Am. J. Hum. Genet. 41:1132–1144, 1987

Globin Gene–associated Restriction-Fragment-Length Polymorphisms in Southern African Peoples

MICHÈLE RAMSAY AND TREFOR JENKINS

MRC Human Ecogenetics Research Unit, Department of Human Genetics, School of Pathology, South African Institute for Medical Research; and University of the Witwatersrand, Johannesburg

SUMMARY

The combination of polymorphic restriction-enzyme sites in the 3′ region of the β-globin gene cluster shows very little variation in southern-African Bantu-speaking black and Kalahari !Kung San populations. The sites of the 5′ region, on the other hand, show marked variation, and two common haplotypes are present—the "Negro" type (− − − − +) and the "San" type (− + − − +)—in frequencies of .404 and .106, respectively, in the Bantu-speakers and .262 and .405, respectively, in the San. Twenty of 23 β^S-associated haplotypes in southern-African Bantu-speaking black subjects were the same as that found commonly in the Central African Republic (CAR)—i.e., the "Bantu" type—a finding providing the first convincing biological evidence for the common ancestry of geographically widely separated speakers of languages belonging to the Bantu family. The ($-\alpha$) haplotype has a frequency of .21 in the Venda, .07 in both the Sotho-Tswana and the Nguni, and .06 among the !Kung San. These data are interpreted in the light of *Plasmodium falciparum* malaria selection and population movements in the African subcontinent.

INTRODUCTION

There is a paucity of written records relating to the history of the early human inhabitants of the southern part of the African continent. The reconstruction of this history is conjectural and based on data obtained from archaeological

Received January 8, 1987; revision received May 4, 1987.

Address for correspondence and reprints: Professor T. Jenkins, Department of Human Genetics, P.O. Box 1038, Johannesburg, 2000, South Africa.

sources and linguistic studies, as well as from studies exploiting classical genetic polymorphisms (Nurse et al. 1985). It is now possible to study extant populations by using DNA markers, and such studies will undoubtedly increase our understanding both of man's origin and of his early history in Africa and elsewhere.

The β-globin-associated haplotypes have been assumed to be evolutionarily neutral and were recently used to construct a phylogeny of man (Wainscoat et al. 1986*a*) that confirmed the finding of previous studies that used classical polymorphic markers (Nei and Roychoudhury 1982). These various studies showed that a primary division or split had probably occurred between African and all other human populations.

More than one-third of the peoples of Africa speak Bantu languages. The term "Bantu" was coined in the 1860s (Bleek 1862–69, quoted in Tobias 1971) to refer to the language family, and as early as 1907 Johnston suggested that the widespread distribution of the Bantu-speaking blacks in sub-Saharan Africa was the result of their differentiation from a single ancestral population (Johnston 1907). The question as to how one family of closely related languages could diffuse over such a vast area in what is thought to be a time span of only 2,000 years remains an intriguing one.

The present study is concerned with globin gene–associated restriction-fragment-length polymorphisms (RFLPs) in several southern-African populations. The main aim was to assess the affinities of these populations.

SUBJECTS AND METHODS

Subjects

The subjects were randomly selected individuals from the San (formerly Bushmen) and southern-African Bantu-speaking black populations, as well as a selected group of patients with sickle cell anemia.

The !Kung San come from Tshumkwe in northeast Namibia, and 30–40-ml venous blood samples were collected from them into ACD vacutainer tubes after voluntary consent had been obtained. Small nuclear families were used for establishing haplotypes. Blood samples were collected from random Venda individuals in the northern Transvaal and from other Bantu-speaking blacks who were resident in and around Johannesburg; Nguni speakers included Zulu, Ndebele, Swazi, and Xhosa, whereas the Sotho-Tswana group included individuals from Tswana, South-Sotho, North-Sotho, and Pedi chiefdoms.

The β^S gene is present in high frequencies in most sub-Saharan African populations, although it does not occur in most of the indigenous populations of Namibia, Botswana, and South Africa. It is present at low frequencies in chiefdoms thought to have been resident for ~1,000 years in the more northerly parts of the subcontinent, in the vicinity of the Cunene, Kavango, and Limpopo rivers. The subjects of the present study were one heterozygous and 11 homozygous Bantu-speaking black patients with sickle cell anemia who were originally from various southern African chiefdoms situated in Zimbabwe, Namibia,

Zambia, and Malawi (although they were not necessarily resident in those countries at the time of the study).

Methods

High-molecular-weight DNA was extracted from buffy-coat samples or whole blood according to the method of Sykes (1983). The DNA was digested with several restriction endonucleases according to the specifications of the manufacturers, except that 2.5 units of enzyme were used per microgram of DNA. Gel electrophoresis was performed in 0.7%–1.0% agarose (Seakem HGT) gels and transferred to nitrocellulose filters by means of the method of Southern (1975). The filters were hybridized to DNA probes that had been radiolabeled by means of nick-translation (Rigby et al. 1977) and were then autoradiographed with Kodak XAR-5 or Trimax 3M type XD X-ray plates.

The β-like globin probes included the following: JW102 β-globin cDNA (Wilson et al. 1977); JW151 γ-globin cDNA (Wilson et al. 1977); ε-globin genomic probe, the 1.3-kb *Eco*RI/*Bam*HI fragment of pε1.3 (Antonarakis et al. 1982); ψβ1 probe, the 1.7-kb *Bgl*II/*Xba*I fragment of pP3.9 (Fritsch et al. 1980); pRK29 from a region 3′ to the β-globin gene (Orkin et al. 1982); the 5′ region of the β-globin gene, the 1.8-kb *Bam*HI fragment of pSS1.8 (Geever et al. 1981); and the large intervening sequence of the β-globin gene, βIVS2 (Baird et al. 1981).

An α-globin cDNA probe, JW101 (Wilson et al. 1977), and a genomic α-globin probe, the 1.6-kb *Pst*I fragment of pDH5 (Higgs et al. 1986), were used to detect the deletion form of α-thalassemia.

RESULTS

β-*Globin Cluster*

Eleven polymorphic restriction-endonuclease sites were tested, and the frequencies of the presence of the different sites are shown in table 1. Two of the sites—the *Hpa*I site 3′ to the β-globin gene and the *Hin*cII site associated with the ε-globin gene—generate three alleles; the remainder of the sites were diallelic.

β^A-Globin haplotypes were determined for 52 randomly selected San and 48 randomly selected black chromosomes, using nuclear families of the parents and two children. Nine sites were included in the haplotype analysis. The haplotypes were analyzed with respect to the 5′ region and the 3′ region, and the results are presented in tables 2 and 3, respectively. A total of 42 5′ and 43 3′ haplotypes were unambiguously determined for the San, and in the blacks complete 5′ haplotypes were determined for 47 chromosomes and complete 3′ haplotypes for 38 chromosomes. Thirteen different haplotypes were represented among the 27 San haplotypes in which all sites were successfully investigated; and 17 were present among the 34 successfully investigated black haplotypes. The two populations had eight haplotypes in common.

The 3′ haplotype containing the 7.0-kb *Hpa*I fragment in the black individuals reported in the present study was the same in all five of those found and

TABLE 1

MEAN ± SE FREQUENCIES OF THE VARIOUS β-GLOBIN GENE CLUSTER POLYMORPHIC RESTRICTION-ENDONUCLEASE SITES IN SOUTHERN-AFRICAN POPULATIONS

RESTRICTION ENDONUCLEASE SITE[a]	POPULATION	
	Bantu-speaking Blacks	!Kung San
*Hinc*II-5′ε	.167 ± .054 (48)[b]	.059 ± .033 (51)
*Hin*dIII-^{G}γ	.438 ± .005 (48)	.680 ± .066 (50)
*Taq*I-inter γ	.773 ± .086 (24)	.739 ± .065 (46)
*Hin*dIII-^{A}γ	.146 ± .051 (48)	.080 ± .038 (50)
*Hinc*II-ψβ1	.106 ± .045 (47)	.275 ± .063 (51)
*Hinc*II-3′ψβ1	.872 ± .049 (47)	.980 ± .020 (50)
*Hinf*I-5′β	.833 ± .108 (12)	.813 ± .097 (16)
*Ava*II-β	.933 ± .037 (45)	.760 ± .060 (50)
*Hpa*I-3′β	.762 ± .119 (42)[c]	1.000
*Hin*dIII-3′β	.563 ± .072 (48)	.765 ± .060 (51)
*Bam*HI-3′β	.826 ± .056 (46)	.788 ± .057 (52)

NOTE.—Numbers in parentheses are number of chromosomes tested.
[a] Shown in fig. 1.
[b] Three alleles = 14.0 kb 5′ε *Hinc*II .042 ± .029 (48).
[c] Three alleles = 7.0 kb 3′β *Hpa*I .119 ± .050 (42).

was (+ * − +) (table 3). The 5′ haplotype differed, however; four were (− − − − +) and one was (− + + − +), indicating that crossovers had occurred subsequent to the mutation that gave rise to the 7.0-kb fragment. The 3′ haplotype containing the 13.0-kb fragment was the same in all five cases tested and was (+ − − +). Again the 5′ haplotypes differed; three were (+ − − − −), one was (+ − − − +), and in one only the first three sites were deduced (+ − − nn) (n = not determined).

The *Taq*I inter-γ-globin RFLP site (Wainscoat et al. 1986*b*) was present at a mean ± SE frequency of .78 ± .05 (N = 69) in the San and .69 ± .07 (N = 42) in the Bantu-speaking black population. This site is polymorphic only in African populations, and the rarer allele (−) occurred in association with the (− − − − +) 5′ haplotype in all cases.

β^S-Associated haplotypes were determined using the 11 polymorphic restriction-endonuclease sites shown in figure 1. Twenty of the 23 β^S-associated haplotypes were the same as the CAR type (Pagnier et al. 1984)—or, in cases in which incomplete haplotypes were obtained, the sites that were determined were compatible with it and are referred to as the "Bantu" type. The remaining three were rare haplotypes. These data, together with β^S-associated haplotype data on African populations, are shown in table 4. The *Xmn*I site 158 bp 5′ to the ^{G}γ-globin gene was consistently absent.

α-*Globin Cluster*

The mean ± SE frequency of the (−α) haplotype was found to be .059 ± .016 in the !Kung San, .206 ± .049 in the Venda, .066 ± .024 in the Sotho-Tswana, and .068 ± .022 in the Nguni (table 5). The San, Sotho-Tswana, and

TABLE 2

NUMBER AND MEAN ± SE FREQUENCIES OF THE 5′ REGION OF THE β-GLOBIN CLUSTER HAPLOTYPE IN SOUTHERN-AFRICAN POPULATIONS

5′ Haplotype Site[a]					Population			
					Bantu-speaking Blacks		!Kung San	
1	3	5	6	7	No.	Frequency	No.	Frequency
−	−	−	−	+	19	.404 ± .072	11	.262 ± .068
−	+	−	−	+	5	.106 ± .045	17	.405 ± .076
+	−	−	−	−	6	.128 ± .049	1	.024 ± .024
−	+	+	−	+	7	.149 ± .052	4	.095 ± .045
−	+	−	+	+	5	.106 ± .045	8	.190 ± .061
0	+	−	−	+	2	.043 ± .030		
+	−	−	−	+	1	.021 ± .021		
−	+	−	−	−	2	.043 ± .030		
−	−	−	+	+			1	.024 ± .024
Total					47		42	

NOTE.—The black and San populations are significantly different from one another ($\chi^2_{[8]} = 19.54$; $P < .02$).

[a] RFLP sites: 1, ε-globin *Hin*cII; 3, ^{G}γ-globin *Hin*dIII; 5, ^{A}γ-globin *Hin*dIII; 6, ψβ1 *Hin*cII; and 7, 3′ψβ1 *Hin*cII. 0 Indicates presence of the 14.0-kb *Hin*cII fragment.

Nguni frequencies were not significantly different from one another, but each differed significantly from that of the Venda ($\chi^2_{[3]} = 13.13$, $P < .01$; $\chi^2_{[3]} = 7.84$, $P < .05$; and $\chi^2_{[3]} = 12.41$, $P < .01$, respectively).

A total of 20 ($-\alpha$) haplotypes—including eight San, three Venda, four Sotho-Tswana, and five Nguni—were shown to possess a 17.0-kb *Bgl*II fragment, indicating that the deletion was of the $-\alpha^{3.7}$ variety.

The ααα haplotype was detected in the heterozygous state in two Venda, two

TABLE 3

NUMBER AND MEAN ± SE FREQUENCIES OF THE 3′ REGION OF THE β-GLOBIN CLUSTER HAPLOTYPE IN SOUTHERN-AFRICAN POPULATIONS

3′ Haplotype Site[a]				Population			
				Bantu-speaking Blacks		!Kung San	
8	9	10	11	No.	Frequency	No.	Frequency
+	+	+	+	15	.395 ± .079	22	.512 ± .076
+	+	+	−	7	.184 ± .063	11	.256 ± .066
+	−	−	+	5	.131 ± .055		
+	+	−	+	3	.079 ± .044	2	.046 ± .032
+	*	−	+	5	.132 ± .055		
−	+	−	+	3	.079 ± .044	6	.140 ± .053
−	+	+	+			2	.046 ± .032
Total				38		43	

NOTE.—The black and San populations are significantly different from one another ($\chi^2_{[6]} = 15.16$; $P < .01$).

[a] RFLP sites: 8, β-globin *Ava*II; 9, 3′β-globin *Hpa*I; 10, 3′β-globin *Hin*dIII; and 11, 3′β-globin *Bam*HI. The asterisk indicates the presence of the 7.0-kb *Hpa*I fragment.

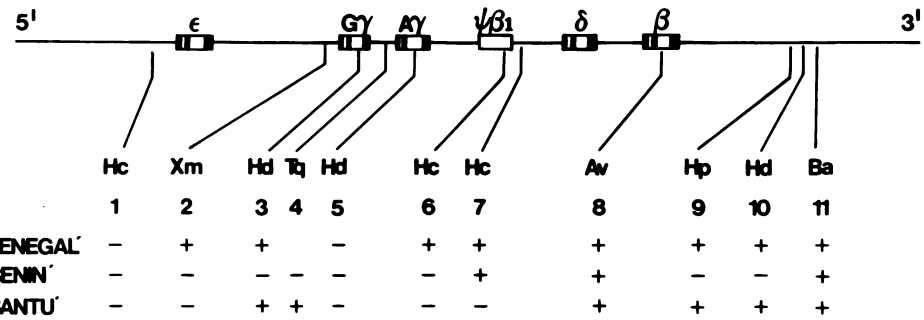

FIG. 1.—Map of the β-globin gene cluster on chromosome 11, showing the African β^S-associated haplotypes. Hc = *Hinc*II; Xm = *Xmn*I; Hd = *Hin*dIII; Tq = *Taq*I; Av = *Ava*II; Hp = *Hpa*I; and Ba = *Bam*HI. The *Taq*I site has not been determined for the "Senegal" haplotype. For β^S-bearing chromosomes, only the "Bantu" haplotype was present in our sample.

Sotho-Tswana, and two San, giving it mean ± SE frequencies of .030 ± .023, .019 ± .013, and .010 ± .007, respectively. The two San ααα haplotypes are of the $\alpha\alpha\alpha^{\text{anti-3.7}}$ type, but both of the Venda and one of the Sotho-Tswana ααα haplotypes are not typical of either the $\alpha\alpha\alpha^{\text{anti-3.7}}$ or the $\alpha\alpha\alpha^{\text{anti-4.2}}$ and have been designated $\alpha\alpha\alpha^{\text{anti-3.7}}$ *Bgl*II(−) (Ramsay and Jenkins 1985*a*). The other Sotho-Tswana ααα haplotype was investigated with only *Bam*HI, so it is not possible to infer the crossover event giving rise to it.

DISCUSSION

Globin gene–associated RFLPs were studied in several southern-African populations to assess the affinities both among them and between them and

TABLE 4

β^S-ASSOCIATED HAPLOTYPES FOUND IN VARIOUS AFRICAN OR AFRICAN-DERIVED POPULATIONS

POPULATION	HAPLOTYPES				
	Total	"Senegal"	"Benin"	"Bantu"	Rare
West Africa:					
Senegal[a]	56	46	8	0	2
Central West Africa:					
Benin[a]	20	0	20	0	0
Algeria[a]	20	0	20	0	0
Nigeria[b]	34	0	33	0	1
Bantu-speaking Africa:					
CAR[a]	28	0	2	24	2
Southern Africa[c]	23	0	0	20	3
American blacks[d]	76	5	45	17	9
Jamaican blacks[d]	94	6	63	15	10
Saudi Arabia:					
Riyadh[b]	22	0	17	0	5

[a] Source: Pagnier et al. 1984.
[b] Source: Kulozik et al. 1986.
[c] Source: Present study.
[d] Source: Antonarakis et al. 1984.

TABLE 5

α-GLOBIN HAPLOTYPES FOUND IN VARIOUS SOUTHERN-AFRICAN POPULATIONS

POPULATION	NO. OF INDIVIDUALS	α-GLOBIN GENOTYPE				HAPLOTYPE FREQUENCY		
		αα/αα	−α/αα	−α/−α	ααα/αα	(αα)	(−α)[a]	(ααα)
San:								
!Kung	101	88	10	1	2[b]	.931	.059	.010
Southern-African blacks:								
Venda	34	20	10	2	2[c]	.764	.206	.030
Sotho-Tswana	53	45	5	1	2[c]	.916	.066	.018
Nguni	66	58	7	1	0	.932	.068	.000

[a] Of the $-\alpha^{3.7}$ type.

[b] Of the $\alpha\alpha\alpha^{anti\text{-}3.7}$ type.

[c] Both Venda and one of the Sotho-Tswana (ααα) haplotypes were of the $\alpha\alpha\alpha^{anti\text{-}3.7}$ *Bgl*II (−) type (Ramsay and Jenkins 1985*a*).

other world populations. Neutral markers—e.g., the β^A-associated haplotypes—provide data that can be used to calculate genetic distances, whereas other markers—i.e., the β^S-associated haplotypes and the α-thalassemia determinants—are ecosensitive and must be used with caution when studying population affinities.

When compared with those of other populations, β^A-associated RFLPs in southern-African populations revealed the following differences: (1) the *Hin*cII, ε-globin site is present at high frequency in Mediterranean and Asian populations but at a much lower frequency in blacks and the San; (2) the *Hin*dIII, $^G\gamma$-globin site has higher frequencies in blacks than in Caucasoids and attains its highest known frequency in the San; (3) the frequency of the *Hin*cII site 3′ to ψβ1 is strikingly higher in blacks and the San than in other populations; and (4) the *Hpa*I RFLP 3′ to the β-globin gene is present in the San and southern-African Bantu-speaking blacks (Ramsay and Jenkins 1985*b*). The 7.0-kb fragment attains its highest frequency in the Venda (mean ± SE = .167 ± .062) but has not been found at all in the San. The reason for the high frequency in the Venda is not known.

Random association occurs between the 3′ and 5′ regions of the haplotype, and it has been proposed that the recombination frequency between these regions is higher than that within them (Chakravarti et al. 1984). Population differences are not very striking when the 3′ region of the haplotype is considered (table 3). The (+ + + +) 3′ haplotype is the most common, whereas the 3′ haplotypes (+ + + −) and (− + − +) are present at lower frequencies in all the populations that have been studied. The Bantu-speakers are distinguished from other populations because of the higher frequencies of the 7.0- and 13-kb *Hpa*I fragments. These fragments are present at significantly lower frequencies in American blacks and in the San and are absent in non-African populations.

Striking population differences are, however, evident in the 5′ region of the haplotype. In peoples of non-African descent, three haplotypes account for the vast majority of haplotypes (Wainscoat et al. 1986*a*). These are (+ − − − −), (− + − + +), and (− + + − +), with the first having the highest frequency in all of them. Two different haplotypes have been shown to be the most common in Africa: (− − − − +) and (− + − − +), which will be referred to as the "black" and "San" types, respectively, because the former is the commonest in the black and the latter the commonest in the San. The black haplotype occurs at a frequency of .400 in the Bantu-speaking blacks and is present at a frequency of .262 in the San. In contrast, the San type occurs at a frequency of .381 in the San and is present at a frequency of .116 in the Bantu-speaking blacks and .207 in American blacks. The southern-African Bantu-speaking population shows, in fact, a distribution of the 5′ haplotype very similar to that of the American-black population studied by Antonarakis et al. (1984). The presence at low frequencies of the San haplotype in the blacks and the presence of the black haplotype in the San may be the result of either common ancestry or recent gene flow. There are at least two differences between the African haplotypes: the extended black haplotype is (− − − − − +), whereas the San type is (− + + − − +), where the third symbol represents the *Taq*I inter-γ-globin

RFLP. The presence of the (− + − − +) haplotype at a frequency of .103 in American blacks suggests that its frequency in southern-African blacks is not the result of admixture with the San but that it is more likely to be a common ancestral African 5′ haplotype.

The three 5′ haplotypes (+ − − − −), (− + + − +), and (− + − + +) are, for the most part, present in all populations and presumably predate raciation. In contrast, the two African 5′ haplotypes, (− − − − +) and (− + − − +), are not found in non-African populations and presumably were lost by the small founder population that migrated out of Africa into Asia and Europe. Where the (− − − − +) haplotype occurs in Melanesian and Polynesian populations, it is clearly distinguished from the African type by the presence of the *Taq*I inter-γ-globin site (Wainscoat et al. 1986*b*). The 5′ β-globin cluster has been used to study the evolution of modern man by Wainscoat et al. (1986*a*), who showed that the data were consistent with the theory that modern man evolved in Africa, that a founder population migrated from Africa and subsequently gave rise to all the non-African populations. Using the 5′-region haplotypes of the present study and others (Antonarakis et al. 1984; Old et al. 1984; Maggio et al. 1986; Wainscoat et al. 1986*a*), we have calculated genetic distances and have used these to construct a dendrogram. We have shown a similar clustering of all the non-African populations that are distinctly separated from the cluster of African populations. These findings will be presented elsewhere.

Jones and Rouhani (1986), however, caution that "phylogeny of individual genes may well be a misleading indication of phylogeny of populations from which they originate," and Giles and Ambrose (1986) point out that the data of Wainscoat et al. (1986*a*) are equally compatible with a European origin of man and that selection together with random drift rather than chance events alone may be a better model for explaining the results. Modern man may well have originated in Africa, but that does not mean that the present peoples on that continent have descended from the relict populations left behind when the ancestors of the rest of humanity migrated from Africa. The presence of the *Taq*I inter-γ-globin RFLP in the San and southern-African blacks (Wainscoat et al. 1986*b*) lends further support to the now generally accepted view that both populations share a common ancestry.

β^S-Associated-haplotype studies have contributed to the understanding of the origin—or, as it now appears, multiple origins—of this mutation on the African continent. Three distinct β^S-associated haplotypes have been found at high frequency in three separate geographical regions: (1) the "Senegal" haplotype in Atlantic West Africa, (2) the "Benin" haplotype in Central West Africa, and (3) the "Bantu" haplotype in the CAR (fig. 1) (Pagnier et al. 1984). The distribution of the β^S-associated haplotypes in Africa is shown in table 4 and figure 2, from which it can be seen that in both the northern and southern extremities of Bantu-speaking Africa the β^S-allele is embedded in the same haplotype. It is postulated, therefore, that the sickle cell mutation arose only once in the Bantu-speakers, presumably in their nuclear area of origin, before the "Bantu expansion" occurred ~2,000 years ago. The common biological

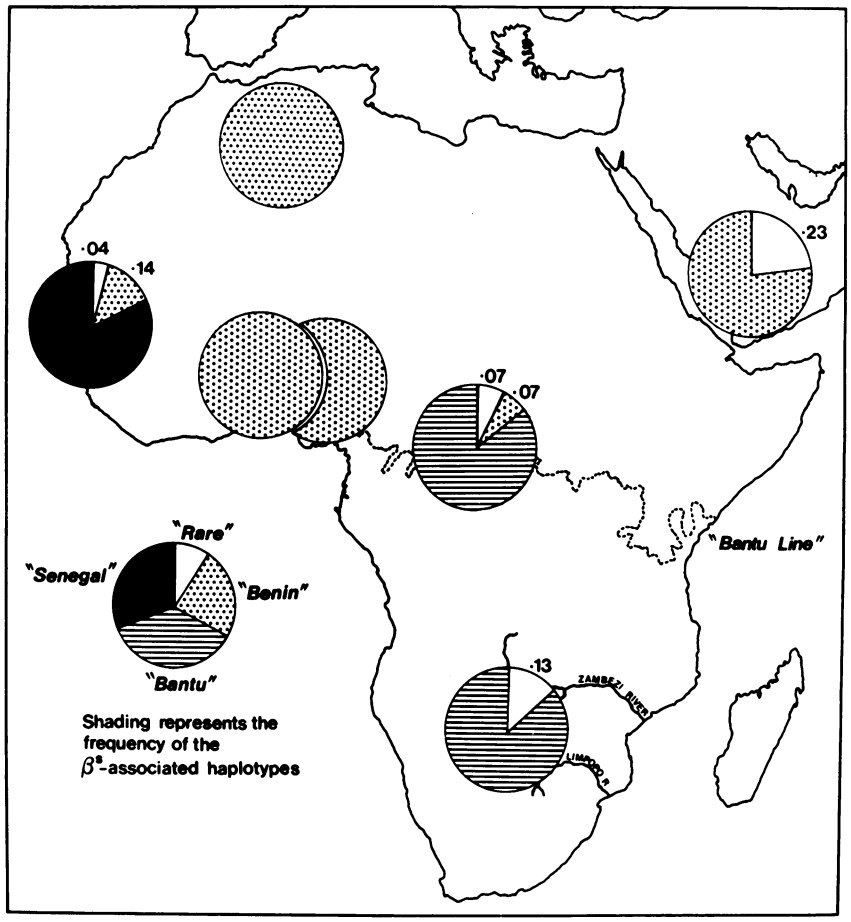

FIG. 2.—Distribution of β^S-associated haplotypes in Africa. The "Bantu Line" is according to Seligman (1961).

origin of the Bantu-speakers that has been suggested by linguistic studies is thus confirmed by these findings.

Little is known about the selective pressures that are responsible for the production and maintenance of the α-thalassemia polymorphisms. It has, however, long been suggested that *Plasmodium falciparum* malaria may be the responsible selective agent, because α^+-thalassemia occurs at its highest frequencies in areas where malaria is hyperendemic (Luzzatto 1979; Oppenheimer et al. 1984; Ramsay and Jenkins 1984). The recent studies in island Melanesia provide the most convincing evidence to date in favor of this hypothesis (Flint et al. 1986). The present study shows that the Venda, who are the most northerly population in South Africa, have a higher frequency of $(-\alpha)$—namely, .21—than do the more southerly Nguni and Sotho-Tswana populations, both of whom have $(-\alpha)$ at a frequency of .07. The presence of the $(-\alpha)$ haplotype in

the San may be explained by proposing that it was present in the ancestral population that extended from the horn of Africa to the Cape of Good Hope and that the population had, in fact, been exposed to malaria infection. The α^+-thalassemia mutation had occurred by then and might have attained the frequency of .06 in the extant !Kung hunter-gatherers found by the present study.

The $\alpha\alpha\alpha$ haplotype in the San ($\alpha\alpha\alpha^{anti\text{-}3.7}$) differs from that in both the Venda and a Sotho-Tswana individual, $\alpha\alpha\alpha^{anti\text{-}3.7}$ *Bgl*II (−), by a single site mutation and is likely a relatively recent mutational event.

The most commonly cited alleles conferring resistance to *P. falciparum* malaria are $Hb\beta^S$, the glucose-6-phosphate dehydrogenase alleles Gd^A and Gd^{A-}, and, more recently, the $(-\alpha)$ haplotype. The decreasing frequency of the $(-\alpha)$ haplotype in the more southerly Bantu-speakers parallels the decrease in Gd^{A-} frequency in these populations and may indicate that the two alleles are of similar antiquity. Gd^A is the most likely candidate for being the oldest *P. falciparum* malaria–protective allele, with Gd^{A-} and α^+-thalassemia being the next oldest, and β^S, which does not occur at all in South-African chiefdoms, being the most recent *P. falciparum* malaria–protective allele in Africa.

The globin gene–associated RFLPs reported in the present paper are unlikely to solve the outstanding problems concerning man's origins and evolution in Africa; answers to these problems will require detailed studies of many more neutral markers from different regions of the genome, and such studies will have to be carried out on many more indigenous populations.

ACKNOWLEDGMENTS

We are most grateful to the San, Venda, and other Bantu-speaking peoples for their willing participation in this project; to Thomas Ramuthaga, who was our guide and interpreter in Vendaland; and to K. Gerdes, commissioner of Tshumkwe, Namibia. Blood samples from the group of patients with sickle cell anemia and from their families were sent to us by Dr. G. Makoni, with the invaluable help of Sister Catherine Mundy and S. M. Magada of Wankie Colliery Hospital, Zimbabwe; by Tony Marinaki from the SAIMR (now SWA State Hospital) laboratory in Rundu, Namibia; by Linda Dodds from the SAIMR (now SWA State Hospital) laboratory in Windhoek, Namibia; and by L. G. MacDougal from Baragwanath Hospital in Johannesburg. We are very grateful to the numerous colleagues who so generously donated the globin probes. Our thanks are extended to David Dunn for help with the analysis of the data, to Bridget Hensman for excellent technical assistance, to Ardyn Erasmus for drawing the maps, and to Shirley Alper for the preparation of the manuscript.

REFERENCES

Antonarakis, S. E., C. D. Boehm, P. J. V. Giardina, and H. H. Kazazian, Jr. 1982. Nonrandom association of polymorphic restriction sites in the β^S-globin gene cluster. Proc. Natl. Acad. Sci. USA **79**:137–141.

Antonarakis, S. E., C. D. Boehm, G. R. Serjeant, C. E. Theisen, G. J. Dover, and H. H. Kazazian, Jr. 1984. Origin of the β^S-globin gene in blacks: the contribution of recurrent mutation or gene conversion or both. Proc. Natl. Acad. Sci. USA **81**:853–856.

Baird, M., C. Driscoll, H. Schreiner, G. V. Sciaratta, G. Sansone, G. Niazi, F. Ramirez, and A. Bank. 1981. A nucleotide change at a splice junction in the human β-globin gene is associated with β^{O}-thalassemia. Proc. Natl. Acad. Sci. USA **78:** 4218–4221.

Bleek, W. H. I. 1862–69. A comparative grammar of South African languages. Juta, Cape Town (quoted in Tobias 1971).

Chakravarti, A., K. H. Buetow, S. E. Antonarakis, P. G. Waber, C. D. Boehm, and H. H. Kazazian, Jr. 1984. Non uniform recombination within the human β-globin gene cluster. Am. J. Hum. Genet. **36:**1239–1258.

Flint, J., A. V. S. Hill, D. K. Bowden, S. J. Oppenheimer, P. R. Sill, S. W. Serjeantson, J. Bana-Koiri, K. Bhatia, M. P. Alpers, A. J. Boyce, D. J. Weatherall, and J. B. Clegg. 1986. High frequencies of α-thalassaemia are the result of natural selection by malaria. Nature **321:**744–750.

Fritsch, E. F., R. M. Lawn, and T. Maniatis. 1980. Molecular cloning and characterization of the human β-like globin gene cluster. Cell **19:**959–972.

Geever, R. F., L. B. Wilson, F. S. Nallaseth, P. F. Milner, M. Bittner, and J. T. Wilson. 1981. Direct identification of sickle cell anemia by blot hybridization. Proc. Natl. Acad. Sci. USA **78:**5081–5085.

Giles, E., and S. H. Ambrose. 1986. Are we all out of Africa? Nature **322:**21–22.

Higgs, D. R., J. S. Wainscoat, J. Flint, A. V. Hill, S. L. Thein, R. D. Nicholls, H. Teal, H. Ayyub, T. E. Peto, and A. G. Falusi. 1986. Analysis of the human alpha-globin gene cluster reveals a highly informative genetic locus. Proc. Natl. Acad. Sci. USA **83:**5165–5169.

Johnston, H. H. 1907. The origin of the Bantu. J. Afr. Soc. **6:**329–340.

Jones, J. S., and S. Rouhani. 1986. Human evolution: how small was the bottle neck? Nature **319:**449–450.

Luzzatto, L. 1979. Genetics of red cells and susceptibility to malaria. Blood **54:**961–976.

Maggio, A., S. Acuto, P. Lo Gioco, R. Di Marzo, A. Giambona, P. Sammarco, and F. Caronia. 1986. β^{A} and β^{thal} DNA haplotypes in Sicily. Hum. Genet. **72:**229–230.

Nei, M., and A. K. Roychoudhury. 1982. Genetic relationship and evolution of human races. Evol. Biol. **14:**1–56.

Nurse, G. T., J. S. Weiner, and T. Jenkins. 1985. The peoples of Southern Africa and their affinities. Clarendon, Oxford.

Old, J. M., M. Petrou, B. Modell, and D. J. Weatherall. 1984. Feasibility of antenatal diagnosis of β thalassemia by DNA polymorphisms in Asian Indian and Cypriot populations. Br. J. Haematol. **57:**255–263.

Oppenheimer, S. J., D. R. Higgs, D. J. Weatherall, J. Barker, and R. A. Spark. 1984. α-Thalassaemia in Papua New Guinea. Lancet **1:**424–426.

Orkin, S. H., H. H. Kazazian, Jr., S. E. Antonarakis, S. C. Goff, C. D. Boehm, J. P. Sexton, P. G. Waber, and P. J. V. Giardina. 1982. Linkage of β-thalassaemia mutations and β-globin gene polymorphisms with DNA polymorphisms in the human β-globin gene cluster. Nature **296:**627–631.

Pagnier, J., J. G. Mears, O. Dunda-Belkhodja, K. E. Schaefer-Rego, C. Beldjord, R. L. Nagel, and D. Labie. 1984. Evidence for the multicentric origin of the sickle cell hemoglobin gene in Africa. Proc. Natl. Acad. Sci. USA **81:**1771–1773.

Ramsay, M., and T. Jenkins. 1984. α-Thalassaemia in Africa: the oldest malaria protective trait? Lancet **2:**410.

———. 1985*a*. The $\alpha\alpha\alpha^{anti\text{-}3.7}$ globin haplotype with an additional Bgl II site mutation [$\alpha\alpha\alpha^{anti\text{-}3.7}$ Bgl II(−)]. Hemoglobin **9:**385–391.

———. 1985*b*. The Hpa I β-globin DNA restriction fragment length polymorphism in Southern Africa. S. Afr. J. Sci. **81:**98–100.

Rigby, P. W. J., M. Dieckmann, C. Rhodes, and P. Berg. 1977. Labeling deoxyribonucleic acid to high specific activity *in vitro* by nick translation with DNA polymerase I. J. Mol. Biol. **113:**237–251.

Seligman, C. G., ed. 1961. Races of Africa. 3d ed. Oxford University Press, Oxford.

Southern, E. M. 1975. Detection of specific sequences among DNA fragments separated by gel electrophoresis. J. Mol. Biol. **98**:503–517.

Sykes, B. G. 1983. DNA in heritable disease. Lancet **2**:787–788.

Tobias, P. V. 1971. The biological invalidity of the term Bantu. S. Afr. J. Sci. **67**:517–520.

Wainscoat, J. S., A. V. S. Hill, A. L. Boyce, J. Flint, M. Hernandez, S. L. Thein, J. R. Old, J. R. Lynch, A. G. Falusi, D. J. Weatherall, and J. B. Clegg. 1986*a*. Evolutionary relationships of human populations from an analysis of nuclear DNA polymorphisms. Nature **319**:491–493.

Wainscoat, J. S., A. E. Kulozik, M. Ramsay, A. G. Falusi, and D. J. Weatherall. 1986*b*. A Taq I γ-globin DNA polymorphism: an African specific marker. Hum. Genet. **74**:90–92.

Wilson, J. T., L. B. Wilson, J. K. Deriel, L. Villa-Komaroff, A. Efstratiadis, B. G. Forget, and S. M. Weissman. 1977. Insertion of synthetic copies of human globin genes into bacterial plasmids. Nucleic Acids Res. **5**:563–581.