

Extensive Genomic Variability of Knops Blood Group Polymorphisms Is Associated with Sickle Cell Disease in Africa

Kimberley C. Duru¹, Jenelle A. Noble¹, Aldiouma Guindo², Li Yi³, Ikhide G. Imumorin⁴, Dapa A. Diallo² and Bolaji N. Thomas¹

¹Department of Biomedical Sciences, College of Health Sciences and Technology, Rochester Institute of Technology, Rochester, NY, USA. ²Centre de Recherche et de Lutte contre la Drepanocytose (CRLD), Bamako, Mali. ³School of Statistics, Shanxi University of Finance and Economics, Shanxi, China. ⁴Animal Genetics and Genomics Lab, Office of International Programs, College of Agriculture and Life Sciences, Cornell University, Ithaca, NY, USA.

ABSTRACT: Sickle cell disease (SCD) is a multisystem disorder characterized by chronic hemolytic anemia, vaso-occlusive crises, and marked variability in disease severity. Patients require transfusions to manage disease complications, with complements, directed by complement regulatory genes (CR1) and its polymorphisms, implicated in the development of alloantibodies. We hypothesize that CR1 polymorphisms affect complement regulation and function, leading to adverse outcome in SCD. To this end, we determined the genomic diversity of complement regulatory genes by examining single nucleotide polymorphisms associated with Knops blood group antigens. Genomic DNA samples from 130 SCD cases and 356 control Africans, 331 SCD cases and 497 control African Americans, and 254 Caucasians were obtained and analyzed, utilizing a PCR–RFLP (polymerase chain reaction–restriction fragment length polymorphism) assay. Analyzing for ethnic diversity, we found significant differences in the genotypic and allelic frequencies of S11/S12 (rs17047661) and McCa/b (rs17047660) polymorphisms between Africans, African Americans, and Caucasians ($P < 0.05$). The homozygote mutant variants had significantly higher frequencies in Africans and African Americans but were insignificant in Caucasians (80.2% and 59.6% vs 5.9% for S11/2; and 36% and 24% vs 1.8% for McCa/b). With SCD, we did not detect any difference among cases and controls either in Africa or in the United States. However, we found significant difference in genotypic ($P < 0.0001$) and allelic frequencies ($P < 0.0001$) of S11/S12 (rs17047661) and McCa/b (rs17047660) polymorphisms between SCD groups from Africa and the United States. There was no difference in haplotype frequencies of these polymorphisms among or between groups. The higher frequency of CR1 homozygote mutant variants in Africa but not United States indicates a potential pathogenic role, possibly associated with complicated disease pathophysiology in the former and potentially protective in the latter. The difference between sickle cell groups suggests potential genetic drift or founder effect imposed on the disease in the United States, but not in Africa, and a possible confirmation of the ancestral susceptibility hypothesis. The lower haplotype frequencies among sickle cell and control populations in the United States may be due to the admixture and the dilution of African genetic ancestry in the African American population.

KEYWORDS: complement receptor 1, sickle cell disease, polymorphisms, pathophysiology, Africa, United States, evolution, ancestry

CITATION: Duru et al. Extensive Genomic Variability of Knops Blood Group Polymorphisms Is Associated with Sickle Cell Disease in Africa. *Evolutionary Bioinformatics* 2015;11 25–33 doi: 10.4137/EBO.S23132.

RECEIVED: December 19, 2014. **RESUBMITTED:** January 27, 2015. **ACCEPTED FOR PUBLICATION:** February 02, 2015.

ACADEMIC EDITOR: Jike Cui, Associate Editor

TYPE: Original Research

FUNDING: College of Health Sciences and Technology, Rochester Institute of Technology (RIT), provided the funding for this project (BNT). Additional support to JAN by the RIT Louis Stokes Alliance for Minority Participation (NSF Award #25901-03200-S01) and College of Health Sciences and Technology Summer Undergraduate Research Fellow is also acknowledged. The authors confirm that the funder had no influence over the study design, content of the article, or selection of this journal.

COMPETING INTERESTS: Authors disclose no potential conflicts of interest.

CORRESPONDENCE: bolaji.thomas@rit.edu

COPYRIGHT: © the authors, publisher and licensee Libertas Academica Limited. This is an open-access article distributed under the terms of the Creative Commons CC-BY-NC 3.0 License.

Paper subject to independent expert blind peer review by minimum of two reviewers. All editorial decisions made by independent academic editor. Upon submission manuscript was subject to anti-plagiarism scanning. Prior to publication all authors have given signed confirmation of agreement to article publication and compliance with all applicable ethical and legal requirements, including the accuracy of author and contributor information, disclosure of competing interests and funding sources, compliance with ethical requirements relating to human and animal study participants, and compliance with any copyright requirements of third parties. This journal is a member of the Committee on Publication Ethics (COPE).

Published by Libertas Academica. Learn more about this journal.

Introduction

Sickle cell disease (SCD) is a complex multisystem disorder that presents with variability in disease pathophysiology and occurs in patients homozygous for the hemoglobin S (HbS) gene. It has a worldwide distribution^{1–4} with differing disease haplotypes, and is largely found in sub-Saharan Africa (Senegal, Benin, Bantu, and Cameroon haplotypes).^{5,6} The Arab-Indian haplotype is found in the Middle East, India, and parts of the Mediterranean.^{7,8} Among the African American population in the United States and populations in the Caribbean, SCD is commonly encountered with multiple and diverse disease haplotypes and clearly distinct and different from those in sub-Saharan Africa,⁹ probably due to generations of admixture and

genetic heterogeneity.⁶ Oxygen is a major factor determining the degree of disease pathophysiology and the formation of the sickle-shaped red blood cells. These abnormally shaped cells obstruct blood flow, causing self-destruction of one's own red cells and the associated anemia.¹⁰ Treating anemia and other disease complications require multiple transfusions, alongside the concomitant alloimmunization, which has been reported in some patient groups.¹¹ Biomarkers that are predictive of susceptibility to alloimmunization before the initiation of a transfusion regimen have been identified,¹² with complements potentially implicated in this process. It is imperative therefore to delineate the mechanism driving their activation and the pathway that directs their participation.



Complement receptor-1 (CR1) is a polymorphic glycoprotein, expressed predominantly on red cells, mainly functioning in the regulation of complement activity and the binding of C3b and C4b found on circulating immune complexes (CICs). Functional impairment of erythrocyte CR1 can interfere with the efficient processing of CICs, whose concentration in circulation is dependent on disease severity.¹³ Due to its involvement in immune complex clearance, it has been associated with the severity of diseases such as malaria and systemic lupus erythematosus.^{13–16} CICs' importance and their relationship to a variety of diseases, including sickle cell pathophysiology as well as intra- and interethnic delineation of disease haplotypes and phenotypes is therefore of great interest. Immune complex formation, under normal circumstances, is a process designed to facilitate normal functioning of the immune system, leading to the removal of CICs from peripheral circulation, which ultimately fails in some disease states, thereby initiating complement-mediated damage. Effective clearance of immune complexes, therefore, can be hampered by CR1 polymorphisms,^{17–20} which differ significantly between Caucasians and Africans and are shown to be associated with susceptibility to malaria infection or disease severity.^{21–24}

Published reports have challenged the degree of protection conferred by Knops blood group polymorphisms in malaria infection.^{16,25,26} The co-evolution of malaria and SCD has been established, as demonstrated by reports showing that there are wide differences in *Plasmodium falciparum* infection rates and multiplicity of infection between carriers of the sickle cell trait (hemoglobin AS) and those with normal hemoglobin.^{27–31} Therefore, the possibility exists of increased selection pressure through single nucleotide polymorphisms (SNPs) that can exacerbate or ameliorate disease outcome.^{13,32} We believe that such protective polymorphisms in malaria could have similar evolutionary advantage in SCD, possibly assisting to delineate SCD pathophysiology. This study was designed to address the following questions: (1) What is the genetic diversity of African alleles of CR1 such as Knops blood group among and between SCD groups in Africa and United States? (2) Does this diversity follow an evolutionary pattern between these population and disease groups or any functional significance on disease severity? (3) If these diversities are significant, can we extrapolate the potential to interfere with the proper functioning of the complement system in SCD? To this end, we evaluated the association of these polymorphisms among SCD and healthy control populations and between Africans and Americans with SCD and analyzed interethnic differences in well-characterized and distinct ethnic control populations. We determined the genotypic, allelic, and haplotype frequencies of the CR1 SNPs (Kn^a/Kn^b , McC^a/McC^b , S11/2) in all groups.

Materials and Methods

Subjects. This study encompasses SCD patients (cases) and healthy control groups from the Centre de Recherche et

de Lutte contre la Drepanocytose (CRLD), an SCD referral center in Mali. Ethical approval was obtained from the Institutional Review Board, Rochester Institute of Technology. Exemption from Federal definition of research with human subjects was granted, as the study uses de-identified, discarded samples. This research was conducted in accordance with the principles of the Declaration of Helsinki. Ethical approval of the original study in Mali, from which samples were obtained, was granted by the National Ethical Review Board, Mali. Patients gave their informed consent to participate in that study.

To be included in this study, patients had to meet criteria including diagnosis with SCD and hospital presentation during crisis or regular patient follow-up. Patient demographics and SCD analysis have been described earlier.⁶ In brief, the SCD population from Mali is made up of 51.5% males and 48.4% females (mean age: 21 years; range: 1–51 years), predominantly of the Bambaran tribe. Healthy control populations from Africa and the United States are as described previously.^{6,33}

Samples. Discarded blood samples collected into ethylenediaminetetraacetic acid (EDTA) tubes from 389 subjects from Mali (145 SCD patients and 244 healthy controls) were spotted onto Whatmann filter papers (GE Healthcare Sciences), and genomic DNA was extracted with the Qiagen Blood Mini Kit (Qiagen Inc., Valencia, CA), with some changes to the manufacturer's instruction.⁶ Final elution volume was 100 μ L, and the DNA samples were stored at -20°C until further analysis. Genomic DNA samples from African American and Caucasian healthy controls (provided by Joann Moulds, Grifols, USA) were obtained from Shreveport, LA. African American SCD patients were recruited as part of the National Institute of Health-funded Cooperative Study of Sickle Cell Disease (CSSCD) (DNA samples provided by Betty Pace, Georgia Regents University, Augusta, GA).

Knops blood group genotyping. We selected the Knops blood group polymorphisms on the basis of their known association/interaction with malaria and disease severity in Africa and the protection conferred from malaria infection on individuals with hemoglobin S. A polymerase chain reaction (PCR) assay, which allowed the detection of all three Knops mutations, was utilized for genotyping.^{20,34} The primer sequences are as published (24 Kn^a 5'-ACCAGTGCCACACTGGACCAGATG GAGAACAGCTGTTTGAGCTT3' and 25 Rb 5'-GGAGGAGTGTGGCAGCTTG 3'), with the forward primer containing a deliberate mismatch for the detection of the Kn^a polymorphism. Two microliters of genomic DNA template was used for PCR amplification, with conditions altered to 50 μ L final volume and amplified using the Lucigen EconoTaq Plus Green 2X Master Mix PCR system (Lucigen Corporation, Middleton, WI), as described.³² The PCR cycling parameters followed published protocol,³⁴ and 5 μ L of the amplified PCR product was evaluated on 2% ethidium bromide-stained



agarose gels, visualized, and photographed. PCR product size (305 bp) was estimated with a TriDye 100 bp DNA ladder (New England Biolabs, Boston, MA), and size analysis was carried out with the Doc-It LS image analysis software (UVP Life Sciences, Upland, CA).

Restriction fragment length polymorphism (RFLP) assay. To detect the Knops blood group polymorphisms carried by CR1 [(Kna/Knb (rs41274768), McCa/McCb (rs17047660), and S11/S12 (rs17047661)], we made some slight modification to an earlier published protocol.³⁴ For Kn^a/Kn^b RFLP detection, 10 μ L of PCR product was mixed with 1 μ L of *Nde* I (20 U) and 5 μ L of 1 \times CutSmart buffer (New England Biolabs) and incubated at 37 °C for two hours. For McC^a/McC^b , 1 μ L of *Bsm* I (20 U) was added to 10 μ L of PCR product and 5 μ L of 1 \times CutSmart buffer and incubated at 65 °C for two hours. For SI^1/SI^2 , 1 μ L of *Mfe* I (20 U) was added to 10 μ L of PCR product and 5 μ L of 1 \times CutSmart buffer and incubated at 37 °C for two hours. The digestion products were analyzed on a 2% ethidium bromide-stained agarose gel, the band size was compared to a TriDye 100 bp DNA ladder (New England Biolabs), and size analysis was done as described above.³² Samples heterozygous for each SNP were run on the gels as a positive restriction enzyme control. All PCR genotyping and restriction digestion were conducted anonymously, as described previously.³³ Briefly, restriction analyses were carried out by two investigators; 50% of amplified products were subjected to repeat digestion (third investigator) with 100% concordance.

Statistical analysis. A PERL script (<http://www.perl.org>) was written to facilitate analysis of genotypic and allelic frequencies of each variant and to convert the original data files to an EH program format. Exact tests for deviation from Hardy–Weinberg equilibrium (HWE) were performed, with SNPs to be rejected based on the recommended threshold of $P < 0.001$ in control individual. Differences in genotype and allele frequencies between populations were assessed using chi-square tests. Haplotype frequencies were estimated and tested for disease differences with the EH program (<http://lab.rockefeller.edu/ott/programs>).^{35,36} Differences in haplotype frequencies among the three populations were assessed with the EH program by likelihood ratio chi-square tests.

Results

We elucidated the genetic diversity of complement regulatory genes by examining the SNPs associated with Knops blood group antigens such as the Knops a and b (rs41274768), determined by a G (Kna) or A (Knb) (V1561M change). McCoy a and b (rs17047660) were determined by A (McCa) or G (McCb) (K1590E change) and Swain–Langley 1 and 2 (rs17047661), defined by an A (S11) or a G (S12) (R1601G change) among SCD and control groups in Africa ($n = 130$ and 356, respectively) and the United States ($n = 331$ and 497, respectively). In addition, we also determined interethnic diversification of

these polymorphisms in healthy control population groups of Africans ($n = 356$), African Americans ($n = 497$), and Caucasians ($n = 254$). The demographics of the African SCD and control populations have been described previously,⁶ as are the American population.³³ We utilized the PCR–RFLP assay, which has produced consistent results in the past.^{13,34} All polymorphisms were in HWE ($P > 0.05$).

The most significant results from our study involved the S11/2 (rs17047661) and McCa/b (rs17047660) polymorphisms. Examining SCD cases and controls from Africa or the United States yielded only wild-type forms of the Kna/b polymorphisms, and hence were not included in our analysis. The distribution of S11/2 (rs17047661) and McCa/b (rs17047660) polymorphisms between SCD and controls showed no difference in genotypic (Table 1) or allelic (Table 2) frequencies among cases and controls, either in Africa or in the United States. However, we found a significant difference in both genotypic and allelic frequencies of the S11/2 (rs17047661) and McCa/b (rs17047660) polymorphisms between African and American sickle cell groups (Table 3). The homozygote mutant variant of the S11/2 (rs17047661) polymorphism occurred at significantly higher frequencies [63.1% and 40.5% ($P = 0.00002$) for African and American SCD, respectively (Fig. 1)]. The same observation was made for the McCa/b (rs17047660) polymorphisms, with genotypic frequencies of 17.7% and 5.1% [($P = 0.00006$) in African and American SCD, respectively (Fig. 2)].

We constructed haplotype groups using SNPs of the three loci [Kna/b (rs41274768), McCa/b (rs17047660), and S11/2 (rs17047661)] under investigation. We found eight haplotype groups that were formed but could find no difference in haplotype frequencies among SCD and control groups, either in Africa or in the United States (Table 4). Among the Africans, the most common haplotype combines the wild-type homozygotes (haplotype 1: S11/1, Kna/a, and McCa/a) for all three polymorphisms (42% and 44% for SCD and healthy controls, respectively), followed by haplotype 3 (36% for both SCD and control groups). Haplotypes containing the homozygote mutant form for all three polymorphisms or a combination of such (haplotypes 2, 4, 6, and 8) were not found. Similar observation were made among Americans, with haplotype 1 found with the highest frequency for all three polymorphisms (39% and 36% for SCD and healthy controls, respectively), followed by haplotypes 5 (37% and 39%, respectively) and 3 (23% for both case and control groups). Haplotypes containing the homozygote mutant form for all three polymorphisms or a combination (haplotype 2, 4, 6, and 8) were not found as well.

To clarify the possibility of ethnic diversification in genotypic and allelic frequencies, we examined the distribution of the Knops blood group polymorphisms in control populations from three ethnically classified groups (Africans, African Americans, and Caucasians). We found a significant difference in the genotypic (Figs. 3 and 4) and allelic (Table 5) frequencies



Table 1. Genotypic frequency of complement receptor-1 (CR1) gene polymorphisms between sickle cell and control groups.

AFRICAN					
POLYMORPHISM	GENOTYPE	SCD: n = 130 (%)	CONTROLS: n = 356 (%)	ODDS RATIO (95% CI)	P-VALUE
SI1/SI2 (rs17047661)	1/1	8 (6.1)	16 (4.5)	1.39 (0.50–3.55)	0.48
	1/2	40 (30.8)	109 (30.6)	1 (0.63–1.59)	1
	2/2	82 (63.1)	231 (64.9)	0.92 (0.60–1.44)	0.74
McCa/b (rs17047660)	a/a	57 (43.9)	146 (41.0)	1.11 (0.72–1.70)	0.60
	a/b	50 (38.4)	160 (45.0)	0.76 (0.49–1.16)	0.21
	b/b	23 (17.7)	50 (14.0)	1.37 (0.76–2.42)	0.25
Kna/b (rs41274768)	a/a	130 (100.0)	356 (100.0)	NS	
	a/b	0 (0.0)	0 (0.0)	NS	
	b/b	0 (0.0)	0 (0.0)	NS	
AFRICAN AMERICAN					
		SCD: n = 331 (%)	CONTROLS: n = 497 (%)	ODDS RATIO (95% CI)	P-VALUE
SI1/SI2 (rs17047661)	1/1	52 (15.7)	87 (17.5)	0.86 (0.58–1.28)	0.45
	1/2	145 (43.8)	225 (45.3)	0.95 (0.71–1.27)	0.72
	2/2	134 (40.5)	185 (37.2)	1.15 (0.86–1.55)	0.34
McCa/b (rs17047660)	a/a	193 (58.3)	286 (57.6)	1.03 (0.77–1.38)	0.88
	a/b	121 (36.6)	182 (36.6)	0.99 (0.74–1.34)	1
	b/b	17 (5.1)	29 (5.8)	0.91 (0.46–1.75)	0.87
Kna/b (rs41274768)	a/a	326 (98.5)	497 (100.0)	NS	
	a/b	5 (1.5)	0 (0.0)	NS	
	b/b	0 (0.0)	0 (0.0)	NS	

Abbreviations: SCD, sickle cell disease; NS, not significant; CI, confidence interval.

Notes: Frequencies of the genotypes at the three loci, Kn, McC and SI, determined among African and American sickle cell disease patients and control groups. a/b and 1/2 denotes the alleles at the three loci. Odds ratio was calculated by Fisher's two-tailed exact test.

of SI1/2 (rs17047661) and McCa/b (rs17047660) polymorphisms between the ethnic groups. The homozygote mutant variants had significantly higher frequencies among Africans, followed by African Americans, but they were insignificant in

Caucasians (64.9% and 37.2% vs 0.8% for SI1/2; rs17047661 and 13.5% and 5.8% vs 0% for McCa/b; rs17047660). Haplotype groups, similar to that constructed for case and controls, were analyzed, and significant interethnic difference in the

Table 2. Allelic frequency of complement receptor-1 (CR1) gene polymorphisms between sickle cell and control groups.

AFRICAN					
POLYMORPHISM	ALLELE	SCD: n = 260 (%)	CONTROLS: n = 712 (%)	ODDS RATIO (95% CI)	P-VALUE
SI1/SI2 (rs17047661)	1	56 (21.5)	141 (19.8)	1.12 (0.77–1.59)	0.59
	2	204 (78.5)	571 (80.2)	0.90 (0.63–1.30)	0.59
McCa/b (rs17047660)	a	164 (63.0)	452 (63.5)	0.97 (0.71–1.32)	0.82
	b	96 (36.9)	260 (36.5)	1.03 (0.76–1.30)	0.82
Kna/b (rs41274768)	a	260 (100.0)	712 (100.0)	NS	–
	b	0 (0.0)	0 (0.0)	NS	–
AFRICAN AMERICAN					
		SCD: n = 662 (%)	CONTROLS: n = 994 (%)	ODDS RATIO (95% CI)	P-VALUE
SI1/SI2 (rs17047661)	1	249 (37.6)	399 (40.1)	0.89 (0.72–1.10)	0.28
	2	413 (62.4)	595 (59.6)	1.12 (0.91–1.38)	0.28
McCa/b (rs17047660)	a	507 (76.6)	754 (75.9)	1.03 (0.81–1.31)	0.81
	b	155 (23.4)	240 (24.1)	0.97 (0.76–1.22)	0.81
Kna/b (rs41274768)	a	657 (99.2)	998 (100.0)	NS	–
	b	5 (0.8)	0 (0.0)	NS	–

Table 3. Genotypic and allelic frequency of complement receptor-1 (CR1) gene polymorphisms between sickle cell groups.

GENOTYPIC FREQUENCY					
POLYMORPHISM	GENOTYPE	Mali: n = 130 (%)	USA: n = 331 (%)	ODDS RATIO (95% CI)	P-VALUE
SI1/SI2 (rs17047661)	1/1	8 (6.1)	52 (15.7)	0.36 (0.14–0.79)	0.008
	1/2	40 (30.8)	145 (43.8)	0.57 (0.36–0.89)	0.01
	2/2	82 (63.1)	134 (40.5)	2.49 (1.61–3.89)	0.00002
McCa/b (rs17047660)	a/a	57 (43.9)	193 (58.3)	0.56 (0.36–0.86)	0.007
	a/b	50 (38.4)	121 (36.6)	1.08 (0.70–1.68)	0.75
	b/b	23 (17.7)	17 (5.1)	3.96 (1.94–8.22)	0.00006
Kna/b (rs41274768)	a/a	130 (100.0)	326 (98.5)	NA	NS
	a/b	0 (0.0)	5 (1.5)	NA	NS
	b/b	0 (0.0)	0 (0.0)	NA	NS
ALLELIC FREQUENCY					
POLYMORPHISM	ALLELE	Mali: n = 260 (%)	USA: n = 662 (%)	ODDS RATIO (95% CI)	P-VALUE
SI1/SI2 (rs17047661)	1	56 (21.5)	249 (37.6)	0.46 (0.32–0.65)	0.000003
	2	204 (78.5)	413 (62.4)	2.18 (1.54–3.10)	0.000003
McCa/b (rs17047660)	a	164 (63.0)	507 (76.6)	0.52 (0.38–0.72)	0.00005
	b	96 (36.9)	155 (23.4)	1.91 (1.39–2.64)	0.00005
Kna/b (rs41274768)	a	260 (100.0)	657 (99.2)	NA	NS
	b	0 (0.0)	5 (0.8)	NA	NS

Abbreviations: SCD, sickle cell disease; NS, not significant; CI, confidence interval.

Note: Percentile frequency at the three loci (Kn, McC and SI) determined between African and American sickle cell disease groups.

haplotype frequencies between Africans, African Americans, and Caucasians was found (Table 6). The most commonly encountered haplotype (haplotype 1) combined the wild-type homozygotes for all three polymorphisms (19%, 40%, and 94% for Africans, African Americans and Caucasians, respectively), followed by haplotype 5 (44%, 37%, and >1% for Africans and African Americans, respectively) and haplotype 7 (36% and 23% for Africans and African Americans, respectively). Haplotypes 5 and 7 were absent among Caucasians, while

haplotypes 2, 3, 4, 6, and 8 were not found among the three ethnic groups.

Discussion

The genetic diversity of Knops blood group antigens determined by examining SNPs associated with these genes among SCD and control groups from Africa and the United States is reported. In addition, the potential role for interethnic variation in the distribution of these polymorphisms and

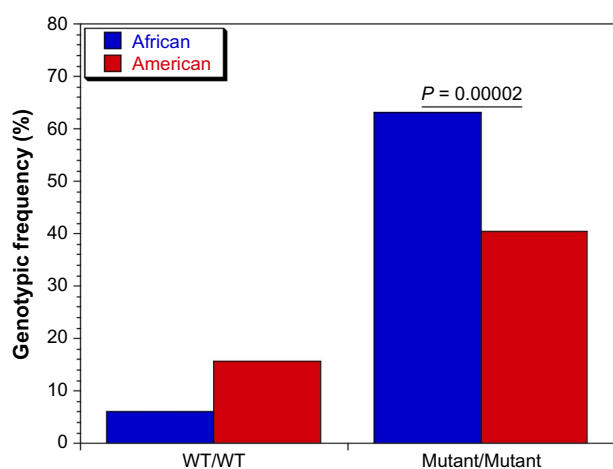


Figure 1. Genotypic frequency of SI1/2 (rs17047661) polymorphisms between African and American sickle cell disease groups. Blue bars: African sickle cell disease; red bars: American sickle cell disease; wild-type homozygotes (SI1/1); mutants (SI2/2).

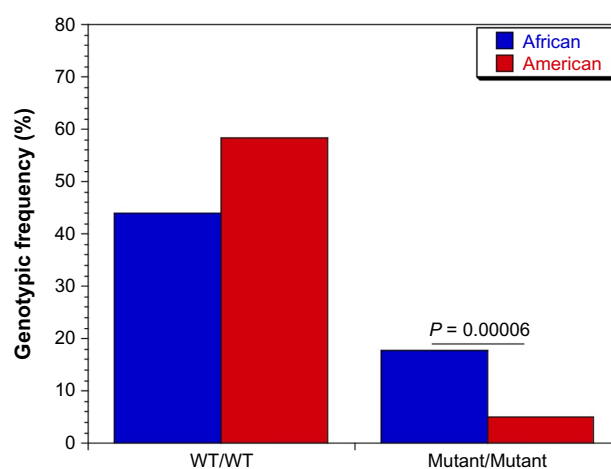


Figure 2. Genotypic frequency of McCa/b (rs17047660) polymorphisms between African and American sickle cell disease groups. Blue bars: African sickle cell disease; red bars: American sickle cell disease; wild-type homozygotes (McCa/a); mutants (McCb/b).



Table 4. Estimated haplotype frequencies of the three SNPs between African and American sickle cell and control groups.

HAPLOTYPE	HAPLOTYPE DEFINITION			HAPLOTYPE FREQUENCY (AFRICAN)			
	SL1/2	McCa/b	Kna/b	SCD	CONTROLS	ODDS RATIO (95% CI)	P-VALUE
H1	SI1	McC ^a	Kn ^a	0.42	0.44	0.93 (0.61–1.42)	0.76
H2	SI1	McC ^a	Kn ^b	0	0	NA	NA
H3	SI1	McC ^b	Kn ^a	0.36	0.36	1 (0.64–1.55)	1
H4	SI1	McC ^b	Kn ^b	0	0	NA	NA
H5	SI2	McC ^a	Kn ^a	0.21	0.19	1.06 (0.62–1.79)	0.80
H6	SI2	McC ^a	Kn ^b	0	0	NA	NA
H7	SI2	McC ^b	Kn ^a	0.007	0	NA	NA
H8	SI2	McC ^b	Kn ^b	0	0	NA	NA

HAPLOTYPE	HAPLOTYPE DEFINITION			HAPLOTYPE FREQUENCY (AFRICAN AMERICAN)			
	SL1/2	McCa/b	Kna/b	SCD	CONTROLS	ODDS RATIO (95% CI)	P-VALUE
H1	SI1	McC ^a	Kn ^a	0.39	0.36	1.12 (0.83–1.50)	0.46
H2	SI1	McC ^a	Kn ^b	0.004	0	NA	NA
H3	SI1	McC ^b	Kn ^a	0.23	0.23	1 (0.71–1.42)	1
H4	SI1	McC ^b	Kn ^b	0	0	NA	NA
H5	SI2	McC ^a	Kn ^a	0.37	0.39	0.89 (0.66–1.19)	0.42
H6	SI2	McC ^a	Kn ^b	0.004	0	NA	NA
H7	SI2	McC ^b	Kn ^a	0.002	0.007	0.37 (0.01–3.80)	0.65
H8	SI2	McC ^b	Kn ^b	0	0	NA	NA

Abbreviations: SCD, sickle cell disease; NS, not significant; CI, confidence interval.

implications for case management and disease pathophysiology are discussed. Our choice of these three sites is based on multiple reports (including ours) that have shown association/interaction of Knops blood group polymorphisms with malaria in Africa.^{13–16,20} SCD is highly prevalent in malaria endemic areas because of selection pressure that favors individuals with the recessive hemoglobin S, a major contributory factor to malaria resistance in this population.^{27–31} It is now accepted that SCD is a genetic mutation that arose as a protective

mechanism against the scourge of malaria in West Africa. Considering this, and the large amount of CICs associated with both disease states²⁰ (SCD and malaria infection), it is imperative and scientifically necessary to examine the role of CR1 polymorphisms in SCD. The emergence of human genome sequencing and large-scale population-based studies have led to an increase in disease association studies utilizing gene polymorphisms that are thought to direct disease outcomes or clinical endpoints.

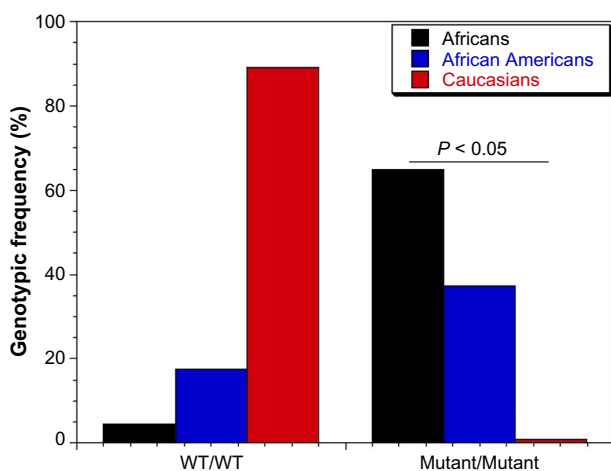


Figure 3. Genotypic frequency of SI1/2 (rs17047661) polymorphisms between ethnic groups (control populations). Black bars: Africans; blue bars: African Americans; red bars: Caucasians; wild-type homozygotes (SI1/1); mutants (SI2/2).

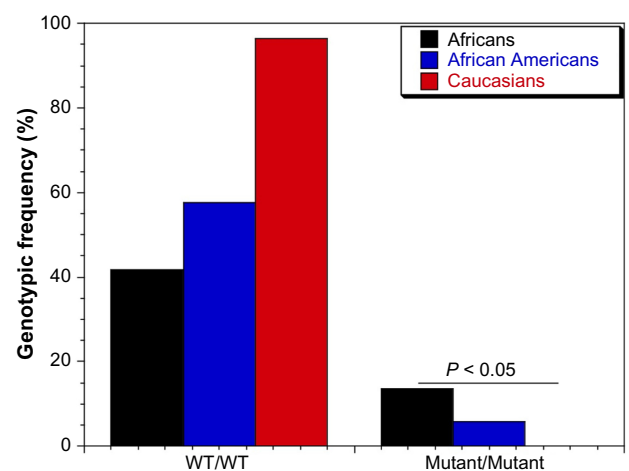


Figure 4. Genotypic frequency of McCa/b (rs17047660) polymorphisms between ethnic groups (control populations). Black bars: Africans; blue bars: African Americans; red bars: Caucasians; wild-type homozygotes (McCa/a); mutants (McCb/b).

Table 5. Allelic frequency of complement receptor-1 (CR1) polymorphisms in ethnic groups.

POLYMORPHISM	ALLELE	ETHNIC GROUPS			Chi SQUARE	P-VALUE
		AFRICAN: n = 712 (%)	AFRICAN AMERICAN: n = 994 (%)	CAUCASIAN: n = 508 (%)		
SI1/SI2 (rs17047661)	1	141 (19.8)	399 (40.1)	478 (94.1)	683.533	3.73E-149
	2	571 (80.2)	595 (59.9)	30 (5.9)	683.533	3.93E-149
McCa/b (rs17047660)	a	452 (63.5)	754 (75.9)	499 (98.2)	199.601	4.54E-44
	b	260 (36.5)	240 (24.1)	9 (1.8)	199.601	4.54E-44
Kna/b (rs41274768)	a	712 (100.0)	998 (100.0)	506 (99.6)	6.738	0.0344
	b	0 (0.0)	0 (0.0)	2 (0.4)	6.738	0.0344

Among the cases and controls, there was no significant difference in either genotypic or allelic frequencies of SI1/2 and McCa/b polymorphisms among SCD and control groups either in Africa or in the United States; nor was there any difference in haplotype frequencies among groups. This is rather surprising, since SCD is expected to have a selective pressure on the CR1 polymorphisms in the African population but not in the American population. This observation implies that these genes could potentially be stable and conserved within a population but not between populations. That is, an intra-population conservation and an inter-population diversification cannot be clarified by observations among SCD and control groups but by the significant diversity in the frequencies of the McC and SI homozygous mutant variations between SCD populations in Africa and the United States. This difference reveals the potential changes imposed on the disease in the United States but not in Africa. This could possibly aid in our understanding of how inherited defects attributable to genetic drift/founder effects can direct disease susceptibility or outcomes in diverse populations, with respect to infectious and chronic disease, and in subpopulations of cases within a broad range.³⁷⁻⁵⁵ This is a confirmation of the ancestral susceptibility hypothesis, which states that disease susceptibility alleles are ancestral while derived variants are protective. That is, SCD in the United States (a new and nonnatural environmental niche) has become adapted to conditions

of the New World, thereby imposing protective founder effects on sufferers not seen in ancestral groups.⁵⁶ The lack of haplotype difference among groups either in Africa or in the United States is noted, but the lower haplotype frequencies found among SCD and control populations in the United States compared to their African counterparts reveal the degree of genetic admixture in the American population^{36,57,58} and resultant dilution of African ancestry in the African American population.

Global distribution of SCD overlaps regions where *P. falciparum* infections are endemic, and hence the assertion that SCD mutation arose as a protective event against malaria infection.^{59,60} Published reports on the association between CR1 polymorphisms and malaria infection in Africa abound, although these studies have come up with conflicting conclusions.^{13,19,23,25,26,61} A recently published report¹⁶ states there is no association between human genetic polymorphisms in the Knops blood groups and a protective advantage against *P. falciparum* malaria in Ghana, and that the theory of selective pressure on CR1 is probably unrelated to malaria but to other infections.⁶² Our report confirms this theory and extends our hypothesis of CR1 gene-selective sweep in new climes.^{36,63-65}

The difference in SI1/2 and McCa/b polymorphisms between African, African American, and Caucasian groups as well as the extensive interethnic haplotype frequency is a confirmation of previous reports,^{20,21,23} summarized

Table 6. Estimated haplotype frequencies of SNPs in the CR1 gene in diverse ethnic groups.

HAPLOTYPE	HAPLOTYPE DEFINITION			ETHNIC GROUP		
	SI1/2	McCa/b	Kna/b	AFRICAN	AFRICAN AMERICAN	CAUCASIAN
H1	SI1	McC ^a	Kn ^a	0.19	0.40	0.94
H2	SI1	McC ^a	Kn ^b	0	0	0
H3	SI1	McC ^b	Kn ^a	0	0.0081	0.002
H4	SI1	McC ^b	Kn ^b	0	0	0
H5	SI2	McC ^a	Kn ^a	0.44	0.37	0.04
H6	SI2	McC ^a	Kn ^b	0	0	0.004
H7	SI2	McC ^b	Kn ^a	0.36	0.23	0.016
H8	SI2	McC ^b	Kn ^b	0	0	0



in Hansson et al.¹⁶ The disparity in haplotype frequencies between the African and African American population vis-a-vis an Afro-Brazilian population⁶⁶ reveals that phenotypic similarity is not sufficient to decipher potential genetic diversities that might have implications for disease outcome. The higher frequency of CR1 mutant variants in Africa but not in the United States reveals a potential pathogenic role, possibly associated with complicated disease pathogenesis in the former and potentially protective in the latter. The possibility that these polymorphic variants could be contributors to alloimmunization^{12,67–70} or post-transfusion hyperhemolysis⁷¹ in American but not African SCD deserves further scrutiny. The mutant variant is almost nonexistent among Caucasians for both McC and Sl polymorphisms, confirming previous reports and conferring advantage seen elsewhere but not in Africa.^{7,72}

The significant difference in estimated haplotype frequencies between the three populations confirms our hypothesis that the observed interethnic disparities have a potential effect on disease outcome and clinical endpoints vis-a-vis SCD, and that the more complicated and severe cases are expected among Africans while milder forms of the disease would be found in the United States. Teasing out these polymorphisms in larger African and American SCD groups, as well as individual or associative clinical endpoints, would be a perfect way to decipher the significance of complement regulatory genes in sickle cell variability and disease pathophysiology.

Acknowledgments

We are very grateful to Joann Moulds, Grifols, USA, for supplying the African American and Caucasian control DNA samples. Special gratitude is due to Betty Pace, Georgia Regents University, who graciously provided the African American sickle cell DNA samples, collected as part of the Cooperative Study of Sickle Cell Disease. Travel awards through the Provost Fund (KCD) and the Federation of American Societies for Experimental Biology Maximizing Access to Research Careers (JAN, BNT) are acknowledged. This work was presented at the Annual Biomedical Research Conference for Minority Students (ABRCMS), San Antonio, TX, USA, during November 12–15, 2014.

Author Contributions

Conceived and designed the experiment, and optimized protocols: BNT. Carried out the clinical sample collection and sickle cell genotyping in Mali: AG, DAD. Carried out DNA extraction, genotyping, and restriction digestion: BNT, KCD, JAN. Drafted the manuscript, designed the tables and figures: BNT, IGI. Carried out the statistical analyses and revised the manuscript: BNT, IGI, YL. All authors read and approved the final version of the manuscript.

REFERENCES

1. Modell B, Darlison M. Global epidemiology of hemoglobin disorders and derived service indicators. *Bull World Health Organ.* 2008;86:480–7.

2. Weatherall DJ. The inherited diseases of hemoglobin are an emerging global health burden. *Blood.* 2010;115:4331–6.
3. Weatherall DJ. The inherited disorders of haemoglobin: an increasingly neglected global health burden. *Indian J Med Res.* 2011;134:493–7.
4. Piel FB, Hay SI, Gupta S, Weatherall DJ, Williams TN. Global burden of sickle cell anemia in children under five, 2010–2050: modelling based on demographics, excess mortality, and interventions. *PLoS Med.* 2013;10:e1001484.
5. Jastaniah W. Epidemiology of sickle cell disease in Saudi Arabia. *Ann Saudi Med.* 2011;31:289–93.
6. Thakur TJ, Guindo A, Cullifer LR, et al. Endothelin-1 but not endothelial nitric oxide synthase gene polymorphism is associated with sickle cell disease in Africa. *Gene Regul Syst Biol.* 2014;8:119–26.
7. Kulozik AE, Wainscoat JS, Serjeant GR, et al. Geographical survey of beta S-globin gene haplotypes: evidence for an independent Asian origin of sickle-cell mutation. *Am J Hum Genet.* 1986;39:239–44.
8. Chui DH, Dover GJ. Sickle cell disease: no longer a single gene disorder. *Curr Opin Pediatr.* 2001;13:22–7.
9. Liu L, Muralidhar S, Singh M, et al. High-density SNP genotyping to define beta-globin locus haplotypes. *Blood Cells Mol Dis.* 2009;42:16–24.
10. Gabriel A, Przybylski J. Sickle-cell anemia: a look at global haplotype distribution. *Nat Educ.* 2010;3:2.
11. Chou ST, Jackson T, Vege S, Smith-Whitley K, Friedman DF, Westhoff CM. High prevalence of red blood cell alloimmunization in sickle cell disease despite transfusion from Rh-matched minority donors. *Blood.* 2013;122:1062–71.
12. Tatar-Calderone Z, Tamouza R, Le Bouder GP, et al. The association of CD81 polymorphisms with alloimmunization in sickle cell disease. *Clin Dev Immunol.* 2013;2013:937846.
13. Thomas BN, Diallo DA, Noumsi GT, Moulds JM. Circulating immune complex levels are associated with disease severity and seasonality in children with malaria from Mali. *Biomark Insights.* 2012;7:81–6.
14. Rowe JA, Moulds JM, Newbold CI, Miller LH. *P. falciparum* rosetting mediated by a parasite-variant erythrocyte membrane protein and complement-receptor 1. *Nature.* 1997;388:292–5.
15. Birmingham DJ, Chen W, Liang G, Schmitt HC, Gavit K, Nagaraja HN. A CR1 polymorphism associated with constitutive erythrocyte CR1 levels affects binding to C4b but not C3b. *Immunology.* 2003;1008:531–8.
16. Hansson HH, Kurtzhals JA, Goka BQ, et al. Human genetic polymorphisms in the Knops blood group are not associated with a protective advantage against *Plasmodium falciparum* malaria in Southern Ghana. *Malar J.* 2013;12:400.
17. Hourcade D, Miesner DR, Bee C, Zeldes W, Atkinson JP. Duplication and divergence of the amino-terminal coding region of the complement receptor 1 (CR1) gene. An example of concerted (horizontal) evolution within a gene. *J Biol Chem.* 1990;265:974–80.
18. Herrera AH, Xiang L, Martin SG, Lewis J, Wilson JG. Analysis of complement receptor type 1 (CR1) expression on erythrocytes and of CR1 allelic markers in Caucasian and African American populations. *Clin Immunol Immunopathol.* 1998;87:176–83.
19. Rowe JA, Raza A, Diallo DA, et al. Erythrocyte CR1 expression level does not correlate with a HindIII restriction fragment length polymorphism in Africans; implications for studies on malaria susceptibility. *Genes Immun.* 2002;3:497–500.
20. Thomas BN, Donvito B, Cockburn I, et al. A complement receptor-1 polymorphisms with high frequency in malaria endemic regions of Asia but not Africa. *Genes Immun.* 2005;6:31–6.
21. Moulds JM, Kassambara L, Middleton JJ, et al. Identification of complement receptor one (CR1) polymorphisms in West Africa. *Genes Immun.* 2000;1:325–9.
22. Rowe JA, Rogerson SJ, Raza A, et al. Mapping of the region of complement receptor (CR) 1 required for *Plasmodium falciparum* rosetting and demonstration of the importance of CR1 in rosetting in field isolates. *J Immunol.* 2000;165:6341–6.
23. Thaty V, Moulds JM, Guyah B, Otieno W, Stoute JA. Complement receptor 1 polymorphisms associated with resistance to severe malaria in Kenya. *Malar J.* 2005;4:54.
24. Doumbo OK, Thera MA, Kone AK, et al. High levels of *Plasmodium falciparum* rosetting in all clinical forms of severe malaria in African children. *Am J Trop Med Hyg.* 2009;81:987–93.
25. Zimmerman PA, Fitness J, Moulds JM, et al. CR1 Knops blood group alleles are not associated with severe malaria in the Gambia. *Genes Immun.* 2003;4:368–73.
26. Tetteh-Quarcoo PB, Schmidt CQ, Tham WH, et al. Lack of evidence from studies of soluble protein fragments that Knops blood group polymorphisms in complement receptor-type 1 are driven by malaria. *PLoS One.* 2012;7:e34820.
27. Williams TN, Mwangi TW, Roberts DJ, et al. An immune basis for malaria protection by the sickle cell trait. *PLoS Med.* 2005;2:e128.
28. Kreuels B, Kreuzberg C, Kobbe R, et al. Differing effects of HbS and HbC traits on uncomplicated falciparum malaria, anemia, and child growth. *Blood.* 2010;115:4551–8.



29. Gong L, Maiteki-Sebuguzi C, Rosenthal PJ, et al. Evidence for both innate and acquired mechanisms of protection from *Plasmodium falciparum* in children with sickle cell trait. *Blood*. 2012;119:3808–14.
30. Taylor SM, Parobek CM, Fairhurst RM. Hemoglobinopathies and the clinical epidemiology of malaria: a systematic review and meta-analysis. *Lancet Infect Dis*. 2012;12:457–68.
31. Gong L, Parikh S, Rosenthal PJ, Greenhouse B. Biochemical and immunological mechanisms by which sickle cell trait protects against malaria. *Malar J*. 2013;12:317.
32. Thomas BN, Petrella CR, Thakur TJ, Crespo SR, Diallo DA. Genetic polymorphism of *Plasmodium falciparum* merozoite surface protein-1 and 2 and diversity of drug resistance genes in blood donors from Bamako, Mali. *Infect Dis Res Treat*. 2012;6:49–57.
33. Noble JA, Duru KC, Guindo A, Yi L, Imumorin IG, Diallo DA, Thomas BN. Interethnic diversity of the CD209 (rs4804803) gene promoter polymorphism in African but not American sickle cell disease. *PeerJ*. 2015;e799.
34. Moulds JM, Thomas BN, Doumbo O, Diallo D, Lyke K, Plowe CV. Identification of the Kna/Knb polymorphism and a method for Knops genotyping. *Transfusion*. 2004;44:164.
35. Xie X, Ott J. Testing linkage disequilibrium between a disease gene and marker loci. *Am J Hum Genet*. 1993;53:1107.
36. Thomas BN, Thakur TJ, Yi L, Guindo A, Diallo DA, Ott JG. Extensive ethnogenomic diversity of endothelial nitric oxide synthase (eNOS) polymorphisms. *Gene Regul Syst Biol*. 2013;7:1–10.
37. Rosenberg NA, Pritchard JK, Weber JL, et al. Genetic structure of human populations. *Science*. 2002;298:2381–5.
38. Chung HW, Seo JS, Hur SE, et al. Association of interleukin-6 promoter variant with bone mineral density in pre-menopausal women. *J Hum Genet*. 2003;48:243–8.
39. Di Rienzo A, Hudson RR. An evolutionary framework for common diseases: the ancestral-susceptibility model. *Trends Genet*. 2005;21:596–601.
40. Young JH, Hang YP, Kim JD, et al. Differential susceptibility to hypertension is due to selection during the out-of-Africa expansion. *PLoS Genet*. 2005;1:e82.
41. Xue Y, Daly A, Yngvadottir B, et al. Spread of an inactive form of caspase-12 in humans is due to recent positive selection. *Am J Hum Genet*. 2006;78:659–70.
42. Packer BR, Yeager M, Burdett L, et al. SNP500Cancer: a public resource for sequence validation, assay development, and frequency analysis for genetic variation in candidate genes. *Nucleic Acids Res*. 2006;34(Database issue):D617–21.
43. Camargo MC, Mera R, Correa P, et al. IL1B polymorphisms and gastric cancer risk. *Cancer Epidemiol Biomarkers Prev*. 2007;16:635.
44. Huang HY, Thuita L, Strickland P, Hoffman SC, Comstock GW, Helzlsouer KJ. Frequencies of single nucleotide polymorphisms in genes regulating inflammatory responses in a community-based population. *BMC Genet*. 2007;8:7.
45. Rana BK, Insel PA, Payne SH, et al. Population-based sample reveals gene-gender interactions in blood pressure in White Americans. *Hypertension*. 2007;49:96–106.
46. Myles S, Davison D, Barrett J, Stoneking M, Timpson N. Worldwide population differentiation at disease-associated SNPs. *BMC Med Genomics*. 2008;1:22.
47. Chang MH, Lindgren ML, Butler MA, et al. Prevalence in the United States of selected candidate gene variants: Third National Health and Nutrition Examination Survey, 1991–1994. *Am J Epidemiol*. 2009;169:54–66.
48. Dadd T, Weale ME, Lewis CM. A critical evaluation of genomic control methods for genetic association studies. *Genet Epidemiol*. 2009;33:290–8.
49. Guan W, Liang L, Boehnke M, Abecasis GR. Genotype-based matching to correct for population stratification in large-scale case-control genetic association studies. *Genet Epidemiol*. 2009;33:508–17.
50. McEvoy BP, Montgomery GW, McRae AF, et al. Geographical structure and differential natural selection among North European populations. *Genome Res*. 2009;19:804–14.
51. Price AL, Helgason A, Palsson S, et al. The impact of divergence time on the nature of population structure: an example from Iceland. *PLoS Genet*. 2009;5:e1000505.
52. Wang K. Testing for genetic association in the presence of population stratification in genome-wide association studies. *Genet Epidemiol*. 2009;33:637–45.
53. Cross DS, Ivacic LC, Stefanski EL, McCarty CA. Population based allele frequencies of disease-associated polymorphisms in the Personalized Medicine Research Project. *BMC Genet*. 2010;11:51.
54. Galvan A, Ioannidis JP, Dragani TA. Beyond genome-wide association studies: genetic heterogeneity and individual predisposition to cancer. *Trends Genet*. 2010;26:132–41.
55. Dettogni RS, Sa RT, Tovar TT, Louro ID. Polymorphic genetic variation in immune system genes: a study of two populations of Espirito Santo, Brazil. *Mol Biol Rep*. 2013;40:4843–9.
56. Biswas S, Akey JM. Genomic insights into positive selection. *Trends Genet*. 2006;22:437–46.
57. Lind JM, Hutcheson HB, Williams SM, et al. Elevated male European and female African contributions to the genomes of African American individuals. *Hum Genet*. 2007;120:713–22.
58. Lao O, Vallone PM, Coble MD, et al. Evaluating self-declared ancestry of US American with autosomal, Y-chromosomal and mitochondrial DNA. *Hum Mutat*. 2010;31:E1875–93.
59. Allison AC. The distribution of the sickle-cell trait in East Africa and elsewhere, and its apparent relationship to the incidence of subtertian malaria. *Trans R Soc Trop Med Hyg*. 1954;48:312–8.
60. Aidoo M, Terlouw DJ, Kolczak MS, et al. Protective effects of the sickle-cell gene against malaria morbidity and mortality. *Lancet*. 2002;359:1311–2.
61. Jallow M, Teo YY, Small KS, et al. Genome-wide and fine-resolution association analysis of malaria in West Africa. *Nat Genet*. 2009;41:657–65.
62. Noumsi GT, Tounkara A, Diallo H, Billingsley K, Moulds JJ, Moulds JM. Knops blood group polymorphism and susceptibility to *Mycobacterium tuberculosis* infection. *Transfusion*. 2011;51:2462–9.
63. Fullerton SM, Bartoszewicz A, Ybazeta G, et al. Geographic and haplotype structure of candidate type 2 diabetes susceptibility variants at the calpain-10 locus. *Am J Hum Genet*. 2002;70:1096–106.
64. Rockman MV, Hahn MW, Soranzo N, Loisel DA, Goldstein DB, Wray GA. Positive selection on MMP3 regulation has shaped heart disease risk. *Curr Biol*. 2004;14:1531–9.
65. Thompson EE, Kuttub-Boulos H, Witonsky D, Yang L, Roe BA, Di Rienzo A. CYP3A variation and the evolution of salt-sensitivity variants. *Am J Hum Genet*. 2004;75:1059–69.
66. Covas DT, de Oliveira FS, Rodrigues ES, Abe-Sandes K, Silva WA Jr, Fontes AM. Knops blood group haplotypes among distinct Brazilian populations. *Transfusion*. 2007;47:147–53.
67. Rosse WF, Gallagher D, Kinney TR, et al. Transfusion and alloimmunization in sickle cell disease. The cooperative study of sickle cell disease. *Blood*. 1990;76:1431–7.
68. Vichinsky EP, Earles A, Johnson RA, Hoag MS, Williams A, Lubin B. Alloimmunization in sickle cell anemia and transfusion of racially unmatched blood. *N Engl J Med*. 1990;322:1617–21.
69. Aygun B, Padmanabhan S, Paley C, Chandrasekaran V. Clinical significance of RBC alloantibodies and autoantibodies in sickle cell patients who received transfusions. *Transfusion*. 2002;42:37–43.
70. Yazdanbakhsh K, Ware RE, Noizat-Pirenne F. Red blood cell alloimmunization in sickle cell disease: pathophysiology, risk factors, and transfusion management. *Blood*. 2012;120:528–37.
71. Taylor JG, Nolan VG, Mendelsohn L, Kato GJ, Gladwin MT, Steinberg MH. Chronic hyper-hemolysis in sickle cell anemia: association of vascular complications and mortality with less frequent vasoocclusive pain. *PLoS One*. 2008;3:e2095.
72. Kan YW, Dozy AM. Evolution of the hemoglobin S and C genes in world populations. *Science*. 1980;209:388–91.