The human myoglobin gene: a third dispersed globin locus in the human genome

Alec J.Jeffreys[1], Victoria Wilson[1], Alain Blanchetot[1], Polly Weller[1], Ad Geurts van Kessel[2], Nigel Spurr[3], Ellen Solomon[3] and Peter Goodfellow[3]

[1]Department of Genetics, University of Leicester, Leicester LE1 7RH, UK, [2]Department of Cell Biology and Genetics, Erasmus University, P.O. Box 1738, 3000 DR Rotterdam, The Netherlands, and [3]Imperial Cancer Research Fund, P.O. Box 123, Lincoln's Inn Fields, London WC2A 3PX, UK

Received 15 Feburary 1984; Accepted 9 March 1984

ABSTRACT

Human myoglobin is specified by a single gene. Unique sequence DNA probes were isolated from the cloned gene and used to test for the presence of the human myoglobin gene in a series of human rodent somatic cell hybrids containing various complements of human chromosomes. The myoglobin gene cosegregated with human chromosome 22. Somatic cell hybrids containing translocation chromosomes carrying part of chromosome 22 were used to locate the myoglobin gene to the region 22q11→22q13. The myoglobin gene is therefore not linked to the α-globin gene cluster on chromosome 16 or the β-globin cluster on chromosome 11, and represents a third dispersed globin locus in the human genome.

INTRODUCTION

Tetrameric haemoglobins of vertebrates have evolved by a succession of gene duplications, commencing with an initial αβ-globin gene duplication early in vertebrate evolution, about 450 million years ago [1-3]. In birds and mammals, the α- and β-related globin gene subfamilies are organised into two unlinked gene clusters, each containing a variety of developmentally regulated genes which specify a range of specialised non-adult and adult haemoglobins [3,4]. In contrast, the amphibian Xenopus tropicalis has a single cluster containing both α- and β-globin genes [5]. In X.laevis, two similar unlinked clusters have been generated by tetraploidization [5-7]. It thus appears that α- and β-globin genes originally arose as a tandem gene duplication, and that this close linkage has been maintained in at least some amphibians, whereas the α- and β-globin genes became dispersed to separate chromosomes in a common ancestor of birds and mammals [5].

We have recently characterised a third diverged member of the globin gene family, the myoglobin gene [8-10]. The myoglobin/haemoglobin duplication preceded the αβ-globin gene duplication, and occurred according to various estimates between 500 and 800 million years ago [1,2]. In contrast to haemoglobin, myoglobin is monomeric and serves to facilitate the

© IRL Press Limited, Oxford, England.

diffusion of oxygen within muscle fibres. In some animals, for example diving mammals such as whales and seals, elevated levels of myoglobin also act as a significant muscle oxygen store [11]. In both seal and man, myoglobin is specified by a single gene which has the same three exon-two intron organization as found in α- and β-globin genes [9,10]. The myoglobin gene differs from its haemoglobin relatives in having long non-coding regions and an atypical promoter region. There is also evidence for additional diverged myoglobin related DNA sequences in both human and seal DNA [8,10].

As a first step in the analysis of linkage relationships between myoglobin and haemoglobin genes, we now report the chromosome assignment of the human myoglobin gene, using man-rodent somatic cell hybrids.

MATERIALS AND METHODS

Isolation of DNA

Maja human cells [12], RAG mouse cells [13] and the human-rodent somatic cell hybrids listed in Table 1 were grown in Dulbecco's modification of Eagles medium supplemented with 10% calf serum. DNA was isolated from cultured cells, and from human placenta, Balb/c mouse liver and Chinese hamster liver, as described elsewhere [14].

Purification of human myoglobin-specific probes

The cloning of the human myoglobin gene into the bacteriophage vector λ L47.1, and the subcloning of exon 1- and exon 2- containing fragments into pAT 153 to give pHM.27.B2.9 and pHM.27.B1.1 respectively, are described elsewhere [10]. A 1.0 kb EcoRI fragment was purified from pHM.27.B2.9, and a 0.8 kb BstEII fragment isolated from pHM.27.B1.1, by restriction endonuclease cleavage and electrophoresis onto DE81 paper [15].

Filter hybridizations

10 μg samples of human, mouse, Chinese hamster and somatic cell hybrid DNAs were digested with EcoRI (Boehringer) and electrophoresed in a 0.7% agarose gel. DNA was denatured *in situ* and transferred to nitrocellulose filters (Schleicher Schüll) by the method of Southern [16]. Human myoglobin DNA fragments were denatured by heating to 100° for 3 min and labelled with ^{32}P in a conventional "nick translation" reaction; we have previously shown that the use of single stranded DNA as a substrate for *E. coli* DNA polymerase I results in very efficient incorporation of α-^{32}P-dCTP into DNA which is suitable for filter hybridization [10]. Southern blot filters were hybridized with ^{32}P-labelled probe DNA in 1xSSC at 65° in the presence of dextran sulphate, as described elsewhere [5]. In most hybridizations, 50

μg/ml alkali-denatured (0.3 M NaOH, 20 mM EDTA, 100° for 20 min) human placental DNA was added to suppress any background labelling resulting from hybridization to repeated DNA sequences. After hybridization, filters were washed in 0.2xSSC at 65° and autoradiographed.

RESULTS

Detection of human myoglobin gene sequences

The organization of the cloned human myoglobin gene is shown in Fig. 1A. A 1.0 kb EcoRI fragment ($MbE_{1.0}$) was purified from the first intervening sequence and tested for its ability to hybridize specifically to the human myoglobin gene and not to mouse DNA sequences. As shown in Fig. 1B, the $MbE_{1.0}$ probe fulfilled these requirements, and only detected its homologous 1.0 kb EcoRI fragment in a Southern blot hybridization with an EcoRI digest of human DNA. No hybridization to mouse or Chinese hamster DNA was found.

For confirmatory studies, a 0.8 kb BstEII fragment ($MbB_{0.8}$) was isolated which contained the central exon of the human myoglobin gene, plus part of the second intron (Fig. 1). This probe hybridized (in 1xSSC at 65°) to a complex smear of fragments in digests of both human and rodent DNA, suggesting the presence of repeated DNA in this fragment (data not shown). We have previously shown that the central exon of the seal myoglobin gene also hybridizes to multiple DNA fragments in restriction endonuclease digests of human DNA, and have suggested that these extra DNA sequences might represent some form of diverged myoglobin gene or exon family [10]. Repeating the hybridization with the $MbB_{0.8}$ fragment in 1xSSC at 65° in the presence of 50 μg/ml competitor human DNA, followed by a stringent posthybridization wash in 0.2xSSC at 65°, reduced the profile of hybridizing human DNA to a single 6.9 kb EcoRI fragment, as predicted from the map of the human myoglobin gene (Fig. 1). Under these conditions, there was little or no hybridization to rodent DNA.

The myoglobin gene is located on human chromosome 22

To determine the chromosome location of the human myoglobin gene, DNAs prepared from a panel of man-rodent somatic cell hybrids containing various complements of human chromosomes were digested with EcoRI and probed with the $MbE_{1.0}$ intron DNA fragment. As shown in Fig. 1, this probe clearly discriminated between somatic cell hybrids which either contained or lacked the human gene. Of 13 hybrid cell lines tested, 5 were positive and shared only human chromosomes 21, 22 and X in common (Table 1). The remaining 8 hybrids lacked the human myoglobin gene, yet between them contained all human

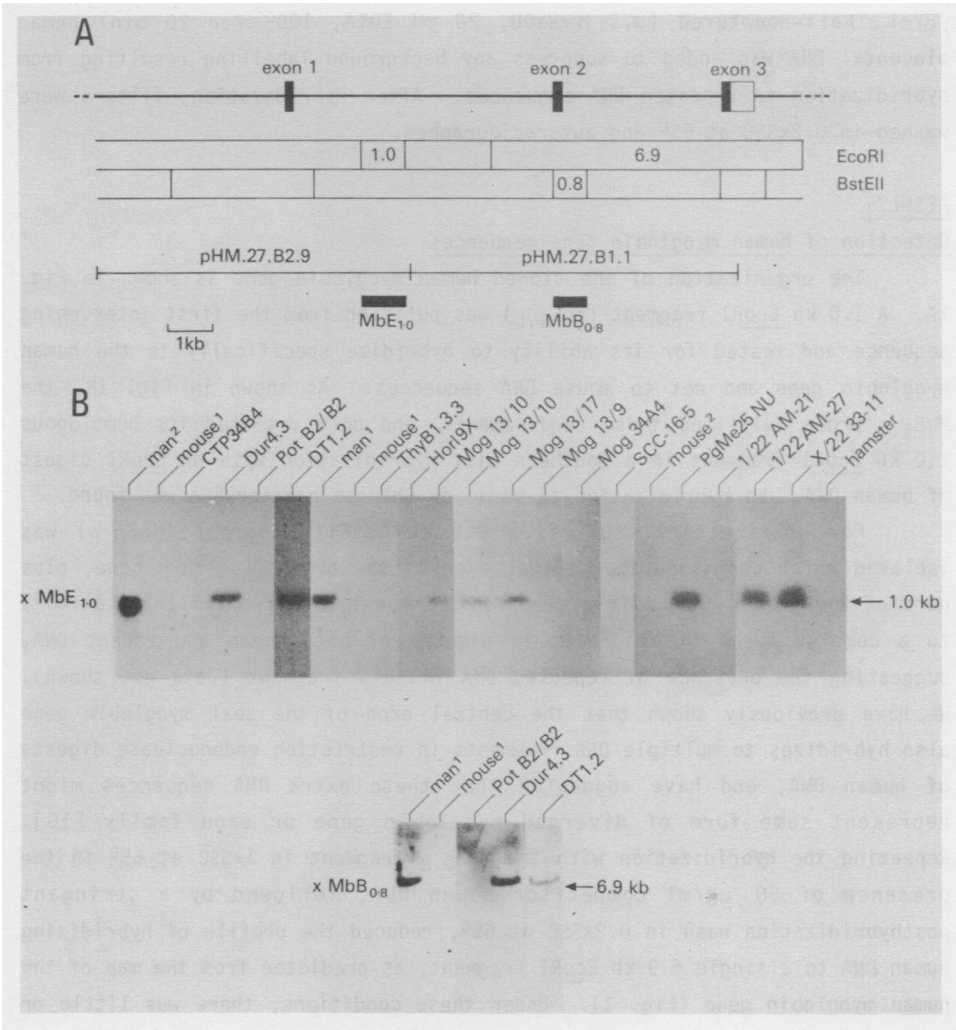

Figure 1. Detection of human myoglobin gene specific sequences in man-rodent somatic cell hybrids.
A. Isolation of DNA probes for the human myoglobin gene. The structure of the myoglobin gene is shown, together with EcoRI and BstEII restriction maps determined by cleavage analysis of the cloned gene. An intron 1 probe was isolated as a 1.0 kb EcoRI fragment ($MbE_{1.0}$) from the subclone pHM.27.B2.9, and an exon 2 probe as a 0.8 kb BstEII fragment ($MbB_{0.8}$) from pHM.27.B1.1.
B. Detection of human myoglobin gene specific sequences in man-rodent somatic cell hybrids. 10 μg samples of human, rodent or somatic cell hybrid DNA were digested with EcoRI, electrophoresed through a 0.7% agarose gel, transferred to nitrocellulose and hybridized with the $MbE_{1.0}$ or $MbB_{0.8}$ probes isolated from the human myoglobin gene. The photograph shows a composite sample of autoradiographs from several separate experiments. Mouse[1], RAG DNA; mouse[2], Balb/c liver DNA; man[1], Maja DNA; man[2], human placental DNA.

TABLE 1: ANALYSIS OF HUMAN-MOUSE SOMATIC CELL HYBRIDS FOR HUMAN MYOGLOBIN SPECIFIC SEQUENCES

Hybrid	Ref	Human Chromosomes																						Myoglobin
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	X	
Mog34A4	29	+	-	+	+	+	+	+	+	-	+	+	+	+	+	-	+	-	+	+	-	+	-	+	-
CTP34B4	30	+	+	+	-	+	+	+	+	+	-	-	+	-	+	-	+	tr	+	-	-	-	-	+	-
CTP41P1	30	-	-	+	-	-	+	+	-	-	-	-	-	-	+	-	+	+	-	-	+	-	-	-	-
PotB2/B2	31	-	-	-	-	-	-	-	-	-	-	-	+	-	-	-	-	+	-	+	-	-	-	-	-
Mog13/9	29	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	+	-
Horl 9X	32	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	+	-
ThyB1.3.3	33	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	+	-	+	-
SCC-16-5	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	+	-	-	-
Dur4.3	34	-	-	+	-	+	-	-	-	-	+	+	+	+	+	+	-	+	+	-	+	+	+	+	+
DT1.2	35	-	-			-	-		-		tr	+	-	+	-	+	-	+	+	-	+	+	+	+	+
Mog13/10	29	+	-	+	-	-	-	-	-	-	-	-	-	-	-	-	+	-	-	-	-	+	+	+	+
Mog13/22	29	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-	+	-	-	-	-	+	+	tr	+
Mog13/17	29	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	+	+	+	+
PgMe25 NU	17	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	+	-	+

The origins and details of the somatic cell hybrid are given in the references indicated in the second column. The human chromosome complements were deduced from a combination of karyotypic, antigenic and enzymatic analyses. Karyotypic analysis was performed by a combination of Gll staining and quinacrine banding. Analysis of hybrid cells for human myoglobin gene sequences was performed as described in Fig. 2. tr = trace; blank = not tested. The concordance between the presence/absence of the human myoglobin gene and chromosome 22 is shown by boxes. +, gift from David Cox, Dept. of Paediatrics, University of California, San Francisco, U.S.A.

chromosomes except 15 and 22. These data locate the myoglobin gene on chromosome 22. To confirm this assignment, DNA from the hybrid cell line PgMe-25 NU, which contains only human chromosome 22 [17], was probed with the $MbE_{1.0}$ fragment (Fig. 1). As predicted, this cell line contained the human myoglobin gene (Table 1).

To confirm this assignment, Southern blots were washed free of the $MbE_{1.0}$ probe and reprobed at high stringency with the $MbB_{0.8}$ DNA fragment isolated from the non-overlapping subclone pHM.27.B1.1 (Fig. 1). As expected, all somatic cell hybrids previously scored as positive with the $MbE_{1.0}$ probe showed the predicted 6.9 kb EcoRI fragment from the human myoglobin gene, and all negative somatic cell hybrids failed to hybridize with this probe.

Regional localization of the human myoglobin gene

The human-Chinese hamster cell hybrid 1/22 AM-21 contains amongst its human complement a human chromosome 1 carrying a translocation of the distal part of the long arm of chromosome 22 [18]. Neither the $MbE_{1.0}$ nor the $MbB_{0.8}$ probe detected human myoglobin gene sequences in 1/22 AM-21 DNA (Fig. 1), which excludes the myoglobin gene from the distal region of 22q13 (Figs. 1,2). As predicted, the somatic cell hybrid 1/22 AM-27 [18] carrying the

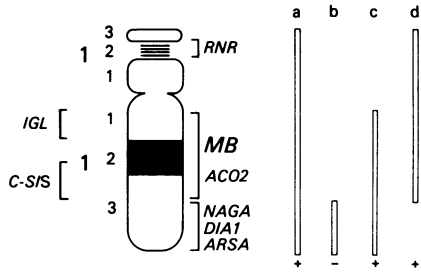

Figure 2. Diagram of human chromosome 22 showing the location of translocation breakpoints used in the regional assignment of the human myoglobin gene (MB). Chromosome 22 segregants used were: a, PgMe-25 NU [17]; b, 1/22 AM-21 [18]; c, X/22 33-11 [18] and d, 1/22 AM-27 [18]. The absence or presence of human myoglobin gene sequences is denoted by - or +. The reported approximate locations of other genes located on chromosome 22 [18, 20, 26-28, 37] are also shown (ACO2, mitochondrial aconitase; ARSA, arylsulphatase A; C-SIS, c-sis oncogene; DIA1, diaphorase; IGL, Ig λ light chain cluster; NAGA, α-N-acetyl-galactosaminidase; RNR, ribosomal RNA).

reciprocal 22/1 translocation product contained the human myoglobin gene. This gene was also detected in the human-Chinese hamster cell hybrid X/22 33-11 containing 22q11 → 22qter translocated onto the X chromosome [18], which therefore excludes the myoglobin gene from 22pter → 22q11 (proximal) (Figs. 1,2).

DISCUSSION

The myoglobin gene represents the final known functional member of the human globin gene superfamily to be characterised. Early evidence from human myoglobin variants suggested that myoglobin is specified by a single gene located on one of the autosomes [19]. This conclusion is now confirmed by direct gene analysis, which demonstrates a single functional myoglobin gene located on chromosome 22 in the region 22q11 → 22q13 (Fig.2).

We have previously shown the existence of an additional diverged family of human DNA sequences preferentially related to the central exon of the human myoglobin gene [10]. A complex profile of hybridizing fragments were also detected by the human exon 2 probe in rodent DNAs, suggesting that this family might be ancient. The complexity of these profiles in both human and rodent DNA precluded the clear detection of human specific sequences in somatic cell hybrids, and the chromosome assignment of these exon-2 related sequences must therefore await their cloning and characterization.

The location of the myoglobin gene on human chromosome 22 identifies a

third dispersed globin locus in the human genome. All other characterized human globin genes and globin pseudogenes are found either in the α-globin gene cluster located on chromosome 16 in the region 16p12 → 16pter [21], or in the β-globin gene cluster on chromosome 11 (position uncertain, but probably in the region 11p15 → 11pter [22, see 36]). The mutual dispersal of myoglobin and α- and β-globin genes obscures the mechanism of the initial globin gene duplication 500-800 million years ago which gave rise to the ancestors of contemporary vertebrate myoglobin and haemoglobin genes. However, by analogy with the more recent αβ-globin gene duplication 450 million years ago [2], it is possible that myoglobin and haemoglobin genes initially arose as a tandem duplication, followed by dispersal via transposition, translocation or chromosome duplication and gene silencing. In view of the fact that diverged haemoglobin genes can remain tightly linked over long evolutionary periods, as witnessed by the αβ-globin gene linkage maintained for 450 million years in the lineage leading to Xenopus [5,7], it is possible that other vertebrate groups might still show linkage between myoglobin and haemoglobin genes. In this respect, it will be interesting to determine the relative locations of the myoglobin gene and the single αβ-globin gene cluster in Xenopus tropicalis [5].

Human chromosome 22 is of considerable interest in view of translocations involving band 22q11 of this chromosome [23] which are associated with chronic myelocytic leukaemia [17] and with Burkitt lymphoma [24,25]. To date, only two cloned genes specific to human chromosome 22 have been identified : the Ig λ light chain cluster in band 22q11 [26,27] and the c-sis oncogene at 22q12 → 22q13 [28, see 37] (Fig. 2). The myoglobin gene therefore represents a useful additional functional gene on this chromosome and provides another well-characterized DNA marker for studies of chromosome 22 translocations.

ACKNOWLEDGEMENTS

A.J.J. is a Lister Institute Research Fellow, and this work was supported by a grant from the MRC and a Fellowship in the European Science Exchange Programme, Royal Society/CNRS (to A.B.).

REFERENCES

1. Hunt, T.L., Hurst-Calderone, S. and Dayhoff, M.O. (1978) In Atlas of protein sequence and structure (ed. M.O. Dayhoff; National Biomedical Research Foundation, Washington D.C.) pp. 229-251.
2. Czelusniak, J., Goodman, M., Hewett-Emmett, D., Weiss, M.L., Venta,

P.J. and Tashian, R.E. (1982) Nature 298, 297-300.
3. Jeffreys, A.J., Harris, S., Barrie, P.A., Wood, D., Blanchetot, A. and Adams, S.M. (1983) In Evolution from Molecules to Men. (ed. D.S. Bendall; Cambridge University Press) pp.175-195.
4. Efstratiadis, A., Posakony, J.W., Maniatis, T., Lawn, R.M., O'Connell, C., Spritz, R.A., DeRiel, J.K., Forget, B.G., Weissman, S.M., Slightom, J.L., Blechl, A.E., Smithies, O.,Baralle, F.E., Shoulders, C.C. and Proudfoot, N.J. (1980) Cell 21, 653-668.
5. Jeffreys, A.J., Wilson, V., Wood, D., Simons, J.P., Kay, R.M. and Williams, J.G. (1980) Cell 21, 555-564.
6. Patient, R.K., Elkington, J.A., Kay, R.M. and Williams, J.G. (1980) Cell 21, 565-573.
7. Hosbach, H.A., Wyler, T. and Weber, R. (1983) Cell 32, 45-53.
8. Wood, D., Blanchetot, A. and Jeffreys, A.J. (1982) Nucleic Acids Res. 10, 7133-7144.
9. Blanchetot, A., Wilson, V., Wood, D. and Jeffreys, A.J. (1983) Nature 301, 732-734.
10. Weller, P., Jeffreys, A.J., Wilson, V. and Blanchetot, A. (1984) EMBO J. 3, 439-446.
11. Wittenberg, J.B. (1970) Physiol. Rev. 50, 559-636.
12. Brodsky, F.M., Parham, P., Barnstable, C.J., Crumpton, M.J. and Bodmer, W.F. (1979). Immunol. Rev. 47, 3-61.
13. Klebe, R.J., Chen, T.R. and Ruddle, F.H. (1970) J. Cell Biol. 45, 74-82.
14. Jeffreys, A.J. (1979) Cell 18, 1-10.
15. Dretzen, G., Bellard, M., Sassone-Corri, P. and Chambon, P. (1981) Anal. Biochem. 112, 295-298.
16. Southern, E.M. (1980) In Methods in Enzymology (ed. R. Wu) 68, 152-176.
17. de Klein, A., Geurts van Kessel, A., Grosveld, G., Bartram, C.R., Hagemeijer, A., Bootsma, D., Spurr, N.K., Heisterkamp, N., Groffen, J. and Stephenson, J.R. (1982) Nature 300, 765-767.
18. Geurts van Kessel, A.H.M., Westerveld, A., de Groot, P.G., Meera Khan, P. and Hagemeijer, A. (1980). Cytogenet. Cell Genet. 28, 169-172.
19. Boyer, S.H., Fainer, D.C. and Naughton, M.A. (1963) Science 140, 1228-1231.
20. Human Gene Map (1983) Am. J. Hum. Genet. 35, 134-156.
21. Barton, P., Malcolm, S., Murphy, C. and Ferguson-Smith, M.A. (1982) J. Mol. Biol. 156, 269-278.
22. de Martinville, B. and Francke, U. (1983) Nature 305, 641-643.
23. Rowley, J.D. (1983) Nature 301, 290-291.
24. Rowley, J.D. (1980) Clin. Haemat. 9, 55-86.
25. Bornheim, A., Berger, R. and Lenoir, G. (1981) Cancer Genet. Cytogenet. 3, 307-316.
26. Erikson, J., Martinis, J. and Croce, C.M. (1981) Nature 294, 173-175.
27. Taub, R., Kirsch, I., Marton, C., Lenoir, G., Swan, D., Tronick, S., Aaronson, S. and Leder, P. (1982) Proc. Nat. Acad. Sci. USA 79, 7837-7841.
28. Dalla-Favera, R., Gallo, R.C., Giallongo, A. and Croce, C. (1982) Science 218, 686-688.
29. Solomon, E., Swallow, D., Burgess, S. and Evans, L. (1979) Ann. Hum. Genet. 42, 273-281.
30. Jones, E.A., Goodfellow, P.N., Kennett, R.H. and Bodmer, W.F. (1976) Somat. Cell Genet. 2, 483-496.
31. Andrews, P.W., Knowles, B.B. and Goodfellow, P.N. (1981) Somat. Cell Genet. 7, 435-443.

32. Heisterkamp, N., Groffen, J., Stephenson, J.R., Spurr, N.K., Goodfellow, P.N., Solomon, E., Carritt, B. and Bodmer, W.F. (1982) Nature 299, 747-749.
33. Goodfellow, P.N., Banting, G., Levy, R., Povey, S. and McMichael, A. (1980) Somat. Cell Genet. 6, 777-787.
34. Solomon, E., Bobrow, M., Goodfellow, P.N., Bodmer, W.F., Swallow, D.M., Povey, S. and Noel, B. (1976) Somat. Cell Genet. 2, 125-140.
35. Swallow, D.M., Solomon, E. and Pajunen, L. (1977) Cytogenet. Cell Genet. 18, 136-148.
36. Gerald, P.S. and Grzeschik, K.H. (1984) Cytogenet. Cell Genet. 37, 103-126.
37. Westerveld, A. and Naylor, S. (1984) Cytogenet. Cell Genet. 37, 155-175.