

Study of Genetic Modifiers of Fetal Hemoglobin and Mechanisms of Hydroxyurea-induced γ -globin Expression in Sickle Cell Disease

By: Gift Dineo Pule
PLXDIN003



SUBMITTED TO THE UNIVERSITY OF CAPE TOWN
In fulfilment of the requirements of the degree

PhD in Human Genetics
Faculty of Health Sciences,
Department of Pathology, Division of Human Genetics

Supervisor: Professor Ambroise Wonkam
Co-supervisors: Dr Shaheen Mowla and Professor Nicolas Novitzky

Date of submission:
02 August 2016

The copyright of this thesis vests in the author. No quotation from it or information derived from it is to be published without full acknowledgement of the source. The thesis is to be used for private study or non-commercial research purposes only.

Published by the University of Cape Town (UCT) in terms of the non-exclusive license granted to UCT by the author.

Declaration

I, Gift Pule, hereby declare that the work on which this dissertation/thesis is based is my original work (except where acknowledgements indicate otherwise) and that neither the whole work nor any part of it has been, is being, or is to be submitted for another degree in this or any other university.

I empower the university to reproduce, for the purpose of research, either the whole or any portion of the contents, in any manner whatsoever.

Signature:

Date: 02 August 2016

Acknowledgements

I would like to express my utmost gratitude,

To my parents, Joseph and Margaret, for their unconditional love and support from elementary school to this very moment.

To my siblings, Omphile; Mango and Bosa, for their encouragement and confidence in me.

To my girlfriend, Cindy, for the undying comfort and support throughout the ups and downs of this journey.

To those who generously funded this research: The National Research Foundation; Oppenheimer Memorial Trust and FirstRand Laurie Dippenaar Scholarship, for the financial support without which this would not be possible.

To the Mellon Mays Foundation and family, specifically Kathy and Gideon, I will be eternally grateful for your support and motivation.

To my co-supervisors, Shaheen and Nicolas, for the encouragement and support.

To my supervisor, Ambroise, for believing in and supporting my ambitions. I will always be indebted to you.

To my colleagues in the Divisions of Human Genetics and Hematology; Mowbray Maternity and Groote Schuur Hospitals and international collaborators, for their input and support of this work.

Table of Contents

Declaration	iii
Acknowledgements	iv
Preface	vii
List of publications included in this thesis.....	x
Summary	xii
Chapter 1. Introduction	1
Chapter 2. Aims and Objectives	5
Chapter 3: Results: Original publications	7
Chapter 3.1. Study of additional fetal hemoglobin-promoting loci	7
3.1.1. Bitoungui VL, Pule DG, Hanchard N, Ngogang J and Wonkam A. 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 Journal of Integrative Biology. 2015 19(3): 171-9.....	7
3.1.2. Pule DG, Ngo Bitoungui VJ, Chemegni BC, Kengne AP, Antonarakis S and Wonkam A. Association between variants at BCL11A erythroid-specific enhancer and fetal hemoglobin levels among Sickle Cell Patients in Cameroon: Implications for future therapeutic interventions. OMICS Journal of Integrative Biology. 2015 19(10): 627-31.....	19
3.1.3. Pule DG, Bitoungui VJN, Chemegni BC, Kengne AP and Wonkam A. Study of novel variants associated with fetal hemoglobin in Sardinians and Tanzanians among Sickle Cell Disease patients from Cameroon. Submitted to Hemoglobin (in press) 2016.	25
3.1.4. Pule DG, Bitoungui VJN, Chemegni BC, Kengne AP and Wonkam A. SAR1a promoter polymorphisms are not associated with fetal hemoglobin in Cameroon SCD patients. Submitted to BMC Research Notes (revised manuscript submitted) 2016..	30
Chapter 3.2. Mechanisms of hydroxyurea-induced γ -globin expression in Sickle Cell Disease.....	45
3.2.1. Pule GD, Mowla S, Novitzky N, Wiysonge C and Wonkam A. Systematic review of known mechanisms of hydroxyurea-induced fetal hemoglobin for treatment of Sickle Cell Disease. Expert Review of Hematology. 2015 8(5): 669-79.....	45
3.2.2. Pule GD, Mowla S, Novitzky N, Wonkam A. Hydroxyurea down-regulates BCL11A, KLF-1 and MYB through miRNA-mediated actions to induce γ -globin expression: implications for new therapeutic approaches of sickle cell disease. Clinical Translational Medicine. 2016 Dec;5(1):15.....	58
3.2.3. Pule GD and Wonkam A. Treatment of Sickle Cell Disease in Africa: should we invest in haematopoietic stem cell transplantation? Pan African Medical Journal. 2014 18:46.....	71

Chapter 3.3. Southern African perspective of Sickle Cell Disease	77
3.3.1. Pule GD, Mnika K, Jourbert M, Mowla S, Novitzky N, Wonkam A. Increasing burden of adult patients living with Sickle Cell Disease in Cape Town. Submitted to South African Medical Journal (in review) 2016.....	77
3.3.2. Pule GD, Mnika K, Chimusa E, Mhandire K, Kampira E, Dandara C, Wonkam A. Beta-globin gene haplotypes and tag variants for malaria resistance among black southern African populations. Submitted to Global Health, Epidemiology and Genomics (in review) 2016.....	100
Chapter 4. Academic Discussion and Conclusions	133
4.1. Genetics of fetal hemoglobin	135
4.2. Alternative treatment approaches to Sickle Cell Disease	138
4.3. Sickle Cell Disease in Southern Africa	143
Conclusions and perspectives.....	149
References (Introduction and Academic Discussion chapters)	151
Annexes.....	163
Approval letter to authorise inclusion of original publications in the thesis.....	163
Published abstract: D.G. Pule, S. Mowla, N. Novitzky, A. Wonkam. Hydroxyurea induces γ -globin expression through microRNAs-mediated actions in human hematopoietic and K562 erythroleukemia cells. American Society of Human Genetics 2015 Congress, Baltimore, USA.....	164
Turnitin® Originality Report	165

Preface

Sickle Cell Disease (SCD) is a worldwide health problem with a burden comparable to that of communicable diseases and other major global diseases such as diabetes and hypertension. Furthermore, over 3/4 of this burden is in sub-Saharan Africa, with limited studies investigating the genetic polymorphisms that affect the severity, hematologic phenotypes, disease course and less so, the mechanism of action of the only approved medication for treatment of SCD, namely hydroxyurea (HU). It was therefore advisable to present the current PhD thesis in this format, “with inclusion of publications”, mostly reported or submitted in international peer-reviewed journals in order to populate this sparse niche in SCD research. Moreover, the multiple peer reviews by international researchers was also intended to improve the quality of the final outcomes in the thesis, which will hopefully, contributed significantly to scholarship in this field. In addition, the inclusion of published work also aimed to: (i) set a path that would bolster the academic career of the candidate and enhance his candidacy for the financial support received from the NRF, Oppenheimer Memorial Trust and FirstRand Laurie Dippenaar Scholarship; (ii) enable participation in and attendance at academic conferences and training workshops: the candidate effectively attended a total of 16 specialised meetings in several cities including Cape Town, Johannesburg, Tunis, Kyoto, London, Washington DC and Chicago; and lastly (iii) the published works would form integral parts of a cohesive body of research that will fit the encouragement policies of the Faculty of Health Sciences (UCT) to publish the findings of research projects, particularly at the doctoral level, in order to continue to disseminate the knowledge generated within the Faculty and enhance the profile of the institution.

With the advice and guidance of the supervisors, the candidate’s contributions to the included publications ranged from conceiving and designing the experiments; recruiting and sampling of biological material from participants; execution of all experiments; analysis of experimental results; drafting the full manuscripts and incorporating revisions from co-authors and journal reviewers. The candidate’s contribution to this project and thesis can be detailed as follows:

Conception of research:

The study and all experiments were conceived by the candidate in conjunction with the supervisors, primary supervisor Professor Ambroise Wonkam and jointly refined with the advice of Dr Shaheen Mowla and Professor Nicolas Novitzky (co-supervisors).

Data collection:

All necessary recruitment of participants, obtaining informed consent/assent and sampling biological material in this project was executed in its entirety by the candidate. All the required DNA, RNA and protein isolation and cataloguing of patient material was done by candidate. The clinical information of participants was also retrospectively collected from hospital records and patient files by the candidate.

Experimentation and analysis:

All experiments; molecular analysis by genotyping and sequencing; and functional analysis by tissue culture and expression analysis of messenger transcript, microRNA and protein was executed in full by the candidate. The primary analysis of all data was performed by the candidate and when additional statistical analysis was required, a collaborator was included by the primary supervisor and their contribution to the work clearly indicated in each publication.

Publications:

Synthesis of all the works and drafting of all manuscripts of the publications included in this thesis were executed in full by the candidate, after which revisions from all co-authors were similarly incorporated before submission to the journal by the primary supervisor. After review, all reviewer comments were addressed by the candidate in conjunction with the primary supervisor. The detailed role of the candidate in the publications included is clearly defined in each chapter.

This project and thesis is presented with three cohesive elements; (i) genetic modifiers of fetal hemoglobin; (ii) mechanisms of HU-induced γ -globin expression and lastly, (iii) the southern African perspective of SCD.

We have met all requirements and abided by the UCT's Doctoral Degrees Board, under Rules GP6.7 as follows:

- (i) The candidate's proposal to include publications in the current thesis was approved by the UCT Faculty of Health Sciences Doctoral Degrees Board.
- (ii) The thesis contains an adequate Introduction; a chapter on the Aims and Objectives; a comprehensive Academic Discussion of the results as a whole, forming the basis of the Conclusions and Perspectives drawn from this research.
- (iii) Each Results chapter with publications included, is preceded by a synopsis of how the publications directly tie to the aims and objectives of the project, as well as to the thesis as a whole.
- (iv) All included publications were written and published during the candidate's tenure as a PhD student.

The candidate,

Gift Dineo Pule

List of publications included in the thesis

1. Bitoungui VL, **Pule DG**, Hanchard N, Ngogang J and Wonkam A. 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 Journal of Integrative Biology*. 2015 19(3): 171-9.
2. **Pule DG**, Ngo Bitoungui VJ, Chemegni BC, Kengne AP, Antonarakis S and Wonkam A. Association between variants at *BCL11A* erythroid-specific enhancer and fetal hemoglobin levels among Sickle Cell Patients in Cameroon: Implications for future therapeutic interventions. *OMICS Journal of Integrative Biology*. 2015 19(10): 627-31.
3. **Pule DG**, Bitoungui VJN, Chemegni BC, Kengne AP and Wonkam A. Study of novel variants associated with fetal hemoglobin in Sardinians and Tanzanians among Sickle Cell Disease patients from Cameroon. Submitted to *Hemoglobin (in press)* 2016.
4. **Pule DG**, Bitoungui VJN, Chemegni BC, Kengne AP and Wonkam A. SAR1a promoter polymorphisms are not associated with fetal hemoglobin in Cameroon SCD patients. Submitted to *BMC Research Notes (revised manuscript submitted)* 2016.
5. **Pule GD**, Mowla S, Novitzky N, Wiysonge C and Wonkam A. Systematic review of known mechanisms of hydroxyurea-induced fetal hemoglobin for treatment of Sickle Cell Disease. *Expert Review of Hematology*. 2015 8(5): 669-79.
6. **Pule GD**, Mowla S, Novitzky N, Wonkam A. Hydroxyurea down-regulates *BCL11A*, *KLF-1* and *MYB* through miRNA-mediated actions to induce γ -globin expression: implications for new therapeutic approaches of sickle cell disease. *Clinical Translational Medicine*. 2016 Dec;5(1):15.
7. **Pule GD** and Wonkam A. Treatment of Sickle Cell Disease in Africa: should we invest in haematopoietic stem cell transplantation? *Pan African Medical Journal*. 2014 18:46.

8. **Pule GD**, Mnika K, Jourbert M, Mowla S, Novitzky N, Wonkam A. Increasing burden of adult patients living with Sickle Cell Disease in Cape Town. Submitted to *South African Medical Journal (in final editing, #10849)* 2016.
9. **Pule GD**, Mnika K, Chimusa E, Mhandire K, Kampira E, Dandara C, Wonkam A. Beta-globin gene haplotypes and tag variants for malaria resistance among black southern African populations. Submitted to *Global Health, Epidemiology and Genomics (revised manuscript submitted)* 2016.

Abstract

Sickle Cell Disease (SCD) is a growing global problem with firm roots in sub-Saharan Africa (SSA) representing over 3/4 of the global burden of the disease. The prevalence of the sickle mutation (HbS) in SSA has been amplified by the partial resistance to *Plasmodium falciparum* malaria, which is endemic along tropical equatorial Africa. Several genetic variants have since been associated with fetal hemoglobin (HbF), the disease-ameliorating globin protein, including variants at three principal loci; *BCL11A*, *HBS1L-MYB* intergenic polymorphisms (*HMIP1/2*) and the β -globin gene cluster, which together account for 10 – 20% HbF variance in SCD patients. Similarly, numerous signalling pathways have been implicated in the regulation of γ -globin expression, however, a complete understanding of the regulation of HbF remains elusive.

The overall aims of this project were: 1a) to investigate the known variants in key HbF-promoting loci such as *BCL11A* erythroid-specific enhancer, *BCL11A*, *HBS1L-MYB* intergenic polymorphism (*HMIP1/2*), the β -globin gene cluster, as well as the influence of the co-inheritance of 3.7kb alpha globin gene deletion in a cohort of SCD patients from Cameroon; and 1b) to validate novel HbF-promoting loci reported in 2 genome-wide association studies (GWAS) carried out in a population of Sardinians (Italy) and SCD patients from Tanzania and explore the influence of known promoter variants in *SAR1* associated with HbF in African American patients amongst Cameroonian SCD patients; 2) to investigate the molecular mechanisms of hydroxyurea (HU)-induced production of HbF using a primary erythroid cell model from hematopoietic stem cells (HSCs) derived from umbilical cord blood and lastly, 3) to investigate the prevalence of SCD-related polymorphisms; β -globin gene haplotype, HbS mutation and malaria-resistance variants in 3 SCD-unaffected (HbAA) cohorts from South Africa, Zimbabwe and Malawi.

The studies revealed the following findings:

Regarding the exploration of HbF-promoting loci: 1) two erythroid-specific enhancer polymorphisms in *BCL11A* (rs1427407 and rs7606173) were shown to account for 8% and 6.2% HbF variance in Cameroon SCD patients, respectively; 2) there is some

evidence from the β -globin gene haplotype studies to support a provocative hypothesis of a single origin of the HbS mutation in the East Africa/Sudan region, after the original study of a Cameroonian cohort, coupled with a systematic review of the global distribution of the β -globin haplotype; 3) no associations among Cameroonian patients were observed between the 18 novel and suggestive SNPs and HbF levels from a Tanzanian GWAS SCD cohort and the whole genome sequencing of a population in Sardinia; 4) Furthermore, no association was found between HbF levels and SAR1a promoter variants previously associated with HbF in African American patients.

Regarding the mechanism of action of HU: 1) A systematic review of the known mechanisms of HU-induced HbF revealed 3 main molecular and regulatory pathways; (i) epigenetic and transcriptional events regulating response to HU, ii) signalling pathways involved in HU-mediated response and iii) post-transcriptional pathways of gene expression by microRNAs (miRNAs) regulation. 2) The erythroid model was successfully generated and used to show that HU promotes HbF production by down-regulating negative regulators; *BCL11A*, *KLF-1* and *MYB*, through miRNA-mediated action with miR-26b at the core of *MYB* silencing.

Regarding the studies of SCD within the Southern African context: Our data confirmed that; 1) the evolutionary dynamics of various African genomes (through selective pressure of malaria), with a decreasing frequency of the sickle mutation and specific malaria-associated genomic variants, when studying populations from central-west to southern Africa; 2) the study prompts a debate on the dissociation of the assumed congruence between genetics and cultural/linguistic/anthropological heritages of various so-called Bantu African populations. It further proposes population classification by geographical location evidenced by the presence of the sickle mutation in the Malawian and not in the Xhosa study participants, despite several cultural similarities. These data indicate that the perspective of using the classical ethnolinguistic label of “Bantu” may be irrelevant for pharmacogenetics and drug efficacy studies; 3) the findings of the β -globin gene haplotypes in the 3 Southern African populations unexposed to malaria, displayed an expected high frequency of the Atypical haplotype, a higher prevalence of Benin instead of Bantu haplotype and seeks

to contribute to understanding the origin and distribution of the sickle mutation and migration path of southern African populations following the Bantu expansion; 4) lastly, the clinical and epidemiological study of SCD confirm (i) the increasing burden of adult SCD patients at Groote Schuur Hospital in Cape Town, as previously reported in paediatric patients at Red Cross War Memorial Hospital, as a result of the recent migration of populations from central/eastern African countries where SCD is prevalent and (ii) seeks to caution national health and academic institutions to adapt policies, health care and professional training, accordingly.

In conclusion, this multifaceted thesis has reported on: 1) the importance and clinical implications of two erythroid-specific enhancer polymorphisms in *BCL11A* (rs1427407 and rs7606173), and provided evidence of the necessity and complexity of replicating findings of HbF-promoting loci across populations, 2) the post-transcriptional mechanism of HU to promote HbF production by down-regulating negative regulators; *BCL11A*, *KLF-1* and *MYB*, through miRNA-mediated action with miR-26b at the core of *MYB* silencing; 3) and the increasing burden of SCD in adult patients in Cape Town and the importance of studying Southern African populations to understand the evolutionary dynamics of Bantu population genomes under the selective pressure of malaria.

Collectively, the studies of predisposing polymorphisms to hereditary persistence of HbF in adulthood or mechanisms of HbF production in response to HU, particularly through miRNAs regulation, reveal some alternative perspectives and routes towards identifying new therapeutic targets and approaches for SCD.

Chapter 1 – Introduction

Sickle Cell Disease (SCD) is the oldest monogenic disease, first reported in 1910 (Herrick, 2001), and is concomitant with anaemia, progressive organ damage and chronic and acute illness (Weatherall et al., 2005). Sickle Cell Anemia (SCA), the most common and severe form of SCD, is caused by a T>A nucleotide substitution in the 17th position of the β -globin gene on chromosome 11p (Bunn, 1997). The mutation causes a change in the 6th amino-acid of the peptide from glutamic acid to valine (β 6Glu>Val). A hydrophobic motif is then created in the deoxygenated SCA-hemoglobin (HbS) resulting in polymerization between the β 1 and β 2 chains of the HbS molecules (Brittenham, Schechter & Noguchi, 1985). The HbS polymers grow and crystalize within the erythrocytes, promoting cellular dehydration and disruption of the intracellular architecture and oxygen-carrying capacity of the erythrocytes. This manifests in the characteristic sickled-shaped erythrocytes (Herrick, 2001). The sickled cells drive 2 main pathophysiological processes i.e. vaso-occlusion and hemolytic anaemia. The aberrant-shaped erythrocytes cause occlusion in micro-vessels via abnormal erythrocyte-endothelial interactions resulting in tissue ischemia and inflammation (Frenette, 2002). Restoration of blood flow promotes tissue and organ damage through the release of free-oxygen radicals that initiate symptoms of acute pain, vasculopathy and endothelial dysfunction due to intravascular and extravascular hemolysis

SCA accounts for nearly 70% of SCD cases in populations of African ethnic origin (Serjeant, Hambleton & Thame, 2005) and has the highest disease burden in Sub-Saharan African with in excess of 300 000 affected new-borns per year, accounting for 0.74% of all births in the region and 80% of annual global affected child births (Modell & Darlison, 2008). The high disease-related mortality is generally ascribed to bacteraemia and malaria (Serjeant, 2001; Serjeant, Hambleton & Thame, 2005). SCD is common and most severe in Africa but not to the point of exclusion of other continents such as North America and Europe with 2 600 and 1 300 affected new-borns annually, respectively (Modell & Darlison, 2008). There is strong correlation between the frequency of the HbS gene and the historical distribution and incidences of malaria

(Williams et al., 2005). This is mainly due to partial carrier-resistance to *Plasmodium falciparum* malaria, albeit the mechanisms of protection are not fully understood. The geographical co-occurrence of SCD and malaria and the partial carrier-resistance have caused the local amplification and positive selection of the HbS allele (Flint et al., 1998). Similarly, genetic mutations that often co-segregate with HbS such as HbC and β -thalassemia, have also undergone selection within similar populations. Furthermore, it is well accepted that the HbS allele exists on diverse genetic haplotype backgrounds (Pagnier et al., 1984; Labie et al., 1985), with five, independent and defined region-specific disease haplotypes based on the combination of five sequence variants across the β -globin gene cluster. Four haplotypes are associated with HbS allele in Africa (Benin, Bantu/Central African Republic (CAR), Senegal and Cameroon) and the fifth is thought to have arisen in India and/or the Arabian Peninsula (Arab/Hindu) (Pagnier et al., 1984; Elion et al., 1992). The β -globin haplotypes have been associated with disease severity and progression (Steinberg, 2009; Alsultan et al., 2012).

Approximately, 10% of children with SCA have cerebrovascular accidents (Ohene-Frempong et al., 1998) and a correlation between silent infarctions and overt strokes in young children has been elucidated (Miller et al., 2000; Miller et al., 2001). Pulmonary hypertension due to vaso-occlusion is another frequent complication in children with SCD (Collins & Orringer, 1982). Early diagnosis, prophylaxis and penicillin have been proven to decrease disease-related mortalities (Gaston et al., 1986; Vichinsky, 1991) and drastically improve (Lee et al., 1995) survival rates in the Caribbean (Lee, 1995), United States (Quinn, Rogers & Buchanan, 2004), the United Kingdom (Telfer et al., 2007) as well as in Africa (Fleming, 1989; Rahimy et al., 2009)(Fleming, 1989). Despite, SCD being vastly heterogeneous and multi-systemic, fetal hemoglobin (HbF) has emerged as has the central disease modifier (Platt et al., 1994), which is amenable to therapeutic manipulation (Bauer & Orkin, 2011). HbF is largely regulated by genetic variants at three principal loci, *BCL11A*, *HBS1L-MYB* and β -globin gene cluster, accounting for 10-20% of HbF variation (Menzel et al., 2007; Creary et al., 2009; Thein et al., 2009); among SCD patients in USA and Brazil (Lettre et al., 2008), Tanzania (Makani et al., 2011) and Cameroon (Wonkam et al., 2014).

Currently, hydroxyurea (HU) is the only Food and Drug Administration (FDA-USA)-approved (Centre of Drug Evaluation and Research App No. 75143, 1998) pharmacologic treatment for induction of HbF in patients with SCD. The major HU benefit is directly related to its HbF-producing effect (Lebensburger et al., 2010; Lebensburger et al., 2011) that leads to significant reduction of pain, acute chest episodes, the need for blood transfusions and mortality (Steinberg et al., 2003; Steinberg, 2009; Voskaridou et al., 2010; Lobo et al., 2013). The first clinical application of HU showed swift and vivid increases in HbF concentration within reticulocytes and insignificant toxicity to bone marrow (Platt et al., 1984b) and subsequent trials further demonstrated clinical efficacy and increase in survival rates and life expectancy (Steinberg et al., 2003; Voskaridou et al., 2010), protection against cerebrovascular disease (Zimmerman et al., 2007), long term drug safety, capacity to prevent organ damage, reduced morbidity and mortality in school-age children (Kinney et al., 1999), toddlers (Hankins et al., 2005; Thornburg et al., 2009) and infants (Wang & Thompson, 2010). Several studies have demonstrated variability in the response to HU treatment, with induced HbF levels ranging from 10% to greater than 30% (Maier-Redelsperger et al., 1998; Kinney et al., 1999; Zimmerman et al., 2004). Some of this variance has been attributed to variants at principal HbF-promoting loci like *BCL11A* (Ware et al., 2011). However, despite this, a complete understanding of the molecular mechanisms by which HU induces HbF are not fully understood.

In this thesis, we investigated such critical HbF-promoting loci; *BCL11A* enhancer, *HBB* globin haplotype, *SAR1a* and several other variants previously associated with HbF (Chapters 3.1.3 and 3.3.2). Secondly, we attempted to elucidate a miRNAs-mediated post-transcriptional mechanism of HbF induction through *ex-vivo* HU treatment of human umbilical cord blood-derived erythroid progenitors. Lastly, we reported the clinical phenotypes and genomic data of a cohort of SCD patients at Groote Schuur Hospital (Cape Town, South Africa) and investigated the prevalence of SCD-related polymorphisms; β -globin gene haplotype, HbS mutation and malaria-resistance variants; in 3 SCD-unaffected (HbAA) cohorts from South Africa, Zimbabwe and Malawi.

Investigation of HbF-promoting loci and the molecular mechanisms of HU carries substantial clinical significance as a better understanding of these genomic variants could reveal novel genomic drug-targets in SCD; improve population-specific drug design, targeting and efficacy, as well as the clinical management of patients based on their genetic predispositions. Most importantly, *ex-vivo* disease modelling will build on the knowledge base of drug mechanisms and elucidate potential alternative therapeutic approaches and targets, whereas an updated perspective of SCD in southern Africa will ensure necessary training to medical personnel, improved patient management and counselling and appropriate adaptation of healthcare legislation and policies.

Chapter 2 – Aims and Objectives

Aim 1:

To investigate variants in key HbF-promoting loci such as *BCL11A*, *HBS1L-MYB*, *SAR1*, α -thalassemia and the β -globin gene cluster, in a cohort of SCD patients from Cameroon and validate novel HbF-promoting loci reported in 2 genome-wide association studies (GWAS) carried out in Sardinia and Tanzania and *SAR1a* promoter variants associated with HbF in African American patients.

Objectives:

1. Design genotyping assays for HbF-promoting loci (including primer design and optimization of PCR, SNaPshot and direct Sanger sequencing assays);
2. Genotype a total of 28 loci associated with HbF including *BCL11A* (intragenic variants and enhancer elements), *SAR1*, α -thalassemia 3.7kb gene deletion allele and gene polymorphism conferring the β -globin gene haplotypes in Cameroon SCD patients. The loci investigated are: rs1427407; rs7606173; rs8176703; rs372091; rs2334880; rs2310991; rs4282891; rs76901216; rs76901220; X12_123681790; X16_391593; rs10468869; rs10756993; rs113267280; rs11754265; rs141494605; rs148706947; rs183437571; rs192197462; rs57001381; rs59329875; rs62573842; rs6466533; rs6590706; rs67104793; rs7163278; α -thalassemia 3.7 and 4.2kb gene deletions.
3. Validate (in SCD patients from Cameroon) results observed in genome-wide association studies (GWAS) performed in Sardinia and SCD patients from Tanzania and *SAR1a* promoter variants associated with HbF in African American patients.
4. Perform analysis using an additive genetic model, under a generalized linear regression framework, to investigate the relationship between these polymorphisms and HbF.

Aim 2:

To investigate the molecular mechanisms of hydroxyurea (HU)-induced production of HbF using a primary erythroid cell model from hematopoietic stem cells (HSCs) derived from umbilical cord blood.

Objectives:

1. Recruit and harvest human umbilical cord blood from elective cesarean healthy controls (HbAA) from Mowbray Maternity Hospital (Cape Town, South Africa);
2. Isolate, expand and differentiate human umbilical cord blood-derived HSCs into erythroid progenitors;
3. Treat erythroid and the K562 cells with HU and investigate the molecular mechanisms of HbF induction through differential gene, miRNAs and protein expression;
4. Perform miRNA inhibition assays for candidate miRNAs associated with HU; protein quantification assays using western blot and flow cytometry.

Aim 3:

To investigate the prevalence of SCD-related polymorphisms; β -globin gene haplotype, HbS mutation and malaria-resistance variants in 3 SCD-unaffected (HbAA) cohorts from South Africa, Zimbabwe and Malawi.

Objectives:

1. Perform genotyping assays for select variants associated with SCD; β -globin gene haplotype, HbS mutation and malaria-resistance variants on 3 cohorts from southern Africa that are unaffected (HbAA) by SCD, from South Africa, Zimbabwe and Malawi.

Chapter 3.1 – Results

Study of additional fetal hemoglobin-promoting loci

Synopsis of Chapter 3.1

Fetal hemoglobin (HbF) is the central modifier of Sickle Cell Disease and is inherited as a quantitative trait under the regulation of variants at principal HbF-promoting loci. Understanding the genetic predispositions to hereditary persistence of HbF in adulthood will explain the vast heterogeneity of the disease and improve management. The present chapter includes four original publications (reporting on a cohort of Cameroon SCD patients) that;

- i) Represent the background of this chapter and describes the global distribution of Sickle Cell Disease haplotypes and as a result, puts forward a provocative but likely hypothesis of a single origin for the sickle mutation in Africa.
- ii) Validates the relationship between HbF and two SNPs (rs1427407 and rs7606173) in a *BCL11A* erythroid-specific enhancer and provides the first report on a West African cohort.
- iii) Shows no association between variants previously associated with favourable pharmacologic response to HU in African American SCD patients.
- iv) Shows no associations between 17 novel SNPs (from a Tanzanian GWAS and whole genome sequencing in Sardinia) and HbF levels, other hematological indices or clinical events.

3.1.1. Bitoungui VL, **Pule DG**, Hanchard N, Ngogang J and Wonkam A. 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*. 2015 19(3): 171-9.

Abstract

Studies of hemoglobin S haplotypes in African subpopulations have potential implications for patient care and our understanding of genetic factors that have shaped the prevalence of sickle cell disease (SCD). We evaluated HBB gene cluster haplotypes in SCD patients from Cameroon, and reviewed the literature for a global distribution. We reviewed medical records to obtain pertinent socio-demographic and clinical features for

610 Cameroonian SCD patients, including hemoglobin electrophoresis and full blood counts. RFLP-PCR was used to determine the HBB gene haplotype on 1082 chromosomes. A systematic review of the current literature was undertaken to catalogue HBB haplotype frequencies in SCD populations around the world. Benin (74%; n = 799) and Cameroon (19%; n = 207) were the most prevalent haplotypes observed among Cameroonian patients. There was no significant association between HBB haplotypes and clinical life events, anthropometric measures, hematological parameters, or fetal hemoglobin (HbF) levels. The literature review of the global haplotype distributions was consistent with known historical migrations of the people of Africa. Previously reported data from Sudan showed a distinctly unusual pattern; all four classical haplotypes were reported, with an exceptionally high proportion of the Senegal, Cameroon, and atypical haplotypes. We did not observe any significant associations between HBB haplotype and SCD disease course in this cohort. Taken together, the data from Cameroon and from the wider literature suggest that a careful reassessment of African HBB haplotypes may shed further light on the evolutionary dynamics of the sickle allele, which could suggest a single origin of the sickle mutation.

Nature of publication: Original full article

Journal/Publisher: OMICS: A Journal of Integrative Biology/Mary Ann Liebert; ISSN: 1536-2310

Candidate contribution: Performed main literature search, analysed the data, wrote the paper, revised and approved the manuscript.

Co-authors contribution: Conceived and designed the experiments: VJNB, AW. Performed the experiments: VJNB, AW. Performed the main literature search: **GP**. Analyzed the data: VJNB, **GP**, AW. Contributed reagents/materials/analysis tools: NH, NJ, AW. Wrote the paper: VJNB, **GP**, AW. Revised and approved the manuscript: VJNB, **GP**, NH, NJ, AW.

Beta-Globin Gene Haplotypes Among Cameroonians and Review of the Global Distribution: Is There a Case for a Single Sickle Mutation Origin in Africa?

Valentina J. Ngo Bitoungui,¹ Gift D. Pule,² Neil Hanchard,³ Jeanne Ngogang,¹ and Ambroise Wonkam²

Abstract

Studies of hemoglobin S haplotypes in African subpopulations have potential implications for patient care and our understanding of genetic factors that have shaped the prevalence of sickle cell disease (SCD). We evaluated *HBB* gene cluster haplotypes in SCD patients from Cameroon, and reviewed the literature for a global distribution. We reviewed medical records to obtain pertinent socio-demographic and clinical features for 610 Cameroonian SCD patients, including hemoglobin electrophoresis and full blood counts. RFLP-PCR was used to determine the *HBB* gene haplotype on 1082 chromosomes.

A systematic review of the current literature was undertaken to catalogue *HBB* haplotype frequencies in SCD populations around the world. Benin (74%; $n=799$) and Cameroon (19%; $n=207$) were the most prevalent haplotypes observed among Cameroonian patients. There was no significant association between *HBB* haplotypes and clinical life events, anthropometric measures, hematological parameters, or fetal hemoglobin (HbF) levels. The literature review of the global haplotype distributions was consistent with known historical migrations of the people of Africa. Previously reported data from Sudan showed a distinctly unusual pattern; all four classical haplotypes were reported, with an exceptionally high proportion of the Senegal, Cameroon, and atypical haplotypes.

We did not observe any significant associations between *HBB* haplotype and SCD disease course in this cohort. Taken together, the data from Cameroon and from the wider literature suggest that a careful reassessment of African *HBB* haplotypes may shed further light on the evolutionary dynamics of the sickle allele, which could suggest a single origin of the sickle mutation.

Introduction

SICKLE CELL DISEASE (SCD) is caused by a point mutation ($A>T$) in the sixth codon of the β -globin gene on chromosome 11, resulting in the substitution of the amino acid valine for glutamic acid.

SCD affects the structure of erythrocytes by altering the normal biconcave shape to a crescent. During this process the hemoglobin S (HbS) mutation leads to polymerization and precipitation of hemoglobin during deoxygenation or dehydration, resulting in sickling, abnormal adhesion of leukocytes and platelets, inflammation, hypercoagulation, hemolysis, and hypoxia, in addition to microvascular obstruction and ultimately organ damage (Bartolucci et al., 2012). SCD is most prevalent among populations in regions

of the world where malaria has been endemic and it is estimated that 305,800 births with SCD occur annually, nearly two-third of which take place in Africa (Piel et al., 2013).

It is well accepted that the sickle mutation exists in Africa on diverse genetic haplotype backgrounds (Labie et al., 1985). Five typical haplotypes have been described across the β -globin gene cluster based upon the pattern of specific RFLPs across the region. Four haplotypes are associated with HbS in Africa (Benin, Bantu/Central African Republic (CAR), Senegal, and Cameroon) and the fifth is thought to have arisen in India and/or the Arabian Peninsula (Arab/Hindu) (Elion et al., 1992; Pagnier et al., 1984). It has been suggested that these haplotypes also have an effect on the severity of the disease (Asultan et al., 2012; Steinberg, 2009) and possibly the clinical response to hydroxyurea (Friedrich et al., 2008)—currently the only

¹Department of Microbiology, Parasitology, and Hematology, Faculty of Medicine and Biomedical Sciences, University of Yaoundé, Yaoundé, Cameroon.

²Division of Human Genetics, Institute of Infectious Disease and Molecular Medicine, Faculty of Health Sciences, University of Cape Town, Cape Town, Republic of South Africa.

³Department of Molecular and Human Genetics, Baylor College of Medicine, Houston, Texas.

available treatment for SCD. Furthermore, these haplotypes have a well-defined geographic distribution in Africa, making it possible to establish the origin of African-descendant populations in America (Pagnier et al., 1984; Serjeant, 1989), but also the origin of SCD among Europeans from Mediterranean regions. Studies of SCD haplotypes in African subpopulations therefore have potential implications for the clinical care of patients and for our understanding of the dynamic population genetic factors that shape the SCD phenotype.

Cameroon is a sub-Saharan African country with a population of about 20 million inhabitants and an annual growth rate of 3%, with a carrier frequency of SCD ranging from 8% to 34% (Wheatherall et al., 2001). Universal medical insurance coverage does not exist, and care of SCD patients is therefore dependent on financial support and care-giving by family members (Wonkam et al., 2013a). Sadly, poverty in Cameroon affects more than 50% of the rural population and up to 30% of the urban population (World Bank, 2010), and the financial burden of medical care for SCD is often not met (Wonkam et al., 2013a, 2014a).

SCD patients in Cameroon frequently suffer exceptionally severe complications such as stroke (Njamnshi et al., 2006) and executive neurocognitive compromise (Ruffieux et al., 2013). Consanguinity, ascertained by pedigree analysis, is observed in 5.7% of the population, most of which is found in the northern Sahelian regions of Cameroon where the Arabo-Muslim influence is stronger (Wonkam et al. 2013b). Because of its central location on the continent, Cameroonians cultural, linguistic, and genetic diversity mimics that of various ethno-linguistic groups found in Africa (Tishkoff et al., 2009). It is therefore anticipated that genetic studies in this population might reveal important insights to other African populations. The aim of this study was to explore the frequency and influence of *HBB* haplotypes amongst a group of Cameroonian SCD patients and to compare our findings to published data from Africa and worldwide.

Materials And Methods

Ethics statement

Ethical approval was granted by the National Ethical Committee Ministry of Public Health, Republic of Cameroon (No 033/CNE/DNM/07), and the University of Cape Town, Faculty of Health Sciences Human Research Ethics Committee (HREC RE: 132/2010).

Patients and assessment of clinical events

The study was conducted at the Yaoundé and Douala, the two largest cities in Southern Cameroon. Socio-demographic and clinical data were collected by means of a structured questionnaire for parents/guardians and patients. Patients' medical records were reviewed to delineate their clinical features over the past 3 years. Specifically, the occurrences of vaso-occlusive painful crisis (VOC), hospital outpatient visits, hospitalizations, overt strokes, blood transfusions, and administration of hydroxyurea were recorded. Anthropomorphic variables (height, weight, body mass index (BMI)) and systolic and diastolic blood pressures (SBP and DBP) were measured in the outpatient setting. Detailed descriptions of sampling methods used in the Cameroonian have been previously reported (Wonkam et al, 2014c).

Hematological phenotypes

Hemoglobin electrophoresis and routine complete blood count were conducted on each patient upon arrival at the hospital. Two methods of fetal hemoglobin (HbF) detection were employed; the alkali denaturation test (ADT) in 55.5% ($n=344$) of the patients, and subsequent high performance liquid chromatography (HPLC) in the remainder. HbF levels measurements done in patients <5 years old were excluded from the analysis, because HbF levels are not yet stable at this age (Wonkam et al, 2014c).

Genotypes

Cameroonian patients. DNA samples were extracted from peripheral blood, following manufacturer's instructions (Puregene blood kit® (Qiagen®, USA) at the molecular diagnostic laboratory, Gyneco-Obstetric and Paediatric Hospital, Yaoundé, Cameroon. Genotype analyses were performed in the Division of Human Genetics, Faculty of Health Sciences, University of Cape Town.

Molecular diagnostic testing for sickle cell anemia (SCA; HbSS). SCA diagnosis was carried out using 200 ng of DNA. A thermocycler (Biorad®, USA) was used to amplify a 770 bp segment of the β -globin gene, followed by Dde I (Gibco-BRL®, USA) restriction analysis of the PCR product, as previously reported (Saiki et al., 1985).

Haplotyping of β -globin gene cluster. Five restriction fragment length polymorphism (RFLP) regions in the β -gene cluster were amplified using published methods designed to analyze the *XmnI* (5'G γ), *HindIII* (G γ), *HindIII* (A γ), *HincII* (3'Ψ β), and *HinfI* (5' β) restriction sites (Steinberg et al., 1998). RFLP sites and the fragments were visualized by agarose gel electrophoresis, and β -globin gene haplotypes were defined by examining the combination of the restriction sites. The β S haplotypes were constructed from the absence (–) and presence (+) of each of the five restriction enzyme sites in 541 individuals with HbSS. Of the samples studied, sickle (β S) haplotypes were identified through a specific algorithm (Supplementary Table S1; supplementary material is available online at www.liebertonline.com/omi). Haplotypes were classified as 'atypical' if they did not match the five known typical β S haplotypes.

Literature review

A review of the current literature on SCD haplotypes was conducted from June 2013 to July 2014 using Pubmed (National Library of Medicine), Medline, and Google Scholar. Key words included a combination of the following: "Haplotype," "Sickle Cell," "Africa," and specific country names were also used. Prior knowledge of research groups working on sickle cell disease in Africa and globally further facilitated the identification and selection of research articles. Only available full-length articles, in English, with the use of at least five (5) RFLP to differentiate the five main types of haplotype, were selected. In cases where multiple studies were reported in the same population, the most recent with the largest sample size was included. The main search was conducted by a PhD student in Human Genetics and the search was reviewed by a Medical/Human Geneticist with expertise in SCD.

BETA-GLOBIN GENE HAPLOTYPES AMONG CAMEROONIANS & REVIEW OF GLOBAL DISTRIBUTION

Statistical analysis

Descriptive statistics were obtained for all quantitative data using SPSS (IBM, USA version 21.0). Normality was confirmed by the Shapiro-Wilk Test followed by the use of parametric (Chi-squared test and *t*-test) or nonparametric tests (Mann-Whitney U test for two samples or the Kruskal-Wallis one-way ANOVA for more than two samples). To correct for the skewed HbF distribution, the data were log₁₀-transformed and normalized to obtain the quantitative trait used in the association analysis (after correcting for age, gender, and electrophoresis technique and history of transfusion).

Results

Socio-demographic of Cameroonian SCD patients

The description of the study sample is presented in Table 1. Among the 610 patients; 50.3% were female with mean age of 17.3±10 years (range: 5–54 years) and the majority of patients were children aged between 5–10 years (32.2%; *n* = 196) and adolescents 11–20 years (50%; *n* = 305). Patients lived mostly in the urban areas of Yaoundé and Douala (93%; *n* = 567), both located in the southern part of Cameroon. The median age at diagnosis of SCD, was 3 years (range: 1 month–29 years). Forty two percent (*n* = 158) of fathers and 23% (*n* = 108) of mothers were formally employed; and in 75%, the monthly direct incomes were <300 USD.

Haplotypes in the beta-globin gene cluster, clinical events, and HbF in Cameroonian patients

Following the full blood count and HbF electrophoresis and RFLP-PCR, the vast majority of patients were determined to have SCA (HbSS) (97.4%; *n* = 594). In addition, 61 randomly selected patients were Sanger sequenced, and SCA diagnosis was confirmed in all of them. There was no evidence of the deletional form of HbS β-thalassemia.

TABLE 1. DESCRIPTION OF SCD CAMEROONIAN PATIENTS

Variables	N	Mean ± SD	Minimum–Maximum
M/F	303/307		
Age (Yrs)	610	17.3 ± 10	5–58
RBC (10 ¹² /L)	610	2.8 ± 0.7	1.2–6.5
Hb (g/dL)	610	7.8 ± 1.6	3.8–16.5
MCV (fL)	610	84.4 ± 9.9	59–117
MCHC (g/dL)	610	34.0 ± 3.8	21.5–54.3
WBC (10 ⁹ /L)	610	13.7 ± 5.6	2.9–49.8
Monocytes (10 ⁹ /L)	610	1.5 ± 1	0.1–11.9
Pl atelelets (10 ⁹ /L)	610	374.6 ± 123	97–837
HbA2 (%) (HPLC)*	244	3.3 ± 1.3	0–9.7
HbF (%) (HPLC)*	244	7.5 ± 4.8	0–22.7
N VOC /Yr	572	3 ± 3	0–40
N hospital attendance (per Yr)	608	2.2 ± 3.2	0–60
N hospital admission (per Yr)	606	2.3 ± 3	0–30
N patients with at least one overt stroke	25/608	4.1%	

*Hb electrophoresis was also obtained from 344 patients (55.5%) using alkali denaturation test (ADT), with a mean of 11.4±9.4 for HbF and 4.1±2.1 for HbA2 levels.

Based on analysis of 1082 chromosomes, the allele frequencies of various haplotypes showed that Benin (74%; *n* = 799) and Cameroon (19%; *n* = 207) were the most prevalent forms. Bantu/CAR (*n* = 5), Senegal (*n* = 3), and Indian-Arab (*n* = 2) haplotype were very rare. Atypical haplotypes accounted for 6% (*n* = 66). Homozygous Benin/Benin haplotypes represented 57% (*n* = 307), Benin/Cameroon 25% (*n* = 137), Benin/atypical 8% (*n* = 45), and Cameroon/Cameroon 5% (*n* = 26) of diploid haplotype frequencies. There were no significant associations between major haplotype combinations and hematological indices (Table 2), clinical events (Table 3), or HbF levels (Fig. 1).

Global distribution of haplotypes

As demonstrated in Figure 2 and Table 4, the Benin haplotype was the most prevalent globally with its highest frequency in West Africa, the Mediterranean region, America, and Europe. The Bantu/CAR haplotype was most prevalent in central, southern, and East Africa, and in relatively high proportion in Central and South America and parts of Europe, as compared to North America. The Senegal haplotype was restricted to the far west regions of Africa, with the exception of Sudan and had a relatively low frequency in America; the Indian-Arab haplotype was almost completely restricted to part of India and the Middle East.

Exceptional HBB haplotype distribution in Sudan?

When data within Africa was compared, the literature review revealed four major features in Sudanese SCD patients: 1) all four classical SCD haplotypes were reported in significant proportions; 2) Cameroon haplotype was reported in a much higher proportion in Sudan than in Cameroon; 3) surprisingly, there was a higher proportion of Senegal haplotypes in Sudan when compared to the rest of East Central and parts of West Africa, and lastly 4) there was a relatively high proportion of atypical haplotypes (Elderderly et al., 2012) (Table 4, Fig. 2).

Discussion

To the best of our knowledge, the present study represents the largest study of *HBB* haplotypes in a single African country, and is the most comprehensive review of the global distribution of *HBB* SCD haplotypes thus far. The data presented confirm the origin and flow of people from Africa through the slave trade and/or migration, and emphasize differences in settlement on the American coasts dependent on whether slaves were brought from the Gulf of Guinea or from Angola (Fong et al., 2013). Similarly, trade and migration routes to the Mediterranean areas and the Middle East from West Africa explain the high prevalence of Benin haplotypes and general pattern of haplotype distribution in those parts of the world (Fig. 2).

The present study also confirms that the majority of the chromosomes with the β(S) globin gene are consistent with one of the five common RFLP haplotypes, designated as Benin, Bantu/CAR, Senegal, Cameroon, and Arab-Indian haplotypes. In most settings, 5%–10% of the chromosomes are less commonly observed haplotypes, usually referred to as “atypical” haplotypes (Table 4). However, on the African continent and in two independent studies, atypical haplotypes

TABLE 2. MEDIAN VALUES OF HEMATOLOGICAL INDICES AMONG CAMEROONIAN SCD PATIENTS CONDITIONED ON *HBB* HAPLOTYPES

Hematological variables	Benin/Benin (n=307)	Benin/Cam (n=137)	Benin/Atypical (n=45)	Cameroon/Cameroon (n=26)	P values
RBC ($10^{12}/L$)	2.7	2.6	2.8	2.6	0.6
Hb (g/dL)	7.7	7.5	7.8	7.6	0.5
MCV (fL)	85	84	86	83.5	0.9
MCHC (g/dL)	33.4	34.1	32.6	34.3	0.5
Platelets ($10^9/L$)	395	374	350	382	0.4
WBC ($10^9/L$)	12.6	13.6	13.1	14.3	0.5
Lymphocytes ($10^9/L$)	5.2	5.6	5.4	6.5	0.1
Monocytes ($10^9/L$)	1.3	1.3	1.3	1.3	1

appear in relatively high frequency in Sudan, in concert with all the other types of African haplotypes, particularly a relatively high proportion of Senegal and Cameroon haplotypes (Elderderly et al., 2012; Mohammed et al., 2006). The relatively high frequency of the Cameroon haplotype in Sudan versus in Cameroon (Elderderly et al., 2012) is not completely anomalous, but rather the reflection that the geographical nomenclature of haplotypes (named according to the first region of observation) does not necessarily reflect their place of origin. The refinement of the molecular structure (i.e., haplotype) in *HBB* and migration patterns is still urgently needed, specifically in Africa.

Two decades ago, the observation of a specific RFLP pattern of the Cameroon haplotype strongly argued for yet another independent origin of the sickle cell mutation in Africa (Lapoumeroulie et al., 1992). The Cameroon haplotype was then said to be restricted to a single ethnic group, the Eton of Central Cameroon (Lapoumeroulie et al., 1992), which is contrary to current data from Sudan (Elderderly et al., 2012; Mohammed et al., 2006). More recently, a specific atypical haplotype in Tunisia was considered by authors as specific to the Tunisian chromosome $\beta(S)$ (Imen et al., 2011). In the case of the Cameroon and Tunisia, these specific haplotypes were likely generated by a variety of genetic mechanisms including (a) isolated nucleotide changes in one of the polymorphic restriction sites, (b) simple and double crossovers between two typical $\beta(S)$ haplotypes, or much more frequently between a typical $\beta(S)$ haplotype and a different $\beta(A)$ -associated haplotype that was present in the

population, and (c) gene conversions (Zago et al., 2000)—all of which have been observed at the beta-globin locus (Borg et al., 2009; Holloway et al., 2006; Liu et al., 2009; Neumann et al., 2010; Patrinos et al., 2005; Sankaran et al., 2011).

Thus, a relatively high frequency of what was originally an atypical $\beta(S)$ haplotype might subsequently be labeled a “new” haplotype supposedly (albeit unlikely) bearing a new independent $\beta(S)$ mutation, as was postulated by the authors of the original globin haplotype report (Wainscoat et al., 1983). To date, the existence of at least five different geographic centers of origin has been postulated on the basis of the predominance of major haplotypes associated with the $\beta(S)$ mutation. The multiple and independent origins of the $\beta(S)$ mutation (Chakravarti et al., 1984; Nagel and Fleming, 1992) also fails to account for the absence of $\beta(S)$ mutation from malarial regions in Southeast Asia and Oceania (Flint et al., 1993), although many other hemoglobin gene mutations also confer malaria protection, including α - and β -thalassemia traits and other structural variants of hemoglobin that are common in Southeast Asia.

In addition, the hypothesis of multiple independent origins raises the question of how several identical mutations could have occurred in a short period of time in Africa after the appearance of malaria, despite the low mutation rate of nuclear DNA (Curat et al., 2002). It is also possible that gene conversion has played a role in moving the $\beta(S)$ mutation to different haplotype backgrounds in sub-Saharan Africa. Resolution of RFLP haplotypes with recently described long-range sickle haplotypes (Ghansah et al., 2012; Hanchard et al.,

TABLE 3. CLINICAL EVENT AMONG CAMEROONIAN SCD PATIENTS CONDITIONED ON *HBB* HAPLOTYPES

	Benin/Benin		Benin/Cameroon		Benin/Atypical		Cameroon/Cameroon		P values
	N	Median (25 th -75 th centiles)	N	Median (25 th -75 th centiles)	N	Median (25 th -75 th centiles)	N	Median (25 th -75 th centiles)	
BMI (kg/m^2)	251	18 (14.9–21)	103	17.2 (15–20.1)	39	17 (14.4–19.7)	23	17.9 (15.1–18.5)	0.12
SBP mmHg	259	108 (100–115)	107	107 (98–115)	40	109 (99–113.5)	21	108 (99–115)	0.56
DBP mmHg	259	58 (54–65)	107	58 (52–64)	40	59 (53.5–64.5)	21	56 (53–60)	
No. of VOC/Year									
Stroke									
Yes	13	–	6	–	–	–	–	–	0.8
No	294	–	131	–	45	–	25	–	
No. of consultation/Year	301	1 (0–4)	135	2 (0–4)	45	1 (0–3)	26	1 (0–4)	0.66
No. of hospitalization/Year	301	1 (0–2)	135	1 (0–2)	45	1 (0–2)	26	1 (1–2)	0.54

BETA-GLOBIN GENE HAPLOTYPES AMONG CAMEROONIANS & REVIEW OF GLOBAL DISTRIBUTION

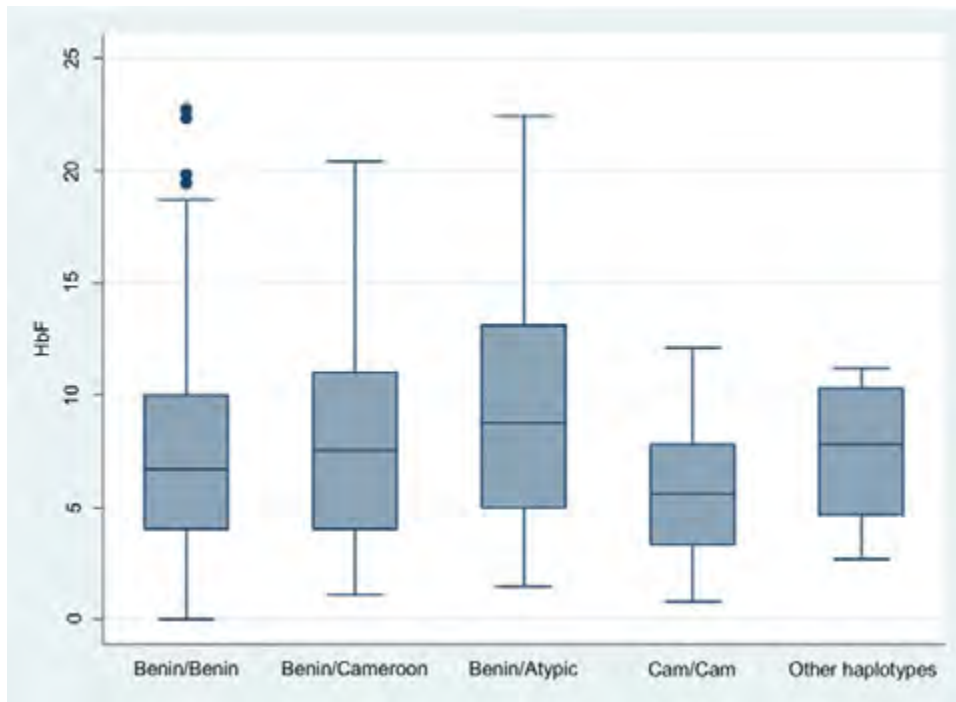


FIG. 1. Distribution of fetal hemoglobin levels conditioned on HBB haplotypes. *Boxes* have lines at the lower quartile, median, and upper quartile. The *plot whiskers* extend up and down from the median a distance 1.5 times the interquartile range of the boxes (truncated at zero where necessary). Outliers are the points outside the whiskers indicated as *circles*. There were not significant differences across various combination of haplotypes ($p = 0.54$) and between them.

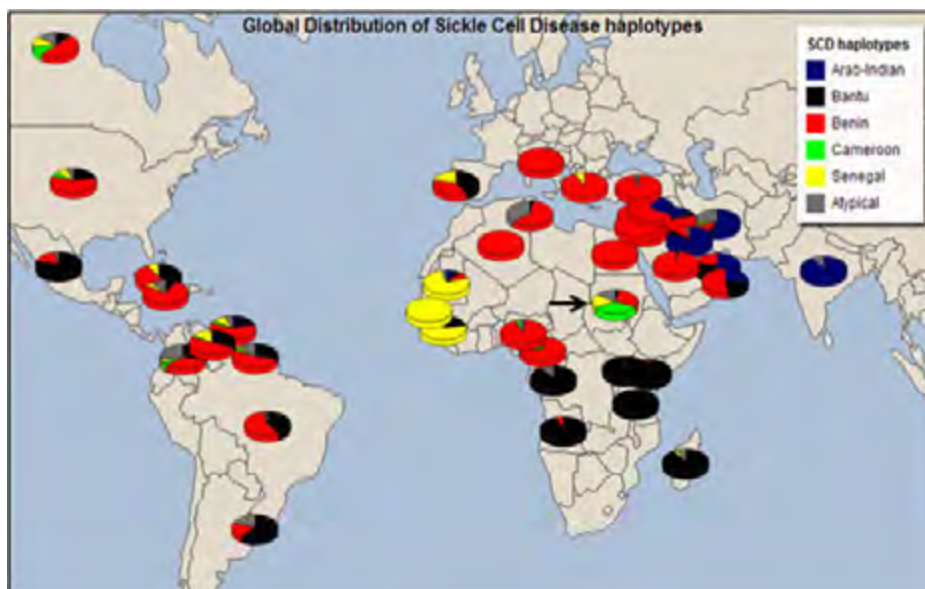


FIG. 2. Global distribution of haplotypes among various world SCD populations. Previously reported data allow drawing a global haplotype distribution that was consistent with known historical migrations of the people of Africa. Among Indigenous African populations, published data from Sudan (*arrow*) showed a distinctly unusual pattern; all four classical haplotypes were reported, with an exceptionally high proportion of the Senegal and Cameroon haplotypes.

TABLE 4. GLOBAL β -S HAPLOTYPE DISTRIBUTION FROM PREVIOUSLY REPORTED DATA

		<i>β-S Haplotypes</i>						
<i>Continents</i>	<i>Countries</i>	<i>Arab-Indian</i> N* (%)	<i>Bantu/Car</i> N (%)	<i>Benin</i> N (%)	<i>Cameroon</i> N (%)	<i>Senegal</i> N (%)	<i>Atypical</i> N (%)	<i>References</i>
Africa	Algeria	0 (0.0)	0 (0.0)	20 (100.0)	0 (0.0)	0 (0.0)	0 (0.0)	Pagnier et al., 1984
	Angola	0 (0.0)	42 (95.5)	2 (4.5)	0 (0.0)	0 (0.0)	0 (0.0)	Lavinha et al., 1992
	Cameroon	3 (0.3)	5 (0.5)	799 (73.8)	207 (19.1)	2 (0.2)	66 (6.1)	The present data
	Congo	0 (0.0)	211 (91.0)	0 (0.0)	0 (0.0)	0 (0.0)	21 (9.1)	Mouele et al., 1999
	Egypt	0 (0.0)	0 (0.0)	28 (100.0)	0 (0.0)	0 (0.0)	0 (0.0)	El-Hazemi et al., 1999
	Guinea	0 (0.0)	9 (22.5)	0 (0.0)	31 (77.5)	0 (0.0)	0 (0.0)	Sow et al., 1995
	Kenya	0 (0.0)	109 (98.2)	2 (1.8)	0 (0.0)	0 (0.0)	0 (0.0)	Ojwang et al., 1987
	Madagascar	0 (0.0)	32 (91.4)	0 (0.0)	0 (0.0)	1 (2.9)	2 (5.7)	Hewitt et al., 1996
	Mauritania	5 (5.6)	4 (4.4)	8 (8.9)	0 (0.0)	70 (77.8)	3 (3.3)	Veten et al., 2012
	Nigeria	0 (0.0)	6 (1.0)	624 (93.2)	23 (3.4)	0 (0.0)	16 (2.4)	Adekile et al., 1992
	Senegal	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	90 (100.0)	0 (0.0)	Curat et al., 2002
	Sudan	0 (0.0)	4 (2.8)	42 (29.4)	50 (35.0)	26 (18.2)	21 (14.6)	Elderderly et al., 2012
	Tanzania	0 (0.0)	41 (100.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	Oner et al., 1992
	Tunisia	0 (0.0)	9 (2.7)	201 (60.5)	0 (0.0)	0 (0.0)	122 (36.7)	Imen et al., 2011
	Uganda	0 (0.0)	207 (99.5)	0 (0.0)	0 (0.0)	1 (0.5)	0 (0.0)	Ndugwa et al., 2012
	Asia	India	128 (91.4)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	12 (8.6)
Bahrain		33 (89.2)	2 (5.4)	1 (2.7)	0 (0.0)	0 (0.0)	1 (2.7)	Al-Arrayed, 1995
Iran		87 (53.7)	5 (3.1)	19 (11.7)	4 (2.5)	6 (3.7)	41 (25.3)	Rahimi et al., 2003
Iraq		16 (12.5)	10 (7.8)	89 (69.5)	0 (0.0)	0 (0.0)	13 (10.2)	Al-Allawi et al., 2012
Jordan		4 (20.0)	0 (0.0)	16 (80.0)	0 (0.0)	0 (0.0)	0 (0.0)	El-Hazemi et al., 1999
Kuwait		101 (80.8)	7 (5.6)	14 (11.2)	0 (0.0)	0 (0.0)	3 (2.4)	Adekile, 2001
Lebanon		10 (10.0)	15 (15.0)	73 (73.0)	0 (0.0)	0 (0.0)	2 (2.0)	Inati et al., 2003
Oman		31 (26.5)	25 (21.4)	61 (52.1)	0 (0.0)	0 (0.0)	0 (0.0)	Daar et al., 2000
Palestine (wb)#		0 (0.0)	6 (5.1)	104 (88.1)	0 (0.0)	0 (0.0)	8 (6.8)	Samarah et al., 2009
Saudi-arabia (se)^		2 (1.6)	0 (0.0)	122 (98.4)	0 (0.0)	0 (0.0)	0 (0.0)	El-Hazemi et al., 1999
Syria		6 (33.3)	0 (0.0)	12 (66.7)	0 (0.0)	0 (0.0)	0 (0.0)	El-Hazemi et al., 1999
United Arab Emirates		49 (52.0)	24 (26.0)	21 (22.0)	0 (0.0)	0 (0.0)	0 (0.0)	El-Kalla & Baysal, 1998
North America	Canada	0 (0.0)	7 (11.5)	30 (49.2)	8 (13.1)	8 (13.1)	8 (13.1)	Oner et al., 1992
	Cuba	0 (0.0)	81 (40.9)	101 (51.0)	0 (0.0)	16 (8.1)	0 (0.0)	Muniz et al., 1995
	Jamaica	0 (0.0)	37 (8.3)	339 (76.0)	0 (0.0)	23 (5.2)	47 (10.5)	Ndugwa et al., 2012
	Mexico	0 (0.0)	26 (78.8)	6 (18.2)	0 (0.0)	0 (0.0)	1 (3.0)	Magana et al., 2002
	Trinidad	9 (3.2)	49 (17.3)	175 (61.8)	10 (3.5)	24 (8.5)	16 (5.7)	Jones-Lecoointe et al., 2008
South America	USA ^s	12 (0.8)	129 (21.6)	503 (58.9)	38 (6.1)	76 (8.0)	48 (4.6)	Crawford et al., 2002
	Uruguay	0 (0.0)	6 (60.0)	2 (20.0)	0 (0.0)	0 (0.0)	2 (20.0)	Luz et al., 2006
	Brazil	1 (0.4)	104 (41.6)	138 (55.2)	3 (1.2)	1 (0.4)	3 (1.2)	Adorno et al., 2008
	Colombia	1 (0.4)	68 (29.7)	76 (33.2)	10 (4.4)	10 (4.4)	64 (27.9)	Fong et al., 2013
	Suriname	0 (0.0)	23 (29.9)	41 (53.2)	2 (2.6)	2 (2.6)	9 (11.7)	Oner et al., 1992
	Venezuela	0 (0.0)	57 (32.2)	90 (50.8)	4 (2.3)	25 (14.1)	1 (0.6)	Arends et al., 2000
Europe	Portugal	0 (0.0)	14 (42.4)	12 (36.4)	0 (0.0)	7 (21.2)	0 (0.0)	Lavinha et al., 1992
	Italy	0 (0.0)	0 (0.0)	64 (100.0)	0 (0.0)	0 (0.0)	0 (0.0)	Schiliro et al., 1992
	Turkey	1 (0.5)	0 (0.0)	206 (96.3)	0 (0.0)	0 (0.0)	7 (3.3)	Oner et al., 1992
	Greece	0 (0.0)	0 (0.0)	13 (92.9)	0 (0.0)	1 (7.1)	0 (0.0)	Oner et al., 1992

*N, number of chromosomes; #WB, West Bank; ^SE, South Eastern; ^sUSA, combined African-American and Hispanic data.

2007), together with next-generation sequencing, should provide definitive answers to this long-standing question.

It is then perhaps provocative and ambitious, but not unreasonable, to hypothesize that, there could be a single origin of the sickle cell mutation, in the region of East Africa/Sudan, with additional haplotypes generated by diverse structural mechanisms, and the spread to the rest of the continent through traditional migratory patterns. Recently, both phylogenetic and network analysis indicate that east Africans possess more ancestral mitochondrial lineages in comparison

to various continental populations placing them at the foot of the human evolutionary tree. Interestingly, the two most ancestral mitochondrial sequences in the NJ tree figure refer to Nubian individuals in Sudan (Elhassan et al., 2014). Moreover, a compound associated with a lithic Middle Stone Age industry was discovered in Dhofar Oman and regarded as evidence of human migration out of Africa through an Arabian route (Rose et al., 2011).

Therefore, the shared rs7482144 SNP in both the Senegal and Indian-Arab haplotypes invites the question of whether

the Indian-Arab haplotype evolved from an east African chromosome with Senegal haplotype. In fact, a specific study of the Senegal haplotype concluded that migration from the east of Senegal was implicated in the spread of this mutation across west African populations and the exact geographic origin of the Senegal $\beta(S)$ mutation was not located (Curat et al., 2002). It could well be even further east of Senegal, perhaps as far east as Sudan/East Africa region. It was suggested more than 2 decades ago, that although $\beta(S)$ is attributed to recent independent mutations occurring approximately 2000 years ago (Curat et al., 2002), the separation and geographical distribution of African populations allows for the possibility of an ancient origin of $\beta(S)$ mutation. The authors even suggested that $\beta(S)$ could have protected selected populations from malaria in rain forest refuges during the most recent Ice Age (Stine et al., 2011).

The data reviewed here provides evidence that the detailed geographic distribution of known sickle haplotypes is still not well established, as many African countries do not have reported data. Were this available, it is conceivable that a gradient of haplotype frequencies from East Africa to other parts of the continent could have emerged differently. Further genetic studies in well-defined populations from across Africa are urgently needed, to allow determination of the exact extent of the spread of various *HBB* haplotypes, as this would help to identify the mutation's geographical origin precisely, and to explore the recent increase in frequency of the $\beta(S)$ mutation, irrespective of its haplotypic background. The data reported here emphasize the need to re-explore the origin of sickle cell mutation as well as past and recent evolutionary dynamics of African human populations, as this could yield important information about the evolutionary history of SCD, which of course hold significant clinical relevance (Adorno et al., 2008).

HbF is the most powerful biological modifier of SCD phenotype and influenced by genomic variations. In a recent study, we confirmed in this group of Cameroonian SCD patients, robust genetic associations between HbF levels and *BCL11A* and *HBSIL-MYB* intergenic locus (Wonkam et al., 2014c), as previously reported in African Americans and Afro-Brazilians (Lette et al., 2008) and Tanzanians (Makani et al., 2011). However, due to the virtual absence of the Senegal and Indian-Arab haplotypes, which contain the XmnI variant (rs7482144), we lacked power to replicate the association of rs7482144 in *HBB* with HbF levels in Cameroonian; this SNP explains 2.2% of the variation in HbF levels in African American SCD patients (Labe et al., 1985; Lette et al., 2008). Indeed, there was no significant difference in HbF levels and various haplotypes combinations among this group of Cameroonian SCD patients (Fig. 1). Contrary to our findings, a recent report argues that the S haplotype itself (beyond HbF regulation) correlates with disease severity (Bean et al., 2013).

The Cameroon haplotype has previously been associated with poor clinical outcome and increased stroke episodes in SCD (Adorno et al., 2008; Steinberg et al., 1998), whereas the Benin haplotype is thought to confer a relatively favorable clinical outcome (Steinberg, 2009). In a previous report, we anticipated that the relatively high prevalence of stroke in SCD patients in Cameroon could be attributed to the high proportion of Cameroonian haplotype (Njamnshi et al., 2006). However, the present data do not confirm this hypothesis, since the most prevalent haplotype in Cameroon is

Benin. The high prevalence of Benin haplotypes rendered the study underpowered to reveal a significant difference in phenotypic outcome with specific haplotype combinations. An alternative explanation could be the influence of environmental factors, specifically the limited access to the already inadequate health care and poverty, illustrated here by the late age at diagnosis, the low rate of formal employment, as well as modest direct monthly direct incomes among parents. In other studies, socioeconomic factors and poor healthcare infrastructure have been shown to be a major factor in the high psychosocial burden of SCD among Cameroonian adult SCD patients (Wonkam et al., 2014a,d) and parents of children with SCD (Wonkam et al., 2013).

Limitations and perspectives

In Cameroon, the data mostly originated from patients living in the southern part of the countries and are thus not fully representative of the general SCD population in Cameroon. Specifically, data from the northern Sahelian part will be needed to complete the picture and possibly provide additional clues of recent population migration from Sudan that brought the 'Cameroon' haplotype to Cameroon. Another limitation of this study was the self-report nature of some clinical variables such as VOC episodes, which can lack precision. In addition, pain tolerance and financial factors could have been limiting factors for hospital attendance (Wonkam et al., 2014c).

Despite the above limitations, this study represents an important step toward the understanding of the molecular structure of SCD in Cameroon, in sub-Saharan African patients and globally. This study also has a capacity-building dimension as it was fully performed from design, funding, molecular analysis, and reporting on the African continent and could in turn create major collaborative research opportunities at the regional and international levels.

Conclusion

This study has 1) confirmed the high frequency of Benin haplotype and lack of significant association between $\beta(S)$ haplotypes and common SCD clinical outcomes in Cameroonian SCD patients; 2) highlighted previous data demonstrating the almost equivalent representation of the four major *HBB* haplotypes in Sudan with a much higher frequency of Cameroon haplotype; 3) provided a comprehensive review of published data illustrating the distribution of $\beta(S)$ haplotypes in world populations, tracking their migration pattern within and outside of Africa. This preliminary report proposes extending the studies on *HBB* haplotype distribution in Africa and advances the opportunity to re-address the question of the origin(s) of the Sickle Cell mutation.

Author Disclosure Statement

The authors declare that there are no conflicting financial interests.

References

- Adekile AD, and Haider MZ. (1996). Morbidity, βS haplotype and α -globin gene patterns among sickle cell anemia patients in Kuwait. *Acta Haematol* 96, 150–154.
- Adorno EV, Zanette Â, Lyra I, et al. (2008). Clinical and molecular characteristics of sickle cell anemia in the northeast of Brazil. *Genet Mol Biol* 31, 621–625.

- Al-Allawi NA, Jalal SD, Nerwey FF, et al. (2012). Sickle cell disease in the Kurdish population of northern Iraq. *Hemoglobin* 36, 333–342.
- Alsultan A, Aleem A, Ghabbour H, et al. (2012). Sickle cell disease subphenotypes in patients from Southwestern Province of Saudi Arabia. *J Pediatr Hematol Oncol* 34, 79–84.
- Arends A, Alvarez M, Velázquez D, et al. (2000). Determination of β -globin gene cluster haplotypes and prevalence of α -thalassemia in sickle cell anemia patients in Venezuela. *Am J Hematol* 64, 87–90.
- Arrayed SSA, and Haltes N. (1995). Features of sickle-cell disease in Bahrain. *East Mediterr Health J* 1, 112–118.
- Bartolucci P, and Galactéros F. (2012). Clinical management of adult sickle-cell disease. *Curr Opin Hematol* 19, 149–155.
- Bean CJ, Boulet SL, Yang G, et al. (2013). Acute chest syndrome is associated with single nucleotide polymorphism-defined beta globin cluster haplotype in children with sickle cell anaemia. *Br J Haematol* 163, 268–276.
- Borg J, Georgitsi M, Aleporou-Marinou V, Kollia P, and Patrinos GP. (2009). Genetic recombination as a major cause of mutagenesis in the human globin gene clusters. *Clin Biochem* 42, 1839–1850.
- Chakravarti A, Buetow KH, Antonarakis SE, et al. (1984). Nonuniform recombination within the human beta-globin gene cluster. *Am J Hum Genet* 36, 1239.
- Crawford DC, Caggana M, Harris KB, et al. (2002). Characterization of the β -globin haplotypes using blood spots from a population-based cohort of newborns with homozygous HbS. *Genet Med* 4, 328–335.
- Currat M, Trabuchet G, Rees D, et al. (2002). Molecular analysis of the β -globin gene cluster in the Niokholo Mandenka population reveals a recent origin of the β -S Senegal mutation. *Am J Hum Genet* 70, 207–223.
- Daar S, Hussain HM, Gravell D, Nagel RL, and Krishnamoorthy R. (2000). Genetic epidemiology of HbS in Oman: Multicentric origin for the β S gene. *Am J Hematol* 64, 39–46.
- Elderdyer AY, Mills J, Mohamed BA, et al. (2012). Molecular analysis of the B-globin gene cluster haplotypes in a Sudanese population with sickle cell anemia. *Int J Lab Hematol* 34, 262–266.
- Elhassan N, Gebremeskel EI, Elnour MA, et al. (2014). The episode of genetic drift defining the migration of humans out of Africa is derived from a large East African population size. *PloS One* 9, e97674.
- El-Hazmi MAF, Warsy AS, Bashir N, et al. (1999). Haplotypes of the β -globin gene as prognostic factors in sickle-cell disease. *East Mediterr Health J* 5, 1154–1158.
- Elion J, Berg PE, Lapoumeroulie C, et al. (1992). DNA sequence variation in a negative control region 5' to the beta-globin gene correlates with the phenotypic expression of the beta s mutation. *Blood* 79, 787–792.
- El-Kalla S, and Baysal E. (1998). Genotype-phenotype correlation of sickle cell disease in the United Arab Emirates. *J Pediatr Hematol Oncol* 15, 237–242.
- Flint J, Harding RM, Boyce AJ, and Clegg JB. (1993). The population genetics of the haemoglobinopathies. *Baillieres Best Pract Res Clin Haematol* 6, 215–262.
- Fong C, Lizarralde-Iragorri MA, Rojas-Gallardo D, and 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, 494–497.
- Friedrich JR, Prá D, Maluf SW, et al. (2008). DNA damage in blood leukocytes of individuals with sickle cell disease treated with hydroxyurea. *Mutat Res Genet Toxicol Environ Mutagen* 649, 213–220.
- Ghansah A, Rockett KA, Clark TG, et al. (2012). Haplotype analyses of haemoglobin C and haemoglobin S and the dynamics of the evolutionary response to malaria in Kassena-Nankana District of Ghana. *PloS one* 7, e34565.
- Hanchard N, Elzein A, Trafford C, et al. (2007). Classical sickle beta-globin haplotypes exhibit a high degree of long range haplotype similarity in African and Afro-Caribbean populations. *BMC Genet* 8, 52.
- Hewitt R, Krause A, Goldman A, Campbell G, and Jenkins T. (1996). Beta-globin haplotype analysis suggests that a major source of Malagasy ancestry is derived from Bantu-speaking Negroids. *Am J Hum Genet* 58, 1303.
- Holloway K, Lawson VE, and Jeffreys AJ. (2006). Allelic recombination and de novo deletions in sperm in the human β -globin gene region. *Hum Mol Genet* 15, 1099–1111.
- Imen, M, Ikbel BMM, Leila C, et al. (2011). Restriction mapping of β S locus among Tunisian sickle cell patients. *Am J Hum Biol* 23, 815–819.
- Inati, A, Taher A, Bou Alawi WW, et al. (2003). β -Globin gene cluster haplotypes and HbF levels are not the only modulators of sickle cell disease in Lebanon. *Eur J Haematol* 70, 79–83.
- Jones-Lecointe A, Smith E, Romana M, et al. (2008). β -Globin gene cluster haplotypes and α -thalassemia in sickle cell disease patients from Trinidad. *Am J Hum Biol* 20, 342–344.
- Labie D, Pagnier J, Lapoumeroulie C, et al. (1985). Common haplotype dependency of high G gamma-globin gene expression and high HbF levels in beta-thalassemia and sickle cell anemia patients. *Proc Natl Acad Sci USA* 82, 2111–2114.
- Lapoumeroulie C, Dunda O, Ducrocq R, et al. (1992). A novel sickle cell mutation of yet another origin in Africa: the Cameroon type. *Hum Genet* 89, 333–337.
- Lavinha J, Goncalves J, Faustino P, et al. (1992). Importation route of the sickle cell trait into Portugal: Contribution of molecular epidemiology. *Hum Biol* 64, 891–901.
- Lette G, Sankaran VG, Bezerra MA, et al. (2008). DNA polymorphisms at the BCL11A, HBS1L-MYB, and beta-globin loci associate with fetal hemoglobin levels and pain crises in sickle cell disease. *Proc Natl Acad Sci USA* 105, 11869–11874.
- Liu L, Muralidhar S, Singh M, et al. (2009). High-density SNP genotyping to define β -globin locus haplotypes. *Blood Cells Mol Dis* 42, 16–24.
- Luz JAD, Sans M, Kimura EM, et al. (2006). Alpha-thalassemia, HbS, and beta-globin gene cluster haplotypes in two Afro-Uruguayan sub-populations from northern and southern Uruguay. *Genet Mol Biol* 29, 595–600.
- Magaña MT, Ongay Z, Tagle J, et al. (2002). Analysis of β -S and β -A genes in a Mexican population with African roots. *Blood Cells Mol Dis* 28, 121–126.
- Makani J, Menzel S, Nkya S, et al. (2011). Genetics of fetal hemoglobin in Tanzanian and British patients with sickle cell anemia. *Blood* 117, 1390–1392.
- Mohammed AO, Attalla B, Bashir FM, et al. (2006). Relationship of the sickle cell gene to the ethnic and geographic groups populating the Sudan. *Community Genet* 9, 113–120.
- Mouélé R, Pambou O, Feingold J, and Galactéros F. (1999). α -Thalassemia in Bantu population from Congo-Brazzaville: Its interaction with sickle cell anemia. *Hum Hered* 50, 118–125.
- Mpalampa L, Ndugwa CM, Ddungu H, and Idro R. (2012). Foetal haemoglobin and disease severity in sickle cell anemia patients in Kampala, Uganda. *BMC Hematol* 12, 11.

BETA-GLOBIN GENE HAPLOTYPES AMONG CAMEROONIANS & REVIEW OF GLOBAL DISTRIBUTION

- Mukherjee MB, Surve RR, Gangakhedkar RR, et al. (2004). β -Globin gene cluster haplotypes linked to the β S gene in western India. *Hemoglobin* 28, 157–161.
- Muniz A, Corral L, Alaez C, et al. (1995). Sick cell anemia and β -gene cluster haplotypes in Cuba. *Am J Hematol* 49, 163–164.
- Nagel RL, and Fleming AF. (1992). Genetic epidemiology of the beta S gene. *Baillieres Best Pract Res Clin Haematol* 5, 331–365.
- Neumann R, Lawson VE, and Jeffreys AJ. (2010). Dynamics and processes of copy number instability in human γ -globin genes. *Proc Natl Acad Sci USA* 107, 8304–8309.
- Njamnshi AK, Mbong EN, Wonkam A, et al. (2006) The epidemiology of stroke in sickle cell patients in Yaounde, Cameroon. *J Neurol Sci* 250, 79–84.
- Ojwang PJ, Ogada T, Beris P, et al. (1987). Haplotypes and α globin gene analyses in sickle cell anaemia patients from Kenya. *Br J Haematol* 65, 211–215.
- Öner C, Dimovski AJ, Olivieri NF, et al. (1992). β S haplotypes in various world populations. *Hum Genet* 89, 99–104.
- Pagnier J, Mears JG, Dunda-Belkhdja O, et al. (1984). Evidence for the multicentric origin of the sickle cell hemoglobin gene in Africa. *Proc Natl Acad Sci USA* 81, 1771–1773.
- Patrinos GP, Kollia P, and Papadakis MN. (2005). Molecular diagnosis of inherited disorders: Lessons from hemoglobinopathies. *Hum Mutat* 26, 399–412.
- Piel FB, Patil AP, Howes RE, et al. (2013). Global epidemiology of sickle haemoglobin in neonates: A contemporary geostatistical model-based map and population estimates. *Lancet* 381, 142–151.
- Rahimi Z, Karimi M, Haghshenass M, and Merat A. (2003). β -Globin gene cluster haplotypes in sickle cell patients from southwest Iran. *Am J Hematol* 74, 156–160.
- Rose JI, Usik VI, Marks AE, et al. (2011). The Nubian complex of Dhofar, Oman: An African Middle Stone Age industry in Southern Arabia. *PLoS One* 6, e28239.
- Ruffieux N, Njamnshi AK, Wonkam A, et al. (2013). Association between biological markers of sickle cell disease and cognitive functioning amongst Cameroonian children. *Child Neuropsychol* 19, 143–160.
- Saiki RK, Scharf S, Faloona F, et al. (1985). Enzymatic amplification of beta-globin genomic sequences and restriction site analysis for diagnosis of sickle cell anemia. *Science* 230, 1350–1354.
- Samarah F, Ayesh S, Athanasiou M, Christakis J, and Vavatsi N. (2009). β S-Globin gene cluster haplotypes in the West Bank of Palestine. *Hemoglobin* 33, 143–149.
- Sankaran VG, Xu J, Byron R, et al. (2011). A functional element necessary for fetal hemoglobin silencing. *N Engl J Med* 365, 807–814.
- Schiliro, G, Samperi P, Consalvo C, et al. (1992). Clinical, hematological, and molecular features in Sicilians with sickle cell disease. *Hemoglobin* 16, 469–480.
- Serjeant GR. (1989). Geography and the clinical picture of sickle cell disease. An overview. *Ann N Y Acad Sci* 565, 109–119.
- Sow A, Peterson E, Josifovska O, et al. (1995). Linkage disequilibrium of the Senegal haplotype with the β s gene in the Republic of Guinea. *Am J Hematol* 50, 301–303.
- Steinberg MH. (2009). Genetic etiologies for phenotypic diversity in sickle cell anemia. *Sci World J* 9, 46–67.
- Steinberg MH, Lu ZH, Nagel RL, et al. (1998). Hematological effects of atypical and Cameroon beta-globin gene haplotypes in adult sickle cell anemia. *Am J Hematol* 59, 121–126.
- Stine OC, Dover GJ, Zhu D, and Smith KD. (1992). The evolution of two West African populations. *J Mol Evol* 34, 336–344.
- Tishkoff SA, Reed FA, Friedlaender FR, et al. (2009). The genetic structure and history of Africans and African Americans. *Science* 324, 1035–1044.
- Veten FM, Abdelhamid IO, Meiloud GM, et al. (2012). Hb S [β 6 (A3) Glu→Val, GAG> GTG] and β -globin gene cluster haplotype distribution in Mauritania. *Hemoglobin* 36, 311–315.
- Wainscoat JS, Bell JI, Thein SL, et al. (1983). Multiple origins of the sickle mutation: Evidence from beta S globin gene cluster polymorphisms. *Mol Biol Med* 1, 191–197.
- Weatherall DJ, and Clegg JB. (2001). Inherited haemoglobin disorders: An increasing global health problem. *Bull World Health Organ* 79, 704–712.
- Wonkam A, Bitoungui VJN, Vorster AA, et al. (2014c). Association of variants at BCL11A and HBS1L-MYB with hemoglobin F and hospitalization rates among sickle cell patients in Cameroon. *PloS one* 9, e92506.
- Wonkam A, de Vries J, Royal CD, Ramesar R, and Angwafo FF. (2014d). Would you terminate a pregnancy affected by sickle cell disease? Analysis of views of patients in Cameroon. *J Med Ethics* 40, 615–620.
- Wonkam A, Mba CZ, Mbanya D, et al. (2014a). Psychosocial stressors of sickle cell disease on adult patients in Cameroon. *J Genet Couns* 23, 948–956.
- Wonkam A, Mba CZ, Mbanya D, et al. (2013). Psychosocial burden of sickle cell disease on parents with an affected child in Cameroon. *J Genet Couns* 23, 192–201.
- Wonkam A, Noubiap JJ, Djomou F, Fieggen K, Njock R, and Toure GB. (2013b) Aetiology of childhood hearing loss in Cameroon (sub-Saharan Africa). *Eur J Med Genet* 56, 20–25.
- World Bank. (2010). Education for all-fast track initiative: Support to the education sector. Report No. 48373-CM, 1–2
- Zago MA, Silva WA, Dalle B, et al. (2000). Atypical beta-S haplotypes are generated by diverse genetic mechanisms. *Am J Hematol* 63, 79–84.

Address correspondence to:
Dr. Ambroise Wonkam
Division of Human Genetics
Faculty of Health Sciences
University of Cape Town
Anzio Road Observatory
7925 Cape Town
Republic of South Africa

E-mail: ambroise.wonkam@uct.ac.za

Supplementary Data

SUPPLEMENTARY TABLE S1. RESTRICTION ENDONUCLEASE CUTTING PATTERNS THAT REPRESENT EACH OF THE FIVE *HBB* GENE HAPLOTYPES

<i>Enzymes</i> <i>Haplotypes</i>	<i>XmnI</i> (5' <i>G</i> γ)	<i>HindIII</i> (<i>G</i> γ)	<i>HindIII</i> (<i>A</i> γ)	<i>HincII</i> (3?' $\Psi\beta$)	<i>HinfI</i> (5' β)
Senegal	+	+	-	+	+
Bantu/Central African Republic	-	+	-	-	-
Cameroon	-	+	+	+	+
Benin	-	-	-	+	-
Arab-Indian	+	+	-	+	-

(+)=Cut by a specific restriction endonuclease; (-)=Is not cut by that specific restriction endonuclease.

3.1.2. Pule DG, Ngo Bitoungui VJ, Chemegni BC, Kengne AP, Antonarakis S and Wonkam A. Association between variants at *BCL11A* erythroid-specific enhancer and fetal hemoglobin levels among Sickle Cell Patients in Cameroon: Implications for future therapeutic interventions. *OMICS*. 2015 19(10): 627-31.

Abstract

Variants in *BCL11A* were previously associated with fetal hemoglobin (HbF) levels among Cameroonian sickle cell disease (SCD) patients, however explaining only ~2% of the variance. In the same patients, we have investigated the relationship between HbF and two SNPs in a *BCL11A* erythroid-specific enhancer (N = 626). Minor allele frequencies in rs7606173 and rs1427407 were 0.42 and 0.24, respectively. Both variants were significantly associated with HbF levels ($p = 3.11e-08$ and $p = 6.04e-06$, respectively) and explained 8% and 6.2% variations, respectively. These data have confirmed a stronger effect on HbF of genomic variations at the *BCL11A* erythroid-specific enhancer among patients with SCD in Cameroon, the first report on a West African population. The relevance of these findings is of prime importance because the disruption of this enhancer would alter *BCL11A* expression in erythroid precursors and thus HbF expression, while sparing the induced functional challenges of any alterations on the expression of this transcription factor in non-erythroid lineages, thus providing an attractive approach for new treatment strategies of SCD.

Nature of publication: Original full article

Journal/Publisher: OMICS: A Journal of Integrative Biology/Mary Ann Liebert; ISSN: 1536-2310

Candidate contribution: Conceived and designed the experiments, performed the experiments, wrote the paper, revised and approved the manuscript.

Co-authors contribution: Conceived and designed the experiments: **GP**, AW. Performed the experiments: **GP**, VJNB AW. Analysed the data: AW, APK, GL, SEA. Contributed reagents/materials/analysis tools: VJNB, BCC, APK, AW. Wrote the paper: **GP**, AW. Revised and approved the manuscript: **GP**, VJNB, BCC, APK, SEA, AW.

Association between Variants at *BCL11A* Erythroid-Specific Enhancer and Fetal Hemoglobin Levels among Sickle Cell Disease Patients in Cameroon: Implications for Future Therapeutic Interventions

Gift Dineo Pule,¹ Valentina Josiane Ngo Bitoungui,² Bernard Chetcha Chemegni,² Andre Pascal Kengne,³ Stylianos Antonarakis,⁴ and Ambroise Wonkam¹

Abstract

Variants in *BCL11A* were previously associated with fetal hemoglobin (HbF) levels among Cameroonian sickle cell disease (SCD) patients, however explaining only ~2% of the variance. In the same patients, we have investigated the relationship between HbF and two SNPs in a *BCL11A* erythroid-specific enhancer ($N=626$). Minor allele frequencies in rs7606173 and rs1427407 were 0.42 and 0.24, respectively. Both variants were significantly associated with HbF levels ($p=3.11e-08$ and $p=6.04e-06$, respectively) and explained 8% and 6.2% variations, respectively. These data have confirmed a stronger effect on HbF of genomic variations at the *BCL11A* erythroid-specific enhancer among patients with SCD in Cameroon, the first report on a West African population. The relevance of these findings is of prime importance because the disruption of this enhancer would alter *BCL11A* expression in erythroid precursors and thus HbF expression, while sparing the induced functional challenges of any alterations on the expression of this transcription factor in non-erythroid lineages, thus providing an attractive approach for new treatment strategies of SCD.

Introduction

ADULT FETAL HEMOGLOBIN (HbF) levels alter the morbidity of sickle cell disease (SCD) (Platt et al., 1994), by inhibiting sickle hemoglobin (HbS) polymerization and therefore reducing hemolytic anemia and tissue injury. The HbF levels are highly variable and inheritable, primarily through genetic variants at three principal loci, *BCL11A*, *HBS1L-MYB*, and *HBB* cluster, accounting for 10%–20% of HbF variation among SCD patients (Thein and Menzel, 2009). Specifically, variations in *BCL11A* have been shown to be amenable to therapeutic manipulation, leading to the reversal of SCD symptoms in mice models (Xu et al., 2011).

Functional studies have shown that the expression of the transcriptional repressor *BCL11A* is regulated by erythroid-specific enhancers that contain 3' DNase hypersensitive sites (DHS) located +62, +58, and +55 kb from the transcription initiation site of *BCL11A* (Bauer et al., 2013). Two SNPs of the enhancer element had a stronger association with HbF levels among African American SCD patients: rs1427407 in DHS +62 and rs7606173 in DHS +55. This is consistent with

the hypothesis that multiple functional SNPs act in combination to influence *BCL11A* regulation (Bauer et al 2013; Sebastiani et al., 2015).

We previously reported that the minor allele frequencies (MAFs) of selected variants in *BCL11A* were similar to those detected in African Americans (Lettre et al., 2008) and the relationships with HbF were significant, but explained less of the variance in expression (~2% vs. 10%) (Wonkam et al., 2014). Other genomic polymorphisms at this locus could better tag the relation with HbF level in some West African SCD patients. In the present study, we have investigated the relationship between HbF levels and rs1427407 and rs7606173 in the *BCL11A* erythroid-specific enhancer among 626 patients with SCD (HbSS) from Cameroon.

Methods

Study population and HbF measurement

The study was performed with the approval of the National Ethical Committee Ministry of Public Health, Republic of Cameroon (No 033/CNE/DNM/07), and the University of

¹Division of Human Genetics, Department of Medicine, Faculty of Health Sciences, University of Cape Town, Cape Town, South Africa.

²Faculty of Medicine and Biomedical Sciences, University of Yaoundé, Cameroon.

³Non-Communicable Diseases Research Unit, South African Medical Research Council, Cape Town, South Africa.

⁴Department of Genetic Medicine and Development, Faculty of Medicine, University of Geneva, Switzerland.

Cape Town, Faculty of Health Sciences Human Research Ethics Committee (HREC REF: 132/2010). Patients' recruitment was conducted at the Yaoundé Central Hospital and Laquintinie Hospital in Douala, Cameroon. Only clinically stable patients, 5 years or older, with no history of blood transfusion, hydroxyurea treatment or hospitalisation in the preceding 6 weeks were included and clinical events were prospectively collected (Table 1).

Complete routine blood counts of patients and hemoglobin electrophoresis were conducted upon arrival at the hospital. Two methods of HbF measurement were employed in this study for successive groups of patients; initially the alkali denaturation test (ADT) in 55.5% ($n = 347$) of the cohort, and subsequently high performance liquid chromatography (HPLC), when it became available.

In a previous report, we disaggregated the SCD patients sample, based on the HbF assessment technique (ADT vs. HPLC) and found that the significant associations with HbF levels, examined independently, were present in both subgroups in rs11886868 (*BCL11A*), rs4671393 (*BCL11A*), rs28384513 (*HMIP 1*), and rs9494142 (*HMIP 2*) (Wonkam et al., 2014). To further avoid any major influence on the outcomes of the present study, the association analyses were also corrected for the electrophoresis techniques.

Genotyping

HbS mutation and HBB haplotypes. DNA samples were extracted from peripheral blood following the manufacturer's instructions (Puregene Blood Kit, Qiagen®, Alameda, CA, USA). Molecular analysis for the presence of the sickle mutation was carried out on 200 ng DNA by PCR to amplify a 770 bp segment of the β -globin gene, followed

by Dde I restriction analysis of the PCR product (Saiki et al., 1985).

Five restriction fragment length polymorphism (RFLP) regions in the β -globin gene cluster were amplified using published primers and methods (Steinberg et al., 1998) to analyze the XmnI (5'G γ), HindIII (G γ), HindIII (A γ), HincII (3'' $\Psi\beta$), and HinfI (5' β) for the *HBB* haplotype background (Bitoungui et al., 2015).

Detection of 3.7 kb α -globin gene deletions. The 3.7 kb alpha-globin gene deletions were successfully screened using the expand-long template PCR (Roche®, UK) following the instructions reported by Tan et al. (2001) with some modifications previously published (Rumaney et al., 2014).

SNPs. For the 626 samples, multiplex SNaPshot PCR and capillary electrophoresis were used to genotype both rs1427407 and rs7606173 at the *BCL11A* erythroid-specific enhancer loci. Quality control of the capillary electrophoresis results was assessed by direct cycle sequencing in a subset of 10% randomly selected samples that confirmed the SNaPshot results. Details of the methods were previously reported (Wonkam et al., 2014).

Data analysis

Descriptive statistics were obtained for all quantitative data using SPSS (IBM, USA version 21.0). A chi-squared test, with one degree of freedom, was used to perform the Hardy-Weinberg Equilibrium (HWE) test on the SNPs genotype with the following results: rs7606173 HWE p -value was 0.534 ($X^2 = 0.845$) and rs1427407 HWE p -value was 0.0042 ($X^2 = 12.7$). To exclude the ascertainment bias due to

TABLE 1. COHORT DESCRIPTION

Variables		Mean \pm SD	Value range	Number of observations
Age (years)		17.3 \pm 9.9	5–53	628
Hematological indices	RBC ($10^{12}/L$)	2.3 \pm 3.1	1.1–2–6.5	626
	Hb (g/dl)	7.8 \pm 1.5	3.5–14.5	626
	MCV (fl)	84 \pm 10.1	59–117	626
	MCHC (g/dl)	34 \pm 3.7	21.5–54.3	626
	WBC ($10^9/l$)	14 \pm 5.9	2.9–49.8	626
	Lymphocytes ($10^9/l$)	5.9 \pm 2.9	1.9–22.6	626
	Monocytes ($10^9/l$)	1.5 \pm 1.0	0.1–11.9	626
	Platelets ($10^9/l$)	373 \pm 122	97–837	626
	HbA2 (%)	3.7 \pm 1.8	0–9.7	626
	HbF (%)	10.1 \pm 8.6	0–26.8	626
Clinical events	Vaso-occlusive crisis (No. /year)	2.9 \pm 3.4	0–30	593
	Consultations (No. /year)	2.8 \pm 3.5	0–29	591
	Hospitalisation (No. /year)	1.4 \pm 2.1	0–15	591
	Stroke	4.2%		25/599
3.7del Alpha-globin gene genotypes	$\alpha\alpha/\alpha\alpha$	59.3%		310/524 ^a
	$\alpha\alpha/\alpha3.7$	29.6%		156/524 ^a
	$\alpha3.7/\alpha3.7$	11.1%		58/524 ^a
<i>HBB</i> haplotypes	Benin/Benin	57%		307/541 ^a
	Benin/Cameroon	25%		137/541 ^a
	Benin/atypical	8%		45/541 ^a
	Cameroon/Cameroon	5%		2/541 ^a

^aNumber of individuals and not alleles.

BCL11A ERYTHROID ENHANCER IN SICKLE DISEASE

TABLE 2. DIFFERENTIAL HbF ASSOCIATION RESULTS FOR SELECTED SNPs AT THE *BCL11A* SPECIFIC ERYTHROID ENHANCER IN CAMEROONIAN SCD HbSS PATIENTS ($N=626$)

Locus Chr. 2	SNP	Position ^{&}	Allele change	MAF [§]	Effect Size [#] (SE)	Variance explained (%)	P values
<i>BCL11A</i> enhancer	rs1427407 (DHS +62)	60718043	C>A	0.26	0.119 (0.024)	6.2	6.04e-06
<i>BCL11A</i> enhancer	rs7606173 (DHS +55)	60725451	C>G	0.42	0.1 (0.025)	8	3.11e-08

[&]Chr, chromosome; Position on NCBI Build 36.1; [§]MAF, minor allele frequency; [#]Effect sizes and standard errors are given in standard deviation units for the minor allele.

technical genotyping reasons, about 10% ($N=60$) of the samples were Sanger-sequenced by an independent laboratory (Central Analytical Facilities, University of Stellenbosch) and confirmed all the SNP genotyping results.

The observed frequencies of the rs1427407 variant could be biased by the presence of a third T allele (MAF=0.048) in 92 samples, suggesting the presence of copy number variations (CNVs) that could be involved in the regulation of *BCL11A* in some individuals. Alternatively, the violation of HWE for rs1427407 could possibly be due to a selective effect of this HbF-promoting SNP that could potentially influence the course of SCD and ultimately the survival of some patients. These various observations and hypotheses deserve further investigations.

Owing to the skewed distribution of HbF level, we \log_{10} -transformed HbF values to approximate normal distribution, and then used these transformed values in the association analysis after correcting for age; gender, electrophoresis technique, history of transfusion, *HBB* haplotypes, and the 3.7 kb alpha-globin gene deletion genotypes. Using an additive genetic model, under a generalized linear regression framework, we investigated the relationship between the SNPs (rs1427407 and rs7606173) and HbF levels, using the R statistical package version 3.0.3 (The R Foundation for statistical computing, Vienna, Austria). Significance was set at 5%.

Results

Patients

Table 1 summarizes the description of the cohort. All patients were homozygous for the HbS mutation. Among

them, 50.3% ($n=316$) were female. After the analysis of 1082 chromosomes, Benin (74%; $n=799$) and Cameroon haplotypes (19%; $n=207$) were the most prevalent. In combination, the Benin/Benin haplotype represented 57% ($n=307$) and the Benin/Cameroon 25% ($n=137$) of the cohort (Table 1). Among 524 patients successfully genotyped for the 3.7 kb alpha-globin gene deletion, 40.7% ($n=214$) had co-inherited alpha-thalassaemia (Table 1).

SNPs in *BCL11A* hypersensitivity sites and HbF levels

In the present study, MAFs in both rs7606173 and rs1427407 variants were 0.42 and 0.24, respectively (Table 2). Both genetic variants were significantly associated with HbF levels ($p=3.11e-08$ and $p=6.04e-06$) (Fig. 1A) and explained 8% and 6.2% variations in HbF levels, respectively (Table 2). The relationship of both SNPs with HbF was not influenced by the electrophoresis techniques, alpha-thalassaemia genotypes, or *HBB* haplotypes. There were no significant associations between the two SNPs and other various hematological indices or clinical events, but rs7606173 that was associated with hemoglobin levels ($p=0.02$) and consultation rates ($p=0.04$).

Discussion

The present article is the first report on the importance of the SNPs rs1427407 and rs7606173 among West African SCD patients. The results are consistent with those reported among African Americans (Bauer et al., 2013) and patients from Saudi Arabia and India (Sebastiani et al., 2015) and Tanzania for rs1427407 (Ntairo et al., 2015). The results of this study are of particular importance for the local and global

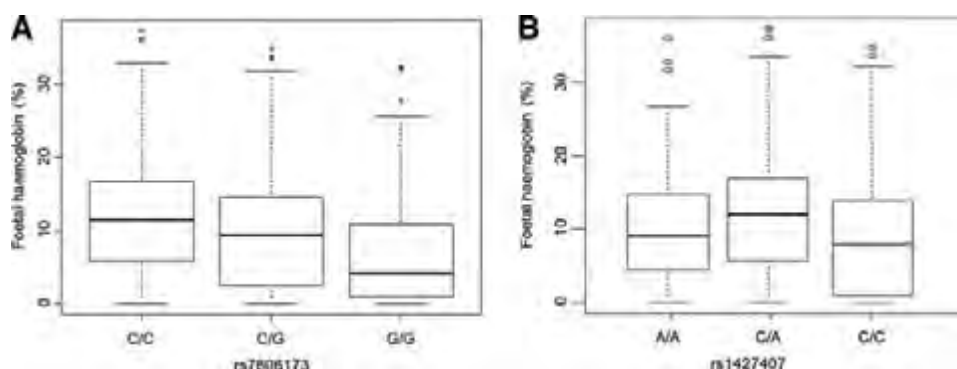


FIG. 1. Box-plots showing the distribution of foetal haemoglobin levels conditioned on SNP genotypes. (A) Fetal haemoglobin levels, conditioned on the rs7606173 SNP genotypes, and (B) conditioned on the rs1427407 SNP genotypes. Boxes have lines at the lower quartile, median, and upper quartile. The plot whiskers extend up and down from the median a distance 1.5 times the interquartile range of the boxes (truncated at zero where necessary). Outliers are the points outside the whiskers indicated as circles.

SCD scientific communities for the following reasons. First, nearly two-thirds of SCD patients globally are born and reside in Africa (Piel et al., 2013) and any potentially important SCD research performed on other populations should be replicated in African populations to assess and confirm their relevance for future care and therapeutic interventions.

Second, it is well known that Africans have the highest genetic variations and diversities (Gurdasani et al., 2015), implying some regional genetic specificities. For instance, we have previously shown that rs9399137, which acts as a tagging HbF-promoting SNP for the *HMIP-2* sub-locus in European populations (Thein et al., 2007), occurs at a low frequency in this Cameroonian cohort (Wonkam et al., 2014). In addition, a 3 bp deletion in perfect linkage disequilibrium (LD) with rs9399137 in Chinese, European, and some African American populations was reported to be the candidate functional motif for the *HMIP-2* region (Farell et al., 2011).

However, in these Cameroonian patients, rs9389269 was not linked to rs9399137 ($r^2 = 0$; $D' = 0.138$) (Wonkam et al., 2014). Moreover, we reported, in the *HMIP-2* sub-locus, a much higher frequency of rs9389269 (0.18) in Cameroonians as compared to Tanzanian SCD patients (0.03), an observation that could indicate a high degree of variation in the MAF of this SNP amongst SCD patients from various African population groups (Wonkam et al., 2014). These results emphasize the need to report on multiple African populations to fully capture the importance of various disease-modifying SNPs.

Lastly, Sebastiani et al. (2015) have reported on the association between variants at *BCL11A* erythroid-specific enhancer and HbF levels in African American SCD patients. However, among self-identified African Americans, analysis of genomic admixture indicates that the median proportion of European ancestry is 18.5% with very large variation, ranging from an estimate of over 99% West African ancestry less than 1% (Bryc et al., 2010). Therefore, any results of genomic studies performed on African Americans are not always appropriate proxies of what could be expected among sub-Saharan Africans. The importance of replicating findings of HbF-promoting SNPs at *BCL11A* in various populations is thus crucial to identifying possible targets for therapeutic intervention to increase HbF levels in patients with SCD.

Nevertheless, a major clinical caution of exploring the decreased expression of *BCL11A*, in order to promote HbF expression, rests on the other deleterious functional implications of this transcription factor, whose reduced expression could have serious developmental and neurological consequences. Indeed, patients described with microdeletions encompassing *BCL11A* display neurodevelopmental abnormalities including intellectual disability and brain malformation (Balci et al., 2015; Funnell et al., 2015; Peter et al., 2014). This emphasizes the importance of the *BCL11A* erythroid-specific enhancer being subject to genetic variations that influence HbF levels as reported in the present study.

Bioinformatics analysis suggests that rs7606173 in DHS +55 changes a binding site for *MZF1*, a gene actively transcribed in hematopoietic cells (Sebastiani et al., 2015) and most importantly, functional analysis have shown that the disruption of this enhancer would impair *BCL11A* expression in erythroid precursors and thus HbF expression, while conserving the “undruggable” nature of this transcription factor in non-erythroid lineages (Bauer et al., 2013).

It is also possible that, in the future, genome editing could allow the introduction of beneficial SNPs at the *BCL11A* erythroid-specific enhancer locus to improve the clinical course of SCD patients. Preliminary proof of concept was shown by experiments that introduced the $-175>$ point mutation and successfully elevated HbF into erythroid cell lines (Wienert et al., 2015). In addition, using a specific guide, RNA and Cas9, authors recently corrected one allele of the SCD *HBB* gene in human erythroid cells (Huang et al., 2015).

A limitation of the present research is the number of SNPs that were studied. Additional genotyping of SNPs will allow the definition of various haplotypes that will best capture the combinatorial importance of *BCL11A* elements (Bae et al., 2012).

Conclusion and Perspectives

The present study validates the relationship between HbF levels and SNPs in a *BCL11A* erythroid-specific enhancer among SCD patients in Cameroon, the first report on a West African cohort. The relevance of this finding is of prime importance because the disruption of this enhancer would impair *BCL11A* expression in erythroid precursors and thus HbF expression, while sparing the induced various functional challenges of any alterations of this transcription factor expression in non-erythroid lineages, thus providing an attractive approach for future treatment of SCD.

Acknowledgments

Conceived and designed the experiments: GP, AW. Performed the experiments: GP, VJNB, AW. Analyzed the data: AW, APK, GL, SEA. Contributed reagents/materials/analysis tools: VJNB, BCC, APK, AW. Wrote the article: GP, AW. Revised and approved the manuscript: GP, VJNB, BCC, APK, SEA, AW.

Funding. The molecular experiments of the study were funded by the National Health Laboratory Services (NHLS), South Africa; and the NIH, USA, Grant Number 1U01HG007459-01. The students' bursaries were supported by the Third World Academy of Sciences (TWAS), Oppenheimer Memorial Trust, the National Research Foundation, and FirstRand Laurie Dippenaar Scholarship, South Africa. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Author Disclosure Statement

The authors have declared that no competing financial interests exist.

References

- Bae HT, Baldwin CT, Sebastiani P, et al. (2012). Meta-analysis of 2040 sickle cell anemia patients: *BCL11A* and *HBS1L-MYB* are the major modifiers of HbF in African Americans. *Blood* 120, 1961–1962.
- Balci TB, Sawyer SL, Davila J, Humphreys P, and Dymont DA. (2015). Brain malformations in a patient with deletion 2p16.1: A refinement of the phenotype to *BCL11A*. *Eur J Med Genet* 58, 351–354.
- Bauer DE, Kamran SC, Lessard S, et al. (2013). An erythroid enhancer of *BCL11A* subject to genetic variation determines fetal hemoglobin level. *Science* 342, 253–257.

BCL11A ERYTHROID ENHANCER IN SICKLE DISEASE

- Bitoungui VJ, Pule GD, Hanchard N, Ngogang J, and 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, 171–179.
- Bryc K, Auton A, Nelson MR, et al. (2010). Genome-wide patterns of population structure and admixture in West Africans and African Americans. *Proc Natl Acad Sci USA* 107, 786–791.
- Farrell JJ, Sherva RM, Chen ZY, et al. (2011). A 3-bp deletion in the *HBS1L-MYB* intergenic region on chromosome 6q23 is associated with HbF expression. *Blood* 117, 4935–4945.
- Funnell AP, Prontera P, Ottaviani V, et al. (2015). 2p15-p16.1 microdeletions encompassing and proximal to *BCL11A* are associated with elevated HbF in addition to neurologic impairment. *Blood* 126, 89–93.
- Gurdasani D, Carstensen T, Tekola-Ayele F, et al. (2015). The African Genome Variation Project shapes medical genetics in Africa. *Nature* 517, 327–332.
- Huang X, Wang Y, Yan W, et al. (2015). Production of gene-corrected adult beta globin protein in human erythrocytes differentiated from patient iPSCs after genome editing of the sickle point mutation. *Stem Cells* 33, 1470–1479.
- Lette G, Sankaran VG, Bezerra MA, et al. (2008). DNA polymorphisms at the *BCL11A*, *HBS1L-MYB*, and beta-globin loci associate with fetal hemoglobin levels and pain crises in sickle cell disease. *Proc Natl Acad Sci USA* 105, 11869–11874.
- Mtatiro SN, Singh T, Rooks H, et al. (2014). Genome wide association study of fetal hemoglobin in sickle cell anemia in Tanzania. *PLoS One* 9, e111464.
- Peter B, Matsushita M, Oda K, and Raskind W. (2014). De novo microdeletion of *BCL11A* is associated with severe speech sound disorder. *Am J Med Genet A* 164A, 2091–2096.
- Piel FB, Patil AP, Howes RE, et al. (2013). Global epidemiology of sickle haemoglobin in neonates: A contemporary geostatistical model-based map and population estimates. *Lancet* 381, 142–151.
- Platt OS, Brambilla DJ, Rosse WF, et al. (1994). Mortality in sickle cell disease. Life expectancy and risk factors for early death. *N Engl J Med* 330, 1639–1644.
- Rumaney MB, Ngo Bitoungui VJ, Vorster AA, et al. (2014). The co-inheritance of alpha-thalassemia and sickle cell anemia is associated with better hematological indices and lower consultations rate in Cameroonian patients and could improve their survival. *PLoS One* 9, e100516.
- Saiki RK, Scharf S, Faloona F, et al. (1985). Enzymatic amplification of beta-globin genomic sequences and restriction site analysis for diagnosis of sickle cell anemia. *Science* 230, 1350–1354.
- Sebastiani P, Farrell JJ, Alsultan A, et al. (2015). *BCL11A* enhancer haplotypes and fetal hemoglobin in sickle cell anemia. *Blood Cells Mol Dis* 54, 224–230.
- Steinberg MH, Lu ZH, Nagel RL, et al. (1998). Hematological effects of atypical and Cameroon beta-globin gene haplotypes in adult sickle cell anemia. *Am J Hematol* 59, 121–126.
- Tan AS, Quah TC, Low PS, and Chong SS. (2011). A rapid and reliable 7-deletion multiplex polymerase chain reaction assay for alpha-thalassemia. *Blood* 98, 250–251.
- Thein SL, and Menzel S. (2009). Discovering the genetics underlying foetal haemoglobin production in adults. *Br J Haematol* 145, 455–467.
- Thein SL, Menzel S, Peng X, et al. (2007). Intergenic variants of *HBS1L-MYB* are responsible for a major quantitative trait locus on chromosome 6q23 influencing fetal hemoglobin levels in adults. *Proc Natl Acad Sci USA* 104, 11346–11351.
- Wienert B, Funnell AP, Norton LJ, et al. (2015). Editing the genome to introduce a beneficial naturally occurring mutation associated with increased fetal globin. *Nat Commun* 6, 7085.
- Wonkam A, Ngo Bitoungui VJ, Vorster AA, et al. (2014). Association of variants at *BCL11A* and *HBS1L-MYB* with hemoglobin F and hospitalization rates among sickle cell patients in Cameroon. *PLoS One* 9, e92506.
- Xu J, Peng C, Sankaran VG, et al. (2011). Correction of sickle cell disease in adult mice by interference with fetal hemoglobin silencing. *Science* 334, 993–996.

Address correspondence to:
Professor Ambroise Wonkam
Division of Human Genetics
Faculty of Health Sciences
University of Cape Town
Anzio Road Observatory
7925 Cape Town
Republic of South Africa

E-mail: ambroise.wonkam@uct.ac.za

3.1.3. Pule DG, Bitoungui VJN, Chemegni BC, Kengne AP and Wonkam A. Study of novel variants associated with fetal hemoglobin in Sardinians and Tanzanians among Sickle Cell Disease patients from Cameroon. Submitted to *Hemoglobin* (in review) 2016.

Abstract

Adult fetal hemoglobin (HbF) has been shown to ameliorate symptoms of Sickle Cell Anemia (SCA) and therefore genetic loci, signalling pathways and therapeutic agents with the potential of increasing HbF levels have been subject to investigation. Collectively, several studies have yielded a myriad of genetic polymorphisms associated with HbF in various populations, however the cross-population association of some of these polymorphisms is yet to be investigated. In this study, seventeen (17) select SNPs previously associated with HbF in a genome-wide association study (GWAS) of SCA patients in Tanzania and whole genome sequencing of the general population in Sardinia were investigated for similar association in Cameroon SCD patients. No associations were observed between any of the SNPs and HbF levels, other hematological indices or clinical events. The present study further demonstrates the population-specificity of the complex, heterogeneous regulation of HbF, thus it is imperative to identify and validate HbF-promoting loci across several populations of varying genetic and environmental backgrounds through replication studies such as this.

Nature of publication: Original full article

Journal/Publisher: Hemoglobin; ISSN: 0363-0269

Candidate contribution: Conceived and designed the experiments, performed the experiments, analysed the data, wrote the paper, revised and approved the manuscript.

Co-authors contribution: Conceived and designed the experiments: **GP**, AW. Performed the experiments: **GP**. Patient recruitment, sample and clinical data collection and processing: VJNB, BCC. Analysed the data: APK, **GP**, AW. Contributed reagents/materials/analysis tools: VJNB, BCC, APK, AW. Wrote the paper: **GP**, AW. Revised and approved the manuscript: **GP**, VJNB, BCC, APK, AW.

SHORT COMMUNICATION

Studies of novel variants associated with Hb F in Sardinians and Tanzanians in sickle cell disease patients from Cameroon

Gift D. Pule^a, Valentina J. Ngo Bitoungui^b, Bernard Chetcha Chemegni^b, Andre P. Kengne^c and Ambroise Wonkam^a

^aDivision of Human Genetics, Department of Pathology, Faculty of Health Sciences, University of Cape Town, Cape Town, South Africa;

^bDepartment of Microbiology, Hematology, Parasitology et de Infectious Pathologies, Faculty of Medicine and Biomedical Sciences, University of Yaoundé, Cameroon; ^cNon-Communicable Diseases Research Unit, South African Medical Research Council, Cape Town, South Africa

ABSTRACT

High level of Hb F has been shown to improve survival in sickle cell disease. Among 453 Cameroonians with sickle cell disease, we have investigated 18 selected single-nucleotide polymorphisms (SNPs) in novel and suggestive loci associated with Hb F level identified through a genomewide association study in sickle cell disease patients in Tanzania, and whole-genome sequencing of a population from Sardinia. Seven of 10 variants reported in Sardinians were either monomorphic or very rare in the Cameroonians. No associations were observed with any SNPs and Hb F levels in Cameroonians affected by sickle cell disease. The present study illustrates the complexity of replicating Hb F-promoting variants association results across populations.

ARTICLE HISTORY

Received 14 June 2016
Revised 28 September 2016
Accepted 28 September 2016
Published online 9 November 2016

KEYWORDS

Cameroon; Hb F; Sardinia; sickle cell disease; Tanzania



Adult Hb F has been shown to be a highly variable quantitative trait that can alter the morbidity of sickle cell disease by reducing hemolytic anemia and organ damage, through the inhibition of Hb S (*HBB*: c.20A > T) polymerization.(1) Hb F level is an inheritable trait associated with genetic polymorphisms at three principal loci, *BCL11A*, *HBSIL-MYB* and the β -globin gene cluster, which accounts for 10.0–20.0% variance in Hb F in sickle cell disease patients.(2–5) Additional loci that account for more variation in Hb F levels are still to be found. Recently, a genomewide association study (GWAS) of Hb F in sickle cell disease patients in Tanzania revealed eight additional loci suggestive of being novel, with associations that reached a genome-wide significance of $p < 10^{-6}$.(6) In addition, 10 other Hb F-associated novel loci were recently identified through whole-genome sequencing studies of a population in Sardinia.(7) In order to possibly replicate these findings in sickle cell disease patients from Cameroon, we investigated in the present study the association with Hb F level with 18 select single-nucleotide polymorphisms (SNPs), from the Hb F-promoting novel loci found in Tanzanians and Sardinians.


The study was approved by the University of Cape Town, Faculty of Health Sciences Human Research Ethics Committee (HREC REF: 132/2010), Cape Town, South Africa and the National Ethics Committee Ministry of Public Health, Yaoundé, Republic of Cameroon (No. 033/CNE/DNM/07). Patients were recruited at the Yaoundé Central Hospital and Laquintinie Hospital in Douala, Cameroon. Only clinically stable patients, aged 5 years or older, with no

history of hydroxyurea treatment, blood transfusion or hospitalization in the preceding 6 weeks were included, and clinical events were prospectively collected (Table 1). Routine blood counts of patients and hemoglobin (Hb) electrophoresis were conducted on arrival at the hospital, initially using the alkali denaturation test (ADT) in 55.5% ($n = 249$) of the cohort and subsequently high-performance liquid chromatography (HPLC), when it became available.(4) In a previous report, we disaggregated the Cameroonian sickle cell disease patients' sample, based on the Hb F assessment technique (ADT vs. HPLC) and found that the significant associations with Hb F levels, examined independently, were present in both subgroups in rs11886868 (*BCL11A*), rs4671393 (*BCL11A*), rs28384513 (*HMIP-1*) and rs9494142 (*HMIP-2*).(4) In the present study, the association analyses were also corrected for the electrophoresis techniques, as stated later.

DNA was extracted from peripheral blood following the manufacturer's instructions (Puregene Blood Kit; Qiagen GmbH, Hilden, Germany). Molecular analysis to determine the presence of the sickle mutation was carried out on 200 ng DNA by polymerase chain reaction (PCR) to amplify a 770 bp segment of the β -globin gene, followed by *DdeI* restriction analysis of the PCR product.(8)

Using published primers and methods, five restriction fragment length polymorphism (RFLP) sites in the β -globin gene cluster were amplified to analyze the *XmnI* ($5^{\prime}G\gamma$), *HindIII* ($^C\gamma$), *HindIII* ($^A\gamma$), *HincII* ($3^{\prime}\psi\beta$) and *HinfI* ($5^{\prime}\beta$) for the *HBB* haplotype background (Supplementary Table 1).(9)

CONTACT Ambroise Wonkam  ambroise.wonkam@uct.ac.za  Division of Human Genetics, Faculty of Health Sciences, University of Cape Town, Anzio Road, Observatory, 7925, Cape Town, Republic of South Africa.

 Supplemental data for this article can be accessed [here](#)

© 2016 Informa UK Limited, trading as Taylor & Francis Group

Table 1. Description of the studied cohort.

Variables	Parameters	Mean \pm SD	Value ranges	Numbers observed
Hematological indices	Age (years)	17.9 \pm 10.5	5.0–58.0	453
	RBC ($10^{12}/L$)	2.7 \pm 0.7	12.0–5.5	453
	Hb (g/dL)	7.8 \pm 1.6	4.0–14.5	453
	MCV (fl)	84.9 \pm 10.0	61.0–112.0	453
	MCHC (g/dL)	34.1 \pm 3.8	23.3–54.0	453
	WBC ($10^9/L$)	14.2 \pm 6.1	3.5–49.8	453
	Lymphocytes ($10^9/L$)	6.0 \pm 3.1	0.87–20.4	453
	Monocytes ($10^9/L$)	1.6 \pm 1.2	0.1–11.9	453
	Platelets ($10^9/L$)	370.4 \pm 130.6	108.0–837.0	453
	Hb A ₂ (%)	3.8 \pm 1.9	0.1–18.2	453
	Hb F (%)	10.4 \pm 8.7	0.0–37.4	453
	Clinical events	VOC crises (n/year)	2.9 \pm 3.5	0–40
Consultations (n/year)		3.1 \pm 3.8	0–40	453
Hospitalization (n/year)		1.5 \pm 2.6	0–30	453
Overt strokes (%)		4.9		22/453 ^a
α Genotypes (%)	$\alpha\alpha/\alpha\alpha$	63.0		285/453 ^a
	$\alpha\alpha/-\alpha^{3.7}$	28.6		130/453 ^a
	$-\alpha^{3.7}/-\alpha^{3.7}$	8.4		38/453 ^a
β Haplotypes (%)	Benin/Benin	50.5		229/453 ^a
	Benin/Cameroon	22.5		103/453 ^a
	Benin/Atypical	4.2		19/453 ^a
	Cameroon/Cameroon	4.2		19/453 ^a

Atypical: the restriction endonuclease cutting patterns were variable and did not fit with the five classical haplotypes; Benin: [- - - + -]; Cameroon: [- + + + +]; Hb: hemoglobin; MCV: mean corpuscular volume; MCHC: mean corpuscular Hb concentration; RBC: red blood cell counts; VOC: vasocclusive crises; WBC: white blood cell counts. (See haplotype classification in Supplementary Table 1).

^aNumber of individuals not alleles.

Using the expand-long template PCR (Roche Diagnostics, Basel, Switzerland), the 3.7 kb α -globin gene deletion was successfully screened following a previously reported protocol.(10)

All variants were analyzed using the iPLEX Gold Sequenom Mass Genotyping Array (Inqaba Biotec, Pretoria, South Africa). The results were validated by direct cycle sequencing of a subset of the sample ($n = 45$; 10.0%) using previously reported methods.(4)

Statistical analyses

Descriptive statistics were obtained for all quantitative data using the Statistical Package for the Social Sciences (version 21.0; SPSS Inc., IBM, Chicago, IL). A χ^2 with one degree of freedom, was used to perform the Hardy–Weinberg equilibrium (HWE) test on the SNP genotype with five SNPs (rs11754265, rs62573842, rs7163278, rs10468869 and rs10756993) out of HWE ($p < .05$). To correct for the skewness of the Hb distribution, we log10-transformed and normalized the data to obtain the quantitative trait used in the association analysis (after correcting for age, gender, and electrophoresis technique, transfusion history, α -thalassemia (α -thal) genotypes and haplotypes in the β -globin-like gene clusters). Statistical models to investigate the relationship between Hb F-associated SNPs and sickle cell disease complications were conducted using an additive genetic model, under a generalized linear regression framework, in the R statistical package version 3.0.3 (The R Foundation for Statistical Computing, Vienna, Austria). Significance was set at 5.0%.

A total of 453 patients were included; Table 1 summarizes their description. All participants had carried homozygous Hb S; 50.7% ($n = 230$) were female; the mean age was

17.9 \pm 10.5 years. Overt stroke was reported in 4.9% ($n = 22$) of the patients and the mean Hb F level was 10.4 \pm 8.7%. The most prevalent β -globin gene cluster haplotype was Benin [- - - + -] (Supplementary Table 1) (73.0%, $n = 661$ chromosomes); 28.6 and 8.4% of patients had co-inherited a single or double 3.7 kb α -globin gene deletion, respectively.

The minor allele frequencies of all the variants investigated are shown in Table 2. The majority of variants (7/10) reported from the Sardinian population were either monomorphic ($n = 4$) or very rare ($n = 3$) in the Cameroonian sickle cell disease patients. One out of eight variants from the Tanzanian cohort was monomorphic. No associations were observed between the SNPs and Hb F levels in Cameroonian affected by sickle cell disease (Table 2) nor were there any associations between these variants and any other hematological indices or clinical events.

This study did not replicate, among sickle cell disease patients from Cameroon, the suggestive novel variants associated with Hb F in populations from Tanzania (6) and novel variants identified in a population from Sardinia.(7) Similarly, the suggestive novel variants reported in the Tanzanians were not replicated in a British cohort of largely Caribbean/West African descent, which the authors attributed to the small sample size or different ancestry;(6) the latter reason further supports the considerable heterogeneity of genetic variants within African populations.(11) This difference was also emphasized by the fact that rs192197462 was present as a monomorphic polymorphism in this study, although significantly associated with Hb F levels in sickle cell disease patients from Tanzania. We also previously reported, in the *HMIP-2* sublocus, a much higher frequency of rs9389269 (0.18) in Cameroonian as compared to the Tanzanian sickle cell disease patients (0.03).(4) Furthermore,

Table 2. Association between selected single nucleotide polymorphisms and Hb F in sickle cell disease patients in Cameroon.

SNPs	Gene	Chromosome loci	Alleles	MAF	Effect size (SE)	p Value
Variants reported in the general population in Sardinia (7)						
X12_123681790	<i>MPHOSPH9</i> intron	X12_123681790	A ^a	NA	NA	NA
X16_391593	α -globin gene cluster	X16_391593	T/C	0.001	0.27 (-1.59–2.14)	.774
rs113267280	<i>CCND3</i>	6:41952511	T/G	0.001	0.34 (-1.54–2.21)	.725
rs117544265	<i>HBS1L</i>	6:135356216	C/G	0.256	0.058 (-0.11–0.23)	.501
rs141494605	<i>HBM</i>	16:216593	T ^a	NA	NA	NA
rs148706947	<i>AXIN1</i>	16:342218	C/T	0.002	0.82 (-0.50–2.15)	.220
rs183437571	<i>NFIX</i>	19:13121899	C ^a	NA	NA	NA
rs570013781	α -globin gene cluster	16:149539	G ^a	NA	NA	NA
rs67104793	<i>XRN1</i>	3:142444839	A/delA	0.243	-0.05 (-0.21–0.10)	.504
rs59329875	<i>CTSA</i>	20:44547672	C/T	0.244	-0.09 (-0.21–0.06)	.257
Variants reported in sickle cell disease patients in Tanzania (6)						
rs62573842	intergenic	9:32264314	A/G	0.216	0.13 (-0.02–0.27)	.089
rs6466533	<i>LOC105375366</i>	7:79163576	T/C	0.237	0.12 (-0.03–0.28)	.112
rs6590706	<i>OPCML</i>	11:133465011	A/G	0.314	0.02 (-0.11–0.15)	.749
rs7163278	<i>LOC105370983</i>	15:93345162	T/C	0.052	-0.06 (-0.30–0.19)	.656
rs10468869	intergenic	18:51795403	A/G	0.360	-0.01 (-0.14–0.12)	.867
rs10756993	<i>ADAMTSL1</i>	9:18839726	C/A	0.174	-0.11 (-0.30–0.08)	.243
rs192197462	intergenic	14:57116065	A ^a	NA	NA	NA
rs28567737	<i>CCDC61/MIR769</i>	19:46531219	C/T	0.471	-0.14 (-0.29–0.10)	1.000

Alleles: major/minor; MAF: minor allele frequency; NA: not applicable. SE: effect size and range; SNPs: single nucleotide polymorphisms.

^aMonomorphic.

Hb F regulation is a highly heterogeneous system and subject to regulatory and signaling pathways that develop differently across different populations.(12)

Indeed, the majority of variants described in Sardinians (7/10) were either monomorphic or very rare in Cameroonians with sickle cell disease (Table 2) and could be attributed to the founder nature of the Sardinian population used. This founder effect could account for some variants that are extremely rare in the rest of the world; in some cases, reaching high enough frequencies to provide unexpected biological insights.(7) For instance, X12_123681790, monomorphic in Cameroonians, was associated with *HBA1* levels in Sardinians. Similarly, rs113267280 located in the *CCND3* gene, a regulator of erythropoiesis particularly cell division during terminal erythroid differentiation,(13) was also associated with *HBA1* levels in Sardinians. Furthermore, rs183437571 was also imputed as monomorphic in the TwinsUK cohort,(14) suggesting that significant variations also exist within European populations. The data further highlight the complexity of Hb F genomic regulation and illustrate the difficulties that will be encountered when attempting to replicate some variant association results.

Taken together, the data suggest that the discovery of new Hb F-promoting loci that are relevant in most populations will require integrated large-scale genome-wide studies across many populations. Specifically, future studies should include multiple African populations, from where the disease burden is highest and also because of the highest genetic diversity among Africans.(11) Moreover, the use in future research of appropriate GWAS panels that capture the full genomic diversity in African populations, one of which is in development within the H3Africa Consortium (www.h3Africa.com), could provide more insight into the missing heritability of the persistence of Hb F production. Indeed, in a recent GWAS study among African-Americans, using the Illumina Omni1-Quad System (San Diego, CA) that contains high density of SNP coverage of the human genome (1,140,419 SNPs) with 76.0% coverage of polymorphisms validated

in people of African ancestry, novel candidate SNPs associated with Hb F levels were reported in the *SPARC*, *GJCI*, *EFTUD2* and *JAZF1* genes.(15) Further exploration of known genetic modifiers of Hb F expression such as *BCL11A*, *HBS1L-MYB*, β -globin gene cluster,(2–5) Krüppel-like factor 1 (*KLF-1*), (haplo-insufficiency),(16,17) *SARI* (18) and *SALL2*,(19) could serve as central nodes for gene network analyses to identify other pathways, genes and variants critical for Hb F production. Possible targets for novel therapeutic agents could be potent Hb F regulators such as *KLF-1*, known activator of *BCL11A*, via transcriptional and post-transcriptional mechanisms.(20)

The present study was a replication, in Cameroonian sickle cell disease patients, of putative novel Hb F-promoting loci identified in Tanzanian sickle cell disease patients and the general Sardinian population: no significant associations with Hb F were observed at these loci. The lack of association further demonstrates the complexity of finding additional Hb F-promoting loci that have multipopulation relevance. These and other similar observations call for more multicenter GWAS with SNPs panels that are representative of interpopulations genetic diversity, particularly that within sub-Saharan Africa, where the disease burden is highest.

Disclosure statement

The authors report no conflicts of interest. The authors alone are responsible for the content and writing of this article.

Funding

The molecular experiments of the study were funded by the National Health Laboratory Services (NHLS), Johannesburg, South Africa, and the National Institutes of Health, Bethesda, MD, USA, 10.13039/1000000051 [grant number 1U01HG007459–01]. The student's bursary was provided by the Oppenheimer Memorial Trust, Johannesburg, South Africa, National Research Foundation, Cape Town, South Africa, and FirstRand Laurie Dippenaar Scholarship, Cape Town, South Africa. These organizations had no role in the study design, data collection and analysis, decision to publish, or preparation of the manuscript.

References

1. Platt OS, Brambilla DJ, Rosse WF, *et al.* Mortality in sickle cell disease. Life expectancy and risk factors for early death. *N Engl J Med.* 1994;330:1639–1644.
2. Thein SL, Menzel S, Lathrop M, Garner C. Control of fetal hemoglobin: new insights emerging from genomics and clinical implications. *Hum Mol Genet.* 2009;18:R216–R223.
3. Lettre G, Sankaran VG, Bezerra MA, *et al.* DNA polymorphisms at the BCL11A, HBS1L-MYB, and beta-globin loci associate with fetal hemoglobin levels and pain crises in sickle cell disease. *Proc Natl Acad Sci USA.* 2008;105:11869–11874.
4. Wonkam A, Bitoungui VJN, Vorster AA, *et al.* Association of variants at BCL11A and HBS1L-MYB with hemoglobin F and hospitalization rates among sickle cell patients in Cameroon. *PLoS One.* 2014;9:e92506.
5. Makani J, Menzel S, Nkya S, *et al.* Genetics of fetal hemoglobin in Tanzanian and British patients with sickle cell anemia. *Blood.* 2011;117:1390–1392.
6. Mtatiro SN, Singh T, Rooks H, *et al.* Genome wide association study of fetal hemoglobin in sickle cell anemia in Tanzania. *PLoS One.* 2014;9:e111464.
7. Danjou F, Zoledziewska M, Sidore C, *et al.* Genome-wide association analyses based on whole-genome sequencing in Sardinia provide insights into regulation of hemoglobin levels. *Nat Genet.* 2015;47:1264–1271.
8. Saiki RK, Scharf S, Faloona F, *et al.* Enzymatic amplification of beta-globin genomic sequences and restriction site analysis for diagnosis of sickle cell anemia. *Science.* 1985;230:1350–1354.
9. Bitoungui VJN, Pule GD, Hanchard N, *et al.* β -Globin gene haplotypes among cameroonians and review of the global distribution: is there a case for a single sickle mutation origin in Africa? *OMICS.* 2015;19:171–179.
10. Rumaney MB, Bitoungui VJN, Vorster AA, *et al.* The co-inheritance of α -thalassemia and sickle cell anemia is associated with better hematological indices and lower consultations rate in Cameroonian patients and could improve their survival. *PLoS One.* 2014;9:e100516.
11. Gurdasani D, Carstensen T, Tekola-Ayele F, *et al.* The African genome variation project shapes medical genetics in Africa. *Nature.* 2015;517:327–332.
12. Solovieff N, Milton JN, Hartley SW, *et al.* Fetal hemoglobin in sickle cell anemia: genome-wide association studies suggest a regulatory region in the 5' olfactory receptor gene cluster. *Blood.* 2010;115:1815–1822.
13. Kozar K, Ciemerych MA, Rebel VI, *et al.* Mouse development and cell proliferation in the absence of D-cyclins. *Cell.* 2014;118:477–491.
14. Moayyeri A, Hammond CJ, Valdes AM, Spector TD. Cohort profile: TwinsUK and healthy ageing twin study. *Int J Epidemiol.* 2013;42:76–85.
15. Liu L, Pertsemlidis A, Ding LH, *et al.* Original research: a case-control genome-wide association study identifies genetic modifiers of fetal hemoglobin in sickle cell disease. *Exp Biol Med (Maywood).* 2016;241:706–718.
16. Borg J, Papadopoulos P, Georgitsi M, *et al.* Haploinsufficiency for the erythroid transcription factor KLF1 causes hereditary persistence of fetal hemoglobin. *Nat Genet.* 2010;42:801–805.
17. Perkins A, Xu X, Higgs DR, *et al.* Krüppeling erythropoiesis: an unexpected broad spectrum of human red blood cell disorders due to KLF1 variants. *Blood.* 2016;127:1856–1862.
18. Zhu J, Chin K, Aerbajinai W, *et al.* Hydroxyurea-inducible SARI gene acts through the $G\alpha/JNK/Jun$ pathway to regulate γ -globin expression. *Blood.* 2014;124:1146–1156.
19. Sheehan VA, Luo Z, Flanagan JM, *et al.* Genetic modifiers of sickle cell anemia in the BABY HUG cohort: influence on laboratory and clinical phenotypes. *Am J Hematol.* 2013;88:571–576.
20. Pule DG, Mowla S, Novitzky N, Wonkam A. Hydroxyurea down-regulates BCL11A, KLF-1 and MYB through miRNA-mediated actions to induce γ -globin expression: implications for new therapeutic approaches in sickle cell disease. *Clin Trans Med.* 2016;5:15–31.

3.1.4. Pule DG, Bitoungui VJN, Chemegni BC, Kengne AP and Wonkam A. SAR1a promoter polymorphisms are not associated with fetal hemoglobin in Cameroon SCD patients. Submitted to *BMC Research Notes (revised manuscript submitted)* 2016.

Abstract

Reactivation of adult hemoglobin (HbF) is currently the dominant therapeutic approach to ameliorating hemoglobinopathies and thalassemias. Hydroxyurea (HU) is the only FDA-approved drug treatment for Sickle Cell Disease (SCD) and has been shown to be effective in HbF induction, increased erythropoiesis, reduced anemia and coagulation and local nitric oxide release. Several studies have collectively identified a myriad of genetic variants associated with baseline and/or HU-induced HbF, however the cross-population validity of some of these novel associations is yet to be elucidated. In this study, the minor allele frequency and relationship with baseline HbF of four variants (rs2310991; rs4282891; rs76901216 and rs76901220) in the HU-inducible small guanosine triphosphate (GTP)-binding protein, secretion-associated and RAS-related (*SAR1a*) protein that were previously associated with HbF after HU treatment in African American patients, were investigated in 484 SCD patients from Cameroon. No associations were observed between any of the promoter variants and baseline HbF, clinical events or other hematological indices. The results of this study demonstrate the population-specificity of the heterogeneous mechanisms regulating HbF production. Therefore, it is critical to validate all novel associations across several populations particularly in sub-Saharan Africa, where the disease burden is highest and the genetic and environmental backgrounds are most varied.

Nature of publication: Original full article

Journal/Publisher: International Journal of Laboratory Hematology; ISSN: 1751-553X

Candidate contribution: Conceived and designed the experiments, performed the experiments, analysed the data, wrote the paper, revised and approved the manuscript.

Co-authors contribution: Conceived and designed the experiments: **GP**, AW. Performed the experiments: **GP**. Patient recruitment, sample and clinical data collection and processing: VJNB, BCC. Analysed the data: **GP**, APK, AW. Contributed reagents/materials/analysis tools: VJNB, BCC, APK, AW. Wrote the paper: **GP**, AW. Revised and approved the manuscript: **GP**, VJNB, BCC, APK, AW.

SAR1a promoter polymorphisms are not associated with fetal hemoglobin in patients with Sickle Cell Disease from Cameroon

Gift Dineo Pule¹, Valentina Josiane Ngo Bitoungui², Bernard Chetcha Chemegni², Andre Pascal Kengne³ and Ambroise Wonkam¹

¹Division of Human Genetics, Department of Pathology, Faculty of Health Sciences, University of Cape Town, Cape Town, South Africa; ²Faculty of Medicine and Biomedical Sciences, University of Yaoundé, Cameroon; ³Non-Communicable Diseases Research Unit, South African Medical Research Council, Cape Town, South Africa.

Corresponding author: Professor Ambroise Wonkam, Division of Human Genetics, Faculty of Health Sciences, University of Cape Town, Anzio Road, Observatory, 7925, Cape Town, Republic of South Africa. Tel: +27214066307; Fax: +27214066826 e-mail: ambroise.wonkam@uct.ac.za

Pule DG: giftpule@outlook.com

Bitoungui VJN: vngobitoungui@yahoo.fr

Chemengi BC: chetcha@yahoo.fr

Pascal AK: andre.kengne@mrc.ac.za

Wonkam A: ambroise.wonkam@uct.ac.za

Short running title: *SAR1a* promoter SNPs and HbF in Cameroonian SCD patients

Word count: 1311

Abstract

Background: Reactivation of adult hemoglobin (HbF) is currently a dominant therapeutic approach to Sickle Cell Disease (SCD). In this study, we have investigated among SCD patients from Cameroon, the association of HbF level and variants in the HU-inducible small guanosine triphosphate (GTP)-binding protein, secretion-associated and RAS-related (*SAR1a*) protein, previously shown to be associated with HbF after HU treatment in African American SCD patients.

Findings: Only patients >5 years old were included; hemoglobin electrophoresis and a full blood count were conducted upon arrival at the hospital. RFLP-PCR was used to describe the *HBB* gene haplotypes and Gap PCR to investigate the 3.7 kb α -globin gene deletion. The iPLEX Gold Sequenom Mass Genotyping Array and cycle sequencing were used for the genotyping of four selected SNPs in *SAR1a* (rs2310991; rs4282891; rs76901216 and rs76901220). Genetic analysis was performed using an additive genetic model, under a generalized linear regression framework. 484 patients were studied. No associations were observed between any of the promoter variants and baseline HbF, clinical events or other hematological indices.

Conclusion: The results of this study could be explained by possible population-specificity of some tagging genomic variants associated with HbF production and illustrated the complexity of replicating HbF-promoting variants association results across African populations.

Key words: *SAR1a* promoter; fetal hemoglobin; Sickle Cell Disease; Cameroon

Background

Reactivation of adult fetal hemoglobin (HbF) is currently the dominant approach for the treatment of Sickle Cell Disease (SCD). The expression of adult HbF is a quantitative trait that is subject to several predisposing loci affecting the persistence of HbF in adulthood, particularly three principal loci; *BCL11A*, *HBS1L-MYB* intergenic variants and the five sequence polymorphisms along the β -globin gene cluster that confer the SCD haplotype (1-3). Taken in sum, these loci have been associated with the disease-ameliorating HbF and account for 10-20% of the variance (4-7). Furthermore, the *BCL11A* erythroid-specific enhancer variants have been shown to account for significant variance in HbF in African American (8,9), Tanzanian (10) and Cameroonian SCD patient cohorts (11).

HbF response to hydroxyurea (HU), has been shown to be subject to a myriad of genetic variations (SNPs, signalling pathways and pharmacogenomics interactions) and environmental factors (socio-economic factors, quality of care, exposure to malaria and infections) (12). Furthermore, some of these variants have been associated with favourable pharmacologic response to HU treatment like the small guanosine triphosphate (GTP)-binding protein, secretion-associated and RAS-related (SAR) protein (13). SAR proteins have been shown to be critical to γ -globin expression (14) via the $G_{i\alpha}$ /JNK/Jun pathway (15) in response to HU. Four variants (rs2310991; rs4282891; rs76901216 and rs76901220) in the *SAR1a* promoter were associated with significant increases in HbF levels in African American SCD patients on HU for 2 years (13). In this present study, we have investigated the relationship between the four *SAR1a* promoter variants and baseline HbF in SCD patients from Cameroon, without any HU treatment.

Findings

Research hypothesis

In this study, we hypothesize that selected variants in *SAR1a* promoter associated with significant increases in HbF levels in African American SCD patients on HU, are also associated with baseline HbF in SCD patients from Cameroon, without any HU treatment.

Methods

Study population and HbF measurement

Patients' recruitment occurred at the Yaoundé Central Hospital and Laquintinie Hospital in Douala, Cameroon. Only clinically stable patients, 5 years and older, with no history of blood transfusion, HU treatment, or hospitalisation in the preceding 6 weeks were included and clinical events were prospectively collected (Table 1). Whole blood counts of patients and Hb electrophoresis were conducted on arrival at the hospital, initially using the alkali denaturation test (ADT) in 55.5% ($n = 266$) of the cohort, and when it became available, high performance liquid chromatography (HPLC). In a previous study of this cohort, the SCD patient cohort was disaggregated based on the technique used for HbF assessment (HPLC vs ADT) and found similar but independently examined associations between HbF levels and *BCL11A* and *HBS1L-MYB* intergenic variants (5).

Genotyping

HbS mutation and *HBB* haplotypes

DNA was extracted from peripheral blood following the manufacturer's instructions (Puregene Blood Kit, Qiagen®, USA). Molecular analysis to determine the presence of the sickle mutation was carried out on 200 ng DNA by PCR to amplify a 770 bp segment of the β -globin gene, followed by Dde I restriction analysis of the PCR product (16).

Using published primers and methods, five restriction fragment length polymorphism (RFLP) sites in the β -globin gene cluster were amplified (17) to analyse the XmnI (5'G γ), HindIII (G γ), HindIII (A γ), HincII (3'' $\Psi\beta$) and HinfI (5' β) for the *HBB* haplotype background (18).

Detection of 3.7 kb α -globin gene deletions

Using the expand-long template PCR (Roche®, UK), the 3.7 kb α -globin gene deletion was successfully screened following the instructions reported (19) with some modifications previously published (20).

SNPs

A total of 484 samples were analysed initially using SNaPshot sequencing and capillary electrophoresis ($n = 234$), then the rest analysed using the iPLEX Gold Sequenom Mass Genotyping Array (Inqaba Biotech, South Africa). The results were validated by direct cycle sequencing of a subset ($n = 48$; 10%) of the samples using previously reported methods (5).

Data analysis

Descriptive statistics were obtained for all quantitative data using SPSS (IBM, USA version 21.0). A chi-squared test, with 1 degree of freedom, was used to perform the Hardy-Weinberg Equilibrium (HWE) test on the SNPs genotype with three SNPs (rs4282891; rs2310991 and rs76901216) out of HWE ($p < 0.05$). Using an additive genetic model, under a generalized linear regression framework, we investigated the relationship between the SNPs and HbF levels, using the R statistical package version 3.0.3 (The R Foundation for statistical computing, Vienna, Austria). Significance was set at 5%.

Results

Patient cohort

Table 1 summarises the main characteristics of the cohort. All patients had confirmed SCA diagnosis (HbSS) among whom 50.2% ($n = 243$) were female and the mean age of the cohort was 17.9 years (± 10.6). After genotyping β -globin gene haplotypes for 968 chromosomes, the most prevalent were Benin (72.5%, $n = 702$) and Cameroon (19.2%, $n = 186$). Haplotypes given in combinations, the Benin/Benin haplotype was 50.2% ($n = 243$) and the Benin/Cameroon haplotype 21.7% ($n = 105$) of the patient cohort (Table 1). The frequency of the 3.7kb α -globin gene deletion ($\alpha 3.7$) was 22.5% ($n = 222$) of 968 chromosomes, where 28.5% and 8.6% of patients had co-inherited a single ($\alpha\alpha/\alpha 3.7$) and double ($\alpha 3.7/\alpha 3.7$) deletions, respectively (Table 1). The average number of annual reported vaso-occlusive crises was 2.8 (± 3.3) with a similar yearly rate of hospital consultation (3.0 ± 3.9) and mean 1.4 (± 2.5) hospitalization per year. Overt stroke was reported in 4.5% ($n = 22$) of the patients.

No association between SNPs and HbF

The minor allele frequency (MAF) of the *SAR1a* promoter SNPs is shown in Table 2. There was no association between the four selected promoter SNPs and HbF levels in the patient cohort neither was there an association with clinical events or hematological indices. The influence of the electrophoretic technique, α -globin genotypes and β -globin haplotypes on the relationship between the SNPs and HbF levels was tested and no significant effect was observed.

Discussion

Four variants (rs2310991; rs4282891; rs76901220 and rs76901216) have previously been associated with higher percent HbF and significant change in HbF levels after HU treatment for two years in African American SCA patients enrolled in the National Institutes of Health (NIH)'s Sickle Cell Pulmonary Hypertension Screening Study (ClinicalTrials.org; NCT00011648). It has been reported that *SAR1* is induced by HU towards the production of

HbF in erythroid cells (14,15) via activation of the $G_i\alpha$ /JNK/Jun pathway. This and other critical signalling pathways in HU-induced HbF as well as genomic variants associated with HbS mutation and the disease course have been reviewed (12). Given the responsiveness of *SAR1* to HU treatment and prior associations with HU-induced HbF, four promoter polymorphisms were investigated in a cohort of SCA patients not treated with HU to determine possible associations between the SNPs and baseline HbF. If these variants result in differential expression of *SAR1* and thus HbF, this loci could become a candidate for therapeutic manipulation for ameliorating SCA and other hemoglobinopathies. However, none of the selected SNPs were associated with HbF, other hematological indices or clinical events. Furthermore, the difference between the minor allele frequencies (MAF) in African American SCA patients (13) and results from the present study may be attributed to the well reported high diversity among African population (21,22) and more specifically to admixture among African Americans, with African ancestry that could vary from 1 to 99 % (23,24). Replication of association studies across different populations, particularly of varied genetic backgrounds and environmental settings, is imperative to understanding the complex processes of HbF production and genetic polymorphisms predisposing to persistence of adult HbF, more so in sub-Saharan Africa where the burden of disease is highest.

Conclusion

The present study did not find any association between selected SNPs in the HU-inducible SAR1a protein, and HbF among SCD patients from Cameroon. Although SAR1a is a known regulator of HbF expression under HU therapy, it is interesting to know, however, that the SAR1a protein does not participate in HbF regulation at steady state in SCD patients. These findings emphasize the genetic heterogeneity of African populations affected by SCD with

regard to HbF-promoting loci and the need to perform replication studies of key association findings in several SCD populations to fully capture their significance.

List of abbreviations

HbF: hemoglobin F; **SCD**: Sickle Cell Disease; **HU**: Hydroxyurea; **GTP**: Guanosine triphosphate; **SAR**: secretion-associated and RAS-related; **RFLP-PCR**: Restriction fragment-length polymorphism polymerase chain reaction; **HBB**: Hemoglobin beta; **SNP**: Single nucleotide polymorphism; **BCL11A**: B-cell lymphoma/leukemia 11A; **HBS1L**: Hemoglobin S-1 like translational protein; **MYB**: myeloblastosis protein family; **G α /JNK/Jun**: mitogen-activated protein kinase pathway; **ADT**: Alkaline denaturation test; **HPLC**: High performance liquid chromatography; **HWE**: Hardy-Weinberg equilibrium; **MAF**: Minor allele frequency; **NIH**: National Institutes of Health.

List of tables

Table 1. Cohort description

Table 2. Association between SAR1a promoter polymorphisms and baseline HbF in Cameroon SCD patients

Ethics approval and consent

Approval for the study was provided by the University of Cape Town, Faculty of Health Sciences Human Research Ethics Committee (HREC REF: 132/2010) and the National Ethical Committee Ministry of Public Health, Republic of Cameroon (No 033/CNE/DNM/07). All participants gave informed and signed consent.

Consent for publication

Not applicable.

Availability of data and materials

Data will not be shared due to ongoing research and analysis.

Competing Interests

The authors have declared that no competing interests exist.

Funding

The molecular experiments of the study were funded by the National Health Laboratory Services (NHLS), South Africa; and the NIH, USA, grant number 1U01HG007459-01. The student's bursary was provided by the Oppenheimer Memorial Trust, National Research Foundation, and FirstRand Laurie Dippenaar Scholarship, South Africa. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Authors' contributions

Conceived and designed the experiments: GP, AW. Performed the experiments: GP. Patient recruitment, sample and clinical data collection and processing: VJNB, BCC. Analysed the data: APK, AW. Contributed reagents/materials/analysis tools: VJNB, BCC, APK, AW. Wrote the paper: GP, AW. Revised and approved the manuscript: GP, VJNB, BCC, APK, AW.

Acknowledgements

We would like to thank the Division of Human Genetics and the University of Cape Town.

References

- (1) Labie D, Pagnier J, Lapoumeroulie C, Rouabhi F, Dunda-Belkhodja O, Chardin P, et al. Common haplotype dependency of high G gamma-globin gene expression and high Hb F levels in beta-thalassemia and sickle cell anemia patients. *Proc Natl Acad Sci U S A* 1985;**82**(7):2111-2114.
- (2) Creary LE, Ulug P, Menzel S, McKenzie CA, Hanchard NA, Taylor V, et al. Genetic variation on chromosome 6 influences F cell levels in healthy individuals of African descent and HbF levels in sickle cell patients. *PLoS One* 2009;**4**(1):e4218.
- (3) Menzel S, Garner C, Gut I, Matsuda F, Yamaguchi M, Heath S, et al. A QTL influencing F cell production maps to a gene encoding a zinc-finger protein on chromosome 2p15. *Nat Genet* 2007;**39**(10):1197-1199.
- (4) Thein SL, Menzel S, Lathrop M, Garner C. Control of fetal hemoglobin: new insights emerging from genomics and clinical implications. *Hum Mol Genet* 2009;**18**(R2):R216-23.
- (5) Wonkam A, Bitoungui VJN, Vorster AA, Ramesar R, Cooper RS, Tayo B, et al. Association of variants at BCL11A and HBS1L-MYB with hemoglobin F and hospitalization rates among sickle cell patients in Cameroon. *PLoS One* 2014;**9**(3):e92506.
- (6) Lettre G, Sankaran VG, Bezerra MA, Araujo AS, Uda M, Sanna S, et al. DNA polymorphisms at the BCL11A, HBS1L-MYB, and beta-globin loci associate with fetal hemoglobin levels and pain crises in sickle cell disease. *Proc Natl Acad Sci U S A* 2008;**105**(33):11869-11874.

- (7) Makani J, Menzel S, Nkya S, Cox SE, Drasar E, Soka D, et al. Genetics of fetal hemoglobin in Tanzanian and British patients with sickle cell anemia. *Blood* 2011;**117**(4):1390-1392.
- (8) Bauer DE, Kamran SC, Lessard S, Xu J, Fujiwara Y, Lin C, et al. An erythroid enhancer of BCL11A subject to genetic variation determines fetal hemoglobin level. *Science* 2013;**342**(6155):253-257.
- (9) Galarneau G, Palmer CD, Sankaran VG, Orkin SH, Hirschhorn JN, Lettre G. Fine-mapping at three loci known to affect fetal hemoglobin levels explains additional genetic variation. *Nat Genet* 2010;**42**(12):1049-1051.
- (10) Mtatiro SN, Mgaya J, Singh T, Mariki H, Rooks H, Soka D, et al. Genetic association of fetal-hemoglobin levels in individuals with sickle cell disease in Tanzania maps to conserved regulatory elements within the MYB core enhancer. *BMC Med Genet* 2015;**16**(1):1.
- (11) Pule GD, Ngo Bitoungui VJ, Chetcha Chemegni B, Kengne AP, Antonarakis S, Wonkam A. 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* 2015;**19**(10):627-631.
- (12) Pule GD, Mowla S, Novitzky N, Wiysonge CS, Wonkam A. A systematic review of known mechanisms of hydroxyurea-induced fetal hemoglobin for treatment of sickle cell disease. *Expert Rev Hematol* 2015;**8**(5):669-679.
- (13) Kumkhaek C, Taylor JG 6, Zhu J, Hoppe C, Kato GJ, Rodgers GP. Fetal haemoglobin response to hydroxycarbamide treatment and sar1a promoter polymorphisms in sickle cell anaemia. *Br J Haematol* 2008;**141**(2):254-259.

- (14) Tang DC, Zhu J, Liu W, Chin K, Sun J, Chen L, et al. The hydroxyurea-induced small GTP-binding protein SAR modulates gamma-globin gene expression in human erythroid cells. *Blood* 2005;**106**(9):3256-3263.
- (15) Zhu J, Chin K, Aerbajinai W, Kumkhaek C, Li H, Rodgers GP. Hydroxyurea-inducible SAR1 gene acts through the Gialpha/JNK/Jun pathway to regulate gamma-globin expression. *Blood* 2014;**124**(7):1146-1156.
- (16) Saiki RK, Scharf S, Faloona F, Mullis KB, Horn GT, Erlich HA, et al. Enzymatic amplification of beta-globin genomic sequences and restriction site analysis for diagnosis of sickle cell anemia. *Science* 1985;**230**(4732):1350-1354.
- (17) Steinberg M, Lu Z, Nagel R, Venkataramani S, Milner P, Huey L, et al. Hematological effects of atypical and cameroon β -globin gene haplotypes in adult sickle cell anemia. *Am J Hematol* 1998;**59**(2):121-126.
- (18) Bitoungui VJN, Pule GD, Hanchard N, Ngogang J, Wonkam A. 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* 2015;**19**(3):171-179.
- (19) Tan AS, Quah TC, Low PS, Chong SS. A rapid and reliable 7-deletion multiplex polymerase chain reaction assay for alpha-thalassemia. *Blood* 2001;**98**(1):250-251.
- (20) Rumaney MB, Bitoungui VJN, Vorster AA, Ramesar R, Kengne AP, Ngogang J, et al. The Co-Inheritance of Alpha-Thalassemia and Sickle Cell Anemia Is Associated with Better Hematological Indices and Lower Consultations Rate in Cameroonian Patients and Could Improve Their Survival. *PLoS One* 2014;**9**(6):e100516.
- (21) Sanchez-Mazas A, Poloni ES. Genetic diversity in Africa. *eLS* 2008.

(22) Gurdasani D, Carstensen T, Tekola-Ayele F, Pagani L, Tachmazidou I, Hatzikotoulas K, et al. The African genome variation project shapes medical genetics in Africa. *Nature* 2015;**517**(7534):327-332.

(23) Shriner D, Kumkhaek C, Doumatey AP, Chen G, Bentley AR, Charles BA, et al. Evolutionary context for the association of γ -globin, serum uric acid, and hypertension in African Americans. *BMC Med Genet* 2015;**16**(1):1.

(24) Bryc K, Auton A, Nelson MR, Oksenberg JR, Hauser SL, Williams S, et al. Genome-wide patterns of population structure and admixture in West Africans and African Americans. *Proc Natl Acad Sci U S A* 2010;**107**(2):786-791.

Table 1: Cohort description

Variables		Mean \pm SD	Value range	Number of observations
Age (years)		17.9 \pm 10.6	5 – 57	484
Haematological indices	RBC ($10^{12}/L$)	2.8 \pm 0.7	1.2 - 5.5	484
	Hb (g/dl)	7.8 \pm 1.6	4.0 - 13.5	484
	MCV(fl)	84.6 \pm 10.0	61 - 112	484
	MCHC (g/dl)	34.2 \pm 3.8	22.3 - 54.0	484
	WBC ($10^9/l$)	14.7 \pm 6.4	3.5 - 48.8	484
	Lymphocytes ($10^9/l$)	6.1 \pm 3.3	0.87 - 20.4	484
	Monocytes ($10^9/l$)	1.7 \pm 1.4	0.1 - 10.9	484
	Platelets ($10^9/l$)	371.5 \pm 131.6	108 - 827	484
	HbA2 (%)	3.6 \pm 2.1	0.1 - 18.2	484
HbF (%)	10.1 \pm 8.5	0 – 37.4	484	
Clinical events	Vaso-Occlusive crisis (No. /year)	2.8 \pm 3.3	0 - 40	484
	Consultations (No. /year)	3.0 \pm 3.9	0 - 24	484
	Hospitalisation (No. /year)	1.4 \pm 2.5	0 - 30	484
	Overt stroke	4.5%		22 / 484
3.7del α-globin gene genotypes	$\alpha\alpha/\alpha\alpha$	62.9%		304 / 484 ^c
	$\alpha\alpha/\alpha3.7$	28.5%		138 / 484 ^c
	$\alpha3.7/\alpha3.7$	8.6%		42 / 484 ^c
β-globin gene haplotypes	Benin/Benin	50.2%		243 / 484 ^c
	Benin/Cameroon	21.7%		105 / 484 ^c
	Benin/Atypical	4.1%		20 / 484 ^c
	Cameroon/Cameroon	4.1%		20 / 484 ^c

^c Number of individuals and not alleles; RBC: red blood cells; Hb: hemoglobin; MCV: mean corpuscular volume; MCHC: mean corpuscular haemoglobin concentration; WBC: white blood cells; HbA2: hemoglobin α 2; HbF: hemoglobin F.

Table 2. No association between SAR1a promoter polymorphisms and baseline HbF in Cameroon SCD patients

SNP	SNP position	Chromosome loci	Alleles [#]	MAF ^{\$}	Effect Size ^{&} (SE)	P-value
<i>rs2310991</i>	-1377	10:70171890	A/C	0.334	-0.05 (-0.14 to 0.05)	0.335
<i>rs4282891</i>	Intron 1 +100	10:70170313	A/C	0.336	0.03 (-0.08 to 0.13)	0.645
<i>rs76901216</i>	-809	71601084	C/G	0.295	-0.07 (-0.20 to 0.05)	0.265
<i>rs76901220</i>	-385	10: 71600660	G*	-	-	-

[#] Alleles: major/minor, ^{\$}MAF: minor allele frequency, *monomorphic, [&]Effect size and range.

Chapter 3.2 – Results

Mechanisms of hydroxyurea-induced γ -globin expression in Sickle Cell Disease

Synopsis of Chapter 3.2

Hydroxyurea (HU) is the only available, FDA approved, drug treatment for Sickle Cell Disease and has demonstrated clinical efficacy and amelioration of disease symptoms, mostly through increasing fetal hemoglobin (HbF). A comprehensive understanding of the mechanisms of HbF-induction by HU treatment is the main focus of this chapter, which includes three publications that;

- i) Represent the background of this chapter and thorough systematic review of all currently known mechanisms by which HU promotes expression of HbF towards treatment and amelioration of SCD.
- ii) Investigates the transcriptional and post-transcriptional pathways involved in the production of HbF in response to HU in primary erythroid cells derived from umbilical cord blood hematopoietic stem cells.
- iii) Comment on the availability of various treatment approaches for SCD in Africa compared to the West and argues for a cross-continent effort from government, private sector and research institutes towards improving access to good medical care for patients.

3.2.1. Pule GD, Mowla S, Novitzky N, Wiysonge C and Wonkam A. Systematic review of known mechanisms of hydroxyurea-induced fetal hemoglobin for treatment of Sickle Cell Disease. *Expert Review of Hematology*. 2015 8(5): 669-79.

Abstract

To report on molecular mechanisms of fetal hemoglobin (HbF) induction by hydroxyurea (HU) for the treatment of sickle cell disease. Study Design: Systematic review. Results: Studies have provided consistent associations between genomic variations in HbF-promoting loci and variable HbF level in response to HU. Numerous signal transduction pathways have been implicated, through the identification of key genomic variants in *BCL11A*, *HBS1L-MYB*, *SAR1* or *Xmnl* polymorphism that predispose the response to the treatment, and signal transduction pathways that modulate γ -globin expression (cAMP/cGMP; $G\alpha/c$ -Jun N-terminal kinase/Jun; methylation and miRNA). Three main molecular pathways have been reported: i) Epigenetic modifications, transcriptional

events and signaling pathways involved in HU-mediated response, ii) Signaling pathways involving HU-mediated response and iii) Post-transcriptional pathways (regulation by miRNAs). Conclusions: The complete picture of HU-mediated mechanisms of HbF production in Sickle Cell Disease remains elusive. Research on post-transcriptional mechanisms could lead to therapeutic targets that may minimize alterations to the cellular transcriptome.

Nature of publication: Original full article

Journal/Publisher: Expert Review of Hematology; ISSN: 1747-4094

Candidate contribution: Performed the main search of the articles, drafted the manuscript and compiled the revisions

Co-authors contribution: **GP** performed the main search of the articles, drafted the manuscript and compiled the revisions. **SM**, **NN** completed the search and checked the content and appropriateness of the molecular mechanisms reported. **CW** contributed to the concept and refined the methodology. **AW** conceived the manuscript concept, performed the secondary search, checked the contents of all the articles included and supervised the revisions. All the authors revised and approved the final version of the manuscript.

EXPERT
REVIEWS

A systematic review of known mechanisms of hydroxyurea-induced fetal hemoglobin for treatment of sickle cell disease

Expert Rev. Hematol. Early online, 1–11 (2015)

Gift D Pule¹,
Shaheen Mowla²,
Nicolas Novitzky²,
Charles S Wiysonge³
and
Ambroise Wonkam*¹

¹Department of Medicine, Division of Human Genetics, Faculty of Health Sciences, University of Cape Town, Cape Town, Republic of South Africa

²Department of Clinical Laboratory Sciences, Division of Haematology, Faculty of Health Sciences, University of Cape Town, Cape Town, Republic of South Africa

³Centre of Evidence-based Health Care, Stellenbosch University, Cape Town, Republic of South Africa

*Author for correspondence:

Tel.: +27 214 066 307

Fax: +27 214 066 826

ambroise.wonkam@uct.ac.za

Aim: To report on molecular mechanisms of fetal hemoglobin (HbF) induction by hydroxyurea (HU) for the treatment of sickle cell disease. **Study Design:** Systematic review. **Results:** Studies have provided consistent associations between genomic variations in HbF-promoting loci and variable HbF level in response to HU. Numerous signal transduction pathways have been implicated, through the identification of key genomic variants in *BCL11A*, *HBS1L-MYB*, *SAR1* or *XmnI* polymorphism that predispose the response to the treatment, and signal transduction pathways that modulate γ -globin expression (cAMP/cGMP; $G\alpha/c$ -Jun N-terminal kinase/Jun; methylation and miRNA). Three main molecular pathways have been reported: i) Epigenetic modifications, transcriptional events and signaling pathways involved in HU-mediated response, ii) Signaling pathways involving HU-mediated response and iii) Post-transcriptional pathways (regulation by miRNAs). **Conclusions:** The complete picture of HU-mediated mechanisms of HbF production in Sickle Cell Disease remains elusive. Research on post-transcriptional mechanisms could lead to therapeutic targets that may minimize alterations to the cellular transcriptome.

KEYWORDS: BCL11A • fetal hemoglobin • HBS1L-MYB and SAR1 • hydroxyurea • molecular mechanism • sickle cell disease

Sickle cell disease (SCD) is a monogenic, hematological and multi-organ disorder associated with progressive organ damage and chronic and acute illness (Wheatherall *et al.*, 2005). SCD is caused by a point mutation (A > T) in the sixth codon of the β -globin gene on chromosome 11, resulting in the substitution of the amino acid glutamic acid to valine. The resulting hemoglobin S (HbS) leads to polymerization and precipitation of hemoglobin during deoxygenation or dehydration resulting in sickling, abnormal adhesion of leukocytes and platelets, inflammation, hypercoagulation, hemolysis and hypoxia, in addition to microvascular obstruction and ultimately organ damage [1].

There is strong correlation between the frequency of the HbS gene and the historical distribution and incidences of malaria [2]. It is

estimated that > 300,000 births with SCD occur annually, nearly two-third of which take place in Africa [3]. SCD is relatively common in other continents such as North America and Europe with 2600 and 1300 affected newborns annually, respectively [4]. It is well accepted that the sickle mutation exists in Africa on diverse genetic haplotype backgrounds [5]. Five typical haplotypes have been described across the β -globin gene cluster based upon the pattern of specific restriction fragment-length polymorphisms across the region. Four haplotypes are associated with HbS in Africa (Benin, Bantu/Central African Republic, Senegal and Cameroon) and the fifth is thought to have arisen in India and/or the Arabian Peninsula (Arab/Hindu) [6,7]. It has been suggested that these haplotypes also have an effect on the severity of the

disease through their genetically determined effect on HbF level [8,9].

Indeed, despite the fact that there are several key phenotypes of SCD (anemia, stroke, infections), fetal hemoglobin (HbF) has emerged as a central disease modifier [10], that is amenable to therapeutic manipulation [11]. Genetic variants at three principal loci, *BCL11A*, *HBSIL-MYB* and *HBB* cluster, account for 10–20% of HbF variation [12–14]; among SCD patients in the US and Brazil [15], Tanzania [16] and Cameroon [17]. Currently, hydroxyurea (HU) is the only US FDA-approved pharmacologic treatment for induction of HbF in patients with SCD. The major HU benefit is directly related to its HbF-producing effect [18,19] that leads to significant reduction of pain, acute chest episodes, the need for blood transfusions and mortality [8,20–22].

A few studies have shown that individual responses to HU treatment are highly variable, with induced HbF levels ranging from 10% to > 30% HbF [23–25]; and sibling pairs analysis has shown that strong genetic component could influence this variable response to HU [20]. Some authors have hypothesized that the effect of HU on HbF level could act through HbF-promoting loci like *BCL11A* [26]. Recent GWAS studies have revealed further sequence variants that could influence the response to HU [27]. However, the precise molecular mode of action of HU, though these genetic variants, remained to be comprehensively determined.

In this paper, it is our aim to provide a comprehensive and systematic review of the current literature on HU's mode of action at the molecular levels and postulate on future work in this field. It is anticipated that adequate knowledge of the mode of action of HU will result in the exploration of alternative therapeutic agents that may minimize alterations to the cellular transcriptome.

Methods

A review of the current literature on reported mechanisms of HbF induction by HU treatment was conducted from October 2014 to December 2014 using Pubmed (National Library of Medicine), Medline and Google scholar. Key words included individual use or a combination of the following: 'Hydroxyurea or HU or hydroxycarbamide', 'Foetal hemoglobin or HbF or gamma globin', 'hemoglobin-induction', 'Sickle Cell treatment' and 'erythroid treatment', and specific author names were also used. Prior knowledge of research groups working on HU and SCD Africa and globally further facilitated the identification and selection of research articles. Only available full-length articles, in English, with the use of HU on erythroid cells (cell lines and/or primary cells) were selected. In cases where multiple studies reported a similar pathway, the most recent with the most detailed mechanisms of HbF induction was included. The main search was conducted by a PhD student in Human Genetics and reviewed by a Medical/Human Geneticist with expertise in SCD and a Hematologist/Cell biologist with expertise in signal transduction pathways. A total of 563 articles were consulted after the search, eliminating on the basis of the

article title and its relevance to the scope of the review. Subsequently, 361 papers were retrieved and their abstracts and results sections perused for further elimination, of which a final total of 129 papers were selected for inclusion in the review.

Results

SCD is highly heterogeneous in clinical manifestation and disease severity even though patients may be diagnosed with the same genetic mutation resulting in the disease. HbF is a major modifier of the SCD disease phenotype that has been associated with reduced risk for vaso-occlusive complications, organ damages and increased life expectancy [10,21,28–30]. During gestation, the γ -globin genes are the predominantly transcribed genes from the β -cluster. After birth the expression of the γ -globin genes is replaced by the adult β -globin genes; this process is known as the 'Fetal Switch' [31]. However, some HbF remains to be heterogeneously transcribed with variable expression levels across HbF red blood cells (F cells) [32]. Initial observations suggested that *HBB* haplotypes also have an effect on the severity of SCD [8,9]. The Bantu haplotype being the less favorable and the Indian-Arab haplotype associated with the least severity; as the latter is associated with higher HbF levels [33], due to a single nucleotide polymorphism (SNP) upstream of the β -globin gene (rs7482144 or Xmn- γ), characteristically present in Indian Arab and Senegal haplotypes [34]. The degree of HbF expression and the blood percentage of F cells at adulthood are heritable quantitative trait loci [35]. Normal individuals have <0.6% HbF distributed among 1–7% F cells in their peripheral blood. However, approximately 2% of the population produce up to 5% HbF and 25% F cells, a trait known as heterocellular Hereditary Persistence of HbF [36]. Heterocellular Hereditary Persistence of HbF is characterized by a persistent high level of HbF in adults [32] and a genome-wide SNP scan of 10 members of a Maltese family with heterocellular Hereditary Persistence of HbF revealed a nonsense mutation in the *KLF-1* gene, a critical activator of *BCL11A* [32]. This mutation was characterized and shown to cause an ablation of the DNA binding domain of this erythroid transcription factor and thus provided a rationale for the effect of *KLF-1* haploinsufficiency on HbF levels. In total, genetic variants at three principal loci, *BCL11A*, *HBSIL-MYB* and *HBB* cluster, account for 10–20% of HbF variation [13,14,37]. *BCL11A* loci have been shown to be amenable to therapeutic manipulation with the aim to increase HbF level [38,39] that have major implications for research on new treatments of SCD.

HU treatment & Hbf induction in SCD

Currently, HU is the only pharmacologic treatment for induction of HbF in patients with SCD both approved by the FDA in 1998 and by the European Medicines Agency in 2007. In addition, the National Institutes of Health Office of Medical Applications of Research and the Agency of Healthcare Research and Quality both declared HU as an effective drug treatment for adults and children with SCD [40–42]. HU is an

oral, S-phase specific cytotoxic, anti-metabolic and anti-neoplastic drug treatment; it is a potent inhibitor of a ubiquitous enzyme called ribonucleotide reductase [43]. In 1984, the first clinical application of HU in hemoglobinopathies demonstrated swift and vivid increases in HbF concentration within reticulocytes and insignificant toxicity to bone marrow [44]. Initial trial showed that HU was associated with decreases in the frequencies of painful episodes, acute chest syndrome, hospitalization and transfusions [28]. Subsequent trials in SCD have demonstrated clinical efficacy and increase in survival rates and life expectancy [20,21], protection against cerebrovascular disease [45], long-term drug safety, capacity to prevent organ damage and reduced morbidity and mortality in school-age children [24], toddlers [46,47] and infants [48]. HU has also been associated with clinical drift, where physicians use the drug for related complications of SCD such as stroke prevention, priapism and pulmonary hypertension [49].

However, potential short and long-term potential adverse effects such as male impotence [49–51], susceptibility to infections [52–55], potential teratogenic effect [28,56] and cutaneous adverse reactions [57] have also been associated with HU. Although sparse and in the majority unsupported, some studies have demonstrated recovery of spermatogenesis, normal counts and motility after the cessation of HU treatment [58]. The fear of such side effects has been a subject of concern to some professionals [59,60], parents as well as patients [40,61–64] and a potential barrier to compliance in some settings [65,66]. As a consequence, HU is still underutilized [40,61], despite the studies that have reported on the overall drug safety [67] and limited evidence regarding potential HU-induced leukemogenic [68] or teratogenic effects [69]. It is however noteworthy that there are no present data validating fertility and teratogenic effects of HU nor on the incidence of abnormal pregnancies as a result of HU treated father. The decision for continuation (or otherwise) of HU in instances of pregnant SCD mothers on HU should thus be taken on a case-to-case basis where the risk of cessation is carefully weighed against potential effects of terminating an effective treatment [21,70]. Furthermore, less evidence is currently available on the persistence over time of responses to HU. Long-term studies such as the HU Study of Long-Term Effects, HU Safety and Organ Toxicity together with BABY HUG follow-up studies continue to reveal more pharmacogenetics data to help define long-term risk of early initiation of HU, potential adverse effects and persistence of response.

It is evident that concomitant to more HU clinical trials, predictive drug-responder indexes, efficacy and toxicity studies,

Table 1. Genomic variants associated with hydroxyurea-induced HbF level.

Gene	SNPs	Chromosome: locus	Ref.
<i>HBB</i> gene cluster haplotype	rs7482144	11:5254939	[5,34,72–81,131]
<i>BCL11A</i>	rs1427407 rs4671393 rs7606173 rs7557939 rs1186868	2:60490908 2:60491212 3:60493111 2:60494212 2:61764103	[132]
<i>ARG1/2</i>	rs2295644 rs17599586 rs28384513	14:67599842 6:131583579 6:135055071	
<i>HBS1L-MYB</i>	rs9399137	6:135097880	
<i>SAR1</i>	rs2310991 rs4282891 rs76901216	3:142444839 10:70171890 10:70170313	[82,83]
<i>SALL2</i>	rs61743453	14:21523209	[27]

HbF: Fetal hemoglobin; SNP: Single nucleotide polymorphism.

educating physicians and/or care providers as well as the global healthcare system is necessary to overcome other barriers such as access to specialty care, insurance coverage as well as further provider-level aspects such as care giver ambivalence. It is also urgent to fully understand HU molecular mode of action, in order to explore alternative and potential less toxic and more acceptable agent that could equally increase the level of HbF.

Genomic variants associated with HU-induced HbF level

Patients' response to HU is highly variable, with induced HbF levels ranging from as low as 2% to > 30% HbF [23–25], potentially due to the pharmacogenomic interactions [67]; but our understanding of the reasons for the wide spectrum of clinical responses to HU treatment is incomplete [71]. It has been initially suggested that haplotypes in the *HBB* gene cluster possibly affect possibly the clinical response to HU, likely mediated by their genetically determined effect on HbF level (TABLE 1) [34,72,73]. There is a positive association between the *XmnI* polymorphism (rs7482144) in the γ -globin promoter and HbF levels in numerous populations [15,16,74–76], as well as accounting for > 2% in HbF variation in patients and 13–32% in F cell variation in non-anemic European population [77]. This polymorphism has also been associated with the less severe disease course of patients with Senegal and/or Indian-Arab haplotype backgrounds as well as improved response to HU treatment [74,75,77–81]. Subsequent research provides some evidence that the effect of HU on HbF level could act on other HbF-promoting loci like *BCL11A* [26]. Indeed, *BCL11A* is central to the 'fetal switch', *BCL11A* is co-expressed, directly interacts and co-occupies the β -globin loci with *SOX-6* in association with the Mi-2/nucleosome remodeling and deacetylase complex for long-range re-conformation of the β -globin cluster for the transcriptional silencing of γ -globin [38]. Besides *BCL11A*, from DNA structural alteration to sequence modification, the

Table 2. Summary of mechanisms of HbF production in response to hydroxyurea.

Signal level	Pathway	Mechanism of γ -globin induction	Ref.
Epigenetic and transcriptional events	DNA remodeling	<i>Mi-2/NuRD</i> ; Acetylation and methylation	[31,38,84,133,134]
	DNA methylation	<i>DNMT-1</i>	[131,132]
Signaling pathways	<i>Gα/JNK/Jun</i>	<i>NFκB</i> induction of <i>SAR1</i>	[82]
	cGMP	sGC-PKG and NO-induced sGC-PKG	[86,88,90,91]
	cAMP	<i>MYB</i> transcriptional control	[90,124,135]
	MAPKs	<i>Erk1-2/p38/JNK/CREB1</i>	[91,97–100,136]
	Nitric oxide	eNOS, flow-induced ATP and anti-sickling	[85,123]
	Phosphorylation	<i>EIF2α</i>	[128]
	Transcription factors	<i>GATA-1/2</i> ; <i>p21</i> ; <i>EPO</i> ; <i>KLF1</i>	[32,101,137]
	Stress signals	ROS (superoxide and hydroxyl radicals)	[104]
Post-transcription	Translational Regulation	miR-494; 15a; 16–1; 151–3p; 148b	[84,124–126,138–141]

eIF2 α : Eukaryotic initiation factor 2 α ; eNOS: Endothelial nitric oxide synthase; NF κ B: Nuclear factor κ B; NuRD: Nucleosome remodeling and deacetylase; ROS: Reactive oxygen species; SAR1: Secretion-associated and ras-related protein.

secretion-associated and ras-related protein (SAR-1) has been shown to play a significant role in γ -globin regulation [82] and three SNPs in the SAR-1a in its promoter sequence have been associated with level of HbF in peripheral blood of SCD patients on HU [83]. Variants in regulatory sequences of *SAR-1* were shown *in vitro* to be associated with HU induced HbF production, underlying *SAR-1* as an alternative therapeutic target for β -globin disorders [82]. Recent GWAS studies have also shown that a coding variant in Spalt-like transcription factor, or *SALL2*, was associated with higher final HbF in response to HU treatment, a new insight into the pharmacological HbF upregulation by HU in patients with SCD [27] that deserves further functional studies.

Molecular pathway of action of HU

Various molecular pathways have been reported to explain the mode of action of HU (TABLE 2 & FIGURE 1): i) Epigenetic modifications, transcriptional events and signaling pathways involved in HU-mediated response, ii) Signaling pathways involving HU-mediated response and iii) Post-transcriptional pathways (regulation by miRNAs).

Epigenetic & transcriptional events regulating HU response

BCL11A, *SOX-6* and the *Mi-2*/nucleosome remodeling and deacetylase complex co-occupy the β -globin loci for long-range re-conformation of the β -globin cluster for the transcriptional silencing of γ -globin [38]. Chromosomal looping mediates the long range interaction between methylation-sensitive regulatory sites in the locus control region and the β -globin gene cluster. It has also been shown that *BCL11A* occupies the LCR HS3 and the intergenic region between $G_{\gamma A}$ and δ -globin genes [31]. Acetylation-based transcriptional activity has been observed in the LCR, δ - and β -globin genes as well as the

intergenic region between A_{γ} and δ -globin genes, whereas methylation was most observed in the CpG islands of the γ -globin proximal promoter [38,84]. Furthermore, HU-induced HbF has previously been inversely correlated with $G\gamma$ -globin methylation in SCD patients [84]. Albeit, the changes in methylation did not reach statistical significance, decreases in methylated CpG islands in the $G\gamma$ -globin promoter were identified and associated with increased γ -globin expression. Such epigenetic modifications have proven to play a significant role in the transcriptional regulation of the β -globin gene cluster and thus may require further investigation to find target-specific chromatin remodeling agents for the re-activation of γ -globin expression to ameliorate SCD symptoms.

Signaling pathways involved in HU-mediated response

Cyclic AMP & GMP pathways

The cGMP pathway has been implicated as a role player in the induction of HbF *in vitro*, in K562 cells and human erythroblasts, as well as *in vivo*, in SCD patients [85–87]. The pivotal role of the cGMP-dependent pathway in γ -globin expression regulation has previously been described [85] with sGC-PKG being involved in γ -globin induction in response to HU [86]. A humanized SCD mouse model of TNF- α -induced acute vaso-occlusion has been used to delineate an HU-induced nitric oxide (NO)-cGMP pathway, where HU-induced NO stimulates intracellular sGC thus increasing cGMP and thus PKG [88]. The observed effects of increased PKG included induction of γ -globin expression in erythroid cells and decreased leukocyte recruitment and adhesion. Although still a mouse model, this pathway is considered true also in humans and provides numerous therapeutic targets for HbF production.

There is cross-talk between the cGMP- and cAMP-dependent pathways and it has been demonstrated in several cell types [89]. Central to the cross-talk is the cAMP-specific

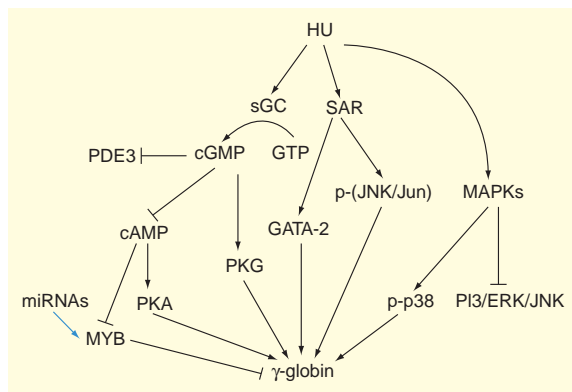


Figure 1. Various HU-activated signal transduction pathways for γ -globin expression.

Summarized pathways previously associated and/or induced by HU treatment in primary erythroid and K562 cells. HU: Hydroxyurea.

phosphodiesterase 3 that is inhibited by cGMP and results in the activation of cAMP-dependent PKA through the increase of intracellular cAMP levels [85]. The role of the cAMP pathway and the mechanisms of HbF regulation have been described previously with c-Myb identified as a key role player in the cAMP-mediated inhibition of γ -globin expression, particularly in K562 cells that significantly express c-Myb when compared to adult erythroblasts [90], possibly suggesting the absence of c-Myb/cAMP regulation of HbF in adult erythroblasts. Although c-Myb reduces the transcriptional activity in the γ -globin promoter, which is concomitant with the low expression of c-Myb in γ -globin expressing-adult erythroblasts, induction of the cAMP pathway is central to erythroid differentiation in K562 cells [90]. The role of the cAMP pathway in γ -globin regulation in erythroblasts appears to be transient and developmental stage-specific, with significantly higher intracellular cAMP levels in fetal liver erythroblasts when compared to the low-to-undetectable levels in adult erythroblasts. These data advocate for further investigation of the cAMP pathway and identification of the 'cAMP-switch' during erythropoiesis as a potential therapeutic target for cAMP-mediated HbF induction in erythroblasts. Furthermore, the polar role of the cAMP pathway in K562, where it inhibits γ -globin expression, and adult erythroblasts, where it induces HbF production, suggests that the role of the cAMP pathway in the regulation of HbF may depend on the global cellular context of erythroid cells.

MAP kinases

MAPKs such as ERK, p38 and c-Jun N-terminal kinase (JNK) appear to be independent of cAMP-dependent pathways regulating HbF [91], even though cross-talk between cAMP-dependent and MAPK pathways has been described in non-erythroid cells [92–96]. Several studies have investigated the role of MAPKs in HbF production using erythroid cells [97–100]. HU has been shown to increase phosphorylation of MAPK

p38, while simultaneously causing dephosphorylation of ERK and JNK [97], effects which are known to affect erythroid differentiation and induce γ -globin expression. Furthermore, specific inhibition of p38 in K562 cells without affecting ERK and JNK resulted in decreased expression of γ -globin, contrary to inhibition of ERK.

JNK/Jun pathway

Recently, the role of *SARI*, a small guanosine triphosphate-binding protein in adult erythroid cells [101], in HU-induced HbF has been shown in bone marrow CD34⁺ and K562 cells [82]. Nuclear factor κ B was shown to be central to the expression of *SARI* through direct binding to the promoter sequence after induction by HU. Furthermore, the involvement of JNK/Jun phosphorylation and $G_i\alpha$ pathways in γ -globin expression was demonstrated through inhibition and silencing experiments in both CD34⁺ and K562 cells, resulting in the reduction of HbF and thus demonstrating that activation of the *G α /JNK/Jun* pathway proteins is necessary for nuclear factor κ B dependent *SARI* expression in the production of HU-mediated HbF.

Other molecular pathways

GATA erythroid transcription factors are widely considered key modulators of γ -globin expression [101]. Both *GATA-1* and *GATA-2* proteins have been associated with HU treatment in K562 cells prompting the possibility of a HU-induced *GATA-1* and/or -2/p21-dependent signal transduction pathway toward γ -globin expression [101]. Reactive oxygen species such as superoxide, hydrogen peroxide (H₂O₂) and NO have been shown to mediate the phosphorylation of p38 MAPK [102,103], whose association with HbF production was discussed above. Inducers of γ -globin such as sodium butyrate and trichostatin A have demonstrated an H₂O₂-dependent activation of p38 through phosphorylation and subsequent HbF production in erythroid progenitors and K562 cells [104]. However, HU-induced γ -globin expression is independent of H₂O₂ formation, albeit other reactive oxygen species such as NO, superoxide and hydroxyl radicals are known to activate MAPKs and cGMP [105,106].

Nitric oxide

HU treatment is known to improve blood flow through reperfusion and decreased sickling; however, the reasons for the improved circulation are not completely understood. NO release has had a long-standing association with HU-induced improved blood flow through vasodilation [107–113]. Recently, a rabbit erythrocyte model was used to demonstrate that HU regulates the production of NO by the enzyme endothelial NO synthase [114], and further showing that the released NO induces blood flow-induced ATP, which is a known mediator of vasodilation through endothelial cell purinergic receptor binding [115–118] and subsequent release of endothelial NO [119–121]. The improved blood flow results from the NO-induced increase in oxygen affinity of sickle erythrocytes in SCD patients because NO exhibits anti-sickling properties and inhibits HbS polymerization [85,122]. NO may be released by

endothelial cells and/or erythrocytes, establishing a feedback loop with flow-derived ATP release from erythrocytes and provide rationale for some of the immediate reported benefits of HU treatment (improved blood flow and attenuated sickling events in SCD patients) [123].

Post-transcriptional pathways: regulation by miRNAs

Transcriptional regulation through short non-coding RNA oligonucleotides (miRNAs) holds significant promise in explaining variances in HbF induction in SCD patients of similar genomic background and environmental influence. Recently, a study demonstrated significant HU-induced miRNAs expression in SCD patients and associated with HbF at both baseline (miR-494) and maximum dose tolerated (miR-26b and miR-151-3p). Furthermore, HU has also been shown to upregulate miR-15a and miR-16-1, whose direct target is *MYB*, a critical regulator of γ -globin expression [124]. Downregulation of miR-148b has also exhibited tumor suppressor properties through inhibition of cell proliferation in gastric cancer [125]. Reticulocytes have also been shown to contain miRNAs, although enucleated and previously thought not to contain any nucleic acids [126]. Furthermore, it was shown that the miRNA profile of terminally differentiated erythrocytes from SCD patients was significantly different from unaffected individuals, thereby prompting an association with disease physiology. Defective terminal differentiation during sickle cell erythropoiesis was supported by an association with poor expression of miR-320 [126]. The role of miRNAs in hematopoiesis, specifically erythropoiesis and erythrocyte physiology has also been previously reviewed [127].

The field of post-transcriptional regulation of HbF has made a recent surge to scientific interest as a potential level along the pathway to γ -globin expression for therapeutic targets. In a recent study, inhibition of dephosphorylation of the eukaryotic initiation factor 2 α , a critical regulator of protein translation, enhanced production of HbF and was devoid of any changes at the levels of gene transcription, cellular proliferation and differentiation [128]. The use of salubrinal, a known inhibitor of eukaryotic initiation factor 2 α phosphorylation, improved γ -globin mRNA turnover to HbF without a globin preference shift (the ratio of $\gamma/[\gamma+\beta]$ remained unchanged), and thereby elucidating another level for potential therapeutic targets. In support of these findings, salubrinal was further shown to increase translation efficiency during the recovery phase of cellular stress response through an increase in the number of selective translating ribosomes on the γ -globin mRNA [129]. Further investigation of translation-altering agents and the global disease-specific miRNA profile in understudied and highly heterogeneous populations such as in Africa promises to yield mechanistic insight into the disease phenotype and differential response to HU treatment.

Discussion & conclusion

HU treatment has demonstrated success in several settings, both in children and adults with SCD. The current dilemma

stems from various clinical responses to the treatment and often the lack thereof. There has been great progress made towards understanding various mechanisms by which HU induces HbF production through identifying key genomic variants that predispose the response to the treatment, various environmental contributors as well as specific signal transduction pathways that modulate γ -globin expression. However, the complete picture of HU-mediated HbF production remains elusive. While the majority of HU benefit on SCD is directly related to the amount of HbF produced, it is also suspected that HU has some disease modulating effects outside of HbF induction, such as stress hematopoiesis, endothelial NO release or the reduction of leukocyte counts, that the present paper has not fully addressed.

In recent times, there has been a shift in the search for therapeutic targets toward the post-transcriptional and post-translational mechanisms that regulate HbF production as a way of minimizing alterations to the global transcriptional system through stress. There is an urgent need to investigate the role of miRNAs in HU-induced HbF production in SCD patients and further explore other potent negative regulators of γ -globin expression and target of miRNAs, with the aim of elucidating a miRNA-mediated signaling transduction pathway induced by HU treatment. The outcomes of such studies would firstly yield post-transcriptional population-specific variants and provide potential explanations for the vast inter-patient variation in response to HU treatment. And, secondly, delineate another regulatory level of gene expression to known pathways and therefore possibly reveal therapeutic targets that may minimize alterations to the cellular transcriptome.

Expert commentary

HU treatment has demonstrated success in several settings, both in children and adults with SCD. Negating the sparse evidence on the potential treatment-related adverse effect and the barriers to use, as the only available drug treatment for SCD, HU has a potential role in ameliorating the global problem of SCD, particularly in high disease burden locales such as Africa. The current dilemma stems from various clinical responses to the treatment and often the lack thereof. This challenge could be overcome in the future with three clinical trials underway in Africa, spanning Angola, Congo, Kenya and Nigeria, as well as North America and Europe, with 21 and 3 open trials, respectively [130]. Furthermore, there has been great progress made towards understanding various mechanisms by which HU induces HbF production through identifying key genomic variants that predispose the response to the treatment, various environmental contributors as well as specific signal transduction pathways that modulate γ -globin expression. However, the complete picture of HU-mediated HbF production remains elusive. The present study represents, to our knowledge, the first systematic review on known molecular mechanism of action of HU-induced HbF in SCD, the only medication available for this condition of global importance. It is anticipated that adequate knowledge of the mode of action of HU will

result in the exploration of alternative therapeutic agents with possibly less toxicity.

Five-year view

The study has provided some specific perspectives in future research, specifically on post-transcriptional regulations by miRNAs that could lead to possible therapeutic targets that minimize alterations to the cellular transcriptome. There is an urgent need to investigate the role of miRNAs in HU-induced HbF production in SCD patients and further explore other potent negative regulators of γ -globin expression and target of miRNAs, with the aim of elucidating a miRNA-mediated signaling transduction pathway induced by HU treatment. The outcomes of such studies would firstly yield post-transcriptional population-specific variants and provide potential explanations for the vast inter-patient variation in response to HU treatment. And, secondly, delineate another regulatory level of gene expression to known pathways and therefore possibly reveal therapeutic targets that may minimize alterations to the cellular transcriptome.

While the majority of HU benefit on SCD is directly related to the amount of HbF produced, it is also suspected

that HU has some disease modulating effects outside of HbF induction, such as stress hematopoiesis, endothelial NO release or the reduction of leukocyte counts; this paper did not explore those mechanisms. It is anticipated that more research on other HU-modulating effects on SCD outside of HbF induction will evolve from these areas to complete a comprehensive profile of the mechanism of action of HU in SCD.

Financial & competing interests disclosure

The students' bursaries were supported by various funding provided by the University of Cape Town, SAMRC, NHLS, NRF, FirstRand Laurie Dippenaar, and Oppenheimer Memorial Trust, South Africa. The National Institute of Health (NIH), USA, funded the report and publication of the present manuscript (grant number 1U01HG007459-01). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript. The authors have no other relevant affiliations or financial involvement with any organization or entity with a financial interest in or financial conflict with the subject matter or materials discussed in the manuscript apart from those disclosed.

Key issues

- The study has summarized various molecular pathways that have been reported to explain the mode of action of hydroxyurea (HU).
- Numerous signal transduction pathways have been implicated, through the identification of key genomic variants in *BCL11A*, *HBS1L-MYB* or *SAR1* that predispose the response to the HU treatment in Sickle Cell Disease.
- Additional signal transduction pathways modulate γ -globin expression (cAMP/cGMP; G α /c-Jun N-terminal kinase/Jun; methylation and miRNA).
- Three main molecular pathways have been reported:
 - Epigenetic modifications, transcriptional events and signaling pathways involved in HU-mediated response
 - Signaling pathways involving HU-mediated responsePost-transcriptional pathways (regulation by miRNAs)

References

Papers of special note have been highlighted as:

- of interest
- of considerable interest

1. Bartolucci P, Galacteros F. Clinical management of adult sickle-cell disease. *Curr Opin Hematol* 2012;19(3):149-55
- **Good article on the management of adult patients with sickle cell disease (SCD).**
2. Williams TN, Mwangi TW, Wambua S, et al. Sickle cell trait and the risk of *Plasmodium falciparum* malaria and other childhood diseases. *J Infect Dis* 2005; 192(1):178-86
3. Piel FB, Patil AP, Howes RE, et al. Global epidemiology of sickle haemoglobin in neonates: a contemporary geostatistical model-based map and population estimates. *Lancet* 2013;381(9861):142-51
4. Modell B, Darlison M. Global epidemiology of haemoglobin disorders and derived service indicators. *Bull World Health Organ* 2008;86(6):480-7
5. Labie D, Pagnier J, Lapoumeroulie C, et al. Common haplotype dependency of high G gamma-globin gene expression and high Hb F levels in beta-thalassemia and sickle cell anemia patients. *Proc Natl Acad Sci USA* 1985;82(7):2111-14
6. Elion J, Berg PE, Lapoumeroulie C, et al. DNA sequence variation in a negative control region 5' to the beta-globin gene correlates with the phenotypic expression of the beta s mutation. *Blood* 1992;79(3): 787-92
7. Pagnier J, Mears JG, Dunda-Belkhdja O, et al. Evidence for the multicentric origin of the sickle cell hemoglobin gene in Africa. *Proc Natl Acad Sci USA* 1984;81(6):1771-3
8. Steinberg MH. Genetic etiologies for phenotypic diversity in sickle cell anemia. *ScientificWorldJournal* 2009;9:46-67
9. Alsultan A, Aleem A, Ghabbour H, et al. Sickle cell disease subphenotypes in patients from Southwestern Province of Saudi Arabia. *J Pediatr Hematol Oncol* 2012; 34(2):79-84
10. Platt OS, Brambilla DJ, Rosse WF, et al. Mortality in sickle cell disease. Life expectancy and risk factors for early death. *N Engl J Med* 1994;330(23):1639-44
11. Bauer DE, Orkin SH. Update on fetal hemoglobin gene regulation in hemoglobinopathies. *Curr Opin Pediatr* 2011;23(1):1-8
12. Thein SL, Menzel S, Lathrop M, Garner C. Control of fetal hemoglobin: new insights emerging from genomics and clinical implications. *Hum Mol Genet* 2009; 18(R2):R216-23
13. Creary LE, Ulug P, Menzel S, et al. Genetic variation on chromosome 6 influences F cell levels in healthy individuals of African

- descent and HbF levels in sickle cell patients. *PLoS One* 2009;4(1):e4218
14. Menzel S, Garner C, Gut I, et al. A QTL influencing F cell production maps to a gene encoding a zinc-finger protein on chromosome 2p15. *Nat Genet* 2007;39(10):1197-9
 15. Lettre G, Sankaran VG, Bezerra MA, et al. DNA polymorphisms at the BCL11A, HBS1L-MYB, and beta-globin loci associate with fetal hemoglobin levels and pain crises in sickle cell disease. *Proc Natl Acad Sci USA* 2008;105(33):11869-74
 - **Great paper showing the association between fetal hemoglobin (HbF) and polymorphisms at key regulators: BCL11A, intergenic region between HBS1L and MYB as well as β -globin cluster.**
 16. Makani J, Menzel S, Nkya S, et al. Genetics of fetal hemoglobin in Tanzanian and British patients with sickle cell anemia. *Blood* 2011;117(4):1390-2
 - **Good comparative study between Europe and Africa on polymorphisms associated with HbF.**
 17. Wonkam A, Bitoungui VJN, Vorster AA, et al. Association of variants at BCL11A and HBS1L-MYB with hemoglobin F and hospitalization rates among sickle cell patients in Cameroon. *PLoS One* 2014;9(3):e92506
 - **Great article showing association of variants at BCL11A and HBS1L-MYB on both HbF and clinical course of African SCD patients.**
 18. Lebensburger JD, Pestina TI, Ware RE, et al. Hydroxyurea therapy requires HbF induction for clinical benefit in a sickle cell mouse model. *Haematologica* 2010;95(9):1599-603
 19. Lebensburger J, Johnson SM, Askenazi DJ, et al. Protective role of hemoglobin and fetal hemoglobin in early kidney disease for children with sickle cell anemia. *Am J Hematol* 2011;86(5):430-2
 20. Steinberg MH, Barton F, Castro O, et al. Effect of hydroxyurea on mortality and morbidity in adult sickle cell anemia: risks and benefits up to 9 years of treatment. *JAMA* 2003;289(13):1645-51
 21. Voskaridou E, Christoulas D, Bilalis A, et al. The effect of prolonged administration of hydroxyurea on morbidity and mortality in adult patients with sickle cell syndromes: results of a 17-year, single-center trial (LaSHS). *Blood* 2010;115(12):2354-63
 22. Lobo CL, Pinto JF, Nascimento EM, et al. The effect of hydroxycarbamide therapy on survival of children with sickle cell disease. *Br J Haematol* 2013;161(6):852-60
 23. Maier-Redelsperger M, de Montalembert M, Flahault A, et al. Fetal hemoglobin and F-cell responses to long-term hydroxyurea treatment in young sickle cell patients. The French Study Group on Sickle Cell Disease. *Blood* 1998;91(12):4472-9
 24. Kinney TR, Helms RW, O'Branski EE, et al. Safety of hydroxyurea in children with sickle cell anemia: results of the HUG-KIDS study, a phase I/II trial. Pediatric Hydroxyurea Group. *Blood* 1999;94(5):1550-4
 25. Zimmerman SA, Schultz WH, Davis JS, et al. Sustained long-term hematologic efficacy of hydroxyurea at maximum tolerated dose in children with sickle cell disease. *Blood* 2004;103(6):2039-45
 26. Ware RE, Despotovic JM, Mortier NA, et al. Pharmacokinetics, pharmacodynamics, and pharmacogenetics of hydroxyurea treatment for children with sickle cell anemia. *Blood* 2011;118(18):4985-91
 - **Great article showing the variable inter-patient variability in hydroxyurea (HU) responses, dosage and disease progression.**
 27. Sheehan VA, Luo Z, Flanagan JM, et al. Genetic modifiers of sickle cell anemia in the BABY HUG cohort: influence on laboratory and clinical phenotypes. *Am J Hematol* 2013;88(7):571-6
 - **Good paper showing genetic modifiers from the BABY HUG study.**
 28. Charache S, Terrin ML, Moore RD, et al. Effect of hydroxyurea on the frequency of painful crises in sickle cell anemia. *N Engl J Med* 1995;332(20):1317-22
 29. Ohene-Frempong K, Weiner SJ, Sleeper LA, et al. Cerebrovascular accidents in sickle cell disease: rates and risk factors. *Blood* 1998;91(1):288-94
 30. Steinberg MH, McCarthy WF, Castro O, et al. The risks and benefits of long-term use of hydroxyurea in sickle cell anemia: A 17.5 year follow-up. *Am J Hematol* 2010;85(6):403-8
 31. Sankaran VG, Menne TF, Xu J, et al. Human fetal hemoglobin expression is regulated by the developmental stage-specific repressor BCL11A. *Science* 2008;322(5909):1839-42
 - **Great paper demonstrating the role of BCL11A during fetal switch of globin expression.**
 32. Borg J, Papadopoulos P, Georgitsi M, et al. Haploinsufficiency for the erythroid transcription factor KLF1 causes hereditary persistence of fetal hemoglobin. *Nat Genet* 2010;42(9):801-5
 - **Good article demonstrating the resultant HPPH from KLF1 haploinsufficiency.**
 33. Ngo D, Bae H, Steinberg MH, et al. Fetal hemoglobin in sickle cell anemia: genetic studies of the Arab-Indian haplotype. *Blood Cells Mol Dis* 2013;51(1):22-6
 34. Gilman J. Chromosomes. *Blood* 1984;64(2):452-7
 35. Garner C, Tatu T, Reittie JE, et al. Genetic influences on F cells and other hematologic variables: a twin heritability study. *Blood* 2000;95(1):342-6
 36. Higgs DR, Wood WG. Genetic complexity in sickle cell disease. *Proc Natl Acad Sci USA* 2008;105(33):11595-6
 37. Thein SL, Menzel S, Peng X, et al. Intergenic variants of HBS1L-MYB are responsible for a major quantitative trait locus on chromosome 6q23 influencing fetal hemoglobin levels in adults. *Proc Natl Acad Sci USA* 2007;104(27):11346-51
 38. Xu J, Sankaran VG, Ni M, et al. Transcriptional silencing of γ -globin by BCL11A involves long-range interactions and cooperation with SOX6. *Genes Dev* 2010;24(8):783-98
 - **Nice article showing the components of the silencing complex and the reconfiguration of DNA to silence HbF.**
 39. Bauer DE, Kamran SC, Lessard S, et al. An erythroid enhancer of BCL11A subject to genetic variation determines fetal hemoglobin level. *Science* 2013;342(6155):253-7
 - **Nice article showing the effect of enhancer polymorphisms in BCL11A.**
 40. Lanzkron S, Hayward C Jr, Hassell KL, Rand C. Provider barriers to hydroxyurea use in adults with sickle cell disease: a survey of the Sickle Cell Disease Adult Provider Network. *J Natl Med Assoc* 2008;100(8):968-73
 41. Strouse JJ, Lanzkron S, Beach MC, et al. Hydroxyurea for sickle cell disease: a systematic review for efficacy and toxicity in children. *Pediatrics* 2008;122(6):1332-42
 - **Great systematic review on efficacy and potential toxic effects of HU.**
 42. Brawley OW, Cornelius LJ, Edwards LR, et al. National Institutes of Health Consensus Development Conference statement: hydroxyurea treatment for sickle

- cell disease. *Ann Intern Med* 2008;148(12):932-8
43. Elford HL. Effect of hydroxyurea on ribonucleotide reductase. *Biochem Biophys Res Commun* 1968;33(1):129-35
 44. Platt OS, Orkin SH, Dover G, et al. Hydroxyurea enhances fetal hemoglobin production in sickle cell anemia. *J Clin Invest* 1984;74(2):652-6
 45. Zimmerman SA, Schultz WH, Burgett S, et al. Hydroxyurea therapy lowers transcranial Doppler flow velocities in children with sickle cell anemia. *Blood* 2007;110(3):1043-7
 46. Thornburg CD, Dixon N, Burgett S, et al. A pilot study of hydroxyurea to prevent chronic organ damage in young children with sickle cell anemia. *Pediatr blood cancer* 2009;52(5):609-15
 47. Hankins JS, Ware RE, Rogers ZR, et al. Long-term hydroxyurea therapy for infants with sickle cell anemia: the HUSOFT extension study. *Blood* 2005;106(7):2269-75
 48. Wang W, Thompson B. Hydroxyurea treatment of infants with sickle cell anemia: results of the BABY HUG study. *Pediatric Blood Cancer* 2010;54(2):250-5
 49. DeBaun MR. Hydroxyurea therapy contributes to infertility in adult men with sickle cell disease: a review. *Exp rev hematol* 2014;7(6):767-73
 50. Smith-Whitley K. Reproductive issues in sickle cell disease. *Blood* 2014;124(24):3538-43
 51. Berthaut I, Guignédoux G, Kirsch-Noir F, et al. Influence of sickle cell disease and treatment with hydroxyurea on sperm parameters and fertility of human males. *Haematologica* 2008;93(7):988-93
 52. Shemisa K, Jafferjee N, Thomas D, et al. *Mycobacterium avium* complex infection in a patient with sickle cell disease and severe iron overload. *Case Rep Infect Dis* 2014;2014:405323
 53. Amoako N, Asante KP, Adjei G, et al. Associations between Red Cell Polymorphisms and *Plasmodium falciparum* Infection in the Middle Belt of Ghana. *PLoS One* 2014;9(12):e112868
 54. Hirst C, Owusu-Ofori S. Prophylactic antibiotics for preventing pneumococcal infection in children with sickle cell disease. *The Cochrane Library* 2012
 55. Owusu ED, Visser BJ, Nagel IM, et al. The interaction between sickle cell disease and HIV infection: a systematic review. *Clin Infect Dis* 2015;60(4):612-26
 56. Flanagan JM, Howard TA, Mortier N, et al. Assessment of genotoxicity associated with hydroxyurea therapy in children with sickle cell anemia. *Mutat ResGenet Toxicol Environ Mutagen* 2010;698(1):38-42
 57. Chaîne B, Neonato M, Girot R, Aractingi S. Cutaneous adverse reactions to hydroxyurea in patients with sickle cell disease. *Arch Dermatol* 2001;137(4):467-70
 58. Grigg A. Effect of hydroxyurea on sperm count, motility and morphology in adult men with sickle cell or myeloproliferative disease. *Intern Med J* 2007;37(3):190-2
 59. Oyeku SO, Driscoll MC, Cohen HW, et al. Parental and other factors associated with hydroxyurea use for pediatric sickle cell disease. *Pediatr blood cancer* 2013;60(4):653-8
 60. Brandow AM, Jirovec DL, Panepinto JA. Hydroxyurea in children with sickle cell disease: practice patterns and barriers to utilization. *Am J Hematol* 2010;85(8):611-13
 61. Zumberg MS, Reddy S, Boyette RL, et al. Hydroxyurea therapy for sickle cell disease in community-based practices: A survey of Florida and North Carolina hematologists/oncologists. *Am J Hematol* 2005;79(2):107-13
 62. Thornburg CD, Calatroni A, Telen M, Kemper AR. Adherence to hydroxyurea therapy in children with sickle cell anemia. *J Pediatr* 2010;156(3):415-19
 63. Strouse JJ, Heeney MM. Hydroxyurea for the treatment of sickle cell disease: efficacy, barriers, toxicity, and management in children. *Pediatr blood cancer* 2012;59(2):365-71
 64. Haywood C Jr, Beach MC, Lanzkron S, et al. A systematic review of barriers and interventions to improve appropriate use of therapies for sickle cell disease. *J Natl Med Assoc* 2009;101(10):1022-33
 65. Lebensburger JD, Sidonio RF, DeBaun MR, et al. Exploring barriers and facilitators to clinical trial enrollment in the context of sickle cell anemia and hydroxyurea. *Pediatr blood cancer* 2013;60(8):1333-7
 66. Hampton T. Strides made in red blood cell disorders, but substantial barriers to care remain. *JAMA* 2008;299(4):395-6
 67. Steinberg MH, Barton F, Castro O, et al. Effect of hydroxyurea on mortality and morbidity in adult sickle cell anemia: risks and benefits up to 9 years of treatment. *JAMA* 2003;289(13):1645-51
 68. Segal JB, Strouse JJ, Beach MC, et al. Hydroxyurea for the treatment of sickle cell disease. *Evid Rep Technol Assess* 2008;165:1-95
 69. Ballas SK, Kesen MR, Goldberg MF, et al. Beyond the definitions of the phenotypic complications of sickle cell disease: an update on management. *The Scientific World Journal* 2012;2012
 70. Ware RE. Hydroxycarbamide: clinical aspects. *C R Biol* 2013;336(3):177-82
 71. Bakanay SM, Dainer E, Clair B, et al. Mortality in sickle cell patients on hydroxyurea therapy. *Blood* 2005;105(2):545-7
 72. Adekile AD. Limitations of Hb F as a phenotypic modifier in sickle cell disease: study of kuwaiti arab patients. *Hemoglobin* 2011;35(5-6):607-17
 73. Friedrichs JR, Prá D, Maluf SW, et al. DNA damage in blood leukocytes of individuals with sickle cell disease treated with hydroxyurea. *Mutat ResGenet Toxicol Environ Mutagen* 2008;649(1):213-20
 74. Pandey S, Pandey S, Mishra RM, Saxena R. Modulating Effect of the -158 gamma (C->T) Xmn1 Polymorphism in Indian Sickle Cell Patients. *Mediterr. J Hematol Infect Dis* 2012;4(1):e2012001
- **Good article showing the proven effect of the XmnI polymorphism on HbF and Indian-Arab haplotype.**
75. Aditya R, Verma I, Saxena R, et al. Relation of Xmn-1 polymorphism & five common Indian mutations of thalassaemia with phenotypic presentation in β -thalassaemia. *JK Science* 2006;8:139-43
 76. Nemati H, Rahimi Z, Bahrami G. The XmnI polymorphic site 5' to the G γ gene and its correlation to the G γ : A γ ratio, age at first blood transfusion and clinical features in β -Thalassaemia patients from Western Iran. *Mol Biol Rep* 2010;37(1):159-64
 77. Garner C, Tatu T, Game L, et al. A candidate gene study of F cell levels in sibling pairs using a joint linkage and association analysis. *GeneScreen* 2000;1(1):9-14
 78. Panigrahi I, Agarwal S, Gupta T, et al. Hemoglobin E-beta thalassaemia: factors affecting phenotype. *Indian Pediatr* 2005;42(4):357-62
 79. Oberoi S, Das R, Panigrahi I, et al. Xmn1-G γ polymorphism and clinical predictors of severity of disease in β -thalassaemia intermedia. *Pediatr blood cancer* 2011;57(6):1025-8
 80. Das S. A study to understand the relation between fetal haemoglobin, the

- hematological parameters and Xmn i gene polymorphism. *Ind J Med Health* 2012; 1(9):206-10
81. Thein S, Wainscoat J, Sampietro M, et al. Association of thalassaemia intermedia with a beta-globin gene haplotype. *Br J Haematol* 1987;65(3):367-73
 82. Zhu J, Chin K, Aerbajinai W, et al. Hydroxyurea-inducible SAR1 gene acts through the Galpha/JNK/Jun pathway to regulate gamma-globin expression. *Blood* 2014;124(7):1146-56
 83. Kumkhaek C, Taylor JG VI, Zhu J, et al. Fetal haemoglobin response to hydroxycarbamide treatment and sar1a promoter polymorphisms in sickle cell anaemia. *Br J Haematol* 2008;141(2):254-9
 - **Nice paper on secretion-associated and ras-related protein 1 promoter polymorphisms associated with HbF response to HU.**
 84. Walker AL, Steward S, Howard TA, et al. Epigenetic and molecular profiles of erythroid cells after hydroxyurea treatment in sickle cell anemia. *Blood* 2011;118(20):5664-70
 85. Ikuta T, Aüsenda S, Cappellini MD. Mechanism for fetal globin gene expression: role of the soluble guanylate cyclase-cGMP-dependent protein kinase pathway. *Proc Natl Acad Sci USA* 2001; 98(4):1847-52
 86. Cokic VP, Smith RD, Beleslin-Cokic BB, et al. Hydroxyurea induces fetal hemoglobin by the nitric oxide-dependent activation of soluble guanylyl cyclase. *J Clin Invest* 2003; 111(2):231-9
 87. Conran N, Oresco-Santos C, Acosta HC, et al. Increased soluble guanylate cyclase activity in the red blood cells of sickle cell patients. *Br J Haematol* 2004;124(4):547-54
 88. Almeida CB, Scheiermann C, Jang JE, et al. Hydroxyurea and a cGMP-amplifying agent have immediate benefits on acute vaso-occlusive events in sickle cell disease mice. *Blood* 2012;120(14):2879-88
 89. Beavo JA, Hansen RS, Harrison SA, et al. Identification and properties of cyclic nucleotide phosphodiesterases. *Mol Cell Endocrinol* 1982;28(3):387-410
 90. Kuroyanagi Y, Kaneko Y, Muta K, et al. cAMP differentially regulates gamma-globin gene expression in erythroleukemic cells and primary erythroblasts through c-Myb expression. *Biochem Biophys Res Commun* 2006;344(3):1038-47
 - **Great paper on the role of cMYB and the cAMP pathway in HbF production.**
 91. Inoue A, Kuroyanagi Y, Terui K, et al. Negative regulation of γ -globin gene expression by cyclic AMP-dependent pathway in erythroid cells. *Exp Hematol* 2004;32(3):244-53
 92. de Rooij J, Zwartkruis FJ, Verheijen MH, et al. Epac is a Rap1 guanine-nucleotide-exchange factor directly activated by cyclic AMP. *Nature* 1998;396(6710):474-7
 93. Italia K, Jain D, Gattani S, et al. Hydroxyurea in sickle cell disease—a study of clinico-pharmacological efficacy in the Indian haplotype. *Blood Cells Mol Dis* 2009;42(1):25-31
 94. Li X, Zarinetchi F, Schrier RW, Nemenoff RA. Inhibition of MAP kinase by prostaglandin E2 and forskolin in rat renal mesangial cells. *Am J Physiol* 1995; 269(4 Pt 1):C986-91
 95. Rao KM. MAP kinase activation in macrophages. *J Leukoc Biol* 2001;69(1): 3-10
 96. Rincon M. MAP-kinase signaling pathways in T cells. *Curr Opin Immunol* 2001;13(3): 339-45
 97. Park JI, Choi HS, Jeong JS, et al. Involvement of p38 kinase in hydroxyurea-induced differentiation of K562 cells. *Cell Growth Differ* 2001;12(9): 481-6
 98. Rivero JA, Adunyah SE. Sodium butyrate induces tyrosine phosphorylation and activation of MAP kinase (ERK-1) in human K562 cells. *Biochem Biophys Res Commun* 1996;224(3):796-801
 99. Yang J, Kawai Y, Hanson RW, Arinze IJ. Sodium butyrate induces transcription from the G alpha(i2) gene promoter through multiple Sp1 sites in the promoter and by activating the MEK-ERK signal transduction pathway. *J Biol Chem* 2001; 276(28):25742-52
 100. Liu K, Xing H, Zhang S, et al. Cucurbitacin D induces fetal hemoglobin synthesis in K562 cells and human hematopoietic progenitors through activation of p38 pathway and stabilization of the gamma-globin mRNA. *Blood Cells Mol Dis* 2010;45(4):269-75
 101. Tang DC, Zhu J, Liu W, et al. The hydroxyurea-induced small GTP-binding protein SAR modulates gamma-globin gene expression in human erythroid cells. *Blood* 2005;106(9):3256-63
 102. Dabrowski A, Boguslawicz C, Dabrowski M, et al. Reactive oxygen species activate mitogen-activated protein kinases in pancreatic acinar cells. *Pancreas* 2000;21(4): 376-84
 103. Dong J, Ramachandiran S, Tikoo K, et al. EGFR-independent activation of p38 MAPK and EGFR-dependent activation of ERK1/2 are required for ROS-induced renal cell death. *Am J Physiol Renal Physiol* 2004;287(5):F1049-58
 104. Hsiao CH, Li W, Lou TF, et al. Fetal hemoglobin induction by histone deacetylase inhibitors involves generation of reactive oxygen species. *Exp Hematol* 2006; 34(3):264-73
 105. Lander HM. An essential role for free radicals and derived species in signal transduction. *FASEB J* 1997;11(2):118-24
 106. Cerimele F, Bartle T, Lynch R, et al. Reactive oxygen signaling and MAPK activation distinguish Epstein-Barr Virus (EBV)-positive versus EBV-negative Burkitt's lymphoma. *Proc Natl Acad Sci USA* 2005;102(1):175-9
 107. Cokic VP, Beleslin-Cokic BB, Tomic M, et al. Hydroxyurea induces the eNOS-cGMP pathway in endothelial cells. *Blood* 2006;108(1):184-91
 108. Glover RE, Ivy ED, Orringer EP, et al. Detection of nitrosyl hemoglobin in venous blood in the treatment of sickle cell anemia with hydroxyurea. *Mol Pharmacol* 1999; 55(6):1006-10
 109. Nahavandi M, Wyche MQ, Perlin E, et al. Nitric Oxide Metabolites in Sickle Cell Anemia Patients after Oral Administration of Hydroxyurea; Hemoglobinopathy. *Hematology* 2000;5(4):335-9
 110. Jiang J, Jordan SJ, Barr DP, et al. In vivo production of nitric oxide in rats after administration of hydroxyurea. *Mol Pharmacol* 1997;52(6):1081-6
 111. Huang J, Yakubu M, Kim-Shapiro DB, King SB. Rat liver-mediated metabolism of hydroxyurea to nitric oxide. *Free Radic Biol Med* 2006;40(9):1675-81
 112. King SB. Nitric oxide production from hydroxyurea. *Free Radic Biol Med* 2004; 37(6):737-44
 113. King SB. The nitric oxide producing reactions of hydroxyurea. *Curr Med Chem* 2003;10(6):437-52
 114. Raththagala M, Karunarathne W, Kryziniak M, et al. Hydroxyurea stimulates the release of ATP from rabbit erythrocytes through an increase in calcium and nitric oxide production. *Eur J Pharmacol* 2010; 645(1-3):32-8
 115. Burnstock G, Kennedy C. A dual function for adenosine 5'-triphosphate in the

- regulation of vascular tone. Excitatory cotransmitter with noradrenaline from perivascular nerves and locally released inhibitory intravascular agent. *Circ Res* 1986;58(3):319-30
116. Abbracchio MP, Burnstock G. Purinoceptors: are there families of P2X and P2Y purinoceptors? *Pharmacol Ther* 1994; 64(3):445-75
117. Fredholm BB, Abbracchio MP, Burnstock G, et al. Nomenclature and classification of purinoceptors. *Pharmacol Rev* 1994;46(2):143-56
118. Motte S, Communi D, Piroton S, Boeynaems JM. Involvement of multiple receptors in the actions of extracellular ATP: the example of vascular endothelial cells. *Int J Biochem Cell Biol* 1995;27(1):1-7
119. Bogle RG, Coade SB, Moncada S, et al. Bradykinin and ATP stimulate L-arginine uptake and nitric oxide release in vascular endothelial cells. *Biochem Biophys Res Commun* 1991;180(2):926-32
120. Busse R, Ogilvie A, Pohl U. Vasomotor activity of diadenosine triphosphate and diadenosine tetraphosphate in isolated arteries. *Am J Physiol* 1988;254(5 Pt 2): H828-32
121. De Mey JG, Claeys M, Vanhoutte PM. Endothelium-dependent inhibitory effects of acetylcholine, adenosine triphosphate, thrombin and arachidonic acid in the canine femoral artery. *J Pharmacol Exp Ther* 1982; 222(1):166-73
122. Ikuta T, Thatte HS, Tang JX, et al. Nitric oxide reduces sickle hemoglobin polymerization: potential role of nitric oxide-induced charge alteration in depolymerization. *Arch Biochem Biophys* 2011;510(1):53-61
123. Lockwood SY, Erkal JL, Spence DM. Endothelium-derived nitric oxide production is increased by ATP released from red blood cells incubated with hydroxyurea. *Nitric Oxide* 2014;38:1-7
124. Sankaran VG, Xu J, Byron R, et al. A functional element necessary for fetal hemoglobin silencing. *N Engl J Med* 2011; 365(9):807-14
125. Song YX, Yue ZY, Wang ZN, et al. MicroRNA-148b is frequently down-regulated in gastric cancer and acts as a tumor suppressor by inhibiting cell proliferation. *Mol. Cancer* 2011;10:1
126. Chen SY, Wang Y, Telen MJ, Chi JT. The genomic analysis of erythrocyte microRNA expression in sickle cell diseases. *PLoS One* 2008;3(6):e2360
- **Nice article showing the micronome of erythrocytes from patients.**
127. Byon JC, Papayannopoulou T. MicroRNAs: Allies or foes in erythropoiesis? *J Cell Physiol* 2012;227(1):7-13
128. Hahn CK, Lowrey CH. Eukaryotic initiation factor 2alpha phosphorylation mediates fetal hemoglobin induction through a post-transcriptional mechanism. *Blood* 2013;122(4):477-85
129. Hahn CK, Lowrey CH. Induction of fetal hemoglobin through enhanced translation efficiency of gamma-globin mRNA. *Blood* 2014;124(17):2730-4
- **Good paper showing the potential therapeutic mode of increasing translational efficiency of HbF.**
130. US National Institutes of Health. *ClinicalTrials.gov*. 2015
131. Bitoungui VJN, Pule GD, Hanchard N, et al. 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* 2015;19(3):171-9
- **Great paper showing the global distribution of the SCD haplotype and questioning the basis of haplotype nomenclature.**
132. Ware RE, Schultz WH, Yovetich N, et al. Stroke With Transfusions Changing to Hydroxyurea (SWITCH): a phase III randomized clinical trial for treatment of children with sickle cell anemia, stroke, and iron overload. *Pediatr Blood Cancer* 2011; 57(6):1011-17
133. Hong W, Nakazawa M, Chen YY, et al. FOG-1 recruits the NuRD repressor complex to mediate transcriptional repression by GATA-1. *EMBO J* 2005; 24(13):2367-78
134. Braconi C, Huang N, Patel T. MicroRNA-dependent regulation of DNA methyltransferase-1 and tumor suppressor gene expression by interleukin-6 in human malignant cholangiocytes. *Hepatology* 2010;51(3): 881-90
- **Good article showing a clear relationship between miRNAs and DNMT1.**
135. Kodeboyina S, Balamurugan P, Liu L, Pace BS. cJun modulates Ggamma-globin gene expression via an upstream cAMP response element. *Blood Cells Mol Dis* 2010;44(1):7-15
136. Ramakrishnan V, Pace BS. Regulation of gamma-globin gene expression involves signaling through the p38 MAPK/ CREB1 pathway. *Blood Cells Mol Dis* 2011;47(1):12-22
- **Nice paper showing the importance of the MAPKs pathway in HbF production.**
137. Munugalavada V, Dore LC, Tan BL, et al. Repression of c-kit and its downstream substrates by GATA-1 inhibits cell proliferation during erythroid maturation. *Mol Cell Biol* 2005;25(15):6747-59
138. Lawrie CH. microRNA expression in erythropoiesis and erythroid disorders. *Br J Haematol* 2010;150(2):144-51
- **Great article discussing the potential role of the micronome in erythroid diseases.**
139. Lawrie CH. microRNA expression in erythropoiesis and erythroid disorders. *Br J Haematol* 2010;150(2):144-51
140. Zhao G, Yu D, Weiss MJ. MicroRNAs in erythropoiesis. *Curr Opin Hematol* 2010; 17(3):155-62
141. Sangokoya C, Telen MJ, Chi JT. microRNA miR-144 modulates oxidative stress tolerance and associates with anemia severity in sickle cell disease. *Blood* 2010; 116(20):4338-48

3.2.2. Pule GD, Mowla S, Novitzky N, Wonkam A. Hydroxyurea down-regulates *BCL11A*, *KLF-1* and *MYB* through miRNA-mediated actions to induce γ -globin expression: implications for new therapeutic approaches of sickle cell disease. *Clinical Translational Medicine*. 2016 Dec;5(1):15.

Abstract

Background: The major therapeutic benefit of hydroxyurea, the only FDA-approved pharmacologic treatment for sickle cell disease (SCD), is directly related to fetal hemoglobin (HbF) production that leads to significant reduction of morbidity and mortality. However, potential adverse effects such as infertility, susceptibility to infections, or teratogenic effect have been subject of concerns. Therefore, understanding HU molecular mechanisms of action, could lead to alternative therapeutic agents to increase HbF with less toxicity. This paper investigated whether HU-induced HbF could operate through post-transcriptional miRNAs regulation of *BCL11A*, *KLF-1* and *MYB*, potent negative regulators of HbF. Both ex vivo differentiated primary erythroid cells from seven unrelated individuals, and K562 cells were treated with hydroxyurea (100 μ M) and changes in *BCL11A*, *KLF-1*, *GATA-1*, *MYB*, β - and γ -globin gene expression were investigated. To explore potential mechanisms of post-transcriptional regulation, changes in expression of seven targeted miRNAs, previously associated with basal γ -globin expression were examined using miScript primer assays. In addition, K562 cells were transfected with miScript miRNA inhibitors/anti-miRNAs followed by Western Blot analysis to assess the effect on HbF protein levels. Direct interaction between miRNAs and the *MYB* 3'-untranslated region (UTR) was also investigated by a dual-luciferase reporter assays.

Results: Down-regulation of *BCL11A* and *MYB* was associated with a sevenfold increase in γ -globin expression in both primary and K562 cells ($p < 0.003$). Similarly, *KLF-1* was down-regulated in both cell models, corresponding to the repressed expression of *BCL11A* and β -globin gene ($p < 0.04$). HU induced differential expression of all miRNAs in both cell models, particularly miR-15a, miR-16, miR-26b and miR-151-3p. An HU-induced miRNAs-mediated mechanism of HbF regulation was illustrated with

the inhibition of miR-26b and -151-3p resulting in reduced HbF protein levels. There was direct interaction between miR-26b with the *MYB* 3'-untranslated region (UTR).

Conclusions: These experiments have shown the association between critical regulators of γ -globin expression (*MYB*, *BCL11A* and *KLF-1*) and specific miRNAs; in response to HU, and demonstrated a mechanism of HbF production through HU-induced miRNAs inhibition of *MYB*. The role of miRNAs-mediated post-transcriptional regulation of HbF provides potential targets for new treatments of SCD that may minimize alterations to the cellular transcriptome

Nature of publication: Original full article

Journal/Publisher: Clinical and Translational Medicine; ISSN: 2001-1326.

Candidate contribution: Conceived and designed the experiments, performed the experiments, analysed the data, wrote the paper, revised and approved the manuscript.

Co-authors contribution: Conceived and designed the experiments: **GP**, SM, AW. Performed the experiments: **GP**. Analysed the data: **GP**, SM, AW. Contributed reagents/materials/analysis tools: SM, NN, AW. Wrote the article: **GP**, AW. Revised and approved the manuscript: **GP**, SM, NN, AW.

RESEARCH

Open Access



Hydroxyurea down-regulates *BCL11A*, *KLF-1* and *MYB* through miRNA-mediated actions to induce γ -globin expression: implications for new therapeutic approaches of sickle cell disease

Gift Dineo Pule¹, Shaheen Mowla², Nicolas Novitzky² and Ambroise Wonkam^{1*}

Abstract

Background: The major therapeutic benefit of hydroxyurea, the only FDA-approved pharmacologic treatment for sickle cell disease (SCD), is directly related to fetal hemoglobin (HbF) production that leads to significant reduction of morbidity and mortality. However, potential adverse effects such as infertility, susceptibility to infections, or teratogenic effect have been subject of concerns. Therefore, understanding HU molecular mechanisms of action, could lead to alternative therapeutic agents to increase HbF with less toxicity. This paper investigated whether HU-induced HbF could operate through post-transcriptional miRNAs regulation of *BCL11A*, *KLF-1* and *MYB*, potent negative regulators of HbF. Both ex vivo differentiated primary erythroid cells from seven unrelated individuals, and K562 cells were treated with hydroxyurea (100 μ M) and changes in *BCL11A*, *KLF-1*, *GATA-1*, *MYB*, β - and γ -globin gene expression were investigated. To explore potential mechanisms of post-transcriptional regulation, changes in expression of seven targeted miRNAs, previously associated with basal γ -globin expression were examined using miScript primer assays. In addition, K562 cells were transfected with miScript miRNA inhibitors/anti-miRNAs followed by Western Blot analysis to assess the effect on HbF protein levels. Direct interaction between miRNAs and the *MYB* 3'-untranslated region (UTR) was also investigated by a dual-luciferase reporter assays.

Results: Down-regulation of *BCL11A* and *MYB* was associated with a sevenfold increase in γ -globin expression in both primary and K562 cells ($p < 0.003$). Similarly, *KLF-1* was down-regulated in both cell models, corresponding to the repressed expression of *BCL11A* and β -globin gene ($p < 0.04$). HU induced differential expression of all miRNAs in both cell models, particularly miR-15a, miR-16, miR-26b and miR-151-3p. An HU-induced miRNAs-mediated mechanism of HbF regulation was illustrated with the inhibition of miR-26b and -151-3p resulting in reduced HbF protein levels. There was direct interaction between miR-26b with the *MYB* 3'-untranslated region (UTR).

Conclusions: These experiments have shown the association between critical regulators of γ -globin expression (*MYB*, *BCL11A* and *KLF-1*) and specific miRNAs; in response to HU, and demonstrated a mechanism of HbF production through HU-induced miRNAs inhibition of *MYB*. The role of miRNAs-mediated post-transcriptional regulation of HbF provides potential targets for new treatments of SCD that may minimize alterations to the cellular transcriptome.

Keywords: Hydroxyurea, γ -Globin expression, Human erythroid, *BCL11A*, *KLF-1*, *MYB*, miRNAs-mediated actions

*Correspondence: ambroise.wonkam@uct.ac.za

¹ Division of Human Genetics, Faculty of Health Sciences, University of Cape Town, Anzio Road, Observatory, Cape Town 7925, Republic of South Africa

Full list of author information is available at the end of the article

Background

Sickle cell disease (SCD) occupies a prominent place in human genetics, as the paradigmatic example of a monogenic disorder, that was described over 100 years ago, and in 1949, it became the first human disease to be deciphered at the molecular level [1]. SCD is a monogenic, hematological and multi-organ disorder associated acute and chronic illness, and progressive organ damage [2]. The disease is due to a single point mutation (Glu6Val) that causes polymerization of the mutant hemoglobin (Hb) S, resulting in sickling of erythrocytes. Inflammation, hemolysis, microvascular obstruction and organ damage characterize the clinical expression of SCD, which is highly variable in individual patients [3]. Sub-Saharan Africa (SSA) has the highest disease burden with approximately 305,800 affected new-borns per year, accounting for 0.74 % of all births in the region (Modell et al. 2008) and 80 % of annual global affected child births [4]. Despite this high incidence, the life-saving public health programs have not been implemented in most SSA countries often associated with limited medical resources and infrastructures. As a consequence, neonatal and childhood mortality due to SCD remains high and estimates suggest that without intervention, up to 90 % of affected children in SSA die by age five from SCD [4, 5]. Nevertheless, SCD patients manifest with wide varying degrees of severity. In combination with environmental factors, several genomic loci, influence the clinical course of SCD. Indeed, despite the fact that there are several key phenotypes of SCD (anemia, stroke, infections), fetal hemoglobin (HbF) has emerged as a central disease modifier [3], that is amenable to therapeutic manipulation [6]. Genetic variants at three principal loci, *BCL11A*, *HBS1L-MYB* and *HBB* cluster account for 10–20 % of HbF variation [7–9]; among SCD patients in USA and Brazil [10], Tanzania [5] and Cameroon [11]. Currently, hydroxyurea (HU) is the only FDA-approved pharmacologic treatment for the induction of HbF in patients with SCD. The major HU benefit is directly related to its HbF-producing effect [12, 13] that leads to significant reduction of pain, acute chest episodes, mortality and the need for blood transfusions [14–17]. HU has also been associated with clinical drift, where physicians use the drug for related complications of SCD such as stroke prevention, priapism and pulmonary hypertension [18]. However potential short and long term adverse effects such as infertility [18–20], susceptibility to infections [21–24], potential teratogenic effect [25, 26], have also been associated with HU. The fear of such side-effects has been a subject of concern to some professionals [27, 28], parents as well as patients [29–33] and a potential barrier to compliance in some settings [34, 35]. As a consequence, HU is still

underutilized [29, 30]. It is then urgent to fully understand HU molecular mechanisms of action, in order to explore alternative and potential less toxic and more acceptable agents that could equally increase the level of HbF.

Several other HU-mediated mechanisms of disease amelioration have been reported including production of nitric oxide, regulation of soluble guanylyl cyclase, cyclic adenosine and guanosine monophosphate [36, 37] as well as erythropoietic stress response [38]. Furthermore, various signalling pathways have been implicated in HU-mediated fetal hemoglobin (HbF) induction such as the $G_{i\alpha}$ /JNK/Jun [39]; p38/MAPK/CREB1 [40]; cAMP-mediated response [41, 42]; erythropoietin (EPO)-induced activation of the ERK-1/ERK-2 MAPK [43]; histone deacetylase (HDAC) and DNA methyl-transferase (DNMT) inhibitor-mediated epigenetic modification of γ -globin expression. Despite this, a complete understanding of HU-mediated HbF production remains incomplete.

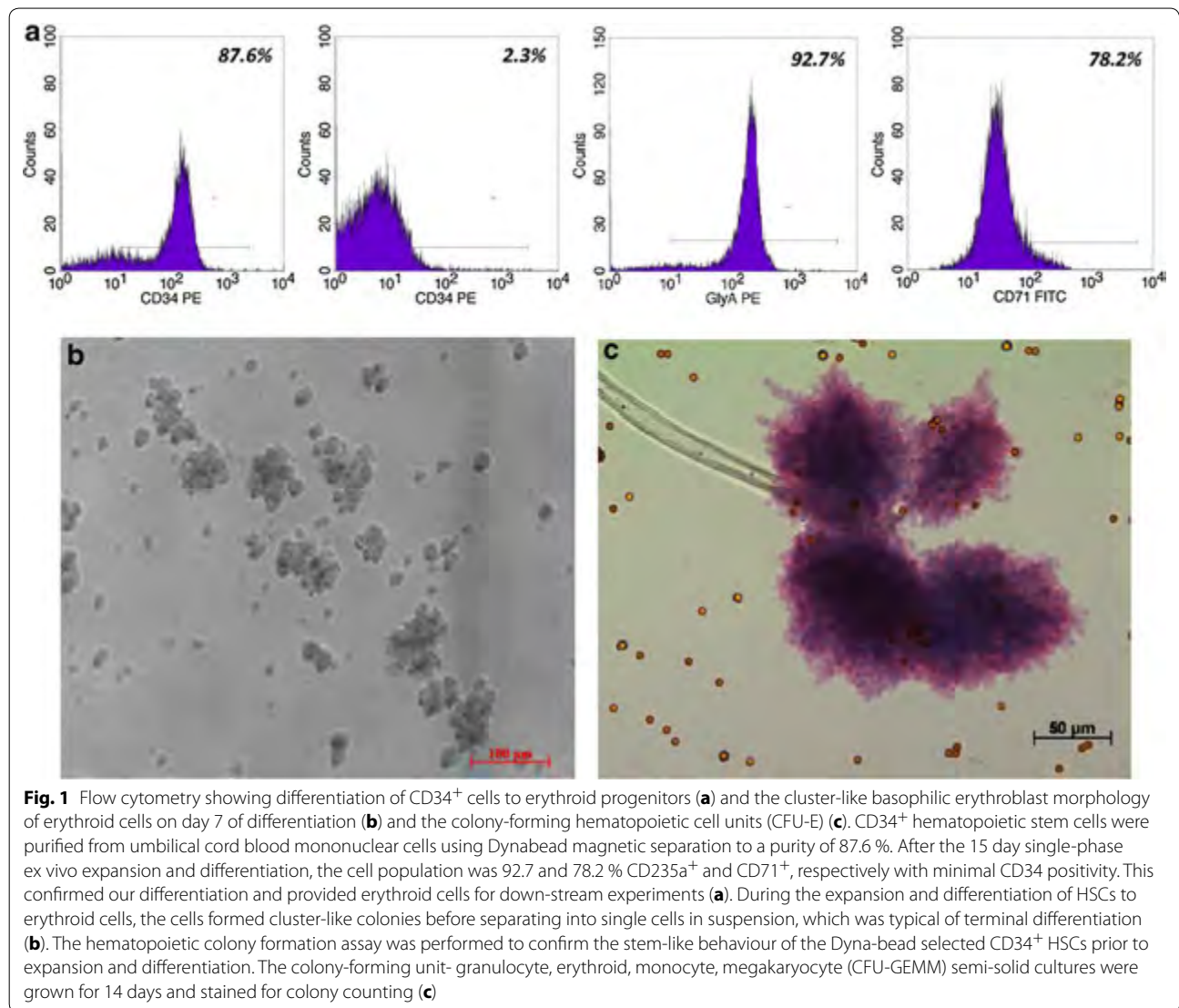
Post-transcriptional regulation of γ -globin expression through micro RNAs (miRNAs) has been shown to play an important role in HU-mediated HbF induction as HU causes differential expression of a suite of miRNAs associated with basal and γ -globin expression at maximum tolerated dose (MTD) in SCD patients [44, 45]. Likewise, DNA methylation has been significantly associated with baseline HbF [38, 46, 47] but provided no substantial explanation for HbF induction in response to HU. Small non-coding RNAs, particularly miRNAs, however, have emerged as powerful regulators and modifiers of gene expression through inhibition of mRNA translation [48], which has implications for hematopoiesis and erythropoiesis [49, 50], particularly in the distinct miRNA expression patterns in SCD patients [51] and the severity of anaemia [52]. Moreover, a few reports have specifically implicated the miRNAs in the regulation of HbF [44, 53, 54].

This study has further investigated whether the induction of HbF by HU treatment could operate through post-transcriptional regulation of *BCL11A*, *KLF-1* and *MYB*, by analysing changes in expression of select miRNAs in erythroid cells derived from umbilical cord blood CD34⁺ hematopoietic stem cells and the K562 cell line.

Results

CD34⁺ hematopoietic cells

The average volume of umbilical cord blood collected was 108 ml (± 25 ml) with a 1.3 % (± 0.5 %) yield of CD34⁺ cells at an average 87 % (± 6 %) percentage purity of the cell population after Dynabead magnetic separation technology (Fig. 1a). The CFU-GEMMs also produced several burst-forming unit-erythroid cells (BFU-E) and erythroid clusters typically formed during expansion (Fig. 1b).



Ex-vivo differentiation of CD34⁺ cells

CD34⁺ cells were cultured using an ex vivo single-phase expansion and differentiation protocol and all primary CD34⁺ primary lines were successfully expanded (mean fold expansion 4.2; \pm 2.7) and differentiated into erythroid progenitors with average expression of 84 % (\pm 3 %) and 83 % (\pm 8 %) for CD235a and CD71, respectively (Fig. 1a). The first half of the expansion and differentiation was characterized by cluster-like colonies of erythroid cells and sparse loose-lying cells (Fig. 1b, c). During early differentiation (days 1–3) large blasts were observed in culture followed by proerythroblast morphology around day 5. The morphology typical of basophilic erythroblasts was observed on days 7–8 (Fig. 1b) and on day 15, the majority of the cells had ortho/polychromatophilic morphology, which was confirmed by the expression of cell-surface markers CD71 and CD235a

(Additional file 1: Figure S1), at which point differentiation was halted and cells treated with HU for various analyses.

WST-1 cell proliferation assay

To determine the optimal concentration and exposure time to HU, K562 cells (8×10^4 cells) were treated with varying doses of HU and for 2, 4, 6, 12 and 24 h. Six hour (6) exposure time and 100 μ M were determined to be optimal as there was minimal initial cytotoxic effect and sufficient cellular metabolic activity to alter gene expression in response to HU (Additional file 2: Figure S2).

Hydroxyurea induces γ -globin expression in both erythroid and K562 cells

Erythroid progenitors from seven unrelated individuals and the K562 cell line were successfully analysed for

differential gene expression in response to HU treatment, with three technical and independent qPCR experimental repeats for K562 cells and one experimental repeat for ex vivo derived erythroid cells due to the limited number of cells. Subsequently, four of the ex vivo derived lines were analysed for corresponding miRNA expression. HU treatment of K562 cells showed a sigmoidal pattern of expression in *BCL11A* transcription with an apparent down-regulation at 6 h, an expression pattern inverse to the γ -globin that showed a sevenfold upregulation at the same HU exposure time (Fig. 2). *KLF-1*, a critical activator of *BCL11A* [55], and *MYB*, a potent activator of *KLF-1* [56, 57] and negative regulator of γ -globin [58] had similar expression patterns and were significantly down-regulated at 12 h ($p < 0.03$). β -Globin and *GATA-1* expression remained largely unaffected by HU treatment with the exception of a slight down-regulation also after 12 h treatment.

The gene expression pattern in the primary erythroid lines was largely similar to that in K562 cells with an inverse relationship between γ -globin expression and its critical regulators; *BCL11A*, *KLF-1* and *MYB*. Some primary lines had a sevenfold ($p < 0.003$) up-regulation of γ -globin expression 6 h after HU treatment associated with the down-regulation of *MYB* ($p < 0.04$) and *BCL11A* expression (Fig. 3). Gene expression analysis was done after successful ex vivo differentiation of primary erythroid cells and HU treatment for 24 h.

Hydroxyurea up-regulates miRNAs associated with inhibition of *BCL11A* and *MYB*

In the primary lines analysed for miRNAs expression 6 h post treatment, there was inter-individual variation in expression however, miR-26b was significantly down-regulated in all but one primary line (Fig. 4). HU treatment also caused sigmoidal time-dependent changes in miRNAs expression in K562 cells (Additional file 3: Figure S3), with significant up-regulation of miR-15a and miR-16-1 at 6 h (p value 0.027 and 0.002, respectively), which are known inhibitors of *MYB* [54]. MiR-151-3p and miR-451 were also significantly up-regulated at 6 h post treatment (0.041 and 0.042, respectively) and although not statistically significant, miR-494 had a twofold increase. All miRNAs were down-regulated 12 h after treatment and returned to baseline after 24 h.

MicroRNA inhibition

The inhibition of miR-26b and miR-151-3p resulted in a significant decrease in γ -globin expression (Fig. 5), which could be partially rescued when cells were treated with HU. This suggests that most miRNAs target negative regulators of HbF as their inhibition causes up-regulation of γ -globin expression. At a higher concentration (25 nM),

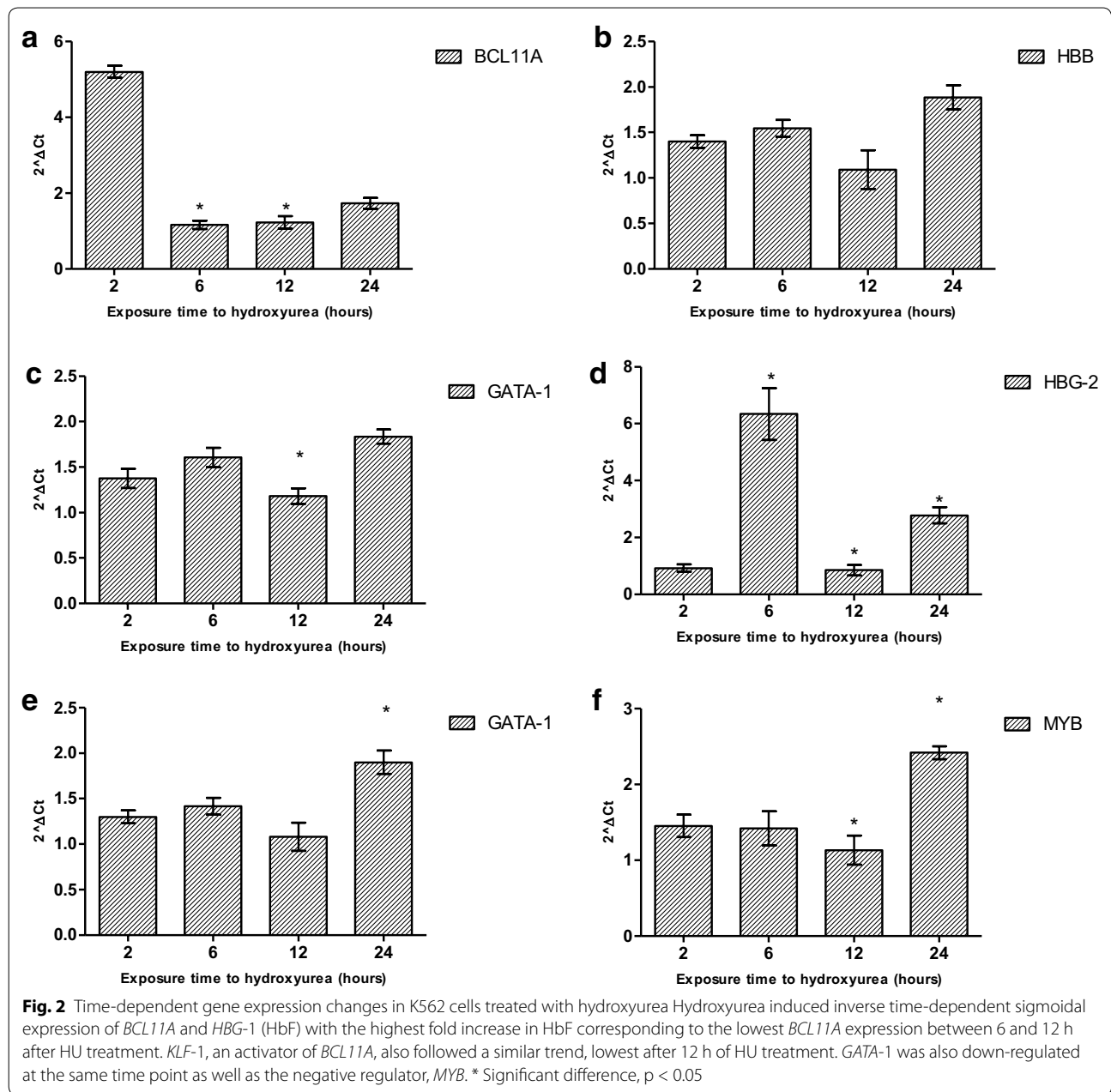
all miRNAs down-regulated HbF (Additional file 4: Figure S4). Cells were co-transfected with select pairing of miRNAs in order to investigate the combinatorial effect of anti-miRNAs. The effect of miR-26b/miR-151-3p and miR-151-3p/miR-494 were comparable in the decrease of HbF. Although the inhibition of miR-494 had no apparent effect on HbF, co-transfection with miR-26b resulted in near-complete transient knock-down of HbF.

MiR-26b directly targets *MYB* 3'-UTR

Given the clear association between (i) HU and *MYB* expression; (ii) HU and our candidate miRNAs and (iii) the miRNAs and *MYB* expression patterns, dual-luciferase reporter assays were conducted using the pGL3-*MYB*-3'-UTR to investigate the potential direct interactions between our candidate miRNAs and the *MYB*-3'-UTR (1191 base pairs). The amount of luciferase activity from the vector is directly proportional to the miRNAs occupancy of the UTR, which is known to induce mRNA transcript degradation via the RNA-induced silencing complex (RISC) [59]. There was a 2.2-fold increase in luciferase activity when cells were co-transfected with anti-miR-26b (Fig. 6a) and similarly, a time-dependent change in luciferase activity in response to HU treatment, with the highest luciferase activity observed after 12 h of HU treatment. Similarly, however less apparent, luciferase activity was observed with the transfection of anti-miR-151-3p, anti-miR-451 and anti-miR-494 (Additional file 5: Figure S5). To confirm that miR-26b directly interacts with the *MYB* 3'-UTR and initiates RISC-mediated mRNA degradation, there was gradual recovery of the luciferase activity with the decrease in concentration of the transfected anti-miR-26b (Fig. 6b). Therefore, at low anti-miR-26b concentration (2.5 nM), endogenous miR-26b is permitted to bind the *MYB* 3'-UTR, which would result in *MYB* mRNA degradation and therefore up-regulation of γ -globin expression.

Discussion

Although HU has demonstrated significant clinical improvements in SCD patients, a complete understanding of the myriad of molecular mechanisms by which HU induces the disease-ameliorating HbF remain elusive [60]. This paper demonstrates the post-transcriptional effect of HU on critical regulators of γ -globin expression and associated miRNAs using two models; erythroid cells derived from umbilical cord blood CD34⁺ HSCs and K562 cells. Cord blood-derived CD34⁺ cells are an abundant, non-invasive and largely unutilised source of HSCs and a good model to investigate erythropoiesis [61, 62] and elucidate the molecular mechanisms of HbF production in response to therapeutic agents [63]. The volume



of cord blood collected in this study was consistent with expected norms [64] although the average CD34⁺ cell yield was slightly lower, possibly due to our stringent cell separation method. Erythroid cells were successfully differentiated using an ex vivo single-phase expansion and differentiation protocol [65] with minor alterations to achieve optimal ex vivo expansion [66]. Given that HU is thought to induce γ -globin via perturbation of erythroid cell maturation [36, 67], in this study HSCs were differentiated to orthochromatophilic erythroblasts for 15 days. It has been previously demonstrated that cells treated

with HU around day 6 and harvested at day 10 of differentiation consist mostly of basophilic and polychromatophilic normoblasts, comparable to untreated cultures at a similar stage [68]. In another study, investigating the effect of HU on *BCL11A*, *KLF-1* and *TAL-1* in bone marrow-derived basophilic erythroblasts, post terminal differentiation, there was no significant difference in the expression of the markers of late erythroid differentiation, CD235a and CD71, between HU treated and untreated cells [54]. Likewise in this study, the comparable levels of CD235a and CD71 between pairwise samples (HU+ and

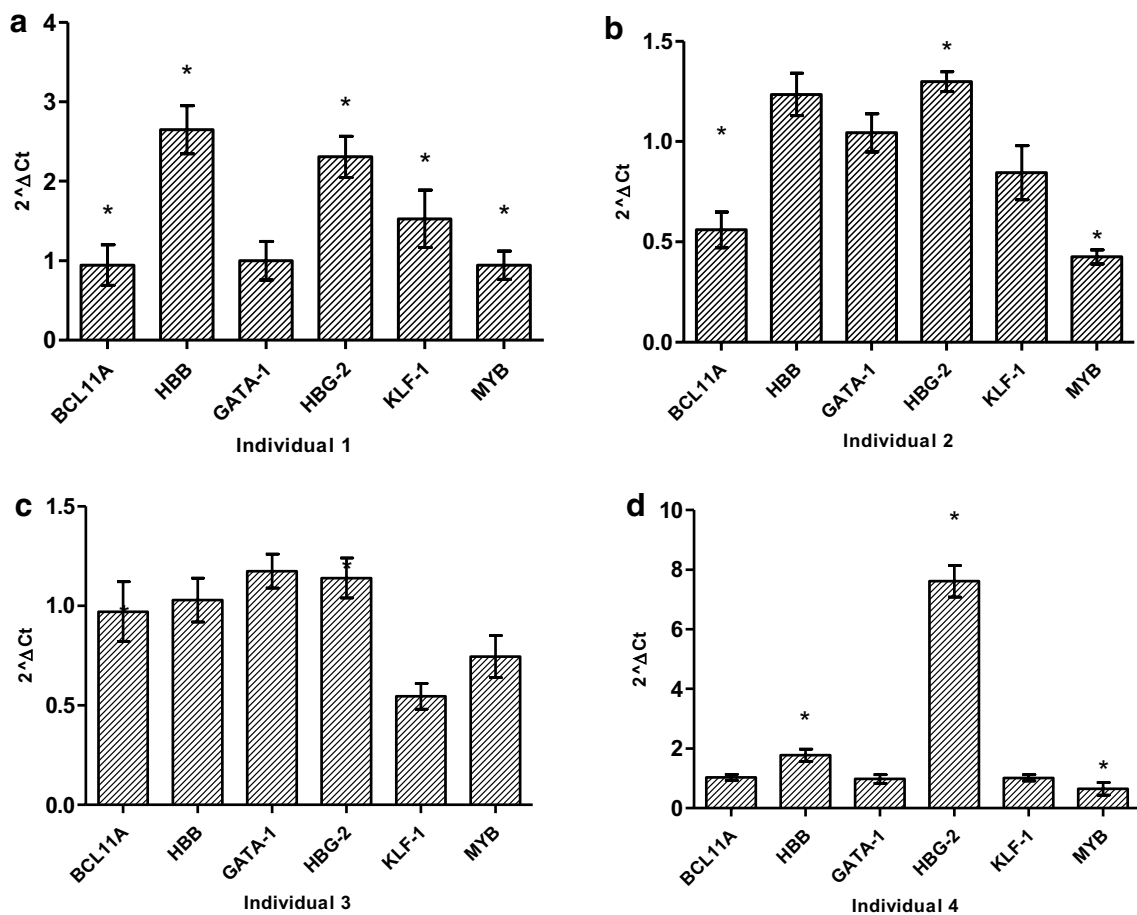


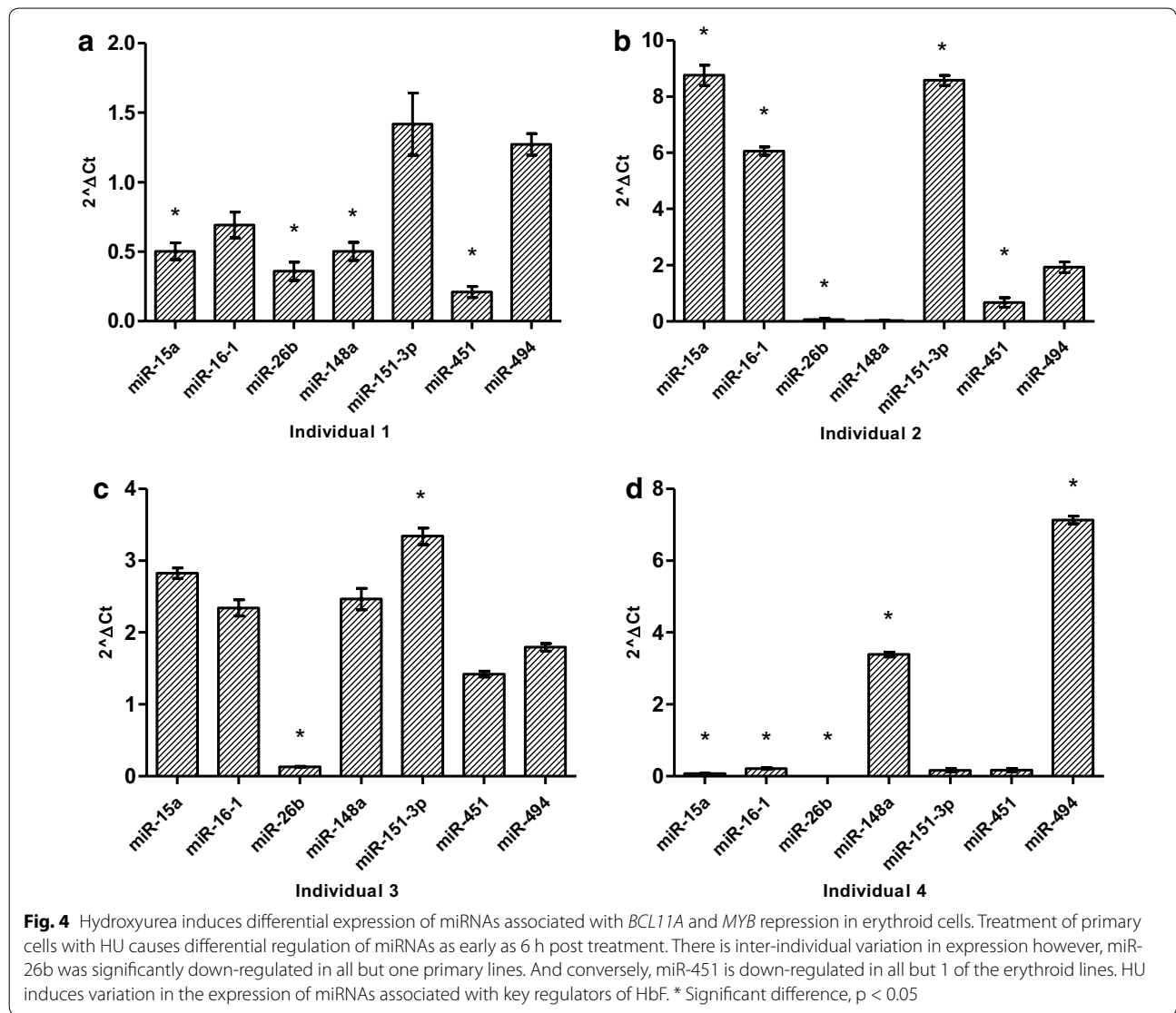
Fig. 3 Hydroxyurea-induced gene expression changes in ex vivo derived erythroid cells. The gene expression profiles of primary erythroid lines after HU treatment for 6 h. Although the expected inter-individual variation in expression was observed, there were several common trends such as the low *BCL11A*, *KLF-1* and *MYB*; and significant upregulation of HbF expression. *HBB* (β -globin) expression was always associated with *BCL11A*, which is expected given the role of this transcription factor in the “fetal switch”. * Significant difference, $p < 0.05$

HU-) confirms that at the point of expression analyses, the paired samples were at a similar maturation stage (Additional file 1: Figure S1), and thus the erythroid cell maturation cannot account for the observed changes in expression. It is thus more likely that the observed induction of γ -globin is a result of the repression of *BCL11A*, *KLF-1* and *MYB*. Therefore, with further development of expansion and differentiation protocols, the use of cord blood stem cells can be extended to routine laboratory disease modelling and testing of therapeutic agents above and beyond the curative transplantation value of this biological material.

The WST-1 assay, a non-radioactive spectrophotometric quantification of cellular metabolic activity and proliferation was used to determine the optimal HU concentration and cellular incubation periods within a 24 h period given that HU is commonly prescribed to

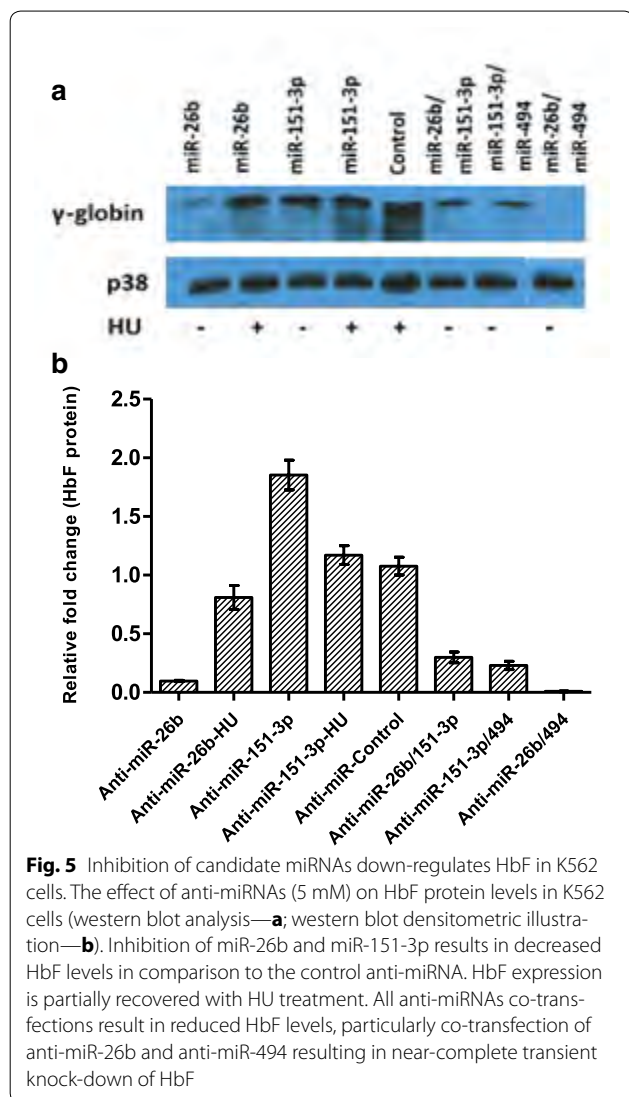
SCD patients as a daily oral drug [69]. The most apparent effects of HU on γ -globin, its critical regulators (*BCL11A*, *MYB*, *KLF-1*) and erythroid transcription factor, *GATA-1*, was between 6 and 12 h after 100 μ M treatment, at which time cellular metabolic activity was sufficient to assess changes in gene expression.

Using paired analysis, we demonstrated a sevenfold ($p < 0.003$) up-regulation of γ -globin expression and corresponding down-regulation of *BCL11A* in both erythroid and K562 cells (Figs. 2, 3). *BCL11A* has been shown to be a critical negative regulator of γ -globin and fundamental to the ‘fetal switch’ from γ - to β -globin expression [58]. Furthermore, it has been successfully shown to be responsive to therapeutic manipulation for potential SCD treatment [65]. The time-dependent patterns of *BCL11A* and γ -globin gene expression are inversely related, with the most apparent effect at 6 h in both erythroid and



K562 cells. Similarly, *KLF-1* was down-regulated in both cell models, corresponding to the repressed expression of *BCL11A* and β -globin gene ($p < 0.04$). This result is consistent with the function of *KLF-1* as an activator of *BCL11A* and its association with haplo-insufficiency-induced hereditary persistence of fetal hemoglobin [55]. *GATA-1* expression was similar to *KLF-1*, which is consistent with evidence that this principal erythroid transcription factor co-occupies the 5' locus control regions and 3' DNase I-hypersensitivity site of the β -globin gene cluster and associates with *BCL11A* in the Mi-2/nucleosome remodelling and deacetylases (NuRD) complex and therefore possibly critical in the repression of HbF production [58, 65, 70]. Down-regulation of *MYB* in both erythroid and K562 cells corresponded with induction of γ -globin expression and was inversely related to miR-26b

as well as miR-15a and miR-16-1, which have been implicated in the elevation of HbF production via directly targeting the *MYB* transcription factor [45]. Although HU induced down-regulation of *BCL11A*; *KLF-1* and *MYB*, the effect on γ -globin expression was variable, which suggests that HU may interact with many genes upstream of γ -globin induction and also induce other post-transcriptional changes of gene expression through miRNAs. Recently similar effects of HU on *BCL11A* and *KLF-1* were demonstrated in late differentiation erythroblasts derived from bone marrow progenitors, also implicating *TAL-1* in the regulation of *BCL11A*; *KLF-1* via promoter-binding and *MYB* repression [54]. Furthermore, K562 cells were previously used to demonstrate the crucial role of *TAL-1* as the DNA-binding component of the *LDB-1* complex regulating the long-range looping of the globin



gene cluster to allow interaction between the 5'-LCR and the β -globin gene [71]. Taken in sum with the results of the present study, HU induces γ -globin through numerous networks comprising *BCL11A*, *KLF-1*, *TAL-1* and *MYB*. The data of this study also demonstrated an association between miR-26b and γ -globin expression, which has previously been shown both at basal and MTD in HU-treated SCD patients [44]. There was a 4.5-fold up-regulation of miR-16-1 and miR-151-3p in K562 cells (Additional file 3: Figure S3), which was related to the most apparent induction of HbF at 6 h after HU treatment. A direct target of DNMT-1, miR-148a [72], was also up-regulated 6 h after HU treatment, which correlated to γ -globin induction. These findings suggest that the observed HbF production may also be through other post-transcriptional regulatory modalities such as methylation inhibition in the β -globin gene cluster.

The inhibition of the miRNAs, particularly miR-26b, unequivocally demonstrates a causative effect by HU on the induction of HbF through miRNA-mediated mechanisms. This is further supported by the near-complete transient knock-down of HbF when miR-26b and miR-494 are co-inhibited. The inhibition of miR-26b, miR-151-3p and miR-451 results in apparent decreases in HbF in comparison to all controls (inhibitor negative control; untransfected and HU treated cells). This effect is likely because miR-26b inhibits *MYB* translation, causing reduced *KLF-1* activation and thus lowering *BCL11A* expression, thereby increasing HbF production (Figs. 5, 6, 7). Therefore these data highlight the critical role played by potent negative regulators of HbF such as *MYB*, *KLF-1* and *BCL11A* and demonstrate that miRNAs provide a viable therapeutically responsive tier of HbF production for SCD treatment (Fig. 7), importantly while sparing the undruggable nature these transcription factors in non-erythroid cells [73]. Overall, the K562 cell experimental repeats were similar and although there was expected inter-individual variation in the degree and timing of up- and/or down-regulation of γ -globin and its key regulators in the erythroid model, the trends remained true and largely comparable. Despite the fact that several studies have and continue to utilise K562 as a model to investigate various components of γ -globin expression [39, 53, 74–77], it could be argued that K562 cells are not an ideal model for studying globin switching because of their bias expression of γ -globin. We however, saw this as an advantage in a number of ways: (1) the aim of this study was not to demonstrate the switching in expression of the globin genes (β to γ) but rather to use a stable cell line model in conjunction with cord blood-derived erythroid cells to demonstrate increases in γ -globin and decreases its regulators (*BCL11A*, *KLF-1* and *MYB*) in response to HU, and for this purpose, K562 cells would suffice as a model; (2) in choosing the whether to use miRNA mimics or anti-miRNAs (miR-inhibitors), the high expression of γ -globin in K562 was considered; if our hypothesis of HU-induced miRNA-mediated repression of *MYB* was correct, the use of mimics would not only introduce unnaturally high levels of miRNAs into this model (which we felt may in fact result in some artefactual results merely because of this unnatural state of aberrant concentration of miRNAs in the cells) but it would also be compounded by any co-treatments with HU in causing γ -globin expression. This rationale led to the choice of anti-miRNAs, which would go toward decreasing the already high levels of γ -globin in K562 (via *MYB*) and also provide an obvious opportunity of the rescue of miR-26b inhibition with HU co-treatment.

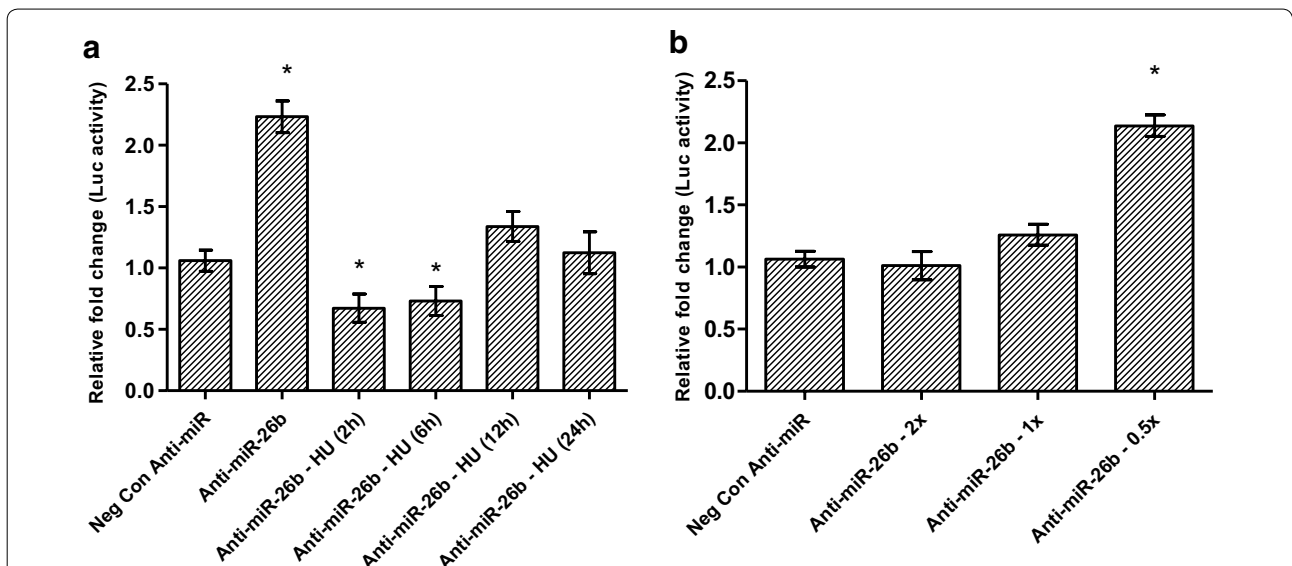
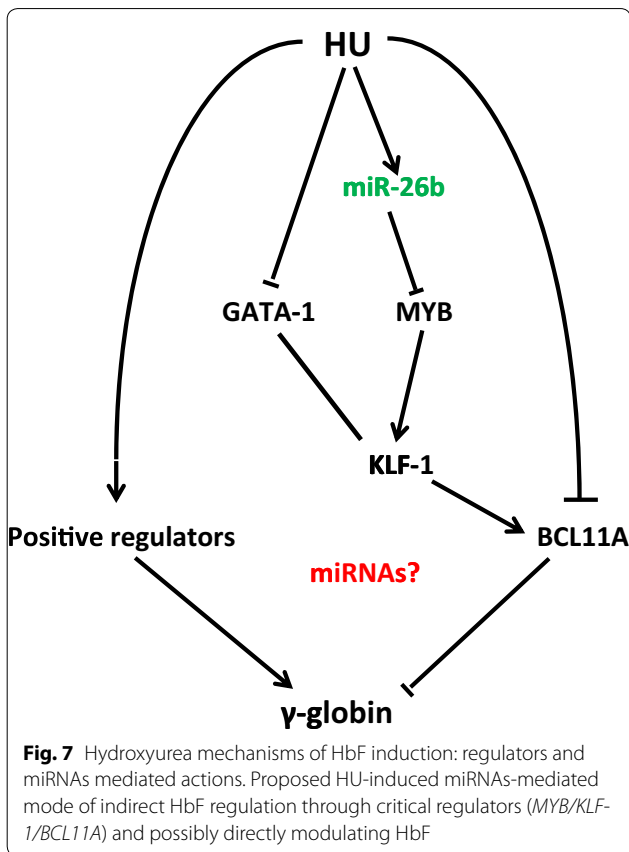


Fig. 6 Time- and concentration-dependent luciferase activity in response to hydroxyurea treatment of anti-miR-26b transfected K562 cells. Inhibition of miR-26b results in a significant ($p < 0.02$) increase in luciferase activity and a sigmoidal pattern of action during a time course with HU treatment, with the highest luciferase activity observed after 12 h of treatment (a). A concentration gradient of anti-miR-26b demonstrates an inverse relationship between anti-miR26b concentration and *MYB*-3'-UTR luciferase activity. This unequivocally demonstrates interaction between miR-26b and the 3'-UTR of *MYB* (b). The likely implication of this result is that the miR-26b-induced translational inhibition of *MYB* (a known activator of *KLF-1*) could result in down-regulation of *KLF-1*, which is known to directly induce *BCL11A* expression. In sum, miR-26b could explain the observed up-regulation of γ -globin via a *MYB/KLF-1/BCL11A* pathway. * Significant difference, $p < 0.05$



Future work: implications for novel therapeutics in SCD

Future work regarding the post-transcriptional mechanisms regulating the expression of various trans-acting factors, such as *BCL11A*, *KLF-1* and *MYB*, could include experimental variations of the level of candidate miRNAs and examination of the effect on HbF levels and additional 3'-UTR luciferase assays on such regulators. The limitation on the number of primary erythroid cells could be overcome by the use of improved primary cell transfection protocols or alternative sources of HSCs such as commercially available HSCs or bone marrow aspirates. Future work will look to replicate these results in primary erythroid cells and investigate other candidate miRNAs for interactions with key regulators of γ -globin expression. Another limitation of the present work the "promiscuous" nature of HU, as it could influence many aspects of cellular functions and almost certainly alter the global cellular transcriptome. Therefore, future studies should continue to evaluate the in vivo impact of efficacious concentrations of HU on the erythroblast transcriptome and/or proteome, as well as the erythroid-specific micro-nome of SCD patients (before treatment and at MTD) as global analysis of these epigenetic mechanisms could highlight multiple components of this complex system and possibly yield alternative (possibly miRNA-based) therapeutic approaches to hemoglobinopathies [78]. This is likely not implausible as several clinical trials in

other diseases have demonstrated a promising future for miRNA-based therapeutics such as the liposome-based human miR-34 mimic (MRX-34) miRNA-based drug against hepatocellular carcinoma (NCT01829971). MiRNAs have also been implicated in many cancers (colorectal, cervical, prostate, breast and lung) as either biomarkers and/or highlighted for potential therapeutic significance [79]. Closer to hemoglobinopathies, several miRNAs have been associated with specific phenotypes or key signalling pathways in SCD; miR-15a and miR-16-1 with hereditary persistence of HbF [53], miR-148a with DNMT-1 [72], miR-144 and tolerance of oxidative stress and anemia severity [52], a mouse model showing the correction of SCD via interference with the “fetal switch” [80, 81], in vivo association of miR-26b and miR-151-3p with HU treatment at MTD [44], miR-320 and down-regulation of CD71 during terminal differentiation of reticulocytes [51], miR-24 and inhibition of erythropoiesis via activin type I receptor *ALK4* [82], a suite of miRNAs including miR-451 in *Plasmodium falciparum* parasite interactions with erythrocytes [83] and now in this study, miR-26b direct interaction with *MYB* 3'-UTR and the rescue of miR-26b inhibition with HU treatment.

It is however noteworthy to consider the myriad of challenges in miRNAs-based therapies including site-specificity, delivery and treatment efficacies, off-targeting and side effects [84], undesired in vivo pharmacokinetics [85] and short half-life in peripheral blood before renal clearance [86]. These challenges may also be compounded by the rapid erythrocyte turnover and anemia in SCD, the heterogeneous multi-organ complications and the numerous genetic polymorphisms affecting disease severity, predisposition to specific phenotypes, the vast variation in response to HU treatment and the overall clinical course of the disease [60]. Although faced with many challenges, novel methodologies such as exosome-based delivery systems provide sufficient evidence in support of the continued efforts to develop improved systems of delivery and target-specificity, desired in vivo stability, reduced side- and off-target effects for miRNA-based therapeutic approaches [87–89].

Conclusion

In the present article, the authors have demonstrated successful expansion and differentiation of umbilical cord blood-derived CD34⁺ HSCs into CD71⁺/CD235a⁺ erythroid cells and showed HU-mediated induction of γ -globin expression concomitant with the down-regulation of key negative regulators, *BCL11A*, *MYB* and *KLF-1*. Importantly, the experiments have demonstrated association between the induction of miR-15a and miR-16-1 and HU-mediated down-regulation of *MYB* as well as up-regulation of miR-148a, which targets DNMT-1.

Furthermore, the data have shown the direct interaction between miR-26b and the *MYB* 3'-UTR, possible indirect modulation of *BCL11A* via *MYB* activation of *KLF-1*, all of which were concomitant with HbF production in both erythroid and K562 cells. The results of this study demonstrate the role of miRNAs in the modulation of HbF, directly and through critical regulators. The elucidation of the post-transcriptional regulation of HbF through miRNAs could incite investigation of therapeutic agents that would avoid global transcriptome changes but rather hone in on critical regulators with lineage-specific miRNA-mediated inhibition of negative regulation of HbF.

Methods

Umbilical cord blood

Umbilical cord blood was harvested during elective caesareans at Mowbray Maternity Hospital (Cape Town, South Africa) from healthy females of black African origins, free of the sickle cell mutation. Cord blood (90–140 ml) was collected in anticoagulant citrate phosphate dextrose adenine (CPDA-1)-containing bags (SSEM Mthembu, South Africa) and processed within 4 h of harvest.

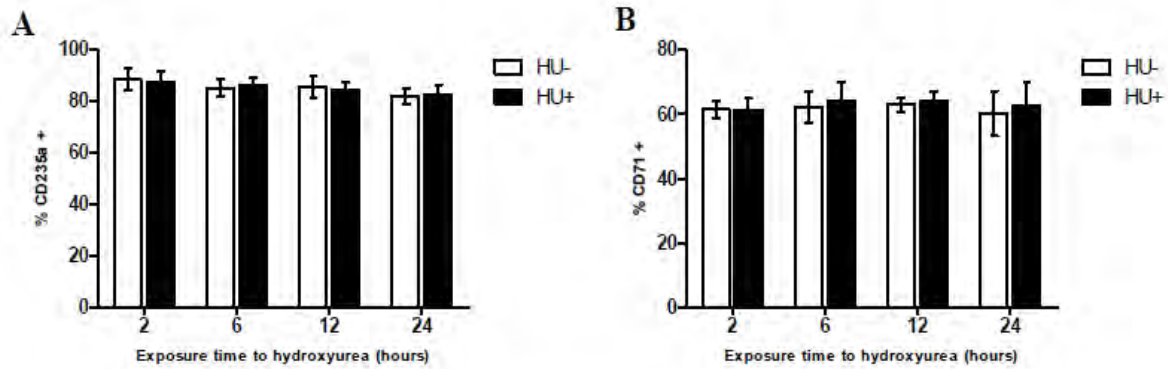
Isolation of CD34⁺ hematopoietic cells

Cord blood samples were diluted using Iscove's Modified Dulbecco Medium (Sigma, SA), layered on Ficoll-Histopaque (1.077 g/ml) (Sigma, SA), centrifuged to collect the mononuclear cell layer (MNCL), then enriched for CD34⁺ hematopoietic stem cells (HSCs) using Dynabead magnetic separation technology according to manufacturer's instructions (Life Technologies, SA).

Flow cytometry

The purity of the Dynabead-selected HSCs and differentiated erythroid cells was determined using cell-surface antibodies and flow cytometry (FACSCalibur, BD Biosciences, USA). Briefly, 5×10^4 HSCs were stained with phycoerythrin-conjugated CD34 (CD34-PE) and fluorescein isothiocyanate-conjugated CD45 (CD45-FITC) (BD Biosciences, USA) according to the manufacturer's instructions and analysed. Subsequent to erythroid differentiation, markers of the derived-erythroid cells were analysed on day 15 using CD34-PE, CD45-APC, CD71-FITC and CD235a-PE (BD Biosciences, USA). The controls included in the flow cytometry experiments were as follows: (i) No cells (cell suspension fluid: 1x PBS), (ii) Cells only, (iii) Cells and primary antibody only, (iv) Cells and secondary antibody only. These were included to ensure specificity in antigen-antibody interactions in the experiments. During optimization, a separate cell line, RAMOS (RA-1 CRL1596, ATCC) was used to

A



Supplementary Figure 1

[40169_2016_92_MOESM1_ESM.pdf](#) **Additional file 1: Figure S1.** Time dependent expression of markers of late erythroid differentiation (CD235a and CD71) in erythroid cells treated with HU. Expression of CD235a (A) and CD71 (B) during late erythroid differentiation show no significant differences between HU treated (+) and untreated (-) cells. The lack of significant changes in the expression of these markers indicates that HU had minimal effect on the processes of erythroid differentiation, thus suggesting that the results, particularly the induction of γ -globin and repression of *BCL11A*, *KLF-1* and *MYB*, were not artefacts of stress erythropoiesis.

3.2.3. Pule GD and Wonkam A. Treatment of Sickle Cell Disease in Africa: should we invest in haematopoietic stem cell transplantation? *Pan African Medical Journal*. 2014 18:46.

Abstract

This commentary paper discusses the state of treatment for Sickle Cell Disease in Africa compared to the West and poses the question of whether we should invest in hematopoietic stem cell transplantation.

Nature of publication: Original full article

Journal/Publisher: Pan African Medical Journal; ISSN: 1937-8688

Candidate contribution: Performed the main search of the articles, drafted the manuscript and compiled the revisions

Co-authors contribution: **GP** performed the main search of the articles, drafted the manuscript and compiled the revisions. **AW** conceived the article and revised the draft, both authors approved the final version of the manuscript.

Commentary

Treatment for sickle cell disease in Africa: should we invest in haematopoietic stem cell transplantation?

Gift Pule¹, Ambroise Wonkam^{1&}

¹Division of Human Genetics, Department of Clinical Laboratory Sciences, Faculty of Health Sciences, University of Cape Town (UCT), Cape Town, South Africa

[&]Corresponding author: Ambroise Wonkam, Division of Human Genetics, Department of Clinical Laboratory Sciences, Faculty of Health Sciences, University of Cape Town (UCT), Cape Town, South Africa

Key words: Sickle Cell Disease, hydroxyurea, haematopoietic stem cells transplantation, Africa

Received: 27/01/2014 - Accepted: 15/02/2014 - Published: 13/05/2014

Pan African Medical Journal. 2014; 18:46 doi:10.11604/pamj.2014.18.46.3923

This article is available online at: <http://www.panafrican-med-journal.com/content/article/18/46/full/>

© Gift Pule et al. The Pan African Medical Journal - ISSN 1937-8688. This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/2.0>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Commentary

Epidemiology and burden of SCD

Sickle cell disease is a life-long genetic disease that begins in childhood, affecting the structure of erythrocytes, altering the healthy biconcave shape to a crescent shape, leading to the blockage of veins, thereby resulting in organ damage [1]. There is strong correlation between the frequency of the HbS gene and the historical distribution and incidences of malaria due to the partial HbS-carrier resistance to *Plasmodium falciparum* malaria [2]. Indeed, Sickle Cell Anaemia mutation (HbS gene) appears to have occurred independently in 4 regions in Africa, defined by four haplotypes (Senegal, Benin, Bantu and Cameroon haplotypes)[3]. SCD is prevalent among indigenous populations in tropical regions of Africa and Asia; 305800 births with SCD are estimated to occur annually, nearly 67% of which take place in Africa. Sickle Cell Anaemia (SCA; the homozygous HbSS state) is by far the most prevalent and severe form of SCD [4]. Many countries in Africa have developed a national control program for SCD, however provisions of neonatal screening are rare [5] and development of specialized centres for lifelong medical care and surveillance have yet to become part of many SCD health systems, and in the absence of universal medical insurance coverage in many African countries, the chronic care of SCD patients is therefore dependent on financial support and care-giving by family member [6]. In addition, vaso-occlusive painful events, silent and overt stroke that occur in SCD could potentially contribute to functional limitations and poor academic achievement of affected children. Indeed, it was reported in Cameroon that up to 37.5 % of participants' SCD-affected children had mild-to-severe cognitive deficits, and there was a significant effect on executive functions and attention [7]. Poor health status of children with SCD could also reduce caregivers' employability and worsen the socioeconomic burden on families. Indeed up to 24.3 % of caregivers in the USA missed two or more days of work per 3 days-hospital admission of their children [8], and the morbidity of a painful event continued after discharge from hospital [8,9]. Similar findings were also recently reported in Cameroon[6].

The mortality rate associated with SCD has remained high in Africa, despite the use of appropriate interventions to manage the various forms of crises [10]. In the USA and Europe, who together account for less than 8% of the global disease burden of SCD, new-born screening, pneumococcal immunization, prophylactic penicillin and most importantly HU treatment, have decreased morbidity and mortality and thus increasing survival rates from childhood diagnoses to over 95% [4,11]. In stark contrast, as of 2010, sub-Saharan Africa accounted for 75.5% of the global number of new-borns with SCD, where most of these children die before age 5 due to a myriad of socio-economic factors and a poor public healthcare system [4]. The limited early detection and treatment initiatives that have been implemented in Africa result in high death rates before the age of 5 [10,12]. These statistics highlight the imperative necessity of research and translational medicine in to improve the burden through better care and potentially a cure of SCD in Africa.

Treatment approaches

There are five treatment approaches for SCD that are tailored to the clinical phenotype of a patient, namely supportive, symptomatic, preventative, abortive and curative approaches [13]. The supportive approach is the most common, aimed at the management of the patient and such an approach includes a balanced diet, hydration and folic acid supplementation. Blood transfusions, analgesia and

antibiotics are typed as symptomatic approaches because their function is to alleviate specific SCD symptoms. The preventative approach is taken to preclude the occurrence of disease complications such as pneumonia and influenza vaccination, hydroxyurea for the induction of foetal haemoglobin (HbF) and blood transfusions to avert primary and secondary stroke episodes [14]. Nitric oxide (NO) is the only accepted agent for the abortive approach, reported to completely terminate of chronic pain episodes in some SCD patients [15]. Lastly, the curative approach is the ultimate goal for all genetic disorders, intended to correct the disease-causing mutation and prevent all complications. Currently, transplantation of haematopoietic stem cells (HSCs) is the only accepted curative treatment for SCD. Below, we briefly describe the 3 current major strategies for effective treatment of SCD, namely blood transfusion, hydroxyurea (HU) and HSC transplantation.

Blood transfusion: Blood transfusions improve the oxygen-carrying capacity and oxygen-delivery efficiency of blood to the tissues and decrease the blood concentration of sickle cells and improve perfusion of tissue microvasculature. Transfusions are typically used to ameliorate chronic anaemia and pain episodes [16], and are highly effective in patients with sporadic episodes of severe anaemia by preventing organ damage. Although transfusion can be applied as a preventative, abortive or curative approach for treatment of certain complications, they have many drawbacks. There are few blood transfusion services across the continent with the capacity to cope with the demand for regular transfusions [17,18]. Furthermore, widespread diseases such as HIV/AIDS and TB reduce the number of possible donors as well as the limited technical and financial support offered by the state's department of health. Although not specific to Africa, graft versus host disease (GVHD), even after alloimmunization [19], is one of the challenges surrounding transfusions. Other possible complications include transfusion-induced haemolytic reaction or autoimmune hyper-haemolysis [20] and blood volume and iron overload [21]. This evidence suggests that although useful and effective in certain circumstances, blood transfusion alone is not a sufficient treatment for SCD nor is it curative. Therefore below, we discuss the use hydroxyurea as a drug treatment for SCD.

Hydroxyurea: Hydroxyurea is an oral, cytotoxic, anti-metabolic and -neoplastic drug for principal haematological disorders. The first clinical application of HU was in 1984 [22] and since then have been supported by numerous clinical trials demonstrating clinical efficacy and increase in survival rates and life expectancy [23], protection against cerebrovascular disease [24], long term drug safety, capacity to prevent organ damage, reduce morbidity and mortality in school-age children [25], toddlers [26] and infants [27]. The most notable evidence for the clinical efficacy of HU came from the Multicentre Study of Hydroxyurea (MSH) clinical trial [28]. The trial showed decreases in the frequency of painful episodes, acute chest syndrome, hospitalization and transfusions. Following FDA approval in 1998, 2007 saw HU receive approval from the European Medicines Agency (EMeA) as treatment for both adults and children with SCD. One year later, the National Institutes of Health Office of Medical Applications of Research (NIH-OMAR) and the Agency of Healthcare Research and Quality (AHRQ) both declared HU as an effective drug treatment for adults and children with SCD [29]. However, despite these results and the National Institutes of Health (NIH) recommendation for the use of HU in adults and children with SCD [30], HU is still underutilized [31]. This was confirmed by the NIH Consensus Development Conference statement [29] leading to studies investigating the barriers to the widespread use of HU. In surveys of the American Society of Paediatric Haematology and Oncology and Florida and North Carolina's Haematologists and

Oncologists, it was reported that the most common barrier to the prescription of HU for SCD is compliance from the patients and their families [32,33]. Similar barriers to the use of HU in children have also been reported [34] with the age of the children being one of the major barriers to prescription, despite data from the Hydroxyurea to Prevent Organ Damage in Children with SCD confirming the safety of HU in young children [35]. Similarly, HU is still not included within national guidelines for use in children below age 5 in some West African countries such as Kenya where the disease burden is highest [36,37], and recent report suggest that less than 5% of SCD patients in Cameroon ever used HU [6]. Furthermore, a recent review on the efficacy of HU in preventing SCD complications revealed that the majority of the studies were conducted in high-income countries with just 2 studies completed low-income countries, Tunisia and India [36]. This is illustrative of the apparent disproportion between regions of high disease burden and investment into medical research and intervention. Clinical trials investigating the effectiveness of HU in Africa are imperative as well as overcoming the barriers to the necessary utilization of HU in both children and adults with SCD and more importantly, further exploring curative modalities such as stem cell transplantation.

Haematopoietic stem cell transplantation. Currently, allogeneic HSC transplantation is the most successful curative treatment for SCD [38]. This method was first reported in 1984 with the HSCs derived from bone marrow, transplanted and successfully curing a patient with SCD and acute myelogenous leukemia [39,40]. To date, approximately 250 individuals have received HSC transplantation for SCD worldwide [41]. Transplantations can be autologous or allogeneic. Autologous HSC transplantation is used depending on the severity of the presenting symptoms as this approach averts immune rejection. After comprehensive screening, the patient receives immunosuppressive and myeloblastic treatment [41]. Although the myeloblastic treatment has been associated with secondary toxicity-related complications such as infertility and graft versus host disease (GVHD) [41], immunosuppressants such as cyclosporine and methotrexate have been used to prevent GVHD [42]. Using the Worldwide Network for Blood and Marrow Transplantation and the World Health Organization (WHO), the total number of first-time HSC transplantation between 2006 and 2008 were reported by WHO regional offices [43]. The findings were indicative of a general increase in the number of HSC transplants worldwide, with the highest in Asia. Of all 146 808 HSC transplants between 2006 and 2008, Europe and the Americas accounted for 50.7 and 28.9% respectively, whereas both Africa and the eastern Mediterranean region accounted for 2.7%, most of which were conducted in the United Arab Emirates, Qatar and Egypt [43]. Using a linear regression, they showed strong correlations between rates of transplantation and government healthcare expenditure, gross national income per capita and overall infrastructure in the country [43]. As encouraging as the general hike in transplant rates is, the above correlations bid developing countries, particularly in Africa, more challenges to meet the extraordinary needs for HSC transplants. In contrast, Egypt has since 1989 initiated a Stem Cell Transplant (SCT) program for all haematological disorders, which by 2007 had performed 1362 transplants, 80% of which were allogeneic [44]. The average 25 - 30% sibling HLA match is generally higher (about 40%) owing to the typical larger family size in most African communities. Egypt has 8 transplant centres performing about 210 transplants annually, the biggest being the Nasser Institute, which has completely shifted from bone marrow to peripheral blood as a stem cell source, a seemingly better option in **the developing world**. β -thalassaemia major, being the most common haemolytic anaemia in Egypt, saw compelling economic support for establishing a SCT program, which now has an overall and disease-free rates of 90 and 85%, respectively [44]. The SCT program in Egypt validates the feasibility of such programs in the third world

and could be the focal point of a regional collaboration to initiate and develop this practice to cure SCD in Sub-Saharan African settings, and its successful transplant rates advocate for the establishment of similar programs across Africa.

Advocacy for Haematopoietic stem cell transplantation centres

Every 2 years, Africa sees the birth of a number of SCD-affected children equal to the sum of the American and European populations affected by SCD, where most of these children die before age 5 [45]. One of the major obstacles to the management of SCD, particularly in developing countries, is the reluctance of governments and international healthcare agencies to accept that SCD is a worldwide health problem, comparable to that of communicable diseases and other major global disease such as diabetes and hypertension [46]. The Nasser Institute in Egypt serves as a model from which we advocate for the establishment of similar transplantation centres across the continent. Although the average cost of SCT in Egypt is 15 000 USD, this cost is completely sponsored by the Ministry of Health or medical insurance [44] due to a collaborative commitment between government and their private sector. We strongly urge such commitments to developing research, science and technology in order to build the necessary scientific and research capacity in Africa. We advocate for a joint effort from Cape and Cairo, to initiate a centre for HSC transplantation in Africa, with the respective governments, members of the private sector and pharmaceutical industry, leading researchers and clinicians as stakeholders. We propose an intra-Africa collaboration and pooling of resources towards developing a transplantation centre that will be mandated with developing cost-effective procedures for HSC transplants for haematological disorders like SCD and β -thalassaemia; large-scale clinical trials and follow-up studies, initiating epidemiological studies of haematological disorders in Africa, studies on health-related quality of life (HRQL) and patient survival rates, commercialization and distribution of treatments and therapies across the continent and most importantly, building technological and research capacity in Africa through teaching and training laboratory technicians and researchers. Such an initiative would also strengthen collaborative relationships across research and academic institutions in Africa, leading to a combined database of patients, information and experimental procedures, which will go towards standardizing practices within the field. The centre could initiate research towards developing disease models and investigating manifestation of diseases and accounting for previously undocumented African polymorphisms, develop an Afrocentric pharmacogenomics knowledge base including methods of drug design, predictive responder indexes and toxicity studies. This centre could also give rise to other initiatives such as umbilical cord blood and stem cell banks and continue building scientific and technological infrastructure in Africa. Furthermore, it will lead to improving blood transfusion services across the continent, incite further investigation of drug treatments like HU through clinical trials such as the NIH's Novel use of HU in an African Region with Malaria (NOHARM) and a better understanding of disease-modifying polymorphisms within various African populations.

Conclusion

The data available on clinical trials and reports conducted outside of Africa is illustrative that a range of treatments are available and have been successful at curing SCD. What is required is the implementation of strategies to affordably avail these treatments in Africa, particularly HU and ultimately, HSC transplantation, through

a collective effort from researchers, physicians, state departments of health and the private sector. This approach is likely to foster collaborative healthcare and research networks across the continent, ensuring the continued development of scientific and technological capacity in Africa to effectively manage the high disease burden of haematological disorders such as SCD.

Competing interests

The authors declare no competing interests.

Authors' contributions

Gift Pule drafted the manuscript and compiled the revisions. Ambrose Wonkam conceived the article and revised the draft, both authors approved the final version of the manuscript.

References

1. Bartolucci P and Galacteros F. Clinical management of adult sickle-cell disease. *Curr Opin Hematol.* 2012; 19(3):149-155. **PubMed | Google Scholar**
2. Flint J, Harding RM, Boyce AJ, Clegg JB. The population genetics of the haemoglobinopathies. *Baillieres Clin Haematol.* 1998; 11(1):1-51. **PubMed | Google Scholar**
3. Nagel RL, Fabry ME, Pagnier J, Zohoun I et al. Hematologically and genetically distinct forms of sickle cell anaemia in Africa - The Senegal type and the Benin type. *N Engl J Med.* 1985; 312(14):880-884. **PubMed | Google Scholar**
4. Piel FB, Patil AP, Howes RE, Nyangiri OA et al. Global epidemiology of sickle haemoglobin in neonates: a contemporary geostatistical model-based map and population estimates. *Lancet.* 2013; 381(9861):142-151. **PubMed | Google Scholar**
5. Rahimy MC, Gangbo A, Ahouignan G and Alihonou E. New born screening for sickle cell disease in the Republic of Benin. *J Clin Pathol.* 2009; 62(1):46-48. **PubMed | Google Scholar**
6. Wonkam A, De Vries J, Royal C et al. Would you terminate a pregnancy affected by Sickle Cell Disease? Analysis of views of patients in Cameroon. *J Med Ethics.* 2013; 0:1-6 doi: 10.1136/medethics-2013-101392. **PubMed | Google Scholar**
7. Ruffieux N, Njamnshi AK, Wonkam A, Hauert CA et al. Association between biological markers of sickle cell disease and cognitive functioning amongst Cameroonian children. *Child Psychol.* 2013; 19(2):143-160. **PubMed | Google Scholar**
8. Moskowitz JT, Butensky E, Harmatz P, Vichinsky E et al. Caregiving time Sickle Cell Disease: Psychological effects in maternal caregivers. *Pediatr Blood Cancer.* 2007; 48(1):64-71. **PubMed | Google Scholar**
9. Brandow AM, Brousseau DC and Panepinto A. Postdischarge pain, functional limitations and impact on caregivers of children with sickle cell disease treated for painful events. *Brit J Haematol.* 2009; 144(5):782-788. **PubMed | Google Scholar**
10. Makani J, Ofori-Acquah SF, Nnodu O, Wonkam A et al. Sickle Cell Disease: New Opportunities and Challenges in Africa. *Scientific World Journal.* 2013; Article ID 193252: 16 pages. **PubMed | Google Scholar**
11. Quinn CT, Rogers ZR, Buchanan GR. Survival of children with sickle cell disease. *Blood.* 2004; 103(11):4023-4027. **PubMed | Google Scholar**
12. Grosse SD, Odame I, Atrash HK, Amendah DD, et al. Sickle Cell Disease in Africa: A neglected cause of early childhood mortality. *Am J Prev Med.* 2011; 41(6):398-405. **PubMed | Google Scholar**
13. Ballas SK, Kesen MR, Goldberg MF, Luty GA et al. Beyond the definitions of the phenotypic complications of Sickle Cell Disease: an update on management. *Scientific World Journal.* 2012; 2012:949535. **PubMed | Google Scholar**
14. Ware RE and Helms RW. Stroke with transfusion changing to hydroxyurea (SWITCH): a phase 3 randomized clinical trial for treatment of children with sickle cell anemia, previous stroke, and iron overload. *Blood.* 2012; 119(17):3925-3932. **PubMed | Google Scholar**
15. Atz AM, Wessel DL. Inhaled nitric oxide in sickle cell disease with acute chest syndrome. *Anesthesiology.* 1997; 87(4):988-990. **PubMed | Google Scholar**
16. Alhashimi D, Fedorowicz Z, Alhashimi F, Dastgiri S. Blood transfusions for treating acute chest syndrome in people with sickle cell disease. *Cochrane Database Syst Rev.* 2010; (1):CD007843. **PubMed | Google Scholar**
17. Osaro E and Charles AT. The challenges of meeting the blood transfusion requirements in Sub-Saharan Africa: the need for the development of alternatives to allogenic blood. *J Blood Med.* 2011; 2:7-21. **PubMed | Google Scholar**
18. Bugge HF, Karlsen NC, Oydna E, Rake MM et al. A study of blood transfusion services at a district hospital in Malawi. *Vox Sang.* 2013; 104(1):37-45. **PubMed | Google Scholar**
19. Vichinsky EP. Current issues with blood transfusions in sickle cell disease. *Semin Hematol.* 2001; 38(1):14-22. **PubMed | Google Scholar**
20. Petz LD, Calhoun L, Shulman IA, Johnson C et al. The sickle cell hemolytic transfusion reaction syndrome. *Transfusion.* 1997; 37(4):382-92. **PubMed | Google Scholar**
21. Ballas SK. Iron overload is a determinant of morbidity and mortality in adult patients with sickle cell disease. *Semin Hematol.* 2001; 38(1):30-6. **PubMed | Google Scholar**
22. Platt OS, Orkin SH, Dover G, et al. Hydroxyurea enhances fetal hemoglobin production in sickle cell anemia. *J Clin Invest.* 1984; 74(2):652-6. **PubMed | Google Scholar**
23. Voskaridou E, Christoulas D, Bilalis A, et al. The effect of prolonged administration of Hydroxyurea on morbidity and mortality in adult patients with sickle cell syndromes: results of a 17-year, single-center trial (LaSHS). *Blood.* 2010; 115(12):2354-2363. **PubMed | Google Scholar**

24. Zimmerman SA, Schultz WH, Burgett S, Mortier NA et al. Hydroxyurea therapy lowers transcranialdoppler flow velocities in children with sickle cell Anaemia. *Blood*. 2007; 110(3):1043-1047. **PubMed | Google Scholar**
25. Kinney TR, Helms RW, O'Branski EE, et al. Safety of hydroxyurea in children with sickle cell anemia: results of the HUG-KIDS study, a phase I/II trial PediatricHydroxyurea Group. *Blood*. 1999; 94(5):1550-1504. **PubMed | Google Scholar**
26. Thornburg CD, Dixon N, Burgett S, et al. A pilot study of hydroxyurea to prevent chronic organ damage in young children with sickle cell anemia. *Pediatr Blood Cancer*. 2009; 52(5):609-615. **PubMed | Google Scholar**
27. Alvarez O, Miller ST, Wang WC. Effect of Hydroxyurea Treatment on Renal Function Parameters: Results from the Multi-Center Placebo-Controlled BABY HUG Clinical Trial for Infants with Sickle Cell Anemia. *Pediatr Blood Cancer*. 2012; 59(4):668-674. **PubMed | Google Scholar**
28. Charache S, Terrin ML, Moore RD, et al. Effect of hydroxyurea on the frequency of painful crises in sickle cell anemia. *N Engl J Med*. 1995; 332(20):1317-1322. **PubMed | Google Scholar**
29. Brawley OW, Cornelius LJ, Edwards LR, et al. National Institutes of Health Consensus Development Conference statement: hydroxyurea treatment for sickle cell disease. *Ann Intern Med*. 2008; 148(12):932-938. **PubMed | Google Scholar**
30. Bethesda MD. National Institutes of Health. The Management of Sickle Cell Disease. 2002; 161-165. **PubMed | Google Scholar**
31. Lanzkron S, Strouse JJ, Wilson R, et al. Systematic review: Hydroxyurea for the treatment of adults with sickle cell disease. *Ann Intern Med*. 2008; 148(12):939-955. **PubMed | Google Scholar**
32. Brandow AM, Jirovec DL, Panepinto JA. Hydroxyurea in children with sickle cell disease: Practice patterns and barriers to utilization. *Am J Hematol*. 2010; 85(8):611-613. **PubMed | Google Scholar**
33. Zumberg MS, Reddy S, Boyette RL, et al. Hydroxyurea therapy for sickle cell disease in community-based practices: A survey of Florida and North Carolina hematologists/oncologists. *Am J Hematol*. 2005; 79(2):107-113. **PubMed | Google Scholar**
34. Thornburg CD, Calatroni A, Telen M, et al. Adherence to hydroxyurea therapy in children with sickle cell anemia. *J Pediatr*. 2010; 156(3):415-419. **PubMed | Google Scholar**
35. Bethesda MD. National Heart, Lung, and Blood Institute (NHLBI). Hydroxyurea to prevent organ damage in children with sickle cell anemia. *Clin Trials*. 2010. **PubMed | Google Scholar**
36. Mulaku N, Opiyo N, Karumbi J, Kitonyi G. Evidence review of hydroxyurea for the prevention of sickle cell complications in low-income countries. *Arch Dis Child*. 2013; 98(11):908-914. **PubMed | Google Scholar**
37. Kenyan Ministry of Health. Clinical guidelines for management and referral of common conditions at levels 4-6. *Hospitals*. 2009; 3:259-261. **PubMed | Google Scholar**
38. Bernaudin F, Socie G, Kuentz M, et al. Long-term results of related myeloablative stem-cell transplantation to cure sickle cell disease. *Blood*. 2007; 110(7):2749-2756. **PubMed | Google Scholar**
39. Johnson FL, Look AT, Gockerman J, Ruggiero MR et al. Bone-marrow transplantation in a patient with sickle-cell Anaemia. *N Engl J Med*. 1984; 311(12):780-783. **PubMed | Google Scholar**
40. Milpied NHJ, Garand R, David A. Bone marrow transplantation for sickle cell anaemia (letter). *Lancet*. 1988; 2(8606):328-329. **PubMed | Google Scholar**
41. Oringanje C, Nemecek E, Oniyangi O. Hematopoietic stem cell transplantation for people with sickle cell disease. *Cochrane Database Syst Rev* 2013; 5:CD007001. **Google Scholar**
42. Artkins RC, Walter MC. Hematopoietic cell transplantation in the treatment of sickle cell disease. *Exp Opin Biol Therapy*. 2003; 3(1):1-10. **PubMed | Google Scholar**
43. Gratwohl A, Baldomero H, Gratwohl M, Aljurf M et al. Quantitative and qualitative differences in use and trends of hematopoietic stem cell transplantation: a Global Observational Study. *Haematologica*. 2013; 98(8):1282-1290. **PubMed | Google Scholar**
44. Mahmoud HK, El-Haddad AM, Fahmy OA, El-Emary M et al. Hematopoietic stem cell transplantation in Egypt. *Bone Marrow Transplant*. 2008; 42(1):76-80. **PubMed | Google Scholar**
45. Modell B, Darlison M. Global epidemiology of hemoglobin disorders and derived service indicators. *Bull World Health Organ*. 2008; 86(6):480-487. **PubMed | Google Scholar**
46. Wheatherall DJ. Hemoglobinopathies worldwide: present and future. *Curr Mol Med*. 2008; 8(7):592-599. **PubMed | Google Scholar**

Chapter 3.3 – Results

Southern African perspective of Sickle Cell Disease

Synopsis of Chapter 3.3

Sickle Cell Disease (SCD) was historically largely restricted to tropical equatorial Africa, concomitant with malaria endemicity and as a result, similarities in anthropology and/or cultural practices have been used as proxy for generalizing genomic data. This chapter includes two publications that;

- i) Argues for the dissociation of the assumed congruence between genetics and anthropology or culture and seeks to contribute to understanding the origin and distribution of the HbS mutation among southern African population.
- ii) Confirms the increasing burden of adult SCD patients at Groote Schuur Hospital, as previously reported in paediatric patients at Red Cross War Memorial Hospital, and seeks to caution national health and academic institutions to adapt policies, health care and professional training, accordingly.

3.3.1. Pule GD, Mnika K, Jourbert M, Mowla S, Novitzky N, Wonkam A. Increasing burden of adult patients living with Sickle Cell Disease in Cape Town. Submitted to *South African Medical Journal* (in final editing, #10849) 2016.

Abstract

Background: An exponential increase of the number of Sickle Cell Disease (SCD) patients in paediatric services in Cape Town has been reported. The trend in adult/adolescent services has not been investigated.

Objectives: The present study aims to evaluate the epidemiological trends and the profile of patients affected by SCD at the Haematology clinic at Groote Schuur Hospital.

Methods: 1) Retrospective review of the number SCD patients over the past 20 years; 2) cross-sectional analysis of clinical and haematological characteristics of SCD patients; 3) Molecular analysis of Hemoglobin S mutation, the haplotype in the β -globin-

like gene cluster, the 3.7kb α -thalassemia gene deletion and 19 selected single nucleotide polymorphisms (SNPs) associated with HbF levels.

Results: From 1995 to 2016, 81 SCD adolescent/adult patients were registered, mostly from others African countries (n = 61; 75.3%). There was over 200% increase in new cases (n = 47) in last quarter of the 2 decades investigated. Of the 58 regular attendees, data from 34 patients (58.6%) were analysed. The mean age was 26.1 ± 9.8 years; 70.6% were male. Except four patients with sickle/ β -thalassemia, all the patients had sickle cell anaemia (HbSS). The co-inheritance of a single 3.7kb α -globin gene deletion was found in 42.3% (n = 11) of cases. The Bantu haplotype was the most observed (65.4% of chromosomes). Most HbF-promoting SNPs were not associated with variable levels of haematological indices.

Conclusion: There is an increasing burden of adult SCD patients at Groote Schuur Hospital. National health and academic institutions need to adapt policies and health care professional training, accordingly.

Nature of publication: Original full article

Journal/Publisher: South African Medical Journal; ISSN: 2078-5135.

Candidate contribution: Conceived and designed the experiments, recruited and sampled the patients, performed the experiments, analysed the data, wrote the paper, revised and approved the manuscript.

Co-authors contribution: Conceived and designed the experiments: **GP**, AW. Recruited and sampled the patients: **GP**, KM, MJ. Performed the experiments: **GP**, KM. Analyzed the data: **GP**, KM, AW. Contributed reagents/materials/analysis tools: AW, MJ, SM, NN. Wrote the paper: **GP**, AW. Revised and approved the manuscript: **GP**, KM, MJ, SM, NN, AW.

***Genotype and Phenotype Profiles of Adult Sickle Cell Patients at Groot
Schoor Hospital***

Gift Dineo Pule¹, Khuthala Mnika¹, Myrtle Jourbert^{2,4}, Shaheen Mowla³, Nicolas Novitzky^{2,3,4}, Ambroise Wonkam¹

¹Division of Human Genetics, Department of Pathology, Faculty of Health Sciences, University of Cape Town, Cape Town, South Africa; ² Haematology Clinic, Groot Schoor Hospital, Cape Town, South Africa; ³ Division of Haematology, Departments of Internal Medicine and Pathology, Faculty of Health Sciences, University of Cape Town, South Africa.

⁴National Health Laboratory Service, Groot Schoor Hospital, Cape Town, South Africa.

Corresponding author: Professor Ambroise Wonkam, Division of Human Genetics, Department of Pathology, Faculty of Health Sciences, University of Cape Town, Anzio Road, Observatory, 7925, Cape Town, Republic of South Africa. Tel: +27214066307; Fax: +27214066826; e-mail: ambroise.wonkam@uct.ac.za.

Abstract

Background: An exponential increase of the number of Sickle Cell Disease (SCD) patients in paediatric services in Cape Town has been reported; the trend in adult/adolescent services has not been investigated.

Objectives: To evaluate the epidemiological trends and the profile of patients affected by SCD at the Haematology clinic at Groote Schuur Hospital.

Methods: 1) Retrospective review of the number of SCD patients over the past 20 years; 2) cross-sectional analysis of clinical and haematological characteristics of SCD patients; 3) Molecular analysis of Hemoglobin S mutation, the haplotype in the β -globin-like genes cluster, the 3.7kb α -thalassemia gene deletion and 19 selected single nucleotide polymorphisms (SNPs) associated with HbF levels.

Results: From 1995 to 2016, 81 SCD adolescent/adult patients were registered, mostly coming from other African countries ($n = 61$; 75.3%). There was over 200% increase in new cases ($n = 47$) in last quarter of the 2 decades investigated. Of the 58 regular attendees, data from 34 patients (58.6%) were analysed. The mean age was 26.1 ± 9.8 years; 70.6% were male. Except four patients with sickle/ β -thalassemia, all the patients had SCD (HbSS). The co-inheritance of a single 3.7kb α -globin deletion was found in 42.3% ($n = 11$) of cases. The Bantu haplotype was the most observed (65.4% of chromosomes). Most HbF-promoting SNPs were not associated with variable levels of haematological indices.

Conclusion: There is an increasing burden of adult SCD patients at Groote Schuur Hospital. National health and academic institutions need to adapt policies and health care professional training, accordingly.

Keywords: Sickle Cell Disease; Groote Schuur Hospital; South Africa; sub-Saharan Africa

Background

Sickle cell disease (SCD) is the first well documented molecular disease ^[1] and the most prevalent monogenic disease in the world. SCD is an accepted worldwide health problem that is comparable to other major global diseases such as diabetes and hypertension, including communicable diseases ^[2]. Sub-Saharan Africa (SSA) has the highest burden of disease with an excess of about 300 000 new affected births annually, which accounts for 80% of all annual global affected child births ^[3]. In spite of the high burden of disease in SSA, this is often associated with limited-to-poor medical resources, infrastructure and quality of care, thus estimates of neonatal and childhood mortality remain high, with up to 90% of affected children dying by five years of age ^[4]. SCD is caused by the polymerization and precipitation of the β -globin chains (HbS) during deoxygenation and dehydration of erythrocytes ^[5]. The altered structure of erythrocytes (normal biconcave shape to a crescent) is the basis of the vascular pathology of the disease, which includes abnormality of platelet and leukocyte adhesion and hyper-coagulation leading to microvascular occlusion, hemolysis, hypoxia, failed nitric oxide production and multi-organ damage ^[5]. The hallmark phenotypes of the disease include vaso-occlusive crises, stroke and acute chest syndrome, which are the basis for the typical episodes of acute pain ^[5-7]. The phenotype of SCD is influenced by both environmental and genetic factors. Variants at three principal loci; *BCL11A*, *HBS1L-MYB* intergenic polymorphism (HMIP) and the β -globin haplotype have been shown to account for 10-20% of the variance of Fetal Hemoglobin (HbF) levels and associated with the amelioration of the SCD symptoms ^[8-10]. Other variants in the *BCL11A* erythroid-specific enhancer (rs1427407 and rs7606173) have been shown to account for 8% and 6.2% in HbF variance, respectively, among SCD patient cohorts in the USA ^[11, 12], Tanzania ^[13] and Cameroon ^[14]. The co-inheritance of α -thalassemia has also been associated with improved clinical manifestation of SCD ^[15-18]. Even though the multiple independent origin of the HbS

mutation has been recently questioned ^[19], the SCD mutation is classically associated with five-region defined β -globin gene haplotypes; Benin, Bantu or Central African (CAR), Cameroon, Senegal and Indian-Arab ^[20-23], four of which are from Africa and associated with malaria incidence ^[24].

Because of the low incidence of malaria, the incidence of SCD in South Africa is equally extremely low; the HbS allele can be found in some indigenous South African ethnic groups (Venda and Shangaan) at an approximated frequency of 0.2% ^[25, 26]. However, this is changing with the socio-economically motivated influx of immigrants from other African countries, especially those within the equatorial malaria-endemic belt, resulting in study period by 300 - 400% increase of new cases of SCD over the past 10 years SCD at Red Cross War Memorial Children's Hospital (RCWMH) in Cape Town ^[27]. The existence of similar trends in adult SCD patient services has not been investigated. Following our previous report at RCWMH, we have reported in the present study the trend of new cases of adolescent and adult patients with SCD over the past 20 years, and studied the clinical, haematological and genetic profiles of a cohort of thirty-four of adolescent and adult SCD patients at the Haematology Unit at Groote Schuur Hospital (GSH), Cape Town.

Methods

Ethical approval

The study was performed in accordance with the Helsinki declaration and with the approval of the University of Cape Town, Faculty of Health Sciences Human Research Ethics Committee (HREC REF: 132/2010). Informed and written consent was obtained from adult participants (18 years and older) and for the one patient (15 years old), informed consent was obtained from the guardian with assent from the participant.

Patients

The haematology clinic runs weekly every Wednesday and most patients are seen at least once a month and every 3 months for the clinically stable patients, with the exception of crises related hospitalization. A retrospective review of the number of SCD patients attending the clinic over the past 20 years and a cross-sectional analysis of patients that regularly attend the clinic were performed. Clinical events and haematological indices were retrospectively collected from hospital records. The haematological measures were that reported at the first visit to the hospital.

Molecular methods

DNA extraction

DNA isolated using from the peripheral blood using the AllPrep DNA/RNA/miRNA universal kit (QIAGEN, USA) according to manufacturer's instructions.

Genotyping

HbS mutation and β -globin haplotypes

PCR and DdeI restriction analysis were used to confirm the presence of the HbS mutation using 100 ng DNA ^[28]. To determine the haplotype background, published primers and methods ^[29] genotyping five restriction fragment length polymorphic (RFLP) regions in the β -globin gene cluster were used to analyse the XmnI (5'G γ), HindIII (G γ), HindIII (A γ), HincII (3'Ψ β) and HinfI (5' β) for the β -globin haplotype background ^[19].

SNPs

Using a reported method ^[9], SNaPshot genotyping, capillary electrophoresis and direct cycle sequencing were used to assay five selected HbF-associated variants: rs8176703, rs372091, rs2334880, rs1427407, rs7606173. In addition, 14 other variants were analysed using the iPLEX Gold Sequenom Mass Genotyping Array (Inqaba Biotec, South Africa): X12_123681790, X16_391593, rs10468869, rs10756993, rs113267280, rs11754265,

rs141494605, rs148706947, rs183437571, rs192197462, rs570013781, rs59329875, rs62573842, rs6466533, rs6590706, rs67104793, rs7163278 and rs76901220.

Statistical analysis

Descriptive statistics were obtained for all quantitative data using SPSS (IBM, USA version 21.0). A chi-squared test, with 1 degree of freedom, was used to perform the Hardy-Weinberg Equilibrium (HWE) test on the SNPs genotypes with all variants in HWE ($p > 0.05$).

Results

Patients' origin and trends

A total of 128 patients' files from 1995 to March 2016 were reviewed. Among them, 47 patients were diagnosed with some form of α - or β -thalassemia (Figure 1). Of the remaining 81 patients affected by SCD, 61 (75.3%) were from other Sub-Saharan African countries. Over the last quarter (2011 – 2016) of the last two decades, there was an approximately 200% increase of new cases of SCD registered at the haematological clinic of GSH ($n = 47$) (Figure 2). Figure 3A and B shows the number of patients seen at GSH and country of origin respectively, with 16.4% ($n=21$) South African patients, most of which are of Mixed/Indian ancestry ($n=15$) and 21.9% ($n=28$) from Congo.

Of the 58 patients that regularly attend the Haematology clinic, 34 (58.6%) consented to be included in the study (Figure 1). Over 75% of the patients at GSH are referrals from Red Cross War Memorial Paediatric Hospital, some coming from neighbouring secondary level hospitals in Cape Town and a minority of internal referrals of relatives of attending patients.

Clinical and haematological profile

The mean age was 26.1 ± 9.8 years (range: 15 – 51 years), of which 70.6% were male. The co-inheritance of a single 3.7kb α -globin gene deletion was 42.3% ($n=11$).

Table 1 summarizes the haematological and clinical events recorded for all patients. Information obtained from the anamnesis indicated that the majority of the patients were diagnosed relatively late (mean age at diagnosis: 6.8 ± 7.1 years; range: 1-39 years), as a result of the presentation of the initial clinical manifestations of SCD, mainly pain episodes.

Clinical management

With regards to treatment, 33% of the patients in the cohort had received at least one blood transfusion; about 8% ($n=5$) of the patients have been enrolled in a hyper-transfusion program, ranging from fortnight to monthly transfusion regimen to manage complications such as stroke, chronic painful crises and non-healing chronic leg ulcer. The frequency of the transfusions is largely dependent on symptom severity and availability of blood units from the Western Cape Blood Transfusion Service (Cape Town, South Africa).

Of the patient cohort, 85% ($n=88$) were at maximum tolerated dose (MTD) with dosage ranging between 500 -1000 mg per day. From our patient survey, 30 – 50% of the participants were fully compliant with HU treatment; 20 – 30% reporting partial compliance (tended to forget to take the treatment 2 – 3 times a week) and some patients refusing the treatment. Reasons for refusal included potential cancer development in the future, family planning particularly for men afraid of treatment-related infertility and self-perceived improvement of symptoms without the treatment.

Genetic characteristics

Sickle cell genotypes; β -globin haplotypes and co-inheritance of α -thalassemia

The description of the HbS allele frequency, β -globin haplotype background and α -globin gene deletion for the patients is given in Table 1. Genotyping for the HbS mutation revealed that 85.2% ($N=23$) of the patients were homozygous for the mutation (HbSS) with the rest being heterozygous (HbAS) with a possibility of β^0 -thalassemia (HbS/ β^0) ($N=4$), all of which were South African with Mixed and Indian ancestry. Figure 4 shows the distribution of the

SCD β -globin gene haplotypes: the Bantu and Atypical haplotypes accounted for 65.4% and 26.9%, respectively, whereas the Senegal and Benin haplotypes each were 3.8% with no observation of the Cameroon and Indian-Arab haplotypes. In combination, the Bantu/Bantu haplotype represented 50% of the patients (Table 1). The heterozygous 3.7kb α -globin gene deletion was observed in 42.3% of the patients.

Frequency of genetic variants associated with HbF levels

Table 2 shows the observed alleles and minor allele frequency (MAF) of genetic variants previously associated with HbF, some recently in a Sardinian population ^[30]. All variants were in Hardy Weinberg Equilibrium ($p > 0.05$) with the exception of 4 loci (X12_123681790, rs141494605, rs183437571 and rs192197462) that presented monomorphic alleles in all patients. Tests of association between the variants and all haematological indices including Hb levels were conducted however, no significant associations were observed except for rs6466533 and the SCD haplotype combinations. The CC genotype (rs6466533) was associated with higher platelet counts compared to the heterozygous TC genotype ($p < 0.05$). Similarly, the Bantu/Bantu haplotype combination was associated with higher platelet counts compared to the Atypical/Atypical genotype ($p < 0.05$). Table 3 shows the MAF of the above variants in African populations; Esan (Nigeria), Luhya (Kenya), Mandinka (Gambia) and Mende (Sierra Leone); American (including African American), Europe and both East and South Asian populations.

Discussion

To the best of our knowledge, this is the first study describing the clinical and genetic backgrounds of SCD patients at Groote Schuur Hospital and reporting on adult patients with SCD in South Africa. The results of this study indicate the similar trend of a rapid increase in the number of cases of SCD that was previously reported in at Red Cross War Memorial Children's Hospital in Cape Town ^[27]. This was also the result of migration from sub-

Saharan Africa countries where SCD is most prevalent. Related to this was a specific administrative difficulty in taking care of some patients that lack up-to-date and correct paperwork for immigrants and asylum seekers. This was an indirect indication that most patients arrived as adults in South Africa, contrary to what was observed in the second part of the last decade at RCWMH, where most patients were South African born. It is therefore expected the adult population of SCD are GSH will continue to grow as a compounded effect of future referrals from neighbouring paediatric hospitals and arrival of new adult patients from migrant populations, as migration particularly into SSA continues to be a reality of many seeking political asylum, economic opportunities and better healthcare. Concomitant with the migration, the improved clinical management and healthcare of paediatric SCD patients is expected to grow the pool of adults patients living with SCD, and increased the number of patients that will survive well beyond the reproductive age, therefore likely increasing, and the frequency of the HbS allele in the population.

Indeed, newborn screening and comprehensive clinical care programs, which are also possible in South Africa, have reduced SCD-related premature childhood deaths by 70% in high-income nations like the USA ^[31, 32] and most patients can survive into adulthood ^[33]. A similar increasing trend of SCD in countries previously not affected by the disease has been observed in Ireland ^[34], Italy ^[35], Germany ^[36], England ^[37] or France ^[38]. Therefore the evidence that SCD burden is comparable to that of communicable diseases and other major global diseases such as hypertension and diabetes ^[2] will have increasing resonance. The marked increase in patients between 2001 and 2005 could also be associated with the ending of a civil war, transitional government and political instability in Congo, which had spread into neighbouring states. The latter increase (2011 – 2016) is more likely due to economic and health-motivated migration. The increase of SCD in both paediatric and adult settings will impose new burden on the healthcare system in South Africa concomitant with the new

need for training at all levels of medical education, as well as the need for policies from health authorities for the prevention, management and care of haemoglobinopathies.

The number of annual vaso-occlusive crises was similar to that reported in Cameroon and likewise the mean number of vaso-occlusive crises per year ^[9]. As specific challenge faced by the health care professional was the compliance with HU treatment, supportive medication such as folic acid and clinic attendance are the major challenges at Groote Schuur Hospital. There are several barriers to HU treatment; the financial implications of taking days off at work to attend the clinic and receive the treatment (maximum one month's supply), misconceptions about the treatment and possible carcinogenic effects as well as potential impotence in male patients ^[39]. Most patients fail to comply with clinic appointments simply because they feel better and thus see no need to go to the hospital.

The novel aspect of this paper is the report on the genetic background of four key modifiers of HbF; variants at the *BCL11A* erythroid-specific enhancer, β -globin haplotypes and α -thalassaemia 3.7kb gene deletion and several other known HbF-promoting polymorphisms. It is imperative to gain a better understanding of genetic variants affecting the predisposition to specific complications such as stroke, acute chest syndrome and polymorphisms affecting susceptibility to pain as well as the pharmacogenomics of the commonly prescribed treatments such as hydroxyurea, malaria prophylaxis and pain medication. The dominance of the Bantu haplotype in this cohort is in accordance with the Congolese origin of most patients. Indeed, the Bantu haplotype is most prevalent below the equatorial malaria belt across Southern African countries ^[19]. Most variants were not associated with haematological indices except the CC genotype at rs6466533 and Bantu/Bantu haplotype combination with platelet counts, likely due to the modest sample size. The differences in MAF of the recently identified HbF-promoting loci in a Sardinian population ^[30] among African populations from the Human 1000 Genome Project (1000G) (Table 3) emphasises the necessity of large-scale

genomic analyses on various populations across the continent as there is vast variations between any two African populations.

It should be underlined that the intention of this paper is not to stigmatize SCD nor immigrant patients, but to inform and prepare medical care providers and healthcare officials of the increasing need for management of haemoglobinopathies in South Africa. This trend is not restricted to South Africa, with countries such as Italy ^[40], the Republic of Ireland ^[34], England ^[41] and Germany ^[36] that have been affected by the reality of population movement and new burden of diseases, developing neonatal screening programmes and establishing SCD centres in response to similar increases in SCD prevalence.

Limitations

The final sample size of patients included in the present study was limited because of poor patients compliance to clinic attendance, self-transfers to other hospitals and the shortened period of recruitment. The sample size did not allow for robust statistical analyses to reveal potent markers of specific phenotypes or clinical measures. A handful of patients had HbF levels performed using HPLC (High performance liquid chromatography) before initiation of hydroxyurea treatment. This haematological measure would have been ideal to check for association with genetic variants as baseline HbF.

Conclusion

Over the past 10 years the number of adult patients living with SCD has increased considerably, imposing the creation of a weekly out-patients service at Groote Schuur Hospital. The genic profile is similar to that of many other SCD patients from other sub-Saharan Africa countries where most patients are originally from. The trend has implication for medical education through academic and training institutions and policy action via the national health ministries for the prevention and care as well as research in hemoglobinopathies in South Africa. This paper demonstrates the changing distribution of

the HbS allele and prevalence of the disease in Southern Africa and aims to enlighten all stakeholders: patients, care providers and health ministries towards better management of SCD.

References

1. **Herrick JB.** Peculiar elongated and sickle-shaped red blood corpuscles in a case of severe anemia. 1910. *Yale J.Biol.Med.* 2001; **74**:179-184.
2. **Weatherall DJ.** Hemoglobinopathies worldwide: present and future. *Curr.Mol.Med.* 2008; **8**:592-599.
3. **Piel FB, Patil,AP, Howes,RE, et al.** Global epidemiology of sickle haemoglobin in neonates: a contemporary geostatistical model-based map and population estimates. *The Lancet.* 2013; **381**:142-151.
4. **Makani J, Menzel,S, Nkya,S, et al.** Genetics of fetal hemoglobin in Tanzanian and British patients with sickle cell anemia. *Blood.* 2011; **117**:1390-1392.
5. **Bartolucci P and Galacteros,F.** Clinical management of adult sickle-cell disease. *Curr.Opin.Hematol.* 2012; **19**:149-155.
6. **Weatherall DJ.** The inherited diseases of hemoglobin are an emerging global health burden. *Blood.* 2010; **115**:4331-4336.
7. **Malowany JI and Butany,J.** Pathology of sickle cell disease. *Conference Proceedings.* 2012; **29**:49-55.
8. **Thein SL, Menzel,S, Peng,X, et al.** Intergenic variants of HBS1L-MYB are responsible for a major quantitative trait locus on chromosome 6q23 influencing fetal hemoglobin levels in adults. *Proc.Natl.Acad.Sci.U.S.A.* 2007; **104**:11346-11351.
9. **Wonkam A, Bitoungui,VJN, Vorster,AA, et al.** Association of variants at BCL11A and HBS1L-MYB with hemoglobin F and hospitalization rates among sickle cell patients in Cameroon. *PLoS One.* 2014; **9**:e92506.
10. **Lettre G, Sankaran,VG, Bezerra,MA, et al.** DNA polymorphisms at the BCL11A, HBS1L-MYB, and beta-globin loci associate with fetal hemoglobin levels and pain crises in sickle cell disease. *Proc.Natl.Acad.Sci.U.S.A.* 2008; **105**:11869-11874.
11. **Bauer DE, Kamran,SC, Lessard,S, et al.** An erythroid enhancer of BCL11A subject to genetic variation determines fetal hemoglobin level. *Science.* 2013; **342**:253-257.
12. **Galarneau G, Palmer,CD, Sankaran,VG, et al.** Fine-mapping at three loci known to affect fetal hemoglobin levels explains additional genetic variation. *Nat.Genet.* 2010; **42**:1049-1051.
13. **Mtatiro SN, Mgaya,J, Singh,T, et al.** Genetic association of fetal-hemoglobin levels in individuals with sickle cell disease in Tanzania maps to conserved regulatory elements within the MYB core enhancer. *BMC Med Genet.* 2015; **16**:1.
14. **Pule GD, Ngo Bitoungui,VJ, Chetcha Chemegni,B, et al.** 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*. 2015; **19**:627-631.

15. **Hsu LL, Miller,ST, Wright,E, et al.** Alpha Thalassemia is associated with decreased risk of abnormal transcranial Doppler ultrasonography in children with sickle cell anemia. *J.Pediatr.Hematol.Oncol.* 2003; **25**:622-628.
16. **Rumaney MB, Bitoungui,VJN, Vorster,AA, et al.** The Co-Inheritance of Alpha-Thalassemia and Sickle Cell Anemia Is Associated with Better Hematological Indices and Lower Consultations Rate in Cameroonian Patients and Could Improve Their Survival. *PLoS One.* 2014; **9**:e100516.
17. **Saleh-Gohari N and Mohammadi-Anaie,M.** Co-inheritance of sickle cell trait and thalassemia mutations in South central iran. *Iran.J.Public.Health.* 2012; **41**:81-86.
18. **Wambua S, Mwacharo,J, Uyoga,S, et al.** Co-inheritance of α -thalassaemia and sickle trait results in specific effects on haematological parameters. *Br.J.Haematol.* 2006; **133**:206-209.
19. **Bitoungui VJN, Pule,GD, Hanchard,N, et al.** 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*. 2015; **19**:171-179.
20. **Flint J, Harding,RM, Boyce,AJ, et al.** 1 The population genetics of the haemoglobinopathies. *Bailliere Clin Hematol.* 1998; **11**:1-51.
21. **Elion J, Berg,PE, Lapoumeroulie,C, et al.** DNA sequence variation in a negative control region 5' to the beta-globin gene correlates with the phenotypic expression of the beta s mutation. *Blood.* 1992; **79**:787-792.
22. **Pagnier J, Mears,JG, Dunda-Belkhodja,O, et al.** Evidence for the multicentric origin of the sickle cell hemoglobin gene in Africa. *Proc.Natl.Acad.Sci.U.S.A.* 1984; **81**:1771-1773.
23. **Labie D, Pagnier,J, Lapoumeroulie,C, et al.** Common haplotype dependency of high G gamma-globin gene expression and high Hb F levels in beta-thalassemia and sickle cell anemia patients. *Proc.Natl.Acad.Sci.U.S.A.* 1985; **82**:2111-2114.
24. **Williams TN, Mwangi,TW, Wambua,S, et al.** Sickle cell trait and the risk of Plasmodium falciparum malaria and other childhood diseases. *J.Infect.Dis.* 2005; **192**:178-186.
25. **Bonafede RP, Botha,MC and Beighton,P.** Inherited anaemias in the Greek community of Cape Town. *J.Med.Genet.* 1979; **16**:197-200.
26. **Beighton P and Botha,M.** Inherited disorders in the black population of southern Africa. *S.Afr.Med.J.* 1986; **69**:247-249.
27. **Wonkam A, Ponde,C, Nicholson,N, et al.** The burden of sickle cell disease in Cape Town. *S.Afr.Med.J.* 2012; **102**:753-754.

28. **Saiki RK, Scharf,S, Faloona,F, et al.** Enzymatic amplification of beta-globin genomic sequences and restriction site analysis for diagnosis of sickle cell anemia. *Science*. 1985; **230**:1350-1354.
29. **Steinberg M, Lu,Z, Nagel,R, et al.** Hematological effects of atypical and cameroon β -globin gene haplotypes in adult sickle cell anemia. *Am.J.Hematol.* 1998; **59**:121-126.
30. **Danjou F, Zoledziewska,M, Sidore,C, et al.** Genome-wide association analyses based on whole-genome sequencing in Sardinia provide insights into regulation of hemoglobin levels. *Nat.Genet.* 2015; **47**:1264-1271.
31. **Yanni E, Grosse,SD, Yang,Q, et al.** Trends in pediatric sickle cell disease-related mortality in the United States, 1983-2002. *J.Pediatr.* 2009; **154**:541-545.
32. **Vichinsky EP.** Comprehensive care in sickle cell disease: its impact on morbidity and mortality. *Semin.Hematol.* 1991; **28**:220-226.
33. **Chaturvedi S and DeBaun,MR.** Evolution of sickle cell disease from a life-threatening disease of children to a chronic disease of adults: The last 40 years. *Am.J.Hematol.* 2016; **91**:5-14.
34. **Gibbons C, Geoghegan,R, Conroy,H, et al.** Sickle cell disease: time for a targeted neonatal screening programme. *Ir.Med.J.* 2015;
35. **Colombatti R, Perrotta,S, Samperi,P, et al.** Organizing national responses for rare blood disorders: the Italian experience with sickle cell disease in childhood. *Orphanet J. Rare Diseases.* 2013; **8**:1.
36. **Zur B.** Increase in genetically determined anemia as a result of migration in Germany. *Internist.* 2016; 1-7.
37. **Pizzo E, Laverty,AA, Phekoo,KJ, et al.** A retrospective analysis of the cost of hospitalizations for sickle cell disease with crisis in England, 2010/11. *J.Public.Health.(Oxf).* 2015; **37**:529-539.
38. **Habibi A, Arlet,JB, Stankovic,K, et al.** French guidelines for the management of adult sickle cell disease: 2015 update. *Rev.Med.Interne.* 2015; **36**:5S3-84.
39. **Pule GD, Mowla,S, Novitzky,N, et al.** A systematic review of known mechanisms of hydroxyurea-induced fetal hemoglobin for treatment of sickle cell disease. *OMICS.* 2015; **8**:669-679.
40. **Colombatti R, Montanaro,M, Guasti,F, et al.** Comprehensive care for sickle cell disease immigrant patients: a reproducible model achieving high adherence to minimum standards of care. *Paed. Blood Cancer.* 2012; **59**:1275-1279.
41. **Daniel YA and Henthorn,J.** Newborn screening for sickling and other haemoglobin disorders using tandem mass spectrometry: A pilot study of methodology in laboratories in England. *J.Med.Screen.* 2016; pii: 0969141316631008.

List of tables and figures

Table 1. Description of hematological indices, clinical events and β -globin markers of the study cohort.

Table 2. Genetic variants previously associated with HbF.

Table 3. Minor allele frequencies of select HbF-promoting SNPs in various populations in 1000G project.

Figure 1. Patients' recruitment flow chart

Figure 2. Incidence trend of Sickle Cell Disease at Groote Schuur Hospital over 20 years.

Figure 3. Number of patients by nationality (A) and the distribution of countries of birth for patients at Groote Schuur Hospital (B).

Figure 4. Distribution of SCD haplotypes by number and percentage of chromosomes.

Acknowledgments

We would like to thank the nursing staff of the Haematology Service at Groote Schuur Hospital for their support and the patients for their participation.

Contribution to authorship

Conceived and designed the experiments: GP, AW. Recruited and sampled the patients: GP, KM, MJ. Performed the experiments: GP, KM. Analyzed the data: GP, KM, AW. Contributed reagents/materials/analysis tools: AW, MJ, SM, NN. Wrote the paper: GP, AW. Revised and approved the manuscript: GP, KM, MJ, SM, NN, AW.

Funding statement

The molecular experiments of the study were funded by the National Health Laboratory Services (NHLS), South Africa; and the NIH, USA, grant number 1U01HG007459-01. The students' bursaries were supported by the Oppenheimer Memorial Trust, the National Research Foundation, and FirstRand Laurie Dippenaar Scholarship, South Africa. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Table 1. Description of hematological indices, clinical events and β -globin gene of the study cohort

<i>Variables</i>	<i>Categories</i>	<i>Mean +/- SD</i>	<i>Value range</i>	<i>Number of observations</i>
<i>Gender</i>	Female	29.4%		10
	Male	70.6%		24
<i>Age (years)</i>		26.1 \pm 9.8	15 - 51	34
<i>Hematological indices</i>	Hb (g/dl)	9.1 \pm 1.8	5.6 – 14.0	27
	MCV (fl)	91.5 \pm 31.9	53.6 – 146.4	27
	WBC ($10^9/l$)	8.1 \pm 3.0	3.0 – 13.6	27
	Platelets ($10^9/l$)	334 \pm 166.5	133 – 737	27
	Neutrophils ($10^9/l$)	4.6 +/- .,1	2.0 – 10.3	27
<i>Clinical events</i>	Age of diagnosis (years)	6.8 \pm 7.1	1 – 39	30
	Vaso-occlusive crisis (No. /year)	1.5 \pm 1.2	0 – 5	30
	Stroke	13.3%		4/30
	Leg ulcers	16.7%		5/30
	Hospitalizations (No. /year)	2.1 \pm 1.4	0 - 5	30
<i>Treatment</i>	Blood transfusions	33.3%		10/30
	Hydroxyurea (mg/day)	544.1 \pm 144	500 – 1000	30
<i>β-globin genotype</i>	HbSS	85.2%		23/27
	HbAS ^{&}	14.8%		4/27
<i>β-globin haplotype</i>	Bantu/Bantu	50.0%		13/26
	Bantu/Senegal	7.7%		2/26
	Bantu/Benin	7.7%		2/26
	Bantu/Atypical	15.4%		4/26
	Atypical	19.2%		5/26
<i>α-globin gene deletion</i>	$\alpha\alpha/\alpha\alpha$	57.7%		15/26
	$\alpha\alpha/\alpha3.7$	42.3%		11/26

[&]: HbAS with possibility of β^0 (HbS β^0); β -globin haplotypes numbers given by combination

Table 2. Genetic variants previous associated with HbF

<i>SNP</i>	<i>Chromosome loci</i>	<i>Alleles</i>	<i>MAF</i>	<i>Number of observations</i>
<i>rs8176703</i>	136135863 ^{&}	G/A	0,130	27
<i>rs372091</i>	11:5496926	G/A	0,148	27
<i>rs2334880</i>	16:71619734	C/T	0,400	25
<i>rs1427407</i>	2:60490908	C/A	0,204	27
<i>rs7606173</i>	2:60498316	C/G	0,426	27
<i>X12_123681790</i>	X12_123681790	A*	-	31
<i>X16_391593</i>	X16_391593	T/C	0,015	33
<i>rs10468869</i>	18:51795403	A/G	0,400	30
<i>rs10756993</i>	9:18839726	C/A	0,016	31
<i>rs113267280</i>	6:41952511	T/G	0,016	31
<i>rs11754265</i>	6:135356216	C/G	0,172	29
<i>rs141494605</i>	16:216593	T*	-	27
<i>rs148706947</i>	16:342218	C/T	0,047	32
<i>rs183437571</i>	19:13121899	C*	-	30
<i>rs192197462</i>	14:57116065	A*	-	30
<i>rs570013781</i>	16:149539	G/A	0,016	31
<i>rs59329875</i>	20:44547672	C/T	0,296	27
<i>rs62573842</i>	9:32264314	A/G	0,250	30
<i>rs6466533</i>	7:79163576	T/C	0,283	30
<i>rs6590706</i>	11:133465011	A/G	0,217	30
<i>rs67104793</i>	3:142444839	A/DEL.A	0,190	29
<i>rs7163278</i>	15:93345162	T/C	0,121	29
<i>rs76901220</i>	ss131769967 ^{&}	G/A	0,016	31

* monomorphic; [&] no chromosome location provided in dbSNP, reference ID provided.

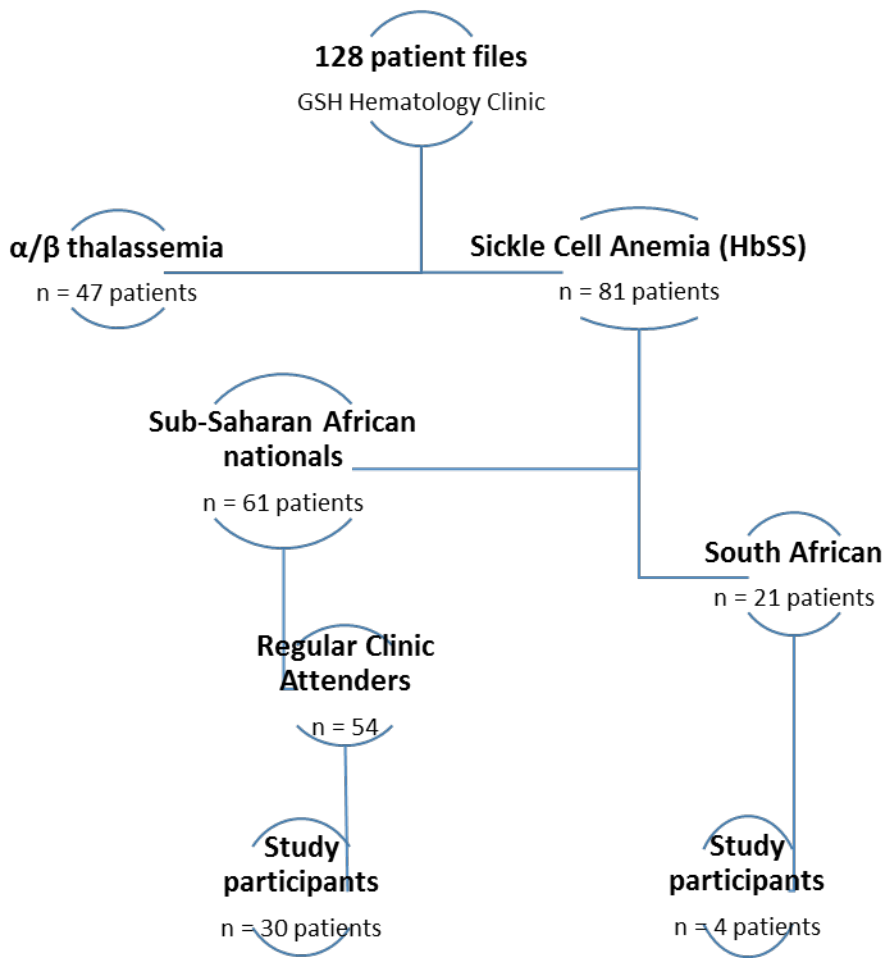


Figure 1. Patient recruitment flow chart

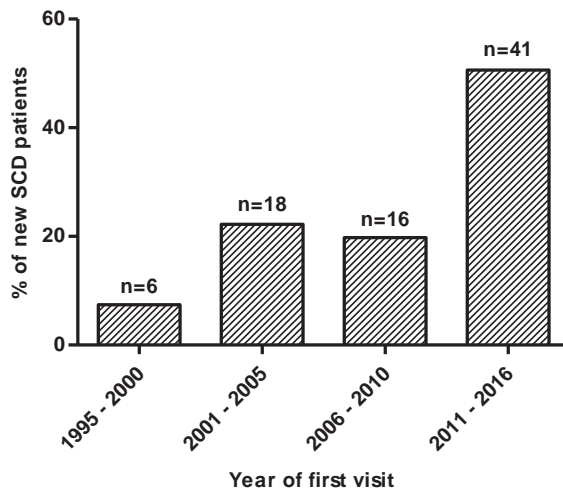


Figure 2. Frequency of Sickle Cell Disease at Groote Schuur Hospital.

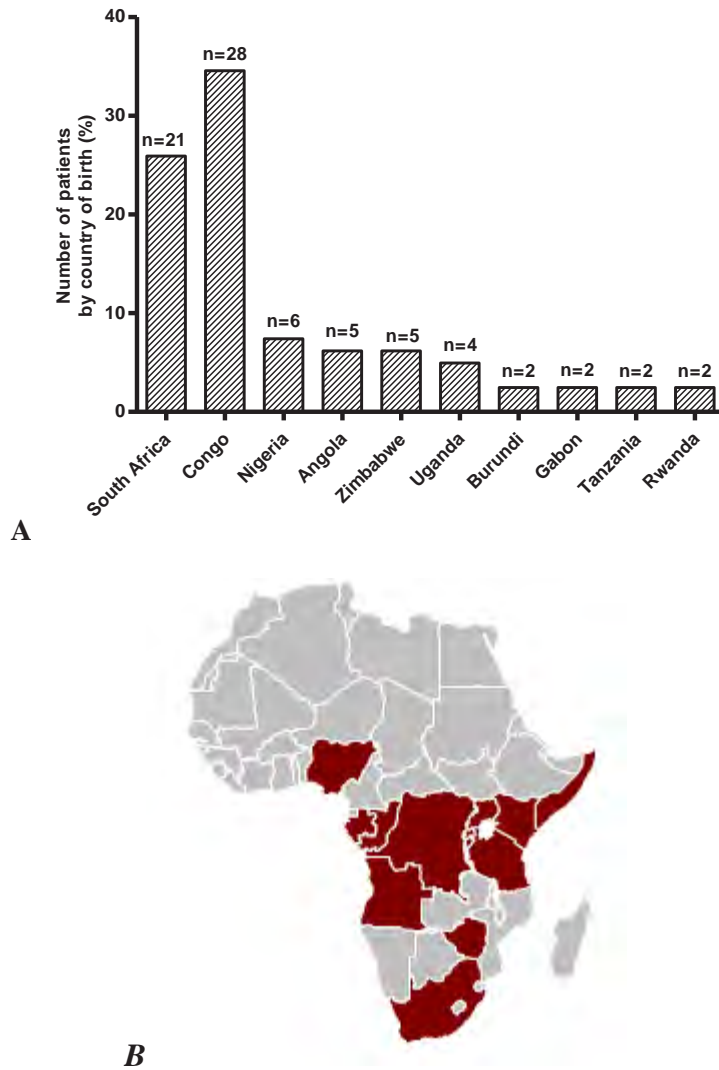


Figure 3. Number of patients by nationality (A) and the distribution of countries of birth for patients at Groote Schuur Hospital (B). There was one patient from each of the following countries; Zaire, Somalia, Hong Kong and the UK.

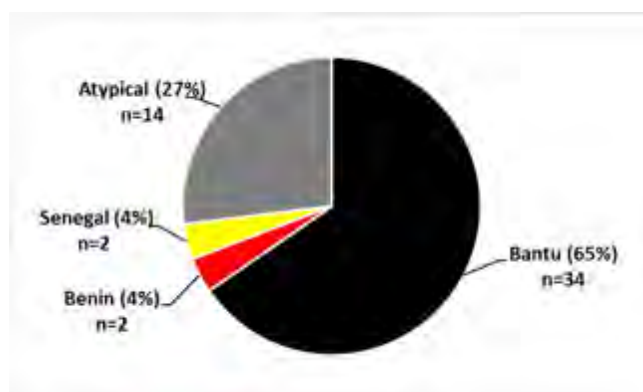


Figure 4. Distribution of SCD haplotypes. Haplotypes given by number and percentage of chromosomes.

Table 3. Minor allele frequencies of select HbF-promoting SNPs in various populations in 1000G project. &: monomorphic

SNPs	Our data	African	Esan (Nigeria)	Luhya (Webuye, Kenya)	Mandinka (The Gambia)	Mende (Sierra Leone)	America	African Caribbean (Barbados)	African American (Southwest US)	Europe	East Asia	South Asia
rs8176703	A=0,13	T= 0.041	T=0.0354	T=0.0202	T= 0.053	T=0.0353	T=0.006	T=0.0573	T=0.0164	T= 0.002	T=0.0000	T=0.0000
rs372091	A=0,148	A= 0.068	A=0.1414	A=0.0707	A= 0.013	A=0.0235	A= 0.001	A=0.0469	A=0.0410	A= 0.000	A= 0.000	A= 0.000
rs2334880	T=0,4	A= 0.398	A=0.3990	A=0.4697	A= 0.403	A=0.4000	A=0.102	A=0.3594	A=0.3115	A= 0.160	A=0.009	A= 0.019
rs1427407	A=0,204	T= 0.238	T=0.1970	T=0.2020	T= 0.301	T=0.3353	T= 0.226	T=0.2135	T=0.2295	T= 0.151	T= 0.256	T= 0.120
rs7606173	G=0,426	C= 0.467	G=0.4697	G=0.4545	C=0.403	C=0.4176	C= 0.290	C=0.4115	G=0.4754	C= 0.435	C= 0.013	C= 0.173
rs10468869	G=0,4	G= 0.341	G=0.3283	G=0.3737	G= 0.274	G=0.3529	A= 0.324	G=0.3542	G=0.4180	A= 0.251	G= 0.377	A= 0.369
rs10756993	A=0,016	A=0.0000	A=0.0000	A=0.0000	A=0.0000	A=0.0000	A=0.000	A=0.0000	A=0.0000	A=0.000	A= 0.003	A=0.000
rs113267280	G=0,016	G=0.0000	G=0.0000	G=0.0000	G=0.0000	G=0.0000	G=0.0000	G=0.0000	G=0.0000	G= 0.008	G=0.0000	G= 0.004
rs11754265	G=0,172	G: 0.234	G=0.3081	G=0.1717	G= 0.195	G=0.1882	G= 0.448	G=0.2552	G=0.2951	C= 0.493	G=0.466	G=0.468
rs141494605	&	C=0.0000	C=0.0000	C=0.0000	C=0.0000	C=0.0000	C= 0.012	C=0.0000	C=0.0000	C= 0.012	C=0.0000	C=0.002
rs148706947	T=0,047	T=0.0000	T=0.0000	T=0.0000	T=0.0000	T=0.0000	T= 0.001	T=0.0000	T=0.0000	T= 0.013	T=0.0000	T=0.003
rs183437571	&	T= 0.008	T=0.0000	T=0.0051	T: 0.027	T=0.0118	T= 0.006	T=0.0000	T=0.0082	T= 0.001	T=0.0000	T=0.0000
rs570013781	A=0,016	A=0.0000	A=0.0000	A=0.0000	A=0.0000	A=0.0000	A= 0.009	A=0.0000	A=0.0000	A= 0.003	A=0.0000	A=0.001
rs59329875	T=0,296	C: 0.191	C=0.2222	C=0.2374	C: 0.133	C=0.1412	C= 0.114	C=0.2292	C=0.1639	C=0.198	C= 0.026	C=0.252
rs62573842	G=0,25	G=0.160	G=0.1313	G=0.1970	G= 0.150	G=0.0941	G= 0.329	G=0.1927	G=0.2295	G= 0.437	A=0.459	A= 0.473
rs6466533	C=0,283	C= 0.188	C=0.2071	C=0.1717	C: 0.146	C=0.1235	T= 0.419	C=0.2552	C=0.3279	T= 0.171	C= 0.486	T= 0.497
rs6590706	G=0,217	G= 0.222	G=0.1667	G=0.1717	G= 0.221	G=0.1588	A= 0.091	G=0.2865	G=0.3934	A= 0.026	A= 0.154	A= 0.143
rs67104793	DEL.A=0,19	T=0.227	T=0.1364	T=0.2828	T=-0.279	T=0.2529	T=0.218	T=0.2031	T=0.2295	T= 0.197	T=0.319	T=0.231
rs7163278	C=0,121	C= 0.137	C=0.1111	C=0.0909	C= 0.168	C=0.1353	C= 0.311	C=0.1719	C=0.2459	T= 0.383	C= 0.300	C=0.221

3.3.2. Pule GD, Mnika K, Chimusa E, Mhandire K, Kampira E, Dandara C, Wonkam A. Beta-globin gene haplotypes and tag variants for malaria resistance among black southern African populations. Submitted to *Global Health, Epidemiology and Genomics (revised manuscript submitted)* 2016.

Abstract

The partial carrier-resistance to *Plasmodium falciparum* malaria conferred by the sickle cell (HbS) mutation has resulted in the local amplification and positive selection of Sickle Cell Disease in malaria-endemic regions and particularly in sub-Saharan Africa (SSA). Recently a genome-wide association study revealed three linkage disequilibrium (LD) tags of known functional mutation associated with resistance to malaria; rs8176703 (9q34.2; *ABO*), rs2334880 (16q22.2; *MARVELD3*) and rs372091 (11p15.5; *HBB*). The present study investigated the β -globin gene haplotypes, and malaria-associated variants among three cohorts of Bantu-speaking individuals from Malawi, Zimbabwe and South Africa compared to reports from other SSA populations. The data confirm a south-ward frequency decrease of malaria-associated variants in SSA linked to the evolutionary dynamics of various African populations genomes through selective pressure of malaria. These selected genomics differences, among “Bantus” from various part of Africa emphasize the evidence of the dissociation between genetics, anthropology and culture, that could have implications in forensic sciences and precision medicine. The present study also shown a prevalent Benin haplotype, which is mostly found in West Africa, among Southern African Blacks and very low Bantu haplotype; and this results could contribute in investigating the origin and age the HbS mutation as well as migration of Southern African populations within Africa.

Nature of publication: Original full article

Journal/Publisher: Global Health, Epidemiology and Genomics; ISSN: 2054-4200.

Candidate contribution: Conceived and designed the experiments, performed the experiments, analysed the data, wrote the paper, revised and approved the manuscript.

Co-authors contribution: Conceived and designed the experiments: **GP**, AW. Performed the experiments: KM, **GP**. Performed bioinformatics analysis: EC. Analyzed the data: **GP**, AW. Contributed reagents/materials/analysis tools: CD, AW. Wrote the paper: **GP**, AW. Revised and approved the manuscript: **GP**, EC, KM, CD, AW.

Beta-Globin Gene Haplotypes And Tag Variants for Malaria-Resistance Among Black Southern African Populations

Gift Dineo Pule¹, Khuthala Mnika¹, Emile R. Chimusa¹, Kudakwashe Mhandire², Elizabeth Kampira³, Collet Dandara¹ and Ambroise Wonkam¹

¹Division of Human Genetics, Department of Pathology, Faculty of Health Sciences, University of Cape Town, Cape Town, South Africa

²Departments of Chemical Pathology, University of Zimbabwe, Harare, Zimbabwe

³Malawi College of Health Sciences, University of Malawi, Blantyre, Malawi.

Corresponding author: Professor Ambroise Wonkam, Division of Human Genetics, Faculty of Health Sciences, University of Cape Town, Anzio Road, Observatory, 7925, Cape Town, Republic of South Africa. Tel: +27214066307; Fax: +27214066826; e-mail: ambroise.wonkam@uct.ac.za

Short running title: β -globin haplotypes and malaria resistance variants among Southern African populations

Keywords: β -globin haplotypes; malaria resistance; Southern Africa; Sickle Cell Disease

Word count: 3500

Abstract

The partial carrier-resistance to *Plasmodium falciparum* malaria conferred by the sickle cell (HbS) mutation has resulted in the local amplification and positive selection of Sickle Cell Disease in malaria-endemic regions and particularly in sub-Saharan Africa (SSA). Recently, genome-wide association study revealed three linkage disequilibrium (LD) tags of known functional mutation associated with resistance to malaria; rs8176703 (9q34.2; ABO), rs2334880 (16q22.2; MARVELD3) and rs372091 (11p15.5; HBB). The present study investigated the β -globin gene haplotypes, and malaria-associated variants among three cohorts of Bantu-speaking individuals from Malawi, Zimbabwe and South Africa compared to reports from other SSA populations. The data illustrate a south-ward frequency decrease of malaria-associated variants in SSA linked to the evolutionary dynamics of various African populations genomes through selective pressure of malaria. These selected genomics differences, among “Bantus” from various part of Africa emphasize the evidence of the dissociation between genetics, anthropology and culture, that may have implications in precision medicine. The present study also shown a prevalent Benin haplotype, which is mostly found in West Africa, among Southern African Blacks and very low Bantu haplotype; and this results illustrate the value of investigating the origin and age of the HbS mutation in Southern African populations within Africa.

Introduction

Sickle Cell Disease (SCD) is a monogenic, hematological and multi-organ disorder affecting the structure of erythrocytes by altering the normal biconcave shape to a crescent [1]. The sickling results from the polymerization and precipitation of the β -globin chains (HbS) during deoxygenation and dehydration of erythrocytes [2]. The vascular pathology of the disease includes platelet and leukocyte adhesion abnormality and hypercoagulation leading to microvascular occlusion, hemolysis and hypoxia and ultimately, multi-organ damage.

There is strong correlation between the frequency of the HbS gene and the historical distribution and incidences of malaria [3] because of the partial carrier-resistance to *Plasmodium falciparum* malaria. The geographical co-occurrence of SCD and malaria and the partial carrier-resistance is believed to have resulted in the local amplification and positive selection of SCD in malaria-endemic regions [4,5]. Besides HbS mutation, three tag SNPs were identified in a genome-wide association study (GWAS) associated with resistance to severe malaria; rs8176703 (9q34.2; *ABO*), rs2334880 (16q22.2; *MARVELD3*) and rs372091 (11p15.5; *HBB*) [6]. These variants were found to have genome-wide significance and be in linkage-disequilibrium (LD) with functional mutations associated with resistance to severe malaria in patients and controls from Ghana. The *ABO* locus has previous indication of a protective effect conferred by the blood group O against severe malaria [7-9]. The variant rs2334880, was one of the novel resistance loci identified and was mapped to 6.4kb upstream of the MARVEL domain-containing protein 3 gene (*MARVELD3*; MIM ID*614094), which forms part of multiple tight-junction of epithelial and vascular endothelial cells [10-12] and is strongly associated with severe malaria [13]. The endemicity of malaria in sub-Saharan Africa (SSA) and the associated HbS mutations, has resulted in the highest SCD burden with nearly 80% of the approximately 300 000 new affected births that occur in SSA annually [14].

The HbS mutation is believed to have evolved independently in 5 regions of the world, classically associated with five region-defined haplotypes, four of which are African, based on conserved patterns of polymorphisms across the β -globin gene cluster, namely Benin, Central African (CAR) or Bantu; Cameroon; Senegal and Indian-Arab [4,15,16]. A recent review of the global distribution and frequencies of these haplotypes has provided a glimpse into population dynamics and migration within and out of Africa that has prompted the hypothesis of a single origin of HbS mutation [17]. In this context, the study of malaria associated variants among Southern African populations, specifically among South African Blacks that have been living outside the malaria-endemic equatorial belt for 3-5000 years [18,19], provides new insight into the within-Africa migration patterns, the origin of the HbS mutation, and some perspectives into the dissociation between genetics and anthropology, with regards to differential allele frequencies related to various conditions such as malaria, susceptibility and resistance among Bantu-speaking groups from various parts of Africa.

In this present study, we investigated the β -globin gene haplotypes and malaria-associated tag variants among three cohorts of healthy Southern African populations from Malawi, Zimbabwe and South Africa and compared the frequencies of these variants to that of other SSA populations.

Methods

Ethics Approval

The study was performed with the approval of the University of Cape Town, Faculty of Health Sciences Human Research Ethics Committee (HREC REF: 132/2010 and HREC REF: 1094/2009).

Populations

One hundred and fifty-eight DNA samples (50 Zimbabweans; 58 Malawians and 50 South Africans) all of Bantu origin were randomly selected for the Division of Human Genetics biorepositories, Faculty of Health Sciences, University of Cape Town. These participants were randomly sampled from a cohort of unrelated and apparently healthy individual, initially recruited for a population genetics study.

Genotyping

HbS mutation and β -globin haplotypes

Using the participants from the three southern African populations, PCR and *Dde I* restriction analysis were used to confirm the absence of the HbS mutation [20] and published primers and methods [21] genotyping five restriction fragment length polymorphic (RFLP) regions in the β -globin gene cluster were used to analyse the *XmnI* (5'G γ), *HindIII* (G γ), *HindIII* (A γ), *HincII* (3'Ψ β) and *HinfI* (5' β) loci for the HbS haplotype background (Suppl. Table S1) [17].

Malaria Associated SNPs

Using a reported method [22], SNaPshot genotyping (based on the incorporation of a single ddNTP to an extension primer designed to anneal 1 bp upstream of the target SNP) and capillary electrophoresis was used to assay the allele frequencies and relative geographical distribution of polymorphisms at novel malaria resistance loci, 1q32 within the ATP2B4 gene and 16q22.2 neighbouring the MARVELD3 gene, rs8176703, rs372091 and rs2334880. [6]. Genotyping results were confirmed using Sanger sequencing.

Data analysis and Bioinformatics analysis using data extracted from the 1000G

Genotyping at the characterised loci conformed to Hardy-Weinberg Equilibrium (HWE) (P-values > 0.05). Leveraging the moderated sample size and the accurate publicly phased data from 1000 Genomes Project, we compared the Minor allele frequencies (MAF) of all SNPs to

those of other African and non-African populations, and 2) analysed the diversity of the beta globin haplotype in five other African populations. We have used a custom python script to extract the data of five African populations from 1000 Genome project phase3 on chromosome 11 in a 100 kb region around *HBB*. The data included 108 samples from Yoruba (YRI) in Nigeria, 99 from Esan (ESN) in Nigeria, 113 from Gambia (GWD) in Western Divisions in the Gambia, 99 Luhya (LWK) in Webuye, Kenya and 85 from Mende (MSL) in Sierra Leone. Plink software [23] was used to compute the haplotype blocks in each of those populations. Each inferred haplotype blocks, was utilized in plink to estimate the haplotype frequency within the specific population. Similarly, the LD blocks were computed using Plink and the LD pattern was visualized using Haploview [24]. From a custom R script, we have made use of 20 haplotypes from each population to plot the haplotype bifurcation at the variant *rs334*.

Results

Sickle Cell genotype frequencies

The description of the HbS allele frequency, β -globin haplotype background and selected malaria-related SNPs for the study cohorts is given in Table 1. All participants from South Africa (100%, $n=50$); and the majority from Zimbabwe (88%, $n=50$) and Malawi (93.5%, $n=62$) were determined to be homozygous unaffected (HbAA), with the rest being heterozygous for the sickle mutation (HbAS).

Haplotypes in the β -globin gene cluster

SCD exists in Africa on disparate haplotype backgrounds [25] and are described by a specific pattern of five SNPs across the β -globin gene cluster [17]. This pattern confers four haplotypes associated with the HbS mutation in Africa; Benin, Bantu/Central African Republic (CAR), Senegal and Cameroon, with the fifth haplotype arising in the

Indian/Arabian peninsula (Arab/Hindu) [26, 27]. Any recombination of the defining SNPs results in recombinant haplotypes referred to as “atypical”. The SCD haplotypes were described using a previously published method and the global distribution of the haplotypes reviewed [17]. The haplotypes were described based on the analysis of chromosomes from the South Africa, Zimbabwe and Malawi cohorts (78, 64 and 70 chromosomes respectively), the most prevalent of the β -globin gene haplotypes was the atypical form; 67.95%, 65.63% and 51.43% respectively. Specifically, Atypical I was common across all 3 populations at similar frequencies, (32.1% South Africa; 38.1% Zimbabwe and 38.9% Malawi) (Suppl. Table S3). The two second most prevalent haplotypes were the Benin and Cameroon forms. In combination, the Atypical/Atypical haplotype was most frequent in the South Africa and Zimbabwe cohorts (41.03% and 37.50%, respectively) whereas the Benin/Atypical was the most frequent combination in Malawi (41.18%). Figure 1 shows the distribution of the SCD β -globin gene haplotypes amongst the study cohorts compared to the haplotypes reported in SCD patients in other African countries [17].

Targeted Malaria-related variants

Variants at novel malaria resistance loci, 1q32 within the ATP2B4 gene and 16q22.2 neighbouring the MARVELD3 gene, were selected to probe the allele frequencies and relative geographic distribution around and below the equatorial malaria belt. This approach was driven by the hypothesis that, such resistance loci even at the level of single nucleotide polymorphisms, which confer clinically significant resistance to severe malaria would undergo strong positive selection in malaria-endemic regions and to a gradual lesser extent, regions around the equatorial belt. Therefore, the allele frequencies of three variants at resistance loci [6] were investigated among three sub-Saharan African populations (Malawi, Zimbabwe and South Africa) at varying proximity to the equatorial malaria endemicity belt. SCD unaffected populations were selected in order to eliminate the possible effect of co-

inheritance of malaria resistance loci and the HbS allele, as a result of the HbS allele-conferred partial resistance to *P. falciparum*. The genotype frequencies for the rs8176703 (GG), rs372091 (GG) and rs2334880 (CT) among South African, Malawian and Zimbabwean populations were largely similar. However, when comparing MAFs at these loci with other populations from the Human 1000 Genome Project 1000Genomes Phase III, there was an apparent gradient of the MAF for rs8176703 and rs372091, highest in countries within the equatorial malaria belt (Gambia, Nigeria and Kenya) and lowest in the sub-equatorial populations investigated in this study (Table 2, Figure 2); however, this pattern was not observed in the MAF at rs2334880. When comparing the measure of frequency differentiation among the genotyped SNPs and the corresponding frequencies of these tag-SNPs in the 1000 Genomes data, the frequency of the genotyped SNPs were highest among the Southern African populations and the African populations (Esan, Luhya, Yoruba, Mende and Mandinka) (Figure 3). The frequencies were lowest between American, Asian and European populations. Table 3 and Figure S1 show the frequency of the HbS allele across African populations [14, 28-36]. When investigating the LD between these variants in the 1000 Genomes data, the variants were found to be in linkage equilibrium with their respective functional mutations.

Pattern of Linkage Disequilibrium and Haplotype blocks at rs334 in African populations

We have computed the haplotype blocks, block of linkage disequilibrium and the haplotype frequency in a 100 kb region around HBB, targeting the variant rs334 in that region which is well known of alleles A/T, encoding the Hb A form of (adult) hemoglobin and the sickling form of hemoglobin, Hb S, respectively. The results in Figures 4 and 5 show differing pattern of LD between Western and Eastern African Bantu.

Discussion

The present data aims to confirm the evolutionary dynamics of various African genomes through selective pressure of malaria, prompt the persecution of the dissociation between genetics and anthropology and culture, and lastly contribute to understanding the origin and evolution of the HbS mutation as well as migration of southern African populations within and out of Africa.

South-ward frequency decrease of malaria-associated SNPs in SSA

The high frequency of the homozygous HbAA genotype in the study cohorts was expected as the participants were neither confirmed nor suspected SCD patients, and were all from Southern African population where malaria has low prevalence or is inexistent. As a result of the known partial resistance conferred by the HbS allele to malaria *Plasmodium falciparum* infection, the HbS allele is highly prevalent in malaria endemic regions particularly around the tropical equatorial belt in sub-Saharan Africa [3,4]. Therefore, this study investigated the effect of the low environmental pressure from malaria in Malawi, Zimbabwe and, to a greater extent, South Africa, on the frequency of three genetic backgrounds associated with SCD; namely the HbS allele, conservation of the five loci in the β -globin gene cluster that confer the SCD haplotype and lastly, variants associated with resistance to malaria [6]. To confirm the accepted notion of low HbS allele frequency in populations outside malaria-endemic regions (figure S1; Table 3), was the frequency gradient of the HbS allele in the Zimbabwe (12%) and Malawi (6%) cohorts compared to the South Africa cohort (0%), in the present study.

In addition to the HbS mutation, whose association with malaria is extensively studied, the second objective was to investigate and validate three variants identified in a genome-wide association study (GWAS) as LD-tags of known functional mutations associated with malaria

resistance; rs8176703 (9q34.2; ABO locus), rs2334880 (16q22.2; MARVELD3 locus) and rs372091 (11p15.5; HBB locus) [6]. LD-tag SNPs were investigated over the known functional mutations as the populations studied were largely unexposed to malaria and devoid of the HbS mutation, a variant that has been positively selected in regions of high malaria endemicity. Furthermore, given that these variants were proposed as LD-tag SNPs, they could be genotyped as proxy for the functional mutations. In this study, the MAFs of these three variants were determined and compared to those in Gambia, Nigeria and Kenya (1000 Genomes data) and showed a decreasing gradient of MAF for two of the loci (rs8176703 and rs372091), which are tags for the functional mutations at *ABO* and β -globin, were highest within the equatorial malaria belt and lowest in all three study cohorts. However, there was no such gradient for the rs2334880 variant, with similar MAF across all six populations, therefore not specifically linked with malaria endemicity, which is concomitant with results from an independent study where this variant failed to replicate its association [2,3]. This trend suggests that although all three variants were highlighted as LD-tags for known functional mutations and have been associated with resistance to severe malaria, only two (rs8176703 and rs372091) are largely restricted to the equatorial malaria belt and possibly confer greater resistance to severe malaria as compared to the less equator-bound variant (rs2334880). The lack of LD between these tag SNPs and their respective functional mutations in other African populations is possibly because LD may differ between populations and therefore may have affected the power of these tag SNPs in association with the causal mutations and thus severe malaria.

It is however noteworthy that the *ATP2B4* [1] and *FREM3* [2] loci were not genotyped in this study due to a limitation of resources and challenges with assay optimization. Furthermore, several other loci have been associated with malaria-resistance and were not included in the present study. Future work should consider determining the frequency and

effect of not only LD-tag SNPs but include functional mutations at several loci to improve our understanding of the malaria-resistance genotype profile of Southern African populations, largely unexposed to malaria.

Haplotype blocks at rs334 in African populations

The differing pattern of LD between Western and Eastern African Bantu can be explained by the fact that Eastern African Bantu undergone admixture due to the various contacts with sea bone migrants in their past history [37]. However, the result in Suppl. Table S4 suggests similar pattern of beta globin haplotype diversity at the variant rs334 across all the five African populations (Western and Eastern Bantu). This illustrates the importance of investigating the origin and age of the HbS mutation following the past southern Bantu migration, admixture and population sub-structure.

Implications for genetics, anthropology and culture

Taken in sum; the frequency of HbS mutation and the decreasing MAF gradient of the targeted SNPs from the tropical malaria endemic regions towards the South suggest that the specific combination and pattern of multiple malaria resistant variants could allow the broad determination of the regional origins of an individual as Western, Central or Southern African. Although the vast majority of differentiated loci among Bantu populations are no more differentiated than would be expected from population drift, the modest data presented here, supports the proposed notion of dissociation between genetic background and ethno-linguistic attributes and classifications as demonstrated with populations that can loosely be defined as “Bantu” having disparate genetic backgrounds, particularly in loci that may be under positive selection due to the fitness benefit they confer, in this case resistance to severe malaria.

Within the context of Africa; that could have major implications in future research and practice on precision medicine; as gene-based treatment could be population specific, but not based on the ethno-linguistic classification, rather on target genomic variants. Despite genetic diversity, the three populations studied here can be loosely labelled as “*Bantu*” or “*Bantu-speaking*”, as is an individual from the Ewondo tribe in Cameroon, and rightly so. This is primarily due to several indicators of a common language root; for instance (i) “muntu” for “human” is the same in Xhosa (South African Bantu language) and Ewondo, (ii) the Ewondo and Xhosa tribes also share similar cultural and rite of passage practices such as the ritual of male circumcision and the burial of the umbilical cord or placenta of new-born as part of welcoming the new-born and introduction to the ancestors; (iii) and religious beliefs such as the “cult of ancestry” and reincarnation are common amongst Bantu-speakers. Despite these and many other shared cultural, linguistic and anthropological attributes, the present data further supports the notion that Bantu-speakers from Central and West Africa are no more genetically similar to those in Southern Africa, as previously illustrated with differential prevalence of HIV resistant genes amongst SSA populations [38, 39] and now with malaria associated variants. Given the vast genetic diversity within the continent and amongst any two SSA populations, the present research further emphasizes the need to redefine the classifications of various groups in Africa by region-defined genomic attributes as this approach could better serve the progress towards gene-based therapies, genetic profiling and diagnoses specific to targeted population groups based on common genetic tag loci, as opposed to the classical ethno-linguistic population classification approach. The present data among Bantu-speakers from Central and West Africa and Southern Africa, genetic relatedness at malaria-associated loci may be very different from genome-wide genetic relatedness, since the latter depends only on demographic processes whereas the former may additionally be driven by selection of specific haplotypes. Therefore, emphasizing the point

that the loose labelling of some African populations as “Bantu-speakers” and more so assuming genetic relatedness, is an inappropriate presumption, with particularly when this notion is stretched within the African continent and even beyond for individuals not of African descent.

SCD β -globin haplotype: insights into the migration of Southern African Blacks

The third question of this study was to investigate the degree of conservation of the five SCD haplotype-conferring loci in populations both largely unaffected by the disease and void of the environmental pressure of malaria. The most apparent, although not surprising result, was the high frequency of the Atypical haplotype in all the study cohorts leading to the hypothesis that in such populations, the five loci of the β -globin gene cluster may be under less evolutionary pressure to remain conserved. This could be due to several reasons; there is no apparent clinical benefit to retaining an otherwise unfavourable haplotype in the absence of malaria and potentially its strongest environmental positive selector, malaria. Furthermore, this could be as a result of genetic drift and recombination at the β -globin gene cluster. The next frequent haplotype in all study cohorts was the Benin form, suggesting that the Southern African Bantu-speakers migrated southwards post-occurrence of the HbS mutation and is consistent with their West African origins [17]. The data showing the classification of the HbS haplotypes in these Southern African populations and the degree of similarity among the haplotype distributions in South Africa, Zimbabwe and Malawi is novel. This is because these populations have historically been largely unaffected by malaria and suggest some insight into the evolutionary dynamics at the β -globin loci with regards to recombination of the classical HbS haplotypes and expansion of the Atypical form in malaria-devoid regions in Africa.

This result also confirms to anthropological data detailing the most significant events of the geographic expansion of the Bantu Niger-Kordofanian-speakers out of Cameroon and Nigeria [18,40]. It was previously hypothesised that the migration path was first through rainforest equatorial Africa and later into Eastern and Southern Africa. This is supported by the widespread distribution of Bantu-related linguistic groups and the presence of Niger-Kordofanian genetic ancestry in many African populations. However, the present result with a prevalent Benin haplotype and very low Bantu haplotype that is characteristic of SCD patients from Central and West Africa [17] could be due to a myriads of possible reasons that remain to be investigated: (i) the migration through East Africa of modern Southern African Bantu-speaking populations was transient with limited admixture with populations found locally in East Africa; (ii) some Bantu haplotypes may have been lost during recombination events at the β -globin gene locus potentially leading to the expansion of the highly prevalent Atypical form; (iii) during the early migration events through the equatorial rainforests, the migrating populations from Central, East and West Africa encountered largely unoccupied regions, therefore expanding the Benin, Cameroon and Senegal haplotypes; (iv) the Bantu haplotype could be a recent haplotype of SCD, only recently expanding in Central Africa and subsequently in some parts of North and South America through slave trade; and lastly (v) the continuous socio-economically motivated migration from Central, East and West Africa into Southern African countries could have led to the relatively higher frequencies of the Benin, Cameroon and Senegal haplotypes although unlikely as this has become a significant migration phenomenon only in the past two to three decades. Beyond the concept of the dissociation between genetic background and ethno-linguistic attributes and classifications, the present data also complements previous studies on migrations of Southern African populations from West and/or Central Africa [41].

Conclusion

This data has alluded to the frequency and distribution of (likely) positively selected polymorphisms at two novel resistance loci for severe malaria in three SCD-unaffected sub-Saharan African populations to probe the possibility of a frequency gradient of such polymorphisms in a decreasing fashion away from the equatorial malaria belt. The data is indicative of the importance of the inclusion of Southern African populations when studying the age and origin of the HbS mutation, whose precise origin remains to be fully elucidated [17]. Future studies should include Khoi and San populations, some of which may not have been exposed to malaria. Furthermore, these data also provide genetic evidence indicating the independent and continuous waves of migration of West and East African Bantu-speaking groups into Southern Africa. Beyond the modest data presented here, the high proportion of atypical haplotypes in Southern African populations, together with the data from diverse populations on the African continents could suggest various level of genetic diversification of African populations, whether attributable to recent and/or more ancient admixture, that did not probably result from a single North to South migration path nor a specific era, but rather through several independent and associated, multi-directional migration events. It can be anticipated that modern day continuous immigration, will further reinforce the African genomic diversity, by allowing the redistribution of gene pools previously restricted to specific geographical location, such as malaria-related mutations, across the continent.

References

1. **Weatherall D, et al.** A case for developing North-South partnerships for research in sickle cell disease. *Blood*. 2005; **105**:921-923.
2. **Bartolucci P , et al.** Clinical management of adult sickle-cell disease. *Current opinion in hematology*. 2012; **19**:149-155.

3. **Williams TN, et al.** Sickle cell trait and the risk of Plasmodium falciparum malaria and other childhood diseases. *The Journal of infectious diseases*. 2005; **192**:178-186.
4. **Flint J, et al.** 1 The population genetics of the haemoglobinopathies. *Baillière's clinical haematology*. 1998; **11**:1-51.
5. **Flint J, et al.** High frequencies of alpha-thalassaemia are the result of natural selection by malaria. *Nature*. 1986; **321**:744-750.
6. **Timmann C, et al.** Genome-wide association study indicates two novel resistance loci for severe malaria. *Nature*. 2012; **489**:443-446.
7. **Jallow M, et al.** Genome-wide and fine-resolution association analysis of malaria in West Africa. *Nature genetics*. 2009; **41**:657-665.
8. **Loscertales M, et al.** ABO blood group phenotypes and Plasmodium falciparum malaria: unlocking a pivotal mechanism. *Advances in Parasitology*. 2007; **65**:1-50.
9. **Fry AE, et al.** Common variation in the ABO glycosyltransferase is associated with susceptibility to severe Plasmodium falciparum malaria. *Human molecular genetics*. 2008; **17**:567-576.
10. **Steed E, et al.** Identification of MarvelD3 as a tight junction-associated transmembrane protein of the occludin family. *BMC cell biology*. 2009; **10**:1.
11. **Miller LH, et al.** The pathogenic basis of malaria. *Nature*. 2002; **415**:673-679.
12. **Idro R, et al.** Cerebral malaria: mechanisms of brain injury and strategies for improved neurocognitive outcome. *Pediatric research*. 2010; **68**:267-274.
13. **Cinel I et al.** Advances in pathogenesis and management of sepsis. *Current opinion in infectious diseases*. 2007; **20**:345-352.

14. **Piel FB, et al.** Global epidemiology of sickle haemoglobin in neonates: a contemporary geostatistical model-based map and population estimates. *The Lancet*. 2013; **381**:142-151.
15. **Elion J, et al.** DNA sequence variation in a negative control region 5' to the beta-globin gene correlates with the phenotypic expression of the beta s mutation. *Blood*. 1992; **79**:787-792.
16. **Pagnier J, et al.** Evidence for the multicentric origin of the sickle cell hemoglobin gene in Africa. *Proceedings of the National Academy of Sciences of the United States of America*. 1984; **81**:1771-1773.
17. **Bitoungui VJN, et al.** 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: a journal of integrative biology*. 2015; **19**:171-179.
18. **Campbell MC et al.** African genetic diversity: implications for human demographic history, modern human origins, and complex disease mapping. *Annual review of genomics and human genetics*. 2008; **9**:403-433.
19. **Blench R et al.** The origins and development of African livestock: archaeology, genetics, linguistics and ethnography. 2006;
20. **Saiki RK, et al.** Enzymatic amplification of beta-globin genomic sequences and restriction site analysis for diagnosis of sickle cell anemia. *Science (New York, N.Y.)*. 1985; **230**:1350-1354.
21. **Steinberg M, et al.** Hematological effects of atypical and Cameroon β -globin gene haplotypes in adult sickle cell anemia. *American Journal of Hematology*. 1998; **59**:121-126.
22. **Wonkam A, et al.** Association of variants at BCL11A and HBS1L-MYB with hemoglobin F and hospitalization rates among sickle cell patients in Cameroon. *PloS one*. 2014; **9**:e92506.
23. **Purcell S, et al.** PLINK: a tool set for whole-genome association and population-based linkage analyses. *American Journal of Human Genetics*. 2007; **81**:559-575.

24. **Barrett JC, et al.** Haploview: analysis and visualization of LD and haplotype maps. *Bioinformatics*. 2005; **21**:263-265.
25. **Labie D, et al.** Common haplotype dependency of high G gamma-globin gene expression and high HbF levels in beta-thalassemia and sickle cell anemia patients. *Proceedings of the National Academy of Sciences*. 1985; **82**:2111-2114.
26. **Elion J, et al.** DNA sequence variation in a negative control region 5'to the beta-globin gene correlates with the phenotypic expression of the beta s mutation. *Blood*. 1992; **79**:787-792.
27. **Pagnier J, et al.** Evidence for the multicentric origin of the sickle cell hemoglobin gene in Africa. *Proceedings of the National Academy of Sciences*. 1984; **81**:1771-1773.
28. **Barclay GP.** A mortality study of sickle cell anaemia in Central Africa. *Journal of clinical pathology*. 1971; **24**:768.
29. **Sarr JB, et al.** Impact of red blood cell polymorphisms on the antibody response to Plasmodium falciparum in Senegal. *Microbes and Infection*. 2006; **8**:1260-1268.
30. **Saurin W.** Natural-selection study in small size populations - The drepanocytosis case in Madinka and Bedik from the Kedougou region (Eastern Senegal). *Bulletins et Memoires de la societe d Anthropologie de Paris*. 1984; **11**:37-56.
31. **Fleming AF, et al.** Abnormal haemoglobins in the Sudan savanna of Nigeria. I. Prevalence of haemoglobins and relationships between sickle cell trait, malaria and survival. *Annals of Tropical Medicine and Parasitology*. 1979; **73**:161-172.
32. **McAuley CF, et al.** High mortality from Plasmodium falciparum malaria in children living with sickle cell anemia on the coast of Kenya. *Blood*. 2010; **116**:1663-1668.
33. **Desai PC, et al.** Alloimmunization is associated with older age of transfused red blood cells in sickle cell disease. *American Journal of Hematology*. 2015; **90**:691-695.

34. **Cox SE, et al.** Haptoglobin genotype, anaemia and malaria in Gambian children. *Tropical Medicine & International Health*. 2008; **13**:76-82.
35. **Simpore J, et al.** Biological and clinical presentation of patients with hemoglobinopathies attending an urban hospital in Ouagadougou: confirmation of the modification of the balance between Hb S and Hb C in Burkina Faso. *Hemoglobin*. 2002; **26**:121-127.
36. **Simpore J, et al.** Modification in the frequency of Hb C and Hb S in Burkina Faso: an influence of migratory fluxes and improvement of patient health care. *Hemoglobin*. 2002; **26**:113-120.
37. Chimusa ER, et al. A genomic portrait of haplotype diversity and signatures of selection in indigenous southern African populations. *PLoS Genetics*. 2015; **11**:p.e1005052.
38. **Kinomoto M, et al.** HIV-1 proteases from drug-naive West African patients are differentially less susceptible to protease inhibitors. *Clinical infectious diseases: an official publication of the Infectious Diseases Society of America*. 2005; **41**:243-251.
39. **Jenabian M, et al.** Pitfalls of antiretroviral drug resistance genotyping of HIV-1 Group M and Group N from Cameroon by sequenced-based assays. *Nigerian Medical Journal*. 2015; **56**:420.
40. **Blench R et al. (eds)** *The origins and development of African livestock: archaeology, genetics, linguistics and ethnography*. Routledge. 2006.
41. **Campbell MC and Sarah A. Tishkoff.** The evolution of human genetic and phenotypic variation in Africa. *Current biology*. 2010; **20**:R166-R173.

List of tables and figures

Table 1. Frequencies of the HbAA; β -globin haplotypes and malaria-related SNPs

Table 2. Minor allele frequency of study cohorts and several populations from the 1000Genomes Project

Table 3. HbS allele frequencies by country in Africa

Figure 1. Distribution of SCD haplotypes amongst three southern African SCD-unaffected populations (South Africa, Zimbabwe and Malawi) and SCD populations across the continent (with adaptation of previously reported from [17])

Figure 2. Minor frequencies of malaria-restriction SNPs amongst southern African populations and three populations from the 1000Genomes Project within the malaria-endemic central Africa. A: rs8176703; B: rs372091; C: rs2334880

Figure 3. Distribution of frequency differentiation of targeted SNPs rs8176703; rs372091 and rs2334880 across various African populations. When comparing the measure of frequency differentiation among the genotyped SNPs and the corresponding frequencies of these tag-SNPs in the 1000Genomes data, the frequency of the genotyped SNPs were highest among the Southern African populations and the African populations (Esan, Luhya, Yoruba, Mende and Mandinka) (Figure 3). The frequencies were lowest between American, Asian and European populations. * Populations studied in from current article (South Africa, Zimbabwe and Malawi); other data were extracted from the 1000G project. The values provided are F-statistics calculated between each MAF for the 3 tag SNPs (rs8176703; rs372091 and rs2334880) and coloured coded grey (genetically proximal) to red (genetically distal). Populations with less genetic distance have lower F-st and shown in grey whereas populations with greater genetic distance have higher F-st and are shown in red.

Figure 4. Haplotype bifurcation diagrams. The root of each diagram is a core haplotype at the variant rs333, identified by a white circle. The diagram is bi-directional, portraying both proximal and distal LD for derived (each top) and ancestral allele (each bottom). The breakdown of LD on the core haplotype background is portrayed at progressively longer distances, depending on whether allele is present or not. The thickness of the lines corresponds to the number of samples with the indicated long-distance haplotype. (A) ESN (B) YRI (C) GWD (D) MSL and (E) LWK.

Figure 5. Pattern of Linkage Disequilibrium. A linkage disequilibrium (LD) block of polymorphisms in a tight region around rs333. (A) ESN (B) YRI (C) GWD (D) MSL and (E) LWK.

Table S1. Restriction endonuclease cutting patterns that represent each of the five β -globin gene haplotypes

Table S2. Restriction endonuclease cutting patterns that represent each of the five most common Atypical β -globin gene haplotypes

Table S3. Frequency of various forms of Atypical β -globin haplotypes in Southern African populations

Table S4. Haplotypes frequency at rs334 with known alleles A/T, encoding the Hb A form of (adult) hemoglobin and the sickling form of hemoglobin, Hb S.

Figure S1. HbS allele frequency in Africa (adapted from [14])

Contribution to authorship

Conceived and designed the experiments: GP, EC, AW. Performed the experiments: EK, KM, GP. Analyzed the data: GP, EC, AW. Contributed reagents/materials/analysis tools:

AW, CD, KM, EK. Wrote the paper: GP, EC, AW. Revised and approved the manuscript: GP, KM, EC, CD, KM, EK, AW.

Funding statement

The molecular experiments of the study were funded by the National Health Laboratory Services (NHLS), South Africa; and the NIH, USA, grant number 1U01HG007459-01. The students' bursaries were supported by the Oppenheimer Memorial Trust, the National Research Foundation, and FirstRand Laurie Dippenaar Scholarship, South Africa. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Table 1. Frequencies of the HbAA; β -globin haplotypes and malaria-related SNPs

		South Africa	Zimbabwe	Malawi
		N (%)	N (%)	N (%)
β-globin mutation	HbAA	50 (100)	44 (88.0)	58 (93.5)
	HbAS	0 (0.0)	6 (12.0)	4 (6.5)
β-globin haplotypes[#]	Atypical	53 (68.0)	42 (65.7)	36 (51.4)
	Benin	13 (16.6)	8 (12.5)	19 (27.1)
	Bantu	4 (5.1)	2 (3.1)	4 (5.7)
	Cameroon	5 (6.4)	10 (15.6)	5 (7.1)
	Senegal	3 (3.9)	2 (3.1)	6 (8.7)
β-globin haplotype recombinants*	Atypical/Atypical	16 (41.0)	12 (38.0)	8 (23.5)
	Benin/Atypical	13 (33.3)	8 (25.0)	14 (41.2)
	Bantu/Atypical	4 (10.3)	2 (6.3)	2 (5.9)
	Senegal/Atypical	2 (5.1)	1 (3.1)	4 (11.8)
rs8176703	GG	35 (0.97)	48 (0.96)	48 (0.98)
	AG	1 (0.03)	2 (0.04)	1 (0.02)
	AA	0 (0.0)	0 (0.0)	0 (0.0)
rs372091	GG	34 (0.94)	42 (0.93)	48 (0.98)
	AG	1 (0.03)	3 (0.07)	1 (0.02)
	AA	1 (0.03)	0 (0.0)	0 (0.0)
rs2334880	CC	7 (0.23)	10 (0.21)	14 (0.29)
	CT	19 (0.63)	23 (0.48)	24 (0.50)
	TT	4 (0.13)	15 (0.31)	10 (0.21)

* β -globin haplotype recombinants: the pair of haplotypes inherited in 2 separate chromosomes in an individual. [#] β -globin haplotype frequencies are given as the number of chromosomes presenting with a specific haplotype

Table 2. Minor allele frequencies of study cohorts and several populations from the 1000Genomes Project

<i>Region</i>	<i>Variants</i>		
	rs8176703	rs372091	rs2334880
<i>South Africa*</i>	0,013	0,023	0,472
<i>Malawi*</i>	0,01	0,01	0,457
<i>Zimbabwe*</i>	0,02	0,033	0,552
<i>African</i>	0.398	0.068	0.398
<i>African Caribbean (Barbados)</i>	0.359	0.047	0.359
<i>Southwest US (African American)</i>	0.311	0.041	0.311
<i>Nigeria (Esan)</i>	0.399	0.141	0.399
<i>Kenya (Luhya)</i>	0.470	0.071	0.470
<i>Kenya (Yoruba)</i>	0.407	0.125	0.407
<i>Mende (Sierra Leone)</i>	0.400	0.024	0.400
<i>Gambia (Mandinka)</i>	0.403	0.013	0.403
<i>America</i>	0.102	0.001	0.102
<i>Europe</i>	0.160	1.000	0.160
<i>East Asia</i>	0.009	1.000	0.009
<i>South Asia</i>	0.019	1.000	0.019

* study populations from current study. Other data was sourced from the 1000G project.

Table 3. HbS allele frequencies by country in Africa

Country	Study years	Population	Age group	HbS allele frequency	Reference
Angola	1950-2010	18,994	All ages	0.137	[14]
Benin	1950-2010	9,219	All ages	0.159	[14]
Botswana	1950-2010	1,977	All ages	0.003	[14]
Burkina Faso	1950-2010 1997-1999	16,250 9201	All ages New-borns Median 9 years	0.056 0.0025 0.0013	[14, 30, 31]
Burundi	1950-2010	8,519	All ages	0.040	[14]
Cameroon	1950-2010	19,957	All ages	0.120	[14]
Cape Verde	1950-2010	513	All ages	0.026	[14]
Central African Republic	1950-2010	4,506	All ages	0.077	[14]
Chad	1950-2010	11,509	All ages	0.051	[14]
Comoros	1950-2010	691	All ages	0.018	[14]
Congo	1950-2010	3,760	All ages	0.145	[14]
Democratic Republic of the Congo	1950-2010	67,829	All ages	0.165	[14]
Djibouti	1950-2010	879	All ages	0.001	[14]
Equatorial Guinea	1950-2010	693	All ages	0.192	[14]
Eritrea	1950-2010	5,204	All ages	0.003	[14]
Ethiopia	1950-2010	84,996	All ages	0.003	[14]
Gabon	1950-2010	1,501	All ages	0.280	[14]
Gambia	1950-2010 2003	1,751 536	All ages New-borns 10-72 months	0.075 0.012 0.003	[14, 29]
Ghana	1950-2010 2002	24,339 1266 842	All ages 0-4 years 5-10 years	0.087 0.0039 0.0012	[14]
Guinea	1950-2010	10,324	All ages	0.168	[14]
Guinea-Bissau	1950-2010	1,647	All ages	0.041	[14]
Kenya	1950-2010 1998-1999 1998-2008	40,835 2,774 782 282 415 3,677	All ages New-borns 0-3 years 0-11 months 12-23 months 3-5 years 6-13 years	0.038 0.016 0.006 0.01 0.0035 0.0024 0.0009	[14, 27, 28]
Lesotho	1950-2010	2,064	All ages	0.001	[14]
Liberia	1950-2010	4,102	All ages	0.046	[14]
Madagascar	1950-2010	20,146	All ages	0.061	[14]

Malawi	1950-2010	15,690	All ages	0.033	[14]
Mali	1950-2010	13,362	All ages	0.057	[14]
Mauritania	1950-2010	3,359	All ages	0.050	[14]
Mozambique	1950-2010	23,418	All ages	0.027	[14]
Namibia	1950-2010	2,212	All ages	0.010	[14]
Niger	1950-2010	15,885	All ages	0.080	[14]
Nigeria	1950-2010 1970-1972	158,255 534 259 637	All ages New-borns 1-4 years 5-14 years	0.171 0.021 0.004 0.002	[14, 26]
Rwanda	1950-2010	10,277	All ages	0.023	[14]
Sao Tome and Principe	1950-2010	165	All ages	0.094	[14]
Senegal	1950-2010 2002-2003	12,866 432	All ages New-borns 2-10 years Newborn	0.067 0.005 None 0.01	[14, 24, 25]
Senegal (rural kegoudou)					
Sierra Leone	1950-2010	5,837	All ages	0.164	[14]
South Africa	1950-2010	50,523	All ages	0.003	[14]
Sudan	1950-2010	43,182	All ages	0.043	[14]
Swaziland	1950-2010	1,195	All ages	0.006	[14]
Tanzania, United Republic of	1950-2010	45,028	All ages	0.074	[14]
Uganda	1950-2010	33,798	All ages	0.082	[14]
Zambia	1950-2010 1967-1971	13,254 2,845 2,200 2,306	All ages 0-11 months 1-3 years 3-12 years	0.112 0.013 0.009 0.005	[14, 23]
Zimbabwe	1950-2010	12,645	All ages	0.021	[14]

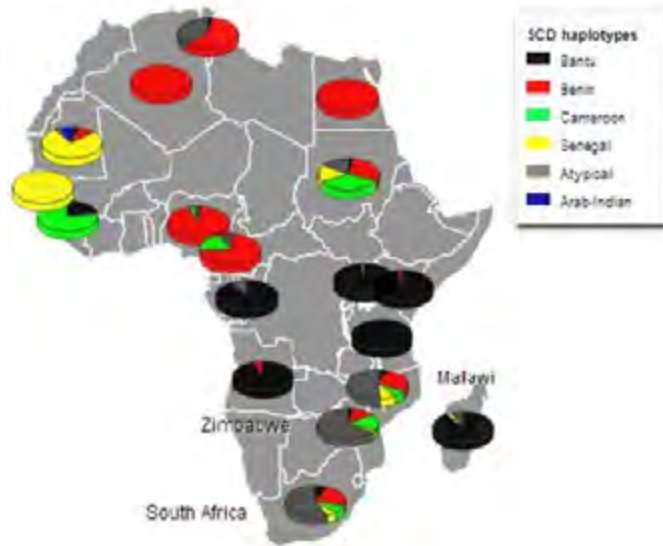


Figure 1. Distribution of SCD haplotypes amongst three southern African SCD-unaffected populations (South Africa, Zimbabwe and Malawi) and SCD populations across the continent (with adaptation of previously reported from [17])

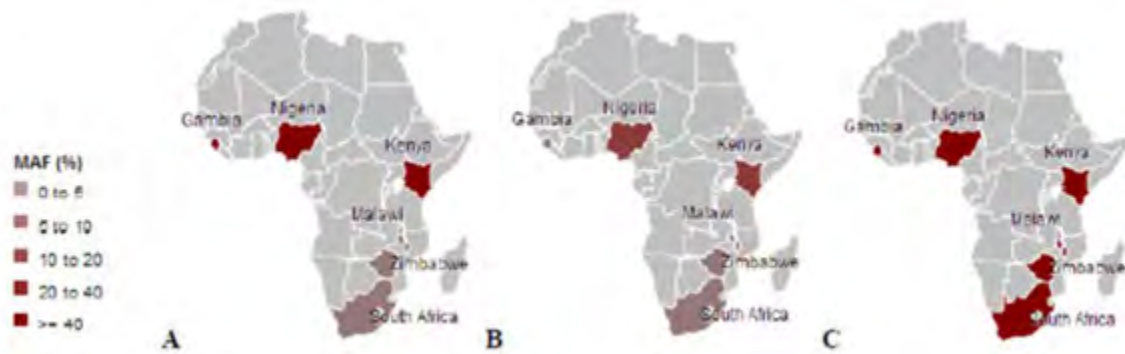


Figure 2. Minor frequencies of malaria-restriction SNPs amongst southern African populations and three populations from the 1000Genomes Project within the malaria-endemic central Africa. A: rs8176703; B: rs372091; C: rs2334880

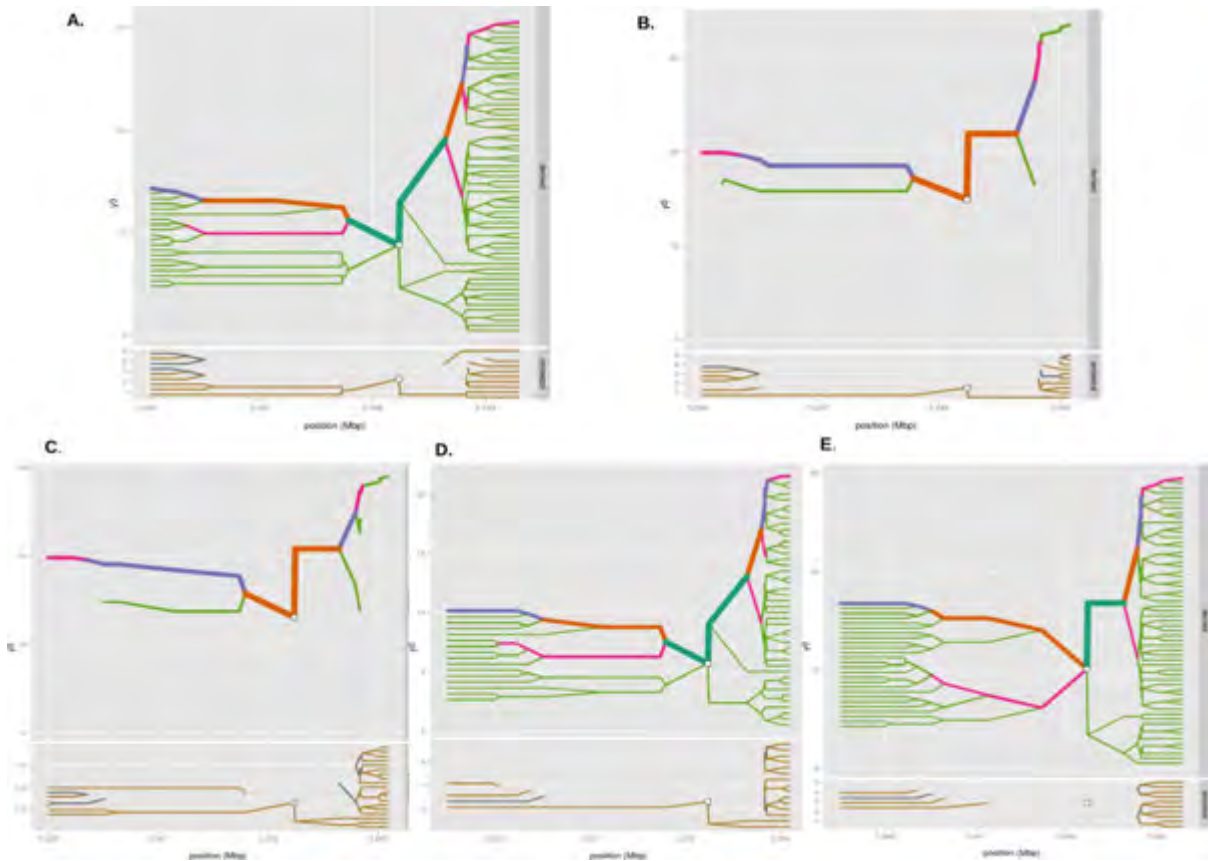


Figure 4. Haplotype bifurcation diagrams.

The root of each diagram is a core haplotype at the variant rs333, identified by a white circle. The diagram is bi-directional, portraying both proximal and distal LD for derived (each top) and ancestral allele (each bottom). The breakdown of LD on the core haplotype background is portrayed at progressively longer distances, depending on whether allele is present or not. The thickness of the lines corresponds to the number of samples with the indicated long-distance haplotype. (A) ESN (B) YRI (C) GWD (D) MSL and (E) LWK.

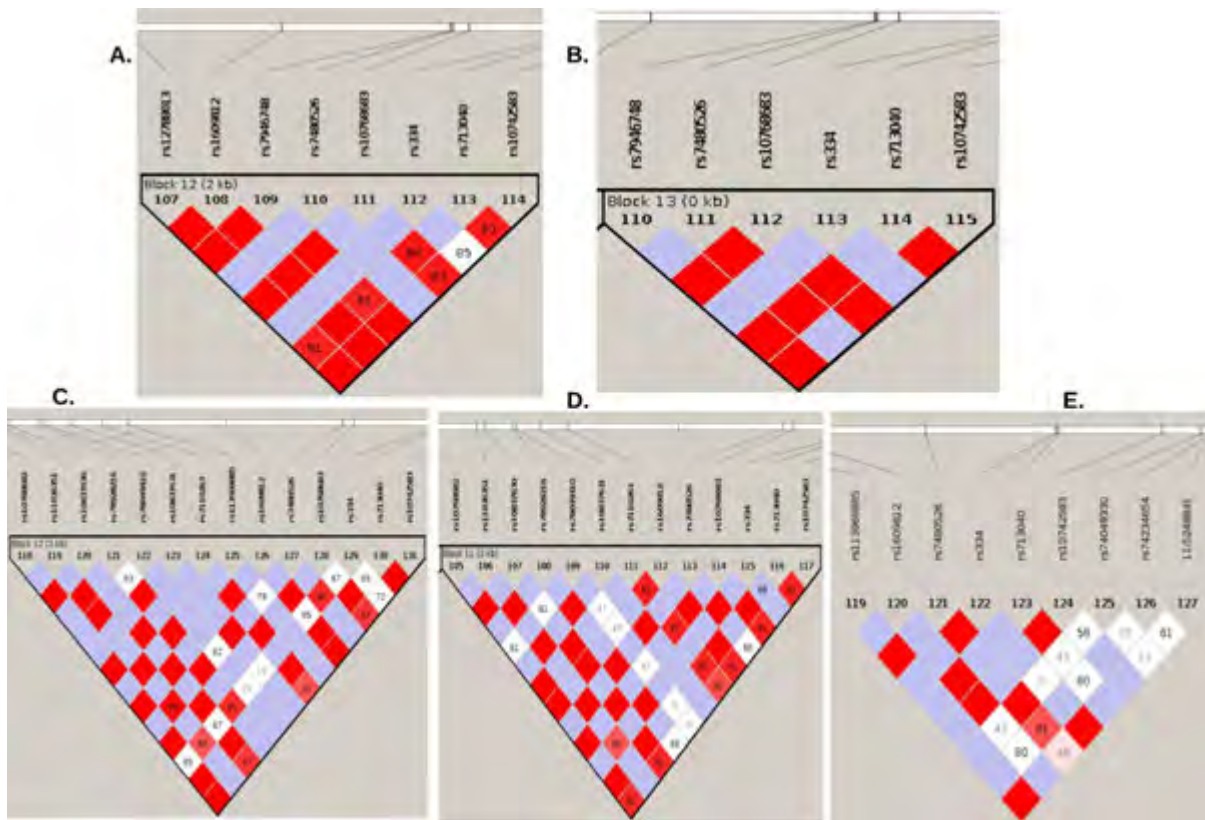


Figure 5. Pattern of Linkage Disequilibrium.

A linkage disequilibrium (LD) block of polymorphisms in a tight region around rs333. (A) ESN (B) YRI (C) GWD (D) MSL and (E) LWK.

Supplementary Table 1

Restriction endonuclease cutting patterns that represent each of the five β -globin gene haplotypes

ENZYMES

HAPLOTYPES	<i>XmnI</i> (5' <i>GY</i>)	<i>HindIII</i> (<i>GY</i>)	<i>HindII</i> (<i>AY</i>)	<i>HincII</i> (3' β)	<i>HinfI</i> (5' β)
ARAB-INDIAN	+	+	-	+	-
BANTU/CAR	-	+	-	-	-
BENIN	-	-	-	+	-
CAMEROON	-	+	+	+	+
SENEGAL	+	+	-	+	

Supplementary Table 2

Restriction endonuclease cutting patterns that represent each of the five most common Atypical β -globin gene haplotypes

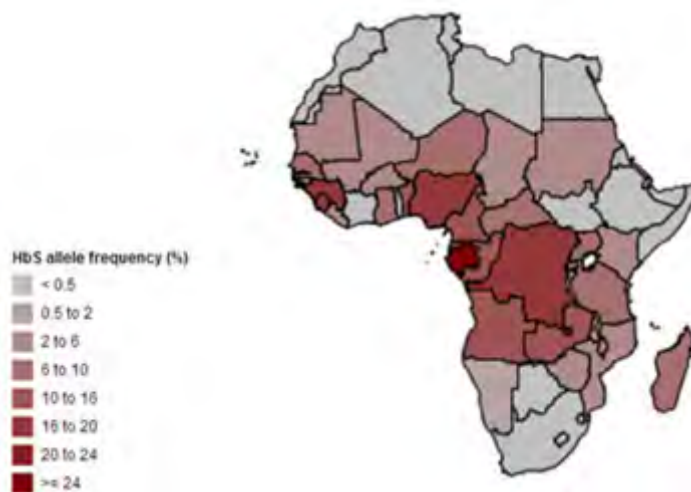
ATYPICAL TYPES	ENZYMES				
	<i>XmnI</i> (5'GY)	<i>HindIII</i> (GY)	<i>HindIII</i> (AY)	<i>HincII</i> (3'φβ)	<i>HinfI</i> (5'β)
I	-	-	-	+	+
II	-	+	-	-	+
III	-	+	-	+	+
IV	-	-	-	-	+
V	-	+	-	+	-

Supplementary Table 3

Frequency of various forms of Atypical β -globin haplotypes in Southern African populations

HAPLOTYPE FREQUENCY BY POPULATION N (%)

ATYPICAL TYPES	South Africa	Zimbabwe	Malawi
I	17 (32.1)	16 (38.1)	14 (38.9)
II	13 (24.5)	10 (23.8)	9 (25.0)
III	4 (7.5)	5 (11.9)	6 (16.7)
IV	4 (7.5)	2 (4.8)	3 (8.3)
V	4 (7.5)	3 (7.1)	2 (5.6)



Supplementary Figure 1

HbS allele frequency in Africa (adapted from [14])

Supplementary Table 4. Haplotypes frequency at *rs334* with known alleles A/T, encoding the Hb A form of (adult) hemoglobin and the sickling form of hemoglobin, Hb S, respectively. Haplotypes frequency obtained from five African populations from 1000 Genome project, including Yoruba (YRI) in Nigeria, Esan (ESN) in Nigeria, Gambia (GWD) in Western Divisions in the Gambia, Luhya (LWK) in Webuye, Kenya and Mende in Sierra Leone.

Populations*	CAGCGTGA	CAGAGAGA	CAGCGAGA	CAGAGTGA	CGGAGTAG	GGAAGTAG	CGGACTAG	CGGACTAA	CGGCCTAA
YRI	0.338	0.1389	0.0	0.4167	0.0	0.06481	0.02778	0.01389	0.0
ESN	0.3182	0.1127	0.0	0.463	0.0	0.06566	0.0	0.01515	0.0
GWD	0.2881	0.0	0.1043	0.4903	0.0	0.04253	0.02248	0.03097	0.0
MSL	0.3353	0.0	0.1176	0.4588	0.0	0.04706	0.02353	0.0	0.0
LWK	0.3081	0.0	0.101	0.5303	0.0202	0.02525	0.0	0.0	0.0101

* study populations was sourced from the 1000G project.

Chapter 4. Academic Discussion and Conclusions

Sickle Cell Disease (SCD) is the oldest monogenic disease, first reported in 1910 (Herrick, 2001) but since then grew into a worldwide health problem with a disease burden comparable to that of communicable diseases and major global diseases such as diabetes and hypertension (Weatherall, 2010). The genetic basis for the disease is a single nucleotide substitution (T>A) in the 17th position of the β -globin gene on chromosome 11p (Bunn, 1997), leading to a glutamine to valine peptide change thereby creating a hydrophobic motif in the deoxygenated sickle hemoglobin (HbS) (Brittenham, Schechter & Noguchi, 1985). This results in the polymerization and growth of the β 1 and β 2 chains of the HbS molecules as they crystalize within the erythrocytes and alter the intracellular architecture, cellular conformation and oxygen-carrying capacity of the cells. This process is the pathophysiological basis of the disease and all associated symptoms as the aberrantly-shaped sickle cells cause occlusion of micro-vessels, resulting in denaturation of hemoglobin, decreased oxygen delivery to tissues, regional hypoxia and increased HbS polymerization and erythrocyte rigidity (Gladwin et al., 2012). Vaso-occlusive crises (VOC) is therefore a hallmark phenotype of the disease, including acute chest syndrome (ACS) and cerebrovascular complications such as silent and overt stroke.

Sickle cell anemia (SCA) is the most common and severe form of SCD, accounting for approximately 70% of the disease cases in African populations (Serjeant et al., 2005). The burden of disease in sub-Saharan Africa well exceeds that of any other region in the world sitting at 80%, with an estimated 305 800 annual new affected births (Modell & Darlison, 2008). Although the disease is most rampant in sub-Saharan Africa, it is not to the exclusion of North America and Europe with 2 600 and 1 300 annual new affected births. The high incidence of SCA in sub-Saharan Africa is largely ascribed to the positive selection conferred by malaria as a result of the geographical co-occurrence of SCD and malaria (Flint et al., 1998; Williams et al., 2005). It is accepted that the primary reason for this amplification of the HbS allele frequency in malaria endemic regions is the partial-carrier resistance to *Plasmodium falciparum* malaria, even though the precise

mechanisms of resistance are not yet fully elucidated. Similarly, other genetic variants associated with SCA such as HbC and β -thalassemia have undergone similar selection.

The disease exists on a variable genetic haplotype background of five independent, region-specific disease haplotypes based on the permutation of five SNPs in the β -globin gene cluster: Benin, Bantu/Central African Republic (CAR), Senegal, Cameroon and the Indian-Arab haplotypes (Pagnier et al., 1984; Labie et al., 1985; Elion et al., 1992). These haplotypes are said to be multi-centric origins of the HbS gene and have also been associated with disease severity and clinical course of patients (Steinberg, 2009; Alsultan et al., 2012). Early child mortality is a common feature of SCD with approximately 10% of affected children suffering from cerebrovascular accidents (Ohene-Frempong et al., 1998). One other frequent complication that is associated with vaso-occlusion is pulmonary hypertension (Collins & Orringer, 1982). Early diagnosis has been shown to decrease disease-related mortalities (Gaston et al., 1986; Vichinsky, 1991) and significantly improve survival rates in the Caribbean (Lee et al., 1995), United States (Quinn et al., 2004), the United Kingdom (Telfer et al., 2007) as well as in Africa (Rahimy et al., 2009).

The key to improved clinical management and therapeutic interventions lies in a comprehensive understanding of the genetic heterogeneity of the disease, the molecular processes involved in drug-induced disease amelioration and a firm knowledge base of the effects of the vast genetic variation in African patients on predispositions to symptoms, drug-metabolism and quantitative trait loci such as HbF. To our knowledge, this thesis is the first to collectively address all these 3 key aspects of SCD in African patients and unequivocally demonstrate a functional model of the miR-26b/*MYB* post-transcriptional signalling pathway for the induction of HbF by HU treatment. The significance of this work will be discussed within the context of the currently available data on the genetic modifiers of HbF and mechanisms of HU-induced HbF expression in SCD in the following parts; (i) genetics of HbF; (ii) alternative treatment approaches for SCD and (iii) SCD in Southern Africa.

4.1. Genetics of fetal hemoglobin

Persistent fetal hemoglobin (HbF) is a heritable quantitative trait locus (QTL) and has been shown to ameliorate symptoms of SCA in adult patients. Research efforts have thus been directed towards building a comprehensive understanding of the genomic modifiers of HbF in the search for therapeutically amenable polymorphisms with significant bearing on this QTL. Three principal loci have been elucidated; *BCL11A*, the intergenic region between *HBS1L* and *MYB* (*HMIP1/2*) and the β -globin haplotype. Together, these have been shown to account for 10 – 20% variance in HbF levels (Menzel et al., 2007; Creary et al., 2009; Thein et al., 2009); amongst various SCD patient populations including the USA and Brazil (Lettre et al., 2008), Tanzania (Makani et al., 2011) and Cameroon (Wonkam et al., 2014).

In this study, we investigated additional HbF-promoting loci with the intention of identifying key regulatory loci of persistence of HbF in adult SCD patients of African descent. The envisaged implications of this approach was to (i) prompt a revision of the central dogma; multiple independent origins of β -globin haplotypes, and interrogation of the geographic nomenclature of haplotypes; (ii) elucidate polymorphisms that are critical in HbF production in an African population and (iii) highlight other variants that may be of significance in European or African American populations but bear neither genomic nor clinical significance in our African study cohort.

This intention was made clear with the work on β -globin haplotypes in a 2-part study of a cohort of SCD patients from Cameroon in which the β -globin haplotypes of the patients were investigated as well as an extensive review of the literature for the global distribution of β -globin haplotypes. After the evaluation of 1082 chromosomes, no significant association between haplotypes and clinical events, anthropometric measures, hematological indices or HbF levels. However, we reported what appeared to be peculiar at first sight; a 74% ($n = 799$) frequency of the Benin haplotype in the cohort and only 19% ($n = 207$) with the Cameroon haplotype background (Bitoungui et al., 2015). This was unusual as the expectation would be the predominance of the Cameroon haplotype in SCD patients from Cameroon. Equally unexpected was the

relatively high frequency of the Cameroon haplotype in Sudan compared to Cameroon (Elderderly et al., 2012) coupled with the wide representation of all African β -globin haplotypes in Sudan. These observations, however unusual, could be attributed to the geographic nomenclature of haplotypes (naming in accordance with the region of the first observation). This system is flawed in that it fails to reflect the accurate locale of origin of the haplotypes. Furthermore, the current central dogma, that is the hypothesis of multiple independent origins, also raises the question of how several identical mutations could have occurred in a short period of time in Africa after the appearance of malaria, despite the low mutation rate of nuclear DNA (Currat et al., 2002). Therefore, taken in sum, it is perhaps provocative and ambitious, but not unreasonable, to hypothesize that there could be a single origin of the HbS mutation in the region of East Africa/Sudan, with additional haplotypes generated by diverse structural mechanisms such as gene conversion, expanded and spread to the rest of the continent through traditional migratory patterns. This assertion may be supported by recent phylogenetic analysis of African ancestry pointing to the Nubian tribe of Sudan as the 2 most ancestral mitochondrial sequences (Elhassan et al., 2014). The expansion out of Africa and rise of the Indian-Arab haplotype could possibly be a product of human migration out of Africa through the Arab peninsula from Dhofar Oman (Rose et al., 2011). To the best of our knowledge, the present study represents the largest study of β -globin haplotypes in a single African country, and is the most comprehensive review of the global distribution of β -globin SCD haplotypes thus far.

HbF is the most significant biological modifier of SCD and influenced by *cis*- and *trans*-acting genomic variations. In a recent study, in this group of Cameroonian SCD patients, robust genetic associations between HbF levels and *BCL11A* and *HBS1L-MYB* intergenic locus were demonstrated (Wonkam et al., 2014). This was congruous with previous reports in African American, Afro-Brazilian (Lettre et al., 2008) and Tanzanian SCD patients (Makani et al., 2011). In this study, we showed significance association between variants at the *BCL11A* enhancer (rs1427407 and rs7606173) and HbF levels in Cameroonian patients (Pule et al., 2015). These variants explained 8% and 6.2% variance in HbF, respectively, and confirmed a stronger effect of these extragenic polymorphisms on HbF levels compared to intragenic variants. The present

article is the first report on the importance of the SNPs rs1427407 and rs7606173 among West African SCD patients and the results are consistent with those reported among African Americans (Bauer et al., 2013) and patients from Saudi Arabia and India (Sebastiani et al., 2015) and Tanzania for rs1427407 (Mtatiro et al., 2015). What is probably of greater significance beyond the associations are the implications on future therapeutic interventions. This is because the major clinical caution of exploring decreased expression of *BCL11A* as a route toward promoting HbF production comes from the deleterious functional implications of altered expression levels of this developmental-stage specific transcription factor as reduced *BCL11A* levels display neurodevelopmental abnormalities including intellectual disabilities and brain (Peter et al., 2014; Balci et al., 2015; Funnell et al., 2015). This thus amplifies the significance of the genetic variation at the *BCL11A* erythroid-specific enhancer and confirmed effect on HbF levels. The on-going pre-clinical trials for hemoglobinopathies (SCD and β -thalassemia) at Sangamo BioSciences (www.sangamo.com/pipeline/hemoglobinopathies) is the exact embodiment of the significance of these results. The therapeutic intervention utilises genomic editing of autologous hematopoietic stem cells to introduce favourable variants at the *BCL11A* enhancer and re-transplantation as a curative approach. The present study has thus contributed to the knowledge base of the effect of the erythroid-specific enhancer on HbF and beyond that, demonstrated the applicability of clinical trials such as this (discussed above) for African SCD patients.

However, not all polymorphisms have clinical significance across many populations. Recently a genome-wide association study (GWAS) of HbF in Tanzania revealed 8 novel loci with genome wide significance in SCA patients (HbS/ β 0) (Mtatiro et al., 2014) and whole genome sequencing of the general population in Sardinia provided further insight into regulation of HbF and several novel variants (Danjou et al., 2015). Similarly, 4 variants (rs2310991; rs4282891; rs76901216 and rs76901220) in the promoter sequence of the HU-inducible small guanosine triphosphate (GTP)-binding protein and secretion-associated and RAS-related (*SAR1a*) protein have been associated with HbF (both baseline and HU-induced) in African American patients (Kumkhaek et al., 2008). We replicated all 3 studies in patients from Cameroon ($n = 484$; 484 and 453,

respectively) and observed no association between any of the variants and clinical events, anthropometric measures, hematological indices or HbF levels; and 6 of a total of 21 loci were monomorphic in our cohort. In fact, the Tanzanian GWAS failed to show association in a British cohort from the same study (Mtatiro et al., 2014) due to the different ancestries between the East Africa (Tanzania) and the largely Caribbean/West African descents (United Kingdom), citing the considerable heterogeneity of both the genetic and environmental factors affecting the manifestation of SCD within African populations (Wonkam et al., 2014), even more so between populations in different continental settings. It is therefore likely that, the differences in ancestry between Tanzanian patients (majority of which carry the Bantu/CAR haplotype) and Sardinian populations (mostly if not all HbAA and without the selection pressure of malaria) compared to Cameroonian SCD patients (largely Benin and Cameroon haplotypes), could explain the incongruence in association results amongst the studies. HbF regulation is a highly heterogeneous system and subject to regulatory and signalling pathways that develop differently across different populations (Solovieff et al., 2010). Therefore, replication of association studies across different populations, particularly of varied genetic backgrounds and environmental settings, is imperative to understanding the complex processes of HbF production and genetic polymorphisms predisposing patients to persistence of adult HbF, more so in sub-Saharan Africa where the burden of disease is highest.

4.2. Alternative treatment approaches to Sickle Cell Disease

Currently, HU is the only pharmacologic drug for treatment of SCD patients. HU received FDA approval for treatment of both adults and children with SCD in 1998 and by the European Medicines agency in 2007. In addition, the National Institutes of Health Officer of Medical Applications of Research and the Agency of Healthcare Research and Quality both declared HU as an effective drug treatment for adults and children with SCD (Brawley et al., 2008; Lanzkron et al., 2008; Strouse et al., 2008). Since its first clinical application for hemoglobinopathies in 1984, HU has demonstrated swift and

vivid increases in HbF levels and low bone marrow toxicity (Platt et al., 1984a). Several clinical trials since have further elucidated the clinical efficacy of HU in decreasing the frequency of painful episodes, acute chest syndrome (ACS), hospitalization and transfusion (Charache et al., 1995), increase in survival rates and life expectancy (Steinberg et al., 2003; Voskaridou et al., 2010), protection against cerebrovascular disease (Zimmerman et al., 2007), long-term drug safety, capacity to prevent organ damage and reduced morbidity and mortality in school-age children (Kinney et al., 1999), toddlers (Hankins et al., 2005; Thornburg et al., 2009) and infants (Wang & Thompson, 2010). Despite this evidence, some potential short and long-term potential adverse effects such as male impotence (Berthaut et al., 2008; DeBaun, 2014; Smith-Whitley, 2014), susceptibility to infections (Shemisa et al., 2014; Amoako et al., 2014; Hirst et al., 2012; Owusu et al., 2015), potential teratogenic effect (Charache et al., 1995; Flanagan et al., 2010) and cutaneous adverse reactions (Chaine et al., 2001) have also been associated with HU. Therefore, the fear of such side effects is concerning to some health professionals (Brandow et al., 2010; Oyeku et al., 2013), parents and patients (Zumberg et al., 2005; Lanzkron et al., 2008; Haywood et al., 2009; Thornburg et al., 2010; Strouse & Heeney, 2012) and constitutes the biggest potential barriers to compliance in some settings (Hampton, 2008; Lebensburger et al., 2013). As a consequence, HU remains underutilized (Zumberg et al., 2005; Lanzkron et al., 2008), despite the studies that have reported on the overall drug safety (Steinberg et al., 2003) and limited evidence regarding potential HU-induced leukemogenic (Segal et al., 2008) or teratogenic effects (Ballas et al., 2012).

Despite this plethora of knowledge on HU and its clinical applications, there remains several questions on its molecular mechanisms of HbF induction. To our knowledge, as part of this thesis, we have provided the first comprehensive systematic review of the known mechanisms of HU-induced HbF production in SCD, citing several key signalling pathways (Pule et al., 2015) and elucidated a novel HU-induced miRNA-mediated post-transcriptional pathway for HbF expression (Pule et al., 2016). Furthermore, to define a perspective on the overall management and treatment of SCD in sub-Saharan Africa, we provided a commentary on the availability of various treatment approaches for SCD in Africa compared to the West, proposing a collaborative cross-continent effort from

government, private sector and research institutes towards improving access to good medical care for patients (Pule & Wonkam, 2014).

To fully understand the progress to date on known mechanisms of HU induction of HbF, we began with an extensive and comprehensive review, consulting a total of 562 articles. After careful scrutiny and strict selection criteria, 361 abstracts were reviewed for further elimination and from this a total of 129 articles were eventually included in the review. This formed the basis of the current knowledge on HU. It is noteworthy that although we recognise other disease modulating effects of HU outside of HbF induction such as stress hematopoiesis, endothelial nitric oxide release or reduction of leukocyte counts, the focus of the review was particular to HbF production including genomic polymorphisms that predispose hereditary persistence of HbF in adulthood. Therefore, 3 signalling levels were reported; (i) epigenetic and transcriptional events, eluding to DNA remodelling and methylation pathways; (ii) signal transduction pathways, including $G_{i\alpha}$ /JNK/Jun, cGMP/cAMP and a myriad of erythroid-related transcriptional factors, and (iii) post-transcriptional and translational regulation through miRNAs. Key epigenetic and transcriptional events regulating response to HU treatment included DNA remodelling events through the *BCLL1A*, *SOX-6* and Mi-2/nucleosome remodelling and deacetylase complex and differential methylation at regulatory sites like the 5' locus control regions and 3' hypersensitivity sites flanking the globin genes in the β -globin cluster. Such epigenetic modifications have proven to play a significant role in the fetal switch (γ - to β -globin) and thus may hold some insight to future therapeutic interventions. Several signal transduction pathways such as the $G_{i\alpha}$ /JNK/Jun, cGMP/cAMP, MAPKs and stress signals like superoxide and hydroxyl radicals were also implicated in HU-induced HbF production. The intertwinement of these pathways were collated into a signalling network of various HU-activated signal transduction pathways for HbF production (Pule et al., 2015). This network also included miRNAs as a pivotal post-transcriptional and translation regulatory tier in HbF production, erythropoiesis and erythrocyte physiology. Further investigation of translation-altering agents and the global disease-specific miRNA profiles of SCD patients in understudied and highly heterogeneous populations, such as in Africa, promises to yield mechanistic insight into the disease phenotype and differential response to HU treatment.

In line with this very hypothesis, we then investigated the potential post-transcriptional regulation of HbF through miRNAs and elucidated a novel HU-induced miR-26b/*MYB*-*BCL11A* pathway for γ -globin expression using erythroid cells derived from umbilical cord blood CD34+ hematopoietic stem cells (HSCs) and K562 cells (Pule et al., 2016). We showed that HU causes down-regulation of *BCL11A*, *KLF-1* and *MYB* was associated with a 7-fold up-regulation in HbF in both cell models. HU also differentially regulated several miRNAs including miR-15a, miR-16-1, miR-148a and miR-151-3p, however the increased production of HbF was particularly associated with miR-26b expression. The inhibition of the miRNAs, specifically miR-26b, unequivocally demonstrated the causative effect of HU on the induction of HbF through miRNA-mediated inhibition of *MYB* expression. The pathway between *MYB* (activator of *KLF-1*), *KLF-1* (activator of *BCL11A*) and *BCL11A* (potent repressor of HbF) was altered by HU-induced expression of miR-26b, which was shown to directly target the 3' untranslated regions (UTR) of *MYB*. Therefore, this demonstrates the potential of miRNAs as a viable therapeutically responsive tier of HbF production for SCD treatment, which may spare the undruggable nature of transcription factors such as *BCL11A* in non-erythroid cells.

The use of K562 cells as an erythroid cell model remains contentious despite the fact that several studies have and continue to utilise this cell line as a model to investigate various components of γ -globin expression (Sankaran et al., 2011; Hahn & Lowrey, 2014; Trakarnsanga et al., 2014; Zhu et al., 2014; Finotti et al., 2015; Aimola et al., 2016). The main bone of contention is their bias expression of γ -globin. We however, saw this as an advantage in 2 ways: (i) the aim of this study was not to demonstrate the switching in expression of the globin genes (β to γ) but rather to use a stable cell line model in conjunction with cord blood-derived erythroid cells to demonstrate the associated increases in γ -globin and decreases in its regulators (*BCL11A*, *KLF-1* and *MYB*) in response to HU, (ii) in choosing the whether to use miRNA mimics or anti-miRNAs (miR-inhibitors), the high expression of γ -globin in K562 was considered. If our hypothesis of HU-induced miRNA-mediated repression of *MYB* was correct, the use of

mimics would not only introduce unnaturally high levels of miRNAs into this model (which we felt may in fact result in some artefactual results merely because of this unnatural state of aberrant concentration of miRNAs in the cells) but it would also be compounded by any co-treatments with HU in causing γ -globin expression. This rationale led to the choice of anti-miRNAs, which would go toward decreasing the already high levels of γ -globin in K562 (via miR-26b/*MYB*) and also provide an obvious opportunity for HbF rescue (with HU co-treatment) after miR-26b inhibition. Therefore, future studies should continue to evaluate the *in-vivo* impact of efficacious concentrations of HU on the erythroblast transcriptome, proteome, as well as the erythroid-specific microne of SCD patients (before and after HU treatment) as the global analysis of these epigenetic mechanisms could highlight multiple components of this complex and heterogeneous regulatory system and possibly yield alternative (possibly miRNA-based) therapeutic approaches to hemoglobinopathies.

Considering the scientific and clinical progress made in the past 106 years since the first report of SCD in 1910, we interrogated the available treatment approaches for SCD in Africa and examining the issue of investing in HSC transplantation in Africa (Pule & Wonkam, 2014). We reported on 5 broad treatment approaches (Ballas et al., 2012): (i) supportive, which is the most common and included a balanced diet, hydration and folic acid; (ii) blood transfusions, analgesia and antibiotics characterise the symptomatic approach as they specifically serve to alleviate a target symptom(s); (iii) the preventative approach is taken to preclude the occurrence of disease complications such as pneumonia and influenza vaccination for infections, HU for low HbF levels and blood transfusions to avert primary and secondary stroke episodes; (iv) although widely contentious, nitric oxide treatment is considered an abortive approach, reported to terminate chronic pain episodes and lastly (v) HSC transplantation as the sole curative treatment for SCD. We reported that despite accounting for >80% of the burden of disease, sub-Saharan Africa still faces major barriers to quality care and management of patients, which include safe and reliable blood transfusion services, clinical and laboratory expertise and equipment for accurate diagnosis, vaccinations and pain management drugs and stigmatization of the disease. SCD is becoming a worldwide health problem, with a burden comparable to that of communicable diseases and other

major global diseases such as diabetes and hypertension (Weatherall, 2008). Therefore, we strongly advocate for a commitment to collaboration among all stakeholders: health ministries, governments, private sector and pharmaceutical companies to develop cost-effective procedures for HSC transplants for hemoglobinopathies like SCD and β -thalassemia; conduct large-scale clinical trials and follow-up studies; perform studies on the epidemiology; health-related quality of life (HRQL) and patient survival rate; and most importantly, build technological and research capacity in Africa through teaching and training laboratory technicians and researchers.

4.3. Sickle Cell Disease in Southern Africa

The frequency of the HbS mutation has a strong correlation with the historical distribution of incidences of malaria, primarily due to the partial carrier-resistance to *Plasmodium falciparum* malaria parasite (Williams et al., 2005). Therefore, the geographical co-occurrence of the HbS mutation and malaria, taken in sum with the conferred partial carrier-resistance is believed to have resulted in the local amplification and positive selection of SCD in malaria-endemic regions (Flint et al., 1986; Flint et al., 1998). Therefore, in sub-Saharan Africa, SCD has largely remained within the confines of the tropical equatorial malaria belt and its burden rarely experienced in southern Africa. As a result, similarities in anthropology and/or cultural practices have been used as proxy for generalizing genomic data.

In this part of the study, we interrogated the commonly assumed congruence between genetics and culture/language/anthropology and sought to understand the origin and distribution of the HbS mutation and other SCD-related polymorphisms among southern African populations. Furthermore, we studied and confirmed the increasing burden of adult SCD patients at Groote Schuur Hospital (Cape Town), as previously reported in paediatric patients at Red Cross War Memorial Hospital (Cape Town) (Wonkam et al.,

2012), with the primary aim being to caution national health and academic institutions to adapt policies, health care and professional training, accordingly.

Within our 3 southern African study cohorts; Malawi, Zimbabwe and South Africa, we investigated the effect of the low environmental pressure from malaria in Malawi, Zimbabwe and, to a greater extent, South Africa, on the frequency of three genetic backgrounds associated with SCD; namely the HbS allele, conservation of the five loci in the β -globin gene cluster that confer the SCD haplotype and lastly, variants associated with resistance to malaria. Although we only interrogated about 100 chromosomes from each cohort, the increasing frequency gradient of the HbS allele from 0% (South Africa), 6% (Malawi) and 12% (Zimbabwe), is proof of the concept that the HbS allele frequency increases with proximity to the tropical equatorial malaria belt. A similar trend was also observed for 2 of the 3 polymorphisms previously associated with malaria-resistance (Timmann et al., 2012), suggesting that although all three variants have been associated with resistance to severe malaria, only two (rs8176703 and rs372091) are largely restricted to the equatorial malaria belt and thus possibly confer greater resistance to severe malaria as compared to the less equator-bound variant (rs2334880). The most apparent, although not surprising result, was the high frequency of the Atypical haplotype in all the study cohorts leading to the hypothesis that in such populations, the five loci of the β -globin gene cluster may be under less evolutionary pressure to remain conserved. This could be due to several reasons: there is no apparent clinical benefit to retaining an otherwise favourable haplotype in the absence of SCD and potentially its strongest environmental positive selector, malaria. Furthermore, this could be as a result of genetic drift and recombination at the β -globin gene cluster

Although the vast majority of differentiated loci among Bantu populations are no more differentiated than would be expected from population drift, the modest data presented here, supports the proposed notion of dissociation between genetic background and ethno-linguistic attributes and classifications as demonstrated with populations that can loosely be defined as “Bantu” having disparate genetic backgrounds, particularly in loci that may be under positive selection due to the fitness benefit they confer, in this case

resistance to severe malaria. Despite the several shared cultural, linguistic and anthropological attributes such as common language roots, rite of passages and religious belief; our data further supports the notion that Bantu-speakers from Central and West Africa are no more genetically similar to those in Southern Africa, as previously illustrated with differential prevalence of HIV resistant genes amongst SSA populations (Kinomoto et al., 2005; Jenabian et al., 2015) and now with malaria associated variants. Given the vast genetic diversity within the continent and amongst any two SSA populations, the present data further emphasizes the need to redefine the classifications of various African groups by region-defined genomic attributes. This approach is likely to better serve the progress towards precision medicine, as opposed to the classical ethno-linguistic population classification approach. What remains apparent is the fact that the evident gene admixture and diversification of African populations did not result from a single North to South migration path nor a specific era, but rather through several independent and associated, multi-directional migration events further fortified by modern day immigration and continentalisation that continues to distribute previously restricted gene pools across the continent.

As a testament of the distribution of previously locale-specific gene polymorphisms, we provided the first report describing the clinical and genetic backgrounds of SCD patients at Groote Schuur Hospital (GSH) and reporting on adult patients in South Africa. This study is a follow-up of a previous report at Red Cross War Memorial Children's Hospital in Cape Town, showing rapid increase in the number of paediatric cases of SCD (Wonkam et al., 2012). From 128 patients' files from 1995 to March 2016 reviewed in this study, 47 patients were diagnosed with some form of α - or β -thalassemia and of the remaining 81 patients affected by SCA (HbSS), 61 (75.3%) were from other Sub-Saharan African countries. A key finding of this study was the marked 200% increase in the number of new cases of SCD patients at the GSH hematology clinic between 2011 and 2016. The marked increase in patients between 2001 and 2005 could be associated with the ending of a civil war, transitional government and political instability in Congo, which had spread into neighbouring states whereas the 2011 – 2016 increase is more likely due to economic and health-motivated migration. This trend is expected to continue, compounded by future referrals from neighbouring paediatric hospitals, arrival

of new migrant patients, and the improved clinical management and healthcare of paediatric patients as the percentage of patients surviving to reproductive age continues to increase. However, the increasing prevalence of SCD is not unique to South Africa as similar trends have been reported in Ireland (Gibbons et al., 2015), Italy (Colombatti et al., 2013), Germany (Zur, 2016), England (Pizzo et al., 2015) and France (Habibi et al., 2015). The novelty of this work is specifically the report on the genetic background of critical modifying loci of HbF; polymorphisms at the *BCL11A* erythroid-specific enhancer, β -globin haplotypes and α -thalassemia 3.7kb gene deletion. To improve clinical management and pharmacological accuracy in treatment, it is imperative to understand the genetic variants affecting the predisposition to specific complications such as stroke, acute chest syndrome, variants associated with susceptibility to pain, pharmacogenomics of the commonly prescribed treatments such as HU, malaria prophylaxis and vaccinations for infections. It warrants declaration that the intention of this study was not to, in any way, stigmatize SCD or immigrant patients but rather to provide necessary awareness to academic and training institutions, national health ministries and agencies about the changing distribution of the HbS allele and prevalence of SCD in Southern African countries, particularly in South African hospitals in order to better inform policies, regulations and management of SCD.

Study limitations

Limitations of molecular studies

A limitation of the molecular studies was the use of two methods for HbF measurement: alkaline denaturation test (ADT) and, later when it became available, high performance liquid chromatography (HPLC) during the prospective patient recruitment in Cameroon. Although significant associations with HbF levels were present in both sub-sets of the patient cohort (after disaggregation of patient samples based on the HbF assessment technique (ADT vs HPLC)) and all association analysis in all the studies reported in this thesis were corrected for the electrophoresis technique, the use of HPLC for the whole cohort would have allowed more precise statistical approximation of the effect of various

polymorphisms on clinical events, disease course and hematological indices of SCD. Although the sample sizes used in the studies reported in this thesis were sufficient for the assertions and conclusions drawn from their results, more robust patient numbers, particularly in the cohorts used for the southern African perspective of SCD, would assist in the development of a more accurate model for the expansion and distribution of the HbS allele in southern Africa. Validation of novel variants associated with HbF levels in non-African populations is clearly important as it has been demonstrated by two of the manuscripts included in this thesis. However, the assessment of tag or targeted SNPs may be insufficient to replicate associations across populations. This may be because LD differs across different populations and therefore tag SNPs may very well be dissimilar between any two populations. The study of the burden of SCD in Cape Town was limited due to its retrospective nature as clinical and phenotype information were collected from patient files; thus the cohort was not emphatically phenotyped. A comparison among other populations within and outside of Africa with (and without) malaria exposure would be the ideal experiment to further demonstrate the inheritance of HbS-related polymorphisms in HbAA individuals. The experiments in this thesis were hypothesis-testing and have succeeded in providing interesting preliminary data that requires exploration in other populations and at scale.

Limitations of functional studies

One of the limitations of the functional studies was the limited number of CD34⁺ hematopoietic stem cells (HSCs) from umbilical cord blood samples. Post expansion and differentiation of primary erythroid cells, meant that there was often a limitation on the number and variety of experiments that could be carried out. Although we were able to demonstrate the inter-individual variation in response to HU, the use of non-clonal populations of erythroid cells introduced variables into the *ex vivo* model. The results from the *ex vivo* model should be validated in CD34⁺ HSCs in peripheral blood as well as bone marrow aspirates from SCD patients with varying clinical severity in order to

develop our understanding of the relation between patients' miRNome and the clinical course of the disease.

Implications of research

There are several implications of the aspects of the research presented here as well as of the thesis as a whole: (i) The interrogation of genetic polymorphisms that modulate the hereditary persistence of the disease-ameliorating HbF in a cohort of African patients with severe clinical events such as stroke and vaso-occlusive crises will improve our understanding of the genomics and clinical management of African patients. Furthermore, the provocative hypothesis of a single origin of the HbS mutation in Africa, expansion and diversification of haplotypes will initiate an academic dialogue on the geographic nomenclature that forms the central dogma of SCD haplotypes; (ii) The work presented on the erythroid specific enhancer of *BCL11A* represents one of the most significant contributions of the work presented to our understanding of the heterogeneity and complexity of HbF regulation. The outcome of the *BCL11A* enhancer pre-clinical trial in the USA could be a giant leap in SCD treatment; (iii) The systematic review and comprehensive synthesis of the known mechanisms of the HbF-promoting effect of HU will serve as a reference for future studies on the progress made to date in building an understanding of the molecular pathways involved in HU treatment, and a guide for future areas of research towards developing novel therapeutic approaches; (iv) In the same way, we developed a HU-induced post-transcriptional miRNAs-mediated pathway of HbF induction through silencing of potent negative regulators such as *MYB*, *KLF-1* and *BCL11A*, and in the process putting forth miRNAs as candidate targets for future therapeutic interventions in SCD; (v) The exploration of the effect of malaria on SCD-related polymorphisms will contribute to our understanding of the epidemiology of SCD in sub-Saharan Africa, particularly in and around the tropical equatorial malaria belt of central Africa; Lastly, the report of the expansion and distribution of the HbS allele in southern Africa, regions that were historically devoid of the mutation, will provide the necessary awareness to governments, healthcare

providers and teaching institutions to adapt health policies and practices to this increasing prevalence of SCD.

Conclusions and perspectives

The present study is the first to collectively address 3 pertinent components of SCD in sub-Saharan Africa: (i) the genomic modifiers of the persistence of HbF in adult patients from Cameroon; (ii) collate the current knowledge base of mechanisms of HU-induced HbF and unequivocally demonstrate a functional model of the miR-26b/*MYB* post-transcriptional signalling pathway for the induction of HbF by HU; and (iii) provide evidence of the increasing prevalence of the HbS allele in southern Africa, particularly South Africa, a country historically devoid of SCD. The sum of the work presented in this thesis has significantly contributed to our understanding of the genetics and treatment of SCD.

This work has also highlighted the clinical significance of the erythroid specific *BCL11A* enhancer and its account for variance in HbF levels. A review of the global distribution of β -globin haplotypes has yielded a provocative hypothesis of a single African origin of the HbS mutation and interrogation of the geographic nomenclature used in defining SCD haplotypes. Epigenetic and transcriptional events, signal transduction pathways and post-transcriptional pathways were highlighted as broad categories of reported molecular processes that involved HU induction of HbF. Furthermore, an umbilical cord blood-derived erythroid model was used to show HU-induced HbF production through down-regulation of negative regulators, *BCL11A*, *KLF-1* and *MYB*, by miRNA-mediated action with miR-26b-silencing of *MYB* central to HbF induction. Closer to home, the data presented in this thesis confirmed the evolutionary dynamics of various African genomics in response to selection pressure from malaria, and strongly prompted the dissociation of the assumed congruence between the genetics and cultural/linguistic/anthropological heritages of African populations. Furthermore, we provided evidence of the increasing burden of adult SCD patients in South Africa, cautioning national health institutions to adapt policies accordingly.

Taken in sum, the data presented is a comprehensive account of SCD in sub-Saharan Africa, comprising the following: components of SCD haplotype distribution, collation of the knowledge base on HU mechanisms of HbF induction; elucidation of a novel post-transcriptional miRNAs-mediated regulatory pathway for HbF induction by HU; interrogation of the evolutionary dynamics of African genomes and congruence between genetics and anthropological heritages; and a report on the expansion of SCD into southern Africa. The anticipated outcomes of this collective body of work are an expanded understanding of the effect of genomic modifiers on the disease-ameliorating HbF, the distribution of such polymorphisms across the continent and the emphasis on the potential role of post-transcriptional therapeutic targets in SCD treatment.

References (Introduction and Academic Discussion chapters)

- Adekile, A.D., Kitundu, M.N., Gu, L.H., Lanclos, K.D., Adeodu, O.O. & Huisman, T.H. 1992. Haplotypes in SS patients from Nigeria; characterization of one atypical beta S haplotype no. 19 (Benin) associated with elevated HbF and high G-gamma level. *Annals of Hematology*. 65(1):41-5.
- Aimola, I.A., Inuwa, H.M., Nok, A.J., Mamman, A.I. & Bieker, J.J. 2016. Cis-vaccenic acid induces differentiation and up-regulates Gamma globin synthesis in K562, JK1 and transgenic mice erythroid progenitor Stem Cells. *European Journal of Pharmacology*. 776:9-18.
- Alsultan, A., Aleem, A., Ghabbour, H., AlGahtani, F.H., Al-Shehri, A., Osman, M.E., Kurban, K., Alsultan, M.S. et al. 2012. Sick cell disease subphenotypes in patients from Southwestern Province of Saudi Arabia. *Journal of Pediatric hematology/oncology*. 34(2):79-84. DOI:10.1097/MPH.0b013e3182422844 [doi].
- Amoako, N., Asante, K.P., Adjei, G., Awandare, G.A., Bimi, L. & Owusu-Agyei, S. 2014. Associations between Red Cell Polymorphisms and Plasmodium falciparum Infection in the Middle Belt of Ghana. *PloS One*. 9(12):e112868.
- Balci, T.B., Sawyer, S.L., Davila, J., Humphreys, P. & Dymont, D.A. 2015. Brain malformations in a patient with deletion 2p16. 1: A refinement of the phenotype to BCL11A. *European Journal of Medical Genetics*. 58(6):351-354.
- Ballas, S.K., Kesen, M.R., Goldberg, M.F., Luty, G.A., Dampier, C., Osunkwo, I., Wang, W.C., Hoppe, C. et al. 2012. Beyond the definitions of the phenotypic complications of sickle cell disease: an update on management. *The Scientific World Journal*. 2012:949535.
- Bauer, D.E., Kamran, S.C., Lessard, S., Xu, J., Fujiwara, Y., Lin, C., Shao, Z., Canver, M.C. et al. 2013. An erythroid enhancer of BCL11A subject to genetic variation determines fetal hemoglobin level. *Science (New York, N.Y.)*. 342(6155):253-257. DOI:10.1126/science.1242088 [doi].
- Bauer, D.E. & Orkin, S.H. 2011. Update on fetal hemoglobin gene regulation in hemoglobinopathies. *Current Opinion in Pediatrics*. 23(1):1-8. DOI:10.1097/MOP.0b013e3283420fd0 [doi].
- Berthaut, I., Guignedoux, G., Kirsch-Noir, F., de Larouziere, V., Ravel, C., Bachir, D., Galacteros, F., Ancel, P.Y. et al. 2008. Influence of sickle cell disease and treatment with hydroxyurea on sperm parameters and fertility of human males. *Haematologica*. 93(7):988-993. DOI:10.3324/haematol.11515 [doi].

- Bitoungui, V.J.N., **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: A Journal of Integrative Biology*. 19(3):171-179.
- Brandow, A.M., Jirovec, D.L. & Panepinto, J.A. 2010. Hydroxyurea in children with sickle cell disease: practice patterns and barriers to utilization. *American Journal of Hematology*. 85(8):611-613.
- Brawley, O.W., Cornelius, L.J., Edwards, L.R., Gamble, V.N., Green, B.L., Inturrisi, C., James, A.H., Laraque, D. et al. 2008. National Institutes of Health Consensus Development Conference statement: hydroxyurea treatment for sickle cell disease. *Annals of Internal Medicine*. 148(12):932-938. DOI:0000605-200806170-00220 [pii].
- Brittenham, G.M., Schechter, A.N. & Noguchi, C.T. 1985. Hemoglobin S polymerization: primary determinant of the hemolytic and clinical severity of the sickling syndromes. *Blood*. 65(1):183-189.
- Bunn, H.F. 1997. Pathogenesis and treatment of sickle cell disease. *New England Journal of Medicine*. 337(11):762-769.
- Chaine, B., Neonato, M., Girot, R. & Aractingi, S. 2001. Cutaneous adverse reactions to hydroxyurea in patients with sickle cell disease. *Archives of Dermatology*. 137(4):467-470.
- Charache, S., Terrin, M.L., Moore, R.D., Dover, G.J., Barton, F.B., Eckert, S.V., McMahon, R.P. & Bonds, D.R. 1995. Effect of hydroxyurea on the frequency of painful crises in sickle cell anemia. *New England Journal of Medicine*. 332(20):1317-1322.
- Collins, F.S. & Orringer, E.P. 1982. Pulmonary hypertension and cor pulmonale in the sickle hemoglobinopathies. *The American Journal of Medicine*. 73(6):814-821.
- Colombatti, R., Perrotta, S., Samperi, P., Casale, M., Masera, N., Palazzi, G., Sainati, L. & Russo, G. 2013. Organizing national responses for rare blood disorders: the Italian experience with sickle cell disease in childhood. *Orphanet Journal of Rare Diseases*. 8(1):1.
- Creary, L.E., Ulug, P., Menzel, S., McKenzie, C.A., Hanchard, N.A., Taylor, V., Farrall, M., Forrester, T.E. et al. 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.

- Curat, M., Trabuchet, G., Rees, D., Perrin, P., Harding, R.M., Clegg, J.B., Langaney, A. & Excoffier, L. 2002. Molecular analysis of the β -globin gene cluster in the Niokholo Mandenka population reveals a recent origin of the β S Senegal mutation. *The American Journal of Human Genetics*. 70(1):207-223.
- Danjou, F., Zoledziewska, M., Sidore, C., Steri, M., Busonero, F., Maschio, A., Mulas, A., Perseu, L. et al. 2015. Genome-wide association analyses based on whole-genome sequencing in Sardinia provide insights into regulation of hemoglobin levels. *Nature Genetics*. 47(11):1264-71.
- DeBaun, M.R. 2014. Hydroxyurea therapy contributes to infertility in adult men with sickle cell disease: a review. *Expert Review of Hematology*. 7(6):767-773.
- Elderderly, A., Mills, J., Mohamed, B., Cooper, A., Mohammed, A., Eltieb, N. & Old, J. 2012. Molecular analysis of the β -globin gene cluster haplotypes in a Sudanese population with sickle cell anaemia. *International Journal of Laboratory Hematology*. 34(3):262-266.
- Elhassan, N., Gebremeskel, E.I., Elnour, M.A., Isabirye, D., Okello, J., Hussien, A., Kwiatkowski, D., Hirbo, J. et al. 2014. The episode of genetic drift defining the migration of humans out of Africa is derived from a large East African population size. *PLoS One*. 9(5):e97674.
- Elion, J., Berg, P.E., Lapoumeroulie, C., Trabuchet, G., Mittelman, M., Krishnamoorthy, R., Schechter, A.N. & Labie, D. 1992. DNA sequence variation in a negative control region 5' to the beta-globin gene correlates with the phenotypic expression of the beta s mutation. *Blood*. 79(3):787-792.
- Finotti, A., Bianchi, N., Fabbri, E., Borgatti, M., Breveglieri, G., Gasparello, J. & Gambari, R. 2015. Erythroid induction of K562 cells treated with mithramycin is associated with inhibition of raptor gene transcription and mammalian target of rapamycin complex 1 (mTORC1) functions. *Pharmacological Research*. 91:57-68.
- Flanagan, J.M., Howard, T.A., Mortier, N., Avlasevich, S.L., Smeltzer, M.P., Wu, S., Dertinger, S.D. & Ware, R.E. 2010. Assessment of genotoxicity associated with hydroxyurea therapy in children with sickle cell anemia. *Mutation Research/Genetic Toxicology and Environmental Mutagenesis*. 698(1):38-42.
- Fleming, A.F. 1989. The aetiology of severe anaemia in pregnancy in Ndola, Zambia. *Annals of Tropical Medicine and Parasitology*. 83(1):37-49.
- Flint, J., Harding, R.M., Boyce, A.J. & Clegg, J.B. 1998. 1 The population genetics of the haemoglobinopathies. *Baillière's Clinical Haematology*. 11(1):1-51.

- Flint, J., Hill, A.V., Bowden, D.K., Oppenheimer, S.J., Sill, P.R., Serjeantson, S.W., Bana-Koiri, J., Bhatia, K. et al. 1986. High frequencies of alpha-thalassaemia are the result of natural selection by malaria. *Nature*. 321(6072):744-750. DOI:10.1038/321744a0 [doi].
- Frenette, P.S. 2002. Sick cell vaso-occlusion: multistep and multicellular paradigm. *Current Opinion in Hematology*. 9(2):101-106.
- Funnell, A.P., Prontera, P., Ottaviani, V., Piccione, M., Giambona, A., Maggio, A., Ciaffoni, F., Stehling-Sun, S. et al. 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. DOI:10.1182/blood-2015-04-638528 [doi].
- Gaston, M.H., Verter, J.I., Woods, G., Pegelow, C., Kelleher, J., Presbury, G., Zarkowsky, H., Vichinsky, E. et al. 1986. Prophylaxis with oral penicillin in children with sickle cell anemia. *New England Journal of Medicine*. 314(25):1593-1599.
- Gibbons, C., Geoghegan, R., Conroy, H., Lippacott, S., O'Brien, D., Lynam, P., Langabeer, L., Cotter, M. et al. 2015. Sick cell disease: time for a targeted neonatal screening programme. *Irish Medical Journal*. 108(2):43-5.
- Gladwin, M.T., Kanas, T. & Kim-Shapiro, D.B. 2012. Hemolysis and cell-free hemoglobin drive an intrinsic mechanism for human disease. *The Journal of Clinical Investigation*. 122(4):1205-1208. DOI:10.1172/JCI62972 [doi].
- Habibi, A., Arlet, J.B., Stankovic, K., Gellen-Dautremer, J., Ribeil, J.A., Bartolucci, P., Lionnet, F. & centre de reference maladies rares syndromes drepanocytaires majeurs. 2015. French guidelines for the management of adult sickle cell disease: 2015 update. *La Revue De Medecine Interne / Fondée ...Par La Societe Nationale Francaise De Medecine Interne*. 36(5 Suppl 1):5S3-84. DOI:10.1016/S0248-8663(15)60002-9 [doi].
- Hahn, C.K. & Lowrey, C.H. 2014. Induction of fetal hemoglobin through enhanced translation efficiency of gamma-globin mRNA. *Blood*. 124(17):2730-2734. DOI:10.1182/blood-2014-03-564302 [doi].
- Hampton, T. 2008. Strides made in red blood cell disorders, but substantial barriers to care remain. *JAMA*. 299(4):395-396.
- Hankins, J.S., Ware, R.E., Rogers, Z.R., Wynn, L.W., Lane, P.A., Scott, J.P. & Wang, W.C. 2005. Long-term hydroxyurea therapy for infants with sickle cell anemia: the HUSOFT extension study. *Blood*. 106(7):2269-2275. DOI:106/7/2269 [pii].

- Haywood, C., Jr, Beach, M.C., Lanzkron, S., Strouse, J.J., Wilson, R., Park, H., Witkop, C., Bass, E.B. et al. 2009. A systematic review of barriers and interventions to improve appropriate use of therapies for sickle cell disease. *Journal of the National Medical Association*. 101(10):1022-1033.
- Herrick, J.B. 2001. Peculiar elongated and sickle-shaped red blood corpuscles in a case of severe anemia. 1910. *The Yale Journal of Biology and Medicine*. 74(3):179-184.
- Hirst, C. & Owusu-Ofori, S. 2012. Prophylactic antibiotics for preventing pneumococcal infection in children with sickle cell disease. *The Cochrane Library*. 12(9):CD003427. DOI: 10.1002/14651858.CD003427.
- Jenabian, M., Talla, F., Talla, P., Mbopi-Kéou, F., Charpentier, C., Kane, C.T. & Bélec, L. 2015. Pitfalls of antiretroviral drug resistance genotyping of HIV-1 Group M and Group N from Cameroon by sequenced-based assays. *Nigerian Medical Journal*. 56(6):420.
- Kinney, T.R., Helms, R.W., O'Branski, E.E., Ohene-Frempong, K., Wang, W., Daeschner, C., Vichinsky, E., Redding-Lallinger, R. et al. 1999. Safety of hydroxyurea in children with sickle cell anemia: results of the HUG-KIDS study, a phase I/II trial. Pediatric Hydroxyurea Group. *Blood*. 94(5):1550-1554.
- Kinomoto, M., Appiah-Opong, R., Brandful, J.A., Yokoyama, M., Nii-Trebi, N., Ugly-Kwame, E., Sato, H., Ofori-Adjei, D. et al. 2005. HIV-1 proteases from drug-naive West African patients are differentially less susceptible to protease inhibitors. *Clinical Infectious Diseases : An Official Publication of the Infectious Diseases Society of America*. 41(2):243-251. DOI:CID35241 [pii].
- Kumkhaek, C., Taylor JG, 6., Zhu, J., Hoppe, C., Kato, G.J. & Rodgers, G.P. 2008. Fetal haemoglobin response to hydroxycarbamide treatment and sar1a promoter polymorphisms in sickle cell anaemia. *British Journal of Haematology*. 141(2):254-259. DOI:10.1111/j.1365-2141.2008.07045.x [doi].
- Labie, D., Pagnier, J., Lapoumeroulie, C., Rouabhi, F., Dunda-Belkhodja, O., Chardin, P., Beldjord, C., Wajcman, H. et al. 1985. Common haplotype dependency of high G gamma-globin gene expression and high Hb F levels in beta-thalassemia and sickle cell anemia patients. *Proceedings of the National Academy of Sciences of the United States of America*. 82(7):2111-2114.
- Lanzkron, S., Haywood, C., Jr, Hassell, K.L. & Rand, C. 2008. Provider barriers to hydroxyurea use in adults with sickle cell disease: a survey of the Sickle Cell Disease Adult Provider Network. *Journal of the National Medical Association*. 100(8):968-973.
- Lebensburger, J.D., Sidonio, R.F., DeBaun, M.R., Safford, M.M., Howard, T.H. & Scarinci, I.C. 2013. Exploring barriers and facilitators to clinical trial enrollment in

the context of sickle cell anemia and hydroxyurea. *Pediatric Blood & Cancer*. 60(8):1333-1337.

Lebensburger, J., Johnson, S.M., Askenazi, D.J., Rozario, N.L., Howard, T.H. & Hilliard, L.M. 2011. Protective role of hemoglobin and fetal hemoglobin in early kidney disease for children with sickle cell anemia. *American Journal of Hematology*. 86(5):430-432. DOI:10.1002/ajh.21994 [doi].

Lebensburger, J.D., Pestina, T.I., Ware, R.E., Boyd, K.L. & Persons, D.A. 2010. Hydroxyurea therapy requires HbF induction for clinical benefit in a sickle cell mouse model. *Haematologica*. 95(9):1599-1603. DOI:10.3324/haematol.2010.023325 [doi].

Lee, A., Thomas, P., Cupidore, L., Serjeant, B. & Serjeant, G. 1995. Improved survival in homozygous sickle cell disease: lessons from a cohort study. *BMJ (Clinical Research Ed.)*. 311(7020):1600-1602.

Lettre, G., Sankaran, V.G., Bezerra, M.A., Araujo, A.S., Uda, M., Sanna, S., Cao, A., Schlessinger, D. et al. 2008. DNA polymorphisms at the BCL11A, HBS1L-MYB, and beta-globin loci associate with fetal hemoglobin levels and pain crises in sickle cell disease. *Proceedings of the National Academy of Sciences of the United States of America*. 105(33):11869-11874. DOI:10.1073/pnas.0804799105 [doi].

Lobo, C.L., Pinto, J.F., Nascimento, E.M., Moura, P.G., Cardoso, G.P. & Hankins, J.S. 2013. The effect of hydroxycarbamide therapy on survival of children with sickle cell disease. *British Journal of Haematology*. 161(6):852-860. DOI:10.1111/bjh.12323 [doi].

Maier-Redelsperger, M., de Montalembert, M., Flahault, A., Neonato, M.G., Ducrocq, R., Masson, M.P., Girot, R. & Elion, J. 1998. Fetal hemoglobin and F-cell responses to long-term hydroxyurea treatment in young sickle cell patients. The French Study Group on Sickle Cell Disease. *Blood*. 91(12):4472-4479.

Makani, J., Menzel, S., Nkya, S., Cox, S.E., Drasar, E., Soka, D., Komba, A.N., Mgaya, J. et al. 2011. Genetics of fetal hemoglobin in Tanzanian and British patients with sickle cell anemia. *Blood*. 117(4):1390-1392. DOI:10.1182/blood-2010-08-302703 [doi].

Menzel, S., Garner, C., Gut, I., Matsuda, F., Yamaguchi, M., Heath, S., Foglio, M., Zelenika, D. et al. 2007. A QTL influencing F cell production maps to a gene encoding a zinc-finger protein on chromosome 2p15. *Nature Genetics*. 39(10):1197-1199.

Miller, S.T., Macklin, E.A., Pegelow, C.H., Kinney, T.R., Sleeper, L.A., Bello, J.A., DeWitt, L.D., Gallagher, D.M. et al. 2001. Silent infarction as a risk factor for overt

- stroke in children with sickle cell anemia: a report from the Cooperative Study of Sickle Cell Disease. *The Journal of Pediatrics*. 139(3):385-390.
- Miller, S.T., Sleeper, L.A., Pegelow, C.H., Enos, L.E., Wang, W.C., Weiner, S.J., Wethers, D.L., Smith, J. et al. 2000. Prediction of adverse outcomes in children with sickle cell disease. *New England Journal of Medicine*. 342(2):83-89.
- Modell, B. & Darlison, M. 2008. Global epidemiology of haemoglobin disorders and derived service indicators. *Bulletin of the World Health Organization*. 86(6):480-487.
- Mtatiro, S.N., Mgaya, J., Singh, T., Mariki, H., Rooks, H., Soka, D., Mmbando, B., Thein, S.L. et al. 2015. Genetic association of fetal-hemoglobin levels in individuals with sickle cell disease in Tanzania maps to conserved regulatory elements within the MYB core enhancer. *BMC Medical Genetics*. 16(1):1.
- Mtatiro, S.N., Singh, T., Rooks, H., Mgaya, J., Mariki, H., Soka, D., Mmbando, B., Msaki, E. et al. 2014. Genome wide association study of fetal hemoglobin in sickle cell anemia in Tanzania. *PloS One*. 9(11):e111464.
- Ohene-Frempong, K., Weiner, S.J., Sleeper, L.A., Miller, S.T., Embury, S., Moohr, J.W., Wethers, D.L., Pegelow, C.H. et al. 1998. Cerebrovascular accidents in sickle cell disease: rates and risk factors. *Blood*. 91(1):288-294.
- Owusu, E.D., Visser, B.J., Nagel, I.M., Mens, P.F. & Grobusch, M.P. 2015. The interaction between sickle cell disease and HIV infection: a systematic review. *Clinical Infectious Diseases : An Official Publication of the Infectious Diseases Society of America*. 60(4):612-626. DOI:10.1093/cid/ciu832 [doi].
- Oyeku, S.O., Driscoll, M.C., Cohen, H.W., Trachtman, R., Pashankar, F., Mullen, C., Giardina, P.J., Velazco, N. et al. 2013. Parental and other factors associated with hydroxyurea use for pediatric sickle cell disease. *Pediatric Blood & Cancer*. 60(4):653-658.
- 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. *Proceedings of the National Academy of Sciences of the United States of America*. 81(6):1771-1773.
- Peter, B., Matsushita, M., Oda, K. & Raskind, W. 2014. De novo microdeletion of BCL11A is associated with severe speech sound disorder. *American Journal of Medical Genetics Part A*. 164(8):2091-2096.
- Pizzo, E., Laverty, A.A., Phekoo, K.J., AlJuburi, G., Green, S.A., Bell, D. & Majeed, A. 2015. A retrospective analysis of the cost of hospitalizations for sickle cell disease with crisis in England, 2010/11. *Journal of Public Health (Oxford, England)*. 37(3):529-539. DOI:10.1093/pubmed/fdu026 [doi].

- Platt, O.S., Brambilla, D.J., Rosse, W.F., Milner, P.F., Castro, O., Steinberg, M.H. & Klug, P.P. 1994. Mortality in sickle cell disease. Life expectancy and risk factors for early death. *The New England Journal of Medicine*. 330(23):1639-1644. DOI:10.1056/NEJM199406093302303 [doi].
- Platt, O.S., Orkin, S.H., Dover, G., Beardsley, G.P., Miller, B. & Nathan, D.G. 1984. Hydroxyurea enhances fetal hemoglobin production in sickle cell anemia. *The Journal of Clinical Investigation*. 74(2):652-656. DOI:10.1172/JCI111464 [doi].
- Pule, G.D.**, Ngo Bitoungui, V.J., Chetcha Chemegni, B., Kengne, A.P., & Wonkam, A. 2016. Studies of novel variants associated with HbF in Sardinians and Tanzanians in Sickle Cell Disease patients from Cameroon. *Hemoglobin*. 10:1-4.
- Pule G.D.**, Mnika K, Jourbert M, Mowla S, Novitzky N, Wonkam A. 2016. Increasing burden of adult patients living with Sickle Cell Disease in Cape Town. Submitted to *South African Medical Journal (in press, #10849)*.
- Pule, G.D.**, Mowla, S., Novitzky, N. & Wonkam, A. 2016. Hydroxyurea down-regulates BCL11A, KLF-1 and MYB through miRNA-mediated actions to induce γ -globin expression: implications for new therapeutic approaches of sickle cell disease. *Clinical and Translational Medicine*. 5(1):1.
- Pule G.D.**, Mowla S, Novitzky N, Wiysonge C and Wonkam A. Systematic review of known mechanisms of hydroxyurea-induced fetal hemoglobin for treatment of Sickle Cell Disease. *Expert Review of Hematology*. 2015 8(5): 669-79.
- Pule, G.D.**, Ngo Bitoungui, V.J., Chetcha Chemegni, B., Kengne, A.P., 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: A Journal of Integrative Biology*. 19(10):627-631.
- Pule, G.D.** & Wonkam, A. 2014. Treatment for sickle cell disease in Africa: should we invest in haematopoietic stem cell transplantation? *The Pan African Medical Journal*. 18:46. DOI:10.11604/pamj.2014.18.46.3923 [doi].
- Quinn, C.T., Rogers, Z.R. & Buchanan, G.R. 2004. Survival of children with sickle cell disease. *Blood*. 103(11):4023-4027. DOI:10.1182/blood-2003-11-3758 [doi].
- Rahimy, M.C., Gangbo, A., Ahouignan, G. & Alihonou, E. 2009. Newborn screening for sickle cell disease in the Republic of Benin. *Journal of Clinical Pathology*. 62(1):46-48. DOI:10.1136/jcp.2008.059113 [doi].
- Rose, J.I., Usik, V.I., Marks, A.E., Hilbert, Y.H., Galletti, C.S., Parton, A., Geiling, J.M., Černý, V. et al. 2011. The Nubian complex of Dhofar, Oman: an African middle stone age industry in southern Arabia. *PLoS One*. 6(11):e28239.

- Sankaran, V.G., Menne, T.F., Scepanovic, D., Vergilio, J.A., Ji, P., Kim, J., Thiru, P., Orkin, S.H. et al. 2011. MicroRNA-15a and -16-1 act via MYB to elevate fetal hemoglobin expression in human trisomy 13. *Proceedings of the National Academy of Sciences of the United States of America*. 108(4):1519-1524. DOI:10.1073/pnas.1018384108 [doi].
- Sebastiani, P., Farrell, J., Alsultan, A., Wang, S., Edward, H., Shappell, H., Bae, H., Milton, J. et al. 2015. BCL11A enhancer haplotypes and fetal hemoglobin in sickle cell anemia. *Blood Cells, Molecules, and Diseases*. 54(3):224-230.
- Segal, J.B., Strouse, J.J., Beach, M.C., Haywood, C., Witkop, C., Park, H., Wilson, R.F., Bass, E.B. et al. 2008. Hydroxyurea for the treatment of sickle cell disease. *Evidence report/technology Assessment*. (165)(165):1-95.
- Serjeant, G.R. 2001. The emerging understanding of sickle cell disease. *British Journal of Haematology*. 112(1):3-18.
- Serjeant, G.R., Hambleton, I. & Thame, M. 2005. Fecundity and pregnancy outcome in a cohort with sickle cell-haemoglobin C disease followed from birth. *BJOG: An International Journal of Obstetrics & Gynaecology*. 112(9):1308-1314.
- Shemisa, K., Jafferjee, N., Thomas, D., Jacobs, G. & Meyerson, H.J. 2014. Mycobacterium avium Complex Infection in a Patient with Sickle Cell Disease and Severe Iron Overload. *Case Reports in Infectious Diseases*. 2014:405323. DOI:10.1155/2014/405323 [doi].
- Smith-Whitley, K. 2014. Reproductive issues in sickle cell disease. *Blood*. 124(24):3538-3543. DOI:10.1182/blood-2014-07-577619 [doi].
- Solovieff, N., Milton, J.N., Hartley, S.W., Sherva, R., Sebastiani, P., Dworkis, D.A., Klings, E.S., Farrer, L.A. 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(9):1815-1822. DOI:10.1182/blood-2009-08-239517 [doi].
- Steinberg, M.H., Barton, F., Castro, O., Pegelow, C.H., Ballas, S.K., Kutlar, A., Orringer, E., Bellevue, R. et al. 2003. Effect of hydroxyurea on mortality and morbidity in adult sickle cell anemia: risks and benefits up to 9 years of treatment. *Jama*. 289(13):1645-1651.
- Steinberg, M.H. 2009. Genetic etiologies for phenotypic diversity in sickle cell anemia. *TheScientificWorldJournal*. 9:46-67. DOI:10.1100/tsw.2009.10 [doi].
- Strouse, J.J. & Heeney, M.M. 2012. Hydroxyurea for the treatment of sickle cell disease: efficacy, barriers, toxicity, and management in children. *Pediatric Blood & Cancer*. 59(2):365-371.

- Strouse, J.J., Lanzkron, S., Beach, M.C., Haywood, C., Park, H., Witkop, C., Wilson, R.F., Bass, E.B. et al. 2008. Hydroxyurea for sickle cell disease: a systematic review for efficacy and toxicity in children. *Pediatrics*. 122(6):1332-1342. DOI:10.1542/peds.2008-0441 [doi].
- Telfer, P., Coen, P., Chakravorty, S., Wilkey, O., Evans, J., Newell, H., Smalling, B., Amos, R. et al. 2007. Clinical outcomes in children with sickle cell disease living in England: a neonatal cohort in East London. *Haematologica*. 92(7):905-912.
- Thein, S.L., Menzel, S., Lathrop, M. & Garner, C. 2009. Control of fetal hemoglobin: new insights emerging from genomics and clinical implications. *Human Molecular Genetics*. 18(R2):R216-23. DOI:10.1093/hmg/ddp401 [doi].
- Thornburg, C.D., Calatroni, A., Telen, M. & Kemper, A.R. 2010. Adherence to hydroxyurea therapy in children with sickle cell anemia. *The Journal of Pediatrics*. 156(3):415-419.
- Thornburg, C.D., Dixon, N., Burgett, S., Mortier, N.A., Schultz, W.H., Zimmerman, S.A., Bonner, M., Hardy, K.K. et al. 2009. A pilot study of hydroxyurea to prevent chronic organ damage in young children with sickle cell anemia. *Pediatric Blood & Cancer*. 52(5):609-615.
- Timmann, C., Thye, T., Vens, M., Evans, J., May, J., Ehmen, C., Sievertsen, J., Muntau, B. et al. 2012. Genome-wide association study indicates two novel resistance loci for severe malaria. *Nature*. 489(7416):443-446.
- Trakarnsanga, K., Wilson, M.C., Lau, W., Singleton, B.K., Parsons, S.F., Sakuntanaga, P., Kurita, R., Nakamura, Y. et al. 2014. Induction of adult levels of beta-globin in human erythroid cells that intrinsically express embryonic or fetal globin by transduction with KLF1 and BCL11A-XL. *Haematologica*. 99(11):1677-1685. DOI:10.3324/haematol.2014.110155 [doi].
- Vichinsky, E.P. 1991. Comprehensive care in sickle cell disease: its impact on morbidity and mortality. *Seminars in Hematology*. 28(3):220-226.
- Voskaridou, E., Christoulas, D., Bilalis, A., Plata, E., Varvagiannis, K., Stamatopoulos, G., Sinopoulou, K., Balassopoulou, A. et al. 2010. The effect of prolonged administration of hydroxyurea on morbidity and mortality in adult patients with sickle cell syndromes: results of a 17-year, single-center trial (LaSHS). *Blood*. 115(12):2354-2363. DOI:10.1182/blood-2009-05-221333 [doi].
- Wang, W. & Thompson, B. 2010. Hydroxyurea treatment of infants with sickle cell anemia: results of the BABY HUG study. *Pediatric Blood & Cancer*. Wiley-Liss Div John Wiley & Sons Inc. 111 River St. Hoboken, NJ 07030 USA. 787.

- Ware, R.E., Despotovic, J.M., Mortier, N.A., Flanagan, J.M., He, J., Smeltzer, M.P., Kimble, A.C., Aygun, B. et al. 2011. Pharmacokinetics, pharmacodynamics, and pharmacogenetics of hydroxyurea treatment for children with sickle cell anemia. *Blood*. 118(18):4985-4991. DOI:10.1182/blood-2011-07-364190 [doi].
- Weatherall, D.J. 2008. Hemoglobinopathies worldwide: present and future. *Current Molecular Medicine*. 8(7):592-599.
- Weatherall, D., Hofman, K., Rodgers, G., Ruffin, J. & Hrynkow, S. 2005. A case for developing North-South partnerships for research in sickle cell disease. *Blood*. 105(3):921-923. DOI:2004-06-2404 [pii].
- Weatherall, D.J. 2010. The inherited diseases of hemoglobin are an emerging global health burden. *Blood*. 115(22):4331-4336. DOI:10.1182/blood-2010-01-251348 [doi].
- Williams, T.N., Mwangi, T.W., Wambua, S., Alexander, N.D., Kortok, M., Snow, R.W. & Marsh, K. 2005. Sickle cell trait and the risk of Plasmodium falciparum malaria and other childhood diseases. *The Journal of Infectious Diseases*. 192(1):178-186. DOI:JID34144 [pii].
- Wonkam, A., Bitoungui, V.J.N., Vorster, A.A., Ramesar, R., Cooper, R.S., 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.
- Wonkam, A., Ponde, C., Nicholson, N., Fieggen, K., Ramesar, R. & Davidson, A. 2012. The burden of sickle cell disease in Cape Town. *South African Medical Journal*. 102(9):753-754.
- Zhu, J., Chin, K., Aerbajinai, W., Kumkhaek, C., Li, H. & Rodgers, G.P. 2014. Hydroxyurea-inducible SAR1 gene acts through the G1alpha/JNK/Jun pathway to regulate gamma-globin expression. *Blood*. 124(7):1146-1156. DOI:10.1182/blood-2013-10-534842 [doi].
- Zimmerman, S.A., Schultz, W.H., Burgett, S., Mortier, N.A. & Ware, R.E. 2007. Hydroxyurea therapy lowers transcranial Doppler flow velocities in children with sickle cell anemia. *Blood*. 110(3):1043-1047. DOI:blood-2006-11-057893 [pii].
- Zimmerman, S.A., Schultz, W.H., Davis, J.S., Pickens, C.V., Mortier, N.A., Howard, T.A. & Ware, R.E. 2004. Sustained long-term hematologic efficacy of hydroxyurea at maximum tolerated dose in children with sickle cell disease. *Blood*. 103(6):2039-2045. DOI:10.1182/blood-2003-07-2475 [doi].

Zumberg, M.S., Reddy, S., Boyette, R.L., Schwartz, R.J., Konrad, T.R. & Lottenberg, R. 2005. Hydroxyurea therapy for sickle cell disease in community-based practices: A survey of Florida and North Carolina hematologists/oncologists. *American Journal of Hematology*. 79(2):107-113.

Zur, B. 2016. Increase in genetically determined anemia as a result of migration in Germany. *Der Internist*. 57(5):444-51.

Search

Fol...

Inbox

Archive

Junk

Drafts

Sent

Deleted

Search R..

Ne...

RE: Application to include publications



Janine Isaacs (janine.isaacs@uct.ac.za) [Add to contacts](#) 2

To: Gift Pule Cc: Adri Winckler

Dear Mr Pule

I confirm that your application/request to include publications was approved by the DDB Chair Professor Visser.

Regards

Janine

From: Gift Pule [mailto:giftpule@outlook.com]

Sent: Wednesday, June 29, 2016 9:30 AM

To: Janine Isaacs <janine.isaacs@uct.ac.za>

Cc: Adri Winckler <adri.winckler@uct.ac.za>

Subject: RE: Application to include publications

Hi Janine

Could you kindly amend the wording on that email.

...inclusion of publications, not "...applications..."As noted currently.

© 2016 Microsoft Terms Privacy & cookies Developers



Hydroxyurea induces γ -globin expression through microRNAs-mediated actions in human hematopoietic and K562 erythroleukemia cells.

D.G. Pule¹, S. Mowla², N. Novitzky², A. Wonkam¹.

1) Human Genetics, Clinical Laboratory Sciences, University of Cape Town, Cape Town, Western Cape, South Africa; 2) Hematology, Clinical Laboratory Sciences, University of Cape Town, Cape Town, Western Cape, South Africa.

Background: Hydroxyurea is the only available drug treatment for Sickle Cell Disease (SCD); but the mechanism of fetal hemoglobin (HbF) induction is poorly understood. In particular, the role of microRNAs in the post-transcriptional regulation of HbF, through potent regulators such as *MYB*, *BCL11A* and *KLF-1*, is understudied.

Objectives: We have investigated the role of microRNAs in the post-transcriptional regulation of HbF in erythroid cells treated with hydroxyurea.

Methods: As models, we used 1) hematopoietic stem cells derived from umbilical cord blood (HSCs) and 2) the K562 human erythroleukemia cell line. Both primary erythroid and K562 cells were treated with hydroxyurea and changes in *BCL11A*, *KLF-1*, *GATA-1*, *MYB*, β - and γ -globin gene expression and 7 targeted miRNAs, previously shown to be modified by hydroxyurea or associated with basal γ -globin expression were also investigated.

Results: We differentiated CD34⁺ stem cells into erythroid precursors using a single-phase expansion and differentiation protocol. *BCL11A* was down-regulated and there was a marked 7-fold increase ($p < 0.003$) in γ -globin expression in both primary human erythroid and K562 cells. Down-regulation of *GATA-1* was also consistent with the above findings. Similarly, *KLF-1* was down-regulated in both cell models, corresponding to the repressed expression of *BCL11A* and β -globin gene ($p < 0.04$). In both cell models, most miRNAs was significantly up-regulated by hydroxyurea with some level of variation. Down-regulation of, another potent negative regulator of γ -globin and direct target of the up-regulated miR-15a and miR-16-1, provides a post-transcriptional tier of HbF regulation. We also demonstrated that the inhibition of miR-26b; miR-151-3p and miR-451 resulted in an apparent decrease in HbF.

Conclusions: The data provide evidence of a microRNAs-mediated post-transcriptional mechanism for γ -globin regulation under hydroxyurea treatment. These findings will add a unique piece to our understanding of HbF regulation and incite further investigation of the post-transcriptional regulatory network for HbF expression.

Website publish date: September 8, 2015

This certifies that Gift Pule attended and presented the abstract below at the 2015 Annual Meeting of the American Society of Human Genetics held at the Baltimore Convention Center in Baltimore, Maryland over the dates of October 6-10, 2015.

ASHG Meeting Staff
475T

07 March 2016

Turnitin Reports - Home

Turnitin

plxdino03:Thesis-Compiled.pdf
By Gift Pule

Processed On: 02-Aug-2016 14:08:54ST
ID: 491290428
Word Count: 32046
Submitted: 1

Originality Report

Document Viewer

Similarity Index: 56%

Similarity by Source	
Internet Sources:	30%
Publications:	53%
Student Papers:	23%

Study of Genetic Modifiers of Fetal Hemoglobin and Mechanisms of
Hydroxyurea-induced γ -globin Expression in Sickle Cell Disease 32

By: Gift Dineo Pule PLXDIN003

SUBMITTED TO THE UNIVERSITY OF CAPE TOWN in fulfillment of the requirements of the degree PhD in Human Genetics Faculty of Health Sciences, Department of Pathology, Division of 7

Human Genetics Supervisor: Professor Ambrose Workam Co-supervisors: Dr Shaheen Mowla and Professor Nicolas Novitsky Date of submission: 02 August 2016 Declaration I, Gift Pule,

hereby declare that the work on which this dissertation/thesis is based is my original work (except where acknowledgements indicate otherwise) and that neither the whole work nor any part of it has been, is being, or is to be submitted for another degree in this or any other university. I empower the university to reproduce, for the purpose of research, either the whole or any portion of the contents, in any manner whatsoever. Signature: _____ Date: 02 August 2016 Acknowledgements I would like to express my utmost gratitude, To my parents, Joseph and Margaret, for their unconditional love and 10

support from elementary school to this very moment, To my siblings, Orpheus, Margo and Bossa, for their encouragement and confidence

Turnitin_Originality_...html

Show all downloads

1 27% match (publications)
Bitoungui, Valentina J., Ngo, Gift D. Pule, Neil Hanchard, Jeanne Ngogang, and Ambrose Workam. "Beta-Globin Gene Haplotypes Among Cameroonians and Review of the Global Distribution: Is There a Case for a Single Sickle Mutation Origin in Africa?". *Omicron: A Journal of Integrative Biology*, 2015.

2 16% match (publications)
Pule, Gift Dineo, Valentina Josiane Ngo Bitoungui, Bernard Chetcha Chemegni, Andre Pascal Kengne, Stylianos Antonarakis, and Ambrose Workam. "Association between Variants at BCL11A Erythrocytic Specific Enhancer and Fetal Hemoglobin Levels among Sickle Cell Disease Patients in Cameroon: Implications for Future Therapeutic Interventions". *Omicron: A Journal of Integrative Biology*, 2015.

3 3% match (Internet from 28-Jul-2016)
[https://clintramed.springeropen.com/article/10.1186-s12916-016-0092-7](https://clintramed.springeropen.com/article/10.1186/s12916-016-0092-7)