British Journal of Haematology, 2003, **123**, 942–947

Determination of the breakpoint and molecular diagnosis of a common α-thalassaemia-1 deletion in the Indian population

R. V. Shaji,[1] S. E. Eunice,[1] S. Baidya,[2] A. Srivastava[1] and M. Chandy[1] *Departments of* [1]*Haematology and* [2]*Clinical Pathology, Christian Medical College, Vellore, India*

Received 26 May 2003; accepted for publication 1 September 2003

Summary. The previously described South African type α-thalassaemia-1 mutation was identified in Indian HbH patients using a polymerase chain reaction (PCR) strategy. A multiplex PCR assay was devised to detect heterozygotes and homozygotes. This α-thalassaemia-1 mutation was found to be the commonest determinant causing HbH disease in this population. In one family this mutation was found in combination with a novel splice donor mutation α2 IVS I-1 (G → A). Characterization of the breakpoint junction sequence revealed, in addition to a 23 kb deletion, that there was an addition of ~160 bp bridging the breakpoints. Similar to other deletions in the α-globin gene cluster, there is an *Alu* repeat-mediated mechanism for the origin of the deletion.

Keywords: α-thalassaemia, PCR, India, *Alu*

The α-thalassaemias are genetic defects characterized by the decrease or complete suppression of synthesis of the α-globin polypeptide chains of haemoglobin. They are the most common single-gene diseases in the world. The α-globin genes are duplicated (α1 and α2) and located within the α gene cluster $\zeta 2-\psi\zeta 1-\psi\alpha 2-\psi\alpha 1-\alpha 2-\alpha 1-\theta 1$ on the distal short arm region of chromosome 16 (Higgs *et al*, 1989). This cluster contains four genes and three pseudogenes. Deletions of one (−α) or both (——) of these *cis*-linked genes are the most common causes of the α-thalassaemias. Around 25 different two-gene *cis* deletions (− −) and eight single-gene deletions (−α) have been reported so far in different populations (Weatherall & Clegg, 2001).

Homozygosity for two gene deletions, which results in the loss of four α globin genes (− −/− −), causes Hb Bart's hydrops fetalis syndrome, a condition that generally results in severe intrauterine anaemia and fetal death in the last trimester of pregnancy or shortly after birth (Lie-Injo, 1962). Such pregnancies are accompanied by increased risks of maternal complications. The loss of three α-globin genes in the compound heterozygotes of double gene and single gene deletions (−α/− −) or in the compound heterozygotes of a double gene deletion and a non-deletional α-thalassaemia mutation (− −/$\alpha^{T}\alpha$) result in a less severe condition known as haemoglobin H (HbH) disease which causes chronic haemolytic anaemia of variable severity in children and adults (Galanello *et al*, 1992); (Tamary *et al*, 1998).

Correspondence: R.V. Shaji, PhD, Department of Haematology, Christian Medical College, Vellore 632004, India. E-mail: rvshaji@cmcvellore.ac.in

Genetic diagnosis of α-thalassaemia is significant, not only for elucidating the molecular pathology of the condition, but also for providing carrier detection and prenatal diagnosis in families at risk of Hb Bart's hydrops fetalis syndrome and HbH disease and its interaction with homozygous β-thalassaemia ameliorating or aggravating the phenotype of these patients.

Measurement of Hb Bart's levels in cord blood samples indicated that α-thalassaemia is extremely common among certain groups in the Indian subcontinent (Misra *et al*, 1991; Hassall *et al*, 1998). In tribal populations it reaches a frequency of 0·9, which is as high as that recorded anywhere in the world (Fodde *et al*, 1988, 1991). The DNA analysis carried out in certain regions of India has shown that $-\alpha^{3.7}$ and $-\alpha^{4.2}$ are the common single gene deletions, as in other populations (Kulozik *et al*, 1988). A 23 kb double gene deletion in the α-globin gene cluster that removes both α2 and α1 genes causing α-thalassaemia-1 ($--^{SA}$), which was originally found in a patient from the 'cape' coloured population of South Africa (Vandenplas *et al*, 1987), was also detected in an individual from Gujarat in combination with heterozygous HbS (Drysdale & Higgs, 1988). Isolated cases of HbH disease and Hb Bart's hydrops fetalis have been reported in the Indian population. Nevertheless, none of the surveys published so far have documented the genetic basis of the disease and the actual frequency of the double gene deletions that are responsible for HbH disease and Hb Bart's hydrops fetalis is not known.

We describe here a polymerase chain reaction (PCR)-based detection of $--^{SA}$ allele in a cohort of HbH disease patients in the Indian population. We have characterized

the breakpoint junction sequence of this deletion and developed a multiplex PCR protocol for the rapid and reliable detection of α-thalassaemia deletions in the Indian population.

PATIENTS AND METHODS

Patients. Seventeen Indian patients with HbH disease, diagnosed on the basis of clinical and haematological investigations, and their parents were the subjects of the present study.

Analysis of seven common deletions. Seven common deletions in the α-globin gene cluster causing α-thalassaemia ($-\alpha^{3.7}$, $-\alpha^{4.2}$, $--^{SEA}$, $--^{THAI}$, $--^{FIL}$, $--^{MED}$ and $--^{20.5}$) were analysed in these samples using a previously described single tube multiplex PCR method (Tan *et al*, 2001).

Detection of $--^{SA}$ *mutation.* Previous restriction mapping studies have shown that the left breakpoint of the $--^{SA}$ deletion lies downstream to that of the $--^{20.5}$ deletion and the right breakpoint lies upstream of those of the $--^{SEA}$ and the $--^{THAI}$ deletions and downstream of $--^{FIL}$ deletion (Fig 1A). Using the primers described earlier for the detection of these deletions (Tan *et al*, 2001), three separate PCR reactions were carried out with the forward primer for $--^{20.5}$ deletion (20·5-F) and the reverse primers, for $--^{SEA}$, $--^{THAI}$ and $--^{FIL}$ deletions (SEA-R, THAI-R and FIL-R) (Fig 1A) with the genomic DNA extracted from one of the patients with HbH disease. The amplified product obtained with 20·5-F and THAI-R was sequenced for at least 200 bases away from the forward and reverse primers to confirm specific amplification of the α-globin gene cluster. For the characterization of the breakpoint, another PCR was carried out with the primers α52 and α53 to obtain a shorter PCR product. This product was sequenced by primer walking with the primers α52, α53, α50, α55 and α56. The primers used and their locations are presented in Tables I and II.

Modified multiplex PCR to detect the common deletions in the Indian population. We incorporated PCR-based detection of the $--^{SA}$ allele into the single tube multiplex assay described by Tan *et al* (2001) to enable simultaneous screening for the α-thalassaemia deletions prevalent in the Indian population. Each 50 μl multiplex PCR reaction contained 200 μmol/l of each dNTP, 1XQ solution (Qiagen, Hilden, Germany), 1·25 U HotStarTaq Polymerase in the supplied buffer containing 1·5 mmol/l $MgCl_2$, 0·2 μmol/l of the primers α52 and α53 were used for the detection of $--^{SA}$. The concentration of the other primers and the cycling conditions for the amplification reaction were as described by Tan *et al* (2001).

Alu repeat analysis at the breakpoint junction. The *Alu* repetitive elements near the breakpoints were analysed using 'Repeat Masker' (http://repeatmasker.genome.washington.edu/cgi-bin/RepeatMasker).

Analysis of point mutations in the α-globin genes. The α2 and α1 genes were amplified as described earlier (Harteveld *et al*, 1996) and the amplified products were subjected to DNA sequencing using the Big Dye terminator kit (Applied Biosystems, Foster City, CA, USA) on the ABI 310 genetic analyzer (Applied Biosystems).

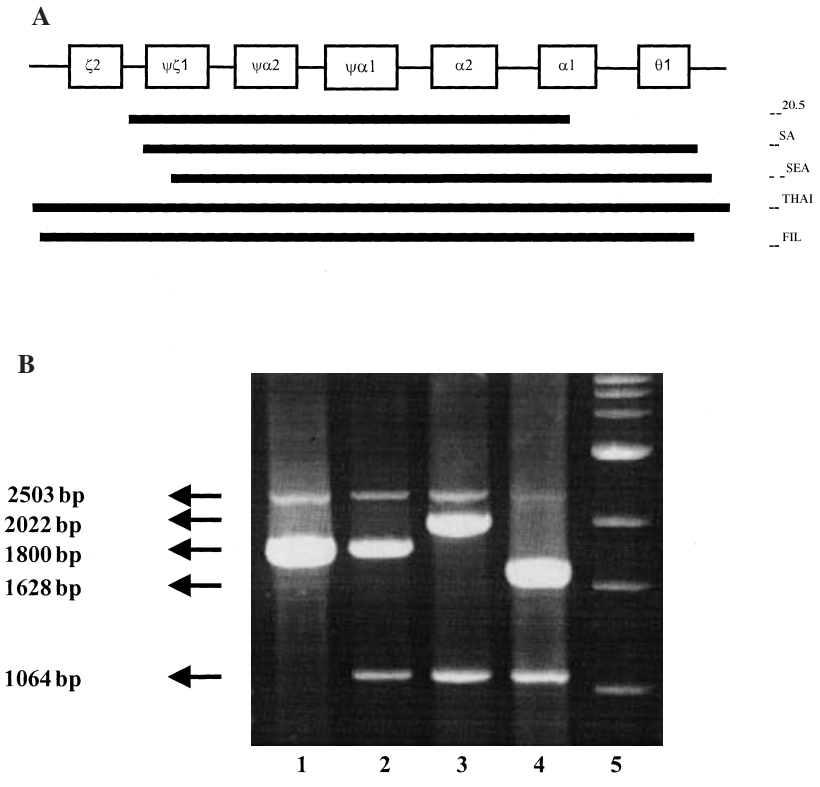

Fig 1. (A) Schematic representation of α-globin gene cluster indicating the extent of the $--^{SA}$ deletion and its neighbouring deletions. (B) Representative results of multiplex-PCR from DNA samples with various α-globin genotypes. Lane 1: αα/αα; Lane 2: αα/$--^{SA}$; Lane 3: $-\alpha^{3.7}$/$--^{SA}$; Lane 4: $-\alpha^{4.2}$/$--^{SA}$ and Lane 5: molecular weight marker. With the primers flanking the breakpoints 2022 bp, 1628 bp and 1064 bp were produced for $-\alpha^{3.7}$, $-\alpha^{4.2}$ and $--^{SA}$ deletions respectively. A 2503 bp fragment from *LIZ* gene was amplified in all the genotypes and a 1800 bp fragment from α2 gene was amplified in normal individuals (lane 1) and in those who were heterozygous for the deletions (lane 2) but not for compound heterozygotes (lanes 3 and 4).

Table I. Sequence and location of primers used for PCR.

Primer	Sequence 5′–3′	Location
20·5-F*	GCCCAACATCCGGAGTACATG	17904–17924 in Z84721
THAI R*	CAAGTGGGCTGAGCCCTTGAG	1241–1221 in Z69706
SEA R*	AGCCCACGTTGTGTTCATGGC	3817–3797 in Z69706
FIL R*	ATAACCTTTATCTGCCACATGTAGC	570–546 in Z69706
α52	CTCTTTCTCTCGGAGCCCTT	19286–19303 in Z84721
α53	ACTCGAGCTACCCCAAGGAT	975–956 in Z69706

*Primers described by Tan *et al* (2001).

Table II. Sequence and location of the primers used for sequencing.

Primer	Sequence 5′–3′	Location
α50	GCAGTGGGACAATCTTGGTT	19656–19675 in Z84721
α56	TGGTAGCCTGTAATCCCAGC	19835–19854 in Z84721
α55	CTCAGCTCACACCAACCTCC	826–807 in Z69706

RESULTS

Initial screening of seven common deletions by the previously described multiplex PCR (Tan *et al*, 2001) showed that most of the patients with HbH disease had a pattern of homozygosity for $-\alpha^{3.7}$ or $-\alpha^{4.2}$ deletions; only one of the parents was heterozygous for either $-\alpha^{3.7}$ or $-\alpha^{4.2}$ deletions. This discrepancy usually occurs when the patient is compound heterozygous for a large deletion causing α-thalassaemia-1 and $-\alpha^{3.7}$ or $-\alpha^{4.2}$ deletions. The search for the $--^{SA}$ deletion in these samples by PCR with the primers used for the detection of the neighbouring deletions resulted in successful amplification with the reverse primers for $--^{SEA}$ and $--^{THAI}$, but not with $--^{FIL}$. These results correlated with the genotyping results of $--^{SEA}$, $--^{SA}$, $--^{20.5}$, $--^{THAI}$ and $--^{FIL}$ deletions, which showed that the 5′ breakpoint of the $--^{SA}$ deletion was inside the deleted region of the $--^{20.5}$ deletion and the 3′ breakpoint was inside the deleted region of the $--^{THAI}$ and $--^{20.5}$ deletions and outside the $--^{FIL}$ deletion. The newly developed multiplex PCR revealed that seven of 17 patients had the genotype $--^{SA}/-\alpha^{3.7}$ while three had $--^{SA}/-\alpha^{4.2}$. The results obtained by the multiplex PCR in different genotypes are presented in Fig 1B. Three patients from one family had $--^{SA}$ in combination with a novel splice donor mutation, IVS 1-1 (G → A) in the α2 gene, which was identified by DNA sequencing (Fig 2). Among the other four patients, two were homozygous for the previously reported polyadenylation signal mutation in the α2 globin gene (AATAAA → AATA− −) and two had the genotypes $--^{MED}/-\alpha^{3.7}$ and $--^{SEA}/-\alpha^{3.7}$.

Haematological data and genotypes of the patients with the $--^{SA}$ deletion are illustrated in Table III.

The DNA sequencing of the PCR product obtained with the primers α52 and α53 revealed that the 5′ and 3′ breakpoints of the $--^{SA}$ deletion were at nucleotides 19 915 and 696 of the sequences with GenBank accession numbers Z84721 and Z69706, respectively, and the two breakpoints were separated by an insertion of 157 bp of DNA that had originated from elsewhere in the genome (Fig 3).

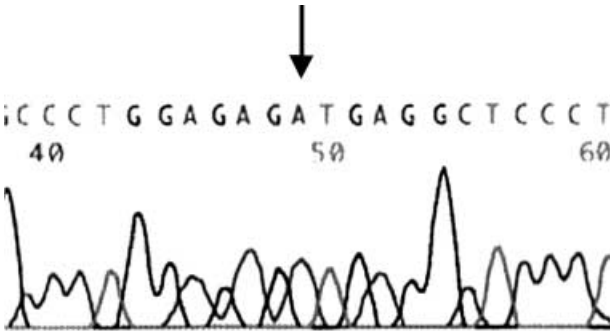

Fig 2. Nucleotide sequencing results showing a novel splice donor site mutation IVS I-1 (G → A), which was found in the compound heterozygous state with $--^{SA}$ deletion in three siblings with HbH disease. The arrow indicates the position of the mutation.

Analysis of *Alu* repeats in the breakpoints revealed that the 5′ breakpoint occurred in a 144 bp *Alu* Sg/x sequence and the 3′ breakpoint occurred in a 307 bp *Alu* Sx sequence. The deletion removed 69 bases of 5′ *Alu* Sgx and 102 bases of 3′*Alu* Sx along with the ∼23 kb intervening sequence and bridged the two breakpoints with 157 bp of DNA. The inserted sequence had two sections of *Alu* sequences; 59 bp at the 5′ region with sequence similarity to the *Alu* Yd3a1 sequence and 98 bp at the 3′ region to *Alu* Sc. This rearrangement of deletion and insertion of DNA changed the pattern of *Alu* sequences at the region where the mutation occurred to *Alu* Sgx in the 3′ region with three bases different from the original *Alu* Sgx present in the normal α-globin gene cluster and *Alu* Sc in the 5′ region.

DISCUSSION

Only two cases of α-thalassaemia with $--^{SA}$ allele have been reported so far. In this report we have described an additional 13 Indian HbH patients with this mutation. This suggests that this mutation is prevalent in the Indian population, is more common than previously reported and

Table III. The origin, haematological and genetic data of patients with the $--^{SA}$ deletion.

Family/patient no.	Relationship	Age (years)	Origin	Hb (g/dl)	MCV (fl)	MCH (pg)	HbF (%)	HbA2 (%)	α-globin genotype
Family									
1	Propositus	38	West Bengal	8·4	62	18·1	0·2	1·6	$-\alpha^{3.7}/--^{SA}$
	Father	60		12·9	80·7	26·3	0·2	3·3	$-\alpha^{3.7}/\alpha\alpha$
	Mother	50		11·2	67·7	19·3	0·3	2·3	$\alpha\alpha/--^{SA}$
	Daughter 1	4		10·6	59·8	21·1	0·1	3·1	$\alpha\alpha/--^{SA}$
	Daughter 2	7		12·7	83·9	25·5	0·2	2·6	$-\alpha^{3.7}/\alpha\alpha$
	Sister 1	35		12·6	85·7	27·9	0·2	2·8	$\alpha\alpha/\alpha\alpha$
	Sister 2	35		13·5	71·8	21·0	0·4	2·6	$\alpha\alpha/--^{SA}$
	Sister 3	29		11·0	84·1	24·8	0·4	2·5	$-\alpha^{3.7}/\alpha\alpha$
	Brother	27		9·8	59·6	17·0	0·2	1·8	$-\alpha^{3.7}/--^{SA}$
2	Propositus	13	Orissa	9·8	66·0	17·6	0·1	1·7	$-\alpha^{3.7}/--^{SA}$
	Mother	42		12·8	87·8	28·3	0·1	2·4	$-\alpha^{3.7}/\alpha\alpha$
3	Propositus	6	West Bengal	8·3	56·8	17·3	na	na	$-\alpha^{4.2}/--^{SA}$
	Father	32		12·6	77·6	25·0	0·1	2·4	$-\alpha^{3.7}/\alpha\alpha$
	Mother	27		9·1	58·1	20·1	0·4	HbE = 18·8	$\alpha\alpha/--^{SA}$
4	Propositus	23	West Bengal	7·7	67·4	19·8	1	1·7	$\alpha^{IVS\ I\text{-}1(G\rightarrow A)}\alpha/--^{SA}$
	Father	52		12·1	69·3	21·3	0·2	3·2	$\alpha^{IVS\ I\text{-}1(G\rightarrow A)}\alpha/\alpha\alpha$
	Mother	48		10·2	66·9	20·2	1·9	2·2	$\alpha\alpha/--^{SA}$
	Brother 1	15		9·1	63·6	18·7	1	1·1	$\alpha^{IVS\ I\text{-}1(G\rightarrow A)}\alpha/--^{SA}$
	Sister 1	20		7·5	63·7	17·7	0·8	1·5	$\alpha^{IVS\ I\text{-}1(G\rightarrow A)}\alpha/--^{SA}$
Individuals									
1		26	Orissa	10·1	61·0	19·4	0·2	2·1	$-\alpha^{3.7}/--^{SA}$
2		15	Punjab	8·2	74·4	20·1	0·3	2·3	$-\alpha^{3.7}/--^{SA}$
3		42	West Bengal	9·2	64·1	18·5	0·7	2·0	$-\alpha^{4.2}/--^{SA}$
4		26	Madhya Pradesh	8·7	63·9	19·7	0·6	1·9	$-\alpha^{4.2}/--^{SA}$
5		33	West Bengal	9·5	59·6	19·1	0·4	1·9	$-\alpha^{3.7}/--^{SA}$
6		32	West Bengal	10·3	69·5	19·9	0·2	1·9	$-\alpha^{3.7}/--^{SA}$

na, not available/not analysed.

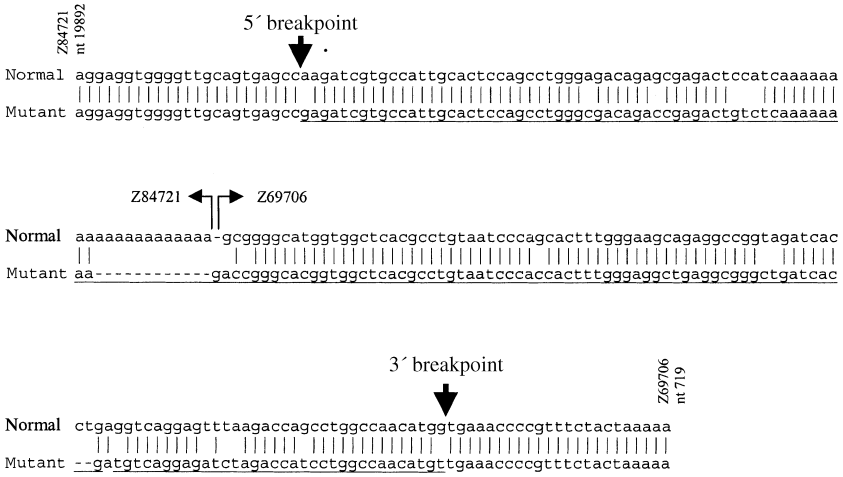

Fig 3. Alignment of the sequences surrounding the breakpoint of $--^{SA}$ deletion with the normal sequences in the GenBank accessions Z84721 and Z69706. The vertical bars show identical nucleotides. The underlined region indicates the inserted section of DNA that bridged the breakpoints in the mutant chromosome.

is, in fact, the most common α-thalassaemia determinant causing HbH disease in India.

The α-globin gene cluster is interspersed with *Alu*-family repeats, which are commonly associated with DNA recombination. Most of the deletions causing α-thalassaemia are simple linear chromosomal deletions between two highly homologous *Alu* repeats in the α-globin gene cluster. In the deletion $--^{MED}$ there was an insertion of a new section of DNA, originating from upstream of the α-globin gene cluster, bridging the two breakpoints (Nicholls *et al*, 1987). The inserted DNA appears to have been incorporated into the junction in a manner which suggests that the upstream

segment lies close to the breakpoint regions during replication. Comparison of the mutant sequence in the $--^{SA}$ allele across the breakpoint and the normal sequence in the α-globin gene cluster showed that there was a deletion of 23 kb and an insertion of 157 bases at the breakpoint junction. Sequence homology search of this inserted sequence in the human genome carried out using FastaA (http://www.dna.affrc.go.jp/htdocs/Blast/fasta.html) found that the reverse complementary sequence, which has a close similarity (155 of the 157 nucleotides) to the inserted sequence in the breakpoints, is present in chromosome 16 in the nucleotides 66097–66253 of the sequence with GenBank accession number AC007604. In this location, these DNA sequences lie in an *Alu* Sc-*Alu* Sx sequence with part of the sequences in the *Alu* Sc and part of the sequence in *Alu* Sx. However, the exact physical location and orientation of this sequence (GenBank accession number AC007604) from the α-globin gene cluster is not yet available. It is possible that a mechanism similar to that involved in $--^{MED}$ is responsible for the origin of $--^{SA}$ mutation. As the inserted fragment has *Alu* sequences, it may also be possible that a homologous recombination between the *Alu* sequences, present near the breakpoint of the deletion and those present elsewhere in chromosome 16, is responsible for the deletion and insertion of DNA in this mutation.

It was confirmed that the amplified product with the primers α52 and α53 was from the chromosome carrying the $--^{SA}$ deletion, as the amplification was only obtained in those patients whose haematological parameters indicated a possible α-thalassaemia-1 mutation. Among those who were compound heterozygous for $--\alpha^{3.7}/--^{SA}$, there was no amplification of the α2 gene, which was used as an internal control to amplify α-globin gene without any deletions. This is well illustrated by the analysis of a large family with 10 members, of whom two had HbH disease (Fig 4).

Isolated cases of HbH disease have been reported in the Indian population. The present study analysed 17 HbH patients in the Indian population for the underlying genotypes and showed that compound heterozygosity for single and double gene deletions in the α-globin gene cluster ($--\alpha/--$) is the most common genotype that causes HbH disease in this population, as in other populations. The modified multiplex PCR should simplify and subsequently reduce the cost of screening for α-thalassaemia deletions in suspected cases, particularly for the molecular diagnosis of HbH disease and prenatal diagnosis of Hb Bart's disease.

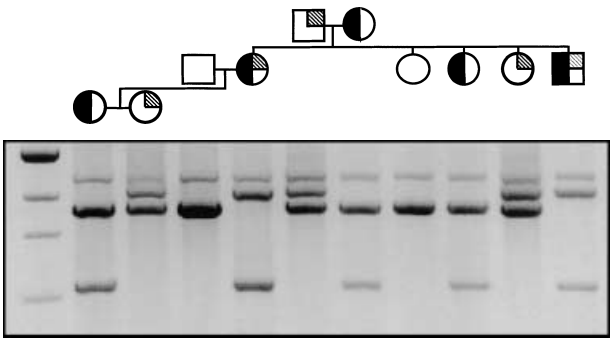

Fig 4. Analysis of α-globin gene deletions with the multiplex PCR in the family with 10 members, two of whom had HbH disease. The band produced by the primers α52 and α53 was present only in those who had α-thalassaemia-1 and it is inherited in a Mendelian fashion. In those who had the genotype $--\alpha^{3.7}/--^{SEA}$, there was no amplification of the α2 gene that confirms that genotype in these individuals.

ACKNOWLEDGMENTS

This study was supported by the Department of Biotechnology, Government of India (grant BT/RO948/Med/13/034/98). We are grateful to Dr G. Garewal for providing the DNA sample of her patient.

REFERENCES

Drysdale, H.C. & Higgs, D.R. (1988) Alpha-thalassaemia in an Asian Indian. *British Journal of Haematology*, **68**, 264–xxx.

Fodde, R., Losekoot, M., van den Broek, M.H., Oldenburg, M., Rashida, N., Schreuder, A., Wijnen, J.T., Giordano, P.C., Nayudu, N.V., Khan, P.M. & Bernini, L.F. (1988) Prevalence and molecular heterogeneity of alfa+ thalassemia in two tribal populations from Andhra Pradesh, India. *Human Genetics*, **80**, 157–160.

Fodde, R., Harteveld, C.L., Losekoot, M., Giordano, P.C., Khan, P.M., Nayudu, N.V. & Bernini, L.F. (1991) Multiple recombination events are responsible for the heterogeneity of alpha(+)-thalassemia haplotypes among the forest tribes of Andhra Pradesh, India. *Annals of Human Genetics*, **55**, 43–50.

Galanello, R., Aru, B., Dessi, C., Addis, M., Paglietti, E., Melis, M. A., Cocco, S., Massa, P., Giagu, N. & Barella, S. (1992) HbH disease in Sardinia: molecular, hematological and clinical aspects. *Acta Haematologica*, **88**, 1–6.

Harteveld, K.L., Heister, A.J., Giordano, P.C., Losekoot, M. & Bernini, L.F. (1996) Rapid detection of point mutations and polymorphisms of the alpha-globin genes by DGGE and SSCA. *Human Mutation*, **7**, 114–122.

Hassall, O.W., Tillyer, M.L. & Old, J.M. (1998) Prevalence and molecular basis of alpha thalassaemia in British South Asians. *Journal of Medical Screening*, **5**, 31–33.

Higgs, D.R., Vickers, M.A., Wilkie, A.O., Pretorius, I.M., Jarman, A.P. & Weatherall, D.J. (1989) A review of the molecular genetics of the human alpha-globin gene cluster. *Blood*, **73**, 1081–1104.

Kulozik, A.E., Kar, B.C., Serjeant, G.R., Serjeant, B.E. & Weatherall, D.J. (1988) The molecular basis of alpha thalassemia in India. Its interaction with the sickle cell gene. *Blood*, **71**, 467–472.

Lie-Injo, L.E. (1962) Alpha-chain thalassemia and hydrops fetalis in Malaya: report of five cases. *Blood*, **20**, 581–590.

Misra, R.C., Ram, B., Mohapatra, B.C., Das, S.N. & Misra, S.C. (1991) High prevalence and heterogenicity of thalassaemias in Orissa. *The Indian Journal of Medical Research*, **94**, 391–394.

Nicholls, R.D., Fischel_Ghodsian, N. & Higgs, D.R. (1987) Recombination at the human alpha-globin gene cluster: sequence features and topological constraints. *Cell*, **49**, 369–378.

Tamary, H., Klinger, G., Shalmon, L., Kirschmann, H., Koren, A., Bennet, M. & Zaizov, R. (1998) The diverse molecular basis and mild clinical picture of HbH disease in Israel. *Annals of the New York Academy of Sciences*, 850.

Tan, A.S., Quah, T.C., Low, P.S. & Chong, S.S. (2001) A rapid and reliable 7-deletion multiplex polymerase chain reaction assay for alpha-thalassemia. *Blood*, **98**, 250–251.

Vandenplas, S., Higgs, D.R., Nicholls, R.D., Bester, A.J. & Mathew, C.G. (1987) Characterization of a new alpha zero thalassaemia defect in the South African population. *British Journal of Haematology*, **66**, 539–542.

Weatherall, D.J. & Clegg, J.B. (2001) *The Thalassaemia Syndromes*. Blackwell Scientific Publications, Oxford, UK.