

## Sickle Cell Trait in Sudanese Populations

Bakhietta Atalla, Department of Pediatrics, University of Bahr El Gazal, Sudan (Pediatrics)

Muntaser E. Ibrahim, Institute of Endemic Diseases, University of Khartoum (Molecular Genetics)

Ahmed M. El Hassan, Institute of Endemic Diseases, University of Khartoum, (Pathology)

Abdelrahim O. Mohammed, Department of Biochemistry, Faculty of Medicine, University of Khartoum, Khartoum, Sudan (Biochemistry)

Gafar Ibn Auf, Federal Ministry of Health, Khartoum, Sudan (Pediatrics)

Zein Al Abdin Karrar, Department of Pediatrics, Faculty of Medicine University of Khartoum, Khartoum, Sudan (Pediatrics)

### ABSTRACT

The presence of a geographical pattern in the distribution of the sickle cell gene (S gene) and its association with malaria was well documented. To study the distribution of S gene among various ethnic and linguistic groups in the Sudan, we analysed a sample of 189 families reporting to the sickle cell disease clinic June 1996 and March 2000 against their ethnic and linguistic affiliations and their geographic origin. The results indicated that the S gene is predominantly distributed among Afroasiatic speaking groups including nomadic groups of Arab and non Arab descent who migrated to the Sudan in various historical epochs (>70%). The proportion of patients reporting from other ethnic groups, including Nilo-Saharan speaking groups who populated the country in previous times is between 20-30% and is disproportionate to their population size and their wider geographic range in the Sudan. These results highlight the strong link of the middle Nile Valley with west Africa through the open plateau of the Sahel and the nomadic cattle herders and probably the relatively new age of the sickle cell trait.

---

Correspondence to:

Mustaser E. Ibrahim

Department of Molecular Biology, Institute of Endemic Diseases,

University of Khartoum, Medical Campus, Qasser Street,

P.O. Box 102 Khartoum, Sudan

Fax: 249 11 779712

Phone: 793263

e-mail:

## INTRODUCTION

Due to a variety of reasons, sickle cell anaemia (SCA) became one of the most studied genetic alterations affecting mankind. The condition could either lead to natural resistance to a major infectious diseases in the tropics (malaria) among heterozygous individuals, or to a debilitating disease that could leads to early death of the homozygous carriers of the gene. The sickle cell trait is known to be confined or occurs in higher frequencies in particular affected populations in the tropics and therefore makes one of the most classical population specific markers (Lanclos et al, 1991, Oner et al, 1992, Goncalves et al, 1994).

Controversy surrounded the origin of the sickle cell gene and the question of whether it arose as a single mutation which then spread to occupy its present distribution or whether it originated from identical mutations in different geographical areas, remained debated until recently. A single mutation occurring in Neolithic times in the fertile Arabian peninsula was favored by Lehmann (1959a, 1959b), who postulated that the changing climatic conditions, and conversion of this area to a desert, caused a migration of peoples that could have carried the gene to India, Eastern Saudi Arabia, and down to Equatorial Africa. A single mutation theory was also favored by Gelpi who considered that the evidence from blood groups and other genetic markers was more compatible with an origin in Equatorial Africa and subsequent diffusion of the gene to India, Arabia, and the Mediterranean by the East African slave trade (Kamel and Awany, 1965).

However, and on the other hand, the multiple mutation theory has recently received more support with the increasing use of new molecular markers which stands in favor of the hypothesis that the mutation occurred independently several times on chromosomes manifesting a variety of different haplotypes, probably between 3000-6000 generations (70 000 - 150 000 years) ago (Kurjnit, 1979).

Five major haplotypes are associated with SS. These have been named the Benin, Bantu or CAR, Cameroon, Senegal, and the Saudi-Indian types, and have also been numbered as 19, 20, 17, 3, and 31 respectively (Oner et al, 1992).

Sudan is one country where genetic diversity could ideally be utilized to illuminate population histories. The constitutional unit known as the Republic of the Sudan is a far smaller area than the geographic ethnic part of Africa which have variably been called Kush, Nubia and Bilad Es Sudan (Land of the blacks). More than 100 languages are spoken in the Sudan. These languages relate to three of the four of Greenberg's African language families: i.e. the Afro-Asiatic, the Nigar-Kordofanian and the Nilo-Saharan families (Greenberg, 1954). There are about 570 tribes who

have been grouped into 56 ethnic groups on the basis of linguistic, cultural and other ethnological characteristics. Malaria is also endemic with varying levels in the whole of the country, and there are several reasons to believe that a long history of malaria extends back in time in this part of the African continent.

We analyzed a sample of Sudanese with sickle cell traits who reported to a central clinic in children emergency. We believe that the sample is fairly representative because the demographic changes in the capital (estimated to be above 4 millions), due to the rural urban migration, displacement because of wars, famine and draught, ensured a high representation of families from south, west, south-east.

## **MATERIALS AND METHODS**

Khartoum, the capital, witnessed a very rapid expansion due more to migratory factors rather than to the local birth-rate. The years 1984-85 showed massive influx of migrants from the West and South as a result of the civil war and natural disasters. The greater Kharotum (in the sense of the three towns: Khartoum, Omdurman and Khartoum North) almost represent all the tribes of the Sudan

This study was carried on one hundred and eighty nine families reporting to the sickle cell clinic at children emergency hospital from June 1996 to March 2000. The clinic is the only specialized referral clinic for children with sickle cell anaemia for patients from different parts of greater Khartoum.

A questionnaire was designed which included questions about the patient's history of disease, father's tribe, mother's tribe and their area of origin. For the question about the tribe the responders sometimes gave the name of a section. We classified the sections which belong to the same tribe after checking in specialized relevant literature. Patients and subjects were examined clinically. Individuals with minor health problems were treated and those requiring further investigation were referred to specialized clinics in Khartoum Teaching Hospital.

Sickling test using Sodium Metabisulphite and electrophoresis by cellulose acetate method were done for all patients and most of the patients. Only those who have evidence of Hb SS/SA or other sickle variants were included. Data was entered into the programmes EP16 and analyzed using Chi square test.

## RESULTS

The results of the distribution of the sickle cell trait among various tribal groups are summarized in table 1 and figure 1. All families affected were grouped according to their geographical location linguistic group, and whether they belonged to the particular tribal group from the side of the father or mother. Table 1 shows the names of the main tribes or groups affected in western, northern, central and southern Sudan, their linguistic affiliation and comparison of the total population census to the percentage of S gene in the same population. Figure 1 shows those tribes in which there are 5 affected individuals (families) or more, and whether the affected belong to the particular group from the mother or the father's side. The histogram indicates the percentages of patients with sickle cell diseases of different tribes from the mother and father's side, showing high concordance between the parenthoods and thus a high degree of within group marriages. The Messeyria tribe possessed the highest frequency of the S gene thus making it the highest risk group in the sample.

Figure 2 shows the percentage of each linguistic group in the sample, indicating the Afro-Asiatic group as by far the largest compared to other linguistic groups accounting for nearly 70% of the sample, followed by the Nilo-Saharan speaking group (20%). Those of mixed ethnic origins accounted for 6.3% and finally the Niger-Kordofanian who accounted for 5.8%. Compared to the actual percentages of the various populations these numbers are quite disproportionate (Table 1, P value = 0.000) except for populations in the centre. The major portion of those speaking Afro-Asiatic languages here are nomadic cattle herders in addition to the Hausa who speaks a Chadic branch of the Afro-Asiatic family. Within Nilo-Saharan group (Figure 2) the Bergu and Fur were the ones with the highest frequencies of the S gene.

Of those tribes and sub tribes with less than 5 individuals in the sample 35% belonged to the Nilo-Saharan group from the mother side. 2% to the Niger-Kordofanian group and 69% to the Afro-Asiatic. Similar percentages were obtained for the side of the father, reinforcing the previous results.

Figure 3 shows the percentage of cases in the sample from each geographic location and the percentage of the S gene. More than half the cases clustered in the state of Kordofan followed by Darfur and finally the Central states. This agrees with the distribution of the Nomadic Bagara tribes in the country.

The map shows the frequency of the S gene in the west, north, south and centre of the Sudan, in addition to the different zones of endemicity for

malaria. It shows the lack of association between the distribution of the S gene and the degree of malaria endemicity.

Table 1. The group/tribal affiliation of the Sickle cell diseases affected population in our sample, their linguistic group, percentage in the sample, approximate percentage in the population, chi square and P value of comparisons between percentage in sample and in general population of the Sudan.

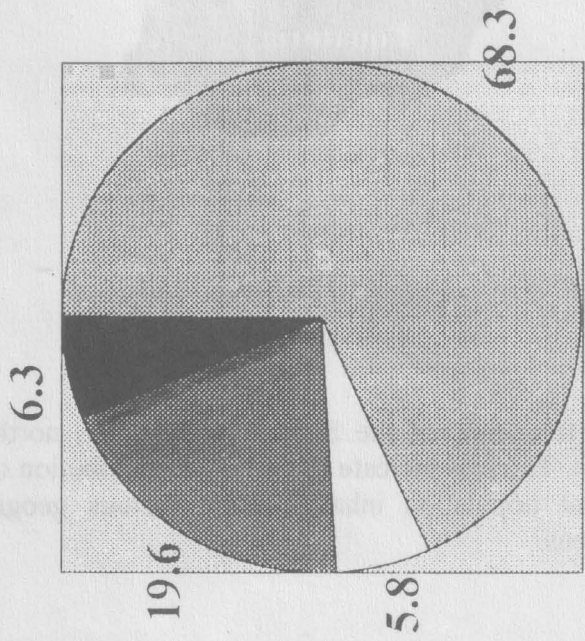
Group/Tribe % in population	Linguistic family	% in sample	Approx Total	
West	Taaysha	Afroasiatic	72	30
	Rezeigat			
	Messeyria			
	Hawazama			
	Fur	Nilo-Saharan		
Bergu				
Masalit				
Nuba				
South	Mundari	Nilo-Saharan	3.1	16
	Dinka			
	Bari	Niger Kordofanian		
	Ferteit			
Center And North	Hesanyia	Afroasiatic	30	23.8
	Bederyia			
	Riverians Kawahla			

$X^2 = 20.6$        $P = 0.000$



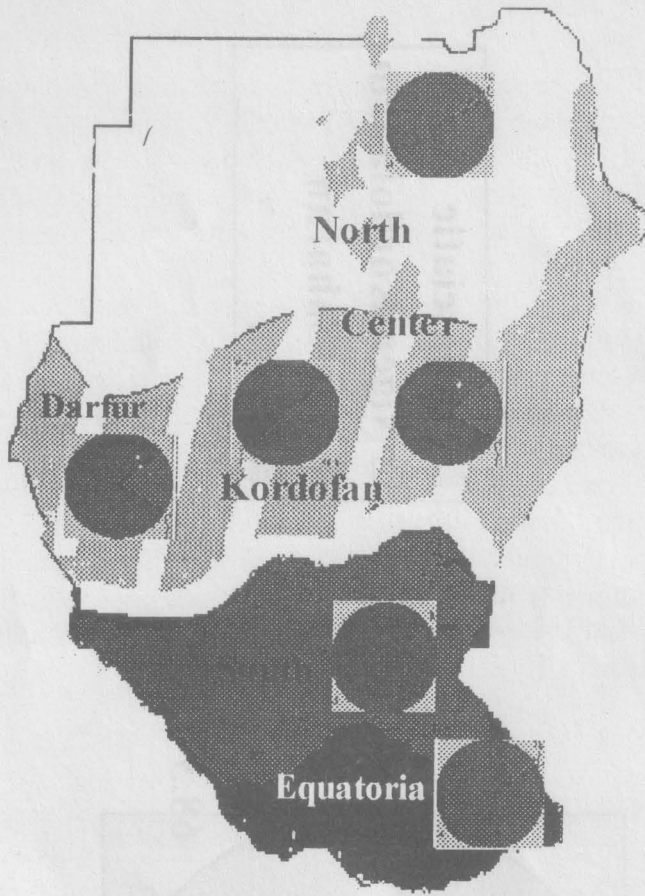
**Figure 1**

Sudanese tribes in which there are at least 5 affected individuals (families) with sickle cell disease in our sample, categorized by belonging to tribe from the mother or father's side. Lesser affected tribes are grouped together under "others" but also showing concordance between father and mother.



**Figure**

The percentage of S families from each of the three main linguistic groups present in the Sudan according to Greenberg classification as well as patients from mixed linguistic backgrounds.



**Map 1**

Map showing the frequency of the S gene in the west, north, south and center of the Sudan. The pies indicate the different distribution of the S gene within the different population inhabiting the various geographical and malaria endemic areas.



## DISCUSSION

The utility of the sickle cell trait as a population specific marker and in tracing population movements is documented (Lanclos 1991; Oner, 1992, Goncalves, 1994). Such utility is undoubtedly of relevance to study populations as diverse as those of the Sudan and to address questions related to the history of the various groups and relationship to endemic disease if any.

It is obvious from the data presented here that the sickle cell trait is clustered among the Afro-Asiatic speaking group and particularly a high prevalence is present in the nomadic Bagara tribes from Western Sudan (Darfur and Kordofan states). Those are cattle herders who migrated to the Sudan from Northern and western Africa, some entering the Sudan as recently as the Nineteenth Century from the area around Lake Chad. Another root of migration is eastern Sudan and Egypt, though probably of lesser prominence. Although Arabic is the official national language today, more than half of the population speaks other languages. Arabic became the official language and prevailed only during the few past centuries with the decline of the Nubian states and culture and the rise of religious and modern education.

Our data matched those of previous surveys in the Sudan SCA was first reported by Archibald (1926). Three foci of the disease were subsequently described Western Sudan, where a prevalence rate of up to 30.4% was reported among Messeyria of Darfur (Vbella, 1966). A survey in Kordofan reported a prevalence rate of 18% in Messeyria Humur tribe and showed that one in every 123 children born to this tribe was in danger of having the disease (Launden and Ibrahim, 1970). The other focus is in Southern where the overall prevalence was 6%, ranging from 2-11.6% in belanda tribe in Bahr Elghazal (Vella, 1964). The third focus in the Blue Nile province, central Sudan reported a prevalence ranging from 0.5% among the indigenous population and up to 16% was found among the immigrant tribes in the area (Ahmed and Baker, 1986). Although SCA was described in the North, yet a prevalence was not reported.

The Messeyria emerges unequivocally as the highest risk group, whether this is due to the size of the tribe compared to others, needs clarification and a proper population based study. The high concordance between the mothers tribe and the father tribe in our sample, indicates considerable intermarriages within tribes and probably a high consanguinity in the affected tribes. Such potential consanguinity may be a crucial factor in the augmentation of the diseases among certain groups. Bayoumi et al, (1985) however, concluded that despite the high consanguinity, there was no departure from Hardy Weinberg Equilibrium among the local tribes studied

probably due to their admixture in the area of Jabal Marra. The finding that these tribes harbored haplotypes of west African origin agrees with the fact that the Bagara Arab had followed a west African route to enter the Sudan. The open plateau of the Sahel-Sahara have always facilitated migration across the Sahel in both directions. The current demographical picture of the Sudan largely undermines the notion of south-north population fluxes and the theory of the Nile valley as corridor for population movements (Krings et al, 1999).

The role of gene flow has been highlighted by Bayoumi et al, (1985, 1987) judging from the low frequency of haemoglobin S among historically indigenous population of the Sudan several inferences could be made on the history of the sickle cell mutation and the history of falciparum malaria. We have shown recently that the haplotypes of the S gene in Sudan is more related to those of west Africa rather than the neighbouring central African haplotype (Mohammed et al, submitted). The Haj rout are the most prominent factors in shaping the population profile of the Sudan. The proportion of west African tribes from the total population rose from tens of thousands in the turn of the century to constitute nearly 10% of the total population (Balamoan, 1976, Encyclopedia Britanica, 1994). This route across the Sahel is known of acting as a medium for transmission of several microbes, most notoriously are the periodic outbreaks of meningitis of related strains in eastern and eastern Sahel (Achtman, 1998).

The Bergu, Fur and Masalit had the highest frequencies of the S gene among the "indigenous" local Sudanic Nilo-Saharan speaking groups. The fact they mostly harbor west African haplotypes and the geographic location of the tribes, makes it safe to assume a gene flow through genetic admixture with population from west Africa. In this respect it appears that the Bergu are probably the most admixed with west Africans populations, which is consistent with history. It is well established that the Bewrgu Sultanate extended through the Sahel during the 17<sup>th</sup> and 18<sup>th</sup> century.

Some of the Nilo-Saharan speaking people of the Sudan, characteristically with lower frequencies of the S gene, harbor some of the most ancient Y chromosome haplotypes (Underhill et al, 2000), thus indicating their long term settlement in the area. This fact in addition to presence of almost exclusively west African S gene haplotypes, might be an indication of the new arrival of in the area. The eastern part of the Sahel is endemic with equivalent proportions to the parasite *Plasmodium falciparum* and thus does not exclude a lower selective pressure in bringing down the frequency of the gene. All this agrees with the data suggesting a recent origin of the parasite *P. falciparum* (Volkman et al, 2001, Tishkoff et al, 2001) probably as recent as the discovery of agriculture. Natural selection

and the relatively young age of the sickle cell mutations may have collaborated to influence the distribution of the S gene and to make it a population specific marker. This fact is well documented in the literature as shown earlier.

#### Acknowledgement

We thank the staff of the Children emergency hospital for their extended help and the 189 families who donated samples to this study.

## REFERENCES

1. Achtman M(1998) Microevolution during epidemic spread of *Neisseria meningitidis* Electrophoresis, 19 (4) 593-6
2. Archibald RG.(1926) A case of sickle cell anemia in Sudan. *Tran R SOC Trop Med. Hyg* 1926, 19:389-393.
3. Balamoan GA (1976) *People and Economis in the Sudan, 1884-1956 (A study in migration policies and economics)*. White friars Press LTD, London.
4. Bayoumi RA, Saha N (1987) Some blood genetic markers of the Nuba and Hawazma tribes of Sudan western *Am J Phys Anthropol* 73, 379-88.
5. Bayoumi RA, Taha TS, (1985) A study of some genetic Characteristicsofhte Fur and Baggara tribes of the Sudan. *AMJ, Phys Anthro*, 67: 363-70.
6. *Encyclopedia Britanica (Chicago) volume 28 1994 pp 256.*
7. Goncalves MS, Nechtman JF, Figueiredo MS, Kerbauy J. Arruda VR, Sonati MF, Saad SO, Costa FF, Stoming TA. (1994) Sickle cell Disease in Brazilian population a study of  $\beta$  haplotypes. *Hum herid*, 44:322-327.
8. Greenberg JH(1955) *Studies in African Linguistic Classification* Bradfor, Conn.: Compass Publishing
9. Hewitt R, Krause A, Goldman A, Casmpbell G, Jenkins T(1996) Beta-globin haplotype analysis suggests that a major source of Malagasy ancestry is derived from Bantu-speaking Negroids. *Am J. Hum Genet*, 58, 1303-8.
10. Kamel, K. and Awny, A.Y.(1965) Origin of the sickling gene *Nature* 205, 919.
11. Krings M, Salem AE, Bauer K, Geisert H, Malek AK, Chaix L, Simon C, Welsby D, Di Rienzo A, utermann G, Sajantila A, Paabo S, Stoneking M.(1999) mtDNA analysis of Nile River Valley populations: A genetic corridor or a barrier to migration. *Am J Hum Genet* 64(4): 1166-76.
12. Kurnit, D.M.(1979) Evolution of sickle variant gene *Lancet*, 1, 104.
13. Lanclos KD, Oner C, Dimovski AJ, Gu YC, Huisman TH(1991) Sequence variations in the 5 flanking and IVS-II regions of the G gamma- and A gamma-globin genes of B chromosomes with five different haplotypes. *Blood* 77(11): 2488-96.
14. Launden RJ, Ibrahim SA,(1970) Sickling in South West Kordofan, *Sud. Med. J.* 1970, 8(4) 207-217.

15. Lehmann, H.(1954a) Distribution of the sickle cell gene. A new light on the origin of the East Africans *Eugenics Review* 46, 101-21.
16. Lehmann, H.(1954b) Origin of the sickle cell. *S. Afr J. Sci* 50, 140-1.
17. Oner C., Dimovski A, Olivieri N, Schiliro G, Codrington J. Fattoum S. et al(1992) B haplotypes in various world population's, *Hum Gent* 89:99-10.
18. Pagnier J, Mears JG, Dunda-Belkhodja O, Schaefer-Rego KE, Beldjord C, Nagel RL, Labie D;(1984) Evidence for the multicentric origin of the sickle cell hemoglobin gene in Africa *Proc natl Acad Sci USA* 81; 1771, 1984
19. Tishkoff SA, Varkonyl R, Cahinhinan N, Abbes S, Argyropoulos G, Destro-Bisol G, Drousiotou A, Dangerfield B, Lefrane G, Loiselet J., Piro A, Stoneking M, Tagarelli A, Tagarelli G, Touma EH, Williams SM, Clark AG(2001) Haplotype diversity and linkage disequilibrium at human G6PD. Recent origin of alleles that confer malaria resistance, *Science*, 293, 455-462.
20. Underhill PA, Shen P, Lin AA, Jin L, Passarino G, Yang WH, Kauffman E, Bonne-Tamir B, Bertranpetit J, Francalacci P, Ibrahim M, Jenkins T, Kidd JR, Mehdi SQ, Seielstad MT, Wells RS, Piazza A, Davis RW, Feldman MW, Cavalli Sforza LL, Oefner PJ(2000) Y chromosome sequence variation and the history of human populations. *Nat Genet*; 26(3) 358-61.
21. Vella F. (1966) Hemoglobin S and sickling in Khartoum province. *Trans Roy SOC Trop. Med Hyg*, 60:48-52.
22. Vella F.(1964) Sickling in the Western Sudan. *Sudan Med j* 1964; 3(1): 16-17.
23. Volkman SK, Barry AE, Lyons EJ, Nielsen KM, Thomas SM, Choi M, Thakore SS, Day KP, Wirth DF, hartl DL(2001) Recent origin of *Plasmodium falciparum* from a single progenitor. *Science*; 293, 482:483.