A peer-reviewed version of this preprint was published in PeerJ on 24 February 2015.

View the peer-reviewed version (peerj.com/articles/799), which is the preferred citable publication unless you specifically need to cite this preprint.

https://doi.org/10.7717/peerj.799
Interethnic diversity of the CD209 (rs4804803) gene promoter polymorphisms in African but not American sickle cell disease

Elucidating the genomic diversity of CD209 gene promoter polymorphisms could assist in clarifying disease pathophysiology as well as contribution to co-morbidities. CD209 gene promoter polymorphisms have been shown to be associated with susceptibility to infection. We hypothesize that CD209 mutant variants occur at a higher frequency among Africans and in sickle cell disease. We analyzed the frequency of the CD209 gene (rs4804803) in healthy control and sickle cell disease (SCD) populations and determined association with disease. We obtained genomic DNA from 145 SCD and 244 control Africans (from Mali), 331 SCD and 379 control African Americans and 159 Caucasians. Comparative analysis among and between groups was carried out by polymerase chain reaction-restriction fragment length polymorphism (PCR-RFLP). Per ethnic diversification, we found significant disparity in genotypic (23.4% versus 16.9% versus 3.2%) and allelic frequencies (36.1% versus 25.1% versus 11.6%) of the mutant variant of the CD209 (snp 309A/G) gene promoter between Africans, African Americans and Caucasians respectively. Surprisingly, there was a wide disparity in the genotypic and allelic frequencies among African SCD versus healthy controls (10.4% versus 23.4% (genotypes) and 25.2% versus 36.1% (alleles), which is completely absent among African Americans. Comparing SCD groups, there was no difference between Africans and Americans, implying a lack of association between CD209 polymorphisms and sickle cell disease in either population. The higher frequency of CD209 mutant variants in the non-SCD group reveals an impaired capacity to mount an immune response to infectious diseases. We conclude that CD209 polymorphism play a major role in susceptibility to infectious pathogens and could potentially delineate susceptibility to and severity of co-morbidities.
Interethnic diversity of the CD209 (rs4804803) gene promoter polymorphisms in African but not American sickle cell disease

Jenelle A. Noble, Kimberley C. Duru, Aldiouma Guindo, Li Yi, Dapa A. Diallo, Bolaji N. Thomas

1Department of Biomedical Sciences, College of Health Sciences and Technology, Rochester Institute of Technology, Rochester NY; 2Centre de Recherche et de Lutte contre la Drepanocytose, Bamako, Mali; 3School of Statistics, Shanxi University of Finance and Economics, Shanxi, China

Jenelle Noble: jxn6742@rit.edu
Kimberley Duru: kcd5949@rit.edu
Aldiouma Guindo: guindo100@yahoo.fr
Yi Li: liyi3344520@hotmail.com
Dapa Diallo: dadiallo@icermali.org
Bolaji Thomas: bolaji.thomas@rit.edu
16 "Corresponding author:

17 Dr Bolaji N. Thomas

18 Department of Biomedical Sciences

19 College of Health Sciences and Technology

20 Rochester Institute of Technology

21 153 Lomb Memorial Drive

22 Rochester NY 14623

23 Office: (585) 475-6382

24 Fax: (585) 475-5809

25 Email: bolaji.thomas@rit.edu
Abstract

Elucidating the genomic diversity of CD209 gene promoter polymorphisms could assist in clarifying disease pathophysiology as well as contribution to co-morbidities. CD209 gene promoter polymorphisms have been shown to be associated with susceptibility to infection. We hypothesize that CD209 mutant variants occur at a higher frequency among Africans and in sickle cell disease. We analyzed the frequency of the CD209 gene (rs4804803) in healthy control and sickle cell disease (SCD) populations and determined association with disease. We obtained genomic DNA from 145 SCD and 244 control Africans (from Mali), 331 SCD and 379 control African Americans and 159 Caucasians. Comparative analysis among and between groups was carried out by polymerase chain reaction-restriction fragment length polymorphism (PCR-RFLP). Per ethnic diversification, we found significant disparity in genotypic (23.4% versus 16.9% versus 3.2%) and allelic frequencies (36.1% versus 25.1% versus 11.6%) of the mutant variant of the CD209 (snp 309A/G) gene promoter between Africans, African Americans and Caucasians respectively. Surprisingly, there was a wide disparity in the genotypic and allelic frequencies among African SCD versus healthy controls (10.4% versus 23.4% (genotypes) and 25.2% versus 36.1% (alleles), which is completely absent among African Americans. Comparing SCD groups, there was no difference between Africans and Americans, implying a lack of association between CD209 polymorphisms and sickle cell disease in either population. The higher frequency of CD209 mutant variants in the non-SCD group reveals an
impaired capacity to mount an immune response to infectious diseases. We conclude that CD209 polymorphism play a major role in susceptibility to infectious pathogens and could potentially delineate susceptibility to and severity of co-morbidities.
Introduction

Sickle cell disease (SCD) is an inherited multisystem disorder, characterized by chronic hemolytic anemia, vaso-occlusive crises and several other disease outcomes such as acute chest syndrome, bacteremia, leg ulcers and priapism (Bunn 1997; Benkerrou et al., 2002). SCD has shown marked variability in severity between individuals, with evidence of extensive differences in both clinical and genotypic presentations, with a global distribution, especially in sub-Saharan Africa, Middle East, parts of the Indian subcontinent, and Americans with an African or Caribbean descent (Hassell, 2010; Piel et al., 2013; Bandeira et al., 2014; Saraf et al., 2014; Thakur et al., 2014). SCD occurs in patients that are homozygous for the hemoglobin S gene, produced by a defective \(\beta\)-globin gene on chromosome 11 and has also been defined as resulting from compound heterozygosity for hemoglobin S and another \(\beta\)-globin chain abnormality (typically hemoglobin C or \(\beta\)-thalassemia), with \(\alpha\)-thalassemia serving as a modifier of the clinical manifestations (Weatherall, 2010; Saraf et al., 2014). Patients commonly require red cell transfusions to manage complications, with alloimmunization a common occurrence (Charache et al., 1983; Rosse et al., 1990; Tatari-Calderone et al., 2013) leaving such multiply transfused patients at risk for delayed hemolytic transfusion reactions (Piomelli et al., 1985; Petz et al., 1997; Taylor et al., 2008; Yazdanbaksh et al., 2012), development of autoimmune hemolytic anemia.
Infectious pathogens are a threat to those individuals with SCD, particularly children, that are prone to frequent and severe attacks (Overturf, 1999; Halasa et al., 2007; Szczepanek et al., 2013). For children in endemic countries, with very high circulating immune complexes due to constant exposure to multiple pathogenic stimuli, the added burden of these comorbidities can severely impact immune response and survival (Thomas et al., 2012). Recent reports showing high mortality rates post-vaccination in transgenic animals demonstrates that a dysregulated immune response might be responsible for such mortality and could be a major drawback to the current push to vaccinate (McCavit et al., 2011; Szczepanek et al., 2013). In fact, other reports have shown that there is an over-stimulation of pro-inflammatory cytokines in sickle cell disease patients, which might be related to vaso-occlusion (Makis et al., 2000; Pathare et al., 2004; Steinberg 2006; Conran et al., 2009; Qari et al., 2012; Bandeira et al., 2014). In fact, this hyperstimulation has been associated with sickle cell haplotype in Brazil, and as such, the obvious consequence of worsening immune response to secondary infectious pathogens or co-morbidities of infection.

Recently published data have shown that there are wide differences in infection rates and multiplicity of infection between children who are carriers of the sickle cell trait (hemoglobin AS) and those patients that possess the normal hemoglobin (HbAA) gene. In addition, extensive
differences in genomic diversity of endothelial nitric oxide synthase (eNOS)

genes, that had been reported to bear clinical significance on sickle cell
pathogenesis, has been reported between Africans and African Americans
(Thomas et al., 2013). These polymorphisms have been shown to be
potential modifiers of clinical disease, with significant differences reported
between Indian and African sickle cell disease patients (Nishank et al.,
2013; Thakur et al., 2014), and these differences could be potentially
linked to disease haplotype. These interethnic differences can be attributed
to the introduction of single nucleotide polymorphisms over a very long
period, which can ultimately influence gene expression, protein structure
and potentially function. Therefore, single nucleotide polymorphisms
located in certain promoter regions can affect transcription thereby altering
variability in the immune response, and contributing to disease
susceptibility or host resistance (Sakuntabhai et al., 2005). Despite the fact
that African Americans can trace their ancestry to sub-Saharan Africa,
recombination and genetic diversity in the African American gene pool has
facilitated the introduction of single nucleotide polymorphisms leading to
differing immune response to infectious pathogens. In addition, they are
exposed to different groups of infectious agents compared to their African
counterparts, which in turn directs immune system development, alongside
circulating antibodies. These phenomena would undergo a similar
diversification in the sickle cell population as well.
One of the most common immunogenetic markers, usually evaluated for immune system response and susceptibility to infectious pathogens is dendritic cell-specific ICAM-3 grabbing non-integrin (DC-SIGN) encoded by CD209. It assists in the migration dendritic cells on endothelium as well as enabling the activation of signal transduction pathways (Rappociolo et al., 2006; Dettogni et al., 2013). They are targets for pathogens, seeking to impair the immune response in early infection, and are known to recognize diverse pathogens, with reports showing association between CD209 gene polymorphisms and infectious agents (Mummidi et al., 2001; Martin et al., 2004). The guanine (G) to adenine (A) transition within the gene promoter (SNP -336 A/G; rs4804803) polymorphism has shown the most significance, demonstrating association with susceptibility to HIV, tuberculosis, leishmaniasis and dengue (Tailleux et al., 2003; Tassaneetrithep et al., 2003; Van Kooyk et al., 2003; Martin et al., 2004; Sakuntabhai et al., 2005; Barreiro et al., 2006). Sickle cell disease presents with variability in clinical severity, alongside genetic diversity and selection pressure imposed on patients by infectious diseases, leading to single nucleotide polymorphisms that can exacerbate or ameliorate disease outcome, especially among Africans, exposed to multiple infectious assaults and co-morbidities (Thomas et al., 2012a, 2012b). We have shown that there is an extensive diversity in the ethnogenomic distribution of endothelial nitric oxide synthase (eNOS) polymorphisms (Thomas et al., 2013). Despite reports to the contrary, we have also demonstrated that endothelin-1 polymorphisms rather than eNOS are the most important in African SCD (Thakur et al.,...
Therefore, since infections are common occurrences in SCD, there is a need to characterize the genomic diversity as well as haplotype frequency of immunogenetic markers and extrapolate their potential role in susceptibility to infectious diseases. This could clarify disease pathophysiology as well as their contribution to co-morbidities. To this end, what is the genotypic and allelic frequency of CD209 gene promoter polymorphisms (SNP -336 A/G; rs4804803) in control groups (African versus African American versus Caucasians) and between sickle cell disease populations (African versus African American), and does this polymorphism ameliorate or exacerbate disease pathophysiology? We will conduct our analyses using a polymerase chain reaction-restriction fragment length polymorphism (PCR-RFLP) assay.

**Materials and Methods**

**Subjects**

This study encompasses sickle cell disease patients (cases) and control groups (Africans versus African Americans), as well as diverse ethnic groups (Africans, African Americans and Caucasians). The African portion was conducted at the Centre de Recherche et de Lutte contre la Drepanocytose (CRLD), a sickle cell disease treatment and referral center in Mali. Approval was received from the national ethical review board, and a written consent obtained before study was initiated. Inclusion criteria include diagnosis with sickle cell disease and presentation during crisis or during regular follow-up. Sickle cell disease demographic data has been
described previously (Thakur et al., 2014). Healthy population controls comprised of family members or those recruited by word of mouth, able to provide informed consent and without a diagnosis of sickle cell disease. In the United States, control groups are African American and Caucasian self-identified individuals, recruited from Shreveport, Louisiana. African American sickle cell disease patients were recruited as part of the National Institute of Health-funded Cooperative Study of Sickle Cell Disease (CSSCD).

**Samples and Genomic DNA Extraction**

Discarded EDTA-anticoagulated blood samples, from 376 subjects (145 sickle cell disease patients and 231 controls) were spotted onto filter papers (GE Healthcare Sciences, Piscataway NJ) and genomic DNA samples extracted from the dried, spotted samples with the Qiagen Blood Mini Kit (Qiagen Inc., Valencia, CA), with some changes to the manufacturer’s instruction (Thakur et al., 2014). Final elution volume was 100 μl and DNA samples were stored at −20 °C until further analysis. Genomic DNA samples from African American sickle cell disease patients as well as African American and Caucasian controls were gratefully provided (Betty Pace, Georgia Regents University and Joann Moulds, Grifols USA respectively).

**Genotyping for CD209 single nucleotide polymorphism**

To genotype for the single nucleotide polymorphisms of the CD209 gene promoter, we utilized a previously published primer and PCR assay (Dettogni et al., 2013), with a slight modification to the protocol. The
primer sequences are 5′- GGATGGTCTGGGGTTGACAG-3′ (forward reaction) and 5′- ACTGGGGGTGCTACCTGGC-3′ (reverse reaction). 1 μl of genomic DNA served as template for PCR amplification, with conditions optimized to 25μl final volume and amplified using the Lucigen EconoTaq Plus Green 2X Master Mix PCR system (Lucigen Corporation, Middleton WI), as described previously (Thomas et al., 2012), and PCR cycling parameters as published (Sakuntabhai et al., 2005). Amplified PCR products (5 μl) was examined on a 2% (w/v) agarose gel and photographed. Positive amplification yielded products of 150 bp, with size estimated with a TriDye 100 bp DNA ladder (New England Biolabs, Boston MA).

**Restriction Fragment Length Polymorphism Assay**

We utilized the MscI (New England Biolabs, Boston MA) restriction endonuclease for restriction fragment length polymorphism analysis of CD209 (DC-SIGN 336A/G) variants. 10 μl of PCR product was mixed with 0.5μl of enzyme (5000U/ml), 5μl of 1X CutSmart buffer and incubated at 37°C for 1 hour. Digested products were analyzed on an ethidium bromide-stained agarose gel, and band analysis carried out with a Doc-It LS Image Analysis Software (UVP Life Sciences, Upland CA). Restriction analysis was conducted by two investigators anonymously and 50% of amplified products subjected to repeat digestion (3rd investigator), with 100% concordance. Wild type variants (-336A/A) were undigested (150 bp) while mutant variants (-336G/G) produced bands of 131 and 19 bp.

**Statistical analysis**

Genotypic and allelic frequencies were determined with a simple PERL
script, as described previously (Thakur et al., 2014). Differences in
genotype and allele frequencies between populations were assessed by
chi-square tests, while differences between sickle cell disease and controls
were assessed by odds ratio. Tests for deviation from Hardy-Weinberg
equilibrium (HWE) were performed, with SNP’s rejected based on the
recommended threshold of p<0.001 in control individuals.

**Results**

We found a wide disparity in the genetic diversity of the promoter variant
of CD209 (DC-SIGN1-336A/G; rs4804803) gene polymorphisms in different
populations. Genotypic frequencies of 23.4%, 16.9% and 3.2% were
observed for the mutant variant between Africans, African Americans and
Caucasians respectively (Table 1). Similar findings were made for the allelic
frequencies (36.1%, 25.1% and 11.6% respectively), with a significant
difference in both genotypic and allelic frequencies (P<0.05) of CD209 gene promoter variants between all population groups. Surprisingly, the mutant variant (GG) is almost absent among Caucasians (3.2%). The genotypic and allelic frequencies of the mutant variant (snp-336GG) had the highest frequency among Africans (23.4% and 36.1% respectively). The wild type and heterozygote variants (AA and AG), that are necessary to facilitate dendritic cell activation and function during immune response, occurred at higher frequencies among African Americans (83.1%) and Caucasians (~97%), and an unprecedented low frequency among Africans (26%) (Fig 1).

We also examined the diversity of CD209 (snp 336A/G) gene promoter polymorphisms between sickle cell disease and healthy control groups in Africa and United States. There was an extensive and significant disparity in the genotypic (Fig 2a, Table 2) frequency of the CD209 mutant variant (snp 336G/G) between sickle cell disease and control populations in Africa (P=0.002). Surprisingly, this was not the case between sickle cell disease and control populations recruited from the United States (Fig 2b) (P=0.54). In addition, the mutant variant has a higher frequency among healthy control groups than sickle cell patients (23.4% versus 10.4% respectively) in Africa, but no difference in the United States (16.9% versus 15.1% for controls and cases respectively). Similar observation was made for the allelic frequencies between controls and cases in Africa and United States (Table 3).
Since sickle cell disease has been known to display disease severity between population groups, we evaluated the diversity of CD209 (snp 336A/G) gene promoter polymorphisms between sickle cell groups recruited from Africa and United States. Surprisingly, there was no difference either in genotypic ($P=0.19$) or allelic frequencies ($P=0.72$) of mutant variants (snp 336G/G) between sickle cell disease groups (Table 4). The similarities in the genotypic and allelic frequencies (10.4% versus 15.1% and 25.2% versus 28.1% for genotypes and alleles respectively) of mutant variants were statistically insignificant.

Discussion

Sickle cell disease is the most commonly inherited hemoglobinopathy with a worldwide distribution. It is a major disease represented in populations of sub-Saharan Africa, the Middle East and several parts of India, and remains
a significant health burden borne by the African American population in the United States, and several Caribbean island nations, whose populations are dominated by ethnicities of African origin. It has recently been classified as a disease that would create a global challenge to the population of three major countries, therefore requiring a need to clarify, elucidate and decipher the various parameters contributing to its severity and diverse clinical pathophysiology among and between populations. To our understanding, this is the first report to elucidate the genomic diversity of CD209 promoter gene (snp-336A/G) polymorphisms in sickle cell disease, with the potential to clarify its role or otherwise in susceptibility to infectious pathogens. We chose three definitively classified populations, and as such permits conclusive inferences based on our finding. The African samples are from Mali facilitating analysis from a homogeneous population in comparison to the heterogeneous nature of the African American group.

Our observation that the CD209 promoter gene wild-type allele (snp-336A/A) occurred at a lower frequency among Africans compared to African Americans and Caucasians is significant, though not unexpected considering the degree of genetic admixture in the African American population. This is similar to our previous finding while examining the genomic diversity of endothelial nitric oxide synthase genes in differing populations (Thomas et al., 2013). Though both populations share a common ancestry, it is expected that the several hundred years of sexual
recombination and both the uncomfortable and under-reported legacies of
slavery would affect the genetic contribution of African genes into the
African American genome. The wild type variant is necessary for dendritic
cell activation and initiation of adaptive immune response. Therefore the
reduced frequency of this allele among Africans potentially is a major
contributing factor to their susceptibility to infectious pathogens.

Unfortunately, sub-Saharan Africa is blessed with geographic and weather
pattern that sustains the endemicity of many pathogens, especially
neglected tropical diseases, and could potentially explain the often-
encountered cases of disease co-morbidities with multiple infectious
agents in a single host. In addition, this could be an evolutionary
disadvantage in the African continent enhancing susceptibility and
infectivity, thereby underscoring the preponderance of infections. The
possibility that these infectious agents may have imposed a selection
pressure on dendritic cells, that are imperative to initiate and exert
immune pressure, is of potential significance and deserves further analysis.

In addition, this observation in Africans is enhanced by the reverse
observation in the Caucasian population of the United States. The wild type
variants (AA, AG) allele is ~97% among Caucasians and 83% among
African Americans, with the mutant variant almost absent in both groups
(3.2% among Caucasians and 16% among African Americans). This low
genotypic frequency of the mutant variant is similar to results from
previous reports, which showed 0%, 3% and 5% in the Taiwan, general
Brazilian and Sao Paulo populations respectively (Kashima et al., 2009; Wang et al., 2011; Dettogni et al., 2013). In fact, in a study conducted among three groups of healthy control populations of Thailand, a similar scenario was observed, with a genotypic frequency of 5%, 1% and 3% (Sakuntabhai et al., 2005). This observation potentially confirms our hypothesis that this immunogenetic marker has undergone evolutionary changes over time, conferring a selective advantage on populations outside of Africa (Miller et al., 1994; Gibbons, 2001; Simmer et al., 2001; Thomas et al., 2005). In other words, populations with the wild type variant are able to fight infections, hence the reduced prevalence of infectious agents, while the reverse is the case in Africa. The ancestral-susceptibility model, which states that disease susceptibility alleles are ancestral while derived variants are protective, has been proposed and validated (Di Rienzo and Hudson, 2005; Biswas and Akey 2006). It further emphasizes that ancestral alleles were adapted to historical environmental conditions, becoming maladaptive based on changes in human lifestyle and dispersal into new environmental niches (Biswas and Akey 2006). In fact, extensive reports of geographically restricted selection have been found in genome-wide studies of humans and human diseases (Carlson et al., 2005; Weir et al., 2005; Voight et al., 2006; Nakajima et al., 2004; Zhou et al., 2004; Sakagami et al., 2004; Di Rienzo and Hudson, 2005; Young et al., 2005). It seems clear therefore that local adaptation in extant populations is a major contributor to this observation (Fullerton et al., 2002; Rockman et al., 2004;
Thompson et al., 2004), and is a confirmation of the out-of-Africa hypothesis (Biswas and Akey 2006; Thomas et al., 2013).

Additionally, contrary to other reports, we conclude that the sickle cell gene potentially confer a protective mechanism against common infectious co-morbidities in Africa, based on our present observation. The higher frequency of CD209 mutants in the non-SCD group reveals an impaired capacity to mount an immune response to infectious diseases, potentially a contributor to the dominance of co-morbidities in this population. The red cell abnormality, which causes sickle cell disease, is probably protective in the present case, compared to normal individuals. We conclude that CD209 polymorphism play a major role in susceptibility to infectious pathogens among Africans and could potentially delineate severity of SCD. The implications of this finding for co-morbidities or as modifiers of SCD pathophysiology, and its significance in African Americans with SCD deserves extensive and detailed elucidation. The next step would be to determine if this protection is due to disease haplotypes and evaluate immunoassays for immunoglobulin E and eosinophilia as markers of common helminthic infections between both disease and control groups. Our endpoint would be to decipher the synergistic or pathogenic advantage of the sickle cell gene in disparate disease and population groups.
Acknowledgements

We are grateful to Joann Moulds, Grifols USA for the African American and Caucasian control DNA samples. Special gratitude 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. Special thanks to Iuri Drummond Louro, Universidad Federal do Espirito Santo, Brazil for technical assistance. College of Health Sciences and Technology, Rochester Institute of Technology provided the funding for this project. Additional support by RIT Louis Stokes Alliance for Minority Participation (LSAMP) to JAN is also acknowledged.
List of Abbreviations

SCD: sickle cell disease; OR: odds ratio; HbAA: hemoglobin AA; HbSS: hemoglobin SS; PCR-RFLP: polymerase chain reaction-restriction fragment length polymorphism; ACS: acute chest syndrome.
368  **Competing interests**

369  The authors declare that they have no competing interests.
Author Contributions

BNT conceived and designed the experiment, and optimized protocols; AG and DAD carried out sample collection and sickle cell genotyping; BNT, KCD and JAN carried out DNA extraction, genotyping and restriction digestion; BNT drafted the manuscript; BNT and YL carried out the statistical analyses. All authors read and approved the final version of the manuscript. There are no conflicts of interest.
Author Details

1 Department of Biomedical Sciences, College of Health Sciences and Technology, Rochester Institute of Technology, Rochester, NY
2 Centre de Recherche et de Lutte contre la Drepanocytose (CRLD), Bamako, Mali
3 School of Statistics, Shanxi University of Finance and Economics, Shanxi, China
References


selection and population history in the human angiotensinogen gene (AGT): 736 complete ATG sequences in chromosomes from around the world. American Journal of Human Genetics, 74:898–916


27. Petz LD, Calhoun L, Shulman IA, Johnson C, Herron RM. The sickle cell hemolytic transfusion reaction syndrome. Transfusion 37:382–392


39. Tailleux L, Schwartz O, Herrmann JL. et al 2003. DC-SIGN is the major


protein-1 and 2 and diversity of drug resistance genes in blood donors from Bamako, Mali. Infectious Diseases: Research and Treatment 6:49–57


56. Young JH, hang YP, Kim JD, Chretien JP, Klag MJ, Levine MA, Ruff CB,


**Figure Legends**

**Fig 1.** Genotypic distribution of CD209 gene promoter polymorphisms (SNP -336 A/G; rs4804803) in Caucasian, African American and African control populations. Wild type variant (snp-336A) showed no digestion (150 bp),
while mutant variant (snp-336G) produced two bands (131 and 19 bp) on digestion (lower band size not shown). Marker: 100 bp ladder, where the 500 bp band stains most intensely (New England Biolabs). Black bars-wild-type homozygotes (AA); blue bars-heterozygotes (AG); red bars-homozygote mutants (GG)

**Figure 1**

**Fig 2**: Genotypic frequency of CD209 gene promoter polymorphisms (SNP -336 A/G; rs4804803) in African (Fig 2a) and African American (Fig 2b) sickle cell disease and control groups. Amplified PCR products were digested with MscI restriction endonuclease (Fisher Scientific), and
expressed on a 2% ethidium bromide-stained agarose gel. Wild type variant (snp-336A) showed no digestion (150 bp), while mutant variant (snp-336G) produced two bands (131 and 19 bp) on digestion (lower band size not shown). Marker: 100 bp ladder, where the 500 bp band stains most intensely (New England Biolabs). Blue bars-sickle cell disease; red bars-control groups.
Table 1. Genotypic and allelic frequency of CD209 polymorphisms in diverse populations
### Table 1. Genotypic and allelic frequency of CD209 polymorphisms in diverse populations

<table>
<thead>
<tr>
<th>Polymorphism</th>
<th>Genotype</th>
<th>Ethnic groups</th>
<th></th>
<th></th>
<th>Chi square</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>African</td>
<td>African American</td>
<td>Caucasian</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>n=244 (%)</td>
<td>n=379 (%)</td>
<td>n=159 (%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CD209 (rs4804803)</td>
<td>A/A</td>
<td>60 (26.0)</td>
<td>191 (50.4)</td>
<td>101 (63.5)</td>
<td>59.9243</td>
<td>9.72E-14</td>
</tr>
<tr>
<td></td>
<td></td>
<td>117 (50.6)</td>
<td>124 (32.7)</td>
<td>53 (33.3)</td>
<td>21.5787</td>
<td>2.06E-05</td>
</tr>
<tr>
<td></td>
<td></td>
<td>54 (23.4)</td>
<td>64 (16.9)</td>
<td>5 (3.2)</td>
<td>29.1326</td>
<td>4.72E-07</td>
</tr>
<tr>
<td></td>
<td>G/G</td>
<td>54 (23.4)</td>
<td>64 (16.9)</td>
<td>5 (3.2)</td>
<td>29.1326</td>
<td>4.72E-07</td>
</tr>
</tbody>
</table>

**Allelic diversity**

<table>
<thead>
<tr>
<th>Allele</th>
<th>n=488 (%)</th>
<th>n=758 (%)</th>
<th>n=318 (%)</th>
<th>Chi square</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>CD209</td>
<td>A</td>
<td>179 (38.7)</td>
<td>444 (58.6)</td>
<td>229 (72.0)</td>
<td>83.7253</td>
</tr>
<tr>
<td></td>
<td>G</td>
<td>167 (36.1)</td>
<td>190 (25.1)</td>
<td>37 (11.6)</td>
<td>83.7253</td>
</tr>
</tbody>
</table>

Percentile frequency of the genotypes and alleles at CD209 locus, determined among African, African American and Caucasian ethnic populations. Africans were recruited from Mali while African American and Caucasian populations were recruited from Louisiana. Odds ratio was calculated by Fisher's two-tailed exact test.
Table 2. Genotypic frequency of CD209 polymorphisms between sickle cell and control groups
Table 2. Genotypic frequency of CD209 polymorphisms between sickle cell and control groups

<table>
<thead>
<tr>
<th>Polymorphism</th>
<th>Genotype</th>
<th>SCD: n=145 (%)</th>
<th>Controls: n=231 (%)</th>
<th>Odds ratio (95% CI)</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>CD209 (rs4804803)</td>
<td>A/A</td>
<td>45 (31.0)</td>
<td>60 (26.0)</td>
<td>1.28 (0.79-2.08)</td>
<td>0.2907</td>
</tr>
<tr>
<td></td>
<td>A/G</td>
<td>85 (58.6)</td>
<td>117 (50.6)</td>
<td>1.38 (0.89-2.15)</td>
<td>0.1382</td>
</tr>
<tr>
<td></td>
<td>G/G</td>
<td>15 (10.4)</td>
<td>54 (23.4)</td>
<td>0.38 (0.19-0.72)</td>
<td>0.002</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Polymorphism</th>
<th>Genotype</th>
<th>SCD: n=331 (%)</th>
<th>Controls: n=379 (%)</th>
<th>Odds ratio (95% CI)</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>CD209 (rs4804803)</td>
<td>A/A</td>
<td>110 (33.2)</td>
<td>191 (50.4)</td>
<td>0.49 (0.36-0.67)</td>
<td>4.70E-06</td>
</tr>
<tr>
<td></td>
<td>A/G</td>
<td>171 (51.7)</td>
<td>124 (32.7)</td>
<td>2.20 (1.60-3.01)</td>
<td>4.33E-07</td>
</tr>
<tr>
<td></td>
<td>G/G</td>
<td>50 (15.1)</td>
<td>64 (16.9)</td>
<td>0.88 (0.57-1.33)</td>
<td>0.54</td>
</tr>
</tbody>
</table>

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

Percentile frequency of the genotypes at CD209 locus, determined among African American sickle cell disease patients and control groups. Sickle cell disease populations were recruited from Mali and Georgia. Control populations (individuals without sickle cell disease) were recruited from Mali and Louisiana. Odds ratio was calculated by Fisher’s two-tailed exact test.
Table 3. Allelic frequency of CD209 polymorphisms between sickle cell and control groups
Table 3. Allelic frequency of CD209 polymorphisms between sickle cell and control groups

<table>
<thead>
<tr>
<th>Polymorphism</th>
<th>Allele</th>
<th>SCD: n=290 (%)</th>
<th>Controls: n=462 (%)</th>
<th>Odds ratio (95% CI)</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>CD209 (rs4804803)</td>
<td>A</td>
<td>133 (45.9)</td>
<td>179 (38.7)</td>
<td>1.70 (1.17-2.47)</td>
<td>0.003432</td>
</tr>
<tr>
<td></td>
<td>G</td>
<td>73 (25.2)</td>
<td>167 (36.1)</td>
<td>0.59 (0.41-0.85)</td>
<td>0.003432</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Polymorphism</th>
<th>Allele</th>
<th>SCD: n=662 (%)</th>
<th>Controls: n=758 (%)</th>
<th>Odds ratio (95% CI)</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>CD209 (rs4804803)</td>
<td>A</td>
<td>306 (46.2)</td>
<td>444 (58.6)</td>
<td>0.70 (0.54-0.91)</td>
<td>0.006167</td>
</tr>
<tr>
<td></td>
<td>G</td>
<td>186 (28.1)</td>
<td>190 (25.1)</td>
<td>1.42 (1.10-1.84)</td>
<td>0.006167</td>
</tr>
</tbody>
</table>

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

Percentile frequency of the genotypes at CD209 locus, determined among African American sickle cell disease patients and control groups. Sickle cell disease populations were recruited from Mali (African) and Augusta GA (African American). Control populations (individuals without sickle cell disease) were recruited from Mali and Louisiana. Odds ratio was calculated by Fisher’s two-tailed exact test.
Table 4 (on next page)

Table 4. Genotypic and allelic frequency of CD209 polymorphisms between sickle cell disease groups
Table 4. Genotypic and allelic frequency of CD209 polymorphisms between sickle cell disease groups

<table>
<thead>
<tr>
<th>Polymorphism</th>
<th>Genotype</th>
<th>Mali: n=145 (%)</th>
<th>USA: n=331 (%)</th>
<th>Odds ratio (95% CI)</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>CD209 (rs4804803)</td>
<td>A/A</td>
<td>45 (31.0)</td>
<td>110 (33.2)</td>
<td>0.90 (0.59-1.40)</td>
<td>0.67</td>
</tr>
<tr>
<td></td>
<td>A/G</td>
<td>85 (58.6)</td>
<td>171 (51.7)</td>
<td>1.32 (0.87-2.00)</td>
<td>0.16</td>
</tr>
<tr>
<td></td>
<td>G/G</td>
<td>15 (10.4)</td>
<td>50 (15.1)</td>
<td>0.65 (0.33-1.23)</td>
<td>0.19</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Polymorphism</th>
<th>Allele</th>
<th>Mali: n=290 (%)</th>
<th>USA: n=662 (%)</th>
<th>Odds ratio (95% CI)</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>CD209 (rs4804803)</td>
<td>A</td>
<td>133 (45.9)</td>
<td>306 (46.2)</td>
<td>1.05 (0.79-1.41)</td>
<td>0.72</td>
</tr>
<tr>
<td></td>
<td>G</td>
<td>73 (25.2)</td>
<td>186 (28.1)</td>
<td>0.95 (0.71-1.27)</td>
<td>0.72</td>
</tr>
</tbody>
</table>

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

Sickle cell disease populations were recruited from Mali (African) and Augusta GA (African American), while control populations, who are individuals without sickle cell disease, were recruited from Mali and Louisiana. A/G denotes the alleles at the locus. Odds ratio was calculated by Fisher’s two-tailed exact test.
Fig 1. Genotypic distribution of CD209 gene promoter polymorphisms (SNP -336 A/G; rs4804803) in Caucasian, African American and African control populations
Fig 2a. Genotypic frequency of CD209 gene promoter polymorphisms (SNP -336 A/G; rs4804803) in African sickle cell disease and control groups
Fig 2b. Genotypic frequency of CD209 gene promoter polymorphisms (SNP -336 A/G; rs4804803) in American sickle cell disease and control groups

![Bar graph showing genotypic frequency of CD209 gene promoter polymorphisms in SCD and Control groups. The graph compares the frequency of wild type (WT) and mutant alleles with a p-value of 0.54.]