High resolution gene mapping of the human α globin locus

R D NICHOLLS*, J A JONASSON†, J O’D MCGEE‡, S PATIL§, V V IONASESCU§, D J WEATHERALL*, AND D R HIGGS*

From *the MRC Molecular Haematology Unit, Nuffield Department of Clinical Medicine, and †the Nuffield Department of Pathology, University of Oxford, John Radcliffe Hospital, Oxford OX3 9DU; ‡the Department of Medical Genetics, Churchill Hospital, Headington, Oxford; and §the Department of Pediatrics, Division of Medical Genetics, The University of Iowa Hospitals and Clinics, Iowa City, Iowa 52242, USA.

SUMMARY A combination of polymorphic DNA markers, cytogenetic analysis, and in situ hybridisation has been used for the high resolution assignment of the human α globin gene cluster on chromosome 16. Multiallelic DNA probes from within the α globin cluster were used to determine the number of copies of this locus in three cell lines containing trisomies of the short arm of chromosome 16 and one with a familial inversion, inv(16). The breakpoints in these rearrangements flank the α globin locus and locate a shortest region of overlap to 16p13-1. A meiotic crossover was also localised to this band. In situ hybridisation of biotinylated DNA probes to normal and inverted chromosomes 16 [inv(16)(p13-1;q22)] showed hybridisation sites at opposite ends of the chromosomes, consistent with this regional localisation.

The precise regional localisation of highly polymorphic DNA markers is of fundamental importance in establishing a human genetic linkage map. With accurately localised markers, the combined use of molecular studies and high resolution karyotypic analysis will provide a greater understanding of the normal structure of the human genome and the process of recombination, ultimately establishing the relationship between physical and genetic distance within chromosomes. Furthermore, this combined approach will provide insight into the chromosomal rearrangements and aneuploidies that underly many genetic diseases.

Accurate localisation of the human α globin complex would enable it to be used as a useful model for this combined cytogenetic and molecular strategy. It is a multiallelic locus localised to chromosome 16 within the broad limits of p12→qter. As yet it is the only cloned segment of DNA on the short arm of chromosome 16 but 200 kb of this region has now been characterised (R D Nicholls and D R Higgs, in preparation). It is an important disease locus; mutations in this region (α thalassaemias) are a major cause of morbidity and mortality in many populations. Moreover, it is genetically linked to other important, poorly localised polymorphic genes and disease loci, such as that for adult polycystic kidney disease (APCKD), and is associated with an uncommon form of mental retardation.

In order to obtain a more precise intrachromosomal localisation of the α globin locus we have determined the number of copies of this region in cell lines obtained from four subjects with well defined cytogenetic abnormalities of chromosome 16, using polymorphic DNA markers and in situ hybridisation studies. The combined molecular and karyotypic data localise the human α globin gene locus to 16p13-1.

Materials and methods

Cell lines and other human DNA samples

Cell lines from two subjects with partial trisomy 16p and one parent of each carrying a balanced translocation involving 16p were obtained from the Human Genetic Cell Repository, Institute for Medical Research, Camden, NJ [GM2324, balanced, 46,XX, t(16;22)(p13;q12) and GM2325, unbalanced, 47,XX, +der(22) (that is, partial trisomy for both 16p13→pter and 22q12→pter); GM6227, balanced, 46,XX,t(1;16)(q44;p13:11) and GM6226, unbalanced, 46,XY,−1,+der(1) (that is, partial trisomy for
16p13-11→qter and partial monosomy for 1q44→qter]. The latter two have αβ globin chain synthesis ratios of 1:1 and 3:2, respectively, in peripheral blood reticulocytes. GM2325 is a fibroblast line from an 11 day old child with multiple congenital anomalies. No further data were reported for the other three, lymphoblast, cell lines. A fibroblast cell line (SJ') was isolated from an asymptomatic carrier of a familial pericentric inversion of chromosome 16: 46,XX.inv(16)(p13:1;q22). DNA was also obtained from the peripheral blood of her son (SJ) who has a severe myopathy and other congenital abnormalities associated with the karyotype 46,XY,rec(16), which results in partial trisomy for 16p13-1→qter and partial monosomy for 16q22→qter. As controls, DNA was obtained from the peripheral blood of normal British subjects.

DNA STUDIES

Southern blot hybridisation studies were carried out essentially as described except that Hybond-N membranes (Amersham Int) were used so that prehybridisation, hybridisation, and washing of membranes were performed with the addition of SDS to 0.5%. Final wash stringency was 68°C in 0.015 mol/l NaCl/0.0015 mol/l Na-citrate/0.1% SDS. The probes used were a 4 kb HindIII fragment of the 3'-HVR (hypervariable region) subclone pSEAl and a 1.8 kb SstI ψC probe (fig 1a). pDH8, a clone containing a 1.5 kb PstI α globin genomic fragment, was used for in situ hybridisation.

IN SITU HYBRIDISATION

Human metaphase cells were prepared from cell lines of the inv(16) heterozygote (SJ') and three other subjects. pDH8 was labelled with biotin by nick translation as described except that the mean biotinylated probe size was 150 bp. The details of the in situ hybridisation protocol and detection system are described elsewhere.

Results

REFINEMENT OF α GLOBIN ASSIGNMENT TO P13.11→qter

The human α globin genes were previously localised to 16p12→qter by in situ hybridisation and indirectly by finding a raised αβ globin chain synthesis ratio in a person trisomic for this distal segment. The availability of further subjects with rearrangements of 16p and the identification of two highly informative, multiallelic genetic markers within the α globin gene cluster (IZHVR and 3'-HVR) provided the means to refine this localisation significantly.

The cell line GM2325 has an unbalanced karyotype derived by a 3:1 type of meiotic disjunction in the mother (GM2324), whose karyotype shows a balanced translocation. The breakpoints 16p13 and 22q12 in this rearrangement have been reanalysed because of an improvement in cytogenetic banding techniques and refined to 16p13-11 and 22q11-21 (M Aronson, 1985, personal communication). In the index case, one rearranged chromosome (the der(22)) carrying distal 16p has been transmitted together with the normal 16 homologue from the mother and therefore this cell line contains three copies of 16p13-11→qter (two maternal and one paternal). The 3'-HVR probe (fig 1a) which usually distinguishes an allelic α globin gene clusters shows that GM2325 has inherited both maternal alleles plus a third (presumably paternal) 3'-HVR allele (fig 1b). Therefore the trisomic segment in this cell line includes the α globin gene locus which is thus localised to 16p13-11→qter.

A second pair of balanced and unbalanced chromosomes 16 with a breakpoint at 16p13-11 is represented by GM2626 and GM6227.2 The cytogenetic findings suggest that GM2626 also contains three copies of 16p13-11→qter and that this trisomy results from an adjacent 1q11 type segregation. That is, one of the two rearranged chromosomes (the der(1)) is transmitted with the normal homologue of the other chromosome (the normal 16).

The mother (GM6227) has two 3'-HVR alleles (fig 1b), one linked to the α complex on the normal chromosome 16 and the other to the α globin gene on either the der(1) or der(16) depending on whether the locus is distal or proximal to the breakpoints respectively. From the results for GM2325 (see above) and the globin chain synthesis data (see Materials and methods) it was predicted that GM6226, with the 16p trisomy, should have three 3'-HVR alleles (two maternal and one paternal). Nevertheless, only one maternal and one paternal 3'-HVR allele could be identified (fig 1b and legend).

However, subsequent gene dosage experiments using a ψC gene probe, which detects the multiallelic IZHVR, showed that there are two identical copies of the maternally inherited α globin gene cluster, as described below. Thus, GM6226 has a total of three copies of the α globin gene cluster, again consistent with the localisation to 16p13-11→qter.

LOCALISATION OF A MEIOTIC CROSSOVER NEAR TO THE α GLOBIN CLUSTER

The presence of three copies of the α globin complex in GM6226 is most easily demonstrated using BglII and the ψC probe (fig 1a), which detects an invariant 12.6 kb band, 3' of ψC, an inter-α fragment of variable size, 10.5 to 11.5 kb, and an invariant 1.8 kb band 5' of ψC which is not seen on this autoradiograph (fig 1c). In subjects with two inter-α bands differing in
size, these represent paternal and maternal alleles. The ratio of upper-middle bands thus shows a 2:2 or 2:1:1 gene dosage effect in normal subjects (for example GM2324). In contrast, for GM2325 this is 3:1:2, consistent with the previous results demonstrating three copies of the α globin locus in this cell line.

Subjects GM6226/GM6227 display an uncommon BgII polymorphism (±2) 3' of the ψζ gene (fig 1a) which results in a 4·6 kb band (+) rather than the usual 12·6 kb band (−) and thus distinguishes maternal (+) and paternal (−) 3' ψζ bands. This allows quantification of bands in GM6226 as 1:1:1:1 and in GM6227 as 1:1:1:1. These findings indicate that GM6226 is monosomic for paternal bands and disomic for maternal bands (fig 1c). This estimation

![Molecular localisation of the human α globin locus by Southern blot analysis of cell lines carrying 16p rearrangements. (a) Restriction map of the α globin genes. Closed boxes represent genes and hypervariable region (HVR) sequences are denoted by the symbols IZHVR (inter-ζ) and 3'-HVR. The two probes used are shown underneath. Polymorphic enzyme sites are indicated by ±. B = BamHI, Bg = BglII, E = EcoRI, Pv = PvuII, S = SstI. (b) Distribution of 3'-HVR alleles using PvuII. Faint background bands result from lower stringency detection of a minisatellite family using the 3' HVR probe. This probe detects long tandem arrays of a 17 bp repeat and thus the hybridisation intensity is dependent on the repeat copy number. This makes the probe unsuitable for the quantification of gene dosage effects. Each allele detected by PvuII has about 700 bp non-HVR sequence. N=normal subject, SJ′=inv(16) heterozygote, SJ=rec dup (16p) (see fig 3). 4=GM2324, 5=GM2325, 6=GM6226, 7=GM6227, λ=size markers (kb). (c) Detection of aneuploidy at the IZHVR and 3'-HVR multiallelic loci. These were identified with the ψζ and 3'-HVR probes, respectively. See text for details of gene dosage effects. Comparison of 3'-HVR BgII digests with EcoRI and PvuII (fig 1b) identifies a BgII site polymorphism near the 3'-HVR (fig 1a) in both of these families, also seen in other unrelated persons.
of gene dosage from band quantification is also consistent with the intensity of the BanHI 3' specific fragments (fig 1c).

These data show that although it is only possible to demonstrate one of the two maternal 3'-HVR (fig 1b) and IZHVR (fig 1c) alleles in GM6226, there are two identical copies of the region of the maternal chromosome containing the α globin cluster. The most likely mechanism to explain this is a genetic crossover between adjacent, non-sister chromatids in the region between the α globin locus and the translocation breakpoint, preceding adjacent 1 type segregation at maternal meiosis (fig 2).

**Molecular analysis of an inversion (16)**

The above data indicate that the α locus lies between 16p13.11 and pter. In order to obtain a more accurate distal limit, we analysed DNA from a subject with a familial inversion of chromosome 16, inv(16)(p13-1;q22), and her son with a recombinant chromosome 16 resulting in partial trisomy 16p13.1→pter and partial monosomy 16q22→qter, associated with mental retardation and multiple congenital abnormalities. The recombinant chromosome can only arise from a crossover within the paired, inverted segment at maternal meiosis (fig 3). This will result in duplicated α globin genes if the α locus is distal to the p13-1 breakpoint, but not if it is proximal, and will thus result in the inheritance of two or one maternal 3'-HVR alleles, respectively. The mother (SJ') is heterozygous at the 3'-HVR and the son (SJ) inherits only one maternal 3'-HVR allele (fig 1b). Gene dosage estimations at other regions of the α complex were also consistent with the inheritance of a single maternal α globin locus (data not shown) and therefore this locus lies proximal to the 16p13.1 breakpoint. The combined data (summarised in fig 4) thus show that the α globin genes lie between p13-1 and p13-2 and assign the shortest region of overlap (SRO) to 16p13.1.

![Diagram of chromosomes involved in the reciprocal translocation t(1;16)(q44;p13-11) at pachytene. During meiosis, a putative crossover (X) between the α globin complex and the translocation breakpoint, followed by alternate 1 type segregation, produces unbalanced gametes that include the two chromosomes der(1) and 16 with the same α globin allele (α'). Alternative crossover events or modes of segregation would generate molecular or karyotypic differences from those found.](image1)

![Schematic illustration of the meiotic behaviour of cells heterozygous for the inv(16)> meiotic pairing with one crossing over within the inversion segment and both resulting recombinants. SJ is the rec(16) used in this study. The inversion loop model shown is that commonly accepted. Alternative models have been suggested since there was no evidence of inversion loops on meiotic analysis of pachytene cells from a single inversion heterozygote, although the outcome of a crossover is the same. Distal 16q is shaded and distal 16p is cross hatched.](image2)
IN SITU HYBRIDISATION

In situ hybridisation studies of an α globin probe to metaphase cells (fig 5) were also consistent with this regional localisation. Chromosome 16 was specifically labelled in several normal karyotypes, as expected, although we also observed several weakly labelled regions which may represent dispersed pseudogenes (fig 5a). Two further subjects were analysed for hybridisation sites on chromosome 16 only. From a total of 16 sites on the karyotypically normal chromosomes of one subject, 10 (63%) were localised at p12→p13-1 (fig 5b). From fig 4 it can be predicted that the α globin genes will be relocated in the inv(16) heterozygote (SJ'), and this was confirmed (fig 5c) by showing that the α globin locus (proximal p13) now abuts the distal q22 segment.

Discussion

The human α globin complex contains two highly polymorphic regions which, together with several dimorphic RFLPs, increase the heterozygosity at this locus to almost 1.0 in many populations. Although previous studies have emphasised the importance of such multiallelic markers in establishing a human genetic linkage map, these loci are also of great value in establishing a physical map of the genome. Their presence or absence can be readily shown in well characterised chromosomal rearrangements and hence their regional localisation can be accurately determined. Using this approach, we have localised the human α globin locus to band 16p13-1. Consistent with this is a recent preliminary study that placed it distal to the fragile site at 16p12-3.28 This localisation increases the scope of further studies using a similar approach; for example, it will allow the identification of subcytogenetic rearrangements in this region29 and provide the means to examine the origin of chromosome 16 trisomies, so frequently found in early spontaneous abortions.29 Using similar multiallelic loci this general strategy will be of value in the identification and characterisation of other common chromosomal abnormalities.21 30 31

It has previously been shown that the α globin locus is linked to the loci for phosphoglycolate phosphatase (PGP) and APCKD and is associated with an unusual form of mental retardation.8-10 The region of chromosome 16 shown to contain the α globin genes (16p13-1) lies about mid-way along the genetic span of 16p (approximately 50 cM in the male32 33), so that many other genes on the short arm of chromosome 16 will be linked within 20 cM either side of this. The identification of translocation and inversion breakpoints closely flanking the α locus will enable determination of the chromosome order...
R D Nicholls, J A Jonasson, J O’D McGee, S Patil, V V Ionasescu, D J Weatherall, and D R Higgs

44

(a) Distribution of chromosome associated hybridisation sites from an analysis of metaphase chromosomes: 337 sites were scored from 71 and 36 cells, respectively, from two normal haploid karyotypes. The major cluster of sites is on chromosome 16 and from a total of 46 sites on chromosome 16, 25 (55%) were localised at 16p12–p13.1. Hybridisation above background is observed at 16q22–23 and 22qter (in both subjects) and possibly at 16pter. This most likely represents sequences related to α globin, possibly α globin pseudogenes, that were previously undetected. Low stringency Southern hybridisation also detects specific fragments that are not on chromosome 16 (R D Nicholls, unpublished data). These related sequences may be analogous to the mouse α globin \( \gamma \) gene \( \gamma_3 \) and \( \gamma_4 \) on mouse chromosomes 15 and 17, respectively.24,25 Indeed, several genes on the former appear to be part of a syntenic linkage group with genes on human chromosome 22qter26 and a human α globin related sequence has recently been found on chromosome 22.27

(b) Regional localisation of α globin on chromosome 16 from a third subject. The bar at the top represents a single hybridisation site. The dot at the heterochromatic region (16qh) represents staining of this region due to the Giemsa banding. This may be misinterpreted as hybridisation sites. It was not observed in the experiments in (a). (c) Relocation of α globin labelled sites on the inverted chromosome inv(16) (p13.1q22). The arrows identify the breakpoints. Of 26 hybridisation sites on the inv(16), 17 (65%) were located at p12–p13.1. Analyses were performed on 55 chromosome spreads from the inv(16) heterozygote SJ'.
and orientation of these linked genes with respect to the centromere. This in turn will provide strategies for establishing a physical linkage of neighbouring genes by long range chromosome ‘walking’ techniques. Unfortunately, none of the chromosomal rearrangements described here have breakpoints within or close to (≤ 25 kb) the α globin locus, as determined by genomic mapping with α, ζ, and 3'-HVR probes (Fig 1 and unpublished data).

The combined use of high resolution karyotype analysis and extremely polymorphic loci provides a way of analysing the recombination events that underly the chromosomal rearrangements that occur in germ cells and some malignant somatic cells. In one of the cases reported here, for example, we identified a meiotic crossover in the relatively small cytogenetic distance between the α globin locus and a translocation breakpoint at 16p13-11. Although we consider a meiotic crossover to be the most likely explanation of the data, we cannot exclude an interchromosomal gene conversion event of the distal portion of 16p including at least the α locus. At the molecular level little is known about the sequences involved in large chromosomal rearrangements. Recently, molecular cloning has enabled characterisation of sequences involved in homologous and illegitimate recombination, associated with gene deletions of between about 4 kb to greater than 65 kb, that cause α thalassaemia. (R D Nichols and D R Higgs, in preparation). These may provide a useful background with which to compare the local sequence events occurring at translocation, inversion, and chiasma breakpoints in this relatively small region of the genome.

The accurate identification and characterisation of genes that are closely associated with translocation breakpoints may be important for understanding the functional effect of changing the position and neighbouring sequences of a gene. Pericentric inversions of chromosome 16 have been observed in phenotypically normal subjects, with breakpoints at p11q12-13, p13q12-13 (D Bianchi and S A Latt, 1986, personal communication), and p13-1q22 (this paper). Indeed, 16p breakpoints in the latter two are proximal (R D Nichols, unpublished data) and distal to the α locus, respectively. An inv(16) (p13-1q22) has also been implicated in the evolution of the leukemic cell clone in many cases of acute non-lymphocytic leukemia (ANLL) with abnormal bone marrow eosinophils. The inversion breakpoint in these cases splits the metallothionine (Mt) gene cluster. It has been suggested that an enhancer element in the Mt-2 gene promoter activates a proto-oncogene at 16p13. Comparison of the familial and leukemic inv(16) will clearly be of practical (see above) and biological interest. Furthermore, isolation of this putative proto-oncogene at 16p13 may be of importance in understanding the molecular basis of diseases closely linked to the α globin gene locus at 16p13-1.

We thank Dr C Potter for culture of the cell lines, H Teal for DNA extractions, Dr M Aronson for communication of unpublished cytogenetic data, and L Roberts for typing the manuscript. We also thank Dr J B Clegg for support and encouragement and Drs K E Davies, S T Readers, W G Wood, and J B Clegg for comments on the manuscript. RDN was initially supported by a Research Scholarship from the Royal Commission for the Exhibition of 1851 and for the latter part of this project was supported by the Rockefeller Foundation.

References


