Malaria continues to select for sickle cell trait in Central Africa

Permalink
https://escholarship.org/uc/item/4t87z58x

Journal
Proceedings of the National Academy of Sciences of the United States of America, 112(22)

ISSN
0027-8424

Authors
Elguero, E
Délicat-Loembet, LM
Rougeron, V
et al.

Publication Date
2015-06-02

DOI
10.1073/pnas.1505665112

Peer reviewed
Malaria continues to select for sickle cell trait in Central Africa


Maladies Infectieuses et Vecteurs, Ecologie, Génétique, Évolution et Contrôle (Unité Mixte de Recherche 5290, Centre National de la Recherche Scientifique, Institut de Recherche pour le Développement, Université de Montpellier), 34295 Montpellier, France; *Evolutionary Parasitology Department, Centre International de Recherches Médicales de Franceville, BP 769 Franceville, Gabon; †Unité de Modélisation Mathématique et Informatique des Systèmes Complexes, Unité Mixte Internationale 209, Institut de Recherche pour le Développement, Université Pierre et Marie Curie, 93142 Bondy Cedex, France; ‡METABIOTA Inc., Emerging Diseases & Biosecurity, Silver Spring, MD 20910; *Department of Ecology and Evolutionary Biology, University of California, Irvine, CA 92697

Contributed by Francisco J. Ayala, March 24, 2015 (sent for review January 2, 2015; reviewed by Richard E. Paul and Stephen M. Rich)

Sickle cell disease (SCD) is a genetic disorder that poses a serious health threat in tropical Africa, which the World Health Organization has declared a public health priority. Its persistence in human populations has been attributed to the resistance it provides to Plasmodium falciparum malaria in its heterozygous state, called sickle cell trait (SCT). Because of migration, SCT is becoming common outside tropical countries: It is now the most important genetic disorder in France, affecting one birth for every 2,400, and one of the most common in the United States. We assess the strength of the association between SCT and malaria, using current data for both SCT and malaria infections. A total of 3,959 blood samples from 195 villages distributed over the entire Republic of Gabon were analyzed. Hemoglobin variants were identified by using HPLC (HPLC). Infections by three species of Plasmodium were detected by PCR followed by sequencing of a 201-bp fragment of cytochrome b. An increase of 10% in P. falciparum malaria prevalence is associated with an increase by 4.3% of SCT carriers. An increase of 10 y of age is associated with an increase by 5.5% of SCT carriers. Sex is not associated with SCT. These strong associations show that malaria remains a selective factor in current human populations, despite the progress of medicine and the actions undertaken to fight this disease. Our results provide evidence that evolution is still present in humans, although this is sometimes questioned by scientific, political, or religious personalities.

Sickle cell disease | Plasmodium falciparum | human evolution | Gabon | natural selection

Sickle cell disease (SCD) is a serious public health concern, present mainly in tropical countries, especially Africa (1, 2), but spreading to many other countries with the increasing population migrations (3, 4). It is due to a mutation of the gene coding for the β-globin, called the hemoglobin (Hb)S mutated allele, whereas the wild type is called the HbA allele (5). In the homozygous condition (hereafter denoted as HbSS), this mutation causes a severe disease, sickle cell anemia, which was almost always lethal before the advent of modern medicine, and still is where modern medical care is not available (6, 7). In the heterozygous condition (HbAS), called sickle cell trait (SCT), the mutation results in a much milder condition to the point that carriers may go unnoticed, although it has been associated with a variety of conditions or diseases like hematuria, splenic infarction, and exercise-related sudden death (8).

The global impact of SCD has been estimated at approximately 275,000 births every year (9) and could reach approximately 400,000 births annually by 2050 according to recent projections (4). The most affected continent is Africa, where approximately 85% of the cases are located, and where the vast majority of SCD children do not reach adulthood (10). SCT prevalence is more than 15% in a great part of Central Africa (7) and reaches 28% in Gabon (11), with the consequence that on average between 1 and 2% of all born children are affected by SCD.

Besides its importance as a public health threat, SCD holds a special place in human population biology as a paradigmatic example of selective advantage of the heterozygotes, leading to balancing selection. Given that persons carrying the homozygous HbSS genotype had almost no chance to reproduce, there should be a steady decrease of the HbS allele frequency with each generation in the absence of a counteracting force. Thus, the question arises as to what was the force that maintained the HbS allele at a high frequency in human populations. Mainly based on the similarity of the geographical distribution of the two diseases (malaria and SCD), Haldane (12) hypothesized that SCT could provide protection against severe forms of malaria. This insight was substantiated thereafter by Allison (13); see refs. 14 and 15 for historical details and ref. 16 for a review of the proposed mechanisms of this protective effect. SCD/SCT is now a member of a set of several other genetic disorders or variants linked to malaria resistance, as showed by recent studies (17, 18).

Although the principle of SCT protection against severe malaria as the factor accounting for the high levels of SCT in some parts of the world is now generally accepted, the details of...
this interaction are much less known. There is a lack of epidemiological studies assessing the strength of the association between malaria and SCT. Global studies, old (13) or recent (19), provide compelling evidence in favor of the so-called malaria hypothesis, but are based on historical data of malaria prevalence, and do not allow a precise estimation of the SCT effect against malaria. At a global scale, human populations differ in several respects, including their genetic backgrounds, lifestyles, and access to healthcare, among others. Moreover, human populations are immersed in diverse environments, especially concerning pathogen communities. Thus, there exist numerous potentially confounding factors, likely to obscure the precise relationship between malaria and SCT.

The aim of this paper is to provide an epidemiological study investigating the relationship between present-day malaria and SCT prevalence, at the scale of an African country, namely the Republic of Gabon. This scale is relevant to address the question raised; it is large enough for there to be variation in malaria circulation, and small enough for its populations to have fairly similar lifestyles in similar environments.

**Results**

Blood samples were obtained from 4,359 participants, 15–80 y old, sampled from 220 randomly selected villages, roughly evenly distributed among the nine Gabonese provinces (Fig. 1A and B). Because this study focuses on the Bantu population, samples from Pygmy villages were not included, thus resulting in a cohort of 3,959 participants, 17–80 y old, from 195 villages. Among the 3,959 blood samples, 859 (21.7%) had the HbAS (SCT) genotype, one had the HbAC genotype, the remaining samples having the HbAA (normal) genotype. No sample had the HbSS genotype, thus confirming the high mortality in childhood associated with SCD. The HbAC individual—a foreigner from Ghana—was subsequently removed from the study, resulting in a sample size of 3,958.

Of the 3,958 samples, 2,061 (52%) were positive for *Plasmodium* cytochrome *b* (cytb). Of these 2,061 samples, 572 (28%) could not be used to determine the species, because of an insufficient cytb amplicon concentration. Among the remaining 1,489 samples, 1,387 (93.1%) had *Plasmodium falciparum* infection, 32 (2.1%) had *Plasmodium malariae* infection, and 32 (2.1%) had *Plasmodium ovale* infection. A total of 234 samples had two infections and 11 samples had three. Table 1 shows the distribution of sample types according to genotype and *Plasmodium* species; Table 2 shows the prevalences of HbAS, HbAC, and *Plasmodium falciparum* for each 10-y age class; Dataset S1 shows detailed information for each of the 195 villages.

Because *P. ovale* infections were rare (10 only *P. ovale* and 11 associated with either *P. falciparum* or *P. malariae*), and because infections by *P. falciparum* and *P. malariae* were strongly associated (Chi-2 = 140 on 1 d.f., *P* value < 1.0×16), we decided to exclude the *P. ovale* association with the HbAS genotype, and to restrict ourselves to two analyses.

**P. Falciparum Malaria.** In this analysis, the 572 samples infected by an undetermined *Plasmodium* species were discarded. The GLM model showed that SCT was associated with *P. falciparum* malaria prevalence: odds-ratio = 1.054, 95% confidence interval (c.i.) = (1.014–1.096); *P* value = 0.008; and with age: odds-ratio = 1.071, 95% c.i. = (1.012–1.133), *P* value = 0.017. Malaria prevalence is taken in units of 10%, and age is taken in units of 10 y. In other words, the odds-ratio for malaria prevalence is given for a 10% increase in prevalence, and the odds-ratio for age is given for a 10-y increase of age. The 1.071 odds-ratio implies that when age changes from 50 y (the sample median) to 60 y, the SCT prevalence changes from 18.4% to 19.5%, an increase by 5.5%. SCT was not associated with sex: odds-ratio M vs. F: 0.997, 95% c.i. = (0.85–1.17), *P* value = 0.97. Fig. 1A and B show the maps of SCT and *P. falciparum* prevalence. Fig. 1C shows the distribution of village SCT prevalence as a function of village malaria prevalence, as well as the SCT prevalence predicted by the statistical model.

**All-Plasmodium Malaria.** Because it is likely that most of the excluded 572 undetermined infections were actually *P. falciparum* infections, removing them could bias the results. Hence, we performed a subsidiary analysis. In this analysis, the *Plasmodium* species was ignored; infection status was taken as positivity to *Plasmodium* cytb. The GLM model showed that SCT was associated with *Plasmodium sp.* malaria prevalence: odds-ratio = 1.058, 95% c.i. = (1.012–1.100), *P* value = 0.005, and with age: odds-ratio = 1.072, 95% c.i. = (1.014–1.134), *P* value = 0.015. SCT was not associated with sex: odds-ratio M vs. F: 0.992, 95% c.i. = (0.85–1.16), *P* value = 0.92.

**Discussion**

The main finding of this study is a strong association between *P. falciparum* malaria and SCT prevalence. The all-*Plasmodium* analysis, yielding almost identical estimates to those obtained for the analysis of *P. falciparum* alone, shows that the removal of the undetermined *Plasmodium* infections did not result in bias, but because 90% of all *Plasmodium* infections are *P. falciparum*, it
does not add new information. Translated into prevalence scale, the observed odds-ratio of 1.054 for *P. falciparum* alone means that when, for example, the malaria prevalence changes from 40% (its average value) to 50%, the number of SCT carriers changes from 21.1 to 22.0%, that is an increase in number of SCT carriers by 4.3%. The same change in malaria prevalence (40–50%) would yield, in the absence of bias introduced by genetic counseling, a change in the prevalence of SCD (homozygotes) newborns from 4.4 to 4.8%, an increase by more than 9%. Using the formalism of Hartl and Clark (20), if WAA, WAS, and WSS denote the fitnesses of the three genotypes, we have here WSS = 0, we can assume conventionally that WAS = 1 and then WAA = 1–s, where s is the selection coefficient against the HbAA genotype. It can be proved that at equilibrium, s = p/(1–p) where p is the frequency of the HbS allele. If we take the 21% HbAS average prevalence in Gabon, it translates to a HbS frequency p = 0.105 and to a selection coefficient s = 0.12, 95% c.i. = (0.110–0.125), a figure comparable to that of 0.11 found by Cavalli-Sforza and Bodmer (21). If we consider only those villages where the estimated prevalence of *P. falciparum* infection is less than 20%, we find s = 0.093, 95% c.i. = (0.070–0.121), and if on the contrary we consider those villages where the prevalence is above 60%, we find s = 0.135, 95% c.i. = (0.116–0.156). If we now consider the effect of age, taking an average 50% of *P. falciparum* prevalence in the population, and considering the selection coefficient varies from s = 0.118, 95% c.i. = (0.117–0.119) for the cohort of individuals who were 20 y old at the time of the survey, to s = 0.121, 95% c.i. = (0.118–0.123) for persons who were 60 y old.

The positive association between asymptomatic *P. falciparum* malaria and SCT prevalence in Bantu villages is consistent with previous findings, in particular with what Piel et al. observed in Africa, in a recent and elaborate study (19). There are, however, a number of differences. First, these authors (19) did not use a continuous measure of malaria prevalence, but instead a classification of endemicity in six levels, from malaria-free to holoendemic. Second, they used malaria data from the preintervention era, which occurred before 1960. Third, having individual data enabled us to take into account the individual potentially confusing factors of age and sex, although it turned out that only age was associated with SCT and malaria.

The most significant difference between the present study and previous ones is that other studies used indices based on clinical malaria, whereas our estimates of malaria prevalence are based partly on asymptomatic infections. Indeed, only healthy individuals were enrolled, but whether we find that the infected individuals were recovering from a clinical episode, or on the contrary would develop symptoms sometime after having been sampled, is not known. One could argue that SCT, by providing protection against clinical malaria (22, 23), should be positively associated with asymptomatic malaria and negatively associated with clinical malaria. However, mathematical modeling (24, 25) has shown that, in the long term, high levels of malaria prevalence select for SCT, and that, in turn, a high SCT prevalence is associated with a higher prevalence of asymptomatic malaria, so that SCT prevalence is positively associated with both asymptomatic and clinical malaria. Hence, our findings demonstrate a selective effect of *P. falciparum* malaria on SCT.

The second finding of our study, namely the significant increase of SCT prevalence with age, could be explained by one of two hypotheses. The first one is that the significant increase of SCT with age could indicate that malaria selective pressure is relaxing with time as a result of improved medical care and prophylaxis, so that elder people experienced stronger selective pressure in their youth than today’s young people. Alternatively, the increase of SCT prevalence with age could indicate that the HbAS genotype continues to provide protection against severe malaria with respect to the HbAA genotype in adult life. Indeed, contrary to the once-prevailing opinion, immunity to severe malaria is not always acquired in the first years of life (26, 27) and severe malaria is not uncommon in adults (28, 29).

Sex was not associated with SCT, a result that is not surprising given that the β-globin gene is autosomal. This lack of association between SCT and sex shows that there is no significant difference in the protection against malaria provided by SCT to men and women.

The association observed at the level of Gabonese villages would be of interest from a health management perspective. Villages with a high level of malaria transmission could be the targets of coupled antimalaria campaigns and information campaigns regarding SCD. In any event, malaria and SCD are two entangled health threats that should be managed synergistically. In Gabon, given that approximately 21% of the population carries the S allele (11), on average 500 (1%) of the approximately 50,000 children born each year in the country (30) will carry the HbSS genotype and, hence, develop SCD.

The impact of our results goes far beyond the public health perspectives. To our knowledge, our results constitute one of the best illustrations that present day humans are still evolving in response to the selective pressures imposed by their environment, notably in the present case, the pathogenic one.

Reports from the scientific, public, and religious communities have repeatedly claimed that evolution is no longer relevant to humans and that our species now mostly or only depends on culture and technology for survival (31). Several studies have provided clear evidence that selection impacted human evolution after the agricultural revolution. However, most of the examples are historical (before the industrial era) (see refs. 31 and 32 for reviews) or concerned populations in the last hundred years (e.g., refs. 33 and 34).

### Table 1. Blood samples classified according to genotype and *Plasmodium* species

<table>
<thead>
<tr>
<th>Genotype</th>
<th>Not infected</th>
<th>Undetermined</th>
<th><em>Plasmodium</em> Pf only</th>
<th>Pf only</th>
<th>Po only</th>
<th>Pf+Po</th>
<th>Pf+Po</th>
<th>Po only</th>
<th>Pf+Pm+Po</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>HbAA</td>
<td>1,492</td>
<td>446</td>
<td>894</td>
<td>67</td>
<td>10</td>
<td>172</td>
<td>8</td>
<td>2</td>
<td>8</td>
<td>3,099</td>
</tr>
<tr>
<td>HbAS</td>
<td>405</td>
<td>126</td>
<td>250</td>
<td>23</td>
<td>0</td>
<td>51</td>
<td>1</td>
<td>0</td>
<td>3</td>
<td>859</td>
</tr>
<tr>
<td>Total</td>
<td>1,897</td>
<td>572</td>
<td>1,144</td>
<td>90</td>
<td>10</td>
<td>223</td>
<td>9</td>
<td>2</td>
<td>11</td>
<td>3,958</td>
</tr>
</tbody>
</table>

**Pf, Plasmodium falciparum; Pm, Plasmodium malariae; Po, Plasmodium ovale.**

### Table 2. Prevalences of HbAS genotype, *Plasmodium sp* infection, and *P. falciparum* infection by 10- y age class

<table>
<thead>
<tr>
<th>Age class</th>
<th>15–24</th>
<th>25–34</th>
<th>35–44</th>
<th>45–54</th>
<th>55–64</th>
<th>65+</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sample size</td>
<td>351</td>
<td>537</td>
<td>725</td>
<td>846</td>
<td>1022</td>
<td>477</td>
<td>3,958</td>
</tr>
<tr>
<td>HbAS</td>
<td>74</td>
<td>100</td>
<td>151</td>
<td>185</td>
<td>239</td>
<td>110</td>
<td>23.1%</td>
</tr>
<tr>
<td><em>Plasmodium sp</em></td>
<td>197</td>
<td>288</td>
<td>407</td>
<td>460</td>
<td>504</td>
<td>205</td>
<td>43.0%</td>
</tr>
<tr>
<td><em>P. falciparum</em></td>
<td>144</td>
<td>206</td>
<td>276</td>
<td>285</td>
<td>344</td>
<td>132</td>
<td>27.7%</td>
</tr>
</tbody>
</table>
The Republic of Gabon is a Central African country, with an area of 270,000 square kilometers and a population of 1.5 million people (30). The Gabonese populations belong mainly to two ethnocultural backgrounds: the Bantu people, longtime established farmers, and the Pygmy people, who traditionally lived a nomadic life in the forest and are now becoming more and more sedentary (35). Because of this difference in lifestyle and environment, we decided to exclude pygmies from our study. We therefore are concerned only with the Bantu populations of Gabon.

Data were collected between June 2005 and September 2008, as part of a project focused on several pathogens in Gabon (36). First, 220 villages were randomly selected, out of the 2,056 Gabonese villages, so as to cover evenly the nine Gabonese provinces. In the 220 villages, all healthy volunteers over the age of 15 who had been residing in the village for more than one year were included in the study, resulting in 4,349 participants. Fig. 1 A and B show the location of the included villages.

Written consent was secured from all participants. In the case of minors, consent was obtained from at least one parent. Our study was approved by the Gabonese Ministry of Health on March 15, 2005, with research authorization No. 00093 for our samples.

Detection of Hemoglobin Variants. The presence of abnormal hemoglobin was ascertained by isoelectric focusing. When an abnormal protein was detected, HPLC was used to identify the variant, Hbs or Hbc (37, 38), according to the protocol described in ref. 39. Details of the procedures are described (11).

Plasmodium Detection and Species Determination. Plasmodium infection data were obtained by a two-step process. We first tested the blood samples for Plasmodium cyt b mitochondrial sequences by nested PCR, as described (40). To identify the Plasmodium species present in each positive sample, we used deep sequencing. More particularly, we used the 454 GS-FLX Titanium technology on pools of several hundreds of samples amplified for a fragment of 201 bp (Table S1 and Table S2). A sample was considered infected with a given species of Plasmodium when the proportion of reads attributed to that species was at least 5% of all reads for that sample. Details are given in SI Materials and Methods.

Statistical Analyses. The association between SCT and the intensity of malaria circulation in the living environment of a given individual was assessed by logistic regression (41). A proxy for malaria circulation was taken as the malaria prevalence in the individual’s village, as estimated by the proportion of carriers among the sampled individuals. To take into account the hierarchical sampling, we introduced two nested random effects: village and province (42). Included potential confounders were sex and age. All computations were performed with the R software and specifically the lme4 package (43, 44).

ACKNOWLEDGMENTS. The authors thank Centre National de la Recherche Scientifique (CNRS), Centre International de Recherche Médicales de Franceville (CIRMF) and Institut de Recherche pour le Développement (IRD) for general support, International Atomic Energy Agency (IAEA) for the funding of the Center for sickle-cell screening and for supporting the national program on neonatal diagnosis in Libreville, Gabon, and Dr. Krishnamoorthy Rajagopal from Institut National de la Santé et de la Recherche Médicale (INSERM) for initiating research on Sickle Cell Disease at CIRMF. This research was funded by Agence Nationale de la Recherche (ANR) Grant ORIGIN, JCIC 012.