Brief report

Genetics of fetal hemoglobin in Tanzanian and British patients with sickle cell anemia

*Julie Makani,1,2 *Stephan Menzel,3 Siana Nkya,1 Sharon E. Cox,1,4 Emma Drasar,3,5 Deogratius Soka,1 Albert N. Komba,1 Josephine Mgaya,1 Helen Rooks,3 Nisha Vasavada,3 Gregory Fegan,2,6 Charles R. Newton,1,4,6,7 Martin Farrall,8 and Swee Lay Thein3,5

1Muhimbili University of Health and Allied Sciences, Dar-es-Salaam, Tanzania; 2University of Oxford, Oxford, United Kingdom; 3King’s College London School of Medicine, Molecular Haematology, London, United Kingdom; 4London School of Hygiene and Tropical Medicine, London, United Kingdom; 5King’s College Hospital, Haematological Medicine, Denmark Hill, London, United Kingdom; 6Centre for Geographic Medical Research–Coast, Kenya Medical Research Institute, Kilifi, Kenya; 7Institute of Child Health, University College London, London, United Kingdom; and 8University of Oxford, Wellcome Trust Centre for Human Genetics, Oxford, United Kingdom

Fetal hemoglobin (HbF, \(\alpha_2\gamma_2\)) is a major contributor to the remarkable phenotypic heterogeneity of sickle cell anemia (SCA). Genetic variation at 3 principal loci (HBB cluster on chromosome 11p, HBS1L-MYB region on chromosome 6q, and BCL11A on chromosome 2p) have been shown to influence HbF levels and disease severity in \(\beta\)-thalassemia and SCA. Previous studies in SCA, however, have been restricted to populations from the African diaspora, which include multiple genealogies. We have investigated the influence of these 3 loci on HbF levels in sickle cell patients from Tanzania and in a small group of African British sickle patients. All 3 loci have a significant impact on the trait in both patient groups. The results suggest the presence of HBS1L-MYB variants affecting HbF in patients who are not tracked well by European-derived markers, such as rs9399137. Additional loci may be identified through independent genome-wide association studies in African populations. (Blood. 2011;117(4):1390-1392)

Methods

Tanzanian patients

The Muhimbili Sickle Cell Collaborative Program was established in 2004 in Dar-es-Salaam, Tanzania. Patients were recruited from the hematology clinic in Muhimbili National Hospital, Dar-es-Salaam. Written informed consent was obtained from patients or parents/guardians of children in accordance with the Declaration of Helsinki. Ethical approval was given by the Muhimbili University Research and Publications Committee (no. MU/RP/AEC/VOL XI/33). Hematologic data were included from nontransfused state only and from patients 5 years of age or older. Pairs of phenotype (HbF) and genotype data could be assembled for 1045 patients (Table 1).

British patients

We have included data on a population of 151 British patients (146 with HbSS and 5 with HbS/β-sickle genotype) of African-Caribbean (Jamaican, Trinidadian) or West African (Nigerian, Ghanaian, Sierra Leonean) descent from King’s College Hospital, London, United Kingdom (Table 1). The patients were recruited through the specialist clinic in the Hematology Outpatient Unit (King’s College Hospital Local Research Ethics Committee, protocol no. 01-083). At the time of study, patients ranged from 11 to 64 years of age (median, 29 years) had not been transfused within 120 days and were not receiving hydroxyurea.

Hemoglobin profiling by high performance liquid chromatography (Variant II Hemoglobin Testing System) is part of the routine clinical procedures for all patients seen at the clinics in Dar-es-Salaam and London. Genomic DNA isolated from EDTA blood samples was used for genotyping. DNA from Tanzanian samples were first genome-wide amplified by random primer amplification.13


*J. Makani and S.M. contributed equally to this study.

The publication costs of this article were defrayed in part by page charge payment. Therefore, and solely to indicate this fact, this article is hereby marked “advertisement” in accordance with 18 USC section 1734.

© 2011 by The American Society of Hematology

1390

BLOOD, 27 JANUARY 2011 • VOLUME 117, NUMBER 4
Table 1. Summary data for both patient populations

<table>
<thead>
<tr>
<th></th>
<th>Tanzanian patients</th>
<th>British patients</th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
<td>529 female, 516 male</td>
<td>82 female, 64 male</td>
</tr>
<tr>
<td>HbS/HbS</td>
<td>0</td>
<td>2 female, 3 male</td>
</tr>
<tr>
<td>Age, y</td>
<td>13</td>
<td>29</td>
</tr>
<tr>
<td>Range</td>
<td>5-45</td>
<td>11-64</td>
</tr>
<tr>
<td>HbF</td>
<td>4.40</td>
<td>5.60</td>
</tr>
<tr>
<td>Interquartile range, %</td>
<td>2.7-7.8</td>
<td>3.2-10.3</td>
</tr>
</tbody>
</table>

Single nucleotide polymorphism (SNP) genotyping was performed by the TaqMan procedure (Applied Biosystems) at King’s College London, as described,1 or by MassARRAY procedure (Sequenom) at The Wellcome Trust Center for Human Genetics, Oxford, United Kingdom. SNP rs7482144 resides within the promoter of the $G^2$-globin gene (HBG2), which is very similar to that of HBG1 ($G^1$). Typing of rs7482144 involves an initial stage of specific amplification of the HBG2 promoter encompassing rs7482144 by polymerase chain reaction14 followed by TaqMan genotyping of the polymerase chain reaction product.

Statistical analysis

For quality control, a Hardy-Weinberg test was performed on all genotype results. The HbF percentage (of total hemoglobin) values were natural log-transformed, and the extreme low tail of the distribution was trimmed to create a nearly normally distributed quantitative trait. Genetic association of this trait with SNP alleles was analyzed through multiple linear regression (SPSS, Version 12, IBM), with age and sex included as covariates. Dominance was tested for, but no significant ($P < .05$) effects were detected.

Results and discussion

All 3 principal HbF loci have a significant impact in Tanzanian patients with SCA (Table 2), the strongest association being seen at the BCL11A locus on chromosome 2. The considerable effect of alleles at this locus ($-0.406$ for rs11886868 and $-0.412$ rs4671393) results in a marked depression of mean HbF values for genotypes containing the minor allele (eg, for rs4671393: 3.7% for G/G and 5.4% for A/G compared with 8.1% for A/A). Together with a high prevalence of the minor alleles (26% and 30% for rs11886868 and rs4671393, respectively) in the Tanzanian population, this leads to an important influence of this locus on the overall phenotype (ie, 12.8% of the trait variance can be explained by genetic variation at rs4671393 alone; Table 2). A similar impact is seen in the African British patient population (Table 2) and has been reported for African American and African Brazilian patients.6-8,15,16 To date, the functional variant causing this strong association signal across most human populations has not been identified.

The largest allelic effect (0.668, Table 2) in the Tanzanian patients was detected at the HMIP locus on chromosome 6, specifically sub-locus HMIP-2 (rs9399137), leading to mean HbF values of 8.8% for C/T versus 4.5% for T/T (C homozygotes were not found; data not shown in Table 2). Variant alleles for this marker are rare in the Tanzanian patients, though, as they are in the African British and most patients or healthy persons of African descent.6-7 Therefore, its overall impact is small (1.6% of the trait variance). Hence, rs9399137, which acts as tagging SNP for the HMIP-2 sub-locus in European populations,17,18 does not track the causative sequence variant at HMIP-2 on African chromosomes very well7 because of its low frequency on the African chromosomes.

In a conditional regression analysis, there is evidence for a more extensive association signal at HMIP-2 (Table 2) that is only partially tracked by rs9399137 and independent of rs9399137. The importance of the HMIP locus in African populations might therefore have been underestimated by this and other datasets using markers tailored to European studies. Our findings also support the presence of shorter haplotype blocks at HMIP in the Tanzanians, which might include the

Table 2. Association of the 3 principal loci (BCL11A, HMIP, and the $\beta$-globin gene cluster) with HbF levels

<table>
<thead>
<tr>
<th>Genetic variants tested</th>
<th>BCL11A</th>
<th>HMIP-1</th>
<th>HMIP-2</th>
<th>HMIP-2</th>
<th>HBG2</th>
<th>ORS51B5/S6</th>
</tr>
</thead>
<tbody>
<tr>
<td>Locus</td>
<td>rs4671393</td>
<td>rs28384513</td>
<td>rs9376090</td>
<td>rs9399137</td>
<td>rs9402686</td>
<td>rs9494142</td>
</tr>
<tr>
<td>Position* on chromosome</td>
<td>60 573 750</td>
<td>135 417 902</td>
<td>135 452 920</td>
<td>135 460 710</td>
<td>135 469 509</td>
<td>135 473 333</td>
</tr>
<tr>
<td>Allele change</td>
<td>C → T</td>
<td>A → C</td>
<td>T → C</td>
<td>T → C</td>
<td>G → A</td>
<td>G → T</td>
</tr>
<tr>
<td>Minor allele frequency</td>
<td>0.26</td>
<td>0.21</td>
<td>0.01</td>
<td>0.03</td>
<td>0.06</td>
<td>0.05</td>
</tr>
<tr>
<td>Effect† of allele change</td>
<td>$-0.406$</td>
<td>$-0.146$</td>
<td>$+0.471$</td>
<td>$+0.668$</td>
<td>$+0.342$</td>
<td>$+0.164$</td>
</tr>
<tr>
<td>$P$</td>
<td>$3.0 \times 10^{-30}$</td>
<td>$1.9 \times 10^{-6}$</td>
<td>$+0.164$</td>
<td>$+0.471$</td>
<td>$+0.217$</td>
<td>$+0.024$</td>
</tr>
<tr>
<td>$n$</td>
<td>1000</td>
<td>1021</td>
<td>1021</td>
<td>975</td>
<td>1014</td>
<td>957</td>
</tr>
</tbody>
</table>

Representative SNPs for BCL11A on chromosome 2, the HBS1L-MYB Intergenic polymorphism (HMIP) on chromosome 6, and HBG2 on chromosome 11, which encodes $G^2$-globin are shown. Rs7482144 has been extensively studied as the Xmn1 $G^2$-polymorphism, a component of the classic sickle haplotypes, where the alternative allele A is detected as positive (site present, Senegal and Arab/Indian haplotypes). Allele frequencies in the patient population were derived from all genotyped individuals ($n > 1300$).6

*The chromosomal position given here is based on the public human genome assembly National Center for Biotechnology Information 36.1 (University of California Santa Cruz, genome browser).†$F^* + $ or $−$ indicates whether the allele change elevates or depresses HbF (expressed as % HbF). The value reported here (the regression coefficient $\beta$) indicates by how much the trait value changes, on average, when one of the alleles in a genotype is changed. This is also referred to as the additive allelic effect.‡Although the allelic effect sizes ($\beta$) are often comparable between the 2 groups, $P$ values are much larger in the British patients because of the smaller number of subjects studied.§Association with these makers remains significant at $P < .001$ after linkage disequilibrium with rs9399137 is taken into account (ie, these markers display HbF association independent of rs9399137).

rs7482144 is also referred to as the Xmn1-$G^2$ site, where “+” indicates the alternative allele A.
biologically active variable sites, but not rs9399137. Shorter blocks would greatly aid further fine-mapping efforts at this important locus.

Similar to HMIp, the β-globin cluster had a muted effect on HbF in the Tanzanian patients, presumably because of low allele frequencies and lack of power. Alleles at rs7482144 exert a strong effect on HbF, but the A allele (also referred to as Xmn1 Gγγγ) is absent in the Central African Republic or Bantu β⁺ haplotype, which seems prevalent in Tanzania. Other African populations with the Senegal β⁺ haplotype that contains the rs7482144 SNP would be better suited to study the effects of this variant. Such differences in haplotype and allele frequencies between populations in Africa provide a strong argument for the necessity of genome-wide association studies carried out in individual African populations.

A strong signal adjacent to the HBF cluster, recently detected in African American patients, is significant (P = .024, Table 2) in SCA patients from Tanzania but disappears (P = .14) when linkage disequilibrium with rs7482144 is taken into account. In conclusion, this additional HbF locus seems absent in Tanzanians.

The number of British patients studied here is too small to detect more than the strongest markers and to statistically compare findings with those from Tanzanians. The results show similar impacts of the HbF loci (Table 2), with the exception of rs7482144 at HBG2, where no association was seen in British patients. rs7482144 failed quality control in this group (Hardy-Weinberg test, P = .02) because of an excess of homozygote (AA) genotypes. The fact that patients with this genotype were subsequently found to originate mostly from Sierra Leone illustrates the potentially confounding influence of hidden heterogeneity or admixture.

To uncover new loci and variants controlling HbF in populations where SCA is endemic, genetic studies focused on individual African populations may be more informative, and patient resources across Africa, such as the Muhimbili Sickle Cell Collaborative Program, can make important contributions toward this goal.

Acknowledgments
The authors thank the patients and staff of Muhimbili National Hospital, Muhimbili University of Health and Allied Sciences and Hematology Outpatient Unit, King’s College Hospital; Dr Kirk Rockett, The Wellcome Trust Center for Human Genetics in Oxford for crucial advice and help with DNA handling, primer-extension preamplification (PEP), and genotyping procedures; Professor Kevin Marsh (Kenya Medical Research Institute, Kelifi, Kenya) for support; and Claire Steward for help in preparation of the manuscript.

This work was supported by The Wellcome Trust, United Kingdom (JKM 072064; project grant 080025; strategic award 084538), Kenya Medical Research Institute, Center for Geographic Medicine Research–Coast, and Medical Research Council, United Kingdom (G000111, ID51640) (S.L.T.).

Authorship

Contribution: J. Makani, S.M., and S.L.T. designed the research and wrote the manuscript; J. Makani, D.S., A.N.K., J. Mgaya, E.D., N.V., G.F., C.R.N., and S.L.T. collected data; S.N., S.E.C., and H.R. performed genotyping and analyzed results; J. Makani, S.M., H.R., S.L.T., and M.F. analyzed and interpreted results; and all authors commented on drafts of the manuscript.

Conflict-of-interest disclosure: The authors declare no competing financial interests.

Correspondence: Swee Lay Thein, King’s College London School of Medicine, James Black Centre, 125 Coldharbour Lane, London, SE5 9NU, United Kingdom; e-mail: sl.thein@kcl.ac.uk; or Stephan Menzel, King’s College London School of Medicine, James Black Centre, 125 Coldharbour Lane, London, SE5 9NU, United Kingdom; e-mail: stephan.menzel@kcl.ac.uk; or Julie Makani, King’s College London School of Medicine, James Black Centre, 125 Coldharbour Lane, London, SE5 9NU, United Kingdom; e-mail: jmakani@muhas.ac.tz.

References

Genetics of fetal hemoglobin in Tanzanian and British patients with sickle cell anemia

Julie Makani, Stephan Menzel, Siana Nkya, Sharon E. Cox, Emma Drasar, Deogratius Soka, Albert N. Komba, Josephine Mgaya, Helen Rooks, Nisha Vasavda, Gregory Fegan, Charles R. Newton, Martin Farrall and Swee Lay Thein