Technical Report Series

Detection in a Japanese Population of a Length Polymorphism in the 5′ Flanking Region of the Human β-globin Gene with Denaturing Gradient Gel Electrophoresis

Norio Takahashi, Ph.D.; Keiko Hiyama, M.D.;
Mieko Kodaira, Ph.D.; Chiyoko Satoh, Ph.D.

Radiation Effects Research Foundation
A Cooperative Japan–United States Research Organization

RERF Technical Report Series

Technical reports are the basic medium for reporting of original research carried out at the Radiation Effects Research Foundation. Reports in this series receive both internal and external peer review and may serve, in part or in toto, as the basis for publication in the open scientific literature. Although they may be quoted and cited, these reports are considered to be internal publications of the Foundation. Copies are available upon request from Publication and Documentation Center, RERF, 5-2 Hijiyama Park, Minami-ku, Hiroshima, 732 Japan.

The Radiation Effects Research Foundation (formerly ABCC) was established in April 1975 as a private nonprofit Japanese foundation, supported equally by the Government of Japan through the Ministry of Health and Welfare, and the Government of the United States through the National Academy of Sciences under contract with the Department of Energy.

業績報告書シリーズ

変性剤濃度勾配ゲル電気泳動法による日本人集団のヒトβグロビン遺伝子の5′側領域に存在する長さ多型性の検索[§]

Detection in a Japanese Population of a Length Polymorphism in the 5′ Flanking Region of the Human β-globin Gene with Denaturing Gradient Gel Electrophoresis

高橋規郎　檜山桂子　小平美江子　佐藤千代子

要　約

βグロビン構造遺伝子の約1,400塩基対上流に存在するATTTT反復配列多型性の分析を，RNA：DNAデュプレックスの変性剤濃度勾配ゲル電気泳動（DGGE）法を用いて行った。染色体およびクローン化DNAは，制限酵素により切断され，^{32}PでラベルされたRNAプローブとハイブリダイゼーションされた。得られたRNA：DNAデュプレックスはDGGEで検査された。反復配列単位の数の差異はDGGEゲル上でデュプレックスの移動度の差として検出された。広島在住の血縁関係のない81人の日本人を検査したところ，この領域中に多型性を示す配列の変異が認められた。既に報告されている5個および6個のATTTT反復単位を有する対立遺伝子は，多型の頻度で認められた。その上，2種類の新しい変異が検出された。そのうち1種は7個の反復を有し，もう一方は，5番目の反復配列中にAからGへの塩基置換を有するものであった。家族調査の結果は，これら4種の変異の分離状況は，優劣のない対立遺伝子による遺伝様式と矛盾しないことを示した。この研究はRNA：DNAデュプレックスのDGGEが，DNA中の変異を検出するための高感度の技法であることも示した。

[§]本業績報告書は研究計画書RP 7-85に基づく。本報告にはこの要約以外に訳文はない。本報告に基づく論文はHum Genet 87:219–20，1991に掲載された。承認1991年7月12日。印刷1992年10月。

注 著者はすべて広島の放影研 遺伝学部。

Technical Report Series

Detection in a Japanese Population of a Length Polymorphism in the 5′ Flanking Region of the Human β-globin Gene with Denaturing Gradient Gel Electrophoresis§

Norio Takahashi, Ph.D.; Keiko Hiyama, M.D.; Mieko Kodaira, Ph.D.; Chiyoko Satoh, Ph.D.

Summary

An analysis of the ATTTT repeat polymorphism located approximately 1,400 base pairs upstream from the β-globin structural gene was carried out by denaturing gradient gel electrophoresis (DGGE) of RNA:DNA duplexes. Genomic or cloned DNAs were digested with restriction enzymes and hybridized with ^{32}P-labeled RNA probes, and resulting RNA:DNA duplexes were examined by DGGE. A difference in the number of repeat units was recognized by differences in duplex mobility on the DGGE gel. In this study of 81 unrelated Japanese from Hiroshima, a sequence heteromorphism was observed at this site. Alleles with 5 and 6 repeats of the ATTTT unit, which had already been reported, were found in polymorphic proportions. In addition, two unreported alleles, one having 7 repeats and the other having an A-to-G nucleotide substitution in the 5th repeat, were detected. Family study data showed that the segregation of these four types of variants is consistent with an autosomal codominant mode of inheritance. This study also demonstrated that DGGE of RNA:DNA duplexes is a sensitive tool for detecting variations in DNA.

Introduction

Tandem repeats of relatively simple sequences are widely distributed in the genomes of eukaryotes and form substantial portions of the genome in many species. An array of short, simple, repeated sequences has been found to be a "hotspot" for spontaneous mutations,[1] with frequent duplication or deletion of single copies of the repeating unit occurring at such a site.

Once the hypervariable, repetitive sequences were reported,[2–6] oligomeric sequences derived from the tandem repeat regions of various genes have been

§*This technical report is based on Research Protocol 7-85. The complete text of this report will not be available in Japanese. A paper based on this report was published in* Hum Genet *87:219–20, 1991. Approved 12 July 1991; printed October 1992.*

NOTE: All the authors are with the Department of Genetics, RERF, Hiroshima.

developed as probes. By using these probes, multiallelic polymorphisms due to differences in the number of tandemly repeated, short DNA sequences have been found (e.g., the "minisatellite DNA" reported by Jeffreys et al.,[6] and the variable number of tandem repeat [VNTR] loci reported by Nakamura et al.[7]).

DNA sequence analysis has demonstrated similar short, simple, tandem repeat sequences approximately 1,400 base pairs (bp) 5′ to the human β-globin gene. Spritz[8] compared a number of cloned β-globin genes and demonstrated a remarkable sequence heteromorphism in this region, consisting of 4, 5, or 6 repeating units of ATTTT. He assumed that duplication/deletion of the repeating unit had probably occurred at least twice in this region. Thus far a large population could not be screened for such a heteromorphism, since DNA sequencing requires many reaction steps and is not suitable for examining many DNA samples. In contrast to sequence analysis, two recently developed scanning methods detect variations with high efficiency and simplicity. One is based on the melting behavior of DNA and uses denaturing gradient gel electrophoresis (DGGE).[9–12] In this procedure, a single-stranded, radioactive DNA probe was hybridized to target DNA molecules and the resulting DNA:DNA duplex was analyzed by DGGE. DGGE was also used to detect variations in influenza virus RNA molecules,[13] which have a much larger number of copies of the target sequences than does genomic DNA, by using either RNA probes or DNA probes. Because it was easier to prepare RNA probes than DNA probes, we recently demonstrated that DGGE of RNA:DNA duplexes was much more practical for screening a large number of samples than DGGE of DNA:DNA heteroduplexes,[14] as originally developed by Lerman et al.[15] The other method is based on the susceptibility of single-base mismatches in RNA:DNA duplexes to ribonuclease (RNase) cleavage, and is called the "RNase cleavage method."[16,17] We exhibited the usefulness of both DGGE of RNA:DNA duplexes and the RNase cleavage method in model experiments in which sequence variations in the β-globin genes were efficiently detected.[14,18] The RNase cleavage method could detect differences in the number of tandem repeats of ATTTT in the cloned β-globin gene fragments[18] and PCR-amplified genomic DNA fragments, an example of which will be described in this report. Clear results were not obtained when genomic DNA samples were examined without amplification. In this report we demonstrate that DGGE of RNA:DNA duplexes can detect a length polymorphism both in cloned DNA and genomic DNA samples obtained from 81 unrelated Japanese. During the course of this study, we found two rare types that have never been reported, one of which showed 7 repeats and another that showed a base substitution in the 5th repeating unit.

Materials and Methods

High-molecular-weight DNAs were extracted from the lymphoblastoid cell lines obtained from 81 unrelated Japanese in Hiroshima by Epstein-Barr virus transformation, using a modification of the techniques described by Maniatis et al.[19] In addition, cell lines from both parents of 13 of the 81 subjects were established and examined. A plasmid containing the normal β-globin gene (YF204) and plasmids (YF226 and YF227) with two types of thalassemia genes[20] (Y. Fukumaki, personal communication), were obtained from Dr. Y. Fukumaki of

Kyushu University. Fragments obtained from the plasmids by digestion with *Hpa*I and *Rsa*I were used as cloned DNA samples in examinations with DGGE. The number of ATTTT repeats in the *Hpa*I-*Rsa*I fragments of these clones was determined by sequence analysis after they were subcloned into pGem-3 blue plasmids (Promega, Madison, Wis.). The plasmids YF204, YF227, and YF226 contained 4, 5, and 6 repeats, respectively (Figure 1).

Preparation of RNA probes

Uniformly labeled, single-stranded RNA probes and RNA:DNA duplexes were prepared as described previously.[18] Briefly, in order to obtain DNA templates for synthesizing RNA probes, *Hpa*I-*Rsa*I fragments from YF204 and YF227 (Figure 1) were ligated into the *Sma*I site of pGEM-3 blue plasmids. The plasmids were linearized by *Eco*RI digestion and used as templates for probe preparation using SP6-polymerase (Promega). The resulting probes were designated as antisense YF204I·384 and antisense YF227I·389, since these probes had nucleotide sequences identical (except with U replacing T) to those of the antisense strands

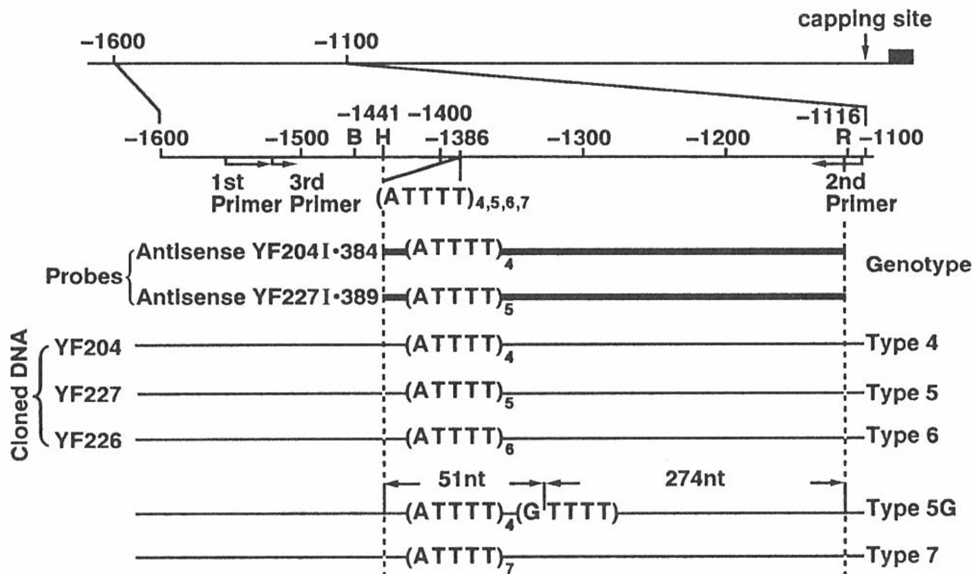

Figure 1. Schematic diagram showing the sequence differences in DNA fragments. Three DNA fragments (YF204, YF227, and YF226) cloned from the human β-globin genes were used for determining the experimental conditions necessary for denaturing gradient gel electrophoresis (DGGE). Four genotypes (5, 6, 5G, and 7) were observed in genomic DNA in the Japanese population (genotype 4 was not observed). Thick lines show the fragments that were ligated into a transcription vector, and the resulting plasmids were used for preparing the RNA probes for DGGE (antisense YF204I) and for the RNase cleavage method (antisense YF227I). The exon (solid block), the capping site, the sequences used as primers (1st and 2nd primer) for polymerase chain reaction and one primer (3rd primer) used for direct sequencing, and the cutting sites of the restriction enzymes used in this experiment are also indicated. Expected sizes of bands obtained after treatment of RNA:DNA heteroduplexes by RNase A and RNase T1 are shown above the diagram of genotype 5G. B: *Bam*HI; H: *Hpa*I; and R: *Rsa*I.

of the DNA fragments between the *Hpa*I and *Rsa*I sites of YF204 and YF227, which were named YF204I·384 and YF227I·389. The antisense YF204I·384 probe has 4 ATTTT repeats and the antisense YF227I·389 probe has 5 ATTTT repeats. The probes made for examination of the cloned DNA samples had a specific activity of approximately 2 to 4 $\times 10^7$ cpm/μg of template, while the probes for genomic DNA samples had a specific activity of approximately 1 to 2 $\times 10^8$ cpm/μg of template.

Preparation of RNA:DNA duplexes for DGGE

The probe with low specific activity (1 $\times 10^5$ to 5 $\times 10^5$ cpm) was hybridized according to the method of Myers et al.[17] with the *Hpa*I-*Rsa*I treated cloned DNA samples (200 ng), while the RNA probe with high specific activity (1 $\times 10^6$ to 2.5 $\times 10^6$ cpm) was hybridized with genomic DNA samples (4.5 μg) treated with *Hpa*I and *Rsa*I. Since antisense YF204I·389 included the transcript of polylinker parts, both ends of the probe could not hybridize with the *Hpa*I-*Rsa*I fragment of a DNA sample (Figure 1). After the hybridization reaction, this single-stranded "over-hang" part of the probe and the unhybridized RNA probe were digested with RNase A (Sigma, St. Louis, Mo.) and RNase T1 (Sankyo, Tokyo) according to the method of Hiyama et al.,[18] with minor modifications. The concentrations of RNase A and RNase T1 were decreased to 0.2 μg/mL and 5 ng/mL for examination of duplexes produced from cloned DNA samples, and to 1.25 μg/mL and 31.25 ng/mL for duplexes derived from genomic DNA samples. In the case of the duplex made from cloned DNA, an aliquot (1 to 2 $\times 10^4$ cpm) of the products was applied to the gel. For genomic DNA, the whole product was applied to the gel.

Electrophoresis

The submerged gel apparatus was made according to Fisher and Lerman[21] and as modified by Takahashi et al.[14] Gels were made of 6.5% polyacrylamide (acrylamide/bisacrylamide = 30/0.8) in a TAE buffer (40 mM Tris, 20 mM sodium acetate, 1 mM EDTA; pH 7.4), with a linear gradient of the denaturant of 0%–25% (100% denaturant = 7 M urea and 40% formamide) parallel to the direction of electrophoresis. The sample wells were prepared with a short stacking gel without denaturant. Electrophoresis was carried out at 200 V for 3 hr submerged in the TAE buffer. A peristaltic pump circulated buffer between the cathodal and anodal chambers. The electrophoresis chambers were maintained at 60°C with a circulating heater. After electrophoresis, the gel was dried and autoradiography was performed at −80°C.

PCR amplification and sequence analyses of DNA fragments

Two primers of 30mer were used to amplify a sequence between positions −1553 and −1133 relative to the capping site of the β-globin gene. The sequence of the first primer shown in Figure 1 lay between positions −1553 and −1524 in the sense strand and the sequence of the second primer lay between positions −1104 and −1133 in the antisense strand of the DNA.

Amplification of the target sequences was carried out following the protocol recommended by Perkin Elmer Cetus Instruments (Norwalk, Conn.). The 100 μL reaction mixture contained approximately 1 μg of genomic DNA in 10 mM Tris

(pH 8.3), 1.5 mM $MgCl_2$, 50 mM KCl, each primer at 1 μM, each dNTP (dATP, dCTP, TTP, dGTP) at 200 μM, and gelatin at 100 μg/mL. The reaction mixture was heated at 94°C for 10 min and rapidly cooled to 4°C. After the addition of 2.5 units of *Thermus aquaticus* (*Taq*) DNA polymerase, the reaction mixture was overlaid with mineral oil and subjected to amplification in the programmable heat block (Perkin Elmer Cetus Instruments). The conditions of thermal cycling were 94°C for 1 min, 60°C for 2 min, 72°C for 3 min, and 30 cycles. In the last cycle, heating at 72°C was extended to 10 min. The PCR products were precipitated with ethanol after phenol-chloroform extraction. Sequence analyses of the amplified DNAs were carried out by two methods. (1) Direct sequencing of the PCR products was carried out as described by Wong et al.[22] and Saiki et al.[23] using approximately 0.7 pmol of the product and 5 pmol of a sequencing primer of 20mer, which has a sequence between positions −1532 and −1513 relative to the capping site of the human β-globin gene in the sense strand (the third primer shown in Figure 1), and AMV reverse transcriptase (Seikagaku Kogyo, Tokyo). (2) Fragments from the PCR products after treatment with *Bam*HI and *Rsa*I were ligated to pGEM-3 blue plasmids, and at least three clones were isolated. Analyses of their sequences were carried out by the chain termination method[24,25] using both SP6- and T7-promoter primers (Promega) and the Klenow fragment of *Escherichia coli* DNA polymerase I.

Gene cloning and DNA sequence analysis

The β-globin gene from one individual who showed a variant band with mobility different from those of the control bands was cloned in a λ-phage as a 5.5-kb *Eco*RI fragment following the method of Maniatis et al.[26,27] A 1.9-kb *Bam*HI fragment was subcloned in a pGem-3 blue plasmid and subjected to partial DNA sequencing by the chain termination method.[24,25]

Analysis of a mismatch by the RNase cleavage method

The analysis was carried out according to the procedure described by Hiyama et al.[18] using the antisense YF227I·389 probe.

Results

Detection of sequence differences in cloned DNA fragments

In order to establish optimum conditions for the detection of base mismatches, we examined three types of duplexes: antisense YF204I·384/YF204, antisense YF204I·84/YF227, and antisense YF204·384/YF226. Antisense YF204I·384/YF204 denotes a duplex produced from the ^{32}P-labeled antisense YF204I·384 probe and the sense strand of the DNA fragment between *Hpa*I and *Rsa*I of YF204.

After a series of preliminary experiments, we determined the optimum electrophoresis conditions to separate three types of duplexes as follows: polyacrylamide concentration = 6.5%, denaturant gradient = 0% to 25%, electrophoresis time = 3 hr, and temperature = 60°C. Typical patterns obtained under these conditions are shown in Figure 2A. Antisense YF204I·384/YF204, which contained no mismatch, migrated farthest. Antisense YF204I·384/YF226, with 10 unpaired nucleotides (nt), showed the slowest mobility. The mobility of antisense

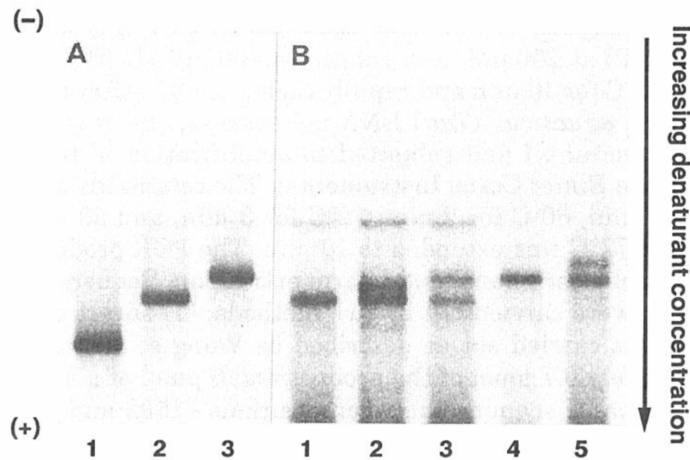

Figure 2. Denaturing gradient gel electrophoresis (DGGE) patterns of RNA:DNA duplexes produced from cloned DNA (A) and genomic DNA samples (B). (A) The duplexes produced from the antisense YF204I·384 and the cloned DNAs. Lane 1: antisense YF204I·384/YF204 (YF204 has 4 repeats). Lane 2: antisense YF204I·384/YF227 (YF227 has 5 repeats). Lane 3: antisense YF204I·384/YF226 (YF226 has 6 repeats). (B) The duplexes with antisense YF204I·384 and genomic DNA. The genotypes are as follows: lane 1, 5/5; lane 2, 5/5G; lane 3, 5/6; lane 4, 6/6; and lane 5, 6/7.

YF204I·384/YF227, with 5 unpaired nt, was intermediate. The conformation of each heteroduplex could not be precisely determined. It has been proposed that the unpaired nucleotides in the DNA strand may loop out of the duplex, thereby decreasing the stability of the duplex. However, these results demonstrated that DGGE under these experimental conditions could detect the length polymorphism of ATTTT.

Screening of genomic DNA

Genomic DNA samples obtained from 81 unrelated persons were examined using DGGE. In the course of screening, four types of bands with different mobilities were clearly separated from each other as shown in Figure 2B. By comparing their mobilities with those of the duplexes made from the cloned DNA fragments with 4, 5, and 6 repeats (Figure 2A), the bands with the fastest mobility, shown in lanes 1, 2, and 3 in Figure 2B, and the bands with the third-fastest mobility, in lanes 3, 4, and 5, were shown to have 5 and 6 repeats. The genotype of an individual showing only the fastest migrating band was designated 5/5, while that of an individual showing only the band with the third-fastest mobility was designated 6/6. Accordingly, the genotype of the individual in lane 3 exhibiting two bands was designated 5/6. In order to confirm the number of repeats, genomic DNA fragments of 450 bp including the tandem repeats from two individuals whose genotypes were termed 5/5 and 6/6 were amplified by PCR. Direct sequencing of the PCR products confirmed that the fragments contained 5 and 6 repeat units, respectively. Then, each PCR product was digested with *Bam*HI and *Rsa*I, and the resulting *Bam*HI-*Rsa*I fragment

that included the repeats was ligated to the pGEM-3 blue plasmid. Appropriately cloned plasmids were obtained, and sequence analyses of the plasmids gave the same results as were shown by direct sequencing.

One individual showed two bands (lane 2, Figure 2B). The mobility of the faster band was identical with that of the antisense YF204I·384/YF226 (type 5), whereas the mobility of the slower band was different from those of the other known bands, suggesting the existence of another sequence abnormality in the *Hpa*I-*Rsa*I fragment. In order to identify the abnormality, a 450-bp sequence of genomic DNA from this individual was amplified by PCR. In addition to performing two types of sequence analyses of the PCR products, an examination by the RNase cleavage method was carried out. An *Hpa*I-*Rsa*I fragment was examined using the antisense YF227I·389 probe with 5 repeats. Three bands with lengths of approximately 320, 270, and 50 nt were observed. The longest band, corresponding to a 320-nt fragment, was derived from the RNA:DNA duplex without a mismatch. However, the two shorter bands seemed to have resulted from RNase -hydrolysis of the 320-nt band at a mismatch. This observation suggested that the sequence abnormality was present at a position approximately 50 bp apart from either one of the restriction sites. The sequence analyses showed that this individual had G at the −1391st nt from the capping site of the β-globin gene on one chromosome 11 and A at the same position on the other chromosome 11 (Figure 1). These observations indicate that an A-to-G substitution occurred at the −1391st nt, which was the first position of the 5th repeating unit. Thus, the genotype of the individual exhibiting two bands in lane 2 was termed 5/5G.

Two individuals exhibited the pattern with two bands shown in lane 5 of Figure 2B. The mobility of the faster band was identical with that of the antisense YF204I·384/YF226 (type 6), while the other band migrated slower than any of the other bands. In order to characterize the abnormality, an *Eco*RI fragment from a DNA sample from one of the two individuals was cloned in the λ-phage vector and then subcloned in the pGEM-3 blue plasmid. In addition, 450-bp fragments from both individuals were amplified by PCR. Sequence analyses of fragments from the plasmid clones and the PCR products revealed that the two individuals had 7 repeats on one chromosome and 6 repeats on the other. Their genotype was termed "6/7." Table 1 presents the distribution of the genotypes and allele frequencies among 81 persons. The frequencies of the alleles "5," "6," "7," and "5G" were 0.6235, 0.3580, 0.0123, and 0.0062, respectively. The distribution of the observed genotypes was in agreement with the allele frequencies predicted by the Hardy-Weinberg equilibrium. The results of family studies carried out on 13 propositi are shown in Table 2. There were no exceptions to the predicted genotypes of a two-allele system among 13 children from 13 matings. It is noteworthy that inheritance of the two types of unreported variants described above was confirmed by the family studies.

Discussion

We previously reported that base substitutions and small deletions in cloned and genomic DNAs could be detected very effectively by DGGE of RNA:DNA duplexes.[14] It was also shown that this method could detect these abnormalities in genomic DNA without amplification. Here we have applied this method to the

Table 1. Distribution of genotypes and their frequencies among 81 unrelated Japanese

Genotype	No. observed	No. expected	Frequency
5/5	31	31.49	
5/6	38	36.16	5 = 0.6235
6/6	9	10.38	6 = 0.3580
6/7	2	0.71	7 = 0.0123
5/5G	1	0.63	5G = 0.0062
Other	0	1.62	
Total	81	80.99	

NOTE: $\chi^2 = 4.46$ ($.2 < p < .3$, $df = 3$).

examination of an ATTTT repeat polymorphism upstream of the β-globin gene. When an RNA probe transcribed from a DNA template with 4 repeats was hybridized with a DNA sample with 4 repeats, the resulting duplex had no single-stranded sequence. When the same probe was hybridized with DNA samples with 5, 6, and 7 repeats, single-stranded DNA sequences remained in the resulting duplexes since the probe was shorter than the samples. The single-stranded sequences must have looped out from the duplexes and made the

Table 2. Distribution of genotypes among 13 families

Genotypes of propositi	No. of propositi	Genotypes of parents: No. of cases	Genotypes of parents: MO FA
5/5	3	2	5/5 × 5/5
		1	5/5 × 5/6
5/6	7	2	5/6 × 5/5
		3	5/6 × 5/6
		2	5/5 × 6/6
6/6	1	1	6/6 × 6/6
5/5G	1	1	5/5G × 5/5
6/7	1	1	5/6 × 6/7

NOTE: MO = mother; FA = father.

duplexes labile, thereby causing them to have less mobility on the DGGE gel. In our experiments using cloned and genomic DNA samples, the difference in the number of repeat units was clearly demonstrated through differential migration of the duplexes on the DGGE gel. In addition, the experiments also demonstrated that the duplex from DNA with 5 repeat units and one base substitution in the 5th repeat unit was distinguishable from DNA with 5 repeats without the substitution. The family studies (Table 2) using this method demonstrated that the sequence difference that produced the polymorphism was inherited in a Mendelian fashion. This indicates that the alteration of melting behavior observed in this study was not caused by artifacts from base methylation or partial digestion of examined DNAs with restriction enzymes. The sequence analyses of this region confirmed the results obtained by DGGE.

In our preliminary experiment, the DGGE of PCR-amplified DNA also identified these polymorphisms. However, in our study, RNA:DNA duplexes were chosen as a target instead of amplified DNA because we wanted to exclude any ambiguity that might result from low-fidelity amplification of the repeat units.[28]

Spritz[8] first demonstrated the ATTTT tandem repeat polymorphism by sequence analyses of cloned human β-globin genes. Subsequently, Kazazian et al.[29] conducted a similar study using chromosomes obtained from Asian Indians. When the cases obtained from these two studies are combined, 9 of the 12 chromosomes had 5 repeat units, 2 had 6, and the remaining one had 4 repeat units. Thus, the frequencies of β-globin genes with 5, 6, and 4 repeat units were 0.75, 0.17, and 0.08, respectively. Our study of 81 persons also showed types 5 and 6 to be polymorphic. In addition, the allele with 7 repeat units (type 7), which had never been reported previously, was detected in two cases. Another case having 5 repeat units showed a base substitution from A to G in the −1391st residue (type 5G). The gene frequencies of types 5, 6, 7, and 5G were 0.624, 0.358, 0.012, and 0.006, respectively. No allele of type 4 was detected in the Japanese tested in this study. If data on this polymorphism in various populations are accumulated, it will be possible to determine whether the frequency of each genotype is peculiar to the Japanese.

In the previous studies[8,29] the length polymorphism was investigated by sequence analysis of cloned DNA fragments from β-globin genes. Therefore, it is possible that events such as duplications and deletions observed in the tandem repeat array of β-globin genes may have occurred in the cloning or subsequent propagation in bacteria. However, this possibility was excluded in our study because genomic DNA samples were examined by DGGE without cloning. Thus, variant sequences observed with this technique accurately reflect a gene heteromorphism among individuals.

Polymorphisms attributable to differences in the number of tandem repeats of short (11–60 bp) nucleotide sequences have been observed in or around several kinds of genes. Two general mechanisms have been proposed to explain the events of duplication/deletion in the repeat array. Strand slippage and mispairing may occur between tandem repeat units on the leading and trailing strands during DNA replication.[30] Alternatively, duplication and deletion of an integral number of repeat units may result from nonhomologous recombination between repeat units on sister chromatids or the homologous chromosomes.[31] The latter mechanism, that is, nonhomologous recombination at a high frequency, may

explain the increase or decrease in the number of tandem repeats in the VNTR[7] sequences and minisatellite DNA.[6] Some "core" elements of VNTR show an apparent relationship to the *chi*-sequence of the λ-phage, which has been implicated as a hotspot for *rec* A-mediated recombination in *E. coli*.[32] The presence of a sequence similar to the *chi*-sequence probably could account for the generation of highly mutable alleles.

Considering the events necessary for the occurrence of the tandem repeat polymorphism, it is extremely likely that the types showing reduced frequencies were produced from the most common type. If chromosomes containing 6 repeats have arisen by unequal crossing-over of chromosomes with 5 repeats, chromosomes containing 4 repeats should be observed at a similar frequency. Unless the possibilities of selection or drift are considered in the determination of the allele frequency, the absence of chromosomes with 4 repeats makes it very unlikely that this polymorphism has been caused by nonhomologous recombination. Moreover, there is no region homologous to the *chi*-sequence in the ATTTT repeats. For these reasons, we think that the "strand slippage and mispairing" mechanism seems most likely to account for duplication or deletion events at the β-globin tandem repeat array that produced the polymorphism observed in this study.

Acknowledgments

We are grateful to N. Kosaka, H. Omine, and Y. Kimura for their excellent technique in DGGE. We also thank J. Kaneko for her technical work throughout this study. We are much indebted to E. S. Frankel and S. Abrahamson for their assistance in the preparation of this manuscript.

References

1. Farabaugh PJ, Schmeissner U, Hofer M, Miller JH: Genetic studies of the *lac* repressor. VII. On the molecular nature of spontaneous hotspots in the *lac*I gene of *Escherichia coli*. J Mol Biol 126:847–63, 1978
2. Wyman AR, White R: A highly polymorphic locus in human DNA. Proc Natl Acad Sci USA 77:6754–8, 1980
3. Bell GI, Selby MJ, Rutter WJ: The highly polymorphic region near the human insulin gene is composed of simple tandemly repeating sequences. Nature 295:31–5, 1982
4. Proudfoot NJ, Gil A, Maniatis T: The structure of the human zeta-globin gene and a closely linked, nearly identical pseudogene. Cell 31:553–63, 1982
5. Capon DJ, Chen EY, Levinson AD, Seeburg PH, Goeddel DV: Complete nucleotide sequences of the T24 human bladder carcinoma oncogene and its normal homologue. Nature 302:33–7, 1983
6. Jeffreys AJ, Wilson V, Thein SL: Hypervariable "minisatellite" regions in human DNA. Nature 314:67–73, 1985
7. Nakamura Y, Leppert M, O'Connell P, Wolff R, Holm T, Culver M, Martin C, Fujimoto E, Hoff M, Kumlin E, White R: Variable number of tandem repeat (VNTR) markers for human gene mapping. Science 235:1616–22, 1987

8. Spritz RA: Duplication/deletion polymorphism 5′ - to the human β globin gene. Nucleic Acids Res 9:5037–47, 1981
9. Myers RM, Maniatis T: Recent advances in the development of methods for detecting single-base substitutions associated with human genetic diseases. Cold Spring Harb Symp Quant Biol LI:275–84, 1986
10. Lerman LS, Silverstein K, Grinfeld E: Searching for gene defects by denaturing gradient gel electrophoresis. Cold Spring Harb Symp Quant Biol LI:285–97, 1986
11. Noll WW, Collins M: Detection of human DNA polymorphisms with a simplified denaturing gradient gel electrophoresis technique. Proc Natl Acad Sci USA 84:3339–43, 1987
12. Myers RM, Maniatis T, Lerman LS: Detection and localization of single base changes by denaturing gradient gel electrophoresis. In *Methods in Enzymology* (Vol. 155). Ed. by R. Wu. New York, Academic Press, 1987. pp 501–27
13. Smith FI, Parvin JD, Palese P: Detection of single base substitution in influenza virus RNA molecules by denaturing gradient gel electrophoresis of RNA–RNA or DNA–RNA heteroduplexes. Virology 150:55–64, 1986
14. Takahashi N, Hiyama K, Kodaira M, Satoh C: An improved method for the detection of genetic variations in DNA with denaturing gradient gel electrophoresis. Mutat Res 234:61–70, 1990
15. Lerman LS, Fischer SG, Hurley I, Silverstein K, Lumelsky N: Sequence-determined DNA separations. Annu Rev Biophys Bioeng 13:399–423, 1984
16. Winter E, Yamamoto F, Almoguera C, Perucho M: A method to detect and characterize point mutations in transcribed genes: Amplification and overexpression of the mutant c-Ki-*ras* allele in human tumor cells. Proc Natl Acad Sci USA 82:7575–9, 1985
17. Myers RM, Larin Z, Maniatis T: Detection of single base substitutions by ribonuclease cleavage at mismatches in RNA:DNA duplexes. Science 230:1242–6, 1985
18. Hiyama K, Kodaira M, Satoh C: Detection of deletions, insertions and single nucleotide substitutions in cloned β-globin genes and new polymorphic nucleotide substitutions in β-globin genes in a Japanese population using ribonuclease cleavage at mismatches in RNA:DNA duplexes. Mutat Res 231:219–31, 1990
19. Maniatis T, Fritsch EF, Sambrook J: Isolation of high-molecular-weight, eukaryotic DNA from cells grown in tissue culture. In *Molecular Cloning; A Laboratory Manual.* Cold Spring Harbor, N.Y., Cold Spring Harbor Laboratory, 1982. pp. 280–1
20. Takihara Y, Matsunaga E, Nakamura T, Lin S-T, Lee H-T, Fukumaki Y, Takagi Y: One base substitution in IVS-2 causes a β^{+}–thalassemia phenotype in a Chinese patient. Biochem Biophys Res Commun 121:324–30, 1984
21. Fischer SG, Lerman LS: Two-dimensional electrophoretic separation of restriction enzyme fragments of DNA. In *Methods in Enzymology* (Vol. 68). Ed. by R. Wu. New York, Academic Press, 1979. pp. 183–91
22. Wong C, Dowling CE, Saiki RK, Higuchi RG, Erlich HA, Kazazian HH Jr: Characterization of β-thalassaemia mutations using direct genomic sequencing of amplified single copy DNA. Nature 330:384–6, 1987

23. Saiki RK, Gelfand DH, Stoffel S, Scharf SJ, Higuchi R, Horn GT, Mullis KB, Erlich HA: Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. Science 239:487–91, 1988

24. Sanger F, Coulson AR, Barrell BG, Smith AJH, Roe BA: Cloning in single-stranded bacteriophage as an aid to rapid DNA sequencing. J Mol Biol 143:161–78, 1980

25. Hattori M, Sakaki Y: Dideoxy sequencing method using denatured plasmid templates. Anal Biochem 152:232–8, 1986

26. Maniatis T, Fritsch EF, Sambrook J: Introduction of plasmid and bacteriophage DNA into *Escherichia coli*. In *Molecular Cloning; A Laboratory Manual*. Cold Spring Harbor, N.Y., Cold Spring Harbor Laboratory, 1982. pp. 247–68

27. Maniatis T, Fritsch EF, Sambrook J: The identification of recombinant clones. In *Molecular Cloning; A Laboratory Manual*. Cold Spring Harbor, N.Y., Cold Spring Harbor Laboratory, 1982. pp. 309–28

28. Jeffreys AJ, Wilson V, Neumann R, Kete J: Amplification of human minisatellites by the polymerase chain reaction: Towards DNA fingerprinting of single cells. Nucleic Acids Res 16:10953–71, 1988

29. Kazazian HH Jr, Orkin SH, Antonarakis SE, Sexton JP, Boehm CD, Goff SC, Waber PG: Molecular characterization of seven β-thalassemia mutations in Asian Indians. EMBO J 3:593–6, 1984

30. Kornberg A, Bertsch LL, Jackson JF, Khorana HG: Enzymatic synthesis of deoxyribonucleic acid. XVI. Oligonucleotides as templates and the mechanism of their replication. Proc Natl Acad Sci USA 51:315–23, 1964

31. Perelson AS, Bell GI: Mathematical models for the evolution of multigene families by unequal crossing over. Nature 265:304–10, 1977

32. Stahl FW: Special sites in generalized recombination. Annu Rev Genet 13:7–24, 1979