The EMBO Journal vol.7 no.9 pp.2795-2799, 1988

Nullisomic deletion of the *mcf.2* transforming gene in two haemophilia B patients

D.S.Anson[1], D.J.Blake, P.R.Winship, D.Birnbaum[2] and G.G.Brownlee

Sir William Dunn School of Pathology, University of Oxford, South Parks Road, Oxford OX1 3RE, UK and [2]Unité 119 de l'INSERM, 27 boulevard Lei Roure, 13009 Marseille, France

[1]Present address: Cancer Biology Laboratory, The Salk Institute, 10010 North Torrey Pines Road, La Jolla, CA 92138, USA

Communicated by G.G.Brownlee

The *mcf.2* transforming gene sequence has been located to the region between 29 and 61 kb 3′ of the factor IX gene. Two unrelated haemophilia B patients who raise antibodies to infused factor IX ('inhibitors') have deletions in excess of 273 kb encompassing the factor IX and *mcf.2* genes and a CG-rich island. We believe these patients show the first nullisomic deletion of a transforming gene to be reported. No clinical condition can be attributed to the loss of the *mcf.2* gene.

Key words: haemophilia B/genomic walk/*mcf.2*/gene deletions

Introduction

The transforming gene *mcf.2* was identified in transformants produced by co-transfection of NIH 3T3 cells with DNA from the human breast carcinoma cell line, MCF7, and the plasmid pKOneo in a tumorigenicity assay (Fasano *et al.*, 1984). *mcf.2* is an X-linked transforming gene showing no homology to previously characterized oncogenes or growth factors (Noguchi *et al.*, 1987). It has been localized to within 270 kb of the factor IX gene by pulsed-field gel electrophoresis (Nguyen *et al.*, 1987). This suggests that haemophilia B patients, having large deletions encompassing the factor IX locus, might also have partial or total deletions of the *mcf.2* gene.

Partial or total factor IX deletions are rare within the population. The incidence of haemophilia B is one in every 60 000 births (for review see Brownlee, 1988). Patients with gene deletions represent $<1\%$ of the affected total, although this may be an underestimate as small deletions may be undetected. Deletions are often (Giannelli *et al.*, 1983; Peake *et al.*, 1984; Bernardi *et al.*, 1985; Hassan *et al.*, 1985) but not always (Chen *et al.*, 1985) associated with the inhibitor phenotype, which is found in $\sim 1\%$ of haemophilia B patients.

Allelic deletions of oncogenes have been observed in a number of sporadic tumours. Loss of one of the c-*ras*^{Ha} alleles has been observed in DNA from many carcinomas, including those in breast and colon (Theillet *et al.*, 1986; Yokata *et al.*, 1986), in patients constitutionally heterozygous for this locus. Allelic deletion of c-*myb* has been reported in primary and metastatic tumours and may be correlated with the progression of malignant disease (Yokata *et al.*, 1986). In familial tumours such as retinoblastoma and Wilms' tumour, non-random deletions of chromosomal sequences are very frequently found and are thought to unmask recessive oncogenes (for review see Nördenskjörd and Lundberg, 1986). In no instance has a nullisomic deletion of either an oncogene or transforming gene been shown to occur in any tumour.

Results

Isolation of the factor IX flanking clones

Figure 1 shows a restriction map of 285 kb of the human X chromosome encompassing the factor IX locus, deduced from the isolation and mapping (see Materials and methods) of a set of overlapping cosmid and phage lambda clones from normal human DNA. The cloned region comprises 58 kb upstream of the factor IX promoter, the 33-kb structural gene encoding factor IX and 194 kb downstream of the factor IX polyadenylation signal. Much of the gene walking was achieved using cosmid technology (see Materials and methods) but this approach proved impossible for the sequence -58 to -35 kb and 162 to 227 kb (Figure 1). Clones from these regions were isolated from lambda EMBL3 (Frischauf *et al.*, 1983) libraries grown in recombination deficient *Escherichia coli* hosts (Wyman *et al.*, 1985). This observation demonstrates that the choice of vector and host is important for the isolation of certain clones.

Localization of the mcf.2 transforming sequence to c1.1

The *mcf.2* cDNA probe, pBSIII-9 (see Materials and methods), was shown by dot-blot hybridization to hybridize to cosmid clone c1.1 (data not shown). Southern blotting (Figure 2) of restriction fragments and summation of the positively hybridizing bands (i.e. 4.7, 3.3 and 3.1 kb) in an *Eco*RI–*Bam*HI digest showed that pBSIII-9 was localized to an 11-kb region of c1.1 (Figure 1). The cDNA clone pBSIII-9 is derived from the non-transforming end of the *mcf.2* gene localizing one end of the *mcf.2* transcription unit to between 50 and 61 kb 3′ of the polyadenylation site of the factor IX gene. The genomic clone pmcf.2-8 (see Materials and methods) was localized to clone lambda 0.1 by Southern blot hybridization (results not shown). By comparison of the *Eco*RI restriction map of the normal *mcf.2* gene (Noguchi *et al.*, 1987) with the *Eco*RI restriction map in Figure 1 the position of the cDNA probe pb51 (see Materials and methods) can be assigned accurately to between 29 and 50 kb 3′ of the factor IX polyadenylation site (see Figure 1). The *mcf.2* gene is transcribed in the opposite direction to factor IX (D.Birnbaum, unpublished results) as indicated by the arrows in Figure 1. In summary, the *mcf.2* transforming gene is located between 29 to 61 kb 3′ of the factor IX polyadenylation site but could extend further in either direction as the promoter and poly(A) site have not yet been defined.

The mcf.2 structural gene is deleted in two unrelated haemophilia B inhibitor patients, Manchester 1 and Manchester 2

Southern blot analysis (Figure 3a) demonstrates that restriction fragments probed by a factor IX exon f probe (probe III, Anson *et al.*, 1984) are deleted in both of the inhibitor patients, thus confirming earlier results (Giannelli *et al.*, 1983) showing that the entire factor IX gene is deleted in both these patients. Figure 3b shows that the deletion extends to the *mcf.2* locus as probe pBSIII-9 fails to hybridize to DNA from either of the inhibitor patients. The genomic *Eco*RI hybridization pattern in Figure 3b is identical to the *Eco*RI pattern for cosmid c1.1 (Figure 2) when probed with pBSIII-9. This confirms that the derivative restriction map around the *mcf.2* locus is correct and rules out the possibility of rearrangement in the cosmid, a phenomenon occasionally encountered in gene walking projects. Failure to detect any hybridization signal with probe 9 (Figure 3c) shows that the gene deletion in these patients extends to more than 215 kb 3′ of the factor IX gene. Probe E fails to hybridize to patient DNA (data not shown), showing that the deletion is >58 kb upstream of the factor IX promoter. The extent of the deletion is therefore >273 kb in each case and encompasses both the factor IX and the *mcf.2* loci.

Characterization of the CG-rich island sequence 163 kb downstream from the factor IX promoter

The presence of unique *Not*I and *Sac*II restriction sites 163 kb downstream of the factor IX gene (see Figure 1) suggests that it may be a CG-rich island (Bird, 1986). Complete dideoxy sequence analysis of a subcloned 2.9-kb restriction fragment from clone c4.1 (see Figure 1) revealed the presence of a cluster of rare-cutting restriction sites (Figure 4). (The complete sequence has been submitted to the EMBL database with the accession number X07398.) Most of the rare-cutting sites are clustered in the central 1 kb of the sequence between the arrows, as indicated in Figure 4. The G+C content within this region is 67.5% and there are 130 GC dinucleotides and 74 CG dinucleotides giving a GC:CG ratio of 1:1.76. These values approximate to those indicative of a classic HTF island (Bird, 1986). The methylation pattern has not been determined for this CG-rich island, but earlier

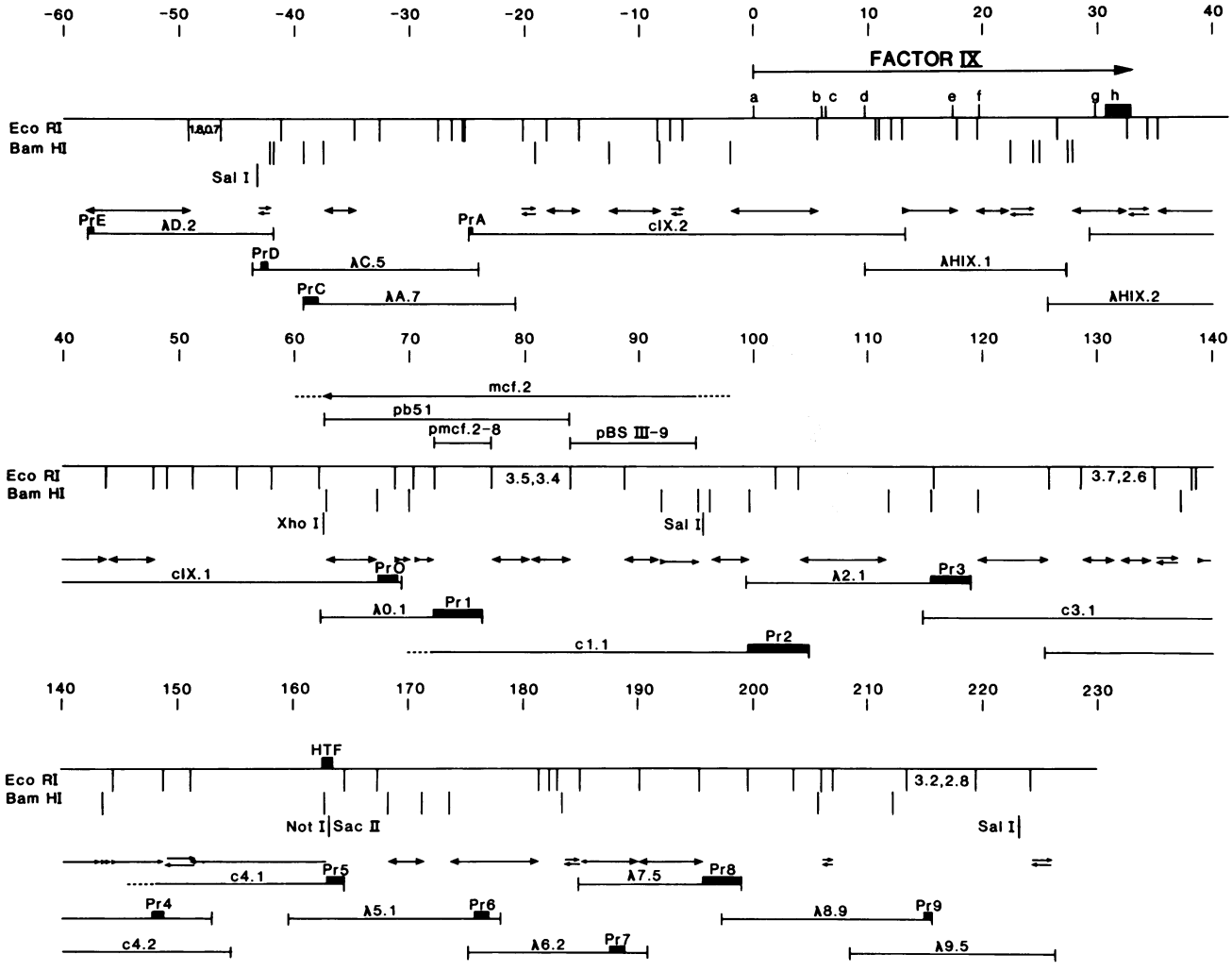

Fig. 1. A restriction map of 285 kb of DNA including and flanking the factor IX gene. The positions of the factor IX gene (0−33 kb), the *mcf.2* gene (62−94 kb) and the CG-rich island, labelled HTF (163 kb), are indicated above the restriction map. Restriction sites for the enzymes *Eco*RI, *Bam*HI, *Sal*I, *Sac*II, *Xho*I and *Not*I are shown as deduced from mapping the various cosmid and lambda clones. ■ = probes for walking. As the exact size of the *mcf.2* gene is uncertain the 5′ and 3′ regions are dashed outside the regions hybridizing to the *mcf.2* probes. The map positions of the three probes used to identify the *mcf.2* locus are also shown. Fragments hybridizing to the *Alu*I repeat sequence are indicated by <-> and those hybridizing to total human DNA, but not to *Alu*I sequences by >-<. The smallest identified hybridizing fragment is indicated in each case.

pulsed field mapping results (Nguyen *et al.*, 1987) suggest that it is hypomethylated, as it is cleaved by the methylation-sensitive restriction enzymes *Sac*II, *Bss*HII and *Eag*I (see Figure 4).

Discussion

We have shown that the locus encoding the *mcf.2* transforming gene is located in a region 29–61 kb 3′ from the factor IX polyadenylation site. The factor IX and *mcf.2* genes and a CG-rich island are completely deleted in two unrelated haemophilia B inhibitor patients, the total size of the deletions being >273 kb. We believe that these inhibitor patients, both of whom are from families with a history of the disease, represent the first documented case of a nullisomic deletion of a transforming gene.

The two inhibitor patients lack common clinical symptoms, other than those of haemophilia B, attributable to the loss of the *mcf.2* gene (Matthews *et al.*, 1988). This implies that the *mcf.2* gene product is facultative in these patients. However, the absence of a deletion phenotype could be explained in other ways. It is possible that the *mcf.2* deletion phenotype is masked by the severity of haemophilia B or by the complications that arise from the anti-factor IX immune response to replacement therapy with factor IX. A number of genetic diseases including Huntington's chorea and Alzheimer's disease show late adult onset of symptoms. Since both of the inhibitor patients are in their early thirties the effect of the *mcf.2* deletion may become apparent as they get older. It is also possible that *mcf.2* has a functional counterpart elsewhere in the genome that will compensate for the loss of function at the *mcf.2* locus. The deletion of

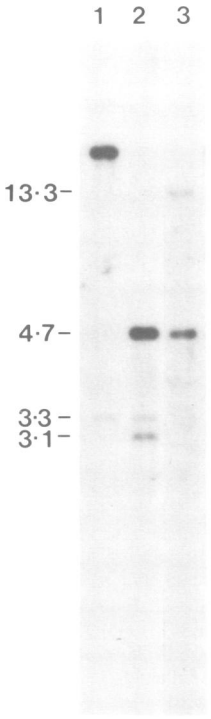

Fig. 2. Localization of *mcf.2* probe pBSIII-9 to clone c1.1. c1.1 was digested with *Bam*HI (**lane 1**), *Bam*HI–*Eco*RI (**lane 2**) and *Eco*RI (**lane 3**) and Southern blots were probed with nick-translated pBSIII-9. The sizes of the smaller fragments are identified in kilobases. The fact that the 13.3-kb *Eco*RI fragment in lane 3 is faint suggests that the cDNA probe pBSIII-9 is predominantly encoded by exons in the 4.7-kb *Eco*RI fragment.

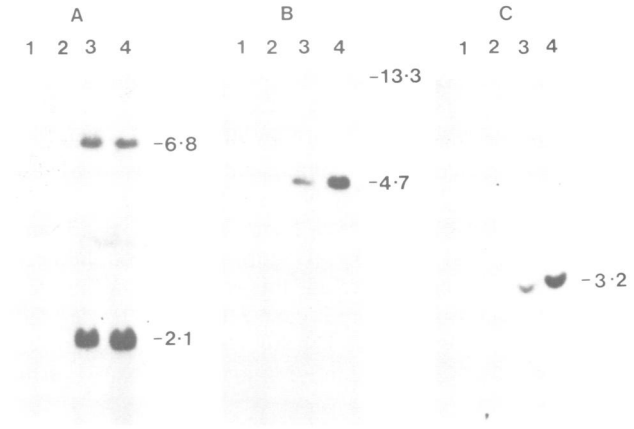

Fig. 3. Southern blot analysis of *Eco*RI-digested patient and control DNA. (**A**) and (**B**) are duplicate blots, probed with factor IX, probe III (see Materials and methods) in (A) and the *mcf.2* probe pBSIII-9 (see Materials and methods) in (B). **Lane 1** is haemophilia $B_{Manchester\ 1}$, **lane 2** is haemophilia $B_{Manchester\ 2}$, **lane 3** is normal male, **lane 4** is normal female. (**C**) is probed with probe 9 (see Figure 1). **Lane 1** is haemophilia $B_{Manchester\ 1}$, **lane 2** is haemophilia $B_{Manchester\ 2}$, **lane 3** is CL-2D, **lane 4** is GM1416B.

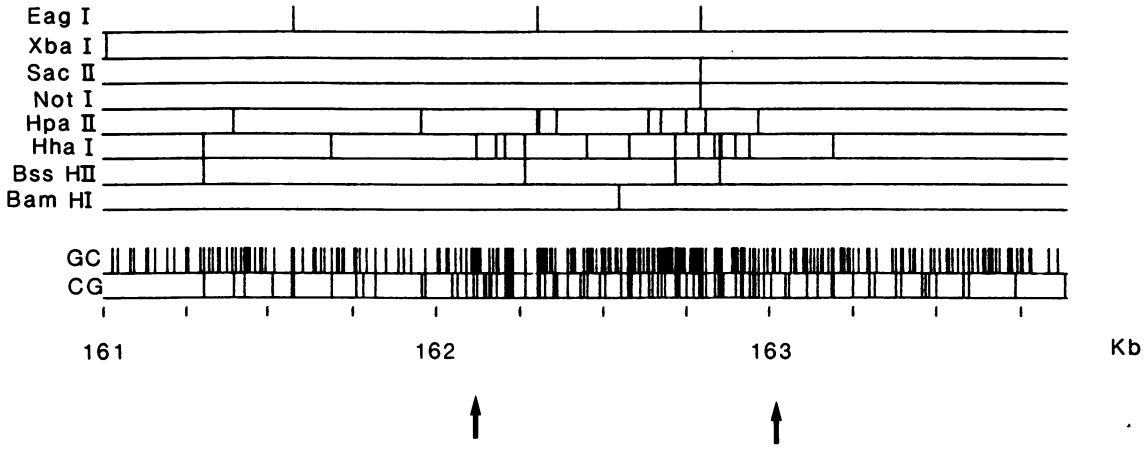

Fig. 4. Restriction map of the sequenced 2.9-kb restriction fragment (between 161 and 164 kb) showing the distribution of CG and GC dinucleotides including the rare-cutting sites.

the adjacent *mcf.2* gene in haemophilia B patients is analogous to the homozygous deletion of adjacent $\theta-1$ and α-globin genes in a child homozygous for the South-East Asian α-thalassemia determinant (Fischel-Ghodsian *et al.*, 1987). The child appears to develop normally implying that the θ-1 gene is non-essential. In summary, although the evidence favours a facultative role for *mcf.2*, it would be premature to conclude it has no function.

Allelic deletions of autosomal oncogenes are known to contribute to neoplastic transformation (Yokata *et al.*, 1986). Since *mcf.2* is a transforming gene, allelic deletion of the locus in females might result in neoplastic transformation. However, in the families of the two inhibitor patients there is no history of tumours. Other haemophilia B patients with deletions encompassing the *mcf.2* locus also have no family history of tumours (Matthews *et al.*, 1988). Constitutional allelic and homozygous deletions of the retinoblastoma (Rb-1) gene have been reported (Friend *et al.*, 1986) in patients with intraocular tumour. The Rb-1 gene product is thought to regulate the expression of structural transforming genes (Comings *et al.*, 1973) and is therefore not strictly an oncogene, but a tumour suppressing gene. *mcf.2* is clearly not a tumour suppressing gene as it has been shown to be transforming *in vitro* (Fasano *et al.* 1984) and deletion results in no tumour phenotype.

CG-rich islands are known to be gene markers in the vertebrate genome (for review see Bird, 1987). The CG-rich island deleted in the inhibitor patients was thought to identify the *mcf.2* gene (Nguyen *et al.*, 1987). However, the distance between the most 5′ region of the *mcf.2* gene and the CG island is 70 kb making this suggestion unlikely. Nevertheless we cannot exclude this suggestion definitively since the 5′ end of the *mcf.2* gene remains to be defined. It seems more likely to us that this CG island is a marker for an as yet undefined gene lying just 5′ or just 3′ of the island.

Materials and methods

Isolation of high mol. wt DNA

DNA was prepared from Epstein–Barr virus transformed lymphoblastoid lines (Y.Boyd, Department of Genetics, University of Oxford) of the two inhibitor patients, haemophilia $B_{Manchester\ 1}$ and haemophilia $B_{Manchester\ 2}$ from GM1416B (human 4X lymphoblastoid cell line, Human Genetic Mutant Repository, NJ, USA) and from CL-2D [human X in the Chinese hamster background constructed by S.J.Goss from a fusion of Chinese hamster cells and irradiated human peripheral blood lymphocytes (Goss and Harris, 1977)], by an adaptation of an existing method (Kunkel *et al.*, 1977).

Construction and screening of lambda and cosmid libraries

Partial *Mbo*I digests of high mol. wt GM1416B DNA were phosphatased and size fractionated on sucrose gradients (Maniatis *et al.*, 1978). Fragments in the 15–24 kb size range were cloned into *Bam*HI–*Eco*RI-cut lambda EMBL3 (Frischauf *et al.*, 1983). Fragments in the range 35–45 kb were cloned into the cosmid vectors pDV or pDVcos (Knott *et al.*, 1988) using an 'arms' procedure (Ish-Horowicz and Burke, 1981). After *in vitro* packaging, cosmids were infected into *E.coli* NM554 (*sup°, recA13, hsdR*) and lambda into *E.coli* CES200 (*recBC, sbc, hsdR*) or CES201 (*recABC, sbc, hsdR*). Libraries were screened using standard methods for both lambda (Benton and Davis, 1977) and cosmids (Grosveld *et al.*, 1981).

Subcloning

Fragments of interest were subcloned into the plasmid vector pAT153/*Pvu*II/8 (Anson *et al.*, 1984) or pBR322 using standard procedures. Alternatively pDVcos clones were digested with either *Eco*RI or *Bam*HI (which do not cut the vector) and recircularized with T4 DNA ligase at low DNA concentration. This effectively subclones the end fragments of the cosmid.

Southern blot hybridization

Restriction-enzyme-digested DNA was transferred to nylon membranes (Genescreen Plus, NEN) according to the manufacturers specifications. Restriction fragment probes were labelled by nick translation and hybridized under the conditions specified by the manufacturer for 18 h. Filters were washed to high stringency (0.1 or 0.2 X SSC, 65°C) and autoradiographed for 1–4 days at −70°C with an intensifying screen.

DNA sequencing

Sonicated DNA (Fuhrman *et al.*, 1981) fragments from the CG-rich island were subcloned into M13mp8 (Messing and Vieira, 1982) and sequenced using the dideoxy method (Sanger *et al.*, 1977). Some sequence was also determined using the method of Maxam and Gilbert (1980).

mcf.2 probes

pBSIII-9 is a 0.5-kb cDNA fragment derived from the 5′ end of the *mcf.2* cDNA. pB51 is a 1.4-kb cDNA fragment (Noguchi *et al.*, 1987) from the 3′ end of the *mcf.2* cDNA. pmcf.2-8 is a 4.8-kb genomic fragment.

Acknowledgements

We wish to thank David Hunter (Genetics Laboratory, University of Oxford) for assisting with the cell culture and Cathy Lee and Philip Thornton-Evison for their help with the preparation of the figures. This work was supported by MRC grants to D.S.A. and G.G.B., a SERC studentship to D.J.B. and a Royal Society Fellowship to P.R.W.

References

Anson,D.S., Choo,K.H., Rees,D.J.G., Giannelli,F., Gould,K., Huddleston, J.A. and Brownlee,G.G. (1984) *EMBO J.*, **3**, 1053–1060.
Benton,W.D. and Davis,R.W. (1977) *Science*, **196**, 180–182.
Bernardi,F., Del Senno,L., Barbieri,R., Buzzani,D., Gambari,R., Marchetti,G., Canconi,F., Panicucci,F., Positano,M. and Pitruzzello,S. (1985) *J. Med. Genet.*, **22**, 305–307.
Bird,A.P. (1986) *Nature*, **321**, 209–213.
Bird,A.P. (1987) *Trends Genet.*, **3**, 342–347.
Brownlee,G.G. (1988) In Hoffbrand,A.V. (ed.), *Recent Advances in Clinical Haemotology*, **5**. Churchill, Edinburgh, in press.
Chen,S.-H., Yoshitake,S., Chance,P.L., Bray,G.L., Thompson,A.R., Scott,C.R. and Kurachi,K. (1985) *J. Clin. Invest.*, **76**, 2161–2164.
Comings,D.E. (1973) *Proc. Natl. Acad. Sci. USA*, **70**, 3324–3328.
Fasano,O., Birnbaum,D., Edlund,L., Fogh,J. and Wigler,M. (1984) *Mol. Cell. Biol.*, **9**, 1695–1705.
Fischel-Ghodsian,N., Higgs,D.R. and Beyer,E.C. (1987) *Nature*, **329**, 397.
Friend,S.H., Bernards,R., Rogelj,S., Weinberg,R.A., Rapaport,J.M., Albert,D.M. and Dryja,T.P. (1986) *Nature*, **323**, 643–646.
Frischauf,A.M., Lehrach,H., Poustka,A. and Murray,N. (1983) *J. Mol. Biol.*, **170**, 827–842.
Fuhrman,S.A., Deininger,P.L., LaPorte,P., Friedmann,T. and Geiduschek,E.P. (1981) *Nucleic Acids Res.*, **9**, 6439–6456.
Giannelli,F., Choo,K.H., Rees,D.J.G., Boyd,Y. Rizza,C.R. and Brownlee,G.G. (1983) *Nature*, **303**, 181–182.
Goss,S.J. and Harris,H. (1977) *J. Cell. Sci.*, **25**, 17–37.
Grosveld,F.G., Dahl,H.-H.M., de Boer,E. and Flavell,R.A. (1981) *Gene*, **13**, 227–237.
Hassan,H.J., Leonardi,A., Guerriero,R., Chelucci,C., Cianetti,L., Ciavarella,N., Ranieri,R., Pilolli,D. and Peschle,C. (1985) *Blood*, **66**, 728–730.
Ish-Horowicz,D. and Burke,J.F. (1981) *Nucleic Acids Res.*, **9**, 2989–2998.
Knott,V., Rees,D.J.G., Cheng,Z. and Brownlee,G.G. (1988) *Nucleic Acids Res.*, **16**, 2601–2612.
Kunkel,L.M., Smith,K.D., Boyer,S.H., Borgaonkar,D.S., Wachtel,S.S., Miller,R.J., Breg,W.R., Jones,H.W. and Rary,J.M. (1977) *Proc. Natl. Acad. Sci. USA*, **74**, 1245–1249.
Maniatis,T., Hardison,R.C., Lacy,E., Lauer,J., O'Connell,C., Quon,D., Sim,G.K. and Efstratiadis,A. (1978) *Cell*, **15**, 687–701.
Matthews,R.J., Peake,I.R., Blake,D.J., Brownlee,G.G. and Bloom,A.L. (1988), in press.
Maxam,A.M. and Gilbert,W. (1980) *Methods Enzymol.*, **65**, 499–559.
Messing,J. and Vieira,J. (1982) *Gene*, **19**, 269–276.
Nguyen,C., Pontarotti,P., Birnbaum,D., Chimini,G., Rey,J.A., Mattei,J.-F. and Jordan,B.R. (1987) *EMBO J.*, **6**, 3285–3289.
Noguchi,T., Mattei,M.G., Oberle,I., Planche,J., Imbert,J., Pelassy,C., Birg,F. and Birnbaum,D. (1987) *EMBO J.*, **6**, 1301–1307.

Nordenskjöld,M. and Lundberg,C. (1986) *Ann. Clin. Res.*, **18**, 307–313.
Peake,I.R., Furlong,B.L., Bloom,A.L. and Brownlee,G.G. (1984) *Lancet*, **i**, 242–243.
Sanger,F., Nicklen,S. and Coulson,A.R. (1977) *Proc. Natl. Acad. Sci. USA*, **74**, 5463–5468.
Theillet,C., Lidereau,R., Escort,C., Hutzell,P., Brunet,M., Gest,J., Schlom,J. and Callahan,R. (1986) *Cancer Res.*, **46**, 4776–4781.
Wyman,A.R., Wolfe,L.B. and Botstein,D. (1985) *Proc. Natl. Acad. Sci. USA*, **82**, 2880–2884.
Yokata,J., Tsunetsugu-Yokata,Y., Battifora,H., Le Fevre,C. and Cline,M.J. (1986) *Science*, **231**, 261–265.

Received on April 12, 1988; revised on June 13, 1988