredo et al. 2017).

βs haplotypes effect

The response to HU treatment according to βs haplotypes is variable in the literature. Our results corroborate with those of Vicari et al (Vicari et al. 2005), Silva et al (Silva et al. 2013), and Ware et al (Ware 2010). Both studies conducted in Brazilian SCA patients reported a significant increase in HbF level after HU treatment within the homozygous CAR haplotype, the most frequent haplotype in their cohort (40.9% and 48.2% respectively).

But in another Brazilian SCA patients cohort, the second more frequent haplotype, BEN, was correlated with an increase in HbF levels (Yahouédéhou et al. 2019). The same results were reported by Bernaudin et al in a cohort of 580 French SCA patients, with an important correlation of BEN haplotype with increased HbF level after HU treatment, despite the highest frequency of the CAR haplotype (Bernaudin et al. 2018). For those authors, this difference in HU-induced HbF levels between CAR and BEN haplotypes may be explained by the positive correlation of the BEN haplotype with the favorable allele "T" of SNP rs 1427407 (BCL11A). Such comparison was not possible in our cohort due to the homogeneous distribution of the CAR haplotype.

Similar to our result, the study by Ware et al did not find an association of HU response at the maximum tolerated dose neither with βs haplotypes and α-thalassemia (Ware 2010).

SNPs associated

Among SNPs significantly associated with HU-increased HbF in our study, three of them rs9483947 (MAP3K5), rs765587 (TOX), and rs 9319428 (FLT1) were also significantly associated with HbF as reported by Ma et al (Ma et al. 2007). In both studies, HbF was expressed in the percentage of total hemoglobin evaluated after two years of HU treatment. Those three candidate genes are involved respectively in HbF level regulation as QTL, DNA transcription regulation, and vascular endothelial cell differentiation and proliferation.

Depending on the criteria used, SNPs associated with HU response differ from ours in the present study. Green et al reported a significant association of SNPs within BCL11A and GLP2R genes with maximum HbF during HU treatment (Green et al. 2013). Interestingly, those SNPs are also associated with baseline HbF levels.

Our results were consistent with those by Sheehan et al who also did not find any association of SNPs in BCL11A or HBSIL-MYB with HU response in terms of maximum HbF at MTD or difference in final HbF level with HbF baseline (Sheehan et al. 2014).

All those studies result in different HU response genetic profiles according to the studied populations as well as the response criteria used.

Our study may be limited by the only parameter considered for HU biological response, the increase of HbF level, as the main mechanism of HU. However, since other hematological and biochemical parameters, as well as clinical data, are involved in HU response (Yahouédéhou et al. 2019), further studies are important to evaluate the association with SNPs.

Conclusion

This study highlights the significant contribution of some SNPs to HU treatment response in our cohort of Congolese SCA patients. However, the study is limited by the low number of patients who achieved the clinical trial and consequently evaluated; and also, the consideration of HbF increased level as the only evaluation parameter. Anyway, the six SNPs significantly associated with HU response can be used as a predictive marker of HU treatment efficacity. But those results need to be confirmed in a larger cohort.

References

Baliga BS, Pace BS, Chen H-H, Shah AK & Yang Y-M. (2000). Mechanism for fetal hemoglobin induction by hydroxyurea in sickle cell erythroid progenitors. *Am J Hematol*. 65(3):227–33

Bernaudin F, Arnaud C, Kamdem A, Hau I, Lelong F, Epaud R, Pondarré C & Pissard S. (2018). Biological impact of α genes, β haplotypes, and G6PD activity in sickle cell anemia at baseline and with hydroxyurea. *Blood Adv*. 2(6):626–37

Charache S, Barton FB, Moore RD, Terrin ML, Steinberg MH, Dover GJ, Ballas SK, McMahon RP, Castro O & Orringer EP. (1996). Hydroxyurea and sickle cell anemia. Clinical utility of a myelosuppressive "switching" agent. The Multicenter Study of Hydroxyurea in Sickle Cell Anemia. *Medicine.* 75(6):300–326

Charache S, Dover GJ, Moore RD, Eckert S, Ballas SK, Koshy M, Milner PF, Orringer EP, Phillips G Jr & Platt OS . (1992). Hydroxyurea: effects on hemoglobin F production in patients with sickle cell anemia. *Blood*. 79(10):2555–65

Charache S, Terrin ML, Moore RD, Dover GJ, Barton FB, Eckert SV, McMahon RP & Bonds DR (1995). Effect of Hydroxyurea on the Frequency of Painful Crises in Sickle Cell Anemia. *New Engl J Med*. 332(20):1317–22

Darbari DS, Onyekwere O, Nouraie M, Minniti CP, Luchtman-Jones L, Rana S, Sable C, Ensing G, Dham N, Campbell A, Arteta M, Gladwin MT, Castro O, Taylor JG 6th, Kato GJ & Gordeuk V. (2012). Markers of Severe Vaso-Occlusive Painful Episode Frequency in Children and Adolescents with Sickle Cell Anemia. *J Pediatr*. 160(2):286–90

Dover GJ, Humphries RK, Young N, Ley T, Boyer S, Charache S, Nienhuis A. (1985). Pharmacologic manipulation of fetal hemoglobin synthesis. *Prog Clin Biol Res.* 191:447–54

Embury SH, Dozy AM, Miller J, Davis JR, Kleman KM, Preisler H, Vichinsky E, Lande WN, Lubin BH, Kan YW, Mentzer WC. (1982). Concurrent Sickle-Cell Anemia and α-Thalassemia: : effect on severity of anemia . *N Engl J Med.* 306(5):270–74

Figueiredo L, Morrone K, Wei C, Ireland K, Cohen HW, Driscoll C, Manwani D. (2017). An age dependent response to hydroxyurea in pediatric sickle cell anemia patients with alpha thalassemia trait. *Blood Cells Mol Dis*. 66:19–23

Green NS, Ender KL, Pashankar F, Driscoll C, Giardina PJ, Mullen CA, Clark LN, Manwani D, Crotty J, Kisselev S, Neville KA, Hoppe C & Barral S. (2013). Candidate Sequence Variants and Fetal Hemoglobin in Children with Sickle Cell Disease Treated with Hydroxyurea. *PLoS One*. 8(2):e55709

Hankins JS, Aygun B, Nottage K, Thornburg C, Smeltzer MP, Ware RE & Wang WC. (2014). From Infancy to Adolescence : fifteen years of continuous treatment with hydroxyurea in sickle cell anemia. *Medicine (Baltimore)*. 93(28):e215

Lebensburger JD, Pestina TI, Ware RE, Boyd KL, Persons DA. (2010). Hydroxyurea therapy requires HbF induction for clinical benefit in a sickle cell mouse model. Haematologica. 95(9):1599–1603

Ma Q, Wyszynski DF, Farrell JJ, Kutlar A, Farrer LA, Kutlar A, Farrer LA, Baldwin CT, Steinberg MH. (2007). Fetal hemoglobin in sickle cell anemia: genetic determinants of response to hydroxyurea. *Pharmacogenomics* J. 7(6):386–94

McGann PT & Ware RE. (2011). Hydroxyurea for sickle cell anemia: what have we learned and what questions still remain? *Curr Opin Hematol*. 18(3):158–65

McGann PT, Ware RE. 2015. Hydroxyurea therapy for sickle cell anemia. *Expert Opin Drug Saf*. 14(11):1749–58

Ofakunrin AOD, Oguche S, Adekola K, Okpe ES, Afolaranmi TO, Diaku-Akinwumi IN, Zoakah AI & Sagay AS. (2020). Effectiveness and Safety of Hydroxyurea in the Treatment of Sickle Cell Anaemia Children in Jos, North Central Nigeria. *J Trop Pediatr*. 66(3):290–98

Okumura J v, Silva DGH, Torres LS, Belini-Junior E, Barberino WM, Oliveira RG, Carrocini GC, Gelaleti GB, Lobo CL & Bonini-Domingos CR. (2016). Inheritance of the Bantu/Benin haplotype causes less severe hemolytic and oxidative stress in sickle cell anemia patients treated with hydroxycarbamide. *J Hum Genet*. 61(7):605–11

Sales RR, Nogueira BL, Tosatti JAG, Gomes KB & Luizon MR. (2022). Do Genetic Polymorphisms Affect Fetal Hemoglobin (HbF) Levels in Patients With Sickle Cell Anemia Treated With Hydroxyurea? A Systematic Review and Pathway Analysis. *Front Pharmacol*. 12: 779497

Santana SS, Pitanga TN, de Santana JM, Zanette DL, Vieira J de J, Yahouédéhou SCMA, Adanho CSA, Viana SM, Luz NF, Borges VM, Goncalves MS. 2020. Hydroxyurea Scavenges Free Radicals and Induces the Expression of Antioxidant Genes in Human Cell Cultures Treated With Hemin. *Front Immunol*. 11: 1488

Shah F & Dwivedi M. (2020). Pathophysiology and recent therapeutic insights of sickle cell disease. *Ann Hematol*. 99(5):925–35

Sheehan VA, Crosby JR, Sabo A, Mortier NA, Howard TA, Muzny DM, Dugan-Perez S, Aygun B, Nottage KA, Boerwinkle E, Gibbs RA, Ware RE & Flanagan JM.(2014). Whole Exome Sequencing Identifies Novel Genes for Fetal Hemoglobin Response to Hydroxyurea in Children with Sickle Cell Anemia. *PLoS One*. 9(10):e110740

Sheehan VA, Luo Z, Flanagan JM, Howard TA, Thompson BW, Wang WC, Kutlar A & Ware RE . (2013). Genetic modifiers of sickle cell anemia in the BABY HUG cohort: influence on laboratory and clinical phenotypes. *Am J Hematol*. 88(7):571–76

Silva DGH, Belini Junior E, Carrocini GC de S, Torres L de S, Ricci Júnior O, Lobo CL, Bonini-Domingos CR & de Almeida EA. (2013). Genetic and biochemical markers of hydroxyurea therapeutic response in sickle cell anemia. *BMC Med Genet*. 14(1):108

Steinberg MH, Lu ZH, Barton FB, Terrin ML, Charache S & Dover GJ. (1997). Fetal hemoglobin in sickle cell anemia: determinants of response to hydroxyurea. Multicenter Study of Hydroxyurea. *Blood*. 89(3):1078–88

Tshilolo L, Tomlinson G, Williams TN, Santos B, Olupot-Olupot , PLane A, Aygun B, Stuber SE, Latham TS, McGann PT, Ware RE & REACH Investigators.(2019). Hydroxyurea for Children with Sickle Cell Anemia in Sub-Saharan Africa. *N Engl J Med*. 380(2):121–31

Ulug P, Vasavda N, Kumar R, Keir L, Awogbade M, Cunningham J, Rees DC, Menzel S &Thein SL. (2008). Hydroxyurea therapy lowers circulating DNA levels in sickle cell anemia. *Am J Hematol*. 83(9):714–16

Vasavda N, Woodley C, Allman M, Drašar E, Awogbade M, Howard J & Thein SL. (2012). Effects of co-existing α-thalassaemia in sickle cell disease on hydroxycarbamide therapy and circulating nucleic acids. Br J Haematol. 157(2):249–52

Vicari P, Barretto de Mello A & Figueiredo MS. (2005). Effects of hydroxyurea in a population of Brazilian patients with sickle cell anemia. *Am J Hematol*. 78(3):243–44

Ware RE. (2010). How I use hydroxyurea to treat young patients with sickle cell anemia. *Blood.* 115(26):5300–5311

Yahouédéhou SCMA, Adorno EV, da Guarda CC, Ndidi US, Carvalho SP, Santiago RP, Aleluia MM, de Oliveira RM & Gonçalves MS. (2018). Hydroxyurea in the management of sickle cell disease: pharmacogenomics and enzymatic metabolism. *Pharmacogenomics J*. 18(6):730–39

Yahouédéhou SCMA, da Guarda CC, Figueiredo CVB, Santiago RP, Carvalho SP, Fiuza LM, Ndidi US, Oliveira RM, Carvalho MOS, Nascimento VML, Rocha LC, Lyra IM, Adorno EV & Goncalves MS (2019). Hydroxyurea alters hematological, biochemical and inflammatory biomarkers in Brazilian children with SCA: Investigating associations with βS haplotype and α-thalassemia. *PLoS One.* 14(7):e0218040

Yawn BP, Buchanan GR, Afenyi-Annan AN, Ballas SK, Hassell KL, James AH, Jordan L, Lanzkron SM, Lottenberg R, Savage WJ, Tanabe PJ, Ware RE, Murad MH, Goldsmith JC, Ortiz E, Fulwood R, Horton A & John-Sowah J. (2014). Management of sickle cell disease: summary of the 2014 evidence-based report by expert panel members. *JAMA*. 312(10):1033–48

Youssry I, Abdel-Salam A, Ismail R, Bou-Fakhredin R, Mohamed Samy R, Ezz El-Deen F & Taher AT. 2017. Enhancing Effect of Hydroxyurea on Hb F in Sickle Cell Disease: Ten-Year Egyptian Experience. *Hemoglobin.* 41(4–6):267–73

4. GENERAL DISCUSSION

This study was designed to contribute to reducing the burden of sickle cell anemia (SCA) on the Congolese population. We did not initially conduct a systematic assessment of the needs: it was obvious that a timely and robust diagnostic confirmation of SCA was the angle where a genetic laboratory could bring improvement. The clinical presentation of SCA is widely variable. Several environmental modifiers of the clinical variability have been explored in the Congolese setting (Kadima et al.2015; Aloni et al. 2017 ; Mikobi et al. 2018) . Using genetic tools, it would be possible to advance such a study with the investigation of genetic modifiers in Congolese patients. Therefore, applying genetics to improve the diagnosis and to gain more insights into the genetic modifiers of the disease presentation, the response to treatment, and disease prognostic were the two major challenges that we decided to take on with this study. SCA patients in DR Congo and in similar settings experience additional challenges such as the general limitation of resources, the additional economic burden imposed by the complications of the disease on the patient and relatives, the limited availability of treatment options including blood exchange program or bone marrow transplant, the weakness of the health system, the stigmatization in society and many more (Wembonyama et al. 2007; Luboya et al. 2014; Mulumba & Wilson 2015; Mukinayi et al. 2018). These challenges were not the topic of the present research but were taken into consideration when discussing on the potential of genotyping and haplotype analysis is a low-income setting.

4.1 Implementation of DNA-based test

Reaching the diagnosis is of utmost importance for any genetic disease in general, and for SCA in particular. A diagnosis launches the start of a disease-specific patient's follow-up and informs the implementation of appropriate treatment when needed and available. This follow-up could prevent complications, including death, and would lower the economic burden on the families and on society. Until the time this work was completed, there was no newborn screening program in the country although a pilot study proved the feasibility and utility of NBS for SCA in DR Congo (Tshilolo et al. 2009). Screening is possible using existing laboratory test, but it should be noted that existing diagnostic options, which are based on the detection of sickle hemoglobin, are known for their lower reliability during the neonatal period

due to the interference of fetal hemoglobin. Congolese SCA children enter life undiagnosed, leading to SCA-related mortality of up to 6 to 15% (Mulumba & Wilson, 2015).

In most Congolese SCA patients, SCA workup is only initiated after an acute event or a chronic complication of the disease (Aloni et al.2017).

Profiling of hemoglobin chains, mainly by electrophoresis (for which the equipment is most available in the country) not only lacks robustness during the neonatal period, these tests also have limited accuracy after transfusion. In addition, a fresh blood sample is the only specimen tested with these technics.

First, we implemented locally a test for SCA that distinguishes whether or not the wild-type allele has been substituted by the mutant allele, by virtue of restriction enzymes (RFLP). This implementation required local challenges to be evaluated and resolution strategies to be developed in order to ensure the uptake of the method. Our evaluation of the DBS proved that FTA Elute cards did not require specialized transportation nor cold chain for storage and that when stored on shelves at room temperature, they can be reused for up to four years (limit of our controls). This is an important advantage in developing countries where the specialized transportation for biospecimen is unavailable, the distance between the field and the laboratory is long, roads are not in good condition, the electricity supply is unreliable, and biobanks consist of fridges where samples and stored next to reagents. In addition, because multiple FTA cards well may be piled up to reduce storage space, FTA cards for the entire country can easily be stored in one facility for the 2 840 000 annual births in DR Congo (Tshilolo et al. 2009).

By skipping the DNA extraction step, the FTA cards reduce the initial cost by half (from 11.5 € to 6.5 €). This makes the DNA-based test about 2 times cheaper than the hemoglobin-based test (15 $). Moreover, skipping DNA extraction also led to a significant reduction in the turnaround time.

In this study, we proposed and validated FTA cards a DNA-based test that allows an accurate SCA diagnosis in both newborns and in post-transfusion periods, using venous blood, umbilical cord blood, or buccal swab samples.

Besides being cheaper than the standard of care (SoC) and robust, our approach is reliable on the one hand on the most challenging specimen such as cord blood, which is enriched with fetal hemoglobin. On the other hand, we demonstrated reliability in using buccal swab which is not

considered by the SoC despite being easily accessible. This is of relevance since using a buccal swab avoids blood sampling in patients with chronic anemia and does not cause pain as opposed to peripheral blood sampling, certainly in young children. Offering these painless biological sampling methods may gain in SCA testing acceptability among the population (Abraham et al.2012). Feasibility of the FTA-RFLP test from buccal swabs as an in vitro diagnostic test for SCA opens the doors for rapid, convenient, and inexpensive diagnosis for SCA.

It is well known that establishing the hemoglobin profile using blood sampled after transfusion is challenging because of the normal hemoglobin from the donor. We have demonstrated that this foreign hemoglobin could not fool a DNA-based test as it does with the Hb-based tests. The implication of this result is that testing for SCA in Congolese children should not be delayed because of a transfusion as suggested in the past (Burton et al. 2014).

Based on its reliability , we recommend FTA-RFLP to be used as (1) a first-tier diagnostic test for newborns from high-risk couples (both partners are heterozygotes) and (2) as the second-tier test for newborns with inconclusive or positive results at the standard newborn screening without having to wait the 3 months period as previously recommended (Tshilolo et al. 2009).

Our approach is limited in differentiating HbS from HbC, since both mutations target the same nucleotide position and any change at this position will remove the restriction site of the DdeI enzyme. Also, our test is not indicated for the diagnosis of compound heterozygous HbS/βthalassemia, which is anyway rare in the DR Congo setting. Finally, the FTA-RFLP is unable to diagnose SCD due to a point mutation in the HBB gene, different from the common E7V mutation. However, HbC and HbS/βthalassemia are rare as well (Tshilolo et al. 2016).

4.2 Genetic factors explaining biological and clinical variability in SCA patients in Congo

Although the causal mutation is unique, SCA patients exhibit heterogeneous presentations in the nature and the severity of clinical and biological manifestations. Previous studies have reported a correlation between the level of fetal hemoglobin (HbF) and the clinical expression of the disease. Variations in the baseline levels of fetal hemoglobin are under the control of three major eQTLs, namely in BCL11A, HBS1L-MYB, and HBB (Makani et al. 2011; Mikobi et al. 2017) and with a number of sickle haplotypes (Piel et al. 2017). Therefore, the

determination of the genetic profile of SCA patients would allow to prioritize treatment for those with predictive more severe phenotypes.

4.2.1 Haplotypes

Data on the distribution of SCA haplotypes in DR Congo are scarce. Until recently, researchers relied on publications performed in neighboring countries (Bitoungui et al. 2015). The first study to report on SCA haplotypes in DR Congo was published while we were sequencing our samples (Tshilolo et al., 2019). This international collaborative study included 600 SCA patients distributed in 150 participants per site (DRC, Kenya, Angola and Uganda). Similar to that study and to the previous estimates, the CAR haplotype has the highest frequency in our study. Over 90% of our patients were homozygous for the CAR haplotype.

The remaining patients shared compound heterozygous CAR/BEN, CAR/SEN, CAR/AI, and CAR/ATYP haplotypes. This indicates a high homogeneity of SCA patients in DR Congo, and would be reflecting a weak admixture with other populations.

We did not identify an association between the CAR alleles and the hematological profile and could not test the effect of homozygous status for the CAR haplotype on the hematological parameters given the almost absence of homozygous for the other haplotypes.

The CAR haplotype is reported to be characterized by the lowest level of HbF (Akinsheye et al. 2011). Consequently, without the protective effect of high HbF level, homozygous CAR patients are expected to be situated in the more severe spectrum of the disease (Piel et al.,2017).

Because of the lack of haplotype diversity in our cohort, we could not evaluate the impact of other haplotypes on HbF levels. However, due to the broad variation range of these levels in our study, we concluded that haplotypes were not likely the explanation for the variability of the baseline values of HbF. In addition, only 11 % of children and 8% of adults presented a severe clinical score as defined by Adegoke (Adegoke et al., 2013) and Mikobi (Mikobi et al., 2015). There may be an ascertainment bias to explain this observation, the possible early deaths of patients harboring the severe phenotypes could be a reasonable explanation.

4.2.2 Variants

The next step in the search for relevant genetic modifiers of the phenotype was to evaluate known SNPs.

In DR Congo, the available knowledge on the distribution of SNPs associated with HbF baseline levels results from the study by Mikobi few years ago (Mikobi et al. 2018). This previous study reported on eight SNPs, we aimed to explore a larger set of SNPs known for a potential association with HbF levels and the hematological parameters involved in the disease severity.

In our cohort of 215 patients , the BCL11A locus (rs 1427407) results as the most influent for HbF baseline level, particularly the T allele . However, other variants rs766432, rs11886868, rs4671393 and rs7606173 in BCL11A locus as well as rs28384513 in HBS1L-MYB locus , were also significatively associated with baseline HbF levels.

Together, variants within the three main quantitative traits loci , i. e. BCL11A, HBS1L-MYB and HBB (especially Xmn1-HBG2) have been reported to explain 20 to 50% of HbF level variability (Thein et al. 2013 ; Gueye et al. 2019) . However, through GWAS , new variants are still being described.

This emphasizes the importance of regular updates with the aim to find more potential therapeutic targets for sickle cell anemia.

High levels of HbF result in SCA phenotype improvement. Consequently, most therapeutic approaches target the increment of HbF level (Paikari et al. 2018). The BCL11A gene is among key regulators of HbF through its silencing action of globin gene expression . So under-expression of BCL11A in erythroid cells in adults SCA patients lead to stimulate HbF expression (Sankaran et al. 2008).

The important contribution of SNPs in the BCL11A loci in HbF level variability has been established by our study and several others, involving African SCA patients in Angola (Delgadinho et al. 2021) () , Cameroon (Wonkam et al. 2014) , Tanzania (Makani et al. 2011) as well as patients in India Bhanushali et al. 2015), Saudi Arabia (Ngo et al. 2013), Brazil (Friedrisch et al. 2016).

The HBG2 locus did not influence HbF variability in our cohort as well as in other sub-saharan SCA patients. In contrast, the rs7481244 in HBG2 has the most influence in HbF variability in SCA patients in other regions including Senegal (Gueye et al. 2019), Iraq (Al-Allawi et al. 2019) , India (Bhanushali et al. 2003) Brazil (Cardoso et al. 2014) and African Americans (Lettre et al. 2008) . The rs7481244 polymorphism "T allele" is reported to be associated with

the Arab Indian and Senegal haplotypes (Rahimi et al. 2003) which are frequent in those regions . So, the scarcity of this haplotype among African SCA patients may explain the observed discrepancy .

None of the HbF-associated SNPs reported in our study exert a simultaneous influence on hematological variables. Contrary to our results, several SNPs in BCL11A and HBS1L-MYB loci correlated to both HbF baseline levels and various hematological parameters including red blood cells, white blood cells and platelets counts as well as Hb level and red blood cells indices, in a Cameroonian and Tanzanian cohort of SCA patients (Wonkam et al. 2014 ; Mtatiro et al. 2015).

In our study , only the OR51B5/6 loci with rs5006884 and rs5024042 were significantly associated with platelet count without significant influence on HbF baseline level. While particularly , the SNP rs 5006884 influenced significatively HbF baseline levels in African Americans SCA patients (Solovieff et al. 2010). But in Cameroonians and Tanzanian SCA cohorts, the contribution of the OR51B5/6 loci was not significant .

This observed diversity requires more collaborative investigations to reach out the genetic background explaining those difference.

4.3 HU pharmacogenomic study

Hydroxyurea is the first and most widely available FDA-approved drug indicated to alleviate manifestations of SCA (Ware, 2015). However, an inter-individual variability in the response to HU treatment has been observed (Hankins et al. 2014). Genetic polymorphisms have been reported to strongly influence the response to hydroxyurea treatment, mainly in term of HU-induced HbF levels.

Pharmacogenomic determinants of HU response in SCA include the polymorphisms associated with baseline HbF levels in three main QTL (Friedrisch et al., 2008 ; Adekile, 2011 ; Ware et al., 2011) as well as other polymorphisms not related to the baseline HbF level (https://www.pharmgkb.org). To that, other genetic markers are also involved such as coinheritance with alpha thalassemia and beta S haplotypes.

This study involving a cohort of Congolese SCA patients was conducted to determine the influence of some candidate SNPs, α thalassemia status and βs haplotypes in HU response.

Coinheritance of α thalassemia did not influence HU response, neither the βs haplotypes as the majority of patients were harboring the CAR haplotype. However, within the CAR haplotype, a significant change was observed between HbF at baseline level and HU-induced HbF after two years of treatment. Six SNPs rs28384513 (HBS1L-MYB), rs9483947 (MAP3K5), rs765587 (TOX) and rs 9319428 (FLT1), rs1480642 and rs487278 (PDE7B)) were significantly associated with HU-induced HbF level.

The coinheritance with α-thalassemia in SCA is quite frequently encountered in SCA patients (Mikobi et al., 2018). Even if our results did not report significative influence in HU response as studies by Sheehan and co-workers (Sheehan et al. 2013) and by Figueiredo et al. (Figueiredo et al. 2017), the influence of α-thalassemia status remains to date controversial.

Several criteria are used to measure the change in HbF level induced by HU treatment : ΔMTD (maximum tolerated dose of HU) as difference between HbF level at MTD and HbF baseline level; ΔHbF as difference between final HbF level and HbF baseline level; ΔHbF as difference between maximum HbF and HbF baseline level. Also, the period of follow-up of HU treatment remains another criteria (Sales et al. 2022). This may influence at a certain degree the SNPs association. However, the set of SNPs found in our study as predictors of HU response in term of HU-induced HbF level is consistent with the report by Ma et al. in US Afro-American SCA patients (Ma et al. 2007). Both studies share similarities in period of HU treatment follow up (21 and 24 months) and change in HbF criteria.

The HU response in SCA involves several other criteria apart from HbF induction, such as other hematological, biochemical and clinical parameters.

As in this study we focused only on the HbF induction, which is the main mechanism, further studies are needed to completely assess the influence of these SNPs.

In practice, our study on genetic association faced limitations at two levels. The first one consists in SNPs selection. The available candidate approach appears to be less effective than a broader genotyping. We consider our study as a preliminary one until we can perform GWAS. The power of our study was also hampered by the fact that with the MIPs method, 11 SNPs failed to be genotyped, thus reducing the number of assessed SNPs.

The second limit is the number of patients included due to logistic limitations.

4.4 Usefulness of polygenic risk score (PRS) in SCA patients from DRC

A polygenic risk score (PRS) is an estimation of an individual's predicted risk to develop a disease based on its genetic profile. It consists in an algorithm based on cumulative effect of several SNPs significatively associated with a disease, rather than considering each variant individually. The benefit of a PRS is early identification of individuals at risk for the disease according to the score obtained. It modulates the risk conferred by the causal disease gene and family history (Mars et al. 2022). Therefore, preventive care can be applied for (high) risk individuals in the context of preventive medicine (Babb et al. 2020).

GWAS has allowed the discovery of several variants associated to HbF level in different SCA populations. The HbF level is one of the main modulators of the disease phenotype variability. Consequently, a PRS could be built based on each SNP's risk allele and effect size.

This could be advantageous, maybe even more in a low resource country, by priority selection of SCA patients likely to present a severe phenotype. For them, HU administration or curative treatment such as bone marrow transplantation or gene therapy would have a greater benefit in terms of better quality of life. We wanted to investigate whether PRS scores could be generated for the Congolese SCA population. Of course, we realized that our cohort was very small. In addition, the data (risk allele and effect size) used to build a PRS are mostly based on candidate SNPs obtained from Europeans or Afro-American populations, and data from African, particularly from sub-Saharan populations where the disease is more frequent, are not available. An illustration of the genetic differences between European and African -American populations with African natives, was recently given by Fatumo et al. who found several variants specific only to Ugandan populations (Fatumo, 2020).

Then the question is to which extent those PRS would be accurate for DR Congolese SCA patients? Certainly, the candidate SNPs approach that we used is less suitable than next generation sequencing methods would be (van Dijk et al. 2014) . These new sequencing technologies, by a complete analysis of the human genome, would allow the discovery of both frequent and rare variants associated with HbF levels variation. Unfortunately, those techniques are not affordable in the DRC setting.

Probably these limits explain the scarcity of PRS in the literature performed among African SCA patients.

In 2021, Rampersaud et al build a PRS for acute vaso-occlusive pain in children SCA patients in the US using the whole exome sequencing (WES)(Rampersaud et al.2021). Another study by Pincez T et al. elaborated a hematological traits PRS for SCA patients from the US cohort (Pincez et al., 2022). The study by Milton et al. constructed a PRS predictive for HbF levels, again based on American SCA patients (Milton et al., 2014). All three studies provided prediction on HbF levels in African -American SCA patients. The question now is to which extent those PRS can be of use for Congolese SCA patients. The prerequisite could be the common ancestral background . The difference in ancestral background with the population from which the PRS was validated, is among factor leading to poor performance of a PRS when applied in a different population (Martin et al. 2019). However this point should be considered in future studies.

References

Fatumo S.(2020). The opportunity in African genome resource for precision medicine. *EBioMedicine.* 54:102721

Martin AR, Kanai M, Kamatani Y, Okada Y, Neale BM & Daly MJ. (2019). Clinical use of current polygenic risk scores may exacerbate health disparities. Nat Genet. 51(4):584-591

Milton JN, Gordeuk VR, Taylor JG 6th, Gladwin MT, Steinberg MH & Sebastiani P. (2014). Prediction of fetal hemoglobin in sickle cell anemia using an ensemble of genetic risk prediction models. *Circ Cardiovasc Genet*. 7(2):110-5

Pincez T, Lo KS, D'Orengiani APHD, Garrett ME, Brugnara C, Ashley-Koch AE, Telen MJ, Galacteros F, Joly P, Bartolucci P & Lettre G. (2023).Variation and impact of polygenic hematologic traits in monogenic sickle cell disease. *Haematologica.* 108(3):870-881

Rampersaud E, Kang G, Palmer LE, Rashkin SR, Wang S, Bi W, Alberts NM, Anghelescu D, Barton M, Birch K, Boulos N, Brandow AM, Brooke RJ, Chang TC, Chen W, Cheng Y, Ding J, Easton J, Hodges JR, Kanne CK, Levy S, Mulder H, Patel AP, Puri L, Rosencrance C, Rusch M, Sapkota Y, Sioson E, Sharma A, Tang X, Thrasher A, Wang W, Yao Y, Yasui Y, Yergeau D, Hankins JS, Sheehan VA, Downing JR, Estepp JH, Zhang J, DeBaun M, Wu G & Weiss MJ. (2021). A polygenic score for acute vaso-occlusive pain in pediatric sickle cell disease. *Blood Adv.* 5(14):2839-2851

Sales RR, Nogueira BL & Luizon MR (2022). Pharmacogenomics of hydroxyurea therapy and fetal hemoglobin levels in sickle cell anemia. *Pharmacogenomics*. 23(7):393-396.

van Dijk EL, Auger H, Jaszczyszyn Y & Thermes C. (2014). Ten years of next-generation sequencing technology. Trends Genet. 30(9):418-26

Mars N, Lindbohm JV, Della Briotta Parolo P, Widén E, Kaprio J, Palotie A, FinnGen & Ripatti S. (2011). Systematic comparison of family history and polygenic risk across 24 common diseases. *Am J Hum Genet.* 109(12):2152-2162.

REFERENCES (to the Introduction and Discussion)

Agasa B, Bosunga K, Opara A, Tshilumba K, Dupont E, Vertongen F, Cotton F, Gulbis B. (2010). Prevalence of sickle cell disease in a northeastern region of the Democratic Republic of Congo: what impact on transfusion policy? *Transfusion Medicine*. 20(1):62–65

Ager JAM, Lehmann H, Vandepitte JM. (1958). ROLE OF PAPER ELECTROPHORESIS IN THE IDENTIFICATION OF HUMAN HÆMOGLOBIN VARIANTS. *The Lancet*. 271(7015):318–19

Agrawal RK, Patel RK, Shah V, Nainiwal L, Trivedi B. (2014). Hydroxyurea in Sickle Cell Disease: Drug Review. *Indian Journal of Hematology and Blood Transfusion*. 30(2):91–96

Akinsheye I, Alsultan A, Solovieff N, Ngo D, Baldwin CT, Sebastiani P, Chui D, Steinberg MH. (2011). Fetal hemoglobin in sickle cell anemia. *Blood*. 118(1):19–27

Alapan Y, Fraiwan A, Kucukal E, Hasan MN, Ung R, Kim M, Odame I, Little JA, Gurkan UA. (2016). Emerging point-of-care technologies for sickle cell disease screening and monitoring. *Expert Rev Med Devices*. 13(12):1073–93

AlDallal SM. (2020). Voxelotor: A Ray of Hope for Sickle Disease. *Cureus. 12(2):e7105*

Al-Allawi N, Qadir SMA, Puehringer H, Chui DHK, Farrell JJ, Oberkanins C.(2019). The association of HBG2, BCL11A, and HMIP polymorphisms with fetal hemoglobin and clinical phenotype in Iraqi Kurds with sickle cell disease. *Int J Lab Hematol*. 41(1):87-93

Aloni MN, Tshimanga BK, Ekulu PM, Ehungu JLG, Ngiyulu RM. (2013). Malaria, clinical features and acute crisis in children suffering from sickle cell disease in resource-limited settings: a retrospective description of 90 cases. *Pathog Glob Health*. 107(4):198–201

Aloni MN, Kadima BT, Ekulu PM, Budiongo AN, Ngiyulu RM, Gini-Ehungu JL. (2017).Acute Crises and Complications of Sickle Cell Anemia Among Patients Attending a Pediatric Tertiary Unit in Kinshasa, Democratic Republic of Congo. *Hematol Rep*.9(2):6952.

Archer NM, Petersen N, Clark MA, Buckee CO, Childs LM, Duraisingh MT. (2018). Resistance to *Plasmodium falciparum* in sickle cell trait erythrocytes is driven by oxygen-dependent growth inhibition. *Proceedings of the National Academy of Sciences*. 115(28):7350–55

Arishi WA, Alhadrami HA, Zourob M. (2021). Techniques for the Detection of Sickle Cell Disease: A Review. *Micromachines (Basel)*. 12(5):519

Ataga KI, Kutlar A, Kanter J, Liles D, Cancado R, Friedrisch J, Guthrie TH, Knight-Madden J, Alvarez OA, Gordeuk VR, Gualandro S, Colella MP, Smith WR, Rollins SA, Stocker JW, Rother RP (2017). Crizanlizumab for the Prevention of Pain Crises in Sickle Cell Disease. *New England Journal of Medicine*. 376(5):429–39

Babb de Villiers C, Kroese M, Moorthie S. (2020).Understanding polygenic models, their development and the potential application of polygenic scores in healthcare. *J Med Genet*. 57(11):725-732

Barabino GA, Platt MO, Kaul DK. (2010). Sickle Cell Biomechanics. *Annu Rev Biomed Eng*. 12(1):345–67

Barbarani G, Labedz A, Stucchi S, Abbiati A, Ronchi AE. 2021. Physiological and Aberrant γ-Globin Transcription During Development. *Front Cell Dev Biol*. 9:

Baruch DI, Ma XC, Singh HB, Bi X, Pasloske BL, Howard RJ. (1997). Identification of a Region of PfEMP1 That Mediates Adherence of Plasmodium falciparum Infected Erythrocytes to CD36: Conserved Function With Variant Sequence. *Blood*. 90(9):3766–75

Battersby AJ, Knox-Macaulay HHM, Carrol ED. (2010). Susceptibility to invasive bacterial infections in children with sickle cell disease. *Pediatr Blood Cancer*. 55(3):401–6

Bhanushali AA, Patra PK, Nair D, Verma H, Das BR. (2015).Genetic variant in BCL11A (rs1427407), but not HBS1-MYB (rs6934903) loci associate with fetal hemoglobin levels in Indian Sickle Cell disease patients. *Blood Cells Mol Dis*.54:4-8.

Bond M, Hunt B, Flynn B, Huhtinen P, Ware R, Richards-Kortum R. (2017). Towards a point-of-care strip test to diagnose sickle cell anemia. *PLoS One*. 12(5):e0177732

Bunn HF. (1997). Pathogenesis and Treatment of Sickle Cell Disease. *New England Journal of Medicine*. 337(11):762–69

Cardoso GL, Diniz IG, Silva AN, Cunha DA, Silva Junior JS, Uchôa CT, Santos SE, Trindade SM, Cardoso Mdo S, Guerreiro JF (2014). DNA polymorphisms at BCL11A, HBS1L-MYB and Xmn1-

HBG2 site loci associated with fetal hemoglobin levels in sickle cell anemia patients from Northern Brazil. *Blood Cells Mol Dis*.53:176-179
Charache S, Terrin ML, Moore RD, Dover GJ, Barton FB, Eckert SV, McMahon RP, Bonds DR . (1995). Effect of Hydroxyurea on the Frequency of Painful Crises in Sickle Cell Anemia. *New England Journal of Medicine*. 332(20):1317–22
Chen FT, Liu CM, Hsieh YZ, Sternberg JC. (1991). Capillary electrophoresis--a new clinical tool. *Clin Chem*. 37(1):14–19
Clark BE & Thein SL. (2004). Molecular diagnosis of haemoglobin disorders. *Clin Lab Haematol*. 26(3):159–76
Clarke GM& Higgins TN. (2000). Laboratory Investigation of Hemoglobinopathies and Thalassemias: Review and Update. *Clin Chem*. 46(8):1284–90
Cotton F, Wolff F, Gulbis B. (2013). Automated Capillary Electrophoresis in the Screening for Hemoglobinopathies. , pp. 227–35
Davis BA, Allard S, Qureshi A, Porter JB, Pancham S, Win N, Cho G, Ryan K. (2017). Guidelines on red cell transfusion in sickle cell disease Part II: indications for transfusion. *Br J Haematol*. 176(2):192–209
Delgadinho M, Ginete C, Santos B, Miranda A, Brito M.(2021).Genotypic Diversity among Angolan Children with Sickle Cell Anemia. *Int J Environ Res Public Health*.18(10):5417
de Rycke M, Berckmoes V. (2020). Preimplantation Genetic Testing for Monogenic Disorders. *Genes (Basel)*. 11(8):871
Depetris-Chauvin E, Weil DN. (2018). Malaria and Early African Development: Evidence from the Sickle Cell Trait. *The Economic Journal*. 128(610):1207–34
Elguero E, Délicat-Loembet LM, Rougeron V, Arnathau C, Roche B, Becquart P, Gonzalez JP, Nkoghe D, Sica L, Leroy EM, Durand P, Ayala FJ, Ollomo B, Renaud F, Prugnolle F. (2015). Malaria continues to select for sickle cell trait in Central Africa. *Proceedings of the National Academy of Sciences*. 112(22):7051–54
Fakher R, Bijan K, Taghi A. (2007). Application of diagnostic methods and molecular diagnosis of hemoglobin disorders in Khuzestan province of Iran. *Indian J Hum Genet*. 13(1):5
Fasano RM, Leong T, Kaushal M, Sagiv E, Luban NLC, Meier ER. (2016). Effectiveness of red blood cell exchange, partial manual exchange, and simple transfusion concurrently with iron chelation therapy in reducing iron overload in chronically transfused sickle cell anemia patients. *Transfusion (Paris)*. 56(7):1707–15
Fatumo S. (2020).The opportunity in African genome resource for precision medicine. *EBioMedicine*.54:102721
Feola M, Simoni J, Angelillo R, Luhruma Z, Kabakele M, Manzombi M, Kaluila M. (1992). Clinical trial of a hemoglobin based blood substitute in patients with sickle cell anemia. *Surg Gynecol Obstet*. 175(5):379–86
Figueiredo L, Morrone K, Wei C, Ireland K, Cohen HW, Driscoll C, Manwani D. (2017).An age dependent response to hydroxyurea in pediatric sickle cell anemia patients with alpha thalassemia trait. *Blood Cells Mol Dis*.66:19-23
Fort R. (2019). Recommendations for the use of red blood cell exchange in sickle cell disease. *Transfusion and Apheresis Science*. 58(2):128–31
Friedrisch JR, Sheehan V, Flanagan JM, Baldan A, Summarell CCG, Bittar CM, Friedrisch BK, Wilke II, Ribeiro CB, Daudt LE, da Rocha Silla LM (2016). The role of BCL11A and HMIP-2 polymorphisms on endogenous and hydroxyurea induced levels of fetal hemoglobin in sickle cell anemia patients from southern Brazil. *Blood Cells Mol Dis*. 62:32–37
Frömmel C. (2018). Newborn Screening for Sickle Cell Disease and Other Hemoglobinopathies: A Short Review on Classical Laboratory Methods—Isoelectric Focusing, HPLC, and Capillary Electrophoresis. *Int J Neonatal Screen*. 4(4):39
Fucharoen S, Winichagoon P. (2011). Haemoglobinopathies in Southeast Asia. *Indian J Med Res*. 134(4):498–506
Gladwin MT, Kato GJ. (2008). Hemolysis-associated hypercoagulability in sickle cell disease: the plot (and blood) thickens! *Haematologica*. 93(1):1–3

Gonçalves GAR, Paiva R de MA. (2017). Gene therapy: advances, challenges and perspectives. *Einstein (São Paulo)*. 15(3):369–75
Gueye Tall, Martin C, Ndour EHM, Renoux C, Ly ID, Connes P, Gueye PM, Diallo RN, Diagne I, Diop PA, Cissé A, Lopez Sall P, Joly P (2019). Combined and differential effects of alpha-thalassemia and HbF-quantitative trait loci in Senegalese hydroxyurea-free children with sickle cell anemia. *Pediatr. Blood Cancer*.66, e27934
Halsey C, Roberts IAG. (2003). The role of hydroxyurea in sickle cell disease. *Br J Haematol*. 120(2):177–86
Herrick JB. (2001). Peculiar elongated and sickle-shaped red blood corpuscles in a case of severe anemia. 1910. *Yale J Biol Med*. 74(3):179–84
Houwing ME, de Pagter PJ, van Beers EJ, Biemond BJ, Rettenbacher E, Rijneveld AW, Schols EM, Philipsen JNJ, Tamminga RYJ, van Draat KF, Nur E, Cnossen MH. (2019). Sickle cell disease: Clinical presentation and management of a global health challenge. *Blood Rev*. 37:100580
Howard J, Hemmaway CJ, Telfer P, Layton DM, Porter J, Awogbade M, Mant T, Gretler DD, Dufu K, Hutchaleelaha A, Patel M, Siu V, Dixon S, Landsman N, Tonda M, Lehrer-Graiwer J. (2019). A phase 1/2 ascending dose study and open-label extension study of voxelotor in patients with sickle cell disease. *Blood*. 133(17):1865–75
Joutovsky A, Hadzi-Nesic J, Nardi MA. (2004). HPLC Retention Time as a Diagnostic Tool for Hemoglobin Variants and Hemoglobinopathies: A Study of 60000 Samples in a Clinical Diagnostic Laboratory. *Clin Chem*. 50(10):1736–47
Kabanova S, Kleinbongard P, Volkmer J, Andrée B, Kelm M, Jax TW. (2009). Gene expression analysis of human red blood cells. *Int J Med Sci*. 156–59
Kabinda JM, Mitashi PM, Chenge FM. (2019). analysis of financing arrangements for health care in the Democratic Republic of Congo: a systematic review. *Annales Africaines de Médecine*. 12(2):
Kadima BT, Gini Ehungu JL, Ngiyulu RM, Ekulu PM, Aloni MN. (2015).High rate of sickle cell anaemia in Sub-Saharan Africa underlines the need to screen all children with severe anaemia for the disease. *Acta Paediatr*.104(12):1269-73
Kambale-Kombi P, Marini Djang'eing'a R, Alworong'a Opara J-P, Minon J-M, Boemer F, Bours V, Tonen-Wolyec S, Kayembe Tshilumba C, Batina-Agasa S . (2021). Management of sickle cell disease: current practices and challenges in a northeastern region of the Democratic Republic of the Congo. *Hematology*. 26(1):199–205
Kambale-Kombi P, Marini Djang'eing'a R, Alworong'a Opara J-P, Tonen-Wolyec S, Kayembe Tshilumba C, Batina-Agasa S. (2020). Students' knowledge on sickle cell disease in Kisangani, Democratic Republic of the Congo. *Hematology*. 25(1):91–94
Kanter J, Falcon C. (2021). Gene therapy for sickle cell disease: where we are now? *Hematology*. 2021(1):174–80
Kasai ET, Boemer F, Djang'eing'a RM, Ntokumunda JK, Agasa SB, Ngbonda ND & Opara JP. (2020). Systematic Screening of Neonatal Sickle Cell Disease with HemoTypeSC^{TM} Kit-Test: Case Study and Literature Review. *Open Journal of Blood Diseases*. 10(01):12–21
Katawandja AL, Nzengu-Lukusa F, Nzongola-Nkasu DK, Muepu LT. (2021). Dépistage néonatal de la drépanocytose dans la ville de Kindu, à l'Est de la République Démocratique du Congo: étude préliminaire dans neuf maternités. *PAMJ Clinical Medicine*. 5:
Kaul DK, Finnegan E, Barabino GA. (2009). Sickle Red Cell–Endothelium Interactions. *Microcirculation*. 16(1):97–111
King L, Fraser R, Forbes M, Grindley M, Ali S, Reid M. (2007). Newborn sickle cell disease screening: the Jamaican experience (1995–2006). *J Med Screen*. 14(3):117–22
Konotey-Ahulu FID. (1968). Hereditary Qualitative and Quantitative Erythrocyte Defects in Ghana: An Historical and Geographical Survey. *Ghana Med J* . 7:118–19
Krishnamurti L, Neuberg DS, Sullivan KM, Kamani NR, Abraham A, Waller EK, Walters MC. (2019). Bone marrow transplantation for adolescents and young adults with sickle cell disease: Results of a prospective multicenter pilot study. *Am J Hematol*. 94(4):446–54
Kumar AA, Chunda-Liyoka C, Hennek JW, Mantina H, Lee SYR, Patton MR, Sambo P, Sinyangwe S, Kankasa C, Chintu C, Brugnara C, Stossel TP, Whitesides GM . (2014). Evaluation of a Density-

Based Rapid Diagnostic Test for Sickle Cell Disease in a Clinical Setting in Zambia. *PLoS One*. 9(12):e114540
Kumar R, Derbigny WA. (2019). Cellulose Acetate Electrophoresis of Hemoglobin. *Methods Mol Biol*.1855:81-85
Lambotte-Legrand J, Lambotte-Legrand C. (1958). [Additional notes on drepanocytosis. I. Sicklemia & malaria]. *Ann Soc Belg Med Trop (1920)*. 38(1):45–53
LaMonte G, Philip N, Reardon J, Lacsina JR, Majoros W.(2012). Translocation of Sickle Cell Erythrocyte MicroRNAs into Plasmodium falciparum Inhibits Parasite Translation and Contributes to Malaria Resistance. *Cell Host Microbe*. 12(2):187–99
Lettre G, Sankaran VG, Bezerra MAC, Araújo AS, Uda M, Sanna S, Cao A, Schlessinger D, Costa FF, Hirschhorn JN, Orkin SH. (2008). DNA polymorphisms at the BCL11A, HBS1L-MYB, and β-globin loci associate with fetal hemoglobin levels and pain crises in sickle cell disease. *Proc Natl Acad Sci.* 2008.105:11869-11874
Luboya E, Tshilonda J-CB, Ekila MB, Aloni MN. (2014). Répercussions psychosociales de la drépanocytose sur les parents d'enfants vivant à Kinshasa, République Démocratique du Congo: une étude qualitative. *Pan African Medical Journal*. 19:
Luzzatto L. (2012). Sickle cell anaemia and malaria. *Mediterr J Hematol Infect Dis*. 4(1):e2012065
Luzzatto L, Nwachuku-Jarrett ES, Reddy S. (1970). Increased sickling of parasitised erythrocytes as mechanism of resistance against malaria in the sickle-cell trait. *The Lancet*. 295(7642):319–22
Ma Q, Wyszynski DF, Farrell JJ, Kutlar A, Farrer LA, Baldwin CT, Steinberg MH. (2007). Fetal hemoglobin in sickle cell anemia: genetic determinants of response to hydroxyurea. *Pharmacogenomics J*. 7(6):386-94
Maier-Redelsperger M, Noguchi C, de Montalembert M, Rodgers G, Schechter A. (1994). Variation in fetal hemoglobin parameters and predicted hemoglobin S polymerization in sickle cell children in the first two years of life: Parisian Prospective Study on Sickle Cell Disease. *Blood*. 84(9):3182–88
Makani J, Ofori-Acquah SF, Nnodu O, Wonkam A, Ohene-Frempong K. (2013). Sickle Cell Disease: New Opportunities and Challenges in Africa. *The Scientific World Journal*. 2013:1–16
Makani J, Menzel S, Nkya S, Cox SE, Drasar E, Soka D, Komba AN, Mgaya J, Rooks H, Vasavda N, Fegan G, Newton CR, Farrall M, Thein SL. (2011). Genetics of fetal hemoglobin in Tanzanian and British patients with sickle cell anemia. *Blood* .117:1390–2
Marengo-Rowe AJ.(2007). The Thalassemias and Related Disorders. *Baylor University Medical Center Proceedings*. 20(1):27–31
Mars N, Lindbohm JV, Della Briotta Parolo P, Widén E, Kaprio J, Palotie A; FinnGen; Ripatti S. (2022). Systematic comparison of family history and polygenic risk across 24 common diseases. *Am J Hum Genet*.109(12):2152-2162.
Mbiya BM, Kalombo DK, Mukendi YN, Daubie V, Mpoyi JK, Biboyi PM, Disashi GT, Gulbis B. (2021). Improvement of SCD morbimortality in children: experience in a remote area of an African country. *BMC Health Serv Res*. 21(1):294
McGann PT. (2014). Sickle Cell Anemia: An Underappreciated and Unaddressed Contributor to Global Childhood Mortality. *J Pediatr*. 165(1):18–22
McGann PT, Ferris MG, Ramamurthy U, Santos B, Oliveira V, Bernardino L, Ware RE. (2013). A prospective newborn screening and treatment program for sickle cell anemia in Luanda, Angola. *Am J Hematol*. 88(12):984–89
McGann PT, Hoppe C. (2017). The pressing need for point-of-care diagnostics for sickle cell disease: A review of current and future technologies. *Blood Cells Mol Dis*. 67:104–13
Mekontso Dessap A, Contou D, Dandine-Roulland C, Hemery F, Habibi A. (2014). Environmental Influences on Daily Emergency Admissions in Sickle-Cell Disease Patients. *Medicine*. 93(29):e280
Mikobi TM, Lukusa PT, Aloni MN, Lumaka A, Akilimali PZ, Devriendt K, Matthijs G, Mbuyi Muamba JM, Race V.(2018=a. Association between sickle cell anemia and alpha thalassemia reveals a high prevalence of the $\alpha^{3.7}$ triplication in congolese patients than in worldwide series. *J Clin Lab Anal*. 32(1):e22186
Mikobi TM, Tshilobo Lukusa P, Aloni MN, Lumaka AZ, Kaba DK, Devriendt K, Matthijs G, Mbuyi Muamba JM, Race V. (2018)b. Protective *BCL11A* and *HBS1L-MYB* polymorphisms in a cohort of 102 Congolese patients suffering from sickle cell anemia. *J Clin Lab Anal*. 32(1):e22207

Mikobi TM, Lukusa Tshilobo P, Aloni MN, Mvumbi Lelo G, Akilimali PZ, Muyembe-Tamfum JJ, Race V, Matthijs G, Mbuyi Mwamba JM. (2015). Correlation between the Lactate Dehydrogenase Levels with Laboratory Variables in the Clinical Severity of Sickle Cell Anemia in Congolese Patients. *PLoS One*. 10(5):e0123568

Milton JN, Gordeuk VR, Taylor JG 6th, Gladwin MT, Steinberg MH, Sebastiani P.(2014). Prediction of fetal hemoglobin in sickle cell anemia using an ensemble of genetic risk prediction models. Circ Cardiovasc Genet. 7(2):110-5.

Minniti CP. (2018). <scp>l</scp> -Glutamine and the Dawn of Combination Therapy for Sickle Cell Disease. *New England Journal of Medicine*. 379(3):292–94

Modell B. (2008). Global epidemiology of haemoglobin disorders and derived service indicators. *Bull World Health Organ*. 2008(6):480–87

Modell B, Darlison M, Birgens H, Cario H, Faustino P. (2007). Epidemiology of haemoglobin disorders in Europe: an overview. *Scand J Clin Lab Invest*. 67(1):39–70

Mpiana PT, Mudogo V, Tshibangu DST, Kitwa EK, Kanangila AB.(2008). Antisickling activity of anthocyanins from Bombax pentadrum, Ficus capensis and Ziziphus mucronata: Photodegradation effect. *J Ethnopharmacol*. 120(3):413–18

Mpiana PT, Ngbolua KNN, Bokota MT, Kasonga TK, Atibu EK. (2010). In vitro effects of anthocyanin extracts from Justicia secunda Vahl on the solubility of haemoglobin S and membrane stability of sickle erythrocytes. *Blood Transfus*. 8(4):248–54

Mpiana PT, Ngbolua K-N, Tshibangu STD. (2016). Les alicaments et la drépanocytose : une mini-revue. *Comptes Rendus Chimie*. 19(7):884–89

Mpiana PT, Tshibangu DST, Shetonde OM, Ngbolua KN. (2007). In vitro antidrepanocytary actvity (anti-sickle cell anemia) of some congolese plants. *Phytomedicine*. 14(2–3):192–95

Mtatiro SN, Makani J, Mmbando B, Thein SL, Menzel S, Cox SE. (2015).Genetic variants at HbF-modifier loci moderate anemia and leukocytosis in sickle cell disease in Tanzania. *Am J Hematol*. 90(1):E1-4

Mukinayi BM, Cibeyibeyi GK, Tumba GD, Gulbis B. (2021). Drépanocytose en République Démocratique du Congo: quels sont les obstacles à un traitement par hydroxyurée? *Pan African Medical Journal*. 38:

Mukinayi BM, Kalenda DK, Mbelu S, Gulbis B.(2018). Connaissances et comportements de 50 familles congolaises concernées par la drépanocytose: une enquête locale. *Pan African Medical Journal*. 29:

Mukinayi Mbiya B, Tumba Disashi G, Gulbis B. 2020. Sickle Cell Disease in the Democratic Republic of Congo: Assessing Physicians' Knowledge and Practices. *Trop Med Infect Dis*. 5(3):127

Mulumba LL, Wilson L. 2015. Sickle cell disease among children in Africa: An integrative literature review and global recommendations. *Int J Afr Nurs Sci*. 3:56–64

Musumari PM, Mbikayi SM, Srithanaviboonchai K, Techasrivichien T, Tangmunkongvorakul A, Tshilolo L. (2021). Symptoms of acute transfusion reactions at a general referral hospital in Kinshasa, Democratic Republic of Congo: a cross-sectional study. *BMJ Open*. 11(7):e045081

Nair S. (2018). Potential Pithfalls in Using HPLC and its Interpretation in Diagnosing HbS. *J Rare Dis Res Treat*. 3(3):9–12

Newton CR, Graham A, Heptinstall LE, Powell SJ, Summers C.(1989). Analysis of any point mutation in DNA. The amplification refractory mutation system (ARMS). *Nucleic Acids Res*. 17(7):2503–16

Ngo D, Bae H, Steinberg MH, Akinsheye I, Gallagher P, Luo HY, Chui DH, Farrell JJ, Al-Ali AK, Alsultan A. (2013).Fetal hemoglobin in sickle cell anemia: genetic studies of the Arab-Indian haplotype. *Blood Cells Mol Dis*.51:22-26

Niihara Y, Miller ST, Kanter J, Lanzkron S, Smith WR. (2018). A Phase 3 Trial of <scp>l</scp> -Glutamine in Sickle Cell Disease. *New England Journal of Medicine*. 379(3):226–35

Nnodu O, Isa H, Nwegbu M, Ohiaeri C, Adegoke S. (2019). HemoTypeSC, a low-cost point-of-care testing device for sickle cell disease: Promises and challenges. *Blood Cells Mol Dis*. 78:22–28

Nussbaum RL. 2007. *Thompson & Thompson Genetics in Medicine*. Philadelphia: SAUNDERS Elsevier. 7th ed.

Odame I. (2014). Perspective: We need a global solution. *Nature*. 515(7526):S10–S10

Odievre M-H, Bony V, Benkerrou M, Lapoumeroulie C, Alberti C. (2008). Modulation of erythroid adhesion receptor expression by hydroxyurea in children with sickle cell disease. *Haematologica*. 93(4):502–10

Ohene-Frempong K, Oduro J, Tetteh H, Nkrumah F. (2008). SCREENING NEWBORNS FOR SICKLE CELL DISEASE IN GHANA. *Pediatrics*. 121(Supplement_2):S120–21

Okwi A, Ocaido M, Byarugaba W, Ndugwa C, Parkes A. (2009). Sickling and solubility tests and the peripheral blood film method for screening for sickle cell disease. *S Afr Med J*. 99(12):887–91

Ou C-N, Rognerud CL. (2001). Diagnosis of hemoglobinopathies: electrophoresis vs. HPLC. *Clinica Chimica Acta*. 313(1–2):187–94

Paikari A & Sheehan VA. (2018).Fetal haemoglobin induction in sickle cell disease. *Br J Haematol* .180:189–200

Pays J-F. (2010). Toutankhamon et sicklanémie. *Bulletin de la Société de pathologie exotique*. 103(5):346–47

Piel FB, Hay SI, Gupta S, Weatherall DJ, Williams TN. (2013)a. Global Burden of Sickle Cell Anaemia in Children under Five, 2010–2050: Modelling Based on Demographics, Excess Mortality, and Interventions. *PLoS Med*. 10(7):e1001484

Piel FB, Howes RE, Patil AP, Nyangiri OA, Gething PW . (2013)b. The distribution of haemoglobin C and its prevalence in newborns in Africa. *Sci Rep*. 3(1):1671

Piel FB, Patil AP, Howes RE, Nyangiri OA, Gething PW, Dewi M, Temperley WH, Williams TN, Weatherall DJ, Hay SI. (2013)c. Global epidemiology of sickle haemoglobin in neonates: a contemporary geostatistical model-based map and population estimates. *The Lancet*. 381(9861):142–51

Piety NZ, Yang X, Kanter J, Vignes SM, George A, Shevkoplyas SS. (2016). Validation of a Low-Cost Paper-Based Screening Test for Sickle Cell Anemia. *PLoS One*. 11(1):e0144901

Pincez T, Lo KS, D'Orengiani APHD, Garrett ME, Brugnara C, Ashley-Koch AE, Telen MJ, Galacteros F, Joly P, Bartolucci P, Lettre G. (2022).Variation and impact of polygenic hematological traits in monogenic sickle cell disease. Haematologica oct 13

Quinn CT. (2018). l-Glutamine for sickle cell anemia: more questions than answers. *Blood*. 132(7):689–93

Quinn CT, Paniagua MC, DiNello RK, Panchal A, Geisberg M. (2016). A rapid, inexpensive and disposable point-of-care blood test for sickle cell disease using novel, highly specific monoclonal antibodies. *Br J Haematol*. 175(4):724–32

Quinn CT, Rogers ZR, McCavit TL, Buchanan GR. (2010). Improved survival of children and adolescents with sickle cell disease. *Blood*. 115(17):3447–52

Rahimy MC, Gangbo A, Ahouignan G, Alihonou E. (2009). Newborn screening for sickle cell disease in the Republic of Benin. *J Clin Pathol*. 62(1):46–48

Rahimi Z, Karimi M, Haghshenass M, Merat A. (2003).Beta-globin gene cluster haplotypes in sickle cell patients from southwest Iran. *Am J Hematol*.74:156–160.

Reddy M, Franciosi R. (1994). Rapid quantitation of hemoglobin S by isoelectric focusing. *Ann Clin Lab Sci*. 24(5):401–6

Rees DC. (2011). The rationale for using hydroxycarbamide in the treatment of sickle cell disease. *Haemaatologica*. 96(4):488–91

Rees DC, Robinson S, Howard J. (2018). How I manage red cell transfusions in patients with sickle cell disease. *Br J Haematol*. 180(4):607–17

Rees DC, Williams TN, Gladwin MT. (2010). Sickle-cell disease. *The Lancet*. 376(9757):2018–31

Ribeil J-A, Hacein-Bey-Abina S, Payen E, Magnani A, Semeraro M, Ross RW, Negre O, Veres G, Sandler L, Soni S, de Montalembert M, Blanche S, Leboulch P, Cavazzana M. (2017). Gene Therapy in a Patient with Sickle Cell Disease. *New England Journal of Medicine*. 376(9):848–55

Roberts I, de Montalembert M. (2007). Sickle cell disease as a paradigm of immigration hematology: new challenges for hematologists in Europe. *Haematologica*. 92(7):865–71

Robinson TM, Fuchs EJ. (2016). Allogeneic stem cell transplantation for sickle cell disease. *Curr Opin Hematol*. 23(6):524–29

Saiki RK, Scharf S, Faloona F, Mullis KB, Horn GT. (1985). Enzymatic Amplification of β-Globin Genomic Sequences and Restriction Site Analysis for Diagnosis of Sickle Cell Anemia. *Science (1979)*. 230(4732):1350–54

Sales RR, Nogueira BL, Tosatti JAG, Gomes KB, Luizon MR. (2022). Do Genetic Polymorphisms Affect Fetal Hemoglobin (HbF) Levels in Patients with Sickle Cell Anemia Treated with Hydroxyurea? A Systematic Review and Pathway Analysis. *Front Pharmacol*. 12:779497.

Sankaran VG, Menne TF, Xu J, Akie TE, Lettre G, Van Handel B, A Mikkola HK, Hirschhorn JN, Cantor AB, Orkin SH. (2008).Human fetal hemoglobin expression is regulated by the developmental stage-specific repressor BCL11A. *Science*.322(5909):1839-42

Santana SS, Pitanga TN, de Santana JM, Zanette DL, Vieira J de J. (2020). Hydroxyurea Scavenges Free Radicals and Induces the Expression of Antioxidant Genes in Human Cell Cultures Treated With Hemin. *Front Immunol*. 11:

Schechter AN. (2008). Hemoglobin research and the origins of molecular medicine. *Blood*. 112(10):3927–38

Serjeant GR. (1997). Sickle-cell disease. *The Lancet*. 350(9079):725–30

Sharma P, Das R. (2016). Cation-exchange high-performance liquid chromatography for variant hemoglobins and HbF/A2: What must hematopathologists know about methodology? *World J Methodol*. 6(1):20

Sheehan VA, Luo Z, Flanagan JM, Howard TA, Thompson BW, Wang WC, Kutlar A, Ware RE; BABY HUG Investigators.(2013).Genetic modifiers of sickle cell anemia in the BABY HUG cohort: influence on laboratory and clinical phenotypes. *Am J Hematol*. 88(7):571-6.

Shenoy S. (2013). Hematopoietic stem-cell transplantation for sickle cell disease: current evidence and opinions. *Ther Adv Hematol*. 4(5):335–44

St John A, Price C. (2014). Existing and Emerging Technologies for Point-of-Care Testing. *Clin Biochem Rev*. 35(3):155–67

Stamatoyannopoulos G. (2005). Control of globin gene expression during development and erythroid differentiation. *Exp Hematol*. 33(3):259–71

Steinberg MH, Sebastiani P. (2012). Genetic modifiers of sickle cell disease. *Am J Hematol*. 87(8):795–803

Solovieff N, Milton JN, Hartley SW, Sherva R, Sebastiani P, et al. 2010.Fetal hemoglobin in sickle cell anemia: genome-wide association studies suggest a regulatory region in the 5 olfactory receptor gene cluster. *Blood*.115:1815–1822

Sundd P, Gladwin MT, Novelli EM. (2019). Pathophysiology of Sickle Cell Disease. *Annual Review of Pathology: Mechanisms of Disease*. 14(1):263–92

Tewari S, Brousse V, Piel FB, Menzel S, Rees DC. (2015). Environmental determinants of severity in sickle cell disease. *Haematologica*. 100(9):1108–16

Thein, S.L. (2013).Genetic association studies in β-hemoglobinopathies. *Hematology*.2013:354–361

Thom CS, Dickson CF, Gell DA, Weiss MJ. (2013). Hemoglobin Variants: Biochemical Properties and Clinical Correlates. *Cold Spring Harb Perspect Med*. 3(3):a011858–a011858

Timson J. (1975). Hydroxyurea. *Mutation Research/Reviews in Genetic Toxicology*. 32(2):115–31

Toye ET, van Marle G, Hutchins W, Abgabiaje O, Okpuzor JO. (2018). Single tube allele specific PCR: a low cost technique for molecular screening of sickle cell anaemia in Nigeria. *Afr Health Sci*. 18(4):995

Tshilolo L, Aissi LM, Lukusa D, Kinsiama C, Wembonyama S, Gulbis B, Vertongen F. (2009). Neonatal screening for sickle cell anaemia in the Democratic Republic of the Congo: experience from a pioneer project on 31 204 newborns. *J Clin Pathol*. 62(1):35–38

Tshilolo L, Kafando E, Sawadogo M, Cotton F, Vertongen F, Ferster A, Gulbis B. (2008). Neonatal screening and clinical care programmes for sickle cell disorders in sub-Saharan Africa: Lessons from pilot studies. *Public Health*. 122(9):933–41

Tshilolo L, Mukendi R, Girot R. (1996). La drépanocytose dans le sud du Zaïre. Étude de deux séries de 251 et 340 malades suivis entre 1988 et 1992. *Archives de Pédiatrie*. 3(2):104–11

Tshilolo L, Tomlinson G, Williams TN, Santos B, Olupot-Olupot P, Lane A, Aygun B, Stuber SE, Latham TS, McGann PT, Ware RE; REACH Investigators. (2019). Hydroxyurea for Children with Sickle Cell Anemia in Sub-Saharan Africa. *New England Journal of Medicine*. 380(2):121–31

Tshilolo L, Zita MN, Ngiyulu R, Kayembe Nzongola D. (2016). Iron status in 72 Congolese patients with sickle cell anemia. *Med Sante Trop*. 26(1):83–87

Tshilolo LM, Mukendi RK, Wembonyama SO. (2007). Blood transfusion rate in congolese patients with sickle cell anemia. *The Indian Journal of Pediatrics*. 74(8):735–38

Tubman VN, Field JJ. (2015). Sickle solubility test to screen for sickle cell trait: what's the harm? *Hematology*.(1):433–35

United Nations. (2009). *United Nations press office Press conference on raising awareness of sickle-cell anaemia*. www.un.org/News/briefings/docs/2009/090619_Anaemia.doc.htm
Vandepite J, Delaisse J. (1920). Sicklemia and malaria; review of the problem and personal contribution. *Ann Soc Belg Med Trop*. 31(5):703–35
Vandepite JM, Zuelzer WW, Neel J v., Colaert J. (1955). Evidence Concerning the Inadequacy of Mutation as an Explanation of the Frequency of the Sickle Cell Gene in the Belgian Congo. *Blood*. 10(4):341–50
Vandepitte JM, Pieters G. (1952). Case of sickle cell anemia in a young adults in the Belgian Congo; possibilities of laboratory diagnosis. *Ann Soc Belg Med Trop*. 32(3):281–88
Wajcman H, Moradkhani K. (2011). Abnormal haemoglobins: detection & characterization. *Indian J Med Res*. 134(4):538–46
Wang Q, Zennadi R. (2021). The Role of RBC Oxidative Stress in Sickle Cell Disease: From the Molecular Basis to Pathologic Implications. *Antioxidants*. 10(10):1608
Wang WC, Ware RE, Miller ST, Iyer R v, Casella JF.(2011). Hydroxycarbamide in very young children with sickle-cell anaemia: a multicentre, randomised, controlled trial (BABY HUG). *The Lancet*. 377(9778):1663–72
Ware RE, de Montalembert M, Tshilolo L, Abboud MR. (2017). Sickle cell disease. *The Lancet*. 390(10091):311–23
Wembonyama S, Mpaka S, Tshilolo L. (2007). Médecine et santé en République Démocratique du Congo : de l'indépendance à la 3e République. *Med Trop*. 67:447–57
WHO Regional Office for Africa. (2010). Sickle-cell disease: a strategy for the WHO African Region: Report of the Regional Director. AFR/RC60/8. Geneva
Williams TN, Mwangi TW, Roberts DJ, Alexander ND, Weatherall DJ.(2005). An Immune Basis for Malaria Protection by the Sickle Cell Trait. *PLoS Med*. 2(5):e128
Wilson JT, Milner PF, Summer ME, Nallaseth FS, Fadel HE. (1982). Use of restriction endonucleases for mapping the allele for beta s-globin. *Proceedings of the National Academy of Sciences*. 79(11):3628–31
Yanni E, Grosse SD, Yang Q, Olney RS. (2009). Trends in Pediatric Sickle Cell Disease-Related Mortality in the United States, 1983-2002. *J Pediatr*. 154(4):541–45
Yue L, Lin M, Chen J-T, Zhan X-F, Zhong D-S, Ehapo JC, Eyi UM, Yang HT, Yang LY. (2014). Rapid screening for sickle cell disease by polymerase chain reaction-high resolution melting analysis. *Mol Med Rep*. 9(6):2479–84
Zago MA, Pinto ACS. (2007). Physiopathology of sickle cell disease: From genetic mutation to insufficiency of multiple organs. *Rev Bras Hematol Hemoter*. 29(3):207–14
Zheng Y, Chou ST. (2021). Transfusion and Cellular Therapy in Pediatric Sickle Cell Disease. *Clin Lab Med*. 41(1):101–19

5. SUMMARY

Sickle cell anemia (SCA) is the most common hemoglobinopathy worldwide. Developing countries like the Democratic Republic of Congo (DRC) suffer a heavy burden and the disease-related mortality remains high, up to 50% to 90% before five years of age. Factors that contribute to high mortality include delayed diagnosis due to the scarcity of appropriate laboratory infrastructure, poor access to adequate medical care due to limited financial resources of patients, inadequate supply of state-of the art medical, prevailing mystical beliefs and use of traditional herbal medicine. Despite the fact that a single mutation in the β globin gene causes SCA, a marked interindividual variability in clinical manifestations and severity as well as the response to hydroxyurea (HU) is observed among SCA patients. Genetic factors are among factors that could explain this variability.

We conducted this study to improve the care to SCA patients though accurate, rapid and efficient diagnostic and a better understanding of the influence of genomic polymorphisms on clinical and biological profile and on the response to HU treatment in Congolese SCA patients.

The overall cohort consisted of 338 patients of whom 102 were neonates, 236 children and adults. Neonates were recruited at 2 maternities of Kinshasa i.e. Maternity of Binza and Maternity of Kingasani. Recruitment for children and adults were performed respectively at the Saint Luc Hospital in Kisantu (Kongo central Province) and the "Centre de Médecine Mixte et d'Anémie SS (CMMASS)" , St Crispin Hospital , and University Hospitals of the University of Kinshasa at Kinshasa (the capital city) . All sites are located in the DRCongo.

In the first part, we aimed to implement a DNA-based SCA test that would be suited for a limited resource setting and overcome limitations of hemoglobin (Hb)-based tests commonly used in the current standard of care. The SCA Hb-based tests are not reliable particularly in neonatal and post-transfusion periods.

Three types of biological samples were collected: venous blood, umbilical cord blood (UCB) and buccal swab. A drop of blood (approximately 50 µl of blood) (venous and UCB) was deposited on a circle of an FTA Elute card (Whatman WB120206).Buccal swabs collected using a homemade sterile swab, was also rubbed on a circle of FTA elute card for 10 seconds in order to transfer the sample. After DNA extraction by the manual salting out method (SO), SCA molecular testing (RFLP) was performed to identify the sickled mutated allele (SO-RFLP

method). We adapted the SO-RFLP protocol by replacing the extracted DNA sample by a 1.2 mm punch of FTA Elute cards, thus avoiding an expensive DNA extraction step.

There was a perfect agreement between both protocols i.e. SO-RFLP and FTA-RFLP. By skipping the DNA extraction step, the FTA-RFLP reduces the cost by half (from 11.5 € to 6.5 €) and turn-around time significantly. FTA Elute cards with blood spots were conserved in a clean closet at room temperature (RT) and retested after 2 years. We obtained the exact same results.

The value and utility of DNA-based test on DBS in post-transfusion period is an ongoing study where FTA-RFLP is performed before and 1 hour after a blood transfusion. Our preliminary results (on 6 patients) show that the yield of FTA-RFLP remains unchanged after packed cells transfusion contrary to the two Hb-based tests.

Thus, the FTA-RFLP can be recommended as (1) a first-tier diagnostic test for newborns from high risk couples (both partners are heterozygotes) and (2) as the second-tier test for newborns with inconclusive or positive results at the standard newborn screening without having to wait the 3 months period. The test can also be offered immediately after transfusion, avoiding the currently long waiting time required for a reliable Hb electrophoresis The easy transportation as well as the proven good conservation of samples suggests that FTA Elute card is a suited collection device in a resource limited setting. Moreover, storage at RT makes access to cold storage facilities unnecessary, which is a critical advantage in developing countries where electricity supply is unreliable

In the second part, we focused on the evaluation of genetic factors involved in modulating the clinical and biological expression of SCA, as well as the pharmacogenetics of HU.

We first assessed the distribution of βS-globin haplotypes and the correlation between haplotypes and baseline hematological and clinical profiles in our cohort of patients recruited at the Saint Luc Hospital in Kisantu and the "Centre de Médecine Mixte et d'Anémie SS (CMMASS)". Patients were classified according to the clinical phenotype score as previously described.

Then we evaluated a set of single nucleotide polymorphisms (SNPs), for their potential association with HbF levels and the hematological parameters involved in the disease severity.

Finally, we determined the contribution of α-thalassemia coinheritance, the βs haplotypes and candidate SNPs in HU response in the set of patients within our initial cohort , who were included in a HU clinical trial.

Prior to be included in this part of the study, SCA homozygous status had to be confirmed by DNA testing as previously described. In addition, we excluded in confirmed SCA patients, those not in a steady state of the disease, those who received a transfusion within 4 months prior to blood sample collection, and those under Hydroxyurea treatment. For each participant, peripheral blood was collected in two EDTA-coated tubes of 4ml for hematologic tests and DNA extraction. Measured hematologic parameters were: hemoglobin (Hb), white blood cells (WBC), platelets (PLT), and reticulocytes (RETIC). The HbF level was quantified by capillary electrophoresis (Minicap, SEBIA, France). Genomic DNA was extracted from blood samples using the salting out method.

We used a candidate SNP approach exploring only SNPs with previously reported associations in the literature and public databases. Candidate SNPs associated with the baseline HbF level were retrieved from literature, searching PubMed using "fetal hemoglobin polymorphisms". We also extended the search to relevant papers from the reference lists of articles retrieved through Pubmed search. We also searched in GWAS Catalog using " fetal hemoglobin" AND "sickle cell anemia". The candidate SNPs associated with Hydroxyurea-induced HbF level were also retrieved through the same process using "fetal hemoglobin" AND "Hydroxyurea". We also extend the research to the PharmGKB database using "fetal hemoglobin hydroxyurea". Our SNP set contained the 4 polymorphic positions defining the five βs- globin haplotypes . In addition, 63 SNPs were included in the study of other genomic modifiers exerting pharmacogenomic variants. All these single nucleotide polymorphisms (SNPs) were genotyped at once using Molecular Inversion Probes (MIP) method. The α-thalassemia status was determined using the multiplex ligation-dependent probe amplification(MLPA) method.

The haplotype CAR was the predominant allele (95.92%), followed by BEN (1.67%), SEN (1.31%), and AI (0.37%). Atypical haplotypes accounted for 0.74 % of alleles. Six haplotypic genotypes were identified, including homozygous CAR (92.2%), homozygous BEN (0.37%), and double heterozygous CAR/BEN (2.59%), CAR/SEN (2.59%), CAR/AI (0.74%) and CAR/ATYP (1.48%). Less than 15% of patients within the CAR/CAR haplotype presented a severe clinical phenotype , representing 11% among children and 8% among adults.

Concerning the HbF baseline levels, the most significant association with HbF level was found for rs1427407 at BCL11A loci, with the highest level of HbF (13.8%) in homozygous TT genotype. Other SNPs within BCL11A locus (rs766432, rs11886868, rs4671393,rs7606173) and HBS1L-MYB intergenic region (rs28384513) were also significantly associated with HbF levels. No simultaneous association with HbF and hematologic variables was observed for any of these SNPs. Only the OR51B5/6 loci with rs5006884 and rs5024042 were associated with platelet count.

We conclude that βs haplotypes are not the major explanation for the clinical and biological variability in Congolese patients as patients still have variable hematologic and clinical profiles. This biological variability is explained by genomic loci that are identical across the sub-Saharan regions.

The coinheritance with α-thalassemia as well as βs haplotypes did not influence the HU response in term of HU-induced HbF level. A set of six SNPs rs 28384513 (HBS1L-MYB), rs 1480642 and rs 487278 (PDE7B), rs 9483947 (MAP3K5), rs 765587 (TOX) and rs 9319428 (FLT1) were significantly associated with HU-induced HbF levels. These set of SNPs may serve as predictors of HU response in our setting.

Implications of the study

We plan to implement the SCA molecular test we validated in this study in the molecular laboratory of the Center for Human Genetics at the University of Kinshasa. This will allow to realize the test for routine diagnostic of SCA. Actually, existing molecular laboratories in the country are dedicated essentially to the diagnostic of infectious diseases. Validation of reliable molecular testing in our setting will lead to the implementation of reference molecular laboratories dedicated to SCA diagnostic, helpful for the achievement of one of the ‘ PNLCD’ program goal.

Also in collaboration with the Authorities of our home university (University of Kinshasa), we plan to organize regular SCA screening among students. Sickle cell trait students will then benefit from genetic counselling. Applied at the national level, this preventive measure can result in the coming years in reducing the disease incidence.

The collaboration with the “PNLCD” program will be crucial for the deployment of the molecular test at the national level not only for routine diagnostic but also for newborn

screening. Systematic NBS has not been set up yet despite the creation of the “PNLCD” more than 20 years ago. The functioning of such national program is closely linked to the national health policy and depend on the budget allocated. If this structure has not worked properly for a long time , we see a relaunch of activities on the ground . Some of them involve the implementation of a national routine NBS program . Also with the growing interest of the government focused on the disease , we believe in an improvement of “PNLCD” activities .

This collaboration will be also helpful as a framework to present our results at a national level, and also to extend our research through GWAS. Through the national program, collaboration will be extended to physicians and local SCA patients and parents’ associations to raise awareness on the diagnostic and carrier testing and HU prescription. Physicians will acquire more knowledge about HU, the only approved molecule available in our local medical arsenal to date. Consequently, the HU prescription will increase, and why not having HU produced locally will reduce the price.

With the genetic data on Congolese SCA patients, collaboration including larger cohorts of patients can be established to verify the replication of our results on a larger scale.

In conclusion, this study has generated data on genetic context of sickle cell disease , particularly the sickle cell anemia , the most severe form and the main type encountered in our DR Congolese cohort. The obtained data will be helpful for the clinical management of SCA patients in the country and other countries. And also, the study provides preliminary data for further studies in the field of genetic context of SCA.

6. SAMENVATTING

De bloedarmoede van de sikkelcel (SCA) is de gemeenschappelijkste hemoglobinopathie wereldwijd. Ontwikkelingslanden als de Democratische Republiek Congo (DRC) hebben een zware last en de ziektegerelateerde sterfte blijft hoog, tot 50% tot 90% voor de leeftijd van vijf jaar. Factoren die bijdragen aan een hoge sterfte zijn onder andere vertraagde diagnose als gevolg van de schaarste van de juiste laboratoriuminfrastructuur, slechte toegang tot adequate medische zorg als gevolg van beperkte financiële middelen van patiënten, ontoereikende levering van state-of-the-art medische, heersende mystieke overtuigingen en het gebruik van traditionele kruidengeneeskunde. Ondanks het feit dat een enkele mutatie in het β globin-gen SCA veroorzaakt, wordt een duidelijke interindividuele variabiliteit waargenomen in klinische manifestaties en ernst, evenals de reactie op hydroxyurea (HU) bij patiënten met een plotselinge hartstilstand. Genetische factoren zijn een van de factoren die deze variabiliteit kunnen verklaren.

We hebben dit onderzoek uitgevoerd om de zorg voor SCA-patiënten te verbeteren, ondanks nauwkeurige, snelle en efficiënte diagnostische gegevens en een beter begrip van de invloed van genomische polymorfismen op het klinische en biologische profiel en op de respons op HU-behandeling bij Congolese SCA-patiënten.

De totale cohort bestond uit 338 patiënten waarvan 102 neonaten, 236 kinderen en volwassenen waren. Neonaten werden geworven bij 2 materiniteiten van Kinshasa, d.w.z. het moederschap van Binza en het moederschap van Kingsani. Rekrutering voor kinderen en volwassenen werd uitgevoerd in het Saint Luc Hospital in Kisantu (centrale provincie Kongo) en het “Centre de Médecine Mixte et d’Anémie SS (CMMASS)”, het St Crispin Hospital en de Universitaire Ziekenhuizen van de Universiteit van Kinshasa in Kinshasa (de hoofdstad) . Alle sites bevinden zich in de DRCongo.

In het eerste deel wilden we een op DNA gebaseerde SCA-test implementeren die geschikt zou zijn voor een beperkte resource setting en beperkingen van op hemoglobine (Hb) gebaseerde tests overwonnen die vaak worden gebruikt in de huidige zorgstandaard. De SCA Hb-gebaseerde tests zijn niet betrouwbaar, vooral in neonatale en post-transfusieperioden.

Er werden drie soorten biologische monsters genomen: Veneus bloed, navelstrengbloed (UCB) en buccaal wattenstaafje. Een druppel bloed (ongeveer 50 µl bloed) (veneus en UCB) werd

afgezet op een cirkel van een FTA-elutekaart (Whatman WB120206).Buccal-wattenstaafjes verzameld met een zelfgemaakt steriel wattenstaafje, werden ook gedurende 10 seconden op een cirkel van FTA-Elute kaart gewreven om het monster over te brengen. Na de DNA-extractie met de handmatige zoutmethode (SO) werd SCA moleculair onderzoek (RFLP) uitgevoerd om het gesikkelde gemuteerde allel (SO-RFLP-methode) te identificeren. We hebben het SO-RFLP-protocol aangepast door het geëxtraheerde DNA-monster te vervangen door een 1.2 mm pons FTA Elute-kaarten, waardoor een dure DNA-extractiestap wordt vermeden.

Er was een perfecte overeenkomst tussen beide protocollen, d.w.z. SO-RFLP en FTA-RFLP. Door de DNA-extractiestap over te slaan, verlaagt de FTA-RFLP de kosten met de helft (van 11.5 € tot 6.5 €) en de turnaround-tijd aanzienlijk. FTA Elute-kaarten met bloedvlekken werden bewaard in een schone kast bij kamertemperatuur (RT) en na 2 jaar opnieuw getest. We hebben precies dezelfde resultaten behaald. De waarde en het nut van DNA-gebaseerde test op DBS in de post-transfusieperiode is een lopend onderzoek waarbij FTA-RFLP vóór en 1 uur na een bloedtransfusie wordt uitgevoerd. Uit onze voorlopige resultaten (bij 6 patiënten) blijkt dat de opbrengst van FTA-RFLP ongewijzigd blijft na transfusie van verpakte cellen, in tegenstelling tot de twee op HB gebaseerde tests.

In het tweede deel concentreerden we ons op de evaluatie van genetische factoren die betrokken zijn bij het moduleren van de klinische en biologische expressie van SCA, evenals de farmacogenetica van HU.

We hebben eerst de distributie van βS-globin haplotypes en de correlatie tussen haplotypes en baseline hematologische en klinische profielen beoordeeld in onze cohort van patiënten die zijn gerekruteerd in het Saint Luc Hospital in Kisantu en het “Centre de Médecine Mixte et d’Anémie SS (CMMASS)”. Patiënten werden geclassificeerd volgens de klinische fenotypenscore zoals eerder beschreven.

Vervolgens evalueerden we een reeks single nucleotide polymorfismen (SNPs), voor hun mogelijke associatie met HbF-niveaus en de hematologische parameters die betrokken zijn bij de ernst van de ziekte.

Tot slot hebben we de bijdrage bepaald van α-thalassemie-coerfenis, de βs-haplotypes en kandidaat-SNPs in HU-respons in de reeks patiënten binnen onze initiële cohort , die deel uitmaakten van een klinische HU-studie.

Voordat het in dit deel van het onderzoek werd opgenomen, moest de homozygoot status van SCA worden bevestigd door DNA-tests, zoals eerder beschreven. Bovendien hebben we patiënten met een bevestigde plotselinge hartstilstand uitgesloten, patiënten die niet in een stabiele toestand verkeren, patiënten die binnen 4 maanden vóór de monsterafname een transfusie hebben ondergaan, en patiënten met een hydroxyureumbehandeling. Voor elke deelnemer werd perifeer bloed verzameld in twee buizen met EDTA-coating van 4 ml voor hematologische tests en DNA-extractie. De gemeten hematologische parameters waren: Hemoglobine (HB), witte bloedcellen (WBC), bloedplaatjes (PLT) en reticulocyten (RETIC). Het HBF-niveau werd gekwantificeerd door capillaire elektroforese (Minicap, SEBIA, Frankrijk). Genomisch DNA werd geëxtraheerd uit bloedmonsters met behulp van de zoutmethode. We gebruikten een kandidaat SNP-benadering die alleen SNP's onderzocht met eerder gerapporteerde associaties in de literatuur en openbare databanken. Kandidaat-SNP's die verband houden met het HBF-basisniveau zijn uit de literatuur opgehaald, waarbij PubMed werd doorzocht met “foetale hemoglobine-polymorfismen”. We breidden de zoekopdracht ook uit naar relevante artikelen uit de referentielijsten van artikelen die via Pubmed-zoekacties zijn opgehaald. We hebben ook gezocht in de GWAS-catalogus met behulp van “foetale hemoglobine” EN “sikkelcelanemie”. De kandidaat-SNP's die gerelateerd zijn aan het door Hydroxyurea geïnduceerde HBF-niveau werden ook via hetzelfde proces opgehaald met behulp van “foetale hemoglobine” EN “Hydroxyureum”. We breiden het onderzoek ook uit naar de PharmGKB-database met behulp van “foetale hemoglobine hydroxyureum”. Onze SNP set bevatte de 4 polymorfe posities die de vijf βs-globin haplotypes definiëren. Bovendien werden 63 SNPs opgenomen in het onderzoek naar andere genomische modifiers die farmacogenomische varianten uitvoeren. Al deze single nucleotide polymorfismen (SNPs) werden onmiddellijk genotyped met behulp van de Moleculaire Inversion sondes (MIP) methode. De α-thalassemia-status werd bepaald met behulp van de multiplex ligation-dependent probe amplification (MLPA)-methode.

De haplotype- CAR was de dominante allel (95.92%), gevolgd door BEN (1.67%), SEN (1.31%) en AI (0.37%). Atypische haplotypes vertegenwoordigden 0.74 % van alle allelen. Er werden zes haplotypische genotypen geïdentificeerd, waaronder homozygoot (92.2%),

homozygoot BEN (0.37%) en dubbel heterozygoot CAR/BEN (2.59%), CAR/SEN (2.59%), CAR/AI (0.74%) en CAR/ATYP (1.48%). Minder dan 15% van de patiënten binnen de CAR/CAR haplotype presenteerde een ernstig klinisch fenotype , dat 11% onder kinderen en 8% onder volwassenen vertegenwoordigde.

Wat de HbF-basisniveaus betreft, werd de meest significante associatie met het HbF-niveau gevonden voor rs1427407 op BCL11A loci, met het hoogste niveau van HbF (13.8%) in homozygoot TT-genotype. Andere SNP's binnen BCL11A locus (rs766432, rs11886868, rs4671393,rs7606173) en HBS1L-MYB intergene regio (rs28384513) werden ook aanzienlijk geassocieerd met HBF-niveaus. Voor geen van deze SNP's werd gelijktijdige associatie met HbF- en hematologische variabelen waargenomen. Alleen de OR51B5/6 loci met rs5006884 en rs5024042 werden geassocieerd met het aantal plaatjes.

We concluderen dat βs haplotypes niet de belangrijkste verklaring zijn voor de klinische en biologische variabiliteit bij Congolese patiënten, omdat patiënten nog steeds variabele hematologische en klinische profielen hebben. Deze biologische variabiliteit wordt verklaard door genomische loci die identiek zijn in de gebieden ten zuiden van de Sahara.

De samenname met α-thalassemie en βs-haplotypes heeft de HU-respons op termijn van het door HU geïnduceerde HBF-niveau niet beïnvloed. Een set van zes SNPs rs 28384513 (HBS1L-MYB), rs 1480642 en rs 487278 (PDE7B), rs 9483947 (MAP3K5), rs 765587 (TOX) en rs 9319428 (FLT1) werden aanzienlijk geassocieerd met door HU geïnduceerde HBF-niveaus. Deze set SNP's kan als voorspellers van HU-respons in onze setting dienen.

7. ACKNOWLEDGEMENTS, PERSONAL CONTRIBUTION and CONFLICT OF INTEREST STATEMENTS

1. Scientific acknowledgement

I thank Professor Chris Van Geet (Department of Cardiovascular Sciences and Pediatrics [Hemato-oncology] , KU Leuven and UZ Leuven) and Professor Koenraad Devriendt (Department of Human Genetics, UZ Leuven) for their involvement in the design of this project and their supervision.

I thank my colleagues Dr Gloire Mbayabo and Dr Paul Lumbala for patients recruitment at respectively Kisantu St Luc Hospital and "Centre de Médecine Mixte et d'Anémie SS" (CMMASS) at Kinshasa ; and also for conducting the Hydroxyurea clinical trial.

I would like to thank Dr Valerie Race (Center of Human Genetics, UZ leuven) for her supervision and guidance in my first contact with the molecular biology field , since the beginning of this research. At the same time, I would like to thank Bruno Vankeirsbilck and the entire team of the Laboratory for Molecular Diagnosis of UZ Leuven for their technical support in the laboratory work all along my research.

Special thanks to Cathy Songo and Chloé Musuamba for assistance in DNA extraction , samples aliquot preparation , evaluation of SCA testing reproducibility and capillary electrophoresis testing.

I express my gratitude to Prof. Isabelle Cleynen (department of Human Genetics, KU Leuven) and Dr Nono Mvuama (Kinshasa School of Public Health, University of Kinshasa) for their important contribution in data statistical analysis and interpretation.

I thank Stephanie Deman (Genomics Core KULeuven – UZLeuven) for SNPs genotyping with the MIP technique and Erika souche (department of Human Genetics, KULeuven) for bioinformatic analysis of sequencing data.

2. Personal contribution

I contributed in the design of the protocol of the study and prepared the thesis manuscript.

Specifically for the first part of the work (SCA DNA testing): I recruited patients at both maternities , “CMMASS” , St Crispin and University Hospital attached to the University of Kinshasa .

I collected samples (umbilical cord blood, , buccal swab and peripherical blood). I spotted samples on FTA cards. I performed both protocol for SCA testing: the SO-RFLP and the FTA-RFLP. I performed the SCA RDT after transfusion. I contributed to data interpretation.

For the second part (Genetic factors associated with HbF levels and HU response): I selected candidate SNPs according to literature. I performed the SCA molecular testing and the α-thalassemia MLPA testing. I manually classified the βglobin haplotypes based on the SNPs genotype. I contributed to data interpretation.

3. Conflict of interest statement

This work was funded by VLIR UOS scholarship under the grant “Initiatives Sud et projets TEAM 2015”.

There is no conflict of interest related to this work.

8. PROFESSIONAL CAREER MAMY NGOLE

Curriculum vitae

IDENTITY

FAMILY NAME	: NGOLE ZITA
FIRST NAME	: MAMY
GENDER	: FEMALE
BIRTHPLACE AND DATE	: KINSHASA, APRIL 23rd 1975
MARITAL STATUS	: MARRIED
NATIONALITY	: DR CONGOLESE
PROFESSION	: BIOLOGIST MEDICAL DOCTOR
E--MAIL ADDRESS	: mmyngole@gmail.com mamy.ngole@unikin.ac.cd

EDUCATION

2016 - 2017	: Predoctorat in Biomedical Sciences (KULeuven)
2007 - 2012	: Specialization in Medical Biology (University of Kinshasa) Grade: Great distinction
1992 - 2003	: Medicine (University Simon Kimbangu) Grade: Distinction
1986 - 1992	: Secondary school at CS Les Loupiots Grade: Distinction
1980 - 1986	: Primary school at CS Les Loupiots Grade: Great distinction

POSTGRADUATE TRAINING

October 2017 to date	: PhD training in Genetics and genomics, KU Leuven, Belgium
December 2022	: Course on NGS analysis for monogenic diseases, Rabat,Morocco
July 2021 to June 2022	: Online training on "Managing Innovation for Sustainable Health (MISH)" co-organized by Karolinska Institute (Sweden), Makerere University (Uganda), Benadir University (Somalia), Public School of University of Kinshasa (DRC)
May 2022	: Residential workshop on MISH program at Kampala, Uganda
September 2019	: Training on Nanopore sequencer, Kinshasa, DRCongo
July 2019	: Summer school, Human genome variation: Applications in human medicine and genetic identification,Ifrane, Morocco
May 2015	: Training on Infectious blood safety at Pasteur Institute of Paris, France
March and June 2015	: University Diploma on sickle cell syndromes at the Medecine, Faculty of Paris-Est Créteil University (UPEC) 1st and 2nd part, France
February 2015	: Training on Management of Public Health Laboratory Systems at Stellenboch University Medicine and Health Sciences, Tygeberg Campus,South Afric
April 2014	: Workshop SLMTA (Strengthening Laboratory Management Towards Accreditation) organized by WHO- AFRO and CDC, at Kinshasa, DRCongo
August 2010	: Training on qualitative research at « Institut National de Recherche Biomédicale » (INRB) organized by Anvers Tropical Medecine Institute, at Kinshasa, DRCongo
November 2009	: Training in "Infectious diseases: updates in Laboratory diagnosis" organized by Suez Canal University, Ismailia, Egypt
May 2009	: Training in clinical genetic organized by the "Programme de Migration pour le développement en Afrique » (MIDA) at Kinshasa, DRCongo

ATTENDED CONGRESSES AND CONFERENCES

December 2022	: 14th Meeting of African Society of Human Genetics and the 2nd International Congress of Société Marocaine de Génétique Humaine, Rabat, Morocco
September 2022	: Workshop to update the sickle cell management protocol and develop the training module for healthcare providers organized at Kinshasa by the DRC Ministry of Health
March 2021	: 3rd Symposium on Birth Defects and Rare Diseases Ghana Initiative, online attendance
December 2017	: Symposium of the Belgian Society of Human Genetics organized at Ghent University Hospital in Belgium
May 2014	: 5th Symposium of "Réseau d'Etudes sur la Drépanocytose en Afrique Centrale" (REDAC), at Kinshasa
May 2013	: 1st Congress of Congolese Human Genetic Society at Kinshasa

ORAL PRESENTATIONS

- Distribution of βs haplotypes among DR Congolese sickle cell anemia patients (Summer school, Human genome variation, Ifrane, Morocco, 2019)

- Cell blood count in sickle cell pediatric patients: a multicentric study (REDAC, Kinshasa, 2014)
- Profile of infectious markers in blood donors at University of Kinshasa Clinic from 2010 to 2013 (workshop for TCTP Egypt-Japan, Cairo, Egypt, 2014)

POSTER PRESENTATION

Correlation of βS globin haplotypes with hematological parameters and clinical severity score in DRCongolese sickle cell anemia patients (14th Meeting of African Society of Human Genetics and the 2nd International Congress of Société Marocaine de Génétique Humaine, Rabat, Morocco, 2022)

SCIENTIFIC SOCIETIES

2022 : Member of African Society of Human Genetics

2017 : Member of Congolese Society of Clinical Biology

LIST OF PUBLICATIONS

Publications included in this thesis

1. **Ngole M**, Race V, Mbayabo G, Lumbala P, Songo C, Lukusa PT, Devriendt K, Matthijs G, Lumaka A. DNA testing for sickle cell anemia in Africa: Implementation choices for the Democratic Republic of Congo. J Clin Lab Anal. 2022 May;36(5):e24398. doi: 10.1002/jcla.24398. Epub 2022 Apr 11. PMID: 35405024

Other Sickle Cell Anemia-related publications

1. Lumbala PK, Mbayabo G, **Ngole MN**, Lumaka A, Race V, Matthijs G, Van Geet C, Lukusa PT, Devriendt K, Mikobi TM. Clinical and laboratory characterization of adult sickle cell anemia patients in Kinshasa.PLoS One. 2022 Dec 16;17(12):e0278478. doi: 10.1371/journal.pone.0278478. eCollection 2022. PMID: 36525434

2. Mbayabo G, Lumbala Kabuyi P, **Ngole M**, Lumaka A, Race V, Maisin D, Gruson D, Matthijs G, Minga TM, Devriendt K, Van Geet C, Tshilobo PL. Value of DNA testing in the diagnosis of sickle-cell anemia in childhood in an environment with a high prevalence of other causes of anemia. J Clin Lab Anal. 2022 Aug;36(8):e24593. doi: 10.1002/jcla.24593. Epub 2022 Jul 12. PMID: 35819088
3. Tshilolo L, **Zita MN**, Ngiyulu R, Kayembe Nzongola D. Le statut martial chez soixante-douze drépanocytaires homozygotes congolais [Iron status in 72 Congolese

patients with sickle cell anemia]. Med Sante Trop. 2016 Jan-Mar;26(1):83-7. French. doi: 10.1684/mst.2016.0535. PMID: 26987842

Miscellaneous publications

1. Mavinga Laetitia M, Kakase Veronique, **Ngole Mamy Z**, Songo Cathy M, Lumaka Aimé, Race V, Lukusa Prosper T, Devriendt K. Molecular genetic characterization of Congolese patients with oculocutaneous albinism. Eur J Med Genet. 2022 Nov;65(11):104611. doi: 10.1016/j.ejmg.2022.104611. Epub 2022 Sep 16. PMID: 36116698.

2. Makay P, Mubungu G, Mupuala A, Bluske K, Brown C, Schmidt SA, **Ngole M**, Fuanani P, Perry DL, Lukusa P, Devriendt K, Taft RJ, Lumaka A. PERCHING syndrome: Clinical presentation in the first African patient confirmed by clinical whole genome sequencing. Am J Med Genet A. 2022 Sep;188(9):2825-2831. doi: 10.1002/ajmg.a.62855. Epub 2022 Jun 7. PMID: 35670385.

3. B. D. Tshibuela, N. N. Kayembe, M. J. Muwonga, N. K. Nganga, **N. Ngole**, J. P. Elongi. Relationship between uricemia and other biochemical markers with the materno-fetal complications during pre-eclampsia. Open Journal of Obstetrics and Gynecology, 2017 : 1255-1261. https://doi.org/10.4236/ojog.2017.713128

ACKNOWLEDGEMENTS

First of all, may all the glory return to God, the King of times and circumstances.

I would like to sincerely thank all the jury members (Prof. Kathleen Freson, Prof. Marc Jacquemin, Prof. Jean-Marie Kayembe, Prof. Solomon Ofori-Acquah and Prof. François Boemer) for improving this manuscript through their relevant comments.

I am grateful to the TEAM project's Managers for giving me the opportunity to work on a project based on a disease that is an important challenge for my country. I think about Prof. Dr. Chris Van Geet, Prof. Dr. Koenraad Devriendt, Prof. Dr. Prosper Lukusa, Prof. Dr. Gert Matthijs, Prof. Dr. Aimé Lumaka, Prof. Dr. Tite Mikobi and Dr. Valerie Race.

I am grateful to the patients (as well as their parents or their guardians), for their participation in our study. I hope that the results of this work will help to improve their care.

To my colleagues, Dr. Gloire Mbayabo and Dr. Paul Lumbala, many thank for this scientific journey we have travelled together, as a real team.

To Prof. Dr. Kayembe Nzongola-Nkasu, you have accepted me at the department and you have trusted me since then, thank you for your continued support.

I thank all people who worked with me on the floor: Marie Eboma, Cathy Songo and Chloe Muswamba.

To Pitchou Sundi, Dr Dophie Tshibuela, Dr Dahlia Pambu, Dr Bizette Bizeti and Dr Ritha Nyembu, thank you for your presence and support at every moment.

To Prof Dr Gerrye Mubungu, Dr Prince Makay, Dr Aline Engo, thank so much for your support.

Dear Professor Gert Matthjis, it is a great privilege to learn from you. I worked with you in an atmosphere that you certainly wanted to be rigorous but above all friendly. I hope to always honor you.

Dear Professor Prosper Lukusa, receive my deepest gratitude for accepting me in the local genetic team and for your guidance. You were always there for each of us, as a father with his children.

Dear Professor Aimé Lumaka, you have always helped me stay the course. The journey was not easy every time, but you have accompanied me. There are always several things to learn from you. I am proud to be your second "scientific child". This is just the beginning; the journey will continue. Thank you so much, and special thanks to Bijou Lumaka who made my stay in Leuven so special!

Dear Professor Koenraad Devriendt, it was an honor to be trained and accompanied by you. You always had many ideas for our work to adapt with local realities on the field. Thank you so much.

To Pr Fons Verdonck, many thanks for the financial support through the scholarship by the KU Leuven Alumni

To Professors Jeremie Muwonga, Joseph Bodi and Gerrye Mubungu, thank you for your encouragement.

I express my deep gratitude to the academic authorities of the University of Kinshasa: The Rector Pr. Dr. Jean-Marie Kayembe, Professors Roger Mbungu, Eric Kamangu and Andy Mbangama and all the staff of the Faculty of Medicine; to Professors Jean-Jacques Muyembe, Steve Ahuka, Octavie Lunguya, Sheila Makiala, Placide Mbala, Mireille Nganga, Edith Nkwembe , Alain chabo, Blaise Sumbu ; to Drs Gustave Ilunga, Jean-Pierre Mufuta, Marie-José Kabedi, Eddy Sokolua, Blanchard Malenga, Didier Nsonso, Gabriel Lukusa, Eric Mukenge, Frank Nzengu, Gaby Lukusa, José Kadima, Mamy Mbelu, Patience Maindo and all staff of the Medical Biology Department (University of Kinshasa).

To Annemie Puttemans and Veerle Mattheus, find the expression of my gratitude for your administrative support throughout these doctoral years.

To my dear mother, Esperance Usuku, I miss you too much Mama, your encouragement has enabled me to start with this thesis and I wish you had seen the end of it. But unfortunately, you left me on the way! Anyway, even if physically you are no longer with me, I dedicate this thesis to you because it is the fruit of so many sacrifices to educate me and to inculcate in me moral, scientific, and Christian values. Thank you, Mama!

To my dear father, Dr Robert Azode Nemchi, I miss you too much Papa. I had to change the words previously written as you are not anymore with me. Thank you for your love, your

encouragement. I will always make you proud of the «Igbo» blood in my veins. Thank you, Papa!

To my husband, my best friend, my all-time partner, Freddy Lokossa, this thesis is also yours. You have made so many sacrifices for the achievement of this thesis. You spent sleepless nights with me when I had to read or write, you consented to me being away from you for several months, and you was there to comfort me every time.

To my brother Josué Usuku, thank you for your love and support .

To Maman Regine, Ma Philo, Dr Chris Azode, thanks for your support!

I am grateful to all those who in one way or another have supported me throughout my PhD. I can’t name all of them. But from the bottom of my heart, thank you so much.