Proc. Natl. Acad. Sci. USA
Vol. 85, pp. 4431–4435, June 1988
Genetics

Structural analysis of the 5′ flanking region of the β-globin gene in African sickle cell anemia patients: Further evidence for three origins of the sickle cell mutation in Africa

(hemoglobinopathies/DNA polymorphisms/evolution)

YAHIA CHEBLOUNE*†, JOSÉE PAGNIER‡, GUY TRABUCHET*, CLAUDINE FAURE*, GÉRARD VERDIER*, DOMINIQUE LABIE§, AND VICTOR NIGON*

*Département de Biologie Générale et Appliquée Unité Associée 92, Université Claude Bernard-Lyon I, 43 Boulevard du 11 Novembre 1918, 69622 Villeurbane Cedex, France; ‡Institut National de la Santé et de la Recherche Médicale Unité 299, Hopital de Bicêtre, 94275 Le Kremlin-Bicêtre, France; and §Institut National de la Santé et de la Recherche Médicale Unité 15, 24 rue du Faubourg Saint-Jacques, 75014 Paris, France

Communicated by Philip Leder, February 19, 1988

ABSTRACT **Haplotype analysis of the β-globin gene cluster shows two regions of DNA characterized by nonrandom association of restriction site polymorphisms. These regions are separated by a variable segment containing the repeated sequences $(ATTTT)_n$ and $(AT)_xT_y$, which might be involved in recombinational events. Studies of haplotypes linked to the sickle cell gene in Africa provide strong argument for three origins of the mutation: Benin, Senegal, and the Central African Republic. Nevertheless, the haplotype determination does not give any information about the variable segment and does not totally exclude the possibility of recombination leading to different haplotypes linked to the mutation. The structure of the variable segment in the three African populations was studied by S1 nuclease mapping of genomic DNA, which allows a comparison of several samples. A 1080-base-pair DNA segment was sequenced for one sample from each population. S1 nuclease mapping confirmed the homogeneity of each population with regard to both $(ATTTT)_n$ and $(AT)_xT_y$ repeats. We found three additional structures for $(AT)_xT_y$ correlating with the geographic origin of the patients. Ten other nucleotide positions, 5′ and 3′ to the $(AT)_xT_y$ copies, were found to be variable when compared to homologous sequences from human and monkey DNAs. These results allow us to propose an evolutionary scheme for the polymorphisms in the 5′ flanking region of the β-globin gene. The results strongly support the hypothesis of three origins for the sickle mutation in Africa.**

Haplotype analysis of the β-globin gene cluster shows two regions of DNA characterized by nonrandom association of restriction site polymorphisms, as first defined by Antonarakis *et al.* (1). One extends over 38 kilobases (kb) 5′ to the δ gene, and the other (3′ to the δ gene) includes the β gene and extends 20 kb downstream. Between these two regions lies a 9-kb DNA segment, which is immediately 5′ to the β-globin gene and includes the δ gene. In this segment, restriction polymorphism sites show random association with both 5′ and 3′ sides of the β-globin gene cluster, which leads to no significant linkage disequilibrium between the 5′ and 3′ sides. This segment is assumed to be a "hot spot" for recombination (1). By population genetic analysis, Chakravarti *et al.* (2) estimated that the recombination rate was 3–30 times higher than that expected for a segment of this size, compared to surrounding DNA sequences. Some curious features are apparent in this 9-kb region. There are areas of tandemly repeated sequences, among them $(ATTTT)_n$ and $(AT)_xT_y$, located 1.4 and 0.5 kb 5′ to the cap site, respectively. The ATTTT motif has been found in four, five, and six copies (3). Three structures of $(AT)_xT_y$ have been described in human DNA: $(AT)_7T_7$, $(AT)_9T_5$, and $(AT)_{11}T_3$ (4–7). When these sequences in human and monkey DNAs are compared, it is obvious that they display a high degree of variability (3–9). These repeats may participate at a high frequency in recombination events, and an example of such a recombination occurring in this DNA segment has already been described (10).

In contrast to the normal β^A chromosome, the 5′ and 3′ subhaplotypes linked to β-gene mutations were found to be nonrandomly associated (11–15). The sickle mutation is associated with more than 30 haplotypes, but most of these are rare. Three types are largely predominant in Black populations (13–15) and are found with a high degree of homogeneity and specificity in three separate areas of Africa (15). The geographical specificity of the three major haplotypes strongly supports the hypothesis of three origins of the mutation. However, this conclusion may be disputed for three reasons: (*i*) the distance between the restriction polymorphic sites defining the haplotype is long enough to allow recombination events; (*ii*) the "framework," defined for normal β-globin genes by polymorphisms and nucleotide sequence data (11), is identical in the three cases (framework 1); and (*iii*) no difference other than the *Hpa* I polymorphism was observed in the 3′ subhaplotype between Benin chromosomes on the one hand and Senegal and Central African ones on the other hand. Senegal and Central African haplotypes can be derived from each other by a single crossing-over event leading to the hypothesis of only two different origins. Moreover, most of the minor haplotypes associated with the mutation in American and Jamaican Black populations could be accounted for by one or two recombination events in the characteristic segment 5′ to the β-globin gene (13). Thus, the possibility of recombination events leading to the three major haplotypes cannot be totally excluded.

We have previously devised a technique of S1 nuclease mapping of genomic DNA suitable for studying polymorphisms deriving from $(ATTTT)_n$ and $(AT)_xT_y$ repeats (6, 16). In an attempt to obtain more information about the origin of the sickle mutation and possible recombinational events within this DNA segment, we first performed S1 nuclease genomic mapping on the DNA of several sickle cell anemia (SCA) patients originating from three different regions of Africa. We found three additional structures of the $(AT)_xT_y$ repeat, which correlate with the geographical origin of the subjects. One β^S chromosome of each origin was analyzed by sequencing, and ten more nucleotide positions were found to

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Abbreviations: SCA, sickle cell anemia; CAR, Central African Republic.

†To whom reprint requests should be addressed.

Table 1. Sequences of the characteristic segments in the single-stranded probes used for S1 nuclease mapping

Probe	Structure of the repetitive segments
ss1 (P9βN2)	$(ATTTT)_4$ $(AT)_7T_7$
ss2 (P9βT2)	$(ATTTT)_5$ $(AT)_{11}T_3$
ss3 (P8βT1)	$(ATTTT)_5$ $(AT)_7T_7$
ss4 (MβSen)	$(ATTTT)_4$ $(AT)_8T_9$

be variable. These results constitute a confirmation of the previous hypothesis of three origins for the sickle mutation (15).

MATERIAL AND METHODS

Source and Preparation of DNA. The patients were selected from previously reported groups of SCA individuals; they were homozygotes for the S mutation and for the associated specific haplotype. They originated from three different countries in Africa: 13 from Benin, 12 from Senegal, and 15 from the Central African Republic (CAR) (15). DNAs were prepared from leukocytes by standard methods as described (17).

S1 Nuclease Mapping of Genomic DNA. Ten to 20 μg of genomic DNA was denatured for 2 hr at 37°C in 0.2 M NaOH, neutralized, and ethanol precipitated. The pellet was dissolved in 30 μl of deionized formamide. The DNA was then annealed to 50–200 ng of a single-stranded DNA probe in a final volume of 40 μl of 40 mM Tris·HCl, pH 7.9/600 mM NaCl/4 mM EDTA. The DNA solution was incubated 5–10 min at 85°C and was quickly cooled in ice before overnight incubation at 37°C. The heteroduplex mixture was diluted 1:10 in 30 mM sodium acetate, pH 4.5/200 mM NaCl/4 mM $ZnSO_4$ and was incubated for 1 hr at 37°C with 20–50 units of S1 nuclease (Bethesda Research Laboratories). DNA was then ethanol precipitated in the presence of 5 μg of sonicated salmon sperm DNA as carrier (6, 16). The resulting fragments were separated on an agarose gel and transferred onto a nitrocellulose filter (18); bands were detected by autoradiography after hybridization with a ^{32}P-labeled β-globin probe as previously described (17). The position of putative S1-nuclease-sensitive sites and the sizes of the expected fragments in the different cases are illustrated in Figs. 1*C* and 2*B*.

Double-Stranded DNA Clones. The 4.4-kb β *Pst* I clone was obtained from T. Maniatis (19). The 5.5-kb βN1 clone was obtained from a normal subject (6) and the 5.5-kb βT1 and βT2 clones were obtained from each of the thalassemic chromosomes from an Algerian patient (6). In the present work, the 5.5-kb *Eco*RI DNA fragment from one SCA patient of each origin was cloned in λ Wes vector. Each of the cloned DNA fragments was isolated from its cloning vector by appropriate endonuclease digestion and purified by either sucrose gradient centrifugation or electroelution from a preparative agarose gel.

Single-Stranded Probes. The 3.7-kb *Pst* I/*Eco*RI and 2-kb *Bam*HI fragments were isolated from the double stranded clones and transferred into the single-stranded vectors pEMBL8+/9+ and M13mp10/mp11, respectively. The P9β-N2 probe (ss1) was obtained starting from the double stranded Maniatis β *Pst* I clone. The characteristics of the single-stranded probes are given in Table 1. Phage growth and isolation of recombinant DNA were performed as described (16, 20).

DNA Sequencing. The 2-kb *Bam*HI DNA fragments containing the 5′ end of the β-globin gene and 1.5 kb of the flanking sequence were either digested by *Sau*3A or sequentially degraded by BAL-31 (21). The resulting fragments were subcloned in the M13mp10/mp11 vector and sequenced by the dideoxy chain termination method (22, 23).

RESULTS

S1 Nuclease Mapping of DNA from SCA Patients. The structures of two arrays of tandemly repeated sequences located immediately 5′ to the β-globin gene locus were found to be highly polymorphic in human DNA. We first investigated the structure of these regions in the DNAs of SCA patients by S1 nuclease mapping by using four different

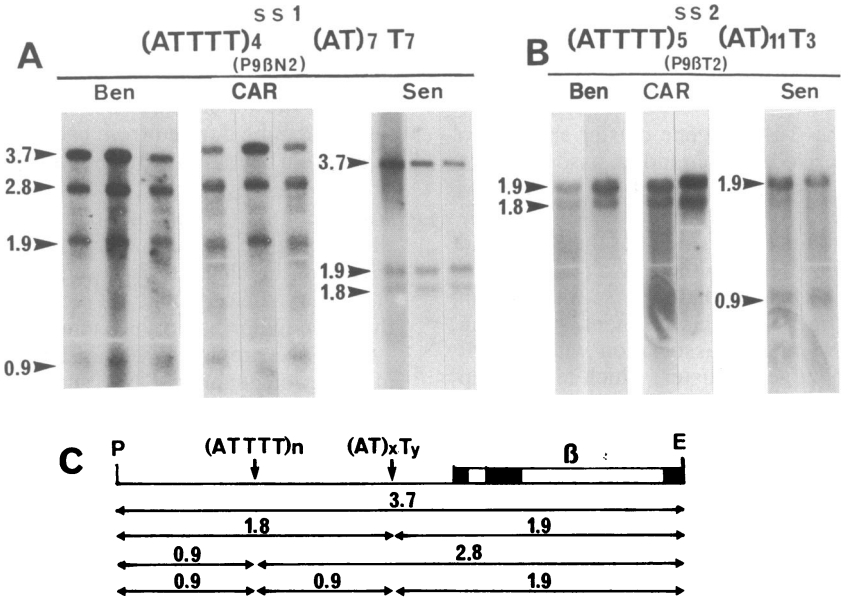

FIG. 1. S1 nuclease mapping of the δ–β intergene region of genomic DNAs from African SCA patients. Genomic DNAs were hybridized with the ss1 (*A*) or ss2 (*B*) probe, which contains a 3.7-kb *Pst* I/*Eco*RI DNA fragment that includes the β-globin gene. Hybrids were submitted to S1 nuclease treatment, electrophoresed on an agarose gel, transferred onto a nitrocellulose filter and hybridized with a ^{32}P-labeled β *Pst* I probe (19). (*C*) Interpretation of the patterns shown in *A* and *B*. Arrows indicate the location of possible mismatches in heteroduplexes. Fragments of 1.9 and 0.9 kb reveal the presence of S1-nuclease-sensitive sites at the $(ATTTT)_n$ and $(AT)_xT_y$ positions, respectively. Partial digestion by the enzyme results in the presence of 3.7- and/or 1.8-kb fragments. P, *Pst* I; E, *Eco*RI; Ben, Benin; Sen, Senegal.

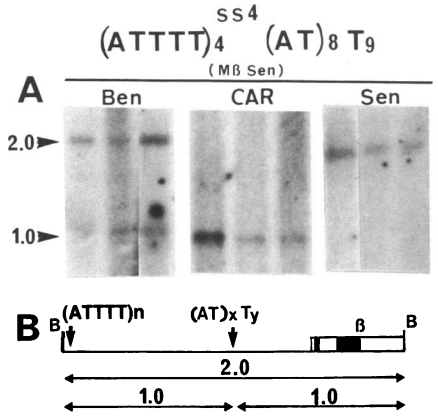

FIG. 2. S1 nuclease analysis of SCA patients' DNA using the ss4 probe prepared from Senegalese patient DNA. (*A*) Genomic DNAs from Benin, CAR, and Senegalese patients were hybridized with the ss4 probe, a 2-kb *Bam*HI DNA fragment cloned in M13 phage DNA. The resulting fragments were analyzed as described in Fig. 1. (*B*) Interpretation of the patterns shown in *A*. Arrows indicate the location of possible mismatches in heteroduplexes. Fragments of 1 kb reveal the presence of a sensitive site at the $(AT)_xT_y$ position. B, *Bam*HI; Ben, Benin; Sen, Senegal.

single-stranded DNA probes containing the variable segments $(ATTTT)_n$ and $(AT)_xT_y$ (Table 1).

When the single-stranded probe and the genomic DNA are identical, the hybrid does not present any sites sensitive to S1 nuclease, which results in a 3.7- or a 2-kb fragment by Southern analysis, depending on the length of the sequence protected by the probe on the target genomic DNA (Figs. 1 and 2). A difference at the $(ATTTT)_n$ or the $(AT)_xT_y$ position generates a single-stranded loop sensitive to S1 nuclease. The cleavage is total or partial, depending to the size of the loop. The lengths of the expected fragments are indicated in Figs. 1 and 2. The presence of a 3.7-kb fragment (when using the ss1, ss2, or ss3 probe) or a 2-kb fragment (when using the ss4 probe) as well as smaller fragments may therefore result either from partial cleavage with S1 nuclease or from heterozygous individuals.

Total genomic DNAs from Benin, Senegalese, and CAR SCA patients were first hybridized with the ss1 and ss2 probes (Fig. 1) and with the ss3 probe (not shown); after S1 nuclease assay, the fragments were mapped according to Southern (18) by using a ^{32}P-labeled β *Pst* I probe. With any of the single-stranded probes, the patterns obtained with CAR and Benin DNAs were identical (Fig. 1). The sizes of the fragments are listed in Table 2. These patterns showed that no S1-nuclease-sensitive site was created when Benin and CAR DNAs were hybridized with an $(ATTTT)_5$ structure, so the number of repeats in the probe and the genomic DNA was identical. The pattern was different with Senegalese DNAs, which revealed an $(ATTTT)_4$ structure (no cleavage with the ss1 probe). In all cases, differences were detected at the $(AT)_xT_y$ locus. Fragments of 3.7 and 2.8 kb are the result of partial digestion during S1 nuclease treatment. When the DNAs were hybridized with the ss2 probe (Fig. 1*B*), it was observed that the cleavage at the $(AT)_xT_y$ site was always complete (no 3.7-kb fragment), which reveals a more extensive mismatch than with the other probes and excludes the possibility of heterozygotic subjects.

DNAs from the three origins were then hybridized to the ss4 probe [prepared from the Senegalese DNA sample sequenced in the next step (see below)]. As expected, no cleavage was observed with any of the Senegalese DNAs studied. The hybrids formed with the Benin DNAs were partially digested during the S1 nuclease treatment at the $(AT)_xT_y$ position, whereas the cleavage was total when CAR DNAs were tested (Fig. 2). These results attest to the differences between the $(AT)_xT_y$ structures of Benin, CAR, and Senegalese DNAs.

DNA Sequence Analysis. A 1080-base-pair (bp) region 5′ to the β-globin gene was sequenced for one sample of DNA from each origin. The results confirmed the structure determined by S1 mapping regarding the ATTTT repeat: $(ATTTT)_4$ for Senegalese DNA and $(ATTTT)_5$ for Benin and CAR DNA. The sequences of the $(AT)_xT_y$ tract are shown in Fig. 3. They are aligned with homologous sequences from human and monkey DNAs by using the human DNA sequence published by Poncz *et al.* (5) as a reference. As expected from S1 nuclease analysis, each of the β^S chromosomes showed a specific structure of the alternating purine/pyrimidine track. Presently six different allelic forms are known to exist at this locus; four of them are presented in Fig. 3. Two others have been described previously; one of them is associated with an Algerian β^0-thalassemic chromosome (6), and the other is associated with an Albanian β-thalassemic gene (7).

Ten other nucleotide substitutions were observed among the four human chromosomes shown in Fig. 3; eight could be aligned without any ambiguity with the homologous sequences of the same region from monkey DNAs (5, 8, 9). Two positions (−543 and −535) were not used in the nucleotide comparison because of the ambiguity in the sequence alignment due to the high variability of the $(AT)_xT_y$ repeat. These results are illustrated in Fig. 3. At seven positions, two possible nucleotides were observed. Only the −491 position displayed three possibilities. These observations, together with the monkey nucleotide at each position, allow us to infer the most probable ancestral nucleotide (see *Discussion*).

Table 2. Sizes of the fragments obtained after S1 nuclease mapping and Southern analysis using various single-stranded probes

Probe	Fragment size, kb			Sequence homology*		
	CAR	Ben	Sen	CAR	Ben	Sen
ss1 [$(ATTTT)_4$ $(AT)_7T_7$]	3.7	3.7	3.7	None	None	$(ATTTT)_4$
	2.8	2.8	1.9			
	1.9	1.9	1.8			
	0.9	0.9				
ss2 [$(ATTTT)_5$ $(AT)_{11}T_3$]	1.9	1.9	1.9	$(ATTTT)_5$	$(ATTTT)_5$	None
	1.8	1.8	0.9			
ss3 [$(ATTTT)_5$ $(AT)_7T_7$]	3.7	3.7	3.7	$(ATTTT)_5$	$(ATTTT)_5$	None
	1.9	1.9	1.9			
	1.8	1.8	0.9			

Ben, Benin; Sen, Senegal.
*Sequence homologies between the single-stranded probe and the target genomic DNA at the level of repeated segments are indicated.

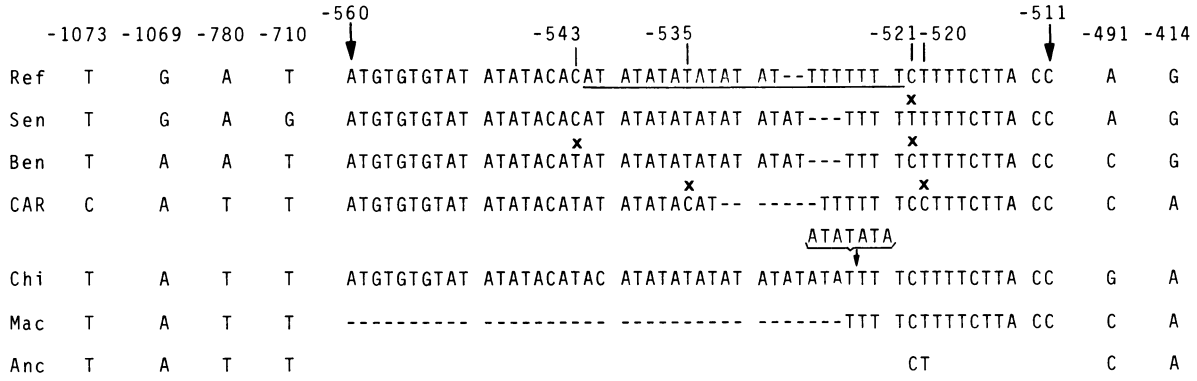

FIG. 3. Nucleotide sequence comparison of a 1080-bp DNA segment 5′ to the β-globin gene from human and monkey chromosomes. Numbering is that of the human sequence used as a reference (5) with the origin (0) at the cap site of the β-globin gene. The central part of the figure displays the nucleotide sequence between positions −560 and −511, including that of the $(AT)_xT_y$ repeat (underlined for the reference DNA). Also shown are the variable positions 5′ and 3′ to the segment. Ref, human sequence used as a reference (5); Sen, Senegal; Ben, Benin; Chi, chimpanzee (8); Mac, macaque (9); Anc, hypothetical ancestral nucleotides at the variable positions; x, substitution; –, deletion or gap.

In conclusion, (*i*) the number of ATTTT repeats was determined by using S1 nuclease mapping and was confirmed by sequencing. The value of 5 found in Benin and CAR DNAs is the more common. Up to now, only one case was described with $n = 4$, which appears to be characteristic of Senegalese SCA chromosomes.

(*ii*) In all cases, the structure of the $(AT)_xT_y$ fragment was different from that of the single-stranded probes used as a reference [i.e., different from $(AT)_7T_7$ and $(AT)_{11}T_3$]. The analysis with the Senegalese probe demonstrated that this structure was different for each population. To date six different $(AT)_xT_y$ structures have been described for human chromosomes, including those found in this work. Moreover, this region is also known to be variable in monkeys.

(*iii*) Specific substitutions were found at 10 positions as compared to homologous sequences from human and monkey DNAs (Fig. 3).

DISCUSSION

Haplotype determinations strongly argue in favor of three origins for the β^S mutation in African populations (15). Nevertheless, this argument is indirect and cannot totally rule out the possibility of recombinational events. Analysis of the 5′ flanking region of the β-globin gene, including repeated sequences, could therefore be informative. Such analysis was performed in two ways: (*i*) S1 nuclease mapping allowing analysis of multiple DNA samples from subjects from three regions of Africa and (*ii*) sequencing of the 1080-bp region 5′ to the β gene from one sample of each origin.

S1 nuclease mapping confirmed the homogeneity and the specificity of each population with regard to the structure of a highly variable region of the DNA characterized by the presence of the (ATTTT)*n* and $(AT)_xT_y$ repeats.

Among the three β^S chromosomes, each originating from a different population, 10 other nucleotide substitutions were detected along a 1080-nucleotide segment containing the $(AT)_xT_y$ sequence. Fig. 3 displays the nucleotides found at the same location in the normal human DNA sequence used as reference (5) as well as those found in the chimpanzee and macaque sequences (8, 9). The nucleotide differences at these positions allow us to characterize three chromosomes, each of which bear the β^S mutation. These observations confirm the geographic specificity of β^S-bearing chromosomes, with regard to a short sequence of repeated segments, known to be highly variable.

Using the method of maximum parsimony (i.e., an analysis based on the smallest number of independent mutations), it is possible to infer the relationship between the six different chromosomes. The ancestral nucleotide at each position can be deduced from the monkey DNA sequences. Among the six sequences compared, two distinct nucleotides are present at each position, with the exception of position −491, where three different possibilities are observed. In the latter case, we assume that there has been a cytidine→guanosine substitution in the chimpanzee lineage, as cytidine is present in both macaque and some human chromosomes. The $(AT)_xT_y$ segment is presumed to be a recombinational hot spot. This hypothesis stems from its particular structure of alternating purine/pyrimidine repeats and its high level of polymorphisms. Accordingly we determined separately the hypothetical phylogeny of the DNA segments lying 5′ and 3′ to this locus. The results are graphically presented in Fig. 4. A parallelism was observed in the diagram obtained with the two segments. In both cases, the CAR and Benin nucleotide sequences differ from the ancestral sequence by only one substitution, whereas the reference and the Senegalese chromosomes are more distant from the ancestral sequence and have two and three substitutions, respectively. These data do not bear any evidence of a recombination event at the $(AT)_xT_y$ locus between two different chromosomes after the occurrence of the β^S mutation. The evolution of these two segments, each about 500 bp long, involves three comparable levels of substitution, which make it possible to propose a hypothetical evolution diagram (Fig. 5).

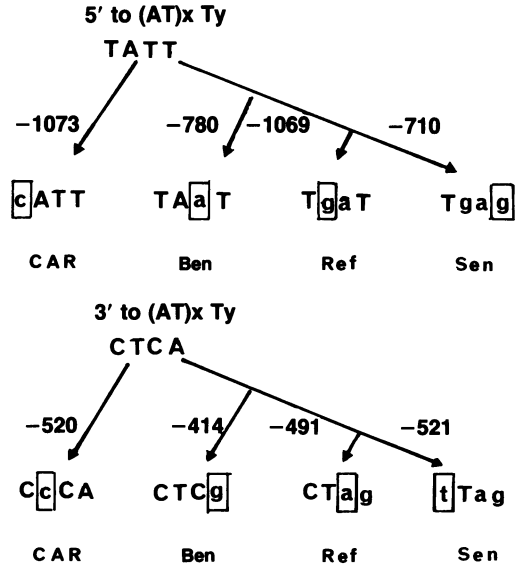

FIG. 4. Hypothetical phylogenetic tree of DNA segments both 5′ and 3′ to the $(AT)_xT_y$ region.

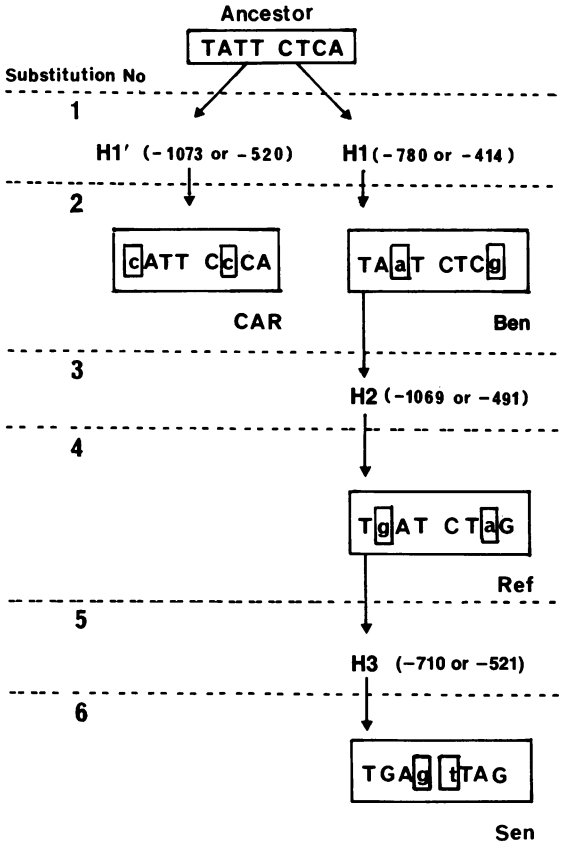

FIG. 5. Rooted parsimonious tree describing the emergence of polymorphisms in the 5' flanking region of the human β-globin gene. H1, H1', H2, and H3 are hypothetical intermediary sequences.

These phylogenetic trees tend to demonstrate that the origin of the polymorphisms associated with the β^S mutation in the 5' flanking region is quite ancient. They can be compared with those obtained for the evolution of the β-globin gene cluster based on sequencing data (24). These authors estimated that the divergence of the β-globin allele occurred 1.5 million years ago. The β^S mutation would have occurred after the divergence of the analyzed region. This is in agreement with different features, among them the geographical distribution of the sickle mutation, which suggest that it is quite recent. Accordingly, it has been estimated that it occurred a few thousand years ago (25).

In conclusion these studies support the hypothesis of three origins of the β^S mutation in African populations. We add more direct arguments based on the nucleotide structure of these alleles to the observations concerning the geographical distribution of the different alleles.

Note Added in Proof. Since this paper has been accepted, the 1080-bp segment 5' to the β-globin gene was sequenced for three samples of DNA from SCA Benin patients and two from Senegalese patients. The results confirm the sequence homogeneity in each population.

We are grateful for the assistance and cooperation of the following colleagues: Dr. J. Teyssier from Hopital Principal, Dakar, Sénégal; Dr. J. Roffi from Institut Pasteur, Dakar; Dr. I. Zohoun from Hopital de Cotonou, République Démocratique du Bénin; and Dr. G. Jaeger from Centre Eurafricain de Biologie Humaine, Université René Descartes, Paris, France. We thank Prof. S. Edelstein for fruitful discussion and Prof. J. Edsall and Dr. S. Satchwell for a critical reading of the manuscript. This work was supported by grants from Institut National de la Santé et de la Recherche Médicale and Centre National de la Recherche Scientifique.

1. Antonarakis, S. E., Boehm, C. D., Giardina, P. J. V. & Kazazian, H. H. (1982) *Proc. Natl. Acad. Sci. USA* **79,** 137–141.
2. Chakravarti, A., Buetow, K. H., Antonarakis, S. E., Waber, P. G., Boehm, C. D. & Kazazian, H. H. (1984) *Am. J. Hum. Genet.* **36,** 1239–1258.
3. Spritz, R. A. (1981) *Nucleic Acids Res.* **9,** 5037–5047.
4. Moschonas, N., deBoer, E. & Flavell, R. A. (1982) *Nucleic Acids Res.* **10,** 2109–2120.
5. Poncz, M., Schwartz, E., Ballantine, M. & Surrey, S. (1983) *J. Biol. Chem.* **258,** 11599–11609.
6. Chebloune, Y., Trabuchet, G., Poncet, D., Cohen-Solal, M., Faure, C., Verdier, G. & Nigon, V. M. (1984) *Eur. J. Biochem.* **142,** 473–480.
7. Semenza, G. L., Mallidi, P., Surrey, S., Delgrosso, K., Poncz, M. & Schwartz, E. (1984) *J. Biol. Chem.* **259,** 6045–6048.
8. Savatier, P., Trabuchet, G., Faure, C., Chebloune, Y., Gouy, M., Verdier, G. & Nigon, V. (1985) *J. Mol. Biol.* **182,** 21–29.
9. Savatier, P., Trabuchet, G., Chebloune, Y., Faure, C., Verdier, G. & Nigon, V. (1987) *J. Mol. Evol.* **24,** 297–308.
10. Gerhard, D. S., Kidd, K. K., Kidd, J. R., Egeland, J. A & Housman, D. E. (1984) *Proc. Natl. Acad. Sci. USA* **81,** 7875–7879.
11. Orkin, S. H., Kazazian, H. H., Antonarakis, S. E., Goff, S. C., Boehm, C. D., Sexton, J. P., Waber, P. G. & Giardina J. V. (1982) *Nature (London)* **296,** 627–631.
12. Antonarakis, S. E., Orkin, S. H., Kazazian, H. H., Goff, S. C., Boehm, C. D., Waber, P. G., Sexton, J. P., Ostrer, H., Fairbanks, V. F. & Chakravarti, A. (1982) *Proc. Natl. Acad. Sci. USA* **79,** 6608–6611.
13. Antonarakis, S. E., Boehm, C. D., Serjeant G. R., Theisen, C. E., Dover, G. J. & Kazazian, H. H. (1984) *Proc. Natl. Acad. Sci. USA* **81,** 853–856.
14. Wainscoat, J. S., Bell, J. I., Thein, S. L., Higgs, D. R., Serjeant, G. R., Peto, T. E. A. & Weatherall, D. J. (1983) *Mol. Biol. Med.* **1,** 191–197.
15. Pagnier, J., Mears, J. G., Dunda-Belkhodja, O., Schaeffer-Rego, K. E., Beldjord, C., Nagel, R. L. & Labie, D. (1984) *Proc. Natl. Acad. Sci. USA* **81,** 1771–1773.
16. Chebloune, Y., Poncet, D. & Verdier, G. (1984) *Biochem. Biophys. Res. Commun.* **120,** 116–123.
17. Whitelaw, E., Pagnier, J., Verdier, G., Henni, T., Godet, J. & Williamson, R. (1980) *Blood* **55,** 511–516.
18. Southern, E. M. (1975) *J. Mol. Biol.* **98,** 503–517.
19. Lawn, R. M., Efstradiatis, A., O'Connel, C. & Maniatis, T. (1980) *Cell* **21,** 647–651.
20. Messing, J. & Vieira, J. (1982) *Gene* **19,** 269–276.
21. Poncz, M., Solowiejczyk, D., Ballantine, M., Schwartz, E. & Surrey S. (1982) *Proc. Natl. Acad. Sci. USA* **79,** 4298–4302.
22. Sanger, F., Nicklen, S. & Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. USA* **74,** 5463–5467.
23. Biggin, M. D., Gibson, T. J. & Hong, G. F. (1983) *Proc. Natl. Acad. Sci. USA* **80,** 3963–3965.
24. Savatier, P., Trabuchet, G., Chebloune, Y., Faure, C., Verdier, G. & Nigon, V. (1987) *J. Mol. Evol.* **24,** 309–318.
25. Edelstein, S. J. (1986) in *The Sickled Cell* (Harvard Univ. Press, Cambridge, MA), pp. 44–64.