

## PHYSICAL MAPPING OF THE GLOBIN GENE DELETION IN $(\delta\beta)^\circ$ - THALASSAEMIA

(Foetal-adult gene switch; genomic blotting; filter hybridization; recombinant DNA; plasmid pCRI vector)

R. BERNARDS, J.M. KOOTER and R.A. FLAVELL

*Section for Medical Enzymology and Molecular Biology, Laboratory of Biochemistry, University of Amsterdam, Jan Swammerdam Institute, P.O. Box 60.000, 1005 GA Amsterdam (The Netherlands)*

(Received May 9, 1979)

(Accepted May 14, 1979)

---

### SUMMARY

We have constructed a physical map of restriction endonuclease cleavage sites in the  $(\delta+\beta)$ -globin gene region in the DNA of patients with  $(\delta\beta)^\circ$ -thalassaemia. This map shows that a 10 kb deletion has occurred in  $(\delta\beta)^\circ$ -thalassaemia to remove the entire  $\beta$ -globin gene and the 3' portion of the  $\delta$ -globin gene. The 5' terminus of the deletion is in the large intron of the  $\delta$ -globin gene and the 3' terminus 1.8 kb to the 3'-side of the  $\beta$ -globin gene.

A similar deletion of about 7 kb has been described previously in the DNA of patients with Hb Lepore; the 5' terminus of the deletion is also in the  $\delta$ -globin gene but the 3' terminus is in the  $\beta$ -globin gene. Comparison of the foetal ( $\gamma$ ) globin gene expression in adults with  $(\delta\beta)^\circ$ -thalassaemia and Hb Lepore suggests that the 3' extragenic regions of the  $\beta$ -globin gene contain DNA sequences involved in the regulation of  $\gamma$ -globin gene expression.

---

### INTRODUCTION

In adult humans, HbA ( $\alpha_2\beta_2$ ) accounts for more than 95% of the total haemoglobin. In addition about 2% HbA<sub>2</sub> ( $\alpha_2\delta_2$ ) and very low levels of the

**Abbreviations:** HPFH, Hereditary Persistence of Foetal Haemoglobin; kb, kilo-base pair(s); cDNA, complementary DNA; mRNA, messenger RNA; SSC, 150 mM NaCl–15 mM sodium citrate (pH 7.0).

Throughout this paper, when reference is made to the 5'- or 3'-side of a gene, this refers to the 5'- or 3'-side of the anti-sense strand (i.e. the sequence homologous to the mRNA). The term  $(\delta\beta)^\circ$ -thalassaemic globin gene is used to refer to the globin gene sequences in DNA from  $(\delta\beta)^\circ$ -thalassaemic patients which form hybrids stable to stringent post-hybridization washes with probes for the  $\delta$ - and  $\beta$ -globin genes.

foetal haemoglobin (HbF,  $\alpha_2\gamma_2$ ) are found. The latter is restricted to a minor fraction of the erythroid cells (about 1%) called F cells; these cells also contain HbA and HbA<sub>2</sub> (Weatherall and Clegg, 1972).

In homozygous  $(\delta\beta)^\circ$ -thalassaemia, a rare inherited anaemia, the  $\delta$ - and  $\beta$ -globin chains are entirely absent and only HbF is found. Previous studies (Ottolenghi et al., 1976; Ramirez et al., 1976) have shown that at least a partial deletion of the  $\delta$ - and  $\beta$ -globin genes has occurred in  $(\delta\beta)^\circ$ -thalassaemia since the plateau level of  $\beta$ -globin cDNA hybridization is reduced in DNA of patients with this disease. The precise extent of this deletion could not be determined from these experiments.

There are several reasons why it is of interest to map the deletion in the  $(\gamma + \delta + \beta)$ -globin gene region in  $(\delta\beta)^\circ$ -thalassaemia. This deletion results in the absence of  $(\beta + \delta)$ -globin mRNA in erythroid cells (Ottolenghi et al., 1975; Old et al., 1978); more important, even heterozygous  $(\delta\beta)^\circ$ -thalassaemics have elevated levels of HbF when compared with either normal humans or  $\beta^\circ$ - or  $\beta^+$ -thalassaemics (Weatherall and Clegg, 1972). In  $(\delta\beta)^\circ$ -thalassaemia the level of the  $\gamma$ -globin chains is lower than that of the  $\alpha$ -chains; this  $\alpha$ - $\gamma$ -chain imbalance results in the abnormal erythroid cells and anaemia found in the thalassaemia phenotype. A deletion of the  $\delta$ - and  $\beta$ -globin gene region has also occurred in another inherited disorder, called HPFH. In this case, however, the  $\gamma$ -globin chain levels almost fully compensate for the absent  $(\beta + \delta)$ -globin chains; HPFH is, therefore, an essentially non-clinical condition. Thus it appears that the switch from  $\gamma$ - to  $(\delta + \beta)$ -globin gene expression, that occurs around birth in normal humans, has not occurred at all in HPFH and only partially in  $(\delta\beta)^\circ$ -thalassaemia.

In the past few years methods have been developed which make it possible to construct physical maps of restriction enzyme cleavage sites in and around any gene for which a hybridization probe is available (Southern, 1975; Botchan et al., 1976; Jeffreys and Flavell, 1977). Recently we have used this approach to analyse the  $\beta$ -related globin gene locus in DNA from normal humans (Flavell et al., 1978; Little et al., 1979; Bernards et al., 1979). In this paper we describe the application of these methods to the mapping of the  $\beta$ - and  $\delta$ -globin genes in the DNA of patients with  $(\delta\beta)^\circ$ -thalassaemia.

In normal human DNA the  $\delta$ -globin gene is found 5.5 kb to the 5'-side of the  $\beta$ -globin gene (Fig.1) (see also Lawn et al., 1978). In  $(\delta\beta)^\circ$ -thalassaemia, however, a 10 kb deletion has occurred in this region to remove DNA extending from the large intron of the  $\delta$ -globin gene to a site 1.8 kb to the 3'-side of the  $\beta$ -globin gene. A summary of part of this work has been presented elsewhere (Flavell et al., 1979a, b).

## MATERIALS AND METHODS

### *Preparation of thalassaemic DNAs*

Lymphocyte cell lines of two sibling homozygous  $(\delta\beta)^\circ$ -thalassaemia patients were obtained from The Human Genetic Mutants Cell Repository



(Camden, NJ, USA: culture numbers 2266 and 2267, respectively). DNA was prepared as described for rabbit-liver DNA except that the 2-methoxyethanol step was omitted (Jeffreys and Flavell, 1977). Blood DNA from an Italian  $\beta^\circ / (\delta\beta)^\circ$  double heterozygote (Ottolenghi et al., 1975) and a Greek  $(\delta\beta)^\circ$  heterozygote was prepared in the same way.

#### *Detection of globin gene fragments by Southern blotting filter hybridizations*

Southern transfers and filter hybridizations were as described (Jeffreys and Flavell, 1977; Flavell et al., 1978) with the following modifications. Hybridization of filters with  $^{32}\text{P}$ -labelled pH $\beta$ G1 DNA was for 16 h at 65°C. Post-hybridization washes were: (a) six washes, each of 5 min, with hybridization solution (without plasmid probe DNA and poly(A)), and (b) sequential washes (each  $2 \times 15$  min at 65°C) of  $1.0 \times \text{SSC}$ ,  $0.3 \times \text{SSC}$  and  $0.1 \times \text{SSC}$ . All SSC solutions contained 0.1% sodium dodecylsulphate and 50  $\mu\text{g}$  denatured salmon sperm DNA per ml. The markers used for molecular weight calibrations were 5'  $^{32}\text{P}$ -end labelled phage lambda DNA, phage lambda DNA + *Eco*RI fragments, phage  $\phi$ 29 DNA + *Eco*RI fragments and phage  $\phi$ X174 RF DNA + *Taq*I fragments.

The  $\beta$ - and  $\delta$ -globin genes were detected with  $^{32}\text{P}$ -labelled pH $\beta$ G1 DNA — a pCRI plasmid containing human  $\beta$ -globin cDNA — as described (Flavell et al., 1978). The  $\gamma$ -globin genes were detected with pH $\gamma$ G1 DNA, a human G $\gamma$  cDNA plasmid which hybridizes to the G $\gamma$  and A $\gamma$  genes (Little et al., 1979). For certain experiments a probe containing the  $\beta$ -globin cDNA regions to the 5'-side of the intragenic *Bam*HI site was prepared as described in Flavell et al. (1978). Briefly, this involves *Bam*HI + *Hind*III cleavage of pH $\beta$ G1 DNA and separation of the two fragments generated, followed by nick-translation in the presence of  $^{32}\text{P}$ -labelled deoxyribonucleoside triphosphates.

TABLE I

#### SIZE OF GLOBIN GENE FRAGMENTS IN RESTRICTION ENDONUCLEASE DIGESTS

The  $(\delta\beta)^\circ$  fragment sizes are from this paper. The  $\beta$ - and  $\delta$ -globin gene fragments have been described previously. The sizes for the larger fragments reported here are somewhat shorter than our published values (Flavell et al., 1978) owing to a re-calibration with an extensive series of markers. These new values agree well with those measured on cloned  $\delta$ - and  $\beta$ -globin genes by Lawn et al. (1978).

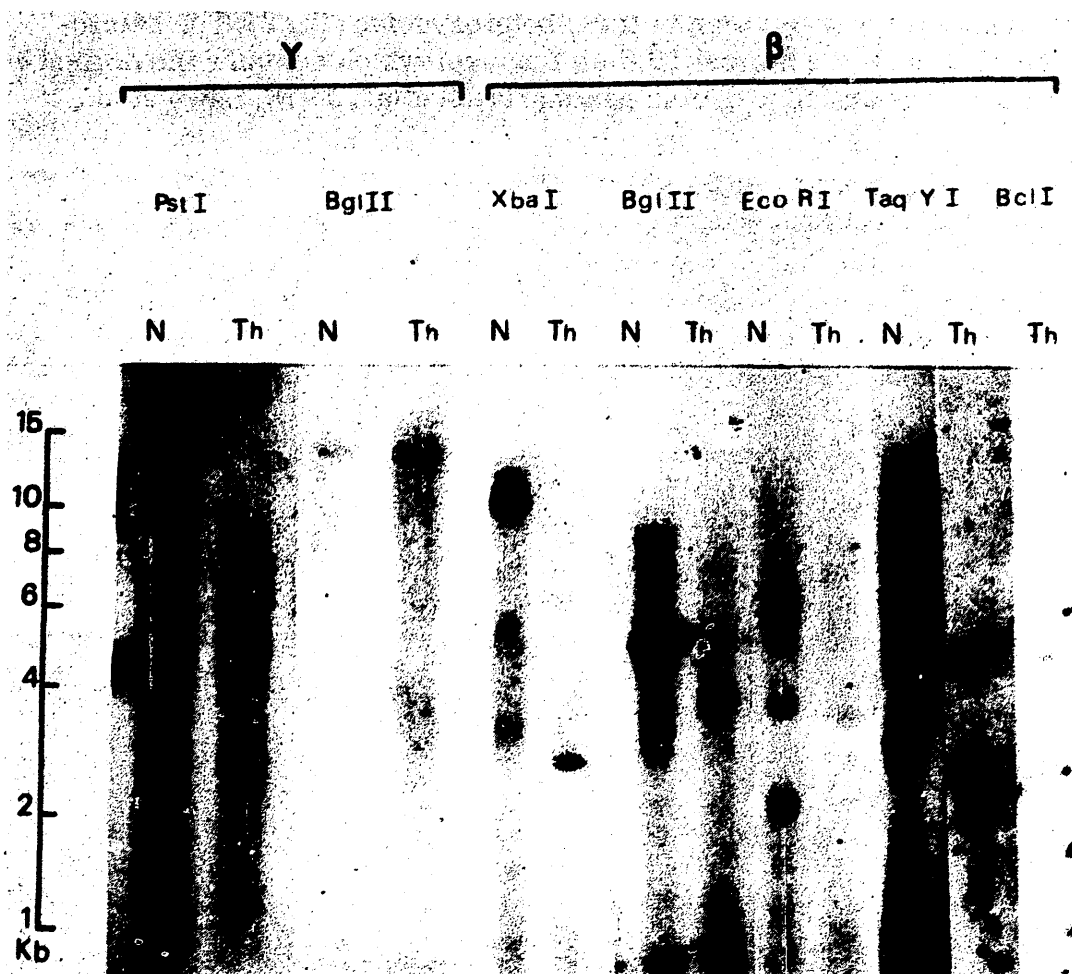
Enzyme	$(\delta\beta)^\circ$	$\beta$	$\delta$
<i>Xba</i> I	2.9	10.3	10.3
<i>Bgl</i> II	3.8	5.0	8.4
<i>Eco</i> RI	3.5	5.8 + 3.7	2.3 + 1.9 <sup>a</sup>
<i>Taq</i> YI	2.4	8.6	3.1
<i>Hind</i> II + III	1.6	5.7	1.4
<i>Bcl</i> I	4.2	9.0	4.7
<i>Bam</i> HI	15	1.8 + 9.0	15 + ? <sup>a</sup>

<sup>a</sup>Fragments containing only the 3' regions of the  $\delta$ -globin gene are poorly detected by our probes (Flavell et al., 1978).

## RESULTS

*Abnormal globin gene DNA fragments in  $(\delta\beta)^\circ$ -thalassaemia*

DNA from cultured lymphocytes derived from a homozygous  $(\delta\beta)^\circ$ -thalassaemic was cleaved with a variety of restriction endonucleases and analysed by Southern (1975) blotting and filter hybridization to  $^{32}\text{P}$ -labelled H $\beta$ G1 DNA. In normal human DNA both the  $\delta$ - and  $\beta$ -globin genes are detected by this probe because of the extensive sequence homology between the  $\delta$ - and  $\beta$ -globin genes. In DNA from the homozygous  $(\delta\beta)^\circ$ -thalassaemic cell-line, fragments were observed which are not found in normal human DNA. Thus, normal human DNA digested with restriction endonuclease *Xba*I gives a 10.3 kb fragment which contains the human  $\delta$ - and  $\beta$ -globin genes; in  $(\delta\beta)^\circ$ -thal-



**Fig. 2.** The  $\gamma$ -,  $\beta$ - and  $\delta$ -globin gene fragments in normal DNA and DNA of patients with  $(\delta\beta)^\circ$ -thalassaemia. DNA prepared from normal (N) placenta DNA or from cultured lymphocytes from a patient with homozygous  $(\delta\beta)^\circ$ -thalassaemia (Th) was digested with the restriction endonuclease indicated, denatured with alkali, applied to a 1.2% agarose gel, electrophoresed and transferred to a nitrocellulose filter as described in METHODS. The  $(\beta + \delta)$ -globin gene fragments were then detected by filter hybridization to  $^{32}\text{P}$ -labelled pH $\beta$ G1 DNA (Flavell et al., 1978). The  $\gamma$ -globin gene fragments were detected using  $^{32}\text{P}$ -labelled pH $\gamma$ G1 DNA (Little et al., 1979).

assaemic DNA a novel DNA fragment of 2.9 kb is found instead. Similarly, in normal DNA *Bgl*III gives DNA fragments of 8.4 kb and 5.0 kb containing the  $\delta$ - and  $\beta$ -globin genes, respectively; in  $(\delta\beta)^\circ$ -thalassaemia DNA only a single fragment of 3.8 kb is found (Table I, Fig.2). The same type of result was found with a variety of other restriction enzymes (Table I, Fig.2).

The abnormal pattern described above was first observed in a transformed cell line derived from a patient with homozygous  $(\delta\beta)^\circ$ -thalassaemia. It could be argued that the fragment pattern that we observe is the result of a cell culture artefact. To exclude this we have examined the  $\beta$ - and  $\delta$ -globin gene fragments in DNA derived from peripheral blood from two additional  $(\delta\beta)^\circ$  cases — a Greek heterozygote and an Italian  $\beta^\circ/(\delta\beta)^\circ$ -thalassaemic, double heterozygote (the latter case has been analysed previously by Ottolenghi et al. (1975)). The 2.9 kb *Xba*I band, which is present in the homozygous  $(\delta\beta)^\circ$ -thalassaemia DNA is also seen in these heterozygotes (Fig.3). The heterozygotes also show the normal 10.3 kb fragment which contains the normal  $\delta$ - and  $\beta$ -, or  $\beta^\circ$ -thalassaemic, globin genes (Flavell et al., 1978; Lawn et al.,

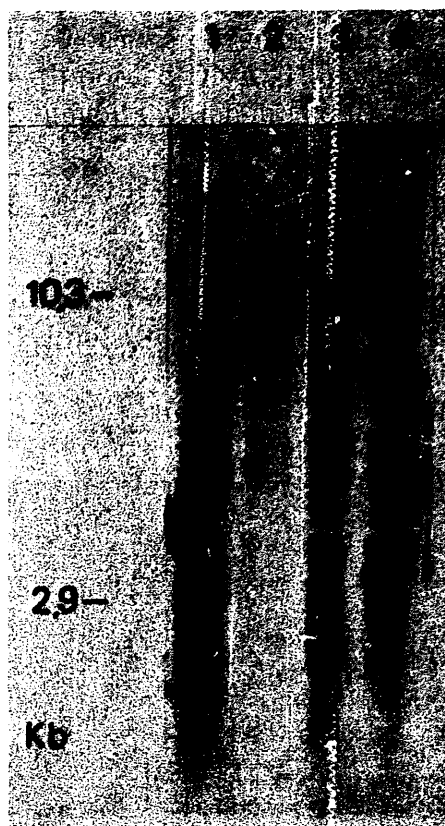


Fig.3. The  $(\delta + \beta)$ -globin gene fragments in different patients with  $(\delta\beta)^\circ$ -thalassaemia. DNA from a control human placental DNA, a lymphocyte cell line of a homozygous  $(\delta\beta)^\circ$ -thalassaemic patient and from peripheral blood from a  $(\delta\beta)^\circ$  heterozygote and a  $(\delta\beta)^\circ/\beta^\circ$  double heterozygote was cleaved with *Xba*I and analysed for the  $(\beta + \delta)$ -globin gene fragments as described in Methods and Fig.2.

- (1) The homozygous  $(\delta\beta)^\circ$ -thalassaemic DNA. (2) Normal placental DNA.  
 (3) Heterozygous  $(\delta\beta)^\circ$ -thalassaemic DNA. (4)  $(\delta\beta)^\circ/\beta^\circ$  double heterozygote DNA.

TABLE II

SIZE OF ( $\delta + \beta$ )-GLOBIN GENE FRAGMENTS IN DOUBLE DIGESTS OF ( $\delta\beta$ )<sup>o</sup>-THALASSAEMIC DNA

Enzyme	Size (kb)
<i>Pvu</i> II + <i>Eco</i> RI	1.3
+ <i>Xba</i> I	0.8
<i>Bq</i> II + <i>Hind</i> II/III	1.3
+ <i>Taq</i> I	1.3
+ <i>Xba</i> I	1.65
+ <i>Pst</i> I	1.7
+ <i>Eco</i> RI	2.2
<i>Bam</i> HI + <i>Hind</i> II/III	0.5
+ <i>Taq</i> I	0.5
+ <i>Xba</i> I	0.85
+ <i>Pst</i> I	0.95
+ <i>Eco</i> RI	1.3
+ <i>Bcl</i> II	2.1
+ <i>Bgl</i> II	3.0
<i>Pst</i> I + <i>Kpn</i> I	3.1
+ <i>Hind</i> III	3.35

1978). Finally, we have observed the same pattern of abnormal globin gene fragments in cultured cells from a second homozygous ( $\delta\beta$ )<sup>o</sup>-thalassaemic patient (a sibling of the first case; not shown).

Despite the fact that abnormalities are seen in the ( $\delta + \beta$ )-globin gene fragments in ( $\delta\beta$ )<sup>o</sup>-thalassaemia, the  $\gamma$ -globin gene fragments appear to be the same as those found in normal DNA. The *Bgl*II fragments and the *Pst*I fragments, which contain the G $\gamma$ - and A $\gamma$ -globin genes (Little et al., 1979) are present in DNA from ( $\delta\beta$ )<sup>o</sup>-thalassaemics (Fig.2). The 15 kb *Bam*HI fragment which links the  $\gamma$ - and  $\delta$ -globin genes (Bernards et al., 1979) is also present in ( $\delta\beta$ )<sup>o</sup>-thalassaemic DNA (Table I).

#### *Mapping the deletion in ( $\delta\beta$ )<sup>o</sup>-thalassaemia: the 5' break point*

The abnormal ( $\beta + \delta$ )-globin gene fragments observed in ( $\delta\beta$ )<sup>o</sup>-thalassaemia is the type of result expected for the partial deletion of these globin genes in this disease. To determine the extent of the deletion and to identify which globin genes remain, we performed double digests of ( $\delta\beta$ )<sup>o</sup>-thalassaemic DNA with a series of restriction endonucleases and *Bam*HI; *Bam*HI cleaves the  $\delta$ - and  $\beta$ -globin genes at the position coding for amino acids 98–100 of the globin chain (Fig.3, Table II). The *Bam*HI double digest fragments obtained from the ( $\delta\beta$ )<sup>o</sup>-thalassaemic gene are identical in size with the corresponding fragments containing the 5' regions of the  $\delta$ -globin gene. These data predict the sequence and relative distance of cleavage sites characteristic of the 5' ex-tragenic regions of the  $\delta$ -globin gene: 5'-*Bam*HI-*Bgl*II-*Bcl*II-*Eco*RI-*Pst*I-*Xba*I-*Hind*II + III and *Taq*I-3'.

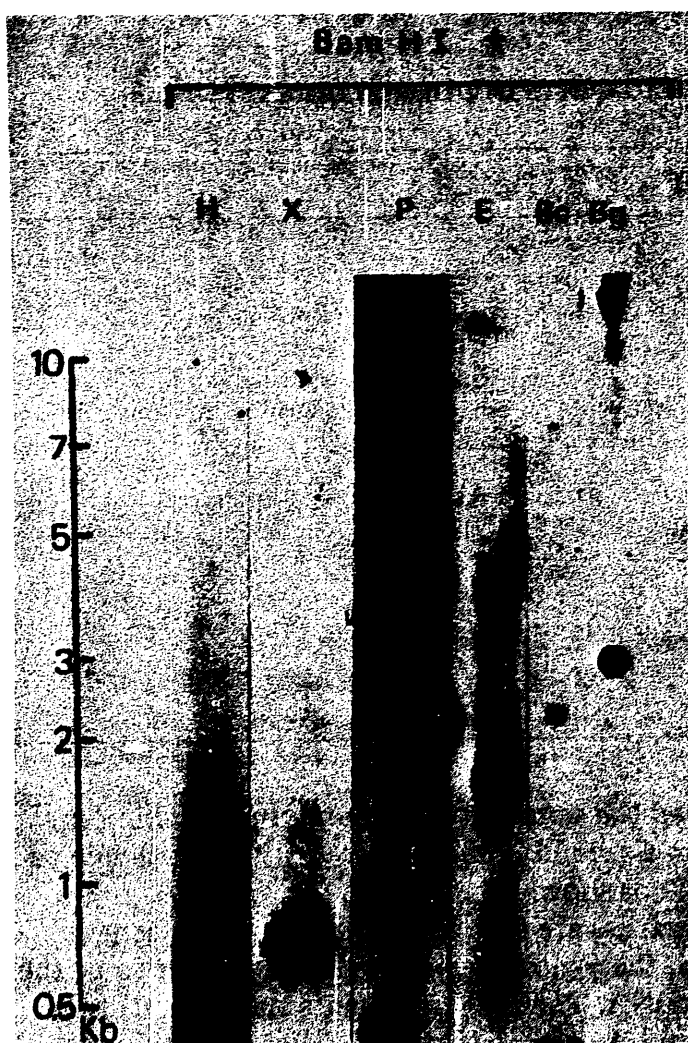


Fig.4. The  $(\delta\beta)^{\circ}$ -globin gene fragments in double digests of  $(\delta\beta)^{\circ}$ -thalassaemic DNA with *Bam*HI and different restriction endonucleases. DNA was digested with *Bam*HI and the restriction enzymes indicated and analysed for the  $(\delta + \beta)$ -globin gene fragments as in Fig.2 and Methods. The restriction endonucleases used were: *Bam*HI + *Hind*II/III (H), *Xba*I (X), *Pst*I (P), *Eco*RI (E), *Bcl*II (Bc) and *Egl*II (Bg). The *Bam*HI + *Bcl*II and *Bam*HI + *Egl*II digests were hybridized with a probe for the regions to the 5'-side of the intragenic *Bam*HI site of the  $\delta$ - and  $\beta$ -globin genes.

Two further pieces of evidence show that the region conserved in  $(\delta\beta)^{\circ}$ -thalassaemia is indeed the 5' region of the  $\delta$ -globin gene. First, the 2.1 kb *Bam*HI + *Bcl*II and 3.0 kb *Bam*HI + *Egl*II (Fig.3) double digest fragments are detected by hybridization probes for the coding regions to the 5'-side of the intragenic *Bam*HI site in the  $\beta$ - and  $\delta$ -globin genes. Second, the *Pvu*II site, which we have previously shown to be present in the 5' segments of the  $\delta$ -globin gene (Flavell et al., 1978) and absent from the  $\beta$ -globin gene, is also present in the same position in  $(\delta\beta)^{\circ}$ -thalassaemia. Thus, digests of *Eco*RI or *Xba*I with *Pvu*II give double digest fragments of 1.3 kb and 0.8 kb, respec-



tively (Table II, Fig.5). This places a *Pvu*II site just 5' from the intragenic *Bam*HI site in the  $(\delta\beta)^{\circ}$ -thalassaemic globin gene.

The restriction endonuclease cleavage sites to the 3'-side of the intragenic *Bam*HI site have been deduced from the position of the 5' sites discussed above and from the size of the respective fragments in single digests (see Tables I and II). For example, *Bgl*II gives a 3.8 kb fragment which is cleaved by *Bam*HI to yield a 3 kb double digest fragment. The 3' *Bgl*II site is, therefore, 0.8 kb to the 3'-side of the intragenic *Bam*HI site. Furthermore, in a number of cases these positions have been verified in additional double digests. Thus, the 3' *Bgl*II site has been verified independently in double digests with *Hind*II + III, *Taq*I, *Xba*I, *Pst*I and *Eco*RI (Table II, Fig.5).

The restriction map deduced from the fragments described above for  $(\delta\beta)^{\circ}$ -thalassaemia is identical to the 5' regions of the  $\delta$ -globin gene in normal DNA

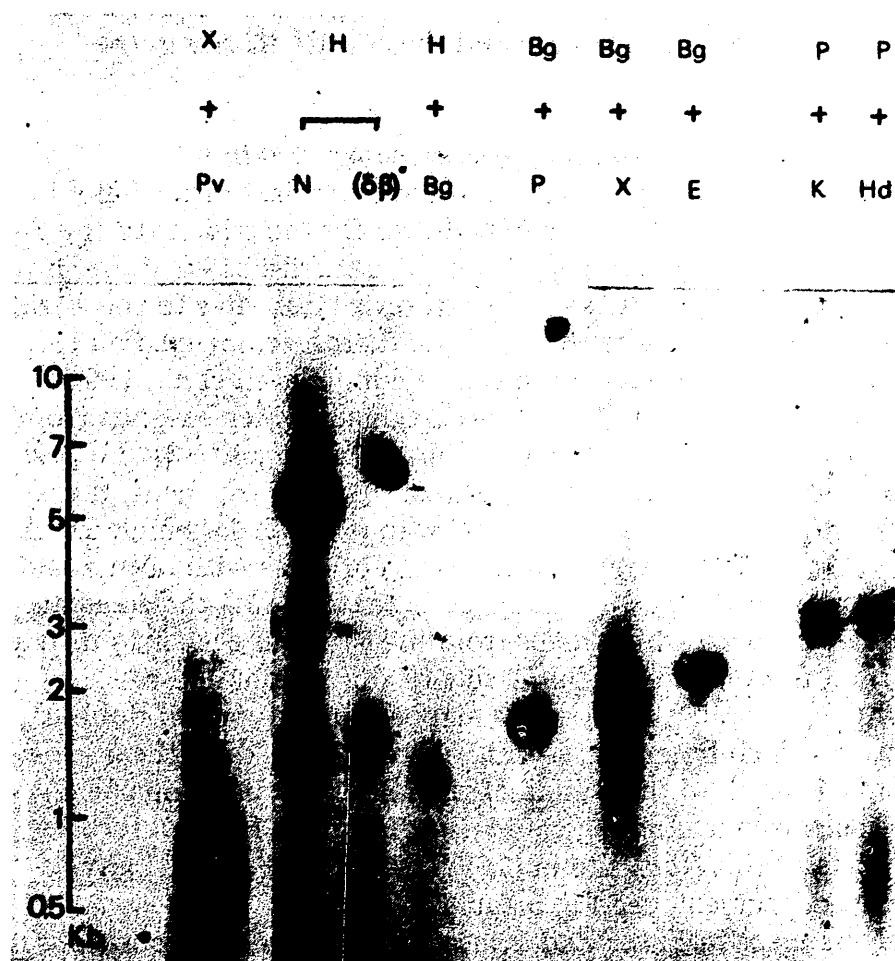


Fig.5. Globin gene fragments from the  $(\delta\beta)^{\circ}$ -thalassaemic globin gene in double digests. DNA from the  $(\delta\beta)^{\circ}$ -thalassaemia cell line was digested with the restriction enzymes indicated and analysed for the  $(\delta + \beta)$ -globin gene fragments as described in Fig.2 and Methoda. The *Hind*II/III digests of normal (N) and  $(\delta\beta)^{\circ}$ -thalassaemic DNA are also shown. The enzymes used were: *Xba*I (X), *Pvu*II (Pv), *Hind*II/III (H), *Bgl*II (Bg), *Pst*I (P), *Eco*RI (E), *Kpn*I (K), *Hind*III (Hd).

up to and including the intragenic *Bam*HI site. The DNA region to the 3'-side of this *Bam*HI site in  $(\delta\beta)^\circ$ -thalassaemic DNA is not present at this position in normal DNA. Thus, the normal  $\delta$ -globin gene contains a *Hind*II/III site close to the 3'-end of the large intron. This is deduced from the 1.4 kb *Hind*II/III  $\delta$ -globin fragment (Fig.5), which is cleaved by *Bam*HI to give a 0.5 kb double digest fragment; the *Hind*II/III fragment is not cleaved by *Eco*RI. These data place the *Hind*II/III site in the  $\delta$ -globin large intron just before the 3'-end. In contrast,  $(\delta\beta)^\circ$ -thalassaemic DNA shows an abnormal 1.6 kb *Hind*II/III fragment which is cleaved by *Bgl*II to give a 1.3 kb double digest fragment (Fig.5); the *Bgl*II site is, therefore, 0.3 kb to the 5'-side of the *Hind*II + III site in  $(\delta\beta)^\circ$ -thalassaemic DNA. These data — and the fact that none of the restriction sites to the 3'-side of the “ $(\delta\beta)^\circ$ -thalassaemia gene” region to be discussed below are found at the 3'-side of the  $\delta$ -globin gene — show that the 5' terminus of the deletion in  $(\delta\beta)^\circ$ -thalassaemia is between the intragenic *Bam*HI site (at the end of the second  $\delta$ -globin gene exon) and the *Hind*II/III site in the  $\delta$ -globin gene large intron.

#### *Mapping of the 3' break point of the $(\delta\beta)^\circ$ -thalassaemia deletion*

In none of the  $(\delta\beta)^\circ$  DNA samples were any of the characteristic  $\beta$ -globin gene fragments visible. The restriction map deduced for the regions to the 3'-side of the  $(\delta\beta)^\circ$ -thalassaemia globin gene from the data discussed above, however, appeared to us to resemble the sequence of restriction sites to the 3'-side of the  $\beta$ -globin gene. In order to establish this we, therefore, mapped all the known restriction sites which are found in this region (Flavell et al., 1978).

Normal human DNA contains a 46 kb *Kpn*I fragment which extends from a site approx. 14 kb to the 3'-side of the  $\gamma$ -globin gene to a site 3.4 kb past the 3'-end of the  $\beta$ -globin gene (Bernards et al., 1979). This fragment, therefore, contains the  $\gamma$ -,  $A\gamma$ -,  $\delta$ - and  $\beta$ -globin genes. In  $(\delta\beta)^\circ$ -thalassaemic DNA a novel 36 kb *Kpn*I fragment is found which also hybridizes with both  $\gamma$  and  $(\delta + \beta)$  hybridization probes (Fig.6). This suggests that in  $(\delta\beta)^\circ$ -thalassaemia a 10 kb deletion has occurred which maps from the  $\delta$ -globin gene intron to a position just before the *Kpn*I site on the 3'-side of the  $\delta$ -globin gene.

In support of this proposal, a *Kpn*I site is mapped in *Pst*I + *Kpn*I double digests (which give a 3.1 kb fragment; see Fig.5) to be 0.25 kb to the 3'-side of the *Taq*I site and 1.35 kb to the 3'-side of the *Bgl*II site found on the 3'-side of the  $(\delta\beta)^\circ$ -thalassaemic globin gene. A *Hind*III site is found a further 0.25 kb to the 3'-side of this *Kpn*I site in  $(\delta\beta)^\circ$  DNA (*Pst*I + *Hind*III gives a 3.35 kb double digest fragment; see Fig.5 and Table II).

Comparison of the maps in Figs.7 and 1 shows that the relative positions of the 3' extragenic *Bgl*II, *Taq*I, *Kpn*I, *Eco*RI and *Hind*III sites of the  $\beta$ -globin gene and the  $(\delta\beta)^\circ$ -thalassaemic globin gene are the same (the distances relative to the 3' extragenic *Bgl*II site are for normal DNA: *Taq*I, 1.1 kb; *Kpn*I, 1.35 kb; *Eco*RI, 1.5 kb; *Hind*III, 1.6 kb. For  $(\delta\beta)^\circ$ -thalassaemic DNA the corresponding values are: *Taq*I, 1.1 kb; *Kpn*I, 1.35 kb; *Eco*RI, 1.4 kb; *Hind*III, 1.6 kb). Two apparent differences are found. First, an *Xba*I site is present in



**Fig.6.** The ' $\gamma\delta\beta$ ' *Kpn*I fragment in normal DNA, Hb Lepore DNA and  $(\delta\beta)^\circ$ -thalassaemic DNA. DNA from normal placental DNA (N), from a patient with homozygous Hb Lepore (Le) and the cell line from the homozygous  $(\delta\beta)^\circ$ -thalassaemia patient was digested to completion with *Kpn*I, electrophoresed in double-stranded form on a 0.5% agarose gel for 48 h (1 V/cm), transferred to a nitrocellulose filter as described by Southern (1975) and hybridized with  $^{32}\text{P}$ -labelled pH $\gamma$ G1 DNA to detect the  $\gamma$ -globin genes. Normal DNA shows a fragment of about 46 kb which contains the G $\gamma$ -, A $\gamma$ -,  $\delta$ - and  $\beta$ -globin genes (Bernards et al., 1979). Hb Lepore DNA gives a fragment of approx. 39 kb. This is 7 kb shorter than the corresponding normal fragment due to the deletion of the 7 kb region consisting of the 3' regions of the  $\delta$ -globin gene, the 5' regions of the  $\beta$ -globin gene and the  $\delta$ - $\beta$  intergenic DNA.

$(\delta\beta)^\circ$  DNA between the *Taq*I and *Kpn*I sites. This site cannot be detected in normal DNA because another *Xba*I site is present, closer to the  $\beta$ -globin gene, which is deleted in  $(\delta\beta)^\circ$ -thalassaemia. In limit digests with a given restriction enzyme, only those sites closest to the gene are detected. This is, therefore, not a discrepancy. Second, a *Hind*II/III site (which is in fact a *Hpa*I site (unpublished data)) is present between the 3' extragenic *Bgl*II and *Taq*I sites. This site is not present in any normal DNA examined up till now. It is not unreasonable to assume that this *Hind*II/III-*Hpa*I site is the result of a polymorphism which was present on the original chromosome in which the  $(\delta\beta)^\circ$ -tha-

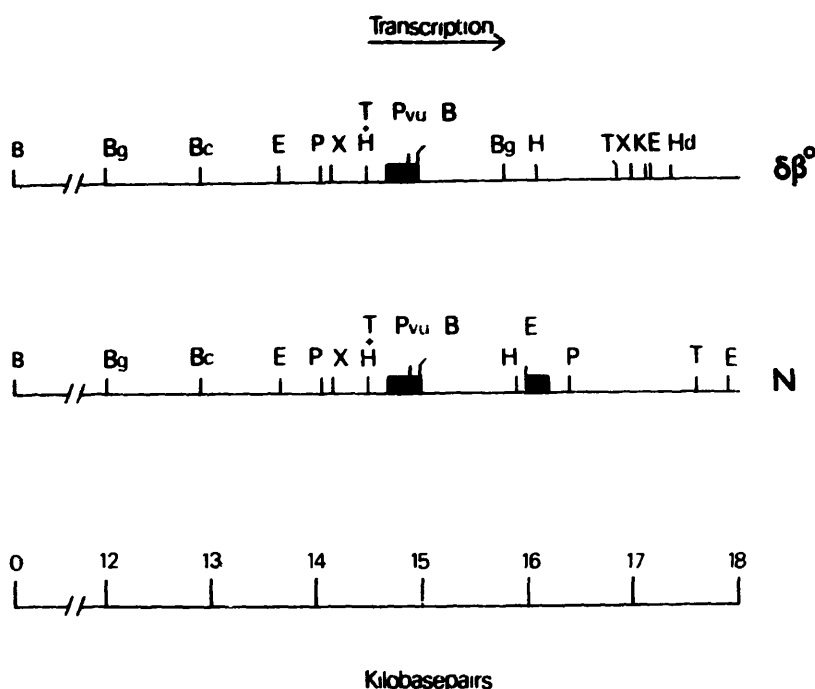


Fig. 7. A physical map of the  $(\delta\beta)^\circ$ -thalassaemic globin gene compared with the map of the normal (N) human  $\delta$ -globin gene. The general comments in the legend to Fig. 1 apply here also. The map of the normal  $\delta$ -globin gene (N) is from Flavell et al. (1978) with minor modifications because of some revised fragment sizes (see Table I). The enzymes used were: *Bam*HI (B), *Bcl*II (Bc), *Bgl*II (Bg), *Eco*RI (E), *Hind*II/III (H), *Hind*III (Hd), *Kpn*I (K), *Pst*I (P), *Pvu*II (Pvu), *Taq*I (T) and *Xba*I (X) (see Roberts, 1978).

lassaemic lesion occurred. Such a polymorphism would remain fixed in the DNA of individuals with  $(\delta\beta)^\circ$ -thalassaemia even though it might be uncommon in other populations. An *Hpa*I polymorphism has recently been shown to be linked to the sickle-cell  $\beta$ -globin gene (Kan and Dozy, 1978). We are currently screening individuals in the affected areas (Southern Italy and Greece) for *Hpa*I polymorphisms.

We conclude that the 3' terminus in  $(\delta\beta)^\circ$ -thalassaemia is between the *Xba*I site and the *Bgl*II site (which are separated by only 150 bp in normal DNA) approx. 1.8 kb to the 3'-side of the  $\beta$ -globin gene.

## DISCUSSION

Our results show that an approx. 10 kb deletion of DNA in the  $(\gamma + \delta + \beta)$ -globin locus has occurred in  $(\delta\beta)^\circ$ -thalassaemia. The  $\gamma$  region and the  $\delta$ -globin gene up to the intragenic *Bam*HI site is present; the region extending from the  $\delta$ -globin intron to a site 1.8 kb past the  $\beta$ -globin gene is deleted. More complex interpretations of our data are also possible. First, if a region of the  $\beta$ -globin gene shorter than 50 bp were present, it would be difficult to detect with our probes. Second, the 5' and 3' extragenic regions that we identify flanking the

$(\delta\beta)^{\circ}$ -thalassaemic globin gene could be derived from DNA regions, located elsewhere in the normal human genome, with a sequence of restriction endonuclease cleavage sites identical to the bona fide 5'- $\delta$  and 3'- $\beta$  extragenic regions. Since the number of sites identified is 8 and 5, respectively, we consider this unlikely (although a reiteration of these regions in the human genome has not been excluded). Finally, more complex models, such as multiple deletions, could be made to fit our data. The 5' terminus of deletion in  $(\delta\beta)^{\circ}$ -thalassaemia has been determined independently by Ottolenghi et al. (1979) and by T. Maniatis and his colleagues (personal communication).

Is it possible to rationalise the  $(\delta\beta)^{\circ}$ -thalassaemia phenotype with the structure of the  $\delta\beta$ -locus in  $(\delta\beta)^{\circ}$ -thalassaemia? The fact that the  $\beta$ -globin gene is entirely deleted, together with the 3' terminal exon of the  $\delta$ -globin chain (which codes for amino acids 105 to the end of the  $\delta$ -globin chain), of course explains the absence of  $\delta$ - and  $\beta$ -globin chains and the absence of  $\beta$ -globin mRNA in  $(\delta\beta)^{\circ}$ -thalassaemia patients. It has also been reported that  $\delta$ -globin mRNA is absent from erythroid cells of  $(\delta\beta)^{\circ}$ -thalassaemics. The failure to detect  $\delta$ -globin mRNA could either result from a lack of transcription of the  $\delta$ -globin gene or in a decrease in the stability of the  $\delta$ -globin mRNA. Since the 3' regions of the  $\delta$ -globin gene have been deleted, we would expect transcription to go unhindered on the basis of bacterial transcription models. The mode of transcription of eukaryotic genes is, however, entirely unknown and it would be premature to rule out a priori a role in transcription for the regions on the 3'-side of a gene (for example, a role in the organization of a transcriptionally active region of chromatin).

More plausible is the expectation that transcripts of the " $\delta$ -half gene" in  $(\delta\beta)^{\circ}$ -thalassaemia would be less stable than their normal counterparts. The 3' untranslated region presumed to be involved in polyadenylation is deleted, as is the 3' splicing acceptor sequence of the large intron. We might, therefore, expect the  $\delta$  mRNA to remain an "unprocessable" pre-mRNA which could be degraded in the nucleus. Recently SV40 deletion mutants, lacking a single intron-exon junction sequence, have been constructed. In this case, mRNA became unstable which supports the above interpretation (Lai and Khoury, 1979). We have also recently described a case of  $\beta^{\circ}$ -thalassaemia where again a deletion of the 3' regions of the gene has occurred. In this case, the 3' terminal exon of the  $\beta$ -globin gene has been deleted (see Flavell et al., 1979a). Interestingly, no transcripts of this globin gene have been detected in this patient either (Tolstoshev, P., personal communication).

The most intriguing — and also the most puzzling — aspect of  $(\delta\beta)^{\circ}$ -thalassaemia is the elevation of HbF, which invariably accompanies this disease.  $(\delta\beta)^{\circ}$ -Thalassaemia is clinically less severe than  $\beta^{\circ}$ -thalassaemia, yet more severe than the — essentially non-pathological — condition of HPFH where balanced synthesis of  $\gamma$ - and  $\alpha$ -chains occurs. The most trivial explanation for this could be that the  $(\delta\beta)^{\circ}$ -thalassaemics are double mutants, the first mutation would be the deletion of the  $\delta$ - and  $\beta$ -globin genes, the second mutation would be in a "gene" affecting HbF levels in an adult background. In this case we might expect these two mutations to segregate: in the family studies performed to date

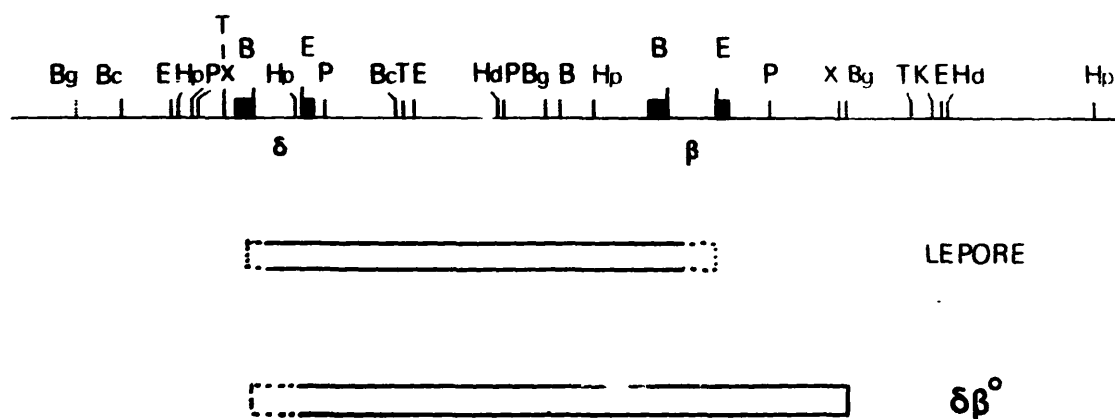


Fig. 8. Schematic map of the ( $\delta + \beta$ )-globin gene region in normal DNA (top drawing), Hb Lepore and ( $\delta\beta$ ) $^\circ$ -thalassaemic DNA. The general comments mentioned in the legend to Fig. 1 apply here also. The deletions (represented by rectangles) mapped in Hb Lepore are from Flavell et al. (1978) and in ( $\delta\beta$ ) $^\circ$ -thalassaemia from this paper. The dotted lines imply uncertainties in the extent of the deletions. The exact cross-over point which gives rise to the Hb Lepore gene has not been determined; the uncertainties of the location of the break points in ( $\delta\beta$ ) $^\circ$ -thalassaemia are described in the text.

this has not been seen. The second possibility is that the minor population of cells (about 1%) containing HbF in normal individuals is strongly selected for in the bone marrow of ( $\delta\beta$ ) $^\circ$ -thalassaemics because of the erythropoietic stress inflicted by the thalassaemia. This might then yield the 30–40% level of F cells found in ( $\delta\beta$ ) $^\circ$ -thalassaemia heterozygotes. This we consider unlikely. The HbF elevation does not occur in  $\beta^\circ$ - or  $\beta^+$ -thalassaemia heterozygotes where the stress is equally great.

The third and more attractive possibility is that the ( $\delta + \beta$ )-gene deletion is itself the causal agent in the elevation of HbF. Fig. 8 compares the structure of the two known deletion mutants in the ( $\delta + \beta$ )-globin region. In Hb Lepore, a 7 kb deletion has occurred with its 5' terminus in the  $\delta$ -gene and the 3' terminus in the  $\beta$ -gene (Flavell et al., 1978; Mears et al., 1978a). The result is a single, active fused  $\delta$ - $\beta$ -globin gene; the HbF level is only slightly elevated in Hb Lepore patients (1–5% in heterozygotes). In ( $\delta\beta$ ) $^\circ$ -thalassaemia the same deletion extends a further 2–3 kb in the 3' direction to remove the 3' regions of the  $\beta$ -globin gene. This results in a significant elevation of HbF (5–20% HbF in heterozygotes). In HPFH the deletion encompasses the entire  $\delta\beta$  region since no  $\delta$ - or  $\beta$ -coding sequences are present (Mears et al., 1978b; Orkin et al., 1978; Maniatis, T., personal communication; our unpublished results). This results in a still greater elevation of HbF (15–30% HbF in heterozygotes).

It appears, therefore, that both the 5' and 3' regions of the ( $\delta + \beta$ )-globin genes exert an influence on the expression of the  $\gamma$ -globin genes found upstream. Since both ( $\delta\beta$ ) $^\circ$  and HPFH heterozygotes still express elevated HbF levels and since the heterozygote HbF levels in these cases are lower than those of the corresponding homozygotes, it seems reasonable to assume that this is a *cis*-acting effect. The causal agent in suppressing  $\gamma$ -gene expression does not,

therefore, seem likely to be a diffusable product from this region, e.g. such that it is produced essentially normally in Hb Lepore at a low level in  $(\delta\beta)^\circ$ -thalassaemia and absent in HPFH.

One surprising aspect is that some of the DNA sequences which affect  $\gamma$ -globin gene expression are located at a great distance on the 3'-side of the  $\gamma$ -globin genes. In prokaryotes, all regulatory elements have been found on the 5'-side of the gene in question. There are several possible ways in which a 3' distal DNA sequence might regulate the expression of a gene. For example, these sequences might influence the organization of a region of a chromosome into an active configuration (cf. puffing in insects and the loops in lamp-brush chromosomes). If this is the case, then it is clear that several sites around the  $(\delta + \beta)$ -globin gene locus would affect this process since deletions between  $\delta$  and  $\beta$  (Hb Lepore), to the 3'-side of  $\beta$  ( $(\delta\beta)^\circ$ -thalassaemia) and to the 5'-side of  $\delta$  (HPFH) influence this. Also, the fact that the HbF levels seem to vary more or less continuously from a low but significant value in Hb Lepore to a high value in HPFH might suggest that the regulatory system consists of an equilibrium between two states, either active  $\gamma$  or active  $(\delta + \beta)$ ; this equilibrium might be influenced by deletions in the  $(\delta + \beta)$ -globin region.

#### ACKNOWLEDGEMENTS

We thank Peter F.R. Little for gifts of the two heterozygous thalassaemic DNAs and discussions, Piet Borst, Henrik H.M. Dahl, Frank G. Grosveld and Bob Williamson for discussions. We are especially grateful to Brigit Schrijver and Ernie De Boer for growing the lymphocyte cell lines, Wietske Oerlemans for art work and Dr. A. Bank for informing us of the availability of these lines. This work was supported in part by a grant to R.A.F. from The Netherlands Foundation for Chemical Research (SON) with financial aid from The Netherlands Organization for the Advancement of Pure Research (ZWO), and also by a NATO Travel grant.

#### REFERENCES

- Bernards, R., Little, P.F.R., Annison, G., Williamson, R. and Flavell, R.A., Structure of the human  $G\gamma$ - $A\gamma$ - $\delta$ - $\beta$ -globin gene locus, *Proc. Natl. Acad. Sci. USA*, 76 (1979) in press.
- Botchan, M., Topp, W. and Sambrook, J., The arrangement of Simian Virus 40 sequences in the DNA of transformed cells, *Cell*, 9 (1976) 269-287.
- Flavell, R.A., Kooter, J.M., De Boer, E., Little, P.F.R. and Williamson, R., Analysis of the  $\beta$ - $\delta$ -globin gene loci in normal and Hb Lepore DNA: Direct determination of gene linkage and intergene distance, *Cell*, 15 (1978) 25-41.
- Flavell, R.A., Grosveld, G.C., Grosveld, F.G., Bernards, R., Kooter, J.M., De Boer, E. and Little, P.F.R., The structure and expression of normal and abnormal globin genes, in 11th Miami Winter Symp., *From Gene to Protein: Information Transfer in Normal and Abnormal Cells*, 1979a, in press.
- Flavell, R.A., Bernards, R., Grosveld, G.C., Hoeijmakers-Van Dommelen, H.A.M., Kooter, J.M., De Boer, E. and Little, P.F.R., The structure and expression of globin genes in rabbit and man, in Axel, R. and Maniatis, T. (Eds.), *Eucaryotic Gene Regulation: ICN-UCLA Symposia on Molecular and Cellular Biology*, Academic Press, New York, 1979b, in press.

- Jeffreys, A.J. and Flavell, R.A., A physical map of the DNA regions flanking the rabbit  $\beta$ -globin gene, *Cell*, 12 (1977) 429–439.
- Kan, Y.W. and Dozy, A.M., Polymorphism of DNA sequence adjacent to the human  $\beta$  structural gene: Relationship to sickle cell mutation, *Proc. Natl. Acad. Sci. USA*, 75 (1978) 5631–5635.
- Lai, C.J. and Khoury, G., Deletion mutants of Simian Virus 40 defective in biosynthesis of late viral mRNA, *Proc. Natl. Acad. Sci. USA*, 76 (1979) 71–75.
- Lawn, R.M., Fritsch, E.F., Parker, R.C., Blake, G. and Maniatis, T., The isolation and characterization of linked  $\delta$ - and  $\beta$ -globin genes from a cloned library of human DNA, *Cell*, 15 (1978) 1157–1174.
- Little, P.F.R., Flavell, R.A., Kooter, J.M., Annison, G. and Williamson, R., Structure of the human foetal globin gene locus, *Nature*, 278 (1979) 227–231.
- Mears, J.G., Ramirez, F., Leibowitz, D. and Bank, A., Organization of human  $\delta$ - and  $\beta$ -globin genes in cellular DNA and the presence of intragenic inserts, *Cell*, 15 (1978a) 15–23.
- Mears, J.G., Ramirez, F., Leibowitz, D., Nakamura, F., Bloom, A., Konotey-Ahulu, F. and Bank, A., Changes in restricted human cellular DNA fragments containing globin gene sequences in thalassaemias and related disorders, *Proc. Natl. Acad. Sci. USA*, 75 (1978b) 1222–1226.
- Old, J.M., Proudfoot, N.J., Wood, W.G., Longley, J.I., Clegg, J.B. and Weatherall, D.J., Characterization of  $\beta$ -globin mRNA in the  $\beta^0$ -thalassaemias, *Cell*, 14 (1978) 289–298.
- Orkin, S.H., Alter, B.P., Altay, C., Mahoney, M.J., Lazarus, H., Hobbins, J.C. and Nathan, D.G., Application of endonuclease mapping to the analysis and prenatal diagnosis of thalassaemias caused by globin gene deletion, *New Engl. J. Med.*, 299 (1978) 166–172.
- Ottolenghi, S., Lanyon, W.G., Williamson, R., Weatherall, D.J., Clegg, J.B. and Pitcher, C.S., Human globin gene analysis for a patient with  $\beta^0/\delta\beta^0$ -thalassaemia, *Proc. Natl. Acad. Sci. USA*, 72 (1975) 2294–2299.
- Ottolenghi, S., Comi, P., Giglioni, B., Tolstoshev, P., Lanyon, W.G., Mitchell, G.J., Williamson, R., Russo, G., Musumeci, S., Schiliro, G., Tsistrakis, G.A., Charache, S., Wood, W.G., Clegg, J.B. and Weatherall, D.J.,  $\delta\beta^0$ -Thalassaemia is due to a gene deletion, *Cell*, 9 (1976) 71–80.
- Ottolenghi, S., Giglioni, B., Comi, P., Gianni, A.M., Polli, E., Acquaye, C.T.A., Oldham, J.H. and Masera, G., Globin gene deletion in HPFH,  $\delta^0\beta^0$  thalassaemia and Hb Lepore disease, *Nature*, 278 (1979) 654–657.
- Ramirez, F., O'Donnell, J.V., Marks, P.A., Bank, A., Musumeci, S., Schiliro, G., Pizzarelli, G., Russo, G., Luppis, B. and Gambino, R., Abnormal or absent  $\beta$  mRNA in  $\beta^0$  Ferrara and gene deletion in  $\delta\beta$ -thalassaemia, *Nature*, 263 (1976) 471–475.
- Roberts, R.J., Restriction and modification enzymes and their recognition sequences, *Gene*, 4 (1978) 183–193.
- Southern, E.M., Detection of specific sequences among DNA fragments separated by gel electrophoresis, *J. Mol. Biol.*, 98 (1975) 503–517.
- Weatherall, D.J. and Clegg, J.B. (Eds.), *The Thalassaemia Syndromes*, 2nd ed., Blackwell, Oxford, 1972.

Communicated by W. Szybalski.