Volume 9 Number 15 1981 Nucleic Acids Research

Two types of triplicated α-globin loci in humans

Luan Eng Lie-Injo*, Alejandro R.Herrera* and Yuet Wai Kan**

*Department of Epidemiology and International Health and **Howard Hughes Medical Institute and Department of Medicine, University of California, San Francisco, CA 94143, USA

Received 28 April 1981

ABSTRACT

DNA from healthy Malaysian newborns was studied on gene maps after digestion with different restriction endonucleases. Of 65 newborns, two were found to be carriers of two different variants of triplicated α-globin loci. In variant no. 1, found in a Malay, the three α-globin genes are in an elongated DNA fragment on digestion with Eco RI and Bam HI. The third α-globin gene was found in an additional 3.7-kb fragment on digestion with Hpa I, Bgl II and Hind III. In variant no. 2, a new type of triplicated α-globin loci, found in a Chinese, the three α-globin genes reside in an elongated DNA fragment longer than that of variant no. 1 on digestion with Eco RI and Bam HI. The third α-globin gene was found in an additional 4.2-kb fragment on digestion with Hpa I and Hind III. Digestion of this variant DNA with Bgl II produced an abnormal 16.7-kb fragment in addition to the normal 7.0-kb Bgl-II fragment. The locations of the restriction sites in the two types of triplicated α-globin loci are compatible with a mechanism of unequal crossing over following two different modes of misalignment.

INTRODUCTION

Normal humans usually have four α-globin genes, two on each chromosome. The severe clinical condition known as Hb Bart's hydrops fetalis, which is the homozygous condition for α-thalassemia (1), is due to deletion of all four α-globin genes (2, 3). Deletion of three α-globin genes leads to Hb-H disease (4), which is the combination of the severe type of αthal-assemia ($\alpha thal_1$) and the mild type of αthalassemia ($\alpha thal_2$)(5). The $\alpha thal_1$ trait condition is due to a deletion of both α-globin genes on one chromosome. The $\alpha thal_2$ trait is due to deletion of one α-globin gene on one chromosome. A nondeletion type αthalassemia has also been reported (6). The organization of the α-like globin genes in the chromosome has been determined as being 5'-ζ2-ζ1-ψα1-α2-α1-3' (7). The distance between the centers of the α1- and α2-globin genes is 3.7 kb and that between the centers of α2- and ψα1-globin gene is 4.2 kb. Recently, Embury et al (8) showed that the deletion $\alpha thal_2$ genotype has two different patterns of gene organization.

The first pattern is the result of a 4.2-kb pair deletion involving the 5'α-globin gene (leftward-deletion $\alpha thal_2$ genotype); the second is probably the result of crossover deletion of a DNA fragment bridging the normal α1- and α2-globin genes (rightward-deletion $\alpha thal_2$ genotype). Goossens et al (9) and Higgs et al (10) described the existence of triplicated α-globin loci in humans which is thought to be the result of an unequal crossover mechanism producing triplicated α-globin loci on one side and the rightward -deletion $\alpha thal_2$ genotype on the other side. We report the finding of a new type of triplicated α-globin loci during a survey of healthy newborn babies. The results described in detail here show that this new type is probably the counterpart for the leftward-deletion $\alpha thal_2$ genotype.

MATERIALS AND METHODS

Study material. Cord blood samples from healthy Malaysian newborns were collected from the Maternity Unit, General Hospital, Kuala Lumpur, Malaysia. The white blood cells were isolated and dissolved in Tris-EDTA buffer, quick-frozen and airfreighted to San Francisco. Hemolysate was prepared from the red blood cells and hemoglobin analysis carried out according to standard methods.

Preparation of DNA and restriction endonuclease digestion. DNA was prepared from white blood cells by phenol chloroform-isoamyl alcohol extraction and ethanol precipitation as described earlier (2). DNA was digested with restriction endonucleases under conditions recommended by the manufacturers from whom they were purchased.

Hybridization and gene mapping. DNA fragments obtained by restriction endonuclease digestion were electrophorsed in 0.8% agarose and transferred to nitrocellulose filter paper according to Southern (11) with modifications. They were hybridized with 32P-labeled cDNA probe synthesized by reverse transcription of purified mRNA obtained from reticulocyte rich adult blood. They were also hybridized with specific 32P-labeled α-globin gene probe prepared by nick translation as described by Maniatis et al (12) from α-globin genes cloned in plasmid JW 101 (13). Hybridization was carried out for 2-3 days, followed by thorough washing ot the nitrocellulose filter paper under stringent conditions and autoradiography for 2-5 days.

RESULTS

A group of 65 white blood cell DNA samples from blood of healthy newborns without Hb Bart's or other abnormal hemoglobin in the red

blood cells was studied. These DNA samples served as controls in our study of cord blood of newborn Malaysians in which we correlated the level of Hb Bart's with deletions of α-globin genes. The results of that survey will be published elsewhere.

The DNA samples were studied after digestion with different restriction endonucleases on gene maps as described under Methods.

Eco RI and Bam HI. The restriction endonucleases Eco RI and Bam HI cut normal human DNA outside the two α-globin genes leaving both α-globin genes intact on the same DNA fragment. On an Eco RI gene map the normal α-globin genes are located at position 23 kb in length and on a Bam HI gene map the normal α-globin genes are at position 14.5 kb in length. Two of the 65 DNA samples examined on Eco RI and Bam HI gene maps showed two α-globin gene bands, one normal band and an additional slower-than-normal-moving band, indicating that the DNA fragments on which they are located are longer than normal. Both variants were slower than normal but variant no. 2, found in a Chinese, was even slower than variant no. 1, found in a Malay. On an Eco RI gene map, variant no. 1 was barely separated from the normal 23-kb fragment whereas variant no. 2 was clearly separated from the normal band (Fig. 1).

Hpa I. The restriction endonuclease Hpa I cuts human DNA outside and between the two α-globin genes, producing two DNA fragments, each carrying an α-globin gene, one 14.5 kb and the other 4.2 kb long. Variant no. 1, described above, on the Hpa I gene map clearly shows an additional, abnormal α-globin gene-carrying fragment, 3.7 kb in length (Fig. 2). The specificity of this α-globin gene band was demonstrated by hybridization with a specific α-globin cDNA probe. This variant is apparently the same as the one described earlier (9, 10), which is a case of triplicated α-globin loci due to unequal crossing over. Variant no. 2, described above, which on Eco RI and Bam HI gene maps was shown to reside on an elongated DNA fragment did not show an abnormal additional band on the Hpa I gene map. However, the usual 4.2-kb Hpa I band appeared definitely stronger than normal, indicating that there are more than the normal 4.2-kb fragments carrying α-globin genes (Fig. 2).

Bgl II. The restriction endonuclease Bgl II cuts human DNA outside and between the two α-globin gene producing two DNA fragments, each carrying an α-globin gene, one 12.5 kb and the other 7.0 kb long. On a Bgl II gene map, the carrier of variant no. 1 showed in addition to

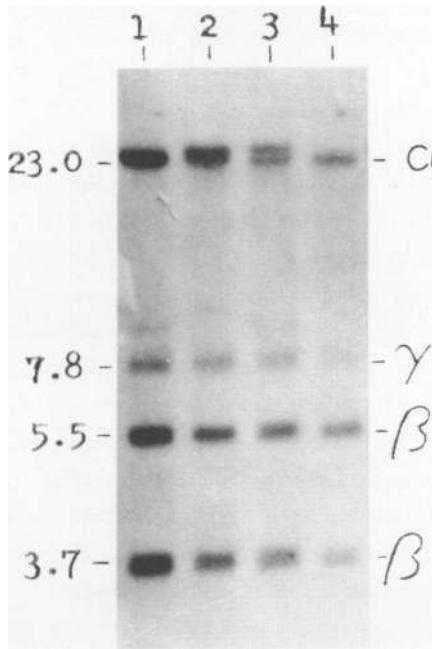

Figure 1. Autoradiogram of Eco RI endonuclease digestion patterns of human DNA. Total globin cDNA is used as probe. Lanes 1 and 4: normal. Lane 2: carrier of triplicated α-globin loci variant no. 1. Lane 3: carrier of triplicated α-globin loci variant no.2

the normal 12.5-kb and 7.0-kb fragments an abnormal α-globin gene band 3.7 kb in length. The carrier of variant no. 2 showed a slow-moving 16.7-kb fragment carrying α-globin genes in addition to the 12.5-kb and 7.0-kb fragments. The specificity of the abnormal bands was demonstrated by hybridization with a specific α-globin cDNA probe (Fig. 3).

Hind III. The restriction endonuclease Hind III cuts human DNA within the coding sequences of the α-globin genes, producing three DNA fragments carrying α-globin genes, one 17 kb long, which contains the 5' portion of α2-globin gene, a 3.7-kb bridging fragment containing the 3' portion of α2- and the 5' portion of α1-globin genes and a 4.5-kb fragment which contains the 3' portion of the α1-globin gene locus. On an Hind III gene map, variant no. 1 showed the normal positions of the α-globin gene bands but the 3.7-kb band is pronounced, indicating that it contains more 3.7-kb fragments carrying α-globin genes than normal.

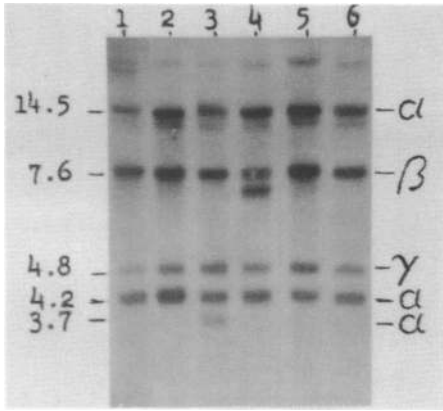

Figure 2. Autoradiogram of HPA I endonuclease digestion patterns of human DNA. Total globin cDNA is used as probe. Lanes 1 and 4: normal (4 shows polymorphism of the β-globin gene). Lane 2: carrier of triplicated α-globin loci variant no. 2. Lane 3: carrier of triplicated α-globin loci variant no. 1. Lane 5: carrier of $\alpha thal_2$. Lane 6: carrier of Hb CoSp.

This is apparently due to overlap of the 3.7-kb fragment carrying the additional α-globin gene with the normal 3.7-kb Hind-III fragment. Variant no. 2, however, shows an additional abnormal fragment 4.2 kb in length instead of 3.7 kb in length and demonstrable on an Hind III gene map slightly separated from the normal 4.5-kb fragment (Fig. 4).

DISCUSSION

Comparison of the locations of restriction sites within and flanking the α1- and α2-globin genes and heteroduplex studies by Lauer et al (7) reveal extensive sequence homology within and flanking the two α-globin genes. The homologous sequences, interrupted by two blocks of nonhomology, span a region of approximately 4 kb. This extensive sequence homology between two genes, which are thought to be the product of an ancient duplication event, suggests the existence of a mechanism for sequence matching during evolution. Two homologous regions for possible crossing over were described; one is a region of homology that begins 100 bp 3' to the coding sequences of each α-globin gene and continues through the genes into the 5' flanking sequences for a total of 1.8 kb. The other homologous region is a stretch of homologous sequence that starts, according to Lauer et al (7) from the intergenic

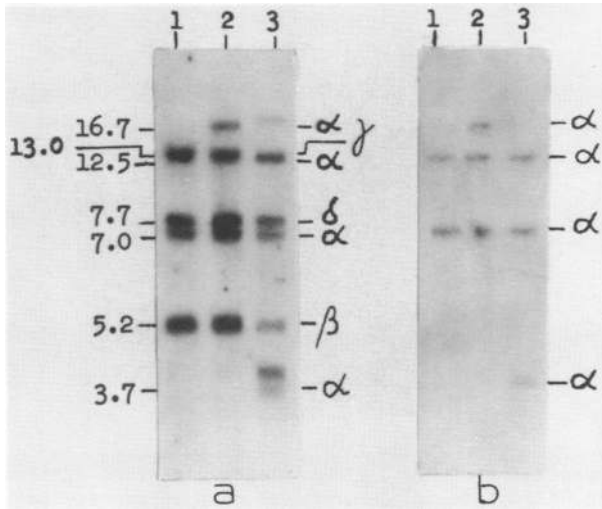

Figure 3. Autoradiogram of Bgl II endonuclease digestion patterns of human DNA.
a. Total globin cDNA is used as probe. Lane 1: normal. Lane 2: carrier of triplicated α-globin loci variant no. 2. Lane 3: carrier of triplicated α-globin loci variant no.1 (the faint slowest band in this sample is not an α-globin gene band, it does not hybridize with specific α-globin gene probe. See b.).
b. The above DNA map rehybridized with specific α-globin gene probe. Only the α-globin gene bands are seen.

Hpa I restriction points 3' to the α2- and ψα1-globin genes and extends for 1 kb 3' ward. Two types of deletions in recombinant phage DNA were found during propagation in E. coli that were compatible with their being the consequence of the particular arrangement of homologous sequences. The locations and sizes of the two types of deletions were indistinguishable from those of the two types of deletions associated with $\alpha thal_2$ as has been reported (8, 14-16). These findings strongly suggest that $\alpha thal_2$ is produced by unequal crossing over between homologous sequences within and or surrounding the α-globin genes. Further support for this assumption was the finding of triplicated α-globin loci (9, 10).

Recent sequence analysis of the α2-globin gene (17) and of the ψα1-globin gene (18) and comparisons of the sequences of α1-, α2-, and ψα1-globin genes determined more precisely the homologous regions between the different α-globin genes (17-19). The

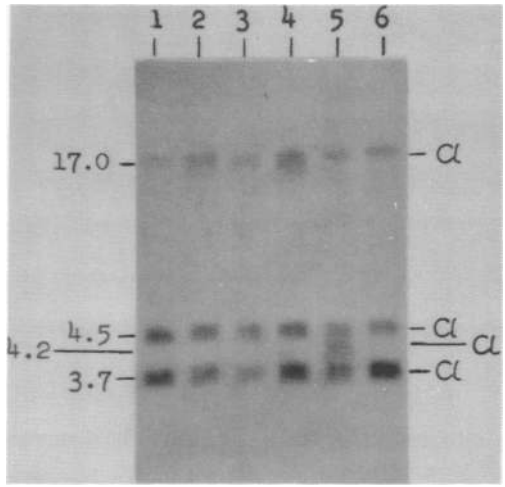

Figure 4. Autoradiogram of Hind III endonuclease digestion patterns of human DNA. Specific α-globin cDNA is used as probe. Lanes 1 and 4: normal. Lane 2: $\alpha thal_2$ trait. Lane 3: $\alpha thal_1$ trait. Lane 5: triplicated α-globin loci variant no. 2. Lane 6: triplicated α-globin loci variant no. 1 (note the pronounced 3.7-kb fragment in this sample).

5' border of the homologous area flanking the ψα 1- and α 2-globin genes is found just 3' to the poly (A) addition site of the ψα 1- and α 2-globin genes (Fig. 4). The 3' border of the second homologous area, that is between α 2- and α 1-globin genes, has been reassigned to a position 14 bases 3' to the UAA termination codon (19).

Detailed studies of the two variants of triplicated α-globin loci, described in our present study, show that variant no. 1 is similar to the one reported earlier (9, 10). This variant is the result of unequal crossing over after misalignment of the α 2-globin gene on one chromosome with the α 1-globin gene on the other chromosome, producing on one side a DNA strand in which one α-globin gene is deleted and which is 3.7 kb shorter than normal (the rightward-deletion $\alpha thal_2$ genotype described earlier)(8) and on the other side a DNA strand with three α-globin genes which is 3.7 kb longer than normal (the triplicated α-globin loci described earlier (9, 10) Digestion of this abnormal elongated DNA containing the three α-globin loci with enzymes that cut the DNA outside the globin genes, such as Eco RI and Bam HI, produces a DNA fragment 3.7 kb longer than normal because the centers of the α 1-

and α2-globin genes are located 3.7 kb apart (Fig. 5). Digestion of the elongated DNA with enzymes that cut it outside and between the α-globin genes, such as Hpa I and Bgl II, produces an additional 3.7-kb fragment which contains the extra α-globin gene.

Our variant no. 2 is the result of unequal crossing over following a misalignment of a homologous region 3' to the α2-globin gene on one chromosome with the homologous region 3' to the ψα1-globin gene on the other chromosome producing on one side a shortened DNA fragment with 4.2-kb deletion involving the 5'α-globin gene (the leftward-deletion $\alpha thal_2$ genotype previously described)(8) and on the other side a chromosome with triplicated α-globin loci 4.2 kb longer than normal (our variant no. 2). Digestion of the abnormal DNA with the triplicated α-globin loci with enzymes that cut the DNA outside the α-globin genes such as Eco RI and Bam HI produces a DNA fragment 4.2 kb longer than normal, which is the distance between the centers of the α2-globin gene and the ψα1-globin gene (Fig. 5). Digestion of this

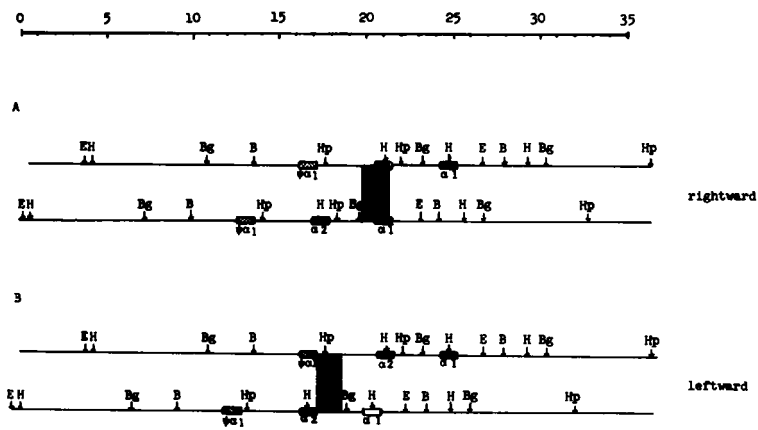

Figure 5. a. Mode of misalignment of homologous regions flanking and involving the α1- and α2-globin genes in the two chromosomes leading to the generation of the rightward-deletion $\alpha thal_2$ genotype on one side and the triplicated α-globin loci variant no. 1 on the other side.
b. Mode of misalignment of homologous regions flanking the 3' side of the α2-globin gene in one chromosome and the ψα1-globin gene in the other chromosomes, leading to the generation of the leftward-deletion $\alpha thal_2$ genotype on one side and the triplicated α-globin loci variant no. 2 on the other side. The shaded areas indicate where the crossover point may reside.

elongated DNA containing the three α-globin genes by Hpa I, which cuts it outside and between the α-globin genes, or by Hind III, which cuts the coding sequences of the α-globin genes, produces an extra DNA fragment 4.2 kb long containing the third α-globin gene. In the Hpa I gene map, this extra 4.2-kb fragment overlaps with the normal 4.2-kb Hpa I fragment while on a Hind III gene map it shows as an extra α-globin gene band at 4.2 kb in length slightly separated from the normal Hind III 4.5-kb fragment. When digested with Bgl II, no additional 4.2-kb fragment is generated, instead a fragment 16.7 kb long is produced. This means that the unequal crossing over took place somewhere in the homologous region 3' to the α2- and ψα1-globin genes (7, 17-19), retaining the Hpa I restriction point in the elongated fragment but eliminating the Bgl II restriction point, which is located 3' to and outside the homologous area (Fig. 5). Digestion with Bgl II therefore generates the usual 7.0-kb fragment and an elongated DNA fragment which is 16.7 kb long, being the sum of the usual 12.5-kb fragment and the additional 4.2-kb fragment (gained from the unequal crossing over). The existence of the new triplicated α-globin loci reported here has been anticipated from the existence of the leftward-deletion $\alpha thal_2$ genotype and from the studies of cloned α-globin genes that show the presence of extensive homologous regions flanking the α-globin genes. Indeed, this new type of triplicated α-globin loci has also just been found in recombinant phage DNA during propagation in *E. coli* (Goossens, M., Lee, K.Y. and Kan, Y.W., in preparation).

Theoretically, the process of unequal crossing over should produce equal numbers of chromosomes with the single α-globin gene ($_{\alpha}thal_2$) and the triplicated α-globin loci. However, the triplicated α-globin loci, probably the counterparts for the righward-deletion type $\alpha thal_2$ and the leftward-deletion type $\alpha thal_2$ are in general much less prevalent than both $\alpha thal_2$ genotypes (except perhaps in Greek Cypriots)(9). Apparently, chromosomes with triplicated α-globin loci did not offer an advantage over the normal duplicated α-globin loci in the areas where they have been found to be rare. The one α-globin gene deletion apparently did offer a selective advantage. This selective advantage is probably given by malaria infection, which is prevalent or has been prevalent, in the areas where αthalassemias are also prevalent.

ACKNOWLEDGENTS

We thank the staff of the Delivery Room of the Maternity Unit, General Hospital, Kuala Lumpur, Malaysia, for their generous help in collecting cord blood samples, Ms. Wong Pui Wan for sending the blood samples to San Fran-

cisco, and Dr. B.G. Forget for the gift of JW 101 plasmid containing α-globin genes.

This work was supported by research grants AM 2157, AM 16666, and by Biochemical Research Support Grant S07 RR05355, Division of Research Resources, all three from the National Institutes of Health, and by grants from the Committees on Research of the University of California and the School of Medicine, University of California, San Francisco. Dr. Y.W. Kan is an Investigator of the Howard Hughes Medical Institute.

The AMV reverse transcriptase was obtained from the Division of Cancer Cause and Prevention, National Cancer Institute.

Address correspondence and reprint requests to Dr. Luan Eng Lie-Injo, Department of Epidemiology and International Health, 1699 HSW, University of California, School of Medicine, San Francisco, CA 94143.

REFERENCES

1. Lie-Injo, L.E. (1962) Alpha-chain thalassemia and hydrops foetalis in Malaya. A report of five cases. Blood 30, 581-590.
2. Taylor, J.M., Dozy, A., Kan, Y.W., Varmus, H.E., Lie-Injo, L.E., Ganesan, J., and Todd, D. (1974) Genetic lesion in homozygous α-thalassemia (hydrops fetalis). Nature 251, 392-393.
3. Ottolenghi, S., Lanyon, W.G., Paul, J., Williamson, R., Weatherall, D.J., Clegg, J.B., Pritchard, J., Pootrakul, S., and Wong, H.B. (1974) The severe form of α-thalassemia is caused by a haemoglobin gene deletion. Nature 251, 389-392.
4. Kan, Y.W., Dozy, A.M., Varmus, H.E., Taylor, J.M., Holland, J.P., Lie-Injo, L.E., Ganesan, J., and Todd, D. (1975) Deletion of α-globin genes in haemoglobin H disease demonstrates multiple α-globin structural loci. Nature 255, 255-256.
5. Wasi, P., Nanakorn, S., and Suingdumrong, A. (1964) Haemoglobin H disease in Thailand. A genetic study. Nature, 204, 907.
6. Kan, Y.W., Dozy, A.M., Trecartin, R. and Todd, D. (1977) Identification of a non-deletion defect in α-thalassemia. N. Engl. J. Med., 297, 1081-1084.
7. Lauer, J., Shea, C.J., and Maniatis, T. (1980) The chromosomal arrangement of human α like globin genes: sequence homology and α-globin gene deletions. Cell 20, 119-130.
8. Embury, S.H., Miller, J.A., Dozy, A.M., Kan, Y.W., Chan, V., and Todd, D. (1980) Two different molecular organizations account for the single α-globin gene of the α-thalassemia-2 genotype. J. Clin. Invest. 66, 1319-1315.
9. Goossens, M., Dozy, A.M., Embury, S.H., Zachariades Z., Hadjiminas, M.C., Stamatoyannoupoulos, G., and Kan, Y.W. (1980) Triplicated α-globin loci in humans. Proc. Natl. Acad. Sci. 77, 518-521.
10. Higgs, D.R., Old, J.M., Pressley, L., Clegg, J.B. and Weatherall, D.J. (1980) A novel α-globin gene arrangement in man. Nature 284, 632-635.
11. Southern, E.M. (1975) Detection of specific sequences among DNA fragments separated by gel electrophoresis. J. Mol. Biol. 98, 503-517.
12. Maniatis, T., Jeffrey, A. and Kleid, D.G. (1975) Nucleotide sequence of the rightward operation of phage λ. Proc. Nat. Acad. Sci. USA 72, 1184-1188.

13. Wilson, J.T., Wilson, L.B., de Riel, J.K., Villa-Komaroff, L., Efstratiadis, A., Forget, B.G. and Weissman, S.M. (1978) Insertion of synthetic copies of human globin genes into bacterial plasmids. Nucleic Acids Res. 5, 563-581.
14. Embury, S.H., Lebo, R.V., Dozy, A.M. and Kan, Y.W. (1979) Organization of the α-globin genes in the Chinese α-thalassemia syndromes. J. CLin. Invest. 63, 1307-1310.
15. Lie-Injo, L.E., Dozy, A.M., Kan, Y.W., Lopes, M., and Todd, D. (1979) The α-globin gene adjacent to the gene for HbQ-α 74 Asp→His is deleted, but not that adjacent to the gene for HbG-α 30 Glu→Gln; three-fourths of the α-globin genes are deleted in HbQ-α-thalassemia. Blood, 54, 1407-1416.
16. Orkin, S.H., Old, J., Lazarus, H., Altay, C., Gurgey, A., Weatherall, D.J. and Nathan, D.G. (1979) The molecular basis of α thalassemia. Frequent occurrences of dysfunctional α loci among non-Asians with HbH disease. Cell 17, 33-42.
17. Liebhaber, S.A., Goossens, M. and Kan, Y.W. (1980) Cloning and complete nucleotide sequence of human 5'-α-globin gene. Proc. Nat. Acad. Sci. USA 77, 7054-7058.
18. Proudfoot, N.J and Maniatis, T. (1980) The structure of a human α-globin pseudogene and its relationship to α-globin gene duplication. Cell 21, 537-544.
19. Liebhaber, S.A., Goossens, M., and Kan, Y.W. (1981) Homology and concerted evolution at the α1 and α2 loci of human α-globin. Nature 290, 26-29.